Evolution of Cryptanalysis: Security Updates on SHA-3 and AES Hashing

Jian Guo

CRYPTANALYSIS TASKFORCE

NANYANG TECHNOLOGICAL UNIVERSITY

ASIACRYPT 2022 @ Taipei
7th December 2022

Acknowledgements

Many thanks go to my collaborators on these topics:

Zhenzhen Bao
Colin Chaigneau
Lin Ding
Le Dong
Xiaoyang Dong
Alexandre Duc
Dengguo Feng
Thomas Fuhr
Henri Gilbert
Jérémy Jean
Guohong Liao
Shun Li
Zheng Li
San Ling
Guozhen Liu
Meicheng Liu
Phuong Pham
Thomas Peyrin
Kexin Qiao
Jean-René Reinhard
Danping Shi
Ling Song
Siwei Sun
Yi Tu
Haoyang Wang
Xiaoyun Wang
Lei Wei
Shuang Wu
Wenling Wu
Wenying Zhang
Jian Zou

Outlines

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

Outline

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

Introduction to Hash Functions

Definition

$h : \{0,1\}^* \longrightarrow \{0,1\}^n$, a function mapping a bit string of arbitrary length to a fixed length n-bit digest.

Security Properties

- **Collision Resistance**: it is computationally difficult to find x, x' such that $h(x) = h(x')$.
- **Preimage Resistance**: given t, it is computationally difficult to find x such that $h(x) = t$.
- **Second-Preimage Resistance**: given x, it is computationally difficult to find x' such that $h(x) = h(x')$.

Applications

digital signatures ($sign_{key}(h(m))$), password storage ($h(salt, password)$), checksum ($h(m)$), Message Authentication Code ($h(key, m)$)

SHA-3

Triggered by the broken of SHA-1 in 2005 by Xiaoyun Wang et al., NIST organized the SHA-3 competition

timeline

- 2008/10: 64 submissions received
- 2009/07: 14 selected for Round 2
- 2010/12: 5 selected for Round 3
- 2012/10: Keccak announced as the winner, designed by Guido Bertoni, Joan Daemen, Michael Peeters, and Gilles Van Assche.
- 2015/08: formally standardized as "SHA-3"

SHA-3 basics

- supports 224, 256, 384, 512-bit digest sizes, and extendable digest sizes by SHAKE-128 and SHAKE-256.
- design strategy different from SHA-1 and SHA-2
- 24 rounds; designers also proposed KangarooTwelve with 12 rounds.

AES Hashing

DES has to be replaced due to linear cryptanalysis by Matsui 1993, and other issues, NIST organized AES competition

timeline

- 1998/06: 15 submissions received
- 1999/08: 5 selected for Round 2
- 2000/10: Rijndael announced as winner, designed by Joan Daemen, Vincent Rijmen
- 2001/11: formally standardized as "AES"

AES basics

- supports 128, 192, 256-bit keys
- wide trail strategy to resist differential cryptanalysis
- 10/12/14 rounds; can be converted to hash function by modes such as Davies-Meyer $H_i = E_M(H_{i-1}) + M$

Outline

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

Introduction to Cryptanalysis

Cryptanalysis methods develop together with designs, serve as a tool to trade-off between simplicity, efficiency, and security for designs, through round number etc.

Cryptanalysis Methods

differential cryptanalysis, linear cryptanalysis, integral cryptanalysis, meet-in-the-middle attack, boomerang attack, cube attack, etc.

Take differential cryptanalysis as example, the search of high probability differentials are done:

Evolution

- by hand according to cryptanalysts' experiences, e.g., SHA-1 in 2005 ?
- by dedicated search programs
- by SAT like solvers (from Mouha et al. Inscrypt 2011)
- by Machine Learning (from Gohr CRYPTO 2019)

Search of Differentials

The Problem

Given the round number r, find

$$\Delta_{IN} \rightarrow \Delta_1 \rightarrow \Delta_2 \cdots \rightarrow \Delta_r = \Delta_{OUT}$$

such that the overall probability from $\Delta_{IN} \longrightarrow \Delta_{OUT}$ is optimal.

Approaches

- Human experiences used to narrow down the search space: truncated differential search then differential; find good differential gradually by round numbers; meet in the middle approaches, etc.
- SAT-like solvers (solving time hard to predict)
- Machine Learning (constrained by practical complexities)

Optimization by SAT-like solvers

It is hard to achieve optimality for any cryptanalytic attacks, which are usually tunable by various parameters. SAT-like solvers, e.g., SAT, Constrained Programming (CP), Mixed-Integer-Linear-Programming (MILP), work better for man-made search programs.

Simple MILP Example

$$-2x + 2y \geq 1$$
$$-8x + 10y \leq 13$$
$$x, y \geq 0$$
$$x, y \in \mathcal{Z}$$

Objective: maximize $x + y$

Optimal solution is $(x, y) = (1, 2)$ with objective 3.

Differential Search

Constraints: $\Delta_{i-1} \to \Delta_i$ has to be compatible for all i.
Optimization Goal: Overall probability.

Quantum Cryptanalysis

Assuming quantum computer in hand, two major algorithms available

Grover's Algorithm, a.k.a Quantum Search Algorithm

It takes $\Theta(\sqrt{N})$ time by quantum computer to search the one target out of a space of size N, v.s. $\Theta(N)$ by classical computers.
Impacts:

- security strength of block ciphers is halved, 128-bit key becomes insufficient
- (second-) preimage resistance of hash functions is halved

Simon's Algorithm

If the function f is periodic, i.e., $f(x) = f(x + p)$ for all x from the domain, it costs polynomial time to find the period p.

Collision Finding

Collision Attack in Classical Setting

One randomly select x and store $h(x)$ in a hash table until a repeated value detected, this costs $2^{n/2}$ with a hash function with n-bit digest.

Collision Attack in Quantum Setting

One randomly pick $2^{n/3}$ x and store the respective $h(x)$ in quantum memory, and carry out Grover search of $h(y)$ against the stored value, this costs $2^{n/3}$.

- 128-bit digest offers collision resistance of $2^{42.7}$, insufficient ...

Impact on differential-based collision attacks

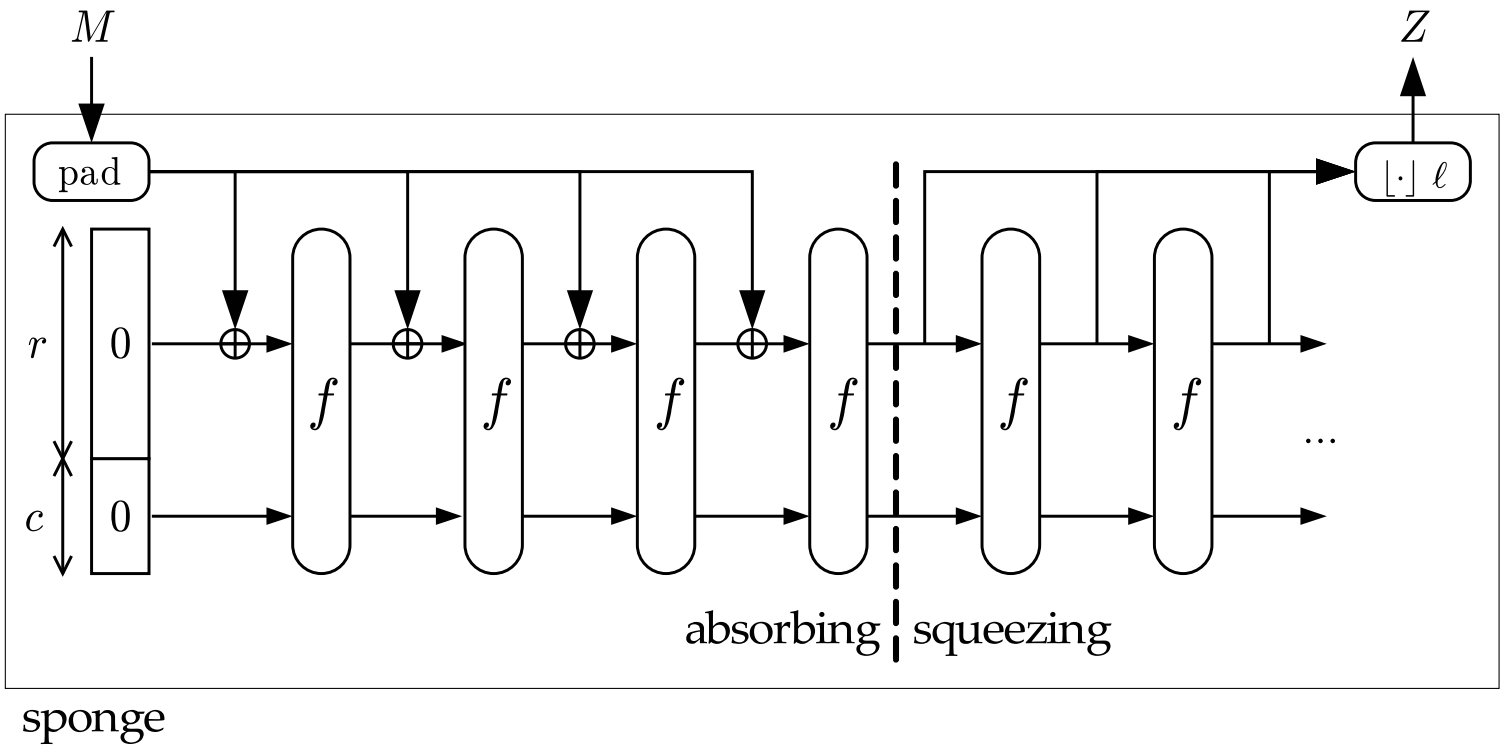
Lowest useful probability is $2^{-n/2}$ (Classical) v.s. $2^{-2n/3}$ (Quantum); Potentially larger search space, and longer attacked rounds.

Outline

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

SHA-3 (Keccak) Hash Function

The sponge construction [BDPV11]



- b-bit permutation f
- Two parameters: bitrate r, capacity c, and $b = r + c$.
- The message is padded and then split into r-bit blocks.

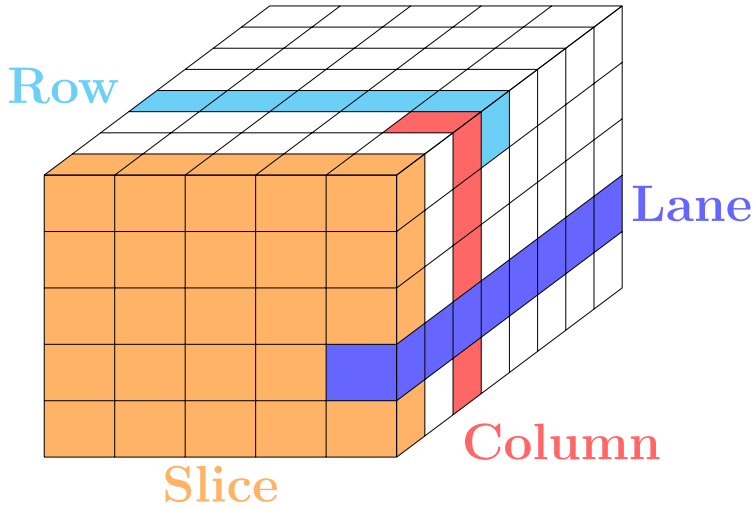
SHA-3 Hash Function

Keccak-f permutation

- 1600 bits: seen as a 5×5 array of 64-bit lanes, $A[x, y], 0 \leq x, y < 5$
- 24 rounds
- each round R consists of five steps:

$$R = \iota \circ \chi \circ \pi \circ \rho \circ \theta$$

- χ : the only nonlinear operation



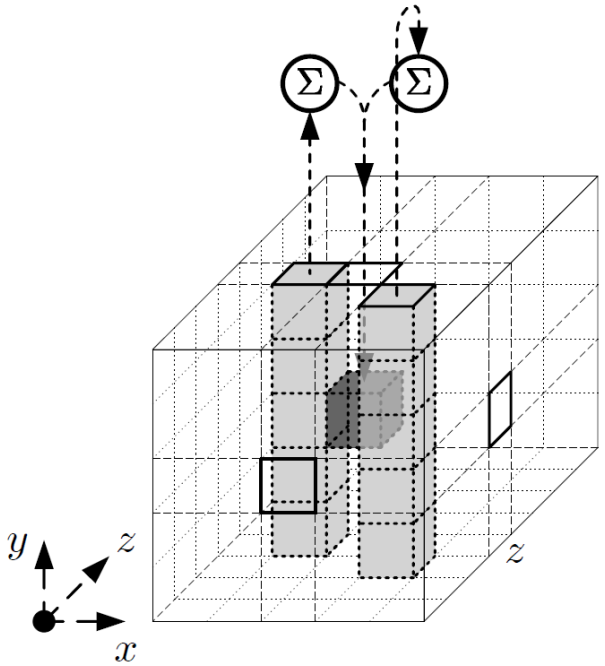
http://www.iacr.org/authors/tikz/

SHA-3 Hash Function

Keccak permutation: $\iota \circ \chi \circ \pi \circ \rho \circ \theta$

θ step: adding two columns to the current bit

$$
\begin{aligned}
C[x] =& A[x,0] \oplus A[x,1] \oplus A[x,2] \oplus \\
& A[x,3] \oplus A[x,4] \\
D[x] =& C[x-1] \oplus (C[x+1] \lll 1) \\
A[x,y] =& A[x,y] \oplus D[x]
\end{aligned}
$$



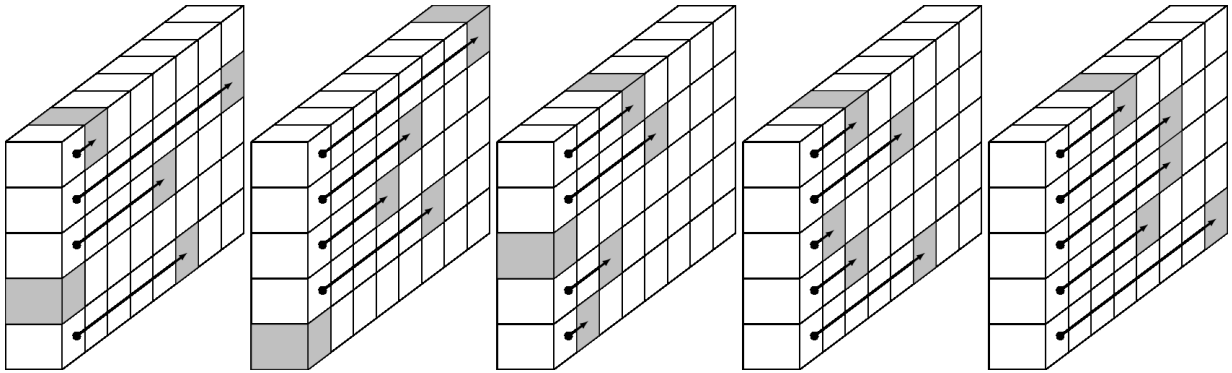
http://keccak.noekeon.org/

- The Column Parity kernel
  - If $C[x] = 0, 0 \leq x < 5$, then the state A is in the CP kernel.

SHA-3 Hash Function

KECCAK permutation: $\iota \circ \chi \circ \pi \circ \rho \circ \theta$

ρ step: lane level rotations, $A[x, y] = A[x, y] \lll r[x, y]$



http://keccak.noekeon.org/

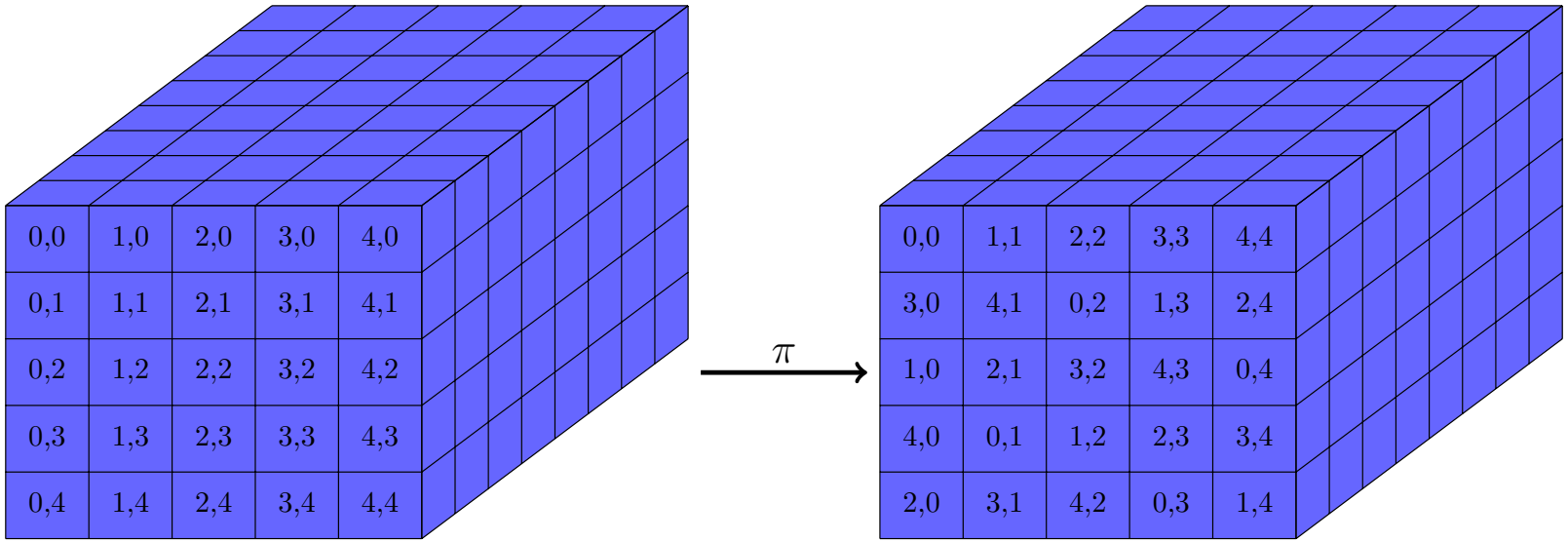
Rotation offsets $r[x, y]$

	$x = 0$	$x = 1$	$x = 2$	$x = 3$	$x = 4$
$y = 0$	0	1	62	28	27
$y = 1$	36	44	6	55	20
$y = 2$	3	10	43	25	39
$y = 3$	41	45	15	21	8
$y = 4$	18	2	61	56	14

SHA-3 Hash Function

Keccak permutation: $\iota \circ \chi \circ \pi \circ \rho \circ \theta$

π step: permutation on lanes



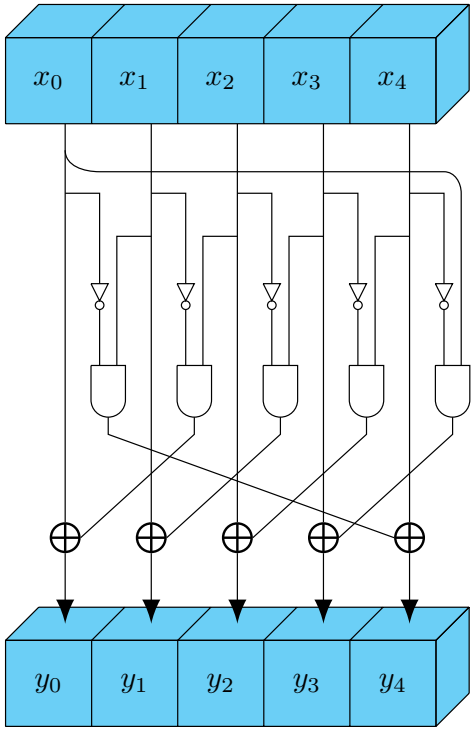
$$A[y, 2 * x + 3 * y] = A[x, y]$$

SHA-3 Hash Function

KECCAK permutation: $\iota \circ \chi \circ \pi \circ \rho \circ \theta$

χ step: 5-bit S-boxes, nonlinear operation on rows

$$
\begin{aligned}
y_0 &= x_0 \oplus (x_1 \oplus 1) \cdot x_2 \\
y_1 &= x_1 \oplus (x_2 \oplus 1) \cdot x_3 \\
y_2 &= x_2 \oplus (x_3 \oplus 1) \cdot x_4 \\
y_3 &= x_3 \oplus (x_4 \oplus 1) \cdot x_0 \\
y_4 &= x_4 \oplus (x_0 \oplus 1) \cdot x_1
\end{aligned}
$$



The algebraic degrees of χ and χ^{-1} are 2 and 3.

SHA-3 Hash Function

Keccak permutation: $\iota \circ \chi \circ \pi \circ \rho \circ \theta$

ι step: adding a round constant to the state

Adding one round-dependent constant to the first "lane", to destroy the symmetry.

Without ι

- The round function would be symmetric.
- All rounds would be the same.
- Fixed points exist.
- Vulnerable to rotational attacks, slide attacks, ...

SHA-3 Hash Function

Round function of Keccak-f

Internal state A: a 5×5 array of 64-bit lanes

θ step $C[x] = A[x,0] \oplus A[x,1] \oplus A[x,2] \oplus A[x,3] \oplus A[x,4]$
$D[x] = C[x-1] \oplus (C[x+1] \lll 1)$
$A[x,y] = A[x,y] \oplus D[x]$

ρ step $A[x,y] = A[x,y] \lll r[x,y]$
- The constants $r[x,y]$ are the rotation offsets.

π step $A[y, 2*x+3*y] = A[x,y]$

χ step $A[x,y] = A[x,y] \oplus ((\ A[x+1,y]) \& A[x+2,y])$

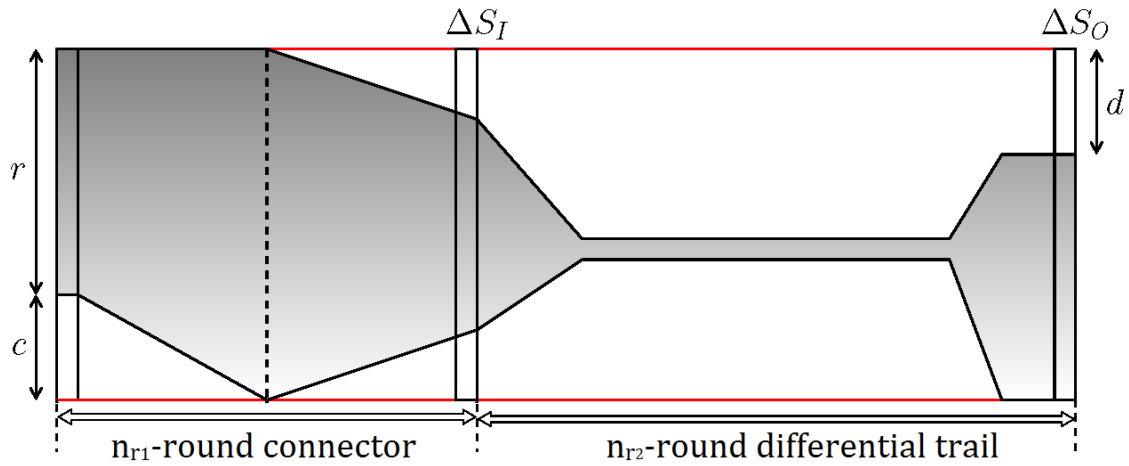
ι step $A[0,0] = A[0,0] \oplus RC$
- $RC[i]$ are the round constants.

$L \triangleq \pi \circ \rho \circ \theta$

The only non-linear operation is χ step.

Collision Attacks — the Framework

$(n_{r_1} + n_{r_2})$-round collision attacks:



- n_{r_1}-round **connector**: produces message pairs (M_1, M_2) s.t.

$$\mathrm{R}^{n_{r_1}}(\overline{M_1}||0^c) + \mathrm{R}^{n_{r_1}}(\overline{M_2}||0^c) = \Delta S_I, \quad (\mathrm{R}^{n_{r_1}} : n_{r_1} \text{ rounds})$$

$$n_{r_1} = 1 \text{ [DDS13]} \longrightarrow n_{r_1} = 2 \text{ [QSLG17]} \longrightarrow n_{r_1} = 3 \text{ [SLG17]} .$$

- n_{r_2}-round **differential**: $\Delta S_I \to \Delta S_O$,
  with first d bits of ΔS_O being 0, *i.e.*, collision.

Collisions Attack — The State of the Art

Without the help of SAT or MILP

Rounds	Target	Setting	Complexity	Reference
5	SHA3-256	Classical	Practical	[GLL+19]
5	SHA3-224	Classical	Practical	[SLG17]
5	SHAKE128	Classical	Practical	[QSLG17]
4	SHA3-384	Classical	2^{147}	[DDS13]
3	SHA3-512	Classical	Practical	[DDS13]

Collisions Attack — The State of the Art

With the help of SAT or MILP

Rounds	Target	Setting	Complexity	Reference
6	SHA3-256	Quantum	$2^{104}/\sqrt{S}$	[GLST22]
6	SHA3-224	Quantum	$2^{98}/\sqrt{S}$	[GLST22]
6	SHAKE128	Quantum	$2^{67}/\sqrt{S}$	[GLST22]
6	SHAKE128	Classical	$2^{123.5}$	[GLST22]
5	SHA3-256	Classical	Practical	[GLL+19]
5	SHA3-224	Classical	Practical	[SLG17]
5	SHAKE128	Classical	Practical	[QSLG17]
4	SHA3-384	Classical	2^{147}	[DDS13]
3	SHA3-512	Classical	Practical	[DDS13]

Outline

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

Description of AES

There are 10/12/14 rounds for AES-128/192/256, respectively.

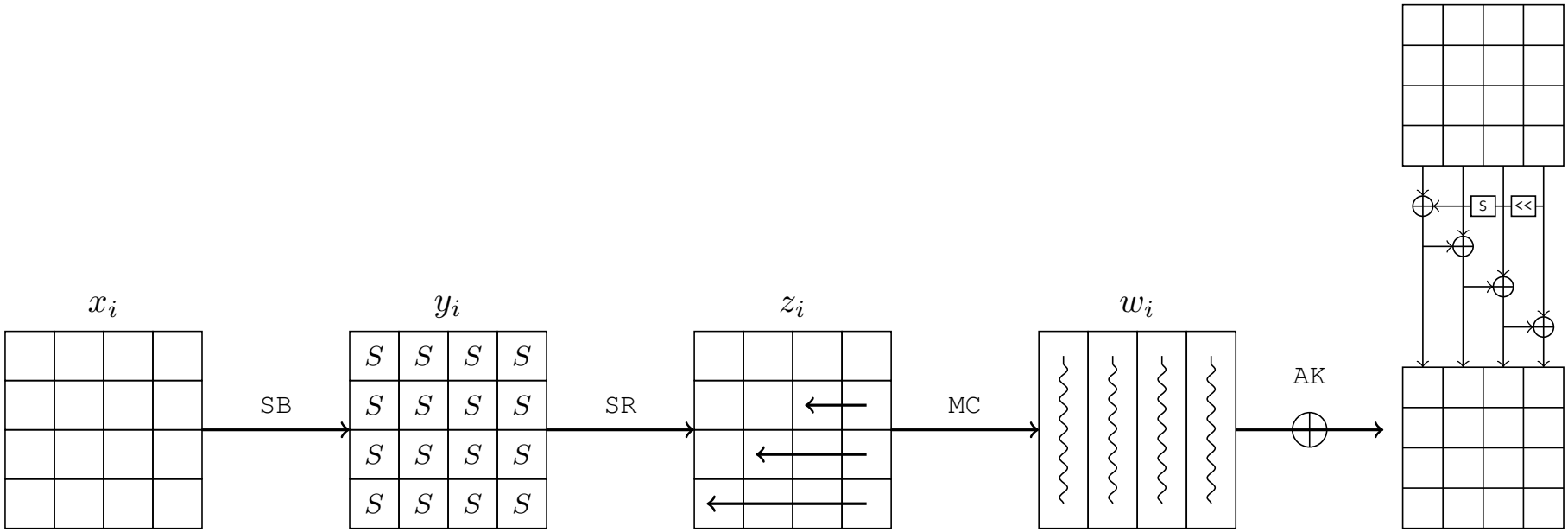


Figure: The round function of AES

(Triangulating) Rebound Attack

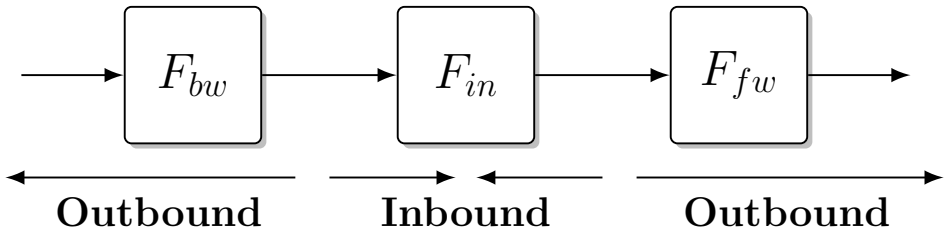


Figure: The Rebound Attack

- introduced by Mendel et al. in FSE 2009
- relies on high probability differentials in F_{bw} and F_{fw}
- constraints: the middle have to match in state and keys (if use)
- optimization goal: overall complexity

Collision Attack on AES Hashing

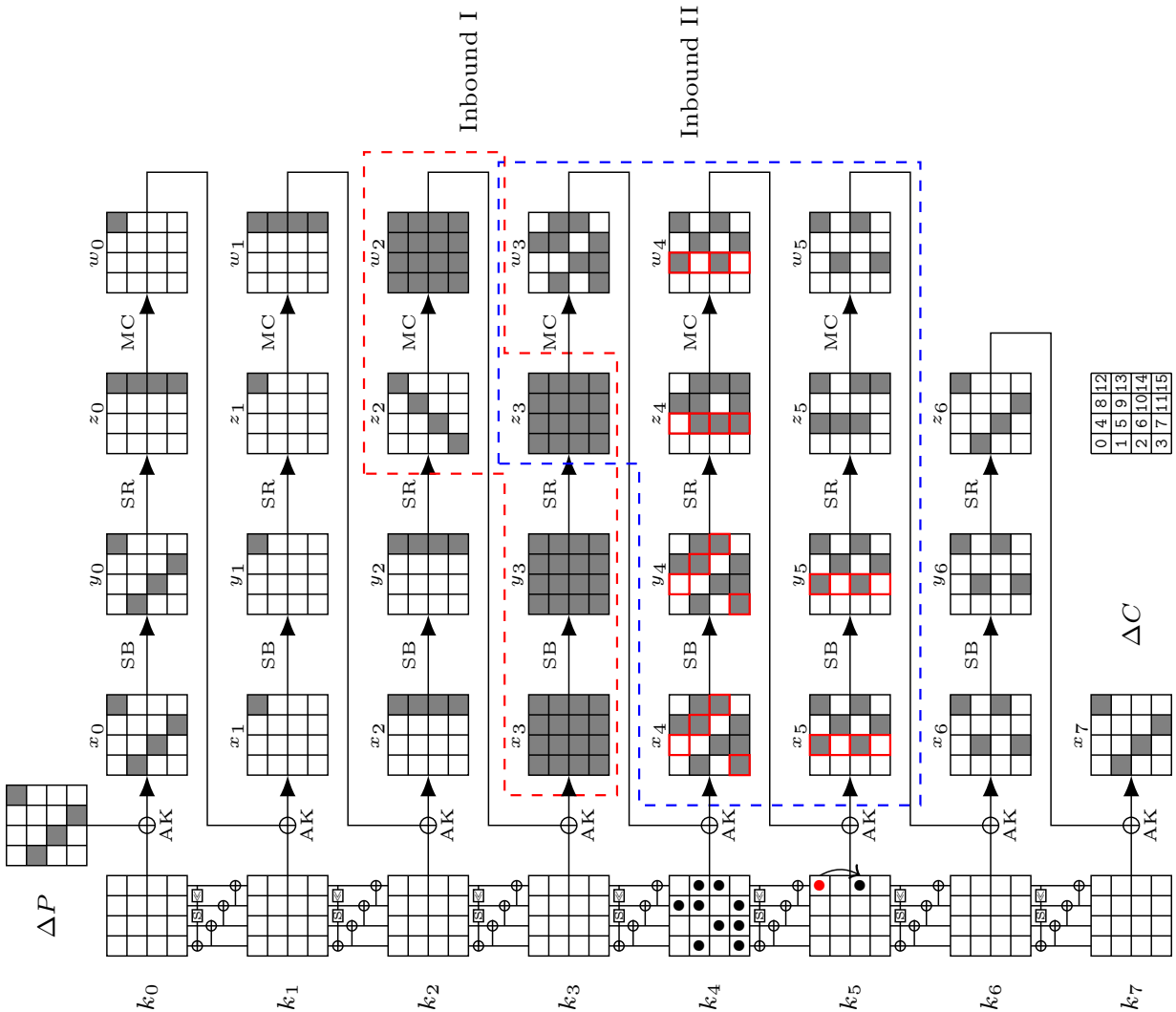


Figure: Semi-free-start collision attack on 7-round AES-128 with complexity 2^{56} in classical setting.

Collision Attack on AES Hashing

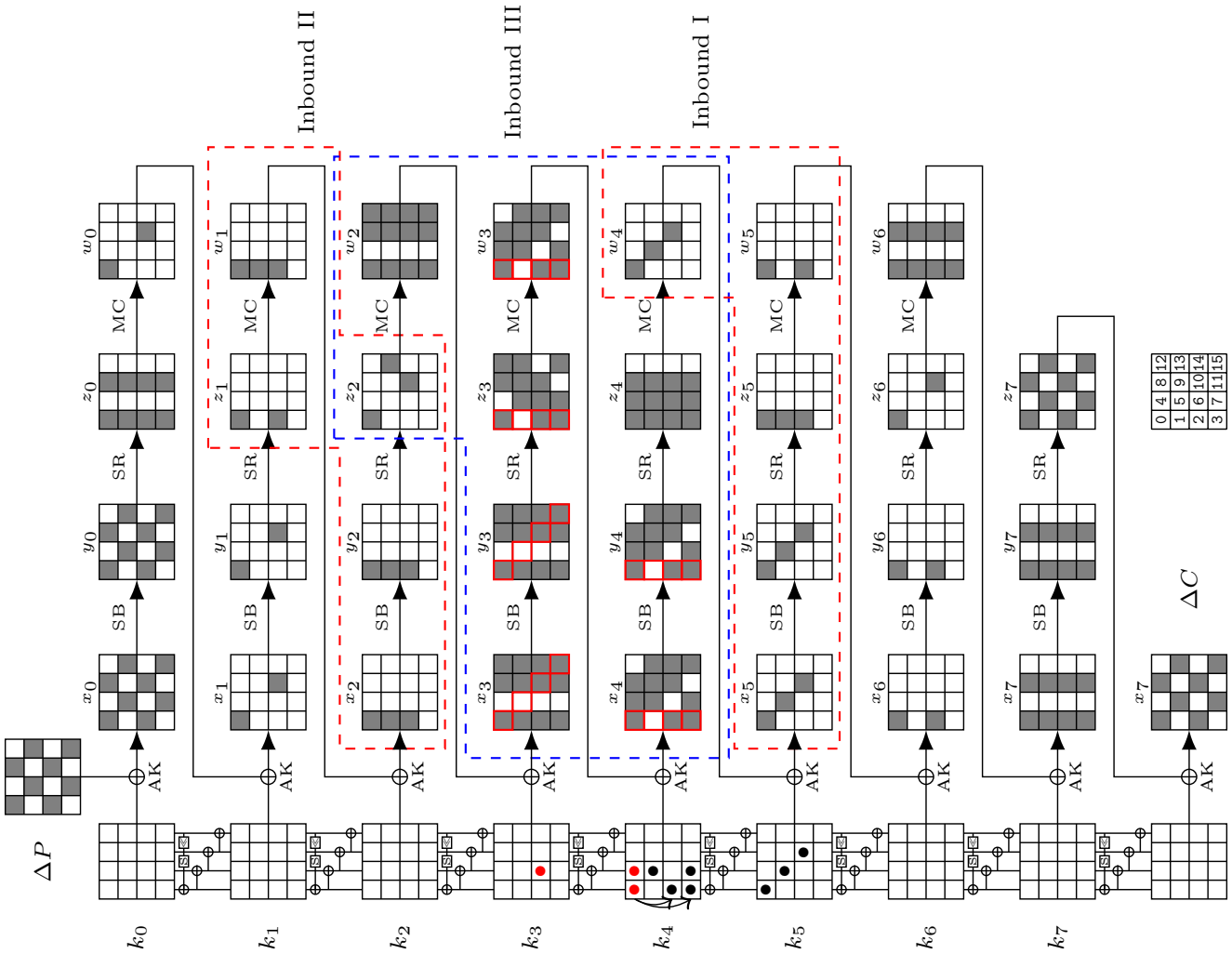


Figure: Semi-free-start collision attack on 8-round AES-128 with complexity 2^{34} in quantum setting.

Collision Attack on AES Hashing — Summary

- Classical: 7-round with complexity 2^{56}
- Quantum: 8-round with complexity 2^{34}

Outline

1. Introductions to Hash Functions and Cryptanalysis
2. Developments of Cryptanalysis
3. Security Status of SHA-3
4. Security Status of AES Hashing
5. Summary and Projections

Cryptanalysis by Machine Learning

The idea is to mimic the differential based key recovery attack: to train the machine to distinguish the reduced rounds with high-probability differential from random ones, and then launch key recovery attack utlizing such distinguishers.

Developments

- First Attempt by Gohr at CRYPTO 2019 with application to SPECK
- Second Attempt by Bao et al. ASIACRYPT 2022 to SPECK and SIMON with better complexities than men-made attacks.

Limitations

Limited by practical complexity, no good theoretical model to treat the complexities, hard to project.

Conclusions

- Collision attack of SHA-3 is up to 6 rounds, lots of security margin left, maybe KangarooTwelve is a good idea (twice as fast).

Conclusions

- Collision attack of SHA-3 is up to 6 rounds, lots of security margin left, maybe KangarooTwelve is a good idea (twice as fast).
- Collision attack of AES is up to 7/8 rounds in classical and quantum settings, still usable, but too low security strength in quantum setting.

Conclusions

- Collision attack of SHA-3 is up to 6 rounds, lots of security margin left, maybe KangarooTwelve is a good idea (twice as fast).
- Collision attack of AES is up to 7/8 rounds in classical and quantum settings, still usable, but too low security strength in quantum setting.
- What to expect from future development of cryptanalysis: deeper look at the core algorithms underlying the solvers and machine learning

Thank You !

Q&A

References I


Itai Dinur, Orr Dunkelman, and Adi Shamir.
Collision Attacks on Up to 5 Rounds of SHA-3 Using Generalized Internal Differentials.
In Shiho Moriai, editor, *FSE 2013*, volume 8424 of *LNCS*, pages 219–240. Springer, 2013.


Jian Guo, Guohong Liao, Guozhen Liu, Meicheng Liu, Kexin Qiao, and Ling Song.
Practical Collision Attacks against Round-Reduced SHA-3.
Journal of Cryptology, 2019.


Jian Guo, Guozhen Liu, Ling Song, and Yi Tu.
Exploring SAT for Cryptanalysis: (Quantum) Collision Attacks against 6-Round SHA-3.
In *ASIACRYPT 2022*, 2022.


Kexin Qiao, Ling Song, Meicheng Liu, and Jian Guo.
New Collision Attacks on Round-Reduced Keccak.
In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *EUROCRYPT 2017 (III)*, volume 10212 of *LNCS*, pages 216–243, 2017.


Ling Song, Guohong Liao, and Jian Guo.
Non-full Sbox Linearization: Applications to Collision Attacks on Round-Reduced Keccak.
In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017 (II)*, volume 10402 of *LNCS*, pages 428–451. Springer, 2017.