Exploring Key-Recovery-Friendly Differential Distinguishers for SM4 and Their Performance in Differential Attacks (Full Version)

Bingqing Li[1,2,3] and Ling Sun(✉)[1,2,3,4]

[1] State Key Laboratory of Cryptography and Digital Economy Security, Shandong University, Qingdao, China
[2] Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan, China
[3] School of Cyber Science and Technology, Shandong University, Qingdao, China
[4] Quan Cheng Shandong Laboratory, Jinan, China
lingsun@sdu.edu.cn

Abstract. In this paper, we focus on SM4, a widely used and standardized Chinese block cipher. After revisiting the previously proposed optimal 19-round differential characteristic, we observe that its applicability in differential attacks is limited by a reduced pre-sieving probability, causing the time complexity to exceed that of brute force. To overcome this issue, we employ an automated search approach to identify more promising optimal 19-round differential characteristics. By translating key properties relevant to key recovery into Boolean expressions, we uncover three structural properties common to all optimal 19-round characteristics. While these properties dictate the overall probability of the resulting 19-round distinguishers, their varying pre-sieving probabilities influence their practical effectiveness in differential attacks. Using Boolean encodings, we identify four representative key-recovery-friendly differential characteristics. We then conduct an in-depth analysis of one such characteristic and demonstrate that, when evaluated under both the hypothesis testing paradigm and the key ranking paradigm, the proposed attack requires slightly more data than existing 23-round attacks. Nonetheless, it achieves lower time and memory complexities and ensures a higher success probability, offering a valuable new avenue for differential cryptanalysis of SM4. We believe our findings enhance the understanding of SM4's differential structure and provide a solid foundation for future research on advanced key-recovery techniques that leverage these newly identified structural properties and differential characteristics.

Keywords: Differential cryptanalysis · Distinguisher · Characteristic · Key-recovery-friendly · SM4.

1 Introduction

SM4 [7] is a block cipher standard widely used in China to safeguard sensitive data. Developed by the Chinese government, it forms a core component of

the nation's cryptographic standards. Originally introduced as SMS4, the algorithm was standardized by the Chinese State Cryptography Administration and later renamed SM4, aligning it with the SM series of cryptographic algorithms. Since its inception, SM4 has been subjected to extensive cryptanalysis, with researchers examining its resilience against a wide range of attacks — most notably differential attacks [9,24,25,19], as well as linear [8,9] and integral attacks [12], among other advanced techniques.

Like many modern block ciphers, SM4 is designed to resist differential cryptanalysis — a powerful method introduced by Biham and Shamir [3,4]. Nevertheless, researchers continue to investigate differential characteristics in pursuit of potential vulnerabilities. Notably, Su *et al.* [19] identified a family of 19-round differential characteristics for SM4, including one with a maximal probability of 2^{-124}. By exploiting this distinguisher, they launched the first 23-round differential attack on SM4.

For nearly a decade, the probability bound of 19-round differential characteristics in SM4 remained unchallenged. In 2019, Liu *et al.* [13,14] leveraged an automated approach based on the Simple Theorem Prover (STP) to search for differential and linear characteristics in S-box-based ciphers. Through this method, they claimed to discover a 19-round SM4 differential characteristic with a probability of 2^{-123}, although they did not provide detailed information about the characteristic or confirm whether 2^{-123} is indeed the tightest bound.

More recently, Li and Sun [10] used a Boolean Satisfiability (SAT) based automated search to address the gap left by Liu *et al.* They confirmed that 2^{-123} is indeed the upper bound for 19-round differential characteristics in SM4 and publicly released the first known optimal 19-round characteristic. However, this newly identified characteristic has not yet been employed in any differential attack on SM4. Consequently, the primary motivation for this work was to determine whether this optimal 19-round differential characteristic could be leveraged to improve differential attacks on SM4.

In this paper, we begin by analyzing the optimal differential characteristic proposed in [10]. Our findings reveal that the output difference of this characteristic reduces the pre-sieving probability in the differential attack, resulting in a time complexity that exceeds that of brute force. Since the existing differential characteristic does not yield an improvement in the differential attack on SM4, we shift our focus to identifying other 19-round differential characteristics with the same probability of 2^{-123} that are more conducive to key recovery.

In view of the SAT-based method's outstanding performance in searching for optimal differential characteristics in SM4, we also adopt an automated SAT-based search strategy to discover new differential characteristics. Given the vast number of 19-round optimal differential characteristics with a probability of 2^{-123} in SM4 — and considering the lengthy runtime required by SAT solvers for multi-solution searches — it is not feasible to enumerate all such characteristics in order to identify those best suited for key-recovery attacks. Instead, we explore the properties these optimal 19-round differential characteristics share, as well as those that set them apart.

By translating the characteristic properties that influence key recovery into Boolean expressions, we successfully identified three structural properties common to all optimal 19-round differential characteristics.

- The total number of active S-boxes cannot be fewer than 19.
- Each round must contain at least one active S-box.
- The number of active S-boxes is exactly 19, with precisely one active S-box present in every round.

Based on these properties, we find that the 19-round differential distinguisher derived from any optimal characteristic has a probability of $2^{-123.99}$.

Although all 19-round distinguishers from the optimal characteristics share this overall probability, their respective pre-sieving probabilities differ, influencing their practical utility in differential attacks. Using Boolean encodings, we confirm the existence of key-recovery-friendly characteristics among the optimal 19-round differential characteristics of SM4. Furthermore, we identify four representative differential characteristics that are particularly favorable for key-recovery attacks.

We conduct an in-depth analysis of one of the four differential characteristics in the context of key-recovery attacks. To comprehensively evaluate its performance, we employ both the hypothesis testing paradigm and the key ranking paradigm, with the resulting performance metrics summarized in Table 1. Under the hypothesis testing paradigm, our attack requires slightly more data than the 23-round attack proposed by Su *et al.* [19], yet it achieves lower time and memory complexities while ensuring at least a 90% success probability. Under the key ranking paradigm, our attack also requires more data than the 23-round attack by Zhao *et al.* [26], but it offers lower time complexity and a higher success probability.

Table 1. Comparison of differential attacks on SM4.

Round	Evaluation Method	Data Complexity	Time Complexity	Memory Complexity	Success Probability	Reference
21 (18 + 3)	Signal-to-noise ratio in [3]	$2^{118.0}$	$2^{126.6}$	$2^{118.0}$	-	[24]
22 (18 + 4)	Poisson limit theorem	$2^{118.0}$	$2^{125.7}$	$2^{118.0}$	60.0%	[9]
22 (18 + 4)	Poisson limit theorem	$2^{117.0}$	$2^{112.3}$	$2^{117.0}$	76.2%	[25]
23 (19 + 4)	Signal-to-noise ratio in [3]	$2^{118.0}$	$2^{126.7}$	$2^{118.0}$	98.9%	[19]
23 (19 + 4)	Key ranking paradigm in [5]	$2^{113.5}$	$2^{126.7}$	negligible	90.0%	[26]
23 (19 + 4)	Hypothesis testing paradigm in [5]	$\mathbf{2^{120.4}}$	$\mathbf{2^{122.8}}$	$\mathbf{2^{116.4}}$	**90.0**%	**Section 4.1**
23 (19 + 4)	Key ranking paradigm in [5]	$\mathbf{2^{116.8}}$	$\mathbf{2^{122.8}}$	negligible	**99.9**%	**Section 4.2**

The remainder of this paper is organized as follows. Section 2 presents an overview of differential cryptanalysis and summarizes essential information on SM4. Section 3 describes our methodology for identifying key-recovery-friendly optimal 19-round differential characteristics of SM4. In Section 4, we evaluate the attack performance using one of these newly identified 19-round characteristics. Finally, Section 5 concludes the paper.

2 Preliminary

2.1 Differential Cryptanalysis

Differential cryptanalysis [3,4] is a cryptanalytic method primarily applied to symmetric-key algorithms, especially block ciphers. It leverages the nonlinear components of a cipher to identify cases where certain *input differences* Δ_{in} yield specific *output differences* Δ_{out} with a higher probability than would occur by chance. The pair $(\Delta_{in}, \Delta_{out})$ is called a *differential*. The *probability* of this differential, for an n-bit primitive E_K, is defined as

$$\Pr_{E_K}(\Delta_{in}, \Delta_{out}) = \frac{\{x \in \mathbb{F}_2^n \mid E_K(x) \oplus E_K(x \oplus \Delta_{in}) = \Delta_{out}\}}{2^n}.$$

The *weight* of the differential is given by

$$\omega_{E_K}(\Delta_{in}, \Delta_{out}) = -\log_2[\Pr_{E_K}(\Delta_{in}, \Delta_{out})].$$

Determining the probability of a valid differential in a multi-round cryptographic algorithm can be quite challenging. Typically, *differential characteristics* are constructed to trace internal differences after each round. Let $(\Delta_{in} = \Delta_0, \Delta_1, \ldots, \Delta_R = \Delta_{out})$ represent an R-round differential characteristic. Suppose the R-round encryption function E_K is composed of R round functions, i.e., $E_K = F_{K_{R-1}} \circ F_{K_{R-2}} \circ \cdots \circ F_{K_0}$. Under the assumption that the round keys K_0, K_1, …, K_{R-1} are uniformly random and independent, the probability of this differential characteristic is the product of the differential transition probabilities at each round:

$$\Pr_{E_K}(\Delta_0, \Delta_1, \ldots, \Delta_R) = \prod_{r=0}^{R-1} \Pr_{F_{K_r}}(\Delta_r, \Delta_{r+1}).$$

When determining differential characteristics for round functions, S-boxes often present the greatest complexity. To simplify their analysis, one typically constructs a *Differential Distribution Table* (DDT). For an s-bit S-box, the DDT is a $2^s \times 2^s$ table, with the entry in the i-th row and j-th column indicating the number of input-output pairs that satisfy the differential (i, j). An S-box is considered *active* if its differential corresponds to a DDT entry that is nonzero yet strictly less than 2^s.

2.2 Data Complexity and Success Probability in Differential Attack

Once an effective differential characteristic for the target cipher is identified, a differential attack can be mounted. Let p_0 denote the probability of the R-round differential characteristic, and let N be the number of plaintext pairs used in the attack. As a form of statistical cryptanalysis, we define the statistic Σ (used in the differential attack) as the number of pairs that satisfy the differential characteristic. Under the correct key guess, Σ follows a binomial distribution with parameters (N, p_0). Conversely, if the probability that a pair satisfies the differential characteristic under an incorrect key guess is p (where $p < p_0$), then Σ follows a binomial distribution with parameters (N, p).

Once the observed value of Σ is obtained from the sample, we construct a list $\mathbb{L}$ of the most likely keys among all κ possible keys. As shown by Blondeau *et al.* [5], two paradigms can be adopted for this process. In both paradigms, the *success probability* P_S is defined as the probability that the correct key is included in $\mathbb{L}$.

Hypothesis Testing Paradigm. *Hypothesis testing* begins by setting a *threshold* τ. Any key whose statistic Σ is at least τ is included in the list $\mathbb{L}$. Two types of errors can arise.

1. *Non-detection error probability* (α): This error occurs if the correct key k_0 is not included in $\mathbb{L}$ (i.e., $k_0 \notin \mathbb{L}$). From its definition,
$$\alpha = \sum_{i<\tau} \binom{N}{i} p_0^i (1-p_0)^{N-i}. \tag{1}$$
In this situation, the success probability P_S is $1-\alpha$.
2. *False alarm error probability* (β): This error occurs if an incorrect key k is included in $\mathbb{L}$ (i.e., $k \in \mathbb{L}$). In this case,
$$\beta = \sum_{i\geqslant\tau} \binom{N}{i} p^i (1-p)^{N-i}. \tag{2}$$

The following theorem can be used to estimate α and β.

Theorem 1 (Blondeau *et al.* [5]). *Let p_0 and p be real numbers such that $0 < p < p_0 < 1$, and let τ satisfy $p < \tau/N < p_0$. Define Σ_0 and Σ_k as binomial random variables with parameters (N, p_0) and (N, p), respectively. Then,*

$$\Pr(\Sigma_0 \leqslant \tau) \overset{N\to\infty}{\sim} \frac{p_0\sqrt{1-\tau/N}}{(p_0-\tau/N)\sqrt{2\pi\tau}} e^{-ND(\frac{\tau}{N}\|p_0)}, \tag{3}$$

$$\Pr(\Sigma_k \geqslant \tau) \overset{N\to\infty}{\sim} \frac{(1-p)\sqrt{\tau/N}}{(\tau/N-p)\sqrt{2\pi N(1-\tau/N)}} e^{-ND(\frac{\tau}{N}\|p)}, \tag{4}$$

where $D(q_1\|q_2) \triangleq q_1 \times \ln\left(\frac{q_1}{q_2}\right) + (1-q_1) \times \ln\left(\frac{1-q_1}{1-q_2}\right)$ is the Kullback-Leibler divergence between two Bernoulli distributions with parameters q_1 and q_2.

Remark 1. As noted in [5], when p and p_0 are not extremely small, the approximation formulas from Theorem 1 can reliably estimate both types of error probabilities. However, when p and p_0 are exceedingly small, directly applying these approximations to compute α and β may cause substantial errors. For instance, in our evaluation of the SM4 differential attack, if we set $p = 2^{-128}$ and $p_0 = 2^{-123.99}$ with N ranging from 2^{123} to 2^{126}, using (3) sometimes produces negative values or values greater than 1 — both of which are invalid probabilities. As a result, in the subsequent analysis, we do not use the approximation formulas from Theorem 1. Instead, we revert to the original definitions of the error probabilities and employ formulas (1) and (2) to evaluate α and β.

Key Ranking Paradigm. *Key ranking* fixes the size of the candidate key list $\mathbb{L}$ to an integer ℓ, retaining the ℓ most probable keys. By leveraging properties of the Beta distribution, the success probability can be evaluated using the following theorem.

Theorem 2 (Blondeau *et al.* [5]). *Let* F *be the cumulative distribution function of a binomial distribution with parameters* (N, p)*, i.e.,* $\mathsf{F}(x) \triangleq \sum_{i \leqslant x} \binom{N}{i} p^i (1-p)^{N-i}$*. Its inverse function* F^{-1} *is* $\mathsf{F}^{-1}(x) = \min\{t \in \mathbb{N} \mid \mathsf{F}(t) \geqslant x\}$*. Denote by* $\mathsf{f}_0(i)$ *the probability that the statistic* Σ *for the correct key takes the value* i*:* $\mathsf{f}_0(i) = \binom{N}{i} p_0^i (1-p_0)^{N-i}$*. Let* $\lambda \triangleq \frac{\ell-1}{\kappa-2}$*. If* $\lambda \leqslant 0.25$*, then the success probability* P_S *satisfies*

$$P_S = 1 - \delta + O\left(\delta(1 + C_\lambda)\sqrt{\frac{\ln(\ell/\delta^2)}{\ell}} + \frac{1}{\ell^2} + \frac{1}{\kappa}\right), \tag{5}$$

where

$$\delta = \sum_{i=0}^{\mathsf{F}^{-1}(1-\lambda)-1} \mathsf{f}_0(i) \text{ and } C_\lambda = \frac{p}{p_0} \frac{p_0(N+1) - \mathsf{F}^{-1}(1-\lambda)}{\mathsf{F}^{-1}(1-\lambda) - p(N+1)}.$$

Through experimental verification, Blondeau *et al.* showed that when N takes the form

$$N = c \frac{-\ln(2\sqrt{\pi}\ell/\kappa)}{D(p_0||p)},$$

the success probability P_S depends primarily on the constant c.

2.3 Description of SM4

SM4 [7] operates on 128-bit blocks, consistent with other modern block ciphers such as the Advanced Encryption Standard (AES) [16]. The cipher uses a 128-bit key K, providing a symmetric encryption scheme in which the same key is employed for both encryption and decryption. SM4 consists of 32 rounds of

encryption. Let $(X_i, X_{i+1}, X_{i+2}, X_{i+3}) \in (\mathbb{F}_2^{32})^4$ denote the 128-bit input to the i-th round, where $0 \leqslant i < 32$. Denote the round keys by RK_i.

In each round, the internal state is updated using the function $\mathcal{T}$. The encryption process for the i-th round is given by

$$X_{i+4} = X_i \oplus \mathcal{T}(X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus RK_i).$$

The function $\mathcal{T}$ is composed of a nonlinear transformation $\mathcal{S}$ and a linear transformation $\mathcal{L}$, i.e., $\mathcal{T} = \mathcal{L} \circ \mathcal{S}$. An illustration for the round function of SM4 can be found in Figure 1.

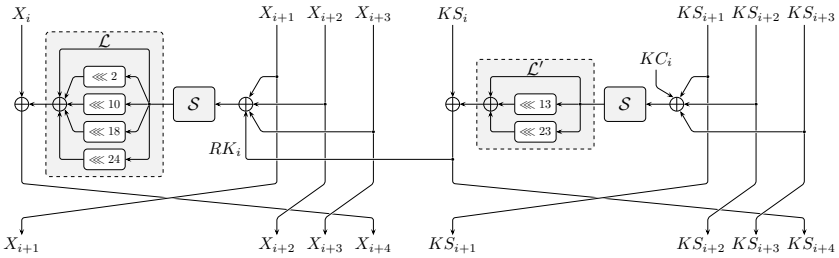


Fig. 1. Round function and key schedule of SM4.

Nonlinear transformation $\mathcal{S}$. SM4 applies the same 8×8 S-box S four times in parallel to a 32-bit input. Let $Y = (Y[0], Y[1], Y[2], Y[3]) \in (\mathbb{F}_2^8)^4$ represent the input to $\mathcal{S}$, and $Z = (Z[0], Z[1], Z[2], Z[3]) \in (\mathbb{F}_2^8)^4$ denote the corresponding output. The transformation $\mathcal{S}$ is defined by

$$(Z[0], Z[1], Z[2], Z[3]) = (S(Y[0]), S(Y[1]), S(Y[2]), S(Y[3])).$$

Linear transformation $\mathcal{L}$. The linear transformation $\mathcal{L}$ is a simple linear function that takes the output of $\mathcal{S}$ as its input. Given Z as the input to $\mathcal{L}$, its output is computed as

$$\mathcal{L}(Z) = Z \oplus (Z \lll 2) \oplus (Z \lll 10) \oplus (Z \lll 18) \oplus (Z \lll 24).$$

A key property of the SM4 S-box, often exploited in its cryptanalysis, is the following.

Property 1. For any nonzero input difference, the SM4 S-box produces 127 possible output differences. Among these, one specific output difference appears with probability 2^{-6}, while each of the remaining 126 output differences occurs with probability 2^{-7}.

As illustrated in Figure 1, the SM4 key schedule operates similarly to its encryption function. The function $\mathcal{T}'$, used in the key schedule, closely resembles $\mathcal{T}$ except for the linear transformation: it uses

$$\mathcal{L}'(Z) = Z \oplus (Z \lll 13) \oplus (Z \lll 23)$$

instead of $\mathcal{L}$. The round key RK_i for the i-th round is then computed as

$$RK_i = KS_{i+4} = KS_i \oplus \mathcal{T}'(KS_{i+1} \oplus KS_{i+2} \oplus KS_{i+3} \oplus KC_i), \quad 0 \leqslant i < 32,$$

where KC_i are certain constants. Moreover, (KS_0, KS_1, KS_2, KS_3) is obtained by XORing the master key K with a fixed system parameter. Additional details on the cipher can be found in [7].

2.4 Previous Most Effective Differential Attacks on SM4

The most effective differential attacks on SM4 so far target 23 rounds. Both known 23-round differential attacks on SM4 [19,26] are based on the 127^2 19-round differential characteristics first introduced by Su *et al.* [19]. These characteristics share the same output difference[5]:

$$(\texttt{3f0000cf}, \texttt{ccf300fc}, \texttt{f3f30033}, \texttt{f3f30033}),$$

while their input differences vary. The input differences take the form

$$(a_0, \texttt{f3f30033}, \texttt{f3f30033}, \texttt{f3000030}),$$

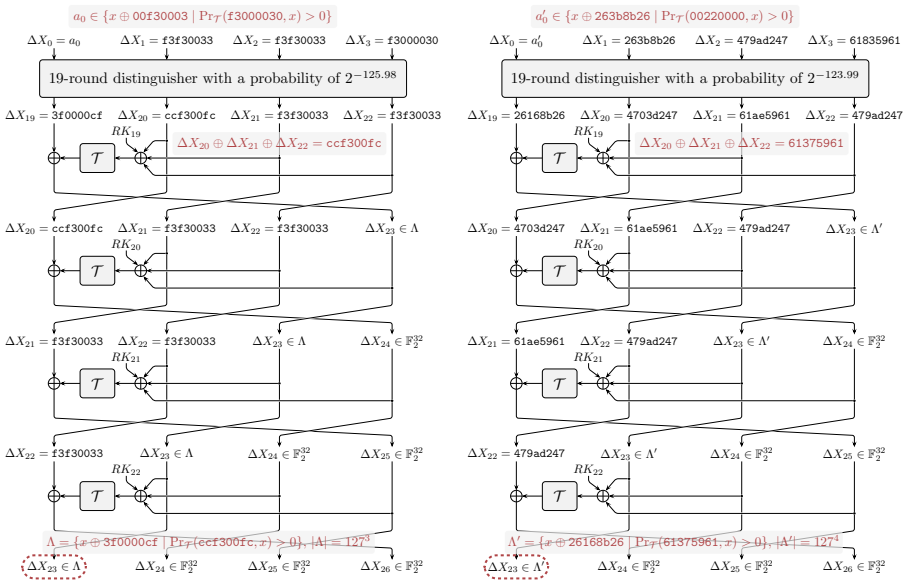
where $a_0 \in \{x \oplus \texttt{00f30003} \mid \Pr_{\mathcal{T}}(\texttt{f3000030}, x) > 0\}$. Among these 127^2 differential characteristics, one has a probability of 2^{-124}, 254 have a probability of 2^{-125}, and the remaining characteristics have a probability of 2^{-126}. The average probability of these differential characteristics is $2^{-125.98}$, representing the probability for the 19-round distinguisher.

In the 23-round differential attacks, this 19-round distinguisher is placed at the beginning (rounds 0 to 18), followed by four additional rounds. As illustrated in Figure 2(a), only the second and third bytes of ΔX_0 are fixed, enabling structures to be constructed at the distinguisher input so that more plaintext pairs can be generated from a fixed number of plaintexts. At the output of the distinguisher, $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22} = \texttt{ccf300fc}$, which matches the input difference of the function $\mathcal{T}$ in the nineteenth round. Since $X_{23} = X_{19} \oplus \mathcal{T}(X_{20} \oplus X_{21} \oplus X_{22} \oplus RK_{19})$, ΔX_{23} is confined to the set

$$\Lambda = \{x \oplus \texttt{3f0000cf} \mid \Pr_{\mathcal{T}}(\texttt{ccf300fc}, x) > 0\},$$

which contains 127^3 elements. Because ΔX_{23} is part of the ciphertext difference, the condition $\Delta X_{23} \in \Lambda$ filters out ciphertext pairs with probability $2^{-11.03}$ prior to key enumeration. This probability is called the *pre-sieving probability*. As the initial number of pairs heavily impacts the time complexity of subsequent key enumeration, the magnitude of this pre-sieving probability is critically important.

[5] `Typewriter` font is used to represent hexadecimal bit vectors.



(a) Key-recovery procedure using the distinguisher proposed by Su *et al.* (b) Key-recovery procedure using the distinguisher proposed by Li and Sun.

Fig. 2. 23-round key-recovery procedures with different distinguishers.

2.5 Automated Search for Optimal Differential Characteristics

Several studies [23,11,13,14,10] have employed automated techniques to explore the differential properties of SM4. Among these, the method based on Boolean satisfiability problems (SAT) has delivered the best performance to date.

In a SAT problem, one seeks an assignment of truth values to variables that satisfies a given Boolean formula — making it evaluate to `true`. SAT problems are widely utilized in areas such as cryptography, hardware and software verification, and automated theorem proving. Due to their importance, many advanced algorithms and tools, known as *SAT solvers*, have been developed to efficiently handle large-scale and complex SAT instances. When given to SAT solvers, SAT problems are generally expressed in *Conjunctive Normal Form* (CNF), where the formula is a conjunction (AND, $\wedge$) of clauses, each *clause* being a disjunction (OR, $\vee$) of literals. A *literal* is defined as a variable or its negation (NOT, $\bar{\cdot}$).

By formulating the search for differential characteristics in SM4 as SAT problems, Li and Sun [10] introduced the first publicly available optimal 19-round differential characteristic with a probability of 2^{-123}. Its input difference is

$$(\texttt{4703d247}, \texttt{263b8b26}, \texttt{479ad247}, \texttt{61835961}),$$

and its output difference is

$$(\texttt{26168b26}, \texttt{4703d247}, \texttt{61ae5961}, \texttt{479ad247}).$$

3 Explorations of Key-Recovery-Friendly Distinguishers

In this section, we first demonstrate that the previously discovered optimal differential characteristic does not yield a stronger differential attack on SM4. Subsequently, we succeed in identifying alternative 19-round differential characteristics that are more conducive to key recovery.

3.1 Main Problem of Employing Differential Characteristic in [10]

Although Li and Sun [10] presented an optimal 19-round differential characteristic for SM4, its potential to strengthen differential attacks has not been examined. In this work, we show that the 19-round distinguisher derived from their characteristic offers no advantage when used to launch 23-round differential attacks on SM4.

As illustrated in Figure 2(b), since $\Delta X_1 \oplus \Delta X_2 \oplus \Delta X_3 = \texttt{00220000}$ serves as the input difference for the function $\mathcal{T}$ in the zeroth round, we can derive a set of 127 differential characteristics based on the original characteristic from [10] (see Section 2.5). These differential characteristics share the same output difference as the optimal one but differ in their input differences, which are of the form

$$(a_0', \texttt{263b8b26}, \texttt{479ad247}, \texttt{61835961}),$$

where $a_0' \in \{x \oplus \texttt{263b8b26} \mid \Pr_{\mathcal{T}}(\texttt{00220000}, x) > 0\}$. Among these 127 characteristics, one achieves the highest probability of 2^{-123}, while the remaining 126 have a probability of 2^{-124}. Consequently, the new 19-round distinguisher has an average probability of $2^{-123.99}$.

To mount a 23-round attack, we append four rounds to the new 19-round distinguisher. Since $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22} = \texttt{61375961}$, the possible values of ΔX_{23} belong to

$$\Lambda' = \{x \oplus \texttt{26168b26} \mid \Pr_{\mathcal{T}}(\texttt{61375961}, x) > 0\},$$

which contains 127^4 elements. When applying the condition $\Delta X_{23} \in \Lambda'$ to filter ciphertext pairs, the pre-sieving probability is $2^{-4.05}$. Given the large number of pairs advancing to the key enumeration phase, the time complexity of the attack will be difficult to manage.

Given that even the 19-round distinguisher with higher differential probability does not yield a more efficient differential attack, we must also consider whether other 19-round differential characteristics with probability 2^{-123} might facilitate stronger attacks. We aim to use a SAT-based automated approach[6] to investigate the existence of more effective differential characteristics. Before presenting our method, we first review the automated techniques from [10] for searching differential characteristics of SM4.

[6] In this work, we employ the SAT solvers Kissat [2] and CryptoMiniSat [18].

3.2 Methods for Finding Differential Characteristics of SM4

To find differential characteristics of SM4, one can translate differential propagation through the cipher's components into Boolean formulas in CNF. This SAT formulation, geared toward identifying SM4 differential characteristics, naturally divides into three parts: SAT models for the linear components, the nonlinear components, and the objective function.

SAT Models for Linear Components. As discussed in [10], characterizing differential propagation over the linear components of SM4 only requires a SAT model for multi-input XOR operations. Because Kissat does not natively support Boolean formulas involving XOR, the SAT model from [21] is employed instead. Throughout this work, for an n-bit vector ΔX, we use $\Delta X^{(i)}$ to denote the i-th bit of ΔX, where $0 \leqslant i < n$.

Model 1 (l-Input XOR, [21]) *Consider the n-bit XOR operation $Y = X_0 \oplus X_1 \oplus \cdots \oplus X_{l-1}$, whose input and output differences are ΔX_0, ΔX_1, ..., ΔX_{l-1} and ΔY, respectively. The differential propagation is valid if and only if ΔX_0, ΔX_1, ..., ΔX_{l-1}, and ΔY satisfy the following equations for every $(l+1)$-bit vector $(\theta^{(0)}, \theta^{(1)}, \ldots, \theta^{(l)})$ such that $\theta^{(0)} \oplus \theta^{(1)} \oplus \cdots \oplus \theta^{(l)} = 1$.*

$$(\Delta X_0^{(i)} \oplus \theta^{(0)}) \vee \cdots \vee (\Delta X_{l-1}^{(i)} \oplus \theta^{(l-1)}) \vee (\Delta Y^{(i)} \oplus \theta^{(l)}) = 1, 0 \leqslant i < n.$$

SAT Models for Nonlinear Components. In SM4, the S-box S is the sole nonlinear component. The SAT model for S, which aims to capture differential probabilities, follows the approach in [20]. Specifically, each S-box is characterized by sixteen Boolean variables, $\Delta X = (\Delta X^{(0)}, \Delta X^{(1)}, \ldots, \Delta X^{(7)})$ and $\Delta Y = (\Delta Y^{(0)}, \Delta Y^{(1)}, \ldots, \Delta Y^{(7)})$, representing the input and output differences. The probability of differential propagation through S can take one of three values: 1, 2^{-6}, or 2^{-7}. To encode these probabilities, two additional Boolean variables $w = (w^{(0)}, w^{(1)})$ are introduced, and the corresponding weight is computed as $w^{(0)} + 6w^{(1)}$. Hence, the 18-bit vector $(\Delta X, \Delta Y, w)$ must belong to the set

$$\Theta = \left\{ (\Delta X, \Delta Y, w) \;\middle|\; \begin{array}{l} \Delta X, \Delta Y \in \mathbb{F}_2^8, w \in \mathbb{F}_2^2, \mathrm{Pr}_S(\Delta X, \Delta Y) > 0 \\ w = \begin{cases} (1,1), & \text{if } \mathrm{Pr}_S(\Delta X, \Delta Y) = 2^{-7} \\ (0,1), & \text{if } \mathrm{Pr}_S(\Delta X, \Delta Y) = 2^{-6} \\ (0,0), & \text{if } \mathrm{Pr}_S(\Delta X, \Delta Y) = 1 \end{cases} \end{array} \right\}.$$

To enforce this, the model introduces a clause

$$\bigvee_{i=0}^{7} \left(\Delta X^{(i)} \oplus \theta^{(i)}\right) \vee \bigvee_{i=0}^{7} \left(\Delta Y^{(i)} \oplus \theta^{(i+8)}\right) \vee \bigvee_{i=0}^{1} \left(w^{(i)} \oplus \theta^{(i+16)}\right) = 1$$

for every 18-bit vector $\theta \notin \Theta$, thus forming a baseline SAT description of the S-box. However, because $\mathbb{F}_2^{18} \backslash \Theta$ has 229758 elements, encoding all such clauses

directly would produce an intractably large SAT problem. To mitigate this complexity, the ESPRESSO logic minimizer [6] is applied, reducing the model to 8599 clauses. This streamlined SAT model effectively captures the S-box's differential probability properties.

SAT Models for the Objective Function. Since the goal of the automatic search is to identify differential characteristics with high probabilities, the objective function in the SAT problem can be represented by a *cardinality inequality*: $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$, where w_i are Boolean variables encoding the weights of possible differential propagations through S-boxes, and ϖ is a predefined upper bound on the characteristic's total weight. A common technique for translating this cardinality constraint into Boolean logic is the sequential encoding method [17]. Wang *et al.* [22] demonstrated that when the upper and lower bounds of each partial sum $\sum_{i=0}^{\gamma'-1} w_i$ (for $0 \leqslant \gamma' < \gamma$) in the inequality $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$ are known, this information can be integrated into the sequential encoding method to construct a more effective objective function model. The enhanced model significantly improves the efficiency of SAT-based searches. Due to space limitations, the two SAT models for the objective function are provided in Appendix A.

3.3 Structural Properties of 19-Round Optimal Characteristics

Theoretically, confirming whether other 19-round differential characteristics with a probability of 2^{-123} exist requires only a multi-solution SAT solver — such as CryptoMiniSat [18] — to enumerate these characteristics in SM4. However, our experiments with CryptoMiniSat suggest that there may be a very large number of such characteristics, and exhaustively exploring them is extremely time-consuming. In one instance, we ran the solver for 15 days and discovered only 77491 characteristics, yet the search remained incomplete.

Although we can select some of these 77491 characteristics that potentially facilitate more efficient key-recovery attacks than the one presented in [10], this approach does not guarantee finding the truly optimal distinguisher. Consequently, rather than enumerating all 19-round optimal differential characteristics of SM4, we focus on analyzing both their shared and distinct properties. These properties — closely tied to the probability of the derived distinguisher and the pre-sieving probability — allow us to estimate an upper limit on the effectiveness of 19-round optimal differential characteristics for key-recovery attacks.

Maximum Number of Derived Characteristics from a Single Optimal Characteristic. As discussed in Sections 2.4 and 3.1, the probability of a distinguisher depends not only on the probability of its optimal characteristic but also on the number of differential characteristics derivable from that characteristic. Although the optimal characteristic in [10] achieves a probability of 2^{-123},

it yields only 127 associated differential characteristics. In contrast, the distinguisher in [19] encompasses 127^2 characteristics. Because a greater number of characteristics can boost the probability of the distinguisher, we first investigate whether it is possible to find a 19-round characteristic with probability 2^{-123} that generates at least 127^2 characteristics.

To explore this question, we revisit the differential characteristics presented in [19] and [10]. We observe that the characteristic proposed in [10] achieves a higher probability but involves 19 active S-boxes, whereas the best characteristic from [19] contains only 18 active S-boxes. Both works [19,10] confirm that 18 is the minimum number of active S-boxes possible for a 19-round differential characteristic in SM4. Thus, we aim to determine whether a 19-round differential characteristic exists in SM4 that offers a probability of 2^{-123} while maintaining exactly 18 active S-boxes.

We address this problem via automated searches. Because we must constrain both the probability of the characteristic and the number of active S-boxes, we employ a dual-objective SAT problem. In particular, we use Model 6 to limit the differential characteristic's weight to 123 and Model 5 to fix the number of active S-boxes at 18. Leveraging the SAT solver Kissat under these constraints yields the following property.

Property 2. The number of active S-boxes in any 19-round differential characteristic of SM4 with a probability of 2^{-123} cannot be fewer than 19.

Because the minimum number of active S-boxes in a 19-round optimal differential characteristic is 19, the arrangement of these active S-boxes directly affects how many differential characteristics can be derived from the optimal one. For example, in the differential characteristic analyzed in Section 2.4 of [19], there are two active S-boxes in the first round; consequently, under fixed output differences, the 127^2 choices of input differences yield 127^2 possible differential characteristics. This observation prompts the question of whether a 19-round differential characteristic with probability 2^{-123} could also feature two active S-boxes in its first round.

To investigate the positions of active S-boxes in a 19-round optimal characteristic more comprehensively, we reformulate the problem into a SAT-based check. Specifically, we test whether a 19-round optimal differential characteristic can feature a round with no active S-boxes. Let $w_{r,i}$ be the 2-bit Boolean variable encoding the weight of the i-th S-box in the r-th round. We then incorporate the following model into our SAT problem to determine if the r-th round can indeed contain zero active S-boxes.

Model 2 *The r-th round of the differential characteristic contains no active S-boxes if and only if the variables $w_{r,0}^{(1)}$, $w_{r,1}^{(1)}$, $w_{r,2}^{(1)}$, and $w_{r,3}^{(1)}$ satisfy the following equations.*

$$\overline{w_{r,i}^{(1)}} = 1, 0 \leqslant i < 4.$$

Our test results indicate that, when these conditions are successively introduced into the SAT problem, none of the 19 resulting SAT instances is solvable. Hence, we arrive at the following conclusion.

Property 3. In SM4, any 19-round differential characteristic with probability 2^{-123} must have at least one active S-box in each round.

Since the maximum differential propagation probability of SM4's S-box is 2^{-6}, and the probability of the optimal 19-round characteristic is 2^{-123}, the number of active S-boxes in the optimal characteristic cannot exceed 20. By Property 3, this implies that the optimal 19-round differential characteristic may contain at most one round with two active S-boxes, with each of the remaining rounds featuring exactly one active S-box. To determine whether the r-th round of the optimal 19-round differential characteristic can accommodate two active S-boxes, we incorporate the following model into our SAT problem.

Model 3 *The r-th round of the SM4 differential characteristic has at least two active S-boxes if and only if the variables $w_{r,0}^{(1)}$, $w_{r,1}^{(1)}$, $w_{r,2}^{(1)}$, and $w_{r,3}^{(1)}$ satisfy the following equations.*

$$
\begin{aligned}
w_{r,0}^{(1)} \vee w_{r,1}^{(1)} \vee w_{r,2}^{(1)} \vee w_{r,3}^{(1)} &= 1,\\
\overline{w_{r,0}^{(1)}} \vee w_{r,1}^{(1)} \vee w_{r,2}^{(1)} \vee w_{r,3}^{(1)} &= 1,\\
w_{r,0}^{(1)} \vee \overline{w_{r,1}^{(1)}} \vee w_{r,2}^{(1)} \vee w_{r,3}^{(1)} &= 1,\\
w_{r,0}^{(1)} \vee w_{r,1}^{(1)} \vee \overline{w_{r,2}^{(1)}} \vee w_{r,3}^{(1)} &= 1,\\
w_{r,0}^{(1)} \vee w_{r,1}^{(1)} \vee w_{r,2}^{(1)} \vee \overline{w_{r,3}^{(1)}} &= 1.
\end{aligned}
$$

The test results show that, when these conditions are successively included in the SAT problem, all 19 instances become unsatisfiable. Therefore, we arrive at the following conclusion.

Property 4. In every optimal 19-round differential characteristic of SM4, there are exactly 19 active S-boxes — one in each round.

According to Property 4, the first round of an optimal 19-round differential characteristic must contain exactly one active S-box. As a result, the total number of differential characteristics derivable from a single optimal characteristic is capped at 127. Furthermore, the probability of any 19-round differential distinguisher constructed from an optimal characteristic is $2^{-123.99}$, matching the probability of the distinguisher discussed in Section 3.1.

Maximum Pre-Sieving Probability for Optimal Characteristics. As discussed in Sections 2.4 and 3.1, the pre-sieving probability is influenced by the XOR of the last three branches of the output difference, $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$.

A smaller number of nonzero bytes in $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ reduces the range of possible values for ΔX_{23} in the first branch of the ciphertext pair, thereby decreasing the pre-sieving probability. In the distinguisher from [19], $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ contains three nonzero bytes, whereas in the distinguisher from [10], it contains four. Since our goal is to identify optimal 19-round differential characteristics for key-recovery attacks, we focus on whether a 19-round characteristic with a probability of 2^{-123} can be found such that $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ has exactly three nonzero bytes. This condition is equivalent to asking whether there exists one zero byte in $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$. To test this, we integrate the following model into the SAT problem, enabling the solver to analyze whether the i-th byte of $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ is zero.

Model 4 *In SM4's 19-round optimal differential characteristic, the i-th byte of $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ is zero if and only if, for every $0 \leqslant j < 8$, the variables $\Delta X_{20}^{(8i+j)}$, $\Delta X_{21}^{(8i+j)}$, and $\Delta X_{22}^{(8i+j)}$ satisfy the following equations.*

$$\begin{aligned}
\overline{\Delta X_{20}^{(8i+j)}} \vee \Delta X_{21}^{(8i+j)} \vee \Delta X_{22}^{(8i+j)} &= 1,\\
\Delta X_{20}^{(8i+j)} \vee \overline{\Delta X_{21}^{(8i+j)}} \vee \Delta X_{22}^{(8i+j)} &= 1,\\
\Delta X_{20}^{(8i+j)} \vee \Delta X_{21}^{(8i+j)} \vee \overline{\Delta X_{22}^{(8i+j)}} &= 1,\\
\overline{\Delta X_{20}^{(8i+j)}} \vee \overline{\Delta X_{21}^{(8i+j)}} \vee \overline{\Delta X_{22}^{(8i+j)}} &= 1.
\end{aligned}$$

Test results indicate that each of the four bytes in $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ can independently be zero. Naturally, this raises the question of whether two of these bytes can be zero simultaneously. To explore this, we adapt the constraints from Model 4 and integrate them into the SAT problem. The solver's outcome reveals that, in SM4, no 19-round differential characteristic with probability 2^{-123} has two zero bytes in the XOR of the last three branches of the output difference. Consequently, the minimum number of nonzero bytes in $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22}$ is three. This finding also establishes the upper bound on the pre-sieving probability for optimal differential characteristics in key-recovery attacks, consistent with the pre-sieving probability of $2^{-11.03}$ reported in [19]. The four 19-round optimal differential characteristics identified via Model 4 are provided in Table 2. In the next section, we examine the effectiveness of employing the fourth characteristic in differential key-recovery attacks.

4 Attack Performance with the New Distinguisher

In this section, we present differential attacks on 23-round SM4 using a 19-round distinguisher derived from the fourth differential characteristic in Table 2. To ensure fair comparisons with previously proposed distinguishers, we reimplement the attacks in [19,26] using our new 19-round distinguisher. By applying the two paradigms introduced in Section 2.2, we accurately determine the complexities of these attacks in each scenario.

Table 2. Four optimal 19-round differential characteristics enabling effective differential attacks on SM4.

Round r	-	ΔX_r	ΔX_{r+1}	ΔX_{r+2}	ΔX_{r+3}	-	ΔX_r	ΔX_{r+1}	ΔX_{r+2}	ΔX_{r+3}	Pr
0	The first differential characteristic	638a2222	b51f7c7c	7e8a2222	a7955e5e	The second differential characteristic	f56d7cf5	25e04825	f5047cf5	d0ea34d0	-
1		b51f7c7c	7e8a2222	a7955e5e	a81f7c7c		25e04825	f5047cf5	d0ea34d0	25894825	2^{-6}
2		7e8a2222	a7955e5e	a81f7c7c	7e8a2222		f5047cf5	d0ea34d0	25894825	f5047cf5	2^{-7}
3		a7955e5e	a81f7c7c	7e8a2222	1d955e5e		d0ea34d0	25894825	f5047cf5	d06934d0	2^{-6}
4		a81f7c7c	7e8a2222	1d955e5e	0f1f7c7c		25894825	f5047cf5	d06934d0	25634825	2^{-6}
5		7e8a2222	1d955e5e	0f1f7c7c	638a2222		f5047cf5	d06934d0	25634825	f56d7cf5	2^{-6}
6		1d955e5e	0f1f7c7c	638a2222	1d955e5e		d06934d0	25634825	f56d7cf5	d06934d0	2^{-6}
7		0f1f7c7c	638a2222	1d955e5e	b51f7c7c		25634825	f56d7cf5	d06934d0	25e04825	2^{-7}
8		638a2222	1d955e5e	b51f7c7c	c48a2222		f56d7cf5	d06934d0	25e04825	f5877cf5	2^{-7}
9		1d955e5e	b51f7c7c	c48a2222	00955e5e		d06934d0	25e04825	f5877cf5	d00034d0	2^{-7}
10		b51f7c7c	c48a2222	00955e5e	b51f7c7c		25e04825	f5877cf5	d00034d0	25e04825	2^{-7}
11		c48a2222	00955e5e	b51f7c7c	7e8a2222		f5877cf5	d00034d0	25e04825	f5047cf5	2^{-7}
12		00955e5e	b51f7c7c	7e8a2222	a7955e5e		d00034d0	25e04825	f5047cf5	d0ea34d0	2^{-7}
13		b51f7c7c	7e8a2222	a7955e5e	a81f7c7c		25e04825	f5047cf5	d0ea34d0	25894825	2^{-7}
14		7e8a2222	a7955e5e	a81f7c7c	7e8a2222		f5047cf5	d0ea34d0	25894825	f5047cf5	2^{-7}
15		a7955e5e	a81f7c7c	7e8a2222	1d955e5e		d0ea34d0	25894825	f5047cf5	d06934d0	2^{-6}
16		a81f7c7c	7e8a2222	1d955e5e	0f1f7c7c		25894825	f5047cf5	d06934d0	25634825	2^{-6}
17		7e8a2222	1d955e5e	0f1f7c7c	638a2222		f5047cf5	d06934d0	25634825	f56d7cf5	2^{-6}
18		1d955e5e	0f1f7c7c	638a2222	1d955e5e		d06934d0	25634825	f56d7cf5	d06934d0	2^{-6}
19		0f1f7c7c	638a2222	1d955e5e	7e8a2222		25634825	f56d7cf5	d06934d0	f5047cf5	2^{-6}
0	The third differential characteristic	1c1c1e07	18182806	1c1c2d07	0404f301	The fourth differential characteristic	6fb9b974	e793932d	6fb9b98f	882a2a9e	-
1		18182806	1c1c2d07	0404f301	18181b06		e793932d	6fb9b98f	882a2a9e	e79393d6	2^{-6}
2		1c1c2d07	0404f301	18181b06	1c1c2d07		6fb9b98f	882a2a9e	e79393d6	6fb9b98f	2^{-7}
3		0404f301	18181b06	1c1c2d07	04043301		882a2a9e	e79393d6	6fb9b98f	882a2afb	2^{-6}
4		18181b06	1c1c2d07	04043301	1818e806		e79393d6	6fb9b98f	882a2afb	e7939348	2^{-6}
5		1c1c2d07	04043301	1818e806	1c1c1e07		6fb9b98f	882a2afb	e7939348	6fb9b974	2^{-6}
6		04043301	1818e806	1c1c1e07	04043301		882a2afb	e7939348	6fb9b974	882a2afb	2^{-6}
7		1818e806	1c1c1e07	04043301	18182806		e7939348	6fb9b974	882a2afb	e793932d	2^{-7}
8		1c1c1e07	04043301	18182806	1c1ced07		6fb9b974	882a2afb	e793932d	6fb9b9ea	2^{-7}
9		04043301	18182806	1c1ced07	04040001		882a2afb	e793932d	6fb9b9ea	882a2a00	2^{-7}
10		18182806	1c1ced07	04040001	18182806		e793932d	6fb9b9ea	882a2a00	e793932d	2^{-7}
11		1c1ced07	04040001	18182806	1c1c2d07		6fb9b9ea	882a2a00	e793932d	6fb9b98f	2^{-7}
12		04040001	18182806	1c1c2d07	0404f301		882a2a00	e793932d	6fb9b98f	882a2a9e	2^{-7}
13		18182806	1c1c2d07	0404f301	18181b06		e793932d	6fb9b98f	882a2a9e	e79393d6	2^{-7}
14		1c1c2d07	0404f301	18181b06	1c1c2d07		6fb9b98f	882a2a9e	e79393d6	6fb9b98f	2^{-7}
15		0404f301	18181b06	1c1c2d07	04043301		882a2a9e	e79393d6	6fb9b98f	882a2afb	2^{-6}
16		18181b06	1c1c2d07	04043301	1818e806		e79393d6	6fb9b98f	882a2afb	e7939348	2^{-6}
17		1c1c2d07	04043301	1818e806	1c1c1e07		6fb9b98f	882a2afb	e7939348	6fb9b974	2^{-6}
18		04043301	1818e806	1c1c1e07	04043301		882a2afb	e7939348	6fb9b974	882a2afb	2^{-6}
19		1818e806	1c1c1e07	04043301	1c1c2d07		e7939348	6fb9b974	882a2afb	6fb9b98f	2^{-6}

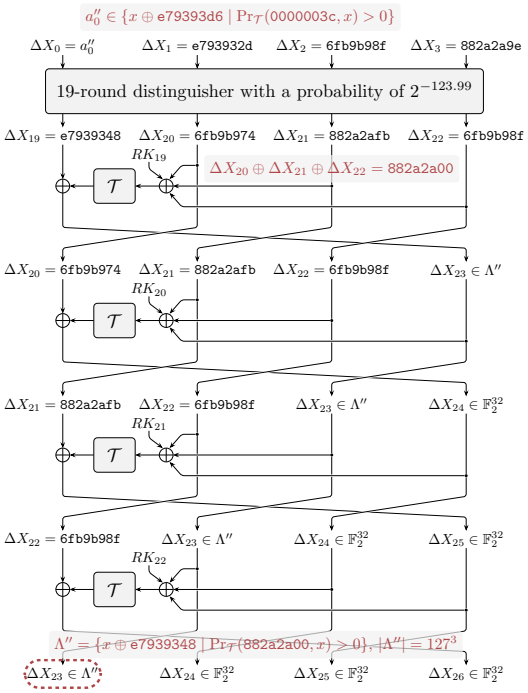


Fig. 3. Key-recovery attack on 23-round SM4 using the new distinguisher.

As depicted in Figure 3, the 19-round distinguisher is applied to the first nineteen rounds of the cipher in both attacks. Its input difference has the form

$$(a_0'', \texttt{e793932d}, \texttt{6fb9b98f}, \texttt{882a2a9e}),$$

where $a_0'' \in \{x \oplus \texttt{e79393d6} \mid \Pr_{\mathcal{T}}(\texttt{0000003c}, x) > 0\}$. The probability of this 19-round distinguisher is $2^{-123.99}$. Because the input difference for the nineteenth $\mathcal{T}$ function is $\Delta X_{20} \oplus \Delta X_{21} \oplus \Delta X_{22} = \texttt{882a2a00}$, the difference ΔX_{23} can only take values from

$$\Lambda'' = \{x \oplus \texttt{e7939348} \mid \Pr_{\mathcal{T}}(\texttt{882a2a00}, x) > 0\},$$

which contains 127^3 possible elements.

To generate plaintext pairs, we construct two pools, M_1 and M_2, each containing 2^{32} plaintexts:

$$M_1 = \{(x, c_1, c_2, c_3) \mid x \in \mathbb{F}_2^{32}\},$$
$$M_2 = \{(y, c_1 \oplus \texttt{e793932d}, c_2 \oplus \texttt{6fb9b98f}, c_3 \oplus \texttt{882a2a9e}) \mid y \in \mathbb{F}_2^{32}\},$$

where c_1, c_2, and c_3 are fixed constants in $\mathbb{F}_2^{32}$. We refer to the pair (M_1, M_2) as a *structure*. Each structure contains 2^{33} distinct plaintexts and can generate approximately $2^{32} \times 127 \approx 2^{38.99}$ plaintext pairs.

4.1 Key-Recovery Attack Using the Method from [19]

Following the approach in [19], the attack proceeds as follows.

Step 1. Select m structures to obtain approximately $m \times 2^{38.99}$ plaintext pairs, generated from $m \times 2^{33}$ plaintexts.

Step 2. For each plaintext pair $(P, \hat{P})$, compute the difference of the corresponding ciphertexts $(C, \hat{C})$. Check whether the first word of the ciphertext difference belongs to the set Λ''. If not, discard the pair. After this filtering process, the expected number of surviving pairs is $m \times 2^{38.99} \times 2^{-11.03} = m \times 2^{27.96}$.

Step 3. For each guess of $RK_{22}[0]$, perform the following steps.

(1) For each remaining ciphertext pair $(C, \hat{C})$, where $C = (C_0, C_1, C_2, C_3)$ and $\hat{C} = (\hat{C}_0, \hat{C}_1, \hat{C}_2, \hat{C}_3)$, partially decrypt to obtain the output difference of the zeroth S-box in the 22-nd round $\mathfrak{d}_{22}[0] = S(C_0[0] \oplus C_1[0] \oplus C_2[0] \oplus RK_{22}[0]) \oplus S(\hat{C}_0[0] \oplus \hat{C}_1[0] \oplus \hat{C}_2[0] \oplus RK_{22}[0])$, and compute $\mathfrak{c}_{22}[0] = (\mathcal{L}^{-1}(C_3 \oplus \hat{C}_3 \oplus \Delta X_{22}))[0]$. If $RK_{22}[0]$ is guessed correctly, then for all correct ciphertext pairs, we must have $\mathfrak{d}_{22}[0] = \mathfrak{c}_{22}[0]$. Discard any pairs that do not satisfy this condition. After this sieving, approximately $m \times 2^{27.96} \times 2^{-8} = m \times 2^{19.96}$ ciphertext pairs remain. Next, repeat this guessing procedure for each of the other three bytes of RK_{22}. Compute the corresponding S-box output and input differences $\mathfrak{d}_{22}[i]$ and $\mathfrak{c}_{22}[i]$ for $i = 1$ to 3. At each step, discard pairs whenever $\mathfrak{d}_{22}[i] \neq \mathfrak{c}_{22}[i]$. After these tests, the expected number of surviving pairs is $m \times 2^{19.96} \times 2^{-24} = m \times 2^{-4.04}$.

(2) For each of the four bytes of RK_{21}, guess its value and perform similar sieving steps on the remaining ciphertext pairs, as in Step 3(1). After applying these tests, the number of surviving ciphertext pairs becomes $m \times 2^{-4.04} \times 2^{-32} = m \times 2^{-36.04}$.

(3) Similarly, guess each of the four bytes of RK_{20} and apply the same sieving process. After these tests, the expected number of remaining ciphertext pairs is $m \times 2^{-36.04} \times 2^{-32} = m \times 2^{-68.04}$.

(4) Next, guess $RK_{19}[0]$ and perform sieving similar to Step 3(1). Because of the chosen set Λ'', the number of surviving pairs is $m \times 2^{-68.04} \times \frac{1}{127} = m \times 2^{-75.03}$. For each of $RK_{19}[1]$ and $RK_{19}[2]$, guess its value and perform the same sieving as in Step 3(1). Each guess further reduces the number of pairs by a factor of $\frac{1}{127}$. After both guesses, the expected number of remaining pairs is $m \times 2^{-75.03} \times \frac{1}{127^2} = m \times 2^{-89.01}$. Since the last S-box in the 19-th round is inactive, there is no information available to recover $RK_{19}[3]$.

Step 4. After completing the sieving steps, any 120-bit subkey guess (RK_{19}[0-2], $RK_{20}, RK_{21}, RK_{22}$) that leaves at least a threshold τ number of ciphertext pairs is considered a candidate key. For each candidate key, guess the missing byte $RK_{19}[3]$. Then, use at most two plaintext-ciphertext pairs to verify each fully reconstructed key by encrypting or decrypting with the guessed subkey values. Finally, the correct key is uniquely determined as the one that reproduces the observed ciphertexts or plaintexts upon encryption or decryption.

Analysis of the Attack Using the Hypothesis Testing Paradigm. Since the probability of the distinguisher is $2^{-123.99}$, the number of remaining pairs for the correct key follows a binomial distribution with parameters (N, p_0), where $N = m \times 2^{38.99}$ and $p_0 = 2^{-123.99}$. For incorrect keys, the number of remaining pairs follows a binomial distribution with parameters (N, p), where $p = 2^{-128}$.

We set the threshold τ to 3 and the success probability to 90%. By applying formulas (1) and (2) for numerical approximation, we obtain $N = 2^{126.40}$ and $\beta = 2^{-7.74}$. Consequently, the number of structures used in the attack is $m = 2^{87.41}$, leading to a data complexity of $2^{120.41}$ chosen plaintexts.

The overall time complexity can be divided into three parts. $T_1 = m \times 2^{33}$ is the cost of generating ciphertexts from the $m \times 2^{33}$ plaintexts. $T_2 = 13 \times m \times 2^{27.96} \times 2^8 \times 2 \times \frac{1}{4} \times \frac{1}{23} + m \times 2^{27.96} \times \frac{256}{127} \times 2^8 \times 2 \times \frac{1}{4} \times \frac{1}{23} + m \times 2^{27.96} \times (\frac{256}{127})^2 \times 2^8 \times 2 \times \frac{1}{4} \times \frac{1}{23}$ accounts for enumerating 120 bits of the round keys. $T_3 = 2^{128} \times \beta$ is the exhaustive search over the remaining key candidates plus the unguessed 8-bit subkey. Summing these yields $T_1 + T_2 + T_3 = 2^{122.77}$ 23-round encryptions. The memory required for the attack is primarily used to store the remaining $m \times 2^{27.96}$ ciphertext pairs in the second step. Therefore, the memory complexity of the attack is approximately $m \times 2^{28.96} = 2^{116.37}$ blocks.

As summarized in Table 1, although our attack uses slightly more data than the 23-round attack of Su *et al.* [19], it achieves lower time and memory complexities while ensuring at least a 90% success probability.

4.2 Key-Recovery Attack Using the Method from [26]

In the attack described in [26], the differential key recovery framework for partial substitution-permutation networks proposed in [1] is employed. Instead of performing exhaustive key enumeration, the primary idea is to generate key suggestions for each ciphertext pair. During the preprocessing phase, two matrices are constructed. The method for constructing matrices A and B is detailed in Appendix B. With these matrices prepared, the key-recovery process proceeds as follows.

Step I. By selecting m structures, we generate approximately $m \times 2^{38.99}$ plaintext pairs from $m \times 2^{33}$ plaintexts.

Step II. For each pair $(P, \hat{P})$, compute the difference of the corresponding ciphertexts $(C, \hat{C})$ and check whether the first word of the ciphertext difference is in Λ''. If it is not, discard the pair. After this filtering, around $m \times 2^{38.99} \times 2^{-11.03} = m \times 2^{27.96}$ pairs are expected to remain.

Step III. For each remaining pair $(C, \hat{C})$, perform the following steps.

(1) Calculate $(\Delta Z_{19}, \Delta Z_{20}, \Delta Z_{21}, \Delta Z_{22})$ using matrix A from equation (6), which enables the computation of the input differences ΔY_i and output differences ΔZ_i of the nonlinear transformation $\mathcal{S}$ in rounds 19 to 22.

(2) For each of the 12 active S-boxes in rounds 20 to 22, check whether the differential transitions are possible according to the DDT; if any transition is impossible, discard this pair and return to Step III to analyze the next pair. In view of Property 1, the sieving probability for each active

S-box is $\frac{127}{256}$, so after these checks, the expected number of surviving ciphertext pairs becomes $m \times 2^{27.96} \times \left(\frac{127}{256}\right)^{12} = m \times 2^{15.82}$.

(3) Based on the DDT, construct a list List of vectors containing all possible values for $(Y_{19}[0\text{-}2], Y_{20}, Y_{21}, Y_{22}, Z_{19}[0\text{-}2], Z_{20}, Z_{21}, Z_{22})$, representing the inputs and outputs of all 15 active S-boxes in rounds 19 to 22. Since each S-box input and output occupies 16 bits, each of these vectors comprises 240 bits.

(4) For each 240-bit vector in List, combine it with the known ciphertext C, apply matrix B (see equation (7)), and derive a candidate 120-bit key $(RK_{19}[0\text{-}2], RK_{20}, RK_{21}, RK_{22})$. Then guess the missing byte $RK_{19}[3]$. Verify each candidate by testing at most two plaintext-ciphertext pairs; if the correct ciphertexts are reproduced, output the 128-bit key.

Analysis of the Attack Using the Key Ranking Paradigm. In the key-ranking framework, let $p = 2^{-128}$, $p_0 = 2^{-123.99}$, and the total number of possible keys be $\kappa = 2^{120}$. Since each possible differential transition of the S-box yields an average of $\frac{2\times126+4}{127} = 2^{1.01}$ correct pairs, the number of remaining key candidates is $\ell = m \times 2^{15.82} \times 2^{1.01\times15} = m \times 2^{30.97}$. By setting $N = 2^{122.83}$ and applying formula (5), we determine that the attack's success probability is approximately $P_S = 99.99\%$. From $N = m \times 2^{38.99}$, we obtain $m = 2^{83.84}$, giving a data complexity of $2^{116.84}$ chosen plaintexts.

The overall time complexity can be divided into four parts. $T_1 = m \times 2^{33}$ is the time required to encrypt $m \times 2^{33}$ plaintexts. T_2 is the time required to use matrix A to derive input and output differences for the nonlinear transformation $\mathcal{S}$. Following the approach in [26], we treat one computation of expression (6) as a single round of SM4 encryption. Thus, $T_2 = m \times 2^{27.96} \times \frac{1}{23}$ 23-round encryptions. T_3 is the time needed to use matrix B to obtain 120-bit round-key candidates. We treat one computation of expression (7) as four rounds of SM4 encryption. Hence, $T_3 = m \times 2^{30.97} \times \frac{4}{23}$ 23-round encryptions. T_4 is the time for brute-forcing the correct key, giving $T_4 = m \times 2^{30.97} \times 2^8$ 23-round encryptions. Summing these contributions, we obtain $T_1 + T_2 + T_3 + T_4 = 2^{122.84}$ 23-round encryptions. The memory complexity is below 2^{10} words of 128 bits, primarily for storing matrices A and B, since elements of List can be generated on-the-fly.

As summarized in Table 1, although our attack uses more data than the 23-round attack of Zhao *et al.* [26], it achieves lower time complexity and a higher success probability.

5 Conclusion

In this work, we revisit the 19-round differential characteristics of SM4, motivated by the observation that previously published optimal characteristics do not necessarily yield practical efficiency in key-recovery attacks. To address this shortcoming, we employ an automated search method to examine the properties of all optimal 19-round differential characteristics of SM4. Through Boolean

encodings and comprehensive exploration, we identify three structural properties common to all such optimal characteristics and highlight four representative characteristics that are particularly well-suited for key recovery. Our detailed analysis of one such characteristic reveals that although it requires slightly more data compared to certain existing 23-round attacks, it achieves lower time and memory complexities under both the hypothesis testing and key ranking paradigms. These findings not only deepen our understanding of SM4's differential structure but also offer a strong foundation for future research.

Acknowledgements The research leading to these results has received funding from the Shan-dong Key Research and Development Program (Grant No. 2024ZLGX05), the National Natural Science Foundation of China (Grant No. 62272273, Grant No. 62002201, Grant No. 62032014), and the National Cryptologic Science Fund of China (Grant No. 2025NCSF02006). Ling Sun gratefully acknowledges the support by the Program of TaiShan Scholars Special Fund for young scholars (Grant No. tsqn202306043) and Xiaomi Young Talents Program.

References

1. Bar-On, A., Dinur, I., Dunkelman, O., Lallemand, V., Keller, N., Tsaban, B.: Cryptanalysis of SP networks with partial non-linear layers. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I. Lecture Notes in Computer Science, vol. 9056, pp. 315–342. Springer (2015). https://doi.org/10.1007/978-3-662-46800-5_13
2. Biere, A., Faller, T., Fazekas, K., Fleury, M., Froleyks, N., Pollitt, F.: CaDiCaL, Gimsatul, IsaSAT and Kissat entering the SAT Competition 2024. In: Heule, M., Iser, M., Järvisalo, M., Suda, M. (eds.) Proc. of SAT Competition 2024 – Solver, Benchmark and Proof Checker Descriptions. Department of Computer Science Report Series B, vol. B-2024-1, pp. 8–10. University of Helsinki (2024)
3. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. In: Menezes, A.J., Vanstone, S.A. (eds.) CRYPTO'90. LNCS, vol. 537, pp. 2–21. Springer, Berlin, Heidelberg (Aug 1991). https://doi.org/10.1007/3-540-38424-3_1
4. Biham, E., Shamir, A.: Differential cryptanalysis of the full 16-round DES. In: Brickell, E.F. (ed.) CRYPTO'92. LNCS, vol. 740, pp. 487–496. Springer, Berlin, Heidelberg (Aug 1993). https://doi.org/10.1007/3-540-48071-4_34
5. Blondeau, C., Gérard, B., Tillich, J.: Accurate estimates of the data complexity and success probability for various cryptanalyses. Des. Codes Cryptogr. **59**(1-3), 3–34 (2011). https://doi.org/10.1007/S10623-010-9452-2
6. Brayton, R.K., Hachtel, G.D., McMullen, C.T., Sangiovanni-Vincentelli, A.L.: Logic Minimization Algorithms for VLSI Synthesis, The Kluwer International Series in Engineering and Computer Science, vol. 2. Springer (1984). https://doi.org/10.1007/978-1-4613-2821-6
7. Diffie, W., (translators), G.L.: SMS4 encryption algorithm for wireless networks. Cryptology ePrint Archive, Report 2008/329 (2008), https://eprint.iacr.org/2008/329

8. Etrog, J., Robshaw, M.J.B.: The cryptanalysis of reduced-round SMS4. In: Avanzi, R.M., Keliher, L., Sica, F. (eds.) SAC 2008. LNCS, vol. 5381, pp. 51–65. Springer, Berlin, Heidelberg (Aug 2009). https://doi.org/10.1007/978-3-642-04159-4_4
9. Kim, T., Kim, J., Hong, S., Sung, J.: Linear and differential cryptanalysis of reduced SMS4 block cipher. Cryptology ePrint Archive, Report 2008/281 (2008), https://eprint.iacr.org/2008/281
10. Li, B., Sun, L.: Exploring the optimal differential characteristics of SM4 (full version): Improving automatic search by including human insights. Cryptology ePrint Archive, Paper 2024/2003 (2024), https://eprint.iacr.org/2024/2003
11. Li, L., Wu, W., Zhang, L., Zheng, Y.: New method to describe the differential distribution table for large S-boxes in MILP and its application. IET Inf. Secur. **13**(5), 479–485 (2019). https://doi.org/10.1049/IET-IFS.2018.5284
12. Liu, F., Ji, W., Hu, L., Ding, J., Lv, S., Pyshkin, A., Weinmann, R.P.: Analysis of the SMS4 block cipher. In: Pieprzyk, J., Ghodosi, H., Dawson, E. (eds.) ACISP 07. LNCS, vol. 4586, pp. 158–170. Springer, Berlin, Heidelberg (Jul 2007). https://doi.org/10.1007/978-3-540-73458-1_13
13. Liu, Y., Liang, H., Li, M., Huang, L., Hu, K., Yang, C., Wang, M.: STP models of optimal differential and linear trail for S-box based ciphers. Cryptology ePrint Archive, Report 2019/025 (2019), https://eprint.iacr.org/2019/025
14. Liu, Y., Liang, H., Li, M., Huang, L., Hu, K., Yang, C., Wang, M.: STP models of optimal differential and linear trail for s-box based ciphers. Sci. China Inf. Sci. **64**(5) (2021). https://doi.org/10.1007/S11432-018-9772-0
15. Matsui, M.: On correlation between the order of S-boxes and the strength of DES. In: Santis, A.D. (ed.) Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings. Lecture Notes in Computer Science, vol. 950, pp. 366–375. Springer (1994). https://doi.org/10.1007/BFB0053451
16. Rijmen, V., Daemen, J.: Advanced encryption standard. Proceedings of federal information processing standards publications, national institute of standards and technology **19**, 22 (2001)
17. Sinz, C.: Towards an optimal CNF encoding of boolean cardinality constraints. In: van Beek, P. (ed.) Principles and Practice of Constraint Programming - CP 2005, 11th International Conference, CP 2005, Sitges, Spain, October 1-5, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3709, pp. 827–831. Springer (2005). https://doi.org/10.1007/11564751_73
18. Soos, M., Nohl, K., Castelluccia, C.: Extending SAT solvers to cryptographic problems. In: Kullmann, O. (ed.) Theory and Applications of Satisfiability Testing - SAT 2009, 12th International Conference, SAT 2009, Swansea, UK, June 30 - July 3, 2009. Proceedings. Lecture Notes in Computer Science, vol. 5584, pp. 244–257. Springer (2009). https://doi.org/10.1007/978-3-642-02777-2_24
19. Su, B., Wu, W., Zhang, W.: Differential cryptanalysis of SMS4 block cipher. Cryptology ePrint Archive, Report 2010/062 (2010), https://eprint.iacr.org/2010/062
20. Sun, L., Wang, M.: SoK: Modeling for large S-boxes oriented to differential probabilities and linear correlations. IACR Trans. Symm. Cryptol. **2023**(1), 111–151 (2023). https://doi.org/10.46586/tosc.v2023.i1.111-151
21. Sun, L., Wang, W., Wang, M.: Accelerating the search of differential and linear characteristics with the SAT method. IACR Trans. Symm. Cryptol. **2021**(1), 269–315 (2021). https://doi.org/10.46586/tosc.v2021.i1.269-315

22. Wang, S., Feng, D., Hu, B., Guan, J., Zhang, K., Shi, T.: New method for combining Matsui's bounding conditions with sequential encoding method. Des. Codes Cryptogr. **91**(11), 3603–3642 (2023). https://doi.org/10.1007/S10623-023-01259-9
23. Zhang, J., Wu, W., Zheng, Y.: Security of SM4 against (related-key) differential cryptanalysis. In: Bao, F., Chen, L., Deng, R.H., Wang, G. (eds.) Information Security Practice and Experience - 12th International Conference, ISPEC 2016, Zhangjiajie, China, November 16-18, 2016, Proceedings. Lecture Notes in Computer Science, vol. 10060, pp. 65–78 (2016). https://doi.org/10.1007/978-3-319-49151-6_5
24. Zhang, L., Zhang, W., Wu, W.: Cryptanalysis of reduced-round SMS4 block cipher. In: Mu, Y., Susilo, W., Seberry, J. (eds.) ACISP 08. LNCS, vol. 5107, pp. 216–229. Springer, Berlin, Heidelberg (Jul 2008). https://doi.org/10.1007/978-3-540-70500-0_16
25. Zhang, W., Wu, W., Feng, D., Su, B.: Some new observations on the SMS4 block cipher in the chinese WAPI standard. In: Bao, F., Li, H., Wang, G. (eds.) Information Security Practice and Experience, 5th International Conference, ISPEC 2009, Xi'an, China, April 13-15, 2009, Proceedings. Lecture Notes in Computer Science, vol. 5451, pp. 324–335. Springer (2009). https://doi.org/10.1007/978-3-642-00843-6_28
26. ZHAO Yan-Min, L.Y., Mei-Qin, W.: Improved differential attack on 23-round SMS4. Journal of Software **29**(9), 2821 (2018). https://doi.org/10.13328/j.cnki.jos.005271

Appendix

A Two SAT Models for the Objective Function

Model 5 is based on the sequential encoding method described in [17], while Model 6 is the enhanced model proposed by Wang *et al.* [22].

Model 5 (Sequential Encoding Method, [17]) *To encode the cardinality inequality* $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$*, auxiliary Boolean variables* $\sigma_{i,j}$ *(*$0 \leqslant i \leqslant \gamma-2$*,* $0 \leqslant j \leqslant \varpi-1$*) must be introduced. The constraint* $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$ *holds if and only if all of the following clauses are satisfied simultaneously.*

$$
\begin{aligned}
&\overline{w_0} \vee \sigma_{0,0} = 1 \\
&\overline{\sigma_{0,j}} = 1,\ 1 \leqslant j \leqslant \varpi - 1 \\
&\left.\begin{aligned}
&\overline{w_i} \vee \sigma_{i,0} = 1 \\
&\overline{\sigma_{i-1,0}} \vee \sigma_{i,0} = 1 \\
&\left.\begin{aligned}
&\overline{w_i} \vee \overline{\sigma_{i-1,j-1}} \vee \sigma_{i,j} = 1 \\
&\overline{\sigma_{i-1,j}} \vee \sigma_{i,j} = 1
\end{aligned}\right\} 1 \leqslant j \leqslant \varpi - 1 \\
&\overline{w_i} \vee \overline{\sigma_{i-1,\varpi-1}} = 1
\end{aligned}\right\} 1 \leqslant i \leqslant \gamma - 2. \\
&\overline{w_{\gamma-1}} \vee \overline{\sigma_{\gamma-2,\varpi-1}} = 1
\end{aligned}
$$

Model 6 (Enhanced Sequential Encoding Method, [22]) *The dominated inequality* $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$ *requires introducing auxiliary Boolean variables* $\sigma_{i,j}$ *(*$0 \leqslant i \leqslant \gamma-2$, $0 \leqslant j \leqslant \varpi-1$*). The constraints* $\sum_{i=0}^{\gamma-1} w_i \leqslant \varpi$ *and* $l_{\gamma'} \leqslant \sum_{i=0}^{\gamma'-1} w_i \leqslant u_{\gamma'}$ *(*$0 \leqslant \gamma' < \gamma$*) are satisfied if and only if all of the following clauses hold.*

if $l_0 = 0$ and $u_0 = 1$ then $\overline{w_0} \vee \sigma_{0,0} = 1$
if $l_0 = u_0 = 0$ then $\overline{w_0} = 1$
if $l_0 = 1$ and $u_0 = 1$ then $w_0 = 1$
if $u_{\gamma'} = 0$ then $\overline{w_{\gamma'}} = 1$
if $u_{\gamma'} > 0$
  if $l_{\gamma'} = 0$ then $\overline{w_{\gamma'}} \vee \sigma_{\gamma',0} = 1$
    if $l_{\gamma'-1} < u_{\gamma'-1}$ then $\overline{\sigma_{\gamma'-1,0}} \vee \sigma_{\gamma',0} = 1$
    if $l_{\gamma'-1} = j$ then
      $\overline{w_{\gamma'}} \vee \sigma_{\gamma',j} = 1$
    if $l_{\gamma'-1} < j$ and $u_{\gamma'-1} \geqslant j$ then
      $\overline{w_{\gamma'}} \vee \overline{\sigma_{\gamma'-1,j-1}} \vee \sigma_{\gamma',j} = 1$
    if $l_{\gamma'-1} \leqslant j$ and $u_{\gamma'-1} \geqslant j+1$ then
      $\overline{\sigma_{\gamma'-1,j}} \vee \sigma_{\gamma',j} = 1$
  } $j \geqslant \max(l_{\gamma'}, 1)$ and $j \leqslant u_{\gamma'} - 1$
  if $u_{\gamma'-1} = u_{\gamma'}$ and $l_{\gamma'-1} < u_{\gamma'}$ then $\overline{w_{\gamma'}} \vee \overline{\sigma_{\gamma'-1,u_{\gamma'-1}}} = 1$
  if $l_{\gamma'-1} = u_{\gamma'}$ then $\overline{w_{\gamma'}} = 1$
} $\gamma' \geqslant 1$ and $\gamma' \leqslant \gamma - 2$.
if $u_{\gamma-1} = 0$ then $\overline{w_{\gamma-1}} = 0$
if $u_{\gamma-1} > 0$
  if $u_{\gamma-2} = u_{\gamma-1}$ and $l_{\gamma-2} < u_{\gamma-1}$ then $\overline{w_{\gamma-1}} \vee \overline{\sigma_{\gamma-2,u_{\gamma-1}-1}} = 1$
  if $l_{\gamma-2} = u_{\gamma-1}$ then $\overline{w_{\gamma-1}} = 1$

According to Wang *et al.*, the upper bound $u_{\gamma'}$ and lower bound $l_{\gamma'}$ for partial sums can be derived by applying Matsui's bounding conditions [15] to solve small-scale mixed integer linear programming (MILP) problems. For more details, see [22].

B Creation of Two Matrices Using the Method from [1]

The first matrix A is used to compute the differential transitions for each ciphertext pair based on the ciphertext differences. Let ΔZ_i denote the output difference of the nonlinear transformation $\mathcal{S}$ in the i-th round. As shown in Figure 3, four linear equations can be derived for rounds $19 \leqslant i < 23$: $\Delta X_i \oplus \mathcal{L}(\Delta Z_i) = \Delta X_{i+4}$. Since the values of ΔX_i for all $19 \leqslant i \leqslant 26$ are known – either from the output difference of the 19-round distinguisher or from

the ciphertext differences – it is possible to calculate ΔZ_{19}, ΔZ_{20}, ΔZ_{21}, and ΔZ_{22} via the equation

$$\begin{bmatrix} \Delta Z_{19} \\ \Delta Z_{20} \\ \Delta Z_{21} \\ \Delta Z_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{L}^{-1} & 0 & 0 & 0 \\ 0 & \mathcal{L}^{-1} & 0 & 0 \\ 0 & 0 & \mathcal{L}^{-1} & 0 \\ 0 & 0 & 0 & \mathcal{L}^{-1} \end{bmatrix}}_{\text{Matrix } A} \times \begin{bmatrix} \Delta C_0 \\ \Delta C_1 \\ \Delta C_2 \\ \Delta C_3 \end{bmatrix} \oplus \begin{bmatrix} \mathcal{L}^{-1}(\Delta X_{19}) \\ \mathcal{L}^{-1}(\Delta X_{20}) \\ \mathcal{L}^{-1}(\Delta X_{21}) \\ \mathcal{L}^{-1}(\Delta X_{22}) \end{bmatrix}, \tag{6}$$

where $\mathcal{L}^{-1}$ is the matrix representation of the inverse of $\mathcal{L}$. Consequently, the matrix A that we seek to construct is the block diagonal matrix in expression (6), where each diagonal element is $\mathcal{L}^{-1}$.

The second matrix B is used to generate key suggestions based on the ciphertexts and the differential transitions. Let ΔY_i denote the input difference to the nonlinear transformation $\mathcal{S}$ in the i-th round, and let Y_i and Z_i be the input and output values of $\mathcal{S}$ in the i-th round, respectively. After determining the differential transitions $(\Delta Y_i, \Delta Z_i)$ for all rounds $19 \leqslant i < 23$, the values of Y_i and Z_i become almost uniquely determined (ignoring the ciphertext order within each pair). Referring to the cipher structure in Figure 3, we have the following equations for $0 \leqslant i < 4$:

$$\begin{aligned} RK_{i+19} &= X_{i+20} \oplus X_{i+21} \oplus X_{i+22} \oplus Y_{i+19}, \\ X_{i+19} &= C_i \oplus \mathcal{L}(Z_{i+19}), \quad X_{i+23} = C_i. \end{aligned}$$

By combining these equations, we can derive expressions for RK_{19}, RK_{20}, RK_{21}, and RK_{22}

$$\begin{bmatrix} RK_{19} \\ RK_{20} \\ RK_{21} \\ RK_{22} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \mathcal{L} & \mathcal{L} & \mathcal{L} & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & \mathcal{L} & \mathcal{L} & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \mathcal{L} & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}}_{\text{Matrix } B} \times \begin{bmatrix} Y_{19} \\ Y_{20} \\ Y_{21} \\ Y_{22} \\ Z_{19} \\ Z_{20} \\ Z_{21} \\ Z_{22} \\ C_0 \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}. \tag{7}$$

The matrix B we construct is the block matrix in expression (7).