

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

Akinori Hosoyamada

NTT Social Informatics Laboratories, Tokyo, Japan
akinori.hosoyamada@ntt.com

Abstract. This paper shows how to achieve a quantum speed-up for multidimensional (zero correlation) linear distinguishers. To understand the post-quantum security of symmetric-key cryptosystems, it is important to study how much quantum speed-up can be obtained for classical cryptanalytic techniques such as differential, linear, and integral cryptanalysis. A previous work by Kaplan et al. has already shown a quantum quadratic speed-up for one-dimensional linear distinguishers. However, classical linear cryptanalysis often exploits multidimensional approximations to achieve more efficient attacks, and in fact it is highly non-trivial whether Kaplan et al.'s technique can be extended into the multidimensional case. To remedy this, we investigate a new quantum technique to speed-up multidimensional linear distinguishers. Firstly, we observe that there is a close relationship between the subroutine of Simon's algorithm and linear correlations via Fourier transform. Specifically, a slightly modified version of Simon's subroutine, which we call Correlation Extraction Algorithm (CEA), can be used to speed-up multidimensional linear distinguishers. CEA also leads to a speed-up for multidimensional zero correlation distinguishers, as well as some integral distinguishers through the correspondence of zero correlation and integral properties shown by Bogdanov et al. and Sum et al. Furthermore, we observe possibility of a more than quadratic speed-ups for some special types of integral distinguishers when *multiple* integral properties exist. Especially, we show a single-query distinguisher on a 4-bit cell SPN cipher with the same integral property as 2.5-round AES. Our attacks are the first to observe such a speed-up for classical cryptanalytic techniques without relying on any algebraic structures such as hidden periods or shifts. By replacing the Hadamard transform in CEA with the general quantum Fourier transform, our technique also speeds-up generalized linear distinguishers on an arbitrary finite abelian group.

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

1 Introduction

Research in the past decade has revealed possible quantum attacks on symmetric cryptosystems are not limited to the exhaustive key search with Grover's algorithm [30] or the collision search by the BHT algorithm [19]. A notable line of

research is the one initiated by Kuwakado and Morii showing that Simon’s algorithm breaks lots of classically secure schemes in polynomial time [45, 46, 41, 13]. Other previous works show how to speed-up classical cryptanalytic techniques such as differential and linear cryptanalysis, MITM, and integral attacks [42, 36, 15]. Some recent papers study dedicated quantum collision attacks on concrete hash functions such as SHA-2 and SHA-3 [37, 28, 38, 31].

Despite of the discovery of many interesting attacks, lots of interesting topics have not yet been well studied in quantum cryptanalysis. In this paper, we investigate the possibility to achieve more quantum speed-up for major classical cryptanalytic techniques than previous works.

Q1 and Q2. For quantum cryptanalysis on symmetric cryptosystems, there are two attack models called Q1 and Q2 [42]. The Q1 model assumes the existence of a quantum computer but oracles of keyed functions are classical. Meanwhile, Q2 assumes that oracles are also quantum¹. For instance, suppose an attack target is a cipher. Then, an attacker in the Q2 model is allowed to query quantum superposition of messages to the encryption oracle. Such attacks are called Quantum Chosen-Plaintext Attacks (QCPAs). This paper studies attacks in the Q2 model.

Significance of Studying Q2 Attacks. The Q1 model is more realistic than the Q2 model in that oracles in the Q1 model are the same as classical ones and attacks in the Q1 model become real threats as soon as a large-scale fault-tolerant quantum computer is available. Still, we argue that it is important to study not only Q1 but also Q2 attacks 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 by Bonnetain et al. [12], which is a Q1 attack, is developed by modifying the Q2 attack by Leander and May [47]. Second, Q2 attacks can be converted into Q1 attacks when the 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, all the (classical) inputs to E_K being queried and the results stored in a qRAM. Then the quantum oracle of E_K can be simulated by accessing the qRAM. Especially, the Q2 attack can be mounted by using the simulated quantum oracle. This is still a valid Q1 attack since now $2n < k$ is assumed 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 quantum superposition queries are required only on some small portion of inputs.

¹ When attack targets are primitives without secret keys, e.g. hash functions, it is reasonable to assume attackers can compute all functions in quantum superposition. (In other words, we are always in the Q2 model, or there is no distinction between Q1 and Q2.)

Quantum Speed-Up for Linear Cryptanalysis. Linear cryptanalysis [49] is one of the most fundamental cryptanalytic techniques in symmetric cryptology. Kaplan et al. [42] has already shown how to achieve a quadratic speed-up for linear attacks. However, their distinguisher uses only one-dimensional linear approximations, while classical linear cryptanalysis often exploit multidimensional linear approximations to mount more efficient attacks [32–34]. In fact, it is unclear whether Kaplan et al.’s distinguisher can be sped-up further even if multiple linear approximations are available, due to the following reason.

Kaplan et al.’s distinguisher relies on the quantum counting algorithm [18], which (approximately) counts the number of x satisfying $F(x) = 1$ for an (efficiently computable) Boolean function F . Since classical one-dimensional linear distinguishers work just by counting the number of messages satisfying a linear approximation, such F is naturally defined in the one-dimensional case, and the quantum counting algorithm can be applied.

Meanwhile, classical multidimensional linear distinguishers are based on sophisticated statistical tests exploiting a relationship between capacity and a sum of squared correlations in a clever way. It is highly unclear whether there exists an efficiently computable Boolean function F such that just counting the number of x satisfying $F(x) = 1$ corresponds to performing such statistical tests.

Thus it is natural to ask whether there exists another quantum technique for linear distinguishers running faster than Kaplan et al.’s when a multidimensional linear approximation is available.

Multidimensional linear cryptanalysis has many variants including (multidimensional) zero correlation linear cryptanalysis [9] and generalized linear cryptanalysis on an arbitrary finite abelian group [4], though, so far no previous work has shown quantum speed-up for these attacks. A technique speeding-up multidimensional linear distinguishers may also lead to a speed-up for these variants, which is of another interest. Moreover, as shown by Bogdanov et al. and Sun et al., integral distinguishers based on balanced functions correspond to a class of multidimensional zero correlation linear distinguishers [8, 59]. Hence we may obtain a new quantum algorithm speeding-up integral distinguishers through this correspondence.

Quadratic Barrier. The complexity of the generic key-recovery attack on a k -bit key cipher drops from 2^k in the classical setting to $\approx \frac{\pi}{2} 2^{k/2}$ encryptions in the quantum setting by the Grover search. However, the generic complexity cannot be decreased further unless some internal structure of the cipher is exploited. This means the following thing: If we find a dedicated quantum attack on a cipher of which the time complexity is less than the square root of the corresponding classical complexity², then the quantum attack is especially interesting. This is because, if the quantum speed-up for dedicated attacks exceeds that for the generic attack, quantum attacks may break more rounds of the cipher than

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

classical ones do. Hence finding such a quantum attack is one of the main goals in quantum cryptanalysis on symmetric-key cryptosystems.

Some previous works have already achieved such a speed-up, though, the types of such attacks are limited. To achieve a 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 a 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 [41, 14], 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 periods or shifts) has been an important open problem for years.

1.1 Our Contributions

This paper shows that 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 find that, with a slight modification made, the subroutine outputs a pair of linear masks of the target function with probability proportional to the squared linear correlation. Since it extracts linear correlations of a function into quantum amplitude, we call the subroutine after the modification the *correlation extraction algorithm*, or CEA for short.

Second, we show that multidimensional linear distinguishers can be sped-up by combining CEA and the Quantum Amplitude Amplification (QAA) technique. As an application example, we show that the multidimensional distinguishers on FEA-1 and FEA-2 by Beyne [6] can be sped-up from $O(2^{(r/4-3/4)n})$ and $O(2^{(r/6-3/4)n})$ to $O(2^{(r/8-1/4)n})$ and $O(2^{(r/12-1/4)n})$, respectively, when messages are n bits the number of rounds is r .

Then we show that CEA also leads to a speeds-up for multidimensional zero correlation linear distinguishers. 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 Type-I/II generalized Feistel structures. (See Table 2 in the appendix for details.)

Finally, we show a speed-up for integral distinguishers. The speed-up is obtained via the correspondence of integral and zero correlation properties observed by Bogdanov et al. and Sun et al. [8, 59], and applicable when integral properties are based on balanced functions. Especially, we observe the possibility of a more than quadratic speed-up when there are *multiple* integral properties on mutually

orthogonal subspaces, which appear in some SPN ciphers such as the 3.5-round AES. As a notable example, we show that 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 a single-query attack is almost impossible in the classical setting (unless another weakness exists), and the example illustrates a new type of qualitative difference between classical and quantum computation 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 (a modified) Simon’s algorithm, which was primarily developed to solve an algebraic problem of hidden periods, leads to non-trivial speed-ups for various classical attacks not relying on algebraic structures.

Our technique extends to generalized linear distinguishers on arbitrary finite abelian groups [4] by replacing the Hadamard transform in CEA with the general Quantum Fourier Transform (QFT). For instance, we can also achieve the generalized linear distinguisher on the FF3-1 structure by Beyne [6]. The amount of speed-up is the same as that for FEA-1.

A drawback of our technique 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. Thus our technique does not directly lead to breaking more rounds than classical attacks. Still, we believe that our techniques are novel and general, and will inspire other new types of quantum attacks on symmetric cryptosystems in both of the Q1 and Q2 models.

1.2 Related Works

Bonnetain et al. [16] showed a more than quadratic speed-up can be obtained even in the Q1 model. The difference between their results and ours is that theirs exploit periodic structures to apply offline Simon’s algorithm [12] in a clever way, while our goal is to speed-up classical cryptanalytic properties without periods or shifts. Bonnetain [11] showed a single query attack on the one-time pad encryption scheme by making quantum registers for messages and ciphertexts disentangled. Quantum speed-up for integral attacks has already been studied in, e.g., [15], but zero-sum properties are used (because they break more rounds than balanced properties) and the distinguisher part itself is not sped-up. Recently, Shi et al. published a paper of which the title is “Quantum zero correlation linear cryptanalysis” [57]. 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.

Schrottenloher’s Key-Recovery Attack. Another recent concurrent and independent work by Schrottenloher [56] showed how to obtain quantum speed-up for linear *key-recovery* attacks.

Classical linear distinguishers are often combined with efficient key-recovery attacks using the FFT [23, 60, 29]. What Schrottenloher did is to combine such

classical techniques with the QFT. Computing convolution of some Boolean functions related to linear approximations in quantum superposition, Schrottenloher’s algorithm produces some quantum superposition of *subkey candidates* in such a way that the quantum amplitude are proportional to their experimental correlations. Then the amplitude of the right key is amplified by QAA.

Note that our main interest is to achieve a speed-up for *multidimensional* (zero correlation) linear distinguishers. Schrottenloher’s work [56] also deals with multiple linear approximations, but the existence of multiple approximations improves only precision of attack by a constant factor, and essentially does not contribute much to reducing the time complexity. Additionally, zero correlation linear or integral attacks are not studied in [56].

One would expect that a more speed-up for key-recovery is obtained by combining our technique and Schrottenloher’s. Still, the mechanism of the two techniques is quite different (Schrottenloher uses the QFT to compute convolution in superposition to obtain a superposition of key candidates, while we use it to extract correlations of multidimensional linear approximations), and so far we do not have any idea on how to combine them. Studying theoretical connection between them and reducing the time complexity of key-recovery exploiting (zero correlation) multidimensional linear approximations is definitely an important and interesting future work.

1.3 Organization

Section 2 introduces basic notions and facts used throughout the paper. Section 3 reviews classical (multidimensional) linear distinguishers and Kaplan et al.’s quantum one-dimensional linear distinguisher. Section 4 studies relationships between the Simon’s subroutine and linear correlations, and introduces CEA. Sections 5, 6, and 7 show how to achieve a quantum speed-up with CEA for multidimensional linear, zero correlation multidimensional linear, and integral distinguishers, respectively. Section 8 shows the extension to generalized linear distinguishers on an arbitrary finite abelian group. Section 9 concludes the paper.

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 . Especially, by “bit string” we denote an element of \mathbb{F}_2^n for some n . By \mathbf{e}_i we denote the n -bit string (for some n) of which the i -th bit is 1 and other bits are 0. $x \oplus y$ denotes the addition of x and y in \mathbb{F}_2^n , and $x||y$ denotes the concatenation as bit strings. For a bit string $x \in \mathbb{F}_2^n$, 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 $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$), by $\text{msb}_u[x]$ (resp., $\text{msb}_u[f]$) we denote the most significant u bits of x (resp., the function that returns $\text{msb}_u[f(x)]$ for each input x). The notations $\text{lsb}_u[x]$ and $\text{lsb}_u[f]$ are similarly defined for least significant u bits. For a unitary operator U , its adjoint is denoted by U^* . In cryptanalysis of a block cipher E , we regard the unit of time as the time to encrypt a message by E . We assume that readers are familiar with Pearson's chi-squared test of goodness-of-fit. For those who are not, we provide a brief overview about the relationship between the test and distinguishers in Section A in the appendix.

2.1 Linear Approximations and Correlations

The (one-dimensional) linear approximation of a function $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ 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 $\text{Cor}(f; \alpha, \beta)$ of this linear approximation is defined by $\text{Cor}(f; \alpha, \beta) := \Pr_x[\alpha \cdot x = \beta \cdot f(x)] - \Pr_x[\alpha \cdot x \neq \beta \cdot f(x)]$. It is well-known that the linear correlation satisfies

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

The following property is useful to analyze 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 use this property without any mention.

In addition, we need the following claim for analysis of attacks.

Claim 1 (Distribution of capacity of a random permutation.) *Let $V \subset \mathbb{F}_2^v \times \mathbb{F}_2^u$ be a vector space and S be an arbitrary basis of V . Then, for a randomly chosen permutation P , the value $2^n \cdot \sum_{(\alpha, \beta) \in V - \{0\}} \text{Cor}(P; \alpha, \beta)^2$ approximately follows the χ^2 distribution with $2^v - 2^u - 2^w + 1$ degrees of freedom. Here, $v := \dim(V)$, $u := \dim(V \cap \mathbb{F}_2^v \times \{0^u\})$ and $w := \dim(V \cap \{0^v\} \times \mathbb{F}_2^u)$.*

We do not have a completely formal proof for the claim, but explain why this claim is plausible in Section D in the appendix.

2.2 Balanced Function and Zero-Sum Property

Integral cryptanalysis [44], which was initially proposed as a dedicated attack on the block cipher SQUARE [24], exploits a *zero-sum property* of (a part of) ciphers. Here, we say that a function $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ has a zero-sum property if $\sum_x f(x) = 0$. An important special version of zero-sum properties is a balanced function, of which the definition is as follows.

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

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

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.

2.3 Quantum Computation

We assume that the readers are familiar with quantum computation and linear algebra (see, e.g., [53] 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 $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$, 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 [18], which is a generalization of Grover’s algorithm [30]. Let $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2$ 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 the state $U|0^m\rangle$ is measured 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 $\mathcal{S}_0|x\rangle = (-1)^{\delta_{0^m,x}}|x\rangle$, where $\delta_{0^m,x}$ is Kronecker’s delta.

Proposition 1 (Quantum amplitude amplification). *With the above setting and notations, let $Q(U, f) := -U\mathcal{S}_0U^*\mathcal{S}_f$. When the state $Q(U, f)^iU|0^m\rangle$ 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, such an x is obtained with probability at least $\max\{p, 1 - p\}$ by setting $i = \lceil \pi/4 \arcsin(\sqrt{p}) \rceil$.*

We obtain Grover’s algorithm when $U = H^{\otimes m}$. In this case $p = |f^{-1}(1)|/2^m$ holds and an x satisfying $f(x) = 1$ is found 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 from³.

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 conditions.

- (1) If U is chosen from D_1 , then the probability $p_U := \Pr \left[x \stackrel{\text{measure}}{\leftarrow} U |0^n\rangle : F(x) = 1 \right]$ is relatively high on average.
- (2) If U is chosen from D_2 , then p_U is relatively low on average.

Specifically, 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 , then QAA with $O(\sqrt{t^{-1}})$ applications of U , U^* , and \mathcal{S}_F will find x satisfying $F(x) = 1$ because $p_U \geq t$. If U is chosen from D_2 , such QAA will not find x because $t \gg p_U$ and the number of iteration is not large enough.

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.

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

1. For $i = 1, \dots, s$, do:
 - (a) Choose i uniformly at random from the set of integers from 0 to $\left\lfloor \frac{1}{\sin(2 \cdot \arcsin(\sqrt{t}))} \right\rfloor$.
 - (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.

Here, s is a positive integer constant chosen depending on applications. We denote the above algorithm by \mathcal{A}_0 .

The idea of randomly choosing the number of iteration follows previous works on Grover's algorithm and QAA without knowing initial success probability [18, 17]. The above algorithm is just a straightforward adaptation of the previous ideas.

Proposition 2. *With the above setting and notions, suppose $1/4 > t > 0$. Then, \mathcal{A}_0 applies U , U^* , and \mathcal{S}_F at most $s(\frac{1}{\sqrt{t}} + 1)$ times and (1) returns 1 with*

³ A typical example is that D_1 corresponds to the quantum encryption oracle U_{E_K} of a block cipher E_K while D_2 to the oracle U_P of a random permutation P (choosing K or P randomly corresponds to sampling according to D_1 or D_2).

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 $s \cdot (\frac{16t'}{t} + \frac{20t'}{\sqrt{t}}) + \Pr_{U \sim D_2} [t' < p_U]$ for any $t' > 0$ satisfying $4\sqrt{t'}/t + 2\sqrt{t'} < \pi/2$ if U is chosen according to D_2 .

The interpretation of the proposition is as follows. Suppose that p_U is distributed around t (resp., t') if U is chosen according to D_1 (resp., 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 [17, 18], though, we provide a proof in Section B in the appendix for completeness.

Simon's algorithm. Simon's quantum algorithm [58] finds a period of a periodic function. More precisely, it solves the following problem.

Problem 1. Let $s \in \mathbb{F}_2^m$ be a (secret) constant, and $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ be a function satisfying the following properties C1 and C2.

- 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 $i = 1, 2, \dots, 2m$, 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 $\dim(\text{Span}_{\mathbb{F}_2}(\alpha_1, \dots, \alpha_{2m})) = m - 1$, compute and output the unique $s' \in \mathbb{F}_2^m \setminus \{0^m\}$ such that $s' \cdot \alpha_i = 0$ for $i = 1, \dots, 2m$. If $\dim(\text{Span}_{\mathbb{F}_2}(\alpha_1, \dots, \alpha_{2m})) \neq m - 1$, output \perp .

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

Many papers (e.g., [45, 46, 41]) 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 $\alpha_i \cdot s = 0$ [41].

3 Classical and Kaplan et al.'s Linear Distinguishers

Here we review classical (multidimensional) linear distinguishers and Kaplan et al.'s quantum one-dimensional linear distinguisher [42].

3.1 Classical One-Dimensional Linear Distinguisher

The linear correlation $\text{Cor}(P; \alpha, \beta)$ of an n -bit random permutation P approximately follows the normal distribution $\mathcal{N}(0, 2^{-n})$ for an arbitrary mask (α, β) with $\alpha, \beta \neq 0^n$ [26]. Thus, if a linear correlation $\text{Cor}(E_K; \alpha, \beta)$ of a block cipher E_K with $\alpha, \beta \neq 0^n$ 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-ciphertext pairs $L = \{(P_1, C_1), \dots, (P_N, C_N)\}$ and checking if the estimated empirical correlation

$$\widehat{\text{Cor}}(E_K; \alpha, \beta) = \frac{\#\{(P, C) \in L \mid \alpha \cdot P = \beta \cdot C\} - \#\{(P, C) \in L \mid \alpha \cdot P \neq \beta \cdot C\}}{N}$$

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

3.2 Classical Multidimensional Linear Distinguishers

A natural idea to enhance the power of linear distinguishers is to utilize multiple linear approximations. Some early works indeed show such attacks, assuming the existence of multiple approximations that are statistically independent [40, 7]. However, the assumption does not necessarily hold in general [50]. Instead, Hermelin et al. [35] proposed to use multidimensional linear approximations, i.e., sets of linear approximations of which the input-output masks form a vector space.

Specifically, let $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ 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 $\text{Lin}_S^f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^\ell$ such that

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

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

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 $V_1 \subset \mathbb{F}_2^m$ and $V_2 \subset \mathbb{F}_2^n$. Moreover, we say that the input and output masks are linearly completely dependent if there exists a basis $\{(\alpha_i, \beta_i)\}_{1 \leq i \leq \dim(V)}$ of V such that both of $\{\alpha_i\}_{1 \leq i \leq \dim(V)}$ and $\{\beta_i\}_{1 \leq i \leq \dim(V)}$ 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^ℓ is the value defined⁴ by

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

⁴ In fact this is the χ^2 -divergence between p and the uniform distribution over \mathbb{F}_2^ℓ . We use the term *capacity* following previous works on linear cryptanalysis.

The important well-known fact is that

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

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 $\text{Cap}(\hat{p}_S^f)$ of the estimated empirical distribution \hat{p}_S^f (defined by $\hat{p}_S^f(z) := \frac{\#\{(P, C) \in L \mid \text{Lin}_S^f(P) = z\}}{N}$) 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}_2^ℓ).

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. (2) 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 Pearson's chi-squared test is larger than a certain threshold. Specifically, given a list of (real) random plaintext-ciphertext pairs $L = \{(P_1, C_1), \dots, (P_N, C_N)\}$, we count the number $\text{num}(z) := \#\{(P_i, C_i) \in L \mid \text{Lin}_S^{E_K}(P_i) = z\}$ for each z , and compute the test statistic $\chi_{\text{real}}^2 := N2^\ell \sum_z (\text{num}(z)/N - 2^{-\ell})^2 = N \cdot \text{Cap}(\hat{p}_S^{E_K})$. Then χ_{real}^2 is approximately distributed around $(2^\ell - 1) + N \sum_{(\alpha, \beta) \in V - \{\mathbf{0}\}} \text{Cor}(E_K; \alpha, \beta)^2$. If the plaintext-ciphertext pairs are generated from a truly random permutation, then $\text{num}(z)$ approximately follows the uniform distribution. Thus, the similarly computed statistic χ_{ideal}^2 approximately follows the χ^2 distribution with $(2^\ell - 1)$ degrees of freedom (denoted by $\chi_{2^\ell - 1}^2$), of which the standard deviation is $\sqrt{2(2^\ell - 1)}$. Hence E_K can be distinguished from a random permutation with a constant advantage when⁵ $N \gg \sqrt{2^\ell} / \sum_{(\alpha, \beta) \in V - \{\mathbf{0}\}} \text{Cor}(E_K; \alpha, \beta)^2 = \sqrt{2^\ell} / \text{Cap}(p_S^{E_K})$, by checking whether the test statistic is larger than $(2^\ell - 1) + \sqrt{2(2^\ell - 1)}$ or not.

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

⁵ The values of squared correlation and capacity can significantly change depending on keys in general, but they are often estimated by their averages under the assumption that they concentrate around the mean. As the first step of achieving quantum speed-up for multidimensional linear attacks, we also focus on the case where this assumption holds. Studying the dependence of complexity on keys in detail is an important future work.

Remark 2. Some early works showed that distinguishers based on the Log Likelihood Ratio (LLR) test [3, 32, 33] requires only $O(1/\text{Cap}(p_S^{E_K}))$ data instead of $O(\sqrt{2^\ell}/\text{Cap}(p_S^{E_K}))$ 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 [22].

3.3 Kaplan et al.’s Quantum One-Dimensional Linear Distinguisher

Kaplan et al. [42] observed that a quadratic quantum speed-up can be obtained for linear distinguishers by using the quantum counting algorithm [18].

Roughly speaking, the quantum counting algorithm achieves a quadratic speed-up to solve the problem of estimating $M := \#\{x|F(x) = 1\}$ for a Boolean function F . Making $O(q)$ quantum queries to F , it returns an approximation \tilde{M} of M satisfying $|\tilde{M} - M| \leq O\left(\frac{\sqrt{M(2^n - M)}}{q} + \frac{2^n}{q^2}\right)$.

Now, suppose that there exists a linear approximation of an n -bit block cipher E_K satisfying $c := |\text{Cor}(E_K; \alpha, \beta)| \gg 2^{-n/2}$ for a random key K , and let F be the Boolean function such that $F(x) = 1$ iff $\alpha \cdot x \oplus \beta \cdot E_K(x) = 0$. Then, E_K can be distinguished by estimating $M = \#\{x|F(x) = 1\}$ and checking whether $|M - \frac{2^n}{2}| \gg 2^{n/2}$. Using the quantum counting algorithm, one can obtain an estimation of \tilde{M} with sufficient precision for distinguisher ($|\tilde{M} - M| \leq \frac{M}{a}$ for a small integer $a > 0$) in time $O(1/c)$. Compared to the classical complexity of $O(1/c^2)$, a quadratic speed-up is achieved.

Extension to Multidimensional Linear Distinguishers? After seeing Kaplan et al.’s work, it is natural to ask whether their technique can be extended to multidimensional linear distinguishers. However, to apply the quantum counting algorithm to solve a problem, one has to construct an efficiently computable Boolean function F in such a way that counting the number of x satisfying $F(x) = 1$ solves the problem. In the one-dimensional case F is obtained in a quite natural way as explained above, but in the multidimensional case essentially we have to construct F in such a way to achieve a quadratic speed-up for Pearson’s chi-squared test applied to the distribution $p_S^{E_K}$. It seems highly unclear whether such F exists, and thus we seek for another technique.

4 New Observation on Simon’s Algorithm

As explained in Section 2, the subroutine of Simon’s algorithm uses only the quantum oracle of a target function and the Hadamard transform, which is the Fourier transform over the group $(\mathbb{Z}/2\mathbb{Z})^n$. Meanwhile, a well-know fact is that linear correlations have strong relationships with Fourier transform. This section shows that a slightly modified version of Simon’s subroutine, which we call CEA, returns input and output masks of a function with a probability proportional to

the linear correlations. Later we show that CEA can be utilized to obtain speed-up for various techniques including multidimensional linear distinguishers. Since the underlying additive group of \mathbb{F}_2^n is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$, in what follows we identify \mathbb{F}_2^n with $(\mathbb{Z}/2\mathbb{Z})^n$.

4.1 Fourier Transform

First, we recall the Fourier transform (over $(\mathbb{Z}/2\mathbb{Z})^n$) and its relationship with linear cryptanalysis and quantum computation. The Fourier transform of a function $F : \mathbb{F}_2^n \rightarrow \mathbb{C}$ is the function $\mathcal{F}F : \mathbb{F}_2^n \rightarrow \mathbb{C}$ defined by $\mathcal{F}F(x) := \sum_{y \in \mathbb{F}_2^n} \frac{(-1)^{x \cdot y} F(y)}{\sqrt{2^n}}$.

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 $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$, let $f_{\text{emb}} : \mathbb{F}_2^m \times \mathbb{F}_2^n \rightarrow \mathbb{C}$ be the function defined by $f_{\text{emb}}(x, y) = 1$ if $f(x) = y$ and $f_{\text{emb}}(x, y) = 0$ otherwise⁶. Then some straightforward calculation shows

$$\mathcal{F}f_{\text{emb}}(\alpha, \beta) = \sqrt{2^{m-n}} \cdot \text{Cor}(f; \alpha, \beta). \quad (3)$$

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 $\psi : \mathbb{F}_2^n \rightarrow \mathbb{C}$ and $|\psi\rangle := \sum_{x \in \mathbb{F}_2^n} \psi(x) |x\rangle$. Then

$$H^{\otimes n} |\psi\rangle = \sum_{y \in \mathbb{F}_2^n} \mathcal{F}\psi(y) |y\rangle \quad (4)$$

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 $O(1)$.

4.2 Extracting Correlations by (Modified) Simon's Subroutine

Here we show that Simon's subroutine with a slight modification returns input and output masks for linear approximations with high correlation. We call the resulting algorithm Correlation Extraction Algorithm (CEA) because it extracts linear correlations into the quantum amplitude of a state, and we denote it by CEA^f when applied to a function f . Specifically, the algorithm runs as follows.

⁶ "emb" is an abbreviation of "embedding".

Algorithm CEA^f .

- (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 $\alpha||\beta$ ($\alpha \in \mathbb{F}_2^m$ and $\beta \in \mathbb{F}_2^n$).

The underlines indicate the parts modified from the original Simon's subroutine on p.10. CEA^f is different from the original Simon's subroutine only in that CEA^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 is just uniformly distributed over the subspace $\{v \in \mathbb{F}_2^m | v \cdot s = 0\}$ if f satisfies the conditions of Problem 1. Thus there is nothing new if we focus only on α . However, we observe that CEA^f shows an interesting link to linear correlations when β is into account, as shown in the following proposition.

Proposition 3. *The quantum state of CEA^f before the final measurement is*

$$\sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\text{Cor}(f; \alpha, \beta)}{\sqrt{2^n}} |\alpha\rangle |\beta\rangle. \quad (5)$$

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

$$\Pr \left[(\alpha, \beta) \leftarrow \text{CEA}^f : (\alpha, \beta) \in S \right] = \sum_{(\alpha, \beta) \in S} \frac{\text{Cor}(f; \alpha, \beta)^2}{2^n} \quad (6)$$

holds.

Proof (of Proposition 3). The quantum state of CEA^f before the final measurement is

$$\begin{aligned} H^{\otimes(m+n)} U_f (H^{\otimes m} \otimes I_n) |0^m\rangle |0^n\rangle &= H^{\otimes(m+n)} U_f \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle |0^n\rangle \\ &= H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle |f(x)\rangle \\ &\stackrel{\text{Def. of } f_{\text{emb}}}{=} H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n} \frac{f_{\text{emb}}(x, y)}{\sqrt{2^m}} |x\rangle |y\rangle \\ &\stackrel{\text{Eq. (4)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\mathcal{F} f_{\text{emb}}(\alpha, \beta)}{\sqrt{2^m}} |\alpha\rangle |\beta\rangle \\ &\stackrel{\text{Eq. (3)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\text{Cor}(f; \alpha, \beta)}{\sqrt{2^n}} |\alpha\rangle |\beta\rangle. \end{aligned}$$

Hence we have Eq. (5). Eq. (6) immediately follows from Eq. (5). \square

Some Remarks. CEA is quite close to the Bernstein-Vazirani algorithm [5] when $n = 1$ and some previous works [20, 62] already observes similar relationships between linear correlations and the Bernstein-Vazirani algorithm. Still, analysis in the 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 an arbitrary finite abelian group and the general quantum Fourier transformation. See Section 8 in the appendix for details.

5 Speed-Up for Multidimensional Linear Distinguishers

By using the CEA in the previous section, here we show quantum linear distinguishers achieving a more speed-up than Kaplan et al.'s when a multidimensional linear approximation with high correlations exists. Recall that what the algorithm CEA^f does is to apply the unitary operator $H^{\otimes(m+n)}U_f(H^{\otimes m} \otimes I_n)$ on $|0^m\rangle|0^n\rangle$ and measure the entire state by the computational basis. By abuse of notation, let CEA^f also denote the operator $H^{\otimes(m+n)}U_f(H^{\otimes m} \otimes I_n)$ itself.

We show three distinguishers⁷ \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). Here are some remarks on notations and assumptions.

- We assume that $\sum_{(\alpha,\beta) \in V - \{\mathbf{0}\}} \text{Cor}(E_K; \alpha, \beta)^2 \geq c$ holds with a high probability for some c , and that we know the value of c . \mathcal{O} denotes the given oracle, which is either E_K for a random K or a random permutation P .
- (Notations for \mathcal{A}_2) When the input and output masks are linearly independent, i.e., $V = V_1 \times V_2$ holds for some subspaces $V_1, V_2 \subset \mathbb{F}_2^n$, we denote $\dim(V_1)$ and $\dim(V_2)$ by u and w , respectively. In addition, $S_1 := \{\alpha_1, \dots, \alpha_u\}$ and $S_2 := \{\beta_1, \dots, \beta_w\}$ denotes basis of V_1 and V_2 . Without loss of generality, we assume $V_2 = \{\beta | |0^{n-w}| \beta \in \mathbb{F}_2^w\}$ and $\beta_i = \mathbf{e}_i$ ⁸. Especially, we regard V as a subspace of $\mathbb{F}_2^n \times \mathbb{F}_2^w$.
- (Notations for \mathcal{A}_3) When the input and output masks are linearly completely dependent, we choose and fix a basis $S := \{(\alpha_i, \beta_i)\}_{1 \leq i \leq \dim(V)}$ of V such that both of $\{\alpha_i\}_{1 \leq i \leq \dim(V)}$ and $\{\beta_i\}_{1 \leq i \leq \dim(V)}$ are linearly independent in \mathbb{F}_2^n .

⁷ In fact the three distinguishers may be unified into a single one by restricting inputs and outputs appropriately. Still we focus on these cases because the three distinguishers are enough for the examples of interest shown later, and to avoid unnecessarily complex analysis.

⁸ 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, we can assume them without loss of generality.

Without loss of generality, we assume $\beta_i = \mathbf{e}_i$ ⁹. Especially, we regard V as a subspace of $\mathbb{F}_2^n \times \mathbb{F}_2^{\dim(V)}$.

Distinguishers \mathcal{A}_1 , \mathcal{A}_2 , and \mathcal{A}_3 . All the three distinguishers are obtained by applying the algorithm \mathcal{A}_0 of Proposition 2. The difference between the distinguishers is the choice of the parameters s and t , the unitary operator U , and the Boolean function¹⁰ F , which is as follows. (Recall that we assume $\beta_i = \mathbf{e}_i$ for \mathcal{A}_2 and \mathcal{A}_3 .)

- \mathcal{A}_1 (**general case**): $(s, t) := (3, c/2^n)$. U is set to be $\text{CEA}^\mathcal{O}$. F is the Boolean function of which the domain is $\mathbb{F}_2^n \times \mathbb{F}_2^n$ and $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$.
- \mathcal{A}_2 (**linearly independent masks**): $(s, t) := (3, c/2^w)$. U is set to be $\text{CEA}^{\text{msb}_w[\mathcal{O}]}$. F is the Boolean function of which the domain is $\mathbb{F}_2^n \times \mathbb{F}_2^w$ and $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$.
- \mathcal{A}_3 (**linearly completely dependent masks**): $(s, t) := (3, c/2^{\dim(V)})$. U is set to be $\text{CEA}^{\text{msb}_{\dim(V)}[\mathcal{O}]}$. F is the Boolean function of which the domain is $\mathbb{F}_2^n \times \mathbb{F}_2^{\dim(V)}$ and $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$.

Here, sampling a unitary U according to D_1 (resp., D_2) in Proposition 2 corresponds to sampling a random key K for a real cipher (resp., choosing an ideally random permutation P).

Analysis. If input and output masks are linearly independent, \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 $\text{CEA}^{\text{msb}_w[E_K]} |0^n\rangle |0^w\rangle$ is approximately lower bounded by $c/2^w$. Hence, QAA on $\text{CEA}^{\text{msb}_w[E_K]}$ 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 $\text{CEA}^{\text{msb}_w[P]} |0^n\rangle |0^w\rangle$ is approximately upper bounded by $2^{\dim(V)}/2^{n+w}$. Especially, the probability that QAA on $\text{CEA}^{\text{msb}_w[P]}$ 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. For similar reasons, \mathcal{A}_1 and \mathcal{A}_3 distinguish E_K and P in time $O(\sqrt{2^n/c})$ and $O(\sqrt{2^{\dim(V)}/c})$, respectively.

More precisely, define parameters \mathbf{c}_i , \mathbf{pb}_i , and \mathbf{T}_i for $i = 1, 2, 3$ as follows.

- $\mathbf{c}_1 := \mathbf{c}_3 = 2^{-n}$, $\mathbf{c}_2 := 2^{-n-w+\dim(V)}$.
- $\mathbf{pb}_1 := \frac{2^{\dim(V)+7}(n+1)}{2^{2n} \cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}$, $\mathbf{pb}_2 := \frac{2^{\dim(V)+7}(n+1)}{2^{n+w} \cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}$, and $\mathbf{pb}_3 := \frac{2^7(n+1)}{2^n \cdot c} + 2^{-\dim(V)+1} \cdot n^{-2}$.
- $\mathbf{T}_1 := 6\sqrt{2^n/c}$, $\mathbf{T}_2 := 6\sqrt{2^w/c}$, and $\mathbf{T}_3 := 6\sqrt{2^{\dim(V)}/c}$.

Then the following proposition holds.

⁹ The reasoning is almost the same as before.

¹⁰ See Section F in the appendix for details on how to efficiently compute F on a quantum circuit.

Proposition 4. *Let $i = 1, 2$, or 3 . Suppose that $c \gg \mathfrak{c}_i$, and that $1/4 > \sum_{(\alpha, \beta) \in V - \{\mathbf{0}\}} \text{Cor}(E_K; \alpha, \beta)^2 \geq c$ holds with a constant probability p when K is randomly chosen. If \mathcal{A}_i 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}_i runs relative to a random permutation P , then the probability that \mathcal{A}_i outputs 1 is approximately upper bounded by pb_i . In addition, the time complexity of \mathcal{A}_i is at most \mathfrak{T}_i . (The probabilities are taken not only over the randomness of \mathcal{A}_i 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. Still, we provide a proof in Section C in the appendix for completeness.

Remark 3. 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).

5.1 Application Example: FEA-1 and FEA-2 Structures

FEA is a Korean standard (TTAK.KO-12.0275) for format preserving encryption [48], which has two variants named FEA-1 and FEA-2. Both variants adopt *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.

At CRYPTO 2021, Beyne showed multidimensional linear distinguishers on these structures [6]. The multidimensional linear approximation¹¹ for FEA-1 is a vector space V of completely linearly dependent input-output masks with $\dim(V) = n/2$ (when n is the block size of Feistel), and the sum of the squared correlations $\sum_{(\alpha, \beta) \in V} \text{Cor}(\alpha, \beta)^2$ is equal to $2^{n(1-r/4)}$. Meanwhile, the approximation for FEA-2 is a vector space V' of linearly independent input-output

¹¹ 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.

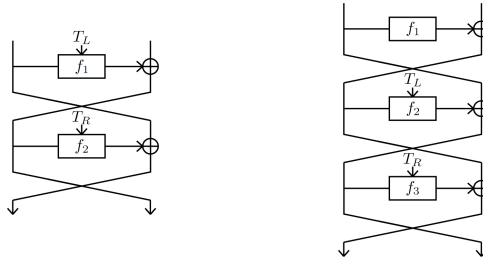


Fig. 1: The FEA-1 structure (left) and FEA-2 structure (right).

masks with $\dim(V) = \dim(V'_2) = n/2$ (here, we assume V' is decomposed as $V' = V'_1 \times V'_2$), and the sum of the squared correlation is equal to $2^{n(1-r/6)}$.

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 $O(2^{(r/12-1/4)n})$.

Remark 4. In [6], linear distinguishers are extended to message recovery attacks and key recovery attacks. Our distinguishers could also be extended to message or key recovery attacks in the quantum setting by just guessing the secret information with the Grover search, though, non-trivial extension of interest (other than just applying Grover) seems to require another new idea and would not be straightforward.

6 Speed-Up for Zero Correlation Linear Distinguishers

This section shows how CEA can be used to speed-up (multidimensional) zero correlation linear distinguishers [9]. We first recall the basic ideas of attacks in the classical setting.

6.1 Classical Zero Correlation Linear Distinguishers

Unlike linear cryptanalysis, zero correlation linear cryptanalysis exploits linear approximations of which the correlation is exactly zero.

For instance, let E_K be an n -bit block cipher and suppose $\text{Cor}(E_K; \alpha, \beta) = 0$ holds for some input and output masks $\alpha, \beta \neq 0^n$. Then, $\text{Cor}(P; \alpha, \beta)$ for a random permutation P is distributed around 2^{-n} and non-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 3.2). Again, let E_K be an n -bit block cipher and $V \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$ be a vector space

such that $\text{Cor}(E_K; \alpha, \beta) = 0$ 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 $p_S^{E_K}(z) := \Pr_x [\text{Lin}_S^f(x) = z]$ 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 setting¹².

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 $O(2^n/\sqrt{2^{\dim(V)}})$ by using the link between zero correlation linear cryptanalysis and integral cryptanalysis, which we will elaborate in Section 7.

6.2 Quantum Speed-Up by CEA

Next, we study how to speed-up (multidimensional) zero correlation linear distinguishers by using CEA and QAA.

As well as linear distinguishers in Section 3.2, we introduce three distinguishers which we denote by \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 when the input and output masks are linearly independent (resp., completely dependent).

In what follows, we assume that $\text{Cor}(E_K; \alpha, \beta)^2 = 0$ holds for all $(\alpha, \beta) \in V - \{\mathbf{0}\}$. \mathcal{O} denotes the oracle, which is either E_K for a random K or a random permutation P . For notations related to \mathcal{B}_2 and \mathcal{B}_3 , we use the same ones as those for \mathcal{A}_2 and \mathcal{A}_3 introduced on p.16.

Distinguisher \mathcal{B}_1 (General Case). When nothing can be assumed on linear dependence of masks, a natural way to mount a distinguisher by using QAA and CEA is to run the following procedure.

1. Let $F : \mathbb{F}_2^n \times \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ be the Boolean function such that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$.
2. Apply QAA on $\text{CEA}^{\mathcal{O}}$ and F with the number of iterations being $\lfloor \frac{\pi}{4} \sqrt{2^{2n - \dim(V)}} \rfloor$.
Namely, let the unitary operator $Q(\text{CEA}^{\mathcal{O}}, F)^i \text{CEA}^{\mathcal{O}}$ 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.

Some analysis shows that this algorithm distinguishes E_K and P with high probability. 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

¹² 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.

does not obtain any speed-up from classical attacks. Meanwhile, we can obtain some quantum speed-up when input-output masks are linearly independent or linearly completely dependent, which we explain below.

Remark 6. \mathcal{B}_1 runs QAA only once, unlike \mathcal{A}_0 of Proposition 2 (or its applications \mathcal{A}_1 - \mathcal{A}_3) running QAA multiple times. This is because the probability $\Pr[F(\alpha, \beta) = 1]$ is exactly zero when (α, β) is obtained by measuring the state $\text{CEA}^{E_K} |0^n\rangle |0^n\rangle$, and thus we can achieve a sufficiently high advantage with a single run of QAA.

Distinguishers \mathcal{B}_2 and \mathcal{B}_3 . Here we show distinguishers \mathcal{B}_2 and \mathcal{B}_3 for linearly independent and completely dependent masks, respectively¹³.

\mathcal{B}_2 is obtained by modifying the unitary operators 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 CEA^O to $\text{CEA}^{\text{msb}_w[O]}$, and
2. the number of iterations from $\lfloor \frac{\pi}{4} \sqrt{2^{2n-\dim(V)}} \rfloor$ to $\lfloor \frac{\pi}{4} \sqrt{2^{n+w-\dim(V)}} \rfloor = \lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \rfloor$.

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

\mathcal{B}_2 distinguishes E_K and P with high probability, roughly for the following reason: If the oracle given to \mathcal{B}_2 is E_K , \mathcal{B}_2 always returns 1. If the oracle given to \mathcal{B}_2 is a random permutation P , Claim 1 guarantees¹⁴ that the probability that we observe $(\alpha, \beta) \in F^{-1}(1)$ when measuring $\text{CEA}^{\text{msb}_w[P]} |0^n\rangle |0^w\rangle$ 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). Similarly, \mathcal{B}_3 distinguishes E_K in time $O(\sqrt{2^n})$. More precisely, the following proposition holds.

Proposition 5. *If \mathcal{B}_2 (resp., \mathcal{B}_3) runs relative to E_K , then the probability that \mathcal{B}_2 (resp., \mathcal{B}_3) always outputs 1. If \mathcal{B}_2 (resp., \mathcal{B}_3) runs relative to a random permutation P , then the probability that \mathcal{B}_2 (resp., \mathcal{B}_3) outputs 1 is approximately upper bounded by $\frac{1}{2} \cdot (1 - 2^{-\dim(V)+1})$. In addition, the running time of \mathcal{B}_2 (resp., \mathcal{B}_3) is at most $2 \lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \rfloor + 1$ (resp., $2 \lfloor \frac{\pi}{4} \sqrt{2^n} \rfloor + 1$) encryptions by E_K . (The probabilities are taken not only over the randomness of \mathcal{B}_2 or \mathcal{B}_3 but also over the randomness of choices of K or P .)*

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

6.3 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

¹³ Recall that we use the same notations as those for \mathcal{A}_2 and \mathcal{A}_3 . See p.16 for details.

¹⁴ Note that the value $2^n \cdot \sum_{(\alpha, \beta) \in V - \{0\}} \text{Cor}(P; \alpha, \beta)^2$ in Claim 1 is equal to $2^n \cdot \text{Cap}(p_S^P)$.

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 outputs 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 general¹⁵ 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 $O(2^{n/2})$.

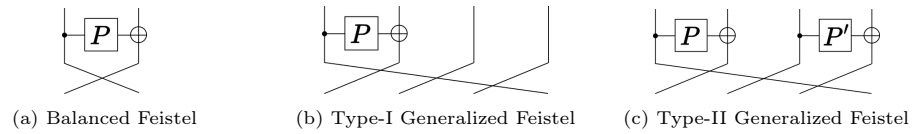


Fig. 2: One round of Balanced and (4-branch) generalized Feistel structures. What we assume is only that P and P' are bijections. Our attacks work regardless of whether P (and P') for different rounds are independent or not.

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

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

Note that the numbers of rounds we attack here are larger than those broken by previous polynomial time attacks using Simon’s algorithm in a black-box way. The number of rounds of balanced (resp., k -branch Type-I and Type-II) Feistel broken by previous polynomial time attacks is 4 [39] (resp., $k^2 - k + 1$ [52] and $k + 1$ [27]). See also Table 2 in the appendix.

In fact, the complexity of our 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, 59, 1] and our \mathcal{B}_2 or \mathcal{B}_3 can be applied to all of them in principle. The amount of quantum speed-up compared to classical distinguish-

¹⁵ k must be even for Type-II structures.

ers depends on linear approximations, and we can achieve at most quadratic speed-up.

7 Speed-Up for Integral Distinguishers

This section shows applications of CEA to integral cryptanalysis. As shown by Bogdanov et al. [8] and Sun et al. [59], balanced property of a cipher is equivalent to multidimensional zero correlation linear properties of which the input-output masks are linearly independent. Specifically, the following proposition holds¹⁶.

Proposition 6 ([8, 59]). *Let $F : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ 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., $\text{Cor}(F; \alpha, \beta) = 0$ 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 7. 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 only if a balanced property exists.

Recall that the distinguisher \mathcal{B}_2 (Proposition 5) is applicable when a multidimensional zero correlation linear approximation exists and the input-output masks are linearly independent. Together with Proposition 6, this implies the following proposition.

Proposition 7. *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 $V' := \{x \mid |0^{n-w}|x \in \mathbb{F}_2^w\}$.) 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 $\frac{1}{2} (1 - 2^{-\dim(V)+1})$.*

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 impossible for integral distinguishers. However, we see possibility of a more than quadratic speed-up in some situations.

¹⁶ This equivalence was first shown by [8] and later refined by [59]. [59] proves the equivalence only in the special case $\dim(V_2) = 1$ but it immediately implies the equivalence for $\dim(V_2) > 1$.

7.1 Possibility of More than Quadratic Speed-Up

Roughly speaking, if a part of the outputs of a cipher (e.g., a specific byte of ciphertexts) is balanced on *multiple mutually orthogonal vector spaces* included in the input space, there exists possibility to achieve a more than quadratic quantum speed-up by using CEA.

Specifically, let $E_K : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be a block cipher, and suppose there exist sub-vector spaces $V_1, \dots, V_s \subset \mathbb{F}_2^n$ 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 the 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 6 we have $\text{Cor}(\text{msb}_w[E_K]; \alpha, \beta) = 0$ if $\alpha \in (V_1)^\perp \cup \dots \cup (V_s)^\perp - \{\mathbf{0}\}$ and $(\alpha, \beta) \neq (0, 0)$. This means

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

Meanwhile, for a random permutation P we can show

$$\Pr_P \left[(\alpha, \beta) \leftarrow \text{CEA}^{\text{msb}_w[P]} : \alpha \perp V_i \text{ for some } i \text{ and } \alpha, \beta \neq 0 \right] \approx \frac{s}{2^d}. \quad (7)$$

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

This can lead to a more than quadratic speed-up compared to the corresponding classical integral distinguisher (when $s \geq 4$) because the classical complexity is 2^d : Even if we have such multiple spaces V_1, \dots, V_s , what we can do in the classical setting is just to choose a single space V_i and check whether (a part of) E_K is balanced on that space, unless some additional properties can be assumed.

Application Examples. To see how the above distinguisher can be applied to concrete ciphers, let us recall the 3.5-round integral property of AES for an example [25]. If a tuple of certain four cells of inputs take all values while others being constant, we can make a single column after the first round take all values while others remain constant, and each cell after 3.5 rounds balanced (see Fig. 3). Since there are four choices on which tuple of four cells to activate (i.e., which column after the first round to activate), we are in the situation of the distinguisher explained above with $d = 32$ and $s = 4$ (V_i corresponds to a tuple of four active cells of inputs).

In fact this example itself is not so significant because more efficient distinguishers exist: If a tuple of four cells of inputs take all values like Fig. 3, actually a certain tuple of four cells of outputs take all values. This means that the integral property specifies a 32-bit permutation between part of inputs and outputs.

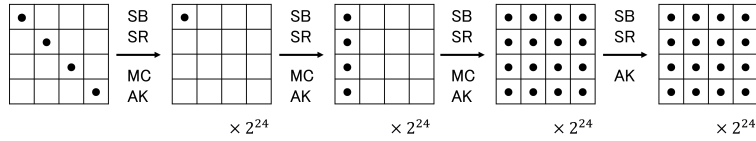


Fig. 3: The integral property of the 2.5-round AES. Cells with filled circles are those taking all values and others are constants. “ $\times 2^{24}$ ” indicates that 2^{24} sets of the same active cell pattern shown in the figure is observed.

Hence the 3.5-round AES is distinguished by checking if this part contains a collision in time $\approx \sqrt[3]{2^{32}}$ with the BHT algorithm [19].

Still, we observe an interesting attack when s is relatively large. For instance, suppose $s = 2^d$ holds. This situation happens if, e.g., E_K is a 4-bit cell SPN cipher with the same integral property as the 2.5-round AES (the latter 2.5 round of Fig. 3). Then $\Pr[F(\alpha, \beta) = 1] \approx s/2^d = 1$ holds when (α, β) is obtained by measuring $\text{CEA}^{\text{msb}_w[P]}$, while the probability is always zero for the real cipher E_K . Thus we can distinguish E_K only with a *single* quantum query, which apparently exhibits a more than quadratic speed-up.

The margin compared to the square-root of the classical complexity is not large, but this example is important in that a new-type example illustrating a qualitative difference between classical and quantum computation is achieved by using a classical cryptanalytic technique.

8 Extension to Generalized Linear Distinguishers

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 $\mathbb{Z}/2^n\mathbb{Z}$), where 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) correlations and quantum computation via Fourier transform. This section shows how the technique of Section 5 extends to generalized linear distinguishers. In this section, the symbol “ \oplus ” denotes the direct sum of groups.

8.1 Fourier Transform on Arbitrary Finite abelian Group

Let G be an arbitrary finite abelian group. Then, w.l.o.g. we can assume $G = \mathbb{Z}/N_1\mathbb{Z} \oplus \dots \oplus \mathbb{Z}/N_m\mathbb{Z}$ for some positive integers N_1, \dots, N_m . Recall that a character of a finite abelian group G is a group homomorphism $\phi : G \rightarrow \mathbb{C}^\times$. 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 $w = (w_1, \dots, w_m) \in G$, the function

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

is a character of G . In fact the map $w \mapsto \text{ch}_w$ 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 $F : G \rightarrow \mathbb{C}$ be a function. Then, the Fourier transform of F over G is a function $\mathcal{F}_G F : G \rightarrow \mathbb{C}$ defined by

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

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

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 $\mathbb{C}^{|G|}$). Moreover, 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.

8.2 Linear Correlations

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

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

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

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

$$((\mathcal{F}_G^* \otimes \mathcal{F}_H) f_{\text{emb}})(\alpha, \beta) = \sqrt{|G|/|H|} \cdot \text{Cor}(f; \alpha, \beta) \quad (8)$$

holds. (This equation corresponds to Eq. (3) for the usual linear cryptanalysis over $(\mathbb{Z}/2\mathbb{Z})^{\oplus n}$.)

8.3 Extension of CEA

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 $\psi : G \rightarrow \mathbb{C}$ 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

$$\text{QFT}_G |\psi\rangle = \sum_x (\mathcal{F}_G^* \psi)(x) |x\rangle. \quad (9)$$

With these notations, the extension of CEA on a function $f : G \rightarrow H$ (G and H are finite abelian groups) is obtained by replacing the Hadamard transform in CEA with the QFT (or its inverse) over G and H . Specifically, the extended algorithm runs as follows.

Extended Version of CEA.

- (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 $\text{QFT}_G \otimes \text{QFT}_H^*$ 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 CEA^f to denote the extended algorithm.

The following proposition is an extension of Proposition 3.

Proposition 8. *The quantum state of CEA^f before the final measurement is*

$$\sum_{\alpha \in G, \beta \in H} \frac{\text{Cor}(f; \alpha, \beta)}{\sqrt{|H|}} |\alpha\rangle |\beta\rangle. \quad (10)$$

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

$$\Pr \left[(\alpha, \beta) \leftarrow \text{CEA}^f : (\alpha, \beta) \in S \right] = \sum_{(\alpha, \beta) \in S} \frac{\text{Cor}(f; \alpha, \beta)^2}{|H|} \quad (11)$$

holds.

Proof. The quantum state of CEA^f before the final measurement is

$$\begin{aligned} & (\text{QFT}_G \otimes \text{QFT}_H^*) U_f (\text{QFT}_G \otimes I_n) |0_G\rangle |0_H\rangle \\ &= (\text{QFT}_G \otimes \text{QFT}_H^*) U_f \sum_{x \in G} \frac{1}{\sqrt{|G|}} |x\rangle |0_H\rangle \\ &= (\text{QFT}_G \otimes \text{QFT}_H^*) \sum_{x \in G} \frac{1}{\sqrt{|G|}} |x\rangle |f(x)\rangle \\ &\stackrel{\text{Def. of } f_{\text{emb}}}{=} (\text{QFT}_G \otimes \text{QFT}_H^*) \sum_{x \in G, y \in H} \frac{f_{\text{emb}}(x, y)}{\sqrt{2^m}} |x\rangle |y\rangle \\ &\stackrel{\text{Def. of QFT}}{=} \sum_{\alpha \in G, \beta \in H} \frac{((\mathcal{F}_G^* \otimes \mathcal{F}_H) f_{\text{emb}})(\alpha, \beta)}{\sqrt{|G|}} |\alpha\rangle |\beta\rangle \\ &\stackrel{\text{Eq. (8)}}{=} \sum_{\alpha \in G, \beta \in H} \frac{\text{Cor}(f; \alpha, \beta)}{\sqrt{|H|}} |\alpha\rangle |\beta\rangle. \end{aligned}$$

Hence we have Eq. (10). Eq. (11) immediately follows from Eq. (10). \square

8.4 Quantum Speed-up for Generalized Linear Distinguishers

Let $f : G \rightarrow H$ be a function, where G and H are finite abelian groups. Here we define linearly independent masks and linearly completely dependent masks.

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 $\phi : G_1 \rightarrow H_1$. 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 5, respectively. We provide only rough ideas and heuristic estimations, and omit detailed analysis.

Distinguisher for Linearly Independent Input-Output Masks. Suppose $f_K : G \rightarrow H$ is a keyed function, G and H are decomposed as $G = G_1 \oplus G_2$, $H = H_1 \oplus H_2$, and $\sum_{\alpha \in G_1, \beta \in H_1} \text{Cor}(f_K; \alpha, \beta)^2 |\alpha\rangle |\beta\rangle \gg \frac{1}{|G_2|}$ holds. Let $f_K^{(1)} : G \rightarrow H_1$ be the projection of f_K onto H_1 . In addition, let $F : G_1 \times H_1 \rightarrow \{0, 1\}$ be the binary function such that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in G_1 \oplus H_1$. Then, from Proposition 8,

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

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

$$\begin{aligned} p_{\text{ideal}} &:= \Pr \left[(\alpha, \beta) \leftarrow \text{CEA}^{\text{RF}^{(1)}} : F(\alpha, \beta) = 1 \right] = \sum_{(\alpha, \beta) \in G_1 \oplus H_1} \frac{\text{Cor}(\text{RF}; \alpha, \beta)^2}{|H_1|} \\ &\approx \sum_{(\alpha, \beta) \in G_1 \oplus H_1} \frac{1}{|G|} \cdot \frac{1}{|H_1|} = \frac{1}{|G_2|}. \end{aligned}$$

(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 $\text{CEA}^{f_K^{(1)}}$ (or $\text{CEA}^{\text{RF}^{(1)}}$) and F with $O(\sqrt{1/p_{\text{real}}})$ iterations.

Distinguisher for Linearly Completely Dependent Input-Output Masks.

Again, let $f_K : G \rightarrow H$ 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 $\phi : G_1 \rightarrow H_1$ and $\sum_{\alpha \in G_1} \text{Cor}(f_K; \alpha, \phi(\alpha))^2 |\alpha\rangle |\phi(\alpha)\rangle \gg \frac{1}{|G|}$ holds. Let $F : G_1 \times H_1 \rightarrow \{0, 1\}$ be the binary function such that $F(\alpha, \beta) = 1$ iff $\alpha \in G_1$ and $\beta = \phi(\alpha)$. Then, from Proposition 8,

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

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

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

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

Since $p_{\text{real}} \gg p_{\text{ideal}}$ holds by assumption, we can distinguish f_K from RF by applying the QAA on $\text{CEA}^{f_K^{(1)}}$ (or $\text{CEA}^{\text{RF}^{(1)}}$) and F with $O(\sqrt{1/p_{\text{real}}})$ iterations.

Application to the FF3-1 Structure. Beyne [6] 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), but 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 $F_K : \mathbb{Z}/2^{n/2}\mathbb{Z} \oplus \mathbb{Z}_2^t \rightarrow \mathbb{Z}/2^{n/2}\mathbb{Z}$ 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 $\{((\alpha, 0), \alpha) \in (\mathbb{Z}/2^{n/2}\mathbb{Z} \oplus \mathbb{Z}_2^t) \oplus \mathbb{Z}/2^{n/2}\mathbb{Z} \mid \alpha \in \mathbb{Z}/2^{n/2}\mathbb{Z}\}$. In particular, the input-output masks are linearly completely dependent. The corresponding sum of the squared correlation is estimated as $\sum_{\alpha \in \mathbb{Z}/2^{n/2}\mathbb{Z}} \text{Cor}(F_K; (\alpha, 0), \alpha) \approx 2^{-n(r/4-1)}$, 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 $O(2^{(r/8-1/4)n})$.

9 Concluding Remarks

This paper showed a quantum speed-up for the multidimensional (zero correlation) linear distinguishers for the first time in such a way to exploit multidimensional approximation in a non-trivial way. Firstly, we observed that there is a close relationship between the subroutine of Simon’s algorithm and linear correlations of functions via Fourier transform. Specifically, a slightly modified version of the subroutine, which we call CEA, returns input and output linear masks of a target function with probability proportional to the squared linear correlation. Combining CEA with QAA, we achieved a quantum speed-up for multidimensional linear distinguishers. It is interesting that, only with a slight modification made, the subroutine of Simon’s algorithm can be used to speed-up such a statistical attack. Our technique is naturally extended to generalized linear distinguishers on arbitrary finite abelian groups by replacing the Hadamard transform in CEA with general QFT. We also showed that CEA similarly speeds-up multidimensional zero correlation linear distinguishers, as well as some integral distinguishers via the correspondence shown by Bogdanov et al. and Sun et

al [8, 59]. Especially, we observe that a more than quadratic speed-up is possible if an integral property holds on multiple mutually orthogonal vector spaces of the same dimension, and even a single-query distinguisher for a toy example of a 4-bit cell SPN cipher with the same integral property as the 2.5-round AES.

Future Directions. An important future work is to investigate how to extend our technique to key-recovery attacks, or combine it with Schrottenloher’s [56].

All the distinguishers in this paper can be extended to 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 the cipher running in time T . In addition, assume that we can apply the distinguisher on the 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 $O(T \cdot 2^{k/2})$.

Still this idea is too naive, compared to classical key-recovery attacks using sophisticated techniques such as the FFT [23, 60, 29]. As mentioned in Section 1.2, the recent work by Schrottenloher [56] has shown how to combine such key-recovery techniques using the FFT with the QFT, taking multiple linear approximations into account. However, in Schrottenloher’s attack, multiple approximations contribute to only the precision of the attack by a constant factor, and does not contribute much to reducing time complexity. It is definitely an important and interesting future work to investigate theoretical relationships between our technique with Schrottenloher’s and study how to reduce the time complexity of key-recovery exploiting multidimensional approximations.

Another important future work is to study quantum speed-up for integral distinguishers based on zero-sum properties. 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, distinguishers based on zero-sum properties often break more rounds than those on balanced functions, especially when extended to key-recovery attacks. Since multiple integral properties sometimes could lead to a more than quadratic speed-up, a quantum attack breaking more rounds of a cipher than classical attacks may be found by investigating this direction.

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 tweak 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.: Hidden structures and quantum cryptanalysis. PhD thesis (2019)
12. 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)
13. 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)
14. 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)
15. Bonnetain, X., Naya-Plasencia, M., Schrottenloher, A.: Quantum security analysis of AES. *IACR Trans. Symmetric Cryptol.* **2019**(2), 55–93 (2019)
16. Bonnetain, X., Schrottenloher, A., Sibleyras, F.: Beyond quadratic speedups in quantum attacks on symmetric schemes. In: Dunkelman, O., Dziembowski, S. (eds.) *EUROCRYPT 2022, Proceedings, Part III*. LNCS, vol. 13277, pp. 315–344. Springer (2022)
17. 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)
18. Brassard, G., Hoyer, P., Mosca, M., Tapp, A.: Quantum amplitude amplification and estimation. *Contemporary Mathematics* **305**, 53–74 (2002)
19. 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)
20. Canale, F., Leander, G., Stennes, L.: Simon’s algorithm and symmetric crypto: Generalizations and automatized applications (2022)

21. Chartoumy, M., Patarin, J., Toulemonde, A.: Quantum cryptanalysis of rounds feistel schemes and benes schemes. IACR Cryptology ePrint Archive 2022/1015 (2022)
22. 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)
23. 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)
24. 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)
25. Daemen, J., Rijmen, V.: Aes proposal: Rijndael (1999)
26. Daemen, J., Rijmen, V.: Probability distributions of correlation and differentials in block ciphers. *J. Mathematical Cryptology* **1**(3), 221–242 (2007)
27. Dong, X., Li, Z., Wang, X.: Quantum cryptanalysis on some generalized feistel schemes. *Sci. China Inf. Sci.* **62**(2), 22501:1–22501:12 (2019)
28. 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)
29. 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)
30. Grover, L.K.: A Fast Quantum Mechanical Algorithm for Database Search. In: ACM STOC 1996. pp. 212–219. ACM (1996)
31. 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
32. 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)
33. 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)
34. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional linear cryptanalysis. *J. Cryptol.* **32**(1), 1–34 (2019)
35. Hermelin, M., Nyberg, K.: Multidimensional linear distinguishing attacks and boolean functions. In: Fourth International Workshop on Boolean Functions: Cryptography and Applications (2008)
36. 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)
37. 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)
38. 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)
39. 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)

40. 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)
41. 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)
42. Kaplan, M., Leurent, G., Leverrier, A., Naya-Plasencia, M.: Quantum differential and linear cryptanalysis. IACR Trans. Symmetric Cryptol. **2016**(1), 71–94 (2016)
43. Knudsen, L.R.: The security of feistel ciphers with six rounds or less. J. Cryptol. **15**(3), 207–222 (2002)
44. 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)
45. Kuwakado, H., Morii, M.: Quantum distinguisher between the 3-round Feistel cipher and the random permutation. In: ISIT 2010. pp. 2682–2685. IEEE (2010)
46. Kuwakado, H., Morii, M.: Security on the quantum-type Even-Mansour cipher. In: ISITA 2012. pp. 312–316. IEEE (2012)
47. Leander, G., May, A.: Grover Meets Simon - Quantumly Attacking the FX-construction. In: ASIACRYPT 2017. LNCS, vol. 10625, pp. 161–178. Springer (2017)
48. 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)
49. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseht, T. (ed.) EUROCRYPT 1993, Proceedings. LNCS, vol. 765, pp. 386–397. Springer (1993)
50. Murphy, S.: The independence of linear approximations in symmetric cryptanalysis. IEEE Trans. Inf. Theory **52**(12), 5510–5518 (2006)
51. Nachev, V., Patarin, J., Volte, E.: Generic Attacks on Generalized Feistel Ciphers, pp. 139–153. Springer International Publishing, Cham (2017)
52. 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)
53. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press (2010)
54. 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)
55. 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)
56. Schrottenloher, A.: Quantum linear key-recovery attacks using the QFT. IACR Cryptol. ePrint Arch. p. 184 (2023)
57. Shi, R., Xie, H., Feng, H., Yuan, F., Liu, B.: Quantum zero correlation linear cryptanalysis. Quantum Inf. Process. **21**(8), 293 (2022)
58. Simon, D.R.: On the power of quantum computation. SIAM J. Comput. **26**(5), 1474–1483 (1997)
59. 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)

60. 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)
61. 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)
62. Xie, H., Yang, L.: Using bernstein-vazirani algorithm to attack block ciphers. Des. Codes Cryptogr. **87**(5), 1161–1182 (2019)

Structure	Rounds	Round Functions	Attack Type	Complexity	Reference
Balanced Feistel	4		CPA	$O(2^{n/2})$	[54]
Balanced Feistel	4		QCCA	$O(\text{poly}(n))$	[39]
Balanced Feistel	5		CPA	$O(2^n)$	[55]
Balanced Feistel	5	bij.	KPA	$O(2^{2n/3})$	[43, 61]
Balanced Feistel	5		QCPA	$O(2^{2n/3})$	[21]
Balanced Feistel	5	bij.	QCPA	$O(2^{n/2})$	Ours
k -branch Type-I Generalized Feistel	$k^2 - k + 1$	bij.	QCCA	$O(\text{poly}(n))$	[52]
k -branch Type-I Generalized Feistel	$k^2 + k - 1$		CPA	$O(2^{(1-\frac{1}{k})n})$	[51]
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	$O(\text{poly}(n))$	[27]
k -branch Type-II Generalized Feistel	$2k + 1$		CPA	$O(2^{(1-\frac{1}{k})n})$	[51]
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., quantum superposition chosen ciphertext attack). All of our attacks appearing in this table are multidimensional zero correlation linear distinguishers.

A Pearson's Chi-Squared Goodness-of-Fit Test and Distinguishers

Given a list L of elements of a finite set S , the test statistic of Pearson's chi-squared goodness-of-fit test (on the uniform distribution over S) is defined as $\chi^2 := |S| \cdot |L| \cdot \sum_{z \in S} \left(\frac{\text{num}(z)}{|L|} - \frac{1}{|S|} \right)^2$, where $\text{num}(z)$ denotes the number of times that z appears in L .

Suppose that entries of L are sampled independently from the uniform distribution or another distribution D , and let χ_U^2 (resp., χ_D^2) be the value of the test statistic when the distribution is uniform (resp., D). Then χ_U^2 approximately follows the chi-squared distribution with $|L| - 1$ degrees of freedom, of which the mean is $\mu_U := |L| - 1$ and the standard deviation is $\sigma_U := \sqrt{2(|L| - 1)}$. If $|L|$ is sufficiently large, we can assume that χ_U^2 approximately follows the normal distribution $\mathcal{N}(|L| - 1, \sqrt{2(|L| - 1)})$, which implies $\Pr[\chi_U^2 \geq \mu_U + \sigma_U] \gtrsim 1/6$. Thus, if we know that the value of χ_D^2 is larger than $\mu_U + \sigma_U$ quite often, then we can distinguish which distribution the entries of L follow with a sufficiently large advantage by checking whether the test statistic is larger than $\mu_U + \sigma_U$ or not.

B 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 $p_U := \Pr \left[x \xleftarrow{\text{measure}} U |0^n\rangle : F(x) = 1 \right]$.

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

1. For $i = 1, \dots, s$, do:
 - (a) Choose i uniformly at random from the set of integers from 0 to $\left\lfloor \frac{1}{\sin(2 \cdot \arcsin(\sqrt{t}))} \right\rfloor$.
 - (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.

Proposition 9 (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 [17].

Lemma 1 (Lemma 2 in [17]). *Let $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ be a Boolean function, U be a unitary operator acting on an n qubit system, and let $p_{\text{init}} := \Pr \left[x \stackrel{\text{measure}}{\leftarrow} U |0^n\rangle : F(x) = 1 \right]$. Assume $0 < p_{\text{init}} < 1/2$, and let N be an integer satisfying $N \geq \frac{1}{\sin(2 \cdot \arcsin(\sqrt{p_{\text{init}}}))}$. Then $\Pr \left[i \stackrel{\$}{\leftarrow} \{0, \dots, N-1\}, x \stackrel{\text{measure}}{\leftarrow} Q(U, F)^i U |0^n\rangle : F(x) = 1 \right] \geq 1/4$ holds.*

In fact Lemma 2 in [17] 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 [18]). We assume $0 < p_{\text{init}} < 1/4$ so that $0 < 2 \cdot \arcsin(\sqrt{p_{\text{init}}}) < \pi/2$ 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

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

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 $Q(U, F)U |0^n\rangle$. 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

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

and thus

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

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

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

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 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. \square

C Proof of Proposition 4

Proof. Here we provide a proof only for \mathcal{A}_2 . The proofs for \mathcal{A}_1 and \mathcal{A}_3 are obtained similarly by changing parameters appropriately. First, note that

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

holds for arbitrary function f by Lemma 3 (in this proof, f is now a truncated version of a block cipher $\text{msb}_w[E_K]$ or a randomly chosen permutation $\text{msb}_w[P]$). 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 $\mu_P := \mathbb{E}_P[p_P] \lesssim 2^{\dim(V) - n - w}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lesssim \sqrt{2^{\dim(V) + 1 - 2n - 2w}}$. Hence we have

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

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 most¹⁷

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

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 $6\sqrt{2^w/c}$. \square

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.

D About Why Claim 1 is Plausible

We argue that Claim 1 is plausible due to the following four facts.

- (i) 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 weakened to “*at most* $2^v - 2^u - 2^w + 1$ degrees of freedom”.
- (ii) The same work conjectures that the above claim holds [2, Conjecture 1], showing some experimental results supporting the conjecture.
- (iii) If a random variable X follows the χ^2 distribution with $2^v - 2^u - 2^w + 1$ degrees of freedom, then $\mathbb{E}[X] = 2^v - 2^u - 2^w + 1$.
- (iv) We can formally prove that $\mathbb{E}_P[2^n \cdot \text{Cap}(p_S^P)]$ is equal to $\frac{2^n}{2^n - 1}(2^v - 2^u - 2^w + 1)$, which is quite close to $2^v - 2^u - 2^w + 1$. Namely, the following proposition holds.

Proposition 10. *For a randomly chosen permutation P , $\mathbb{E}[2^n \cdot \text{Cap}(p_S^P)] = \frac{2^n}{2^n - 1}(2^v - 2^u - 2^w + 1)$ holds.*

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

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

Proposition 11. *Let $f : \{0, 1\}^m \rightarrow \{0, 1\}^n$ 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 $\beta \in \{0, 1\}^n$,*

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

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

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

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

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 $\text{Perm}(m)$ be the set of all permutations over $\{0, 1\}^m$ and $\text{Reg}(m, n)$ 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

$$\begin{aligned}
 \text{Cor}(f; \alpha, \beta)^2 &= \left(\Pr_x[\alpha \cdot x = \beta \cdot f(x)] - \Pr_x[\alpha \cdot x \neq \beta \cdot f(x)] \right)^2 \\
 &= \left(2 \Pr_x[\alpha \cdot x = \beta \cdot f(x)] - 1 \right)^2 \\
 &= \left(\frac{2 \cdot \#\{x \in \{0, 1\}^m \mid \alpha \cdot x = \beta \cdot f(x)\} - 2^m}{2^m} \right)^2 \\
 &= \frac{1}{2^{2m}} \left(2 \sum_{x \in \{0, 1\}^m} \chi_{\alpha, \beta}(f, x) - 2^m \right)^2 \\
 &= \frac{1}{2^{2m}} \left(4 \sum_{x, x' \in \{0, 1\}^m} \chi_{\alpha, \beta}(f, x) \chi_{\alpha, \beta}(f, x') - 2^{m+2} \sum_{x \in \{0, 1\}^m} \chi_{\alpha, \beta}(f, x) + 2^{2m} \right) \\
 &= \frac{1}{2^{2m}} \left(4 \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x'}} \chi_{\alpha, \beta}(f, x) \chi_{\alpha, \beta}(f, x') - (2^{m+2} - 4) \sum_{x \in \{0, 1\}^m} \chi_{\alpha, \beta}(f, x) + 2^{2m} \right)
 \end{aligned} \tag{12}$$

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

$$\mathbf{E}_f [\chi_{\alpha, \beta}(f, x)] = \Pr_f [\chi_{\alpha, \beta}(f, x) = 1] = \frac{1}{2} \tag{13}$$

because, for each fixed tuple (α, β, x) with $\beta \neq 0^n$, the number of $f \in \text{Reg}(m, n)$ 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 $P \in \text{Perm}(m)$, 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

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

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

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

Therefore

$$\begin{aligned} & \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x'}} \mathbf{E}_f [\chi_{\alpha, \beta}(f, x) \chi_{\alpha, \beta}(f, x')] \\ &= \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x' \wedge \alpha \cdot x = \alpha \cdot x'}} \mathbf{E}_f [\chi_{\alpha, \beta}(f, x) \chi_{\alpha, \beta}(f, x')] + \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x' \wedge \alpha \cdot x \neq \alpha \cdot x'}} \mathbf{E}_f [\chi_{\alpha, \beta}(f, x) \chi_{\alpha, \beta}(f, x')] \\ &= \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x' \wedge \alpha \cdot x = \alpha \cdot x'}} \frac{2^{m-1} - 1}{2^m - 1} \cdot \frac{1}{2} + \sum_{\substack{x, x' \in \{0, 1\}^m \\ x \neq x' \wedge \alpha \cdot x \neq \alpha \cdot x'}} \frac{2^{m-1}}{2^m - 1} \cdot \frac{1}{2} \\ &= \frac{2^m (2^{m-1} - 1)^2}{2 \cdot (2^m - 1)} + \frac{2^{3m-2}}{2 \cdot (2^m - 1)} \\ &= \frac{2^{3m-2} - 2^{2m-1} + 2^{m-1}}{2^m - 1} \tag{14} \end{aligned}$$

holds. From Eq. (12)-(14), it follows that

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

which completes the proof. \square

E Proof of Proposition 3

Proof (of Proposition 3). The quantum state of CEA^f before the final measurement is

$$\begin{aligned} H^{\otimes(m+n)} U_f (H^{\otimes m} \otimes I_n) |0^m\rangle |0^n\rangle &= H^{\otimes(m+n)} U_f \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle |0^n\rangle \\ &= H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m} \frac{1}{\sqrt{2^m}} |x\rangle |f(x)\rangle \\ &\stackrel{\text{Def. of } f_{\text{emb}}}{=} H^{\otimes(m+n)} \sum_{x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n} \frac{f_{\text{emb}}(x, y)}{\sqrt{2^m}} |x\rangle |y\rangle \\ &\stackrel{\text{Eq. (4)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\mathcal{F} f_{\text{emb}}(\alpha, \beta)}{\sqrt{2^m}} |\alpha\rangle |\beta\rangle \\ &\stackrel{\text{Eq. (3)}}{=} \sum_{\alpha \in \mathbb{F}_2^m, \beta \in \mathbb{F}_2^n} \frac{\text{Cor}(f; \alpha, \beta)}{\sqrt{2^n}} |\alpha\rangle |\beta\rangle. \end{aligned}$$

Hence we have Eq. (5). Eq. (6) immediately follows from Eq. (5). \square

F How to Efficiently Compute F in Distinguishers

This section explains how to compute Boolean functions F used in our quantum multidimensional (zero-correlation) linear distinguishers on a quantum circuit F .

The functions F in the distinguishers satisfy that $F(\alpha, \beta) = 1$ iff $(\alpha, \beta) \in V - \{\mathbf{0}\}$ for some vector space $V \subset \mathbb{F}_2^m \times \mathbb{F}_2^n$. We explain how to judge whether a given (α, β) belongs to V in the following two cases.

Case 1: V is the set of n -bit strings of which specific bit positions are zero (e.g., the set of strings of the form $0^{n/2}||x$ for some $x \in \mathbb{F}_2^{n/2}$).

Case 2: V is a general sub-vector-space of \mathbb{F}_2^n .

For Case 1, whether $(\alpha, \beta) \in V$ can be judged just by checking whether some specific bits are zero.

For Case 2, we choose and fix a \mathbb{F}_2 -linear isomorphism (invertible matrix) M such that $M(V)$ becomes a space of Case 1. To check whether $(\alpha, \beta) \in V$, we perform the following procedure.

- (i) compute $v := M(\alpha, \beta)$
- (ii) check whether $v \in M(V)$
- (iii) uncompute (i)

Note that (ii) is exactly the judgement for Case 1 (because $M(V)$ is a vector space of Case 1). (i) and (iii) may require additional computational costs, but they will be less than the cost for a few computations of a linear layer of a typical block cipher. (In any case, we can implement a quantum circuit by using some CNOT/CCNOT gates.)

G Proof of Proposition 5

Proof (of Proposition 5). Here we give a proof only for \mathcal{B}_2 . The proof for \mathcal{B}_3 can be obtained just by replacing the value w appearing in the proof for \mathcal{B}_2 with $\dim(V)$.

It immediately follows from the definition of \mathcal{B}_2 , Proposition 2, and Proposition 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

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

holds.

First, note that

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

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

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 $\mu_P := \mathbb{E}_P[p_P] \lesssim 2^{\dim(V)-n-w} = 2^{u-n}$ and $\sigma_P := \sqrt{\mathbf{Var}_P[p_P]} \lesssim \sqrt{2^{\dim(V)+1-2n-2w}} = 2^{\frac{-\dim(V)+1}{2}}$. Hence we have

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

Especially,

$$\Pr_P \left[\frac{1}{2} 2^{u-n} \leq p_P \leq 2 \cdot 2^{u-n} \right] \geq 1 - 2^{-\dim(V)+1} \quad (15)$$

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

$$\begin{aligned} &\Pr \left[(\alpha, \beta) \xleftarrow{\text{measure}} Q(\text{CEA}^{\text{msb}_w[P]}, F)^{\lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \rfloor} \text{CEA}^{\text{msb}_w[P]} |0^n\rangle |0^w\rangle \right] \\ &= \sin^2 \left(\left(2 \left\lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \right\rfloor + 1 \right) \arcsin(\sqrt{p_P}) \right) \\ &\approx \sin^2 \left(\frac{\pi}{2} \sqrt{2^{n-u}} \cdot \sqrt{p_P} \right) \gtrsim \frac{1}{2}. \end{aligned} \quad (16)$$

Therefore

$$\begin{aligned} &\Pr \left[(\alpha, \beta) \xleftarrow{\text{measure}} Q(\text{CEA}^{\text{msb}_w[P]}, F)^{\lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \rfloor} \text{CEA}^{\text{msb}_w[P]} |0^n\rangle |0^w\rangle \right] \\ &\geq \Pr \left[(\alpha, \beta) \leftarrow Q(\text{CEA}^{\text{msb}_w[P]}, F)^{\lfloor \frac{\pi}{4} \sqrt{2^{n-u}} \rfloor} \text{CEA}^{\text{msb}_w[P]} |0^n\rangle |0^w\rangle \left| \frac{1}{2} 2^{u-n} \leq p_P \leq 2 \cdot 2^{u-n} \right. \right] \\ &\quad \cdot \Pr_P \left[\frac{1}{2} 2^{u-n} \leq p_P \leq 2 \cdot 2^{u-n} \right] \\ &\gtrsim \frac{1}{2} \cdot \left(1 - 2^{-\dim(V)+1} \right) \end{aligned} \quad (17)$$

holds, which completes the proof. \square

H About Eq. (7)

Eq. (7) holds because we have

$$\begin{aligned}
 & \Pr_P \left[(\alpha, \beta) \leftarrow \text{CEA}^{\text{msb}_w[P]} : \alpha \perp V_i \text{ for some } i \text{ and } \alpha \neq 0 \text{ and } \beta \neq 0 \right] \\
 & \stackrel{\text{Prop. 3}}{=} \sum_{\substack{\alpha \neq 0, \beta \neq 0 \\ \alpha \perp V_i \text{ for some } i}} \mathbb{E}_P \left[\frac{\text{Cor}(\text{msb}_w[P]; \alpha, \beta)^2}{2^w} \right] = \sum_{\substack{\alpha \neq 0, \beta \neq 0 \\ \alpha \perp V_i \text{ for some } i \\ \text{lsb}_{n-w}[\beta]=0}} \mathbb{E}_P \left[\frac{\text{Cor}(P; \alpha, \beta)^2}{2^w} \right] \\
 & \stackrel{\text{Prop. 11}}{=} \# \{ \alpha \in \mathbb{F}_2^n - \{0\} \mid \alpha \perp V_i \text{ for some } i \} \cdot \frac{2^w - 1}{2^w(2^n - 1)} \\
 & \geq \left(\sum_{1 \leq i \leq s} |V_i^\perp| - \sum_{1 \leq i < j \leq s} |V_i^\perp \cap V_j^\perp| - 1 \right) \cdot \frac{2^w - 1}{2^w(2^n - 1)} \\
 & \stackrel{V_i \perp V_j \text{ for } i \neq j}{\geq} (s2^{n-d} - s^2 2^{n-2d} - 1) \cdot \frac{2^w - 1}{2^w(2^n - 1)} \approx \frac{s}{2^d}.
 \end{aligned}$$