Adaptive Recovery Attack on Cyclic Shift Under Uncertain Ciphertexts

Xu Liu[1], Hongmin Liu[2], Yihui Zhou[3, 4], and Laifeng Lu[1]
[1]School of Mathematics and Statistics, Shaanxi Normal University, Xi'an, 710119, China
[2]School of Innovation and Entrepreneurship, North University of China, Taiyuan, 030051, China
[3]Shanxi Key Laboratory of Network and System Security, Xidian University, Xi'an, 710071, China
[4]School of Artificial Intelligence and Computer Science, Shaanxi Normal University, Xi'an, 710119, China

With the rapid development of lightweight network scenarios such as low-orbit satellite communication, the Internet of Things (IoT) and edge computing, block ciphers adopting cyclic shift as the basic permutation operation are widely deployed on resource-constrained devices. Conventional cryptanalysis relies on the ideal noise-free ciphertext assumption, under which adversaries can obtain complete and accurate ciphertexts. Nevertheless, during practical transmission, channel noise and environmental interference cause random bit flips in ciphertexts, generating uncertain ciphertexts and rendering deterministic recovery methods invalid. To address this issue, we propose an adaptive cyclic shift recovery attack based on Hamming distances, which realizes key recovery without any prior knowledge of noise parameters. We first design a lightweight single-column anchor recovery scheme. This scheme calculates the Hamming distances between the target plaintext anchor column and all ciphertext columns, and automatically identifies the noise bias direction by comparing the deviation magnitudes of the minimum distance and the maximum distance relative to the random expectation. Furthermore, we further develop a robust multi-column voting recovery scheme. Each column performs independent matching and votes for candidate shift offsets, and the candidate offset with the largest number of votes is taken as the recovered key. We then rigorously derive closed-form expressions for the key recovery success rates of the two schemes. Monte Carlo simulations are conducted to verify the theoretical analysis.

***Index Terms*—chosen-plaintext attack, cyclic shift, noise interference, Hamming distance, transition probability matrix.**

I. Introduction

WITH the rapid development of network scenarios such as low-orbit satellite Internet [1], the Industrial Internet of Things (IIoT) and edge computing, guaranteeing the security and confidentiality of data transmission has become the core requirement of network systems. As the cornerstone of modern cryptography, block ciphers are widely adopted to protect data confidentiality [2]. Among them, block ciphers such as DES, AES and SM4 algorithms [3] which adopt cyclic shifts as their basic permutation operations, feature low computational overhead and high hardware implementation efficiency, so they are extensively deployed on resource-constrained devices including satellite communication terminals and IoT sensors.

In practical block ciphers, the cyclic shift primitive is widely deployed as a fundamental building block: in DES, cyclic left shifts of the 28-bit key registers are a core operation in the round key schedule; in AES, the ShiftRows operation cyclically shifts each row of the state matrix by a fixed number of bytes, serving as an essential component of the diffusion layer; in SM4, the linear transformation $\mathcal{L}$ is composed of multiple cyclic rotations of a 32-bit word by fixed amounts. Although the full round structures and secret keys of these ciphers are far more complex than a single shift offset k, and the shift amounts in these ciphers are typically fixed constants rather than secret keys, the cyclic shift remains a fundamental permutation primitive whose statistical behavior under noise directly affects the security of the overall cipher. This paper takes the most basic form of cyclic shift (a variable shift offset k applied to an entire n-bit block) as its research subject, with the aim of establishing a general methodology and analytical framework that can be extended to more complex multi-round ciphers in future work.

As a linear permutation operation, cyclic shift essentially shifts an n-bit binary sequence left or right by k positions as a whole, where k serves as the shift key. This operation plays a vital role in the diffusion layer of block ciphers: by rearranging the positions of plaintext bits, cyclic shift propagates the influence of a single plaintext bit to multiple ciphertext bits, thus strengthening the cipher's resistance against statistical attacks [4]. Cyclic shift serves as a core building block of the diffusion layer in multi-round block ciphers[5]. This paper first takes the fundamental cyclic shift as its research subject, investigates attack methods targeting this most universally applicable low-level permutation primitive, and refines the general key recovery logic and attack boundaries thereof. The proposed attack methodology and analytical framework enjoy sound extensibility: on this basis, follow-up research can be further extended to attack studies on multi-round iterative block ciphers, providing general reference and methodological support for the security analysis of similar cryptographic algorithms incorporating cyclic shift structures.

Conventional block cipher cryptanalysis has formed a relatively mature theoretical framework. Biham and Shamir proposed differential cryptanalysis, which constructs distinguishers by propagating plaintext differentials and has been applied to the cryptanalysis of DES [6]. Matsui put forward linear cryptanalysis, which recovers secret keys by exploiting the bias of linear approximations between plaintexts and ciphertexts

Manuscript received July 31, 2026; revised November 16, 2026. Corresponding author: Laifeng Lu (email: lulaifeng@snnu.edu.cn).

[7]. Since then, truncated differential cryptanalysis, impossible differential cryptanalysis, integral cryptanalysis and other approaches have emerged one after another. Conventional block cipher cryptanalytic methods are all built upon the ideal noise-free ciphertext assumption: adversaries can obtain complete, correct and undisturbed ciphertexts, and there exists a deterministic correspondence among ciphertexts, plaintexts and keys. Nevertheless, in practical transmission scenarios, ionospheric scintillation in satellite communications, multipath fading of wireless channels, and environmental noise in IoT deployments all trigger random bit flips in ciphertexts, which generate uncertain ciphertexts. Such uncertainty directly disrupt the statistical features required by traditional cryptanalysis, resulting in invalid distinguishers and a dramatic drop in attack success rate.

Against this backdrop, cryptanalysis under noisy environments has attracted growing research attention. Xi et al. constructed an XOR encryption-perturbation model and investigated the key recovery method for XOR encryption under noisy environments with known noise parameters [8]. Zhou et al. proposed two algorithms for recovering permutation operations under noise interference [9]. Considering three scenarios including noise-free condition, scenario with known noise parameters and scenario with unknown noise parameters, Li et al. conducted key recovery research on cyclic shift components using the correlation coefficient method [10]. Nevertheless, this method relies on sophisticated floating-point statistical computations, which incurs non-negligible computational overhead. Moreover, it is incapable of identifying the perturbation direction, presenting distinct limitations in direction-sensitive noise analysis scenarios.

To address the aforementioned limitations, this paper constructs a hierarchical attack system ranging from lightweight schemes to robust schemes. The main contributions of this work are summarized as follows: a lightweight single-column anchor recovery method is proposed. It only matches Hamming distances between one target plaintext column and all ciphertext columns. It then automatically identifies the noise direction by comparing the deviation of minimum and maximum distances from the random expectation, without any prior knowledge of p and q. This method is applicable to extremely resource-limited scenarios such as low-power IoT nodes. A robust multi-column voting recovery method is presented. Leveraging the structural property of cyclic shifts that all columns share an identical shift offset, each column performs independent matching and votes for candidate shift offsets, and the offset with the largest number of votes is selected as the recovered key. The recovery success rate retains the advantage of low computational overhead brought by integer operations. A complete theoretical framework for success rate analysis is established. Based on the Central Limit Theorem, closed-form expressions for single-column recovery success rates of the two schemes are rigorously derived. For finite sample cases, Gumbel extreme value correction is introduced to align theoretical curves with experimental results across the entire sample range. Systematic experimental validations are carried out. Extensive Monte Carlo simulations verify the accuracy of theoretical derivations.

II. Preliminaries

This section introduces the basic mathematical concepts and theorems used in the subsequent derivations of this paper.

A. Notation

For subsequent analysis, we define the following notation, as summarized in Table I.

TABLE I: Mathematical Notations

Symbol	Description
n	Bit length in operations
N	Number of plaintext samples
p	Probability that a ciphertext bit "0" remains "0" after noise perturbation
q	Probability that a ciphertext bit "1" remains "1" after noise perturbation
$\rho = p + q - 1$	Correlation factor
$D_{i,j}$	The Hamming distance between the i-th column of plaintext matrix and the j-th column of the noise-perturbed ciphertext matrix
$M \in \{0,1\}^{N\times n}$	Plaintext matrix, each bit follows a Bernoulli distribution with parameter $1/2$
$C' \in \{0,1\}^{N\times n}$	Noise-perturbed ciphertext matrix
k	Correct key
k_{rec}	Recovered key
$j_{\text{min}}^{(i)}(j_{\text{max}}^{(i)})$	index of ciphertext column with minimum (maximum) Hamming distance to plaintext column i

B. Fundamental Mathematical Concepts

Definition 1 (Transition Probability Matrix)[11] Let S be a state space consisting of n distinct states. Given that the system is currently in state s_i, the conditional probability of transitioning to state s_j at the next time step is denoted by a_{ij}:

$$a_{ij} = \Pr(X_{t+1} = s_j \mid X_t = s_i). \tag{1}$$

Organize the transition probabilities a_{ij} between all states in the system into an $n \times n$ square matrix (where n is the number of states). This matrix is called the transition probability matrix and is conventionally denoted by

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \tag{2}$$

where A has the following characteristics:

$$a_{ij} \geq 0, \tag{3}$$

$$\sum_{j=1}^{n} a_{ij} = 1. \tag{4}$$

Definition 2 (Binomial Distribution)[12] Let n denote the number of independent and identical Bernoulli trials, and let $p \in (0,1)$ be the success probability of each single trial. A random variable X, which counts the total number of successes among n trials, is said to follow the *binomial distribution* with parameters n and p, denoted by

$$X \sim B(n,p). \tag{5}$$

The probability mass function (PMF) of X is formulated as

$$\mathbb{P}(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\quad k=0,1,2,\ldots,n, \quad (6)$$

where

$$\binom{n}{k}=\frac{n!}{k!\,(n-k)!}, \quad (7)$$

is defined as the binomial coefficient.

The core statistical properties are listed below:

- Expectation: $\mathbb{E}[X]=np$.
- Variance: $\mathrm{Var}(X)=np(1-p)$.

Definition 3 (Standard Normal Distribution)[12] When $\mu=0$ and $\sigma^2=1$, the distribution is called the standard normal distribution, denoted $Z\sim\mathcal{N}(0,1)$. Its PDF is

$$\phi(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}},\quad z\in\mathbb{R}, \quad (8)$$

and its cumulative distribution function is denoted by $\Phi(z)=\mathbb{P}(Z\le z)$.

Definition 4 (Central Limit Theorem)[12] Let X_1,X_2 X_3, $\cdots,X_n$ be a sequence of independent and identically distributed random variables with finite expectation

$$\mu=\mathbb{E}[X_i], \quad (9)$$

and finite variance

$$\sigma^2=\mathrm{Var}(X_i)>0. \quad (10)$$

Denote the sample mean by

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i. \quad (11)$$

Then the standardized sample mean converges in distribution to the standard normal distribution as $n\to\infty$, i.e.,

$$\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\xrightarrow{d}\mathcal{N}(0,1). \quad (12)$$

Equivalently, the cumulative distribution function satisfies

$$\lim_{n\to\infty}\mathbb{P}\left(\frac{\sum_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}\le x\right)=\Phi(x),\quad \forall x\in\mathbb{R}, \quad (13)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

Definition 5 (Subgaussian Distribution)[13] If X has the distribution $\mathcal{N}(0,\sigma^2)$, then an easy computation shows that for any $t\in\mathbb{R}$,

$$\mathbb{E}e^{tX}=e^{\sigma^2t^2/2}. \quad (14)$$

Thus X is subgaussian with parameter σ.

Definition 6 (The Fisher–Tippett–Gnedenko theorem)[14] The Fisher–Tippett–Gnedenko theorem is the fundamental theorem of extreme value theory, which characterizes the asymptotic distribution of the extremes (maxima or minima) of a large number of random variables. Let $X_1,X_2,\ldots,X_n$ be independent and identically distributed random variables, and denote the sample maximum as $M_n=\max(X_1,\ldots,X_n)$. After proper linear normalization:

$$\frac{M_n-b_n}{a_n}\xrightarrow{d}G(x), \quad (15)$$

where $a_n>0$ and b_n are normalization constants, and $G(x)$ stands for the extreme value limit distribution. The theorem states that the limiting distribution of normalized sample maxima can only fall into three families: the Gumbel, Fréchet, or Weibull distributions, which can be unified under the generalized extreme value distribution.

Definition 7 (Gumbel Distribution)[15] The Gumbel distribution is a type of extreme value distribution. Its cumulative distribution function and probability density function are given by:

$$F(x;\mu,\beta)=e^{-e^{-(x-\mu)/\beta}}, \quad (16)$$

$$f(x;\mu,\beta)=\frac{1}{\beta}e^{-\left[e^{-(x-\mu)/\beta}+(x-\mu)/\beta\right]}. \quad (17)$$

For the standard Gumbel distribution with $\mu=0$ and $\beta=1$:

$$F(x)=e^{-e^{-x}},\quad f(x)=e^{-\left(x+e^{-x}\right)}. \quad (18)$$

Its expectation:

$$\mathbb{E}[x]=\mu+\gamma\cdot\beta, \quad (19)$$

where γ denotes the Euler–Mascheroni constant.

Definition 8(Mills' ratio)[16] Mills' ratio is defined as

$$R_x=e^{\frac{1}{2}x^2}\int_x^{\infty}e^{-\frac{1}{2}t^2}dt,\quad x\ge0, \quad (20)$$

which is also called failure rate for the standard normal law.

III. Construction of Cyclic Shift Perturbation Model and Attack Methods

Throughout this paper, we adopt the *Chosen-Plaintext Attack* (CPA) model: the cryptanalyst can select arbitrary plaintext vectors, construct a plaintext matrix M, and obtain the corresponding noise-perturbed ciphertext matrix C'. We assume that each bit of the plaintext matrix is independently and identically distributed (i.i.d.) according to a Bernoulli distribution with parameter $1/2$. Under this i.i.d. uniform plaintext assumption, the CPA model is equivalent to the *Known-Plaintext Attack* (KPA) model in terms of the statistical distributions of the observables, since the plaintext distribution is identical to that of randomly observed plaintexts. We therefore use the CPA terminology consistently throughout the paper. In this paper, we use right cyclic shift encryption as an illustrative example. For left cyclic shifts, only the offset calculation formula needs to be modified, while the single-anchor adaptive recovery and multi-column voting methods remain completely unchanged.

A. Construction of Cyclic Shift Perturbation Model

Under noisy conditions, the cryptanalysis is divided into two phases: cyclic Shift and perturbation.

Stage 1: Cyclic Shift Encryption We construct the plaintext matrix by taking N plaintext vectors with block length n. The plaintext matrix satisfies $M\in\{0,1\}^{N\times n}$, where

$$M=(m_1,m_2,\cdots,m_n), \quad (21)$$

and the key k of right cyclic shift encryption is defined as the shift offset k for right cyclic shifting of plaintext columns. We perform a right cyclic shift operation on each row of the

plaintext to obtain the original ciphertext matrix : $R_k(m_i) = c_{(i+k) \bmod n}$ (Left cyclic shift $L_k(m_i)$ can be defined similarly and it holds that $L_k(m_i) = R_{n-k}(m_i)$),

$$C_{r,j} \in \{0,1\}^{N \times n}, r = 1, 2, \cdots, N. \tag{22}$$

The core property of cyclic shift is that all plaintext columns are offset by the same number of positions as a whole, and thus the global shift amount is consistent.

Stage 2: Channel Noise Perturbation Each bit of the original ciphertext matrix passes independently through an asymmetric binary noise channel. Adversary finally observes the noisy ciphertext matrix C'. The flipping of each ciphertext bit follows the transition probability matrix

$$\begin{aligned} \boldsymbol{A} &= \begin{pmatrix} \Pr(c'=0 \mid c=0) & \Pr(c'=1 \mid c=0) \\ \Pr(c'=0 \mid c=1) & \Pr(c'=1 \mid c=1) \end{pmatrix} \\ &= \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}. \end{aligned} \tag{23}$$

In this transition probability matrix, p stands for the probability that a ciphertext bit equal to "0" stays "0" under noisy perturbation, while q stands for the probability that a ciphertext bit equal to "1" stays "1" under noisy perturbation. All bits are mutually independent.

For the scenario of key recovery, we utilize the differences in output features generated by key bits during perturbation under particular plaintexts, and then achieve the determination of key bits by comparing Hamming statistical distances.

This paper assumes that we are capable of mounting a chosen-plaintext attack: we can select plaintext vectors to construct a plaintext matrix and obtain the corresponding ciphertext matrix superimposed with perturbations. Complete knowledge of the cyclic shift encryption algorithm can be obtained yet remain completely unaware of the exact values of channel noise parameters p and q, which aligns with the constraints imposed by real-world dynamic scenarios. This paper aims to recover the correct cyclic shift key k based on observed perturbed ciphertext matrices, with the minimum feasible sample size and computational cost. We use right cyclic shift encryption as an illustrative example.

B. Cryptographic Analysis Methods

Consider two completely unrelated, statistically independent binary sequences compared bit by bit. The probability that a single bit position matches is $1/2$, and the probability of mismatch is also $1/2$.

Let $d_{i,j}^{(r)} = \mathbf{1}\{M_{r,i} \neq C'_{r,j}\} \in \{0,1\}$ denote the per-sample mismatch indicator between plaintext column i and ciphertext column j. The Hamming distance $D_{i,j}$ is the number of bit positions with mismatched bits, i.e., $D_{i,j} = \sum_{r=1}^{N} d_{i,j}^{(r)}$. The expectation of the per-sample mismatch indicator satisfies

$$\mathbb{E}[d_{i,j}^{(r)}] = \frac{1}{2}, \tag{24}$$

so the overall expectation yields:

$$\mathbb{E}[D_{i,j}] = \sum_{r=1}^{N} \mathbb{E}[d_{i,j}^{(r)}] = \frac{N}{2}. \tag{25}$$

The Hamming distance of unrelated random pairs fluctuates slightly around $N/2$.

Under the correct cyclic shift, plaintext columns and their corresponding ciphertext columns exhibit statistical correlation. These columns originate from an identical bit sequence that has only undergone cyclic shifting and channel noise perturbation. As a result, for correct cyclic shift, the expectation of the Hamming distance deviates markedly from the random baseline of $N/2$. In contrast, for incorrect cyclic shift, plaintext and ciphertext columns are derived from statistically independent bit sequences, where the expected Hamming distance equals exactly $N/2$.

Given that the noise direction is unknown, we cannot prejudge whether the correctly matched pairs correspond to the minimum or maximum Hamming distance. The core idea of this paper is as follows: by comparing the deviation magnitudes of the minimum-distance set and maximum-distance set relative to the random expectation, the scheme automatically identifies the noise direction. Then select the correct extremum side to locate the cyclic shift key, which completely removes the reliance on prior knowledge of noise statistics.

Meanwhile, leveraging the globally consistent property of cyclic shifts, multiple plaintext columns can be matched separately as independent anchor points. The voting mechanism is then adopted to aggregate all matching results, which further improves the robustness of key recovery.

IV. Single Anchor Point Adaptive Recovery Attack

The single anchor point scheme selects only one column from the plaintext matrix as the fixed anchor point. It completes shift offset estimation by computing the Hamming distance between this column and every column in the perturbed ciphertext matrix. This scheme features low computational and storage overhead, making it applicable to scenarios with extremely limited resources.

A. Construction of Hamming Distance Vector

Let the plaintext matrix be $M \in \{0,1\}^{N \times n}$, and the observed ciphertext matrix be $C' \in \{0,1\}^{N \times n}$. The adversary extracts the i-th column of the plaintext matrix as the anchor point. For each column $j = 1, \ldots, n$ of the observed ciphertext, compute the Hamming distance between the anchor column and this ciphertext column:

$$D_{i,j} = \sum_{r=1}^{N} \mathbf{1}\left\{M_{r,i} \neq C'_{r,j}\right\}, \quad i, j = 1, \ldots, n. \tag{26}$$

B. Adaptive Direction Judgment

Lemma 1 (Distribution for Correct Matching) When the j-th column of perturbed ciphertext is the shifted counterpart of the i-th plaintext column, the single-bit mismatch probability is:

$$p_d^* = \frac{1-\rho}{2}. \tag{27}$$

The total Hamming distance D_c follows a binomial distribution:

$$D_{\mathrm{c}} \sim \text{Binomial}\left(N,\ \frac{1-\rho}{2}\right). \tag{28}$$

Its expectation and variance are respectively:

$$\begin{aligned} \mathbb{E}\left[D_{\mathrm{c}}\right] &= N \cdot \frac{1-\rho}{2}, \\ \text{Var}\left(D_{\mathrm{c}}\right) &= \frac{N}{4}\left[1-\rho^2\right]. \end{aligned} \tag{29}$$

Proof: Under the correct pairing, the probability that plaintext bit 0 maps to perturbed ciphertext bit 0 is p, with mismatch probability $1-p$; the probability that plaintext bit 1 maps to perturbed ciphertext bit 1 is q, with mismatch probability $1-q$. Plaintext bits 0 and 1 each occur with probability $1/2$. Thus the average mismatch probability is:

$$\begin{aligned} p_d^* &= P(M_{r,i} \neq C'_{r,j}) \\ &= P(M_{r,i}=0)P(C'_{r,j}=1 \mid M_{r,i}=0) \\ &+ P(M_{r,i}=1)P(C'_{r,j}=0 \mid M_{r,i}=1) \\ &= \frac{1}{2}(1-p) + \frac{1}{2}(1-q) \\ &= \frac{2-p-q}{2} \\ &= \frac{1-\rho}{2}. \end{aligned} \tag{30}$$

The expectation and variance of the binomial distribution can be obtained by direct substitution into the formulas.

Lemma 2 (Distribution for Incorrect Matching) When the j-th column of the perturbed ciphertext and the i-th column of plaintext share no matching correspondence, the plaintext and ciphertext are statistically independent. The single-bit mismatch probability is a constant $1/2$, and the total Hamming distance D_w follows:

$$D_w \sim \text{Binomial}\left(N,\ \frac{1}{2}\right). \tag{31}$$

Its expectation and variance are respectively:

$$\begin{aligned} \mathbb{E}\left[D_w\right] &= \frac{N}{2}, \\ \text{Var}\left(D_w\right) &= \frac{N}{4}. \end{aligned} \tag{32}$$

Proof: When perturbed ciphertext bits and plaintext bits are mismatched in incorrect pairs, they are statistically independent.

$$\begin{aligned} P(C'_{r,j}=1) &= P(C_{r,j}=1)P(C'_{r,j}=1 \mid C_{r,j}=1) \\ &+ P(C_{r,j}=0)P(C'_{r,j}=1 \mid C_{r,j}=0) \\ &= \frac{1}{2}q + \frac{1}{2}(1-p) \\ &= \frac{1+q-p}{2}. \end{aligned} \tag{33}$$

Similarly,

$$P(C'_{r,j}=0) = \frac{1+p-q}{2}. \tag{34}$$

The probability of mismatch can be calculated and verified via the law of total probability.

$$\begin{aligned} p_d &= P(M_{r,i} \neq C'_{r,j}) \\ &= P(M_{r,i}=0)P(C'_{r,j}=1) \\ &+ P(M_{r,i}=1)P(C'_{r,j}=0) \\ &= \frac{1}{2}(\frac{1+q-p}{2}) + \frac{1}{2}(\frac{1+p-q}{2}) \\ &= \frac{1}{2}. \end{aligned} \tag{35}$$

The expectation and variance of the binomial distribution can be obtained by direct substitution into the formulas.

Define the correlation factor as $\rho = p+q-1$. From Lemma 1, its physical interpretation is given as follows:

- When $\rho > 0$ (i.e., $p+q > 1$), $\mathbb{E}[D_c] < N/2$. The Hamming distance of correctly matched pairs is smaller than the random expectation;
- When $\rho < 0$ (i.e., $p+q < 1$), $\mathbb{E}[D_c] > N/2$. The Hamming distance of correctly matched pairs is larger than the random expectation;
- When $\rho = 0$ (i.e., $p+q = 1$), $\mathbb{E}[D_c] = N/2$. Correct pairs and mismatched pairs are statistically indistinguishable.

Based on the above analysis, without knowing the specific values of p and q, we can automatically judge the noise direction by comparing the respective deviation magnitudes of the minimum and maximum values in the distance vector from the random expectation $N/2$:

$$\begin{aligned} \text{dev}_{\min} &= \left|\min(D_{i,j}) - \frac{N}{2}\right|, \\ \text{dev}_{\max} &= \left|\max(D_{i,j}) - \frac{N}{2}\right|. \end{aligned} \tag{36}$$

- If $\text{dev}_{\min} > \text{dev}_{\max}$, the minimum-value side has a larger deviation, i.e., $\rho > 0$. The correctly matched pairs correspond to the minimum Hamming distance.
- If $\text{dev}_{\max} > \text{dev}_{\min}$, the maximum-value side deviates further, i.e., $\rho < 0$. The correctly matched pairs correspond to the maximum Hamming distance.

C. Key Recovery Attack Procedure

Based on the above analysis, the pseudocode for the key recovery algorithm is presented below.

This algorithm only requires one single traversal to compute the Hamming distance, with a time complexity of $O(Nn)$ and a space complexity of $O(N)$. It involves no floating-point arithmetic or complex matrix operations, leading to an extremely low hardware implementation cost.

D. Success Rate Analysis of Key Recovery

The recovery success of the fully adaptive scheme has two ingredients: the correct column must lie on the correct side of $N/2$, and the rule (36) must select that side.

Lemma 3 (Asymptotic Distribution of Difference Statistic) Define the difference statistic

$$\Delta = D_c - D_w, \tag{37}$$

Algorithm 1 Single Anchor Adaptive Cyclic Shift Recovery Algorithm

Require: Sample length N, block size n, noisy ciphertext matrix C', anchor plaintext column m_i
Ensure: Estimated shift offset k_{rec}
1: Extract the i-th column of plaintext as anchor
2: **for** $j = 0$ to $n-1$ **do**
3: $\quad D_{i,j} \leftarrow \sum_{r=1}^{N} \mathbf{1}\left\{M_{r,i} \neq C'_{r,j}\right\}$
4: **end for**
5: Compute minimum distance index and maximum distance index

$$j_{\min}^{(i)} \leftarrow \arg\min D_{i,j},$$
$$j_{\max}^{(i)} \leftarrow \arg\max D_{i,j}$$

6: Compute deviation magnitudes

$$\text{dev}_{\min} \leftarrow \left| D_{i,j_{min}} - \frac{N}{2} \right|,\ \text{dev}_{\max} \leftarrow \left| D_{i,j_{max}} - \frac{N}{2} \right|$$

7: **if** $\text{dev}_{\min} > \text{dev}_{\max}$ **then**
8: $\quad k_{\text{rec}} \leftarrow (j_{\min} - i) \bmod n$

Judge $\rho > 0$, take the minimum side

9: **else**
10: $\quad k_{\text{rec}} \leftarrow (j_{\max} - i) \bmod n$

Judge $\rho < 0$, take the maximum side

11: **end if**
12: **return** k_{rec}

where D_c denotes the Hamming distance of correctly matched pairs and D_w denotes the Hamming distance of any mismatched pair. For asymptotic analysis we use $\mathbb{E}[D_w] = N/2$ and treat the extremal fluctuation of the $n-1$ incorrect candidates separately (see Section III-E). Then we have:

$$\begin{aligned} \mathbb{E}[\Delta] &= -\frac{N\rho}{2}, \\ \text{Var}(\Delta) &= \frac{N}{4}\left[2 - \rho^2 - 2(p-q)^2\right]. \end{aligned} \tag{38}$$

Moreover, as $N \to \infty$, Δ converges in distribution to a normal distribution.

Proof: The expectations and variances of D_c and D_w can be obtained from Lemma 1, Lemma 2.

$$\begin{aligned} D_c &= \sum_{r=1}^{N} \mathbf{1}\{M_{r,i} \neq C'_{r,j^*}\}, \\ D_w &= \sum_{r=1}^{N} \mathbf{1}\{M_{r,i} \neq C'_{r,j}\}. \end{aligned} \tag{39}$$

For the correctly matched column, C'_{r,j^*} is the noisy version of the anchor bit $M_{r,i}$ itself; for an incorrectly matched column, $C'_{r,j}$ is the noisy version of an independent plaintext bit. Since the N rows are independent, the total covariance is N times the per-row covariance.

Fix a row r, conditional on $M_{r,i} = 0$,

$$\begin{aligned} &\Pr(M_{r,i} \neq C'_{r,j^*}) = 1 - p, \\ &\Pr(M_{r,i} \neq C'_{r,j}) = (1 - p + q)/2; \end{aligned} \tag{40}$$

conditional on $M_{r,i} = 1$,

$$\begin{aligned} &\Pr(M_{r,i} \neq C'_{r,j^*}) = 1 - q, \\ &\Pr(M_{r,i} \neq C'_{r,j}) = (1 + p - q)/2; \end{aligned} \tag{41}$$

Hence Write

$$I_c = \mathbf{1}\{M_{r,i} \neq C'_{r,j^*}\}, I_w = \mathbf{1}\{M_{r,i} \neq C'_{r,j}\}. \tag{42}$$

Then

$$\begin{aligned} \mathbb{E}[I_c I_w] &= \Pr(M_{r,i} = 0)\,\mathbb{E}[I_c I_w \mid M_{r,i} = 0] \\ &\quad + \Pr(M_{r,i} = 1)\,\mathbb{E}[I_c I_w \mid M_{r,i} = 1] \\ &= \tfrac{1}{4}\left[(1-p)(1-p+q) + (1-q)(p+1-q)\right] \\ &= \frac{2 - (p+q) + (p-q)^2}{4}. \end{aligned} \tag{43}$$

Since

$$\mathbb{E}[I_c] = (1-\rho)/2, \mathbb{E}[I_w] = 1/2, \tag{44}$$

$$\text{Cov}(I_c, I_w) = \mathbb{E}[I_c I_w] - \mathbb{E}[I_c]\,\mathbb{E}[I_w]. \tag{45}$$

The per-row covariance is $(p-q)^2/4$. As the N rows are independent, the total covariance is N times this value, yielding (46). The covariance between them arises from the shared plaintext column, and calculation yields

$$\text{Cov}(D_c, D_w) = N(p-q)^2/4. \tag{46}$$

Substituting this into the variance formula, we obtain $\text{Var}(\Delta)$. The asymptotic normality of Δ holds by the Central Limit Theorem.

Theorem 1 (Recovery success probability conditional on the judged direction) Under the condition that the adaptive rule (36) selects the correct side, the recovery success probability of the single-column anchor scheme satisfies:

$$P_{\text{single}} = \Phi(Z) + o(1), \tag{47}$$

where

$$Z = \frac{|\rho|\sqrt{N}}{\sqrt{2 - \rho^2 - 2(p-q)^2}}. \tag{48}$$

$\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, and $o(1)$ represents an infinitesimal of higher order that tends to 0 as $N \to \infty$.

Proof: When $\rho < 0$, we define the success event as the Hamming distance of the correct matching being larger than the maximum Hamming distance of all incorrect matchings.

For $\rho < 0$, the condition $D_c > \max D_{i,j},\ c_j \neq R_k(m_i)$,

$$\begin{aligned} D_c - \max D_{i,j} &= D_c - \frac{N}{2} - \left(\max D_{i,j} - \frac{N}{2}\right) \\ &= \Delta - \left(\max D_{i,j} - \frac{N}{2}\right). \end{aligned} \tag{49}$$

From the property of sub-Gaussian distributions, we obtain that the order of magnitude of the extreme value fluctuation for the $n-1$ incorrect candidates is $O_p\big(\sqrt{N \ln n}\big)$ under large-sample conditions. While the expected shift of the correct pair is of order $O(N)$, where the former is asymptotically far smaller than the latter. Therefore, the extreme-value competition term degenerates into a higher-order infinitesimal, and the success probability is dominated by the differential statistic condition $\Delta > 0$ (or $\Delta < 0$, depending on the direction).

$$\begin{aligned} P(\Delta > 0) &= \Phi\left(\frac{\Delta - \mathbb{E}(\Delta)}{\sqrt{\mathrm{Var}(\Delta)}} > \frac{-\mathbb{E}(\Delta)}{\sqrt{\mathrm{Var}(\Delta)}}\right) \\ &= \Phi\left(\frac{\mathbb{E}(\Delta)}{\sqrt{\mathrm{Var}(\Delta)}}\right). \end{aligned} \tag{50}$$

Thus

$$\begin{aligned} P\big(D_c - \max D_{ij} > 0\big) &= P(\Delta > 0) + o(1) \\ &= \Phi\left(\frac{\mathbb{E}(\Delta)}{\sqrt{\mathrm{Var}(\Delta)}}\right) + o(1). \end{aligned} \tag{51}$$

Similarly, When $\rho > 0$,we define the success event as the Hamming distance of the correct matching being smaller than the minimum Hamming distance of all incorrect matchings. The success probability is dominated by

$$P\big(D_c - \min D_{ij} < 0\big) = \Phi\left(\frac{-\mathbb{E}(\Delta)}{\sqrt{\mathrm{Var}(\Delta)}}\right) + o(1). \tag{52}$$

Substitute the expectation and variance derived in Lemma 3 into the expression, define $Z = \frac{|\rho|\sqrt{N}}{\sqrt{2-\rho^2-2(p-q)^2}}$, which yields $P_{\text{single}} = \Phi(Z) + o(1)$. When the discriminant signal dominates the Gumbel-scale fluctuation, i.e. $|\rho|\sqrt{N} \gg \sqrt{\ln(n-1)}$, the side selected by (36) coincides with the side containing the correct column, so the closed form above also describes the fully adaptive Algorithm 1; the Monte Carlo curves in Figs. 2–7 confirm this.

E. Extreme-Value Correction from the Incorrect Candidates

The conclusion of Theorem 1 presumes that the correct column is distinguished from the incorrect ones by its mean deviation alone. In fact, the scheme selects the minimum (or maximum) among $n-1$ incorrect Hamming distances, and the extremum of $n-1$ identically distributed variables is itself subject to a Gumbel-type fluctuation, this fluctuation originates from the number of competing incorrect columns $n-1$.

The $n-1$ incorrect distances are not fully independent: they share the anchor column , and by the same shared-anchor argument as in (46), so any two incorrect distances carry correlation $(p-q)^2$. Conditional on the anchor column, however, the $n-1$ distances are independent, since their underlying plaintext columns and channel-noise bits are mutually independent. This weak equicorrelation modifies only the extremal index at order $(p-q)^2$ and does not change the leading Gumbel normalization $\sqrt{2\ln(n-1)}$; it vanishes entirely when $p = q$.

Using the Gumbel extreme value distribution approximation and Mills' ratio, the expectation of the extreme Hamming distance among $n-1$ identically distributed incorrect pairs exhibits a statistical shift of order $\sqrt{2\ln(n-1)}$ relative to the mean $N/2$. The correction arises from the maximum of $n-1$ approximately standard normal incorrect distances. For m i.i.d. standard normal variables, the maximum converges after standardization to a Gumbel variable, and its leading fluctuation is $\sqrt{2\ln m}$ (Fisher–Tippett–Gnedenko; see also Def. 6). Replacing $m = n-1$ and rescaling by the standard deviation of Δ, the largest incorrect deviation contributes $\sqrt{2\ln(n-1)}/\sqrt{2-\rho^2-2(p-q)^2}$ in the Z-scale, which is exactly the subtraction in (53).

Incorporating this shift into the effective signal-to-noise ratio yields the corrected closed-form expression:

$$Z' = Z - \frac{\sqrt{2\ln(n-1)}}{\sqrt{2-\rho^2-2(p-q)^2}}, \tag{53}$$

$$P'_{\text{single}} \approx \Phi(Z'). \tag{54}$$

This corrected formula accounts for the extremum selected among the $n-1$ incorrect candidates. When the discriminant signal $|\rho|\sqrt{N}$ dominates the Gumbel-scale fluctuation $\sqrt{2\ln(n-1)}$, the correction becomes negligible and

$$P'_{\text{single}} \to \Phi(Z). \tag{55}$$

V. Multi-Column Voting Adaptive Recovery Attack

The single-column anchor attack only exploits information of N bits. When the sample size N is small or the correlation factor $|\rho|$ approaches 0, the success rate of the single-column scheme remains low. A core characteristic of cyclic shifting is that all columns share an identical shift offset k. Hence every plaintext column can participate in the "guessing" procedure: each plaintext column independently searches for its matching column in the observed ciphertext matrix by the methods of single-anchor recovery attack, computes the candidate shift offset k_{rec}, and casts a corresponding vote. All correctly matched columns converge to the genuine shift offset, resulting in concentrated votes. In contrast, mismatched columns randomly point to arbitrary candidate offsets, leading to dispersed votes.

A. Key Recovery Analysis

After the i-th column of plaintext undergoes a right cyclic shift by k bits, it appears at the $(i+k) \bmod n$-th column of ciphertext.

Therefore, if the i-th plaintext column regards the j-th ciphertext column as its matching position, the candidate shift offset it deduces is given by:

$$k_{\text{rec}} = (j - i) \bmod n. \tag{56}$$

Based on the above mapping rule, the two votes cast by the i-th column are defined as follows.

Vote for the minimum-distance case:

$$k_{\min}^{(i)} = \big(j_{\min}^{(i)} - i\big) \bmod n. \tag{57}$$

Vote for the maximum-distance case:

$$k_{\max}^{(i)} = \left(j_{\max}^{(i)} - i\right) \bmod n. \tag{58}$$

Votes cast by all columns of the plaintext are accumulated into two separate vote bins:

- Minimum-distance vote bin min_votes$[k_{\min}]$: Counts the number of "minimum-distance-side" votes cast for candidate shift offset k_{rec}.
- Maximum-distance vote bin max_votes$[k_{\max}]$: Counts the number of "maximum-distance-side" votes cast for candidate shift offset k_{rec}.

Similar to the single-column anchor model, the multi-column voting scheme also requires judging the noise direction, determining whether the minimum-distance side or the maximum-distance side should be trusted. The judgment criterion states that the side with more concentrated votes is the correct direction.

- Under random conditions, n columns cast votes across n candidates, and each candidate has an expected vote count of $n/n = 1$, leading to dispersed vote distribution.
- In the correct direction, all valid columns cast votes for the identical genuine shift offset k_{rec}, resulting in highly concentrated votes (the peak value is close to the total number of valid columns).
- In the incorrect direction, votes from mismatched columns scatter randomly, and the peak value is far lower than that of the correct direction.

B. Key Recovery Attack Procedure

Based on the above principles, the complete workflow of the multi-column voting robust recovery algorithm is shown in Algorithm 2.

Algorithm 2 Multi-column Voting Key Recovery

Require: Sample length N, block size n, noisy ciphertext matrix $\boldsymbol{C'}$, plaintext matrix $\boldsymbol{M}$
Ensure: Estimated shift offset k_{rec}

Initialize minimum-side vote vector min_votes $\leftarrow \mathbf{0}_{1\times n}$, maximum-side vote vector max_votes $\leftarrow \mathbf{0}_{1\times n}$
for $i = 0$ to $n-1$ **do**
  $D_{i,j} \leftarrow \sum_{r=1}^{N} \mathbf{1}\left\{M_{r,i} \neq C'_{r,j}\right\}, j = 0, 1, \ldots, n-1$
  Hamming-distance vector: i-th plaintext col. vs. all ciphertext cols.
  $j_{\min}^{(i)} \leftarrow \arg\min_j D_{i,j}, \quad j_{\max}^{(i)} \leftarrow \arg\max_j D_{i,j}$
  $k_{\min}^{(i)} \leftarrow \left(j_{\min}^{(i)} - i\right) \bmod n$, $k_{\max}^{(i)} \leftarrow \left(j_{\max}^{(i)} - i\right) \bmod n$
  min_votes$[k_{\min}^{(i)}]$ += 1, max_votes$[k_{\max}^{(i)}]$ += 1
end for
$p_{\min} \leftarrow \max(\text{min_votes}), \ p_{\max} \leftarrow \max(\text{max_votes})$
if $p_{\min} > p_{\max}$ **then**
  $k_{\text{rec}} \leftarrow \arg max(\text{min_votes})$
else
  $k_{\text{rec}} \leftarrow \arg max(\text{max_votes})$
end if
return k_{rec}

Time Complexity: The algorithm computes the Hamming distance between each of the n plaintext columns and all n ciphertext columns. Each distance computation traverses N bits. Thus the overall time complexity is $O(Nn^2)$.

C. Success Rate Analysis of Key Recovery

Theorem 2(Vote Distribution Characteristics) Assume the matching events of all columns are mutually independent. The probability that a single column achieves successful matching (i.e., casting a vote for the genuine shift offset k) is $P_c = P_{\text{single}} \approx \Phi(Z')$. Let T denote the total number of votes accumulated by the genuine shift offset. Then T follows a binomial distribution with parameters (n, P_c):

$$T \sim \text{Binomial}(n, P_c). \tag{59}$$

The probability that exactly t columns are matched correctly is given by:

$$\Pr(T = t) = \binom{n}{t} P_c^t (1 - P_c)^{n-t}. \tag{60}$$

remark 1 (On the independence of the per-column voting process). Theorem 2 treats the n per-column matching events underlying (59) as mutually independent, although every plaintext column is compared against the same ciphertext matrix. This is justified by conditioning on C': once the ciphertext matrix is fixed, the Hamming distances computed from each plaintext column depend only on that column, and the plaintext columns are mutually independent under the chosen-plaintext model. Hence the per-column voting events are independent conditionally on C'. The only unconditional coupling is the random channel noise realized in C', which perturbs all columns' statistics by a common random shift; since the channel asymmetry enters only through $p - q$, the induced correlation is of order $(p-q)^2$ rather than $O(1)$. A direct measurement at $n = 16$, $N = 20$, $p = 0.4$, $q = 0.3$ gives an empirical correlation of only 0.012. To further examine violation, we let plaintext bits share a row-common bit, inducing equicorrelation γ between columns; Fig. 1 shows that as γ grows from 0 to 0.5, the success rate shifts only mildly downward and all curves approach 1 as N grows. The binomial law (59) thus remains a valid leading-order approximation.

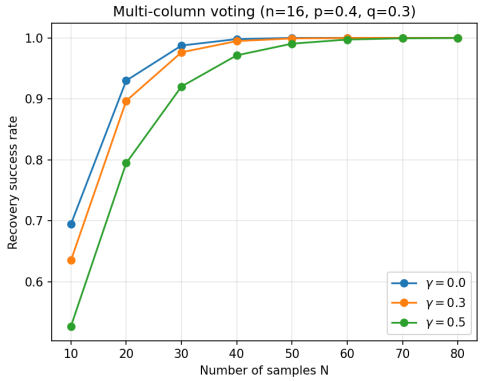


Fig. 1: Multi-column voting success rate under plaintext-column equicorrelation.

If a single column fails to match correctly, its vote is uniformly and randomly assigned to the remaining $n-1$ incorrect shift candidates, and no statistical bias exists in the vote count of any erroneous candidate offset.

1) Strict Conservative Lower Bound

The worst-case interference scenario: all erroneous votes are concentrated on a single incorrect shift offset. Under this scenario, the incorrect offset can receive at most $n-t$ votes. For the genuine shift offset to guarantee victory, its vote count must exceed half of all votes:

$$t > n-t \iff t > \frac{n}{2}. \tag{61}$$

As long as the number of correctly matched columns exceeds $\lfloor n/2 \rfloor + 1$, the genuine shift offset will inevitably obtain the maximum vote count regardless of how erroneous votes aggregate, which guarantees a successful attack. Therefore, the global success probability of key recovery satisfies the strict lower bound:

$$P_{\text{vote}} \geq \sum_{t=\lfloor n/2 \rfloor+1}^{n} \binom{n}{t} P_c^t (1-P_c)^{n-t}. \tag{62}$$

2) Approximate Closed-Form for Success Probability

In practical scenarios, the vote cast by a mismatched column is equally likely to fall on every incorrect key, uniformly distributed across the $n-1$ candidate offsets; the probability that all erroneous votes converge to a single shift offset is extremely low. The dominant event leading to attack failure is the case where all columns produce incorrect matches ($T=0$). Under this condition, the genuine shift offset receives zero votes, which inevitably results in recovery failure.

Thus the global failure probability can be approximated as the probability that every column fails to match correctly:

$$P_{\text{fail}} \approx (1-P_c)^n. \tag{63}$$

The corresponding approximate closed-form expression for the success probability is:

$$P_{\text{vote}} \approx 1-(1-P_c)^n. \tag{64}$$

When $n=1$, the multi-column voting scheme degenerates into the single-column anchor model, satisfying $P_{\text{vote}} = P_c = \Phi(Z')$, and the two schemes yield identical results.

VI. Experimental Verification

A. Experimentation Analysis of Single Anchor Point Adaptive Recovery Attack

To verify the correctness of the theoretical model for single anchor point adaptive recovery attack success rate derived in this paper, we conduct three sets of Monte Carlo simulations for the single anchor point adaptive recovery attack with block length $n=16$ in Figs. 2-4. The horizontal axis denotes the number of samples N and the vertical axis represents the key recovery success rate. The blue solid lines stand for the predicted values from the extremum-corrected closed-form theoretical model, and the red dashed lines correspond to experimental simulation results.

For the two parameter configurations (p=0.4, q=0.3) and (p=0.8, q=0.7), the experimental curves lie slightly above the theoretical curves in the small-sample region; as the sample size increases, the experimental and theoretical curves gradually coincide and finally converge to 100%. In contrast, the setting (p=0.7, q=0.3) corresponds to $\rho=0$. In this case, the theoretical success rate is fixed at the random-guess baseline of 6.25%, and the experimental curve fluctuates slightly around this constant value without growing as the sample size rises.

The three sets of results jointly verify that the extremum-corrected closed-form formula derived in this paper can accurately characterize the relationship between sample size and recovery success rate under various noisy conditions, and also validate the boundary conclusion that the attack completely fails when $\rho=0$.

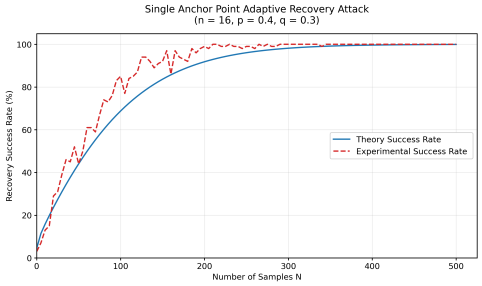


Fig. 2: $p=0.4, q=0.3$.

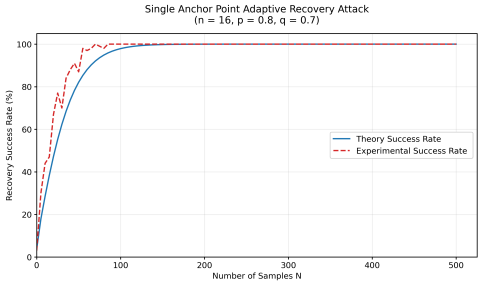


Fig. 3: $p=0.8, q=0.7$.

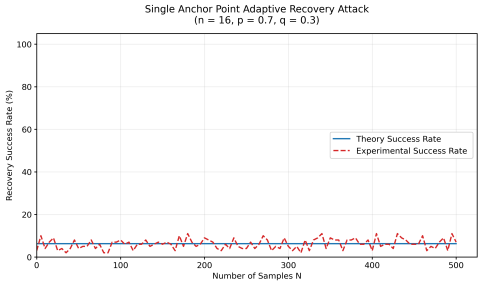


Fig. 4: $p=0.7, q=0.3$.

B. Experimentation Analysis of Multi-Column Voting Adaptive Recovery Attack

We conduct three sets of Monte Carlo simulations for the multi-column voting adaptive recovery attack with block length $n=16$. The horizontal axis denotes the number of samples N and the vertical axis represents the key recovery success rate. The blue solid lines correspond to the predicted

success rates derived from the extremum-corrected closed-form theoretical formula, and the red dashed lines stand for the average experimental results from repeated independent simulations in Figs. 5-7.

For the two distinguishable channel configurations (p=0.4, q=0.3) and (p=0.8, q=0.7) satisfying $\rho = p + q - 1 \neq 0$, the experimental curves lie slightly above the theoretical curves in the small-sample region; as the sample size N increases, the experimental and theoretical curves gradually overlap and eventually converge to 100%, and a larger absolute value of ρ enables faster convergence to full recovery.

In contrast, the parameter setting (p=0.7, q=0.3) yields $\rho = 0$, which means the correct key cannot be statistically distinguished under such a noisy channel. In this case, the theoretical success rate is fixed at the random-guess baseline of 6.25%, and the experimental curve fluctuates mildly around this constant value without increasing with growing sample size. The three sets of curves collectively verify that our proposed approximate closed-form expression for success probability can accurately characterize the relationship between sample size and recovery performance under asymmetric noisy channels, and also validate the theoretical boundary that the attack degenerates into random guessing with no effective discriminant information when $\rho = 0$.

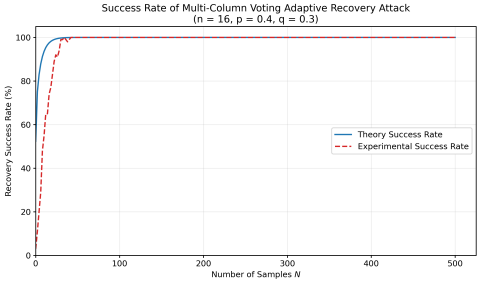


Fig. 5: $p = 0.4, q = 0.3$.

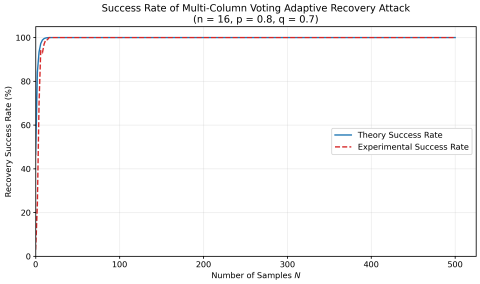


Fig. 6: $p = 0.8, q = 0.7$.

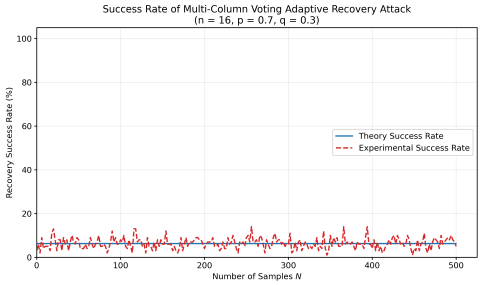


Fig. 7: $p = 0.7, q = 0.3$.

C. Comparison with Existing Methods

Compared with the existing correlation coefficient method proposed by Li [10], the adaptive recovery attack method combining single anchor point and multi-column voting proposed in this paper has the following advantages:

1) The correlation coefficient only characterizes the degree of linear correlation between two sequences and cannot identify the perturbation direction. In contrast, the adaptive recovery scheme based on multi-column voting proposed in this work is capable of recovering the genuine key. Furthermore, it can identify the noise direction from voting statistics and distinguish between strong and weak perturbation scenarios.
2) The core bottleneck of the correlation coefficient method lies in the repeated computation of floating-point statistics. The covariance and standard deviation of the entire sequence must be recalculated every time a candidate shift magnitude is traversed. By comparison, the proposed method only requires computing the Hamming distance for all candidate shifts, involving merely integer XOR operations and maximum value comparison. Its computational overhead is far lower than that of the floating-point statistical scheme, making it more adaptable to long-sequence and multi-round shift key recovery scenarios.

It is worth noting that the $\rho = 0$ boundary is fundamental rather than a limitation of the Hamming-distance statistic. When $p = q$, the channel is symmetric: each ciphertext bit is flipped with the same probability regardless of whether the original plaintext bit is zero or one. Under the i.i.d. plaintext model, the resulting ciphertext entries are therefore indistinguishable from independent fair random bits and carry no information about the unknown shift at all. No more sophisticated statistic or alternative distinguisher can break this symmetry; the random-guessing floor is an information-theoretic limit.

VII. Conclusions

To address the issue that cyclic shift recovery relies on prior knowledge of noise under uncertain ciphertext scenarios, this paper proposes an adaptive Hamming voting recovery attack. An architecture combining lightweight single-anchor processing and voting of multiple anchors is constructed, which eliminates the requirement for prior knowledge of noise parameters entirely.

The theoretical framework of this work is established as follows. Based on the asymmetric binary channel model, we introduce covariance correction for shared plaintext columns. The Gumbel extreme value distribution is adopted to compensate for the multi-candidate competition effect for finite samples, yielding the closed-form expression for the single-column recovery success rate. Furthermore, rigorous conservative lower bounds and closed-form approximations for multi-column voting are derived via the binomial distribution, forming a complete theoretical system. Experiments demonstrate that the theoretical predictions are highly consistent with the Monte Carlo simulation results.

This paper isolates a single cyclic shift under a memoryless asymmetric binary channel; two concrete generalizations follow directly.

1) The derivations above rely on independent bit flips, whereas ionospheric scintillation and multipath fading produce contiguous bursts of corrupted bits rather than scattered independent errors. A natural model is the two-state Gilbert–Elliott channel, in which a hidden good/bad Markov state modulates the flip probability. Under such a channel, the per-bit mismatch probabilities in Lemmas 1–2 and the shared-column covariance (46) must be re-derived with state persistence, and the Gumbel normalization $\sqrt{2\ln(n-1)}$ in (53) must be recomputed from the effective number of independent trials rather than N itself. A specific open question is how the effective independent sample size decays as burst length grows, and whether a run-length-weighted Hamming statistic can recover the lost discriminant signal.
2) The target primitive here is a single n-bit shift, whereas real ciphers embed shifts inside larger linear layers—byte-wise ShiftRows in AES, the fixed 32-bit rotations composing the SM4 linear transformation L, and the 28-bit rotations in the DES round-key schedule. Extending the framework means handling noise accumulation across rounds and its interleaving with nonlinear S-box layers; the clean column-to-column correspondence exploited in Algorithm 2 then becomes a multi-round statistical trail rather than a direct match. Combining the integer-voting distinguisher derived here with truncated-differential or linear approximations is the natural next step.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62672293, 62572371, 62572221, 62372076) and the Natural Science Foundation of Shaanxi Province, China (Grant No. 2024JC-YBMS-543).

REFERENCES

[1] P. Yue, J. An, J. Zhang, J. Ye, G. Pan, S. Wang, P. Xiao and L. Hanzo, Low Earth Orbit Satellite Security and Reliability: Issues, Solutions, and the Road Ahead, in *IEEE Communications Surveys and Tutorials*, vol. 25, no. 3, pp. 1604–1652, 2023.

[2] V. Agate, F. Concone, A. De Paola, P. Ferraro, G. L. Re and M. Morana, Bayesian Modeling for Differential Cryptanalysis of Block Ciphers: A DES Instance, in *IEEE Access*, vol. 11, pp. 4809–4820, 2023.

[3] J. Daemen, *The Design of Rijndael: AES–the Advanced Encryption Standard*, Springer, Berlin, 2002.

[4] Y. Zhu, Y. Zhou, Y. Li and Z. Wu, Differential Cryptanalysis of the Improved-DLBCA Lightweight Block Cipher, in *Journal of Networking and Network Applications*, vol. 5, no. 4, pp. 158–164, Dec. 2025.

[5] R. Feng, Y. Xi, H. Liu and L. Lu, SAT-Based Differential Analysis of the Feistel Block Cipher GRANULE, in *Journal of Networking and Network Applications*, vol. 5, no. 3, pp. 130–136, Oct. 2025.

[6] E. Biham and A. Shamir, Differential Cryptanalysis of DES-like Cryptosystems, in *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[7] M. Matsui, Linear Cryptanalysis Method for DES Cipher, in *Proceedings of Advances in Cryptology – EUROCRYPT 1993*, vol. 765, pp. 386–397, Springer, Berlin, Heidelberg, 1994.

[8] Y. Xi, R. Feng, Y. Zhou, H. Liu and L. Lu, Analysis Methods for Block Ciphers Under Noise Interference: A Case Study of XOR Operations, in *2025 International Conference on Satellite Internet (SAT-NET)*, pp. 67–72, 2025.

[9] J. Zhou, Y. Zhu, M. Tian, L. Lu and Y. Zhou, Cryptanalysis of Block Ciphers with Noise: A Case Study on Permutation Operations, in *2025 International Conference on Satellite Internet (SAT-NET)*, Hangzhou, China, pp. 19–24, 2025.

[10] J. Li, Z. Yang, Y. Zhou and L. Lu, Cryptanalysis of Cyclic Shift Block Ciphers in Noisy Environment, in *2026 International Conference on Networking and Network Applications (NaNA)*, accepted.

[11] Y. Zhou, W. Wang, J. Yan, Z. Wu and L. Lu, Optimization of Binary Randomized Response Based on Lanke Privacy and Utility Analysis, in *Chinese Journal of Electronics*, vol. 34, no. 1, pp. 242–256, 2025.

[12] Z. Sheng, S. Xie and C. Pan, *Probability Theory and Mathematical Statistics* (in Chinese), 4th ed., Higher Education Press, Beijing, 2008.

[13] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, 2nd ed., Cambridge University Press, 2026.

[14] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events: for Insurance and Finance*, vol. 33, *Stochastic Modelling and Applied Probability*, Springer Science & Business Media, 1997.

[15] L. L. Fan and Y. Y. Wang, Parameter estimation and case analysis of generalized extreme value distribution, in *Journal of Capital Normal University (Natural Science Edition)*, vol. 38, no. 3, pp. 13–18, 2017.

[16] D. Lu, L. Song and G. Tang, Some new approximations and proofs for Mills' ratio, in *Results in Mathematics*, vol. 73, no. 1, p. 27, 2018.



Xu Liu was born in 2003. She is currently pursuing her master's degree at the School of Mathematics and Statistics, Shaanxi Normal University. Her research interests lie in cryptography and information security.

Hongmin Liu received the M.S. degree in computer application technology from North University of China, Taiyuan, China in 2007. Her research interests include computer application and computer network security

Yihui Zhou received the B.S. degree in mathematics and applied mathematics, the M.S. degree in fundamental mathematics, and the Ph.D. degree in fundamental mathematics from Shaanxi Normal University, Xi'an, China, in 2003, 2006, and 2009, respectively. She is currently a Lecturer of the School of Artificial Intelligence and Computer Science, Shaanxi Normal University, Xi'an, China. Her research interests include information security and privacy preserving.

Laifeng Lu received the Ph.D. degree from Xidian University, Xi'an, China, in 2012. She is currently a Professor of the School of Mathematics and Statistics, Shaanxi Normal University. Her research interests include cryptography and information security .