Black-Box Separations for Differentially Private Protocols

Dakshita Khurana*, Hemanta K. Maji*, and Amit Sahai*

Department of Computer Science, UCLA and Center for Encrypted Functionalities.
{dakshita,hmaji,sahai}@cs.ucla.edu

Abstract. We study the maximal achievable accuracy of distributed differentially private protocols for a large natural class of boolean functions, in the computational setting.

In the information theoretic model, McGregor et al. [FOCS 2010] and Goyal et al. [CRYPTO 2013] demonstrate several functionalities whose differentially private computation results in much lower accuracies in the distributed setting, as compared to the client-server setting.

We explore lower bounds on the computational assumptions under which this accuracy gap can possibly be reduced for two-party boolean output functions. In the distributed setting, it is possible to achieve optimal accuracy, i.e. the maximal achievable accuracy in the client-server setting, for any function, if a semi-honest secure protocol for oblivious transfer exists. However, we show the following strong impossibility results:

- For *any* general boolean function and fixed level of privacy, the maximal achievable accuracy of any (fully) black-box construction based on existence of key-agreement protocols is at least a constant smaller than optimal achievable accuracy. Since key-agreement protocols imply the existence of one-way functions, this separation also extends to one-way functions.
- Our results are tight for the AND and XOR functions. For AND, there exists an accuracy threshold such that any accuracy up to the threshold can be information theoretically achieved; while no (fully) black-box construction based on existence of key-agreement can achieve accuracy beyond this threshold. An analogous statement is also true for XOR (albeit with a different accuracy threshold).

Our results build on recent developments in black-box separation techniques for functions with private input [1, 16, 27, 28]; and translate information theoretic impossibilities into black-box separation results.

Keywords: Differentially Private Protocols, Computational Complexity, Random Oracle, Key-agreement Protocols, Black-box Separation.

* Research supported in part from a DARPA/ONR PROCEED award, NSF grants 1228984, 1136174, 1118096, and 1065276, a Xerox Faculty Research Award, a Google Faculty Research Award, an equipment grant from Intel, and an Okawa Foundation Research Grant. This material is based upon work supported by the Defense Advanced Research Projects Agency through the U.S. Office of Naval Research under Contract N00014-11- 1-0389. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense, the National Science Foundation, or the U.S. Government.

1 Introduction

Differential privacy [7] provides strong input privacy guarantees to individuals participating in a statistical query database. Consider the quintessential example of trying to publish some statistic computed on a database holding confidential data hosted by a trusted server [31]. For example, consider a query that checks if there is an empirical correlation between smoking and lung cancer instances from the medical records of patients stored at a hospital. The server wants to provide *privacy* guarantees to each record holder as well as help the client compute the statistic *accurately*. Even in this setting, where privacy concerns lie at the server's end only, it is clear that privacy and accuracy are antagonistic to each other. The tradeoff between accuracy and privacy is non-trivial and well understood only for some classes of functions (for e.g. [30, 15]). For any level of privacy, we refer to the maximal achievable accuracy in the client-server setting for a particular functionality, as the *optimal accuracy*.

In the distributed setting, where multiple mutually distrusting servers host parts of the database, privacy concerns are further aggravated. Continuing the previous example, consider the case of two hospitals interested in finding whether a correlation exists between smoking and lung cancer occurrences by considering their combined patient records. In such a setting, we want the servers to engage in a protocol, at the end of which the privacy of each record of both the servers is guaranteed without a significant loss in accuracy. Note that the privacy requirements must be met for both servers, *even given their view of the protocol transcript, not just the computed output*; thus, possibly, necessitating an additional loss in accuracy.

At a basic level, we wish to study privacy-accuracy tradeoffs that arise in the distributed setting. Following [15], in order to obtain results for a wide class of functions, we focus on the computation of functions with Boolean output, with accuracy defined (very simply) as the probability that the answer is correct. The intuition that privacy in the distributed setting is more demanding is, in fact, known to be true in the information theoretic setting: For any fixed level of privacy, it was shown that for all boolean functions that the maximal achievable accuracy in the distributed setting is significantly lower than the optimal accuracy achievable in the client-server setting [15], as long as the boolean function depends on both server's inputs. But in the computational setting, this gap vanishes if a (semi-honest[1]) protocol for oblivious-transfer exists. The two servers would then be able to use secure multi-party computation [14] to simulate the client-server differentially private computation, thereby achieving optimal accuracy on the union of their databases. Although this computational assumption suffices, it is not at all clear whether this assumption is *necessary* as well.

Indeed, this is a fascinating question because even for very simple functions, like XOR, that require no computational assumptions to securely compute in

[1] In this work, as in previous works on distributed differential privacy, we restrict ourselves to the semi-honest setting where all parties follow the specified protocol, but remember everything they have seen when trying to break privacy.

the semi-honest setting, the question of differentially private computation is non-trivial. Could there be any simple functions that can be computed differentially privately with weaker assumptions? For the general class of boolean output functions, our paper considers the following problem:

"What are the computational assumptions under which there exist distributed differentially private protocols for boolean f with close to optimal accuracy?"

Goyal et al. [15] showed that for any boolean function such that both parties' inputs influence the outcome, achieving close to optimal accuracy would imply the existence of one-way functions. Could one-way functions also be *sufficient* to achieve optimal accuracy for certain simple functions?

Our results give evidence that the answer is *no*. Indeed, we provide evidence that achieving optimal accuracy for *any* boolean function that depends on both parties' inputs is not possible based on one-way functions. We go further and provide similar evidence that this goal is not possible even based on the existence of key-agreement protocols (which also implies one-way functions; and, thus, is a stronger computational assumption). More precisely, we show a (fully black-box) separation [35] of the computational assumptions necessary to bridge the accuracy gap from the existence of key-agreement protocols. A black-box separation between two cryptographic primitives has been widely acknowledged as strong evidence that they are distinct [23]. Indeed, we note that a black-box separation is particularly meaningful in the context of protocols with guarantees only against *semi-honest* adversaries, like the differentially private protocols we consider in this work. (Recall that an impossibility result like ours is strongest when it applies to the *weakest* security setting possible – this is why we focus on just semi-honest security.) This is because the most common non-black-box techniques used in cryptography typically apply only to the setting of malicious adversaries: for example, cryptographic proof systems like zero-knowledge proofs are sometimes applied in a non-black-box manner in order for a party to prove that it behaves honestly. However, in the semi-honest security context, such proofs are never needed since even adversarial parties must follow the protocol as specified. We crucially employ recently developed separation techniques for protocols with private inputs from key-agreement protocols [27, 28].

Our work is reminiscent of, but also quite different from, the work of Haitner et al. [16], who proved that the information theoretic impossibility of accurate distributed differentially private evaluation of the inner-product functionality [30] could be extended to a black-box separation result from one-way functions. Our results are different both qualitatively and technically: Qualitatively, our results differ in that they apply to the wide class of all boolean functions where the output of the function is sensitive to both parties' inputs. Furthermore, we show separations from key-agreement protocols as well. Moreover, our separation results for extremely simple binary functions like AND and XOR show that differentially private distributed computation even of very simple functions may also require powerful computational assumptions.

At a technical level, a crucial ingredient of our proofs is the recently developed toolset of [27, 28] which deal with *private inputs of parties* even in presence of

the "idealized key-agreement oracle," while Haitner et al. [16] *adapt* the analysis of McGregor et al. [30] to a setting where the input is part of the local random tape of parties, i.e. parties have no private inputs.

1.1 Our Contribution

Before we elaborate upon our results, we briefly summarize what is known so far about accuracy gaps in boolean distributed differentially private computation.

Suppose Alice and Bob have inputs x and y, respectively; and they are interested in computing $f(x, y)$ in a differentially private manner in the distributed setting. An ε-differentially private protocol for some functionality f ensures that the probability of Alice's views conditioned on y and y' are $\lambda := \exp(\varepsilon)$ multiplicatively-close to each other, where y and y' represented as bit-strings differ only in one coordinate (i.e. they are *adjacent* inputs). Let x and y be the private inputs of parties Alice and Bob respectively. A protocol between them is α-accurate if for any x and y, the output of the protocol agrees with $f(x, y)$, with probability at least α.

For boolean functions, the optimal accuracy (in the client-server model) is $\alpha^*_\varepsilon := \frac{\lambda}{(\lambda+1)}$, where $\lambda = \exp(\varepsilon)$.[2] Goyal et al. [15] showed that, in the information theoretic setting, $f = \mathsf{AND}$ can only be computed ε-differentially privately up to accuracy $\alpha^{(\mathsf{AND})}_\varepsilon := \frac{\lambda(\lambda^2+\lambda+2)}{(\lambda+1)^3}$. Similarly, when $f = \mathsf{XOR}$ the maximal achievable accuracy is $\alpha^{(\mathsf{XOR})}_\varepsilon := \frac{(\lambda^2+1)}{(\lambda+1)^2}$. Note that $\alpha^{(\mathsf{XOR})}_\varepsilon < \alpha^{(\mathsf{AND})}_\varepsilon < \alpha^*_\varepsilon$, for any finite $\varepsilon > 0$. By observing that any boolean function f which is sensitive to both parties' inputs either contains an embedded XOR or AND[3] [3], the maximal achievable accuracy is bounded by:

$$\alpha^{(f)}_\varepsilon := \begin{cases} \alpha^{(\mathsf{XOR})}_\varepsilon, & \text{if } f \text{ contains an embedded } \mathsf{XOR} \\ \alpha^{(\mathsf{AND})}_\varepsilon, & \text{otherwise.} \end{cases} \tag{1}$$

Note that in the computational setting, if semi-honest secure protocol for oblivious-transfer exists then we can achieve accuracy $\alpha = \alpha^*_\varepsilon$ for any boolean f. We explore the necessary computational assumptions for which this gap in accuracy in the distributed and client-server setting vanishes. Although Goyal et al. [15] showed that achieving close to optimal accuracy implies one-way functions, we show that it is highly unlikely that such constructions can solely be based on one-way functions. In fact, we show a (fully) black-box separation from a weaker variant of differential privacy, namely *computational differential privacy* (see Section 2).

---

[2] In the client-server setting, any boolean function f can be computed ε-differentially privately by evaluating a suitably noisy version of f.

[3] We say that f contains an embedded XOR if there exists $x_0, x_1, y_0, y_1, z_0, z_1$ such that $f(x_a, y_b) = z_{\mathsf{XOR}(a,b)}$ for all $a, b \in \{0, 1\}$. Similarly, we define en embedded AND. Note that embedded OR is identical to embedded AND (by interchanging z_0 and z_1).

Informal Theorem 1. *For any boolean f and privacy threshold $\varepsilon > 0$, there exists a constant $c > 0$ such that any ε-differentially private α-accurate evaluation of f (in the distributed setting) which uses key-agreement protocols in fully black-box manner cannot have accuracy $\alpha > (\alpha^*_\varepsilon - c)$, where $\alpha^*_\varepsilon = \frac{\lambda}{(\lambda+1)}$ and $\lambda = \exp(\varepsilon)$.*

Further, our result is tight for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ and, in fact, a stronger lower bound is exhibited. We show that for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$: 1) In the information theoretic setting, it is possible to ε-differentially privately α-accurately evaluate f in the distributed setting [15], if $\alpha \leqslant \alpha^{(f)}_\varepsilon$, and 2) In the computational setting, it is impossible to construct (by using key-agreement protocols in black-box manner) an ε-differentially private α-accurate evaluation of f, for $\alpha \geqslant \alpha^{(f)}_\varepsilon + 1/\mathsf{poly}(\kappa)$ (where, κ is the statistical security parameter). In fact, this gives a (fully) black-box separation of a weaker notion of differential privacy, namely *computational differential privacy* (see Section 2). Note that it suffices to just consider $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ because the maximal achievable accuracy for a general boolean function is bounded in terms of $\alpha^{(\mathsf{AND})}_\varepsilon$ and $\alpha^{(\mathsf{XOR})}_\varepsilon$. As a primer, we begin with the separation result from existence of one-way functions.

Separation from One-way Functions. Random oracles serve as an idealization of one-way functions because they cannot be inverted at non-negligible fraction of their image by any algorithm whose query complexity is polynomial in query-length of the random oracle [23, 12].

Suppose there exists a purported ε-differentially private α-accurate protocol for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ in the random oracle world, where parties have unbounded computational power and their query complexity is at most n. We show that if $\alpha \geqslant \alpha^{(f)}_\varepsilon + \sigma$ then one of the parties could perform additional $\mathsf{poly}(n/\sigma\varepsilon)$ queries to the random oracle and break the ε-differential privacy of the protocol. The existence of this strategy relies on the recent progress of "Eavesdropper strategies in the random oracle setting" for protocols with private inputs [27]. For more details, refer to Imported Theorem 1.

This impossibility result easily translates into a fully black-box separation as defined in [35]. This translation of impossibility in the random-oracle model into a black-box separation uses techniques introduced in [23, 13, 1, 5, 16, 27].

Informal Theorem 2 (Separation from One-way Functions). *For $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, $\varepsilon > 0$ and $\alpha \geqslant \alpha^{(f)}_\varepsilon + 1/\mathsf{poly}(\kappa)$, where κ is the security parameter, there cannot exist an ε-differentially private α-accurate protocol for f in the distributed setting which uses one-way functions in fully black-box manner.*

Note that this separation also extends to primitives which can be constructed from one-way functions in black-box manner, like pseudorandom generators [21, 18, 19] and digital signatures/universal one-way hash functions [33, 36, 26]. Moreover, it is also applicable to other computational primitives like ideal-ciphers [4, 20] (which are indifferentiable [29] from random oracles) and one-way permutations (which themselves cannot be based on one-way function [37, 25]).

Separation from Public-key Encryption. To show a similar separation result from key-agreement protocols, it suffices to show a separation from public-key encryption; because public-key encryption is equivalent to two-round key agreement which in turn directly implies (any round) key-agreement protocols. Before we proceed further, we introduce the idealization of public-key encryption as an oracle [13].

Our public-key encryption oracle is a triplet of correlated oracles $\mathbb{PKE} \equiv (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$. The key-generation oracle Gen is a length tripling random oracle which maps $sk \in \{0,1\}^n$ to $pk \in \{0,1\}^{3n}$, i.e. $\mathsf{Gen}(sk) = pk$. The encryption oracle, is a collection of 2^{3n} independent length-tripling oracles which maps a message m, using a public-key $pk \in \{0,1\}^{3n}$, to a cipher text c, i.e. $\mathsf{Enc}(m; pk) = c$. The decryption oracle Dec decrypts a cipher text $c \in \{0,1\}^{3n}$ using a secret key $sk \in \{0,1\}^n$. It maps it to (the lexicographically first) m such that $\mathsf{Gen}(sk) = pk$ and $\mathsf{Enc}(m; pk) = c$; otherwise outputs $\perp$, i.e. $\mathsf{Dec}(c, sk) \in \{m, \perp\}$.

This oracle is too powerful and yields a semi-honest secure protocol for oblivious-transfer (see discussion in [13]). Thus, it cannot be used to show the intended separation result. An additional Test oracle is provided, which allows testing of whether pk lies in the range of the Gen oracle, and whether c lies in the range of the Enc oracle with public key pk. Intuitively, the Test oracle can be thought of as part of Gen and Enc oracles themselves. Such oracles with *image-testability* are referred to as *image-testable random oracles* (ITRO) [28].

To tackle the decryption oracle, we follow the technique introduced by [28]. Suppose there exists a purported ε-differentially private α-accurate protocol for f in the PKE-oracle world. Then there exists an $(\varepsilon + \gamma)$-differentially private $(\alpha - \gamma)$-accurate protocol for f in the "PKE minus decryption oracle" world, i.e. in the $(\mathsf{Gen}, \mathsf{Enc})$ oracle world (with implicitly included Test oracles), with query complexity $\mathsf{poly}(n/\gamma\varepsilon)$ and identical round complexity. The slight loss in parameter γ can be made arbitrarily small $1/\mathsf{poly}(n)$.

Finally, similar to the separation from one-way functions, we show that if $(\alpha - \gamma) \geqslant \alpha^{(f)}_{\varepsilon+\gamma} + (\sigma/2)$ then one of the parties can perform $\mathsf{poly}(n/\sigma\gamma\varepsilon)$ queries and violate the $(\varepsilon + \gamma)$-differential privacy of this protocol. This part of the result crucially relies on the recently proven result of [28] which shows that image-testable random oracles *mimics* several properties of random-oracles and the "eavesdropper strategies" in the random oracle model extend to (collections of) image-testable random oracles as well. Hence, we have the following result.

Informal Theorem 3 (Separation from Key-Agreement). *For $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, $\varepsilon > 0$ and $\alpha \geqslant \alpha^{(f)}_{\varepsilon} + 1/\mathsf{poly}(\kappa)$, where κ is the security parameter, there cannot exist an ε-differentially private α-accurate protocol for f in the distributed setting which uses key-agreement protocols in fully black-box manner.*

We emphasize that our negative results not only hold for ε-differential privacy, but also hold for a weaker (ε, δ)-indistinguishability based computational differential privacy (see Section 2 for definition). For a precise statement refer to our main theorem, Theorem 1.

1.2 Related Work

Differential Privacy. Differential privacy [8, 7, 10, 11, 6] has been popular as a strong privacy guarantee to participants of statistical databases. In settings where the database could possibly be split among various parties, Dwork et al. [9] obtained distributed differential privacy via SFE and secure noise generation. Subsequently, [2] studied trade-offs between distributed privacy and SFE. A computational relaxation of differential privacy was defined by Mironov et al. [31], that would help improve the range of achievable accuracies while still maintaining this relaxed notion of privacy.

A gap in the maximal achievable accuracy of differentially private protocols, between the client-server and distributed settings, was first observed by McGregor et al. [30] for specific large functions such as the inner product and hamming distance. Recently, Goyal et al. [15] showed the existence of a constant information theoretic gap between the accuracies of boolean output functions, in the client-server and distributed settings. They also showed that any hope of bridging this gap necessitates the assumption that one-way functions exist.

Black-box Separations. Impagliazzo and Luby [22] showed that most non-trivial cryptographic primitives imply existence of one-way functions. Subsequently, it turned out that several primitives like pseudorandom generators [21, 18] and digital signatures/universal one-way hash functions [33, 36] can indeed be constructed from one-way functions; thus, establishing equivalence of these primitives to existence of one-way functions. It is highly unlikely, on the other hand, that primitives like key-agreement [23] protocols and semi-honest secure oblivious-transfer protocol [13] can be securely constructed from one-way functions using black-box construction. A black-box separation result between two cryptographic primitives is widely acknowledged as an evidence that they should be treated as separate computational assumptions.

Reingold et al. [35] formally defined (several variants of) black-box separations. And Gertner et al. [12] provided a technique to translate information theoretic impossibility results in random oracle model into unconditional black-box separation results.

Recently, there has been significant progress in black-box separation techniques where parties have private inputs due to [27, 28]. They show that if semi-honest secure function evaluation of any two-party deterministic function exists by using one-way functions or key-agreement protocols in black-box manner then there exists a semi-honest secure protocol for that function in the information theoretic plain model itself. Haitner et al. [16] show that the information theoretic impossibility of evaluating the inner-product functionality both differentially privately and accurately [30], in the client-server model, can be translated into a black-box separation result from one-way functions.

1.3 Technical Outline

Our black-box separation results are a consequence of amalgamation of the following techniques: 1) Information theoretic lower bounds for ε-differentially pri-

vate α-accurate protocols for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ in the distributed setting [15], and 2) Recent progress in black-box separation techniques as introduced in [1, 16, 27, 28]. Our separation from key-agreement protocols especially relies on the recent results of [28]. We essentially show that based on computational assumptions like "existence of one-way functions" and "existence of (any round) key-agreement protocol" it is highly unlikely to construct ε-differentially private α-accurate protocols for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, if $\alpha \geqslant \alpha_{\varepsilon}^{(f)}$.

Henceforth, we shall assume that $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ and understand the computational assumptions necessary to realize ε-differentially private α-accurate protocols for f, where $\alpha > \alpha_{\varepsilon}^{(f)}$.

Information theoretic result. Before we begin, we sketch an intuitive summary of the proof technique of Goyal et al. [15]. They leveraged the Markov-chain property of distribution of next-message function in the information theoretic setting, i.e. the next message sent by a party is *solely* a (deterministic) function of its current view. Suppose the public transcript generated thus far is m. Then, using this Markov-chain property of protocols in the information theoretic setting and the fact that they begin with independent views, one can obtain the following *protocol compatibility constraint*: $\Pr[m|x, y] \cdot \Pr[m|x', y'] = \Pr[m|x, y'] \cdot \Pr[m|x', y]$, for private inputs $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$. By considering every complete transcript m, the protocol compatibility constraint implies a set of constraints. For every privacy parameter $\varepsilon \geqslant 0$, they show that there exists an ε-differentially private α-accurate protocol for f, if $\alpha \in [0, \alpha_{\varepsilon}^{(f)}]$.

Separation from one-way functions. Although the result presented in this section is subsumed by our main theorem, we feel that an independent presentation of this result adds clarity to the overall proof.

Suppose we have a (purportedly) ε-differentially private α-accurate protocol for f in the random oracle model, where each party performs at most n private queries to the random oracle. A random oracle randomly maps κ-length bit-strings to κ-length bit-strings, where κ is the statistical security parameter. Assume that $\alpha \geqslant \alpha_{\varepsilon}^{(f)} + \sigma$, where $\sigma = 1/\mathsf{poly}(\kappa)$. To show a black-box separation result from one-way functions, we need to show that if α is significantly larger than $\alpha_{\varepsilon}^{(f)}$, then differential privacy must be violated by one of the parties.

But, the Markov-chain property (upon which the information theoretic characterization crucially relies) is not a priori guaranteed in the random oracle model. So, a logical starting point is to consider an algorithm which perform additional queries to the random oracle to kill correlations between parties and ensures this property (with high probability), cf. [23, 1, 5, 16, 27]. For any $\rho > 0$, there exists a (deterministic) algorithm Eve_{ρ} which performs additional $\mathsf{poly}(n/\rho)$ queries to the random oracle based on the public transcript; and appends the sequence of query-answer pairs to the current transcript. This Eve_{ρ} ensures that when she stops, the joint view of Alice and Bob is ρ-close to a product distribution with $(1 - \rho)$ probability. Being agnostic to the private inputs used by the parties, Eve_{ρ} can ensure this Markov-chain property only when,

for any complete transcript m, the probabilities $\Pr[y|m]$ and $\Pr[x|m]$, for every $x \in \{0,1\}$ and $y \in \{0,1\}$, is at least a constant [27].

Note that the ε-differential privacy constraint implies that $\Pr[x|m]$ and $\Pr[x'|m]$ are $\lambda = \exp(\varepsilon)$ (multiplicative) approximations of each other for all adjacent x and x'. If f is a function such that both parties' inputs influence the output, then it has an embedded AND or XOR minor in its truth table. Let $\mathcal{X}$ and $\mathcal{Y}$ be the respective input sets of Alice and Bob such that f restricted to $\mathcal{X} \times \mathcal{Y}$ is an AND or XOR minor. Given such a minor, our negative result shall exhibit violation of the differential privacy guarantee. So, for all our negative results we have $|\mathcal{X}| = |\mathcal{Y}| = 2$. Consequently, $\Pr[x|m]$ is a constant for every $x \in \mathcal{X}$; otherwise the complete transcript m is a witness to violation of ε-differential privacy. Analogously, the same holds for every $y \in \mathcal{Y}$.

For Alice inputs in $\mathcal{X}$ and Bob inputs in $\mathcal{Y}$, for any $\rho > 0$, there exists Eve_ρ with query complexity $\mathsf{poly}(n/\rho)$ such that, with probability $(1-\rho)$ over the generated public transcript, the joint view of Alice-Bob is ρ-close to a product distribution. Now, consider the *augmented protocol* where the original ε-differentially private α-accurate protocol is augmented with Eve_ρ, who adds her sequence of query-answer pairs to the public transcript. In this augmented protocol, we show that ε-differentially private α-accurate protocol implies $\alpha \leqslant \alpha^{(f)}_{\varepsilon,\rho}$, which can be made arbitrarily close to $\alpha^{(f)}_{\varepsilon}$ by choosing suitably small value of ρ. Intuitively, this result relies on the fact that the polytope of feasible solutions to the constraints in the information theoretic setting cannot change significantly if each of them has bounded slope and is weakened slightly (see full version for details). When $\alpha = \alpha^{(f)}_{\varepsilon} + \sigma$, where $\sigma = 1/\mathsf{poly}(n)$, by choosing suitably small $\rho = \mathsf{poly}(\sigma\varepsilon)$, one of the parties can violate the ε-differential privacy guarantee by performing $\mathsf{poly}(n/\rho)$ additional queries to the random oracle.

This technique is applied in a significantly sophisticated manner to show the separation from key-agreement protocols.

Separation from key-agreement. We show a separation from public-key encryption, which is equivalent to a 2-round key-agreement protocol. Separation from 2-round key-agreement implies separation from (any round) key-agreement protocols. This separation relies on the recent results pertaining to the "ideal public-key encryption oracle" (PKE-oracle, introduced by [13]) as shown in [28].

Our result depends on two technical results proven in [28]. First, they show that, against semi-honest adversaries, queries to the decryption-oracle of PKE-oracle are (nearly) useless; and, finally, the PKE-oracle minus the decryption-oracle (closely) mimics properties of (collection of) random oracles.

The first part shows that if there exists an ε-differentially private α-accurate protocol for f in the PKE-oracle world, then there exists another (closely related) $(\varepsilon+\gamma)$-differentially private $(\alpha-\gamma)$-differentially private protocol for f in the "PKE-oracle minus the decryption-oracle" world with query complexity $\mathsf{poly}(n/\gamma)$. Here, the parameter γ can be made arbitrarily small $1/\mathsf{poly}(n)$.

Finally, we use the property that "PKE-oracle minus decryption-oracle" is *similar* to the random oracle world [28]. We use the fact that, relative to this

oracle, there exists an Eve_ρ which can make the joint distribution of Alice-Bob joint views ρ-close to product with high probability. Since $(\alpha-\gamma) > \alpha^{(f)}_{\varepsilon+\gamma,\rho}$, one of the parties can violate the $(\varepsilon+\gamma)$-differential privacy of the protocol.

Overall, if δ is at least $\alpha^{(\mathsf{AND})}_\varepsilon + \sigma$, where $\sigma = 1/\mathsf{poly}(n)$, we can choose $\gamma, \rho = \mathsf{poly}(\sigma\varepsilon)$ to show that the ε-differential privacy is violated by performing only $\mathsf{poly}(n/\sigma\varepsilon)$ queries to the PKE-oracle. In fact, our final theorem rules out a stronger form of differentially private protocols, namely, (ε,δ)-computational differential privacy (see Section 2 for definitions). Intuitively, $\delta = 0$ corresponds to the previously discussed notion of ε-differential privacy. Our final theorem is:

Theorem 1. *For any boolean function f whose output is sensitive to both parties' inputs, $\varepsilon > 0$ and $\lambda = e^\varepsilon$, define $\alpha^{(f)}_\varepsilon$ as follows:*

$$\alpha^{(f)}_\varepsilon := \begin{cases} \alpha^{(\mathsf{XOR})}_\varepsilon = \frac{\lambda^2+1}{(\lambda+1)^2}, & \text{if } f \text{ contains an embedded } \mathsf{XOR} \\ \alpha^{(\mathsf{AND})}_\varepsilon = \frac{\lambda(\lambda^2+\lambda+2)}{(\lambda+1)^3}, & \text{otherwise.} \end{cases}$$

Then for any $\alpha \geqslant \alpha^{(f)}_\varepsilon + \sigma$, where $\sigma = 1/\mathsf{poly}(\kappa)$ and κ is the statistical security parameter, there exists a $\hat{\delta} = \mathsf{poly}(\sigma\varepsilon)$ such that any (ε,δ)-computational differentially private α-accurate protocol for f in the distributed setting constructed in a fully black-box manner from key-agreement protocols must have $\delta \geqslant \hat{\delta}$. Further, when $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ and $\varepsilon > 0$, there exists an ε-differentially private α-accurate protocol for f, if $\alpha \leqslant \alpha^{(f)}_\varepsilon$.

The negative result rules out fully-BB constructions of ε indistinguishable computationally differentially private (ε-IND-CDP) α-accurate protocols with $\alpha > \alpha^{(f)}_\varepsilon$, based on existence of key agreement. The second part of the theorem (the positive result) is with respect to the stronger notion of ε-differential privacy.

An overview of the separation from one-way functions is provided in Section 3. An overview of the proof of Theorem 1 is presented in Section 4. Compplete proofs are deferred to the full version.

2 Preliminaries

We introduce important definitions in this section, with details in the full version.

Differential Privacy. The following definitions of differential privacy are provided for the distributed setting:

Definition 1 ((ε,δ)-Differential Privacy). *A two-party protocol Π is* (ε,δ)-differentially private*, referred to as (ε,δ)-DP, if for any subset $\mathcal{S}$ of Alice-views, for all Alice inputs x and for any pair of adjacent[4] Bob inputs y, y', we have:*

$$\Pr[\mathcal{S}|x,y] \leqslant \exp(\varepsilon)\cdot\Pr[\mathcal{S}|x,y'] + \delta$$

The same condition also holds for adjacent Alice inputs x, x' and all Bob's inputs y with respect to Bob private views.

[4] Two inputs are adjacent if they differ only in one coordinate.

Definition 2 ((ε,δ)-(IND)-Computational Differential Privacy). *A two-party protocol Π is* (ε,δ)-computational differentially private*, referred to as* (ε,δ)-IND-CDP*, if for any efficient adversary $\mathcal{A}$, for all Alice inputs x and any pair of adjacent Bob inputs y, y', we have:*

$$\Pr[\mathcal{A}(V_A, 1^\kappa) = 1|x, y] \leqslant \exp(\varepsilon) \cdot \Pr[\mathcal{A}(V_A, 1^\kappa) = 1|x, y'] + \delta$$

The same condition also holds for adjacent Alice inputs x, x' and all Bob's inputs y, with respect to Bob private views.

We refer $(\varepsilon, \mathsf{negl}(\kappa))$-IND-CDP as ε-IND-CDP, defined first in [31]. We note that this indistinguishability based definition is weaker than the simulation based one (SIM-CDP privacy [31]). Our separations hold even for this weaker differential privacy definition. In the above definition, the protocol Π, ε and δ are parameterized by the security parameter κ as well, but is not explicitly mentioned for ease of presentation. Without loss of generality, we assume that ε is not an increasing function (of κ); and in all our analysis we shall have δ as a decreasing function.

Accuracy. Following [15] we measure the accuracy of two-party protocols in evaluating a boolean function as follows:

Definition 3 (α-Accuracy). *A two party protocol Π evaluates a function f α-accurately, if, for every private input x and y of Alice and Bob respectively, the output of the protocol is identical to $f(x, y)$ with probability at least α.*

Information theoretic bounds on the maximal achievable accuracy for ε-DP protocols computing the AND and XOR functions, are known in the Plain Model [15]. Define $\lambda = \exp(\varepsilon)$, then $\alpha_\varepsilon^{(\mathsf{AND})} = \frac{\lambda(\lambda^2+\lambda+2)}{(\lambda+1)^3}$ is the maximal achievable accuracy of any protocol for the AND function, and $\alpha_\varepsilon^{(\mathsf{XOR})} = \frac{\lambda^2+1}{(\lambda+1)^2}$, is the maximal achievable accuracy of any protocol for the XOR function.

Black-box Separations. We use the definition of fully black-box construction as introduced by Reingold et al. [35]. To show a separation of (ε,δ)-IND-CDP α-accurate protocol from key-agreement protocols, we need to show existence of an oracle relative to which key-agreement protocol exists but there exists an adversary which violates the (purported) (ε,δ)-IND-CDP guarantee.

3 Separation from One-way Functions

Our main result shows a separation from key-agreement protocols. Despite the fact that the separation from one-way functions will be subsumed by our separation from key-agreement protocols, we present this result separately because it is conceptually simpler and captures several of the crucial ideas required to show such black-box separation results.

For $\varepsilon > 0$ differential privacy parameter, suppose $\alpha \in [\alpha_{\varepsilon}^{(f)} + 1/\mathsf{poly}(\kappa), \alpha_{\varepsilon}^{*}]$. We shall show that, for such choices of α, we cannot construct ε-IND-CDP α-accurate protocols for boolean f, in the information theoretic random oracle world. It suffices to show this result for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$. This is done by showing an impossibility result in the random oracle model against information theoretic adversaries but with polynomially bounded query complexity. However, we shall show existence of an adversary who can break the ε-IND-CDP.

3.1 Notations and Definitions

We introduce some notations for our separation result. For security parameter κ, let $\mathbb{O}_{\kappa}$ denote the set of all functions from $\{0,1\}^{\kappa} \to \{0,1\}^{\kappa}$.

We will consider *private-input randomized two party protocols* Π, such that Alice and Bob have access to a common random oracle $O \xleftarrow{\$} \mathbb{O}_{\kappa}$. As in the plain model, parties send messages to each other in alternate rounds, starting with Alice in the first round. However, they have (private) access to a common random oracle.

For odd i, at the beginning of the i^{th} round, Alice queries the random oracle multiple times based on her current view (private input x, local randomness r_A, private query-answer pairs and the transcript $m^{(i-1)}$ so far). She appends the new set of query-answer pairs $P_{A,i}$ to her partial sequence of query-answers. The complete set of private query-answers at this point is denoted by $P_A^{(i)}$. She then computes her next-message m_i as a function of her current view, $(x, r_A, m^{(i-1)}, P_A^{(i)})$. The i^{th} round ends when she sends message m_i. Her view at the end of round i is $V_A^{(i)} \equiv (x, r_A, m^{(i)}, P_A^{(i)})$. Similarly, Bob queries the oracle followed by computing and sending his message in even rounds as a function of his view. His view at the end of round i is (analogously) defined to be $V_B^{(i)} \equiv (y, r_B, m^{(i)}, P_B^{(i)})$. At the end of n rounds, both parties locally obtain outputs as an efficiently computable deterministic function out of their view, $z_A = \mathsf{out}(V_A^{(n)})$ and $z_B = \mathsf{out}(V_B^{(n)})$. We note at this point, that we our analysis will only be over functions with boolean output, such that $z_A, z_B \in \{0,1\}$. Our underlying sample space in the random oracle world is the joint distribution over Alice-Bob views when $r_A, r_B \sim \mathbf{U}$ and $O \xleftarrow{\$} \mathbb{O}_{\kappa}$.

Two-party protocols in the Random Oracle World Before we present our separation result, we need to introduce the notion of *public-query* strategy and *augmentation of a protocol* with a public-query strategy.

Definition 4 (Public Query strategy). *A public query strategy is a deterministic algorithm, which, after every round of the protocol, queries the oracle multiple times based on the transcript generated thus far. It then adds this sequence of query-answers to the transcript being generated.*

Definition 5 (Augmented Protocol). *Given a protocol Π, the augmented protocol Π^{+}:=(Π, Eve) denotes Π augmented with a* public query strategy *"Eve"*

which generates public query-answer sequences after every message in Π and appends them to the protocol transcript after the messages in Π.

Now, we define the views of parties (Alice, Bob and Eve) in an augmented protocol $\Pi^+ := (\Pi, \text{Eve})$. The protocol Π proceeds with parties sending messages in alternate rounds and Eve appending query-answer pairs after the message of the underlying protocol Π is sent.

Formally, consider an odd i. Alice is supposed to generate the message m_i in round i. Round i *begins* with Alice querying the random oracle based on her view $(x, r_A, m^{(i-1)}, P_A^{(i-1)}, P_E^{(i-1)})$, where $P_E^{(i-1)}$ is the sequence of query-answer pairs added by Eve thus far. Alice performs additional queries $P_{A,i}$ and sends the next message m_i. Thereafter, the public query strategy Eve performs additional queries to the random oracle and adds the corresponding sequence of query-answer pairs $P_{E,i}$ to the transcript. This marks the end of round i. At this point, the views of parties Alice, Bob and Eve are: $V_A^{(i)} \equiv (x, r_A, m^{(i)}, P_A^{(i)}, P_E^{(i)})$, $V_B^{(i)} \equiv (y, r_B, m^{(i)}, P_B^{(i)}, P_E^{(i)})$ and $V_E^{(i)} \equiv (m^{(i)}, P_E^{(i)})$, respectively.

(ε, δ)-IND-CDP in the Random Oracle Model

Definition 6 ((ℓ, n) Two-party Protocol). *An (ℓ, n) two-party protocol is a two-party protocol of round complexity at most n such that both parties have query complexity at most ℓ.*

Definition 7 ((ε, δ)-IND-CDP (ℓ, n) Protocol).
A two-party protocol Π is (ε, δ)-IND-CDP if for any computationally unbounded adversary (but polynomial-query complexity) $\mathcal{A}$ and any pair of adjacent Bob inputs y, y', we have:

$$\Pr[\mathcal{A}^O(V_A, 1^\kappa) = 1 | y] \leqslant \exp(\varepsilon) \cdot \Pr[\mathcal{A}^O(V_A, 1^\kappa) = 1 | y'] + \delta$$

The same condition also holds for adjacent Alice inputs x, x' with respect to Bob private views.

We emphasize that the adversary $\mathcal{A}$ gets access to an oracle O with respect to which the view V_A is generated. Accuracy is defined identically as in the plain model.

Remark: We briefly motivate the reasons behind choosing $\mathcal{A}$ as computationally unbounded adversary with polynomially bounded query complexity. Consider a world where "random oracle plus PSPACE" oracle is provided. A computationally bounded adversary in that oracle world shall correspond to an unbounded computational power adversary with polynomially bounded query complexity in the random oracle world. Therefore, we define ε-IND-CDP with respect to such adversaries because we shall exhibit such an adversary to show the separation from one-way functions. Note that we allow the adversary $\mathcal{A}$ to perform additional queries to the random oracle, because, in the computational setting, a computationally bounded adversary can perform additional queries to the one-way function itself.

We shall use the following definition on "closeness to product distribution."

Definition 8 (Close to Product Distribution). *A joint distribution* $(\mathbf{X}, \mathbf{Y})$ *is* ρ*-close to product distribution if* $\boldsymbol{\Delta}((\mathbf{X}, \mathbf{Y}), \mathbf{X} \times \mathbf{Y}) \leqslant \rho$*. Here,* $\mathbf{X}$ *and* $\mathbf{Y}$ *are the respective marginal distributions.*

3.2 Imported Results

The crux of the information theoretic bounds derived by [15] was the leveraging of an important Markov-chain property of the distribution of the next-message function of parties in the information theoretic setting. More specifically, the next message sent by a party is *solely* a deterministic function of its current view. Then, using the Markov chain property of protocols in the information theoretic plain model, it is easy to conclude that if the views of both parties were independent before protocol execution, they remain independent conditioned on the public transcript $m^{(n)}$. For any private inputs $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$, the following *protocol compatibility constraint* can be obtained directly:

$$\Pr[m^{(n)}|x, y] \cdot \Pr[m^{(n)}|x', y'] = \Pr[m^{(n)}|x', y] \cdot \Pr[m^{(n)}|x, y']$$

We begin with the observation that this constraint is not guaranteed a-priori in the information theoretic random oracle world. Intuitively, the views of both parties may be correlated via the common random oracle and not just the transcript. However, there are algorithms which query the random oracle polynomially many times to obtain independent views [23, 1, 5, 16, 27]. The state of the art (where parties have private inputs) is due to [27], from where we import the following theorem.

Imported Theorem 1 (Independence of Views in RO World [27]). *Given any two-party* (ℓ, n) *protocol* Π *(where parties have private inputs), there exists a public query strategy* Eve_ρ *which performs at most* $\mathsf{poly}(n\ell/\rho)$ *queries such that in the augmented protocol* $\Pi^+ := (\Pi, Eve_\rho)$*, with probability* $(1 - \rho)$ *over* $V_E \sim \mathbf{V}_E$*, we have: For all* $(x, y) \in \mathcal{X} \times \mathcal{Y}$*, if* $\Pr[x, y|V_E] > \rho$*, then* $(\mathbf{V}_A, \mathbf{V}_B|V_E, x, y)$ *is* ρ*-close to product distribution, i.e.*

$$\boldsymbol{\Delta}((\mathbf{V}_A, \mathbf{V}_B|V_E, x, y), (\mathbf{V}_A|V_E, x) \times (\mathbf{V}_B|V_E, y)) \leqslant \rho$$

3.3 Impossibility in the RO World

Instead of a key agreement enabling oracle, if we just have a random oracle, it suffices to show the following lemma:

Lemma 1 (Key Lemma for RO-Separation). *Suppose* $f \in \{\mathsf{AND}, \mathsf{XOR}\}$*. Consider any* $\varepsilon > 0$*,* $\alpha \in [\alpha_\varepsilon^{(f)} + \sigma, \alpha_\varepsilon^*]$ *and (positive) decreasing* δ*. If there exists an* (ε, δ)-$\mathsf{IND\text{-}CDP}$ α*-accurate protocol for* f *in the information theoretic random oracle world, then there exists a public query strategy* Eve_ρ *with query complexity* $\mathsf{poly}(n\ell/\rho)$*, where* $\rho = \sigma^2\varepsilon/\exp(2\varepsilon)$*, such that in the augmented protocol* $\Pi^+ := (\Pi, Eve_\rho)$*, (at least) one of the following is true:*

1. *There exists* $(\hat{y}, \hat{y}', \hat{x})$ *so that: With probability* $\tilde{\delta} = \mathsf{poly}(\sigma)$ *over* $V_E \sim \mathbf{V}_E$ *we have:*
$$\Pr[V_E|\hat{x}, \hat{y}] > \exp(\varepsilon) \cdot \Pr[V_E|\hat{x}, \hat{y}'] + \delta' \Pr[V_E] \ ,$$
where $\delta' = \mathsf{poly}(\sigma)$.
2. *There exists* $(\hat{x}, \hat{x}', \hat{y})$ *so that: With probability* $\tilde{\delta} = \mathsf{poly}(\sigma)$ *over* $V_E \sim \mathbf{V}_E$ *we have:*
$$\Pr[V_E|\hat{y}, \hat{x}] > \exp(\varepsilon) \cdot \Pr[V_E|\hat{y}, \hat{x}'] + \delta' \Pr[V_E] \ ,$$
where $\delta' = \mathsf{poly}(\sigma)$.

Proof Overview: Let p_{V_E} denote the probability of obtaining public transcript V_E over the sample space. Let $p_{V_E|x,y}$ denote the probability of obtaining public transcript V_E from Π, when inputs of Alice and Bob are $x \in \mathcal{X}$, $y \in \mathcal{Y}$.

We first observe that if some input occurs with very low probability, then ε-IND-CDP can be trivially broken. Therefore, we can directly invoke Imported Theorem 1 such that Eve_ρ generates a close-to product distribution on the views of both parties with high probability. This gives an approximate protocol compatibility constraint on most transcripts.

Next, we observe that if the views of parties are nearly independent, then with high probability, for any inputs $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$ the distributions $p_{V_E|x,y} \cdot p_{V_E|x',y'}$ and $p_{V_E|x,y'} \cdot p_{V_E|x',y}$ must be close. We obtain the following equation (refer full version for proof),

$$p_{V_E|x,y} \cdot p_{V_E|x',y'} = p_{V_E|x,y'} \cdot p_{V_E|x',y} \pm 96\rho p_{V_E}^2$$

Next, using the differential privacy constraint we mimic the proof of Goyal et al. [15] to obtain that for some transcript V_E, for some adjacent (x, y, y'), there are $\hat{\delta} = \mathsf{poly}(\sigma)$ transcripts such that for $\delta' = \mathsf{poly}(\sigma)$:

$$p_{V_E|x,y} > \lambda p_{V_E|x,y'} + \delta' p_{V_E}$$

Using averaging arguments, it is possible to show the existence of a tuple $(\hat{y}, \hat{y}', \hat{x})$ or $(\hat{x}, \hat{x}', \hat{y})$ satisfying the conditions of the lemma. □

4 Separation from Key-agreement Protocols

For $\varepsilon > 0$ differential privacy parameter, suppose $\alpha \in [\alpha_\varepsilon^{(f)} + 1/\mathsf{poly}(\kappa), \alpha_\varepsilon^*]$. In this section, we shall show that, for such choices of α, there exists an oracle relative to which public-key encryption exists but ε-IND-CDP α-accurate protocols for boolean f do not exist. It suffices to show this result for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$. This is done by showing an impossibility result in the key agreement world against information theoretic adversaries but with polynomially bounded query complexity. However, we shall show existence of an adversary who can break the ε-IND-CDP.

Note that public-key encryption is equivalent to 2-round key-agreement protocols and hence this separation translates into a separation of non-trivial (ε, δ)-differentially private protocols for AND or XOR from (any round) key-agreement protocols.

4.1 Notations and Definitions

We give some notation and definitions. These definitions were introduced in [28].

Oracle Classes

Image-testable Random Oracle Class. This is the set $\mathbb{O}_\kappa$ consisting of all possible pairs of correlated oracles $O \equiv (R, T)$ of the form:

- $R : \{0,1\}^\kappa \to \{0,1\}^{3\kappa}$ which is a *length-tripling (injective) random oracle.*
- $T : \{0,1\}^{3\kappa} \to \{0,1\}$ which is a *test oracle* defined by: $T(\beta) = 1$ if there exists $\alpha \in \{0,1\}^\kappa$ such that $R(\alpha) = \beta$; otherwise $T(\beta) = 0$.

The length of a query uniquely determines whether it is a query to the R oracle (called R-query) or to the T oracle (called T-query).

Keyed version of Image-testable Random Oracle Class. Given a set $\mathbb{K}$ of keys, consider oracle $O^{(\mathbb{K})}$ such that for every $k \in \mathbb{K}$, $O^{(k)} \in \mathbb{O}_k$ (the class of image-testable random oracles). A query is parsed as $\langle k, q \rangle$, the answer to which is $O^{(k)}(q)$. Let $\mathbb{O}_k^{(\mathbb{K})}$ denote the set of all possible oracles $O^{(\mathbb{K})}$. Then, $\mathbb{O}_k^{(\mathbb{K})}$ is the keyed version of the class of image-testable random oracles.

Public Key Encryption Oracle Class. We define a class of "PKE-enabling" oracles, from [28]. With access to this oracle, a semantically secure PKE scheme can be readily constructed, yet we shall show that it does not give (ε, δ)-IND-CDP protocols with any better than information theoretic accuracy. This oracle, called $\mathbb{PKE}_\kappa$ is a collection of oracles $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Test}_1, \mathsf{Test}_2, \mathsf{Dec})$ defined as follows:

- Gen: A length-tripling injective random oracle $\{0,1\}^\kappa \to \{0,1\}^{3\kappa}$ that takes as input a secret key sk and returns the corresponding public key pk, i.e., $\mathsf{Gen}(sk) = pk$. A public key pk is *valid* only if it is in the range of Gen.
- Enc: A collection of keyed length-tripling injective random oracles, with keys in $\{0,1\}^{3\kappa}$. For each $pk \in \{0,1\}^{3\kappa}$, the oracle implements a random injective function $\{0,1\}^\kappa \to \{0,1\}^{3\kappa}$. When queried on any (possibly invalid) random public key pk, the oracle provides the corresponding ciphertext $c \in \{0,1\}^{3\kappa}$.
- Test_1: This is a function that tests if a public key pk is *valid*, that is, it returns 1 if and only if pk is in the range of Gen
- Test_2: This is a function that tests if a public key and ciphertext pair is *valid*, i.e., it returns 1 if and only if c is in the range of the Enc oracle keyed by pk.
- Dec: This is the decryption oracle, $\{0,1\}^\kappa \times \{0,1\}^{3\kappa} \to \{0,1\}^\kappa \cup \{\perp\}$, which takes as input a secret key, ciphertext pair and returns the unique m, such that $\mathsf{Enc}(\mathsf{Gen}(\mathsf{sk}), \mathsf{m}) = c$. If such an m does not exist, it returns $\perp$.

We note that Enc produces ciphertexts for public key pk irrespective of whether there exists sk satisfying Gen $(sk) = pk$. This is crucial because we want the key set $\mathbb{K}$ to be defined independent of the Gen oracle.

We also note that $\mathbb{PKE}_\kappa$ without Dec is exactly the same as the image-testable random oracle $\mathbb{O}_k^{(\mathbb{K})}$, with $\mathbb{K} = \{0,1\}^{3\kappa} \cup \{\perp\}$. This fact will be used very crucially in the sections that follow, where we compile out the Decryption oracle and work with the resulting image-testable random oracle $\mathbb{O}_k^{(\mathbb{K})}$.

Our Setting. We will consider *private-input randomized two party protocols Π*, such that Alice and Bob have access to a common oracle PKE_κ. As in the plain model, parties send messages to each other in alternate rounds, starting with Alice in the first round. However, they have (private) access to a the common PKE_κ oracle consisting of $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Test}_1, \mathsf{Test}_2, \mathsf{Dec})$.

The views with respect to the $\mathbb{PKE}_\kappa$ oracle remain the same as views in the random oracle world. Our underlying sample space in the random oracle world is the joint distribution over Alice-Bob views when $r_A, r_B \sim \mathbf{U}$ and $\mathsf{PKE}_\kappa \sim \mathbb{PKE}_\kappa$.

We use the definition of (ε, δ)-IND-CDP protocols in the oracle world and accuracy of protocols as introduced in previous section.

4.2 Compiling out the Decryption Oracle

Using the query techniques of [28], for any arbitrarily small γ, it is possible to construct an $(\varepsilon + \gamma, \delta + \gamma)$ differentially private protocol with accuracy $\alpha - \gamma$, that uses only the family of image testable random oracles $\mathbb{O}_k^{(\mathbb{K})}$ oracle from an (ε, δ) differentially private protocol that uses the $\mathbb{PKE}_k$ oracle.

Imported Theorem 2 (Decryption Queries are Useless [28]). *Suppose Π is an (ℓ, n) (ε, δ)-differentially private α-accurate protocol for f in the $\mathbb{PKE}_\kappa$ oracle world. For every $\gamma > 0$, there exists a protocol Π' in the $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Test}_1, \mathsf{Test}_2)$ oracle world which is an $(\varepsilon + \gamma, \delta + \gamma)$ differentially private $(\alpha - \gamma)$-accurate $(\mathsf{poly}(n\ell/\gamma), n)$ protocol for f.*

4.3 Impossibility in ITRO World

Recall that the PKE-oracle without the decryption oracle is in fact a collection of keyed image-testable random oracles, where the key-set is $\mathbb{K} = \{0,1\}^{3\kappa} \cup \{\perp\}$. We import the following result of eavesdropper strategy:

Imported Theorem 3 (Independence of Views in ITRO World [28]). *For any key-set $\mathbb{K}$ and any (ℓ, n) protocol Π (where parties have private inputs), there exists a public query strategy Eve_ρ which performs at most $\mathsf{poly}(\ell/\rho)$ queries such that in the augmented protocol $\Pi^+ := (\Pi, Eve_\rho)$, the following holds over the views of Eve_ρ, when $V_E \sim \mathbf{V}_E$, with probability at least $(1 - \rho)$: For all $(x, y) \in \mathcal{X} \times \mathcal{Y}$, if $\Pr[x, y | V_E] > \rho$ then $(\mathbf{V}_A, \mathbf{V}_B | V_E, x, y)$ is ρ-close to product distribution, i.e.*

$$\boldsymbol{\Delta}\left((\mathbf{V}_A, \mathbf{V}_B | V_E, x, y), (\mathbf{V}_A | V_E, x) \times (\mathbf{V}_B | V_E, y)\right) \leqslant \rho$$

This gives us exactly the same independence characterization as Section 3.3, and we can obtain the following Lemma for the ITRO world (analogously to the random oracle world).

Lemma 2 (Key Lemma for ITRO-Separation). *Suppose $f \in \{\mathsf{AND}, \mathsf{XOR}\}$. Consider any $\varepsilon > 0$, $\alpha \in [\alpha_\varepsilon^{(f)} + \sigma, \alpha_\varepsilon^*]$ and (positive) decreasing δ. For any key-set $\mathbb{K}$, if there exists an (ε, δ)-IND-CDP α-accurate protocol for f in the image-testable random oracle world with respect to key-set $\mathbb{K}$, then there exists a public*

query strategy Eve_ρ with query complexity $\mathsf{poly}(n\ell/\rho)$, where $\rho = \sigma^2\varepsilon/\exp(2\varepsilon)$, such that in the augmented protocol $\Pi^+ := (\Pi, Eve_\rho)$, (at least) one of the following is true:

1. *There exists $(\hat{y}, \hat{y}', \hat{x})$ so that: With probability $\tilde{\delta} = \mathsf{poly}(\sigma)$ over $V_E \sim \mathbf{V}_E$ we have:*
$$\Pr[V_E|\hat{x}, \hat{y}] > \exp(\varepsilon) \cdot \Pr[V_E|\hat{x}, \hat{y}'] + \delta' \Pr[V_E] \ ,$$
where $\delta' = \mathsf{poly}(\sigma)$.
2. *There exists $(\hat{x}, \hat{x}', \hat{y})$ so that: With probability $\tilde{\delta} = \mathsf{poly}(\sigma)$ over $V_E \sim \mathbf{V}_E$ we have:*
$$\Pr[V_E|\hat{y}, \hat{x}] > \exp(\varepsilon) \cdot \Pr[V_E|\hat{y}, \hat{x}'] + \delta' \Pr[V_E] \ ,$$
where $\delta' = \mathsf{poly}(\sigma)$.

4.4 Impossibility in Key Agreement World

To prove Theorem 1, it suffices to show the following Lemma:

Lemma 3 (Key Lemma for KA-Separation). *Suppose $f \in \{\mathsf{AND}, \mathsf{XOR}\}$. Consider any $\varepsilon > 0$, $\alpha \in [\alpha_\varepsilon^{(f)} + \sigma, \alpha_\varepsilon^*]$ and (positive) decreasing δ. If there exists an (ε, δ)-$\mathsf{IND\text{-}CDP}$ α-accurate protocol for f in the $\mathbb{PKE}_\kappa$ world, then for $\gamma = \sigma^3$, the corresponding protocol Π' as defined in Imported Theorem 2 is an $(\varepsilon+\gamma, \delta+\gamma)$-$\mathsf{IND\text{-}CDP}$ $(\alpha-\gamma)$-accurate $(\mathsf{poly}(n\ell/\gamma), n)$ protocol in $\mathbb{O}_k^{(\mathbb{K})}$, where $\mathbb{K} = \{0,1\}^{3^\kappa} \cup \{\perp\}$. Then, there exists a public query strategy Eve_ρ with query complexity $\mathsf{poly}(n\ell/\gamma\rho)$, where $\rho = \sigma^2\varepsilon/\exp(2\varepsilon)$, such that in the augmented protocol $\Pi'^+ := (\Pi', Eve_\rho)$, $\delta + \gamma > \gamma^{5/6}$.*

Proof Overview: Note that we can use Imported Theorem 2 to compile any given two-party (ε, δ)-$\mathsf{IND\text{-}CDP}$ (ℓ, n) protocol Π in the key agreement world with accuracy $\alpha > \alpha_\varepsilon^{(\mathsf{AND})} + \sigma$ for the AND function (resp. $\alpha > \alpha_\varepsilon^{(\mathsf{XOR})} + \sigma$ for the XOR function), to an $(\varepsilon+\gamma, \delta+\gamma)$-$\mathsf{IND\text{-}CDP}$ (ℓ, n) protocol Π' with accuracy $(\alpha - \gamma)$ in the ITRO world (which closely mimics the RO world).

In fact, while moving from the key agreement to the ITRO world, there is a γ-loss in protocol accuracy and a corresponding (say γ') increase in maximal achievable accuracy. These parameters can be carefully tied to σ such that setting $\gamma + \gamma' = \sigma^6$, helps obtain $\delta + \gamma > \gamma^{5/6}$, thereby giving $\delta = poly(\sigma)$ transcripts violating the differential privacy constraint.

In fact, we can show (refer full version) that if $\sigma = 1/\mathsf{poly}(\kappa)$, it is possible to construct an adversary that breaks $(\varepsilon+\gamma)$-$\mathsf{IND\text{-}CDP}$ of the $(\varepsilon+\gamma, \delta+\gamma)$-$\mathsf{IND\text{-}CDP}$ (ℓ, n) protocol Π' in the ITRO world, with accuracy $(\alpha - \gamma)$ according to Definition 2. This gives a contradiction and completes the proof. □

References

1. Boaz Barak and Mohammad Mahmoody. Merkle puzzles are optimal - an $O(n^2)$-query attack on any key exchange from a random oracle. In Halevi [17], pages 374–390.

2. Amos Beimel, Kobbi Nissim, and Eran Omri. Distributed private data analysis: Simultaneously solving how and what. In David Wagner, editor, *CRYPTO*, volume 5157 of *Lecture Notes in Computer Science*, pages 451–468. Springer, 2008.
3. Benny Chor and Eyal Kushilevitz. A zero-one law for boolean privacy (extended abstract). In Johnson [24], pages 62–72.
4. Jean-Sébastien Coron, Jacques Patarin, and Yannick Seurin. The random oracle model and the ideal cipher model are equivalent. In 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*, pages 1–20. Springer, 2008.
5. Dana Dachman-Soled, Yehuda Lindell, Mohammad Mahmoody, and Tal Malkin. On black-box complexity of optimally-fair coin-tossing. In *Theory of Cryptography Conference - TCC 2011*, 2011.
6. Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In Frank Neven, Catriel Beeri, and Tova Milo, editors, *PODS*, pages 202–210. ACM, 2003.
7. Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *ICALP (2)*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.
8. Cynthia Dwork. A firm foundation for private data analysis. *Commun. ACM*, 54(1):86–95, 2011.
9. Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In Serge Vaudenay, editor, *EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 486–503. Springer, 2006.
10. Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer, 2006.
11. Cynthia Dwork and Kobbi Nissim. Privacy-preserving datamining on vertically partitioned databases. In Matthew K. Franklin, editor, *CRYPTO*, volume 3152 of *Lecture Notes in Computer Science*, pages 528–544. Springer, 2004.
12. Rosario Gennaro, Yael Gertner, Jonathan Katz, and Luca Trevisan. Bounds on the efficiency of generic cryptographic constructions. *SIAM J. Comput.*, 35(1):217–246, 2005.
13. Yael Gertner, Sampath Kannan, Tal Malkin, Omer Reingold, and Mahesh Viswanathan. The relationship between public key encryption and oblivious transfer. In *FOCS*, pages 325–335. IEEE Computer Society, 2000.
14. 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.
15. Vipul Goyal, Ilya Mironov, Omkant Pandey, and Amit Sahai. Accuracy-privacy tradeoffs for two-party differentially private protocols. In Ran Canetti and Juan A. Garay, editors, *CRYPTO (1)*, volume 8042 of *Lecture Notes in Computer Science*, pages 298–315. Springer, 2013.
16. Iftach Haitner, Eran Omri, and Hila Zarosim. Limits on the usefulness of random oracles. *Theory of Cryptography Conference (TCC, to appear)*, 2013.
17. Shai Halevi, editor. *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, volume 5677 of *Lecture Notes in Computer Science*. Springer, 2009.

18. Johan Håstad. Pseudo-random generators under uniform assumptions. In Ortiz [34], pages 395–404.
19. Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudo-random generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.
20. Thomas Holenstein, Robin Künzler, and Stefano Tessaro. Equivalence of the random oracle model and the ideal cipher model, revisited. In *STOC*, 2011.
21. Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions (extended abstracts). In Johnson [24], pages 12–24.
22. Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *FOCS*, pages 230–235. IEEE, 1989.
23. Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In Johnson [24], pages 44–61.
24. David S. Johnson, editor. *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing, 15-17 May 1989, Seattle, Washington, USA*. ACM, 1989.
25. Jeff Kahn, Michael E. Saks, and Clifford D. Smyth. A dual version of Reimer's inequality and a proof of Rudich's conjecture. In *IEEE Conference on Computational Complexity*, pages 98–103, 2000.
26. Jonathan Katz and Chiu-Yuen Koo. On constructing universal one-way hash functions from arbitrary one-way functions. *IACR Cryptology ePrint Archive*, 2005:328, 2005.
27. Mohammad Mahmoody, Hemanta K. Maji, and Manoj Prabhakaran. Limits of random oracles in secure computation. In *ITCS*, 2014.
28. Mohammad Mahmoody, Hemanta K. Maji, and Manoj Prabhakaran. On the power of public-key encryption in secure computation. In *TCC*, 2014.
29. Ueli M. Maurer, Renato Renner, and Clemens Holenstein. Indifferentiability, impossibility results on reductions, and applications to the random oracle methodology. In Naor [32], pages 21–39.
30. Andrew McGregor, Ilya Mironov, Toniann Pitassi, Omer Reingold, Kunal Talwar, and Salil P. Vadhan. The limits of two-party differential privacy. In *FOCS*, pages 81–90. IEEE Computer Society, 2010.
31. Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil P. Vadhan. Computational differential privacy. In Halevi [17], pages 126–142.
32. Moni Naor, editor. *Theory of Cryptography, First Theory of Cryptography Conference, TCC 2004, Cambridge, MA, USA, February 19-21, 2004, Proceedings*, volume 2951 of *Lecture Notes in Computer Science*. Springer, 2004.
33. Moni Naor and Moti Yung. Universal one-way hash functions and their cryptographic applications. In Johnson [24], pages 33–43.
34. Harriet Ortiz, editor. *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*. ACM, 1990.
35. Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In Naor [32], pages 1–20.
36. John Rompel. One-way functions are necessary and sufficient for secure signatures. In Ortiz [34], pages 387–394.
37. Steven Rudich. *Limits on the Provable Consequences of One-way Functions*. PhD thesis, University of California at Berkeley, 1988.