Electronic Colloquium on Computational Complexity, Report No. 146 (2010)

Homomorphic Encryption: from Private-Key to Public-Key

Ron Rothblum *

September 21, 2010

Abstract

We show that any private-key encryption scheme that is weakly homomorphic with respect to addition modulo 2, can be transformed into a public-key encryption scheme. The homomorphic feature referred to is a minimalistic one; that is, the length of a homomorphically generated encryption should be independent of the number of ciphertexts from which it was created. We do not require anything else on the distribution of homomorphically generated encryptions (in particular, we do not require them to be distributed like real ciphertexts). Our resulting public-key scheme is homomorphic in the following sense. If $i+1$ repeated applications of homomorphic operations can be applied to the private-key scheme, then i repeated applications can be applied to the public-key scheme.

*Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. E-mail: ron.rothblum@weizmann.ac.il. This research was partially supported by the Israel Science Foundation (grant No. 1041/08).

ISSN 1433-8092

1 Introduction

Homomorphic encryption is a paradigm that refers to the ability, given encryptions of some messages, to generate an encryption of a value that is related to the original messages. Specifically, this ability means that from encryptions of k messages $m_1, \ldots, m_k$ it is possible to generate an encryption of $m^* = f(m_1, \ldots, m_k)$ for some (efficiently computable) function f. Ideally, one may want the homomorphically generated encryption of m^* to be distributed identically (or statistically close) to a standard encryption of m^*. We call schemes that have this property strongly homomorphic. Indeed, some proposed encryption schemes are strongly homomorphic w.r.t some algebraic operations such as addition or multiplication (e.g. Goldwasser-Micali [GM84], El-Gamal [Gam84]).

For some applications, it seems as though strongly homomorphic encryption is an overkill. There are weaker notions of homomorphic encryption that might be easier to construct and still suffice for these applications. The very minimal requirement is that a homomorphically generated encryption decrypts correctly to the corresponding message. Alas, this weak requirement does not seem to be useful as is, because it captures schemes that we do not really consider to be homomorphic: Actually, *any* encryption scheme can be slightly modified to satisfy this weak requirement w.r.t *any* efficient operation[1]. A more meaningful notion is obtained by restricting the length of the homomorphically generated encryption. Specifically, we call an encryption scheme weakly homomorphic if homomorphically generated encryptions properly decrypt to the correct message *and* their lengths depend *only* on the security parameter and the message length (and not on the number of input ciphertexts).

1.1 Private-Key vs. Public-Key

When presenting homomorphic encryption, we did not specify whether we consider private-key or public-key encryption schemes. Indeed, one can define strong/weak homomorphic encryption in both settings (with only minor differences). The focus of this paper is showing the connection between public-key and private-key homomorphic encryption.

The easy direction is showing that a public-key homomorphic encryption scheme can be transformed into a private-key homomorphic scheme. This transformation is quite simple and involves only a minor issue. Intuitively, it seems as though any public-key homomorphic scheme *is* a private-key homomorphic scheme. The only problem is that in the public-key setting (in contrast to the private-key one), the homomorphic evaluation algorithm is also given the encryption-key. A simple transformation that addresses this issue is to append the encryption-key to each ciphertext. The resulting private-key scheme clearly retains the homomorphic properties of the public-key scheme (this holds for both strongly and weakly homomorphic schemes).

The harder direction is showing that a private-key homomorphic encryption scheme implies a public-key one. This direction will be addressed by our main result, Theorem 3.1, which shows how to construct a public-key encryption scheme from any private-key scheme that is weakly homomorphic w.r.t addition modulo 2. The resulting public-key scheme partially retains the homomorphic properties of the private-key scheme (see Section 1.2).

We note that it is quite easy to transform a *strongly homomorphic* private-key scheme into a strongly homomorphic public-key one. In fact, this transformation was used by Barak [Bar10] in his exposition of the work of van Dijk et al. [vDGHV10]. For further discussion, see Section 1.3.

[1] Consider implementing the homomorphic evaluation algorithm as the identity function. That is, given ciphertexts and a description of an operation, just output both. Then, modify the decryption algorithm to first decrypt all the ciphertexts and then apply the operation to the decrypted messages. Thus, homomorphic evaluation is delegated to the decryption algorithm that, using the decryption key, can trivially evaluate the required operation.

1.2 Homomorphic Properties of the Public-Key Scheme

So far we have described homomorphic evaluation as a one-shot process, however one can consider repeated application of the homomorphic evaluation algorithm. For *strongly* homomorphic encryption it is possible to do this because homomorphically generated values are identical (or statistically close) to real ciphertexts. For *weakly* homomorphic encryption, the homomorphically generated values can completely differ from real ciphertexts, hence it is unclear that it is possible to keep computing on such homomorphically generated data. Gentry et al. [GHV10] called a scheme that supports i such repeated applications an i-hop homomorphic encryption scheme.

The public-key scheme that we construct is homomorphic in the following sense. If the original private-key scheme is $(i+1)$-hop homomorphic w.r.t some set of operations (which must include addition modulo 2), then the public-key scheme is i-hop homomorphic w.r.t the same set of operations. That is, we lose one application of the homomorphic operation in the construction.

1.3 Technique

The intuition for how to move from private to public key can be seen in a more straightforward manner in the case of *strongly* homomorphic schemes. The following construction was suggested implicitly in [Bar10].

Let E and D be the respective encryption and decryption algorithm of a private-key encryption scheme. Suppose that this encryption scheme is *strongly* homomorphic w.r.t the identity function. That is, it is possible to "re-randomize"[2] ciphertexts. Such a scheme can be used to construct a public-key bit-encryption scheme[3] as follows. The (private) decryption-key is a key k of the private-key scheme and the (public) encryption-key consists of an encryption of 0 and an encryption of 1 (i.e. $E_k(0)$ and $E_k(1)$). To encrypt a bit σ just re-randomize the ciphertext corresponding to σ. To decrypt, apply the private-key decryption algorithm using k (i.e. D_k).

The security of this construction follows from the fact that after re-randomization, all information on the original ciphertext, which was re-randomized, is completely lost. Since *weakly* homomorphic encryption does not guarantee this property, this transformation does not work and we use a more complicated construction, outlined next.

We construct a public-key bit-encryption scheme based on any private-key scheme that is weakly homomorphic w.r.t addition modulo 2. Our decryption key is also a key k of the private-key scheme but the public-key is no longer a single encryption of 0 and 1, but rather a sequence of *many* encryptions of each. Specifically, the public-key consists of two lists of ciphertexts; the first is a list of ℓ encryptions of 0 and the second is a list of ℓ encryptions of 1. To encrypt a bit σ we choose a *random subset* $S \subseteq [\ell]$ that has parity σ (i.e. $|S| \equiv \sigma \bmod 2$). We use S to select ℓ ciphertexts from the public key by selecting the i-th ciphertext from the first list if $i \notin S$ (and from the second if $i \in S$). By *homomorphically adding* the selected ciphertexts modulo 2, we obtain a ciphertext that correctly decrypts to σ.

Most of this work deals with showing that the construction is indeed semantically-secure. To prove security we consider, as a mental experiment, setting both lists in the public-key to be encryptions of 0. Because the mental experiment is computationally indistinguishable from the actual scheme, proving that the original scheme is secure reduces to showing that when *both* lists consist of encryptions of 0, it is essentially impossible to find the parity of the random subset used in the homomorphic encryption process.

[2] This means that there exists an algorithm RR such that for *any* encryption c of a bit b, the output of $RR(c)$ is distributed identically to $E_e(b)$.

[3] A bit-encryption scheme is a public-key encryption scheme that only handles single-bit messages. Such schemes suffice to construct full-fledged public-key encryption schemes (see [Gol04]).

We prove the latter via an information-theoretic theorem that may be of independent interest: Let $X_1, \ldots, X_\ell$ and $Y_1, \ldots, Y_\ell$ be independent and identically distributed over a finite set Ω and let S be a random subset of $[\ell]$. We consider the list Z, defined as $Z_i = X_i$ for $i \notin S$ and $Z_i = Y_i$ for $i \in S$. The theorem states that it is essentially impossible to guess the parity of S based on X, Y and m bits of information on Z. That is, any such guess will be correct with probability that is bounded by (roughly) $\frac{1}{2} + 2^{\ell - m}$. The proof of the information-theoretic theorem makes use of the Efron-Stein decomposition [ES81], an extension of Fourier analysis for product distributions.

We mention that our construction is secure even if we use a slightly weaker definition of homomorphic encryption. Specifically, the length of homomorphically generated encryptions can be a mildly increasing function of the number of input ciphertexts.

1.4 Application of our Construction to Fully-Homomorphic Encryption

Our generic transformation from private-key to public-key encryption can be used as a general methodology for constructing (weakly) homomorphic public-key encryption. One application of this methodology, which actually motivated this work, is to simplify the presentation of the [vDGHV10] fully-homomorphic encryption scheme.

A fully-homomorphic encryption scheme is an encryption scheme that is homomorphic w.r.t any (efficiently computable) function. The concept of fully-homomorphic encryption was first proposed by Rivest et al. [RAD78] in the 70's, but the first concrete proposal was only made recently in the breakthrough work of Gentry [Gen09].

Building on the work of Gentry [Gen09], van Dijk et al. [vDGHV10], proposed a simpler fully-homomorphic public-key scheme. Actually, they propose several variants of the same scheme. Barak [Bar10] noted that one of these variants is in fact fully-homomorphic in the *strong* sense, that is, homomorphically evaluated encryptions are distributed statistically close to actual encryptions. However, this variant requires a stronger assumption than the other variants that are only weakly homomorphic.

From a high-level point of view, both the weak and strong variants of the fully homomorphic scheme are constructed by first proposing a simple *private-key* homomorphic scheme that is only "somewhat" homomorphic (that is, homomorphic w.r.t some restricted functions) and then showing how to modify this scheme into a somewhat homomorphic *public-key* one. The last step uses the bootstrapping technique of [Gen09] to transform the somewhat homomorphic scheme into a fully-homomorphic one.

The aforementioned modification, from private-key to public-key, uses specific properties of the [vDGHV10] scheme. We suggest to use our transformation as an alternative, where the advantage is that our transformation is generic and does not use specific properties of their scheme. Our transformation can be applied to both the strong and weak variants of the somewhat homomorphic *private-key* scheme to obtain a correspondingly strong/weak somewhat homomorphic *public-key* scheme. Note that although the somewhat homomorphic public-key scheme produced by our transformation is slightly different from the one of [vDGHV10], the last step of bootstrapping (see [Gen09]) and reducing the (multiplicative) depth of the decryption circuit can still be applied to our construction.

An alternative, and perhaps more intuitive way to present the [vDGHV10] scheme was taken by Barak [Bar10] for the strongly homomorphic variant of [vDGHV10]. Barak focuses only on presenting the simpler fully-homomorphic *private-key* scheme, since the transformation to a public-key one is easy (as described in Section 1.3). Using our result, it is possible to extend Barak's approach to the weakly homomorphiv variant of the [vDGHV10] scheme. Thus, we suggest to simplify the presentation of the [vDGHV10] scheme by focusing only on showing a (weakly) fully-

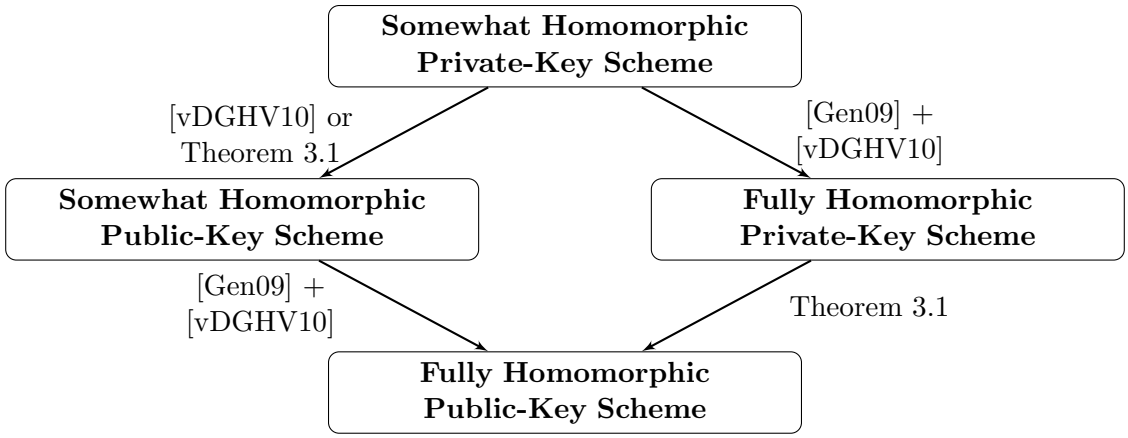

Figure 1: Constructing the weakly homomorphic variant of the [vDGHV10] fully-homomorphic public-key scheme.

homomorphic *private-key* scheme and then, using our generic transformation, to obtain a (weak) fully-homomorphic public-key one. The two approaches to presenting the weakly homomorphic variant of the [vDGHV10] scheme, that were outlined in this section, are depicted in Figure 1.

2 Preliminaries

For a set S, we denote by $x \in_R S$ a uniformly distributed element $x \in S$. Similarly we denote by $X \subseteq_R S$ a uniformly distributed random subset of S.

Non-Standard Notation For every $\ell \in \mathbb{N}$, random variables $X = X_1, \ldots, X_\ell$ and $Y = Y_1, \ldots, Y_\ell$ and set $S \subseteq [\ell]$, we denote by $X_{\overline{S}} Y_S$, the random variable $Z = Z_1, \ldots, Z_\ell$ where $Z_i = X_i$ for $i \notin S$ and $Z_i = Y_i$ for $i \in S$.

2.1 Encryption Schemes

We follow notations and definitions of [Gol01, Gol04]. In particular we use their definition of semantically secure encryption schemes, both in the private-key and public-key settings. Throughout this paper we restrict our attention to bit-encryption schemes, i.e., schemes that encrypt a single bit. For simplicity, we say public-key (resp. private-key) encryption when we actually mean public-key (resp. private-key) bit-encryption.

2.2 Homomorphic Encryption

Since we only consider *weakly* homomorphic encryption, from here on, when we say homomorphic we always mean in the *weak* sense as defined next.

Definition 2.1. *(G, E, D, H) is a* homomorphic private-key encryption scheme *with respect to a set of families of polynomial-sized circuits $\mathcal{C}$ if (G, E, D) are a private-key encryption scheme, H is a probabilistic polynomial-time algorithm and there exists a polynomial $m(\cdot)$ such that for every circuit family $\{C_k\}_{k \in \mathbb{N}} \in \mathcal{C}$, $n \in \mathbb{N}$, polynomial $\ell(\cdot)$, keys $(e, d) \leftarrow G(1^n)$, and $\ell = \ell(n)$ single bit messages $b_1, \ldots, b_\ell \in \{0, 1\}$ the following holds:*

- *Correct decryption of homomorphically generated encryptions:*

$$D_d\left(H\left(C_\ell, E_e(b_1), \ldots, E_e(b_\ell)\right)\right) = C_\ell\left(b_1, \ldots, b_\ell\right).$$

- *The length of homomorphically generated encryptions is independent of ℓ:*

$$\left|H\left(C_\ell, E_e(b_1), \ldots, E_e(b_\ell)\right)\right| \leq m(n).$$

Homomorphic *public-key* encryption is defined analogously (with the modification that H gets the public encryption-key as an additional input).

2.3 i-Hop Homomorphic Encryption

The homomorphic evaluation algorithm in Definition 2.1 is only required to operate on ciphertexts that were output by the encryption algorithm. The definition does not specify what happens if the homomorphic evaluation algorithm is applied to its own output. Gentry et al. [GHV10] defined an i-hop homomorphic encryption scheme as a scheme for which it is possible to apply the homomorphic evaluation algorithm consecutively i times.

Let G, E, D, H be a homomorphic encryption scheme w.r.t to a set of circuit families $\mathcal{C}$. For a given encryption key e, we denote by $W_0(e)$ the set of all valid ciphertexts of the encryption scheme, i.e., all possible outputs of the encryption algorithm E_e applied to a single bit message. For $j \geq 1$, we define $W_j(e)$ to be the set of all possible outputs of the homomorphic evaluation algorithm H when applied to elements in $W_{j-1}(e)$ and a circuit $C \in \mathcal{C}$. We say that elements in $W_j(e)$ are j-th level ciphertexts.

Definition 2.2. *(G, E, D, H) is an i-*hop homomorphic private-key encryption scheme *with respect to a set of families of polynomial-sized circuits $\mathcal{C}$ if (G, E, D) are a private-key encryption scheme, H is a probabilistic polynomial-time algorithm and there exists a polynomial $m(\cdot)$ such that for every circuit family $\{C_k\}_{k\in\mathbb{N}} \in \mathcal{C}$, $n \in \mathbb{N}$, polynomial $\ell(\cdot)$, keys $(e, d) \leftarrow G(1^n)$, $0 \leq j \leq i$, and $\ell = \ell(n)$, ciphertexts $w_1, \ldots, w_\ell \in W_j(e)$ of level j the following holds:*

- *Correct decryption of homomorphically generated encryptions:*

$$D_d\left(H\left(C_\ell, w_1, \ldots, w_\ell\right)\right) = C_\ell\left(D_d(w_1), \ldots, D_d(w_\ell)\right). \tag{2.1}$$

- *The length of homomorphically generated encryptions is independent of ℓ:*

$$\left|H\left(C_\ell, w_1, \ldots, w_\ell\right)\right| \leq m(n). \tag{2.2}$$

Homomorphic *public-key* encryption is defined analogously, with the modification that H receives the encryption-key as an additional input.

3 Constructing a Public-Key Scheme from a Homomorphic Private-Key Scheme

In this section we show how to construct a public-key scheme based on any private-key scheme that is homomorphic w.r.t addition modulo 2.

Theorem 3.1. *Any multiple-message semantically secure private-key encryption scheme that is homomorphic with respect to addition modulo 2 can be transformed into a semantically secure public-key encryption scheme. Furthermore, if the private-key scheme is $(i+1)$-hop homomorphic w.r.t to a set of circuit families, then the constructed public-key scheme is i-hop homomorphic w.r.t to the same set.*

The discussion on the homomorphic properties of the scheme (i.e. the furthermore part) is presented in Section 5. To prove Theorem 3.1, we assume the existence of a homomorphic private-key scheme and use it to construct a public-key scheme (Construction 3.2). The main part of the proof is showing that this public-key scheme is indeed semantically secure.

Construction 3.2. *Let (G, E, D, H) be a homomorphic private-key scheme with respect to addition modulo 2 and let $m(\cdot)$ be the polynomial as in Definition 2.1. We denote by $H_{\oplus}$ the algorithm H when applied to the circuit family that computes addition modulo 2. The encryption scheme (G', E', D', H') is defined as follows:*

Key Generation - $G'(1^n)$: *Set $\ell = 10m(n)$. Select $k \leftarrow G(1^n)$, $X = (X_1, \ldots, X_\ell)$ and $Y = (Y_1, \ldots, Y_\ell)$ such that $X_i \leftarrow E_k(0)$ and $Y_i \leftarrow E_k(1)$ (with fresh random coins for each i). Output X, Y as the public-key and k as the private-key.*

Encryption - $E'_{X,Y}(\sigma)$: *Select a random subset $S \subseteq_R [\ell]$ that has size of parity σ (i.e. $|S| \equiv \sigma \bmod 2$) and output $H_{\oplus}(X_{\overline{S}}Y_S)$ (recall that $X_{\overline{S}}Y_S$ is a list of ℓ ciphertexts that are encryptions of 1 for coordinates in S and encryptions of 0 elsewhere).*

Decryption - $D'_k(c)$: *Output $D_k(c)$.*

Homomorphic Evalutation - $H'(C, (X, Y), c_0, \ldots, c_\ell)$**:** *Output $H(C, c_0, \ldots, c_\ell)$.*

We start by showing that the decryption algorithm correctly decrypts proper ciphertexts. We then proceed to the main part of the proof, showing that Construction 3.2 is indeed semantically secure. In Section 5 we discuss the homomorphic properties of the scheme.

Proposition 3.3. *For every $n \in \mathbb{N}$, $\sigma \in \{0,1\}$ and $((X,Y), k) \leftarrow G'(1^n)$:*

$$D'_k\left(E'_{X,Y}(\sigma)\right) = \sigma.$$

Proof. Based on the first property of homomorphic encryption (Definition 2.1),

$$D'_k\left(E'_{X,Y}(\sigma)\right) = D_k\left(H_{\oplus}\left(X_{\overline{S}}Y_S\right)\right) = \bigoplus_{i=1}^{\ell} D_k(C_i)$$

where $\oplus$ denotes addition modulo 2, $C_i = Y_i$ for $i \in S$ and $C_i = X_i$ otherwise. Since D decrypts correctly, $D_k(X_i) = 0$ and $D_k(Y_i) = 1$. Therefore, $D'_k\left(E'_{X,Y}(\sigma)\right) = \oplus_{i \in S} 1 = |S| \bmod 2 = \sigma$. □

We proceed to the main part of the proof, showing that Construction 3.2 is semantically secure.

Proposition 3.4. *If (G, E, D) is a multiple-message semantically secure private-key scheme then (G', E', D') is a semantically secure public-key scheme.*

Proof. Assume toward a contradiction that (G', E', D') is not semantically secure. This means that there exists a probabilistic polynomial-time adversary A' and a polynomial $p(\cdot)$ such that for infinitely many $n \in \mathbb{N}$:

$$\Pr_{\substack{(X,Y),k\leftarrow G'(1^n)\\ \sigma\in_R\{0,1\}}}\left[A'\left(X,Y,E'_{X,Y}\left(\sigma\right)\right)=\sigma\right] > \frac{1}{2}+\frac{1}{p(n)}\ . \tag{3.1}$$

To derive a contradiction, we consider n from this infinite set and construct a probabilistic polynomial-time adversary A for the underlying private-key scheme. A receives 2ℓ ciphertexts $(\alpha_1,\ldots,\alpha_\ell,\beta_1,\ldots,\beta_\ell)$ and will be shown to distinguish between the following two cases:

- $\alpha_1,\ldots,\alpha_\ell$ are encryptions of 0 and $\beta_1,\ldots,\beta_\ell$ are encryptions of 1.
- $\alpha_1,\ldots,\alpha_\ell,\beta_1,\ldots,\beta_\ell$ are encryptions of 0.

A operates as follows:

1. Set $X=(\alpha_1,\ldots,\alpha_\ell)$ and $Y=(\beta_1,\ldots,\beta_\ell)$.
2. Select $S\subseteq_R[\ell]$.
3. Output 1 if $A'(X,Y,H_\oplus(X_{\overline{S}}Y_S))=|S| \bmod 2$ and 0 otherwise.

Accordingly,

$$\Pr_{\substack{k\leftarrow G(1^n)\\ \alpha_j,\beta_j}}\left[A\left(\alpha_1,\ldots,\alpha_\ell,\beta_1,\ldots,\beta_\ell\right)=1\right]=\Pr_{\substack{k\leftarrow G(1^n)\\ X,Y,S}}\left[A'\left(X,Y,H_\oplus\left(X_{\overline{S}}Y_S\right)\right)=|S| \bmod 2\right].$$

We proceed by analyzing A's behavior in the two different cases. In the first case, $\alpha_i=E_k(0)$ and $\beta_i=E_k(1)$. Consequently, $H_\oplus(X_{\overline{S}}Y_S)$ is distributed identically to an encryption of a random bit under E' and so, by Eq. (3.1), it holds that

$$\Pr_{\substack{k\leftarrow G(1^n)\\ X,Y,S}}\left[A'\left(X,Y,H_\oplus\left(X_{\overline{S}}Y_S\right)\right)=|S| \bmod 2\right]=\Pr_{\substack{(X,Y),k\leftarrow G'(1^n)\\ \sigma\in_R\{0,1\}}}\left[A'\left(X,Y,E'_{X,Y}\left(\sigma\right)\right)=\sigma\right] > \frac{1}{2}+\frac{1}{p(n)}\ .$$

In the second case, $\alpha_i=\beta_i=E_k(0)$. We argue that in this case for every $n\in\mathbb{N}$ and even for an unbounded adversary A',

$$\Pr_{\substack{k\leftarrow G(1^n)\\ X,Y,S}}\left[A'\left(X,Y,H_\oplus\left(X_{\overline{S}},Y_S\right)\right)=|S| \bmod 2\right] < \frac{1}{2}+2^{-0.2\ell+m(n)+1}. \tag{3.2}$$

Equation (3.2) follows from an information-theoretic theorem (Theorem 3.5) that will be stated next and proved in Section 4.

Using Theorem 3.5, we conclude that A distinguishes between the two cases with non-negligible probability, in contradiction to the multiple-message security of (G,E,D), □

Information-Theoretic Theorem. Let Ω be a finite non-empty set and $\ell \in \mathbb{N}$. Let $\mu_1, \ldots, \mu_\ell$ be distributions over Ω and $\mu = \mu_1 \times \cdots \times \mu_\ell$ be a product distribution over Ω^ℓ. Let X and Y be independent random variables identically distributed according to μ over Ω^ℓ.

Theorem 3.5. *For any $\ell, m \in \mathbb{N}$ and any functions $h\colon \Omega^\ell \to \{0,1\}^m$ and $g\colon \Omega^\ell \times \Omega^\ell \times \{0,1\}^m \to \{0,1\}$, it holds that*

$$\Pr_{X,Y,S \subseteq_R [\ell]} \left[g\left(X, Y, h(X_{\overline{S}} Y_S)\right) = |S| \bmod 2 \right] < \frac{1}{2} + 2^{-0.2\ell + m + 1}.$$

Equation (3.2) seems to follow immediately from Theorem 3.5 by setting A' as g, $H_\oplus$ as h and having X and Y distributed as ℓ independent encryptions of 0 each. However, there is a small subtlety - Theorem 3.5 addresses g and h that are *deterministic* functions, in contrast to A' and H that are *probabilistic* algorithms. Additionally, since X and Y are distributed w.r.t to the same randomly chosen key, they are not product distributions as required by Theorem 3.5.

Both issues are resolved by an averaging argument. If Eq. (3.2) does not hold for some $n \in \mathbb{N}$, then there exist random coins for A', H and a fixed private key k for which it does not hold. Once we fix these coins, A' and H become deterministic functions. Additionally, we set X and Y to each be distributed as ℓ encryptions of 0 under the fixed key k, which is in particular a product distribution. Thus, the hypothesis that Eq. (3.2) does not hold contradicts Theorem 3.5.

4 Proof of Theorem 3.5

Theorem 3.5 considers a game in which a computationally unbounded adversary sees X, Y and m bits of information on $X_{\overline{S}} Y_S$ and needs to decide whether S is of even or odd cardinality. In other words, the adversary specifies a function $h : \Omega^\ell \to \{0,1\}^m$ and based on $X, Y, h(X_{\overline{S}} Y_S)$ needs to find $|S| \bmod 2$. Theorem 3.5 states that winning this game with probability noticeably better than $\frac{1}{2}$ is impossible as long as m is sufficiently smaller than ℓ. Note that winning the game becomes easy if m is very large w.r.t ℓ[4] (as long as the probability of a collision in each coordinate, i.e. $\Pr[X_i = Y_i]$, is sufficiently small). Thus, we are interested in the case $m \ll \ell$.

Organization. The proof of Theorem 3.5 uses the Efron-Stein decomposition, an extension of Fourier analysis for general product distributions. We begin by presenting this decomposition, together with the relevant facts. We then turn to the actual proof of Theorem 3.5.

4.1 Efron-Stein Decomposition

Recall that X and Y are independent random variables identically distributed by μ, a product distribution over Ω^ℓ. We consider the inner-product space of functions from Ω^ℓ to $\mathbb{R}$, where the inner product of f and g is $\langle f, g \rangle \stackrel{\text{def}}{=} \mathbf{E}_X[f(X)g(X)]$. We stress that the expectation is over X (which is distributed according to μ). We use the convention that lowercase x and y refer to elements in Ω^ℓ (in contrast to uppercase X and Y which are random variables over Ω^ℓ).

Theorem 4.1 (Efron-Stein Decomposition [ES81])**.** *Any function $f : \Omega^\ell \to \mathbb{R}$ can be decomposed to $f = \sum_{S \subseteq [\ell]} f^S$, where $f^S : \Omega^\ell \to \mathbb{R}$ satisfy:*

1. *f^S only depends on the coordinates of x that reside in S (i.e. x_S).*

[4] If $m \geq \ell \log(|\Omega|)$ just take h to be the identity function.

2. For any $x \in \Omega^\ell$ and $S \not\supseteq U$ it holds that $E_Y[f^U(x_S Y_{\overline{S}})] = 0$.

Note that if $\Omega = \{\pm 1\}$ it is easy to verify that the Fourier representation of the function is also its Efron-Stein decomposition (taking $f^S = \hat{f}(S)\chi_S$ where $\chi_S(x) = \prod_{i \in S} x_i$). In our general setting we denote $\hat{f}(S)^2 \stackrel{\text{def}}{=} \langle f^S, f^S \rangle$ (indeed, when $\Omega = \{\pm 1\}$ this notation agrees with the standard interpretation of $\hat{f}(S)$ in Fourier analysis of Boolean functions).

One of the important properties of this decomposition is that it is orthogonal and therefore Parseval's Equality holds.

Fact 4.2 (Orthogonality)**.** *For any $S \neq U$, f^S and f^U are orthogonal.*

Proof. Assume without loss of generality that $S \not\supseteq U$. Since $X_S Y_{\overline{S}}$ is identically distributed to X,

$$\langle f^S, f^U \rangle = \mathop{\mathbf{E}}_{X}[f^S(X) f^U(X)] = \mathop{\mathbf{E}}_{X,Y}\left[f^S(X_S Y_{\overline{S}}) f^U(X_S Y_{\overline{S}})\right].$$

Based on the fact that f^S only depends on coordinates in S, we can replace $f^S(X_S Y_{\overline{S}})$ with $f^S(X_S X_{\overline{S}}) = f^S(X)$. Thus,

$$\langle f^S, f^U \rangle = \mathop{\mathbf{E}}_{X,Y}\left[f^S(X) f^U(X_S Y_{\overline{S}})\right] = \mathop{\mathbf{E}}_{X}\left[f^S(X) \mathop{\mathbf{E}}_{Y}\left[f^U(X_S Y_{\overline{S}})\right]\right].$$

But by the second property of the decomposition (Theorem 4.1), for every $x \in \Omega^\ell$, $\mathbf{E}_Y[f^U(x_S Y_{\overline{S}})] = 0$ and so we have $\langle f^S, f^U \rangle = 0$. □

Theorem 4.3 (Parseval's Equality)**.**

$$\sum_{S \subseteq [\ell]} \hat{f}(S)^2 = \mathop{\mathbf{E}}_{X}[f(X)^2].$$

Proof.

$$\sum_{S \subseteq [\ell]} \hat{f}(S)^2 = \sum_{S \subseteq [\ell]} \langle f^S, f^S \rangle = \sum_{S,T \subseteq [\ell]} \langle f^S, f^T \rangle = \langle \sum_{S \subseteq [\ell]} f^S, \sum_{T \subseteq [\ell]} f^T \rangle = \langle f, f \rangle,$$

where the second equality follows from orthogonality. □

The Efron-Stein decomposition has proved to be extremely useful in giving explicit expressions for the noise sensitivity of a function or the influence of a subset of its coordinates. We will use it to express the "stability" of a subset of coordinates, which is in a sense the complement of the influence for this set. The fact that we use is summarized in Proposition 4.4 (a similar analysis has been applied previously to give an explicit expression for influence, e.g., in [Bla09]).

Proposition 4.4. *If f is Boolean valued* (i.e. $f \colon \Omega^\ell \to \{0,1\}$)*, then for every $S \subseteq [\ell]$ it holds that:*

$$\Pr_{X,Y}[f(X) = f(X_{\overline{S}} Y_S) = 1] = \sum_{T \subseteq \overline{S}} \hat{f}(T)^2.$$

Proof. Using the fact that f is Boolean, the Efron-Stein decomposition, and linearity of expectation we have:

$$\begin{aligned}\Pr_{X,Y}\left[f(X)=f(X_{\overline{S}}Y_S)=1\right] &= \mathop{\mathbf{E}}_{X,Y}\left[f(X)f(X_{\overline{S}}Y_S)\right] \\ &= \mathop{\mathbf{E}}_{X,Y}\left[\sum_{T\subseteq[\ell]} f^T(X)\sum_{U\subseteq[\ell]} f^U(X_{\overline{S}}Y_S)\right] \\ &= \sum_{U,T\subseteq[\ell]} \mathop{\mathbf{E}}_{X}\left[f^T(X)\mathop{\mathbf{E}}_{Y}\left[f^U(X_{\overline{S}}Y_S)\right]\right]. \end{aligned} \tag{4.1}$$

From the Efron-Stein decomposition we have that if $U \not\subseteq \overline{S}$ then $E_Y[f^U(X_{\overline{S}}Y_S)] = 0$, whereas if $U \subseteq \overline{S}$ then $E_Y[f^U(X_{\overline{S}}Y_S)] = f^U(X)$. Thus, Eq. (4.1) yields that:

$$\Pr_{X,Y}\left[f(X)=f(X_{\overline{S}}Y_S)=1\right] = \sum_{T\subseteq[\ell]}\sum_{U\subseteq\overline{S}} \mathop{\mathbf{E}}_{X}[f^T(X)f^U(X)] = \sum_{T\subseteq[\ell]}\sum_{U\subseteq\overline{S}}\langle f^T, f^U\rangle = \sum_{T\subseteq\overline{S}}\langle f^T, f^T\rangle$$

where the last equality follows from orthogonality. □

4.2 Proof of Theorem 3.5

We would like to show that for a *typical* $\gamma \in \{0,1\}^m$, the number of odd S that map to γ (that is $h(X_{\overline{S}}Y_S) = \gamma$) and the number of even such S are roughly the same. This would imply that any adversary, which sees only X, Y and γ, cannot guess whether γ was produced from an odd or even S, which is exactly what we are looking to prove. To formalize this, we introduce the following notation; for $\gamma \in \{0,1\}^m$, we define:

$$I_{\text{odd}}(X,Y,\gamma) \stackrel{\text{def}}{=} \left|\{\, T\subseteq[\ell] : h(X_{\overline{T}}Y_T)=\gamma \text{ and } |T| \text{ is odd}\,\}\right| \tag{4.2}$$

$$I_{\text{even}}(X,Y,\gamma) \stackrel{\text{def}}{=} \left|\{\, T\subseteq[\ell] : h(X_{\overline{T}}Y_T)=\gamma \text{ and } |T| \text{ is even}\,\}\right| \tag{4.3}$$

Organization. We begin by presenting some basic facts. The proof will be composed of two lemmas, Lemma 4.7 (which is the main lemma) states that for *every* $\gamma \in \{0,1\}^m$, w.h.p, the number of odd T that map to γ is fairly close to the number of even T (in absolute terms). Lemma 4.11 states that for a *typical* γ the total number of T that map to it is very large. Combining these two lemmas we prove Theorem 3.5.

4.2.1 Basic Facts

We first present two basic facts that follow immediately from the structure of $X_{\overline{S}}Y_S$.

Fact 4.5. *For every $\gamma \in \{0,1\}^m$, there exists a constant $\mu_\gamma \in [0,1]$ such that for every $S \subseteq [\ell]$:*

$$\Pr_{X,Y}\left[h(X_{\overline{S}}Y_S)=\gamma\right] = \mu_\gamma.$$

Proof. Define $\mu_\gamma \stackrel{\text{def}}{=} \Pr\left[h(X)=\gamma\right]$ and note that $\Pr\left[h(X_{\overline{S}}Y_S)=\gamma\right] = \Pr\left[h(X)=\gamma\right]$ (because $X_{\overline{S}}Y_S$ and X are identically distributed). □

Fact 4.6. *For every $S, T \subseteq [\ell]$ and $\gamma \in \{0,1\}^m$,*

$$\Pr_{X,Y}\left[h(X_{\overline{S}}Y_S) = h(X_{\overline{T}}Y_T) = \gamma\right] = \Pr_{X,Y}\left[h(X) = h(X_{\overline{S \oplus T}}Y_{S \oplus T}) = \gamma\right]$$

where $S \oplus T$ denotes the symmetric difference of two sets, i.e., $S \oplus T \stackrel{def}{=} (S \backslash T) \cup (T \backslash S)$.

Proof. Using the fact that $X_{\overline{S}}Y_S$ is identically distributed to X, we can swap Y_S and X_S in the expression $\Pr\left[h(X_{\overline{S}}Y_S) = h(X_{\overline{T}}Y_T)\right]$. Hence, $X_{\overline{S}}Y_S$ becomes X. For $X_{\overline{T}}Y_T$ we use X for coordinates that are in $\overline{T} \backslash S$ or in $T \cap S$ and use Y for coordinates that are in $\overline{T} \cap S$ or in $T \backslash S$. Therefore, $X_{\overline{T}}Y_T$ becomes $X_{\overline{S \oplus T}}Y_{S \oplus T}$. □

4.2.2 The Main Lemma

Lemma 4.7. *For every $\gamma \in \{0,1\}^m$, it holds that:*

$$\Pr_{X,Y}\left[|I_{\text{odd}}(X,Y,\gamma) - I_{\text{even}}(X,Y,\gamma)| \geq 2^{0.6\ell}\right] \leq 2^{-0.2\ell}.$$

Throughout the proof of this lemma, in all probabilistic statements, the probability is always over X and Y. Additionally, since X and Y are clear from the context, we use the shorthand $I_{\text{odd}}(\gamma)$ (resp. $I_{\text{even}}(\gamma)$) for $I_{\text{odd}}(X,Y,\gamma)$ (resp. $I_{\text{even}}(X,Y,\gamma)$).

Foreseeing that we will prove Lemma 4.7 by an application of Chebyshev's inequality, we proceed by bounding the expectation and variance of $I_{\text{odd}}(\gamma) - I_{\text{even}}(\gamma)$.

Proposition 4.8. *For every $\gamma \in \{0,1\}^m$, it holds that:*

$$\mathbf{E}[I_{\text{odd}}(\gamma) - I_{\text{even}}(\gamma)] = 0.$$

Proof. $I_{\text{odd}}(\gamma)$ can be expressed as a sum of indicator variables: $I_{\text{odd}}(\gamma) = \sum_{\text{odd } T} I_T(\gamma)$, where $I_T(\gamma)$ is an indicator for the event $h(X_{\overline{T}}Y_T) = \gamma$. Thus,

$$\mathbf{E}\left[I_{\text{odd}}(\gamma)\right] = \mathbf{E}\left[\sum_{\text{odd } T} I_T(\gamma)\right] = \sum_{\text{odd } T} \mathbf{E}\left[I_T(\gamma)\right] = \sum_{\text{odd } T} \Pr\left[h(X_{\overline{T}}Y_T) = \gamma\right] = 2^{\ell-1}\mu_\gamma$$

where the last equality follows from Fact 4.5. Similarly, it is easy to see that $\mathbf{E}\left[I_{\text{even}}(\gamma)\right] = 2^{\ell-1}\mu_\gamma$ and thus $\mathbf{E}\left[I_{\text{odd}}(\gamma) - I_{\text{even}}(\gamma)\right] = 0$. □

Proposition 4.9. *For every $\gamma \in \{0,1\}^m$, it holds that*

$$\mathbf{V}\text{ar}[I_{\text{odd}}(\gamma) - I_{\text{even}}(\gamma)] \leq 2^\ell.$$

Proof. Recall that I_{odd} and I_{even} can be expressed as the sum of the indicator variables I_T (as defined in the proof of Proposition 4.8). Thus, using Proposition 4.8 and some manipulations we

have:

$$\begin{aligned}
\mathbf{V}\mathrm{ar}\left[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)\right] &= \mathbf{E}\left[\left(I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)\right)^2\right] \\
&= \mathbf{E}\left[I_{\mathrm{odd}}(\gamma)^2\right] + \mathbf{E}\left[I_{\mathrm{even}}(\gamma)^2\right] - 2\,\mathbf{E}\left[I_{\mathrm{odd}}(\gamma) I_{\mathrm{even}}(\gamma)\right] \\
&= \mathbf{E}\left[\left(\sum_{\text{odd } T} I_T(\gamma)\right)^2\right] + \mathbf{E}\left[\left(\sum_{\text{even } T} I_T(\gamma)\right)^2\right] \\
&\quad - 2\,\mathbf{E}\left[\left(\sum_{\text{odd } T} I_T(\gamma)\right)\left(\sum_{\text{even } T} I_T(\gamma)\right)\right] \\
&= \sum_{\substack{T,U\subseteq[\ell] \text{ s.t.} \\ |T|=|U| \bmod 2}} \mathbf{E}\left[I_T(\gamma) I_U(\gamma)\right] - \sum_{\substack{T,U\subseteq[\ell] \text{ s.t.} \\ |T|\neq|U| \bmod 2}} \mathbf{E}\left[I_T(\gamma) I_U(\gamma)\right] \\
&= \sum_{T,U} (-1)^{|T \oplus U|}\, \mathbf{E}\left[I_T(\gamma) I_U(\gamma)\right] \\
&= \sum_{T,U} (-1)^{|T \oplus U|} \Pr\left[h(X_{\overline{T}} Y_T) = h(X_{\overline{U}} Y_U) = \gamma\right].
\end{aligned}$$

Now using Fact 4.6 we have:

$$\begin{aligned}
\mathbf{V}\mathrm{ar}[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)] &= \sum_{T,U} (-1)^{|T \oplus U|} \Pr\left[h(X) = h(X_{\overline{T \oplus U}} Y_{T \oplus U}) = \gamma\right] \\
&= \sum_{T,U} (-1)^{|T|} \Pr\left[h(X) = h(X_{\overline{T}} Y_T) = \gamma\right] \\
&= 2^\ell \sum_{i=0}^{\ell} (-1)^i \sum_{T\colon |T|=i} \Pr\left[h(X) = h(X_{\overline{T}} Y_T) = \gamma\right].
\end{aligned}$$

Let $f\colon \Omega^\ell \to \{0,1\}$ be the indicator function for $h(X) = \gamma$. Clearly, for every T, it holds that $\Pr\left[h(X) = h(X_{\overline{T}} Y_T) = \gamma\right] = \Pr\left[f(X) = f(X_{\overline{T}} Y_T) = 1\right]$ and so by using Proposition 4.4 we derive:

$$\begin{aligned}
\mathbf{V}\mathrm{ar}[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)] &= 2^\ell \sum_{i=0}^{\ell} (-1)^i \sum_{T\colon |T|=i} \Pr\left[f(X) = f(X_{\overline{T}} Y_T) = 1\right] \\
&= 2^\ell \sum_{i=0}^{\ell} (-1)^i \sum_{T\colon |T|=i} \left(\sum_{U\subseteq\overline{T}} \hat{f}(U)^2\right) \\
&= 2^\ell \sum_{i=0}^{\ell} (-1)^i \sum_{R\colon |R|=\ell-i} \left(\sum_{U\subseteq R} \hat{f}(U)^2\right).
\end{aligned}$$

Note that each $\hat{f}(U)^2$ in the sum appears $\binom{\ell-|U|}{i}$ times with respect to each i (and this holds even

when $i > \ell - |U|$). Thus:

$$\begin{aligned}\mathbf{V}\mathrm{ar}[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)] &= 2^{\ell}\sum_{i=0}^{\ell}(-1)^i \sum_{U\subseteq[\ell]} \binom{\ell-|U|}{i} \hat{f}(U)^2 \\ &= 2^{\ell}\sum_{U\subseteq[\ell]} \hat{f}(U)^2 \sum_{i=0}^{\ell}(-1)^i \binom{\ell-|U|}{i} \\ &= 2^{\ell}\sum_{U\subseteq[\ell]} \hat{f}(U)^2 (1-1)^{\ell-|U|} \\ &= 2^{\ell}\hat{f}([\ell])^2.\end{aligned}$$

Finally, using Parseval's Equality (Theorem 4.3) and the fact that range of f is $\{0,1\}$:

$$\mathbf{V}\mathrm{ar}[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)] = 2^{\ell}\hat{f}([\ell])^2 \le 2^{\ell}\sum_{S\subseteq[\ell]}\hat{f}(S)^2 = 2^{\ell}\,\mathop{\mathbf{E}}_{X}[f(X)^2] \le 2^{\ell}.$$

□

Deriving Lemma 4.7. Applying Chebyshev's inequality, while using Propositions 4.8 and 4.9, we get that

$$\Pr\left[|I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)| \ge 2^{0.6\ell}\right] \le \frac{\mathbf{V}\mathrm{ar}[I_{\mathrm{odd}}(\gamma) - I_{\mathrm{even}}(\gamma)]}{2^{1.2\ell}} \le \frac{2^{\ell}}{2^{1.2\ell}} = 2^{-0.2\ell}.$$

□

4.2.3 Completing the Proof

Lemma 4.7 addresses the case where γ is fixed. However, we need to handle γ that are chosen according to a specific distribution ($\gamma \sim h(X_{\overline{S}}Y_S)$). Since we consider such γ, it is convenient to define:

$$\tilde{I}_{\mathrm{even}}(X,Y,S) = I_{\mathrm{even}}\left(X,Y,h(X_{\overline{S}}Y_S)\right) \tag{4.4}$$

$$\tilde{I}_{\mathrm{odd}}(X,Y,S) = I_{\mathrm{odd}}\left(X,Y,h(X_{\overline{S}}Y_S)\right) \tag{4.5}$$

$$\Delta_{X,Y}(S) = \left|\tilde{I}_{\mathrm{even}}(X,Y,S) - \tilde{I}_{\mathrm{odd}}(X,Y,S)\right| \tag{4.6}$$

Corollary 4.10.

$$\Pr_{X,Y,S\subseteq_R[\ell]}\left[\Delta_{X,Y}(S) \ge 2^{0.6\ell}\right] \le 2^{-0.2\ell+m}.$$

Proof. If $\Delta_{X,Y}(S) \ge 2^{0.6\ell}$ then for $\gamma = h(X_{\overline{S}}Y_S)$ it holds that $|I_{\mathrm{odd}}(X,Y,\gamma) - I_{\mathrm{even}}(X,Y,\gamma)| \ge 2^{0.6\ell}$. Thus:

$$\Pr_{X,Y,S}\left[\Delta_{X,Y}(S) \ge 2^{0.6\ell}\right] \le \Pr_{X,Y}\left[\exists\,\gamma\in\{0,1\}^m \text{ s.t. } |I_{\mathrm{odd}}(X,Y,\gamma) - I_{\mathrm{even}}(X,Y,\gamma)| \ge 2^{0.6\ell}\right].$$

The corollary follows by applying a union bound and Lemma 4.7. □

Consider all $T \subseteq [\ell]$ that map (via h) to the same value as S. Corollary 4.10 bounds the difference between the number of even and odd such T. However, since it does so only in absolute terms, it is meaningless if the number of such T is small. Lemma 4.11 shows that for a typical γ, w.h.p, this is not the case.

Notation. Recall our convention that lowercase x and y refer to elements in Ω^ℓ. For fixed x and y, we define $I_{x,y}(\gamma)$ to be the total number of $T \subseteq [\ell]$ that h maps to γ, i.e.,

$$I_{x,y}(\gamma) \stackrel{\text{def}}{=} I_{\text{odd}}(x,y,\gamma) + I_{\text{even}}(x,y,\gamma) = \left|\{\, T \subseteq [\ell] :\ h(x_{\overline{T}} y_T) = \gamma \,\}\right|. \tag{4.7}$$

Since we are sometimes interested in typical γ's, we also define

$$\tilde{I}_{x,y}(S) \stackrel{\text{def}}{=} I_{x,y}\left(h(x_{\overline{S}}, y_S)\right). \tag{4.8}$$

Lemma 4.11. *For every $x, y \in \Omega^\ell$,*

$$\Pr_S\left[\tilde{I}_{x,y}(S) \leq 2^{0.8\ell}\right] \leq 2^{-0.2\ell+m}.$$

Proof.

$$\begin{aligned}
\Pr_S\left[\tilde{I}_{x,y}(S) \leq 2^{0.8\ell}\right] &= \sum_{\gamma\in\{0,1\}^m} \Pr_S\left[\tilde{I}_{x,y}(S) \leq 2^{0.8\ell} \bigwedge h(x_{\overline{S}} y_S) = \gamma\right] \\
&= \sum_{\gamma\in\{0,1\}^m} \Pr_S\left[I_{x,y}(\gamma) \leq 2^{0.8\ell} \bigwedge h(x_{\overline{S}} y_S) = \gamma\right] \\
&= \sum_{\gamma:\ I_{x,y}(\gamma) \leq 2^{0.8\ell}} \Pr_S\left[h(x_{\overline{S}} y_S) = \gamma\right] \\
&\leq 2^m \cdot \frac{2^{0.8\ell}}{2^\ell}\ .
\end{aligned}$$

□

Lemma 4.11 together with Corollary 4.10 imply, that w.h.p, $\tilde{I}_{\text{odd}}(X,Y,S)$ and $\tilde{I}_{\text{even}}(X,Y,S)$ are very close (since their sum is big and their difference is small). Intuitively, this implies that an adversary that tries to find $|S| \bmod 2$ from X, Y and $h(X_{\overline{S}} Y_S)$ can not do much better than a fair coin toss. Proposition 4.12 formalizes this intuitive connection.

Proposition 4.12. *For every $x, y \in \Omega^\ell$:*

$$\Pr_S\left[g\left(x, y, h(x_{\overline{S}} y_S)\right) = |S| \bmod 2\right] \leq \frac{1}{2} + \frac{1}{2} \cdot \mathbf{E}\left[\frac{\Delta_{x,y}(S)}{\tilde{I}_{x,y}(S)}\right]$$

where $\Delta_{x,y}(S)$ and $\tilde{I}_{x,y}(S)$ are as defined in Eq. (4.6) *and Eq.* (4.8) *respectively.*

Proof. Since x and y are fixed, and we quantify over *all* g and h, we can just consider functions that depend on x and y. Thus, we denote $g_{x,y}(\gamma) \stackrel{\text{def}}{=} g(x,y,\gamma)$ and $h_{x,y}(S) \stackrel{\text{def}}{=} h(x_{\overline{S}} y_S)$.

Choosing a random subset $S \subseteq [\ell]$ is equivalent to first choosing $\gamma = h_{x,y}(S)$ and then choosing uniformly over all $T \subseteq [\ell]$ that h maps to γ. Formally, let S be a uniformly distributed subset of $[\ell]$ and let T_S be distributed uniformly over $\{\, T \subseteq [\ell] :\ h_{x,y}(T) = h_{x,y}(S) \,\}$. Since S and T_S are identically distributed (by the uniform distribution) it holds that

$$\begin{aligned}
\Pr_S\left[g_{x,y}\left(h_{x,y}(S)\right) = |S| \bmod 2\right] &= \Pr_{S,T_S}\left[g_{x,y}\left(h_{x,y}(S)\right) = |T_S| \bmod 2\right] \\
&= \mathbf{E}_S\left[\Pr_{T_S}\left[g_{x,y}\left(h_{x,y}(S)\right) = |T_S| \bmod 2\right]\right].
\end{aligned}$$

For fixed S, by definition, $\Pr_{T_S}\left[g_{x,y}(h_{x,y}(S)) = |T_S| \bmod 2\right]$ is just

$$\frac{|\{\, T :\ (|T| \bmod 2) = g_{x,y}(h_{x,y}(S)) \text{ and } h_{x,y}(T) = h_{x,y}(S)\,\}|}{|\{\, T :\ h_{x,y}(T) = h_{x,y}(S)\,\}|}.$$

The numerator of this expression equals the number of T's that map to the same value as S whose size is of some fixed parity (note that $g_{x,y}(h_{x,y}(S))$ is fixed) and thus is at most $\max\left(\tilde{I}_{\text{odd}}(x,y,S), \tilde{I}_{\text{even}}(x,y,S)\right)$. Likewise, the denominator is exactly $\tilde{I}_{x,y}(S)$ and so we have:

$$\begin{aligned}\Pr_S\left[g_{x,y}\left(h_{x,y}(S)\right) = |S| \bmod 2\right] &\leq \mathop{\mathbf{E}}_S\left[\frac{\max\left(\tilde{I}_{\text{odd}}(x,y,S), \tilde{I}_{\text{even}}(x,y,S)\right)}{\tilde{I}_{x,y}(S)}\right] \\ &= \frac{1}{2} + \frac{1}{2}\cdot\mathop{\mathbf{E}}_S\left[\frac{\Delta_{x,y}(S)}{\tilde{I}_{x,y}(S)}\right].\end{aligned}$$

□

Deriving Theorem 3.5. Corollary 4.10 and Lemma 4.11 imply that:

$$\Pr_{X,Y,S\subseteq_R[\ell]}\left[\frac{\Delta_{X,Y}(S)}{\tilde{I}_{X,Y}(S)} < 2^{-0.2\ell}\right] > 1 - 2\cdot 2^{-0.2\ell+m}.$$

Therefore,

$$\mathop{\mathbf{E}}_{X,Y,S\subseteq_R[\ell]}\left[\frac{\Delta_{X,Y}(S)}{\tilde{I}_{X,Y}(S)}\right] < \left(1 - 2^{-0.2\ell+m+1}\right)\cdot 2^{-0.2\ell} + 2^{0.2\ell+m+1}\cdot 1 < 2^{-0.2\ell+m+2}.$$

And so, by Proposition 4.12,

$$\Pr_{X,Y,S\subseteq_R[\ell]}\left[g\left(X,Y,h(X_{\overline{S}}Y_S)\right) = |S| \bmod 2\right] < \frac{1}{2} + 2^{-0.2\ell+m+1}.$$

□

5 Homomorphic Properties of the Public-Key Scheme

In this section, we discuss the homomorphic properties of the public-key scheme presented in Construction 3.2. Specifically, we shall show that if the private-key scheme supports $i+1$ repeated homomorphic operations then the public-key scheme supports i such operations. Intuitively, this follows by the fact that the public-key encryption algorithm applies a single homomorphic operation (see Fact 5.2).

Proposition 5.1. *Suppose G, E, D, H are an $(i+1)$-hop homomorphic private-key scheme w.r.t to a set of circuit families $\mathcal{C}$ that includes addition modulo 2. Then G', E', D', H' as defined in Construction 3.2 are an i-hop homomorphic public-key scheme w.r.t the set $\mathcal{C}$.*

Theorem 3.1 shows that (G', E', D', H') is indeed a public-key encryption scheme and so, we only need to show that the scheme supports i repeated evaluations of circuits from $\mathcal{C}$.

Let $(X, Y), k$ be a pair of encryption/decryption keys of the public scheme (w.r.t to the security parameter n). We denote the j-th level ciphertexts of the private-key scheme by $W_j(k)$ and the j-th level ciphertexts of the public-key scheme by $W'_j(X, Y)$.

Fact 5.2. *For every $j \in \mathbb{N}$, $W'_j(X,Y) \subseteq W_{j+1}(k)$.*

Proof. By induction on j. □

Let $\{C_k\}_k \in \mathcal{C}$, $0 \leq j \leq i$, $\ell = \ell(n)$ and $w_1, \ldots, w_\ell$ be j-th level ciphertexts of the public-key scheme (i.e., in $W'_j(X,Y)$). We proceed by showing that the first property of Definition 2.2 (Eq. 2.1) holds. By Fact 5.2, it holds that $w_1, \ldots, w_\ell \in W_{j+1}(k)$ and thus,

$$\begin{aligned} H'(C_\ell, (X,Y), w_1, \ldots, w_\ell) &= H(C_\ell, w_1, \ldots, w_\ell) \\ &= C_\ell(D_d(w_1), \ldots, D_d(w_\ell)) \qquad = C_\ell(D'_d(w_1), \ldots, D'_d(w_\ell)). \end{aligned}$$

where the first and third equalities follow from the definition of H' and D' respectively and the second equality follows from the first requirement of Definition 2.2, noting that $w_1, \ldots, w_\ell$ are ciphertexts of level $j+1 \leq i+1$ of the private-key scheme.

A similar argument shows that the second property of Definition 2.2 (Eq. 2.2) holds. Indeed, since $w_1, \ldots, w_\ell \in W'_j(X,Y) \subseteq W_{j+1}(k)$ it holds that,

$$|H'(C_\ell, (X,Y), w_1, \ldots, w_\ell)| = |H(C_\ell, w_1, \ldots, w_\ell)| \leq m(n)$$

for every $0 \leq j \leq i$.

Acknowledgments

I would like to express my thanks and appreciation to my M.Sc. advisor, Oded Goldreich, for his encouragement and guidance in completing this work. In particular, I would like to thank him for many helpful discussions and constructive comments that helped present this work in a more coherent way.

References

[Bar10] Boaz Barak. Cryptography course - Lecture notes , COS 433. Princeton University, Computer Science Department. Available at `http://www.cs.princeton.edu/courses/archive/spring10/cos433`, Spring 2010.

[Bla09] Eric Blais. Testing juntas nearly optimally. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009*, pages 151–158. ACM, 2009.

[ES81] Brad Efron and Charles Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9(3):586–596, 1981.

[Gam84] Taher El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *CRYPTO*, pages 10–18, 1984.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009*, pages 169–178. ACM, 2009.

[GHV10] Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. i-hop homomorphic encryption and rerandomizable yao circuits. In *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference*, pages 155–172. Springer, 2010.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.

[Gol01] Oded Goldreich. *Foundations of Cryptography. Volume I: Basic Tools.* Cambridge University Press, 2001.

[Gol04] Oded Goldreich. *Foundations of Cryptography: Volume 2: Basic Applications.* Cambridge University Press, 2004.

[RAD78] Ronald L. Rivest, Leonard Adleman, and Michael L. Dertouzos. On data banks and privacy homomorphisms. In *Foundations of Secure Computation*, pages 169–180. Academic Press, 1978.

[vDGHV10] Marten van Dijk, Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. Fully homomorphic encryption over the integers. In *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 24–43. Springer, 2010.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de