Conditional Linear Cryptanalysis

Stav Perle

Joint work with Eli Biham

Technion – Israel Institute of Technology

Cryptanalytic Techniques

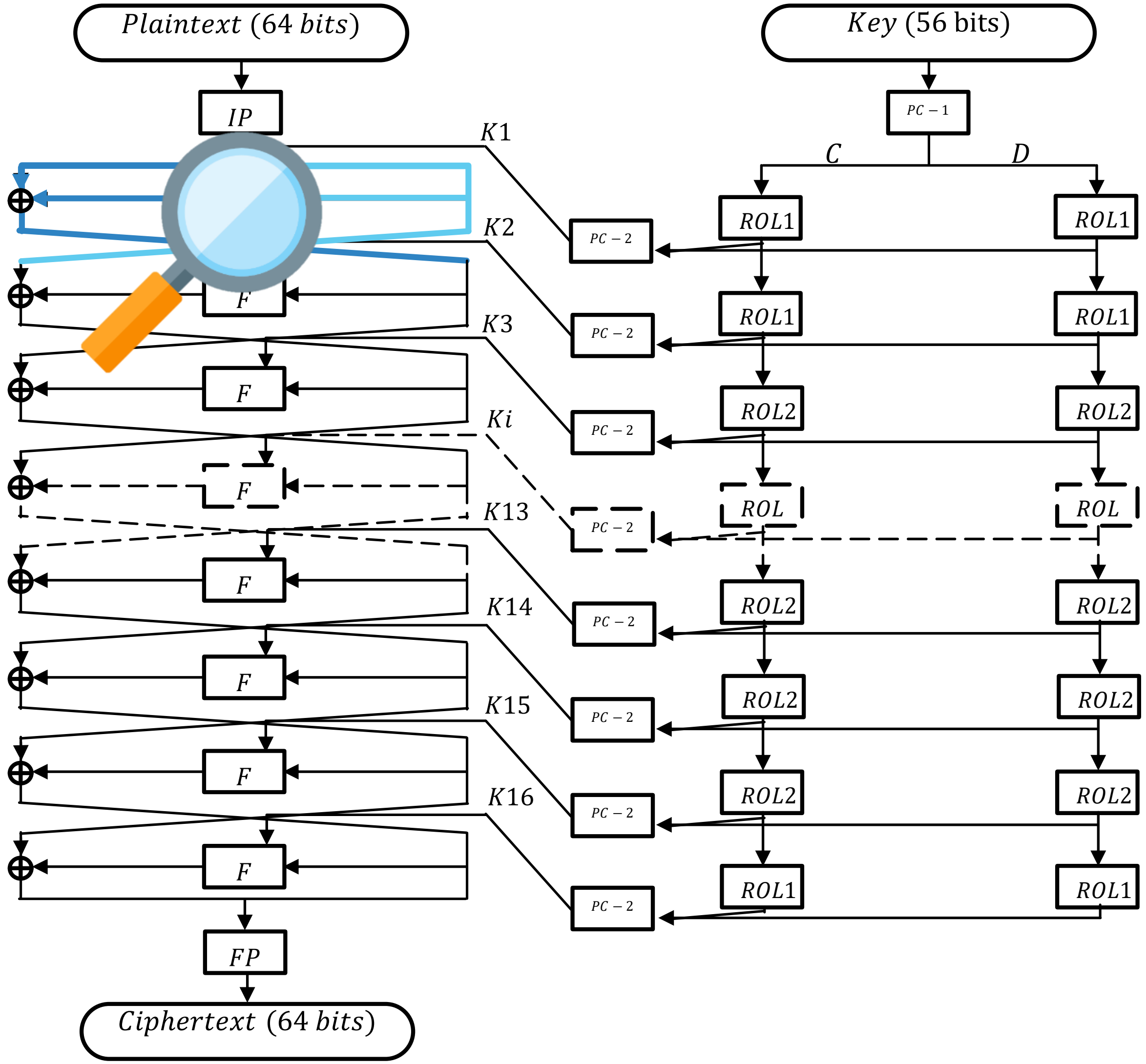
- Differential and linear cryptanalysis are two major generic techniques for assessing the strength and vulnerabilities of block ciphers
- These techniques have various extensions which can improve their success in various cases
- Along with Davies' attack, they are the best attacks against the Data Encryption Standard (DES)

Technique	Complexity
Differential Cryptanalysis	2^{47}
Linear Cryptanalysis	2^{43}
Improved Davies' Attack	2^{50}

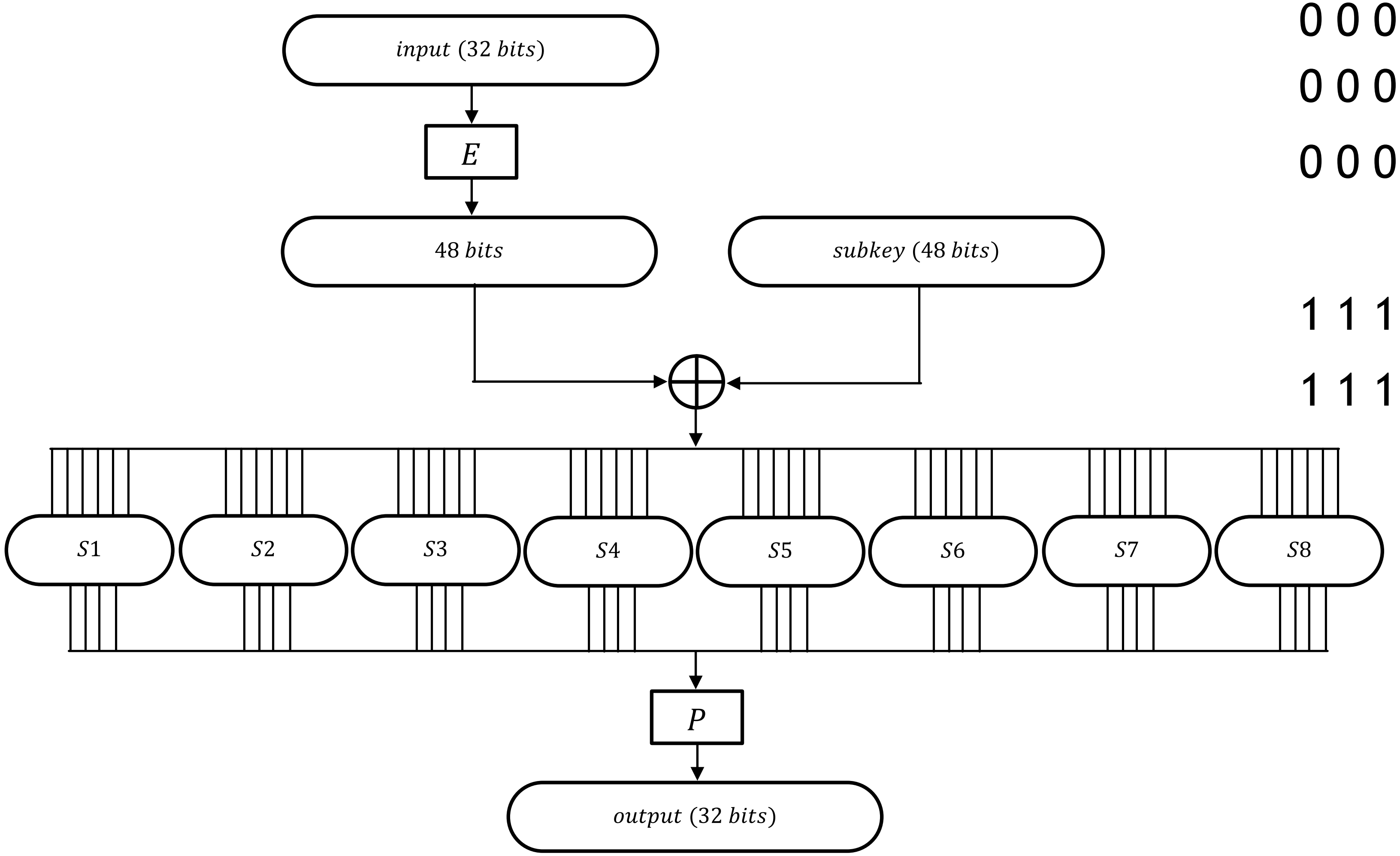
- Today, we will show a new extension that reduces this complexity further

Conditional Linear Cryptanalysis	$\leq 2^{42}$

DES – Example of a Block Cipher



The F Function of DES



0 0 0 0 0 0 → 0 0 1 0

0 0 0 0 0 1 → 1 1 1 0

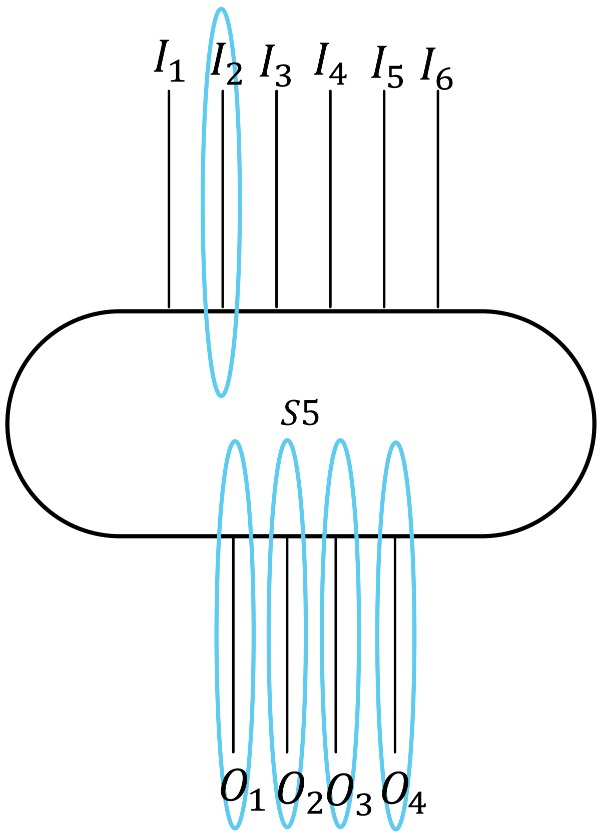
0 0 0 0 1 0 → 1 1 0 0

…

1 1 1 1 1 0 → 0 1 0 0

1 1 1 1 1 1 → 0 0 1 0

The Best Non-Trivial Approximation of S5



- It approximates the second bit of input to the XOR of the four output bits
  - In 12 cases: $I_2 \oplus O_1 \oplus O_2 \oplus O_3 \oplus O_4 = 0$
  - In 52 cases: $I_2 \oplus O_1 \oplus O_2 \oplus O_3 \oplus O_4 = 1$

Linear Cryptanalysis

- Linear Cryptanalysis uses statistical approximations that approximate parity of subsets of bits of the plaintext, ciphertext, and the subkeys
  - E.g., (second bit of the plaintext) XOR (fifth bit of ciphertext) XOR (keys bits) = 0

$$P_1\ P_2\ P_3\ P_4\ P_5\ \ldots\ P_{60}\ P_{61}\ P_{62}\ P_{63}\ P_{64} \qquad C_1\ C_2\ C_3\ C_4\ C_5\ \ldots\ C_{60}\ C_{61}\ C_{62}\ C_{63}\ C_{64} \qquad K_1\ K_2\ K_3\ K_4\ K_5\ \ldots\ K_{60}\ K_{61}\ K_{62}\ K_{63}\ K_{64}$$

$$P \qquad\qquad C \qquad\qquad K$$

- Each approximation has a probability, p, to hold
  - Which is the fraction of plaintexts whose encryption follow the approximation
  - In random cases, the probability is expected to be ½, or close to ½

Linear Cryptanalysis

- The ability if distinguish whether an approximation holds highly depends on the distance of the probability from ½
- Let the bias be $\varepsilon = p - \frac{1}{2}$
  - Range: -½ to +½
  - The higher the (absolute value of the) bias, the easier to distinguish
  - $\varepsilon = 0$ means that the approximation is mostly useless

Linear Approximations

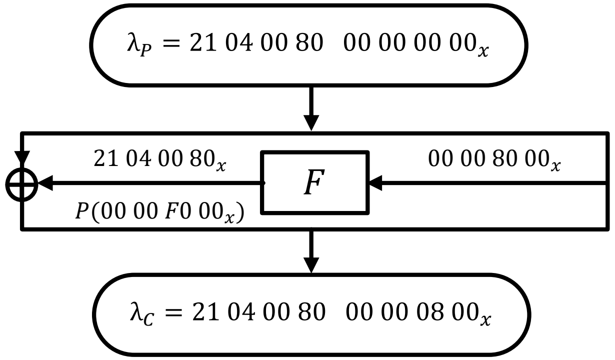
- A linear approximation is a tuple $(\lambda_P, \lambda_C, \lambda_K)$
  - λ_P is a subset of bits of the plaintext
  - λ_C is a subset of bits of the ciphertext
  - λ_K is a subset of bits of the key (or the subkeys)

0 0 0 1 1 …1 0 0 1 0	0 1 0 1 1 …0 0 0 0 1	0 1 1 1 1 …0 0 1 1 0
λ_P	λ_C	λ_K
$P_1\ P_2\ P_3\ P_4\ P_5\ \dots\ P_{60}\ P_{61}\ P_{62}\ P_{63}\ P_{64}$	$C_1\ C_2\ C_3\ C_4\ C_5\ \dots\ C_{60}\ C_{61}\ C_{62}\ C_{63}\ C_{64}$	$K_1\ K_2\ K_3\ K_4\ K_5\ \dots\ K_{60}\ K_{61}\ K_{62}\ K_{63}\ K_{64}$
P	C	K

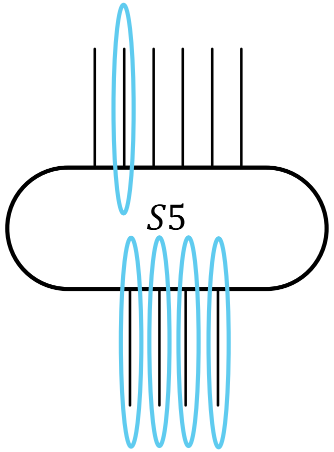
- The **probability of the approximation** is the probability that $\boldsymbol{P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0}$

Linear Approximations – Examples

- The best non-trivial approximation

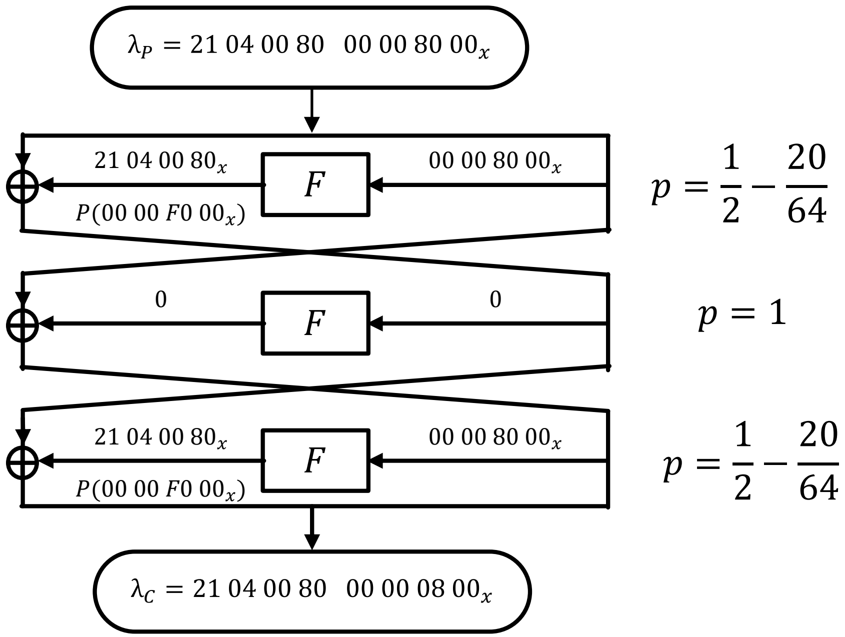


- It approximates the second bit of input to S5 to the XOR of the four output bits of S5
  - In 12 cases: $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0$
  - In 52 cases: $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 1$
  - $\frac{1}{2} + \varepsilon = \frac{12}{64}$
  - $\varepsilon = \frac{-20}{64}$



Linear Approximations – Examples

- This linear approximation has probability $\frac{1}{2}+2(\frac{-20}{64})^2=\frac{1}{2}+\frac{25}{128}$



Algorithm 1

- Given λ=(λ_P,λ_C,λ_K), we know that $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0$ holds with probability $p = \frac{1}{2} + \varepsilon$
- Given plaintext and the corresponding ciphertext, we can calculate the value of $P\lambda_P \oplus C\lambda_C$

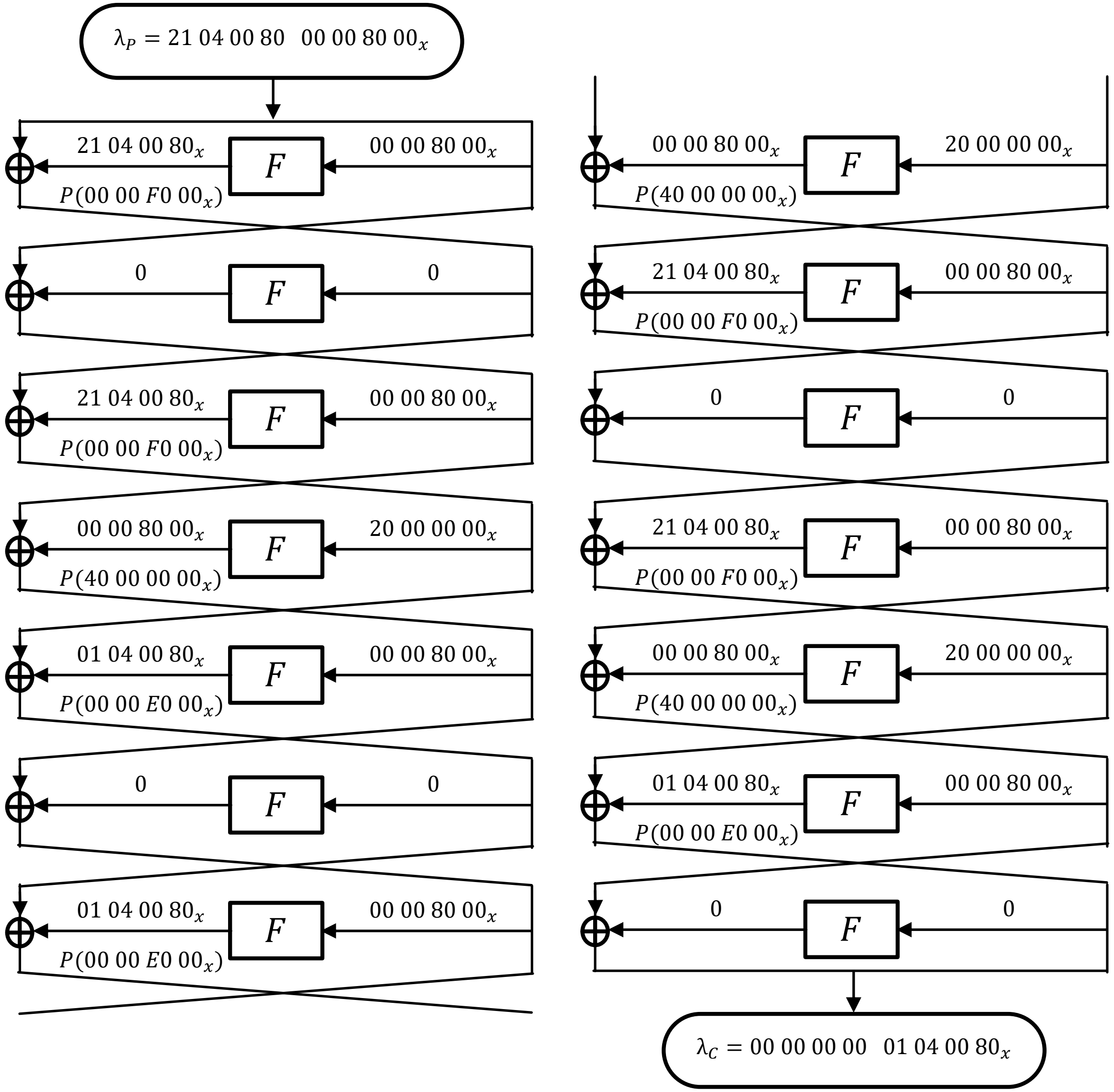
Algorithm 1

- Given λ=(λ_P,λ_C,λ_K), $\varepsilon(\lambda)$, and N plaintexts and their ciphertexts, the algorithm counts the number M of plaintexts satisfying
$$P\lambda_P \oplus C\lambda_C = 0$$
  - Recall that $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0$ holds with probability $p = \frac{1}{2} + \varepsilon$
- The algorithm guesses that the parity of the key bits $K\lambda_K$ is

	$\varepsilon > 0$	$\varepsilon < 0$
$M > \frac{N}{2}$	0	1
$M < \frac{N}{2}$	1	0

- This algorithm finds only one parity bit of the key
- The success rate of the algorithm grows as the number of plaintexts N increases, and as the value of $|\varepsilon|$ increases
- For a high probability of success, $N \approx \frac{1}{\varepsilon^2}$ or higher

Matsui's Best Approximation (14 rounds)



Linear Cryptanalysis of the Full DES

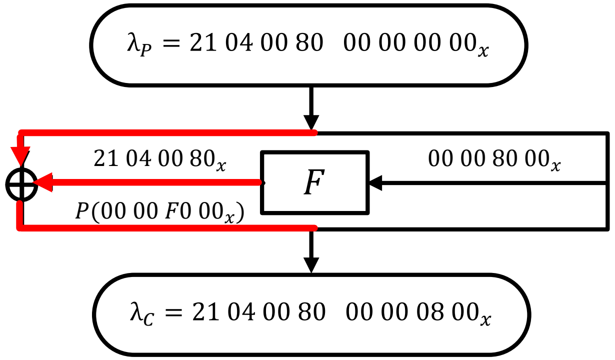
- Matsui uses the best 14-round approximation with probability $\approx \frac{1}{2} - 2^{-20.75}$
- The attack requires about 2^{43} known plaintexts

Conditional Linear Cryptanalysis

- Using conditions to discard data that reduces that bias
  - So the bias of the remaining data increase or decrease
- Conditions can be by any observable data available to the cryptanalyst
  - Plaintexts, ciphertexts, and formulae on them
- In Feistel ciphers we can compute various linear combinations of internal bits directly from the plaintext and ciphertext

A Case of Single Round

- The best non-trivial approximation



- It approximates the second bit of input to S5 to the XOR of the four output bits of S5
  - Probability $\frac{1}{2} + \frac{-20}{64}$
  - I.e., 12 cases with equality (parity 0 of the 5 bits), 52 cases with inequality (parity 1 of the 5 bits)

A Case of Single Round

- Conditioning on all the four output bits of S5 (16 cases) we get

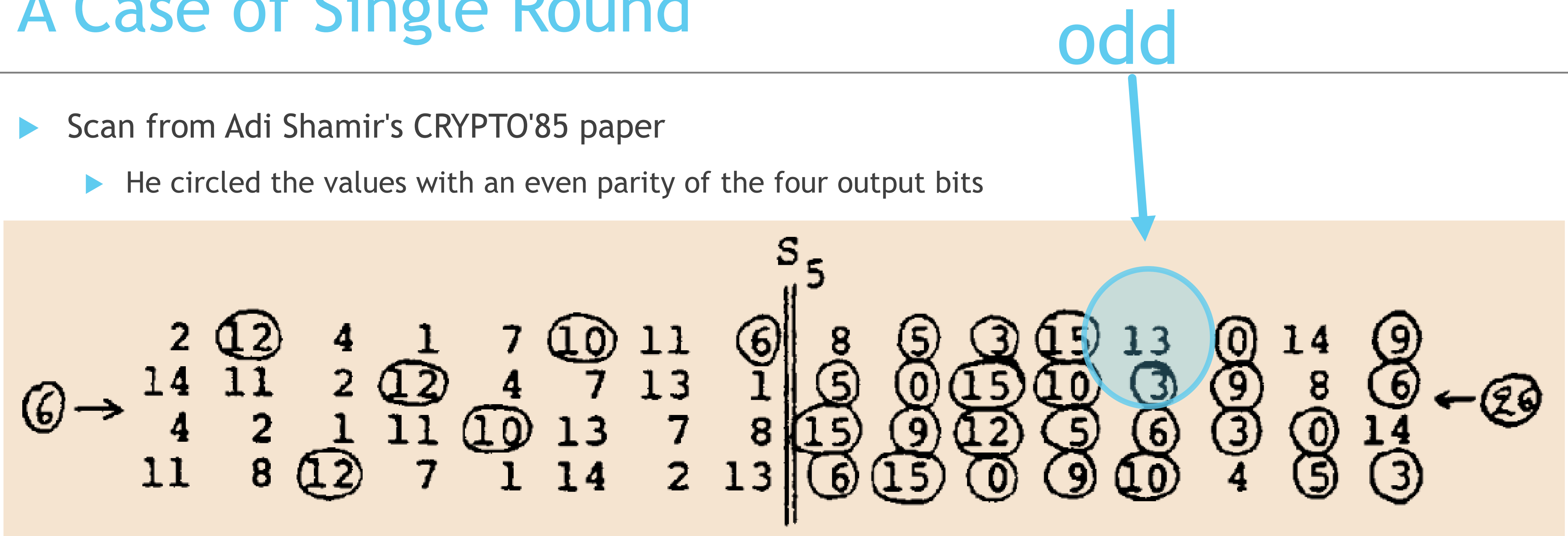
0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0	0	0	0	1	0	1	0	2	0	2	0	3	1	2	0
-0.5	-0.5	-0.5	-0.5	-0.25	-0.5	-0.25	-0.5	0	-0.5	0	-0.5	0.25	-0.25	0	-0.5

- Consider a condition on the LSB of the four output bits of S5 (a single bit)

Condition	$P\lambda_P \oplus C\lambda_C \oplus K\lambda_K =$		Bias
	0	1	
none	12	52	-20/64
LSB=0	11	21	-5/32=-10/64
LSB=1	1	31	-15/32=-30/64

A Case of Single Round

- Scan from Adi Shamir's CRYPTO'85 paper
  - He circled the values with an even parity of the four output bits



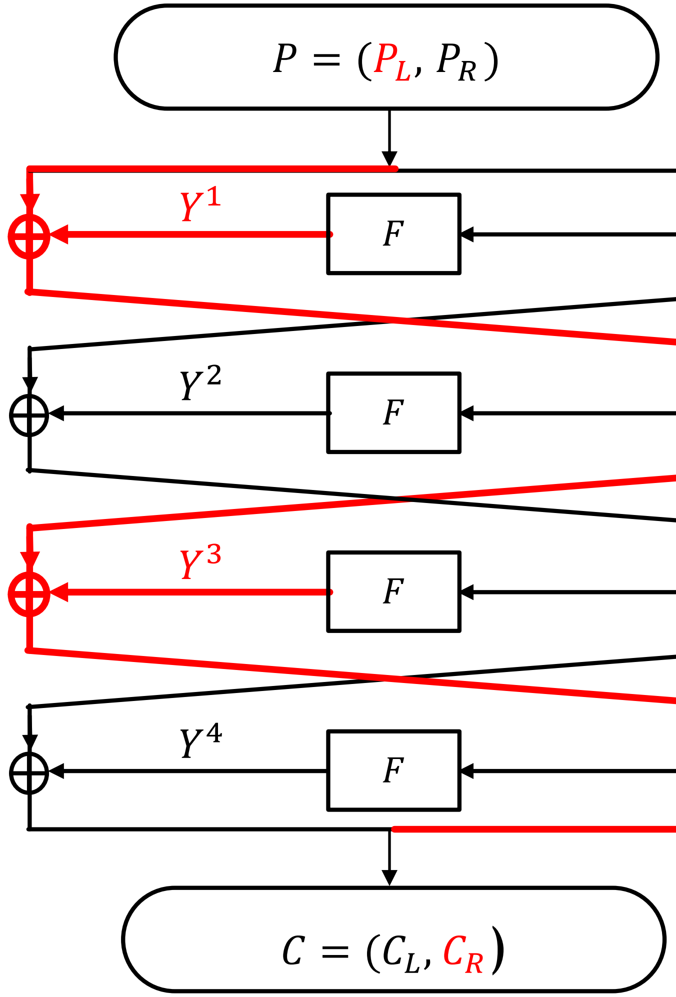
2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3

The 12 VS. the 52

- ➔ 1 vs. 31 for LSB=1, and 11 vs. 21 for LSB=0

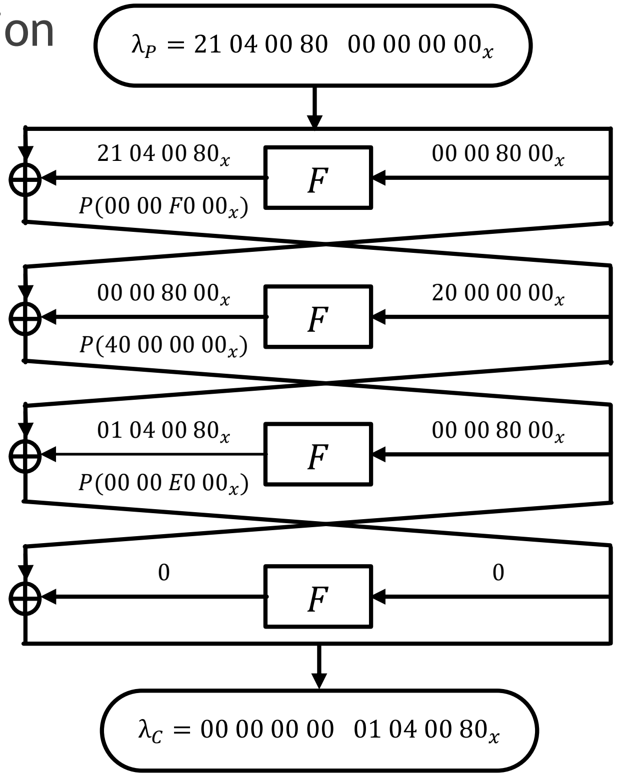
Conditional Linear Cryptanalysis

- We can condition on the XOR of plaintext and ciphertext bits
  - even more than one bit at a time
- For example, on $P_L \oplus C_R = \oplus_{\mathrm{r}\ is\ odd} Y^{\mathrm{r}}$
  - which is the XOR of the output of F in all odd rounds
- Consider any one of these bits as a linear approximation
  - E.g., $P_{L,17} \oplus C_{R,17}$=0
    - Equivalent to $Y^1_{17} \oplus Y^3_{17}$=0
  - Such approximations are expected to have bias 0
- But they are very useful as conditions to other approximations



A Four-Round Example

- Consider four successive rounds taken from Matsui's best linear approximation
- This approximation uses three active S boxes:
  - S5 on the first and third rounds, and
  - S1 on the fourth round
- Both odd rounds have the same active S box



A Four-Round Example

- Conditioning on all the four XOR output bits of S5 (16 cases) we get

0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
0.008	0	0.008	0	0.009	0	0.009	0	0.014	0	0.014	0	0.015	0	0.015	0

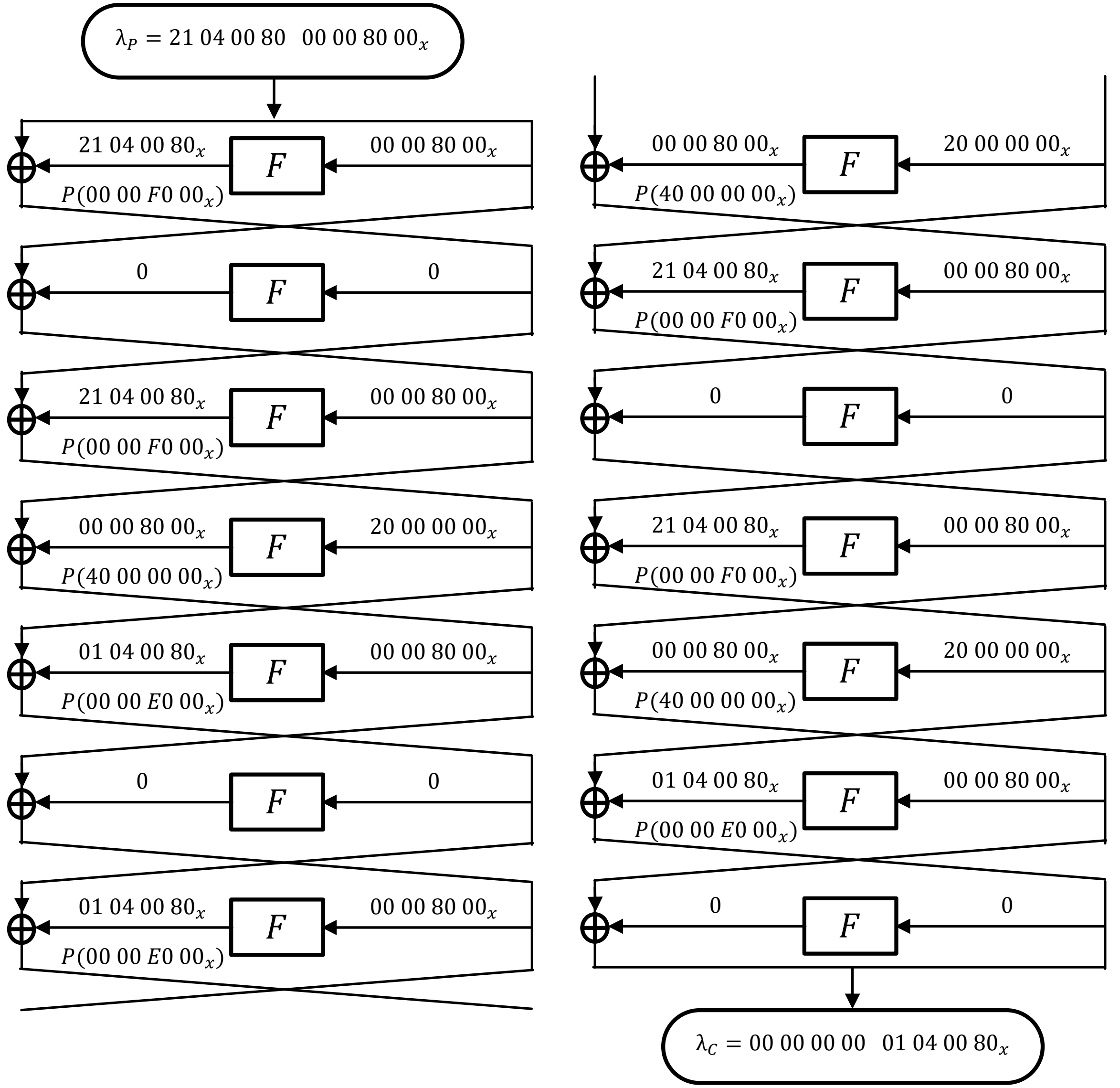
  - Notice that this condition is based on the XOR of both odd rounds
    - Not just on one of them
- For applying Matsui's Algorithm1 with our observation, we discard half of the known plaintexts, and use only the plaintexts in which the XOR of the LSB bits of S5 is zero
  - Their average bias is 0.0115
  - While the bias over all cases is 0.0057
- Using only these plaintexts increases the bias by a factor of two

A Four-Round Example

- We need a quarter of the data
  - Compared to a regular linear attack with the same approximation
  - But this is after we discard half of the given data that fails the condition
- We need half of the original data
  - We discard half of it, and get the required quarter

Conditional Linear Cryptanalysis of the Full DES

λ_1



Conditional Linear Cryptanalysis of the Full DES

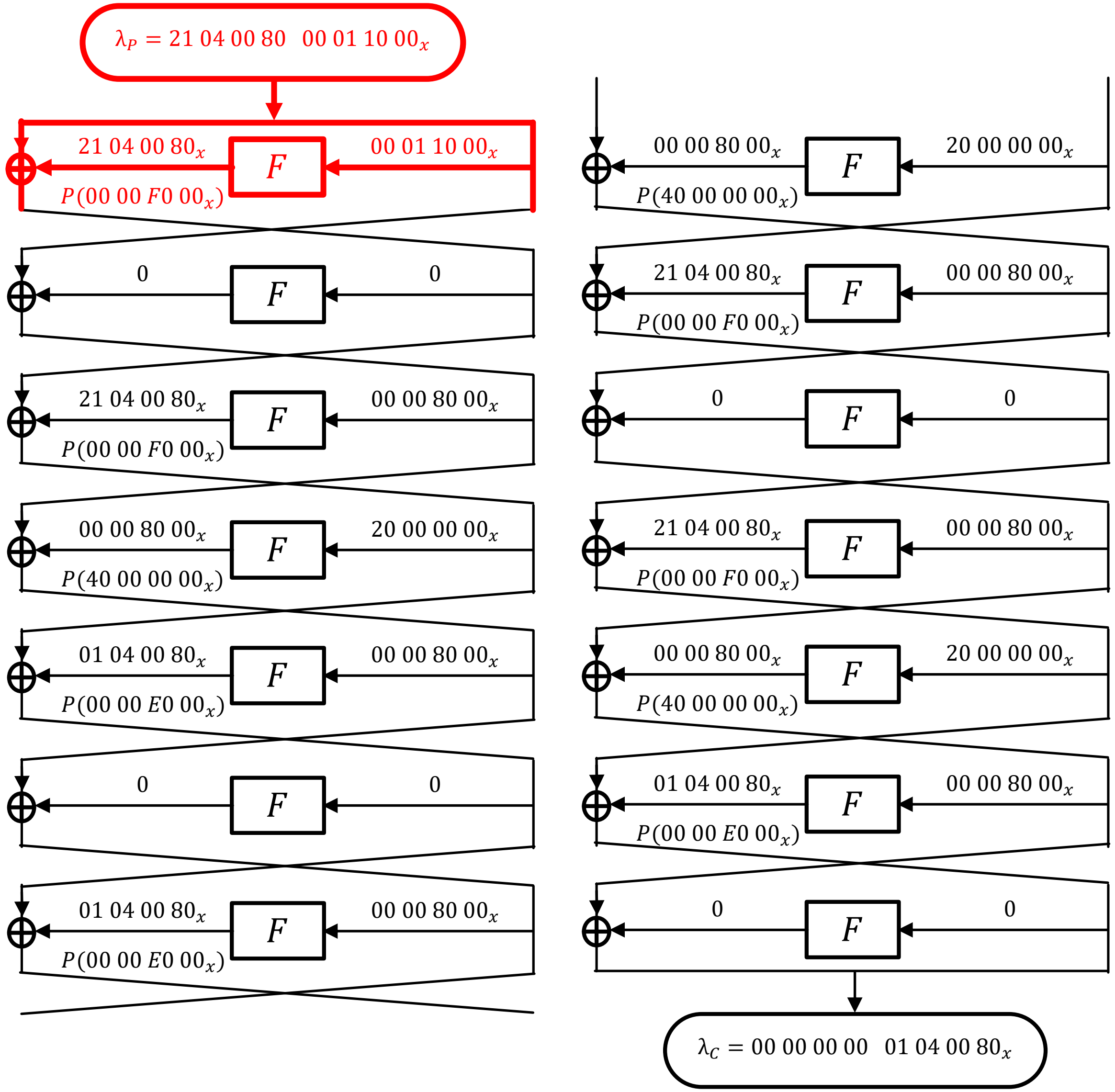
- Conditioning λ_1 on all the four XOR output bits of S5 (16 cases) we get

0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$

Condition	Bias
none	$\sim -2^{-20.75}$
XOR LSB=0	$\sim -2^{-21.74}$
XOR LSB=1	$\sim -2^{-20.16}$

Conditional Linear Cryptanalysis of the Full DES

λ_1
λ_2



Conditional Linear Cryptanalysis of the Full DES

- Conditioning λ_1 on all the four XOR output bits of S5 (16 cases) we get

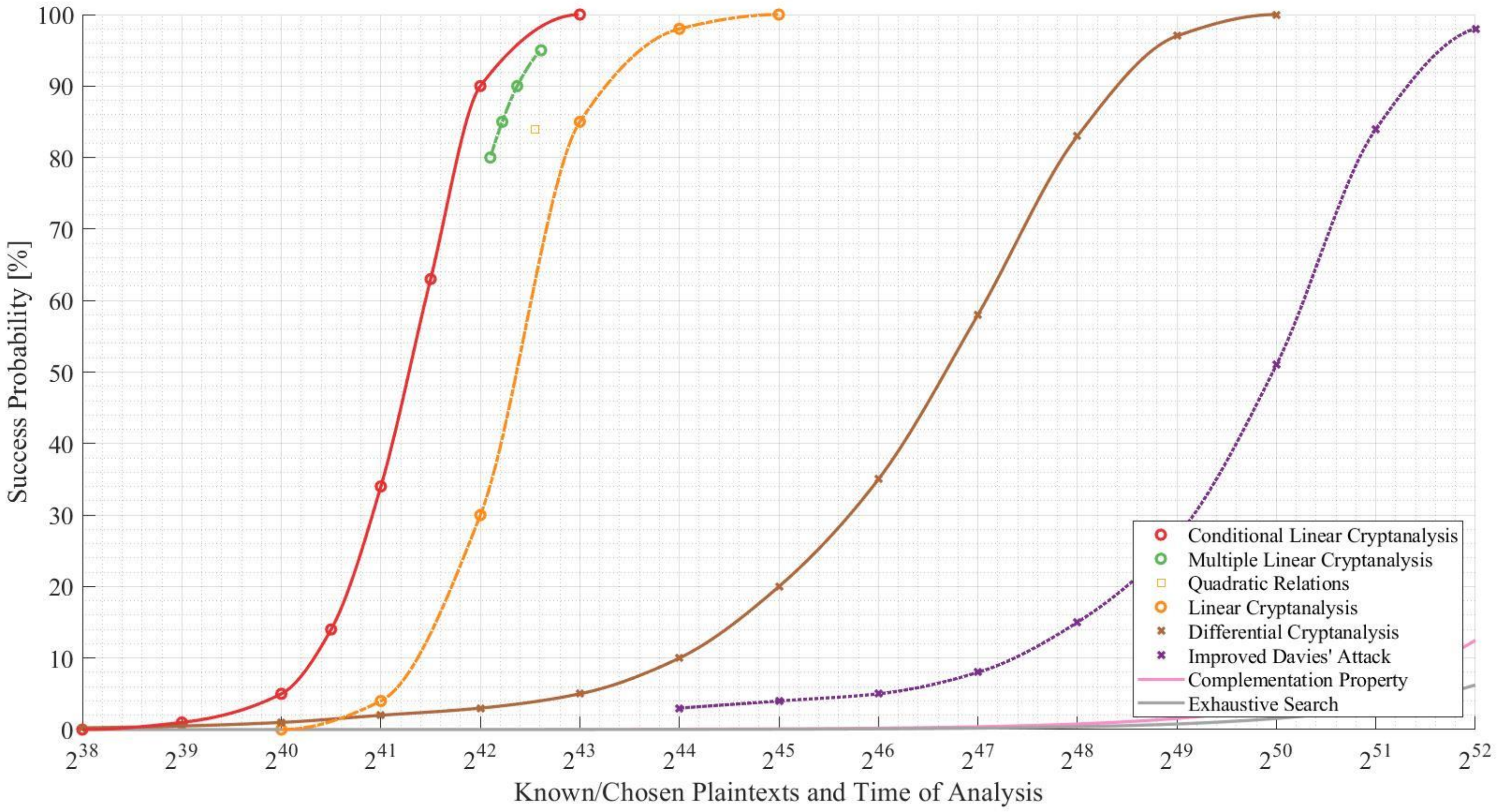
0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.77}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$	$-2^{-21.71}$	$-2^{-20.16}$

- Conditioning λ_2 on all the four XOR output bits of S5 (16 cases) we get

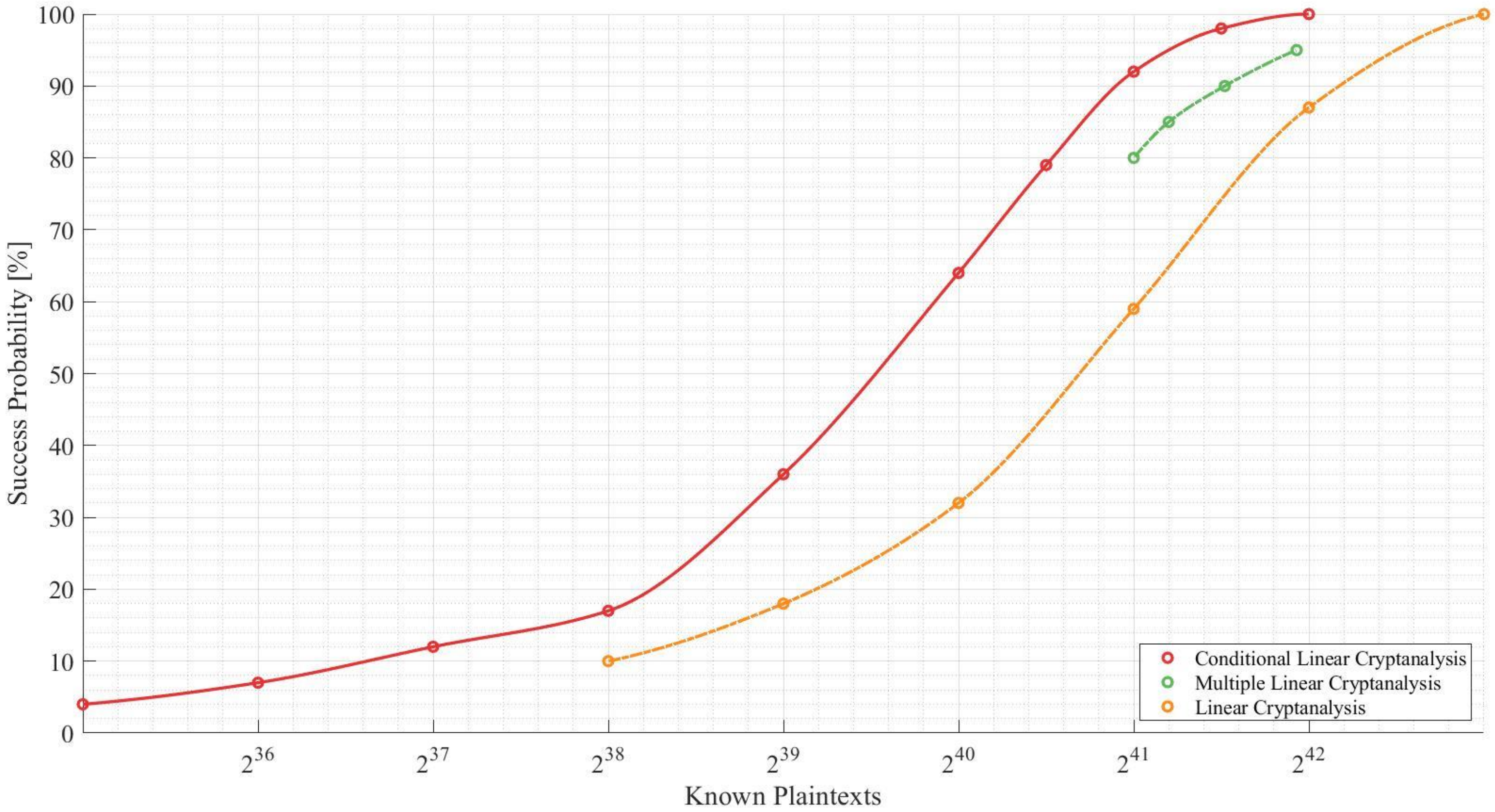
0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
$-2^{-20.26}$	$2^{-23.13}$	$-2^{-20.26}$	$2^{-23.13}$	$-2^{-20.26}$	$2^{-23.13}$	$-2^{-20.26}$	$2^{-23.13}$	$-2^{-20.26}$	2^{-23}	$-2^{-20.26}$	2^{-23}	$-2^{-20.26}$	2^{-23}	$-2^{-20.26}$	2^{-23}

Condition	Bias	
	λ_1	λ_2
none	$\sim -2^{-20.75}$	$\sim -2^{-21.48}$
LSB=0	$\sim -2^{-21.74}$	$\sim -2^{-20.26}$
LSB=1	$\sim -2^{-20.16}$	$\sim 2^{-23.06}$

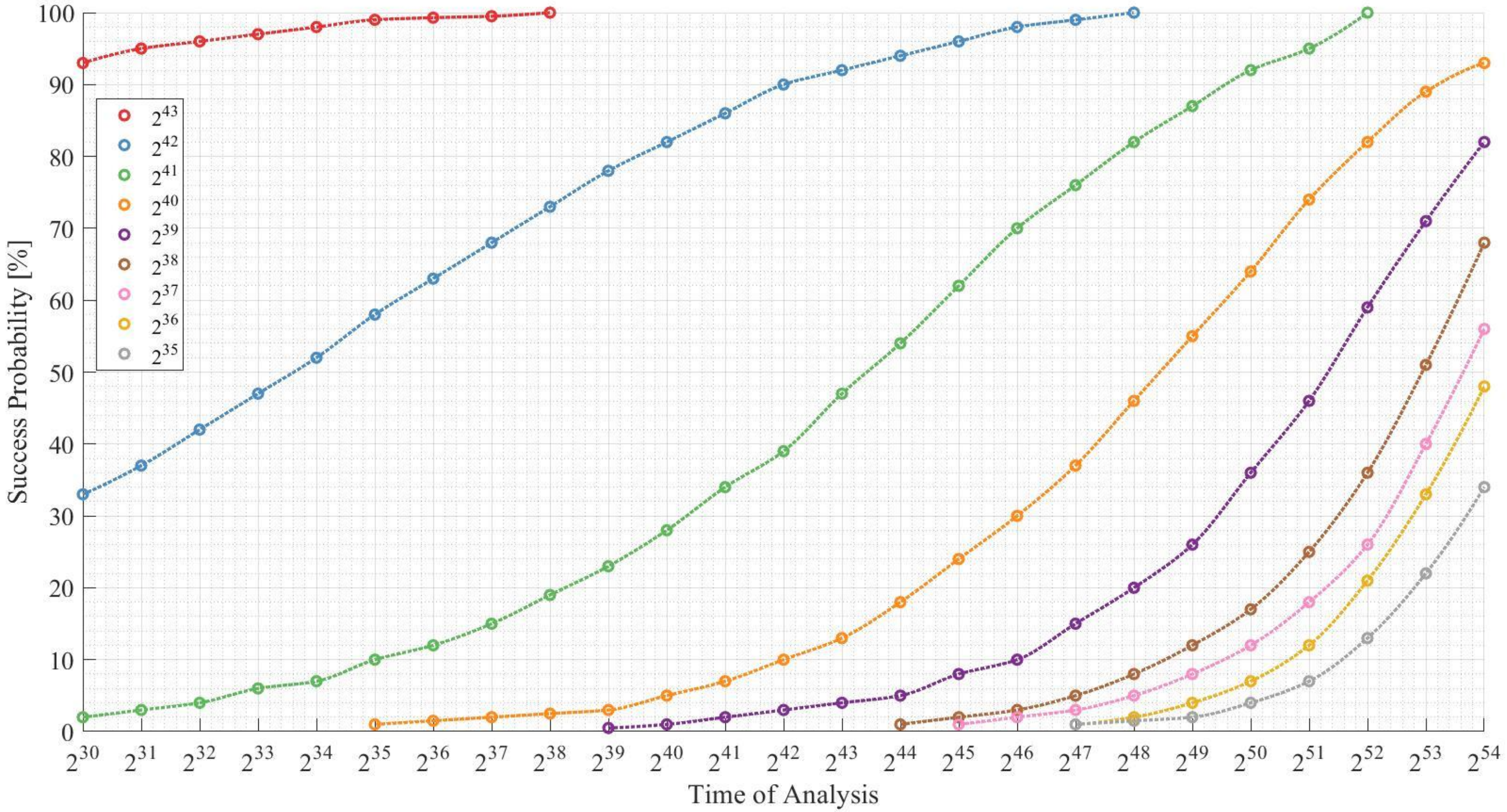
Success Probability by Complexity (#Ps&Time)



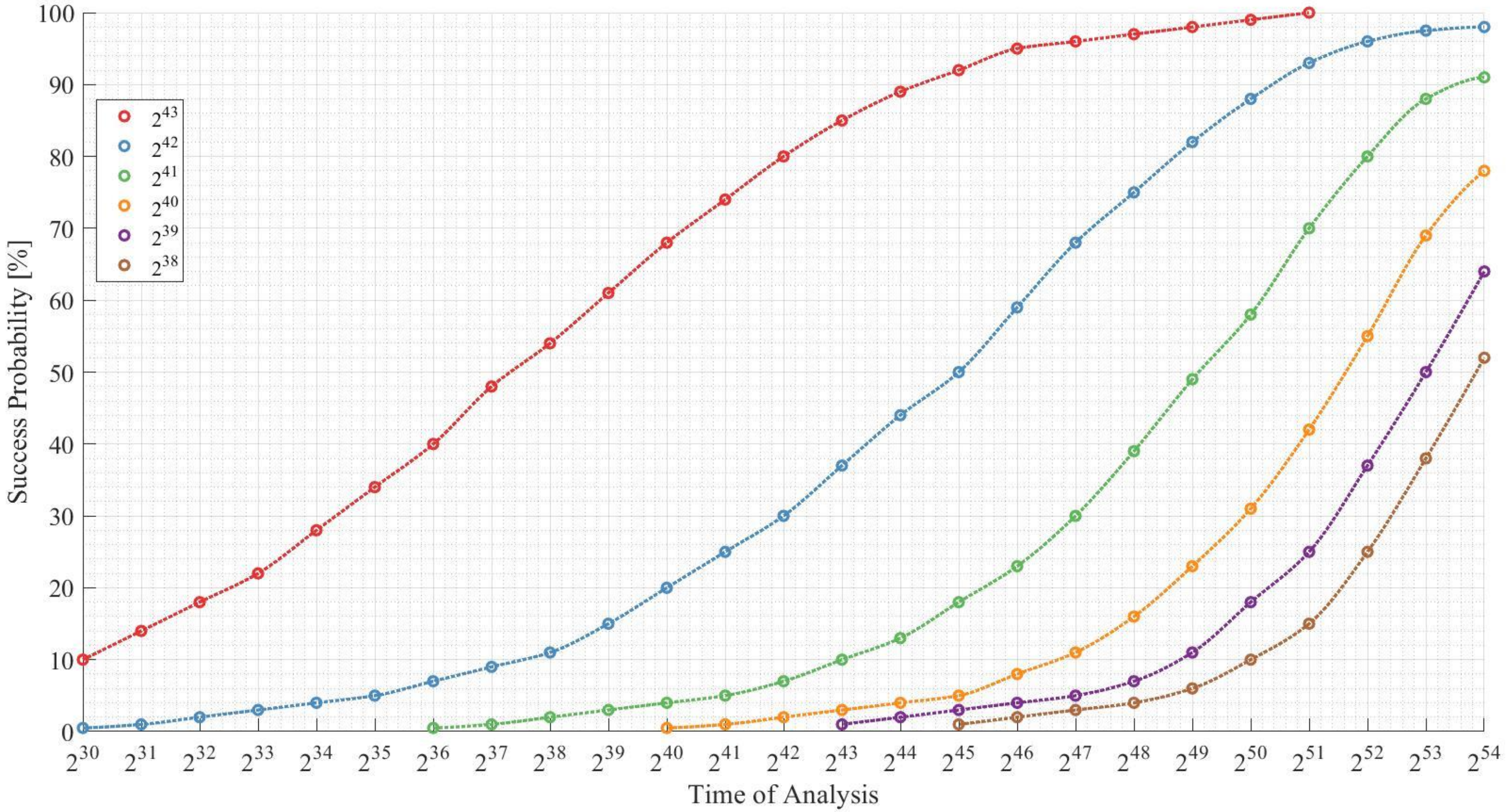
Known Plaintexts when Time is Fixed to 2^{50}



Success Probability by Time for Various #KPs



Matsui’s Success Prob. by Time for Various #KPs



The End