Feistel Schemes and Bi-Linear Cryptanalysis[⋆]

(Long extended version of Crypto 2004 paper)

Nicolas T. Courtois
Axalto Smart Cards Crypto Research,
36-38 rue de la Princesse, BP 45, F-78430 Louveciennes Cedex, France,
courtois@minrank.org

Abstract. In this paper we introduce the method of bi-linear cryptanalysis (BLC), designed specifically to attack Feistel ciphers. It allows to construct periodic biased characteristics that combine for an arbitrary number of rounds. In particular, we present a practical attack on DES based on a 1-round invariant, the fastest known based on such invariant, and about as fast as the best Matsui's attack. For ciphers similar to DES, based on small S-boxes, we claim that BLC is very closely related to LC, and we do not expect to find a bi-linear attack much faster than by LC. Nevertheless we have found bi-linear characteristics that are strictly better than the best Matsui's result for 3, 7, 11 and more rounds of DES. We also study s^{5}DES [22], substantially stronger than DES against LC, yet for BLC we exhibit several unexpectedly strong biases, stronger even than existing for DES itself.

For more general Feistel schemes there is no reason whatsoever for BLC to remain only a small improvement over LC. We present a construction of a family of practical ciphers based on a big Rijndael-type S-box that are strongly resistant against linear cryptanalysis (LC) but can be easily broken by BLC, even with 16 or more rounds.

Key Words: Block ciphers, Feistel schemes, S-box design, inverse-based S-box, DES, s^{5}DES, linear cryptanalysis, generalised linear cryptanalysis, I/O sums, correlation attacks on block ciphers, multivariate quadratic equations.

1 Introduction

In spite of growing importance of AES, Feistel schemes and DES remain widely used in practice, especially in financial/banking sector. The linear cryptanalysis (LC), due to Gilbert and Matsui is the best known plaintext attack on DES, see [5, 28, 30, 19, 24]. (For chosen plaintext attacks, see [24, 2]).

A straightforward way of extending linear attacks is to consider nonlinear multivariate equations. Exact multivariate equations can give a tiny improvement to the last round of a linear attack, as shown at Crypto'98 [21]. A more powerful idea is to use probabilistic multivariate equations, for every round, and replace Matsui's biased linear I/O sums by nonlinear I/O sums as proposed by Harpes, Kramer, and Massey at Eurocrypt'95 [12]. This is known as Generalized Linear Cryptanalysis (GLC). In [13, 14] Harpes introduces partitioning cryptanalysis (PC) and shows that it generalizes both LC and GLC. The correlation cryptanalysis (CC) introduced in Jakobsen's master thesis [16] is claimed even more general. Moreover, in [15] it is shown that all these attacks, including also Differential Cryptanalysis are closely related and can be studied in terms of the

[⋆] Work supported by the French Ministry of Research RNRT Project "X-CRYPT".

Fast Fourier Transform for the cipher round function. Unfortunately, computing this transform is in general infeasible for a real-life cipher and up till now, non-linear multivariate I/O sums played a marginal role in attacking real ciphers. Accordingly, these attacks may be excessively general and there is probably no substitute to finding and studying in details interesting special cases.

At Eurocrypt'96 Knudsen and Robshaw consider applying GLC to Feistel schemes [23], and affirm that (cf. page 226 in [23]) in this case non-linear characteristics cannot be joined together. In this paper we will demonstrate that GLC can indeed be applied to Feistel ciphers. This is made possible with our new "Bi-Linear Cryptanalysis" (BLC) attack.

2 Feistel Schemes and Bi-Linear Functions

Feistel schemes are a construction that allows to build a pseudo-random permutation from a pseudo-random function, see [27]. In theory, neither their periodic connection scheme nor the fact that the round functions are usually more or less identical will be a weakness in itself. Indeed, when this round function is [pseudo-]random, it is still possible to prove very strong security results on such schemes, see for example [27] and [32]. However unfortunately, the round functions used in all practical Feistel ciphers are not pseudo-random. In this case, the regularity properties do help the attacker: if an interesting periodic property is found for a few rounds of the cipher, it will be extended to a general attack for an arbitrary number of rounds. Hence we may hope to find a good attack without having to explore all possible combinations for an arbitrary number of rounds. Thus differential [2] and linear attacks on DES [28, 1] have periodic patterns with invariant equations for some 1, 3 or 8 rounds. In this paper we will present several new practical attacks with periodic structure for DES, including new 1-round invariants.

2.1 The Principle of the Bi-Linear Attack on Feistel Schemes

In one round of a Feistel scheme, one half is unchanged, and one half is linearly combined with the output of the component connected to the other half. This will allow bi-linear I/O expressions on the round function to be combined together. First we will give an example with one product, and extend it to arbitrary bi-linear expressions. Then in Section 3 we explain the full method in details (with linear parts present too) for an arbitrary Feistel schemes. Later we will apply it to get concrete working attacks for DES and other ciphers.

In this paper we represent Feistel schemes in a completely "untwisted" way, allowing to see more clearly the part that is not changed in one round. As a consequence, the orientation changes compared to most of the papers and we obtain an apparent (but extremely useful) distinction between odd and even rounds of a Feistel scheme. Otherwise, our notations are very similar to these used for DES in [26, 21]. For example $L_0[\alpha]$ denotes a sum (XOR) of some subset α of bits of the left half of the plaintext. Combinations of inputs (or outputs) of round function number $r = 1, 2, \ldots$ are denoted by $I_r[\alpha]$ (or $O_r[\beta]$). Our exact notations for DES will be explained in more details when needed, in Section 6.1. For the time being, we start with a simple rather self-explaining example (cf. Figure 1 below) that works for any Feistel cipher.

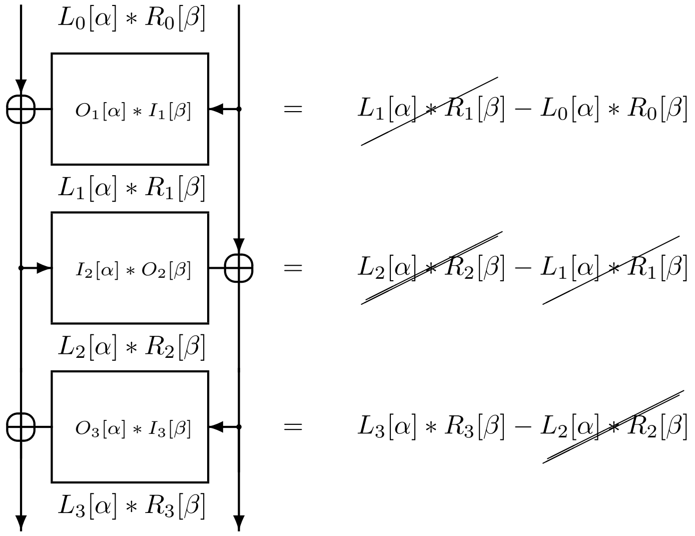
$L_0[\alpha] * R_0[\beta]$
$O_1[\alpha] * I_1[\beta]$
$=$
$\cancel{L_1[\alpha] * R_1[\beta]} - L_0[\alpha] * R_0[\beta]$
$L_1[\alpha] * R_1[\beta]$
$I_2[\alpha] * O_2[\beta]$
$=$
$\cancel{L_2[\alpha] * R_2[\beta]} - \cancel{L_1[\alpha] * R_1[\beta]}$
$L_2[\alpha] * R_2[\beta]$
$O_3[\alpha] * I_3[\beta]$
$=$
$L_3[\alpha] * R_3[\beta] - \cancel{L_2[\alpha] * R_2[\beta]}$
$L_3[\alpha] * R_3[\beta]$

Fig. 1. Fundamental remark: combining bi-linear expressions in a Feistel cipher

Proposition 2.1.1 (Combining bi-linear expressions in a Feistel cipher). For all (even unbalanced) Feistel ciphers operating on $n + n'$ bits with arbitrary round functions we have: $\forall \alpha \subset \{1, \ldots, n\}, \forall \beta \subset \{1, \ldots, n'\},\ \forall r \geq 0$:

$$L_r[\alpha]R_r[\beta] \oplus L_0[\alpha]R_0[\beta] = \sum_{i=1}^{\lceil r/2 \rceil} O_{2i-1}[\alpha]I_{2i-1}[\beta] \ \oplus \sum_{i=1}^{\lfloor r/2 \rfloor} I_{2i}[\alpha]O_{2i}[\beta] \qquad \square$$

From one product this fundamental result extends immediately, by linearity, to arbitrary bi-linear expressions. Moreover, we will see that these bi-linear expressions do not necessarily have to be the same in every round, and that they can be freely combined with linear expressions (BLC contains LC).

3 Bi-linear Characteristics

For simplicity let $n = n'$. In this section we construct a completely general bi-linear characteristic for one round of a Feistel cipher. Then we show how it combines for the next round. Here we study bits locally and denote them by A_i, B_j etc. Later for constructing attacks for many rounds of practical Feistel ciphers we will use (again) the notations $L_i[j_1, \ldots, j_k]$ (cf. Section 6.1).

3.1 Constructing a Bi-linear Characteristic for One Round

Let $\mathcal{S}$ be a homogeneous bi-linear Boolean function $GF(2^n) \times GF(2^n) \to GF(2)$. Let $\mathcal{S}(A_1, \ldots, A_n; B_1, \ldots, B_n) = \sum s_{ij} A_i B_j$.

Let f_K be the round function of a Feistel cipher. We assume that there exist two linear combinations u and v such that the function:

$$(B_1, \ldots, B_n) \mapsto \begin{cases} \sum s_{ij} O_i B_j \oplus \sum u_i O_i \oplus \sum v_i B_i \\ \text{with } \ (O_1, \ldots, O_n) = f_K(B_1, \ldots, B_n) \end{cases}$$

is biased and equal to 0 with some probability $p \neq 1/2$ with $p = p(K)$ depending in some way on the round key K.

We have $C_i = A_i \oplus O_i$. By bi-linearity (or from Proposition 2.1.1) the following holds:

$$\sum s_{ij} A_i B_j \oplus \sum s_{ij} O_i B_j = \sum s_{ij} C_i B_j$$

From this, for the first round, (could be also any odd-numbered round), we obtain the following characteristic:

$$\left.\begin{array}{l} \sum s_{ij} A_i B_j \oplus \sum u_i A_i \oplus \sum v_i B_i = \\ \sum s_{ij} C_i B_j \oplus \sum u_i C_i \end{array}\right\} \quad \text{with probability } \; p(K)$$

Finally, we note that, the part linear in the B_i can be arbitrarily split in two parts: $\sum v_i B_i = \sum v_i^{(1)} B_i \oplus \sum v_i^{(2)} B_i$ with $v_i = v_i^{(1)} \oplus v_i^{(2)}$ for all $i = 1, \ldots, n$.

All this is summarized on the following picture:

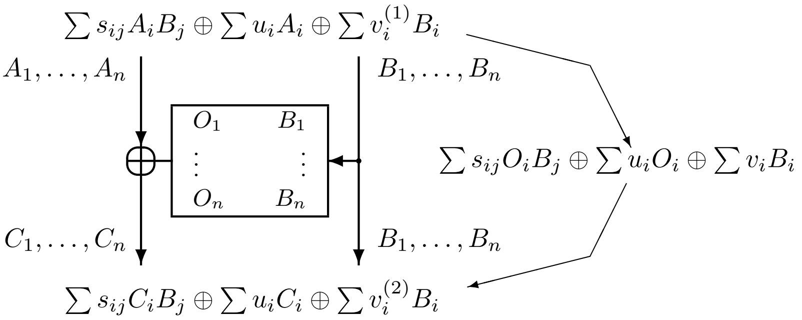
$\sum s_{ij} A_i B_j \oplus \sum u_i A_i \oplus \sum v_i^{(1)} B_i$
$A_1, \ldots, A_n$
$B_1, \ldots, B_n$
O_1
B_1
O_n
B_n
$\sum s_{ij} O_i B_j \oplus \sum u_i O_i \oplus \sum v_i B_i$
$C_1, \ldots, C_n$
$B_1, \ldots, B_n$
$\sum s_{ij} C_i B_j \oplus \sum u_i C_i \oplus \sum v_i^{(2)} B_i$

Fig. 2. Constructing a bi-linear characteristic for an odd round of a Feistel cipher

3.2 Application to the Next (Even) Round

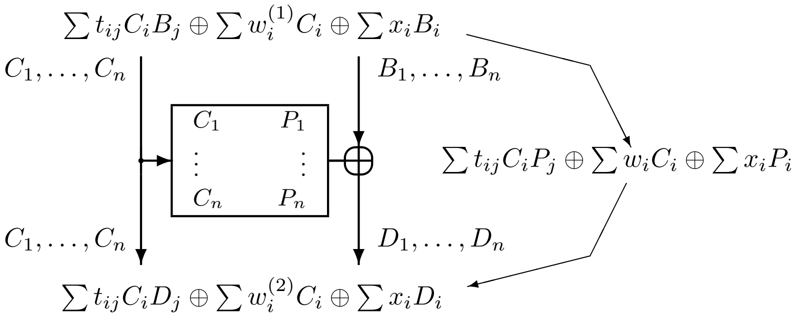
$\sum t_{ij} C_i B_j \oplus \sum w_i^{(1)} C_i \oplus \sum x_i B_i$
$C_1, \ldots, C_n$
$B_1, \ldots, B_n$
C_1
P_1
C_n
P_n
$\sum t_{ij} C_i P_j \oplus \sum w_i C_i \oplus \sum x_i P_i$
$C_1, \ldots, C_n$
$D_1, \ldots, D_n$
$\sum t_{ij} C_i D_j \oplus \sum w_i^{(2)} C_i \oplus \sum x_i D_i$

Fig. 3. Constructing a bi-linear characteristic for an even round of a Feistel cipher

The same method can be applied to the next, even, round of a Feistel scheme, with the only difference that the round function is connected in the inverse direction. In this case, to obtain a characteristic true with probability $\neq 1/2$, we need to have a bias in the function:

$$(C_1, \ldots, C_n) \mapsto \left\{ \begin{array}{c} \sum t_{ij} C_i P_j \oplus \sum w_i C_i \oplus \sum x_i P_i \\ \text{with } \; (P_1, \ldots, P_n) = f_K(C_1, \ldots, C_n) \end{array} \right.$$

3.3 Combining Approximations to Get a Bi-Linear Attack for an Arbitrary Number of Rounds

It is obvious that such I/O sums as specified above can be combined for an arbitrary number of rounds (contradicting [23] page 226). To combine the two characteristics specified above, we require the following three conditions:

1. We need $u = w^{(1)}$.
2. We need $v^{(2)} = x$.
3. We need the homogenous quadratic parts s et t to be correlated (seen as Boolean functions). They do **not** have to be the same (though in many cases they will). In linear cryptanalysis (LC), a correlation between two linear combinations means that these linear combinations have to be the same. In generalized linear cryptanalysis (GLC) [12], and in particular here, for bi-linear I/O sums, it is no longer true. Correlations between quadratic Boolean functions are frequent, and does not imply that $s = t$. For these reasons the number of possible bi-linear attacks is potentially very large.

Summary: We observe that bi-linear characteristics combine exactly as in LC for their linear parts, and that their quadratic parts should be either identical (with orientation that changes in every other round), or correlated.

4 Predicting the Behaviour of Bi-linear Attacks

The behaviour of LC is simple and the heuristic methods of Matsui [28] are known to be able to predict the behaviour of the attacks with good precision (see below). Some attacks work even better than predicted. As already suggested in [12, 23] the study of generalised linear cryptanalysis is **much harder**.

4.1 Computing the Bias of Combined Approximations

A bi-linear attack will use an I/O sum for the whole cipher, being a sum of I/O sums for each round of the cipher such that the terms in the internal variables do cancel. To compute the probability the resulting equation is true, is in general not obvious. Assuming that the I/O sum uses balanced Boolean functions, (otherwise it will be even harder to analyse) one can apply the Matsui's Piling-up Lemma from [28]. This however **can fail**. It is known from [12] that a sum of two very strongly biased characteristics can have a bias much weaker than expected. The resulting bias can even be exactly zero: an explicit example can be found in Section 6.1. of [12]. Such a problem can arise when the connecting characteristics are not independent. This will happen more frequently in BLC than in LC: two linear Boolean functions are perfectly independent unless equal, for non-linear Boolean functions, correlations are frequent. Accordingly, we do not sum independent random variables and the Matsui's lemma may fail.

At this stage there are two approaches: one can try to define a class of attacks that can be proved to work, and restrict oneself only to studying such attacks, or try to explore all possible attacks, including those that do work experimentally without proof. This first approach is adopted in [12]: the Lemma 6 gives a sufficient condition to guarantee that the Piling-up lemma will apply. For this the probability, that the characteristic is true, for a random partial key, should be independent of the input (e.g. the input of the whole round). This explains why Matsui's attacks indeed work well. In [12] it allows to prove that the proposed family of GLC attacks based on homomorphic properties will work as predicted. We will also use this argument in Section 5.

In this paper we frequently adopt rather the second approach: try find as many working attacks as possible, even if current theory does not allow to predict their behaviour with accuracy. We claim that studying only attacks that satisfy the assumption of Lemma 6 in [12] is fairly limitative: for an attack to work there is obviously no necessity for the equations to be independent from their inputs (it would however be nice if the sign of the bias remained the same with good probability). Without this relaxation we would never find bi-linear characteristics that are better than Matsui (Section 6.5): they are not independent on the input. Thus we can go beyond homomorphic attacks related to the given group operation that is used to combine the key in the cipher. We gain in freedom and can find better attacks than previously known. A price to pay for this is that each application of Matsui's Lemma, and similar heuristic deductions will be systematically questioned and confronted to experimental results.

4.2 Key Dependence in Bi-Linear Attacks

Another important property of bi-linear cryptanalysis is that the existence of a bias for one characteristic does frequently depend on the key. This does not really happen for LC applied DES, because in DES all key bits are combined linearly and a linear equation will be true with probability either p or $1 - p$ depending on the key. However it will happen for LC and other ciphers, if key bits are involved in a more complex way, for example for ICE [25].

In bi-linear cryptanalysis, the behaviour becomes complex already when the key bits are combined linearly as in DES. Adding a constant (a key bit) to an input of an S-box, does not only modify the constant part in a bi-linear characteristic, but also the linear part. (We note that for DES only the linear part in the output variables will be modified when the key changes). From this, quite frequently two bi-linear characteristics for two parts of a cipher (e.g. for S-boxes) will only connect together for some keys. Such attacks are still very interesting and frequently also do work, with only a slightly weaker bias, for all the other keys. For simplicity, no key bits are displayed in bi-linear characteristics for one or several rounds of a cipher that are studied/displayed in this paper. The values of biases we will present (unless otherwise stated) are given for the reference key being zero. Yet typically we observed that they exist, and slightly vary in value, also for **any** other key (chosen at random). In rare cases, the bias works well only for a fraction of keys (e.g. 25 %): this happens in Appendix B.1.

4.3 Exploring Bi-linear Cryptanalysis

There are different approaches to finding interesting bi-linear attacks to block ciphers. In few cases one can construct attacks that will provably or arguably work (see [12] and later Section 5). Another method is to construct characteristics "by hand" around some particularly strong bias found for one S-box. This will allow us to find attacks on DES better than Matsui in Section 6.5 and in the Appendix. Unfortunately we have to check if it really works at every stage.

We also propose a third method of "periodic constructions": we build a large class of plausible attacks exploring all the existing biases in the S-boxes that can be connected in some specific way to another S-box in the next round, and finally to itself after a few rounds. Then we explore this class by computer simulations to find the best experimental characteristic for the whole/reduced cipher. This method is used in Appendix E works well for up to about 5 rounds. For more rounds it produces mostly attacks that does not work well, see Appendix E.

We noted the two major difficulties: predicting the bias of combined characteristics, and huge number of possible characteristics (including fragmentation due to the fact they the bias does in general depend on the key). These make it very difficult to have a systematic method (a computer program) that would compute the best bi-linear characteristic for a given cipher. To check if an attack indeed works requires to be able to generate as many plaintexts as for the real attack. To find the best attack is even much harder. It requires to exhaustively search and reject lots of other combinations that should work well but they don't. Each of them has to be tested on an equally large set of plaintexts.

5 The Killer Example for Bi-Linear Cryptanalysis

We will construct a practical cipher that is very secure w.r.t. all known attacks for block ciphers, in particular for LC, yet broken by BLC. It mixes two group operations: the XOR and the multiplication in $GF(2^n)$ e.g. $n = 32$ or 64. It uses the inverse in $GF(2^n)$ (cf. Rijndael): let $Inv(X) = X^{-1}$ in $GF(2^n)$ when $X \neq 0$ and 0 otherwise. We build a $2n$-bit Feistel cipher with the i-th round function being:

$$f_i(X) = Inv(X) \cdot (K_i \oplus G(X)) \quad \text{in } GF(2^n), \tag{1}$$

with K_i being the partial key, and G being some function with S-boxes and arbitrary components $\{0,1\}^n \to \{0,1\}^n$. In order to get an insecure cipher, we need to assume that some linear combination of outputs of G is biased. For example, let $Y_1 \oplus Y_5 = 0$ with probability 3/4. Building a cipher with G alone would be insecure for LC, however here G is composed by a group operation $\cdot$ with $Inv(X)$. The $Inv(X)$ assures global diffusion and very high non-linearity (cf. [4]). Accordingly our round function has very good resistance to linear and differential cryptanalysis for most G, even when $G = 0$. But not against BLC.

First, we can consider a bi-linear attack with bi-linear equations over $GF(2^n)$: $\forall r \geq 0$:

$$L_r \cdot R_r \oplus L_0 \cdot R_0 = \sum_{i=1}^{\lceil r/2 \rceil} O_{2i-1} \cdot I_{2i-1} \;\oplus \sum_{i=1}^{\lfloor r/2 \rfloor} I_{2i} \cdot O_{2i} = \sum_{i=1}^{r} I_i \cdot O_i \tag{2}$$

We can, at any time, transform these to multivariate bi-linear equations with n variables over $GF(2)$. Any bi-linear function over $GF(2^n)$, e.g. the multiplication, can be re-written as n bi-linear multivariate functions $GF(2)^n \to GF(2)$. Let $X \cdot Y = (Z_1, \ldots, Z_n)$ with $Z_k = \sum_{ij} M_k^{ij} X_i Y_j$. From (2), or if we prefer, directly from Proposition 2.1.1 and by using (very useful) the symmetry $M_k^{ij} = M_k^{ji}$, we get:

$$\forall k \in \{1, \ldots, n\}, \forall r \geq 0 \qquad \sum_{ij} M_k^{ij} \left(L_{ri} R_{rj} \oplus L_{0i} R_{0j}\right) = \sum_{l=1}^{r} \sum_{ij} M_k^{ij} I_{li} O_{lj} \tag{3}$$

Now, $\forall l \geq 1$, $I_l \cdot O_l = K_l \oplus G(I_l)$ with probability $(1 - 1/2^n)$. We rewrite it:

$$\forall k \in \{1, \ldots, n\}, \forall l \geq 0 \qquad \sum_{ij} M_k^{ij} I_{li} O_{lj} = K_{ik} \oplus G_k(I_i) \tag{4}$$

Then we use the linear output bias of G: $G_1 \oplus G_5 = 0$ with probability 3/4.

$$\forall l \geq 0 \quad \sum_{ij} M_1^{ij} I_{li} O_{lj} \oplus \sum_{ij} M_5^{ij} I_{li} O_{lj} = K_{i1} \oplus G_1(I_i) \oplus K_{i5} \oplus G_5(I_i) \approx C_l \tag{5}$$

The last expression is equal to come constant denoted C_l with probability 3/4. Finally, we combine with (3) (or equivalently sum these bi-linear expressions over the whole cipher with r rounds).

$$\sum_{ij} \left(M_1^{ij} \oplus M_5^{ij} \right) \left(L_{ri} R_{rj} \oplus L_{0i} R_{0j} \right) = \sum_{l=1}^{r} C_l \text{ with probability } \frac{1}{2} + \frac{1}{2^{r+1}} \tag{6}$$

What we obtained is a biased bi-linear I/O sum for the whole cipher. We can distinguish this cipher from a random permutation given about 2^{2r+2} plaintexts. For example 16 rounds will be broken on a laptop PC.

Does it work as predicted ? In general, as we explain in Section 4.1, it is hard to predict accurately the behaviour of a composed bi-linear attack. However we have little doubt it will work: the $Inv(X)$ should render possible correlation between approximations being combined negligible. By Lemma 6 in [12], if the characteristics are independent of the input (the input of the whole round), the Piling-up lemma does apply. In real life the dependencies do exist, but due to very good properties of Inv function they must be very weak: In equation (5) the complex bi-linear part that comes from $Inv()$ will assure that. This argument is valid for most functions G. In some cases we can even prove that this attack works: when $G = 0$, and also when one fixed linear combination of output bits of G is 0, (the other parts can be arbitrary functions). In these cases, dependencies cannot be a problem: we add equations (5) true with probability 1 to get the equation (6) true with probability 1.

Related work: Similar results were previously obtained for some substitution-permutation network (SPN) ciphers. In [12] Harpes, Kramer and Massey give an example of 8-bit SPN that is secure against LC and DC, but insecure for generalised linear cryptanalysis due to a probabilistic homomorphic property of each round relative to quadratic residuosity function modulo $2^8 + 1$. The Jakobsen attack for substitution ciphers that uses probabilistic univariate polynomials from [18] can also be seen as a special case of GLC. However, it is the first time that GLC allows to break a Feistel cipher, which contradicts the impossibility professed by Knudsen and Robshaw [23]. This cipher is built with state-of-art components (inverse in $GF(2^n)$) and can in addition incorporate any additional fashionable component with lots of theory and designer tricks, as a part of G. Due to G it will not have homomorphic properties. Moreover, by adjusting the bias in G, the security of this cipher against BLC will be freely adjusted between (nearly) zero and infinity. It can therefore be arbitrarily weak for BLC, and this even for a very large number of rounds. Yet, the security against the usual attacks (LC, DC) should remain equally good (due to the big Inv S-box).

6 Bi-Linear Attacks on DES

6.1 Notation

We ignore the initial and final permutations of DES that have no incidence on the attacks. We use the "untwisted method" of representing DES, as on the right-hand figure, page 254 in [31]. The bit numbering is compatible with the FIPS standard [11], and [26, 21], and differs from the notations of Biham, Shamir [2] or Matsui [28, 30, 20].

We denote the bits of the left hand side of the plaintext by $L_0[1] \ldots L_0[n]$. The bits of the right hand side are $R_0[1] \ldots R_0[n]$. Similarly, as in other papers, the plaintext after i rounds will be L_i, R_i, except that we felt it necessary to have our notations completely "untwisted" which implies that our L_i and R_i for an odd $i = 1, 3, \ldots$ will be inversed compared to [26, 21, 31], Then, we apply the popular convention $X[i_1, \ldots, i_n]$being $X[i_1] \oplus \ldots \oplus X[i_n]$. For example $L_0[9, 7, 23, 31]$ is the XOR of 4 bits of the left half of the plaintext that are added to the outputs of S1 in the first round. We denote the input bits to the $i-$th round function by $I_i[1], \ldots, I_i[32]$. Similarly the output bits will be $O_i[1], \ldots, O_i[1]$.
For odd i we have $I_i[j] = R_{i-1}[j] = R_i[j]$ and $O_i[j] = L_{i-1}[j] \oplus L_i[j]$.
For even i we have $I_i[j] = L_{i-1}[j] = L_i[j]$ and $O_i[j] = R_{i-1}[j] \oplus R_i[j]$.

For individual S-boxes, we will denote the inputs/outputs by respectively $O[i]$ and $J[j]$ with i, j being directly the numbers 1..32 in the round function of DES. For example $O[8], O[14], O[25], O[3]$ are the outputs of S-box S5, and $J[16], \ldots, J[21]$ are the inputs of this S-box S5.

Finally, we call $K_j^{(Sx,i)}$ the key bit that is XORed at the entry of S-box Sx, in round $i = 1, 2, 3, \ldots$, at the bit numbered j, with $j = 1..32$ coded as above (i.e. corresponding to the position in the round function input). Depending on the key in round i, we have $I_i[k] = J_i[k]$ or $I_i[k] = J_i[k] + 1$. For better readability, in most cases we avoid naming precisely the key bits involved.

6.2 First Example of Bi-Linear Cryptanalysis of DES

Our simulations on DES S-boxes (cf. Appendix A) show that the following two bi-linear characteristics exist for DES S-boxes S1 and S5:

$$O[8, 14, 25, 3] \oplus J[17] \cdot O[3] = 0 \quad \text{for } S5 \text{ with probability } 17/64$$

$$O[17] \oplus J[3] \cdot O[17] = 0 \quad \text{for } S1 \text{ with probability } 47/64$$

From these, acting as if all the key bits were zero ($I_i[k] = J_i[k]$), we deduce the following bi-linear characteristic for two rounds:

$$(*) \begin{array}{l} L_0[3, 8, 14, 25] \oplus L_0[3]R_0[17] \oplus R_0[17] \oplus \\ L_2[3, 8, 14, 25] \oplus L_2[3]R_2[17] \oplus R_2[17] = K[sth] \end{array} \Bigg\} \quad \frac{1}{2} - 1.76 \cdot 2^{-4}$$

The explanation is given on the following picture (Figure 4):

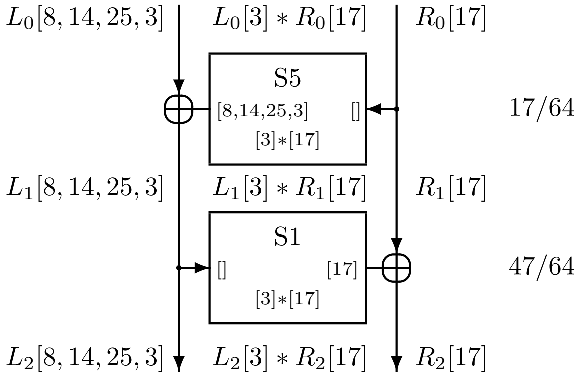

Fig. 4. Our first example - an invariant bi-linear attack on DES ($*$)

We verified this bias experimentally, and the probability is (we were lucky) equal to the probability that is predicted by Matsui's Piling-Up Lemma.

Key Dependence: Very surprisingly, the above equation ($*$) is biased, not only when all key bits are 0, but for every DES key. This can be seen to come from a couple of other (different) bi-linear characteristics from Appendix A.

More rounds: It is easy to see from the picture, and we verified it experimentally, that ($*$) is also biased for $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots$ rounds of DES, and all this happens to work about equally well for an arbitrary key.

Relation to LC: The bias of ($*$) is closely related to some prominent equations of Matsui, their difference is "naturally" biased, see Appendix F.

6.3 Invariant Attacks on DES

The equation ($*$) is an invariant equation, i.e. the input and the output bi-linear expressions are the same. We have found a simple invariant bi-linear I/O sum for DES that is biased for any key and for any number of rounds. For LC and DES, such simple invariant characteristics do exist, have been found by Biham (page 347 in [1]) in close relation to Davies-Murphy attack. The example ($*$) above is one of the best we found for DES, and so far it also **the only known** non-linear 1-round invariant attack on DES that works really well in practice. Our invariant on DES is stronger than Biham's. We recall that Biham uses a bias on a sum of some outputs for two successive DES S-boxes. The best bias obtained by Biham (also exhibited by Matsui in [29] and contained unnoticed in the earlier Davies-Murphy attack [9, 10]) is equal to $(35/64 - 1/2)$ for 2 rounds and for S-boxes S7-S8. This gives $1.4 \cdot 2^{-22}$ for 12 rounds. Instead, ($*$) gives experimentally only about $1.3 \cdot 2^{-18}$. Accordingly, ($*$) is **the strongest known 1-round invariant attack on DES**.

To break full DES requires a bias for 14 rounds (Matsui's 2R method) and the Biham's invariant requires then 2^{50} plaintexts. Our invariant attack requires about 2^{43} plaintexts (the bias of ($*$) for 14 rounds is expected to be about 2^{-22}, we did not dispose of a sufficient computing power to compute it exactly).

6.4 How Good is Our First Example, BLC vs. LC

These new properties of DES give a chosen-plaintext attack on an arbitrary number of rounds of DES, somewhat simpler than Matsui's laborious search for the best linear characteristic. If we try here to predict the resulting bias for 14 rounds by applying the Matsui's Piling-up formula, we would get for 14 rounds the bias of: $1.63 \cdot 2^{-17}$ which means an attack on full DES with only $2^{32.6}$ known plaintexts (!?). Unfortunately, unlike for LC in DES, such predictions are frequently not valid for BLC. Starting from 3 rounds, the bias of our invariant does not follow the prediction at all, yet remains significative (cf also Table 6). For example if we apply Matsui's Piling-Up Lemma to predict the bias for 4 rounds as 2+2 rounds, we obtain $1.55 \cdot 2^{-6}$, while in practice it is about $1.80 \cdot 2^{-8}$ (cf. Table 6). Our invariant attack seems very bad for 4 rounds, and unfortunately with $(*)$ we never get a bias better than obtained by Matsui. Yet, it is the best invariant attack on DES known, and for more than 4 rounds the results are again not so bad. Only slightly worse than Matsui. For example for 12 rounds the best result of Matsui from [28] gives $1.19 \cdot 2^{-17}$, while for $(*)$ and a random key our simulation gives $1.3 \cdot 2^{-18}$, To break full DES Matsui requires about 2^{43} plaintexts, and with $(*)$ we also need about 2^{43} (and both are related). In general, in Appendix F we give a heuristic argumentation why for DES (but not in general !) the complexity of the best bi-linear attack should be roughly the same than for LC.

For DES and 1-round invariants attacks extended to an arbitrary number of rounds, BLC gives strictly better results than LC. It is also so for more complex periodic constructions and we are going to see that BLC attacks can also be strictly better than any existing linear attack.

6.5 Second Example of Bi-Linear Cryptanalysis of DES

In order to exhibit biases really better than Matsui we looked what is the best bi-linear characteristic that exists in DES:

$$J[16,20] \oplus O[8,14,25,3] \oplus J[16,17,20] \cdot O[3] = 0 \quad \text{for } S5 \text{ with probability } \; 61/64.$$

We note that this equation can be seen as "causing" the existence of the Matsui's best equation (A) for S5: their difference is highly biased. Based mainly on this, we constructed a periodic characteristic for 3,7, 11 and more rounds that is strictly better than the best results of Matsui for the same number of rounds.

Proposition 6.5.1 (Our Best Attack on 11 Rounds of DES). For all keys, the following equation is biased for 11 rounds of DES:

$$(**) \quad \left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[17] \oplus \\ L_{11}[3,8,14,25] \oplus L_{11}[3]R_{11}[16,17,20] \oplus R_{11}[17] = \\ K[sth] + K[sth']L_0[3] + K[sth'']L_{11}[3] \end{array}\right\} \; \frac{1}{2} \pm \text{around } 1.2 \cdot 2^{-15}$$

The exact construction to achieve this is a bit complicated. (cf. Appendix B). The bias of this equation is strictly better than the best linear characteristic for 11 rounds obtained by Matsui (which gives $1.91 \cdot 2^{-16}$ for 11 rounds). It has been verified by computer simulations at every stage. We note also that both are closely related: their difference, is a biased Boolean function.

Our second example allows us to give an attack strictly better than Matsui for 11+2=13 rounds of DES. For the full 16-round DES our results are roughly as good as Matsui (but we hope to improve this soon too). For 17 rounds of DES, as the construction of our second example ($**$) is periodic, we expect that for 11+4=15 rounds it should also be better than the best bias of Matsui, which would allow to break 15+2=17 rounds of DES faster than by LC. We do not dispose of a sufficient computing power to fully confirm this fact.

6.6 Bi-Linear Cryptanalysis of s^{5}DES

S^{5}DES is a version of DES with DES-boxes modified to be substantially more secure than DES for all known attacks on DES [22] and in particular LC. It is not so for bi-linear attacks. The S-boxes of s^{5}DES have some bi-linear characteristics true with probability 1 (best result for DES is 61/64), and for more rounds we have already found a few examples (e.g. with 3 rounds), with biases for s^{5}DES being stronger than for DES itself. However so far we did not found very good attacks for more than 4 rounds and it is likely that full s^{5}DES remains secure against BLC. Our results on s^{5}DES are given in Appendix D.

7 Conclusion

It was stated that for Feistel ciphers non-linear characteristics cannot be joined together for several rounds, see [23]. In this paper we show that generalised linear cryptanalysis (GLC) is in fact possible for Feistel schemes. To achieve this goal, we introduced bi-linear cryptanalysis (BLC). It gives a new (and the fastest known) 1-round invariant attack on DES. Though more powerful, generalized linear cryptanalysis is unfortunately much harder to study than LC. At present heuristic constructions, to be confirmed (or not) by computer simulations are the only method known to explore it. BLC is related to LC in multiple important ways. It contains LC as a sub-set. LC can be used to construct good bi-linear characteristics and vice-versa. BLC also contains LC as an extension: a combination of biased bi-linear characteristics may extend a concrete combination of biased linear characteristics by adding quadratic polynomials. Yet BLC can be strictly better than any (existing) linear attack. This was demonstrated for 3, 7, 11 and more rounds of DES, and also for s^{5}DES.

In this paper we only initiate the study of bi-linear cryptanalysis. BLC and GLC extend the role of LC as an essential tool to evaluate the real-life security of many practical ciphers. An interesting contribution of this paper is to point out that, though GLC is excessively general to be systematically explored, the properties of the top-level structure of a cryptographic scheme (e.g. being a Feistel scheme) will determine the type of the attacks (e.g. BLC) that may indeed work. Our new attack can be quite devastating: we constructed a large family of practical ciphers based on big Rijndael-type S-box, that are strongly resistant against LC and all previously known attacks on Feistel ciphers, yet can be broken in practice with BLC for an important number of rounds. Fortunately, for DES, BLC gave only slight improvements over LC and does not cause excessive trouble.

References

1. Eli Biham: *On Matsui's Linear Cryptanalysis,* Eurocrypt'94, LNCS 950, Springer-Verlag pp. 341-355, 1994.
2. Eli Biham and Adi Shamir, *Differential Cryptanalysis of DES-like Cryptosystems,* Journal of Cryptology, vol. 4, pp. 3-72, IACR, 1991.
3. Alex Biryukov, Christophe de Cannière and Michael Quisquater: *On Multiple Linear Approximations,* In Crypto 2004, LNCS 3152, Springer. Available also at `eprint.iacr.org/2004/057/`.
4. Anne Canteaut, Marion Videau: *Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis,* Eurocrypt 2002, LNCS 2332, Springer, 2002.
5. Anne Tardy-Corfdir, Henri Gilbert: *A Known Plaintext Attack of FEAL-4 and FEAL-6,* Crypto'91, LNCS 576, Springer, pp. 172-181, 1992.
6. Nicolas Courtois, Guilhem Castagnos and Louis Goubin: *What do DES S-boxes Say to Each Other ?* Available on `eprint.iacr.org/2003/184/`.
7. Nicolas Courtois: *The Inverse S-box, Non-linear Polynomial Relations and Cryptanalysis of Block Ciphers,* in AES 4 Conference, Bonn May 10-12 2004, LNCS 3373, pp. 170-188, Springer.
8. Nicolas Courtois: *General Principles of Algebraic Attacks and New Design Criteria for Components of Symmetric Ciphers,* in AES 4 Conference, Bonn May 10-12 2004, LNCS 3373, pp. 67-83, Springer.
9. D.W. Davies, *Some Regular Properties of the Data Encryption Standard,* Crypto'82, pp. 89-96, Plenum Press, New-York, 1982.
10. D. Davies and S. Murphy, *Pairs and Triplets of DES S-Boxes,* Journal of Cryptology, vol. 8, Nb. 1, pp. 1-25, 1995.
11. *Data Encryption Standard (DES),* Federal Information Processing Standards Publication (FIPS PUB) 46-3, National Bureau of Standards, Gaithersburg, MD (1999). `http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf`
12. C. Harpes, G. Kramer, and J. Massey: *A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-up Lemma,* Eurocrypt'95, LNCS 921, Springer, pp. 24-38. `http://www.isi.ee.ethz.ch/~harpes/GLClong.ps`
13. Carlo Harpes: *Cryptanalysis of iterated block ciphers,* PhD thesis, No 11625, Swiss Federal Int. of Tech., ETH Series in Information Processing, Ed. J. L. Massey, Hartung-Gorre Verlag Konstanz, 1996, ISBN 3-89649-079-6, ISSN 0942-3044.
14. Carlo Harpes: *Partitioning Cryptanalysis,* Post-Diploma Thesis, Signal and Information Processing Lab., Swiss Federal Institute of Technology, Zurich, March 1995. `http://www.isi.ee.ethz.ch/~harpes/pc.ps`
15. Thomas Jakobsen, Carlo Harpes: *Non-Uniformity Measures for Generalized Linear Cryptanalysis and Partitioning Cryptanalysis,* Pragocrypt'96, 1996.
16. Thomas Jakobsen: *Correlation Attacks on Block Ciphers,* Master's Thesis, Dept. of Mathematics, Technical University of Denmark, January 1996.
17. Thomas Jakobsen: *Higher-Order Cryptanalysis of Block Ciphers.* Ph.D. thesis, Dept. of Math., Technical University of Denmark, 1999.
18. Thomas Jakobsen: *Cryptanalysis of Block Ciphers with Probabilistic Non-Linear Relations of Low Degree,* Crypto 98, LNCS 1462, Springer, pp. 212-222, 1998.
19. Pascal Junod: *On the complexity of Matsui's attack,* Selected Areas in Cryptography (SAC'01), Toronto, Canada, LNCS 2259, pp. 199-211, Springer, 2001.
20. Burton S. Kaliski Jr, and M.J.B. Robshaw. *Linear Cryptanalysis Using Multiple Approximations,* Crypto'94, LNCS, Springer, pp. 26-39, 1994.

21. Toshinobu Kaneko and Takeshi Shimoyama: *Quadratic Relation of S-box and Its Application to the Linear Attack of Full Round DES,* In Crypto 98, LNCS 1462, p. 200-211, SPringer, 1998.
22. Kwangjo Kim. Sangjin Lee, Sangjoon Park, Daiki Lee: *Securing DES S-boxes against Three Robust Cryptanalysis,* SAC'95, pp.145-157, 1995.
23. Lars R. Knudsen, Matthew J. B. Robshaw: *Non-Linear Characteristics in Linear Cryptoanalysis,* Eurocrypt'96, LNCS 1070, Springer, pp. 224-236, 1996.
24. Lars R. Knudsen, John Erik Mathiassen: *A Chosen-Plaintext Linear Attack on DES.* FSE'2000, LNCS 1978, Springer, pp. 262-272, 2001.
25. Matthew Kwan: *The Design of the ICE Encryption Algorithm,* FSE'97, 4th International Workshop, Haifa, Israel, Springer, LNCS 1267, pp. 69-82, 1997. Available from `http://www.darkside.com.au/ice/ice.ps.gz`.
26. Susan K. Langford, Martin E. Hellman: *Differential-linear cryptanalysis,* Crypto 94, LNCS 839, pp. 17-25, Springer, 1994.
27. Michael Luby, Charles W. Rackoff, *How to construct pseudorandom permutations from pseudorandom functions*, SIAM Journal on Computing, vol. 17, n. 2, pp. 373-386, April 1988.
28. M. Matsui: *Linear Cryptanalysis Method for DES Cipher,* Eurocrypt'93, LNCS 765, Springer, pp. 386-397, 1993.
29. M. Matsui, *On correlation between the order of S-boxes and the strength of DES,* Eurocrypt'94, LNCS 950, pp. 366-375, Springer, 1995.
30. M.Matsui: *The First Experimental Cryptanalysis of the Data Encryption Standard,* Crypto'94, LNCS 839, Springer, pp. 1-11, 1994.
31. Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone: *Handbook of Applied Cryptography*; CRC Press, 1996.
32. J. Patarin, *How to construct pseudorandom and super pseudorandom permutations from one single pseudorandom function. Eurocrypt'92, Springer, pp. 256-266, 1992.*
33. *Adi Shamir: On the security of DES,* Crypto'85, LNCS 218, Springer, pp. 280-281, 1985.

A Selected Bi-Linear Characteristics of DES S-boxes

In this section we give some bi-linear characteristics for DES S-boxes. Our results are not exhaustive: the number of possible bi-linear characteristics is huge and we do not have a fast method to find all interesting characteristics. Accordingly we are not certain to have found the best existing characteristics. It is certain that there is no characteristics true with probability 1, as these are easy to check algebraically. Otherwise we explored all cases that use up to two products of linear combinations of variables. We conjecture that the other does not have practical relevance for the security of DES, (see Section F). We give here some interesting results we have found.

Table 1. Selected bi-linear characteristics for DES S-boxes (extends on 2 pages)

		equation			remarks and comments
		input	output	input*output	
$S5$	12/64	17	8, 14, 25, 3		Matsui's equation A
$S5$	27/64	17	8, 14, 25, 3	[17] ∗ [3]	not very good
$S5$	6/64	17	8, 14, 25, 3	[17] ∗ [8, 14, 25, 3]	gets better
$S5$	58/64			[17] ∗ [8, 14, 25, 3]	
$S5$	56/64			[16, 20] ∗ [8, 14, 25]	
$S5$	8/64	17	8, 14, 25, 3	[16, 17, 20] ∗ [8]	
$S5$	8/64	16, 20	8, 14, 25	[16, 20] ∗ [8, 14, 25]	
$S5$	**61/64**	16, 20	8, 14, 25, 3	[16, 17, 20] ∗ [3]	the best in DES
$S5$	47/64		8, 14, 25	17 ∗ 3	
$S5$	17/64		8, 14, 25, 3	17 ∗ 3	
$S5$	47/64			17 ∗ 3	
$S5$	49/64		3	17 ∗ 3	
$S5$	49/64	17		17 ∗ 3	
$S5$	17/64	17	3	17 ∗ 3	
$S5$	25/64	17	14, 25	17 ∗ 3	
$S5$	25/64	17	14, 25, 3	17 ∗ 3	
$S5$	27/64	17	8, 14, 25	17 ∗ 3	
$S5$	47/64		8, 14, 25	17 ∗ 3	
$S5$	25/64	19, 20, 21	8, 14, 25	17 ∗ 3	
$S5$	25/64	17, 18	8, 14, 25	17 ∗ 3	
$S5$	19/64	16, 20	8, 14, 25	17 ∗ 3	
$S5$	25/64	16, 17, 18, 20	8, 14, 25	17 ∗ 3	
$S5$	25/64	16, 17, 18, 19	8, 14, 25	17 ∗ 3	
$S5$	27/64	17	8, 14, 25, 3	17 ∗ 3	
$S5$	17/64		8, 14, 25, 3	17 ∗ 3	
$S5$	17/64	16, 20	8, 14, 25, 3	17 ∗ 3	
$S5$	23/64	18, 21	8, 14, 25, 3	17 ∗ 3	
$S5$	25/64	16, 18, 20, 21	8, 14, 25, 3	17 ∗ 3	

		equation			remarks and comments
		input	output	input*output	
$S1$	$30/64$	3	17		Matsui's equation C
$S1$	$15/64$	3	17	$3*17$	gets better
$S1$	$47/64$		17	$3*17$	
$S1$	$47/64$	3		$3*17$	
$S1$	$49/64$			$3*17$	
$S1$	$25/64$	$1,2,4,5$	17	$3*17$	
$S1$	$25/64$	$1,2,3,4,5$	17	$3*17$	
$S1$	$22/64$	$32,1,2,3,4,5$	17	$3*17$	
$S1$	$39/64$	$32,1,2,4$		$3*17$	
$S1$	$39/64$	$32,1,2,3,4$		$3*17$	
$S1$	$39/64$		$9,23$	$3*17$	
$S1$	$25/64$		$9,17,23$	$3*17$	
$S1$	$27/64$		$9,23,31$	$3*17$	
$S1$	$37/64$		$9,17,23,31$	$3*17$	
$S1$	$29/64$		$9,17,31$	$3*17$	
$S1$	$35/64$		$9,31$	$3*17$	
$S1$	$45/64$			$[1,3]*[9,23,31]$	
$S1$	$57/64$			$[1]*[9,17,23,31]$	

		equation			remarks and comments
		input	output	input*output	
$S2$	$8/64$	5	$13,28,18$	$8*2$	
$S4$	$56/64$			$[12,16]*[26,20,10,1]$	
$S4$	$56/64$			$[14,17]*[26,20,10,1]$	
$S4$	$56/64$			$[12,14,16,17]*[26,1]$	(there are many similar)
$S6$	$48/64$			$21*29$	
$S6$	$48/64$		29	$21*29$	
$S6$	$38/64$		$11,19$	$21*29$	
$S6$	$24/64$		$4,11,19$	$21*29$	
$S7$	$11/64$	$25,28$	$32,12,7$	$28*12,27*22$	
$S7$	$12/64$	$24,28$	$32,12,22$	$27*12,29*7$	
$S8$	$48/64$			$29*21$	
$S8$	$48/64$		21	$29*21$	
$S8$	$40/64$		$5,27,15$	$29*21$	
$S8$	$24/64$		$5,27,15,21$	$29*21$	

B Improved Bi-Linear Attacks for DES

The goal of this section is to find or construct examples where bi-linear cryptanalysis gives strictly better bias on DES than the best Matsui's result.

We look at the best Matsui's characteristic on 3 rounds given at the last page of [28]. By itself, it can be considered as very good, even compared to other Matsui's characteristics: it uses twice the best element (A) of Matsui, and nothing between them. Moreover, this element (A) is in itself the best linear characteristic that exist in DES, first discovered and described by Shamir in 1985 [33]:

$$\textbf{(A)} \quad J[17] \oplus O[8,14,25,3] = 0 \quad \text{for } S5 \text{ with probability } \ 12/64$$

From this we get immediately, using Matsui's Piling-Up Lemma from [28], that for 3 rounds, and for any key, the following equation is biased:

$$\left.\begin{array}{r} L_0[8,14,25,3] \oplus R_0[17] \oplus \\ L_3[8,14,25,3] \oplus R_3[17] = K[sth] \end{array}\right\} \ \frac{1}{2} - 1.56 \cdot 2^{-3}$$

We call Matsui-3 this equation.

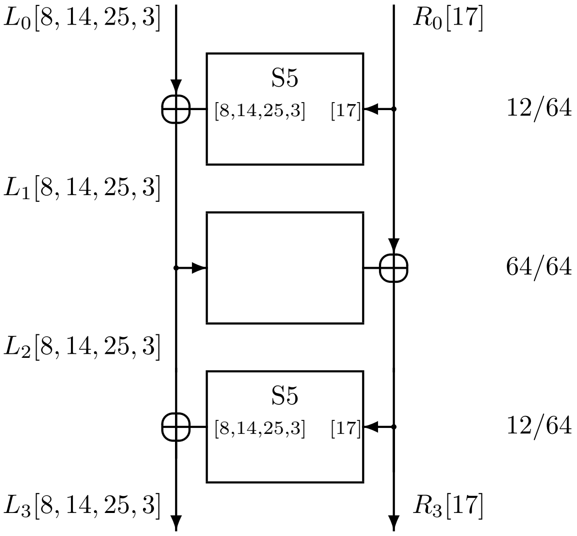

Fig. 5. Matsui's Best Linear Approximation on 3 Rounds of DES

B.1 Improving on Matsui-3

We will show that with bi-linear characteristics, there are strictly better equations than Matsui-3. Our simulations looking for the best bi-linear characteristics for DES S-boxes (cf. Appendix A), showed that the best one is the following:

$$J[16,20] \oplus O[8,14,25,3] \oplus J[16,17,20] \cdot O[3] = 0 \quad \text{for } S5 \text{ with probability } \ 61/64$$

Remark: It is clearly related to, and can be seen as "causing" the existence of the Matsui's equation (A): their difference is naturally biased.

We will use this characteristic. Let KS5 denote the combination of the S-box S5 and the key bits XORed to its inputs. It is easy to see that for KS5, if we denote by $K[sth]$ some constant linear combination of key bits, for any key, one of the following equations is always strongly biased:

$$\left\{\begin{array}{ll}\textbf{(a1)}\ I[16,20] \oplus O[8,14,25,3] \oplus I[16,17,20] \cdot O[3] = K[sth] & \\ \qquad\qquad\qquad\qquad \text{or} & |\text{bias}| = 1/2{-}3/64 \\ \textbf{(a2)}\ I[16,20] \oplus O[8,14,25] \oplus I[16,17,20] \cdot O[3] = K[sth] & \end{array}\right.$$

Remark: More precisely, in the $i - th$ round we have (a1) exactly when $K_{16}^{(S5,i)} \oplus K_{17}^{(S5,i)} = 0 \oplus K_{20}^{(S5,i)} = 0$, and we have (a2) otherwise. (We recall that $K_j^{(S5,i)}$ is the key bit XORed in round i to S5 input corresponding to the number j in the order of inputs of DES round function.)

In our construction, we will use one of the above, and we will also use another, naturally biased equation, which will be one of the following:

$$\left\{\begin{array}{ll}\textbf{(b)}\ O[16,17,20] \oplus I[3] \cdot O[16,17,20] = 0 & \\ \qquad\qquad\qquad\qquad \text{and} & |\text{bias}| = 1/2 - 1/4 \\ \textbf{(c)}\ I[3] \oplus O[16,17,20] \oplus I[3] \cdot O[16,17,20] \cdot O[3] = 0 & \end{array}\right.$$

Now we are ready to construct characteristics for 3 rounds of DES.

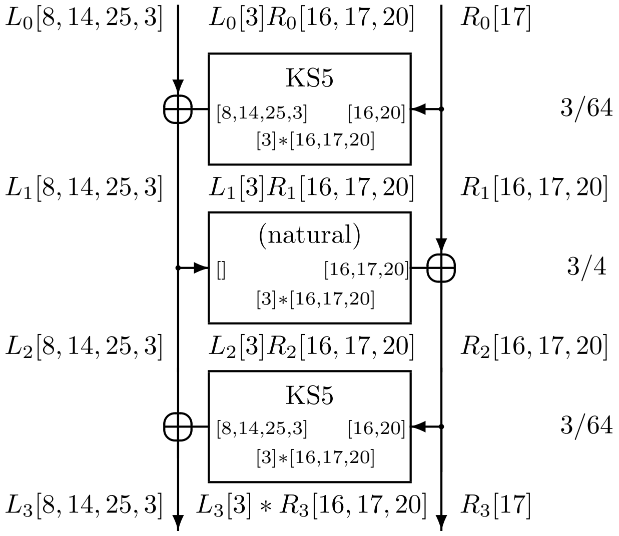

Fig. 6. Combining a1-b-a1 to get a characteristic for 3 rounds of DES

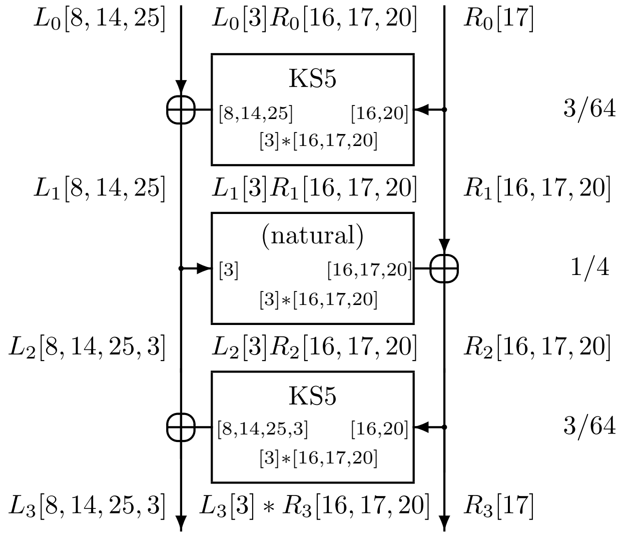

Fig. 7. Combining a2-c-a1 to get a characteristic for 3 rounds of DES

As one should expect, our construction goes as follows:

◇ In round 1 and 3, depending on the key either a1 or a2 is strongly biased.

◇ To connect a1 to a1, or a2 with a2, we can use b, as in Figure 6.

◇ To connect a1 with a2 and the reverse, we use c, as in Figure 7.

◇ For 3 rounds and for any key, we always have a strong bias on one of the four possibilities: a1-b-a1, a1-c-a2, a2-c-a1, a2-b-a2.

◇ From Matsui's Piling-Up Lemma, we expect that the whole characteristic will be true with probability $\frac{1}{2} \pm 1.64 \cdot 2^{-3}$. Our simulations show that it is between $\frac{1}{2} \pm 1.65 \cdot 2^{-3}$ and $\frac{1}{2} \pm 1.67 \cdot 2^{-3}$.

◇ Since, the choice of a1/a2 depends on a linear combination of key bits, We can combine all these into one equation and we get the following result:

Proposition B.1.1 (Our Best Attack on 3 Rounds of DES). For all keys, the following equation is biased for 3 rounds of DES: :

$$(**)\quad \left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[17]\oplus \\ L_3[3,8,14,25] \oplus L_3[3]R_3[16,17,20] \oplus R_3[17] = \\ K[sth] + K[sth']L_0[3] + K[sth'']R_3[3] \end{array}\right\} \; \frac{1}{2} \pm 1.66 \cdot 2^{-3}$$

In comparison, Matsui-3 gives $\frac{1}{2} - 1.56 \cdot 2^{-3}$. Bi-linear cryptanalysis works better than LC. In the next section we will extend this result (and again beat Matsui) to 7, 11 and more rounds.

Remark: The equation above can be seen as 4 different equations, each of them is highly biased for 1/4 of all keys. We observed that each of the 4 equations is also biased for all DES keys, except that for 3/4 of them the bias is much weaker, we get about $\frac{1}{2} \pm 1.6 \cdot 2^{-7}$.

B.2 Extending the Result for 7, 11 and More Rounds

The idea is to find an element (maybe not very good in itself) that will allow to connect together our (very good) characteristics on 3 rounds. For example, to connect Figure 6 with Figure 7 we use the following element:

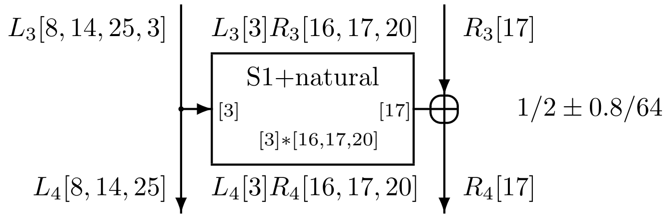

Fig. 8. Connecting the output of a1 to the input of a2

Simulations show that, for any key, this characteristic is true with probability about 1/2 ± 0.8/64. The explanation is as follows: the bias is due to to the combination of Matsui's equation (C)

$$\textbf{(C)} \quad J[3] \oplus O[17] = 0 \quad \text{for } S1 \text{ with probability } \ 30/64$$

and of the fact that $I[3] \cdot O[16,17,20]$ is naturally biased. The same element (Figure 8) does also work to connect a2 to a1.

It remains to be seen how the connection between a1 and a1 or a2 and a2. This is done in a very similar way: we combine (C) with $I[3] \oplus I[3] \cdot O[16,17,20]$ that is also naturally biased.

Summary: In every of 4 possible cases, there is a connecting element based on (C). This means that, also for 7 rounds and for any key, again one of the four possibilities is quite biased: a1-b-a1, a1-c-a2, a2-c-a1, a2-b-a2. Again we can recompose it in a single attack:

Proposition B.2.1 (Extension to 7 Rounds of DES). For all keys, the following equation is biased for 7 rounds of DES:

$$\left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[17] \oplus \\ L_7[3,8,14,25] \oplus L_7[3]R_7[16,17,20] \oplus R_3[17] = \\ K[sth] + K[sth']L_0[3] + K[sth'']L_7[3] \end{array}\right\} \ \frac{1}{2} \pm \text{about } 2^{-9}$$

This bias is, depending on the key, sometimes better, sometimes worse than Matsui-7 that gives $\frac{1}{2} - 1.95 \cdot 2^{-10}$.

Finally, it is now obvious, that our construction works also for 11, 15, 19 rounds etc. We verified experimentally that for 11 rounds we have:

Proposition B.2.2 (Our Best Attack on 11 Rounds of DES). For all keys, the following equation is biased for 11 rounds of DES: :

$$\left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[17] \oplus \\ L_{11}[3,8,14,25] \oplus L_{11}[3]R_{11}[16,17,20] \oplus R_{11}[17] = \\ K[sth] + K[sth']L_0[3] + K[sth'']L_{11}[3] \end{array}\right\} \ \frac{1}{2} \pm \text{around } 1.2 \cdot 2^{-15}$$

For a few different keys we have tried (long computation on a PC) the bias was **always strictly better than Matsui-11** that gives $\frac{1}{2} - 1.91 \cdot 2^{-16}$.

Remark: The best characteristics found by Matsui for 3 and 11 rounds [28] are closely related to those presented here: their difference is a biased Boolean function. BLC contains LC not only as a subset, but also as an extension allowing to strictly improve the best linear attacks on DES by adding higher degree monomials.

B.3 Beyond Bi-Linear Attacks: Using Cubic Equations

We observed that, for 3 rounds, even better results can be achieved using cubic partially bi-linear characteristics, instead of quadratic bi-linear (**) from Proposition B.1.1. Our simulations show that, for an important fraction of keys:

$$(***)\ \left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20]R_0[17,18,19,20] \oplus \\ L_3[3,8,14,25] \oplus L_3[3]R_3[16,17,20]R_3[17,18,19,20] \oplus \\ R_0[17] \oplus R_3[17] = K[sth] \end{array}\right\} \ \frac{1}{2} - 1.82 \cdot 2^{-3}$$

The explanation why this works is quite similar. Though the non-linear part of this equation is not bi-linear, it is well correlated with a truly bi-linear function:

$$L[3]R[16,17,20]R[17,18,19,20] = L[3]R[16,17,20] \quad \text{with probability } 7/8$$

Unfortunately, the bias of $(***)$ is worse for other keys. On average, the best bias we know for 3 rounds remains $(**)$ from Proposition B.1.1. We also observed that that $(***)$ works for any number of DES rounds and for any key, but again the results are not as good as with $(**)$.

B.4 Some Other Examples of Bi-Linear Characteristics for DES

Table 2. Selected bi-linear characteristics for DES found by various methods. Biases are given for key = 0, for other keys they are usually very close. Results strictly better than the best result of Matsui are marked with a "!").

scheme	rounds	characteristic	result	
DES	3	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.84\cdot 2^{-5}$	
DES	3	$L[3,8,14,25]\oplus R[16,20]\oplus L[3,14]R[16,17,20]$	$1.04\cdot 2^{-3}$	
DES	3	$L[3,8,14,25]\oplus R[17]\oplus L[3,7]R[16,17,20]$	$1.05\cdot 2^{-3}$	
DES	3	$L[3,8,14,25]\oplus R[17]\oplus L[3,19]R[16,17,20]$	$1.08\cdot 2^{-3}$	
DES	3	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$\mathbf{1.65\cdot 2^{-3}}$	!
DES	3	$L[3,8,14,25]\oplus R[16,20]\oplus L[3]R[16,17,20]$	$\mathbf{1.67\cdot 2^{-3}}$	!
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.83\cdot 2^{-8}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[2]R[17]$	$1.16\cdot 2^{-7}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$1.92\cdot 2^{-8}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3]*R[16,17,20]\oplus L[1]*R[32]$	$1.35\cdot 2^{-7}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3]*R[16,17,20]\oplus L[1,3]*R[31]$	$1.57\cdot 2^{-7}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3]*R[16,17,20]\oplus L[1,3]*R[16,20]$	$1.65\cdot 2^{-7}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3,4,32]*R[16,17,20]$	$1.27\cdot 2^{-6}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[3,4,5,32]*R[16,17,20]$	$1.27\cdot 2^{-6}$	
DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[1,3]*R[16,17,20]$	$\mathbf{1.30\cdot 2^{-6}}$	
DES	5	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.59\cdot 2^{-9}$	
DES	5	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$1.4\cdot 2^{-11}$	
DES	5	$L[8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.70\cdot 2^{-9}$	
DES	5	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]\oplus L[1,25]R[16,20]$	$1.12\cdot 2^{-8}$	
DES	5	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]\oplus L[1]R[12,15,20]$	$1.30\cdot 2^{-8}$	
DES	5	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]\oplus L[1]R[12,14,15,17,20]$	$\mathbf{1.34\cdot 2^{-8}}$	
DES	6	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.3\cdot 2^{-10}$	
DES	6	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$1.6\cdot 2^{-12}$	
DES	6	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]\oplus L[4,32]R[16,20]$	$1.9\cdot 2^{-11}$	
DES	6	$L[8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\mathbf{1.65\cdot 2^{-10}}$	
DES	7	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.2\cdot 2^{-11}$	
DES	7	$L[3,8,14,25]\oplus L[3]R[17]$	$1.45\cdot 2^{-11}$	
DES	7	$L[3,8,14,25]\oplus R[16,20]\oplus L[3]R[16,17,20]$	$\mathbf{1.15\cdot 2^{-9}}$	!
DES	7	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$\mathbf{1.42\cdot 2^{-9}}$	!
DES	8	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$1.8\cdot 2^{-14}$	
DES	8	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.9\cdot 2^{-14}$	
DES	8	$L[3,8,14,25]\oplus R[17]\oplus L[3]*R[16,17,20]\oplus L[1,3]*R[16,20]$	$\mathbf{1.42\cdot 2^{-13}}$	
DES	9	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.8\cdot 2^{-15}$	
DES	9	$L[3,8,14,25]\oplus R[17]\oplus L[3]*R[16,17,20]\oplus L[1,3]*R[16,20]$	$1.9\cdot 2^{-15}$	
DES	9	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]\oplus L[1]R[12,14,15,17,20]$	$\mathbf{1.3\cdot 2^{-14}}$	
DES	10	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.2\cdot 2^{-16}$	
DES	11	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.7\cdot 2^{-18}$	
DES	11	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	$\mathbf{1.2\cdot 2^{-15}}$	!
DES	12	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$1.1\cdot 2^{-18}$	
DES	13	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\approx 2^{-20}$	
DES	14	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\approx 2^{-21}$	
DES	14	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[16,17,20]$	2^{-20}	

C Complementation Theorem for Bi-Linear Cryptanalysis of DES

Theorem C.0.1 (Complementation Result for DES). Let $X_i, i = 1 \dots m$ be a freely chosen subset of input and output bits of DES (or a reduced version of DES for k rounds). Let $K_i, i = 1 \dots n$ be some key bits. If the equation $F(X_1, \dots, X_m; K_1 \dots K_n)$ is biased for some fraction of keys, then the equation $F(X_1 + 1, \dots, X_m + 1; K_1 + 1 \dots K_n + 1)$ has exactly the same bias for the same fraction of keys.

Proof: Follows immediately from the complementation property of DES. □

For LC this result is trivial and gives no information. For BLC it is non-trivial, and gives interesting information. We see that the complementation property of DES does not only have security implications on the exhaustive search, but can be used much more frequently to derive bi-linear attacks from the existing ones.

Example: Our equation $(**)$ is the best known bi-linear characteristic for 11 rounds of DES.

$$(**) \quad \left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[17] \oplus \\ L_3[3,8,14,25] \oplus L_3[3]R_3[16,17,20] \oplus R_3[17] = \\ K[sth] + K[sth']L_0[3] + K[sth'']R_3[3] \end{array}\right\} \quad \frac{1}{2} \pm \text{around } 1.2 \cdot 2^{-15}$$

From our theorem we deduce that the following equation is equally good (the parts $K[sth]$ are the same than in $(**)$):

$$(\overline{**}) \quad \left.\begin{array}{l} L_0[3,8,14,25] \oplus L_0[3]R_0[16,17,20] \oplus R_0[16,20] \oplus \\ L_{11}[3,8,14,25] \oplus L_{11}[3]R_{11}[16,17,20] \oplus R_{11}[16,20] = \\ (K[sth]+K[sth']+K[sth'']) + K[sth']L_0[3] + K[sth'']L_{11}[3] \end{array}\right\} \quad \frac{1}{2} \pm \text{around } 1.2 \cdot 2^{-15}$$

Remark: Looking at these equations, and knowing that their bias depends (slightly) on the key, presumably in a linear way, the following question can be asked: is it possible to combine $(**)$ and $(\overline{**})$ in an even better characteristic ? The answer is negative. We have verified experimentally (100 random keys tried) that their current dependence on the key bits already allows to achieve the best possible bias for every key, and the "mirror" equation allows to achieve the same bias and not a better one. It turns out that, for any fixed key, the best bi-linear characteristic we know **is simultaneously attained by two different equations**. Such a situation, as far as we know, did not happen so far in LC (except the usual rigt/left and up/down symmetries).

D Bi-Linear Attacks on s^{5}DES

S^{5}DES is a version of DES with DES-boxes modified to resist all known attacks on DES [22]. In particular it is a good deal more resistant than DES against linear cryptanalysis. However, for bi-linear cryptanalysis we have already found a few examples (for now with a small number of rounds), where s^{5}DES is worse than even DES itself (!). This work is still in progress.

D.1 Bi-Linear Properties of s^{5}DES S-boxes

In this section we give some interesting bi-linear characteristics for s^{5}DES S-boxes. Surprisingly, s^{5}DES, designed to be much more secure than DES against LC and other known attacks, is much weaker than DES at this point. Indeed, for each of the S-boxes, there is exactly one bi-linear equation that is true with probability 1:

Table 3. Selected bi-linear characteristics for s^{5}DES S-boxes

		equation		
		input	output	input*output
$S1$	0/64	1	$17, 23, 31$	$5 * 9$
$S2$	64/64	5	$28, 2, 18$	$9 * 13$
$S3$	0/64	9	$24, 30, 6$	$13 * 16$
$S4$	64/64	13	$20, 10, 1$	$17 * 26$
$S5$	0/64	$17, 21$	$8, 25, 3$	$21 * 14$
$S6$	0/64	21	$4, 29, 11$	$25 * 19$
$S7$	64/64	25	$32, 22, 7$	$29 * 12$
$S8$	0/64	29	$27, 15, 21$	$1 * 5$

These equations have a lot in common. If we change the notations and denote the inputs of the S-box $x_1, \ldots, x_6$ and the outputs by $y_1, \ldots, y_4$, the same equations become:

```
S1. 1+x[2]+          y[2]+y[3]+y[4]+x[6]*y[1] = 0
S2.   x[2]+          y[2]+y[3]+y[4]+x[6]*y[1] = 0
S3. 1+x[2]+     y[1]+     y[3]+y[4]+x[6]*y[2] = 0
S4.   x[2]+          y[2]+y[3]+y[4]+x[6]*y[1] = 0
S5. 1+x[2]+x[6]+y[1]+     y[3]+y[4]+x[6]*y[2] = 0
S6. 1+x[2]+     y[1]+y[2]+y[3]+     x[6]*y[4] = 0
S7.   x[2]+     y[1]+     y[3]+y[4]+x[6]*y[2] = 0
S8. 1+x[2]+          y[2]+y[3]+y[4]+x[6]*y[1] = 0
```

We see that the input bit is (almost) always x_2, the same that correspond to left/right half in the DES tables. The input bit in the product is always x_6, and $x_6 = 1$ corresponds to the second and last line in DES S-boxes. This gives the impression that the authors of s^{5}DES have derived all its S-boxes from a single (and apparently not so good) S-box. This is very closely related to strange properties of s^5DES discovered by Guilhem Castagnos and described in the Appendix of [6].

D.2 Interesting Approximations for 3 Rounds of s^{5}DES

It is possible to see, that each of the above 8 equations, allows to build a characteristic for 3 rounds of s^{5}DES as follows:

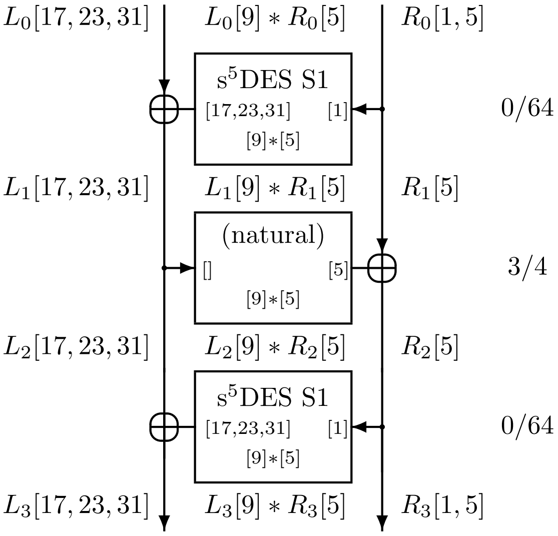

Fig. 9. A bi-linear attack on 3 rounds of s^{5}DES

We see that, at least when all key bits are 0, we have 8 bi-linear characteristics for s^{5}DES that are true with probability about $1/2 \pm 1/4$. This is better than Matsui-3 for DES, and better than the best bi-linear characteristic we found for 3 rounds of DES itself.

For some keys, one of these biases is even better than $1/2 \pm 1/4$:

$$\left.\begin{array}{l} L_0[4,11,29] \oplus L_0[19]R_0[25] \oplus R_0[21,25] \oplus \\ L_3[4,11,29] \oplus L_3[19]R_3[25] \oplus R_3[21,25] = K[sth] \end{array}\right\} \ \frac{1}{2} - 1.06 \cdot 2^{-2}$$

In comparison, the best linear characteristic for DES gives $1.56{\cdot}2^{-3}$ (improved to $1.82 \cdot 2^{-3}$ in Section B). Clearly, for 3-round I/O sums, s^{5}DES is weaker than DES itself.

D.3 Bi-Linear Attacks for s^{5}DES - 4 rounds

Unfortunately, it is possible to see that the above characteristics work only for one half of all keys. Moreover, we were not able to extend it for 4 rounds. This fails in fact for all the 8 "good" characteristics on 3 rounds that can be built (in the same way) from table 3.

For s^{5}DES the best characteristics for 3 rounds are better than for DES. However, very surprisingly, for 4 rounds, it is very difficult to find many biased characteristics for s^{5}DES, it is clearly harder than for DES. The method of combining known characteristics exploited by Matsui with success [28, 30], fails quite badly here. We tried many combinations of 2+2 or 3+1 rounds, that in theory should work... yet in practice they don't at all: the combined characteristic is not biased or the bias was so small that we could not detect it. Finally we have found some biased bi-linear characteristics for 4 rounds. Currently, (one of) the best we know is the following:

$$\left.\begin{array}{l} L_0[29] \oplus R_0[15,21,27] \oplus L_0[29]R_0[20,21,22,24,25] \oplus \\ R_4[29] \oplus L_4[15,21,27] \oplus R_4[29]L_4[20,21,22,24,25] = K[sth] \end{array}\right\}$$

With this, for a fraction of keys, we get $1/2 - 1.85 * 2^{-7}$ for 4 rounds, and $1/2 - 1.54 * 2^{-8}$ for 5 rounds.

D.4 Bi-Linear Attacks for More than 4 Rounds of s^{5}DES

The later bias seems to disappear for 6 rounds (10 billions of plaintexts tried). However, we have found some other similar characteristics that work for 6 rounds... These are given in Table 4, Appendix E. It may be possible to find better results. Our work on s^5 DES is still in progress but at present at seems that the whole cipher should have a sufficient security margin to remain secure against BLC, as it is quite secure for LC, and this probably for reasons we give in Appendix F.

D.5 Cubic GLC Attacks for s^{5}DES

As for DES, we have also found some very strongly biased cubic I/O sums, for example for 3 rounds and for a fraction of keys:

$$\left.\begin{array}{l} L_0[15,21,27] \oplus L_0[5]R_0[1]R_0[28,29,31,32] \oplus R_0[1,29] \oplus \\ L_3[15,21,27] \oplus L_3[5]L_3[1]R_3[28,29,31,32] \oplus R_3[1,29] = K[sth] \end{array}\right\} \quad \frac{1}{2} - 1.02 \cdot 2^{-2}$$

$$\left.\begin{array}{l} L_0[15,21,27] \oplus L_0[5]R_0[1]R_0[1,28,29,31,32] \oplus R_0[29] \oplus \\ L_3[15,21,27] \oplus L_3[5]R_3[1]R_3[1,28,29,31,32] \oplus R_3[29] = K[sth] \end{array}\right\} \quad \frac{1}{2} - 1.01 \cdot 2^{-2}$$

E A General Heuristic Construction for Feistel Ciphers

We present here a very general, rather surprising and even controversial result: we construct a large family of simple invariant bi-linear attacks that work for 1,2,3,4 etc.. rounds for a large family of Feistel ciphers, and with a fairly weak requirement on the S-boxes. Later we will see however that that only a tiny fraction of them will work well for an important number of rounds.

To begin with, let us assume that the block cipher is as follows:

1. It is a Feistel cipher.
2. Every even round is arbitrary, (could even be a truly random function).
3. Every odd round has "almost the same" round function. More precisely we require that every odd round function is built as follows:
   a. A fixed (identical in every odd round) injective multivariate linear operation (for example expansion in DES).
   b. Then a different session key is XORed (as in DES),
   c. Then we apply an arbitrary fixed non-linear component (e.g. the lot of DES S-boxes).
4. The non-linear part of the round functions used for all odd rounds have one common linear characteristic that is biased (one will be sufficient).

Then **the following invariant bi-linear characteristic will be biased for 1, 2, 3, and more rounds:**

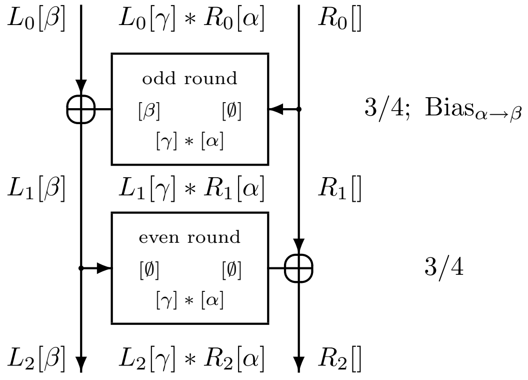

Fig. 10. How to use **any** linear characteristic in every other round in a bi-linear attack

Explanation: In the even rounds we use the fact that $I[\gamma]O[\alpha]$ is naturally biased. In the odd rounds we use the analogous bias on $I[\alpha](O[\gamma]+1)$ combined with the linear bias on $I[\alpha] \oplus O[\beta]$. In LC $I[\alpha]$ may be in general difficult to connect with the previous/next rounds, especially if α is a large subset of bits. Here it gets simply eliminated.

E.1 Extending the General Construction

We construct 2^4 similar characteristics as shown on Figure 11. In the even rounds we use the fact that each of these four Boolean functions is always biased: $I[\gamma]O[\alpha] \oplus I[\gamma] \oplus O[\alpha]$, or $I[\gamma]O[\alpha] \oplus I[\gamma]$ or $I[\gamma]O[\alpha] \oplus O[\alpha]$ or $I[\gamma]O[\alpha]$. In the odd rounds we use the analogous fact that is combined with the linear bias on $I[\alpha] \oplus O[\beta]$.

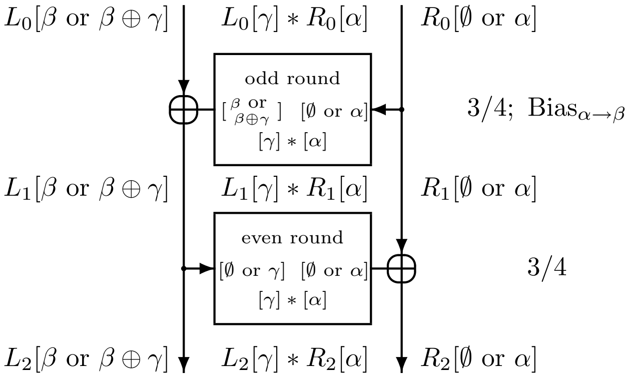
$L_0[\beta$ or $\beta \oplus \gamma]$
$L_0[\gamma] * R_0[\alpha]$
$R_0[\emptyset$ or $\alpha]$
odd round
$[\beta$ or $\beta\oplus\gamma]$
$[\emptyset$ or $\alpha]$
$[\gamma] * [\alpha]$
3/4; $\text{Bias}_{\alpha\to\beta}$
$L_1[\beta$ or $\beta \oplus \gamma]$
$L_1[\gamma] * R_1[\alpha]$
$R_1[\emptyset$ or $\alpha]$
even round
$[\emptyset$ or $\gamma]$
$[\emptyset$ or $\alpha]$
$[\gamma] * [\alpha]$
3/4
$L_2[\beta$ or $\beta \oplus \gamma]$
$L_2[\gamma] * R_2[\alpha]$
$R_2[\emptyset$ or $\alpha]$

Fig. 11. Deriving more bi-linear attacks given one linear characteristic in every other round

In this extended version all the choices are not allowed: when composing characteristics together, the quadratic parts do not change, linear parts combine as in linear cryptanalysis. As a consequence all the masks connected to a $\oplus$ have to be the same. Still it is possible to see that for 2 rounds, with an arbitrary choice of any of 2^4 possibilities, we may adjust the internal masks for the round functions in such a way that the whole will still be biased (based on any of the biased equations in $I[\alpha]$ etc. listed above). Extension for more than 2 rounds is immediate, and for 1 round it also works except that the masks on the left side must be identical.

Thus, given a Feistel cipher with 64-bit block size, from one single linear characteristic $\alpha \to \beta$ for the even rounds, this construction gives as many as $2^4 \cdot 2^{32}$ or $2^3 \cdot 2^{32}$ bi-linear characteristics. Our construction applies to DES and many other similar ciphers (for some other ciphers, such as ICE [25], our initial requirements can be relaxed and our still applies with good probability). There are also other similar constructions. For example clearly if the output γ is be different from the input γ, the attack should still work (both product will be correlated). We expect that all these characteristics are always or (almost always) biased for any number for rounds > 1.

Remark: This construction is quite nice and simple. It is easy to see that if this construction worked as predicted by Matsui's Piling-up lemma from [28], then with the best linear approximation known in DES used 7 times, and the natural bias 3/4 used 14 times, we get a characteristic for 14 rounds (one example is displayed on Figure 4) with a theoretical bias of $1.19 \cdot 2^{-20}$ which would allow ot break full DES given about $2^{39.5}$ plaintexts. It turns out that for our example of Figure 4 the actual bias is not much worse than this for 14 rounds, yet in most other cases the biases obtained by our method work as predicted only for a small number of rounds.

E.2 Real-Life Behaviour of Our Heuristic Attack on Feistel Ciphers

For 2 rounds the method works perfectly well for any α, β and γ. The bias is with good precision equal to combining with Matsui's Piling-up rule, the 3 biases $(3/4, 3/4, \text{Bias}_{\alpha\to\beta})$ with $\text{Bias}_{\alpha\to\beta}$ being the bias of the initial linear characteristic. In general very good results are obtained for $1, 2$ or 3 rounds.

For 4 rounds the bias is very frequently substantially lower than expected. For 5,6 and more rounds in most cases our construction does **not** give good results at all. This construction is a general framework, and one should **not** use this construction to build many good attacks for ciphers that have many rounds However, this (and similar) constructions proved to be a great tool for finding very good attacks on practical ciphers. This is because, we expect that when exploring by computer simulations a large number (here 2^{36}) of very weak attacks we occasionally find a very good attack. Besides, since in most bi-linear attacks the expected bias turns out to be very different from Matsui's predictions, this method is as good as any other method. In fact it is about **the only method we know** to systematically find working attacks on practical ciphers (it allowed to find our invariant attack on DES).

We did a lot of simulations for this (and similar) constructions, to see what will be the best values for (α, β, γ) for DES, s^{5}DES and ICE. Below we give the best results obtained. Only invariant characteristics are displayed (input and output expressions the same). **Remark:** For DES we have found better attacks, by different methods. These are displayed in Table 2 in Appendix B.4.

For DES and s^{5}DES the bias values given here are true when all the key bits are zero. For other keys they are usually very close but not always identical. All the displayed characteristics are systematically biased for any number of rounds (but if one characteristic may be the best for 2 rounds, in general a different one will be the leading result for a different number of rounds.) Exceptionally for ICE [25], that has key-dependent bit permutations the biases displayed work only for a (substantial) fraction of keys. (This cipher has big S-boxes that have no interesting bi-linear approximations and we did not find very interesting bi-linear attacks on it for ≥ 6 rounds.)

Table 4. Some Experimental Results on Our General Construction

scheme	rounds	characteristic	result
DES	2	$L[8,14,25]\oplus R[]\oplus L[3]R[17]$	$\frac{1}{2} - 1.00 \cdot 2^{-3}$
DES	2	$L[8,14,25,3]\oplus R[]\oplus L[3]R[17]$	$\frac{1}{2} + 1.00 \cdot 2^{-3}$
DES	2	$L[4,29,11,19]\oplus R[]\oplus L[4]R[21]$	$\frac{1}{2} - 1.50 \cdot 2^{-4}$
s^{5}DES	2	$L[5,15,21,27]\oplus R[29]\oplus L[20,21,22,24,25]R[29]$	$\frac{1}{2} - 1.47 \cdot 2^{-4}$
s^{5}DES	2	$L[3,8,25]\oplus R[17]\oplus L[14]R[17]$	$\frac{1}{2} + 1.00 \cdot 2^{-3}$
ICE	2	$L[2,22,27]\oplus R[] \oplus L[9]R[18]$	$\frac{1}{2} - 1.91 \cdot 2^{-7}$
DES	4	$L[5,15,21,27]\oplus R[29]\oplus L[20,21,22,23,24,25]R[29]$	$\frac{1}{2} - 1.56 \cdot 2^{-7}$
DES	4	$L[9,17,23,31]\oplus R[1]\oplus L[12,13,14,16,17]R[1]$	$\frac{1}{2} - 1.55 \cdot 2^{-7}$
DES	4	$L[8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} - 1.87 \cdot 2^{-8}$
DES	4	$L[3,8,14,25]\oplus L[1,2,4,5]R[17]$	$\frac{1}{2} - 1.95 \cdot 2^{-7}$
s^{5}DES	4	$L[3,8,14,25]\oplus R[17]\oplus L[1,2,3,5,32]R[17]$	$\frac{1}{2} - 1.55 \cdot 2^{-7}$
s^{5}DES	4	$L[15,21,27]\oplus R[29]\oplus L[29]R[20,21,22,24,25]R[29]$	$\frac{1}{2} - 1.85 \cdot 2^{-7}$
ICE	4	$L[2,22,27]\oplus R[]\oplus L[9]R[18]$	$\frac{1}{2} - 1.05 \cdot 2^{-14}$
DES	5	$R[17]\oplus L[8,14,25]\oplus L[3]R[17]$	$\frac{1}{2} + 1.70 \cdot 2^{-9}$
s^{5}DES	5	$L[1,10,20]\oplus R[13]\oplus L[5,6,9]R[13]$	$\frac{1}{2} - 1.12 \cdot 2^{-8}$
s^{5}DES	5	$L[15,21,27]\oplus R[29]\oplus L[20,21,22,24,25]R[29]$	$\frac{1}{2} - 1.54 \cdot 2^{-8}$
DES	6	$L[3]\oplus R[9,23]\oplus L[3]R[17]$	$\frac{1}{2} + 1.6 \cdot 2^{-13}$
DES	6	$L[3]\oplus R[9,17,23]\oplus L[3]R[17]$	$\frac{1}{2} + 1.5 \cdot 2^{-13}$
DES	6	$L[6,16,24,30]\oplus R[10]\oplus L[16]R[10]$	$\frac{1}{2} - 1.7 \cdot 2^{-13}$
DES	6	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1.30 \cdot 2^{-10}$
DES	6	$L[8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} - 1.65 \cdot 2^{-10}$
s^{5}DES	6	$L[17,23,31]\oplus R[1]\oplus L[13,14,15]R[1]$	$\frac{1}{2} - 1.46 \cdot 2^{-13}$
s^{5}DES	6	$L[5,15,21,27]\oplus R[29]\oplus L[20,22,23,25]R[29]$	$\frac{1}{2} - 1.30 \cdot 2^{-13}$
DES	7	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1.2 \cdot 2^{-11}$
DES	7	$L[3,8,14,25]\oplus R[]\oplus L[3]R[17]$	$\frac{1}{2} + 1.45 \cdot 2^{-11}$
DES	11	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1.7 \cdot 2^{-18}$
DES	12	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1.1 \cdot 2^{-18}$
DES	13	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1 \cdot 2^{-19}$
DES	14	$L[3,8,14,25]\oplus R[17]\oplus L[3]R[17]$	$\frac{1}{2} + 1.? \cdot 2^{-20}$

E.3 Downsizing Our General Construction and Gaining Insights

In this section we will give some conjectured necessary conditions for our construction to work well for a substantial number of rounds of a Feistel cipher similar to DES. This result is heuristic but of great practical value: allows to construct realistic bi-linear attacks on practical ciphers such as DES or s^{5}DES. In the general case, we don't think that such restrictions exist and there are probably many block ciphers for which specific special cases of our construction will work well.

By looking at the leading results obtained by intensive computer simulations, we observed that for $4, 5, 6$ and more rounds all really "good" results are such that:

(A) $\alpha \rightarrow \beta$ is a good linear approximation for some S-box Si.
This is required already be the construction, we assume in addition that it uses only one S-box, since the bias of $\alpha \rightarrow \beta$ is used many times in every second round of the attack, it is very reasonable to think that using several S-boxes for this leads to much weaker attacks)

(B) $\gamma \rightarrow \alpha$ is a biased linear approximation for another S-box Sj (the bias has not be as good as (A)).

With these two the construction uses one linear approximation in each round.

We remark that for DES, since the P-box always distribute that outputs of one S-box to 4 different S-boxes in the next round, these two conditions (A) and (B) can be satisfied only if $i \neq j$ and the set α has one element.

E.4 Additional Criteria (More Heuristic)

We also observed that in DES, for these few examples we know that work really well for 6, 8 and more rounds, the following two conditions are also satisfied:

(C) $\alpha \rightarrow \beta \oplus \gamma$ is a biased linear approximation for the same S-box Si.
(C') $\alpha \rightarrow \gamma$ is a biased linear approximation for the same S-box Si.

Equivalence of (C) and (C'): Each of these two properties (C) and (C'), when combined with (A), implies the other property.

Strictly speaking, the best examples that we know for 6 rounds do satisfy $(A - C)$. This is not any longer an exact science, very few pairs of strong linear characteristic in DES exist and very few applications of our construction give biases for 6 and more rounds being large enough to be tested in practice on a regular basis. Yet we know that (at least for 6 rounds), **the condition (C) is not at all necessary**. In Table 4 the first three examples for 6 rounds do not satisfy it at all. Finally it is probably possible (open problem) to construct contrived ciphers to see that the condition (B) is not absolutely necessary either.

E.5 Summary and Further Research

Our construction always works well for a small number of rounds. It is an open problem to give a necessary and sufficient condition for our construction to work for an important number of rounds. At present testing experimentally all possibilities that satisfy (A) and (B) and selecting the best results seems to be the best method. It applied to nearly arbitrary Feistel schemes and allows (at least for DES) to find some interesting attacks on them.

F Links Between BLC and LC

It is natural to assume that equations on different round of a cipher are (at least to some extent) independent and to estimate the bias of composed BLC characteristics using the Matsui's Piling-up Lemma. Our experience on real-life ciphers shows however that, though this does allow to construct very few efficient and interesting attacks, very frequently it will give both misleading and disappointing results, especially when the number of rounds is large. In this section we will explain that there are somewhat deep reasons for this. We claim that beyond composition, there is another (heuristic) law that (apparently) does govern bi-linear attacks. It seems that, due to the construction of DES, BLC should always remain correlated to LC, and thus can be only slightly better, but probably not be much better. The argument however does not apply to other Feistel ciphers, and in general there is no doubt that BLC can be much faster than LC (see Section 5).

F.1 Preliminaries: Observations on DES and Similar Ciphers

The idea is as follows: if we want to have good bi-linear characteristic on one round of DES, we need to use good bi-linear characteristics for one S-box in a round (attacks using two or more S-boxes per round will probably be worse). Let A be the set of inputs and B set of outputs of this S-box. Then in the following round, in which the S-boxes are connected in the opposite direction, let C and D be the set of input/outputs used. Following Section 3.3, the homogenous quadratic parts in the successive rounds should be correlated. This, again heuristically, means that for "really efficient" attacks they will probably be equal, and therefore we probably have $A = D$ and $B = C$. We know however that the DES P-box scatters outputs of one S-box over inputs of different S-boxes in the next round. Thus, assuming that these sets are nonempty, by inspection we see that this implies that each set A and B contains one element. We come to the preliminary conclusion than to have a "very good" attack on DES with period 2, we need two positions a and b such that one S-box in even round a connected to an input and b to an output, and for another S-box in odd rounds, b connected to an input and a to an output.

Table 5. Pairs of S-boxes in DES that are connected to one another in the previous/next round by the DES P-box

1 15		$S8 - S4$	8 18		$S2 - S5$	16 10		$S4 - S3$	24 12		$S7 - S3$
1 17		$S1 - S4$	9 2		$S2 - S1$	17 1		$S4 - S1$	25 19		$S6 - S5$
2 9		$S1 - S2$	10 16		$S3 - S4$	17 3		$S5 - S1$	26 12		$S7 - S4$
3 17		$S1 - S5$	11 24		$S3 - S6$	18 8		$S5 - S2$	27 32		$S7 - S8$
4 23		$S1 - S6$	12 24		$S3 - S7$	19 25		$S5 - S6$	28 5		$S8 - S2$
5 28		$S2 - S8$	12 26		$S4 - S7$	20 14		$S5 - S4$	28 7		$S7 - S2$
5 31		$S1 - S8$	13 6		$S3 - S2$	21 29		$S6 - S8$	29 21		$S8 - S6$
6 13		$S2 - S3$	14 20		$S4 - S5$	22 29		$S6 - S7$	29 22		$S7 - S6$
7 28		$S2 - S7$	15 1		$S4 - S8$	23 4		$S6 - S1$	31 5		$S8 - S1$
8 16		$S3 - S5$	16 8		$S5 - S3$	24 11		$S6 - S3$	32 27		$S8 - S7$

This is possible, and we verified that for DES there are exactly 20 such couples. We we list them in Table 5. Accordingly, it seems that the best bi-linear attack on DES should probably use only one product in the homogenous quadratic part. This may not always be true, but we would expect that:

Conjecture F.1.1 (Heuristic Remark). The best bi-linear attack on ciphers similar to DES with small S-boxes and a P-box with "good diffusion properties" will probably use a small number (e.g. 1 or 2) of products of variables.

Remark: For general Feistel ciphers other than DES, things are very different as demonstrated by our "Killer Example" of Section 5.

F.2 "Gravity Law" for Bi-Linear Cryptanalysis of DES

Our Conjecture F.1.1 above suggests that for DES, and similar ciphers, the best bi-linear attack is likely to use only a few products. We fix the key and look at the best bi-linear characteristic on, for example 12 rounds. It is composed of a linear part, and of a Boolean function with a "very strong" bias (only few products are present). Yet, if the whole is strongly biased, the linear part of it can be seen as a sum of two "strongly" biased expressions, and in turn should still be "strongly" biased. We come to the conclusion that:

Conjecture F.2.1 (Gravity Law for Bi-Linear Cryptanalysis). A Bilinear attack with a quadratic part using only few products **should not** be much better than a linear attack with its linear part alone. Consequently it will not be much better either than the best linear attack.

For example if we assume that there is one input and one output product in a bi-linear characteristic on 12 rounds having a bias of about 2^{-15}, then by Matsui's Piling-up Lemma [28] we expect that the bias of its linear part should not be worse than about 2^{-17}. (We expect that the bias of the linear part will be between $2^{-15-2} = 2^{-13}$ and $2^{-15+2} = 2^{-17}$).

It seems that the best bi-linear attack on ciphers similar to DES can only be slightly better than the best linear attack. **However:**

- Still, a bi-linear attack can be strictly better than the best existing linear attack for the same number of rounds. In Appendix B we show such examples for $3, 7, 11, \ldots$ rounds of DES.
- It is possible to see that with the average bias per round that is achieved by Matsui in his best characteristics, a bi-linear attack with one product can potentially allow a gain of about 2 rounds. Examples we exhibit in this paper are better than Matsui but the gain rather compares to removing one round. This suggests that even better examples may exist.
- A bi-linear attack with two products can potentially allow a gain of up to 4 rounds compared to the best linear attack. The best non-linear approximation we know for 5 rounds, given in Table 2, does indeed use 2 products. However so far we did not find so far an example where such an approximation would be better than the best linear attack.

- For ciphers other than DES, we expect that a bi-linear attack that is much better than LC will use many products. Following Conjecture F.1.1 this could probably happen only for ciphers using large S-boxes (or a poor P-box). It seems that for large S-boxes the designer will be able to prevent BLC by making sure that there is no good bi-linear approximations at all. In Section 5 we show however than even then he could be unlucky and though at first sight the cipher would look stronger than any other practical block cipher, it can contain a subtle but fatal flaw (cf. Section 5).

Simulations on the "Gravity Law" (Conjecture F.2.1): One should understand that this "Gravity Law" (Conjecture F.2.1) is heuristic, and though we observed that it is usually verified for DES, we have also found many counter-examples. In our simulations, the masks at the input and at the output of the cipher are the same. We call $\mathcal{L}$ a linear part of a characteristic and $\mathcal{Q}$ the quadratic part. Then we compare the bias for the two probabilities: for the linear part only $Pr(\mathcal{L})$, and $Pr(\mathcal{L} \oplus \mathcal{Q})$ in which both the quadratic and the linear parts are present. Some results are given in Table 6.

Table 6. Some examples for the "Gravity Law" (Conjecture F.2.1)

DES	linear part	quad. part	*biases*	
rounds	$\mathcal{L}$	$\mathcal{Q}$	$\lvert Pr(\mathcal{L}\oplus\mathcal{Q})-\frac{1}{2}\rvert$	$\lvert Pr(\mathcal{L})-\frac{1}{2}\rvert$
1	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.88 \cdot 2^{-3}$	$< 2^{-10}$
1	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$1.00 \cdot 2^{-6}$	$< 2^{-10}$
2	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.75 \cdot 2^{-4}$	$< 2^{-15}$
2	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$1.0 \cdot 2^{-11}$	$< 2^{-15}$
3	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.84 \cdot 2^{-5}$	$1.56 \cdot 2^{-3}$
3	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$\mathbf{1.65 \cdot 2^{-3}}$	$1.56 \cdot 2^{-3}$
4	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.83 \cdot 2^{-8}$	$< 2^{-17}$
4	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$1.92 \cdot 2^{-8}$	$< 2^{-17}$
5	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.59 \cdot 2^{-9}$	$< 2^{-17}$
5	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$1.4 \cdot 2^{-11}$	$< 2^{-17}$
6	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.3 \cdot 2^{-10}$	$1.3 \cdot 2^{-14}$
6	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$1.2 \cdot 2^{-13}$	$1.3 \cdot 2^{-14}$
7	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.2 \cdot 2^{-11}$	$1.8 \cdot 2^{-13}$
7	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$\mathbf{1.41 \cdot 2^{-9}}$	$1.8 \cdot 2^{-13}$
11	$L[3,8,14,25]\oplus R[17]$	$L[3]R[17]$	$1.7 \cdot 2^{-18}$	$1.91 \cdot 2^{-16}$
11	$L[3,8,14,25]\oplus R[17]$	$L[3]R[16,17,20]$	$\mathbf{1.2 \cdot 2^{-15}}$	$1.91 \cdot 2^{-16}$

We used here two interesting bi-linear characteristics from our first example (Section 6.2) and from our second example (Appendix B). The "Gravity Law" is confirmed for 3 and 11 rounds, and partially for 7 and 6 rounds. The bias of the linear part is sometimes better, sometimes worse than for the full bi-linear characteristics. We distinguish in bold characters biases that are strictly stronger than the best existing linear attack found by Matsui. For 3 and 11 rounds their linear part is precisely the best characteristic of Matsui.

Counter-examples for the "Gravity Law". Nothing should be taken for granted with generalized linear attacks. For the first example for 6 and 11 rounds, and in the second example for 7 rounds, we observe that the gap between the two probabilities is too big to be explained by the bias of the difference (the quadratic part). It gets worse for 2, 4 and 5 rounds: here the "Gravity Law" is not confirmed at all. For example, for 4 rounds the bi-linear bias is $1.83 \cdot 2^{-8}$ and since we have only one product, by an application of Matsui's Piling-up Lemma the bias of the linear part should **not** be worse than $1.83 \cdot 2^{-10}$. Yet we found that it is rather negligible, at most 2^{-17} and probably even much less. We see that **though our heuristic law says that these two bi-linear attacks cannot work for 2, 4 and 5 rounds, clearly they do always work, and for any number of rounds**.

In fact the examples in which the difference in the bias of two correlated expressions is higher than expected are frequent, and it has to be so. A non-linear function is usually correlated to many linear functions and the bias of the function applied to the (plaintext, ciphertext) pairs cannot be at close distance to all of them at the same time (there is usually no correlation between two different linear characteristics and their biases can differ in order of magnitude quite arbitrarily).

G Paradoxes and Pitfalls of Generalised Linear Cryptanalysis

Bi-linear cryptanalysis is harder than it looks. In Section F we explain that it is usually in some way "tied" to LC. Attacks with a small number of products are in most cases strongly correlated to linear attacks and they cannot be much better than the best linear attack. This remark holds for all ciphers, not specifically Feistel ciphers. We will explain now that there is another reason, very specific to the structure of Feistel schemes, why BLC attack cannot work as well as one would sometime expect.

We consider three consecutive rounds an arbitrary Feistel cipher, it can be DES, but the round functions can be also random functions. We choose two arbitrary linear combinations of bits, for example bit 3 and bit 17. We assume that input bit 3 and the output bit 17 are statistically independent (for example they are not connected in the round function or they are connected by a be a very good S-box). We also assume the same independence for the reverse direction: for the input bit 17 and the output bit 3.

For 1 round we have the following characteristic:

$$\left.\begin{array}{l} L_0[3]R_0[17]\oplus \\ L_1[3]R_1[17] = 0 \end{array}\right\} \quad \frac{1}{2} + 1.00 \cdot 2^{-2}$$

For 2 rounds we have the following characteristic:

$$\left.\begin{array}{l} L_0[3]R_0[17]\oplus \\ L_2[3]R_2[17] = 0 \end{array}\right\} \quad \frac{1}{2} + 1.00 \cdot 2^{-3}$$

We observe that this later equation can be written in two different ways as a sum of two products:

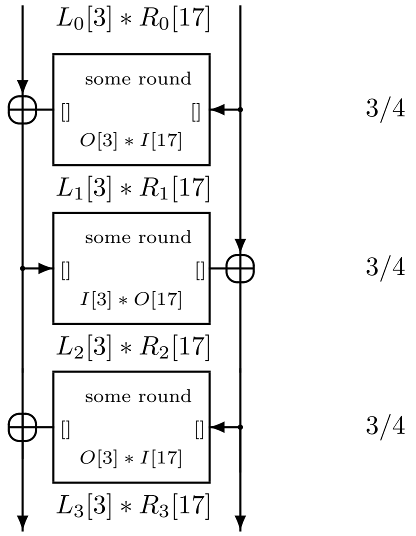

Fig. 12. Generic bi-linear attack that cannot work in general

$$(a) \quad L_0[3]R_0[17] \oplus L_2[3]R_2[17] = I_1[3]O_1[17] \oplus O_2[3]I_2[17]$$

Each of them suggests that it should be equal to 0 with probability $3/8 = 1/2 + 1.00 \cdot 2^{-3}$. For a perfect cipher with random round functions will give a prediction of exactly $3/8 = 1/2 + 1.00 \cdot 2^{-3}$ (though strictly speaking it is not yet a proof that the bias will indeed be such, see below).

Now, if for our cipher the round functions are not perfect the predictions of the bias by the two methods (both heuristic) may differ.

Assume that, due to the weakness of one S-box in the round function both products $I_1[3]O_1[17]$ and $O_2[3]I_2[17]$ are equal to 0 not with probability 3/4 but with, say probability 9/10, from the Matsui's lemma we expect that $L_0[3]R_0[17] \oplus L_2[3]R_2[17] = 0$ with probability 0.82, stronger than 0.75 expected looking at the left side of the equation. Obviously the stronger bias will prevail and our characteristic will indeed be true with probability 0.82 instead of 0.75 for this specific cipher.

Now if the biases of $I_1[3]O_1[17]$ and $O_2[3]I_2[17]$ are in turn weaker than 3/4 something very different will happen. Our characteristic will remain true with probability about 0.75 (as for a random cipher). Though the right side of our equation does suggest a weaker bias, this in fact is not true for other reasons (left side) and we have to infer that the two products on the right side cannot be independent. In general the bias of the characteristics and their correlation will vary depending on the key, and we expect that the right side of our equation still has some but very weak incidence on the actually observed bias for 2 rounds. Yet we have here clearly a highly non-linear behaviour: the bias of an equation will depend on the comparison of two biases.

If one of the values is fixed, as in our example, we will have a type of "threshold behaviour". In general, in a generalised linear attack there may be two or several ways of decomposing an equation as a sum of biased expressions, the bias of each method will depend on the key in a complex way. All this will make the analysis of the actual behaviour of the attack fairly intricate, if not impossible. These phenomena do obviously exist also in LC, however in LC of DES the bias of an equation does never depend on the key, and this is in fact the main reason why linear cryptanalysis of DES behaves well w.r.t. predictions (related work: [12]). For other ciphers, such as ICE, the existence of linear approximations depends on some key bits, and then the threshold behaviours will appear and it will become really harder to analyse the complexity of linear attacks.

3 and more rounds. Having gained some insights, we can modestly try to predict the behaviour of some attack for 3 rounds, maybe not beyond. We go back to our example that is very interesting. We have:

$$(a) \quad L_0[3]R_0[17] \oplus L_3[3]R_3[17] = I_1[3]O_1[17] \oplus O_2[3]I_2[17] \oplus I_3[3]O_3[17]$$

From the left side we expect the bias of the characteristic to be about $3/8 = 1/2 + 1.00 \cdot 2^{-3}$, as for a perfect cipher, from the right hand the Matsui's lemma gives $1/2 + 1.00 \cdot 2^{-4}$, (though since the Boolean functions used here are not balanced, it is possible to see that it should not be applied here, and the real result is probably even less).

The attack cannot work. The threshold behaviour will make that our generic attack on Feistel ciphers cannot work. We can even mathematically prove that it cannot work for reasonable good ciphers. For example, if the probability distribution of round functions for a random key of our cipher is indistinguishable from a randomly chosen function, then by the Luby-Rackoff theorem [27] we have that:

$$\left.\begin{array}{l} L_0[3]R_0[17]\oplus \\ L_3[3]R_3[17] = 0 \end{array}\right\} \quad \frac{1}{2} + 1.00 \cdot 2^{-3}$$

and this exactly, i.e. we are not able to distinguish if the probability is different than $\frac{1}{2}+1.00\cdot 2^{-3}$ with non-negligible advantage, because otherwise we would distinguish this cipher from a random permutation with a small (polynomial) number of queries.

The attack can work after all. Nevertheless, even this very naive and badly designed attack will work for some ciphers. It is sufficient for this that due to a weakness of S-boxes the biases for the two products $I[3]O[17]$ and $O[3]I[17]$ are very strong for some fraction of keys, and the resulting bias of $I_1[3]O_1[17] \oplus O_2[3]I_2[17] \oplus I_3[3]O_3[17]$ will occasionally be lower than our threshold of 3/8. Then we get an attack that will work for a fraction of keys.

Relation to the design criteria for DES S-boxes The known design criteria for DES S-boxes already prevent such attacks. The equations such as $I[3]O[17]$ and $O[3]I[17]$ cannot be strongly biased, because this would imply that one input is strongly correlated with one output. Nevertheless in DES we have for example:

(X) $(J[17]+1)(O[8,14,25,3]+1)=0$ for $S5$ with probability $58/64$

which can be used in a very similar way to mount attack (if we had a matching equation for the next round and if the biases were strong enough). Similarly, and this is even more tricky, we could conduct a very fast attack on a Feistel cipher in which, for example $O[\alpha] \oplus I[17]O[3] = 0$ is very highly biased for some keys and for all the odd rounds, and $O[\beta] \oplus O[17]I[3] = 0$ is very highly biased for some keys and for all the even rounds, for some well chosen sets of outputs α and β. This attack is not prevented by the historically known design criteria of DES, but it is arguably (but not provably) prevented by the (still relatively good) resistance of DES to linear cryptanalysis. Indeed, if the equation $O[\alpha] \oplus I[17]O[3] = 0$ is very highly biased, it is also presumably so for $O[\alpha] \oplus I[17] = 0$ and a few other equations. Thus basically only ciphers with S-boxes having very bad S-boxes for linear cryptanalysis could be broken by this attack. (Yet it is possible to construct very special S-boxes such that $O[\alpha] \oplus I[17]O[3] = 0$ is very highly biased and $O[\alpha] \oplus I[17] = 0$ just isn't.)

Summary: There are many non-linearities and threshold behaviours in predicting the bias of composed characteristic in a real cipher. Some biases are due to the structure of the Feistel scheme itself, and alone cannot allow to distinguish the cipher from a random permutation. Sufficiently strong biases will always combine well and allow working attacks (no doubt about this), other biases will prevail otherwise.

This may explain why in Appendix E we found very few examples that work as predicted for an important number of rounds: these examples use one I/O product for every round of the Feistel scheme and for reasons very similar than in (a) above, except the equations are also combined with linear parts, have multiple threshold behaviours in which we are in fact systematically below the threshold.

This leads to think that for ciphers such as DES with small S-boxes, interesting constructions of bi-linear attacks for DES with a small number of products such as found in this paper, are in a very limited number. They do depend on the existence of some bi-linear characteristics stronger than some threshold that could (in theory) be computed exactly. Our research shows that DES has “reasonably good” S-boxes for which such constructions exist but do not significantly lower the security of DES w.r.t. previously known attacks. We came to the same conclusion that in Appendix F.

H Beyond BLC - Multi-Linear Cryptanalysis

The Bi-Linear Cryptanalysis described in this paper applies not only to regular Feistel ciphers with two identical halves, and unbalanced Feistel schemes with 2 halves of different size, but also to generalised Feistel schemes with more than two branches, as used in SHA_x and related block ciphers, or Skipjack. In this latter case, a much more general attack, called Multi-Linear Cryptanalysis (MLC) is possible, as briefly outlined in [7]. However, we believe that in practice the gain achieved with MLC compared to LC should be rather disappointing. We expect that, as in this paper, though it is possible to construct contrived ciphers that are very weak for Multi-Linear Cryptanalysis (some are given in [7], more in the extended version of this paper available form the authors) for most real-life ciphers our heuristic law should hold: this type of attack will not be much better than the best linear attack.