Round-Based Approximation of (Higher-Order) Differential-Linear Correlation
A Geometric Approach Perspective

Kai Hu[1,3,4], Zhongfeng Niu[2(✉)], and Meiqin Wang[1,3,5(✉)]

[1] School of Cyber Science and Technology, Shandong University, Qingdao, China
kai.hu@sdu.edu.cn
[2] School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore
zhongfeng.niu@ntu.edu.sg
[3] State Key Laboratory of Cryptography and Digital Economy Security, Shandong University, Qingdao, China
mqwang@sdu.edu.cn
[4] Suzhou Research Institute, Shandong University, Suzhou 215123, China
[5] Quan Cheng Shandong Laboratory, Jinan, China

Abstract. This paper presents a new method for approximating the correlations of differential-linear distinguishers. From the perspective of Beyne's geometric approach, the differential-linear correlation is a corresponding coordinate of the *correlation vector* associated with the ciphertext multiset, which can be obtained by using the correlation matrix of the *2-wise form* of the cipher. The composite nature of the correlation matrix leads to a round-based approach to approximate the correlation vector. This simple approximation is remarkably precise, yielding the most accurate estimation for differential-linear correlations in `Ascon` thus far and the first DL distinguishers for 6-round `Ascon`-128a initialization. For `Present`, we present 17-round DL distinguishers, 4 rounds longer than the current record. To apply the round-based approach to ciphers with the large Chi (χ) function as nonlinear functions, we derive theorems to handle the correlation propagation for χ and its 2-wise form. Strong DL distinguishers for up to 6 rounds of `Subterranean`-2.0 and `Koala`-p are provided, 2 rounds longer than the previous differential and linear distinguishers. Furthermore, the round-based approximation idea is also extended to the higher-order differential-linear distinguishers. We give the first second-order DL distinguisher for 6-round `Ascon`-128 initialization and the first second-order DL distinguishers for up to 7 rounds of `Subterranean`-2.0 and `Koala`-p.

Keywords: Differential-linear · Geometric approach · Round-based · `Ascon` · `Subterranean`-2.0 · `Koala`-p

All appendices can be found in the full vesion https://eprint.iacr.org/2026/358.

J. Daemen and E. Thomé (Eds.): EUROCRYPT 2026, LNCS 16546, pp. 152–180, 2026.
https://doi.org/10.1007/978-3-032-25333-0_6

1 Introduction

Differential-linear (DL) cryptanalysis was proposed by Langford and Hellman in 1994 [25] to combine the two most effective cryptanalysis techniques: differential [12] and linear cryptanalysis [28]. In DL cryptanalysis, the adversary first finds a DL distinguisher where the correlation of a linear combination of the ciphertext difference bits under a fixed plaintext difference is significantly away from zero.

The classical method of detecting the DL distinguisher is to split the target cipher into two parts and combine a strong differential characteristic for the first part and a strong linear approximation for the second. This method is far from precise and sometimes even fails; thus, experimental computations are actually largely used to compute the real DL correlations for short rounds before. Since 2014, many theoretical methods have been proposed to enhance the precision of the DL correlation detection.

Related Works. In [13,14], Blondeau, Leander, and Nyberg introduced a precise formula to compute the DL correlation, where the sole assumption is the independence between the two parts of the cipher. However, it requires enumerating all trails, which is computationally infeasible. Later, Bar-On et al. proposed the Differential-Linear Connectivity Table (DLCT) to handle the independence assumption [2]. The DLCT is applied to one layer of S-boxes to switch the difference to linear masks. Recently, Hadipour, Derbez and Eichlseder generalized the DLCT framework to multiple inner rounds, with building several kinds of tables [22]. Meanwhile, Peng et al. also proposed a similar approach, which propagates the input difference forward for several rounds with truncated differences and then connects them to a linear trail through the DLCT [31].

In [27], Liu, Lu and Lin found another way called Differential Algebraic Transitional Form (DATF) to detect the DL correlations. The correlations of intermediate state difference bits are estimated according to the algebraic normal forms (ANFs) of the round functions. The final DL correlation is estimated based on the intermediate difference bits. This technique was later applied to the higher-order differential-linear (HDL) attacks [11] by Hu et al. [23]. Most recently, Che and Tian improved the DATF precision by proposing new bias estimation algorithms [16].

These methods have been mainly applied to S-box based ciphers. For Addition-Rotation-XOR (ARX) ciphers, Niu et al. derived an exact differential-linear propagation for the modular addition operation [30]. Recently, Gong et al. proposed an hourglass(-like) structure to capture the effective DL distinguishers [34]. Regarding ciphers that utilize large Chi (χ) functions, there is a limited number of research approaches in DL cryptanalysis, let alone the HDL cryptanalysis.

Motivation. Beyne's geometric approach [3–6] is a meta-methodology for cryptanalysis, transforming cryptanalysis into linear algebra in high-dimensional spaces, thereby simplifying their conceptual complexity. This geometric approach

framework has resolved many long-standing challenges. For instance, quasidifferential [7] techniques enable the study of differential probabilities under fixed-key settings; ultrametric integral cryptanalysis [9,10] allows probing whether the sum of ciphertext bits is a multiple of p^t (p is a prime and t is an integer); and a mixed-basis geometric approach [24] can automatically search for distinguishers based on multiple-of-8 property [21] with tools.

DL cryptanalysis has been incorporated into the geometric approach framework using a mixed-basis approach [24]. Theoretically, the exact DL correlation can be obtained by enumerating all differential-linear trails. However, due to the exceedingly weak correlation of individual trails, exhaustive enumeration is practically infeasible even for relatively simple ciphers. For example, Hu et al. performed trail enumeration for 7-round Simeck-32 and 8-round Simeck-48, confirming two DL distinguishers with correlation 1 previously reported by Hadipour et al. in [22]. Yet, for ciphers with more rounds or larger state sizes (e.g., Simeck-64), their method becomes computationally prohibitive. This reveals the inadequacy of trail-based geometric approaches for DL attacks, necessitating the development of new approximation techniques for estimating DL correlations.

Our Contributions. We begin by proposing a novel perspective for understanding DL correlations from the viewpoint of the geometric approach: the DL correlation corresponds to a coordinate of the *correlation vector* (Definition 2) of the ciphertext multiset. This correlation vector can be derived by multiplying the correlation matrix [5,17] of the 2-wise form (Definition 1) of the cipher with the correlation vector of the plaintext multiset.

Leveraging the properties of correlation matrices within the geometric approach framework, we adopt a round-based approximation strategy to estimate the ciphertext correlation vector, thereby enabling the estimation of DL correlations. This round-based approach offers several advantages:

1. It is highly straightforward, requiring only simple matrix and vector operations. For an Substitution-Permutation Network (SPN) cipher with t parallel m-bit S-boxes, the complexity of the basic method is $\mathcal{O}(t \times 2^{3m})$, allowing efficient correlation estimation within a short time.
2. It achieves high accuracy. With only minimal optimizations, we obtain the most precise and longest DL distinguishers known to date for both `Ascon` and `Present`. A detailed comparison with previous results is provided in Table 1.
3. It is theoretically sound, facilitating the estimation of DL correlations under fixed-key settings – an aspect often overlooked in prior work.

To extend this round-based approach to ciphers that utilize large χ, we derive exact propagation rules for the correlation of χ and its 2-wise form with a recursive matrix multiplication method. This enables the estimation of DL correlations for ciphers like `Subterranean`-2.0 and `Koala`, which employ large (257-bit) χ as nonlinear layers. We present 5-round and 6-round DL distinguishers for `Subterranean`-2.0 and `Koala`, surpassing the best previous differential and linear distinguishers by 2 rounds.

Furthermore, a similar methodology can be applied to HDL cryptanalysis. For an SPN cipher with t parallel m-bit S-boxes, the complexity of a d-th-order

Table 1. DL and HDL correlations detected by previous methods and our round-based method. All results here are without conditions. In [31], a 6-round Ascon-128 initialization DL correlation was reported, but it has been found invalid [16,33]. The entries marked with a ★ indicate that we obtained valid DL correlations for more rounds than prior work. TDT is short for *Truncated Difference Distribution Table*, the technique proposed in [31]. GDLCT is short for generalized DLCT, the technique presented in [22].

Target	Rounds	Exp. Cor.	Th. Cor.	Method	Ref.
Differential-linear distinguisher					
Ascon-128 init.	4	2^{-1}	2^{-5}	DLCT	[2]
			$2^{-1.36}$	ATF	[27]
			$2^{-1.09}$	HATF	[23]
			2^{-1}	TDT	[31]
			$\mathbf{2^{-1}}$	**Round-based**	Section 4.1
	5	$2^{-8.94}$	$2^{-9.1}$	TDT	[27]
			$\mathbf{2^{-8.94}}$	**Round-based**	Section 4.1
		$2^{-7.94}$	$\mathbf{2^{-7.94}}$	**Round-based**	Section 4.1
Ascon-128a init.	6 ★	–	$\mathbf{2^{-23.89}}$	**Round-based**	Section 4.1
Ascon-p	5	$2^{-4.33}$	$2^{-4.0}$	GDLCT	[22]
			$\mathbf{2^{-4.21}}$	**Round-based**	Section 4.1
		$2^{-7.61}$	$2^{-6.83}$	GDLCT	[22]
			$\mathbf{2^{-7.58}}$	**Round-based**	Section 4.1
	6 ★		$\mathbf{-2^{-21.89}}$	**Round-based**	Section 4.1
Present	13		$2^{-27.01}$	GDLCT	[22]
			$\mathbf{2^{-22.43}}$	**Round-based**	Section 4.2
	17 ★		$\mathbf{2^{-29.76}}$	**Round-based**	Section 4.2
Subterranean-2.0	5 ★	$-2^{-7.88}$	$\mathbf{-2^{-7.98}}$	**Round-based**	Section 6.1
	6 ★		$\mathbf{2^{-20.09}}$	**Round-based**	Section 6.1
Koala-p	5 ★	$2^{-6.73}$	$\mathbf{2^{-6.87}}$	**Round-based**	Section 6.2
	6 ★		$\mathbf{2^{-21.42}}$	**Round-based**	Section 6.2
Second-order differential-linear distinguisher					
Ascon-128 init.	5	$2^{-5.60}$	$2^{-6.05}$	HATF	[23]
			$\mathbf{2^{-5.63}}$	**Round-based**	Section 7.1
	6 ★		$\mathbf{2^{-45.35}}$	**Round-based**	Section 7.1
Subterranean-2.0	5 ★	$2^{-6.05}$	$\mathbf{2^{-6.05}}$	**Round-based**	Section 7.2
	7 ★		$\mathbf{2^{-90.99}}$	**Round-based**	Section 7.2
Koala-p	5 ★	$2^{-5.89}$	$\mathbf{2^{-6.09}}$	**Round-based**	Section 7.3
	7 ★		$\mathbf{2^{-86.24}}$	**Round-based**	Section 7.3

DL attack is $\mathcal{O}(t \times 2^{(2^d+1)m})$. We report the first second-order DL distinguisher targeting the 6-round initialization phase of `Ascon`-128. The technique is also applied to large χ, yielding the first 7-round second-order DL distinguishers for both `Subterranean`-2.0 and `Koala` (in the black-box setting). These results are also summarized in Table 1.

Source Codes. All sourse codes that can reproduce this paper's results are provided at https://github.com/hukaisdu/Round-based-approximation.git.

Organization of this Paper. The remaining part of this paper is organized as follows. Section 2 introduces necessary background knowledge for understanding this paper. In Sect. 3, we present our round-based approximation approach. Section 4 gives the applications of the new approach to S-box based `Ascon` and `Present`. We show how to use the round-based approximation for ciphers with large χ in Sect. 5 and Sect. 6. In Sect. 7, we give the results of our second-order DL distinguishers. Section 8 concludes the paper. In Appendix A, we give the theorems for HDL correlation approximation for both S-box and χ functions.

2 Preliminaries

2.1 Notations

In this paper, math italic alphabets represent vectors $x \in \mathbb{S}^n$ for a positive integer n, where $\mathbb{S}$ can be $\mathbb{F}_2$ (the finite field containing 0 and 1) or $\mathbb{R}$ (the real numbers) in this paper. The i-th coordinate of x is written as $x[i]$, where $x[0]$ is the leftmost element. The inner product of two vectors in $\mathbb{F}_2^n$ is written as $x^\top y = x^\top \cdot y = \sum_{i=0}^{n-1} x[i] \cdot y[i] \bmod 2$. The coordinate at the v-th row and u-th column of a matrix $\mathsf{M} \in \mathbb{S}^{n\times n}$ is written as $\mathsf{M}[v,u]$, also called the (v,u)-coordinate. If the index of a vector or a matrix consists of d parts, we write the index in a bracket, like (v_0, v_1). Thus, a coordinate of a matrix can be written like $\mathsf{M}[(v_0,v_1),(u_0,u_1)]$. In this paper, we do not distinguish between addition in $\mathbb{R}$ and addition in $\mathbb{F}_2$ in writing, as they can be naturally derived from the context.

Let n, m and t are three integers, where $n = mt$. If a vector $\gamma \in \mathbb{R}^{2^n}$ can be decomposed into a Kronecker product of t small vectors in $\mathbb{R}^{2^m}$, we write these small vectors as $\gamma^{[0]}, \gamma^{[1]}, \ldots, \gamma^{[t-1]} \in \mathbb{R}^{2^m}$, i.e.,

$$\gamma = \bigotimes_{i=0}^{t-1} \gamma^{[i]} \qquad \text{or} \qquad \gamma[(v_0, v_1, \ldots, v_{t-1})] = \prod_{i=0}^{t-1} \gamma^{[i]}[v_i].$$

For the sake of convenience, we introduce two maps to handle the change of indices.

Projection Maps. For $v = (v_0, v_1, \ldots, v_{t-1}) \in \mathbb{F}_2^{mt}$, a projection map is defined as

$$\varpi_i^t : \mathbb{F}_2^{mt} \to \mathbb{F}_2^m; \quad v \mapsto v_i.$$

Inclusion Maps. An inclusion map is defined as

$$\tau_i^t : \mathbb{F}_2^m \to \mathbb{F}_2^{mt}; \quad v \mapsto (0, \ldots, v, \ldots, 0) \quad (v \text{ is the } i\text{-th element, the length is } t).$$

If the information about t is clear from the context, we can omit t.

When the index element is a tuple, ϖ and τ work parallelly for each element in the tuple. For example, for $v = ((v_{0,0}, v_{0,1}), (v_{1,0}, v_{1,1}), \ldots, (v_{t-1,0}, v_{t-1,1}))$,

$$\varpi_i^t(v) = (v_{i,0}, v_{i,1})$$

and

$$\tau_i^t(\varpi_i^t(v)) = ((0,0), (0,0), \ldots, (v_{i,0}, v_{i,1}), \ldots, (0,0)).$$

Furthermore, the Kronecker delta function is used in this paper, which is defined as

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise.} \end{cases}$$

2.2 Differential-Linear and Higher-Order Differential-Linear Cryptanalysis

Differential-linear cryptanalysis was proposed by Langford and Hellman in 1994 [25]. For a cipher $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$, let $c = \mathsf{F}(p)$ and $c' = \mathsf{F}(p')$. Given a difference-mask pair (Δ, λ) where $p' = p + \Delta$, the correlation of a DL approximation, denoted by $\mathsf{DL}[\Delta \xrightarrow{\mathsf{F}} \lambda]$, can be derived from the following equation

$$\mathsf{DL}[\Delta \xrightarrow{\mathsf{F}} \lambda] = \mathsf{Cor}[\lambda^\top \cdot (c \oplus c')] := 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{\lambda^\top (\mathsf{F}(x) + \mathsf{F}(x+\Delta))}. \tag{1}$$

If $\mathsf{DL}[\Delta \xrightarrow{\mathsf{F}} \lambda]$ is significantly different from 0, we can distinguish the cipher from a random permutation.

Inspired by DL cryptanalysis, Biham, Dunkelman, and Keller studied other kinds of combined attacks including the higher-order differential-linear cryptanalysis [11] that combines a higher-order differential distinguisher and a linear approximation. Suppose $\boldsymbol{\Delta} = (\Delta_0, \Delta_1, \ldots, \Delta_{d-1})$ where all Δ_i's are linearly independent differences. $\mathsf{Span}(\boldsymbol{\Delta})$ represents the space that contains all linear combinations of components of $\boldsymbol{\Delta}$. The HDL correlation for F with the input difference $\boldsymbol{\Delta}$ and output mask λ, denoted by $\mathsf{HDL}[\boldsymbol{\Delta} \xrightarrow{\mathsf{F}} \lambda]$, is defined as

$$\mathsf{HDL}[\boldsymbol{\Delta} \xrightarrow{\mathsf{F}} \lambda] = \mathsf{Cor}\left[\lambda^\top \left(\sum_{\theta \in \mathsf{Span}(\boldsymbol{\Delta})} \mathsf{F}(x+\theta)\right)\right] := 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{\lambda^\top \sum_{\theta \in \mathsf{Span}(\boldsymbol{\Delta})} \mathsf{F}(x+\theta)} \tag{2}$$

2.3 Geometric Approach

The geometric approach in cryptanalysis was proposed by Beyne [5,6] as a meta approach to think about various cryptanalyses. With this approach, the linear

cryptanalysis [5], differential cryptanalysis [7], and integral cryptanalysis [8–10] have been reconsidered as special forms after the change-of-basis of a linear extension of the target cipher. Given a function $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ (this paper does not consider ciphers in characteristics larger than 2), the geometric approach first regards elements in $\mathbb{F}_2^n$ as a set of basis vectors, denoted by δ_u for $u \in \mathbb{F}_2^n$, and constructs a linear space with a field $\mathbb{K}$. The linear space is called a free vector space, defined as

$$\mathbb{K}[\mathbb{F}_2^n] := \left\{ \sum_{u \in \mathbb{F}_2^n} k_u \delta_u : k_u \in \mathbb{K} \right\}.$$

The linear extension of F is defined as

$$\mathsf{T}^{\mathsf{F}} : \mathbb{K}[\mathbb{F}_2^n] \to \mathbb{K}[\mathbb{F}_2^n]; \quad \sum_{u \in \mathbb{F}_2^n} k_u \delta_u \mapsto \sum_{u \in \mathbb{F}_2^n} k_u \delta_{\mathsf{F}(u)}.$$

If we regard $\{\delta_u, u \in \mathbb{F}_2^n\}$ as the standard basis, the matrix of T^{F} can be written out. For the sake of simplicity, we also write this matrix as T^{F} which would not cause any problem in this paper. By choosing another basis, for example, $\{\beta_u, u \in \mathbb{F}_2^n\}$, that satisfies

$$(\delta_0, \delta_1, \ldots, \delta_{2^n-1}) = (\beta_0, \beta_1, \ldots, \beta_{2^n-1})\mathsf{H}$$

where H is the change-of-basis matrix, we can do the change-of-basis operation on T^{F}. The new matrix similar to T^{F} is

$$\mathsf{C}^{\mathsf{F}} = \mathsf{H} \cdot \mathsf{T}^{\mathsf{F}} \cdot \mathsf{H}^{-1}.$$

When $\mathbb{K} := \mathbb{R}$ and H is the Walsh-Hadamard matrix, $\{\beta_u, u \in \mathbb{F}_2^n\}$ is called the linear basis [5], and C^{F} is called the *correlation matrix*. Note that the correlation matrix was for the first time introduced by Daemen, Govaerts, and Vandewalle in [17]. The (v, u)-coordinate of C^{F} is the famous expression used in linear cryptanalysis,

$$\mathsf{C}^{\mathsf{F}}[v, u] = 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{u^\top x + v^\top \mathsf{F}(x)}. \tag{3}$$

The above analysis is applicable to any function. In [24], Hu et al. used the *order* of an attack to describe the number of messages in an attack sample. For example, in differential cryptanalysis, a pair of two messages is used for the attack, so it is a second-order attack. However, the term *order* has been heavily used for describing some attacks like the higher-order differential attack. To avoid such ambiguity, we rename it and reformulate it as follows,

Definition 1 (d-wise form of a function [24]). *The d-wise form of a function $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is defined as*

$$F^{\times d} : \mathbb{F}_2^{n \times d} \to \mathbb{F}_2^{n \times d}$$
$$(x, \theta_1, \theta_2, \ldots, \theta_{d-1}) \mapsto (F(x), D_{\theta_1} F(x), \ldots, D_{\theta_{d-1}} F(x)).$$

$D_\theta F(x)$ is the derivative of F in the direction of θ at the point x, i.e., $D_\theta F(x) = F(x) + F(x+\theta)$.

Since $\mathsf{F}^{\times d}$ is also a function, we can apply the geometric approach to it. For example, considering the change-of-basis matrix as $\bigotimes_{i=1}^{d} H$, we obtain

$$\mathsf{C}^{\mathsf{F}^{\times d}} = \left(\bigotimes_{i=0}^{d-1} H\right) \cdot \mathsf{T}^{\mathsf{F}^{\times d}} \cdot \left(\bigotimes_{i=0}^{d-1} H\right)^{-1} = \left(\bigotimes_{i=0}^{d-1} H\right) \cdot \mathsf{T}^{\mathsf{F}^{\times d}} \cdot \left(\bigotimes_{i=0}^{d-1} H^{-1}\right),$$

where $\bigotimes$ is the Kronecker product. $\mathsf{C}^{\mathsf{F}^{\times d}}$ is called the correlation matrix of $\mathsf{F}^{\times d}$ or the d-wise correlation matrix of F. The corresponding basis is called the d-wise linear basis.

The coordinates of $\mathsf{C}^{\mathsf{F}^{\times d}}$ are indexed by d-tuples, such as

$$\mathsf{C}^{\mathsf{F}^{\times d}}[(v_0,\ldots,v_{d-1}),(u_0,\ldots,u_{d-1})] = 2^{-dn} \sum_{x,\theta_1,\ldots,\theta_{d-1}} (-1)^{u_0^\top x + v_0^\top \mathsf{F}(x) + \sum_{j=1}^{d-1} u_j^\top \theta_j + v_j^\top \mathsf{D}_{\theta_j}\mathsf{F}(x)}$$

For the sake of convenience, we sometimes use the bold alphabet $\boldsymbol{u} = (u_0,\ldots,u_{d-1})$ and $\boldsymbol{v} = (v_0,\ldots,v_{d-1})$ to represent the indices.

Theorem 1 ([5]). *The correlation matrix has the following properties.*

(1) If F is a r-round composite function $F = F_{r-1} \circ F_{r-2} \circ \cdots \circ F_0$, $C^F = \prod_{i=0}^{r-1} C^{F_i}$
(2) If F is a parallel application of t small functions $F = f||f||\cdots||f$, $C^F = \bigotimes_{i=0}^{t-1} C^f$, where $\bigotimes$ is the Kronecker product.

3 Round-Based Approximation

This section presents our main idea for efficiently and precisely approximating DL correlations. From the perspective of the geometric approach, the correlation of a DL distinguisher can be understood as a coordinate of a vector associated with the ciphertext under the 2-wise linear basis.

Previous work, notably by Hu et al. [24], utilized a mix-basis framework – employing a quasidifferential basis for the plaintext and a 2-wise linear basis for the ciphertext. A major drawback of this approach is its reliance on clustering a vast number of DL trails, a process that is often inefficient and computationally expensive.

In contrast, we propose a more efficient same-basis approach. We apply the 2-wise linear basis consistently to both the plaintext and ciphertext. This allows us to derive a 2-wise correlation matrix that linearly maps the (correlation) vector associated with the plaintext to the vector associated with the ciphertext. Leveraging the composite property of this correlation matrix (Theorem 1), we can efficiently approximate the correlation for each intermediate round state.

3.1 Correlation Vector and Correlation Matrix

We begin by defining the correlation vector of a multiset, which is the vector that represents the multiset under the linear basis.

Definition 2 (Correlation vector of a multiset). *The correlation vector of a multiset* $\mathbb{S}$ *whose values are taken from* $\mathbb{F}_2^n$ *is a vector* $\mathsf{CV}(\mathbb{S}) \in \mathbb{R}[\mathbb{F}_2^n]$ *whose* u*-th coordinate is*

$$\mathsf{CV}(\mathbb{S})[u] = \frac{1}{|\mathbb{S}|} \sum_{x \in \mathbb{S}} (-1)^{u^\top x}.$$

Consider a function $\mathsf{F} : \mathbb{F}_2^n \to \mathbb{F}_2^n$. Let the input multiset be $\mathbb{P}$ and the output multiset be $\mathbb{C}$, denoted by $\mathbb{C} = \mathsf{F}(\mathbb{P})$, and C^{F} be the correlation matrix of F.

Lemma 1. *The correlation vectors of* $\mathbb{P}$ *and* $\mathbb{C}$ *satisfy*

$$\mathsf{CV}(\mathbb{C}) = \mathsf{C}^{\mathsf{F}} \cdot \mathsf{CV}(\mathbb{P}).$$

Remark. Lemma 1 is not completely new. In [17], Daemen et al. introduced the correlation matrix to transform correlations of boolean functions. In [3], the correlation matrix was used by Beyne to transform two probability mass functions. All three forms including Lemma 1 are essentially the same. It is crucial to note that Lemma 1 applies to any multiset and function, including the function's d-wise forms. This generality is vital to our new approach, yet it was not stressed in [3,17].

Let the input difference and output mask of a DL attack be Δ and λ, respectively. The input multiset in the DL circumstance is $\mathbb{P} = \mathbb{F}_2^n \times \{\Delta\}$, and $\mathbb{C} = \mathsf{F}^{\times 2}(\mathbb{P})$. It is easy to check that the DL correlation of F is the $(0, \lambda)$-coordinate of $\mathsf{CV}(\mathbb{C})$, i.e.,

$$\mathsf{DL}[\Delta \xrightarrow{\mathsf{F}} \lambda] = \mathsf{CV}(\mathbb{C})[(0, \lambda)] = 2^{-n} \sum_{x \in \mathbb{F}_2^n, \theta = \Delta} (-1)^{\lambda^\top \mathsf{D}_\theta \mathsf{F}(x)}$$

In terms of modern ciphers, usually F as well as $\mathsf{F}^{\times 2}$ is composite, i.e.,

$$\mathsf{F}^{\times 2} = \mathsf{F}_{r-1}^{\times 2} \circ \cdots \circ \mathsf{F}_1^{\times 2} \circ \mathsf{F}_0^{\times 2}.$$

Thus, according to Theorem 1(1),

$$\mathsf{C}^{\mathsf{F}^{\times 2}} = \mathsf{C}^{\mathsf{F}_{r-1}^{\times 2}} \cdots \mathsf{C}^{\mathsf{F}_1^{\times 2}} \cdot \mathsf{C}^{\mathsf{F}_0^{\times 2}}.$$

Denoting $\sigma_0 = \mathsf{CV}(\mathbb{P})$ and $\sigma_i = \mathsf{C}^{\mathsf{F}_{i-1}^{\times 2}} \cdots \mathsf{C}^{\mathsf{F}_0^{\times 2}} \mathsf{CV}(\mathbb{P})$ for $1 \le i \le r$, we introduce a method to approximate σ_i from σ_{i-1}, with which we can ultimately approximate $\mathsf{CV}(\mathbb{C})$.

3.2 DL Correlation Approximation

We, by default, assume that our target is a typical SPN cipher with block size n. The round approximation is naturally applicable to other cipher structures, but left as future work.

A typical round function of SPN ciphers F_i consists of an S-box layer $\mathsf{S} : \mathbb{F}_2^{m\times t} \to \mathbb{F}_2^{m\times t} = \mathsf{s}||\cdots||\mathsf{s}$ where $\mathsf{s} : \mathbb{F}_2^m \to \mathbb{F}_2^m$ and a linear operation $\mathsf{L} : \mathbb{F}_2^n \to \mathbb{F}_2^n$, denoted by $\mathsf{F}_i = \mathsf{L} \circ \mathsf{S}$. The 2-wise form of F is

$$\mathsf{F}_i^{\times 2} = \mathsf{L}^{\times 2} \circ \mathsf{S}^{\times 2} = \mathsf{L}^{\times 2} \circ (\mathsf{s}^{\times 2}||\cdots||\mathsf{s}^{\times 2}).$$

In the following, we let $\gamma_{i+1} = \mathsf{C}^{\mathsf{S}^{\times 2}}\sigma_i$ for $i \geq 0$ and $\sigma_i = \mathsf{C}^{\mathsf{L}^{\times 2}}\gamma_i$ for $i \geq 1$, i.e.,

$$\mathsf{CV}(\mathbb{P}) = \sigma_0 \xrightarrow{\mathsf{C}^{\mathbf{s}^{\times 2}}} \gamma_1 \xrightarrow{\mathsf{C}^{\mathsf{L}^{\times 2}}} \sigma_1 \to \cdots \to \sigma_{r-1} \xrightarrow{\mathsf{C}^{\mathbf{s}^{\times 2}}} \gamma_r = \mathsf{CV}(\mathbb{C}).$$

$$\mathsf{CV}(\mathbf{X}) = \sigma_0 \xrightarrow{\mathsf{C}^{\mathbf{F}_1}} \sigma_1 \xrightarrow{\mathsf{C}^{\mathbf{F}_2}} \cdots \xrightarrow{\mathsf{C}^{\mathbf{F}_r}} \sigma_r = \mathsf{CV}(\mathbf{C})$$

We describe how to compute/approximate γ_i and σ_i round by round until $\mathsf{CV}(\mathbb{C})$.

Precise Computation for $\mathbf{S}^{\times 2}$. Since $\mathbb{P} = \mathbb{F}_2^n \times \{\Delta\}$, $\mathsf{CV}(\mathbb{P})$ can be represented by a Kronecker product of t component vectors,

$$\sigma_0 = \mathsf{CV}(\mathbb{P}) = \bigotimes_{i=0}^{t-1} \mathsf{CV}(\mathbb{F}_2^m \times \{\varpi_i(\Delta)\}) = \bigotimes_{i=0}^{t-1} \sigma_0^{[i]}.$$

According to Theorem 1(2), $\mathsf{C}^{\mathsf{S}^{\times 2}} = \bigotimes_{i=0}^{t-1} \mathsf{C}^{\mathsf{s}^{\times 2}}$, thus

$$\gamma_1 = \bigotimes_{i=0}^{t-1} \mathsf{C}^{\mathsf{s}^{\times 2}} \cdot \bigotimes_{i=0}^{t-1} \sigma_0^{[i]} = \bigotimes_{i=0}^{t-1} \mathsf{C}^{\mathsf{s}^{\times 2}} \cdot \sigma_0^{[i]}. \tag{4}$$

Note that all coordinates of γ_1 are precisely obtained in this step.

A straightforward implementation for computing γ_1 with Eq. (4) requires $\mathcal{O}(t \times 2^{4m})$ computations. However, the coordinates of $\mathsf{C}^{\mathsf{s}^{\times 2}}$ have the following property,

Proposition 1. $\mathcal{C}^{\mathsf{s}^{\times 2}}[(v_0, v_1), (u_0, u_1)] = \mathcal{C}^{\mathsf{s}}[v_0 + v_1, u_0 + u_1]\mathcal{C}^{\mathsf{s}}[v_1, u_1]$.

A proof can be found in the full version. For $0 \leq i < t$,

$$\begin{aligned}\gamma_1^{[i]}[(v_0, v_1)] &= \sum_{u_0,u_1} \mathsf{C}^{\mathsf{s}^{\times 2}}[(v_0, v_1), (u_0, u_1)] \cdot \sigma_0^{[i]}[(u_0, u_1)] \\ &= \sum_{u_1} \mathsf{C}^{\mathsf{s}}[v_1, u_1] \sum_{u_0} \mathsf{C}^{\mathsf{s}}[v_0 + v_1, u_0 + u_1] \cdot \sigma_0^{[i]}[(u_0, u_1)].\end{aligned}$$

We can use a vector with $(v_0 + v_1, u_1)$ as the index, denoted by ξ, to store the result of $\sum_{u_0} \mathsf{C}^{\mathsf{s}}[v_0 + v_1, u_0 + u_1] \cdot \sigma_0^{[i]}[(u_0, u_1)]$ for all possibilities of $v_0 + v_1$ and u_1 at the first step. In the second step, we use a vector ξ' with $(v_0 + v_1, v_1)$ as the

index to store $\sum_{u_1} \mathsf{C}^{\mathsf{s}}[v_1, u_1]\xi[v_0 + v_1, u_1]$. Finally, $\gamma_1^{[i]}[(v_0, v_1)] = \xi'[(v_0 + v_1, v_1)]$. The time complexity is reduced to $\mathcal{O}(t \times 2^{3m})$.

Heuristic Approximation for $\mathsf{L}^{\times 2}$. Now we compute $\sigma_1 = \mathsf{C}^{\mathsf{L}^{\times 2}} \cdot \gamma_1$, where

$$\sigma_1 = \mathsf{C}^{\mathsf{L}^{\times 2}} \cdot \left(\bigotimes_{i=0}^{t-1} \gamma_1^{[i]} \right)$$

Unfortunately, $\mathsf{C}^{\mathsf{L}^{\times 2}}$ cannot be written as a Kronecker product of smaller matrices with the same alignment of the S-boxes, so we have to handle $\mathsf{C}^{\mathsf{L}^{\times 2}} \in \mathbb{R}^{2^{2n} \times 2^{2n}}$ as a whole. Actually, we have no precise means to compute σ_1 due to the huge time complexity, so we have to introduce a heuristic approximation.

From the correlation matrix property [5,17], we know that $\sigma_1[v] = \gamma_1[(\mathsf{L}^{\times 2})^\top v]$[1]. To approximate the value of $\gamma_1[(\mathsf{L}^{\times 2})^\top v]$, we need the following assumption.

Assumption 1. For any $v \in \mathbb{F}_2^{2m \times t}$ and $\varphi \in \mathbb{R}[\mathbb{F}_2^{2m \times t}]$

$$\varphi[v] \approx \prod_{i=0}^{t-1} \varphi[\tau_i(\varpi_i(v))]$$

is a good approximation. Equivalently, there exist t vectors $\varphi^{[0]}, \varphi^{[1]}, \ldots, \varphi^{[t-1]} \in \mathbb{R}[\mathbb{F}_2^{2m}]$ that satisfy

$$\varphi \approx \bigotimes_{i=0}^{t-1} \varphi^{[i]}.$$

Remark. Assumption 1 holds if we assume that the S-boxes in a round are all independent. This independence is itself a consequence of the popular Markov assumption in cryptanalysis. Thus, the introduction of Assumption 1 is natural. In fact, Assumption 1 is weaker than the Markov assumption, as it is also a consequence of the Markov assumption.

With Assumption 1, for each index v of σ_1 we have

$$\sigma_1[v] = \gamma_1[(\mathsf{L}^{\times 2})^\top v] \approx \prod_{i=0}^{t-1} \gamma_1 \left[\tau_i(\varpi_i((\mathsf{L}^{\times 2})^\top v))\right] = \prod_{i=0}^{t-1} \gamma_1^{[i]} \left[\varpi_i((\mathsf{L}^{\times 2})^\top v)\right].$$

Since $\gamma_1^{[i]} \left[\varpi_i((\mathsf{L}^{\times 2})^\top v))\right]$ has been obtained from the S-box layer, $\sigma_1[v]$ can be obtained too. However, the problem is that it is impossible to compute all indices $v \in \mathbb{F}_2^{2n}$. Our method is to compute coordinates for all $\sigma_1^{[i]}$ as follows,

$$\mathbb{I} = \left\{\sigma_1[\tau_i(w)] : w \in \mathbb{F}_2^{2^m}, 0 \le i < t\right\}.$$

[1] This property means that the input mask equals the transpose of the matrix multiplied by the output mask.

Algorithm 1. Approximate the DL bias $\mathsf{DL}[\Delta \xrightarrow{\mathsf{F}} \lambda]$ for $\mathsf{F} = \mathsf{F}_{r-1} \cdots \mathsf{F}_0$, $\mathsf{F}_i = \mathsf{K} \circ \mathsf{L} \circ \mathsf{S}$, where $\mathsf{S} = \mathsf{s}||\cdots||\mathsf{s}$ is an S-box layer with t m-bit S-boxes.

procedure ApproxDL(Δ, λ)
  for j from 0 to $t-1$ **do**
    $\sigma_0^{[j]} = \mathsf{CV}(\mathbb{F}_2^m \otimes \{\varpi_j(\Delta)\})$
  for i from 1 to r **do**
    for j from 0 to $t-1$ **do**
      $\gamma_i^{[j]} = \mathsf{C}^{\mathsf{s}^{\times 2}} \cdot \sigma_{i-1}^{[j]}$ ▷ For $\mathsf{S}^{\times 2} = \mathsf{s}^{\times 2}||\cdots||\mathsf{s}^{\times 2}$
    for j from 0 to $t-1$ **do**
      for $w \in \mathbb{F}_2^m \times \mathbb{F}_2^m$ **do**
        $\sigma_i^{[j]}[w] = \prod_{k=0}^{t-1} \gamma_i^{[k]}[\varpi_k((\mathsf{L}^{\times 2})^\top(\tau_j(w)))]$ ▷ For the linear layer $\mathsf{L}^{\times 2}$
    for j from 0 to $t-1$ **do**
      for $(v_0, v_1) \in \mathbb{F}_2^m \times \mathbb{F}_2^m$ **do**
        $\sigma_i^{[j]}[(v_0, v_1)] = (-1)^{v_0^\top k} \sigma_i^{[j]}[(v_0, v_1)]$ ▷ Fixed-key/constant XOR/
    for j from 0 to $t-1$ **do**
      for $(v_0, v_1) \in \mathbb{F}_2^m \times \mathbb{F}_2^m$ **do**
        $\sigma_i^{[j]}[(v_0, v_1)] = \begin{cases} 0, & \text{if } v_0 \neq 0 \\ \sigma_i^{[j]}[(v_0, v_1)], & \text{if } v_0 = 0 \end{cases}$ ▷ Average-key key-XOR
  $\sigma_r[(0, \lambda)] = \prod_{j=0}^{t-1} \sigma_i^{[j]}[(0, \varpi_j(\lambda))]$
  return $\sigma_r[(0, \lambda)]$

and use these coordinates to approximate any $\sigma_1[v]$ based on Assumption 1. The reason for selecting these coordinates is that they allow us to proceed with the calculation of the S-boxes in the next step. Concretely,

$$\sigma_1 = \bigotimes_{i=0}^{t-1} \sigma_1^{[i]} \quad \text{or} \quad \sigma_1[v] = \prod_{i=0}^{t-1} \sigma_1^{[i]}[\varpi_i(v)].$$

Since $\mathsf{L}^{\times 2} : (x, \theta) \mapsto (\mathsf{L}(x), \mathsf{L}(\theta))$, $\mathsf{L}^{\times 2} = \mathsf{L} \otimes \mathsf{L}$. Let $v = (v_0, v_1)$, $(\mathsf{L}^{\times 2})^\top v = (\mathsf{L}^\top v_0, \mathsf{L}^\top v_1)$. Thus, computing $(\mathsf{L}^{\times 2})^\top v$ is easy. Consequently, the complexity of this step is to compute all coordinates in $\mathbb{I}$, which requires $\mathcal{O}(t \times 2^{2m+1})$ computations.

The Whole Algorithm. The approach introduced above enables us to iteratively compute σ_i and γ_i until $\mathsf{CV}(\mathbb{C})$. The whole approximation process is given in Algorithm 1. Assumption 1 holds for γ_1, so the computation of $\sigma_1^{[i]}$ as well as $\gamma_2^{[i]}$ is precise. However, since

$$\sigma_2^{[i]}[w] = \sigma_2[\tau_i(w)] = \gamma_1[(\mathsf{L}^{\times 2})^\top \tau_i(w)] \approx \prod_{j=0}^{t-1} \gamma_1^{[j]} \left[\varpi_j((\mathsf{L}^{\times 2})^\top \tau_i(w))\right],$$

the computation after σ_2 (inclusive) is approximate.

3.3 Influence of the Key-XOR

For block ciphers, between two consecutive round functions there is usually a key-XOR operation denoted by K. In the fixed-key setting, the XORed key k is a constant, thus the 2-wise form of K can be written as

$$\mathsf{K}_k^{\times 2} : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n; (x, \theta) \to (x + k, \theta)$$

The correlation matrix of $\mathsf{K}_k^{\times 2}$ has the element

$$\begin{aligned}
\mathsf{C}^{\mathsf{K}_k^{\times 2}}[(v_0, v_1)(u_0, u_1)] &= 2^{-2n} \sum_{x \in \mathbb{F}_2^n, \theta \in \mathbb{F}_2^n} (-1)^{u_0^\top x + u_1^\top \theta + v_0^\top (x+k) + v_1^\top \theta} \\
&= (-1)^{v_0^\top k} \left(2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{(u_0^\top + v_0^\top)x} \right) \left(2^{-n} \sum_{\theta \in \mathbb{F}_2^n} (-1)^{(u_1^\top + v_1^\top)\theta} \right) \\
&= (-1)^{v_0^\top k} \delta(u_0 + v_0) \delta(u_1 + v_1).
\end{aligned}$$

Let $\varphi = \mathsf{C}^{\mathsf{K}_k^{\times 2}} \sigma$, then

$$\varphi[(v_0, v_1)] = \sum_{u_0 \in \mathbb{F}_2^n, u_1 \in \mathbb{F}_2^n} (-1)^{v_0^\top k} \delta(v_0 + u_0) \delta(v_1 + u_1) \sigma[(u_0, u_1)] = (-1)^{v_0^\top k} \sigma[(v_0, v_1)].$$

This is to say, the values of key would solely cause a possible sign flip when copying the coordinates. Although this paper does not consider the fixed-key setting for block ciphers, this calculation is useful when handling the constant-XOR operation in cryptographic permutations like `Subterranean`-2.0 and `Koala`-p, like what we will present in Sect. 6.

In the average-key setting, the key is regarded as a uniform variable.

$$\mathsf{K}^{\times 2} : \mathbb{F}_2^n \times \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n \times \mathbb{F}_2^n; (x, \theta, k) \to (x + k, \theta).$$

The correlation matrix has the elements as

$$\begin{aligned}
\mathsf{C}^{\mathsf{K}^{\times 2}}[(v_0, v_1), (u_0, u_1, u_k)] &= 2^{-3n} \sum_{x \in \mathbb{F}_2^n, \theta \in \mathbb{F}_2^n, k \in \mathbb{F}_2^n} (-1)^{u_0^\top x + u_1^\top \theta + u_k^\top k + v_0^\top (x+k) + v_1^\top \theta} \\
&= \delta(v_0 + u_k) \delta(v_0 + u_0) \delta(v_1 + u_1).
\end{aligned}$$

Let $\varphi = \mathsf{C}^{\mathsf{K}^{\times 2}} \sigma$. Note that the round key is uniform and independent, thus $\sigma = \sigma' \otimes \sigma''$ where σ' and σ'' are the correlation vectors of the data path and the round key, respectively. Therefore,

$$\varphi[(v_0, v_1)] = \sigma'[(v_0, v_1)] \cdot \sigma''[v_0].$$

Under the assumption that k is uniform,

$$\sigma''[v_0] = 2^{-n} \sum_{k \in \mathbb{F}_2^n} (-1)^{v_0^\top k} = \begin{cases} 1, & v_0 = 0, \\ 0, & v_0 \neq 0. \end{cases}$$

Consequently,

$$\varphi[(v_0, v_1)] = \begin{cases} 0, & v_0 \neq 0, \\ \sigma'[(v_0, v_1)], & v_0 = 0. \end{cases}$$

That is to say, when considering the DL attack, we do not need to consider the influence of values *under the average-key setting* for block ciphers.

We call v_0 the value mask, it is equivalent to say that we can safely set $v_0 = 0$ in our approximation.

3.4 Optimization for Efficiency and Precision

We have demonstrated that for block ciphers, the value mask (v_0) can be consistently set to zero under the assumption of uniformly random and independent round keys. For permutations, however, no analogous assertion exists. Nevertheless, here we present an observation-based assumption: During the approximation process *in our applications*, the non-zero value mask exhibits a minimal effect on the approximation results.

In fact, most prior works on DL correlation detection rely on the Markov assumption, which is equivalent to setting the value mask to zero. As far as we know, only DATF and HATF incorporate the impact of values. However, upon inspecting their source code, we observed that the DL correlation remains unchanged even when value-related correlations in their implementations are forcibly set to zero. Furthermore, in our experiments, we observe that the correlations of the values are always zero (i.e., if the value mask is non-zero, the difference correlation is always zero). We therefore conclude this assumption is empirically justified. For clarity, we formally state this assumption:

Assumption 2 (Zero value-mask assumption). At each step of our approximation, setting the value mask to zero exhibits no discernible impact on the approximation results.

Under Assumption 2, for both block ciphers and permutations, we can approximate the DL correlation by setting the value mask to zero. Consequently, for the correlation vector of each intermediate multiset, Assumption 1 is easier to be true. If the intermediate multiset of the i-th round is $\mathbb{X}_i = \mathbb{S} \times \{\Delta_i\}$, we can always obtain

$$\mathsf{CV}(\mathbb{X}_i)[0, v] = \prod_{i=0}^{t-1} \mathsf{CV}(\mathbb{S} \times \{\Delta_i\})^{[i]}[\varpi_i(0, v)].$$

Hence, we can extend the input difference Δ over the first several rounds, to delay the failure timing of Assumption 1.

In practice, we split the r-round F into two parts: the first r_1-round part and the second $(r - r_1)$-round part, denoted by $\mathsf{F} = \mathsf{F}_2 \circ \mathsf{F}_1$. The DL correlation of F is then calculated by

$$\mathsf{DL}[\Delta \xrightarrow{F} \lambda] = \sum_{\Delta'} \mathsf{Prob}[\Delta \xrightarrow{\mathsf{F}_1} \Delta'] \cdot \mathsf{DL}[\Delta' \xrightarrow{\mathsf{F}_2} \lambda] \tag{5}$$

For each Δ', $\mathsf{Prob}[\Delta \xrightarrow{\mathsf{F}_1} \Delta']$ is the differential probability from Δ to Δ' over F_1. For F_2, we use our basic method to approximate the correlation of $\mathsf{DL}[\Delta' \xrightarrow{\mathsf{F}_2} \lambda]$.

4 Applications to S-Box Based Primitives Ascon and Present

In this section, we apply the round-based approximation approach to `Ascon` [20] and `Present` [15]. The DL properties of both primitives have been extensively checked in previous papers [22,31]. With our method, for `Ascon` we have verified the previous results, with improvements in several DL correlation estimations. Considering that recently, NIST standardized `Ascon` based on `Ascon`-128a, we also provide DL distinguishers for `Ascon`-128a. Notably, we find a 6-round DL distinguisher for `Ascon`-128a with a correlation of $2^{-23.89}$. For `Present`, we detect 17-round effective DL correlations, which is 4 rounds longer than previous best results in [22].

4.1 More Precise DL Correlations of Ascon

`Ascon` Initialization. In [19], the designers of `Ascon` detected by experiments a 4-round DL correlation with 2^{-1} correlation and a 5-round one with 2^{-9} correlation[2]. Later, several theoretical approaches [23,27,31] have been developed to give approximations to the two experimental values, as shown in Table 1.

For all cases, our results are obtained by optimization for the first 2 rounds (Eq. (5)). For `Ascon`-p, the probability for the differential part is computed in a normal way where we assume that all input bits are uniform. However, for `Ascon` initialization, the differential probability for the first round is computed after considering the initial value (IV) and the actual values.

Details of all DL distinguishers we obtained for `Ascon` are provided in Table 2 (Appendix C). Next, we introduce them one by one.

DL Distinguisher I. For the simpler case of the 4-round `Ascon` initialization where the correlation is 1, most previous methods can provide meaningful estimations for this value. Until very recently, at CRYPTO 2024 [31], the TDT technique was the first one that could fully match the experimental values. Our approximation also matches it, giving the correlation 1.

DL Distinguisher II. The 5-round `Ascon` initialization is more complicated. Dobraunig et al. reported a 2^{-9} correlation for this distinguisher. However, our experiments show that this correlation is not stable until we use 2^{34} samples. Our experimental correlation is $2^{-8.94}$. Before the TDT technique [31], all techniques cannot precisely approximate this DL correlation. The TDT technique estimated it as $2^{-9.1}$, while our approach approximates it as $2^{-8.94}$, showing advantages in precision.

[2] With 2^{34} samples, our experiment shows that the correlation should be $2^{-8.94}$.

In [19], Dobraunig et al. further reported that if they force the two nonce bits corresponding to the difference to be equal, the correlation would be improved to 2^{-8}. With 2^{34} samples, we checked that this correlation is $2^{-7.94}$. Our round-based approach well matches these experimental values, which is the only public theoretical result for this DL correlation.

DL Distinguisher V. Previous papers usually focused solely on Ascon-128, where only the first 64 bits can be observed after the initialization phase. Recently, NIST finished the standardization of Ascon, where the Ascon-128a is taken. Different from Ascon-128, the first 128 bits of Ascon-128a can be observed, giving the adversary more flexibility. We apply our techniques to Ascon-128a, and detect a 6-round DL correlation with $2^{-23.89}$.

Ascon-p. In [22], Hadipour, Derbez and Eichlseder applied their general DLCT detection technique to Ascon-p. Different from the Ascon initialization case, the input difference can be selected arbitrarily. For 4-round Ascon-p, DL distinguishers with correlation 1 are found. Our approach can replay these results. For 5-round Ascon-p, we obtained more precise approximations compared to [22], as shown in Distinguishers III and IV in Table 2 (Appendix C).

DL Distinguisher VI. We also checked the DL correlations for 6-round Ascon-p. The input values are assumed to be uniform values. With the same input difference as Distinguisher IV, we checked all output masks for single S-boxes. The best correlation we detected is $2^{-21.89}$, slightly better than the Ascon-128a case.

4.2 Longer DL Distinguishers of Present

In [22], Hadipour, Derbez and Eichlseder gave the best DL distinguishers for Present up to 13 rounds. By our round-based approximation, we verified their results up to 13 rounds, which confirms the correctness of our approach and codes. For up to 10 rounds, we did the experiments and found that our results are closer to the experimental values. For 11, 12 and 13 rounds, the correlation is too small to be verified with experiments. Our results are slightly larger than [22], which presents better DL distinguishers for Present.

More importantly, our tool allows us to detect DL correlations up to 17 rounds, which are 4 rounds longer than [22]. These distinguishers are found by examining all single S-box cases for the input differences and output masks. All these distinguishers are provided in Table 4 (Appendix C).

5 Approximate Correlation Propagation for χ and $\chi^{\times 2}$

The n-bit χ (n is an odd number larger than 1) function is defined as

$$\chi_n : \mathbb{F}_2^n \to \mathbb{F}_2^n; \quad (x_0, x_1, \ldots, x_{n-1}) \mapsto (y_0, y_1, \ldots, y_{n-1}),$$

where $y_i = x_i + (x_{i+1} + 1)x_{i+2}$, and the indices are modulo n.

To apply our round-based approximation approach to ciphers with the χ function, we need to figure out the correlation propagation property of the χ function. When n is small, we can compute all coordinates of the input and output correlation vectors. When n is very large, however, it is infeasible. Thus, we have to use Assumption 1, i.e., we will only compute out a fraction of coordinates, and use them to approximate any coordinate.

We decompose the input and output of χ into bits, and assume their correlation vectors are the Kronecker product of small correlation vectors corresponding to 1 bit. In other words, each bit is treated like one S-box.

Let $\gamma = \mathsf{C}^{\chi^{\times 2}}\sigma$. With Assumption 1, we will assume that

$$\gamma = \bigotimes_{j=0}^{n-1} \gamma^{[j]} \qquad \text{and} \qquad \sigma = \bigotimes_{j=0}^{n-1} \sigma^{[j]} \tag{6}$$

Similar to the discussions on the S-box layer, here we want to compute the following coordinates of γ

$$\mathbb{I} = \left\{\gamma[\tau_i(w)] : w \in \mathbb{F}_2^2, i = 0, \ldots, n-1\right\}$$

from the coordinates of σ in

$$\mathbb{J} = \left\{\sigma[\tau_i(w)] : w \in \mathbb{F}_2^2, i = 0, \ldots, n-1\right\}.$$

Note that the $\tau_i(\cdot)$ (as well as the later $\varpi_i(\cdot)$) operation in this section will take each bit as an S-box.

Unfortunately, χ_n cannot be decomposed into n small parallel functions, thus we cannot handle the χ_n function like the S-box layer directly. Inspired by a recursive matrix multiplication method for the modular addition [30], we introduce a new recursive matrix multiplication method to handle the χ_n function, based on which we can handle the problem of approximating γ's coordinates in $\mathbb{I}$ from $\mathbb{J}$.

5.1 Correlation Computation for χ and $\chi^{\times 2}$

Denote the correlation matrix of χ_n by C^{χ_n}. We first give an approach for computing $\mathsf{C}^{\chi_n}[v, u]$ based on a supportive function g_n. There is a linear term in χ_n's output bit. Omitting this linear term will bring convenience in our computation. Denote by 1^n the vector whose all coordinates are 1. Let $g_n(x) = (x + 1^n) \bigwedge (x \lll 1)$. We have

$$\chi_n(x) = x + (g_n(x) \lll 1).$$

There is a direct connection between the coordinates between C^{χ_n} and C^{g_n}.

Lemma 2. *The coordinates of C^{χ_n} and C^{g_n} have the following relationship,*

$$\mathsf{C}^{\chi_n}[v, u] = \mathsf{C}^{g_n}[v \ggg 1, v + u].$$

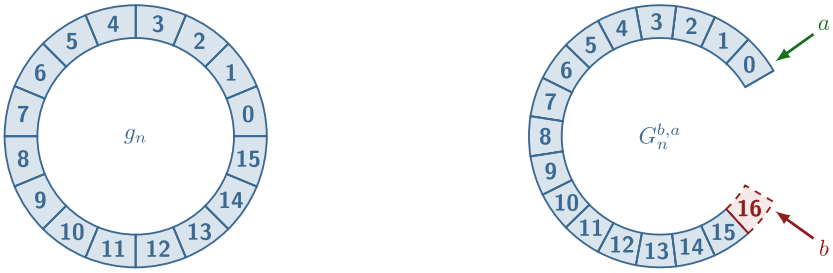

Fig. 1. Breaking th circle of g_n for $n = 16$, we can address the troublesome circle structure by defining $G_n^{b,a}$.

Proof. Let $\mathsf{Cor}(f)$ represent the correlation of f.

$$\begin{aligned}\mathsf{C}^{\chi_n}[v,u] &= \mathsf{Cor}[u^\top x + v^\top \chi_n(x)] = \mathsf{Cor}[u^\top x + v^\top (x + (g(x) \lll 1))] \\ &= \mathsf{Cor}[u^\top x + v^\top x + (v \ggg 1)^\top g_n(x)] = \mathsf{Cor}[(u^\top + v^\top)x + (v \ggg 1)^\top g_n(x)] \\ &= \mathsf{C}^{g_n}[v \ggg 1, u + v].\end{aligned}$$

□

As a result, if we can compute the coordinates of C^{g_n}, we can compute the coordinate of C^{χ_n}. Handling g_n is easier than handling χ_n.

When dealing with the g_n function, the most troublesome aspect is that each bit is multiplied by the subsequent one, and they are cyclically linked. To address this, we first sever this connection by fixing the first bit and the bit following the last bit of a g_n function as constants. We define a circle-free version of g_n.

Definition 3. *We define*

$$G_n^{b,a} : \mathbb{F}_2^n \to \mathbb{F}_2^n; \quad y = G_n(x)$$

where $y[i] = (x[i] + 1)x[i + 1], 0 \le i < n$ *and* $x[0] = a, x[n] = b$.

Note that the indices are not modulo n for $G_n^{b,a}$. *An illustration of* $G_n^{b,a}$ *is provided in Fig. 1. The reason we regard the input length of* $G_n^{b,a}$ *as n rather than* $n + 1$ *or* $n - 1$ *is that we want it to have the same input as* g_n *in form.*

Proposition 2. *Let* $C^{G_n^{b,a}}$ *be the correlation matrix of* $G_n^{b,a}$.

$$C^{g_n}[v,u] = C^{G_n^{0,0}}[v,u] + C^{G_n^{1,1}}[v,u].$$

Proof. For a pair of masks (v, u),

$$\begin{aligned}\mathsf{C}^{g_n}[v,u] &= 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{u^\top x + v^\top g_n(x)} \\ &= \sum_{c \in \mathbb{F}_2} \frac{1}{2^n} \sum_{\substack{y[0]=y[n]=c \\ (y[1],\ldots,y[n-1]) \in \mathbb{F}_2^{n-1}}} (-1)^{\sum_{i=0}^{n-1} u[i]y[i] + \sum_{i=0}^{n-1} v[i](y[i]+1)y[i+1]} \\ &= \mathsf{C}^{G_n^{0,0}}[v,u] + \mathsf{C}^{G_n^{1,1}}[v,u].\end{aligned}$$

□

$\mathsf{C}^{G_n^{b,a}}$ has a chained property.

Proposition 3. *Let $n > 1$, $v'' = v||v', u'' = u||u' \in \mathbb{F}_2^{n+1}$ where $u', v' \in \mathbb{F}_2$. From $\mathcal{C}^{G_n^{c,a}}[v,u]$ we can get $\mathcal{C}^{G_{n+1}^{b,a}}[v'',u'']$ from the following formula,*

$$\mathcal{C}^{G_{n+1}^{b,a}}[v'',u''] = \frac{1}{2}\sum_{c\in\mathbb{F}_2}(-1)^{u'c+v'\bar{c}b}\mathcal{C}^{G_n^{c,a}}[v,u].$$

Proof. According to Definition 3, when $n=1$, $y = G_1^{b,a}(x) = (a+1)b$. Thus,

$$\mathsf{C}^{G_n^{b,a}}[v,u] = \frac{1}{2}(-1)^{ua+v(a+1)b}.$$

For $n > 1$,

$$\begin{aligned}
\mathsf{C}^{G_{n+1}^{b,a}}[v'',u''] &= 2^{-(n+1)}\sum_{\substack{x[0]=a,x[n+1]=b\\(x[1],\ldots,x[n])\in\mathbb{F}_2^n}}(-1)^{\sum_{i=0}^{n}u''[i]x[i]+\sum_{i=0}^{n}v''[i](x[i]+1)x[i+1]}\\
&= \frac{1}{2}\sum_{c\in\mathbb{F}_2}2^{-n}\sum_{\substack{x[0]=a,x[n]=c,x[n+1]=b\\(x[1],\ldots,x[n-1])\in\mathbb{F}_2^{n-1}}}(-1)^{\sum_{i=0}^{n-1}u[i]x[i]+\sum_{i=0}^{n-1}v[i](x[i]+1)x[i+1]}(-1)^{u'c+v'(c+1)b}\\
&= \frac{1}{2}\sum_{c\in\mathbb{F}_2}(-1)^{u'c+v'(c+1)b}\mathsf{C}^{G_n^{c,a}}[v,u].
\end{aligned}$$

□

According to different values of u' and v', we can write $\frac{1}{2}\sum_{c\in\mathbb{F}_2}(-1)^{u'c+v'(c+1)b}$ into a matrix with c and b being the column and row indices. Therefore, we have the following lemma.

Lemma 3. *For the four possible values of $v', u', a, b \in \mathbb{F}_2$, we define a matrix $M_{v',u'}$ whose (b,a)-coordinate is $M_{v',u'}[b,a] = \frac{1}{2}(-1)^{u'a+v'(a+1)b}$. According to the four possibilities of (v',u'), we obtain four matrices:*

$$M_{0,0} = \frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}, M_{0,1} = \frac{1}{2}\begin{bmatrix}1&-1\\1&-1\end{bmatrix}, M_{1,0} = \frac{1}{2}\begin{bmatrix}1&1\\-1&1\end{bmatrix}, M_{1,1} = \frac{1}{2}\begin{bmatrix}1&-1\\-1&-1\end{bmatrix}.$$

$\mathcal{C}^{G_n^{b,a}}[v,u]$ can be computed as

$$\mathcal{C}^{G_n^{b,a}}[v,u] = e_b^\top\left(\prod_{i=0}^{n-1}M_{v[i],u[i]}\right)e_a$$

where $e_0 = [1,0]^\top$ and $e_1 = [0,1]^\top$.

Proof. It is obviously true for $n=1$ from the definition of $M_{v',u'}$. For larger n, assume that the lemma is true for the $n-1$ case. Let $v,u \in \mathbb{F}_2^{n-1}$, according to Proposition 3,

$$\begin{aligned}
\mathsf{C}^{G_n^{b,a}}[v||v',u||v'] &= \frac{1}{2}\sum_{c\in\mathbb{F}_2}(-1)^{u'c+v'(c+1)b}\mathsf{C}^{G_{n-1}^{c,a}}[v,u] = \frac{1}{2}\sum_{c\in\mathbb{F}_2}e_b^\top M_{v',u'}e_c\mathsf{C}^{G_{n-1}^{c,a}}[v,u]\\
&= e_b^\top M_{v',u'}\sum_c e_c e_c^\top\left(\prod_{i=0}^{n-2}M_{v[i],u[i]}\right)e_a = e_b^\top\left(\prod_{i=0}^{n-1}M_{v[i],u[i]}\right)e_a.
\end{aligned}$$

Note that $\sum_c e_c e_c^\top$ is an identity matrix.

Lemma 4. *For the four possible values of $v', u', a, b \in \mathbb{F}_2$, we define a matrix $M_{v',u'}$ whose (b,a)-coordinate is $M_{v',u'}[b,a] = \frac{1}{2}(-1)^{u'a+v'(a+1)b}$. According to the four possibilities of (v',u'), we obtain four matrices:*

$$M_{0,0} = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}, M_{0,1} = \frac{1}{2}\begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix}, M_{1,0} = \frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}, M_{1,1} = \frac{1}{2}\begin{bmatrix}1 & -1\\ -1 & -1\end{bmatrix}.$$

$\mathsf{C}^{g_n}[v,u]$ can be computed as follows.

$$\mathsf{C}^{g_n}[v,u] = \begin{bmatrix}1 & 0\end{bmatrix}\prod_{i=0}^{n-1} M_{v[i],u[i]}\begin{bmatrix}1\\ 0\end{bmatrix} + \begin{bmatrix}0 & 1\end{bmatrix}\prod_{i=0}^{n-1} M_{v[i],u[i]}\begin{bmatrix}0\\ 1\end{bmatrix}.$$

Proof. This lemma is a natural corollary of Proposition 2 and Lemma 3.

Proposition 4. *Denote $e_0 = [1,0]^\top$ and $e_1 = [0,1]^\top$. Given $\lambda_0, \lambda_1, \xi_0, \xi_1 \in \mathbb{F}_2^n$, we have*

$$\begin{aligned}&\mathsf{C}^{g_n}[\lambda_0,\lambda_1]\cdot\mathsf{C}^{g_n}[\xi_0,\xi_1]\\ =&\sum_{i=0}^{1}\sum_{i'=0}^{1}(e_i\otimes e_{i'})^\top\prod_{j=0}^{n-1}\left(M_{\lambda_0[j],\lambda_1[j]}\otimes M_{\xi_0[j],\xi_1[j]}\right)(e_i\otimes e_{i'})\end{aligned}$$

Proof. Since $\mathsf{C}^{g_n}[\lambda_0,\lambda_1]$ and $\mathsf{C}^{g_n}[\xi_0,\xi_1]$ are two values, the multiplication of them is equivalent to the Kronecker product of them:

$$\mathsf{C}^{g_n}[\lambda_0,\lambda_1]\cdot\mathsf{C}^{g_n}[\xi_0,\xi_1] = \mathsf{C}^{g_n}[\lambda_0,\lambda_0]\otimes\mathsf{C}^{g_n}[\xi_0,\xi_1].$$

According to Lemma 4,

$$\begin{aligned}&\mathsf{C}^{g_n}[\lambda_0,\lambda_1]\otimes\mathsf{C}^{g_n}[\xi_0,\xi_1]\\ =&\sum_{i=0}^{1}e_i^\top\left(\prod_{j=0}^{n-1}M_{\lambda_0[j],\lambda_1[j]}\right)e_i\bigotimes\sum_{i'=0}^{1}e_{i'}^\top\left(\prod_{j=0}^{n-1}M_{\xi_0[j],\xi_1[j]}\right)e_{i'}\\ =&\sum_{i=0}^{1}\sum_{i'=0}^{1}(e_i\otimes e_{i'})^\top\prod_{j=0}^{n-1}\left(M_{\lambda_0[j],\lambda_1[j]}\otimes M_{\xi_0[j],\xi_1[j]}\right)(e_i\otimes e_{i'}).\end{aligned}$$

□

Theorem 2. *Denote $e_0 = [1,0]^\top$ and $e_1 = [0,1]^\top$. The coordinates of $C^{\chi^{\times 2}}$ can be computed as*

$$C^{\chi_n^{\times 2}}[(v_0,v_1),(u_0,u_1)] = \sum_{i=0}^{1}\sum_{i'=0}^{1}(e_i\otimes e_{i'})^\top\prod_{j=0}^{n-1}\left(M^{\otimes}_{v_0,v_1,u_0,u_1,j}\right)(e_i\otimes e_{i'}),$$

where $M^{\otimes}_{v_0,v_1,u_0,u_1,j} = M_{((v_0+v_1)\ggg 1)[j],(u_0+u_1+v_0+v_1)[j]}\otimes M_{(v_1\ggg 1)[j],(u_1+v_1)[j]}$.

Proof. According to Proposition 1 and Lemma 2, we have

$$\mathsf{C}^{\chi_n^{\times 2}}[(v_0,v_1),(u_0,u_1)] = \mathsf{C}^{g_n}[(v_0+v_1)\ggg 1, u_0+u_1+v_0+v_1]\mathsf{C}^{g_n}[v_1\ggg 1, u_1+v_1].$$

Replacing $\lambda_0,\lambda_1,\xi_0,\xi_1$ with $(v_0+v_1)\ggg 1, u_0+u_1+v_0+v_1, v_1\ggg 1, u_1+v_1$ respectively in Proposition 4, we finish the proof. □

5.2 Compute Coordinates in $\mathbb{J}$ from $\mathbb{I}$

Based on Assumption 1,

$$\sigma[(u_0,u_1)] \approx \prod_{j=0}^{n-1} \sigma[\tau_j(\varpi_j(u_0,u_1)].$$

The coordinate $\gamma[(v_0,v_1)]$ can be approximated by

$$\begin{aligned}
\gamma[(v_0,v_1)] &= \sum_{(u_0,u_1)\in\mathbb{F}_2^{2n}} \mathsf{C}_{\chi_2}^{\times 2}[(v_0,v_1),(u_0,u_1)]\cdot\sigma[(u_0,u_1)] \\
&= \sum_{(u_0,u_1)\in\mathbb{F}_2^{2n}} \sum_{i=0}^{1}\sum_{i'=0}^{1} (e_i\otimes e_{i'})^\top \prod_{j=0}^{n-1} M^{\otimes}_{v_0,v_1,u_0,u_1,j}\cdot(e_i\otimes e_{i'})\cdot\sigma[(u_0,u_1)] \\
&= \sum_{i=0}^{1}\sum_{i'=0}^{1} (e_i\otimes e_{i'})^\top \sum_{(u_0,u_1)\in\mathbb{F}_2^{2n}} \prod_{j=0}^{n-1} M^{\otimes}_{v_0,v_1,u_0,u_1,j}\sigma[(u_0,u_1)]\cdot(e_i\otimes e_{i'}) \\
&= \sum_{i=0}^{1}\sum_{i'=0}^{1} (e_i\otimes e_{i'})^\top \sum_{(u_0,u_1)\in\mathbb{F}_2^{2n}} \prod_{j=0}^{n-1} M^{\otimes}_{v_0,v_1,u_0,u_1,j} \prod_{j=0}^{n-1}\sigma[\tau_j(\varpi_j((u_0,u_1)))]\cdot(e_i\otimes e_{i'}) \\
&= \sum_{i=0}^{1}\sum_{i'=0}^{1} (e_i\otimes e_{i'})^\top \sum_{(u_0,u_1)\in\mathbb{F}_2^{2n}} \prod_{j=0}^{n-1} (\sigma[\tau_j(\varpi_j((u_0,u_1)))]M^{\otimes}_{v_0,v_1,u_0,u_1,j})\cdot(e_i\otimes e_{i'}).
\end{aligned}$$

Note that for any two different indices j and j', the two related coordinates $\sigma[\tau_j(\varpi_j((u_0,u_1)))]$ and $\sigma[\tau_{j'}(\varpi_{j'}((u_0,u_1)))]$ are indexed by two different unit vector tuples, thus independent of each other according to Eq. 6. Consequently, we can swap the "$\sum$" and "$\prod$" in the above formula. Finally, we obtain

$$\begin{aligned}
&\gamma[(v_0,v_1)] \\
&= \sum_{i=0}^{1}\sum_{i'=0}^{1} (e_i\otimes e_{i'})^\top \prod_{j=0}^{n-1} \sum_{u_0[j],u_1[j]\in\mathbb{F}_2} (\sigma[\tau_j(\varpi_j((u_0,u_1)))]M^{\otimes}_{v_0,v_1,u_0,u_1,j})\cdot(e_i\otimes e_{i'}).
\end{aligned}$$

The above formula becomes handy to calculate. Therefore, from the knowledge of the coordinates in $\mathbb{J}$, we can approximate any coordinates in $\mathbb{I}$ under Assumption 1. The complexity of computing each coordinate in $\mathbb{J}$ is a small constant; thus, the whole time complexity is $\mathcal{O}(n)$.

6 Applications to χ Based Ciphers

In this section, we apply the round-based approximation approach to Subterranean-2.0 [18] and Koala-p [1]. Both permutations take χ_{257} as their non-linear functions. Subterranean-2.0 is used in Subterranean-SAE, one of the second round candidates of the NIST LWC competition[3]. The designers gave

[3] https://csrc.nist.gov/Projects/Lightweight-Cryptography.

differential characteristics of Subterranean-2.0 for up to 4 rounds. For up to 8 rounds, differential probability upper bounds were provided as arguments for its security. These bounds had been improved in [29] with a smart tree search strategy. As the algebraic degree of Subterranean-2.0 round functions is only 2, there are trivial cube distinguishers with 2^{2^r+1} data complexity for r rounds. Thus, the trivial cube distinguishers can reach at most 7 rounds (8-round distinguisher will require 2^{257} data complexity). If we set a restrict that only disjoint input bits are selected as cube variables, there are cube distinguishers with $2^{2^{r-1}+1}$ data complexity for r rounds, but the data complexity in this case is at most 2^{128}. Therefore, the cube distinguisher can only reach 7 rounds, too. Koala-p is adapted from Subterranean-2.0 by changing the order of round function components and parameters. Differential and linear characteristics and bounds were provided in their design document up to 8 rounds. There are also analyses of Subterranean-2.0 used in mode [26,32], but not relevant to this work.

With our round-based approximation approach, for Subterranean-2.0 and Koala-p, we find 6-round DL distinguishers, which are 2 rounds more than the previous best differential and linear distinguishers in [1,18]. All results in this section consider the effects of values, i.e., we work without considering Assumption 2.

Description of Subterranean-2.0 and Koala-p. Subterranean-2.0 and Koala-p are both $\mathbb{F}_2^{257} \to \mathbb{F}_2^{257}$ permutations with 8 rounds. Their round functions ρ are defined in Fig. 2. For convenience, in this section, we use the support of a vector $x \in \mathbb{F}_2^n$ to represent x. For example, if the support of x is $\{i_1, i_2, \cdots i_n\}$, we write it as $x := [i_1, i_2, \cdots i_n]$.

Subterranean-2.0

$\rho_j = \pi \circ \theta \circ \iota_j \circ \chi$

$\chi : s_i \leftarrow s_i + (s_{i+1} + 1) \wedge s_{i+2}$

$$\iota_j : s_i \leftarrow \begin{cases} s_i + 1 & i = 0, \\ s_i & \text{otherwise} \end{cases}$$

$\theta : s_i \leftarrow s_i + s_{i+3} + s_{i+8}$,

$\pi : s_i \leftarrow s_{12i}$.

Koala-p

$\rho_j = \chi \circ \iota_j \circ \theta \circ \pi$,

$\pi : s_i \leftarrow s_{121i}$,

$\theta : s_i \leftarrow s_i + s_{i+3} + s_{i+10}$,

$$\iota_j : s_i \leftarrow \begin{cases} s_i + 1 & i = 0 \text{ and } j \notin \{2, 5, 6\} \\ s_i & \text{others} \end{cases}$$

$\chi : s_i \leftarrow s_i + (s_{i+1} + 1) \wedge s_{i+2}$.

Fig. 2. Round functions of Subterranean-2.0 and Koala-p.

6.1 Applications to Subterranean-2.0

5-round DL Distinguisher. The best DL distinguisher for 5-round Subterranean-2.0 detected by our method is as follows,

$$[0] \xrightarrow[-2^{-7.98}]{\chi \circ \rho_3 \circ \rho_2 \circ \rho_1 \circ \rho_0} [128].$$

With 2^{25} random samples, the experimental correlation is $-2^{-7.88}$.

6-round DL Distinguisher. We did not find any DL distinguishers for 5 rounds with single difference bit with our basic approximation approach. Thus, we first derive a 2-round differential characteristic with a similar method in [18] as follows,

$$[0] \xrightarrow[2^{-2}]{\rho_0} [0, 64, 85] \xrightarrow[2^{-6}]{\rho_1} [0, 64, 85, 91, 155, 157, 176, 221, 242].$$

The reason for choosing this differential characteristic is to minimize the hamming weight of the difference after two rounds. Next, we apply the round-based approximation for the following 4 rounds, and found the following DL distinguisher

$$[0, 64, 85, 91, 155, 157, 176, 221, 242] \xrightarrow[2^{-12.09}]{\chi \circ \rho_4 \circ \rho_3 \circ \rho_2} [228].$$

Finally, We obtain a 6-round DL distinguisher,

$$[0] \xrightarrow[2^{-20.09}]{\chi \circ \rho \circ \rho \circ \rho \circ \rho \circ \rho} [228].$$

6.2 Application to Koala-*p*

The first several operations of Koala-p round functions are linear ones, but we start with χ similarly to what we did for Subterranean-2.0. This is reasonable as we are analyzing a permutation.

5-round DL Distinguisher. The best DL distinguisher for 5-round Subterranean-2.0 detected by our method is as follows,

$$[0] \xrightarrow[2^{-6.87}]{\rho_4 \circ \rho_3 \circ \rho_2 \circ \rho_1 \circ \chi} [105].$$

With 2^{26} random samples, the experimental correlation is $2^{-6.73}$.

6-round DL Distinguisher. Similarly to Subterranean-2.0, to construct a 6-round DL distinguisher, we first find a 2-round differential characteristic as follows:

$$[0] \xrightarrow[2^{-2}]{\iota_1 \circ \theta \circ \pi \circ \chi} [0, 247, 254] \xrightarrow[2^{-6}]{\iota_2 \circ \theta \circ \pi \circ \chi} [0, 77, 84, 87, 196, 203, 206, 247, 254].$$

We then apply the round-based approximation to the following 4 rounds, and obtain a DL distinguisher,

$$[0, 77, 84, 87, 196, 203, 206, 247, 254] \xrightarrow[2^{-13.42}]{\rho_5 \circ \rho_4 \circ \rho_3 \circ \chi} [232].$$

Finally, we obtain a 6-round DL distinguisher,

$$[0] \xrightarrow[2^{-21.42}]{\rho_5 \circ \rho_4 \circ \rho_3 \circ \rho_2 \circ \rho_1 \circ \chi} [232]$$

7 Applications of Second-Order DL Attacks

The HDL attack was proposed for the first time by Biham, Dunkelman, and Keller at FSE 2005 [11], but it had not attracted much attention from the community until recently. One of the reasons is that there are no convenient means for detecting the HDL distinguishers. At Asiacrypt 2023 [23], Hu et al. applied the HATF technique to HDL attacks, proposing the first theoretical approach to detecting HDL distinguishers. In this section, we show that the round-based approximation approach can be extended for the HDL attacks. The round-based approach is more precise and easier to use than the HATF technique. The round-based approximation for the theoretical aspects of HDL attacks is similar to those of DL attacks, including the theories related to the χ function. They are essentially corollaries of Sects. 3 and 5 in the higher-order case, and therefore, we present them in Appendix A.

Here, we present the applications about the second-order DL distinguishers for `Ascon` initialization, `Ascon`-p, `Subterranean`-2.0 and `Koala`-p. For `Ascon`-128 initialization, the HATF can find at most 5-round HDL distinguishers (we do not consider the conditional case in this paper). But for 6 rounds, the precision of HATF decreases sharply, so it cannot find any distinguishers. For χ-based ciphers such as `Subterranean`-2.0 and `Koala`, there was no method to detect the HDL correlations until now.

7.1 Second-Order DL Attacks on `Ascon`-128

For 4-round `Ascon`-128, the second-order DL correlations reported in [23] are usually large. The HATF can precisely detect those large correlations for 4-round `Ascon`-128. Our round-based approach can give similar results.

However, for 5 rounds and onwards, the precision of the HATF decreases. The authors of [23] found the best 5-round second-order DL correlation is with a correlation of $2^{-6.05}$ [23, Section 5.1], and the experimental correlation is $2^{-5.60}$. With the round-based approach, the correlation of this second-order DL distinguisher is approximated as $2^{-5.63}$, which is closer to the experimental result. Furthermore, we also tested all single-bit output masks for the same input difference, and compared the theoretical correlations of the round-based approximation and the HATF results as well as the experimental results. The precision comparison is shown in Fig. 3.

Finally, for the 6-round `Ascon`-128 initialization as shown in Table 5 (Appendix C), the second-order DL correlation is estimated as $2^{-45.35}$. This is the first time that a valid DL-like distinguisher is found for the 6-round `Ascon`-128 initialization. The results for the second-order DL distinguishers have been presented in Table 1.

Remark. We also tried the second-order DL distinguishing attack on `Ascon`-128a and `Present`, but no better results than the DL attacks were found.

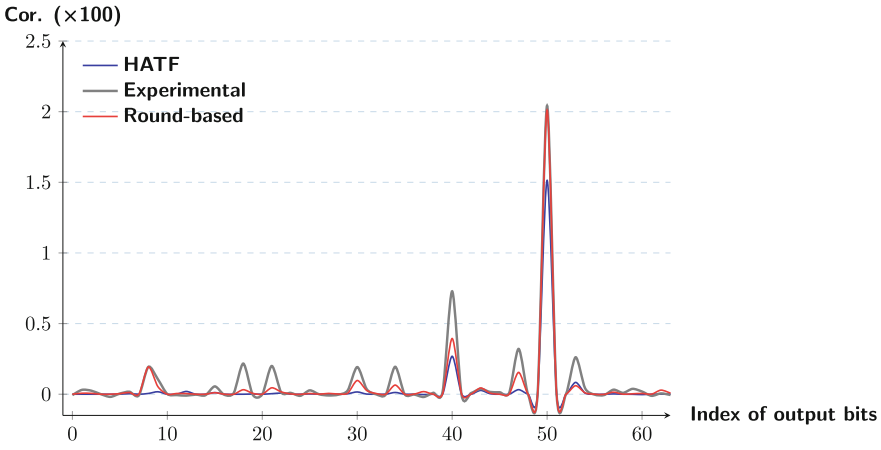

Fig. 3. The theoretical correlations of HATF and the round-based approach, and the experimental correlations. It can be seen that the round-based approximation is generally better than the HATF.

7.2 Applications to Subterranean-2.0

5-round Second-Order DL Distinguisher. The best second-order DL distinguisher for 5-round Subterranean-2.0 detected by our method is as follows,

$$([0],[32]) \xrightarrow[2^{-6.05}]{\rho_4\circ\rho_3\circ\rho_2\circ\rho_1\circ\chi} [144].$$

With 2^{26} random samples, the experimental correlation is $2^{-6.05}$.

6- and 7-round Second-Order DL Distinguisher. With the same difference as the 5-round second-order DL distinguisher, we also detected valid second-order DL distinguishers for 6 and 7 rounds.

$$([0],[32]) \xrightarrow[2^{-22.07}]{\rho_5\circ\cdots\circ\rho_1\circ\chi} [56], ([0],[32]) \xrightarrow[2^{-90.99}]{\rho_6\circ\cdots\circ\rho_1\circ\chi} [255]..$$

7.3 Applications to Koala-p

5-round Second-Order DL Distinguisher. A valid second DL distinguisher for 5-round Koala-p detected by our method is as follows,

$$([0],[11]) \xrightarrow[2^{-6.09}]{\rho_4\circ\rho_3\circ\rho_2\circ\rho_1\circ\chi} [192].$$

With 2^{26} random samples, the experimental correlation is $2^{-5.89}$.

6- and 7-round Second-Order DL Distinguisher. With the same difference as the 5-round second-order DL distinguisher, we also detected valid second-order DL distinguishers.

$$([0],[11]) \xrightarrow[2^{-22.20}]{\rho_5\circ\cdots\circ\rho_1\circ\chi} [2], ([0],[11]) \xrightarrow[2^{-86.24}]{\rho_6\circ\cdots\circ\rho_1\circ\chi} [255].$$

In Appendix B, we provide two curves for theoretical and experimental correlations for 5-round Subterranean-2.0 and Koala-p. The theoretical values match well with the experimental ones.

7.4 Applications to 8-Round Subterranean-2.0 and Koala-p

When we apply our method to 8-round Subterranean-2.0 and Koala-p, the algorithm returns the following results

$$\texttt{Subterranean} - 2.0 : ([0],[32]) \xrightarrow[2^{-63.84}]{\rho_7 \circ \cdots \circ \rho_1 \circ \chi} [193], \texttt{Koala}p : ([0],[11]) \xrightarrow[2^{-63.75}]{\rho_7 \circ \cdots \circ \rho_1 \circ \chi} [138].$$

The correlations are even larger than the 7-round cases. This suggests that when our approximation method predicts very small correlations, the impact of error may exceed the signal. Therefore, we consider the 8-round prediction results to be unreliable. How to measure the error of our approximation method is an important direction for future work.

8 Conclusion

This paper presents a round-based approach to approximate the DL and HDL correlations for primitives based either on S-boxes or on large χ such as Ascon, Present, Subterranean-2.0 and Koala-p. We obtained more precise HDL/DL correlations for these applications. For Subterranean-2.0 and Koala-p, the second-order DL distinguishers work for full rounds.

The round-based approximation starts with a new perspective on DL attacks from the viewpoint of Beyne's geometric approach. The previous method of applying the geometric approach to DL attacks is to enumerate all DL trails [24] within a mixed-basis geometric approach framework, which seems computationally infeasible in most cases. Our new approach reveals new possibilities for utilising the geometric approach in classical cryptanalysis. For differential, linear, and integral attacks, the currently dominant method is also trail-based. it is interesting to explore whether the round-based idea can be applied to more classical cryptanalysis.

The optimisation of our search algorithm is still straightforward. With more dedicated optimizations, the precision of the approximation has the potential to increase, allowing us to obtain valid DL/HDL distinguishers for more rounds of our target ciphers. In terms of the HDL attacks on χ based ciphers, it is easy to use the current framework for up to 4-th order DL attacks, which has the potential to bring better results or provide new attacks on other ciphers with χ being the nonlinear layer. We leave these interesting questions for future work.

Acknowledgements. The authors thank Tim Beyne for his kind help and fruitful discussions. This research is supported by the National Key R&D Program of China (Grant No. 2024YFA1013000, 2023YFA1009500), the National Natural Science Foundation of China (Grant No. U2336207, 62032014), Department of Science & Technology

of Shandong Province (No. SYS202201), the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM114). Kai Hu is supported by National Cryptologic Science Fund of China (2025NCSF02007), the National Natural Science Foundation of China (62402283), Shandong Provincial Natural Science Foundation (No.2025HWYQ-025), the Natural Science Foundation of Jiangsu Province (BK20240420), Program of TaiShan Scholars Special Fund for young scholars (Grant No. tsqn202507063) and Program of Qilu Young Scholars of Shandong University. Zhongfeng Niu is supported by the Singapore NRF-NRFI08-2022-0013 grant. Meiqin Wang is also supported by the National Cryptologic Science Fund of China (2025NCSF01013).

References

1. Amiri-Eliasi, P., et al.: Koala: a low-latency pseudorandom function. In: Eichlseder, M., Gambs, S. (eds.) SAC 2024, Part II. LNCS, vol. 15517, pp. 239–266. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-82841-6_10
2. Bar-On, A., Dunkelman, O., Keller, N., Weizman, A.: DLCT: a new tool for differential-linear cryptanalysis. In: Ishai, Y., Rijmen, V. (eds.) EUROCRYPT 2019, Part I. LNCS, vol. 11476, pp. 313–342. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17653-2_11
3. Beyne, T.: Block cipher invariants as eigenvectors of correlation matrices. In: Peyrin, T., Galbraith, S. (eds.) ASIACRYPT 2018, Part I. LNCS, vol. 11272, pp. 3–31. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03326-2_1
4. Beyne, T.: Block cipher invariants as eigenvectors of correlation matrices. J. Cryptol. **33**(3), 1156–1183 (2020). https://doi.org/10.1007/S00145-020-09344-1
5. Beyne, T.: A geometric approach to linear cryptanalysis. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021, Part I. LNCS, vol. 13090, pp. 36–66. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92062-3_2
6. Beyne, T.: A geometric approach to symmetric-key cryptanalysis (Ph.D. thesis). https://lirias.kuleuven.be/retrieve/713998
7. Beyne, T., Rijmen, V.: Differential cryptanalysis in the fixed-key model. In: Dodis, Y., Shrimpton, T. (eds.) CRYPTO 2022, Part III. LNCS, vol. 13509, pp. 687–716. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15982-4_23
8. Beyne, T., Verbauwhede, M.: Integral cryptanalysis using algebraic transition matrices. IACR Trans. Symmetric Cryptol. **2023**(4), 244–269 (2023). https://doi.org/10.46586/TOSC.V2023.I4.244-269
9. Beyne, T., Verbauwhede, M.: Ultrametric integral cryptanalysis. In: Chung, K., Sasaki, Y. (eds.) ASIACRYPT 2024, Part VII. LNCS, vol. 15490, pp. 392–423. Springer, Cham (2024). https://doi.org/10.1007/978-981-96-0941-3_13
10. Beyne, T., Verbauwhede, M.: Integral cryptanalysis in characteristic p. IACR Cryptol. ePrint Arch. 932 (2025). https://eprint.iacr.org/2025/932
11. Biham, E., Dunkelman, O., Keller, N.: New combined attacks on block ciphers. In: Gilbert, H., Handschuh, H. (eds.) FSE 2005. LNCS, vol. 3557, pp. 126–144. Springer, Heidelberg (2005). https://doi.org/10.1007/11502760_9
12. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. In: Menezes, A.J., Vanstone, S.A. (eds.) CRYPTO 1990. LNCS, vol. 537, pp. 2–21. Springer, Heidelberg (1991). https://doi.org/10.1007/3-540-38424-3_1

13. Blondeau, C., Leander, G., Nyberg, K.: Differential-linear cryptanalysis revisited. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 411–430. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46706-0_21
14. Blondeau, C., Leander, G., Nyberg, K.: Differential-linear cryptanalysis revisited. J. Cryptol. **30**(3), 859–888 (2017). https://doi.org/10.1007/S00145-016-9237-5
15. Bogdanov, A., et al.: PRESENT: an ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) CHES 2007. LNCS, vol. 4727, pp. 450–466. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74735-2_31
16. Che, C., Tian, T.: Enhancing the DATF technique in differential-linear cryptanalysis. IACR Cryptol. ePrint Arch. 1518 (2025). https://eprint.iacr.org/2025/1632.pdf
17. Daemen, J., Govaerts, R., Vandewalle, J.: Correlation matrices. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 275–285. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-60590-8_21
18. Daemen, J., Massolino, P.M.C., Mehrdad, A., Rotella, Y.: The Subterranean 2.0 cipher suite. IACR Trans. Symmetric Cryptol. **2020**(S1), 262–294 (2020). https://doi.org/10.13154/TOSC.V2020.IS1.262-294
19. Dobraunig, C., Eichlseder, M., Mendel, F., Schläffer, M.: Cryptanalysis of Ascon. In: Nyberg, K. (ed.) CT-RSA 2015. LNCS, vol. 9048, pp. 371–387. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16715-2_20
20. Dobraunig, C., Eichlseder, M., Mendel, F., Schläffer, M.: Ascon v1.2: lightweight authenticated encryption and hashing. J. Cryptol. **34**(3), 33 (2021). https://doi.org/10.1007/S00145-021-09398-9
21. Grassi, L., Rechberger, C., Rønjom, S.: A new structural-differential property of 5-round AES. In: Coron, J.-S., Nielsen, J.B. (eds.) EUROCRYPT 2017. LNCS, vol. 10211, pp. 289–317. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56614-6_10
22. Hadipour, H., Derbez, P., Eichlseder, M.: Revisiting differential-linear attacks via a boomerang perspective with application to AES, Ascon, CLEFIA, SKINNY, PRESENT, KNOT, TWINE, WARP, LBlock, Simeck, and SERPENT. In: Reyzin, L., Stebila, D. (eds.) CRYPTO 2024, Part IV. LNCS, vol. 14923, pp. 38–72. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-68385-5_2
23. Hu, K., Peyrin, T., Tan, Q.Q., Yap, T.: Revisiting higher-order differential-linear attacks from an algebraic perspective. In: Guo, J., Steinfeld, R. (eds.) ASIACRYPT 2023, Part III. LNCS, vol. 14440, pp. 405–435. Springer, Cham (2023). https://doi.org/10.1007/978-981-99-8727-6_14
24. Hu, K., Zhang, C., Chang, C., Zhang, J., Wang, M., Peyrin, T.: Unlocking mix-basis potential: Geometric approach for combined attacks. In: Kalai, Y.T., Kamara, S.F. (eds.) CRYPTO 2025, Part V. LNCS, vol. 16004, pp. 293–334. Springer, Cham (2025). https://doi.org/10.1007/978-3-032-01901-1_10
25. Langford, S.K., Hellman, M.E.: Differential-linear cryptanalysis. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 17–25. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48658-5_3
26. Liu, F., Isobe, T., Meier, W.: Cube-based cryptanalysis of subterranean-SAE. IACR Trans. Symmetric Cryptol. **2019**(4), 192–222 (2019). https://doi.org/10.13154/TOSC.V2019.I4.192-222
27. Liu, M., Lu, X., Lin, D.: Differential-linear cryptanalysis from an algebraic perspective. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021. LNCS, vol. 12827, pp. 247–277. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84252-9_9

28. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48285-7_33
29. Mehrdad, A., Mella, S., Grassi, L., Daemen, J.: Differential trail search in cryptographic primitives with big-circle chi: application to Subterranean. IACR Trans. Symmetric Cryptol. **2022**(2), 253–288 (2022). https://doi.org/10.46586/TOSC.V2022.I2.253-288
30. Niu, Z., Sun, S., Liu, Y., Li, C.: Rotational differential-linear distinguishers of ARX ciphers with arbitrary output linear masks. In: Dodis, Y., Shrimpton, T. (eds.) CRYPTO 2022, Part I. LNCS, vol. 13507, pp. 3–32. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15802-5_1
31. Peng, T., Zhang, W., Weng, J., Ding, T.: New approaches for estimating the bias of differential-linear distinguishers. In: Reyzin, L., Stebila, D. (eds.) CRYPTO 2024, Part IV. LNCS, vol. 14923, pp. 174–205. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-68385-5_6
32. Song, L., Tu, Y., Shi, D., Hu, L.: Security analysis of Subterranean 2.0. Des. Codes Cryptogr. **89**(8), 1875–1905 (2021). https://doi.org/10.1007/S10623-021-00892-6
33. Tezcan, C., Leander, G., Hadipour, H.: Cryptanalysis: theory versus practice. IACR Trans. Symmetric Cryptol. **2025**(3), 729–754 (2025). https://doi.org/10.46586/tosc.v2025.i3.729-754. https://tosc.iacr.org/index.php/ToSC/article/view/12484/12196
34. Gong, X., Wang, Q., Hao, Y., Jiao, L., Hu, X.: Persistence of hourglass(-like) structure: improved differential-linear distinguishers for several ARX ciphers. IACR Cryptol. ePrint Arch. 1518 (2025). https://eprint.iacr.org/2025/1667.pdf