Collision Attacks on Up to 5 Rounds of SHA-3 Using Generalized Internal Differentials

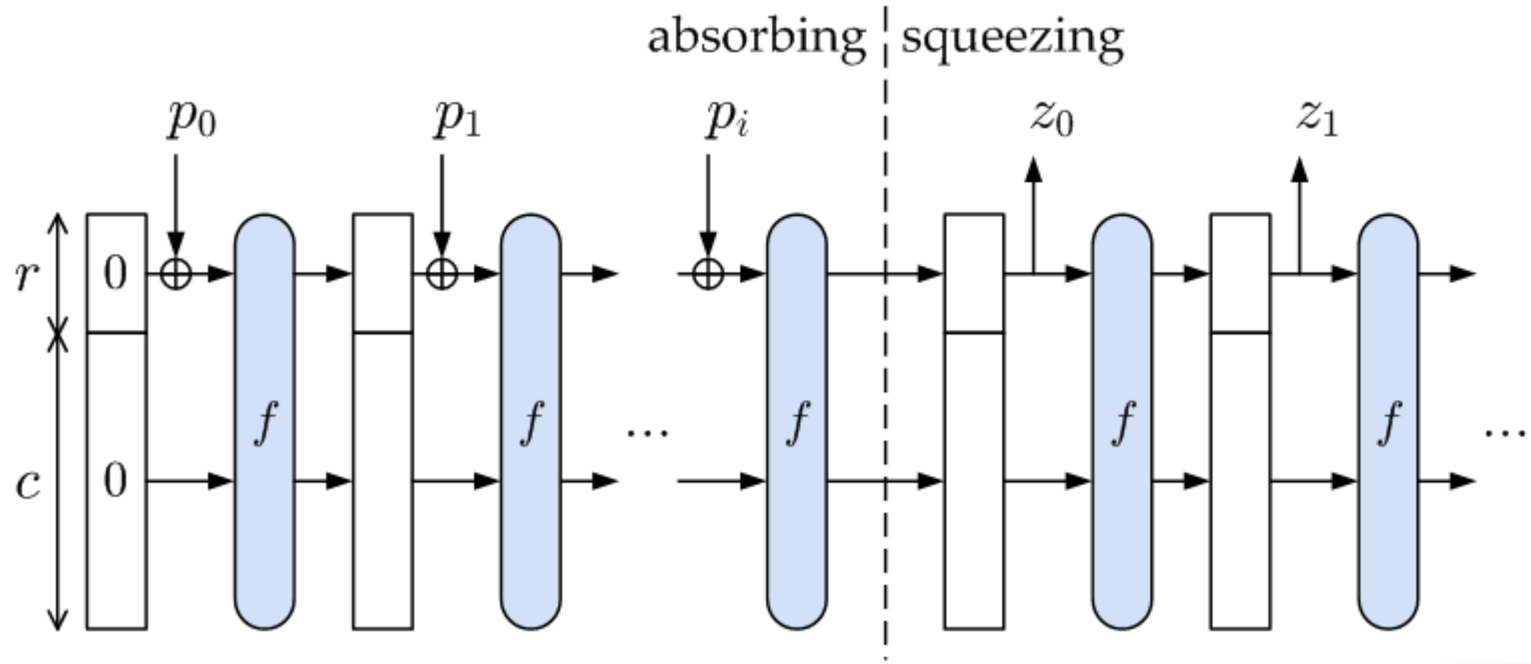
Itai Dinur[1], Orr Dunkelman[1,2] and Adi Shamir[1]

[1]The Weizmann Institute, Israel

[2]University of Haifa, Israel

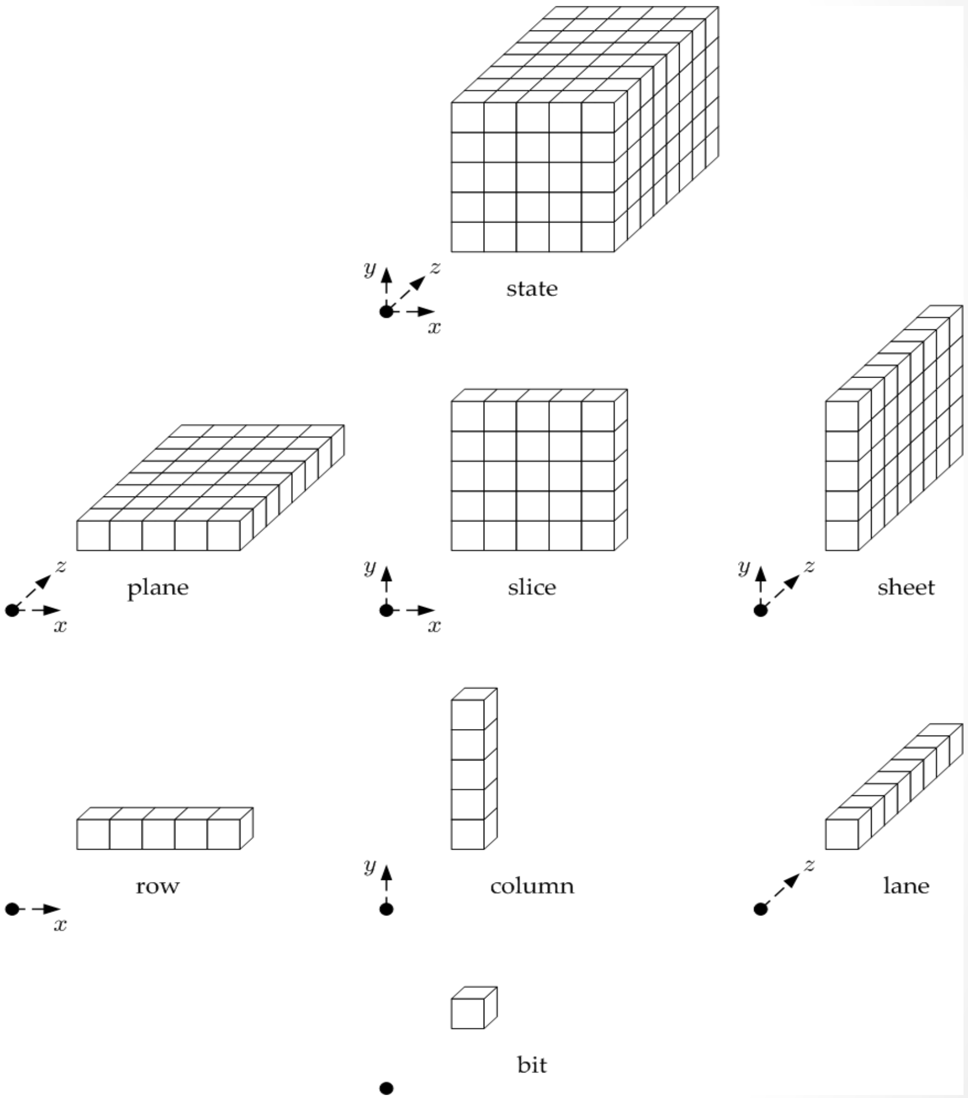
Keccak
(Bertoni, Daemen, Peeters and Van Assche)

- Uses the **sponge construction**



Keccak
The Inner State

- Can be viewed as a 5x5x64-bit cube
- Or as a 5x5 matrix, where each cell is a 64-bit lane in the direction of the z axis



Keccak
The function f

- f is a **24-round** permutation on the 1600-bit state

- Each round consists of 5 mappings $R = \iota \circ \chi \circ \pi \circ \rho \circ \Theta$

- We denote $L = \pi \circ \rho \circ \Theta$ and refer to L as a "**half-round**", where $\iota \circ \chi$ make up the other half

Keccak

The function f

- χ is the only **non-linear** mapping of Keccak
  - It has an algebraic degree of 2
- ι adds a **low Hamming-weight** round constant to the state
- The state is initialized to **zero** before the XOR with the first message block

Keccak
Collision Attacks on Round-Reduced Keccak

- “Practical analysis of reduced-round Keccak” by Naya-Plasencia, Röck and Meier (Indocrypt 2011)
  - Collisions in 2 rounds of **Keccak-224** and **Keccak-256**
- “New attacks on Keccak-224 and Keccak-256” by Dinur, Dunkelman and Shamir (FSE 2012)
  - Collisions in 4 rounds of **Keccak-224** and **Keccak-256**
- **No** published collision attack on **Keccak-384** and **Keccak-512**

Keccak

Our New Results

- **Keccak-512**: A 3-round **practical** collision attack
- **Keccak-384**: A 3-round **practical** collision attack
- A 4-round collision attack (faster than the birthday bound by 2^{45})
- **Keccak-256**: A 5-round collision attack (faster than the birthday bound by 2^{13})

	Keccak-224	Keccak-256	Keccak-384	Keccak-512
Previous	4 (practical)	4 (practical)	-	-
New	-	5 (2^{115})	3 (practical) 4 (2^{147})	3 (practical)

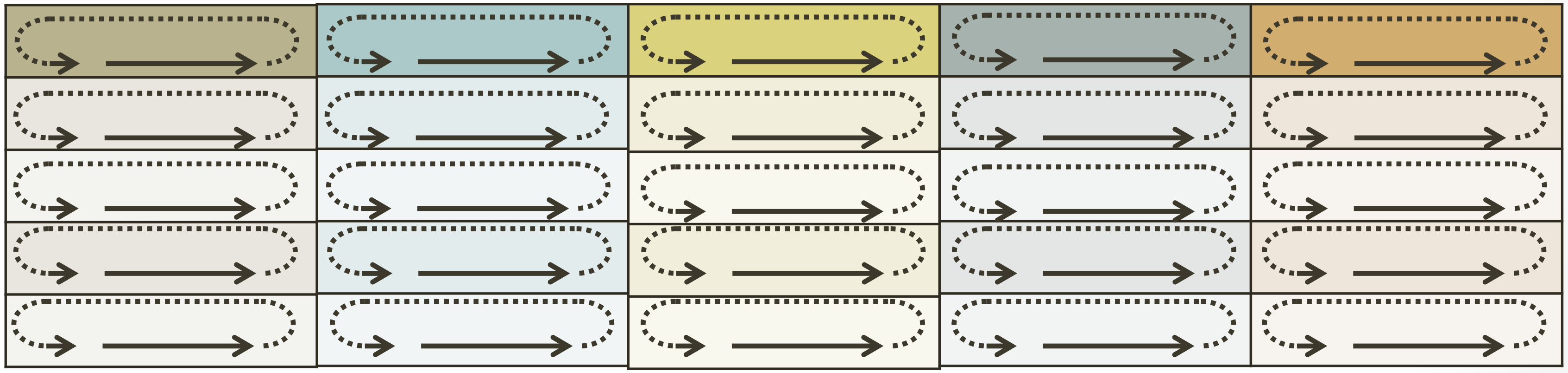
Keccak

The Translation-Invariance Property

- Defined in the Keccak submission document
- 4 out of the 5 internal mappings (all but ι) are translation invariant in the direction of the z axis (of length 64)

Keccak
The Translation-Invariance Property

- If one state is the rotation of the other with respect to the z-axis, then applying to them any of the Θ,ρ,π,χ operations, **maintains this property**



Symmetric States

- A state which is **rotation-invariant** in the direction of the z axis by some **rotation index** i is called a **symmetric state**
- i can attain non-trivial values that divide the lane size 64 ($i \in \{1,2,4,8,16,32\}$)

Consecutive Slice Sets

An example

- For i=16 we split the state into **4 consecutive slice sets (CSS)**

a_1				b_1				c_1				d_1				e_1			
f_1				g_1				h_1				i_1				j_1			
k_1				l_1				m_1				n_1				o_1			
p_1				q_1				r_1				s_1				t_1			
u_1				v_1				w_1				x_1				y_1			

	a_2				b_2				c_2				d_2				e_2		
	f_2				g_2				h_2				i_2				j_2		
	k_2				l_2				m_2				n_2				o_2		
	p_2				q_2				r_2				s_2				t_2		
	u_2				v_2				w_2				x_2				y_2		

Symmetric States
An Example

- In symmetric states all CSS's are equal
- In a symmetric state with i=16, each 64-bit lane is composed of a 4-repetition of a 16-bit value

a_1	**a_1**	**a_1**	**a_1**	**b_1**	**b_1**	**b_1**	**b_1**	**c_1**	**c_1**	**c_1**	**c_1**	**d_1**	**d_1**	**d_1**	**d_1**	**e_1**	**e_1**	**e_1**	**e_1**
f_1	f_1	f_1	f_1	g_1	g_1	g_1	g_1	h_1	h_1	h_1	h_1	i_1	i_1	i_1	i_1	j_1	j_1	j_1	j_1
k_1	k_1	k_1	k_1	l_1	l_1	l_1	l_1	m_1	m_1	m_1	m_1	n_1	n_1	n_1	n_1	o_1	o_1	o_1	o_1
p_1	p_1	p_1	p_1	q_1	q_1	q_1	q_1	r_1	r_1	r_1	r_1	s_1	s_1	s_1	s_1	t_1	t_1	t_1	t_1
u_1	u_1	u_1	u_1	v_1	v_1	v_1	v_1	w_1	w_1	w_1	w_1	x_1	x_1	x_1	x_1	y_1	y_1	y_1	y_1

Symmetric states remain symmetric after applying the Θ,ρ,π,χ operations

a_1	a_1	a_1	a_1	b_1	b_1	b_1	b_1	c_1	c_1	c_1	c_1	d_1	d_1	d_1	d_1	e_1	e_1	e_1	e_1
f_1	f_1	f_1	f_1	g_1	g_1	g_1	g_1	h_1	h_1	h_1	h_1	i_1	i_1	i_1	i_1	j_1	j_1	j_1	j_1
k_1	k_1	k_1	k_1	l_1	l_1	l_1	l_1	m_1	m_1	m_1	m_1	n_1	n_1	n_1	n_1	o_1	o_1	o_1	o_1
p_1	p_1	p_1	p_1	q_1	q_1	q_1	q_1	r_1	r_1	r_1	r_1	s_1	s_1	s_1	s_1	t_1	t_1	t_1	t_1
u_1	u_1	u_1	u_1	v_1	v_1	v_1	v_1	w_1	w_1	w_1	w_1	x_1	x_1	x_1	x_1	y_1	y_1	y_1	y_1

↓ Θ,ρ,π,χ

a_2	a_2	a_2	a_2	b_2	b_2	b_2	b_2	c_2	c_2	c_2	c_2	d_2	d_2	d_2	d_2	e_2	e_2	e_2	e_2
f_2	f_2	f_2	f_2	g_2	g_2	g_2	g_2	h_2	h_2	h_2	h_2	i_2	i_2	i_2	i_2	j_2	j_2	j_2	j_2
k_2	k_2	k_2	k_2	l_2	l_2	l_2	l_2	m_2	m_2	m_2	m_2	n_2	n_2	n_2	n_2	o_2	o_2	o_2	o_2
p_2	p_2	p_2	p_2	q_2	q_2	q_2	q_2	r_2	r_2	r_2	r_2	s_2	s_2	s_2	s_2	t_2	t_2	t_2	t_2
u_2	u_2	u_2	u_2	v_2	v_2	v_2	v_2	w_2	w_2	w_2	w_2	x_2	x_2	x_2	x_2	y_2	y_2	y_2	y_2

The Fifth Mapping

- ι **destroys** the perfect symmetry of the state by adding a non-symmetric round constant

An Overview of the Basic Attack

- Pick a single-block message such that the initial state is **symmetric**
- The state **remains symmetric** after the first 4 mappings
- The symmetry is **slightly perturbed** by the ι mapping since the constants added are of **low Hamming-weight** (between 1 and 5)
- The **diffusion** is sufficiently **slow** such that the state remains "close" to symmetric for the first few rounds

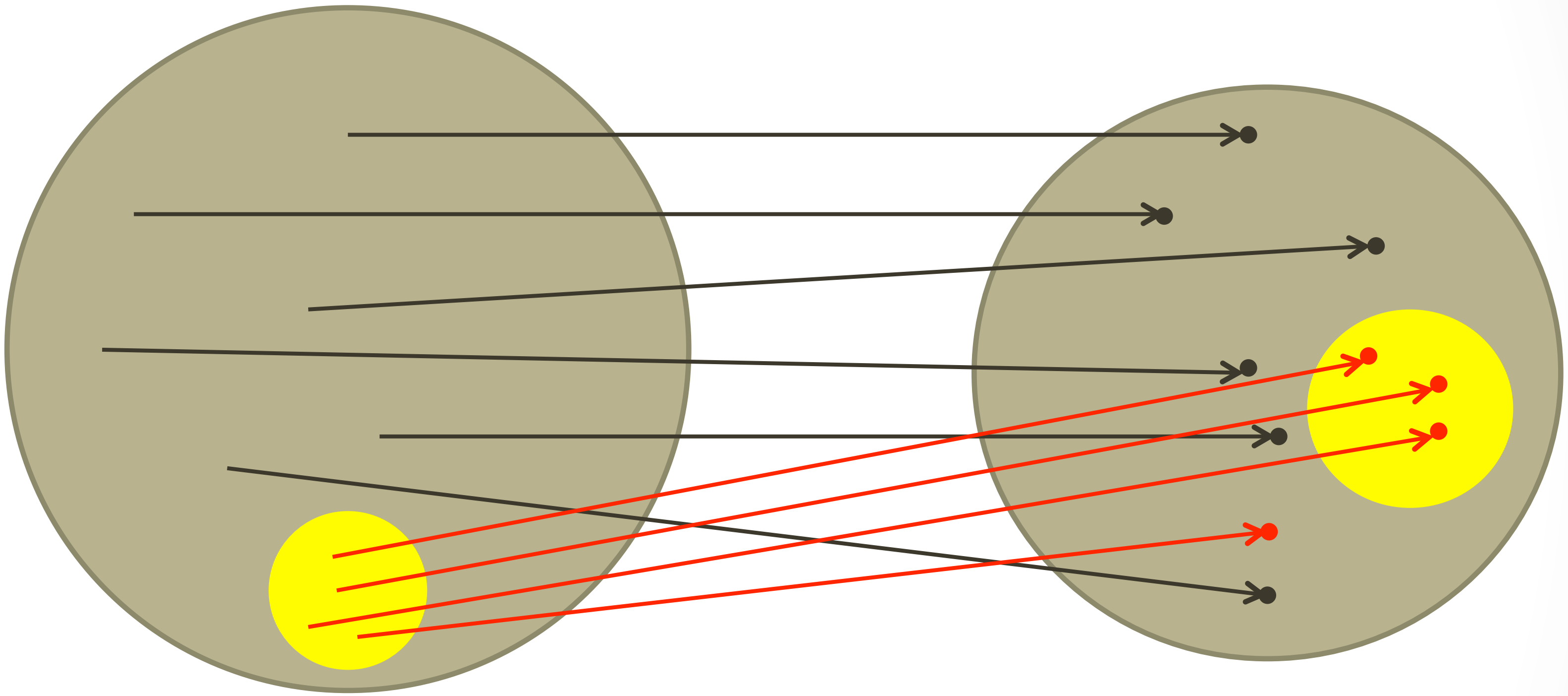
An Overview of the Basic Attack

The Squeeze Attack

- The **effective output size** for symmetric messages is **reduced**
- We use a natural attack (called the **squeeze attack**) that exploits this property
- We force a larger than expected number of inputs to squeeze into a **small subset** of possible outputs in which collisions are **more likely**

An Overview of the Basic Attack

The Squeeze Attack



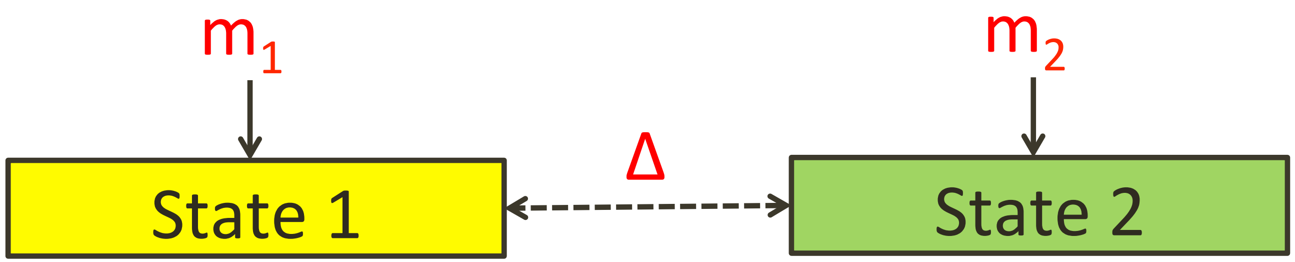
- A member of the input set is mapped with probability p to the output set of size D
- The time complexity of the attack is 1/p·√D

Subset Cryptanalysis

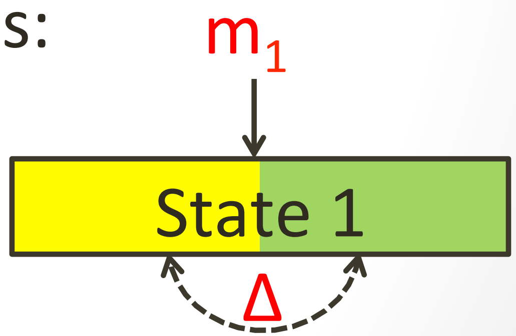
- In order to devise and analyze the attack we use a very common cryptanalysis framework which we call **subset cryptanalysis**
- Uses **subset characteristics** to track the evolution of subsets through the internal state of the cryptosystem
  - Associate a triplet (input subset, output subset, transition probability) to each internal operation

Internal Differential Cryptanalysis

- Introduced by Thomas Peyrin (Crypto 2010) in the analysis of Grostl
- Standard differential cryptanalysis:



- Internal differential cryptanalysis:



Generalized Internal Differential Cryptanalysis

- We generalize and extend it:
  - Shown to be applicable only to hash functions built using **separate data-paths**, whereas Keccak has only one **data-path**
  - The differences considered were between 2 parts of the state, whereas we consider more complex differential relations between **multiple parts of the state**

Internal Differences
Definitions

- In symmetric states all CSS's are equal
- In states which are almost symmetric the **differences** between the first CSS and the other 3 CSS's ($\Delta_1,\Delta_2,\Delta_3$) are of **low Hamming weight**
- We group all states with a **fixed** ($\Delta_1,\Delta_2,\Delta_3$) into an **internal difference set**

Internal Differences

Definitions

- Given a state *u*, the set

{*v*| *v=u+w* and *w* is symmetric} is an **internal difference set**

- The differences between the CSS's is specified by *u* which is a **representative state**
- A state *v* of a **lowest Hamming weight** defines the **weight** of the internal difference
- The **zero internal difference** contains the symmetric states and has a weight of 0

Internal Differential Characteristics

The Evolution of Internal Differences

- Any **symmetric** state chosen from the zero self-difference **remains symmetric** after applying Θ,ρ,π,χ
- Internal Differences are **affine subspaces**
- Their evolution through the 4 linear maps can be **easily tracked**

Internal Differential Characteristics

A 1.5-round Example

```
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|  R0
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
```

$$\downarrow R$$

```
|---------------1|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|  R1
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
```

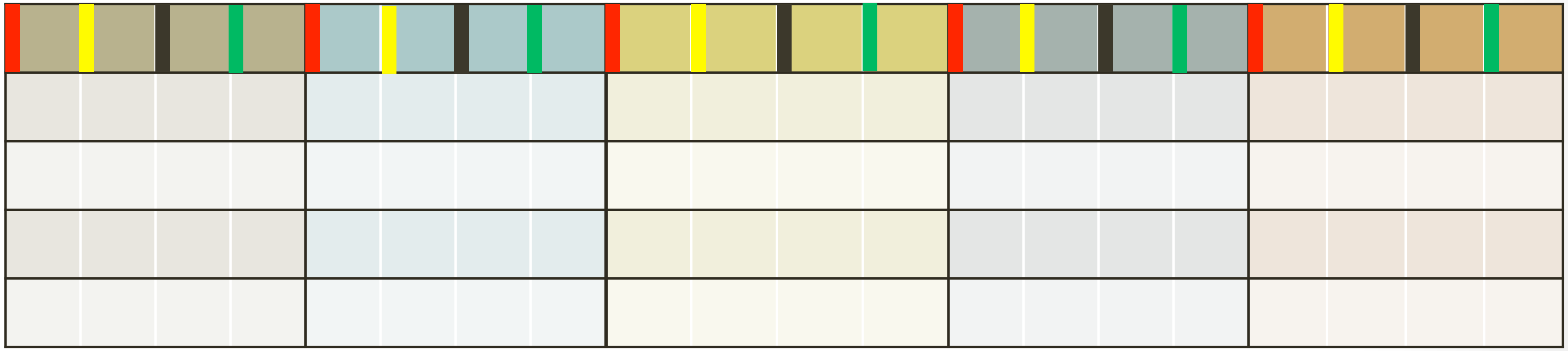
$$\downarrow L$$

```
|---------------1|----1-----------|----------------|----------------|------------8---|
|----------------|----------2-----|----------------|----2-----------|----------------|
|---------------2|----------------|----------------|-------------2--|----------------|  R1.5
|--------1-------|----------------|-------------4--|----------------|----------------|
|----------------|----------------|-----1----------|----------------|---------------4|
```

Internal Differential Characteristics

The Evolution of Internal Differences

- The evolution through χ is analyzed by considering **rotated row sets (RRS)**
- A RRS contains an Sbox (row) in the first CSS and its 3 **symmetric counterparts** in the other CSS's



Internal Differential Characteristics

The Evolution of Internal Differences

- The input internal difference specifies the **differences** between the Sboxes of **each RRS**
- Each RRS can assume **exactly** 32 values
- The distribution of the output internal difference can be **computed exhaustively**
- For i=32 the output internal difference can be analyzed in a similar way to **standard differential cryptanalysis**

Internal Differential Characteristics

The Evolution of Internal Differences

- A RRS with a zero input difference is called **non-active**
- A non-active RRS passes through χ with probability 1
- A **low-weight** internal difference passes through χ with high probability
- We look for internal differential characteristics composed of **low-weight** internal differences

Internal Differential Characteristics

Another 1.5-round Example

```
|--8------------|--4------------|---------------|---------------|---------------|
|--8------------|---------------|---------2-----|---------------|---------------|
|---------------|--4------------|---------2-----|---------------|---------------|  R0
|---------------|---------------|---------------|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
```

$$\downarrow \mathrm{L}$$

```
|--8------------|---------------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
|--8------------|--------8------|---------------|---------------|---------------|
|---------------|--------8------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
```

$$\downarrow \chi \quad (p = 2^{-12})$$

```
|--8-----------1|---------------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
|--8------------|--------8------|---------------|---------------|---------------|
|---------------|--------8------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
```

$$\downarrow \iota$$

```
|--8------------|---------------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
|--8------------|--------8------|---------------|---------------|---------------|  R1
|---------------|--------8------|--------------1|---------------|---------------|
|---------------|---------------|---------------|---------------|---------------|
```

$$\downarrow \mathrm{L}$$

```
|--8------------|---------------|---------------|---------------|---------------|
|---------------|---------------|-4-------------|------------1--|---------------|
|---------------|---------------|---------------|---------------|---------------|  R1.5
|---------------|---------------|------2--------|-----------8---|---------------|
|4--------------|---------------|---------------|---------------|---------------|
```

Internal Differential Characteristics

Extension to 2.5 rounds

```
|--8-------------|----------------|----------------|----------------|----------------|
|----------------|----------------|-4--------------|-------------1--|----------------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|------2---------|------------8---|----------------|
|4---------------|----------------|----------------|----------------|----------------|
```

$$\downarrow \chi \quad (p = 2^{-12})$$

```
|--8-------------|----------------|----------------|----------------|----------------|
|----------------|----------------|-4--------------|-------------1--|----------------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|------2---------|------------8---|----------------|
|4---------------|----------------|----------------|----------------|----------------|
```

$$\downarrow \iota$$

```
|--8---------8-82|----------------|----------------|----------------|----------------|
|----------------|----------------|-4--------------|-------------1--|----------------|
|----------------|----------------|----------------|----------------|----------------| R2
|----------------|----------------|------2---------|------------8---|----------------|
|4---------------|----------------|----------------|----------------|----------------|
```

$$\downarrow \mathrm{L}$$

```
|8-8------------8|-8-82488---4----|-81-------------|-4----1-----8---|--------6--12-4-|
|----------4----2|------18--481---|----------------|1-1-491----8----|------------2-4-|
|91----8----1-1-4|----------4-8--1|4----------8----|---------18--481|-----------1----| R2.5
|-----C--24-8----|----------------|---1-----2-2-922|--1-----81------|---4----2-------|
|------------4-8-|8--2----1-------|--C--24-8-------|----------------|22---1-----2-2-9|
```

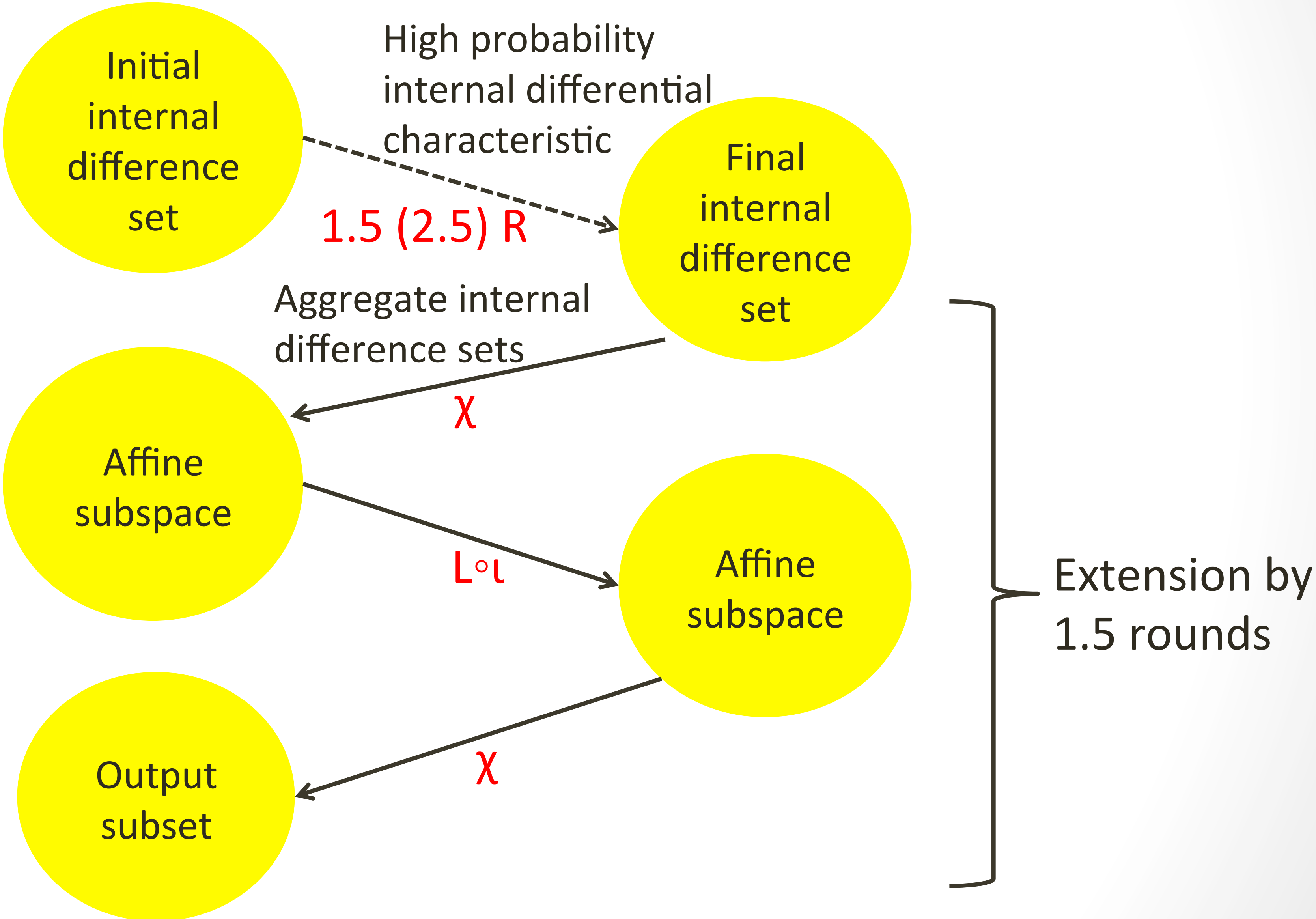
Extending Internal Differential Characteristics

- Given a characteristic that ends before the χ layer with a (relatively) **high weight** internal difference
- We extend it by 1.5 **additional** rounds to a **subset characteristic**
  - **Do not restrict** its subsets to specific internal differences
  - **Avoid** the reduction in probability

Extending Internal Differential Characteristics

- Exploit the **low algebraic degree** of χ to **aggregate** all the possible internal differences at the output of χ to a single **affine subspace**
- Extend the characteristic by 1 round using an **affine subspace**
- Use the **low diffusion** of χ to bound the output subset size after an additional 0.5 round

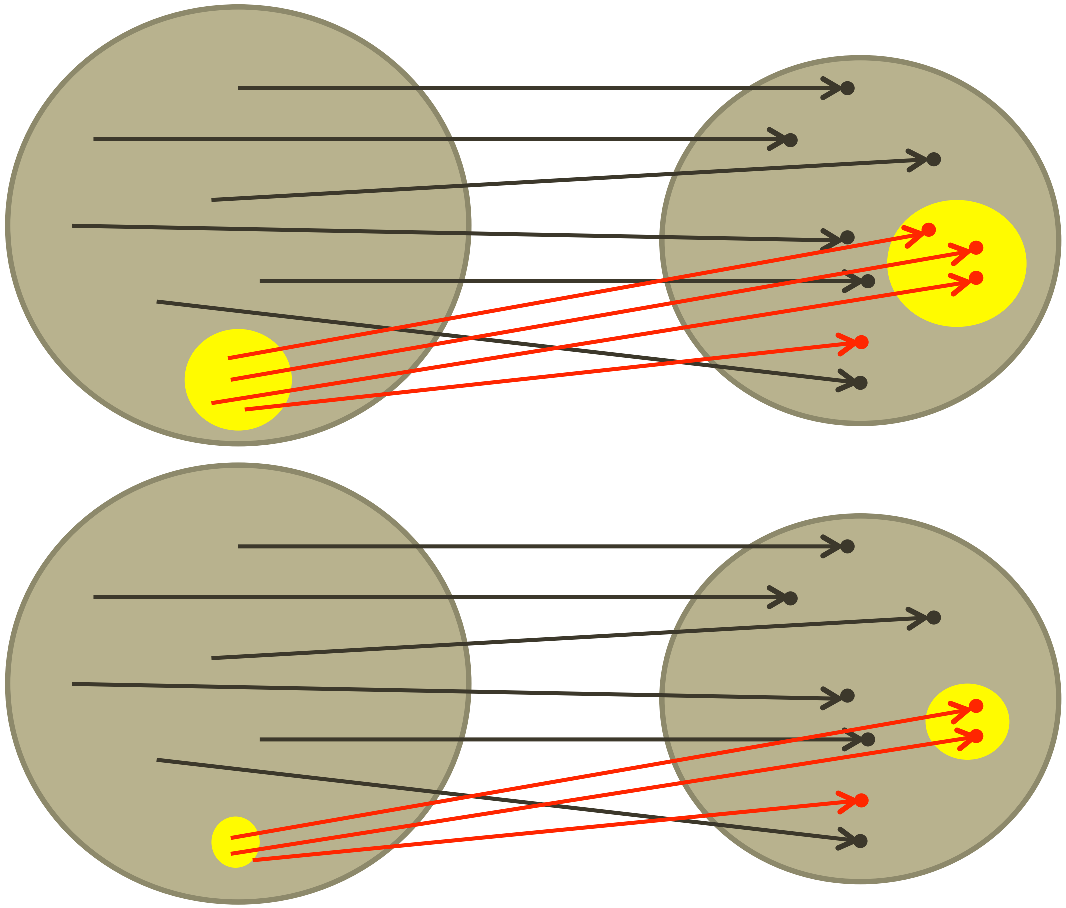
The Evolution of Subsets



Choosing the Rotation Index

- A **smaller** rotation index enforces **more symmetry relations**
  - **Reduces** the size of the **output** set
  - **Reduces** the size of the **input** set

Choosing the Rotation Index



- Choose the **smallest** value of i∈{1,2,4,8,16,32} for which the **input set** is **large** enough to find a **collision**

Collision Attacks
Practical Attacks

- A 3-round collision in **Keccak-512** (with rotation index i=4)

M1=

88888888 88888888 66666666 66666666 AAAAAAAA AAAAAAAA 77777777 77777777 BBBBBBBB BBBBBBBB BBBBBBBB BBBBBBBB 11111111 11111111 88888888 88888888 CCCCCCCC CCCCCCCC

M2=

AAAAAAAA AAAAAAAA 88888888 88888888 EEEEEEEE EEEEEEEE 99999999 99999999 99999999 99999999 99999999 99999999 88888888 88888888 CCCCCCCC CCCCCCCC CCCCCCCC CCCCCCCC

Output=

56BCC94B C4445644 D7655451 5DD96555 71FA7332 3BA30B23 958408C5 64407664 41805414 11190901 6ABAA8BA A8ABAEFA 7EF8AEEE ECCE68DC 4EC8ACEC DD5D5CCC

Collision Attacks
Practical Attacks

- A 3-round collision in **Keccak-384** (with rotation index i=4)

M1=

FFFFFFFF FF7FFFFF BBBBBBBB BBFBBBBB 44444444 44444444 FFFFFFFF FFFFFFFF 99999999 99999999 44444444 44C44444 44444444 44444444 44644444 44444444 AAAAAAAA AAAAAAAA 66666666 66666666 44444444 44444444 DDDDDDDD DD9DDDDD DDFDDDDD DDDDDDDD

M2=

33333333 33B33333 55555555 55155555 AAAAAAAA AAAAAAAA 77777777 77777777 44444444 44444444 66666666 66E66666 EEEEEEEE EEEEEEEE 11311111 11111111 CCCCCCCC CCCCCCCC FFFFFFFF FFFFFFFF 11111111 11111111 99999999 99D99999 DDFDDDDD DDDDDDDD

Output=

99999991 11199999 4440C444 405C60DC 00000000 0C100010 777677F7 73F77767 3550F597 55D57155 66666664 66666666

Collision Attacks

- The 2.5-round characteristic is used in a 4-round collision attack on **Keccak-384**
- The time complexity is 2^{147} (faster than the birthday bound by 2^{45})

Collision Attacks

- A 5-round collision attack on **Keccak-256**
- Based on a **target internal difference algorithm**
  - An extension of the **target difference algorithm** (FSE 2012)

The Target Internal Difference

Extending a Characteristic Backwards

```
|BCE-1EDC--------|68A-49EE--------|4A6-2999--------|916-3-39--------|68A--D8---------|
|BCE-9E54--------|68A-49EA--------|4A6-2999--------|916-3-99--------|6CA--D8---------|
|FCE-9E54--------|68A-49EE--------|4A2-2999--------|916-3-99--------|68A--D8---------|
|BCE-9E54--------|68A-49EE--------|4A6-2999--------|916-3-99--------|68A--D8---------|
|BCE-9E54--------|68A-49EE--------|4A6-2B99--------|916-3-99--------|69A--D8---------|
```

$\uparrow$ **R** (p=2^{-16})

```
|-------8--------|----4-----------|-------2--------|----------------|------4---------|
|-------8--------|----4-----------|-------2--------|----------------|------4---------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
|----------------|----------------|----------------|----------------|----------------|
```

$\downarrow \mathrm{L}$

```
|-------8--------|-4--------------|----------------|----------------|----------------|
|----------------|-4--------------|----------------|----------------|----------------|
|----8-----------|------8---------|----------------|----------------|----------------|
|-------2--------|------8---------|----------------|----------------|----------------|
|8---------------|----------------|----------------|----------------|----------------|
```

$\downarrow \chi \quad (p = 2^{-16})$

```
|-------8--------|-4--------------|----------------|----------------|----------------|
|----------------|-4--------------|----------------|----------------|----------------|
|----8-----------|------8---------|----------------|----------------|----------------|
|------82--------|------8---------|----------------|----------------|----------------|
|8---------------|----------------|----------------|----------------|----------------|
```

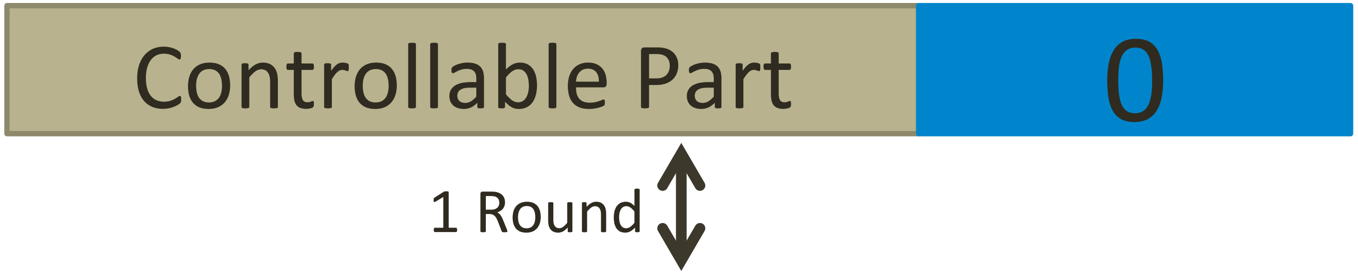
$\downarrow \iota$

```
|8---8-82--------|-4--------------|----------------|----------------|----------------|
|----------------|-4--------------|----------------|----------------|----------------|
|----8-----------|------8---------|----------------|----------------|----------------|
|------82--------|------8---------|----------------|----------------|----------------|
|8---------------|----------------|----------------|----------------|----------------|
```

$\downarrow \mathrm{L}$

```
|8---8-82--------|------4---------|----------------|----------------|----------------|
|----------------|----------------|---4------------|--1-------------|----------------|
|-8--------------|----------------|----------------|----------------|---2------------|
|----------------|----------------|---2------------|----------------|----------------|
|----------------|----------------|----------------|---1-4----------|----------------|
```

The Target Internal Difference Algorithm
Linking a Characteristic Form an Initial State




```
|BCE-1EDC--------|68A-49EE--------|4A6-2999--------|916-3-39--------|68A--D8---------|
|BCE-9E54--------|68A-49EA--------|4A6-2999--------|916-3-99--------|6CA--D8---------|
|FCE-9E54--------|68A-49EE--------|4A2-2999--------|916-3-99--------|68A--D8---------|
|BCE-9E54--------|68A-49EE--------|4A6-2999--------|916-3-99--------|68A--D8---------|
|BCE-9E54--------|68A-49EE--------|4A6-2B99--------|916-3-99--------|69A--D8---------|
```

- The time complexity of the optimized attack is about 2^{115}
  - Faster than the birthday bound by 2^{13}

Conclusions and Future Work

- We presented the **first collision attacks** on round reduced **Keccak-384** and **Keccak-512**
  - Some of them are practical
- For **Keccak-256** we **increased** the number of rounds that can be attacked from 4 to 5
  - We are still **very far** from attacking the full 24 rounds
- An interesting future work item is to find **better** internal differential characteristics for Keccak **or** to **prove** that they do not exist

Thank you for your attention!