Practical Order-Revealing Encryption with Limited Leakage

Nathan Chenette, Kevin Lewi, Stephen A. Weis, and David J. Wu

Fast Software Encryption

March, 2016

Order-Revealing Encryption [BLRSZZ15]

sk

Client

$$
\begin{aligned}
\mathrm{ct}_1 &= \mathrm{Enc}(\mathrm{sk}, 123) \\
\mathrm{ct}_2 &= \mathrm{Enc}(\mathrm{sk}, 512) \\
\mathrm{ct}_3 &= \mathrm{Enc}(\mathrm{sk}, 273)
\end{aligned}
$$

Server

secret-key encryption scheme

Order-Revealing Encryption [BLRSZZ15]

$$\begin{aligned} ct_1 &= \mathrm{Enc}(sk, 123) \\ ct_2 &= \mathrm{Enc}(sk, 512) \\ ct_3 &= \mathrm{Enc}(sk, 273) \end{aligned}$$

Which is greater: the value encrypted by ct_1 or the value encrypted by ct_2?

Server

Application: range queries / binary search on encrypted data

Order-Revealing Encryption [BLRSZZ15]

given any two ciphertexts

$\mathrm{ct}_1 = \mathrm{Enc}(\mathrm{sk}, x)$

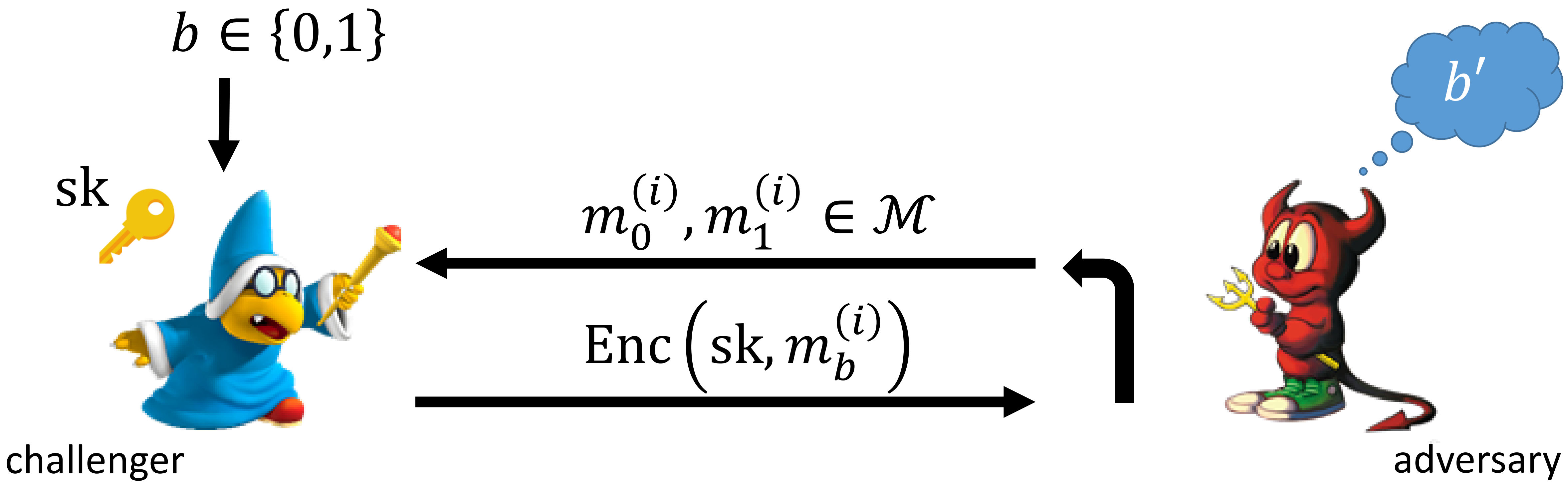
$\mathrm{ct}_2 = \mathrm{Enc}(\mathrm{sk}, y)$

$x > y$

there is a <u>publically</u> evaluatable function that evaluates the comparison function

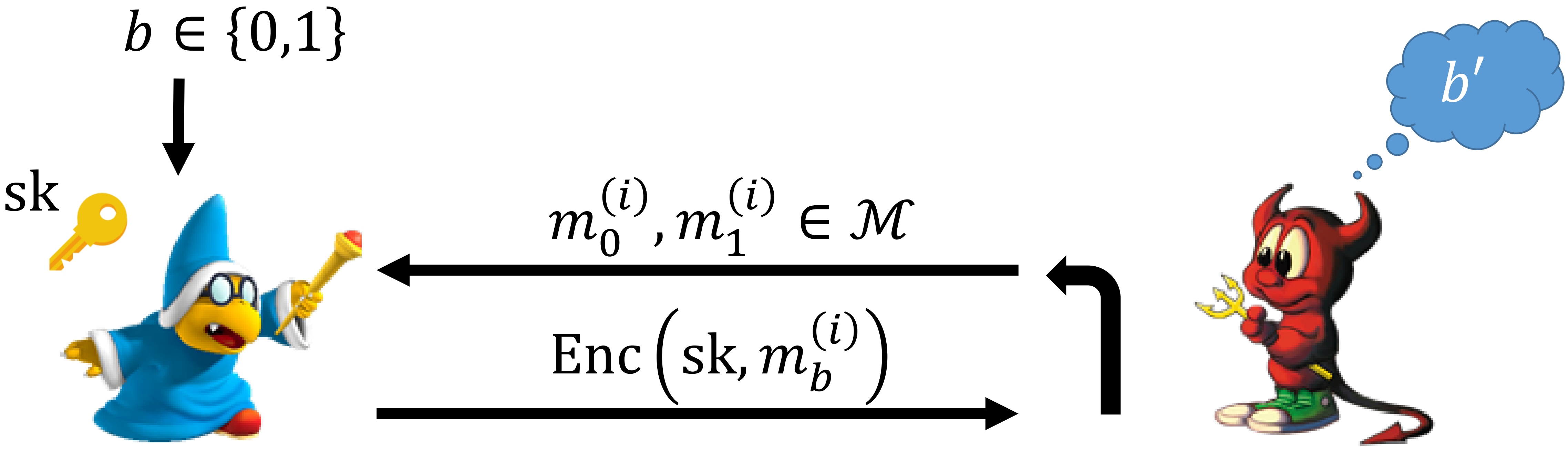
Defining Security

Starting point: semantic security (IND-CPA) [GM84]



semantic security: adversary cannot guess b (except with probability negligibly close to 1/2)

Best-Possible Security [BCLO09]



must impose restriction on messages: otherwise trivial to break semantic security using comparison operator

Best-Possible Security [BCLO09]

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

sk

$m_0^{(i)}, m_1^{(i)} \in \mathcal{M}$

$\mathrm{Enc}\left(\mathrm{sk}, m_b^{(i)}\right)$

$$\forall i, j: m_0^{(i)} < m_0^{(j)} \Leftrightarrow m_1^{(i)} < m_1^{(j)}$$

Best-Possible Security [BCLO09]

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

sk

$m_0^{(i)}, m_1^{(i)} \in \mathcal{M}$

$\mathrm{Enc}\left(\mathrm{sk}, m_b^{(i)}\right)$

order of "left" set of messages same as order of "right" set of messages

Existing Approaches

General-Purpose Multi-Input Functional Encryption [GGGJKLSSZ14, BV15, AJ15]

- Powerful cryptographic primitive that fully subsumes ORE
- Achieves best-possible security
- Impractical (requires obfuscating a PRF)

Existing Approaches

Multilinear-map-based Solution [BLRSZZ15]

- Much more efficient than general purpose indistinguishability obfuscation
- Achieves best-possible security
- Security of multilinear maps not well-understood
- Still quite inefficient (e.g., ciphertexts on the order of GB)

Existing Approaches

Order-preserving encryption (OPE) [BCLO09, BCO11]:

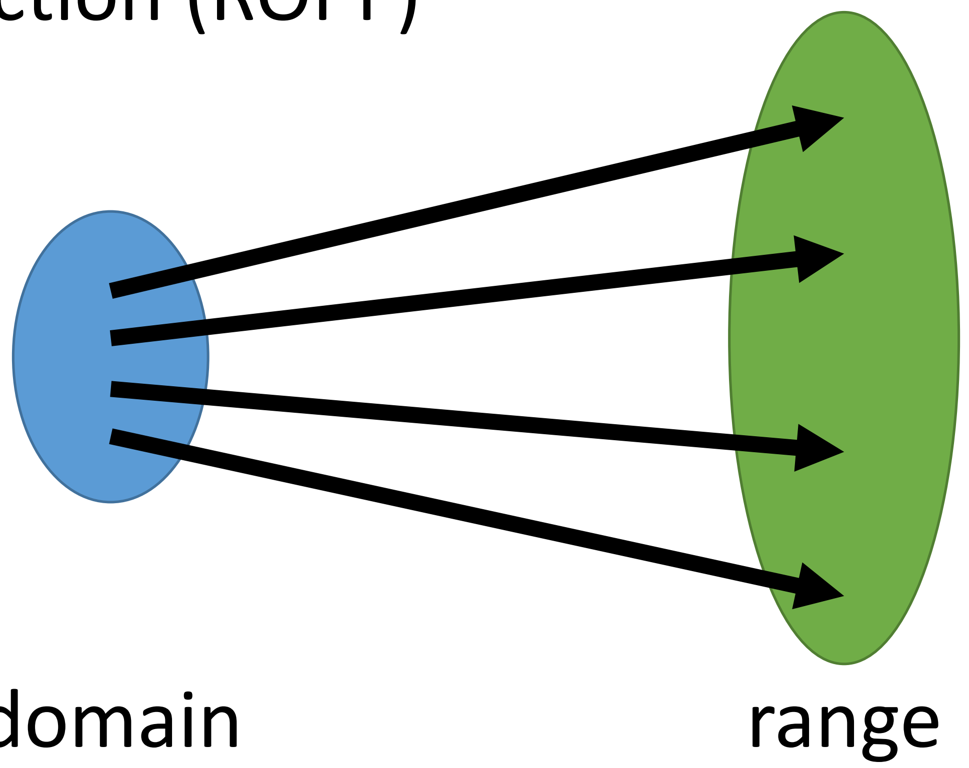
- Comparison operation is <u>direct</u> comparison of ciphertexts:
$$x > y \Leftrightarrow \mathrm{Enc}(\mathrm{sk}, x) > \mathrm{Enc}(\mathrm{sk}, y)$$
- Lower bound: no OPE scheme can satisfy “best-possible” security unless the size of the ciphertext space is <u>exponential</u> in the size of the plaintext space

Existing Approaches

Order-preserving encryption (OPE) [BCLO09, BCO11]:

- No "best-possible" security, so instead, compare with <u>random</u> order-preserving function (ROPF)

encryption function implements a <u>random</u> order-preserving function



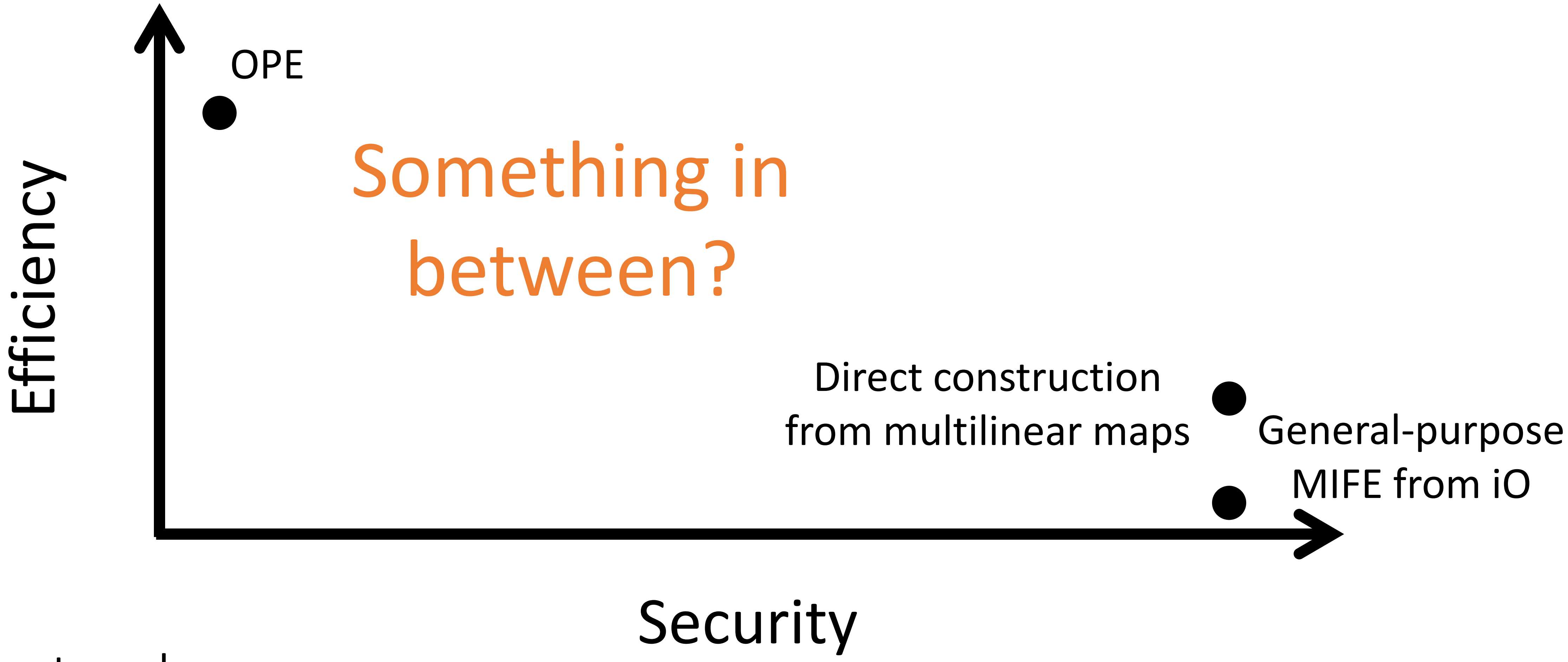
Existing Approaches

Properties of a random order-preserving function [BCO’11]:

- Each ciphertext roughly leaks half of the most significant bits
- Each pair of ciphertexts roughly leaks half of the most significant bits of their difference

No semantic security for
even a single message!

Existing Approaches



Not drawn to scale

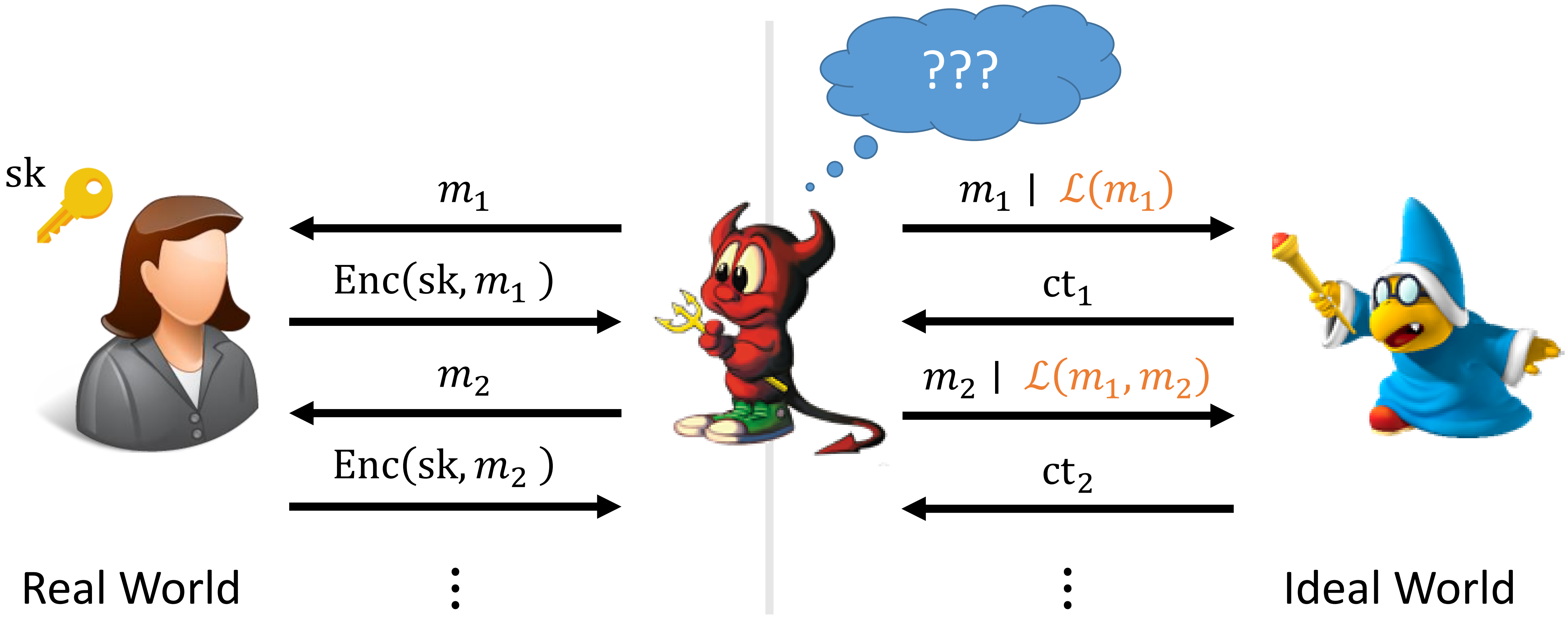
A New Security Notion

Two existing security notions:

- IND-OCPA: strong security, but hard to achieve efficiently
- ROPF-CCA: efficiently constructible, but lots of leakage, and difficult to precisely quantify the leakage

A New Security Notion: SIM-ORE

Idea: augment "best-possible" security with a leakage function $\mathcal{L}$



A New Security Notion: SIM-ORE

Similar to SSE definitions [CM05, CGKO06]

Leakage functions specifies exactly what is leaked

“Best-possible” simulation security:

$$\mathcal{L}(m_1, \ldots, m_q) = \{\, (i, j, \mathbf{1}\{m_i < m_j\}) \mid 1 \le i < j \le q \,\}$$

A New Security Notion: SIM-ORE

"Best-possible" simulation security:

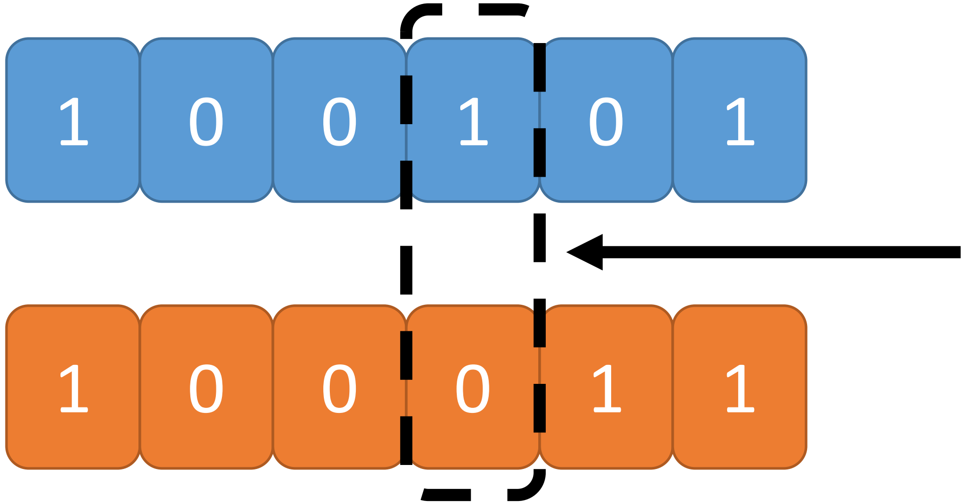
$$\mathcal{L}\left(m_1, \ldots, m_q\right) = \left\{ \left(i, j, \mathbf{1}\left\{m_i < m_j\right\}\right) \mid 1 \leq i < j \leq q \right\}$$

Anything that can be computed given the ciphertexts can be computed given the ordering on the messages

Our Construction

Leak a little more than just the ordering:

$$\mathcal{L}(m_1, \dots, m_q) = \left\{ \left(i, j, \mathbf{1}\{m_i < m_j\}, \mathrm{ind}_{\mathrm{diff}}(m_i, m_j) \right) \,\middle|\, 1 \le i < j \le q \right\}$$



$\mathrm{ind}_{\mathrm{diff}}(m_1, m_2)$: index of first bit that differs

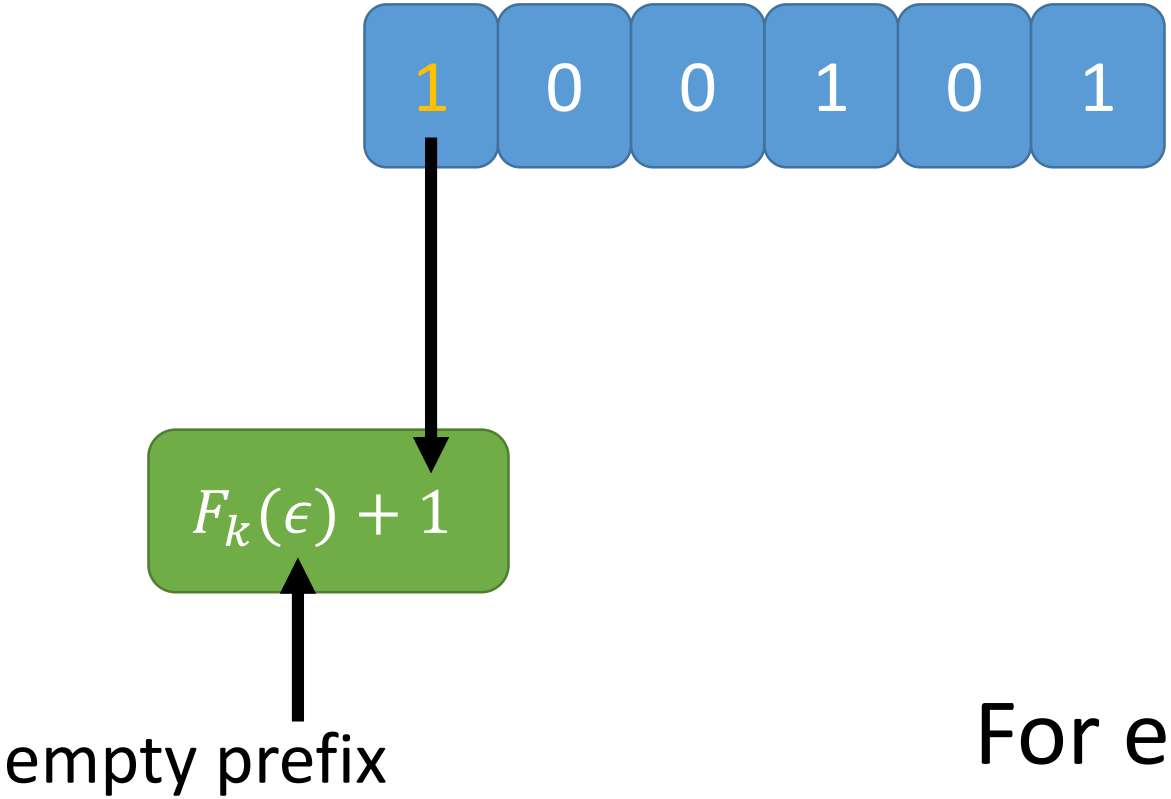
Our Construction

1	0	0	1	0	1

$F: \mathcal{K} \times \{0,1\}^* \rightarrow \mathbb{Z}_3$

For each index i, apply a PRF to the first $i-1$ bits, then add $b_i \pmod 3$

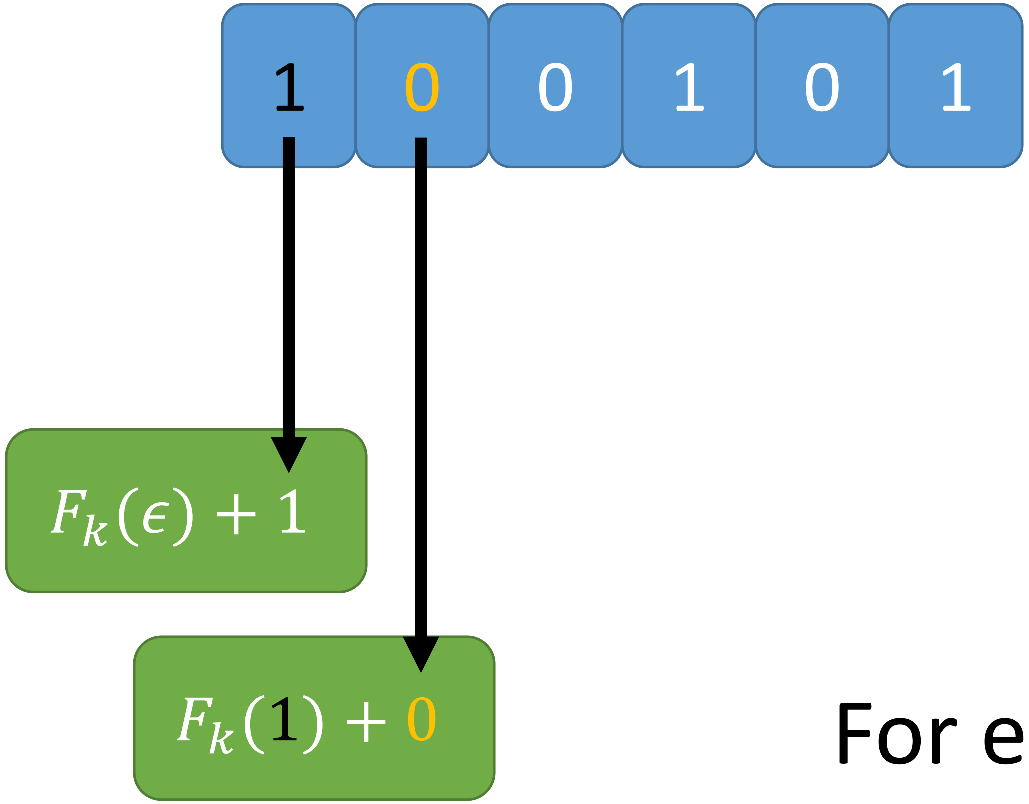
Our Construction



For each index i, apply a PRF to the first $i - 1$ bits, then add $b_i \pmod 3$

$F\colon \mathcal{K} \times \{0,1\}^* \to \mathbb{Z}_3$

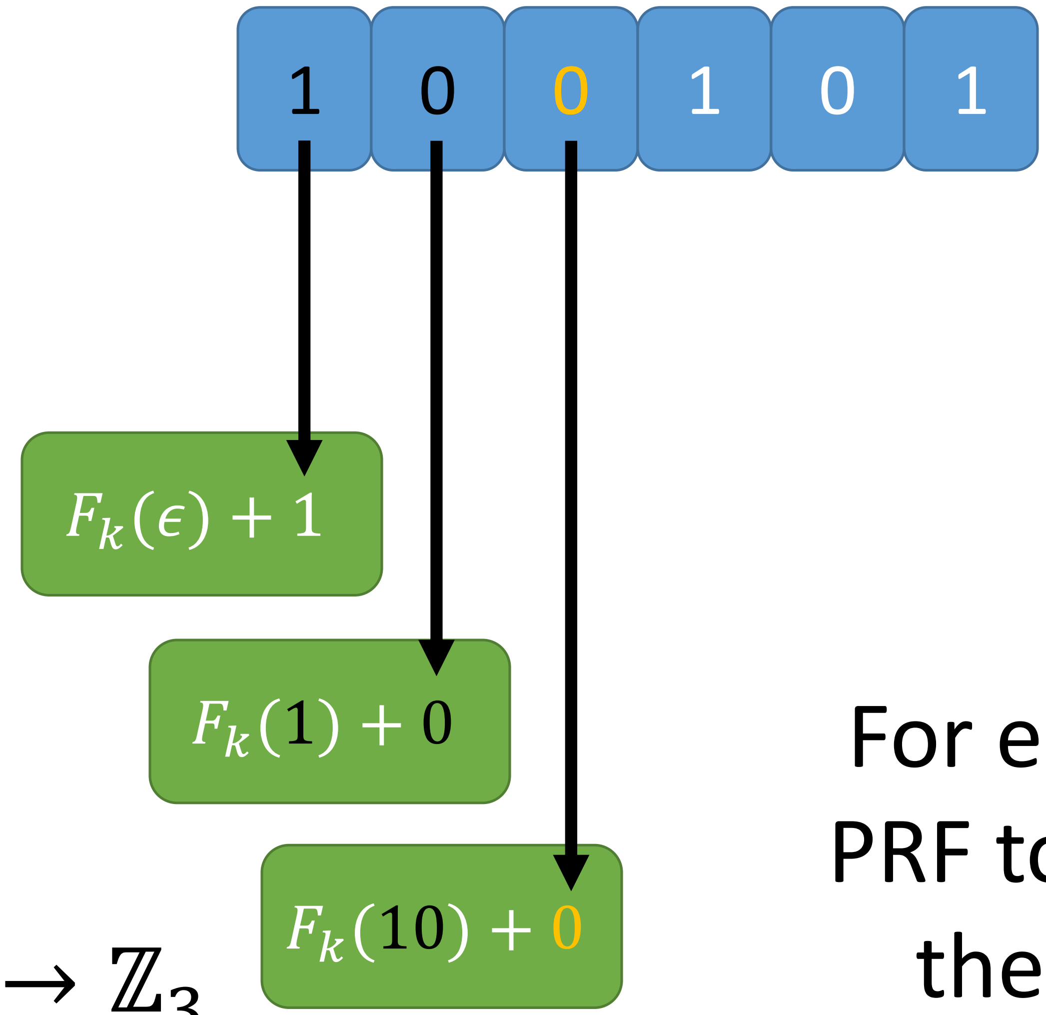
Our Construction



For each index i, apply a PRF to the first $i-1$ bits, then add $b_i \pmod 3$

$F\colon \mathcal{K} \times \{0,1\}^* \to \mathbb{Z}_3$

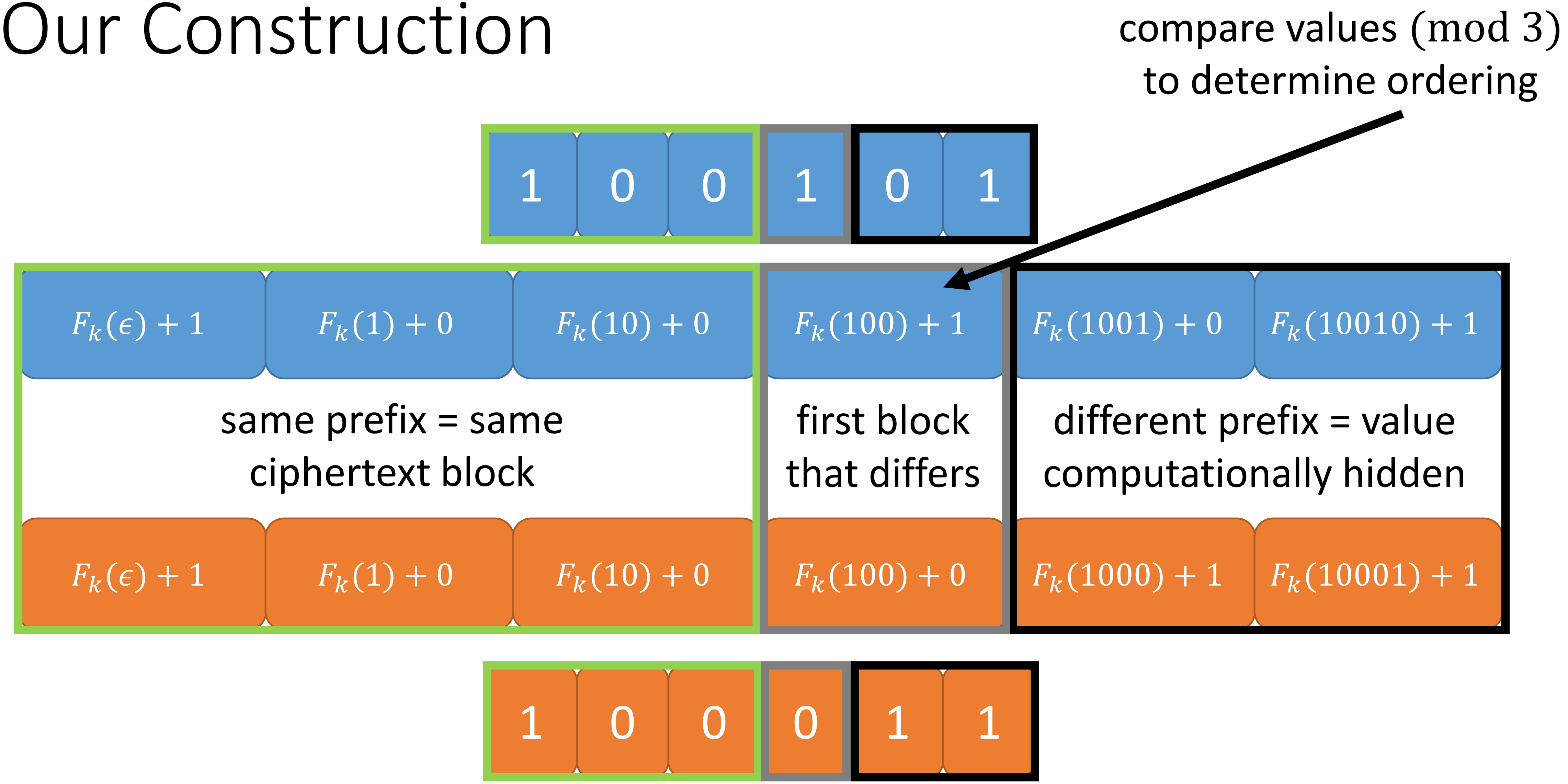
Our Construction



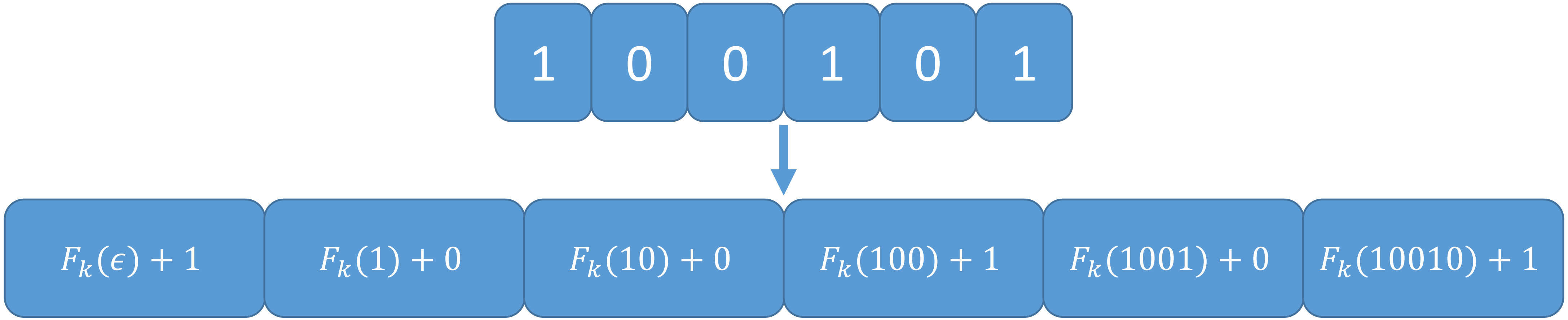
$F\colon \mathcal{K} \times \{0,1\}^* \to \mathbb{Z}_3$

For each index i, apply a PRF to the first $i-1$ bits, then add $b_i \pmod 3$

Our Construction



Our Construction: Efficiency

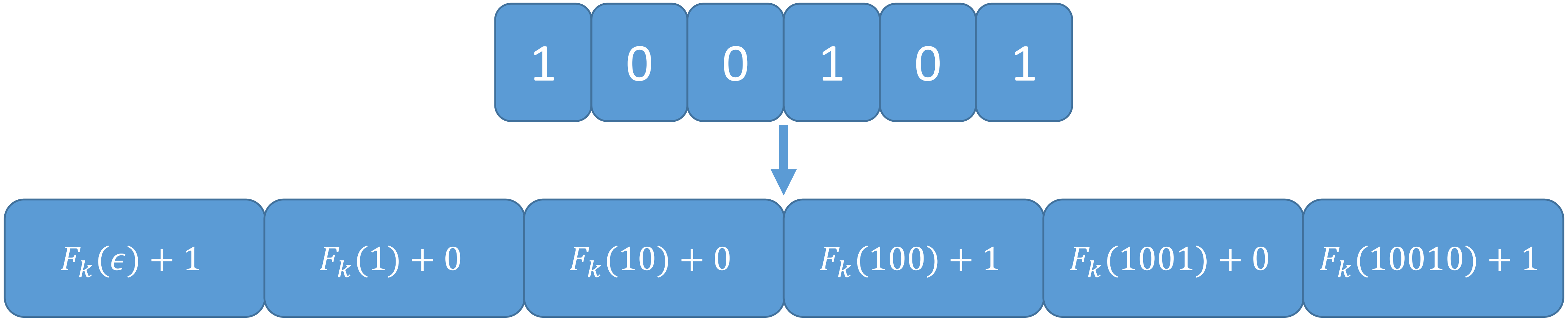


Each ciphertext block is an element in $\mathbb{Z}_3$

For n-bit messages, can obtain ciphertexts of length $\approx 1.6n$

Encryption only requires PRF evaluations while decryption just requires bitwise comparisons

Our Construction: Security

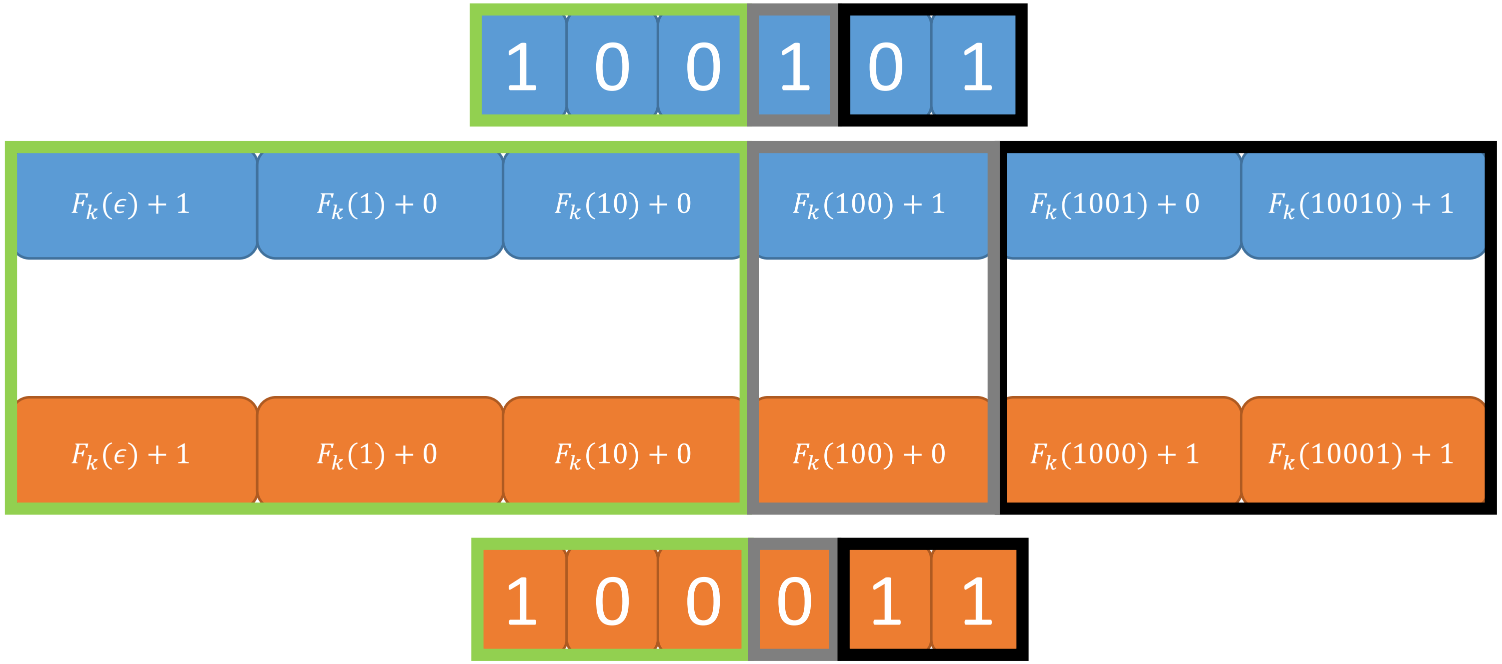


Security follows directly from security of the PRF

<u>Proof sketch</u>. Simulator responds to encryption queries using random strings. Maintains consistency using leakage information (first bit that differs).

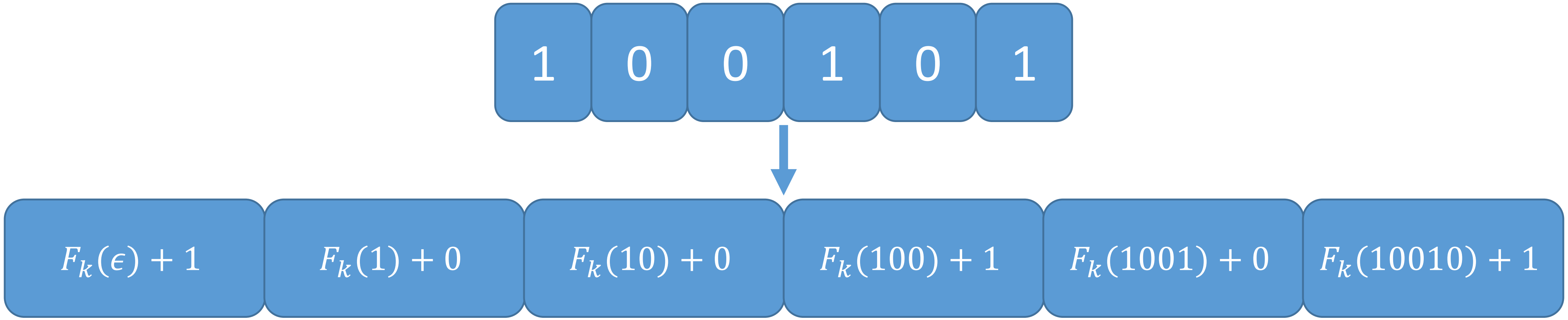
OPE Conversion

In database applications, OPE preferred over ORE since it does not require changes to the DBMS (e.g., supporting custom comparator)



Work modulo n instead of modulo 3 and view ciphertext blocks as digits of a base n number

OPE Conversion

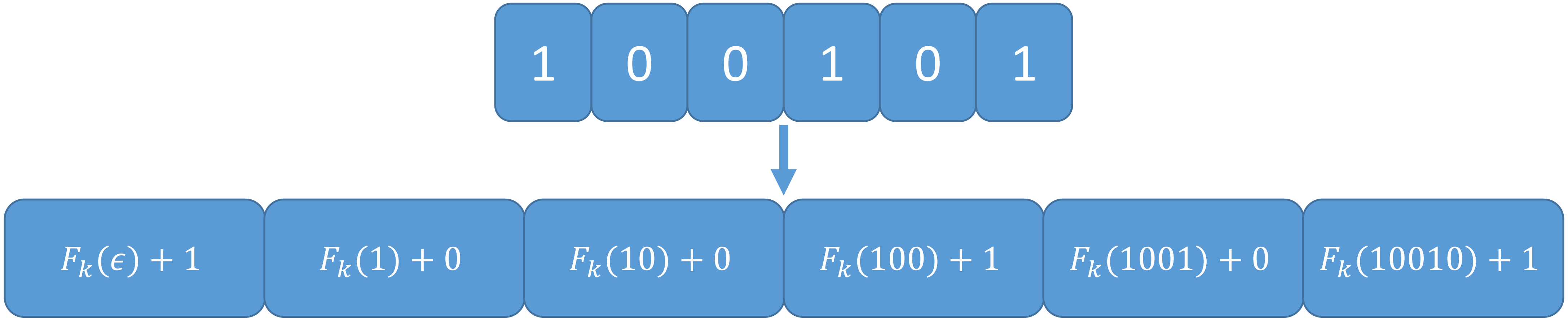


But sacrifice some correctness (when the values “wrap around”):

- If $F_k(z) = n - 1$, then $F_k(z) + 1 = 0 \pmod{n}$

Happens with negligible probability if n is $\omega(\log \lambda)$ bits long

OPE Conversion



Note: unlike most existing OPE schemes, this OPE scheme is <u>not</u> a ROPF, and does <u>not</u> suffer from many of the security limitations of ROPFs

Comparison to Previous OPE Schemes

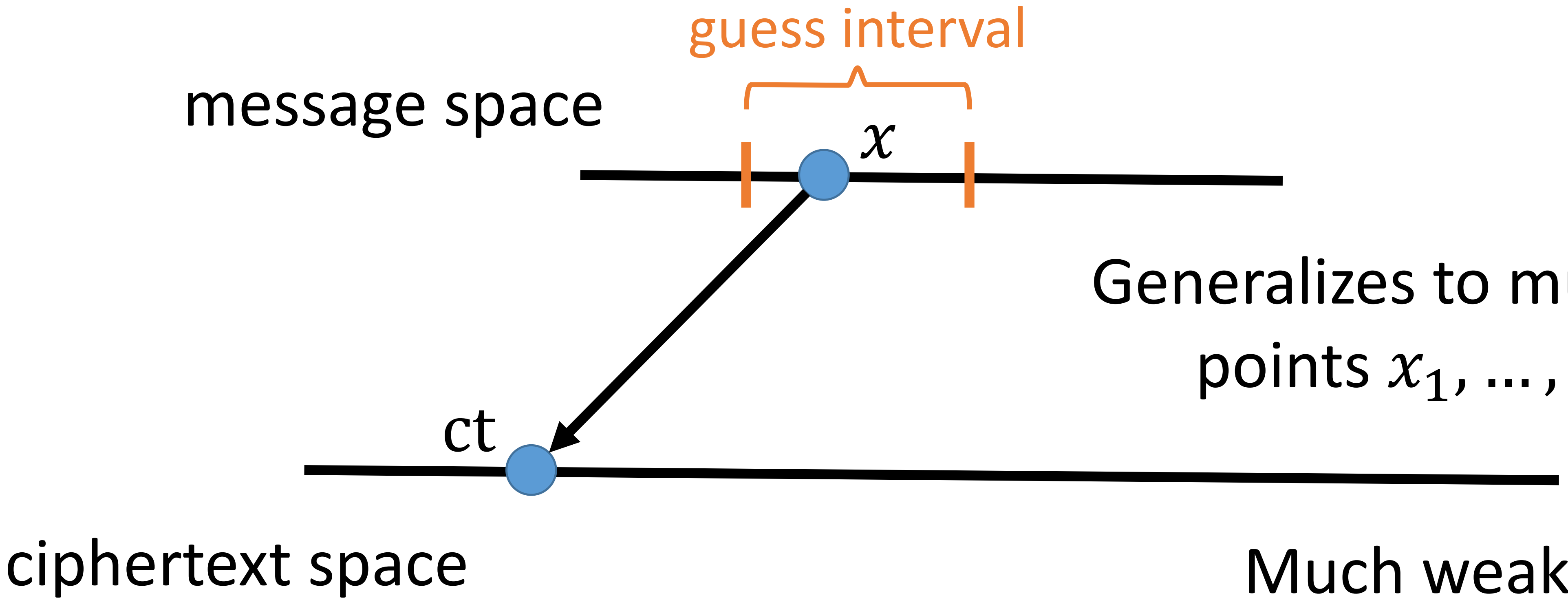
One metric: window one-wayness [BCO11]

Let message space be $\{0,1,\dots,M\}$

Given an encryption of a random message x, adversary outputs an interval I in $\{0,1,\dots,M\}$, and wins if $x \in I$

Comparison to Previous OPE Schemes

Window one-wayness:



Generalizes to multiple points $x_1, \dots, x_z$

Much weaker than semantic security!

Comparison to Previous OPE Schemes

Theorem (Informal) [BCO11]: For an ROPF, if the size of the guess interval $r = O(\sqrt{M})$, then there is an efficient adversary whose window one-wayness advantage is close to 1.

Each ciphertext alone <u>reveals half</u> of the most significant bits of the plaintext!

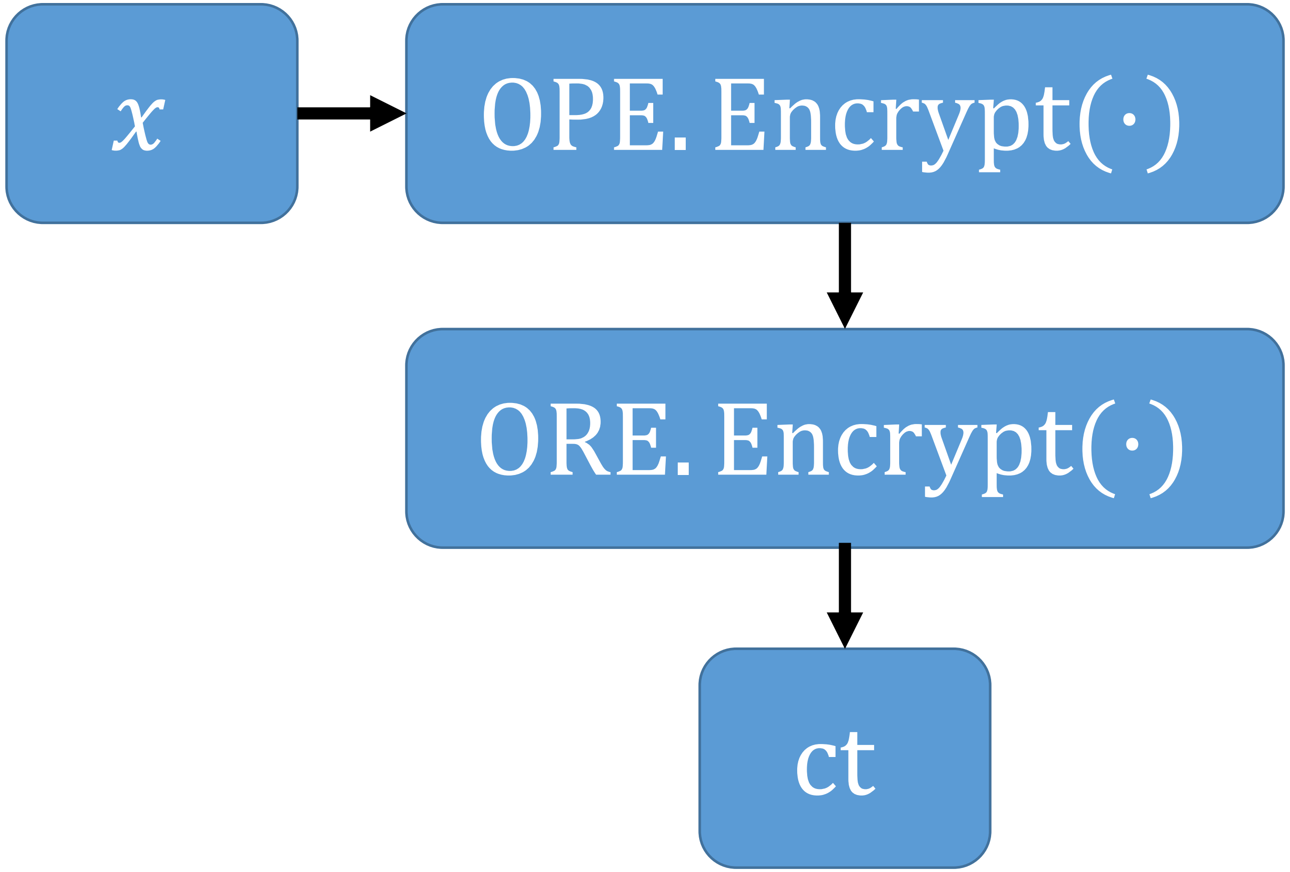
Comparison to Previous OPE Schemes

Theorem (Informal). For our OPE scheme, if the size of the guess interval $r = M^{1-\epsilon}$ for any constant $\epsilon > 0$, then for all efficient adversaries, their (generalized) window one-wayness advantage is negligible.

No constant fraction ϵ of the bits of the plaintexts are revealed.

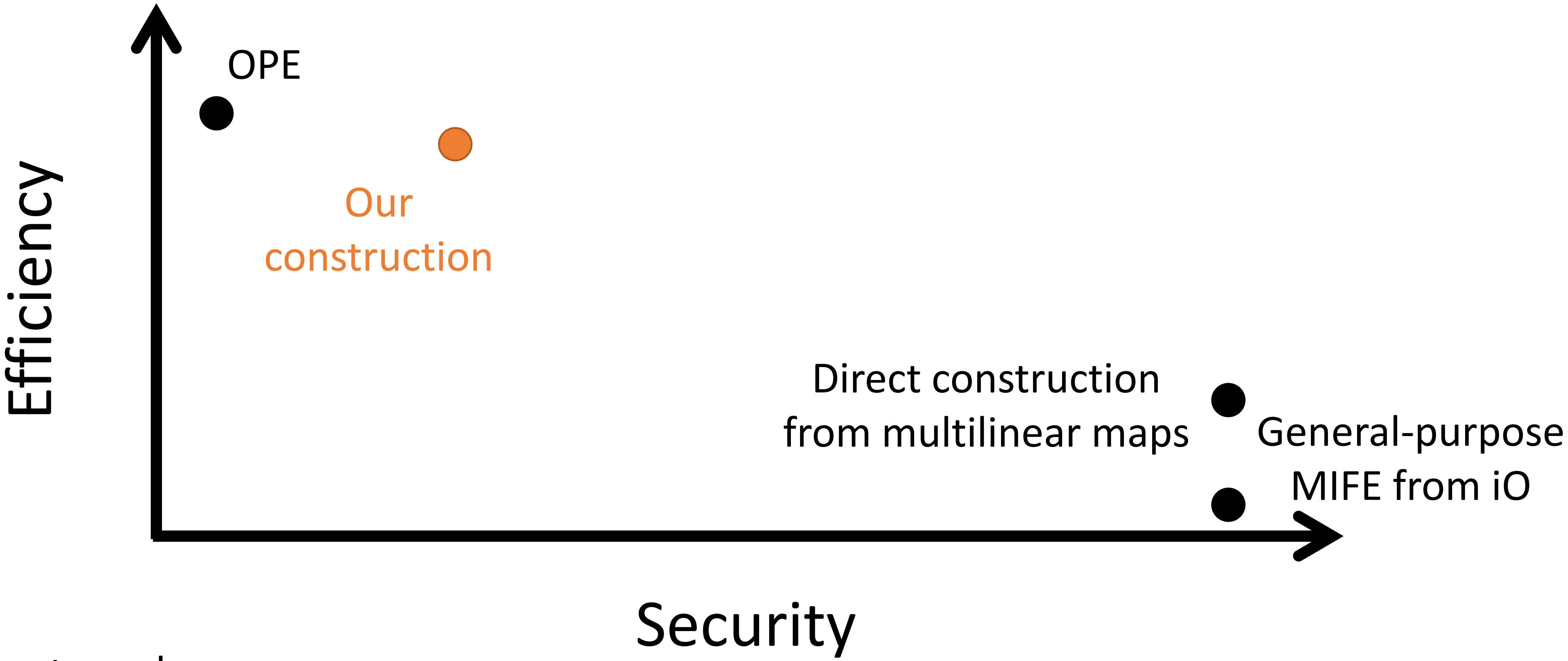
Composing OPE with ORE

Possible to compose OPE with ORE to achieve more secure OPE scheme:



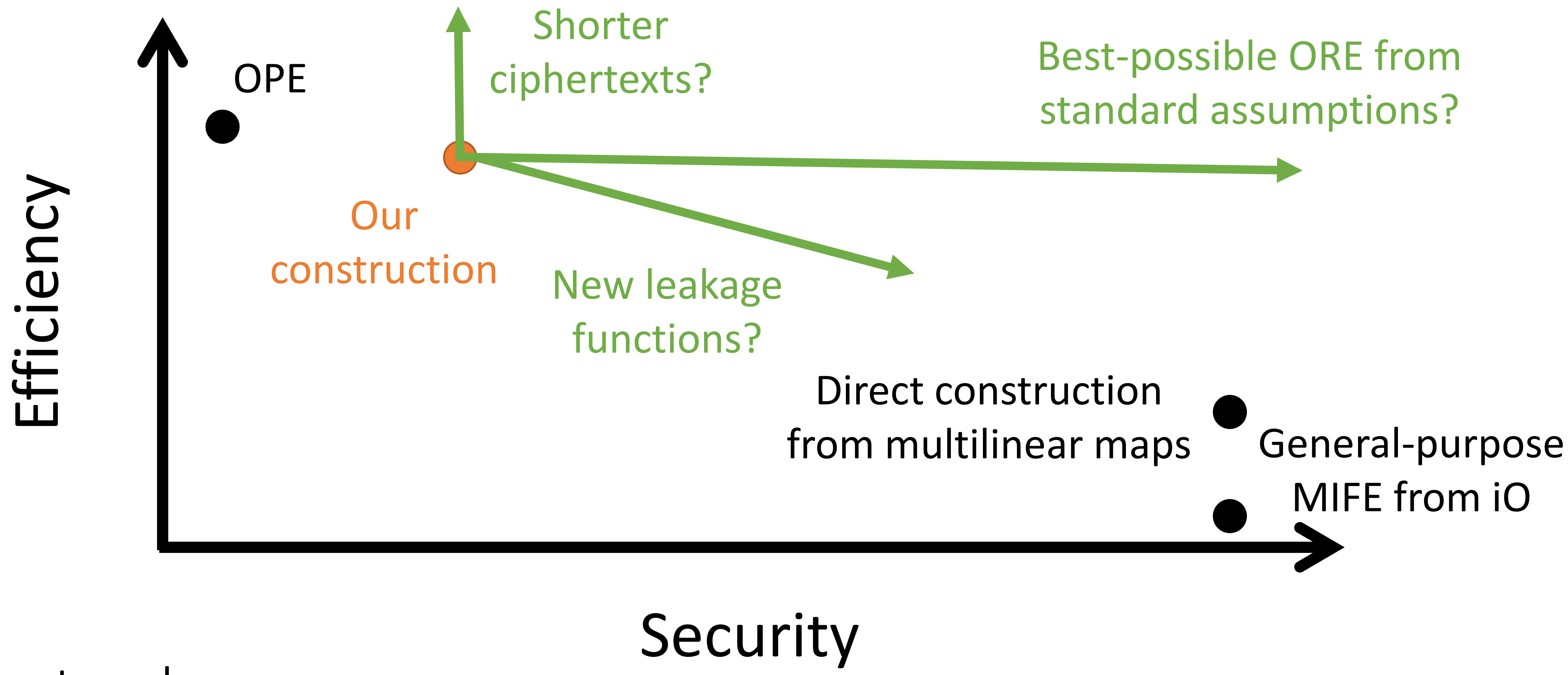
Resulting construction strictly stronger than inner OPE scheme, but may not be more secure than directly applying ORE to plaintext

The Landscape of OPE/ORE



Not drawn to scale

Directions for Future Research



Not drawn to scale

Sample Implementation:

`https://github.com/kevinlewi/fastore`

Project Website:

`https://crypto.stanford.edu/ore/`

Questions?

http://eprint.iacr.org/2015/1125