PROBABILISTIC CRYTOSYSTEMS

JINHO KIM AND JONATHAN ZHAO

Introduction

These are notes from an undergraduate summer research project at Western University on probabilistic cryptography. The notes were prepared by Jinho Kim and Jonathan Zhao, under the supervision of Chris Kapulkin.

Probabilistic encryption was introduced in 1982 by Goldwasser and Micali [GM82]. The key idea behind probabilistic cryptosystems is that each message may be encrypted into different ciphertexts. In the case of the Goldwasser–Micali cryptosystem, the message space is simply $\{0, 1\}$. If the standard RSA system were used to encrypt these messages, the adversary could simply compute the encryptions of 0 and 1, and compare them to the publicly available ciphertext. The Goldwasser–Micali system works around this by randomizing the encryption, i.e., allowing half of the elements in $\mathbb{Z}_n^*$ to be a valid encryption of a single bit.

These notes present the four probabilistic cryptosystem:

- Goldwasser–Micali [GM82];
- Okamoto–Uchiyama [OU98];
- Paillier [Pai99];
- Groth [Gro05].

We are presenting them in chronological order to show how each can be thought of an improvement on the previous ones.

The style of our presentation roughly follows the textbook "Introduction to Mathematical Cryptography" by Hoffstein, Pipher, and Silverman [HPS14], e.g., we display each cryptosystem in a table. We also assume that the reader is familiar with the content of Chapters 1–3 of this textbook [HPS14].

Organization. We begin in Section 1 by summarizing the necessary background from cryptography and elementary number theory. We then survey each cryptosystem in a separate section: Goldwasser–Micali in Section 2, Okamoto–Uchiyama in Section 3, Paillier in Section 4, and Groth in Section 5.

Acknowledgement. We are grateful to Jefrey Judes for helping with earlier revisions, Chris Kapulkin for providing mathematical guidance, and Jens Groth for helpful correspondence. We also thank the Natural Science and Engineering Research Council (NSERC) for providing the funding for this project.

The latest version of these notes can be found at:

sites.google.com/view/westerncrypto/papers

If you have found any typos or mistakes, please let us know at jkim2492@uwo.ca or jzhao293@uwo.ca.

1. Background

We begin by introducing some notations. We use $x \xleftarrow{\mathsf{R}} X$ to represent the fact that $x \in X$ is chosen randomly. The order of an element g in the group $\mathbb{Z}_n^*$ is denoted $\mathrm{ord}_n\ g$. We assume all algorithms to be probabilistic and polynomial time unless stated otherwise. Finally, given $x \in \mathbb{Z}_n^*$, we write $\overline{x}$ for the inverse of x in $\mathbb{Z}_n^*$.

There is a common structure to all of our cryptosystems. In each case, we begin by fixing an integer k, the bit-length of our primes (e.g., in RSA, one takes $k = 512$). The set of primes of bit-length k will be denoted $\mathcal{P}_k$, and p and q will be used to denote its typical elements. We will use n to represent the encryption modulus (for instance, $n = pq$ in RSA).

Each encryption function raises a certain element to the power of the message before randomizing. We call the set of suitable such elements the encryption base and denote it by $\mathcal{B}$. With these notations in place, the public key for each of our cryptosystems will be of the form (n, g) where $g \in \mathcal{B}$.

Crytosystems are proven to be secure by showing that "breaking" them would be harder than mathematical problems generally considered difficult. We make the related notions precise as follows.

Let $X = X_1 \times ... \times X_t$ be a finite product of subsets of $\mathbb{N}$ and let Y be a set. A *problem* refers to a function $\mathsf{P}\colon X \to Y$ where X represents the set of inputs, and $\mathsf{P}(\mathbf{x})$ is the set of solutions corresponding to input $\mathbf{x}$. If Y has two elements, then P is called a *decisional problem*.

Next, we define when a problem can be considered hard.

Definition 1.1. A function $f\colon \mathbb{N} \to \mathbb{R}$ is *negligible* with respect to $k \in \mathbb{N}$ if for all $c \in \mathbb{N}$, there exists $M \in \mathbb{Z}$ such that $f(k) < \frac{1}{k^c}$ whenever $k > M$.

Definition 1.2. A problem $\mathsf{P}\colon X \to Y$ is *intractable* with respect to an input $x \in X_i$ if for all algorithms A, when the inputs in $\{X_j \mid j \neq i\}$ are held fixed, the probability function

$$\Pr\left[\, \mathsf{A}(\mathbf{x}) \in \mathsf{P}(\mathbf{x}) \,\right]$$

is non-negligible with respect to $x \in X_i$.

We consider the following two notions of security: one-way encryption and semantic security. Informally, an encryption is one-way if it is not possible to invert the encryption, i.e., to compute the original message given the ciphertext and public keys with non-negligible probability over guessing. And it is semantically secure if it is not possible to distinguish between encryptions of any two distinct messages given the public keys with non-negligible advantage over guessing.

To formalize these notions, consider a fixed cryptosystem. Let $\mathcal{M}$ be the message space of the cryptosystem and $\mathsf{Enc}(n, g, m)$ be the encryption of $m \in \mathcal{M}$ using a parameter $g \in \mathcal{B}$.

Definition 1.3. Let $m \in \mathcal{M}$. *Inverting the encryption* refers to the following computational problem of, given $g \xleftarrow{\mathsf{R}} \mathcal{B}$ and $\mathsf{Enc}(n, g, m)$, computing m. If inverting the encryption is intractable, then Enc is called a *one-way function*.

Definition 1.4.

(1) Let $b \xleftarrow{\mathsf{R}} \{0, 1\}$. A *semantic security algorithm* is a pair of algorithms $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ where
- $\mathsf{A}_1(\mathcal{M})$ returns a pair of distinct messages $m_0, m_1 \in \mathcal{M}$;
- $\mathsf{A}_2(n, g, \mathsf{Enc}(n, g, m_b))$ returns 0 or 1.

(2) The cryptosystem is *semantically secure* if for all semantic security algorithms A,

$$\Pr\left[\, \mathsf{A}_2(n, g, m_0, m_1, \mathsf{Enc}(n, g, m_b)) = b \mid g \xleftarrow{\mathsf{R}} \mathcal{B} \,\right]$$

is negligible with respect to k. From here on, we omit input arguments n, m_0, and m_1 i.e., A_2 will only take in $g \in \mathcal{B}$ and $\mathsf{Enc}(n, g, m_b)$.

Next, we introduce the number theoretic concepts used in the cryptosystems we cover.

Definition 1.5. Let $x, n \in \mathbb{Z}$. We say x is a *quadratic residue* $\bmod n$ if there is $y \in \mathbb{Z}$ such that $x \equiv y^2 \bmod n$. Otherwise, x is a *quadratic non-residue* $\bmod n$.

Lemma 1.6 ([HPS14, Section 3.9])**.** *If* QR *denotes a quadratic residue and* NR *denotes a non-residue, then*

$$\mathsf{QR} \cdot \mathsf{QR} = \mathsf{QR},\ \mathsf{QR} \cdot \mathsf{NR} = \mathsf{NR}, \textit{ and } \mathsf{NR} \cdot \mathsf{NR} = \mathsf{QR}.$$

It follows that half of the elements in $\mathbb{Z}_p^*$ are residues while the other half are non-residues. We generalize this property as follows.

Definition 1.7. Let p be an odd prime and $a \in \mathbb{Z}$. The *Legendre symbol of* a is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p \\ -1 & \text{if } a \text{ is a non-residue modulo } p \\ 0 & \text{if } p \mid a. \end{cases}$$

To extend the idea of multiplicative property to composite modulus, we define the Jacobi symbol.

Definition 1.8. Let $a \in \mathbb{Z}$ and n be an odd positive integer with a prime factorization $\prod p_i^{e_i}$. The *Jacobi symbol of* a is defined as the product of legendre symbols of a with respect to prime factors of n i.e.,

$$\left(\frac{a}{n}\right) = \prod \left(\frac{a}{p_i}\right)^{e_i}.$$

For computing Legendre and Jacobi symbols, see [HPS14, Section 3.9].

2. Goldwasser–Micali

The Goldwasser–Micali cryptosystem is introduced in [HPS14, Section 3.10] but we describe it here in greater detail. In order to introduce the parameters efficiently, we write $\mathbb{QR}_n$ for the subset of quadratic residues in $\mathbb{Z}_n^*$ and $\mathbb{J}_n$ for the subset whose elements have Jacobi symbol of 1. Clearly, $\mathbb{QR}_n \subset \mathbb{J}_n$.

This cryptosystem uses modulus $n = pq$, encryption base $\mathcal{B} = \mathbb{J}_n \setminus \mathbb{QR}_n$, and message space $\mathcal{M} = \{0, 1\}$.

Alice	Bob
Chooses $p, q \xleftarrow{\mathsf{R}} \mathcal{P}_k$. Chooses $g \xleftarrow{\mathsf{R}} \mathcal{B}$. Computes $n = pq$. Publishes (n, g).	
	Chooses $r \xleftarrow{\mathsf{R}} \{2, ..., n-1\}$. Chooses $m \in \mathcal{M}$. Computes and sends $C = g^m r^2 \bmod n$.
Computes $\left(\frac{C}{p}\right)$. $m = \begin{cases} 0 & \text{if } \left(\frac{C}{p}\right) = 1. \\ 1 & \text{if } \left(\frac{C}{p}\right) = -1. \end{cases}$	

The primes p and q are generated just as in the RSA, by choosing random numbers and testing for primality. Elements $g \in \mathcal{B}$ are chosen in a similar way. Indeed, a randomly selected $g \in \mathbb{Z}_n$ is a valid encryption base with probability $\frac{1}{4}$ by Lemma 1.6.

The correctness of the decryption can be verified as

$$\left(\frac{C}{p}\right) = \begin{cases} \left(\frac{r^2}{p}\right) = \left(\frac{r}{p}\right)^2 = 1 & \text{if } m = 0 \\ \left(\frac{gr^2}{p}\right) = \left(\frac{g}{p}\right)\left(\frac{r^2}{p}\right) = \left(\frac{g}{p}\right)\left(\frac{r}{p}\right)^2 = (-1)(1) = -1 & \text{if } m = 1. \end{cases}$$

Security of encryption follows from the fact that an eavesdropper cannot compute $\left(\frac{C}{p}\right)$ without knowing the factorization of n. She is, however, able to compute $\left(\frac{C}{n}\right)$ which is always 1. Thus, the problem of distinguishing quadratic residues from arbitrary elements with Jacobi symbol 1 is the underlying hardness assumption of our cryptosystem. Now, we formally define the problem.

Definition 2.1. Let $X \xleftarrow{\mathsf{R}} \{\mathbb{QR}_n, \mathbb{J}_n \setminus \mathbb{QR}_n\}$. The *quadratic residuosity problem* is the decisional problem of, given $x \xleftarrow{\mathsf{R}} X$, deciding whether $X = \mathbb{QR}_n$ or $X = \mathbb{J}_n \setminus \mathbb{QR}_n$.

Theorem 2.2. *If the quadratic residuosity problem is intractable, then the Goldwasser-Micali cryptosystem is semantically secure.*

Proof. Let $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ be a semantic security algorithm with a non-negligible advantage $\varepsilon > 0$. Suppose $X \xleftarrow{\mathsf{R}} \{\mathbb{QR}_n\,,\, \mathbb{J}_n \setminus \mathbb{QR}_n\}$ and we are given $x \xleftarrow{\mathsf{R}} X$. Let a deciding algorithm B be constructed as:

```
Algorithm B(x)
      Call A_1({0,1}) to get (m_0, m_1).
      Choose b ←R {0,1}.
      Compute C = x^{m_b} 2^2 mod n
      Compute y = A_2(x, C).
      If y ≠ b, then return QR_n.
      If y = b, then return J_n \ QR_n.
```

To see that B has a non-negligible advantage over guessing, consider two cases.

If $X = \mathbb{QR}_n$, then $x \notin \mathcal{B}$, and so A_2 has no advantage over guessing. Consequently,

$$\Pr\left[\,\mathsf{B}(x) = \mathbb{QR}_n \mid X = \mathbb{QR}_n\,\right] = \Pr\left[\,\mathsf{A}_2(x, C) \neq b \mid x \xleftarrow{\mathsf{R}} \mathbb{QR}_n\,\right] = \frac{1}{2}.$$

If however $X = \mathbb{J}_n \setminus \mathbb{QR}_n$, then $x \in \mathcal{B}$, and so A_2 has an advantage over guessing. Since C is an encryption of m_b,

$$\Pr\left[\,\mathsf{B}(x) = \mathbb{J}_n \setminus \mathbb{QR}_n \mid X = \mathbb{J}_n \setminus \mathbb{QR}_n\,\right] = \Pr\left[\,\mathsf{A}_2(x, C) = b \mid x \xleftarrow{\mathsf{R}} \mathbb{J}_n \setminus \mathbb{QR}_n\,\right] = \frac{1}{2} + \varepsilon.$$

Therefore, suppressing the arguments of A_2, the overall probability of A_2 being correct is

$$\begin{aligned}
\Pr\left[\,\mathsf{B}(x) = X | X \xleftarrow{\mathsf{R}} \{\mathbb{QR}_n\,,\, \mathbb{J}_n \setminus \mathbb{QR}_n\}\,\right] &= \Pr\left[\,\mathsf{B}(x) = \mathbb{QR}_n \mid X = \mathbb{QR}_n\,\right] \cdot \Pr\left[\,X = \mathbb{QR}_n\,\right] \\
&\quad + \Pr\left[\,\mathsf{B}(x) = \mathbb{J}_n \setminus \mathbb{QR}_n \mid X = \mathbb{J}_n \setminus \mathbb{QR}_n\,\right] \cdot \Pr\left[\,X = \mathbb{J}_n \setminus \mathbb{QR}_n\,\right] \\
&= \Pr\left[\,\mathsf{A}_2 \neq b\,\right] \cdot \tfrac{1}{2} + \Pr\left[\,\mathsf{A}_2 = b\,\right] \cdot \tfrac{1}{2} \\
&= \tfrac{1}{2} \cdot \tfrac{1}{2} + \left(\tfrac{1}{2} + \varepsilon\right) \cdot \tfrac{1}{2} \\
&= \tfrac{1}{2} + \tfrac{\varepsilon}{2},
\end{aligned}$$

which indeed is non-negligible. □

Remark 2.3. Usually, the hardness assumption assumed for proving that the encryption is one-way is harder than the hardness assumption used for proving semantic security. However, in the case of Goldwasser–Micali, we see that the proof above can be modified slightly to show that the intractability of the quadratic residuosity problem also implies that the encryption function is one-way.

3. Okamoto-Uchiyama

We begin by introducing the parameters of the cryptosystem. It uses modulus $n = p^2 q$, encryption base $\mathcal{B} = \left\{g \in \mathbb{Z}_n^* \mid \operatorname{ord}_{p^2}\ \left(g \bmod p^2\right)^{p-1} = p\right\}$, and message space $\mathcal{M} = \mathbb{Z}_{2^{k-1}}$.

To efficiently define decryption, consider the group $\Gamma = \{ap + 1 \mid a \in \mathbb{Z}_p\}$ and the group homomorphism $L \colon \Gamma \to \mathbb{Z}_p$ given by

$$L(ap + 1) = a.$$

Alice	Bob
Chooses $p, q \xleftarrow{R} \mathcal{P}_k$. Chooses $g \xleftarrow{R} \mathcal{B}$. Computes $n = p^2 q$. Publishes (n, g).[1]	
	Chooses $r \xleftarrow{R} \mathbb{Z}_n$. Chooses $m \in \mathcal{M}$. Computes and sends $C = g^{m+nr} \bmod n$.
Computes $L_1 = L(C^{p-1} \bmod p^2)$. Computes $L_2 = L(g^{p-1} \bmod p^2)$. Computes $m = L_1 \cdot \overline{L}_2 \bmod p$.	

Now we describe how $g \in \mathcal{B}$ can be generated. Given $g \in \mathbb{Z}_p^*$, it is clear that $g \in \mathcal{B}$ if $g^{p-1} \bmod p^2 \neq 1$. Next, we prove how often such a g can be chosen.

Lemma 3.1. *If $g \xleftarrow{R} \mathbb{Z}_n^*$, then $g \in \mathcal{B}$ with a probability $\frac{p-1}{p}$.*

Proof. Since $p^2 \mid n$, we have that $g \bmod p^2$ is uniformly distributed in $\mathbb{Z}_{p^2}^*$. Using the isomorphism $\mathbb{Z}_{p^2}^* \cong \mathbb{Z}_p \times \mathbb{Z}_{p-1}$, we see that there are $(p-1)^2$ elements $h \in \mathbb{Z}_{p^2}^*$ such that h^{p-1} has order p. □

To verify correctness of decryption, we first compute

$$\begin{aligned} C^{p-1} &\equiv (g^{m+nr})^{p-1} \\ &\equiv g^{m(p-1)} \cdot g^{p(p-1)(pqr)} \\ &\equiv g^{m(p-1)} \bmod p^2. \end{aligned}$$

[1] One may compute and publish $g^n \bmod n$ for computational efficiency

Then since L is a homomorphism, it follows that

$$\begin{aligned} L(C^{p-1} \bmod p^2) \cdot \overline{L(g^{p-1} \bmod p^2)} \bmod p &= L(g^{m(p-1)} \bmod p^2) \cdot \overline{(p-1) \cdot L(g \bmod p^2)} \bmod p \\ &= m(p-1) \cdot \overline{p-1} \bmod p \\ &= m \bmod p. \end{aligned}$$

We will show that inverting the encryption of this scheme is as hard as factoring n. It is clear that knowing the factorization will reveal m so we prove that inverting the encryption leads to an efficient factorization method. We formally state the factoring problem.

Definition 3.2. The *factoring problem* is the computational problem of, given n, computing p and q.

Theorem 3.3. *If the factoring problem is intractable, then the encryption function is one-way.*

To prove the hardness of inverting the encryption we verify that the distribution of ciphertext in the cryptosystem can be simulated by an algorithm. For the next lemma, let p' and q' be such that $\operatorname{ord}_{p^2} g = p'p$ and $\operatorname{ord}_q g = q'$ and so

$$\operatorname{ord}_n g = \operatorname{lcm}(\operatorname{ord}_{p^2} g, \operatorname{ord}_q g) = p \cdot \operatorname{lcm}(p', q').$$

Lemma 3.4. *Given*

$$\begin{aligned} (p, q) &\xleftarrow{\mathsf{R}} \mathcal{P}_k \times \mathcal{P}_k, \\ (g, m, r) &\xleftarrow{\mathsf{R}} \mathcal{B} \times \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_n, \quad \textit{and} \\ z &\xleftarrow{\mathsf{R}} \mathbb{Z}_n, \end{aligned}$$

the distributions of $g^{m+nr} \bmod n$ *and* $g^z \bmod n$ *are close. Formally, for all* $X \in \langle g \rangle$,

$$\frac{1}{2} \cdot \left| \Pr\left[g^{m+nr} \equiv X \bmod n \right] - \Pr\left[g^z \equiv X \bmod n \right] \right|$$

is indeed negligible with respect to k.

Proof. Let $l = \operatorname{lcm}(p', q')$ and $\rho \colon \mathbb{Z}_l \to \mathbb{R}$ be given by $\rho(y) = \Pr\left[x \equiv y \bmod l \mid x \xleftarrow{\mathsf{R}} \mathbb{Z}_n \right]$. Since $X \in \langle g \rangle$, there exist integers a and b such that $X = g^{a+nb}$.

Clearly, $g^z \equiv X \bmod n$ implies $z \equiv a + nb \bmod pl$ so consequently, $z \equiv a \bmod p$ and $z \equiv nb \bmod l$. Then $z \xleftarrow{\mathsf{R}} \mathbb{Z}_p$ and consequently,

$$\begin{aligned} \Pr\left[g^z \equiv X \bmod n \right] &= \Pr\left[z \equiv a \bmod p \right] \cdot \Pr\left[z \equiv b \bmod l \right] \\ &= \tfrac{1}{p} \cdot \Pr\left[x \equiv b \bmod l \mid x \xleftarrow{\mathsf{R}} \mathbb{Z}_n \right] \\ &= \tfrac{1}{p} \cdot \rho(b) \end{aligned}$$

Similarly, $g^{m+nr} \equiv X \bmod n$ implies $m + nr \equiv a + nb$ so $m \equiv a \bmod p$ and $m + nr \equiv b \bmod l$. For a fixed m, we have $m + nr \equiv b \bmod l$ whenever $r + \overline{n}(m - a) \equiv b \bmod l$. Since $r \xleftarrow{\mathsf{R}} \mathbb{Z}_n$,

consequently, $r + \overline{n}(m-a) \xleftarrow{R} \mathbb{Z}_n$. Hence,

$$\begin{aligned}
\Pr[\, g^{m+nr} \equiv X \bmod n \,] &= \Pr[\, m \equiv a \bmod p \,] \cdot \Pr[\, m + nr \equiv a + nb \bmod l \,] \\
&= \Pr[\, m \equiv a \bmod p \,] \cdot \Pr[\, r + \overline{n}(m-a) \equiv b \bmod l \,] \\
&= \tfrac{1}{2^{k-1}} \cdot \Pr\left[\, x \equiv b \bmod l \mid x \xleftarrow{R} \mathbb{Z}_n \,\right] \\
&= \tfrac{1}{2^{k-1}} \cdot \rho(b)
\end{aligned}$$

Therefore, the statistical distance,

$$\begin{aligned}
&\tfrac{1}{2} \cdot |\Pr[\, g^{m+nr} \equiv X \bmod n \,] - \Pr[\, g^z \equiv X \bmod n \,]| \\
&= \tfrac{\rho(b)}{2} \cdot \left| \tfrac{1}{p} - \tfrac{1}{2^{k-1}} \right| < \tfrac{1}{2} \cdot \left(\tfrac{1}{2^{k-1}} - \tfrac{1}{2^k} \right) \\
&= \tfrac{1}{2^{k+1}}
\end{aligned}$$

is negligible with respect to k. □

Proof of Theorem 3.3. Let A be an algorithm that inverts the encryption with a non-negligible advantage ε. Construct a factoring algorithm B that factors $n = p^2 q$:

```
Algorithm B(n):
    Choose g' ←R Z_n.
    Choose z ←R Z_n.
    Compute C' = g^z mod n.
    Compute m = A(n, g', C').
    Compute d = gcd(z − m, n).
    If √d ∈ Z, then d = √d.
    If d < 2^k, then return (d, n/d).
    Else, return (n/d, d^2/d).
```

By Lemma 3.1 and Lemma 3.4, we have that distributions of g', C' are similar to distributions of g, C chosen in the cryptosystem. Hence, A returns the correct m corresponding to C' with a non-negligible advantage.

Then $g^z \equiv g^m \bmod n$ so $g^z \equiv g^m \bmod p^2$. Since $\mathrm{ord}_{p^2}\, g = p$ this implies $m \equiv z \bmod p$ so $p \mid z - m$. Note that $m \xleftarrow{R} \mathbb{Z}_{2^{k-1}}$ and $z \xleftarrow{R} \mathbb{Z}_n$. So $z \equiv m \bmod p$ implies that $z = m$ which occurs with a negligible probability of $\frac{1}{n}$. Hence, $n \nmid z - m$ with an overwhelming probability. In this case, $\gcd(z-m, n)$ is a non-zero factor of n that p divides but n does not. Hence $\gcd(z-m, n)$ is one of $\{p, p^2, pq\}$, which implies that B will return the correct pair (p, q). This breaks the factoring assumption. □

To prove semantic security we introduce the p-subgroup assumption. The problem states that it is hard to distinguish $\mathbb{Z}_n^*$ from its Sylow p-subgroup[2]. We first introduce the necessary notions to prove semantic security. Note that the hardness assumption is stated differently here than in the paper. To state the precise definition of the p-subgroup problem, consider the group $\Omega = \{ypq + 1 \mid y \in \mathbb{Z}_p\}$ which is the p-subgroup of $\mathbb{Z}_n^*$.

[2] i.e., the unique p-subgroup of maximal order

Definition 3.5. Let $n = p^2q$ and $X \xleftarrow{R} \{\Omega, \mathbb{Z}_n^* \setminus \Omega\}$. The *p-subgroup problem* is the decisional problem of, given n and $x \xleftarrow{R} X$, deciding whether $X = \Omega$ or $X = \mathbb{Z}_n^* \setminus \Omega$.

Theorem 3.6. *If p-subgroup problem is intractable, then the Okamoto-Uchiyama cryptosystem is semantically secure.*

We first prove a relationship between Ω and $\mathcal{B}$. Observe that if $x \in \Gamma \setminus \{1\}$, then x has order p.

Lemma 3.7.

(1) If $x \in \Omega$, then $x^{x-1} \notin \mathcal{B}$.
(2) If $x \in \mathcal{B}$ and $x \not\equiv 1 \bmod p$, then $x^{x-1} \notin \mathcal{B}$.

Proof. Suppose $x = ypq + 1 \in \Omega$. Then $x^{(x-1)(p-1)} \equiv x^{ypq(p-1)} \equiv 1 \bmod p^2$. So $x^{(x-1)(p-1)}$ does not have order p in $\mathbb{Z}_{p^2}^*$. Hence $x^{x-1} \notin \mathcal{B}$.

Let $x \in \mathcal{B}$ and $x \not\equiv 1 \bmod p$ and assume $x^{x-1} \in \mathcal{B}$. Since $\mathrm{ord}_{p^2}\, x^{p-1} = p$, if $x^{(x-1)(p-1)} \equiv 1 \bmod p^2$, then $p \mid (x-1)$ which contradicts that $x \not\equiv 1 \bmod p$. □

Before the proof of semantic security, we compute the cardinality of $\{x \in \mathcal{B} \mid x \not\equiv 1 \bmod p\}$. There are $pq - 1$ elements $h \in \mathbb{Z}_n^*$ such that $h \equiv 1 \bmod p$. Since $\#\mathcal{B} = (p-1)^2q$ by Lemma 3.1, the set $\{x \in \mathcal{B} \mid x \not\equiv 1 \bmod p\}$ has $(p-1)^2q - (pq-1) = p^2q - 3pq + 2$ elements.

Proof of Theorem 3.6. Let $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ be a semantic security algorithm with a non-negligible advantage $\varepsilon > 0$. Suppose $X \xleftarrow{R} \{\Omega, \mathbb{Z}_n^* \setminus \Omega\}$ and we are given $x \xleftarrow{R} X$. Let a deciding algorithm B be constructed as:

```
Algorithm B(x)
    Call A1(Z_{2^{k-1}}) to get (m0, m1).
    Choose b ←R {0,1}.
    Compute g = x^{x-1} mod n
    Compute C = g^{m_b} mod n
    Compute y = A2(g, C).
    If y ≠ b, then return Ω.
    If y = b, then return Z_n^* \ Ω.
```

To see that B has a non-negligible advantage over guessing, consider two cases.

If $X = \Omega$, then by Lemma 3.7, we have $g \notin \mathcal{B}$, and so A_2 has no advantage over guessing. Consequently,

$$\Pr[\mathsf{B}(x) = \Omega \mid X = \Omega] = \Pr\left[\mathsf{A}_2(g, C) \neq b \mid x \xleftarrow{R} \Omega\right] = \frac{1}{2}.$$

If $X = \mathbb{Z}_n^* \setminus \Omega$, then $x \in \{x \in \mathcal{B} \mid x \not\equiv 1 \bmod p\}$ with an overwhelming probability of

$$\frac{p^2q - 3pq + 2}{p^2q - p} > \frac{p-3}{p}.$$

In this case, by Lemma 3.7, $g \in \mathcal{B}$ so A_2 has an advantage. Consequently,

$$\Pr[\mathsf{B}(x) = \mathbb{Z}_n \setminus \Omega \mid X = \mathbb{Z}_n \setminus \Omega] = \Pr\left[\mathsf{A}_2(g, C) = b \mid x \xleftarrow{R} \mathbb{Z}_n \setminus \Omega\right] = \frac{1}{2} + \varepsilon.$$

Therefore, suppressing the arguments of A_2, the overall probability of A_2 being correct is

$$\begin{aligned}
\Pr\left[\mathsf{B}(x) = X | X \xleftarrow{\mathsf{R}} \{\Omega\,,\, \mathbb{Z}_n^* \setminus \Omega\}\right] &= \Pr\left[\mathsf{B}(x) = \Omega \mid X = \Omega\right] \cdot \Pr\left[X = \Omega\right] \\
&\quad + \Pr\left[\mathsf{B}(x) = \mathbb{Z}_n^* \setminus \Omega \mid X = \mathbb{Z}_n^* \setminus \Omega\right] \cdot \Pr\left[X = \mathbb{Z}_n^* \setminus \Omega\right] \\
&= \Pr\left[\mathsf{A}_2 \neq b\right] \cdot \tfrac{1}{2} + \Pr\left[\mathsf{A}_2 = b\right] \cdot \Pr\left[x \in \mathcal{B} \text{ and } x \not\equiv 1 \bmod p \mid x \xleftarrow{\mathsf{R}} \mathbb{Z}_n^* \setminus \Omega\right] \cdot \tfrac{1}{2} \\
&> \tfrac{1}{2} \cdot \tfrac{1}{2} + \left(\tfrac{1}{2} + \varepsilon\right) \cdot \left(1 - \tfrac{3}{p}\right) \cdot \tfrac{1}{2} \\
&= \tfrac{1}{2} + \tfrac{\varepsilon}{2} - \tfrac{3+6\varepsilon}{4p},
\end{aligned}$$

which indeed is non-negligible. □

4. Paillier

To define the parameters of the Paillier cryptosystem, we must first define $\lambda\colon \mathbb{Z} \to \mathbb{Z}$ by $\lambda(a) = \min\{y \in \mathbb{Z} \mid \text{for all } x \in \mathbb{Z}_a^*\colon x^y \equiv 1 \bmod a\}$. We will denote $\lambda(n)$ by λ for clarity.

This scheme considers $n = pq$ for $p, q \in \mathcal{P}_k$ and uses n^2 as the modulus. It uses encryption base

$$\mathcal{B} = \left\{x \in \mathbb{Z}_{n^2}^* \mid \operatorname{ord}_{n^2} x = \alpha n \text{ for } \alpha = 1, 2, \ldots, \lambda\right\}.$$

Finally, the message space is $\mathcal{M} = \mathbb{Z}_n$.

Next, let us discuss the encryption function used in this cryptosystem. For a fixed $g \in \mathbb{Z}_{n^2}^*$, the encryption function $\mathsf{Enc}_g\colon \mathbb{Z}_n \times \mathbb{Z}_n^* \to \mathbb{Z}_{n^2}^*$ is given by

$$\mathsf{Enc}_g(m, r) = g^m r^n \bmod n^2$$

Lemma 4.1. *The encryption function Enc_g is bijective.*

Proof. Since the sets $\mathbb{Z}_n \times \mathbb{Z}_n^*$ and $\mathbb{Z}_{n^2}^*$ have equal cardinalities, it suffices to show that Enc_g is injective.

Let $(a, b), (c, d) \in \mathbb{Z}_n \times \mathbb{Z}_n^*$ such that $g^n b^n \equiv g^c d^n \bmod n^2$. Dividing both sides by $g^a b^n$, we have $g^{c-a}(d\bar{b})^n \equiv 1 \bmod n^2$. Since $\lambda(n^2) = n\lambda$, we have

$$g^{(c-a)\lambda}(d\bar{b})^{n\lambda} \equiv g^{(c-a)\lambda} \equiv 1 \bmod n^2.$$

Thus, $\operatorname{ord}_{n^2} g \mid \lambda(c-a)$ and so $n \mid \lambda(c-a)$. But p, q are primes of equal bit-length, so p, q do not divide $\lambda = (p-1)(q-1)$. Then $n \mid (c-a)$ which implies $a \equiv c \bmod n$.

Then from $(d\bar{b})^n \equiv 1 \bmod n^2$ so we have $(d\bar{b})^n \equiv 1 \bmod n$. Then $\operatorname{ord}_n(d\bar{b})$ divides n and $\lambda = (p-1)(q-1)$. Clearly, $\operatorname{ord}_n(d\bar{b}) = 1$, so $d\bar{b} \equiv 1 \bmod n$ which implies $b \equiv d \bmod n$. □

If we let $c = g^m r^n$, when $g \in \mathcal{B}$, we can define the function $\mathsf{Dec}\colon \mathbb{Z}_{n^2}^* \times \mathcal{B} \to \mathbb{Z}_n$ which inverts the encryption by $\mathsf{Dec}(c, g) = m$.

Similar to Okamoto-Uchiyama, to efficiently describe the encryption, we introduce $\mathcal{S} = \{yn + 1 \mid y \in \mathbb{Z}_n\}$ and the function $L\colon \mathcal{S} \to \mathbb{Z}_n$ given by

$$L(yn + 1) = y.$$

We prove a logarithmic property of this function which will be used in decryption.

Lemma 4.2. *For any $c \in \mathbb{Z}_{n^2}^*$, $L(c^\lambda \bmod n^2) = \lambda \mathsf{Dec}(c, 1+n) \bmod n$.*

Proof. Since $1+n \in \mathcal{B}$, there is a unique (a,b) in $\mathbb{Z}_n \times \mathbb{Z}_n^*$ such that $c = (1+n)^a b^n \bmod n^2$. Then $a = \mathsf{Dec}(c, 1+n)$ and

$$c^\lambda \equiv (1+n)^{a\lambda} b^{n\lambda} \equiv (1+n)^{a\lambda} = 1 + a\lambda n \bmod n^2.$$ □

Alice	**Bob**
Chooses $p, q \xleftarrow{\mathsf{R}} \mathcal{P}_k$. Chooses $g \xleftarrow{\mathsf{R}} \mathcal{B}$. Computes $n = pq$. Publishes (n, g).	
	Chooses $r \xleftarrow{\mathsf{R}} \mathbb{Z}_n^*$. Chooses $m \in \mathcal{M}$. Computes and sends $C = g^m r^n \bmod n^2$.
Computes $L_1 = L(C^\lambda \bmod n^2)$. Computes $L_2 = L(g^\lambda \bmod n^2)$. Computes $m = L_1 \cdot \overline{L}_2 \bmod n$.	

First, we show how one can efficiently check if an element of $\mathbb{Z}_{n^2}$ is $\mathcal{B}$.

Lemma 4.3. *Let $x \in \mathbb{Z}_{n^2}$. If $\gcd\left(L\left(x^\lambda \bmod n^2\right), n\right) = 1$, then $x \in \mathcal{B}$.*

Proof. Since $n+1 \in \mathcal{B}$, there exists $(a,b) \in \mathbb{Z}_n \times \mathbb{Z}_n^*$ such that $x \equiv (n+1)^a b^n \bmod n^2$. Then denoting $\mathrm{ord}_{n^2}\, x$ by d,

$$x^{\lambda d} \equiv (n+1)^{\lambda a d} y^{\lambda n^2 d} \equiv (n+1)^{\lambda a d} \equiv 1 \bmod n^2.$$

Hence, the order of $n+1$ must divide $\lambda a d$. Recall that by Lemma 4.2, $L\left(x^\lambda \bmod n^2\right) = \lambda a$. Thus n does not divide λa and consequently n divides d. □

Now, we prove that most elements of $\mathbb{Z}_n^2$ belong to $\mathcal{B}$.

Lemma 4.4. *If $g \xleftarrow{\mathsf{R}} \mathbb{Z}_{n^2}$, then $g \in \mathcal{B}$ with an overwhelming probability*

$$\Pr\left[\, g \in \mathcal{B} \mid g \xleftarrow{\mathsf{R}} \mathbb{Z}_{n^2} \,\right] = \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right).$$

Proof. Since $p^2 \mid n$, we have that $g \bmod p^2$ is uniformly distributed in $\mathbb{Z}^*_{p^2}$. Using the isomorphism $\mathbb{Z}^*_{p^2} \cong \mathbb{Z}_p \times \mathbb{Z}_{(p-1)}$, we see that there are $(p-1)^2$ elements with order that is a nonzero multiple p. Similarly, there are $(q-1)^2$ elements in $\mathbb{Z}^*_{q^2}$ with order that is a nonzero multiple q. □

To verify the correctness of decryption, note that by the definition of Dec, for all $g_1, g_2 \in \mathcal{B}$, we have

$$\mathsf{Dec}(C, g_1) \equiv \mathsf{Dec}(C, g_2) \cdot \mathsf{Dec}(g_2, g_1) \bmod n.$$

Then correctness of decryption follows from the logarithmic property of L

$$\begin{aligned} L(C^\lambda \bmod n^2) \cdot \overline{L(g^\lambda \bmod n^2)} \bmod n &= \lambda \mathsf{Dec}(C, 1+n) \cdot \overline{\lambda \mathsf{Dec}(g, 1+n)} \bmod n \\ &= \mathsf{Dec}(C, 1+n) \cdot \overline{\mathsf{Dec}(g, 1+n)} \bmod n \\ &= \mathsf{Dec}(C, g) \bmod n \\ &= m \bmod n. \end{aligned}$$

Next, we describe the computational and decisional problems we assume to be hard, that make this cryptosystem secure. To prove that the function is one-way, we introduce n^{th} residuosity which is similar to quadratic residuosity but involves higher exponents.

Definition 4.5. An integer z is an n^{th} *residue modulo* n^2 if there exists $y \in \mathbb{Z}^*_{n^2}$ such that

$$z = y^n \bmod n^2.$$

We denote the set of n^{th} residues mod n^2 as $\mathbb{CR}_{n^2}$. Next, we describe the decisional problem associated with composite residuosity.

Definition 4.6. Let $X \xleftarrow{\mathsf{R}} \{\mathbb{CR}_{n^2}, \mathbb{Z}_n \setminus \mathbb{CR}_{n^2}\}$. The *composite residuosity problem* is the decisional problem of, given $x \xleftarrow{\mathsf{R}} X$, deciding whether $X = \mathbb{CR}_{n^2}$ or $X = \mathbb{Z}_n \setminus \mathbb{CR}_{n^2}$.

We also formally state the problem of verifying the correctness of inverting the encryption.

Definition 4.7. Let $(c, g) \in \mathbb{Z}^*_{n^2} \times \mathcal{B}$. The *decisional inverting problem* is the decisional problem of, given $x \in \mathbb{Z}_n$ and (c, g), deciding whether $x = \mathsf{Dec}(c, g)$.

The next lemma establishes a relationship between composite residuosity and the inverting function.

Lemma 4.8. *For all $c \in \mathbb{Z}^*_{n^2}$, and $g \in \mathcal{B}$, $\mathsf{Dec}(c, g) = 0$ if and only if c is an n^{th} residue modulo n^2.*

Proof. We have that:

$\mathsf{Dec}(c, g) = 0$ iff

$\mathsf{Enc}(0, g) \equiv c \bmod n^2$ iff

there exists $y \in \mathbb{Z}^*_n \subset \mathbb{Z}^*_{n^2}$ such that $c = g^0 y^n$ iff

c is an n^{th} residue mod n^2.

□

Theorem 4.9. *If the composite residuosity problem is intractable, then the encryption function is one-way.*

We will prove that if composite residuosity is intractable, then the decisional inverting problem is intractable. Since the decisional inverting problem is verifying the solution to inverting the function, hardness in inverting will follow from the general fact that computing a solution is at least as hard than verifying it.

Proof. Suppose we have an algorithm $\mathsf{A}(x, c, g)$ that solves the decisional inverting problem with a non-negligible advantage over guessing. Let $x \xleftarrow{\mathsf{R}} \mathbb{Z}_n$. Consider an algorithm $\mathsf{B}(x)$ which returns $\mathsf{A}(0, x, 1+n)$. Then, by Lemma 4.8, x is an n^{th} residue if and only if $\mathsf{Dec}(x, 1+n) = 0$. □

Finally, we show that this cryptosystem is semantically secure relative to the composite residuosity. To do so, we establish a relationship between $\mathbb{CR}_{n^2}$ and $\mathcal{B}$.

Theorem 4.10.

(1) If $x \in \mathbb{CR}_{n^2}$, then $x^\lambda \notin \mathcal{B}$.
(2) If $x \in \mathcal{B} \setminus \mathbb{CR}_{n^2}$, then $x^\lambda \in \mathcal{B}$.

Proof.

(1) If $x \in \mathbb{CR}_{n^2}$, then there exists $y \in \mathbb{Z}^*_{n^s}$ such that $x \equiv y^n \bmod n^2$. Then $x^\lambda \equiv y^{\lambda n} \equiv 1$ which clearly has order 1 in $\mathbb{Z}_{n^2}$. Thus $x^\lambda \notin \mathcal{B}$.
(2) Let $x \in \mathcal{B} \setminus \mathbb{CR}_{n^2}$ and suppose there exists t such that $x^{\lambda t} \equiv 1 \bmod n^2$. From $x \in \mathcal{B}$ we have $a \mid \lambda t$ and consequently $n \mid t$ which shows that x^λ has order n. Thus $x^\lambda \in \mathcal{B}$.

□

Theorem 4.11. *If the composite residuosity problem is intractable then the cryptosystem is semantically secure.*

Proof. Let $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ be a semantic security algorithm with a non-negligible advantage $\varepsilon > 0$. Suppose $X \xleftarrow{\mathsf{R}} \{\mathbb{CR}_{n^2}, \mathbb{Z}_{n^2} \setminus \mathbb{CR}_{n^2}\}$ and we are given $x \xleftarrow{\mathsf{R}} X$. Let a deciding algorithm B be constructed as:

```
Algorithm B(x):
    Call A1(Z_n) to get (m0, m1).
    Choose b <-R {0,1}.
    Compute g = x^λ mod n^2
    Compute C = g^{m_b} mod n^2
    Compute y = A2(g, C).
    If y ≠ b, then return CR_{n^2}.
    If y = b, then return Z*_{n^2} \ CR_{n^2}.
```

To see that B has a non-negligible advantage over guessing, consider two cases.

If $X = \mathbb{CR}_{n^2}$, then by Theorem 4.10, we have $g \notin \mathcal{B}$, and so A_2 has no advantage over guessing. Consequently,

$$\Pr\left[\mathsf{B}(x) = \mathbb{CR}_{n^2} \mid X = \mathbb{CR}_{n^2}\right] = \Pr\left[\mathsf{A}_2(g, C) \neq b \mid x \xleftarrow{\mathsf{R}} \mathbb{CR}_{n^2}\right] = \frac{1}{2}.$$

If $X = \mathbb{Z}^*_{n^2} \backslash \mathbb{CR}_{n^2}$, then by Lemma 4.4 $x \in \mathcal{B}$ with an overwhelming probability $(1-\frac{1}{p})(1-\frac{1}{q}) > 1-\frac{4}{p}$. In this case, by Theorem 4.10, $g \in \mathcal{B}$ so A_2 has an advantage. Consequently,

$$\Pr\left[\,\mathsf{B}(x) = \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2} \mid X = \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2}\,\right] = \Pr\left[\,\mathsf{A}_2(g, C) = b \mid x \xleftarrow{\mathsf{R}} \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2}\,\right] = \frac{1}{2} + \varepsilon.$$

Therefore, suppressing the arguments of A_2, the overall probability of A_2 being correct is

$$\begin{aligned}
\Pr\left[\,\mathsf{B}(x) = X | X \xleftarrow{\mathsf{R}} \left\{\mathbb{CR}_{n^2}\,,\, \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2}\right\}\right] &= \Pr\left[\,\mathsf{B}(x) = \mathbb{CR}_{n^2} \mid X = \mathbb{CR}_{n^2}\,\right] \cdot \Pr\left[\,X = \mathbb{CR}_{n^2}\,\right] \\
&\quad + \Pr\left[\,\mathsf{B}(x) = \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2} \mid X = \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2}\,\right] \cdot \Pr\left[\,X = \mathbb{Z}^*_{n^2} \setminus \mathbb{CR}_{n^2}\,\right] \\
&= \Pr\left[\,\mathsf{A}_2 \neq b\,\right] \cdot \tfrac{1}{2} + \Pr\left[\,\mathsf{A}_2 = b\,\right] \cdot \Pr\left[\,x \in \mathcal{B}\,\right] \cdot \tfrac{1}{2} \\
&> \tfrac{1}{2} \cdot \tfrac{1}{2} + \left(\tfrac{1}{2} + \varepsilon\right) \cdot \left(1 - \tfrac{4}{p}\right) \cdot \tfrac{1}{2} \\
&= \tfrac{1}{2} + \tfrac{\varepsilon}{2} - \tfrac{2\varepsilon+1}{p},
\end{aligned}$$

which indeed is non-negligible. □

5. Groth

The Groth cryptosystem uses the modulus of $n = pq$ for primes p, q are of the form $p = 2p'r_p + 1$ where p' is a prime and r_p is a product of distinct odd prime numbers less than a fixed integer B. It is necessary that B is small (less than 15 bits) such that discrete logarithms in a group of B smooth order can be computed efficiently.

Primes p and q are generated by fixing a number B (usually 2^{15}), choosing odd primes $p_1, ..., p_s$ smaller than B, choosing a prime p' that is larger than B, and then testing the expression $2p'(p_1...p_s) + 1$ for primality.

In this scheme, the encryption base $\mathcal{B}$ is the set $\{x \in \mathbb{QR}_n \mid \mathrm{ord}_n\, x = p'q'r_g$ where r_g divides $r_p r_q\}$. Encryption also requires the unique subgroup of $\mathbb{Z}^*_n$ with order $p'q'$ which we denote by $\mathbb{H}$. Finally, we specify the message space $\mathcal{M} = \mathbb{Z}_{2^{k-1}}$.

Alice	**Bob**
Chooses $p, q \in \mathcal{P}_k$ as described above. Chooses $g \in \mathcal{B}$. Computes $n = pq$. Publishes (n, g).	
	Chooses $r \xleftarrow{\mathsf{R}} \mathbb{Z}_{2^{2k-1}}$. Computes a generator h of $\mathbb{H}$. Chooses $m \in \mathcal{M}$. Publishes $C = g^m h^r \bmod n$.
Computes $m = \log_{g^{p'q'}} C^{p'q'} \bmod n$	

To describe how g, h are chosen, consider the set $\mathcal{B}' = \{x \in \mathbb{Z}_n^* \mid p'q' \text{ divides } \operatorname{ord}_n x\}$. We see that given an element $x \in \mathbb{Z}_n^* \setminus \{1\}$, if $x^{2r_p r_q p'} \not\equiv 1$ and $x^{2r_p r_q q'} \not\equiv 1$ then $x \in \mathcal{B}'$. Next we describe how frequently $x \in \mathcal{B}$ can be selected.

Lemma 5.1. *If $x \xleftarrow{\mathsf{R}} \mathbb{Z}_n^*$, then $x \in \mathcal{B}'$ with an overwhelming probability of*

$$\left(1 - \frac{1}{p'}\right)\left(1 - \frac{1}{q'}\right).$$

Proof. From the isomorphism $\mathbb{Z}_p^* \cong \mathbb{Z}_{2r_p} \times \mathbb{Z}_{p'}$ we see that there are $2r_p(p'-1)$ elements in $\mathbb{Z}_p^*$ with order that is a multiple of p'. Similarly, there are $2r_q(q'-1)$ elements in $\mathbb{Z}_q^*$ with order that is a multiple of q'. □

Once an element x in $\mathcal{B}'$ has been chosen, $x^2 \in \mathcal{B}$ and x^P generates $\mathbb{H}$ where P is the product of all prime numbers less than B.

We verify the correctness of the decryption:

$$\begin{aligned} C^{p'q'} &\equiv (g^m h^r)^{p'q'} \\ &\equiv (g^{mp'q'})(h^{p'q'})^r \\ &\equiv (g^{p'q'})^m \qquad \bmod n. \end{aligned}$$

Note that $\operatorname{ord}_n g^{p'q'} = r_g$ where $r_g \mid r_p r_q$, so its order $\bmod n$ is B-smooth. Then the discrete logarithm problem $C^{p'q'} = (g^{p'q'})^m$ can be solved quickly using the Pohlig-Hellman algorithm for instance, and a unique solution modulo r_g may be returned for sufficiently small values of B.

To verify semantic security, we first establish a relationship between $\mathbb{QR}_n$ and $\mathbb{H}$.

Lemma 5.2. $\mathbb{H} \subseteq \mathbb{QR}_n$.

Proof. Let $x \in \mathbb{H}$. If $x = 1$, then $x \in \mathbb{QR}_n$ trivially. Otherwise, $\operatorname{ord}_n x \mid p'q'$, so $\operatorname{ord}_n x$ is odd. Thus, $x \equiv 1 \cdot x \equiv x^{\operatorname{ord}_n x + 1} \equiv (x^{\frac{\operatorname{ord}_n x + 1}{2}})^2 \pmod{n}$, so $x \in \mathbb{QR}_n$. □

The semantic security follows from the fact that distinguishing elements of $\mathbb{H}$ from elements of $\mathbb{QR}_n$ is difficult.

Definition 5.3. Let $X \xleftarrow{\mathsf{R}} \{\mathbb{QR}_n, \mathbb{QR}_n \setminus \mathbb{H}\}$. The *decisional strong RSA problem* is the decisional problem of, given $x \xleftarrow{\mathsf{R}} X$, deciding whether $X = \mathbb{QR}_n$ or $X = \mathbb{QR}_n \setminus \mathbb{H}$.

Theorem 5.4. *If the decisional strong RSA problem is intractable, then the cryptosystem is semantically secure.*

Proof. Let $\mathsf{A} = (\mathsf{A_1}, \mathsf{A_2})$ be a semantic security algorithm with a non-negligible advantage $\varepsilon > 0$. Suppose $X \xleftarrow{\mathsf{R}} \{\mathbb{H}, \mathbb{QR}_n \setminus \mathbb{H}\}$ and we are given $x \xleftarrow{\mathsf{R}} X$. Let a deciding algorithm B be constructed as:

```
Algorithm B(x):
    Call A_1(Z_{2^{k-1}}) to get (m_0, m_1).
    Choose b ←R {0,1}.
    Choose y ←R Z_n^* and compute g = y^2 mod n.
    Compute C = g^{m_b} x mod n^2.
    Compute y = A_2(x, C).
    If y ≠ b, then return H.
    If y = b, then return QR_n \ H.
```

For $\mathsf{A_2}$ to have an advantage, we need that $g \in \mathcal{B}$ which occurs with $(1 - \frac{1}{p'})(1 - \frac{1}{q'}) > 1 - \frac{4}{p'}$ probability. To see that B has a non-negligible advantage over guessing, consider two cases.

If $x \in \mathbb{H}$, the encryption function is not injective and so $\mathsf{A_2}$ has no advantage over guessing. Consequently,

$$\Pr[\,\mathsf{B}(x) = \mathbb{H} \mid X = \mathbb{H}\,] = \Pr\left[\,\mathsf{A_2}(x, C) \neq b \mid x \xleftarrow{\mathsf{R}} \mathbb{H}\,\right] = \frac{1}{2}.$$

If $x \notin \mathbb{H}$, then $\mathsf{A_2}$ retains its advantage. Consequently,

$$\Pr[\,\mathsf{B}(x) = \mathbb{QR}_n \setminus \mathbb{H} \mid X = \mathbb{QR}_n \setminus \mathbb{H}\,] = \Pr\left[\,\mathsf{A_2}(x, C) = b \mid x \xleftarrow{\mathsf{R}} \mathbb{QR}_n \setminus \mathbb{H}\,\right] = \frac{1}{2} + \varepsilon.$$

Therefore, suppressing the arguments of $\mathsf{A_2}$, the overall probability of $\mathsf{A_2}$ being correct is

$$\begin{aligned}
\Pr\left[\,\mathsf{B}(x) = X \mid X \xleftarrow{\mathsf{R}} \{\mathbb{H}, \mathbb{QR}_n \setminus \mathbb{H}\}\,\right] &= \Pr[\,g \in \mathcal{B}\,] \cdot \big(\Pr[\,\mathsf{B}(x) = \mathbb{H} \mid X = \mathbb{H}\,] \cdot \Pr[\,X = \mathbb{H}\,] \\
&\qquad + \Pr[\,\mathsf{B}(x) = \mathbb{QR}_n \setminus \mathbb{H} \mid X = \mathbb{QR}_n \setminus \mathbb{H}\,] \cdot \Pr[\,X = \mathbb{QR}_n \setminus \mathbb{H}\,]\big) \\
&> \left(1 - \tfrac{4}{p'}\right) \cdot \big(\Pr[\,\mathsf{A_2} \neq b\,] \cdot \tfrac{1}{2} + \Pr[\,\mathsf{A_2} = b\,] \cdot \tfrac{1}{2}\big) \\
&= \tfrac{1}{2} + \tfrac{\varepsilon}{2} - \tfrac{2\varepsilon + e}{p'},
\end{aligned}$$

which indeed is non-negligible. □

References

[GM82] Shafi Goldwasser and Silvio Micali, *Probabilistic encryption & how to play mental poker keeping secret all partial information*, Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '82, ACM, 1982, pp. 365–377, doi:10.1145/800070.802212, http://doi.acm.org/10.1145/800070.802212.

[Gro05] Jens Groth, *Cryptography in subgroups of* $\mathbb{Z}_n^*$, Theory of cryptography, Lecture Notes in Comput. Sci., vol. 3378, Springer, Berlin, 2005, pp. 50–65, doi:10.1007/978-3-540-30576-7_4, https://doi.org/10.1007/978-3-540-30576-7_4.

[HPS14] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman, *An introduction to mathematical cryptography*, second ed., Undergraduate Texts in Mathematics, Springer, New York, 2014, doi:10.1007/978-1-4939-1711-2, https://doi.org/10.1007/978-1-4939-1711-2.

[OU98] Tatsuaki Okamoto and Shigenori Uchiyama, *A new public-key cryptosystem as secure as factoring*, Advances in cryptology—EUROCRYPT '98 (Espoo), Lecture Notes in Comput. Sci., vol. 1403, Springer, Berlin, 1998, pp. 308–318, doi:10.1007/BFb0054135, https://doi.org/10.1007/BFb0054135.

[Pai99] Pascal Paillier, *Public-key cryptosystems based on composite degree residuosity classes*, Advances in cryptology—EUROCRYPT '99 (Prague), Lecture Notes in Comput. Sci., vol. 1592, Springer, Berlin, 1999, pp. 223–238, doi:10.1007/3-540-48910-X_16, https://doi.org/10.1007/3-540-48910-X_16.