# Linear Distinguishers in the Key-less Setting: Application to PRESENT

Martin M. Lauridsen and Christian Rechberger

DTU Compute, Technical University of Denmark, Denmark
{mmeh,crec}@dtu.dk

**Abstract.** The application of the concept of linear cryptanalysis to the domain of key-less primitives is largely an open problem. In this paper we, for the first time, propose a model in which its application is meaningful for distinguishing block ciphers.
Combining our model with ideas from message modification and rebound-like approaches, we initiate a study of cryptographic primitives with respect to this new attack vector and choose the lightweight block cipher PRESENT as an example target. This leads to known-key distinguishers over up to 27 rounds, whereas the best previous result is up to 18 rounds in the chosen-key model.

**Keywords:** hash function, block cipher, linear cryptanalysis, distinguisher, PRESENT

## 1 Introduction

We start off with a simple, clearly undesirable property of a block cipher and generalize it; suppose there is an $n$-bit block cipher which allows, for a particular known or chosen key, to determine a plaintext, such that the plaintext is the same as the ciphertext. For a good block cipher, accomplishing this should be very unlikely with much less than $2^n$ trials. It would, for example, allow preimage attacks in fully preimage-secure compression function constructions that use this block cipher.

Now, consider an $n$-bit block cipher where the key is known or chosen by the attacker and let us focus on a single bit at position $i$ of the plaintext $p_i$ and ciphertext $c_i$ in this setting. We would expect that the equation $p_i = c_i$ holds in exactly half the cases. In fact, any statistically significant deviation from this expectation can be interpreted as a sign of non-randomness in the cipher.

Such an attack would be in the so-called *key-less model*, which covers both the *known-key* and *chosen-key* models, and is hence of relevance if the cipher is used as part of a hash function construction. More generally, it allows to make meaningful statements and differentiate between ciphers beyond what is possible

This version is a minor revision of the FSE 2015 paper [30]

in other models. Should we consider such a cipher as a good building block for a compression function? Not if there would be an alternative cipher with similar implementation characteristics that does not allow for such a distinguisher!

### 1.1 Contributions

We discuss the two types of contributions in this paper. One is of a more conceptual/modeling nature, while the other is a concrete cryptanalytic application of the former.

**A New Way of Formulating Key-less Distinguishers.** The property described in the beginning resembles properties used in linear cryptanalysis to recover secret keys. The problem with the above line of reasoning was that so far there did not exist a meaningful model to properly express the setting. By this, we mean a model which has a proper characterization of the power of generic attackers and a clear distinction as to when a dedicated attack in fact can be considered a valid distinguisher, i.e. outperforms generic attackers. In this paper, after starting off by giving notation and preliminary notions of block ciphers and linear cryptanalysis in Section 2, we put in Section 3 the above very informal description of a possible demonstration of non-randomness on more rigorous grounds.

The usual requirement for a distinguisher to be valid is, that one must compare the cost of satisfying a specific property, which varies from case to case, for a concrete permutation $\pi$, with achieving the same property for an ideal permutation. In our model, we expand on this by posing the problem of determining for a concrete permutation $\pi$: i) a linear relation over $\pi$ in the form of an input/output mask and ii) a set of inputs to $\pi$, such that the number of inputs satisfying the linear relation is *expected* to deviate from what one expects of an ideal permutation, by a significant amount. A property which should not be attainable for an ideal primitive.

Our proposed key-less linear distinguisher model captures the possibility of distinguishing a cipher using any previous linear cryptanalysis, in the sense that the attacker needs only a linear hull and the probability distribution on the absolute correlation, to perform his analysis. To amplify the distinguisher to either cover more rounds or to need less computation, approaches inspired by message modification [43] and rebound attacks [36,28] are used.

**Application to PRESENT.** We can find concrete results in the new model in round-reduced versions of the leading lightweight-cipher PRESENT [10] (used in compression function designs advocated e.g. in [11]). In Section 4 we describe the relevant aspects of the PRESENT block cipher and give results on linear hulls and keys pertaining to it. Section 5 details the application of the key-less linear distinguisher to PRESENT. We fix a bit position $i$, devise an algorithm for determining up to $2^{61.97}$ key-dependent plaintexts in a very efficient manner, and study the expected number of plaintext and ciphertext pairs where $p_i = c_i$.

What we claim to be able to find is a deviation from the expectation that the equation $p_i = c_i$ is fulfilled with probability $\frac{1}{2}$. Depending on the size of the allowable key-set, this will work for up to 27 rounds of PRESENT. Detailed results are summarized in Table 4, before our conclusions and a discussion of open problems in Section 6. We confirm the results with experimental verifications (see Appendix C and [29]).

### 1.2 Related Work

Linear cryptanalysis, a technique to recover keys in ciphers, was pioneered by Matsui from 1992 on [33,35], with extensions or variants such as multiple linear approximations [5,20], linear hulls [39], multidimensional variants [16], zero-correlations [12] and considerations of a general statistical framework [3,31,38].

The application of linear cryptanalysis to key-less constructions, i.e. in models where the key is either known or chosen by the attacker, is largely an open problem. Sometimes, designs are evaluated with respect to standard linear cryptanalysis [2,32]. Some designers of SHA-3 candidates state properties with respect to this class of attacks (such as linear probability) without ever mentioning specific models. The reason is that there simply was no model, a situation that we address in this paper.

In all cases of linear cryptanalysis applied in a key-less setting, the analysis done is exactly the same as in a setting with a secret key: a linear approximation with a non-zero correlation is presented. The only known exception to us is a linear analysis of Cubehash by Ashur and Dunkelman [2]. There, an 11-round linear approximation with bias $2^{-235}$ is used to describe a standard distinguisher with $2^{470}$ queries. Then, inspired by a chosen-plaintext variant of linear cryptanalysis of DES by Knudsen and Mathiassen [23], the authors fix 80 bits of the plaintext input of modular additions, thereby gaining the first round for free, arriving at a 12-round result with a complexity below $2^{512}$. This can be seen as a predecessor to our deterministic technique of Section 5.2.

The only analysis of PRESENT in a setting without secret keys we know of is by Koyama, Sasaki, and Kunihiro [25]. In their work, differential chosen-key distinguishers (a setting that gives the attacker more freedom than in our known-key model) for up to 18 rounds are obtained.

At its core is a differential rebound attack with an inbound phase of 5 rounds that needs 100 degrees of freedom[1]. In the method we propose, we allow the key to be fixed arbitrarily, and out of the remaining 64 degrees of freedom from the plaintext input more than 61 degrees of freedom remain. Hence our results, that cover more rounds, and use our deterministic phase over 3 rounds that needs only 3 degrees of freedom, compare favorably to this result.

[1] Authors mention that 92 degrees of freedom out of 192 (from key and plaintext input) are left for the outbound phase

## 2 Preliminaries

In this section we introduce our notation, give basic definitions and recall known properties related to our analysis throughout the paper.

**Notation.** For an $n$-bit block cipher with key space $\mathcal{K}$, let $E : \mathbb{F}_2^n \times \mathcal{K} \to \mathbb{F}_2^n$ and $D : \mathbb{F}_2^n \times \mathcal{K} \to \mathbb{F}_2^n$ denote encryption and decryption functions, respectively. For convenience, we also use the notation that $E_K(x) := E(x, K)$ and $D_K(c) := D(c, K)$. We use $\sharp X$ to denote the size of a set $X$. For a real number $w$, $|w|$ denotes the absolute value of $w$. We let $\mathsf{Perm}(n)$ denote the set of all permutations on $n$-bit inputs and we let $x \xleftarrow{\$} X$ denote the assignment of $x$ by an element of $X$ chosen uniformly at random. We use $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{B}(n, p)$ to denote the normal- and binomial distributions respectively. For a distribution $D$ we use $\Phi(D, x)$ to denote the cumulative distribution function of $D$ at point $x$. We use the notation that $\mathbf{e}_i$ is a binary string with a 1 in position $i$ and zeroes elsewhere.

In this paper, when we talk about the key-less setting, we implicitly mean adversarial assumptions where the key $K \in \mathcal{K}$ is either known or chosen by the attacker.

**Trails and Hulls.** In the following, let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an iterated function of the form $F = F_R \circ \cdots \circ F_1$. We borrow to a large extent the notation from Leander's treatment on linear cryptanalysis [31]. We define a *mask* as a vector $\alpha \in \mathbb{F}_2^n$. For two masks $\alpha, \beta$, we denote by $\langle \alpha, \beta \rangle$ the inner product of the two masks:

$$\langle (\alpha_0, \ldots, \alpha_{n-1}), (\beta_0, \ldots, \beta_{n-1}) \rangle := \bigoplus_{i=0}^{n-1} \alpha_i \beta_i.$$

We define an $R$-round *trail* as an element $(\delta, \alpha_1, \ldots, \alpha_{R-1}, \gamma) \in (\mathbb{F}_2^n)^{R+1}$, where $\delta$ and $\gamma$ are the *input* and *output* masks, respectively. The $\alpha_i$ are called the *intermediate* masks. For a randomly chosen $x \in \mathbb{F}_2^n$, and for $i = 1, \ldots, R$ (letting $\alpha_0 = \delta$ and $\alpha_R = \gamma$), we have

$$\Pr\left[\langle x, \alpha_{i-1} \rangle = \langle F_i(x), \alpha_i \rangle\right] = \frac{1}{2} + \frac{\mathbf{C}_{F_i}(\alpha_{i-1}, \alpha_i)}{2},$$

where $\mathbf{C}_{F_i}(\alpha_{i-1}, \alpha_i)$ is the *correlation* over $F_i$. The *trail correlation* over $F$ is defined in terms of the $\mathbf{C}_{F_i}$ as

$$\mathbf{C}_F(\delta, \alpha_1, \ldots, \alpha_{R-1}, \gamma) = \mathbf{C}_{F_1}(\delta, \alpha_1) \left( \prod_{i=2}^{R-1} \mathbf{C}_{F_i}(\alpha_{i-1}, \alpha_i) \right) \mathbf{C}_{F_R}(\alpha_{R-1}, \gamma). \quad (1)$$

We say that a trail is *valid* if and only if each constituent correlation of (1) is non-zero.

We define an $R$-round *linear hull* $\mathrm{LH}_R(\delta, \gamma)$ as the union of all valid linear trails with input mask $\delta$ and output mask $\gamma$. As such, we use the notation that

$t \in \mathrm{LH}_R(\delta, \gamma)$ for an $R$-round trail $t$. Note that a linear hull $\mathrm{LH}_R(\delta, \gamma)$ defines an $R$-round *linear relation* between $x$ and $F(x)$, which we denote $\mathcal{R}^F_{\delta,\gamma} : \mathbb{F}_2^n \to \mathbb{F}_2$, where

$$\mathcal{R}^F_{\delta,\gamma}(x) = \begin{cases} 1 & , \langle x, \delta \rangle = \langle F(x), \gamma \rangle \\ 0 & , \langle x, \delta \rangle \neq \langle F(x), \gamma \rangle \end{cases}.$$

When $\mathcal{R}^F_{\delta,\gamma}(x) = 1$ we say the relation is *satisfied* for input $x$ and otherwise it is not. The *linear hull correlation* [17, Theorem 7.8.1] is given by

$$\mathbf{C}_F(\mathrm{LH}_R(\delta, \gamma)) = \sum_{t \in \mathrm{LH}_R(\delta,\gamma)} \mathbf{C}_F(t)$$

$$= \sum_{t \in \mathrm{LH}_R(\delta,\gamma)} (-1)^{\mathrm{sgn}(t)} \cdot |\mathbf{C}_F(t)|, \quad \mathrm{sgn}(t) = \begin{cases} 0 & , \mathbf{C}_F(t) \geq 0 \\ 1 & , \mathbf{C}_F(t) < 0 \end{cases}.$$

When the trail or hull is understood, we write $\mathbf{C}_F$ for simplicity to mean the correlation of the trail or hull over $F$. For a block cipher, the value of $\mathrm{sgn}(t)$ for $t \in \mathrm{LH}_R(\delta, \gamma)$ depends on the secret key $K \in \mathcal{K}$, and hence the value of $|\mathbf{C}_F(\mathrm{LH}_R(\delta, \gamma))|$ depends on the difference between the number of trails with $\mathrm{sgn}(t) = 1$ and those with $\mathrm{sgn}(t) = 0$. In this paper, we use the following assumption.

**Assumption 1** *For any fixed key $K \in \mathcal{K}$, we assume that for any two trails $t, t' \in LH_R(\delta, \gamma)$, where $t \neq t'$, the signs $\mathrm{sgn}(t)$ and $\mathrm{sgn}(t')$ are independent Bernoulli random variables with $p = \frac{1}{2}$.*

We note that Assumption 1 has been experimentally verified for PRESENT, see e.g. [13,31].

For readers familiar with differential-type attacks in the known-key setting, we offer the following loose analogy. We say that $x \in \mathbb{F}_2^n$ *follows* an $R$-round trail over $F$ if and only if

$$\langle x, \delta \rangle = \langle F_1(x), \alpha_1 \rangle = \cdots = \langle (F_{R-1} \circ \cdots \circ F_1)(x), \alpha_{R-1} \rangle = \langle F(x), \gamma \rangle.$$

This notion will be used in Section 5, when we describe how to use a technique similar to message modification, to extend a presented distinguisher in the key-less setting.

## 3 Key-Less Linear Distinguishers for Block Ciphers

Even though block ciphers have been used for a very long time, either implicitly or explicitly, to construct hash functions, a separate study of the security of block ciphers where the key is either known or under control of the adversary, has started only recently. Knudsen and Rijmen proposed so-called known-key distinguishers [24]. Later Biryukov, Khovratovich, and Nikolic [8] and Lamberger,

Mendel, Schläffer, Rechberger and Rijmen [27] proposed open- or chosen-key models to evaluate the security of block ciphers.

Even though these models often exhibit a rather contrived looking property, and evade a formally rigorous definition[2] (a property they share with collision attacks), cryptanalysts largely agree that these distinguishers are useful and interesting. Indeed, techniques developed to improve the original known-key distinguishers from [24], such as the rebound attack later led to collision attacks on various hash functions [21,37,27]. Also, the findings in the open-key model from [8] were later used to find the first related-key key-recovery attacks on AES-256 and AES-192 [6,7].

### 3.1 Motivation for our Distinguisher

Sometimes distinguisher descriptions are merely motivated by the fact that they *can* be formulated, as e.g. the 7-round known-key distinguisher on AES from [24], where byte-level zero-sums are used as a distinguishing property. Another example is the rotational rebound attack on reduced Skein [22], where the existence of *rotational collisions with errors* is defined as a distinguishing property. Sometimes, however, they are better motivated, e.g. by the construction of near-collisions, or the subspace- and limited-birthday distinguishers [19,27,28] that resemble some generalization of the concept of near-collisions.

The distinguisher we propose below comes with a new motivation that stems from preimage attacks on hash functions or compression functions[3]. As an example, consider the compression function construction using a single call to a block cipher in Matyas-Meyer-Oseas mode. The $i$th message block $m_i$ is compressed by using it as the plaintext input when computing the next chaining value $H_{i+1}$ using $H_i$ as the cipher key, i.e. $H_{i+1} = E_{H_i}(m_i) \oplus m_i$. If an attacker can determine a relation stating that the $j$th bit of $m_i$ equals the $j$th bit of $E_{H_i}(m_i)$ with a high probability, then it is likely that the $j$th bit of $H_{i+1}$ equals zero. In a preimage attack, if the target preimage is zero at position $j$, this then leads to an advantage over brute-force search.

Motivated by this example, we proceed with our new key-less linear distinguisher model for block ciphers that we will use throughout the paper.

### 3.2 The Key-less Linear Distinguisher Model

In the following, we give our definition of a model for key-less linear distinguishers. Essentially, the model captures the possibility of distinguishing any block cipher in the key-less setting, given that a linear relation (in the form of a linear hull) of sufficiently high absolute correlation for a reasonable fraction of the key space $\mathcal{K}$, is available. The notions of Definitions 1 and 2 are largely inspired by the recent work of Gilbert on pushing known-key attacks further on the AES [18].

[2] One exception being [1]

[3] We emphasize here that the application to PRESENT later in the paper will not be a preimage attack

The following definition of $\alpha$-*separability* formalizes how a linear relation, combined with a set of inputs for a permutation $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$, can exhibit a *significant* deviation from the behavior of a random permutation.

**Definition 1 ($\alpha$-separability).** *Let $\mathcal{P}$ be a set of permutations from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$ and let $\pi \in \mathcal{P}$ denote a particular, fixed permutation from $\mathcal{P}$. Let $\mathcal{S} \subseteq \mathbb{F}_2^n$ with size $\mathcal{M}$ and let $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$.*

*Without checking each input, each $x_i \in \mathcal{S}$ has an (a priori) associated probability $p_i = \Pr\left[\mathcal{R}_{\delta,\gamma}^{\pi}(x_i) = 1\right]$ that the linear relation is satisfied for that particular input. Let $\mathcal{X} = \sharp\{x \in \mathcal{S} \mid \mathcal{R}_{\delta,\gamma}^{\pi}(x) = 1\}$, then $\mathbb{E}[\mathcal{X}] = \sum_{i=1}^{\mathcal{M}} p_i$. We say that the tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{\pi})$ is* $\alpha$-separable *if and only*

$$\Pr\left[\left|\mathbb{E}[\mathcal{X}] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}\right] \geq \alpha,$$

*where the probability is taken over $\pi \in \mathcal{P}$.*

**Definition 2 ($(T, \mathcal{M}, \alpha)$-intractability).** *Let $\mathcal{P}$ be a set of permutations from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$ and let $\pi \in \mathcal{P}$ denote a particular, fixed permutation from $\mathcal{P}$. Let $\mathcal{S} \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$ and let $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$. We say that the tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{\pi})$ is* $(T, \mathcal{M}, \alpha)$-intractable *if and only if it is impossible, for any algorithm $\mathcal{A}$ to*

1. *Commit to a choice of $\delta', \gamma' \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ and*
2. *When given access to a fixed pair $\Pi, \Pi^{-1}$ with $\Pi \xleftarrow{\$} \mathsf{Perm}(n)$, construct a set $\mathcal{S}'$ of size $\mathcal{M}$ in time $T$, s.t. the tuple $(\mathsf{Perm}(n), \Pi, \mathcal{S}', \mathcal{R}_{\delta',\gamma'}^{\Pi})$ is $\alpha$-separable.*

*Note 1.* For our distinguisher model, the notion of *one time unit* corresponds to a single evaluation of the respective permutation.

With the definition of $\alpha$-separability and $(T, \mathcal{M}, \alpha)$-intractability in hand, we are ready to formulate our proposed key-less linear distinguisher.

**Definition 3 (Key-less linear distinguisher).** *Let $E : \mathbb{F}_2^n \times \mathcal{K} \to \mathbb{F}_2^n$ be a block cipher and let $\mathcal{E}$ to denote the set of permutations due to choices of the key $K \in \mathcal{K}$. Let $E_K$ denote* some *fixed permutation from $\mathcal{E}$.*

*Fix $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ and let $\mathcal{A}$ be an algorithm producing in time $T$ a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$. Then the tuple $(\mathcal{A}, \mathcal{E}, E_K, \mathcal{S}, T, \mathcal{R}_{\delta,\gamma}^{E_K}, \alpha)$ is said to be a* key-less linear distinguisher *if and only if $(\mathcal{E}, E_K, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{E_K})$ is both $\alpha$-separable and $(T, \mathcal{M}, \alpha)$-intractable.*

*Note 2.* In all of the definitions above, the fixed linear masks $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ are *chosen* by the algorithm $\mathcal{A}$, but the choice *must be made before* the production of the input set $\mathcal{S}$ commences.

In the context of distinguishing a block cipher, the adversary commits to $\delta$ and $\gamma$ and then obtains access to $E_K$ upon which the production of $\mathcal{S}$ in time $T$ begins. The parameter $\alpha$ directly expresses a lower bound on the fraction of the permutations $\pi \in \mathcal{P}$ for which the key-less linear distinguisher is valid. The time $T$ allowed to construct $\mathcal{S}$ is a parameter chosen by the adversary.

**Analysis.** In the following, we analyze and argue that the key-less linear distinguisher is meaningful. First, informally, the notion of $\alpha$-separability expresses that for a concrete permutation $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$, one can provide a linear relation which captures, for some constructed set of inputs, a *significant non-random behavior* in a permutation which is supposed to behave randomly. The *significant* part is captured by the requirement that the number of inputs satisfying the relation $\mathcal{R}^{\pi}_{\delta,\gamma}$ should deviate from what is expected in the ideal case by at least $\sqrt{\mathcal{M}}$. This reflects the usual requirement in linear cryptanalysis, that the data complexity is inversely proportional to the squared correlation. Second, on top of that, Definition 2 captures the notion that for a random permutation $\Pi \xleftarrow{\$} \mathsf{Perm}(n)$, it should not be possible, in the same amount of time, to provide such a relation with a set of inputs which exhibits the same significant non-random behavior.

With respect to Definition 2, one of the components to analyzing our proposed key-less linear distinguisher is to answer the following question: what is the *upper bound* on the probability $\alpha'$ that an algorithm $\mathcal{A}$, when given access to the fixed pair $\Pi$ and $\Pi^{-1}$, can produce in time $T$ a set $\mathcal{S}' \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$, together with a pre-determined relation $\mathcal{R}^{\Pi}_{\delta,\gamma}$, such that $(\mathsf{Perm}(n), \Pi, \mathcal{S}', \mathcal{R}^{\Pi}_{\delta,\gamma})$ is $\alpha'$-separable? Our analysis answers this question in the following, and it implicitly provides a *lower bound* on $\alpha$ for when a concrete permutation $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n \in \mathcal{P}$ (in the notation of Definitions 1 and 2) can be shown to be $(T, \mathcal{M}, \alpha)$-intractable, for fixed $T$ and $\mathcal{M}$. We begin our analysis with Lemma 1.

**Lemma 1.** *In the notation of Definition 2, let $\delta', \gamma' \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ be fixed non-zero linear masks, and let then an algorithm $\mathcal{A}$ be given access to $\Pi, \Pi^{-1}$, where $\Pi \xleftarrow{\$} \mathsf{Perm}(n)$. The optimal way for $\mathcal{A}$ to construct $\mathcal{S}' \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$ in time $T$ is the following:*

1. *Construct an arbitrarily chosen set $\mathcal{Q} \subseteq \mathbb{F}_2^n$ of size $T$.*
2. *Partition $\mathcal{Q}$ into $\mathcal{Q}_1 = \{x \in \mathcal{Q} \mid \mathcal{R}^{\Pi}_{\delta',\gamma'}(x) = 1\}$ and $\mathcal{Q}_0 = \{x \in \mathcal{Q} \mid \mathcal{R}^{\Pi}_{\delta',\gamma'}(x) = 0\}$ by querying $\Pi(x)$ for all $x \in \mathcal{Q}$ (this has time complexity $T$).*
3. *Set $\mathcal{S}'$ equal to the larger of the sets $\mathcal{Q}_0$ and $\mathcal{Q}_1$.*
4. *Fill up $\mathcal{S}'$ with arbitrarily chosen inputs from $\mathbb{F}_2^n \backslash \mathcal{Q}$ until $\sharp\mathcal{S}' = \mathcal{M}$.*

*Proof.* As $\Pi \xleftarrow{\$} \mathsf{Perm}(n)$, the particular choice of $\delta', \gamma' \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ does not affect the analysis. The most information $\mathcal{A}$ can learn about $\Pi$ in time $T$ is to obtain $T$ pairs $(x, \Pi(x))$, as is done when determining $\mathcal{Q}$ and its image under $\Pi$. In order to optimally shift the balance of the expected number of inputs of $\mathcal{S}'$ satisfying $\mathcal{R}^{\Pi}_{\delta',\gamma'}$ away from $\mathcal{M}/2$, $\mathcal{A}$ should take the larger of $\mathcal{Q}_1$ and $\mathcal{Q}_0$ and pool it with randomly chosen inputs $x$ for which the value of $\mathcal{R}^{\Pi}_{\delta',\gamma'}(x)$ is not known. □

Continuing our analysis, assuming an algorithm $\mathcal{A}$ constructs $\mathcal{S}'$ as in Lemma 1, we determine an upper bound on the value $\alpha'$ as a function of $\mathcal{M}$ and $T$, such that the resulting tuple $(\mathsf{Perm}(n), \Pi, \mathcal{S}', \mathcal{R}^{\Pi}_{\delta',\gamma'})$ is $\alpha'$-separable. We give this result in Theorem 1.

**Theorem 1 (Generic success probability).** *Let $\mathcal{A}, \Pi, \delta', \gamma', \mathcal{S}'$ and $T$ be as in Lemma 1, where $T \leq 4\sqrt{\mathcal{M}}$, and let $\mathcal{X} := \sharp\{x \in \mathcal{S}' \mid \mathcal{R}^{\Pi}_{\delta,\gamma}(x) = 1\}$. Then*

$$\Pr\left[\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}\right] = 2^{-T} \cdot \left[\sum_{k=0}^{T-2\sqrt{\mathcal{M}}} \binom{T}{k} + \sum_{k=2\sqrt{\mathcal{M}}}^{T} \binom{T}{k}\right].$$

*Proof.* First, note that $\sharp\mathcal{Q}_1 \sim \mathcal{B}(T, \frac{1}{2})$. We want to determine the probability that we have $\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$. The consideration is split into two cases depending on whether or not $\sharp\mathcal{Q}_1 \geq T/2$.

*Case $\sharp\mathcal{Q}_1 \geq T/2$.* In this case, we know that at least $\sharp\mathcal{Q}_1$ of the $\mathcal{M}$ inputs satisfy the relation. Thus, $\mathbb{E}\left[\mathcal{X}\right] = \mathbb{E}\left[Z\right] + \sharp\mathcal{Q}_1$ where $Z \sim \mathcal{B}\left(\mathcal{M} - \sharp\mathcal{Q}_1, \frac{1}{2}\right)$. Thus, $\mathbb{E}\left[\mathcal{X}\right] = \frac{\mathcal{M}+\sharp\mathcal{Q}_1}{2}$, and the requirement $\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$ is equivalent to either $\sharp\mathcal{Q}_1 \geq 2\sqrt{\mathcal{M}}$ or $\sharp\mathcal{Q}_1 \leq -2\sqrt{\mathcal{M}}$, the latter not being possible as $\sharp\mathcal{Q}_1$ is non-negative.

*Case $\sharp\mathcal{Q}_1 < T/2$.* In this case, we know that there are at least $T - \sharp\mathcal{Q}_1$ of the $\mathcal{M}$ inputs that *do not* satisfy the relation. Thus, $\mathbb{E}\left[\mathcal{X}\right] = \mathbb{E}\left[Z\right]$ where $Z \sim \mathcal{B}\left(\mathcal{M} - T + \sharp\mathcal{Q}_1, \frac{1}{2}\right)$. Thus, $\mathbb{E}\left[\mathcal{X}\right] = \frac{\mathcal{M}-T+\sharp\mathcal{Q}_1}{2}$, and the requirement $\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$ is equivalent to either $\sharp\mathcal{Q}_1 \geq T + 2\sqrt{\mathcal{M}}$ or $\sharp\mathcal{Q}_1 \leq T - 2\sqrt{\mathcal{M}}$, the former not being possible as $\sharp\mathcal{Q}_1 \leq T$.

In both cases considered, there is one event which makes the inequality $\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$ true. The combined probability of those two events is

$$\begin{aligned}&\Pr\left[\sharp\mathcal{Q}_1 \geq 2\sqrt{\mathcal{M}}\right] + \Pr\left[\sharp\mathcal{Q}_1 \leq T - 2\sqrt{\mathcal{M}}\right] \\ &= 2^{-T} \cdot \left[\sum_{k=0}^{T-2\sqrt{\mathcal{M}}} \binom{T}{k} + \sum_{k=2\sqrt{\mathcal{M}}}^{T} \binom{T}{k}\right].\end{aligned}$$

From this, the result follows. □

*Note 3.* The requirement $T \leq 4\sqrt{\mathcal{M}}$ in the statement of Theorem 1 arises because otherwise the two sums would overlap and add the same terms twice. The probability which is derived as a function of $\mathcal{M}$ and $T$ provides a lower bound on $\alpha$ for when, in the notation of Definition 2, a tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}^{\pi}_{\delta,\gamma})$ can be $(T, \mathcal{M}, \alpha)$-intractable. By using the normal approximation of $\sharp\mathcal{Q}_1$, i.e. $\sharp\mathcal{Q}_1 \sim \mathcal{N}\left(\frac{T}{2}, \frac{T}{4}\right)$, one obtains a very precise and easily-computable approximation of the probability as

$$1 - \Phi\left(\mathcal{N}\left(\frac{T}{2}, \frac{T}{4}\right), 2\sqrt{\mathcal{M}}\right) + \Phi\left(\mathcal{N}\left(\frac{T}{2}, \frac{T}{4}\right), T - 2\sqrt{\mathcal{M}}\right).$$

**Corollary 1.** *Let $\mathcal{A}$ be an algorithm which, after a choice of $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ is fixed, is given access to some permutation $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n \in \mathcal{P}$.*

*When $T < 2\sqrt{\mathcal{M}}$ and $\mathcal{P} = \mathsf{Perm}(n)$, it is impossible for $\mathcal{A}$ to produce in time $T$ a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$ s.t. the tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}^{\pi}_{\delta,\gamma})$ is $\alpha$-separable for any $\alpha > 0$.*

*On the other hand, when $T \geq 4\sqrt{\mathcal{M}}$ and $\mathcal{P} = \mathcal{E}$ (in the notation of Definition 3), then it is impossible for $\mathcal{A}$ to produce in time $T$ a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$ s.t. the tuple $(\mathcal{A}, \mathcal{P}, \pi, \mathcal{S}, T, \mathcal{R}^{\pi}_{\delta,\gamma}, \alpha)$ is a key-less linear distinguisher for any $\alpha > 0$.*

*Proof.* The first result follows directly from Theorem 1 when observing that the both sums are zero when $T < 2\sqrt{\mathcal{M}}$. The second result follows from Theorem 1 when observing that the sums equal one when $T = 4\sqrt{\mathcal{M}}$. This makes $(T, \mathcal{M}, \alpha)$-intractability impossible. □

*Note 4.* The key-less linear distinguisher specified in Definition 3 does not ask to provide outputs. Thus, it is not ruled out to give a valid key-less linear distinguisher without pre-computation, i.e. to have $T = 0$. Indeed, one of the concrete applications we show to the block cipher PRESENT does not need any computations.

From Corollary 1 it follows that when no pre-computation is allowed, i.e. when $T = 0$, any algorithm $\mathcal{A}$ producing a set $\mathcal{S} \subseteq \mathbb{F}_2^n$ together with any relation $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$ for a permutation $E_K \in \mathcal{E}$, yields a key-less linear distinguisher $(\mathcal{A}, \mathcal{E}, E_K, \mathcal{S}, T, \mathcal{R}^{E_K}_{\delta,\gamma}, \alpha)$ for *some* $\alpha > 0$. Note, however, that the parameter $\alpha$ measures how likely such a distinguisher is to succeed for a specific key. For example, when $\alpha$ is very low, one might have a valid key-less linear distinguisher for many rounds, but for a tiny fraction of the key space. As such, when $T = 0$, such a key-less linear distinguisher is to be taken with a grain of salt, depending on the value $\alpha$. In the following dicussions, we always provide together with our distinguishers the parameter $\alpha$, to make clear the lower bound on the fraction of the key space for which it is valid.

Having analyzed the generic case, we move on to stating in Theorem 2 a necessary condition for when, for a particular fixed $\pi \in \mathcal{P}$ and non-zero linear masks $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$, an algorithm $\mathcal{A}$ can construct $\mathcal{S} \subseteq \mathbb{F}_2^n$ of size $\mathcal{M}$ in time $T$, s.t. the tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}^{\pi}_{\delta,\gamma})$ is a $\alpha$-separable.

**Theorem 2.** *Let $\pi \in \mathcal{P}$ and fix non-zero linear masks $\delta, \gamma \in \mathbb{F}_2^n \backslash \{(0, \ldots, 0)\}$. Let $\mathcal{S} \subseteq \mathbb{F}_2^n$ have size $\mathcal{M}$. Then the tuple $(\mathcal{P}, \pi, \mathcal{S}, \mathcal{R}^{\pi}_{\delta,\gamma})$ can be $\alpha$-separable for $\alpha > 0$ if and only if the absolute correction $|\mathbf{C}_\pi|$ of $\mathcal{R}^{\pi}_{\delta,\gamma}$ satisfies $|\mathbf{C}_\pi| \geq 2/\sqrt{\mathcal{M}}$. Furthermore, the largest $\alpha$ for which $\alpha$-separability is obtained, is given by $\alpha = \Pr\left[|\mathbf{C}_\pi| \geq 2/\sqrt{\mathcal{M}}\right]$.*

*Proof.* Let $\mathcal{X} := \{x \in \mathcal{S} \mid \mathcal{R}^{\pi}_{\delta,\gamma}(x) = 1\}$. Then $\mathcal{X} \sim \mathcal{B}\left(\mathcal{M}, \frac{1}{2} + \frac{\mathbf{C}_\pi}{2}\right)$. We have $\alpha$-separability if and only if $\Pr\left[\left|\mathbb{E}\left[\mathcal{X}\right] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}\right] \geq \alpha$. Thus, we require either $\mathbb{E}\left[\mathcal{X}\right] \geq \frac{\mathcal{M}}{2} + \sqrt{\mathcal{M}}$ or $\mathbb{E}\left[\mathcal{X}\right] \leq \frac{\mathcal{M}}{2} - \sqrt{\mathcal{M}}$. Since $\mathbb{E}\left[\mathcal{X}\right] = \frac{\mathcal{M}}{2} + \mathcal{M} \cdot \frac{\mathbf{C}_\pi}{2}$, this happens exactly when $|\mathbf{C}_\pi| \geq 2/\sqrt{\mathcal{M}}$. From this, the results follow. □

## 4 The Block Cipher PRESENT, Keys and Linear Hulls

PRESENT is a 64-bit iterated block cipher [10] for use in lightweight applications such as RFID tags and wireless sensor networks. Its use in compression function designs is e.g. studied and advocated for in [11]. The key space is $\mathcal{K} = \mathbb{F}_2^\kappa$ with $\kappa$ either 80 or 128 bits. The respective block ciphers are denoted PRESENT-80 and PRESENT-128. Both ciphers have 31 rounds. The PRESENT key-schedule (see Appendix A for details) produces 32 $\kappa$-bit round keys, but only the 64 most significant bits are used in the key addition of each round. We refer to these 64-bit round keys as $K_i$ with $i = 0, \ldots, 31$.

The structure of PRESENT is a substitution-permutation network, repeating the round function

$$R_i(x) = P \circ S(x \oplus K_i),$$

where $x$ is the 64-bit state input to round $i$, $S$ is the parallel application of sixteen identical 4-bit S-boxes and $P$ is a fixed bitwise permutation[4]. The full cipher is composed of 31 applications of the round function followed by addition of a post-whitening key, i.e.

$$E_K = (R_{30} \circ \cdots \circ R_0)(x) \oplus K_{31}.$$

An illustration of a single round of PRESENT is given in Figure 1. For the specification of the PRESENT S-box and permutation $P$, see Appendix A.

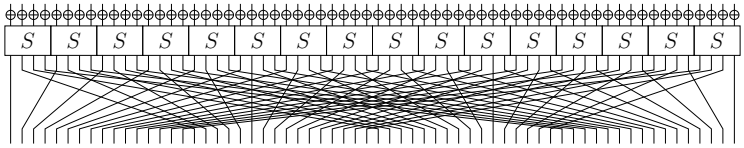


Fig. 1: Top-to-bottom illustration of a single round of PRESENT

### 4.1 Keys and Linear Hulls in PRESENT

One of the first thorough treatments of linear cryptanalysis on PRESENT is by Ohkuma [40]. This work defines *optimal linear trails* using solely masks of Hamming weight one. Furthermore, 64 *optimal hulls* using these trails are determined, along with the number of trails in each hull.

The absolute correlation for one of Ohkuma's $R$-round optimal trails $t$ is $|\mathbf{C}_{E_K}(t)| = 2^{-2R}$. Considering a particular $R$-round optimal hull $\mathrm{LH}_R(\delta, \gamma)$, let $T_R^+$ (respectively $T_R^-$) denote the number of trails $t$ in the hull for which $\mathrm{sgn}(t) = 0$ (respectively $\mathrm{sgn}(t) = 1$). We also let $T_R := \sharp\mathrm{LH}_R(\delta, \gamma)$, i.e. $T_R = T_R^+ + T_R^-$.

[4] $S$ and $P$ are called `sBoxLayer` and `pLayer`, respectively, in the specification.

By Assumption 1, for a fixed key $K \in \mathcal{K}$, we have $T_R^+ \sim \mathcal{B}\left(T_R, \frac{1}{2}\right)$, which for sufficiently large $T_R$ is well approximated by $T_R^+ \sim \mathcal{N}\left(\frac{T_R}{2}, \frac{T_R}{4}\right)$. Let $Z = T_R^+ - T_R^- = 2T_R^+ - T_R$. Thus, $Z$ is normally distributed with $\mu = 2 \cdot \frac{T_R}{2} - T_R = 0$ and $\sigma^2 = 2^2 \cdot \frac{T_R}{4} = T_R$, so $Z \sim \mathcal{N}(0, T_R)$. When $|Z| \geq N$, for some $N$, where $0 \leq N \leq T_R$, the absolute linear hull correlation is

$$|\mathbf{C}_{E_K}| \geq N \cdot 2^{-2R}.$$

Thus, there is a clear trade-off between the lower bound on $|\mathbf{C}_{E_K}|$ and the probability that a randomly chosen $K \in \mathcal{K}$ yields such a lower bound.

For the $T_R$ values, we refer to [40] or Table 6 in Appendix B. For a fixed number of rounds $R$, using the analysis above, $T_R$ can be used directly to determine i) a lower bound on $|\mathbf{C}_{E_K}|$ and ii) the probability that for a random $K \in \mathcal{K}$, this bound is obtained. Table 1 gives, for various probabilities $\alpha$ and number of rounds $R$ the value $\beta$ such that $\alpha = \Pr\left[|\mathbf{C}_{E_K}| \geq \beta\right]$. Table 7 in Appendix B gives the same data points for $R \in \{1, \ldots, 31\}$.

Table 1: Values $\log_2 \beta$ s.t. $\alpha = \Pr\left[|\mathbf{C}_{E_K}| \geq \beta\right]$ for $R$-round PRESENT

| | $\alpha$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0.01 | 0.05 | 0.10 | 0.30 | 0.50 | 0.70 | 0.90 | 0.95 | 0.99 |
| 7 | −9.55 | −9.94 | −10.20 | −10.86 | −11.48 | −12.29 | −13.91 | −14.91 | −17.23 |
| 11 | −14.74 | −15.14 | −15.39 | −16.06 | −16.68 | −17.48 | −19.10 | −20.10 | −22.43 |
| 16 | −21.27 | −21.66 | −21.92 | −22.58 | −23.20 | −24.01 | −25.63 | −26.63 | −28.95 |
| 24 | −31.71 | −32.11 | −32.36 | −33.03 | −33.65 | −34.46 | −36.07 | −37.07 | −39.40 |
| 26 | −34.33 | −34.72 | −34.97 | −35.64 | −36.26 | −37.07 | −38.68 | −39.69 | −42.01 |
| 28 | −36.94 | −37.33 | −37.58 | −38.25 | −38.87 | −39.68 | −41.30 | −42.30 | −44.62 |
| 31 | −40.85 | −41.25 | −41.50 | −42.17 | −42.79 | −43.60 | −45.21 | −46.22 | −48.54 |

*Example 1.* For $R = 28$, we have $T_{28} = 45170283840$. Thus, with probability $\alpha = 0.30$, a randomly chosen $K \in \mathcal{K}$ yields that one of Ohkuma's optimal hulls has $|\mathbf{C}_{E_K}| \geq 2^{-38.25}$.

## 5 Application to PRESENT

In this section we give key-less linear distinguishers on PRESENT for varying parameters; the number of rounds $R$; the pre-computation time $T$; the size $\mathcal{M}$ of the set $\mathcal{S}$ produced and the lower bound $\alpha$ on the fraction of the key space for which they are valid. PRESENT has previously received attention in the context of key-recovery attacks, especially with respect to linear cryptanalysis [13,15,31,40] on which our results build. The attack described is completely independent of the key size used, and hence also of the key schedule.

### 5.1 Probabilistic Phase

In this section we present key-less linear distinguishers on PRESENT using the model introduced in Section 3. We refer to approach described here as the *probabilistic phase*, which in Section 5.2 is combined with a *deterministic phase* to extend the distinguishers for three more rounds. The distinguishers we present here do not use any pre-computation, i.e. in the notation of the model, we have $T = 0$. Corollary 1 implies in this case that when $|\mathbf{C}_{E_K}| > 0$, the tuple produced by any algorithm $\mathcal{A}$ is always $(T, \mathcal{M}, \alpha)$-intractable for some $\alpha > 0$, and hence a valid distinguisher. The results match those of distinguishers used in key-recovery attacks and are as such of limited interest. We hope the discussion below makes it easier to follow (and appreciate) the real use of the model introduced, namely the case described in Section 5.2 when we do some, albeit very little, pre-computation.

In the following, let $\mathcal{R}^{E_K}_{\delta,\gamma}$ be the linear relation used, where $\delta = \gamma = \mathbf{e}_{21}$, which is one of the optimal linear hulls for PRESENT identified by Ohkuma. Also, let $\mathcal{A}$ be an algorithm constructing $\mathcal{S} \subseteq \mathbb{F}_2^n$ by picking $\mathcal{M}$ arbitrary $x \in \mathbb{F}_2^n$. In Table 2 we give, for various $\mathcal{M}$ and number of rounds $R$, lower bounds $\alpha$ on the fraction of the key space, s.t. $(\mathcal{A}, \mathcal{E}, E_K, \mathcal{S}, T = 0, \mathcal{R}^{E_K}_{\delta,\gamma}, \alpha)$ are key-less linear distinguishers.

Table 2: Lower bounds $\alpha$ on the fraction of the key space $\mathcal{K}$ susceptible to key-less linear distinguishers using $T = 0$, and the specified $\mathcal{M}$ and number of rounds $R$. A dash indicates that $\alpha < 0.00$.

| | Rounds $R$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{M}$ | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| $2^{40}$ | 0.96 | 0.89 | 0.74 | 0.41 | 0.04 | – | – | – | – | – | – | – | – | – |
| $2^{44}$ | 0.99 | 0.97 | 0.93 | 0.84 | 0.61 | 0.21 | – | – | – | – | – | – | – | – |
| $2^{46}$ | 0.99 | 0.99 | 0.97 | 0.92 | 0.80 | 0.53 | 0.12 | – | – | – | – | – | – | – |
| $2^{52}$ | 1.00 | 1.00 | 1.00 | 0.99 | 0.97 | 0.94 | 0.85 | 0.63 | 0.24 | – | – | – | – | – |
| $2^{54}$ | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.97 | 0.92 | 0.81 | 0.55 | 0.14 | – | – | – | – |
| $2^{56}$ | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.98 | 0.96 | 0.90 | 0.77 | 0.46 | 0.07 | – | – | – |
| $2^{62}$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.97 | 0.93 | 0.82 | 0.58 | 0.17 | – |
| $2^{63}$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.98 | 0.95 | 0.87 | 0.69 | 0.33 | 0.02 |
| $2^{64}$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.96 | 0.91 | 0.78 | 0.49 | 0.09 |

Note, that the $\alpha$ parameter from Table 2 gives immediately the probability that such an $R$-round key-less linear distinguisher without pre-computation for PRESENT is valid in practice, for a fixed chosen- or known key $K \in \mathcal{K}$. As examples, we see that with $\mathcal{M} = 2^{40}$, the probability of having a valid key-less linear distinguisher for 13-round PRESENT with a fixed key $K$ is *at least* $\alpha = 0.41$. Another example is a key-less linear distinguisher on 22-round PRESENT which is valid for a fraction of at least $\alpha = 0.33$ of the key space, using $\mathcal{M} = 2^{63}$.

### 5.2 Extension by Deterministic Phase

Next, we describe how one can use pre-computation to extend the key-less linear distinguishers from Section 5.1 to cover three more rounds with no degradation to the valid key space fraction $\alpha$. In the notation of the model, we now have $T > 0$, which in turn means that $(T, \mathcal{M}, \alpha)$-intractability is no longer granted for free by Corollary 1, unless below $T < 2\sqrt{\mathcal{M}}$. In Appendix D we outline an approach for a deterministic phase over 6 rounds, reminiscent of the rebound approach [36,28], which however has a computational complexity too high to fit into our model.

We describe in the following the algorithm $\mathcal{A}$ which will construct the set of inputs $\mathcal{S}$. The algorithm we give will construct $\mathcal{S}$ such that each $x \in \mathcal{S}$ is guaranteed to follow the linear trail $\mathcal{T} = (\mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21})$ over the first three rounds. We remark that this choice of trail is not unique; several others choices are possible, this is but one example. We refer to the approach we describe as the *deterministic phase*.

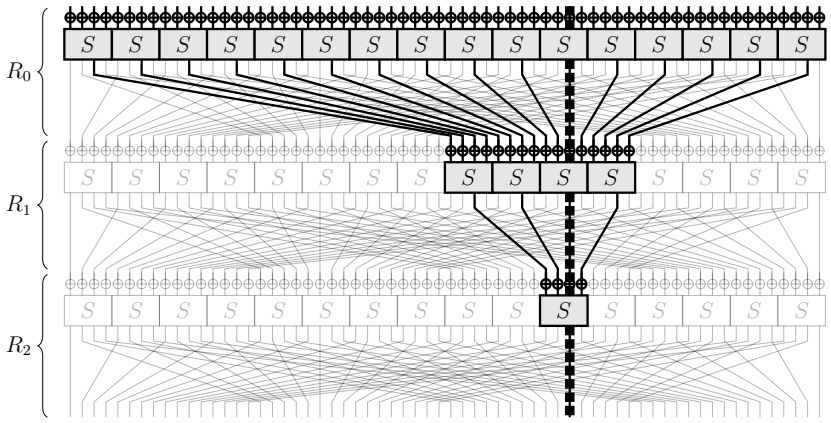


Fig. 2: Construction of $\mathcal{S}$ for 3-round PRESENT using the trail $\mathcal{T} = (\mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21})$. The highlighted parts show the S-boxes and key bits involved in the construction. The trail is indicated by the thick dotted line.

For notation, in round $r \in \{0, 1, 2\}$, let $S_{r,j}$ denote the $j$th S-box of round $R_r$ (counting from right to left) and let $K_{r,j}$ denote the $j$th least significant bit of the round key $K_r$, where all indices start from zero. Consider then $S_{2,5}$ which is highlighted in Figure 2. By inspection, the PRESENT S-box has 10 inputs $x$ which satisfy $\langle x, (0,0,1,0)\rangle = \langle S(x), (0,0,1,0)\rangle$ and hence follow the trail $(\mathbf{e}_{21}, \mathbf{e}_{21})$ over the round $R_2$, no matter what the inputs on the other S-boxes are. By adding the key bits $(K_{2,23} \| \cdots \| K_{2,20})$ to each $x$, we can trace those back through the permutation layer of the round $R_1$. For each value of $x \oplus (K_{2,23} \| \cdots \| K_{2,20})$, we now have a particular value on output bit 1 of each

of the S-boxes $S_{1,7}, \ldots, S_{1,4}$, as indicated in Figure 2. By the bijectivity of the S-box, it holds that for each of these S-boxes, half the inputs will give the desired output bit. However, for the S-box $S_{1,5}$ we have the extra requirement that the input bit on position 1 should equal the output bit on position 1, and only 5 inputs satisfy both properties simultaneously. As such, we can trace each of the ten values for $x$ back through $R_1$ and also adding the key bits $(K_{1,31} \| \cdots \| K_{1,16})$ to obtain $10 \cdot 8^3 \cdot 5 = 25600$ inputs to $R_2 \circ R_1$ which follow the trail $(\mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21})$ by construction. By tracing each of these values back through $R_0$ the same way, and adding the full round key $K_0$, algorithm $\mathcal{A}$ has a construction of the set $\mathcal{S}$ which consists of inputs which follow $\mathcal{T}$ over three rounds with probability 1. Using this approach to constructing $\mathcal{S}$, the size of the set can be *up to* $\mathcal{M} = 25600 \cdot 8^{15} \cdot 5 = 4503599627370496000 \approx 2^{61.97}$. As such, if one should wish to use a smaller $\mathcal{M}$ for the key-less linear distinguisher, this is also possible, simply by leaving out elements in the construction of $\mathcal{S}$.

Table 3: Tight values $\alpha$ such that $(\mathcal{E}, E_K, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{E_K})$ is $\alpha$-separable, where $E_K$ is $R$-round PRESENT for a fixed, known $K \in \mathcal{K}$ (and thus $E_K \in \mathcal{E}$)

| Rounds $R$ | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | 0.998 | 0.995 | 0.988 | 0.970 | 0.926 | 0.819 | 0.571 | 0.162 | 0.001 |

Consider $E_K$ being $R$-round PRESENT for a particular fixed $K \in \mathcal{K}$, and thus $E_K \in \mathcal{E}$. Let $\mathcal{A}$ be an algorithm for constructing $\mathcal{S}$ using the 3-round deterministic phase described, with $\mathcal{M} \approx 2^{61.97}$ for one of Ohkuma's optimal linear hull relations $\mathcal{R}_{\delta,\gamma}^{E_K}$. Table 3 gives, for various number of rounds $R$, the highest possible $\alpha$ s.t. $(\mathcal{E}, E_K, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{E_K})$ is $\alpha$-separable as per Definition 1. Of course, in order for the key-less linear distinguisher $(\mathcal{A}, \mathcal{E}, E_K, \mathcal{S}, T, \mathcal{R}_{\delta,\gamma}^{E_K}, \alpha)$ to be valid, it also has to hold that the tuple $(\mathcal{E}, E_K, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{E_K})$ is $(T, \mathcal{M}, \alpha)$-intractable as per Definition 2, where $T$ is the time required by $\mathcal{A}$ to construct the set $\mathcal{S}$.

In Section 5.3, we show that the time $T$ required to construct $\mathcal{S}$ by $\mathcal{A}$ is equivalent to $T = \frac{409641}{16R}$ calls to an $R$-round PRESENT encryption oracle. As such, we have that $T < 2\sqrt{\mathcal{M}}$, and from Corollary 1, it follows that $(\mathcal{E}, E_K, \mathcal{S}, \mathcal{R}_{\delta,\gamma}^{E_K})$ is $(T, \mathcal{M}, \alpha)$-intractable.

In Appendix C, we give examples of experimental verification of the key-less linear distinguishers presented on 9-round PRESENT. The code for this experimental verification is available as [29].

### 5.3 Computational Complexity

In this section we analyze the computational complexity, i.e. the time $T$ required by $\mathcal{A}$ to construct $\mathcal{S}$ in the deterministic phase of Section 5.2. In order to measure the time $T$ spent in this phase, we determine the number of S-box lookups

performed by $\mathcal{A}$ and then compare this to the number of S-box applications for a full call to the encryption oracle.

Let us consider all S-boxes as being different for generality, as the complexity in this case will certainly upper bound the case where they are all equal. In particular, since the key is known, this allows us to consider the key addition as part of the S-boxes. The analysis follows the construction of $\mathcal{S}$ by $\mathcal{A}$ itself, starting from round $R_2$ and working its way up (referring again to Figure 2). To determine the 10 inputs to $S_{2,5}$, $\mathcal{A}$ performs one lookup into this S-box. For each of these 10 values, one bit is traced back to an S-box of $R_1$, so this adds $10 \cdot 4$ S-box lookups. Finally, $\mathcal{A}$ has 25600 inputs to round $R_1$ for which it traces one bit back to each of the 16 S-boxes of round $R_0$, contributing by $25600 \cdot 16$ S-box lookups.

In total, the number of lookups is $1+10\cdot 4+25600\cdot 16 = 409641$. Now, comparing to the number of S-box lookups involved with a call to an $R$-round PRESENT oracle, the number of lookups would be $16R$, not counting key scheduling. As such, we find that the time $T$ spent by $\mathcal{A}$ for constructing $\mathcal{S}$ is $T = \frac{409641}{16R}$.

**Memory Complexity.** The memory complexity, though not a formal part of the key-less linear distinguisher model, is at a practical level. The storage of the set $\mathcal{S}$ can be encoded efficiently as follows. We define three sets

$$\begin{aligned} Q &= \{X \mid X_1 = S(X)_1\}, \\ S_0 &= \{X \mid S(X)_1 = 0\}, \quad \text{and} \\ S_1 &= \{X \mid S(X)_1 = 1\}. \end{aligned} \tag{2}$$

In a set $L$ we store the 25600 inputs which follow the trail $(\mathbf{e}_{21}, \mathbf{e}_{21})$ over $R_2 \circ R_1$. Let $X = X_{15} \| \cdots \| X_0$ denote one such 16-bit from $L$. The corresponding set of inputs to $R_0$ is now determined as the Cartesian product

$$S_{X_{15}} \times \cdots \times S_{X_6} \times Q \cap S_{X_5} \times S_{X_4} \times \cdots \times S_{X_0}. \tag{3}$$

The storage of $Q$, $S_0$ and $S_1$ take up 5 bytes, 4 bytes and 4 bytes, respectively. The storage of $L$ takes up 50 KB.

Using these simple observations, we give in Table 4 an overview of selected results for key-less linear distinguishers on $R$-round PRESENT. We give the size $\mathcal{M}$ of $\mathcal{S} \subseteq \mathbb{F}_2^n$ constructed by $\mathcal{A}$, the time $T$ required to do so, and the parameter $\alpha$ (implicitly, as we give $\alpha \cdot 2^{128}$) for the distinguisher, i.e. the lower bound on the fraction of the key space for which the distinguisher is valid. As such, the table is representative for PRESENT-128. Numbers for PRESENT-80 can be directly determined with the same $T$ and $\alpha \cdot 2^{80}$. Note, however, that for 27-round PRESENT-80 using $\mathcal{M} = 2^{61.97}$, $\alpha \cdot 2^{80} < 0$, so one can distinguish at most 26 rounds of PRESENT-80.

What is evident from Table 4 is, that there is a clear limit to how many rounds can be distinguished using a particular $\mathcal{M}$. This shows in the diagonal line through the table. Another observation is that for a fixed $\mathcal{M}$, there is a clear drop in the fraction of the key space $\alpha$ for which the distinguisher works

Table 4: Overview of parameters for key-less linear distinguishers on PRESENT. The entries give, for each $\mathcal{M}$ and each total number of rounds $R$ a pair $(\log_2 T, \log_2(\alpha \cdot 2^{128}))$ s.t. algorithm $\mathcal{A}$ can construct $\mathcal{S}$ in time $T$ and result in a distinguisher for *at least* a fraction $\alpha$ of the key space. Here, we indicate for PRESENT-128 the number of keys supporting the distinguisher. The equivalent number for PRESENT-80 is obtained as $\alpha \cdot 2^{80}$. A dash indicates that $\alpha \cdot 2^{128} < 0$.

| | Rounds $R$ | | | | | | |
|---|---|---|---|---|---|---|---|
| $\mathcal{M}$ | 14 | 18 | 22 | 25 | 26 | 27 | 28 |
| $2^{22}$ | – | – | – | – | – | – | – |
| $2^{25}$ | – | – | – | – | – | – | – |
| $2^{28}$ | $(-3.4, 70.9)$ | – | – | – | – | – | – |
| $2^{31}$ | $(-3.4, 119.2)$ | – | – | – | – | – | – |
| $2^{34}$ | $(-3.4, 126.2)$ | – | – | – | – | – | – |
| $2^{37}$ | $(-3.4, 127.5)$ | – | – | – | – | – | – |
| $2^{40}$ | $(-3.4, 127.8)$ | $(-3.8, 107.1)$ | – | – | – | – | – |
| $2^{43}$ | $(-3.4, 127.9)$ | $(-3.8, 124.3)$ | – | – | – | – | – |
| $2^{46}$ | $(-3.4, 128.0)$ | $(-3.8, 127.1)$ | – | – | – | – | – |
| $2^{49}$ | $(-1.7, 128.0)$ | $(-2.1, 127.7)$ | $(-2.4, 75.1)$ | – | – | – | – |
| $2^{52}$ | $(0.9, 128.0)$ | $(0.5, 127.9)$ | $(0.3, 119.8)$ | – | – | – | – |
| $2^{55}$ | $(3.9, 128.0)$ | $(3.5, 128.0)$ | $(3.2, 126.3)$ | – | – | – | – |
| $2^{58}$ | $(6.9, 128.0)$ | $(6.5, 128.0)$ | $(6.2, 127.5)$ | $(6.0, 103.1)$ | – | – | – |
| $2^{61}$ | $(9.9, 128.0)$ | $(9.5, 128.0)$ | $(9.2, 127.8)$ | $(9.0, 123.7)$ | $(9.0, 108.5)$ | $(8.9, 21.0)$ | – |
| $2^{61.97}$ | $(10.8, 128.0)$ | $(10.5, 128.0)$ | $(10.2, 127.9)$ | $(10.0, 125.4)$ | $(9.9, 117.1)$ | $(9.9, 71.8)$ | – |

between $R$ and $R+1$ rounds. For example, with $\mathcal{M} = 2^{61}$, we see a drop from $2^{108.5}$ keys supporting the distinguisher for 26 rounds to just $2^{21}$ for 27 rounds. What is also apparent is that in all cases, $T \ll 2\sqrt{\mathcal{M}}$, indeed sometimes $T < 1$, so by Corollary 1, $(T, \mathcal{M}, \alpha)$-intractability is for granted.

One thing worth discussion is the time complexity $T$. This is the time, converted to equivalent calls to an $R$-round encryption oracle, required by the key-less linear distinguisher algorithm $\mathcal{A}$ to construct the set $\mathcal{S}$. In a scenario where one would verify the distinguisher for a concrete block cipher $E_K$, i.e. for a particular value of $K$, one would need to determine the value of the random variable $\mathcal{X}$ of Definition 1. What we denote as the *verifying complexity* in this case is dominated by $\mathcal{M}$, because this is the number of inputs to the permutation that needs to be evaluated in order to determine $\mathcal{X}$.

## 6 Conclusion and Open Problems

In this paper we have formalized the notion of distinguishers for block ciphers using linear cryptanalysis in the key-less setting, i.e. where the block cipher is instantiated with a single known or chosen key.

The introduced key-less statistical distinguisher based on linear cryptanalysis led to a wide variety of results on PRESENT, for example a linear distinguisher of up to 26 and 27 rounds of PRESENT-80 and PRESENT-128, with respective computational complexities of about $2^9$ and $2^{10}$, and verifying complexities of about $2^{61}$ and $2^{61.97}$, for both PRESENT variants. The very low computational complexity made a practical verification possible for a reduced number of rounds, but also leaves room for improvements. For example, it is an open question whether it is possible to extend the deterministic phase to cover more rounds, while still keeping the work factor in the order of $2^{30}$ to allow for a valid distinguisher c.f. Corollary 1.

While PRESENT was chosen because it is a relatively high profile cryptanalytic target and the fact that relatively long useful linear hulls exist, we point out that the new distinguisher model is not specifically tailored for it. KATAN, a cipher with a very different round transformation and design philosophy, exhibits linear effects as described in [14] that makes it another interesting target for an application of the techniques introduced in this paper.

More research is needed on the relations between the use of degrees of freedom and the number of rounds that can be sidestepped, e.g. in our deterministic phase. Even though there is no good theoretical understanding of this yet, the literature already contains many data points for differential properties. The linear counterpart seems different and interesting enough to warrant a separate study, see also Appendix D.

The techniques we developed for the presented distinguisher might also have applications to preimage attacks that are inspired by linear cryptanalysis, or at least to somewhat speed-up brute-force preimage search. It will be interesting to see how this approach compares to other such methods [9,42]. Also, the approach naturally and directly applies to permutations, which become an increasingly

important primitive in their own right, also due to the popularization of the Sponge [4] construction.

## Acknowledgments

We would like to thank Mohamed Ahmed Abdelraheem, Dmitry Khovratovich, Gregor Leander, and Tyge Tiessen for helpful discussions on the paper.

## References

1. Elena Andreeva, Andrey Bogdanov, and Bart Mennink. Towards Understanding the Known-Key Security of Block Ciphers. In Shiho Moriai, editor, *Fast Software Encryption - 20th International Workshop, FSE 2013, Singapore, March 11-13, 2013. Revised Selected Papers*, volume 8424 of *Lecture Notes in Computer Science*, pages 348–366. Springer, 2013.
2. Tomer Ashur and Orr Dunkelman. Linear Analysis of Reduced-Round CubeHash. In Javier Lopez and Gene Tsudik, editors, *Applied Cryptography and Network Security - 9th International Conference, ACNS 2011, Nerja, Spain, June 7-10, 2011. Proceedings*, volume 6715 of *Lecture Notes in Computer Science*, pages 462–478, 2011.
3. Thomas Baignères, Pascal Junod, and Serge Vaudenay. How Far Can We Go Beyond Linear Cryptanalysis? In Pil Joong Lee, editor, *Advances in Cryptology - ASIACRYPT 2004, 10th International Conference on the Theory and Application of Cryptology and Information Security, Jeju Island, Korea, December 5-9, 2004, Proceedings*, volume 3329 of *Lecture Notes in Computer Science*, pages 432–450. Springer, 2004.
4. Guido Bertoni, Joan Daemen, Michael Peeters, and Gilles Van Assche. On the Indifferentiability of the Sponge Construction. In Nigel P. Smart, editor, *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, volume 4965 of *Lecture Notes in Computer Science*, pages 181–197. Springer, 2008.
5. Alex Biryukov, Christophe De Cannière, and Michaël Quisquater. On Multiple Linear Approximations. In Matthew K. Franklin, editor, *Advances in Cryptology - CRYPTO 2004, 24th Annual International CryptologyConference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings*, volume 3152 of *Lecture Notes in Computer Science*, pages 1–22. Springer, 2004.
6. Alex Biryukov, Orr Dunkelman, Nathan Keller, Dmitry Khovratovich, and Adi Shamir. Key Recovery Attacks of Practical Complexity on AES-256 Variants with up to 10 Rounds. In Henri Gilbert, editor, *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, May 30 - June 3, 2010. Proceedings*, volume 6110 of *Lecture Notes in Computer Science*, pages 299–319. Springer, 2010.
7. Alex Biryukov and Dmitry Khovratovich. Related-Key Cryptanalysis of the Full AES-192 and AES-256. In Matsui [34], pages 1–18.
8. Alex Biryukov, Dmitry Khovratovich, and Ivica Nikolic. Distinguisher and Related-Key Attack on the Full AES-256. In Shai Halevi, editor, *Advances in Cryptology*

- *CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, volume 5677 of *Lecture Notes in Computer Science*, pages 231–249. Springer, 2009.
9. Andrey Bogdanov, Dmitry Khovratovich, and Christian Rechberger. Biclique Cryptanalysis of the Full AES. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology - ASIACRYPT 2011 - 17th International Conference on the Theory and Application of Cryptology and Information Security, Seoul, South Korea, December 4-8, 2011. Proceedings*, volume 7073 of *Lecture Notes in Computer Science*, pages 344–371. Springer, 2011.
10. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings*, volume 4727 of *Lecture Notes in Computer Science*, pages 450–466. Springer, 2007.
11. Andrey Bogdanov, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, and Yannick Seurin. Hash Functions and RFID Tags: Mind the Gap. In Elisabeth Oswald and Pankaj Rohatgi, editors, *Cryptographic Hardware and Embedded Systems - CHES 2008, 10th International Workshop, Washington, D.C., USA, August 10-13, 2008. Proceedings*, volume 5154 of *Lecture Notes in Computer Science*, pages 283–299. Springer, 2008.
12. Andrey Bogdanov and Vincent Rijmen. Linear hulls with correlation zero and linear cryptanalysis of block ciphers. *Des. Codes Cryptography*, 70(3):369–383, 2014.
13. Stanislav Bulygin. More on linear hulls of PRESENT-like ciphers and a cryptanalysis of full-round EPCBC-96. *IACR Cryptology ePrint Archive*, 2013:28, 2013.
14. Christophe De Cannière, Orr Dunkelman, and Miroslav Knezevic. KATAN and KTANTAN - A Family of Small and Efficient Hardware-Oriented Block Ciphers. In Christophe Clavier and Kris Gaj, editors, *Cryptographic Hardware and Embedded Systems - CHES 2009, 11th International Workshop, Lausanne, Switzerland, September 6-9, 2009, Proceedings*, volume 5747 of *Lecture Notes in Computer Science*, pages 272–288. Springer, 2009.
15. Joo Yeon Cho. Linear Cryptanalysis of Reduced-Round PRESENT. In Pieprzyk [41], pages 302–317.
16. Joo Yeon Cho, Miia Hermelin, and Kaisa Nyberg. A New Technique for Multidimensional Linear Cryptanalysis with Applications on Reduced Round Serpent. In Pil Joong Lee and Jung Hee Cheon, editors, *Information Security and Cryptology - ICISC 2008, 11th International Conference, Seoul, Korea, December 3-5, 2008, Revised Selected Papers*, volume 5461 of *Lecture Notes in Computer Science*, pages 383–398. Springer, 2008.
17. Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Information Security and Cryptography. Springer, 2002.
18. Henri Gilbert. A Simplified Representation of AES. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014. Proceedings, Part I*, volume 8873 of *Lecture Notes in Computer Science*, pages 200–222. Springer, 2014.
19. Henri Gilbert and Thomas Peyrin. Super-Sbox Cryptanalysis: Improved Attacks for AES-Like Permutations. In Seokhie Hong and Tetsu Iwata, editors, *Fast Software Encryption, 17th International Workshop, FSE 2010, Seoul, Korea, February*

*7-10, 2010, Revised Selected Papers*, volume 6147 of *Lecture Notes in Computer Science*, pages 365–383. Springer, 2010.
20. Burton S. Kaliski Jr. and Matthew J. B. Robshaw. Linear Cryptanalysis Using Multiple Approximations. In Yvo Desmedt, editor, *Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings*, volume 839 of *Lecture Notes in Computer Science*, pages 26–39. Springer, 1994.
21. Dmitry Khovratovich, María Naya-Plasencia, Andrea Röck, and Martin Schläffer. Cryptanalysis of *Luffa* v2 Components. In Alex Biryukov, Guang Gong, and Douglas R. Stinson, editors, *Selected Areas in Cryptography - 17th International Workshop, SAC 2010, Waterloo, Ontario, Canada, August 12-13, 2010, Revised Selected Papers*, volume 6544 of *Lecture Notes in Computer Science*, pages 388–409. Springer, 2010.
22. Dmitry Khovratovich, Ivica Nikolic, and Christian Rechberger. Rotational Rebound Attacks on Reduced Skein. In Masayuki Abe, editor, *Advances in Cryptology - ASIACRYPT 2010 - 16th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 5-9, 2010. Proceedings*, volume 6477 of *Lecture Notes in Computer Science*, pages 1–19. Springer, 2010.
23. Lars R. Knudsen and John Erik Mathiassen. A Chosen-Plaintext Linear Attack on DES. In Bruce Schneier, editor, *Fast Software Encryption, 7th International Workshop, FSE 2000, New York, NY, USA, April 10-12, 2000, Proceedings*, volume 1978 of *Lecture Notes in Computer Science*, pages 262–272. Springer, 2000.
24. Lars R. Knudsen and Vincent Rijmen. Known-Key Distinguishers for Some Block Ciphers. In Kaoru Kurosawa, editor, *Advances in Cryptology - ASIACRYPT 2007, 13th International Conference on the Theory and Application of Cryptology and Information Security, Kuching, Malaysia, December 2-6, 2007, Proceedings*, volume 4833 of *Lecture Notes in Computer Science*, pages 315–324. Springer, 2007.
25. Takuma Koyama, Yu Sasaki, and Noboru Kunihiro. Multi-differential Cryptanalysis on Reduced DM-PRESENT-80: Collisions and Other Differential Properties. In Kwon et al. [26], pages 352–367.
26. Taekyoung Kwon, Mun-Kyu Lee, and Daesung Kwon, editors. *Information Security and Cryptology - ICISC 2012 - 15th International Conference, Seoul, Korea, November 28-30, 2012, Revised Selected Papers*, volume 7839 of *Lecture Notes in Computer Science*. Springer, 2013.
27. Mario Lamberger, Florian Mendel, Christian Rechberger, Vincent Rijmen, and Martin Schläffer. Rebound Distinguishers: Results on the Full Whirlpool Compression Function. In Matsui [34], pages 126–143.
28. Mario Lamberger, Florian Mendel, Martin Schläffer, Christian Rechberger, and Vincent Rijmen. The Rebound Attack and Subspace Distinguishers: Application to Whirlpool. *J. Cryptology*, 28(2):257–296, 2015.
29. Martin M. Lauridsen and Christian Rechberger. Source code for experimental validation. `https://github.com/mmeh/present-keyless`.
30. Martin M. Lauridsen and Christian Rechberger. Linear Distinguishers in the Keyless Setting: Application to PRESENT. In Gregor Leander, editor, *Fast Software Encryption - 22nd International Workshop, FSE 2015, Istanbul, Turkey, March 8-11, 2015, Revised Selected Papers*, volume 9054 of *Lecture Notes in Computer Science*, pages 217–240. Springer, 2015.
31. Gregor Leander. On Linear Hulls, Statistical Saturation Attacks, PRESENT and a Cryptanalysis of PUFFIN. In Kenneth G. Paterson, editor, *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory*

*and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings*, volume 6632 of *Lecture Notes in Computer Science*, pages 303–322. Springer, 2011.
32. Yunqiang Li and Wang Ailan. Linear Cryptanalysis for the Compression Function of Hamsi-256. In *Proceedings of the 2011 International Conference on Network Computing and Information Security - Volume 01*, NCIS '11, pages 302–306, Washington, DC, USA, 2011. IEEE Computer Society.
33. Mitsuru Matsui. Linear Cryptoanalysis Method for DES Cipher. In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer, 1993.
34. Mitsuru Matsui, editor. *Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings*, volume 5912 of *Lecture Notes in Computer Science*. Springer, 2009.
35. Mitsuru Matsui and Atsuhiro Yamagishi. A New Method for Known Plaintext Attack of FEAL Cipher. In Rainer A. Rueppel, editor, *Advances in Cryptology - EUROCRYPT '92, Workshop on the Theory and Application of of Cryptographic Techniques, Balatonfüred, Hungary, May 24-28, 1992, Proceedings*, volume 658 of *Lecture Notes in Computer Science*, pages 81–91. Springer, 1992.
36. Florian Mendel, Christian Rechberger, Martin Schläffer, and Søren S. Thomsen. The Rebound Attack: Cryptanalysis of Reduced Whirlpool and Grøstl. In Orr Dunkelman, editor, *Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers*, volume 5665 of *Lecture Notes in Computer Science*, pages 260–276. Springer, 2009.
37. Florian Mendel, Christian Rechberger, Martin Schläffer, and Søren S. Thomsen. Rebound Attacks on the Reduced Grøstl Hash Function. In Pieprzyk [41], pages 350–365.
38. Sean Murphy. The effectiveness of the linear hull effect. *J. Mathematical Cryptology*, 6(2):137–147, 2012.
39. Kaisa Nyberg. Linear Approximation of Block Ciphers. In Alfredo De Santis, editor, *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 439–444. Springer, 1994.
40. Kenji Ohkuma. Weak Keys of Reduced-Round PRESENT for Linear Cryptanalysis. In Michael J. Jacobson Jr., Vincent Rijmen, and Reihaneh Safavi-Naini, editors, *Selected Areas in Cryptography, 16th Annual International Workshop, SAC 2009, Calgary, Alberta, Canada, August 13-14, 2009, Revised Selected Papers*, volume 5867 of *Lecture Notes in Computer Science*, pages 249–265. Springer, 2009.
41. Josef Pieprzyk, editor. *Topics in Cryptology - CT-RSA 2010, The Cryptographers' Track at the RSA Conference 2010, San Francisco, CA, USA, March 1-5, 2010. Proceedings*, volume 5985 of *Lecture Notes in Computer Science*. Springer, 2010.
42. Christian Rechberger. On Bruteforce-Like Cryptanalysis: New Meet-in-the-Middle Attacks in Symmetric Cryptanalysis. In Kwon et al. [26], pages 33–36.
43. Xiaoyun Wang and Hongbo Yu. How to Break MD5 and Other Hash Functions. In Ronald Cramer, editor, *Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings*, volume 3494 of *Lecture Notes in Computer Science*, pages 19–35. Springer, 2005.

# A PRESENT Components

Table 5 gives the 4-bit S-box used in PRESENT. The bit-permutation $P$ is defined s.t. bit $i$ is moved to bit $P(i)$ where

$$P(i) = 16 \cdot (i \bmod 4) + 4 \cdot \left\lfloor \frac{i}{16} \right\rfloor + \left\lfloor \frac{i \bmod 16}{4} \right\rfloor .$$

Algorithms 1 and 2 give pseudo-code for the key-scheduling algorithms for PRESENT-80 and PRESENT-128, respectively.

Table 5: The 4-bit PRESENT S-box in hexadecimal notation

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S[x]$ | C | 5 | 6 | B | 9 | 0 | A | D | 3 | E | F | 8 | 4 | 7 | 1 | 2 |

**Algorithm 1:** PRESENT-80 key schedule

**Data**: 80-bit master key $K$
**Result**: PRESENT-80 round key array $k_i, 0 \leq i \leq 31$
**for** $i \leftarrow 0$ **to** 31 **do**
  $k_i^\ell \leftarrow [K_{79} \| \ldots \| K_{16}]$;
  $k_i^r \leftarrow [K_{15} \| \ldots \| K_0]$;
  $[K_{79} \| K_{78} \| \ldots \| K_1 \| K_0] \leftarrow [K_{18} \| K_{17} \| \ldots \| K_{20} \| K_{19}]$;
  $[K_{79} \| K_{78} \| K_{77\|} K_{76}] \leftarrow S[K_{79} \| K_{78} \| K_{77} \| K_{76}]$;
  $[K_{19} \| K_{18} \| K_{17} \| K_{16} \| K_{15}] \leftarrow [K_{19} \| K_{18} \| K_{17} \| K_{16} \| K_{15}] \oplus (i+1)$;
**end**

**Algorithm 2:** PRESENT-128 key schedule

**Data**: 128-bit master key $K$
**Result**: PRESENT-128 round key array $k_i, 0 \leq i \leq 31$
**for** $i \leftarrow 0$ **to** 31 **do**
  $k_i^\ell \leftarrow [K_{127} \| \ldots \| K_{64}]$;
  $k_i^r \leftarrow [K_{63} \| \ldots \| K_0]$;
  $[K_{127} \| K_{126} \| \ldots \| K_1 \| K_0] \leftarrow [K_{66} \| K_{65} \| \ldots \| K_{68} \| K_{67}]$;
  $[K_{127} \| K_{126} \| K_{125} \| K_{124}] \leftarrow S[K_{127} \| K_{126} \| K_{125} \| K_{124}]$;
  $[K_{123} \| K_{122} \| K_{121} \| K_{120}] \leftarrow S[K_{123} \| K_{122} \| K_{121} \| K_{120}]$;
  $[K_{66} \| K_{65} \| K_{64} \| K_{63} \| K_{62}] \leftarrow [K_{66} \| K_{65} \| K_{64} \| K_{63} \| K_{62}] \oplus (i+1)$;
**end**

# B Data Pertaining to Correlation Bounding

Table 6 is the one determined by Ohkuma in [40], giving the number of optimal trails in an optimal hull for $R$ rounds with $R \in \{1, \ldots, 31\}$. Table 7 gives values $\log_2 \beta$ such that $\Pr[|\mathbf{C}_{E_K}| \geq \beta] = \alpha$ for various $\alpha$ and number of rounds $R$.

Table 6: Number of trails $T_R$ in optimal hull for $R$-round PRESENT, $R \in \{1, \ldots, 31\}$

| $R$ | $T_R$ | $R$ | $T_R$ | $R$ | $T_R$ | $R$ | $T_R$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 9 | 512 | 17 | 1140480 | 25 | 2517252696 |
| 2 | 1 | 10 | 1344 | 18 | 2985984 | 26 | 6590254272 |
| 3 | 1 | 11 | 3528 | 19 | 7817472 | 27 | 17253512704 |
| 4 | 3 | 12 | 9261 | 20 | 20466576 | 28 | 45170283840 |
| 5 | 9 | 13 | 24255 | 21 | 53582633 | 29 | 118257341400 |
| 6 | 27 | 14 | 63525 | 22 | 140281323 | 30 | 309601747125 |
| 7 | 72 | 15 | 166375 | 23 | 367261713 | 31 | 810547899975 |
| 8 | 192 | 16 | 435600 | 24 | 961504803 | | |

Table 7: Values $\log_2 \beta$ s.t. $\alpha = \Pr\left[|\mathbf{C}_{E_K}| \geq \beta\right]$ for $R$-round of PRESENT

| | $\alpha$ | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R$ | 0.01 | 0.05 | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | 0.95 | 0.99 |
| 1 | −0.63 | −1.03 | −1.28 | −1.64 | −1.95 | −2.25 | −2.57 | −2.93 | −3.38 | −3.98 | −4.99 | −6.00 | −8.32 |
| 2 | −2.63 | −3.03 | −3.28 | −3.64 | −3.95 | −4.25 | −4.57 | −4.93 | −5.38 | −5.98 | −6.99 | −8.00 | −10.32 |
| 3 | −4.63 | −5.03 | −5.28 | −5.64 | −5.95 | −6.25 | −6.57 | −6.93 | −7.38 | −7.98 | −8.99 | −10.00 | −12.32 |
| 4 | −5.84 | −6.24 | −6.49 | −6.85 | −7.16 | −7.46 | −7.78 | −8.14 | −8.58 | −9.19 | −10.20 | −11.20 | −13.53 |
| 5 | −7.05 | −7.44 | −7.70 | −8.06 | −8.36 | −8.66 | −8.98 | −9.35 | −9.79 | −10.40 | −11.41 | −12.41 | −14.73 |
| 6 | −8.26 | −8.65 | −8.90 | −9.26 | −9.57 | −9.87 | −10.19 | −10.55 | −11.00 | −11.60 | −12.61 | −13.62 | −15.94 |
| 7 | −9.55 | −9.94 | −10.20 | −10.56 | −10.86 | −11.16 | −11.48 | −11.85 | −12.29 | −12.90 | −13.91 | −14.91 | −17.23 |
| 8 | −10.84 | −11.24 | −11.49 | −11.85 | −12.16 | −12.46 | −12.78 | −13.14 | −13.58 | −14.19 | −15.20 | −16.20 | −18.53 |
| 9 | −12.13 | −12.53 | −12.78 | −13.14 | −13.45 | −13.75 | −14.07 | −14.43 | −14.88 | −15.48 | −16.49 | −17.50 | −19.82 |
| 10 | −13.44 | −13.83 | −14.09 | −14.45 | −14.75 | −15.05 | −15.37 | −15.74 | −16.18 | −16.78 | −17.80 | −18.80 | −21.12 |
| 11 | −14.74 | −15.14 | −15.39 | −15.75 | −16.06 | −16.36 | −16.68 | −17.04 | −17.48 | −18.09 | −19.10 | −20.10 | −22.43 |
| 12 | −16.05 | −16.44 | −16.69 | −17.05 | −17.36 | −17.66 | −17.98 | −18.34 | −18.79 | −19.39 | −20.40 | −21.41 | −23.73 |
| 13 | −17.35 | −17.75 | −18.00 | −18.36 | −18.67 | −18.97 | −19.29 | −19.65 | −20.09 | −20.70 | −21.71 | −22.71 | −25.04 |
| 14 | −18.66 | −19.05 | −19.30 | −19.66 | −19.97 | −20.27 | −20.59 | −20.95 | −21.40 | −22.00 | −23.01 | −24.02 | −26.34 |
| 15 | −19.96 | −20.36 | −20.61 | −20.97 | −21.28 | −21.58 | −21.90 | −22.26 | −22.70 | −23.31 | −24.32 | −25.32 | −27.65 |
| 16 | −21.27 | −21.66 | −21.92 | −22.28 | −22.58 | −22.88 | −23.20 | −23.56 | −24.01 | −24.61 | −25.63 | −26.63 | −28.95 |
| 17 | −22.57 | −22.97 | −23.22 | −23.58 | −23.89 | −24.19 | −24.51 | −24.87 | −25.32 | −25.92 | −26.93 | −27.93 | −30.26 |
| 18 | −23.88 | −24.27 | −24.53 | −24.89 | −25.19 | −25.49 | −25.81 | −26.18 | −26.62 | −27.23 | −28.24 | −29.24 | −31.56 |
| 19 | −25.19 | −25.58 | −25.83 | −26.19 | −26.50 | −26.80 | −27.12 | −27.48 | −27.93 | −28.53 | −29.54 | −30.55 | −32.87 |
| 20 | −26.49 | −26.89 | −27.14 | −27.50 | −27.80 | −28.11 | −28.42 | −28.79 | −29.23 | −29.84 | −30.85 | −31.85 | −34.17 |
| 21 | −27.80 | −28.19 | −28.44 | −28.80 | −29.11 | −29.41 | −29.73 | −30.09 | −30.54 | −31.14 | −32.15 | −33.16 | −35.48 |
| 22 | −29.10 | −29.50 | −29.75 | −30.11 | −30.42 | −30.72 | −31.04 | −31.40 | −31.84 | −32.45 | −33.46 | −34.46 | −36.79 |
| 23 | −30.41 | −30.80 | −31.06 | −31.42 | −31.72 | −32.02 | −32.34 | −32.71 | −33.15 | −33.75 | −34.77 | −35.77 | −38.09 |
| 24 | −31.71 | −32.11 | −32.36 | −32.72 | −33.03 | −33.33 | −33.65 | −34.01 | −34.46 | −35.06 | −36.07 | −37.07 | −39.40 |
| 25 | −33.02 | −33.41 | −33.67 | −34.03 | −34.33 | −34.63 | −34.95 | −35.32 | −35.76 | −36.37 | −37.38 | −38.38 | −40.70 |
| 26 | −34.33 | −34.72 | −34.97 | −35.33 | −35.64 | −35.94 | −36.26 | −36.62 | −37.07 | −37.67 | −38.68 | −39.69 | −42.01 |
| 27 | −35.63 | −36.03 | −36.28 | −36.64 | −36.95 | −37.25 | −37.57 | −37.93 | −38.37 | −38.98 | −39.99 | −40.99 | −43.31 |
| 28 | −36.94 | −37.33 | −37.58 | −37.94 | −38.25 | −38.55 | −38.87 | −39.23 | −39.68 | −40.28 | −41.30 | −42.30 | −44.62 |
| 29 | −38.24 | −38.64 | −38.89 | −39.25 | −39.56 | −39.86 | −40.18 | −40.54 | −40.98 | −41.59 | −42.60 | −43.60 | −45.93 |
| 30 | −39.55 | −39.94 | −40.20 | −40.56 | −40.86 | −41.16 | −41.48 | −41.85 | −42.29 | −42.90 | −43.91 | −44.91 | −47.23 |
| 31 | −40.85 | −41.25 | −41.50 | −41.86 | −42.17 | −42.47 | −42.79 | −43.15 | −43.60 | −44.20 | −45.21 | −46.22 | −48.54 |

## C Experimental Verification

In this section we describe experiments performed to verify the validity of the proposed key-less linear distinguishers. Concretely, we describe a key-less linear distinguisher $\mathcal{A}$ trying to distinguish 9-round PRESENT (regardless of key size). We let $\mathcal{R}_{\delta,\gamma}^{E_K}$ be the linear relation used in Section 5.

We know that to distinguish 9-round PRESENT, we can do a 3-round deterministic phase to construct $\mathcal{S}$ as described in Section 5. In this case, the probabilistic phase is 6 rounds. In the following, we give two examples using two different values $\alpha \in \{0.33, 0.75\}$.

For the first example, we have $\alpha = 0.33$. From Theorem 2, we know that to have 0.33-separability, we require that the event $|\mathbf{C}_{E_K}| \geq 2/\sqrt{\mathcal{M}}$ happens with probability (at least) $\alpha = 0.33$. From the analysis of Section 4.1, we find that for 6-round PRESENT we have $\Pr[|\mathbf{C}_{E_K}| \geq \beta] = \alpha$ for $\beta = 2^{-9.66}$. When using the inequality $|\mathbf{C}_{E_K}| \geq 2/\sqrt{\mathcal{M}}$, we find this bound is tight when $\mathcal{M} = 2619369$. As such, $\mathcal{A}$ is an algorithm for constructing an $\mathcal{S}$ of this size in the 3-round deterministic phase described in Section 5.2. For this key-less linear distinguisher $\mathcal{A}$, we now have 0.33-separability, because we expect that $\left|\mathbb{E}[\mathcal{X}] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$, where $\mathcal{X}$ is the number of inputs satisfying the linear relation.

The experimental part comes now from actually encrypting each $x \in \mathcal{S}$ under a fixed, known key $K \in \mathcal{K}$ using 9-round PRESENT. We then check each $\mathcal{R}_{\delta,\gamma}^{E_K}(x) = 1$ by checking the relation on the input/output pair. For our experiment, we repeated 1000 times the experiment of computing $\mathcal{X}$ for a random key $K$ and the corresponding set $\mathcal{S}$. We found that 389 keys satisfied $\left|\mathbb{E}[\mathcal{X}] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$, and as such we see that this fits with $\frac{389}{1000} \geq \alpha = 0.33$.

We repeated the same experiment again with $\alpha = 0.75$. In this case, we found that we require $\mathcal{M} = 24480331$. In the same way, we did 1000 experiments with random keys $K$ and found that 764 keys satisfied $\left|\mathbb{E}[\mathcal{X}] - \frac{\mathcal{M}}{2}\right| \geq \sqrt{\mathcal{M}}$. Again, this fits with $\frac{764}{1000} \geq \alpha = 0.75$.

## D 6-round Deterministic Phase

By combining the 3-round deterministic phase of Section 5.2 with another 3 rounds appearing before, it is possible to construct a 6-round deterministic phase, reminiscent of the rebound approach [36,28] (see Figure 3). The idea is, that for rounds 3 to 5, the same approach as in Section 5.2 is used. Also, the same approach is used, but going in the other direction, for rounds 0 to 2.

This describes a construction to *independently* obtain (i) a set of *outputs* from round $R_2$, for which the inputs follow the trail over the *first* three rounds and (ii) a set of *inputs* to $R_3$ which follow the trail over the *last* three rounds. These two sets meet at the same point: Right around the addition of the round key of round $R_3$. Thus, one can use said round key to determine a matching between the two sets, to obtain a set which has the desirable property of following the trail over both the top and bottom part. However, as the approaches are independent, there are constraints put on the round key of round $R_3$ due to both parts, and this loss in degrees of freedom must be taken into account.

While the technique described here is not directly applicable with our model, as it by nature needs to use several different keys to match the two sets, it could potentially be useful in chosen-key models which allow an adversary to make a statement using multiple different keys.

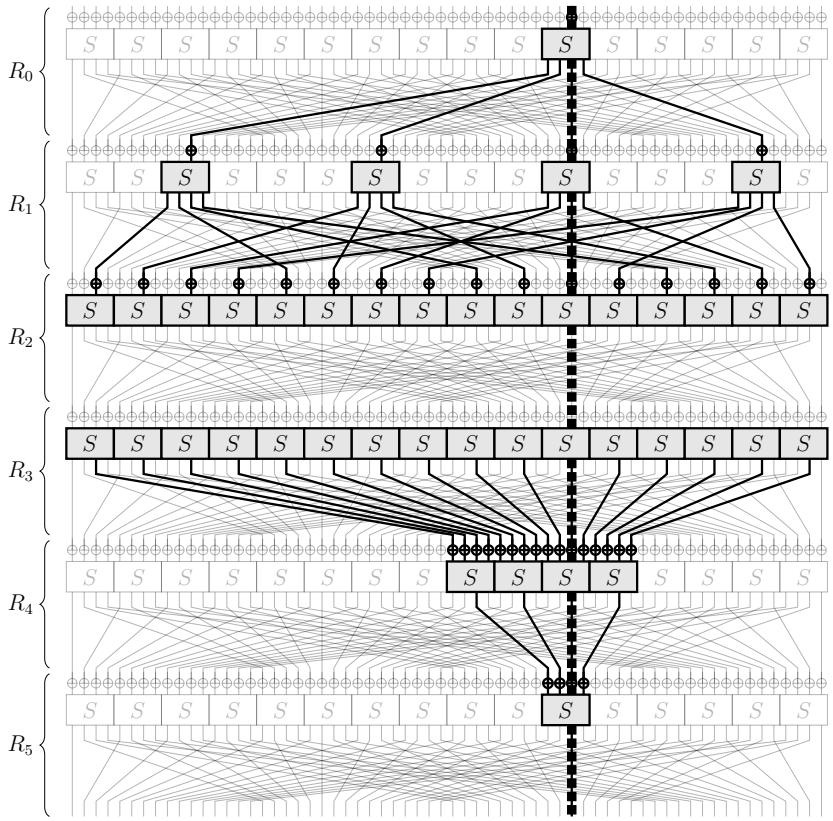


Fig. 3: 6-round deterministic phase for PRESENT using the trail $(\mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21}, \mathbf{e}_{21})$