Black-Box Separations for Differentially Private Protocols*

Dakshita Khurana[†] Hemanta K. Maji[†] Amit Sahai[†]

November 23, 2014

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] have demonstrated 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 particular accuracy gap can possibly be reduced for general 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* 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 [BM09, HOZ13, MMP14a, MMP14b]; and consequently translate information theoretic impossibilities into black-box separation results.

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

[*] ©IACR 2014. This article is the full version of the paper submitted by the authors to the IACR and to Springer-Verlag on 14 Sept, 2014. The version published by Springer-Verlag is available at http://dx.doi.org/10.1007/978-3-662-45608-8_21

[†] UCLA and Center for Encrypted Functionalities. dakshita, hmaji, sahai@cs.ucla.edu. 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.

Contents

1 Introduction **1**

1.1 Our Contribution . . . 2

1.2 Related Work . . . 5

1.3 Technical Outline . . . 6

2 Preliminaries **9**

3 Separation from One-way Functions **10**

3.1 Notations and Definitions . . . 10

3.1.1 Two-party protocols in the Random Oracle World . . . 11

3.1.2 (ε, δ)-IND-CDP in the Random Oracle Model . . . 11

3.2 Imported Results . . . 12

3.3 Impossibility in the RO World . . . 12

4 Separation from Key-agreement Protocols **14**

4.1 Notations and Definitions . . . 14

4.1.1 Oracle Classes . . . 14

4.1.2 Our Setting . . . 15

4.2 Compiling out the Decryption Oracle . . . 15

4.3 Impossibility in ITRO World . . . 16

4.4 Impossibility in Key Agreement World . . . 16

References **18**

A Preliminaries **21**

A.1 Two-party Differentially private Protocols . . . 21

A.2 Black-box Separations . . . 21

B Outline of Black-box Separation **23**

C Technical Results **25**

C.1 Property A . . . 25

C.2 Near Protocol Compatibility . . . 25

C.3 Bad Transcripts are Significant in number . . . 26

C.4 Bounded slopes for Key Agreement . . . 32

C.5 Key Agreement Impossibility - Details . . . 32

1 Introduction

Differential privacy [Dwo06] 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 [MPRV09]. 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. [MMP$^+$10, GMPS13]). 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 [GMPS13], 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 [GMPS13], 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 [GMW87] 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 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?"

[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.

Goyal et al. [GMPS13] 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 [RTV04] 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 [IR89]. 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 [MMP14a, MMP14b].

Our work is reminiscent of, but also quite different from, the work of Haitner et al. [HOZ13], who proved that the information theoretic impossibility of accurate distributed differentially private evaluation of the inner-product functionality [MMP$^+$10] could be extended to a black-box separation result from one-way functions. Our results are different both qualitatively and technically: Qualitatively, our results differ in that they apply to the wide class of all boolean functions where the output of the function is sensitive to both parties' inputs. Furthermore, we show separations from key-agreement protocols as well. Moreover, our separation results for extremely simple binary functions like AND and XOR show that differentially private distributed computation even of very simple functions may also require powerful computational assumptions. At a technical level, a crucial ingredient of our proofs is the recently developed toolset of [MMP14a, MMP14b] which deal with *private inputs of parties* even in presence of the "idealized key-agreement oracle," while Haitner et al. [HOZ13] *adapt* the analysis of McGregor et al. [MMP$^+$10] 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 the distributed differentially private computation of boolean functions.

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. [GMPS13] showed that, in the information theoretic setting, $f = \mathsf{AND}$ can only be computed ε-differentially privately up to accuracy $\alpha_\varepsilon^{(\mathsf{AND})} {:=} \frac{\lambda(\lambda^2+\lambda+2)}{(\lambda+1)^3}$. Similarly, when $f = \mathsf{XOR}$ the maximal achievable accuracy is $\alpha_\varepsilon^{(\mathsf{XOR})} {:=} \frac{(\lambda^2+1)}{(\lambda+1)^2}$. Note that $\alpha_\varepsilon^{(\mathsf{XOR})} < \alpha_\varepsilon^{(\mathsf{AND})} < \alpha_\varepsilon^*$, 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] [CK89], the maximal achievable accuracy is bounded by:

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

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. [GMPS13] 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 an weaker variant of differential privacy, namely *computational differential privacy* (see Section 2).

Informal Theorem 1. *For any boolean f and privacy threshold $\varepsilon > 0$, there exists a constant c 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 [GMPS13], if $\alpha \leqslant \alpha_\varepsilon^{(f)}$, 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 $\alpha \geqslant \alpha_\varepsilon^{(f)} + 1/\mathsf{poly}(\kappa)$ (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 $\alpha_\varepsilon^{(\mathsf{AND})}$ and $\alpha_\varepsilon^{(\mathsf{XOR})}$. 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 [IR89, GGKT05].

[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).

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 $\delta \geqslant \alpha_\varepsilon^{(f)} + \sigma$ 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 [MMP14a]. For more details, refer to Imported Theorem 1.

This impossibility result easily translates into a fully black-box separation result as defined in [RTV04]. This translation of impossibility in the random-oracle model into a black-box separation is using techniques introduced in [IR89, GKM$^+$00, BM09, DLMM11, HOZ13, MMP14a].

Informal Theorem 2 (Separation from One-way Functions)**.** *For $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, $\varepsilon > 0$ and $\alpha \geqslant \alpha_\varepsilon^{(f)} + 1/\mathsf{poly}(\kappa)$, where κ is the statistical 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 [ILL89, Hås90, HILL99] and digital signatures/universal one-way hash functions [NY89, Rom90, KK05]. Moreover, it is also applicable to other computational primitives like ideal-ciphers [CPS08, HKT11] (which are indifferentiable [MRH04] from random oracles) and one-way permutations (which themselves cannot be based on one-way function [Rud88, KSS00]).

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 [GKM$^+$00].

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. $\mathsf{Gen}(sk) \mapsto pk$. 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. $\mathsf{Enc}(m; pk) \mapsto c$. 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. $\mathsf{Dec}(c, sk) \mapsto m$ or $\perp$.

This oracle is too powerful and yields a semi-honest secure protocol for oblivious-transfer (see discussion in [GKM$^+$00]). 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) [MMP14b].

To tackle the decryption oracle, we follow the technique introduced by [MMP14b]. 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 [MMP14b] 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 statistical 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 [Dwo11, Dwo06, DMNS06, DN04, DN03] 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. [DKM$^+$06] obtained distributed differential privacy via SFE and secure noise generation. Subsequently, [BNO08] studied trade-offs between distributed privacy and SFE. A computational relaxation of differential privacy was defined by Mironov et al. [MPRV09], that would help improve the range of achievable accuracies while still maintaining this relaxed notion of privacy.

A gap in the maximal obtainable accuracy of differentially private protocols, between the client-server and distributed settings, was first observed and explored by McGregor et al. [MMP$^+$10]. They demonstrated such gaps for specific large functions such as the inner product and hamming distance, under natural notions of accuracy specific to these functions. Very recently, Goyal et al. [GMPS13] 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 [IL89] showed that most non-trivial cryptographic primitives imply existence of one-way functions. Subsequently, it turned out that several primitives like pseudorandom generators [ILL89, Hås90] and digital signatures/universal one-way hash functions [NY89, Rom90] 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 [IR89] protocols and semi-honest secure oblivious-transfer protocol [GKM$^+$00] 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. [RTV04] formally defined (several variants of) black-box separations. And Gertner et al. [GGKT05] 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 [MMP14a, MMP14b]. 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. [HOZ13] show that the information theoretic impossibility of evaluating the inner-product functionality both differentially privately and accurately [MMP+10], 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 private α-accurate protocols for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$ in the distributed setting [GMPS13], and 2) Recent progress in black-box separation techniques as introduced in [BM09, HOZ13, MMP14a, MMP14b]. Our separation from key-agreement protocols especially relies on the recent results of [MMP14b]. 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. [GMPS13]. 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 the differential privacy constraint 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. [IR89, BM09, DLMM11, HOZ13, MMP14a]. 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 [MMP14a].

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 adjacent x and x'. 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 $y \in \mathcal{Y}$.

Therefore, 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 $\alpha \leqslant \alpha_{\varepsilon,\rho}^{(f)}$, which can be made arbitrarily close to $\alpha_{\varepsilon}^{(f)}$ 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 Appendix C.4). When $\alpha = \alpha_{\varepsilon}^{(f)} + \sigma$, 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. Thus, we defer our main theorem statement to the following section.

Separation from key-agreement. We show a separation from public-key encryption, which is equivalent to a 2-round key-agreement protocol. Separation from a 2-round key-agreement protocol 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 [GKM+00]) as shown in [MMP14b].

Our result depends on two technical results proven in [MMP14b]. 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 [MMP14b]. 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 $\alpha_\varepsilon^{(\mathsf{AND})} + \sigma$, where $\sigma = 1/\mathsf{poly}(n)$, then 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 detailed definitions). Intuitively, $\delta = 0$ corresponds to the previously discussed notion of ε-differential privacy. 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 $\alpha_\varepsilon^{(f)}$ as follows:*

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

Then for any $\alpha \geqslant \alpha_\varepsilon^{(f)} + \sigma$, 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 $\alpha \leqslant \alpha_\varepsilon^{(f)}$.

Note that the negative result rules out fully-BB constructions of ε indistinguishability-based computationally differentially private (ε-IND-CDP) α-accurate protocols with $\alpha > \alpha_\varepsilon^{(f)}$, 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. Due to lack of space we defer the full proofs to Appendix B.

2 Preliminaries

We introduce some important definitions in this section. For more details refer to Appendix A.

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.*

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 [MPRV09]. We note that this indistinguishability based definition is weaker than the simulation based one (SIM-CDP privacy [MPRV09]). 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 [GMPS13] 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 [GMPS13]. 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. [RTV04]. 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.

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

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-$\mathsf{IND\text{-}CDP}$ 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-$\mathsf{IND\text{-}CDP}$.

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 $O \xleftarrow{\$} \mathbb{O}_\kappa$.

3.1.1 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* $\Pi^{+}{:=}(\Pi, \mathit{Eve})$ *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 $\Pi^{+}{:=}(\Pi, \mathrm{Eve})$. 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.

3.1.2 (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.

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.

Accuracy is defined identically as in the plain model.

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* $\mathbf{\Delta}\left((\mathbf{X}, \mathbf{Y}), \mathbf{X} \times \mathbf{Y}\right) \leqslant \rho$*. Here,* $\mathbf{X}$ *and* $\mathbf{Y}$ *are the respective marginal distributions.*

3.2 Imported Results

The crux of the information-theoretic bounds derived by [GMPS13] 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 [IR89, BM09, DLMM11, HOZ13, MMP14a]. The state of the art (where parties have private inputs) is due to [MMP14a], from where we import the following theorem.

Imported Theorem 1 (Independence of Views in RO World [MMP14a])**.** *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.*

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

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 (Refer Appendix B for a complete proof):

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)-IND-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 the inputs of Alice and Bob are $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

We first observe that if some input occurs with very low probability, then ε-IND-CDP can be trivially broken (See Appendix C.1 for a formal proof). 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 a near 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. That is,

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

For a detailed proof, see Lemma 4 in Appendix C.2.

Now, mimicing the proof of Goyal et al. [GMPS13], we use differential privacy to obtain the following equation for a transcript V_E, for all adjacent (x, x', y, y'):

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

Using the modified protocol compatibility constraint, we can obtain a bound on the minimum number of public transcripts $\tilde{\delta}$ that violate this condition in the augmented protocol Eve_ρ for $\rho = \varepsilon\sigma^2/\exp(2\varepsilon)$ and δ. This bound is significant, that is, $\tilde{\delta} \geqslant \varepsilon\sigma^2 - \frac{\delta'\lambda}{8}$. For a detailed derivation, refer to Appendix C.3.

Such significant "bad" transcripts give a contradiction as explained in Appendix B. □

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 for our separation. These definitions were introduced in [MMP14b].

4.1.1 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:

- **Length-tripling Random Oracles** $R : \{0,1\}^\kappa \mapsto \{0,1\}^{3\kappa}$ which is a length-tripling (injective) random oracle
- **Test Oracles** $T : \{0,1\}^{3\kappa} \mapsto \{0,1\}$ 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 or to the T oracle. These queries are, respectively, called R-queries and T-queries.

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 [MMP14b]. 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 $\{0,1\}^\kappa \mapsto \{0,1\}^{3\kappa}$ 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: This is 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 $\{0,1\}^\kappa \mapsto \{0,1\}^{3\kappa}$. When queried with any (possibly invalid) random public key pk, this 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, $\{0,1\}^\kappa \times \{0,1\}^{3\kappa} \mapsto \{0,1\}^\kappa \cup \{\perp\}$, 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 $\mathbb{O}_k^{(\mathbb{K})}$.

4.1.2 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 [MMP14b], 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. We import a theorem from their result:

Imported Theorem 2 (Decryption Queries are Useless [MMP14b])**.** *Suppose* Π *is an* (ℓ, n)*-protocol in the* $\mathbb{PKE}_\kappa$ *oracle world which is* (ε, δ)*-differentially private* α*-accurate protocol for* f *in the* $\mathbb{PKE}_\kappa$ *world. For every* $\gamma > 0$*, there exists exists a protocol* Π' *which is an* $(\varepsilon + \gamma, \delta + \gamma), (\mathsf{poly}(n\ell/\gamma), n)$*-differentially private* $(\alpha - \gamma)$*-accurate protocol for* f *in the* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Test}_1, \mathsf{Test}_2)$ *oracle world.*

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 [MMP14b])**.** *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.*

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

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)*-*$\mathsf{IND\text{-}CDP}$ 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 worldAppendix B).

However, while moving from the key agreement to the ITRO world, there is a γ-loss in protocol accuracy and a corresponding (bounded, say γ') increase in maximal achievable accuracy (Refer Appendix C.4). However, these parameters can be carefully tied to σ such that setting $\gamma + \gamma' = \sigma^6$ helps obtain a contradiction when $\sigma = 1/\mathsf{poly}(\kappa)$.

Specifically, in the ITRO analysis, the fraction of "bad" transcripts $\tilde{\delta}$, so crucial to the violation of $(\varepsilon + \gamma)$-$\mathsf{IND\text{-}CDP}$ in Appendix B (that is, transcripts which do not satisfy Property A defined in Appendix C.1) for $\delta' = \varepsilon\sigma^2$, reduces by $2\gamma + \gamma'$(refer Appendix C.5.) But, since $\gamma + \gamma' = \sigma^6$, this property is still violated for $\delta' = \varepsilon\sigma^2$ on more than $\tilde{\delta} = \sigma^3$ fraction of the transcripts. Therefore, $\delta + \gamma \geqslant \delta'\tilde{\delta} = \sigma^5 > \gamma^{5/6}$, and thus δ cannot be $\mathsf{negl}(\kappa)$.

In fact, we can show 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

[BM09] 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 [Hal09], pages 374–390. 1, 4, 6, 7, 12

[BNO08] 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. 5

[CK89] Benny Chor and Eyal Kushilevitz. A zero-one law for boolean privacy (extended abstract). In Johnson [Joh89], pages 62–72. 3, 24

[CPS08] 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. 4

[DKM$^+$06] 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. 5

[DLMM11] 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. 4, 7, 12

[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, *TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer, 2006. 5

[DN03] 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. 5

[DN04] 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. 5

[Dwo06] 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. 1, 5

[Dwo11] Cynthia Dwork. A firm foundation for private data analysis. *Commun. ACM*, 54(1):86–95, 2011. 5

[GGKT05] 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. 3, 6, 23

[GKM$^+$00] 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. 4, 5, 7

[GMPS13] 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. 1, 2, 3, 5, 6, 9, 12, 13

[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. 1

[Hal09] 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, 20

[Hås90] Johan Håstad. Pseudo-random generators under uniform assumptions. In Ortiz [Ort90], pages 395–404. 4, 5

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999. 4

[HKT11] Thomas Holenstein, Robin Künzler, and Stefano Tessaro. Equivalence of the random oracle model and the ideal cipher model, revisited. In *STOC*, 2011. 4

[HOZ13] Iftach Haitner, Eran Omri, and Hila Zarosim. Limits on the usefulness of random oracles. *Theory of Cryptography Conference (TCC, to appear)*, 2013. 1, 2, 4, 6, 7, 12

[IL89] Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *FOCS*, pages 230–235. IEEE, 1989. 5

[ILL89] Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions (extended abstracts). In Johnson [Joh89], pages 12–24. 4, 5

[IR89] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In Johnson [Joh89], pages 44–61. 2, 3, 4, 5, 7, 12

[Joh89] 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. 18, 19, 20

[KK05] 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. 4

[KSS00] 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. 4

[MMP$^+$10] 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. 1, 2, 5, 6

[MMP14a] Mohammad Mahmoody, Hemanta K. Maji, and Manoj Prabhakaran. Limits of random oracles in secure computation. In *ITCS*, 2014. 1, 2, 4, 6, 7, 12

[MMP14b] Mohammad Mahmoody, Hemanta K. Maji, and Manoj Prabhakaran. On the power of public-key encryption in secure computation. In *TCC*, 2014. 1, 2, 4, 5, 6, 7, 8, 14, 15, 16, 24

[MPRV09] Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil P. Vadhan. Computational differential privacy. In Halevi [Hal09], pages 126–142. 1, 5, 9

[MRH04] Ueli M. Maurer, Renato Renner, and Clemens Holenstein. Indifferentiability, impossibility results on reductions, and applications to the random oracle methodology. In Naor [Nao04], pages 21–39. 4

[Nao04] 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. 20

[NY89] Moni Naor and Moti Yung. Universal one-way hash functions and their cryptographic applications. In Johnson [Joh89], pages 33–43. 4, 5

[Ort90] Harriet Ortiz, editor. *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*. ACM, 1990. 19, 20

[Rom90] John Rompel. One-way functions are necessary and sufficient for secure signatures. In Ortiz [Ort90], pages 387–394. 4, 5

[RTV04] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In Naor [Nao04], pages 1–20. 2, 4, 6, 10, 21

[Rud88] Steven Rudich. *Limits on the Provable Consequences of One-way Functions.* PhD thesis, University of California at Berkeley, 1988. 4

A Preliminaries

For $\theta \geqslant 0$, we use $a = b \pm \theta$ to denote $|a - b| \leqslant \theta$.

A.1 Two-party Differentially private Protocols

In a *private-input randomized two-party protocol* Π, Alice has private input $x \in \mathcal{X}$ and samples local randomness $r_A \sim \mathbf{U}$; and Bob has private input $y \in \mathcal{Y}$ and samples local randomness $r_B \sim \mathbf{U}$. Alice sends the first message in the protocol; and thereafter they send messages in alternate rounds. That is, messages $m_1, m_3, \dots$ are sent by Alice in rounds $1, 3, \dots$, respectively, and Bob sends $m_2, m_4, \dots$ in rounds $2, 4, \dots$, respectively. The partial transcript till round i is $m_1 \dots m_i$ and is represented by $m^{(i)}$. The *view* of Alice at the end of round i is represented by $V_A^{(i)} \equiv (x, r_A, m^{(i)})$; and, similarly, the view of Bob at the end of round i is represented by $V_B^{(i)} \equiv (y, r_B, m^{(i)})$. Suppose the protocol Π has n rounds, then at the culmination of n rounds Alice and Bob output z_A and z_B as, respective, deterministic function of their views $V_A^{(n)}$ and $V_B^{(n)}$. The protocol Π is a *symmetric-output* protocol if $z_A = z_B$; and this (common) output denoted by $\mathsf{out}_\Pi(m^{(n)})$. We shall write the random variable corresponding to it as out_Π.

Our underlying sample space is the joint distribution over complete Alice-Bob views when $r_A, r_B \sim \mathbf{U}$. The random variable $\mathbf{V}_A^{(i)}$, thus, represents the following distribution: Sample a joint view $(V_A, V_B) \sim (\mathbf{V}_A^{(n)}, \mathbf{V}_B^{(n)})$ and output $V_A^{(i)}$, which is the restriction of V_A till end of round i. Further, we also write $\mathbf{V}_A^{(n)}$ as $\mathbf{V}_A$, if the round complexity of the protocol is implicit.

A.2 Black-box Separations

For completeness, we include the definition of fully black-box reduction between primitives as introduced by Reingold et al. [RTV04]. A primitive is described by a pair $\mathcal{P} = \langle F_\mathcal{P}, R_\mathcal{P} \rangle$, such that $F_\mathcal{P}$ is a set of functions $g : \{0,1\}^* \mapsto \{0,1\}^*$, of which at least one is computable by a PPT machine. $R_\mathcal{P}$ is a relation over pairs $\langle g, M \rangle$ of a function $g \in F_\mathcal{P}$ and a machine M, such that the machine M $\mathcal{P}$-breaks g if $\langle g, M \rangle \in R_\mathcal{P}$. An efficient and secure implementation of the primitive $\mathcal{P}$ is a function $g \in F_\mathcal{P}$ which is efficiently computable by a PPT machine, moreover no PPT machine $\mathcal{P}$-breaks g. A primitive exists if it has an efficient and secure implementation.

Definition 9 (Fully Black-box Reduction [RTV04])**.** *There exists a* fully black-box *(fully-BB) reduction from a primitive $\mathcal{P} = \langle F_\mathcal{P}, R_\mathcal{P} \rangle$ to $\mathcal{Q} = \langle F_\mathcal{Q}, R_\mathcal{Q} \rangle$, if there exist PPT oracle machines G (read as, construction) and S (read as, security reduction) such that:*

- *Correctness: For every implementation $g \in F_\mathcal{Q}$ it holds that $G^g \in F_\mathcal{P}$.*
- *Security: For every implementation $g \in F_\mathcal{Q}$ and every machine $\mathcal{A}$, if $\mathcal{A}$ $\mathcal{P}$-breaks G^g then $S^{\mathcal{A}^g, g}$ $\mathcal{Q}$-breaks g.*

Consider, for example, some fully-BB reduction from primitive $\mathcal{P}$ to $\mathcal{Q}$, where $\mathcal{P}$ is (ε, δ)-DP for a function f and $\mathcal{Q}$ is one-way functions.

- For any implementation g of one-way functions, G^g implements an (ε, δ)-DP for f, and

- For any implementation of one-way function and every adversary $\mathcal{A}$ such that $\mathcal{A}^g$ α-breaks the ε-DP experiment, the reduction $S^{A^g,g}$ breaks the one-way function g.

Thus, to show a *fully black-box separation* of $\mathcal{P}$ from $\mathcal{Q}$, it suffices to consider a construction G such that $G^g \in \mathcal{P}$ for every $g \in \mathcal{Q}$ and exhibit an $g^* \in \mathcal{Q}$ and an adversary $\mathcal{A}$ which $\mathcal{P}$-breaks G^{g^*} but $S^{A^{g^*},g^*}$ cannot $\mathcal{Q}$-break g^*.

Consider the implementation one-way functions using random oracles. Then, in our example, to show a fully-BB separation of two-party (ε, δ)-differentially private α-accurate protocol for f and one-way functions, it suffices to show the following. Consider any construction G which implements two-party (ε, δ)-differentially private α-accurate protocol for f with oracle access to a random oracle and then exhibit an *efficient* (uniform) adversary $\mathcal{A}$ such that $\mathcal{A}^g$ breaks the (ε, δ)-differential privacy, if α-accuracy holds. Since S is efficient, the query complexity of $S^{\mathcal{A}^\circ,\circ}$, i.e. the composition of the reduction S with the adversary $\mathcal{A}$, is polynomial in the length of random oracle queries. And we know that any algorithm with polynomial query complexity cannot invert a random oracle, except at negligible fraction of outputs. So, there exists g^* such that the (ε, δ)-differential privacy of the protocol is broken and g^* cannot be inverted by the reduction. Thus, showing a fully black-box separation.

B Outline of Black-box Separation

We present a high level overview of the simpler black-box separation from one-way functions. Our oracle is a random oracle which maps κ-bit strings to κ-bit strings. Additionally, a PSPACE oracle is also provided. The PSPACE oracle ensures that all parties have unbounded computational power. It is known that even in presence of a PSPACE oracle, a random oracle cannot be inverted by any algorithm with $\mathsf{poly}(\kappa)$ query complexity at more than a negligible fraction of its image [GGKT05]. Thus, with respect to the "random-oracle + PSPACE" oracle one-way functions exist.

It suffices to consider the information theoretic setting and prove our results in the random oracle model. Henceforth, we shall assume that parties have unbounded computational power but $\mathsf{poly}(\kappa)$ query complexity.

Suppose we are provided with an (ε, δ)- computationally differentially private α-accurate protocol for $f = \mathsf{AND}$. (An analogous argument will also work for $f = \mathsf{XOR}$.) Let $\alpha = \alpha_\varepsilon^{(f)} + \sigma$, where $\sigma = 1/\mathsf{poly}(\kappa)$. We execute Eve_ρ, where $\rho = \sigma^2\varepsilon/\lambda^2$, at the end of the protocol. This ensures that with probability $(1-\rho)$ over Eve view V_E, the protocol compatibility constraint is approximately satisfied, namely

$$\Pr[V_E|x,y] \cdot \Pr[V_E|x'y'] = \Pr[V_E|xy'] \cdot \Pr[x'y] \pm 96\rho \cdot \Pr[V_E]^2$$

Next, of all of Eve's views, we consider the set of Eve views V_E where the following is satisfied: There exists a tuple (x, x', y) or (y, y', x) such that,

$$\Pr[V_E|x,y] > \exp(\varepsilon) \cdot \Pr[V_E|x,y'] + \delta' \Pr[V_E] \ ,$$

or the analogous statement for (y, y', x) holds. Let $\mathcal{Z}$ denote this set of Eve's views.

We have, from Corollary 2 and Corollary 3, that with probability at least $\tilde{\delta} = \varepsilon\sigma^2 - \frac{\delta'\lambda}{8}$, a public transcript $V_E \sim \mathbf{V_E}$ lies in $\mathcal{Z}$. On any such V_E, there is a *witness* tuple (x, x', y) or (y, y', x) for which the condition mentioned above holds at V_E. Overall, of the 8 possible witnesses, we choose the most probable one. We call is the *representative witness*. Suppose $(\hat{y}, \hat{y}', \hat{x})$ has the highest overall probability, i.e. it is a witness with probability at least $\tilde{\delta}/8$.

Now consider $\mathcal{A}$ which outputs 1 only at those transcripts V_E such that

$$\Pr[V_E|\hat{x}, \hat{y}] > \exp(\varepsilon) \cdot \Pr[V_E|\hat{x}, \hat{y}'] + \delta' \Pr[V_E]$$

And $\mathcal{A}$ outputs 0 everywhere else.

We note that since $(\hat{y}, \hat{y}', \hat{x})$ is a witness with probability at least $\tilde{\delta}/8$, over transcripts $V_E \in \mathcal{Z}$, $\Pr[\mathcal{A}(V_E, x)|y]$ outputs 1 (with the input of Bob being y) is greater than $\exp(\varepsilon) \cdot \Pr[\mathcal{A}(V_E, x)|y']$ by $\delta'\tilde{\delta}/8$.

With respect to this tuple, the protocol is not $(\varepsilon, \hat{\delta})$-computationally differentially private, where $\hat{\delta} = \delta'\tilde{\delta}/8 \geqslant \frac{7\sigma^4\varepsilon^2}{64\lambda}$ and $\delta' = \frac{\sigma^2\varepsilon}{\lambda}$. Therefore, the original protocol cannot have $\delta > \hat{\delta}$. In other words, if $\alpha \geqslant \alpha_\varepsilon^{(f)} + 1/\mathsf{poly}(\kappa)$ for all $\mathsf{poly}(\kappa)$, the original protocol cannot be ε-IND-CDP.

General Functionalities f. Let f be a function where both parties' inputs influence the output. Let (x_0, x_1) and (y_0, y_1) be input pairs such that they represent an XOR minor in f. If f does not contain any XOR minor, then choose (x_0, x_1) and (y_0, y_1) to correspond to an AND minor in f. Note that, because both parties' inputs influence the output of f, it necessarily contains at least one AND or XOR minor [CK89].

For this minor, compute the representative witness as defined in the previous section. Suppose the tuple for which privacy breaks is $(\hat{y}, \hat{y}', \hat{x})$.

Now, in the computational differential privacy experiment corresponding to this tuple, note that we do *not* have access to the input $\{x_0, x_1\} \setminus \{\hat{x}\}$. So, in the computational differential privacy experiment, our adversary $\mathcal{A}$ does the following: Let η denote the size of Alice's input space $\mathcal{X}$, and $\hat{x}_i$ represent $\hat{x}$ with its i-th bit flipped, for $i \in [\eta]$. Then for every $\tilde{x} \in \{\hat{x}_1, \ldots, \hat{x}_\eta\}$, the adversary $\mathcal{A}$ runs Eve_ρ with respect to the input pairs $(\hat{x}, \tilde{x})$ and $(\hat{y}, \hat{y}')$; and outputs 1 if and only if the following condition is satisfied for (at least) one value of $\tilde{x}$:

$$\Pr[V_E | \hat{x}, \hat{y}] > \exp(\varepsilon) \cdot \Pr[V_E | \hat{x}, \hat{y}'] + \delta' \Pr[V_E]$$

So, we can claim that if the original protocol is (ε, δ)-differentially private and α-accurate, then $\delta > \frac{7\sigma^4\varepsilon^2}{8\lambda}$. This contradicts the ε-computational differential privacy of the original protocol.

Separation from Key-agreement. The separation from Key-agreement protocols follows a similar outline. Suppose we are provided with an (ε, δ)-$\mathsf{IND\text{-}CDP}$ α-accurate (ℓ, n)-protocol Π for f in the $\mathbb{PKE}_\kappa$ world, where $\varepsilon > 0$, $\delta > 0$, $\alpha \in [\alpha_\varepsilon^{(f)} + \sigma, \alpha_\varepsilon^*]$ and $\sigma = 1/\mathsf{poly}(\kappa)$. We apply a result of [MMP14b] to obtain the following result. We construct a (related) protocol Π' in the $\mathbb{O}_\kappa^{(\mathbb{K})}$ world, where $\mathbb{K} = \{0,1\}^{3\kappa} \cup \{\perp\}$, which is $(\varepsilon + \gamma, \delta + \gamma)$-$\mathsf{IND\text{-}CDP}$ $(\alpha - \gamma)$-accurate $(\mathsf{poly}(n\ell/\gamma), n)$-protocol for f, where $\gamma = \sigma^6$.

Now, we apply a result of [MMP14b] which says that collection of image-testable random oracles mimics several properties on random oracle. In particular, there exists a dependence killing public query algorithm. Finally, we show that (similar to the previous separation from one-way functions) that there exists a constant $c \in (0,1)$ such that: $\delta + \gamma \geqslant \gamma^{1-c}$. The adversary performs $\mathsf{poly}(n\ell/\sigma\gamma)$ additional queries to the oracle. This implies that δ cannot be $\mathsf{negl}(\kappa)$.

C Technical Results

For brevity, we define some additional notation that will help prove our impossibility result.

Let $p_{V_E|x,y}$ denote the probability of obtaining transcript V_E from protocol Π, when the inputs of Alice and Bob are $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Let $p_{\mathsf{out}|x,y}$ denote the probability of obtaining boolean output $\mathsf{out} \in \{0,1\}$ when the inputs of Alice and Bob are $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Let $p_{x,y|V_E}$ denote the probability of Alice and Bob having private inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, conditioned on transcript V_E.

C.1 Property A

For any (λ, δ'), a transcript V_E is said to satisfy Property A on (λ, δ') if, for all inputs x, x', y, y'

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

.

Claim 1. *Consider a transcript V_E which satisfies Property A on (λ, δ'). Then, for any adjacent pair of inputs x, x', y, y' and any $X, Y \in \{x, x'\} \times \{y, y'\}$,*

$$\frac{1}{\lambda^2 + 3\lambda + 4} \leqslant p_{X,Y|V_E} \leqslant \frac{\lambda^2 + \lambda + 1}{(\lambda+1)^2}$$

Proof. It follows directly from Property A, that for adjacent inputs (x, x', y, y'):

$$p_{x,y|V_E} \leqslant \lambda p_{x,y'|V_E} + (\delta'/4), \text{ and}$$

$$p_{x,y'|V_E} \leqslant \lambda p_{x',y'|V_E} + (\delta'/4)$$

Let p denote the probability $p_{x,y|V_E}$, then we have for any $p_{X,Y|V_E}$ where $X, Y \in \{x, x'\} \times \{y, y'\}$:

$$\frac{1}{(\lambda+1)^2 + (\delta'/4p)\cdot(\lambda+3)} \leqslant p_{X,Y|V_E} \leqslant \frac{\lambda^2 + (\delta'/4p)\cdot(\lambda'+1)}{(\lambda+1)^2 + (\delta'/4p)\cdot(\lambda+3)}$$

The claim follows by substituting $0 \leqslant (\delta'/4) \leqslant p_{x,y|V_E}$ [5]. □

C.2 Near Protocol Compatibility

The following lemma shows 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 $p_{V_E|x,y'} \cdot p_{V_E|x',y}$ are close.

Definition 10 ($(1-\beta)$ - Protocol Compatibility)**.** *A public transcript V_E generated by protocol Π is said to be $(1-\beta)$ protocol compatible, if the following constraint holds for all $(x, x') \in \mathcal{X}$ and $(y, y') \in \mathcal{Y}$:*

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

[5] Note that we will work with arbitrarily small values of δ, therefore, it is okay to assume that $\delta' \leqslant 1/(\lambda+3)^2$

Lemma 4. *In the augmented protocol* $\Pi^+ := (\Pi, Eve_\rho)$*, if the views of parties are* $(1-\rho)$*-close to product distribution conditioned on a public transcript* $V_E \sim \mathbf{V_E}$*, then* V_E *is* $(1-96\rho)$ *- protocol compatible (according to Definition 10).*

Proof. In our randomized experiment, with inputs (X, Y) picked uniformly at random in $(\{x, x'\} \times \{y, y'\})$,

$$\Pr[X, Y] = 1/4 \tag{2}$$

Now consider a public transcript V_E, such that the views of both parties are independent conditioned on V_E in the augmented protocol Π^+,

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

Restricting the views of both parties to just their inputs, we obtain

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

Using the definition of statistical distance along with equation (2), we have for transcript V_E,

$$\begin{aligned} \Pr[x, y | V_E] &= \Pr[x | V_E] \cdot \Pr[y | V_E] \pm \rho, \text{ equivalently} \\ p_{V_E | x, y} &= \frac{\Pr[V_E | x] \Pr[V_E | y]}{\Pr[V_E]} \pm \frac{\rho \Pr[V_E]}{\Pr[x] \Pr[y]} \end{aligned}$$

This gives us directly, $p_{V_E|x,y} \cdot p_{V_E|x',y'} = p_{V_E|x,y'} p_{V_E|x',y} \pm 96\rho \cdot p^2_{V_E}$ for transcript V_E such that the views of parties are ρ-independent conditioned on V_E. □

C.3 Bad Transcripts are Significant in number

Lemma 5. *In the augmented protocol* $\Pi^+ := (\Pi, Eve_\rho)$ *if*

- *With probability* $(1-\rho)$*, a public transcript* $V_E \sim \mathbf{V_E}$ *is such that the views of parties are* ρ*-independent conditioned on* V_E*, and,*
- *With probability* $(1-\tilde{\delta})$*, a public transcript* $V_E \sim \mathbf{V_E}$ *is such that Property A holds for* V_E*,*

Then for all choices of adjacent *inputs* $(x, x') \in \mathcal{X}$ *and* $(y, y') \in \mathcal{Y}$*, the following constraints hold, for* $\hat{\rho} = 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$ *and* $\lambda = \exp(\varepsilon)$*:*

$$\begin{aligned} p_{z|x,y'} + p_{z|x',y} &\leqslant \lambda^{-1} \cdot p_{z|x,y} + \lambda \cdot p_{z|x',y'} + \hat{\rho} \\ p_{z|x,y'} + p_{z|x',y} &\leqslant \lambda \cdot p_{z|x,y} + \lambda^{-1} \cdot p_{z|x',y'} + \hat{\rho} \\ p_{z|x,y} + p_{z|x',y'} &\leqslant \lambda^{-1} \cdot p_{z|x,y'} + \lambda \cdot p_{z|x',y} + \hat{\rho} \\ p_{z|x,y'} + p_{z|x',y} &\leqslant \lambda \cdot p_{z|x,y} + \lambda^{-1} \cdot p_{z|x',y'} + \hat{\rho} \end{aligned}$$

(Recall that $p_{z|x,y}$ *denotes the probability of obtaining symmetric output* z *from* Π *conditioned on the choice of private input being* $x, y \in \mathcal{X} \times \mathcal{Y}$*.)*

Proof. Consider a transcript V_E, then Property A gives us that with probability $(1 - \tilde{\delta})$,

$$p_{V_E|x',y}, p_{V_E|x,y'} \in [\lambda^{-1} \cdot (p_{V_E|x,y} - \delta' p_{V_E}),\ \lambda \cdot p_{V_E|x',y'} + \delta' p_{V_E}]$$

Since $(\lambda^{-1} \cdot (p_{V_E|x,y} - \delta' p_{V_E})) \leqslant p_{V_E|x,y'} \leqslant (\lambda \cdot p_{V_E|x',y'} + \delta' p_{V_E})$ and probabilities are positive, we can rewrite this as a quadratic inequality,

$$\begin{gathered} {p_{V_E|x,y'}}^2 - \left(\lambda^{-1} \cdot p_{V_E|x,y} + \lambda \cdot p_{V_E|x',y'} + \delta' p_{V_E}(1 - \lambda^{-1})\right) p_{V_E|x,y'} + \\ p_{V_E|x,y} \cdot p_{V_E|x',y'} + \delta' p_{V_E}(\lambda^{-1} \cdot p_{V_E|x,y} - p_{V_E|x',y'}) - \lambda^{-1} \cdot\ (delta')^2 p_{V_E}^2 \leqslant 0 \end{gathered}$$

Let $\tilde{\rho} = 96\rho$, then Lemma 4 gives us that with probability $(1 - \rho)$, $V_E \sim \mathbf{V_E}$ is $(1 - \tilde{\rho})$ protocol compatible. Then, also using (2),

$$\begin{aligned} p_{V_E|x',y} &= \frac{p_{V_E|x,y} \cdot p_{V_E|x',y'}}{p_{V_E|x,y'}} \pm \frac{p_{V_E}^2 \tilde{\rho}}{p_{V_E|x,y'}} && (3) \\ &= \frac{p_{V_E|x,y} \cdot p_{V_E|x',y'}}{p_{V_E|x,y'}} \pm \frac{p_{V_E} \tilde{\rho}}{4 p_{x,y'|V_E}} && (4) \end{aligned}$$

Rewriting the quadratic inequality, dividing by $p_{V_E|x,y'}$ and using (4), we have with probability $(1 - \rho - \tilde{\delta})$:

$$\begin{aligned} & p_{V_E|x,y'} + p_{V_E|x',y} \\ & \leqslant \lambda^{-1} \cdot_{V_E|x,y} + \lambda \cdot p_{V_E|x',y'} + \left(\frac{p_{V_E} \tilde{\rho}}{4 p_{x,y'|V_E}} \right) + \delta' p_{V_E} \left((1 - \lambda^{-1}) + \lambda^{-1} \cdot \delta' p_{V_E} + \left(\frac{\lambda^{-1} p_{x,y|V_E}}{p_{x,y'|V_E}} - \frac{p_{x',y'|V_E}}{p_{x,y'|V_E}} \right) \right) \end{aligned}$$

Using Claim 1, we have that with probability $(1 - \rho - \tilde{\delta})$, a public view $V_E \sim \mathbf{V_E}$ is such that:

$$\begin{aligned} p_{V_E|x,y'} + p_{V_E|x',y} &\leqslant \lambda^{-1} \cdot p_{V_E|x,y} + \lambda \cdot p_{V_E|x',y'} + p_{V_E} \left(\frac{\tilde{\rho}(\lambda^2 + 3\lambda + 4)}{4} + \frac{\delta'(\lambda^2 + 7\lambda + 4)}{2\lambda} \right) && (5) \\ &\leqslant \lambda^{-1} \cdot p_{V_E|x,y} + \lambda \cdot p_{V_E|x',y'} + p_{V_E} \left(2\tilde{\rho}\lambda^2 + 6\delta'\lambda \right) && (6) \end{aligned}$$

Let $\Pr[z]$ denote the probability of obtaining output z from function f on inputs (X, Y) sampled uniformly in $\{x, x'\} \times \{y, y'\}$. In this paper, we study Boolean protocols such that $z \in \{0, 1\}$. In the specific case of AND and XOR functionalities, we know that $\Pr[z] \geqslant 1/4$.

Consider an (ε, δ)-$\mathsf{IND\text{-}CDP}$ protocol with respect to functionality $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, then we observe that $\delta' = 1/2$ is trivially achievable by sampling outputs uniformly at random in $\{0, 1\}$. Then, a loose lower bound on the probability of obtaining output z in an (ε, δ)-$\mathsf{IND\text{-}CDP}$ protocol for $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, is $1/8$.

Note that for any *symmetric* protocol, the output is a deterministic function of the transcript generated by the underlying protocol Π.

Let us denote the output corresponding to a particular view V_E by out_{V_E}. Then, $\Pr[\mathsf{out}_{V_E} = z] \geqslant 1/8$, if the (ε, δ)-$\mathsf{IND\text{-}CDP}$ protocol is with respect to functionality $f \in \{\mathsf{AND}, \mathsf{XOR}\}$.

Let $\mathbb{S}_z$ denote the set of all transcripts V_E, such that $\mathsf{out}_{V_E} = z$ and equation (6) holds.

Then, $\Pr[V_E \in \mathbb{S}_z] \geqslant (1 - 8(\rho + \tilde{\delta}))$. Then,

$$\begin{aligned}\sum_{V_E \in \mathbb{S}_z} p_{V_E|x,y'} + \sum_{V_E \in \mathbb{S}_z} \lambda^{-1} \cdot p_{V_E|x',y} \\ \leqslant \sum_{V_E \in \mathbb{S}_z} p_{V_E|x,y} + \sum_{V_E \in \mathbb{S}_z} \lambda \cdot p_{V_E|x',y'} + \sum_{V_E \in \mathbb{S}_z} p_{V_E}(2\tilde{\rho}\lambda^2 + 6\delta'\lambda)\end{aligned}$$

Let $\mathbb{T}_z$ be the set of all transcripts V_E, such that $\mathsf{out}_{V_E} = z$ and equation (6) does not hold. Then, for all $V_E \in \mathbb{T}_z$,

$$p_{V_E|x,y'} + \lambda^{-1} \cdot p_{V_E|x',y} \leqslant 2p_{V_E}$$

Then,

$$\sum_{V_E \in \mathbb{T}_z} p_{V_E|x,y'} + \sum_{V_E \in \mathbb{T}_z} \lambda^{-1} \cdot p_{V_E|x',y} \leqslant \sum_{V_E \in \mathbb{T}_z} 4p_{V_E}(1 + \lambda^{-1}) \leqslant 64(\rho + \tilde{\delta})$$

Summing up over all transcripts V_E such that $\mathsf{out}_{V_E} = z$, we obtain

$$p_{z|x,y'} + \lambda^{-1} \cdot p_{z|x',y} \leqslant p_{z|x,y} + \lambda \cdot p_{z|x',y'} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$$

All other inequalities can be proven similarly.

□

Lemma 6. *Let Π be an (ε, δ)-IND-CDP protocol for $f = \mathsf{AND}$, with accuracy α. Then, for any ρ, δ', in the augmented protocol $\Pi^+ := (\Pi, Eve_\rho)$ (as per Definition 5), suppose with probability $1 - (\rho + \tilde{\delta})$ over $V_E \sim \mathbf{V}_E$:*

1. *For all $x, y \in \{0, 1\}$, we have $\Pr[x, y|V_E] \geqslant \rho$,*
2. *$\mathbf{\Delta}((\mathbf{V}_A, \mathbf{V}_B|V_E), (\mathbf{V}_A|V_E) \times (\mathbf{V}_B|V_E)) \leqslant \rho$, and*
3. *Property A holds for V_E.*

Then, $\alpha = \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}$ is upper bounded by $\alpha^{(\mathsf{AND})}_{\varepsilon} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$.

Proof. By the accuracy constraints, we have

$$p_{0|1,1} \leqslant 1 - \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}$$
$$p_{0|0,1} + p_{0|1,0} \geqslant 2\alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}$$

Using Lemma 5, we also obtain the following constraints on $p_{0|x,y}$ for adjacent $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$.

$$p_{0|0,1} + p_{0|1,0} \leqslant \lambda^{-1} \cdot p_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho}$$
$$p_{1|0,1} + p_{1|1,0} \leqslant \lambda \cdot p_{1|0,0} + \lambda^{-1} \cdot p_{1|1,1} + \hat{\rho}$$

Substituting $p_{0|0,1} + p_{0|1,0} = q$, we obtain the following three linear inequalities:

$$\begin{aligned} q &\leqslant \lambda^{-1} \cdot p_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho} && (7) \\ q &\geqslant 2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} && (8) \\ q &\geqslant 2 - (\lambda + \lambda^{-1}) + \lambda \cdot p_{0|0,0} + \lambda^{-1} \cdot p_{0|1,1} - \hat{\rho} && (9) \end{aligned}$$

We observe that the two lines bounding the half-planes specified by (8) and (9), with $(q, p_{0|0,0})$ being the free variables and $(\lambda, \alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}}, p_{0|1,1})$ the parameters, intersect at $(p^*_{0|0,0}, q^*)$ such that

$$p^*_{0|0,0} = 1 + \lambda^{-2} - 2\lambda^{-1} + \lambda^{-1}(2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} + \hat{\rho}) - \lambda^{-2} p_{0|1,1} \text{ and } q^* = 2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}}$$

Claim 2. *$(p^*_{0|0,0}, q^*)$ satisfies (7), i.e.* $2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \lambda^{-1} \cdot p^*_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho}$

Proof. Assume to the contrary, that

$$2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} > \lambda^{-1} \cdot p^*_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho} \tag{10}$$

Consider two cases:

- ○ $p_{0|0,0} \leqslant p^*_{0|0,0}$. Then

$$\begin{aligned} q &\overset{(7)}{\leqslant} \lambda^{-1} \cdot p_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho} \\ &\leqslant \lambda^{-1} \cdot p^*_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho} \\ &\overset{(10)}{<} 2\alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} \end{aligned}$$

  contradicting (8)

- ○ $p_{0|0,0} > p^*_{0|0,0}$. Then

$$q \overset{(7)}{\leqslant} \lambda^{-1} \cdot p_{0|0,0} + \lambda \cdot p_{0|1,1} + \hat{\rho} \overset{(10)}{<} \lambda^{-1} \cdot (p_{0|0,0} - p^*_{0|0,0}) + 2 \cdot \alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}}$$

  Moreover,

$$\begin{aligned} q &\overset{(9)}{\geqslant} 2 - (\lambda + \lambda^{-1}) + \lambda \cdot p_{0|0,0} + \lambda^{-1} \cdot p_{0|1,1} - \hat{\rho} \\ &= \lambda \cdot (p_{0|0,0} - p^*_{0|0,0}) + \lambda \cdot p^*_{0|0,0} + 2 - (\lambda + \lambda^{-1}) + \lambda^{-1} \cdot p_{0|1,1} - \hat{\rho} \\ &\overset{\text{def of } p^*_{0|0,0}}{=} \lambda \cdot (p_{0|0,0} - p^*_{0|0,0}) + 2 \cdot \alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}}, \end{aligned}$$

  This gives

$$\lambda \cdot (p_{0|0,0} - p^*_{0|0,0}) + 2 \cdot \alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}} < \lambda^{-1} \cdot (p_{0|0,0} - p^*_{0|0,0}) + 2 \cdot \alpha^{\textsf{(AND)}}_{\varepsilon,\rho,\delta',\tilde{\delta}}$$

  which is a contradiction since $\lambda \geqslant 1$ and $p_{0|0,0} > p^*_{0|0,0}$.

□

By substituting $p^*_{0|0,0}$ and q^* in (7), and using $p_{0|1,1} \leqslant 1 - \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}$, we obtain

$$\begin{aligned} 2\alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} &\leqslant \lambda^{-1} \cdot (1 + \lambda^{-2} - 2\lambda^{-1} + 2\alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}\lambda^{-1} + \hat{\rho}\lambda^{-1} \cdot -\lambda^{-2} \cdot p_{0|1,1}) + \lambda \cdot p_{0|1,1} + \hat{\rho} \\ &\leqslant \lambda + 1/\lambda - 2 \cdot \lambda^{-2} + \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \cdot (2\lambda^{-2} - \lambda + \lambda^{-3}) + \hat{\rho}(1 + \lambda^{-2}) \end{aligned}$$

This gives us

$$\begin{aligned} \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} &\leqslant \lambda(\lambda^2 + \lambda + 2)/(1+\lambda)^3 + \hat{\rho}(1 + \lambda^{-2})/(1+\lambda)^3(\lambda - 1) \\ &\leqslant \alpha^{(\mathsf{AND})}_{\varepsilon} + \hat{\rho}(1 + \lambda^{-2})/(1+\lambda)^3(\lambda - 1) \end{aligned}$$

Substituting $\hat{\rho}$ and λ, we obtain

$$\alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \alpha^{(\mathsf{AND})}_{\varepsilon} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda \qquad (11)$$

□

Corollary 2. *Let Π be an (ε, δ)-$\mathsf{IND\text{-}CDP}$ protocol for $f = \mathsf{AND}$, with accuracy $\alpha^{(\mathsf{AND})}_{\varepsilon} + \sigma$. Then, there exist ρ, δ' and an augmented protocol $\Pi^{+} {:=} (\Pi, Eve_\rho)$ (by Imported Theorem 1), such that with probability $1 - (\rho + \tilde{\delta})$ over $V_E \sim \mathbf{V}_E$:*

1. *For all $x, y \in \{0,1\}$, we have $\Pr[x, y|V_E] \geqslant \rho$,*
2. *$\mathbf{\Delta}\left((\mathbf{V}_A, \mathbf{V}_B|V_E), (\mathbf{V}_A|V_E) \times (\mathbf{V}_B|V_E)\right) \leqslant \rho$, and*
3. *Property A holds for V_E.*

And, $\tilde{\delta} \geqslant \varepsilon\sigma^2 - \frac{\delta'\lambda}{8}$.

Proof. Using Claim 1 and Imported Theorem 1, we can set $\rho = \frac{\sigma^2 \cdot \varepsilon}{\lambda^2}$, such that

$$\alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \alpha^{(\mathsf{AND})}_{\varepsilon} + 64\sigma^2 + \frac{16\tilde{\delta} + 2\delta'\lambda}{\varepsilon}$$

meaning that

$$\begin{aligned} \tilde{\delta} &\geqslant \varepsilon\left(\frac{\sigma}{16} - 4\sigma^2\right) - \frac{\delta'\lambda}{8} \\ &\geqslant \varepsilon\sigma^2 - \frac{\delta'\lambda}{8} \end{aligned}$$

□

Lemma 7. *Let Π be an (ε, δ)-$\mathsf{IND\text{-}CDP}$ protocol for $f = \mathsf{XOR}$, with accuracy α. Then, for any ρ, δ', in the augmented protocol $\Pi^{+} {:=} (\Pi, Eve_\rho)$ (as per Definition 5), suppose with probability $1 - (\rho + \tilde{\delta})$ over $V_E \sim \mathbf{V}_E$:*

1. *For all $x, y \in \{0,1\}$, we have $\Pr[x, y|V_E] \geqslant \rho$,*

2. *$\mathbf{\Delta}\left((\mathbf{V}_A, \mathbf{V}_B|V_E), (\mathbf{V}_A|V_E) \times (\mathbf{V}_B|V_E)\right) \leqslant \rho$, and*

3. *Property A holds for V_E.*

Then, $\alpha = \alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}}$ is upper bounded by $\alpha^{(\mathsf{XOR})}_{\varepsilon} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$.

Proof. Using Lemma 5, we directly obtain the following bounds on the values of $p_{0|x,y}$ for adjacent $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$.

$$p_{0|0,0} + p_{0|1,1} \leqslant \lambda^{-1} \cdot p_{0|1,0} + \lambda \cdot p_{0|0,1} + \hat{\rho}$$
$$p_{0|0,0} + p_{0|1,1} \leqslant \lambda \cdot p_{0|1,0} + \lambda^{-1} \cdot p_{0|0,1} + \hat{\rho}$$

Summing the inequalities and dividing by two, we have

$$p_{0|0,0} + p_{0|0,1} \leqslant \frac{\lambda^2+1}{2\lambda}(p_{0|1,0} + p_{0|0,1}) + \hat{\rho} \tag{12}$$

Observe that

$$\begin{aligned}\alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} &= \min(p_{0|0,0}\ , 1 - p_{0|0,1}\ , 1 - p_{0|1,0}\ , p_{0|1,1}) \\ &\leqslant \min\left(\frac{p_{0|0,0} + p_{0|1,1}}{2}, 1 - \frac{p_{0|1,0} + p_{0|0,1}}{2}\right)\end{aligned}$$

Denoting $\frac{p_{0|0,0}+p_{0|1,1}}{2}$ by x and $\frac{p_{0|1,0}+p_{0|0,1}}{2}$ by y, we have

$$\begin{aligned}\alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} &= \min(x, 1-y) \\ &\overset{(12)}{\leqslant} \min(x, 1 + \hat{\rho} - \frac{2\lambda}{1+\lambda^2}x),\end{aligned}$$

which attains its maximal value when $x = 1 + \hat{\rho} - \frac{2\lambda}{1+\lambda^2}x$.

Solving this for x and substituting in the above expression, we have that

$$\alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \frac{1+\lambda^2}{(1+\lambda)^2}(1+\hat{\rho})$$

Again, substituting $\hat{\rho}$ and λ, we obtain

$$\alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \alpha^{(\mathsf{XOR})}_{\varepsilon} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$$

□

Corollary 3. *Let Π be an (ε, δ)-IND-CDP protocol for $f = \mathsf{XOR}$, with accuracy $\alpha^{(\mathsf{XOR})}_{\varepsilon} + \sigma$. Then, there exist ρ, δ' and an augmented protocol $\Pi^{+}{:=}(\Pi, Eve_\rho)$ (by Imported Theorem 1), such that with probability $1 - (\rho + \tilde{\delta})$ over $V_E \sim \mathbf{V}_E$:*

1. *For all $x, y \in \{0, 1\}$, we have $\Pr[x, y|V_E] \geqslant \rho$,*

2. *$\mathbf{\Delta}\left((\mathbf{V}_A, \mathbf{V}_B | V_E), (\mathbf{V}_A | V_E) \times (\mathbf{V}_B | V_E)\right) \leqslant \rho$, and*
3. *Property A holds for V_E.*

And, $\tilde{\delta} \geqslant \varepsilon\sigma^2 - \frac{\delta'\lambda}{8}$.

Proof. Using Claim 1 and Imported Theorem 1, we can set $\rho = \frac{\sigma^2 \cdot \varepsilon}{\lambda^2}$, such that

$$\alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} \leqslant \alpha^{(\mathsf{XOR})}_{\varepsilon} + 256\rho\lambda^2 + 64\tilde{\delta} + 6\delta'\lambda$$

meaning that

$$\begin{aligned} \tilde{\delta} &\geqslant \varepsilon\left(\frac{\sigma}{64} - 256\sigma^2\right) - \frac{\delta'\lambda}{8} \\ &\geqslant \varepsilon\sigma^2 - \frac{\delta'\lambda}{8} \end{aligned}$$

□

C.4 Bounded slopes for Key Agreement

Claim 3. *For all $\gamma = \frac{1}{\mathsf{poly}(\kappa)}$, there exists $\gamma' = \frac{1}{\mathsf{poly}(\kappa)}$, such that $\alpha^{(\mathsf{AND})}_{\varepsilon+\gamma,\rho,\delta',\tilde{\delta}} = \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}} + \gamma'$ and $\alpha^{(\mathsf{XOR})}_{\varepsilon+\gamma,\rho,\delta',\tilde{\delta}} = \alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}} + \gamma'$ for $\varepsilon \in [0, 1]$.*

Proof. We first note that $\alpha^{(\mathsf{AND})}_{\varepsilon}$ and $\alpha^{(\mathsf{XOR})}_{\varepsilon}$ are continuous functions with bounded slope for $\varepsilon \in [0, 1]$. This implies that for all binary boolean functions f, there exist positive constants a, b, c such that for all $\varepsilon \in [0, 1], \sigma \leqslant 1, \delta' \leqslant 1$,

$$\frac{\partial \alpha^{(\mathsf{AND})}_{\varepsilon,\rho,\delta',\tilde{\delta}}}{\partial \varepsilon} = \frac{\partial(\alpha^{(\mathsf{AND})}_{\varepsilon} + c\varepsilon\sigma^2)}{\partial \varepsilon} \in [a, b]$$

and

$$\frac{\partial \alpha^{(\mathsf{XOR})}_{\varepsilon,\rho,\delta',\tilde{\delta}}}{\partial \varepsilon} = \frac{\partial(\alpha^{(\mathsf{XOR})}_{\varepsilon} + c\varepsilon\sigma^2)}{\partial \varepsilon} \in [a, b]$$

. Then, $\gamma' \in [a\gamma, b\gamma]$. □

C.5 Key Agreement Impossibility - Details

For $f \in \{\mathsf{AND}, \mathsf{XOR}\}$, going back to $\alpha^{(f)}_{\varepsilon+\gamma,\rho,\delta',\tilde{\delta}}$ as defined in Lemma 6 and Lemma 7, we let the accuracy of the compiled protocol in ITRO world

$$\alpha - \gamma = \alpha^{(f)}_{\varepsilon+\gamma,\rho,\delta',\tilde{\delta}} = \alpha^{(f)}_{\varepsilon,\rho,\delta',\tilde{\delta}} + \gamma'$$

Note that $\hat{\delta}$ (and therefore, $\tilde{\delta}$) is also allowed to increase by γ. Set $\gamma + \gamma' \leqslant \sigma^3$ to conclude, from Equation 11, that the equation

$$p_{V_E|x,y} \leqslant \lambda p_{V_E|x,y'} + ((\varepsilon\sigma^2/\lambda) + \sigma^3) \cdot p_{V_E}$$

will no longer be satisfied for all adjacent (x, x', y, y'), on $\varepsilon\sigma^2$ fraction of public views.