A New Aspect for Security Notions: Secure Randomness in Public-Key Encryption Schemes

Takeshi Koshiba

Secure Computing Lab., Fujitsu Laboratories Ltd.,
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki 211-8588, Japan
koshiba@acm.org

Abstract. In this paper, we introduce a framework in which we can uniformly and comprehensively discuss security notions of public-key encryption schemes even for the case where some weak generator producing seemingly random sequences is used to encrypt plaintext messages. First, we prove that indistinguishability and semantic security are not equivalent in general. On the other hand, we derive some sufficient condition for the equivalence and show that polynomial-time pseudo-randomness is not always necessary for the equivalence.

1 Introduction

One of the important goals in computational cryptography is to provide a public-key encryption scheme that achieves a security level as strong as possible under various circumstances. For this purpose, several security notions have been introduced. In particular, we will discuss in this paper the notions of "semantic security" and "indistinguishability of encryptions" introduced by Goldwasser and Micali [12], which have shown to be equivalent [12,18]. For another major security notion, we have "non-malleability" introduced in [8]. These notions are defined in terms of an adversary who is given only a challenge ciphertext. This attack model is called *ciphertext only attack* (abbreviated COA). Besides COA, three major attack models have been studied in the literature. One is called *chosen plaintext attack* (abbreviated CPA) model, in which the adversary can encrypt any plaintext messages of his choice. For more stronger attack models, *chosen ciphertext attack* [19] and *adaptive chosen ciphertext attack* [20] have been also considered.

Although these security notions have been studied quite well (see, e.g., [1,2,4]), we think that there are still some important issues that have not been addressed in the previous research. Security when used with a "pseudo-random" resource is one of such issues. Usually, security notions are defined assuming that ideal (i.e., true) random resource is available. Furthermore, it has been shown that one can safely use any "polynomial-time pseudo-random" generator for the substitute of the true random resource; that is, most security notions do not change by using the polynomial-time pseudo-randomness for the true randomness. Although we have several "theoretically guaranteed" polynomial-time pseudo-random generators, they are unfortunately not fast enough for practical use, and much faster

K. Kim (Ed.): PKC 2001, LNCS 1992, pp. 87–103, 2001.

but less reliable pseudo-random generators have been used in many practical situations. Then the above security notions (and their relations) may be no longer valid with such weak pseudo-randomness. In fact, it has been shown [3] that if DSS is used with a linear congruential generator, then its secret key can be easily detected after seeing a few signatures. (See also [7,9,15,22].) Though this result indicates that the linear congruential generator is unsuitable for cryptographic purposes, it does not mean that the linear congruential generator is useless at all for *all* cryptographic systems. It is certainly important to study more carefully which aspect of the randomness is indeed important for discussing several security levels.

In this paper, we will introduce a framework in which we can uniformly and comprehensively discuss "semantic security" and "indistinguishability" notions even for the case where some weak generator producing seemingly random sequences is used. (In order to avoid confusion, we will use throughout the paper the term "quasi-random" for referring pseudo-randomness including one without any guarantee.) While most of corresponding notions are easily restated from the original definitions, we have to be a bit careful for choosing right definitions in relation to attack models. In the context of public-key encryption scheme, CPA is equivalent to COA because once the adversary obtains the public-key, he can compute ciphertext messages easily. This may not be true any more when quasi-random generators are involved. In this paper, for the model corresponding to COA, we consider the situation in which an adversary cannot access to the quasi-random generator, it is still possible for the adversary to make use of the encryption algorithm. On the other hand, in the model corresponding to CPA, it may be more natural that adversary can invoke the encryption oracle which, given a plaintext message, uses the quasi-random generator in encrypting the message and replies with its ciphertext message. Throughout this paper, we will discuss under this revised COA model.

Next we will study the relationships between these security notions: semantic security and indistinguishability. We first prove that they are not equivalent in general. On the other hand, while the well-known fact can be restated in our framework as the polynomial-time pseudo-randomness is *sufficient* to have the equivalence between semantic security and indistinguishability, the polynomial-time pseudo-randomness is not necessary for the equivalence. It is easy to see that "polynomial-time pseudo-randomness" has two aspects: "efficient samplability" and theoretically guaranteed "pseudo-randomness." We call the former property *samplability* simply and the latter *semi-randomness* to distinguish from both pseudo-randomness and quasi-randomness. Then we derive that semi-randomness *or* samplability is better sufficient condition for the equivalence between semantic security and indistinguishability.

In Section 2, we review the definitions of security notions of public-key encryption scheme. In Section 3, we introduce our new framework and reformulate well-known security notions to fit for the new framework. In Section 4, we discuss the relation between semantic security and indistinguishability in our new framework.

2 Preliminaries

2.1 Notations and Conventions

We introduce some useful notations and conventions for discussing probabilistic algorithms. If A is a probabilistic algorithm, then for any input x, the notation $A(x)$ refers to the probability space which assigns to the string y the probability that A, on input x, outputs y. If S is a probability space, denote by $\Pr_{e \leftarrow S}[e]$ (or $\Pr_S[e]$) the probability that S associates with element e. When we consider a finite probability space, it is convenient to consider separately the corresponding sample set and probability distribution on the set. If S is a finite set and D is a probability distribution on S, denote by $\Pr_{e \in_D S}[e]$ the probability that element $e \in S$ is chosen according to D. If S is a finite set, denote by $\Pr_{e \in_U S}[e]$ the probability that element $e \in S$ is chosen uniformly.

By 1^n we denote the unary representation of the integer n. A function $f : \{0,1\}^* \to \{0,1\}^*$ is *polynomially-bounded* if there exists a polynomial $p(\cdot)$ such that $|f(x)| \leq p(|x|)$ for all $x \in \{0,1\}^*$.

2.2 True Randomness Framework

In this paper, we will introduce a framework in which we can uniformly and comprehensively discuss "semantic security" and "indistinguishability" notions even for the case where some weak generator producing seemingly random sequences is used. (In order to avoid confusion, we will use throughout the paper the term "quasi-random" for referring the pseudo-randomness including one without any guarantee.) Usually any cryptographic notions are defined in the "true randomness framework," where the true randomness is available. We review the notions of semantic security and indistinguishability (in the true randomness framework) and the relation between them. We begin with the definition of public-key encryption schemes (in the true randomness framework).

Definition 1 (public-key encryption scheme). A *public-key encryption scheme* is a quadruple (G, M, E, D), where the following conditions hold.

1. G, called the *key generator*, is a probabilistic polynomial-time algorithm which, on input 1^n, outputs a pair of binary strings.
2. $M = \{M_n\}_{n \in N}$ is a family of message spaces from which all plaintext messages are drawn. In order to make our notation simpler (but without loss of generality), we assume that $M_n = \{0,1\}^n$.
3. For every n, for every pair (e, d) in the support of $G(1^n)$, and for any $\alpha \in M_n$, probabilistic polynomial-time (*encryption*) algorithm E and deterministic polynomial-time (*decryption*) algorithm D satisfy

$$\Pr_E[D(d, E(e, \alpha)) = \alpha] = 1,$$

where the probability is over the internal coin tosses of the algorithm E.

The integer n serves as the *security parameter* of the scheme. Each (e, d) in the range of $G(1^n)$ constitutes a pair of corresponding *encryption/decryption keys*. The string $E(e, \alpha)$ is the *encryption of the plaintext message* $\alpha \in \{0,1\}^*$ using the encryption key e, whereas $D(d, \beta)$ is the *decryption of the ciphertext message* β using the decryption key d.

Hereafter, we write $E_e(\alpha)$ instead of $E(e, \alpha)$ and $D_d(\beta)$ instead of $D(d, \beta)$. We also write $E_e(\alpha; r)$ when we want to express explicitly the randomness r of the encryption algorithm. Also, we let $G_1(1^n)$ denote the first element in the pair $G(1^n)$.

Since Goldwasser and Micali defined semantic security and polynomial security (a.k.a. indistinguishability), several ways to define such notions are shown. In this paper, we adopt a non-uniform formulation as in [10] in order to simplify the exposition. We note that employing such a non-uniform formulation (rather than a uniform one) may strengthen the definitions; yet, it does weaken the implications proven between the definitions, since proofs make free usage of non-uniformity.

A transformation is a uniform algorithm which, on input $\overline{C_n}$, outputs $\overline{C'_n}$, where $\overline{C_n}$ (resp., $\overline{C'_n}$) is the representation of a circuit C_n (resp., C'_n) in some standard encoding. Without loss of generality, we identify a circuit with its representation (in the standard encoding).

Definition 2 (semantic security). An encryption scheme (G, M, E, D) is *semantically secure* if there exists a probabilistic polynomial-time transformation T so that for every polynomial-size circuit family $\{C_n\}_{n\in N}$, for every probability ensemble $\{X_n\}_{n\in N}$ satisfying that X_n is a probability distribution on M_n, for every pair of polynomially-bounded functions $f, h : \{0,1\}^* \to \{0,1\}^*$, every polynomial $p(\cdot)$ and sufficiently large n,

$$\Pr_{G,E,X_n}\Big[C_n(G_1(1^n), E_{G_1(1^n)}(X_n), 1^n, h(X_n)) = f(X_n)\Big] < \Pr_{T,G,X_n}\Big[C'_n(G_1(1^n), 1^n, h(X_n)) = f(X_n)\Big] + \frac{1}{p(n)}$$

where $C'_n = T(C_n)$ is the circuit produced by T on input C_n. (The probability in the above terms is taken over X_n as well as over the internal coin tosses of the algorithms G and E.)

Definition 3 (indistinguishability). An encryption scheme (G, M, E, D) has *indistinguishable encryptions* if for every polynomial-size circuit family $\{C_n\}_{n\in N}$, for every polynomial $p(\cdot)$, for all sufficiently large n, and for every $x, y \in M_n$

$$\left|\Pr_{G,E}\big[C_n(G_1(1^n), E_{G_1(1^n)}(x)) = 1\big] - \Pr_{G,E}\big[C_n(G_1(1^n), E_{G_1(1^n)}(y)) = 1\big]\right| < \frac{1}{p(n)}.$$

The probability in the above terms is taken over the internal coin tosses of the algorithms G and E.

We have seen two notions of security for public-key encryption schemes. The above definitions only refer to the security of a scheme that is used to encrypt a single plaintext message (per key generated). Clearly, in reality, we want to encrypt many massages with the same key. The corresponding definitions of security notions in the multiple message setting have been given and discussed in [10]. Although it is important to consider the security in the multiple message setting, we will omit to discuss them on account of space constraints. The following theorem has been already shown [12,18].

Theorem 1 ([12,18]). *Let (G, M, E, D) be an encryption scheme. In the true randomness framework, (G, M, E, D) is semantically secure if and only if (G, M, E, D) has indistinguishable encryptions.*

The reductions that are used in the proof of the above theorem are randomized. Care must obviously be taken to confirm that the reductions still work for the case where some quasi-random generator is used instead of the true random resource.

3 Quasi-randomness Framework

In this section, we prepare a framework — quasi-randomness framework — in which we can uniformly and comprehensively discuss "semantic security" and "indistinguishability" notions even for the case where some weak quasi-random generator producing seemingly random sequences is used.

3.1 Public-Key Encryption Scheme for Quasi-random Set Family

We begin with introducing the notion of "quasi-randomness" and some notations. In this paper, a quasi-random string is just a string (of certain length) drawn from some subset of strings (of this length) uniformly at random. More specifically, we consider the following family of sets of strings.

Definition 4. Let $q(\cdot)$ be a polynomial. A *$q(n)$-quasi-random set family* (abbreviated QRSF) $\{R_n\}_{n\in N}$ is a family of sets of strings of length $q(n)$.

Below we usually use $\{R_n\}$ to denote some quasi-random set family. On the other hand, *$q(n)$-true-random set family* (abbreviated TRSF) is just a collection of sets $T_n = \{0,1\}^{q(n)}$. We use $\{T_n\}$ to denote some true-random set family.

Our ultimate purpose is to give a taxonomy of quasi-random set families from a viewpoint of the security of public-key encryption schemes. We have to enumerate some properties over quasi-random set families to begin with.

While the well-known fact can be restated in our framework as the polynomial-time pseudo-randomness is *sufficient* to have the equivalence between semantic security and indistinguishability, we show that the polynomial-time pseudo-randomness is not necessary to have the equivalence. This implies that there may

be more usable sufficient conditions for the equivalence. It is easy to consider separately "efficient samplability" and "pseudo-randomness" as some properties on quasi-randomness. We call the former property *samplability* simply and the latter *semi-randomness* to distinguish from both pseudo-randomness and quasi-randomness. "Samplability" is quite a natural property because generators without samplability is, in general, difficult to use. Especially in Monte-Carlo simulation, efficient samplability is required. On the other hand, "semi-randomness" is also one of important properties. Semi-random sequences pass many statistical tests. Some sequences that are obtained from physical sources such as electronic noise or the quantum effects in a semiconductor. When the sequences pass all known statistical tests, it is often that we use the sequences as "random sequences." We may consider that such sequences may have the semi-randomness property. So, in this paper, we study these two properties on QRSF.

We begin with definition of "semi-randomness." Semi-random sequences are ones that are not distinguished by any polynomial-size circuit. More specifically, we consider the following definition.

Definition 5. A $q(n)$-QRSF $\{R_n\}$ is said to be *semi-random* if for every polynomial-size circuit family $\{C_n\}_{n\in N}$, every polynomial $p(\cdot)$, all sufficiently large n,

$$\left| \Pr_{r\in_U R_n}\left[C_n(r)=1\right] - \Pr_{r'\in_U T_n}\left[C_n(r')=1\right] \right| < \frac{1}{p(n)},$$

where $\{T_n\}$ is $q(n)$-TRSF.

We note that semi-random sequences are different from output sequences by polynomial-time pseudo-random generators. Semi-random sequences need not to be recursive nor generated efficiently.

Next, we give a definition of "samplability." For any samplable sequence, there exists a (polynomial-size) generator $\{S_n\}_{n\in N}$ whose output is statistically close to the samplable sequence. More specifically, we consider the following definition.

Definition 6. A $q(n)$-QRSF $\{R_n\}$ is said to be *samplable* if there exists a polynomial-size circuit family $\{S_n\}_{n\in N}$ so that for every polynomial $p(\cdot)$ and all sufficiently large n,

$$\max_A \left\{ \left| \Pr_{r\in_U\{0,1\}^{q(n)}}\left[S_n(r)\in A\right] - \Pr_{r\in_U R_n}\left[r\in A\right] \right| \right\} < \frac{1}{p(n)},$$

where the maximum is taken all over the subsets of $\{0,1\}^{q(n)}$.

We note that the maximum value in the above definition is so called "statistical difference" between two probability distributions: $\{S_n(r)\}_{r\in_U\{0,1\}^{q(n)}}$ and the uniform distribution on R_n.

We extend the notion of public-key encryption scheme in order to cope with QRSF instead of the true randomness.

Definition 7 (public-key encryption scheme, revisited). A *public-key encryption scheme* is a quadruple (G, M, E, D), where the following conditions hold.

1. G, called the *key generator*, is a probabilistic polynomial-time algorithm which, on input 1^n, outputs a pair of binary strings. (Although the key generator also uses randomness, we disregard it here in order to cast light on roles of randomness in encrypting. So, we assume that randomness in key generator is always ideal.)
2. $M = \{M_n\}_{n\in N}$ is a family of message spaces from which all plaintext messages will be drawn. In order to make our notation simpler (but without loss of generality), we will assume that $M_n = \{0,1\}^n$.
3. For every $q(n)$-QRSF $\{R_n\}$, for every n, for every pair (e, d) in the support of $G(1^n)$ and for any $\alpha \in M_n$, "deterministic" polynomial-time (*encryption*) algorithm E and deterministic polynomial-time (*decryption*) algorithm D satisfy
$$\Pr_{r\in_U R_n}\Big[D_d(E_e(\alpha; r)) = \alpha\Big] = 1,$$
where the probability is over the uniform distribution on R_n.

We note that we treat the encryption algorithm as deterministic one fed with a plaintext message and a (random) supplementary input of length $q(n)$.

3.2 Security Notions for Quasi-random Set Family

In this subsection, we reformulate the notions of semantic security and indistinguishability to suit the framework of quasi-random set family.

Definition 8 (semantic security, revisited). An encryption scheme (G, M, E, D) is *semantically secure w.r.t. $q(n)$-quasi-random set family* $\{R_n\}$ if there exists a probabilistic polynomial-time transformation T so that every polynomial-size circuit family $\{C_n\}_{n\in N}$, for every probability ensemble $\{X_n\}_{n\in N}$ satisfying that X_n is a probability distribution on M_n, every pair of polynomially-bounded functions $f, h : \{0,1\}^* \to \{0,1\}^*$, every polynomial $p(\cdot)$ and all sufficiently large n,
$$\Pr_{G,X_n;r\in_U R_n}\Big[C_n(G_1(1^n), E_{G_1(1^n)}(X_n; r), 1^n, h(X_n)) = f(X_n)\Big] < \Pr_{T,G,X_n}\Big[C'_n(G_1(1^n), 1^n, h(X_n)) = f(X_n)\Big] + \frac{1}{p(n)}$$
where $C'_n = T(C_n)$.

Some explanation on the attack model is needed here. In the above definition, an adversary C_n is given only an encryption key $G_1(1^n)$ and a ciphertext message $E_{G_1(1^n)}(X_n; r)$ (and some supplementary information $h(X_n)$). Thus, it is considered as ciphertext only attack (COA) model. But note here that we may consider any polynomial-size circuit C_n for the adversary; hence, we may

assume that the encryption algorithm is also included in C_n. In the true randomness framework, this immediately includes the chosen plaintext attack (CPA) in which model the adversary can encrypt any plaintext messages of his choice. This is not true any more in the quasi-randomness framework because there is no guarantee that some (randomized) polynomial-size circuit can generate quasi-random strings in R_n uniformly at random.

Moreover, we consider our revised COA model. For our COA model, we consider the situation in which an adversary cannot access to the quasi-random generator. The situation means that those who use public-key encryption scheme have their *private* quasi-random generators. In general, they do not have to publicize their quasi-random generators which are used in public-key encryption scheme. In addition, the case where *private* quasi-random generators are used is more secure than the case where *public* quasi-random generators are used. Thus, we can say that our COA model makes sense.

Definition 9 (indistinguishability, revisited). An encryption scheme (G, M, E, D) has *indistinguishable encryptions w.r.t. $q(n)$-quasi-random set family* $\{R_n\}$ if for every polynomial-size circuit family $\{C_n\}_{n\in N}$, every polynomial $p(\cdot)$, all sufficiently large n and every $x, y \in M_n$,

$$\left| \Pr_{G;r\in_U R_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r)) = 1\right] - \Pr_{G;r'\in_U R_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(y;r')) = 1\right] \right| < \frac{1}{p(n)}.$$

We also note that C_n, in the above definition, cannot directly access to QRSF $\{R_n\}$.

The following notion is somewhat artificial. However, it is useful to characterize the notions of semantic security and indistinguishability.

Definition 10 (skew-indistinguishability). An encryption scheme (G, M, E, D) has *skew-indistinguishable encryptions w.r.t. $q(n)$-quasi-random set family* $\{R_n\}$ if for every polynomial-size circuit family $\{C_n\}_{n\in N}$, every polynomial $p(\cdot)$, all sufficiently large n and every $x, y \in M_n$,

$$\left| \Pr_{G;r\in_U R_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r)) = 1\right] - \Pr_{G;r'\in_U T_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(y;r')) = 1\right] \right| < \frac{1}{p(n)}$$

where $\{T_n\}$ is $q(n)$-TRSF.

In this paper, we do not consider the non-malleability, However, we only give the corresponding definition. We note that, in the definition below, "$r' \in R_n$" is optional, since the definition without it is alternative.

Definition 11 (non-malleability, revisited). An encryption scheme (G, M, E, D) is *non-malleable w.r.t. $q(n)$-quasi-random set family* $\{R_n\}$ if there exists

a probabilistic polynomial-time transformation T so that every polynomial-size circuit family $\{C_n\}_{n\in N}$, for every relation V that is decidable by a polynomial-size circuit family, for every probability ensemble $\{X_n\}_{n\in N}$ satisfying that X_n is a probability distribution on M_n, every polynomially-bounded function $h : \{0,1\}^* \to \{0,1\}^*$, every polynomial $p(\cdot)$ and all sufficiently large n,

$$\begin{aligned}\Pr_{G,X_n;r\in_U R_n}\Big[&C_n(G_1(1^n), E_{G_1(1^n)}(X_n;r), 1^n, h(X_n)) = E_{G_1(1^n)}(X'_n;r') \\ &\text{such that } V(X_n,X'_n)=1,\ X_n \neq X'_n \text{ and } r' \in R_n\Big] \\ < \Pr_{T,G,X_n}\Big[&C'_n(G_1(1^n), 1^n, h(X_n)) = E_{G_1(1^n)}(X'_n;r') \\ &\text{such that } V(X_n,X'_n)=1,\ X_n \neq X'_n \text{ and } r' \in R_n\Big] + \frac{1}{p(n)}\end{aligned}$$

where $C'_n = T(C_n)$.

We note that, in four definitions above, any adversary does not directly access to QRSF $\{R_n\}$ but gets ciphertext messages encrypted using QRSF $\{R_n\}$ as challenge inputs.

4 Properties in Quasi-randomness Framework

4.1 Relations among Security Notions

In this subsection, we consider classes of pairs of QRSF and public-key encryption schemes w.r.t. the QRSF. We will especially show that semantic security and indistinguishability (in the quasi-randomness framework) are separable from each other.

We denote by $\mathcal{SS}_q$ the class of pairs of encryption scheme (G,M,E,D) and QRSF $\{R_n\}$ satisfying that (G,M,E,D) w.r.t. $\{R_n\}$ is semantically secure. We also denote $\langle (G,M,E,D), \{R_n\}\rangle \in \mathcal{SS}_q$ if an encryption scheme (G,M,E,D) which is semantically secure w.r.t. a QRSF $\{R_n\}$. We denote by $\mathcal{IND}_{qq}$ the class of pairs of encryption schemes (G,M,E,D) and QRSF $\{R_n\}$ satisfying that (G,M,E,D) w.r.t. QRSF $\{R_n\}$ has indistinguishable encryptions. We denote by $\mathcal{IND}_{qt}$ the class of pairs of encryption scheme (G,M,E,D) and QRSF $\{R_n\}$ satisfying that (G,M,E,D) w.r.t. QRSF $\{R_n\}$ has skew-indistinguishable encryptions.

Theorem 2. $\mathcal{IND}_{qt} \subsetneq \mathcal{SS}_q \subsetneq \mathcal{IND}_{qq}$.

The above theorem follows the four lemmas below.

Lemma 1. $\mathcal{SS}_q \subseteq \mathcal{IND}_{qq}$.

Lemma 2. $\mathcal{IND}_{qt} \subseteq \mathcal{SS}_q$.

Lemma 3. $\mathcal{IND}_{qq} \setminus \mathcal{SS}_q \neq \emptyset$.

Lemma 4. $\mathcal{SS}_q \setminus \mathcal{IND}_{qt} \neq \emptyset$.

Proof. **(Lemma 1)** We show that if an encryption scheme (G, M, E, D) w.r.t. $\{R_n\}$ is semantically secure then (G, M, E, D) w.r.t. $\{R_n\}$ has indistinguishable encryptions.

Now, we assume that (G, M, E, D) w.r.t. $\{R_n\}$ does not have indistinguishable encryptions; namely, there exist a polynomial-size circuit family $\{D_n\}_{n\in N}$ and a polynomial $p(\cdot)$ such that for infinitely many n, there exist x_n and $\tilde{x}_n$ satisfying

$$\left| \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x_n;r)) = 1\right] - \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x}_n;r)) = 1\right] \right| > \frac{1}{p(n)}.$$

Without loss of generality, for infinitely many n, there exist x_n and $\tilde{x}_n$ satisfying

$$\Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x_n;r)) = 1\right] - \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x}_n;r)) = 1\right] > \frac{1}{p(n)}.$$

Let X_n be a random variable such that $\Pr[X_n = x_n] = \Pr[X_n = \tilde{x}_n] = 1/2$. Let f be a function such that $f(x_n) = 1$ and $f(\tilde{x}_n) = 0$. Now, we consider the following circuit C_n. On input $(e, E_e(x;r))$, the new circuit C_n feeds D_n with input $(e, E_e(x;r))$ and output 1 if D_n outputs 1; otherwise, C_n outputs 0. It is left to estimate the probability that $C_n(e, E_e(x;r)) = f(x)$ when x is drawn according to X_n.

$$\begin{aligned}
&\Pr_{\substack{G;x\in_{X_n}\{x_n,\tilde{x}_n\}\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r)) = f(x)\right] \\
&\quad= \frac{1}{2}\cdot \Pr_{G;r\in_U R_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x_n;r)) = f(x_n)\right] \\
&\qquad + \frac{1}{2}\cdot \Pr_{G;r\in_U R_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x}_n;r)) = f(\tilde{x}_n)\right] \\
&\quad= \frac{1}{2}\Big(\Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x_n;r)) = 1\right] + 1 \\
&\qquad\qquad - \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x}_n;r)) = 1\right]\Big) \\
&\quad\geq \frac{1}{2} + \frac{1}{2p(n)}.
\end{aligned}$$

In contrast, for every (randomized) circuit C'_n, $\Pr[C'_n(G_1(1^n)) = f(X_n)] \leq 1/2$. This contradicts the hypothesis that the scheme is semantically secure. □

Proof. **(Lemma 2)** We show that if (G, M, E, D) w.r.t. $q(n)$-QRSF $\{R_n\}$ has skew-indistinguishable encryptions then (G, M, E, D) w.r.t. $q(n)$-QRSF $\{R_n\}$ is semantically secure.

Now, we assume that there exist a polynomial-size circuit family $\{C_n\}_{n\in N}$, a polynomial $p(\cdot)$, and polynomially-bounded functions f, h such that for infinitely many n,

$$\Pr_{\substack{T,G;x\in_{X_n}M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r), 1^n, h(x)) = f(x)\right] - \Pr_{\substack{G;x\in_{X_n}M_n\\ r\in_U R_n}}\left[C'_n(G_1(1^n), 1^n, h(x)) = f(x)\right] > \frac{1}{p(n)}.$$

Now, we consider the following circuit $C'_{n,r}$. $C'_{n,r}$ feeds C_n with input $(e, E_e(1^n; r), 1^n, h(x))$ and outputs a value that C_n outputs. Thus it is easy to transform C_n to $C'_{n,r}$ in probabilistic polynomial time. Then

$$\Pr_{\substack{G;x\in_{X_n}M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r), 1^n, h(x)) = f(x)\right] - \Pr_{\substack{G;x\in_{X_n}M_n\\ r\in_U T_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), 1^n, h(x)) = f(x)\right] > \frac{1}{p(n)},$$

where $\{T_n\}$ is $q(n)$-TRSF. Let x_n be a string for which the difference above is maximum over X_n. Using this x_n, we construct a new circuit D_n as follows. On input $(e, E_e(\alpha; r))$, D_n feeds C_n with input $(e, E_e(\alpha; r), 1^n, h(x_n))$ and outputs 1 if C_n outputs $f(x_n)$; otherwise D_n outputs 0. Then

$$\Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x_n;r)) = 1\right] - \Pr_{G;r\in_U T_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(1^n;r)) = 1\right] > \frac{1}{p(n)}.$$

This contradicts the hypothesis that the encryption scheme has skew-indistinguishable encryptions. □

Proof. **(Lemma 3)** Suppose that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{SS}_q$, where $\{T_n\}$ is $q(n)$-TRSF. Then there exists an encryption scheme (G, M, E', D') such that $\langle (G, M, E', D'), \{T'_n\}\rangle \in \mathcal{SS}_q$, where $\{T'_n\}$ is $(q(n) + 1)$-TRSF, $E'_e(\alpha; r) = E_e(\alpha; r_1)r_2$, $D'_d(\beta) = D_d(\beta')$, $r = r_1r_2$, $|r_2| = 1$, and β' is the prefix of β of length $|\beta| - 1$. Let V be a non-BPP subset (i.e., tally set) of 1^*. We consider a QRSF $\{R_n\} = \{\{0,1\}^{q(n)}b\}$, where $b = 1$ if $1^n \in V$; $b = 0$ otherwise. Since the last bit of the ciphertext message is always constant of all ciphertext messages specified by security parameter n, any distinguishing circuit cannot use the last bit of the ciphertext message. ¿From Lemma 1, it follows that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{INTD}_{qq}$. Thus $\langle (G, M, E', D'), \{R_n\}\rangle \in \mathcal{INTD}_{qq}$.

On the other hand, $\langle (G, M, E', D'), \{R_n\}\rangle \notin \mathcal{SS}_q$. We will show this by contradictory. We assume that $\langle (G, M, E', D'), \{R_n\}\rangle \in \mathcal{SS}_q$. In other words, there exists a probabilistic polynomial transformation T such that for every polynomial-size circuit $\{C_n\}_{n\in N}$, for every probability ensemble $\{X_n\}_{n\in N}$ satisfying that X_n is a probability distribution on M_n, for every pair of polynomially-bounded functions $f, h : \{0,1\}^* \to \{0,1\}^*$, every polynomial $p(\cdot)$ and sufficiently

large n,

$$\Pr_{G,X_n;r\in_U R_n}\Big[C_n(G_1(1^n), E_{G_1(1^n)}(X_n;r), 1^n, h(X_n)) = f(X_n)\Big] < \Pr_{T,G,X_n}\Big[C'_n(G_1(1^n), 1^n, h(X_n)) = f(X_n)\Big] + \frac{1}{p(n)}$$

where $C'_n = T(C_n)$. Here, we consider a circuit family $\{C_n\}_{n\in N}$, probability ensemble $\{X_n\}_{n\in N}$, polynomially-bounded functions f, h satisfying the following: C_n outputs the last bit of the ciphertext message; $\Pr[X_n = 1^n] = 1$; h is constant; and $f(X_n) = 1$ if $1^n \in V$, $f(X_n) = 0$ otherwise. Then C_n always computes $f(X_n)$ correctly. We can say that there exists a probabilistic polynomial transformation T such that for sufficiently large n,

$$1 - \Pr_{G,T}\Big[T(C_n)(G_1(1^n), 1^n) = f(1^n)\Big] < \frac{1}{p(n)}.$$

Since $\{C_n\}_{n\in N}$ can be implemented by a (uniform) constant size circuit family, $T(\{C_n\}_{n\in N})$ can be also implemented by a (uniform) probabilistic polynomial-time algorithm B. Thus, we can say that B computes the membership of a non-BPP tally set. This is a contradiction. Therefore $\langle (G, M, E', D'), \{R_n\}\rangle \notin \mathcal{SS}_q$. □

Proof. **(Lemma 4)** Suppose that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{IND}_{qt}$, where $\{T_n\}$ is $q(n)$-TRSF. Then there exists an encryption scheme (G, M, E', D') such that $\langle (G, M, E', D'), \{T'_n\}\rangle \in \mathcal{IND}_{qt}$, where $\{T'_n\}$ is $(q(n) + 1)$-TRSF, $E'_e(\alpha; r) = E_e(\alpha; r_1)r_2$, $D'_d(\beta) = D_d(\beta')$, $r = r_1r_2$, $|r_2| = 1$, and β' is the prefix of β of length $|\beta| - 1$. We consider a QRSF $\{R_n\} = \{\{0,1\}^{q(n)}1\}$. It is easy to see that $\langle (G, M, E', D'), \{R_n\}\rangle \notin \mathcal{IND}_{qt}$ because a distinguisher can use the last bit of the ciphertext message.

On the other hand, from Lemma 2, it follows that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{SS}_q$. Since, in the scheme (G, M, E', D'), the last bit of the ciphertext message gives no information on the plaintext message, $\langle (G, M, E', D'), \{R_n\}\rangle \in \mathcal{SS}_q$. □

4.2 Properties of QRSF and Their Effects on the Security

In this subsection, we consider how properties of QRSF affect on the security of encryption schemes. We will especially give a sufficient condition that semantic security and indistinguishability become equivalent in the quasi-randomness framework.

Theorem 3. *Suppose that* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{qq}$. *If* $\{R_n\}$ *is semi-random, then* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{qt}$.

We note that since the true randomness is semi-random, the equivalence between semantic security and indistinguishability (w.r.t. the true randomness) can be shown as a corollary of Lemmas 1, 2 and Theorem 3. The above theorem

says that if an encryption scheme is semantically secure in the true-randomness framework and we use "semi-random" sequence as random inputs to the encryption algorithm then the encryption scheme is still semantically secure in the quasi-randomness framework.

Proof. We show that if an encryption scheme (G, M, E, D) w.r.t. $\{R_n\}$ has indistinguishable encryptions and $\{R_n\}$ is semi-random, then (G, M, E, D) w.r.t. $\{R_n\}$ has skew-indistinguishable encryptions.

Now, we assume that there exist a polynomial-size circuit family $\{D_n\}_{n\in N}$ and a polynomial $p(\cdot)$ such that for infinitely many n and for some $x, \tilde{x} \in M_n$

$$\left| \Pr_{G;r'\in_U T_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x;r')) = 1\right] - \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x};r)) = 1\right] \right| > \frac{1}{p(n)}.$$

If $x = \tilde{x}$ then it is easy to construct a polynomial-size circuit from D_n and E_e to distinguish r' and r using the circuit. This contradicts that $\{R_n\}$ is semi-random. Thus we have only to consider the case $x \neq \tilde{x}$.

Since $\{R_n\}$ is semi-random, for any polynomial $p'(\cdot)$ such that $p(n) < p'(n)$,

$$\left| \Pr_{G;r'\in_U T_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x;r')) = 1\right] - \Pr_{G;r\in_U R_n}\left[D_n(E_{G_1(1^n)}(x;r)) = 1\right] \right| < \frac{1}{p'(n)}.$$

Therefore, there exists a polynomial $p''(\cdot)$ such that

$$\left| \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x;r)) = 1\right] - \Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(\tilde{x};r)) = 1\right] \right| > \frac{1}{p(n)} - \frac{1}{p'(n)} > \frac{1}{p''(n)}.$$

This contradicts the hypothesis that the scheme has indistinguishable encryptions. □

Theorem 4. *Suppose that* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{qq}$. *If* $\{R_n\}$ *is samplable, then* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_q$.

The above theorem says that if the combination of the encryption scheme and quasi-randomness sequences as random inputs to the encryption algorithm has indistinguishable encryptions then the combined encryption scheme is semantically secure in the quasi-randomness framework. Namely, the above theorem offers us another way to show that the encryption scheme is semantically secure (in the quasi-randomness framework). The above theorem also says that the property of "semi-randomness" for random inputs to the encryption algorithm is not essential. It is open to further discussion whether or not the combined encryption schemes are semantically secure in the quasi-randomness framework even though the quasi-random sequences are not semi-random or even though the quasi-random sequences have not been proved to be semi-random yet.

Proof. We show that if (G, M, E, D) w.r.t. $q(n)$-QRSF $\{R_n\}$ has indistinguishable encryptions and $\{R_n\}$ is samplable then (G, M, E, D) w.r.t. $q(n)$-QRSF $\{R_n\}$ is semantically secure.

Now, we assume that, for any transformation T, there exist a polynomial-size circuit family $\{C_n\}$, a polynomial $p(\cdot)$, and polynomially-bounded functions f, h such that for infinitely many n,

$$\Pr_{\substack{G;x\in_X M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r), 1^n, h(x)) = f(x)\right] - \Pr_{\substack{T,G;x\in_X M_n\\ r\in_U R_n}}\left[C'_n(G_1(1^n), 1^n, h(x)) = f(x)\right] > \frac{1}{p(n)},$$

where $C'_n = T(C_n)$. Now, we consider the following circuit $C'_{n,r'}$. $C'_{n,r'}$ feeds C_n with input $(e, E_e(1^n; r))$ and outputs a value that C_n outputs. Since $\{R_n\}$ is samplable, $r \in R_n$ is samplable in polynomial time using the truly random r'. Thus it is easy to transform C_n to $C'_{n,r'}$ in probabilistic polynomial time. Then

$$\Pr_{\substack{G;x\in_X M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r), h(x)) = f(x)\right] - \Pr_{\substack{G;x\in_X M_n\\ r'\in_U\{0,1\}^{q(n)};r\leftarrow S_n(r')}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), h(x)) = f(x)\right] > \frac{1}{p(n)},$$

where S_n is the sampling circuit. Since the statistical difference between $\{S_n(r)\}$ and the uniform distribution on R_n is less than $1/4p(n)$ (actually it is less than $1/p'(n)$ for any polynomial $p'(\cdot)$), we have,

$$\begin{aligned}
&\Pr_{\substack{G;x\in_X M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(x;r), 1^n, h(x)) = f(x)\right]\\
&\qquad - \Pr_{\substack{G;x\in_X M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), 1^n, h(x)) = f(x)\right]\\
&> \Pr_{\substack{G;x\in_X M_n\\ r'\in_U\{0,1\}^{q(n)};r\leftarrow S_n(r')}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), 1^n, h(x)) = f(x)\right]\\
&\qquad - \Pr_{\substack{G;x\in_X M_n\\ r\in_U R_n}}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), 1^n, h(x)) = f(x)\right] + \frac{1}{p(n)}\\
&> \frac{1}{p(n)} - \sum_r\left(\Pr_{G;x\in_X M_n}\left[C_n(G_1(1^n), E_{G_1(1^n)}(1^n;r), 1^n, h(x)) = f(x)\right]\cdot \left|\Pr[r\leftarrow S_n(r')] - \Pr[r\in_U R_n]\right|\right)\\
&> \frac{1}{p(n)} - \sum_r\left|\Pr[r\leftarrow S_n(r')] - \Pr[r\in_U R_n]\right|\\
&= \frac{1}{p(n)} - 2\cdot\max_A\left\{\left|\Pr_{r'\in_U\{0,1\}^{q(n)}}\left[S_n(r')\in A\right] - \Pr_{r\in_U R_n}\left[r\in A\right]\right|\right\} > \frac{1}{2p(n)}.
\end{aligned}$$

Let x_n be a string for which the difference above is maximum over X_n. Using this x_n, we construct a new circuit D_n as follows. On input $(e, E_e(\alpha; r))$, D_n

feeds C_n with input $(e, E_e(\alpha; r), 1^n, h(x_n))$ and outputs 1 if C_n outputs $f(x_n)$; otherwise D_n outputs 0. Then

$$\Pr_{G;r\in_U R_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(x_n; r)) = 1\right] - \Pr_{G;r\in_U T_n}\left[D_n(G_1(1^n), E_{G_1(1^n)}(1^n; r)) = 1\right] > \frac{1}{2p(n)}.$$

This contradicts the hypothesis that the scheme has indistinguishable encryptions. □

As a corollary, we have the following. The below gives us a better sufficient condition for the equivalence between semantic security and indistinguishability.

Corollary 1. *Suppose that* $\{R_n\}$ *is semi-random or samplable. Then* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{qq}$ *if and only if* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{SS}_q$.

Theorem 5. *There exists* $\langle (G, M, E, D), \{R_n\}\rangle \in \mathcal{IND}_{qt}$ *such that* $\{R_n\}$ *is not semi-random.*

Although we have a better sufficient condition for the equivalence between semantic security and indistinguishability, the condition is not necessary for the equivalence. The above theorem actually says that neither semi-randomness nor polynomial-time pseudo-randomness is necessary for the equivalence.

Proof. Suppose that $\langle (G, M, E, D), \{T_n\}\rangle \in \mathcal{IND}_{qt}$, where $\{T_n\}$ is $q(n)$-TRSF. Then there exists an encryption scheme (G, M, E', D') such that $\langle (G, M, E', D'), \{T'_n\}\rangle \in \mathcal{IND}_{qt}$, where $\{T'_n\}$ is $(q(n)+1)$-TRSF, $E'_e(\alpha; r) = E_e(\alpha; r_1)$, $D'_d(\beta) = D_d(\beta)$, $r = r_1 r_2$ and $|r_2| = 1$. We consider a QRSF $\{R_n\} = \{\{0,1\}^{q(n)}1\}$. It is easy to see that $\langle (G, M, E', D'), \{R_n\}\rangle \in \mathcal{IND}_{qt}$, because the last bit of the supplementary random input is not used in encrypting.

On the other hand, it is easy to see that $\{R_n\}$ and $\{T'_n\}$ are distinguishable. In other words, $\{R_n\}$ is not semi-random. □

5 Concluding Remarks

We have introduced a framework in which we can uniformly and comprehensively discuss security notions of public-key encryption schemes even for the case where some weak generator producing seemingly random sequences is used to encrypt plaintext messages. Since the new framework separates chosen plaintext attack and ciphertext only attack, we consider the security under the COA model in the framework. We have proved that indistinguishability and semantic security are not equivalent in general. On the other hand, we have derived some sufficient condition for the equivalence and shown that polynomial-time pseudo-randomness is not always necessary for the equivalence.

The discussion has been restricted on the case of ciphertext only attack, so we will consider the case of chosen plaintext attack and chosen ciphertext attack. We will also consider non-malleability [8] in the new framework.

Acknowledgments

I would like to thank Osamu Watanabe for helpful comments and suggestions on this paper drafts.

References

1. M. Bellare, A. Boldyreva, and S. Micali. Public-key encryption in a multi-user setting: Security proofs and improvements. In B. Preneel, editor, *Advances in Cryptology — EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 259–274. Springer-Verlag, 2000.
2. M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway. Relations among notions of security for public-key encryption schemes. In H. Krawczyk, editor, *Advances in Cryptology — CRYPTO'98*, volume 1462 of *Lecture Notes in Computer Science*, pages 26–45. Springer-Verlag, 1998.
3. M. Bellare, S. Goldwasser, and D. Micciancio. Pseudo-random number generation within cryptographic algorithms: The DSS case. In B. S. Kaliski Jr., editor, *Advances in Cryptology — CRYPTO'97*, volume 1294 of *Lecture Notes in Computer Science*, pages 277–291. Springer-Verlag, 1997.
4. M. Bellare and A. Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In M. Wiener, editor, *Advances in Cryptology — CRYPTO'99*, volume 1666 of *Lecture Notes in Computer Science*, pages 519–536. Springer-Verlag, 1999.
5. L. Blum, M. Blum, and M. Shub. A simple unpredictable pseudo-random number generator. *SIAM Journal on Computing*, 15(2):364–383, 1986.
6. M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM Journal on Computing*, 13(4):850–864, 1984.
7. J. Boyar. Inferring sequences produced by pseudo-random number generators. *Journal of the Association for Computing Machinery*, 36(1):129–141, 1989.
8. D. Dolev, C. Dwork, and M. Naor. Non-malleable cryptography. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, pages 542–552. ACM Press, 1991.
9. A. M. Frieze, J. Hastad, R. Kannan, J. C. Lagarias, and A. Shamir. Reconstructing truncated integer variables satisfying linear congruences. *SIAM Journal on Computing*, 17(2):262–280, 1988.
10. O. Goldreich. *Foundation of Cryptography (Fragment of a Book – Version 2.03)*, 1998.
11. O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *Journal of the Association for Computing Machinery*, 33(4):792–807, 1986.
12. S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.
13. D. E. Knuth. *The Art of Computer Programming*, volume 2. Seminumerical Algorithms. Addison-Wesley, 3rd edition, 1998.
14. T. Koshiba. A theory of randomness for public key cryptosystems: The ElGamal cryptosystem case. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E83-A(4):614–619, 2000.
15. H. Krawczyk. How to predict congruential generators. *Journal of Algorithms*, 13(4):527–545, 1992.
16. M. Luby. *Pseudorandomness and Cryptographic Applications*. Princeton Univ. Press, 1996.

17. A. J. Menezes, P. C. van Oorschot, and S. A. Vanestone. *Handbook of Applied Cryptography*. CRC Press, 1997.
18. S. Micali, C. Rackoff, and B. Sloan. The notion of security for probabilistic cryptosystems. *SIAM Journal on Computing*, 17(2):412–426, 1988.
19. M. Naor and M. Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, pages 427–437. ACM Press, 1990.
20. C. Rackoff and D. R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In J. Feigenbaum, editor, *Advances in Cryptology — CRYPTO'91*, volume 576 of *Lecture Notes in Computer Science*, pages 433–444. Springer-Verlag, 1992.
21. A. Shamir. On the generation of cryptographically strong pseudorandom sequences. *ACM Transactions on Computer Systems*, 1(1):38–44, 1983.
22. J. Stern. Secret linear congruential generators are not cryptographically secure. In *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science*, pages 421–426. IEEE Computer Society Press, 1987.
23. A. C. Yao. Theory and applications of trapdoor functions. In *Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pages 80–91. IEEE Computer Society Press, 1982.