# Security Analysis of SIMD

Charles Bouillaguet, Pierre-Alain Fouque, Gaëtan Leurent

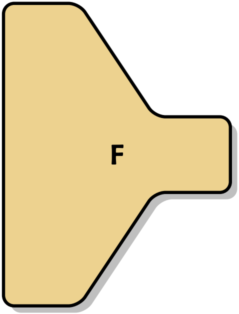
École Normale Supérieure
Paris, France

SAC 2010 – University of Waterloo

# Hash functions

- A public function with no structural properties.
  - Cryptographic strength without keys!

- $F : \{0,1\}^* \to \{0,1\}^n$

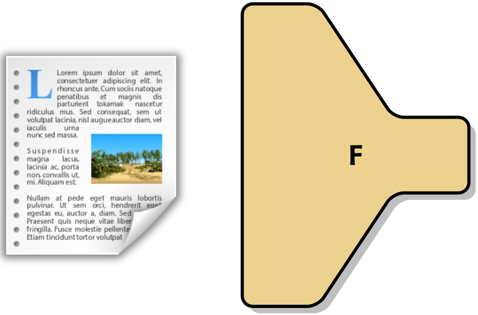




0x1d66ca77ab361c6f

# *Hash functions*

- A public function with no structural properties.
  - Cryptographic strength without keys!

- $F : \{0,1\}^* \to \{0,1\}^n$




0x1d66ca77ab361c6f

# The SHA-3 competition

- Similar to the AES competition
- Organized by NIST

- Submission dead-line was October 2008: 64 candidates
- 51 valid submissions

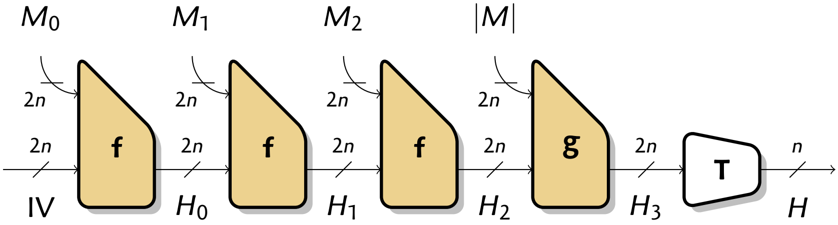
- 14 in the second round (July 2009)
- 5 finalists in September 2010?
- Winner in 2012?

# SIMD

- Merkle-Damgård with a Davies-Meyer compression function
- Strong message expansion
- Several Parallel MD-like Feistel

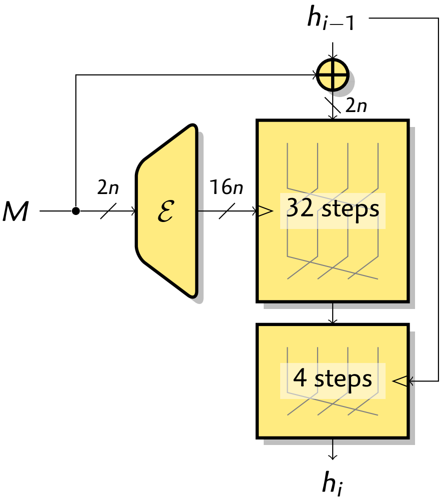
Gaëtan Leurent, Charles Bouillaguet, Pierre-Alain Fouque
SIMD Is a Message Digest
Submission to the NIST SHA-3 competition

# SIMD Iteration Mode



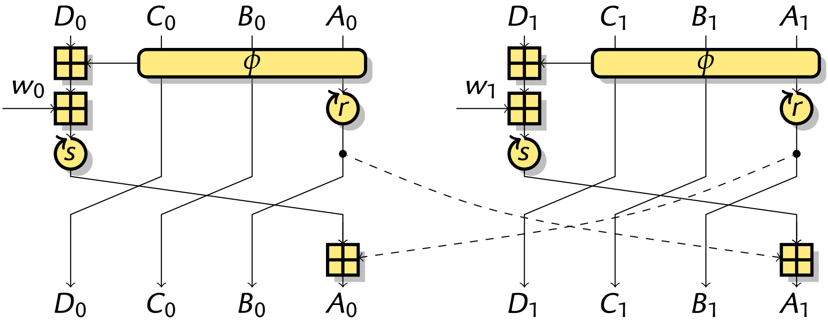

- Wide-pipe
- Finalisation function
- Use only the message length as input in the last block
  - Acts as a kind of blank round
  - Can break unexpected properties

# SIMD Compression Function




- Block cipher based
  - Well understood
- Davies-Meyer
  - Allows a strong message expansion
- Add the message at the start
  - Prevents some message modifications
- Modified feed-forward: Feistel rounds instead of XOR
  - Avoids some fixed point and multi-block attacks

# SIMD Feistel Rounds




- Follows the SHA/MD legacy
  - Additions, rotations, boolean functions
- 4 Parallel lanes for SIMD-256, 8 for SIMD-512
- Parallel Feistel rounds allow vectorized implementation

# Outline

New distinguisher for SIMD

Security proof with distinguishers

Analysis of differential paths

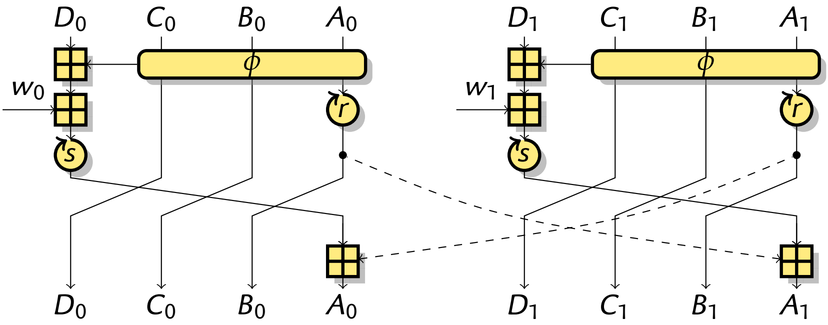
# Outline

New distinguisher for SIMD

Security proof with distinguishers

Analysis of differential paths

# Symmetry based distinguisher




- Put the same values in two lanes
- Put the same message
- Need a special message...

## Message pairs

- Let $\overleftrightarrow{\bullet}$ be a symmetry relation swapping pairs of lanes
- Let $M$, $M'$ be such that $\mathcal{E}(M') = \overleftrightarrow{\mathcal{E}(M)}$
- Let $\mathcal{S}^{(0)}$, $\mathcal{S}'^{(0)}$ be such that $\mathcal{S}'^{(0)} = \overleftrightarrow{\mathcal{S}^{(0)}}$
- Then $\mathcal{S}'^{(31)} = \overleftrightarrow{\mathcal{S}^{(31)}}$
- We can use a single message
- We can use a single state

# Message pairs

- Let $\overleftrightarrow{\bullet}$ be a symmetry relation swapping pairs of lanes
- Let $M$, $M'$ be such that $\mathcal{E}(M') = \overleftrightarrow{\mathcal{E}(M)}$
- Let $\mathcal{S}^{(0)}$, $\mathcal{S}'^{(0)}$ be such that $\mathcal{S}'^{(0)} = \overleftrightarrow{\mathcal{S}^{(0)}}$
- Then $\mathcal{S}'^{(31)} = \overleftrightarrow{\mathcal{S}^{(31)}}$
- We can use a single message
- We can use a single state

# Message pairs

- Let $\overleftrightarrow{\bullet}$ be a symmetry relation swapping pairs of lanes
- Let $M$, $M'$ be such that $\mathcal{E}(M') = \overleftrightarrow{\mathcal{E}(M)}$
- Let $\mathcal{S}^{(0)}$, $\mathcal{S}'^{(0)}$ be such that $\mathcal{S}'^{(0)} = \overleftrightarrow{\mathcal{S}^{(0)}}$
- Then $\mathcal{S}'^{(31)} = \overleftrightarrow{\mathcal{S}^{(31)}}$

- We can use a single message
- We can use a single state

# Message expansion

1. FFT transform over $\mathbb{F}_{257}$ doubles the size of the message.
2. Make two copies of the FFT output.
3. Multiply by 185/233 (from $\mathbb{F}_{257}$ to 16-bit words).
4. Permute and pack into 32-bit words.

- Constant are only in the first layer.

- FFT is linear: easy to enforce linear conditions.
- Enough degrees of freedom for equality constraints.
- Equality is preserved by the remaining steps.
- Permutations are nice wrt. to this.
- We can easily generate those messages.

- Obvious fix: add constants at the end of the expansion.

# Message expansion

1. FFT transform over $\mathbb{F}_{257}$ doubles the size of the message.
2. Make two copies of the FFT output.
3. Multiply by 185/233 (from $\mathbb{F}_{257}$ to 16-bit words).
4. Permute and pack into 32-bit words.

- Constant are only in the first layer.

- FFT is linear: easy to enforce linear conditions.
- Enough degrees of freedom for equality constraints.
- Equality is preserved by the remaining steps.
- Permutations are nice wrt. to this.
- We can easily generate those messages.

- Obvious fix: add constants at the end of the expansion.

# Message expansion

1. FFT transform over $\mathbb{F}_{257}$ doubles the size of the message.
2. Make two copies of the FFT output.
3. Multiply by 185/233 (from $\mathbb{F}_{257}$ to 16-bit words).
4. Permute and pack into 32-bit words.

- Constant are only in the first layer.

- FFT is linear: easy to enforce linear conditions.
- Enough degrees of freedom for equality constraints.
- Equality is preserved by the remaining steps.
- Permutations are nice wrt. to this.
- We can easily generate those messages.

- Obvious fix: add constants at the end of the expansion.

# Message expansion

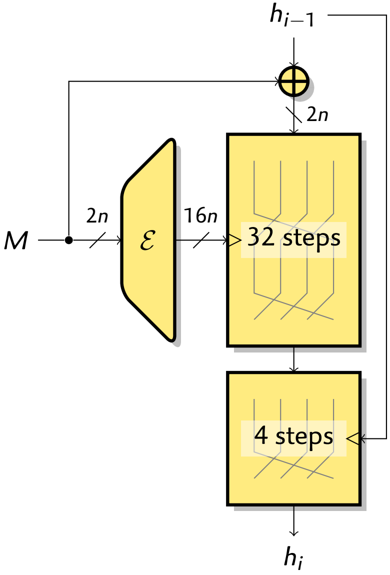
1. FFT transform over $\mathbb{F}_{257}$ doubles the size of the message.
2. Make two copies of the FFT output.
3. Multiply by 185/233 (from $\mathbb{F}_{257}$ to 16-bit words).
4. Permute and pack into 32-bit words.

- Constant are only in the first layer.

- FFT is linear: easy to enforce linear conditions.
- Enough degrees of freedom for equality constraints.
- Equality is preserved by the remaining steps.
- Permutations are nice wrt. to this.
- We can easily generate those messages.

- Obvious fix: add constants at the end of the expansion.

# Application to the Compression Function




- There are a few messages giving a symmetric expanded message

- Symmetric expanded message
- Symmetric state in the Feistel

- Message not symmetric
- Almost symmetric input
- Somewhat symmetric output

## Important properties

- $2^{16}$ weak messages ($2^{32}$ for SIMD-512)
  - $2^{256+16}$ weak chaining values ($2^{512+32}$ for SIMD-512)
- $2^{32}$ weak pairs of messages ($2^{64}$ for SIMD-512)
  - $2^{512+32}$ pairs of weak chaining values ($2^{1024+64}$ for SIMD-512)

- Wide-pipe: It is hard to get into a symmetric state / pair of states
  - Takes time $2^{256-16}$ ($2^{512-32}$ for SIMD-512)

- There is no intersection between the symmetry classes
- Each pair only works with a single message pair
- An output pair can not be used as input pair
- It cannot be used in the final transform
- Getting into a symmetric state is not really useful...

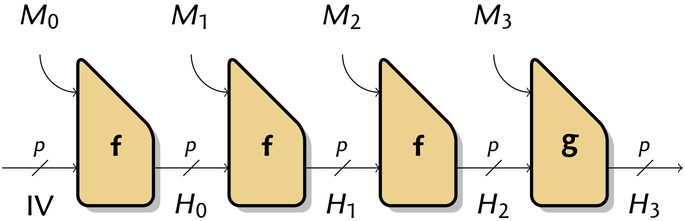
# Outline

New distinguisher for SIMD

Security proof with distinguishers

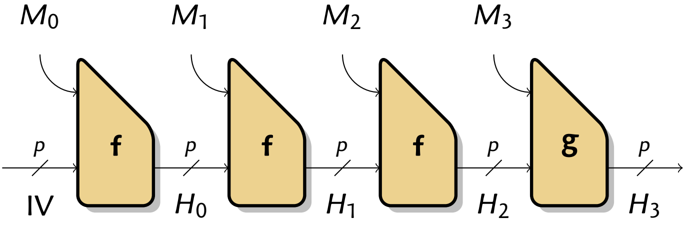
Analysis of differential paths

# Prefix-free Merkle-Damgård




- Used by several SHA-3 candidates
- Indistinguishable up to $2^{p/2}$ queries [Coron, Dodis, Malinaud, and Puniya]
  - Assuming that the compression function is perfect

# Prefix-free Merkle-Damgård




- Used by several SHA-3 candidates
- Indistinguishable up to $2^{p/2}$ queries [Coron, Dodis, Malinaud, and Puniya]
  - Assuming that the compression function is perfect

# Weak random oracle

- Random oracle with some efficient distinguishers
- We model the compression function as a random oracle, constrained to satisfy some relations:

$$\begin{aligned} \forall(h, m): &\quad \mathcal{R}_1(h, m, F(h, m)) = 1 \\ \forall(h_1, h_2, m_1, m_2): &\quad \mathcal{R}_2(h_1, m_1, h_2, m_2, F(h_1, m_1), F(h_2, m_2)) = 1 \end{aligned}$$

- Examples:
  - Symmetric states:
    $\mathcal{R}_1 := \left(m = \overleftrightarrow{m} \wedge h = \overleftrightarrow{h}\right) \Rightarrow h' = \overleftrightarrow{h'}$
  - Deterministic differential path
    $\mathcal{R}_2 := (m_1 \oplus m_2 = \Delta_m \wedge h_1 \oplus h_2 = \Delta_{\text{in}}) \Rightarrow h'_1 \oplus h'_2 = \Delta_{\text{out}}$

# Proof of Security

**Definition (Weak states)**

$$\mathcal{W} = \{h \mid \exists m, h' \text{ s.t. } \mathcal{R}_1(h, m, h') = 0\}$$

**Definition (Weak pairs of states)**

$$\mathcal{WP} = \{h_1 \leftrightarrow h_2 \mid \exists m_1, m_2, h'_1, h'_2 \text{ s.t. } \mathcal{R}_2(h_1, m_1, h_2, m_2, h'_1, h'_2) = 0\}$$

- In order to distinguish the weak RO from a real RO, the adversary needs to reach $\mathcal{W}$ or $\mathcal{WP}$.
- If they are small enough, we can simulate the weakness.

# Proof of Security

*Definition (Weak states)*

$$\mathcal{W} = \{h \mid \exists m, h' \text{ s.t. } \mathcal{R}_1(h, m, h') = 0\}$$

*Definition (Weak pairs of states)*

$$\mathcal{WP} = \{h_1 \leftrightarrow h_2 \mid \exists m_1, m_2, h_1', h_2' \text{ s.t. } \mathcal{R}_2(h_1, m_1, h_2, m_2, h_1', h_2') = 0\}$$

*Definition (Weak pairs of state+message)*

$$\mathcal{WP}' = \{(h_1, m_1) \leftrightarrow (h_2, m_2) \mid \exists m_1, m_2, h_1', h_2' \text{ s.t. } \mathcal{R}_2(\ldots) = 0\}$$

- Connected components in $\mathcal{WP}'$ must be of size 2 at most
  - Evaluation on one input gives information about a single extra input

# Proof of Security

**Definition (Weak states)**

$$\mathcal{W} = \{h \mid \exists m, h' \text{ s.t. } \mathcal{R}_1(h, m, h') = 0\}$$

**Definition (Weak pairs of states)**

$$\mathcal{WP} = \{h_1 \leftrightarrow h_2 \mid \exists m_1, m_2, h'_1, h'_2 \text{ s.t. } \mathcal{R}_2(h_1, m_1, h_2, m_2, h'_1, h'_2) = 0\}$$

**Definition (Weak pairs of state+message)**

$$\mathcal{WP}' = \{(h_1, m_1) \leftrightarrow (h_2, m_2) \mid \exists m_1, m_2, h'_1, h'_2 \text{ s.t. } \mathcal{R}_2(\dots) = 0\}$$

- Connected components in $\mathcal{WP}'$ must be of size 2 at most
  - Evaluation on one input gives information about a single extra input
- $\mathrm{Adv} \leq 16 \cdot \frac{q^2}{2^p} + 4 \cdot |\mathcal{W}| \cdot \frac{q}{2^p} + 4 \cdot |\mathcal{WP}| \cdot \frac{q^2}{(2^p - q)^2}$

# Results

## Iterating a random oracle [Coron, Dodis, Malinaud, and Puniya]

$$\mathrm{Adv} = \mathcal{O}\left(\frac{q^2}{2^p}\right)$$

- Secure up to $q = \mathcal{O}(2^{p/2})$

## Iterating a weak random oracle

$$\mathrm{Adv} = \mathcal{O}\left(\frac{q^2}{2^p} + |\mathcal{W}| \cdot \frac{q}{2^p} + |\mathcal{WP}| \cdot \frac{q^2}{(2^p - q)^2}\right)$$

- Secure up to $q = \mathcal{O}(2^{p/2})$
  if $|\mathcal{W}| = \mathcal{O}(2^{p/2})$ and $|\mathcal{WP}| = \mathcal{O}(2^p)$

- Indifferentiability proofs are quite resilient:
  many defects in the compression function have a small impact
- Can we extent this result by allowing other kinds of weaknesses?

# Application

- Symmetry based distinguishers
  - *Lesamnta*-256 is secure up to $2^{127}$ queries
  - *Lesamnta*-512 is secure up to $2^{255}$ queries
  - SIMD-256 is secure up to $2^{256-16}$ queries
  - SIMD-512 is secure up to $2^{512-32}$ queries
- Free-start differential paths
  - A differential path with a non-zero difference in $h$ costs one bit of security
- Rotational distinguisher, …

# Wide-pipe vs Narrow-pipe

- In a wide-pipe design, the indifferentiability proof implies:
  - Collision resistance
  - Preimage resistance (up to a small loss)
  - No other attack (up to a small loss)
- In a narrow-pipe design, the indifferentiability proof implies:
  - Collision resistance (up to a small loss)
  - Some distinguishers can be used for non-standard attack:
    - Herding attack on *Lesamnta* with a symmetry based distinguisher
    - Distinguishing-H attack on HMAC-MD5 with a free-start differential path

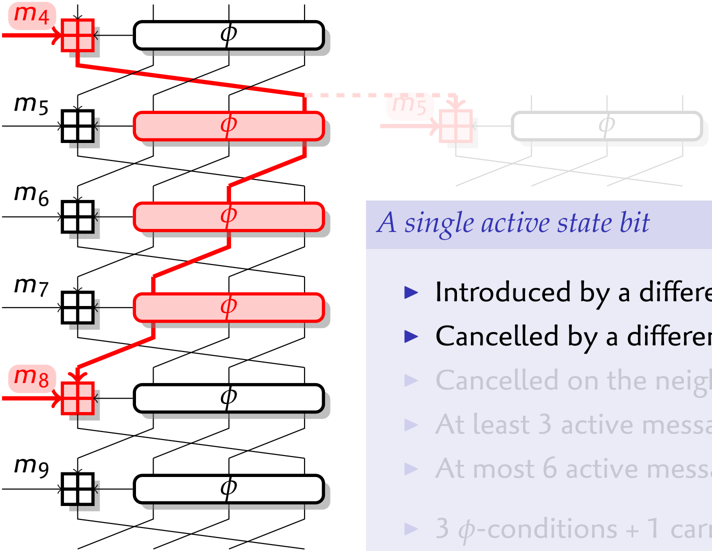
# Outline

New distinguisher for SIMD

Security proof with distinguishers

Analysis of differential paths

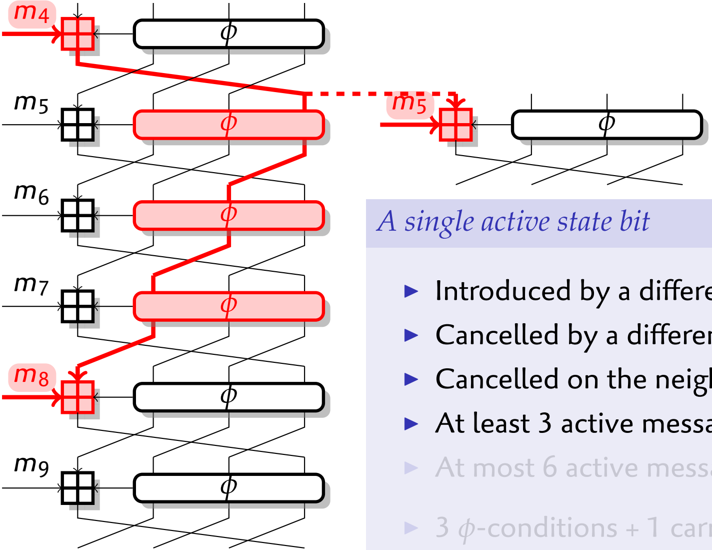
# Local Collisions




## A single active state bit

- Introduced by a difference in $m_4$
- Cancelled by a difference in $m_8$
- Cancelled on the neighbour lane
- At least 3 active messages
- At most 6 active messages
- 3 $\phi$-conditions + 1 carry condition

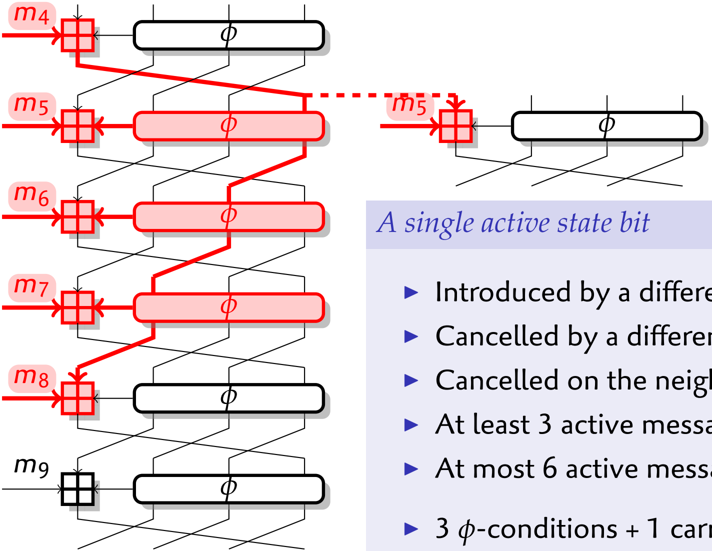
# Local Collisions




## A single active state bit

- Introduced by a difference in $m_4$
- Cancelled by a difference in $m_8$
- Cancelled on the neighbour lane
- At least 3 active messages
- At most 6 active messages
- 3 $\phi$-conditions + 1 carry condition

# Local Collisions




## A single active state bit

- Introduced by a difference in $m_4$
- Cancelled by a difference in $m_8$
- Cancelled on the neighbour lane
- At least 3 active messages
- At most 6 active messages
- 3 $\phi$-conditions + 1 carry condition

# Differential Attacks

- We assume that the adversary builds a differential path with a signed difference.
- We consider paths with a non-zero message difference
  - paths with no message difference only give free-start attacks
- Each active state bit lowers the probability
  - Minimize active state bits
- The message expansion gives many message differences
  - 520 for SIMD-256
  - 1032 for SIMD-512

# Heuristic

## Heuristic

The adversary can build an expanded message of minimal weight

- such that the differences create local collisions
- but without extra properties

- Optimal path: all Boolean function transmit differences
  - Minimizes the number of active state bits
- 6 active message bits per active state bit
  - 87 active state bits for SIMD-256 / 172 for SIMD-512
- 4 conditions per active state bit
  - 348 conditions for SIMD-256 / 688 for SIMD-512

# Comparison with SHA-1

- Differential attacks on SHA-1 use local collisions.
- Use the fact that the code is linear and circulant
  - Start with an expanded message of minimal weight
  - Make 6 shifted copy to create local collisions
  - The final expanded message has weight 6 times the minimal distance
- Our heuristic is quite weak.
- The message expansion of SIMD is neither circulant nor linear

# Comparison with SHA-1

- Differential attacks on SHA-1 use local collisions.
- Use the fact that the code is linear and circulant
  - Start with an expanded message of minimal weight
  - Make 6 shifted copy to create local collisions
  - The final expanded message has weight 6 times the minimal distance
- Our heuristic is quite weak.
- The message expansion of SIMD is neither circulant nor linear

# Weaker assumptions

## Strong adversary

The adversary can build an expanded message with any difference pattern

- If active state words are adjacent, some $\phi$ conditions disappear
  - If two inputs of the MAJ function are active we know the output

- 1 active state bit gives
  - 4.5 active message bits
  - 1 conditions

- SIMD-256: 116 conditions
- SIMD-512: 230 conditions

# Modeling Differential Paths

- Impossible to have two active inputs for all active function
- Hard to proof any usefull bound...

- We model the this problem as an Integer Linear Program
  - about 30,000 variables, 80,000 equations

- Solver computes a lower bound,
  and tries to improve the lower bound

  *SIMD*-256 $p \leq 2^{-132}$
  *SIMD*-512 $p \leq 2^{-253}$ (several weeks of computation)

# Conclusion

- SIMD security
  - Differential paths with a difference in the message are unlikely
  - Differential paths with a difference in the chaining value do not affect the iterated hash function.
- Security with distinguishers
  - Not specific to SIMD
  - A class of distinguishers does not affect the indifferentiability proof
  - Interesting for wide-pipe design
- *Full version*: ePrint report 2010/323.