Quantum Speed-Up for Multidimensional (Zero Correlation) Linear and Integral Distinguishers

Akinori Hosoyamada

NTT Social Informatics Laboratories, Tokyo, Japan
akinori.hosoyamada.bh@hco.ntt.co.jp

Abstract. This paper shows how to achieve quantum speed-up for multidimensional (zero correlation) linear and integral distinguishers. To understand post-quantum security of symmetric-key cryptosystems, it is important to study how much quantum speed-up we can obtain for classical cryptanalytic techniques such as differential, linear, and integral cryptanalysis. A previous work by Kaplan et al. already showed a quantum quadratic speed-up for one-dimensional linear distinguishers, but it is unclear how to extend their technique to multidimensional linear distinguishers. To remedy this, we investigate how to speed-up multidimensional linear distinguishers in the quantum setting. Firstly, we observe that there is a close relationship between the subroutine of Simon's algorithm and linear correlations via Fourier transform, and a slightly modified version of Simon's subroutine can be used to speed-up multidimensional linear distinguishers. The modified Simon's subroutine also leads to speed-ups for multidimensional zero correlation and some integral distinguishers. Surprisingly, our technique achieves more-than-quadratic speed-ups for some special types of integral distinguishers. This is because the modified Simon's subroutine can exploit the existence of *multiple* multidimensional zero correlation linear approximations. Our attacks are the first examples achieving such speed-up on classical cryptanalytic techniques without relying on any algebraic structures such as hidden periods or shifts. The speed-ups for multidimensional (zero correlation) linear distinguishers are at-most-quadratic, and all of our attacks require quantum superposition queries.

Keywords: symmetric-key cryptography · quantum cryptanalysis · linear cryptanalysis · integral cryptanalysis · more-than-quadratic speed-up

1 Introduction

Researches in the past decade have revealed possible quantum attacks on symmetric cryptosystems are not limited to the exhaustive key search with Grover's algorithm [28] or the collision search by the BHT algorithm [17]. A notable line of researches is the one initiated by Kuwakado and Morii showing that Simon's algorithm breaks lots of classically secure schemes in polynomial time [43, 44, 39, 12]. Other previous works shows how to speed-up classical cryptanalytic techniques such as differential and linear cryptanalysis, MITM, and integral attacks [40, 34, 14], and some recent papers study dedicated quantum collision attacks on concrete hash functions such as SHA-2 and SHA-3 [35, 26, 36, 29].

Although many interesting attacks have been found in recent studies, the field of quantum cryptanalysis is still far from mature and many questions are yet to be resolved, especially for symmetric-key cryptosystems.

Q1 and Q2 Models. For quantum cryptanalysis on symmetric cryptosistems, there are two attack models called Q1 and Q2 [40]. The Q1 model assumes the existence of a quantum computer but the oracles given to attackers are classical. Meanwhile, Q2 assumes that not only attackers' computers but also oracles are quantum. Namely, the quantum encryption oracle of a target cipher is given to an attacker and the attacker can make quantum superposition queries to the oracle. Such attacks are called Quantum Chosen-Plaintext Attacks (QCPAs). If an attack assumes not only the quantum encryption oracle but also the quantum decryption oracle, it is called a Quantum Chosen-Ciphertext Attack (QCCA). All the quantum attacks in this papers are in the Q2 model unless otherwise noted.

Significance of Studying Q2 Attacks. Q1 model is more realistic than Q2 model in that oracles in the Q1 model are the same as classical ones and attacks in Q1 model becomes a real threat as soon as a large-scale fault-tolerant quantum computer is available. Still, we argue studying not only Q1 model but also Q2 model is significant due to the following two reasons. First, a new non-trivial Q1 attack may be found based on Q2 attacks. For instance, the so-called offline Simon's algorithm presented by Bonnetain et al. [11] is a Q1 attack but is developed by carefully modifying the Q2 attack by Leander and May [45]. Second, Q2 attacks can be converted into Q1 attacks when key length is sufficiently long. Let E_K be an n-bit block cipher with k-bit keys. Suppose that $k > 2n$, and that there is a Q2 attack on E_K with time complexity $T < 2^{k/2}$. Now, assume we are in the Q1 model, and query all the (classical) inputs to E_K and store the results in a qRAM. Then we can simulate the quantum oracle of E_K by accessing the qRAM. Especially, we can run the Q2 attack by using the simulated quantum oracle. This is still a valid Q1 attack since now we are assuming $2n < k$ and the resulting complexity $T' = \max\{T, 2^n\}$ is less than $2^{k/2}$, the complexity of the exhaustive key-search by Grover's algorithm. Even if $2n \geq k$, some Q2 attacks may similarly be converted into Q1 if they require quantum superposition query only on a part of inputs.

Multidimensional Linear Cryptanalysis. Linear cryptanalysis [47] is one of the most fundamental cryptanalytic techniques in symmetric cryptology, and Kaplan et al. [40] already showed how to achieve a quadratic speed-up for linear attacks. However, Kaplan et al.'s approach is applicable only for one-dimensional linear approximations, while it is common to exploit multidimensional linear approximations in the classical setting [30–32]. To be more precise, it is unclear whether we can speed-up Kaplan et al.'s one-dimensional attack further even if multiple linear approximations are available. Thus it is natural to ask whether there exists a quantum linear attack that performs better than Kaplan et al.'s when a multidimensional linear approximation is available.

Quadratic Barrier. As mentioned before, the Grover search provides a quadratic speed-up for the complexity of exhaustive key search. More specifically, the search on a k-bit key of a block cipher can be done with time complexity equivalent to $\approx \frac{\pi}{2} 2^{k/2}$ encryptions[1]. This means that, if we find a dedicated quantum attack on a cipher and its time complexity is less than the square root of the corresponding classical complexity[2], then the quantum attack is especially interesting because it must exploit internal structure of the cipher in a non-trivial and non-classical way. Finding such a quantum attack achieving a more-than-quadratic speed-up in some sense has been one of the main goals in studying quantum cryptanalysis on symmetric-key cryptosystems.

Indeed, some previous works achieve such non-trivial speed-up. However, the types of such attacks are limited: To achieve more-than-quadratic speed-up, all of them exploit algebraic structures such as *hidden periods* or *shifts* of target ciphers by using Simon's algorithm or related algorithms to solve algebraic problems.

Moreover, few previous works have succeeded to achieve more-than-quadratic speed-up on classical cryptanalytic techniques such as differential, linear, or integral cryptanalysis. The only one exception is the quantum versions of (advanced) slide attacks [39, 13], but the speed-up also relies on special algebraic structures like hidden periods. Whether a more-than-quadratic speed-up is possible for other major classical techniques without relying on algebraic structures has been an important open problem for years.

1.1 Our Contributions

This paper shows quantum speed-up for multidimensional (zero correlation) linear and integral distinguisher can be achieved by using a modified version of the subroutine of Simon's algorithm, without exploiting algebraic structures such as hidden periods or shifts. Especially, we show that some special versions of integral distinguishers achieve more-than-quadratic speed-up.

First, we observe that Simon's algorithm has a close relationship with linear correlations of functions via Fourier transform. Simon's algorithm iterates a subroutine, which is composed of the Hadamard transform and an oracle query to the target function. We observe that, after a slight modification is made, the subroutine outputs a pair of linear masks of the target function with probability proportional to the squared linear correlation. We call the subroutine after the modification the *modified Simon's subroutine.*

[1] Here we are assuming the block length n matches k. If $k > n$, more operations are required.

[2] The complexity of the classical attack that the quantum attack is based on, or jsut the best classical complexity.

Second, we show multidimensional linear distinguishers can be sped-up by the modified Simon's subroutine. By combining the Quantum Amplitude Amplification (QAA) technique, we achieve an at-most-quadratic speed-up from classical complexity. As an application example, we see how much speed-up we can obtain for the multidimensional linear distinguishers on FEA-1 and FEA-2 by Beyne [6].

Then we show that an at-most-quadratic speed-up for multidimensional zero correlation linear distinguishers can be obtained similarly. Our technique leads to quantum distinguishers on 5-round balanced Feistel running in time $O(2^{n/2})$ when round functions are bijections and the entire width of the cipher is n, and distinguishers on some Type-I/II generalized Feistel structures. (See Table 2 in the appendix for details.)

Finally, we show how to speed-up integral distinguishers. In fact, our technique is applicable only when distinguishers are based on balanced functions and not zero-sum properties. As shown by Bogdanov et al. and Sun et al. [8, 57], distinguishers based on balanced functions correspond to a class of multidimensional zero correlation linear distinguishers. Our speed-up for integral distinguishers is obtained via this correspondence. Moreover, we observe that some integral distinguishers including the ones on 2.5 or 3.5-round AES yield *multiple* mutually orthogonal multidimensional zero correlation linear approximations. By exploiting such approximations with the modified Simon's subroutine, we can achieve a more-than-quadratic speed-up. As a notable example, a toy 4-bit-cell SPN cipher having the same integral property as the 2.5-round AES is distinguished only by a *single* quantum query. Such single-query attack seems almost impossible in the classical setting, and our technique can be regarded as a new type of quantum speed-up exploiting linear correlations that has not been observed before.

Note that all of our attacks do not require the target cipher to algebraic structures such as hidden periods or shifts. It is somewhat surprising that Simon's algorithm, which is developed to solve algebraic problem of hidden periods, can be used to obtain a super-quadratic speed-up for classical attacks that do not rely on algebraic structures.

Our technique extends to generalized linear distinguishers on arbitrary finite groups [4] in a straightforward manner by replacing the Hadamard transform in the (modified) Simon's subroutine with general quantum Fourier transform. For insatance, we can also achieve at-most-quadratic speed-up for the generalized linear distinguisher on the FF3-1 structure by Beyne [6].

A drawback of our techniques is that integral distinguishers based on zero-sum properties are not sped-up, although usually zero-sum properties are used to extend distinguishers into key-recovery attacks on more rounds. Especially, it seems hard to achieve a more-than-quadratic speed-up for integral *key-recovery* attacks with our techniques. Still, we believe our techniques are novel and will inspire other new types of quantum attacks on symmetric cryptosystems in both of the Q1 and Q2 models.

1.2 Related Works

Recently, Shi et al. published a paper titled with “Quantum zero correlation linear cryptanalysis” [55]. Their work is mainly on how to find zero correlation linear approximations of ciphers by using quantum computers, and does not have much overlap with our results.

1.3 Organization

Section 2 introduces basic notions used throughout the paper. Section 3 studies relationships between the (modified) Simon’s subroutine and linear correaltions. Sections 4 and 5 show how to apply the modified Simon’s subroutine to gain quantum speed-up for multidimensional linear and zero correlation linear distinguishers. Section 6 shows speed-up for integral distinguishers. Secion 7 discusses on extensions to key-recovery attacks and limitations of our techniques.

2 Preliminaries

$\mathbb{F}_2$ denotes the Galois field of order two. We identify the set of n-bit strings $\{0,1\}^n$ and the n-dimensional $\mathbb{F}_2$-vector space $\mathbb{F}_2^n$. By $\mathbf{e}_i$ we denote the vector of $\mathbb{F}_2^n$ of which i-th entry is 1 and other entries are 0. $x \oplus y$ denotes the addition of x and y in $\mathbb{F}_2^n$. The additive group of $\mathbb{F}_2^n$ is isomorphic to $\mathbb{Z}_2^n := (\mathbb{Z}/2\mathbb{Z})^{\oplus n}$. We sometimes use the symbol $\mathbb{Z}_2^n$ instead of $\mathbb{F}_2^n$ to emphasize we focus on the additive structure. For a bit string $x \in \mathbb{F}_2^n = \{0,1\}^n$, we denote the i-th bit (from the left) by x_i. Namely we represent x as $x = x_1 || \cdots || x_n$. For $x, y \in \mathbb{F}_2^n$, the dot product of x and y is defined by $x \cdot y := (x_1 \cdot y_1) \oplus \cdots \oplus (x_n \cdot y_n)$. For a vector space $V \subset \mathbb{F}_2^n$ (resp., vector x), $V^\perp$ (resp., $x^\perp$) denotes the subspace that is composed of y satisfying $y \cdot x = 0$ for all $x \in V$ (resp., y satisfying $y \cdot x = 0$). For two vector spaces $V_1, V_2 \subset \mathbb{F}_2^n$, we write $V_1 \perp V_2$ if $v_1 \cdot v_2 = 0$ for all $v_1 \in V_1$ and $v_2 \in V_2$. The event that a (classical or quantum) algorithm $\mathcal{A}$ outputs a classical bit string x is denoted by $x \leftarrow \mathcal{A}$. For a bit string $x \in \mathbb{F}_2^n$ (resp., function $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$), by $\mathsf{msb}_u[x]$ (resp., $\mathsf{msb}_u[f]$) we denote the most significant u bits of x (resp., the function that returns $\mathsf{msb}_u[f(x)]$ for each input x). The noations $\mathsf{lsb}_u[x]$ and $\mathsf{lsb}_u[f]$ are similarly defined for least significant u bits. In cryptanalysis of a block cipher E, we regard the unit of time is the time to encrypt a message by E.

2.1 Linear Approximations and Correlations

The (one-dimensional) linear approximation of a function $f : \mathbb{F}_2^m \to \mathbb{F}_2^n$ for an input mask $\alpha \in \mathbb{F}_2^m$ and output mask $\beta \in \mathbb{F}_2^n$ is the Boolean function defined by $x \mapsto (\alpha \cdot x) \oplus (\beta \cdot f(x))$. The correlation $\mathrm{Cor}(f; \alpha, \beta)$ of this linear approximation is defined by $\mathrm{Cor}(f; \alpha, \beta) := \Pr_x[\alpha \cdot x = \beta \cdot f(x)] - \Pr_x[\alpha \cdot x \neq \beta \cdot f(x)]$. It is well-known that the linear correlation satisfies

$$\mathrm{Cor}(f; \alpha, \beta) = \sum_{x \in \mathbb{F}_2^m} \frac{(-1)^{\alpha \cdot x \oplus \beta \cdot f(x)}}{2^m}. \tag{1}$$

A multidimensional linear approximation of f is the set of 2^d linear approximations of which masks form a d-dimensional linear subspace of $\mathbb{F}_2^m \times \mathbb{F}_2^n$ $(d \geq 0)$.

The following property is useful to analyse linear correlations.

$$\sum_{x \in \{0,1\}^n} (-1)^{\alpha \cdot x} = \begin{cases} 2^n & \text{if } \alpha = 0^n \\ 0 & \text{if } \alpha \neq 0^n \end{cases}$$

Throughout the paper, we will use this property without any mention.

2.2 Quantum Computation

We assume that the readers are familiar with quantum computation and linear algebra (see, e.g., [52] for basics of quantum computation). We adopt the standard quantum circuit model and do not take the cost of quantum error correction into account. I_m denotes the identity operator on an m-qubit system and H denotes the (1-qubit) Hadamard transform. For a function $f : \{0,1\}^m \to \{0,1\}^n$, U_f denotes the unitary operator defined by $U_f : |x\rangle |y\rangle \mapsto |x\rangle |y \oplus f(x)\rangle$. Namely, U_f is the quantum oracle of f. All quantum attacks in this paper are Quantum Chosen-Plaintext Attacks (QCPAs, in the Q2 model) assuming that the quantum encryption oracle U_{E_K} of a target cipher E_K is available. If E_K is a tweakable block cipher, we assume adversaries query tweaks also in quantum superposition.

Quantum Amplitude Amplification. Here we recall the Quantum Amplitude Amplification (QAA) technique [16], which is a generalization of Grover's algorithm [28]. Let $f : \{0,1\}^m \to \{0,1\}$ be a Boolean function, U be a unitary operator acting on an m-qubit system, and p denote the probability that we observe a bit string x satisfying $f(x) = 1$ when we measure the state $U |0^m\rangle$ by the computational basis. In addition, let $\mathcal{S}_f$ and $\mathcal{S}_0$ be the unitary operators acting on an m-qubit quantum system defined by $\mathcal{S}_f |x\rangle = (-1)^{f(x)} |x\rangle$ and $\mathcal{S} |x\rangle = (-1)^{\delta_{0^m,x}} |x\rangle$, where $\delta_{0^m,x}$ is Kronecker's delta.

Proposition 1 (Quantum amplitude amplification). *Given the above situation, let* $Q(U, f) := -U\mathcal{S}_0 U^* \mathcal{S}_f$*. When the state* $Q(U, f)^i U |0^m\rangle$ *is measured by the computational basis for some* $i > 0$*, a bit string* x *satisfying* $f(x) = 1$ *is obtained with probability* $\sin^2((2i + 1) \cdot \arcsin(\sqrt{p}))$*. Especially, we obtain such* x *with probability at least* $\max\{p, 1 - p\}$ *by setting* $i = \lceil \pi/4 \arcsin(\sqrt{p}) \rceil$*.*

We obtain Grover's algorithm when $\mathcal{A} = H^{\otimes m}$. In this case $p = |f^{-1}(1)|/2^m$ holds and we can find an x satisfying $f(x) = 1$ by applying $Q(H^{\otimes m}, f)$ at most $\sqrt{2^m/|f^{-1}(1)|}$ times.

Applications to Distinguishers. A typical task in cryptanalysis is to distinguish two distributions of functions. That is, under the assumption that a function f is chosen from a distribution D_1 or D_2, an adversary tries to judge which distribution f is chosen from. A typical example is a linear distinguisher where

D_1 corresponds to a linear approximation of a real block cipher and D_2 to the linear approximation a random permutation.

A counterpart of such a task in the quantum setting is to distinguish two distributions of unitary operators. That is, under the assumption that a unitary operator U is an unitary oracle chosen according to a distribution D_1 or D_2, an adversary tries to judge which distribution U is chosen from.

Sometimes QAA is useful to solve such a task. Assume that an adversary has access to not only U but U^*, and that U acts on an n-qubit system. Moreover, suppose that we know a Boolean function $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ satisfying the following condition when we measure the state $U|0^n\rangle$ by the computational basis and observe $x \in \{0,1\}^n$. (1) If U is chosen from D_1, the probability that $F(x) = 1$ is relatively high on average. (2) If U is chosen from D_2, the probability that x such that $F(x) = 1$ is relatively low on average. Specifically, letting $p_U := \Pr\left[x \xleftarrow{\text{measure}} U|0^n\rangle : F(x) = 1\right]$, assume that we know a threshold t satisfying $\mathbb{E}_{U\sim D_1}[p_U] \geq t \gg \mathbb{E}_{U\sim D_2}[p_U]$. Then we can distinguish D_1 and D_2 by using QAA on U and F: Roughly speaking, if U is chosen from D_1, we can expect that QAA with $O(\sqrt{t^{-1}})$ applications of U, U^*, and $\mathcal{S}_F$ will find x satisfying $F(x) = 1$ because $p_U \geq t$ holds on average. If U is chosen from D_2, QAA with only $O(\sqrt{t^{-1}})$ applications of U, U^*, and $\mathcal{S}_F$ will not find such x because $t \gg p_U$ holds on average.

More precisely, since we know only the lower bound of $\mathbb{E}_{U\sim D_1}[p_U]$, we use multiple instances of QAAs with the number of iteration randomized as follows[3]. Let s be any positive integer constant.

QAA for Distinguisher.

1. For $i = 1, \ldots, s$, do:
   (a) Choose i uniformly at random from the set of integers from 0 to $\left\lfloor \frac{1}{\sin\left(2\cdot\arcsin\left(\sqrt{t}\right)\right)} \right\rfloor$.
   (b) Apply $Q(U,F)^i U$ to $|0^n\rangle$ and measure the entire state by the computational basis, and let x be the outcome.
   (c) Compute $F(x)$. If $F(x) = 1$, return 1 and abort.
2. Return 0.

We denote the above algorithm by $\mathcal{A}_0$.

Proposition 2. *With the above setting and notions, suppose* $1/4 > t > 0$. *Then, for any constant s, $\mathcal{A}_0$ applies U, U^*, and $\mathcal{S}_F$ at most $s(\frac{1}{\sqrt{t}} + 1)$ times and (1) returns* 1 *with probability at least* $(1-(\frac{3}{4})^s)\cdot\Pr_{U\sim D_1}[1/4 > p_U \geq t]$ *if U is chosen according to D_1 and (2) returns* 1 *with probability at most* $s\cdot(16t'/t + 20t'/\sqrt{t}) + \Pr_{U\sim D_2}[t' < p_U]$ *for any* $t' > 0$ *satisfying* $4\sqrt{t'/t} + 2\sqrt{t'} < \pi/2$ *if U is chosen according to D_2.*

[3] The idea of randomly choosing the number of iteration follows previous works on the Grover search and QAA without knowing initial success probability [16, 15]. Our algorithm is just a straightforward adaptation of the ideas in these previous works.

The interpretation of the proposition is as follows. Suppose that p_U distributes around t if U is chosen according to D_1 and distribututes around t' if U is chosen according to D_2, and $1/4 > t \gg t'$ holds. Then, for a sufficiently large constant s (e.g., $s = 3$), the proposition guarantees that $\mathcal{A}_0$ returns 1 with probability $\geq 1/2$ when U is chosen according to D_1 while $\mathcal{A}_0$ outputs 1 only with a negligibly small probability when U is chosen according to D_2. Hence D_1 is distinguished from D_2. The proof of Proposition 2 is a straightforward application of some lemmas in previous works [15, 16], though, we provide a proof in Section A in the appendix for completeness.

Simon's algorithm. Simon's quantum algorithm [56] finds a period of a periodic function. More precisely, it solves the following problem.

Problem 1. Let $s \in \{0,1\}^m$ be a (secret) constant, and $f : \{0,1\}^m \rightarrow \{0,1\}^n$ be a function satisfying the following properties.

C1 $f(x \oplus s) = f(x)$ for all x. Namely, f is a periodic function with period s.
C2 $f(x) \neq f(y)$ if $x \neq y$ and $x \oplus s \neq y$.

Given the (quantum) oracle of f, find s.

The classical complexity to solve the problem is $\Theta(2^{m/2})$ but Simon's algorithm, which runs as follows, solves it in polynomial time with high probability.

1. For $i = 1, 2, \ldots, 2m$, execute the following subroutine (a)-(e).
   (a) Prepare the initial state $|0^m\rangle\,|0^n\rangle$.
   (b) Apply the m-qubit Hadamard transform $H^{\otimes m}$ on the first m qubits.
   (c) Apply U_f on the state (i.e., make a quantum query to f).
   (d) Apply the $H^{\otimes m} \otimes I_n$ on the state.
   (e) Measure the first m qubits by the computational basis, discard the remaining n-qubits, and return the observed m-bit string (denoted by α_i).
2. If $\mathrm{Span}_{\mathbb{F}_2^n}(\alpha_1, \ldots, \alpha_{2m}) = m - 1$, compute and output the unique $s' \in \mathbb{F}_2^m$ such that $s' \cdot \alpha_i = 0$ for $i = 1, \ldots, 2m$. If $\mathrm{Span}_{\mathbb{F}_2^n}(\alpha_1, \ldots, \alpha_{2m}) \neq m - 1$, output $\perp$.

Simon showed that each α_i uniformly distributes over the subspace $\{v \in \mathbb{F}_2^m | v \cdot s = 0\}$, and thus the algorithm returns the secret s with high probability. We refer to the subroutine (a)-(e) as Simon's subroutine.

Many papers (e.g., [43, 44, 39]) showed polynomial-time quantum attacks on symmetric cryptosystems by using Simon's algorithm. In fact only C1 is satisfied in those applications and C2 is not necessarily satisfied. Still, C1 guarantees that the subroutine (a)-(e) always returns an α_i satisfying $\alpha_i \cdot s = 0$ [39].

3 New Observation on Simon's Algorithm

As explained in the previous section, the subroutine of Simon's algorithm uses only the quantum oracle of a target function and Hadamard transform, which is a Fourier transform on $\mathbb{Z}_2^n$. Meanwhile, a well-know fact is that linear correlations have strong relationships with Fourier transform. This section observes a link between Simon's subroutine and linear cryptanalysis via Fourier transform.

3.1 Fourier Transform on $\mathbb{Z}_2^n$

First, we recall the Fourier transform on $\mathbb{Z}_2^n$ and its relationship with linear cryptanalysis and quantum computation. The Fourier transform[4] over $\mathbb{Z}_2^n$ of a function $F : \mathbb{F}_2^n \to \mathbb{C}$ is the function $\mathcal{F}F : \mathbb{F}_2^n \to \mathbb{C}$ defined by

$$\mathcal{F}F(x) := \sum_{y \in \mathbb{F}_2^n} \frac{(-1)^{x \cdot F(y)}}{\sqrt{2^n}}.$$

Relationship with Linear Correlations. It is well-known that the linear correlation of an arbitrary function f is obtained by applying the Fourier transform on a function naturally defined from f.

For arbitrary function $f : \mathbb{F}_2^m \to \mathbb{F}^n$, let $f_{\mathrm{emb}} : \mathbb{F}_2^m \times \mathbb{F}_2^n \to \mathbb{C}$ be the function defined by $f_{\mathrm{emb}}(x, y) = 1$ if $f(x) = y$ and $f_{\mathrm{emb}}(x, y) = 0$ otherwise[5]. Then a straightforward calculation shows that

$$\mathcal{F}f_{\mathrm{emb}}(\alpha, \beta) = \sqrt{2^{m-n}} \cdot \mathrm{Cor}(f; \alpha, \beta) \tag{2}$$

holds. This relation plays an important role in distinguishers exploiting multidimensional (zero-correlation) linear approximations [30, 8].

Relationship with Quantum Computation. The relationship with quantum computation is quite clear. The Fourier transform on $\mathbb{F}_2^n$ exactly corresponds to the Hadamard operator $H^{\otimes n}$. For instance, let $\psi : \mathbb{F}_2^n \to \mathbb{C}$ and $|\psi\rangle := \sum_{x \in \mathbb{F}_2^n} \psi(x) |x\rangle$. Then

$$H^{\otimes n} |\psi\rangle = \sum_{y \in \mathbb{F}_2^n} \mathcal{F}\psi(y) |y\rangle \tag{3}$$

holds. (Note that this property holds regardless of the norm of $|\psi\rangle$.) In fact this is one of the most important sources of quantum speed-up: While the classical FFT requires time $O(n2^n)$ to compute the Fourier transform of a function, an application of the Hadamard transform to a quantum state requires time $O(1)$.

3.2 (Modified) Simon's Subroutine and Linear Correlations

Here we show that what (a slightly modified version of) Simon's subroutine does is to return input and output masks for linear approximations with high correlation. First, we show a modified version of Simon's subroutine as follows. The modified parts are underlined. We name the resulting algorithm $\mathcal{L}^f$.

[4] We call this transform "Fourier transform on $\mathbb{Z}_2^n$" but not " Fourier transform on $\mathbb{F}_2^n$" because the latter refers to another operation.

[5] "emb" is an abbreviation of "embedding".

Algorithm $\mathcal{L}^f$: A Modified Simon's Subroutine.

(a) Prepare the initial state $|0^m\rangle\,|0^n\rangle$.
(b) Apply the m-qubit Hadamard transform $H^{\otimes m}$ on the first m qubits.
(c) Apply U_f on the state (i.e., make a quantum query to f).
(d) Apply the $(m+n)$-qubit Hadamard transform $H^{\otimes(m+n)}$ on the state.
(e) Measure the entire $(m+n)$ qubits by the computational basis and return the observed $(m+n)$-bit string $\alpha\|\beta$ ($\alpha \in \{0,1\}^m$ and $\beta \in \{0,1\}^n$).

$\mathcal{L}^f$ is different from the original Simon's subroutine only in that $\mathcal{L}^f$ does not discard the last n qubits and measure them after applying $H^{\otimes n}$.

Note that this change does not affect the distribution of α in Step (e). Especially, α just uniformly distributes over the subspace $\{v \in \mathbb{F}_2^m | v \cdot s = 0\}$ if f satisfies the condition of Problem 1. Thus there is nothing new if we focus only on α. However, we observe that $\mathcal{L}^f$ shows an interesting link to linear cryptanalysis when β is into account, as shown in the following proposition.

Proposition 3. *The quantum state of $\mathcal{L}^f$ before the final measurement is*

$$\sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\mathrm{Cor}(f;\alpha,\beta)}{\sqrt{2^n}} |\alpha\rangle\,|\beta\rangle\,. \tag{4}$$

In particular, for any subset $S \subset \{0,1\}^m \times \{0,1\}^n$,

$$\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^f : (\alpha,\beta) \in S\right] = \sum_{(\alpha,\beta)\in S} \frac{\mathrm{Cor}(f;\alpha,\beta)^2}{2^n} \tag{5}$$

holds.

Proof (of Proposition 3). The quantum state of $\mathcal{L}^f$ before the final measurement is

$$\begin{aligned}
H^{\otimes(m+n)} U_f \left(H^{\otimes m} \otimes I_n\right) |0^m\rangle\,|0^n\rangle &= H^{\otimes(m+n)} U_f \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle\,|0^n\rangle \\
&= H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle\,|f(x)\rangle \\
&\overset{\text{Def. of } f_{\mathrm{emb}}}{=} H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n} \frac{f_{\mathrm{emb}}(x,y)}{\sqrt{2^m}} |x\rangle\,|y\rangle \\
&\overset{\text{Eq. (3)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\mathcal{F}f_{\mathrm{emb}}(\alpha,\beta)}{\sqrt{2^m}} |\alpha\rangle\,|\beta\rangle \\
&\overset{\text{Eq. (2)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\mathrm{Cor}(f;\alpha,\beta)}{\sqrt{2^n}} |\alpha\rangle\,|\beta\rangle\,.
\end{aligned}$$

Hence we have Eq. (4). Eq. (5) immediately follows from Eq. (4). □

Later, by using the above proposition we will show that $\mathcal{L}^f$ leads to speed-up for multidimensional (zero correlation) linear and integral distinguishers.

Some Remarks. $\mathcal{L}^f$ is quite close to the Bernstein-Vazirani algorithm [5] when $n = 1$. Thus $\mathcal{L}^f$ can also be regarded as a generalization of the Bernstein-Vazirani algorithm. Moreover, some previous works [18, 60] already observes similar relationships between linear correlations and the Bernstain-Vazirani algorithm. Still, analysis in previous works is done only in the case of $n = 1$. To obtain speed-up for multidimensional (zero correlation) linear and integral distinguishers, our analysis for general n involving both input and output masks is essential. Furthermore, we observe that a similar relationship holds for generalized linear correlations over arbitrary finite abelian groups and general quantum Fourier transformations. See Section H in the appendix for details.

4 Speeding-Up Multidimensional Linear Distinguishers

This section shows that the modified Simon's subroutine $\mathcal{L}^f$ can be used to achieve at-most-quadratic speed-up for (multidimensional) linear distinguishers. We begin with briefly reviewing the basics of classical linear distinguishers.

Linear Distinguishers. The linear correlation $\mathrm{Cor}(P; \alpha, \beta)$ of an n-bit random permutation P approximately follows the normal distribution $\mathcal{N}(0, 2^{-n})$ for an arbitrary mask (α, β) with $\alpha \neq 0^m$ and $\beta \neq 0^n$ [24]. Thus, if a linear correlation $\mathrm{Cor}(E_K; \alpha, \beta)$ of a block cipher E_K with $\alpha \neq 0^m$ and $\beta \neq 0^n$ significantly deviates from the segment $[-2^{-n/2}, 2^{-n/2}]$ for a random key K, E_K can be distinguished by collecting a list of random plaintext-ciphertex pairs $L = \{(P_1, C_1), \ldots, (P_N, C_N)\}$ and checking if the estimated empirical correlation

$$\widehat{\mathrm{Cor}}(f; \alpha, \beta) = \frac{\#\{(P,C) \in L | \alpha \cdot P = \beta \cdot C\} - \#\{(P,C) \in L | \alpha \cdot P = \beta \cdot C\}}{N}$$

is significantly larger than $2^{-n/2}$ or smaller than $-2^{-n/2}$. Here, the data complexity required for a constant advantage is about $N \approx 1/\mathbb{E}_K[\mathrm{Cor}(E_K; \alpha, \beta)^{-2}]$. This is the basic principle of classical (one-dimensional) linear distinguishers.

4.1 Quantum Linear Distinguisher by Kaplan et al [40].

Kaplan et al already observed that a quadratic quantum speed-up can be obtained for linear distinguishers. Let E_K be an n-bit block cipher and suppose there exists a linear approximation satisfying $|\mathrm{Cor}(E_K; \alpha, \beta)| \gg 2^{-n/2}$ for a random key K. Then, E_K can be distinguished by estimating $M := \#\{x | \alpha \cdot x \oplus \beta \cdot E_K(x) = 0\}$ and testing whether $|M - \frac{2^n}{2}| \gg 2^{n/2}$. To mount a quantum attack based on this idea, Kaplan et al. suggested to use the quantum approximate counting algorithm [16]. The counting algorithm returns an approximation $\tilde{M}$ of M satisfying $|\tilde{M} - M| \leq O\left(\frac{\sqrt{M(2^n - M)}}{q} + \frac{2^n}{q^2}\right)$ in time $O(q)$, by making $O(q)$ quantum queries to E_K In particular, an estimation of $\tilde{M}$ with sufficeint precision for distinguisher ($|\tilde{M} - M| \leq \frac{M}{a}$ for a small integer

$a > 0$) can be obtained in time $O(1/c)$. Compared to the classical complexity of $O(1/c^2)$, a quadratic speed-up is achieved. However, it is unclear whether this approach can be extended to multidimensional linear distinguishsers in such a way that multiple linear approximations lead to further speed-ups.

4.2 Application of the Modfied Simon's Subroutine $\mathcal{L}^f$

Here we show quantum linear distinguisher based on the modified Simon's subroutine $\mathcal{L}^f$. For one-dimensional case, we achieve the same speed-up as Kaplan et al's. We also show quantum versions of multidimensional linear distinguisher for the first time, which achieve at-most-quadratic speed-up from classical attacks.

Recall that what $\mathcal{L}^f$ does is to apply the unitary operator $H^{\otimes(m+n)}U_f H^{\otimes(m+n)}$ on $|0^m\rangle\,|0^n\rangle$ and measure the entire state by the computational basis. By abuse of notation, let $\mathcal{L}^f$ also denote the unitary operator $H^{\otimes(m+n)}U_f H^{\otimes(m+n)}$.

One-dimensional Case. Let E_K be a block cipher and suppose a linear approximation of input-output mask (α, β) has a high correlation c. Consider to apply the modified Simon's subroutine on the Boolean function $\beta \cdot E_K(x)$. When we measure the state $\mathcal{L}^{\beta\cdot E_K}\,|0^m\rangle\,|0^n\rangle$, Proposition 3 guarantees that we observe $(\alpha, 1)$ with probability $c^2/2$. Now, apply the QAA (Proposition 1) with $\mathcal{A} = \mathcal{L}^{\beta\cdot E_K}$ and the Boolean function $F : \mathbb{F}_2^m \times \mathbb{F}_2 \to \mathbb{F}_2$ such that $F(x, y) = 1$ iff $(x, y) = (\alpha, 1)$. When measuring the state $Q(\mathcal{L}^{\beta\cdot E_K}, F)^i\mathcal{L}^{\beta\cdot E_K}\,|0^m\rangle\,|0^n\rangle$ with $i = \lceil \pi/4 \arcsin(c)\rceil \leq O(1/c)$, we obtain $(\alpha, 1)$ with an overwhelming probability.

On the other hand, if the given oracle is a random permutation P instead of E_K, the probability that we observe $(\alpha, 1)$ when we measure $\mathcal{L}^{\beta\cdot P}$ is $O(2^{-n})$. Thus, when applying the QAA and measure the state $Q(\mathcal{L}^{\beta\cdot P}, F)^i\mathcal{L}^{\beta\cdot P}\,|0^m\rangle\,|0^n\rangle$ with $i = \lceil \pi/4 \arcsin(c)\rceil \leq O(1/c)$, the probability that we obtain $(\alpha, 1)$ is negligibly small.

Hence E_K can be distinguished by applying the QAA on $\mathcal{L}^{\beta\cdot E_K}$ and F as above. The number of queries made to E_K is $2 \cdot \lceil \pi/4 \arcsin(c)\rceil + 1 \leq O(1/c)$ and the time complexity is also $O(1/c)$. Thus the complexity of our distinguisher is the the same as that of Kaplan et al.'s. Though the complexity does not change, we believe our distinguisher is theoretically more natural than Kaplan et al.'s using approximate counting because ours directly exploits the linear correlation through the modified Simon's subroutine.

Remark 1. Strictly speaking, the above attack works only if a precise estimation of c is known in advance. This is because the success probability of QAA can be negligibly small not only when the number of iteration i is too small but also when i is too large. If only a lower bound of c is known, a modified version of the QAA applying multiple $Q(\mathcal{L}^{E_K}, F)^i$ for different parameters i should be used [16]. Still, this modification adds only a constant overhead to the query and time complexity. A more precise analysis is given later.

4.3 Extension to Multidimensional Linear Distinguishers

Basics in the Classical Setting. A natural idea to enhance the power of linear cryptanalysis is to utilize multiple linear approximations. Some early works indeed show such attacks, assuming the existence of multiple approximations that are statistically independet [38, 7]. However, the assumption does not necessarily hold in general [48]. Instead, Hermelin et al. [33] proposed to use mulidimensioanl linear appproximations, i.e., sets of linear approximations of which input-output masks form a vector space.

Specifically, let $f : \{0,1\}^m \to \{0,1\}^n$ be a function, $V \subset \mathbb{F}_2^m \times \mathbb{F}_2^n$ be a set of input-output masks for f that is a vector space, and $S := \{(\alpha_1, \beta_1), \dots, (\alpha_\ell, \beta_\ell)\}$ be a basis of V. Then the multidimensional linear approximation of f (w.r.t. (V, S)) is defined as the function $\mathsf{Lin}_S^f : \mathbb{F}_2^m \to \mathbb{F}_2^\ell$ such that

$$\mathsf{Lin}_S^f(x) = (\alpha_1 \cdot x \oplus \beta_1 \cdot \oplus f(x), \dots, \alpha_\ell \cdot x \oplus \beta_\ell \cdot \oplus f(x)).$$

Define a distribution p_S^f on $\mathbb{F}_2^\ell$ by $p_S^f(z) := \Pr_{x \xleftarrow{\$} \mathbb{F}_2^m}\left[\mathsf{Lin}_S^f(x) = z\right]$.

Below we denote the zero vector $(0^m, 0^n)$ by $\mathbf{0}$. We say that the input and output masks are linearly independent if $V = V_1 \times V_2$ holds for some $V_1 \in \mathbb{F}_2^m$ and $V_2 \in \mathbb{F}_2^n$. Moreover, we say that the input and output masks are linearly completely dependent if there exists a basis $\{(\alpha_i, \beta_i)\}_{1 \le i \le \dim(V)}$ of V such that both of $\{\alpha_i\}_{1 \le i \le \dim(V)}$ and $\{\beta_i\}_{1 \le i \le \dim(V)}$ are linearly independent in $\mathbb{F}_2^n$.

The advantage of considering a set of masks forming a vector space is that we can utilize a link of the sum of the squared correlations to the *capacity* of p_S^f and Pearson's chi-squared test: Here, the capacity of a probability function (distribution) p over $\mathbb{F}_2^\ell$ is the value defined[6] by

$$\mathrm{Cap}(p) := 2^\ell \sum_{z \in \mathbb{F}_2^\ell} (p(z) - 2^{-\ell})^2. \tag{6}$$

The important well-known fact is that

$$\mathrm{Cap}(p_S^f) = \sum_{(\alpha,\beta) \in V - \{\mathbf{0}\}} \mathrm{Cor}(f; \alpha, \beta)^2 \tag{7}$$

holds for the multidimensional approximation of f. Moreover, suppose a list of random input-output pairs $L = \{(P_1, C_1), \dots, (P_N, C_N)\}$ is given. Then the capacity $\mathrm{Cap}(\hat{p}_S^f)$ of the estimated empirical distribution $\hat{p}_S^f$ (defined by $\hat{p}_S^f(z) := \frac{\#\{(P,C) \in L | \mathsf{Lin}_S^f(P) = z\}}{N}$) multiplied by N is equal to the test statistic of the Pearson's chi-squared goodness-of-fit test (for testing the goodness-of-fit of p_S^f and the uniform distribution on $\mathbb{F}_n^\ell$).

A rough idea of multidimensional linear distinguishers for a block cipher E_K is that the distribution $p_S^{E_K}$ is far from uniform if the right hand side of Eq. (7)

[6] In fact this is the χ^2-divergence between p and the uniform distribution over $\mathbb{F}_2^\ell$. We use the term *capacity* following previous works on linear cryptanalysis.

with $f = E_K$ is sufficiently large for random K, and thus E_K can be distinguished from random by checking whether the test statistic of the Peason's chi-squared test is larger than a certain threshold. More concretely, given a list of (real) random plaintext-ciphertext pairs $L = \{(P_1, C_1), \ldots, (P_N, C_N)\}$, we count the number $\mathsf{num}(z) := \{(P_i, C_i) \in L | \mathsf{Lin}_S^{E_K}(P_i) = z\}$ for each z, and compute the test statistic $\chi^2_{\mathrm{real}} := N2^\ell \sum_z (\mathsf{num}(z)/N - 2^{-\ell})^2 = N \cdot \mathrm{Cap}(\hat{p}_S^{E_K})$. Then χ^2_{real} approximately distributes around $(2^\ell - 1) + N \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathbb{E}_K \left[\mathrm{Cor}(E_K; \alpha, \beta)^2\right]$. If the plaintext-ciphertext pairs are generated from a truly random permutation, then $\mathsf{num}(z)$ approximately follows the uniform distribution. Thus, the similarly computed statistic χ^2_{ideal} approximately follows the χ^2 distribution with $(2^\ell - 1)$ degrees of freedom (denoted by $\chi^2_{2^\ell - 1}$), of which standard deviation is $\sqrt{2(2^\ell - 1)}$. Hence E_K can be distinguished from a random permutation with a constant advantage when $N \gg \sqrt{2^\ell} / \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathbb{E}_K \left[\mathrm{Cor}(E_K; \alpha, \beta)^2\right] = \sqrt{2^\ell}\mathrm{Cap}(p_S^{E_K})$.

Remark 2. The arguments in the above paragraph are mainly based on [6, Section 4.3]. Strictly speaking, the statistic in the ideal world χ^2_{ideal} does not follow $\chi^2_{2^\ell - 1}$ actually because the squared correlation $\mathrm{Cor}(P; \alpha, \beta)^2$ is not zero on average even for a random permutation P for $\alpha, \beta \neq 0^n$. Still, the difference of χ^2_{ideal} and $\chi^2_{2^\ell - 1}$ is very small compared to the difference of χ^2_{real} and $\chi^2_{2^\ell - 1}$, and it is usually (and implicitly) assumed that the above arguments heuristically work in practice. Meanwhile, zero-correlation linear cryptanalysis *does* exploit difference between χ^2_{ideal} and $\chi^2_{2^\ell - 1}$, which we will explain later.

Remark 3. Some early works showed that distinguishers based on the Log Likelihood Ratio (LLR) test [3, 30, 31] requires only $O(1/\mathrm{Cap}(p_S^{E_K}))$ data instead of $O(\sqrt{2^\ell}/\mathrm{Cap}(p_S^{E_K}))$ of the χ^2-test-based distinguishers, and the LLR-test-based distinguishers perform better. However, the LLR test requires accurate knowledge on key-dependent distributions of multidimensional linear approximations, which is not often the case as pointed out by Cho [20].

Quantum Multidimensional Linear Distinguisher Based on $\mathcal{L}^f$. Next, we show how multidimensional linear distinguishers can be extended into the quantum setting by using the modified Simon's subroutine $\mathcal{L}^f$. We show three distinguishers $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{A}_3$. $\mathcal{A}_1$ is a general distinguisher applicable to arbitrary multidimensional linear approximations. $\mathcal{A}_2$ (resp., $\mathcal{A}_3$) is applicable only when the input and output masks are linearly independent (resp., completely dependent). The quantum speed-up by the distinguishers are at-most-quadratic compared to classical ones.

Before describing the distinguishers, we show an important claim on the distribution of multidimensional linear approximations of an ideally random permutation. The claim is used to analyze the behavior of the distinguishers when running relative to a random permutation.

Claim 1 (Distribution of capacity on a random permutation.) *Let $V \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$ be a vector space and S be an arbitrary basis of V. Then, for a randomly chosen permutation P, the value $2^n \cdot \mathrm{Cap}(p_S^P) = 2^n \cdot \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(P; \alpha, \beta)^2$*

approximately follows the χ^2 *distribution with* $2^v - 2^u - 2^w + 1$ *degrees of freedom. Here,* $v := \dim(V)$*,* $u := \dim(V \cap \mathbb{F}_2^n \times \{0^n\})$ *and* $w := \dim(V \cap \{0^n\} \times \mathbb{F}_2^n)$*.*

We argue that this claim is plausible due to the following four facts. (1) A previous work [2, Theorem 4] proves a weaker statement where "$2^v - 2^u - 2^w + 1$ degrees of freedom" in the above claim is weaken to "*at most* $2^v - 2^u - 2^w + 1$ degrees of freedom". (2) The same work conjectures that the above claim holds [2, Conjecture 1], showing some experimental results supporting the conjecture. (3) If a random variable X follows the χ^2 distribution with $2^v - 2^u - 2^w + 1$ degrees of freedom, then $\mathbb{E}[X] = 2^v - 2^u - 2^w + 1$. (4) We can formally prove that $\mathbb{E}_P[2^n \cdot \mathrm{Cap}(p_S^P)]$ is equal to $\frac{2^n}{2^n-1}(2^v - 2^u - 2^w + 1)$, which is quite close to $2^v - 2^u - 2^w + 1$ (see Proposition 12 in the appendix).

Distinguisher for General Case ($\mathcal{A}_1$). Here we show a quantum multidimensional linear distinguisher based on the modified Simon's subroutine $\mathcal{L}^f$ that is applicable without any assumptions on dependence between input and output masks. We denote the distinguisher by $\mathcal{A}_1$ and assume that we know $\sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(E_K;\alpha,\beta)^2 \geq c$ holds with a high probability for a certain value c.

The distinguisher $\mathcal{A}_1$ is obtained just by applying the algorithm $\mathcal{A}_0$ of Proposition 2 on $\mathcal{L}^{E_K}$ (or $\mathcal{L}^P$) and the Boolean function $F : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ such that $F(\alpha,\beta) = 1$ iff $(\alpha,\beta) \in V - \{\mathbf{0}\}$. Here, choosing a unitary U according to D_1 (resp., D_2) of Proposition 2 corresponds to choosing a random key K (resp., random permutation P) and define $U := \mathcal{L}^{E_K}$ (resp., $U := \mathcal{L}^P$). We set the parameters s and t of $\mathcal{A}_0$ as $s := 3$ and $t := c/2^n$.

$\mathcal{A}_1$ distinguishes E_K and P in time $O(\sqrt{2^n/c})$. The reason is roughly as follows. If the oracle given to $\mathcal{A}_1$ is E_K, the probability that we observe $(\alpha,\beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{E_K}|0^n\rangle|0^n\rangle$ is approximately lower bounded by $c/2^n$. Hence, QAA on $\mathcal{L}^{E_K}$ and F with $O(\sqrt{2^n/c})$ iterations returns $(\alpha,\beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_1$ returns 1) with high probability. On the other hand, if the oracle given to $\mathcal{A}_1$ is a random permutation P, from Claim 1 it follows that the probability that we observe $(\alpha,\beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^P|0^n\rangle|0^n\rangle$ is approximately upper bounded by $2^{\dim(V)}/2^{2n}$. Especially, the probability that QAA on $\mathcal{L}^P$ and F with $O(\sqrt{2^n/c})$ iterations returns $(\alpha,\beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_1$ returns 1) is negligibly small. Hence $\mathcal{A}_1$ distinguishes E_K and P. More precisely, the following proposition holds.

Proposition 4. *Suppose* $1/4 > \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(E_K;\alpha,\beta)^2 \geq c$ *holds with a constant probability* p *when* K *is randomly chosen, and assume* $c \gg 2^{-n}$*. If* $\mathcal{A}_1$ *runs relative to the real cipher* E_K*, then the probability that* $\mathcal{A}_1$ *outputs* 1 *is at least* $p/2$*. If* $\mathcal{A}_1$ *runs relative to a random permutation* P*, then the probability that* $\mathcal{A}_1$ *outputs* 1 *is approximately upper bounded by* $\frac{2^{\dim(V)+7}(n+1)}{2^{2n}\cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}$*. In addition,* $\mathcal{A}_1$ *makes at most* $6\sqrt{2^n/c}$ *queries to* E_K *or* P*. (The probabilities are taken not only over the randomness of* $\mathcal{A}_1$ *but also over the randomness of choices of* K *or* P*.)*

This proposition can be proven by applying Proposition 2 and Claim 1 in a straightforward manner, though, we provide a proof in Section B in the appendix for completeness.

Distinguisher for Independent Input-Output Masks ($\mathcal{A}_2$). Next, we show a distinguisher applicable if the input and output masks are linearly independent. That is, $V = V_1 \times V_2$ for some subspaces $V_1, V_2 \subset \mathbb{F}_2^n$. We denote the distinguisher by $\mathcal{A}_2$ and assume that we know $\sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(E_K;\alpha,\beta)^2 \geq c$ holds with a high probability for a certain value c.

We denote $\dim(V_1)$ and $\dim(V_2)$ by u and w, respectively. Let $S_1 := \{\alpha_1,\dots,\alpha_u\}$ and $S_2 := \{\beta_1,\dots,\beta_w\}$ be basis of V_1 and V_2. Without loss of generality we can assume $V_2 = \{\beta||0^{n-w}|\beta \in \mathbb{F}_2^w\}$ and $\beta_i = \mathbf{e}_i$[7].

$\mathcal{A}_2$ is defined by using $\mathcal{A}_0$ in almost the same way as $\mathcal{A}_1$, but here the unitary operator U is set as $U = \mathcal{L}^{\mathsf{msb}_w[E_K]}$ for the real cipher E_K (resp., $U = \mathcal{L}^{\mathsf{msb}_w[P]}$ for an ideally random permutation P) and the Boolean function $F : \mathbb{F}_2^n \times \mathbb{F}_2^w \rightarrow \mathbb{F}_2$ is such that $F(\alpha,\beta) = 1$ iff $(\alpha,\beta) \in V - \{\mathbf{0}\}$. The parameters s and t in $\mathcal{A}_0$ are set as $s := 3$ and $t := c/2^w$.

$\mathcal{A}_2$ distinguishes E_K and P in time $O(\sqrt{2^w/c})$ roughly due to the following reasoning. If the oracle given to $\mathcal{A}_2$ is E_K, the probability that we observe $(\alpha,\beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{\mathsf{msb}_w[E_K]}|0^n\rangle|0^w\rangle$ is approximately lower bounded by $c/2^w$. Hence, QAA on $\mathcal{L}^{\mathsf{msb}_w[E_K]}$ and F with $O(\sqrt{2^w/c})$ iterations returns $(\alpha,\beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_2$ returns 1) with high probability. On the other hand, if the oracle given to $\mathcal{A}_2$ is a random permutation P, from Claim 1 it follows that the probability that we observe $(\alpha,\beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{\mathsf{msb}_w[P]}|0^n\rangle|0^w\rangle$ is approximately upper bounded by $2^{\dim(V)}/2^{n+w}$. Especially, the probability that QAA on $\mathcal{L}^{\mathsf{msb}_w[P]}$ and F with $O(\sqrt{2^w/c})$ iterations returns $(\alpha,\beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_2$ returns 1) is negligibly small. More precisely, the following proposition holds.

Proposition 5. *Suppose* $1/4 > \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(E_K;\alpha,\beta)^2 \geq c$ *holds with a constant probability* p *when* K *is randomly chosen, and assume* $c \gg 2^{-n-w+\dim(V)}$. *If* $\mathcal{A}_2$ *runs relative to the real cipher* E_K*, then the probability that* $\mathcal{A}_2$ *outputs* 1 *is at least* $p/2$. *If* $\mathcal{A}_2$ *runs relative to a random permutation* P*, then the probability that* $\mathcal{A}_2$ *outputs* 1 *is approximately upper bounded by* $\frac{2^{\dim(V)+7}(n+1)}{2^{n+w}\cdot c} + 2^{-\dim(V)+1}\cdot n^{-2}$. *In addition, the running time of* $\mathcal{A}_2$ *is at most* $6\sqrt{2^w/c}$. *(The probabilities are taken not only over the randomness of* $\mathcal{A}_2$ *but also over the randomness of choices of* K *or* P*.)*

A proof of the proposition is given in Section C in the appendix.

[7] Let M be an arbitrary full-rank $n \times n$ matrix over $\mathbb{F}_2$ satisfying $M^T\mathbf{e}_i = \beta_i$. Then we have $\beta_i \cdot E_K(x) = (M^T\mathbf{e}_i)\cdot E_K(x) = \mathbf{e}_i \cdot M(E_K(x))$, and thus distinguishing E_K by using output mask β_i is equivalent to distinguishing $M \circ E_K$ by using output mask $\mathbf{e}_i$. Since E_K can be distinguished from a random permutation P iff $M \circ E_K$ can be distinguished from P, we can assume $V_2 = \{\beta||0^{n-w}|\beta \in \mathbb{F}_2^w\}$ and $\beta_i = \mathbf{e}_i$ without loss of generality.

Distinguisher for Completely Dependent Input-Output Masks ($\mathcal{A}_3$). Next, we show a distinguisher applicable if the input and output masks are linearly completely dependent. That is, there exists a basis $S := \{(\alpha_i, \beta_i)\}_{1 \le i \le \dim(V)}$ of V such that both of $\{\alpha_i\}_{1 \le i \le \dim(V)}$ and $\{\beta_i\}_{1 \le i \le \dim(V)}$ are linearly independent in $\mathbb{F}_2^n$. Without loss of generality we can assume $\beta_i = \mathbf{e}_i$[8]. We denote the distinguisher by $\mathcal{A}_3$.

$\mathcal{A}_3$ is defined by using $\mathcal{A}_0$ in almost the same way as $\mathcal{A}_1$, but here the unitary operator U is set as $U = \mathcal{L}^{\mathsf{msb}_{\dim(V)}[E_K]}$ for the real cipher E_K (resp., $U = \mathcal{L}^{\mathsf{msb}_{\dim(V)}[P]}$ for an ideally random permutation P) and the Boolean function $F : \mathbb{F}_2^n \times \mathbb{F}_2^{\dim(V)} \to \mathbb{F}_2$ is such that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$. The parameters s and t in $\mathcal{A}_0$ are set as $s := 3$ and $t := c/2^{\dim(V)}$.

$\mathcal{A}_3$ distinguishes E_K and P in time $O(\sqrt{2^{\dim(V)}/c})$ roughly due to the following reasoning. If the oracle given to $\mathcal{A}_3$ is E_K, the probability that we observe $(\alpha, \beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{\mathsf{msb}_{\dim(V)}[E_K]} |0^n\rangle |0^{\dim(V)}\rangle$ is approximately lower bounded by $c/2^{\dim(V)}$. Hence, QAA on $\mathcal{L}^{\mathsf{msb}_{\dim(V)}[E_K]}$ and F with $O(\sqrt{2^{\dim(V)}/c})$ iterations returns $(\alpha, \beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_3$ returns 1) with high probability. On the other hand, if the oracle given to $\mathcal{A}_3$ is a random permutation P, from Claim 1 it follows that the probability that we observe $(\alpha, \beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{\mathsf{msb}_{\dim(V)}[P]} |0^n\rangle |0^{\dim(V)}\rangle$ is approximately upper bounded by $1/2^n$. Especially, the probability that QAA on $\mathcal{L}^{\mathsf{msb}_{\dim(V)}[P]}$ and F with $O(\sqrt{2^{\dim(V)}/c})$ iterations returns $(\alpha, \beta) \in F^{-1}(1)$ (i.e., $\mathcal{A}_3$ returns 1) is negligibly small. More precisely, the following proposition holds.

Proposition 6. *Suppose* $1/4 > \sum_{(\alpha,\beta) \in V - \{\mathbf{0}\}} \mathrm{Cor}(E_K; \alpha, \beta)^2 \ge c$ *holds with a constant probability* p *when* K *is randomly chosen, and assume* $c \gg 2^{-n}$*. If* $\mathcal{A}_3$ *runs relative to the real cipher* E_K*, then the probability that* $\mathcal{A}_3$ *outputs* 1 *is at least* $p/2$*. If* $\mathcal{A}_3$ *runs relative to a random permutation* P*, then the probability that* $\mathcal{A}_3$ *outputs* 1 *is approximately upper bounded by* $\frac{2^7(n+1)}{2^n \cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}$*. In addition, the running time of* $\mathcal{A}_3$ *is at most* $6\sqrt{2^{\dim(V)}/c}$*. (The probabilities are taken not only over the randomness of* $\mathcal{A}_3$ *but also over the randomness of choices of* K *or* P*.)*

A proof of the proposition is given in Section D in the appendix.

Remark 4. So far we have discussed how to distinguish block ciphers from random permutations, but we expect the above distinguishers are also applicable to distinguish keyed functions from random functions of n-bit inputs, without changing the asymptotic complexity (in the same way as classical linear distinguishers work not only for permutations but also for functions). In the next section we give some application examples, but they are essentially distinguishers on keyed functions from random functions (rather than block ciphers from random permutations).

[8] As before, let M be an arbitrary full-rank $n \times n$ matrix over $\mathbb{F}_2$ satisfying $M^T \mathbf{e}_i = \beta_i$. Then we have $\beta_i \cdot E_K(x) = \mathbf{e}_i \cdot M(E_K(x))$, and thus distinguishing E_K by using output mask β_i is equivalent to distinguishing $M \circ E_K$ by using output mask $\mathbf{e}_i$. Hence we can assume S is of the form $\{(\alpha_i, \mathbf{e}_i)\}_{1 \le i \le \dim(V)}$ without loss of generality.

4.4 Appliction Example: FEA-1 and FEA-2 Structures

FEA is a Korean standard (TTAK.KO-12.0275) for format preserving encryption [46], which has two variants named FEA-1 and FEA-2. Both variants adopts *tweakable* Feistel structures. Here we study linear distinguishers on these structures when round functions are ideally random.

The FEA-1 and FEA-2 structures look like Fig. 1. As in usual Feistel structures, plaintexts are divided into two parts. We focus on the case when the widths of the two branches are equal. A tweak T is also divided into two parts, denoted by T_L and T_R, and processed in an alternate manner. In FEA-1, the i-th round function takes T_L (resp., T_R) when $i \equiv 1$ (resp., $i \equiv 0$) mod 2. In FEA-2, the i-th round function takes T_L (resp., T_R) when $i \equiv 2$ (resp., $i \equiv 0$) mod 3. The $(3j+1)$-th round function of FEA-1 does not take any tweak (or equivalently, take a constant value instead). For simplicity, we assume the tweak length is sufficiently large.

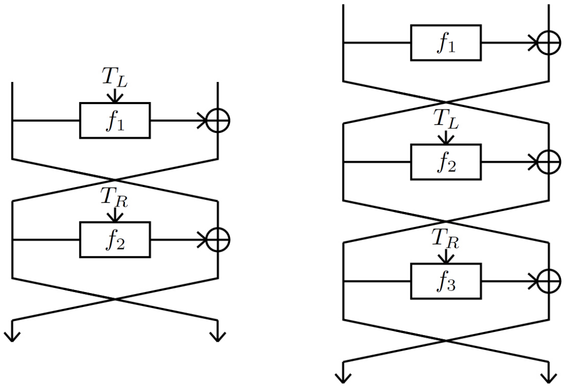


Fig. 1: The FEA-1 structure (left) and FEA-2 structure (right).

At CRYPTO 2021, Beyne showed multidimensional linear distiguishers on these structures [6]. The multidimensional linear approximation[9] for FEA-1 is a vector space V of completely linearly dependent input-output masks with $\dim(V) = 2^{n/2}$ (when n is the block size of Feistel), and the sum of the squared correlations $\sum_{(\alpha,\beta)\in V} \mathrm{Cor}(\alpha,\beta)^2$ is equal to $2^{-n(1-r/4)}$. Meanwhile, the approximation for FEA-2 is a vector space V' of linearly independent input-output masks with $\dim(V) = \dim(V_2') = 2^{n/2}$ (here, we assume V' is decomposed as $V' = V_1' \times V_2'$), and the sum of the squared correlation is equal to $2^{-n(1-r/6)}$.

[9] See the original paper [6] on details of linear approximations. Here what is significant for us is only the dimensions of the approximations and the sum of the squared correlations.

The classical distinguishing complexity is $O(2^{(r/4-3/4)n})$ for FEA-1 and $O(2^{(r/6-3/4)n})$ for FEA-2, respectively. By applying our quantum distinguishers above, the complexity is reduced to $O(2^{(r/8-1/4)n})$ and $O(2^{(r/12-1/4)n})$. In both structures, we obtain an at-most-quadratic speed-up.

In [6], Beyne also showed a linear distinguisher on FF3-1 (a NIST standard for format preserving encryption [50]). For the distinguisher on FF3-1 we can also achieve the same speed-up as that for FEA-1 by using the relationships between generalized linear correlations and quantum Fourier transform on general finite abelian groups. See Section H in the appendix for details.

5 Speed-Up for Zero Correlation Linear Distinguishers

This section shows how the modified Simon's subroutine $\mathcal{L}^f$ can be used to speed-up (multidimensional) zero correlation linear distinguishers [9]. We first recall the basic ideas of attacks in the classical setting.

5.1 Classical Zero Correlation Linear Distinguishers

Unlike linear cryptanalysis that exploits linear approximations with high correlation, zero correlation linear cryptanalysis exploits linear approximations of which correlation is exactly zero.

For instance, let E_K be an n-bit block cipher and suppose $\mathrm{Cor}(E_K;\alpha,\beta)=0$ holds for some input and output masks $\alpha,\beta\neq 0^n$. Then, for a random permutation P, $\mathrm{Cor}(P;\alpha,\beta)$ distributes around 2^{-n} and it is not equal to zero with high probability. Hence we can distinguish E_K from P if we have sufficiently many ($\approx 2^n$) plaintext-ciphertext pairs by checking whether the estimated empirical correlation is zero or not.

This idea naturally extends to attacks exploiting multidimensional linear approximations of correlation zero (below we follow the notations of Section 4.3). Again, let E_K be an n-bit block cipher and $V\subset\mathbb{F}^n\times\mathbb{F}_2^n$ be a vector space such that $\mathrm{Cor}(E_K;\alpha,\beta)=0$ for all $(\alpha,\beta)\in V$. Moreover, let S be an arbitrary basis of V. Then the distribution $p_S^{E_K}$ over $\mathbb{F}_2^{\dim(V)}$ defined by $p_S^{E_K}(z):=\Pr_z\left[\mathsf{Lin}_S^f(x)=z\right]$ exactly matches the uniform distribution. On the other hand, the distribution p_S^P similarly defined for a random permutation P is slightly different from the uniform distribution. Hence E_K and P can be distinguished by using suitable statistical tests. Indeed, Bogdanov et al. [8] showed that E_K can be distinguished in time $O(2^n/\sqrt{2^{\dim(V)}})$ in such a setting[10].

Remark 5. In the special case where the input-output masks are independent and $V=V_1\times V_2$ holds, we can achieve the time complexity $O(2^n/2^{\dim(V_1)})$ instead of $O(2^n/\sqrt{2^{\dim(V)}})$. This case is related to integral cryptanalysis, which we will elaborate in Section 6.

[10] Bogdanov and Wang showed a similar result assuming the existence of many linear approximations that are statistically independent to each other [10], but the assumption often does not hold.

5.2 Quantum Speed-Up by the Modified Simon's Subroutine

Next, we show how to mount (multidimensional) zero correlation distinsuishers by using the modified Simon's subroutine. As well as linear distinguishers in Section 4.3, we introduce three distinguishers which we name $\mathcal{B}_1$, $\mathcal{B}_2$, and $\mathcal{B}_3$. $\mathcal{B}_1$ is a general distinguisher applicable to arbitrary multidimensional linear approximations. $\mathcal{B}_2$ (resp., $\mathcal{B}_3$) is applicable only when the input and output masks are linearly independent (resp., completely dependent). In what follows, by O we denote the quantum encryption oracle given to a distinguisher, which is either of E_K or a random permutation P.

Distinguisher for General Case ($\mathcal{B}_1$). Here we show a quantum multidimensional zero correlation linear distinguisher based on the modified Simon's subroutine $\mathcal{L}^f$ that is applicable without any assumptions on dependence between input and output masks, assuming that we know $\mathrm{Cor}(E_K; \alpha, \beta) = 0$ for any $(\alpha, \beta) \in V - \{\mathbf{0}\}$. The distinguisher, denoted by $\mathcal{B}_1$, runs as follows.

1. Let $F : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ be the Boolean function such that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$.
2. Apply QAA on $\mathcal{L}^O$ and F with the number of iterations being $\lfloor \frac{\pi}{4}\sqrt{2^{2n-\dim(V)}} \rfloor$. Namely, let the unitary operator $Q(\mathcal{L}^O, F)^i \mathcal{L}^O$ act on $|0^n\rangle|0^n\rangle$ with $i = \lfloor \frac{\pi}{4}\sqrt{2^{2n-\dim(V)}} \rfloor$. Then, measure the resulting state by the computational basis and let (α, β) be the observed bit string.
3. If $F(\alpha, \beta) = 0$, return 1. Otherwise, return 0.

This $\mathcal{B}_1$ distinguishes E_K and P with high probability. The reason is roughly as follows. If the oracle given to $\mathcal{B}_1$ is E_K, the probability that we observe $(\alpha, \beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{E_K}|0^n\rangle|0^n\rangle$ is exactly zero by Proposition 3. Thus we always observe (α, β) such that $F(\alpha, \beta) = 0$ at Step 2, and $\mathcal{B}_1$ always returns 1. On the other hand, if the oracle given to $\mathcal{B}_1$ is a random permutation P, from Claim 1 it follows that the probability that we observe $(\alpha, \beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^P|0^n\rangle|0^n\rangle$ is approximately equal to $2^{\dim(V)}/2^{2n}$. Hence the QAA with $O(\sqrt{2^{2n}/\dim(V)})$ iterations in Step 2 of $\mathcal{B}_1$ returns $(\alpha, \beta) \in F^{-1}(1)$ with high probability, and $\mathcal{B}_1$ returns 0. Thus $\mathcal{B}_1$ distinguishes E_K and P.

However, the running time of $\mathcal{B}_1$ is $O(\sqrt{2^{2n-\dim(V)}}) = O(2^n/\sqrt{2^{\dim(V)}})$, which is the same as the complexity of the classical distinguisher. Namely, $\mathcal{B}_1$ does not obtain any speed-up from classical attacks. On the other hand, we can obtain at-most-quadratic quantum speed-up when input-output masks are linearly independent or linearly completely dependent, which we explain below.

Distinguisher for Independent Input-Output Masks ($\mathcal{B}_2$). Assume again that $\mathrm{Cor}(E_K; \alpha, \beta) = 0$ holds for any $(\alpha, \beta) \in V - \{\mathbf{0}\}$. Here we show a distinguisher, denoted by $\mathcal{B}_2$, applicable if the input and output masks are linearly independent. That is, $V = V_1 \times V_2$ for some subspaces $V_1, V_2 \subset \mathbb{F}_2^n$. Let $u := \dim(V_1)$ and $w := \dim(V_2)$. As well as the discussions on $\mathcal{A}_2$, without loss of generality we assume $V_2 = \{\beta || 0^{n-w} | \beta \in \mathbb{F}_2^w\}$ holds.

$\mathcal{B}_2$ is obtained by modifying the unitary operations and the number of iterations for QAA in $\mathcal{B}_1$. Specifically, we change

1. the unitary operator for QAA of $\mathcal{B}_1$ from $\mathcal{L}^{O}$ to $\mathcal{L}^{\mathsf{msb}_w[O]}$, and
2. the number of iterations from $\lfloor\frac{\pi}{4}\sqrt{2^{2n-\dim(V)}}\rfloor$ to $\lfloor\frac{\pi}{4}\sqrt{2^{n+w-\dim(V)}}\rfloor = \lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor$.

This $\mathcal{B}_2$ distinguishes E_K and P with high probability. The reason is roughly as follows. If the oracle given to $\mathcal{B}_2$ is E_K, $\mathcal{B}_2$ always returns 1 as well as $\mathcal{B}_1$. If the oracle given to $\mathcal{B}_2$ is a random permutation P, from Claim 1 it follows that the probability that we observe $(\alpha,\beta) \in F^{-1}(1)$ when measuring $\mathcal{L}^{\mathsf{msb}_w[P]}\left|0^n\right\rangle\left|0^w\right\rangle$ is approximately equal to $2^{\dim(V)}/2^{n+w} = 2^{u-n}$. Hence the QAA with $O(\sqrt{2^{n-u}})$ iterations in Step 2 of $\mathcal{B}_2$ returns $(\alpha,\beta) \in F^{-1}(1)$ with high probability, and $\mathcal{B}_2$ returns 0. Thus $\mathcal{B}_2$ distinguishes E_K and P. Especially, $\mathcal{B}_2$ achieves a quadratic speed-up in the special case where $w = 1$ (see Remark 5). More precisely, the following proposition holds.

Proposition 7. *If $\mathcal{B}_2$ runs relative to the real cipher E_K, then the probability that $\mathcal{B}_2$ always outputs* 1*. If $\mathcal{B}_2$ runs relative to a random permutation P, then the probability that $\mathcal{B}_2$ outputs* 1 *is approximately upper bounded by $\frac{1}{2}\cdot\left(1-2^{-\dim(V)+1}\right)$. In addition, the running time of $\mathcal{B}_2$ is at most $2\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor+1$ encryptions by E_K. (The probabilities are taken not only over the randomness of $\mathcal{B}_2$ but also over the randomness of choices of K or P.)*

A proof of the propositoin is given in Section F in the appendix.

Distinguisher for Independent Input-Output Masks ($\mathcal{B}_3$). Assume again that $\mathrm{Cor}(E_K;\alpha,\beta) = 0$ holds for any $(\alpha,\beta) \in V - \{\mathbf{0}\}$. Here we show a distinguisher $\mathcal{B}_3$ that is applicable if the input and output masks are linearly completely dependent. That is, V has a basis $\{(\alpha_i,\beta_i)\}_{1\le i\le\dim(V)}$ such that both of $\{\alpha_i\}_{1\le i\le\dim(V)}$ and $\{\alpha_i\}_{1\le i\le\dim(V)}$ are independent in $\mathbb{F}_2^n$. As well as the discussions on $\mathcal{A}_3$, without loss of generality we assume $\beta_i = \mathbf{e}_i$ holds for each i.

$\mathcal{B}_3$ is obtained just by changing the parameter w appeared in $\mathcal{B}_2$ to $\dim(V)$. Specifically,

1. the unitary operator for QAA is $\mathcal{L}^{\mathsf{msb}_{\dim(V)}[O]}$, and
2. the number of iterations for QAA is $\lfloor\frac{\pi}{4}\sqrt{2^{n+\dim(V)-\dim(V)}}\rfloor = \lfloor\frac{\pi}{4}\sqrt{2^{n}}\rfloor$.

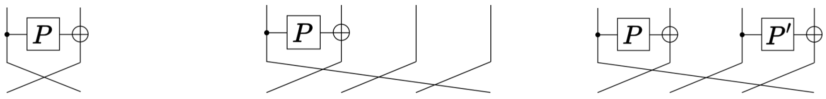
This $\mathcal{B}_3$ distinguishes E_K and P with high probability. The reason is the same as that for $\mathcal{B}_2$. Especially, $\mathcal{B}_3$ achieves an at-most-quadratic speed-up compared to the classical distinguisher, of which time complexity is $O(2^n/\sqrt{2^{\dim(V)}})$. More precisely, the following proposition holds.

Proposition 8. *If $\mathcal{B}_3$ runs relative to the real cipher E_K, then the probability that $\mathcal{B}_3$ always outputs* 1*. If $\mathcal{B}_3$ runs relative to a random permutation P, then the probability that $\mathcal{B}_3$ outputs* 1 *is approximately upper bounded by $\frac{1}{2}\cdot\left(1-2^{-\dim(V)+1}\right)$. In addition, the running time of $\mathcal{B}_3$ is at most $2\lfloor\frac{\pi}{4}\sqrt{2^{n}}\rfloor+1$ encryptions by E_K. (The probabilities are taken not only over the randomness of $\mathcal{B}_3$ but also over the randomness of choices of K or P.)*

A proof of the proposition can be obtained just by replacing the value w appearing in the proof of Propositoin 8 (Section F) with $\dim(V)$.

5.3 Applications

Both of $\mathcal{B}_2$ and $\mathcal{B}_3$ have various immediate applications. For instance, Bogdanov and Rijmen showed multidimensional zero correlation linear approximations on the 5-round balanced Feistel structure, 18-round 4-branch Type-I generalized Feistel structure, and 9-round 4-branch Type-II generalized Feistel structure (see Fig. 2 and Table 1 in the appendix) when round functions are bijections. The input-output masks of the linear approximations are linearly completely dependent. Thus $\mathcal{B}_3$ distinguish these constructions in time $O(2^{n/2})$ (when inputs and outpus are n bits). In fact the linear approximations on the 4-branch Type-I/II generalized Feistel structures can be extended to k-branch structures for general[11] k in a straightforward manner, and $\mathcal{B}_3$ distinguishes (k^2+k-2)-round (resp., $(2k+1)$-round) k-branch Type-I (resp., Type-II) generalized Feistel structure in time $O(2^{n/2})$. (See Section G for details on the extension for Type-I. The generalization for Type-II can be obtained similarly.)



(a) Balanced Feistel (b) Type-I Generalized Feistel (c) Type-II Generalized Feistel

Fig. 2: One round of Balanced and (4-branch) genelalized Feistel structures. What we assume is only that P and P' are bijections. Our attacks work regardless of whether P (and P') for different rounds are independent or not.

Balanced	k-branch Type-I	k-branch Type-II
$(\alpha \| 0^{n/2}, 0^{n/2} \| \alpha)$	$(\beta \| 0 \| \cdots \| 0, 0 \| \beta \| 0 \| \cdots \| 0$	$(\alpha \| 0 \| \cdots \| 0, 0 \| \cdots \| 0 \| \beta)$

Table 1: Input-output mask patterns for balanced and generalized Feistel structures. $\alpha \in \mathbb{F}_2^{n/2}$ and $\beta \in \mathbb{F}^{n-\frac{n}{k}}$ are non-zero values. "0" for generalized Feistel structures denotes $0^{n/k} \in \mathbb{F}_2^{n/k}$.

Remark 6. In fact, the complexity of these distinguishers may also be achieved just by speeding-up a one-dimensional zero-correlation linear distinguisher with simpler techniques. Still, to the authors' best knowledge, we are the first to point out the existence of attacks with such complexity.

There also exist lots of other previous works showing zero correlation approximations [10, 9, 8, 57, 1] and our $\mathcal{B}_2$ or $\mathcal{B}_3$ can be applied to all of them in principle.

[11] k must be even for Type-II structures.

The amount of quantum speed-up compared to classical distinguishers depends on linear approximations, and we can achieve at-most-quadratic speed-up.

6 Speed-Up for Integral Distinguishers

This section shows applications of the modified Simon's subroutine to integral cryptanalysis. Integral cryptanalysis [42], which was initially proposed as a dedicated attack on the block cipher SQUARE [22], exploits a *zero-sum property* of (a part of) ciphers. Here, we say a function $H : \mathbb{F}_2^m \to \mathbb{F}_2^n$ has a zero-sum property if $\sum_x H(x) = 0$. An important special version of zero-sum properties is a balanced function, of which definition is as follows.

Definition 1. *We say that a function H is balanced if $|H^{-1}(y)| = |H^{-1}(y')|$ holds for any y, y' in the codomain of H.*

A balanced function has a zero-sum property but the converse doe not necessarily hold.

Remark 7. Sometimes the zero-sum property is referred to as "balanced" in previous works, but this paper uses the words "balance" or "balanced" only when a balanced function (in the sense of Definition 1) appears.

For instance, let E_K be an n-bit block cipher and suppose that an output bit (say, the most significant bit) is balanced over a vector space $V \subset \mathbb{F}_2^n$. Then we can distinguish E_K from a random permutation by checking whether $\sum_{x \in V} \mathsf{msb}[E_K] = 0$ holds or not. The time and query complexity of this distinguisher is $2^{\dim(V)}$.

As shown by Bogdanov et al. [8] and later revisited by Sun et al. [57], balanced property of a cipher is equivalent to multidimensional zero correlation linear properties of which input-output masks are linearly independent. Specifially, the following proposition holds[12].

Proposition 9 ([8, 57]). *Let $F : \mathbb{F}_2^m \to \mathbb{F}_2^n$ be a function. Let $V_1 \subset \mathbb{F}_2^m, V_2 \subset \mathbb{F}_2^n$ be sub-vector spaces, and $V := V_1 \times V_2$. Then the following conditions are equivalent.*

1. *V is the set of input-output masks of a multidimensional zero correlation linear approximation of F, i.e.,* $\mathrm{Cor}(F; \alpha, \beta) = 0$ *for all* $(\alpha, \beta) \in V - \{\mathbf{0}\}$.
2. *The function $G : x \mapsto \beta \cdot F(x \oplus \lambda)$ is balanced over $V_1^{\perp}$ for all $\lambda \in \mathbb{F}_2^m$ and $\beta \in V_2 - \{\mathbf{0}\}$.*

Remark 8. Note that this equivalence holds only for balanced property but not for zero-sum property. Our quantum attacks below also rely on the above equivalence. Especially, our attacks are applicable if a cipher has a balanced property, but not necessarily applicable if the cipher has only a zero-sum property.

[12] This equivalence was first shown by [8] and later refined by [57]. [57] proves the equivalence only in the special case $\dim(V_2) = 1$ but it immediately implies the equivalence for $\dim(V_2) > 1$.

Recall that the distinguisher $\mathcal{B}_2$ (Proposition 7) is applicable when a multidimensional zero correlation linear approximation exists and the input-output masks are linearly independent. Together with Proposition 9, this implies the following proposition.

Proposition 10. *Let E_K be an n-bit block cipher. Suppose some output bits of E_K are balanced over a vector space $V \subset \mathbb{F}_2^n$. (W.l.o.g., we assume the most significant w bits are balanced, and let $V' := \{x \| 0^{n-w} | x \in \mathbb{F}_2^w\}$.) Then, by applying $\mathcal{B}_2$ on the zero correlation multidimensional linear approximations of $V^{\perp} \times V'$, we can distinguish E_K from P with time and query complexity at most $2\lfloor \frac{\pi}{4}\sqrt{2^{\dim(V)}} \rfloor + 1$. $\mathcal{B}_2$ always outputs 1 if the given encryption oracle is the real cipher E_K. If the oracle is a random permutation P, the probability that $\mathcal{B}_2$ outputs 1 is at most $\frac{1}{2}\left(1 - 2^{-\dim(V)+1}\right)$.*

This proposition shows that we can obtain (almost) quadratic speed-up for integral distinguisher because the complexity of $\mathcal{B}_2$ is $\approx 1.6\sqrt{2^{\dim(V)}}$ while the complexity of the classical integral distinguisher is $2^{\dim(V)}$.

Still, this is at-most-quadratic speed-up. At first glance, achieving a more-than-quadratic speed-up may seem hard also for integral distinguishers. However, we can actually achieve a more-than-quadratic speed-up in some situations.

Roughly speaking, if a part of outputs of a cipher (e.g., a specific byte of ciphertexts) is balanced on multiple mutually orthogonal vector spaces included in the input space, we obtain *multiple* multidimensional zero correlation linear approximations of the cipher such that the vector spaces of the input-output spaces are mutually orthogonal. This situation often occurs if an AES-like cipher has an integral property. Such a structure can be exploited to achieve more-than-quadratic quantum speed-up by using the modified Simon's subroutine. To see this, we first revisit the 2.5-round integral distinguisher for AES.

6.1 Case Study on 2.5-round AES

Recall that each cell after 2.5 rounds of AES takes all values when the first cell of inputs take all values while others are fixed to some constants [23] (see Fig. 3). This in turn means that the 2.5-round AES has multidimensional zero-correlation linear approximations by Proposition 9. Specifically, $V_1 := \{\alpha \in \mathbb{F}_2^n | \text{the 1st byte of } \alpha \text{ is } 0\} \times$

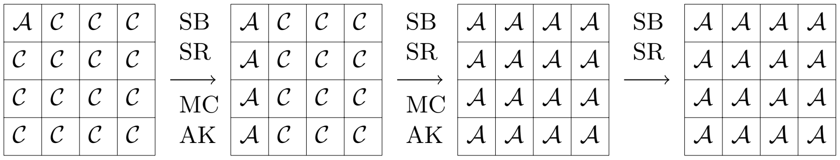


Fig. 3: The integral property of the 2.5-round AES.

$\{\beta \in \mathbb{F}_2^n | \text{all bytes except for the 1st byte of } \beta \text{ are } 0\}$ yields a zero correlation multidimensional linear approximation. By applying $\mathcal{B}_2$ to V_1 we can obtain an (almost) quadratic quantum speed-up.

Now, the important point is that integral properties of 2.5-round AES are not limited to the one shown in Fig. 3. We can obtain 15 similar integral properties due to the degrees of freedom on which input cell to activate. This means that there exist other multidimensional zero correlation linear approximations different from V_1. The choice of which input cell to activate in 2.5-round integral properties corresponds to the choice of which byte of input mask α to be zero. For instance, the vector space $V_2 := \{\alpha \in \mathbb{F}_2^n | \text{the second byte of } \alpha \text{ is zero}\} \times \{\beta \in \mathbb{F}_2^n | \text{all bytes except for the 1st byte of } \beta \text{ are } 0\}$ corresponds to the integral property activating the second input cell instead of the first.

Let us similarly define V_i for $i = 3, \ldots, 16$ by setting the i-th byte of α to be zero. Then we have that $\mathrm{Cor}(E_K; \alpha, \beta) = 0$ holds for all $(\alpha, \beta) \in V_0 \cup \cdots \cup V_{15} - \{\mathbf{0}\}$. This means that, if we run the modified Simon's subroutine on the 2.5-round AES with outputs being truncated to the first byte (denoted by $\mathsf{msb}_8[\mathrm{AES}_{2.5}]$), we never observe (α, β) such that a byte of α is zero (among 16 bytes) and $(\alpha, \beta) \neq (0^{128}, 0^8)$. Namely,

$$\Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{\mathsf{msb}_8[\mathrm{AES}_{2.5}]} : \text{a byte of } \alpha \text{ is zero and } (\alpha, \beta) \neq (0^{128}, 0^8)\right] = 0 \quad (8)$$

holds. Meanwhile, if we run the modified Simon's subroutine on a random permutation P similarly (with outputs being truncated to the first byte, denoted by $\mathsf{msb}_8[P]$) and measure the final state, the observed bit string (α, β) is just random[13]. In particular, roughly we have

$$\begin{aligned}
&\Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{\mathsf{msb}_8[P]} : \text{a byte of } \alpha \text{ is zero and } (\alpha, \beta) \neq (0^{128}, 0^8)\right] \\
&= \sum_{1 \leq i \leq 16} \Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{\mathsf{msb}_8[P]} : \text{the } i\text{-th byte of } \alpha \text{ is zero and } (\alpha, \beta) \neq (0^{128}, 0^8)\right] \\
&\approx \sum_{1 \leq i \leq 16} \frac{1}{2^8} = \frac{1}{2^4}. \quad (9)
\end{aligned}$$

Now, define a Boolean function $F : \mathbb{F}_2^{128} \times \mathbb{F}_2^8 \to \mathbb{F}_2$ by $F(\alpha, \beta) = 1$ iff a byte of α (among 16 bytes) is zero and $(\alpha, \beta) \neq (0^{128}, 0^8)$. Then, by applying QAA on the Boolean function F and $\mathcal{L}^{\mathsf{msb}_8[\mathrm{AES}_{2.5}]}$ or $\mathcal{L}^{\mathsf{msb}_8[P]}$ with the number of iterations about $\frac{\pi}{4}\sqrt{2^4} \approx 3$, we can distinguish the 2.5-round AES from a random permutation with a high probability. The total time and query of the distinguisher is $2\lfloor \frac{\pi}{4}\sqrt{2^4} \rfloor + 1 \lessapprox 2^3$, which is less than the square root of the complexity of the classical integral distinguisher ($\sqrt{2^8} = 2^4$). Hence we obtain a super-quadratic speed-up for the integral distinguisher on the 2.5-round AES.

The margin (compared to the square-root of the classical complexity) is not large, but still we obtain the first example of a non-trivial super-quadratic speed-

[13] Strictly speaking, $\alpha \neq 0^{128}$ and $\beta \neq 0^8$ always hold while the probability of observing $(0^{128}, 0^8)$ is $\frac{1}{2^8}$. Still, such details do not significantly matter the arguments here.

up for classical cryptanalytic technique on symmetric cryptosystems without relying on algebraic structures such as hidden periods or shifts.

A Single-Query Distinguisher on 4-bit Cell Toy Ciphers. The quantum speed-up for the integral distinguisher by our technique becomes more evident if the size of each cell is smaller. For instance, suppose there is a 4-bit cell SPN cipher E_K that has the same integral property as the 2.5-round AES (the block size is 64-bit). To distinguish E_K from a random permutation P, we apply the modified Simon's subroutine on $\mathsf{msb}_4[E_K]$ or $\mathsf{msb}_4[P]$. Then, since the size of cells is now 4-bit, the probability corresponding to Eq. (9) becomes

$$\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^{\mathsf{msb}_4[P]} : \text{a cell of } \alpha \text{ is zero and } (\alpha,\beta) \neq (0^{64},0^4)\right] \approx \sum_{1\leq i\leq 16} \frac{1}{2^4} = 1.$$

Meanwhile, the corresponding probability for E_K is again zero. That is,

$$\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^{\mathsf{msb}_4[E_K]} : \text{a cell of } \alpha \text{ is zero and } (\alpha,\beta) \neq (0^{64},0^4)\right] = 0.$$

These equations show that E_K can be distinguished from P by a *single-query* quantum attack applying the modified Simon's subroutine and checking whether one of the 16 cells of α is zero and $(\alpha,\beta) \neq (0^{64},0^4)$.

It seems impossible to obtain such a single-query distinguisher in the classical setting, and our attack exhibits a new type of quantum algorithm exploiting linear correlations of target functions.

6.2 Generalization

From the above arguments, we observe that we can obtain a more-than-quadratic speed-up on integral distinguishers based on balanced functions if there are multiple choices on which part of inputs (e.g., cells) to take all values.

Specifically, let $E_K : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a block cipher, and suppose there exist sub-vector spaces $V_1,\dots,V_s \subset \mathbb{F}_2^m$ satisfying the following conditions.

1. $V_1,\dots,V_s$ are mutually orthogonal, i.e., $V_i \perp V_j$ for $i \neq j$.
2. There exists some $d \leq n/2$ and $\dim(V_i) = d$ holds for all i.
3. A part of outputs of E_K is balanced on $V_i \oplus \lambda$ for all $1 \leq i \leq s$ and arbitrary $\lambda \in \mathbb{F}_2^m$. (For ease of explanation, below we assume the most significant w bits of outputs of E_K are balanced.)

Then, by Proposition 9 we have $\mathrm{Cor}(\mathsf{msb}_w[E_K];\alpha,\beta) = 0$ if $\alpha \in (V_1)^\perp \cup \cdots \cup (V_s)^\perp - \{\mathbf{0}\}$ and $(\alpha,\beta) \neq (0,0)$. This means

$$\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^{\mathsf{msb}_w[E_K]} : \alpha \perp V_i \text{ for some } i \text{ and } \alpha \neq 0 \text{ and } \beta \neq 0\right] = 0.$$

Meanwhile, for a random permutation P we have

$$\begin{aligned}
&\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^{\mathsf{msb}_w[P]} : \alpha \perp V_i \text{ for some } i \text{ and } \alpha \neq 0 \text{ and } \beta \neq 0\right] \\
&\overset{\text{Prop. 3}}{=} \sum_{\substack{\alpha\neq 0, \beta\neq 0 \\ \alpha\perp V_i \text{ for some } i}} \frac{\mathrm{Cor}(\mathsf{msb}_w[P];\alpha,\beta)^2}{2^w} = \sum_{\substack{\alpha\neq 0, \beta\neq 0 \\ \alpha\perp V_i \text{ for some } i \\ \mathsf{lsb}_{n-w}[\beta]=0}} \frac{\mathrm{Cor}(P;\alpha,\beta)^2}{2^w} \\
&\overset{\text{Prop.13}}{=} \#\left\{\alpha \in \mathbb{F}_2^n - \{\mathbf{0}\} | \alpha \perp V_i \text{ for some } i\right\} \cdot \frac{2^w - 1}{2^w(2^n-1)} \\
&\geq \left(\sum_{1\leq i\leq s} |V_i^{\perp}| - \sum_{1\leq i<j\leq s} |V_i^{\perp} \cap V_j^{\perp}| - 1\right) \cdot \frac{2^w-1}{2^w(2^n-1)} \\
&\overset{V_i\perp V_j \text{ for } i\neq j}{\geq} \left(s2^{n-d} - s^2 2^{n-2d} - 1\right) \cdot \frac{2^w-1}{2^w(2^n-1)} \quad \approx \quad \frac{s}{2^d}.
\end{aligned}$$

Therefore, E_K can be distinguished from P in time about $\frac{\pi}{2}\sqrt{2^d/s}$ by applying QAA on $\mathcal{L}^{\mathsf{msb}_w[E_K]}$ (or $\mathcal{L}^{\mathsf{msb}_w[P]}$) and the Boolean function $F : \mathbb{F}_2^n \times \mathbb{F}_2^w \to \mathbb{F}_2$ such that $F(\alpha,\beta) = 1$ iff $\alpha \perp V_i$ for some $i = 1, \ldots, s$ and $\alpha \neq 0$ and $\beta \neq 0$.

This is a more-than-quadratic speed-up compared to the corresponding classical integral distinguisher (when $s \geq 4$) because the classical complexity is 2^{n-d}.

The attacks on the 2.5-round AES and 4-bit cell toy cipher mentioned in the previous section are special cases of the above technique, and the technique can also be applied to, e.g., the 3.5-round AES integral distinguisher [23] (for the 3.5 round distinguisher, there are 4 choices on which 32-bit set to choose).

7 Discussions

On Extension to Key-Recovery. All the distinguishers in previous sections can be extended into key-recovery attacks just by guessing sub-keys of additional rounds using Grover's algorithm. Suppose we would like to recover the key of an $(r + r')$-round cipher and there is a (quantum) r-round distinguisher on a cipher running in time T. In addition, assume that we can apply the distinguisher on an intermediate r rounds if we know a k-bit subkey K' in the remaining r'-rounds. Then, roughly speaking, by just guessing the subkey K' with the Grover search while checking if a key-guess is correct with the distinguisher, we achieve an $(r + r')$-round quantum key-recovery attack of time complexity $O(T \cdot 2^{k/2})$[14].

However, this attack idea is a very naive one. In the classical setting, key-recovery attacks are efficiently performed by using sophisticated techniques such as FFT [21, 58, 27]. Since FFT clearly relates to Fourier transform, it is natural to expect a non-trivial speed-up for such sophisticated key-recovery techniques

[14] This is a very rough analysis. In fact the complexity may be much higher depending on the power of the distinguisher to filter wrong sub-keys, but we omit details here.

by using quantum algorithms. Still, currently we think how to speed-up such techniques in the quantum setting is highly non-trivial and leave this as an open question.

On Limitations of Our Quantum Integral Distinguishers. As mentioned before, our quantum integral distinguishers are applicable only if the distinguishers are based on a balanced functions and not a zero-sum property. However, zero-sum properties are often more useful than balanced functions when extending distinguihsers to key-recoveries. Especially, we are currently not aware of any example such that our distinguisher leads to a more-than-quadratic speed-up for existing classical key-recovery attacks based on integral properties. So far we do not have any idea on how to achieve quantum speed-up for integral distinguishers based on zero-sum properties.

Acknowledgements We thank María Naya-Plasencia for reminding us of some previous works on quantum speed-up for classical cryptanalytic techniques.

References

1. Ankele, R., Dobraunig, C., Guo, J., Lambooij, E., Leander, G., Todo, Y.: Zero-correlation attacks on tweakable block ciphers with linear tweakey expansion. IACR Trans. Symmetric Cryptol. **2019**(1), 192–235 (2019)
2. Ashur, T., Khan, M., Nyberg, K.: Structural and statistical analysis of multidimensional linear approximations of random functions and permutations. IEEE Trans. Inf. Theory **68**(2), 1296–1315 (2022)
3. Baignères, T., Junod, P., Vaudenay, S.: How far can we go beyond linear cryptanalysis? In: Lee, P.J. (ed.) ASIACRYPT 2004, Proceedings. LNCS, vol. 3329, pp. 432–450. Springer (2004)
4. Baignères, T., Stern, J., Vaudenay, S.: Linear cryptanalysis of non binary ciphers. In: Adams, C.M., Miri, A., Wiener, M.J. (eds.) SAC 2007, Revised Selected Papers. LNCS, vol. 4876, pp. 184–211. Springer (2007)
5. Bernstein, E., Vazirani, U.V.: Quantum complexity theory. SIAM J. Comput. **26**(5), 1411–1473 (1997)
6. Beyne, T.: Linear cryptanalysis of FF3-1 and FEA. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021, Proceedings, Part I. LNCS, vol. 12825, pp. 41–69. Springer (2021)
7. Biryukov, A., Cannière, C.D., Quisquater, M.: On multiple linear approximations. In: Franklin, M.K. (ed.) CRYPTO 2004, Proceedings. LNCS, vol. 3152, pp. 1–22. Springer (2004)
8. Bogdanov, A., Leander, G., Nyberg, K., Wang, M.: Integral and multidimensional linear distinguishers with correlation zero. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012, Proceedings. LNCS, vol. 7658, pp. 244–261. Springer (2012)
9. Bogdanov, A., Rijmen, V.: Linear hulls with correlation zero and linear cryptanalysis of block ciphers. Des. Codes Cryptogr. **70**(3), 369–383 (2014)
10. Bogdanov, A., Wang, M.: Zero correlation linear cryptanalysis with reduced data complexity. In: Canteaut, A. (ed.) FSE 2012, Revised Selected Papers. LNCS, vol. 7549, pp. 29–48. Springer (2012)

11. Bonnetain, X., Hosoyamada, A., Naya-Plasencia, M., Sasaki, Y., Schrottenloher, A.: Quantum attacks without superposition queries: The offline simon's algorithm. In: Galbraith, S.D., Moriai, S. (eds.) ASIACRYPT 2019, Proceedings, Part I. LNCS, vol. 11921, pp. 552–583. Springer (2019)
12. Bonnetain, X., Leurent, G., Naya-Plasencia, M., Schrottenloher, A.: Quantum linearization attacks. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021, Proceedings, Part I. LNCS, vol. 13090, pp. 422–452. Springer (2021)
13. Bonnetain, X., Naya-Plasencia, M., Schrottenloher, A.: On quantum slide attacks. In: Paterson, K.G., Stebila, D. (eds.) SAC 2019, Revised Selected Papers. LNCS, vol. 11959, pp. 492–519. Springer (2019)
14. Bonnetain, X., Naya-Plasencia, M., Schrottenloher, A.: Quantum security analysis of AES. IACR Trans. Symmetric Cryptol. **2019**(2), 55–93 (2019)
15. Boyer, M., Brassard, G., Høyer, P., Tapp, A.: Tight bounds on quantum searching. Fortschritte der Physik: Progress of Physics **46**(4-5), 493–505 (1998)
16. Brassard, G., Hoyer, P., Mosca, M., Tapp, A.: Quantum amplitude amplification and estimation. Contemporary Mathematics **305**, 53–74 (2002)
17. Brassard, G., Høyer, P., Tapp, A.: Quantum cryptanalysis of hash and claw-free functions. In: LATIN 1998. LNCS, vol. 1380, pp. 163–169. Springer (1998)
18. Canale, F., Leander, G., Stennes, L.: Simon's algorithm and symmetric crypto: Generalizations and automatized applications (2022)
19. Chartouny, M., Patarin, J., Toulemonde, A.: Quantum cryptanalysis of rounds feistel schemes and benes schemes. IACR Cryptology ePrint Archive 2022/1015 (2022)
20. Cho, J.Y.: Linear cryptanalysis of reduced-round PRESENT. In: Pieprzyk, J. (ed.) CT-RSA 2010, Proceedings. LNCS, vol. 5985, pp. 302–317. Springer (2010)
21. Collard, B., Standaert, F., Quisquater, J.: Improving the time complexity of matsui's linear cryptanalysis. In: Nam, K., Rhee, G. (eds.) ICISC 2007, Proceedings. LNCS, vol. 4817, pp. 77–88. Springer (2007)
22. Daemen, J., Knudsen, L.R., Rijmen, V.: The block cipher square. In: Biham, E. (ed.) FSE 1997, Proceedings. LNCS, vol. 1267, pp. 149–165. Springer (1997)
23. Daemen, J., Rijmen, V.: Aes proposal: Rijndael (1999)
24. Daemen, J., Rijmen, V.: Probability distributions of correlation and differentials in block ciphers. J. Mathematical Cryptology **1**(3), 221–242 (2007)
25. Dong, X., Li, Z., Wang, X.: Quantum cryptanalysis on some generalized feistel schemes. Sci. China Inf. Sci. **62**(2), 22501:1–22501:12 (2019)
26. Dong, X., Sun, S., Shi, D., Gao, F., Wang, X., Hu, L.: Quantum collision attacks on aes-like hashing with low quantum random access memories. In: Moriai, S., Wang, H. (eds.) ASIACRYPT 2020, Part II. LNCS, vol. 12492, pp. 727–757. Springer (2020)
27. Flórez-Gutiérrez, A., Naya-Plasencia, M.: Improving key-recovery in linear attacks: Application to 28-round PRESENT. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020, Proceedings, Part I. LNCS, vol. 12105, pp. 221–249. Springer (2020)
28. Grover, L.K.: A Fast Quantum Mechanical Algorithm for Database Search. In: ACM STOC 1996. pp. 212–219. ACM (1996)
29. Guo, J., Liu, G., Song, L., Tu, Y.: Exploring sat for cryptanalysis: (quantum) collision attacks against 6-round sha-3. To appear at ASIACRYPT 2022
30. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional linear cryptanalysis of reduced round serpent. In: Mu, Y., Susilo, W., Seberry, J. (eds.) ACISP 2008, Proceedings. LNCS, vol. 5107, pp. 203–215. Springer (2008)

31. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional extension of matsui's algorithm 2. In: Dunkelman, O. (ed.) FSE 2009, Revised Selected Papers. LNCS, vol. 5665, pp. 209–227. Springer (2009)
32. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional linear cryptanalysis. J. Cryptol. **32**(1), 1–34 (2019)
33. Hermelin, M., Nyberg, K.: Multidimensional linear distinguishing attacks and boolean functions. In: Fourth International Workshop on Boolean Functions: Cryptography and Applications (2008)
34. Hosoyamada, A., Sasaki, Y.: Cryptanalysis against symmetric-key schemes with online classical queries and offline quantum computations. In: CT-RSA. LNCS, vol. 10808, pp. 198–218. Springer (2018)
35. Hosoyamada, A., Sasaki, Y.: Finding hash collisions with quantum computers by using differential trails with smaller probability than birthday bound. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020, Part II. LNCS, vol. 12106, pp. 249–279. Springer (2020)
36. Hosoyamada, A., Sasaki, Y.: Quantum collision attacks on reduced SHA-256 and SHA-512. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021, Proceedings, Part I. LNCS, vol. 12825, pp. 616–646. Springer (2021)
37. Ito, G., Hosoyamada, A., Matsumoto, R., Sasaki, Y., Iwata, T.: Quantum chosen-ciphertext attacks against feistel ciphers. In: Matsui, M. (ed.) CT-RSA 2019, Proceedings. LNCS, vol. 11405, pp. 391–411. Springer (2019)
38. Jr., B.S.K., Robshaw, M.J.B.: Linear cryptanalysis using multiple approximations. In: Desmedt, Y. (ed.) CRYPTO 1994, Proceedings. LNCS, vol. 839, pp. 26–39. Springer (1994)
39. Kaplan, M., Leurent, G., Leverrier, A., Naya-Plasencia, M.: Breaking symmetric cryptosystems using quantum period finding. In: CRYPTO 2016, Part II. LNCS, vol. 11693, pp. 207–237. Springer (2016)
40. Kaplan, M., Leurent, G., Leverrier, A., Naya-Plasencia, M.: Quantum differential and linear cryptanalysis. IACR Trans. Symmetric Cryptol. **2016**(1), 71–94 (2016)
41. Knudsen, L.R.: The security of feistel ciphers with six rounds or less. J. Cryptol. **15**(3), 207–222 (2002)
42. Knudsen, L.R., Wagner, D.A.: Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002, Revised Papers. LNCS, vol. 2365, pp. 112–127. Springer (2002)
43. Kuwakado, H., Morii, M.: Quantum distinguisher between the 3-round Feistel cipher and the random permutation. In: ISIT 2010. pp. 2682–2685. IEEE (2010)
44. Kuwakado, H., Morii, M.: Security on the quantum-type Even-Mansour cipher. In: ISITA 2012. pp. 312–316. IEEE (2012)
45. Leander, G., May, A.: Grover Meets Simon - Quantumly Attacking the FX-construction. In: ASIACRYPT 2017. LNCS, vol. 10625, pp. 161–178. Springer (2017)
46. Lee, J., Koo, B., Roh, D., Kim, W., Kwon, D.: Format-preserving encryption algorithms using families of tweakable blockciphers. In: Lee, J., Kim, J. (eds.) ICISC 2014, Revised Selected Papers. LNCS, vol. 8949, pp. 132–159. Springer (2014)
47. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993, Proceedings. LNCS, vol. 765, pp. 386–397. Springer (1993)
48. Murphy, S.: The independence of linear approximations in symmetric cryptanalysis. IEEE Trans. Inf. Theory **52**(12), 5510–5518 (2006)
49. Nachef, V., Patarin, J., Volte, E.: Generic Attacks on Generalized Feistel Ciphers, pp. 139–153. Springer International Publishing, Cham (2017)

50. National Institute of Standards and Technology: Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption. SP 800-38G Rev. 1, U.S. Department of Commerce (Feb 2019)
51. Ni, B., Ito, G., Dong, X., Iwata, T.: Quantum attacks against type-1 generalized feistel ciphers and applications to CAST-256. In: Hao, F., Ruj, S., Gupta, S.S. (eds.) INDOCRYPT 2019, Proceedings. LNCS, vol. 11898, pp. 433–455. Springer (2019)
52. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press (2010)
53. Patarin, J.: New results on pseudorandom permutation generators based on the DES scheme. In: Feigenbaum, J. (ed.) CRYPTO 1991, Proceedings. LNCS, vol. 576, pp. 301–312. Springer (1991)
54. Patarin, J.: Security of random feistel schemes with 5 or more rounds. In: Franklin, M.K. (ed.) CRYPTO 2004, Proceedings. LNCS, vol. 3152, pp. 106–122. Springer (2004)
55. Shi, R., Xie, H., Feng, H., Yuan, F., Liu, B.: Quantum zero correlation linear cryptanalysis. Quantum Inf. Process. **21**(8), 293 (2022)
56. Simon, D.R.: On the power of quantum computation. SIAM J. Comput. **26**(5), 1474–1483 (1997)
57. Sun, B., Liu, Z., Rijmen, V., Li, R., Cheng, L., Wang, Q., AlKhzaimi, H., Li, C.: Links among impossible differential, integral and zero correlation linear cryptanalysis. In: Gennaro, R., Robshaw, M. (eds.) CRYPTO 2015, Proceedings, Part I. LNCS, vol. 9215, pp. 95–115. Springer (2015)
58. Todo, Y., Aoki, K.: Fast fourier transform key recovery for integral attacks. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **98-A**(9), 1944–1952 (2015)
59. Treger, J., Patarin, J.: Generic attacks on feistel networks with internal permutations. In: Preneel, B. (ed.) AFRICACRYPT 2009, Proceedings. LNCS, vol. 5580, pp. 41–59. Springer (2009)
60. Xie, H., Yang, L.: Using bernstein-vazirani algorithm to attack block ciphers. Des. Codes Cryptogr. **87**(5), 1161–1182 (2019)

A Proof of Proposition 2

First, we restate the algorithm and the statement of the proposition. Recall that s is an arbitrary positive integer constant and $p_U := \Pr\left[x \xleftarrow{\text{measure}} U\left|0^n\right\rangle : F(x) = 1\right]$.

QAA for Distinguisher (Algorithm $\mathcal{A}_0$).

1. For $i = 1, \ldots, s$, do:
   (a) Choose i uniformly at random from the set of integers from 0 to $\left\lfloor \frac{1}{\sin\left(2\cdot\arcsin\left(\sqrt{t}\right)\right)} \right\rfloor$.
   (b) Apply $Q(U, F)^i U$ to $\left|0^n\right\rangle$ and measure the entire state by the computational basis, and let x be the outcome.
   (c) Compute $F(x)$. If $F(x) = 1$, return 1 and abort.
2. Return 0.

Structure	Rounds	Round Functions	Attack Type	Complexity	Reference
Balanced Feistel	4		CPA	$O(2^{n/2})$	[53]
Balanced Feistel	4		QCCA	$O(\mathrm{poly}(n))$	[37]
Balanced Feistel	5		CPA	$O(2^n)$	[54]
Balanced Feistel	5	bij.	KPA	$O(2^{2n/3})$	[41, 59]
Balanced Feistel	5		QCPA	$O(2^{2n/3})$	[19]
Balanced Feistel	5	bij.	QCPA	$O(2^{n/2})$	**Ours**
k-branch Type-I Generalized Feistel	k^2-k+1	bij.	QCCA	$O(\mathrm{poly}(n))$	[51]
k-branch Type-I Generalized Feistel	k^2+k-1		CPA	$O(2^{(1-\frac{1}{k})n})$	[49]
4-branch Type-I Generalized Feistel	18	bij.	KPA	$O(2^{3n/4})$	[9] (combined with [8])
k-branch Type-I Generalized Feistel	k^2+k-2	bij.	QCPA	$O(2^{n/2})$	**Ours**
k-branch Type-II Generalized Feistel	$k+1$	bij.	QCPA	$O(\mathrm{poly}(n))$	[25]
k-branch Type-II Generalized Feistel	$2k+1$		CPA	$O(2^{(1-\frac{1}{k})n})$	[49]
4-branch Type-II Generalized Feistel	9		KPA	$O(2^{3n/4})$	[9] (combined with [8])
k-branch Type-II Generalized Feistel	$2k+1$	bij.	QCPA	$O(2^{n/2})$	**Ours**

Table 2: Comparison of classical and quantum attacks on balanced and Type-I/II generalized Feistel structures. "bij." means that the attack assumes round functions are bijective. The parameters k for Type-II generalized Feistel are even numbers. QCPA (resp., QCCA) denotes quantum superposition (Q2) chosen plaintext attack (resp., quanutm superposition chosen ciphertext attack). All of our attacks appearing in this table are multidimensional zero correlation linear distinguishers.

Proposition 11 (Restatement of Proposition 2). *Suppose $1/4 > t > 0$. Then, for any constant s, $\mathcal{A}_0$ applies U, U^*, and $\mathcal{S}_F$ at most $s(\frac{1}{\sqrt{t}}+1)$ times and (1) returns 1 with probability at least $(1-(\frac{3}{4})^s)\cdot \Pr_{U\sim D_1}[1/4 > p_U \geq t]$ if U is chosen according to D_1 and (2) returns 1 with probability at most $s\cdot(16t'/t + 20t'/\sqrt{t}) + \Pr_{U\sim D_2}[t' < p_U]$ for any $t' > 0$ satisfying $4\sqrt{t'/t}+2\sqrt{t'} < \pi/2$ if U is chosen according to D_2.*

We use the following lemma from [15].

Lemma 1 (Lemma 2 in [15]). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function, U be a unitary operator acting on an n qubit system, and let* $p_{\mathrm{init}} := \Pr\left[x \xleftarrow{measure} U\left|0^n\right\rangle : F(x) = 1\right]$.

Assume $0 < p_{\mathrm{init}} < 1/2$*, and let* N *be an integer satisfying* $N \geq \frac{1}{\sin\left(2\cdot\arcsin\left(\sqrt{p_{\mathrm{init}}}\right)\right)}$*. Then* $\Pr\left[i \xleftarrow{\$} \{0,\dots,N-1\}, x \xleftarrow{\textit{measure}} Q(U,F)^i U\left|0^n\right\rangle : F(x)=1\right] \geq 1/4$ *holds.*

In fact Lemma 2 in [15] proves the claim only when $U = H^{\otimes n}$ but it is straightforward to check that the proof is valid for arbitrary U (due to Lemma 1 in [16]). We assume $0 < p_{\mathrm{init}} < 1/4$ so that $0 < 2\cdot\arcsin\left(\sqrt{p_{\mathrm{init}}}\right) < \pi/2$ will hold.

Proof (of Proposition 2). The claim for the number of applications of U, U^*, and $\mathcal{S}_F$ immediately follows from the definition of the algorithm and $Q(U,F)$ because

$$\frac{1}{\sin\left(2\cdot\arcsin\left(\sqrt{t}\right)\right)} \leq \frac{1}{\sin(2\sqrt{t})} \leq \frac{1}{2\sqrt{t}-\frac{4}{3}t\sqrt{t}} \leq \frac{1}{\sqrt{t}}$$

holds, where we used $\arcsin(x) \leq x$, $\sin(x) \leq x - x^3/6$, and $t < 1/4$.

Next, we lower bound the success probability when U is chosen according to D_1. Recall that the algorithm iteratively picks a random i and measure the state $Q(U,F)U\left|0^n\right\rangle$. Now, assume $1/4 > p_U \geq t$. Then, by Lemma 1, the probability that the algorithm fails to find x satisfying $F(x)=1$ at the j-th iteration of the algorithm is at most $3/4$ for each $j=1,\dots,s$. Thus, assuming $1/4 > p_U \geq t$, the probability that the algorithm succeeds to find x satisfying $F(x)=1$ after s iteration is at least $(1-(3/4)^s)$.

Hence, when U is chosen according to D_1, the probability that the algorithm finds x satisfying $F(x)=1$ is lower bounded by $(1-(3/4)^s)\cdot\Pr_{U\sim D_1}[1/4 > p_U \geq t]$.

Next, we upper bound the success probability when U is chosen according to D_2. Recall that t' is a positive value satisfying $4\sqrt{t'/t}+2\sqrt{t'} \leq \pi/2$. Now, assume $p_U \leq t'$. Then, for any i between 0 and $\frac{1}{\sin(2\cdot\arcsin(\sqrt{t}))}$ we have

$$i \leq \frac{1}{\sin(2\cdot\arcsin(\sqrt{t}))} \overset{(x/2\leq\sin(x))}{\leq} \frac{1}{\arcsin(\sqrt{t})} \overset{(x\leq\arcsin(x))}{\leq} \frac{1}{\sqrt{t}}$$

and thus

$$(2i+1)\cdot\arcsin(\sqrt{p_U}) \leq (\frac{2}{\sqrt{t}}+1)\cdot\arcsin(\sqrt{t'}) \overset{(2x\geq\arcsin(x))}{\leq} 4\sqrt{\frac{t'}{t}}+2\sqrt{t'} \quad (\leq \pi/2)$$

holds. Hence, by Proposition 1, the probability that the algorithm finds x satisfying $F(x)=1$ at the j-th iteration of the algorithm is at most

$$\begin{aligned}\sin^2((2i+1)\arcsin\sqrt{p_U}) &\leq \sin^2\left(4\sqrt{t'/t}+2\sqrt{t'}\right)\\ &\overset{(\sin(x)\leq x)}{\leq} 16t'/t+16t'/\sqrt{t}+4t'\\ &\leq 16t'/t+20t'/\sqrt{t}.\end{aligned}$$

Thus, assuming $t' \geq p_U$, the probability that the algorithm succeeds to find x satisfying $F(x)=1$ after s iteration is at most $s\cdot(16t'/t+20t'/\sqrt{t})$.

Therefore, when U is chosen according to D_2, the probability that the algorithm finds x satisfying $F(x) = 1$ is upper bounded bounded by

$$\begin{aligned} s \cdot (16t'/t + 20t'/\sqrt{t}) \Pr_{U \sim D_2}[t' \geq p_U] + \Pr_{U \sim D_2}[t' < p_U] \\ \leq s \cdot (16t'/t + 20t'/\sqrt{t}) + \Pr_{U \sim D_2}[t' < p_U], \end{aligned}$$

which completes the proof. □

B Proof of Proposition 4

Proof (of Proposition 4). First, note that

$$p_\pi := \Pr\left[(\alpha, \beta) \xleftarrow{\text{measure}} \mathcal{L}^\pi \left|0^m\right\rangle \left|0^n\right\rangle : F(x) = 1\right] = \sum_{(\alpha,\beta) \in V - \{\mathbf{0}\}} \frac{\mathrm{Cor}(\pi; \alpha, \beta)^2}{2^n}$$

holds for arbitrary permutation π by Lemma 3 (in this proof, π is now a block cipher E_K or a randomly chosen permutation P). p_U in Proposition 2 corresponds to p_π here.

If the oracle given to $\mathcal{A}_1$ is the real cipher E_K, then Proposition 2 guarantees that $\mathcal{A}_1$ returns 1 (i.e., it judges the oracle is the real cipher E_K) with probability at least $(1 - (3/4)^s)p \geq p/2$.

If the oracle given to $\mathcal{A}_1$ is a random permutation P, then Claim 1 guarantees that $\mathbb{E}_P[p_P]$ and $\mathbf{Var}_P[p_P]$ are approximately upper bounded as $\mu_P := \mathbb{E}_P[p_P] \lessapprox 2^{\dim(V)-2n}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lessapprox \sqrt{2^{\dim(V)+1-4n}}$. Hence we have

$$\begin{aligned} \Pr_P\left[p_P > (n+1) \cdot \frac{2^{\dim(V)}}{2^n}\right] &\lessapprox \Pr_P\left[p_P > \mu_P + n \cdot 2^{\frac{\dim(V)-1}{2}} \sigma_P\right] \\ &\overset{\text{Chebyshev's inequality}}{\leq} 2^{-\dim(V)+1} \cdot n^{-2}. \end{aligned}$$

By the claim on D_2 in Proposition 2 with $t' = (n+1) \cdot \frac{2^{\dim(V)}}{2^{2n}}$, the probability that $\mathcal{A}_1$ returns 1 is at most[15]

$$3 \cdot (16t'/c + 20t'/\sqrt{c}) + 2^{-\dim(V)+1} \cdot n^{-2} \lessapprox \frac{2^{\dim(V)+7}(n+1)}{2^{2n} \cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}.$$

In addition, by definition of $\mathcal{A}_0$ and $\mathcal{A}_1$, $\mathcal{A}_1$ makes at most $6\sqrt{2^n/c}$ quantum queries to E_K or P and the costs for other operations are negligibly small. Hence the running time of $\mathcal{A}_1$ is at most $6\sqrt{2^n/c}$. □

[15] We need the condition $c \gg 2^{-n}$ so that $4\sqrt{t'/c} + 2\sqrt{t'} \ll \pi/2$ will hold and we can apply the claim on D_2 here.

C Proof of Proposition 5

Proof (of Proposition 5). First, note that

$$p_f := \Pr\left[(\alpha, \beta) \xleftarrow{\text{measure}} \mathcal{L}^f \left|0^m\right\rangle \left|0^n\right\rangle : F(x) = 1\right] = \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \frac{\mathrm{Cor}(f;\alpha,\beta)^2}{2^w}$$

holds for arbitrary function f by Lemma 3 (in this proof, f is now a truncated version of a block cipher $\mathsf{msb}_w[E_K]$ or a randomly chosen permutation $\mathsf{msb}_w[P]$). p_U in Proposition 2 corresponds to p_f here.

If the oracle given to $\mathcal{A}_2$ is the real cipher E_K, then Proposition 2 guarantees that $\mathcal{A}_2$ returns 1 (i.e., it judges the oracle is the real cipher E_K) with probability at least $(1-(3/4)^s)p \geq p/2$.

If the oracle given to $\mathcal{A}_2$ is a random permutation P, then Claim 1 guarantees that $\mathbb{E}_P[p_P]$ and $\mathbf{Var}_P[p_P]$ are approximately upper bounded as $\mu_P := \mathbb{E}_P[p_P] \lessapprox 2^{\dim(V)-n-w}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lessapprox \sqrt{2^{\dim(V)+1-2n-2w}}$. Hence we have

$$\begin{aligned} \Pr_P\left[p_P > (n+1)\cdot\frac{2^{\dim(V)}}{2^{n+w}}\right] &\lessapprox \Pr_P\left[p_P > \mu_P + n\cdot 2^{\frac{\dim(V)-1}{2}}\sigma_P\right] \\ &\overset{\text{Chebyshev's inequality}}{\leq} 2^{-\dim(V)+1}\cdot n^{-2}. \end{aligned}$$

By the claim on D_2 in Proposition 2 with $t' = (n+1)\cdot\frac{2^{\dim(V)}}{2^{n+w}}$, the probability that $\mathcal{A}_2$ returns 1 is at most[16]

$$3\cdot(16t'/c + 20t'/\sqrt{c}) + 2^{-\dim(V)+1}\cdot n^{-2} \lessapprox \frac{2^{\dim(V)+7}(n+1)}{2^{n+w}\cdot c} + 2^{-\dim(V)+1}\cdot n^{-2}.$$

In addition, by definition of $\mathcal{A}_0$ and $\mathcal{A}_2$, $\mathcal{A}_2$ makes at most $6\sqrt{2^w/c}$ quantum queries to E_K or P and the costs for other operations are negligibly small. Hence the running time of $\mathcal{A}_2$ is at most $6\sqrt{2^w/c}$. □

D Proof of Proposition 6

Proof (of Proposition 6). First, note that

$$p_f := \Pr\left[(\alpha, \beta) \xleftarrow{\text{measure}} \mathcal{L}^f \left|0^m\right\rangle \left|0^n\right\rangle : F(x) = 1\right] = \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \frac{\mathrm{Cor}(f;\alpha,\beta)^2}{2^{\dim(V)}}$$

holds for arbitrary function f by Lemma 3 (in this proof, f is now a truncated version of a block cipher $\mathsf{msb}_{\dim(V)}[E_K]$ or a randomly chosen permutation $\mathsf{msb}_{\dim(V)}[P]$). p_U in Proposition 2 corresponds to p_f here.

[16] We need the condition $c \gg 2^{-n-w+\dim(V)}$ so that $4\sqrt{t'/c} + 2\sqrt{t'} \ll \pi/2$ will hold and we can apply the claim on D_2 here.

If the oracle given to $\mathcal{A}_3$ is the real cipher E_K, then Proposition 2 guarantees that $\mathcal{A}_3$ returns 1 (i.e., it judges the oracle is the real cipher E_K) with probability at least $(1-(3/4)^s)p \geq p/2$.

If the oracle given to $\mathcal{A}_3$ is a random permutation P, then Claim 1 guarantees that $\mathbb{E}_P[p_P]$ and $\mathbf{Var}_P[p_P]$ are approximately upper bounded as $\mu_P := \mathbb{E}_P[p_P] \lessapprox 2^{-n}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lessapprox \sqrt{2^{-\dim(V)+1-2n}}$. Hence we have

$$\begin{aligned}\Pr_P\left[p_P > (n+1)\cdot\frac{1}{2^n}\right] &\lessapprox \Pr_P\left[p_P > \mu_P + n\cdot 2^{\frac{\dim(V)-1}{2}}\sigma_P\right]\\ &\overset{\text{Chebyshev's inequality}}{\leq} 2^{-\dim(V)+1}\cdot n^{-2}.\end{aligned}$$

By the claim on D_2 in Proposition 2 with $t' = (n+1)\cdot\frac{1}{2^n}$, the probability that $\mathcal{A}_1$ returns 1 is at most[17]

$$3\cdot(16t'/c + 20t'/\sqrt{c}) + 2^{-\dim(V)+1}\cdot n^{-2} \lessapprox \frac{2^7(n+1)}{2^n\cdot c} + 2^{-\dim(V)+1}\cdot n^{-2}.$$

In addition, by definition of $\mathcal{A}_0$ and $\mathcal{A}_3$, $\mathcal{A}_3$ makes at most $6\sqrt{2^{\dim(V)}/c}$ quantum queries to E_K or P and the costs for other operations are negligibly small. Hence the running time of $\mathcal{A}_2$ is at most $6\sqrt{2^{\dim(V)}/c}$. □

E On Expected Values of Capacity

The goal of the section is to show the following proposition. In this section, we follow the notations used in Claim 1. That is, $V \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$ is a vector space and S be an arbitrary basis of V, and paremeters v, u, and w are defined as $v := \dim(V)$, $u := \dim(V \cap \mathbb{F}_2^n \times \{0^n\})$, and $w := \dim(V \cap \{0^n\} \times \mathbb{F}_2^n)$.

Proposition 12. *For a randomly chosen permutation P, $\mathbf{E}[2^n \cdot \mathrm{Cap}(p_S^P)] = \frac{2^n}{2^n-1}(2^v - 2^u - 2^w + 1)$ holds.*

Since $2^n \cdot \mathrm{Cap}(p_S^P) = 2^n \cdot \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \mathrm{Cor}(P;\alpha,\beta)^2$ holds, it suffices to show the following proposition.

Proposition 13. *Let $f : \{0,1\}^m \to \{0,1\}^n$ be a random balanced function. (Namely, f is chosen uniformly at random from the set of all balanced functions from $\{0,1\}^m$ to $\{0,1\}^n$. If $m = n$, f is just a random permutation.) Then, for arbitrary $\alpha \in \{0,1\}^m$ and $\beta \in \{0,1\}^n$,*

$$\mathbf{E}_f\left[\mathrm{Cor}(f;\alpha,\beta)^2\right] = \begin{cases} 1 & \text{if } \alpha = 0^m \text{ and } \beta = 0^n \\ 0 & \text{if } \alpha \neq 0^m \text{ and } \beta = 0^n, \text{ or } \alpha = 0^m \text{ and } \beta \neq 0^m \\ \frac{1}{2^m-1} & \text{if } \alpha \neq 0^m \text{ and } \beta \neq 0^n \end{cases}$$

holds. Here, the expectation value is taken over the random choice of f.

[17] We need the condition $c \gg 2^{-n-w+\dim(V)}$ so that $4\sqrt{t'/c} + 2\sqrt{t'} \ll \pi/2$ will hold and we can apply the claim on D_2 here.

Proof. When $\alpha = 0^m$ and $\beta = 0^n$ (resp., $\alpha \neq 0^m$ and $\beta = 0^n$), the correlation $\mathrm{Cor}(f; \alpha, \beta)$ is zero (resp., 1) for arbitrary function f. When $\alpha = 0^m$ and $\beta \neq 0^n$,

$$\begin{aligned}
\mathrm{Cor}(f; \alpha, \beta) &= \sum_{x \in \{0,1\}^m} \frac{(-1)^{\alpha \cdot x \oplus \beta \cdot f(x)}}{2^m} = \sum_{x \in \{0,1\}^m} \frac{(-1)^{\beta \cdot f(x)}}{2^m} \\
&= \sum_{y \in \{0,1\}^n} \sum_{x \in f^{-1}(y)} \frac{(-1)^{\beta \cdot y}}{2^m} = \sum_{y \in \{0,1\}^n} 2^{m-n} \frac{(-1)^{\beta \cdot y}}{2^m} = 0
\end{aligned}$$

always holds for arbitrary balanced function f (we used the balancedness of f for the second last equality).

Next, we show the claim when $\alpha \neq 0^m$ and $\beta \neq 0^m$. Let $\mathsf{Perm}(m)$ be the set of all permutations over $\{0,1\}^m$ and $\mathsf{Reg}(m,n)$ denote the set of all balanced functions from $\{0,1\}^m$ to $\{0,1\}^n$. $\chi_{\alpha,\beta}(f,x)$ denote the indicator function such that $\chi_{\alpha,\beta}(f,x) = 1$ iff $\alpha \cdot x = \beta \cdot f(x)$. Then

$$\begin{aligned}
\mathrm{Cor}(f; \alpha, \beta)^2 &= \left(\Pr_x[\alpha \cdot x = \beta \cdot f(x)] - \Pr_x[\alpha \cdot x \neq \beta \cdot f(x)]\right)^2 \\
&= \left(2 \Pr_x[\alpha \cdot x = \beta \cdot f(x)] - 1\right)^2 \\
&= \left(\frac{2 \cdot \#\left\{x \in \{0,1\}^m | \alpha \cdot x = \beta \cdot f(x)\right\} - 2^m}{2^m}\right)^2 \\
&= \frac{1}{2^{2m}} \left(2 \sum_{x \in \{0,1\}^m} \chi_{\alpha,\beta}(f,x) - 2^m\right)^2 \\
&= \frac{1}{2^{2m}} \left(4 \sum_{x,x' \in \{0,1\}^m} \chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x') - 2^{m+2} \sum_{x \in \{0,1\}^m} \chi_{\alpha,\beta}(f,x) + 2^{2m}\right) \\
&= \frac{1}{2^{2m}} \left(4 \sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x'}} \chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x') - (2^{m+2} - 4) \sum_{x \in \{0,1\}^m} \chi_{\alpha,\beta}(f,x) + 2^{2m}\right)
\end{aligned} \tag{10}$$

holds. Now, for each $x \in \{0,1\}^m$ we have

$$\mathbf{E}_f\left[\chi_{\alpha,\beta}(f,x)\right] = \Pr_f\left[\chi_{\alpha,\beta}(f,x) = 1\right] = \frac{1}{2} \tag{11}$$

because, for each fixed tuple (α, β, x) with $\beta \neq 0^n$, the number of $f \in \mathsf{Reg}(m,n)$ satisfying $\alpha \cdot x = \beta \cdot f(x)$ is equal to the number of f satisfying $\alpha \cdot x \neq \beta \cdot f(x)$. For any permutation $P \in \mathsf{Perm}(m)$, let P_{tr} be the truncated function of P obtained by discarding the rightmost $(m-n)$ bits. Then, for arbitrary distinct

$x, x' \in \{0,1\}^m$ we have

$$\begin{aligned}
&\mathbf{E}_f\left[\chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x')\right]\\
&\quad= \Pr_{f \xleftarrow{\$} \mathsf{Reg}(m,n)}\left[\chi_{\alpha,\beta}(f,x) = 1 \wedge \chi_{\alpha,\beta}(f,x') = 1\right]\\
&\quad= \Pr_{P \leftarrow \mathsf{Perm}(m)}\left[\chi_{\alpha,\beta||0^{m-n}}(P_{tr},x) = 1 \wedge \chi_{\alpha,\beta||0^{m-n}}(P_{tr},x') = 1\right]\\
&\quad= \Pr_{P \xleftarrow{\$} \mathsf{Perm}(m)}\left[\chi_{\alpha,\beta||0^{m-n}}(P_{tr},x') = 1 | \chi_{\alpha,\beta||0^{m-n}}(P_{tr},x) = 1\right]\\
&\qquad\qquad \cdot \Pr_{P \xleftarrow{\$} \mathsf{Perm}(m)}\left[\chi_{\alpha,\beta||0^{m-n}}(P_{tr},x) = 1\right]\\
&\quad= \begin{cases} \frac{2^{m-1}-1}{2^m-1} \cdot \frac{1}{2} \text{ if } \alpha \cdot x = \alpha \cdot x' \\ \frac{2^{m-1}}{2^m-1} \cdot \frac{1}{2} \text{ if } \alpha \cdot x \neq \alpha \cdot x'. \end{cases}
\end{aligned}$$

In addition, since $\alpha \neq 0^m$ we have

$$\begin{aligned}
\{(x,x') \in \{0,1\}^m \times \{0,1\}^m | x \neq x' \wedge \alpha \cdot x = \alpha \cdot x'\} = 2^m \cdot (2^{m-1} - 1),\\
\{(x,x') \in \{0,1\}^m \times \{0,1\}^m | x \neq x' \wedge \alpha \cdot x \neq \alpha \cdot x'\} = 2^m \cdot 2^{m-1}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x'}} \mathbf{E}_f\left[\chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x')\right]\\
&= \sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x' \wedge \alpha \cdot x = \alpha \cdot x'}} \mathbf{E}_f\left[\chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x')\right] + \sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x' \wedge \alpha \cdot x \neq \alpha \cdot x'}} \mathbf{E}_f\left[\chi_{\alpha,\beta}(f,x)\chi_{\alpha,\beta}(f,x')\right]\\
&= \sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x' \wedge \alpha \cdot x = \alpha \cdot x'}} \frac{2^{m-1}-1}{2^m-1} \cdot \frac{1}{2} + \sum_{\substack{x,x' \in \{0,1\}^m \\ x \neq x' \wedge \alpha \cdot x \neq \alpha \cdot x'}} \frac{2^{m-1}}{2^m-1} \cdot \frac{1}{2}\\
&= \frac{2^m(2^{m-1}-1)^2}{2 \cdot (2^m-1)} + \frac{2^{3m-2}}{2 \cdot (2^m-1)}\\
&= \frac{2^{3m-2} - 2^{2m-1} + 2^{m-1}}{2^m-1} \qquad (12)
\end{aligned}$$

holds. From Eq. (10)-(12), it follows that

$$\begin{aligned}
\mathbf{E}_f\left[\mathrm{Cor}(f;\alpha,\beta)\right] &= \frac{1}{2^{2m}}\left(4 \cdot \frac{2^{3m-2} - 2^{2m-1} + 2^{m-1}}{2^m-1} - (2^{m+2}-4) \cdot \frac{2^m}{2} + 2^{2m}\right)\\
&= \frac{1}{2^m-1},
\end{aligned}$$

which completes the proof. □

F Proof of Proposition 7

Proof (of Proposition 7). It immediately follows from the definition of $\mathcal{B}_2$, Proposition 2, and Proposiion 3 that $\mathcal{B}_2$ always returns 1 when the given oracle is E_K and the running time of $\mathcal{B}_2$ is at most $2\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor + 1$ encryptions by E_K.

Below we prove that the claim on the probability that $\mathcal{B}_2$ returns 1 when the given oracle is a random permutation P. To show this, it suffices to prove that

$$\Pr_P\left[(\alpha,\beta) \xleftarrow{\text{measure}} Q(\mathcal{L}^{\mathsf{msb}_w[P]}, F)^{\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor}\mathcal{L}^{\mathsf{msb}_w[P]}\left|0^n\right\rangle\left|0^w\right\rangle\right] \gtrapprox \frac{1}{2}\cdot\left(1 - 2^{-\dim(V)+1}\right)$$

holds.

First, note that

$$p_f := \Pr\left[(\alpha,\beta) \xleftarrow{\text{measure}} \mathcal{L}^f\left|0^m\right\rangle\left|0^n\right\rangle : F(x) = 1\right] = \sum_{(\alpha,\beta)\in V-\{\mathbf{0}\}} \frac{\mathrm{Cor}(f;\alpha,\beta)^2}{2^w}$$

holds for arbitrary function f by Lemma 3 (in this proof, f is now a truncated version of a randomly chosen permutation $\mathsf{msb}_w[P]$).

If the oracle given to $\mathcal{B}_2$ is a random permutation P, then Claim 1 guarantees that $\mathbb{E}_P[p_P]$ and $\mathbf{Var}_P[p_P]$ are approximately upper bounded as $\mu_P := \mathbb{E}_P[p_P] \lessapprox 2^{\dim(V)-n-w} = 2^{u-n}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lessapprox \sqrt{2^{\dim(V)+1-2n-2w}} = 2^{\frac{-\dim(V)+1}{2}} \cdot \mu_P$. Hence we have

$$\begin{aligned}\Pr_P\left[|p_P - \mu_P| > \mu_P\right] &\lessapprox \Pr_P\left[|p_P - \mu_P| > 2^{\frac{\dim(V)-1}{2}}\sigma_P\right]\\ &\overset{\text{Chebyshev's inequality}}{\leq} 2^{-\dim(V)+1}.\end{aligned}$$

Expecially,

$$\Pr_P\left[\frac{1}{2}2^{u-n} \leq p_P \leq 2\cdot 2^{u-n}\right] \geq 1 - 2^{-\dim(V)+1} \tag{13}$$

holds. In addition, for each P such that $\frac{1}{2}2^{u-n} \leq p_P \leq 2\cdot 2^{u-n}$, we have

$$\begin{aligned}&\Pr\left[(\alpha,\beta) \xleftarrow{\text{measure}} Q(\mathcal{L}^{\mathsf{msb}_w[P]}, F)^{\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor}\mathcal{L}^{\mathsf{msb}_w[P]}\left|0^n\right\rangle\left|0^w\right\rangle\right]\\ &= \sin^2\left(\left(2\left\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\right\rfloor + 1\right)\arcsin(\sqrt{p_P})\right)\\ &\approx \sin^2\left(\frac{\pi}{2}\sqrt{2^{n-u}}\cdot\sqrt{p_P}\right) \gtrapprox \frac{1}{2}.\end{aligned} \tag{14}$$

Therefore

$$\begin{aligned}&\Pr\left[(\alpha,\beta) \xleftarrow{\text{measure}} Q(\mathcal{L}^{\mathsf{msb}_w[P]}, F)^{\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor}\mathcal{L}^{\mathsf{msb}_w[P]}\left|0^n\right\rangle\left|0^w\right\rangle\right]\\ &\geq \Pr\left[(\alpha,\beta) \leftarrow Q(\mathcal{L}^{\mathsf{msb}_w[P]}, F)^{\lfloor\frac{\pi}{4}\sqrt{2^{n-u}}\rfloor}\mathcal{L}^{\mathsf{msb}_w[P]}\left|0^n\right\rangle\left|0^w\right\rangle\middle|\frac{1}{2}2^{u-n} \leq p_P \leq 2\cdot 2^{u-n}\right]\\ &\quad\cdot\Pr_P\left[\frac{1}{2}2^{u-n} \leq p_P \leq 2\cdot 2^{u-n}\right]\\ &\gtrapprox \frac{1}{2}\cdot\left(1 - 2^{-\dim(V)+1}\right)\end{aligned} \tag{15}$$

holds, which completes the proof. □

G On Zero Correlation Linear Approximations for Type-I Generalized Feistel

This section shows the following proposition, which is a natural extension of the zero correlation linear approximations for 18-round 4-branch Ttype-I generalized Feistel structure by Bogdanov and Rijmen [9].

Proposition 14. *For any* $\alpha \in \mathbb{F}_2^{\frac{n}{k}}$ *and any* $k \geq 2$*, the input-output mask* $(\alpha || 0^{n-\frac{n}{k}}, 0^{\frac{n}{k}} || \alpha || 0^{n-\frac{2n}{k}})$ *yields a zero correlation linear approximation of* $(k^2 + k - 2)$*-round* k*-branch Type-I generalized Feistel structure of which round functions are bijections.*

In what follows, by Ψ_k^r we denote an r-round k-branch Type-I generalized Feistel structure of which round functions are bijections (over $\mathbb{F}_2^{n/k}$). The block length of Ψ_k^r is n. For each input or output mask $X \in \mathbb{F}_2^n$, by X_i we denote the (n/k)-bit sub-mask for the i-th leftmost branch. (The sub-mask for the leftmost (resp., rightmost) branch is X_1 (resp., X_k).)

Lemma 2. *Let* A *and* B *be an input and output mask for* Ψ_k^k*, and let* i *a positive integer such that* $3 \leq i \leq k-1$*. Assume that (1) the sub-input-mask for the* j*-th branch* A_j *is* $0^{n/k}$ *for* $j = 1, \ldots, i$*, and (2) the sub-input-mask for the rightmost branch* A_k *is equal to a non-zero value* $\alpha \in \mathbb{F}_2^{n/k}$*. Then, for the correlation* $\mathrm{Cor}(\Psi_k^k; A, B)$ *to be non-zero,* B *must satisfy the following two conditions. (1) The sub-output-mask for the* j*-th branch* B_j *is* $0^{n/k}$ *for* $j = 1, \ldots, i-1$*. (2) The sub-output-mask for the rightmost branch* B_k *is equal to* α*.*

Proof. We prove the claim when $k = 5$ and $i = 3$ (see Fig. 4).

For the correlation $\mathrm{Cor}(\Psi_k^k; A, B)$ to be non-zero, firstly the input-output masks for the first round function must be 0 since $A_2 = 0$. (Together with the condition $A_1 = 0$,) this implies that the sub-mask of the rightmost branch after the first round must be zero. Especially, the input-output masks for the last round function must be 0. This implies that $B_1 = 0$ and $B_5 = \alpha$ must hold (because $\mathcal{A}_5 = \alpha$).

Similarly, the sub-mask of the rightmost branch after the second round must be zero because $A_2 = A_3 = 0$. Hence B_2 must be zero.

The proofs for other k and i are similar. □

By applying Lemma 2 $(k-2)$ times, we obtain the following lemma.

Lemma 3. *Let* A *and* B *be an input and output mask for* $\Psi_k^{k(k-2)}$*. Assume that (1) the sub-input-mask for the* j*-th branch* A_j *is* $0^{n/k}$ *for* $j = 1, \ldots, k-1$*, and (2) the sub-input-mask for the rightmost branch* A_k *is equal to a non-zero value* $\alpha \in \mathbb{F}_2^{n/k}$*. Then, for the correlation* $\mathrm{Cor}(\Psi_k^k; A, B)$ *to be non-zero, the sub-output-mask for the rightmost branch* B_k *must be equal to* α*.*

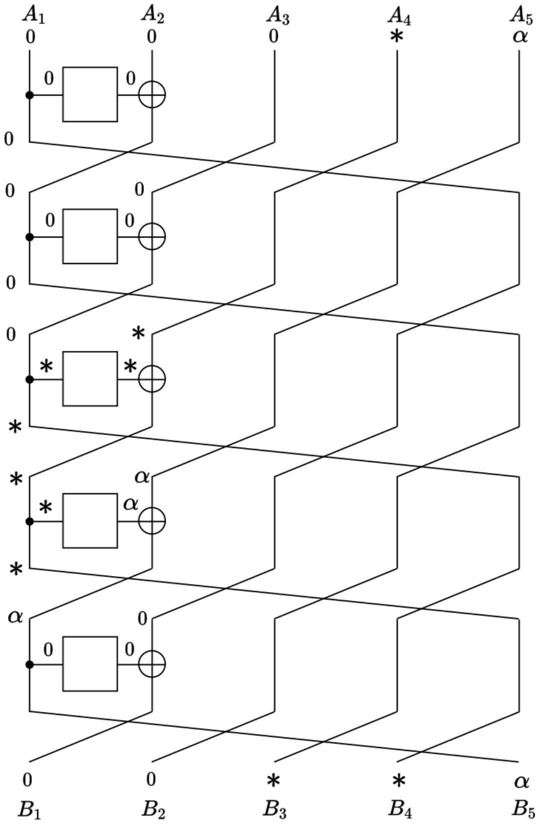


Fig. 4: Masks when $k = 5$ and $i = 3$.

In addition, since the rightmost branch will not be updated in successive $(k-2)$ rounds during encryption, we have the following lemma.

Lemma 4. *Let A and B be an input and output mask for Ψ_k^{k-1}. Assume that the sub-input-mask for the rightmost branch A_k is equal to a non-zero value $\alpha \in \mathbb{F}_2^{n/k}$. Then, for the correlation $\mathrm{Cor}(\Psi_k^k; A, B)$ to be non-zero, the sub-output-mask for the leftmost branch B_1 must be equal to α.*

Then, the lemma below follows from Lemmas 3 and 4.

Lemma 5. *Let A and B be an input and output mask for $\Psi_k^{k^2-k-1}$. Assume that (1) the sub-input-mask for the j-th branch A_j is $0^{n/k}$ for $j = 1, \ldots, k-1$, and (2) the sub-input-mask for the rightmost branch A_k is equal to a non-zero value $\alpha \in \mathbb{F}_2^{n/k}$. Then, for the correlation $\mathrm{Cor}(\Psi_k^k; A, B)$ to be non-zero, the sub-output-mask for the leftmost branch B_1 must be equal to α.*

So far we have studies some conditions that output mask must satisfy when input masks are given. In what follows, we show some lemmas on some conditions that *input* masks must satisfy when output masks are given.

Lemma 6. *Let A and B be an input and output mask for Ψ_k^k. Assume that (1) the sub-output-mask for the j-th branch B_j is $0^{n/k}$ for $j = 2, \ldots, k$, and (2) the sub-output-mask for the leftmost branch B_1 is equal to a non-zero value $\alpha \in \mathbb{F}_2^{n/k}$. Then, for the correlation $\mathrm{Cor}(\Psi_k^k; A, B)$ to be non-zero, the sub-input-mask for the leftmost branch A_1 must be different from α.*

Proof. We prove the statemnt when $k = 5$ (see Fig. 5).

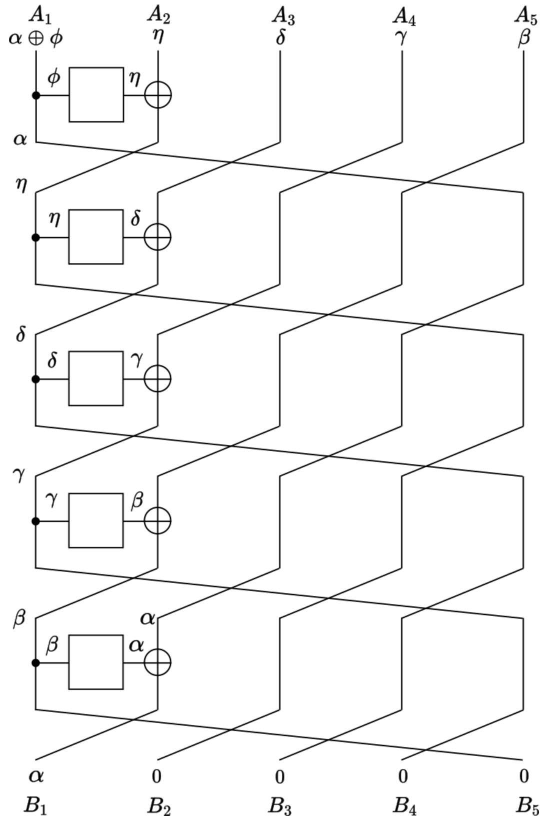


Fig. 5: Masks when $k = 5$. $\alpha, \beta, \gamma, \eta, \phi$ are non-zero values in $\mathbb{F}_2^{n/k}$.

For the correlation $\mathrm{Cor}(\Psi_k^k; A, B)$ to be non-zero, firstly the input mask for the final round function must be a non-zero value β since the sub-output-mask of the rightmost branch B_5 is 0. We can similarly deduce that the input mask for the 4th, 3rd, 2nd, and 1st round function must be a non-zero value. Since $B_1 = \alpha$ and the input mask for the first round function is non-zero, A_1 must be different from α.

The statement for other k can be proven in the same way. □

Since the second-left branch will not be updated for sucessive $(k-2)$ rounds in decryption, the following lemma holds.

Lemma 7. *Let A and B be an input and output mask for Ψ_k^{k-1}. Assume that (1) the sub-output-mask for the j-th branch B_j is $0^{n/k}$ for $j \neq 2$, and (2) the sub-output-mask for the second-left branch B_2 is equal to a non-zero value $\alpha \in \mathbb{F}_2^{n/k}$. Then, for the correlation $\mathrm{Cor}(\Psi_k^{k-1}; A, B)$ to be non-zero, A must be $\alpha || 0^{n-\frac{n}{k}}$.*

The lemma below follows from Lemmas 6 and 7.

Lemma 8. *Let A and B be an input and output mask for Ψ_k^{2k-1}. Assume that (1) the sub-output-mask for the j-th branch B_j is $0^{n/k}$ for $j \neq 2$, and (2) the sub-output-mask for the second-left branch B_2 is equal to a non-zero value $\alpha \in \mathbb{F}_2^{n/k}$. Then, for the correlation $\mathrm{Cor}(\Psi_k^{2k-}; A, B)$ to be non-zero, the sub-input mask for the leftmost branch A_1 must be different from α.*

Proof (of Proposition 14). Apply Lemma 5 and Lemma 8 for the first (k^2-k-1) rounds and last $(2k-1)$ rounds of $\Psi_k^{k^2+k-2}$, respectively. Then Proposition 14 immediately follows. □

H Extensions to Other Groups

Linear cryptanalysis is useful when group operations are done in $\mathbb{Z}_2^n$, but some ciphers use other group operations such as modular additions (i.e., additions in $\mathbb{Z}/2^n\mathbb{Z}$). In such situations, generalized linear cryptanalysis on arbitrary finite groups [4] is more useful. Generalized linear cryptanalysis uses group characters instead of bit masks, but we observe again there exists a close relationship between (generalized) linear correlations and quantum computation via Fourier transform.

In what follows, we show how the arguments on quantum speed-up for multi-dimensional (zero-correlation) linear distinguishers extend to generalized linear distinguishers where group operations may be an addition in an arbitrary finite abelian group. Note that the symbol "$\oplus$" denotes the direct sum of groups in this section.

H.1 Fourier Transform on Arbitrary Finite abelian Group

Let G be an arbitrary finite abelian group. Then, w.l.o.g. we can assume $G = \mathbb{Z}/N_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/N_m\mathbb{Z}$ for some positive integers $N_1, \ldots, N_m$. Recall that a character of a finite abelian group G is a group homomorphism $\phi : G \to \mathbb{C}^\times$. The set of characters of G is denoted by $\hat{G}$, which forms a group by point-wise multiplication. It is well-known that $\hat{G}$ is isomorphic to G as a group.

Specifically, for each $w = (w_1, \ldots, w_m) \in G$, the function

$$\mathsf{ch}_w : (x_1, \ldots, x_m) \mapsto \exp\left(2\pi i \frac{x_1 w_1}{N_1}\right) \cdots \exp\left(2\pi i \frac{x_m w_m}{N_m}\right)$$

is a character of G. In fact the map $w \mapsto \mathsf{ch}_w$ defines a group isomorphism from G to $\hat{G}$. We identify G with $\hat{G}$ by this isomorphism.

Let G be a finite abelian group and $F : G \to \mathbb{C}$ be a function. Then, the Fourier transform of F over G is a function $\mathcal{F}_G F : G \to \mathbb{C}$ defined by

$$\mathcal{F}_G F(w) := \sum_{x \in G} \frac{1}{\sqrt{|G|}} \cdot \overline{\mathsf{ch}_w(x)} \cdot F(x).$$

The inverse transform of $\mathcal{F}_G$, denoted by $\mathcal{F}_G^*$, is given by $\mathcal{F}_G^* F(x) = \sum_{x \in G} \frac{1}{\sqrt{|G|}} \cdot \mathsf{ch}_w(x) \cdot F(x)$.

In what follows, we naturally identify a function from G to $\mathbb{C}$ (resp., the set of the functions from G to $\mathbb{C}$) with a vector in the $|G|$-dimensional vector space $\mathbb{C}^{|G|}$ (resp., the vector space $\mathbb{C}^{|G|}$). We assume that $\mathbb{C}^{|G|}$ is endowed with the standard Hermitian inner product. Then $\mathcal{F}_G$ can be regarded as a unitary operator.

H.2 Linear Correlations

Let G, H be finite abelian groups and $f : G \to H$ be a function. For $\alpha \in G$ and $\beta \in H$, the (generalized) linear correlation $\mathrm{Cor}(f; \alpha, \beta)$ is defined as

$$\mathrm{Cor}(f; \alpha, \beta) := \sum_{x \in G} \frac{1}{|G|} \overline{\mathsf{ch}_\beta(f(x))} \cdot \mathsf{ch}_\alpha(x).$$

We call α (resp., β) an input mask (resp., output mask).

Let $f_{\mathrm{emb}} : G \times H \to \mathbb{C}$ be the function defined by $f_{\mathrm{emb}}(x, y) = 1$ if $y = f(x)$ and $f_{\mathrm{emb}}(x, y) = 0$ if $y \neq f(x)$. Then, a straightforward calculation shows that

$$\left(\left(\mathcal{F}_G^* \otimes \mathcal{F}_H \right) f_{\mathrm{emb}} \right) (\alpha, \beta) = \sqrt{|G|/|H|} \cdot \mathrm{Cor}(f; \alpha, \beta) \tag{16}$$

holds. (This equation corresponds to Eq. (2) for the usual linear cryptanalysis over $(\mathbb{Z}/2\mathbb{Z})^{\oplus n}$.)

H.3 Extension of Modified Simon's subroutine

For an arbitrary finite abelian group G, we assume that elements of G are appropriately encoded into n-bit strings for some n s.t. $|G| \leq 2^n$. Let $\psi : G \to \mathbb{C}$ be a function satisfying $\sum_{x \in G} |\psi(x)|^2 = 1$, and $|\psi\rangle := \sum_x \psi(x) |x\rangle$ be a quantum state. Recall that the Quantum Fourier Transform (QFT) over an abelian group G, denoted by QFT_G, is defined by

$$\mathrm{QFT}_G |\psi\rangle = \sum_x (\mathcal{F}_G^* \psi)(x) |x\rangle . \tag{17}$$

With these notations, the extension of the modified Simon's subroutine $\mathcal{L}$ on a function $f : G \to H$ (G and H are finite abelian groups) is obtained by replacing the Hadamard transform in $\mathcal{L}$ with the QFT (or its inverse) over G and H. Specifically, the extended algorithm runs as follows.

Extended Version of Modified Simon's Subroutine.

(a) Prepare the initial state $|0_G\rangle\,|0_H\rangle$.
(b) Apply QFT_G on the first (left) register.
(c) Apply U_f on the state (i.e., make a quantum query to f).
(d) Apply $\mathrm{QFT}_G \otimes \mathrm{QFT}_H^*$ on the state.
(e) Measure the entire state by the computational basis and return the observed result $(\alpha, \beta) \in G \oplus H$.

We also use the symbol $\mathcal{L}^f$ to denote the extended algorithm.

The following proposition is an extension of Proposition 3.

Proposition 15. *The quantum state of $\mathcal{L}^f$ before the final measurement is*

$$\sum_{\alpha\in G,\beta\in H} \frac{\mathrm{Cor}(f;\alpha,\beta)}{\sqrt{|H|}}\,|\alpha\rangle\,|\beta\rangle\,. \tag{18}$$

In particular, for any subset $S \subset G \oplus H$,

$$\Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^f : (\alpha,\beta) \in S\right] = \sum_{(\alpha,\beta)\in S} \frac{\mathrm{Cor}(f;\alpha,\beta)^2}{|H|} \tag{19}$$

holds.

Proof. The quantum state of $\mathcal{L}^f$ before the final measurement is

$$\begin{aligned}
&(\mathrm{QFT}_G \otimes \mathrm{QFT}_H^*)U_f\,(\mathrm{QFT}_G \otimes I_n)\,|0_G\rangle\,|0_H\rangle \\
&\quad= (\mathrm{QFT}_G \otimes \mathrm{QFT}_H^*)U_f \sum_{x\in G} \frac{1}{\sqrt{|G|}}\,|x\rangle\,|0_H\rangle \\
&\quad= (\mathrm{QFT}_G \otimes \mathrm{QFT}_H^*) \sum_{x\in G} \frac{1}{\sqrt{|G|}}\,|x\rangle\,|f(x)\rangle \\
&\overset{\text{Def. of } f_{\mathrm{emb}}}{=} (\mathrm{QFT}_G \otimes \mathrm{QFT}_H^*) \sum_{x\in G, y\in H} \frac{f_{\mathrm{emb}}(x,y)}{\sqrt{2^m}}\,|x\rangle\,|y\rangle \\
&\overset{\text{Def. of QFT}}{=} \sum_{\alpha\in G,\beta\in H} \frac{((\mathcal{F}_{\mathcal{G}}^* \otimes \mathcal{F}_{\mathcal{H}})f_{\mathrm{emb}})(\alpha,\beta)}{\sqrt{|G|}}\,|\alpha\rangle\,|\beta\rangle \\
&\overset{\text{Eq. (16)}}{=} \sum_{\alpha\in G,\beta\in H} \frac{\mathrm{Cor}(f;\alpha,\beta)}{\sqrt{|H|}}\,|\alpha\rangle\,|\beta\rangle\,.
\end{aligned}$$

Hence we have Eq. (18). Eq. (19) immediately follows from Eq. (18). □

H.4 Quantum Speed-up for Generalized Linear Distinguishers

This section shows how quantum speed-up for generalized linear distinguishers can be obtained. Before showing distinguishers, we define linearly independent masks and linearly completely dependent masks.

Linear (In)dependence of Input and Output Masks. Let $f : G \to H$ be a function, where G and H are finite abelian groups.

1. Suppose G and H are decomposed as $G = G_1 \oplus G_2$ and $H = H_1 \oplus H_2$. Then, we say the set $G_1 \oplus H_1 (\subset G \oplus H)$ is a set of linearly independent input-output masks.
2. Suppose again the decomposition $G = G_1 \oplus G_2$ and $H = H_1 \oplus H_2$, and assume that there is a group isomorphism $\phi : G_1 \to H_1$. Then we say that the set $\{(g, \phi(g)) | g \in G_1\}$ is a set of linearly completely dependent input-output masks.

We show distinguishers when input-output masks are linearly independent or completely dependent, which correspond to $\mathcal{A}_2$ and $\mathcal{A}_3$ in Section 4.3, respectively. We provide only rough ideas and heuristic estimations, and omit detailed analysis.

Distinguisher for Linearly Independent Input-Output Masks. Suppose $f_K : G \to H$ is a keyed function, G and H are decomposed as $G = G_1 \oplus G_2$, $H = H_1 \oplus H_2$, and $\sum_{\alpha \in G_1, \beta \in H_1} \mathrm{Cor}(f_K; \alpha, \beta)^2 |\alpha\rangle |\beta\rangle \gg \frac{1}{|G_2|}$ holds. Let $f_K^{(1)} : G \to H_1$ be the projection of f_K onto H_1. In addition, let $F : G_1 \times H_1 \to \{0, 1\}$ be the binary function such that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in G_1 \oplus H_1$. Then, from Proposition 15,

$$p_{\mathrm{real}} := \Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{f_K^{(1)}} : F(\alpha, \beta) = 1\right] = \sum_{(\alpha,\beta) \in G_1 \oplus H_1} \frac{\mathrm{Cor}(f_K; \alpha, \beta)^2}{|H_1|}$$

follows. On the other hand, for a random function $\mathsf{RF} : G \to H$ we have

$$p_{\mathrm{ideal}} := \Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{\mathsf{RF}^{(1)}} : F(\alpha, \beta) = 1\right] = \sum_{(\alpha,\beta) \in G_1 \oplus H_1} \frac{\mathrm{Cor}(\mathsf{RF}; \alpha, \beta)^2}{|H_1|}$$
$$\approx \sum_{(\alpha,\beta) \in G_1 \oplus H_1} \frac{1}{|G|} \cdot \frac{1}{|H_1|} = \frac{1}{|G_2|}.$$

(We heuristically assume the third equality approximately holds due to [6, Theorem 3.2].) Since $p_{\mathrm{real}} \gg p_{\mathrm{ideal}}$ holds by assumption, we can distinguish f_K from RF by applying the QAA on $\mathcal{L}^{f_K^{(1)}}$ (or $\mathcal{L}^{\mathsf{RF}^{(1)}}$) and F with $O(\sqrt{1/p_{\mathrm{real}}})$ iterations.

Distinguisher for Linearly Completely Dependent Input-Output Masks. Again, let $f_K : G \to H$ be a keyed function, G and H are decomposed as $G = G_1 \oplus G_2$, $H = H_1 \oplus H_2$. Moreover, assume there is a group isomorphism $\phi : G_1 \to H_1$ and $\sum_{\alpha \in G_1} \mathrm{Cor}(f_K; \alpha, \phi(\alpha))^2 |\alpha\rangle |\alpha\rangle \gg \frac{1}{|G|}$ holds. Let $F : G_1 \times H_1 \to \{0, 1\}$ be the binary function such that $F(\alpha, \beta) = 1$ iff $\alpha \in G_1$ and $\beta = \phi(\alpha)$. Then, from Proposition 15,

$$p_{\mathrm{real}} := \Pr\left[(\alpha, \beta) \leftarrow \mathcal{L}^{f_K^{(1)}} : F(\alpha, \beta) = 1\right] = \sum_{\alpha \in G_1} \frac{\mathrm{Cor}(f_K; \alpha, \phi(\alpha))^2}{|H_1|}$$

follows. On the other hand, for a random function $\mathsf{RF} : G \to H$ we have

$$p_{\text{ideal}} := \Pr\left[(\alpha,\beta) \leftarrow \mathcal{L}^{\mathsf{RF}^{(1)}} : F(\alpha,\beta) = 1\right] = \sum_{\alpha \in G_1} \frac{\mathrm{Cor}(\mathsf{RF}; \alpha, \phi(\alpha))^2}{|H_1|}$$

$$\approx \sum_{\alpha \in G_1} \frac{1}{|G|} \cdot \frac{1}{|H_1|} = \frac{1}{|G|}.$$

(We heuristically assume the third equality approximately holds due to [6, Theorem 3.2].) Since $p_{\text{real}} \gg p_{\text{ideal}}$ holds by assumption, we can distinguish f_K from RF by applying the QAA on $\mathcal{L}^{f_K^{(1)}}$ (or $\mathcal{L}^{\mathsf{RF}^{(1)}}$) and F with $O(\sqrt{1/p_{\text{real}}})$ iterations.

Application to the FF3-1 Structure. As mentioned in Section 4.4, Beyne [6] also showed generalized linear distinguishers on the FF3-1 structure in addition to linear distinguishers on FEA. The FF3-1 structure is almost the same as the FEA-1 structure (see Fig. 1). However, the XOR operations in FEA-1 are replaced with modular additions in FF3-1. Thus, generalized linear distinguisher is more suitable for the FF3-1 structure.

The (generalized) linear approximation for FF3-1 in [6] is similar to the multidimensional linear approximation for FEA-1, but underlying groups are different from $\mathbb{Z}_2^n$. In fact, firstly a keyed function $F_K : \mathbb{Z}/2^{n/2}\mathbb{Z} \oplus \mathbb{Z}_2^t \to \mathbb{Z}/2^{n/2}\mathbb{Z}$ is built from the FF3-1 structure by fixing some inputs (here, input means plaintext and tweak) and truncating some outputs, and the distinguisher is applied F_K. The set (sub-group) of input-output masks is given by $\big\{((\alpha,0),\alpha) \in (\mathbb{Z}/2^{n/2}\mathbb{Z} \oplus \mathbb{Z}_2^t) \oplus \mathbb{Z}/2^{n/2}\mathbb{Z} \big| \alpha \in \mathbb{Z}/2^{n/2}\mathbb{Z}\big\}$. In particular, the input-output masks are linearly completely dependent. The corresponding sum of the squared correlation is estimated as $\sum_{\alpha \in \mathbb{Z}/2^{n/2}\mathbb{Z}} \mathrm{Cor}(F_K; (\alpha,0), \alpha) \approx 2^{-n(1-r/4)}$, and the classical distinguishing complexity is $O(2^{(r/4-3/4)n})$.

On the other hand, if we apply the quantum distinguisher explained above, we achieve the complexity $O(2^{(r/8-1/4)n})$.