Published in P. Y. A. Ryan, D. Naccache, and J.-J. Quisquater, Eds., *The New Codebreakers (Kahn Festschrift)*, vol. 9100 of *Lecture Notes in Computer Science*, pp. 470-478, Springer, 2016.

Secure ElGamal-type Cryptosystems Without Message Encoding

Marc Joye

Technicolor
175 S. San Antonio Road, Los Altos, CA 94022, USA
marc.joye@technicolor.com

Abstract. ElGamal cryptosystem is one of the oldest public-key cryptosystems. It is known to be semantically secure for arbitrary messages in the random oracle model under the decisional Diffie-Hellman assumption. Semantic security also holds in the standard model when messages are encoded as elements in the group for which the decisional Diffie-Hellman assumption is defined. This paper introduces a setting and companion cryptosystem where semantic security can be proved in the standard model *without* message encoding. Extensions achieving security against chosen-ciphertext attacks are also provided.

1 Introduction

The classical ElGamal cryptosystem [7,8] is closely related to the Diffie-Hellman key exchange protocol. It was introduced by Taher ElGamal in 1984. In retrospect, it is somewhat surprising that this scheme was not discovered before the RSA cryptosystem.

The original scheme goes as follows. Let p be a large prime and let g be a primitive element modulo p, that is, g generates all of $\mathbb{F}_p^*$. The public key is $\mathsf{pk} = \{g, y, p\}$ where $y = g^x \bmod p$ for a random integer r with $0 \leq x < p$. The secret key is $\mathsf{sk} = \{x\}$. The encryption of a message m, with $1 \leq m < p$, is given by the pair (c_1, c_2) where $c_1 = g^r \bmod p$ and $c_2 = m\, y^r \bmod p$ for a random integer $0 \leq r < p$. Message m is then recovered using secret key x as $m = c_2 \cdot {c_1}^{-x} \bmod p$.

As already pointed out in [7], breaking the above scheme and the Diffie-Hellman key exchange protocol are equally difficult. Using modern terminology, ElGamal showed that his cryptosystem is *one-way* against chosen-plaintext attacks (OW-CPA) under the *computational Diffie-Hellman (*CDH*) assumption* . Informally, the CDH assumption says that given two random elements $A = g^a$ and $B = g^b$ in $\mathbb{F}_p^*$ the value of $g^{ab} \in \mathbb{F}_p^*$ cannot be recovered. It is now easy to see that an attacker against the one-wayness of the ElGamal cryptosystem can be used to solve a CDH problem in $\mathbb{F}_p^*$; namely, obtaining $g^{ab} \pmod p$ from (p, g, g^a, g^b). One gives the attacker the public key $\mathsf{pk} = \{g, g^a, p\}$ and the challenge ciphertext $c = (g^b, c_2)$ for some random value $c_2 \in \mathbb{F}_p^*$. The attacker answer with the corresponding message $m = c_2 \cdot (g^b)^{-a} \pmod p$, which yields the value of g^{ab} as $c_2/m \pmod p$. The other direction is immediate.

Semantic security [10] captures a stronger notion of data privacy. Basically, it requires that an adversary should not learn anything about a plaintext given its encryption beyond the length of the plaintext. An equivalent notion is that of indistinguishable encryptions. ElGamal cryptosystem is known to meet the IND-CPA security level (i.e., *indistinguishability against chosen-plaintext attacks*) under the *decisional Diffie-Hellman (*DDH*) assumption* [15]. Formal definitions are given in the next section.

Unfortunately, the DDH assumption does not hold in $\mathbb{F}_p^*$. If $\mathbb{F}_p^* = \langle g \rangle$ (that is, g is a generator of $\mathbb{F}_p^*$) then it follows that $\left(\frac{g}{p}\right) = -1$, where $\left(\frac{g}{p}\right)$ denotes the Legendre symbol of g modulo p. Hence, for any $0 \leq x < p-1$, $\left(\frac{g^x}{p}\right) = (-1)^x$ leaks the value of x mod 2. A more sophisticated attack against the original ElGamal cryptosystem is presented in [3, Section 4].

These issues can easily be circumvented by working in a (large) prime-order subgroup $\mathbb{G}$ of $\mathbb{F}_p^*$. The ElGamal encryption proceeds exactly in the same way except that g is now an element of order q where q is a large prime factor of $p-1$. Messages being encrypted must also be elements of the subgroup generated by g; i.e., the message space is $\mathcal{M} = \mathbb{G}$ with $\mathbb{G} = \langle g \rangle$. Representing messages as group elements is known as *message encoding*.

In [15], Tsiounis and Yung propose the following message encoding when $\mathbb{G}$ is the subgroup of quadratic residues modulo p and $(p-1)/2$ is prime. Let ε be a security parameter (typically $\varepsilon = 64$). A κ-bit message m (with $\kappa = |p|_2 - \varepsilon$) is encoded as $\widehat{m} = \rho\, 2^\kappa + m$ for $\rho = 0, 1, 2, \ldots$ until $\widehat{m}$ is a quadratic residue so that $\widehat{m} \in \mathbb{G}$. Note that message m can be obtained from $\widehat{m}$ as $m = \widehat{m} \bmod 2^\kappa$. There are several drawbacks in this approach. First, it requires computing Legendre symbols to test the quadratic residuosity. Second, the potential message space is not fully exploited as several bits are provisioned to store ρ. Third, as presented, the message encoding is not time-constant and so may reveal information on m (timing attack).

An alternative message encoding for the previous setting is mentioned in [6, § 5.1]. The messages being encrypted are restricted to be elements in the set $\{1, \ldots, q\}$ with $q = (p-1)/2$ prime. A message m is then encoded by squaring it modulo p, yielding $\widehat{m} = m^2 \bmod p$ in $\mathbb{G}$. For the decryption, message m can be recovered from its encoding $\widehat{m}$ by computing its unique square root in the set $\{1, \ldots, q\}$ —observe that since q is an [odd] prime, it follows that $p \equiv 3 \pmod 4$. On the down side, the total decryption time to get plaintext (decoded) message m is longer.

In [5], Chevallier-Mames *et al.* suggest a modification to the classical ElGamal cryptosystem so as to avoid message encoding. The semantic security of the resulting cryptosystem relies on a new, specifically introduced assumption; namely, the *decisional class Diffie-Hellman assumption*. However, its connection with the standard DDH assumption is unclear and was left as an open problem in [5].

Another way to avoid message encoding is to invoke the random oracle model [1] when proving the security. The second part of an ElGamal ciphertext is modified as $c_2 = m \oplus \mathcal{H}(y^r \bmod p)$, where $\mathcal{H} : \mathbb{G} \to \mathcal{M}$ and $\mathcal{M} = \{0,1\}^\kappa$. The

random oracle model assumes that the hash function $\mathcal{H}$ behaves as a random function. While the resulting cryptosystem can be shown to achieve semantic security, the security proof only stands in the idealized random-oracle model. In particular, there are no guarantees that the proof holds in the standard model when function $\mathcal{H}$ is concretely instantiated, as demonstrated by Canetti *et al.* [4].

This paper presents a variation of the ElGamal cryptosystem meeting the IND-CPA security notion without message encoding, nor random oracles. Extensions to deal with stronger scenario attacks are also presented. Unlike [15] the message space is optimal and unlike [6] the decryption time is roughly the same as for the original ElGamal cryptosystem. Further, the ciphertext components are one bit shorter, which can be useful for super-encryption. The security of all presented schemes relies on a standard DDH assumption.

2 Background

In this section, we review well-known definitions and notions for public-key encryption. We also introduce some useful notation.

Public-key encryption A *public-key encryption scheme* [10] is a tuple of three algorithms $(\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$:

Key generation The key generation algorithm KeyGen is a randomized algorithm that takes as input some security parameter 1^λ and returns a matching pair of public key and secret key for some user: $(\mathsf{pk}, \mathsf{sk}) \xleftarrow{R} \mathsf{KeyGen}(1^\lambda)$. The message space is denoted by $\mathcal{M}$.

Encryption The encryption algorithm Enc is a randomized algorithm that takes as input a public key pk and a plaintext $m \in \mathcal{M}$, and returns a ciphertext c. We write $c \leftarrow \mathsf{Enc}_{\mathsf{pk}}(m)$.

Decryption The decryption algorithm Dec takes as input secret key sk (matching pk) and ciphertext c and returns the corresponding plaintext m or a special symbol $\perp$ indicating that the ciphertext is invalid. We write $m \leftarrow \mathsf{Dec}_{\mathsf{sk}}(c)$ if c is a valid ciphertext and $\perp \leftarrow \mathsf{Dec}_{\mathsf{sk}}(c)$ if it is not.

We require that $\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{pk}}(m)) = m$ for any message $m \in \mathcal{M}$, where $(\mathsf{pk}, \mathsf{sk}) \xleftarrow{R} \mathsf{KeyGen}(1^\lambda)$.

Complexity assumptions Let $\mathbb{G} = \langle g \rangle$ denote a (multiplicatively written) cyclic group of order q. Given $g^a, g^b \xleftarrow{R} \mathbb{G}$, the *computational* Diffie-Hellman (CDH) problem is to compute g^{ab}. Likewise, the *decisional* Diffie-Hellman (DDH) problem is to distinguish between the two distributions (g, g^a, g^b, g^{ab}) and (g, g^a, g^b, g^c) for $a, b, c \xleftarrow{R} \mathbb{Z}/q\mathbb{Z}$.

More formally, the DDH assumption is defined as follows.

Definition 1. *The* DDH assumption in $\mathbb{G}$ *requires that for any probabilistic polynomial-time adversary* $\mathcal{A}$ *the advantage*

$$\left|\Pr\left[\mathcal{A}(\mathbb{G}, q, g, g^a, g^b, g^{ab}) = 1\right] - \Pr\left[\mathcal{A}(\mathbb{G}, q, g, g^a, g^b, g^c) = 1\right]\right|$$

is negligible in the security parameter λ*, where the probabilities are taken over the experiment of generating a group* $\mathbb{G} = \langle g \rangle$ *of order* q *on input* 1^λ *and choosing* $a, b, c \xleftarrow{R} \mathbb{Z}/q\mathbb{Z}$*.*

Examples of groups $\mathbb{G}$ for which the DDH problem is hard include a prime-order subgroup of $\mathbb{F}_p^*$ or a prime-order subgroup of the points of an elliptic curve over a finite field. An excellent survey on the DDH problem is provided in [2].

Security notions In order to properly define the notion of *indistinguishability of encryptions*, we view an adversary $\mathcal{A}$ as a pair $(\mathcal{A}_1, \mathcal{A}_2)$ of probabilistic algorithms. This corresponds to adversary $\mathcal{A}$ running in two stages. In the "find" stage, algorithm $\mathcal{A}_1$ takes as input a public key pk and outputs two equal-size messages m_0 and $m_1 \in \mathcal{M}$ and some state information s. In the "guess" stage, algorithm $\mathcal{A}_2$ receives a challenge ciphertext c which is the encryption of m_b under pk and where b is chosen at random in $\{0, 1\}$. The goal of $\mathcal{A}_2$ is to recover the value of b from s and c.

A public-key encryption scheme is said *indistinguishable* (or *semantically secure*) if

$$\Pr\left[\begin{array}{l}(\mathsf{pk}, \mathsf{sk}) \xleftarrow{R} \mathsf{KeyGen}(1^\lambda), (m_0, m_1, s) \leftarrow \mathcal{A}_1(\mathsf{pk}), \\ b \xleftarrow{R} \{0,1\}, c \leftarrow \mathsf{Enc}_{\mathsf{pk}}(m_b)\end{array} : \mathcal{A}_2(s, c) = b\right] - \frac{1}{2}$$

is negligible in the security parameter for any polynomial-time adversary $\mathcal{A}$; the probability is taken over the random coins of the experiment according to the distribution induced by KeyGen and over the random coins of the adversary.

As we are in the public-key setting, the adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ is given the public key pk and so can encrypt any message of its choice. In other words, the adversary can mount chosen-plaintext attacks (CPA). Hence, we write IND-CPA the security notion achieved by a semantically secure encryption scheme.

A stronger scenario is to give the adversary an adaptive access to a decryption oracle. The previous definition readily extends to this model. Adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ is allowed to submit any ciphertext of its choice and receives the corresponding plaintext (or $\perp$); the sole exception is that $\mathcal{A}_2$ may not query the decryption oracle on challenge ciphertext c [14]. We write IND-CCA*2* the corresponding security notion; it stands for indistinguishability under *adaptive* chosen-ciphertext attacks. A weaker security notion is when only $\mathcal{A}_1$ is given access to the decryption oracle [13]. The corresponding security notion is written IND-CCA1 and stands for indistinguishability under *non-adaptive* chosen-ciphertext attacks.

3 A DDH-type Group

Let $q \neq 2$ be a Sophie Germain prime; that is, both q and $2q+1$ are prime. We let $p = 2q+1$. Consider the set $\{1, 2, \ldots, q\}$. This set can be endowed with the structure of a group under the group law $\star$ given by

$$a \star b = |ab \bmod p|$$

where $ab \bmod p$ represents the absolute smallest residue of ab modulo p (namely, the complete set of absolute smallest residues are: $-(p-1)/2, \ldots, -1, 0, 1, \ldots, (p-1)/2$), and where $|ab \bmod p|$ represents the absolute value of $ab \bmod p$. We let $\mathbb{H}_q$ denote the set $\{1, \ldots, q\}$ equipped with the group law $\star$. A similar setting was considered in [9,11] for RSA composites.

Let g be a generator of $(\mathbb{Z}/p\mathbb{Z})^*$ (i.e., $\langle g \rangle = (\mathbb{Z}/p\mathbb{Z})^*$) and $h = |g \bmod p|$. It is easily verified that h generates the group $\mathbb{H}_q$:

$$\mathbb{H}_q = \left\{ |h^j \bmod p| \text{ for } 0 \leq j < q \right\} \ .$$

Indeed, we have $h^q \bmod p = -1$ and thus $|h^q \bmod p| = 1$. Further, for $0 \leq j_1, j_2 < q$, $|h^{j_1} \bmod p| = |h^{j_2} \bmod p|$ implies $h^{2j_1} \equiv h^{2j_2} \pmod{p}$, which in turn implies $j_1 \equiv j_2 \pmod{q}$ and thus $j_1 = j_2$.

Remarkably, as will be stated, the DDH assumption in $\mathbb{H}_q$ is equivalent to the DDH in the subgroup of quadratic residues in $\mathbb{F}_p^*$, which is believed to be hard when $p = 2q+1$ for some prime q. This latter assumption is a standard intractability assumption that has been used in proving the security of a variety of cryptographic schemes.

Theorem 1. *Let $q \neq 2$ be a Sophie Germain prime and let $p = 2q+1$. Let also g be a generator of $(\mathbb{Z}/p\mathbb{Z})^*$, $s = g^2 \bmod p$, and $h = |g \bmod p|$. Define the groups $\mathrm{QR}(p) = \left\{ s^j \bmod p \text{ for } 0 \leq j < q \right\}$ and $\mathbb{H}_q = \left\{ |h^j \bmod p| \text{ for } 0 \leq j < q \right\}$. Then the groups $\mathrm{QR}(p)$ and $\mathbb{H}_q$ are isomorphic; we have*

$$\psi : \mathrm{QR}(p) \xrightarrow{\sim} \mathbb{H}_q, x \longmapsto |\sqrt{x} \bmod p|$$

and

$$\psi^{-1} : \mathbb{H}_q \xrightarrow{\sim} \mathrm{QR}(p), y \longmapsto y^2 \bmod p \ .$$

Proof. Note that the map ψ is well defined. Since q is odd, it follows that $p \equiv 3 \pmod{4}$ and thus $-1 \notin \mathrm{QR}(p)$. Hence square roots exist and are unique in $\mathrm{QR}(p)$: if $x \in \mathrm{QR}(p)$ then $\sqrt{x}$ denotes the unique element $z \in \mathrm{QR}(p)$ such that $z^2 \equiv x \pmod{p}$ —observe that $-z \notin \mathrm{QR}(p)$.

Consider two arbitrary elements $x_1, x_2 \in (\mathbb{Z}/p\mathbb{Z})^*$. For $i \in \{1, 2\}$, letting $z_i = \sqrt{x_i} \in \mathrm{QR}(p)$, we have $\psi(x_i) = |z_i \bmod p|$.

Define $x_3 = x_1 x_2 \bmod p$ and let $z_3 = \sqrt{x_3} \in \mathrm{QR}(p)$. Then we obtain $\psi(x_3) = |z_3 \bmod p| = |z_1 z_2 \bmod p| = \psi(x_1) \star \psi(x_2)$. Map ψ is a group homomorphism.

We have to show that ψ is bijective. Suppose that $\psi(x_1) = \psi(x_2)$. This means that $|z_1 \bmod p| = |z_2 \bmod p|$ and thus $z_1 \equiv \pm z_2 \pmod{p}$. This implies $z_1 = z_2$ since they are both elements of $\mathrm{QR}(p)$. Moreover, for each $y \in \mathbb{H}_q$, there exists an $x \in \mathrm{QR}(p)$; namely, $x = \psi^{-1}(y)$, and $\psi\big(\psi^{-1}(y)\big) = \psi(y^2 \bmod p) = \left| \left(\frac{y}{p}\right) y \bmod p \right| = y$. □

Corollary 1. *The* DDH *in* $\mathbb{H}_q$ *is equivalent to the* DDH *in* $\mathrm{QR}(p)$.

Proof. Let $\mathbb{H}_q = \langle h \rangle$ and $\mathrm{QR}(p) = \langle s \rangle$ with $s = \psi(h)$. Since the isomorphisms between $\mathbb{H}_q$ and $\mathrm{QR}(p)$ are efficiently computable, it is easy to transform a DDH challenge $(h, |h^a \bmod p|, |h^b \bmod p|, |h^c \bmod p|) \in {\mathbb{H}_q}^4$ into a DDH challenge $(s, s^a \bmod p, s^b \bmod p, s^c \bmod p) \in \mathrm{QR}(p)^4$ as $s = \psi(h)$, $s^a \bmod p = \psi(|h^a \bmod p|)$, $s^b \bmod p = \psi(|h^b \bmod p|)$, $s^c \bmod p = \psi(|h^c \bmod p|)$; and vice-versa using ψ^{-1}. □

4 A Variant of ElGamal Cryptosystem

From Corollary 1, we obtain an ElGamal-type cryptosystem that is IND-CPA secure under the standard DDH assumption in $\mathrm{QR}(p)$.

Key generation On input some security parameter 1^λ the key generation algorithm generates a Sophie Germain prime q. It also defines $p = 2q + 1$, a generator g of $(\mathbb{Z}/p\mathbb{Z})^*$, and $h = |g \bmod p|$. Finally it picks at random an element x in $\mathbb{Z}/q\mathbb{Z}$ and computes $y = |h^x \bmod p|$. The public key is $\mathsf{pk} = \{h, p, q, y\}$ while the secret key is $\mathsf{sk} = \{x\}$. The message space is $\mathcal{M} = \{1, \ldots, q\}$.

Encryption The encryption of a message $m \in \mathcal{M}$ is given by $c = (c_1, c_2)$ where

$$c_1 = |h^r \bmod p| \quad \text{and} \quad c_2 = |m\, y^r \bmod p| \ .$$

Decryption Given a ciphertext $c = (c_1, c_2) \in {\mathbb{H}_q}^2$, the corresponding plaintext message m can be recovered using secret key x as

$$m = |c_2 / {c_1}^x \bmod p| \ .$$

5 Chosen-Ciphertext Security

The previous cryptosystem is "malleable". Given the encryptions of messages m and m' in $\mathcal{M}$, say (c_1, c_2) and (c'_1, c'_2), anyone can derive the encryption of message $m'' = m \star m' \in \mathcal{M}$ as $(c''_1, c''_2) = (c_1 \star c'_1, c_2 \star c'_2)$.

While malleability is sometimes useful for certain applications (e.g., for blind decryption), it also rules out the security against chosen-ciphertext attacks. From Corollary 1 and [6, § 5.4], it is possible to get an ElGamal-type cryptosystem that is IND-CCA1 secure under the standard DDH assumption in $\mathrm{QR}(p)$.

Key generation On input some security parameter 1^λ the key generation algorithm generates a Sophie Germain prime q. It also defines $p = 2q + 1$, two generators g and $\bar{g}$ of $(\mathbb{Z}/p\mathbb{Z})^*$, and sets $h = |g \bmod p|$ and $\bar{h} = |\bar{g} \bmod p|$. Finally it picks at random three elements $x, \xi, \bar{\xi}$ in $\mathbb{Z}/q\mathbb{Z}$ and computes $y = |h^x \bmod p|$ and $X = |h^\xi \bar{h}^{\bar{\xi}} \bmod p|$. The public key is $\mathsf{pk} = \{h, \bar{h}, p, q, y, X\}$ while the secret key is $\mathsf{sk} = \{x, \xi, \bar{\xi}\}$. The message space is $\mathcal{M} = \{1, \ldots, q\}$.

Encryption The encryption of a message $m \in \mathcal{M}$ is given by $c = (c_1, \bar{c}_1, c_2, v)$ where

$$c_1 = |h^r \bmod p| \ , \ \ \bar{c}_1 = |\bar{h}^r \bmod p| \ , \ \ c_2 = |m\, y^r \bmod p| \ , \\ \text{and } \ v = |X^r \bmod p| \ .$$

Decryption Given a ciphertext $c = (c_1, \bar{c}_1, c_2, v) \in {\mathbb{H}_q}^4$, the decryption algorithm first checks whether $v = |{c_1}^{\xi}\ {\bar{c}_1}^{\bar{\xi}} \bmod p|$. If so, the corresponding plaintext message m can be recovered using secret key x as

$$m = |c_2 / {c_1}^x \bmod p| \ ;$$

otherwise the decryption algorithm returns $\perp$.

Security against *adaptive* chosen-ciphertext attacks can be achieved by assuming in addition the existence of a hash function $\mathcal{H}$ chosen from a universal one-way family [12]. We note that this requirement is weaker than that of collision-resistance. Doing so, we obtain from Corollary 1 and [6, Section 4] a Cramer-Shoup like cryptosystem that is IND-CCA2 under the standard DDH assumption in $\mathrm{QR}(p)$.

Key generation On input some security parameter 1^λ the key generation algorithm generates a Sophie Germain prime q. It also defines $p = 2q + 1$, two generators g and $\bar{g}$ of $(\mathbb{Z}/p\mathbb{Z})^*$, and sets $h = |g \bmod p|$ and $\bar{h} = |\bar{g} \bmod p|$. It picks at random five elements $x, \xi, \bar{\xi}, \eta, \bar{\eta}$ in $\mathbb{Z}/q\mathbb{Z}$ and computes $y = |h^x \bmod p|$, $X = |h^\xi\, \bar{h}^{\bar{\xi}} \bmod p|$ and $Y = |h^\eta\, \bar{h}^{\bar{\eta}} \bmod p|$. Finally it selects a hash function $\mathcal{H}$ from a family of universal one-way hash functions that map bit string to elements of $\mathbb{Z}/q\mathbb{Z}$. The public key is $\mathsf{pk} = \{h, \bar{h}, p, q, y, X, Y, \mathcal{H}\}$ while the secret key is $\mathsf{sk} = \{x, \xi, \bar{\xi}, \eta, \bar{\eta}\}$. The message space is $\mathcal{M} = \{1, \ldots, q\}$.

Encryption The encryption of a message $m \in \mathcal{M}$ is given by $c = (c_1, \bar{c}_1, c_2, v)$ where

$$c_1 = |h^r \bmod p| \ , \ \ \bar{c}_1 = |\bar{h}^r \bmod p| \ , \ \ c_2 = |m\, y^r \bmod p| \ , \\ \text{and } \ v = |X^r\, Y^{r\alpha} \bmod p| \quad \text{where } \alpha = \mathcal{H}(c_1, \bar{c}_1, c_2) \ .$$

Decryption Given a ciphertext $c = (c_1, \bar{c}_1, c_2, v) \in {\mathbb{H}_q}^4$, the decryption algorithm first computes $\alpha = \mathcal{H}(c_1, \bar{c}_1, c_2)$. Next, using α, it checks whether $v = |{c_1}^{\xi+\eta\alpha}\ {\bar{c}_1}^{\bar{\xi}+\bar{\eta}\alpha} \bmod p|$. If so, the corresponding plaintext message m can be recovered using secret key x as

$$m = |c_2 / {c_1}^x \bmod p| \ ;$$

otherwise the decryption algorithm returns $\perp$.

From Corollary 1 and [6, § 5.3], it is also possible to obtain a Cramer-Shoup like cryptosystem that is IND-CCA2 under the *sole* standard DDH assumption in $\mathrm{QR}(p)$. Hash function $\mathcal{H}$ is eliminated, at the expense of longer keys and slightly increased processing time.

Key generation On input some security parameter 1^λ the key generation algorithm generates a Sophie Germain prime q. It also defines $p = 2q + 1$, two generators g and $\bar{g}$ of $(\mathbb{Z}/p\mathbb{Z})^*$, and sets $h = |g \bmod p|$ and $\bar{h} = |\bar{g} \bmod p|$. Finally it picks at random nine elements $x, \xi, \bar{\xi}, \eta_1, \bar{\eta}_1, \eta_2, \bar{\eta}_2, \eta_3, \bar{\eta}_3$ in $\mathbb{Z}/q\mathbb{Z}$ and computes $y = |h^x \bmod p|$, $X = |h^\xi \, \bar{h}^{\bar{\xi}} \bmod p|$, $Y_1 = |h^{\eta_1} \, \bar{h}^{\bar{\eta}_1} \bmod p|$, $Y_2 = |h^{\eta_2} \, \bar{h}^{\bar{\eta}_2} \bmod p|$ and $Y_3 = |h^{\eta_3} \, \bar{h}^{\bar{\eta}_3} \bmod p|$. The public key is $\mathsf{pk} = \{h, \bar{h}, p, q, y, X, Y_1, Y_2, Y_3, \mathcal{H}\}$ while the secret key is $\mathsf{sk} = \{x, \xi, \bar{\xi}, \eta_1, \bar{\eta}_1, \eta_2, \bar{\eta}_2, \eta_3, \bar{\eta}_3\}$. The message space is $\mathcal{M} = \{1, \ldots, q\}$.

Encryption The encryption of a message $m \in \mathcal{M}$ is given by $c = (c_1, \bar{c}_1, c_2, v)$ where

$$c_1 = |h^r \bmod p| \ , \ \ \bar{c}_1 = |\bar{h}^r \bmod p| \ , \ \ c_2 = |m \, y^r \bmod p| \ ,$$
$$\text{and} \ \ v = |X^r \, {Y_1}^{r c_1} \, {Y_2}^{r \bar{c}_1} \, {Y_3}^{r c_2} \bmod p| \ .$$

Decryption Given a ciphertext $c = (c_1, \bar{c}_1, c_2, v) \in {\mathbb{H}_q}^4$, the decryption algorithm first checks whether $v = |{c_1}^{\xi + \eta_1 c_1 + \eta_2 \bar{c}_1 + \eta_3 c_2} \ {\bar{c}_1}^{\bar{\xi} + \bar{\eta}_1 c_1 + \bar{\eta}_2 \bar{c}_1 + \bar{\eta}_3 c_2} \bmod p|$. If so, the corresponding plaintext message m can be recovered using secret key x as

$$m = |c_2 / {c_1}^x \bmod p| \ ;$$

otherwise the decryption algorithm returns $\perp$.

6 Conclusion

This paper described a simple modification to the classical ElGamal cryptosystem which, at the same time,

- provably meets the IND-CPA security notion (a.k.a. semantic security) in the standard model under the standard DDH assumption, and
- enables the encryption of messages without prior encoding as group elements.

Efficient extensions meeting the stronger security notions of IND-CCA1 and IND-CCA2 (security against chosen-ciphertext attacks) were also presented.

References

1. Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *1st ACM Conference on Computer and Communications Security*, pages 62–73. ACM Press, 1993.
2. Dan Boneh. The decision Diffie-Hellman problem. In J. Buhler, editor, *Algorithmic Number Theory (ANTS-III)*, volume 1423 of *Lecture Notes in Computer Science*, pages 48–63. Springer, 2006.
3. Dan Boneh, Antoine Joux, and Phong Q. Nguyen. Why textbook ElGamal and RSA encryption are insecure. In T. Okamoto, editor, *Advances in Cryptology – ASIACRYPT 2000*, volume 1976 of *Lecture Notes in Computer Science*, pages 30–43. Springer, 2000.

4. Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *Journal of the ACM*, 51(4):557–594, 2004.
5. Benoît Chevallier-Mames, Pascal Paillier, and David Pointcheval. Encoding-free ElGamal encryption without random oracles. In M. Yung et al., editors, *Public Key Cryptography − PKC 2006*, volume 3958 of *Lecture Notes in Computer Science*, pages 91–104. Springer, 2006.
6. Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In H. Krawczyk, editor, *Advances in Cryptology − CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*, pages 13–25. Springer, 1998.
7. Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In G. R. Blakley and D. Chaum, editors, *Advances in Cryptology − Proceedings of CRYPTO 84*, volume 196 of *Lecture Notes in Computer Science*, pages 10–18. Springer, 1985.
8. Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985.
9. Roger Fischlin and Claus P. Schnorr. Stronger security proofs for RSA and Rabin bits. *Journal of Cryptology*, 13(2):221–244, 2000.
10. Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.
11. Dennis Hofheinz and Eike Kiltz. The group of signed quadratic residues and applications. In S. Halevi, editor, *Advances in Cryptology − CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 637–653. Springer, 2009.
12. Moni Naor and Moti Yung. Universal one-way hash functions and their cryptographic applications. In *21st Annual ACM Symposium on Theory of Computing*, pages 33–43. ACM Press, 1989.
13. Moni Naor and Moti Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *22nd Annual ACM Symposium on Theory of Computing*, pages 427–437. ACM Press, 1990.
14. Charles Rackoff and Daniel R. Simon. Noninteractive 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, 1992.
15. Yiannis Tsiounis and Moti Yung. On the security of ElGamal based encryption. In H. Imai and Y. Zheng, editors, *Public Key Cryptography (PKC '98)*, volume 1431 of *Lecture Notes in Computer Science*, pages 117–134. Springer, 1998.