Lecture 6 – Hash functions

TEK4500

29.09.2020

Håkon Jacobsen

hakon.jacobsen@its.uio.no

Hash function applications

- File storage verification
- Data authentication (MACs and digital signatures)
- Certificates
- Randomness extraction and key derivation (PRFs)
- Password hashing
- Quantum-resistant signatures
- Hash chains (Bitcoin)

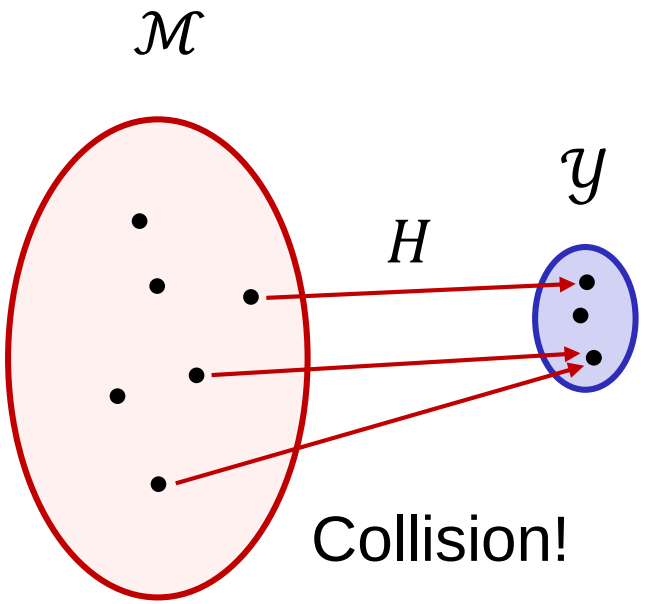
Hash functions

$H : \mathcal{M} \to \mathcal{Y}$

Keyless function

$|\mathcal{M}| \gg |\mathcal{Y}|$

Compressing



Examples:

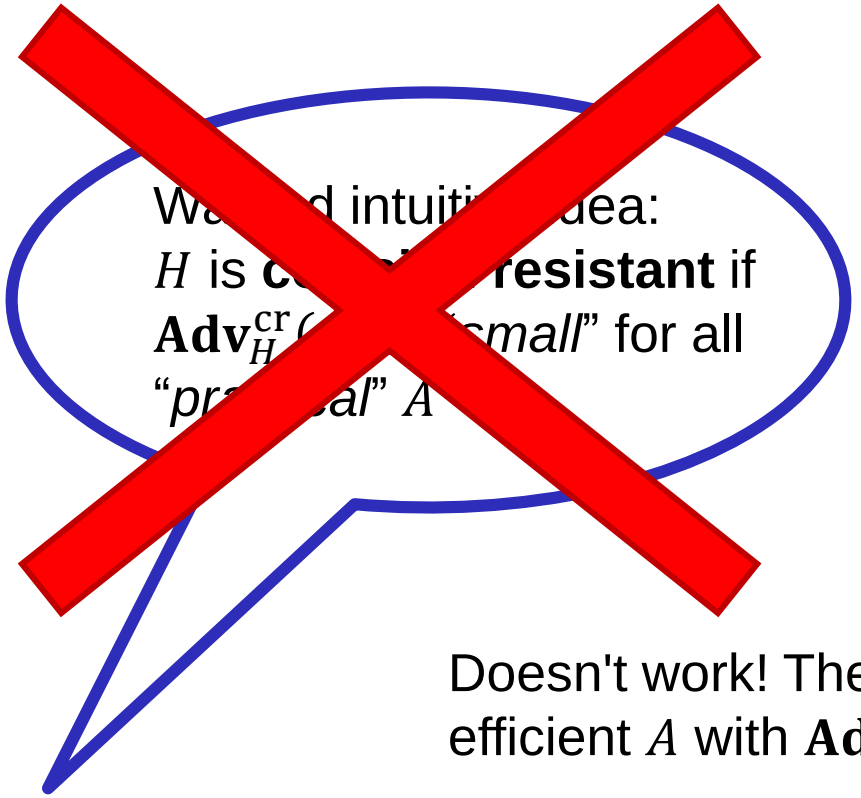
- $\mathrm{MD5} : \{0,1\}^{<2^{64}} \to \{0,1\}^{128}$
- $\mathrm{SHA1} : \{0,1\}^{<2^{64}} \to \{0,1\}^{160}$
- $\mathrm{SHA2-256} : \{0,1\}^{<2^{64}} \to \{0,1\}^{256}$
- $\mathrm{SHA2-512} : \{0,1\}^{<2^{128}} \to \{0,1\}^{512}$

$\{0,1\}^{<2^{64}} \approx \{0,1\}^{*}$

Collision resistance

$\mathbf{Exp}_H^{\mathrm{cr}}(A)$

1. $(X_1, X_2) \leftarrow A_H$
2. **if** $X_1 \neq X_2$ **and** $H(X_1) = H(X_2)$ **then**
3. **return** 1
4. **else**
5. **return** 0



Doesn't work! There always *exist* a very efficient A with $\mathbf{Adv}_H^{\mathrm{cr}}(A) = 1$!

Definition: The **CR-advantage** of an adversary A against H is

$$\mathbf{Adv}_H^{\mathrm{cr}}(A) = \Pr[\mathbf{Exp}_H^{\mathrm{cr}}(A) \Rightarrow 1]$$

Collision resistance

$\mathbf{Exp}_H^{\mathrm{cr}}(A)$

1. $(X_1, X_2) \leftarrow A_H$
2. **if** $X_1 \neq X_2$ **and** $H(X_1) = H(X_2)$ **then**
3. **return** 1
4. **else**
5. **return** 0

A

1. Output (X_1, X_2) where X_1, X_2 is a collision for H

X_1, X_2 must *exist* since $|\mathcal{M}| \gg |\mathcal{Y}|$

hence $\mathbf{Adv}_H^{\mathrm{cr}}(A) = 1$

…but how do we actually find X_1, X_2?!

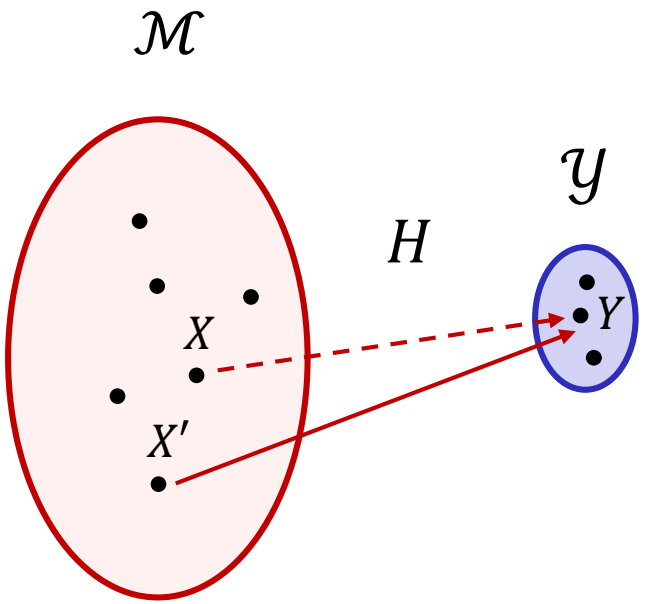
Definition: The **CR-advantage** of an adversary A against H is

$$\mathbf{Adv}_H^{\mathrm{cr}}(A) = \Pr[\mathbf{Exp}_H^{\mathrm{cr}}(A) \Rightarrow 1]$$

One-way security

$\mathbf{Exp}_H^{\mathrm{ow}}(A)$

1. $X \xleftarrow{\$} \mathcal{M}$
2. $Y \leftarrow H(X)$
3. $X' \leftarrow A_H(Y)$
4. **return** $H(X') \stackrel{?}{=} Y$



Definition: The **OW-advantage** of an adversary A against H is

$$\mathbf{Adv}_H^{\mathrm{ow}}(A) = \Pr[\mathbf{Exp}_H^{\mathrm{cr}}(A) \Rightarrow 1]$$

Relation between notions

$\mathbf{Exp}_H^{\mathrm{cr}}(A)$

1. $(X_1, X_2) \leftarrow A_H$
2. **if** $X_1 \neq X_2$ **and** $H(X_1) = H(X_2)$ **then**
3. **return** 1
4. **else**
5. **return** 0

$\mathbf{Exp}_H^{\mathrm{ow}}(A)$

1. $X \xleftarrow{\$} \mathcal{M}$
2. $Y \leftarrow H(X)$
3. $X' \leftarrow A_H(Y)$
4. **return** $H(X') \stackrel{?}{=} Y$

Collision-resistance $\Longrightarrow$ One-wayness

Proof idea: suppose A_{ow} is an algorithm that breaks one-wayness

1. Pick $X \xleftarrow{\$} \mathcal{M}$ and give $Y \leftarrow H(X)$ to A_{ow}
2. A_{ow} outputs X'
3. output (X, X') as a collision ($H(X') = Y = H(X)$)

Problem: what if $X' = X$? Very unlikely assuming $|\mathcal{M}| \gg |\mathcal{Y}|$

Relation between notions

$\mathbf{Exp}_H^{\mathrm{cr}}(A)$

1. $(X_1, X_2) \leftarrow A_H$
2. **if** $X_1 \neq X_2$ **and** $H(X_1) = H(X_2)$ **then**
3. **return** 1
4. **else**
5. **return** 0

$\mathbf{Exp}_H^{\mathrm{ow}}(A)$

1. $X \xleftarrow{\$} \mathcal{M}$
2. $Y \leftarrow H(X)$
3. $X' \leftarrow A_H(Y)$
4. **return** $H(X') \stackrel{?}{=} Y$

Collision-resistance $\Longrightarrow$ One-wayness

Collision-resistance $\not\Longleftarrow$ One-wayness

Suppose $H : \mathcal{M} \to \{0,1\}^{256}$ is one-way. Define

$$H'(X) = \begin{cases} 0^{256} & \text{if } X = 0 \text{ or } X = 1 \\ H(X) & \text{otherwise} \end{cases}$$

H' is one-way

H' is **not** collision-resistant

Collision resistance application – MAC domain extension

$\mathrm{MAC} : \mathcal{K} \times \{0,1\}^n \to \mathcal{T}$ $\quad\quad$ $H : \{0,1\}^* \to \{0,1\}^n$

$\mathrm{MAC}' : \mathcal{K} \times \{0,1\}^* \to \mathcal{T}$

$\mathrm{MAC}'(K, M) = \mathrm{MAC}\big(K, H(M)\big)$ ← Hash-then-MAC paradigm

Theorem: If H is collision-resistant and MAC is UF-CMA secure, then MAC' is UF-CMA secure

Example:

$\mathrm{SHA2-256} : \{0,1\}^* \to \{0,1\}^{256}$
$\mathrm{AES-CBC-MAC} : \mathcal{K} \times \{0,1\}^{2\times 128} \to \{0,1\}^{128}$

$\mathrm{MAC}'(K, M) = \mathrm{AES-CBC-MAC}\big(K, \mathrm{SHA2-256}(M)\big)$

Collision-resistance is *necessary*:

Suppose you can find collision (X_1, X_2) for H

1. Ask for MAC tag on X_1 (receive back T)
2. Output (X_2, T) as forgery

Attacks on hash functions

- Specific attacks; exploit internal design of hash function

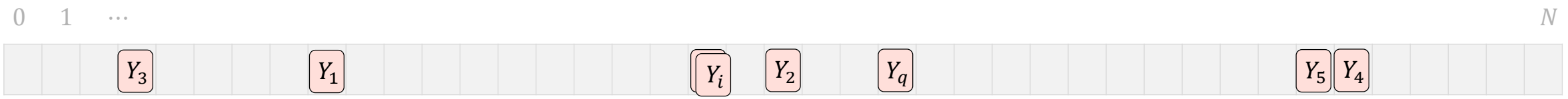
- Generic attacks; work for all hash functions $H : \{0,1\}^* \rightarrow \{0,1\}^n$

  - Brute-force: hash $H(0), H(1), H(2), \ldots, H(2^n)$
    - Success guaranteed ($\mathbf{Adv}_H^{\mathrm{cr}}(A) = 1$)
    - Output must be long enough; e.g., $n = 10$ requires only $2^{10} + 1 = 1025$ values
    - $n = 100$ enough?

  - Attacker B:
    1. pick $q \ll 2^n$ distinct values $X_1, X_2, \ldots, X_q$ at *random*
    2. hash every value
    3. look for collisions

    What's $\mathbf{Adv}_H^{\mathrm{cr}}(B)$?

Birthday attack



$Y_1 \leftarrow H(X_1)$

$Y_2 \leftarrow H(X_2)$

$Y_3 \leftarrow H(X_3)$

$Y_4 \leftarrow H(X_4)$

$Y_5 \leftarrow H(X_5)$

$Y_6 \leftarrow H(X_6)$

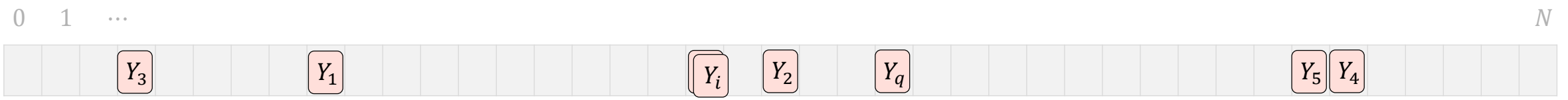
$Y_i \leftarrow H(X_i)$

$Y_q \leftarrow H(X_q)$

$$\mathbf{Adv}^{\mathbf{cr}}_{H}(B) = \Pr[\text{coll}] = \Pr\left[\exists i \neq j : H(X_i) = H(X_j)\right]$$

$X_1, X_2, \ldots, X_q$ are *distinct*

Birthday attack



$Y_1 \leftarrow\$\ \mathcal{Y}$
$Y_2 \leftarrow\$\ \mathcal{Y}$
$Y_3 \leftarrow\$\ \mathcal{Y}$
$Y_4 \leftarrow\$\ \mathcal{Y}$
$Y_5 \leftarrow\$\ \mathcal{Y}$
$Y_6 \leftarrow\$\ \mathcal{Y}$

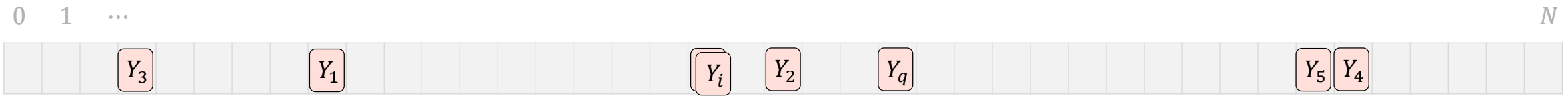
$Y_i \leftarrow\$\ \mathcal{Y}$

$Y_q \leftarrow\$\ \mathcal{Y}$

$$\mathbf{Adv}_{\rho}^{\mathrm{cr}}(B) = \Pr[\mathrm{coll}] = \Pr\left[\exists i \neq j : \rho(X_i) = \rho(X_j)\right] = \Pr\left[\exists i \neq j : Y_i = Y_j\right]$$

$X_1, X_2, \ldots, X_q$ are *distinct*

Birthday attack



$$\Pr[\text{coll}] = \Pr\left[\exists i \neq j : Y_i = Y_j\right]$$

$$= 1 - \Pr\left[\forall i \neq j : Y_i \neq Y_j\right]$$

$$= 1 - 1 \times \frac{N-1}{N} \times \frac{N-2}{N} \times \frac{N-3}{N} \quad \cdots \quad \times \frac{N-(q-1)}{N}$$

$$= 1 - 1 \times \left(1 - \frac{1}{N}\right) \times \left(1 - \frac{2}{N}\right) \times \left(1 - \frac{3}{N}\right) \cdots \times \left(1 - \frac{q-1}{N}\right)$$

$$= 1 - \prod_{i=1}^{q-1} \left(1 - \frac{i}{N}\right)$$

$$\geq 1 - \prod_{i=1}^{q-1} e^{-i/N} = 1 - e^{-(1+2+\cdots+q-1)/N} = 1 - e^{-q(q-1)/2N} \geq 0.3 \times \frac{q(q-1)}{2N}$$

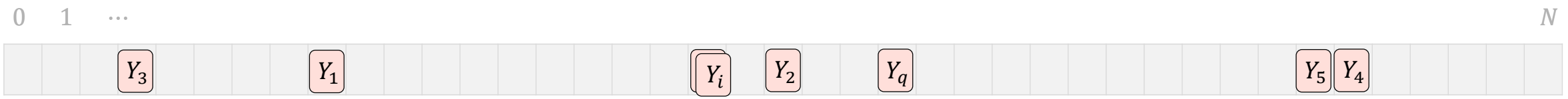
$$e^{-x} \geq 1 - x \qquad x \geq 0$$

$$x = i/N$$

$$e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$q \leq \sqrt{2N}$$

Birthday attack



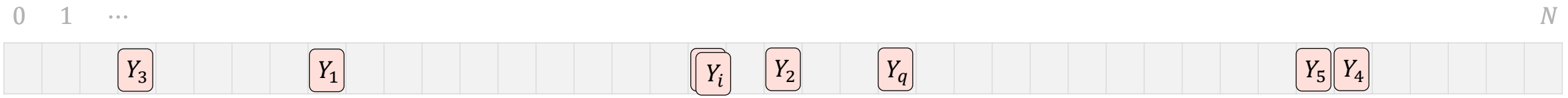
$$0.3 \times \frac{q(q-1)}{2N} \leq \Pr[\text{coll}] = \Pr[Y_1 \vee Y_2 \vee \cdots \vee Y_q]$$

$$\leq \Pr[Y_1] + \Pr[Y_2] + \cdots + \Pr[Y_q]$$

$$\leq \frac{0}{N} + \frac{1}{N} + \cdots + \frac{q-1}{N}$$

$$= \frac{q(q-1)}{2N}$$

Birthday attack



$$0.3 \times \frac{q(q-1)}{2N} \leq \Pr[\text{coll}] \leq \frac{q(q-1)}{2N}$$

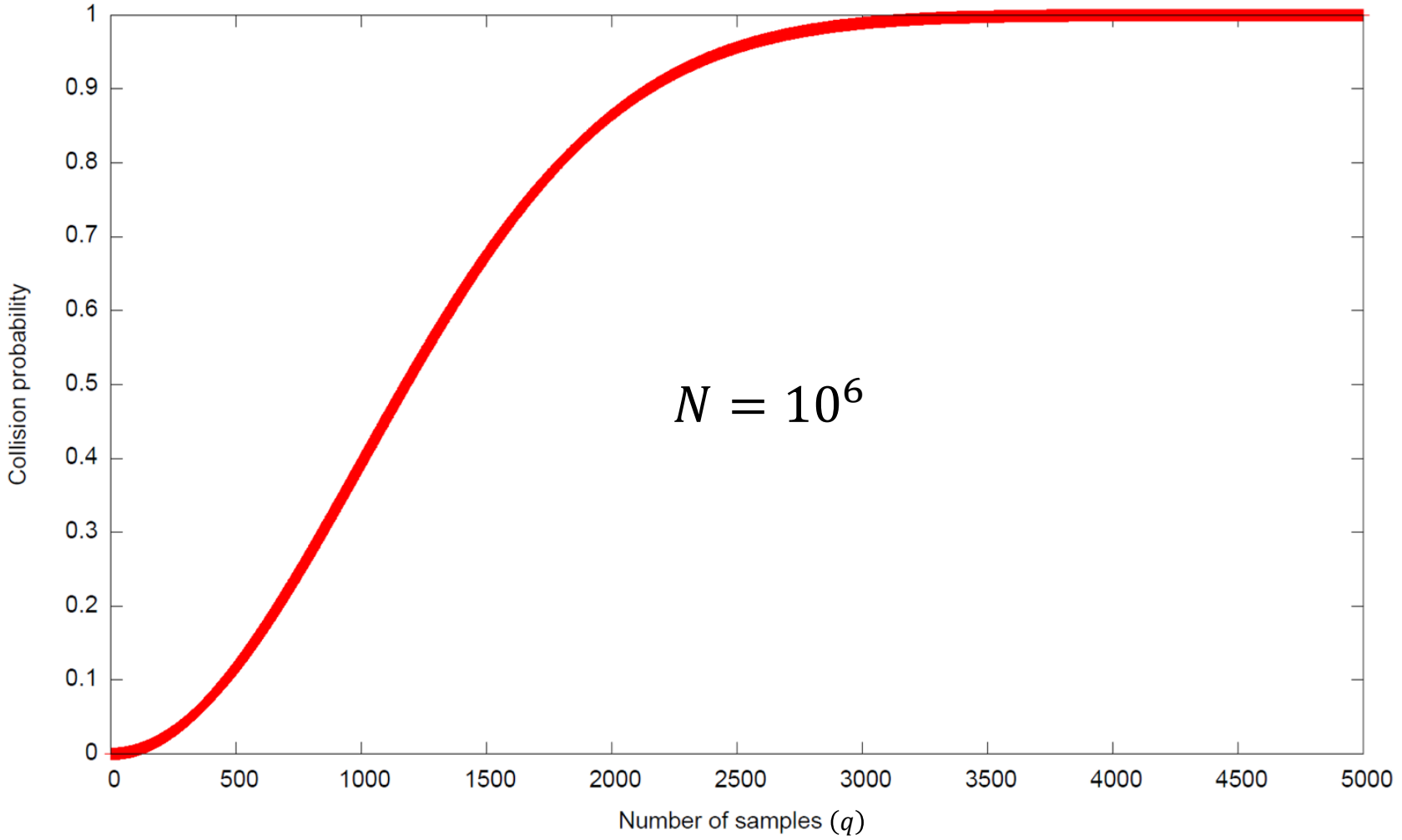
$\Pr[\text{coll}] = 0.5$:

$$q \approx \sqrt{N}$$

N	q
365	20
10^6	1000
2^{64}	2^{32}
2^{80}	2^{40}
2^{256}	2^{128}
2^{512}	2^{256}

Birthday bound: for $q \leq \sqrt{2N}$

$$\Pr[\text{coll}] = \Theta\left(\frac{q^2}{2N}\right) \approx \frac{q^2}{2N}$$


Collision probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
Number of samples (q)
$N = 10^6$

Attacks on hash functions

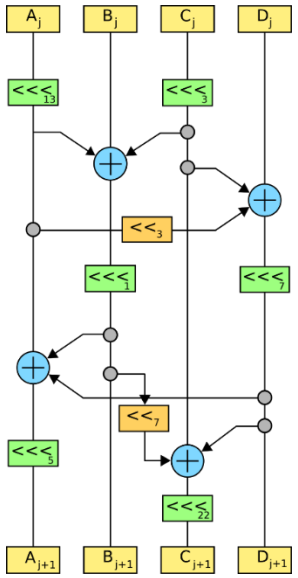
- Specific attacks; exploit internal design of hash function

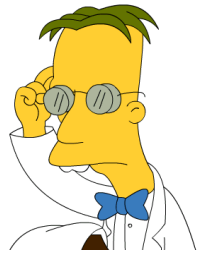
- Generic attacks; work for all hash functions $H : \{0,1\}^* \rightarrow \{0,1\}^n$

  - Brute-force: hash $H(0), H(1), H(2), \ldots, H(2^n)$
    - Success guaranteed ($\mathbf{Adv}_H^{\mathrm{cr}}(A) = 1$)
    - Output must be long enough; e.g., $n = 10$ requires only $2^{10} + 1 = 1025$ values
    - $\boldsymbol{n = 100}$ **enough?** No! Must take birthday attack into account: n / 2 must be large enough

  - Attacker B:
    1. pick $q \ll 2^n$ distinct values $X_1, X_2, \ldots, X_q$ at *random*
    2. hash every value
    3. look for collisions

    What's $\mathbf{Adv}_H^{\mathrm{cr}}(B)$?





DESIGNING HASH FUNCTIONS

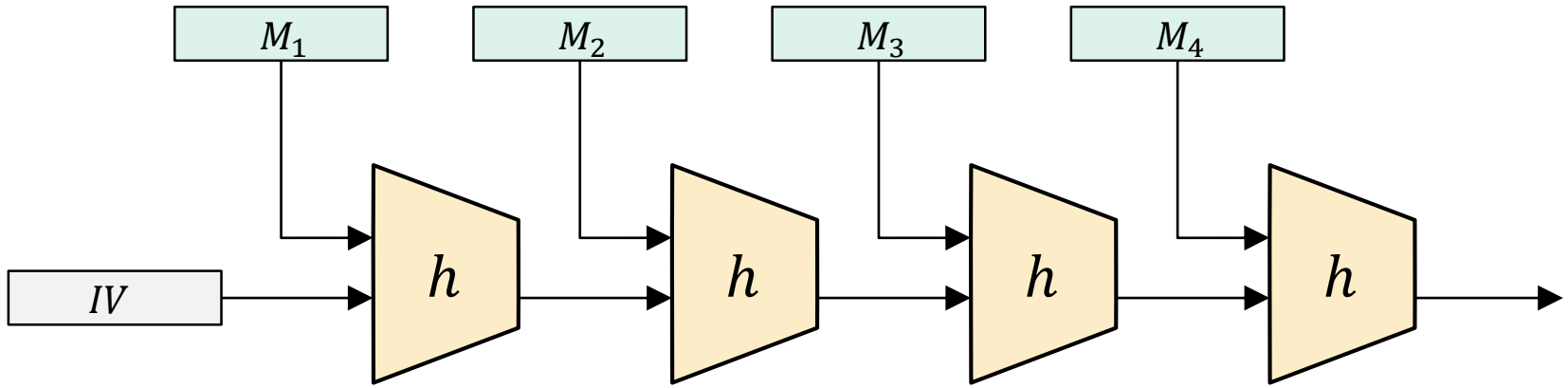
Merkle-Damgård

- Suppose we have a **compression function** $h : \{0,1\}^{n+b} \to \{0,1\}^n$ which is collision-resistant
- Want to create a hash function $H : \{0,1\}^* \to \{0,1\}^n$
- Solution: iterate h

Merkle-Damgård

$H : \{0,1\}^* \rightarrow \{0,1\}^n$

$$H(M) = h(h(\cdots h(h(IV||M_1) \,||\, M_2) \cdots \,||\, M_{n-1}) \,||\, M_n)$$

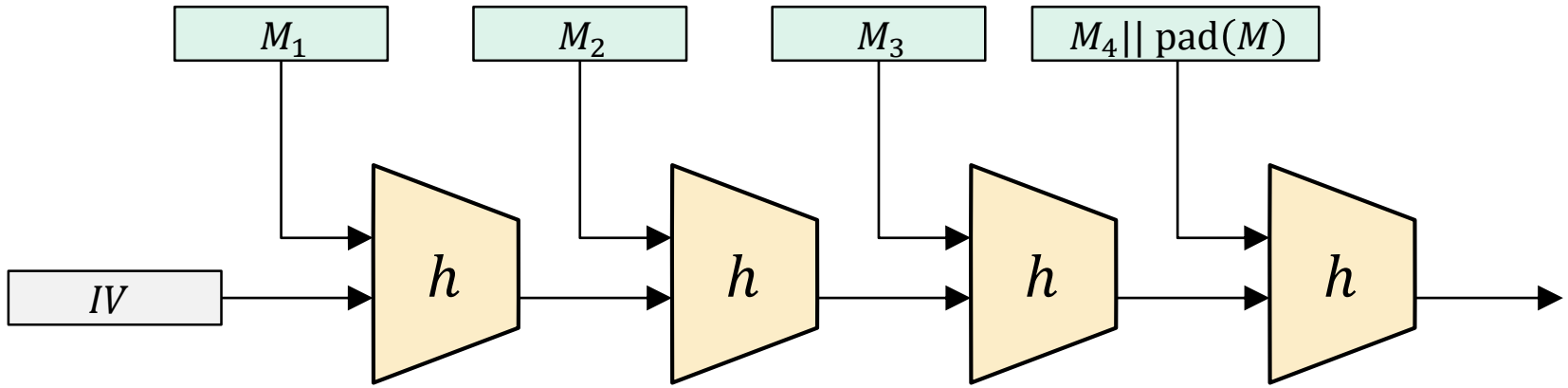


$h : \{0,1\}^{n+b} \rightarrow \{0,1\}^n$

Merkle-Damgård

$H : \{0,1\}^* \rightarrow \{0,1\}^n$

$$H(M) = h(h(\cdots h(h(IV||M_1) \,||\, M_2) \cdots \,||\, M_{n-1}) \,||\, M_n)$$

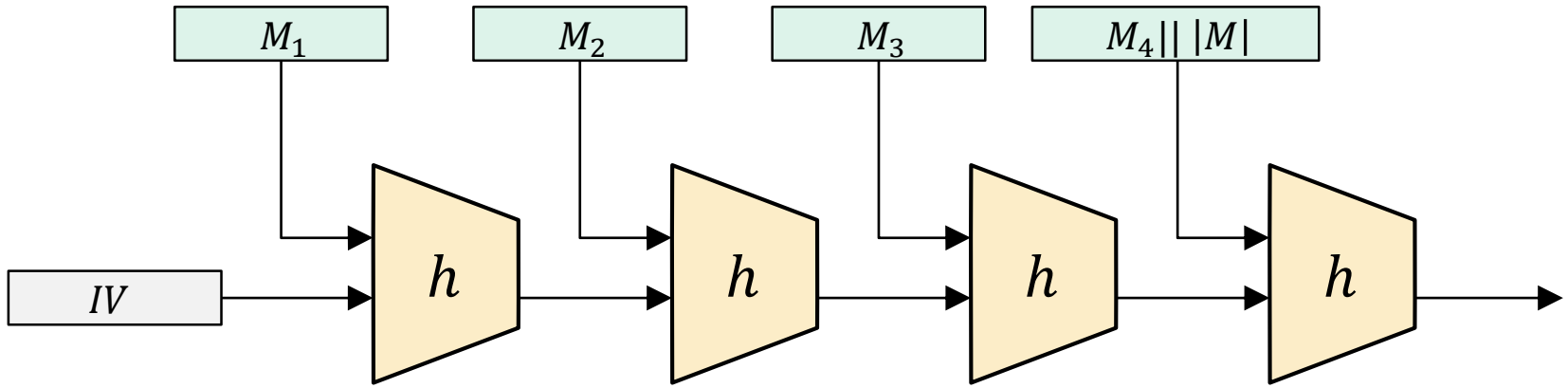


$h : \{0,1\}^{n+b} \rightarrow \{0,1\}^n$

Merkle-Damgård

$H : \{0,1\}^* \rightarrow \{0,1\}^n$

$$H(M) = h(h(\cdots h(h(IV||M_1) \,||\, M_2) \cdots \,||\, M_{n-1}) \,||\, M_n)$$



$h : \{0,1\}^{n+b} \rightarrow \{0,1\}^n$

Merkle-Damgård – security

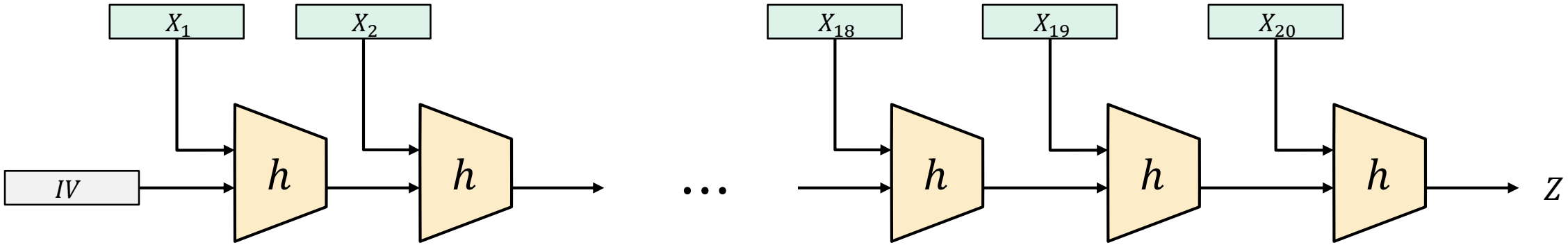
Theorem: If $h : \{0,1\}^{n+b} \to \{0,1\}^n$ is collistion resistant then $H : \{0,1\}^* \to \{0,1\}^n$ is collisition resistant

Proof idea:

Collison on $H \implies$ collision on h

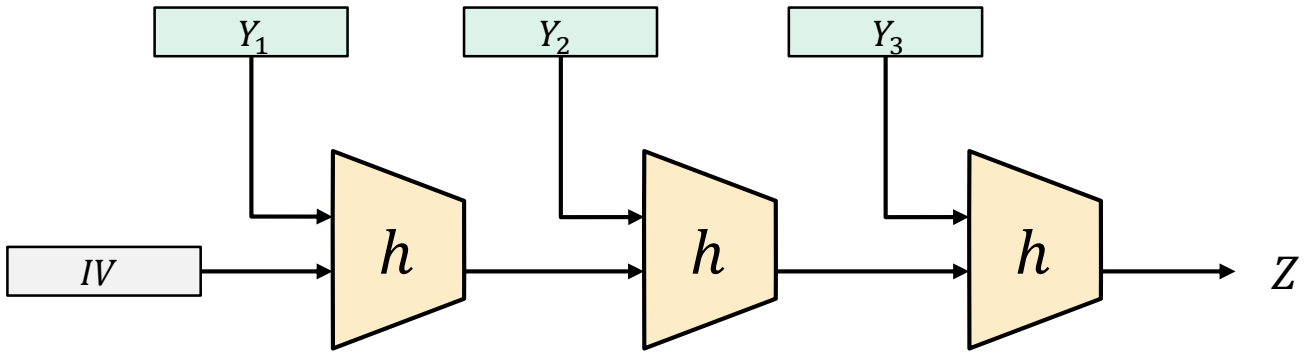
If $H(X) = H(Y)$ then we can construct X', Y' such that $h(X') = h(Y')$

Merkle-Damgård – security

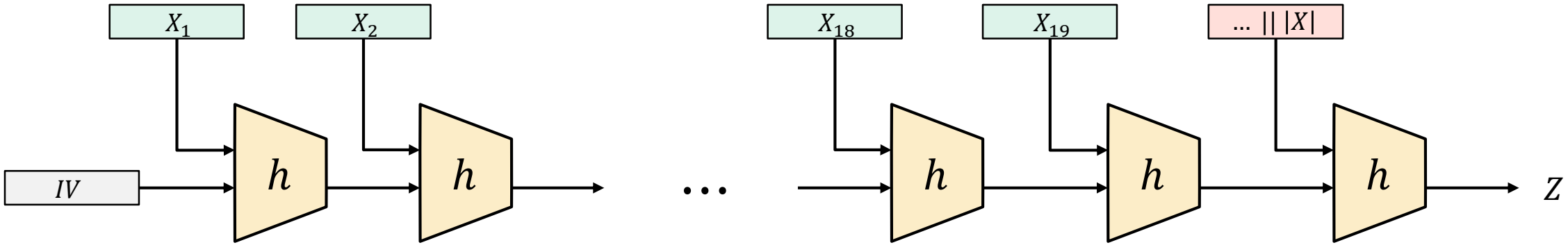


Assumption:

$$H(X) = H(Y) = Z$$

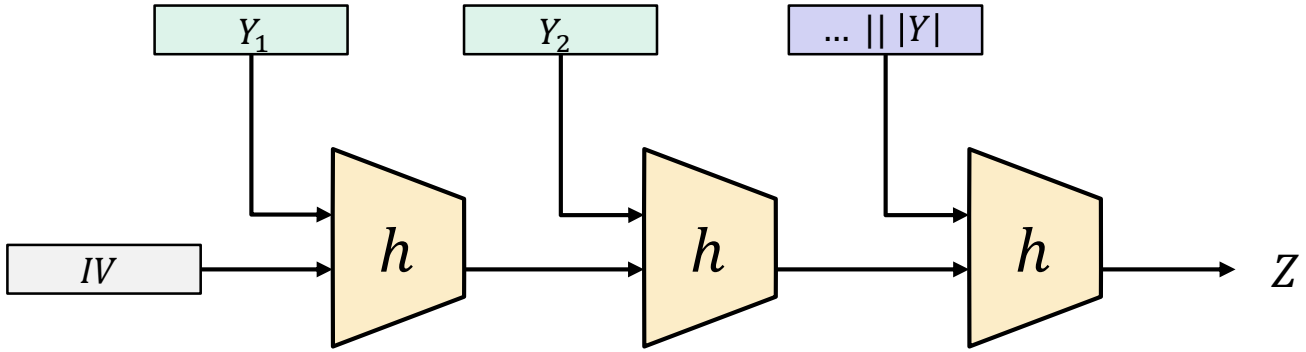


Merkle-Damgård – security

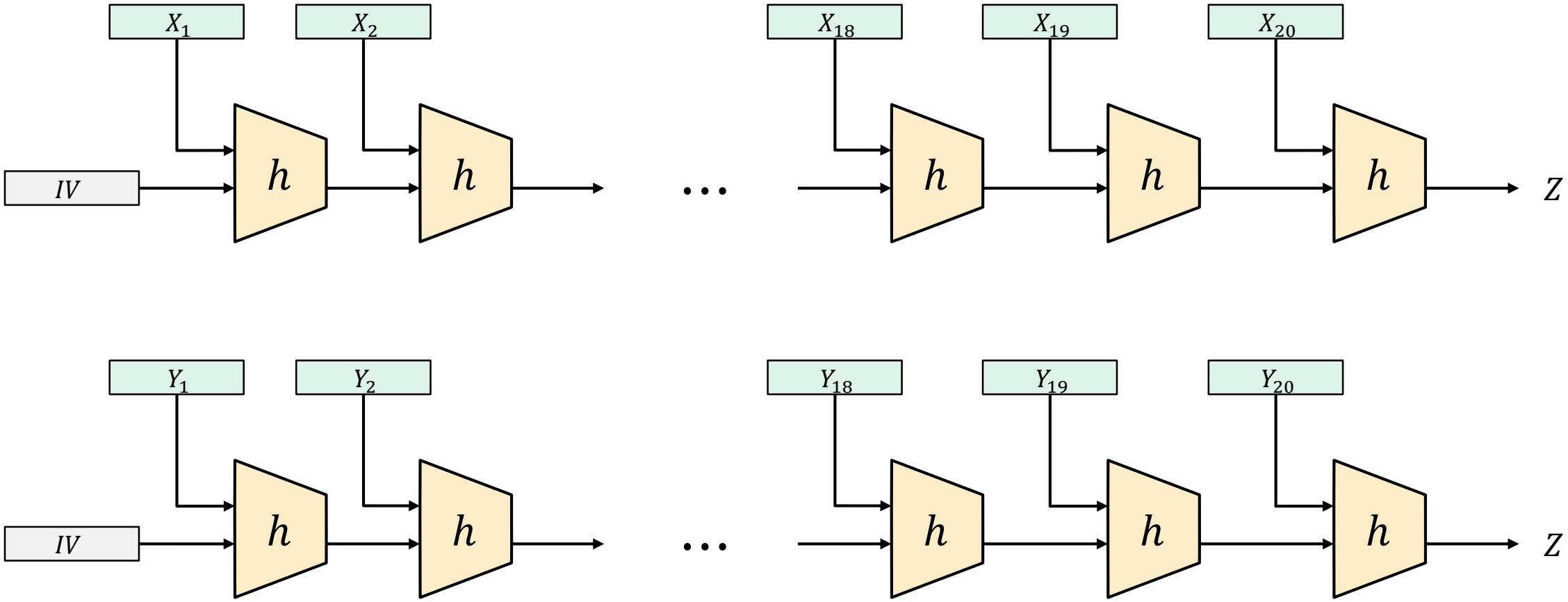


Assumption:

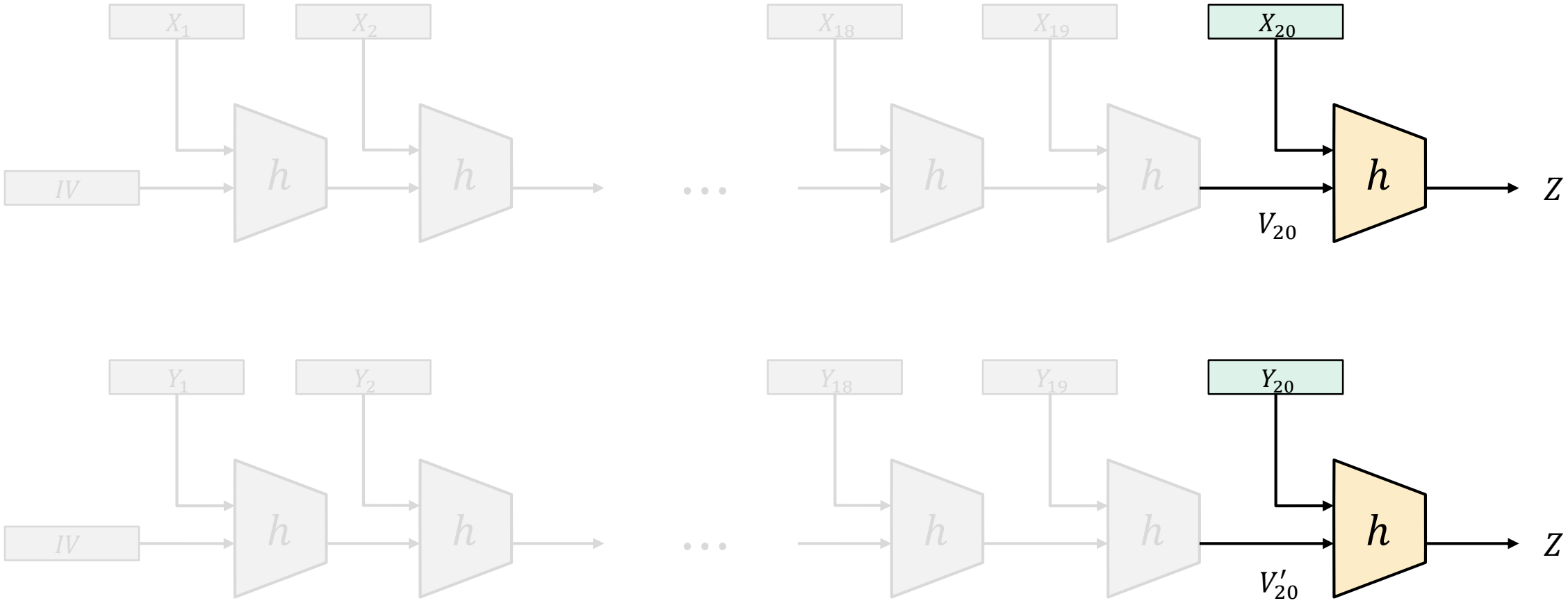
$$H(X) = H(Y) = Z$$



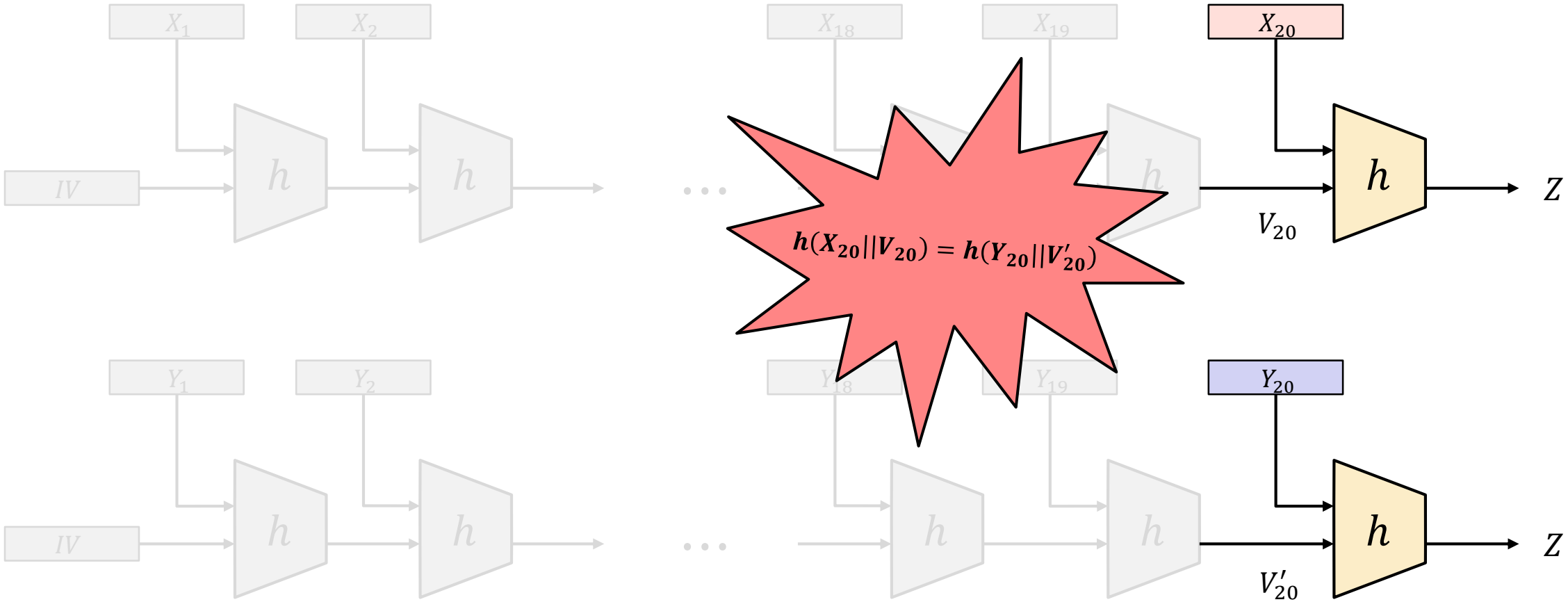
Merkle-Damgård – security



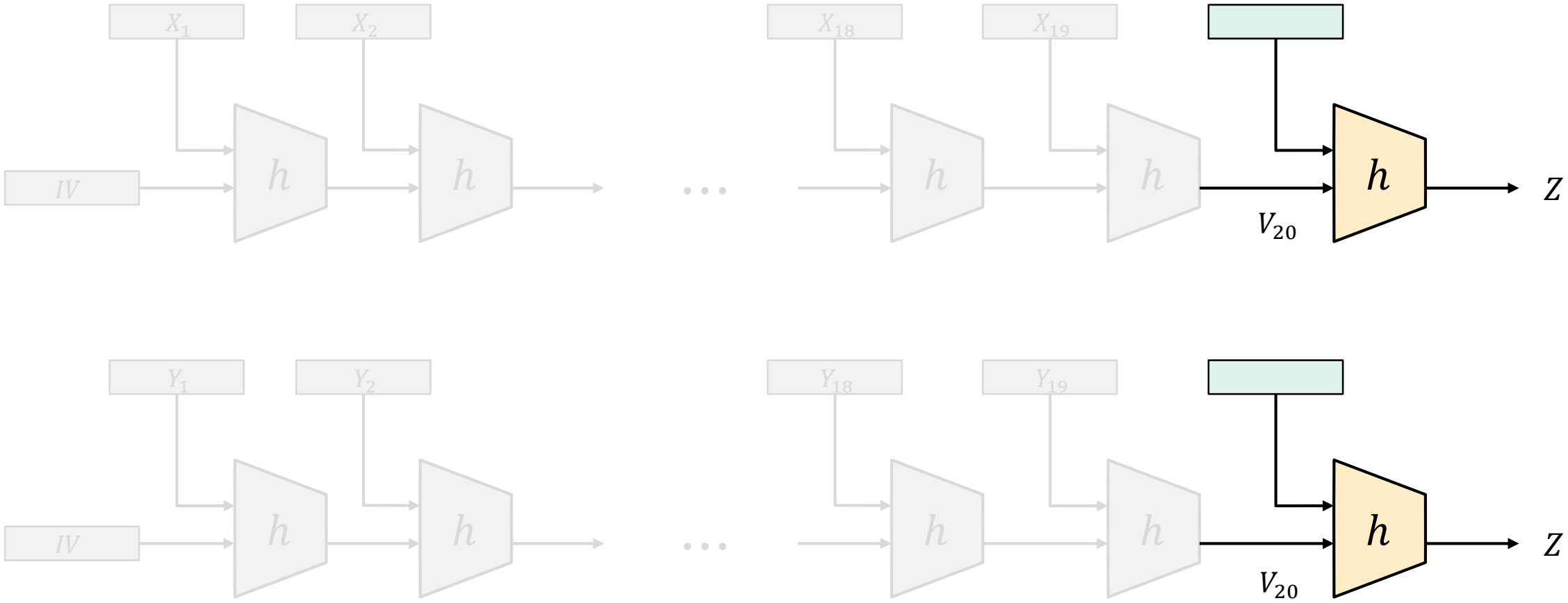
Merkle-Damgård – security



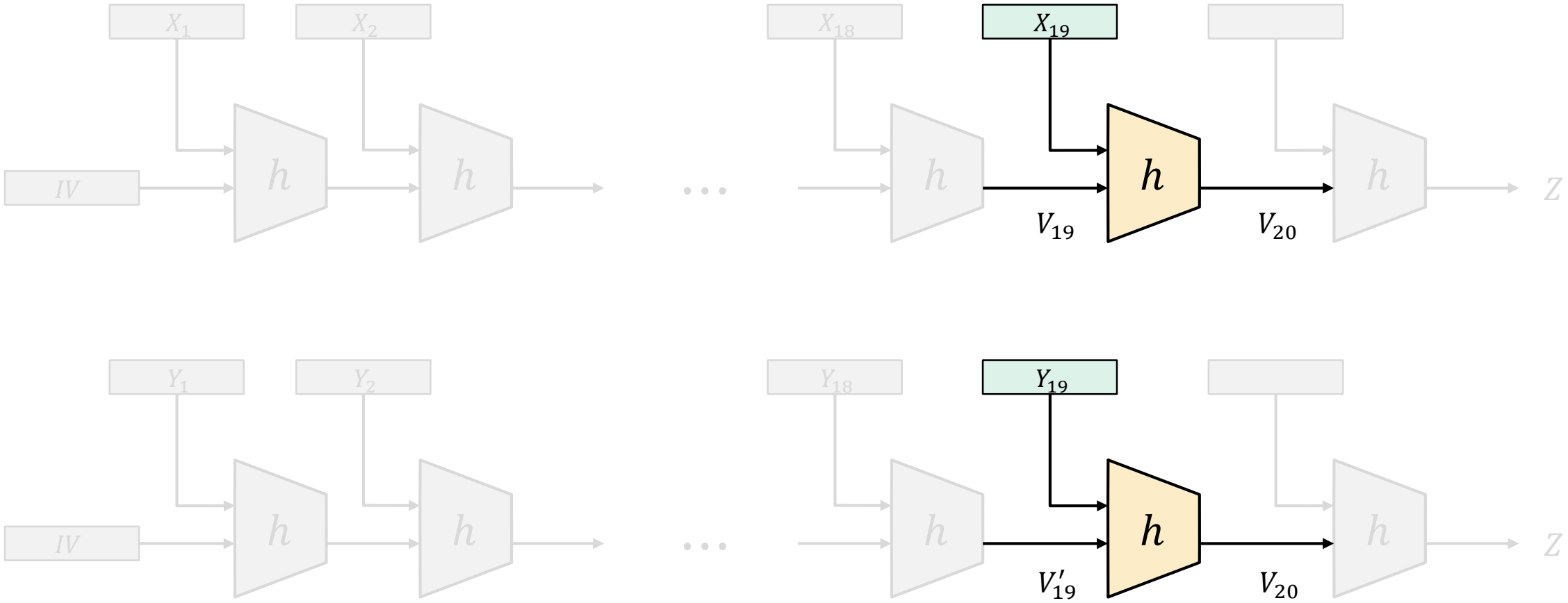
Merkle-Damgård – security



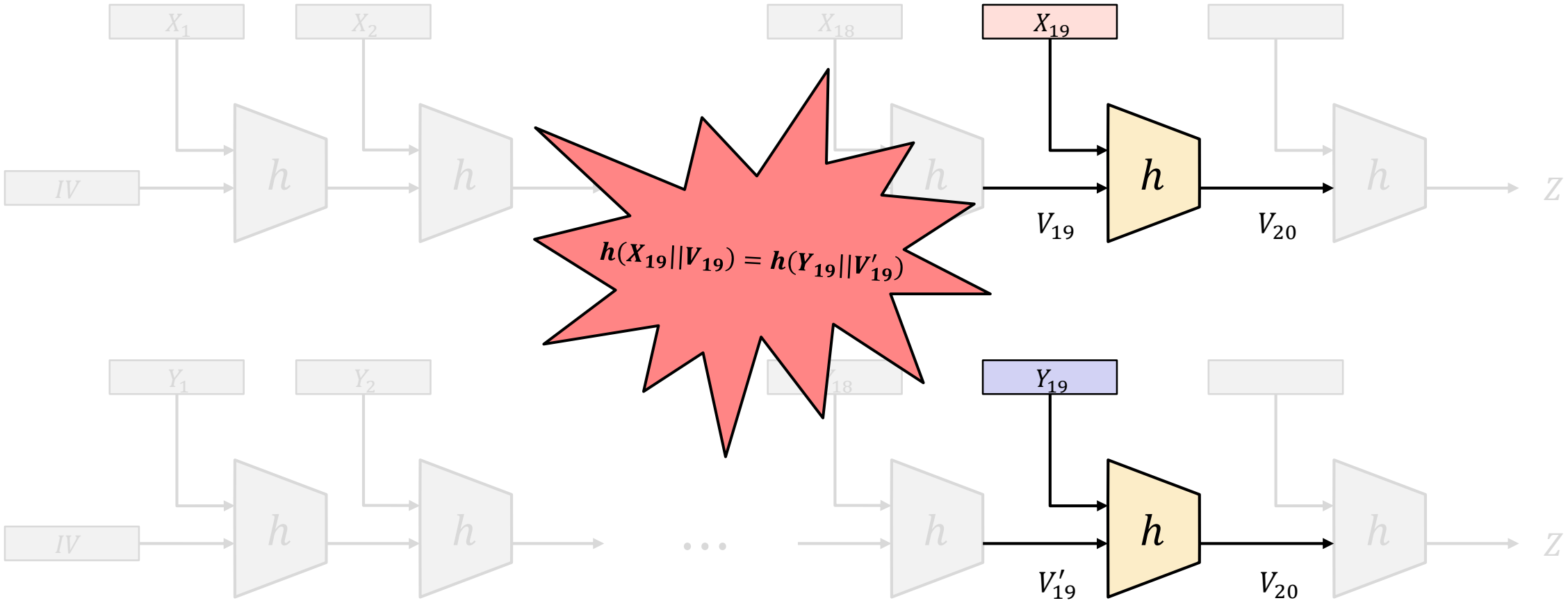
Merkle-Damgård – security



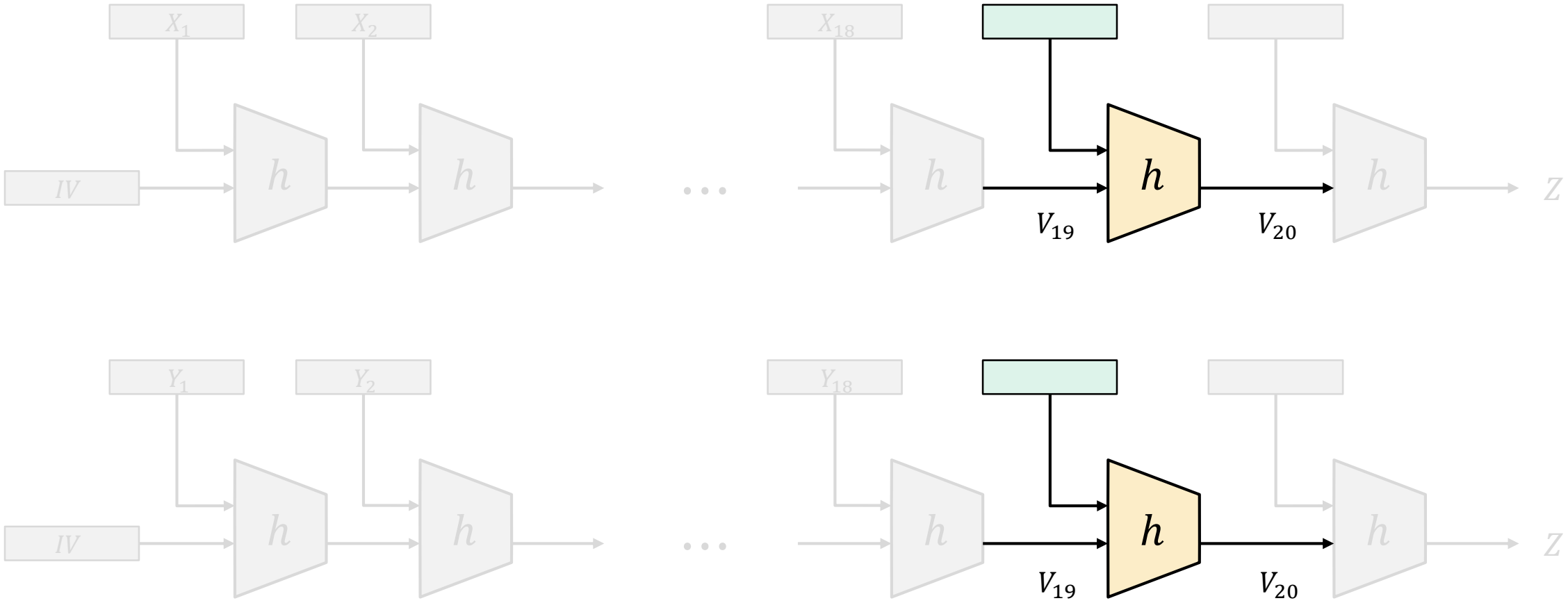
Merkle-Damgård – security



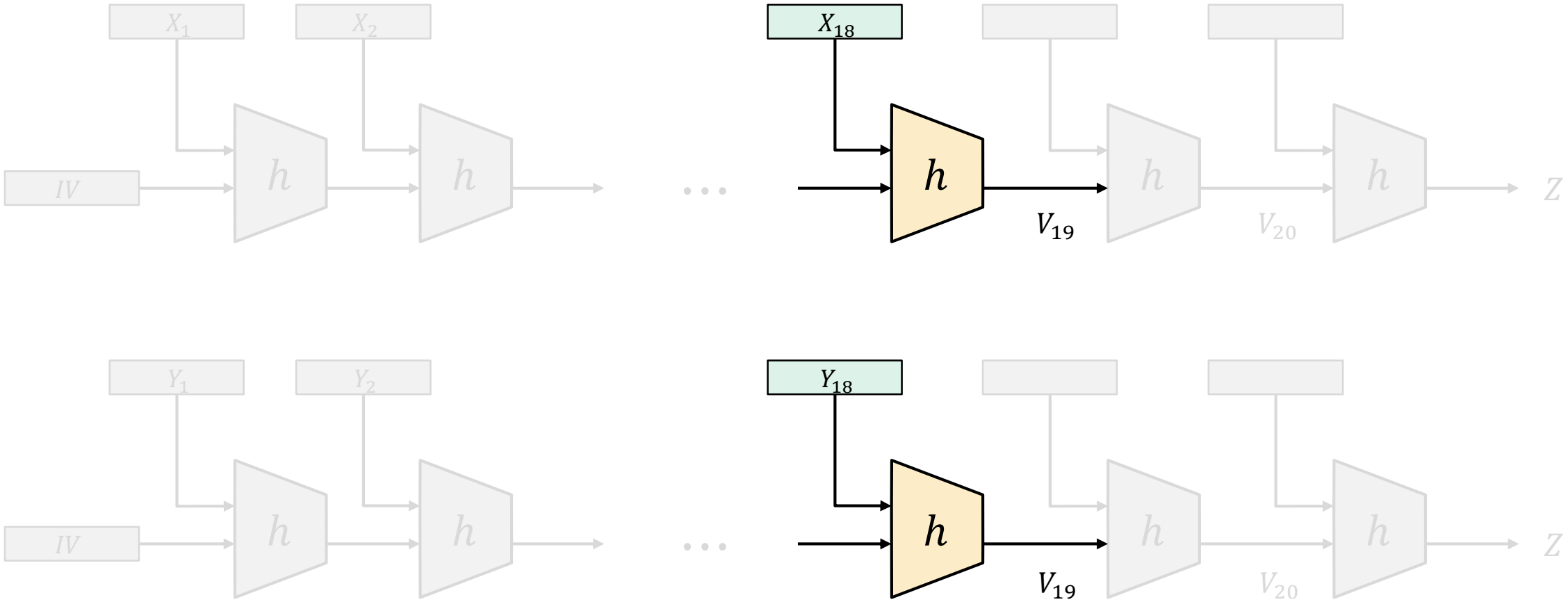
Merkle-Damgård – security



Merkle-Damgård – security



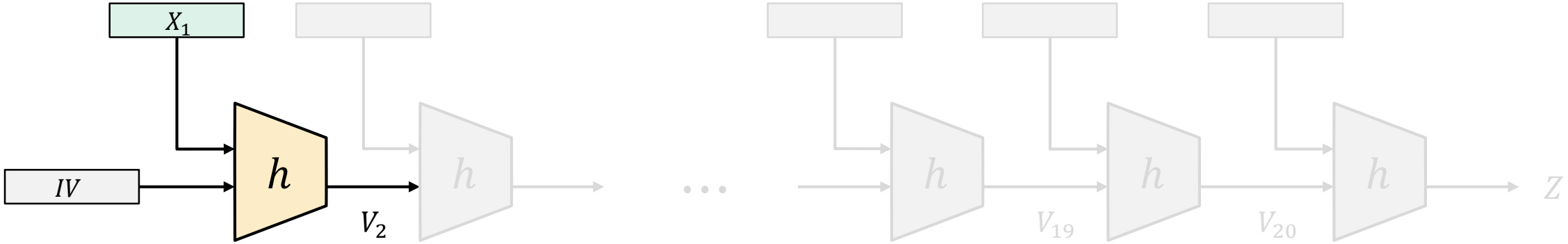
Merkle-Damgård – security

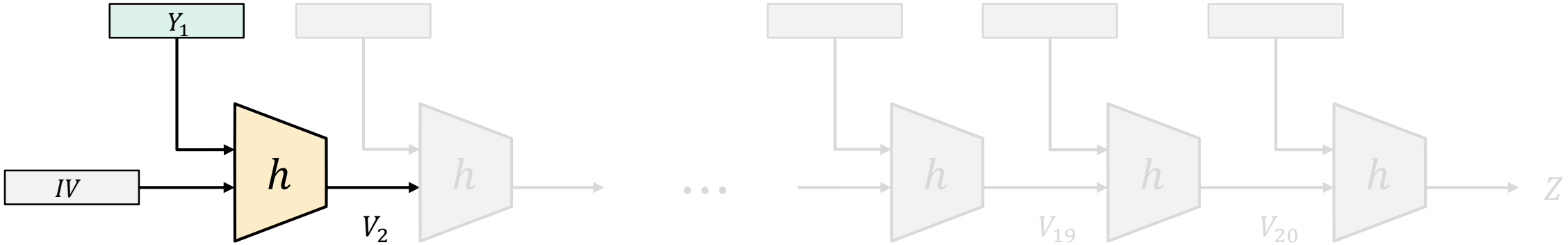


Merkle-Damgård – security

$$X_1 = Y_1 \implies \overbrace{X_1||X_2||\cdots||X_n}^{X} = \overbrace{Y_1||Y_2||\cdots||Y_n}^{Y}$$

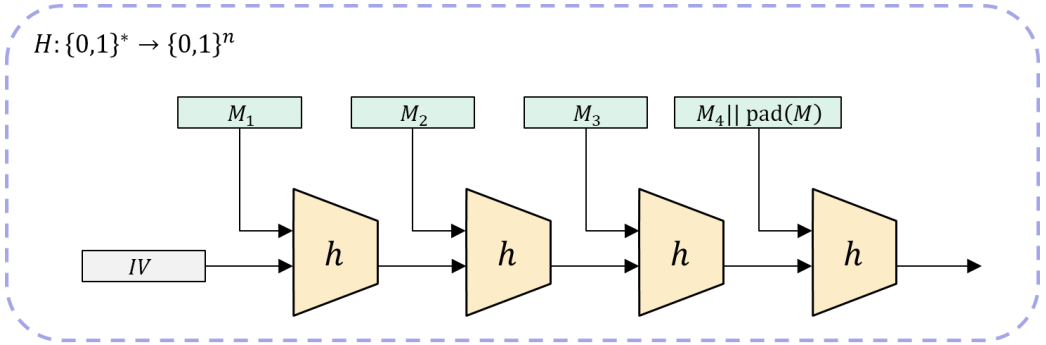
Contradiction! Assumed $X \neq Y$





Compression function designs

- Merkle-Damgård creates $H : \{0,1\}^* \rightarrow \{0,1\}^n$ from $h : \{0,1\}^{b+n} \rightarrow \{0,1\}^n$

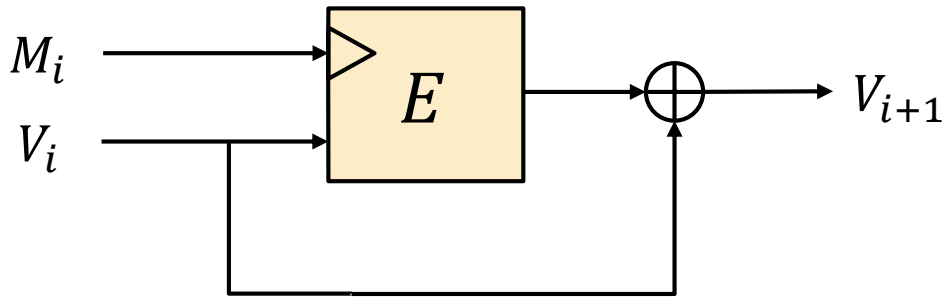


- Need only to focus on compression function

$$h : \{0,1\}^{b+n} \rightarrow \{0,1\}^n$$

- Many design options
  - Ad-hoc
  - Structured

Compression functions from block ciphers – Davies-Meyer

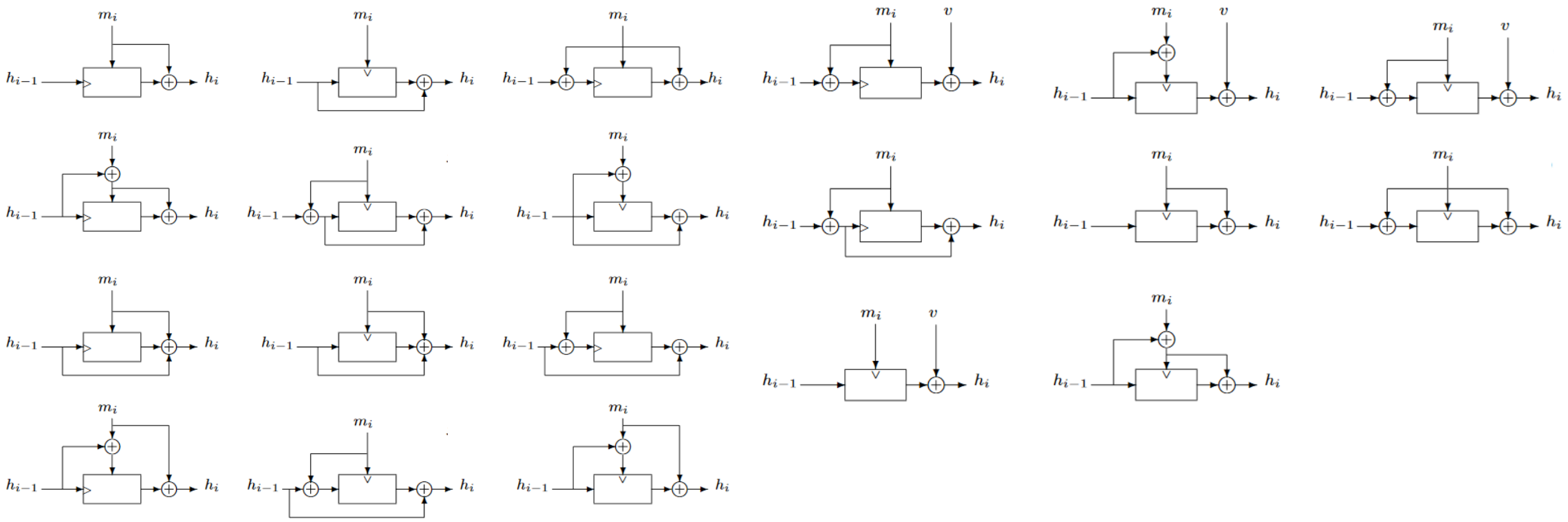


$$\mathrm{DM}(V_i \,||\, M_i) = E(M_i, V_i) \oplus V_i = V_{i+1}$$

Theorem: $\mathrm{DM} : \{0,1\}^{n+b} \rightarrow \{0,1\}^n$ is a collision resistance compression function in the ***ideal cipher model***. In particular, for *all* A making at most q hash queries:

$$\mathbf{Adv}^{\mathrm{cr}}_{\mathrm{DM}}(A) \leq \frac{(q+1)^2}{2^n}$$

Alternatives…



Black, Rogaway & Shrimpton: Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV; CRYPTO'02

SHA – Secure Hash Algorithm

- Family of hash functions published by NIST
  - SHA1 in 1995
  - SHA2-256 and SHA2-512 in 2001
    - Also truncated versions: SHA2-224, SHA2-384
  - SHA3-256 and SHA3-512 in 2015

- SHA1 and SHA2 designed by NSA
  - Merkle-Damgård + Davies-Meyer designs

- SHA3 designed by Belgian cryptographers
  - Sponge design

SHA1

SHA1(M)

1. $Y = M_1 || M_2 \cdots || M_n \leftarrow \text{pad}(M)$ // $|M_i| = 512$
2. $V_1 \leftarrow A_0 || B_0 || C_0 || D_0 || E_0$
3. **for** $i = 1 \ldots n$ **do**
4. $V_{i+1} \leftarrow E^{\text{sha1}}(M_i, V_i) \oplus V_i$ // Davies-Meyer
5. **return** V_{n+1}

pad(M)

---

1. $Y \leftarrow M \,||\, 100 \cdots 0 \,||\, |M|$
2. **return** Y

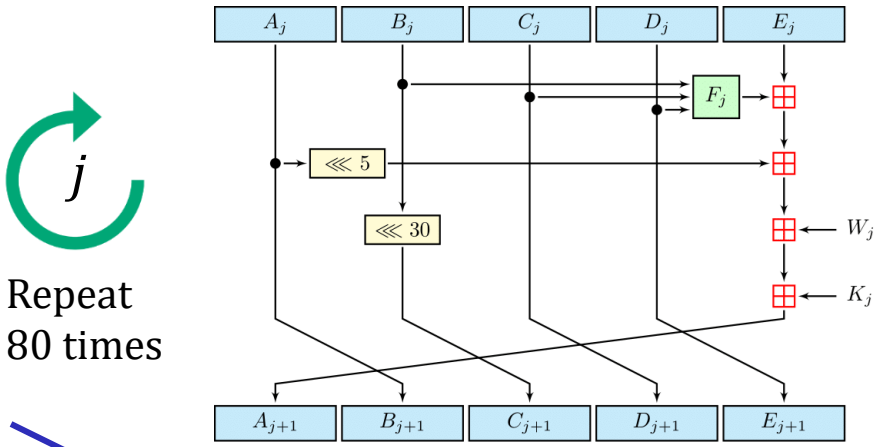
$A_0 = 67452301$
$B_0 = \text{EFCDAB89}$
$C_0 = 98\text{BADCFE}$
$D_0 = 10325476$
$E_0 = \text{C3D2E1F0}$

$$E^{\text{sha1}} : \{0,1\}^{512} \times \{0,1\}^{160} \rightarrow \{0,1\}^{160}$$

$E^{\text{sha1}}(M_i, V_i)$

$A_1 \leftarrow V_{i,1}$
$B_1 \leftarrow V_{i,2}$
$C_1 \leftarrow V_{i,3}$
$D_1 \leftarrow V_{i,4}$
$E_1 \leftarrow V_{i,5}$

for $j = 1 \ldots 16$ **do**
$W_j \leftarrow M_{i,j}$
for $j = 17 \ldots 80$ **do**
$W_j \leftarrow (W_{j-16} \oplus W_{j-14} \oplus W_{j-8} \oplus W_{j-3})_{\lll 1}$



Repeat
80 times

$\text{frac}_{32}(\sqrt{2})$
$\text{frac}_{32}(\sqrt{3})$
$\text{frac}_{32}(\sqrt{5})$
$\text{frac}_{32}(\sqrt{10})$

j	K_j	$F_j(B, C, D)$
1 ... 20	5A827999	$(B \wedge C) \vee (\bar{B} \wedge D)$
21 ... 40	6ED9EBA1	$B \oplus C \oplus D$
41 ... 60	8F1BBCDC	$(B \wedge C) \vee (B \wedge D) \vee (C \wedge D)$
61 ... 80	CA62C1D6	$B \oplus C \oplus D$

SHA0

SHA0(M)

1. $Y = M_1||M_2 \cdots ||M_n \leftarrow \text{pad}(M)$ // $|M_i| = 512$
2. $V_1 \leftarrow A_0||B_0||C_0||D_0||E_0$
3. **for** $i = 1 \ldots n$ **do**
4. $V_{i+1} \leftarrow E^{\text{sha0}}(M_i, V_i) \oplus V_i$ // Davies-Meyer
5. **return** V_{n+1}

pad(M)

---

1. $Y \leftarrow M \,||\, 100 \cdots 0 \,||\, |M|$
2. **return** Y

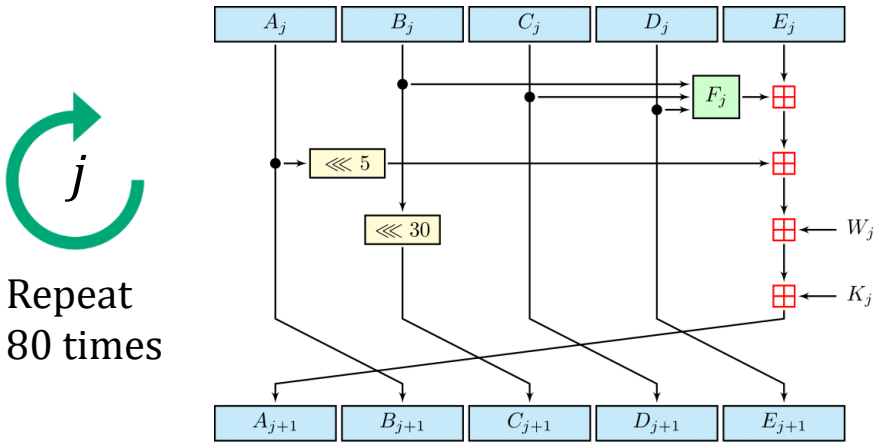
$A_0 = 67452301$
$B_0 = \text{EFCDAB89}$
$C_0 = \text{98BADCFE}$
$D_0 = 10325476$
$E_0 = \text{C3D2E1F0}$

$$E^{\text{sha0}} : \{0,1\}^{512} \times \{0,1\}^{160} \rightarrow \{0,1\}^{160}$$

$E^{\text{sha0}}(M_i, V_i)$

$A_1 \leftarrow V_{i,1}$
$B_1 \leftarrow V_{i,2}$
$C_1 \leftarrow V_{i,3}$
$D_1 \leftarrow V_{i,4}$
$E_1 \leftarrow V_{i,5}$

for $j = 1 \ldots 16$ **do**
  $W_j \leftarrow M_{i,j}$
for $j = 17 \ldots 80$ **do**
  $W_j \leftarrow (W_{j-16} \oplus W_{j-14} \oplus W_{j-8} \oplus W_{j-3})$



j	K_j	$F_j(B, C, D)$
1 … 20	5A827999	$(B \wedge C) \vee (\bar{B} \wedge D)$
21 … 40	6ED9EBA1	$B \oplus C \oplus D$
41 … 60	8F1BBCDC	$(B \wedge C) \vee (B \wedge D) \vee (C \wedge D)$
61 … 80	CA62C1D6	$B \oplus C \oplus D$

SHA – Secure Hash Algorithm

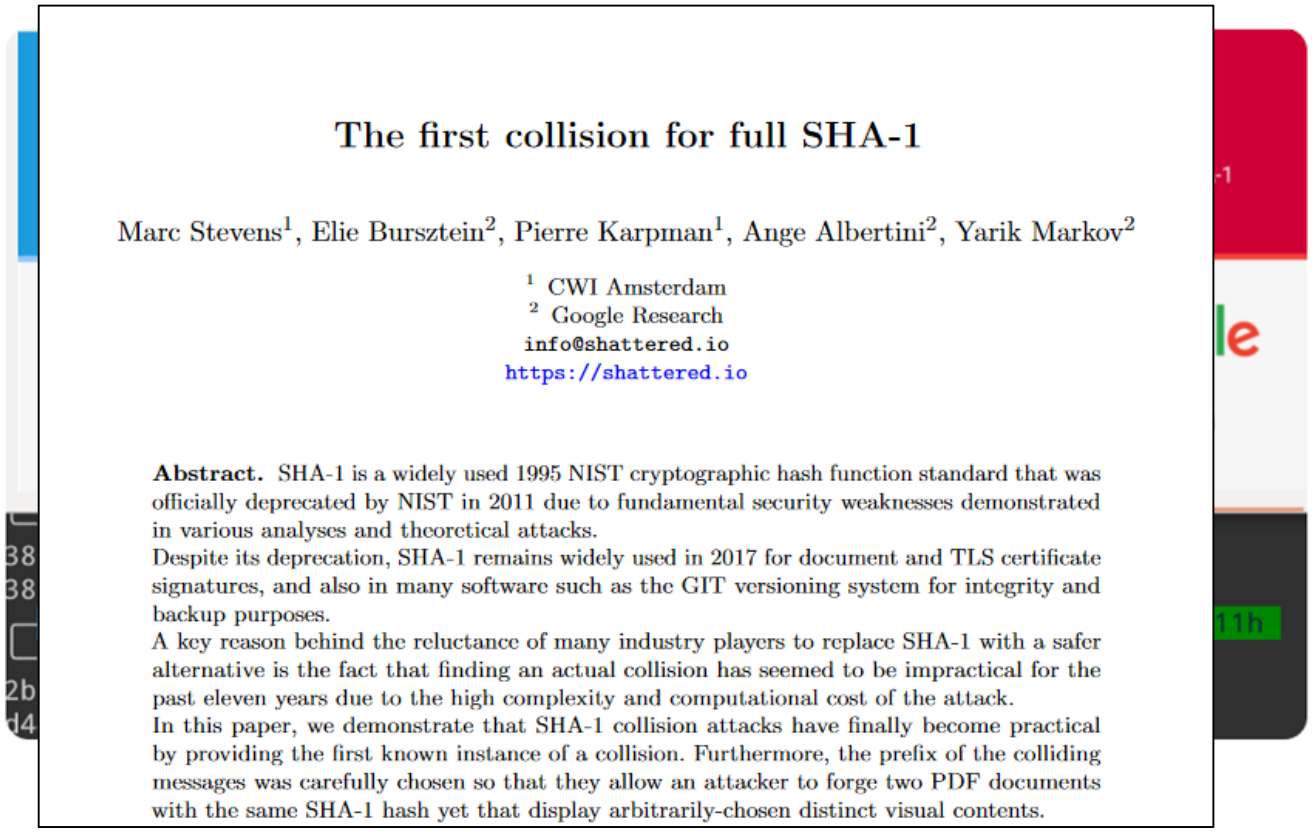
- Family of hash functions published by NIST
  - **SHA0 in 1993**
  - SHA1 in 1995
  - SHA2-256 and SHA2-512 in 2001
    - Also truncated versions: SHA2-224, SHA2-384
  - SHA3-256 and SHA3-512 in 2015
- SHA1 and SHA2 designed by NSA
  - Merkle-Damgård + Davies-Meyer designs
- SHA3 designed by Belgian cryptographers
  - Sponge design

Attacks against SHA0

- *Chabaud & Joux '98*: SHA0 can be attacked in 2^{61} operations
- *Biham & Chen '04*: Broke 62 / 80 rounds
- *Joux, et al. '04*: **First collision found** taking 2^{51} operations
- *Wang et al. '04*: **Second collision found** taking 2^{40} operations
- *Wang et al. '05*: Collisions in 2^{39} operations
- *Manuel & Peyrin '08*: Collisions in $2^{33.6}$ operations

Attacks against SHA1

- *Rijemen & Oswald '05*: Broke 53 / 80 rounds
- *Wang et al. '05*: "Collisions can be found in 2^{63} operations" (*'04*: SHA0 collision found)
- *'06-'10*: various theoretical improvements
- *Stevens, Karpman & Peyrin '15*: Freestart collision in 2^{57} operations
- ***Stevens & Karpman + Google* 2017: First SHA1 collision found** ($2^{63.1}$ operations)



The first collision for full SHA-1

Marc Stevens[1], Elie Bursztein[2], Pierre Karpman[1], Ange Albertini[2], Yarik Markov[2]

[1] CWI Amsterdam
[2] Google Research
info@shattered.io
https://shattered.io

Abstract. SHA-1 is a widely used 1995 NIST cryptographic hash function standard that was officially deprecated by NIST in 2011 due to fundamental security weaknesses demonstrated in various analyses and theoretical attacks.
Despite its deprecation, SHA-1 remains widely used in 2017 for document and TLS certificate signatures, and also in many software such as the GIT versioning system for integrity and backup purposes.
A key reason behind the reluctance of many industry players to replace SHA-1 with a safer alternative is the fact that finding an actual collision has seemed to be impractical for the past eleven years due to the high complexity and computational cost of the attack.
In this paper, we demonstrate that SHA-1 collision attacks have finally become practical by providing the first known instance of a collision. Furthermore, the prefix of the colliding messages was carefully chosen so that they allow an attacker to forge two PDF documents with the same SHA-1 hash yet that display arbitrarily-chosen distinct visual contents.

SHA2-256

SHA2–256(M)

1. $Y = M_1 || M_2 \cdots || M_n \leftarrow \text{pad}(M)$ // $|M_i| = 512$
2. $V_1 \leftarrow A_0 || B_0 || C_0 || D_0 || E_0$
3. **for** $i = 1 \ldots n$ **do**
4. $V_{i+1} \leftarrow E^{\text{sha2}}(M_i, V_i) \boxplus V_i$ // Davies-Meyer
5. **return** V_{n+1}

pad(M)

1. $Y \leftarrow M \,||\, 100 \cdots 0 \,||\, |M|$
2. **return** Y

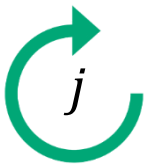
$A_0 = 6A09E667$ $E_0 = 510E527F$
$B_0 = BB67AE85$ $F_0 = 9B05688C$
$C_0 = 3C6EF372$ $G_0 = 1F83D9AB$
$D_0 = A54FF53A$ $H_0 = 5BE0CD19$

$$E^{\text{sha2}} : \{0,1\}^{512} \times \{0,1\}^{256} \rightarrow \{0,1\}^{256}$$

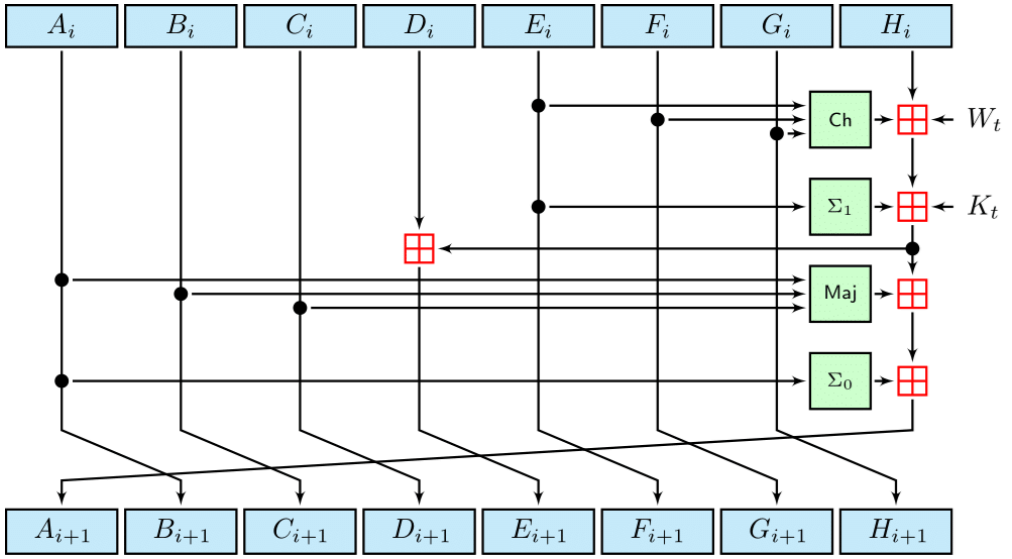
$E^{\text{sha2}}(M_i, V_i)$

$A_1 \leftarrow V_{i,1}$
$B_1 \leftarrow V_{i,2}$
$C_1 \leftarrow V_{i,3}$
$D_1 \leftarrow V_{i,4}$
$E_1 \leftarrow V_{i,5}$
$F_1 \leftarrow V_{i,6}$
$G_1 \leftarrow V_{i,7}$
$H_1 \leftarrow V_{i,8}$

for $j = 1 \ldots 16$ **do**
  $W_j \leftarrow M_{i,j}$
for $j = 17 \ldots 64$ **do**
  $\sigma_0 \leftarrow \text{ROTR}^7(W_{j-15}) \oplus \text{ROTR}^{18}(W_{j-15}) \oplus \text{SHR}^3(W_{j-15})$
  $\sigma_1 \leftarrow \text{ROTR}^{17}(W_{j-2}) \oplus \text{ROTR}^{19}(W_{j-2}) \oplus \text{SHR}^{10}(W_{j-2})$
  $W_j \leftarrow W_{j-16} + \sigma_0 + W_{i-7} + \sigma_1$

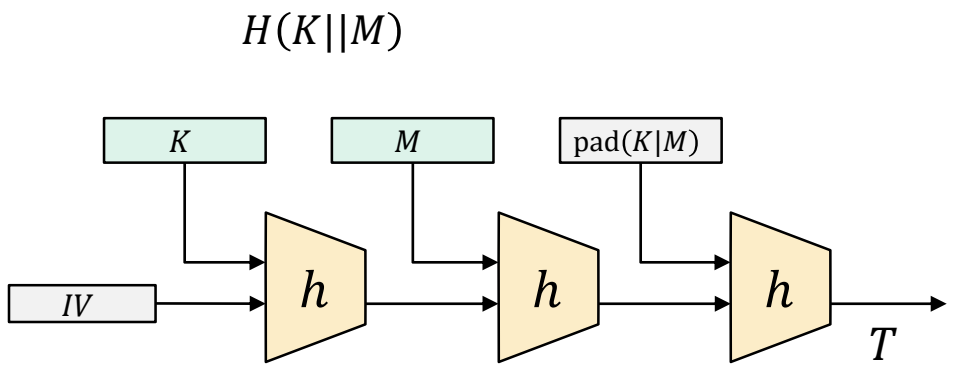


Repeat
64 times



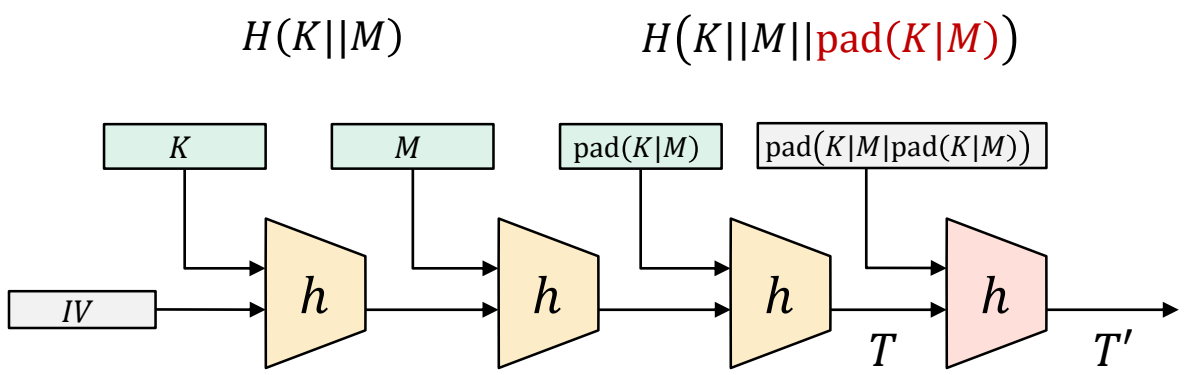
MACs from hash functions – $H(K \,||\, M)$

- Doesn't work in general
- **Length-extension attacks** on Merkle-Damgård hash functions



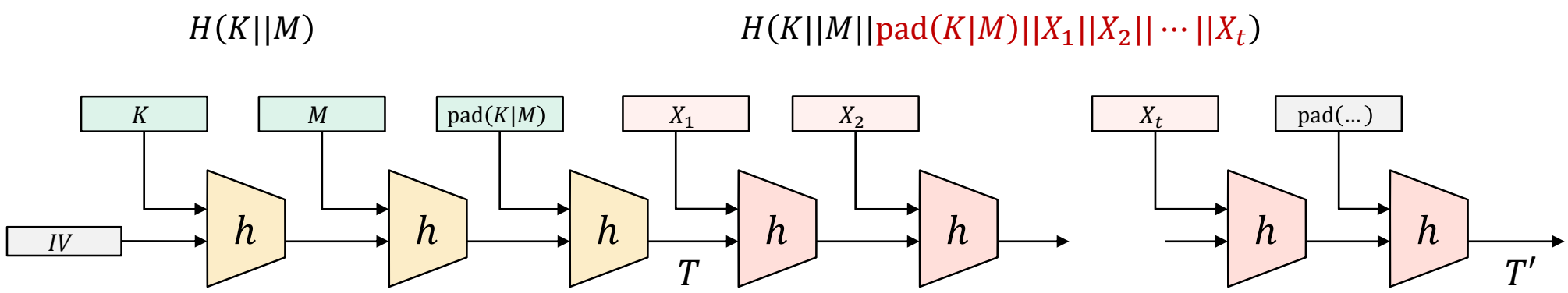
MACs from hash functions – $H(K \,||\, M)$

- Doesn't work in general
- **Length-extension attacks** on Merkle-Damgård hash functions

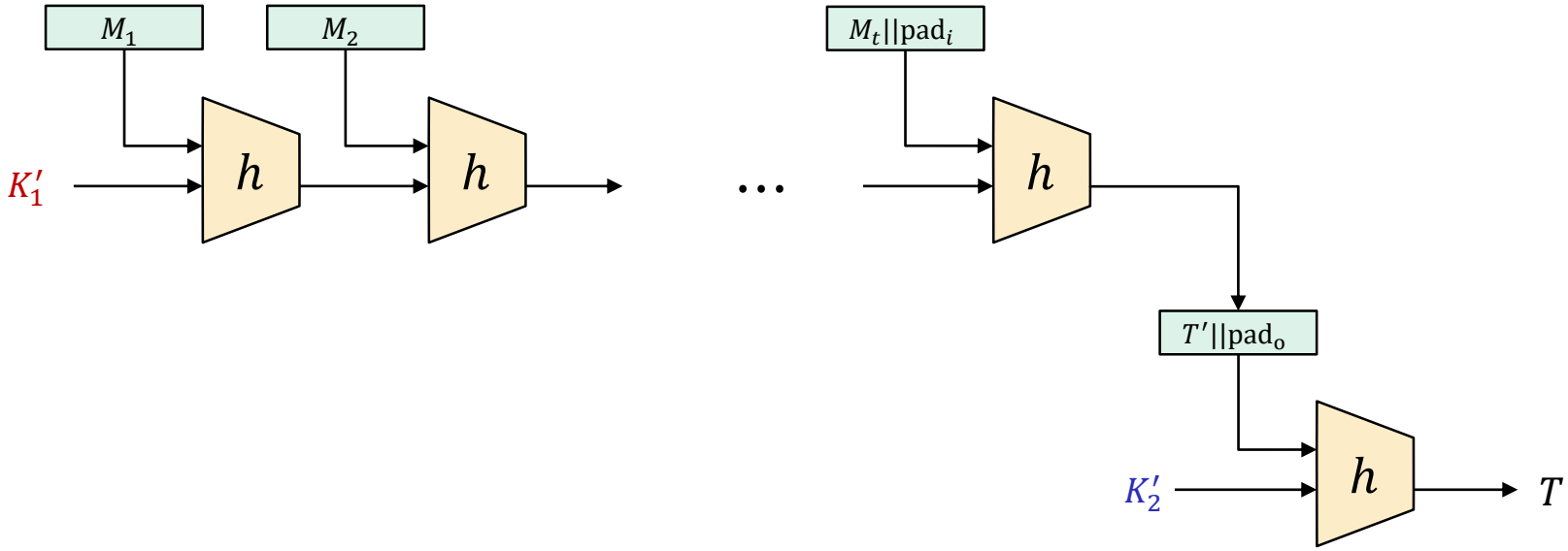


MACs from hash functions – $H(K \,||\, M)$

- Doesn't work in general
- **Length-extension attacks** on Merkle-Damgård hash functions



NMAC

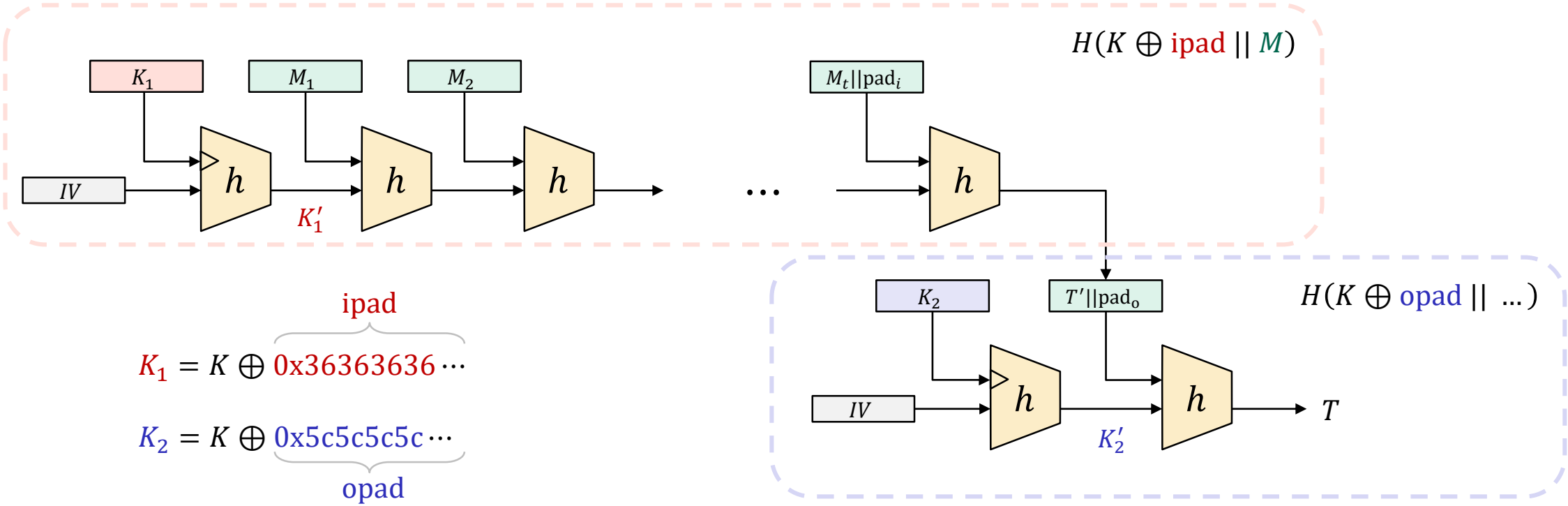


Theorem (Gaži, Pietrzak, Rybár '2014):
If $h : \{0,1\}^n \times \{0,1\}^b \to \{0,1\}^n$ is a secure PRF then $\text{NMAC} : \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^n$ is a secure PRF.

Detailed: for *any* adversary A against NMAC making q PRF oracle queries of at most ℓ blocks, there is an adversary B against h such that:

$$\mathbf{Adv}^{\text{prf}}_{\text{NMAC}}(A) \leq (q\ell + q + 1) \cdot \mathbf{Adv}^{\text{prf}}_{h}(B) + \frac{q^2}{2^n}$$

HMAC



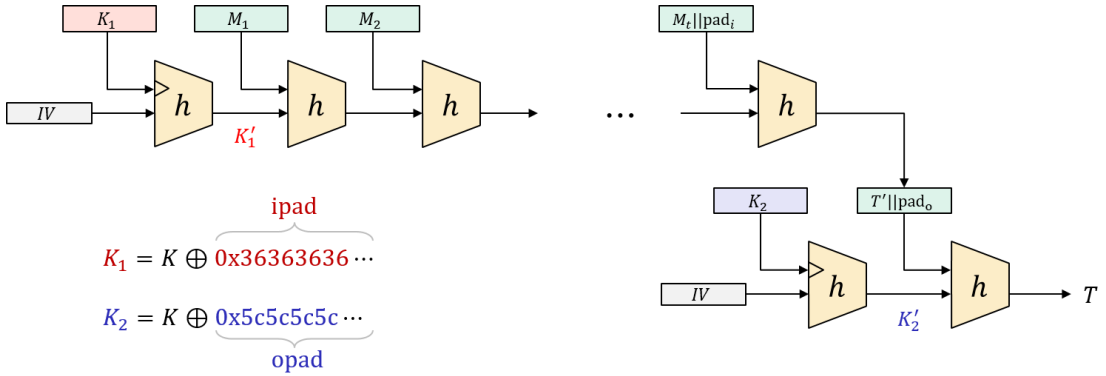
$$K_1 = K \oplus \overbrace{0\text{x}36363636\cdots}^{\text{ipad}}$$

$$K_2 = K \oplus \underbrace{0\text{x}5c5c5c5c\cdots}_{\text{opad}}$$

$$\text{HMAC}(K, M) = H\big(K \oplus \text{opad} \;||\; H(K \oplus \text{ipad} \,||\, M)\big)$$

HMAC

- *Very* widely used
  - TLS
  - IPsec
  - SSH
  - LTE
  - …



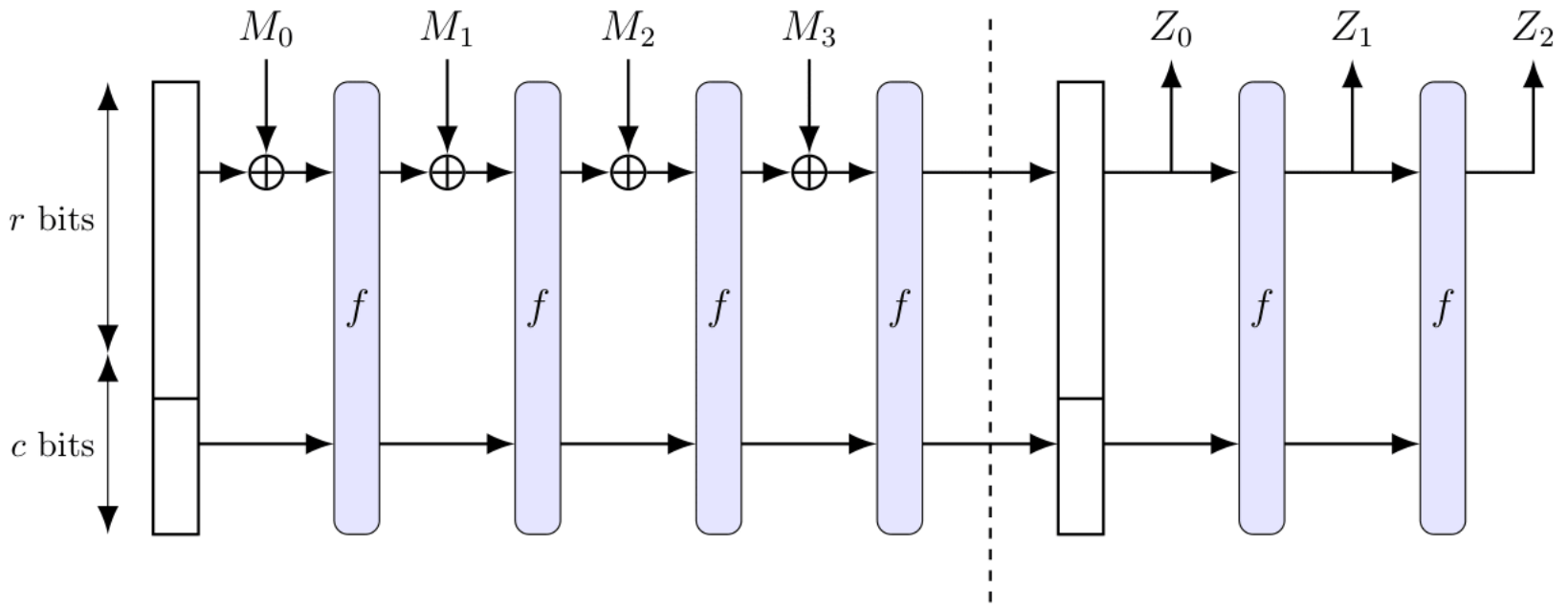
$$\text{HMAC}(K, M) = H\big(\, K \oplus \text{opad} \;\; || \;\; H(K \oplus \text{ipad} \,||\, M)\big)$$

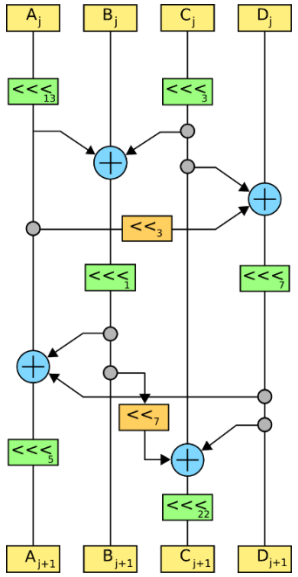
- Standardized by NIST and IETF (RFC 2104)

- Security proof for NMAC can be lifted to HMAC:
  - ++ assume h is a secure **dual** PRF (h keyed through the message is also a secure PRF)
  - ++ assume h is secure against **related-key attacks** (since K_1, K_2 are derived from K)
  - very tricky proof; controversies

SHA3

$f : \{0,1\}^{1600} \rightarrow \{0,1\}^{1600}$





END OF PART 1
(SYMMETRIC CRYPTO)

Summary of symmetric cryptography

Primitive	Functionality + syntax	Security goal	Acronym	Examples
Pseudorandom function	Keyed function mapping fixed-length input to fixed-length output $F : \mathcal{K} \times \{0,1\}^{\text{in}} \to \{0,1\}^{\text{out}}$	Indistinguishability from random function	PRF	AES HMAC
Block cipher / pseudorandom permutation	Encrypt fixed-length block $E : \mathcal{K} \times \{0,1\}^{n} \to \{0,1\}^{n}$	Indistinguishability from random permutation	PRP	AES
Encryption	Encrypt variable-length input $\text{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ $\text{Enc} : \mathcal{K} \times \mathcal{N} \times \mathcal{M} \to \mathcal{C}$ (nonce-based)	Confidentiality: attacker should learn nothing about plaintext (except length) from ciphertexts	IND-CPA, IND$-CPA IND-CCA	CTR$ CBC$
MAC	Produce fixed-length tag on variable-length message $\text{Tag} : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$ $\text{Vrfy} : \mathcal{K} \times \mathcal{M} \times \mathcal{T} \to \{\text{Valid}, \text{Invalid}\}$	Integrity: attacker shouldn't be able to forge messages, i.e., create new messages with valid tags	UF-CMA	CBC-MAC CMAC HMAC
Authenticated encryption	Encrypt variable-length input $\text{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ $\text{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M} \cup \{\perp\}$ With associated data + nonces (AEAD) $\text{Enc} : \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M} \to \mathcal{C}$ $\text{Enc} : \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{C} \to \mathcal{M} \cup \{\perp\}$	Confidentiality + ciphertext integrity Confidentiality (message) + ciphertext integrity + AD integrity	AE	EtM GCM OCB CCM
Hash function	Keyless function mapping variable-length messages to fixed-length tags $H : \mathcal{M} \to \mathcal{Y}$ $H : \{0,1\}^{*} \to \{0,1\}^{n}$	Collision-resistance + one-wayness		SHA1 SHA2-256 SHA2-512 SHA3

Midterm exam

- Home-exam
- Published on Canvas after next week’s lecture (Tuesday 6. October)
- Due: Tuesday 20. October (2 weeks)
- Pass/fail
- Must pass in order to take final exam

Next week – October 6

- Guest lecture (I’m away)