

Related-Key Multi-Pair Neural Distinguishers: Analysis and Applications to Lightweight Block Ciphers

Thanh-Phong Nguyen, Nguyen Tan Cam, Thanh-Hien Vu, Van-Thân Huynh, Hieu-Minh Nguyen

Abstract—Neural differential cryptanalysis has recently been extended to related-key and multi-pair settings, enabling neural distinguishers to aggregate weak statistical bias across multiple ciphertext pairs, yet the statistical origin of the signal exploited by such models remains poorly understood. In this work, we present a signal-centric study of related-key multi-pair neural distinguishers across four block ciphers representing distinct design paradigms: the substitution–permutation network cipher PRESENT-80, the Boolean Feistel cipher SIMECK-32/64, the ARX-based cipher LEA-128, and the generalized Feistel cipher HIGHT. We quantitatively characterise ciphertext distributions using model-independent geometric metrics derived from Principal Component Analysis (PCA) embeddings and silhouette scores, and relate these data-level measurements to neural distinguishing accuracy under increasing aggregation levels. Across all evaluated ciphers, multi-pair aggregation systematically amplifies weak ciphertext-level statistical bias through variance reduction, extending effective distinguishability beyond the single-pair regime in favourable configurations. Our analysis further reveals clear cipher-dependent difficulty boundaries at which ciphertext-level separability collapses and neural distinguishing performance rapidly degrades. Interestingly, neural distinguishers can retain non-trivial accuracy even when linear geometric separability becomes weak, suggesting the exploitation of higher-order distributional structure beyond simple mean-shift effects. Overall, these results establish a principled connection between ciphertext-level statistical signal and neural distinguisher performance, providing a unified interpretation of the capabilities and limitations of multi-pair aggregation in related-key neural cryptanalysis. Under a unified experimental framework, our approach achieves distinguishing depths of up to 14 rounds for PRESENT-80, 17 rounds for SIMECK-32/64, 16 rounds for HIGHT, and 11 rounds for LEA-128, demonstrating consistent improvements over prior neural distinguishers.

I. INTRODUCTION

Differential cryptanalysis, introduced by Biham and Shamir [1], is a fundamental technique for evaluating the security of symmetric-key primitives. It provides a principled framework for analysing how structured input differences propagate through the internal round transformations of block

ciphers. As modern designs increasingly emphasise strong diffusion and non-linearity, however, classical differential techniques rapidly lose effectiveness beyond a limited number of rounds. This limitation has motivated the development of more powerful data-driven distinguishing methods.

Recent advances in machine learning, particularly in deep learning, have enabled models to learn hierarchical representations directly from raw data [2]. These developments have led to the emergence of neural differential cryptanalysis, in which deep neural networks are trained to distinguish ciphertext pairs generated under fixed input differences from random data. The seminal work of Gohr [3] demonstrated that neural distinguishers can outperform carefully constructed classical distinguishers on reduced-round block ciphers, and subsequent studies have explored improved architectures, input representations, and training strategies. In particular, Benamira *et al.* [4] conducted a detailed analysis of Gohr’s neural distinguisher and showed that much of its effectiveness can be explained by statistical biases related to differential propagation in earlier cipher rounds, suggesting that neural models implicitly approximate classical differential structures during training. Despite these advances, the statistical mechanisms underlying neural distinguishers remain insufficiently understood. In particular, it is often unclear which properties of the ciphertext distribution are exploited by neural models and how the availability of such signal evolves as cipher depth increases.

A key limitation of many existing studies is that neural performance is often reported in isolation, without explicitly analysing the structure of the underlying ciphertext distribution. Consequently, improvements in neural accuracy may reflect either genuine cryptanalytic signal or artefacts of model capacity, dataset generation, or training procedures. Recent work has begun to highlight the importance of dataset quality and statistical structure in neural differential cryptanalysis [5], yet a systematic connection between ciphertext-level statistical signal and neural distinguishing performance remains largely unexplored.

Further improvements to neural distinguishers have also been explored through automated differential search and enhanced data representations. Hou *et al.* [6] proposed a SAT-based method for identifying high-quality differential characteristics and introduced a neural distinguisher model that leverages multiple output differences, allowing neural networks to learn richer statistical features from ciphertext data and achieve

Thanh-Phong Nguyen and Nguyen Tan Cam are with University of Information Technology, Ho Chi Minh City, Vietnam (e-mail: phongnt.ncs@grad.uit.edu.vn; camnt@uit.edu.vn). Corresponding author: Thanh-Phong Nguyen.

Thanh-Phong Nguyen and Nguyen Tan Cam are also with Vietnam National University, Ho Chi Minh City, Vietnam.

Thanh-Hien Vu is with HUTECH University, Ho Chi Minh City, Vietnam (e-mail: vt.hien@hutech.edu.vn).

Van-Thân Huynh is with Owentis Vietnam, Ho Chi Minh City, Vietnam (e-mail: hthan401@gmail.com).

Hieu-Minh Nguyen is with Academy of Cryptography Techniques, Hanoi, Vietnam (e-mail: hieuminh@actvn.edu.vn).

improved distinguishing accuracy across several lightweight block ciphers.

Another recent research direction focuses on automating the discovery of effective input differences for neural distinguishers. In classical differential cryptanalysis, identifying suitable differences typically requires extensive manual analysis or combinatorial search techniques. To address this limitation, Bellini et al. [7] proposed a cipher-agnostic neural cryptanalysis pipeline that integrates an evolutionary algorithm for automatically identifying high-quality input differences together with a generic neural distinguisher architecture, enabling neural cryptanalysis to be applied to a wide range of block ciphers with minimal manual intervention. Building upon this direction, Wang et al. [8] introduced a framework that constructs neural distinguishers using ciphertext pairs generated from multiple input differences, combining automated difference selection, dataset generation, and neural training into a unified pipeline. These approaches highlight the growing importance of automated difference discovery and training pipelines in improving the effectiveness and generality of neural differential cryptanalysis.

Another promising direction involves extending neural distinguishers to related-key (RK) and single-key (SK) settings. The related-key attack model, first introduced by Biham [9], studies cryptanalytic scenarios in which an attacker can observe encryptions under multiple keys with known relationships. Introducing controlled key differences can expose additional statistical structure in ciphertext distributions, while aggregating multiple ciphertext pairs allows weak biases to accumulate through statistical averaging.

Recent frameworks combining the RK model with multi-pair neural distinguishers (ND_k) have demonstrated substantial empirical gains in distinguishing performance. However, it remains unclear why such aggregation succeeds in some regimes yet fails in others, and how these behaviours relate to the intrinsic statistical structure of the ciphertext data.

In our prior work [10], we explored the use of lightweight architectural enhancements for neural differential cryptanalysis in the related-key multi-pair setting. Specifically, we proposed a neural distinguisher for the PRESENT-80 cipher combining Inception-style convolutional modules with Efficient Channel Attention to improve feature extraction from structured ciphertext inputs. While this study demonstrated that multi-pair inputs and lightweight attention mechanisms can enhance distinguishing performance, its focus was primarily on architectural design and empirical evaluation on a single cipher. The present work extends this direction by investigating the statistical structure of ciphertext distributions and adopting a broader data-centric perspective on neural distinguishers.

In this work, we adopt a signal-centric perspective on neural cryptanalysis. Rather than analysing neural distinguishers in isolation, we first examine the geometry of ciphertext distributions generated under fixed plaintext and related-key differences, and then study how neural models exploit this structure. Using model-independent geometric metrics derived from PCA embeddings and silhouette scores, we quantitatively

characterise ciphertext-level statistical signal and relate it to neural distinguishing performance under increasing multi-pair aggregation levels.

To assess the generality of the observed phenomena, we conduct experiments on four block ciphers representing distinct design paradigms: the substitution-permutation network (SPN) cipher PRESENT-80, the Boolean Feistel cipher SIMECK-32/64, the add-rotate-xor (ARX) cipher LEA-128, and the generalized Feistel structure (GFS) cipher HIGHT. This architectural diversity allows us to investigate whether neural distinguishers rely on cipher-specific artefacts or capture more general statistical signals that persist across heterogeneous cipher constructions.

Across these ciphers, we observe that multi-pair aggregation systematically amplifies weak ciphertext-level statistical bias through variance reduction, extending effective distinguishability beyond the single-pair regime in favourable configurations. However, this amplification remains fundamentally limited by the intrinsic decay of ciphertext-level signal as the number of encryption rounds increases. Our analysis reveals clear cipher-dependent difficulty boundaries at which geometric separability collapses and neural distinguishing performance rapidly degrades.

The main contributions of this work are as follows:

- **A signal-centric framework for neural cryptanalysis.** We introduce a data-driven analysis pipeline that characterises ciphertext-level statistical signal using geometric metrics, enabling neural distinguishing performance to be interpreted in terms of the underlying data distribution rather than model capacity.
- **Quantitative geometry analysis of multi-pair aggregation.** We analyse how aggregating multiple ciphertext pairs reshapes dataset geometry across rounds and cipher families, identifying regimes of signal amplification, stabilisation, saturation, and collapse.
- **Cross-architecture evaluation of related-key neural distinguishers.** We evaluate related-key multi-pair neural distinguishers across four block ciphers spanning SPN, Boolean Feistel, ARX, and generalized Feistel designs, providing a unified analysis of how cipher structure influences the persistence of ciphertext-level statistical signal.
- **Improved distinguishing depth without cipher-specific tuning.** Under a fixed $(\Delta P, \Delta K)$ configuration and a generic neural architecture, our approach extends the distinguishing boundary across several ciphers, achieving up to five additional rounds compared with prior neural distinguishers in the case of PRESENT-80.

Table I summarises representative distinguishing results from prior work and those obtained in this study, highlighting the deepest distinguishable rounds across different cipher families and key models.

The remainder of this work is organised as follows. Section II reviews relevant background and related work. Section III presents the threat model and methodology. Section IV-A details the experimental setup. Section IV-B anal-

TABLE I
DEEPEST DISTINGUISHABLE ROUNDS REPORTED IN PRIOR WORK AND
ACHIEVED IN THIS STUDY.

Cipher	Des.	Key	Agg.	Rds	Work
PRESENT-80	SPN	RK	ND _k	14	This work
		RK	ND _k	9	Su <i>et al.</i> [11]
		SK	Single	10	Wang <i>et al.</i> [12]
SIMECK-32/64	Feistel	RK	ND _k	17	This work
		SK	Single	15	Lu <i>et al.</i> [13]
		RK	Single	15	Ebrahimi <i>et al.</i> [14]
HIGHT	GFS	RK	ND _k	16	This work
		RK	Single	9	Bellini <i>et al.</i> [7]
LEA-128	ARX	RK	ND _k	11	This work
		SK	Single	11	Bellini <i>et al.</i> [7]

yses ciphertext-level statistical signal using geometric characterisation. Neural distinguisher performance is examined in Section IV-C, and Section V concludes the work.

II. BACKGROUND AND RELATED WORK

A. Block Cipher Families and Cryptanalytic Diversity

Lightweight block ciphers follow several distinct structural paradigms, including substitution–permutation networks (SPN), ARX-based constructions, and various Feistel architectures. These design choices influence how statistical dependencies propagate across rounds and therefore affect the behaviour of both classical differential cryptanalysis and neural distinguishers.

PRESENT-80 [15] is a representative SPN cipher designed for ultra-lightweight hardware implementations. It introduces non-linearity through small 4-bit S-boxes and achieves diffusion via a fixed bit permutation, making differential propagation largely determined by S-box activity patterns and permutation-induced diffusion.

SIMECK-32/64 [16], [17] follows a lightweight Boolean Feistel design derived from the SIMON family. Its round function combines simple AND, XOR, and rotation operations, distributing non-linearity across Boolean operations within the Feistel structure.

LEA-128 [18] represents the ARX design paradigm, where encryption relies solely on modular addition, rotation, and XOR operations. In such constructions, statistical dependencies arise primarily from carries in modular addition.

HIGHT [19] adopts a generalized Feistel structure optimised for lightweight environments. Its byte-level operations and alternating transformations provide diffusion behaviour distinct from both SPN and ARX-based designs.

Together, these four ciphers cover several major architectural families of lightweight symmetric cryptography, enabling a cross-architecture evaluation of neural distinguishers under diverse statistical propagation behaviours.

B. Neural Differential Cryptanalysis

Differential cryptanalysis, introduced by Biham and Shamir [1], studies how input differences propagate through

iterated block ciphers and whether the resulting ciphertext differences deviate from the behaviour of a uniformly random permutation. For an n -bit block cipher E_K , the XOR difference between two plaintexts P and P' is defined as

$$\Delta P = P \oplus P'.$$

After encryption under the same key K , the corresponding ciphertext difference is

$$\Delta C = E_K(P) \oplus E_K(P').$$

Any statistically significant deviation from the behaviour of an ideal permutation may indicate exploitable structure.

A major development in this direction was the work of Gohr [3], who demonstrated that deep neural networks can serve as powerful data-driven differential distinguishers. Training convolutional models on ciphertext pairs of round-reduced Speck32/64 under a fixed input difference, Gohr showed that neural distinguishers can outperform carefully optimised classical differential distinguishers for several rounds.

Since then, neural differential cryptanalysis has been applied to a wide range of lightweight block ciphers, with research exploring improved data generation, input-difference selection, and neural architectures. Despite these advances, the statistical origin of the signal exploited by neural models remains insufficiently understood.

In this work, we adopt a signal-centric interpretation of neural differential cryptanalysis. Rather than focusing on attack construction, we study how ciphertext-level statistical signal emerges across different cipher families and data regimes, and how neural distinguishers exploit this signal under multi-pair aggregation.

C. Multi-Pair Neural Distinguishers

Early neural differential distinguishers operated on a single ciphertext pair generated under a fixed input difference and attempted to decide whether the pair originated from the real cipher or from a uniformly random permutation. Although this setting demonstrated the feasibility of neural differential cryptanalysis, it limits the distinguisher to the statistical information contained in a single ciphertext pair. In practice, the differential bias induced by a fixed input difference is often weak at the ciphertext level, especially for lightweight ciphers and for round configurations close to the security margin.

To address this limitation, Chen *et al.* [20] introduced multi-pair neural distinguishers. In this setting, each sample consists of a group of k ciphertext pairs generated under the same input difference:

$$S = \{(C_1, C_1^*), (C_2, C_2^*), \dots, (C_k, C_k^*)\}.$$

The neural model is therefore trained to distinguish between distributions defined on k -tuples of ciphertext pairs rather than on individual pairs.

The key intuition is that aggregating multiple samples drawn from a weakly biased distribution amplifies the detectable

deviation from the random baseline. While single-pair distinguishers observe only a small bias in the ciphertext distribution, multi-pair distinguishers can implicitly learn group-level statistics and dependencies across the k pairs, resulting in stronger separation between real-cipher and random data.

The multi-pair paradigm has subsequently been adopted in broader neural differential frameworks, including recent work in the related-key setting [11]. This observation directly motivates the signal-centric analysis undertaken in this work, where we study how ciphertext-level statistical signal evolves as the number of aggregated ciphertext pairs increases across different cipher families.

D. Related-Key Differential Cryptanalysis

Let $E_K : \{0, 1\}^n \rightarrow \{0, 1\}^n$ denote a block cipher under a secret key $K \in \{0, 1\}^\kappa$. In the related-key model, the adversary is allowed to query the cipher under multiple unknown keys whose pairwise relations are partially controlled. For a chosen key difference ΔK , a related key is defined as

$$K' = K \oplus \Delta K.$$

Compared with the classical single-key setting, the related-key model introduces structured perturbations to the key schedule. Appropriate choices of $(\Delta P, \Delta K)$ may therefore induce stronger statistical biases than plaintext differences alone, enabling distinguishers or attacks that are infeasible in the single-key model.

Several works have extended neural differential distinguishers to this setting. Wang *et al.* [12] demonstrated that machine-learning based related-key distinguishers can outperform classical differential distinguishers on lightweight ciphers such as SPECK32/64 and PRESENT. Lu *et al.* [13] subsequently introduced systematic related-key neural distinguishers (RKNDs) for SIMON and SIMECK, highlighting the importance of ciphertext representation and difference consistency in extracting related-key signal.

Other works have explored alternative uses of key differences, including low-weight search for effective key relations [21], analysis of conditional differential distributions in ARX-based ciphers [22], and distinguishers in rotational difference spaces [14].

Most recently, Su *et al.* [11] proposed a unified framework that combines the related-key model with multi-pair ciphertext aggregation and dataset-quality analysis. Their results show that aggregation can substantially improve neural distinguishing performance under favourable conditions and represent the closest conceptual prior to the present work.

In this paper, we adopt a simplified related-key setting in which ΔK is treated as a fixed structured perturbation. We focus on analysing how related-key signal is captured and amplified under uniform multi-pair aggregation, independently of cipher-specific tuning.

E. Attention Mechanisms in Neural Cryptanalysis

Recent work has explored the use of attention mechanisms to improve feature selection in neural differential

cryptanalysis. Channel attention modules allow neural models to assign adaptive importance weights to ciphertext features, highlighting components that carry cryptanalytic signal while suppressing noise.

Deng *et al.* [23] introduced channel-wise attention into CNN-based distinguishers for SPECK, demonstrating improved distinguishing accuracy. Guo *et al.* [24] subsequently extended this idea with attribution-based attention mechanisms for PRESENT and SKINNY.

Beyond convolutional architectures, recent work has also investigated sequence modelling approaches for neural cryptanalysis. Bose *et al.* [25] proposed sequence-based neural distinguishers using LSTM and transformer-based encoder architectures, motivated by the observation that ciphertext sequences may contain long-range dependencies that are difficult to capture with standard convolutional models. Their results demonstrate that sequence-aware neural architectures can successfully learn statistical dependencies across ciphertext tokens and achieve competitive distinguishing performance across several lightweight block ciphers. [25]

These studies suggest that attention and sequence-based architectures can act as useful architectural biases for isolating informative statistical dependencies in ciphertext data.

Table II summarises representative neural differential cryptanalysis studies across several dimensions, including cipher coverage, key model, aggregation strategy, neural architecture, and achieved round depth.

III. METHODOLOGY

A. Threat Model and Problem Setting

a) Threat model and task formulation.: Following the related-key model introduced in Section II-D, we consider encryption queries under a fixed key difference ΔK . Plaintexts are sampled uniformly at random and are not adaptively chosen.

The cryptanalytic task is formulated as a binary distinguishing problem. Each sample consists of k ciphertext pairs generated under a common plaintext difference ΔP and related-key difference ΔK . The two classes are defined as follows: (i) in the positive class, ciphertext pairs follow the structured relation $P^* = P \oplus \Delta P$ under related keys; (ii) in the negative class, the related-key condition is preserved but P^* is sampled independently and uniformly at random, breaking the alignment with ΔP .

Datasets are balanced and no key-recovery attacks are attempted. Throughout this work we focus exclusively on the distinguishing problem.

b) Cipher families and experimental scope.: To evaluate the generality of the observed phenomena, experiments are conducted on four lightweight block ciphers spanning distinct design paradigms: the SPN cipher PRESENT-80, the Boolean Feistel cipher SIMECK-32/64, the ARX-based cipher LEA, and the generalized Feistel cipher HIGHT.

This structural diversity enables us to examine whether neural distinguishers rely on cipher-specific artefacts or instead

TABLE II

METHODOLOGICAL COMPARISON OF REPRESENTATIVE NEURAL DIFFERENTIAL CRYPTANALYSIS STUDIES IN SINGLE-KEY AND RELATED-KEY SETTINGS.

Work	Cipher(s)	Key Model	Pair Aggregation	Architecture	Rounds
Gohr [3]	SPECK-32/64	Single-key	Single-pair	CNN	8 at SPECK-32/64
Deng <i>et al.</i> [23]	SPECK-32/64	Single-key	Multi-pair	ResNet + Multi-head Attention	7 at SPECK-32/64
Chen <i>et al.</i> [20]	SPECK-32/64, PRESENT-80, DES, SHA3	Single-key	Multi-pair (ND_k)	Residual CNN (1D + 2D conv)	7 at SPECK-32/64 7 at PRESENT-80
Wang <i>et al.</i> [12]	PRESENT-80, SPECK-32/64	RK (XOR) [†]	Single-pair	MLP / CNN	10 at PRESENT-80 7 at SPECK-32/64
Lu <i>et al.</i> [13]	SIMON, SIMECK	RK (XOR) [†]	Single-pair	CNN	13 at SIMECK-32/64
Tcydenova <i>et al.</i> [21]	SPECK-32/64, HIGHT, GOST	RK (regulated) [‡]	Single-pair	Cipher-specific CNN	9 at SPECK-32/64
Bao <i>et al.</i> [22]	SPECK-32/64	RK-assisted [‡]	Score-level aggregation	CNN + conditional features	10 at SPECK-32/64
Ebrahimi <i>et al.</i> [14]	SIMON, SIMECK	RK (RX) [§]	Score-level aggregation	RX-specific CNN	15 at SIMECK-32/64
Su <i>et al.</i> [11]	PRESENT-80, DES	RK (XOR) [†]	Multi-pair (ND_k)	CNN	9 at PRESENT-80
This work	PRESENT-80, SIMECK-32/64, HIGHT, LEA-128	RK (XOR) [†]	Multi-pair (ND_k)	Generic CNN (Inception + ECA)	14 at PRESENT-80 17 at SIMECK-32/64 16 at HIGHT 11 at LEA-128

[†] Fixed XOR-based related-key model with a predefined ΔK .[‡] Related-key differences used as structural or conditional constraints to regulate differential propagation or expose cipher-specific statistical structure.[§] Rotational-XOR (RX) related-key model.

capture more general statistical signals that persist across heterogeneous cipher architectures.

c) Related-key multi-pair setting.: A central methodological element of this work is the use of multi-pair samples. Rather than operating on a single ciphertext pair, each sample aggregates k independently generated ciphertext pairs sharing the same $(\Delta P, \Delta K)$ and class label. Within a sample, plaintexts and master keys are sampled independently across pairs. This construction ensures that any performance gain arises from statistical aggregation and variance reduction, rather than from shared randomness or trivial correlations.

Each ciphertext pair is represented by the feature triple

$$[\Delta C \parallel C \parallel C^*], \quad \text{where } \Delta C = C \oplus C^*,$$

and a multi-pair sample is formed by concatenating the k triples into a single input representation. From a statistical perspective, the neural distinguisher therefore implements a group-level hypothesis test, operating on aggregated observations drawn from the underlying ciphertext distribution.

d) Difference selection and experimental control.: To prevent tuning-driven bias, we enforce a disciplined difference-selection protocol. For each cipher, a single plaintext difference ΔP is fixed across all rounds, training runs, and experimental conditions, based on prior literature. The related-key difference ΔK is restricted to low-complexity perturbations and is selected using a model-independent, data-level separability criterion rather than neural classifier performance. Once selected, $(\Delta P, \Delta K)$ remain fixed for the entire experimental configuration.

Training, validation, and evaluation datasets are generated independently, differing only by fresh random sampling. No reuse of plaintexts, keys, or ciphertexts occurs across splits or

repetitions. These controls ensure that reported results reflect intrinsic properties of the ciphertext distribution rather than artefacts of data leakage or adaptive tuning.

B. Neural Architecture Design

The neural distinguisher follows a lightweight and cipher-agnostic architecture designed to expose weak ciphertext-level statistical signal while avoiding excessive model capacity. The model consists of three components: (i) multi-scale convolutional feature extraction, (ii) lightweight channel-wise attention, and (iii) a compact classification head.

Recent studies have shown that architectural design plays a critical role in improving the effectiveness of neural distinguishers. In particular, Zhang *et al.* [26], [27] proposed differential-neural distinguishers that incorporate parallel convolutional layers with multiple kernel sizes inspired by the Inception architecture [28]. This design allows the model to capture multi-dimensional statistical features from ciphertext pairs while enlarging the receptive field of the network, leading to improved distinguishing accuracy and the ability to analyse additional cipher rounds across several lightweight block ciphers.

a) Multi-scale feature extraction.: The front-end employs parallel one-dimensional convolutional branches inspired by the Inception architecture [28]. These branches probe statistical dependencies at multiple receptive-field scales, enabling the model to capture both localised nonlinear effects and broader diffusion patterns.

b) Channel-wise reweighting.: A lightweight channel-attention module is applied to adaptively reweight feature channels. This mechanism emphasises informative statistical

features while suppressing noise-dominated activations, following the Efficient Channel Attention paradigm [29].

c) *Statistical decision head.*: Finally, the aggregated representation is mapped to a scalar decision through a compact classification head, ensuring that distinguishing performance reflects extracted signal rather than model capacity.

IV. EVALUATION

This section evaluates neural distinguishers under the proposed related-key multi-pair setting from a signal-centric perspective. Rather than reporting neural performance in isolation, we first characterise the statistical structure of ciphertext distributions induced by fixed $(\Delta P, \Delta K)$ under increasing aggregation levels, and then analyse how neural distinguishers exploit this structure across cipher families and round depths. By separating data-level signal analysis from model-level performance evaluation, this section enables a direct examination of the relationship between ciphertext-level statistical bias, multi-pair aggregation, and neural distinguishability.

A. Experimental Setup

1) *Input Difference Selection.*: For each cipher, a fixed plaintext difference ΔP is used to generate structured plaintext pairs of the form (P, P^*) with $P^* = P \oplus \Delta P$. To avoid tuning-driven bias, ΔP is fixed for all rounds, aggregation levels, and training runs within each cipher.

For PRESENT-80, we adopt the difference $\Delta P = 0 \times 000000000000000080$ from [11]. For SIMECK-32/64, LEA-128, and HIGHT, the differences 0×40 , 0×800000000080 , and $0 \times 4000000000000000$ respectively are obtained using the automated input-difference discovery method proposed in Improved (Related-Key)

Once selected, ΔP remains fixed throughout the entire experimental pipeline.

2) *Related-Key Difference Selection.*: This subsection defines a disciplined, data-level procedure for calibrating a minimal related-key perturbation prior to all signal and neural evaluations. The objective is not to optimise neural distinguishing performance, but to identify a controlled key perturbation under which ciphertext-level statistical bias is maximally observable, while avoiding artefacts introduced by adaptive tuning.

a) *Restriction to minimal key perturbations.*: To isolate the effect of key-schedule perturbation and ensure interpretability, the related-key difference ΔK is restricted to have Hamming weight one, i.e., exactly one key bit is flipped. This restriction enforces a minimal and controlled perturbation model, ensuring that any observed statistical bias arises from intrinsic cipher behaviour rather than from an overly strong or artificially engineered key difference.

b) *Bit-level search space.*: Rather than selecting ΔK directly, the procedure operates at the bit level. Let κ denote the key length. The search space consists of all key-bit indices $b \in \{0, 1, \dots, \kappa - 1\}$, where each candidate induces a unit key-difference vector $\Delta K^{(b)}$ with $\Delta K^{(b)}[b] = 1$ and $\text{HW}(\Delta K^{(b)}) = 1$. The task is therefore to identify the key-bit

index whose perturbation produces the strongest ciphertext-level separability.

This bit-level selection is performed once per cipher under a fixed reference configuration, and is not repeated adaptively across rounds or aggregation levels.

c) *Bit selection criterion and construction of ΔK .*: To quantify the separability of ciphertext distributions, we first project the raw ciphertext features into a low-dimensional representation using principal component analysis (PCA) [30]. PCA is a multivariate statistical technique that transforms a set of correlated variables into a new set of orthogonal components that capture the dominant variance structure of the data. This transformation provides a compact geometric embedding in which structural differences between ciphertext classes can be more easily analysed.

Within this embedding space, the optimal bit index b^* is selected as the maximiser of the silhouette score [31] over all candidate bits. The related-key difference ΔK is then constructed deterministically as the unit vector corresponding to b^* . Algorithm 1 summarises this bit-level selection procedure. To ensure statistical relevance, the selected configuration is retained only if the corresponding silhouette score exceeds a minimum threshold, below which ciphertext distributions are empirically indistinguishable from noise.

Algorithm 1: Bit-Level Selection of Related-Key Difference ΔK

Input: Block cipher $E_K(\cdot)$ with key size κ and block size n ;

Fixed plaintext difference ΔP ;

Reference round number r_0 ;

Multi-pair size k_0 ;

Number of validation samples N

Output: Optimal key-bit index b^* ; corresponding silhouette score σ^*

bestScore $\leftarrow -\infty$;

$b^* \leftarrow \perp$;

for $b = 0$ **to** $\kappa - 1$ **do**

 Construct unit key-difference vector

$\Delta K^{(b)} \in \{0, 1\}^\kappa$ such that $\Delta K^{(b)}[b] = 1$ and $\text{HW}(\Delta K^{(b)}) = 1$;

 Generate a balanced validation dataset S_b of N multi-pair samples under $(\Delta P, \Delta K^{(b)}, r_0, k_0)$;

 Construct raw ciphertext features

$X_b = [\Delta C \parallel C \parallel C^*]$;

 Apply PCA to X_b to obtain a fixed low-dimensional embedding Z_b ;

 Compute the true-label silhouette score

$\sigma_b \leftarrow \text{Silhouette}(Z_b)$;

if $\sigma_b > \text{bestScore}$ **then**

 bestScore $\leftarrow \sigma_b$;

$b^* \leftarrow b$;

return $b^*, \sigma^* = \text{bestScore}$;

d) *Fixing policy and diagnostic analysis.*: Once selected, ΔK is fixed for all subsequent signal characterisation and neural evaluation experiments. No re-selection or adjustment based on neural distinguisher accuracy is performed. The selection procedure may be repeated across round numbers only for diagnostic purposes, in order to study the stability of the selected bit and the eventual collapse of ciphertext-level signal. Table III reports the resulting silhouette profiles, showing that the selected key bit is typically stable across rounds until geometric separability collapses, reflecting cipher-specific key-schedule sensitivity. For each cipher, the highlighted row corresponds to the deepest round configuration for which the silhouette score remains above the selection threshold, indicating the last regime where ciphertext-level separability is still statistically observable.

TABLE III
SILHOUETTE PROFILES OF SINGLE-BIT KEY PERTURBATIONS ACROSS ROUNDS, ILLUSTRATING THE STABILITY OF THE SELECTED KEY-BIT INDEX UNTIL SIGNAL COLLAPSE.

Cipher	Round	Best Key Bit	Silhouette Score
PRESENT-80	1	42	0.9089
	2	56	0.9077
	3	56	0.9079
	4	56	0.9049
	5	56	0.8959
	6	56	0.7631
	7	56	0.4770
	8	56	0.0060
LEA-128	1	116	0.8345
	2	94	0.8267
	3	94	0.8074
	4	94	0.7626
	5	94	0.6710
	6	94	0.5127
	7	94	0.2172
	8	94	0.0072
HIGHT	1	120	0.7796
	2	61	0.7676
	3	53	0.7518
	4	53	0.7217
	5	49	0.6904
	6	96	0.6478
	7	48	0.5746
	8	57	0.4414
	9	97	0.2406
	10	97	0.2242
	11	97	0.1856
	12	97	0.0419
SIMECK-32/64	1	13	0.7026
	2	57	0.6968
	3	57	0.6997
	4	57	0.6987
	5	57	0.6928
	6	57	0.6739
	7	57	0.6584
	8	57	0.6081
	9	57	0.4973
	10	57	0.3143
	11	57	0.0617

e) *From bit-level selection to related-key difference.*: Although the selection procedure operates at the level of individual key-bit indices, the outcome is not analysed at the bit level. Once the optimal index b^* is identified, the related-

key difference is defined as

$$\Delta K = e_{b^*} \in \{0, 1\}^\kappa,$$

where e_{b^*} denotes the unit vector with a single non-zero entry at position b^* . From this point onward, ΔK is treated as a fixed related-key relation that specifies the experimental condition, and its specific bit index is not further interpreted.

3) *Input Encoding and Data Representation*: This subsection describes how multi-pair ciphertext data are encoded and structured prior to neural processing. The input representation is designed to preserve differential information while enabling statistical aggregation across multiple ciphertext pairs, without introducing cipher-specific assumptions or pair-dependent identifiers.

a) *Multi-pair sample construction.*: Each input sample aggregates k ciphertext pairs generated under a common plaintext difference ΔP , related-key difference ΔK , and class label. For each pair, the triple $[\Delta C \parallel C \parallel C^*]$ is formed, and the k triples are concatenated into a single input vector of length $k \cdot 3n$, where n denotes the block size. This construction ensures that the neural distinguisher operates on grouped observations drawn from the ciphertext distribution, effectively implementing a group-level hypothesis test.

b) *Structured reshaping and channel semantics.*: Prior to convolutional processing, the concatenated input vector is reshaped into a three-dimensional tensor that exposes semantic structure to the network. Each ciphertext triple is partitioned into six logical channels of equal length, corresponding to half-word segments of ΔC , C , and C^* . After aggregation over k pairs, the resulting tensor has spatial dimension $k \cdot (n/2)$ and channel dimension 6.

This layout assigns semantic meaning to channels while allowing convolutional filters to slide across bit- and word-level positions spanning multiple pairs. As a result, local dependencies within individual ciphertext pairs and global statistical patterns across pairs can be learned jointly using standard one-dimensional convolutions.

4) *Multi-Pair Dataset Construction*: All experiments are conducted using a custom multi-pair dataset generator, denoted `NDCMultiPairGenerator`, designed to support related-key neural differential distinguishers in a cipher-agnostic manner. The generator produces balanced binary classification datasets on the fly and supports GPU-accelerated data generation for large-scale experiments.

Each sample consists of $k \in \{1, 2, 4, 8, 16\}$ independently generated ciphertext pairs. Within a sample, all pairs share the same plaintext difference ΔP , the same related-key difference ΔK , and the same class label, while plaintexts and master keys are sampled independently across pairs.

For the positive class, plaintext pairs satisfy $P^* = P \oplus \Delta P$ and are encrypted under related keys (K, K^*) with $K^* = K \oplus \Delta K$. For the negative class, P^* is sampled uniformly at random and independently of P , while the related-key condition is preserved. All datasets are strictly balanced, with equal numbers of positive and negative samples.

Algorithm 2: Multi-Pair Related-Key Dataset Generation (ND_k)

Input: Block cipher $E_K(\cdot)$ with block size n and key size κ ;
number of rounds r ; plaintext difference ΔP ; key difference ΔK ;
number of ciphertext pairs k ; batch size B
Output: Feature matrix $X \in \mathbb{R}^{B \times (3kn)}$; label vector $Y \in \{0, 1\}^B$
Initialize Y with $B/2$ ones and $B/2$ zeros;
Randomly shuffle Y ;
for $i \leftarrow 1$ **to** B **do**
 Sample master key $K_i \leftarrow \{0, 1\}^\kappa$;
 Compute related key $K_i^* \leftarrow K_i \oplus \Delta K$;
 for $j \leftarrow 1$ **to** k **do**
 Sample plaintext $P_{i,j} \leftarrow \{0, 1\}^n$;
 if $Y_i = 1$ **then**
 $P_{i,j}^* \leftarrow P_{i,j} \oplus \Delta P$;
 else
 $P_{i,j}^* \leftarrow \{0, 1\}^n$;
 $C_{i,j} \leftarrow E_{K_i}(P_{i,j})$;
 $C_{i,j}^* \leftarrow E_{K_i^*}(P_{i,j}^*)$;
 $\Delta C_{i,j} \leftarrow C_{i,j} \oplus C_{i,j}^*$;
 Construct feature vector:
 $X_i \leftarrow \left\|_{j=1}^k [\Delta C_{i,j} \parallel C_{i,j} \parallel C_{i,j}^*]\right\|$

For each ciphertext pair, the feature representation is constructed as the concatenation $[\Delta C \parallel C \parallel C^*]$, where $\Delta C = C \oplus C^*$. The k feature triples are concatenated to form a single input vector of dimension $k \cdot 3n$, where n denotes the block size.

Data generation is stateless: no plaintexts, keys, or ciphertexts are reused across samples or dataset splits. Training, validation, and test sets are generated independently, ensuring that no information leakage occurs between experimental phases.

5) *Neural Architecture Instantiation:* This subsection specifies how the conceptual neural architecture described in Section III is instantiated for the experiments. A single lightweight architecture is used across all ciphers, round configurations, and multi-pair sizes k , ensuring that observed performance differences reflect data-level effects rather than architectural tuning.

a) *Feature extraction.*: The model begins with a single Inception-style block composed of four parallel one-dimensional convolutional branches with kernel sizes (1, 3, 5, 7), providing receptive fields ranging from local bit interactions to broader diffusion patterns. Each branch uses the same number of filters and operates on the multi-pair input tensor constructed from the concatenated ciphertext triples described in Section III. Branch outputs are concatenated along the channel dimension, followed by batch normalisation and

a ReLU activation.

b) *Channel attention.*: An Efficient Channel Attention (ECA) module is applied after the Inception block to adaptively reweight feature channels. This stage emphasises informative statistical features while keeping computational overhead low.

c) *Residual convolutional tower.*: The attention-weighted features are processed by a shallow residual convolutional tower. Each block consists of two Conv1D layers with batch normalisation and ReLU activations combined with identity skip connections. The residual depth is fixed across all experiments.

d) *Classification head.*: The final representation is flattened and processed by a small fully connected stack with batch normalisation, ReLU activations, and dropout. The output layer produces a scalar value interpreted as the probability that the sample originates from the structured related-key class.

e) *Parameter stability.*: The number of trainable parameters depends on the multi-pair size k only through the input dimensionality; convolutional and dense layer widths remain unchanged. Thus increasing k affects input resolution rather than model capacity, aligning with the goal of analysing statistical amplification rather than architectural scaling effects.

All architectural hyperparameters used throughout the experiments are summarised in Table IV.

TABLE IV
MODEL ARCHITECTURE HYPERPARAMETERS USED IN ALL EXPERIMENTS.

Component	Setting
Inception kernel sizes	$\{1, 3, 5, 7\}$
Filters per branch	32
Inception output channels	128
Attention module	Efficient Channel Attention
ECA parameters	$\gamma = 2, b = 1$
Residual depth	5 blocks
Initial residual kernel size	3 (incremented by +2 per block)
Residual filters	128
Dense layers	$512 \rightarrow 64 \rightarrow 64$
Activation	ReLU
Output activation	Sigmoid
Dropout rate	0.3
Weight regularisation	$\ell_2 = 10^{-4}$
Batch normalisation	Applied throughout

6) *Training Protocol:* All models are trained and evaluated under a fixed and uniform protocol to ensure reproducibility and to avoid tuning-driven bias. Training, validation, and test datasets are generated independently using fresh random sampling, with no reuse of keys or plaintexts across splits.

Training datasets contain 10^6 samples, and evaluation is performed on independently generated test sets of the same size. Models are trained using the Adam optimiser with binary cross-entropy loss.

For round configurations where training remains stable, models are trained for a fixed number of epochs (15). When direct training fails to converge at higher round counts, a staged curriculum strategy is employed in which models are progressively fine-tuned from lower-round checkpoints with a

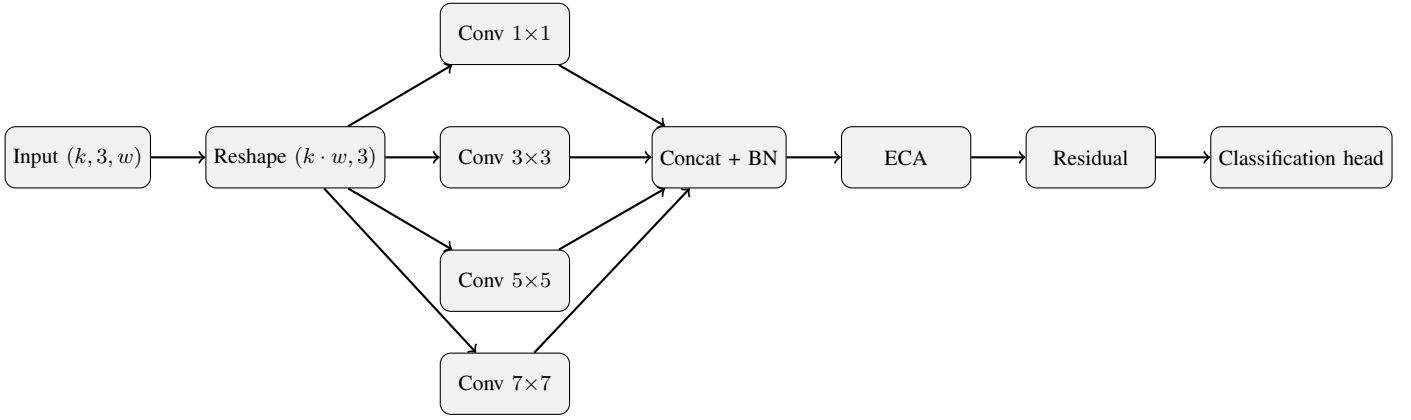


Fig. 1. Instantiation-level topology of the neural distinguisher used in all experiments. The figure illustrates how multi-pair inputs are reshaped and processed by a fixed Inception-based front-end, followed by channel-wise attention, residual refinement, and a compact decision head. This diagram specifies the concrete execution flow used for evaluation and complements the conceptual architecture presented in Section III.

reduced learning rate, following the standard practice in neural differential cryptanalysis.

Evaluation is performed using classification accuracy. Each trained model is tested on five independently generated datasets to assess robustness of the results. A round is considered distinguishable only when the observed accuracy consistently exceeds the random-guessing baseline.

7) *Computational Environment and Hardware Configuration:* All experiments were conducted on a GPU-accelerated environment using an NVIDIA A100 GPU and the TensorFlow 2.19 framework, running on an Ubuntu 22.04 system with Intel Xeon CPU and 80 GB system memory.

B. Ciphertext-Level Signal Characterisation

1) *Embedding Extraction and PCA-Based Characterisation:* To analyse the structure of related-key multi-pair datasets independently of neural training, we perform a geometric characterisation using PCA [30]. PCA provides a low-dimensional embedding that preserves the dominant variance structure of the ciphertext feature space, allowing the geometry of structured and random samples to be examined directly.

For each experimental configuration (cipher, round depth, and aggregation level k), PCA is fitted on a balanced dataset of structured and random samples constructed following the dataset-generation methodology commonly used in neural differential cryptanalysis [5]. The resulting embedding is reused for all analyses under that configuration. Low-dimensional projections (two or three principal components) are used for visualisation, while higher-dimensional embeddings are retained for quantitative geometry analysis. PCA is employed solely as an analytical tool and is not used in neural distinguisher training or inference.

2) *Quantitative Geometry Metrics:* To quantify class separability in the PCA embedding space, we compute silhouette scores [31] using the ground-truth class labels for each dataset. The silhouette score captures both inter-class distance and intra-class compactness, providing a model-independent indicator of ciphertext-level statistical bias. For each cipher,

we analyse how silhouette scores evolve with increasing round depth and multi-pair aggregation level k , thereby identifying regimes where statistical signal persists, weakens, or collapses.

a) *PRESENT-80:* We first examine ciphertext-level separability for PRESENT-80 under the fixed related-key multi-pair setting, focusing on Rounds 6 to 8. Table V reports true-label silhouette scores across aggregation levels.

At Round 6, PRESENT exhibits a weak but clearly non-zero single-pair signal, with a silhouette score of 0.3548 at $k = 1$. As k increases, silhouette scores grow smoothly and monotonically, reaching 0.7727 at $k = 16$ and further increasing to 0.8334 at $k = 32$. The approximately linear increase with respect to $\log_2(k)$ indicates that multi-pair aggregation acts as an effective variance-reduction mechanism, progressively enhancing separability without introducing abrupt structural changes. These values confirm that Round 6 lies within a regime where ciphertext-level bias is faint at the pair level but becomes statistically robust under aggregation.

At Round 7, separability degrades substantially. The single-pair silhouette score drops to 0.1005, placing the signal close to the noise floor. Although aggregation still improves separability, the maximum score rises to 0.5344 at $k = 16$ and reaches 0.6460 at $k = 32$, which remains significantly lower than the corresponding values at Round 6. This behaviour indicates that while some bias persists, it weakens rapidly with depth and requires increasing aggregation to remain detectable. Round 7 therefore represents a transitional regime near the upper boundary of effective signal survival.

At Round 8, ciphertext-level structure largely collapses. Silhouette scores remain extremely small for aggregation levels up to $k = 16$, with values close to zero and a maximum of 0.0644. Although aggregation to $k = 32$ increases the score slightly to 0.1601, the resulting separation remains weak and unstable compared with earlier rounds. From a quantitative geometry perspective, the ciphertext distribution at this depth is therefore close to that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ setting, and the observed increase under large aggregation reflects limited variance reduction rather

than the recovery of a strong underlying differential signal.

Overall, PRESENT exhibits a clear three-phase progression: a reliably separable regime at Round 6, a marginal and aggregation-dependent regime at Round 7, and near-complete signal collapse at Round 8. While multi-pair aggregation extends distinguishability by approximately one round beyond the single-pair limit, it cannot compensate indefinitely for the diffusion introduced by additional rounds.

TABLE V
TRUE-LABEL SILHOUETTE SCORES FOR PRESENT-80 UNDER VARYING
ROUND COUNTS AND MULTI-PAIR AGGREGATION LEVELS k .

Round	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
6	0.3548	0.4618	0.5857	0.6924	0.7727	0.8334
7	0.1005	0.1721	0.2664	0.4074	0.5344	0.6460
8	0.0083	0.0118	0.0256	0.0547	0.0644	0.1601

b) *SIMECK-32/64*: We next analyse ciphertext-level separability for SIMECK-32/64, a Boolean-based Feistel construction, under the same related-key multi-pair setting. Table VI reports true-label silhouette scores across round depths and aggregation levels.

At Round 11, SIMECK exhibits measurable ciphertext-level bias even in the single-pair regime, with a silhouette score of 0.0448 at $k = 1$. Multi-pair aggregation strengthens this signal in a smooth and nearly monotonic manner, with silhouette scores increasing to 0.5269 at $k = 16$ and further to 0.6344 at $k = 32$. The steady growth across aggregation levels indicates that differential structure remains statistically accessible at this depth. In terms of absolute magnitude, these values are comparable to those observed for PRESENT at moderate round depths, indicating that meaningful ciphertext-level signal persists even relatively deep within the cipher.

At Round 12, single-pair separability collapses to near-zero, with silhouette scores of 0.0037 at both $k = 1$ and $k = 2$, indicating that the ciphertext distribution is already close to random under low aggregation levels. Nevertheless, increasing the number of aggregated pairs gradually restores weak but consistent separability, with silhouette scores rising to 0.1369 at $k = 16$ and 0.2425 at $k = 32$. This behaviour reflects a gradual decay of statistical bias rather than an abrupt disappearance. In this regime, multi-pair aggregation acts primarily as a variance-reduction mechanism, allowing weak residual structure in the ciphertext distribution to become detectable through statistical accumulation.

At Round 13, ciphertext-level separability disappears entirely. True-label silhouette scores remain effectively zero for all aggregation levels, and increasing k fails to produce any measurable improvement. From a geometric perspective, the ciphertext distribution at this depth is indistinguishable from that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ configuration, indicating that the underlying differential signal has fully vanished.

Overall, SIMECK exhibits a clear three-stage progression in ciphertext-level signal behaviour: a statistically separable regime at Round 11, a weak but aggregation-recoverable regime at Round 12, and a complete signal collapse at

Round 13. While multi-pair aggregation can significantly enhance separability when residual bias is present, the results demonstrate that once the intrinsic statistical structure disappears, increasing the number of ciphertext pairs cannot recover meaningful separability. This observation highlights the existence of a strict data-level boundary beyond which multi-pair neural distinguishers cannot extract usable signal.

TABLE VI
TRUE-LABEL SILHOUETTE SCORES FOR SIMECK-32/64 UNDER VARYING
ROUND COUNTS AND MULTI-PAIR AGGREGATION LEVELS k .

Round	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
11	0.0448	0.1350	0.2477	0.3873	0.5269	0.6344
12	0.0037	0.0037	0.0068	0.0462	0.1369	0.2425

c) *HIGHT*: We next analyse ciphertext-level separability for HIGHT, a lightweight generalized Feistel cipher, under the same related-key multi-pair setting. Table VIII reports true-label silhouette scores across round depths and aggregation levels.

At Round 11, HIGHT exhibits a clearly measurable ciphertext-level signal even in the single-pair regime, with a silhouette score of 0.1398 at $k = 1$. Multi-pair aggregation strengthens this signal steadily, with silhouette scores increasing to 0.5669 at $k = 16$ and further to 0.6753 at $k = 32$. The smooth and nearly monotonic increase across aggregation levels indicates that multi-pair aggregation effectively amplifies existing statistical bias through variance reduction. These values demonstrate that differential structure remains strongly detectable at this depth.

At Round 12, ciphertext-level separability weakens but remains observable. The single-pair silhouette score drops to 0.0398, indicating that the signal has become weak at the pair level. Nevertheless, aggregation progressively restores separability, with silhouette scores increasing from 0.0714 at $k = 2$ to 0.3325 at $k = 16$ and 0.4549 at $k = 32$. This behaviour suggests that while diffusion has significantly reduced the strength of the underlying differential bias, residual statistical structure persists and can still be revealed through sufficient aggregation.

At Round 13, ciphertext-level separability collapses almost completely. Silhouette scores remain extremely close to zero across all aggregation levels, with a maximum of only 0.0085 at $k = 32$. Increasing the number of aggregated pairs fails to produce meaningful separation between structured and random samples. From a geometric perspective, the ciphertext distribution at this depth is statistically indistinguishable from that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, HIGHT exhibits a clear progression in ciphertext-level signal behaviour: a strongly separable regime at Round 11, a weakened but aggregation-recoverable regime at Round 12, and a near-complete signal collapse at Round 13. While multi-pair aggregation substantially enhances distinguishability when residual statistical bias is present, these results confirm that aggregation alone cannot compensate once the intrinsic ciphertext-level structure has vanished.

d) *LEA-128*: We next analyse ciphertext-level separability for LEA-128, an ARX-based block cipher, under the same related-key multi-pair setting. Table VII reports true-label silhouette scores across round depths and aggregation levels.

At Round 7, LEA exhibits a clearly measurable ciphertext-level signal even in the single-pair regime, with a silhouette score of 0.1826 at $k = 1$. Multi-pair aggregation strengthens this signal in a smooth and monotonic manner, with silhouette scores increasing to 0.5166 at $k = 16$ and further to 0.5565 at $k = 32$. The steady growth across aggregation levels indicates that aggregation effectively amplifies existing statistical bias through variance reduction, allowing the underlying differential structure to become increasingly detectable.

At Round 8, ciphertext-level separability weakens substantially. The single-pair silhouette score drops to 0.0166, placing the signal close to the noise floor. Nevertheless, aggregation progressively restores weak separability, with silhouette scores increasing to 0.1714 at $k = 16$ and 0.2565 at $k = 32$. This behaviour indicates that while the underlying differential signal has become extremely weak at the pair level, residual statistical structure persists and can still be revealed through sufficient aggregation.

At Round 9, ciphertext-level separability disappears entirely. Silhouette scores remain exactly zero across all aggregation levels, and increasing the number of aggregated pairs fails to produce any measurable separation between structured and random samples. From a geometric perspective, the ciphertext distribution at this depth is therefore indistinguishable from that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, LEA exhibits a clear progression in ciphertext-level signal behaviour: a strongly separable regime at Round 7, a weak but aggregation-recoverable regime at Round 8, and complete signal collapse at Round 9. While multi-pair aggregation can significantly enhance separability when residual bias is present, the results confirm that aggregation alone cannot recover meaningful signal once the intrinsic ciphertext-level statistical structure has vanished.

TABLE VII
TRUE-LABEL SILHOUETTE SCORES FOR LEA-128 UNDER VARYING
ROUND COUNTS AND MULTI-PAIR AGGREGATION LEVELS k .

Round	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
7	0.1826	0.2783	0.3814	0.4543	0.5166	0.5565
8	0.0166	0.0310	0.0596	0.1125	0.1714	0.2565

TABLE VIII
TRUE-LABEL SILHOUETTE SCORES FOR HIGHT UNDER VARYING ROUND
COUNTS AND MULTI-PAIR AGGREGATION LEVELS k .

Round	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
11	0.1398	0.2312	0.3198	0.4368	0.5669	0.6753
12	0.0398	0.0714	0.1310	0.2207	0.3325	0.4549
13	0.0004	0.0002	0.0000	0.0002	0.0002	0.0085

3) *Effect of Multi-Pair Aggregation on Data Geometry*: Multi-pair aggregation induces systematic but cipher-dependent transformations in the geometric structure of ciphertext distributions. Figure 2 illustrates how increasing the number of aggregated ciphertext pairs k reshapes the data manifold in PCA space for representative operating points of PRESENT-80, SIMECK-32/64, HIGHT, and LEA-128.

At low aggregation levels ($k = 1$), ciphertext distributions are largely noise-dominated across all evaluated ciphers. The two classes exhibit substantial overlap and form diffuse clouds in PCA space, indicating that single ciphertext pairs provide only weak statistical evidence of differential structure. In this regime, sampling variance dominates the geometry of the dataset, and geometric separation between the classes is limited.

As the number of aggregated pairs increases, multi-pair aggregation progressively reshapes the data geometry through two complementary effects: reduction of intra-class variance and increased inter-class displacement. These effects correspond to the statistical accumulation of weak differential bias across multiple independent ciphertext pairs.

For PRESENT-80, aggregation produces a clear displacement between the two classes. As k increases, the clusters separate along dominant PCA directions while remaining relatively compact internally. The resulting geometry exhibits well-defined and increasingly separated manifolds, reflecting the amplification of differential signal propagated through the SPN structure.

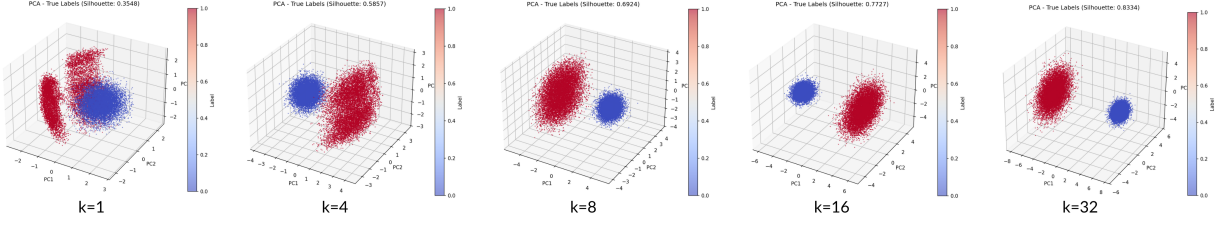
SIMECK-32/64 shows a smoother but similarly consistent evolution. Aggregation simultaneously reduces intra-class variance and induces gradual inter-class displacement. The clusters become progressively more compact while separating steadily in PCA space, indicating that differential structure remains statistically accessible and accumulates across aggregated pairs.

HIGHT exhibits a comparable but slightly more variance-dominated behaviour. At small k , the two classes remain partially overlapped. However, as k increases, clusters tighten significantly and gradually move apart, producing a well-separated geometry at larger aggregation levels. This pattern indicates that aggregation stabilises residual statistical bias while reducing the noise inherent in single-pair observations.

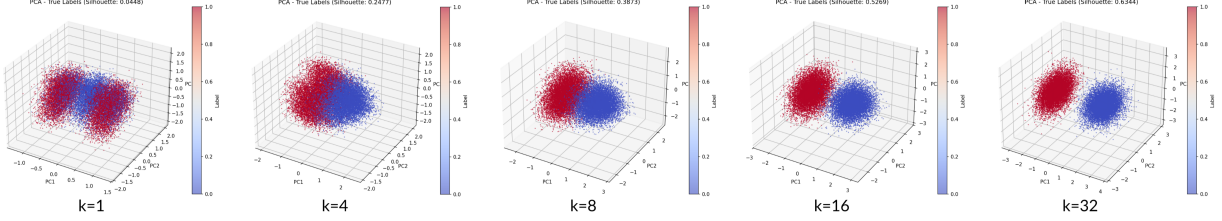
For the ARX-based cipher LEA-128, aggregation initially produces moderate cluster separation, but the geometry evolves differently as diffusion increases. While variance reduction still improves cluster compactness, the underlying signal weakens rapidly with additional rounds, resulting in weaker geometric separation compared with the other ciphers.

Across all evaluated ciphers, a consistent pattern emerges: multi-pair aggregation does not create new geometric structure, but rather reveals or stabilises statistical bias already present in the ciphertext distribution. When residual differential signal exists, aggregation improves geometric separability through variance reduction and cluster displacement. Conversely, once diffusion removes the underlying statistical structure, increasing the number of aggregated pairs fails to recover meaningful separation. These geometric observations establish clear data-

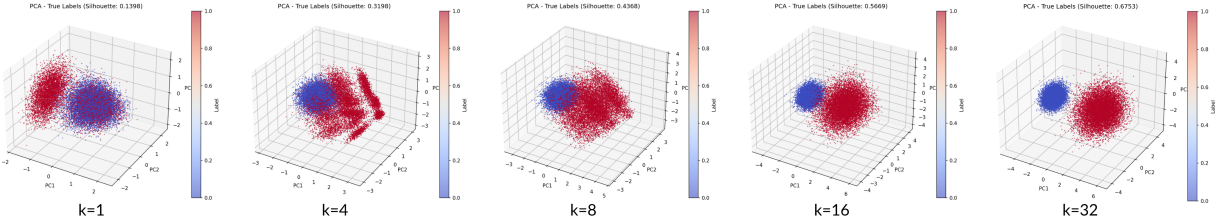
PRESENT-80: Rounds 6



SIMECK-32/64: Rounds 11



HIGHT: Rounds 11



LEA-128: Rounds 7

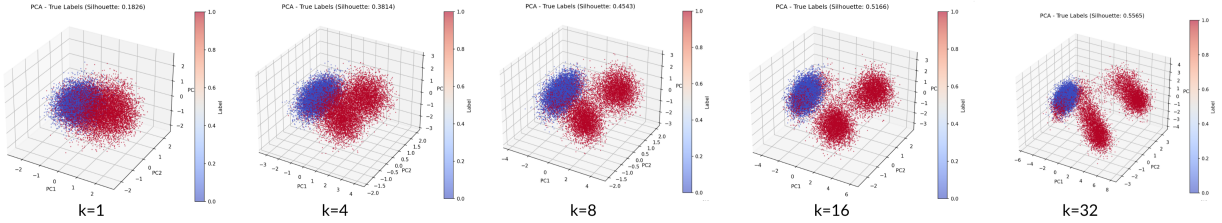


Fig. 2. Effect of multi-pair aggregation on ciphertext distribution geometry. Each panel visualizes PCA projections of aggregated ciphertext embeddings for increasing values of k , highlighting inter-class separation, variance reduction, and saturation effects across different cipher structures.

level difficulty boundaries that ultimately constrain the effectiveness of neural distinguishers trained on the same datasets.

C. Neural Distinguisher Performance Analysis

The geometric analysis in Section V establishes strict data-level constraints on the availability of ciphertext-side statistical signal. In particular, silhouette scores and PCA geometry reveal well-defined regimes of signal presence, saturation, and collapse for each cipher and round depth.

In this section, we examine how neural distinguishers respond to these regimes. Rather than treating neural performance in isolation, we analyse accuracy trends as a function of aggregation level and round number, and explicitly relate learning success or failure to the underlying data geometry. This perspective allows us to identify neural difficulty boundaries that are consistent with, and constrained by, ciphertext-level signal availability.

1) *Accuracy Across Rounds and Aggregation Levels:* We analyse neural distinguisher accuracy as a function of round depth and aggregation level k across four representative block ciphers. Accuracy is evaluated under a fixed related-key configuration and compared across single-pair and multi-pair settings.

The following paragraphs report consolidated test accuracies across round depths and aggregation levels for each cipher.

a) *PRESENT-80:* We begin by analysing neural distinguisher accuracy for PRESENT-80 as a function of the encryption round number and the aggregation level k . Table IX reports consolidated test accuracies averaged over independent test sets.

At shallow depth (Round 7), PRESENT-80 remains easily distinguishable even in the single-pair setting. The test accuracy reaches 0.9516 at $k = 1$ and rapidly exceeds 0.99 under small aggregation levels, indicating that strong ciphertext-level statistical bias is already present. Increasing the number of

aggregated pairs in this regime does not qualitatively change the outcome, but primarily stabilises performance by reducing variance across test repetitions.

At Round 8, the effect of multi-pair aggregation becomes pronounced. Single-pair accuracy drops to 0.6766, reflecting a substantial weakening of the underlying bias. However, aggregation restores distinguishability in a nearly monotonic manner: accuracy increases from 0.7556 at $k = 2$ to 0.9783 at $k = 16$, and remains above 0.98 for larger aggregation levels. These results indicate that although ciphertext-level signal weakens at this depth, it remains statistically exploitable when multiple independent ciphertext pairs are aggregated.

At Round 9, single-pair distinguishers operate close to the random-guessing baseline, with accuracy 0.5161 at $k = 1$. Nevertheless, aggregation continues to provide measurable gains. Accuracy grows steadily with increasing k , reaching 0.6763 at $k = 16$ and 0.7552 at $k = 32$. This behaviour places Round 9 in a marginal regime where distinguishability exists but is weak and strongly dependent on aggregation.

At Round 10, neural distinguishability collapses. Even with aggressive aggregation, test accuracy remains close to the random-guessing baseline, reaching only 0.5515 at $k = 32$. These values indicate that the ciphertext distribution at this depth becomes effectively indistinguishable from that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, PRESENT-80 exhibits a clear aggregation-dependent shift in distinguishability. Multi-pair aggregation extends the effective neural distinguishing range by approximately one round—from Round 8 in the single-pair setting to Round 9 under large aggregation levels—while the fundamental difficulty boundary remains unchanged.

TABLE IX

NEURAL DISTINGUISHER TEST ACCURACY FOR PRESENT-80 ACROSS ROUND DEPTHS AND MULTI-PAIR AGGREGATION LEVELS k . RESULTS ARE AVERAGED OVER INDEPENDENT TEST SETS.

Rounds	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
7	0.9516	0.9919	0.9954	0.9683	0.9808	0.9920
8	0.6766	0.7556	0.8438	0.9235	0.9783	0.9822
9	0.5161	0.5396	0.5674	0.6170	0.6763	0.7552
10	—	—	—	0.4999	0.5356	0.5515

b) *SIMECK-32/64*: We next analyse neural distinguisher accuracy for SIMECK-32/64 under the same related-key multi-pair setting. Table X reports consolidated test accuracies across round depths and aggregation levels k .

At Round 12, SIMECK exhibits a strong ciphertext-level signal even in the single-pair regime, achieving a test accuracy of 0.8787 at $k = 1$. Multi-pair aggregation rapidly amplifies this bias: accuracy exceeds 0.99 for $k \geq 8$ and reaches 0.9987 at $k = 16$. This behaviour indicates that the underlying statistical signal remains strong at this depth and can be exploited reliably by the neural distinguisher once multiple ciphertext pairs are aggregated.

At Round 13, neural distinguishability weakens but remains clearly above the random-guessing baseline. While single-pair

accuracy drops to 0.6754, aggregation substantially improves performance. Accuracy increases from 0.7568 at $k = 2$ to 0.9425 at $k = 16$, and reaches 0.9510 at $k = 32$. This places Round 13 in an aggregation-dependent regime where the ciphertext-level bias persists but becomes increasingly masked by diffusion noise.

At Round 14, distinguishability becomes marginal. Low aggregation levels yield accuracies only slightly above the random-guessing baseline, while larger aggregation levels partially restore performance. Accuracy reaches 0.6702 at $k = 16$ and 0.7810 at $k = 32$, indicating that although residual statistical bias remains detectable, it is substantially weaker than in previous rounds.

At Round 15, neural distinguishability collapses. Only the largest aggregation levels are evaluable, and the resulting accuracies remain close to the random-guessing baseline, reaching 0.5933 at $k = 32$. This behaviour indicates that the ciphertext distribution at this depth approaches pseudorandom permutation behaviour under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, SIMECK-32/64 exhibits a deeper aggregation-supported distinguishing range than PRESENT-80. While aggregation substantially amplifies weak statistical bias at intermediate depths, the effective distinguishing boundary remains approximately one round beyond the single-pair regime.

TABLE X

NEURAL DISTINGUISHER TEST ACCURACY FOR SIMECK-32/64 ACROSS ROUND DEPTHS AND MULTI-PAIR AGGREGATION LEVELS k . RESULTS ARE AVERAGED OVER INDEPENDENT TEST SETS.

Rounds	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
12	0.8787	0.9408	0.9883	0.9976	0.9987	0.9917
13	0.6754	0.7568	0.8098	0.8149	0.9425	0.9510
14	—	0.5662	0.5777	0.5875	0.6702	0.7810
15	—	—	—	—	0.5489	0.5933

c) *HIGHT*: We next analyse neural distinguisher accuracy for HIGHT under the same related-key multi-pair setting. Table XI reports consolidated test accuracies across round depths and aggregation levels k .

At Round 12, HIGHT exhibits a clear ciphertext-level signal even in the single-pair regime, with a test accuracy of 0.7355 at $k = 1$. Multi-pair aggregation substantially amplifies this signal, and accuracy increases steadily with larger aggregation levels, reaching 0.9911 at $k = 16$ and 0.9987 at $k = 32$. This behaviour indicates that although the single-pair signal is weaker than that observed for some other ciphers, aggregation effectively stabilises and amplifies the underlying statistical bias.

At Round 13, neural distinguishability weakens noticeably. Single-pair accuracy drops to 0.5798, placing the distinguisher close to the random-guessing baseline. However, aggregation continues to improve performance, with accuracy rising gradually from 0.5992 at $k = 2$ to 0.8426 at $k = 32$. This behaviour places Round 13 in a clear aggregation-dependent regime where statistical bias remains detectable but requires multiple ciphertext pairs to be reliably exploited.

At Round 14, neural distinguishability becomes marginal. Low aggregation levels fail to produce measurable signal, and only large aggregation levels yield accuracies above the random baseline. Even at $k = 32$, test accuracy reaches only 0.6252, indicating that the remaining ciphertext-level bias is extremely weak. These results suggest that the ciphertext distribution at this depth is close to pseudorandom under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, HIGHT exhibits a moderate aggregation-supported distinguishing range. While multi-pair aggregation significantly improves neural performance at intermediate depths, the effective distinguishing boundary remains limited, with signal rapidly fading beyond Round 13.

TABLE XI
NEURAL DISTINGUISHER TEST ACCURACY FOR HIGHT ACROSS ROUND DEPTHS AND MULTI-PAIR AGGREGATION LEVELS k . RESULTS ARE AVERAGED OVER INDEPENDENT TEST SETS.

Rounds	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
12	0.7355	0.8507	0.9195	0.9664	0.9911	0.9987
13	0.5798	0.5992	0.6361	0.6907	0.7572	0.8426
14	–	–	–	–	0.5820	0.6252

d) LEA: We finally analyse neural distinguisher accuracy for LEA under the same related-key multi-pair setting. Table XII reports consolidated test accuracies across round depths and aggregation levels k .

At Round 7, LEA exhibits a clear ciphertext-level signal in the single-pair setting, achieving a test accuracy of 0.9244 at $k = 1$. Aggregation provides moderate improvements, with accuracy gradually increasing to 0.9624 at $k = 16$ and 0.9701 at $k = 32$. These results indicate that the underlying statistical bias is already detectable without aggregation, while additional ciphertext pairs primarily improve stability and robustness.

At Round 8, neural distinguishability weakens but remains clearly above the random-guessing baseline. Single-pair accuracy decreases to 0.7598, while aggregation steadily improves performance, reaching 0.8942 at $k = 16$ and 0.9134 at $k = 32$. This behaviour indicates that aggregation effectively amplifies the remaining statistical signal as diffusion begins to obscure the differential structure.

At Round 9, the neural distinguisher operates close to the random baseline in the single-pair regime, with accuracy 0.5780 at $k = 1$. However, aggregation substantially improves performance, with accuracy rising to 0.8046 at $k = 16$ and 0.8454 at $k = 32$. This places Round 9 in an aggregation-dependent regime where the remaining bias is weak but still statistically exploitable when sufficient ciphertext pairs are aggregated.

At Round 10, neural distinguishability collapses. Even under the largest aggregation level, test accuracy remains at 0.5009, which is statistically indistinguishable from random guessing. This behaviour indicates that the ciphertext distribution at this depth has effectively converged to pseudorandom permutation behaviour under the fixed $(\Delta P, \Delta K)$ configuration.

Overall, LEA exhibits a moderate aggregation-supported distinguishing range. Multi-pair aggregation substantially improves neural performance at intermediate depths, but does not extend the effective distinguishing boundary beyond approximately one additional round beyond the single-pair regime.

TABLE XII
NEURAL DISTINGUISHER TEST ACCURACY FOR LEA ACROSS ROUND DEPTHS AND MULTI-PAIR AGGREGATION LEVELS k . RESULTS ARE AVERAGED OVER INDEPENDENT TEST SETS.

Rounds	$k = 1$	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$
7	0.9244	0.9132	0.9320	0.9320	0.9624	0.9701
8	0.7598	0.8086	0.8644	0.8644	0.8942	0.9134
9	0.5780	0.6191	0.6310	0.6310	0.8046	0.8454
10	–	–	–	–	–	0.5009

Overall, these results demonstrate that neural distinguishing performance is closely aligned with the availability of ciphertext-level statistical signal. Multi-pair aggregation can substantially improve empirical accuracy when weak bias persists, but it cannot recover distinguishability once the underlying ciphertext distribution becomes statistically indistinguishable from a pseudorandom permutation.

2) Generalization to Higher Rounds: To further investigate the relationship between neural distinguishing performance and ciphertext-level statistical signal, we evaluate whether models trained near the empirical distinguishing boundary retain any predictive capability when applied to deeper cipher configurations.

For each cipher, the neural distinguisher trained at the highest distinguishable round identified in the previous experiments is reused without further training. All evaluations use the same ND_k configuration with $k = 32$ ciphertext pairs per sample. For each evaluation round, test accuracy is measured across five independent runs, and we report the average accuracy together with the standard deviation.

a) PRESENT-80: For PRESENT-80, the model trained at Round 10 achieves an accuracy of 0.5515 at the training depth. When evaluated at Round 11, accuracy drops sharply to 0.5077, indicating that the statistical bias exploited by the distinguisher rapidly weakens after one additional round of diffusion. Beyond Round 11, the accuracy converges to the random-guessing baseline. From Round 12 onward, the measured accuracy fluctuates around 0.50, suggesting that the ciphertext distribution becomes effectively indistinguishable from that of a pseudorandom permutation under the fixed $(\Delta P, \Delta K)$ configuration.

Table XIII reports the generalization behaviour for PRESENT-80.

For visualization purposes, we illustrate the behaviour using PRESENT-80 as a representative example in Fig. 3.

b) SIMECK-32/64: A similar pattern is observed for SIMECK-32/64. The model trained at Round 15 achieves 0.5933 accuracy at the training depth, indicating strong distinguishability. However, when evaluated at Round 16, accuracy decreases rapidly to 0.5109. At Round 17 and beyond, accuracy approaches the random baseline, indicating that the

TABLE XIII
GENERALIZATION OF THE NEURAL DISTINGUISHER TRAINED AT ROUND 10 OF PRESENT-80 AND EVALUATED ON DEEPER ROUNDS.

Evaluation Round	Test Accuracy
10	0.5515 ± 0.0013
11	0.5077 ± 0.0016
12	0.5012 ± 0.0013
13	0.5001 ± 0.0003
14	0.5001 ± 0.0004
15	0.4995 ± 0.0015

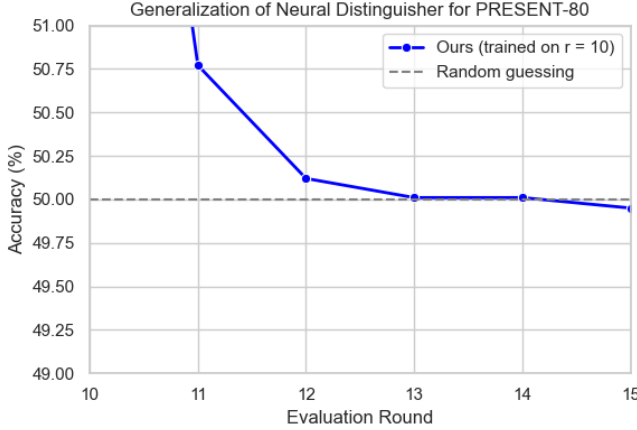


Fig. 3. Generalization behaviour of the neural distinguisher trained at Round 10 of PRESENT-80 and evaluated on deeper rounds. The dashed line indicates the random guessing baseline (50%).

residual statistical signal is almost entirely suppressed after two additional rounds.

The corresponding results for SIMECK-32/64 are summarized in Table XIV.

TABLE XIV
GENERALIZATION OF THE NEURAL DISTINGUISHER TRAINED AT ROUND 15 OF SIMECK-32/64 AND EVALUATED ON DEEPER ROUNDS.

Evaluation Round	Test Accuracy
15	0.5933 ± 0.0011
16	0.5109 ± 0.0007
17	0.5011 ± 0.0012
18	0.4997 ± 0.0015

c) *HIGHT*: For HIGHT, the model trained at Round 14 achieves 0.6252 accuracy at the training round. However, distinguishability collapses immediately when evaluated at Round 15, where accuracy falls to 0.5000. Further increases in the number of rounds yield accuracies indistinguishable from random guessing.

The numerical results are reported in Table XV.

d) *LEA-128*: For LEA-128, evaluation results remain at the random-guessing baseline across all tested rounds. Accuracy values remain extremely close to 0.50 for Rounds 10–12, indicating that no measurable residual signal is detectable beyond the training configuration.

The detailed results are shown in Table XVI.

e) *Cross-Cipher Observation*: Across all evaluated ciphers, a consistent pattern emerges: neural distinguishing

TABLE XV
GENERALIZATION OF THE NEURAL DISTINGUISHER TRAINED AT ROUND 14 OF HIGHT AND EVALUATED ON DEEPER ROUNDS.

Evaluation Round	Test Accuracy
14	0.6252 ± 0.0011
15	0.5000 ± 0.0012
16	0.5001 ± 0.0012
17	0.4995 ± 0.0006

TABLE XVI
GENERALIZATION OF THE NEURAL DISTINGUISHER TRAINED AT ROUND 10 OF LEA-128 AND EVALUATED ON DEEPER ROUNDS.

Evaluation Round	Test Accuracy
10	0.5007 ± 0.0001
11	0.5001 ± 0.0001
12	0.4999 ± 0.0003

accuracy drops rapidly once the evaluation round exceeds the empirical distinguishing boundary. In some cases, a small residual bias is observable immediately after the training round (e.g., PRESENT Round 11 or SIMECK Round 16), but this effect disappears quickly as additional diffusion rounds are introduced.

These results suggest that neural distinguishers primarily exploit existing ciphertext-level statistical structure rather than learning round-invariant representations. Once cipher diffusion suppresses this statistical signal, neural accuracy converges to the random-guessing baseline regardless of model architecture or aggregation level.

Overall, this experiment reinforces the signal-centric interpretation adopted in this work: neural distinguishing performance is fundamentally constrained by the availability of statistical signal in the ciphertext distribution.

V. CONCLUSION AND FUTURE WORK

This work studied neural distinguishers in a related-key multi-pair setting from a signal-centric perspective. By combining ciphertext-level geometric analysis with neural performance evaluation, we showed that neural distinguishability is fundamentally constrained by the availability of statistical bias in ciphertext distributions. Multi-pair aggregation acts primarily as a variance-reduction mechanism that stabilises and amplifies existing bias, extending the effective distinguishing range by approximately one round, but it does not create new structure. While moderate geometric separability is sufficient for high neural accuracy, it is not strictly necessary: in some cases, neural models successfully exploit residual higher-order structure not captured by low-dimensional PCA embeddings. Nevertheless, once ciphertext distributions collapse across aggregation levels, neural distinguishers fail consistently, confirming that neural performance remains bounded by data-level signal rather than model capacity.

Future work will focus on refining data-level diagnostics beyond linear PCA, in order to capture higher-order or non-linear structure that may be exploited by neural models. Additional directions include evaluating robustness across multiple

plaintext and key differences, analysing internal neural representations to better understand learned features, and extending the proposed framework to more complex attack settings. Overall, this study highlights the importance of principled, model-independent signal analysis as a foundation for rigorous and interpretable neural cryptanalysis.

REFERENCES

- [1] E. Biham and A. Shamir, "Differential cryptanalysis of des-like cryptosystems," *Journal of CRYPTOLOGY*, vol. 4, no. 1, pp. 3–72, 1991.
- [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [3] A. Gohr, "Improving attacks on round-reduced speck32/64 using deep learning," in *Annual International Cryptology Conference*. Springer, 2019, pp. 150–179.
- [4] A. Benamira, D. Gerault, T. Peyrin, and Q. Q. Tan, "A deeper look at machine learning-based cryptanalysis," in *Annual international conference on the theory and applications of cryptographic techniques*. Springer, 2021, pp. 805–835.
- [5] B. Seok and C. Lee, "A novel approach to construct a good dataset for differential-neural cryptanalysis," *IEEE Transactions on Dependable and Secure Computing*, vol. 22, no. 1, pp. 246–262, 2024.
- [6] Z. Hou, J. Ren, and S. Chen, "Improve neural distinguisher for cryptanalysis," *Cryptology ePrint Archive*, 2021.
- [7] E. Bellini, D. Gerault, A. Hambitzer, and M. Rossi, "A cipher-agnostic neural training pipeline with automated finding of good input differences," *IACR Transactions on Symmetric Cryptology*, vol. 2023, no. 3, pp. 184–212, 2023.
- [8] G. Wang, G. Wang, and S. Sun, "A new (related-key) neural distinguisher using two differences for differential cryptanalysis," *IET Information Security*, vol. 2024, no. 1, p. 4097586, 2024.
- [9] E. Biham, "New types of cryptanalytic attacks using related keys," *Journal of Cryptology*, vol. 7, no. 4, pp. 229–246, 1994.
- [10] T.-P. Nguyen, N. T. Cam, V.-T. Huynh, T. Nguyen, and H.-M. Nguyen, "Improved neural distinguisher for present-80 using inception and efficient channel attention in related-key multi-pair setting," in *2025 2nd International Conference On Cryptography And Information Security (VCRIS)*. IEEE, 2025, pp. 1–8.
- [11] R. Su, J. Ren, and S. Chen, "Enhanced related-key differential neural distinguishers with data-driven insights: Breakthroughs in the standard ciphers," *IEEE Internet of Things Journal*, 2026.
- [12] G. Wang, G. Wang, and Y. He, "Improved machine learning assisted (related-key) differential distinguishers for lightweight ciphers," in *2021 IEEE 20th international conference on trust, security and privacy in computing and communications (trustCom)*. IEEE, 2021, pp. 164–171.
- [13] J. Lu, G. Liu, B. Sun, C. Li, and L. Liu, "Improved (related-key) differential-based neural distinguishers for simon and simeck block ciphers," *The Computer Journal*, vol. 67, no. 2, pp. 537–547, 2024.
- [14] A. Ebrahimi, D. Gerault, and P. Palmieri, "Deep learning-based rotational-xor distinguishers for and-rx block ciphers: Evaluations on simeck and simon," in *International Conference on Selected Areas in Cryptography*. Springer, 2023, pp. 429–450.
- [15] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. Robshaw, Y. Seurin, and C. Vikkelsoe, "Present: An ultra-lightweight block cipher," in *International workshop on cryptographic hardware and embedded systems*. Springer, 2007, pp. 450–466.
- [16] G. Yang, B. Zhu, V. Suder, M. D. Aagaard, and G. Gong, "The simeck family of lightweight block ciphers," in *International workshop on cryptographic hardware and embedded systems*. Springer, 2015, pp. 307–329.
- [17] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, and L. Wingers, "The simon and speck lightweight block ciphers," in *Proceedings of the 52nd annual design automation conference*, 2015, pp. 1–6.
- [18] D. Hong, J.-K. Lee, D.-C. Kim, D. Kwon, K. H. Ryu, and D.-G. Lee, "Lea: A 128-bit block cipher for fast encryption on common processors," in *international workshop on information security applications*. Springer, 2013, pp. 3–27.
- [19] D. Hong, J. Sung, S. Hong, J. Lim, S. Lee, B.-S. Koo, C. Lee, D. Chang, J. Lee, K. Jeong *et al.*, "Hight: A new block cipher suitable for low-resource device," in *International workshop on cryptographic hardware and embedded systems*. Springer, 2006, pp. 46–59.
- [20] Y. Chen, Y. Shen, H. Yu, and S. Yuan, "A new neural distinguisher considering features derived from multiple ciphertext pairs," *The Computer Journal*, vol. 66, no. 6, pp. 1419–1433, 2023.
- [21] E. Tcydenova, B. Seok, and C. Lee, "Related-key neural distinguisher on block ciphers speck-32/64, hight and gost," *Journal of Platform Technology*, vol. 11, no. 1, pp. 72–84, 2023.
- [22] Z. Bao, J. Lu, Y. Yao, and L. Zhang, "More insight on deep learning-aided cryptanalysis," in *International conference on the theory and application of cryptology and information security*. Springer, 2023, pp. 436–467.
- [23] H. Deng, X. Cao, and Y. Cheng, "Attention in differential cryptanalysis on lightweight block cipher speck," in *2023 20th Annual International Conference on Privacy, Security and Trust (PST)*. IEEE, 2023, pp. 1–9.
- [24] Y. Guo, Y. Lu, W. Liu, W. Chen, and B. Yu, "Improved differential neural distinguishers for present and skinny," *Physica Scripta*, vol. 100, no. 6, p. 066003, 2025.
- [25] A. Bose, D. Pal, and D. Roy Chowdhury, "Deep learning-based differential distinguishers for cryptographic sequences," in *International Conference on Cryptology in India*. Springer, 2024, pp. 114–133.
- [26] L. Zhang, Z. Wang *et al.*, "Improving differential-neural cryptanalysis," *Cryptology ePrint Archive*, 2022.
- [27] L. Zhang, Z. Wang, and Y. Chen, "Improving the accuracy of differential-neural distinguisher for des, chaskey, and present," *IEICE TRANSACTIONS on Information and Systems*, vol. 106, no. 7, pp. 1240–1243, 2023.
- [28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [29] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 534–11 542.
- [30] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [31] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.