How to Compress Encrypted Data

Nils Fleischhacker, Kasper Green Larsen, and Mark Simkin

Lyon, 26. April 2023

The Problem

$\mathsf{Enc}(m_1)$

$\mathsf{Enc}(m_2)$

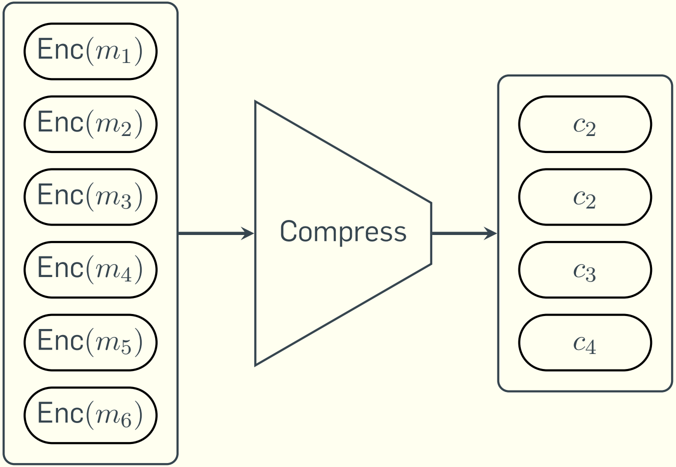
$\mathsf{Enc}(m_3)$

$\mathsf{Enc}(m_4)$

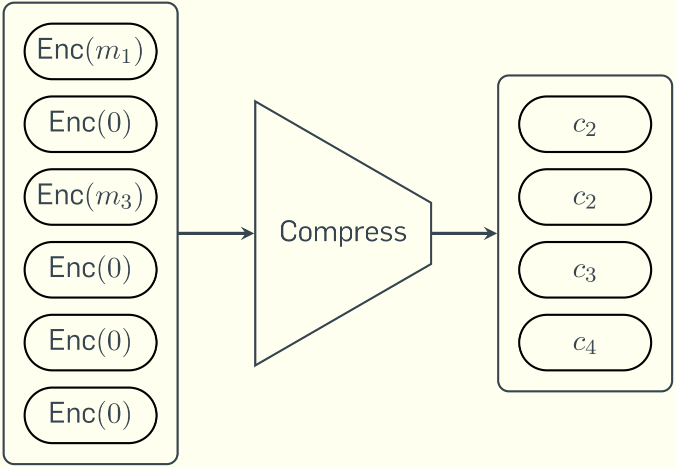
$\mathsf{Enc}(m_5)$

$\mathsf{Enc}(m_6)$

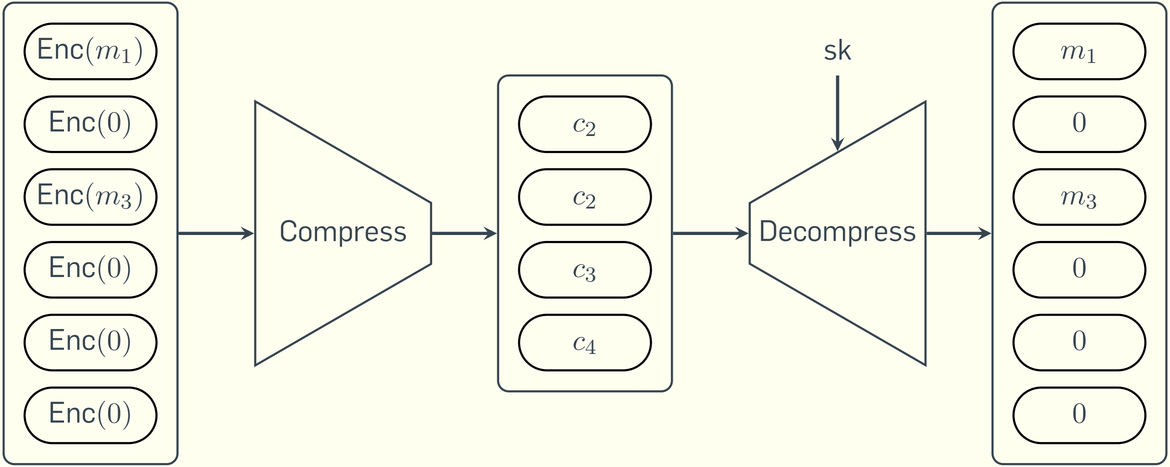
The Problem


$\text{Enc}(m_1)$
$\text{Enc}(m_2)$
$\text{Enc}(m_3)$
$\text{Enc}(m_4)$
$\text{Enc}(m_5)$
$\text{Enc}(m_6)$
Compress
c_2
c_2
c_3
c_4

The Problem



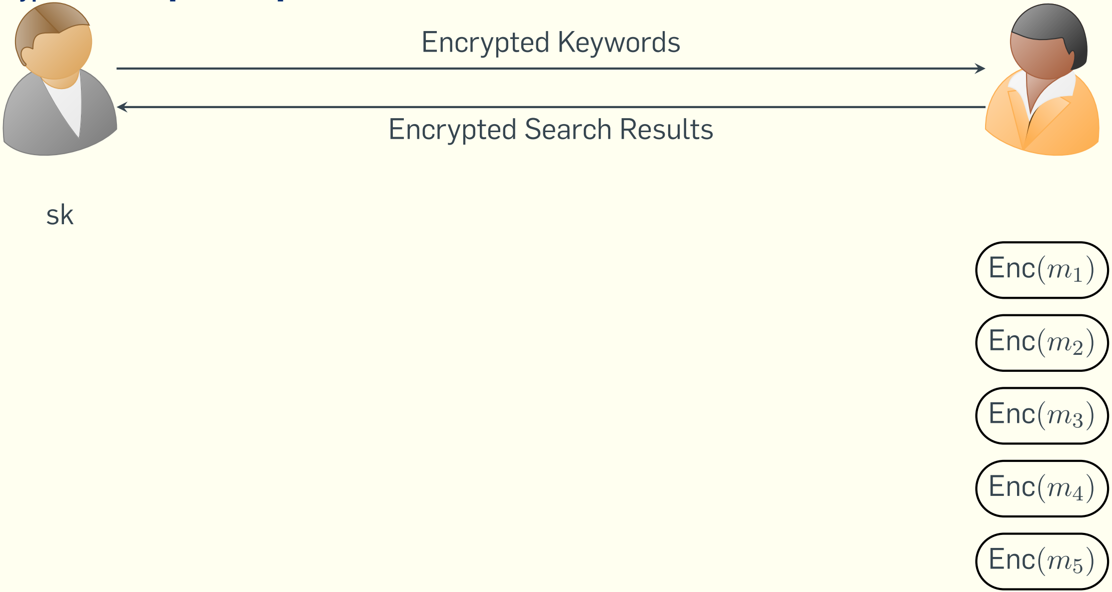
The Problem



$\mathsf{Enc}(m_1)$
$\mathsf{Enc}(0)$
$\mathsf{Enc}(m_3)$
$\mathsf{Enc}(0)$
$\mathsf{Enc}(0)$
$\mathsf{Enc}(0)$
Compress
c_2
c_2
c_3
c_4
sk
Decompress
m_1
0
m_3
0
0
0

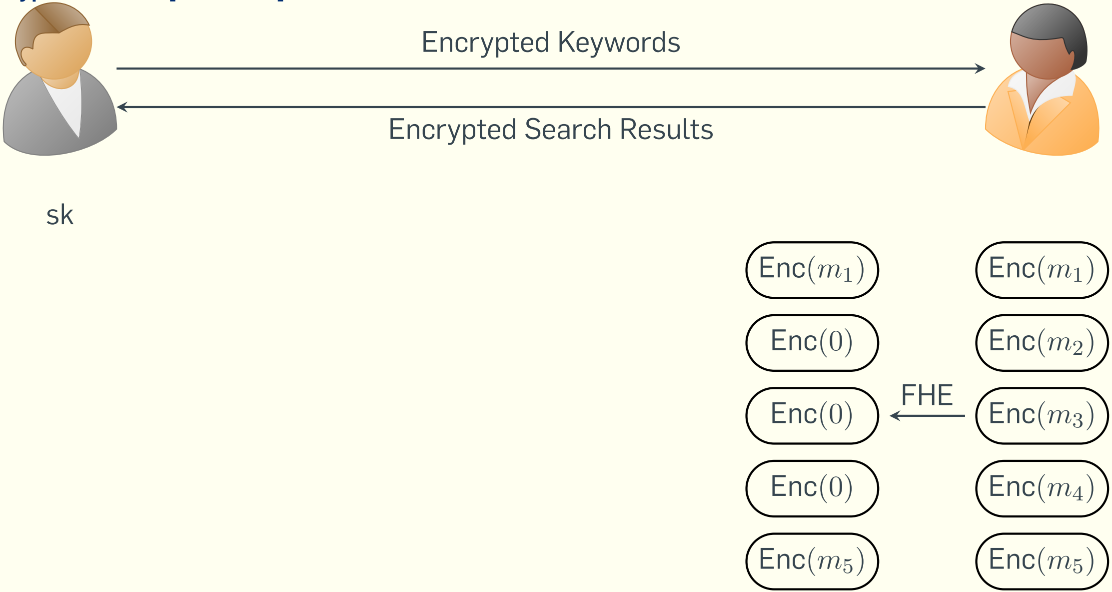
Motivation

Encrypted Search [CDGLY21]



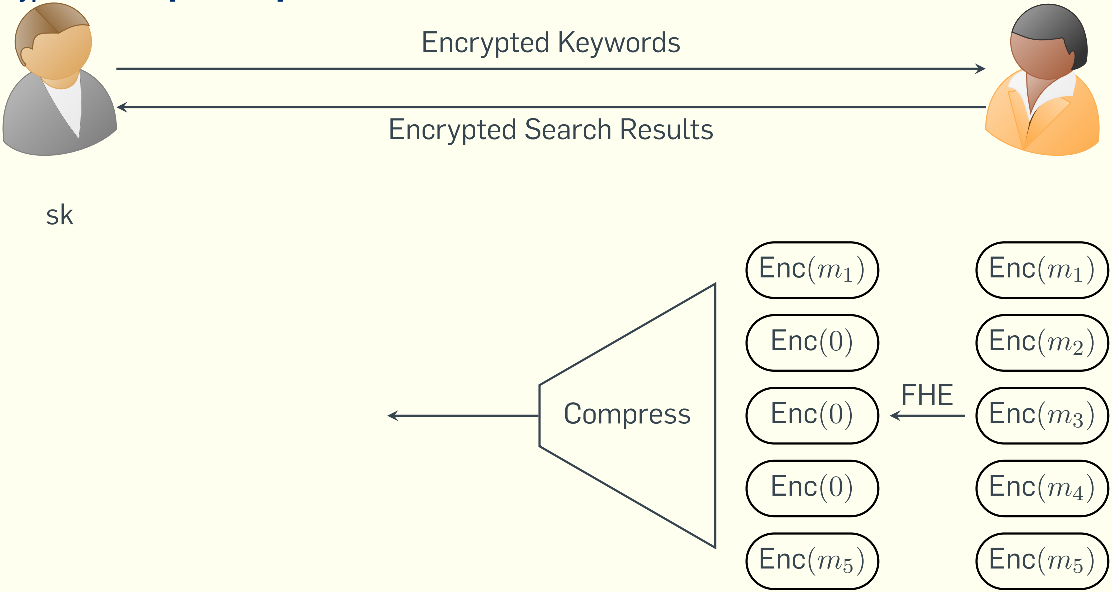
Motivation

Encrypted Search [CDGLY21]



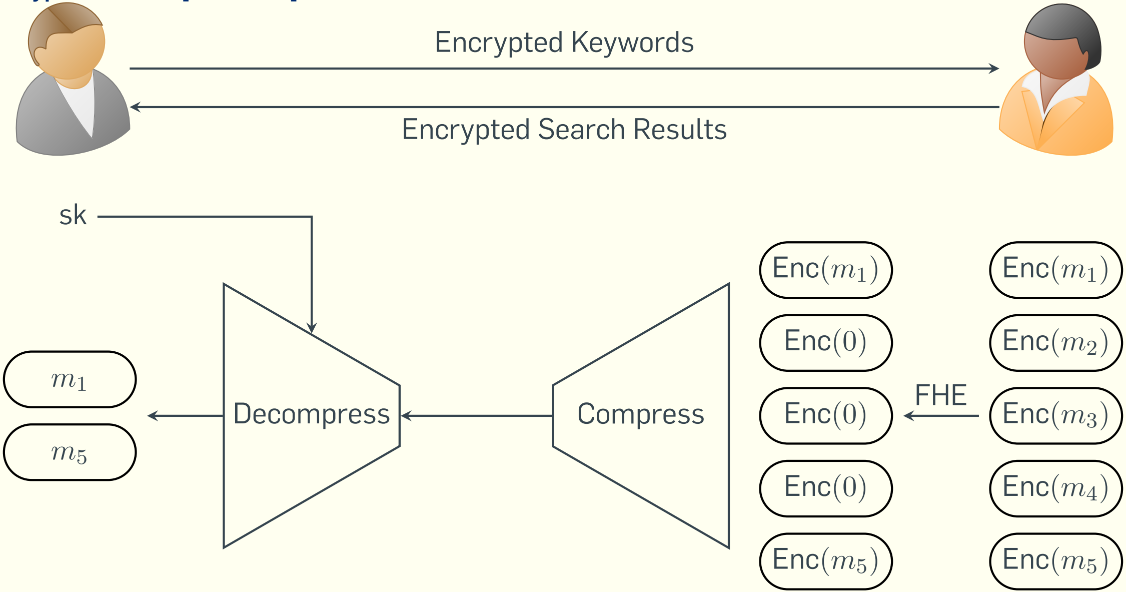
Motivation

Encrypted Search [CDGLY21]



Motivation

Encrypted Search [CDGLY21]



Construction from Sparse Polynomials

Enc(m_1) Enc(0) Enc(m_3) Enc(0) $\cdots$

Construction from Sparse Polynomials

$$f(x) = \mathsf{Enc}(m_1) \cdot x^0 + \mathsf{Enc}(0) \cdot x^1 + \mathsf{Enc}(m_3) \cdot x^2 + \mathsf{Enc}(0) \cdot x^3 + \cdots$$

Construction from Sparse Polynomials

$$f(x) = \mathsf{Enc}(m_1) \cdot x^0 + \mathsf{Enc}(0) \cdot x^1 + \mathsf{Enc}(m_3) \cdot x^2 + \mathsf{Enc}(0) \cdot x^3 + \cdots$$

Polynomials of degree t can be uniquely interpolated from $t+1$ points.

Construction from Sparse Polynomials

$$f(x) = \text{Enc}(m_1) \cdot x^0 + \text{Enc}(0) \cdot x^1 + \text{Enc}(m_3) \cdot x^2 + \text{Enc}(0) \cdot x^3 + \cdots$$

Polynomials of degree t can be uniquely interpolated from $t+1$ points.

Polynomials of sparsity t can be uniquely interpolated from $2t+2$ points.

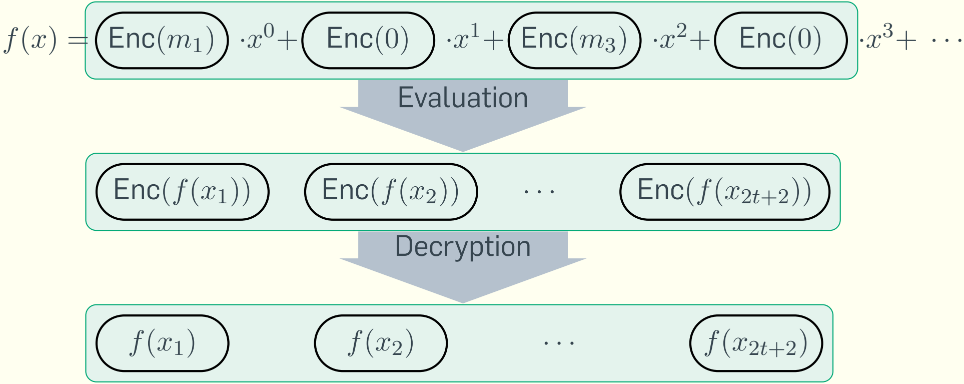
Construction from Sparse Polynomials

$$f(x) = \mathsf{Enc}(m_1) \cdot x^0 + \mathsf{Enc}(0) \cdot x^1 + \mathsf{Enc}(m_3) \cdot x^2 + \mathsf{Enc}(0) \cdot x^3 + \cdots$$

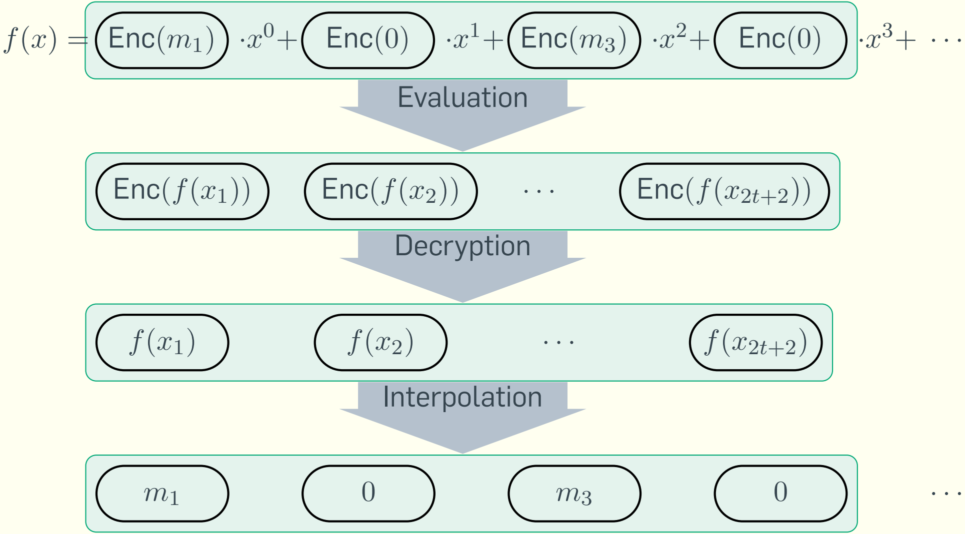
Evaluation

$$\mathsf{Enc}(f(x_1)) \quad \mathsf{Enc}(f(x_2)) \quad \cdots \quad \mathsf{Enc}(f(x_{2t+2}))$$

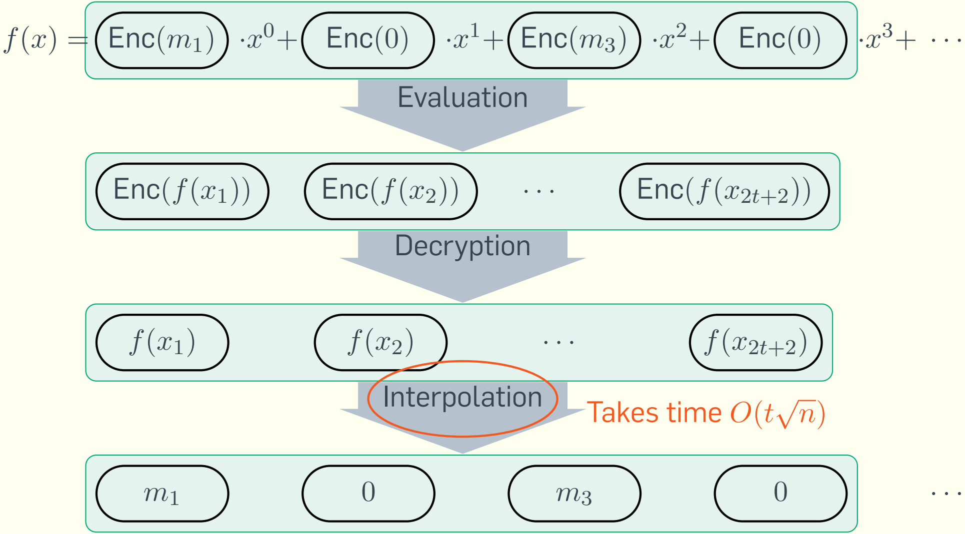
Construction from Sparse Polynomials



Construction from Sparse Polynomials



Construction from Sparse Polynomials



Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

t

H_1	0	0	0	0	0	0	
H_2	0	0	0	0	0	0	$\frac{\varepsilon}{\log t}$
H_3	0	0	0	0	0	0	

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

$H_1(m_1)$	0	m_1	0
$H_2(m_1)$	0	m_1	0
$H_3(m_1)$	0	0	m_1

t

0	1	0
0	1	0
0	0	1

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

$H_1(m_2)$	0	m_1	m_2
$H_2(m_2)$	0	$m_1 + m_2$	0
$H_3(m_2)$	m_2	0	m_1

t

0	1	1
0	2	0
1	0	1

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

$H_1(m_3)$	m_3	m_1	m_2
$H_2(m_3)$	0	$m_1 + m_2$	m_3
$H_3(m_3)$	$m_2 + m_3$	0	m_1

t

1	1	1
0	2	1
2	0	1

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

$H_1(m_4)$	m_3	$m_1 + m_4$	m_2
$H_2(m_4)$	0	$m_1 + m_2$	$m_3 + m_4$
$H_3(m_4)$	$m_2 + m_3$	0	$m_1 + m_4$

t

1	2	1
0	2	2
2	0	2

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

H_1	m_3	$m_1 + m_4$	m_2
H_2	0	$m_1 + m_2$	$m_3 + m_4$
H_3	$m_2 + m_3$	0	$m_1 + m_4$

t

1	2	1
0	2	2
2	0	2

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

$$(m_1, m_2, m_3, m_4)$$

H_1	m_3	$m_1 + m_4$	m_2
H_2	0	$m_1 + m_2$	$m_3 + m_4$
H_3	$m_2 + m_3$	0	$m_1 + m_4$

t

1	2	1
0	2	2
2	0	2

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

$H_1(m_3)$	m_3	$m_1 + m_4$	m_2
$H_2(m_3)$	0	$m_1 + m_2$	$m_3 + m_4$
$H_3(m_3)$	$m_2 + m_3$	0	$m_1 + m_4$

t

1	2	1
0	2	2
2	0	2

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

H_1	0	$m_1 + m_4$	m_2
H_2	0	$m_1 + m_2$	m_4
H_3	m_2	0	$m_1 + m_4$

t

0	2	1
0	2	1
1	0	2

$\frac{\varepsilon}{\log t}$

Invertible Bloom Lookup Tables [GM11]

(m_1, m_2, m_3, m_4)

H_1	0	$m_1 + m_4$	m_2
H_2	0	$m_1 + m_2$	m_4
H_3	m_2	0	$m_1 + m_4$

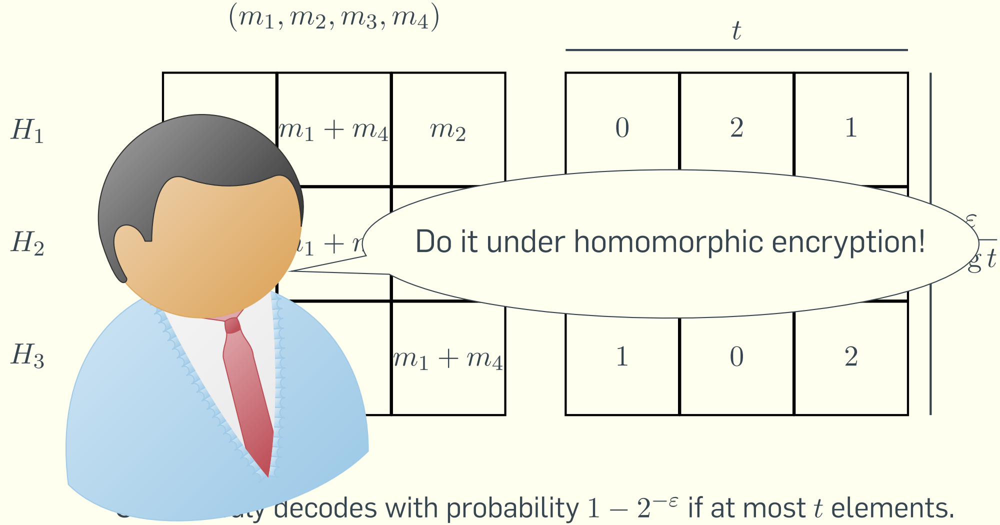
t

0	2	1
0	2	1
1	0	2

$\frac{\varepsilon}{\log t}$

Successfuly decodes with probability $1 - 2^{-\varepsilon}$ if at most t elements.
Only requires a handful of additions.

Invertible Bloom Lookup Tables [GM11]



decodes with probability $1 - 2^{-\varepsilon}$ if at most t elements.

Only requires a handful of additions.

Construction from IBLTs

A First Attempt

H_1	0	0	0
H_2	0	0	0
H_3	0	0	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

Construction from IBLTs

A First Attempt

$H_1(c_1)$	c_1	0	0
$H_2(c_1)$	c_1	0	0
$H_3(c_1)$	0	c_1	0

$$c_1 = \mathsf{Enc}(m_1)$$

$$c_2 = \mathsf{Enc}(0)$$

$$c_3 = \mathsf{Enc}(m_3)$$

$$c_4 = \mathsf{Enc}(m_4)$$

Construction from IBLTs

A First Attempt

H_1	$c_1 + c_2$	0	$c_3 + c_4$
H_2	$c_1 + c_4$	$c_2 + c_3$	0
H_3	$c_3 + c_4$	c_1	c_2

$$c_1 = \mathsf{Enc}(m_1)$$

$$c_2 = \mathsf{Enc}(0)$$

$$c_3 = \mathsf{Enc}(m_3)$$

$$c_4 = \mathsf{Enc}(m_4)$$

Construction from IBLTs

A First Attempt

H_1	$m_1 + 0$	0	$m_3 + m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0
H_3	$m_3 + m_4$	m_1	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

Construction from IBLTs

A First Attempt

$H_1(c_3) = ??$	$m_1 + 0$	0	$m_3 + m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0
$H_3(c_3) = ??$	$m_3 + m_4$	m_1	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

Construction from IBLTs

A First Attempt

H_1	$m_1 + 0$	0	$m_3 + m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0
H_3	$m_3 + m_4$	m_1	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

How to identify "one-entries"?

Need to evaluate predicate "$m_i \neq 0$" to maintain count matrix.

Construction from IBLTs

Wunderbar Pseudorandom Vectors

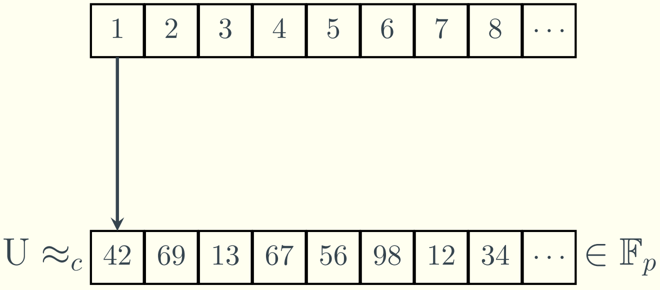
1	2	3	4	5	6	7	8	$\cdots$

$$\mathrm{U} \approx_c \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline 42 & 69 & 13 & 67 & 56 & 98 & 12 & 34 & \cdots \\ \hline\end{array} \in \mathbb{F}_p$$

Description of size $O(\lambda)$

Construction from IBLTs

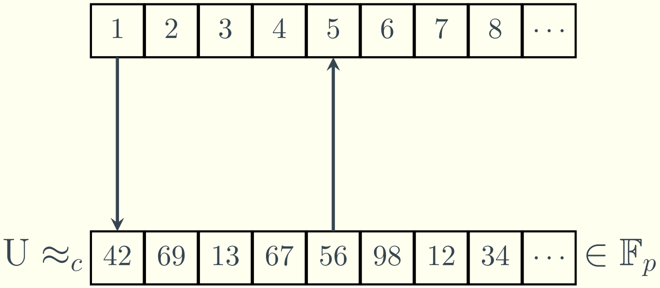
Wunderbar Pseudorandom Vectors



Description of size $O(\lambda)$

Construction from IBLTs

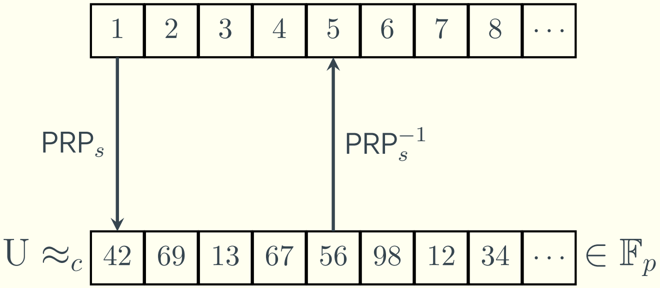
Wunderbar Pseudorandom Vectors



Description of size $O(\lambda)$

Construction from IBLTs

Wunderbar Pseudorandom Vectors



Description of size $O(\lambda)$

Construction from IBLTs

The Real Deal

H_1	0	0	0	0	0	0
H_2	0	0	0	0	0	0
H_3	0	0	0	0	0	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

Construction from IBLTs

The Real Deal

$H_1(1)$	c_1	0	0
$H_2(1)$	c_1	0	0
$H_3(1)$	0	c_1	0

$k_1 c_1$	0	0
$k_1 c_1$	0	0
0	$k_1 c_1$	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

$$k_1 := \mathsf{PRP}_s(1)$$

Construction from IBLTs

The Real Deal

H_1	$c_1 + c_2$	0	$c_3 + c_4$
H_2	$c_1 + c_4$	$c_2 + c_3$	0
H_3	$c_3 + c_4$	c_1	c_2

$k_1 c_1 + k_2 c_2$	0	$k_3 c_3 + k_4 c_4$
$k_1 c_1 + k_4 c_4$	$k_2 c_2 + k_3 c_3$	0
$k_3 c_3 + k_4 c_4$	$k_1 c_1$	$k_2 c_2$

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

$$k_i := \mathsf{PRP}_s(i)$$

Construction from IBLTs

The Real Deal

H_1	$m_1 + 0$	0	$m_3 + m_4$	$k_1 m_1 + 0$	0	$k_3 m_3 + k_4 m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0	$k_1 m_1 + k_4 m_4$	$0 + k_3 m_3$	0
H_3	$m_3 + m_4$	m_1	0	$k_3 m_3 + k_4 m_4$	$k_1 m_1$	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

$$k_i := \mathsf{PRP}_s(i)$$

Construction from IBLTs

The Real Deal

H_1	$m_1 + 0$	0	$m_3 + m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0
H_3	$m_3 + m_4$	m_1	0

$k_1 m_1 + 0$	0	$k_3 m_3 + k_4 m_4$
$k_1 m_1 + k_4 m_4$	$0 + k_3 m_3$	0
$k_3 m_3 + k_4 m_4$	$k_1 m_1$	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

$$k_i := \mathsf{PRP}_s(i)$$

Construction from IBLTs

The Real Deal

H_1	$m_1 + 0$	0	$m_3 + m_4$
H_2	$m_1 + m_4$	$0 + m_3$	0
H_3	$m_3 + m_4$	[illegible]	[illegible]

$k_1 m_1 + 0$	0	$k_3 m_3 + k_4 m_4$
$k_1 m_1 + k_4 m_4$	$0 + k_3 m_3$	0
$k_3 m_3$ [illegible]	[illegible]	0

$c_1 = \text{Enc}(m_1)$

$c_2 = \text{Enc}(0)$

$c_3 = \text{Enc}(m_3)$

$c_4 = \text{Enc}(m_4)$

$$\text{PRP}_s^{-1}\left(\frac{k_3 m_3}{m_3}\right) = 3 \overset{?}{\in} [n]$$

$k_i :=$ [illegible]

Construction from IBLTs

The Real Deal

$H_1(3)$	$m_1 + 0$	0	$m_3 + m_4$	$k_1 m_1 + 0$	0	$k_3 m_3 + k_4 m_4$
$H_2(3)$	$m_1 + m_4$	$0 + m_3$	0	$k_1 m_1 + k_4 m_4$	$0 + k_3 m_3$	0
$H_3(3)$	$m_3 + m_4$	[illegible]	0	$k_3 m_3$ [illegible]	[illegible]	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

$$\mathsf{PRP}_s^{-1}\left(\frac{k_3 m_3}{m_3}\right) = 3 \overset{?}{\in} [n]$$

$k_i :=$ [illegible]

Construction from IBLTs

The Real Deal

$H_1(3)$	$m_1 + 0$	0	m_4	$k_1 m_1 + 0$	0	$k_4 m_4$	$c_1 = \text{Enc}(m_1)$
$H_2(3)$	$m_1 + m_4$	0	0	$k_1 m_1 + k_4 m_4$	0	0	$c_2 = \text{Enc}(0)$
$H_3(3)$	m_4	m_1	0	$k_4 m_4$	$k_1 m_1$	0	$c_3 = \text{Enc}(m_3)$
							$c_4 = \text{Enc}(m_4)$

$$k_i := \text{PRP}_s(i)$$

Construction from IBLTs

The Real Deal

$H_1(3)$	$m_1 + 0$	0	m_4	$k_1 m_1 + 0$	0	$k_4 m_4$
$H_2(3)$	$m_1 + m_4$	0	0	$k_1 m_1 + k_4 m_4$	0	0
$H_3(3)$	m_4	m_1	0	$k_4 m_4$	$k_1 m_1$	0

$c_1 = \mathsf{Enc}(m_1)$

$c_2 = \mathsf{Enc}(0)$

$c_3 = \mathsf{Enc}(m_3)$

$c_4 = \mathsf{Enc}(m_4)$

Can be replaced by a single PRF.

Comparison with Previous Work

	Size	Compression	Decompression
[AFS19]	$O(t^2 \log n)$	🙁	🙁
[LT21]	$O(\varepsilon t \log t)$	🙁	🙁
[CDGLY21]	$O(\varepsilon t)$	🙂	🙂
Polynomials	$O(t)$	😐	🙁
IBLTs	$O(\varepsilon t / \log t)$	🙂	🙂

Vectors of length n and sparsity t decompress correctly with probability $1 - 2^{-\varepsilon}$.

Comparison with Previous Work

	Size	Compression	Decompression
[AFS19]	$O(t^2 \log n)$	☹	☹
[LT21]	$O(\varepsilon t \log t)$	☹	☹
[CDGLY21]	$O(\varepsilon t)$	☺	☺
Polynomials	$O(t)$	😐	☹
IBLTs	$O(\varepsilon \log \varepsilon + t)$	☺	☺

Coming soon to an ePrint near you.

Vectors of length n and sparsity t decompress correctly with probability $1 - 2^{-\varepsilon}$.