Noname manuscript No.
(will be inserted by the editor)

Revisiting Iterated Attacks in the Context of Decorrelation Theory

Aslı Bay · Atefeh Mashatan · Serge Vaudenay

Received: / Accepted:

Abstract Iterated attacks are comprised of iterating adversaries who can make d plaintext queries, in each iteration to compute a bit, and are trying to distinguish between a random cipher C and the perfect cipher C^* based on all bits.

Vaudenay showed that a $2d$-decorrelated cipher resists to iterated attacks of order d when iterations have almost no common queries. Then, he first asked what the necessary conditions are for a cipher to resist a non-adaptive iterated attack of order d. I.e., whether decorrelation of order $2d-1$ could be sufficient. Secondly, he speculated that repeating a plaintext query in different iterations does not provide any advantage to a non-adaptive distinguisher. We close here these two long-standing open problems negatively. For those questions, we provide two counter-intuitive examples.

We also deal with adaptive iterated adversaries who can make both *plaintext* and *ciphertext* queries in which the future queries are dependent on the past queries. We show that decorrelation of order $2d$ protects against these attacks of order d. We also study the generalization of these distinguishers for iterations making non-binary outcomes.

Finally, we measure the resistance against two well-known statistical distinguishers, namely, differential-linear and boomerang distinguishers and show that 4-decorrelation degree protects against these attacks.

Keywords block ciphers · decorrelation theory · iterated attacks · differential-linear distinguishers · boomerang distinguishers

1 Introduction

Unlike asymmetric cryptography, in which the security of a cryptosystem is provably reduced to a mathematical problem and guaranteed by an intractability assumption, the focus in symmetric cryptography is often statistical cryptanalysis and, in the absence of a successful attack, a cryptosystem is believed to be secure. For instance, once the crypto community has spent enough time scrutinizing a block cipher and has found no successful attacks against its full round version, the block cipher is believed to be secure. However, a different approach against block cipher cryptanalysis was pioneered by Nyberg [17]. She

Part of the results in this work were published in [5,6]. In this paper, the study of generalization of adaptive iterated plaintext-ciphertext distinguishers is added; also the resistance against boomerang and differential-linear distinguishers are analyzed by the techniques in Decorrelation Theory.

Aslı Bay
EPFL, Switzerland
Tel.: +41-21-6937617
Fax: +41-21-6937689
E-mail: asli.bay@epfl.ch

Atefeh Mashatan
Security Engineering, Canadian Imperial Bank of Commerce (CIBC), Canada

Serge Vaudenay
EPFL, Switzerland

formalized the notion of strength against differential cryptanalysis. Her work was followed by Chabaud and Vaudenay [9] formalizing the notion of strength against linear cryptanalysis.

Decorrelation Theory, introduced by Vaudenay [22,25], encapsulates the techniques that guarantee the *provable* resistance of block ciphers against a wide range of statistical cryptanalysis, including the seminal differential and linear attacks, as well as their variants, for example the boomerang attack, truncated differential attacks, and impossible differential attacks. The beauty of this theory is that it can even guarantee resistance against some not-yet-discovered attacks that meet a certain broad criteria in the model presented by Luby and Rackoff [14,15]. They prove the security of Feistel schemes by assuming that the round function is random. However, their approach needs a very long secret key and is not suitable in practice. Carter and Wegman [7,8], on the other hand, use derandomization techniques for sampling pairwise independent numbers, which has inspired the notion of decorrelation in that it measures the pseudorandomness with smaller keys and examines its effects against the adversaries.

It is worth mentioning here that *perfect* decorrelation of order d is equivalent to *d-wise independence* [13]. Moreover, decorrelation of order d is also referred to as *almost* d-wise independence [1,16]. Furthermore, the concept of decorrelation is somewhat related to the notion of pseudorandom functions and pseudorandom permutations except that we do not limit the time complexity of the distinguisher, but only the number of queries are restricted.

The adversaries considered here can query d plaintexts and receive their corresponding ciphertexts, but are unlimited in terms of computational power. When these plaintext/ciphertext pairs are chosen randomly and independently from each other, we are dealing with *non-adaptive* d-limited adversaries, as opposed to *adaptive* d-limited adversaries. These adversaries give rise to distinguishers of order d, whether adaptive or otherwise, who are trying to distinguish between a random cipher C and the perfect cipher C^*.

More formally, to a random cipher C, we define a d-wise distribution matrix $[C]^d$ (see Definition 3). The decorrelation of C is the distance $\|[C]^d - [C^*]^d\|$ between $[C]^d$ and $[C^*]^d$ defined by a matrix-norm $\|\cdot\|$ (see Definition 2).

Several block ciphers have been designed, whose security is proven by decorrelation techniques, see for example DFC [18], NUT (n-Universal Transformation) families of block ciphers [10,19,21,25]. Using similar techniques, Baignères and Finiasz propose two provably secure block ciphers to use in practice called the block cipher C [3] and KFC [2]. Decorrelation Theory has been used in other results as well, see for instance [4,20,22–24].

Vaudenay [25] shows how differential and linear attacks fit in the d-limited adversarial model by introducing *iterated* attacks, which are simply constructed by iterating a non-adaptive d-limited distinguisher (see Algorithm 3). Linear and differential cryptanalysis can be formulated as non-adaptive iterated attacks of order 1 and order 2, respectively, and the boomerang attack is an adaptive (chosen plaintext and ciphertext) iterated attack of order 4. Moreover, he computes a bound on the advantage of the d-limited adversaries by decorrelation techniques in the Luby-Rackoff model. This result is expressed in the following theorem.

Theorem 1 *[25] Let C be a cipher on a message space of cardinality M such that $\|[C]^{2d} - [C^*]^{2d}\|_\infty \leq \varepsilon$, for some given integer $d \leq M/2$, where C^* is the perfect cipher. Let us consider a non-adaptive iterated distinguisher of order d between C and C^* with n iterations. We assume that a set of d plaintexts is generated in each iteration in an independent way and following the same distribution. Moreover, we define δ as the probability that two sets drawn with this distribution have a nonempty intersection. Then, we bound the advantage of the adversary as*

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{NAI}(d)}} \leq 5\sqrt[3]{\Big(2\delta + \frac{5d^2}{2M} + \frac{3\varepsilon}{2}\Big)n^2} + n\varepsilon.$$

This theorem shows that, in order to resist a non-adaptive iterated attack of order d with seldom common queries, it is *sufficient* for a cipher to have the decorrelation of order $2d$. However, whether or not this is a necessary condition has not been addressed. Moreover, the bound given in the theorem can be interpreted to imply that, perhaps, a high probability δ of having a common query increases the bound of the attack. Despite this hint, Vaudenay in his EUROCRYPT '99 paper [22] speculates that having the same query to the oracle does not provide any advantage, but whether or not this is true has been left open. We will settle both of these open questions.

Our Contribution. Firstly, we show that the decorrelation of order $2d-1$ is not sufficient. We achieve this by proposing a counterexample consisting of a 3-round Feistel cipher which is decorrelated to the order $2d-1$ and, *yet*, we are able to mount a successful non-adaptive iterated distinguisher of order d against it. Secondly, we propose another set of counterexamples where a higher probability of having common queries surprisingly increases the advantage of the distinguisher. In particular, we show that there is an iterated distinguisher of order 1 on a $2d$-decorrelated cipher when the probability of having at least one query in common in any two iterations is high, which is counterintuitive.

As a second contribution, we concentrate on *adaptive* iterated distinguishers who can adaptively make plaintext and ciphertext queries. Allowing the adversary to make both plaintext and ciphertext queries extends the security model and has already appeared in the literature. Indeed, the boomerang attack [26] is an example of such an adversary. Studying these general distinguishers making adaptive plaintext-ciphertext queries allows us to, for example, interpret Wagner's boomerang attack [26] on COCONUT98 [19,25], a perfect 2-decorrelated block cipher and *provably secure* against differential and linear cryptanalyses and iterated attacks of order 1. Indeed, it could have resisted to Wagner's attack with a decorrelation of order 8.

A bound for the advantage of an *adaptive* iterated distinguisher of order d, who can make both *plaintext* and *ciphertext* queries has not been computed yet. The significance of studying these distinguishers is not hidden to anyone. Therefore, we prove the bound for the advantage of adaptive iterated distinguisher of order d against a $2d$-decorrelated cipher. It comes with no surprise that using this metric, we get a looser, i.e., higher, upper bound for adaptive distinguishers than that for non-adaptive distinguishers (see Theorem 1). We then generalize these distinguishers in a way that the outputs of iterations are not binary anymore, take values from a set of integers, and prove the bound for this generalized version, as well.

Finally, we measure the resistance against two well-known statistical distinguishers, namely, differential-linear and boomerang distinguishers. We show that a cipher resisting these distinguishers needs to be exactly 4-decorrelated. These results are better than the results obtained in Theorems 1 and 5 since a cipher requires to be at most 4-decorrelated and 8-decorrelated according to these theorems, respectively. So, a 4-decorrelation of COCONUT98 would have been enough, actually, to defeat Wagner's attack.

The rest of the paper is organized as follows. Section 2 gives basics of Decorrelation Theory. We dedicate Sections 3, 4, 5 and 6 to our main contributions.

2 Basics of Decorrelation Theory

In this paper, F denotes a random function (or equivalently a function set up with a random key) from $\mathcal{M}_1$ to $\mathcal{M}_2$ and F^* denotes the ideal random function from $\mathcal{M}_1$ to $\mathcal{M}_2$, that is, a function drawn uniformly at random among all $|\mathcal{M}_2|^{|\mathcal{M}_1|}$ functions on the given sets. Similarly, C denotes a random cipher (or equivalently, the encryption function set up with a random key) over $\mathcal{M}_1$ and C^* denotes the ideal random cipher over $\mathcal{M}_1$, that is, a permutation drawn uniformly at random among all $|\mathcal{M}_1|!$ permutations. We use the following standard notations: $|S|$ denotes the cardinality of the set S; $\mathcal{M}^d$ is the set of all sequences of d tuples over the set $\mathcal{M}$; $\mathsf{GF}(q)$ is the finite field with q elements; $\mathsf{GF}(q)[x]$ is the set of polynomials defined over $\mathsf{GF}(q)$; $\mathbb{E}(X)$ denotes the expected value of the random variable X; $V(X)$ is the variance of the random variable X; $\mathsf{gcd}(p(x), q(x))$ denotes the greatest common divisor of $p(x)$ and $q(x)$; and "$\oplus$" denotes addition modulo 2.

We consider the *Luby-Rackoff model* [15] in which an adversary $\mathcal{A}$ is unbounded in terms of *computational power*. It is bounded to d number of plaintext/ciphertext queries to an oracle Ω implementing a random function (resp. a random cipher). The goal of the adversary $\mathcal{A}$ is to guess whether this function (resp. cipher) is drawn following the distribution of F (resp. C) or of F^* (resp. C^*). When queries are chosen randomly and at once, such an adversary is exactly a *non-adaptive* d-limited distinguisher. However, when queries are chosen depending on the answers to the previous queries, it is referred to as an *adaptive* d-limited distinguisher. In both distinguishers, the measure of success of $\mathcal{A}$ is computed by means of the *advantage* of the adversary.

Definition 1 Let F_0 and F_1 be two random functions. The *advantage* of an adversary $\mathcal{A}$ distinguishing F_0 from F_1 is defined by

$$\mathrm{Adv}_{\mathcal{A}}(F_0, F_1) = \big|\Pr[\mathcal{A}(F_0) = 1] - \Pr[\mathcal{A}(F_1) = 1]\big|.$$

Another measure is the *best advantage* of the distinguisher which is formulated as

$$\mathrm{BestAdv}_{\zeta}(F_0, F_1) = \max_{\mathcal{A}\in\zeta} \mathrm{Adv}_{\mathcal{A}}.$$

Here, the maximum is taken over adversaries in a class ζ. For instance, ζ can consist of all non-adaptive or all adaptive d-limited distinguishers, denoted by $\mathcal{A}_{\mathsf{NA}(d)}$ and $\mathcal{A}_{\mathsf{A}(d)}$, respectively, between F_0 and F_1 depending on $\mathcal{A}$ being non-adaptive or adaptive.

Decorrelation Theory has a link with Linear and Differential Cryptanalyses (see Algorithms 1 and 2) which are the essential cryptanalysis methods of both block ciphers and pseudorandom functions. Both methods have iterative analysis of an instance of a block cipher and refer to the set of attacks called *iterated attacks*. More explicitly, iterated attacks are defined as iterations of d-limited distinguishers. When *non-adaptive* d-limited distinguishers are iterated, we obtain *non-adaptive* iterated distinguishers of order d. When *adaptive* d-limited distinguishers are iterated, we get *adaptive* iterated distinguishers of order d. A generic non-adaptive iterated distinguisher of order d is illustrated in Algorithm 3. Briefly, a test $\mathcal{T}$ generates the binary output T_i of each iteration i, and then the distinguisher outputs his final decision based on the acceptance set $\mathcal{A}cc$ and the tuple $(T_1, \ldots, T_n)$. Linear and differential cryptanalyses are the examples for non-adaptive iterated attacks of order 1 and 2, respectively. Their combinations called differential-linear distinguishers, described in Algorithm 4, are non-adaptive iterated attacks of order 2. In addition, the boomerang attack is an adaptive iterated attack of order 4 (with chosen plaintexts and ciphertexts).

Algorithm 1 Linear Distinguisher

Input: an integer n, a set X, a set I, masks a and b
Oracle: an oracle Ω implementing a permutation c
for $i = 1$ to n **do**
  pick x_1 uniformly at random from X
  set $y_1 = c(x_1)$
  set $T_i = a \cdot x_1 \oplus b \cdot y_1$
end for
if $T_1 + \cdots + T_n \in I$ **then**
  output 1
else
  output 0
end if

Algorithm 2 Differential Distinguisher

Input: an integer n, a set X, differences α and β
Oracle: an oracle Ω implementing a permutation c
for $i = 1$ to n **do**
  pick x_1 uniformly at random over X
  set $x_2 = x_1 \oplus \alpha$
  set $y_1 = c(x_1)$, $y_2 = c(x_2)$
  set $T_i = \mathbf{1}_{y_1 \oplus y_2 = \beta}$
end for
if $T_1 + \cdots + T_n \neq 0$ **then**
  output 1
else
  output 0
end if

We now recall two matrix-norms which are used in Decorrelation Theory.

Definition 2 Let $\mathrm{M} \in \mathbb{R}^{|\mathcal{M}_1|^d \times |\mathcal{M}_2|^d}$ be a matrix. Then, two matrix-norms are defined by

$$\|M\|_{\infty} = \max_{x_1,\ldots,x_d} \sum_{y_1,\ldots,y_d} |M_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)}|$$

and

$$\|M\|_{A} = \max_{x_1} \sum_{y_1} \cdots \max_{x_d} \sum_{y_d} |M_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)}|.$$

Algorithm 3 A generic non-adaptive iterated distinguisher of order d

Input: an integer n, a set X, a distribution $\mathcal{X}$ on X, a test $\mathcal{T}$, a set $\mathcal{A}cc$
Oracle: an oracle Ω implementing a permutation c

for $i = 1$ to n **do**
  pick $x = (x_1, \ldots, x_d)$ at random from a distribution $\mathcal{X}$
  set $y = (c(x_1), \ldots, c(x_d))$
  set $T_i = 0$ or 1 such that $T_i = \mathcal{T}(x, y)$
end for
if $(T_1, \ldots, T_n) \in \mathcal{A}cc$ **then**
  output 1
else
  output 0
end if

Algorithm 4 Differential-Linear Distinguisher

Input: an integer n, a set X, a set I, masks a and b, differences α and β
Oracle: an oracle Ω implementing a permutation c

for $i = 1$ to n **do**
  pick x_1 uniformly at random over X
  set $x_2 = x_1 \oplus \alpha$
  set $y_1 = c(x_1)$ and $y_2 = c(x_2)$
  set $T_i = b \cdot y_1 \oplus b \cdot y_2$
end for
if $T_1 + \cdots + T_n \in I$ **then**
  output 1
else
  output 0
end if

The use of these matrix-norms depends on the type of attacks. The first norm which is known as L_∞-matrix-norm is used for non-adaptive distinguishers, while the second matrix-norm which is called the *adaptive matrix-norm* or $\|\cdot\|_A$-matrix-norm by Vaudenay [23], is for adaptive distinguishers. The reasonings are obvious as the former matrix-norm maximizes the sum over all input tuples when they are chosen at once, while the latter one maximizes the sum according to the inputs which are chosen dependently on the previous inputs.

We now recall a fundamental notion of Decorrelation Theory, the *d-wise distribution matrix* of a random function or a random permutation.

Definition 3 ([25]) Let F be a random function from $\mathcal{M}_1$ to $\mathcal{M}_2$. The *d-wise distribution matrix* $[F]^d$ of F is a $|\mathcal{M}_1|^d \times |\mathcal{M}_2|^d$-matrix and is defined by

$$[F]^d_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)} = \Pr_F[F(x_1) = y_1, \ldots, F(x_d) = y_d],$$

where $x = (x_1, \ldots, x_d) \in \mathcal{M}_1^d$ and $y = (y_1, \ldots, y_d) \in \mathcal{M}_2^d$.

Intuitively, a distribution matrix of an arbitrary function F (resp. cipher C) helps to analyze the function and show how far it is from its ideal version F^* (resp. C^*). Mainly, the distribution of $(F(x_1), \ldots, F(x_d))$ is corresponding to each row of the d-wise distribution matrix.

The distribution matrix of a random function lets us define the *d-wise decorrelation or equivalently the decorrelation of order d*. Intuitively, the *decorrelation of order d of a random function* F is defined as the distance $\mathcal{D}([F]^d, [F^*]^d)$ between its d-wise distribution matrix and the d-wise distribution matrix of the perfect function F^*. If the decorrelation distance is zero, that is $\mathcal{D}([F]^d, [F^*]^d) = 0$, we call F as the *perfect* d-decorrelated function. Intuitively, the distribution of both $(F(x_1), \ldots, F(x_d))$ and $(F^*(x_1), \ldots, F^*(x_d))$ are equal. If the decorrelation distance is small enough, then we call this function a d-decorrelated function. Note that we use a d-decorrelated function and a function decorrelated to the order d, interchangeably in the paper. To generalize this, we consider $\mathcal{D}([F_1]^d, [F_2]^d)$ which is called *d-wise decorrelation distance between* F_1 *and* F_2, given two random functions F_1 and F_2.

Next, the advantage of the best distinguisher is computed.

Theorem 2 (Theorems 10 and 11 in [25]) *Let F and F^* be a random function and the ideal random function, respectively. The respective advantages of the* best *non-adaptive and adaptive d-limited distinguishers, $\mathcal{A}_{\mathsf{NA}(d)}$ and $\mathcal{A}_{\mathsf{A}(d)}$, are*

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{NA}(d)}}(F, F^*) = \frac{1}{2}\|[F]^d - [F^*]^d\|_\infty$$

and

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{A}(d)}}(F, F^*) = \frac{1}{2}\|[F]^d - [F^*]^d\|_A.$$

Furthermore, Luby and Rackoff prove the security of a 3-round Feistel scheme when the round functions are perfect functions as follows.

Theorem 3 *([15]) Let F_1, F_2 and F_3 be independent and uniformly distributed functions over a set $\mathcal{M}_1$. Consider a 3-round Feistel cipher $C = \psi(F_1, F_2, F_3)$ on $\mathcal{M}_1^2$ as in Figure 1 and the perfect cipher C^*. The advantage of any distinguisher $\mathcal{A}$ distinguishing between C and C^* which is limited to d queries, $d > 0$ is an integer, is*

$$\mathrm{Adv}_{\mathcal{A}} \leq d^2/\sqrt{M},$$

where $M = |\mathcal{M}_1|^2$.

Lastly, Theorem 1 provides a bound for the advantage of a distinguisher against random permutations. We provide the following theorem for the case of random functions which proves a a tighter bound for the advantage.

Theorem 4 *Let F be a random function from $\mathcal{M}_1$ to $\mathcal{M}_2$, where $d \leq |\mathcal{M}_1|/2$ and $|\mathcal{M}_2| = N$. Assume that F is decorrelated to the order $2d$ by $\|[F]^{2d} - [F^*]^{2d}\|_\infty \leq \varepsilon$, where F^* is the perfect function. We consider a non-adaptive iterated distinguisher of order d between F and F^* with n iterations. We assume that a set of d plaintexts is generated in each iteration in an independent way and following the same distribution. Moreover, we define δ as the probability that two sets drawn with this distribution have a nonempty intersection. Then, we bound the advantage of the adversary as*

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{NAI}(d)}} \leq 5\sqrt[3]{\left(2\delta + \frac{3\varepsilon}{2}\right)n^2} + n\varepsilon.$$

Proof This proof is exactly the same as the proof of Theorem 1 [25] except for the computation of $V(T(F^*))$ which results in a tighter bound for the distinguisher. Given $F = f$ (resp. $F^* = f$), let $T(f)$ be the probability that test function $\mathcal{T}$ outputs 1 when $(X, f(X))$ is its input, i.e. $T(f) = \mathbb{E}_X(\mathcal{T}(X, f(X)))$, where X is input to the oracle. Let p (resp. p^*) be the probability that the distinguisher outputs 1, i.e. $p = \mathsf{Pr}_F[(T_1(F), \ldots, T_n(F)) \in \mathcal{A}cc]$, where $\mathcal{A}cc$ is the acceptance set and $T_i(F)$ (resp. $T_i(F^*)$) is the output of iteration i.

As $T_i(f)$'s are all independent once f is fixed and have the same expected value $T(f)$, we get

$$p = \mathbb{E}_F\left(\sum_{(t_1,\ldots,t_n)\in\mathcal{A}cc} T(F)^{t_1+\cdots+t_n}(1 - T(F))^{n-(t_1+\cdots+t_n)}\right).$$

Then, p can be rewritten as

$$p = \sum_{i=0}^{n} a_i \mathbb{E}_F(T(F)^i(1 - T(F))^{n-i})$$

for some integers a_i such that $0 \leq a_i \leq \binom{n}{i}$. Therefore, the advantage $|p - p^*|$ is maximal when all a_i's are either 0 or $\binom{n}{i}$ depending on the distributions $T(F)$ and $T(F^*)$. This implies that the acceptance set of the best distinguisher is of the form $\mathcal{A}cc = \{(t_1, \ldots, t_n) | \sum_{i=1}^n t_i \in \mathcal{B}\}$ for some set $\mathcal{B} \subseteq \{0, \ldots, n\}$. Therefore, we have $p = \mathbb{E}_F(s(T(F)))$, where $s(x) = \sum_{i\in\mathcal{B}} \binom{n}{i} x^i(1 - x)^{n-i}$.

Now, we compute the derivative of s as

$$s'(x) = \sum_{i\in\mathcal{B}} \binom{n}{i} \frac{i - nx}{x(1 - x)} x^i(1 - x)^{n-i}.$$

Notice that, as the sum over all i, such that $0 \leq i \leq n$, is the derivative of $(x + (1-x))^n$, the total sum is zero.

Hence, we get

$$|s'(x)| \leq \sum_{nx \leq i \leq n} \binom{n}{i} \frac{i - nx}{x(1-x)} x^i (1-x)^{n-i} \leq \frac{n}{x} \sum_{nx \leq i \leq n} \binom{n}{i} x^i (1-x)^{n-i},$$

as $nx \leq i \leq n$. Notice that when $x \geq 1/2$, we have $|s'(x)| \leq 2n$. Similarly, when $x < 1/2$, we have $|s'(x)| \leq 2n$. Thus, we obtain $|s'(x)| \leq 2n$, for every x. So, according to the Mean Value Theorem, we have

$$|s(T(F)) - s(T(F^*))| \leq 2n|T(F) - T(F^*)|.$$

According to Theorem 2, we have $|\mathbb{E}_F(T(F)) - \mathbb{E}_{F^*}(T(F^*))| \leq \varepsilon/2$, $|\mathbb{E}_F(T^2(F)) - \mathbb{E}_{F^*}(T^2(F^*))| \leq \varepsilon/2$ and $|V(T(F)) - V(T(F^*))| \leq 3\varepsilon/2$. Then, the advantage of the distinguisher is

$$|p - p^*| = |\mathbb{E}(T(F)) - \mathbb{E}(T(F^*))| \leq \mathbb{E}(|T(F) - T(F^*)|).$$

By applying Tchebichev's inequality for both $T(F)$ and $T(F^*)$ which is

$$\Pr[|T(F) - \mathbb{E}(T(F))| > \lambda] \leq V(T(F))/\lambda^2$$

for any $\lambda > 0$ (same for $T(F^*)$), we get

$$\begin{aligned} |p - p^*| &\leq \frac{V(T(F))}{\lambda^2} + \frac{V(T(F^*))}{\lambda^2} + 2n\big(|\mathbb{E}(T(F)) - \mathbb{E}(T(F^*))| + 2\lambda\big) \\ &\leq \frac{2V(T(F^*)) + \frac{3}{2}\varepsilon}{\lambda^2} + 2n\Big(\frac{\varepsilon}{2} + 2\lambda\Big). \end{aligned}$$

We have

$$|p - p^*| \leq 5\Big((2V(T(F^*)) + \frac{3\varepsilon}{2})n^2\Big)^{\frac{1}{3}} + n\varepsilon, \tag{1}$$

when $\lambda = \Big(\frac{2V(T(F^*)) + \frac{3\varepsilon}{2}}{n}\Big)^{\frac{1}{3}}$. Up to this point, the proof was the same with the proof of Theorem 1. However, the bound for $V(T(F^*))$ is different from the bound for $V(T(C^*))$, where C^* is the perfect cipher. We get $V(T(F^*))$ to be equal to

$$\sum_{(x,y),(x',y') \in \mathcal{T}} \Pr[X = x] \Pr[X = x'] \Big(\Pr_{F^*}[(x, x') \xrightarrow{F^*} (y, y')] - \Pr_{F^*}[x \xrightarrow{F^*} y] \Pr_{F^*}[x' \xrightarrow{F^*} y'] \Big).$$

In order to bound this sum, we divide pairs of iterations (x, x') into two groups such that the first group has no common queries, i.e. $\forall i, j \; x_i \neq x'_j$, but the second one has. As a remark, we assume that the adversary does not pick the same query in a single iteration, i.e. $x_i \neq x_j$, when $i \neq j$. As all x_i's are distinct in $x = (x_1, \ldots, x_d)$, then $[F^*]^d_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)} = \prod_{i=1}^d \Pr[F^*(x_i) = y_i] = N^{-d}$. When inputs x and x' have no common queries, then

$$[F^*]^{2d}_{(x_1,\ldots,x_d,x'_1,\ldots,x'_d),(y_1,\ldots,y_d,y'_1,\ldots,y'_d)} = \prod_{i=1}^d \Pr[F^*(x_i) = y_i] \prod_{i=1}^d \Pr[F^*(x'_i) = y'_i] = N^{-2d}.$$

Therefore, when input tuples x and x' have no common queries, the sum will be 0. Otherwise, when the plaintext tuples x and x' have common queries with probability δ, then the sum over all these plaintext tuples will be less than δ. Hence, we have $V(T(F^*)) \leq \delta$.

When we substitute δ for $V(T(F^*))$ in the inequality in Line (1), we get

$$|p - p^*| \leq 5\sqrt[3]{\Big(2\delta + \frac{3\varepsilon}{2}\Big)n^2} + n\varepsilon.$$

□

Two followings are useful for the rest of the paper.

Definition 4 The *trace* $\mathsf{Tr}(\beta)$ of an element $\beta \in \mathsf{GF}(2^k)$, is defined as

$$\mathsf{Tr}(\beta) = \beta + \beta^2 + \cdots + \beta^{2^{k-1}}.$$

Note that it is well known that the trace is linear and the trace of an element of $\mathsf{GF}(2^k)$ is either 0 or 1.

Lemma 1 *(Hoeffding's bound [11]): Let $X_1, X_2, \ldots, X_n$ be independent random variables and $0 \leq X_i \leq 1$, for $i \in \{1, \ldots, n\}$. Define $\bar{X} = \frac{1}{n}\sum_{i=0}^{n} X_i$ and let $\mu = \mathbb{E}(\bar{X})$. Then, for ε, $0 \leq \varepsilon \leq 1 - \mu$, we have*

$$\Pr[\bar{X} \geq \mathbb{E}(\bar{X}) + \varepsilon] \leq e^{-2n\varepsilon^2} \text{ and } \Pr[\bar{X} \leq \mathbb{E}(\bar{X}) - \varepsilon] \leq e^{-2n\varepsilon^2}.$$

In addition, two-sided Hoeffding's bound is stated by

$$\Pr[|\bar{X} - \mathbb{E}(\bar{X})| \geq \varepsilon] \leq 2e^{-2n\varepsilon^2}.$$

Definition 5 Let S be the sample space and $E \subseteq S$ be an event. The *indicator function* of the event E, denoted by $\mathbf{1}_E$, is a random variable defined as

$$\mathbf{1}_E(s) = \begin{cases} 1, & \text{if } \ s \in E, \\ 0, & \text{if } \ s \notin E. \end{cases}$$

The indicator function can shortly be denoted as $\mathbf{1}_E$ instead of $\mathbf{1}_E(s)$.

In the sequel, we concentrate on the solutions of two aforementioned open problems.

3 Addressing the Two Open Problems

We deal with two open problems in Decorrelation Theory. In [25], Vaudenay proposes Theorem 1 proving that the decorrelation of order $2d$ is *sufficient* for a cipher in order to resist a non-adaptive iterated attack of order d. We show here that the decorrelation of order $2d - 1$ is not sufficient by providing a counterexample. Secondly, the same theorem can be interpreted to imply that probability of having common queries increases the bound of the attack. To see the effect of this probability, we provide another counterexample showing that when this probability is high, the advantage of the distinguisher can be high as well.

3.1 A 3-round Feistel Scheme

We create a three round Feistel scheme C to be used in the following two subsections. This cipher C consists of three perfect κ-decorrelated functions F_1, F_2, and F_3 on $\mathcal{M}_1 = \mathsf{GF}(q)$. Each F_i is defined by

$$F_i(x) = a^i_{\kappa-1}x^{\kappa-1} + a^i_{\kappa-2}x^{\kappa-2} + \cdots + a^i_0$$

over a finite field $\mathsf{GF}(q)$, where $(a^i_{\kappa-1}, a^i_{\kappa-2}, \ldots, a^i_0)$ is distributed uniformly at random over $\mathsf{GF}(q)^\kappa$, for $i \in \{1, 2, 3\}$. According to Theorem 3, we have $\|[C]^\kappa - [C^*]^\kappa\|_A \leq 2\kappa^2/q$. Notice that this theorem bounds the advantage which is half of the decorrelation distance.

3.2 Decorrelation of Order $2d - 1$ is NOT Sufficient

In this section, we put forward a counterexample on the 3-round Feistel cipher C defined in the previous section, which is decorrelated to the order $\kappa = 2d - 1$. This implies that $\|[C]^{2d-1} - [C^*]^{2d-1}\|_A \leq 2(2d-1)^2/q$ due to Theorem 3. We provide a successful non-adaptive iterated distinguisher of order d against C showing that the decorrelation of order $2d - 1$ is *not* enough to resist a non-adaptive iterated distinguisher of order d.

We first start with explaining the input distribution that the adversary uses. Let $(x_1, x_2, \ldots, x_d)$ be the plaintext tuple and $(y_1, y_2, \ldots, y_d)$ be the ciphertext tuple such that $C(x_i) = y_i$, where $1 \leq i \leq d$. We will pick plaintexts with specific properties. Every plaintext x_i can be written as $x_i = x_i^L \| x_i^R$, where

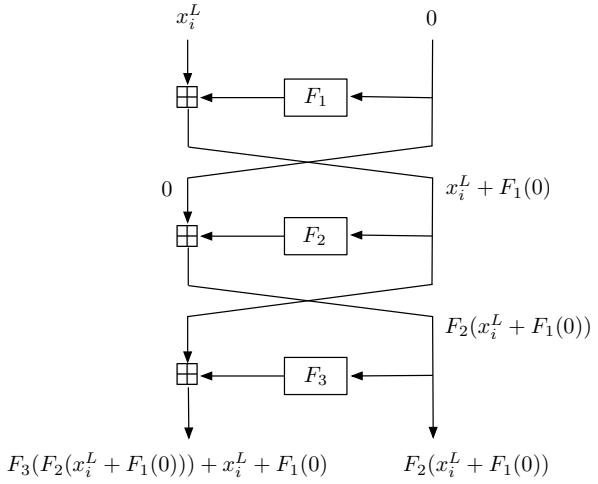

Fig. 1 Our 3-round Feistel scheme indicating the structure of the distinguishers defined in subsections 3.2 and 3.3

x_i^L and x_i^R, both in $\mathsf{GF}(q)$, are left and right halves of x_i. For each i, we let $x_i^R = 0$, i.e., $x_i = x_i^L \| 0$. Moreover, we choose a random c_1 and plaintexts $(x_1^L, x_2^L, \ldots, x_d^L)$ satisfying $\prod_{i=1}^{d} x_i^L = c_0$ and

$$\sum_{i_1 \le d} x_{i_1}^L = c_{d-1}, \sum_{i_1 < i_2 \le d} x_{i_1}^L x_{i_2}^L = c_{d-2}, \ldots, \sum_{i_1 < \cdots < i_{d-1} \le d} x_{i_1}^L x_{i_2}^L \cdots x_{i_{d-1}}^L = c_1,$$

where all c_i's, except c_1, are previously chosen constants and x_i^L's are pairwise distinct. The left half of these inputs is chosen by Algorithm 5.

Elaborating Algorithm 5. Given d and $c_0, c_2, \ldots, c_{d-1} \in \mathsf{GF}(q)$, this algorithm first constructs $h(x)$, where $h(x) = x^d - c_{d-1}x^{d-1} + \cdots + (-1)^{d-2}c_2x^2 + (-1)^d c_0$. It picks a random c_1 from $\mathsf{GF}(q)$ to construct $g(x)$ which is defined as $g(x) = h(x) + (-1)^{d-1}c_1 x$. It checks if $g(x)$ divides $x^q - x$ in order to be sure that all roots are in $\mathsf{GF}(q)$. Afterwards, it confirms that all roots are distinct by verifying that $g(x)$ and its derivative $g'(x)$ have no common divisors. Once these two conditions are satisfied, the algorithm outputs the roots of the polynomial g and gets the desired plaintext tuple. The number of iterations in the algorithm to get the desired plaintext tuple is approximately $q^d / \binom{q}{d} \le d!$, that is, one over the probability that a random monic polynomial of degree d has d distinct roots in $\mathsf{GF}(q)$. To be more precise, since there are q possible irreducible factors of degree 1 in $\mathsf{GF}(q)[x]$, we compute their d possible combinations in $\binom{q}{d}$ ways to construct polynomials of degree d and we divide it by the number of total monic polynomials of degree d which is q^d.

Algorithm 5 Generating the left half of the plaintext tuples

1: **Input:** d,$c_0, c_2, \ldots, c_{d-1}$, q
2: **Output:** $(x_1, \ldots, x_d)$
3: construct $h(x) = x^d - c_{d-1}x^{d-1} + \cdots + (-1)^{d-2}c_2x^2 + (-1)^d c_0$
4: **repeat**
5: pick $c_1 \in \mathsf{GF}(q)$ at random and construct $g(x) = h(x) + (-1)^{d-1}c_1 x$
6: **until** $x^q \equiv x \mod g(x)$ and $\gcd(g(x), g'(x)) = 1$
7: find the roots $(x_1, \ldots, x_d)$ of $g(x)$ by using a factorization algorithm for polynomials
8: **return** $(x_1, \ldots, x_d)$

Consider the encryption of each round when x_i's are satisfying the above properties. After the first round, we have $0 \| (x_i^L + F_1(0))$. Then, the output of the second round encryption is $(x_i^L + F_1(0)) \| F_2(x_i^L + F_1(0))$. Finally, the corresponding ciphertext y_i will be $y_i = y_i^L \| y_i^R = (F_3(F_2(x_i^L + F_1(0))) + x_i^L + F_1(0)) \| F_2(x_i^L + F_1(0))$. However, we will only be interested in the right part of the ciphertext, i.e. y_i^R, which can be seen as the output of a random polynomial function of degree at most $2d-2$. More explicitly,

as $y_i^R = F_2(x_i^L + F_1(0))$, we can write y_i^R as a function of x_i^L such that $F(x_i^L) = F_2(x_i^L + a_0^1)$. Obviously, each coefficient of the polynomial F is a function of coefficients of F_2 and the constant coefficient of F_1, namely $f_i(a_0^1, a_{2d-2}^2, \ldots, a_0^2)$. Because the coefficients of F depend on the coefficients of random functions F_1 and F_2, F is also a random function.

As we use the input distribution defined above, we can get some fixed bits by interpolating the function F, the right part of the output of the cipher. In more detail, in every iteration we interpolate a polynomial r, which will appear in the equation in Line (2). We expect that for the perfect function F^* the constant coefficient of the polynomial r would be random, but for F it would be fixed. We prove this in Lemma 2. After formally writing this argument, by defining the test function $\mathcal{T}$ and the acceptance set $\mathcal{A}cc$, we can distinguish the cipher from its ideal counterpart with only two iterations.

The distinguisher has d plaintext-ciphertext pairs in each iteration and F is indeed a polynomial. Moreover, we know d points on F and we then use the *underdetermined interpolation technique* to determine F. We write F such that $F(x) = a_{2d-2}x^{2d-2} + a_{2d-3}x^{2d-3} + \cdots + a_0$ over $\mathsf{GF}(q)$, then we can write F as

$$F(x) = r(x) + s(x)g(x). \tag{2}$$

Here, r is a unique polynomial of degree at most $d-1$ which interpolates d given points, s is a polynomial of degree at most $d-2$ over $\mathsf{GF}(q)$ and g is a polynomial of degree d with the x_i^L's as its roots $g(x) = (x - x_1^L)\cdots(x - x_d^L)$. Let $r(x) = r_{d-1}x^{d-1} + \cdots + r_0$, $g(x) = x^d - c_{d-1}x^{d-1} + \cdots + (-1)^{d-1}c_1 + (-1)^d c_0$, and $s(x) = s_{d-2}x^{d-2} + \cdots + s_0$, where $r_i, c_j, s_k \in \mathsf{GF}(q)$, $0 \le i, j \le d-1$, and $0 \le k \le d-2$. We note that $g(x)$ can be written as $g(x) = h(x) + (-1)^{d-1}c_1 x$, where h is a fixed polynomial of degree d with zero coefficient for the term x.

The aim is to get some fixed bits related to the function F in each iteration to have a distinguisher. The following lemma shows that, when the input is picked according to Algorithm 5, the constant coefficient r_0 of polynomial r is fixed in each iteration.

Lemma 2 *Let F be the polynomial of degree at most $2d-2$ over $\mathsf{GF}(q)$ as defined above. Let $x_1^L, \ldots, x_d^L \in \mathsf{GF}(q)$ be the left half of the plaintexts generated by Algorithm 5 and $F(x_i^L) = y_i^R$, $0 \le i \le d$. Then, the constant coefficient r_0 of the polynomial r, which is obtained by the Lagrange interpolation of the given d points, is fixed in each iteration.*

Proof We write $g(x) = h(x) + (-1)^{d-1}c_1 x$ for $(x_1^L, \ldots, x_d^L)$ and for some $c_1 \in \mathsf{GF}(q)$, where h is a fixed polynomial. Therefore, $F(x) = r(x) + s(x)g(x) = r(x) + s(x)(h(x) + (-1)^{d-1}c_1 x)$ as in Equation (2). For another input tuple $(x_1'^L, \ldots, x_d'^L)$ and some $c_1' \in \mathsf{GF}(q)$, we have $F(x) = r'(x) + s'(x)g'(x) = r'(x) + s'(x)(h(x) + (-1)^{d-1}c_1' x)$. Hence, we obtain

$$\begin{aligned}
&(r(x) + s(x)(h(x) + (-1)^{d-1}c_1 x)) - (r'(x) + s'(x)(h(x) + (-1)^{d-1}c_1' x)) \\
&= \underbrace{r(x) - r'(x) + ((-1)^{d-1}(c_1 s(x) - c_1' s'(x)))x}_{\text{Polynomial 1}} + \underbrace{(s(x) - s'(x))h(x) = 0}_{\text{Polynomial 2}} \\
&\Rightarrow s(x) = s'(x).
\end{aligned}$$

Polynomial 1 has degree at most $d-1$ and Polynomial 2 has at least degree d (the degree of h), unless $s(x) = s'(x)$. Therefore, in order to have zero on the right side of the equation, $s(x) - s'(x)$ has to be zero. This shows that the polynomial s is a fixed polynomial, i.e. independent from the plaintext tuple that is queried. Therefore, we can write F as $F(x) = r(x) + s(x)(h(x) + (-1)^{d-1}c_1 x)$, for fixed polynomials s and h and for some $c_1 \in \mathsf{GF}(q)$. Hence, when $x = 0$, we have $F(0) = r(0) + s(0)h(0)$ which implies that $r_0 = a_0 - s_0 h_0$ is always fixed. □

We can now deduce that r_0 is always fixed and independent from the plaintext tuples due to the carefully chosen plaintext tuples.

Now, we use Lemma 2 to construct a distinguisher between C and C^*. We denote the derived value of r_0 as a function $f((x_1, \ldots, x_d), (y_1, \ldots, y_d))$. Let D be a subset of distinguished values of $\mathsf{GF}(q)$ with a given cardinality q/μ, where $\mu > 1$ is a positive divisor of q. Define the test function as

$$\mathcal{T}((x_1, \ldots, x_d), (y_1, \ldots, y_d)) = \begin{cases} 1, & \text{if } f((x_1, \ldots, x_d), (y_1, \ldots, y_d)) \in D, \\ 0, & \text{otherwise,} \end{cases}$$

and the acceptance set as

$$\mathcal{A}cc[t_1,\ldots,t_n] = \begin{cases} 1, & \text{if } (t_1,\ldots,t_n) \neq (0,\ldots,0), \\ 0, & \text{otherwise.} \end{cases}$$

All iterations will reply the same answer for the function F and a random answer for F^*. Let p (resp. p^*) be the probability that the distinguisher outputs 1 when it is fed with F (resp. F^*). Hence, according to the acceptance set defined above, we get $p = \frac{1}{\mu}$ and $p^* = 1 - (1 - \frac{1}{\mu})^n$. If we consider $n = 2$, two iterations only, then the advantage of the distinguisher will be $|p - p^*| = |\frac{1}{\mu} - (1 - (1 - \frac{1}{\mu})^2)| = \frac{1}{\mu}(1 - \frac{1}{\mu})$ which is high. By this way, we can distinguish the cipher C from the perfect cipher C^* by distinguishing the function F, which defines the right part of the output of the cipher C, from the perfect function F^*.

Illustration of the attack for $d = 2$. We consider the Feistel scheme for $\kappa = 3$ over $\mathsf{GF}(p^k)$ which makes the write half of the output y_R a random polynomial $F(x) = F_2(x + F_1(0))$ of degree at most 2. The input distribution is chosen as $(x_1, x_2) = (x_1^L \| 0, x_2^L \| 0)$, $x_1^L + x_2^L = 0$ and $x_1^L \neq x_2^L$. If we write F as $F(x) = a_2x^2 + a_1x + a_0$, then we can recover a_1 as $a_1 = (y_1^R - y_2^R)(x_1^L - x_2^L)^{-1}$ by solving $a_2(x_1^L)^2 + a_1x_1^L + a_0 = y_1^R$ and $a_2(x_2^L)^2 + a_1x_2^L + a_0 = y_2^R$. The adversary will notice that a_1 is fixed for F in every iteration while it is random for a truly random cipher F^*. For this attack two iterations are enough to carry out the attack.

3.3 Assuming a Low δ is NECESSARY

Theorem 1 shows that if a cipher is decorrelated to the order $2d$, then it resists to an iterated attack of order d. Moreover, it is speculated that a high probability δ of having a common query does not provide any advantage to the adversary. However, we give a counterintuitive example showing that there is an iterated distinguisher of order 1 on a $2d$-decorrelated cipher when the probability of having at least one query in common in any two iterations is high. This shows that introducing this probability δ in Theorem 1 is necessary.

In our distinguisher, we use C, depicted in Figure 1, for $\kappa = 2d$ and $q = 2^k$ which has the property that $\|[C]^{2d} - [C^*]^{2d}\|_A \leq 8d^2/2^k$ due to Theorem 3. We are going to prove that C is not resisting an iterated attack of order 1 when the set of plaintexts of adversary's choice is small. Let S be a set of plaintexts $S = \{x_1, x_2, \ldots, x_{2d+2}\}$, where $x_i = x_i^L \| x_i^R$, x_L and x_R are the left and the right halves of x_i, respectively. These plaintexts satisfy $x_i^R = 0$ and $\sum_{i=1}^{2d+2} (x_i^L)^j = 0$, $1 \leq j \leq 2d-1$ and all x_i^L's are pairwise distinct elements of $\mathsf{GF}(2^k)$. How to generate the left part of the plaintexts in S is provided in Algorithm 6. In each iteration, we pick one element of S at random. As the adversary's choice of input set has $2d + 2$ elements, we have $\delta = 1/(2d + 2)$.

Analyzing Algorithm 6. This algorithm finds $p(x) = x^{2d+2} + ax^2 + bx + c$ with distinct roots in $\mathsf{GF}(2^k)$, where $a, b, c \in \mathsf{GF}(2^k)$. Note that $p(x)$ has roots satisfying $\sum_{i=1}^{2d+2} (x_i^L)^j = 0$, $1 \leq j \leq 2d-1$. This is proved in Lemma 3 when $n = 2d + 2$. The expected number of iterations in the algorithm can be computed heuristically as $2^{k(2d+2)}/\binom{2^k}{2d+2} \leq (2d + 2)!$, that is, one over the probability that a random monic polynomial of degree $2d + 2$ has $2d + 2$ distinct roots in $\mathsf{GF}(2^k)$. As there are 2^k possible irreducible factors of degree 1 in $\mathsf{GF}(2^k)[x]$, we compute their $2d + 2$ possible combinations $\binom{2^k}{2d+2}$ to construct polynomials of degree $2d+2$ and we divide it by the total number of monic polynomials of degree $2d+2$ which is $2^{k(2d+2)}$.

Lemma 3 *Let f be a polynomial of the form $f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0$ over $\mathsf{GF}(2^k)$ and $x_1, x_2, \ldots, x_n$ be its roots. If $a_{n-1} = a_{n-2} = \cdots = a_3 = 0$, then its roots satisfy $s_k = \sum_{i=1}^n x_i^j = 0$, where $1 \leq j \leq n - 3$ and $n \geq 4$.*

Proof First, we recall the *Newton formulas*. Let f be a polynomial of the form $f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x + a_0$ with the roots $x_1, x_2, \ldots, x_n$ so that $f(x) = (x - x_1)(x - x_2)\cdots(x - x_n)$

over a ring. Then, we define the elementary symmetric functions of the roots as

$$\begin{aligned}
&\sum_{1\le i\le n} x_i = -a_{n-1},\\
&\sum_{1\le i<j\le n} x_i x_j = a_{n-2},\\
&\sum_{1\le i<j<k\le n} x_i x_j x_k = -a_{n-3},\\
&\cdots,\\
&\prod_{i=1}^{n} x_i = (-1)^n a_0.
\end{aligned}$$

In addition, the kth power sums of the roots are defined as $s_k = \sum_{1\le i\le n} x_i^k$. Then, *Newton formulas* give recursion relations between a_i's and s_i's as

$$\begin{aligned}
&s_1 + a_{n-1} = 0,\\
&s_2 + a_{n-1}s_1 + 2a_{n-2} = 0,\\
&\cdots,\\
&s_n + a_{n-1}s_{n-1} + a_{n-2}s_{n-2} + \cdots + na_0 = 0,\\
&s_{n+1} + a_{n-1}s_{n-1} + a_{n-2}s_{n-2} + \cdots + s_1 a_0 = 0.
\end{aligned}$$

Note that Newton formulas are valid over finite fields. If we assume that $a_{n-1} = a_{n-2} = \cdots = a_3 = 0$, then according to the Newton formulas given above we will show that $s_k = 0$ for $1 \le k \le n-3$. As $s_1 + a_1 = 0$ and $a_1 = 0$, we have $s_1 = 0$. By induction, assume that $s_1 = s_2 = \cdots = s_{k-1} = 0$ and $a_{n-1} = a_{n-2} = \cdots = a_{n-k} = 0$, then we have $s_k = -(a_{n-1}s_{k-1} + a_{n-2}s_{k-2} + \cdots + ka_{n-k}) = 0$. Therefore, the polynomial $f(x) = x^n + a_2x^2 + a_1x + a_0$ satisfies $s_k = \sum_{i=1}^{n} x_i^j = 0$, where $1 \le j \le n-3$ and $n \ge 4$. □

Algorithm 6 Generating the left half of the plaintexts in S

1: **Input:** k, d
2: **Output:** $(x_1, \ldots, x_{2d+2})$
3: **repeat**
4: pick $a, b, c \in \mathsf{GF}(2^k)$ at random and construct $p(x) = x^{2d+2} + ax^2 + bx + c$
5: **until** $x^{2^k} \equiv x \mod p(x)$ and $\mathsf{gcd}(p(x), p'(x)) = 1$
6: find the roots $(x_1, \ldots, x_{2d+2})$ of $p(x)$ by using a factorization algorithm for polynomials
7: **return** $(x_1, \ldots, x_{2d+2})$

Like in Section 3.2, in order to distinguish this cipher C from the perfect cipher C^*, we distinguish the right half of the output of the cipher which is itself a random function of the form $F(x) = a_{2d-1}x^{2d-1} + a_{2d-2}x^{2d-2} + \cdots + a_0$ over $\mathsf{GF}(q)$.

To explain how the distinguisher works briefly, when we consider that the plaintext space has the special form mentioned before, the polynomial F can be distinguished by using the trace (see Definition 4). This is because, by the following Lemma, the sum of the trace of all elements of S behaves differently when F is considered than when F^* is considered. Now, the following lemma proves this distinguishing property of F.

Lemma 4 *Let F be a random function and S be the input set defined as above. For $x_i = x_i^L \| 0 \in S$, $1 \le i \le 2d+2$, we have*

$$\sum_{i=1}^{2d+2} \mathsf{Tr}(F(x_i^L)) = 0.$$

Proof As $\sum_{i=1}^{2d+2}(x_i^L)^j = 0$, $1 \le j \le 2d-1$, we have

$$\sum_{i=1}^{2d+2} \mathsf{Tr}(F(x_i^L)) = \sum_{i=1}^{2d+2} \mathsf{Tr}(a_{2d-1}(x_i^L)^{2d-1} + a_{2d-2}(x_i^L)^{2d-2} + \cdots + a_0) =$$
$$\mathsf{Tr}\Big(a_{2d-1}\sum_{i=1}^{2d+2}(x_i^L)^{2d-1}\Big) + \mathsf{Tr}\Big(a_{2d-2}\sum_{i=1}^{2d+2}(x_i^L)^{2d-2}\Big) + \cdots + \mathsf{Tr}\Big(a_0\sum_{i=1}^{2d+2}(x_i^L)^{0}\Big).$$

Which is equal to 0 due the linearity of trace, the characteristic of this field being two, and $\mathsf{Tr}(0) = 0$. □

Lemma 4 implies that there is an *even* number of $F(x_i^L)$'s which satisfy the property that $\mathsf{Tr}(F(x_i^L)) = 1$ as $\sum_{i=1}^{2d+2}\mathsf{Tr}(F(x_i^L)) = 0$ and the characteristic of this field is zero.

Now, we explain how the iterated distinguisher of order 1 with n iterations works, where the input is distributed independently and identically over the set S. We use the property of the polynomial function F which is stated in Lemma 4 in a way that in each iteration, we pick a plaintext x from S at random and compute the trace of $F(x^L)$, i.e. $t = \mathsf{Tr}(F(x^L))$ being $F(x^L) = y^R$. Then, we compute the average $\bar{T} = \frac{1}{n}(t_1 + \cdots + t_n)$, where t_i is the output of iteration i. We decide whether the oracle implements F (equivalently C) or F^* (equivalently C^*) by simply checking that the average value $\bar{T}$ is in the specified set K which is determined according to the expected values of both $\bar{T}$ and $\bar{T}^*$ given in the following lemma.

Lemma 5 *Let S have the aforementioned property. Then, depending on S, the expected value of $\bar{T}$ takes any value from the set $S_1 = \{2m/(2d+2)|\ 0 \le m \le d+1\}$, and that of $\bar{T}^*$ takes any value from the set $S_2 = \{m/(2d+2)|0 \le m \le 2d+2\}$.*

Proof Assume that there are $2m$ number of x_i's in S such that $F(x_i^L)$'s have $\mathsf{Tr}(F(x_i^L)) = 1$ (from Lemma 4). Then, the number of x_i's satisfying $\mathsf{Tr}(F(x_i^L)) = 1$ in n iterations is expected to be $n(2m)/(2d+2)$, where $0 \le m \le d+1$. Therefore, the expected value of $\bar{T}$ will be $2m/(2d+2)$, where $0 \le m \le d+1$. In a similar way, we can find the expected value of $\mathbb{E}(\bar{T}^*)$ for the perfect function F^*. □

Now, using Lemma 5, we define the acceptance set as

$$\mathcal{A}cc[t_1, \ldots, t_n] = \begin{cases} 1, & \text{if } \bar{T} \in K = \bigcup_{m=0}^{d+1}\left(\frac{2m}{2d+2} - \varepsilon, \frac{2m}{2d+2} + \varepsilon\right), \\ 0, & \text{otherwise.} \end{cases}$$

Typically, $\varepsilon = 1/(4d+4)$. Let p (resp. p^*) be the probability that the distinguisher outputs 1 when it is fed with F (resp. F^*). The following lemma states the bounds for both p and p^*.

Lemma 6 *The probabilities of the distinguisher outputs 1 when C is considered and C^* is considered respectively are*

$$p \ge 1 - 2e^{-2n\varepsilon^2}$$

and

$$p^* \le \frac{1}{2} + e^{-2n\varepsilon^2}.$$

Proof For the function F, according to the acceptance set defined previously, p is expressed by

$$p = \sum_{x \in S_1} \Pr[\mathbb{E}(\bar{T}) = x] \Pr[\bar{T} \in K | \mathbb{E}(\bar{T}) = x].$$

As $\Pr[\bar{T} \in K|\mathbb{E}(\bar{T}) = x] \ge 1 - 2e^{-2n\varepsilon^2}$ by Hoeffding's bound from Lemma 1, we have $p \ge 1 - 2e^{-2n\varepsilon^2}$. Similarly, p^* is computed as

$$p^* = \sum_{x \in S_2} \Pr[\mathbb{E}(\bar{T}^*) = x] \Pr[\bar{T}^* \in K | \mathbb{E}(\bar{T}^*) = x].$$

The computation of p^* is not straightforward; hence, we first compute the probability that each expected value of $\bar{T}^*$ from S_2 occurs with probability $\Pr[\mathbb{E}(\bar{T}^*) = x] = \binom{2^{d+2}}{x(2d+2)} 2^{-(2d+2)}$. In detail, we are picking

$x(2d+2)$ places for 1's among $2d+2$ possible places and dividing the total number of possible choices which is 2^{2d+2}. Furthermore, the probabilities $\Pr[\bar{T}^* \in K|\mathbb{E}(\bar{T}^*) = x]$ are different according to the expected value of different $\bar{T}^*$. When $\mathbb{E}(\bar{T}^*) = 2m/(2d+2)$ for $0 \leq m \leq d+1$, we have $\Pr[\bar{T}^* \in K|\mathbb{E}(\bar{T}^*) = x] \leq 1$. Similarly, when $\mathbb{E}(\bar{T}^*) = (2m'+1)/(2d+2)$, $0 \leq m' \leq d$, we get $\Pr[\bar{T}^* \in K|\mathbb{E}(\bar{T}^*) = x] \leq 2e^{-2n\varepsilon^2}$ by Hoeffding's bound. Then, p^* can be computed as

$$\begin{aligned}
p^* &= \sum_{x \in \{2m/(2d+2)|0 \leq m \leq d+1\}} \Pr[\mathbb{E}(\bar{T}^*) = x] \Pr[\bar{T}^* \in K|\mathbb{E}(\bar{T}^*) = x] \\
&+ \sum_{x \in \{(2m'+1)/(2d+2)|0 \leq m' \leq d\}} \Pr[\mathbb{E}(\bar{T}^*) = x] \Pr[\bar{T}^* \in K|\mathbb{E}(\bar{T}^*) = x] \\
&\leq \sum_{x \in \{2m/(2d+2)|0 \leq m \leq d+1\}} \binom{2d+2}{x(2d+2)} 2^{-(2d+2)} \\
&+ \sum_{x \in \{(2m'+1)/(2d+2)|0 \leq m' \leq d\}} \binom{2d+2}{x(2d+2)} 2^{-(2d+2)} 2e^{-2n\varepsilon^2}.
\end{aligned}$$

Note that the sum of even and odd indices of binomial coefficients are $\sum_{i \geq 0} \binom{n}{2i} = 2^{n-1}$ and $\sum_{i \geq 0} \binom{n}{2i+1} = 2^{n-1}$, respectively. Hence, we have

$$\sum_{x \in \{2m/(2d+2)|0 \leq m \leq d+1\}} \binom{2d+2}{x(2d+2)} 2^{-(2d+2)} = \frac{1}{2}.$$

Then, we get

$$\sum_{x \in \{(2m'+1)/(2d+2)|0 \leq m' \leq d\}} \binom{2d+2}{x(2d+2)} 2^{-(2d+2)} 2e^{-2n\varepsilon^2} \leq \frac{1}{2} 2e^{-2n\varepsilon^2} = e^{-2n\varepsilon^2}.$$

Therefore, we get $p^* \leq \frac{1}{2} + e^{-2n\varepsilon^2}$. $\square$

Finally, the advantage of the distinguisher is

$$|p - p^*| \geq \left|\left(1 - 2e^{-2n\varepsilon^2}\right) - \left(\frac{1}{2} + e^{-2n\varepsilon^2}\right)\right| = \left|\frac{1}{2} - 3e^{-2n\varepsilon^2}\right|. \tag{3}$$

When the distinguisher has a large number of iterations, we have $|p - p^*| \approx 1/2$ which is quite high. This way we manage to distinguish the cipher C from the ideal random cipher C^*. When the distinguisher has a large number of iterations, we have $|p - p^*| \approx 1/2$ which is quite high. This way, we manage to distinguish the cipher C from the perfect cipher C^*. Hence, *in specific situations*, having common queries *can* increase the advantage. Essentially, if the images of $2d+2$ points sum to zero, by taking $\varepsilon = 1/(4d+4)$ and $n \approx \Omega(d^2)$, we obtain an efficient iterated distinguisher of order 1.

As a final remark, in Decorrelation Theory, Vaudenay considers block ciphers in the context of deterministic symmetric-key encryption. Therefore, for some input distribution, the probability δ that two iterations have at least one query in common can be high. However, if we consider probabilistic symmetric-key encryption, then δ will always be small. This is because, in this scheme, the oracle picks the random coins, and even if the same plaintext is picked by the adversary, the random coins picked by the oracle for two plaintexts would be different which cause two different inputs to the encryption.

Illustration of the attack for $d = 1$. We consider the Feistel scheme for $\kappa = 2$ over $\mathsf{GF}(2^k)$. Hence, the write half of the output y_R becomes a random polynomial $F(x) = F_2(x + F_1(0))$ of degree at most 1. The attacker picks the set of plaintexts $S = \{x_1, x_2, x_3, x_4\}$ in a way that x_i's are pairwise distinct, $x_1^L + x_2^L + x_3^L + x_4^L = 0$ and $x_i = x_i^L \| 0$, $1 \leq i \leq 4$. In each iteration, a chosen plaintext x is taken from S, hence, $\delta = 1/4$. The adversary computes $t_i = \mathrm{Trace}(y_i^R)$ in each iteration and then calculate the experimental average $\bar{T}$. Since there are even number of $F(x_i^L)$'s for F such that $\mathrm{Trace}(F(x_i^L)) = 1$, the expected value of $\bar{T}$ of F is concentrated around the values in $\{0, 2/4, 4/4\}$ while the expected value of $\bar{T}^*$ of F^* is concentrated around the values in $\{0, 1/4, 2/4, 3/4, 4/4\}$. According to the advantage of the attack given in Line 3, the attack is feasible with 1000 iterations.

4 Resistance Against Adaptive Iterated Distinguishers

As mentioned before, we explore the maximum success of the adaptive iterated distinguishers.

4.1 Adaptive Plaintext-Ciphertext Iterated Distinguishers of Order d

We recall two generic distinguishers, namely an *adaptive* plaintext-ciphertext d-limited distinguisher (see Algorithm 7) and an *adaptive* plaintext-ciphertext iterated distinguisher of order d (see Algorithm 8). Both distinguishers are adaptive in a way that the adversary adaptively asks for both encryption and decryption of the queries. Herein we formalize these distinguishers.

We first define a *compact* function G to be distinguished. The goal of defining this function is that the input to the oracle is able to be specified being either encrypted or decrypted (as the adversary makes either the plaintext queries or the ciphertext queries in a specific order depending on his type of attack).

Let $\mathcal{G}$ be the set of functions G such that $G : \mathcal{M} \times \{0,1\} \to \mathcal{M}$ satisfying $G(G(x,0),1) = x$ and $G(G(x,1),0) = x$, for all x. We denote $G_0(x) = G(x,0)$ and $G_1(x) = G(x,1)$ and point out $G_1^{-1} = G_0$ and $G_0^{-1} = G_1$. In what follows, G denotes a random element of $\mathcal{G}$ and G^* is a uniformly distributed element of $\mathcal{G}$.

Algorithm 7 A generic adaptive plaintext-ciphertext d-limited distinguisher

Input: a function $\mathcal{F}$, a test $\mathcal{T}$, a distribution $\mathcal{R}$ on $\{0,1\}^*$
Oracle: an oracle Ω implementing either an instance of G or an instance of G^*
1: pick $r \in \{0,1\}^*$ at random from $\mathcal{R}$
2: set $u_1 = (a_1, b_1) \leftarrow \mathcal{F}(\cdot; r)$
3: set $v_1 = \Omega(u_1)$
4: set $u_2 = (a_2, b_2) \leftarrow \mathcal{F}(v_1; r)$
5: set $v_2 = \Omega(u_2)$
6: ...
7: set $u_d = (a_d, b_d) \leftarrow \mathcal{F}(v_1, \ldots, v_{d-1}; r)$
8: set $v_d = \Omega(u_d)$
9: output $\mathcal{T}(v_1, \ldots, v_d; r)$

An adaptive d-limited distinguisher. The adversary $\mathcal{A}_{\mathsf{A}(d)}$, detailed in Algorithm 7, has access to an oracle Ω which implements either an instance of G or an instance of G^*, such that functions G_0 and G_1 perform encryption and decryption, respectively. He picks a random coin r from $\{0,1\}^*$ according to a given distribution $\mathcal{R}$ and queries a function $\mathcal{F}$ which is fed with r and the output of the previous queries $(v_1, v_2, \ldots, v_{i-1})$, where $v_k = \Omega(u_k)$ for all $k \in \{1, 2, \ldots, i-1\}$, and $1 \leq i \leq d$. He then receives a new query u_i. He sends this input u_i to the oracle to receive the output v_i, where –as explained– $v_i = \Omega(u_i)$. Finally, using a test $\mathcal{T}$, he outputs a decision bit "1" if he guesses that Ω implements an instance of the random function G, or "0" if he guesses that Ω implements an instance of the perfect function G^*.

Algorithm 8 A generic adaptive plaintext-ciphertext iterated distinguisher of order d

Input: an integer n, a function $\mathcal{F}$, a test $\mathcal{T}$, a set $\mathcal{A}cc$, a distribution $\mathcal{R}$ on $\{0,1\}^*$
Oracle: an oracle Ω implementing a function G or G^*
1: **for** $k = 1$ to n **do**
2: set T_k (with independent coins) $\leftarrow$ output of Distinguisher in Algorithm 7
3: **end for**
4: output $1_{\mathcal{A}cc}(T_1, \ldots, T_n)$

An adaptive iterated distinguisher of order d. The iterated distinguisher given in Algorithm 8 is simply the iteration of the d-limited distinguisher (see Algorithm 5) in a way that the adversary $\mathcal{A}_{\mathsf{AI}(d)}$ repeats the distinguisher n times, then he checks whether the output of n iterations are accepted or not with respect to a set $\mathcal{A}cc$. This gives his final decision.

Algorithm 9 Boomerang Distinguisher

Input: an integer n, a set X, differences Δ and ∇
Oracle: an oracle Ω implementing a permutation c

for $k = 1$ to n **do**
  pick x_1 uniformly at random over X
  set $x_2 = x_1 \oplus \Delta$
  set $y_1 = c(x_1)$, $y_2 = c(x_2)$
  set $y_3 = y_1 \oplus \nabla$, $y_4 = y_2 \oplus \nabla$
  set $x_3 = c^{-1}(y_3)$, $x_4 = c^{-1}(y_4)$
  set $T_k = \mathbf{1}_{x_3 \oplus x_4 = \Delta}$
end for
if $T_1 + \cdots + T_n \neq 0$ **then**
  output 1
else
  output 0
end if

The boomerang attack [26] defined in Algorithm 9 is an example for an adaptive plaintext-ciphertext iterated distinguisher of order d (see Algorithm 8) for the case $d = 4$. The adversary queries two (chosen) plaintexts and receives their corresponding ciphertexts, he then constructs two ciphertexts depending on the previous ciphertexts and asks for their decryption. The adaptively chosen queries to the oracle in each iteration of the boomerang attack [26] can be written as $(u_1, u_2, u_3, u_4) = ((x_1, 0), (x_1 \oplus \Delta, 0), (c(x_1) \oplus \nabla, 1), (c(x_1 \oplus \Delta) \oplus \nabla, 1))$, where x_1 is selected uniformly at random over the set X, and Δ and ∇ denote non-zero differences.

4.2 Advantage of Adaptive Plaintext-Ciphertext Iterated Distinguishers of Order d

Vaudenay [25] found a bound for the advantage of non-adaptive iterated distinguishers of order d, which is not applicable to the adaptive adversaries. We extend his result and provide a bound for the advantage of *adaptive* plaintext-ciphertext iterated distinguishers of order d. Strictly speaking, we compute the maximum success of the adversary who is making d adaptive queries to the oracle in each iteration to distinguish a $2d$-decorrelated random cipher upon using the $\|\cdot\|_A$ norm.

Theorem 5 *Let $G \in \mathcal{G}$ be a random function from $\mathcal{M} \times \{0,1\}$ to $\mathcal{M}$ such that $||[G]^{2d} - [G^*]^{2d}||_A \leq \varepsilon$, for some given integer $d \leq M/2$, where G^* is the perfect function and $|\mathcal{M}| = M$. Let us consider* an adaptive *iterated distinguisher of order d $\mathcal{A}_{\mathsf{AI}(d)}$ which is trying to distinguish G from G^* by performing n iterations (see Algorithm 6). Then, the advantage* $\mathrm{Adv}_{\mathcal{A}_{\mathsf{AI}(d)}}$ *of $\mathcal{A}_{\mathsf{AI}(d)}$ is*

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{AI}(d)}} \leq 5\sqrt[3]{\left(2\theta + e^{8d^2/M} + \frac{2d^2}{M} + \frac{3\varepsilon}{2} - 1\right)n^2} + n\varepsilon,$$

where θ is the probability that any two different iterations have at least one query in common.

Proof Let one iteration consist of the input queries $u = (u_1, u_2, \ldots, u_d)$ and the output queries $v = (v_1, v_2, \ldots, v_d)$, where $u_i = (a_i, b_i)$ and $v_i = \Omega(u_i)$, for $1 \leq i \leq d$.

We first make two conventional *assumptions* about the adaptive adversary.

Assumption 1: *Inner-collisions* in input queries, i.e. $u_i = u_j$, are not allowed, as calling the same query twice in the same iteration will not give any advantage to the adversary.

Assumption 2: Let $(u_i = (a_i, b_i), v_i)$ and $(u_j = (a_j, b_j), v_j)$ be two queries in the same iteration. *Cross inner-collisions* are not allowed, that is, we *never* have $a_i = v_j$ or $a_j = v_i$ when $b_i \neq b_j$, as getting the same information will not give any advantage to the adversary.

Notice that these assumptions do not hold between *different* iterations.

We begin similarly to the proof of Theorem 1 provided in [25]. We first define $T(g)$ to be the probability that the test function $\mathcal{T}$ outputs 1 when $G = g$ (resp. $G^* = g$), i.e.

$$T(g) = \mathbb{E}_r(\mathcal{T}(v_1, \ldots, v_d; r) | G = g).$$

We let p (resp. p^*) be the probability of the distinguisher outputting 1, let $\mathcal{A}cc$ be the acceptance set, and $T_k(G)$ (resp. $T_k(G^*)$) be the output of iteration k using coins r_k's. Then, we have

$$p = \mathsf{Pr}_{G,r_1,\ldots,r_n}[(T_1(G),\ldots,T_n(G)) \in \mathcal{A}cc].$$

Notice that all $T_k(g)$'s are *pairwise independent* once g is fixed and that $\mathbb{E}_{r_k}(T_k(g)) = T(g)$. Hence, we obtain

$$p = \mathbb{E}_G\left(\sum_{(t_1,\ldots,t_n)\in\mathcal{A}cc} T(G)^{t_1+\cdots+t_n}(1-T(G))^{n-(t_1+\cdots+t_n)}\right).$$

Then, p can be rewritten as

$$p = \sum_{k=0}^{n} a_k \mathbb{E}_G(T(G)^k(1-T(G))^{n-k}),$$

for some integers a_k such that $0 \le a_k \le \binom{n}{k}$. Similarly, we have the same argument for p^*, i.e.

$$p^* = \sum_{k=0}^{n} a_k \mathbb{E}_{G^*}(T(G^*)^k(1-T(G^*))^{n-k}).$$

The advantage of the distinguisher $|p-p^*|$ is maximal, when all a_k's are either 0 or $\binom{n}{k}$ depending on the distributions $T(G)$ and $T(G^*)$. Hence, we assume that $\mathcal{A}cc$ of the best distinguisher is of the form

$$\mathcal{A}cc = \left\{(t_1,\ldots,t_n)\Big|\sum_{k=1}^{n} t_k \in \mathcal{B}\right\},$$

for some set $\mathcal{B} \subseteq \{0,\ldots,n\}$. Thus, we rewrite $p = \mathbb{E}_G(s(T(G)))$, where $s(x) = \sum_{k\in\mathcal{B}}\binom{n}{k}x^k(1-x)^{n-k}$.

Now, consider the derivative of s which can be written as

$$s'(x) = \sum_{k\in\mathcal{B}}\binom{n}{k}\frac{k-nx}{x(1-x)}x^k(1-x)^{n-k}.$$

Notice that as the sum over all k, such that $0 \le k \le n$, is the derivative of $(x+(1-x))^n$, then the total sum is zero. Hence, we obtain

$$|s'(x)| \le \sum_{nx\le k\le n}\binom{n}{k}\frac{k-nx}{x(1-x)}x^k(1-x)^{n-k} \le \frac{n}{x}\sum_{nx\le k\le n}\binom{n}{k}x^k(1-x)^{n-k},$$

because $nx \le k \le n$. We note that when $x \ge 1/2$, we have $|s'(x)| \le 2n$. Similarly, when $x < 1/2$, we have $|s'(x)| \le 2n$. Hence, we get $|s'(x)| \le 2n$, for every x. So, according to the Mean Value Theorem, we have

$$|s(T(G)) - s(T(G^*))| \le 2n|T(G) - T(G^*)|. \tag{4}$$

Furthermore, Theorem 2 gives the exact advantage for the best *adaptive* d-limited distinguisher. Hence, $|\mathbb{E}_G(T(G)) - \mathbb{E}_{G^*}(T(G^*))| \le \varepsilon/2$ is obtained.

We now define a new random variable $T^2(G)$ which is the output of another test with $2d$ entries, that is,

$$\mathcal{T}(v_1,\ldots,v_d;r) \times \mathcal{T}(v'_1,\ldots,v'_d;r').$$

Thanks to Theorem 2, we have $|\mathbb{E}_G(T^2(G)) - \mathbb{E}_{G^*}(T^2(G^*))| \le \varepsilon/2$. Hence, we get $|V(T(G)) - V(T(G^*))| \le 3\varepsilon/2$ which is obtained by $|\mathbb{E}_G(T(G)) - \mathbb{E}_{G^*}(T(G^*))| \le \varepsilon/2$ and $|\mathbb{E}_G(T^2(G)) - \mathbb{E}_{G^*}(T^2(G^*))| \le \varepsilon/2$. More precisely, we have

$$\begin{aligned}
|V(T(G)) - V(T(G^*))| &= |\mathbb{E}_G(T^2(G)) - \mathbb{E}^2_G(T(G)) - \mathbb{E}_{G^*}(T^2(G^*)) + \mathbb{E}^2_{G^*}(T(G^*))| \\
&\le |\mathbb{E}_G(T^2(G)) - \mathbb{E}_{G^*}(T^2(G^*))| + |\mathbb{E}^2_G(T(G)) - \mathbb{E}^2_{G^*}(T(G^*))| \\
&\le \frac{3\varepsilon}{2}. \qquad (5)
\end{aligned}$$

To obtain the result in Line (5), we use $|\mathbb{E}_G(T(G)) + \mathbb{E}_{G^*}(T(G^*))| \le 2$, as $0 \le T(G), T(G^*) \le 1$.

Afterwards, the advantage of the distinguisher is

$$|p - p^*| = |\mathbb{E}_G(s(T(G))) - \mathbb{E}_{G^*}(s(T(G^*)))| \leq \mathbb{E}_{G,G^*}(|s(T(G)) - s(T(G^*))|).$$

By applying Tchebichev's inequality for both $T(G)$ and $T(G^*)$, i.e.

$$\Pr[|T(G) - \mathbb{E}_G(T(G))| > \lambda] \leq V(T(G))/\lambda^2 \text{ and } \Pr[|T(G^*) - \mathbb{E}_{G^*}(T(G^*))| > \lambda] \leq V(T(G^*))/\lambda^2$$

for any $\lambda > 0$, we get

$$\begin{aligned} |p - p^*| &\leq \frac{V(T(G))}{\lambda^2} + \frac{V(T(G^*))}{\lambda^2} + 2n\big(|\mathbb{E}(T(G)) - \mathbb{E}(T(G^*))| + 2\lambda\big) && (6)\\ &\leq \frac{2V(T(G^*)) + \frac{3}{2}\varepsilon}{\lambda^2} + 2n\big(\frac{\varepsilon}{2} + 2\lambda\big) && (7) \end{aligned}$$

Note that s maps the elements of $[0,1]$ to $[0,1]$, Line (6) is due to Inequality (4). We have

$$|p - p^*| \leq 5\sqrt[3]{\Big(2V(T(G^*)) + \frac{3\varepsilon}{2}\Big)n^2} + n\varepsilon, \tag{8}$$

when $\lambda = \sqrt[3]{(2V(T(G^*)) + (3\varepsilon/2))/n}$.

So far, everything works similarly to [25]. However, the rest is different as the function implemented in the oracle has new properties. For further details of the proof up to now, refer to [25]. Now, it is left to bound $V(T(G^*))$.

Bounding $V(T(G^*))$. We now bound $V(T(G^*))$ by expanding it as

$$V(T(G^*)) = \sum_S \Pr_R[r] \Pr_R[r'] \Big(\Pr_{G^*}\big[(u,u') \xrightarrow{G^*} (v,v')\big] - \Pr_{G^*}[u \xrightarrow{G^*} v] \Pr_{G^*}[u' \xrightarrow{G^*} v'] \Big), \tag{9}$$

where $S = \{(v,r),(v',r') \in \mathcal{T}\}$ and u (resp. u') is defined by both r and v (resp. r' and v'). For the sake of simplicity, we denote inside the sum in Line (9) as P.

We first divide the expression in Line (9) into two *disjoint* sums depending on whether or not u and u' (from two different iterations) are colliding, i.e. if there exist i and j such that $u_i = u'_j$ with u and u' defined from (v,r) and (v',r'), respectively. In detail, we have $S_1 = \big\{(v,r),(v',r') \in \mathcal{T} |\ \exists i,j \text{ s.t. } u_i = u'_j\big\}$ and $S_2 = \big\{(v,r),(v',r') \in \mathcal{T} |\ \forall i,j \text{ s.t. } u_i \neq u'_j\big\}$ such that $S = S_1 \cup S_2$. Thus, we write

$$\sum_S P = \sum_{S_1} P \ + \sum_{S_2} P.$$

We now bound each sum separately.

The sum $\sum_{S_1} P$ over S_1 is bounded as

$$\begin{aligned} \sum_{S_1} P &\leq \sum_{v,v'} \sum_{r,r'} \Pr_R[r] \Pr_R[r'] \Pr_{G^*}\big[(u,u') \xrightarrow{G^*} (v,v')\big]\ \mathbf{1}_{S_1} \\ &= \sum_g \Pr[G^* = g] \sum_{v,v'} \sum_{r,r'} \Pr_R[r] \Pr_R[r']\ \mathbf{1}_{(u,u') \xrightarrow{g} (v,v')}\ \mathbf{1}_{S_1} \\ &= \sum_g \Pr[G^* = g] \sum_{r,r'} \Pr_R[r] \Pr_R[r'] \mathbf{1}_{\exists i,j \text{s.t. } u_i = u'_j} \\ &= \mathbb{E}_{G^*}\big(\Pr_{r,r'}[\exists i,j \text{ s.t. } u_i = u'_j]\big) \stackrel{def}{=} \theta. \end{aligned} \tag{10}$$

Here, we denote $\mathbb{E}_{G^*}(\Pr_{r,r'}[\exists i,j \text{ s.t. } u_i = u'_j])$ by θ. This can be interpreted as the probability that any two different iterations have at least one query in common. Notice that the sum $\sum_{v,v'} \mathbf{1}_{(u,u') \xrightarrow{g} (v,v')} \mathbf{1}_{S_1}$ is equal to 1 if (v,r) and (v',r') are in S_1.

Now, we provide a bound for the sum $\sum_{S_2} P$ over S_2 which is for non-colliding inputs u and u'. We first note that, as both G_0^* and G_1^* are from $\mathcal{M}$ to $\mathcal{M}$, and, hence, bijective, they are indeed a uniformly distributed permutation C^*. From now on, we use the notation C^* for both G_0^* and G_1^*.

We define $x = (x_1, x_2, \ldots, x_d)$ and $y = (y_1, y_2, \ldots, y_d)$ as

$$x_i = \begin{cases} a_i, & \text{if } \ b_i = 0, \\ v_i, & \text{if } \ b_i = 1, \end{cases} \quad \text{and} \quad y_i = \begin{cases} v_i, & \text{if } \ b_i = 0, \\ a_i, & \text{if } \ b_i = 1, \end{cases}$$

where $u = ((a_1, b_1), (a_2, b_2), \ldots, (a_d, b_d))$, with $b_i \in \{0,1\}$, is the input tuple and $v = (v_1, v_2, \ldots, v_d)$ is its corresponding output tuple. This is basically collecting the plaintexts and ciphertexts into two separate tuples. Now, the sum over S_2 can be rewritten into three *disjoint* sums as

$$\sum_{S_2} P = \sum_{S_3} A + \sum_{S_4} A + \sum_{S_5} A.$$

Here, S_3, S_4 and S_5 are the three partitions of S_2, i.e. $S_2 = S_3 \cup S_4 \cup S_5$, where
$S_3 = \Big\{(v,r),(v',r') \in \mathcal{T} \,|\, \forall i,j,k,m,e,f \ u_i \neq u'_j, x_k \neq x'_m, y_e \neq y'_f \Big\}$,
$S_4 = \Big\{(v,r),(v',r') \in \mathcal{T} \,|\, (\forall i,j,k,m \ u_i \neq u'_j, x_k \neq x'_m) \wedge (\exists e,f \ y_e = y'_f) \Big\}$,
$S_5 = \big\{(v,r),(v',r') \in \mathcal{T} | (\forall i,j \ u_i \neq u'_j) \wedge (\exists k,m \ x_k = x'_m) \big\}$, and A is
$\Pr_R[r] \Pr_R[r'] \Big(\Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] - \Pr_{C^*}[x \xrightarrow{C^*} y] \Pr_{C^*}[x' \xrightarrow{C^*} y'] \Big)$. Note that A and P are the same but written in different ways. We now deal with these three sums.

The sum $\sum_{S_3} A$ over S_3 (all non-colliding u's and u''s, all non-colliding x's and x''s, and all non-colliding y's and y''s) can be rewritten as

$$\begin{aligned} &\sum_{S_3} A \\ &\leq \frac{1}{2} \sum_{v,v'} \sum_{r,r'} \Big| \Pr_R[r] \Pr_R[r'] \Big(\Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] - \Pr_{C^*}[x \xrightarrow{C^*} y] \Pr_{C^*}[x' \xrightarrow{C^*} y'] \Big) \mathbf{1}_{S_3} \Big| \qquad (11) \\ &= \frac{1}{2} \Big| \Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] - \Pr_{C^*}[x \xrightarrow{C^*} y] \Pr_{C^*}[x' \xrightarrow{C^*} y'] \Big| \sum_{v,v'} \sum_{r,r'} \Pr_R[r] \Pr_R[r'] \ \mathbf{1}_{S_3}. \qquad (12) \end{aligned}$$

Here, as $\Big| \Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] - \Pr_{C^*}[x \xrightarrow{C^*} y] \Pr_{C^*}[x' \xrightarrow{C^*} y'] \Big|$ is constant when there is no collision between x and x' and between y and y', in the equality Line (12), we take it out from the sum. Afterwards, as we never have $a_i = v_j$ and $b_i \neq b_j$ according to Assumption 2, there will not be any inner-collisions in x. Therefore, when $x_i \neq x'_j$ and $y_i \neq y'_j$ for all i and j, we have

$$P_1 = \Pr\left[(x,x') \xrightarrow{C^*} (y,y')\right] = M^{-1}(M-1)^{-1} \cdots (M-2d+1)^{-1}$$

and

$$P_2 = \Pr\left[x \xrightarrow{C^*} y\right] = M^{-1}(M-1)^{-1} \cdots (M-d+1)^{-1}.$$

Then, we have

$$|P_1 - P_2^2| = \frac{1}{M} \frac{1}{M-1} \cdots \frac{1}{M-2d+1} - \frac{1}{M^2} \frac{1}{(M-1)^2} \cdots \frac{1}{(M-d+1)^2},$$

because $P_1 \geq P_2^2$.

Now, we bound the equality in Line (12) as

$$\frac{1}{2}\Big|\Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] - \Pr_{C^*}[x \xrightarrow{C^*} y]\Pr_{C^*}[x' \xrightarrow{C^*} y']\Big| \sum_{v,v'}\sum_{r,r'} \Pr_R[r]\Pr_R[r']\ \mathbf{1}_{S_3}$$

$$\leq \frac{1}{2}\left(\frac{1}{M(M-1)\cdots(M-2d+1)} - \frac{1}{M^2(M-1)^2\cdots(M-d+1)^2}\right)M^{2d} \tag{13}$$

$$\leq \frac{e^{8d^2/M}}{2} - \frac{d(d-1)}{2M} - \frac{1}{2}. \tag{14}$$

Note that the inequality in Line (13) is due to fact that the sum in Line (12) is bounded by the total number of v and v' which is M^{2d} and $P_1 \geq P_2^2$.

In order to prove the inequality in Line (14), we rewrite it as

$$\frac{1}{2}\left(\frac{1}{1-\frac{1}{M}}\frac{1}{1-\frac{2}{M}}\cdots\frac{1}{1-\frac{2d-1}{M}}\right) - \frac{1}{2}\left(\frac{1}{(1-\frac{1}{M})^2}\frac{1}{(1-\frac{2}{M})^2}\cdots\frac{1}{(1-\frac{d-1}{M})^2}\right).$$

Now, we maximize $(1-1/M)^{-1}(1-2/M)^{-1}\cdots(1-(2d-1/M))^{-1}$. We use two inequalities such that $(1-1/x)^{-1} \leq 1+2/x$ when $|x| \geq 2$, which holds for $x = M$ as $M \geq 2$ (according to the statement of Theorem 5) and $\big(1+r/k\big)^k \leq e^r$, when $1+r/k \geq 0$, then, the upper bound is

$$\frac{1}{1-\frac{1}{M}}\frac{1}{1-\frac{2}{M}}\cdots\frac{1}{1-\frac{2d-1}{M}} \leq e^{8d^2/M}.$$

In addition, we get

$$\frac{1}{(1-\frac{1}{M})^2}\frac{1}{(1-\frac{2}{M})^2}\cdots\frac{1}{(1-\frac{d-1}{M})^2} \geq 1 + \frac{d(d-1)}{M}.$$

by using *geometric series formula*, i.e. $(1-x)^{-1} = \sum_{n=0}^{\infty} x^n$ for $|x| < 1$ implying $(1-1/x)^{-1} \geq 1+1/x$ for $|x| > 1$. Hence, we get the desired upper bound for the expre

Furthermore, the sum over S_4, $\sum_{S_4} A$, will be the sum over all colliding y's and y''s, all non-colliding x's and x''s, and all non-colliding u's and u''s. When x and x' are non-colliding, it is not possible to have colliding y and y'. Hence, we have $\Pr_{C^*}\left[(x,x') \xrightarrow{C^*} (y,y')\right] = 0$. Therefore, the sum over S_4 will be negative, i.e.

$$\sum_{S_4} A \leq 0. \tag{15}$$

Finally, we provide a bound for the sum S_5, $\sum_{S_5} A$, as

$$\begin{aligned}
\sum_{S_5} A &\leq \sum_{v,v'}\sum_{r,r'} \Pr_R[r]\Pr_R[r']\Pr_{G^*}\left[(u,u') \xrightarrow{G^*} (v,v')\right]\ \mathbf{1}_{S_5} \\
&= \sum_g \Pr[G^*=g]\sum_{v,v'}\sum_{r,r'} \Pr_R[r]\Pr_R[r']\ \ \mathbf{1}_{(u,u')\xrightarrow{g}(v,v')}\ \mathbf{1}_{S_5} \qquad (16)\\
&= \mathbb{E}_{G^*}(\Pr_{r,r'}[\exists i,j \text{ s.t. } x_i = x'_j \mid \forall k,m \text{ s.t. } u_k \neq u'_m]) \\
&\overset{def}{=} \gamma \\
&\leq \frac{d^2}{M}. \qquad (17)
\end{aligned}$$

Here, we define $\gamma = \mathbb{E}_{G^*}(\Pr_{r,r'}[\exists i,j \text{ s.t. } x_i = x'_j \mid \forall k,m \text{ s.t. } u_k \neq u'_m])$ to be the probability that x and x' collide when there is no collision between u and u'.

We find $\gamma \leq d^2/M$ as follows. There is only one way for x and x' to collide when there is no collision between u and u'. This happens when a same query is from both u (resp. u') and v' (resp. v). In detail, let $u_i = (a_i, b_i)$ and $u'_j = (a'_j, b'_j)$ be two respective entries from u and u', and v_i and v'_j be their corresponding output. When $b_i = 0$, $b'_j = 1$ and $a_i = v'_j$, then there is a collision in x and x' such that

$x_i = x'_j$. As u and v' are independent, the probability that u and v' collide is less than $d^2/2M$. Similarly, we have the same result for u' and v.

Now, if we sum up all the results given in Lines (10), (14), (15) and (17), then we have

$$V(T(G^*)) \leq \theta + \frac{e^{8d^2/M}}{2} + \frac{d^2}{M} - \frac{1}{2}$$

by setting $d/2M \leq d^2/2M$.

When we substitute $V(T(G^*))$ in Line (8), then we have

$$|p - p^*| \leq 5\sqrt[3]{\Big(2\theta + e^{8d^2/M} + \frac{2d^2}{M} + \frac{3\varepsilon}{2} - 1\Big)n^2} + n\varepsilon.$$

□

Allowing $\theta \approx \delta$ to compare Theorem 1 with Theorem 5, we observe that the bound for adaptive attacks is *higher* than the bound for non-adaptive attacks. This fact comes with no surprise. Adaptive adversaries are stronger than non-adaptive adversaries, in general, and adaptive queries can provide the adversary with some advantages.

4.3 Some Important Remarks

It is worth mentioning that Theorem 5 of Section 4.2 poses two questions. Not surprisingly, similar questions were posed by Theorem 1 and we answered these questions in Section 3 by providing two counterexamples that are not intuitive. These counter-intuitive examples can also be applied to the adaptive case since the Feistel ciphers used in the solution to both questions are decorrelated by the adaptive norm, and non-adaptive attacks are a subset of adaptive attacks.

5 Generalization of Adaptive Iterated Distinguishers

Adaptive plaintext-ciphertext iterated distinguishers can be generalized in a way that the distinguisher in each iteration produces an information T_i which is not binary, that is, it takes values from a finite set $\{1, 2, \ldots, k\}$. Here, we find a bound for these generalized distinguishers as follows.

Theorem 6 *Let $G \in \mathcal{G}$ be a random function from $\mathcal{M} \times \{0,1\}$ to $\mathcal{M}$ such that $||[G]^{2d} - [G^*]^{2d}||_A \leq \varepsilon$, for some given $d \leq M/2$, where G^* is the ideal random function and $|\mathcal{M}| = M$. Let us consider* an adaptive *iterated distinguisher of order d, $\mathcal{A}_{\mathsf{AI}(d)}$, who is trying to distinguish G from G^* by performing n iterations (see Algorithm 6). We let T_i take values from the set $\{1, 2, \ldots, k\}$. Then, the advantage* $\mathrm{Adv}_{\mathcal{A}_{\mathsf{AI}(d)}}$ *of $\mathcal{A}_{\mathsf{AI}(d)}$ is*

$$\mathrm{Adv}_{\mathcal{A}_{\mathsf{AI}(d)}} \leq 5k\sqrt[3]{\Big(2\theta + e^{8d^2/M} + \frac{2d^2}{M} + \frac{3\varepsilon}{2} - 1\Big)n^2} + nk\varepsilon,$$

where θ is the probability that any two different iterations have at least one query in common.

Proof Let us define $T^j(g) = \Pr[T_i = j]$ for $j = 1, \ldots, k$, where T_i denotes the output of iteration i. As in Theorem 5, the advantage of the distinguisher will be

$$|p - p^*| = \Big| \Pr_{G, r_1, \ldots, r_n}[(T_1(G), \ldots, T_n(G)) \in \mathcal{A}cc] - \Pr_{G^*, r_1, \ldots, r_n}[(T_1(G^*), \ldots, T_n(G^*)) \in \mathcal{A}cc] \Big|.$$

We write

$$\Pr_{G, r_1, \ldots, r_n}[(T_1(G), \ldots, T_n(G)) \in \mathcal{A}cc] = \mathbb{E}_G\left(\sum_{(t_1, \ldots, t_n) \in \mathcal{A}cc} \prod_{i=1}^{n} T^{t_i}(G) \right) = \mathbb{E}_G(s(T^1(G), \ldots, T^k(G))),$$

where $s(\alpha_1, \ldots, \alpha_k) = \sum_{(t_1, \ldots, t_n) \in \mathcal{A}cc} \prod_{i=1}^{n} \alpha_{t_i}$.

Therefore, we get

$$|p - p^*| = |\mathbb{E}_G(s(T^1(G), \ldots, T^k(G))) - \mathbb{E}_{G^*}(s(T^1(G^*), \ldots, T^k(G^*)))|.$$

Let us define the hybrid as

$$s_j = s(T^1(G), \ldots, T^j(G), T^{j+1}(G^*), \ldots, T^k(G^*)),$$

where $s_k = s(T^1(G), \ldots, T^k(G))$ and $s_0 = s(T^1(G^*), \ldots, T^k(G^*))$. We obtain

$$\begin{aligned} |s(T^1(G), \ldots, T^k(G)) - s(T^1(G^*), \ldots, T^k(G^*))| &= |s_k - s_0| \\ &\leq \sum_{j=1}^{k} |s_j - s_{j-1}| \qquad (18) \\ &\leq 2n \sum_{j=1}^{k} |T^j(G) - T^j(G^*)|. \qquad (19) \end{aligned}$$

The inequality in Line (19) is same as the result in Line (4). Furthermore, the advantage of the distinguisher is

$$\begin{aligned} |p - p^*| &\leq \mathbb{E}(|s(T^1(G), \ldots, T^k(G)) - s(T^1(G^*), \ldots, T^k(G^*))|) \\ &\leq \sum_{j=1}^{k} \left(\frac{2V(T^j(G^*)) + \frac{3}{2}\varepsilon}{\lambda^2} + 2n\big(\frac{\varepsilon}{2} + 2\lambda\big) \right) \qquad (20) \\ &\leq 5k \sqrt[3]{\Big(2V(T(G^*)) + \frac{3\varepsilon}{2}\Big)n^2} + nk\varepsilon. \qquad (21) \end{aligned}$$

The inequality in Line (20) is analogous to Line (7). In order to prove inequality in Line (21), we consider another distinguisher with a single iteration such that its test function outputs 1 when $T_i = j$ otherwise (when $T_i \neq j$) 0. In detail, $\mathbb{E}_G(T^j) = \Pr[T^j = 1] = \Pr[T_i = j]$ by considering (w.l.o.g) $j = 1$ for every j, we notice that $\mathbb{E}_G(T^j) = \Pr[T_i = 1] = T(g)$ which is defined previously in Section 4 for binary-valued test functions. Therefore, we can replace $T^j(g)$ by $T(g)$.

Finally, by substituting the value of $V(T(G^*))$, found in the proof of Theorem 5, in to Line (21), we obtain the desired result. □

6 Resistance against Well-Known Statistical Distinguishers

In this section, we analyze two well-known statistical distinguishers , namely boomerang and differential-linear distinguishers, in the context of Decorrelation Theory. and provide security results dedicated for these attacks which improve the general result of Theorem 5. More explicitly, we here measure the highest success of an adversary with the need of the decorrelation distance of a cipher to the perfect cipher. Although, computing this distance is not practical for most of the block ciphers, it is a remarkable step in terms of provable security of block ciphers against these distinguishers. We show that having the decorrelation of order 4 protects ciphers against these attacks.

6.1 Differential-Linear Distinguishers

Differential-linear distinguishers, introduced by Langford and Hellman [12], are the combination of two seminal cryptanalysis techniques, namely differential and linear cryptanalysis. Basically, the adversary considers the cipher as a cascade of two sub-ciphers such that he uses a truncated differential characteristics for a sub-cipher and a linear approximation for the other sub-cipher. The attack is described in Algorithm 4.

The adversary aims at finding an efficient distinguisher with high differential-linear probability (DLP) which is defined by

$$\mathrm{DLP}^c(\alpha, b) = (2\Pr[b \cdot c(X) = b \cdot c(X \oplus \alpha)] - 1)^2,$$

where $\alpha, b \in \{0,1\}^m$, $m \geq 1$, and c is an instance of a block cipher C which is defined on a message space $\{0,1\}^m$ with cardinality $M = 2^m$, X is random variable on $\{0,1\}^m$.

In order to express differential probability in another way, we notice that

$$\mathbb{E}_X\big((-1)^{b\cdot c(X)+b\cdot c(X\oplus\alpha)}\big) = 2\Pr[b \cdot c(X) = b \cdot c(X \oplus \alpha)] - 1,$$

and, we get

$$\mathrm{DLP}^c(\alpha, b) = \mathbb{E}\Big((-1)^{b\cdot c(X_1)+b\cdot c(X_1\oplus\alpha)+b\cdot c(X_3)+b\cdot c(X_3\oplus\alpha)}\Big),$$

where X_1 and X_3 are independent and uniformly distributed random variables. We obtain the average DLP as

$$\begin{aligned}
&\mathbb{E}_C(\mathrm{DLP}^C(\alpha, b)) \\
&= M^{-2} \sum_{\substack{x_1,x_2,x_3,x_4 \\ y_1,y_2,y_3,y_4}} (-1)^{b\cdot y_1+b\cdot y_2+b\cdot y_3+b\cdot y_4} \Pr[(x_1, x_2, x_3, x_4) \xrightarrow{C} (y_1, y_2, y_3, y_4)] \mathbf{1}_{\substack{x_2=x_1\oplus\alpha \\ x_4=x_3\oplus\alpha}}.
\end{aligned}$$

For the computation of $\mathbb{E}_C(\mathrm{DLP}^C(\alpha, b))$ given above, we first divide the sums into two groups. The first group considers the terms $x_1 = x_3$ (which implies that $x_2 = x_4$, $y_1 = y_3$ and $y_2 = y_4$) whose contribution is M^{-1}. The second group is when $x_1 \neq x_3$ and $y_1 \neq y_3$ (which implies $x_2 \neq x_4$ and $y_2 \neq y_4$). We can split this group into four sums according to the two bits $(b \cdot y_1 \oplus b \cdot y_2,\ b \cdot y_3 \oplus b \cdot y_4)$. Let $\sum_{b_1,b_2}$ be the sum of all probabilities for (b_1, b_2), when $x_1 \neq x_3$ and $y_1 \neq y_3$. Then, we have

$$\mathbb{E}_C(\mathrm{DLP}^C(\alpha, b)) = 2^{-m} + 2^{-2m} \sum_{0,0} -2^{-2m} \sum_{1,0} -2^{-2m} \sum_{0,1} +2^{-2m} \sum_{1,1}.$$

The sum of four sums is $\sum_{0,0} + \sum_{0,1} + \sum_{1,0} + \sum_{1,1} = 2^m(2^m - 1)$. Therefore, we have

$$\begin{aligned}
\mathbb{E}_C(\mathrm{DLP}^C(\alpha, b)) &= 2^{-m} + 2^{-2m}(2^m(2^m - 1)) - 4 \times 2^{-2m} \sum_{0,1} \qquad (22) \\
&= 1 - 2^{2-2m} \sum_{\substack{x_1\neq x_3,x_2,x_4 \\ y_1\neq y_3,y_2,y_4}} 1_{\substack{b\cdot(y_1\oplus y_2)=0 \\ b\cdot(y_3\oplus y_4)\neq 0}} \Pr[(x_1, x_2, x_3, x_4) \xrightarrow{C} (y_1, y_2, y_3, y_4)] \mathbf{1}_{\substack{x_2=x_1\oplus\alpha \\ x_4=x_3\oplus\alpha}}.
\end{aligned}$$

Notice that in (22) we use the fact that $\sum_{1,0} = \sum_{0,1}$.

Notice that the expression in Line (22), we use the fact that $\sum_{1,0} = \sum_{0,1}$.

Now, we compute $\mathbb{E}_{C^*}(\mathrm{DLP}^{C^*}(\alpha, b))$ for using later in the proof of advantage of differential-linear distinguishers. We need to compute the sum $\sum_{0,1}$ which can be divided into two following cases.

1. Where each pair of tuples (x_1, x_2, x_3, x_4) and (y_1, y_2, y_3, y_4) contain pairwise different elements. Therefore, we have $\Pr[(x_1, x_2, x_3, x_4) \xrightarrow{C} (y_1, y_2, y_3, y_4)] = 1/(M(M-1)(M-2)(M-3))$. In this case, the number of tuples of pairwise different elements (y_1, y_2, y_3, y_4) satisfying both $b \cdot (y_1 \oplus y_2) = 0$ and $b \cdot (y_3 \oplus y_4) \neq 0$ is $4(M/2)^2(M/2-1)(M/2-2)$, and the number of tuples (x_1, x_2, x_3, x_4) satisfying the sum is $M(M-2)$.
2. Where $x_1 \neq x_3$, $x_2 \neq x_4$, $x_1 = x_4$, $x_2 = x_3$, $y_1 \neq y_3$, $y_2 \neq y_4$, $y_1 = y_4$, and $y_2 = y_3$. Here, there is no such tuples satisfying both $b \cdot (y_1 \oplus y_2) = 0$ and $b \cdot (y_3 \oplus y_4) \neq 0$ since $y_1 \oplus y_2 = y_3 \oplus y_4$.

Therefore, we obtain the expected DLP for C^* as

$$\mathbb{E}_{C^*}(\mathrm{DLP}^{C^*}(\alpha, b)) = \frac{2M - 5}{(M-1)(M-3)}.$$

Lemma 7 *Let $p(c)$ be the probability that differential-linear distinguisher, depicted in Figure 4, outputs 1. We let p_0 be probability that it outputs 1 when the counter is incremented with probability $\frac{1}{2}$ in each iteration instead of querying the oracle. We have*

$$|p(c) - p_0| \leq 2\sqrt{n\mathrm{DLP}^C(\alpha, b)}.$$

Proof Similar to Lemma 15 in [25]. □

Lemma 8 *Let C be a cipher on a message space $\mathcal{M} = \{0,1\}^m$ with $M = |\mathcal{M}|$, $m \geq 1$. For any differential-linear distinguisher between C and C^*, we get*

$$\mathrm{Adv}_{\mathrm{DL}} \leq 3\sqrt[3]{\mathrm{DLP}^{C}(\alpha,b)} + 3\sqrt[3]{n\mathbb{E}_{C^*}(\mathrm{DLP}^{C^*}(\alpha,b))}.$$

Proof Similar to Lemma 16 in [25]. □

Theorem 7 *Let C be a cipher on a message space $\mathcal{M} = \{0,1\}^m$ with $M = |\mathcal{M}|$, $m \geq 1$. For any differential-linear distinguisher between C and C^* with n iterations, we have*

$$\mathrm{Adv}_{\mathrm{DL}} \leq 3\sqrt[3]{n\|[C]^4 - [C^*]^4\|_\infty + \frac{n(2M-5)}{(M-1)(M-3)}} + 3\sqrt[3]{\frac{n(2M-5)}{(M-1)(M-3)}}.$$

Proof We previously compute $\mathbb{E}_{C^*}(\mathrm{DLP}^{C^*}(\alpha,b)) = (2M-5)/(M-1)(M-3)$. We know that

$$\left|\mathbb{E}_{C}(\mathrm{DLP}^{C}(\alpha,b)) - \frac{2M-5}{(M-1)(M-3)}\right| \leq \|[C]^4 - [C^*]^4\|_\infty.$$

This result is from Theorem 2. When iterated attacks have only one iteration, for example, for 2-limited attacks, the advantage will be bounded by the 2-decorrelation degree. However, due to the DLP which contains four inputs due to squaring of the probability, we consider the 4-decorrelation degree here. By using the previous lemma, we get the desired result. □

6.2 Boomerang Distinguishers

We bound the advantage of conventional boomerang distinguishers (see Algorithm 9). Although it is proven in Theorem 5 that in order to resist to the boomerang attack, a cipher needs to be 8-decorrelated. When we compute this bound specifically for the boomerang distinguisher, we show that the 4 decorrelation degree is enough. This is not surprising because a similar result has been found for differential distinguishers. Vaudenay proves in [25] that if a cipher is decorrelated to the order 2, it can resist to differential attacks although Theorem 1 says that decorrelation degree of 4 might be necessary.

Theorem 8 *Let C be a random cipher over the message space $\mathcal{M}$ with $|\mathcal{M}| = M \geq 2$, and C^* be the perfect cipher. The advantage of the boomerang distinguisher, $\mathrm{Adv}_{\mathrm{Boo}}$, depicted in Figure 9, is*

$$\mathrm{Adv}_{\mathrm{Boo}} \leq \frac{n(2M-5)}{(M-1)(M-3)} + \frac{n}{2}\|[C]^4 - [C^*]^4\|_A.$$

Proof First of all, let us define the differential probability of the boomerang distinguisher as

$$\mathrm{DP}^{c}_{\mathrm{Boo}}(\Delta,\nabla) = \Pr_X[c^{-1}(c(X)\oplus\nabla)\oplus c^{-1}(c(X\oplus\Delta)\oplus\nabla) = \Delta],$$

where Δ and ∇ are two strings denoting the input and the output non-zero differences, respectively. Here, $\mathrm{DP}^{c}_{\mathrm{Boo}}(\Delta,\nabla)$ is defined with a fixed key, that is, it is defined for an instance c of C. In order to generalize this for any key, we consider its expected value (average value) as

$$\mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta,\nabla)) = \frac{1}{M}\sum_{\substack{x_1,x_2,x_3,x_4\\ y_1,y_2,y_3,y_4}} \mathbf{1}_{\substack{x_1\oplus x_2=\Delta\\ x_3\oplus x_4=\Delta\\ y_1\oplus y_3=\nabla\\ y_2\oplus y_4=\nabla}} \Pr[(x_1,x_2,x_3,x_4) \xrightarrow{C} (y_1,y_2,y_3,y_4)].$$

Now, let us define $p(c)$ be the probability that the distinguisher outputs 1 for a fixed key (when $C = c$) which is

$$p(c) = 1 - (1 - \mathrm{DP}^{c}_{\mathrm{Boo}}(\Delta,\nabla))^n, \tag{23}$$

when the distinguisher has n iterations.

We compute the probability p (resp. p^*) that the distinguisher outputs 1 which is $p = \mathbb{E}_C(p(C))$ (resp. $p^* = \mathbb{E}_{C^*}(p(C^*))$). From the expression in Line (23), we have $p(c) \leq n\ \mathrm{DP}^{c}_{\mathrm{Boo}}(\Delta,\nabla))$ which implies that $p \leq n\ \mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta,\nabla))$.

Furthermore, $\mathbb{E}_{C^*}(\mathrm{DP}^{C^*}_{\mathrm{Boo}}(\Delta,\nabla))$ can be found as follows.

(a) When all x_i's are pairwise different and all y_i's are pairwise different, we have

$$\Pr[(x_1, x_2, x_3, x_4) \xrightarrow{C^*} (y_1, y_2, y_3, y_4)] = \frac{1}{M(M-1)(M-2)(M-3)}.$$

Here, there are $M(M-2)$ tuples for (x_1, x_2, x_3, x_4) and $M(M-2)$ tuples for (y_1, y_2, y_3, y_4) in the sum.

(b) When $x_1 = x_4$, $x_2 = x_3$, $x_1 \neq x_3$ and $x_2 \neq x_4$, $y_1 = y_4$, $y_2 = y_3$, $y_1 \neq y_3$ and $y_2 \neq y_4$, in this case,

$$\Pr[(x_1, x_2, x_3, x_4) \xrightarrow{C^*} (y_1, y_2, y_3, y_4)] = \frac{1}{M(M-1)}.$$

And, there are M possible tuples both for (x_1, x_2, x_3, x_4) and (y_1, y_2, y_3, y_4).
Note that the case where $x_1 = x_3$ results zero in the sum since it implies $y_1 = y_3$ which contradicts that ∇ is nonzero.

By combining the results found in Entries (a) and (b), we have

$$\mathbb{E}_{C^*}(\mathrm{DP}^{C^*}_{\mathrm{Boo}}(\Delta, \nabla)) = \frac{M-2}{(M-1)(M-3)} + \frac{1}{M-1} = \frac{2M-5}{(M-1)(M-3)}.$$

which implies that $p^* \leq n(2M-5)/(M-1)(M-3)$.

Therefore, we have

$$|p - p^*| \leq n \max\left(\mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta, \nabla)), \frac{2M-5}{(M-1)(M-3)}\right). \tag{24}$$

As the value $\mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta, \nabla))$ is not known for a random cipher, we write it in terms of the values that we know. Hence, we consider the case when $n = 1$ and then we switch to the general case. As the boomerang distinguisher is a 4-limited adaptive distinguisher when there is only one iteration, we know from Theorem 2 that

$$|p - p^*| \leq \frac{1}{2}\|[C]^4 - [C^*]^4\|_A. \tag{25}$$

When $n = 1$, we also have

$$|p - p^*| = \left|\mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta, \nabla)) - \frac{2M-5}{(M-1)(M-3)}\right|. \tag{26}$$

By combining the inequality in Line (25) and the equality in Line (26), we get

$$\mathbb{E}_C(\mathrm{DP}^{C}_{\mathrm{Boo}}(\Delta, \nabla)) \leq \frac{2M-5}{(M-1)(M-3)} + \frac{1}{2}\|[C]^4 - [C^*]^4\|_A.$$

For the case n, we have the desired result according to the inequality in Line (24). □

Notice that $\mathbb{E}_{C^*}(\mathrm{DLP}^{C^*}(\alpha, b)) = \mathbb{E}_{C^*}(\mathrm{DP}^{C^*}_{\mathrm{Boo}}(\Delta, \nabla))$. This is not surprising, a similar result appears in differential and linear attacks (see [25]), as well.

7 Conclusions

We settled an open problem and disproved a claim, both of which are raised by the EUROCRYPT '99 work of Vaudenay in Decorrelation Theory. In particular, we proved that in order for a cipher C to resist a non-adaptive iterated attack of order d, it is not sufficient to have a decorrelation of order $2d-1$. We showed this by providing a cipher decorrelated to the order $2d-1$ and a successful non-adaptive iterated attack against it which has order d. Hence, we concluded that the minimal order of decorrelation to ensure resistance is $2d$. Furthermore, we illustrated that when the probability of having a common query between different iterations increases, the advantage of the distinguisher *can* increase.

In this work, we also study the resistance against adaptive plaintext-ciphertext iterated distinguishers of order d which has not been explored before. We prove the bound for this distinguisher in which the adversary is making adaptive plaintext and ciphertext queries to the oracle depending on the previous queries. This work contributes to proving the security of previous and future designs based on Decorrelation Theory since, previously, there was no clue with adaptive iterated adversaries in this context. We then generalize these distinguishers and obtain a bound for the advantage of them.

Finally, we investigate boomerang and differential-linear distinguishers.

Acknowledgements This work was supported by the European Commission through the ICT program under contract ICT-2007-216646 ECRYPT II and the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center of the SNF under grant number 5005-67322.

References

1. Alon, N., Goldreich, O., Mansour, Y.: Almost k-wise independence versus k-wise independence. Electronic Colloquium on Computational Complexity (ECCC) **9**(048) (2002)
2. Baignères, T., Finiasz, M.: KFC - The Krazy Feistel Cipher. In: X. Lai, K. Chen (eds.) ASIACRYPT'06, *Lecture Notes in Computer Science*, vol. 4284. Springer (2006)
3. Baignères, T., Finiasz, M.: Dial C for Cipher. In: E. Biham, A.M. Youssef (eds.) SAC'06, *Lecture Notes in Computer Science*, vol. 4356, pp. 76–95. Springer (2007)
4. Baignères, T., Vaudenay, S.: Proving the Security of AES Substitution-Permutation Network. In: B. Preneel, S.E. Tavares (eds.) SAC'05, *Lecture Notes in Computer Science*, vol. 3897, pp. 65–81. Springer (2006)
5. Bay, A., Mashatan, A., Vaudenay, S.: Resistance against Adaptive Plaintext-Ciphertext Iterated Distinguishers. In: S.D. Galbraith, M. Nandi (eds.) INDOCRYPT'12, *Lecture Notes in Computer Science*, vol. 7668, pp. 528–544. Springer (2012)
6. Bay, A., Mashatan, A., Vaudenay, S.: Resistance against Iterated Attacks by Decorrelation Revisited. In: R. Safavi-Naini, R. Canetti (eds.) CRYPTO'12, *Lecture Notes in Computer Science*, vol. 7417, pp. 741–757. Springer (2012)
7. Carter, L., Wegman, M.N.: Universal Classes of Hash Functions. Journal of Computer and System Sciences **18**(2), 143–154 (1979)
8. Carter, L., Wegman, M.N.: New Hash Functions and Their Use in Authentication and Set Equality. Journal of Computer and System Sciences **22**(3), 265–279 (1981)
9. Chabaud, F., Vaudenay, S.: Links Between Differential and Linear Cryptoanalysis. In: A.D. Santis (ed.) EUROCRYPT'94, *Lecture Notes in Computer Science*, vol. 950, pp. 356–365. Springer (1995)
10. Cheon, D.H., Lee, S., Lim, J.I., Lee, S.J.: New Block Cipher DONUT Using Pairwise Perfect Decorrelation. In: B.K. Roy, E. Okamoto (eds.) INDOCRYPT'00, *Lecture Notes in Computer Science*, vol. 1977, pp. 262–270. Springer (2000)
11. Hoeffding, W.: Probability Inequalities For Sums Of Bounded Random Variables. jasa **58**, 13–30 (1963)
12. Langford, S.K., Hellman, M.E.: Differential-Linear Cryptanalysis. In: Y. Desmedt (ed.) CRYPTO'94, *Lecture Notes in Computer Science*, vol. 839, pp. 17–25. Springer (1994)
13. Luby, M.: A Simple Parallel Algorithm for the Maximal Independent Set Problem. SIAM J. Comput. **15**(4), 1036–1053 (1986)
14. Luby, M., Rackoff, C.: Pseudo-random Permutation Generators and Cryptographic Composition. In: J. Hartmanis (ed.) STOC'86, pp. 356–363. ACM (1986)
15. Luby, M., Rackoff, C.: How to Construct Pseudorandom Permutations from Pseudorandom Functions. SIAM J. Comput. **17**(2), 373–386 (1988)
16. Naor, J., Naor, M.: Small-bias Probability Spaces: Efficient Constructions and Applications. In: H. Ortiz (ed.) STOC'90, pp. 213–223. ACM (1990)
17. Nyberg, K.: Perfect Nonlinear S-Boxes. In: D.W. Davies (ed.) EUROCRYPT'91, *Lecture Notes in Computer Science*, vol. 547, pp. 378–386. Springer (1991)
18. Poupard, G., Vaudenay, S.: Decorrelated Fast Cipher: An AES Candidate Well Suited for Low Cost Smart Card applications. In: J.J. Quisquater, B. Schneier (eds.) CARDIS'98, *Lecture Notes in Computer Science*, vol. 1820, pp. 254–264. Springer (2000)
19. Vaudenay, S.: Provable Security for Block Ciphers by Decorrelation. In: M. Morvan, C. Meinel, D. Krob (eds.) STACS'98, *Lecture Notes in Computer Science*, vol. 1373, pp. 249–275. Springer (1998)
20. Vaudenay, S.: Feistel Ciphers with L_2-Decorrelation. In: S.E. Tavares, H. Meijer (eds.) SAC'98, *Lecture Notes in Computer Science*, vol. 1556, pp. 1–14. Springer (1999)
21. Vaudenay, S.: On the Lai-Massey Scheme. In: K.Y. Lam, E. Okamoto, C. Xing (eds.) ASIACRYPT'99, *Lecture Notes in Computer Science*, vol. 1716, pp. 8–19. Springer (1999)
22. Vaudenay, S.: Resistance Against General Iterated Attacks. In: J. Stern (ed.) EUROCRYPT'99, *Lecture Notes in Computer Science*, vol. 1592, pp. 255–271. Springer (1999)
23. Vaudenay, S.: Adaptive-Attack Norm for Decorrelation and Super-Pseudorandomness. In: H.M. Heys, C.M. Adams (eds.) SAC'99, *Lecture Notes in Computer Science*, vol. 1758, pp. 49–61. Springer (2000)
24. Vaudenay, S.: On Probable Security for Conventional Cryptography. In: J. Song (ed.) ICISC'99, *Lecture Notes in Computer Science*, vol. 1787, pp. 1–16. Springer (2000)
25. Vaudenay, S.: Decorrelation: A Theory for Block Cipher Security. Journal of Cryptology **16**(4), 249–286 (2003)
26. Wagner, D.: The Boomerang Attack. In: L.R. Knudsen (ed.) FSE'99, *Lecture Notes in Computer Science*, vol. 1636, pp. 156–170. Springer (1999)