Breaking Tweakable Enciphering Schemes using Simon's Algorithm

Sebati Ghosh and Palash Sarkar
Indian Statistical Institute
203, B.T.Road, Kolkata, India - 700108.
{sebati_r,palash}@isical.ac.in

May 5, 2021

Abstract

We show the applicability of Simon's period finding quantum algorithm to the cryptanalysis of several tweakable enciphering schemes (TESs), namely, CMC, EME, XCB, TET and FAST. For all of the five TESs, we show distinguishing attacks, while for XCB, TET and FAST, the attacks reveal portions of the secret keys.
Keywords: tweakable enciphering scheme, Simon's algorithm.

1 Introduction

The eventual availability of large-scale quantum computers appears to be a certainty. This will have major impact on cryptography. Public key cryptography based on factoring and the discrete logarithm problem will be completely broken by Shor's algorithm [17].

For symmetric key ciphers, exhaustive key search will be speeded up by a quadratic factor using Grover's algorithm [7]. Symmetric key primitives such as block and stream ciphers are often used in modes of operation to build versatile cryptographic functionalities. A series of works [13, 14, 12, 2, 5] have shown how to apply Simon's period finding quantum algorithm [18] to break the security of certain modes of operation. These attacks require quantum access to the cryptographic algorithm. More recently, there has been work [3] on developing attacks based on offline Simon's algorithm which do not need to make quantum queries.

In the present work, we continue the line of work on using Simon's algorithm to attack modes of operation. Our target modes of operation are tweakable enciphering schemes (TESs) [10]. These provide several important cryptographic functionalities including that of full disk encryption. We refer to [10] for a description of how a TES can be used for disk encryption and to [4] for more general functionalities. Some TESs have also been standardised [1].

We consider five TESs, namely, CMC [10], EME [11, 8], TET [9], XCB [15, 16] and FAST [4]. CMC was the first TES to be proposed; IEEE has standardised [1] XCB and EME; TET uses invertible universal hash; and presently FAST provides the most recent development.

Following Kaplan et al. [12], the attacks that we describe are essentially based on an algorithm to solve the following problem.

Simon's problem: Given a function $f : \{0,1\}^{\mathtt{m}} \rightarrow \{0,1\}^n$ and the promise that there exists $s \in \{0,1\}^{\mathtt{m}} \setminus 0^{\mathtt{m}}$ such that for all $x \neq y$, $f(x) = f(y)$ if and only if $x \oplus y = s$, find s.

The quantity s is called the period of the function. Simon [18] described a quantum algorithm which, with high probability, finds the period of f with $O(\mathfrak{m})$ quantum queries to the function f and additional polynomial time classical computation. The description in [18] required f to be a 2-to-1 function which was later modified to a looser condition in [12].

For each of the TES that we consider, we construct a function f based on the encryption algorithm of the TES. The function f has a period which is based on variables that are used during the computation, but is not revealed as part of the ciphertext. Applying Simon's algorithm to f uncovers the period and reveals the internal secret variable. In the cases of TET, XCB and FAST, obtaining the period of f reveals a portion of the secret key of the TES resulting in a partial key recovery attack. For all the five TESs, we show that using the period, it is possible to construct two distinct plaintexts such that designated portions of the corresponding ciphertexts are equal. This results in distinguishing attacks on all the TESs under consideration.

2 Preliminaries

The concatenation of two strings x_1 and x_2 will be denoted as $x_1||x_2$. Given an integer i in the range $0 \leq i < 2^k$, $\mathsf{bin}_k(i)$ denotes the k-bit binary representation of i.

We fix a positive integer n. A block cipher is a function $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$, where $\mathcal{K}$ is a finite non-empty set and for each $K \in \mathcal{K}$, $E_K(\cdot) \stackrel{\Delta}{=} E(K, \cdot)$ is a permutation of $\{0,1\}^n$. The integer n denotes the block size and K is the key of the block cipher. The corresponding decryption function is $D : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$, where for each $K \in \mathcal{K}$, $D_K(\cdot) \stackrel{\Delta}{=} D(K, \cdot)$ is the inverse of $E_K(\cdot)$, i.e., for any $x \in \{0,1\}^n$, $D_K(E_K(x)) = x$.

2.1 Tweakable Enciphering Scheme

A tweakable enciphering scheme is a pair $\mathsf{TES} = (\mathsf{TES.Encrypt}, \mathsf{TES.Decrypt})$ where

$$\mathsf{TES.Encrypt}, \mathsf{TES.Decrypt} : \mathcal{K} \times \mathcal{T} \times \mathcal{P} \to \mathcal{P}$$

for finite non-empty sets $\mathcal{K}, \mathcal{T}$ and $\mathcal{P}$. The set $\mathcal{K}$ is called the key space, $\mathcal{T}$ is called the tweak space and $\mathcal{P}$ is called the message/ciphertext space. We write $\mathsf{TES.Encrypt}_K(\cdot, \cdot)$ (resp. $\mathsf{TES.Decrypt}_K(\cdot, \cdot)$) to denote $\mathsf{TES.Encrypt}(K, \cdot, \cdot)$ (resp. $\mathsf{TES.Decrypt}(K, \cdot, \cdot)$). The functions $\mathsf{TES.Encrypt}$ and $\mathsf{TES.Decrypt}$ satisfy the following two properties. For $K \in \mathcal{K}$, $T \in \mathcal{T}$ and $P \in \mathcal{P}$,

1. $\mathsf{TES.Decrypt}_K(T, \mathsf{TES.Encrypt}_K(T, P)) = P$;
2. $\mathsf{len}\,(\mathsf{TES.Encrypt}_K(T, P)) = \mathsf{len}(P)$.

The first property states that the encryption and the decryption functions are inverses of each other while the second property states that the length of the ciphertext is equal to the length of the plaintext. In other words, $\mathsf{TES.Encrypt}_K(T, \cdot)$ is a length preserving permutation of $\mathcal{P}$.

We do not provide the formal definition of security of TES since this will not be required for our work. The notion of security that we consider is that of indistinguishability from a random oracle which returns independent and uniform random strings of appropriate lengths. This implies other notions of security (see [10]).

Four of the five TESs that we consider, namely, XCB, TET, CMC and EME are built using n-bit block ciphers. The security proofs of the TESs assume the underlying block cipher to be strong pseudo-random permutation (SPRP). The other TES that we consider, namely FAST, is built using

a n-bit to n-bit pseudo-random function (PRF). Assuming the underlying primitive to be a secure SPRP (for XCB, TET, CMC, EME) or a secure PRF (for FAST), the security proofs of all the five TESs provide an upper bound on the advantage of an adversary in distinguishing the TES from a random oracle. The upper bound is essentially of the form $c\sigma_n^2/2^n$, where c is a small constant and σ_n is the number of n-bit blocks in all the queries made by the adversary. Ignoring the constant c, at a broad level the proofs show that the TESs are secure up to about $2^{n/2}$ adversarially chosen n-bit blocks.

2.2 Simon's Algorithm with Spurious Collisions

Simon's problem is a promise problem, i.e., the function f has to satisfy the stated condition for Simon's algorithm to work. There may be functions for which there is an $s \in \{0,1\}^{\mathfrak{m}} \setminus 0^{\mathfrak{m}}$, such that for all $(x,y) \in \{0,1\}^{\mathfrak{m}} \times \{0,1\}^{\mathfrak{m}}, x \oplus y \in \{0^{\mathfrak{m}}, s\} \Rightarrow f(x) = f(y)$, but $f(x) = f(y)$ does not necessarily imply $x \oplus y \in \{s, 0^{\mathfrak{m}}\}$, i.e., there could be a t different from s and $0^{\mathfrak{m}}$, such that for some x, $f(x) = f(x \oplus t)$. Such a collision is called a spurious collision. This issue was considered in [12], which defined the notion of approximate promise problem. For $f : \{0,1\}^{\mathfrak{m}} \rightarrow \{0,1\}^n$ such that $f(x \oplus s) = f(x)$ for all x, the following quantity was defined in [12].

$$\varepsilon(f,s) \quad = \quad \mathsf{max}_{t \in \{0,1\}^{\mathfrak{m}} \setminus \{0,s\}} \Pr_x[f(x) = f(x \oplus t)]. \tag{1}$$

If f satisfies the promise in Simon's problem and has period s, then $\varepsilon(f,s) = 0$. We say that a function $f : \{0,1\}^{\mathfrak{m}} \rightarrow \{0,1\}^n$ satisfies the promise in Simon's problem approximately, if there is an s such that $f(x) = f(x \oplus s)$ for all x and $0 < \varepsilon(f,s) < 1$. A modification of Simon's algorithm to solve the approximate promise problem has been considered in [12] where the following result was proved.

Theorem 1 (Kaplan et al. [12]). *If $\varepsilon(f,s) \leq p_0 < 1$, then for constant c, Simon's algorithm returns s with $c\mathfrak{m}$ quantum queries, with probability at least $1 - (2(\frac{1+p_0}{2})^c)^{\mathfrak{m}}$.*

If f satisfies the approximate promise problem, then Theorem 1 shows that s can be recovered with high probability.

Remarks:

1. A function satisfying the promise in Simon's problem is a 2-to-1 function. Simon had considered a slightly different problem. Given a function $f : \{0,1\}^{\mathfrak{m}} \rightarrow \{0,1\}^n$ which is known to be either injective or 2-to-1, determine the correct condition and if f is 2-to-1, then determine its period.

2. In his formulation, Simon required $n \geq \mathfrak{m}$. The analysis of Simon's algorithm, on the other hand, goes through without the condition $n \geq \mathfrak{m}$ and later works [12, 2, 6] have indeed also considered $n < \mathfrak{m}$.

3 Outline of the Attacks

The attacks that we describe are based on Simon's algorithm and are distinguishing attacks. For three of the TESs, namely XCB, TET and FAST, the attacks also reveal part of the secret key.

Suppose the adversary is provided black-box access to Π, where Π is either the encryption algorithm of a TES (which we write as Π is real) or Π is a random oracle (which we write as Π

is random). The goal of the adversary is to determine whether Π is real or random. Using Π we define a function f. If Π is random, then f is a random function. On the other hand, if Π is real, then either f satisfies the promise in Simon's problem, or the approximate promise mentioned in Section 2.2. Simon's algorithm is applied to f which requires making quantum queries to the given black-box. This is a strong attack model and it has been adopted in previous works [12, 2, 6].

If Π is random, then the output of Simon's algorithm on f will be a random string. On the other hand, if Π is real, then with high probability Simon's algorithm will return the period of f. So, given the output of Simon's algorithm, some further work is required to determine whether Π is real or random. This work consists of making two classical queries to the black box. These queries are built from the output of Simon's algorithm. If Π is real, then we show that the outputs of the two classical queries satisfy a pre-defined relation, while if Π is random, then the outputs of the two classical queries satisfy the same relation with very low probability. So, looking at the outputs of the two classical queries, it becomes possible to determine whether Π is real or random.

The above provides the broad outline of the attacks on the TESs. In the subsequent sections, we do not repeat the above strategy. Instead, we provide the definition of f when Π is real, the two classical queries and the pre-defined relation that their outputs satisfy when Π is real. Plugging these two tools into the above attack strategy provides the complete attacks on the individual TESs.

As mentioned above, for some of the TESs, we show that f satisfies an approximate promise. The proof of approximate promise requires upper bounding the probability of spurious collisions. Since the definition of f is based on the block cipher, to bound the probability of spurious collisions of f, we need to make an assumption on the underlying block cipher. The assumption that we make is to consider the block cipher to behave like a uniform random function. Since a block cipher is an injective map, it would be appropriate to assume the block cipher to behave like a uniform random permutation. If the number of inputs on which the block cipher is invoked is below the (quantum) birthday bound, then it is reasonable to consider the block cipher to behave like a uniform random function. In our applications, we will consider the application of the block cipher to only a few (at most six) inputs.

The analyses of the probabilities of spurious collisions for the various TESs have a common structure. Suppose s is the period of f. We start by considering a non-zero $t \neq s$ which maximises the probability of spurious collisions. The requirement is to bound the probability $f(x) = f(x \oplus t)$. Then for any event $\mathbf{E}$, we have

$$\Pr[f(x) = f(x \oplus t)] \quad \leq \quad \Pr[f(x) = f(x \oplus t) | \mathbf{E}] + \Pr[\overline{\mathbf{E}}]. \tag{2}$$

In the analyses of the individual TESs, we identify a suitable event $\mathbf{E}$ and obtain upper bounds on the two terms in the right hand side of (2).

In Section 4, we describe the attacks on XCB, TET and FAST. The attacks on XCB, TET and FAST require the key of the underlying universal hash function to be non-zero. Since the hash key is a random n-bit quantity, it is zero with probability $1/2^n$ which is negligible for $n = 128$ or larger. These attacks also recover the hash key. In Section 5, we describe the distinguishing attacks on CMC and EME. For EME, we require the internal variable L to be non-zero. Since it is the output of a block cipher instantiated with a random key, it is zero with probability $1/2^n$ which is negligible for $n = 128$ or larger.

Offline Simon's Algorithm

The attacks that we describe require quantum access to the encryption algorithms of the respective TESs. A recent work [3] has shown that for some symmetric key algorithms, it is possible to do away with the requirement of quantum access to the encryption algorithms. The quantum computations are done in an offline manner while all the queries to the encryption algorithms are classical. In particular, Simon's algorithm is applied in an offline mode. Such attacks are more practical than attacks which require quantum access to the encryption algorithms.

The core observation in [3] is that it is possible to determine whether a function of the form $f_1 \oplus f_2$ has a period without any quantum query to f_2 if there is a suitable quantum state corresponding to f_2. Various examples of the idea are provided in [3]. For the TESs that we have considered, we tried to apply the idea from [3] to obtain attacks which do not require quantum access to the encryption algorithms. Our efforts were not successful. It was not clear how to modify the functions with period that we constructed for the various TESs to the form $f_1 \oplus f_2$ which seems to be required to apply the technique of [3]. Our inability does not mean that offline Simon's algorithm is not applicable to these TESs. There could be other ways of constructing the functions in the desired form. This though seems to require more work.

4 Partial Key Recovery Attacks

4.1 XCB

XCB was proposed by McGrew and Fluhrer [15]. A later variant [16] was standardised by IEEE [1]. We describe the quantum attack on the standardised version [16] of XCB. A similar attack also works on the previous version.

XCB is built using a block cipher and a polynomial hash function. The key K of XCB is the same as the key of the underlying block cipher. XCB defines a tweak space. In our attack, we will fix the tweak to be the empty string $\mathbf{e}$.

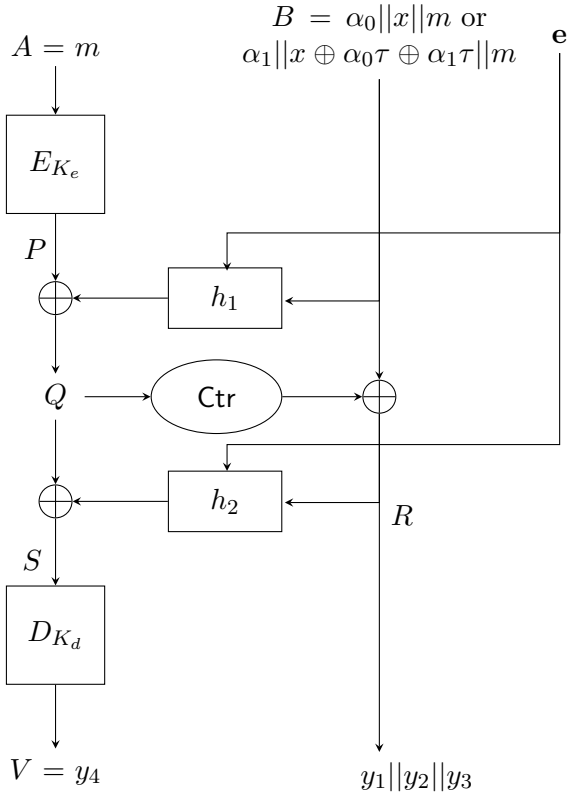
XCB can be used with an n-bit block cipher. For the sake of convenience, we fix $n = 128$. Let E_K denote the encryption function of the underlying block cipher instantiated with the key K. Using E_K, XCB derives the keys K_e, K_d, K_c and τ. Here τ is used as the key to a polynomial hash function called GHASH, K_c is the key to the counter mode of encryption and K_e and K_d are used as shown in Figure 1. The counter mode Ctr uses the function incr to obtain successive values to be encrypted.

Our attack considers 4-block messages. So, we briefly describe the encryption of 4-block messages with reference to Figure 1. The message is partitioned into a single block and a 3-block message. As per the specification of XCB, the quantity A is equal to the last block of the message and the quantity V is the last block of the ciphertext. In more details, if $y_1||y_2||y_3||y_4$ is the 4-block ciphertext corresponding to a 4-block message, then $V = y_4$ and $R = y_1||y_2||y_3$. The functions h_1 and h_2 in Figure 1 are polynomial hash functions using the key τ. The counter mode Ctr uses Q as the initialisation vector. The rest of the encryption algorithm can be understood from Figure 1. We provide more details as part of the attack.

Fix $m, \alpha_0, \alpha_1 \in \{0,1\}^n$, such that $\alpha_0 \neq \alpha_1$; let b denote a bit. For the standardised version [16], we define the following function.

$$\begin{aligned} f : \{0,1\} \times \{0,1\}^n &\rightarrow \{0,1\}^n \\ (b, x) &\stackrel{f}{\mapsto} y_3, \text{ where } y_1||y_2||y_3||y_4 \longleftarrow \mathsf{XCB.Encrypt}_K(\mathbf{e}, \alpha_b||x||m||m). \end{aligned} \tag{3}$$

Figure 1: Enciphering a 4-block message $\alpha_0||x||m||m$ or $\alpha_1||x \oplus \alpha_0\tau \oplus \alpha_1\tau||m||m$ with tweak $\mathbf{e}$ under XCB.



The function f defined in (3) satisfies the following property.

Proposition 1. *Let $b, b' \in \{0,1\}, x, x' \in \{0,1\}^n$. Suppose that the hash key τ is non-zero. Then, $f(b,x) = f(b',x') \Leftrightarrow x \oplus x' = \alpha_b\tau \oplus \alpha_{b'}\tau$, where α_0 and α_1 are as fixed before.*

Proof. Let γ be a 128-bit string which is formed by concatenating the 64-bit binary representation of 128 and the 64-bit binary representation of 512. For the input $\alpha_b||x||m||m$,

$$\begin{aligned}
A &= m;\\
B &= \alpha_b||x||m;\\
P &= E_{K_e}(m);\\
Q &= E_{K_e}(m) \oplus \alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus \gamma\tau;\\
R &= \alpha_b \oplus E_{K_c}(Q)||x \oplus E_{K_c}(\mathsf{incr}(Q))||m \oplus E_{K_c}(\mathsf{incr}(\mathsf{incr}(Q)));
\end{aligned}$$

For the input $\alpha_{b'}||x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau||m||m$,

$$\begin{aligned}
A' &= m;\\
B' &= \alpha_{b'}||x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau||m;\\
P' &= E_{K_e}(m);\\
Q' &= E_{K_e}(m) \oplus \alpha_{b'}\tau^4 \oplus x\tau^3 \oplus \alpha_b\tau^4 \oplus \alpha_{b'}\tau^4 \oplus m\tau^2 \oplus \gamma\tau\\
&= E_{K_e}(m) \oplus x\tau^3 \oplus \alpha_b\tau^4 \oplus m\tau^2 \oplus \gamma\tau;\\
R' &= \alpha_{b'} \oplus E_{K_c}(Q')||x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau \oplus E_{K_c}(\mathsf{incr}(Q'))||m \oplus E_{K_c}(\mathsf{incr}(\mathsf{incr}(Q')));
\end{aligned}$$

We observe, $Q = Q'$ results in equality of last blocks of R and R'. So, the third blocks of the outputs are same, establishing one direction of the result.

For the other direction, we have

$$\begin{aligned}
y_3 = y_3' &\Rightarrow m \oplus E_{K_c}(\mathsf{incr}(\mathsf{incr}(Q))) = m \oplus E_{K_c}(\mathsf{incr}(\mathsf{incr}(Q')))\\
&\Rightarrow Q = Q'\\
&\Rightarrow E_{K_e}(m) \oplus \alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus \gamma\tau = E_{K_e}(m) \oplus \alpha_{b'}\tau^4 \oplus x'\tau^3 \oplus m\tau^2 \oplus \gamma\tau\\
&\Rightarrow \alpha_b\tau^4 \oplus x\tau^3 = \alpha_{b'}\tau^4 \oplus x'\tau^3\\
&\Rightarrow x \oplus x' = \alpha_b\tau \oplus \alpha_{b'}\tau.
\end{aligned}$$

□

Classical queries: Given the period $1||s = 1||\tau(\alpha_0 \oplus \alpha_1)$, the two classical queries required in Section 3 are the following. The first query is $\alpha_0||x||m||m$ with output $y_1||y_2||y_3||y_4$ and the second query is $\alpha_1||x \oplus s||m||m$ with output $y_1'||y_2'||y_3'||y_4'$. From the proof of Proposition 1 we have that $y_3 = y_3'$ which defines the relation between the outputs of the two classical queries.

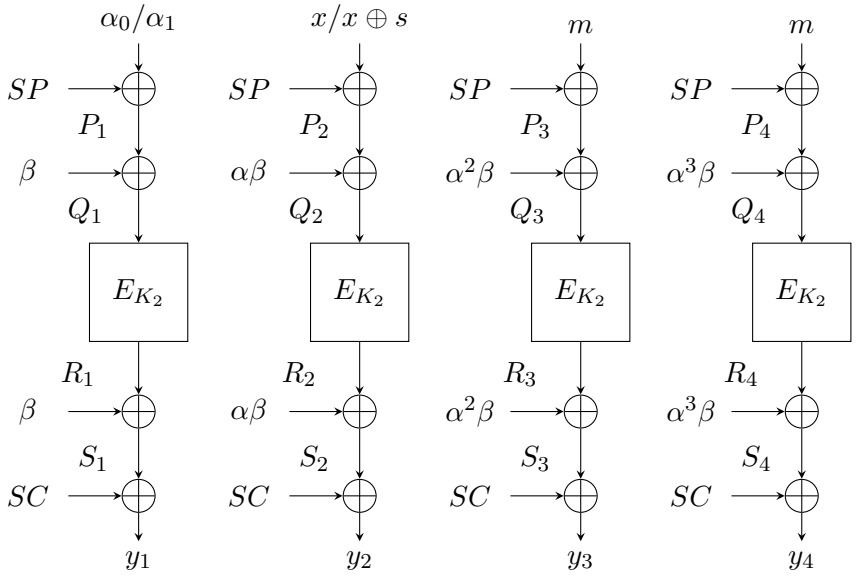
Partial key recovery: Once $s = \tau(\alpha_0 \oplus \alpha_1)$ has been obtained, since α_0 and α_1 are distinct, from s, one obtains the hash key τ as $\tau = s(\alpha_0 \oplus \alpha_1)^{-1}$.

4.2 TET

TET [9] has key space $\mathcal{K} \times \mathcal{K}$, where $\mathcal{K}$ is the key space for the underlying block cipher having block size n. The tweak space is $\mathcal{T} = \{0,1\}^*$. The message space is $\mathcal{P} = \{0,1\}^{\mathfrak{m}}$ where $\mathfrak{m} \in [n, 2^n - 1]$. Fix arbitrary (K_1, K_2) from the key-space and arbitrary T from the tweak-space.

The attack against TET also considers 4-block messages. Hence, we briefly explain the encryption of 4-block messages with reference to Figure 2. E_{K_2} is the encryption function of the underlying block cipher instantiated with the key K_2. The encryption consists of five layers; the first, second, fourth and fifth being masking layers and the third layer being application of E_{K_2}. For a 4-block message $x_1||x_2||x_3||x_4$ and hash key τ, $SP = \sigma^{-1}\left(x_1\tau^4 \oplus x_2\tau^3 \oplus x_3\tau^2 \oplus x_4\tau\right)$, where $\sigma = 1 \oplus \tau \oplus \tau^2 \oplus \tau^3 \oplus \tau^4$ which is assumed to be non-zero. The exact definitions of α, β and SC are not required for our purpose, so, we skip these details and refer to [9] for their definitions. We only note that the tweak T is used in determining β.

Figure 2: Enciphering a 4-block message $\alpha_0||x||m||m$ or $\alpha_1||x \oplus \alpha_0\tau \oplus \alpha_1\tau||m||m$ under TET.



Fix $m, \alpha_0, \alpha_1 \in \{0,1\}^n$, such that $\alpha_0 \neq \alpha_1$; let b denote a bit and we define the following function.

$$\begin{aligned} f : \{0,1\} \times \{0,1\}^n &\rightarrow \{0,1\}^n \\ (b, x) &\overset{f}{\mapsto} y_3 \oplus y_4, \text{ where } y_1||y_2||y_3||y_4 \longleftarrow \mathsf{TET.Encrypt}_{K_1,K_2}(T, \alpha_b||x||m||m). \end{aligned} \tag{4}$$

The function f defined in (4) satisfies the following property.

Proposition 2. *Let $b, b' \in \{0,1\}, x \in \{0,1\}^n$. Suppose that the hash key τ is non-zero. Then, $f(b, x) = f(b', x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau)$, where α_0, α_1 and τ are as described earlier.*

Proof. For the input $\alpha_b||x||m||m$, we have,

$$\begin{aligned} SP &= \sigma^{-1}(\alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus m\tau); \\ Q_3 &= m \oplus SP \oplus \alpha^2\beta; \\ Q_4 &= m \oplus SP \oplus \alpha^3\beta; \\ S_3 &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus \alpha^2\beta; \\ S_4 &= E_{K_2}(m \oplus SP \oplus \alpha^3\beta) \oplus \alpha^3\beta; \\ y_3 \oplus y_4 &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP \oplus \alpha^3\beta) \oplus \alpha^3\beta; \end{aligned}$$

For the input $\alpha_{b'}||x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau||m||m$, we have,

$$\begin{aligned} SP' &= \sigma^{-1}(\alpha_{b'}\tau^4 \oplus x\tau^3 \oplus \alpha_b\tau^4 \oplus \alpha_{b'}\tau^4 \oplus m\tau^2 \oplus m\tau) = \sigma^{-1}(\alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus m\tau); \\ Q'_3 &= m \oplus SP' \oplus \alpha^2\beta; \\ Q'_4 &= m \oplus SP' \oplus \alpha^3\beta; \\ S'_3 &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus \alpha^2\beta; \\ S'_4 &= E_{K_2}(m \oplus SP' \oplus \alpha^3\beta) \oplus \alpha^3\beta; \\ y'_3 \oplus y'_4 &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP' \oplus \alpha^3\beta) \oplus \alpha^3\beta. \end{aligned}$$

Now, $SP = SP'$ implies $y_3 \oplus y_4 = y'_3 \oplus y'_4$.
Hence, the proposition is proved. □

The above proposition establishes that $1||\alpha_0\tau \oplus \alpha_1\tau$ is a period of f. Proposition 2 falls short of showing that f satisfies the promise of Simon's problem. We show below, that f satisfies an approximate promise.

Proposition 3. *Assume that the block cipher E instantiated with a uniform random key, behaves like a uniform random function. Suppose that the hash key τ is non-zero. Then, for f defined in (4), $\varepsilon(f, 1||\alpha_0\tau \oplus \alpha_1\tau) \leq 5/2^n$.*

Proof. Let $\eta||t \notin \{0||0^n, 1||\alpha_0\tau \oplus \alpha_1\tau\}$ be such that the probability of $f(b,x) = f(b \oplus \eta, x \oplus t)$ is maximised.

Case 1: Suppose $\eta = 0$. Then t is necessarily non-zero. Let,

$$SP = \sigma^{-1}(\alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus m\tau); \qquad SP' = \sigma^{-1}(\alpha_b\tau^4 \oplus x\tau^3 \oplus t\tau^3 \oplus m\tau^2 \oplus m\tau). \tag{5}$$

As t is necessarily non-zero, $SP \neq SP'$.
We have,

$$\begin{aligned} f(b,x) &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP \oplus \alpha^3\beta) \oplus \alpha^3\beta; \\ f(b,x \oplus t) &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP' \oplus \alpha^3\beta) \oplus \alpha^3\beta. \end{aligned}$$

Then $f(b,x) = f(b,x \oplus t)$ if and only if $X_1 = X_2$, where

$$\begin{aligned} X_1 &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus E_{K_2}(m \oplus SP \oplus \alpha^3\beta); \\ X_2 &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus E_{K_2}(m \oplus SP' \oplus \alpha^3\beta). \end{aligned}$$

Let $\mathbf{E}$ be the event that $(m \oplus SP \oplus \alpha^2\beta)$, $(m \oplus SP \oplus \alpha^3\beta)$, $(m \oplus SP' \oplus \alpha^2\beta)$ and $(m \oplus SP' \oplus \alpha^3\beta)$ are distinct. As $SP \neq SP'$, clearly $(m \oplus SP \oplus \alpha^2\beta)$ and $(m \oplus SP' \oplus \alpha^2\beta)$ are distinct and so are $(m \oplus SP \oplus \alpha^3\beta)$ and $(m \oplus SP' \oplus \alpha^3\beta)$. As β is generated through the application of a PRF, the probability that $(\alpha^2\beta = \alpha^3\beta)$ is $1/2^n$. With similar reasoning probability that $SP \oplus SP' \oplus \alpha^2\beta \oplus \alpha^3\beta = 0$ is $1/2^n$. So, we have $\Pr[\overline{\mathbf{E}}] = 4/2^n$. Conditioned on $\mathbf{E}$, and under the assumption that E behaves like a uniform random function, the probability that $X_1 = X_2$ is at most $1/2^n$. Using (2) we have $\Pr_{b,x}[f(b,x) = f(b \oplus \eta, x \oplus t)] = \Pr[X_1 = X_2] \leq 5/2^n$.

Case 2: Suppose $\eta = 1$. Then $t \neq \alpha_0\tau \oplus \alpha_1\tau$. Let $b' = b \oplus 1$. Let,

$$SP = \sigma^{-1}(\alpha_b\tau^4 \oplus x\tau^3 \oplus m\tau^2 \oplus m\tau); \qquad SP' = \sigma^{-1}(\alpha_{b'}\tau^4 \oplus x\tau^3 \oplus t\tau^3 \oplus m\tau^2 \oplus m\tau); \tag{6}$$

As $t \neq \alpha_0\tau \oplus \alpha_1\tau$, $SP \neq SP'$.
We have,

$$\begin{aligned} f(b,x) &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP \oplus \alpha^3\beta) \oplus \alpha^3\beta; \\ f(b \oplus \eta, x \oplus t) &= f(b', x \oplus t) \\ &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus \alpha^2\beta \oplus E_{K_2}(m \oplus SP' \oplus \alpha^3\beta) \oplus \alpha^3\beta. \end{aligned}$$

So, again $f(b,x) = f(b \oplus \eta, x \oplus t)$ if and only if $X_1 = X_2$, where

$$\begin{aligned} X_1 &= E_{K_2}(m \oplus SP \oplus \alpha^2\beta) \oplus E_{K_2}(m \oplus SP \oplus \alpha^3\beta); \\ X_2 &= E_{K_2}(m \oplus SP' \oplus \alpha^2\beta) \oplus E_{K_2}(m \oplus SP' \oplus \alpha^3\beta). \end{aligned}$$

A reasoning similar to Case 1 shows that $\Pr_{b,x}[f(b,x) = f(b \oplus \eta, x \oplus t)]$ is at most $5/2^n$. □

Classical queries: Given the period $1||s = 1||\tau(\alpha_0 \oplus \alpha_1)$, the two classical queries required in Section 3 are the following. The first query is $\alpha_0||x||m||m$ with output $y_1||y_2||y_3||y_4$ and the second query is $\alpha_1||x \oplus s||m||m$ with output $y'_1||y'_2||y'_3||y'_4$. From the proof of Proposition 2 we have that $y_3 \oplus y_4 = y'_3 \oplus y'_4$ which defines the relation between the outputs of the two classical queries.

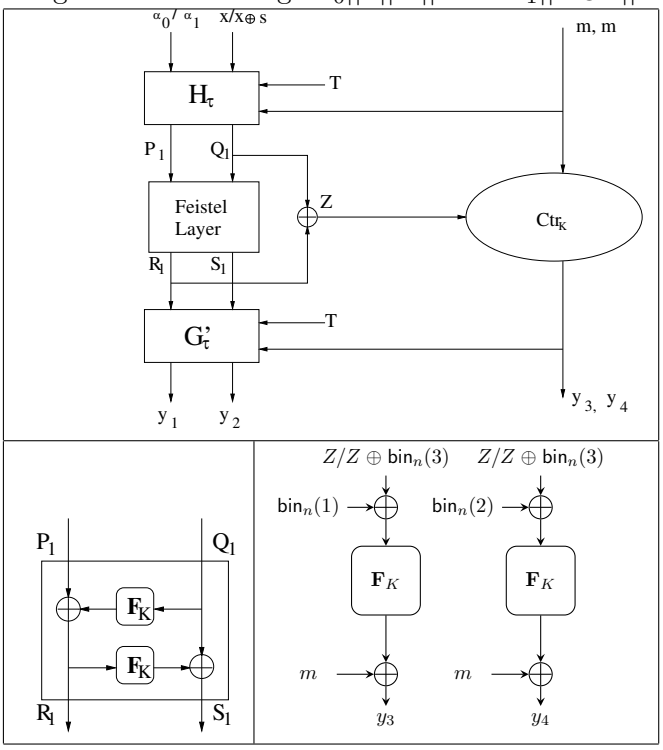
Partial key recovery: Once $s = \tau(\alpha_0 \oplus \alpha_1)$ has been obtained, since α_0 and α_1 are distinct, from s, one obtains the hash key τ as $\tau = s(\alpha_0 \oplus \alpha_1)^{-1}$.

4.3 FAST

FAST was proposed by Chakraborty, Ghosh, López and Sarkar [4]. It is built using a fixed input length pseudo-random function and an appropriate hash function. The key K of FAST is the same as the key of the underlying pseudo-random function. The pseudo-random function maps n-bit strings to n-bit strings. For the sake of concreteness, we fix $n = 128$. Let $\mathbf{F}_K$ denote the pseudo-random function instantiated with the key K. FAST is targeted towards two application scenarios. We describe the quantum attack on the instantiation targeted towards the specific task of disk encryption. In this case, the tweak space is $\mathcal{T} = \{0,1\}^n$ and the message space is $\mathcal{P} = \{0,1\}^{\mathfrak{m}n}$, where $\mathfrak{m} > 2$ is determined by the size of a disk sector. In our attack we will fix the tweak to be an arbitrary $T \in \mathcal{T}$.

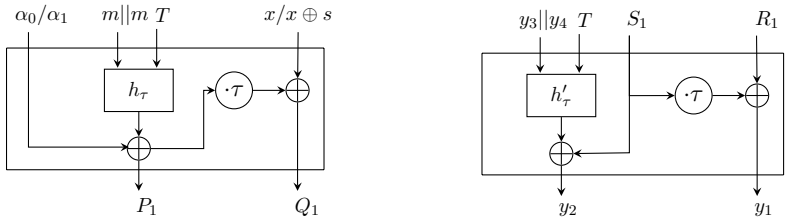
Our attack considers 4-block messages. So, we briefly describe the encryption of 4-block messages with reference to Figures 3 and 4. From a top level view, FAST consists of three distinct layers

Figure 3: Enciphering a 4-block message $\alpha_0||x||m||m$ or $\alpha_1||x \oplus s||m||m$ under FAST.



- hash-encrypt-hash. The hashing layers $\mathbf{H}$ and $\mathbf{G}'$ are based on two universal hash functions h and h', both having the key τ and $h' = \tau h$. For more details of these functions we refer to [4]. The encryption layer consists of a two-round Feistel network and a counter mode Ctr. The two-round Feistel is built using the PRF $\mathbf{F}_K$ and processes the first two blocks of the plaintext. The third and fourth blocks are encrypted in a counter mode built using $\mathbf{F}_K$. The offset for the counter mode is derived from the input and output of the Feistel layer. The input of the Feistel layer is obtained by processing the plaintext and the tweak through the first hash layer. The second hash layer generates the first two blocks of the ciphertext by processing the output of the Feistel layer and the third and fourth blocks of the ciphertext. Some more details are provided as part of the attack.

Figure 4: The hash functions $\mathbf{H}$ (left) and $\mathbf{G}'$ (right).



Fix $m, \alpha_0, \alpha_1 \in \{0,1\}^n$, such that $\alpha_0 \oplus \alpha_1 = 0^{126}11$; let b denote a bit. We define the following

function.

$$\begin{aligned} f : \{0,1\} \times \{0,1\}^n &\rightarrow \{0,1\}^n \\ (b,x) &\stackrel{f}{\mapsto} y_3 \oplus y_4, \text{ where } y_1||y_2||y_3||y_4 \longleftarrow \mathsf{FAST.Encrypt}_K(T, \alpha_b||x||m||m). \quad (7) \end{aligned}$$

The function f defined in (7) satisfies the following property.

Proposition 4. *Let $b, b' \in \{0,1\}, x \in \{0,1\}^n$. Suppose that the hash key τ is non-zero. Then, $f(b,x) = f(b', x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau)$. α_0, α_1 and τ are as described before.*

Proof. For the input $(T, \alpha_b||x||m||m)$,

$$\begin{aligned} P_1 &= \alpha_b \oplus h_\tau(T, m||m); \\ Q_1 &= x \oplus \tau(\alpha_b \oplus h_\tau(T, m||m)); \\ R_1 &= \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1); \\ Z &= Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1); \\ y_3 &= m \oplus \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1)); \\ y_4 &= m \oplus \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(2)); \\ y_3 \oplus y_4 &= \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1)) \\ &\quad \oplus \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(2)); \end{aligned}$$

For the input $(T, \alpha_{b'}||x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau||m||m)$,

$$\begin{aligned} P'_1 &= \alpha_{b'} \oplus h_\tau(T, m||m); \\ Q'_1 &= x \oplus \alpha_b\tau \oplus \alpha_{b'}\tau \oplus \tau(\alpha_{b'} \oplus h_\tau(T, m||m)) \\ &= x \oplus \alpha_b\tau \oplus \tau h_\tau(T, m||m); \\ R'_1 &= \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1); \\ Z' &= Q'_1 \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1); \\ y'_3 &= m \oplus \mathbf{F}_K(Q'_1 \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(1)); \\ y'_4 &= m \oplus \mathbf{F}_K(Q'_1 \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(2)); \\ y'_3 \oplus y'_4 &= \mathbf{F}_K(Q'_1 \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(1)) \oplus \mathbf{F}_K(Q'_1 \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(2)) \\ &= \mathbf{F}_K(Q'_1 \oplus \alpha_b \oplus \mathsf{bin}_n(3) \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(1)) \oplus \mathbf{F}_K(Q'_1 \oplus \alpha_b \oplus \mathsf{bin}_n(3) \oplus h_\tau(T, m||m) \\ &\quad \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(2)), \text{ as } \alpha_b \oplus \alpha_{b'} = \mathsf{bin}_n(3) \\ &= \mathbf{F}_K(Q'_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(2)) \\ &\quad \oplus \mathbf{F}_K(Q'_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q'_1) \oplus \mathsf{bin}_n(1)). \end{aligned}$$

We observe that $Q_1 = Q'_1$, which implies $y_3 \oplus y_4 = y'_3 \oplus y'_4$. This proves the proposition. □

The above discussion establishes that $1||\alpha_0\tau \oplus \alpha_1\tau$ is a period of f. Proposition 4 falls short of showing that f satisfies the promise of Simon's problem. We show below, that f satisfies an approximate promise.

Proposition 5. *Assume that the PRF $\mathbf{F}$ instantiated with a uniform random key, behaves like a uniform random function. Suppose that the hash key τ is non-zero. Then, for f defined in (7), $\varepsilon(f, 1||\alpha_0\tau \oplus \alpha_1\tau) \leq \frac{3}{2^n}$.*

Proof. Let $\eta||t \notin \{0||0^n, 1||\alpha_0\tau \oplus \alpha_1\tau\}$ be such that the probability of $f(b,x) = f(b \oplus \eta, x \oplus t)$ is maximised.

- Case 1: Suppose $\eta = 0$. Then t is necessarily non-zero. We have

$$\begin{aligned} f(b,x) &= \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1)) \\ &\quad \oplus \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(2)); \\ f(b, x \oplus t) &= \mathbf{F}_K(Q_1 \oplus t \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1 \oplus t) \oplus \mathsf{bin}_n(1)) \\ &\quad \oplus \mathbf{F}_K(Q_1 \oplus t \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1 \oplus t) \oplus \mathsf{bin}_n(2)). \end{aligned}$$

  Let

$$\begin{aligned} a_1 &= Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1), \\ a_2 &= Q_1 \oplus t \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1 \oplus t) \oplus \mathsf{bin}_n(1), \\ u &= \mathsf{bin}_n(1) \oplus \mathsf{bin}_n(2); \end{aligned}$$

  Hence

$$\begin{aligned} f(b,x) &= \mathbf{F}_K(a_1) \oplus \mathbf{F}_K(a_1 \oplus u); \\ f(b, x \oplus t) &= \mathbf{F}_K(a_2) \oplus \mathbf{F}_K(a_2 \oplus u). \end{aligned}$$

  Let $\mathbf{E}$ be the event that a_1, $a_1 \oplus u$, a_2 and $a_2 \oplus u$ are distinct. Clearly a_1 and $a_1 \oplus u$ are distinct and so are a_2 and $a_2 \oplus u$. Now we consider the following events.

  - $\mathbf{E}_1 := a_1 = a_2$ or, equivalently $\mathbf{F}_K(Q_1) \oplus \mathbf{F}_K(Q_1 \oplus t) = t$. As $t \neq 0$ and $\mathbf{F}_K$ is assumed to be a uniform random function, hence $\Pr[\mathbf{E}_1] = \frac{1}{2^n}$.
  - $\mathbf{E}_2 := a_1 = a_2 \oplus u$ or, equivalently $\mathbf{F}_K(Q_1) \oplus \mathbf{F}_K(Q_1 \oplus t) = t \oplus u$. As in the previous case, $\Pr[\mathbf{E}_2] = \frac{1}{2^n}$.

  Hence, $\overline{\mathbf{E}} := \mathbf{E}_1 \cup \mathbf{E}_2$. Note that since u is a non-zero string $\mathbf{E}_1$ and $\mathbf{E}_2$ are disjoint. Hence, $\Pr[\overline{\mathbf{E}}] = \Pr[\mathbf{E}_1] + \Pr[\mathbf{E}_2] = \frac{2}{2^n}$. Conditioned on $\mathbf{E}$, and under the assumption that $\mathbf{F}_K$ behaves like a uniform random function, the probability that $f(b,x) = f(b, x \oplus t)$ is at most $1/2^n$. Using (2) we have $\Pr_{b,x}[f(b,x) = f(b \oplus \eta, x \oplus t)] \leq 3/2^n$.

- Case 2: Suppose $\eta = 1$. Then $t \neq \alpha_0\tau \oplus \alpha_1\tau$. Let $b' = b \oplus 1$, $Q_1' = x \oplus t \oplus \tau(\alpha_{b'} \oplus h_\tau(T, m||m))$. We have

$$\begin{aligned} f(b,x) &= \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1)) \\ &\quad \oplus \mathbf{F}_K(Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(2)); \\ f(b \oplus \eta, x \oplus t) &= f(b', x \oplus t) \\ &= \mathbf{F}_K(Q_1' \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1') \oplus \mathsf{bin}_n(1)) \\ &\quad \oplus \mathbf{F}_K(Q_1' \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1') \oplus \mathsf{bin}_n(2)); \end{aligned}$$

  Let

$$\begin{aligned} a_1 &= Q_1 \oplus \alpha_b \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1) \oplus \mathsf{bin}_n(1) \\ a_2 &= Q_1' \oplus \alpha_{b'} \oplus h_\tau(T, m||m) \oplus \mathbf{F}_K(Q_1') \oplus \mathsf{bin}_n(1), \\ u &= \mathsf{bin}_n(1) \oplus \mathsf{bin}_n(2); \end{aligned}$$

Hence

$$\begin{aligned} f(b,x) &= \mathbf{F}_K(a_1) \oplus \mathbf{F}_K(a_1 \oplus u); \\ f(b \oplus \eta, x \oplus t) &= \mathbf{F}_K(a_2) \oplus \mathbf{F}_K(a_2 \oplus u). \end{aligned}$$

A reasoning similar to Case 1 shows that $\Pr_{b,x}[f(b,x) = f(b \oplus \eta, x \oplus t)]$ is at most $3/2^n$.

□

Classical queries: Given the period $1||s = 1||\tau(\alpha_0 \oplus \alpha_1)$, the two classical queries required in Section 3 are the following. The first query is $\alpha_0||x||m||m$ with output $y_1||y_2||y_3||y_4$ and the second query is $\alpha_1||x \oplus \tau(\alpha_0 \oplus \alpha_1)||m||m$ with output $y_1'||y_2'||y_3'||y_4'$. From the proof of Proposition 4 we have that $y_3 \oplus y_4 = y_3' \oplus y_4'$ which defines the relation between the outputs of the two classical queries.

Partial key recovery: Once $s = \tau(\alpha_0 \oplus \alpha_1)$ has been obtained, since α_0 and α_1 are distinct, from s, one obtains the hash key τ as $\tau = s(\alpha_0 \oplus \alpha_1)^{-1}$.

5 Distinguishing Attacks

5.1 CMC

CMC was proposed by Halevi and Rogaway [10], in 2003. It is based on the CBC mode of operation of a block cipher. The block length of the block cipher can be assumed to be n-bit. CMC has the key space $\mathcal{K} \times \mathcal{K}$, where $\mathcal{K}$ is the key space for the underlying block cipher and the tweak space $\mathcal{T} = \{0,1\}^n$. The message space of CMC is $\mathcal{P} = \bigcup_{i \in I} \{0,1\}^i$ for some non-empty index set $I \subseteq \mathbb{N}$. Let E_K denote the encryption function of the underlying block cipher instantiated with the key K.

Our attack considers 3-block messages. Hence, we briefly describe the encryption of 3-block messages with reference to Figure 5. Let CMC be instantiated with the key $(K, \tilde{K}) \in \mathcal{K} \times \mathcal{K}$. In our attack we will fix the tweak to be an arbitrary $T \in \mathcal{T}$. $\tilde{K}$ is used as the key to the block cipher E only to produce $\mathbb{T}$ from T. At a conceptual level, the CMC encryption function consists of three layers. The first layer is essentially CBC encryption on the message blocks, followed by a layer of masking and the third layer is CBC decryption. The rest of the encryption algorithm can be understood from Figure 5. We provide more details as part of the attack.

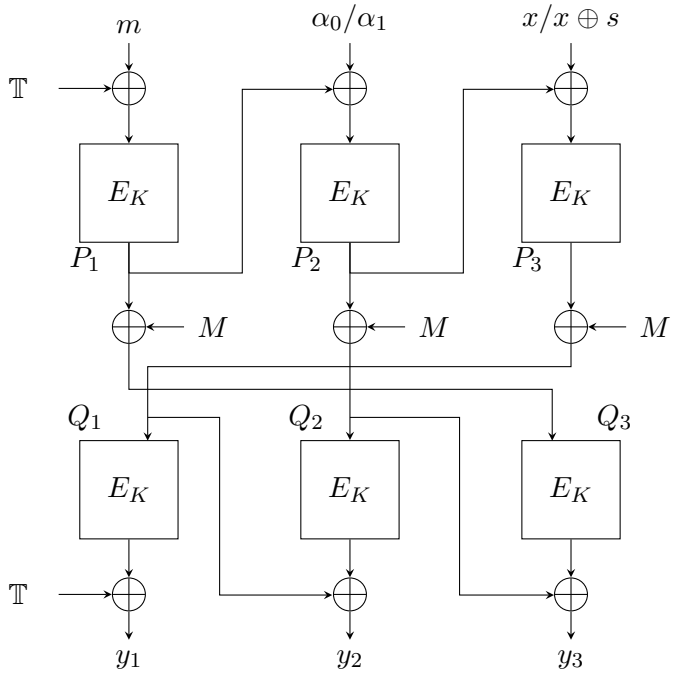
Fix $m, \alpha_0, \alpha_1 \in \{0,1\}^n$, such that $\alpha_0 \neq \alpha_1$; let b denote a bit and we define the following function.

$$\begin{aligned} f : \{0,1\} \times \{0,1\}^n &\rightarrow \{0,1\}^n \\ (b,x) &\overset{f}{\mapsto} y_1, \text{ where } y_1||y_2||y_3 \longleftarrow \mathsf{CMC.Encrypt}_{K,\tilde{K}}(T, m||\alpha_b||x). \end{aligned} \tag{8}$$

The function f defined in (8) satisfies the following property.

Proposition 6. *Let $b, b' \in \{0,1\}, x, x' \in \{0,1\}^n$. Then, $f(b,x) = f(b',x') \Leftrightarrow x \oplus x' = E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_b) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_{b'})$, where the constants α_0 and α_1 are as fixed before.*

Figure 5: Enciphering a 3-block message $m||\alpha_0||x$ or $m||\alpha_1||x \oplus s$ under CMC. Correspondingly, $M = 2(P_1 \oplus P_3)$ and $M' = 2(P'_1 \oplus P'_3)$.



Proof. Let $s = E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_1)$. For the input $m||\alpha_0||x$,

$$\begin{aligned} P_1 &= E_K(m \oplus \mathbb{T}); \\ P_2 &= E_K(\alpha_0 \oplus E_K(m \oplus \mathbb{T})); \\ P_3 &= E_K(x \oplus E_K(\alpha_0 \oplus E_K(m \oplus \mathbb{T}))); \\ M &= 2(P_1 \oplus P_3); \\ Q_1 &= P_3 \oplus M; \\ y_1 &= E_K(Q_1) \oplus \mathbb{T}; \end{aligned} \tag{9}$$

For the input $m||\alpha_1||x \oplus s$,

$$\begin{aligned} P'_1 &= E_K(m \oplus \mathbb{T}); \\ P'_2 &= E_K(\alpha_1 \oplus E_K(m \oplus \mathbb{T})); \\ P'_3 &= E_K(x \oplus s \oplus E_K(\alpha_1 \oplus E_K(m \oplus \mathbb{T}))) \\ &= E_K(x \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_1) \oplus E_K(\alpha_1 \oplus E_K(m \oplus \mathbb{T}))) \\ &= E_K(x \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0)); \\ M' &= 2(P'_1 \oplus P'_3); \\ Q'_1 &= P'_3 \oplus M'; \\ y'_1 &= E_K(Q'_1) \oplus \mathbb{T}. \end{aligned} \tag{10}$$

We see $P_1 = P_1'$ and $P_3 = P_3'$ implying $M = M'$; $P_3 = P_3'$ and $M = M'$ together imply $Q_1 = Q_1'$, which finally establishes $y_1 = y_1'$. This proves one direction of the proposition. Now we see the other direction.

$$\begin{aligned}
f(b,x) = f(b',x') &\Rightarrow y_1 = y_1' \\
&\Rightarrow E_K(Q_1) \oplus \mathbb{T} = E_K(Q_1') \oplus \mathbb{T} \\
&\Rightarrow Q_1 = Q_1' \\
&\Rightarrow P_3 \oplus M = P_3' \oplus M' \\
&\Rightarrow P_3 \oplus 2(P_1 \oplus P_3) = P_3' \oplus 2(P_1' \oplus P_3') \\
&\Rightarrow P_3 \oplus 2(P_1 \oplus P_3) = P_3' \oplus 2(P_1 \oplus P_3') \text{ (as } P_1' = P_1) \\
&\Rightarrow P_3 = P_3' \\
&\Rightarrow E_K(x \oplus E_K(\alpha_b \oplus E_K(m \oplus \mathbb{T}))) = E_K(x' \oplus E_K(\alpha_{b'} \oplus E_K(m \oplus \mathbb{T}))) \\
&\Rightarrow x \oplus E_K(\alpha_b \oplus E_K(m \oplus \mathbb{T})) = x' \oplus E_K(\alpha_{b'} \oplus E_K(m \oplus \mathbb{T})) \\
&\Rightarrow x \oplus x' = E_K(\alpha_b \oplus E_K(m \oplus \mathbb{T})) \oplus E_K(\alpha_{b'} \oplus E_K(m \oplus \mathbb{T})).
\end{aligned}$$

□

The above proposition proves that $1||E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_1)$ is a period for the function f and f is a 2-to-1 function. So, Simon's algorithm applied to f uncovers this period with high probability.

Obtaining the period $1||s$, where $s = E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_1)$ provides a distinguishing attack against CMC.

Classical queries: Given the period $1||s = 1||E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_0) \oplus E_K(E_K(m \oplus \mathbb{T}) \oplus \alpha_1)$, the two classical queries required in Section 3 are the following. The first query is $m||\alpha_0||x$ with output $y_1||y_2||y_3$ and the second query is $m||\alpha_1||x \oplus s$ with output $y_1'||y_2'||y_3'$. From the proof of Proposition 6 we have that $y_1 = y_1'$ which defines the relation between the outputs of the two classical queries.

5.2 EME

EME also was proposed by Halevi and Rogaway [11]. It was later extended to handle arbitrary length messages by Halevi [8] and the resulting scheme was called EME*. EME* has been standardised as a TES by IEEE [1] in the name EME2.

Our attack considers 3-block messages. For this message length the constructions EME and EME2 are identical, with only the minor replacement of the tweak by a function of the tweak in the latter. Hence, we will describe the attack in the context of EME only.

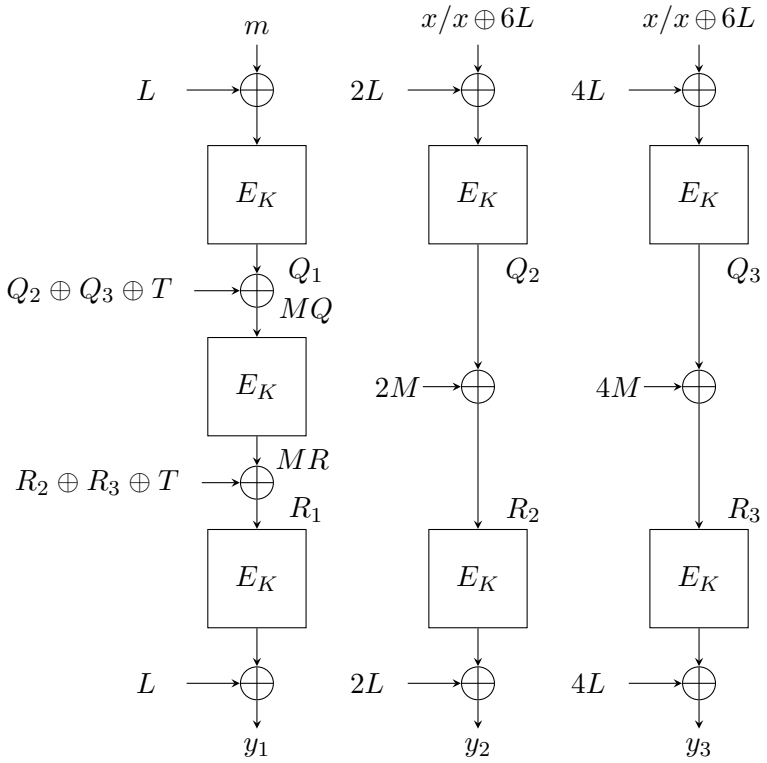
EME has the key space $\mathcal{K}$, the same as the underlying block cipher having block size n; the tweak space is $\mathcal{T} = \{0,1\}^n$. The message space of EME is $\mathcal{P} = \{0,1\}^n \cup \{0,1\}^{2n} \cup \cdots \cup \{0,1\}^{n^2}$. The key K of EME is the same as the key of the underlying block cipher. Let E_K denote the encryption function of the underlying block cipher instantiated with the key K. In our attack we fix the tweak to be an arbitrary $T \in \mathcal{T}$.

As our attack considers 3-block messages, we will briefly describe the encryption of 3-block messages with reference to Figure 6. The encryption of each of the message blocks consists of five layers: initial masking followed by an application of E_K, then another masking followed by another

application of E_K and the final masking. One of the masking elements is $L = 2E_K(0^n)$. The middle layer of masking for the first block is of different form from that for second and third blocks. The rest of the encryption algorithm can be understood from Figure 6. We provide more details as part of the attack.

Notation: Let α denote a root of the primitive polynomial used to represent $GF(2^n)$. Note that, $2L$ represents L times the field element denoted by α. Similarly, $4L$ represents L times the field element denoted by α^2; $6L$ represents L times the field element denoted by $\alpha^2 \oplus \alpha$. Hence, $6L = (\alpha^2 \oplus \alpha)L = 4L \oplus 2L$. This notation has been used in [11] and so we follow this notation.

Figure 6: Enciphering a 3-block message $m||x||x$ or $m||x \oplus 6L||x \oplus 6L$ under EME. Correspondingly, $M = MQ \oplus E_K(MQ)$, where $MQ = E_K(m \oplus L) \oplus E_K(x \oplus 2L) \oplus E_K(x \oplus 4L) \oplus T$ and $M' = MQ' \oplus E_K(MQ')$ where $MQ' = E_K(m \oplus L) \oplus E_K(x \oplus 4L) \oplus E_K(x \oplus 2L) \oplus T$.



Fix $m \in \{0,1\}^n$ and we define the following function.

$$\begin{aligned} f : \{0,1\}^n &\rightarrow \{0,1\}^n \\ x &\xmapsto{f} y_1, \text{ where } y_1||y_2||y_3 \longleftarrow \mathsf{EME.Encrypt}_K(T, m||x||x). \end{aligned} \tag{11}$$

The function f defined in (11) satisfies the following property.

Proposition 7. *Let $x \in \{0,1\}^n$. Then, $f(x) = f(x \oplus 6L)$, where L is as defined before and suppose it is non-zero.*

Proof. Consider the two inputs $m||x||x$ and $m||(x \oplus 6L)||(x \oplus 6L)$.
For the input $m||x||x$,

$$\begin{aligned}
Q_1 &= E_K(m \oplus L);\\
Q_2 &= E_K(x \oplus 2L);\\
Q_3 &= E_K(x \oplus 4L);\\
MQ &= E_K(m \oplus L) \oplus E_K(x \oplus 2L) \oplus E_K(x \oplus 4L) \oplus T;\\
MR &= E_K(MQ);\\
M &= MQ \oplus MR;\\
R_2 &= E_K(x \oplus 2L) \oplus 2M;\\
R_3 &= E_K(x \oplus 4L) \oplus 4M;\\
R_1 &= MR \oplus R_2 \oplus R_3 \oplus T;\\
y_1 &= L \oplus E_K(R_1) = E_K(MR \oplus R_2 \oplus R_3 \oplus T) \oplus L; \qquad (12)
\end{aligned}$$

For the input $m||(x \oplus 6L)||(x \oplus 6L)$,

$$\begin{aligned}
Q'_1 &= E_K(m \oplus L);\\
Q'_2 &= E_K(x \oplus 4L);\\
Q'_3 &= E_K(x \oplus 2L);\\
MQ' &= E_K(m \oplus L) \oplus E_K(x \oplus 4L) \oplus E_K(x \oplus 2L) \oplus T;\\
MR' &= E_K(MQ');\\
M' &= MQ' \oplus MR';\\
R'_2 &= E_K(x \oplus 4L) \oplus 2M';\\
R'_3 &= E_K(x \oplus 2L) \oplus 4M';\\
R'_1 &= MR' \oplus R'_2 \oplus R'_3 \oplus T;\\
y'_1 &= L \oplus E_K(R'_1) = E_K(MR' \oplus R'_2 \oplus R'_3 \oplus T) \oplus L. \qquad (13)
\end{aligned}$$

From above we see that $MQ = MQ'$; hence, $MR = MR'$ and $M = M'$; $M = M'$ implies $R_2 \oplus R_3 = R'_2 \oplus R'_3$. Hence, we see, $y_1 = y'_1$, proving the proposition. □

The above discussion establishes that $6L$ is a period of f. Proposition 7 falls short of showing that f satisfies the promise of Simon's problem. We show below, that f satisfies an approximate promise.

Proposition 8. *Assume that the block cipher E instantiated with a uniform random key, behaves like a uniform random function. Suppose L is non-zero. Then, for f defined in (11), $\varepsilon(f, 6L) \leq 1/2^{n-1}$.*

Proof. Let $t \notin \{0^n, 6L\}$ be such that the probability of $f(x) = f(x \oplus t)$ is maximised.

We have

$$\begin{aligned}
f(x) &= L \oplus E_K(MR \oplus R_2 \oplus R_3 \oplus T);\\
f(x \oplus t) &= L \oplus E_K(MR' \oplus R'_2 \oplus R'_3 \oplus T);\ \text{where } MR = E_K(MQ), MR' = E_K(MQ'),\\
MQ &= E_K(m \oplus L) \oplus E_K(x \oplus 2L) \oplus E_K(x \oplus 4L) \oplus T,\\
MQ' &= E_K(m \oplus L) \oplus E_K(x \oplus t \oplus 2L) \oplus E_K(x \oplus t \oplus 4L) \oplus T,\\
R_2 &= E_K(x \oplus 2L) \oplus 2M, R_3 = E_K(x \oplus 4L) \oplus 4M,\\
R'_2 &= E_K(x \oplus t \oplus 2L) \oplus 2M', R'_3 = E_K(x \oplus t \oplus 4L) \oplus 4M',\\
M &= MQ \oplus MR, M' = MQ' \oplus MR'.
\end{aligned}$$

Then $f(x) = f(x \oplus t) \Leftrightarrow E_K(MR \oplus R_2 \oplus R_3 \oplus T) = E_K(MR' \oplus R'_2 \oplus R'_3 \oplus T) \Leftrightarrow MR \oplus R_2 \oplus R_3 = MR' \oplus R'_2 \oplus R'_3$. So,

$$\begin{aligned}
&\Pr[f(x) = f(x \oplus t)]\\
&= \Pr[MR \oplus MR' = R_2 \oplus R_3 \oplus R'_2 \oplus R'_3]\\
&= \Pr[MR \oplus MR' = MQ \oplus MQ' \oplus 6M \oplus 6M'] \text{ (as } R_2 \oplus R_3 \oplus R'_2 \oplus R'_3 = MQ \oplus MQ' \oplus 6M \oplus 6M')\\
&= \Pr[MQ \oplus MR \oplus MQ' \oplus MR' = 6M \oplus 6M']\\
&= \Pr[M \oplus M' = 6M \oplus 6M'] \qquad (14)\\
&= \Pr[M = M'] \qquad (15)\\
&= \Pr[E_K(MQ) \oplus E_K(MQ') = MQ \oplus MQ'] \qquad (16)
\end{aligned}$$

The explanation for obtaining (15) from (14) is the following. From (14), we have $7M = 7M'$. Recall that here 7 is a shorthand for the polynomial $\alpha^2 \oplus \alpha \oplus 1$, where α is the root of the degree n primitive polynomial used to represent the field. So, $\alpha^2 \oplus \alpha \oplus 1$ is an invertible element in the field and hence the equality $M = M'$ follows.

Let $\mathbf{E}$ be the event $MQ \neq MQ'$. Under the assumption that E_K behaves like a uniform random function, $\Pr[E_K(MQ) \oplus E_K(MQ') = MQ \oplus MQ' | \mathbf{E}] \leq 1/2^n$. Since $t \notin \{0^n, 6L\}$ and L is non-zero, $x \oplus 2L$, $x \oplus 4L$, $x \oplus t \oplus 2L$ and $x \oplus t \oplus 4L$ are distinct. As a result, under the assumption that E_K behaves like a uniform random function, we have, $\Pr[\overline{\mathbf{E}}] \leq 1/2^n$. From (2) and (16) we have $\Pr[f(x) = f(x \oplus t)] \leq 2/2^n$. □

Classical queries: Given the period $6L$, the two classical queries required in Section 3 are the following. The first query is $m||x||x$ with output $y_1||y_2||y_3$ and the second query is $m||x \oplus 6L||x \oplus 6L$ with output $y'_1||y'_2||y'_3$. From the proof of Proposition 7 we have that $y_1 = y'_1$ which defines the relation between the outputs of the two classical queries.

6 Conclusion

This work showed that some of the well known TESs which are secure in the classical world are broken in the quantum world. This brings up the following question. Can some simple modifications of these schemes make them quantum secure? Perhaps future research will answer this question.

Acknowledgement

We are grateful to the reviewers for their kind comments which have helped in improving the paper.

References

[1] IEEE Std 1619.2-2010: IEEE standard for wide-block encryption for shared storage media. `http://standards.ieee.org/findstds/standard/1619.2-2010.html`, March 2011.

[2] Xavier Bonnetain. Quantum key-recovery on full AEZ. In Carlisle Adams and Jan Camenisch, editors, *Selected Areas in Cryptography - SAC 2017 - 24th International Conference, Ottawa, ON, Canada, August 16-18, 2017, Revised Selected Papers*, volume 10719 of *Lecture Notes in Computer Science*, pages 394–406. Springer, 2017.

[3] Xavier Bonnetain, Akinori Hosoyamada, María Naya-Plasencia, Yu Sasaki, and André Schrottenloher. Quantum attacks without superposition queries: The offline Simon's algorithm. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology - ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security, Kobe, Japan, December 8-12, 2019, Proceedings, Part I*, volume 11921 of *Lecture Notes in Computer Science*, pages 552–583. Springer, 2019.

[4] Debrup Chakraborty, Sebati Ghosh, Cuauhtemoc Mancillas López, and Palash Sarkar. FAST: disk encryption and beyond. *Advances in Mathematics of Communications*. `https://www.aimsciences.org/article/doi/10.3934/amc.2020108`.

[5] Xiaoyang Dong, Bingyou Dong, and Xiaoyun Wang. Quantum attacks on some Feistel block ciphers. *Des. Codes Cryptogr.*, 88(6):1179–1203, 2020.

[6] Xiaoyang Dong and Xiaoyun Wang. Quantum key-recovery attack on Feistel structures. *Sci. China Inf. Sci.*, 61(10):102501:1–102501:7, 2018.

[7] Lov K. Grover. A fast quantum mechanical algorithm for database search. In Gary L. Miller, editor, *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 212–219. ACM, 1996.

[8] Shai Halevi. EME*: Extending EME to handle arbitrary-length messages with associated data. In Anne Canteaut and Kapalee Viswanathan, editors, *Progress in Cryptology - INDOCRYPT 2004, 5th International Conference on Cryptology in India, Chennai, India, December 20-22, 2004, Proceedings*, volume 3348 of *Lecture Notes in Computer Science*, pages 315–327. Springer, 2004.

[9] Shai Halevi. Invertible universal hashing and the TET encryption mode. In Alfred Menezes, editor, *Advances in Cryptology - CRYPTO 2007, 27th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2007, Proceedings*, volume 4622 of *Lecture Notes in Computer Science*, pages 412–429. Springer, 2007.

[10] Shai Halevi and Phillip Rogaway. A tweakable enciphering mode. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, volume 2729 of *Lecture Notes in Computer Science*, pages 482–499. Springer, 2003.

[11] Shai Halevi and Phillip Rogaway. A parallelizable enciphering mode. In Tatsuaki Okamoto, editor, *Topics in Cryptology - CT-RSA 2004, The Cryptographers' Track at the RSA Conference 2004, San Francisco, CA, USA, February 23-27, 2004, Proceedings*, volume 2964 of *Lecture Notes in Computer Science*, pages 292–304. Springer, 2004.

[12] Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part II*, volume 9815 of *Lecture Notes in Computer Science*, pages 207–237. Springer, 2016.

[13] Hidenori Kuwakado and Masakatu Morii. Quantum distinguisher between the 3-round Feistel cipher and the random permutation. In *IEEE International Symposium on Information Theory, ISIT 2010, June 13-18, 2010, Austin, Texas, USA, Proceedings*, pages 2682–2685. IEEE, 2010.

[14] Hidenori Kuwakado and Masakatu Morii. Security on the quantum-type even-mansour cipher. In *Proceedings of the International Symposium on Information Theory and its Applications, ISITA 2012, Honolulu, HI, USA, October 28-31, 2012*, pages 312–316. IEEE, 2012.

[15] David A. McGrew and Scott R. Fluhrer. The extended codebook (XCB) mode of operation. *IACR Cryptol. ePrint Arch.*, 2004:278, 2004.

[16] David A. McGrew and Scott R. Fluhrer. The security of the extended codebook (XCB) mode of operation. In Carlisle M. Adams, Ali Miri, and Michael J. Wiener, editors, *Selected Areas in Cryptography, 14th International Workshop, SAC 2007, Ottawa, Canada, August 16-17, 2007, Revised Selected Papers*, volume 4876 of *Lecture Notes in Computer Science*, pages 311–327. Springer, 2007.

[17] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997.

[18] Daniel R. Simon. On the power of quantum computation. *SIAM J. Comput.*, 26(5):1474–1483, 1997.