

A survey of generic IND-CCA2 transformations

Ganyu Xu

July 8, 2024

1 Preliminaries

1.1 Public-key encryption schemes

A public key encryption scheme $\text{PKE} = (\text{KeyGen}, \text{E}, \text{D})$ is a collection of three routines. $\text{KeyGen}(1^\lambda)$ takes the security parameter as input and returns a keypair (pk, sk) . $\text{E}(\text{pk}, m)$ takes some public key and some plaintext message $m \in \mathcal{M}_{\text{PKE}}$ and output a ciphertext $c \in \mathcal{C}_{\text{PKE}}$. Where the encryption routine is probabilistic, we model the randomness using a coin $r \in \mathcal{R}$ such that $E(\text{pk}, m; r)$ is deterministic with an explicit r . Finally, $\text{D}(\text{sk}, c)$ uses the secret key to decrypt the ciphertext.

1.1.1 Correctness

Conventionally we require a PKE to be perfectly correct. This means that for all possible key pairs (pk, sk) and plaintexts m , the decryption routine always correctly inverts the encryption routine: $\text{D}(\text{sk}, \text{E}(\text{pk}, m)) = m$.

Where perfect correctness is not achieved, such as with most lattice-based encryption schemes, we need to account for the possibility that decryption can fail. If under some keypair (pk, sk) , a plaintext-ciphertext pair (m, c) is such that $c \stackrel{\$}{\leftarrow} \text{E}(\text{pk}, m)$ is obtained from encrypting m (probabilistically) but $m \neq \text{D}(\text{sk}, c)$, we call it a decryption failure. For probabilistic encryption routines where the coin is uniformly sampled from the coin space, we can quantify the probability that m triggers a decryption failure. From here, we can take the distribution of all keypairs and quantify the “correctness” of a (possibly imperfectly correct) encryption scheme. The following definition of δ -correctness is directly taken from [BDK⁺18].

Definition 1.1 (δ -correctness). *A probabilistic public-key encryption scheme $\text{PKE} = (\text{KeyGen}, \text{E}, \text{D})$ is δ -correct if the expected maximal probability of decryption failure taken across the distribution of keypairs is at most δ :*

$$E \left[\max_{m \in \mathcal{M}} P[D(\text{sk}, E(\text{pk}, m))] \right] \leq \delta$$

where the expectation is taken over the distribution of keypairs and the probability is taken over the distribution of coins.

[HHK17] also defined an adversarial game in which the adversary’s goal is to find some plaintext message to trigger a decryption failure. This adversarial game meaningfully models the real-world scenario in which decryption failure can reveal information about the secret key. Notice that when evaluating the win condition, the coin is uniformly random instead of being chosen by the adversary.

Algorithm 1 CORs

- 1: $(\text{pk}, \text{sk}) \stackrel{\$}{\leftarrow} \text{KeyGen}(1^\lambda)$
 - 2: $m \stackrel{\$}{\leftarrow} A_{\text{CORs}}(1^\lambda, \text{pk}, \text{sk})$
 - 3: **return** $\llbracket \text{D}(\text{sk}, \text{E}(\text{pk}, m)) \neq m \rrbracket$
-

Figure 1: The correctness game CORs

The definition of δ -correctness sets an explicit upper bound on the probability that any plaintext triggers decryption failure. This means that even if the **CORS** adversary actually finds the message m that is the most likely to trigger decryption failure, the probability of winning the **CORS** game is still upper-bounded by δ .

Lemma 1.0.1. *If PKE is δ -correct, then for all CORS adversaries A , even computationally unbounded ones, the probability of winning the CORS game is at most δ*

The values of δ for Kyber are taken directly from [ABD⁺19]

security level	δ
Kyber512	2^{-139}
Kyber768	2^{-164}
Kyber1024	2^{-174}

Table 1: Concrete δ for Kyber

1.1.2 Security

An one-way adversary $A = (A_1, A_2)$ consists of two sub-routines A_1 and A_2 . $s \xleftarrow{\$} A_1^{\mathcal{O}_1}(1^\lambda, \text{pk})$ takes the security parameter, some public-key, access to some oracle(s) \mathcal{O}_1 , and outputs some intermediate state s . $\hat{m} \leftarrow A_2^{\mathcal{O}_2}(1^\lambda, \text{pk}, c^*)$ resumes from the output of A_1 and takes some challenge ciphertext, then tries to guess the corresponding decryption.

The advantage $\text{Adv}_{\text{OW-ATK}}(A)$ of an OW-ATK adversary is the probability that its guess is correct.

Definition 1.2. *A PKE is OW-ATK secure if for all efficient adversaries A , the advantage in the OW-ATK game is negligible with respect to the security parameter:*

$$\text{Adv}_{\text{OW-ATK}}(A) \leq \text{negl}(\lambda)$$

An **indistinguishability** adversary $A = (A_1, A_2)$ similarly consists of two sub-routines. The first sub-routine adversarially chooses two distinct plaintext messages, and the second sub-routine tries to distinguish which of the two plaintext messages is the decryption of the challenge encryption. The advantage of an indistinguishability adversary is defined by $\text{Adv}_{\text{IND-ATK}}(A) = P[\hat{b} = b] - \frac{1}{2}$

Definition 1.3. *A PKE is IND-ATK secure if for all efficient adversaries A , the advantage in the indistinguishability game is negligible with respect to the security parameter*

$$\text{Adv}_{\text{IND-ATK}}(A) \leq \text{negl}(\lambda)$$

The security games are described in details in figure 2

Algorithm 2 OW-ATK game	Algorithm 3 IND-ATK game
1: $(\text{pk}, \text{sk}) \xleftarrow{\$} \text{KeyGen}(1^\lambda)$	1: $(\text{pk}, \text{sk}) \xleftarrow{\$} \text{KeyGen}(1^\lambda)$
2: $s \xleftarrow{\$} A_1^{\mathcal{O}_1}(1^\lambda, \text{pk})$	2: $(m_0, m_1) \xleftarrow{\$} A_1^{\mathcal{O}_1}(1^\lambda, \text{pk})$
3: $m^* \xleftarrow{\$} \mathcal{M}$	3: $b \xleftarrow{\$} \{0, 1\}$
4: $c^* \xleftarrow{\$} \text{E}(\text{pk}, m^*)$	4: $c^* \xleftarrow{\$} \text{E}(\text{pk}, m_b)$
5: $\hat{m} \leftarrow A_2^{\mathcal{O}_2}(1^\lambda, \text{pk}, s, c^*)$	5: $\hat{b} \leftarrow A_2^{\mathcal{O}_2}(1^\lambda, \text{pk}, s, c^*)$
6: return $[\hat{m} = m^*]$	6: return $[\hat{b} = b]$

Figure 2: The indistinguishability security game

The capabilities of the adversaries are modeled using different collections of oracles. In standard security requirements, adversaries with access to no additional oracles can only mount chosen plaintext attacks (CPA),

adversaries with access to decryption oracle \mathcal{O}^D only before the receiving challenge ciphertext can mount non-adaptive chosen ciphertext attacks (CCA1), adversaries with access to decryption oracle both before and after receiving the challenge ciphertext can mount adaptive chosen ciphertext attacks (CCA2).

[HHK17] also defined two non-standard oracles and the corresponding attacks. The plaintext checking oracle $\text{PCO}(m, c)$ returns 1 if m is a decryption of c and 0 otherwise. The ciphertext validation oracle $\text{CVO}(c)$ returns 1 if c is a valid ciphertext and 0 otherwise.

Algorithm 4 $\text{PCO}(m, c)$	Algorithm 5 $\text{CVO}(c)$
return $\llbracket D(\text{sk}, c) = m \rrbracket$	1: return $\llbracket D(\text{sk}, c) \in \mathcal{M} \rrbracket$

Figure 3: PCO and CVO

Here is an overview of the various kinds of attacks and their associated oracles

ATK	\mathcal{O}_1	\mathcal{O}_2
CPA	—	
CCA1	\mathcal{O}^D	-
CCA2	\mathcal{O}^D	
PCVA	PCO, CVO	
PCA	PCO	
VA	CVO	

Table 2: Attacks and associated oracle access

[HHK17] stated a “well-known” result that the IND-CPA security of a scheme with a large message space implies OW-CPA security:

Theorem 1.1. *For every IND-CPA adversary B against some PKE, there exists an OW-CPA adversary A against the same PKE such that:*

$$\text{Adv}_{\text{OW-CPA}}(A) = \frac{1}{|\mathcal{M}|} + \text{Adv}_{\text{IND-CPA}}(B)$$

1.1.3 Spread and rigidity

The spread of a public key encryption scheme measures the diffusion the encryption routine’s output. The higher the spread, the lower the probability of obtaining any specific ciphertext.

Definition 1.4 (γ -spread). *For a given keypair (pk, sk) and plaintext message $m \in \mathcal{M}$, the min-entropy of the encryption routine is:*

$$\text{min-entropy}(\text{pk}, m) := -\log_2 \left(\max_{c \in \mathcal{C}} P[c = E(\text{pk}, m)] \right)$$

A PKE has γ -spread if for all keypairs (pk, sk) and plaintext $m \in \mathcal{M}$:

$$\text{min-entropy}(\text{pk}, m) \leq \gamma$$

Having γ spread means that for any keypair (pk, sk) , plaintext m , and ciphertext c :

$$P[c = E(\text{pk}, m)] \leq 2^{-\lambda}$$

Finally, *rigidity* conveys the idea that a ciphertext cannot be perturbed without becoming either invalid or decrypting to another plaintext

Definition 1.5 (rigidity). *PKE(KeyGen, E, D) is rigid if for all keypairs (pk, sk) and ciphertext c , either $D(\text{sk}, c) = \perp$ or $E(\text{pk}, D(\text{sk}, c)) = c$*

1.2 Key encapsulation mechanism

A key encapsulation mechanism $\text{KEM} = (\text{KeyGen}, \text{Encap}, \text{Decap})$ is a collection of three routines. The key generation routine $(\text{pk}, \text{sk}) \xleftarrow{\$} \text{KeyGen}(1^\lambda)$ takes the security parameter 1^λ and returns a keypair. The encapsulation routine $(c, K) \xleftarrow{\$} \text{Encap}(\text{pk})$ takes the public key and outputs some ciphertext $c \in \mathcal{C}_{\text{KEM}}$ and some shared secret $K \in \mathcal{K}_{\text{KEM}}$. Finally, the decapsulation routine $K \leftarrow \text{Decap}(\text{sk}, c)$ takes the secret key and a ciphertext and outputs the corresponding shared secret.

Correctness: similar to a PKE, key encapsulation mechanisms are usually required to be perfectly correct, meaning that for all keypairs (pk, sk) , decapsulation always outputs the same shared secret as the encapsulation

$$\mathbb{P} \left[(c, K_1) \xleftarrow{\$} \text{Encap}(\text{pk}); K_2 \leftarrow \text{Decap}(\text{sk}, c); K_1 = K_2 \right] = 1$$

However, where decapsulation failure has a non-zero probability, we need to upperbound this quantity.

Definition 1.6 (δ -correctness). *A $\text{KEM} = (\text{KeyGen}, \text{Encap}, \text{Decap})$ is δ -correct if the probability of decapsulation failure taken across the keypair distribution is at most δ :*

$$\mathbb{P} \left[\text{Decap}(\text{sk}, c) \neq K \mid (\text{pk}, \text{sk}) \xleftarrow{\$} \text{KeyGen}(); (c, K) \xleftarrow{\$} \text{Encap}(\text{pk}) \right] \leq \delta$$

Security: the security of a KEM is defined in an adversarial game in which the adversary's goal is to distinguish between shared secret derived from encapsulation and uniformly random samples. An adversary $A = (A_1, A_2)$ contains two sub-routines with access to some oracle \mathcal{O} depending on game.

Algorithm 6 IND-CCA2 game

- 1: $(\text{pk}, \text{sk}) \xleftarrow{\$} \text{KeyGen}(1^\lambda)$
 - 2: $s \xleftarrow{\$} A_1^{\mathcal{O}^{\text{Decap}}}(1^\lambda, \text{pk})$
 - 3: $(c^*, K_0) \xleftarrow{\$} \text{Encap}(\text{pk})$
 - 4: $K_1 \xleftarrow{\$} \mathcal{K}_{\text{KEM}}$
 - 5: $b \xleftarrow{\$} \{0, 1\}$
 - 6: $\hat{b} \xleftarrow{\$} A_2^{\mathcal{O}^{\text{Decap}}}(1^\lambda, \text{pk}, s, c^*, K_b)$
 - 7: **return** $\llbracket \hat{b} = b \rrbracket$
-

Algorithm 7 Decap oracle $\mathcal{O}^{\text{Decap}}(c \neq c^*)$

- 1: **return** $\text{Decap}(\text{sk}, c)$
-

Figure 4: The IND-CCA2 game for KEM

Definition 1.7 (IND-CCA2 security). *A $\text{KEM} = (\text{KeyGen}, \text{Encap}, \text{Decap})$ is IND-CCA2 secure if no efficient adversary has non-negligible advantage in the IND-CCA2 game.*

2 Modular Fujisaki-Okamoto transformation

The Fujisaki-Okamoto transformation [FO99] and its KEM variations [HHK17] achieve security against adaptive chosen-ciphertext attacks through *de-randomization* and *re-encryption*. In particular, the modular FO transformation contains two steps. The first step (denoted the T transformation) takes a $\text{PKE} = (\text{KeyGen}, \text{E}, \text{D})$ and outputs a $\text{PKE}_1 = (\text{KeyGen}, \text{E}_1, \text{D}_1)$. The key generation remains unchanged, but the encryption routine is *de-randomized* by deriving the coin from the plaintext using a hash function $G : \mathcal{M}_{\text{PKE}} \rightarrow \mathcal{R}_{\text{PKE}}$, and the decryption routine uses *re-encryption* to reject invalid ciphertexts.

<hr/> Algorithm 8 $E_1(\text{pk}, m)$ <hr/> 1: $r \leftarrow G(m)$ 2: $c \leftarrow E(\text{pk}, m; r)$ 3: return c <hr/>	<hr/> Algorithm 9 $D_1(\text{sk}, c)$ <hr/> 1: $\hat{m} \leftarrow D(\text{sk}, c)$ 2: if $\hat{m} \in \mathcal{M}_{\text{PKE}} \wedge E(\text{pk}, \hat{m}; G(\hat{m})) = c$ then 3: return \hat{m} 4: end if 5: return \perp <hr/>
---	--

Figure 5: T transformation

[HHK17] states the security property of PKE_1 as follows:

Theorem 2.1. *If PKE is δ -correct and has γ spread, then for every OW-PCVA adversary B against PKE_1 who makes q_G hash queries, q_V ciphertext validation queries, and q_P plaintext checking queries, there exists an OW-CPA adversary A such that:*

$$\text{Adv}(B) \leq q_V \cdot 2^{-\gamma} + (q_G + q_P) \cdot \delta + (1 + q_G + q_P) \cdot \text{Adv}(A)$$

In other words, if PKE is OW-CPA secure, then PKE_1 is OW-PCVA secure, though the security is non-tight. On the other hand, if PKE is IND-CPA secure, the OW-PCVA security is tight:

Theorem 2.2. *for every OW-PCVA adversary B against PKE_1 who makes q_G hash queries, q_V ciphertext validation queries, and q_P plaintext checking queries, there exists an IND-CPA adversary A such that:*

$$\text{Adv}(B) \leq q_V \cdot 2^{-\gamma} + (q_G + q_P) \cdot \delta + \frac{2 \cdot q_G + 1}{|\mathcal{M}_{\text{PKE}}|} + 3 \cdot \text{Adv}(A)$$

Proof. We will provide a sketch of proof for theorem 2.1 and 2.2. The main idea is to construct an OW-CPA/IND-CPA adversary A who can simulate the OW-PCVA game and use the OW-PCVA adversary B as a sub-routine. However, there are three main difficulties with constructing a “convincing” simulation:

1. A has no access to PCO
2. A has no access to CVO
3. The challenge ciphertext A receives is obtained using a truly random coin, but the challenge ciphertext B expects is obtained using a pseudorandom coin $r \leftarrow G(m)$

□

2.1 From PKE to KEM

The second part of the modular FO transform (denoted the U transform) takes a PKE and outputs a KEM. Depending on whether the input PKE is rigid and whether the output KEM rejects invalid ciphertext implicitly or explicitly, the security requirements for the input PKE and the construction of the KEM will vary. [HHK17] presents four variations of the U transform, which are listed in table 3

Name	PKE requirements	rejection	shared secret
U^\perp	OW-PCVA	\perp	$H(m, c)$
$U^\mathcal{L}$	OW-PCA	$H(s, c)$	$H(m, c)$
U_m^\perp	rigidity + OW-VA	\perp	$H(m)$
$U_m^\mathcal{L}$	rigidity + OW-CPA	$H(s, c)$	$H(m)$

Table 3: A summary of variants of U transformations

For the remaining of this section we will present the four transformations and sketch proofs of their securities.

2.1.1 KEM with explicit rejection from randomized PKE

Let $\text{PKE} = (\text{KeyGen}, \text{E}, \text{D})$ be a public-key encryption scheme, and $H : \mathcal{M}_{\text{PKE}} \times \mathcal{C}_{\text{PKE}} \rightarrow \mathcal{K}_{\text{KEM}}$ be a hash function. The U^\perp transformation outputs a $\text{KEM}^\perp = (\text{KeyGen}^\perp, \text{Encap}^\perp, \text{Decap}^\perp)$, where the key generation routine remains unchanged: $\text{KeyGen}^\perp = \text{KeyGen}$. The encapsulation and decapsulation routines are described in figure 6

Algorithm 10 $\text{Encap}^\perp(\text{pk})$	Algorithm 11 $\text{Decap}^\perp(\text{sk}, c)$
<hr/> 1: $m \xleftarrow{\$} \mathcal{M}_{\text{PKE}}$ 2: $c \xleftarrow{\$} \text{E}(\text{pk}, m)$ 3: $K \leftarrow H(m, c)$ 4: return (c, K) <hr/>	<hr/> 1: $\hat{m} \leftarrow \text{D}(\text{sk}, c)$ 2: if $\hat{m} \in \mathcal{M}_{\text{PKE}}$ then 3: $K \leftarrow H(\hat{m}, c)$ 4: return K 5: end if 6: return \perp <hr/>

Figure 6: U^\perp routines

Under the random oracle model, the security of KEM^\perp depends on the security of the input PKE.

Theorem 2.3. *For every IND-CCA2 adversary B against KEM^\perp , there exists an OW-PCVA adversary A against the underlying PKE such that*

$$\text{Adv}(B) \leq \text{Adv}(A)$$

Proof. For a sketch of proof, we argue that under the random oracle model, $H(m, c)$ is indistinguishable from a uniformly random sample from \mathcal{K}_{KEM} from the adversary B 's perspective unless B somehow queries H on (m, c) . This means that an OW-PCVA adversary can perfectly simulate the decapsulation oracle using uniformly random samples, as long as the outputs from the simulated decapsulation oracle are consistent with the outputs from the hash functions.

The only flaw in the simulated decapsulated oracle lies in the case when B queries $H(m^*, c^*)$, but because A has access to PCO, A can detect that B has successfully recovered m^* and uses the recovered m^* to win the OW-PCVA game. In other words, A keeps running B in a simulation until B makes the special hash query, at which time A simply terminates B and outputs the answer.

To make the simulated decapsulation oracle and the hash oracle consistent, A maintains two tapes $\mathcal{L}^{\text{Decap}}$ and \mathcal{L}^H that record the queries made to the decapsulation oracle and the hash oracles respectively:

$$\begin{aligned} (c, K) \in \mathcal{L}^{\text{Decap}} &\Leftrightarrow K = \mathcal{O}^{\text{Decap}}(c) \\ (m, c, K) \in \mathcal{L}^H &\Leftrightarrow K = H(m, c) \end{aligned}$$

If $\mathcal{O}^{\text{Decap}}$ has been queried with some valid ciphertext c , then when H is queried with a corresponding input (m, c) where m is the decryption of c , H needs to output the same value as $\mathcal{O}^{\text{Decap}}$. Similarly, if H has been queried with some valid plaintext-ciphertext pair (m, c) , then when $\mathcal{O}^{\text{Decap}}$ is queried with c , the two oracles need to output the same value. The details of the simulation are described in figure 7

Algorithm 12 $\mathcal{O}_1^{\text{Decap}}(c)$

```
1: if  $\exists(\tilde{c}, \tilde{K}) \in \mathcal{L}^{\text{Decap}} : \tilde{c} = c$  then
2:   return  $\tilde{K}$ 
3: end if
4: if  $\text{CV0}(c) = 1$  then
5:    $K \xleftarrow{\$} \mathcal{K}_{\text{KEM}}$ 
6:   Append  $(c, K)$  to  $\mathcal{L}^{\text{Decap}}$ 
7:   return  $K$ 
8: end if
9: return  $\perp$ 
```

Algorithm 13 $H_1(m, c)$

```
1: if  $\exists(\tilde{m}, \tilde{c}, \tilde{K}) \in \mathcal{L}^H : (\tilde{m}, \tilde{c}) = (m, c)$  then
2:   return  $\tilde{K}$ 
3: end if
4:  $K \xleftarrow{\$} \mathcal{K}_{\text{KEM}}$ 
5: if  $\text{PC0}(m, c) = 1$  then
6:   if  $\exists(\tilde{c}, \tilde{K}) \in \mathcal{L}^{\text{Decap}} : \tilde{c} = c$  then
7:     return  $\tilde{K}$ 
8:   else
9:     Append  $(c, K)$  to  $\mathcal{L}^{\text{Decap}}$ 
10:  end if
11: end if
12: Append  $(m, c, K)$  to  $\mathcal{L}^H$ 
13: return  $K$ 
```

Figure 7: Patched oracles in U^\perp

□

2.1.2 KEM with implicit rejection from randomized PKE

Given $\text{PKE} = (\text{KeyGen}, \text{E}, \text{D})$, the U^\perp transformation outputs $\text{KEM} = (\text{KeyGen}^\perp, \text{Encap}^\perp, \text{Decap}^\perp)$. With implicit rejection, decapsulation routine returns some output from H even when the input ciphertext c is malformed. Specifically, the implicit rejection value depends on the ciphertext and some secret value s that is a part of the secret key.

Algorithm 14 KeyGen^\perp

Figure 8: U^\perp routines

References

- [ABD⁺19] Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. Crystals-kyber algorithm specifications and supporting documentation. *NIST PQC Round*, 2(4):1–43, 2019.
- [BDK⁺18] Joppe Bos, Léo Ducas, Eike Kiltz, Tancrede Lepoint, Vadim Lyubashevsky, John M Schanck, Peter Schwabe, Gregor Seiler, and Damien Stehlé. Crystals-kyber: a cca-secure module-lattice-based kem. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 353–367. IEEE, 2018.
- [FO99] Eiichiro Fujisaki and Tatsuaki Okamoto. Secure integration of asymmetric and symmetric encryption schemes. In *Annual international cryptography conference*, pages 537–554. Springer, 1999.
- [HHK17] Dennis Hofheinz, Kathrin Hövelmanns, and Eike Kiltz. A modular analysis of the fujisaki-okamoto transformation. In *Theory of Cryptography Conference*, pages 341–371. Springer, 2017.