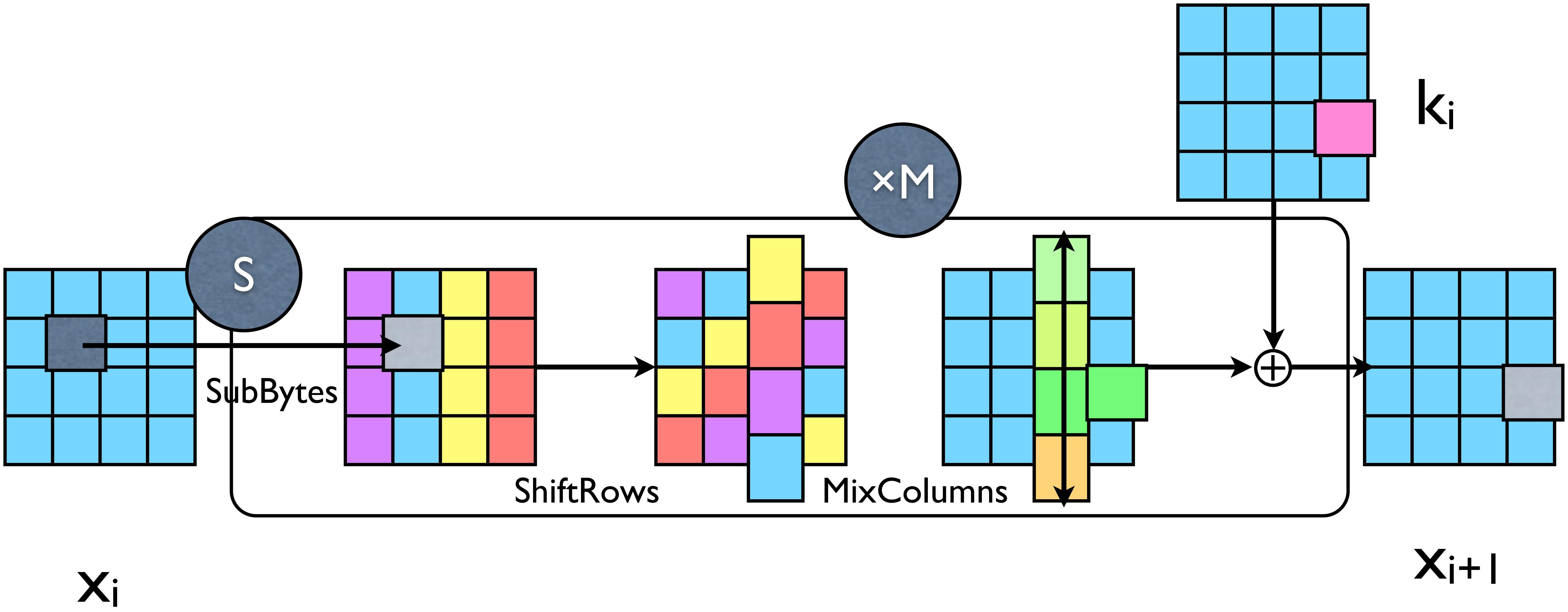
Sécurité de AES

Pierre-Alain Fouque

Advanced Encryption Standard

- Substitution / Permutation Network
- Key Length: 128 / 192 / 256 bits
- Round Number: 10 / 12 / 14
- Block Length: 128 bits
- Designed by Daemen and Rijmen
- Standardized by NIST in 2000
- Encryption scheme widely used and US communication «Top Secret» Information

AES Round

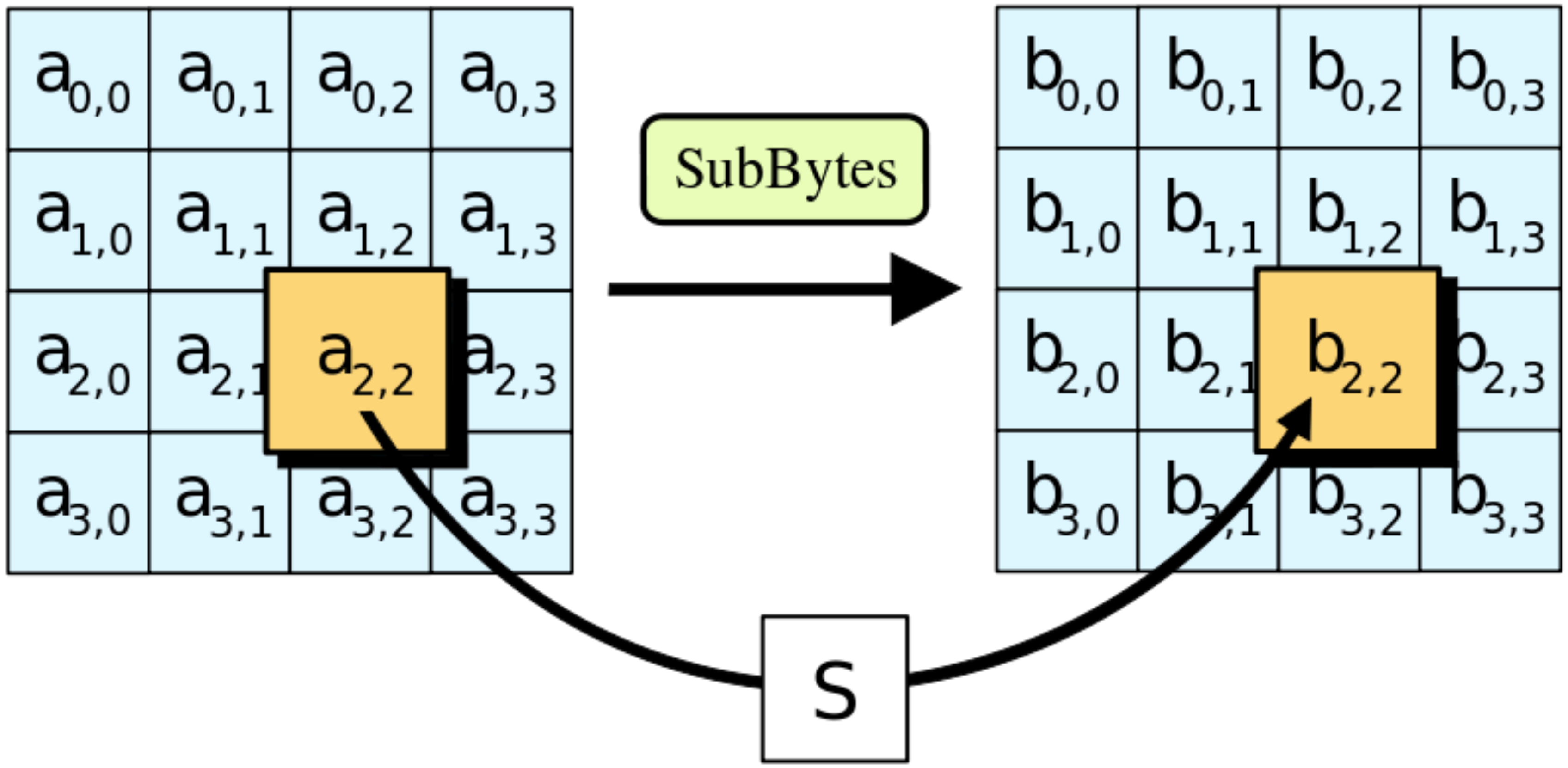


AES state

- 10 rounds = 9 full rounds + 1 final round (w/o MC)
- K0=K is first xored to the state (11 Round Keys)

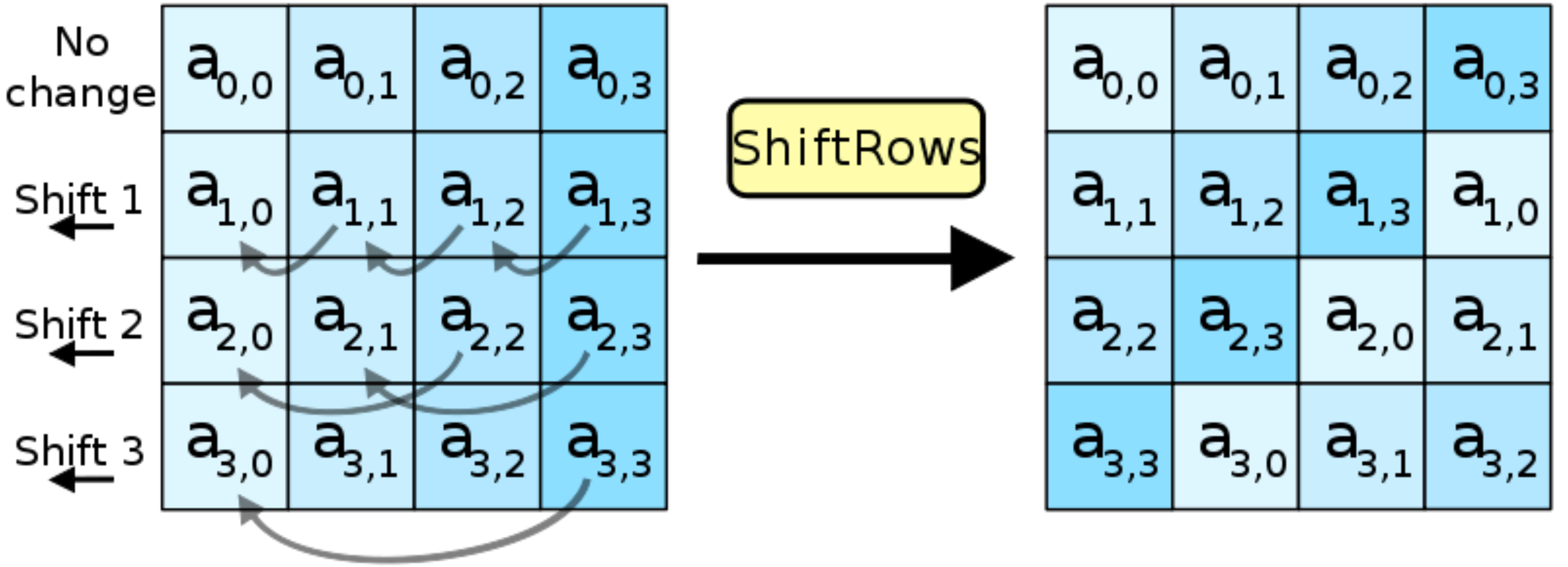
AES SubByte

- Non-linear function resistant against Diff/Lin attack : max DP(a,b)=2^{-6} and max LP(a,b)=2^{-5}
- S(x)=1/x in GF(256) followed by an affine function


$a_{0,0}$ $a_{0,1}$ $a_{0,2}$ $a_{0,3}$
$a_{1,0}$ $a_{1,1}$ $a_{1,2}$ $a_{1,3}$
$a_{2,0}$ $a_{2,1}$ $a_{2,2}$ $a_{2,3}$
$a_{3,0}$ $a_{3,1}$ $a_{3,2}$ $a_{3,3}$
SubBytes
$b_{0,0}$ $b_{0,1}$ $b_{0,2}$ $b_{0,3}$
$b_{1,0}$ $b_{1,1}$ $b_{1,2}$ $b_{1,3}$
$b_{2,0}$ $b_{2,1}$ $b_{2,2}$ $b_{2,3}$
$b_{3,0}$ $b_{3,1}$ $b_{3,2}$ $b_{3,3}$
S

ShiftRows

- Part of the Permutation layer of the SPN
- Responsible for the diffusion property



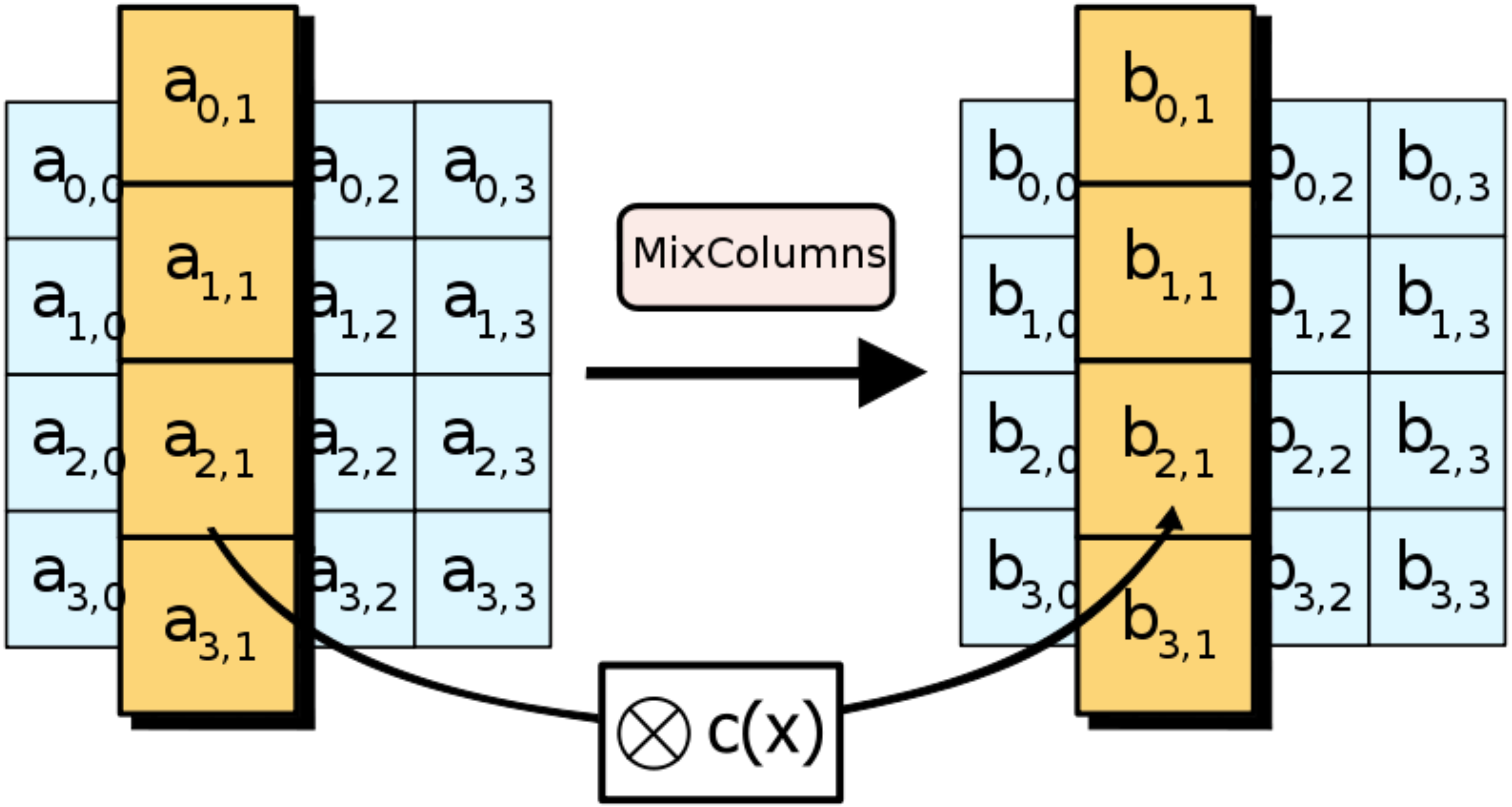
No change	$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$
Shift 1	$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$
Shift 2	$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
Shift 3	$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$

ShiftRows

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$
$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	$a_{1,0}$
$a_{2,2}$	$a_{2,3}$	$a_{2,0}$	$a_{2,1}$
$a_{3,3}$	$a_{3,0}$	$a_{3,1}$	$a_{3,2}$

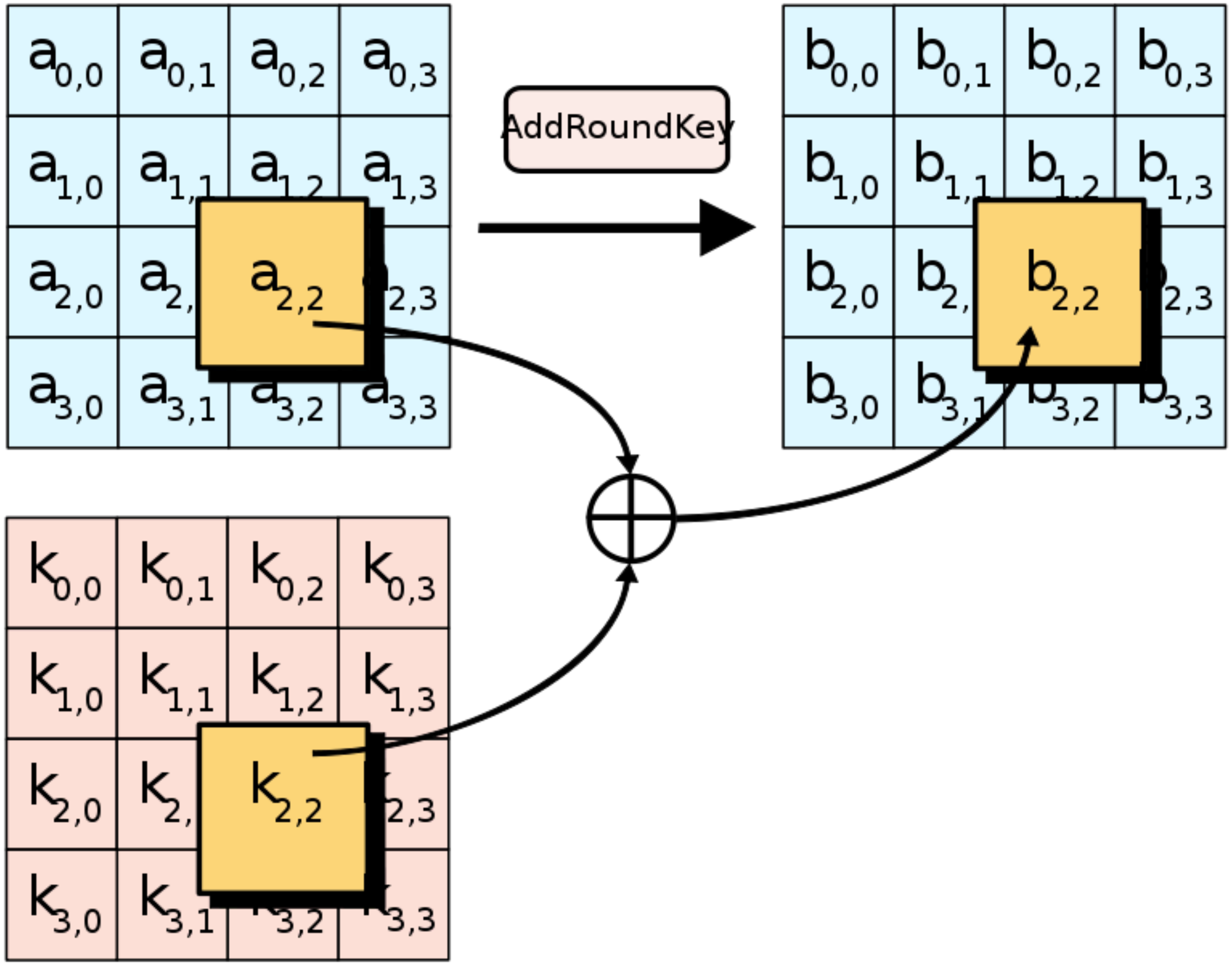
MixColumn

- Part of the Permutation Layer (diffusion)
- MDS matrix (Maximal Decoding Distance)

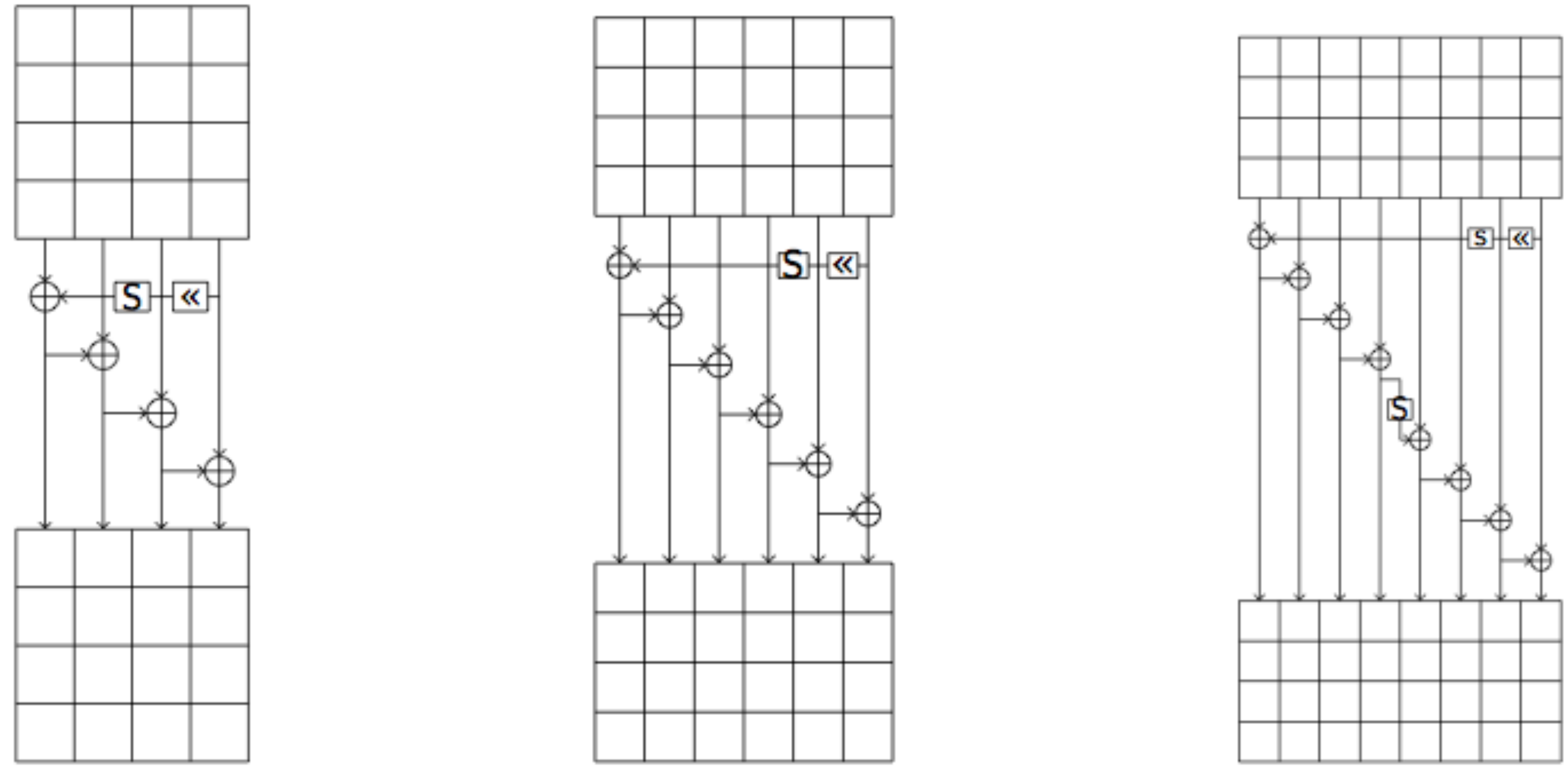

$a_{0,0}$ $a_{0,1}$ $a_{0,2}$ $a_{0,3}$
$a_{1,0}$ $a_{1,1}$ $a_{1,2}$ $a_{1,3}$
$a_{2,0}$ $a_{2,1}$ $a_{2,2}$ $a_{2,3}$
$a_{3,0}$ $a_{3,1}$ $a_{3,2}$ $a_{3,3}$
MixColumns
$\otimes c(x)$
$b_{0,0}$ $b_{0,1}$ $b_{0,2}$ $b_{0,3}$
$b_{1,0}$ $b_{1,1}$ $b_{1,2}$ $b_{1,3}$
$b_{2,0}$ $b_{2,1}$ $b_{2,2}$ $b_{2,3}$
$b_{3,0}$ $b_{3,1}$ $b_{3,2}$ $b_{3,3}$

AddRoundKey

- Add the Key simply by xoring the subkeys


$a_{0,0}$ $a_{0,1}$ $a_{0,2}$ $a_{0,3}$
$a_{1,0}$ $a_{1,1}$ $a_{1,2}$ $a_{1,3}$
$a_{2,0}$ $a_{2,1}$ $a_{2,2}$ $a_{2,3}$
$a_{3,0}$ $a_{3,1}$ $a_{3,2}$ $a_{3,3}$
$k_{0,0}$ $k_{0,1}$ $k_{0,2}$ $k_{0,3}$
$k_{1,0}$ $k_{1,1}$ $k_{1,2}$ $k_{1,3}$
$k_{2,0}$ $k_{2,1}$ $k_{2,2}$ $k_{2,3}$
$k_{3,0}$ $k_{3,1}$ $k_{3,2}$ $k_{3,3}$
AddRoundKey
$\oplus$
$b_{0,0}$ $b_{0,1}$ $b_{0,2}$ $b_{0,3}$
$b_{1,0}$ $b_{1,1}$ $b_{1,2}$ $b_{1,3}$
$b_{2,0}$ $b_{2,1}$ $b_{2,2}$ $b_{2,3}$
$b_{3,0}$ $b_{3,1}$ $b_{3,2}$ $b_{3,3}$

AES Key Schedule



(a) AES-128

(b) AES-192

(c) AES-256

AES Diffusion

- MDS Property: If I=(I_1,I_2,I_3,I_4) and O=(O_1,O_2,O_3,O_4) the input / output vectors of MC, then if we have a difference in x input bytes ($x \neq 0$), we have 5-x output differences
- Diffusion Property: After 2 rounds of AES, if we inject a difference in one byte, then every byte will have a difference
- Scheme

AES Diff / Lin attack

- Daemen and Rijmen show that after two AES rounds, the best linear and differential characteristic path has probability $<2^{-150}$!
- 4 rounds of AES is a perfect substitution
- http://homes.esat.kuleuven.be/~fvercaut/papers/diff_prob.pdf

Square Attack

- Attack discovered by Daemen, Knudsen and Rijmen on the block cipher Square (cipher designed two years before the AES very similar to Rijndael = AES)
- Distinguishing Property : If you encrypt a delta-set (set of 256 plaintexts where the byte takes the 256 different values), then 3 rounds later every byte of the state are balanced (the xor of the 256 ciphertexts will result in the null state)
- Square Attack: using meet-in-the-middle ideas, it is possible to invert (one, two rounds at the end), pass one more round in the beginning (using structure ideas)
  - Complexity: 2^{48} for 6 rounds ! (Fergusson et al. attacks)

Impossible Differential Attack

- Distinguishing property:There is a differential path on 4 rounds from an initial set of plaintext to a set of ciphertext that cannot be true
- Attack: Meet-in-the-middle (structure, and partially decrypting the ciphertext to check the relation)

Meet-in-the-Middle attack

- First attack : Demirci and Selçuk (FSE 2008)
- Distinguishing property: The function that associate one byte of the state (the others are constant) to one byte of the state (4 rounds later) depends only on 25 byte parameters
- Attack: If we tabulate these 2^{200} different functions (among the set of 2^{256*8}), then if we perform a meet-in-the-middle by partially decrypting the 256 ciphertexts (and using structure in the beginning), we have an attack on 7 rounds for AES-256
- Using, Time/Memory tradeoff, by partially storing the different functions, and increasing the time, it is possible to have an attack also for AES-192

DKS improvements

- Dunkelman, Keller and Shamir show that we can reduce the memory by storing only some functions which have a special differential property
- Differential Property: on 4 AES-rounds, it is possible to have a differential path with proba 2^{-120}, if we store only these functions, the table depends only on 16 byte parameters

- Derbez, Jean and myself have shown that the table depends only on 10 byte parameters (and show that it is optimal)

Side Channel attacks

- DPA attack :
  - Combine power trace (or EM trace) with some prediction of the some key bytes
- Timing attack :
  - Cache behaviour of the Sbox table allows to recover high order bits of the indices
- Fault attack :
  - inject some differences in the state and difference of the faulty and correct ciphertext allows to recover the key

DPA Attack

- Target the first xor between the plaintext block and the first round key (master key)
- According to the power trace on one hand, and a model of the leakage ($HW(K_0^i \oplus P_0^i)$) on the other hand, we can perform a correlation between these two value sets
- The higher correlation coefficients for the different values of K_0^i, will expose the correct key bytes
- http://people.rit.edu/kjm5923DPA_attacks_on_AES.pdf

Cache Attack

- Timing attack: Bernstein and Bonneau papers but do not explain where the timing difference comes from
- Cache monitoring: Osvik, Shamir Tromer 2006 allow to recover the high order bits of the indices of the table which are related to key bytes (easy for the first round)

Fault Attack

- Inject a byte difference 3 or 4 rounds before the end of the cipher
- According to the difference between the faulty ciphertext and the correct one, we can partially invert the cipher and discover the last subkey (equivalent to the knowledge of the first one)
- Scheme on 3 rounds

Countermeasures

- Masking ideas based on multiparty computation and secret sharing scheme ideas
- Share a value x as (r,r⊕x) so that if the adversary knows one of this share, he has no information on x (multiple shares if more efficient adversary)
- Masking linear operation is easy: L(x)=L(r)⊕L(r⊕x)
- Masking non-linear operation is the most difficult part:
    - Masking the inversion in GF(256) : express the power to 254 as multiplication and squaring (squaring is a linear op)
    - Masking the table: compute on-line T'(x)=T(r⊕x)⊕s and evaluate T'(r⊕x) to get a share of T(x)

Conclusion

- AES has been designed to resist against differential and linear attack
- Today, the best «academic» attack only works:
  - for 7 rounds of AES-128 (T=M=D$\approx 2^{100}$)
  - Related-key attacks on AES-192 and AES-256 for the full rounds (key schedule weaknesses)
- Today, the more efficient attacks in practice are based on side-channel ideas

Future

- Many new primitives (MAC, stream cipher, lightweight encryption scheme, hash functions) have been designed based on AES building block (SB, SR, MC)
- Breaking these primitives sometimes help to understand new properties of this cipher (cf. NIST SHA-3 competition)
- It is difficult to protect AES implementations against side channel attacks
- Performances of AES-GCM for authenticated cipher are not as efficient as expected for very high speed network
- New CAESAR competition to define authenticated cipher

LED and Dinur's talk

- LED uses 4 AES rounds which act as a big substitution table
- Consequently, the permutation (which guarantees the diffusion) is not longer useful
- 4 AES rounds w/o constant are however very weak (internal differential property)