Public-Key Watermarking Schemes for Pseudorandom Functions

Rupeng Yang

HKU ⇒ UOW

Zuoxia Yu

HKU

Man Ho Au

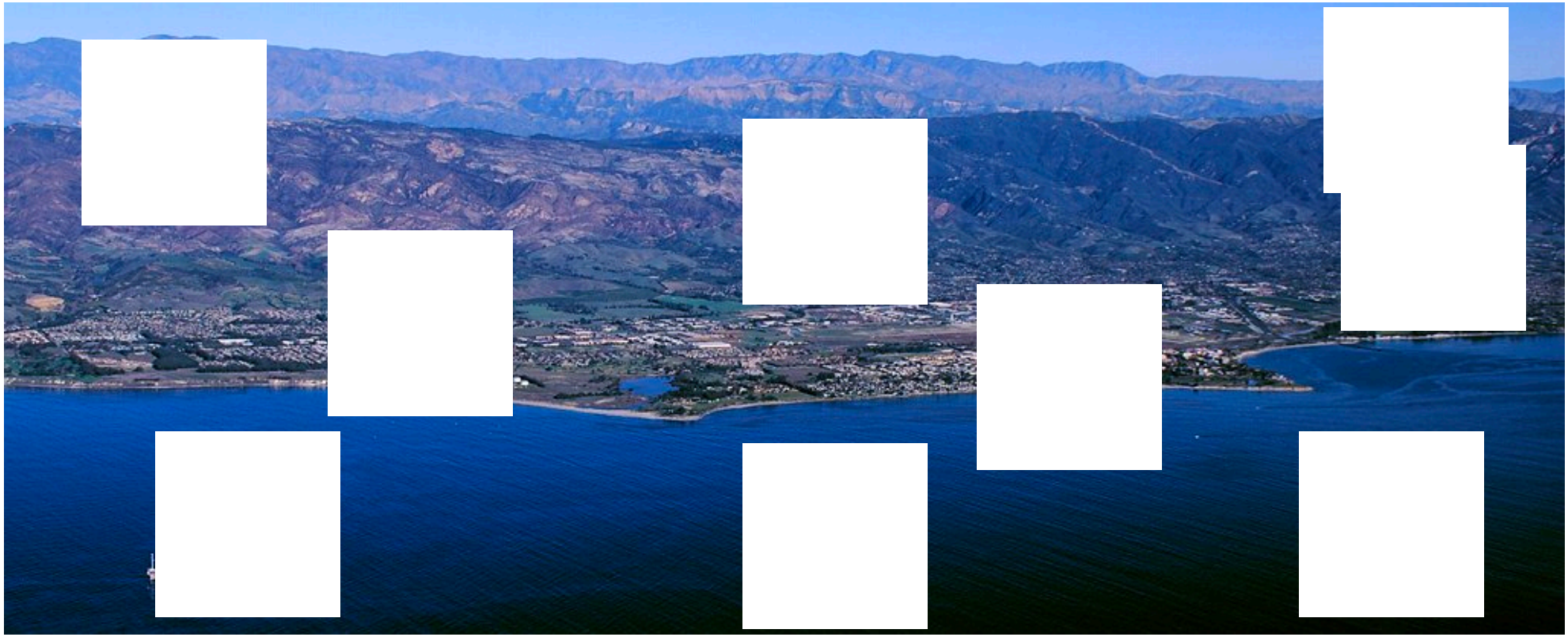
HKU

Willy Susilo

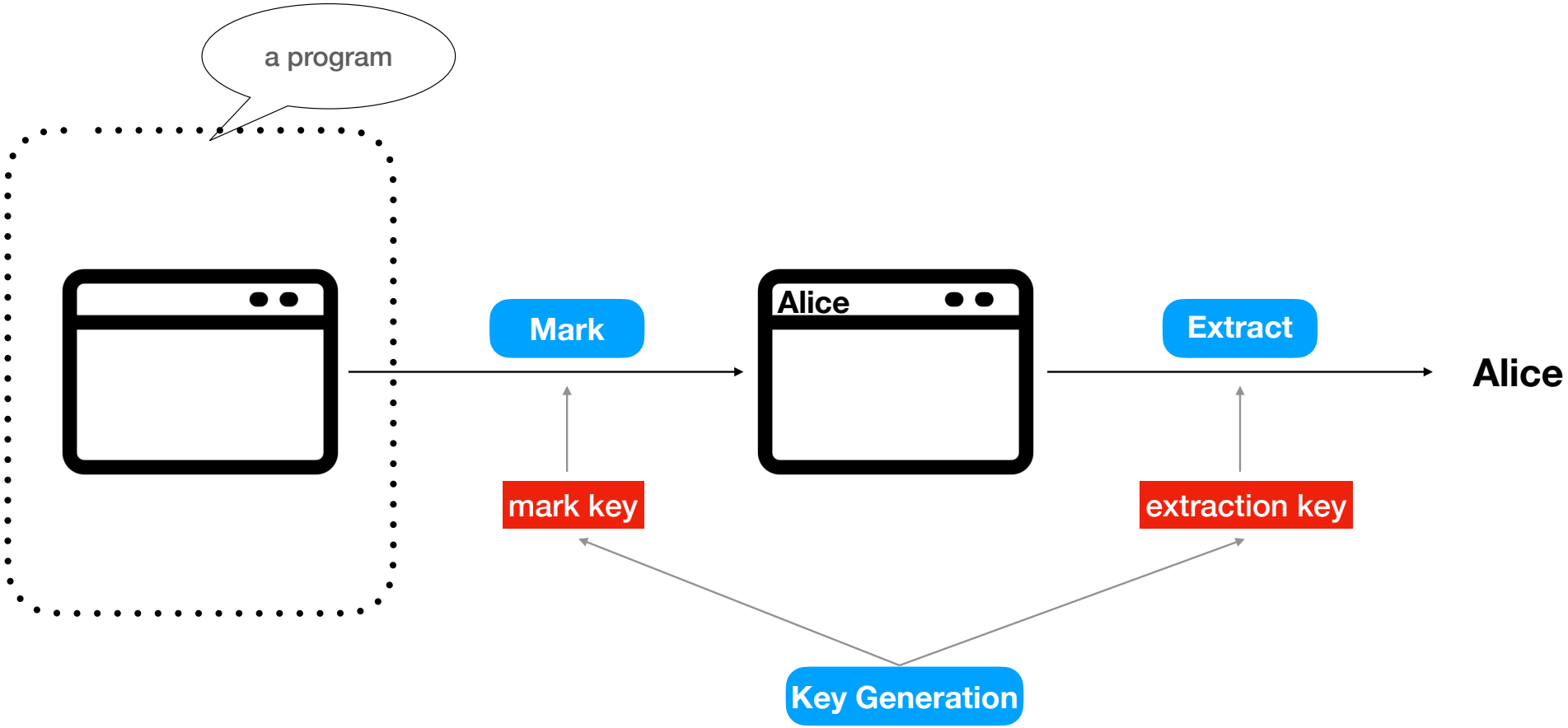
UOW

Watermarking A Cryptographic Program

Watermarking A Cryptographic Program



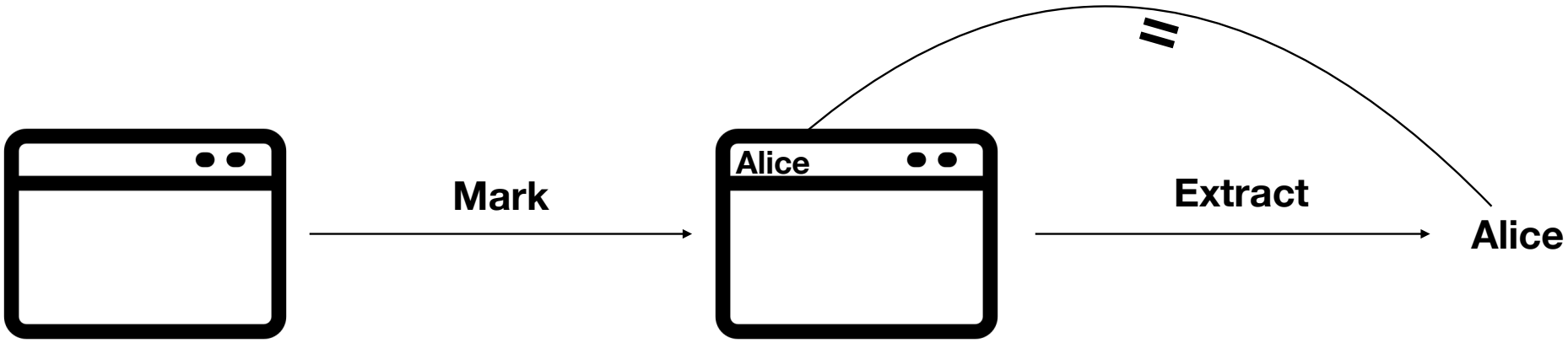
Watermarking A Cryptographic Program



Watermarking A Cryptographic Program

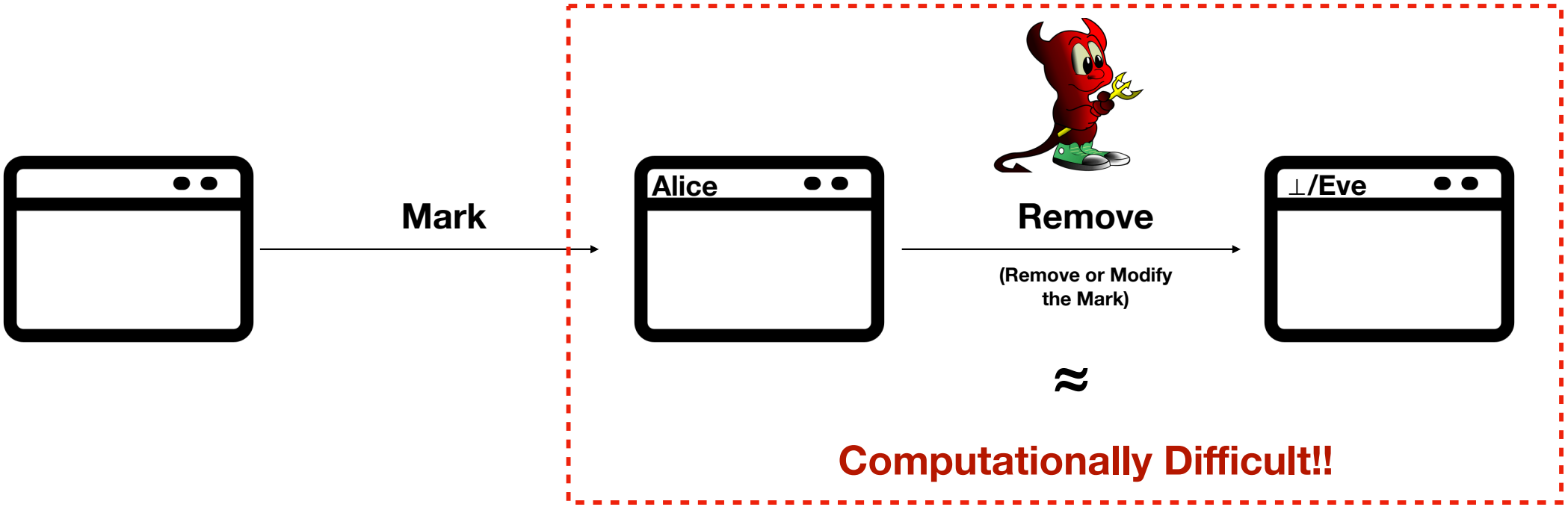
Correctness Requirement: Functionality Preserving

Watermarking A Cryptographic Program



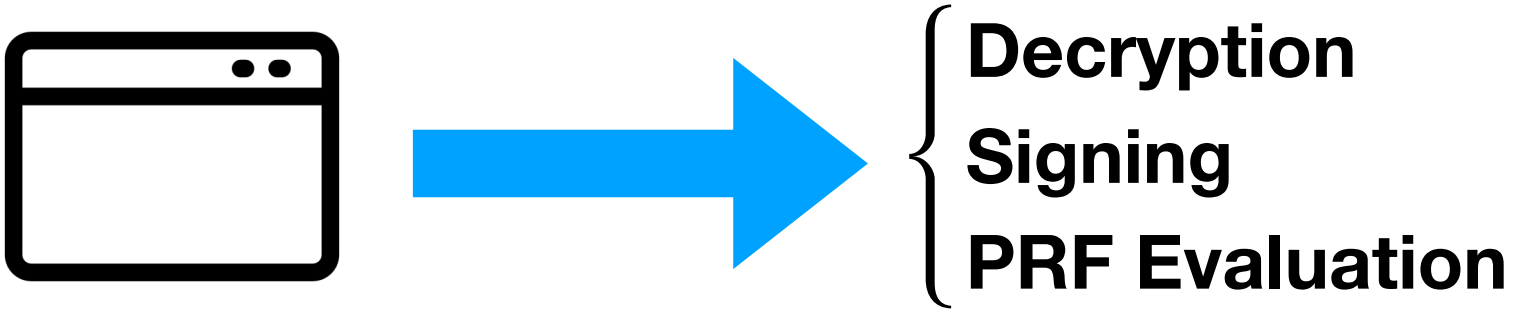
Correctness Requirement: Extraction Correctness

Watermarking A Cryptographic Program



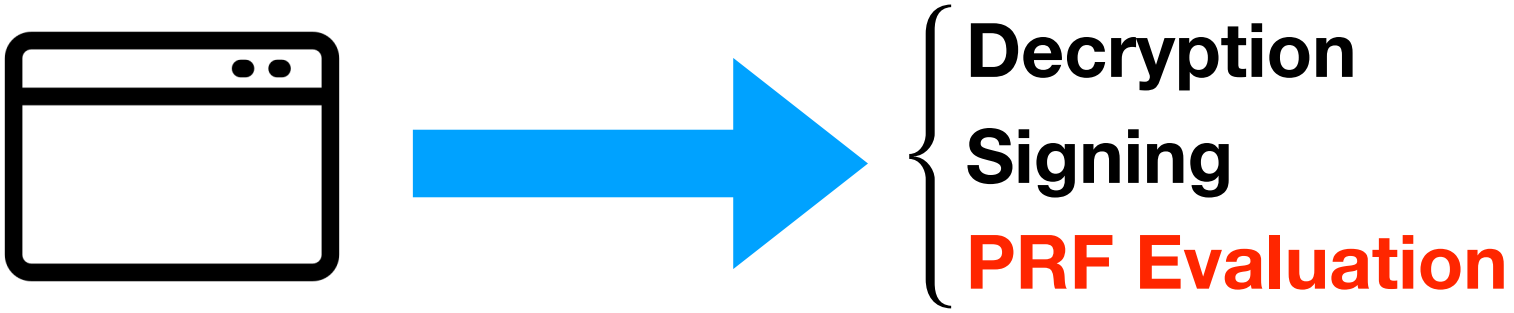
Security Requirement: Unremovability

Watermarking A Cryptographic Program



It is *impossible* to watermark a learnable functionality.

Watermarking A Cryptographic Program



It is *impossible* to watermark a learnable functionality.

Security Definitions of Watermarkable PRF

$(MK, EK) \leftarrow \texttt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$

→ $\texttt{C}^*$

← $\tilde{\texttt{C}}$

The adversary wins if:

1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

$(MK, EK) \leftarrow \texttt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathsf{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←

C^* →

$\tilde{\mathsf{C}}$ ←

The adversary wins if:

1. $\mathsf{C}^* \approx \tilde{\mathsf{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security

$(MK, EK) \leftarrow \texttt{KeyGen}$

$K \leftarrow \mathscr{K}$

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$ → $\texttt{C}^*$

$\tilde{\texttt{C}}$ ←

The adversary wins if:

1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security

$(MK, EK) \leftarrow \texttt{KeyGen}$

$K \leftarrow \mathscr{K}$

MK →

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$ → $\texttt{C}^*$

$\tilde{\texttt{C}}$ ←

The adversary wins if:

1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security

$(MK, EK) \leftarrow \texttt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$ → $\texttt{C}^*$

$\tilde{\texttt{C}}$ ←

The adversary wins if:
1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security

$(MK, EK) \leftarrow \texttt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK →

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$ → $\texttt{C}^*$

$\tilde{\texttt{C}}$ ←

The adversary wins if:
1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security

$(MK, EK) \leftarrow \texttt{KeyGen}$
$K \leftarrow \mathscr{K}$

EK →

$\mathfrak{m}$ ←

$\texttt{C}^* \leftarrow \texttt{Mark}(MK, K, \mathfrak{m})$ → $\texttt{C}^*$

$\tilde{\texttt{C}}$ ←

The adversary wins if:
1. $\texttt{C}^* \approx \tilde{\texttt{C}}$
2. $\texttt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Extraction Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Extraction Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK, EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Key Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Extraction Security

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK, EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Key Security

👍: **No authority (holding secret mark key and/or extraction key) is needed.**

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security: [BLW17,KW17]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security: [QWZ18,KW19,YAYX20]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Extraction Security: [CHN+16,YAL+19]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK, EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Key Security: ???

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Secret-Key Security: [BLW17,KW17]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Marking Security: [QWZ18,KW19,YAYX20]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Extraction Security: [CHN+16,YAL+19]

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK, EK →

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

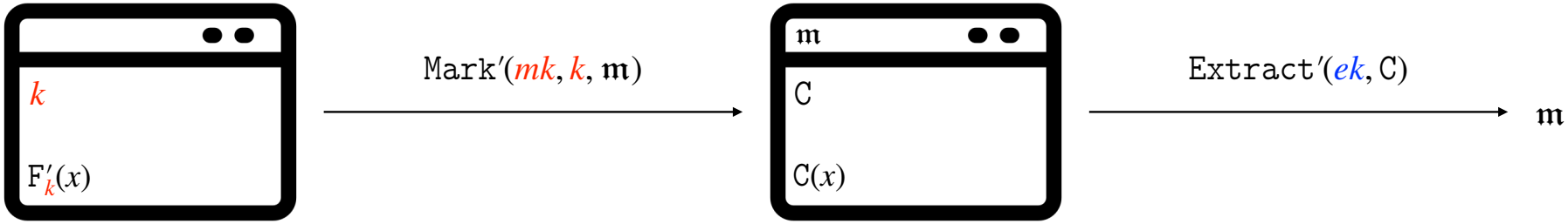
$\mathfrak{m}$ ←
$\mathtt{C}^*$ →
$\tilde{\mathtt{C}}$ ←

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

Public-Key Security: This Work

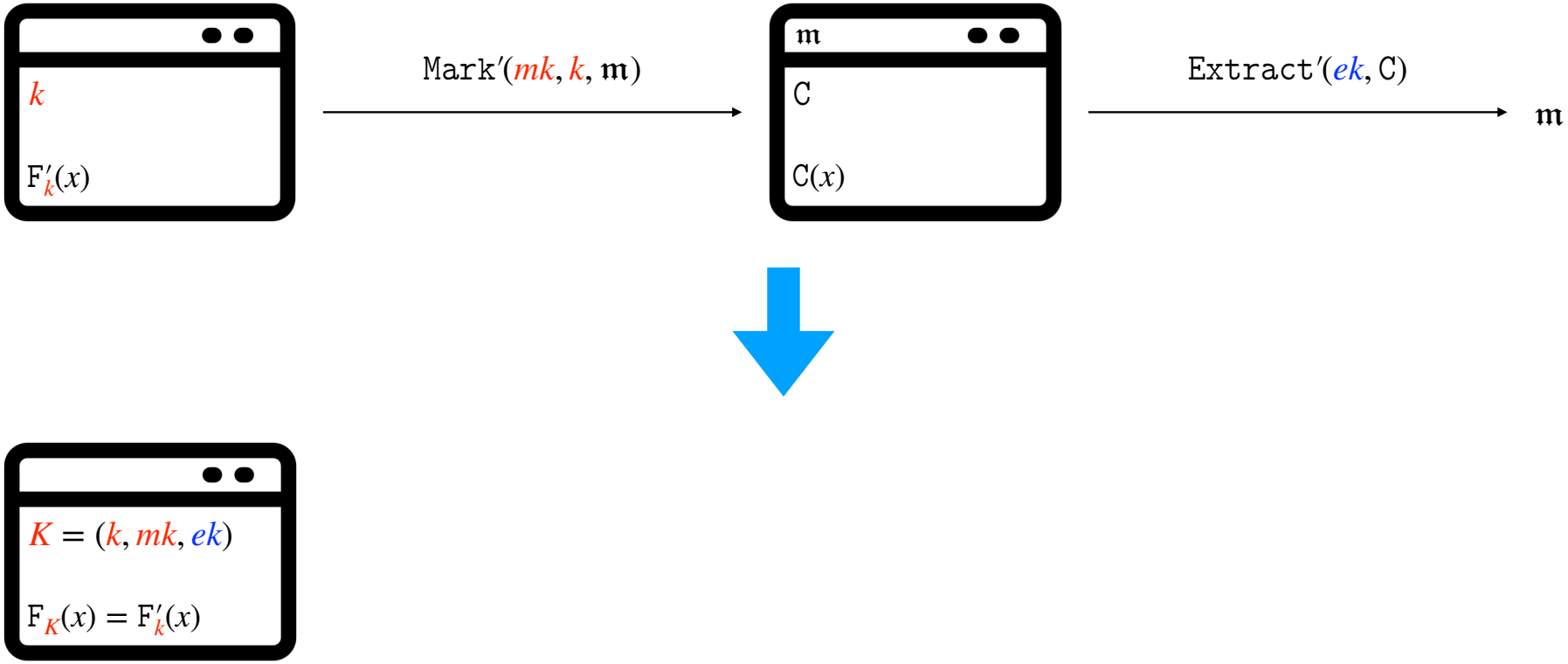
Constructing Public-Key Watermarkable PRF

A Watermarkable PRF with Public-Extraction Security:



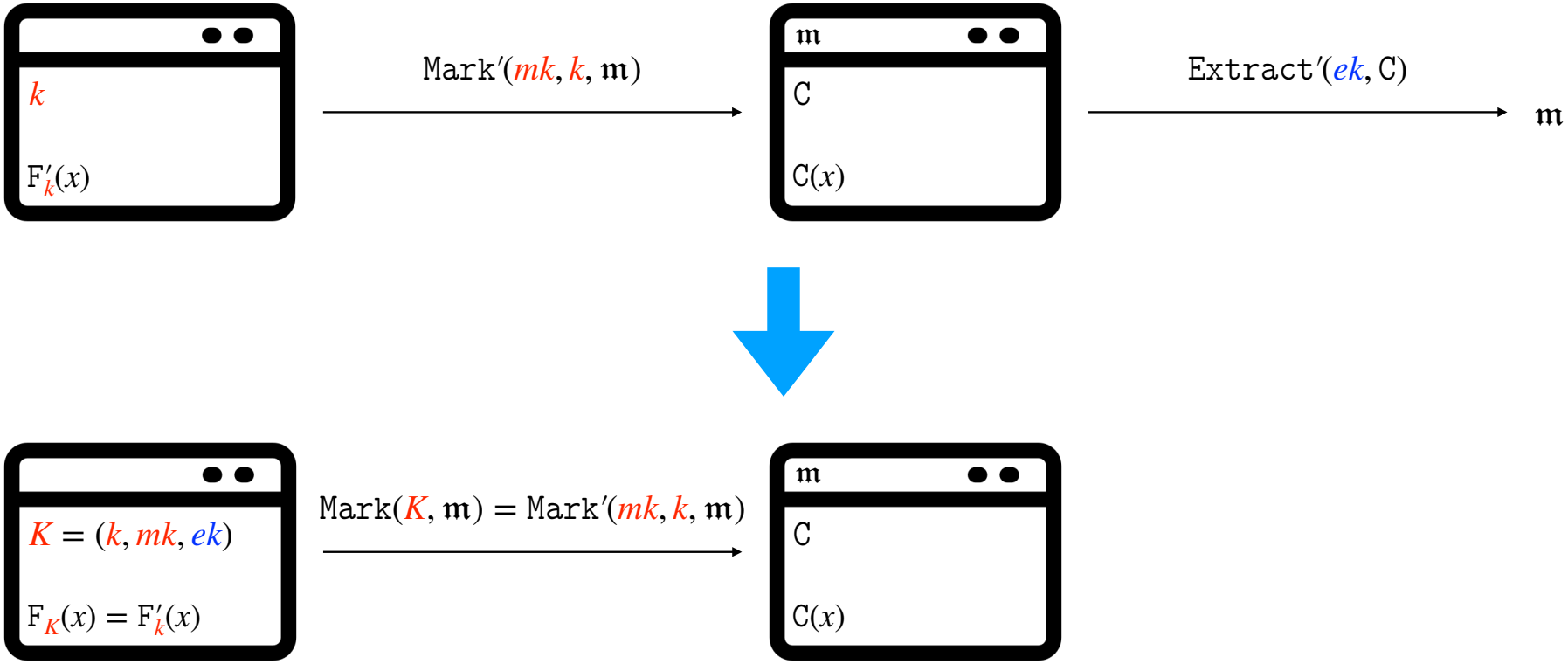
Constructing Public-Key Watermarkable PRF

A Watermarkable PRF with Public-Extraction Security:



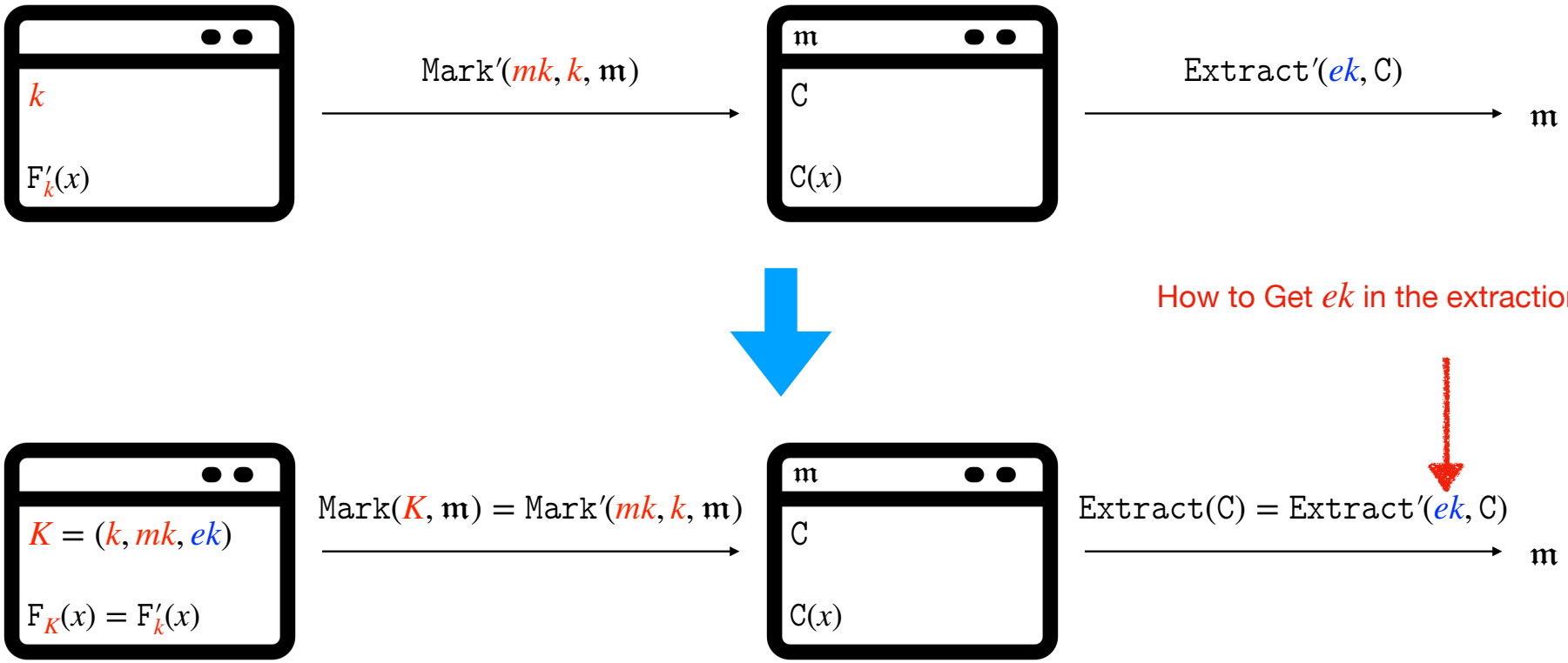
Constructing Public-Key Watermarkable PRF

A Watermarkable PRF with Public-Extraction Security:



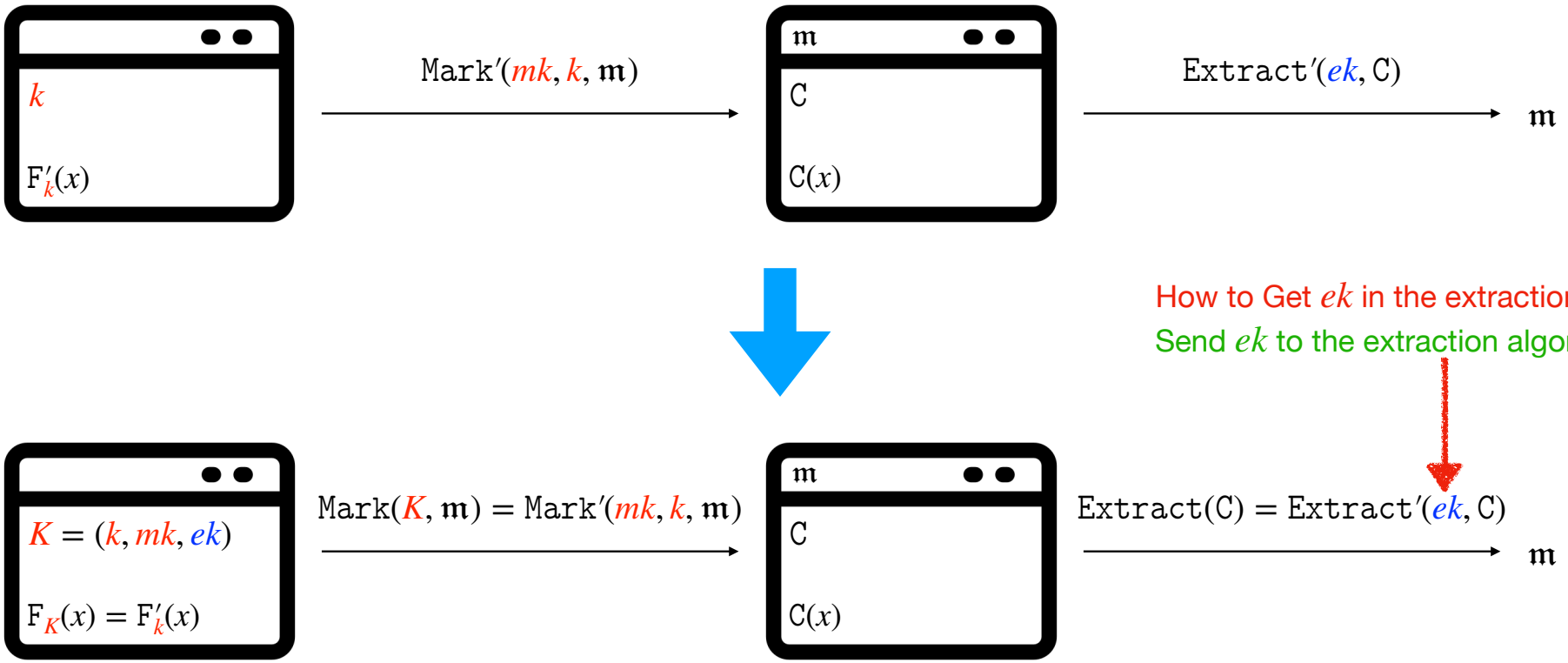
Constructing Public-Key Watermarkable PRF

A Watermarkable PRF with Public-Extraction Security:

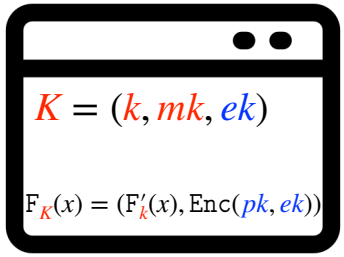


Constructing Public-Key Watermarkable PRF

A Watermarkable PRF with Public-Extraction Security:

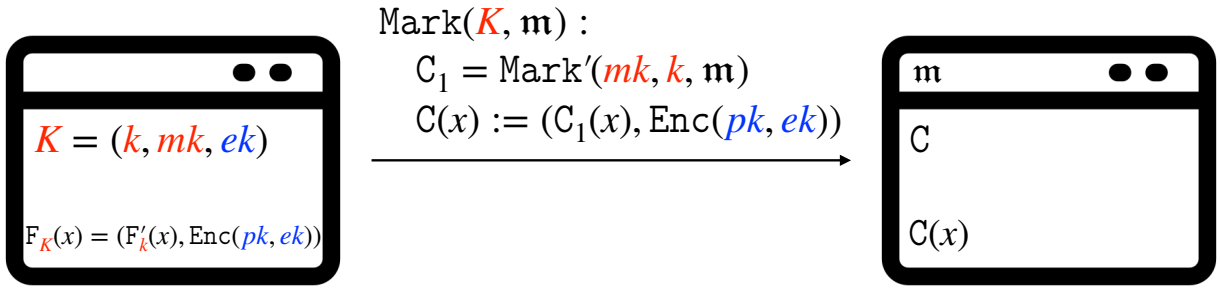


Constructing Public-Key Watermarkable PRF



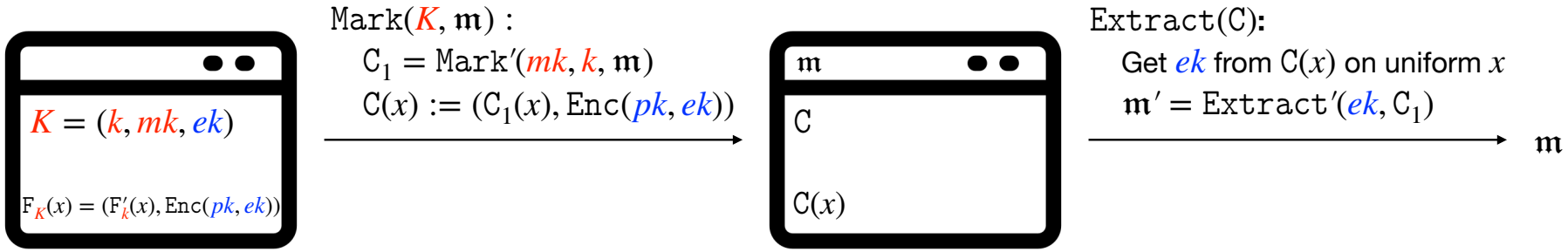
(pk, sk) is a key pair of a PKE scheme and is included in the public parameter of the watermarking scheme.

Constructing Public-Key Watermarkable PRF



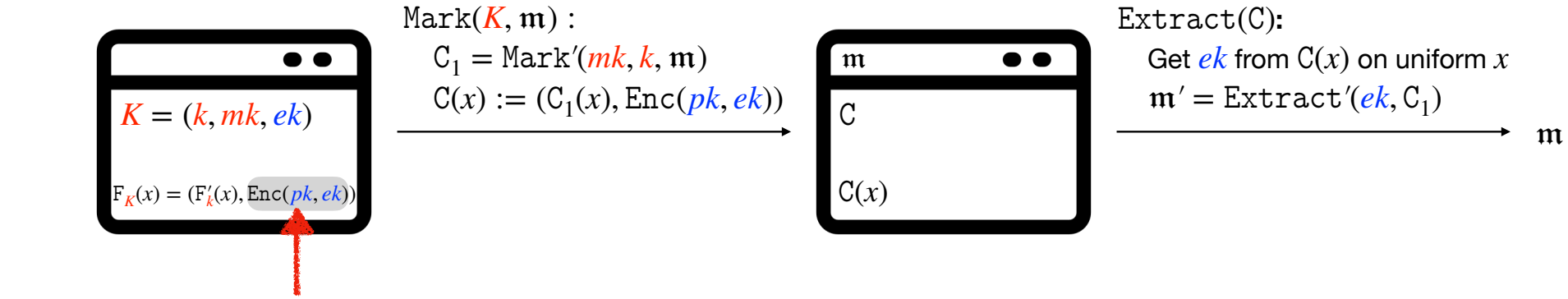
(pk, sk) is a key pair of a PKE scheme and is included in the public parameter of the watermarking scheme.

Constructing Public-Key Watermarkable PRF



(pk, sk) is a key pair of a PKE scheme and is included in the public parameter of the watermarking scheme.

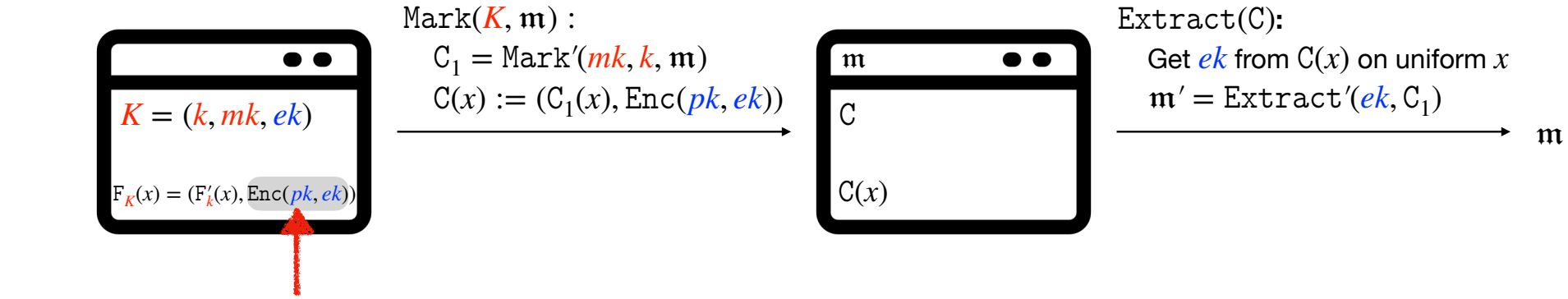
Constructing Public-Key Watermarkable PRF



Distinguishable from random string given sk !

(pk, sk) is a key pair of a PKE scheme and is included in the public parameter of the watermarking scheme.

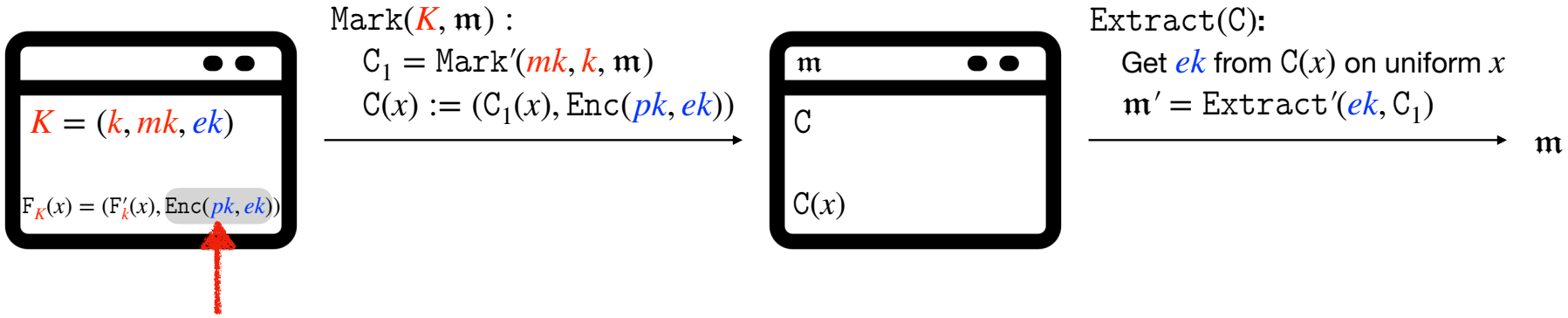
Constructing Public-Key Watermarkable PRF



Distinguishable from random string given sk !

Use robust unobfuscatable PRF instead of PKE!

Constructing Public-Key Watermarkable PRF



Distinguishable from random string given sk !
Use robust unobfuscatable PRF instead of PKE!

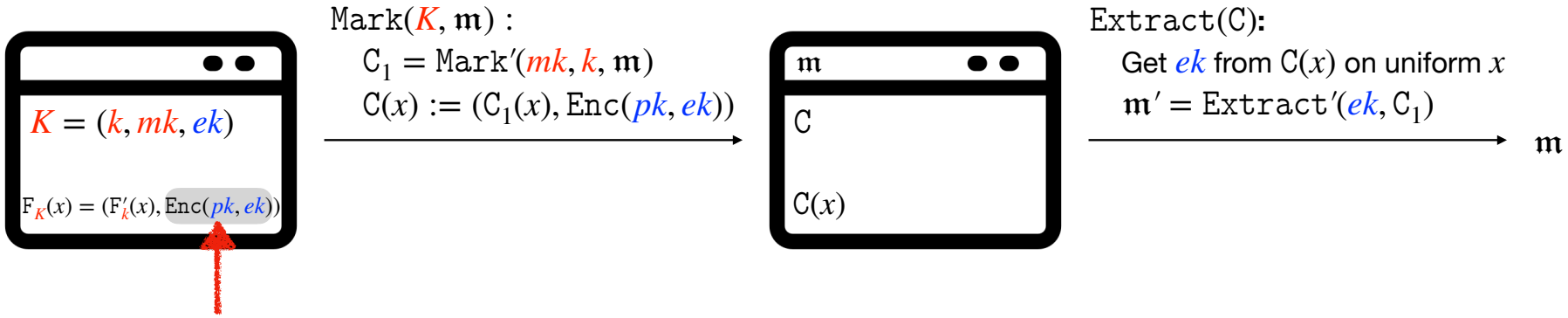
A PRF $\mathrm{UF}_{k_s}(\cdot)$

A PRF key k_s is associated with a secret s.

Pseudorandomness: $\mathrm{UF}_{k_s}(\cdot)$ is pseudorandom given oracle access to it.

Learnability: It is easy to get the secret s given a circuit $C(\cdot) \equiv \mathrm{UF}_{k_s}(\cdot)$.

Constructing Public-Key Watermarkable PRF



Distinguishable from random string given sk !
Use robust unobfuscatable PRF instead of PKE!

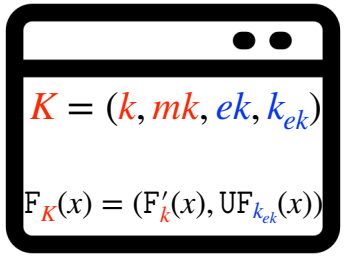
A PRF $\mathtt{UF}_{k_s}(\cdot)$

A PRF key k_s is associated with a secret s.

Pseudorandomness: $\mathtt{UF}_{k_s}(\cdot)$ is pseudorandom given oracle access to it.

Robust Learnability: It is easy to get the secret s given a circuit $C(\cdot) \approx \mathtt{UF}_{k_s}(\cdot)$.

Constructing Public-Key Watermarkable PRF



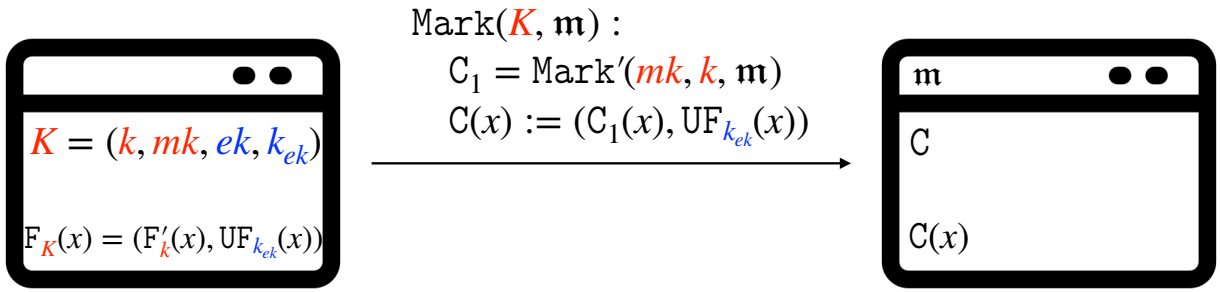
A PRF $\mathtt{UF}_{k_s}(\cdot)$

A PRF key k_s is associated with a secret s.

Pseudorandomness: $\mathtt{UF}_{k_s}(\cdot)$ is pseudorandom given oracle access to it.

Robust Learnability: It is easy to get the secret s given a circuit $C(\cdot) \approx \mathtt{UF}_{k_s}(\cdot)$.

Constructing Public-Key Watermarkable PRF



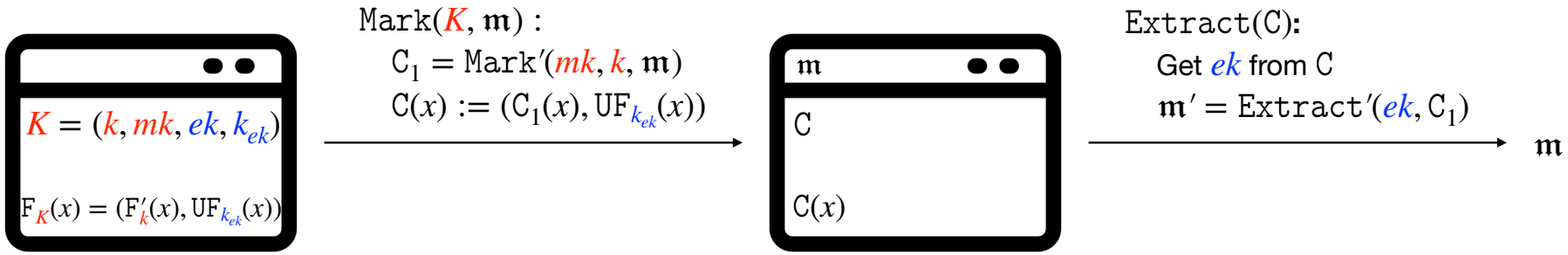
A PRF $\mathsf{UF}_{k_s}(\cdot)$

A PRF key k_s is associated with a secret s.

Pseudorandomness: $\mathsf{UF}_{k_s}(\cdot)$ is pseudorandom given oracle access to it.

Robust Learnability: It is easy to get the secret s given a circuit $C(\cdot) \approx \mathsf{UF}_{k_s}(\cdot)$.

Constructing Public-Key Watermarkable PRF



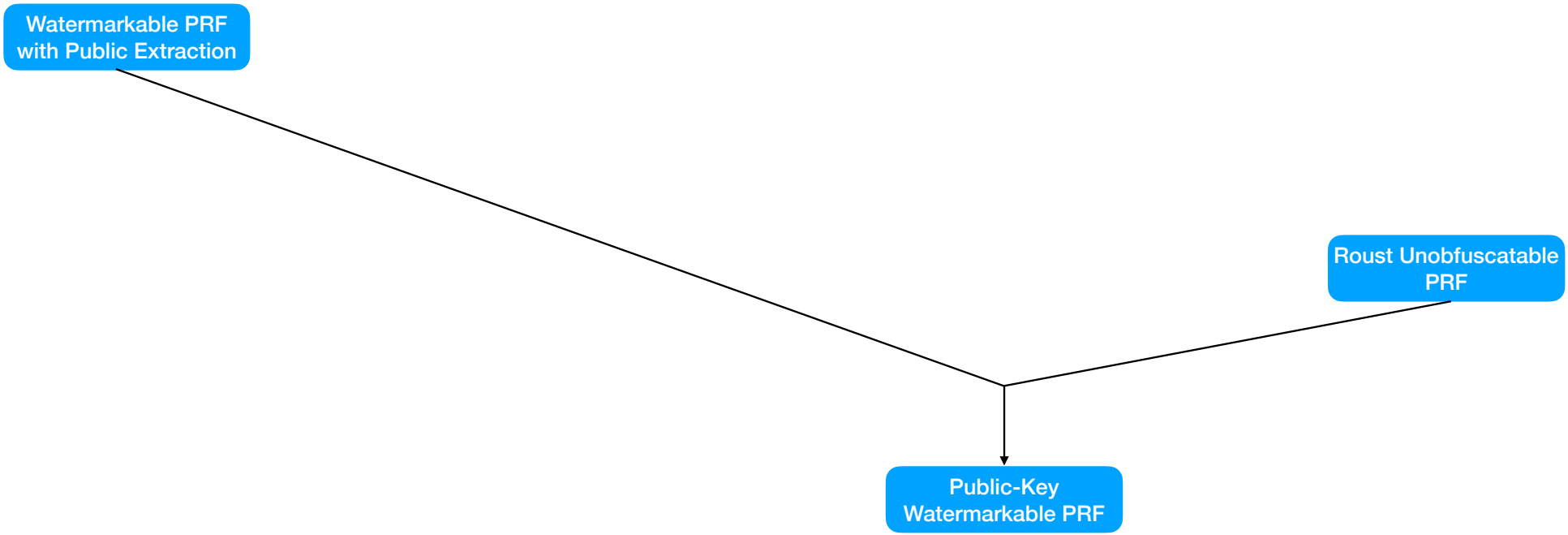
A PRF $\mathsf{UF}_{k_s}(\cdot)$

A PRF key k_s is associated with a secret s.

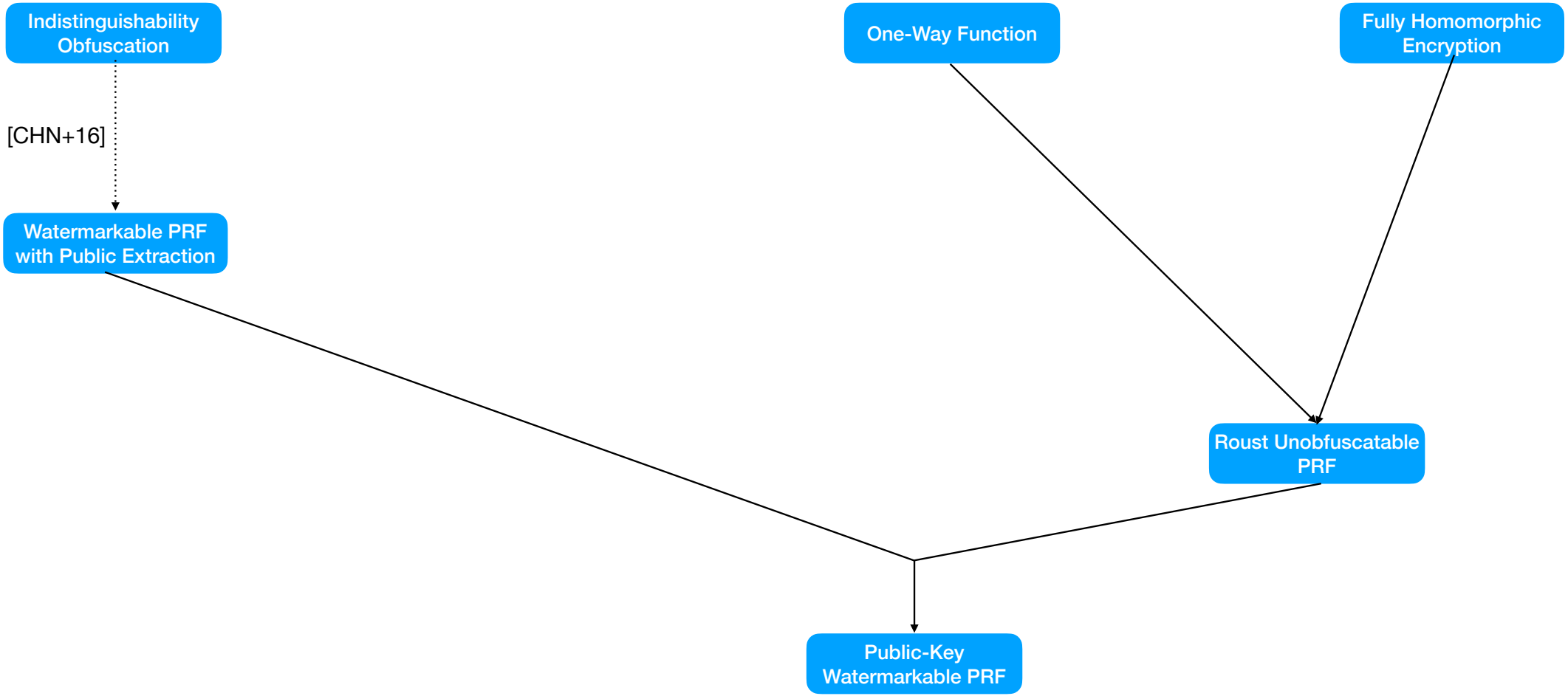
Pseudorandomness: $\mathsf{UF}_{k_s}(\cdot)$ is pseudorandom given oracle access to it.

Robust Learnability: It is easy to get the secret s given a circuit $C(\cdot) \approx \mathsf{UF}_{k_s}(\cdot)$.

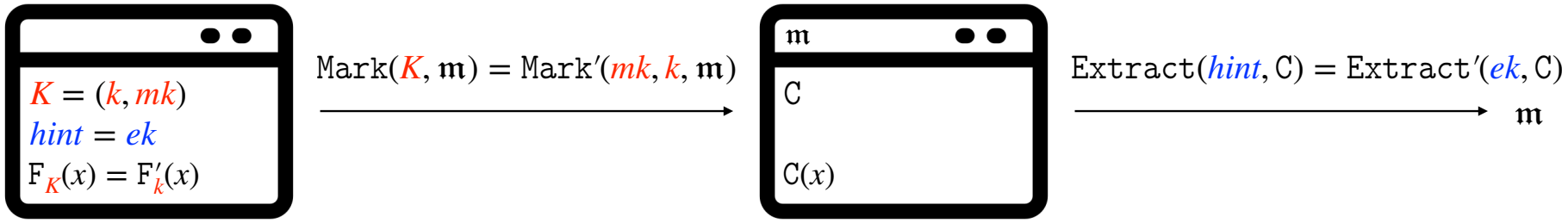
Instantiating Public-Key Watermarkable PRF



Instantiating Public-Key Watermarkable PRF



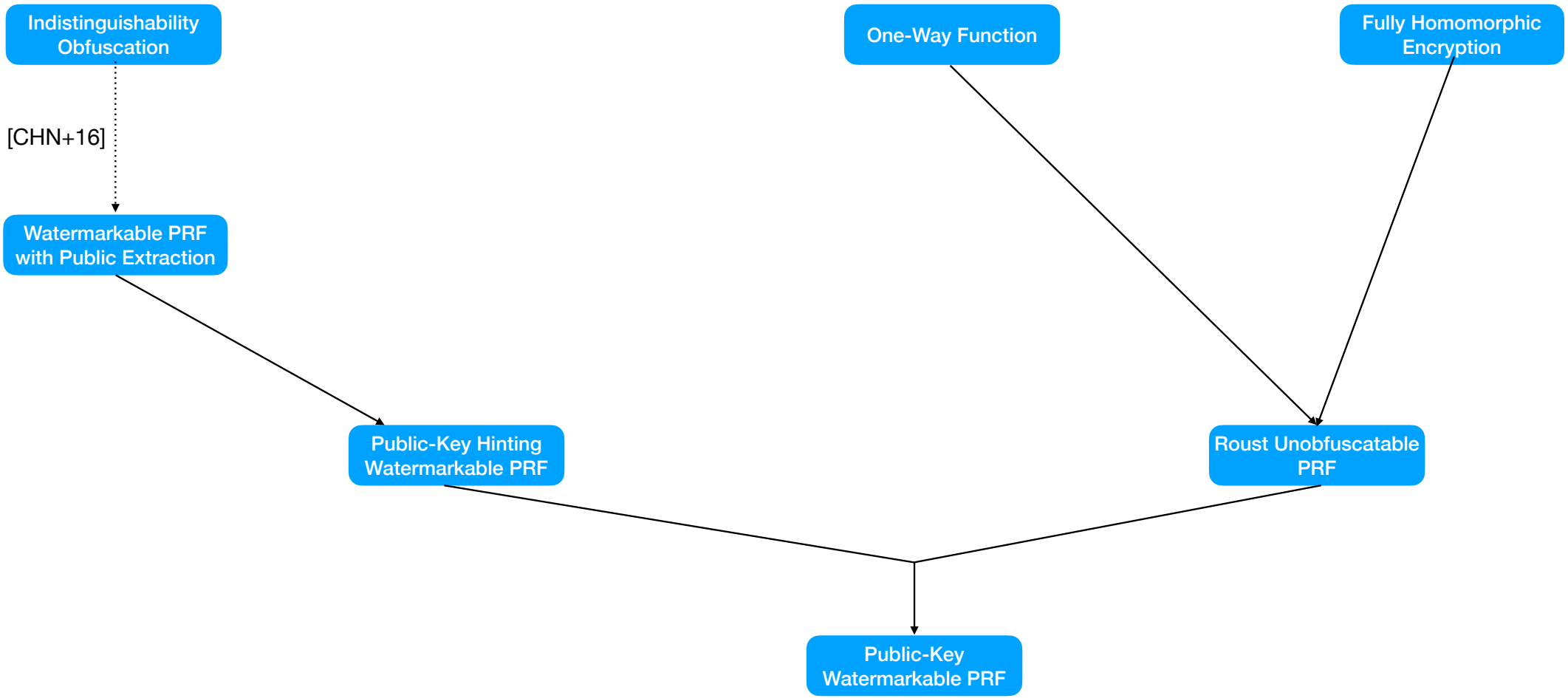
Hinting Watermarkable PRF



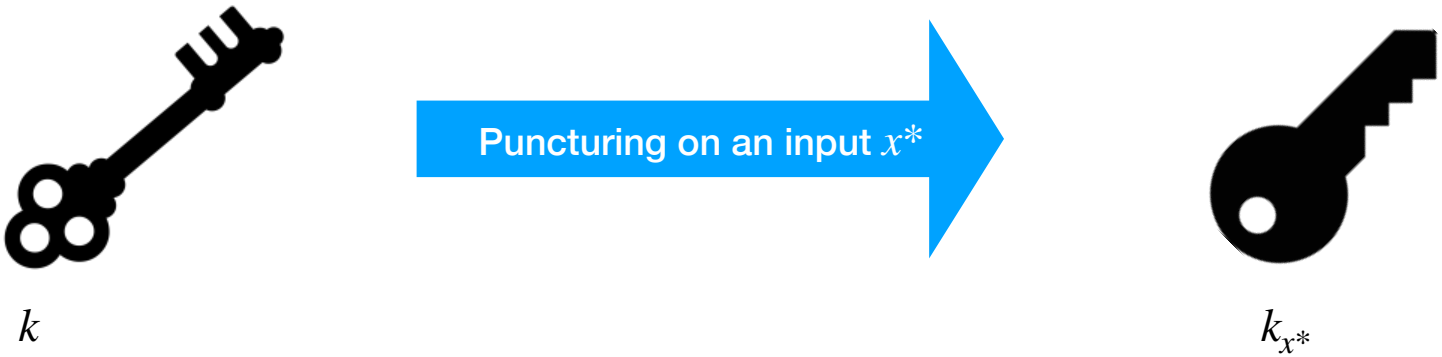
Hinting Watermarkable PRF:

A hint associated with the PRF key can be used in the extraction algorithm.

Instantiating Public-Key Watermarkable PRF

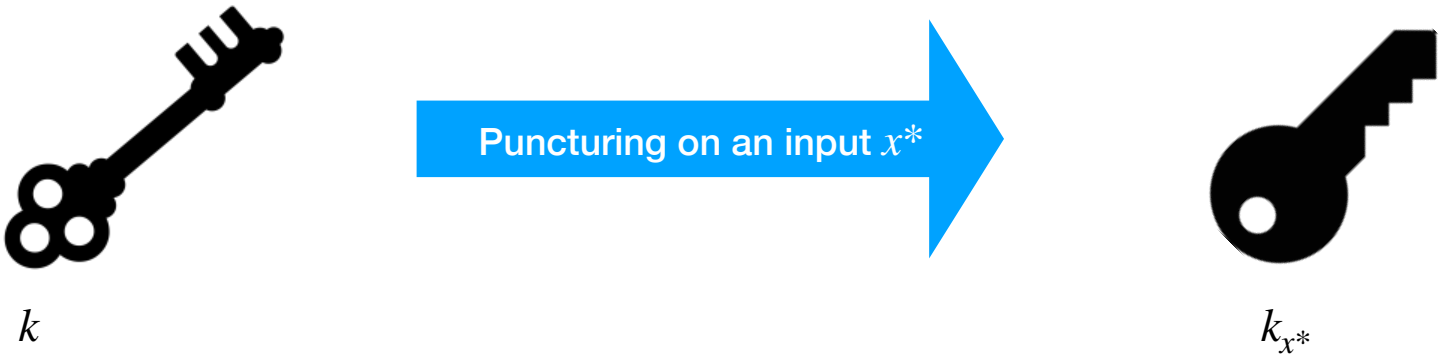


Constructing Public-Key Hinting Watermarkable PRF from Puncturable PRF



Correctness: if $x \neq x^*, \mathsf{F}_k(x) = F_{k_{x^*}}(x)$

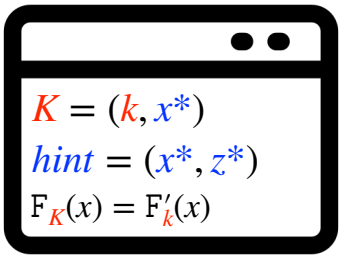
Constructing Public-Key Hinting Watermarkable PRF from Puncturable PRF



Correctness: if $x \neq x^*, \mathsf{F}_k(x) = F_{k_{x^*}}(x)$

Pseudorandomness: $\mathsf{F}_k(x^*)$ **is hidden given** k_{x^*}

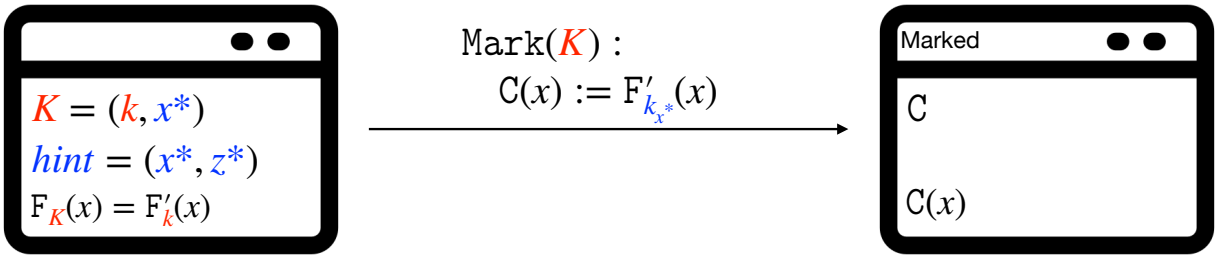
Constructing Public-Key Hinting Watermarkable PRF from Puncturable PRF



k is a PRF key of F'.
x^* is a random input of F'.
$y^* = \mathrm{F}'_k(x^*)$
$z^* = \mathrm{g}(y^*)$

F' is a puncturable PRF and g is an injective one-way function.

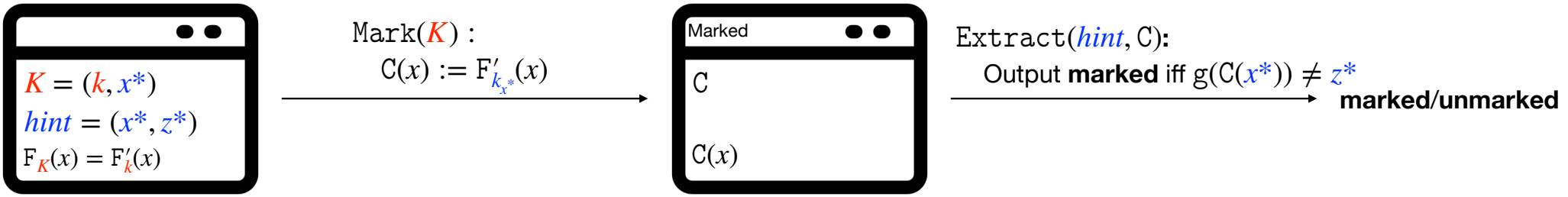
Constructing Public-Key Hinting Watermarkable PRF from Puncturable PRF



k is a PRF key of F'.
x^* is a random input of F'.
$y^* = \mathrm{F}'_k(x^*)$
$z^* = \mathrm{g}(y^*)$

F' is a puncturable PRF and g is an injective one-way function.

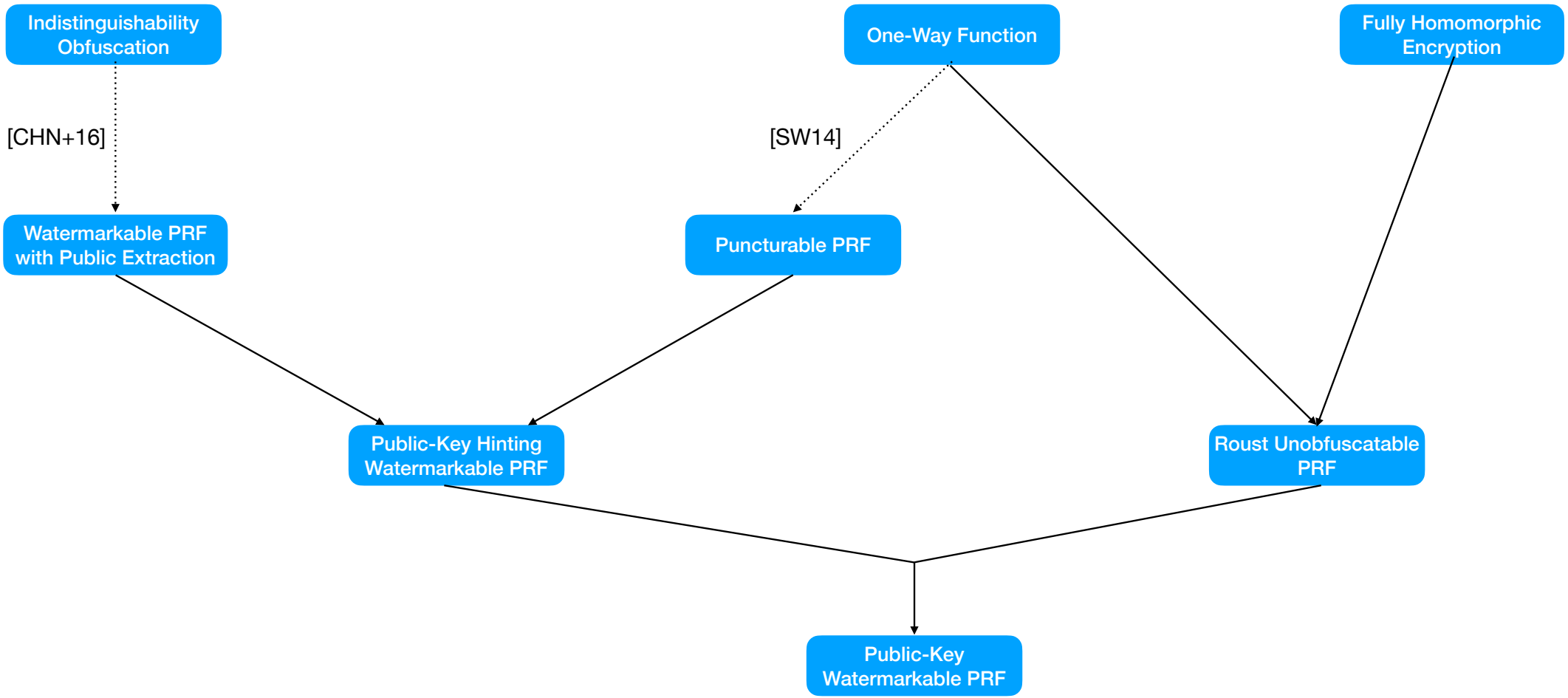
Constructing Public-Key Hinting Watermarkable PRF from Puncturable PRF



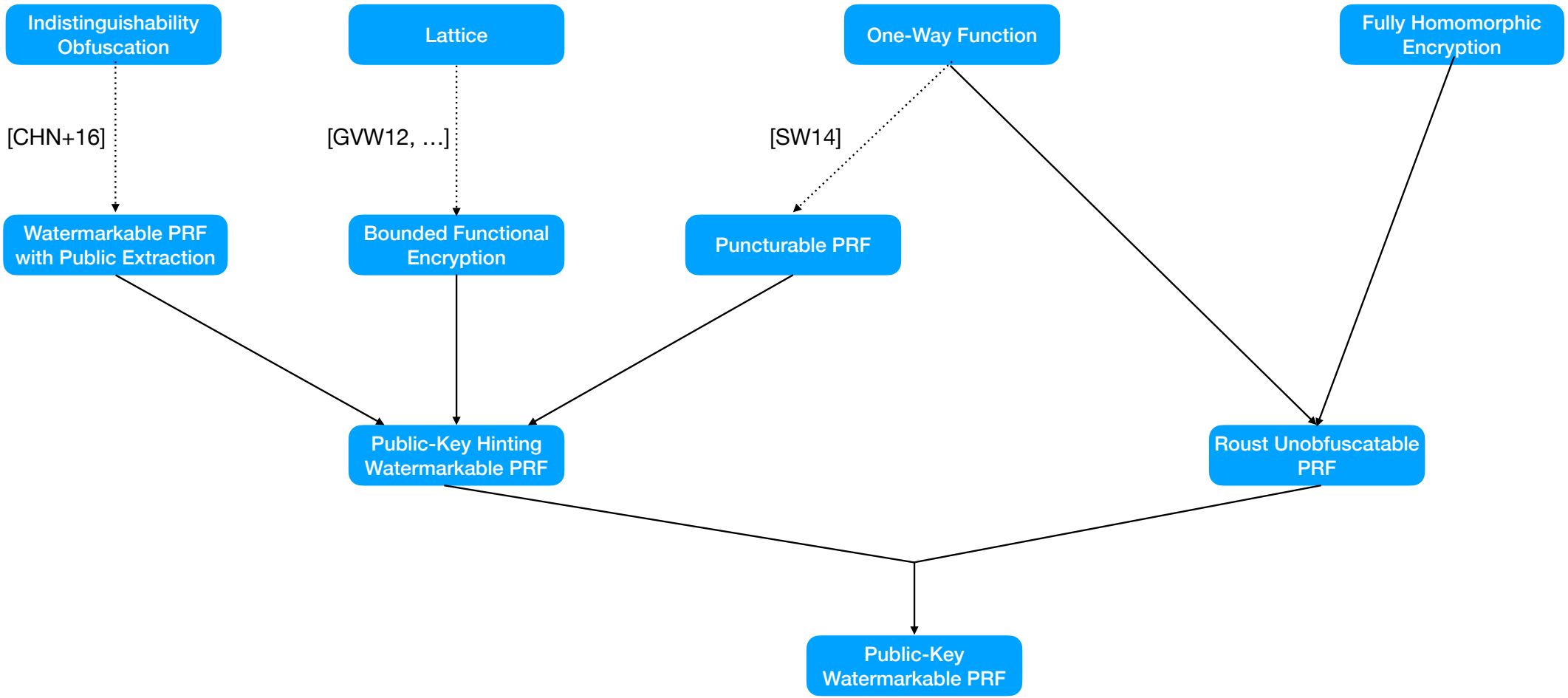
k is a PRF key of F'.
x^* is a random input of F'.
$y^* = \mathrm{F}'_k(x^*)$
$z^* = \mathrm{g}(y^*)$

F' is a puncturable PRF and g is an injective one-way function.

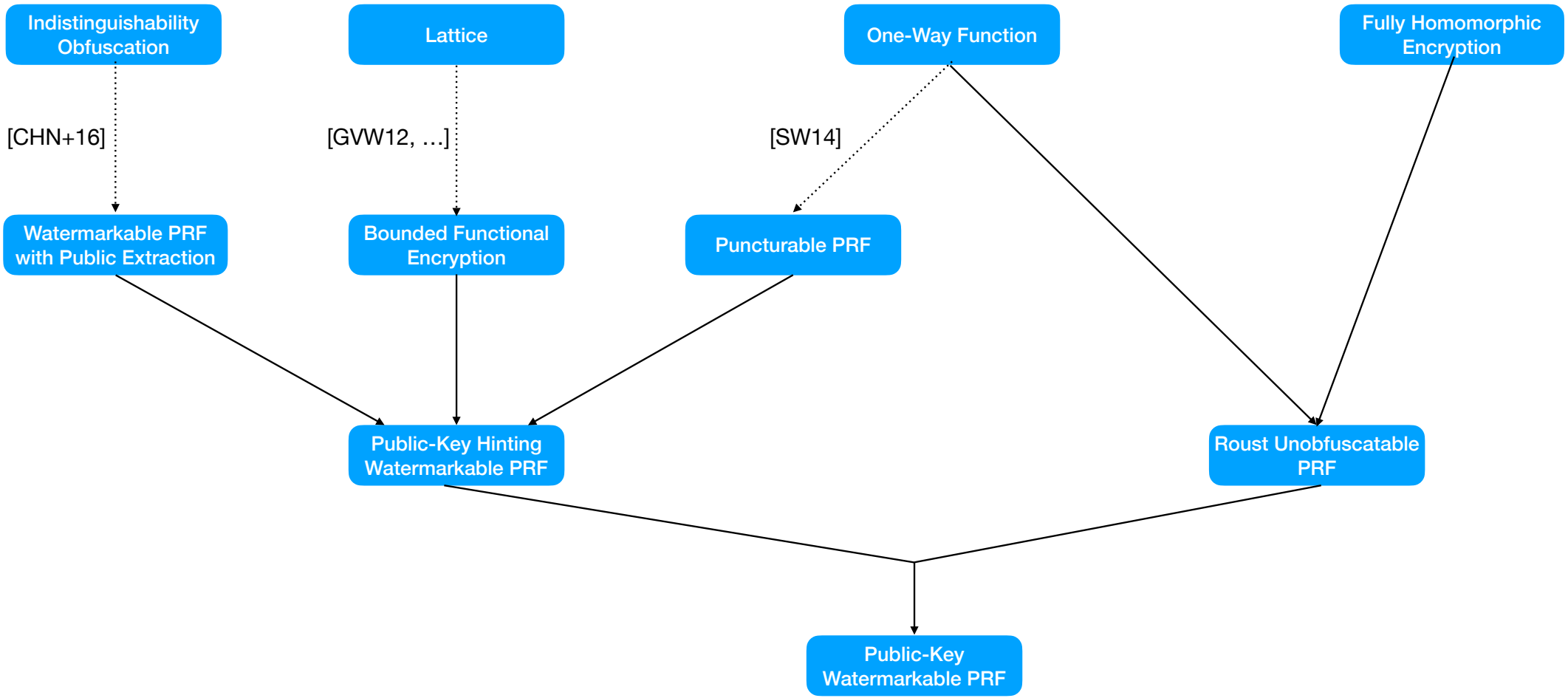
Instantiating Public-Key Watermarkable PRF



Instantiating Public-Key Watermarkable PRF



Conclusion


Indistinguishability Obfuscation
Lattice
One-Way Function
Fully Homomorphic Encryption
[CHN+16]
[GVW12, …]
[SW14]
Watermarkable PRF with Public Extraction
Bounded Functional Encryption
Puncturable PRF
Public-Key Hinting Watermarkable PRF
Roust Unobfuscatable PRF
Public-Key Watermarkable PRF

Conclusion

Assumption	Message-Embedding	ε
Lattice	✗	negl
Lattice + FHE	✗	1/6
Lattice	✓	1/exp
iO	✓	negl
iO+FHE	✓	1/6

Conclusion

Assumption	Message-Embedding	ε
Lattice	✗	negl
Lattice + FHE	✗	1/6
Lattice	✓	1/exp
iO	✓	negl
iO+FHE	✓	1/6

$(MK, EK) \leftarrow \mathtt{KeyGen}$
$K \leftarrow \mathscr{K}$

MK, EK

$\mathfrak{m}$

$\mathtt{C}^* \leftarrow \mathtt{Mark}(MK, K, \mathfrak{m})$

$\mathtt{C}^*$

$\tilde{\mathtt{C}}$

The adversary wins if:
1. $\mathtt{C}^* \approx \tilde{\mathtt{C}}$
2. $\mathtt{Extract}(EK, \tilde{C}) \neq \mathfrak{m}$

$$\varepsilon = \frac{|\{x \in \mathscr{X} : \mathtt{C}^*(x) \neq \tilde{\mathtt{C}}(x)\}|}{|\mathscr{X}|}$$

Open Problems

Assumption	Message-Embedding	ε
Lattice	✗	negl
Lattice + FHE	✗	1/6
Lattice	✓	1/exp
iO	✓	negl
iO+FHE	✓	1/6

Construct Public-Key Watermarkable PRFs with

- message embedding and $\varepsilon \geq negl$ from lattice.
- constant ε without using FHE.
- optimal ε ($\varepsilon \approx 1/2$)

Open Problems

Assumption	Message-Embedding	ε
Lattice	✗	negl
Lattice + FHE	✗	1/6
Lattice	✓	1/exp
iO	✓	negl
iO+FHE	✓	1/6

Construct Public-Key Watermarkable PRFs with

- message embedding and $\varepsilon \geq negl$ from lattice.
- constant ε without using FHE.
- optimal ε ($\varepsilon \approx 1/2$)

Thanks for your Attention!