Linear Cryptanalysis of Reduced-Round Versions of the SAFER Block Cipher Family

Jorge Nakahara Jr, Bart Preneel*, Joos Vandewalle

Katholieke Universiteit Leuven, Dept. Electrical Engineering–ESAT
Kardinaal Mercierlaan 94, B–3001 Heverlee, Belgium
{Jorge.Nakahara,Bart.Preneel,Joos.Vandewalle}@esat.kuleuven.ac.be

Abstract. This paper presents a linear cryptanalytic attack against reduced round variants of the SAFER family of block ciphers. Compared with the 1.5 round linear relations by Harpes *et al.*, the following new linear relations were found: a 3.75-round non-homomorphic linear relation for both SAFER-K and SAFER-SK with bias $\epsilon = 2^{-29}$; a 2.75 round relation for SAFER+ with bias $\epsilon = 2^{-49}$. For a 32-bit block mini-version of SAFER a 4.75-round relation with bias $\epsilon = 2^{-16}$ has been identified. These linear relations apply only to certain weak key classes. The results show that by considering non-homomorphic linear relations, more rounds of the SAFER block cipher family can be attacked. The new attacks pose no threat to any member of the SAFER family.

1 Introduction

SAFER (Secure And Fast Encryption Routine) is a family of block ciphers, designed by Massey, which comprises 64-bit block ciphers like SAFER-K64 [11], SAFER-K128 [12], SAFER-SK40, SAFER-SK64 and SAFER-SK128 [13]. The number that follows each cipher name indicates the key size. The newest member of this family is the AES candidate SAFER+ [10] designed jointly with Khachatrian and Kuregian; SAFER+ has a 128-bit block size and variable key size versions of 128, 192 and 256 bits. We will also analyze a 32-bit block mini-version, called SAFER-K32.

The more widespread, easy-to-deploy and better-understood an encryption algorithm is, the more attractive it becomes as a target for cryptanalysts. All SAFER family members have publicly available descriptions, are unpatented, royalty-free, with plenty of flexibility for different key sizes and block sizes, and are designed to be efficiently implementable in software [13]. These are key features to make SAFER+ widely deployed. An example is the inclusion of SAFER+ for authentication purposes in Bluetooth [1, p. 149]; this is the codename for a technology specification for low-cost, short range radio links between mobile PC's, mobile phones and other portable devices.

* F.W.O. research associate, sponsored by the Fund for Scientific Research, Flanders (Belgium).

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

Several theoretical attacks have been published on the ciphers of the SAFER family (in most cases versions were considered with a reduced number of rounds): differential cryptanalysis by Massey [12], truncated differentials by Knudsen and Berson [17], later improved by Wu *et al.* [7], an algebraic attack by Murphy [20], key schedule attacks by Knudsen [15] and by Kelsey *et al.* [9], and observations on the PHT design by Vaudenay [21] and Brincat *et al.* [2]. Linear cryptanalysis has been considered by Harpes *et al.* in [4] (see also [3]); they show that a generalized linear attack becomes infeasible for three or more rounds of SAFER-K64. This paper proposes an improved linear analysis by considering a wider class of linear relations; it also identifies certain classes of keys that are 'weak' w.r.t. linear cryptanalysis.

This paper is organized as follows. Section 2 describes the structure of SAFER-K64 and its key-schedule algorithm. Section 3 describes a 32-bit-block mini-version of SAFER-K. Section 4 introduces principles of linear cryptanalysis and some terminology for our attack. Section 5 gives particular features of a new type of linear relation for SAFER ciphers. Section 6 contains our results for the SAFER cipher family; their further use in an attack is described in Sect. 7. Section 8 discusses the methodology used to obtain the new linear relations and Sect. 9 summarizes the analysis results. Annex A presents a ciphertext-only attack.

2 Description of SAFER-K64

SAFER-K64 is a 64-bit-block iterated cipher with $r = 6$ rounds and a 64-bit user-selected key K. The key K is expanded into $2r + 1$ subkeys, that is, two subkeys per round plus one subkey for an output transformation. The following description of the round structure of SAFER-K64 also applies to SAFER-SK40, SAFER-SK64 and SAFER-SK128, because their ciphers only differ in the key schedule. Therefore, SAFER-K/-SK will be used as a notation when the analysis applies to both ciphers.

2.1 The Round Structure

In each encryption round, the input block B is first split into 8 bytes: $B = (b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8)$, $b_i \in \mathbb{Z}_{256}, 1 \leq i \leq 8$. Each byte b_j is combined with the first round-subkey K_{2i}: $Y = B + K_{2i} = (b_1 \oplus K_{2i}^1, b_2 \boxplus K_{2i}^2, b_3 \boxplus K_{2i}^3, b_4 \oplus K_{2i}^4, b_5 \oplus K_{2i}^5, b_6 \boxplus K_{2i}^6, b_7 \boxplus K_{2i}^7, b_8 \oplus K_{2i}^8)$ where $\oplus$ denotes bitwise XOR and $\boxplus$ represents ADD(ITION) modulo 256. Each byte of $Y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8)$ is input to an S-box: $Z = (\mathrm{X}(y_1), \mathrm{L}(y_2), \mathrm{L}(y_3), \mathrm{X}(y_4), \mathrm{X}(y_5), \mathrm{L}(y_6), \mathrm{L}(y_7), \mathrm{X}(y_8))$, where X(.) is an eXponentiation S-box and L(.) a Logarithm S-box, described later. This S-box layer will be referred to as the non-linear or NL layer. Subsequently, $Z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8)$ is combined with the second round-subkey K_{2i+1}: $T = Z + K_{2i+1} = (z_1 \boxplus K_{2i+1}^1, z_2 \oplus K_{2i+1}^2, z_3 \oplus K_{2i+1}^3, z_4 \boxplus K_{2i+1}^4, z_5 \boxplus K_{2i+1}^5, z_6 \oplus K_{2i+1}^6, z_7 \oplus K_{2i+1}^7, z_8 \boxplus K_{2i+1}^8)$. Finally, the bytes of T

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

are input to a linear transformation called Pseudo-Hadamard Transform or PHT layer.

The alternating XOR/ADD layer of input data with the first subkey bytes, together with the NL layer will be referred to as the NL half-round; similarly, the alternating ADD/XOR layer of intermediate data with the second subkey, together with the PHT layer will be called the PHT half-round.

There are two S-boxes: an eXponentiation $X(a) = (45^a \bmod 257) \bmod 256$ (X-box, for short), and a Logarithm $L(a) = \log_{45}(a) \bmod 257$ (or L-box, for short) for $a \neq 0$, with the special case $L(0) = 128$. They are each other's inverses, that is, $X(L(a)) = L(X(a)) = a, \forall a \in \mathbb{Z}_{256}$.

The PHT layer denotes a network of twelve 2-PHT boxes, where the latter is defined as $\text{2-PHT}(a, b) = (2 \cdot a \boxplus b, a \boxplus b)$, for $a, b \in \mathbb{Z}_{256}$. Denoting the input to a PHT layer by $Y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8)$ and its output by $Z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8)$, where $y_i, z_i \in \mathbb{Z}_{256}$, $1 \leq i \leq 8$, this transformation can be described by $Z = Y^T \cdot M$, where M is called the PHT matrix:

$$M = \begin{pmatrix} 8 & 4 & 4 & 2 & 4 & 2 & 2 & 1 \\ 4 & 2 & 4 & 2 & 2 & 1 & 2 & 1 \\ 4 & 2 & 2 & 1 & 4 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 \\ 4 & 4 & 2 & 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} .$$

Let $T = (t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8)$ be the output after r rounds. There is an output transformation which mixes T with the last subkey, giving the ciphertext: $C = T + K_{2r+1} = (t_1 \oplus K^1_{2r+1}, t_2 \boxplus K^2_{2r+1}, t_3 \boxplus K^3_{2r+1}, t_4 \oplus K^4_{2r+1}, t_5 \oplus K^5_{2r+1}, t_6 \boxplus K^6_{2r+1}, t_7 \boxplus K^7_{2r+1}, t_8 \oplus K^8_{2r+1})$. Decryption involves the application of the inverse of each round with reverse order for the subkeys. More details can be found in [11, 12].

The round structure of SAFER+ uses the same S-boxes and 2-PHT primitives found in 64-bit block members, but the former uses a different PHT layer composed of four 2-PHT layers, and a particular fixed permutation between 2-PHT layers, called Armenian Shuffle (see Fig. 1).

2.2 The Key Schedule

The key schedule of SAFER-K64 accepts a 64-bit user-selected key K and generates 64-bit subkeys $K_i, 1 \leq i \leq 2r + 1$, that is, two subkeys per round plus one subkey for the output transformation. K itself is used (unchanged) as the first subkey K_1. Subsequently, K is split into eight bytes, $(K^1, K^2, K^3, K^4, K^5, K^6, K^7, K^8)$, and each byte is left rotated by three bits. Next, fixed byte values called key bias $B^1_2, \ldots, B^8_2$ are added to bytes $K^1, \ldots, K^8$ respectively, where

$$B^j_i = (45^{45^{9i+j} \bmod 257} \bmod 257) \bmod 256 \ , \quad 2 \leq i \leq 2r + 1 \ , \quad 1 \leq j \leq 8 \ .$$

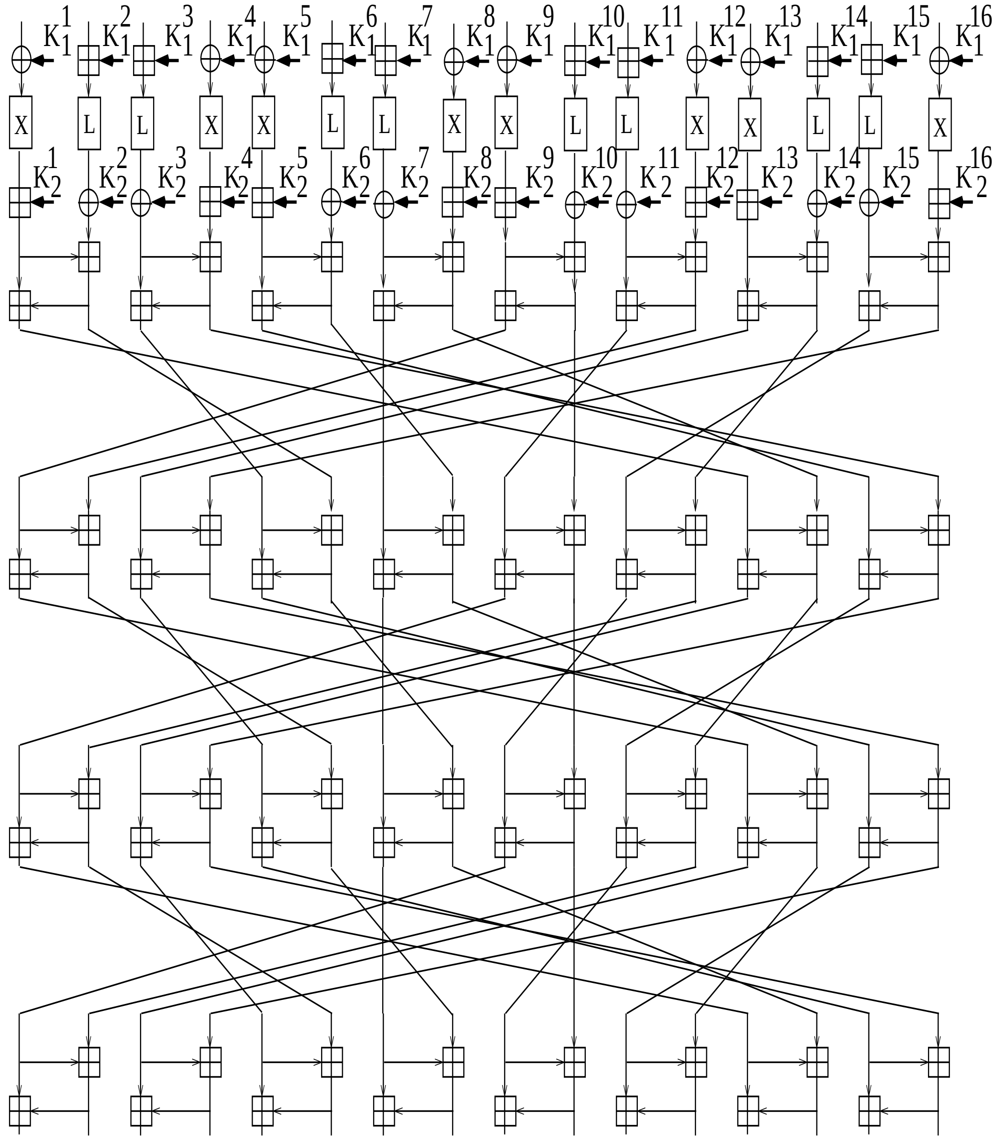

Fig. 1. One round of SAFER+

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer
Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

The result is the second subkey $K_2 = (ROL_3(K^1) \boxplus B_2^1, \ldots, ROL_3(K^8) \boxplus B_2^8)$. The other subkeys are generated by following the same steps using the previous subkey as input: rotate each input byte left by 3 bits and add the next key bias.

Key schedule weaknesses in SAFER-K64 were demonstrated by Knudsen [15], resulting in the improved key schedule of SAFER-SK64 [12]. Kelsey *et al.* have pointed out a weakness in the key schedule of SAFER+ for long keys [9]. Key schedule weaknesses will not be considered in this paper, but our analysis will point out that some keys are weak w.r.t. linear cryptanalysis.

3 A Mini-Version of SAFER-K64

Some block ciphers allows all of their individual components to be reduced to a half, a quarter or even smaller sizes, while the security level relative to the block size remains similar. This is also the case for the SAFER cipher family. This paper analyzes one such reduced version which will be called SAFER-K32. This is a 32-bit block cipher with a 32-bit user key, $r = 8$ rounds, and with S-boxes defined as $X(a) = (g^a \bmod 17) \bmod 16$, and $L(a) = \log_g a \mod 17$, for $a \neq 0$ and $L(0) = 8$. There are eight degrees of freedom in choosing g such that $GF(17) =< g >$, namely $g \in \{3, 5, 6, 7, 10, 11, 13, 14\}$ (see [17]). The value $g = 11$ was chosen arbitrarily for this mini-version.

The emphasis of the current analysis is not to attack the key schedule but the cipher itself; therefore it will be assumed that the key schedule for SAFER-K32 has a structure similar to that of SAFER-K64. The scale is reduced: the key schedule generates $(2r + 1)$ 32-bit subkeys and it uses the same generator as the cipher.

The main reasons to consider reduced versions of ciphers are:

- the reduced dimensions allow a more comprehensive (exhaustive) analysis, to be carried out which is not always possible in the original cipher;
- it is hoped that weaknesses found in the mini-version can be extended to the larger cipher, or at least that they may provide some insight in potential weaknesses in the original cipher.

4 Linear Cryptanalysis of SAFER

4.1 Linear Cryptanalysis

Linear cryptanalysis is a statistical, known-plaintext attack introduced by Matsui and Yamagishi in 1992 in an attack against FEAL [19]. It was extended to DES in 1993 [18]. The attack explores (approximate) linear relations between plaintext, ciphertext and subkey bits. Linear approximations for an iterated cipher are usually made by combining approximations for each round.

If $X_i = (x_n, x_{n-1}, \ldots, x_2, x_1)$ is an n-bit input to a round, $R(X_i)$ is its output, and K_i the round subkey, then a linear relation can be expressed as

$$X_i \cdot \mathit{\Gamma I} \oplus R(X_i) \cdot \mathit{\Gamma O} = K_i \cdot \mathit{\Gamma K}_i \ , \tag{1}$$

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

where $\Gamma I, \Gamma O$ and ΓK_i are n-bit masks which specify the bits of X_i, $R(X_i)$ and K_i involved in the linear relation. For example, $X_i \cdot \Gamma I = X \cdot 45_x = x_1 \oplus x_3 \oplus x_7$ (the subscript '`x`' indicates hexadecimal values).

The left-hand side of equation (1) provides an estimate for the xor of the subkey bits on the right-hand side. Without loss of generality, the following simplified equation is employed

$$X_i \cdot \Gamma I \oplus R(X_i) \cdot \Gamma O = 0 \ . \tag{2}$$

Two numerical values can be associated with (2). First, a probability $p = \Pr(X_i{\cdot}\Gamma I = R(X_i){\cdot}\Gamma O)/2^n$ that expresses the frequency with which equation (2) holds (relation (2) is also called a linear approximation). Second, the deviation of parity of (2) from a random relation, or $p' = p - \frac{1}{2}$. It is clear that $-\frac{1}{2} \leq p' \leq \frac{1}{2}$ and the approximation is useful only if $p' \neq 0$. The absolute value $\epsilon = |p'|$ is called bias [8]. The larger the bias the more useful the linear relation is, that is, the more unbalanced the parity of (2) from a random distribution the less plaintext is needed to estimate the value of $K_i \cdot \Gamma K_i$ (with high degree of assurance). The number N of known plaintexts required for an attack using a linear relation with bias ϵ equals $N = c \cdot \epsilon^{-2}$, where c is a small constant, which depends on the algorithm used for the estimation [8, 18]. In case $p' < 0$, the value obtained for $K_i \cdot \Gamma K_i$ is actually $\overline{K_i \cdot \Gamma K_i} = (K_i \cdot \Gamma K_i) \oplus 1$.

The following notation will be used to represent a binary-valued linear relation for one round of an iterated (n-bit block) cipher:

$$\Gamma = (\Gamma I, \Gamma O, \epsilon) \ . \tag{3}$$

One-round linear relations can be concatenated or stacked in order to approximate more rounds. If $\Gamma_1 = (\Gamma X_1, \Gamma Y_1, \epsilon_1)$, $\Gamma_2 = (\Gamma X_2, \Gamma Y_2, \epsilon_2)$ are r_1-round and r_2-round independent linear relations, respectively and $\Gamma Y_1 = \Gamma X_2$, then it is possible to combine them to form an $(r_1 + r_2)$-round linear relation $\Gamma_3 = (\Gamma X_1, \Gamma Y_2, \epsilon)$ with bias $\epsilon = 2 \cdot \epsilon_1 \cdot \epsilon_2$ (Matsui's Piling-Up lemma [18]). Note however that this assumes that the subkeys are mutually independent and uniformly distributed which is not the case for any member of the SAFER cipher family, when the key schedule algorithms are used. Nonetheless, practical experiments show that the subkeys generated through the respective key schedules of each cipher are adequately randomized in order for the approximations to hold.

As an example of linked relation, a one-round linear relation for SAFER-K64 can be viewed as the concatenation of two half-round linear relations: $\Gamma_{NL} = (\Gamma X, \Gamma M, \epsilon_1)$, and $\Gamma_{PHT} = (\Gamma M, \Gamma Y, \epsilon_2)$, where ΓM denotes a bit-mask applied to the intermediate value in the middle of a round, between the output of the NL and the input to the PHT layers.

4.2 Homomorphic Linear Relations

Definition 1. *Let G_1 and G_2 be groups with operations $\otimes$ and $\boxdot$, respectively. A mapping M from G_1 into G_2 is called a* homomorphism *if*

$$M(y \otimes z) = M(y) \boxdot M(z) \ , \ \forall y, z \in G_1 \ . \tag{4}$$

Definition 2. *A binary-valued function f is* balanced *if it outputs the value 0 for exactly half of its inputs.*

Definition 3 (Harpes-Kramer-Massey [3, 4]). *An* I/O sum $S^{(i)}$ *for a round is a modulo-two sum of a balanced binary-valued function f_i of the round input $Y^{(i-1)}$ and a balanced binary-valued function g_i of the round output $Y^{(i)}$, namely*

$$S^{(i)} = f_i(Y^{(i-1)}) \oplus g_i(Y^{(i)}) \ . \tag{5}$$

The functions f_i and g_i are called input function and output function, respectively, of the I/O sum $S^{(i)}$. I/O sums for successive rounds will be called linked if the output function of each I/O sum except the last coincides with the input function of the following I/O sum: $g_i = f_{i+1}$. When $S^{(1)}, S^{(2)}, \ldots, S^{(r)}$ are linked, then their sum is also an I/O sum:

$$S^{(1\ldots r)} = \bigoplus_{i=1}^{r} S^{(i)} = f_0(Y^{(0)}) \oplus g_r(Y^{(r)}) \ . \tag{6}$$

which will be called an r-round I/O sum.

Harpes *et al.* report in [3, 4] that SAFER-K64 is immune to a generalization of linear cryptanalysis [18] which involves only homomorphic I/O sums after 1.5 rounds. Namely, the best homomorphic I/O sum is stated as the concatenation of the following NL-PHT-NL half-rounds:

$$\begin{array}{llll} (\texttt{000000zz000000zz}_{\texttt{x}} & , \texttt{0000000100000001}_{\texttt{x}} & , 2 \cdot (\frac{28}{256})^2) & \text{(NL half-round)} \qquad (7) \\ (\texttt{0000000100000001}_{\texttt{x}} & , \texttt{0001000000000000}_{\texttt{x}} & , \quad 2^{-1}) & \text{(PHT half-round)} \\ (\texttt{0001000000000000}_{\texttt{x}} & , \texttt{00zz000000000000}_{\texttt{x}} & , \quad \frac{28}{256}) & \text{(NL half-round)} \ , \end{array}$$

where $\texttt{zz} \in \{\texttt{cd}_{\texttt{x}}, \texttt{ff}_{\texttt{x}}\}$, and the overall bias (using Matsui's Piling Up Lemma) is $\epsilon = 2^2 \cdot (\frac{28}{256})^3 \approx 2^{-7.58}$. Using homomorphic linear relations, the bit-masks which are used to approximate one round of SAFER-K64 take into account the group operations used to mix subkey bits in each round. In this way the effect of carry bits is avoided when the group operation is addition modulo 256.

The homomorphicity of the masking function is important for the approximation of subkey bits mixed in a round, or more specifically, for the group operations used to mix subkey bits in a round. For example, let $G_1 = (\mathbb{Z}_{256}, \boxplus)$ and $G_2 = (\mathbb{Z}_{256}, \oplus)$ be groups. The only homomorphic masking function in this setting is $M_1(y) = \Gamma \cdot y = \texttt{01}_{\texttt{x}} \cdot y$, that is, the mask which takes only the least significant bit (LSB). In [3] it is stated that (7) is the best homomorphic linear relation achievable for SAFER-K/-SK. In the next sections, it will be shown that approximations using non-homomorphic bit-masks result in improved linear relations.

4.3 Non-Homomorphic Linear Relations

Let $M_2(y) = \texttt{02}_{\texttt{x}} \cdot y$ be a masking function and K be a subkey byte. The M_2 mapping is non-homomorphic, because

$$M_2(y \boxplus K) = \texttt{02}_{\texttt{x}} \cdot (y \boxplus K) \neq M_2(y) \oplus M_2(K) = \texttt{02}_{\texttt{x}} \cdot y \oplus \texttt{02}_{\texttt{x}} \cdot K \ .$$

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

This happens because of a possible carry bit that can propagate from the LSB to the second LSB. Assuming that the intermediate data values in a round and the subkey bits are uniformly distributed, this carry bit only exists with probability 1/4. Therefore, a bias penalty of 2^{-2} is to be accounted for.

If one considers the bit-masks for the subkey bytes applied only to a fraction of a round, like a subkey-mixing layer, then one can split a one round approximation into quarters of a round. Therefore, the bit-masks that make up the approximation only of the mixing layer of subkey bits in a round, or only of the NL layer, or only of the PHT layer will be called quarter-round approximations. As an example, $(\mathtt{0002020100000201_x}, \mathtt{0001010200000102_x}, 2^{-16})$ is an NL quarter-round approximation of the S-box layer in a round of SAFER-K64. Such partial approximations will be used later for (fractional) linear attacks.

5 Key-Dependent Linear Relations

The non-homomorphicity of some bit-masks has two important consequences for the corresponding linear relations:

(1) Let $X, S, K_i \in \mathbb{Z}_{256}$ be the input, the output and the subkey bytes in a subkey addition operation in a round of SAFER-K/-SK, that is, $S = X \boxplus K_i$. An approximation for the addition operation, using bit-masks $\mathtt{02_x}$, take the form $S \cdot \mathtt{02_x} = X \cdot \mathtt{02_x} \boxplus K_i \cdot \mathtt{02_x}$. This approximation is non-homomorphic. Besides, it assumes that there is no carry bit into the second LSB position, otherwise the approximation would be void. This carry bit can be avoided if the least significant bit of subkey K_i is zero. Other non-homomorphic masks present similar dependencies on the subkey bits.

(2) If the carry bit restrictions are satisfied for the non-homomorphic bit-masks as in the item (1) above, then the masks become homomorphic. In this way, the overall bias of the corresponding linear relation is increased, as there is no more bias penalty to account for. Therefore, one cannot only control the carry propagation (assuring the approximations for the addition of subkey bytes) but also improve the overall bias. The restricted validity of the approximation to subkeys which possess a certain bit-pattern is only apparent. By specifying bit-masks for each key class according to the different approximations of the two LSBs in the addition, all keys in the key space can be attacked. For example, let $X = (x_{n-1}, \ldots, x_1, x_0)$ and K be the input and $S = (s_{n-1}, \ldots, s_1, s_0) = X \boxplus K$, and $\mathit{\Gamma} K$ be the key bit-mask. Then, for the bit-masks which explore only the two LSBs of addition of subkey bytes, there are the following possibilities:
   (a) Let the approximation be $S \cdot \mathtt{02_x} = X \cdot \mathtt{02_x} \oplus K \cdot \mathit{\Gamma} K$ or $s_1 = x_1 \oplus K \cdot \mathit{\Gamma} K$. As the expression of addition is $s_1 = x_1 \oplus k_1 \oplus x_0 \cdot k_0$, it follows that $k_0 = 0$ and $\mathit{\Gamma} K = \mathtt{02_x}$.
   (b) Let the approximation be $S \cdot \mathtt{02_x} = X \cdot \mathtt{03_x} \oplus K \cdot \mathit{\Gamma} K$ or $s_1 = x_1 \oplus x_0 \oplus K \cdot \mathit{\Gamma} K$. As the expression of addition is $s_1 = x_1 \oplus k_1 \oplus x_0 \cdot k_0$, it follows that $k_0 = 1$ and $\mathit{\Gamma} K = \mathtt{02_x}$.

(c) Let the approximation be $S \cdot \texttt{03}_{\texttt{x}} = X \cdot \texttt{02}_{\texttt{x}} \oplus K \cdot \Gamma K$ or $s_1 \oplus s_0 = x_1 \oplus K \cdot \Gamma K$. As the expression of addition is $s_1 \oplus s_0 = x_1 \oplus k_1 \oplus x_0 \cdot k_0 \oplus x_0 \oplus k_0$, it follows that $k_0 = 1$ and $\Gamma K = \texttt{03}_{\texttt{x}}$.

(d) Let the approximation be $S \cdot \texttt{03}_{\texttt{x}} = X \cdot \texttt{03}_{\texttt{x}} \oplus K \cdot \Gamma K$ or $s_1 \oplus s_0 = x_1 \oplus x_0 \oplus K \cdot \Gamma K$. As the expression of addition is $s_1 \oplus s_0 = x_1 \oplus k_1 \oplus x_0 \cdot k_0 \oplus x_0 \oplus k_0$, it follows that $k_0 = 0$ and $\Gamma K = \texttt{03}_{\texttt{x}}$.

Therefore, each bit-mask imposes a different restriction on the key bit pattern but also includes all possibilities for the LSB of the key. For the bit-masks in (a) and (d) the key bit k_0 might be 0, and for the bit-masks in (b) and (c), k_0 is required to be 1. Although the bit masks in the linear relations in the next section are valid for certain specific key bit patterns, they can easily be changed to cover each different key class in the key space.

6 Search Results

We now present the results of our search for non-homomorphic linear relations for the different members of the SAFER family.

Definition 4. *In a linear approximation, an S-box is said to be* active *if the approximation applies a non-zero output bit-mask to that S-box. The number of active S-boxes in a linear relation will be denoted with $\mathcal{S}$.*

A search for linear relations of SAFER-K/-SK resulted in a 3.75-round linear relation with $\mathcal{S} = 7$ and theoretical bias $\epsilon_1 = 2^{-39}$:

$$\begin{array}{lll}
(\texttt{0102010201020102}_{\texttt{x}} & ,\texttt{0000000000020000}_{\texttt{x}} & ,\ 2^{-5})\ \text{(PHT half-round)} \\
(\texttt{0000000000020000}_{\texttt{x}} & ,\texttt{0100000001000000}_{\texttt{x}} & ,\ 2^{-5})\ \text{(one round)} \\
(\texttt{0100000001000000}_{\texttt{x}} & ,\texttt{0002000200020003}_{\texttt{x}} & ,\ 2^{-11})\ \text{(one round)} \\
(\texttt{0002000200020003}_{\texttt{x}} & ,\texttt{0002000100000000}_{\texttt{x}} & ,\ 2^{-18})\ \text{(one round)} \\
(\texttt{0002000100000000}_{\texttt{x}} & ,\texttt{0002000100000000}_{\texttt{x}} & ,\ 2^{-2})\ \text{(subkey quarter-round)}\ .
\end{array} \qquad (8)$$

Recalling the key dependency discussed in Sect. 5, item (1), the following restrictions on subkey bits are necessary for the approximation of subkey addition in a 1.25R attack on five rounds SAFER-K/-SK to hold:

$$\mathrm{LSB}(K_2^4, K_2^8, K_3^6, K_6^1, K_6^5, K_7^2, K_7^6, K_8^4, K_8^8, K_9^2) = 0 \qquad (9)$$

where the notation $\mathrm{LSB}(\cdot, \ldots, \cdot) = 0$ means that the least significant bit of each argument is zero. Therefore, the actual bias of relation (8) is $\epsilon_1^* = 2^{-29}$. These keys are called weak keys w.r.t. relation (8). Incidentally, these ten key bits in (9) map to exactly ten different user key bits, according to the key schedule of SAFER-K64 [11] which means that one in 1024 user keys is weak. For the key schedule of SAFER-SK64 (see [13]), these ten key bits imply conditions on 16 different user key bits.

Recalling the discussion in Sect. 5, item (2), the bit-masks in (8) can be adapted accordingly to satisfy the other 1023 subkey classes. For example,

$$\begin{array}{llll}
(\mathtt{0102010301020102_x} & ,\mathtt{0000000000020000_x} & ,\ 2^{-5}) & \text{(PHT half-round)} \\
(\mathtt{0000000000020000_x} & ,\mathtt{0100000001000000_x} & ,\ 2^{-5}) & \text{(one round)} \\
(\mathtt{0100000001000000_x} & ,\mathtt{0002000200020003_x} & ,\ 2^{-11}) & \text{(one round)} \\
(\mathtt{0002000200020003_x} & ,\mathtt{0002000100000000_x} & ,\ 2^{-18}) & \text{(one round)} \\
(\mathtt{0002000100000000_x} & ,\mathtt{0002000100000000_x} & ,\ 2^{-2}) & \text{(subkey quarter-round)}
\end{array} \tag{10}$$

has the same theoretical bias as (8), but the weak key restrictions are:

$$\mathrm{LSB}(K_2^4, K_2^8, K_3^6, K_6^1, K_6^5, K_7^2, K_7^6, K_8^4, K_8^8, K_9^2) = 0 \quad , \tag{11}$$

which imply a different weak key class. The actual bias though, is the same as before, $\epsilon = 2^{-29}$. Similarly, changing each bit-mask in the addition of subkey bytes, in each round, one can get relations which hold for each key in the key space. The same observation holds for the linear relations of SAFER+ and SAFER-K32 below.

For SAFER+ the following linear relation with $\mathcal{S} = 12$ and 2.75 rounds was found:

$$\begin{array}{rlll}
(\mathtt{00020102010000020002000202000100_x} & ,\alpha & ,\ 2^{-11}) & \text{(PHT half-round)} \\
(\alpha & ,\beta & ,\ 2^{-29}) & \text{(one round)} \\
(\beta & ,\gamma & ,\ 2^{-27}) & \text{(one round)} \\
(\gamma & ,\gamma & ,\ \ 2^{-5} & \text{(subkey quarter round)}
\end{array} \tag{12}$$

with theoretical bias $\epsilon_3 = 2^{-69}$, $\alpha = \mathtt{00000200000202000002020000020200_x}$, $\beta = \mathtt{00000001010000000100000001000001_x}$, $\gamma = \mathtt{02000002000203010203020102000100_x}$. The following key bit conditions are necessary for the approximation of subkey addition in a 1.25R attack on four rounds of SAFER+ to hold:

$$\begin{array}{r}
\mathrm{LSB}(K_2^4, K_2^8, K_2^{12}, K_2^{13}, K_3^3, K_3^6, K_3^7, K_3^{10}, K_3^{11}, K_3^{14}) = 0\ , \\
\mathrm{LSB}(K_3^{15}, K_6^4, K_6^5, K_6^9, K_6^{13}, K_6^{16}, K_7^6, K_7^7, K_7^{10}, K_7^{11}) = 0\ .
\end{array} \tag{13}$$

Therefore, the actual bias of (12) is $\epsilon_3^* = 2^{-49}$.

Linear cryptanalysis of SAFER-K32 resulted in a 4.75 round linear relation with $\mathcal{S} = 9$:

$$\begin{array}{llll}
(\mathtt{12121212_x} & ,\mathtt{00000200_x} & ,\ 2^{-5}) & \text{(PHT half-round)} \\
(\mathtt{00000200_x} & ,\mathtt{10001000_x} & ,\ 2^{-3}) & \text{(one round)} \\
(\mathtt{10001000_x} & ,\mathtt{02020203_x} & ,\ 2^{-7}) & \text{(one round)} \\
(\mathtt{02020203_x} & ,\mathtt{02010000_x} & ,\ 2^{-8}) & \text{(one round)} \\
(\mathtt{02010000_x} & ,\mathtt{32110000_x} & ,\ 2^{-6}) & \text{(one round)} \\
(\mathtt{32110000_x} & ,\mathtt{32110000_x} & ,\ 2^{-2}) & \text{(subkey quarter-round)}
\end{array} \tag{14}$$

with theoretical bias $\epsilon_4 = 2^{-28}$. The following restrictions are needed for the approximation of subkey addition in a 1.25R attack on six rounds of SAFER-K32 to hold:

$$\mathrm{LSB}(K_2^4, K_2^8, K_3^6, K_6^1, K_6^5, K_7^2, K_7^6, K_8^4, K_8^8, K_9^2, K_{10}^4, K_{11}^2) = 0 \tag{15}$$

and the actual bias of (14) is $\epsilon_4^* = 2^{-16}$.

7 Fractional Linear Attacks

In Sect. 4.3 linear approximations were described that covered only part of a SAFER-K/-SK round, for example, a half- or a quarter-round. Linear attacks using such fractional linear relations include fractions of a round only at the beginning and end of the cipher, and will be denoted fractional attacks. In the following the subscript $_\mathtt{x}$ will sometimes be omitted from bit masks to simplify notation, for example, 02 instead of $\mathtt{02}_\mathtt{x}$; the interpretation should be clear from the context.

As an example of fractional attack, relation (8) can be used in a 1.25R attack. This is an analogy with the usual 1R or 2R attacks which only discard full rounds [6, 18]. This 1.25R attack does not include the first half of the first round neither the last three quarters of the last round in the approximation, that is, the linear relation (8) does not cover 1.25 rounds (see Fig. 2). The idea for our attacks is to place the linear relations between two subkey layers, such that subkeys at both ends of the cipher are identified. A similar description applies for other fractional values.

This attack covers five rounds of SAFER-K/-SK, without the output transformation, and identifies 81 subkey bits as follows, assuming the weak-key conditions (9) are satisfied:

- let $P_1, \ldots, P_8$ be plaintext bytes, $D_1, \ldots, D_8$ be the result of applying the inverse of the PHT layer (which is unkeyed) to the ciphertext bytes, and $X(.)$ and $L(.)$ the S-boxes. The following linear relation, derived from (8), can be used:

$$\begin{aligned} X(P_1 \oplus K_1^1) \cdot \mathtt{01} \oplus L(P_2 \boxplus K_1^2) \cdot \mathtt{02} \oplus L(P_3 \boxplus K_1^3) \cdot \mathtt{01} \;\oplus \\ X(P_4 \oplus K_1^4) \cdot \mathtt{02} \oplus X(P_5 \oplus K_1^5) \cdot \mathtt{01} \oplus L(P_6 \boxplus K_1^6) \cdot \mathtt{02} \;\oplus \\ L(P_7 \boxplus K_1^7) \cdot \mathtt{01} \oplus X(P_8 \oplus K_1^8) \cdot \mathtt{02} \oplus X(D_2 \boxplus K_8^2) \cdot \mathtt{02} \;\oplus \\ L(D_4 \boxminus K_8^4) \cdot \mathtt{01} = K_i \cdot \mathit{\Gamma} K_i \quad , \end{aligned} \tag{16}$$

  where $\boxminus$ denotes subtraction in $\mathbb{Z}_{256}$.
- the $K_i \cdot \mathit{\Gamma} K_i$ bit is the following: $(K_2^2 \oplus K_2^4 \oplus K_2^6 \oplus K_2^8 \oplus K_3^6 \oplus K_6^1 \oplus K_6^5 \oplus K_7^2 \oplus K_7^4 \oplus K_8^4 \oplus K_8^8 \oplus K_9^2) \cdot \mathtt{02} \oplus (K_2^1 \oplus K_2^3 \oplus K_2^5 \oplus K_2^7 \oplus K_4^6 \oplus K_5^1 \oplus K_5^5 \oplus K_8^2 \oplus K_8^6 \oplus K_9^4) \cdot \mathtt{01} \oplus K_7^8 \cdot \mathtt{03}$, and
- the other 80 subkey bits are $K_1^1 \cdot \mathtt{ff}$, $K_1^2 \cdot \mathtt{ff}$, $K_1^3 \cdot \mathtt{ff}$, $K_1^4 \cdot \mathtt{ff}$, $K_1^5 \cdot \mathtt{ff}$, $K_1^6 \cdot \mathtt{ff}$, $K_1^7 \cdot \mathtt{ff}$, $K_1^8 \cdot \mathtt{ff}$, $K_9^2 \cdot \mathtt{ff}$, $K_9^4 \cdot \mathtt{ff}$. They can be identified with about $N \approx (2^{-29})^{-2} = 2^{58}$ known plaintext blocks using the maximum likelihood methods from [18].

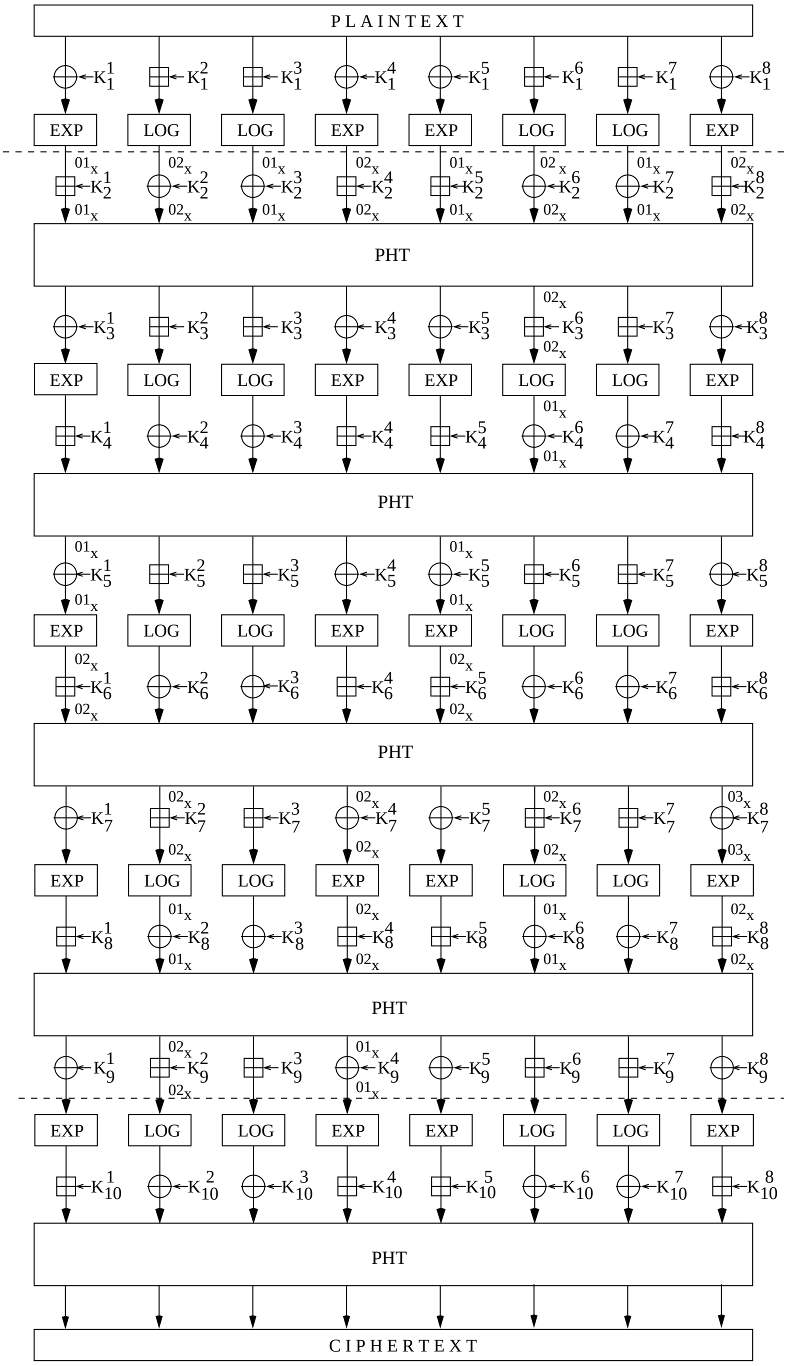
PLAINTEXT
K_1^1 K_1^2 K_1^3 K_1^4 K_1^5 K_1^6 K_1^7 K_1^8
EXP LOG LOG EXP EXP LOG LOG EXP
01_x 02_x 01_x 02_x 01_x 02_x 01_x 02_x
K_2^1 K_2^2 K_2^3 K_2^4 K_2^5 K_2^6 K_2^7 K_2^8
01_x 02_x 01_x 02_x 01_x 02_x 01_x 02_x
PHT
02_x
K_3^1 K_3^2 K_3^3 K_3^4 K_3^5 K_3^6 K_3^7 K_3^8
02_x
EXP LOG LOG EXP EXP LOG LOG EXP
01_x
K_4^1 K_4^2 K_4^3 K_4^4 K_4^5 K_4^6 K_4^7 K_4^8
01_x
PHT
01_x 01_x
K_5^1 K_5^2 K_5^3 K_5^4 K_5^5 K_5^6 K_5^7 K_5^8
01_x 01_x
EXP LOG LOG EXP EXP LOG LOG EXP
02_x 02_x
K_6^1 K_6^2 K_6^3 K_6^4 K_6^5 K_6^6 K_6^7 K_6^8
02_x 02_x
PHT
02_x 02_x 02_x 03_x
K_7^1 K_7^2 K_7^3 K_7^4 K_7^5 K_7^6 K_7^7 K_7^8
02_x 02_x 02_x 03_x
EXP LOG LOG EXP EXP LOG LOG EXP
01_x 02_x 01_x 02_x
K_8^1 K_8^2 K_8^3 K_8^4 K_8^5 K_8^6 K_8^7 K_8^8
01_x 02_x 01_x 02_x
PHT
02_x 01_x
K_9^1 K_9^2 K_9^3 K_9^4 K_9^5 K_9^6 K_9^7 K_9^8
02_x 01_x
EXP LOG LOG EXP EXP LOG LOG EXP
K_{10}^1 K_{10}^2 K_{10}^3 K_{10}^4 K_{10}^5 K_{10}^6 K_{10}^7 K_{10}^8
PHT
CIPHERTEXT

Fig. 2. A 1.25R attack on five rounds of SAFER-K/-SK (only non-zero bit-masks are shown)

Similarly, (12) can be used in a 1.25R-attack on four rounds of SAFER+, assuming weak key conditions (13) hold, as follows:

- let $P_1, \ldots, P_{16}$ be plaintext bytes, $D_1, \ldots, D_{16}$ be the result of applying the inverse of the PHT layer (which is unkeyed) to the ciphertext bytes, and $X(.)$ and $L(.)$ the S-boxes. The linear relation has the form:

$$\begin{aligned} L(P_2 \boxplus K_1^2) \cdot \mathtt{02} \oplus L(P_3 \boxplus K_1^3) \cdot \mathtt{01} \oplus X(P_4 \oplus K_1^4) \cdot \mathtt{02}\ \oplus \\ X(P_5 \oplus K_1^5) \cdot \mathtt{01} \oplus X(P_8 \oplus K_1^8) \cdot \mathtt{02} \oplus L(P_{10} \boxplus K_1^{10}) \cdot \mathtt{02}\ \oplus \\ X(P_{12} \oplus K_1^{12}) \cdot \mathtt{02} \oplus X(P_{13} \oplus K_1^{13}) \cdot \mathtt{02} \oplus L(P_{15} \boxplus K_1^{15}) \cdot \mathtt{01}\ \oplus \\ L(D_1 \boxminus K_8^1) \cdot \mathtt{02} \oplus L(D_4 \boxminus K_8^4) \cdot \mathtt{02} \oplus X(D_6 \oplus K_8^6) \cdot \mathtt{02}\ \oplus \\ X(D_7 \oplus K_8^7) \cdot \mathtt{03} \oplus L(D_8 \boxminus K_8^8) \cdot \mathtt{01} \oplus L(D_9 \boxminus K_8^9) \cdot \mathtt{02}\ \oplus \\ X(D_{10} \oplus K_8^{10}) \cdot \mathtt{03} \oplus X(D_{11} \oplus K_8^{11}) \cdot \mathtt{02} \oplus L(D_{12} \boxminus K_8^{12}) \cdot \mathtt{01}\ \oplus \\ L(D_{13} \boxminus K_8^{13}) \cdot \mathtt{02} \oplus X(D_{15} \oplus K_8^{15}) \cdot \mathtt{01} = K_i \cdot \mathit{\Gamma} K_i\ \ . \end{aligned} \tag{17}$$

- the $K_i \cdot \mathit{\Gamma} K_i$ bit is: $(K_2^2 \oplus K_2^4 \oplus K_2^8 \oplus K_2^{10} \oplus K_2^{12} \oplus K_2^{13} \oplus K_3^3 \oplus K_3^6 \oplus K_3^7 \oplus K_3^{10} \oplus K_3^{11} \oplus K_3^{14} \oplus K_3^{15} \oplus K_6^4 \oplus K_6^5 \oplus K_6^9 \oplus K_6^{13} \oplus K_6^{13} \oplus K_6^{16} \oplus K_7^1 \oplus K_7^4 \oplus K_7^6 \oplus K_7^9 \oplus K_7^{11} \oplus K_7^{13}) \cdot \mathtt{02} \oplus (K_7^7 \oplus K_7^{10}) \cdot \mathtt{03} \oplus (K_2^3 \oplus K_2^5 \oplus K_2^{15} \oplus K_4^3 \oplus K_4^6 \oplus K_4^7 \oplus K_4^{10} \oplus K_4^{11} \oplus K_4^{14} \oplus K_4^{15} \oplus K_5^4 \oplus K_5^5 \oplus K_5^9 \oplus K_5^{13} \oplus K_5^{16} \oplus K_7^8 \oplus K_7^{12} \oplus k_7^{15}) \cdot \mathtt{01}$ and
- the other 160 subkey bits are: $K_1^2 \cdot \mathtt{ff}$, $K_1^3 \cdot \mathtt{ff}$, $K_1^4 \cdot \mathtt{ff}$, $K_1^5 \cdot \mathtt{ff}$, $K_1^8 \cdot \mathtt{ff}$, $K_1^{10} \cdot \mathtt{ff}$, $K_1^{12} \cdot \mathtt{ff}$, $K_1^{13} \cdot \mathtt{ff}$, $K_1^{15} \cdot \mathtt{ff}$, $K_8^1 \cdot \mathtt{ff}$, $K_8^4 \cdot \mathtt{ff}$, $K_8^6 \cdot \mathtt{ff}$, $K_8^7 \cdot \mathtt{ff}$, $K_8^8 \cdot \mathtt{ff}$, $K_8^9 \cdot \mathtt{ff}$, $K_8^{10} \cdot \mathtt{ff}$, $K_8^{11} \cdot \mathtt{ff}$, $K_8^{12} \cdot \mathtt{ff}$, $K_8^{13} \cdot \mathtt{ff}$, $K_8^{15} \cdot \mathtt{ff}$. They can be identified using maximum likelihood techniques with about $N \approx (2^{-49})^{-2} = 2^{98}$ known plaintext blocks.

Finally, (14) leads to the following 1.25R attack on six rounds of SAFER-K32 (using weak keys) without the output transformation:

- let $P_1, \ldots, P_8$ be plaintext nibbles (4 bits), $D_1, \ldots, D_8$ be the result of applying the inverse of the PHT layer (which is unkeyed) to the ciphertext nibbles, and $X(.)$ and $L(.)$ the S-boxes. The linear relation has the form:

$$\begin{aligned} X(P_1 \oplus K_1^1) \cdot \mathtt{1} \oplus L(P_2 \boxplus K_1^2) \cdot \mathtt{2} \oplus L(P_3 \boxplus K_1^3) \cdot \mathtt{1}\ \oplus \\ X(P_4 \oplus K_1^4) \cdot \mathtt{2} \oplus X(P_5 \oplus K_1^5) \cdot \mathtt{1} \oplus L(P_6 \boxplus K_1^6) \cdot \mathtt{2}\ \oplus \\ L(P_7 \boxplus K_1^7) \cdot \mathtt{1} \oplus X(P_8 \oplus K_1^8) \cdot \mathtt{2} \oplus L(D_1 \boxplus K_8^1) \cdot \mathtt{3}\ \oplus \\ X(D_2 \oplus K_8^2) \cdot \mathtt{2} \oplus X(D_3 \oplus K_8^3) \cdot \mathtt{1} \oplus L(D_4 \boxminus K_8^4) \cdot \mathtt{1} = K_i \cdot \mathit{\Gamma} K_i\ \ . \end{aligned} \tag{18}$$

- the $K_i \cdot \mathit{\Gamma} K_i$ bit is: $(K_2^1 \oplus K_2^3 \oplus K_2^5 \oplus K_2^7 \oplus K_4^6 \oplus K_5^1 \oplus K_5^5 \oplus K_8^2 \oplus K_8^6 \oplus K_9^4 \oplus K_{10}^2 \oplus K_{11}^3 \oplus K_{11}^4) \cdot \mathtt{1} \oplus (K_7^8 \oplus K_{11}^1) \cdot \mathtt{3} \oplus (K_2^2 \oplus K_2^4 \oplus K_2^6 \oplus K_2^8 \oplus K_3^6 \oplus K_6^1 \oplus K_6^5 \oplus K_7^2 \oplus K_7^4 \oplus K_7^6 \oplus K_8^4 \oplus K_8^8 \oplus K_9^2 \oplus K_{10}^4 \oplus K_{11}^2) \cdot \mathtt{2}$ and then
- the other 48 subkey bits to be found are: $K_1^1 \cdot \mathtt{f}$, $K_1^2 \cdot \mathtt{f}$, $K_1^3 \cdot \mathtt{f}$, $K_1^4 \cdot \mathtt{f}$, $K_1^5 \cdot \mathtt{f}$, $K_1^6 \cdot \mathtt{f}$, $K_1^7 \cdot \mathtt{f}$, $K_1^8 \cdot \mathtt{f}$, $K_{12}^1 \cdot \mathtt{f}$, $K_{12}^2 \cdot \mathtt{f}$, $K_{12}^3 \cdot \mathtt{f}$, $K_{12}^4 \cdot \mathtt{f}$ using maximum likelihood methods with $N \approx (2^{-16})^{-2} = 2^{32}$ known plaintext blocks.

8 Methodology

The following procedure was used in order to obtain relations (8), (12) and (14):

- Initially, a Linear Approximation Table (LAT) for the S-boxes X and L was generated, containing for all possible input and output bit-masks the corresponding deviation value. Denoting by ΓI and ΓO general input and output bit-masks, and by $S(.)$ an S-box, each entry in the LAT contains
$$\mathrm{LAT}[\Gamma I, \Gamma O] = \Pr(I \cdot \Gamma I = S(I) \cdot \Gamma O) - \frac{1}{2} \tag{19}$$
for all possible inputs $I \in \mathbb{Z}_{256}$. Only one table is actually needed, because the X-box is the inverse of the L-box, and for the latter one can swap the input and output masks to obtain the corresponding linear approximations.
- The approximation of subkey addition and xor was made separately and then evaluated together with the approximations for the S-boxes, confirming the key dependency (for addition). Indeed, the bias decreases if the subkeys do not exhibit a pattern that allows the expected bit-mask approximation. For example, the mask $M_2(x) = \mathtt{02_x} \cdot x$, in the output of an addition operation with odd-valued key bytes gives zero bias; otherwise, the overall bias is the one provided by the xor of subkey and the X-box approximations.
- The approach taken for generating linear relations for the PHT layer was not exhaustive as was done for the S-boxes. Due to the addition operation performed in the 2-PHT boxes, it was observed that the most biased linear relations through the PHT layers would explore preferably the LSB(s) of the 2-PHT because they are least affected by carry bits. Besides, exploring few LSBs would require less weak-key restrictions. It was decided, arbitrarily, to concentrate efforts in the two LSBs only of each 2-PHT. A linear hull approximation was made for the PHT layer, because it was observed that linking together local (non-zero bias) approximations for the 2-PHT boxes could sometimes result in linear approximations for the PHT layer with zero bias. Indeed, some component relations of the linear hull have positive and others negative deviation with the same absolute value, which can cancel the effect of each other.
  Note that a (basic) linear relation tracks a single (approximation) path between input and output bits of a round component. A linear hull [14] corresponds to a set of linear relations all of which share the same input and output bit-masks, but each relation takes different paths across the component.
- The next step consisted in combining linear hulls for the (NL+subkey) layers with others for the PHT layer, in order to generate one-round linear approximations (hulls). Further, these one-round relations were combined either on top or at the bottom end of each other, in order to get as long a linear relation as possible. Such stacking strategy of combining one-round approximations was based on the idea of the inside-out attack of Wagner [5]. While constructing the final linear hull an important restriction was to try to keep the number of active S-boxes as small as possible from one

round to the next, both in order to control the overall bias as well as to avoid attacking too many subkey bits at both ends of the cipher. For a linear relation of bias ϵ, the known-plaintext requirements for an effective (high success rate) linear attack on $2n$-bit block ciphers is $N \geq (\epsilon)^{-2}$, that is, $\epsilon \geq (\sqrt{2^{2n}})^{-1} = 2^{-n}$. Therefore, an immediate restriction for 64-bit-block cipher versions is $\epsilon \geq 2^{-32}$. Similarly, for SAFER+, $\epsilon \geq 2^{-64}$, and for SAFER-K32, $\epsilon \geq 2^{-16}$.

9 Conclusion

In this paper, SAFER-K32, SAFER-K/-SK and SAFER+ ciphers were analyzed using non-homomorphic linear cryptanalysis. Table 1 summarizes our results.

Table 1. Linear relations found for the SAFER cipher family

	Cipher		
	SAFER-K32	SAFER-K/-SK	SAFER+
# rounds lin. rel.	4.75	3.75	2.75
Bias (weak keys)	2^{-16}	2^{-29}	2^{-49}
Attack type	1.25R	1.25R	1.25R
# subkey bits	48+1 †	80+1 †	160+1 †
Time complexity (parity computation)	$2^{48+2\cdot16}$	$2^{80+2\cdot29}$	$2^{160+2\cdot49}$
Space complexity	$\approx 2^{32}$	$\approx 2^{58}$	$\approx 2^{98}$

† That is the worst case, i.e. assuming all the subkeys are independent.

The algorithm used for our attack uses Matsui's idea of keeping only the highest parity counter(s) (say, for the best ten key candidates). The advantage is that we do not need to keep separate counters for each key candidate. But, on the other hand, we need to store all the plaintext samples (therefore our space requirement: $N \approx \epsilon^{-2}$).

The 3.75 round linear relation (8) found for SAFER-K/-SK does not contradict the results in [3] (the 1.5 round relation (7)) because the former is non-homomorphic. Besides, the non-homomorphicity of linear relations (8)–(14) caused them to be key-dependent for the bit-mask approximations to hold, while (7) is key-independent. Another interesting observation is that relation (8) actually holds for any of the 128 possible S-box generators (of GF(257)), not only for 45 as used in SAFER-K/-SK and SAFER+. That is because, only few approximations are actually used, namely, the ones which explore the two LSB's in both the input and output masks. Therefore, changing the S-boxes' generator would not help protect these ciphers against our particular linear attack.

Table 2. Plaintext requirements of DC attacks by Knudsen–Berson (KB) and Wu, Boa, Deng and Ye (WBDY) and LC attacks on SAFER-K64

	Differential (chosen/known texts)			Linear (known texts)	
#rounds	KB [17]	WBDY [7]		Harpes [3]	this paper
	chosen	chosen	known		(weak keys)
2	—	—	—	$\approx 2^{15.2}$	$\approx 2^{8}$ †
3	—	—	—	$(> 2^{64})$	$\approx 2^{12}$ ‡
4	—	—	—	—	$\approx 2^{28}$ §
5	$\approx 2^{45}$	$\approx 2^{38}$	$\approx 2^{51}$	—	$\approx 2^{58}$
6	$(> 2^{64})$	$\approx 2^{53}$	$\approx 2^{59}$	—	$(> 2^{64})$
7	—	$(> 2^{64})$	$(> 2^{64})$	—	—

† Any homomorphic-only approximation can be used here
‡ This approximation comes from the first 1.75 rounds of relation (8)
§ This approximation comes from the first 2.75 rounds of relation (8)

Nonetheless, our relation (14) is only valid for SAFER-K32 with generator $g = 11$. For the other seven possible generators of $GF(17)$ the linear relation (14) does not hold.

Table 2 compares the current analysis to other attacks on SAFER-K64. We conclude that the attack based on truncated differentials by Wu *et al.* [7] (which improves the original attack by Knudsen and Berson [17]) is still the best shortcut attack on SAFER-K64. Moreover, while differential attacks are typically chosen plaintext attacks, they can be converted to known plaintext attacks (see Biham and Shamir [6, p. 31]).

Theoretically it was predicted that for five rounds one key in 1024 is 'weak' (restrictions (9)), which means that the relation (8) with theoretical bias 2^{-37} can actually be used, restricted to weak keys, with bias 2^{-29}. Nonetheless, practical implementations of the attack show that only one out of eight keys is actually weak (only three of the subkey bits K_2^4, K_2^8, K_9^2, at the beginning and end of the linear hull need to have a certain bit-pattern). This may be another consequence of the linear-hull effect.

The main conclusion however is that, while the analysis of Harpes *et al.* could be improved, linear cryptanalysis does not seem a serious threat, even to SAFER-K64 with its nominal number of rounds.

The main contribution of this paper towards better block cipher design is the issue of key-dependency in linear cryptanalysis, through non-homomorphic bitmasks. This improved on previous linear attacks on all SAFER family members by specifying linear relations valid for particular key classes; this analysis can have the same effect on other similar designs.

Acknowledgments. It is a pleasure to acknowledge the fruitful discussions with Johan Borst and Vincent Rijmen during this research and the preparation of all previous drafts of this paper.

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer
Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

References

1. Bluetooth Specification version 1.0B.
   Available at `http://www.bluetooth.com/link/spec/bluetooth_b.pdf`
2. Brincat, K., Meijer, A., "On the SAFER cryptosystem," *Cryptography and Coding, Proceedings of 6th IMA Conference, LNCS 1355,* M. Darnell, Ed., Springer-Verlag, 1997, pp. 59-68.
3. C. Harpes, *"Cryptanalysis of Iterated Block Ciphers,"* ETH series in Information Processing, J.L. Massey, Ed., Vol. 7, Hartung-Gorre Verlag, Konstanz, 1996.
4. C. Harpes, G. Kramer, J.L. Massey, "A generalization of linear cryptanalysis and the applicability of Matsui's piling-up lemma," *Advances in Cryptology, Proceedings Eurocrypt'95, LNCS 921*, L.C. Guillou and J.-J. Quisquater, Eds., Springer-Verlag, 1995, pp. 24–38.
5. D. Wagner, "The boomerang attack," *Fast Software Encryption, LNCS 1636*, L.R. Knudsen, Ed., Springer-Verlag, 1999, pp. 201–214.
6. E. Biham, A. Shamir, *"Differential Cryptanalysis of the Data Encryption Standard,"* Springer-Verlag, 1993.
7. H. Wu, F. Bao, R.H. Deng, Q.-Z. Ye, "Improved truncated differential attacks on SAFER," *Advances in Cryptology, Proceedings Asiacrypt'98, LNCS 1514*, K. Ohta, D. Pei, Eds., Springer-Verlag, 1998, pp. 133–147.
8. J. Borst, B. Preneel, J. Vandewalle, "Linear Cryptanalysis of RC5 and RC6," *Fast Software Encryption, LNCS 1636*, L.R. Knudsen, Ed., Springer-Verlag, 1999, pp. 16–30.
9. J. Kelsey, B. Schneier, D. Wagner, "Key schedule weaknesses in SAFER+," *Proceedings 2nd Advanced Encryption Standard Candidate Conference*, March 22–23, 1999, Rome (I), pp. 155–167.
10. J.L. Massey, G.H. Khachatrian, M.K. Kuregian, *"Nomination of SAFER+ as candidate algorithm for the Advanced Encryption Standard (AES),"* June 12, 1998. Available at `http://www.ii.uib.no/~larsr/aes.html`
11. J.L. Massey, "SAFER-K64: a byte-oriented block ciphering algorithm," *Fast Software Encryption, LNCS 1039*, D. Gollmann, Ed., Springer-Verlag, 1996, pp. 1–17.
12. J.L. Massey, "SAFER-K64: one year later," *Fast Software Encryption, LNCS 1008*, B. Preneel, Ed., Springer-Verlag, 1995, pp. 212–241.
13. J.L. Massey, "Strengthened key schedule for the cipher SAFER," *posted to the USENET newsgroup sci.crypt*, September 1995.
    Available at `ftp://ftp.cert.dfn.de/pub/tools/crypt/SAFER/`
14. K. Nyberg, "Linear approximation of block ciphers," *Advances in Cryptology, Proceedings Eurocrypt'94, LNCS 950*, A. De Santis, Ed., Springer-Verlag, 1995, pp. 439–444.
15. L.R. Knudsen, "A key schedule weakness in SAFER-K64," *Advances in Cryptology, Proceedings Crypto'95, LNCS 963*, D. Coppersmith, Ed., Springer-Verlag, 1995, pp. 274–286.
16. L.R. Knudsen, "Why SAFER K changed its name," *Technical Report LIENS 96-13*, Laboratoire d'Informatique, Ecole Normale Supérieure, Paris, France, April 1996. Available at `http://www.ii.uib.no/~larsr/aes.html`
17. L.R. Knudsen, T.A. Berson, "Truncated differentials of SAFER," *Fast Software Encryption, LNCS 1039*, D. Gollmann, Ed., Springer-Verlag, 1996, pp. 15–26.
18. M. Matsui, "Linear cryptanalysis method for DES cipher," *Advances in Cryptology, Proceedings Eurocrypt'93, LNCS 765*, T. Helleseth, Ed., Springer-Verlag, 1994, pp. 386–397.

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.

19. M. Matsui, A. Yamagishi, "A new method for known plaintext attack on FEAL cipher," *Advances in Cryptology, Proceedings Eurocrypt'92, LNCS 658*, R.A. Rueppel, Ed., Springer-Verlag, 1993, pp. 81–91.
20. S. Murphy, "An analysis of SAFER," *Journal of Cryptology*, Vol. 11, No. 4, 1998, pp. 235–251.
21. S. Vaudenay, "On the need for multipermutations: Cryptanalysis of MD4 and SAFER," *Fast Software Encryption, LNCS 1039*, D. Gollmann, Ed., Springer-Verlag, 1996, pp. 286–297.

A A Ciphertext-Only Attack

In all previous linear attacks we did not make any assumption on the plaintext distribution. In many cases the plaintext consists (mostly) of printable ASCII characters, that is, characters with values between 20_x and $7E_x$. Matsui developed for this case a ciphertext only attack on DES with a reduced number of rounds [18]. For SAFER-K/SK, there exist linear hulls which allow for a 3-round ciphertext-only attack when the most significant bit of all plaintext bytes are equal to zero. One such hull with 2.25 rounds and $\mathcal{S} = 2$ is

$$\begin{aligned}
&(\texttt{0000000000000080}_\texttt{x}\ ,\texttt{0200000000000000}_\texttt{x}\ ,2^{-6})\ \text{(one-round)}\\
&(\texttt{0200000000000000}_\texttt{x}\ ,\texttt{0202020202020202}_\texttt{x}\ ,2^{-9})\ \text{(one round)}\\
&(\texttt{0202020202020202}_\texttt{x}\ ,\texttt{0202020202020202}_\texttt{x}\ ,2^{-5})\ \text{(subkey quarter-round)}
\end{aligned} \tag{20}$$

which has bias $\epsilon_4^* = 2^{-18}$. The key dependency conditions for the validity of (20) in a 0.75R attack on three rounds of SAFER-K/-SK are

$$\mathrm{LSB}(K_4^1, K_5^2, K_5^3, K_5^6, K_5^7) = 0\ , \tag{21}$$

The actual bias of (20) is therefore $\epsilon_4 = 2^{-13}$.

Let $P_1, \ldots, P_8$ denote the plaintext bytes, $D_1, \ldots, D_8$ the result of applying the inverse of the (unkeyed) PHT layer to the ciphertext bytes, and $X(.)$ and $L(.)$ the S-boxes. We use the following linear relation based on (20):

$$\begin{aligned}
P_8 \cdot \texttt{80} \oplus L(D_1 \boxminus K_6^1) \cdot \texttt{02} \oplus X(D_2 \oplus K_6^2) \cdot \texttt{02} \oplus\\
X(D_3 \oplus K_6^3) \cdot \texttt{02} \oplus L(D_4 \boxminus K_6^4) \cdot \texttt{02} \oplus\\
L(D_5 \boxminus K_6^5) \cdot \texttt{02} \oplus X(D_6 \oplus K_6^6) \cdot \texttt{02} \oplus\\
X(D_7 \oplus K_6^7) \cdot \texttt{02} \oplus L(D_8 \boxminus K_6^8) \cdot \texttt{02} \oplus = K_i \cdot \varGamma K_i\ .
\end{aligned} \tag{22}$$

If the plaintext is composed mostly of ASCII characters then equation (22) reduces to

$$\begin{aligned}
L(D_1 \boxminus K_6^1) \cdot \texttt{02} \oplus X(D_2 \oplus K_6^2) \cdot \texttt{02} \oplus\\
X(D_3 \oplus K_6^3) \cdot \texttt{02} \oplus L(D_4 \boxminus K_6^4) \cdot \texttt{02} \oplus\\
L(D_5 \boxminus K_6^5) \cdot \texttt{02} \oplus X(D_6 \oplus K_6^6) \cdot \texttt{02} \oplus\\
X(D_7 \oplus K_6^7) \cdot \texttt{02} \oplus L(D_8 \boxminus K_6^8) \cdot \texttt{02} \oplus = K_i \cdot \varGamma K_i\ .
\end{aligned} \tag{23}$$

keeping the same bias $\epsilon_4 = 2^{-13}$ because $P_8 \cdot 80 = 0$ has bias $\epsilon_5 = 2^{-1}$.

The actual plaintext does not need to be composed of ASCII only characters, like in .HTML files. Some experiments show that even .JPG, .MP3, and .WAV files contain some small bias in the most significant bit of each byte, like $\epsilon_5 = 2^{-10}$; combined with the bias of (20) this results in a linear hull with bias $\epsilon_4 = 2^{-22}$ requiring about $N = 2^{44}$ ciphertext (only) blocks. Note that some popular (UNIX) file compression utilities like "compress" and "gzip" can destroy the redundancy of the MSB byte in ASCII files, but apparently they cannot destroy the bias of .JPG, .MP3 or .WAV files. Other possible biased distributions of (combinations of) plaintext bits can also be explored, that is, there is no need to consider only the most significant bit.

Appeared in *Fast Software Encryption, FSE 2000*, Lecture Notes in Computer Science 1978, B. Schneier (ed.), Springer-Verlag, pp. 244–261, 2000.