Dedicated Attacks on Popular Hash Functions

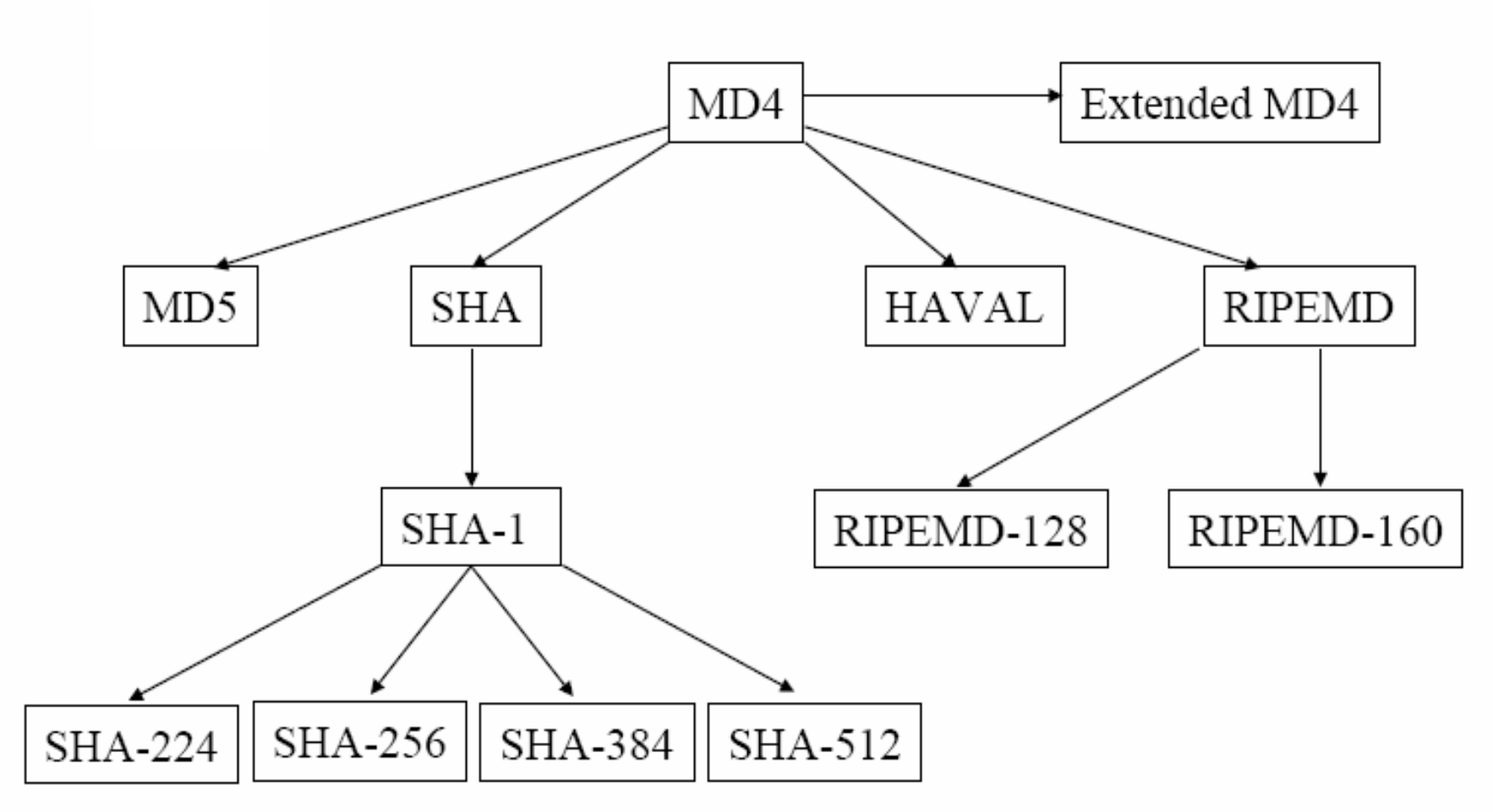
Christian Rechberger

ECRYPT Summer School, May 02, 2007

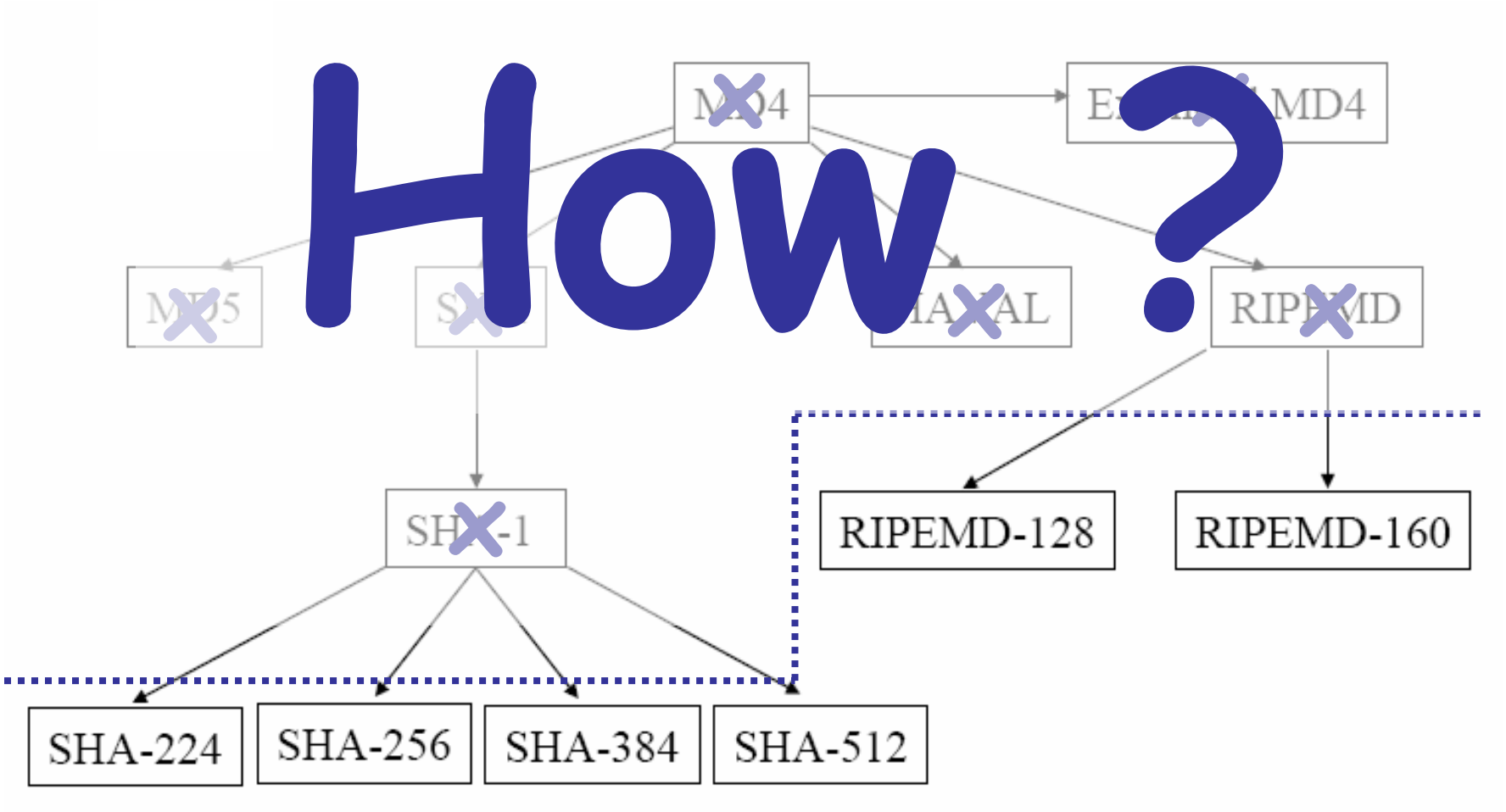
Institute for Applied Information Processing and Communications (IAIK) - Krypto Group

Faculty of Computer Science
Graz University of Technology

Motivation


MD4
Extended MD4
MD5
SHA
HAVAL
RIPEMD
SHA-1
RIPEMD-128
RIPEMD-160
SHA-224
SHA-256
SHA-384
SHA-512

Motivation


How?
MD4
MD4
MD5
SHA-1
HAVAL
RIPEMD
RIPEMD-128
RIPEMD-160
SHA-224
SHA-256
SHA-384
SHA-512

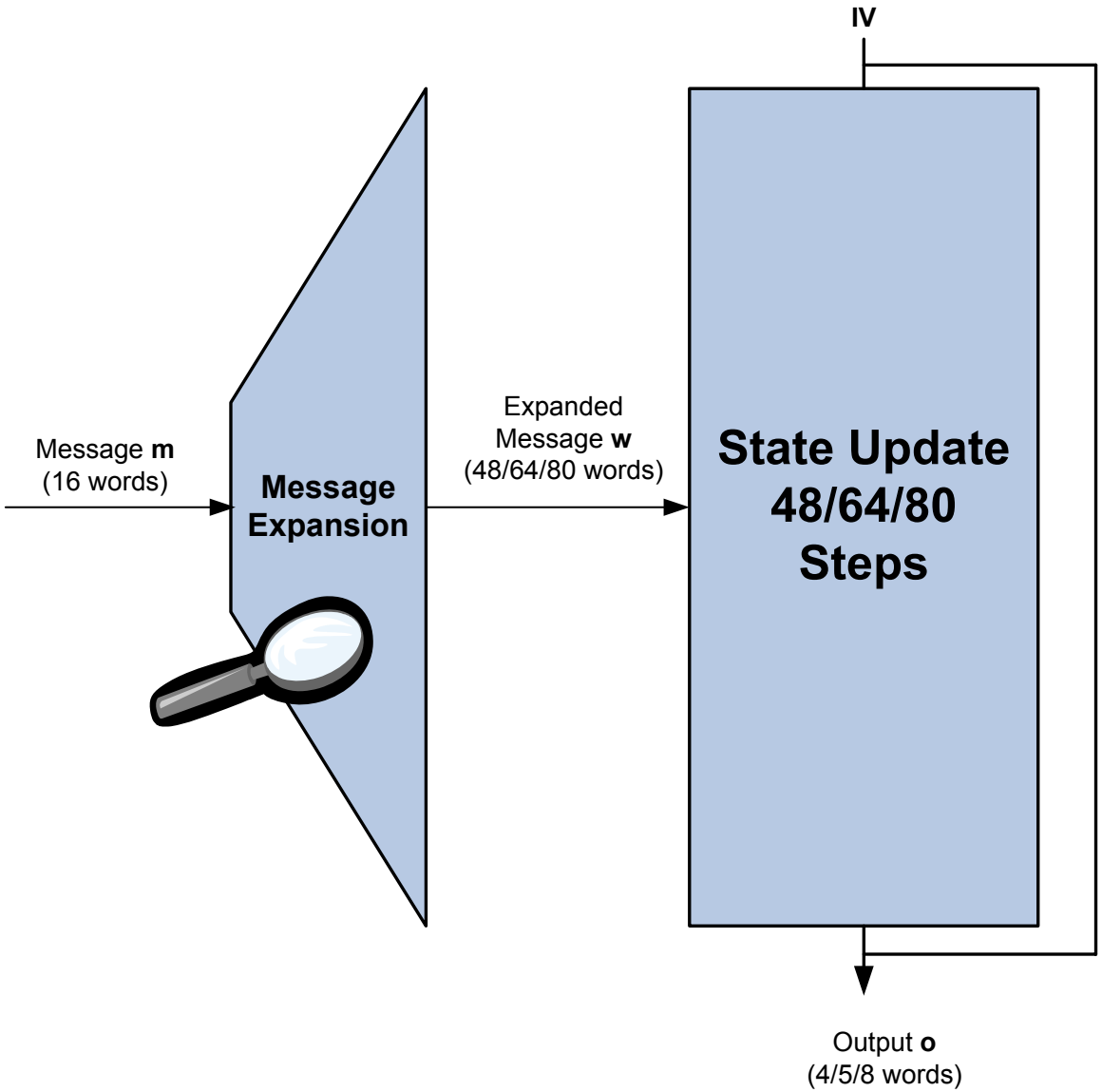
Agenda

- The MD4 “family”

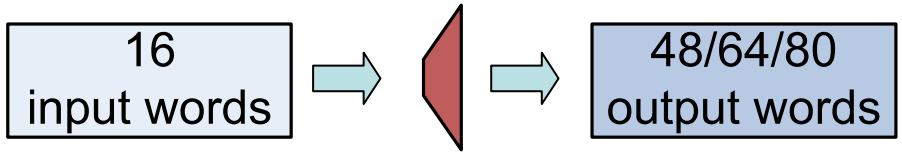
- Basic attack on SHA
- Advanced methods for fast collision search (example SHA-1) and applications
- SHA-2?

- Conclusions

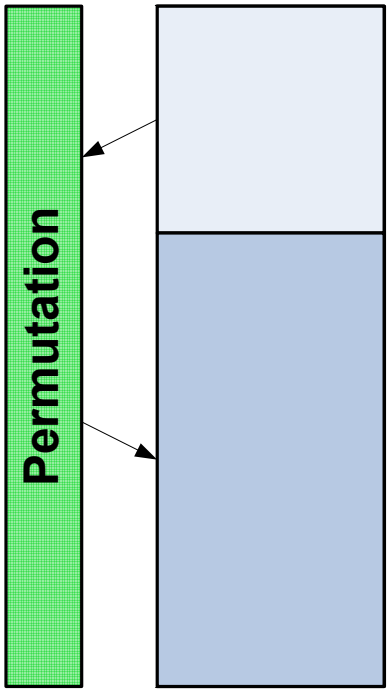
Outline of MD4-style Hash Functions



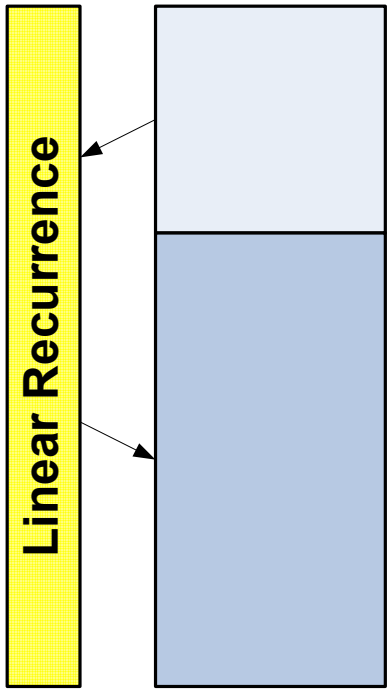
Message Expansions in the MD4 family



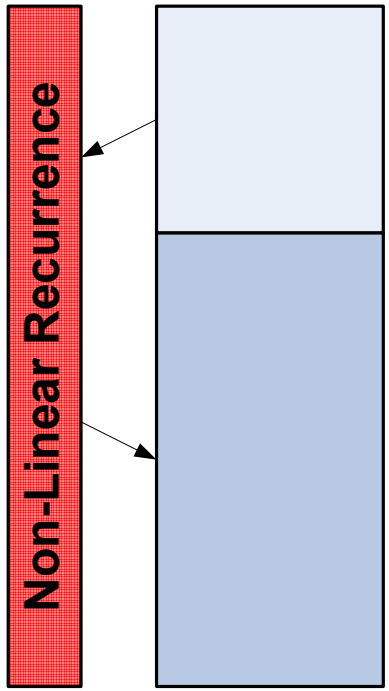
MD4/5, RIPEMD



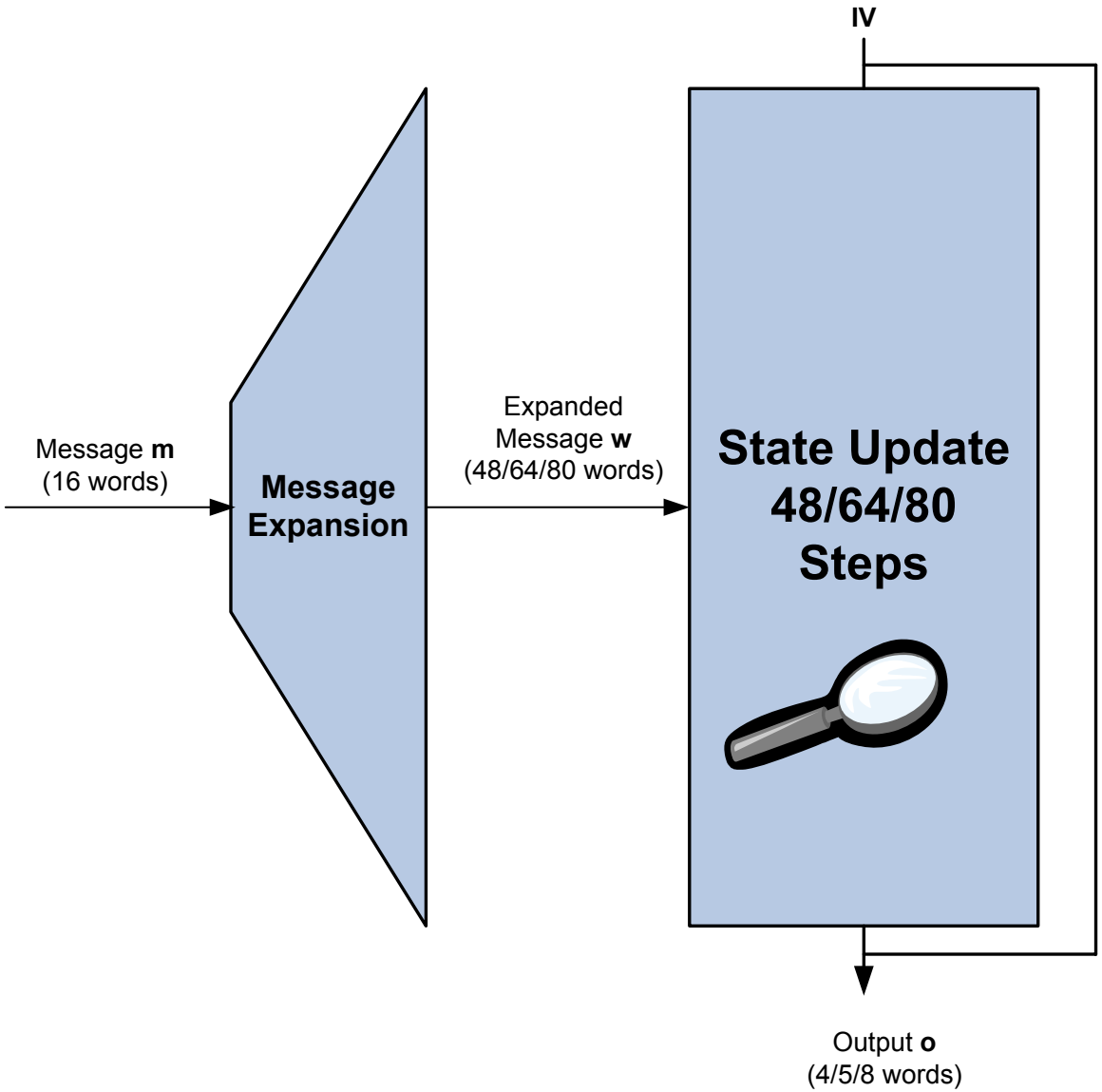
SHA / SHA-1



SHA-2 members

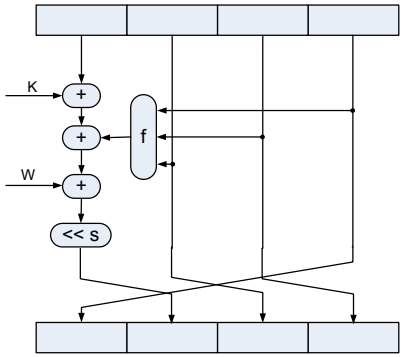


Outline of MD4-style Hash Functions

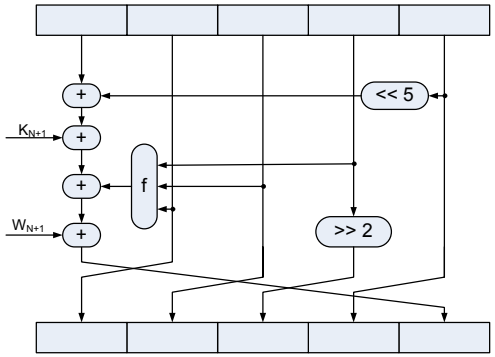


Evolution of the State Updates in the MD4 Family

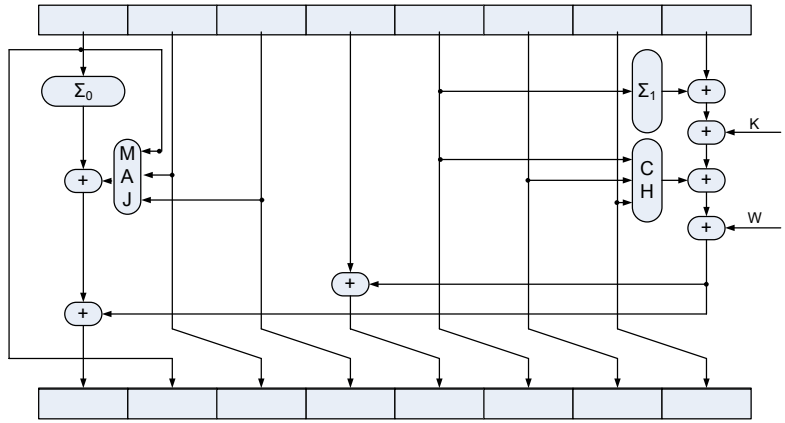
MD4



SHA/SHA-1



SHA-2 members



Design Complexity

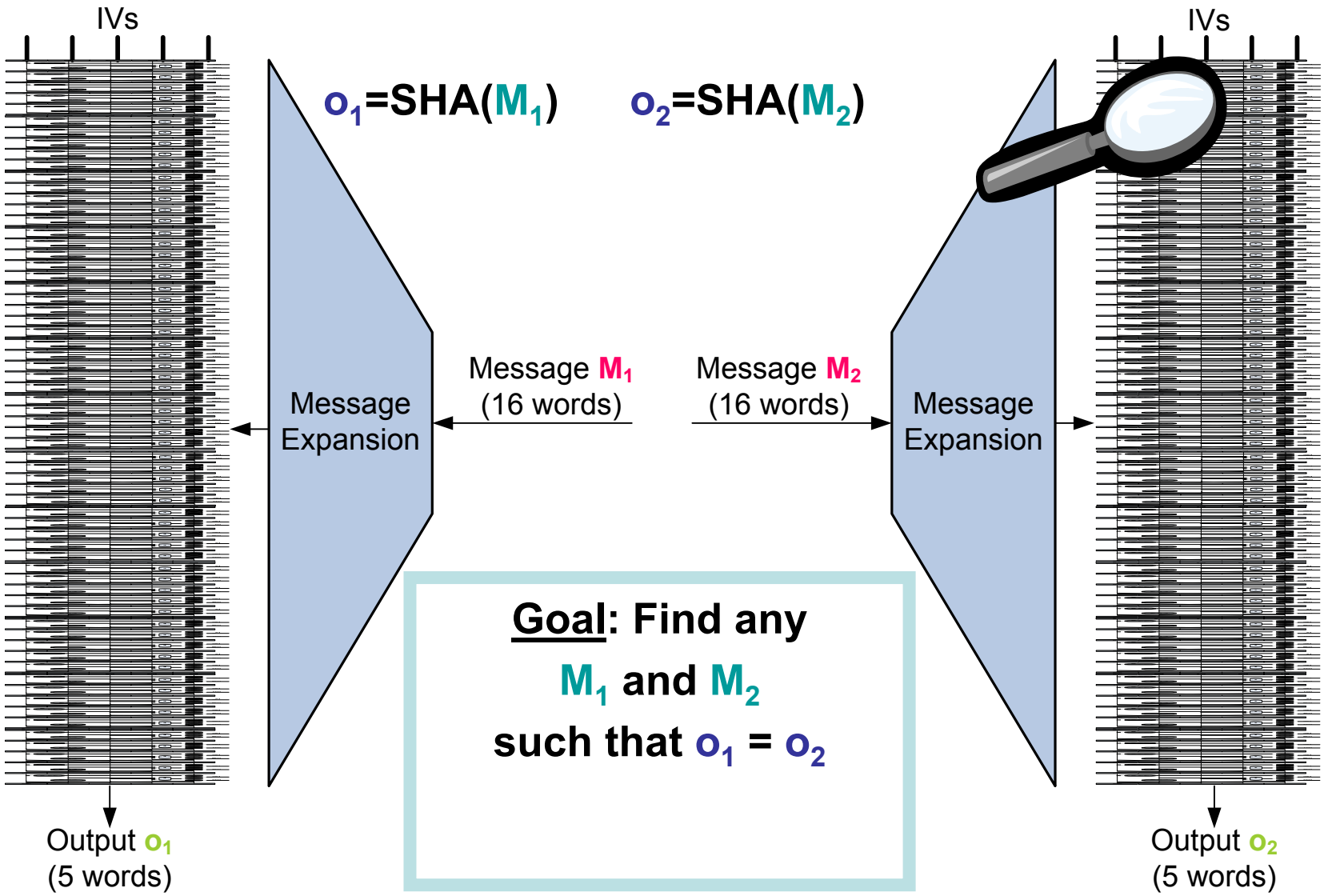
Cryptanalysis of MD4

- Several people have shown weaknesses in the compression function of MD4
  - Merkle [Mer90]
  - Bosselaers and den Boer [BB91]
  - Vaudenay [Vau94]
  - Dobbertin [Dob96]
- Dobbertin's techniques received the most attention in the scientific community [Dob98]
  - Lead to a fast collision-producing algorithm
  - Even (partially) meaningful collisions can be produced

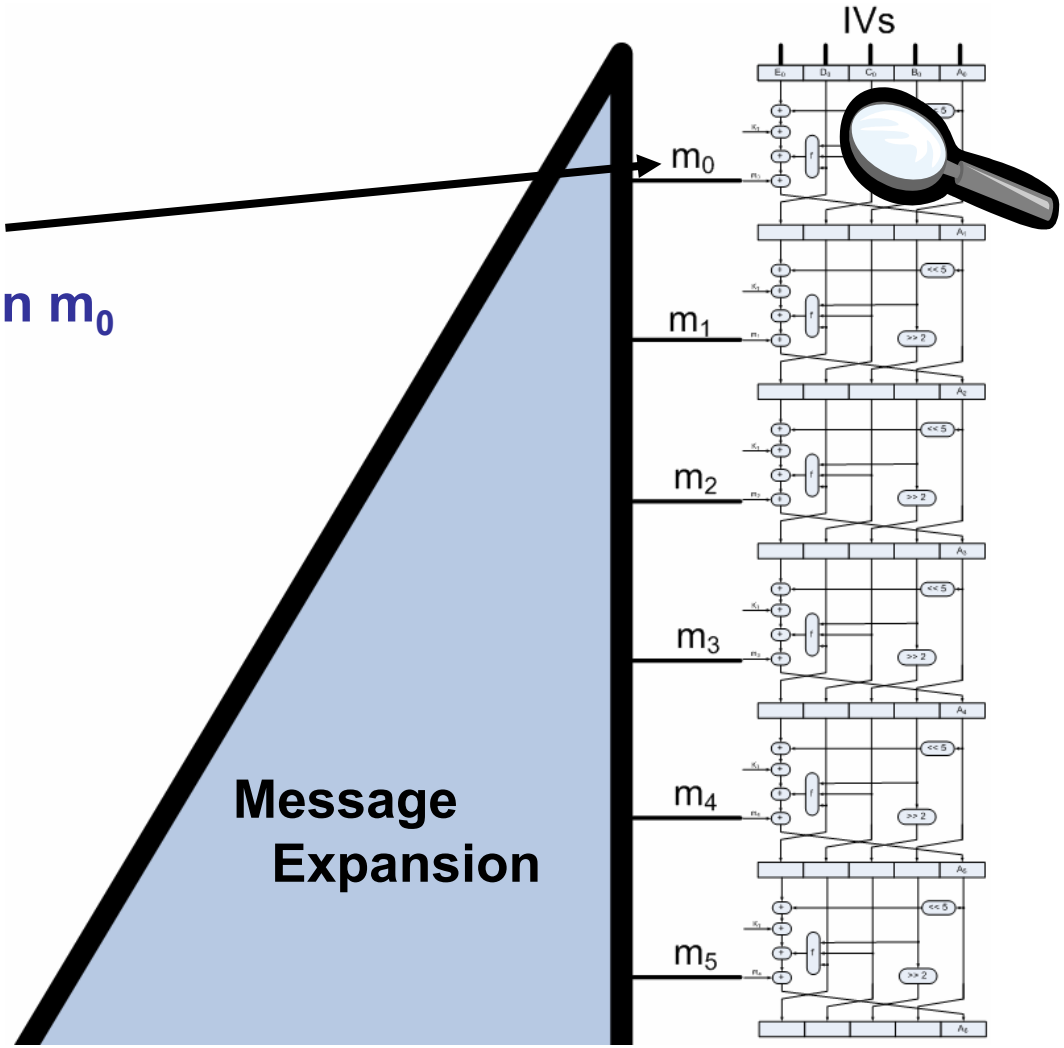
Cryptanalysis of SHA

- SHA
  - Chabaud and Joux [CJ98]
  - Biham and Chen [BC04]
  - Joux et al. [BCJ+05]
  - Wang et al. [WYY05a]
- SHA-1
  - Rijmen and Oswald [RO05]
  - Biham and Chen [BCJ+05]
  - Wang et al. [WYY05b]
  - De Cannière and Rechberger [DR06]

How to produce a collision?

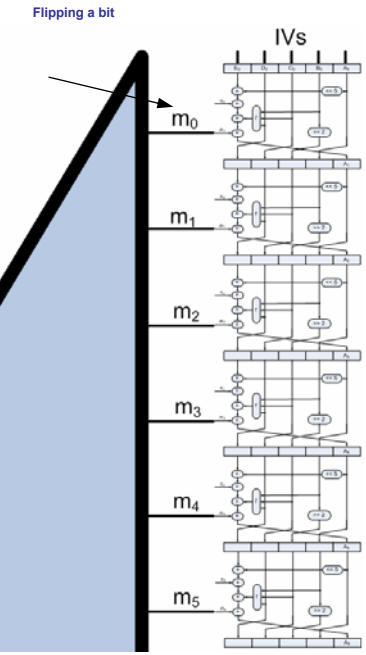


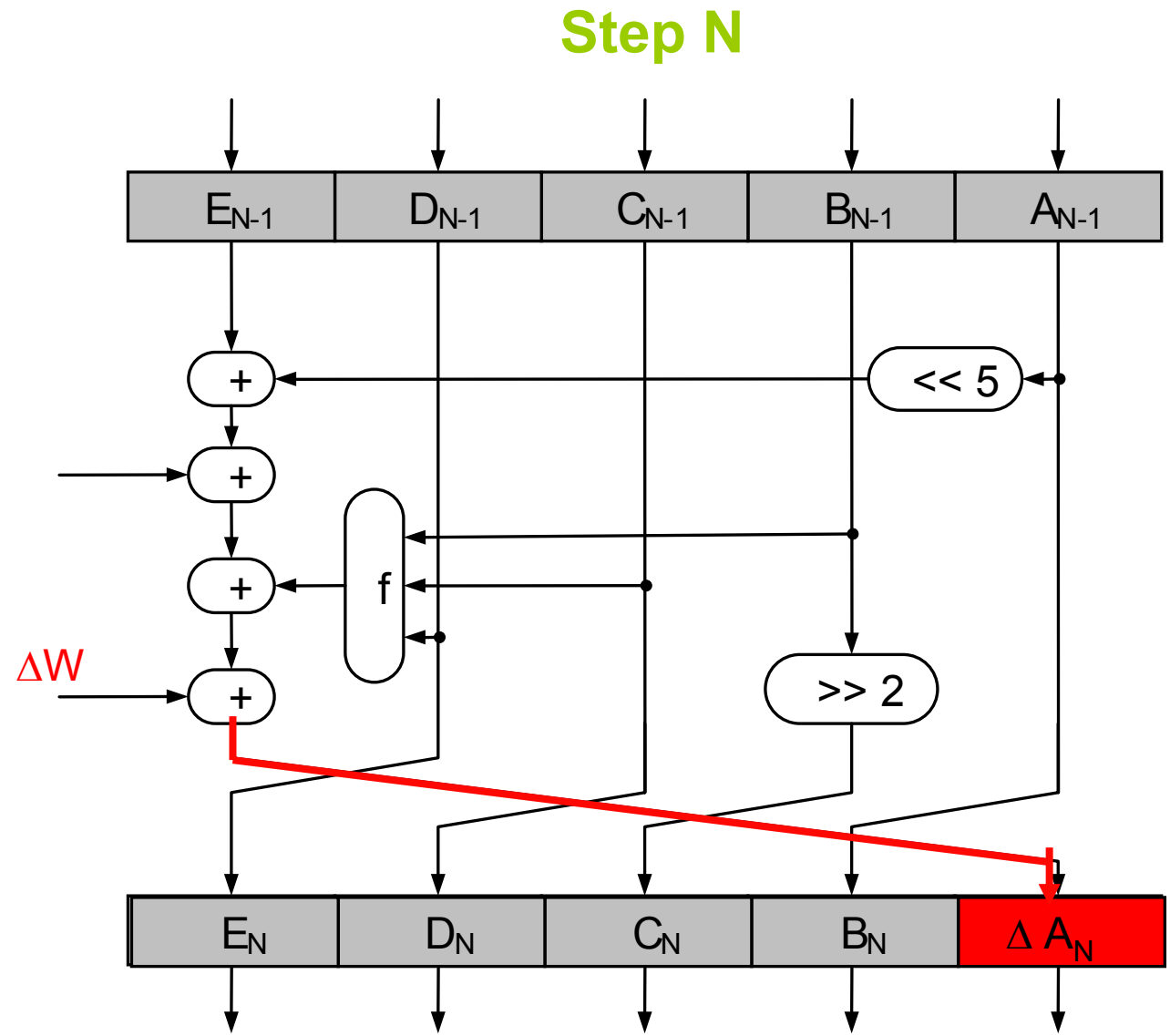
Propagation of a small difference



Flip a bit in m_0

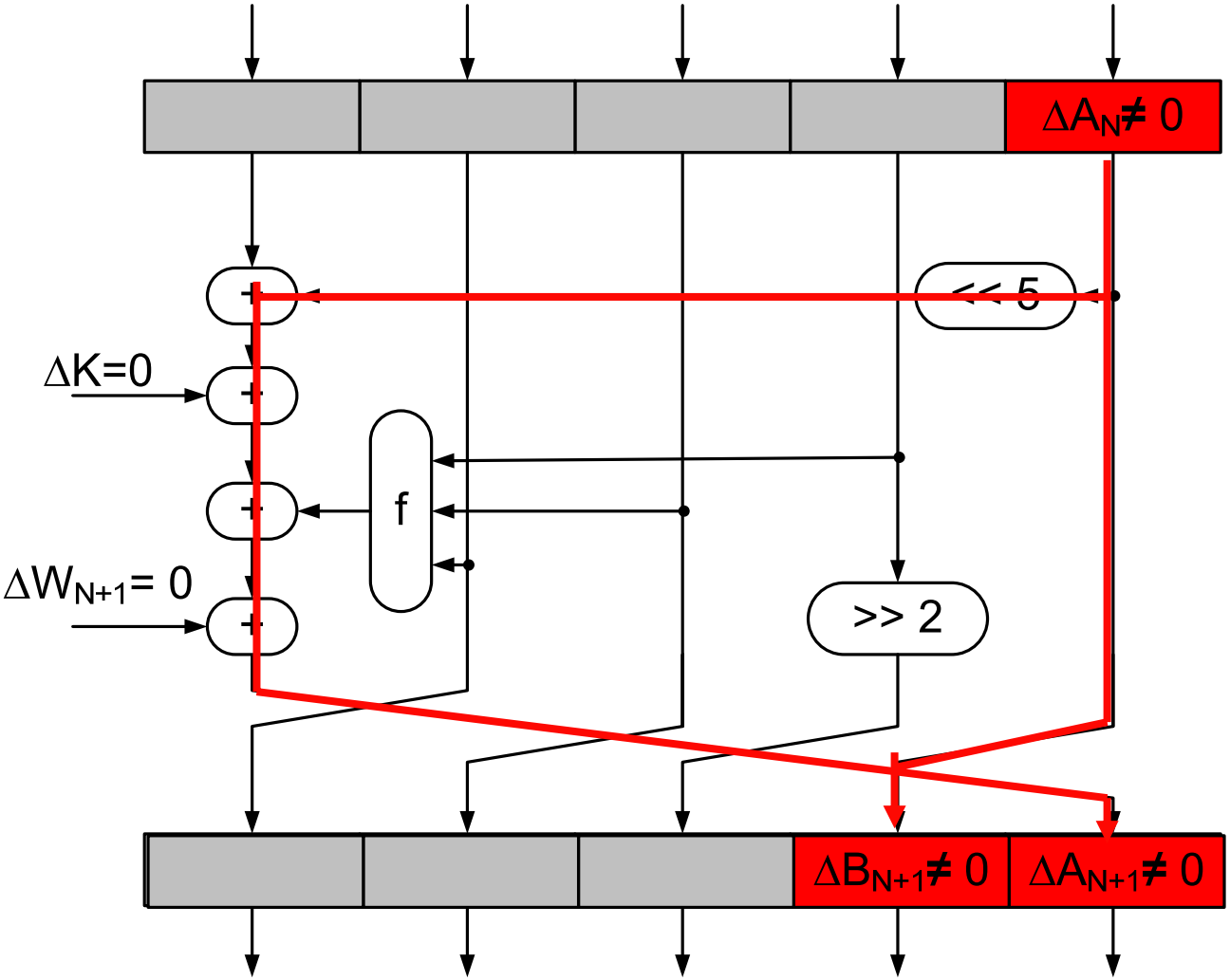
Propagation of a small difference





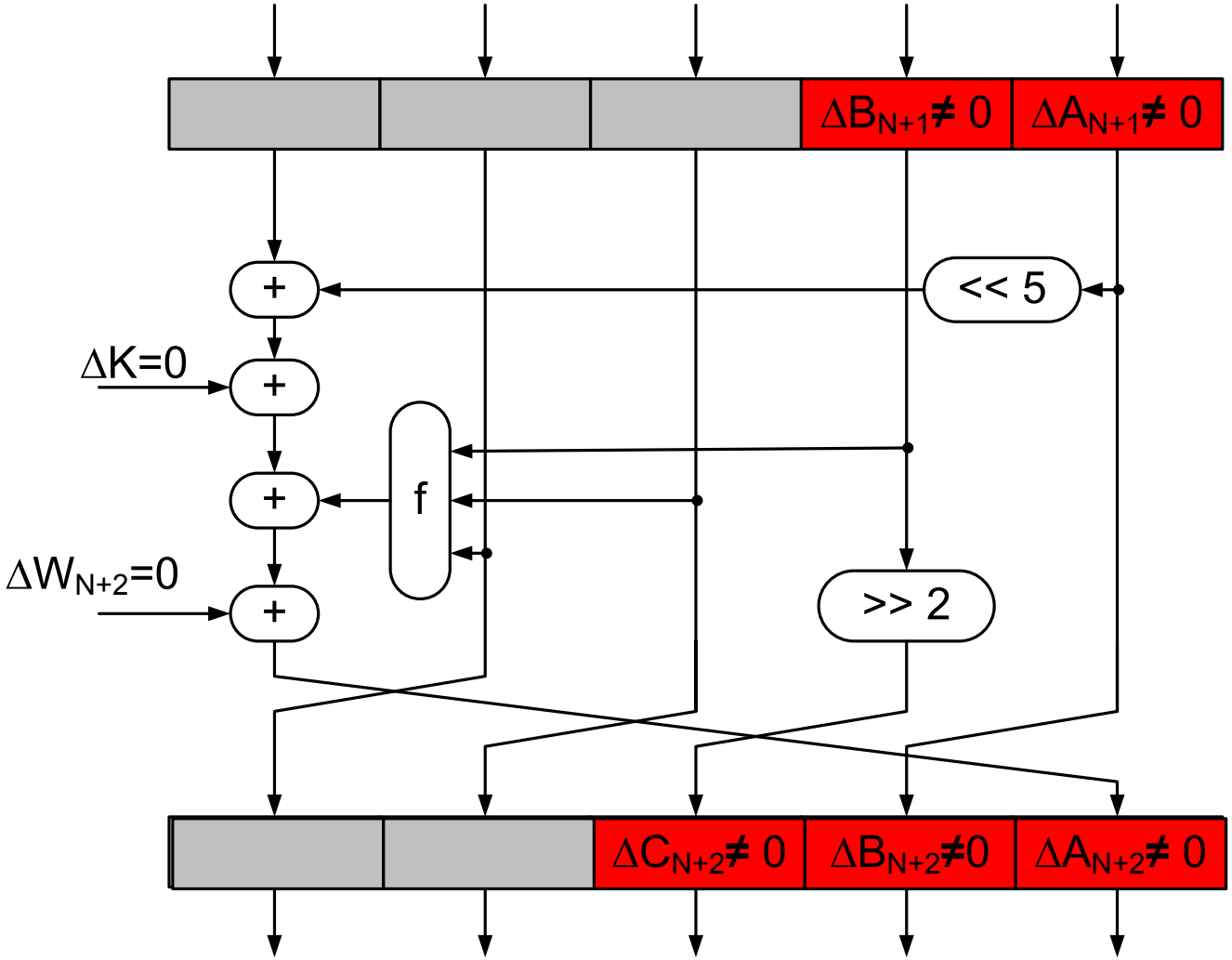
Propagation of a small difference

Step N+1



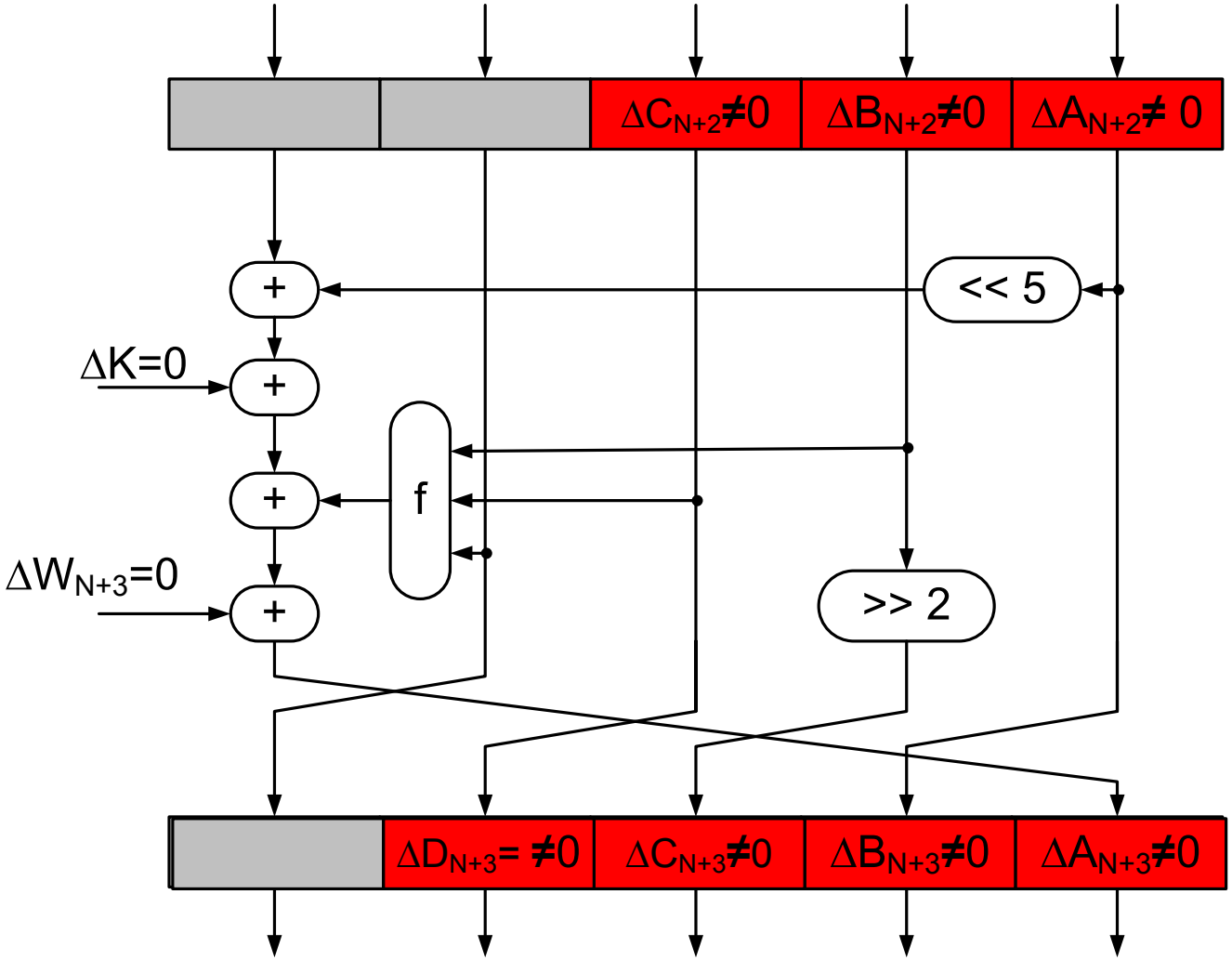
Propagation of a small difference

Step N+2

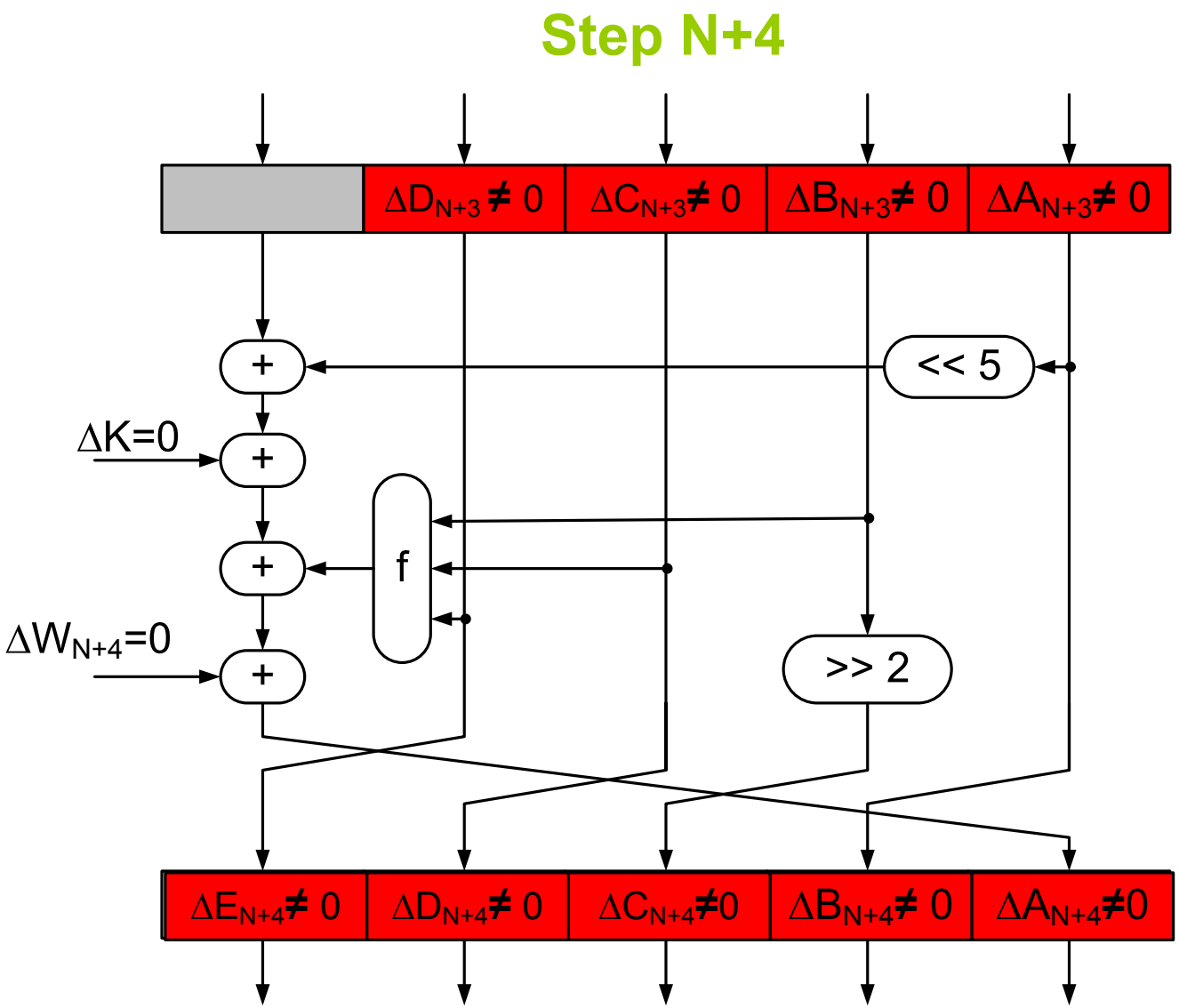


Propagation of a small difference

Step N+3



Propagation of a small difference



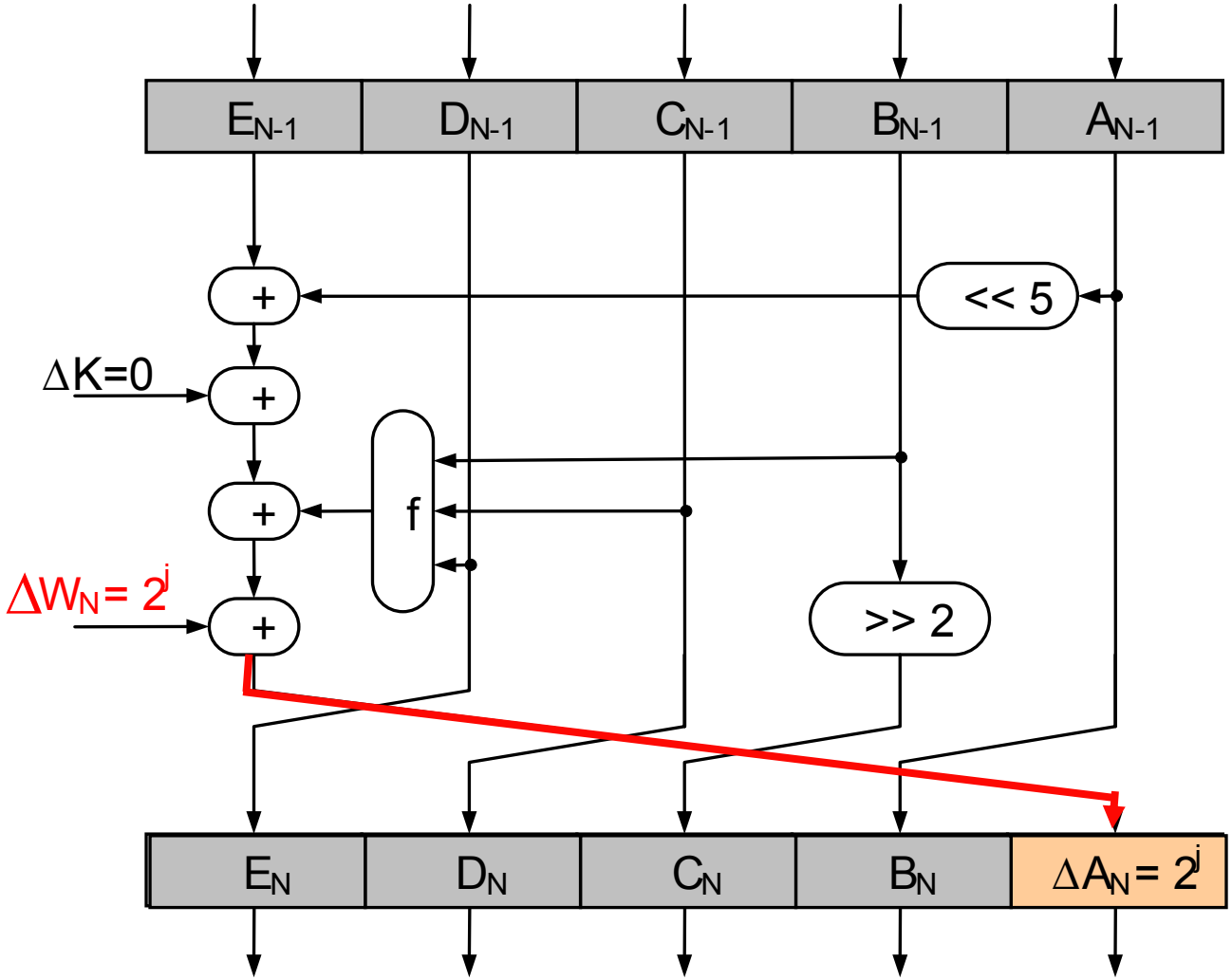
Differential cryptanalysis of SHA

- Small differences quickly expand
- Approach by Chabaud & Joux [CJ98]
  - Perturbations and corrections
  - Theoretical attack on SHA
  - Later on improved to practical attack [BCJ+05]

- Difference = XOR operation

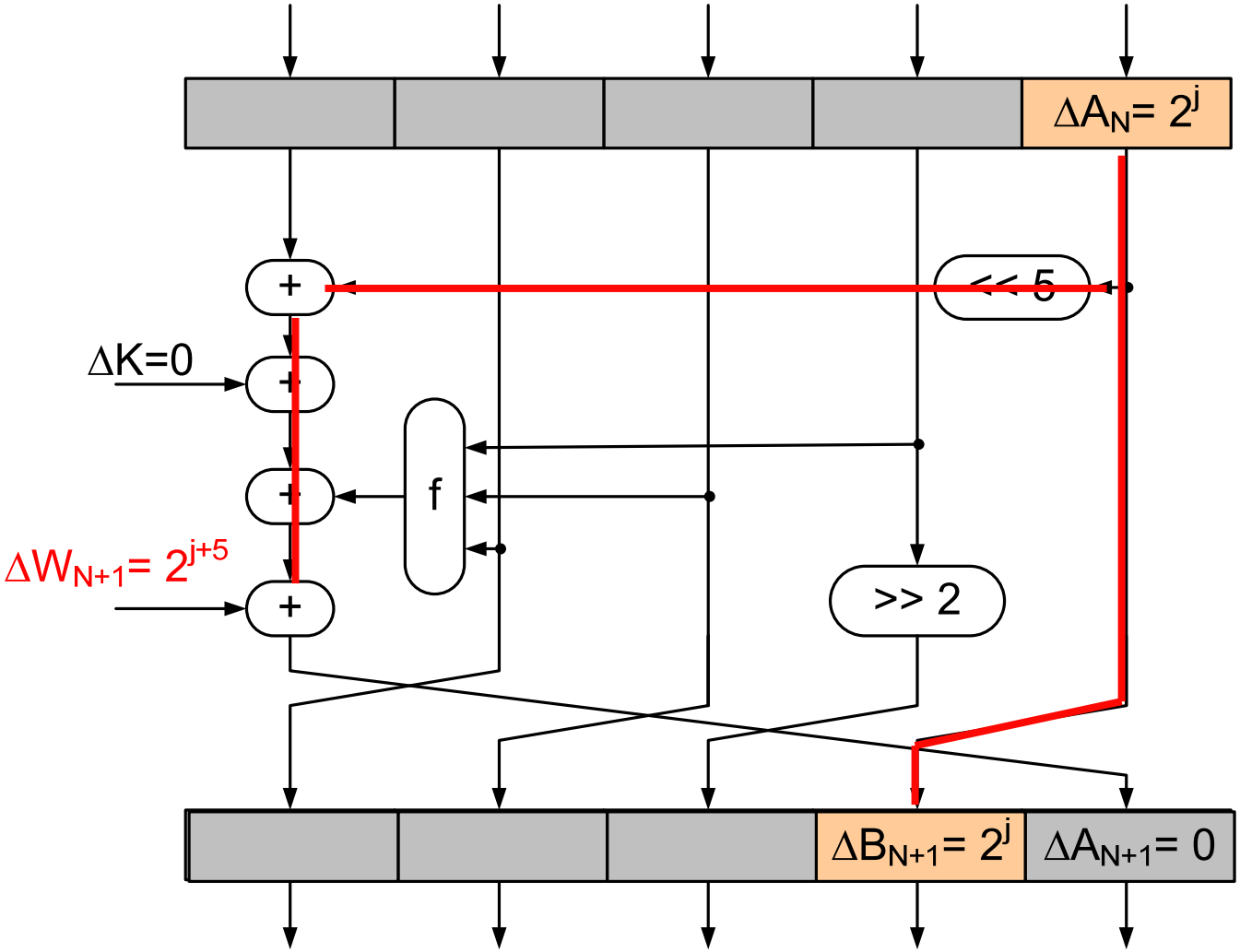
Perturbation

Step N



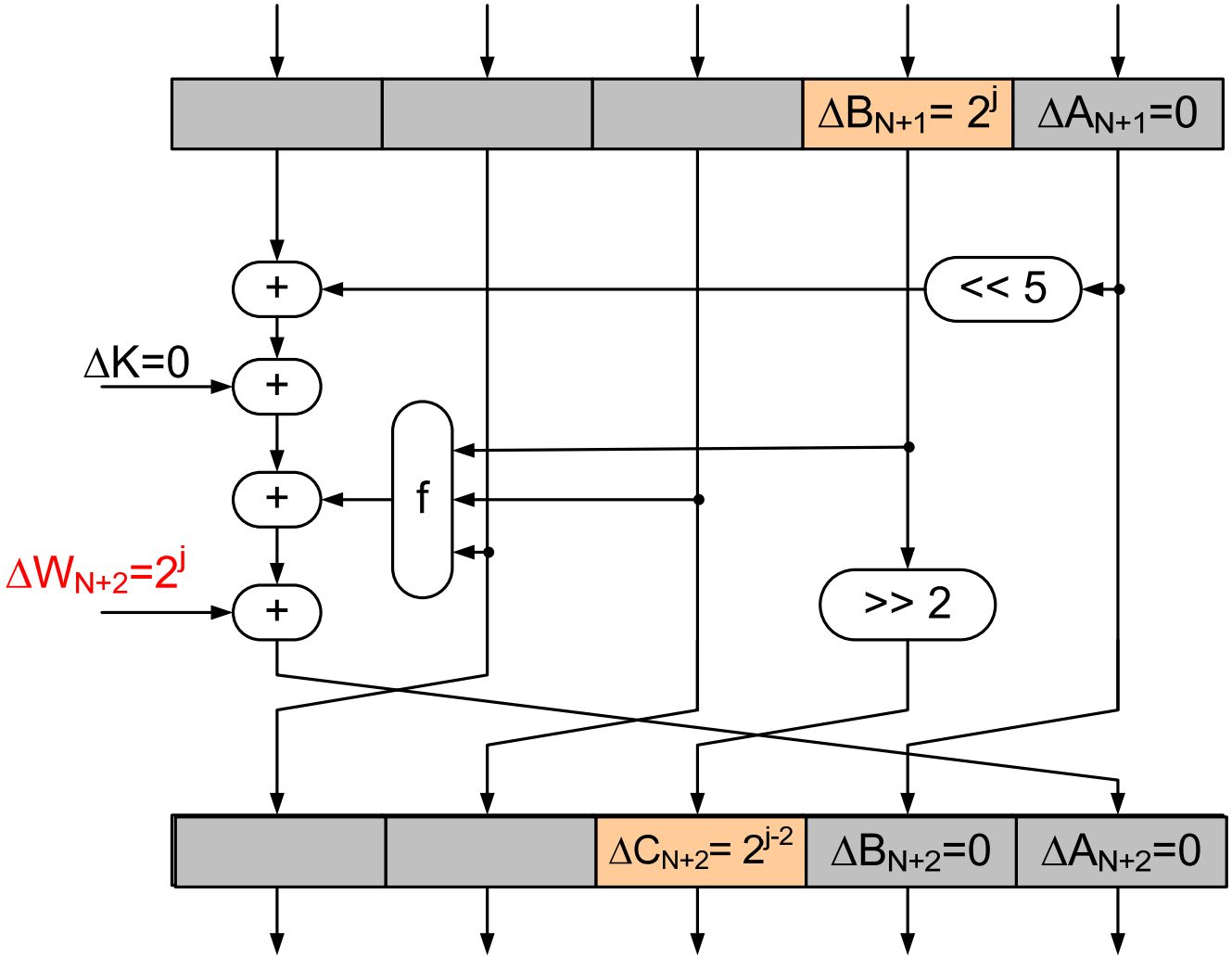
Correction 1

Step N+1



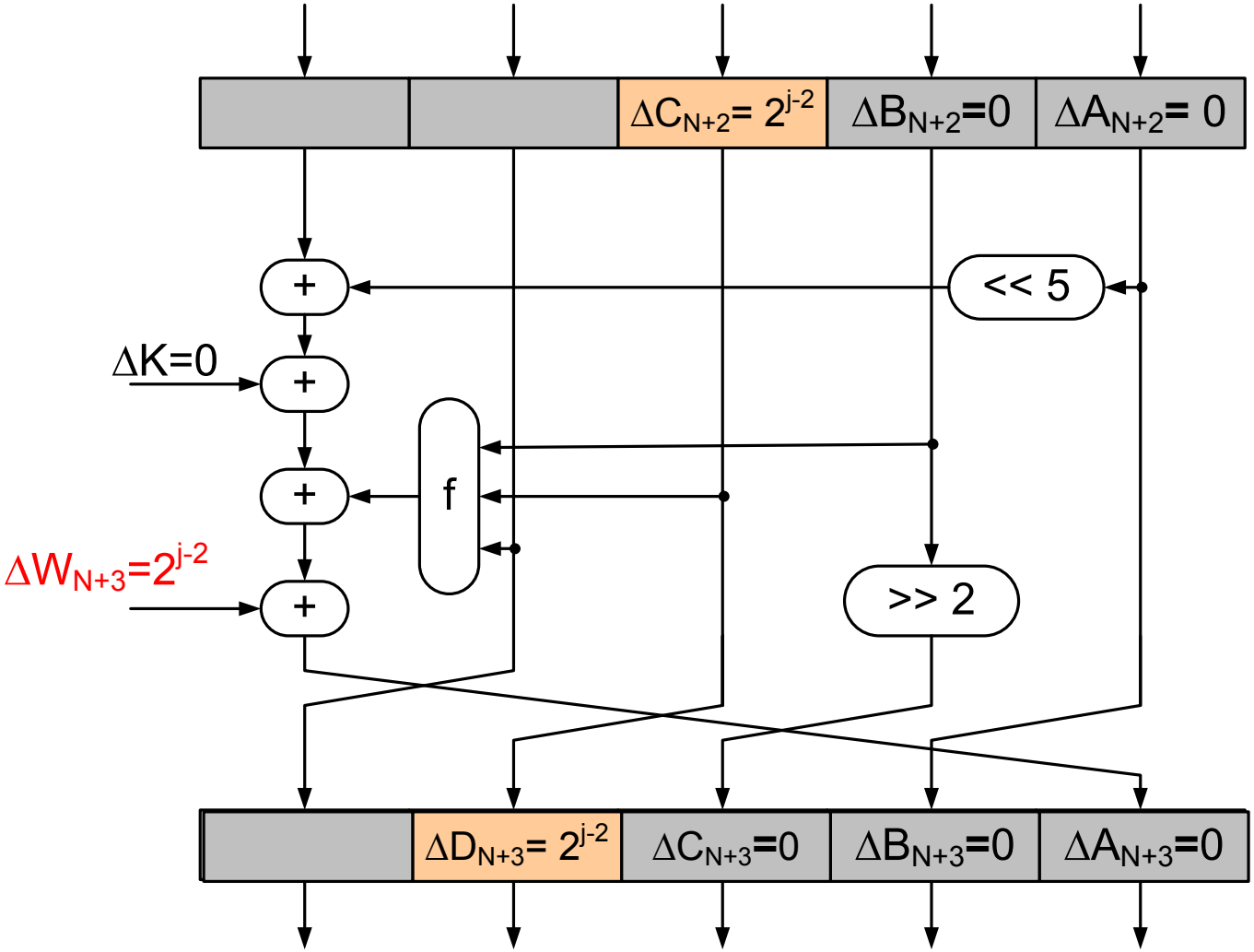
Correction 2

Step N+2



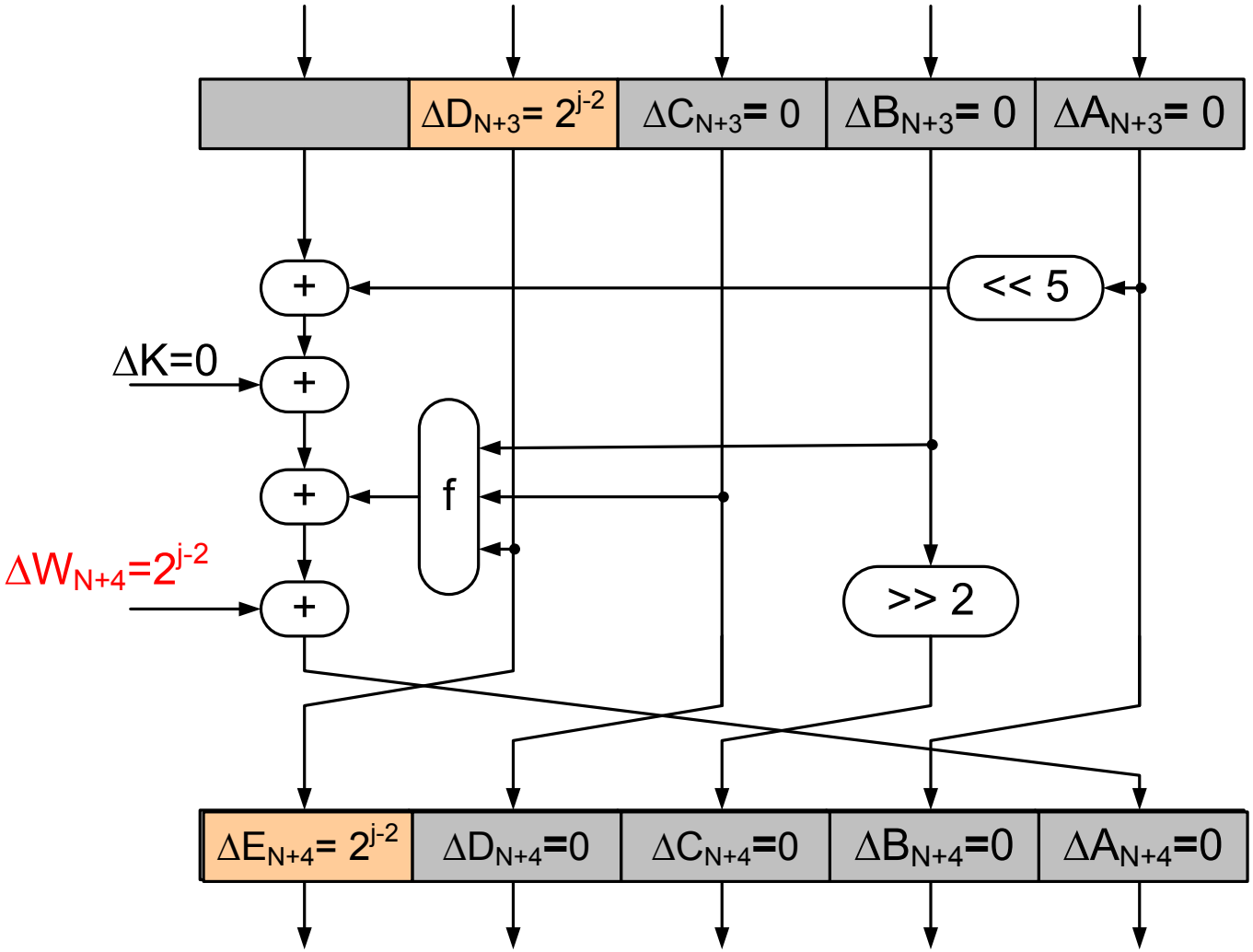
Correction 3

Step N+3



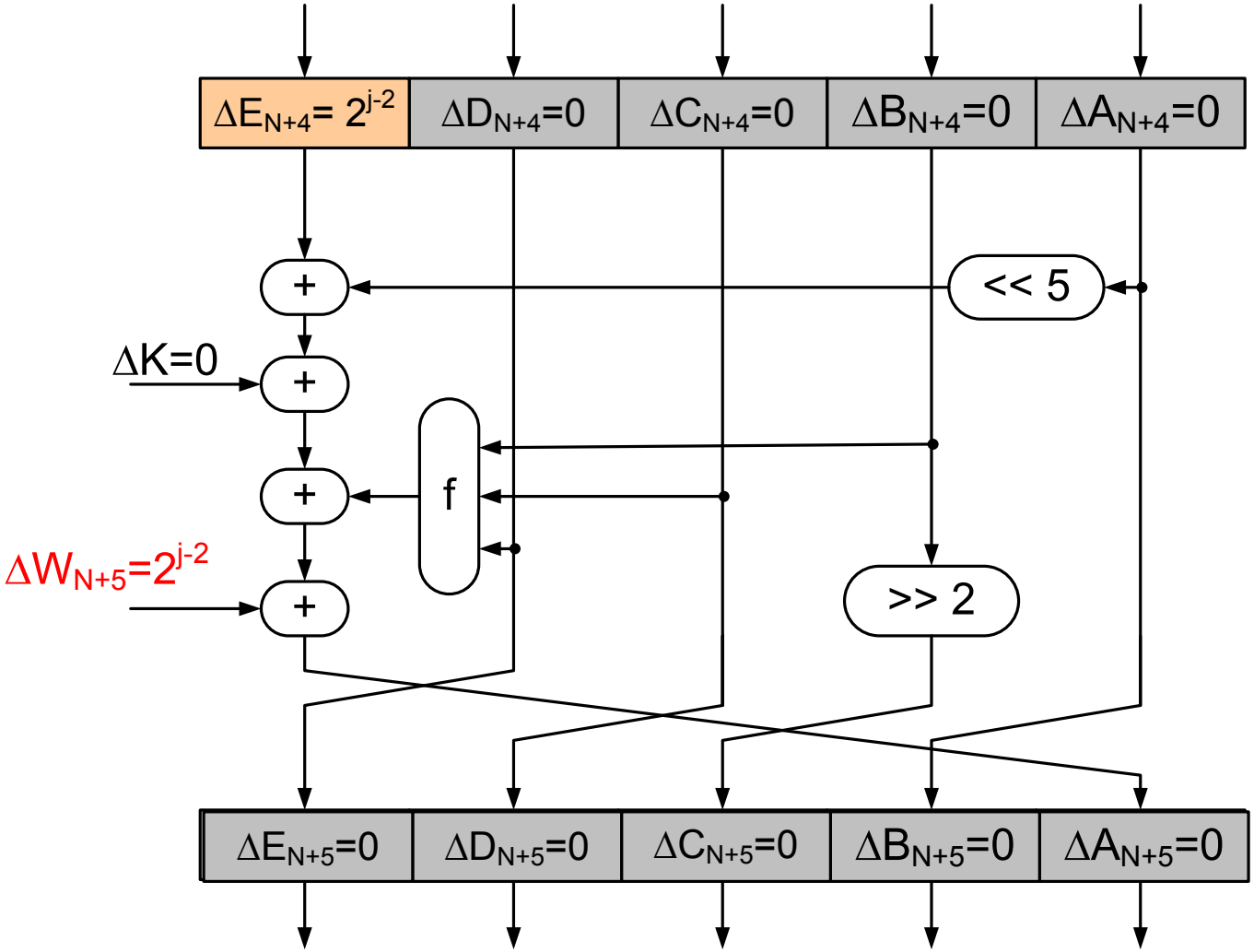
Correction 4

Step N+4



Correction 5

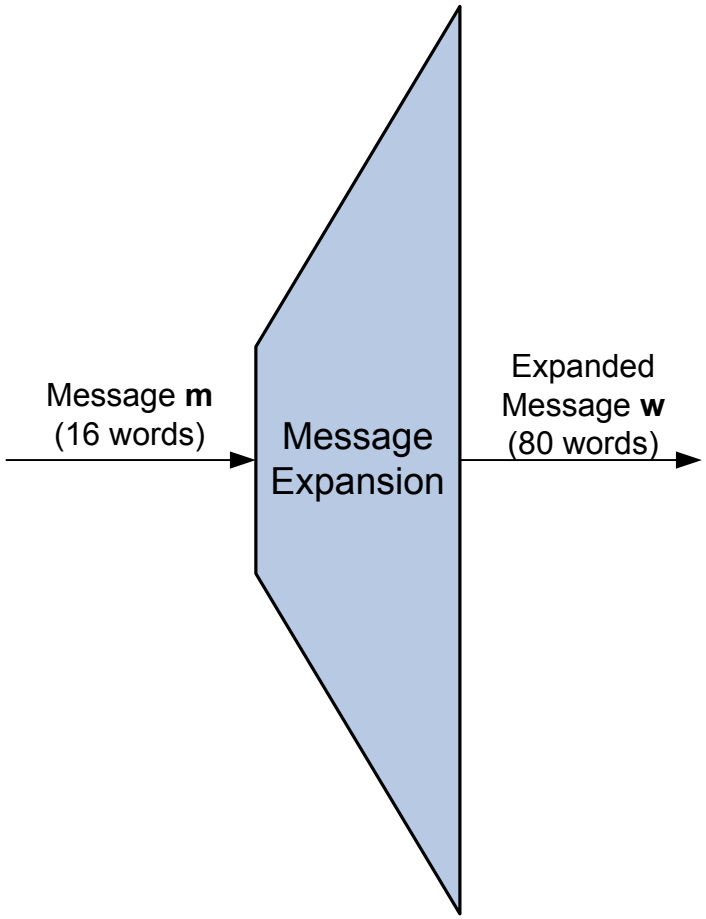
Step N+5



Local collision

- Resynchronisation of internal state
- One perturbation and 2-5 corrections
- Creating local collisions is not so difficult
- Problem: message expansion

Outline of SHA – Message Expansion



SHA

$$W_t = \begin{cases} M_t & for\,(0 \le t \le 15) \\ W_{t-3} \oplus W_{t-8} \oplus W_{t-14} \oplus W_{t-16} & for\,(16 \le t \le 79) \end{cases}$$

SHA-1

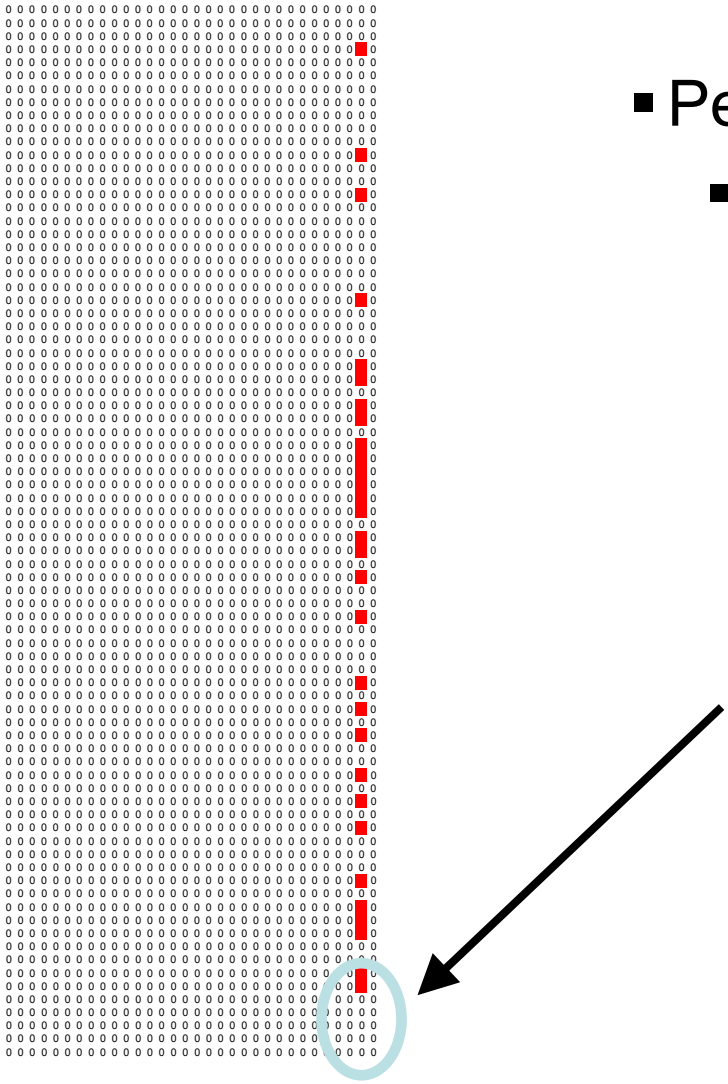
$$W_t = \begin{cases} M_t & for\,(0 \le t \le 15) \\ (W_{t-3} \oplus W_{t-8} \oplus W_{t-14} \oplus W_{t-16}) <<1 & for\,(16 \le t \le 79) \end{cases}$$

Message expansion

- Every bit changed influences other bits
- Impossible to find m_1, m_2 such that ME(m_1), ME(m_2) differ in 3-6 bits only

- Constructing a global collision = finding good characteristic

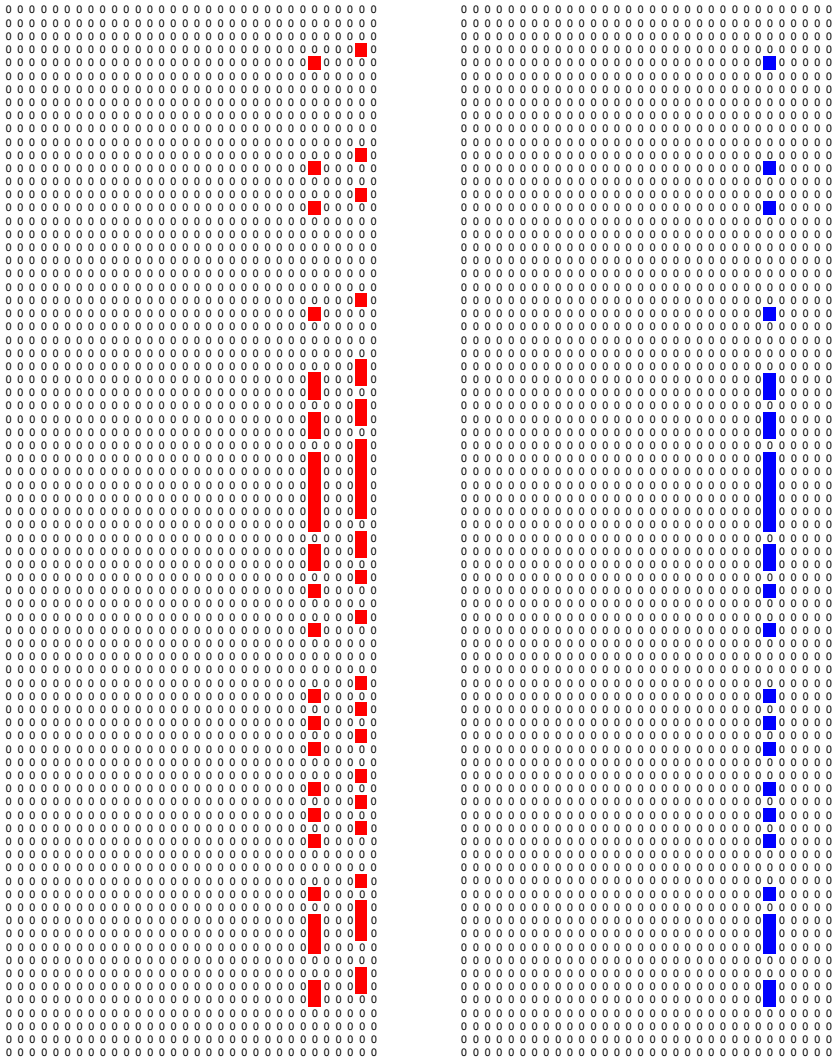
Building a collision for SHA



- Perturbation pattern
  - Low weight

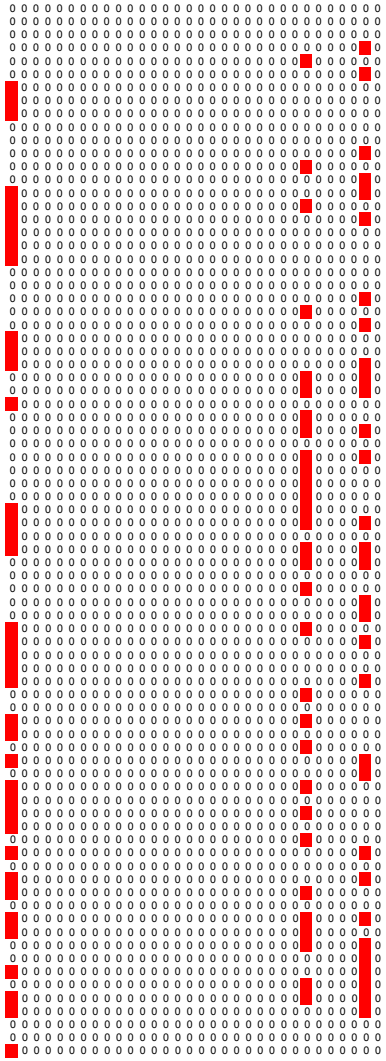
$$W_t = W_{t-3} \oplus W_{t-8} \oplus W_{t-14} \oplus W_{t-16}$$

A collision-producing difference pattern



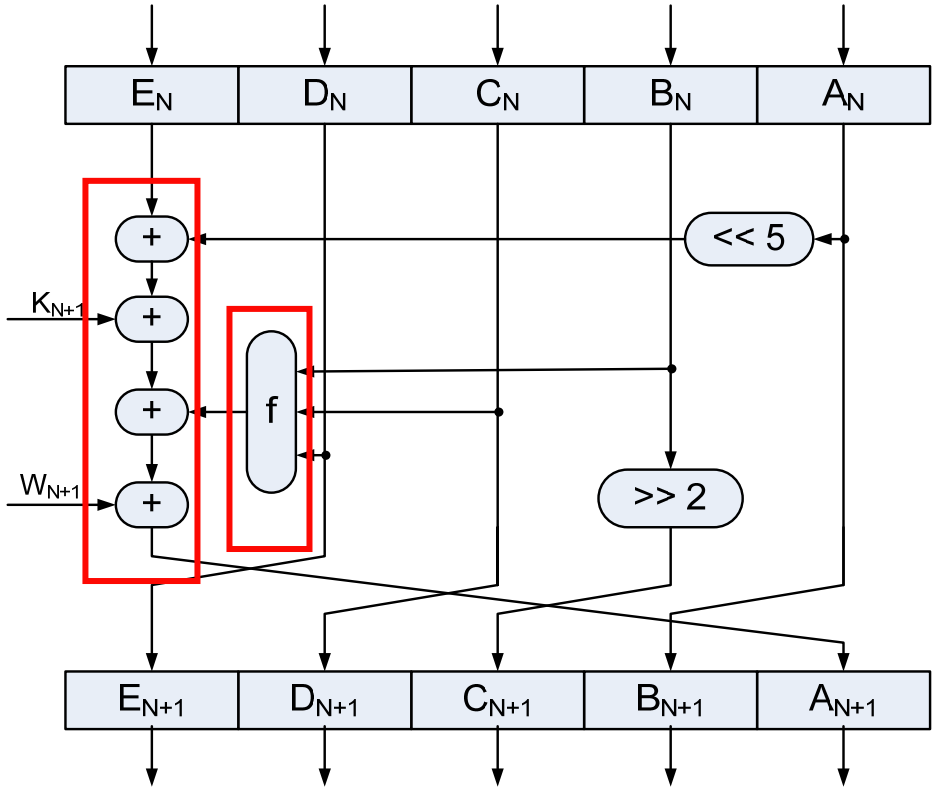
- Apply 5 corrections with the same pattern
  - displaced over steps
  - rotated over bit positions

A collision-producing expanded-message difference pattern



- Completed difference pattern consisting of
  - 1 perturbation pattern
  - 5 correction patterns

Conditions imposed by nonlinear elements



- Boolean function f
- Modular addition

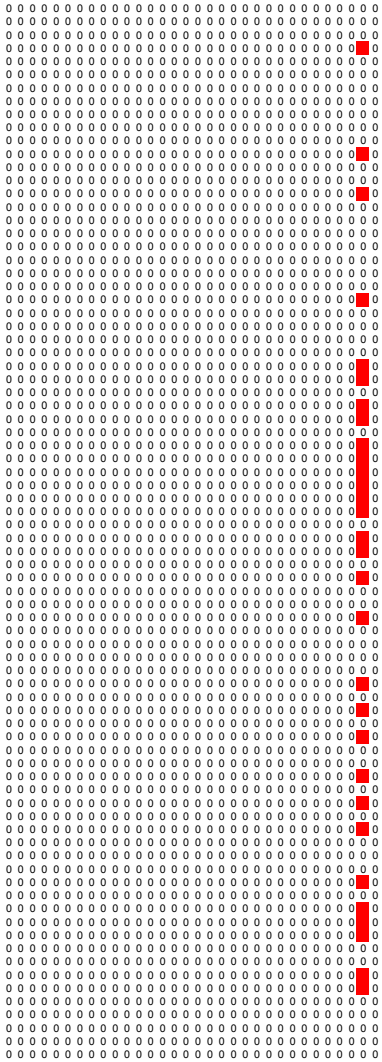
Modular addition

- Linear except for carry effects
  - Carry = 0 with probability 1/2
  - Carry moves upwards only
  - No carry from MSB
- Requirement: difference propagation as with XOR

Boolean functions

- Linear in 40 out of 80 steps
- Bitwise parallel: every input bit affects 1 output bit
- We set as requirement: difference propagations as with XOR
  - High probability
  - Easy to find good characteristics

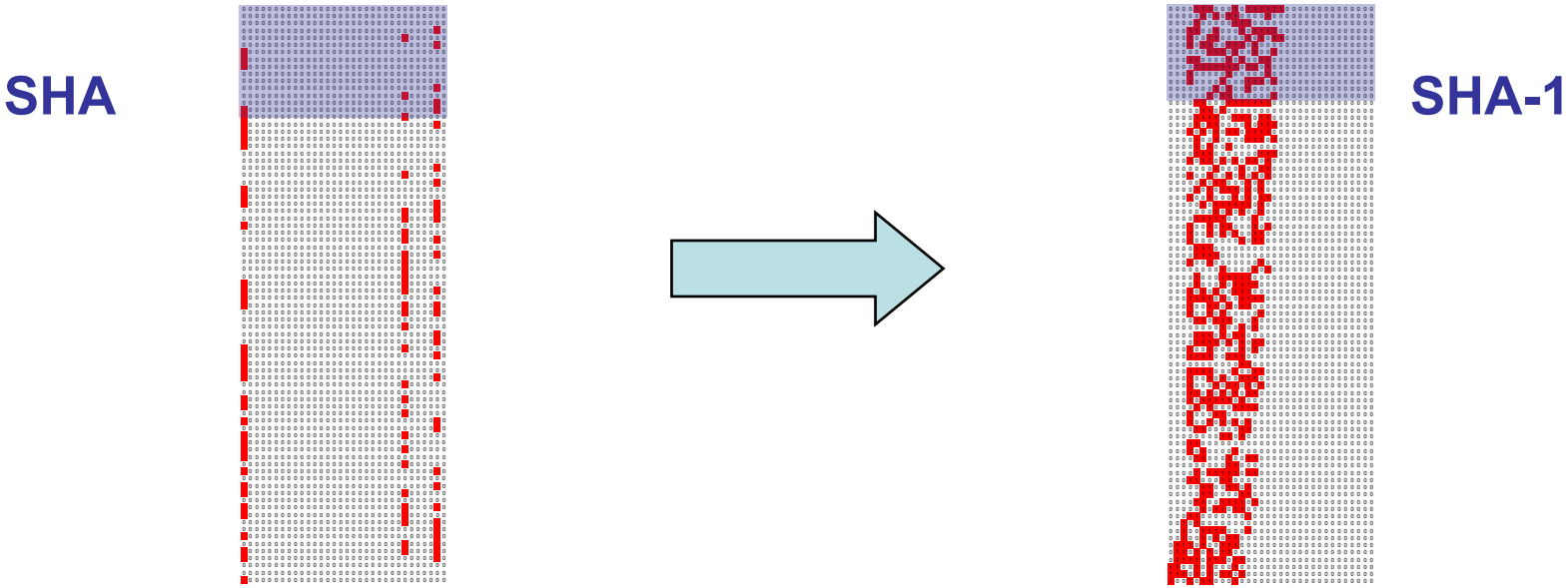
Conditions



- Every perturbation gives 2-5 conditions on the message
- Most conditions are nonlinear and complicated to express in terms of the input message
- Goal is to minimize these conditions (to make final search easier)

Results of CJ98

- Low- weight patterns exist for SHA => break

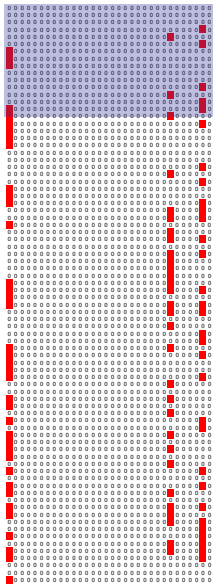


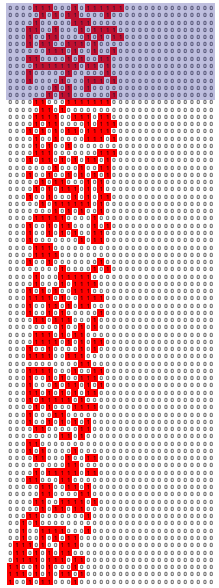
- For SHA-1: weight is too high

Results of CJ98

- Low- weight patterns exist for SHA => break

SHA





SHA-1

- For SHA-1: weight is too high

Improvements

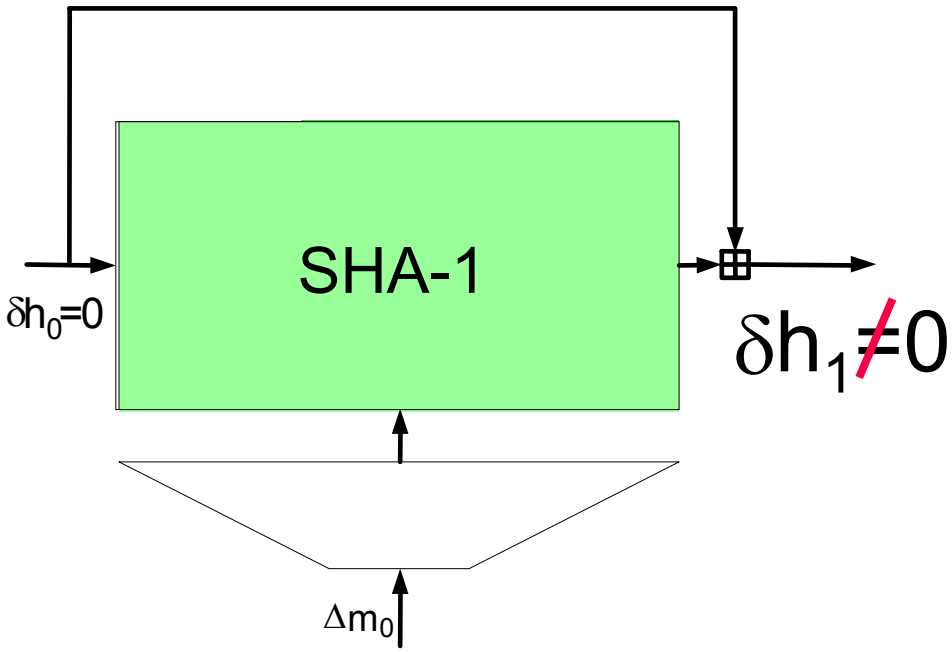
- Better characteristics
  - 1-block ➔ multi-block
  - Better suited for hash functions

- Better ways to construct right pairs
  - Message modification

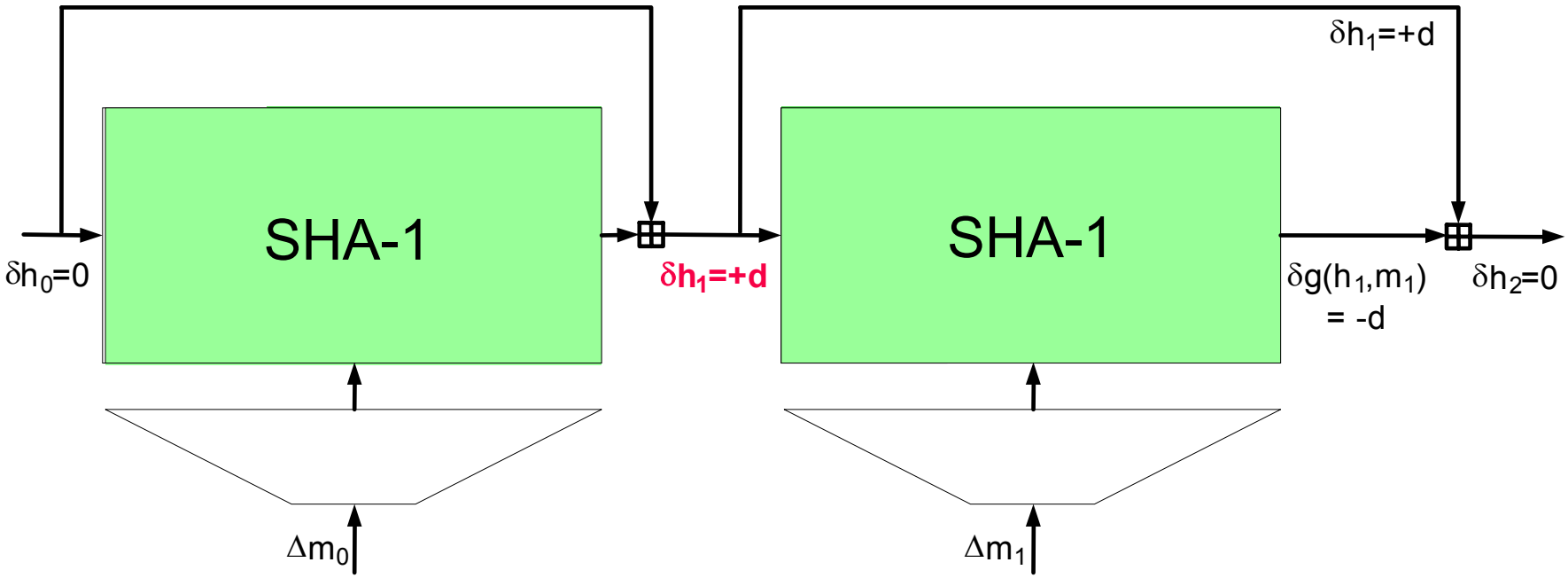
Multi-block collisions

- Near-collision: outputs differ in only a few bits
- Observation: much easier than collisions
  - Characteristics with higher probability

Use of near-collisions



Use of near-collisions

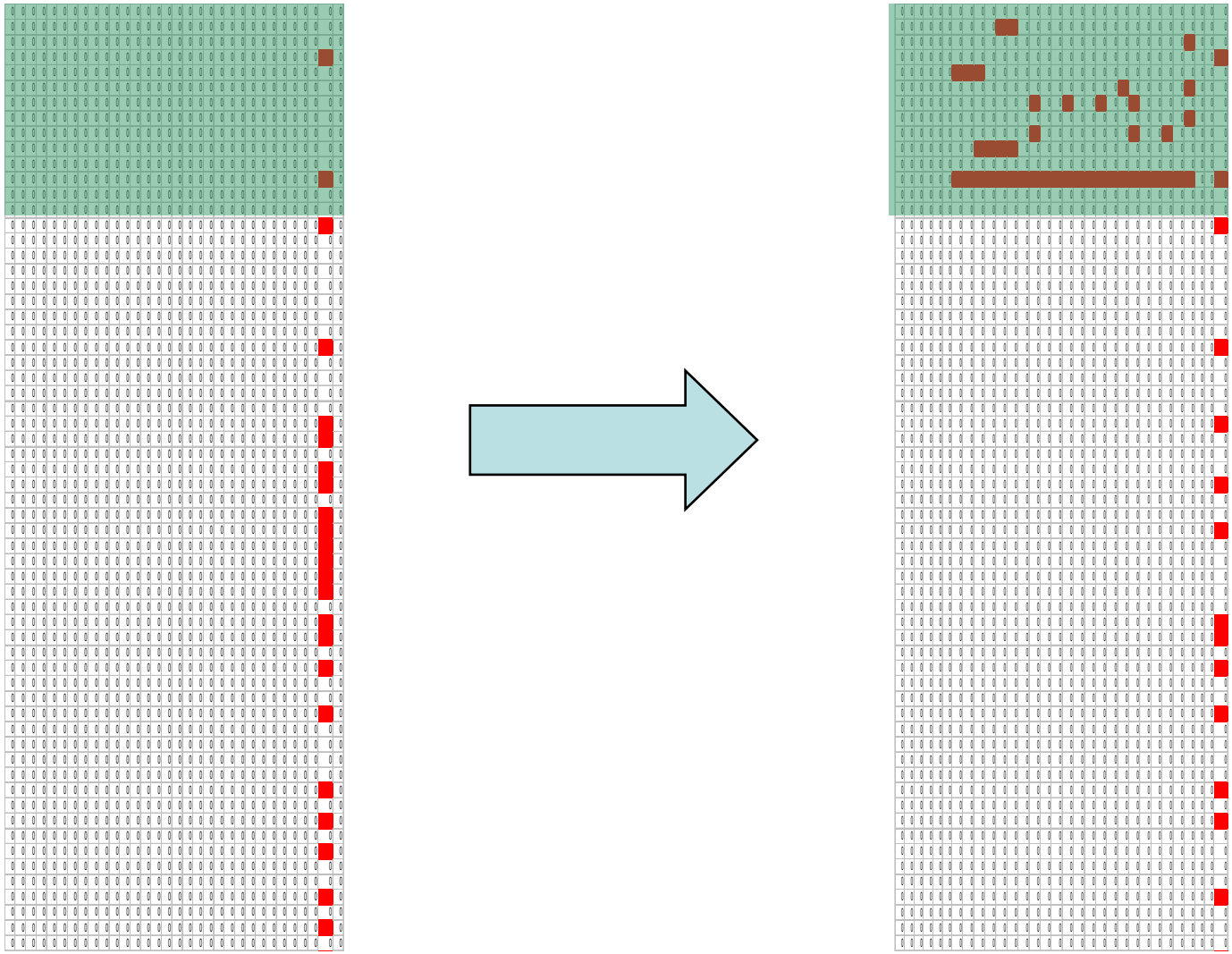


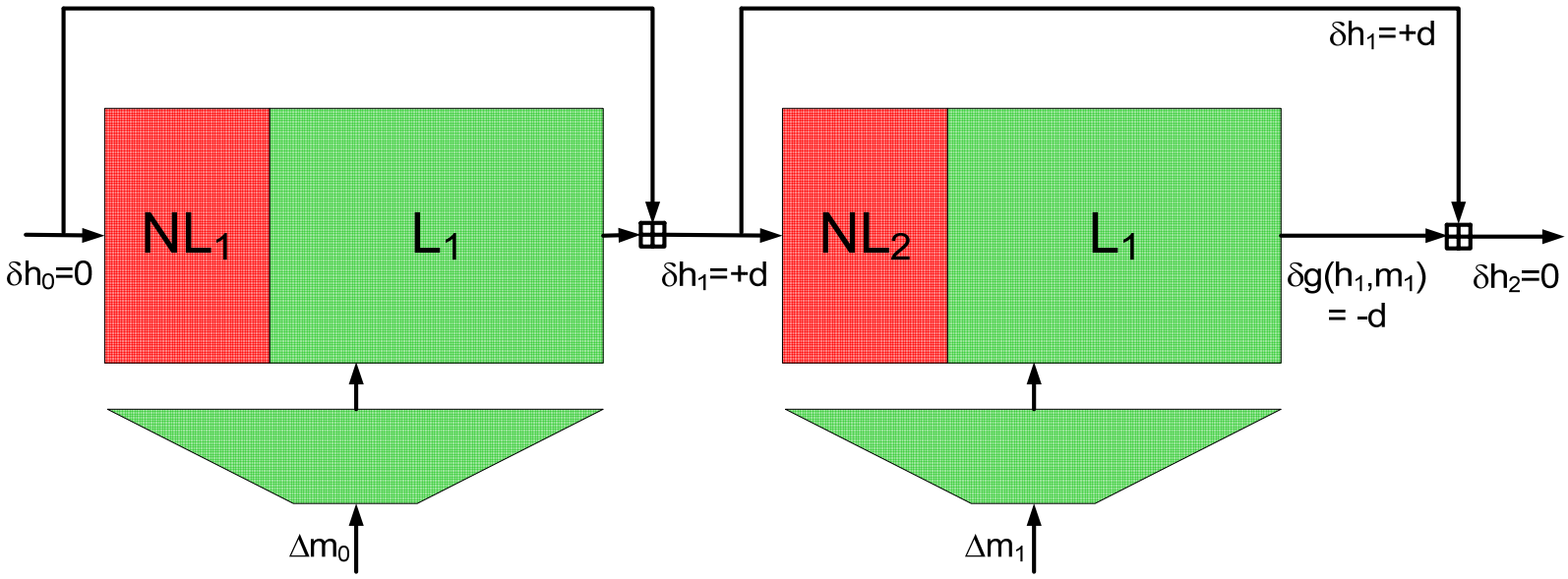
- Two related near-collisions give a 2-block collision
- Work effort of two blocks is only double of one block

Differential attacks on hash functions

- Good characteristics for block ciphers:
  - Optimise probability
  - Minimise number of chosen plaintexts
- Good characteristics for hash functions
  - Optimise probability
  - Minimise effort to solve equations
  - Equations in first steps are always easy
    - Only a small part of the message involved
    - Inputs are known
    - Late start / Early stop

Good characteristics





- Two key techniques of Wang et al.:
  - Manually find suitable complex characteristic NL_1 and NL_2
  - Advanced message modification to improve work factor

- Methods are rather ad hoc (manual)
- Optimization?

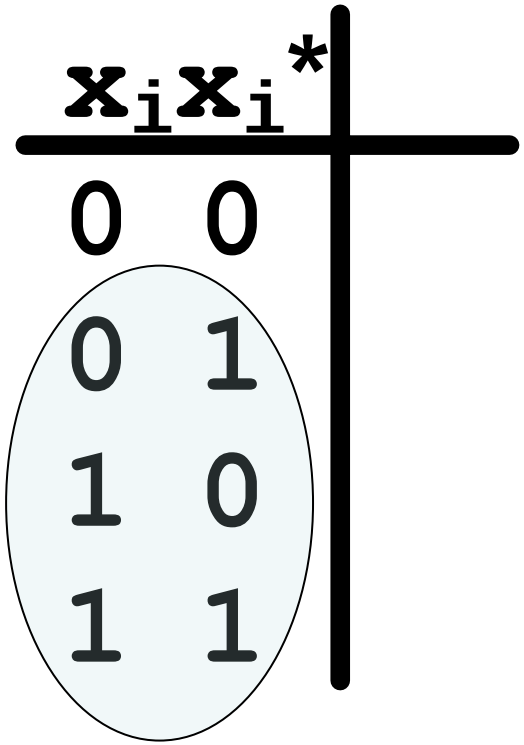
Optimising characteristic

- Optimise probability after step 16
- Before step 16
  - Concentrate low probability in few steps
  - More difficult search problem
    - Manual construction [WYY05b]
    - Automatic tools [DR06]

How does it work?

Principles

- Generalized conditions



Type	Possibilities
XOR	2
Signed-bit	4-6
Generalized:	**16**

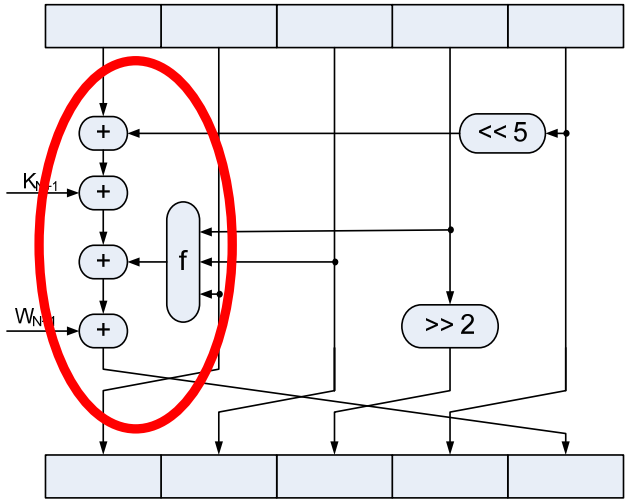
Generalized Conditions - Notation

(x_i, x_i^*)	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
?	✓	✓	✓	✓
-	✓	-	-	✓
x	-	✓	✓	-
0	✓	-	-	-
u	-	✓	-	-
n	-	-	✓	-
1	-	-	-	✓
#	-	-	-	-

Principles

- Generalized conditions

- Use “bit-sliced design” to efficiently
  - Propagate conditions *within one* step transformation
  - Propagate conditions *among all* step transformations



Animated Illustration

Key properties of new approach [DR06]

- Looks for (parts of) the colliding pair and characteristic at the same time
- *Precise* calculation of probabilities
  instead of approx. by HW and counting conditions
- Available degrees of freedom can be used to
  - Direct improvements of probability (Greedy approach)
  - Facilitate final search method for a right pair
  - Fix part of the colliding message on beforehand and/or during final search (meaningful)

Constructing right pairs

- Equations following from nonlinear operations
  - Every step increases the complexity
- First 15 steps: easily solvable
- Next steps: mostly guess and verify
- Solving of eqs. in steps 16 and ff.
  - Neutral-bit technique [Biham and Chen]
  - Advanced message modification [Wang et al.]
  - Symbolic computation [Sugita et al.]
  - Greedy method [De Cannière and Rechberger]
  - Boomerang method [Joux and Peyrin]

Neutral bits

- Weak diffusion
- After small number of steps, not all output bits depend on all input bits

- When trying pairs, only vary bits that don't change the outputs which are already right

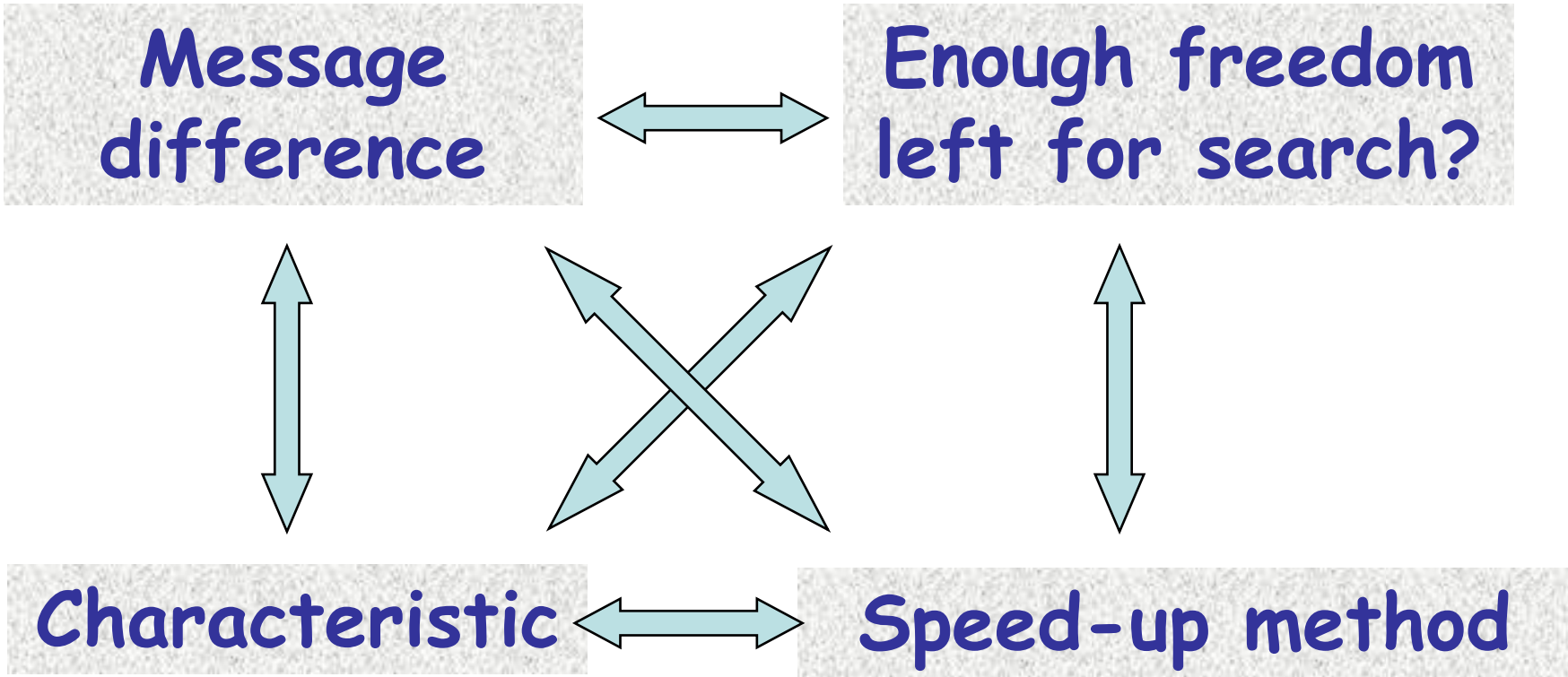
Advanced message modification

- Solve equations deterministically
  - Iterative solving strategy
- Again helped by weak diffusion

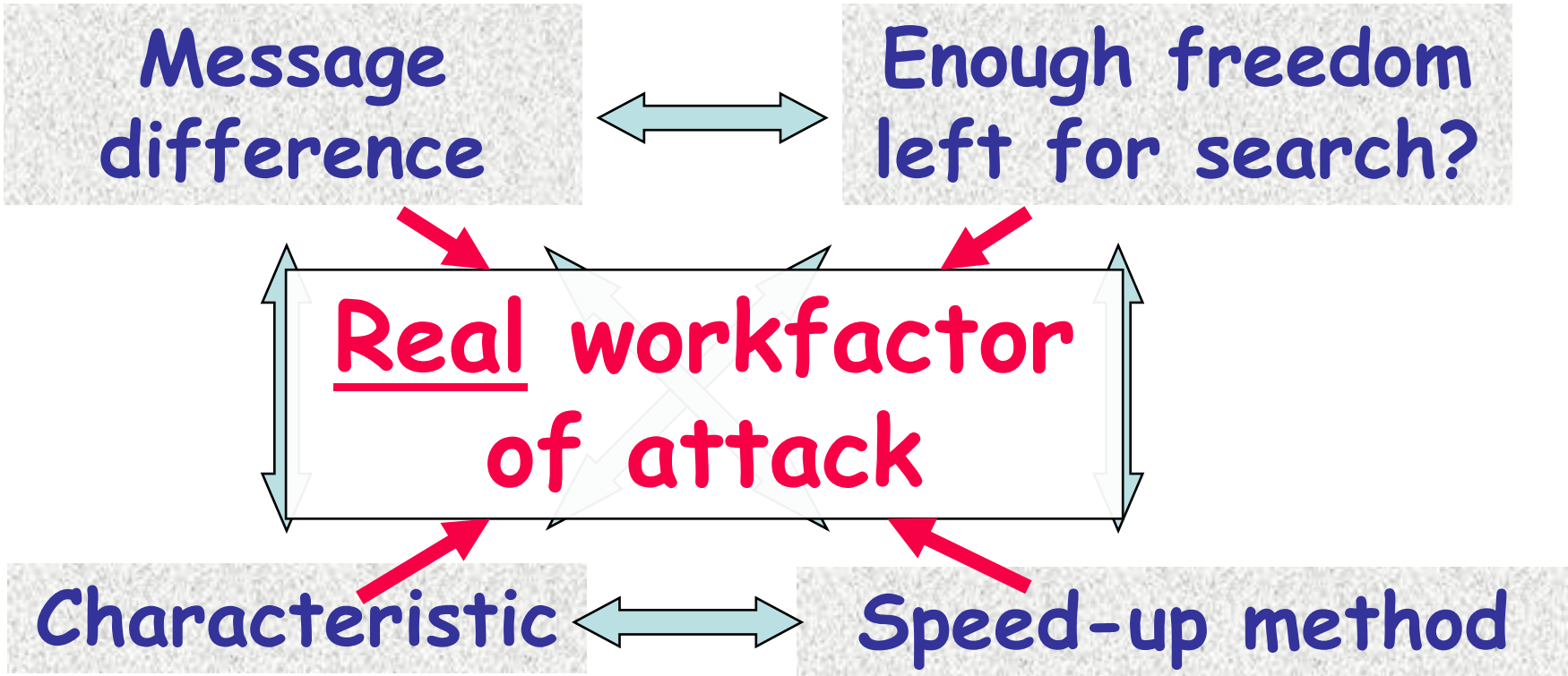
- Results for SHA/SHA-1
  - Up to 25 steps with manual optimization of characteristic (Wang et al., unpublished)
  - Up to 31 steps with especially restricted computer generated characteristic (ongoing work)

- Boundary: we need to leave some degrees of freedom for last phase

Problem of optimization



Problem of optimization



Some results

- SHA
  - 1998: 2^{61} [CJ98]
  - 2004: 2^{51} [BCJ+05]
  - 2005: 2^{39} [WYY05a]
  - 2006: 2^{36} [N+06]
- SHA-1
  - 2005: 58 steps 2^{33} [WYY05b]
  - 2006: 64 steps 2^{35} [DR06]
  - 2007: 70 steps 2^{44} [DMR07]

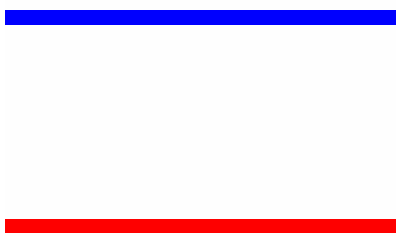
Meaningful Collisions - Motivation

- Setting: Collisions for a hash function can be constructed

- Cryptanalyst perspective: Some more interesting things to find out w.r.t. collision resistance?
  - Constructing collisions **faster**
  - Finding and exploiting degrees of freedom to construct (partially) **meaningful** collisions

  Practically relevant if hash function is widely deployed

Color Code



Under control, attacker can freely choose → meaningful

Not under direct control, determined by the collision search algorithm → not meaningful

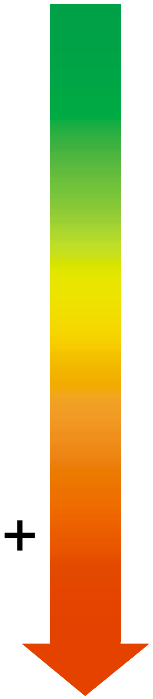
Meaningful Collisions: Challenges for MD4-style Hash Functions

I. One Commonly Chosen Prefix

II. One Commonly Chosen Prefix +
Partial Control over Colliding Blocks

III. Two Arbitrary Different Chosen Prefixes

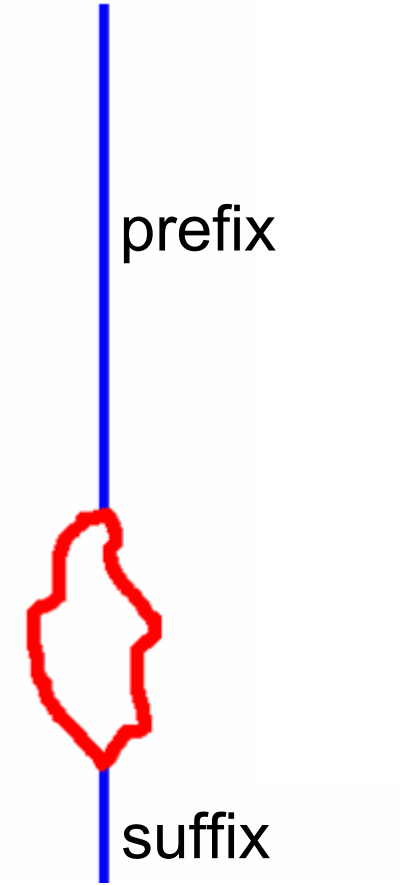
IV. Two Arbitrary Different Chosen Prefixes +
Partial Control over Colliding Blocks



"easier"

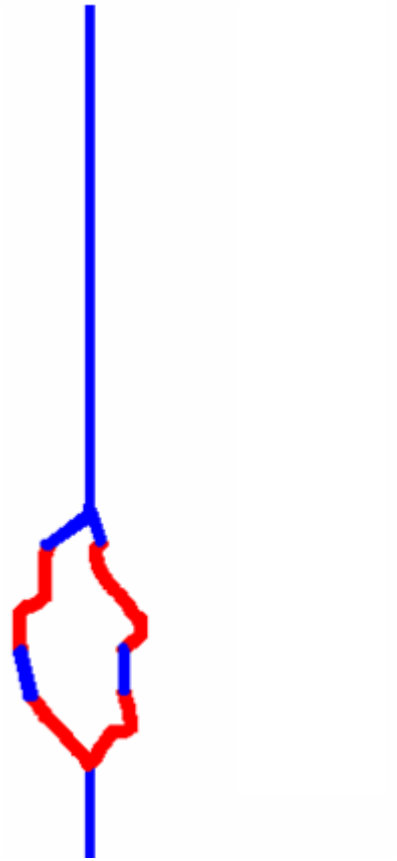
"harder"

I. One Commonly Chosen Prefix



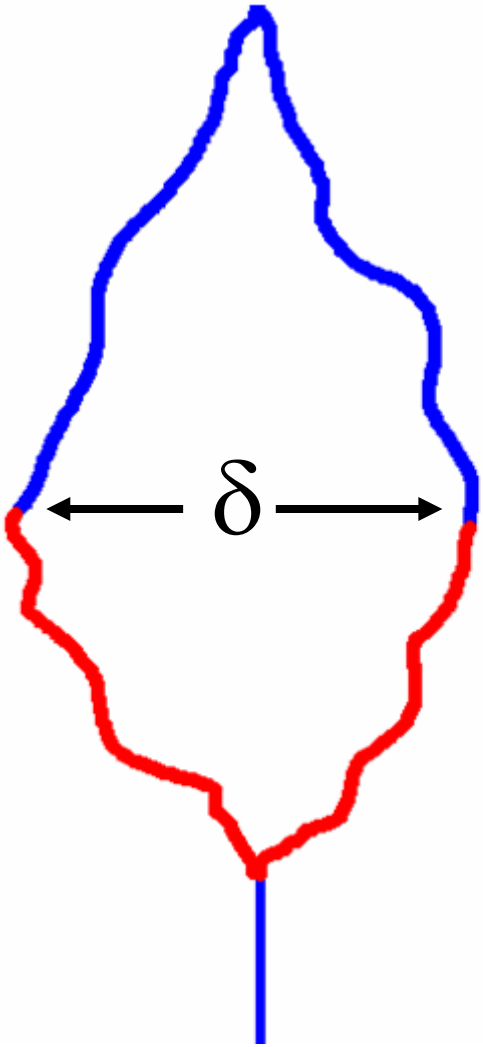
- Small number of colliding blocks
- Enough for colliding meaningful postscript files, etc… (see demo of Lars)

II. One Commonly Chosen Prefix + Partial Control over Colliding Blocks



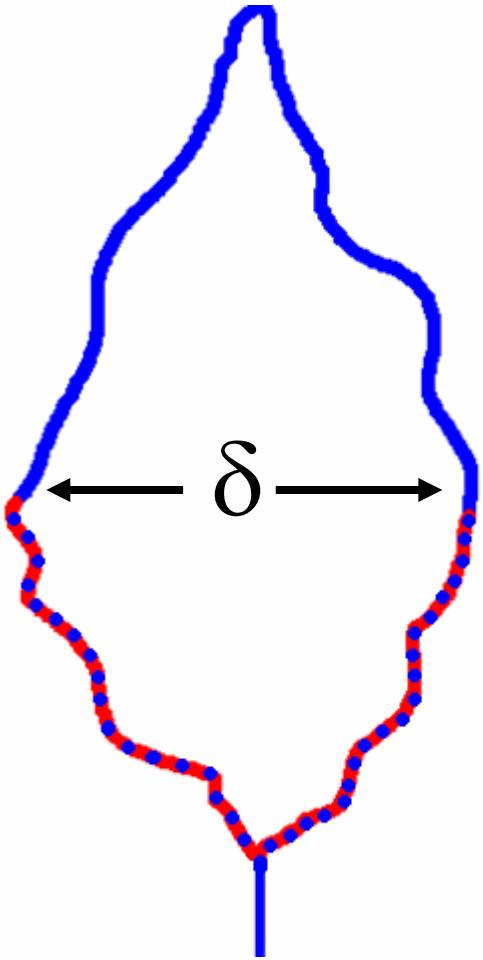
- Small number of colliding blocks
- Application in areas where format restrictions apply
- Example for 64-step SHA-1 in [DR06b]

III. Two Arbitrary Different Chosen Prefixes



- Using feed-forward operation, iteratively cancel out chaining differences with selected near-collision paths
- Usually much more than two message blocks needed
- Speedup: birthday phase before
- Example for MD5: see [SLdW07]

IV. Two Arbitrary Different Chosen Prefixes + Partial Control over Colliding Blocks



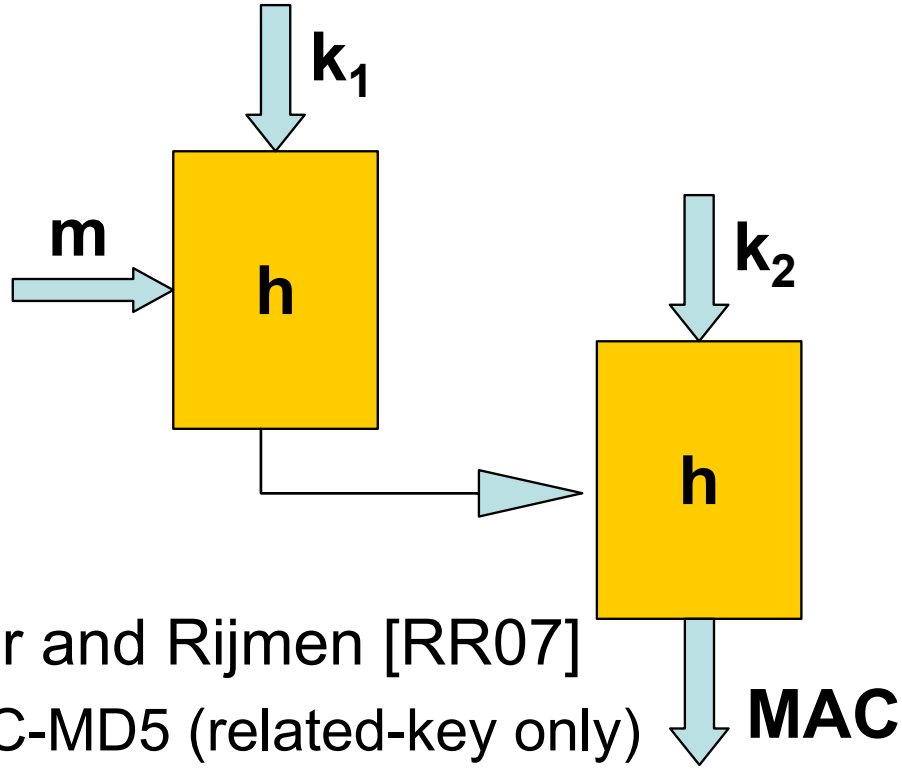
- Using feed-forward operation, iteratively cancel out chaining differences with selected near-collision paths
- Combination of methods

Even that is possible: Example Characteristic for type IV

i	∇A_i	∇W_i
-4:	0000uuu1nu001n1u100nn111u1nn00u1	
-3:	01000n0n110nunnun1nu00n1u1n1un00	
-2:	0uu0nn10uu1nunu10111n01uu1u11n1n	δ
-1:	1u1n1111110n11n110101n1u1n001001	
0:	01u0nu1unu0n01010n10001un0n0n00u	n0n00000000000000000000000000nn
1:	u001uuu1u011n1nn1001un0nu0u10n1u	10n0100010001100000011011 0un0011
2:	uuu0nu0uuu0nu1uuu11u0001n0u0001u	0uu11111010100----------------0
3:	101nn110n1un00nn1uu-0un1uu0-10-1	nnn11010101011-10------1--u1n1u1
4:	1011u01u00n11111n000u0-n0100011n	00u00101100110------100--n1001un
5:	00n111uu101111nn1u0u10u0-1n00010	x1un010010-110--011--0101u-11-10
6:	0nn01n0nn0-1uu--01n1-11u0--u0n0n	xu-n0-11------0-00-0-----x-u--uu
7:	n0-nu-0110n0--1101-0u10-00-011nu	xu1u010-------1-------1---x---u0
8:	00n10001n0u10u101--u0n01u1n--0u1	-1n0------0----1---------1-0----
9:	-11111n---100n-10----0n0u0001---	-nn-------0---------------uu--u-
10:	0--n1-1-0-010n-0--u--1-01u1n0---	-nu----------------------n----uu
11:	1---110-0--11n------1---0-01nu0-	--n--------------------1-n------
12:	--nu-1n-n-n--uun--n-1--nu-u1u010	xnu----------------------u----u-
13:	u---01--0-0-0--100--1-------0-10	-nn--------------------------n-
14:	x-0-11--1-1-1--011--1-----0-1-1-	x----------------------------u
15:	------------------------------	------------------------x-----

Attacks on NMAC/HMAC

Earlier results by
Kim et al. [KBPH06]
Contini and Yin [CY06]

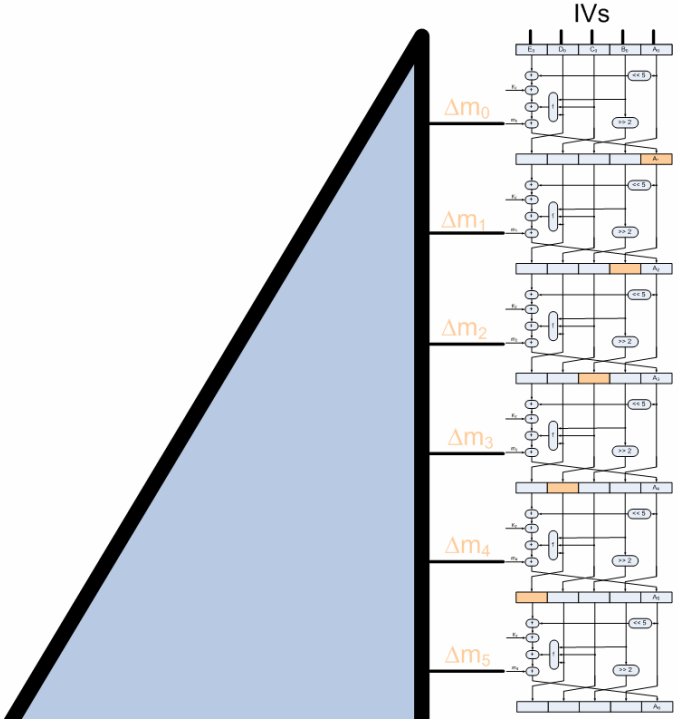


Recent Results by Rechberger and Rijmen [RR07]

- **Full** Key-Recovery for NMAC-MD5 (related-key only)
- **Full** Key-Recovery for NMAC-SHA-1 for 34/80 steps
- Distinguisher and Partial-Key recovery for NMAC/HMAC-SHA-1 for up to **62**/80 steps

What about SHA-2 members?

Probabilities of local collisions

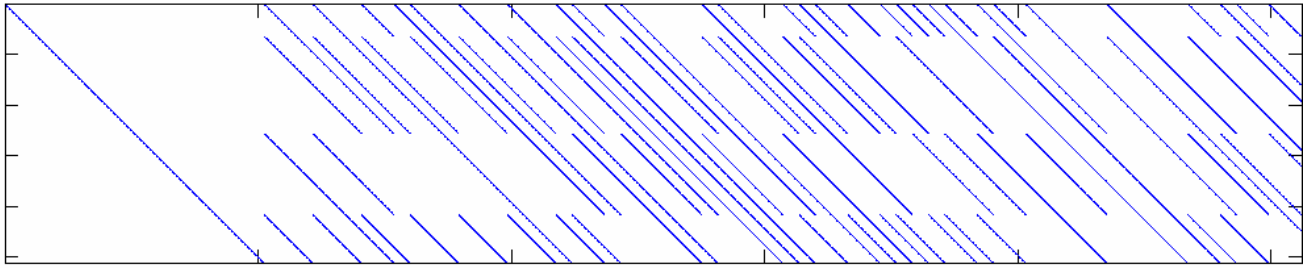


SHA/SHA-1: 2^{-2} to 2^{-5}

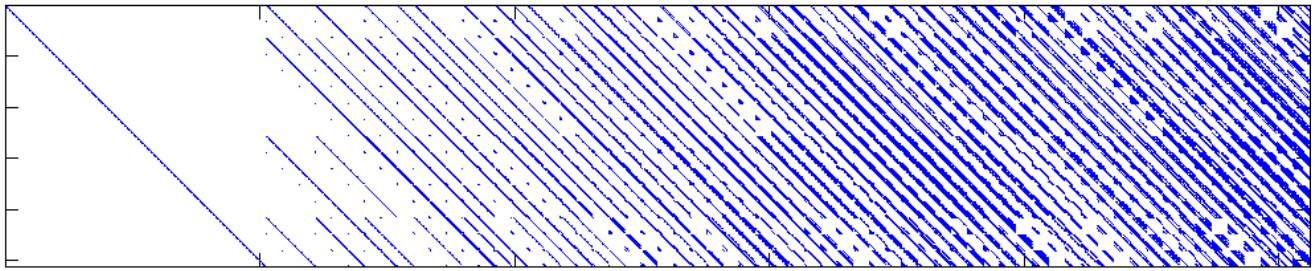
SHA-2: 2^{-38} to 2^{-41}

Comparing the message expansions of the SHA family

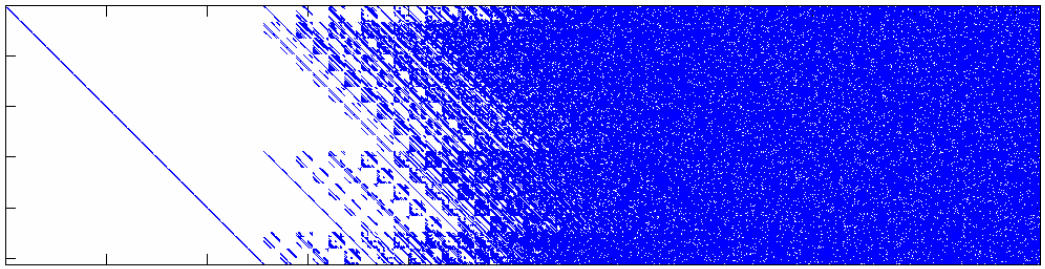
SHA



SHA-1



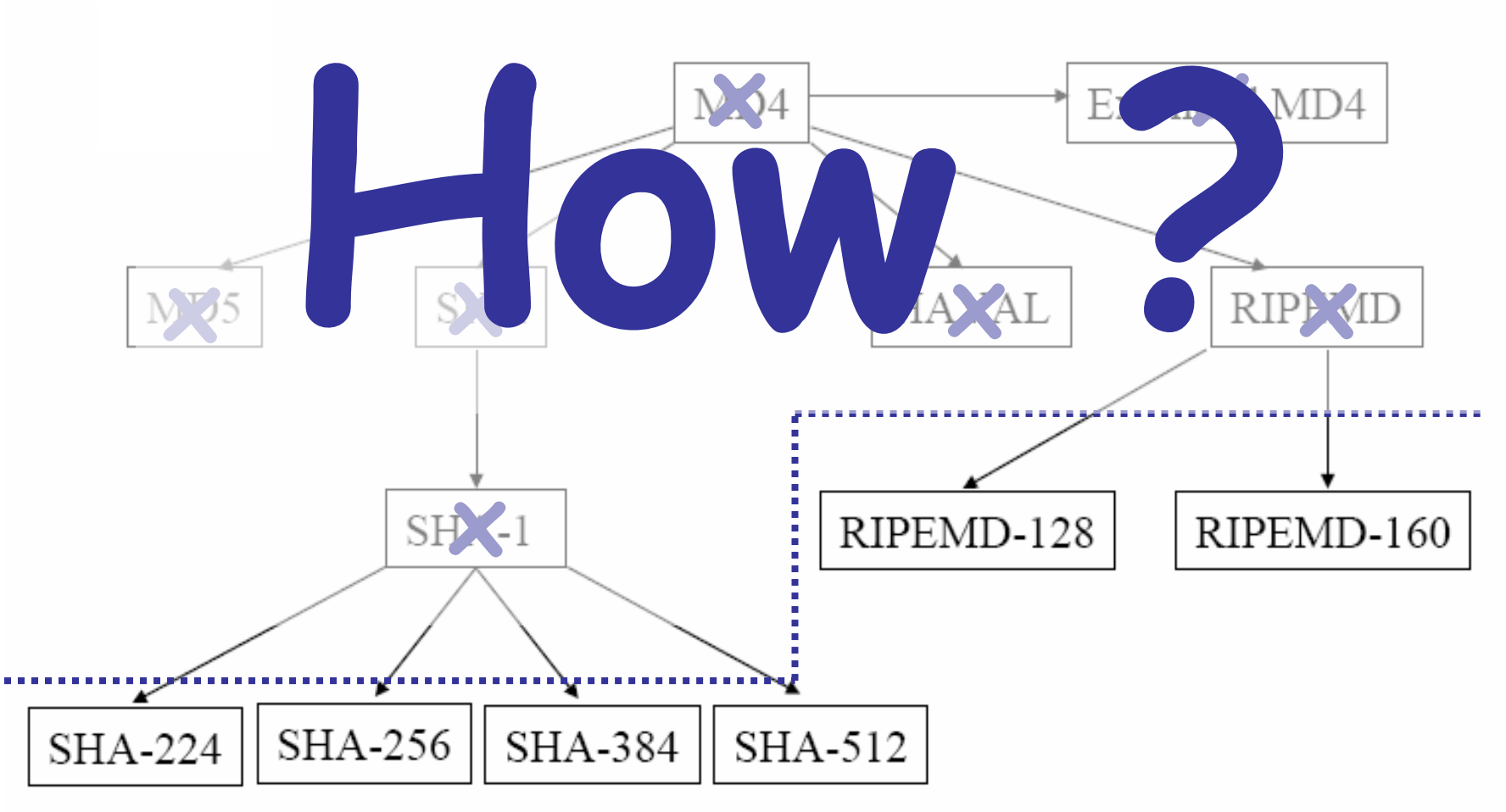
SHA-256
(linearized)



What about SHA-2 members?

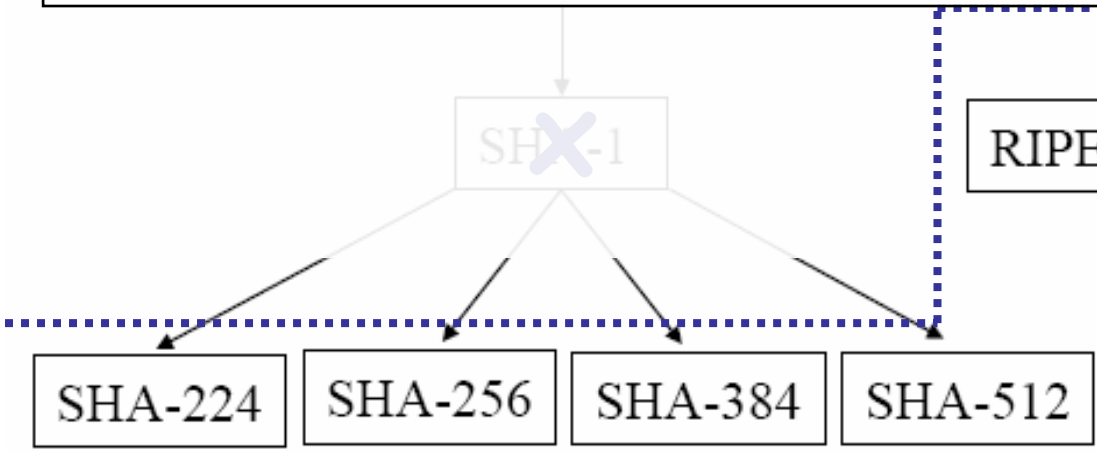
- Collision characteristic for 19-step SHA-224 [MPRR06]
- All publicly known attacks on SHA/SHA-1 since 1997 are not directly applicable to any SHA-2 member
- Message modification more difficult – cyclic dependencies
- New methods for characteristic search not applied yet

Conclusions



Conclusions

1) Special characteristics
2) Clever ways of solving equations (fast)



Conclusions

- Collisions for MD4, MD5, SHA
- Collision for full (80-step) SHA-1 is getting closer
- Optimization is ongoing
  - 2006: 2^{62} * x (unpublished work, estimates)
  - Advanced techniques as used for partial meaningful collisions can also be turned into **faster** collision search
  - 2007: ?

- Apply new insights to other hash functions like RIPEMD-160, SHA-2, new proposals?
- Improved Attacks on NMAC/HMAC?
- Results on (2nd-)preimage resistance?

Some References for MD4/MD5

[BB91] Den Boer, Booselaers, “An Attack on the Last Two Rounds of MD4”, CRYPTO 1991

[Dob98] Dobbertin: “Cryptanalysis Of MD4”, JoC, 1998

[Mer90] Merkle: “Note on MD4”, unpublished, 1990

[SLdW07] Stevens, Lenstra, De Weger, “Chosen-prefix Collisions for MD5 and Colliding X.509 Certificates for Different Identities”, EUROCRYPT 2007

[Vod94] Vaudenay, “On the Need for Multipermutations: Cryptanalysis of MD4 and SAFER”, FSE 1994

[WY05] Wang, Yu, “How to Break MD5 and Other Hash Functions”, EUROCRYPT 2005

Some References for SHA/SHA-1/SHA-256

[BC04] Biham, Chen: “Near-Collisions for SHA-0”, CRYPTO 2004

[BCJ+05] Biham, Chen, Joux et al.: “Collisions of SHA-0 and Reduced SHA-1”, EUROCRYPT 2005

[CJ98] Chabaud, Joux: “Differential Collisions in SHA-0”, CRYPTO 1998

[DR06] De Cannière, Rechberger: “Finding SHA-1 Characteristics: General Results and Applications”, ASIACRYPT 2006

[DR06b] De Cannière, Rechberger: “Meaningful Collisions for SHA-1 at no extra cost?”, CRYPTO 2006 Rump Session

[MPRR06] Mendel, Pramstaller, Rechberger, Rijmen: “Analysis of Step-Reduced SHA-256”, FSE 2006

[RO05] Rijmen, Oswald: “Update on SHA-1”, CT-RSA 2005

[WYY05a] Wang, Yu, Yin: “Efficient Collision Search Attacks on SHA-0”, CRYPTO 2005

[WYY05b] Wang, Yin, Yu: “Finding Collisions in the Full SHA-1”, CRYPTO 2005

Dedicated Attacks on NMAC/HMAC

[KBPH06] Kim, Biryukov, Preneel, Hong: “On the Security of HMAC and NMAC Based on HAVAL, MD4, MD5, SHA-0 and SHA-1”, SCN 2006

[CY06] Contini, Yin: “Forgery and Partial Key-Recovery Attacks on HMAC and NMAC Using Hash Collisions”, ASIACRYPT 2006

[RR07] Rechberger, Rijmen: “On Authentication with HMAC and Non-Random Properties”, Financial Cryptography 2007

Credits

Some slides are by curtsey of

Christophe De Cannière
Florian Mendel
and Vincent Rijmen

Not enough?

Want more?

ECRYPT

Hash Function Workshop

Barcelona, May 24-25, 2007

events.iaik.tugraz.at/hashworkshop07

Dedicated Attacks on Popular Hash Functions

Q&A

Christian.Rechberger@iaik.tugraz.at
www.iaik.tugraz.at/aboutus/people/rechberger

ECRYPT Summer School, May 02, 2007

Institute for Applied Information Processing and Communications (IAIK) - Krypto Group

Faculty of Computer Science
Graz University of Technology