UNIVERSITY OF BERGEN

# Linear dependencies between non-uniform distributions in DES

by

Stian Fauskanger

Supervisor: Prof. Igor A. Semaev

Thesis for the degree Master of Science in Informatics

May 2014

in the
Faculty of Mathematics and Natural Sciences
Department of Informatics
Selmer Centre

## Abstract

Davies and Murphy[1] explained some non-uniform distributions of the output from pairs and triplets of S-boxes in DES, and how they are completely dependent on some key bits. There are linear dependencies between these distributions. In this thesis, we describe these linear dependencies. We also describe linear dependencies between the distributions of the output from three adjacent S-boxes after $n$ rounds in DES. We have found all linear dependencies between the distributions of the output from 5 of the S-box triplets in full DES. The dependencies originates from properties common to all S-boxes in DES.

# *Acknowledgements*

The author would like to thank Prof. Igor A. Semaev for supervising this project. Semaev had already found linear relations between the rows in the tables in Appendix A and realised that the output distributions of the triples of DES adjacent S-boxes are linearly dependent. He also helped with formalizing ideas and contributed with good discussions and advice.

The author would also like to thank Researcher Håvard Raddum from Simula@UiB, for his assistance; the Selmer Center, for an office space; and all the people at the Selmer Center for a good social environment.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

In this thesis we study The Data Encryption Standard (DES). Davies and Murphy[1] found some statistical properties of the S-boxes in DES that lead to non-uniform distributions on fixed output-bits from the round function. The distributions are completely determined by fixed key bits. As a consequence, there is a correlation between fixed plaintext/ciphertext-bits and key bits.

We have found linear dependencies between those non-uniform distributions. All linear dependencies between the distributions of the output from pairs and triplets of adjacent S-boxes are found. All dependencies between the distributions of the XOR of all such outputs in full DES are found for all pairs and for 5 out of the 8 triplets. The triplet with S-box 4, 5 and 6 is found to have an abnormal large amount of linear dependencies between the distributions. We have not managed to explain this abnormality at this time.

In an attempt to compute all linear dependencies between the distributions of the output from all S-boxes in 1-round DES, a modified version of DES (denoted by QDES) is studied. QDES is equal to DES, but with fewer S-boxes and correspondingly smaller block size. All dependencies in QDES with up to 4 S-boxes are found. A lower bound on the number of dependencies for QDES with up to 7 S-boxes is found. The complexity for computing linear dependencies between the distributions for QDES with 8 S-boxes (full DES) is too high.

# Chapter 2

# Background

Part of my work has been to do a survey on Matsui's work on linear cryptanalysis[2] of DES and on Davies and Murphy's analysis[1] of DES. These are described in Chapters 3 and 4. Linear dependencies between distributions are discussed in Chapter 5. This chapter will give a short introduction to The Data Encryption Standard, explain the notations used in this thesis and briefly present linear cryptanalysis of DES and Davies and Murphy's attack.

## 2.1 The Data Encryption Standard

The Data Encryption Standard[3] (DES) is a symmetric block cipher standardized by the National Bureau of Standards (now known as National Institute of Standards and Technology) in 1979. It has block size of 64 bits, and a 56-bit key. It is a 16 round Feistel network, which is depicted in Figure 2.1.

DES uses a key-scheduling algorithm to produce 16 48-bit round keys, from the 56-bit key. The algorithm permutes the 56 bits and selects 48 of them for each round in such a way that each of the 56 bits is used in approximately 14 of the 16 rounds.

Encryption in DES is done by first applying an initial permutation (IP) to the plain-text. The resulting string is divided into two 32-bit strings, $L_0$ and $R_0$. The cipher-text is the inverse of the initial permutation (IP$^{-1}$) applied to $R_{16}L_{16}$, which is computed by 16 iterations of this recursive equation:

$$L_{i+1} = R_i \qquad \text{and} \qquad R_{i+1} = L_i \oplus F(R_i, K_{i+1}) \qquad \text{for} \qquad i = 0, 1, ..., 15.$$

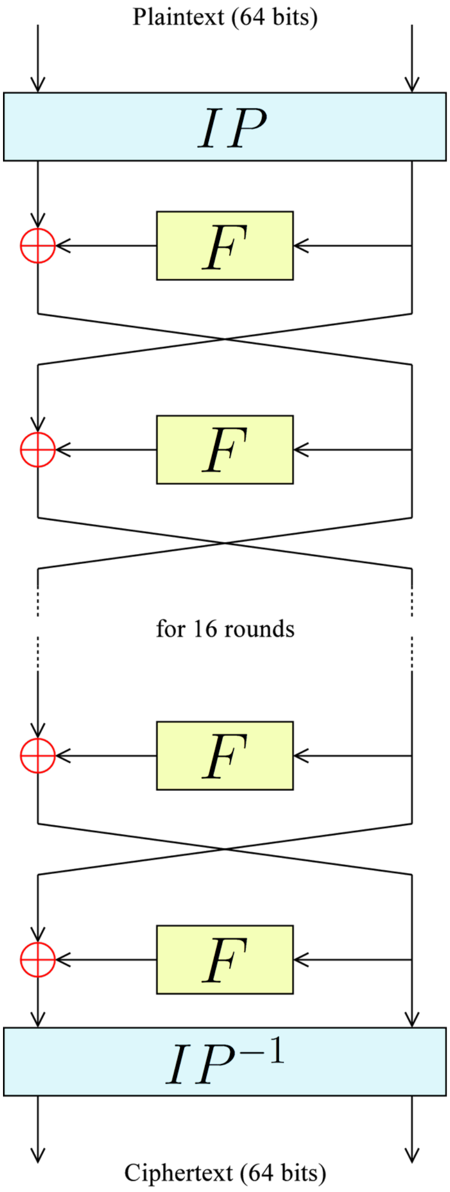


FIGURE 2.1: Feistel network

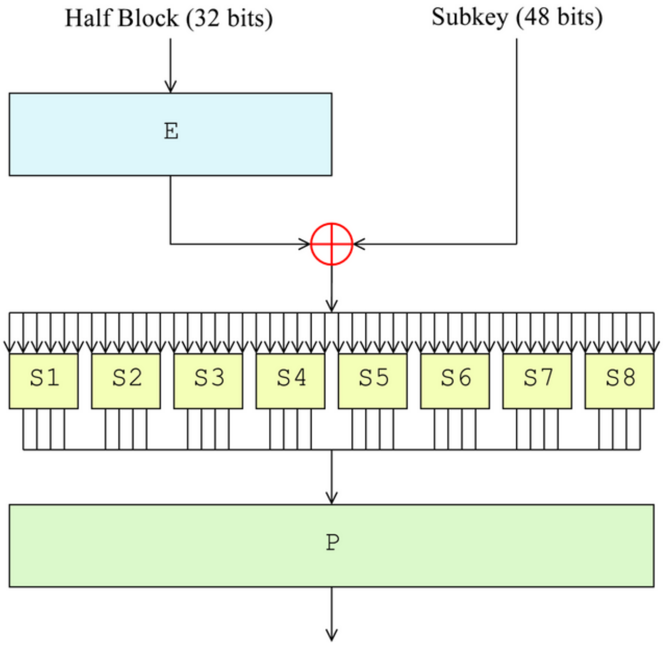


FIGURE 2.2: Round function in DES

(Source: wikipedia.org)

The round function, $F(R_i, K_{i+1})$, depicted in Figure 2.2, is built using these three building blocks: an expansion function $E$; 8 different substitutions S-boxes $S_1$ to $S_8$ and a permutation $P$. $F(R_i, K_{i+1})$ takes a 32-bit string, and a 48-bit round-key as input, and outputs 32 bits. The expansion function expands the 32-bit $R_i$ into 48 bits by repeating some of the bits as shown by the grey cells in Table 2.3 (the numbers are bit indices). The resulting 48-bit string is XORed with the round key, and then divided into eight 6-bit strings. The eight S-boxes map each of these 6-bit strings to 4-bit strings. The result of $F(R_i, K_{i+1})$ is a permutation of the 32 output-bits from the S-boxes.

## 2.2 Notations

This thesis will use the same numbering rule for bit positions as Matsui used in his description of linear cryptanalysis; in an $n$-bit string, the right-most bit is the zero-th bit and the left-most bit is the $(n-1)$-th bit. Matsui assumed that inputs to different rounds are independent and uniformly distributed. This is also assumed in this thesis. The following notations are used on symbols.

| | |
|---|---|
| $X[i]$ | The $i$-th bit of symbol $X$. |
| $X[i, j, ..., k]$ | $X[i] \oplus X[j] \oplus ... \oplus X[k]$. |
| $X[i \sim j]$ | Substring of symbol $X$ starting with $X[i]$ and ending with $X[j]$. |
| $X\|\|Y$ | Concatenation of $X$ and $Y$. |
| $X \oplus Y$ | Bitwise XOR between $X$ and $Y$. |

## 2.3 Linear cryptanalysis of DES

Linear cryptanalysis[2] of DES is a known-plaintext attack found by Mitsuru Matsui. The attack is impractical, in the sense that it requires a huge amount ($\approx 2^{44.5}$) of known plaintexts and their corresponding ciphertexts, but it is still of great theoretical interest.

The basic concept of linear cryptanalysis is to find a linear approximation to S-boxes. That is the parity of some input- and output-bits for S-boxes in the round function, that is zero with probability $p \neq \frac{1}{2}$. An approximation of an S-box leads to a linear approximation of the round function which includes some key bits that is zero with probability $p$ as well. Different linear approximations is then used for different rounds to get a linear approximation of the DES cipher.

The cryptanalytic problem is to determine the key bits in the linear approximation, given enough plaintext/ciphertext pairs. Matsui gave an algorithm for this which is presented in Chapter 3.

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 31 | 30 | 29 | 28 | 27 | 28 | 27 | 26 | 25 | 24 | 23 | 24 | 23 | 22 | 21 | 20 | 19 | 20 | 19 | 18 | 17 | 16 | 15 |
| 16 | 15 | 14 | 13 | 12 | 11 | 12 | 11 | 10 | 9 | 8 | 7 | 8 | 7 | 6 | 5 | 4 | 3 | 4 | 3 | 2 | 1 | 0 | 31 |

TABLE 2.3: Expansion function in $F(R_i, K_{i+1})$

## 2.4 Davies and Murphy's attack

Donald Davies and Sean Murphy[1] found some statistical properties of S-boxes in DES. Let $S(x_5, ..., x_0) = y_3, y_2, y_1, y_0$. So the output from an S-box is uniformly distributed. However, the distribution of $(x_1, x_0, y_3, y_2, y_1, y_0)$ and $(x_5, x_4, y_3, y_2, y_1, y_0)$ is non-uniform.

The expansion function in the round function in DES repeats some of the function's 32 input-bits so that there are 6 bits for each S-box. The bits are repeated in such a way that neighbouring S-boxes share two of the function's input bits.

The non-uniform distribution of $(x_1, x_0, y_3, y_2, y_1, y_0)$ and $(x_5, x_4, y_3, y_2, y_1, y_0)$ combined with the way how input-bits in the round function are repeated, make the output from multiple S-boxes non-uniform. The distribution of this output is easily computed, and depends only on some key bits. The distribution of the XOR of the output from multiple rounds is also non-uniform and easily computed.

The output from two adjacent S-boxes has only two different distributions, and is determined by the XOR of some key bits. The number of different distributions for 3 adjacent S-boxes is even for 16-round DES manageable. Observing a large number of plaintext/ciphertext pairs with a fixed key give us some information about the key. The complexity of the attack is, however, about the same as brute-force attack on the key space.

# Chapter 3

# Linear Cryptanalysis of DES

Mitsuru Matsui[2] describes a known-plaintext attack, known as linear cryptanalysis, that theoretically breaks DES. It's a theoretical attack because it requires a huge amount ($\approx 2^{44.5}$) of known plaintexts and their corresponding ciphertexts.

Linear cryptanalysis looks at correlations between the parity of some fixed input-bits and the parity of fixed output-bits of the round function in DES that hold with probability $p \neq \frac{1}{2}$ for a fixed key. This is extended through multiple rounds of DES, and is used to deduce information about key bits.

## 3.1 Principle of Linear Cryptanalysis

The central part of linear Cryptanalysis is to find what Matsui called an "effective" linear approximate expression which holds with probability $p \neq \frac{1}{2}$ for a random plaintext $P$, the corresponding cipher text $C$ and fixed secret key $K$:

$$P[i_1, i_2, ..., i_a] \oplus C[j_1, j_2, ..., j_b] = K[k_1, k_2, ..., k_c]. \tag{3.1}$$

$i_1, i_2, ..., i_a, j_1, j_2, ..., j_b, k_1, k_2, ..., k_c$ denote bit indices. Both sides of the equation represents one bit of information. Linear Cryptanalisys is a known plaintext attack, so we know all the bits on the left side of the equation. When we have such an "effective" linear approximate expression, we can derive one key bit, $K[k_1, k_2, ..., k_c]$, from the plaintext/ciphertext pairs. Matsui gives the following algorithm, based on maximum likelihood, to find the most probable value for the derived key bit, given $N$ plaintext/ciphertext pairs:

**Algorithm 1**

**Step 1:** Let $T$ be the number of plain texts such that the left hand side of (3.1) is equal to zero.

**Step 2:** If $T > N/2$:
then guess $K[k_1, k_2, ..., k_c] = 0$ (when $p > \frac{1}{2}$) or 1 (when $p < \frac{1}{2}$)
else guess $K[k_1, k_2, ..., k_c] = 1$ (when $p > \frac{1}{2}$) or 0 (when $p < \frac{1}{2}$)

(3.1) and Algorithm 1 only gives us one bit of information about the key. To increase the number of bits, one can find a linear approximation for ($n$-2)-round DES which includes the remaining two rounds.

$$\begin{aligned} &P[i_1, i_2, ..., i_a] \oplus C[j_1, j_2, ..., j_b] \oplus F_1(P_L, K_1)[u_1, u_2, ..., u_d] \\ &\quad \oplus F_n(C_L, K_n)[v_1, v_2, ..., v_e] = K[k_1, k_2, ..., k_c], \end{aligned} \tag{3.2}$$

where $F_1(P_L, K_1)[u_1, u_2, ..., u_d]$ and $F_n(C_L, K_n)[v_1, v_2, ..., v_e]$ depend on some of the bits in the round keys $K_1$ and $K_n$. For incorrect values for these bits, the probability (3.2) holds is much closer to $\frac{1}{2}$. Matsui gave the following algorithm to derive $K_1$, $K_2$ and $K[k_1, k_2, ..., k_c]$ from known plaintext/ciphertext pairs.

**Algorithm 2**

**Step 1:** Let $K_1^{(i)}$ ($i = 1, 2, ...$) and $K_n^{(j)}$ ($j = 1, 2, ...$) be possible candidates for the "effective" bits in $K_1$ and $K_n$ respectively. Then for each pair ($K_1^{(i)}$, $K_n^{(j)}$), let $T_{i,j}$ be the number of plain texts such that the left side of (3.2) is equal to zero.

**Step 2:** Let $T_{max}$ be the maximal value and $T_{min}$ be the minimal value of all $T_{i,j}$'s.

- If $|T_{max} - \frac{N}{2}| > |T_{min} - \frac{N}{2}|$, then adopt the key candidate corresponding to $T_{max}$, and guess $K[k_1, k_2, ..., k_c] = 0$ (when $p > \frac{1}{2}$) or 1 (when $p < \frac{1}{2}$)
- If $|T_{max} - \frac{N}{2}| < |T_{min} - \frac{N}{2}|$, then adopt the key candidate corresponding to $T_{min}$, and guess $K[k_1, k_2, ..., k_c] = 1$ (when $p > \frac{1}{2}$) or 0 (when $p < \frac{1}{2}$)

## 3.2 Linear Approximation of S-boxes

Matsui defined a measure of linearity for S-boxes. One counts how many times the XOR of fixed input bits agrees with the XOR of fixed output bits out of all 64 possible inputs.

For example, the most effective linear approximation is $S_5(x)[4] = x[3,2,1,0]$ which holds for 12 out of 64 possible values for $x$.

**Definition 3.3.** For a given S-box $S_a$ $(a = 1, 2, ..., 8)$, $1 \leq \alpha \leq 63$ and $1 \leq \beta \leq 15$, let

$$NS_a(\alpha, \beta) = \#\{x \mid 0 \leq x \leq 64, (\bigoplus_{s=0}^{5} x[s] \bullet \alpha[s]) = (\bigoplus_{t=0}^{3} S_a(x)[t] \bullet \beta[t])\},$$

where the symbol $\bullet$ denotes bitwise AND.

There is a correlation between the input bits and output bits of $S_a$ if $NS_a(\alpha, \beta) \neq 32$. $|NS_a(\alpha, \beta) - 32|$ gives a measure of the effectiveness of the linearity. The linear approximation, previously mentioned as the most effective, can be written as $NS_5(16, 15)$. The effectiveness of this approximation is $|NS_5(16, 15) - 32| = 20$. Because this is the most effective linear approximation of an S-box, the following equations are the best approximations of the round function $F$.

$$X[15] \oplus F(X, K)[7, 18, 24, 29] = K[22], \tag{3.4}$$

$$X[15] \oplus F(X, K)[7, 18, 24, 29] = K[22] \oplus 1. \tag{3.5}$$

(3.4) was computed from $NS_5(16, 15)$ by tracing the 4-th input-bit to $S_5$ through the expansion function to the 15-th bit of $X$. $X[15]$ is XORed with the 22-th bit in the round key before it reaches the S-box, which is why we have $K[22]$. The four output-bits from $S_5$ is traces through the permutation to output-bits 7, 18, 24 and 29. The first equation holds with probability $p = NS_5(16, 15)/64 \approx 0.19$, and the second equation holds with probability $q = 1 - p \approx 0.81$.

## 3.3 Linear Approximation of DES

This section explains how linear approximations of the round function can be extended to a linear approximation of multiple rounds in DES. First, an approximation of 3 and 5-round DES is given. An approximation of $n$-round DES is then described and used to approximate 16-round DES. Matsui gave a lemma for computing the probability the XOR of multiple random variables is equal to zero.

**Lemma 3.6** (**Piling-up Lemma**). *Let $X_i$ $(1 \leq i \leq n)$ be independent random variables whose values are 0 with probability $p_i$. Then the probability that $X_1 \oplus X_2 \oplus ... \oplus X_n = 0$ is*

$$\frac{1}{2} + 2^{n-1} \prod_{i=1}^{n} \left(p_i - \frac{1}{2}\right).$$

### 3.3.1 3-round DES

Applying (3.4) to the first round we get the following equation:

$$F_1(P_L, K_1)[7, 18, 24, 29] \oplus P_L[15] = K_1[22].$$

We can substitute $F_1$ by $X_2 \oplus P_H$ to get

$$X_2[7, 18, 24, 29] \oplus P_H[7, 18, 24, 29] \oplus P_L[15] = K_1[22].$$

Similarly, applying (3.4) to the third round gives us the following equation:

$$X_2[7, 18, 24, 29] \oplus C_H[7, 18, 24, 29] \oplus C_L[15] = K_3[22].$$

Combining these, and canceling the common term $X_2[7, 18, 24, 29]$, we get a linear approximation (3.7) of 3-round DES. All bits on the left hand side of the equation are known, and we can thus compute $K_1[22] \oplus K_3[22]$ by Algorithm 1.

$$P_H[7, 18, 24, 29] \oplus C_H[7, 18, 24, 29] \oplus P_L[15] \oplus C_L[15] = K_1[22] \oplus K_3[22]. \tag{3.7}$$

(3.7) will be correct if both, or none, of the linear approximations of the first and third round is correct. Since $NS_5(16, 15) = 12$, the probability for this is $(\frac{12}{64})^2 + (1 - \frac{12}{64})^2 = 0.70$, or by the **Piling-up Lemma**:

$$\frac{1}{2} + 2^{2-1} \times (\frac{12}{64} - \frac{1}{2}) \times (\frac{12}{64} - \frac{1}{2}) = 0.70.$$

### 3.3.2 5-round DES

For 5-round DES we will make use of the following equation derived from $NS_1(27, 4) = 22$:

$$X[27, 28, 30, 31] \oplus F(X, K)[15] = K[42, 43, 45, 46]. \tag{3.8}$$

Applying the above linear approximation on the first round and (3.4) to the second round gives us the following equation:

$$X_3[7, 18, 24, 29] \oplus P_H[15] \oplus P_L[7, 18, 24, 27, 28, 29, 30, 31] = K_5[42, 43, 45, 46] \oplus K_4[22].$$

Use the same linear approximations on the fifth and forth rounds respectively, to get

$$X_3[7, 18, 24, 29] \oplus C_H[15] \oplus C_L[7, 18, 24, 27, 28, 29, 30, 31] = K_1[42, 43, 45, 46] \oplus K_2[22].$$

Each of these two holds with probability $(\frac{12}{64})(\frac{22}{64}) + (1 - \frac{12}{64})(1 - \frac{22}{64}) = 0.598$. Combining these, and canceling the common term $X_3[7, 18, 24, 29]$, we get the following linear approximation of 5-round DES:

$$P_H[7, 18, 24, 29] \oplus C_H[7, 18, 24, 29] \oplus P_L[15] \oplus C_L[15] = K_1[22] \oplus K_3[22]. \tag{3.9}$$

(3.9) holds with probability $0.598^2 + (1 - 0.598)^2 = 0.519$ since the approximation of the first and last two rounds both hold with probability 0.598. Calculating the same probability using the **Piling-up Lemma** gives us the same result.

$$\frac{1}{2} + 2^3 \left(\frac{12}{64} - \frac{1}{2}\right)^2 \left(\frac{22}{64} - \frac{1}{2}\right)^2 = 0.519.$$

### 3.3.3 $n$-round DES

To make a linear approximation for $n$-round DES we need to introduce two more linear approximations which come from $NS_1(4, 4) = 30$ and $NS_5(16, 14) = 42$, respectively.

$$X[29] \oplus F(X, K)[15] = K[44], \tag{3.10}$$

$$X[15] \oplus F(X, K)[7, 18, 24] = K[22]. \tag{3.11}$$

Matsui built a linear approximation (3.12) to any intermediate 5 rounds in DES containing $X_i$ and $X_{i+4}$ only, by applying (3.4), (3.10) and (3.11) to round $i + 1$, $i + 2$ and $i + 3$, respectively.

$$\begin{aligned}
K_{i+1}[22] &= X_{i+1}[15] \oplus F_{i+1}(X, K)[7, 18, 24, 29],\\
K_{i+2}[44] &= X_{i+2}[29] \oplus F_{i+2}(X, K)[15],\\
K_{i+3}[22] &= X_{i+3}[15] \oplus F_{i+3}(X, K)[7, 18, 24].
\end{aligned}$$

$$\begin{aligned}
K_{i+1}[22] &= X_{i+1}[15] \oplus X_i[7, 18, 24, 29] \oplus X_{i+2}[7, 18, 24, 29],\\
K_{i+2}[44] &= X_{i+2}[29] \oplus X_{i+1}[15] \oplus X_{i+3}[15],\\
K_{i+3}[22] &= X_{i+3}[15] \oplus X_{i+2}[7, 18, 24] \oplus X_{i+4}[7, 18, 24].
\end{aligned}$$

Combining these we get:

$$\begin{aligned}
&X_i[7, 18, 24, 29] \oplus \cancel{(X_{i+1}[15] \oplus X_{i+1}[15])}\\
&\quad \oplus \cancel{(X_{i+3}[15] \oplus X_{i+3}[15])} \oplus X_{i+4}[7, 18, 24]\\
&\quad \oplus \cancel{(X_{i+2}[7, 18, 24, 29] \oplus X_{i+2}[7, 18, 24] \oplus X_{i+2}[29])}\\
&= X_i[7, 18, 24, 29] \oplus X_{i+4}[7, 18, 24]\\
&= K_{i+1}[22] \oplus K_{i+2}[44] \oplus K_{i+3}[22]
\end{aligned} \tag{3.12}$$

(3.12) holds with probability 0.506 (by **Piling-up Lemma**). This is less than the previous linear approximation for 5-round DES (3.9), which hold with probability 0.519. However, (3.12) can be used repeatedly to build a linear approximation of $n$-round DES.

Matsui built an example to show how to use this repeatedly, by making a linear approximation of 16-round DES. For this, (3.12) is used repeatedly to approximate 12 of the round functions, and (3.4), (3.8) and (3.13) approximate the remaining round functions, where (3.13) is derived from $NS_5(34, 14) = 16$.

$$X[7, 18, 24] \oplus F(X, K)[12, 16] = K[19, 23]. \tag{3.13}$$

(3.14), a linear approximation of 16 round DES, was built by approximating the intermediate rounds as listed below. Terms containing bits from round 2, 6, 10 and 14 got canceled out by common terms in the approximation of the rounds before and after them.

| | |
|---|---|
| Round 1: | (3.13) |
| Round 3, 4 and 5: | (3.12) |
| Round 7, 8 and 9: | (3.12) |
| Round 11, 12 and 13: | (3.12) |
| Round 15: | (3.4) |
| Round 16: | (3.8) |

$$\begin{aligned} & P_H[7, 18, 24] \oplus P_L[12, 16] \oplus C_H[15] \oplus C_L[7, 18, 24, 27, 28, 29, 30, 31] \\ = {} & K_1[19, 23] \oplus K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus K_9[22] \\ & \oplus K_{11}[22] \oplus K_{12}[44] \oplus K_{13}[22] \oplus K_{15}[22] \oplus K_{16}[42, 43, 45, 46]. \end{aligned} \tag{3.14}$$

The probability that (3.14) holds is:

$$\frac{1}{2} + 2^{11}\left(\frac{16}{64} - \frac{1}{2}\right)\left(\frac{12}{64} - \frac{1}{2}\right)^4\left(\frac{30}{64} - \frac{1}{2}\right)^3\left(\frac{42}{64} - \frac{1}{2}\right)^3\left(\frac{22}{64} - \frac{1}{2}\right) = \frac{1}{2} - 1.49 \times 2^{-24}.$$

## 3.4 The best expression and the best probability of DES

Matsui gave the following lemma for the probability of success when using Algorithm 1. It's based on approximating binomial distributions by normal distributions.

**Lemma 3.15.** *Let $N$ be the number of given random plaintext and $p$ be the probability that (3.1) holds. Assuming that $|p - \frac{1}{2}|$ is sufficiently small, the success rate of Algorithm 1 is*

$$\int_{-2\sqrt{N}|p-\frac{1}{2}|}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

He noted that the success rate of Algorithm 1 only depends on $\sqrt{N}|p - \frac{1}{2}|$ and calculated that if $N = |p - \frac{1}{2}|^{-2}$, the success probability would be 97.7%.

Matsui also made a list of the best linear approximations together with the best probabilities for up to 20 rounds. He found this for 3 rounds by brute force, and used induction on the best $i$-round ($0 \leq i < n$) probabilities to find the best $n$-round approximation and probability[4].

For some $n$, there are two linear approximations with the best probability. For these cases, one can compute one of the best linear approximations from the other by switching $P$ and $C$ and substituting the round index $i$ with $(n + 1 - i)$ for $F_i$ and $K_i$. For example, the following two best linear approximations of 6-round DES can be computed from each other:

$$\begin{aligned} P_L[\alpha] \oplus P_H[15] \oplus F_1(P_L, K_1)[15] \oplus C_L[15] \oplus C_H[\alpha, \beta] \oplus F_6(C_L, K_6)[\alpha, \beta] \\ = K_2[22] \oplus K_4[22] \oplus K_5[\gamma] \\ \longleftrightarrow \\ C_L[\alpha] \oplus C_H[15] \oplus F_6(C_L, K_6)[15] \oplus P_L[15] \oplus P_H[\alpha, \beta] \oplus F_1(P_L, K_1)[\alpha, \beta] \\ = K_5[22] \oplus K_3[22] \oplus K_2[\gamma], \end{aligned} \tag{3.16}$$

where $\alpha = (7, 18, 24, 29)$, $\beta = (27, 28, 30, 31)$ and $\gamma = (42, 43, 45, 46)$.

## 3.5 Known-plaintext attack of DES

Matsui described a practical method for a known-plaintext attack on DES. First, he described an attack on 8-round DES. The same method was used for 12 and 16-round DES. These practical attacks used Algorithm 2 and thus used the best linear approximations for 6, 10 and 14 rounds DES.

Lemma 3.15 can not be used to calculate the success rate for these attacks, since Algorithm 2 is used. The complexity of running the attack on 16-round DES was too high, so Matsui used experimental results from 8 and 12-round DES to predict the success rate for 16-round DES.

### 3.5.1 8-round DES

The following expression is a linear approximation of 8-round DES that holds with the best 6-round probability $\frac{1}{2} + 1.95 \times 2^{-9}$, given that we substitute the correct keys $K_1$ and $K_8$.

$$\begin{aligned} &P_H[7,18,24] \oplus F_1(P_L, K_1)[7,18,24] \oplus C_H[15] \\ &\quad \oplus C_L[7,18,24,29] \oplus F_8(C_L, K_8)[15] \\ &= K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22]. \end{aligned} \tag{3.17}$$

The 4 bits, $F_1(P_L, K_1)[7,18,24]$ and $F_8(C_L, K_8)[15]$, on the left hand side of (3.17) is unknown, but can be determined by 6 bits from $K_1$ and 6 bits from $K_8$. The left hand side can thus be divided into:

- 13 known bits from plaintext: $P_L[11] \sim P_L[16], C_L[0], C_L[27] \sim C_L[31]$, $P_H[7,18,24] \oplus C_H[15] \oplus C_L[7,18,24,29]$.
- 12 unknown bits from subkeys: $K_1[18] \sim K_1[23], K_8[42] \sim K_8[47]$.

Matsui calls these 13 known, and 12 unknown bits the effective text bits and the effective key bits respectively. He also presented a practical implementation of Algorithm 2, which compute the effective key bits and the right hand side of (3.17) using the effective text bits. The implementation is given in Algorithm 2-A.

**Algorithm 2-A**

**[Data counting phase]**

**Step 1:** Prepare $2^{13}$ counters $U_i$ $(0 \leq i \leq 2^{13})$ and initialize them to zero, where $i$ corresponds to each value on the 13 effective text bits of (3.17).

**Step 2:** For each plaintext $P$ and the corresponding ciphertext $C$, compute the value $i$ of **Step 1** and count up the counter $U_i$ by one.

**[Key counting phase]**

**Step 3:** Prepare $2^{12}$ counters $T_j$ $(0 \leq j \leq 2^{12})$ and initialize them to zero, where $j$ corresponds to each value on the 12 effective text bits of (3.17).

**Step 4:** For each $j$ of **Step 3**, let $T_j$ be the sum of $U_i$'s such that the left side of (3.17), whose value can be uniquely determined by $i$ and $j$, is equal to zero.

**Step 5:** Let $T_{max}$ be the maximal value and $T_{min}$ be the minimal value of all $T_{i,j}$'s.

- If $|T_{max} - \frac{N}{2}| > |T_{min} - \frac{N}{2}|$, then adopt the subkey value $j$ corresponding to $T_{max}$ and guess that the right hand side of (3.17) is 0.

- If $|T_{max} - \frac{N}{2}| < |T_{min} - \frac{N}{2}|$, then adopt the subkey value $j$ corresponding to $T_{min}$ and guess that the right hand side of (3.17) is 1.

The complexity of the data counting phase is $\mathrm{O}(N)$ as it looks at every plaintext once. **Step 4** in the key counting phase computes the left side of (3.17) $2^{12+13}$ times. The total complexity of the algorithm is thus $\mathrm{O}(N) + \mathrm{O}(2^{t+k})$, where $t$ and $k$ is the number of effective text bits and effective key bits, respectively.

Using Algorithm 2-A on (3.17) gives us 13 subkey bits. There is another linear approximation of 8-round DES, with the same probability. It is found by using the same method as used in (3.16) on (3.17).

$$\begin{aligned} &C_H[7,18,24] \oplus F_8(C_L, K_8)[7,18,24] \oplus P_H[15] \\ &\quad \oplus P_L[7,18,24,29] \oplus F_1(P_L, K_8)[15] \\ &= K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[22]. \end{aligned} \tag{3.18}$$

Solving (3.17) and (3.18) gives us the following 26 subkey bits: $K_1[18] \sim K_1[23]$, $K_1[42] \sim K_1[47]$, $K_8[18] \sim L_8[23]$, $K_8[42] \sim K_8[47]$, $K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[22]$ and $K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22]$. Tracing these through the key-schedule algorithm gives us the following 23 key bits:

$$\begin{gathered} 0,1,3,5,8,11,14,15,18,20,23,24,28,31,37 \\ 38,41,44,46,50,53,54,2 \oplus 22 \oplus 26 \oplus 52. \end{gathered}$$

Matsui created and ran a program implementing Algorithm 2-A, solving (3.17) and (3.18) at the same time, and found the above 23 key bits in a few seconds using 400KB memory. The remaining 33 key bits where found by brute force in 5 hours. This was done with $2^{20}$ plaintext/ciphertext pairs with 99.9% success rate.

Matsui showed another implementation of Algorithm 2 as well. The goal was to shorten the computational time, while accepting that more plaintext/ciphertext pairs would be required or accepting a lower success rate. For this purpose, he used the second-best linear approximations of 6-round DES to create the following 8-round linear approximations:

$$\begin{aligned} &P_H[7,18,24,29] \oplus F_1(P_L, K_1)[7,18,24,29] \\ &\quad \oplus C_H[12,16] \oplus C_L[7,18,24] \oplus F_8(C_L, K_8)[12,16] \\ &= K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[19,23], \end{aligned} \tag{3.19}$$

and

$$\begin{aligned}&C_H[7,18,24,29] \oplus F_8(C_L,K_8)[7,18,24,29]\\&\quad\oplus P_H[12,16] \oplus P_L[7,18,24] \oplus F_1(P_L,K_1)[12,16]\\&= K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[19,23].\end{aligned} \tag{3.20}$$

In contrary to previous expressions with $F(X,K)$ where output from only one S-box is considered, $F_i(X,K)[12,16]$ is the XOR of 2 output-bits from two adjacent S-boxes. Therefore, $F_8(C_L,K_8)[12,16]$ and $F_1(P_L,K_1)[12,16]$ depends on 12 subkey bits each. (3.19) thus has the following effective text/key bits:

- 17 effective text bits: $P_L[11] \sim P_L[16], C_L[0], C_L[15] \sim C_L[24]$, $P_H[7,18,24,29] \oplus C_H[12,16] \oplus C_L[7,18,24]$.
- 18 effective key bits: $K_1[18] \sim K_1[23], K_8[24] \sim K_8[35]$.

The execution of Algorithm 2 will take longer time since there are more effective key bits to try, while the brute-force attack on the rest of the keys will take less time. The overall time used on the attack is much shorter.

Using Algorithm 2-A on (3.19) would cause problems because **Step 4** would take too long. Matsui gave the following implementation of Algorithm 2, which solves (3.17) and (3.19) or (3.17) and (3.20). Matsui's implementation in software solved all three of them at the same time.

**Algorithm 2-B**

**[Data counting phase 1]**

**Step 1:** Prepare $2^{13}$ counters $U_i$ $(0 \leq i \leq 2^{13})$ and $2^{17}$ counters $V_j$ $(0 \leq j \leq 2^{17})$, and initialize them to zero, where $i$ and $j$ correspond to each value on the 13 effective text bits of (3.17) and the 17 effective text bits of (3.19), respectively.

**Step 2:** For each plaintext $P$ and the corresponding ciphertext $C$, compute $i$ and $j$ of **Step 1**, and count up the counters $U_i$ and $V_j$ by one.

**[Key counting phase 1]**

**Step 3:** Solve (3.17) using $U_i$'s. We then have the 12 effective key bits and one subkey bit of the right hand side of (3.17).

In this stage, we are able to calculate the exact value of $F_1(P_L,K_1)$. It is therefore possible to regard the effective text bits and the effective key bits of (3.19) as follows:

- 11 effective text bits: $C_L[15] \sim C_L[24]$,
  $F_1(P_L, K_1)[7, 18, 24, 29] \oplus P_H[7, 18, 24, 29] \oplus \ C_H[12, 16] \oplus C_L[7, 18, 24]$.
- 12 effective key bits: $K_8[24] \sim K_8[35]$.

This enables us to ”pack” $V_j$’s into the following new counters $W_k$:

[**Data counting phase 2**]

**Step 4:** Prepare $2^{11}$ counters $W_k$ $(0 \leq k \leq 2^{11})$ and initialize them by zeros, where $k$ corresponds to each value on the 11 effective text bits above.

**Step 5:** For each $j$ $(0 \leq j \leq 2^{17})$, compute $k$ of **Step 4**, whose value is uniquely determined by $j$, and add $V_j$ to $W_k$.

[**Key counting phase 2**]

**Step 6:** Solve (3.19) using $W_k$’s. We then have the 12 effective key bits above and the right hand side of (3.19).

Solving (3.17), (3.19) and (3.20), and tracing the subkey bits through the key-schedule algorithm, gives us the following 38 key bits:

$$0, 1, 3, 5, 8, 11, 14, 15, 18, 20, 23, 24, 25, 28, 29, 30, 31,$$
$$32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47,$$
$$48, 50, 51, 53, 54, 2 \oplus 7 \oplus 13, 2 \oplus 22 \oplus 26 \oplus 52.$$

Matsui’s instance of Algorithm 2-B found the above 38 key bits and the remaining 18 key bits, by brute force, in less than 10 seconds using 1MB of memory. This was done with $2^{20}$ plaintext/ciphertext pairs with 96.2% success rate.

### 3.5.2 12-round DES

For 12-round DES, Matsui used the same procedure as for 8-round DES. The following linear approximations, which holds with probability $\frac{1}{2} - 1.53 \times 2^{-15}$, was used with Algorithm 2.

$$\begin{aligned} & P_H[7, 18, 24, 29] \oplus F_1(P_L, K_1)[7, 18, 24, 29] \oplus C_H[15] \\ & \quad \oplus \ C_L[7, 18, 24, 29] \oplus \ F_{12}(C_L, K_{12})[15] \\ & = K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus K_9[22] \oplus K_{11}[22], \end{aligned} \tag{3.21}$$

and

$$\begin{aligned}
&C_H[7,18,24,29] \oplus F_{12}(C_L,K_{12})[7,18,24,29] \oplus P_H[15] \\
&\quad \oplus\ P_L[7,18,24,29] \oplus\ F_1(P_L,K_1)[15] \\
&= K_{10}[22] \oplus K_9[44] \oplus K_8[22] \oplus K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[22]. \qquad (3.22)
\end{aligned}$$

Both (3.21) and (3.22) have 13 effective text bits and 12 effective key bits. Solving them both gives us 26 subkey bits which corresponds to the following 25 key bits:

$$\begin{gathered}
0,3,4,8,11,14,16,18,22,24,26,30,31,34,38,39, \\
41,44,46,49,50,52,54,2\oplus 15\oplus 45,13\oplus 17\oplus 20.
\end{gathered}$$

Matsui's implementation of Algorithm 2 found the 25 key bits in just above 4 hours using 400KB of memory. This was done with $2^{32}$ plaintext/ciphertext pairs with a success rate of 94%. The remaining 31 bits was found by brute force in 1.5 hours.

### 3.5.3 16-round DES

The procedure for full, 16-round, DES is exactly the same as for 8 and 12-round DES. The following linear approximations both hold with probability $\frac{1}{2} - 1.19 \times 2^{-21}$.

$$\begin{aligned}
&P_H[7,18,24] \oplus F_1(P_L,K_1)[7,18,24] \oplus C_H[15] \oplus\ C_L[7,18,24,29] \oplus F_{16}(C_L,K_{16})[15] \\
&= K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus K_9[22] \\
&\quad \oplus\ K_{11}[22] \oplus K_{12}[44] \oplus K_{13}[22] \oplus K_{15}[22], \qquad (3.23)
\end{aligned}$$

and

$$\begin{aligned}
&C_H[7,18,24] \oplus F_{16}(C_L,K_{16})[7,18,24] \oplus P_H[15] \oplus\ P_L[7,18,24,29] \oplus F_1(P_L,K_1)[15] \\
&= K_{14}[22] \oplus K_{13}[44] \oplus K_{12}[22] \oplus K_{10}[22] \oplus K_9[44] \oplus K_8[22] \\
&\quad \oplus\ K_6[22] \oplus K_5[44] \oplus K_4[22] \oplus K_2[22]. \qquad (3.24)
\end{aligned}$$

Both (3.23) and (3.24) has 13 effective text bits and 12 effective key bits. Solving them both gives us 26 subkey bits which corresponds to the following 25 key bits:

$$\begin{gathered}
0,1,3,4,8,9,14,15,18,19,24,25,31,32,38,39,41,42,44, \\
45,50,51,54,55,5\oplus 13\oplus 17\oplus 20\oplus 46,2\oplus 7\oplus 11\oplus 22\oplus 26\oplus 37\oplus 52.
\end{gathered}$$

Matsui did not run Algorithm 2 on (3.23) and (3.24) because it would require too much computations. Instead, he used the experimental results from 8 and 12-round DES to calculate the number of required plaintext/ciphertext pairs, and the probability of success.

He computed the efficiency of the attack on 8 and 12 round DES when $N = a|p - \frac{1}{2}|^{-2}$ for (a=2, 4, 8), where $p$ is the best probability for 6 and 10-round DES, $\frac{1}{2} + 1.95 \times 2^{-9}$ and $\frac{1}{2} - 1.53 \times 2^{-15}$, respectively. Table 3.25 lists his results.

| $N$ | $2\|p-1/2\|^{-2}$ | $4\|p-1/2\|^{-2}$ | $8\|p-1/2\|^{-2}$ |
|---|---|---|---|
| (3.17) | $N = 1.05 \times 2^{17}$ | $N = 1.05 \times 2^{18}$ | $N = 1.05 \times 2^{19}$ |
| | 17.9% | 53.7% | 94.8% |
| (3.21) | $N = 1.72 \times 2^{29}$ | $N = 1.72 \times 2^{30}$ | $N = 1.72 \times 2^{31}$ |
| | 13% | 46% | 91% |

Table 3.25: Experimental results to solve (3.17) and (3.21)

Matsui estimated from Table 3.25 that the attack on 16-round DES with (3.23) should be successful with high probability when the number of plaintext/ciphertext pairs is $N = 8|1.19 \times 2^{-21}|^{-2} = 1.41 \times 2^{44}$.

This chapter has described linear cryptanalysis. Linear approximations of the S-boxes in DES was found and used to build linear approximations of the round function. These were extended to multiple rounds and resulted in a best linear approximation of DES. A known-plaintext attack on 8, 12 and 16-round DES was described, where it was shown that full DES can be broken with high success rate, given $1.41 \times 2^{44}$ plaintext/ciphertext pairs.

# Chapter 4

# Davies and Murphy's attack

Donald Davies and Sean Murphy[1] found, in 1993, some statistical properties of S-boxes in DES. The distribution for fixed input-bits and all output-bits is non-uniform and completely depends on some of the key bits, and can be used in a known-plaintext attack on DES.

The complexity of the attack described in the Davies and Murphy's paper is about the same as brute-force attack on the key space. The attack was later improved by Biham and Biryukov[5] to a complexity of $2^{50}$. Later, Kunz-Jacques and Muller[6] improved the results to a chosen-plaintext attack with the cost of $2^{45}$ plaintexts.

This chapter explain Davies and Murphy's original analysis, with a different notation.

## 4.1 Principle of Davies and Murphy's attack

A property of DES is that $E$ causes two adjacent S-boxes in $F_i(X_i, K_i)$ to share two bits from $X$. This is caused by the way that $E$ replicates half of the bits in $X_i$ (see Figure 4.1). If we name the 2 right-most input bits to S-box 5 '$x$', and the 2 left-most input bits to S-box 6 '$y$', then we will have that $x \oplus y = (K_i[17, 19], K_i[16, 18])$, which is part of the round key. We will refer to this as the common key bits from now on. This will be true for any pair of adjacent S-boxes, where S-box 8 and 1 is considered to be adjacent because of the way $E$ "wraps around" the first and last bit of $X$.

Another property of DES is that the output from any single S-box is uniformly distributed. This comes from the fact that the S-box is a permutations on the 4 middle bits, where the two outer bits selects the permutation. It turns out however, that output from two or more adjacent S-boxes is not uniformly distributed and the common key bits

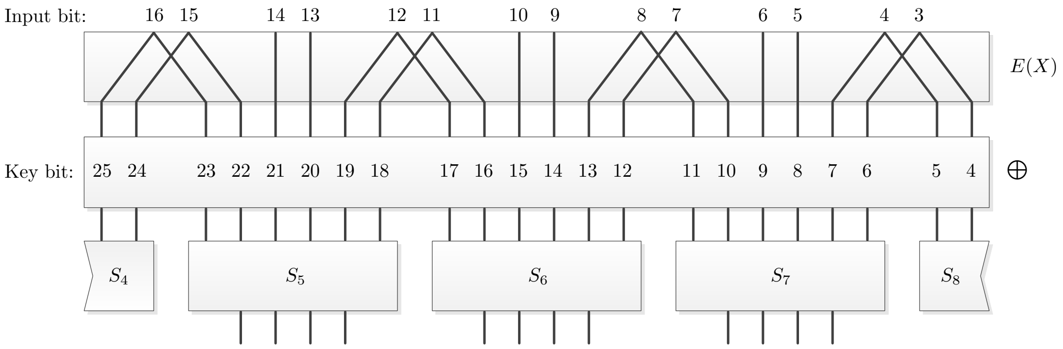


FIGURE 4.1: $E$ replicates input bits to adjacent S-boxes

determines which distribution the output follows. This non-uniform distribution and the correlation between the distribution and the common key bits is what the Davies and Murphy's attack exploits.

Davies and Murphy used a different notation than in this paper. We use the same notation as in our work on linear dependencies between such distributions (next chapter). They also represented the distributions differently.

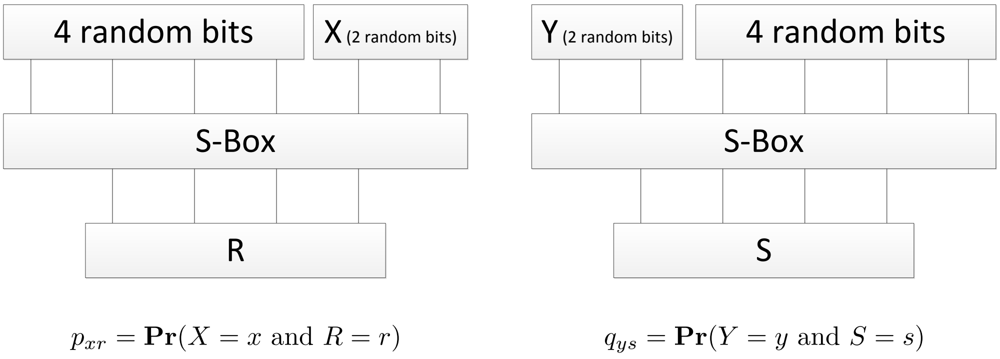
## 4.2 2 adjacent S-boxes

We define two distributions related to one S-box, illustrated in Figure 4.4, and use them to compute the distribution of the output from two adjacent S-boxes.

**Definition 4.2.** An S-box is a mapping $S(x_5, x_4, x_3, x_2, x_1, x_0) = (y_3, y_2, y_1, y_0)$.
The **right distribution** of $S_i$ is the distribution of $(x_1, x_0, y_3, y_2, y_1, y_0)$ given uniformly random input to $S_i$. We denote $p^{(i)}_{xr} = \mathbf{Pr}(x_1x_0 = x \text{ and } y_3y_2y_1y_0 = r)$.
The **left distribution** of $S_i$ is the distribution of $(x_5, x_4, y_3, y_2, y_1, y_0)$ given uniformly random input to $S_i$. We denote $q^{(i)}_{xr} = \mathbf{Pr}(x_5x_4 = x \text{ and } y_3y_2y_1y_0 = r)$.

If the S-box index is unimportant, we may use the notation $p_{xr}$ or $q_{ys}$. Appendix A lists $p_{xr}$ and $q_{xr}$ for all S-boxes. The row index in the tables represents the value for $x$ (the 2 right/left-most input-bits) and the column index represent $r$. The tables show that the left and right distributions are non-uniform and the following equations hold for all S-boxes.

$$p_{(x\oplus 2)r} + p_{xr} = \frac{1}{32}, \qquad q_{(x\oplus 1)r} + q_{xr} = \frac{1}{32}, \tag{4.3}$$



$p_{xr} = \mathbf{Pr}(X = x \text{ and } R = r)$ $\quad q_{ys} = \mathbf{Pr}(Y = y \text{ and } S = s)$

FIGURE 4.4: Right and left distribution of an S-box

where we use the following notation. Let $x = x_1||x_0$, then $x \oplus 2 = (x_1 \oplus 1)||x_0$ and $x \oplus 1 = x_1||(x_0 \oplus 1)$.

Adding the first and the second (or third and forth) row of any of the tables representing $q$ produces a constant row. Similarly adding the first and third (or second and forth) row of any of the tables representing $p$ produces a constant row as well.

The distribution on the 8-bit output from two adjacent S-boxes can be calculated by using the left and right distributions on each S-box. The probability for output $rs$ given random input is the combined probability of $p_{xr}$ and $q_{ys}$, where $x \oplus y = A$, and A are common key bits of the two S-boxes. We have

$$\begin{aligned}\mathbf{Pr}(rs \mid A) &= \frac{\sum_{[x\oplus y=A]} p_{xr}^{(i)} q_{ys}^{(i+1)}}{\mathbf{Pr}(A)} = 4 \times \sum_{x\oplus y=A} p_{xr}^{(i)} q_{ys}^{(i+1)} \\ &= 4 \times \sum_x p_{(x\oplus A)r}^{(i)} q_{xs}^{(i+1)} = 4 \times \sum_x p_{xr}^{(i)} q_{(x\oplus A)s}^{(i+1)}. \end{aligned} \tag{4.5}$$

Davies and Murphy proved

$$2^{10} \times \mathbf{Pr}(rs \mid A) = 4 + (-1)^{A[0,1]}(p_{0r} - p_{1r})(q_{0s} - q_{2s})2^{12}.$$

That implies that $rs$ can have only two distributions, determined by the XOR of the two bits in $A$. We show this is true as well.

**Lemma 4.6.**

$$\mathbf{Pr}(rs \mid A) = \mathbf{Pr}(rs \mid A \oplus 3), \quad \textit{for any } rs.$$

*Proof:*

$$
\begin{aligned}
\mathbf{Pr}(rs \mid A) &= 4 \times \sum_{x \oplus y = A} p_{xr}^{(i)} q_{ys}^{(i+1)} \\
&= 4 \times \sum_{x \oplus y = A} (\frac{1}{32} - p_{(x\oplus 2)r}^{(i)})(\frac{1}{32} - q_{(y\oplus 1)s}^{(i+1)}) \quad \text{(by (4.3))} \\
&= 4 \times \sum_{x\oplus 2 \oplus y \oplus 1 = A\oplus 3} (\frac{1}{32} - p_{(x\oplus 2)r}^{(i)})(\frac{1}{32} - q_{(y\oplus 1)s}^{(i+1)}) \\
&= 4 \times \sum_{x \oplus y = A\oplus 3} (\frac{1}{32} - p_{xr}^{(i)})(\frac{1}{32} - q_{ys}^{(i+1)}) \\
&= 4 \times \sum_{x \oplus y = A\oplus 3} (\frac{1}{32^2} - \frac{p_{xr}^{(i)}}{32} - \frac{q_{ys}^{(i+1)}}{32} + q_{xr}^{(i)} q_{ys}^{(i+1)}) \\
&= 4 \times (\frac{4}{32^2} - \frac{1}{32 \times 16} - \frac{1}{32 \times 16}) + 4 \times \sum_{x \oplus y = A\oplus 3} q_{xr}^{(i)} q_{ys}^{(i+1)} \quad (\text{by} \sum_x p_{xr} = \frac{1}{16}) \\
&= 4 \times \sum_{x \oplus y = A\oplus 3} q_{xr}^{(i)} q_{ys}^{(i+1)} = \mathbf{Pr}(rs \mid A \oplus 3).
\end{aligned}
$$

□

The distribution of the XOR of two outputs from pairs of S-boxes is computed by the self-convolution of (4.5), and the distribution for the XOR of $n$ outputs from pairs of S-boxes is computed by the $n$-fold self-convolution of (4.5). We assume independent input to each round function in DES, so we have for 2 rounds

$$
\mathbf{Pr}(rs \mid AB) = \sum_{ab} \mathbf{Pr}(ab \mid A) \mathbf{Pr}(rs \oplus ab \mid B), \tag{4.7}
$$

and for $n$ rounds

$$
\mathbf{Pr}(rs \mid K_1 ... K_n) = \sum \mathbf{Pr}(I_1 \mid K_1) \times ... \times \mathbf{Pr}(I_n \mid K_n), \tag{4.8}
$$

where the sum is over all $I_i$ such that $\bigoplus_{i=1}^{n} I_i = rs$ and $K_i$ is the common key bits for the $i$-th pair of S-boxes.

Davies and Murphy showed that the XOR of $n$ outputs can not have more than two distributions. To show this with our representation of the distributions, we first need to show the following 2 lemmas about $n$-fold convolutions of $p_{xr}$ and $q_{yt}$.

**Lemma 4.9.**

$$\sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{x_2}a_2}\cdots p_{x_na_n} + \sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_na_n} = 2^{-(2n+3)},$$

*and*

$$\sum_{\oplus c_i=t} q_{y_1c_1}q_{\underline{y_2}c_2}\cdots q_{y_nc_n} + \sum_{\oplus c_i=t} q_{y_1c_1}q_{\underline{(y_2\oplus 1)}c_2}\cdots q_{y_nc_n} = 2^{-(2n+3)}.$$

*Proof:*

$$\sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{x_2}a_2}\cdots p_{x_na_n} + \sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_na_n}$$

$$= \sum_{a_1a_2\ldots a_{n-1}} p_{x_1a_1}p_{\underline{x_2}a_2}\cdots p_{x_{n-1}a_{n-1}}p_{x_n(a'\oplus r)} + \sum_{a_1a_2\ldots a_{n-1}} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_{n-1}a_{n-1}}p_{x_n(a'\oplus r)}$$

$$= \sum_{a_1a_2\ldots a_{n-1}} p_{x_1a_1}p_{x_3a_3}p_{x_4a_4}\cdots p_{x_{n-1}a_{n-1}}p_{x_n(a'\oplus r)} \left(p_{\underline{x_2}a_2} + p_{\underline{(x_2\oplus 2)}a_2}\right)$$

$$= \frac{1}{32} \times \sum_{a_1a_3a_4\ldots a_{n-1}} p_{x_1a_1}p_{x_3a_3}p_{x_4a_4}\cdots p_{x_{n-1}a_{n-1}} \left(\sum_{a_2} p_{x_n(a'\oplus r)}\right)$$

$$= \frac{1}{32\times 4} \times \sum_{a_1a_3a_4\ldots a_{n-2}} p_{x_1a_1}p_{x_3a_3}p_{x_4a_4}\cdots p_{x_{n-2}a_{n-2}} \left(\sum_{a_{n-1}} p_{x_{n-1}a_{n-1}}\right)$$

. . .

$$= \frac{1}{32\times 4^{n-4}} \times \sum_{a_1a_3} p_{x_1a_1}p_{x_3a_3}\left(\sum_{a_4} p_{x_4a_4}\right)$$

$$= \frac{1}{32\times 4^{n-3}} \times \sum_{a_1} p_{x_1a_1}\left(\sum_{a_3} p_{x_3a_3}\right)$$

$$= \frac{1}{32\times 4^{n-2}} \times \left(\sum_{a_1} p_{x_1a_1}\right)$$

$$= \frac{1}{32\times 4^{n-1}} = 2^{-(2n+3)}$$

where $a' = a_1 \oplus a_2 \oplus \ldots \oplus a_{n-1}$.

Similarly, one proves the second equality.

□

**Lemma 4.10.**

$$\sum_{\oplus a_i=r} p_{\underline{(x_1\oplus 2)}a_1}p_{x_2a_2}\cdots p_{x_na_n} = \sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_na_n},$$

*and*

$$\sum_{\oplus c_i=t} q_{\underline{(y_1\oplus 1)}c_1}q_{y_2c_2}\cdots q_{y_nc_n} = \sum_{\oplus c_i=t} q_{y_1c_1}q_{\underline{(y_2\oplus 1)}c_2}\cdots q_{y_nc_n}.$$

*Proof:*

By Lemma 4.9 we have that

$$\begin{aligned}&\sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{x_2}a_2}\cdots p_{x_na_n} + \sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_na_n}\\ =&\sum_{\oplus a_i=r} p_{\underline{x_1}a_1}p_{x_2a_2}\cdots p_{x_na_n} + \sum_{\oplus a_i=r} p_{\underline{(x_1\oplus 2)}a_1}p_{x_2a_2}\cdots p_{x_na_n}.\end{aligned}$$

Canceling the common term (the two left-most sums) we get

$$\sum_{\oplus a_i=r} p_{\underline{(x_1\oplus 2)}a_1}p_{x_2a_2}\cdots p_{x_na_n} = \sum_{\oplus a_i=r} p_{x_1a_1}p_{\underline{(x_2\oplus 2)}a_2}\cdots p_{x_na_n}.$$

Similarly, one proves the second equality.

□

From Lemma 4.10 we have

**Corollary 4.11.** *For $C = 1$ or $2$*

$$\boldsymbol{Pr}(rs \mid K_1 ... K_i ... K_j ... K_n) = \boldsymbol{Pr}(rs \mid K_1...(K_i \oplus C)...(K_j \oplus C)...K_n).$$

*Proof:*

$$\mathbf{Pr}(rs \mid K_1...K_i...K_j...K_n) = \sum_{\bigoplus I_i = rs} \mathbf{Pr}(I_1 \mid K_1) \ ... \ \mathbf{Pr}(I_n \mid K_n),$$

after changing the summation order

$$= \sum_{x_1...x_n} \left( \sum_{\oplus a_i=r} p_{(x_1\oplus K_1)a_1}\cdots p_{(x_n\oplus K_n)a_n} \right) \left( \sum_{\oplus b_i=s} q_{x_1b_1}\cdots q_{x_nb_n} \right) \tag{4.12}$$

$$= \sum_{x_1...x_n} \left( \sum_{\oplus a_i=r} p_{x_1a_1}\cdots p_{x_na_n} \right) \left( \sum_{\oplus b_i=s} q_{(x_1\oplus K_1)b_1}\cdots q_{(x_n\oplus K_n)b_n} \right), \tag{4.13}$$

where $I_i = a_i || b_i$.

By Lemma 4.10, we can XOR $K_i$ and $K_j$ in (4.12) or (4.13) with $C$, depending on whether $C$ equals 2 or 1, respectively.

□

We can now show what we have been working towards, namely that the XOR of $n$ outputs can not have more than two different distributions.

**Lemma 4.14.** *The XOR of n outputs can have only two different distributions, depending on the XOR of all common key bits. That is*

$$\boldsymbol{Pr}(rs \mid K_1...K_n) = \boldsymbol{Pr}(rs \mid 0...0k),$$

*where* $k = K_1[0] \oplus K_1[1] \oplus ... \oplus K_n[0] \oplus K_n[1]$.

*Proof:*
By Corollary 4.11

$$\begin{aligned}\mathbf{Pr}(rs \mid K_1...K_n) &= \mathbf{Pr}(rs \mid 0(K_1 \oplus K_2)K_3...K_n) \\ &= \cdots \\ &= \mathbf{Pr}(rs \mid 0...0(K_1 \oplus ... \oplus K_n)).\end{aligned}$$

By the same procedure as in the proof of Lemma 4.6

$$\mathbf{Pr}(rs \mid 0...0(K_1 \oplus ... \oplus K_n)) = \mathbf{Pr}(rs \mid 0...0(K_1 \oplus ... \oplus K_n \oplus 3)),$$

which is the same as to say that

$$\mathbf{Pr}(rs \mid 0...0(K_1 \oplus ... \oplus K_n)) = \mathbf{Pr}(rs \mid 0...0k).$$

□

## 4.3 3 adjacent S-boxes

Davies and Murphy extended their results for pairs of S-boxes in DES to triplets of S-boxes. The principle is the same, namely, the output $rst$ from 3 adjacent S-boxes is non-uniform and is determined by the common key bits. For triplets there are two sets of common key bits, one for each pair of adjacent S-boxes in the triplet. We will denote these four common key bits by $AB$ ($A$ and $B$ are 2-bit values), or $K_i$ (4-bit value) if we look at multiple outputs. To compute the distributions, we need to define another distribution on fixed input-bits and all output-bits for an S-box.

**Definition 4.15.** An S-box is a mapping $S(x_5, x_4, x_3, x_2, x_1, x_0) = (y_3, y_2, y_1, y_0)$.
The **Q distribution** of $S_i$ is the distribution of $(x_5, x_4, x_1, x_0, y_3, y_2, y_1, y_0)$ given uniformly random input to $S_i$. We denote $Q^{(i)}_{xyr} = \mathbf{Pr}(x_5x_4 = x,\ x_1x_0 = y,\ y_3y_2y_1y_0 = r)$.

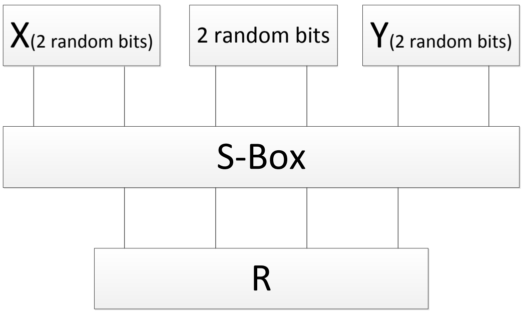


FIGURE 4.16: $Q$ distribution: $Q_{xyr} = \mathbf{Pr}(X = x, Y = y, R = r)$

If the S-box index is unimportant, we may use the notation $Q_{xyr}$. See Figure 4.16 for an illustration.

Appendix B lists $Q_{xyr}$ for all S-boxes. The row index in the tables represent the values for $x, y$ (2 bit each) and the column index represent $r$. The tables show that the $Q$ distributions are non-uniform.

The distribution on the 12-bit output from 3 adjacent S-boxes can be calculated by using the left, $Q$, and right distribution of each S-box. We have

$$\begin{aligned}\mathbf{Pr}(rst \mid AB) &= \frac{1}{\mathbf{Pr}(AB)} \times \sum_{\begin{bmatrix} x\oplus x'=A \\ y\oplus y'=B \end{bmatrix}} p^{(i)}_{xr} Q^{(i+1)}_{x'y's} q^{(i+2)}_{yt} \\ &= 16 \times \sum_{x,y} p^{(i)}_{(x\oplus A)r} Q^{(i+1)}_{xys} q^{(i+2)}_{(y\oplus B)t}, \end{aligned} \tag{4.17}$$

where A is the common key bits of $S_i, S_{i+1}$ and B is the common key bits of $S_{i+1}S_{i+2}$. Figure 4.18 illustrates 3 adjacent S-boxes and the variables in $\mathbf{Pr}(rst \mid AB)$.

(4.17) is an expression for the distribution on $rst$ given key bits $AB$. We can use self-convolution or $n$-fold self-convolution on (4.17) to calculate the distribution for the XOR of 2 or $n$ such outputs, respectively. We thus have

$$\begin{aligned}\mathbf{Pr}(rst \mid AB, CD) &= \sum_{abc} \mathbf{Pr}(abc \mid AB)\mathbf{Pr}(rst \oplus abc \mid CD), \\ \mathbf{Pr}(rst \mid K_1 ... K_n) &= \sum_{\bigoplus I_i = rst} \mathbf{Pr}(I_1 \mid K_1) \; ... \; \mathbf{Pr}(I_n \mid K_n),\end{aligned}$$

where $I_i$ are 12-bit intermediate values and $K_i$ is the 4 common key bits for the $i$-th S-box triplet.

For triplets there can be more than 2 different distributions. Davies and Murphy showed that $rst$, the result of an $n$-fold convolution on (4.17), can have at most $4 \times \frac{(n+3)!}{6n!}$ distributions. We will now show this using our representation of the distributions.

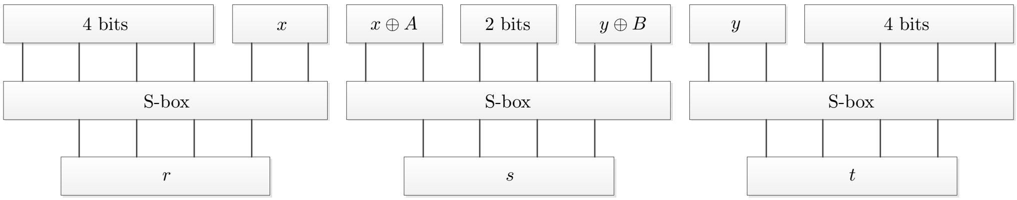


FIGURE 4.18: The variables in $\mathbf{Pr}(rst \mid AB)$

**Lemma 4.19.**
*Let*

- *XOR of the left-most bit in all $K_i$ equals the XOR of the left-most bit in all $K_i'$*
- *XOR of the right-most bit in all $K_i$ equals the XOR of the right-most bit in all $K_i'$*
- *The two middle bits in $K_i$ equals the two middle bits in $K_i'$*

*then*

$$\boldsymbol{Pr}(rst \mid K_1 ... K_n) = \boldsymbol{Pr}(rst \mid K_1' ... K_n').$$

That is, $\mathbf{Pr}(rst \mid K_1...K_n)$ does not change if we set the left-most bit in $K_i$ (for $i \leq n-1$) to zero, and we set the left-most bit in $K_n$ to the left-most bit of $K_1 \oplus ... \oplus K_n$. The same is true for the right-most bits.

*Proof:*
For simplicity, we prove the lemma for a 2-fold convolution of $\mathbf{Pr}(rst \mid AB)$

$$\begin{aligned}\mathbf{Pr}(rst \mid AB, CD) &= \sum_{abc} \mathbf{Pr}(abc \mid AB)\mathbf{Pr}(abc \oplus rst \mid CD) \\ &= \sum_{\substack{x_1,y_1\\x_2,y_2}} (\sum_a p^{(i)}_{(x_1 \oplus A)a} p^{(i)}_{(x_2 \oplus C)(a \oplus r)})(\sum_c q^{(i+2)}_{(y_1 \oplus B)c} q^{(i+2)}_{(y_2 \oplus D)(c \oplus t)})(\sum_b Q^{(i+1)}_{x_1 y_1 b} Q^{(i+1)}_{x_2 y_2 (b \oplus s)}),\end{aligned}$$

let $A = [k_1 k_2], B = [k_3 k_4], C = [k_5 k_6], D = [k_7 k_8]$,

$$\begin{aligned}= \sum_{\substack{x_1,y_1\\x_2,y_2}} &(\sum_a p^{(i)}_{(x_1 \oplus [k_1 k_2])a} p^{(i)}_{(x_2 \oplus [k_5 k_6])(a \oplus r)}) \\ &\times (\sum_c q^{(i+2)}_{(y_1 \oplus [k_3 k_4])c} q^{(i+2)}_{(y_2 \oplus [k_7 k_8])(c \oplus t)})(\sum_b Q^{(i+1)}_{x_1 y_1 b} Q^{(i+1)}_{x_2 y_2 (b \oplus s)}).\end{aligned}$$

We fix the middle bits $(k_2, k_3, k_6, k_7)$ and "move" the right/left-most key-bits $k_1, k_4, k_5, k_8$ to the last term by Lemma 4.10. We get

$$\sum_a p^{(i)}_{(x_1 \oplus \underline{[k_1 k_2]})a} p^{(i)}_{(x_2 \oplus \underline{[k_5 k_6]})(a \oplus r)} = \sum_a p^{(i)}_{(x_1 \oplus \underline{[0 k_2]})a} p^{(i)}_{(x_2 \oplus \underline{[(k_1 \oplus k_5) k_6]})(a \oplus r)},$$

and

$$\sum_c q^{(i+2)}_{(y_1 \oplus \underline{[k_3 k_4]})c} q^{(i+2)}_{(y_2 \oplus \underline{[k_7 k_8]})(c \oplus t)} = \sum_c q^{(i+2)}_{(y_1 \oplus \underline{[k_3 0]})c} q^{(i+2)}_{(y_2 \oplus \underline{[k_7 (k_4 \oplus k_8)]})(c \oplus t)}.$$

The proof can be generalized to an $n$-fold convolution of $\mathbf{Pr}(rst \mid AB)$.

□

**Lemma 4.20.** *Let $W_i$ be the two middle bits in $K_i$. **Pr**(rst $\mid$ $K_1 ... K_n$) will not change if one swaps $W_i$ and $W_j$.*

*Proof:*
Because XOR is commutative, the order in which the outputs from the triplets are XORed does not affect the distribution. Changing the order of the keys thus give the same distribution as changing the order of the outputs, since the distribution of an output is fully dependent on the common key bits. Swapping the two middle bits of $K_i$ and $K_j$, is the same as first swapping the order of the $K_i$ and $K_j$, and then reordering the outer bits back to the original order (which does not affect the distribution according to Lemma 4.19). Therefore, swapping the two middle bits of $K_i$ and $K_j$ cannot change the distribution of $rst$.

□

A naive bound on the maximum number of different distributions, $\mathbf{Pr}(rst \mid K_1 ... K_n)$, would be $2^{4n}$ (the number of values $K_1 ... K_n$ can have). Using Lemma 4.19 and 4.20 we can give a much lower bound. The XOR of all left/right-most bits in $K_i$ can only have 2 values each. So if we fix the 2 middle bits in each $K_i$, we can have 4 different distributions.

For each of these 4, we can permute the set of $n$ middle-bit tuples without changing the distribution. Each of the tuples can have 4 different values (0-3). The number of different combinations when choosing $n$ elements (with repetition) from 4 elements is $\frac{(n+3)!}{6n!}$. The maximum number of different distributions is thus the same as Davies and Murphy found, namely, $4 \times \frac{(n+3)!}{6n!}$.

## 4.4 Practical attack

The two previous sections have described how to compute the distributions on the output of two and three adjacent S-boxes after $n$ rounds, given the common key bits from each round. Davies and Murphy described a practical attack using this information. The problem is to find which distribution a known set of plaintext/ciphertext pairs follows.

For 2 adjacent S-boxes we have only 2 distributions to distinguish between, where the XOR of all the common key bits decides which distribution the output follows. For 3 adjacent S-boxes we have up to $4 \times \frac{(n+3)!}{6n!}$ different distributions, and can thus divide the possible keys into the same amount of classes. A successive attack will give information on which class the correct key originates from.

### 4.4.1 2 adjacent S-boxes

The 8-bit output from two adjacent S-boxed can be regarded as a number $i$ between 0 and 255. There are, as we have seen, 2 different distributions for this output. Which distribution the output follows is decided from the XOR of all common key bits $k$. The probability for output $i$ given $k$ is:

$$P_i(k) = \mathbf{Pr}(rst = i \mid k),$$

where $k = \bigoplus_{j=1}^{n} K_j[0,1]$. Davies and Murphy proved

$$\begin{aligned} \mathbf{Pr}(rs \mid K_1 ... K_n) = 2^{-8} + (-1)^k RS,& \\ R = \sum_{\oplus x_i = r} (p_{0x_1} - p_{1x_1})(p_{0x_2} - p_{1x_2}) \cdots (p_{0x_n} - p_{1x_n}),& \\ S = \sum_{\oplus y_i = s} (q_{0y_1} - q_{2y_1})(q_{0y_2} - q_{2y_2}) \cdots (q_{0y_n} - q_{2y_n}).& \end{aligned}$$

We thus have

$$P_i(k) = d + (-1)^k d_i,$$

where $d = 2^{-8}$ and $d_i$ is easily computed. An attacker is given $m$ plaintext/ciphertext pairs where the value $i$ occurs $m_i$ times. We denote our observations by $M = (m_0, ..., m_{255})$. This give us the following likelihood function:

$$L(M;k) = \prod_{i=0}^{255} P_i(k)^{m_i}.$$

The problem is to find out if $k = 0$ or $k = 1$. The Neyman-Pearson Lemma[7] state that if we fix the probability of guessing $k = 1$ when $k = 0$, the likelihood ratio test is the most powerful test, so that the probability of guessing $k = 0$ when $k = 1$ is the smallest.

$$\lambda = \frac{L(M; k = 0)}{L(M; k = 1)} = \prod_{i=0}^{255} \left( \frac{P_i(0)^{m_i}}{P_i(1)^{m_i}} \right).$$

If the correct value for $k$ is 0, then $\lambda$ will likely be bigger than one. If the correct value for $k$ is 1, then $\lambda$ will likely be smaller than one. $\log(\lambda)$ will, however, be either positive or negative, respectively. So if we can decide the sign of $\log(\lambda)$, we can decide if $k = 0$ or $k = 1$.

$$\begin{aligned}\log \lambda &= \sum_{i=0}^{255} m_i \log\left(\frac{P_i(0)}{P_i(1)}\right) = \sum_{i=0}^{255} m_i \log\left(\frac{d+d_i}{d-d_i}\right) = \sum_{i=0}^{255} m_i \log\left(1 + \frac{2d_i}{d-d_i}\right)\\ &\approx \sum_{i=0}^{255} m_i \frac{2d_i}{d-d_i} \approx \frac{2}{d}\sum_{i=0}^{255} m_i d_i,\end{aligned}$$

$$I = \sum_{i=0}^{255} m_i d_i \approx \frac{d}{2}\log\lambda.$$

The logarithm was approximated by the first term in the Taylor expansion of $\log(1+x)$, and the other approximations is justified by $d_i$ being much smaller than $d$. We can thus use the sign of $I$ to decide if $k = 0$ or $k = 1$.

To find the probability of success with a given number of plaintext/ciphertext pairs, we need the expected value $E[I]$ and the variance $Var[I]$ for $I$.

$$\begin{aligned}E[I] &= \sum_{i=0}^{255} d_i E[m_i] \approx \sum_{i=0}^{255} d_i m(d + (-1)^k d_i)\\ &= \sum_{i=0}^{255} d_i m d + \sum_{i=0}^{255} d_i m(-1)^k d_i = (-1)^k mT,\\ Var[I] &\approx \sum_{i=0}^{255} d_i^2 Var[m_i] \approx md\sum_{i=0}^{255} d_i^2 = mdT.\end{aligned}$$

Where $T = \sum_{i=0}^{255} d_i^2$. Davies and Murphy stated that $I$ is approximately normally distributed when $m$ is large. To show this, we introduce a random vector $M_i$, with 256 binary entries and of weight 1 and each vector has probability $P_i(k)$

$$M_i = \begin{cases} (1,0,0,\dots,0), & P_0(k)\\ (0,1,0,\dots,0), & P_1(k)\\ \quad\cdots\cdots & \\ (0,0,\dots,0,1), & P_{255}(k). \end{cases}$$

So

$$M_i = (x_0, ..., x_{255}),$$

where

$$E[x_i] = P_i(k).$$

We thus have

$$
\begin{aligned}
M = (m_0, ..., m_{255}) &= \sum_{j=0}^{m-1} M_j,\\
E[M_i] &= (P_0(k), ..., P_{255}(k)),\\
E[M] &= m \times (P_0(k), ..., P_{255}(k)),\\
Var[M_i] &= V = (v_{ij}) \text{ (Covariance matrix)},\\
(v_{ij}) &= E[x_i x_j] - E[x_i]E[x_j]\\
&= E[x_i x_j] - P_i(k)P_j(k),\\
E[x_i x_j] &= \begin{cases} 0 & i \neq j,\\ P_i(k) & i = j. \end{cases}
\end{aligned}
$$

The Central limit theorem[8] state that as $m$ grows, the following distributions converges towards the multivariate normal distribution:

$$
\begin{gathered}
\frac{M - E[M]}{\sqrt{m}} \to \mathcal{N}(0, Var[M_i])\\
\Longrightarrow\\
\frac{\sum_{i=0}^{255} m_i d_i - E[I]}{\sqrt{m}} \to \mathcal{N}(0, d' \times Var[M_i] \times d'^{\top}).
\end{gathered}
$$

We can thus approximate $I$ by the normal distribution, $\mathcal{N}(0, Var[I])$. We fix $\alpha$ and calculate $m$, the number of needed plaintext/ciphertext, such that $\alpha$ is the probability of choosing $k = 0$ when $k = 1$.

$$
\begin{aligned}
\alpha = \mathbf{Pr}(I \leq 0) &= \Phi\left(\frac{0 - E[I]}{\sqrt{Var[I]}}\right)\\
&= \Phi^{-1}(\alpha) = \frac{-E[I]}{\sqrt{Var[I]}},\\
E[I] &= -\Phi^{-1}(\alpha) \times \sqrt{Var[I]},\\
(-1)^k mT &\approx -\Phi^{-1}(\alpha)\sqrt{mdT},\\
m^2T^2 &\approx \Phi^{-1}(\alpha)^2 \times mdT,\\
m &\approx \frac{\Phi^{-1}(\alpha)^2 \times d}{T}.
\end{aligned}
$$

The S-boxes with the largest value for $T$ is the S-box pair 78. Davies and Murphy calculated $T = 1.32^{-63}$. If we choose $\alpha = 0.0228$ we get

$$
m = \frac{\Phi^{-1}(0.0228)^2 \times 2^{-8}}{1.32^{-63}} = 1.51 \times 2^{56}.
$$

An attack to find two key bits, one for odd rounds, one for even rounds, with success probability $(1-\alpha)^2 = (1-0.0228)^2 = 95.5\%$ will require $1.51 \times 2^{56}$ plaintext/ciphertext pairs. This is about the same complexity as a brute-force attack on the key in DES.

### 4.4.2 3 adjacent S-boxes

We saw in Section 4.3 that the 12-bit output from three adjacent S-boxes can follow a manageable number of distributions. Which distribution the output follows is determined by which class the key originated from. Let us denote the key class by

$$\Psi = (\oplus s, W, \oplus v),$$

where $\oplus s$ is the XOR of the left-most common key bit in each of the $n$ rounds, $\oplus v$ is the XOR of the right-most common key bits, and $W$ is the set of the 2 middle common bits from each of the $n$ rounds, sorted in ascending order. All possible values for $\Psi$ thus represent each of the $4 \times \frac{(n+3)!}{6n!}$ different distributions by Lemmas 4.19 and 4.20.

If the key for DES is chosen randomly, the different values for $\Psi$ are not equally probable. For 2-round DES, for example, there are 4 key bits in $W$. Only one of the $2^4$ possible 4-bit string give $W = \{0, 0\}$, namely $(0, 0, 0, 0)$. There are however, two possible values which give $W = \{1, 3\}$, and that is $(0, 1, 1, 1)$ and $(1, 1, 0, 1)$. We can thus calculate a prior distribution of $\Psi$.

We denote the initial distribution of $W$ by $q(w) = \mathbf{Pr}(W = w)$. We denote the prior distribution of $\oplus s$ and $\oplus v$ by $f(c) = \mathbf{Pr}(\oplus s = c)$ and $g(d) = \mathbf{Pr}(\oplus v = d)$, which is both $\frac{1}{2}$. We thus have the following prior distribution for $\Psi$:

$$p(c, w, d) = \mathbf{Pr}(\Psi = (c, w, d)) = f(c)q(w)g(d).$$

The 12-bit output from 3 adjacent S-boxes can be regards as a number i between 0 and $N - 1 = 4095$. The probability for output $i$, given $\Psi$ is

$$\begin{aligned} P_i(\Psi) &= \mathbf{Pr}(rst = i \mid K = \Psi) \\ &= d + d_i(\Psi), \end{aligned}$$

where $d = 2^{-12}$ and $K$ is any key that, according to Lemmas 4.19 and 4.20, falls into the class $\Psi$. Since we know how to compute $\mathbf{Pr}(rst \mid K_1 ... K_n)$, we can easily compute the bias, $d_i(\Psi)$, for all $\Psi$.

An attacker is given $m$ plaintext/ciphertext pairs where the value $i$ occurs $m_i$ times. We denote our observations by $M = (m_0, ..., m_{N-1})$ We have the following likelihood function:

$$l(M;\Psi) = \prod_{i=0}^{N-1} P_i(\Psi)^{m_i},$$

and the log-likelihood function

$$\begin{aligned} L(M;\Psi) &= \sum_{i=0}^{N-1} m_i \times \log(P_i(\Psi)) \\ &= \sum_{i=0}^{N-1} m_i \times \log(d + d_i(\Psi)) \\ &= \sum_{i=0}^{N-1} m_i \times (\log(1 + \frac{d_i(\Psi)}{d}) + \log(d)) \\ &\approx \sum_{i=0}^{N-1} m_i \times \frac{d_i(\Psi)}{d} + m\log(d) \\ &= \frac{1}{d}\sum_{i=0}^{N-1} m_i d_i(\Psi) - c, \end{aligned}$$

where $c = 12m\log(2)$.

We denote the correct key by $\psi$ and any possible key by $\theta$. We define $I(\theta)$, which is an approximation of the log-likelihood.

$$I(\theta) = \sum_{i=0}^{N-1} m_i d_i(\theta) \approx dL(M;\theta) + c.$$

Davies and Murphy's idea was to do repeated usage of Bayes' Theorem with the same observation $M$, using the old posterior as the new prior. Since $M$ is fixed, the posterior is proportional to the likelihood times prior. Taking the logarithm of both sides of the proportionality, we get the following (up to an additive constant).

$$\begin{aligned} \text{posterior} &\propto \text{likelihood} \times \text{prior}, \\ \log(\text{posterior}) &\approx \log(\text{likelihood}) + \log(\text{prior}), \\ \log(\mathbf{Pr}(\theta \mid M)) &\approx \log(l(M;\theta)) + \log(p(\theta)), \\ \log(\mathbf{Pr}(\theta \mid M)) &\approx I(\theta) + \log(p(\theta)), \end{aligned}$$

where $\propto$ denotes proportional to. Doing this for each possible key candidate, we alter the belief for each key cancidate $\theta$ with $I(\theta)$. We have the following expectation for $I(\theta)$.

$$E[I(\theta)] = \sum_{i=0}^{N-1} d_i(\theta)E[m_i] = \sum_{i=0}^{N-1} d_i(\theta)m(d + d_i(\psi)) = m\sum_{i=0}^{N-1} d_i(\theta)d_i(\psi).$$

The amount Bayes' Theorem will alter our belief for $\theta$ will thus increase as the correlation between $d_i(\theta)$ and $d_i(\psi)$ increases. We then choose the $\theta$ with the highest posterior as our key candidate.

By the same argument as for $I$ in the case for 2 adjacent S-boxes in Section 4.4.1, $I(\theta)$ is approximately normally distributed when $m$ is large. We want to compute how many plaintext/ciphertext pairs we need to get a probability $(1-\alpha)$ for $I(\psi)$ to be positive. For this, we need the the expectation and variance for $I(\psi)$.

$$\begin{aligned} E[I(\psi)] &= m\sum_{i=0}^{N-1} d_i(\psi)d_i(\psi) = m\Gamma(\psi), \\ Var[I(\psi)] &\approx md\Gamma(\psi), \end{aligned}$$

where $\Gamma(\psi) = \sum_{i=0}^{N-1} d_i(\psi)^2$. We can now compute $m$ for every possible $\psi$, by the following calculations. We then choose the highest $m$ to ensure that $I(\psi)$ is likely to be positive regardless of $\psi$.

$$\begin{aligned} \alpha = \mathbf{Pr}(I(\psi) \leq 0) &= \Phi\left(\frac{0 - E[I(\psi)]}{\sqrt{Var[I(\psi)]}}\right) \\ &= \Phi^{-1}(\alpha) = \frac{-E[I(\psi)]}{\sqrt{Var[I(\psi)]}}, \\ E[I(\psi)] &= -\Phi^{-1}(\alpha) \times \sqrt{Var[I(\psi)]}, \\ m\Gamma(\psi) &= -\Phi^{-1}(\alpha)\sqrt{md\Gamma(\psi)}, \\ m^2\Gamma(\psi)^2 &= \Phi^{-1}(\alpha)^2 md\Gamma(\psi), \\ m &= \frac{\Phi^{-1}(\alpha)^2 \times d}{\Gamma(\psi)}. \end{aligned}$$

Davies and Murphy computed $m$ for all triplets, and found that the S-box triplets 678 and 781 require $1.51 \times 2^{56}$ plaintext/ciphertext pairs. This is the same amount that is required for S-box pair 78, and about the same complexity as a brute-force attack on the key in DES.

# Chapter 5

# Linear dependencies between distributions

(4.3) in Chapter 4 shows two relations between the rows in the tables for left and right distributions. They turn out to have a great impact on relations between other distributions in DES. In this chapter we will explore such relations. We repeat (4.3) as a lemma.

**Lemma 5.1.**

$$p_{(x\oplus 2)r} + p_{xr} = \frac{1}{32}, \qquad q_{(x\oplus 1)r} + q_{xr} = \frac{1}{32}.$$

More specifically, this chapter will discuss linear dependencies between distributions. A distribution on a $n$-bit output can be represented by a $2^n$ vector, $v = (v_0, ..., v_{2^n-1})$, where each entry $v_i$ represents the probability that the output equals $i$. By linear dependencies between distributions, we mean linear dependencies between such vectors.

For a given output, we usually have multiple possible distributions. For example, the output from two adjacent S-boxes, given the common key bits, have 4 different distributions. Which distribution the output follows depend on the common key bits. Since we represent distributions by vectors, we can represent the set of all distributions for a given output by a matrix (**distribution matrix**). In the example above, the row index represents the value for the common key bits and the row vector represents the distribution.

Representing all distributions by matrices gives us the ability to effortlessly use terms like rank, kernel, nullspace, etc. on distributions. When we refer to the rank of distributions, we mean the rank of the distribution matrix.

## 5.1 Relations in left and right distributions

Lemma 5.1 trivially gives us one relation each for the left and right distributions. The same relation can be found by looking at the tables for left/right distributions in Appendix A. Take, for example, the table for the right distribution for an S-box. If we add the first two numbers in each column and subtract the last two numbers in the same column, we get an all zero row. We thus have:

$$\begin{aligned} p_{xr} - p_{(x\oplus 1)r} + p_{(x\oplus 2)r} - p_{(x\oplus 3)r} &= 0, \\ q_{xr} + q_{(x\oplus 1)r} - q_{(x\oplus 2)r} - q_{(x\oplus 3)r} &= 0. \end{aligned} \tag{5.2}$$

The above equations hold for $x = 0, 1, 2, 3$, so they are actually 4 relations each. All of them are however, either equal to or the negative of the following relations:

$$p_{0r} - p_{1r} + p_{2r} - p_{3r} = 0 \quad \text{and} \quad q_{0r} + q_{1r} - q_{2r} - q_{3r} = 0. \tag{5.3}$$

The tables for the left/right distributions can be seen as matrices. We know that there exists at least one linear relation between the rows in these tables. The rank of each matrix must therefore be at most 3. Computing the rank of the 16 matrices, one for the left and one for the right distribution for each of the 8 S-boxes, we see that it is 3 for all of them. The linear relations in (5.2) can be represented by row vectors.

$$C^1 = (1, -1, 1, -1) \qquad \text{and} \qquad C^2 = (1, 1, -1, -1).$$

We have that

$$\sum_A C^1_A \times p_{(x\oplus A)r} = 0 \qquad \text{and} \qquad \sum_A C^2_A \times q_{(x\oplus A)r} = 0. \tag{5.4}$$

If we multiply $C^1$ or $C^2$ with the matrix for the left or right distributions, respectively, the result is a zero vector.

## 5.2 Dependencies in distributions of 2 adjacent S-boxes

As described in Section 4.2, the 2 right most input bits of $S_i$ XORed with the 2 left most input bits of $S_{i+1}$ is the XOR of some of the round key bits, called common key bits. The probability for output $rs$ from $S_i$ and $S_{i+1}$ given the common key bits $A$ is

$$\mathbf{Pr}(rs \mid A) = 4 \times \sum_x p^{(i)}_{(x\oplus A)r} q^{(i+1)}_{xs} = 4 \times \sum_x p^{(i)}_{xr} q^{(i+1)}_{(x\oplus A)s}.$$

From Lemma 4.6, we can find two linear dependencies. The following equations show these dependencies represented as sums of distributions.

$$\mathbf{Pr}(rs \mid 0) - \mathbf{Pr}(rs \mid 3) = 0,$$
$$\mathbf{Pr}(rs \mid 1) - \mathbf{Pr}(rs \mid 2) = 0,$$

for any $rs$. Represented as vectors, we have

$$(1, 0, 0, -1) \qquad \text{and} \qquad (0, 1, -1, 0). \tag{5.5}$$

We can also use (5.4) to get two linear dependencies between distributions on $rs$. The vectors representing these dependencies will be the same as $C^1$ and $C^2$.

$$\begin{aligned}\sum_A C_A \times \mathbf{Pr}(rs \mid A) &= \sum_A C_A \times \left(4 \times \sum_x p^{(i)}_{(x\oplus A)r} q^{(i+1)}_{xs}\right) \\ &= 4 \times \sum_A \sum_x C_A \times p^{(i)}_{(x\oplus A)r} q^{(i+1)}_{xs} \\ &= 4 \times \sum_x q^{(i+1)}_{xs} \left(\sum_A C_A \times p^{(i)}_{(x\oplus A)r}\right) \\ &= 4 \times \sum_x q^{(i+1)}_{xs} \times 0 = 0,\end{aligned} \tag{5.6}$$

for $C = C^1, C^2$ and any $rs$. We now have four linear dependencies, two from (5.5) and two from (5.6). All these dependencies can be represented by one of the row vectors in the matrix $R$ below.

$$R = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$$

We want to use $R$ to compute an upper bound on the rank of the distribution matrix. For this, we introduce the following lemma.

**Lemma 5.7.**
*Let $M$ be a matrix with $k$ rows and $R$ be any matrix such that $R \times M = 0$. Then*

$$\mathit{rank}(M) \leq k - \mathit{rank}(R).$$

*Proof:*
By the Rank-nullity theorem we have that

$$\operatorname{rank}(M^\top) + \operatorname{nul}(M^\top) = k,$$

where $M^\top$ denote the transpose of $M$. The row vectors in the basis of $R$ must be in the left null space of $M$. So

$$\begin{aligned}&\mathrm{nul}(M^\top) \geq \mathrm{rank}(R),\\ &\mathrm{nul}(M^\top) = k - \mathrm{rank}(M) \geq \mathrm{rank}(R),\\ &\mathrm{rank}(M) \leq k - \mathrm{rank}(R).\end{aligned}$$

□

$\mathbf{Pr}(rs \mid A)$ has four distributions, so the distribution matrix thus have 4 rows. The rank of $R$ is 2, so the rank of the distribution matrix must be $\leq 4 - 2 = 2$. We have computed the rank to be exactly 2, so we know that we have found all dependencies in the distribution matrix. We already knew from the result of Chapter 4 that there are only two different distributions, so the rank could not have been larger than 2.

## 5.3 Relations in $Q$ distributions

Definition 4.15 defines the $Q$ distribution as the distribution of $(x_5, x_4, x_1, x_0, y_3, y_2, y_1, y_0)$. It also defines

$$Q^{(i)}_{xyr} = \mathbf{Pr}(x_5x_4 = x,\ x_1x_0 = y,\ y_3y_2y_1y_0 = r),$$

where $x_i$ and $y_i$ are input and output bits of the S-box $S_i$. Appendix B lists the $Q$ distribution for each S-box, where each row index represent $x||y$ and the column index represent $r$. These tables can be seen as matrices. We will find relations between its rows. By definition we have

**Lemma 5.8.**

$$\sum_x Q_{xyr} = p_{yr} \qquad \textit{and} \qquad \sum_y Q_{xyr} = q_{xr}.$$

By Lemma 5.1 and Lemma 5.8:

$$\sum_x Q_{xyr} + Q_{x(y\oplus 2)r} = \frac{1}{32} \qquad \sum_x Q_{x(y\oplus 1)r} + Q_{x(y\oplus 3)r} = \frac{1}{32} \qquad \text{for any } y, \tag{5.9}$$

and

$$\sum_y Q_{xyr} + Q_{(x\oplus 1)yr} = \frac{1}{32} \qquad \sum_y Q_{(x\oplus 2)yr} + Q_{(x\oplus 3)yr} = \frac{1}{32} \qquad \text{for any } x.$$

Subtracting one of the above from another will be a linear relation in a $Q$ distribution. There are 6 ways to choose two of the above sums (unordered and with no repetition). We computed 6 linear relations and the rank of the corresponding relation matrix which was found to be 3. We have selected three independent relations, out of the 6, and listed them as vectors in Table 5.10.

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C^3$ | = | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | -1 | 0 | -1 | 0 | -1 | 0 | -1 | 0 |
| $C^4$ | = | 1 | -1 | 1 | -1 | 1 | -1 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $C^5$ | = | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |

TABLE 5.10: $C^3$, $C^4$ and $C^5$

Since the rank of the relation matrix is 3, the rank of the tables in Appendix B can be at most $16 - 3 = 13$. We have computed the rank for each table which indeed is at most 13. The rank of $Q_{xyr}$ for each S-box is listed in Table 5.11. For some of the S-boxes the rank is lower than 13. Some of these ”extra” relations comes from equal rows in the tables, but not necessarily all of them. We have not investigated this any further.

| | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ | $S_8$ |
|---|---|---|---|---|---|---|---|---|
| **# different rows** | 15 | 12 | 16 | 12 | 16 | 16 | 16 | 15 |
| **Rank** | 12 | 10 | 13 | 10 | 13 | 13 | 13 | 12 |

TABLE 5.11: Rank and number of different rows in each $Q$ distribution

We now show how the vector $C^3$ was computed. The same method was used for $C^4$ and $C^5$. First, we selected two of the relations in (5.9).

$$\sum_{x'} Q_{x'yr} + Q_{x'(y\oplus 2)r} = \frac{1}{32}, \qquad \sum_{y'} Q_{xy'r} + Q_{(x\oplus 1)y'r} = \frac{1}{32}.$$

We subtract them from each-other to get a linear dependency.

$$0 = \left(\sum_{y'} Q_{xy'r} + Q_{(x\oplus 1)y'r}\right) - \left(\sum_{x'} Q_{x'yr} + Q_{x'(y\oplus 2)r}\right),$$

$$0 = \sum_{x',y'} Q_{xy'r} + Q_{(x\oplus 1)y'r} - Q_{x'yr} - Q_{x'(y\oplus 2)r}.$$

Substitute $x' = x \oplus A$, $y' = y \oplus B$ and get

$$0 = \sum_{A,B} Q_{x(y\oplus B)r} + Q_{(x\oplus 1)(y\oplus B)r} - Q_{(x\oplus A)yr} - Q_{(x\oplus A)(y\oplus 2)r}, \tag{5.12}$$

for any $x$ and $y$.

Set $x = y = 0$

$$\begin{aligned} 0 &= \sum_{A,B} Q_{0Br} + Q_{1Br} - Q_{A0r} - Q_{A2r}, \\ 0 &= Q_{01r} + Q_{03r} + Q_{11r} + Q_{13r} - Q_{20r} - Q_{22r} - Q_{30r} - Q_{32r}. \end{aligned} \tag{5.13}$$

The coefficient vector for (5.13) is $C^3$, which is what we wanted to show. Each entry, $C^3_{AB}$, is the coefficient at $Q_{ABr}$, and also $Q_{(x\oplus A)(y\oplus B)r}$ (since we derived (5.13) from (5.12) by setting $x = y = 0$). The same procedure was used to find $C^4$ and $C^5$. So

$$\sum_{A,B} C_{AB} \times Q_{(x\oplus A)(y\oplus B)r} = 0, \qquad \text{for} C = C^3, C^4, C^5. \tag{5.14}$$

## 5.4 Dependencies in QDES

In this section we study multiple adjacent S-boxes in a modified version of DES, called QDES. The round function $F$ in QDES consists of $n$ adjacent S-boxes (instead of 8 as in DES). The number of input-bits to $F$ is thus $4n$ and the number of key-bits is $6n$. The expansion function $E' : \mathbb{Z}_2^{4n} \to \mathbb{Z}_2^{6n}$ "wraps around" the first and last bit in the same way that the original $E$ does in DES (see Table 2.3).

We want to find the distributions of the output from all S-boxes in 1-round QDES, given common key bits for all pairs of neighbouring S-boxes. We also want to find linear dependencies between these distributions and thus get an upper bounds on the rank of the distribution matrix. (5.16) and (5.17) show how to compute the distribution of the output for QDES with 2 and $n$ S-boxes. Figure 5.15 illustrates the S-boxes in QDES where $n = 3$.

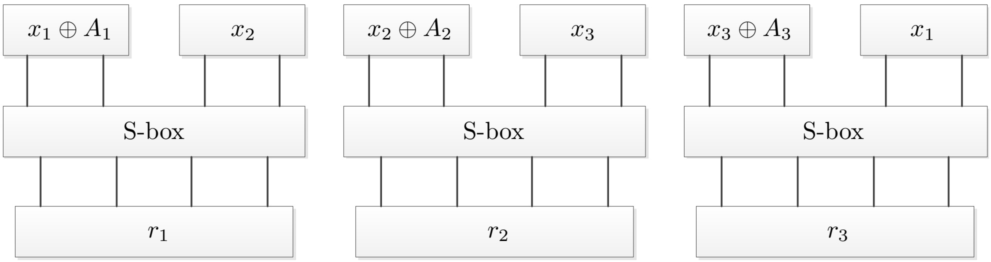


FIGURE 5.15: Modified version of DES with 3 S-boxes

$$\begin{aligned}
\mathbf{Pr}(rs \mid AB) &= \frac{\sum_{x,y} Q^{(1)}_{(x\oplus A)yr} Q^{(2)}_{(y\oplus B)xs}}{\mathbf{Pr}(AB)} \qquad (5.16)\\
&= 16 \times \sum_{x,y} Q^{(1)}_{\underline{(x\oplus A)yr}} Q^{(2)}_{\underline{(y\oplus B)xs}}\\
&= 16 \times \sum_{x,y} Q^{(1)}_{\underline{(x\oplus A)(y\oplus B)r}} Q^{(2)}_{\underline{yxs}}\\
&= 16 \times \sum_{x,y} Q^{(1)}_{\underline{xyr}} Q^{(2)}_{\underline{(y\oplus B)(x\oplus A)s}},\\
\mathbf{Pr}(r_1...r_n \mid K_1...K_n) &= 2^{2n} \times \sum_{x_1...x_n} Q^{(1)}_{(x_1\oplus K_1)x_2r_1} \cdots Q^{(n)}_{(x_n\oplus K_n)x_1r_n}. \qquad (5.17)
\end{aligned}$$

To compute the distributions for QDES, we only need the $Q$ distribution for each S-box. We can use (5.14) to compute dependencies between the distributions by the following lemma.

**Lemma 5.18.**

$$\sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times \boldsymbol{Pr}(r_1...r_n \mid K_1...K_n) = 0, \qquad \textit{for } C = C^3, C^4, C^5,$$

*where $K_i$ are the common bits between $S_{i-1}$ and $S_i$, and $K_{i+1}$ are the common bits between $S_i$ and $S_{i+1}$.*

*Proof:*

$$\begin{aligned}
&\sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times \mathbf{Pr}(r_1...r_n \mid K_1...K_n)\\
&= 2^{2n} \times \sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times \sum_{x_1...x_n} \left(Q^{(1)}_{(x_1\oplus K_1)x_2r_1} \cdots Q^{(n)}_{(x_n\oplus K_n)x_1r_n}\right)\\
&= 2^{2n} \times \sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times \sum_{x_1...x_n} \left( \cdots Q^{(i)}_{\underline{(x_i\oplus K_i)x_{i+1}}r_i} Q^{(i+1)}_{\underline{(x_{i+1}\oplus K_{i+1})x_{i+2}}r_{i+1}} \cdots \right)\\
&= 2^{2n} \times \sum_{x_1...x_n} \sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times \left( \cdots Q^{(i)}_{\underline{(x_i\oplus K_i)(x_{i+1}\oplus K_{i+1})}r_i} Q^{(i+1)}_{\underline{x_{i+1}x_{i+2}}r_{i+1}} \cdots \right)\\
&= 2^{2n} \times \sum_{x_1...x_n} (\ \cdots\ ) \times \left( \sum_{K_iK_{i+1}} C_{K_iK_{i+1}} \times Q^{(i)}_{\underline{(x_i\oplus K_i)(x_{i+1}\oplus K_{i+1})}r_i} \right)\\
&= 2^{2n} \times \sum_{x_1...x_n} (\ \cdots\ ) \times 0 = 0.
\end{aligned}$$

□

The equation in Lemma 5.18 hold for all possible values of $K_j$ $(j \neq i, i+1)$. There is $n-2$ such $K_j$, with a total of $2^{2n-4}$ different values. There are three vectors $C$, and $n$ choices of $i$. This gives a total of $3n \times 2^{2n-4}$ linear dependencies in QDES that comes from (5.14).

We have computed all $3n \times 2^{2n-4}$ linear dependencies from Lemma 5.18 and the rank of the corresponding relation matrix for $2 \leq n \leq 7$. The complexity for computing the relation matrix for $n = 8$ (full DES) was too high. The following table lists the result. $M$ is the distribution matrix and $R$ is the relation matrix.

| **# S-boxes, $n$** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|
| # rows $M$ | 16 | 64 | 256 | 1024 | 4096 | 16384 |
| # rows $R$ | 6 | 36 | 192 | 960 | 4608 | 21504 |
| rank($R$) | 6 | 33 | 158 | 715 | 3123 | 9091 |
| rank($M$) $\leq$ | 10 | 31 | 98 | 309 | 973 | 7293 |

Table 5.19: Rank of and number of rows in $M$ and $R$

We have computed the rank of the distribution matrix for $2 \leq n \leq 4$, and found that the relations described above are all the linear dependencies for these values of $n$. rank($M$) for $5 \leq n \leq 7$ are only upper bounds.

## 5.5 Dependencies in distributions of 3 adjacent S-boxes

In this section, and for the rest of the thesis, we will work with full DES again. We are going to study 3 adjacent S-boxes like in Section 4.3. (4.17) shows how to compute the distribution on output $rst$ given the common key bits $AB$. We repeat (4.17) here for the convenience of the reader.

$$\begin{aligned}\mathbf{Pr}(rst \mid AB) &= 16 \times \sum_{x,y} p^{(i)}_{(x\oplus A)r} Q^{(i+1)}_{xys} q^{(i+2)}_{(y\oplus B)t} \\ &= 16 \times \sum_{x,y} p^{(j)}_{xr} Q^{(j+1)}_{(x\oplus A)(y\oplus B)s} q^{(j+2)}_{yt}.\end{aligned} \tag{4.17}$$

To find linear dependencies between distributions for the output of 3 adjacent S-boxes, we can use (5.4) and (5.14). We have derived (5.21), (5.22) and (5.23) from them. $C^1$ and $C^2$ are defined by (5.4) and $C^3$, $C^4$ and $C^5$ are listed in Table 5.10. We can derive 11 linear dependencies from the three equations.

**Lemma 5.20.**

*for any $B$*
$$\sum_A C^1_A \times \boldsymbol{Pr}(rst \mid AB) = 0, \tag{5.21}$$

*for any $A$*
$$\sum_B C^2_B \times \boldsymbol{Pr}(rst \mid AB) = 0, \tag{5.22}$$

*for $C \in \{C^3, C^4, C^5\}$*
$$\sum_{AB} C_{AB} \times \boldsymbol{Pr}(rst \mid AB) = 0. \tag{5.23}$$

*Proof:*

We will prove (5.21):

$$\begin{aligned}\sum_A C_A^1 \times \mathbf{Pr}(rst \mid AB) &= 16 \times \sum_A C_A^1 \times \left(\sum_{x,y} p^{(j)}_{(x\oplus A)r} Q^{(j+1)}_{xys} q^{(j+2)}_{(y\oplus B)t}\right)\\ &= 16 \times \sum_{x,y}\sum_A C_A^1 \times \left(p^{(j)}_{(x\oplus A)r} Q^{(j+1)}_{xys} q^{(j+2)}_{(y\oplus B)t}\right)\\ &= 16 \times \sum_{x,y} Q^{(j+1)}_{xys} q^{(j+2)}_{(y\oplus B)t} \times \left(\sum_A C_A^1 \times p^{(j)}_{(x\oplus A)r}\right)\\ &= 16 \times \sum_{x,y} Q^{(j+1)}_{x(y\oplus B)s} q^{(j+2)}_{yt} \times 0 = 0.\end{aligned}$$

Similarly we prove (5.22). We will prove (5.23).

$$\begin{aligned}\sum_{AB} C_{AB} \times \mathbf{Pr}(rst \mid AB) &= 16 \times \sum_{AB} C_{AB} \times \left(\sum_{x,y} p^{(j)}_{xr} Q^{(j+1)}_{(x\oplus A)(y\oplus B)s} q^{(j+2)}_{yt}\right)\\ &= 16 \times \sum_{x,y}\sum_{AB} C_{AB} \times \left(p^{(j)}_{xr} Q^{(j+1)}_{(x\oplus A)(y\oplus B)s} q^{(j+2)}_{yt}\right)\\ &= 16 \times \sum_{x,y} p^{(j)}_{xr} q^{(j+2)}_{yt} \times \left(\sum_{AB} C_{AB} \times Q^{(j+1)}_{(x\oplus A)(y\oplus B)s}\right)\\ &= 16 \times \sum_{x,y} p^{(j)}_{xr} q^{(j+2)}_{yt} \times 0 = 0.\end{aligned}$$

□

We have computed 11 linear dependencies from (5.21), (5.22) (5.23). The rank of the relation matrix is 10. We have also computed the rank of the distribution matrix which is 6. Since there are 16 distributions in total, we know that we have found all the linear relations between the distributions for the output of 3 adjacent S-boxes.

## 5.6 3 adjacent S-boxes after multiple rounds

So far in this chapter, we have only discussed dependencies between distributions of 1-round DES. In this section we will study dependencies between the distributions of the output from 3 adjacent S-boxes after multiple rounds.

The ciphertext in DES is the XOR of outputs from multiple round functions, $F(X, K)$. We ignore the initial permutation, final permutation and the permutations in the round function. The XOR of the left-most 32 bits in the plaintext and the right-most 32 bits in the ciphertext is thus an XOR of eight outputs from the round function. Similarly, the XOR of the right-most 32 bits in the plaintext and the left-most 32 bits in the ciphertext is also an XOR of eight outputs from the round function.

As discussed in Section 4.3, the distribution of the XOR of $n$ outputs from the round function is the $n$-fold convolution of $\mathbf{Pr}(rst \mid K)$. The following equations show how to compute the distribution for 2 and $n$ rounds.

$$\mathbf{Pr}(rst \mid AB, CD) = \sum_{abc} \mathbf{Pr}(abc \mid AB)\mathbf{Pr}(rst \oplus abc \mid CD), \tag{5.24}$$

$$\mathbf{Pr}(rst \mid K_1 ... K_n) = \sum_{\bigoplus I_i = rst} \mathbf{Pr}(I_1 \mid K_1) \text{ ... } \mathbf{Pr}(I_n \mid K_n), \tag{5.25}$$

where $abc$ and $I_i$ are 12-bit intermediate values.

### 5.6.1 Dependencies from $\mathbf{C^1}$, ..., $\mathbf{C^5}$

(5.4) and (5.14) are used to compute dependencies between the distributions of $rst$ after $n$ rounds. The following equation hold when $R$ is any of the 4-bits $K_i$, the left-most 2 bits in any $K_i$ or the right-most 2 bits in any $K_i$ $(i = 1, ..., n)$.

**Lemma 5.26.**

$$\sum_R C_R \times \boldsymbol{Pr}(rst \mid K_1 ... K_n) = 0,$$

*where $C_R = C^1$ if $R$ is the two left-most bits in $K_i$, $C_R = C^2$ if $R$ is the two right-most bits in $K_i$ and $C_R = C^3, C^4$, or $C^5$ if $R$ is any $K_i$.*

*Proof:*

$$\begin{aligned}
&\sum_R C_R \times \mathbf{Pr}(rst \mid K_1 ... K_n) \\
&= \sum_R C_R \times \left( \sum_{\bigoplus I_i = rst} \mathbf{Pr}(I_1 \mid K_1) \text{ ... } \mathbf{Pr}(I_n \mid K_n) \right) \\
&= \sum_{\bigoplus I_i = rst} \sum_R C_R \times \mathbf{Pr}(I_1 \mid K_1) \text{ ... } \mathbf{Pr}(I_n \mid K_n).
\end{aligned}$$

Assume $R = K_j$, without loss of generality

$$\begin{aligned}
&= \sum_{\bigoplus I_i = rst} \sum_{K_j} C_{K_j} \times (\text{ ... } \mathbf{Pr}(I_j \mid K_j) \text{ ... }) \\
&= \sum_{\bigoplus I_i = rst} (\text{ ... }) \times \left( \sum_{K_j} C_{K_j} \times \mathbf{Pr}(I_j \mid K_j) \right) \\
&= \sum_{\bigoplus I_i = rst} (\text{ ... }) \times 0 = 0.
\end{aligned}$$

The last step comes from Lemma 5.20.

□

### 5.6.2 Number of different distributions after $n$ rounds

Davies and Murphy found that there are at most $4 \times \frac{(n+3)!}{6n!}$ (660 in the case of full DES) different distributions after $n$ founds. We want to find and compute exactly how many different distributions there are.

The key schedule in DES might duplicate key bits in $K_i$ ($0 \leq i < n$). If, for example, $\mathbf{Pr}(rst \mid K_1...K_8)$ is the distributions for the S-box triplet 812 in 8-round DES, then $K_1[2] = K_2[3]$ and $K_4[2] = K_5[3]$. The common key bits for each round are thus dependent for some triplets. We do not know if it is possible to use this to reduce the upper bound on the number of different distributions.

As Davies and Murphy pointed out, the $n$-fold convolution of $\mathbf{Pr}(rst \mid K)$ can efficiently be computed using fast Walsh-Hadamard transform. We computed all $4 \times \frac{(n+3)!}{6n!}$ distributions for $n = 2, ..., 8$. Table 5.27 list the number of different distributions for all triplets of S-boxes for 1 to 8 rounds.

As the table shows, the number of different distributions for S-box triplet 456 is for some reason only increasing by 8 for each extra round, whereas the other triplets has exactly $4 \times \frac{(n+3)!}{6n!}$ different rows.

| n | Max bound | 123 | 234 | 345 | 456 | 567 | 678 | 781 | 812 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 2 | 40 | 40 | 40 | 40 | **24** | 40 | 40 | 40 | 40 |
| 3 | 80 | 80 | 80 | 80 | **32** | 80 | 80 | 80 | 80 |
| 4 | 140 | 140 | 140 | 140 | **40** | 140 | 140 | 140 | 140 |
| 5 | 224 | 224 | 224 | 224 | **48** | 224 | 224 | 224 | 224 |
| 6 | 336 | 336 | 336 | 336 | **56** | 336 | 336 | 336 | 336 |
| 7 | 480 | 480 | 480 | 480 | **64** | 480 | 480 | 480 | 480 |
| 8 | 660 | 660 | 660 | 660 | **72** | 660 | 660 | 660 | 660 |

TABLE 5.27: Number of different distributions for 3 S-boxes after $n$ rounds

### 5.6.3 Dependencies between distributions after $n$ rounds

How many linear dependencies can we calculate from Lemma 5.26? First consider the dependencies from $C^3$, $C^4$ and $C^5$. $\sum_{K_i}[C_{K_i} \times \mathbf{Pr}(rst \mid K_1...K_n)] = 0$ hold for all possible values for the $(n-1)$ other $K_j$ ($j \neq i$). There is $2^{4n-4}$ such values. There are $n$ choices for $i$, and three choices for $C$. This gives a total of $3n \times 2^{4n-4}$ linear dependencies.

Next, we consider linear dependencies from $C^1$ and $C^2$. These vectors have only 4 entries, so we only sum over the two left or right-most bits in $K_i$. There are $n$ possible choices of $i$. Lemma 5.26 holds for the four possible values for the remaining 2 bits in $K_i$ and for all possible values for the $(n-1)$ other $K_j$ $(j \neq i)$. The total number of linear dependencies is thus $8n \times 2^{4n-4}$.

The total number of linear dependencies from Lemma 5.26 is $11n \times 2^{4n-4}$ which becomes impractical after 3 rounds as the relation matrix for 4 rounds would have $\approx 2^{33.5}$ entries. To find a upper bound on the rank of the distribution matrix, we need to compute the rank of the relation matrix. So we want to find linear dependencies between the at most $N = 4 \times \frac{(n+3)!}{6n!}$ different distributions.

Lemmas 4.19 and 4.20 gives a structured way to select $N$ distributions such that each different distributions are selected at least once. Algorithm 3 generates a set $D$ that contains $N$ vectors of length $n$, where each entry in the vectors are a 4 bit value. One vector can be used as $K_1, ..., K_n$ and thus represents one distribution.

**Algorithm 3**

**Step 1:** Let $S = \{0, 1, 2, 3\}$, $N = \frac{(n+3)!}{6n!}$ and let G be a generator which generates all unordered tuples of $S$ of length $n$. Each tuple should be sorted in ascending order.

**Step 2:** Prepare a list $D$ where we will store the common key bits representing the distributions.

**Step 3:** Use G to generate the $N$ possible tuples.

**Step 4:** For each tuple $t$ from **Step 3**:

- Extend each of the two bit values $t_i$, to four bit values $T_i$ by appending a zero bit left to the left-most bit and a zero bit to the right of the right-most bit in $t_i$. That is, we have a vector $T$ with quaternary elements and $T_i = (0, t_i[1], t_i[0], 0)$.
- Compute $K^\alpha$ by copying $T$ to $K^\alpha$ and set $K^\alpha_n = T_n \oplus \alpha$ for each $\alpha \in \{0, 1, 8, 9\}$ (each possible value for the two outer bits) and add $K^\alpha$ to $D$.

The two underlined statements in Algorithm 3 give us a structured way to compute linear dependencies between the distributions represented by the vectors in $D$. Algorithm 4 show how we computed linear dependencies between the different distributions.

**Algorithm 4**

**Step 1:** Let $D$ be the set of common key bits representing the different distributions from Algorithm 3

**Step 2:** Prepare a storage space $R$ where we will store the dependencies.

**Step 3:** For each $K_1...K_n$ in $D$:

- Repeat the next 5 steps for all $(0 \leq i < n)$
- Make a list of pairs,
  $L = \{(C_{K_i}, K_1...K_n) \mid \text{for all 16 values for } K_i\}$.
- Because we have changed the value of $K_i$, the 16 different distribution represented in $L$ is not necessarily in $D$. The second co-ordinate in $L_j$ can be mapped to different value $K'_1....K'_n$ representing an equal distribution by sorting the set of all middle bits in each $K_i$ and "moving" the outer bits in any $K_k$ to $K_n$. (Lemmas 4.19 and Lemma 4.20). We know from how we designed $D$ that $K'_1....K'_n$ will be in $D$. We now have:
  $L = \{(C_{K_i}, K'_1...K'_n) \mid \text{for all 16 values for } K_i\}$.
- Make a vector, $V$, of length $N$ and set all entries to zero.
- $L$ is a list of distributions in $D$ and coefficients to these distributions such that the sum of all these is the all zero vector. For each of these pairs in $L$, find the index of $K'_1...K'_n$ in $D$ and set the value at the same index in $V$ to the value in the fist co- ordinate in the pair from $L$ (the coefficient).
- Check if $V$ is in $R$ and append it to $R$ if not already there.

The 2nd and 3rd points in step 3 in Algorithm 4 assumes that $C$ is one of $C^3$, $C^4$ or $C^5$. The procedure is easily adjusted for $C^1$ and $C^2$. The maximum bound on the number of relation in $R$ is $4 \times \frac{(n+3)!}{6n!} \times 56n$. The total number of dependencies we can calculate from Lemma 5.26 is more than $10^{n-3.3}$ times as much.

It turns out that in practice, many of the dependencies generated in step 3 are equal to dependencies already in $R$, and therefore not added to $R$. For example, $R$ contains 3342 different dependencies for $n = 8$ instead of the maximum bound, $\approx 3 \times 10^5$, or the number of possible dependencies from Lemma 5.26 which is $\approx 10^{10}$. We ran both algorithms for $n = 2, ..., 10$ simultaneously and computed the rank of all relation matrices in less than five minutes using approximately 1GB of memory.

The maximum number of different distributions minus the rank of the relation matrix computed by Algorithms 3 and 4 is a maximum bound on the rank of the distribution matrix. Table 5.28 lists the upper bound and the actual rank for each triplet for 1-8 rounds. The upper bound on the rank of the relation matrices for 9 and 10 rounds is 58 and 69, respectively.

Again we see that the triplet 456 is abnormal. There are also a few unknown dependencies between distributions for triplet 123 and 345 after 6 rounds or more, which we have not investigated any further.

| **n** | **Upper bound** | **123** | **234** | **345** | **456** | **567** | **678** | **781** | **812** |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 2 | 9 | 9 | 9 | 9 | **7** | 9 | 9 | 9 | 9 |
| 3 | 13 | 13 | 13 | 13 | **8** | 13 | 13 | 13 | 13 |
| 4 | 18 | 18 | 18 | 18 | **9** | 18 | 18 | 18 | 18 |
| 5 | 24 | 24 | 24 | 24 | **10** | 24 | 24 | 24 | 24 |
| 6 | 31 | 30 | 31 | 29 | **11** | 31 | 31 | 31 | 31 |
| 7 | 39 | 36 | 39 | 34 | **12** | 39 | 39 | 39 | 39 |
| 8 | 48 | 42 | 48 | 39 | **13** | 48 | 48 | 48 | 48 |

Table 5.28: Rank of pQq after $n$ rounds

The distribution matrices for triplets 234, 567, 678, 781, and 812 have the same rank as the upper bound. For these triplet, Algorithms 3 and 4 have thus found all linear dependencies between distributions on $rst$, for $\leq 8$ rounds. They originate from properties common to all S-boxes.

# Chapter 6

# Conclusion

In this thesis, we started with a survey of Matsui's linear cryptanalysis and Davies and Murphy's analysis of pairs and triplets in DES. Davies and Murphy showed that the distributions of the output from two and three adjacent S-boxes are non-uniform, and easily computed given some of the key bits.

The output from two adjacent S-boxes in DES has only two different distributions, and is determined by the XOR of what we called the common key bits in each round. The number of different distributions of the output from three adjacent S-boxes is low and determined by the common key bits.

Davies and Murphy gave a known-plaintext attack exploiting the non-uniform distributions. The attack on S-box pairs and triplets both has about the same complexity, as brute-force attack on the key in DES as $\approx 2^{56}$ known plaintext/ciphertext pairs is needed.

There are linear dependencies between those distributions. We defined a distribution matrix and relation matrix. These were used to compute an upper bound on the number of linear independent distributions and to show that, in some cases, all linear dependencies were found.

All linear dependencies for all S-box pairs and triplets in 1-round DES were found. We looked at a different version of DES (QDES) with fewer S-boxes, in an attempt to compute the rank of all distributions for the output of all S-boxes. All linear dependencies for QDES with up to 4 S-boxes were found, and an upper bound on the rank of the distributions for QDES with up to 7 S-boxes was computed. The complexity for computing dependencies between the distributions for QDES with 8 S-boxes (full DES) was too high.

Davies and Murphy showed that the maximum number of different distributions for the output of any S-box triplet in full DES is manageable and that they are easily computed. We found an upper bound on the rank of these distributions, and computed linear dependencies between them.

It was found that S-box triplet 456 has a lower number of different distributions, and thus has a greater number of linear dependencies than the other 7 triplets. All linear dependencies were found for 5 of the S-box triplets in full DES. The linear dependencies originate from properties common to of DES' S-boxes.

We do not know whether or not the linear dependencies described in this thesis can be used to enhance an attack on DES. We find the results very interesting, and hope to answer this question some day in the future.

# Bibliography

[1] Donald Davies and Sean Murphy. Pairs and Triplets of DES S-boxes. *Journal of Cryptology*, 8(1):1–25, 1995.

[2] M. Matsui. Linear Cryptanalysis of DES Cipher (I). *Journal of Cryptology.*

[3] NIST FIPS PUB. 46-3. Data Encryption Standard. *Federal Information Processing Standards, National Bureau of Standards, US Department of Commerce*, 1977.

[4] Mitsuru Matsui. On Correlation Between the Order of S-boxes and the Strength of DES. In *Advances in Cryptology-EUROCRYPT'94*, pages 366–375. Springer, 1995.

[5] Eli Biham and Alex Biryukov. An Improvement of Davies' Attack on DES. *Journal of Cryptology*, 10(3):195–205, 1997.

[6] Sebastien Kunz-Jacques and Frederic Muller. New Improvements of Davies-Murphy Cryptanalysis. In *Advances in Cryptology-ASIACRYPT 2005*, pages 425–442. Springer, 2005.

[7] Maurice George Kendall and Stuart Alan. The Advanced Theory of Statistics. 2, 1961.

[8] Willliam Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. John Wiley & Sons, 2008.

# Appendix A

# Right and left distribution for each S-box

The following tables show the normalized right/left distributions. That is, each entry is $2^6$ times it's original value to make it more readable. The 4 first rows in each table show $p_{xr}$ and the last 4 rows show $q_{xr}$.

| x \ r | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 0 | 2 | 2 | 1 | 2 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 2 | 1 |
| **1** | 1 | 1 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 1 | 1 | 1 |
| **2** | 1 | 2 | 0 | 0 | 1 | 0 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 0 | 0 | 1 |
| **3** | 1 | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 1 | 1 |
| **0** | 1 | 2 | 2 | 0 | 2 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 2 | 2 | 2 |
| **1** | 1 | 0 | 0 | 2 | 0 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 0 | 0 | 0 |
| **2** | 0 | 2 | 2 | 0 | 2 | 0 | 1 | 1 | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **3** | 2 | 0 | 0 | 2 | 0 | 2 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |

Table A.1: Right and left distribution for $S_1$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 1 | 1 | 2 | 1 | 0 | 2 | 1 | 0 | 1 |
| **1** | 1 | 1 | 0 | 2 | 2 | 0 | 1 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 |
| **2** | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 1 | 1 | 0 | 1 | 2 | 0 | 1 | 2 | 1 |
| **3** | 1 | 1 | 2 | 0 | 0 | 2 | 1 | 1 | 1 | 2 | 1 | 0 | 1 | 1 | 1 | 1 |
| **0** | 0 | 1 | 1 | 2 | 2 | 0 | 1 | 1 | 2 | 0 | 0 | 1 | 0 | 1 | 2 | 2 |
| **1** | 2 | 1 | 1 | 0 | 0 | 2 | 1 | 1 | 0 | 2 | 2 | 1 | 2 | 1 | 0 | 0 |
| **2** | 1 | 2 | 1 | 1 | 2 | 0 | 0 | 1 | 1 | 0 | 2 | 1 | 0 | 2 | 1 | 1 |
| **3** | 1 | 0 | 1 | 1 | 0 | 2 | 2 | 1 | 1 | 2 | 0 | 1 | 2 | 0 | 1 | 1 |

Table A.2: Right and left distribution for $S_2$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| **1** | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 0 | 1 | 0 | 0 | 1 | 1 | 2 | 1 | 1 |
| **2** | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| **3** | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 2 | 1 | 2 | 2 | 1 | 1 | 0 | 1 | 1 |
| **0** | 2 | 0 | 0 | 2 | 1 | 1 | 2 | 1 | 0 | 2 | 2 | 0 | 0 | 1 | 1 | 1 |
| **1** | 0 | 2 | 2 | 0 | 1 | 1 | 0 | 1 | 2 | 0 | 0 | 2 | 2 | 1 | 1 | 1 |
| **2** | 2 | 1 | 0 | 1 | 1 | 0 | 2 | 1 | 2 | 2 | 1 | 0 | 0 | 2 | 0 | 1 |
| **3** | 0 | 1 | 2 | 1 | 1 | 2 | 0 | 1 | 0 | 0 | 1 | 2 | 2 | 0 | 2 | 1 |

Table A.3: Right and left distribution for $S_3$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 2 | 2 | 2 | 1 | 1 | 1 | 0 | 1 | 1 |
| **1** | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 0 |
| **2** | 1 | 1 | 2 | 1 | 1 | 1 | 2 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 1 |
| **3** | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 1 | 2 |
| **0** | 2 | 0 | 0 | 2 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 |
| **1** | 0 | 2 | 2 | 0 | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 1 |
| **2** | 2 | 1 | 0 | 1 | 0 | 0 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 0 | 1 |
| **3** | 0 | 1 | 2 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 1 |

Table A.4: Right and left distribution for $S_4$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 2 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| **1** | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| **2** | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| **3** | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| **0** | 0 | 2 | 2 | 0 | 2 | 0 | 1 | 2 | 0 | 0 | 1 | 2 | 2 | 1 | 1 | 0 |
| **1** | 2 | 0 | 0 | 2 | 0 | 2 | 1 | 0 | 2 | 2 | 1 | 0 | 0 | 1 | 1 | 2 |
| **2** | 0 | 2 | 2 | 0 | 1 | 0 | 0 | 2 | 2 | 0 | 1 | 2 | 1 | 2 | 1 | 0 |
| **3** | 2 | 0 | 0 | 2 | 1 | 2 | 2 | 0 | 0 | 2 | 1 | 0 | 1 | 0 | 1 | 2 |

Table A.5: Right and left distribution for $S_5$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 1 | 1 |
| **1** | 1 | 1 | 1 | 1 | 2 | 0 | 2 | 1 | 1 | 2 | 1 | 1 | 0 | 1 | 0 | 1 |
| **2** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 1 | 0 | 2 | 1 | 1 |
| **3** | 1 | 1 | 1 | 1 | 0 | 2 | 0 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 2 | 1 |
| **0** | 0 | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 0 | 2 | 0 | 0 | 2 |
| **1** | 2 | 1 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 0 | 2 | 2 | 0 |
| **2** | 0 | 0 | 2 | 2 | 1 | 2 | 0 | 0 | 1 | 2 | 1 | 0 | 2 | 0 | 1 | 2 |
| **3** | 2 | 2 | 0 | 0 | 1 | 0 | 2 | 2 | 1 | 0 | 1 | 2 | 0 | 2 | 1 | 0 |

Table A.6: Right and left distribution for $S_6$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 0 | 0 | 1 |
| **1** | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 2 | 2 | 0 |
| **2** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 1 |
| **3** | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 0 | 0 | 2 |
| **0** | 2 | 1 | 1 | 0 | 2 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 0 | 2 | 1 | 1 |
| **1** | 0 | 1 | 1 | 2 | 0 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 2 | 0 | 1 | 1 |
| **2** | 0 | 2 | 0 | 1 | 2 | 0 | 1 | 2 | 1 | 0 | 1 | 2 | 1 | 2 | 1 | 0 |
| **3** | 2 | 0 | 2 | 1 | 0 | 2 | 1 | 0 | 1 | 2 | 1 | 0 | 1 | 0 | 1 | 2 |

Table A.7: Right and left distribution for $S_7$

| **r**<br>**x** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 0 | 0 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |
| **1** | 1 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 0 | 1 | 1 | 1 | 1 |
| **2** | 1 | 2 | 2 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **3** | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 |
| **0** | 0 | 2 | 1 | 1 | 2 | 0 | 1 | 1 | 2 | 0 | 1 | 1 | 0 | 2 | 0 | 2 |
| **1** | 2 | 0 | 1 | 1 | 0 | 2 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 2 | 0 |
| **2** | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 0 |
| **3** | 2 | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 2 |

Table A.8: Right and left distribution for $S_8$

# Appendix B

# $Q$ distributions for each S-box

The following tables show the normalized $Q$ distributions. That is, each entry is $2^6$ times it's original value to make it more readable.

| r<br>x y | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **0 1** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **0 2** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 3** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **1 0** | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1 1** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **1 2** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **1 3** | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **2 0** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **2 1** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **2 2** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **2 3** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **3 0** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 1** | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| **3 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **3 3** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |

TABLE B.1: $Q$ distribution for $S_1$

| r<br>x y | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 1** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 2** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **0 3** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **1 0** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **1 1** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **1 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **1 3** | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **2 0** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **2 1** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **2 2** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **2 3** | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 0** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **3 1** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **3 2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 3** | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |

Table B.2: $Q$ distribution for $S_2$

| r<br>x y | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| **0 1** | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **0 2** | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **0 3** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **1 0** | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **1 1** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **1 2** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **1 3** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **2 0** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **2 1** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **2 2** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| **2 3** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **3 0** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **3 1** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **3 2** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **3 3** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |

Table B.3: $Q$ distribution for $S_3$

| **r**<br>**x y** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **0 1** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **0 2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **0 3** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **1 0** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **1 1** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **1 2** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **1 3** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **2 0** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **2 1** | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **2 2** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **2 3** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 0** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 1** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **3 2** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **3 3** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |

Table B.4: $Q$ distribution for $S_4$

| **r**<br>**x y** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **0 1** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **0 2** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **0 3** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **1 0** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **1 1** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **1 2** | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| **1 3** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **2 0** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| **2 1** | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **2 2** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **2 3** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **3 0** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **3 1** | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| **3 2** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **3 3** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Table B.5: $Q$ distribution for $S_5$

| r<br>x y | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **0 1** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **0 2** | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 3** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **1 0** | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **1 1** | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **1 2** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **1 3** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **2 0** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| **2 1** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| **2 2** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **2 3** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **3 0** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **3 1** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **3 2** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **3 3** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |

TABLE B.6: $Q$ distribution for $S_6$

| r<br>x y | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 1** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **0 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **0 3** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **1 0** | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1 1** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **1 2** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **1 3** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **2 0** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **2 1** | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **2 2** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| **2 3** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| **3 0** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| **3 1** | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **3 2** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 3** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |

TABLE B.7: $Q$ distribution for $S_7$

| **x y** \ **r** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0 0** | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **0 1** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **0 2** | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **0 3** | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **1 0** | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| **1 1** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| **1 2** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **1 3** | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| **2 0** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| **2 1** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **2 2** | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **2 3** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| **3 0** | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| **3 1** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| **3 2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| **3 3** | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |

Table B.8: $Q$ distribution for $S_8$