Related-Key Boomerang and Rectangle Attacks: Theory and Experimental Analysis

Jongsung Kim, Seokhie Hong, Bart Preneel, Eli Biham, Orr Dunkelman, and Nathan Keller

***Abstract*— The related-key differential attack and the boomerang attack are two of the classical techniques in crypt-analysis of block ciphers. In 2004, we introduced the related-key boomerang and related-key rectangle attacks, which allow to enjoy the benefits of these two techniques simultaneously. The new techniques proved to be very powerful, and were used to devise the best known attacks against numerous block ciphers, culminating with the first attack on the full AES presented in 2009 and a practical-time attack on KASUMI (the cipher used in GSM and 3G telephony) presented in 2010.**

While the claimed applications of the related-key boomerang/rectangle technique are significant, most of them have a major drawback: due to the extremely high complexity of the attacks, their validity cannot be verified experimentally. Together with the lack of rigorous justification of the probabilistic assumptions underlying the technique, it was claimed that these assumptions cannot be relied upon, and thus, attacks using the related-key boomerang/rectangle technique are not legitimate. These claims were formalized in a recent paper by Murphy [32] who presented scenarios in which the probabilistic assumptions fail, and questioned their validity.

In this paper we present a rigorous treatment of the related-key boomerang/rectangle technique. In the first part of the paper, we devise optimal algorithms for the related-key boomerang/rectangle distinguishers using the Logarithmic Likelihood Ratio statistics. We study the exact independence assumptions the attacks rely upon, and compute the success probability of the attacks under these independence assumptions.

In the second part of the paper, we address the claims against the validity of the related-key boomerang/rectangle technique by an extensive experimental analysis. We consider a specific case — the block cipher KASUMI — and perform an experimental verifications (with more than 2^{48} encryptions) of a related-key boomerang distinguisher against it. The analysis shows that in all attacks, the overall probability of the distinguisher (when averaged over different choices of plaintexts and keys) is close to the theoretically predicted probability. However, it seems that the probability depends on the key, such that for some portion of the keys, the distinguisher holds with a higher probability than expected, while for the rest of the keys, the distinguisher fails completely. We conclude that the probability assumptions underlying the technique make sense in real-life ciphers, and thus, related-key boomerang/rectangle attacks on block ciphers are valid in general. On the other hand, due to the dependence of the probabilities on the key, it is important to verify the validity of the attack experimentally whenever possible in order to measure its success probability.

***Index Terms*— Related-key Boomerang Attack, Related-Key Rectangle Attack, Experimental Analysis, KASUMI.**

I. INTRODUCTION

THE *related-key differential* attack, introduced by Kelsey et al. [23] in 1996, is an extension of differential crypt-analysis [5] in which it is assumed that the adversary has control over the key difference, along with the control over the plaintext/ciphertext differences. Since its introduction, the related-key differential attack was used to break reduced-round variants of various block ciphers, including a practical-time attack on 10-round AES-256 [15]. Moreover, although an attack model in which the adversary has control over the key difference may seem unrealistic, a related-key differential attack on the block cipher TEA [42] was used to devise a practical attack on Microsoft's Xbox architecture [43].[1]

The *boomerang* attack, introduced by Wagner [40] in 1999, is a differential-based attack in which the block cipher E is treated as a cascade: $E = E_1 \circ E_0$, and differentials of E_0 and E_1 are combined into a distinguisher for the entire cipher E in an adaptive chosen plaintext and ciphertext process. The boomerang attack shows that bounding (from above) the probability of differential characteristics through E does not assure immunity of E to differential-type attacks, and the boomerang technique was indeed used to devise practical-time attacks against ciphers which are provably immune to conventional differential attacks, e.g., COCONUT98.

The adaptive chosen plaintext/ciphertext nature of the boomerang attack makes it less realistic in practical scenarios. As a partial remedy of this issue, the attack was transformed into a chosen plaintext variant named the amplified boomerang attack [25] and later renamed as the rectangle attack [7]. The transformation is done by a birthday-paradox argument, which leads to an increase in the data complexity of the attack.

In 2004, Kim et al. [26], and independently, Biham et al. [9], introduced the *related-key boomerang* (RK-boomerang) and *related-key rectangle* (RK-rectangle) attacks – a combination

Manuscript received ?? ?? ??; revised ?? ?? ??. This paper is partially based on the papers [26], [9], [10], [27] which appeared at ACISP 2004, EUROCRYPT 2005, ASIACRYPT 2005 and FSE 2007, respectively.

Jongsung Kim is with the Division of e-Business, Kyungnam University, 449, Wolyoung-dong, Masan, Kyungnam, Korea.

Seokhie Hong is with the Center for Information Security and Technologies(CIST), Korea University, Anam Dong, Sungbuk Gu, Seoul, Korea.

Bart Preneel is with the Department of Electrical Engineering ESAT/SCD-COSIC, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium.

Eli Biham is with Computer Science Department, Technion, Haifa 32000, Israel.

Orr Dunkelman is with Computer Science Department, University of Haifa, 31905 Haifa, Israel.

Nathan Keller is with Faculty of Mathematics and Computer Science, Weizmann Institute of Science, P.O. Box 26, Rehovot 76100, Israel.

[1] In the Xbox architecture, the block cipher TEA was used in a Davies-Meyer mode as a compression function. In such cases, the key difference in the block cipher is transformed into the message difference of the compression function, which indeed can be controlled by the adversary. In general, almost any related-key attack on a block cipher can be converted into a chosen message attack on a compression function based on it.

of the boomerang technique with related-key differentials.[2] It turns out that the combination allows to enjoy the strength of the related-key model twice, by using high-probability related-key differentials in both subciphers, E_0 and E_1. This makes the RK-boomerang/rectangle techniques much more effective than other combined techniques, such as the related-key impossible differential [22] and the related-key differential-linear [11] attacks.

Since its introduction, the RK-boomerang/rectangle technique was used to attack reduced-round variants of various block ciphers (e.g., IDEA, MISTY1, SHACAL-1, SHACAL-2, and XTEA), and even full versions of widely used block ciphers such as AES [14] and KASUMI [19].

In parallel with the increasing popularity of the RK-boomerang/rectangle technique, several researchers raised concerns about its theoretical validity. The main concern is that the technique relies on randomness assumptions which are much stronger than the assumptions relied upon in standard differential attacks (i.e., that the cipher is Markovian, see [6]), and thus they can be inappropriate in real block ciphers. Indeed, while the "Markovity" assumption was treated rigorously and verified experimentally in many practical cases, the exact randomness assumptions underlying the boomerang attack and the RK-boomerang/rectangle attack were never treated rigorously, and in most practical cases, there was no possibility to verify them experimentally, due to the high complexity of the attacks.

These concerns are supported by a paper of Wang et al. [41] published in 2008, that showed that all previously published boomerang and related-key boomerang attacks on SHACAL-1 fail, due to a failure of the randomness assumptions in the specific case of SHACAL-1. In that case, the attacks fail because of *local inconsistency*: while the attacks assume that differential characteristics for different rounds are independent and the probability of their concatenation is the product of their probabilities, it appeared that some characteristics used in subsequent rounds contradict each other, and thus, they never co-occur.[3]

In a recent paper [32], Murphy presented several examples based on a 4-round variant of DES [33] and a 2-round variant of AES [34], in which such local inconsistencies occur in the transition between the two sub-ciphers E_0 and E_1. He concluded that there is no reason to assume that the randomness assumptions underlying the boomerang attack hold in real ciphers, and thus, any boomerang (or RK-boomerang) attack should be viewed extremely skeptically, unless it is verified experimentally.[4]

[2]We note that this paper is written jointly by the two research groups who independently introduced the related-key boomerang/rectangle technique. This explains the term "we" which is used in the abstract when referring to the inventors of the technique.

[3]We note that actually, the flaw detected in [41] is not a special feature of the boomerang attack but rather a failure of the Markovity assumption for the specific type of differential characteristics used in the attacks. However, it still demonstrates the possibility of failure of the randomness assumptions underlying the attack.

[4]We note that the claims of [32] are addressed not only to RK-boomerang/rectangle attacks, but also to the boomerang and the rectangle attacks in the single-key model.

A. Our Contributions – Theoretical Results

In the first part of this paper we present the first rigorous treatment of the (related-key) boomerang and rectangle distinguishers. We devise the optimal distinguishing algorithms using the Logarithmic Likelihood Ratio metric, and compute their success rate. We obtain and prove the following theorem:

Theorem *1:* Let $E = E_1 \circ E_0 : \{0,1\}^n \to \{0,1\}^n$ be a block cipher. Consider encryptions under a secret key K and related-keys whose differences are chosen by the adversary. Let

$$\hat{p} = \max_{\alpha \neq 0, \Delta K_0} \sqrt{\sum_{\beta} \left(\Pr_{P} \left[E_{0,K}(P) \oplus E_{0,K \oplus \Delta K_0}(P \oplus \alpha) = \beta \right] \right)^2},$$

$$\hat{q} = \max_{\delta \neq 0, \Delta K_1} \sqrt{\sum_{\gamma} \left(\Pr_{C} \left[E^{-1}_{1,K}(C) \oplus E^{-1}_{1,K \oplus \Delta K_1}(C \oplus \delta) = \gamma \right] \right)^2}$$

$$= \max_{\delta \neq 0, \Delta K_1} \sqrt{\sum_{\gamma} \left(\Pr_{X} \left[E_{1,K}(X) \oplus E_{1,K \oplus \Delta K_1}(X \oplus \gamma) = \delta \right] \right)^2}$$

where $E_{0,K}(P)$ denotes the partial encryption of P through E_0 under the key K and $E^{-1}_{1,K}(C)$ denotes the partial decryption of C through E_1 under the key K. Let $0 < c < 1$. Under certain independence assumptions between the differentials that will be discussed below, given either

- $4c/(\hat{p}\hat{q})^2$ unique adaptively chosen plaintexts and ciphertexts, or
- $\sqrt{c} \cdot 2^{n/2+2}/\hat{p}\hat{q}$ unique chosen plaintexts,

encrypted under four related-keys of the form $K, K \oplus \Delta K_0, K \oplus \Delta K_1, K \oplus \Delta K_0 \oplus \Delta K_1$,[5] the RK-boomerang/rectangle technique allows to distinguish E from a random permutation. The probability of success of the distinguisher is approximately $1 - e^{-c}/2$ (when $\hat{p}\hat{q}$ is sufficiently high).

We state explicitly the randomness assumptions required for Theorem 1 to hold, and examine their soundness in various scenarios.

After the theoretical treatment, we consider several improvements of the related-key boomerang and rectangle attacks:

1) **The Use of Structures of Keys:** We use structures of keys to overcome a wider range of key schedule algorithms. In ciphers with a nonlinear key schedule, a given key difference may cause many subkey differences, thus interfering with the construction of related-key differentials. Structures of keys can be used to reduce the effects of this event on the differentials.
2) **The Use of Other Relations between the Keys:** While XOR relations are common and inherent to the majority of differential-based related-key attacks, in some cases there are more suitable key relations (either due to the environment of the attack or in order to gain higher probabilities). We show that the proposed attacks are applicable when the XOR relations between the keys

[5]In some cases $\Delta K_0 = \Delta K_1$. In these cases, there are small changes in the analysis, most notably the use of only two related keys.

are replaced with different kinds of relations and discuss which relations induce feasible attacks.

We then compare the RK-boomerang/rectangle attacks with previously proposed related-key techniques. We explore the advantages of the new attacks, and show that in many cases the RK-boomerang/rectangle attacks are significantly more effective than other related-key techniques, even if in the single-key scenario the boomerang and the rectangle attacks are inferior to the respective non-related-key techniques.

B. Our Contributions – Experimental Results

In the second part of this paper, we examine experimentally the validity of the randomness assumptions underlying the RK-boomerang/rectangle attack, in a specific case of a widely used block cipher. As stated above, most of the existing RK-boomerang/rectangle attacks (including all the attacks on reduced-round AES, e.g., [9], [13], [14], [27]) cannot be verified experimentally due to their high data complexity. Even reduced-round variants of these attacks in which the number of rounds is not very small, have too high complexity for being verified. On the other hand, variants with a very small number of rounds, such as the examples studied in [32], are non-representative, since it is clear that the rate of randomness of variants with a very small number of rounds is much smaller than that of the entire ciphers.

In order to obtain representative experimental results, we choose the block cipher KASUMI [38], used in GSM and 3G telephony, and examine various RK-boomerang/rectangle on it. We start with verifying a RK-boomerang distinguisher on 6-round KASUMI (out of the total 8 rounds) in which the probabilities of both differentials are relatively high. This choice follows the intuition that the precision of the randomness assumptions is better when the probability of the differentials is not very low. As we expected, among the 10,000 random keys we sample, the probability of the distinguisher is remarkably close to the probability predicted by Theorem 1.

Then, we experimentally verify our RK-boomerang distinguisher on 7-round KASUMI, which include the distinguishers used in the attacks on the full KASUMI presented in [10], [19]. Among these distinguishers, we check those in which the probability of one of the differentials in the transition round is as low as the probability of a random differential through that round.[6] By the intuition stated above, these are the cases where the validity of the randomness assumptions can be more problematic.

First, we check theoretically, whether there exists a *local inconsistency* in the transition between the differentials (like the inconsistencies presented in [32]). We found that indeed, for many choices of the differentials such inconsistencies exist and lead to failure of the attack.[7]

Then we choose the differentials carefully, such that we cannot detect any more inconsistencies, and perform an experiment checking 215 random quartets of keys, with 2^{39} encrypted plaintexts under each key.[8] We find out that the overall probability of the distinguisher, when averaged over all the $2^{48.7}$ (plaintexts,keys) choices, is very close to the theoretical probability. On the other hand, it appears that the probability depends very much on the key, such that for about $5/6$ of the keys, the distinguisher fails completely, while for $1/6$ of the keys, it holds with an increased probability.

We conclude that in cases where the probabilities of the differentials are not very low, it seems reasonable to assume that the RK-boomerang/rectangle distinguisher holds with the theoretically predicted probability. In cases where some of the probabilities are low, one has to make his best to check that the distinguisher does not have local inconsistencies, and then it is reasonable to assume that it does hold, at least for a significant portion of the keys. However, it is clearly desirable to check the attack experimentally in each specific case, in order to verify its validity and to compute its success probability.

C. The Organization of the Paper

The paper is organized as follows: In Section II we present the related-key boomerang and rectangle attacks and discuss them theoretically. In Section III we present the experimental results on reduced-round variants of KASUMI. Finally, Section IV summarizes the paper.

II. The Related-Key Boomerang and Rectangle Attacks

In this section we introduce the RK-boomerang and the RK-rectangle attacks. We start with a brief description of the boomerang and the rectangle attacks in the single key model. We then introduce and analyze rigorously the RK-boomerang and RK-rectangle attacks. We follow and examine the randomness assumptions used in the attacks. We conclude this section with several generalizations and comparison of the newly proposed attacks with other techniques.

A. Boomerang and Amplified Boomerang (Rectangle) Attacks

The main idea behind the boomerang attack [40] is to use two short differentials with high probabilities instead of one long differential with a low probability. We assume that a block cipher $E:\{0,1\}^n\times\{0,1\}^k\rightarrow\{0,1\}^n$ can be described as a cascade $E=E_1\circ E_0$, such that for E_0 there exists a differential $\alpha\rightarrow\beta$ with probability p, and for E_1 there exists a differential $\gamma\rightarrow\delta$ with probability q.[9]

The distinguisher is based on the following boomerang process:

1) Ask for the encryption of a pair of plaintexts (P_1,P_2) such that $P_1\oplus P_2=\alpha$ and denote the corresponding ciphertexts by (C_1,C_2).

[6]We note that in [19], the probability of the distinguisher on 7-round KASUMI was verified experimentally, and the results were very close to the theoretical prediction. This gives additional evidence to the claim that the randomness assumptions are sound, at least in the cases where the probabilities of the differentials are high, like in the distinguisher checked in [19].

[7]We note that this dependence issue was overlooked in [10], and indeed, the probability of the distinguisher used in [10] is far from the theoretical value used in that paper.

[8]We note that the number of plaintexts for each quartet of keys cannot be smaller, since the theoretical probability of the distinguisher is 2^{-39}.

[9]We note that in the attack, the differentials are used both in the forward (i.e., encryption), and in the backward (i.e., decryption) directions. As the considered differentials are not truncated differentials, the direction does not affect the probability of the differentials.

2) Calculate $C_3 = C_1 \oplus \delta$ and $C_4 = C_2 \oplus \delta$, and ask for the decryption of the pair (C_3, C_4). Denote the corresponding plaintexts by (P_3, P_4).
3) Check whether $P_3 \oplus P_4 = \alpha$.

The boomerang attack uses the first differential ($\alpha \to \beta$) for E_0 with respect to the pairs (P_1, P_2) and (P_3, P_4), and the second differential ($\gamma \to \delta$) for E_1 with respect to the pairs (C_1, C_3) and (C_2, C_4).

For a random permutation the probability that the last condition is satisfied is 2^{-n}, where n is the block size.[10] For E, the probability that the pair (P_1, P_2) is a right pair with respect to the first differential (i.e., the probability that the intermediate difference after E_0 equals β, as predicted by the differential) is p. The probability that both pairs (C_1, C_3) and (C_2, C_4) are right pairs with respect to the second differential is q^2. If all these are right pairs, then $E_1^{-1}(C_3) \oplus E_1^{-1}(C_4) = \beta = E_0(P_3) \oplus E_0(P_4)$. Thus, with probability p, $P_3 \oplus P_4 = \alpha$. Hence, the total probability of this quartet of plaintexts and ciphertexts to satisfy the condition $P_3 \oplus P_4 = \alpha$ is at least $(pq)^2$.

The attack can be mounted for all possible β's and γ's simultaneously (as long as $\beta \neq \gamma$). Therefore, a right quartet for E is encountered with probability not less than $(\hat{p}\hat{q})^2$, where:

$$\hat{p} = \sqrt{\sum_{\beta} \Pr^2[\alpha \to \beta]}, \qquad \text{and} \qquad \hat{q} = \sqrt{\sum_{\gamma} \Pr^2[\gamma \to \delta]}.$$

Using the boomerang process described above, the cipher E can be distinguished from a random permutation given $O((\hat{p}\hat{q})^{-2})$ adaptively chosen plaintexts and ciphertexts, provided that $\hat{p}\hat{q} \gg 2^{-n/2}$. The complete analysis is given in [7], [8], [40]. We omit the analysis here since it is essentially included in the analysis of the related-key boomerang attack presented in Section II-B.

As the boomerang distinguisher requires adaptively chosen plaintexts and ciphertexts, it cannot be combined with many of the standard techniques for using distinguishers in key recovery attacks. This led to the introduction of a chosen plaintext variant of the boomerang attack called the *amplified boomerang attack* [25], and later renamed as the *rectangle attack* [7]. The transformation of the boomerang attack into a chosen plaintext attack relies on standard birthday-paradox arguments. The key idea behind the transformation is to encrypt many plaintext pairs with input difference α, and to look for quartets (i.e., pairs of pairs) that conform to the requirements of the boomerang process.

In the rectangle distinguisher, the adversary considers quartets of plaintexts of the form $((P_1, P_2 = P_1 \oplus \alpha), (P_3, P_4 = P_3 \oplus \alpha))$. A quartet is called a "right quartet" if the following conditions are satisfied:

1) $E_0(P_1) \oplus E_0(P_2) = \beta = E_0(P_3) \oplus E_0(P_4)$.
2) $E_0(P_1) \oplus E_0(P_3) = \gamma$ (which leads to $E_0(P_2) \oplus E_0(P_4) = \gamma$ if previous condition holds as well).
3) $C_1 \oplus C_3 = \delta = C_2 \oplus C_4$.

[10]For the analysis of E we rely on some independence assumptions, addressed in Section II-D.

The probability of a quartet to be a right quartet is a lower bound on the probability of the event

$$C_1 \oplus C_3 = \delta = C_2 \oplus C_4. \tag{1}$$

The usual assumption is that each of the above conditions is independent of the rest, and hence the probability that a given quartet $((P_1, P_2), (P_3, P_4))$ is a right quartet is $p^2 \cdot 2^{-n-1} \cdot q^2$. Since for a random permutation, the probability of Condition (1) is 2^{-2n}, the rectangle process can be used to distinguish E from a random permutation if $pq \gg 2^{-n/2}$ (like in the boomerang distinguisher).

The data complexity of the distinguisher is $O(2^{n/2}(pq)^{-1})$, which is much higher than the complexity of the boomerang distinguisher. The higher data complexity follows from the fact that the event $E_0(P_1) \oplus E_0(P_3) = \gamma$ occurs with a "random" probability of 2^{-n} (actually, this is the birthday-paradox argument used in the construction). The identification of right quartets is also more complicated than in the boomerang case, as instead of checking a condition on pairs, the adversary has to go over all the possible quartets. At the same time, the chosen plaintext nature allows using stronger key recovery techniques. An optimized method of finding the right rectangle quartets is presented in [8].

Like the boomerang attack, the rectangle attack can use all the possible β's and γ's simultaneously. This reduces the data complexity of the attack to $O(2^{n/2}(\hat{p}\hat{q})^{-1})$, where $\hat{p}$ and $\hat{q}$ are as defined above. The complete analysis of the rectangle attack is given in [7], [8].

B. The Related-Key Boomerang Attack

We now present the RK-boomerang distinguisher, and determine the conditions required for the distinguisher to succeed. Following a rigorous treatment, we compute the optimal value of the threshold used in the distinguisher using the Logarithmic Likelihood Ratio (LLR) method. Then we compute the success rate of the distinguisher using a Poisson approximation. In order to keep this section readable, we refrain from presenting a detailed analysis of the key-recovery attack algorithm. The reader is referred to [8] for a generic key-recovery attack algorithm exploiting the boomerang distinguisher (which is easily adapted to the related-key model).
First, we recall the definition of related-key differentials and introduce a shorthand used throughout this paper to denote them:

Definition *1:* We say that a related-key differential $\alpha \to \beta$ with key difference ΔK holds for E with probability p, if

$$\Pr_{P,K}[E_K(P) \oplus E_{K \oplus \Delta K}(P \oplus \alpha) = \beta] = p,$$

where $E_K(\cdot)$ denotes encryption through E with the key K. For the ease of exposition, we denote this event by $\Pr\left(\alpha \xrightarrow[\Delta K]{E} \beta\right) = p$. For sake of simplicity, we shall denote the related-key differential by $\left(\alpha \xrightarrow[\Delta K]{E} \beta\right)$.

In order to present the independence assumption used in the paper, we need another definition:

Definition *2:* For each related-key differential $\left(\alpha \xrightarrow[\Delta K]{E} \beta\right)$, we denote the set of right pairs with respect to the differential (for the given key K) by $G_K\left(\alpha \xrightarrow[\Delta K]{E} \beta\right)$. Formally, for a block cipher E and a given key K,

$$G_K\left(\alpha \xrightarrow[\Delta K]{E} \beta\right) = \left\{P \middle| E_K(P) \oplus E_{K\oplus\Delta K}(P\oplus\alpha) = \beta\right\}.$$

Similarly, we define the set of good ciphertexts:

$$G_K^{-1}\left(\alpha \xrightarrow[\Delta K]{E} \beta\right) = \left\{E_K(P) \middle| P \in G\left(\alpha \xrightarrow[\Delta K]{E} \beta\right)\right\}$$
$$= \left\{C \middle| E_K^{-1}(C) \oplus E_{K\oplus\Delta K}^{-1}(C\oplus\beta) = \alpha\right\}.$$

Our independence assumption asserts that the sets of the form $G\left(\alpha \xrightarrow[\Delta K]{E} \beta\right)$ are independent, in the following sense:

Assumption 1: For the block cipher $E = E_1 \circ E_0$ under consideration, for any fixed key K, and for any set of differences $\alpha, \gamma_1, \delta, \Delta K_0$, and ΔK_1, we assume that the event $\left(X \in G_K\left(\gamma_1 \xrightarrow[\Delta K_1]{E_1} \delta\right)\right)$ is independent of any combination of these three events:

1) $\left(X \oplus \beta_1 \in G_{K\oplus\Delta K_0}\left(\gamma_2 \xrightarrow[\Delta K_1]{E_1} \delta\right)\right)$ for all β_1, γ_2.
2) $\left(X \in G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta_1\right)\right)$ for all β_1.
3) $\left(X \oplus \gamma_1 \in G_{K\oplus\Delta K_1}^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta_2\right)\right)$ for all β_2.

For example, our independence assumption asserts that

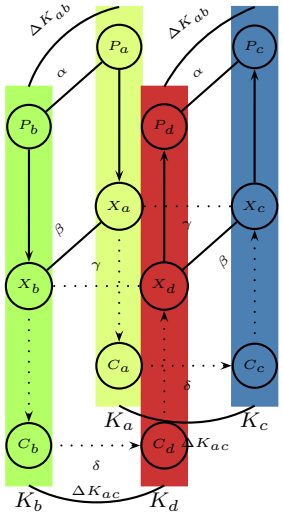
$$\Pr\left[X \in G_K\left(\gamma_1 \xrightarrow[\Delta K_1]{E_1} \delta\right) \middle|\right.$$
$$\left(X \oplus \beta_1 \in G_{K\oplus\Delta K_0}\left(\gamma_2 \xrightarrow[\Delta K_1]{E_1} \delta\right)\right) \bigwedge$$
$$\left(X \in G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta_1\right)\right) \bigwedge$$
$$\left.\left(X \oplus \gamma_1 \in G_{K\oplus\Delta K_1}^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta_2\right)\right)\right]$$
$$= \Pr\left[X \in G_K\left(\gamma_1 \xrightarrow[\Delta K_1]{E_1} \delta\right)\right].$$

This assumption is used implicitly in all the statements in the sequel. We discuss the assumption and its relation to the randomness assumptions used in other techniques, such as differential and linear cryptanalysis, in Section II-D.

1) The Related-Key Boomerang Distinguisher: Now we are ready to present the RK-boomerang distinguisher. Similarly to the boomerang attack, we treat the cipher E as a cascade of sub-ciphers: $E = E_1 \circ E_0$. The distinguisher involves four different unknown (but related) keys — K_a, $K_b = K_a \oplus \Delta K_{ab}$, $K_c = K_a \oplus \Delta K_{ac}$, and $K_d = K_a \oplus \Delta K_{ab} \oplus \Delta K_{ac}$. For fixed values α and δ, the attack algorithm is the following:

1) Choose M plaintexts at random, and set a counter C to zero. For each plaintext P_a, perform the following:

Fig. 1. A Related-Key Boomerang Quartet



a) Compute $P_b = P_a \oplus \alpha$.
b) Ask for the ciphertexts $C_a = E_{K_a}(P_a)$ and $C_b = E_{K_b}(P_b)$.
c) Compute $C_c = C_a \oplus \delta$ and $C_d = C_b \oplus \delta$.
d) Ask for the plaintexts $P_c = E_{K_c}^{-1}(C_c)$ and $P_d = E_{K_d}^{-1}(C_d)$.
e) Check whether $P_c \oplus P_d = \alpha$. If yes, increase the value of the counter C by 1.

2) If $C > Threshold$, output "The cipher E". Otherwise, output "Random Permutation".

The value of $Threshold$ will be specified later in this section. See Figure 1 for an outline of a right RK-boomerang quartet.

It is easy to see that for a random permutation, the probability that the condition $P_c \oplus P_d = \alpha$ is satisfied is 2^{-n}. The probability that the condition is satisfied for E is given in the following proposition:

Proposition *1:* Consider a quartet (P_a, P_b, P_c, P_d) constructed by the algorithm described above. We have

$$\Pr[P_c \oplus P_d = \alpha] =$$
$$\sum_{\beta_1\oplus\beta_2\oplus\gamma_1\oplus\gamma_2=0} \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right] \cdot$$
$$\Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right]. \quad (2)$$

In particular,

$$\Pr[P_c \oplus P_d = \alpha] \geq (\hat{p}\hat{q})^2, \quad (3)$$

where

$$\hat{p} = \sqrt{\sum_{\beta'} \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta'\right]^2} \text{ and } \hat{q} = \sqrt{\sum_{\gamma'} \Pr\left[\gamma' \xrightarrow[\Delta K_{ac}]{E_1} \delta\right]^2}.$$

Proof: Consider a quartet (P_a, P_b, P_c, P_d) constructed by the algorithm. Denote the intermediate values

$(E_0(P_a), E_0(P_b), E_0(P_c), E_0(P_d))$ (where the encryption is under the respective keys) by (X_a, X_b, X_c, X_d), respectively. For all $\beta_1, \gamma_1, \gamma_2$, we say that the event $S_{\beta_1,\gamma_1,\gamma_2}$ occurs, if the following three conditions are satisfied:

$$X_a \oplus X_b = \beta_1, \qquad X_a \oplus X_c = \gamma_1, \qquad X_b \oplus X_d = \gamma_2.$$

Since the events $\{S_{\beta_1,\gamma_1,\gamma_2}\}$ for different values of $(\beta_1, \gamma_1, \gamma_2)$ are disjoint and their union is the entire space, we have

$$\Pr[P_c \oplus P_d = \alpha] =$$

$$\sum_{\beta_1,\gamma_1,\gamma_2} \Pr[P_c \oplus P_d = \alpha \big| S_{\beta_1,\gamma_1,\gamma_2}] \cdot \Pr[S_{\beta_1,\gamma_1,\gamma_2}]. \quad (4)$$

If the event $S_{\beta_1,\gamma_1,\gamma_2}$ occurs, then

$$X_c \oplus X_d = (X_c \oplus X_a) \oplus (X_a \oplus X_b) \oplus (X_b \oplus X_d) = \gamma_1 \oplus \beta_1 \oplus \gamma_2.$$

Hence, by the independence assumptions,

$$\Pr\left[P_c \oplus P_d = \alpha \Big| S_{\beta_1,\gamma_1,\gamma_2}\right] = \Pr[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2], \quad (5)$$

where $\beta_2 = \gamma_1 \oplus \beta_1 \oplus \gamma_2$. Similarly, the three conditions forming the event $S_{\beta_1,\gamma_1,\gamma_2}$ are independent, and hence

$$\Pr[S_{\beta_1,\gamma_1,\gamma_2}] = \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot$$

$$\Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right]. \quad (6)$$

Substituting Equations (5) and (6) into Equation (4) yields Equation (2).

$$\sum_{\beta_1 \oplus \beta_2 \oplus \gamma_1 \oplus \gamma_2 = 0} \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right] \cdot$$

$$\Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \geq$$

$$\geq \sum_{\substack{\beta_1 \oplus \beta_2 = 0, \\ \gamma_1 \oplus \gamma_2 = 0}} \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right] \cdot$$

$$\Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] =$$

$$\sum_{\beta'} \left(\Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta'\right]^2\right) \sum_{\gamma'} \left(\Pr\left[\gamma' \xrightarrow[\Delta K_{ac}]{E_1} \delta\right]^2\right) = (\hat{p}\hat{q})^2,$$

and thus, Inequality (3) follows from Equation (2). ■

Proposition 1 shows that if $\hat{p}\hat{q} > 2^{-n/2}$, then the probability that the condition $P_c \oplus P_d = \alpha$ holds, is higher for E than for a random permutation, i.e., we expect more quartets in the case of E. We next compute the optimal choice of the value $Threshold$ used in the distinguisher.

2) The Optimal Choice of $Threshold$: The optimal value of $Threshold$ can be found using the Likelihood Ratio test for the distributions representing $\Pr[P_c \oplus P_d = \alpha]$ for E and for a random permutation. We use the following standard result:

Proposition *2 ([1], Proposition 1):* Consider two distributions D_0 and D_1 assuming values in a finite set Z, and a sample z^m of m independent elements of Z (represented as a vector in Z^m). The optimal test for deciding whether the sample is distributed according to D_0 or to D_1 is the test having acceptance region

$$A_{D_0} = \{z^m \in Z^m : LLR(z^m) \geq 0\},$$

where

$$LLR(z^m) = \sum_{a \in Z} N(a|z^m) \log \frac{\Pr_{D_0}[a]}{\Pr_{D_1}[a]}$$

is the logarithmic likelihood ratio (with the convention that $\log(0/p) = -\infty$ and $\log(p/0) = \infty$), and where $N(a|z^m)$ is the number of times a occurs in the sequence z^m.

Denote $p_0 = \Pr[P_c \oplus P_d = \alpha]$ (where P_c and P_d are constructed by the boomerang process). We apply Proposition 2, where D_0 and D_1 are the distributions representing $\Pr[P_c \oplus P_d = \alpha]$ for E and for a random permutation, respectively. In this case, $Z = \{0, 1\}$, $m = M$, and both distributions represent Bernoulli random variables, where $D_0 = Ber(p_0)$ and $D_1 = Ber(2^{-n})$. Hence,

$$LLR(z^M) = N(0|z^M) \log \frac{1 - p_0}{1 - 2^{-n}} + N(1|z^M) \log \frac{p_0}{2^{-n}}. \quad (7)$$

Since in our distinguisher, the acceptance region of the test is $\{z^M \in Z^M : N(1|z^M) \geq Threshold\}$, the optimal value of $Threshold$ is $\min\{k : f(k) \geq 0\}$, where

$$f(k) = (M - k) \log \frac{1 - p_0}{1 - 2^{-n}} + k \log \frac{p_0}{2^{-n}}.$$

A simple computation shows that the optimal value is

$$Threshold = \left\lceil \frac{-\log \frac{1-p_0}{1-2^{-n}}}{\log \frac{p_0(1-2^{-n})}{(1-p_0)2^{-n}}} M \right\rceil. \quad (8)$$

3) The Success Probability of the Distinguisher: We use the following standard definition of the success probability of a distinguisher (see, e.g., [1]):

Definition *3:* Let A be a distinguisher between distributions D_0 and D_1, such that for $j = 0, 1$, the statement $[A(D) = j]$ corresponds to "D is distributed like D_j". The probability of success of A is

$$\Pr{}_s(A) = \frac{\Pr[A(D) = 0 | D = D_0] + \Pr[A(D) = 1 | D = D_1]}{2}.$$

Since the distinguisher counts the number of successes amongst M trials, it actually distinguishes between the Binomial distributions $Bin(M, p_0)$ and $Bin(M, 2^{-n})$. Hence,

given the value $Threshold$ (as computed in Equation 8), the success probability of the distinguisher is given by the formula:

$$\Pr[Success] = \frac{1}{2}\Big[\Pr[Bin(M,p_0) \geq Threshold] + \Pr[Bin(M,2^{-n}) < Threshold]\Big] =$$

$$= \frac{1}{2}\left[\sum_{k=Threshold}^{M}\binom{M}{k}p_0^k(1-p_0)^{M-k} + \sum_{k=0}^{Threshold-1}\binom{M}{k}2^{-nk}(1-2^{-n})^{M-k}\right]. \quad (9)$$

For a large value of M (like the values usually used in attacks as M has to be at least $1/p_0$, and p_0 is in most cases very small), the Binomial distributions can be approximated by the Poisson distributions $Poi(p_0M)$ and $Poi(2^{-n}M)$. Using this approximation, Equation (9) is simplified to:

$$\Pr[Success] \approx \frac{1}{2}\left[1 - e^{-p_0M}\sum_{k=0}^{Threshold-1}\frac{(p_0M)^k}{k!} + e^{-2^{-n}M}\sum_{k=0}^{Threshold-1}\frac{(2^{-n}M)^k}{k!}\right]. \quad (10)$$

Denote $c = Mp_0$, and $x = p_0/2^{-n}$. Equation (10) can be reformulated into:

$$\Pr[Success] \approx \frac{1}{2}\left[1 - \left(e^{-c}\cdot\sum_{k=0}^{Threshold-1}\frac{c^k}{k!}\right) + \left(e^{-c/x}\cdot\sum_{k=0}^{Threshold-1}\frac{(c/x)^k}{k!}\right)\right]. \quad (11)$$

We note that in actual attacks, c usually satisfies $1 \leq c \leq 100$, while the value x varies significantly between different attacks. In Table I, we give the optimal threshold and success rate for several common values of c and x.

When x tends to infinity, Equation (11) can be simplified, as $e^{-c/x}$ tends to 1. In other words, when $x \gg 1$, given $M = c \cdot p_0^{-1}$ quartets, a threshold of 1 is sufficient to achieve the following success rate:

$$\Pr[success] \approx \frac{1}{2}\left(1 - e^{-c} + 1\right) = 1 - \frac{e^{-c}}{2}.$$

We note that while for attacks based on linear cryptanalysis, the probability of success can be approximated using the Normal distribution (see, e.g., [1], [37]) in attacks based on differential cryptanalysis (like the attacks discussed in this paper) the Normal approximation may be inaccurate. The reason for the difference is that while in linear-based attacks, the value of the measured random variable is big (close to $M/2$), in differential-based attacks the value is usually very small (e.g., $1 \leq Threshold \leq 10$). For such small values, the approximation of a random variable assuming only integer values by a Normal distribution is inaccurate, and hence approximation using a Poisson random variable is preferable.[11]

4) Practical Lower Bounds for $\hat{p}$ and $\hat{q}$: In practical attacks, the probability of the RK-boomerang distinguisher (given by Equation 2) cannot be computed. Moreover, even the computation of the lower bound given by Inequality (3) is infeasible in most of the cases. Instead, the adversary finds high-probability differential characteristics $\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta\right)$ and $\left(\gamma \xrightarrow[\Delta K_{ac}]{E_1} \delta\right)$. Then, the adversary computes lower bounds for $\hat{p}$ and $\hat{q}$ by considering only part of the possible β' and γ' values. For example, she can take into consideration all the characteristics $\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta'\right)$ that coincide with the characteristic $\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta\right)$ in all the rounds except for the last one, and take all possible values in the output difference of the last round.

In certain cases, especially when a good differential cannot be found, the following simple proposition is useful as a generic lower bound for $\hat{p}$ and $\hat{q}$.

Proposition *3:* Consider related-key differentials through E_0 with input difference α and key difference ΔK. If there exist only m differences β' such that $\Pr\left[\alpha \xrightarrow[\Delta K]{E_0} \beta'\right] > 0$, then $\hat{p} \geq \sqrt{1/m}$. Moreover, equality holds if and only if all the m differentials $(\alpha \xrightarrow[\Delta K]{} \beta')$ with non-zero probability have probability $1/m$ each.

Proof: Recall that the Cauchy-Schwarz inequality asserts that for any two sequences $\{a_1, a_2, \ldots, a_m\}$ and $\{b_1, b_2, \ldots, b_m\}$ of non-negative numbers, we have

$$\sum_{i=1}^{m} a_i \cdot b_i \leq \sqrt{\sum_{i=1}^{m} a_i^2} \cdot \sqrt{\sum_{i=1}^{m} b_i^2}.$$

Denote the probabilities of the differentials of the form $\left(\alpha \xrightarrow[\Delta K]{E_0} \beta'\right)$ by $p_1, p_2, \ldots, p_m$ (ignoring the differentials with zero probability). Clearly, we have

$$\sum_{i=1}^{m} p_i = 1, \qquad \text{and} \qquad \hat{p} = \sqrt{\sum_{i=1}^{m} p_i^2}.$$

We apply the Cauchy-Schwarz inequality for the sequences $\{p_1, p_2, \ldots, p_m\}$ and $\{1, 1, \ldots, 1\}$ and obtain

$$1 = \sum_{i=1}^{m} p_i \cdot 1 \leq \sqrt{\sum_{i=1}^{m} p_i^2} \cdot \sqrt{\sum_{i=1}^{m} 1} = \hat{p}\sqrt{m},$$

and hence $\hat{p} \geq \sqrt{1/m}$, as asserted. Furthermore, since equality in the Cauchy-Schwarz inequality holds if and only if the sequences $\{a_i\}_{i=1}^m$ and $\{b_i\}_{i=1}^m$ are proportional (i.e., there

[11] In [37] the success probabilities of both a linear attack and a differential attack are approximated using the Normal distribution. The experiments presented in [37] show that the approximation is much more accurate in the case of linear cryptanalysis. It is possible that using a Poisson approximation yields a better accuracy in the differential case, as explained above.

TABLE I
OPTIMAL THRESHOLDS AND SUCCESS RATES FOR COMMON PARAMETERS

x	$c=1$	$c=2$	$c=3$	$c=4$	$c=6$	$c=8$	$c=16$
2	1 (61.9%)	2 (66.5%)	Imp	Imp	Imp	Imp	Imp
4	1 (70.5%)	2 (75.2%)	2 (81.4%)	3 (84.1%)	Imp	Imp	Imp
10	1 (76.8%)	1 (84.2%)	2 (88.2%)	2 (92.3%)	3 (95.7%)	4 (97.4%)	Imp
16	1 (78.6%)	1 (87.4%)	2 (89.3%)	2 (94.1%)	3 (96.6%)	3 (98.6%)	6 (99.9%)
100	1 (81.1%)	1 (92.2%)	1 (96.0%)	1 (97.1%)	2 (99.0%)	2 (99.7%)	4 (99.99%)
200	1 (81.4%)	1 (92.7%)	1 (96.8%)	1 (98.1%)	2 (99.1%)	2 (99.8%)	4 (99.995%)
1000	1 (81.6%)	1 (93.1%)	1 (97.4%)	1 (98.9%)	1 (99.6%)	2 (99.8%)	3 (99.999%)
10000	1 (81.6%)	1 (93.2%)	1 (97.5%)	1 (99.1%)	1 (99.8%)	1 (99.9%)	2 (99.9998%)

The entry $X(Y\%)$ means that the optimal threshold is X and the success rate is Y.
Imp — it is impossible to gather the amount of data required in this case.

exists c such that $a_i = c \cdot b_i$ for all i), in our case equality holds if and only if all the p_i's are equal. ■

The generic lower bound given by Proposition 3 can be combined with a "good" differential for part of the rounds.

Proposition *4:* Consider related-key differentials through E_0 with input difference α and key difference ΔK. Assume that there exists a decomposition $E_0 = E_{01} \circ E_{00}$, and a difference α', such that:

1) $\Pr\left[\alpha \xrightarrow[\Delta K]{E_{00}} \alpha'\right] = p'$, and
2) There exist only m differences β' such that $\Pr[\alpha' \xrightarrow[\Delta K]{E_{01}} \beta'] > 0$.

Then $\hat{p} \geq p'\sqrt{1/m}$.

Proof: We compute a lower bound on $\hat{p}$ by considering only the characteristics $\left(\alpha \xrightarrow[\Delta K]{E_0} \beta'\right)$ for E_0 whose restriction to E_{00} is $\left(\alpha \xrightarrow[\Delta K]{E_{00}} \alpha'\right)$. By the assumptions, there are only m such differentials (ignoring differentials with probability zero), and the sum of their probabilities is p'. The assertion follows from the Cauchy-Schwarz inequality by the same argument as used in the proof of Proposition 3. ■

Clearly, the same arguments apply also for the computation of $\hat{q}$.

C. The Related-Key Rectangle Attack

The transformation of the RK-boomerang attack into the RK-rectangle attack is similar to the transformation of the boomerang attack to the rectangle attack in the single-key model. The RK-rectangle distinguisher involves four different unknown (but related) keys — K_a, $K_b = K_a \oplus \Delta K_{ab}$, $K_c = K_a \oplus \Delta K_{ac}$, and $K_d = K_a \oplus \Delta K_{ab} \oplus \Delta K_{ac}$. For fixed values α and δ, the algorithm of the distinguisher is as follows:

1) Choose M plaintexts P_a, and compute $P_b = P_a \oplus \alpha$. Ask for the ciphertexts $C_a = E_{K_a}(P_a)$ and $C_b = E_{K_b}(P_b)$.
2) Choose M plaintexts P_c, and compute $P_d = P_c \oplus \alpha$. Ask for the ciphertexts $C_c = E_{K_c}(P_c)$ and $C_d = E_{K_d}(P_d)$.
3) Set a counter C to zero.
4) For each of the M^2 choices for (P_a, P_c) (and the corresponding (P_b, P_d)):
   a) Check whether both conditions $C_a \oplus C_c = \delta$ and $C_b \oplus C_d = \delta$ are satisfied. If yes, increase the value of the counter C by 1.
5) If $C > Threshold$, output "The cipher E". Otherwise, output "Random Permutation".

The value of $Threshold$ will be specified later in this section.

It is easy to see that for a random permutation, the probability that both conditions $C_a \oplus C_c = \delta$ and $C_b \oplus C_d = \delta$ are satisfied is 2^{-2n}. The probability that the conditions are satisfied for E is given in the following proposition:

Proposition *5:* Consider a quartet of plaintexts and their corresponding ciphertexts $((P_a, C_a), (P_b, C_b), (P_c, C_c), (P_d, C_d))$ constructed by the algorithm described above. We have

$$\Pr\left[(C_a \oplus C_c = \delta) \wedge (C_b \oplus C_d = \delta)\right] \approx$$

$$\approx 2^{-n} \sum_{\beta_1 \oplus \beta_2 \oplus \gamma_1 \oplus \gamma_2 = 0} \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right] \cdot$$

$$\Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right]. \quad (12)$$

In particular,

$$\Pr\left[\left(C_a \oplus C_c = \delta\right) \wedge \left(C_b \oplus C_d = \delta\right)\right] \geq 2^{-n}(\hat{p}\hat{q})^2, \quad (13)$$

where

$$\hat{p} = \sqrt{\sum_{\beta'} \Pr[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta']^2}, \text{ and } \hat{q} = \sqrt{\sum_{\gamma'} \Pr[\gamma' \xrightarrow[\Delta K_{ac}]{E_1} \delta]^2}.$$

Proof: The proof is similar to the proof of Proposition 1. Consider a quartet $((P_a, C_a), (P_b, C_b), (P_c, C_c), (P_d, C_d))$ constructed by the algorithm. Denote the intermediate values $(E_0(P_a), E_0(P_b), E_0(P_c), E_0(P_d))$ (where the encryption is under the respective keys) by (X_a, X_b, X_c, X_d). For all $\beta_1, \beta_2, \gamma_1$, we say that the event $S_{\beta_1,\beta_2,\gamma_1}$ occurs, if the following three conditions are satisfied:

$$X_a \oplus X_b = \beta_1, \qquad X_c \oplus X_d = \beta_2, \qquad X_a \oplus X_c = \gamma_1.$$

Since the events $\{S_{\beta_1,\beta_2,\gamma_1}\}$ for different values of $(\beta_1, \beta_2, \gamma_1)$ are disjoint and their union is the entire space, we have

$$\Pr\left[\left(C_a \oplus C_c = \delta\right) \wedge \left(C_b \oplus C_d = \delta\right)\right] =$$

$$= \sum_{\beta_1,\beta_2,\gamma_1} \Pr\left[\Big(C_a \oplus C_c = \delta\Big) \wedge \Big(C_b \oplus C_d = \delta\Big)\Big| S_{\beta_1,\beta_2,\gamma_1}\right] \cdot \Pr[S_{\beta_1,\beta_2,\gamma_1}]. \quad (14)$$

If the event $S_{\beta_1,\beta_2,\gamma_1}$ occurs, then

$$X_b \oplus X_d = (X_b \oplus X_a) \oplus (X_a \oplus X_c) \oplus (X_c \oplus X_d) = \beta_1 \oplus \gamma_1 \oplus \beta_2.$$

Hence, by the independence assumption,

$$\Pr\left[\Big(C_a \oplus C_c = \delta\Big) \wedge \Big(C_b \oplus C_d = \delta\Big)\Big| S_{\beta_1,\beta_2,\gamma_1}\right] = \Pr\left[\gamma_1 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right] \cdot \Pr\left[\gamma_2 \xrightarrow[\Delta K_{ac}]{E_1} \delta\right], \quad (15)$$

where $\gamma_2 = \beta_1 \oplus \gamma_1 \oplus \beta_2$. Applying again the independence assumption, we have

$$\Pr[S_{\beta_1,\beta_2,\gamma_1}] = \Pr\left[X_a \in G^{-1}_{K_a}\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right) \Big| \left(X_c \in G^{-1}_{K_c}\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right)\right) \bigwedge \Big(X_a \oplus X_c = \gamma_1\Big)\right] \cdot \cdot \Pr\left[X_c \in G^{-1}_{K_c}\left(\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right) \Big| X_a \oplus X_c = \gamma_1\right] \cdot \Pr\left[X_a \oplus X_c = \gamma_1\right] =$$

$$= \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_1\right] \cdot \Pr\left[\alpha \xrightarrow[\Delta K_{ab}]{E_0} \beta_2\right] \cdot \Pr(X_a \oplus X_c = \gamma_1). \quad (16)$$

Since P_a and P_c are chosen independently, then

$$\Pr[X_a \oplus X_c = \gamma_1] \approx 2^{-n}. \quad (17)$$

Note that for any fixed value of $P_a \oplus P_c$ and γ_1, this approximation is rather inaccurate. For an ideal cipher, it is expected that for a fraction $e^{-1/2}$ of the possible values of γ_1, we have $\Pr[X_a \oplus X_c = \gamma_1] = 0$, and for the other values, the probability is at least 2^{-n+1}. However, when the probability is averaged over many different pairs (P_a, P_c), the approximation becomes reasonable.

Substituting Equations (15), (16), and (17) into Equation (14) yields Equation (12). The proof of Equation (13) given Equation (12) is identical to the derivation of Equation (3) from Equation (2) in the proof of Proposition 1. ■

Proposition 5 shows that if $\hat{p}\hat{q} > 2^{-n/2}$, then the probability that the conditions $(C_a \oplus C_c = \delta)$ and $(C_b \oplus C_d = \delta)$ hold simultaneously, is higher for E than for a random permutation, and hence Step 2 of the distinguisher makes sense.

The optimal choice of $Threshold$ and the computation of the success probability of the distinguisher given the probability

$$p_0 = \Pr\left[\Big(C_a \oplus C_c = \delta\Big) \wedge \Big(C_b \oplus C_d = \delta\Big)\right]$$

are very similar to the respective steps for the related-key boomerang distinguisher presented in Section II-B, and hence are omitted here. A key recovery attack based on the RK-rectangle distinguisher is more complex than the respective RK-boomerang attack, due to the abundance of quartets the adversary has to examine. We do not describe the key-recovery algorithm here, and refer the reader to the algorithm of the rectangle key-recovery attack in [8], that can be easily adapted to the related-key model. We note that Table I can also be applied to the case of the rectangle attack, with a different value for p_0, c and x: For the RK-rectangle attack, $p_0 = 2^{-n}(\hat{p}\hat{q})^2$, $x = p_0/2^{-2n}$, and c is the number of expected quartets (i.e., given $M = \sqrt{c/p_0}$ pairs).

D. The Independence Assumptions

All statistical cryptanalytic techniques require various randomness assumptions. For example, the construction of differential characteristics uses the assumption that the cipher is a *Markov cipher* (see [6]), which implies that the characteristics of single rounds are independent of each other and can be combined. Linear cryptanalysis is based on Matsui's Piling up Lemma [31], which essentially asserts that linear approximations of single rounds are independent. These randomness assumptions allow a rigorous treatment of the techniques, as well as better applicability (since the search of differentials and linear approximations can be done for each round separately). It is easy to construct examples of ciphers that do not satisfy the randomness assumptions, which would result in failure of the differential or the linear attacks.[12] However, based on many experimental results, it is reasonable to assume that most of the ciphers satisfy the randomness assumptions. Moreover, if some cipher does not satisfy these assumptions, then this non-randomness is probably exploitable in some other attack, e.g., an impossible differential attack. Nevertheless, it is important to verify the attacks experimentally whenever possible in order to assure that the assumptions indeed hold in the specific case of interest.

The randomness assumption used in the RK-boomerang and RK-rectangle attacks (i.e., Assumption 1) has two parts. The second part of the assumption, that essentially asserts that differentials of different parts of the cipher are independent, is similar to the standard assumption that the cipher is Markovian, which is used in differential cryptanalysis. However, the first part of Assumption 1 is relatively stronger than the assumptions used in differential cryptanalysis.

Differential cryptanalysis is based on the assumption that for any fixed key K and any (related-key) differential $(\alpha \rightarrow \beta)$, the set $G_K(\alpha \rightarrow \beta)$ is distributed randomly and uniformly in the plaintext space.[13] In the RK-boomerang and RK-rectangle attacks, the assumption deals with the distribution of *pairs of sets* of the class $G_K\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right)$. We assume that any two pairs of such sets are independent, i.e., the events $X \in G_K\left(\gamma_1 \xrightarrow[\Delta K_1]{E_1} \delta\right)$ and $X \oplus \beta_1 \in G_{K \oplus \Delta K_0}\left(\gamma_2 \xrightarrow[\Delta K_1]{E_1} \delta\right)$ are independent, for any value of $\gamma_1, \gamma_2, \beta_1, \delta$, and K.

[12]The flaw in the attacks on SHACAL-1, pointed out in [41], is such an example for differential cryptanalysis.

[13]There are cases in which this cannot be satisfied even in a regular cipher as shown in [18], where the behavior of differential characteristics with probability lower than 2^{-n} is shown to be dependent on the key. This is also the case for weak key classes, i.e., classes of keys which behave significantly different than random.

To show the problem that may exist in the independence assumptions, we give the following simple artificial example.

Assume that for given $K, \alpha,$ and β, for which $MSB(\beta) = 0$ (i.e., the most significant bit of β is 0), we have $G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right) = \{X|MSB(X) = 1\}$ and $G_{K\oplus\Delta K_1}^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right) = \{X|MSB(X) = 0\}$ (in particular, it follows that $\Pr\left[\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right] = 1/2$). Further assume that for some γ such that $MSB(\gamma) = 0$ and for some δ, $\Pr\left[\gamma \xrightarrow[\Delta K_1]{E_1} \delta\right] = 1/2$. By the independence assumptions, it is expected that the probability in a RK-boomerang distinguisher based on the differentials $\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right)$ and $\left(\gamma \xrightarrow[\Delta K_1]{E_1} \delta\right)$ is at least $(1/4)^2 = 1/16$. However, consider a right quartet with respect to this distinguisher and denote the intermediate encryption values by (X_a, X_b, X_c, X_d). Since $X_a \in G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right)$, we have $MSB(X_a) = 1$, and thus, since $MSB(\gamma) = 0$, necessarily $MSB(X_c) = 1$. This implies that $X_c \notin G_{K\oplus\Delta K_1}^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right)$, and thus, the actual probability of the distinguisher is zero![14]

This example demonstrates failure of the first part of Assumption 1 (independence inside the same sub-cipher). Similarly, the second part of the assumption fails if we assume that for some K, α, β, γ and δ, we have $G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right) = \{X|MSB(X) = 1\}$ and $G_K\left(\gamma \xrightarrow[\Delta K_1]{E_1} \delta\right) = \{X|MSB(X) = 0\}$, since in this case X_a cannot be element in both $G_K^{-1}\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta\right)$ and $G_K\left(\gamma \xrightarrow[\Delta K_1]{E_1} \delta\right)$ simultaneously.

In [32], Murphy presented several non-artificial examples, based on a 4-round variant of DES [33] (out of the 16 rounds) and on a 2-round variant of AES [34] (out of the 10 rounds), in which the randomness assumptions fail due to *local inconsistencies* in the transition between E_0 and E_1. In some of the examples, the boomerang distinguisher fails completely, while in other examples, its probability is much higher than the theoretically predicted value. Furthermore, in several specific cases, deviations from the prediction of the independence assumptions were detected in "real" ciphers. Such an example is the *ladder switch* described in [14], where *higher* probability for the RK-boomerang distinguisher is obtained using dependence between the differential used in E_0 and the differential used in E_1.[15]

However, these examples are still not sufficiently representative. As for the examples discussed in [32], they can be misleading due to the small number of rounds in the analyzed variant. In a RK-boomerang attack on an entire cipher (or on a variant with a significant number of rounds), the overall probability of the distinguisher is an average taken over many possible differentials, while in the reduced-round variant, only a small subset of the differentials is considered. It is possible that while the reduced-round attack does not satisfy the independence assumption, the full attack does satisfy it, since the deviations from independence for different characteristics cancel each other. As for the *ladder switch* and related examples, they all refer to cases in which there is a clear dependence between the differential used in E_0 and the differential used in E_1 (e.g., the output difference of the E_0-differential is equal to the input difference of the E_1-differential in one half of the state in a Feistel network), and thus, they may not apply to cases where such dependence does not exist.

Moreover, in the RK-boomerang and RK-rectangle attacks, there are several mechanisms which may overcome dependence problems. The first is the fact that in the attack we count over many differentials (all $\beta_1, \beta_2, \gamma_1, \gamma_2$ such that $\beta_1 \oplus \beta_2 \oplus \gamma_1 \oplus \gamma_2 = 0$), which ensures that even if there is a problem in some combination of differentials, it is expected that other combinations still succeed. The second one is the fact that four different keys are used (in the case $\Delta K_0 \neq \Delta K_1$),and thus, even if there is a dependence between the differentials, it may be countered by the different keys.

In view of the above, it appears that the only possible way to decide whether Assumption (1) makes sense in realistic attacks on block ciphers is to check it experimentally, like the way in which the Markov assumption was checked for differential cryptanalysis. The best possibility would be to test the validity of the assumption for each specific cipher and each specific choice of differentials, but this is usually impossible, due to the high complexity of the attacks.[16] Therefore, we chose a single block cipher – KASUMI – and performed an extensive experimental analysis of various RK-boomerang attacks on its reduced-round variants. Our analysis, presented in Section III, suggests that when averaged over different keys, the probability assumptions do hold, unless there is a local inconsistency between the differentials that can be detected manually. We thus conclude that it is reasonable to assume that the assumptions hold in realistic scenarios; however, it will be of course beneficial to further check the validity of the assumptions in other block ciphers.

E. Generalizations of the Related-Key Boomerang and Related-Key Rectangle Attacks

In this section we briefly present two generalizations of the basic RK-boomerang and RK-rectangle attacks.

[14]Actually, the probability of the distinguisher may be higher due to differentials of the form $\left(\alpha \xrightarrow[\Delta K_0]{E_0} \beta'\right)$ for $\beta' \neq \beta$. However, if there are no high-probability differentials of this form, the probability of the distinguisher is still significantly lower than the predicted value $1/16$.

[15]The cases of such dependence were recently formalized in the *sandwich* framework [19], and thus, we refrain from analyzing them in this paper.

[16]We note that for the rectangle attack such verification is inherently impossible: the data complexity of the attack is lower-bounded by $2^{n/2}$, and thus, its verification is infeasible for any block cipher with block size of 128 bits or more, like AES.

1) Using Structures of Keys: The related-key differentials used in the attack are usually based on fixed *subkey* differences. If the key schedule of the attacked cipher is linear, such differences can be achieved by choosing the appropriate key difference. However, if the key schedule is nonlinear, a fixed key difference does not ensure fixed subkey differences. Instead, the adversary can apply differential cryptanalysis to the key schedule. By studying the differential properties of the key schedule, the adversary can find a key difference that leads to the required subkey differences with a relatively high probability. Then, the adversary can repeat the attack for many pairs of related-keys and expect that in one of the pairs, the required subkey differences are satisfied and the basic RK-boomerang/rectangle attack can be applied.

Furthermore, we observe that the number of keys used in the attack can be reduced by using *structures of keys*. Instead of finding a single key difference leading to the required subkey differences with a high probability, the adversary can find many such key differences (possibly with lower probabilities). Then, the adversary can use structures of keys such that each structure contains many pairs of keys corresponding to different "key characteristics", and thus reduce the number of keys required for the attack. Such an approach is demonstrated in the RK-rectangle attack on AES-256 in [9].

2) Generalizing the Key Relation: While XOR relations are common and inherent to the majority of differential-based related-key attacks, in some cases other key relations are more suitable (either due to the environment of the attack or in order to obtain higher probabilities of the differentials). The RK-boomerang and RK-rectangle attacks can be applied almost without a change when the XOR key relations are replaced by any relation satisfying a condition specified below.

Denote the relation between the keys K and K' by $R(K, K')$. We note that R can be any relation which is symmetric, and covers all keys. At the same time, we recall the fact that the more complex the relation R is, the applicability of the related-key attack may be affected. For example, in the basic RK-boomerang and RK-rectangle attacks we can set $R(K, K') = K \oplus K'$.

The RK-boomerang and RK-rectangle attacks can be applied whenever the key relation satisfies the following condition:

$$\forall (K_a, K_b, K_c, K_d), \Big(R(K_a, K_c) = R(K_b, K_d)\Big) \Longrightarrow \Big(R(K_a, K_b) = R(K_c, K_d)\Big). \quad (18)$$

Condition (18) ensures that in each of the sub-ciphers, the same key relation is used in both differentials. For example, for XOR differences

$$(K_a \oplus K_c = K_b \oplus K_d) \Longrightarrow (K_a \oplus K_b = K_c \oplus K_d),$$

and hence the condition holds.

Condition (18) is satisfied for a wide variety of key relations, including additive differences (e.g., $R(K, K') = (K - K') \bmod 2^n$) and rotations. On the other hand, the condition does not hold if the relation used in the first sub-cipher (i.e., between (K_a, K_b) and (K_c, K_d)) and the relation used in the second sub-cipher (i.e., between (K_a, K_c) and (K_b, K_d)) are of different classes (e.g., XOR differences in the first sub-cipher and modular differences in the second sub-cipher).

We note that the basic RK-boomerang and RK-rectangle attacks can be extended to use different values α, α' in the related-key differentials of E_0, and δ, δ' in the related-key differentials of E_1. Similarly, the attack can use distinct key differences $\Delta K_0, \Delta K_0'$ and $\Delta K_1, \Delta K_1'$ in the differentials of E_0 and E_1, respectively. This allows to extend Condition (18) to the following:

Proposition *6:* The RK-boomerang attack can be applied with two key relations R_1, R_2, as long as for every quadruple (K_a, K_b, K_c, K_d) the relations $R_1(K_a, K_b)$, and $R_2(K_a, K_c), R_2(K_b, K_d)$ imply the relation $R_1(K_c, K_d)$. The RK-rectangle attack can be applied if the relations $R_1(K_a, K_b), R_1(K_c, K_d)$ and $R_2(K_a, K_c)$ imply the relation $R_2(K_b, K_d)$.

Finally, even if the condition of Proposition 6 is not satisfied, in some cases the attack can be still applied using structures of keys, as described earlier.

F. Comparison With Other Related-Key Attacks

For any new technique constructed as a combination of existing techniques, a natural question to ask is whether there are cases in which the combined attack is better than each of its components taken separately. In this section we briefly describe several important cases in which the RK-boomerang and RK-rectangle attacks are expected to outperform each of their components. Concrete examples of the advantage of related-key boomerang and rectangle attacks over other attack techniques are the attack on the full AES [14] and the practical-time attack on KASUMI [19].

The main advantage of the related-key differential attacks over ordinary differential attacks is the ability of the adversary to use the subkey differences to cancel the plaintext difference in the input of one (or more) of the non-linear parts of the cipher. As a result, the adversary obtains one (or more) rounds in the differential that hold with probability 1, allowing the extension of the differential by one (or more) rounds.

In the RK-boomerang and RK-rectangle attacks, the adversary can enjoy this advantage twice, once in each of the sub-ciphers. As a result, the overall distinguisher can be extended by two (and in some cases even more) rounds. This is a significant advantage of the RK-boomerang/rectangle attack over all other differential-based related-key attacks (e.g., related-key differential attack, related-key impossible differential attack and related-key differential-linear attack) that can enjoy the advantage of the related-key model only once.

The advantage of gaining a single additional round (or two rounds) to the distinguisher is significant in ciphers in which the number of rounds is small and each round function is relatively strong. Hence, the gain of the RK-boomerang/rectangle attack is expected to be significant in ciphers like AES [34] and KASUMI [38].

Another property of the cipher required for the success of RK-boomerang and RK-rectangle attacks is simplicity of the

key schedule. The basic version of the attack is applicable only to ciphers with a linear key schedule, but using structures of keys, the attack can be applied to ciphers with a nonlinear key schedule as well. However, if the key schedule of the cipher is complex enough and does not have "good" differential properties, then the number of keys required for the attack becomes infeasibly large.

Summarizing the discussion above, the RK-boomerang and RK-rectangle attacks are expected to be successful if the attacked cipher has the following properties:

- A small number of relatively strong rounds.
- A relatively simple key schedule.

The class of ciphers satisfying these properties includes widely used ciphers such as AES [34], KASUMI [38], and IDEA [30]. These three ciphers can be indeed attacked efficiently using the RK-boomerang/rectangle attack technique.

III. Experimental Analysis of the Related-Key Boomerang Attack: A Case Study

In this section we present experimental analysis of the RK-boomerang attack, in the specific case of the block cipher KASUMI [38]. First we describe the structure of KASUMI and the differentials used in the RK-boomerang attacks on KASUMI presented in [10], [19]. Then, we check experimentally a RK-boomerang distinguisher of 6-round KASUMI, in which the probability of both differentials is relatively high (and thus, performing an extensive experiment is an easy task). Finally, we check RK-boomerang distinguishers of 7-round KASUMI that are similar to the distinguishers used in the attacks on the full KASUMI presented in [10], [19].

We note that our choice of KASUMI is motivated by several reasons:

- It is one of the most widely used ciphers that were attacked by the RK-boomerang technique.
- The distinguisher used in the attacks on the full KASUMI has probability of 2^{-38} which allows to verify it experimentally with a big precision.
- There exists a large set of distinguishers that are similar to the original distinguisher used in [10], [19]. This allows to check all of them and find out which ones lead to local inconsistencies.

A. The KASUMI Block Cipher

KASUMI [38] is a 64-bit block cipher with 128-bit keys, and a recursive Feistel structure, following its ancestor, MISTY1. The cipher has eight Feistel rounds, where each round is composed of two functions: the FO function which is in itself a 3-round 32-bit Feistel construction, and the FL function that mixes a 32-bit subkey with the data in a linear way. The order of the two functions depends on the round number: in the even rounds the FL function is applied first, and in the odd rounds the FO function is applied first.

The FO function also has a recursive structure: its F-function, called FI, is a four-round Feistel construction. The FI function uses two non-linear S-boxes $S7$ and $S9$ (where $S7$ is a 7-bit to 7-bit permutation and $S9$ is a 9-bit to 9-bit permutation), and accepts an additional 16-bit subkey, that is mixed with the data. In total, a 96-bit subkey enters FO in each round — 48 subkey bits are used in the FI functions and 48 subkey bits are used in the key mixing stages.

TABLE II
KASUMI'S KEY SCHEDULE ALGORITHM

Round	$KL_{i,1}$	$KL_{i,2}$	$KO_{i,1}$	$KO_{i,2}$	$KO_{i,3}$	$KI_{i,1}$	$KI_{i,2}$	$KI_{i,3}$
1	$K_1 \lll 1$	K'_3	$K_2 \lll 5$	$K_6 \lll 8$	$K_7 \lll 13$	K'_5	K'_4	K'_8
2	$K_2 \lll 1$	K'_4	$K_3 \lll 5$	$K_7 \lll 8$	$K_8 \lll 13$	K'_6	K'_5	K'_1
3	$K_3 \lll 1$	K'_5	$K_4 \lll 5$	$K_8 \lll 8$	$K_1 \lll 13$	K'_7	K'_6	K'_2
4	$K_4 \lll 1$	K'_6	$K_5 \lll 5$	$K_1 \lll 8$	$K_2 \lll 13$	K'_8	K'_7	K'_3
5	$K_5 \lll 1$	K'_7	$K_6 \lll 5$	$K_2 \lll 8$	$K_3 \lll 13$	K'_1	K'_8	K'_4
6	$K_6 \lll 1$	K'_8	$K_7 \lll 5$	$K_3 \lll 8$	$K_4 \lll 13$	K'_2	K'_1	K'_5
7	$K_7 \lll 1$	K'_1	$K_8 \lll 5$	$K_4 \lll 8$	$K_5 \lll 13$	K'_3	K'_2	K'_6
8	$K_8 \lll 1$	K'_2	$K_1 \lll 5$	$K_5 \lll 8$	$K_6 \lll 13$	K'_4	K'_3	K'_7

$(X \lll i)$ — X rotated to the left by i bits

The FL function accepts a 32-bit input and two 16-bit subkey words. One subkey word affects the data using the OR operation, while the second one affects the data using the AND operation. We outline the structure of KASUMI and its parts in Fig. 2.

The key schedule of KASUMI is very simple and the subkeys are derived from the key linearly. The 128-bit key K is divided into eight 16-bit words: $K_1, K_2, \ldots, K_8$. Each K_i is used to compute $K'_i = K_i \oplus C_i$, where the C_i's are fixed constants (we omit these from the paper as they have no effect on our results). We denote the bits of the subkeys by $K_i = (K_i^{15}, K_i^{14}, \ldots, K_i^0)$, where K_i^{15} is the most significant bit.

In each round, eight words are used as the round subkey (up to some in-word rotations). Therefore, the 128-bit subkey of each round is linearly dependent on the secret key in a very simple way. We give the key schedule algorithm of KASUMI in Table II.

B. Related-Key Differentials of KASUMI

The RK-boomerang distinguishers we examine are based on three related-key differentials: a 4-round differential for rounds 1–4, and 3-round differentials for rounds 4–6 and rounds 5–7.

1) A 4-Round Related-Key Differential for Rounds 1–4: The differential of rounds 1–4 of KASUMI is an extension by one round of the related-key differential presented in [16]. The input difference of this differential is $\alpha = (0_x, 0020\ 0000_x)$, and the key difference is $\Delta K_{ab} = (0, 0, 1_x, 0, 0, 0, 0, 0)$, i.e., only the third key word has a non-zero difference $\Delta K_3 = 0001_x$. The first three rounds of the characteristic have probability 1/4, and due to the Feistel structure, the α difference can propagate to at most 2^{32} differences after round 4. Hence, by Proposition 4, we have

$$\hat{p} \geq \frac{1}{4} \cdot \sqrt{2^{-32}} = 2^{-18}.$$

We outline the differential in Figure 3. Along with this differential, we consider a family of 15 similar differentials, in which the key difference ΔK_{ab} assumes all values in which only a single bit in ΔK_3 is non-zero, and the plaintext difference is shifted accordingly. Examples of such

Fig. 2. Outline of KASUMI

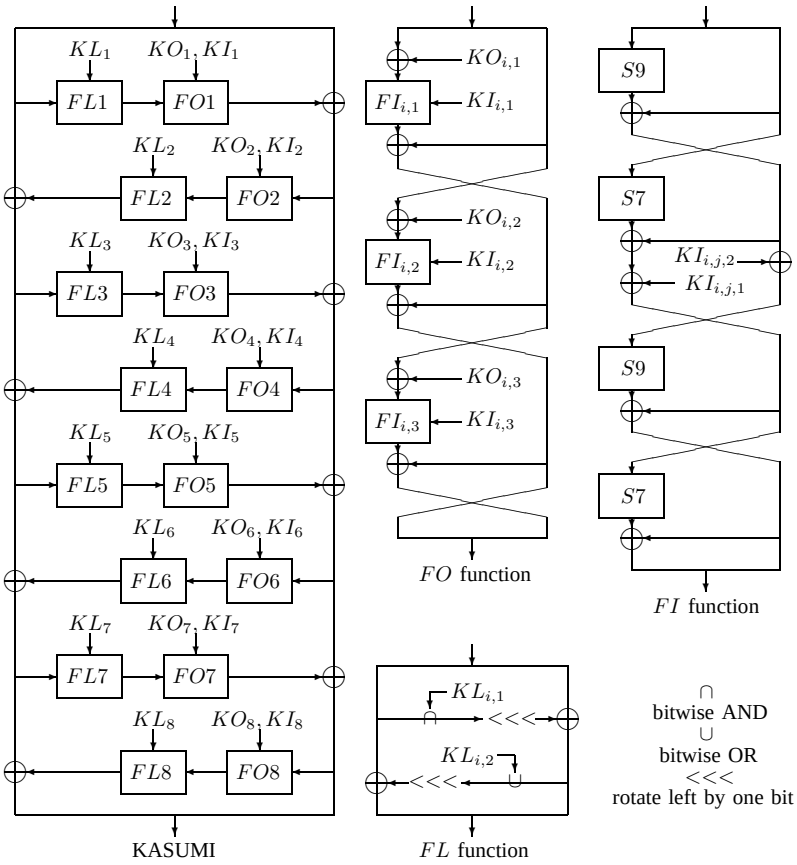
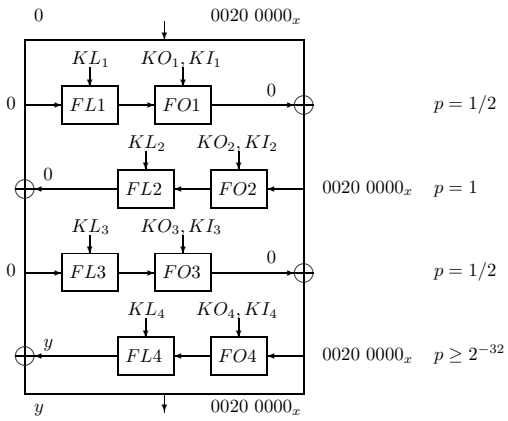


Fig. 3. 4-Round Related-Key Differential Characteristic of KASUMI



differentials are: $\alpha' = (0_x, 0010\ 0000_x)$ with $\Delta K_{ab} = (0, 0, 8000_x, 0, 0, 0, 0, 0)$, and $\alpha'' = (0_x, 0040\ 0000_x)$ with $\Delta K_{ab} = (0, 0, 2_x, 0, 0, 0, 0, 0)$.

It was further observed in [16] that the probability of these differentials can be increased by controlling two plaintext bits. If the adversary assigns one bit of the plaintexts to be one (thus fixing one bit of the output of the OR operation in $FL1$) and one bit of the plaintexts to be zero (thus fixing one bit of the output of the AND operation in $FL1$), then the probability of the differential described in [16] is increased to $1/2$. As a result, for our 4-round differentials we have $\hat{p} \geq 2^{-17}$.[17]

2) 3-Round Related-Key Differential for Rounds 5–7 : The 3-round related-key differential used in rounds 5–7 is the 3-round differential of [16] shifted by four rounds. The key difference is $\Delta K_{ac} = (0, 0, 0, 0, 0, 0, 1_x, 0)$, and the data differences are $\gamma = (0_x, 0020\ 0000_x) \rightarrow (0_x, 0020\ 0000_x) = \delta$. Since we use a single differential (and not count over other possibilities), we have

$$\hat{q} = q = 1/4.$$

As before, along with this differential we consider 15 similar differentials in which the key difference in ΔK_7 is rotated, and the data differences are rotated correspondingly.

3) 3-Round Related-Key Differential for Rounds 4–6.: In rounds 4–6 we use conditional related-key differential characteristics [3], i.e., characteristics that depend on some unknown key bit.

Let $\delta_0 = (0010\ 0000_x, 0_x)$, $\delta_1 = (0010\ 0040_x, 0_x)$, and $\delta' = (0001\ 0000_x, 0_x)$. If $K_5^4 = 0$ (i.e., the fifth least significant bit of K_5 equals zero), we use the two differentials $\delta_0 \rightarrow \delta_0$ and $\delta_0 \oplus \delta' \rightarrow \delta_0$. If $K_5^4 = 1$, we use the differentials $\delta_1 \rightarrow \delta_1$ and $\delta_1 \oplus \delta' \rightarrow \delta_1$. The key difference of all the characteristics is $\Delta K_{ac} = (0, 0, 0, 0, 0, 1_x, 0, 0)$. Each of the four characteristics has probability $1/4$, if K_5^4 has the corresponding value.

[17]Note that if one of the differentials is used in the backward direction, the lower bound remains 2^{-18}.

For example, we describe the difference propagation in the backward direction of the characteristics $\delta_0 \rightarrow \delta_0$ and $\delta_0 \oplus \delta' \rightarrow \delta_0$. Consider a pair with ciphertext difference $\delta_0 = (0010\ 0000_x, 0_x)$. In round 6 the zero difference is preserved with probability 1/2 (i.e., the key difference is cancelled with probability 1/2). In round 5, we need a difference of $0010\ 0000_x$ after $FL5$, which is then cancelled with the key difference in $KO_{5,1}$. If $K_5^4 = 0$, then this is indeed the case with probability 1. In round 4, the zero difference is preserved by the $FO4$ function. As in round 6, it has probability 1/2 to be preserved also by $FL4$, and probability 1/2 to evolve into $\delta_0 \oplus \delta'$. Thus, the input difference of the differential characteristic is either δ_0 or $\delta_0 \oplus \delta'$, with probability $1/4$ each.

In the attack, we apply the distinguisher twice, once with each pair of characteristics, and expect that in one of the applications, both differentials hold with probability $1/4$.[18] For that application, we have

$$\hat{q} = \sqrt{(1/4)^2 + (1/4)^2} = 1/\sqrt{8}.$$

We note that the four conditional differential characteristics we use can be rotated along with the key difference, to produce 15 similar sets of differential characteristics with the same probabilities.

C. First Experiment – Related-Key Boomerang Distinguisher on 6-Round KASUMI

In this subsection we examine a RK-boomerang distinguisher for 6-round KASUMI that we presented in [10]. Let E_0 be rounds 1–3, and let E_1 be rounds 4–6. In E_0 we use the differential $\alpha = (0_x, 0020\ 0000_x) \rightarrow (0_x, 0020\ 0000_x)$ with key difference $\Delta K_{ab} = (0,0,1,0,0,0,0,0)$. As shown in Section III-B.1, the probability of the differential in the forward direction is 1/2 (after adding constraints on the plaintexts), and the probability in the backward direction is 1/4. In E_1, we use the two pairs of differentials $(\delta_0 \rightarrow \delta_0, \delta_0 \oplus \delta' \rightarrow \delta_0)$, and $(\delta_1 \rightarrow \delta_1, \delta_1 \oplus \delta' \rightarrow \delta_1)$, both with key difference $\Delta K_{ac} = (0,0,0,0,0,1,0,0)$. As shown in Section III-B.3, one of the pairs of differentials yields overall probability of $\hat{q} = 1/\sqrt{8}$ (where the "successful" pair depends on the value of the key bit K_5^4).

The attack essentially performs two standard related-key boomerang distinguishers, one for each possible value of the key bit K_5^4. To reduce the data complexity of the attack, we share some of the chosen plaintexts between the two distinguishers. The attack algorithm requires four keys:

$$K_a;\ K_b = K_a \oplus \Delta K_{ab};\ K_c = K_a \oplus \Delta K_{ac};\ K_d = K_b \oplus \Delta K_{ac}.$$

The algorithm of the distinguisher is as follows:

1) Choose M pairs of plaintexts $(P_{a,i}, P_{b,i})$ (for $1 \leq i \leq M$) such that $P_{a,i} \oplus P_{b,i} = \alpha$. For each pair, ask for the encryption of $P_{a,i}$ and $P_{b,i}$ under the keys K_a and K_b, respectively, and denote the corresponding ciphertexts by $C_{a,i}$ and $C_{b,i}$.
2) For $1 \leq i \leq M$, calculate $C_{c,i} = C_{a,i} \oplus \delta_0$ and $C_{d,i} = C_{b,i} \oplus \delta_0$. For all i, ask for the decryption of $C_{c,i}$ and $C_{d,i}$ under the keys K_c and K_d, respectively, and denote the corresponding plaintexts by $P_{c,i}$ and $P_{d,i}$.
3) For $1 \leq i \leq M$, calculate $C_{e,i} = C_{a,i} \oplus \delta_1$ and $C_{f,i} = C_{b,i} \oplus \delta_1$. For all i, ask for the decryption of $C_{e,i}$ and $C_{f,i}$ under the keys K_c and K_d, respectively, and denote the corresponding plaintexts by $P_{e,i}$ and $P_{f,i}$.
4) Check whether $P_{c,i} \oplus P_{d,i} = \alpha$ and count the number of such occurrences.
5) Check whether $P_{e,i} \oplus P_{f,i} = \alpha$ and count the number of such occurrences.
6) If one of the two counters from Steps 4 and 5 is greater than zero, then output "6-Round KASUMI". Otherwise, output "Not 6-Round KASUMI".

The total probability of the boomerang process of this distinguisher is[19] $(1/2) \cdot (1/4) \cdot (1/\sqrt{8})^2 = 1/64$, either for quartets counted in Step 4 or for quartets counted in Step 5. Thus, for $M = 128$ we expect to find two right quartets in Step 4 or in Step 5 (either for the quartets $(P_{a,i}, P_{b,i}, P_{c,i}, P_{d,i})$ or for the quartets $(P_{a,i}, P_{b,i}, P_{e,i}, P_{f,i})$). Moreover, by the analysis presented in Section II, it is expected that the number of right quartets is distributed like a Poisson random variable $Poi(2)$.

In the experiment, we sampled 10,000 random keys, and ran the above distinguisher with $M = 128$. By the analysis presented above, we expected that in 86.5% of the experiments there will be at least one right quartet. Our experiments revealed that in 87% there was at least one such quartet. Moreover, the distribution of the numbers of right quartets obtained in the experiments is indeed very close to the $Poi(2)$ distribution. A comparison between the experimental results and the theoretical prediction is outlined in Table III.

This extensive experiment along with the experimental verification of the high-probability 7-round distinguisher used in [19] (see Section III-D), demonstrate the validity of the probability assumptions in cases where the probabilities of both differentials used in the distinguishers are high. However, it will be beneficial to add more empirical evidence by checking experimentally other boomerang-type distinguishers, such as the boomerang distinguisher of the full COCONUT98 presented in [40] in which the probability of each of the differentials is close to $1/4$.

D. Second Experiment – Related-Key Boomerang Distinguishers on 7-Round KASUMI

The basic RK-boomerang distinguisher on 7-round KASUMI one may consider is the distinguisher used in [10], that is based on the main differentials presented in Sections III-B.1 and III-B.2. Let $\Delta K_{ab} = (0,0,1_x,0,0,0,0,0)$ and $\Delta K_{ac} = (0,0,0,0,0,0,1_x,0)$, and let K_a, $K_b = K_a \oplus \Delta K_{ab}$, $K_c = K_a \oplus \Delta K_{ac}$, and $K_d = K_c \oplus \Delta K_{ab}$ be the unknown related keys. In rounds 1–4, the related-key differential is the one presented in Section III-B.1 that has an input difference $\alpha = (0_x, 0020\ 0000_x)$, a key difference ΔK_{ab} and for which $\hat{p} = 2^{-17}$ in the forward direction (due to fixing plaintext bits properly, as explained in Section III-B.1) and $\hat{p} = 2^{-18}$ in the

[18] We note that the knowledge of the "successful" pair of characteristics reveals the value of the key bit K_5^4.

[19] Recall that the first differential has probability 1/2 for the pair (P_a, P_b) due to fixing the plaintexts correctly.

TABLE III
THE NUMBER OF FOUND QUARTETS IN 10,000 EXPERIMENTS

Quartets	0	1	2	3	4	5	6	7	8	9	10
Experiments	1302	2695	2692	1879	907	348	127	27	9	4	0
Poisson (mean = 2)	1353.3	2706.7	2706.7	1804.5	902.2	360.9	120.3	34.4	8.6	1.9	0.4

backward direction. In rounds 5–7, the related-key differential is the one presented in Section III-B.2 that has an output difference $\delta = (0_x, 0020\ 0000_x)$, a key difference ΔK_{ac} and for which $\hat{q} = 2^{-2}$.

However, as was observed in [19], in this distinguisher there is a clear dependence between the differentials used in E_0 and in E_1. Indeed, it is easy to see that these differentials are equal in the entire 32-bit value that enters the F-function of round 4. As the detailed (and experimentally verified) analysis presented in [19] shows, this leads to a much higher probability of the distinguisher than predicted.

In order to avoid this dependence issue, we consider the 255 distinguishers that can be obtained from the original one by replacing one of the differentials (or both of them) with one of their 15 rotated variants.

1) Checking for local inconsistencies: First we examine whether the distinguisher contains a local inconsistency. A natural candidate for such inconsistency is round 4, since it is the only round in which one of the differentials has a very low (i.e., "random") probability, and since it lies in the transition between E_0 and E_1. In order to concentrate on round 4, we compute the probability of the distinguisher in a slightly different way, that is clearly equivalent to the computation presented in Section II.

We divide the differential of rounds 1–4 into a differential of rounds 1–3 with probability 2^{-2} (or 2^{-1} in the forward direction), and round 4 in which we assume the worst-case assumption that all differentials are equiprobable and count over all the differentials. In order to isolate round 4, we compute the probabilities of the differentials in all other rounds, and only then compute the "cost of the transition" in round 4.[20] By the theoretical analysis, the probability of the distinguisher (for the differentials we examine) is: $2^{-17} \cdot 2^{-2} \cdot 2^{-2} \cdot 2^{-18} = 2^{-39}$. This is equivalent to the claim that given the differentials in rounds 1–3 and 5–7 (whose total probability is 2^{-7}), the cost of the transition is 2^{-32}. This fact is the one that needs verification.

Formally, let (P_a, P_b, P_c, P_d) be a plaintext quartet, denote the corresponding ciphertexts by (C_a, C_b, C_c, C_d), and denote the intermediate values before round 4 by (X_a, X_b, X_c, X_d). Assume that (P_a, P_b) is a right pair for the differential $\alpha \to \beta$ of rounds 1–3, and that $(C_a, C_c), (C_b, C_d)$ are right quartets with respect to the differential $\gamma \to \delta$ of E_1. Due to the Feistel structure of KASUMI, this implies that the right halves of X_a, X_b, X_c, and X_d satisfy:

$$X_a^R \oplus X_b^R \oplus X_c^R \oplus X_d^R = 0,$$

and thus,

$$X_c^R \oplus X_d^R = X_a^R \oplus X_b^R = \beta^R.$$

We would like to check whether the event that indicates that the transition in the middle occurs:

$$X_a^L \oplus X_b^L \oplus X_c^L \oplus X_d^L = 0,$$

holds with the random probability of 2^{-32}, as predicted by the independence assumptions.

This check can be performed by examining solely the function $FO4$. Indeed, since the FL functions are linear (for a fixed key), the condition $X_a^L \oplus X_b^L \oplus X_c^L \oplus X_d^L = 0$ is equivalent to the condition that the XOR of the four intermediate values after the function $FO4$ is zero.

The function $FO4$ (depicted in Figure 4) is a 3-round Feistel construction whose 32-bit values after round j are denoted by $(X_a^j, X_b^j, X_c^j, X_d^j)$. The functions $FI_{4,1}, FI_{4,2}$, and $FI_{4,3}$ are 4-round Feistel constructions, and the 16-bit outputs of $FI_{4,j}$ are denoted by $(I_a^j, I_b^j, I_c^j, I_d^j)$.

First, we observe that in all 255 pairs of differentials we consider, we have $\beta^{RR} = \gamma^{LR} = 0$ (where β^{RR} denotes the 16 rightmost bits of β, and γ^{LR} denotes the right 16 bits of the left half of γ). Hence, if $(P_a, P_b), (C_a, C_c)$, and (C_b, C_d) are right pairs w.r.t. the respective differentials, as assumed, then we have

$$X_a^{RR} = X_b^{RR} = X_c^{RR} = X_d^{RR}. \tag{19}$$

Moreover, since in the second round of $FO4$, there is no key difference inside the key pairs (K_a, K_b) and (K_c, D_d), we have

$$(I_a^2 = I_b^2) \wedge (I_c^2 = I_d^2). \tag{20}$$

Thus,

$$X_a^{3R} \oplus X_b^{3R} \oplus X_c^{3R} \oplus X_d^{3R} = I_a^1 \oplus I_b^1 \oplus I_c^1 \oplus I_d^1. \tag{21}$$

Therefore, if we show that for some pair of differentials, the assumption that $(P_a, P_b), (C_a, C_c)$, and (C_b, C_d) are right pairs implies

$$I_a^1 \oplus I_b^1 \oplus I_c^1 \oplus I_d^1 \neq 0, \tag{22}$$

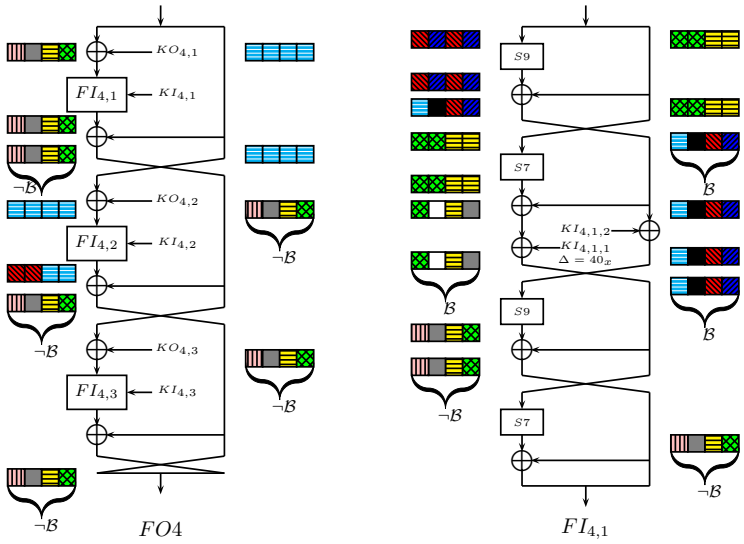
then the probability of the transition in the middle, and thus also of the entire distinguisher, is zero.

Now, we consider the function $FI_{4,1}$. Here we have two possibilities:

1) The unique non-zero difference bits in β^{RL} and γ^{LL}, are in the same half of the input to $FI_{4,1}$.
2) The non-zero difference bits are not in the same half of the input to $FI_{4,1}$. Without loss of generality, in β^{RL} the non-zero bit is in the left half, and in γ^{LL}, the non-zero bit is in the right half.

[20] We note that this kind of computation is performed in [19] in order to compute the probability of the distinguisher in cases of dependence between the differentials.

Fig. 4. Example of a Failing Related-Key Boomerang



In order to reduce the amount of technicalities, we check all the $2 \cdot 63 = 126$ distinguishers that correspond to the second possibility. The analysis of the remaining differentials is, presumably, similar.

By the structure of the pairs of differentials that belong to the second class, the corresponding quartets of inputs to $FI_{4,1}$ (divided into the two "halves") are of the form:

$$(x, y, x, y), (z, z, w, w),$$

where x, y, z, w are mutually distinct (see Figure 4). It follows that the inputs to the S-box $S9$ in the first round of $FI_{4,1}$ are of the form (x, x, y, y), and the inputs to the S-box $S7$ in the second round of $FI_{4,1}$ are of the form (z, w, z, w). Hence, the corresponding outputs are of the forms (x', x', y', y') and (z', w', z', w'), respectively. Since both these quadruples are *balanced* (i.e., sum up to zero), and there is no key difference in $FI_{4,1}$, this implies that in both halves of the intermediate value after the key addition, the quadruples are balanced. Therefore, due to the 4-round Feistel structure, if for some pair of differentials, the outputs of the S-box $S9$ in the third round of $FI_{4,1}$ are *unbalanced*, this implies that the right half of the output of $FI_{4,1}$ is unbalanced, thus proving that inequality (22) holds and that the distinguisher fails.

Consider the four inputs to the S-box $S9$ in the third round of $FI_{4,1}$. By the Feistel structure, they are of the form

$$(x', x', y', y') \oplus (z, w, z, w) \oplus (KI_{4,1,2}, KI_{4,1,2}, KI_{4,1,2}, KI_{4,1,2}),$$

and hence, they are balanced.

The balancedness assures that the XOR of all values is indeed zero. At the same time, the values themselves can be distinct (i.e., when $x' \oplus z \neq y' \oplus w$), then four different values enter $S9$. As a design criteria, $S9$ is an *almost perfect non-linear permutation*, a property which implies that the four outputs are necessarily unbalanced, which lead to the failure of the distinguisher. These four inputs are not distinct only if $x' \oplus z = y' \oplus w$, or equivalently, $x' \oplus y' = z \oplus w$. By the definition of x', y', this can happen only if the differential $x \oplus y \to z \oplus w$ is possible through the S-box S9. This leads us to our first conclusion:

Conclusion *1:* Denote the basic differentials of rounds 1–3 and 5–7 used in the distinguisher by $\alpha \to \beta$ and $\gamma \to \delta$, respectively. Let the only nonzero bit in β^{RL} be bit $7+i$, and let the only nonzero bit in γ^{LL} be j. If the differential $e_i \to e_j$ (where e_i is the 9-bit vector in each the only nonzero position is the ith position) is impossible for the S-box S9, then the entire distinguisher fails.

Now assume that the four inputs to the S-box $S9$ in the third round of $FI_{4,1}$ are not distinct. In such case, the right half of the output of $FI_{4,1}$ is balanced. Hence, if the output of S7 in the fourth round of $FI_{4,1}$ is unbalanced, then the left half of the output of the entire $FI_{4,1}$ is unbalanced, and the distinguisher fails.

Consider the four inputs to the S-box S7 in the fourth round of $FI_{4,1}$. By the Feistel structure, they are of the form

$$(z', w', z', w') \oplus (x', x', y', y'), \oplus (z, w, z, w) \oplus (KI_{4,1,1}, KI_{4,1,1}, KI_{4,1,1}, KI_{4,1,1}),$$

and hence, they are balanced. As in the previous case, if they are distinct, and since $S7$ is an *almost perfect non-linear permutation* as well, the four outputs are necessarily unbalanced, and the distinguisher fails. On the other hand, since by the assumption, $x' \oplus z = y' \oplus w$, these inputs are not distinct only if $z' \oplus x' \oplus z = w' \oplus x' \oplus w$, or equivalently, $z' \oplus w' = z \oplus w$. However, by the definition of z', w', this can happen only if the differential $z \oplus w \to z \oplus w$ is possible through the S-box S7. This leads us to the second conclusion:

Conclusion *2:* Denote the basic differentials of rounds 1–3 and 5–7 used in the distinguisher by $\alpha \to \beta$ and $\gamma \to \delta$, respectively. Let the only nonzero bit in β^{RL} be bit $7+i$, and let the only nonzero bit in γ^{LL} be j. If the differential $e_j \to e_j$ (where e_j is the 7-bit vector in each the only nonzero position is the ith position) is impossible for the S-box S7, then the entire distinguisher fails.

We checked exhaustively all 126 pairs of differences (β, γ) that belong to the class we study, and found out that the only pairs of differences that satisfy the restrictions of the propositions are:

$$\beta = (0_x, 0001\ 0000_x), \qquad \gamma = (0_x, 0400\ 0000_x),$$

and:

$$\beta = (0_x, 0040\ 0000_x), \qquad \gamma = (0_x, 0080\ 0000_x),$$

with appropriately chosen key differences (and the pairs of differentials obtained from them by interchanging the roles of β and γ).

This means that out of the 126 checked pairs of differentials, only 4 can work theoretically! But on the other hand, as demonstrated above, an adversary can check the local inconsistencies manually, and then choose one of the "possible" pairs of differentials.

2) Verifying the transition of round 4 experimentally: As a second step in our analysis, we choose one of the four "possible" differentials, and verified whether there are no further inconsistencies. An easy analysis shows that such inconsistencies can occur only inside $FI_{4,3}$, and the inputs of the round undergo too many changes before that point, so that the values cannot be followed easily. Instead, we observe that if the boomerang passes the filter of $FI_{4,1}$ successfully, then inside $FI_{4,1}$, the differentials $e_i \rightarrow e_j$ through $S9$ and $e_j \rightarrow e_j$ through $S7$ are satisfied. Since $S9$ and $S7$ are *almost perfect nonlinear permutations*, there is only a single pair of inputs to $S9$ that satisfies the differential $e_i \rightarrow e_j$, and there is only a single pair of inputs to $S7$ that satisfies the differential $e_j \rightarrow e_j$. This allows us to *choose* the four inputs to round 4 such that the function $FI_{4,1}$ is passed "for free".

At this stage, we were ready to perform an experiment. We fixed the quartet (X_a, X_b, X_c, X_d) to be one of the quartets for which $FI_{4,1}$ is passed for free, and checked the probability of the condition

$$X_a^L \oplus X_b^L \oplus X_c^L \oplus X_d^L = 0,$$

when averaged over random values of the subkeys used in round 4.

The result of the experiment was that on average, the probability was indeed 2^{-16} as expected (since another 2^{-16} are "gained" by fixing the inputs), which proves that for the "correct" choice of differentials, the distinguisher does work. On the other hand, the experiment revealed that the probabilities depend quite heavily on the exact choices of the subkeys, which leads to a conjecture that the overall probability of the distinguisher is also key-dependent.

3) The full 7-round verification experiment: After verifying that the transition in the fourth round is feasible, we reached the point that we were ready to conduct a verification of the full 7-round distinguisher. We ran a full experiment simulating the RK-boomerang distinguisher for 7-round KASUMI, expecting that the probability of obtaining a right quartet is 2^{-39}.

For each of the 215 keys we have checked, we took 2^{39} quartets, and counted how many of them were right quartets. The results of the experiment are given in Table IV that lists how many right quartets were found.

It is interesting to note that only in 35 out of the 215 experiments we encountered right quartets. This may seem like a failure of the entire boomerang approach, but a further analysis shows that this is not the case. First of all, the distinguisher was found to be useful for about one out of six keys. These keys may be considered as a set of weak keys (i.e., a weak key class) of the cipher, but given its size, this set cannot be disregarded.

Moreover, we note that for the keys for which right quartets were found, significantly more quartets than expected were found. Namely, it seems that while there are keys for which the distinguisher fails, for the keys for which it works, it works significantly better than predicted. For comparison, the probability of an experiment that follows the Poisson distribution with mean value of 1 to obtain 39 "successes" is $\frac{e^{-39}}{39!} \approx 2^{-153.8}$.

We also note that given the restrictions on the computational power we had at our disposal (obtaining the required $215 \cdot 2^{39}$ quartets, which are equivalent to $2^{48.7}$ encryptions using the official KASUMI reference implementation took over a month in three different clusters), we expect some experiments to obtain no quartets, following the Poisson distribution (when the mean value is 1, about $1/e$ of the experiments are expected to have no quartets following the randomness).

Finally, we note that checking what is the exact number of keys for which the attack succeeds requires testing significantly more quartets per key guess to overcome the random nature of the process. However, given the huge computation requirements, this task seems out of our reach at the moment.

IV. Conclusions

In the first part of this paper we presented a rigorous treatment of the related-key boomerang and related-key rectangle attacks. We devised optimal algorithms for the RK-boomerang/rectangle distinguishers and computed their success probability under explicitly stated and analyzed independence assumptions.

In the second part of this paper we presented an extensive experimental analysis of the RK-boomerang attack in the specific case of the block cipher KASUMI. Our experiments (along with previous experimentally verified results) suggest the following heuristics:

- Boomerang-type attacks can fail due to local inconsistencies, especially in the transition between the subciphers. Hence, the designers of attacks should do their best to check that the distinguisher used in the attack does not contain any inconsistency.
- If the probabilities of any round in the differentials used in the distinguisher are not extremely low, it is reasonable to assume that the independence assumptions underlying the boomerang-type attacks are valid.
- If the probabilities of some part of the differentials is very low, then the overall probability of the distinguisher can depend heavily on the key, such that the distinguisher

TABLE IV
THE NUMBER OF FOUND QUARTETS IN 215 EXPERIMENTS

Quartets	0	2	4	5	6	7	9	10	11	12	13	14	15	16	17	22	23	28	34	35	39
Experiments	180	2	3	1	2	1	1	1	2	1	3	3	2	3	2	2	1	2	1	1	1

applies only for a relatively small portion of the keys.[21]

- In any case, it is very important to check the probability of the RK-boomerang/rectangle distinguisher used in each specific attack, whenever possible.

Apart from the immediate attacks, another outcome of the related-key boomerang and rectangle techniques is a better understanding of the importance of a well designed key schedule algorithm for the security of block ciphers. While it is commonly believed that a linear key schedule (or one close to it), is of no security concern to a well designed block cipher, the related-key boomerang and rectangle attacks, along with the concept of structures of keys (that allows to bypass nonlinear key schedule algorithms) show that this belief is dangerous and at times may be faulty.

ACKNOWLEDGEMENTS

We would like to thank Charles Bouillaguet and École Normale Supérieure in France for granting us access to their computing services. We would also like to thank the anonymous referees for their valuable comments.

REFERENCES

[1] Thomas Baignères, Pascal Junod, and Serge Vaudenay, *How Far Can We Go Beyond Linear Cryptanalysis*, Advances in Cryptology, proceedings of ASIACRYPT 2004, Lecture Notes in Computer Science 3329, pp. 432–450, Springer-Verlag, 2004.

[2] Mihir Bellare and Tadayoshi Kohno, *A Theoretical Treatment of Related-Key Attacks: RKA-PRPs, RKA-PRFs, and Applications*, Advances in Cryptology, proceedings of EUROCRYPT 2003, Lecture Notes in Computer Science 2656, pp. 491–506, Springer-Verlag, 2003.

[3] Ishai Ben-Aroya and Eli Biham, *Differential Cryptanalysis of Lucifer*, Advances in Cryptology, proceedings of EUROCRYPT 1993, Lecture Notes in Computer Science 773, pp. 187–199, Springer-Verlag, 1994.

[4] Eli Biham, *New Types of Cryptanalytic Attacks Using Related Keys*, Journal of Cryptology, vol. 7, no. 4, pp. 229–246, Springer-Verlag, 1994.

[5] Eli Biham and Adi Shamir, *Differential Cryptanalysis of DES-like Cryptosystems*, Advances in Cryptology, proceedings of CRYPTO 1990, Lecture Notes in Computer Science 537, pp. 2–21, Springer-Verlag, 1990.

[6] Eli Biham and Adi Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.

[7] Eli Biham, Orr Dunkelman, and Nathan Keller, *The Rectangle Attack — Rectangling the Serpent*, Advances in Cryptology, proceedings of EUROCRYPT 2001, Lecture Notes in Computer Science 2045, pp. 340–357, Springer-Verlag, 2001.

[8] Eli Biham, Orr Dunkelman, and Nathan Keller, *New Results on Boomerang and Rectangle Attacks*, proceedings of Fast Software Encryption 2002, Lecture Notes in Computer Science 2365, pp. 1–16, Springer-Verlag 2002.

[9] Eli Biham, Orr Dunkelman, and Nathan Keller, *Related-Key Boomerang and Rectangle Attacks*, Advances in Cryptology, proceedings of EUROCRYPT 2005, Lecture Notes in Computer Science 3494, pp. 507–525, Springer-Verlag, 2005.

[10] Eli Biham, Orr Dunkelman, and Nathan Keller, *A Related-Key Rectangle Attack on the Full KASUMI*, Advances in Cryptology, proceedings of ASIACRYPT 2005, Lecture Notes in Computer Science 3788, pp. 443–461, Springer-Verlag, 2005.

[21]This fact is supported by similar results concerning differential cryptanalysis [18].

[11] Eli Biham, Orr Dunkelman, and Nathan Keller, *New Cryptanalytic Results on IDEA*, Advances in Cryptology, proceedings of ASIACRYPT 2006, Lecture Notes in Computer Science 4284, pp. 412–427, Springer-Verlag, 2006.

[12] Eli Biham, Orr Dunkelman, and Nathan Keller, *A Unified Framework for Related-Key Attacks*, proceedings of Fast Software Encryption 2008, Lecture Notes in Computer Science 5086, pp. 73–96, Springer-Verlag, 2008.

[13] Alex Biryukov, *The Boomerang Attack on 5 and 6-Round AES*, Proceedings of Advance Encryption Standard Fourth Workshop, Lecture Notes in Computer Science 3373, pp. 11–16, Springer-Verlag, 2005.

[14] Alex Biryukov and Dmitry Khovratovich, *Related-key Cryptanalysis of the Full AES-192 and AES-256*, Advances in Cryptology, proceedings of ASIACRYPT 2009, Lecture Notes in Computer Science 5912, pp. 1–18, Springer-Verlag, 2009.

[15] Alex Biryukov, Orr Dunkelman, Nathan Keller, Dmitry Khovratovich, and Adi Shamir, *Key Recovery Attacks of Practical Complexity on AES-256 Variants with up to 10 Rounds*, Advances in Cryptology, proceedings of EUROCRYPT 2010, Lecture Notes in Computer Science 6110, pp. 299–319, Springer-Verlag, 2010.

[16] Mark Blunden and Adrian Escott, *Related Key Attacks on Reduced Round KASUMI*, proceedings of Fast Software Encryption 2001, Lecture Notes in Computer Science 2355, pp. 277–285, Springer-Verlag, 2002.

[17] Joan Daemen and Vincent Rijmen, *The design of Rijndael: AES — the Advanced Encryption Standard*, Springer-Verlag, 2002.

[18] Joan Daemen and Vincent Rijmen, *Understanding Two-Round Differentials in AES*, proceedings of Security and Cryptography for Networks 2006, Lecture Notes in Computer Science 4116, pp. 78–94, Springer-Verlag, 2006.

[19] Orr Dunkelman, Nathan Keller, and Adi Shamir, *A Practical-Time Related-Key Attack on the KASUMI Cryptosystem Used in GSM and 3G Telephony*, Advances in Cryptology, proceedings of CRYPTO 2010, Lecture Notes in Computer Science 6223, pp. 393–410, Springer-Verlag, 2010.

[20] Michael Gorski and Stefan Lucks, *New Related-Key Boomerang Attacks on AES*, proceedings of INDOCRYPT 2008, Lecture Notes in Computer Science 5365, pp. 266–278, Springer-Verlag, 2008.

[21] Seokhie Hong, Jongsung Kim, Guil Kim, Sangjin Lee, and Bart Preneel, *Related-Key Rectangle Attacks on Reduced Versions of SHACAL-1 and AES-192*, proceedings of Fast Software Encryption 2005, Lecture Notes in Computer Science 3557, pp. 368–383, Springer-Verlag, 2005.

[22] Goce Jakimoski and Yvo Desmedt, *Related-Key Differential Cryptanalysis of 192-bit Key AES Variants*, proceedings of Selected Areas in Cryptography 2003, Lecture Notes in Computer Science 3006, pp. 208–221, Springer-Verlag, 2004.

[23] John Kelsey, Bruce Schneier, and David Wagner, *Key Schedule Cryptanalysis of IDEA, G-DES, GOST, SAFER, and Triple-DES*, Advances in Cryptology, proceedings of CRYPTO 1996, Lecture Notes in Computer Science 1109, pp. 237–251, Springer-Verlag, 1996.

[24] John Kelsey, Bruce Schneier, and David Wagner, *Related-Key Cryptanalysis of 3-WAY, Biham-DES, CAST, DES-X, NewDES, RC2, and TEA*, proceedings of Information and Communications Security 1997, Lecture Notes in Computer Science 1334, pp. 233–246, Springer-Verlag, 1997.

[25] John Kelsey, Tadayoshi Kohno, and Bruce Schneier, *Amplified Boomerang Attacks Against Reduced-Round MARS and Serpent*, proceedings of Fast Software Encryption 2001, Lecture Notes in Computer Science 1978, pp. 75–93, Springer-Verlag, 2002.

[26] Jongsung Kim, Guil Kim, Seokhie Hong, and Dowon Hong, *The Related-Key Rectangle Attack — Application to SHACAL-1*, proceedings of Australasian Conference on Information Security and Privacy 2004, Lecture Notes in Computer Science 3108, pp. 123–136, Springer-Verlag, 2004.

[27] Jongsung Kim, Seokhie Hong, and Bart Preneel, *Related-Key Rectangle Attacks on Reduced AES-192 and AES-256*, proceedings of FSE 2007, Lecture Notes in Computer Science 4593, pp. 225–241, Springer-Verlag, 2007.

[28] Lars R. Knudsen, *Cryptanalysis of LOKI91*, proceedings of Auscrypt

1992, Lecture Notes in Computer Science 718, pp. 196–208, Springer-Verlag, 1993.
[29] Ulrich Kühn, *Cryptanalysis of Reduced-Round MISTY*, Advances in Cryptology, proceedings of EUROCRYPT 2001, Lecture Notes in Computer Science 2045, pp. 325–339, Springer-Verlag, 2001.
[30] Xuejia Lai, James L. Massey, and Sean Murphy, *Markov Ciphers and Differential Cryptanalysis*, Advances in Cryptology, proceedings of EUROCRYPT 1991, Lecture Notes in Computer Science 547, pp. 17–38, Springer-Verlag, 1992.
[31] Mitsuru Matsui, *Linear Cryptanalysis Method for DES Cipher*, Advances in Cryptology, proceedings of EUROCRYPT 1993, Lecture Notes in Computer Science 765, pp. 386–397, Springer-Verlag, 1994.
[32] Sean Murphy, *The Return of the Cryptographic Boomerang*, IEEE Transactions on Information Theory vol. 57, no. 4, pp. 2517–2521, 2011.
[33] National Bureau of Standards, *Data Encryption Standard*, Federal Information Processing Standards Publications no. 46, 1977.
[34] National Institute of Standards and Technology, *Advanced Encryption Standard*, Federal Information Processing Standards Publications no. 197, 2001.
[35] Kaisa Nyberg, *Perfect Nonlinear S-boxes*, Advances in Cryptology, proceedings of EUROCRYPT 1991, Lecture Notes in Computer Science 547, pp. 378–386, Springer-Verlag, 1991.
[36] Kaisa Nyberg and Lars R. Knudsen, *Provable Security Against Differential Cryptanalysis*, Advances in Cryptology, proceedings of CRYPTO 1992, Lecture Notes in Computer Science 740, pp. 566–578, Springer-Verlag, 1993.
[37] Ali Aydin Selçuk, *On Probability of Success in Linear and Differential Cryptanalysis*, Journal of Cryptology, vol. 21 no. 1, pp. 131–147, Springer-Verlag, 2008.
[38] 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, 3G Security, *Specification of the 3GPP Confidentiality and Integrity Algorithms; Document 2: KASUMI Specification*, V.3.1.1, 2001.
[39] Gene Tsudik and Els Van Herreweghen, *On simple and secure key distribution*, Conference on Computer and Communications Security, Proceedings of the 1st ACM conference on Computer and communications security, pp. 49–57, ACM, 1993.
[40] David Wagner, *The Boomerang Attack*, proceedings of Fast Software Encryption 1999, Lecture Notes in Computer Science 1636, pp. 156–170, Springer-Verlag, 1999.
[41] Gaoli Wang, Nathan Keller, and Orr Dunkelman, *The Delicate Issues of Addition with Respect to XOR Differences*, proceedings of Selected Areas in Cryptography 2007, Lecture Notes in Computer Science 4876, pp. 212–231, Springer-Verlag, 2007.
[42] David J. Wheeler and Roger M. Needham, *TEA, a Tiny Encryption Algorithm*, proceedings of Fast Software Encryption 1994, Lecture Notes in Computer Science 1008, pp. 363–366, Springer-Verlag, 1995.
[43] ZDNet, New Xbox security cracked by Linux fans, 2002. Available online at `http://news.zdnet.co.uk/software/developer/0,39020387,2123851,00.htm`.