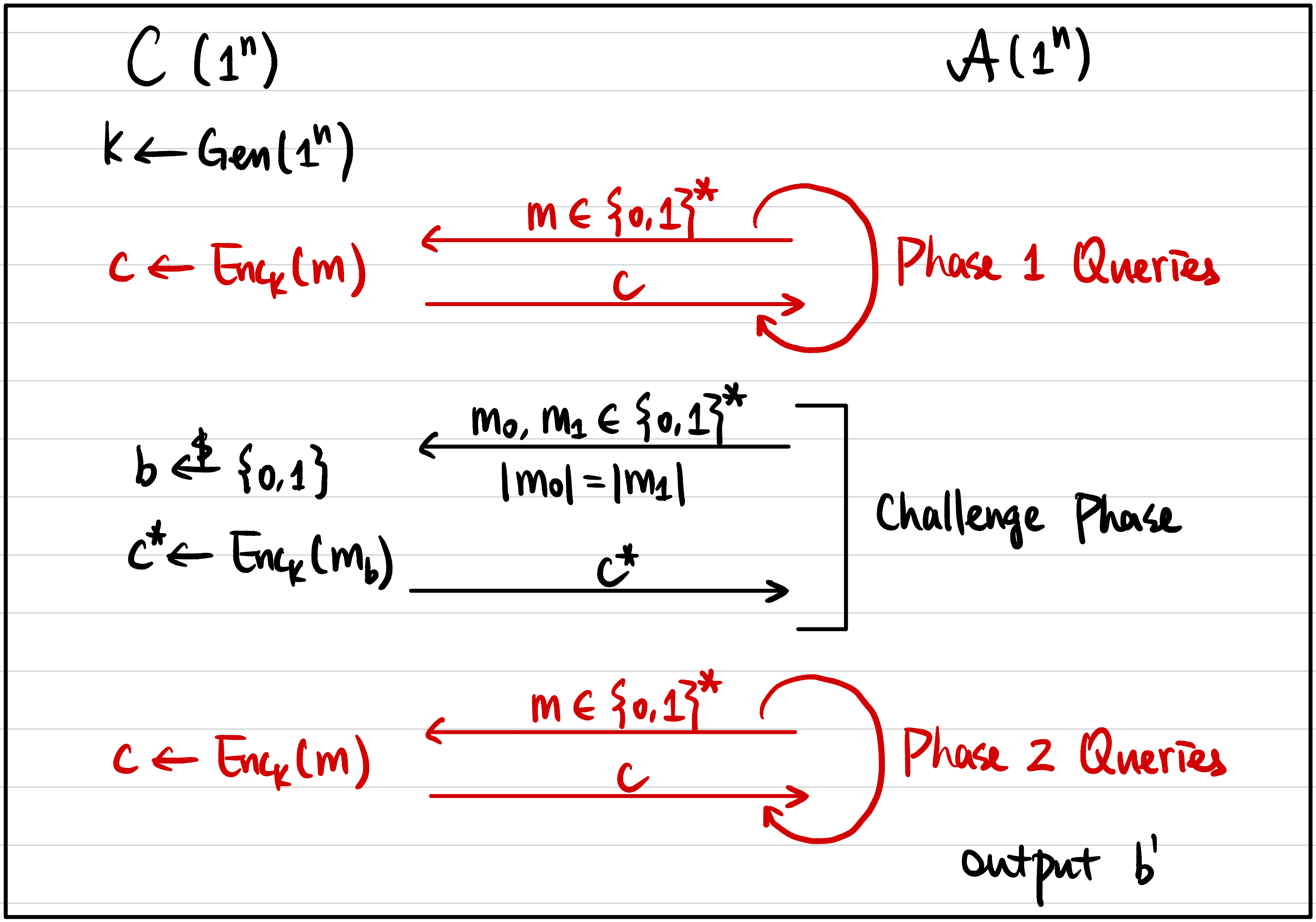
CSCI 1510

- Pseudorandom Function (continued)
- CPA-Secure Encryption from PRF
- Hybrid Argument
- Message Authentication Code (MAC)

Chosen Plaintext Attack (CPA) Security

Def A symmetric-key encryption scheme (Gen, Enc, Dec) is secure against chosen plaintext attacks, or CPA-secure, if $\forall$ PPT $\mathcal{A}$, $\exists$ negligible function $\varepsilon(\cdot)$ s.t. $\Pr[b = b'] \leq \frac{1}{2} + \varepsilon(n)$



Pseudorandom Function (PRF)

Def Let $F: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a deterministic, poly-time, keyed function. F is a pseudorandom function (PRF) if $\forall$ PPT $\mathcal{A}$, $\exists$ negligible function $\varepsilon(\cdot)$ s.t. $\Pr[b=b'] \le \frac{1}{2} + \varepsilon(n)$

$C(1^n)$ | $\mathcal{A}(1^n)$

$b \xleftarrow{\$} \{0,1\}$

If $b=0$, then $k \xleftarrow{\$} \{0,1\}^n$

If $b=1$, then $f \xleftarrow{\$} \mathrm{Func}_n$ — $\{F \mid F: \{0,1\}^n \to \{0,1\}^n\}$

$\xleftarrow{\quad x \quad}$

If $b=0$, then $y := F_k(x)$

If $b=1$, then $y := f(x)$

$\xrightarrow{\quad y \quad}$

output b'

Exercises

Let $F: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a PRF.

Define $F': \{0,1\}^n \times \{0,1\}^{n-1} \to \{0,1\}^{2n}$ as follows.

Is F' necessarily a PRF?

a) $F'_k(x) = F_k(0\|x) \,\|\, F_k(0\|x)$ $\qquad F_k(0\,\boxed{\;x\;}) \,\|\, F_k(0\,\boxed{\;x\;})$

b) $F'_k(x) = F_k(0\|x) \,\|\, F_k(1\|x)$ $\qquad F_k(0\,\boxed{\;x\;}) \,\|\, F_k(1\,\boxed{\;x\;})$

c) $F'_k(x) = F_k(0\|x) \,\|\, F_k(x\|0)$ $\qquad F_k(0\,\boxed{\;x\;}) \,\|\, F_k(\boxed{\;x\;}\,0)$

d) $F'_k(x) = F_k(0\|x) \,\|\, F_k(x\|1)$ $\qquad F_k(0\,\boxed{\;x\;}) \,\|\, F_k(\boxed{\;x\;}\,1)$

a) C $\qquad\qquad$ A

$C \xleftarrow{\;x\;} A$

$C \xrightarrow{\;y_1 \| y_2\;} A$

$y_1 \overset{?}{=} y_2$

c) C $\qquad\qquad$ A

$C \xleftarrow{\;x = 0\cdots0\;} A$

$C \xrightarrow{\;y_1 \| y_2\;} A$

$y_1 \overset{?}{=} y_2$

d) C $\qquad\qquad$ A

$C \xleftarrow{\;x_1 = 0\cdots0\;} A$

$C \xrightarrow{\;y_1 \| y_2\;} A$

$C \xleftarrow{\;x_2 = 0\cdots1\;} A$

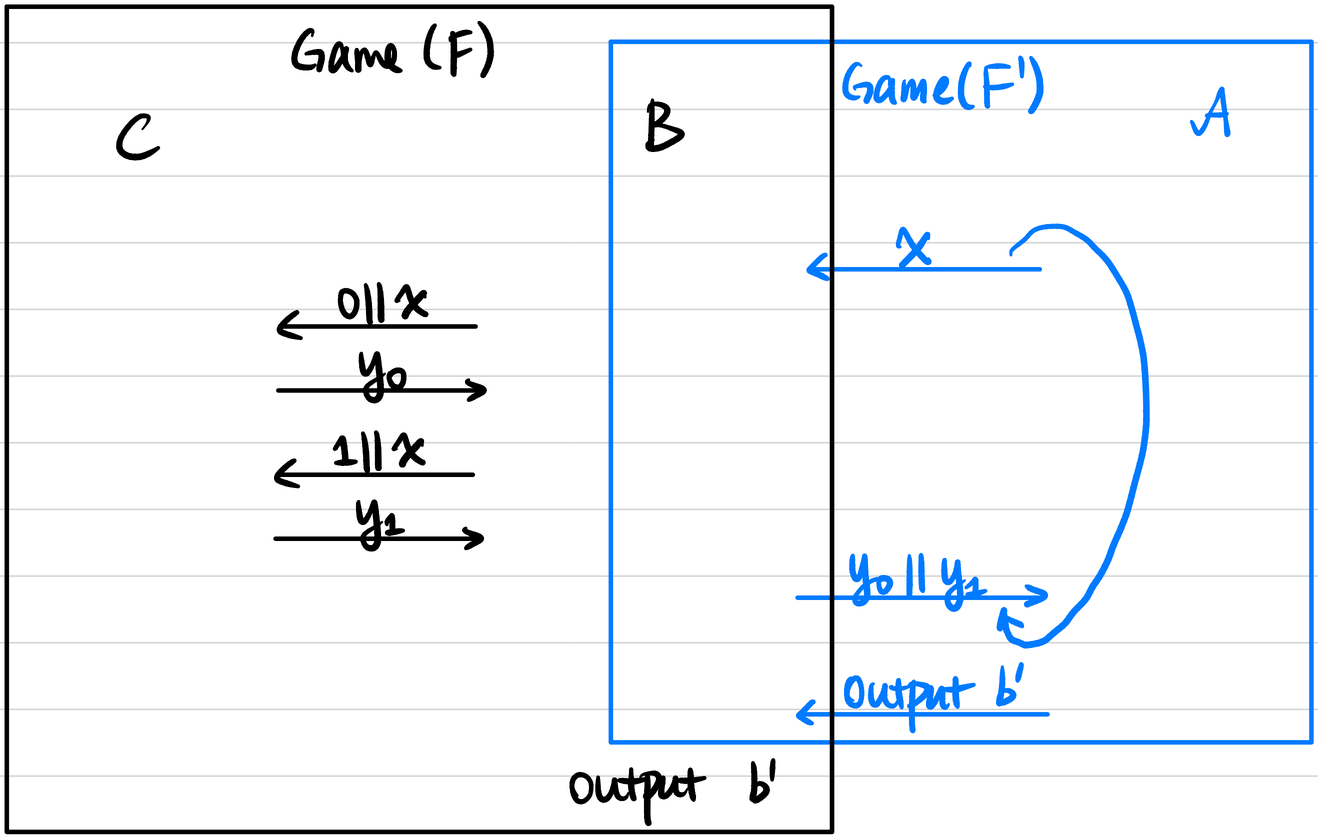
$C \xrightarrow{\;y_3 \| y_4\;} A$

$y_2 \overset{?}{=} y_3$

b) $F'_k(x) = F_k(0\|x) \,\|\, F_k(1\|x)$ is a PRF

Proof Assume not, then $\exists$ PPT $\mathcal{A}$ that breaks the pseudorandomness of F'.

We construct PPT B to break the pseudorandomness of F.



If C uses F_k, then $y_0 \| y_1 = F_k(0\|x) \| F_k(1\|x) = F'_k(x)$

$\checkmark A$ is interacting with F'_k

If C uses a random function, then $y_0 \| y_1 = f(0\|x) \| f(1\|x)$

For distinct inputs x, $0\|x$ & $1\|x$ are all distinct,

y_0 & y_1 are independent random strings,

$\checkmark A$ is interacting with a random function.

$\Pr[B$ guesses correctly for $F] = \Pr[\checkmark A$ guesses correctly for $F'] \geq \frac{1}{2} +$ non-negl(n).

PRF $\Leftrightarrow$ PRG

"$\Rightarrow$": Let $F: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a PRF,

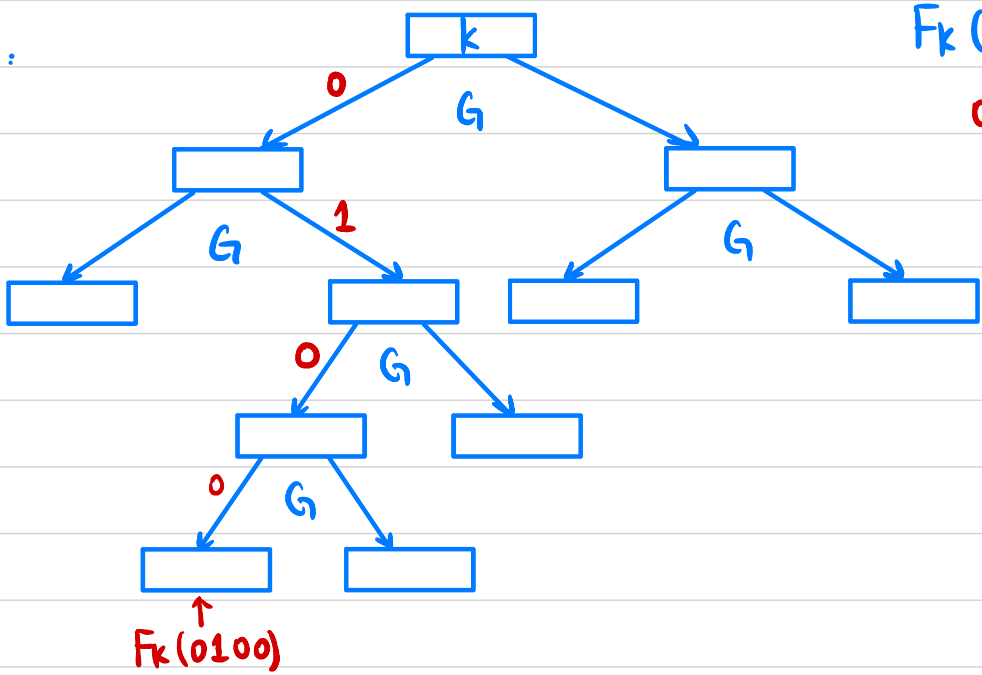
Construct $G: \{0,1\}^n \to \{0,1\}^{2n}$

$G(s) := F_s(0\cdots0) \,\|\, F_s(0\cdots01)$

"$\Leftarrow$": Let $G: \{0,1\}^n \to \{0,1\}^{2n}$ be a PRG,

Construct $F: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$

GGM tree:



$F_k(x)$

x: 0100

Constructing CPA-Secure Encryption

Pseudorandom Function (PRF)

$$\Downarrow$$

CPA-Secure Encryption

CPA-Secure Encryption Scheme

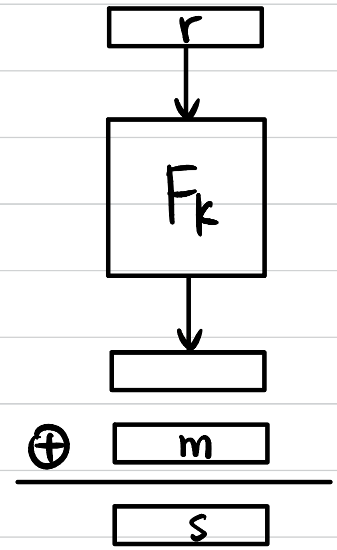
Let $F: \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a PRF,

- $Gen(1^n)$: sample $k \xleftarrow{\$} \{0,1\}^n$, output k.
- $Enc_k(m)$: $m \in \{0,1\}^n$

  $r \xleftarrow{\$} \{0,1\}^n$

  output $c := \langle r, F_k(r) \oplus m \rangle$
- $Dec_k(c)$: $c = \langle r, s \rangle$

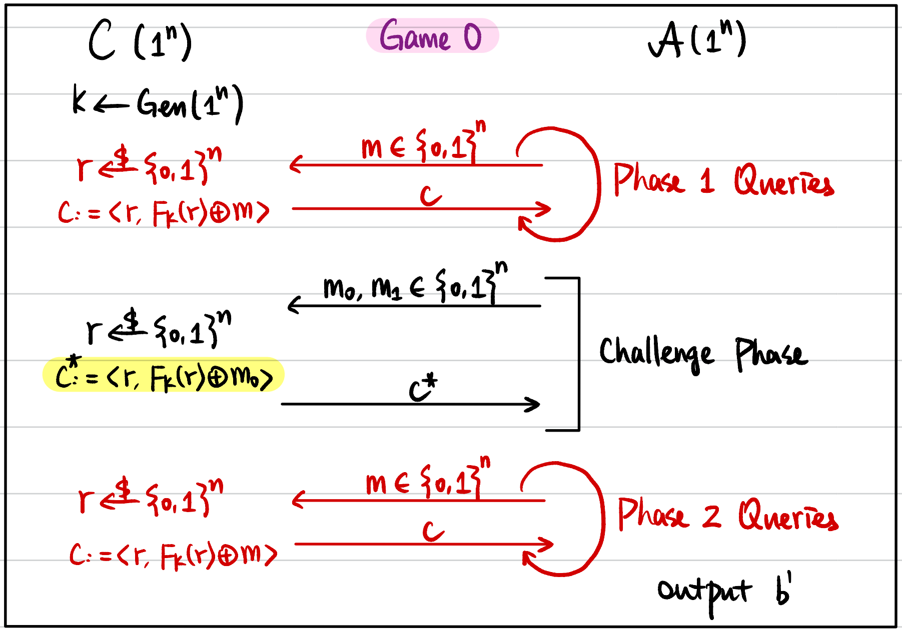
  output $m := F_k(r) \oplus s$

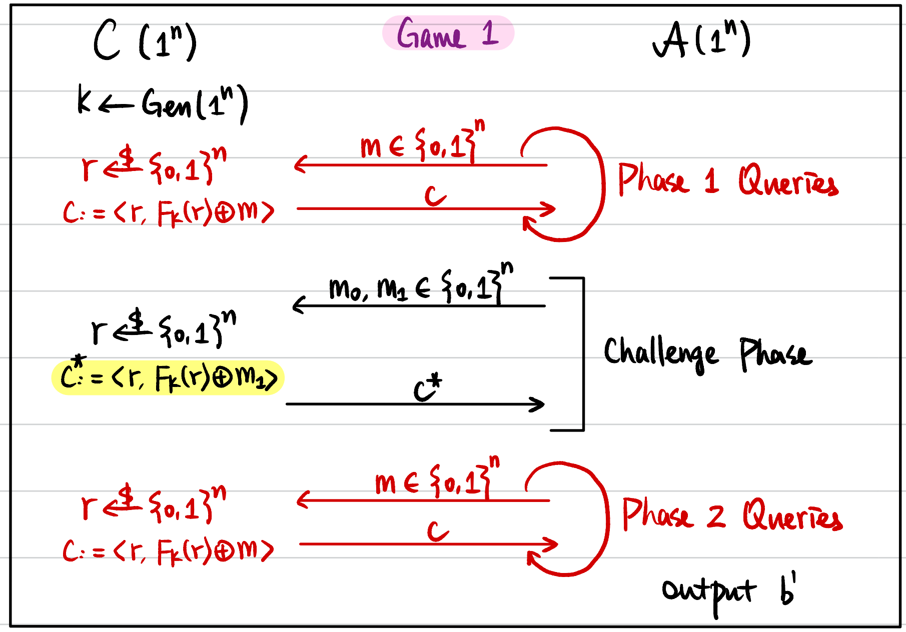


Theorem If F is a PRF, then $\Pi = (Gen, Enc, Dec)$ is CPA-Secure.

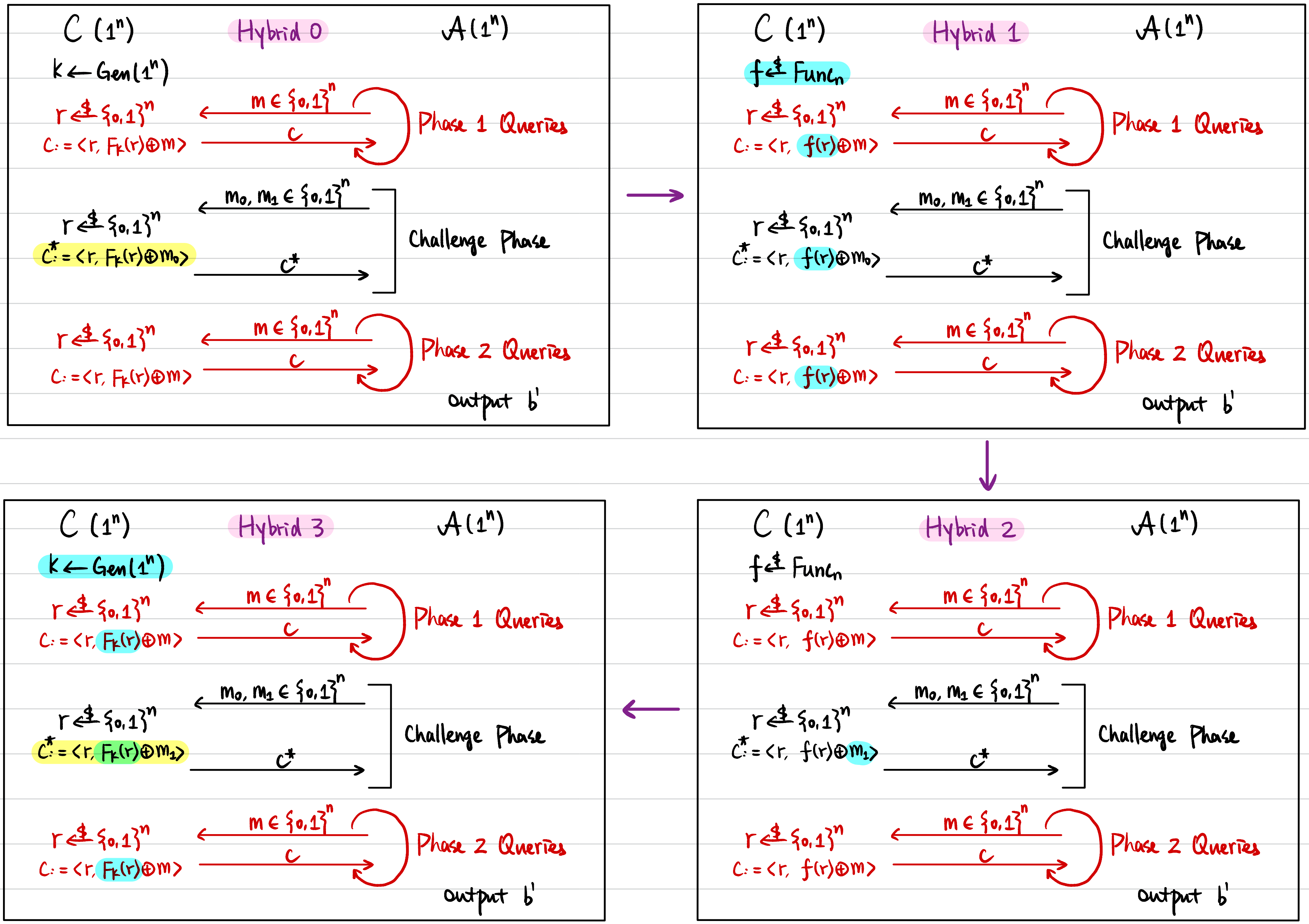
Theorem If F is a PRF, then $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is CPA-Secure.

Proof $\forall$ PPT $\mathcal{A}$,





$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Game } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Game } 1] \right| \le \mathrm{negl}(n) \text{ ?}$$

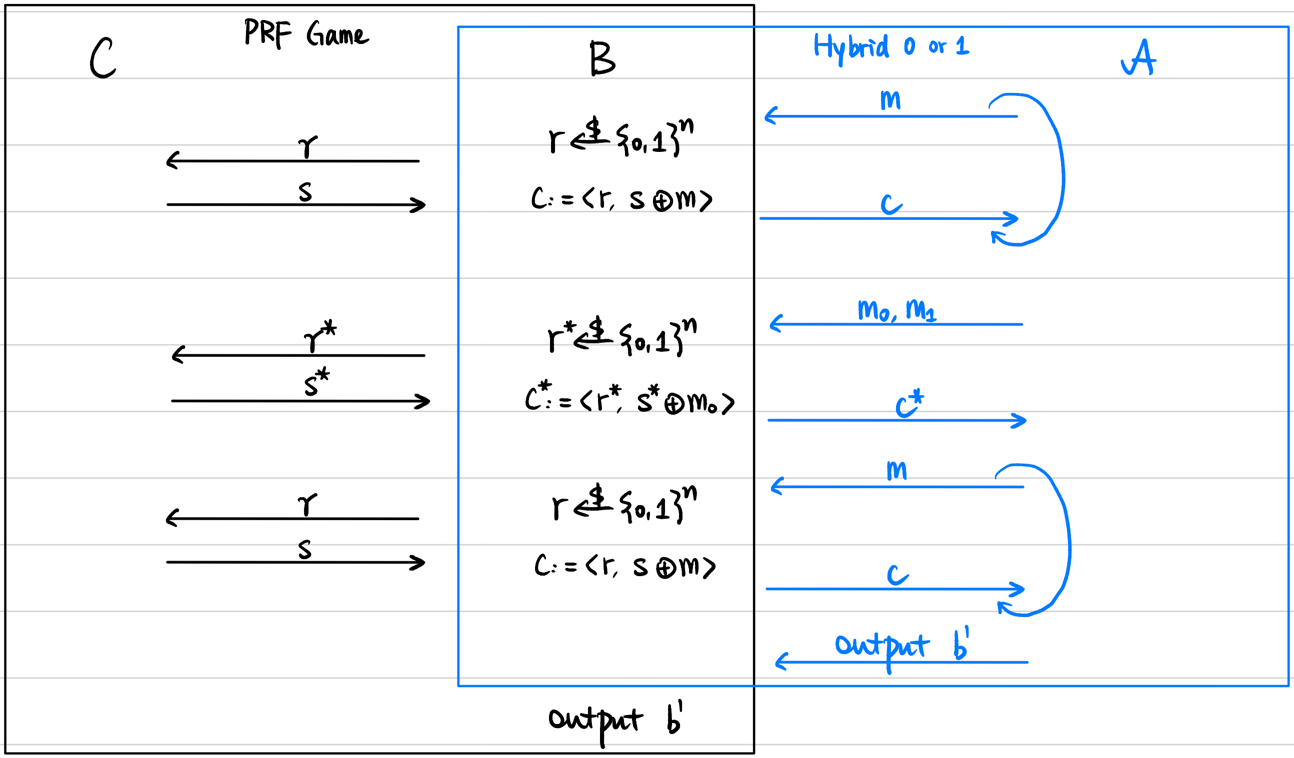


$\mathcal{C}(1^n)$
Hybrid 0
$\mathcal{A}(1^n)$
$k \leftarrow Gen(1^n)$
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, F_k(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 1 Queries
$m_0, m_1 \in \{0,1\}^n$
$r \xleftarrow{\$} \{0,1\}^n$
$c^* := \langle r, F_k(r) \oplus m_0 \rangle$
c^*
Challenge Phase
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, F_k(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 2 Queries
output b'
$\mathcal{C}(1^n)$
Hybrid 1
$\mathcal{A}(1^n)$
$f \xleftarrow{\$} Func_n$
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, f(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 1 Queries
$m_0, m_1 \in \{0,1\}^n$
$r \xleftarrow{\$} \{0,1\}^n$
$c^* := \langle r, f(r) \oplus m_0 \rangle$
c^*
Challenge Phase
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, f(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 2 Queries
output b'
$\mathcal{C}(1^n)$
Hybrid 3
$\mathcal{A}(1^n)$
$k \leftarrow Gen(1^n)$
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, F_k(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 1 Queries
$m_0, m_1 \in \{0,1\}^n$
$r \xleftarrow{\$} \{0,1\}^n$
$c^* := \langle r, F_k(r) \oplus m_1 \rangle$
c^*
Challenge Phase
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, F_k(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 2 Queries
output b'
$\mathcal{C}(1^n)$
Hybrid 2
$\mathcal{A}(1^n)$
$f \xleftarrow{\$} Func_n$
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, f(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 1 Queries
$m_0, m_1 \in \{0,1\}^n$
$r \xleftarrow{\$} \{0,1\}^n$
$c^* := \langle r, f(r) \oplus m_1 \rangle$
c^*
Challenge Phase
$r \xleftarrow{\$} \{0,1\}^n$
$c := \langle r, f(r) \oplus m \rangle$
$m \in \{0,1\}^n$
c
Phase 2 Queries
output b'

<u>Lemma 1</u> $\forall$ PPT $\mathcal{A}$, $\exists$ negligible function $\varepsilon_1(\cdot)$ s.t.

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] \right| \leq \varepsilon_1(n)$$

<u>Proof</u> Assume not, then $\exists$ PPT $\mathcal{A}$ that distinguishes Hybrid 0 & Hybrid 1.
We construct PPT B to break the pseudorandomness of F.



If C uses F_k, then $\mathcal{A}$ is in Hybrid 0

If C uses a random function, then $\mathcal{A}$ is in Hybrid 1

$$\left| \Pr[B \text{ outputs } 1 \text{ on } F_k] - \Pr[B \text{ outputs } 1 \text{ on a random function}] \right|$$

$$= \left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] \right| \geq \text{non-negl}(n).$$

<u>Lemma 2</u> $\forall$ PPT $\mathcal{A}$, $\exists$ negligible function $\varepsilon_2(\cdot)$ s.t.

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] \right| \leq \varepsilon_2(n).$$

<u>Proof</u> Let r_i be the r value sampled for the i-th query in Phase 1 & 2.

Let r^* be the r value sampled for the Challenge Phase.

$\forall i$, $\Pr[r_i = r^*] = 2^{-n}$.

$\Pr[\exists i \text{ s.t. } r_i = r^*] \leq 2^{-n} \cdot Q$ $\quad\quad Q :=$ total # queries in Phase 1 & 2.

$\mathcal{A}$ is PPT $\Rightarrow$ $Q \leq p(n)$ polynomial

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] \right|$$

$\leq \Pr[\exists i \text{ s.t. } r_i = r^*] \leq 2^{-n} \cdot p(n) \rightarrow$ negligible

<u>Lemma 3</u> $\forall$ PPT $\mathcal{A}$, $\exists$ negligible function $\varepsilon_3(\cdot)$ s.t.

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 3] \right| \leq \varepsilon_3(n)$$

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Game } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Game } 1] \right|$$

$$= \big| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] +$$

$$\Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] +$$

$$\Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 3] \big|$$

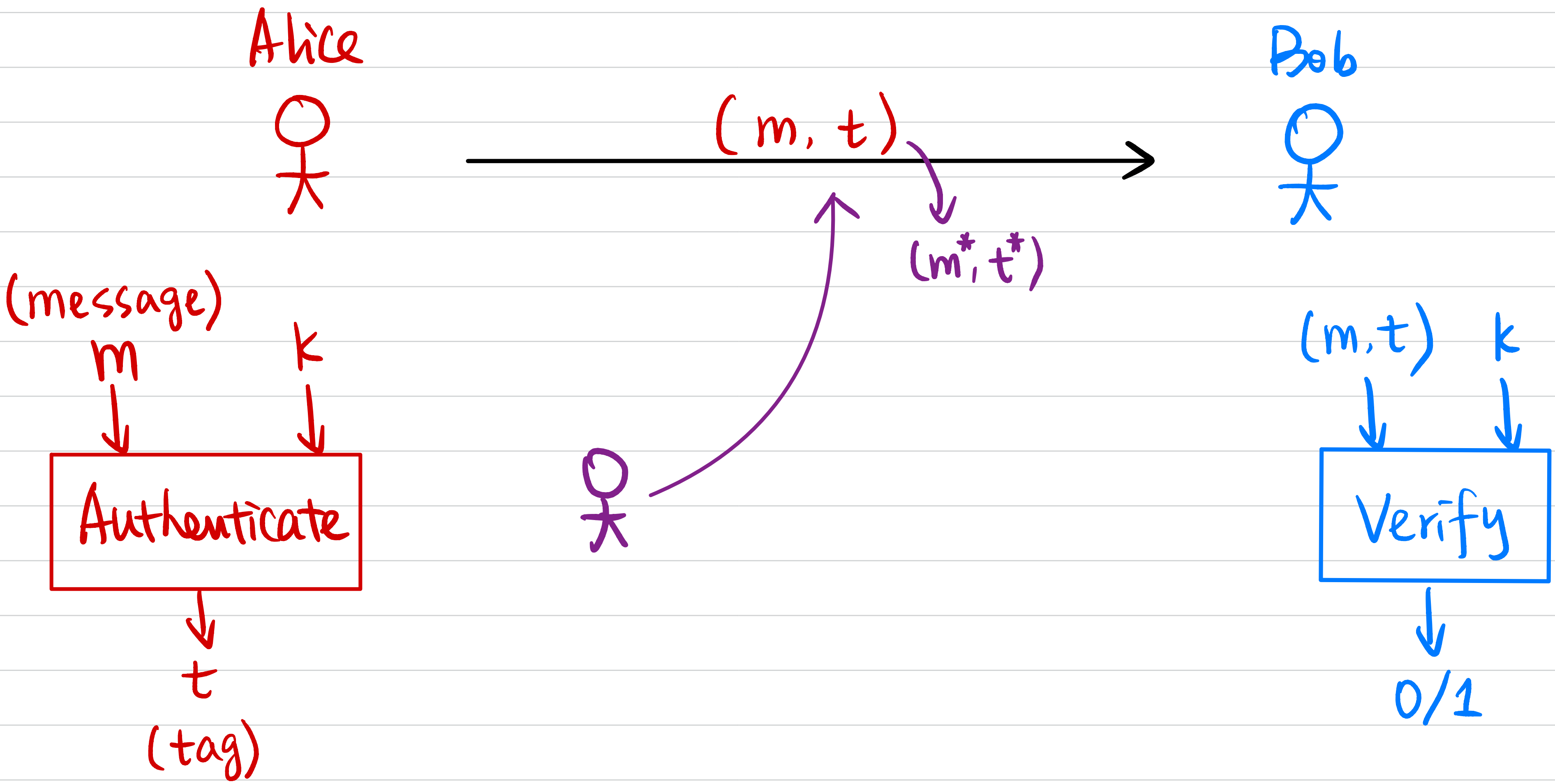
$$\leq \left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 0] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] \right| +$$

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 1] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] \right| +$$

$$\left| \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 2] - \Pr[\mathcal{A} \text{ outputs } 1 \text{ in Hybrid } 3] \right|$$

$$\leq \varepsilon_1(n) + \varepsilon_2(n) + \varepsilon_3(n) \rightarrow \text{negligible}$$

Message Integrity



Message Integrity vs. Secrecy

Does encryption solve the problem? $t \leftarrow Enc_k(m)$

- OTP?

$t = k \oplus m$

$(m, t) \Rightarrow k \Rightarrow (m^*, t^*)$

- Pseudo OTP?

$t = G(k) \oplus m$

$(m, t) \Rightarrow G(k) \Rightarrow (m^*, t^*)$

- CPA-secure encryption from PRF?

$t = \langle r, F_k(r) \oplus m \rangle$

$(m, t) \Rightarrow F_k(r) \Rightarrow (m^*, t^*)$

$t^* = \langle r, F_k(r) \oplus m^* \rangle$