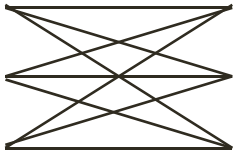
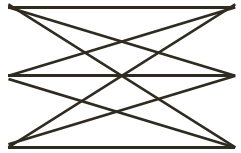
Bicliques for Preimages: Attacks on Skein-512 and the SHA-2 family

Dmitry Khovratovich[1] Christian Rechberger[2]
Alexandra Savelieva[3]





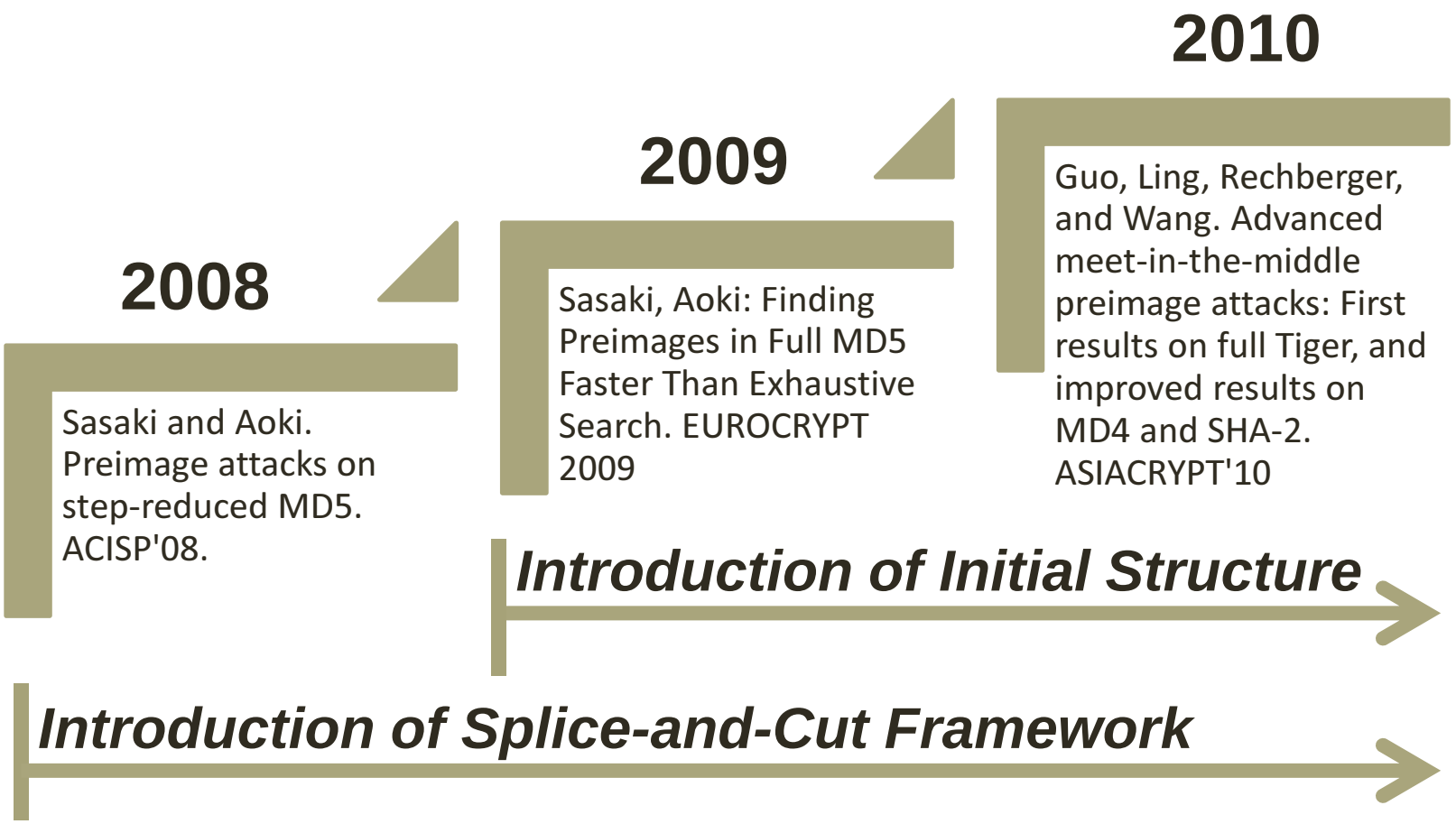
[1]Microsoft Research Redmond, USA

[2]DTU MAT, Denmark

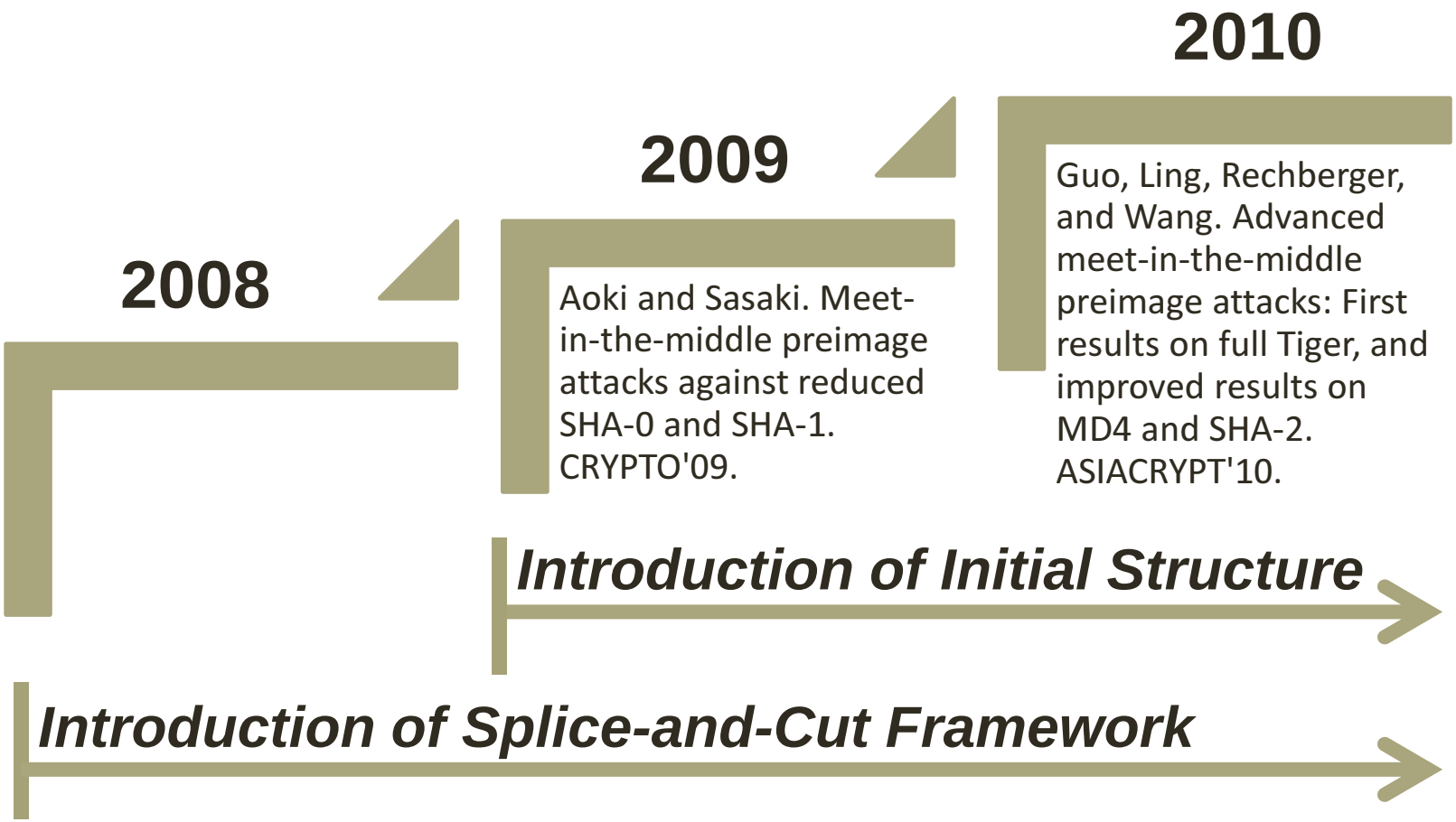
[3]National Research University Higher School of Economics, Russia

19th International Workshop on Fast Software Encryption - FSE 2012
March 19-21, 2012

Recent Progress in Preimage Attacks – MD4, MD5, and Tiger



Recent Progress in Preimage Attacks – SHA-x Family



Problem

- Concrete examples of the initial structure are extremely sophisticated and hard to generalize.
- Many ad-hoc / not formalized techniques are used to build initial structures
- While the other elements of splice-and-cut framework seem exhausted already, the concept behind initial structure has large potential and few boundaries.

Purpose of our Research

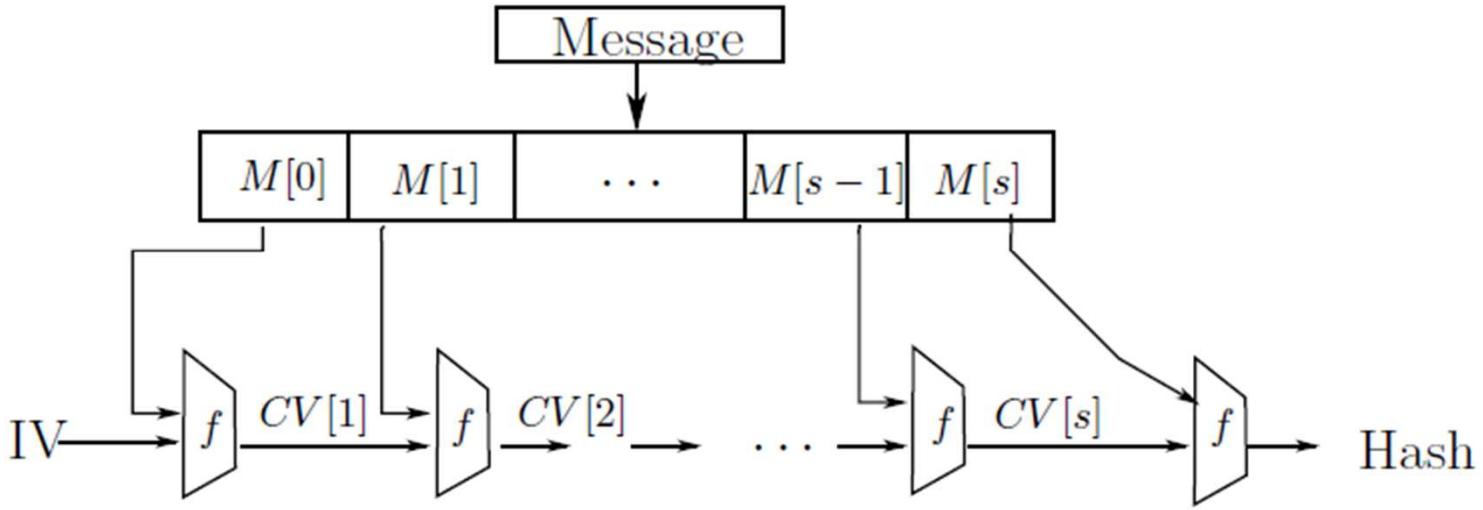
- To replace the idea of the initial structure with a more formal and generic concept
- To design generic algorithms for constructing the initial structure
- To reduce manual efforts and time to build the initial structure

New attacks on Skein-512 and the SHA-2 family

Reference	Target	Steps	Complexity, 2^x			Memory
			Pseudo-preimage	Second preimage	Preimage	
Our results	**Skein-512**	**22**	**508**	**511**	**-**	**2^6**
Our results	**Skein-512**	**37**	**511.2**	**-**	**-**	**2^{64}**
Our results	**Skein-512**	**72**	**-**	**511.7**	**-**	**Negl.**
Aoki et.al.'09	SHA-256	43	251.9	254.9	254.9	2^6
Our results	**SHA-256**	**45**	**253**	**255.5**	**255.5**	**2^6**
Our results	**SHA-256**	**52**	**255**	**-**	**-**	**2^6**
Aoki et.al.'09	SHA-512	46	509	511.5	511.5	2^6
Our results	**SHA-512**	**50**	**509**	**511.5**	**511.5**	**2^4**
Our results	**SHA-512**	**57**	**511**	**-**	**-**	**2^6**

Hash Functions with Merkle-Damgård Structure

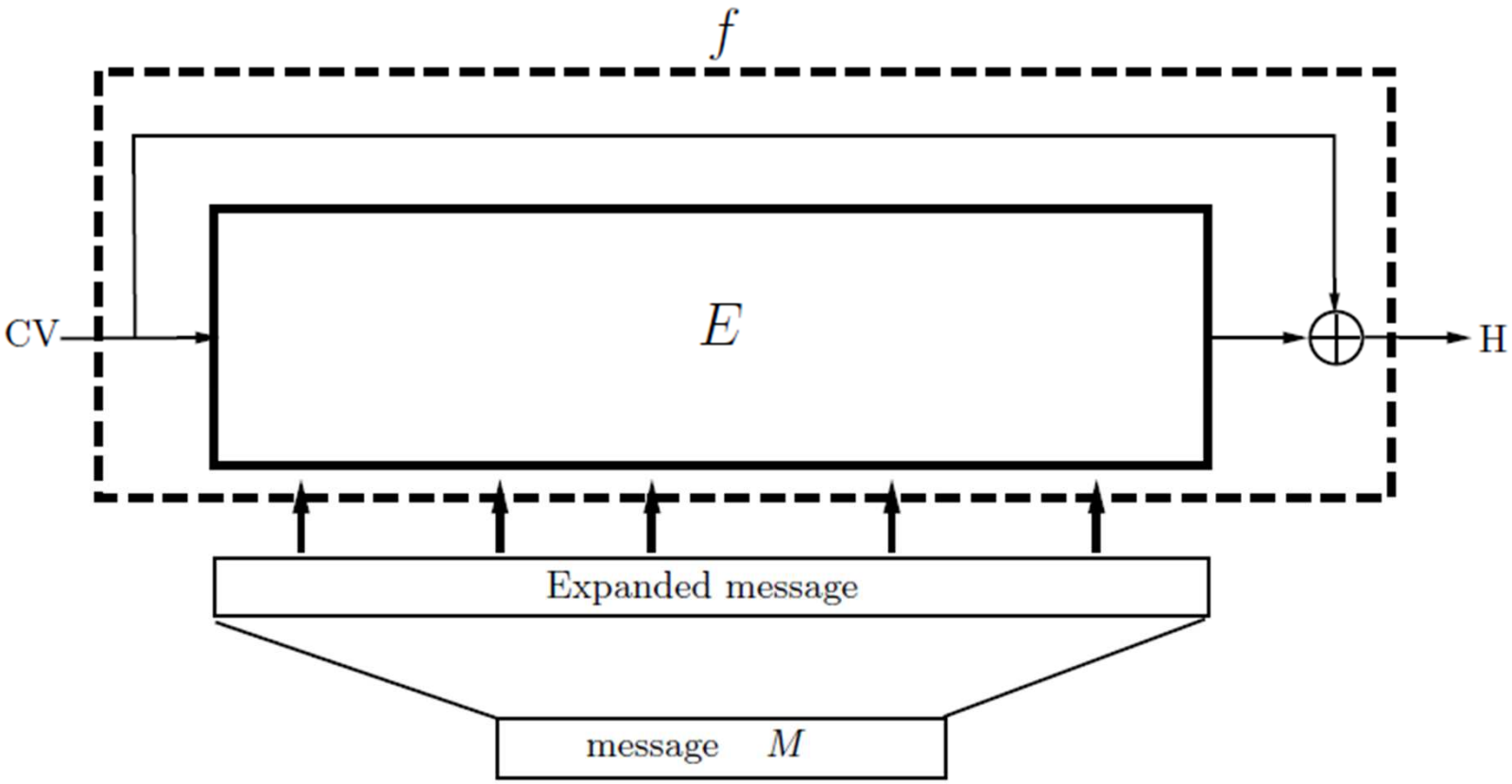
- M is arbitrarily long
- Iterative design
- $H(M) = f(M[s]; CV[s])$
- $CV[i_{+1}] = f(M[i]; CV[i])$



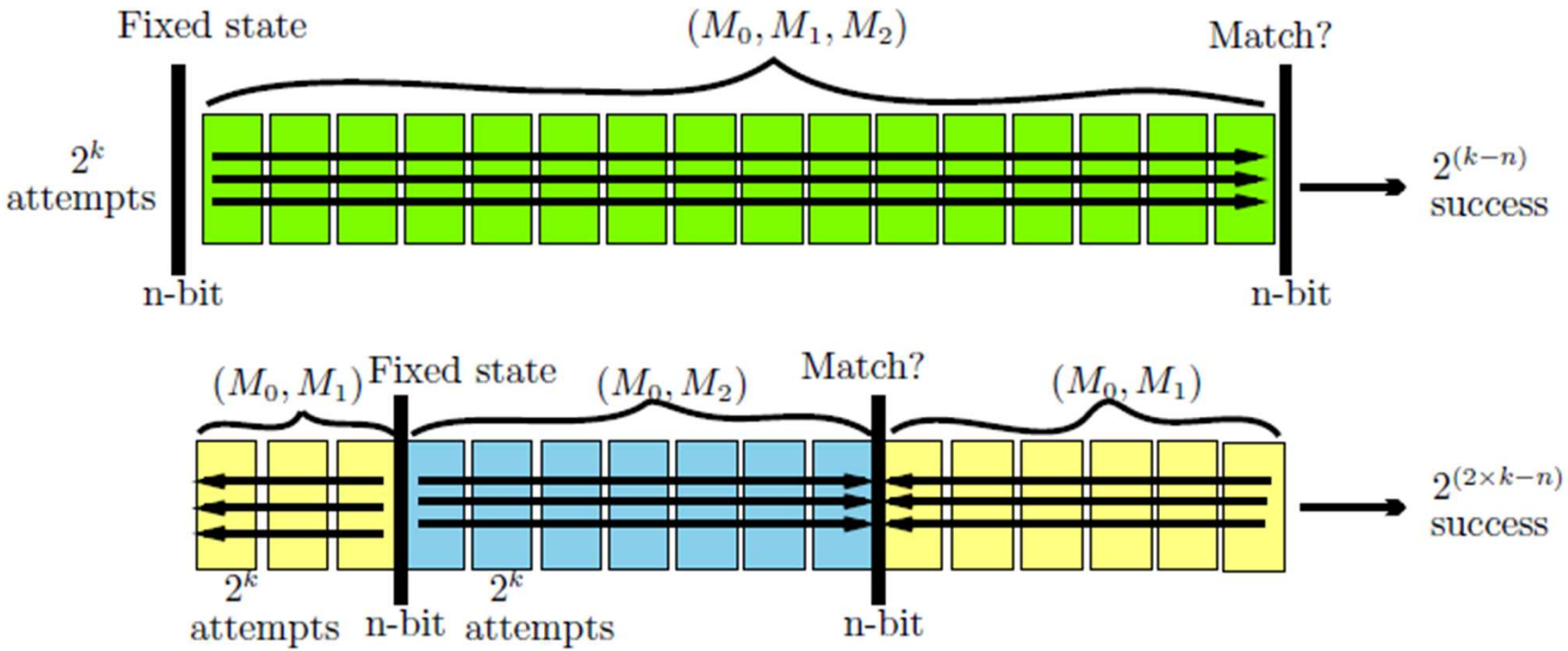
Compression Functions in Davies-Meier Mode

- Blockcipher-based compression function:

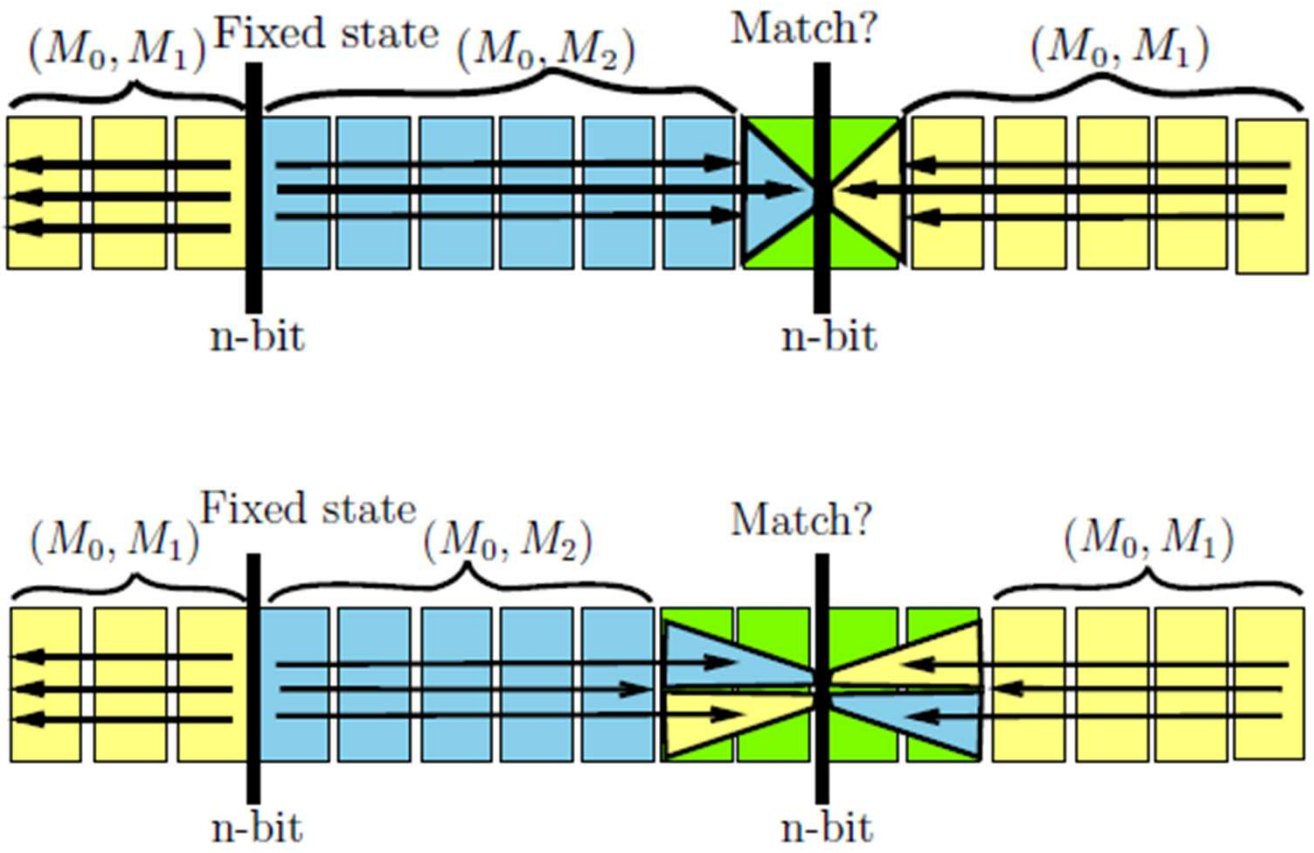
$$f(M; CV) = E_M(CV) \oplus CV;$$



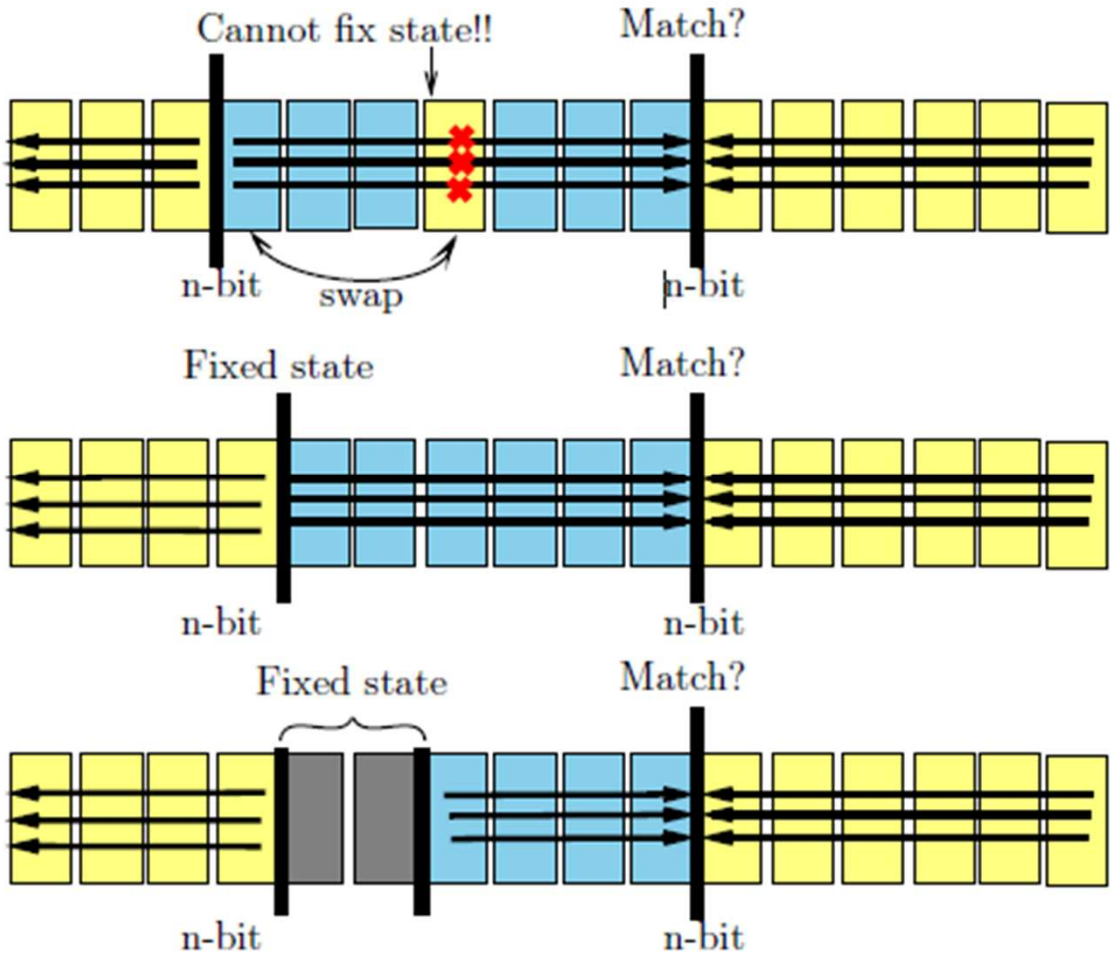
Splice-and-Cut Attacks: Basic Strategy



Partial Matching and ~ Fixing



Initial Structure



Biclique – Formal Definition

Let ϵ be a sub-cipher of E, and $\mathcal{M} = \{M[i,j]\}$ be a group of parameters for ϵ. Then a *biclique of dimension d* over ϵ for $\mathcal{M}$ is a pair of sets $\{Q_i\}$ and $\{P_j\}$ of 2^d states each such that

$$Q_i \xrightarrow[\epsilon]{M[i,j]} P_j.$$

If $M[i,j]$ is a preimage, then

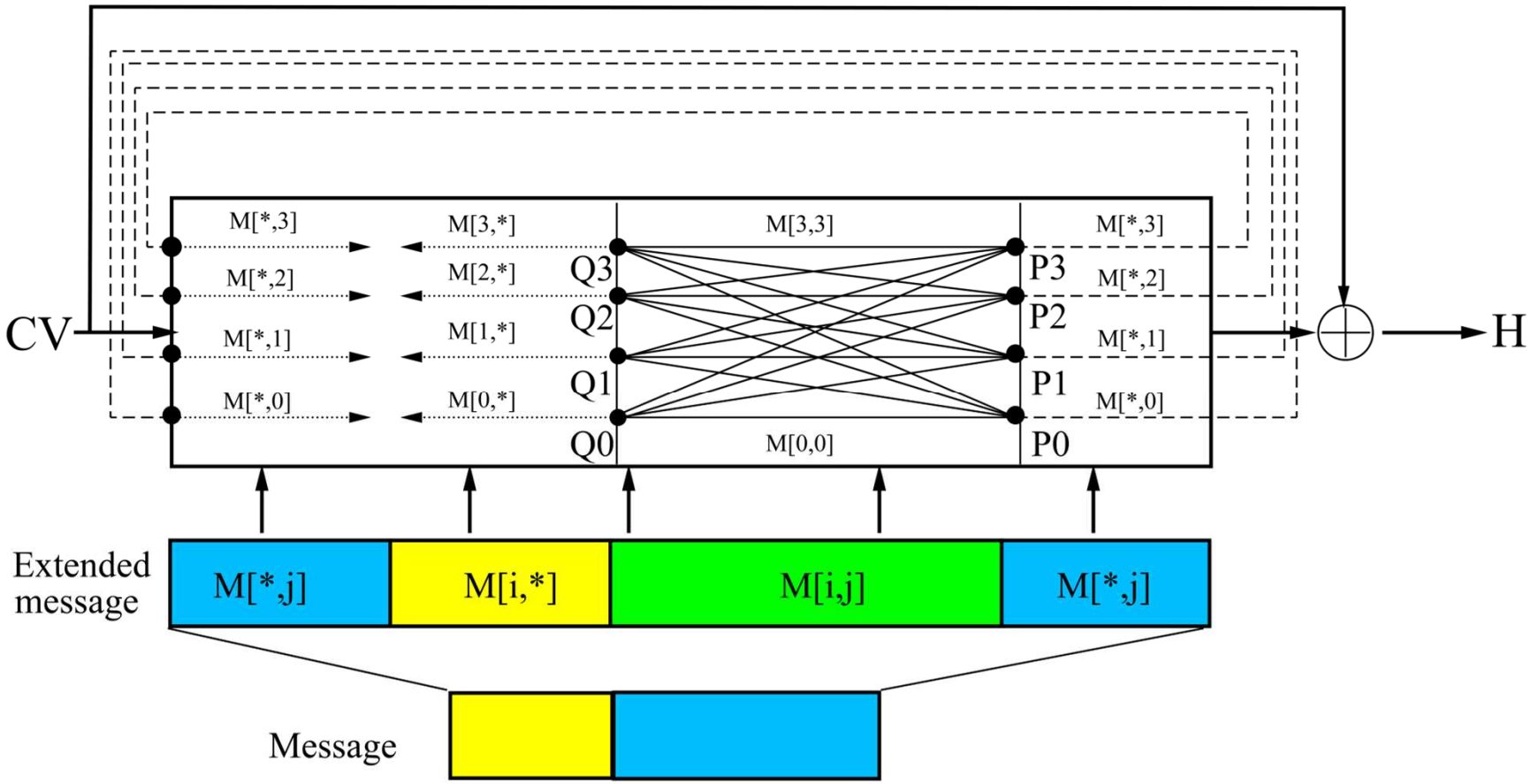
$$E: \quad CV \xrightarrow{M[i,j]} Q_i \xrightarrow[f]{M[i,j]} P_j \xrightarrow{M[i,j]} H.$$

An adversary selects a variable v outside of ϵ (w.l.o.g. between P_j and H) and checks, for appropriate choices of sub-ciphers g_1 and g_2, if

$$\exists i,j: \; P_j \xrightarrow[g_1]{M[i,j]} v \stackrel{?}{=} v \xleftarrow[g_2]{M[i,j]} Q_i.$$

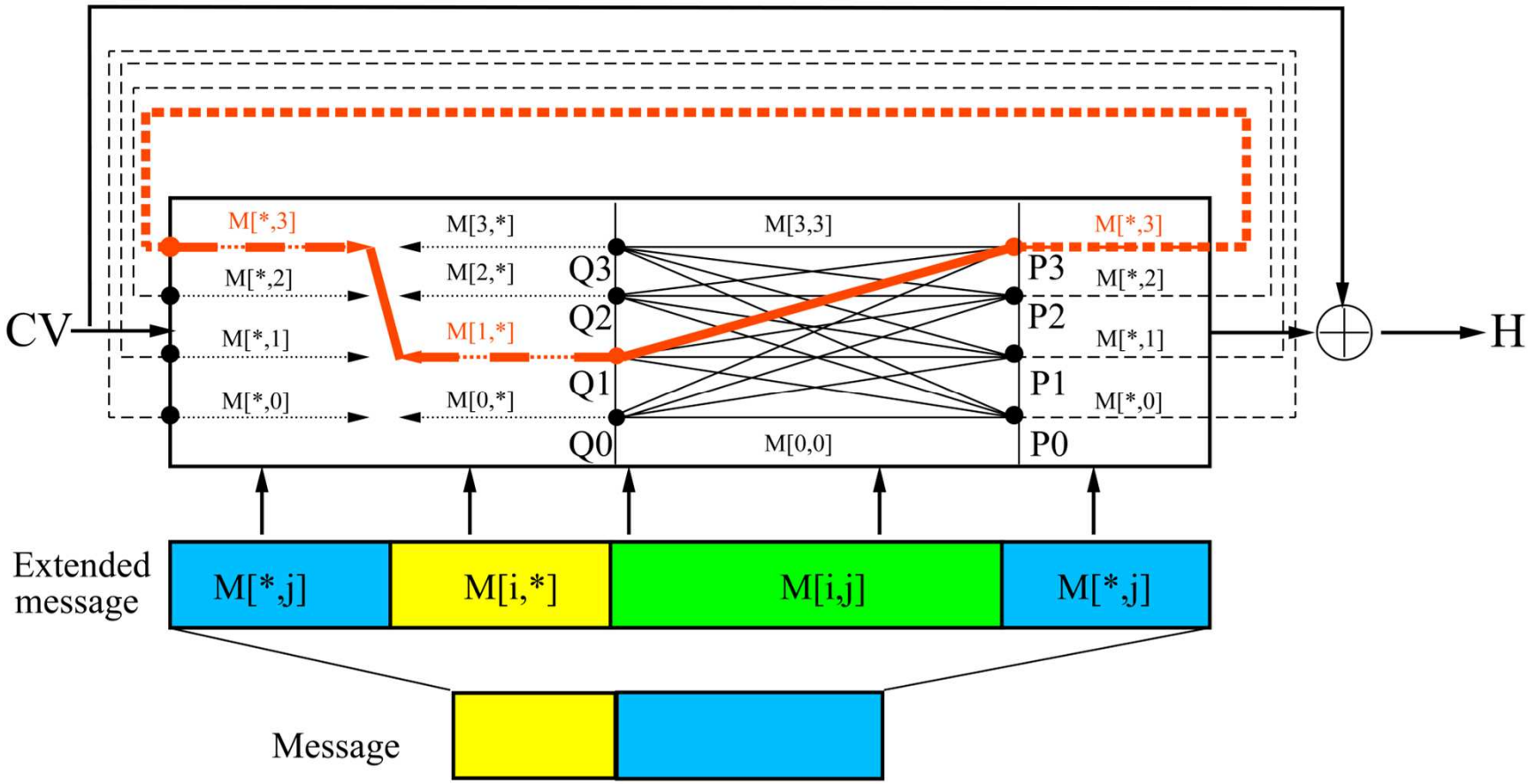
A positive answer yields a candidate preimage. To compute v from Q_i, the adversary computes CV and then derives the output of E as $CV \oplus H$.

Biclique of dimension 2 in the MITM attack on a DM compression function



How it works

- Suppose message M[1,3] is a preimage



Advantage

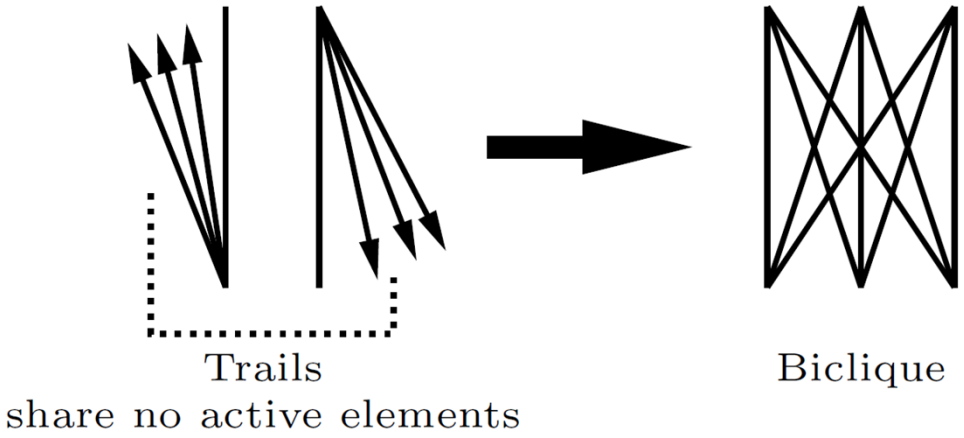
- The complexity of testing 2^{2d} messages for preimages:

$$C = 2^{d}(C_{backward} + C_{forward}) + C_{bicl} \quad [+ C_{recheck}]$$

- One needs 2^{n-2d} bicliques of dimension d to test 2^{n} preimage candidates.

Differential Perspective on Bicliques

- Vast pool of already existing tools when it comes to finding differential trails in hash functions
- Very precise and economic use of degrees of freedom in the resulting attacks



Biclique Construction Algorithms

#	Main idea	Application	Attacks
1	Fully specified or truncated differential trails	Bicliques of arbitrary dimension	Reduced Skein hash function
2	Modification of Algorithm 1 for hash functions in DM mode	For the case when we control internal state and message injections within the biclique	Reduced SHA-2 hash and compression functions
3	Use rebound approach to get more rounds	For bicliques of dimension 1	Reduced Skein compression function

Number of Attacked SHA-2 Hash Function Rounds - Our Improvements

Hash function	Chunks	Partial matching	Partial fixing	Initial structure	Total
SHA-256	29	7	3	4**+2**	43**+2**
SHA-512	29	7	8	2**+4**	46**+4**

- Compared to:

Aoki, Guo, Matusiewicz, Sasaki, and Wang. Preimages for step-reduced SHA-2. In ASIACRYPT'09.

Summary of Our Contributions

- Formalization of Initial Structure technique as a ‘Biclique’
  - 3 generic and flexible algorithms for constructing bicliques
  - differential perspective that allows for application of differential trails, message modification techniques etc. in splice-and-cut framework
- SHA-2 family
  - attack on 45-round SHA-256 and 50-round SHA-512 in the hash mode
  - attack on 52-round SHA-256 and 57-round SHA-512 compression function
- SHA-3 finalist Skein
  - attack on 22 rounds of Skein-512 hash function
  - attack on 37 rounds of Skein-512 compression function
  - MITM speed-up of brute force attack on 72 rounds of Skein-512

Results in Perspective

- We targeted a main security property, not some artificial distinguishing property.

- We have results on the real hash, not some pseudo-attacks, or results that only work with full access to compression function, cipher or permutation

Follow-up Work

- Biclique Cryptanalysis of the Full AES by Bogdanov, Khovratovich, and Rechberger (2011) - *First application to block ciphers*
- A Meet-in-the-Middle Attack on the Full KASUMI by Jia, Yu, and Wang (2011) - *Exploits a new property of the cipher*
- Narrow Bicliques: Cryptanalysis of Full IDEA by Khovratovich, Leurent, and Rechberger (2012) - *Variants of attacks that are many million times faster than brute-force*
- *Even more results on:* SQUARE (by Mala, 2011), IDEA (by Biham, Dunkelman, Keller, and Shamir, 2011), and ARIA (by Chen and Xu, 2012)

Future Work

- Application of the Biclique framework to other hash functions and block ciphers
- Generalization of the Biclique technique, e.g. identifying situations where a graph can be used that deviates from the Biclique definition.
- New design criteria for hash-functions based on their ability to resist meet-in-the-middle attacks

Questions?

Dmitry Khovratovich Christian Rechberger
Alexandra Savelieva