Markov Truncated Differential Cryptanalysis of Skipjack

Ben Reichardt and David Wagner

UC Berkeley, Berkeley CA 94720, USA,
{breic,daw}@cs.berkeley.edu

Abstract. Using Markov chains, we systematically compute all the truncated differentials of Skipjack, assuming the nonlinear G boxes are random permutations. We prove that an attacker with one random truncated differential from each of 2^{128} independently-keyed encryption oracles has advantage of less than 2^{-16} in distinguishing whether the oracles are random permutations or the Skipjack algorithm.

1 Introduction

Skipjack encrypts a 64-bit plaintext by applying eight type-A rounds, followed by eight type-B rounds, then eight more A rounds and eight more B rounds [8]. The A and B rounds, given in Fig. 1, are feedback shift registers on four 16-bit words, with a nonlinear permutation G (a four-round Feistel network). B rounds, when keyed correctly and with w_1 swapped with w_2, w_3 swapped with w_4, invert the A rounds. Compare the "Fibonacci" form of the A-round shift register to that of a Feistel cipher shown in Fig. 2; they differ in the placement of the nonlinear function.

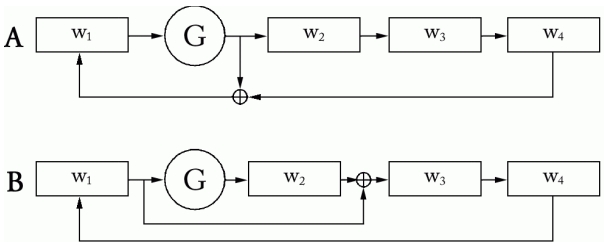

Fig. 1. Skipjack consists of eight rounds through the type A shift register, followed by eight rounds of the reversed shift register type B, then eight further rounds each of types A and B. G is a keyed bijection. A round counter is also input to the exor

A differential for a cipher consists of a set Δ of differences of plaintexts, a set Δ^* of differences of ciphertexts, and a probability p that two randomly chosen

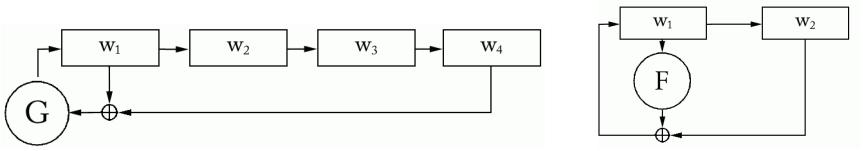

Fig. 2. The "Fibonacci" form of the Skipjack A round (left) has its nonlinear permutation G in a different position than the nonlinear function F of a Feistel shift register (right)

plaintexts with difference in Δ will be encrypted to ciphertexts with difference in Δ^*, for a fixed key. For Skipjack, the appropriate definition of the "difference" between texts t and t' is the exor $t \oplus t'$. Differentials with p significantly different – either more or less – than the same probability for a random permutation can be used as distinguishers, and often to build key-recovery attacks [4]. For traditional differentials, the sets Δ, Δ^* are singletons [2]; so-called truncated differentials deal with larger sets [5].

There are no known attacks on all 32 rounds of Skipjack better than exhaustive search over the 80-bit key space. However, the unbalanced structure of the Skipjack network, in which the nonlinear permutation affects only a quarter of the bits at each round, slows diffusion of nonlinearities, and makes it easy to follow differentials across multiple rounds. The best known attacks on up to 31 rounds of Skipjack are truncated differential attacks by Biham et al. [1] and Knudsen et al. [6]. They found differentials either by tracing through by hand an unrolled diagram of Skipjack rounds, or by having a computer search for differentials across a similar form cipher with fewer bits per word – e.g., a 32-bit cipher with 8-bit words. However, Granboulan [4] found several errors in the probability calculations of [6].

We here apply the Markov techniques of [7] to systematically and efficiently calculate correct probabilities for all the truncated differentials of Skipjack, under the assumption that the G boxes are random permutations. We verify Granboulan's corrected differential probabilities. We verify the 24-round impossible differential (i.e., a differential with $p = 0$) used in the attack of [1], and find three mistaken differentials they state but do not use in their attack. We also describe several new "distinguishers", for instance a 30-round truncated differential. We show that there are no good truncated differentials for the full 32 rounds of Skipjack. Additionally, we estimate the best distinguishing advantage gained by considering multiple differentials, instead of just a single one. We prove that an attacker with one random truncated differential from each of 2^{128} independently-keyed encryption oracles has advantage of less than 1.0003×2^{-17} in distinguishing whether the oracles are random permutations or the Skipjack algorithm. Since no attack on Skipjack can obtain more than $2^{64}(2^{64} - 1)/2$ plaintext pairs, this provides heuristic evidence that Skipjack may be secure against truncated differential distinguishing attacks.

The contributions of this paper are twofold. First, we carefully develop a framework for the analysis of truncated differential attacks, and we introduce new mathematical tools for precisely characterizing the strength of a cipher against truncated differential attacks. Notably, we give methods for calculating the exact probability of truncated differentials (many trails), in contrast to many previous works which only looked at truncated differential characteristics (a single trail), and we show how to bound the distinguishing advantage of any truncated differential attack, even one that uses several truncated differentials simultaneously. Second, we apply these methods to Skipjack, and we characterize its strength against truncated differential cryptanalysis. We hope that these investigations will yield new insight into the structure of Skipjack and more generally into the analysis of security against truncated differential attacks.

In Sect. 2 we define Markov ciphers, for which round applications correspond to a Markov process. In Sect. 3, we introduce our methods and use them to compute truncated differentials for a simple two-word Skipjack variant. In Sect. 4, we compute truncated differentials for a three-word Skipjack variant and give a criterion for weak key classes. In Sect. 5, we give the Markov transition matrices for full, four-word Skipjack's A and B rounds. We correct some differential probability calculations. In Sect. 6, we bound the advantage an attacker gains from using multiple (independent) differentials. Finally, in Sect. 7 we list the best truncated differentials for Skipjack. We also show that it is unlikely that there exist any weak key classes for which a truncated differential attack would be significantly improved.

2 Markov Ciphers

Let Λ be a collection of nonempty, pairwise-disjoint subsets of $\Gamma \times \Gamma$ covering $\Gamma \times \Gamma$. For $f : \Gamma \to \Gamma$, define $\tilde{f} : \Gamma \times \Gamma \to \Gamma \times \Gamma$ by $\tilde{f}(x, y) = (f(x), f(y))$. If F is a random distribution over functions $\Gamma \to \Gamma$, we call F *Markov* with respect to Λ, or Λ-Markov, if for all $\Delta, \Delta' \in \Lambda$, for f sampled according to F and δ uniformly distributed in Δ such that $\tilde{f}(\delta) \in \Delta'$, $\tilde{f}(\delta)$ is uniformly distributed in Δ'. That is, using $\in_R$ to mean sampling according to the uniform distribution, $\Pr[\tilde{f}(\delta) = (x, y) | \delta \in_R \Delta, f(\delta) \in \Delta']$ is independent of $(x, y) \in \Delta'$.

For F, G distributions over functions $\Gamma \to \Gamma$, we define the distribution $G \circ F$ as given by that of $g \circ f$, where g, f are sampled from G, F, respectively. Then for F, G independent Λ-Markov function distributions, the distribution $G \circ F$ is also Λ-Markov, and for $\Delta, \Delta', \Delta'' \in \Lambda$, $\Pr_{G \circ F}[(\tilde{g} \circ \tilde{f})(\delta) \in \Delta'', \tilde{f}(\delta) \in \Delta' | \delta \in_R \Delta] = \Pr_G[\tilde{g}(\delta') \in \Delta'' | \delta' \in_R \Delta'] \cdot \Pr_F[\tilde{f}(\delta) \in \Delta' | \delta \in_R \Delta]$.

We can associate to any Λ-Markov distribution F a corresponding transition matrix $[F]$ with a row and a column for each $\Delta \in \Lambda$, defined by $[F]_{\Delta', \Delta} = \Pr_F[\tilde{f}(\delta) \in \Delta' | \delta \in_R \Delta]$. Then the above property implies $[G \circ F] = [G][F]$, i.e., functional composition corresponds to matrix multiplication.

When Γ is the set of texts, an iterated round block cipher is a Λ-Markov cipher if its round functions are independent Λ-Markov function distributions, where the probability is taken over the round key. For example, if its rounds

were keyed independently, DES would be a Markov cipher with respect to the exor operation $\oplus$, i.e., with respect to Λ consisting of sets $\{(x, y)|x \oplus y = z\}$ for every z [7], [2].

3 Two-Word Skipjack Variant

To introduce our techniques, we'll consider an adaptation of Skipjack into a two-word cipher with A and B rounds as shown in Fig. 3. As with full Skipjack, the A round is inverted by the B round with G^{-1} and swapped words. By analogy with full Skipjack, two-word Skipjack begins with four A rounds, then has four B rounds, then four more of each of the A and B rounds: $SJ_2(x) = B^4A^4B^4A^4(x)$ for $x \in \{0,1\}^{2n}$. Assume the G boxes are uniformly random permutations $\{0,1\}^n \to \{0,1\}^n$, chosen independently for each round.

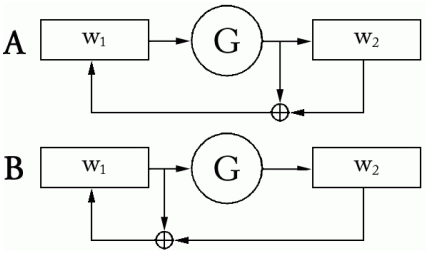

Fig. 3. Our two-word Skipjack variant consists of four A rounds (above), then four B rounds (below), then four more A rounds and four more B rounds

If α is an arbitrary plaintext, then the encryption $SJ_2(\alpha)$ will be drawn uniformly at random from $\{0,1\}^{2n}$. We ask the following natural question: for any given β, what is the distribution of $SJ_2(\beta)$ given $SJ_2(\alpha)$? Since $SJ_2 : \{0,1\} \to \{0,1\}^{2n}$ is a permutation, $SJ_2(\alpha) = SJ_2(\beta)$ if and only if $\alpha = \beta$. But even for $\alpha \neq \beta$, the distribution of $SJ_2(\alpha) \oplus SJ_2(\beta)$ is *not* uniformly distributed across the nonzero elements of $\{0,1\}^{2n}$. If this distribution – of course depending on the relationship between α and β – is sufficiently non-uniform, it may indicate exploitable weaknesses in the cipher. In practice, a noticeably non-uniform distribution of β's encryption given α's encryption across fewer than all the rounds can often give attacks.

To make our calculations, we'll consider each round separately, and then combine the rounds using the round-independence assumption. For $x \in \{0,1\}^{2n}$ we call its first word $x_1 \in \{0,1\}^n$ and its second word x_2, so $x = (x_1, x_2)$. Note $A(x) = (G(x_1) \oplus x_2, G(x_1))$ and $B(x) = (x_1 \oplus x_2, G(x_1))$.

For a pair of words α_1 and β_1, the joint distribution of their images $G(\alpha_1)$, $G(\beta_1)$ through a G box depends only on whether the two words are the same. Indeed, if $\alpha_1 = \beta_1$ then $G(\alpha_1) = G(\beta_1)$ is uniformly distributed across $\{0,1\}^n$.

If $\alpha_1 \neq \beta_1$ then $G(\alpha_1) \in_R \{0,1\}^n$ and $G(\alpha_1) \oplus G(\beta_1) \in_R \{0,1\}^n \setminus \{0\}$. The minimum amount of information needed to keep track of whether the input difference to a G box is zero or nonzero is given by the *truncated differential class* of the pair α, β. For two-word Skipjack, there are five truncated differential classes of text pairs:

$$
\begin{aligned}
[(a,b)] &\equiv \{(\alpha,\beta) | \alpha_1 \oplus \beta_1, \alpha_2 \oplus \beta_2, (\alpha_1 \oplus \beta_1) \oplus (\alpha_2 \oplus \beta_2) \neq 0\}, \\
[(a,a)] &\equiv \{(\alpha,\beta) | \alpha_1 \oplus \beta_1 = \alpha_2 \oplus \beta_2 \neq 0\}, \\
[(a,0)] &\equiv \{(\alpha,\beta) | \alpha_2 \oplus \beta_2 \neq 0, \alpha_2 \oplus \beta_2 = 0\}, \\
[(0,b)] &\equiv \{(\alpha,\beta) | \alpha_1 \oplus \beta_1 = 0, \alpha_2 \oplus \beta_2 \neq 0\}, \\
[(0,0)] &\equiv \{(\alpha,\beta) | \alpha = \beta\} \ .
\end{aligned}
$$

We will henceforth omit the brackets around an equivalence class, and write, e.g., (a,b) for $[(a,b)]$. Our notation deviates slightly from previous work in that we require truncated differential classes to be pairwise disjoint. Earlier authors would typically consider the classes (a,a), $(a,0)$ and $(0,b)$ to be subclasses of (a,b), whereas in our setting every text pair belongs to exactly one truncated differential class.

We treat the evolution of a differential across cipher rounds as a Markov process on the truncated differential classes. Let $\delta = (\alpha_\delta, \beta_\delta) \in \{0,1\}^{2n} \times \{0,1\}^{2n}$ denote a pair of texts, and define $\tilde{A} : \{0,1\}^{2n} \to \{0,1\}^{2n}$ by $\tilde{A}(\delta) = (A(\alpha_\delta), A(\beta_\delta))$. Then if $\delta \in_R (0,0)$, $\tilde{A}(\delta) \in_R (0,0)$. If $\delta \in_R (0,a)$, $\tilde{A}(\delta) \in_R (a,0)$. If $\delta \in_R (a,0)$, $\tilde{A}(\delta) \in_R (a,a)$. If $\delta \in_R (a,a)$ or $\delta \in_R (a,b)$, then with probability $\frac{1}{2^n-1}$, $\tilde{A}(\delta) \in_R (a,0)$, and with probability $1 - \frac{1}{2^n-1}$, $\tilde{A}(\delta) \in_R (a,b)$. Figure 4 shows the matrix $[A]$, as well as $[B]$, which can be calculated similarly.

$$
\begin{pmatrix}
1-\frac{1}{2^n-1} & 1-\frac{1}{2^n-1} & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
\frac{1}{2^n-1} & \frac{1}{2^n-1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\qquad
\begin{pmatrix}
1-\frac{1}{2^n-1} & 0 & 1-\frac{1}{2^n-1} & 0 & 0 \\
\frac{1}{2^n-1} & 0 & \frac{1}{2^n-1} & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
$$

Fig. 4. The transition matrices for two-word A (left) and B (right). The row and column order is $[(a,b)], [(a,a)], [(a,0)], [(0,b)], [(0,0)]$

The rounds A, B are Markov with respect to $\{(a,b),(a,a),(a,0),(0,b),(0,0)\}$, so we can compute the transition probabilities for any sequence of A and B rounds simply by multiplying the matrices appropriately. So, for example, the entry in row (a,a) and column (a,b) of $([B]^4[A]^4)^2$ gives the probability that a pair of plaintexts $\delta \in_R (a,b)$ is sent by the sixteen rounds of two-word Skipjack to a pair of ciphertexts in (a,a); and all output differences in (a,a) are equally likely.

For comparison, a random permutation takes a nonzero difference to a random nonzero difference. Since $|(0,0)| = 2^{2n}$, $|(0,b)| = |(\alpha,0)| = |(a,a)| =$

$2^{3n} - 2^{2n}$, $|(a,b)| = 2^{4n} - 3 \cdot 2^{3n} + 2 \cdot 2^{2n}$, the probability that a nonzero difference is sent to (a,b) is $|(a,b)|/(2^{4n} - 2^{2n}) = \frac{2^n - 2}{2^n + 1}$, and the probability that a nonzero difference is sent to, e.g., (a,a) is $|(a,a)|/(2^{4n} - 2^{2n}) = \frac{1}{2^n + 1}$.

This technique of computing the transition matrix can be applied to any cipher consisting of exors, linear operations, and components that can be modeled as a random permutation or random function – in particular to Feistel ciphers. Figure 5 shows the transition matrices between these equivalence classes for a two-word Feistel network in which the nonlinear function is either a random function, or a random permutation.

$$\begin{pmatrix} 1-\frac{2}{2^n} & 1-\frac{2}{2^n} & 1-\frac{2}{2^n} & 0 & 0 \\ \frac{1}{2^n} & \frac{1}{2^n} & \frac{1}{2^n} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \frac{1}{2^n} & \frac{1}{2^n} & \frac{1}{2^n} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1-\frac{2}{2^n} & 1-\frac{1}{2^n} & 1-\frac{1}{2^n} & 0 & 0 \\ \frac{1}{2^n} & 0 & \frac{1}{2^n} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \frac{1}{2^n} & \frac{1}{2^n} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Fig. 5. The transition matrix for a two-word Feistel cipher $(x_1, x_2) \mapsto (F(x_1) \oplus x_2, x_1)$, for $F : \{0,1\}^n \to \{0,1\}^n$ a random function (left), and for F is a random permutation (right). Once again, the row and column order is $(a,b), (a,a), (a,0), (0,b), (0,0)$

4 Three-Word Skipjack Variant

We adapt Skipjack into a three-word cipher with A and B rounds as shown in Fig. 6. Similarly to full Skipjack, the A round is inverted by the B round with G^{-1} and swapped words w_1 and w_2. By analogy with full Skipjack, three-word Skipjack begins with six A rounds, then has six B rounds, then six more of each of the A and B rounds: $SJ_3(x) = B^6 A^6 B^6 A^6(x)$ for $x \in \{0,1\}^{3n}$. Assume the G boxes are uniformly random permutations $\{0,1\}^n \to \{0,1\}^n$, chosen independently for each round.

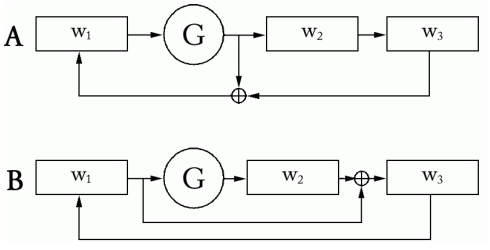

Fig. 6. Our three-word Skipjack variant consists of six A rounds (above), then six B rounds (below), then six more A rounds and six more B rounds

As for two-word Skipjack, to compute long enough differentials we can restrict our attention to sixteen truncated differential classes, corresponding to the sixteen linear subspaces of $GF(2)^3$:

$$(a,b,c),$$
$$(0,b,c),(a,0,c),(a,b,0),(a,a,c),(a,b,a),(a,b,b),(a,b,a\oplus b),$$
$$(a,0,0),(0,b,0),(0,0,c),(a,a,0),(a,0,a),(0,b,b),(a,a,a),$$
$$(0,0,0)\ .$$

This notation captures exactly which exor relationships between the words hold. For example, $(a,b,a\oplus b)$ contains all plaintext pairs of the form $\big((x_1,x_2,x_3),(x_1\oplus a',x_2\oplus b',x_3\oplus(a'\oplus b'))\big)$ with $x\in\{0,1\}^{3n}$ and $a',b',a'\oplus b'\neq 0$. Every pair of texts in $\{0,1\}^{6n}$ belongs to exactly one truncated differential equivalence class. The equivalence class (a,b,c) contains $2^{3n}(2^n-1)(2^n-2)(2^n-4)$ text pairs, the next seven classes each contain $2^{3n}(2^n-1)(2^n-2)$ text pairs, the next four classes each contain $2^{3n}(2^n-1)$ pairs, and $|(0,0,0)|=2^{3n}$ of course.

It is an easy exercise to compute the matrices $[A]$, $[B]$ as for two-word Skipjack, so we'll compute them in a slightly different manner here. Define $G_3:\{0,1\}^{3n}\to\{0,1\}^{3n}$ by $G_3(x_1,x_2,x_3)=(G(x_1),x_2,x_3)$. We'll compute $[G_3]$. Then, since wordwise permutations and exors just permute the differential classes and A, B just differ from G_3 by wordwise permutations and xors, $[G_3]$ differs from $[A]$, $[B]$ only by left- and right-multiplication by some permutation matrices.

Consider for example a text pair in (a,a,a). If G_3 fixes the difference in the first word, then $\tilde{G}_3$ will map the text pair into another pair in (a,a,a); otherwise it will be sent to (a,b,b). Hence the total number of text pairs mapped from (a,a,a) to (a,a,a) is exactly $2^{2n}\theta$, where $\theta=|\{(x,y)\in\{0,1\}^n\times\{0,1\}^n|G(x)\oplus G(y)=x\oplus y\}|$ denotes the number of exor differences fixed by G (the factor of 2^{2n} comes in because it doesn't matter what the last two words of the texts are, as long as their differences are correct). The total number of text pairs mapped from (a,a,a) to (a,b,b) is $2^{3n}(2^n-1)-2^{2n}\theta=2^{2n}(2^n(2^n-1)-\theta)$. Similar arguments show that the number of text pairs mapped from any of $(0,b,b)$, $(a,0,a)$, or $(a,a,0)$ to themselves is also $2^{2n}\theta$.

Consider next a text pair in (a,a,c). G_3 maps the text pair to another in (a,a,c) if and only if the difference in the first words is fixed, so the number of text pairs mapped from (a,a,c) to itself is exactly $2^{2n}(2^n-2)\theta$; the factor of 2^n-2 comes in because the difference in the third words can be anything except 0 or the difference in the first words. If, however, the difference in the first words is not fixed, then we can *choose* the third words so that their difference is the difference in the first words after applying G_3. Hence exactly $2^{2n}(2^n(2^n-1)-\theta)$ text pairs are mapped from (a,a,c) to (a,b,a), and, similarly, the same number is mapped to $(a,b,a\oplus b)$. All the remaining pairs are mapped to (a,b,c).

With similar arguments, we compute for any pair of truncated differential classes the exact number of difference pairs mapped between them, depending only on θ. By dividing through by the size of the source truncated differential classes, we get the matrix $[G_3]$ conditional on θ, shown in Fig. 7.

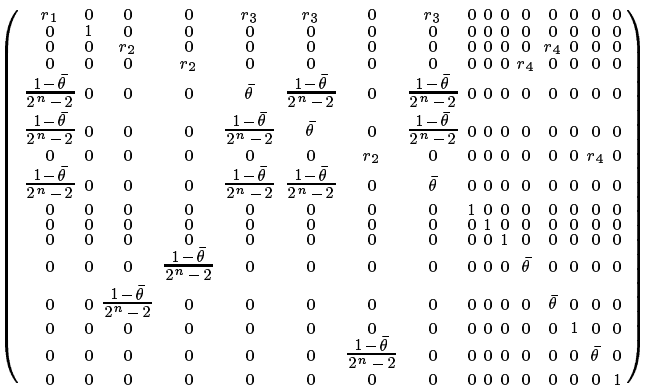

Fig. 7. $[G_3]$ conditional on the number θ of pairs in $\{0,1\}^n \times \{0,1\}^n$ whose exor difference is fixed by G ($0 \le \theta \le 2^{2n}$). For convenience, in this figure we define $\bar{\theta} \equiv \frac{\theta/2^n}{2^n-1}$ and the remainders r_1, r_2, r_3, r_4 are such that the columns sum to one. The row and column order is (a,b,c), $(0,b,c)$, $(a,0,c)$, $(a,b,0)$, (a,a,c), (a,b,a), (a,b,b), $(a,b,a \oplus b)$, $(a,0,0)$, $(0,b,0)$, $(0,0,c)$, $(a,a,0)$, $(a,0,a)$, $(0,b,b)$, (a,a,a), $(0,0,0)$

The expected value of θ is 2^n; this follows since, for example, the probability that a difference in (a,a,a) is mapped to (a,a,a) is $\frac{1}{2^n-1} = \frac{2^n \mathrm{E}[\theta]}{|(a,a,a)|}$. Round functions distributed so that θ is significantly above or below its mean spread the differential classes less uniformly. Knowing $[G_3]$ conditioned on θ lets us compute exactly what cipher weaknesses this may imply. Letting $n_i = |\{x \in \{0,1\}^n | G(x) = x \oplus i\}|$, for $i \in \{0,1\}^n$, we see that $\theta/2 = \sum_{i \in \{0,1\}^n} \binom{n_i}{2}$. Some intensive calculation using indicator variables and taking advantage of the many symmetries among the n_i gives $\mathrm{Var}[\theta] = 2^{n+1}\left(1 + \frac{1}{2^n-3}\right)$. This is a tight enough distribution that weak key classes seem unlikely when the G boxes are random permutations; we'll examine this further for the actual G boxes of full Skipjack in Sect. 7.2.

5 Four-Word Skipjack

Recall the specification of full four-word Skipjack from Fig. 1. Assume the G boxes are uniformly random permutations $\{0,1\}^n \to \{0,1\}^n$, chosen independently for each round; we'll consider the actual 16-bit G boxes in Sect. 7.2.

Corresponding to the 67 linear subspaces of $GF(2)^4$, we consider 67 different truncated differential classes for four n-bit words, listed in Fig. 8. A truncated differential represents the set of pairs in $\{0,1\}^{4n} \times \{0,1\}^{4n}$ where the wordwise exor differences satisfy a given pattern. For example, the class $(0,b,c,b \oplus c)$ contains $\big(w, w \oplus (0,b',c',b' \oplus c')\big)$ for all $w \in \{0,1\}^{4n}$ and all $b', c' \in \{0,1\}^n$ with $b', c' \neq 0$ and $b' \neq c'$. Every pair of texts belongs to exactly one truncated differential class.

(a,b,c,d);

$(0,b,c,d), (a,0,c,d), (a,b,0,d), (a,b,c,0), (a,a,c,d), (a,b,a,d), (a,b,c,a), (a,b,b,d), (a,b,c,b),$
$(a,b,c,c), (a,b,a\oplus b,d), (a,b,c,a\oplus b), (a,b,c,a\oplus c), (a,b,c,b\oplus c), (a,b,c,a\oplus b\oplus c)$;

$(0,0,c,d), (0,b,0,d), (0,b,c,0), (a,0,0,d), (a,0,c,0), (a,b,0,0),$
$(a,a,a,d), (a,a,c,a), (a,b,a,a), (a,b,b,b), (a,a,0,d), (a,a,c,0), (a,0,a,d), (a,b,a,0),$
$(a,0,c,a), (a,b,0,a), (0,b,b,d), (a,b,b,0), (0,b,c,b), (a,b,0,b), (0,b,c,c), (a,0,c,c),$
$(0,b,c,b\oplus c), (a,0,c,a\oplus c), (a,b,0,a\oplus b), (a,b,a\oplus b,0), (a,a,c,c), (a,b,a,b), (a,b,b,a)$;

$(a,a,c,a\oplus c), (a,b,a,a\oplus b), (a,b,a\oplus b,a), (a,b,b,a\oplus b), (a,b,a\oplus b,b), (a,a\oplus c,c,c),$
$(0,0,0,d), (0,0,c,0), (0,b,0,0), (a,0,0,0), (0,b,b,b), (a,0,a,a), (a,a,0,a), (a,a,a,0),$
$(a,a,0,0), (a,0,a,0), (a,0,0,a), (0,b,b,0), (0,b,0,b), (0,0,c,c), (a,a,a,a)$;

$(0,0,0,0)$

Fig. 8. There are 67 different truncated differential classes for four n-bit words, listed here in the basis order we use for our transition matrices. There are $2^{4n}(2^n-1)(2^n-2)(2^n-4)(2^n-8)$ ordered difference pairs of the first type, $2^{4n}(2^n-1)(2^n-2)(2^n-4)$ of each of the next 15 types, $2^{4n}(2^n-1)(2^n-2)$ of each of the next 35 types, $2^{4n}(2^n-1)$ of each of the next 15 types, and, of course, 2^{4n} trivial difference pairs

Figure 9 shows the truncated differential transition matrices $[A]$ and $[B]$ for the A and B rounds, computed as in Sect. 4. The entry in column j and row i is the probability that a truncated differential of type j, as ordered in Fig. 8, is brought to one of type i. For example, the entry in the eighth row and first column indicates that with probability $\frac{1}{2^{16}-1}$ a difference in (a,b,c,d) is taken to (a,b,c,a). The columns sum to one, and the random distribution is an eigenvector.

Composing these matrices $([B]^8[A]^8)^2$ gives the truncated differential transition matrix for the entire 32-round Skipjack cipher. We used Mathematica to calculate the matrix product to 256 decimal digit precision for $n=16$. As a sanity check we also verified the twenty-four round impossible differential through $[B]^4[A]^8[B]^8[A]^4$ [1], which in fact continues to give a good, improbable differential through the last round.

The matrix for the full 32-round Skipjack is too large to show here, but it appears to be quite close to the matrix obtained from a random permutation, for which any nonzero differential is sent to a random nonzero differential. To quantify the "closeness" between the two transition distributions, Fig. 10 shows the columnwise total variation distances. An optimal distinguishing algorithm given the truncated differential equivalence classes of exactly *one* pair of plaintexts and their encryptions has distinguishing advantage less than 1.5×2^{-79}, and is best for plaintext differences in $(0,b,c,0)$ or $(0,b,b,0)$ (the 19th and 63rd columns in the figure). Unfortunately, dependencies invalidate the transition matrices as soon as more than two encryptions are sent through the same keyed cipher. ("Impossible" differentials remain valid, however, and for small numbers of dif-

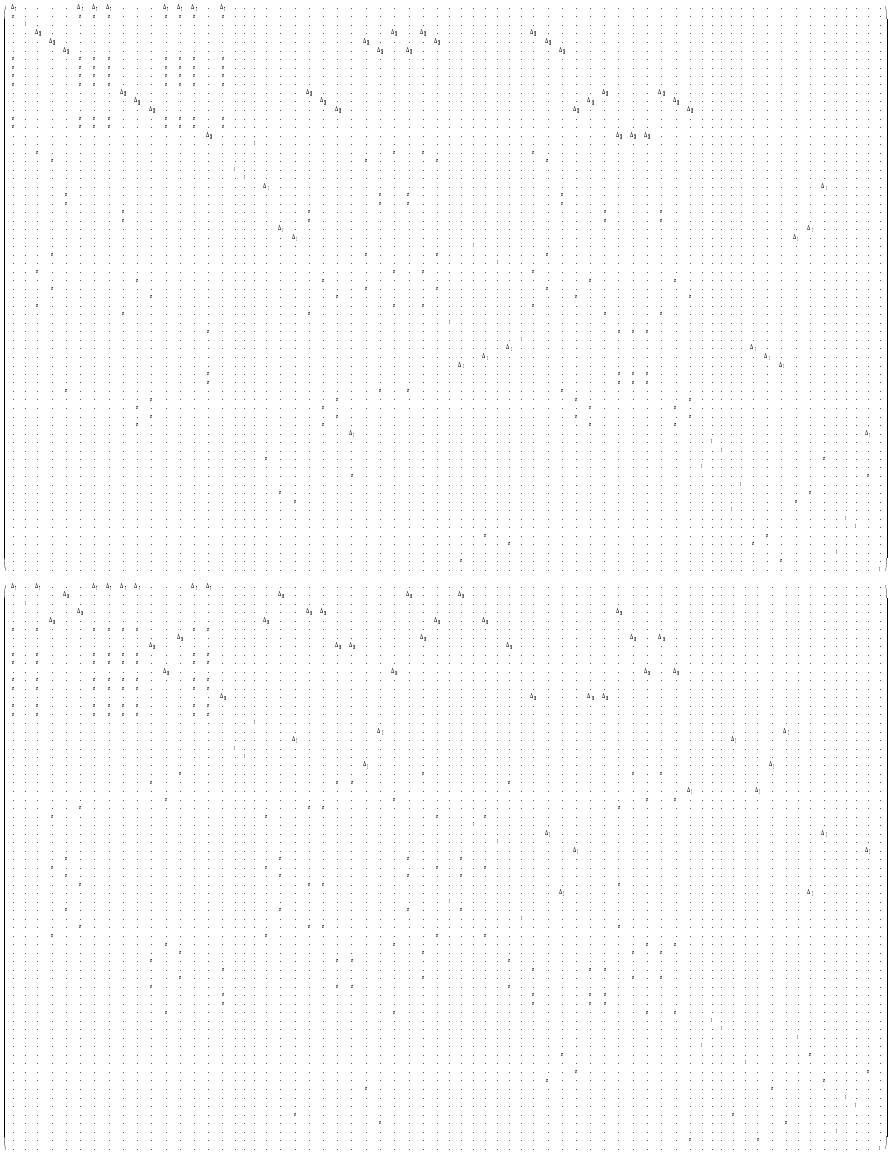

Fig. 9. The truncated differential matrices for the A (top) and the B rounds. $\Delta_j \equiv 1 - \frac{j}{2^n - 1}$ and $x \equiv \frac{1}{2^n - 1}$, while 0 entries are written "." to expose the structure

ferentials the transition probabilities are probably still approximately accurate.) Our model requires the independence of each differential, equivalent to re-keying the cipher for every pair of encryptions. We turn now to some statistics in order to bound an attacker's advantage using *many* truncated differentials with this independence assumption.

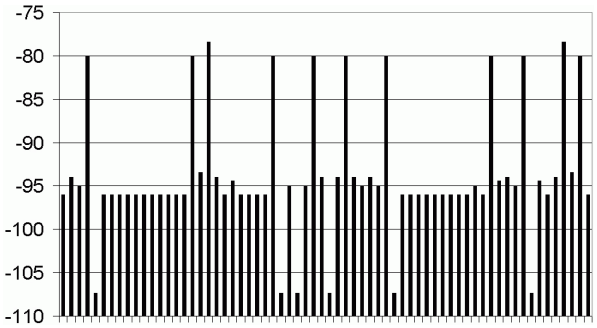

Fig. 10. The total variation norm distance, shown here on a $\log_2$ scale, between columns of the Skipjack truncated differential transition matrix and those for a truly random permutation is always less than 1.5×2^{-79}. The column order is given in Fig. 8

6 Statistical Bounds on Distinguishing Multiple Multinomial Distributions

Definition 1 (Decision Multiple Multinomial Game). *Let SJ be m' different multinomial distributions over m categories, with probabilities $p_1^{(i)}, \ldots, p_m^{(i)}$, $i = 1, \ldots, m'$. Let R be a multinomial distribution over the same m categories, with probabilities $p_1, p_2, \ldots, p_m$, $\sum_{i=1}^{m} p_i = 1$. With equal probabilities $\frac{1}{2}$, take N_i independent samples over the ith multinomial distribution of SJ, or take N_i samples over R, $i = 1, \ldots, m'$; $\sum_{i=1}^{m'} N_i = N$. The results of the samples are $n_1^{(i)}, n_2^{(i)}, \ldots, n_m^{(i)}$, $\sum_{j=1}^{m} n_j^{(i)} = N_i$, $i = 1, \ldots, m'$, denoted individually $S^{(i)}$ or collectively S. An adversary $\mathcal{A}$ has advantage ϵ in this game if $\mathcal{A}$ chooses the correct multiple multinomial distribution when given S with probability at least $\frac{1}{2}(1 + \epsilon)$.*

We note that the problem of distinguishing Skipjack from an ideal cipher, given the truncated differential classes of many known text pairs, is exactly given by the Decision Multiple Multinomial Game. The multinomial distributions SJ for Skipjack are given by $p_j^{(i)} = M_{ji}$, where $M = ([B]^8[A]^8)^2$ and R is the multinomial distribution for an ideal cipher. We turn now to the problem of giving good bounds on the maximum advantage of any attacker at the Decision

Multiple Multinomial Game. This will provide the mathematical tools needed to analyze cryptanalysis with multiple truncated differentials. Those readers uninterested in the statistical details may safely skip ahead to Theorem 5 at the end of this section.

Assume $p_1, p_2, \ldots, p_m > 0$, $p_1^{(i)}, p_2^{(i)}, \ldots, p_m^{(i)} > 0$, $i = 1, \ldots, m'$, so the ratio $\frac{\Pr[S|SJ]}{\Pr[S|R]}$ is well-defined. By Neyman's Lemma, the optimum algorithm for the decision multiple multinomial game decides SJ if this ratio is greater than one, and R if the ratio is less than one. Since

$$\Pr[SJ|S] - \frac{1}{2} = \frac{\Pr[S|SJ]\Pr[SJ]}{\Pr[S|SJ]\Pr[SJ] + \Pr[S|R]\Pr[R]} - \frac{1}{2} = \frac{1}{2}\frac{\frac{\Pr[S|SJ]}{\Pr[S|R]} - 1}{\frac{\Pr[S|SJ]}{\Pr[S|R]} + 1} \quad , \tag{1}$$

this is equivalent to deciding based on the sign of $\Pr[SJ|S] - \frac{1}{2}$.

Lemma 2. *The advantage of an optimal algorithm for the Decision Multiple Multinomial Game is* $\epsilon = 2\,\mathrm{E}_S\left|\Pr[SJ|S] - \frac{1}{2}\right|$.

Proof. Let $Q = \{S : \Pr[SJ|S] \geq \frac{1}{2}\}$.

$$\begin{aligned} \Pr[\text{correct}] &= \overbrace{\Pr[SJ]\sum_S \Pr[S|SJ]\,\chi_Q(S)}^{\text{prob. correctly guesses } SJ} + \overbrace{\Pr[R]\sum_S \Pr[S|R]\,(1-\chi_Q(S))}^{\text{prob. correctly guesses } R} \\ &= \sum_S \Pr[S]\left(\Pr[SJ|S]\chi_Q(S) + (1-\Pr[SJ|S])(1-\chi_Q(S))\right) \quad , \end{aligned} \tag{2}$$

since $\Pr[S] - \Pr[S, SJ] = \Pr[S, R]$. Now $\Pr[SJ|S] = \frac{1}{2} + |\Pr[SJ|S] - \frac{1}{2}|$ when $\chi_Q(S) = 1$ and $1 - \Pr[SJ|S] = \frac{1}{2} + |\Pr[SJ|S] - \frac{1}{2}|$ when $\chi_Q(S) = 0$, so

$$\begin{aligned} \Pr[\text{correct}] &= \sum_S \Pr[S]\left(\frac{1}{2} + \left|\Pr[SJ|S] - \frac{1}{2}\right|\right)\left(\chi_Q(S) + (1-\chi_Q(S))\right) \\ &= \frac{1}{2} + \mathrm{E}_S\left|\Pr[SJ|S] - \frac{1}{2}\right| . \end{aligned} \tag{3}$$

□

Lemma 3. $|\ln x| \geq 2\,|\frac{x-1}{x+1}|$, *with equality only at* $x = 1$.

Proof. $\lim_{x\to 1} \ln x\,\frac{x+1}{x-1} = 2$, by l'Hôpital's rule. Differentiating the ratio gives the equivalent condition that $\frac{1}{x} + \ln\left(\frac{1}{x+1} - \frac{1}{x-1}\right) > 0$. This simplifies to a sign condition on $x - 2\ln x - \frac{1}{x}$, which another differentiation directly proves. □

Proposition 4. *The advantage of an optimal algorithm for the Decision Multiple Multinomial Game is bounded by*

$$\begin{aligned}
&\frac{1}{2}\,\mathrm{E}_S\left|\ln\left(\frac{\Pr[S|SJ]}{\Pr[S|R]}\right)\right| \\
&\quad\le \frac{1}{4}\sqrt{\sum_{i=1}^{m'} N_i\Big(\sum_{j=1}^{m}{\epsilon_j^{(i)}}^2 p_j^{(i)} - \big(\sum_{j=1}^{m}\epsilon_j^{(i)} p_j^{(i)}\big)^2\Big) + \Big(\sum_{i=1}^{m'} N_i \sum_{j=1}^{m}\epsilon_j^{(i)} p_j^{(i)}\Big)^2} \\
&\qquad + \frac{1}{4}\sqrt{\sum_{i=1}^{m'} N_i\Big(\sum_{j=1}^{m}{\epsilon_j^{(i)}}^2 p_j - \big(\sum_{j=1}^{m}\epsilon_j^{(i)} p_j\big)^2\Big) + \Big(\sum_{i=1}^{m'} N_i \sum_{j=1}^{m}\epsilon_j^{(i)} p_j\Big)^2}\ , \qquad (4)
\end{aligned}$$

where $\epsilon_j^{(i)} = \ln\frac{p_j^{(i)}}{p_j}$.

Proof. The bound on the advantage

$$\epsilon \le \frac{1}{2}\,\mathrm{E}_S\left|\ln\left(\frac{\Pr[S|SJ]}{\Pr[S|R]}\right)\right| \qquad (5)$$

follows directly from Lemmas 2 and 3. We need to compute the bound on $\mathrm{E}_S\left|\ln\left(\frac{\Pr[S|SJ]}{\Pr[S|R]}\right)\right|$.

Since

$$\Pr[S|SJ] = \prod_{i=1}^{m'}\Pr[S^{(i)}|SJ] = \prod_{i=1}^{m'} N_i! \prod_{j=1}^{m}(p_j^{(i)})^{n_j^{(i)}}/n_j^{(i)}!\ , \qquad (6)$$

$$\Pr[S|R] = \prod_{i=1}^{m'}\Pr[S^{(i)}|R] = \prod_{i=1}^{m'} N_i! \prod_{j=1}^{m} p_j^{n_j^{(i)}}/n_j^{(i)}!\ , \qquad (7)$$

the log of their ratio is

$$\ln\frac{\Pr[S|SJ]}{\Pr[S|R]} = \sum_{i=1}^{m'}\sum_{j=1}^{m}\epsilon_j^{(i)} n_j^{(i)}\ . \qquad (8)$$

Let $\Delta_k^{(i)} = \sum_{j=1}^{k}\epsilon_j^{(i)} n_j^{(i)}$ and $\Delta^{(i)} = \Delta_m^{(i)}$. Take the multinomial probabilities to be $q_j^{(i)}$, either $p_j^{(i)}$ from SJ or p_j from R. Then $\mathrm{E}\Delta^{(i)} = N_i\sum_{j=1}^{m}\epsilon_j^{(i)} q_j^{(i)}$. We remark that $\mathrm{Var}(n_j^{(i)}) = N_i q_j^{(i)}(1-q_j^{(i)})$ and $\mathrm{Cov}(n_j^{(i)}, n_k^{(i)}) = -N_i q_j^{(i)} q_k^{(i)}$, $j \ne k$. For $k > 1$,

$$\begin{aligned}
\mathrm{Var}\Delta_k^{(i)} &= \mathrm{Var}(\epsilon_k n_k) + 2\,\mathrm{Cov}(\epsilon_k n_k, \Delta_{k-1}) + \mathrm{Var}\Delta_{k-1} \\
&= \epsilon_k^2 N_i q_k(1-q_k) - 2\epsilon_k N_i q_k \sum_{j=1}^{k-1}\epsilon_j q_j + \mathrm{Var}\Delta_{k-1}\ , \qquad (9)
\end{aligned}$$

where we have begun to suppress excessive i superscripts. By induction,

$$\begin{aligned}\frac{\mathrm{Var}\Delta^{(i)}}{N_i} &= \sum_{j=1}^{m} \epsilon_j^2 q_j(1-q_j) - 2\sum_{j<k} \epsilon_j \epsilon_k q_j q_k \\ &= \sum_{j=1}^{m} \epsilon_j^2 q_j - \left(\sum_{j=1}^{m} \epsilon_j q_j\right)^2 .\end{aligned} \tag{10}$$

Note that $\mathrm{E}\Delta = \sum_{i=1}^{m'} \mathrm{E}\Delta^{(i)}$ and $\mathrm{Var}\Delta = \sum_{i=1}^{m'} \mathrm{Var}\Delta^{(i)}$ since $\mathrm{Cov}(\Delta^{(i)}, \Delta^{(i')}) = 0$ for $i \neq i'$ by independence. Now for a random variable X, $\mathrm{Var}X = \mathrm{E}X^2 - (\mathrm{E}X)^2 \geq 0$, so $\mathrm{E}X = \sqrt{\mathrm{E}X^2 - \mathrm{Var}X} \leq \sqrt{\mathrm{E}X^2}$. Hence,

$$\begin{aligned}\mathrm{E}\left(\Big|\sum_{i=1}^{m'} \Delta^{(i)}\Big|\right) &= \mathrm{E}\left(\sqrt{\big(\sum_{i=1}^{m'} \Delta^{(i)}\big)^2}\right) \\ &\leq \sqrt{\mathrm{E}\Big(\big(\sum_{i=1}^{m'} \Delta^{(i)}\big)^2\Big)} \\ &= \sqrt{\mathrm{Var}\Big(\sum_{i=1}^{m'} \Delta^{(i)}\Big) + \Big(\mathrm{E}\sum_{i=1}^{m'} \Delta^{(i)}\Big)^2} \\ &= \sqrt{\sum_{i=1}^{m'} N_i \Big(\sum_{j=1}^{m} {\epsilon_j^{(i)}}^2 q_j^{(i)} - \big(\sum_{j=1}^{m} \epsilon_j^{(i)} q_j^{(i)}\big)^2\Big) + \Big(\sum_{i=1}^{m'} N_i \sum_{j=1}^{m} \epsilon_j^{(i)} q_j^{(i)}\Big)^2} .\end{aligned} \tag{11}$$

The result follows from averaging the above bound over the $q_j^{(i)}$ being $p_j^{(i)}$ from SJ or p_j from R. □

Proposition 4 gives a bound for distinguishing the multiple multinomials of a block cipher's truncated differentials from those of a random permutation's truncated differentials. From the transition matrix calculated in Sect. 5 we obtain a bound for Skipjack, with G boxes given by independent random permutations.

Theorem 5. *An adversary who sees the truncated differential classes for one pair of plaintexts and corresponding ciphertexts from each of 2^{128} independently keyed oracle ciphers, such that each input truncated differential class appears exactly the size of that class times (sizes given in Fig. 8), has advantage less than 7.63103×10^{-6} in distinguishing whether the oracle ciphers are truly random permutations or Skipjack (with G boxes independent random permutations).*

7 Truncated Differentials for Full Skipjack

7.1 Best Differentials

Table 1 shows the longest impossible differentials starting at each round from 1 to 16. The longest impossible differential goes through 24 rounds, from rounds

5 to 28, inclusive: $(0, b, 0, 0) \not\rightarrow (a, 0, 0, 0)$. Biham et al. use this differential to mount an attack on 31-round Skipjack which is slightly faster than exhaustive key-search [1]. On the same differential path is one impossible differential from rounds 5 to 27, $(0, b, 0, 0) \not\rightarrow (0, 0, 0, d)$, and also four from rounds 5 to 26:

$$(0, b, 0, 0) \not\rightarrow (a, 0, c, 0), \qquad (0, b, 0, 0) \not\rightarrow (0, 0, c, 0),$$
$$(0, b, 0, 0) \not\rightarrow (a, 0, 0, 0), \qquad (0, b, 0, 0) \not\rightarrow (a, 0, a, 0) \ .$$

While [1] found these latter four impossible differentials, they did not find the impossible differential from rounds 5 to 26. Additionally, the impossible differentials they list from rounds 5 to 27, $(0, b, 0, 0) \not\rightarrow (a, 0, 0, 0)$ and $(0, b, 0, 0) \not\rightarrow (0, b, 0, 0)$, are *incorrect*, as is another impossible differential they list from rounds 5 to 26, $(0, b, 0, 0) \not\rightarrow (0, b, 0, 0)$.

Along the same differential path are several other long impossible differentials, one from rounds 6 to 28 – $(0, 0, c, 0) \not\rightarrow (a, 0, 0, 0)$ – four from rounds 7 to 28, and seven from rounds 5 to 25. Other long impossible differentials include $(0, b, 0, 0) \not\rightarrow (a, 0, 0, 0)$ from rounds 4 to 24, and $(0, b, 0, 0) \not\rightarrow (a, 0, 0, 0)$ from rounds 9 to 29.

Table 2 shows the best differentials for distinguishing attacks based on an attacker seeing all 2^{64} possible plaintext encryptions. The distinguishing bound is from Proposition 4, using just the stated differential ($m' = 1$, $m = 2$). We assume that the G boxes are independent random permutations, and that each text pair (2^{128} total) is encrypted independently.

For example, the best differential through all of Skipjack, rounds 1 to 32, is $(a, b, 0, d) \rightarrow (a, b, c, 0)$. An attacker with access to the encryptions of all 2^{64} plaintexts has a distinguishing advantage of at most 7.62980×10^{-6} from considering just this differential. By Theorem 5, the distinguishing advantage an attacker gains from considering *all* truncated differentials is only 7.63103×10^{-6}, so most of that advantage is from just this one differential. In general, it appears that counting other differentials besides the single best one does not often significantly increase the distinguishing power.

The differential $(0, b, c, 0) \rightarrow (a, 0, 0, d)$ through rounds 2 to 31 has probability about $(1 - 2^{-32})2^{-32}$, compared to a probability of about 2^{-32} for a random permutation. The $2^{64}(2^{16} - 1)(2^{16} - 2) \approx 2^{96}$ difference pairs of this type can conceivably distinguish between Skipjack and a random permutation using this differential, since the difference in the expected number of pairs satisfying the differential for Skipjack versus for a random permutation is about 2^{32}, and the standard deviation of the number of satisfying pairs is also about $\sqrt{2^{64}} = 2^{32}$.

7.2 Existence of Weak Key Classes

As shown in Sect. 4, G boxes giving an unusually low or high θ – the number of n-bit input pairs whose exor difference is fixed by G – can lead to poor mixing of the truncated differential classes. Figure 11 shows the distribution of θ for the actual 16-bit G boxes of official Skipjack. θ is quite tightly distributed and it seems unlikely that there exist *any* weak keys or key classes for this level of

Table 1. Longest impossible differentials starting at each round from 1 to 16

Start	End	Impossible differentials
1	14	$(0,b,c,0) \nrightarrow (a,0,0,0), (0,0,0,d) \nrightarrow (a,0,0,0), (0,0,c,0) \nrightarrow (a,0,0,0),$ $(0,b,0,0) \nrightarrow (a,0,0,0), (0,b,b,0) \nrightarrow (a,0,0,0)$
2	14	$(0,b,c,d) \nrightarrow (a,0,0,0), (0,0,c,d) \nrightarrow (a,0,0,0), (0,b,0,d) \nrightarrow (a,0,0,0),$ $(0,b,c,0) \nrightarrow (a,0,0,0), (0,b,b,d) \nrightarrow (a,0,0,0), (0,b,c,b) \nrightarrow (a,0,0,0),$ $(0,b,c,c) \nrightarrow (a,0,0,0), (0,b,c,b \oplus c) \nrightarrow (a,0,0,0), (0,0,0,d) \nrightarrow (a,0,0,0),$ $(0,0,c,0) \nrightarrow (a,0,0,0), (0,b,0,0) \nrightarrow (a,0,0,0), (a,0,0,0) \nrightarrow (a,0,0,0),$ $(0,b,b,b) \nrightarrow (a,0,0,0), (0,b,b,0) \nrightarrow (a,0,0,0), (0,b,0,b) \nrightarrow (a,0,0,0),$ $(0,0,c,c) \nrightarrow (a,0,0,0)$
3	21	$(0,b,0,0) \nrightarrow (a,a,0,0)$
4	24	$(0,b,0,0) \nrightarrow (a,0,0,0)$
5	28	$(0,b,0,0) \nrightarrow (a,0,0,0)$
6	28	$(0,0,c,0) \nrightarrow (a,0,0,0)$
7	28	$(0,b,0,d) \nrightarrow (a,0,0,0), (0,0,0,d) \nrightarrow (a,0,0,0), (0,b,0,0) \nrightarrow (a,0,0,0),$ $(0,b,0,b) \nrightarrow (a,0,0,0)$
8	28	$(a,0,c,0) \nrightarrow (a,0,0,0), (0,0,c,0) \nrightarrow (a,0,0,0), (0,b,0,0) \nrightarrow (a,0,0,0),$ $(a,0,0,0) \nrightarrow (a,0,0,0), (a,0,a,0) \nrightarrow (a,0,0,0)$
9	29	$(0,b,0,0) \nrightarrow (a,0,0,0)$
10	29	$(0,0,c,0) \nrightarrow (a,0,0,0)$
11	29	$(a,a,0,d) \nrightarrow (a,0,0,0), (0,0,0,d) \nrightarrow (a,0,0,0), (a,a,0,a) \nrightarrow (a,0,0,0),$ $(a,a,0,0) \nrightarrow (a,0,0,0)$
12	30	$(a,a,0,0) \nrightarrow (a,0,0,0)$
13	30	$(0,b,0,0) \nrightarrow (a,0,0,0)$
14	30	$(0,0,c,0) \nrightarrow (a,0,0,0), (a,a,0,0) \nrightarrow (a,0,0,0)$
15	30	$(a,a,0,d) \nrightarrow (a,0,0,0), (0,0,0,d) \nrightarrow (a,0,0,0), (0,b,0,0) \nrightarrow (a,0,0,0),$ $(a,a,0,a) \nrightarrow (a,0,0,0), (a,a,0,0) \nrightarrow (a,0,0,0)$
16	30	$(a,b,0,0) \nrightarrow (a,0,0,0), (0,0,c,0) \nrightarrow (a,0,0,0), (0,b,0,0) \nrightarrow (a,0,0,0),$ $(a,0,0,0) \nrightarrow (a,0,0,0), (a,a,0,0) \nrightarrow (a,0,0,0)$

Table 2. Best differentials starting from rounds 1–5 (rows) and ending in rounds 28–32 (columns). Beneath each differential is the full-codebook, distinguishing-advantage bound, and also one minus the ratio between the differential probability for Skipjack and that for a random permutation on $\{0,1\}^{64}$

	28	29	30	31	32
1	$(a,b,0,d) \to (a,0,0,0)$ $.56, 2^{-32}$	$(a,b,0,d) \to (0,b,c,0)$ $.0020, 2^{-48}$	$(a,b,0,d) \to (0,0,c,d)$ $.0020, 2^{-48}$	$(a,b,0,d) \to (a,0,0,d)$ $.0020, 2^{-48}$	$(a,b,0,d) \to (a,b,c,0)$ $7.63 \cdot 10^{-6}, 2^{-64}$
2	$(0,b,c,0) \to (a,0,0,0)$ $16385, 2^{-16}$	$(0,b,c,0) \to (0,b,c,0)$ $.56, 2^{-32}$	$(0,b,c,0) \to (0,0,c,d)$ $56, 2^{-32}$	$(0,b,c,0) \to (a,0,0,d)$ $.56, 2^{-32}$	$(0,b,c,0) \to (a,b,c,0)$ $.0020, 2^{-48}$
3	$(0,0,c,d) \to (a,0,0,0)$ $16385, 2^{-16}$	$(0,0,c,d) \to (0,b,c,0)$ $.56, 2^{-32}$	$(0,0,c,d) \to (0,0,c,d)$ $.56, 2^{-32}$	$(0,0,c,d) \to (a,0,0,d)$ $.56, 2^{-32}$	$(0,0,c,d) \to (a,b,c,0)$ $.0020, 2^{-48}$
4	$(a,0,0,d) \to (a,0,0,0)$ $16385, 2^{-16}$	$(a,0,0,d) \to (0,b,c,0)$ $..56, 2^{-32}$	$(a,0,0,d) \to (0,0,c,d)$ $.56, 2^{-32}$	$(a,0,0,d) \to (a,0,0,d)$ $.56, 2^{-32}$	$(a,0,0,d) \to (a,b,c,0)$ $.0020, 2^{-48}$
5	$(0,b,0,0) \nrightarrow (a,0,0,0)$ $\infty, 1$	$(0,b,0,0) \to (0,b,c,0)$ $16385, 2^{-16}$	$(0,b,0,0) \to (0,0,c,d)$ $16385, 2^{-16}$	$(0,b,0,0) \to (a,0,0,d)$ $16385, 2^{-16}$	$(0,b,0,0) \to (a,b,c,0)$ $.56, 2^{-32}$

attack. With our independence assumptions, a successful attack will not be able to assume that the G boxes are random permutations, nor that they are random permutations with a given θ distribution, but will need to look more closely at the exact distribution of G-box permutations.

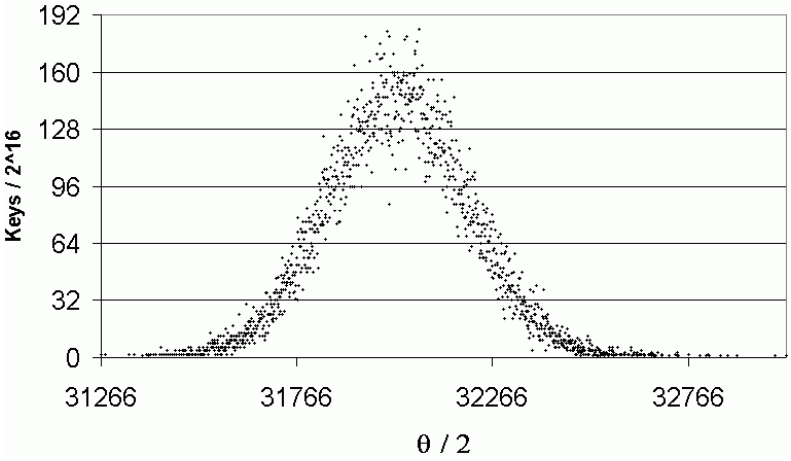

Fig. 11. The distribution of θ, the number of fixed differences, is tighter for Skipjack's true G boxes than for random permutations $\{0,1\}^{16} \to \{0,1\}^{16}$. This histogram shows for each θ value the number of keys which give that θ (always a multiple of 2^{16})

7.3 Interpretation

Our calculations were intended to give us insight into the power of truncated differential attacks on Skipjack. However, to enable explicit calculation, we were forced to abstract away a few features of Skipjack in our theoretical model. Our theoretical model is still fairly close to the real Skipjack, but the differences could potentially affect the correctness of our figures. Consequently, our results must be interpreted with care.

We will try to enumerate to all potential ways that our abstract model could fail to accurately represent the behavior of the real Skipjack, and so Theorem 5 could overestimate Skipjack's security.

1. We only consider standard truncated differential attacks. Obviously, other attacks – exact differentials, higher-order differentials, integrals, boomerang attacks, the Yoyo game, and so on – might have lower complexity.
2. We consider only the possibility of building distinguishers, and ignore methods for guessing key material at the outer rounds. In particular, our model does not account for the possibility of attacks that guess, say, the 32 key bits determining the G permutation used in the last round. Consequently, any

attack that succeeds in distinguishing Skipjack for r rounds in our model can likely be extended to a key-recovery attack on the real Skipjack for $r + r'$ rounds for some small r' (the best attack on Skipjack currently known achieves $r' = 7$ [1]).

3. Our model ignores the Skipjack key schedule. In our model we assume that round keys are independent, while in the real Skipjack round subkeys have a simple relationship to each other. For instance, G boxes separated by a multiple of five rounds are identical. This might make our truncated differential distinguishing bound too optimistic, and – since the first two round subkeys match the last two round subkeys – could also aid key-recovery attacks.

4. Our model ignores the internal structure of the G box, and simply treats it as a random permutation. There might plausibly be some way to take advantage of the internal structure of the G box to build a more efficient attack. For instance, it is known that there are four differential characteristics with probability $2^{-10.42}$ for the Skipjack G box [3], while for a random permutation the chances of encountering such characteristics is remote. As another example, one can see that the parameter θ for the real Skipjack G boxes differs what one would expect for a random permutation: the difference is small but statistically significant (see Fig. 11). It is conceivable that it might be possible to exploit these or other properties of the real G box somehow, but we do not see any obvious way to do so. More convincing is that the strongest existing attacks on Skipjack all treat the G box as though it were a random permutation, and ignore its internal structure.

5. Our model is overly generous with the text pairs given to the attacker. We are trying to model the case of an attacker who is given the entire codebook for Skipjack, i.e., the encryption of all 2^{64} possible plaintexts. Such an attacker can obtain $2^{64} \times (2^{64} - 1)/2 \approx 2^{127}$ text pairs, and this is obviously an upper bound on the number of pairs available to the attacker. If Skipjack is secure against all truncated differential attacks using 2^{127} pairs, then it will also be secure against all truncated differential attacks using less than the whole codebook.

   For technical reasons, to rigorously analyze the advantage of such an attack, we needed to assume that each pair behaves independently. This assumption is embodied by modeling each pair as coming from an independently keyed instantiation of the cipher. This seems to be a clean way to model the heuristic assumptions made in previous work on truncated differential attacks on Skipjack. However, this assumption may be too generous to the attacker. The assumption implies that each of the 2^{127} pairs gives new information to the attacker. However, when we obtain 2^{127} text pairs from 2^{127} texts, the pairs definitely do not behave independently, and some are redundant. For instance, if c, c', c'' represent three ciphertexts with the differences $c \oplus c'$ and $c' \oplus c''$ both in the truncated class $(a, 0, 0, 0)$, then the difference $c \oplus c''$ is surely in the same class, and hence the third pair gives no new information.

The independence issue may well be the most troubling aspect of our model. Characterizing the appropriateness of this assumption is an interesting problem for future research.

References

1. Biham, E., Biryukov, A., Shamir, A.: Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. Technical report CS0947, Computer Science Department, Technion – Israel Institute of Technology (1998)
2. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology, **4** (1991) 3–72
3. Biham, E., Biryukov, A., Dunkelman, O., Richardson, E., Shamir, A.: Initial observations on Skipjack: Cryptanalysis of Skipjack-3XOR. Technical report CS0946, Computer Science Department, Technion – Israel Institute of Technology (1998)
4. Granboulan, L.: Flaws in differential cryptanalysis of Skipjack. Fast Software Encryption, 8th International Workshop (2001)
5. Knudsen, L.: Truncated and higher-order differentials. Fast Software Encryption, 2nd International Workshop Proceedings, Springer LNCS 1008 (1995)
6. Knudsen, L., Robshaw, M., Wagner, D.: Truncated differentials and Skipjack. Advances in Cryptology – CRYPTO'99, Springer LNCS 1666 (1999)
7. Lai, X., Massey, J., Murphy, S.: Markov ciphers and differential cryptanalysis. Advances in Cryptology – EUROCRYPT'91, Springer LNCS 547 (1991)
8. Skipjack and KEA algorithm specifications, Version 2.0, 29 May 1998. Available from the National Institute of Standards and Technology, `http://csrc.nist.gov/encryption/skipjack/skipjack.pdf`