Semi-Honest to Malicious Oblivious Transfer
The Black-Box Way

Iftach Haitner

Dept. of Computer Science and Applied Math., Weizmann Institute of Science, Rehovot, Israel. iftach.haitner@weizmann.ac.il

Abstract. Until recently, all known constructions of oblivious transfer protocols based on general hardness assumptions had the following form. First, the hardness assumption is used in a black-box manner (i.e., the construction uses only the input/output behavior of the primitive guaranteed by the assumption) to construct a *semi-honest* oblivious transfer, a protocol whose security is guaranteed to hold only against adversaries that follow the prescribed protocol. Then, the latter protocol is "compiled" into a (malicious) oblivious transfer using non-black techniques (a Karp reduction is carried in order to prove an NP statement in zero-knowledge).

In their recent breakthrough result, Ishai, Kushilevitz, Lindel and Petrank (STOC '06) deviated from the above paradigm, presenting a black-box reduction from oblivious transfer to enhanced trapdoor permutations and to homomorphic encryption. Here we generalize their result, presenting a black-box reduction from oblivious transfer to semi-honest oblivious transfer. Consequently, oblivious transfer can be black-box reduced to each of the hardness assumptions known to imply a semi-honest oblivious transfer in a black-box manner. This list currently includes beside the hardness assumptions used by Ishai et al., also the existence of families of dense trapdoor permutations and of non trivial single-server private information retrieval.

1 Introduction

Since most cryptographic tasks are impossible to achieve with absolute, information-theoretic security, modern cryptography tries to design protocols that are *infeasible* to break. Namely, their security is based on computational hardness assumptions. These assumptions typically come in two flavors: ***specific hardness assumptions*** like discrete log, factoring and RSA, and ***general hardness assumptions*** like the existence of one-way functions. In this paper we refer to general hardness assumptions and how they are used. Primitives assumed to carry some hardness assumption can be used to construct a provably secure cryptographic tasks in two possible ways: "black-box usage", where the construction uses only the input/output behavior of the primitive, and "non-black-box usage", where the construction uses the internal structure of the primitive, e.g., its code. The above is formalized via the notion of ***black-box reductions***. A black-box reduction from a primitive P to a primitive Q, is an efficient construction of P out of Q that ignores the internal structure of the implementation of Q and merely uses it as a "subroutine" (i.e., as a black-box). Such a reduction is ***fully-black-box*** [25] if the proof of security (showing that an adversary that breaks the implementation of P implies an efficient adversary that breaks the implementation of Q), is black-box as well. That is, the internal structure of the adversary that breaks the implementation of P is ignored. See Section 2.2 for more details.

Staring from the seminal paper of Impagliazzo and Rudich [16], a rich line of works tries to draw the border between possibility and impossibility for black-box reductions in cryptography. Currently, for most cryptographic tasks we either have a black-box reduction to a commonly believed hardness assumption, or have shown the impossibility of such a reduction. There are several important tasks, however, for which we have failed to apply the above black-box classification. Very interestingly, for most of those tasks we do have non-black-box reductions (typical examples are the reductions from oblivious transfer to semi-honest oblivious transfer [12], and from public-key encryption schemes secure against chosen cipher-text attack to semantically-secure encryption schemes [8, 20, 26]). In their recent breakthrough result, Ishai et al. [17] presented the first black-box reduction from oblivious transfer to "low-level" primitives (to homomorphic encryption and to enhanced trapdoor permutations). Yet, the question whether there exists a black-box reduction from oblivious transfer to semi-honest oblivious transfer, remained open.

A better understanding of the above might help up to resolve the intriguing question whether non-black-box techniques are superior to black-box ones also in the setting of reductions between cryptographic primitives. [1] On a more practical level, we mention that the non-black-box reductions of the above tasks are using Karp reductions for the purpose of using a (general) zero-knowledge proof/argument. Such reductions are highly inefficient and unlikely to be used in practice. Furthermore, in most cases the communication complexity in the resulting protocols depends on the complexity of computing the underlying primitive (i.e., of the trapdoor permutations), where black-box reductions, unaware of the inner structure of the underlying primitive, do not suffer from this phenomenon (see [17] for more details).

In this paper, we study the above issues w.r.t. oblivious transfer. Oblivious transfer, introduced by Rabin [24], is a fundamental primitive in cryptography and has several equivalent formulations [3, 5, 4, 6, 9, 24]. The version we study here, defined by Even, Goldreich and Lempel [9], is that of ***one-out-of-two oblivious transfer***. This version is an interactive protocol between a ***sender*** and a ***receiver***. The sender gets as an input two secret bits: σ_0 and σ_1 and the receiver gets an index $i \in \{0, 1\}$. At the end of the protocol, R learns σ_i. Informally, the security of the oblivious transfer states that the receiver does not learn σ_{1-i} and the sender does not learn i. Oblivious transfer is known to

[1] The superiority of non-black-box techniques was demonstrated by Barak [1] in the settings of zero-knowledge arguments for NP. In these settings, however, the black-box access is to the, possibly cheating, verifier and not to any underlying primitive.

imply key-agreement signing contracts protocols [2, 9, 24] and, more generally, secure multiparty computation in the presence of malicious majority [12, 18, 28]. We sometimes add the term ***malicious*** to the above definition, to differentiate it from definitions that guarantee weaker security.

1.1 Defensible Privacy

The notion of defensible privacy, introduced by Ishai et al. [17], is a natural bridging step between semi-honest privacy and fully-fledged one. Informally, a two-party protocol (A,B) is ***defensibly private*** w.r.t. A and a function f defined over the parties' inputs (denoted as (A, f)-defensibly-private), if at the end of the interaction even a cheating A^* cannot simultaneously prove that it has acted honestly (i.e., as the honest party would) and learn the value of f. [2] [17] showed how to use enhanced trapdoor permutation (or homomorphic encryption) to construct ***defensible oblivious transfer***. Where the latter is a protocol with the oblivious transfer functionally, which is defensibly-private w.r.t. to the sender and the input bit of the receiver, and w.r.t. to the receiver and the other secret of the sender. That is, it is $(\mathsf{S}, f_{\mathsf{S}})$ and $(\mathsf{R}, f_{\mathsf{R}})$ defensibly-private, where S and R stand for sender and the receiver respectively, $f_{\mathsf{S}}(\sigma_0, \sigma_1, i) \stackrel{\text{def}}{=} i$ and $f_{\mathsf{R}}(\sigma_0, \sigma_1, i) \stackrel{\text{def}}{=} \sigma_{1-i}$. [17] then show how to use such a defensible oblivious transfer to derive their main result.

1.2 Our Result

A two-party protocol (A,B) is (A, f)***-semi-honest-private***, if at the end of the interaction the semi-honest A does not learn the value of f. Our main technical contribution is the following theorem.

Theorem 1. *Let $\pi = (\mathsf{A},\mathsf{B})$ be a two-party protocol and let $f_{\mathsf{A}}, f_{\mathsf{B}} : \{0,1\}^k \times \{0,1\}^k \mapsto \{0,1\}^*$ be two functions defined over the parties' inputs. Assume that π is $(\mathsf{A}, f_{\mathsf{A}})$ and $(\mathsf{B}, f_{\mathsf{B}})$ semi-honest private. Then there exists a fully-black-box reduction from a protocol $\pi' = (\mathbb{A}, \mathbb{B})$ that has the same functionality as π and is $(\mathbb{A}, f_{\mathsf{A}})$ and $(\mathbb{B}, f_{\mathsf{B}})$ defensibly-private, to π and one-way functions.*

Since one-way functions can be black-box reduced to semi-honest oblivious transfer (see Theorem 4), we obtain the following corollary.

Corollary 1. *There exists a fully-black-box reduction from defensible oblivious transfer to semi-honest oblivious transfer.*

Combining the above with the reduction of [17] from malicious oblivious transfer to derisible one, we derive our main result.

Theorem 2. *There exists a fully-black-box reduction from oblivious transfer to semi-honest oblivious transfer.*

[2] The above generalizes the definition of [17], which was only stated w.r.t. oblivious transfer protocols.

As a corollary of Theorem 2, we have that there exists a fully-black-box reduction from oblivious transfer to each of the assumptions that known to imply semi-honest oblivious transfer in a fully-black-box manner. This list currently includes families of dense/enhanced trapdoor permutations [9, 13], homomorphic encryption [19, 27] and non-trivial single-server private-information retrieval [7]. In addition, Kilian [18] tells us that secure multiparty computation can be black-box reduced to oblivious transfer. Hence, we also have the following corollary.

Corollary 2. *There exist fully-black-box reductions from protocols for securely computing any multiparty functionality with an honest-minority and in the presence of static malicious adversaries, to semi-honest oblivious transfer.*

1.3 Our Technique - From Semi-honest to Defensible Privacy

Given a protocol $\pi = (\mathsf{A}, \mathsf{B})$ that is $(\mathsf{A}, f_{\mathsf{A}})$ and $(\mathsf{B}, f_{\mathsf{B}})$ semi-honest-private, and assuming that one-way functions exist, we create the protocol $\pi' = (\mathbb{A}, \mathbb{B})$ that is $(\mathbb{A}, f_{\mathsf{A}})$ and $(\mathbb{B}, f_{\mathsf{B}})$ defensibly-private. Our reduction is carried out in two steps. First, we create a protocol $(\mathbb{A}, \mathbb{B})$ with the same functionality as (A, B), which is $(\mathbb{A}, f_{\mathsf{A}})$-defensibly-private and $(\mathbb{B}, f_{\mathsf{B}})$-semi-honest-private. Then, we apply the same transformation on $(\mathbb{A}, \mathbb{B})$, to strengthen also the privacy w.r.t. f_{B}. In what follows we describe how to obtain the first step (the second step is analogous), but first let us describe what a commitment scheme is. In a commitment scheme the ***sender*** interacts with the ***receiver*** to commit to a private value; informally the commitment is ***binding*** if the sender cannot open the commitment into a different value than the one it had committed to, where the commitment is ***hiding*** if before the decommitment stage the receiver does not learn the committed value. Fully-black-box reductions from commitment schemes to one-way functions were given by [15, 21] and [14, 23].

In the new protocol $(\mathbb{A}, \mathbb{B})$, we embed an execution of (A, B) while using a commitment scheme in order to enforce the "defensible behavior" of $\mathbb{A}$. Let i_{A}, i_{B} and r_{A}, r_{B} be the inputs and random-coins of A and B respectively. We define $(\mathbb{A}(i_{\mathbb{A}}, r_{\mathbb{A}}), \mathbb{B}(i_{\mathbb{B}}, r^1_{\mathbb{B}}, r^2_{\mathbb{B}}))$ as follows. First, $\mathbb{A}$ commits to $(i_{\mathbb{A}}, r_{\mathbb{A}})$ using a commitment scheme, followed by $\mathbb{B}$ sending $r^1_{\mathbb{B}}$ over to $\mathbb{A}$. Then the two parties execute $(\mathsf{A}(i_{\mathbb{A}}, r_{\mathbb{A}} \oplus r^1_{\mathbb{B}}), \mathsf{B}(i_{\mathbb{B}}, r^2_{\mathbb{B}}))$, where $\mathbb{A}$ and $\mathbb{B}$ act as A and B respectively. The hiding property of the commitment scheme yields that before the embedded execution of (A, B) starts, $\mathbb{B}$ does not learn any information about the input and random-coins that $\mathbb{A}$ uses in this execution. Thus, the semi-honest privacy of $(\mathbb{A}, \mathbb{B})$ w.r.t. $\mathbb{B}$ and f_{B}, follows by the semi-honest privacy of (A, B) w.r.t. B and f_{B}. In order to prove that $(\mathbb{A}, \mathbb{B})$ is $(\mathbb{A}, f_{\mathsf{A}})$-defensibly-private, we first note that a valid defense of $\mathbb{A}$ must include a valid opening of the commitment. Thus, the binding property of the commitment scheme yields that even a dishonest $\mathbb{A}^*$ can only provide a valid defense if it has acted in the embedded execution of (A, B) as A whose input and random-coins are set to $i_{\mathbb{A}}$ and $r_{\mathbb{A}} \oplus r^1_{\mathbb{B}}$ would. Namely, if it has acted as A whose input was decided *before* the execution has started, and its random-coins are chosen *at random* would. Hence, the defensible privacy of the protocol w.r.t. $\mathbb{A}$ and f_{A}, follows by the semi-honest privacy of (A, B) w.r.t. A and f_{A}. [3]

1.4 Paper Organization

Section 2 contains the notations and definitions used in this paper. In Section 3 we present our general transformation from semi-honest privacy to defensible one (Theorem 1) and in Section 4 we use this transformation to derive our main result (Theorem 2).

2 Preliminaries

2.1 Notation

We denote by U_n the random variable uniformly chosen in $\{0,1\}^n$. Given a distribution D, we denote its support by $\mathrm{Supp}(D)$. We adopt the convention that when the same random variable occurs several times in an expression, all occurrences refer to a single sample. For example, $\Pr[f(U_n) = U_n]$ is defined to be the probability that when $x \leftarrow U_n$, we have $f(x) = x$. Given a vector v of dimension n, we denote by $v[i_1, ..., i_k]$, where $i_1, \ldots, i_k \in [n]$, the vector $(v[i_1], \ldots, v[i_k])$. A function $\mu : \mathbb{N} \to [0,1]$ is negligible, denoted $\mu = \mathrm{neg}$, if for every polynomial p we have that $\mu(n) < 1/p(n)$ for large enough n. Two distribution ensembles D_n and ξ_n are computationally-indistinguishable (denoted $D_n \approx_c \xi_n$), if no efficient algorithm distinguishes between them with more then negligible probability. Given a two-party protocol $\pi = (\mathsf{A}, \mathsf{B})$, we denote the inputs and random-coins of A and B by i_{A} and i_{B}, and by r_{A} and r_{B} respectively. We denote by $\mathrm{View}^{\pi}_{\mathsf{A}}((i_{\mathsf{A}}, r_{\mathsf{A}}), (i_{\mathsf{B}}, r_{\mathsf{B}}))$ the view of A after the execution of π on $((i_{\mathsf{A}}, r_{\mathsf{A}}), (i_{\mathsf{B}}, r_{\mathsf{B}}))$. This view consists on i_{A}, r_{A} and the messages A received thought the protocol. We denote by $\mathrm{View}^{\pi}_{\mathsf{A}}(i_{\mathsf{A}}, i_{\mathsf{B}})$, the random variable $\mathrm{View}^{\pi}_{\mathsf{A}}((i_{\mathsf{A}}, R_{\mathsf{A}}), (i_{\mathsf{B}}, R_{\mathsf{B}}))$, where R_{A} and R_{B} are uniformly chosen among all strings of the right length.

2.2 Black-Box Reductions

A reduction from a primitive P to a primitive Q consists of showing that if there exists an implementation C of Q, then there exists an implementation M_C of P. This is equivalent to showing that for every adversary that breaks M_C, there exists an adversary that breaks C. Such a reduction is ***semi-black-box*** if it ignores the internal structure of Q's implementation, and it is ***fully-black-box*** if the proof of correctness is black-box as well, i.e., the adversary for breaking Q

[3] In their construction of defensible oblivious transfer from enhanced families of trapdoor permutations, [17] are using (perfectly-binding) commitment schemes for a similar purpose. More specifically, they employ the semi-honest oblivious transfer of [9] and use a commitment scheme for forcing the receiver to sample one of the two random elements it has to choose in the permutation domain honestly, i.e., choosing it as a random output of the domain sampler.

ignores the internal structure of both Q's implementation and of the (alleged) adversary breaking P. A taxonomy of black-box reductions was provided by [25], and the reader is referred to their paper for a more complete and formal view of these notions. All the reduction considered in this paper are fully-black-box ones.

2.3 Different Notions of privacy

In the following we present the two privacy measures we use in this paper.

Semi-honest privacy

In the standard definitions of semi-honest privacy (c.f, [11]), it is required that the semi-honest party does not learn *any information* about the other party's input, save but the part it suppose to get according to the prescribed functionality. Here we present a natural relaxation to the above, defining the notion of ***semi-honest privacy w.r.t. a function***. Namely, we only require that the semi-honest party does not learn a predefined function of the parties' inputs. [4]

Definition 1 (semi-honest privacy w.r.t. a function). *Let $\pi = (\mathsf{A}, \mathsf{B})$ be a two-party protocol getting security parameter 1^n and let $f : \{0,1\}^k \times \{0,1\}^k \mapsto \{0,1\}^*$ be a function defined over the parties' inputs. We say that π is (A, f)-*semi-honest-private*, if for every efficiently samplable input $i_\mathsf{A} \in \{0,1\}^k$ it holds that*

$$(\mathrm{View}^{\pi}_{\mathsf{A}}(i_\mathsf{A}, U_k), f(i_\mathsf{A}, U_k)) \approx_c (\mathrm{View}^{\pi}_{\mathsf{A}}(i_\mathsf{A}, U_k), f(i_\mathsf{A}, U'_k))$$

Defensible privacy

Definition 2 (defense). *Let $\pi = (\mathsf{A}, \mathsf{B})$ be a two-party protocol and let t be a transcript of an interaction between some party A^* and B. We say that d is a* good defense *for t (w.r.t. A's role in π), if A whose input, including its random-coins, is set to d would have sent the same messages that A^* does in t. We use the following notations: given $v = \mathrm{View}^{(\mathsf{A}^*,\mathsf{B})}_{\mathsf{A}^*}(\cdot)$, we let $\mathrm{Defense}(v)$ be the defense that A^* locally output in the end of the interaction (set to $\perp$ is no such defense is given) and let the predicate $\mathrm{IsGoodDef}^{\pi,\mathsf{A}}(v)$ to be one if $\mathrm{Defense}(v)$ is a good defense for (the transcript embedded in) v.*

Definition 3 (defensible privacy w.r.t. a function). *Let π and f be as in Definition 1. We say that π is (A, f)-*defensibly-private*, if the following holds for every* PPT A^*:

$$\Gamma(\mathrm{View}^{(\mathsf{A}^*,\mathsf{B})}_{\mathsf{A}^*}(U_k), f(i^d_\mathsf{A}, U_k)) \approx_c \Gamma(\mathrm{View}^{(\mathsf{A}^*,\mathsf{B})}_{\mathsf{A}^*}(U_k), f(i^d_\mathsf{A}, U'_k)) \ ,$$

where $\Gamma(x, y)$ equals (x, y) if $\mathrm{IsGoodDef}^{\pi,\mathsf{A}}(x) = 1$ and equals $\perp$ otherwise, and i^d_A is the value of A's input in $\mathrm{Defense}(\mathrm{View}^{(\mathsf{A}^*,\mathsf{B})}_{\mathsf{A}^*}(U_k))$. [5]

---

[4] We have chosen to work with this weaker form of semi-honest privacy, since we have found it simpler to handle and yet strong enough when considering semi-honest oblivious transfer protocols.

[5] It immediately follows that being (A, f)-defensibly-private implies being (A, f)-semi-honest-private. In Section 3, we show that the other direction is also true.

Remark 1. It seems natural to extend the above definition to a simulation based one. Namely, a protocol is ***defensibly private*** if a party that gives a valid defense learns nothing (in the simulation sense) other than the prescribed functionality. It is then seems tempting to try to reduce the above defensible privacy to semi-honest privacy (according to [11]). Namely, to prove that any semi-honest private protocol implies a defensibly private version of this protocol. We hope to address this issue in the full version.

2.4 Oblivious transfer

Oblivious transfer is an interactive protocol between a sender, S, and a receiver, R. The sender gets as an input two secret bits: σ_0 and σ_1 and the receiver gets an index $i \in \{0,1\}$, in the end of the protocol R locally outputs a single bit. We make the following correctness requirement: for all n and all valid values of σ_0, σ_1 and i, with save but negligible probability the output of R in the interaction $(\mathsf{S}(\sigma_0, \sigma_1), \mathsf{R}(i))$ is σ_i.

Let (S, R) be a protocol that computes the oblivious transfer functionality, let $f_{\mathsf{S}}(\sigma_0, \sigma_1, i) \stackrel{\text{def}}{=} i$ and let $f_{\mathsf{R}}(\sigma_0, \sigma_1, i) \stackrel{\text{def}}{=} \sigma_{1-i}$. We say that (S, R) is a ***semi-honest [resp. defensible]*** oblivious transfer if it is $(\mathsf{S}, f_{\mathsf{S}})$ and $(\mathsf{R}, f_{\mathsf{R}})$ semi-honest private [resp. defensibly private]. The protocol (S, R) is ***(malicious) oblivious transfer*** if its computation is secure according to the *real/ideal simulation paradigm* (see [11, Chapter 7] for formal definition). The following is implicit in [17].

Theorem 3 ([17]). *There exists a fully-black-box reduction from oblivious transfer to defensible oblivious transfer.*

2.5 Commitment Schemes

A commitment scheme is a two-stage protocol between a sender and a receiver. In the first stage, called the ***commit stage***, the sender commits to a private string σ. In the second stage, called the ***reveal stage***, the sender reveals σ and *proves* that it was the value to which she committed in the first stage. We require two properties of commitment schemes. The hiding property says that the receiver learns nothing about σ in the commit stage. The binding property says that after the commit stage, the sender is bound to a particular value of σ; that is, she cannot successfully open the commitment to two different values in the reveal stage. See [10] for a more formal definition. Fully-black-box reductions from commitment schemes to one-way functions were given by [15, 21] and [14, 23].

2.6 One-way Functions

Definition 4. *Let $f : \{0,1\}^* \mapsto \{0,1\}^*$ be a polynomial-time computable function. f is* one-way *if the following is negligible for every* PPT *A,*

$$\Pr[A(1^n, U_n) \in f^{-1}(f(U_n))].$$

3 Reducing Semi-Honest Protocols to Defensible Ones

Our transformation from semi-honest privacy to defensible privacy (Theorem 1) immediately follows by applying the next lemma twice. The lemma informally states that it is possible to "upgrade" the security of a protocol w.r.t. one of its parties while maintaining the initial security w.r.t. the other party.

Lemma 1. *Let* $\pi = (\mathsf{A}, \mathsf{B})$ *be a two-party protocol and let* $f_{\mathsf{A}}, f_{\mathsf{B}} : \{0,1\}^k \times \{0,1\}^k \mapsto \{0,1\}^*$ *be two functions defined over the parties' inputs. Assume that* π *is* $(\mathsf{A}, f_{\mathsf{A}})$*-semi-honest-private and* $(\mathsf{B}, f_{\mathsf{B}})$*-*$x$*-private, where* x *stands for 'semi-honest' or 'defensibly'. Then there exists a fully-black-box reduction from a protocol* $\pi' = (\mathbb{A}, \mathbb{B})$ *that has the same functionality as* π *and is* $(\mathbb{A}, f_{\mathsf{A}})$*-defensibly-private and* $(\mathbb{B}, f_{\mathsf{B}})$*-*$x$*-private, to* π *and one-way functions.*

Proof. In the following definition of π' we are using a commitment scheme, Com. Recall that by [15, 22] and by [14, 23], there exists a fully-black-box reduction from Com to one-way functions.

Protocol 1 *[The defensible protocol* $\pi' = (\mathbb{A}, \mathbb{B})$*]*

Common input: 1^n.
$\mathbb{A}$'s inputs: $i_{\mathbb{A}} \in \{0,1\}^k$ *and* $r_{\mathbb{A}} = (r^1_{\mathbb{A}}, r^2_{\mathbb{A}})$.
$\mathbb{B}$'s inputs: $i_{\mathbb{B}} \in \{0,1\}^k$ *and* $r_{\mathbb{B}} = (r^1_{\mathbb{B}}, r^2_{\mathbb{B}}, r^3_{\mathbb{B}})$.

1. $\mathbb{A}$ *commits using* Com *to* $(i_{\mathbb{A}}, r^2_{\mathbb{A}})$*, where the security parameter of the commitment is set to* 1^n *and* $\mathbb{A}$ *and* $\mathbb{B}$ *are using the random-coins* $r^1_{\mathbb{A}}$ *and* $r^1_{\mathbb{B}}$ *respectively.*
2. $\mathbb{B}$ *sends* $r^3_{\mathbb{B}}$ *to* $\mathbb{A}$*.*
3. *The two parties execute the protocol* $(\mathsf{A}(1^n, i_{\mathbb{A}}, r^2_{\mathbb{A}} \oplus r^3_{\mathbb{B}}), \mathsf{B}(1^n, i_{\mathbb{B}}, r^2_{\mathbb{B}}))$*, with* $\mathbb{A}$ *and* $\mathbb{B}$ *acting as* A *and* B *respectively.*

Clearly π' has the same functionality as π. Lemma 2 states that π' maintains the *same* privacy w.r.t. $\mathbb{B}$ and f_{B}. The heart of our proof is in Lemma 3, where we show that π' has defensible privacy w.r.t. $\mathbb{A}$ and f_{A}.

Lemma 2. *Assume that* π *is* $(\mathsf{B}, f_{\mathsf{B}})$*-*$x$*-private, then* π' *is* $(\mathbb{B}, f_{\mathsf{B}})$*-*$x$*-private.*

Proof. We assume that π is $(\mathsf{B}, f_{\mathsf{B}})$-semi-honest-private and prove that π' is $(\mathbb{B}, f_{\mathsf{B}})$-semi-honest-private (the proof for the defensibly-private case is analogous). We first note that the hiding property of Com yields that for every $i_{\mathbb{B}} \in \{0,1\}^k$, the distribution $(\mathrm{View}^{\pi'}_{\mathbb{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U_k, i_{\mathbb{B}}))$ is computationally indistinguishable from $(\mathrm{View}^{\mathsf{Com}}_{\mathbb{B}}(0^\ell), \mathrm{View}^{\pi}_{\mathsf{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U_k, i_{\mathbb{B}}))$. By the semi-honest privacy of π w.r.t. B and f_{B}, we have that $(\mathrm{View}^{\pi'}_{\mathbb{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U_k, i_{\mathbb{B}}))$ is computationally indistinguishable from $(\mathrm{View}^{\mathsf{Com}}_{\mathbb{B}}(0^\ell), \mathrm{View}^{\pi}_{\mathsf{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U'_k, i_{\mathbb{B}})))$. Using the hiding property of Com once more, we have that $(\mathrm{View}^{\pi'}_{\mathbb{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U_k, i_{\mathbb{B}}))$ is computationally indistinguishable from $(\mathrm{View}^{\pi'}_{\mathbb{B}}(U_k, i_{\mathbb{B}}), f_{\mathsf{B}}(U'_k, i_{\mathbb{B}}))$. Namely, we have proved that π' is $(\mathbb{B}, f_{\mathsf{B}})$-semi-honest-private.

Lemma 3. *Assume that π is $(\mathsf{A}, f_\mathsf{A})$-semi-honest-private, then π' is $(\mathbb{A}, f_\mathsf{A})$-defensibly-private.*

Proof. Assume toward a contradiction the existence of an efficient adversary $\mathbb{A}^*$ and a distinguisher $\mathbb{D}$ that violate the defensible privacy of π w.r.t. $\mathbb{A}$ and f_A. Namely, there exists a polynomial p such that for infinitely many n's $\mathbb{D}$ distinguishes with advantage at least $\frac{1}{p(n)}$ between $\Gamma(\mathrm{View}_{\mathbb{A}^*}^{(\mathbb{A}^*,\mathbb{B})}(U_k), f_\mathsf{A}(i^d_\mathbb{A}, U_k))$ and $\Gamma(\mathrm{View}_{\mathbb{A}^*}^{(\mathbb{A}^*,\mathbb{B})}(U_k), f_\mathsf{A}(i^d_\mathbb{A}, U'_k))$, where $\Gamma(x, y)$ equals (x, y) if $\mathrm{IsGoodDef}^{\pi',\mathbb{A}}(x) = 1$ and equals $\perp$ otherwise, and $i^d_\mathbb{A}$ is the value of $\mathbb{A}$'s input in $\mathrm{Defense}(\mathrm{View}_{\mathbb{A}^*}^{(\mathbb{A}^*,\mathbb{B})}(U_k))$. In the following we use $\mathbb{A}^*$ and $\mathbb{D}$ to present an efficient distinguisher D with oracle access to $\mathbb{A}^*$ and $\mathbb{D}$ that violates the semi-honest privacy of (A, B) w.r.t. A and f_A. Recall that in order to violate the semi-honest privacy of (A, B), algorithm D should first sample an input element i_A for A. Then upon getting A's view from the execution of $(\mathsf{A}(i_\mathsf{A}), \mathsf{B}(U_k))$, algorithm D has to distinguish between $f_\mathsf{A}(i_\mathsf{A}, U_k)$ and $f_\mathsf{A}(i_\mathsf{A}, U'_k)$. In order to make the dependencies between its two stages explicit, D uses the variable z to transfer information from its first stage to its second stage.

Algorithm 1 *[The distinguisher D]*

Sampling stage:
Input: 1^n

1. *Choose uniformly at random $r_{\mathbb{A}^*}$ and $r^1_\mathbb{B}$ and fix $\mathbb{A}^*$'s random-coins to $r_{\mathbb{A}^*}$.*
2. *Simulate the first line of $(\mathbb{A}^*, \mathbb{B})$ (i.e., the execution of Com), where $\mathbb{B}$ uses $r^1_\mathbb{B}$ as its random coins.*
3. *Do the following $np(n)$ times:*
   (a) *Simulate the last two lines of $(\mathbb{A}^*, \mathbb{B})$, choosing $\mathbb{B}$'s input and random-coins (i.e., $i_\mathbb{B}$, $r^2_\mathbb{B}$ and $r^3_\mathbb{B}$) uniformly at random.*
   (b) *If $\mathbb{A}^*$ outputs a valid defense d, set $i_\mathsf{A} = i_\mathbb{A}$ and $z = (r_{\mathbb{A}^*}, r^1_\mathbb{B}, r^2_\mathbb{A})$, where $i_\mathbb{A}$ and $r^2_\mathbb{A}$ are the values of these inputs variables in d, and return.*
4. *Set $z = \perp$ and an arbitrary value for i_A.*

Predicting stage:
Input: z, v^π_A - *randomly chosen from* $\mathrm{View}^\pi_\mathsf{A}(i_\mathsf{A}, U_k)$, *and* $c \in Im(f_\mathsf{A})$

1. *If $z = \perp$, output a random coin and return.*
2. *Fix the random-coins of $\mathbb{A}^*$ to $z[r_{\mathbb{A}^*}]$.*
3. *Simulate the first line of $(\mathbb{A}^*, \mathbb{B})$ (i.e., the execution of Com), where $\mathbb{B}$ uses $z[r^1_\mathbb{B}]$ as its random coins.*
4. *Simulate the second line of $(\mathbb{A}^*, \mathbb{B})$, where $\mathbb{B}$ sends $r^3_\mathbb{B} = v^\pi_\mathsf{A}[r_\mathsf{A}] \oplus z[r^2_\mathbb{A}]$ to $\mathbb{A}^*$.*
5. *Simulate the last line of $(\mathbb{A}^*, \mathbb{B})$, where $\mathbb{B}$ sends the same messages that B sends in v^π_A.*
6. *Let $v_{\mathbb{A}^*}$ be the view of $\mathbb{A}^*$ at the end of above simulation,*
   if $\mathrm{IsGoodDef}^{\pi',\mathbb{A}}(v_{\mathbb{A}^*}) = 1$ *output* $\mathbb{D}(v_{\mathbb{A}^*}, c)$,
   otherwise output a random coin.

It is easy to verify that D is efficient given oracle access to $\mathbb{A}^*$ and $\mathbb{D}$, in the following we prove that D violates the semi-honest privacy of π w.r.t. A and f_A. We consider a random execution of $(\mathsf{A}, \mathsf{B}, \mathsf{D})$ with security parameter 1^n and define the random variable $Sim_n = (i_\mathsf{A}, i_\mathsf{B}, r_{\mathbb{A}^*}, r_\mathbb{B}, \mathsf{trans})$ as A and B's inputs in the real execution of π, concatenated with $\mathbb{A}^*$ and $\mathbb{B}$'s views in the simulation of π' done in D's *predicting stage*. More precisely, $i_\mathsf{A} = v^\pi_\mathsf{A}[i_\mathsf{A}]$, $i_\mathsf{B} = v^\pi_\mathsf{A}[i_\mathsf{B}]$, $r_{\mathbb{A}^*} = z[r_{\mathbb{A}^*}]$, $r_\mathbb{B} = (z[r^1_\mathbb{B}], v^\pi_\mathsf{A}[r_\mathsf{B}], v^\pi_\mathsf{A}[r_\mathsf{A}] \oplus z[r^2_\mathbb{A}])$ (set to $\perp$ if $z =\perp$) and finally trans is the transcript of the simulation of π' done in the D's predicting stage (set to $\perp$ if no such simulation occurs).

Let $\mathrm{Defense}(x)$ and $\mathrm{IsGoodDef}(x)$ be $\mathrm{Defense}(x[r_{\mathbb{A}^*}, \mathsf{trans}])$ and $\mathrm{IsGoodDef}^{\pi',\mathbb{A}}(x[r_{\mathbb{A}^*}, \mathsf{trans}])$ respectively. For $c \in Im(f_\mathsf{A})$ let $\mathrm{Out}_\mathsf{D}(x, c)$ be the output bit of D given x and c, note that $\mathrm{Out}_\mathsf{D}(x, c)$ is a random variable that depends on the random-coins used by D to invoke $\mathbb{D}$. Finally, let $\mathrm{Adv}_\mathsf{D}(x)$ be the advantage of D in predicting f_A given x. That is, $\mathrm{Adv}_\mathsf{D}(x) \stackrel{\text{def}}{=} \Pr[\mathrm{Out}_\mathsf{D}(x, f_\mathsf{A}(x[i_\mathsf{A}], x[i_\mathsf{B}])) = 1] - \Pr[\mathrm{Out}_\mathsf{D}(x, f_\mathsf{A}(x[i_\mathsf{A}], U_k)) = 1]$. It is easy to verify that $|\mathsf{Ex}_{x\leftarrow Sim_n}[\mathrm{Adv}_\mathsf{D}(x)]|$ is exactly the advantage of D in breaking the semi-honest privacy of π w.r.t. A and f_A.

We would like the relate the above success probability to that of $\mathbb{D}$ in predicting f_A after a random execution of π'. We define the distribution $Real_n = \left(i^d_\mathbb{A}, i_\mathbb{B}, r_{\mathbb{A}^*}, r_\mathbb{B}, \mathsf{trans}\right)$ induced by a random execution of $(\mathbb{A}^*, \mathbb{B})$ with security parameter 1^n, where $i^d_\mathbb{A}$ is the value of this variable in the defense of $\mathbb{A}^*$ (set to $\perp$ is no good defense is given). Let $\mathrm{Out}_\mathbb{D}(x, c)$ be the output bit of $\mathbb{D}$ given x and c, and let $\mathrm{Adv}_\mathbb{D}(x)$ be the advantage of $\mathbb{D}$ in predicting f_A given x. It is easy to verify that $|\mathsf{Ex}_{x\leftarrow Real_n}[\mathrm{Adv}_\mathbb{D}(x)]|$ is exactly the advantage of $\mathbb{D}$ in breaking the defensible privacy of π' w.r.t. $\mathbb{A}$ and f_A. The following claim helps up to relate the advantage of D in breaking the semi-honest privacy of π to that of $\mathbb{D}$ in breaking the defensible privacy of π'.

Claim. The following hold:

1. For every $n \in \mathbb{N}$ and $x \in \mathrm{Supp}(Real_n)$, it holds that $Sim_n(x) \leq Real_n(x)$
2. For large enough n there exists a set $L_n \subseteq \{x \in \mathrm{Supp}(Real_n) : \mathrm{IsGoodDef}(x) = 1\}$ for which the following hold:
   (a) $\Pr_{x\leftarrow Real_n}[\mathrm{IsGoodDef}(x) \wedge x \notin L_n] \leq \frac{1}{4p(n)}$
   (b) For every $x \in L_n$ it holds that $Sim_n(x) \geq (1 - \frac{1}{4p(n)}) \cdot Real_n(x)$

Proof. When drawing a random $X_R = (i^d_\mathbb{A}, i_\mathbb{B}, r_{\mathbb{A}^*}, r_\mathbb{B}, \mathsf{trans})$ from $Real_n$, its value is fully determined by the value of $X_R[i_\mathbb{B}, r_{\mathbb{A}^*}, r_\mathbb{B}]$, where the latter value is uniformly distributed over all strings of the right length. On the other hand, when drawing a random X_S from Sim_n, the value of $X_S[i_\mathsf{B}, r_{\mathbb{A}^*}, r_\mathbb{B}]$ is uniformly distributed over all strings, only when conditioning that $\mathrm{IsGoodDef}(X_S) = 1$. Where otherwise, $X_S[i_\mathbb{B}, r_{\mathbb{A}^*}, r_\mathbb{B}] = (*, *, \perp)$, a value that is never obtained by an element in $\mathrm{Supp}(Real_n)$. In particular, for every $x \in \mathrm{Supp}(Real_n)$ it holds that

$$\begin{aligned}
&Sim_n(x) = \Pr\left[X_S[i_\mathsf{A}, i_\mathsf{B}, r_{\mathbb{A}^*}, r_\mathbb{B}, \mathsf{trans}] = x[i^d_\mathbb{A}, i_\mathbb{B}, r_{\mathbb{A}^*}, r_\mathbb{B}, \mathsf{trans}]\right] \\
&\leq \Pr\left[X_S[r_{\mathbb{A}^*}, i_\mathbb{B}, r_\mathbb{B}] = x[r_{\mathbb{A}^*}, i_\mathbb{B}, r_\mathbb{B}]\right] \\
&\leq \Pr\left[X_R[r_{\mathbb{A}^*}, i_\mathbb{B}, r_\mathbb{B}] = x[r_{\mathbb{A}^*}, i_\mathbb{B}, r_\mathbb{B}]\right] = Real_n(x) \ ,
\end{aligned}$$

proving the first part of the claim. For $x \in \mathrm{Supp}(Real_n)$, let $\mathrm{Decom}(x)$ be the decommitment of Com given in $\mathrm{Defense}(x)$ (we set it to $\perp$ if no valid defense is given). For $S \subseteq \{0,1\}^*$, we let $W_x(S)$ be the probability that the commitment embedded in x is decommitted to a value in S, conditioned *only* on the random-coins in x used for the commitment (and not on all x). That is, $W_x(S) = \Pr\big[\mathrm{Decom}(X_R) \in S \mid X_R[r_{\mathbb{B}}^1, r_{\mathbb{A}^*}] = x[r_{\mathbb{B}}^1, r_{\mathbb{A}^*}]\big]$. Finally, let $\mathrm{Heaviest}(x) = \mathrm{argmax}_{\sigma\in\{0,1\}^*} W_x(\alpha)$, breaking ties arbitrarily (say, by choosing the lexicographic smallest α) and let $\mathrm{Others}(x) = \{0,1\}^* \setminus \{\mathrm{Heaviest}(x)\}$. We define $L_n = \{x \in \mathrm{Supp}(Real_n) : \mathrm{IsGoodDef}(x) = 1 \wedge W_x(\mathrm{Others}(x)) < \frac{1}{8np(n)^2} \wedge W_x(\mathrm{Heaviest}(x)) > \frac{1}{8p(n)} \wedge \mathrm{Decom}(x) = \mathrm{Heaviest}(x)\}$. In the following we prove the two properties of L_n.

Proving $2(a)$. We first observe that for every polynomial q, it holds that $\Pr[W_{X_R}(\mathrm{Others}(X_R)) > \frac{1}{q(n)}] < \frac{1}{q(n)}$. Assume otherwise, then we can design an adversary for breaking the binding Com. In the commit stage, the adversary acts as $\mathbb{A}^*$ does in the first line of Protocol 1. Then it simulates the rest of the protocol twice (with the same prefix) and outputs the two decommitments implied by $\mathbb{A}^*$'s defenses. Thus, whenever $\Pr[W_{X_R}(\mathrm{Others}(X_R)) > \frac{1}{q(n)}] > \frac{1}{q(n)}$, our adversary breaks the binding of Com with probability $\Omega(\frac{1}{q(n)^3})$.

Since $\mathrm{Decom}(x) \neq \perp$ only if x yields a good defense, it follows that $\Pr[\mathrm{IsGoodDef}(X_R) \wedge (W_x(\mathrm{Heaviest}(x)) + W_x(\mathrm{Others}(x))) < \frac{1}{q(n)}] < \frac{1}{q(n)}$ for every polynomial q. We conclude that

$$\begin{aligned}
&\Pr\left[\mathrm{IsGoodDef}(X_R) \wedge X_R \notin L_n\right] \\
&\leq \Pr\left[W_{X_R}(\mathrm{Others}(X_R)) > \frac{1}{8np(n)^2}\right] + \Pr\Bigg[\mathrm{IsGoodDef}(X_R) \\
&\wedge\ (W_x(\mathrm{Heaviest}(x)) + W_x(\mathrm{Others}(x))) < \left(\frac{1}{8p(n)} + \frac{1}{8np(n)^2}\right)\Bigg] \\
&+ \Pr\Bigg[\mathrm{Decom}(x) \neq \mathrm{Heaviest}(x) \mid \mathrm{IsGoodDef}(X_R) \wedge W_x(\mathrm{Heaviest}(x)) > \frac{1}{8p(n)} \\
&\wedge\ W_{X_R}(\mathrm{Others}(X_R)) \leq \frac{1}{8np(n)^2}\Bigg] \\
&< \frac{1}{8np(n)^2} + \frac{1}{7p(n)} + \frac{8p(n)}{8np(n)^2} < \frac{1}{4p(n)}
\end{aligned}$$

Proving $2(b)$. Let $x \in L_n$, and let X be a random variable drawn from Sim_n conditioned that $X[r_{\mathbb{A}^*}, r_{\mathbb{B}}^1] = x[r_{\mathbb{A}^*}, r_{\mathbb{B}}^1]$. Recall that in order to sample X, algorithm D keeps sampling (up to $np(n)$ times) a random element x' in $Real_n$ conditioned that $x'[r_{\mathbb{A}^*}, r_{\mathbb{B}}^1] = x[r_{\mathbb{A}^*}, r_{\mathbb{B}}^1]$, until $\mathrm{Decom}(x') \neq \perp$. It then set $(X[i_{\mathsf{A}}], z[r_{\mathbb{A}}^2])$ to $\mathrm{Decom}(x')$, where z is the "state" that D transfers from its sampling stage to its predicting stage (the stage where the other parts of X are chosen). In order to keep notations simple, we define $X[r_{\mathbb{A}}^2]$ as $z[r_{\mathbb{A}}^2]$. By the above description it follows that

$$\begin{aligned}
&\Pr[X[i_{\mathsf{A}}, r_{\mathbb{A}}^2] \neq \mathrm{Decom}(x)] \\
\leq\ &\Pr[\mathrm{Decom}(X) = \perp] + \Pr[\mathrm{Decom}(X) \notin \{\mathrm{Decom}(x) \cup \perp\}] \\
\leq\ &\mathrm{neg}(n) + \frac{np(n)}{8np(n)^2} < \frac{1}{4p(n)} \quad ,
\end{aligned} \tag{1}$$

where the second inequality holds since $x \in L_n$. Since the value of $X[i_{\mathsf{B}}, r_{\mathbb{B}}^2, r_{\mathbb{B}}^3]$ is induced by a the parties' inputs and random-coins in a random execution of π, it follows that $X[i_{\mathsf{B}}, r_{\mathbb{B}}^2, r_{\mathbb{B}}^3]$ is uniformly distributed conditioned on $X[i_{\mathsf{A}}, r_{\mathbb{A}}^2] \neq \perp$ and every value of $X[i_{\mathsf{A}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}^1, r_{\mathbb{A}}^2]$. Recall that the value of X_R is fully determined by the value of $X_R[i_{\mathbb{B}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}]$ and that the latter is uniformly distributed over all possible strings. Hence,

$$\begin{aligned}
&\Pr[X_S[i_{\mathsf{B}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}] = x[i_{\mathbb{B}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}] \wedge X_S[i_{\mathsf{A}}, r_{\mathbb{A}}^2] = \mathrm{Decom}(x)] \\
\geq\ &(1 - \frac{1}{4p(n)}) \cdot \Pr\left[X_R[i_{\mathbb{B}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}] = x[i_{\mathbb{B}}, r_{\mathbb{A}^*}, r_{\mathbb{B}}]\right] \\
=\ &(1 - \frac{1}{4p(n)}) \cdot \mathit{Real}_n(x)
\end{aligned} \tag{2}$$

Let $X[r_{\mathsf{A}}]$ be the value of r_{A} in v_{A}^{π} as chosen in the sampling process of X and let $x[r_{\mathbb{A}}^2]$ be the value of $r_{\mathbb{A}}^2$ in $\mathrm{Defense}(x)$. Since $\mathrm{IsGoodDef}(x) = 1$, $\mathbb{A}^*$ acts in the embedded execution of π in x, as $\mathsf{A}(x[i_{\mathbb{A}}^d], x[r_{\mathbb{A}}^2] \oplus x[r_{\mathbb{B}}^3])$ would. Thus, $X_S[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}] = x[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}]$ and $X_S[i_{\mathsf{A}}, r_{\mathbb{A}}^2] = \mathrm{Decom}(x)$ implies that $\mathbb{A}^*$ acts in the embedded execution of π as $\mathsf{A}(X_S[i_{\mathsf{A}}], X_S[r_{\mathbb{A}}^2] \oplus X_S[r_{\mathbb{B}}^3])$ would, that is as $\mathsf{A}(X_S[i_{\mathsf{A}}], X_S[r_{\mathsf{A}}])$. Hence, $X_S[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}] = x[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}]$ and $X[i_{\mathsf{A}}, r_{\mathbb{A}}^2] = \mathrm{Decom}(x)$ implies that $X_S[\mathsf{trans}] = x[\mathsf{trans}]$, and we conclude that

$$\begin{aligned}
&\Pr[X_S = x] \\
=\ &\Pr[X_S[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}] = x[r_{\mathbb{A}^*}, i_{\mathbb{B}}, r_{\mathbb{B}}] \wedge X_S[i_{\mathsf{A}}, r_{\mathbb{A}}^2] = \mathrm{Decom}(x)] \\
\geq\ &(1 - \frac{1}{4p(n)}) \cdot \mathit{Real}_n(x)
\end{aligned}$$

□

Back to the proof the lemma. Let n be large enough be large enough so that Claim 3 holds and assume w.l.o.g. that $\mathsf{Ex}[\mathrm{Adv}_{\mathbb{D}}(X_R)] > \frac{1}{p(n)}$. Since $\mathbb{D}$ gains no advantage when $\mathrm{IsGoodDef}(X_R) = 0$, it follows that $\mathsf{Ex}[\mathrm{Adv}_{\mathbb{D}}(X_R) \cdot \mathrm{IsGoodDef}(X_R)] > \frac{1}{p(n)}$ as well. We first observe that

$$\begin{aligned}
&\mathsf{Ex}[\mathrm{Adv}_{\mathsf{D}}(X_S)] = \mathsf{Ex}[\mathrm{Adv}_{\mathsf{D}}(X_S) \cdot \mathrm{IsGoodDef}(X_S)] \\
\geq\ &\mathsf{Ex}[\mathrm{Adv}_{\mathsf{D}}(X_S) \cdot 1_{X_S \in L_n}] - \Pr[\mathrm{IsGoodDef}(X_S) \wedge X_S \notin L_n] \\
=\ &\Pr\big[\mathrm{Out}_{\mathsf{D}}(X_S, f_{\mathsf{A}}(X_S[i_{\mathsf{A}}, i_{\mathsf{B}}])) = 1) \wedge X_S \in L_n\big] \\
-\ &\Pr\big[\mathrm{Out}_{\mathsf{D}}(X_S, f_{\mathsf{A}}(X_S[i_{\mathsf{A}}], U_k)) = 1) \wedge X_S \in L_n\big] \\
-\ &\Pr[\mathrm{IsGoodDef}(X_S) \wedge X_S \notin L_n] \quad ,
\end{aligned}$$

where $1_{x \in L_n}$ is one if $x \in L_n$ and zero otherwise, and the first equality holds since $\mathrm{Out}_{\mathsf{D}}(x, c)$ is a random coin if $\mathrm{IsGoodDef}(x) = 0$. By Claim 3 we have that

$$\begin{aligned} &\Pr[\mathrm{IsGoodDef}(x) \wedge X_S \notin L_n] \\ \leq\ &\Pr[\mathrm{IsGoodDef}(X_R) \wedge X_R \notin L_n] \leq \frac{1}{4p(n)} \end{aligned} \tag{3}$$

Since $\mathrm{Out}_{\mathsf{D}}(x, c) = \mathrm{Out}_{\mathbb{D}}(x, c)$ for every $x \in \mathrm{Supp}(\mathit{Real}_n)$ such that $\mathrm{IsGoodDef}(x) = 1$, Claim 3 also yields that

$$\begin{aligned} &\Pr[\mathrm{Out}_{\mathsf{D}}(X_S, f_{\mathsf{A}}(X_S[i_{\mathsf{A}}], U_k)) = 1 \wedge X_S \in L_n] \\ \leq\ &\Pr[\mathrm{Out}_{\mathbb{D}}(X_R, f_{\mathsf{A}}(X_R[i^d_{\mathbb{A}}], U_k)) = 1 \wedge X_R \in L_n] \end{aligned} \tag{4}$$

and that

$$\begin{aligned} &\Pr[\mathrm{Out}_{\mathsf{D}}(X_S, f_{\mathsf{A}}(X_S[i_{\mathsf{A}}, i_{\mathsf{B}}])) = 1 \wedge X_S \in L_n] \\ \geq\ &(1 - \frac{1}{4p(n)}) \cdot \Pr[\mathrm{Out}_{\mathbb{D}}(X_R, f_{\mathsf{A}}(X_R[i^d_{\mathbb{A}}, i_{\mathbb{B}}])) = 1 \wedge X_R \in L_n] \end{aligned} \tag{5}$$

We conclude that

$$\begin{aligned} &\mathsf{Ex}[\mathrm{Adv}_{\mathsf{D}}(X_S)] \\ \geq\ &(1 - \frac{1}{4p(n)}) \cdot \Pr[\mathrm{Out}_{\mathbb{D}}(X_R, f_{\mathsf{A}}(X_R[i^d_{\mathbb{A}}, i_{\mathbb{B}}])) = 1 \wedge X_R \in L_n] \\ -\ &\Pr[\mathrm{Out}_{\mathbb{D}}(X_R, f_{\mathsf{A}}(X_R[i^d_{\mathbb{A}}], U_k)) = 1 \wedge X_R \in L_n] - \frac{1}{4p(n)} \\ \geq\ &(1 - \frac{1}{4p(n)}) \cdot \mathsf{Ex}[\mathrm{Adv}_{\mathbb{D}}(X_R) \cdot \mathrm{IsGoodDef}(X_R)] - \frac{1}{4p(n)} - \frac{1}{4p(n)} \\ \geq\ &(1 - \frac{1}{4p(n)}) \cdot \frac{1}{p(n)} - \frac{1}{2p(n)} > \frac{1}{4p(n)} \end{aligned}$$

Since the above holds for infinitely many n's, it concludes the proof of Lemma 3 and thus the proof of Theorem 1.

4 Achieving the Main Result

In the following we prove Theorem 2, the main result of this paper. As corollary of Theorem 1, we have that there exists a fully-black-box reduction from defensible oblivious transfer to semi-honest oblivious transfer and one-way functions. This corollary together with Theorem 3, yields the existence of a fully-black-box reduction from malicious oblivious transfer to semi-honest oblivious transfer and one-way functions. Thus, the proof of the Theorem 2 is concluded by the following folklore theorem (proof given in the full version).

Theorem 4. *There exists a fully-black-box reduction from one-way functions to semi-honest oblivious transfer.*

Acknowledgment

I am very grateful to Yuval Ishai and Omer Reingold for very helpful conversations and for reading early versions of this paper. I am also grateful to Oded Goldreich, Gil Segev and to the anonymous TCC 2008 referees for their useful comments regarding the write-up of the paper.

References

1. B. Barak. How to go beyond the black-box simulation barrier. In *42nd FOCS*, pages 106–115, 2001.
2. M. Blum. How to exchange (secret) keys. *ACM Transactions on Computer Systems*, 1983.
3. G. Brassard, C. Crépeau, and J.-M. Robert. Information theoretic reductions among disclosure problems. In *27th FOCS*, 1986.
4. C. Crépeau. Equivalence between two flavours of oblivious transfers. In *Advances in Cryptology – CRYPTO '87*, 1987.
5. C. Crépeau and J. Kilian. Weakening security assumptions and oblivious transfer. In *CRYPTO '88*, 1988.
6. C. Crépeau and M. Sántha. On the reversibility of oblivious transfer. In *EUROCRYPT '91*, 1991.
7. G. Di Crescenzo, T. Malkin, and R. Ostrovsky. Single database private information retrieval implies oblivious transfer. In *EUROCRYPT '00*, 2000.
8. D. Dolev, C. Dwork, and M. Naor. Nonmalleable cryptography. *JACM*, 30(2):391–437, 2000.
9. S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *Communications of the ACM*, 28(6):637–647, 1985.
10. O. Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, 2001.
11. O. Goldreich. *Foundations of Cryptography – Volume 2: Basic Applications*. Cambridge University Press, 2004.
12. O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority. In *19th STOC*, pages 218–229, 1987.
13. I. Haitner. Implementing oblivious transfer using collection of dense trapdoor permutations. In *1st TCC*, pages 394–409, 2004.
14. I. Haitner and O. Reingold. Statistically-hiding commitment from any one-way function. In *39th STOC*, 2007.
15. J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SICOMP*, 28(4):1364–1396, 1999.
16. R. Impagliazzo and S. Rudich. Limits on the provable consequences of one-way permutations. In *21st STOC*, pages 44–61. ACM Press, 1989.
17. Y. Ishai, E. Kushilevitz, Y. Lindell, and E. Petrank. Black-box constructions for secure computation. In *38th STOC*, 2006.
18. J. Kilian. Founding cryptography on oblivious transfer. pages 20–31, 1988.
19. E. Kushilevitz and R. Ostrovsky. Replication is NOT needed: SINGLE database, computationally-private information retrieval. In *38th FOCS*, pages 364–373, 1997.
20. Y. Lindell. A simpler construction of CCA2-secure public-key encryption under general assumptions. *J. Cryptology*, 19(3):359–377, 2006.

21. M. Naor. Bit commitment using pseudorandomness. *J. of Crypto.*, 4(2):151–158, 1991.
22. M. Naor, R. Ostrovsky, R. Venkatesan, and M. Yung. Perfect zero-knowledge arguments for NP using any one-way permutation. *J. of Crypto.*, 11(2):87–108, 1998.
23. M. Nguyen, S. J. Ong, and S. Vadhan. Statistical zero-knowledge arguments for NP from any one-way function. In *47th FOCS*, pages 3–14, 2006.
24. M. O. Rabin. How to exchange secrets by oblivious transfer. TR-81, Harvard, 1981.
25. O. Reingold, L. Trevisan, and S. P. Vadhan. Notions of reducibility between cryptographic primitives. In *1st TCC*, pages 1–20, 2004.
26. A. Sahai. Non-malleable non-interactive zero knowledge and adaptive chosen-ciphertext security. In *40th FOCS*, pages 543–553, 1999.
27. J. P. Stern. A new and efficient all-or-nothing disclosure of secrets protocol. In *ASIACRYPT '98*, 1998.
28. A. C. Yao. How to generate and exchange secrets. In *27th FOCS*, pages 162–167, 1986.