

**Problem 8):** For a public key encryption scheme  $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ , we define **CPA** security according to the probability obtaining a secure result, as defined in the privacy experiment  $\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}$ . This experiment goes as follows

**The LR-oracle experiment**  $\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n)$

1.  $\text{Gen}(1^n)$  is run to obtain keys  $(pk, sk)$ .
2. A uniform bit  $b \in \{0, 1\}$  is chosen.
3. The adversary  $\mathcal{A}$  is given input  $pk$  and oracle access to  $\text{LR}_{pk, b}(\cdot, \cdot)$ .
4. The adversary  $\mathcal{A}$  outputs a bit  $b'$ .
5. The adversary  $\mathcal{A}$  is defined to be 1 if  $b' = b$ , and 0 otherwise. If  $\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n) = 1$ , we say that  $\mathcal{A}$  **succeeds**.

Using this definition for the experiment  $\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}$ , we say that the encryption scheme  $\Pi$  is secure if the probability of  $\mathcal{A}$  succeeding,  $\Pr[\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n) = 1]$  satisfies the condition

$$\Pr[\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n) = 1] \leq \frac{1}{2} + \text{negl}(n) \quad (8.1)$$

where  $\text{negl}(n)$  is a function/value which is negligible on the order of  $n$ .

In detail, what we are seeking is indistinguishability of multiple encryptions. That is to say, if we have the plain-text of two different messages (denote them  $m_1$  and  $m_2$ ), which we encrypt using a public key (denote it  $pk$ ), then an adversary  $\mathcal{A}$  having access to the cipher-text of both messages **and** the public key should not be able to distinguish the cipher-text of the messages under any circumstances. Using  $pk$ , the encryption algorithm (denoted  $\text{Enc}_{pk}$ ) generates cipher-text from messages  $m_1$  and  $m_2$ . We use

$$\text{Enc}_{pk}(m_1) \quad \text{and} \quad \text{Enc}_{pk}(m_2)$$

to denote the cipher-text generated for these messages, respectively.

We denote both the information  $(pk, \text{Enc}_{pk}(m_1), \& \text{Enc}_{pk}(m_2))$  available/provided to the adversary  $\mathcal{A}$  by

$$\mathcal{A}(pk, \text{Enc}_{pk}(m_1), \text{Enc}_{pk}(m_2)) \quad (8.2)$$

furthermore, we also use this notation to represent the outcome of running PubK on  $\mathcal{A}$ . When  $\mathcal{A}$  succeeds, then the expression in 8.2 yields the result

$$\mathcal{A}(pk, \text{Enc}_{pk}(m_1), \text{Enc}_{pk}(m_2)) = 1 \quad (8.3)$$

The expression in 8.2 yields

$$\mathcal{A}(pk, \text{Enc}_{pk}(m_1), \text{Enc}_{pk}(m_2)) = 0 \quad (8.4)$$

otherwise.

Since **CPA** security requires security over multiple encryptions using the same public key, we will formally define this security using **two** pairs of messages that are all being encrypted using the same public key. We denote the first pair of messages by  $m_{1,0}$  and  $m_{2,0}$ . Similarly, the second pair of messages are denoted by  $m_{1,1}$  and  $m_{2,1}$ . We now use the same notation as in 8.2 with these message pairs (*and their associated public key  $pk$* ) to represent the attack by  $\mathcal{A}$ . This gives

$$\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})), \quad (8.2 \text{ a})$$

for the first message pair; and

$$\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})), \quad (8.2 \text{ b})$$

for the second message pair.

Before proceeding, we point out that we can equivalently use the expression from 8.3 in place of the  $\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n) = 1$  term from 8.1. More clearly, we may formally write this equivalence as

$$\text{PubK}_{\mathcal{A}, \Pi}^{\text{LR-cpa}}(n) = 1 \quad \longleftrightarrow \quad \mathcal{A}(pk, \text{Enc}_{pk}(m_1), \text{Enc}_{pk}(m_2)) = 1$$

This allows us to write a version of 8.1 for both and . For the first message pair (*represented in*), this gives the result

$$\Pr[\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})) = 1] \leq \frac{1}{2} + \text{negl}_0(n), \quad (8.3)$$

where  $\text{negl}_0$  represents the negligible function required to satisfy this expression as applied to this message pair (*we are making allowances in case the results in and use different negl functions*). Writing our expression for the second message pair In a similar fashion yields

$$\Pr[\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})) = 1] \leq \frac{1}{2} + \text{negl}_1(n) \quad (8.4)$$

where  $\text{negl}_1$  represents the negligible function required to satisfy this expression as applied to this message pair just as before (*we will see later that any difference between these negl functions is inconsequential; however differentiating between the negl functions used in either case is required for mathematical rigor*).

To continue the equation in 8.4 is subtracted from the equation in 8.3, after which the result *difference* will be simplified, thereby allowing us to obtain the following expressions for the initial and then the simplified results

$$\begin{aligned} & \left\{ \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})) = 1] - \right. \\ & \quad \left. - \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})) = 1] \right\} \leq \left( \frac{1}{2} + \text{negl}_0(n) \right) - \left( \frac{1}{2} + \text{negl}_1(n) \right) \\ & \left\{ \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})) = 1] - \right. \\ & \quad \left. - \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})) = 1] \right\} \leq \text{negl}_0(n) - \text{negl}_1(n) \end{aligned}$$

Taking the absolute value of this simplified expression allows us to obtain the result

$$\begin{aligned} & \left| \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})) = 1] - \right. \\ & \quad \left. - \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})) = 1] \right| \leq \left| \text{negl}_0(n) - \text{negl}_1(n) \right| \end{aligned} \quad (8.5)$$

Considering the right-hand-side of 8.5, we see that  $\left| \text{negl}_0(n) - \text{negl}_1(n) \right|$  also negligible itself. Therefore, we may define another negligible function, of order  $n$ , that satisfies the relation

$$\left| \text{negl}_0(n) - \text{negl}_1(n) \right| = \text{negl}(n),$$

where  $\text{negl}$  is another negligible function, of order  $n$ . Applying this to the expression in 8.5, we obtain the final result

$$\begin{aligned} & \left| \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,0}), \text{Enc}_{pk}(m_{2,0})) = 1] - \right. \\ & \quad \left. - \Pr [\mathcal{A}(pk, \text{Enc}_{pk}(m_{1,1}), \text{Enc}_{pk}(m_{2,1})) = 1] \right| \leq \text{negl}(n) \end{aligned} \quad (8.6)$$

which provides a formal definition for **CPA** security. In simple terms, the expression in 8.6 formally

describes the requirement that a **CPA** secure encryption scheme be **non-deterministic**. That is to say, the expression in 8.6 mathematically quantifies the requirement that the cipher-text generated by any **CPA** secure encryption scheme be indistinguishable for any arbitrary pair of messages. It is the arbitrary nature of the messages that give rise to the requirement for non-determinism because the result in 8.6 must hold when the messages are **identical**. The only way for identical messages to be indistinguishably enciphered is for the encryption scheme used to encipher them to allow, with some non-zero probability, every possible message in the message space ( $\mathcal{M}$ ) to be encrypted into any cipher-text in the cipher-text space ( $\mathcal{C}$ ).

We now