Practical Analysis of Reduced-Round KECCAK

María Naya-Plasencia, Andrea Röck and Willi Meier

Indocrypt 2011

Overview

- Sponge construction and KECCAK
- Previous analysis results
- Differentials in KECCAK
- Differential distinguisher on 4-round reduced hash
- Collisions/near collisions on reduced-round KECCAK
- Preimages in practical time for 2 rounds
- Conclusions

Sponges and KECCAK

KECCAK is family of sponge hash functions.

In sponge hash function message block of r bits is absorbed into its internal state, and internal permutation P is applied to the state.

This step is applied repeatedly, until all message blocks have been treated.

In sqeezing phase, a subset of r state bits is deduced before each new permutation application, until desired number ℓ of output bits are generated.

Sponges and KECCAK

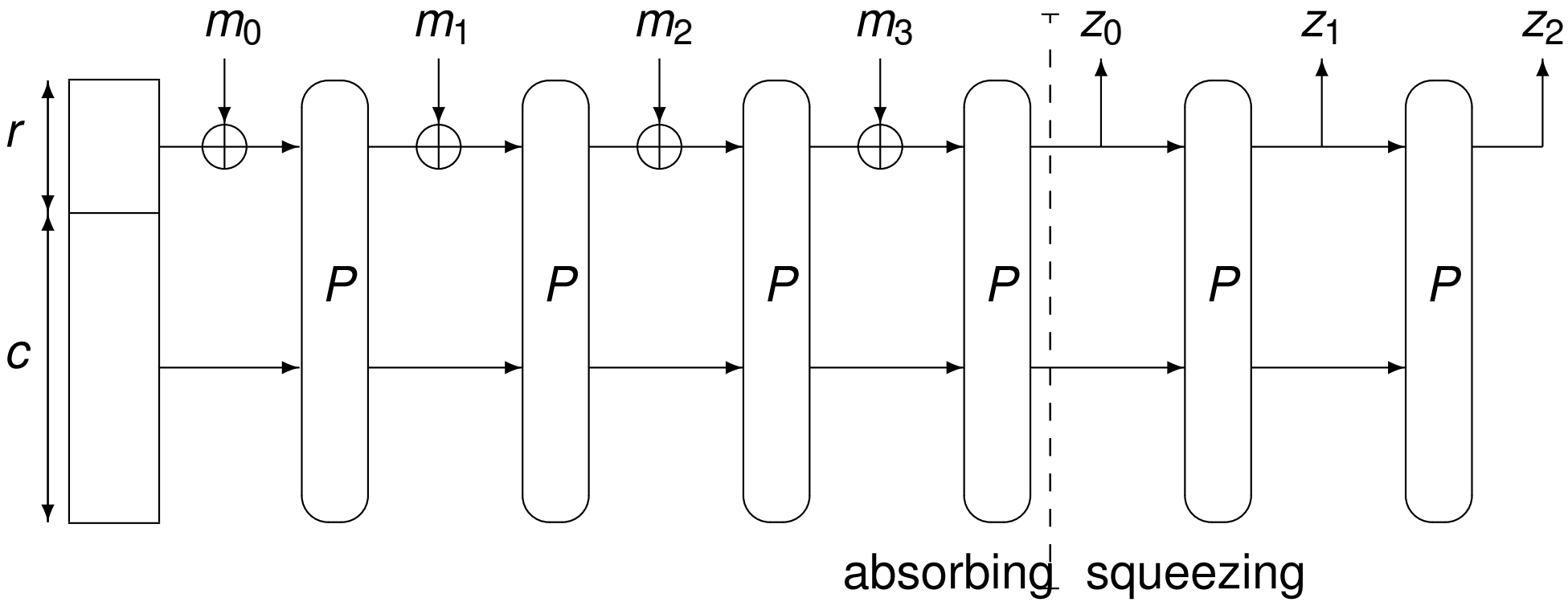
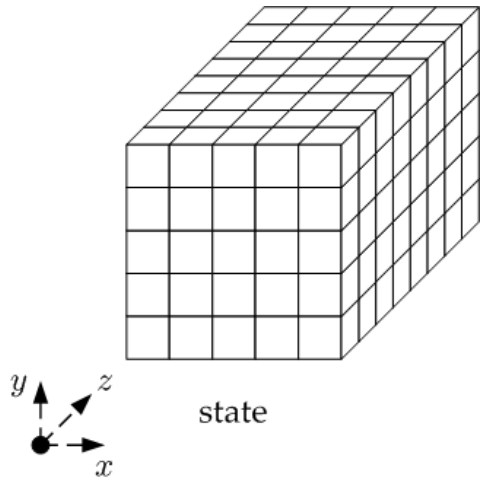


Figure: Sponge construction, for a 4-block message.

KECCAK (Bertoni-Daemen-Peeters-Van Assche 08)



KECCAK: SHA-3 finalist.

- 1600-bit state, viewed as 64 slices of 5×5 bits: 5 rows and 5 columns.
- Nonlinear layer: 320 parallel applications of a 5×5-bit S-box χ of degree 2.
- Internal permutation P, denoted KECCAK-f[1600], consists of 24 iterations of the round function.

KECCAK

Round function composed of five steps:

1. θ: XOR to each bit the XOR of two columns. First column in same slice as the updated bit, second column in slice before updated bit.
2. ρ: Translates bits in z-direction.
3. π: Permute the bits within a slice.
4. χ: Apply S-box on each row ($x = 0, \ldots, 4$, y and z fixed).
5. ι: Addition of a constant.

KECCAK

Capacity c: Difference of sizes of state and message block.

Capacity dependent on output size.
In case of output size $\ell = 256$ bits, capacity is $c = 512$ bits, and message size is $r = 1088$ bits.

Hash output: First 256 bits of the state after absorbing all messsage blocks.

Capacity $c = 2 \cdot \ell$: Security claim for resulting hash function $\mathcal{H}$ against collision and preimage finding is as required, i.e.,

$2^{\ell/2}$ for collisions and

2^{ℓ} for (second) preimages.

Previous Analysis Results

Preimages:

D. Bernstein: Preimage attacks on 6, 7 and 8 rounds, marginally better than generic attacks.

P. Morawiecki - M. Srebrny: Practical preimage attack on 3 rounds of weakened variants of KECCAK (e.g., hash size 1024 bit).

Previous analysis results

Distinguishing internal permutation P from random:

Zero-sum distinguishers (AM), reach considerable number of rounds.

Zero-sum based distinguishers of permutation P by Boura-Canteaut-De Cannière: Reach full 24-round 1600-bit permutation P. Complexity huge: 2^{1575}.

Zero-sums hard to exploit for collisions or preimages.

Rebound attack by Duc-Guo-Peyrin-Wei: Study differential paths for up to 5 rounds, to give distinguisher on permutation P for up to 8 rounds, with complexity about 2^{491}. (Simultaneous and independent from our results.)

Differentials in KECCAK

Aim: Search for low-weight differential paths.

Input difference zero outside message part of state of hash function.

State difference is column parity kernel or CP-kernel, abr. kernel, if it is invariant under function θ, e.g., if in each column difference is in even number of bits.

If in a column a difference is in odd number of bits, θ spreads this difference to 10 bits.

Strategy: Keep state differences within kernel as long as possible.

Shown by designers: No low weight differentials possible that are kernel for 3 consecutive rounds.

Differentials in KECCAK

Search for two consecutive kernels: Double kernels

Property of S-box: Every 1-bit difference within a row before application of χ stays the same after χ with probability 2^{-2}.

Path (with transformation ι ignored in difference):

$$\Delta_1 \overbrace{\xrightarrow{\theta,\rho,\pi,} \Delta_2 \xrightarrow{\chi} \Delta_2}^{\text{round}} \overbrace{\xrightarrow{\theta,\rho,\pi,} \Delta_3 \xrightarrow{\chi} \Delta_3}^{\text{round}}$$

Δ_1 and Δ_2 are kernels.

Highest differential probability $2^{-12} \cdot 2^{-12} = 2^{-24}$ achieved with a characteristic 6-6-6 of active S-boxes.

Differentials in KECCAK

For description of differentials, need to address bits in $5 \times 5 \times 64 = 1600$-bit state.

Coordinates of state bits: (x, y, z), $0 \leq x \leq 4$, $0 \leq y \leq 4$, $0 \leq z \leq 63$.

Alternatively, state bits numbered from 0 to 1599. Conversion from (x, y, z) to global bit position:

$$\text{global pos} = 64(5y + x) + z.$$

Differentials in KECCAK

Assignment of (x, y)-coordinates is as Table:

Table: Bit notation in a slice.

	$x = 3$	$x = 4$	$x = 0$	$x = 1$	$x = 2$
$y = 2$	bit 1	bit 2	bit 3	bit 4	bit 5
$y = 1$	bit 6	bit 7	bit 8	bit 9	bit 10
$y = 0$	bit 11	bit 12	bit 13	bit 14	bit 15
$y = 4$	bit 16	bit 17	bit 18	bit 19	bit 20
$y = 3$	bit 21	bit 22	bit 23	bit 24	bit 25

Differentials in KECCAK

Best path found:

Δ_1:	(x, y, z)	Δ_2:	(x, y, z)	Δ_3:	(x, y, z)
	$(0, 0, 0)$		$(0, 0, 0)$		$(0, 0, 0)$
	$(0, 1, 0)$		$(0, 2, 0)$		$(2, 1, 3)$
	$(2, 1, 30)$		$(2, 0, 9)$		$(0, 4, 7)$
	$(2, 2, 30)$		$(2, 3, 9)$		$(3, 1, 17)$
	$(1, 0, 63)$		$(1, 2, 36)$		$(3, 3, 24)$
	$(1, 2, 63)$		$(1, 3, 36)$		$(2, 3, 46)$

First difference Δ_1 fits into a 1088-bit message:

global pos largest for $(x, y, z) = (2, 2, 30)$: 798 (message is put into state from pos 0 to $msgSize - 1$).

Duc. et. al. independently found similar differentials.

Distinguishing 4 Rounds of the Hash Function

Notations:

f_R: One round of KECCAK-f[1600] function.

X_M: Internal state after absorbing a partial message M.

Offline step:

Find message $M||m$ such that $(X_M \oplus m, X_M \oplus m \oplus \Delta_1)$ satisfies differential path as before:

$$f_R^2(X_M \oplus m) \oplus f_R^2(X_M \oplus m \oplus \Delta_1) = \Delta_3.$$

m, $m \oplus \Delta_1$: last message blocks with correct padding.

Find such compatible message $M||m$ in 2^{24} trials.

Distinguishing 4 Rounds of the Hash Function

Neutral bit:

A bit that can be flipped in m so that differential path is still followed.

Check number of neutral bits and their positions within range of $r = 1088$ bits of message block: 81 neutral bits.

Consider A: vector space of all binary vectors of size r which are 0 outside neutral bit positions.

For any compatible message $M||m$ and any difference $\alpha \in A$, pair of states

$$(X_M \oplus m, X_M \oplus \alpha, X_M \oplus m \oplus \Delta_1 \oplus \alpha)$$

satisfies differential path.

Distinguishing 4 Rounds of the Hash Function

$\mathcal{H}_i$: i-th bit of hash of KECCAK-256 reduced to 4 rounds.

$S_N = (\alpha_1, \ldots, \alpha_N)$: Set of N distinct nonzero differences in A.

Bias ϵ_i of i-th bit defined as:

$$\frac{\#\{1 \leq j \leq N : \mathcal{H}_i\Big(M\|(m \oplus \alpha_j)\Big) \oplus \mathcal{H}_i\Big(M\|(m \oplus \alpha_j \oplus \Delta)\Big) = 1\}}{N} - \frac{1}{2}$$

Distinguishing 4 Rounds of the Hash Function

Distinguishing feature of 4-round KECCAK-hash:

For any compatible message M, and any set S_N of differences, there are 18 positions i in the hash, so that the absolute value of the bias is $|\epsilon_i| = 2^{-1}$:

The bits of the hash at these 18 positions always flip or always stay constant.

For a random function this would happen with probability only 2^{-18N} (where N is cardinality of set S_N).

Near-Collisions on 3 Rounds

Use previous differential path for constructing near-collisions on the 3-round reduced 256-bit hash function.

Tradeoff:

Near-collisions with difference in hash of Hamming weight 29 with complexity 2^{24}, or

weight 9 with increased complexity 2^{44}, by controlling 20 additional bit conditions.

Collisions on 2 Rounds

Find collision on 2-round reduced hash function by means of appropriate differential:

Path with nonzero difference entirely in message part, and with zero difference in the hash.

Impossible by double kernel on 3 slices only, but find such a path with double kernel on 4 slices.

Path (with transformation ι ignored in difference):

$$\overbrace{\Delta_1 \xrightarrow{\theta,\rho,\pi,} \Delta_2 \xrightarrow{\chi} \Delta_2}^{\text{round}} \overbrace{\xrightarrow{\theta,\rho,\pi,} \Delta_3 \xrightarrow{\chi} \Delta_3}^{\text{round}}$$

Collisions on 2 Rounds

Δ_1: (x, y, z)	Δ_2: (x, y, z)	Δ_3: (x, y, z)
$(1, 2, 0)$	$(2, 1, 7)$	$(2, 1, 1)$
$(1, 3, 0)$	$(2, 3, 7)$	$(4, 1, 7)$
$(0, 2, 4)$	$(2, 3, 10)$	$(1, 2, 13)$
$(0, 3, 4)$	$(2, 4, 10)$	$(3, 3, 22)$
$(4, 0, 35)$	$(3, 1, 45)$	$(3, 3, 25)$
$(4, 2, 35)$	$(3, 4, 45)$	$(1, 4, 36)$
$(1, 0, 61)$	$(0, 2, 62)$	$(4, 3, 37)$
$(1, 2, 61)$	$(0, 3, 62)$	$(3, 4, 39)$

Differences Δ_2, Δ_3 have each 8 rows with a 1-bit difference in input and output of χ.

Total probability: $2^{-16} \cdot 2^{-16} = 2^{-32}$ of following characteristic.

Using conditions and free (neutral) bits, can find practical collisions in 2^{13} steps.

Preimages on 2 Rounds

Construct preimages for 2 rounds of KECCAK, with

time complexity 2^{33}, and

2^{29} memory.

Algorithm works for different parameters, but we give description for hash size $\ell = 256$.

Preimages on 2 Rounds

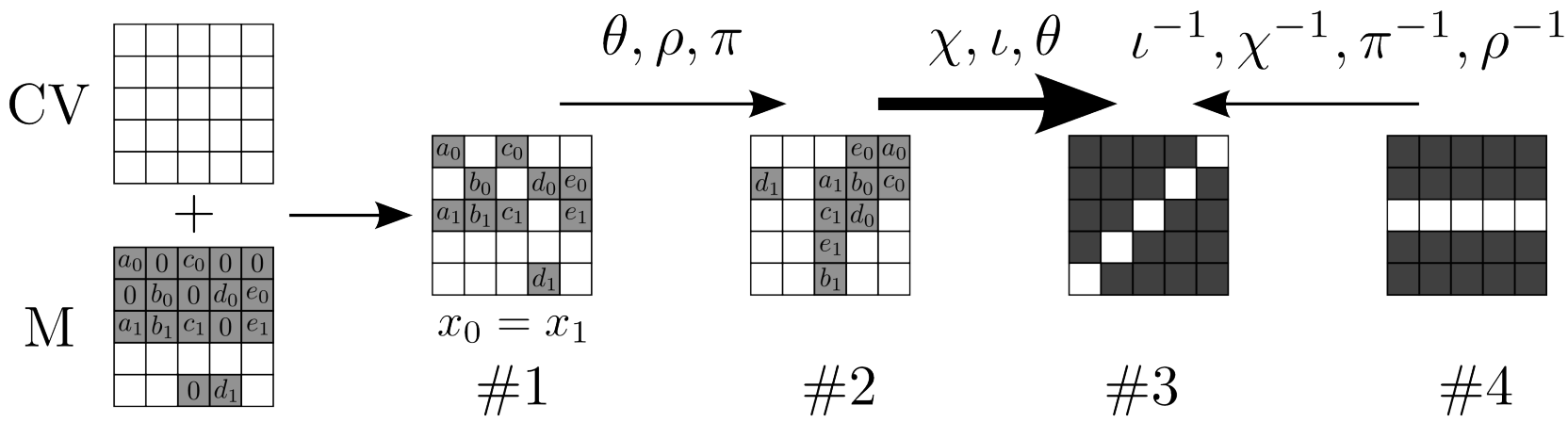


Figure: Diagram of the 2-round preimage attack. Each square represents a 64 bit lane. Each white lane is a lane known and fixed, each colored one, a not-yet-fixed lane.

For simplicity: ι transformation omitted in description of attack (but taken into account for implementation).

Preimages on 2 Rounds

Given:

- A hash value by 4 out of 5 white lanes in rightmost slice #4.
- A chaining value, e.g. the initial all zero one.

Fifth lane unknown. Fix it to a random value.

Problem: Find a message block that produces these 5 lanes, and so fits the given hash value.

Gray lanes show into which lanes of chaining value the message is XORed.

Lanes marked with 0 are fixed to 0.

Lanes marked with $(a_0, a_1, b_0, b_1, \ldots, e_0, e_1)$ are variable and suitably adapted during search, apart from conditions $a_0 = a_1; \ldots; e_0 = e_1$ ($x_0 = x_1$ condition).

Preimages on 2 Rounds

Conditions effect that operation θ will not change the unknown lanes.

Out of initial state #1 compute known lanes in #2 after θ, ρ and π together with their positions.

Imposing previous conditions, still $5 \cdot 64$ degrees of freedom for message remain, to finally agree with the given output.

In backward direction, invert χ from white row of final state #4.

Apply inverse of π and ρ to obtain values and positions of 5 known lanes in #3.

Preimages on 2 Rounds

Problem:

Find values of the 10 64-bit words $(a_0, a_1, b_0, b_1, \ldots, e_0, e_1)$ in #2 so that the two actions/conditions fit:

- transition by χ, θ from #2 forwards
- bits fixed in #3 from the backwards computation

Strategy:

Find partial solutions on suitable subsets of slices

Preimages on 2 Rounds

Start by finding subsets of bits that verify relations for 3 slices.

Step by step increase to partial solutions for 12, 24, 48 slices.

Find partial solutions for remaining 16 slices.

Solutions for 48 slices and solutions for 16 slices have to be matched.

Delicate part: In each step check compatibility regarding conditions $x_0 = x_1$, and check number of available solutions.

Actual preimages on 2 rounds found in 2^{33} time and 2^{29} memory.

Conclusions

- Cryptanalysis on a few rounds of KECCAK hash function, rather than on building blocks only.
- Parameters same as in SHA-3 submission, except number of rounds.
- Methods apply to 256-bit and 224-bit versions.
- Very recent results: Collisions for KECCAK reduced to 4 rounds, and near-collisions for 5 rounds by Dinur-Dunkelman-Shamir.
- Results practical and experimentally verified.
- Number of rounds reached far from total: Results no threat.
- Problem: How to find useful differentials for more than 5 rounds?