Post-Quantum Security of the (Tweakable) FX Construction, and Applications

Gorjan Alagic[1], Chen Bai[2], Jonathan Katz[3], Christian Majenz[4], and Patrick Struck[5]

[1] QuICS, University of Maryland, and NIST
galagic@umd.edu
[2] Dept. of Electrical and Computer Engineering, University of Maryland
cbai1@umd.edu
[3] Dept. of Computer Science, University of Maryland
jkatz2@gmail.com
[4] Dept. of Applied Mathematics and Computer Science, Technical University of Denmark
chmaj@dtu.dk
[5] Faculty of Informatics and Data Science, University of Regensburg
patrick.struck@ur.de

Abstract. The FX construction provides a way to increase the effective key length of a block cipher E. We prove security of a tweakable version of the FX construction in the post-quantum setting, i.e., against a quantum attacker given only *classical* access to the secretly keyed construction while retaining *quantum* access to E. We then use our results to prove post-quantum security—in the same model—of the (plain) FX construction, Elephant (a finalist of NIST's lightweight cryptography standardization effort), and Chaskey (an ISO-standardized lightweight MAC).

1 Introduction

The development of large-scale quantum computers would have a significant impact on cryptography. For example, for symmetric-key cryptosystems— even ideal ciphers—one must at least double the key length in order to achieve the same security against quantum attackers as is enjoyed against classical adversaries, due to the possibility of using Grover's search algorithm [7] to carry out a key-recovery attack. In general, however, doubling the key length may not be sufficient [11,12,4], and it is therefore critical to understand the security of various symmetric-key constructions against quantum attackers.

One can consider two models of quantum attacks [3]. In the so-called Q2 model, the attacker is given quantum access to any underlying public primitives (e.g., a block cipher) as well as the secretly keyed construction itself. In contrast, the Q1 model assumes the adversary has quantum access to all public primitives but only *classical* access to the secretly keyed construction. The distinction between Q1 and Q2 is significant: for many symmetric-key constructions, polynomial-query attacks are known in the Q2 model [11,12,9] but not in the Q1 model. At the same time, however, the Q2 model appears to be highly unrealistic, particularly for real-world applications where the honest parties only run the construction on classical inputs, and do not expose any quantum interface to an attacker (which is necessarily the case whenever the honest devices implementing the construction are entirely classical). The Q1 model is a much better fit for realistic quantum attacks—a quantum adversary would easily have quantum access to any public primitives—and, indeed, recent work [8,1,4] has focused on that model. From here on, by "post-quantum security" we will mean the Q1 model by default.

Proving security in the Q1 model (without just assuming the stronger Q2 model) is challenging since it requires reasoning about a combination of classical and quantum oracles. Indeed, there are at present only a limited number of positive results about security in this model. Jaeger et al. [8] recently showed partial positive results for the FX construction, which provides a mechanism for key-length extension of an ideal cipher; their results imply security either for non-adaptive adversaries or for a variant of the FX construction using a public keyed *function* in place of a public keyed *permutation*. The FX construction degenerates to the Even-Mansour scheme [5] when the public primitive is unkeyed, and so their work also implies security for the Even-Mansour construction either for non-adaptive adversaries or for a variant of the construction based on a public random function. Subsequent work by Alagic et al. [1] was able to show post-quantum security of the full Even-Mansour construction (i.e., using a random permutation) against adaptive adversaries. However, their work left open the question of extending this result to the FX construction.

1.1 Our Results

Let $E : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher that we will treat in our analysis as ideal. We consider a tweakable version of the FX construction $\mathsf{TFX}^{f_1,f_2}[E] : (\{0,1\}^m \times \{0,1\}^\kappa) \times (\mathcal{T} \times \{0,1\}^n) \to \{0,1\}^n$, defined as

$$\mathsf{TFX}^{f_1,f_2}_{k,k'}[E](t,x) = E_k(x \oplus f_1(t,k')) \oplus f_2(t,k')\,,$$

where $\mathcal{T}$ is a tweak space, $\kappa \geq n$, and f_1, f_2 are functions satisfying some technical conditions we omit here. As our main result, we prove that the above is a secure (post-quantum) tweakable block cipher. Concretely (cf. Theorem 1), we show that an adaptive adversary making q_C classical queries to $\mathsf{TFX}^{f_1,f_2}_{k,k'}[E]$ and q_Q quantum queries to E (where we allow queries in both the forward and inverse directions) can distinguish the former from an ideal tweakable cipher only with probability $O\left(2^{-(m+n)/2} \cdot (q_C\sqrt{q_Q} + q_Q\sqrt{q_C})\right)$. A key building block of our result is a generalization of existing "resampling lemmas" [6,1] to cover ideal ciphers (cf. Lemma 2), something that may be of independent interest.

We use our result to derive various corollaries regarding the post-quantum security of other symmetric-key constructions:

1. By taking $\kappa = 2n$, $\mathcal{T} = \emptyset$, $f_1(k_1, k_2) = k_1$, and $f_2(k_1, k_2) = k_2$, the TFX construction degenerates to the FX construction. Our result thus implies

post-quantum security of the full FX construction against adaptive adversaries, answering the open question from Jaeger et al. [8].
2. If we take $m = 0$ (so E is now a public random permutation) and choose the tweak space $\mathcal{T}$ and the functions f_1, f_2 appropriately, TFX becomes the tweakable block cipher at the core of (a slightly simplified variant of) Elephant [2], a lightweight authenticated encryption scheme currently under consideration for standardization by NIST [14]. Our main result implies post-quantum security for this variant of Elephant.
3. Taking $m = 0$ again, we can set $\mathcal{T}$, f_1, f_2 such that TFX captures the three pseudorandom permutations used in Chaskey [13], a lightweight MAC that is an ISO standard. We thus prove post-quantum security of Chaskey.

To our knowledge, these are the first proofs of security for any versions of Elephant and Chaskey against quantum adversaries.

Paper organization. In Section 2, we establish some notation, recall a "reprogramming lemma" from prior work [1], and establish a "resampling lemma" for the ideal-cipher model that will be useful for proving our main result. We introduce the tweakable FX construction, and prove it secure in the post-quantum setting, in Section 3. Finally, in Section 4 we describe the applications of our main result to the post-quantum security of FX, Elephant, and Chaskey.

2 Preliminaries

Notation and basic definitions. We let $\mathcal{P}(n)$ denote the set of all permutations on $\{0,1\}^n$. In the *public permutation model*, a permutation $P \leftarrow \mathcal{P}(n)$ is sampled uniformly and then provided as an oracle (in both the forward and inverse directions) to all parties. A block cipher $E : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$ is a keyed permutation, i.e., $E_k(\cdot) = E(k, \cdot)$ is a permutation of $\{0,1\}^n$ for all $k \in \{0,1\}^m$. We say E is a *pseudorandom permutation* if E_k (for uniform $k \in \{0,1\}^m$) is indistinguishable from a uniform permutation in $\mathcal{P}(n)$, where indistinguishability is required to hold even when the adversary may query its oracle in both the forward and inverse directions.

For a set S, let $\mathcal{E}(S, n)$ be the set of all functions $E : S \times \{0,1\}^n \to \{0,1\}^n$ such that E_k is a permutation on $\{0,1\}^n$ for all $k \in S$. When $S = \{0,1\}^m$ we also write $\mathcal{E}(m, n)$. In the *ideal-cipher model* a cipher $E \leftarrow \mathcal{E}(m, n)$ is sampled uniformly and then provided as an oracle, in both the forward and inverse directions, to all parties. (When $m = 0$ this defaults to the public permutation model.) A tweakable block cipher $\tilde{E} : \{0,1\}^m \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ is a family of permutations indexed by both a key $k \in \{0,1\}^m$ and a tweak $t \in \mathcal{T}$, i.e., we now require that $\tilde{E}_k(t, \cdot) = \tilde{E}(k, t, \cdot)$ is a permutation of $\{0,1\}^n$ for all $k \in \{0,1\}^m$ and $t \in \mathcal{T}$. A tweakable block cipher $\tilde{E}$ is *secure* if $\tilde{E}_k$ (for uniform choice of $k \in \{0,1\}^m$) is indistinguishable from a uniform $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$.

In all the security notions mentioned above we consider algorithms having only classical access to secretly keyed primitives. When we consider constructions of keyed primitives (e.g., a tweakable block cipher) from ideal public primitives (e.g., an ideal cipher), however, we provide the distinguisher with *quantum* oracle access to the public primitive. Thus, for example, a distinguisher in the ideal-cipher model has the ability to apply the unitary operators

$$|k\rangle|x\rangle|y\rangle \mapsto |k\rangle|x\rangle|E_k(x) \oplus y\rangle$$
$$|k\rangle|x\rangle|y\rangle \mapsto |k\rangle|x\rangle|E_k^{-1}(x) \oplus y\rangle$$

to quantum registers of the adversary's choice. (We emphasize that this includes evaluating E/E^{-1} on arbitrary superpositions of both keys and inputs.) This is well-motivated, as in practice E would be instantiated by a public block cipher; adversaries with quantum computers would thus be able to coherently execute the reversible circuit for computing E. On the other hand, as discussed in the introduction, secretly keyed primitives would be implemented by honest parties; if they only evaluate the primitive on classical inputs then the attacker has no way to obtain quantum access to that primitive.

A reprogramming lemma. For a function $F : \{0,1\}^m \to \{0,1\}^n$ and a set $B \subset \{0,1\}^m \times \{0,1\}^n$ such that each $x \in \{0,1\}^m$ is the first element of at most one tuple in B, define

$$F^{(B)}(x) := \begin{cases} y & \text{if } (x,y) \in B \\ F(x) & \text{otherwise.} \end{cases}$$

We rely on the following lemma, taken verbatim from [1]:

Lemma 1. *Let $\mathcal{D}$ be a distinguisher in the following experiment:*

Phase 1: *$\mathcal{D}$ outputs descriptions of a function $F_0 = F : \{0,1\}^m \to \{0,1\}^n$ and a randomized algorithm $\mathcal{B}$ whose output is a set $B \subset \{0,1\}^m \times \{0,1\}^n$ where each $x \in \{0,1\}^m$ is the first element of at most one tuple in B. Let $B_1 = \{x \mid \exists y : (x,y) \in B\}$ and $\epsilon = \max_{x\in\{0,1\}^m} \{\Pr_{B\leftarrow\mathcal{B}}[x \in B_1]\}$.*

Phase 2: *$\mathcal{B}$ is run to obtain B. Let $F_1 = F^{(B)}$. A uniform bit b is chosen, and $\mathcal{D}$ is given quantum access to F_b.*

Phase 3: *$\mathcal{D}$ loses access to F_b, and receives the randomness r used to invoke $\mathcal{B}$ in phase 2. Then $\mathcal{D}$ outputs a guess b'.*

For any $\mathcal{D}$ making q queries in expectation when its oracle is F_0, it holds that

$$|\Pr[\mathcal{D} \text{ outputs } 1 \mid b=1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b=0]| \leq 2q \cdot \sqrt{\epsilon}\,.$$

A resampling lemma for ideal ciphers. As a building block for our main result, we prove a resampling lemma for ideal ciphers that generalizes earlier results for random functions [6] and permutations [1]. We consider the experiment in which a distinguisher $\mathcal{D}$ is first given quantum access to an ideal cipher $E : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$. Then, a key $k_0 \in \{0,1\}^m$ and two points $s_0, s_1 \in \{0,1\}^n$ are chosen according to some distribution, and in a second phase $\mathcal{D}$ is given access either to the original function $E^{(0)} = E$ or a modified function $E^{(1)}$ that is the same as E except that the values of $E_{k_0}(s_0)$ and $E_{k_0}(s_1)$

are swapped. (See below for details.) We show, roughly speaking, that so long as the distribution of k_0, s_0, s_1 has high min-entropy and $\mathcal{D}$ makes only a bounded number of queries in the first phase of its execution, $\mathcal{D}$ cannot distinguish these two possibilities. Compared to prior work of Alagic et al. [1], our proof handles the case $m > 0$ (i.e., ideal ciphers and not just random permutations) and also allows for distributions over k_0, s_0, s_1 other than the uniform distribution.

For $s_0, s_1 \in \{0,1\}^n$, define $\mathsf{swap}_{s_0,\,s_1} \in \mathcal{P}(n)$ as

$$\mathsf{swap}_{s_0,\,s_1}(x) = \begin{cases} s_1 & \text{if } x = s_0 \\ s_0 & \text{if } x = s_1 \\ x & \text{otherwise.} \end{cases}$$

Lemma 2 (Ideal-cipher resampling). *Fix a probability distribution D on $\{0,1\}^{m+2n}$, and let*

$$\epsilon = \max_{\substack{k^* \in \{0,1\}^m \\ s^* \in \{0,1\}^n}} \Pr_{(k,s_0,s_1) \sim D}[(k^*, s^*) \in \{(k, s_0), (k, s_1)\}].$$

Consider the following experiment involving a distinguisher $\mathcal{D}$:

Phase 1: *Choose uniform $E \in \mathcal{E}(m,n)$, and give $\mathcal{D}$ quantum access to E.*
Phase 2: *Choose $k \in \{0,1\}^m$ and $s_0, s_1 \in \{0,1\}^n$ according to D. Let $E^{(0)} = E$ and define $E^{(1)} : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$ by*

$$E^{(1)}_{k^*}(x) = \begin{cases} E_{k^*}(x) & \text{if } k^* \neq k \\ E_{k^*} \circ \mathsf{swap}_{s_0,\,s_1}(x) & \text{if } k^* = k\,. \end{cases}$$

A uniform bit $b \in \{0,1\}$ is chosen, and $\mathcal{D}$ is given k, s_0, s_1, and quantum access to $E^{(b)}$. Then $\mathcal{D}$ outputs a guess b'.

For any $\mathcal{D}$ making at most q queries to E in phase 1,

$$|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]| \leq 2\sqrt{2q\epsilon}\,.$$

The proof is given in Appendix A.

3 Post-Quantum Security of Tweakable FX

Let $E : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher and $\mathcal{T}$ a finite set, and fix two functions $f_1, f_2 : \mathcal{T} \times \{0,1\}^\kappa \to \{0,1\}^n$. We consider a tweakable version of the FX construction $\mathsf{TFX}^{f_1,f_2}[E] : (\{0,1\}^m \times \{0,1\}^\kappa) \times (\mathcal{T} \times \{0,1\}^n) \to \{0,1\}^n$ defined as

$$\mathsf{TFX}^{f_1,f_2}_{k,k'}[E](t,x) = E_k(x \oplus f_1(t,k')) \oplus f_2(t,k')\,.$$

We require the tweak functions f_1, f_2 to satisfy some structural properties:

Definition 1. *We say that $f : \mathcal{T} \times \{0,1\}^\kappa \to \{0,1\}^n$ is* ***proper*** *with respect to $\mathcal{T}$ if it satisfies the following two properties:*

Uniformity: *For all $t \in \mathcal{T}$ and all $y \in \{0,1\}^n$,*

$$\mathrm{Pr}_{k' \leftarrow \{0,1\}^\kappa}[f(t,k') = y] = 2^{-n}.$$

XOR-uniformity: *For all distinct $t, t' \in \mathcal{T}$ and all $y \in \{0,1\}^n$,*

$$\mathrm{Pr}_{k' \leftarrow \{0,1\}^\kappa}[f(t,k') \oplus f(t',k') = y] = 2^{-n}.$$

Note uniformity implies $f(t,\cdot)$ is surjective for any $t \in \mathcal{T}$, and $\kappa \geq n$.

Theorem 1. *Let TFX be as above and let $\mathcal{A}$ be an adversary making q_C classical queries to its first oracle and q_Q quantum queries to its second oracle. Then if f_1, f_2 are proper with respect to $\mathcal{T}$, it holds that*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^m; k' \leftarrow \{0,1\}^\kappa; \\ E \leftarrow \mathcal{E}(m,n)}} \left[\mathcal{A}^{\mathsf{TFX}^{f_1,f_2}_{k,k'}[E], E} = 1 \right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T},n); \\ E \leftarrow \mathcal{E}(m,n)}} \left[\mathcal{A}^{\tilde{E}, E} = 1 \right] \right|$$
$$\leq (3 + 2\sqrt{2}) \cdot 2^{-(m+n)/2} \cdot \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right).$$

Our high-level proof is similar to the proof of security for the Even-Mansour construction by Alagic et al. [1]. However, our proof of Lemma 4 differs substantially from the proof of the corresponding lemma in their work. In particular, by modifying the sequence of hybrid experiments, we are able to avoid a certain "bad event" whose probability is difficult to compute in our setting.

Proof. As noted, the high-level structure of our proof is similar to the proof of security for the Even-Mansour construction by Alagic et al. [1]; for that reason, we copy some portions of their proof (with appropriate updates for our setting).

Without loss of generality, we assume $\mathcal{A}$ never makes a redundant classical query; that is, once it learns an input/output pair (x, y) associated with some tweak t by making a query to its classical oracle, it never again submits the query (t, x) (resp., (t, y)) to the forward (resp., inverse) direction of that oracle. We divide an execution of $\mathcal{A}$ into $q_C + 1$ stages $0, \ldots, q_C$, where the jth stage corresponds to the time between the jth and $(j+1)$st classical queries of $\mathcal{A}$. (The 0th stage is the period of time before $\mathcal{A}$ makes its first classical query, and the q_Cth stage is the period of time after $\mathcal{A}$ makes its last classical query.) $\mathcal{A}$ may adaptively distribute its q_Q quantum queries between these stages arbitrarily, and we let $q_{Q,j}$ be the expected number of quantum queries that $\mathcal{A}^{\tilde{E},E}$ makes in the jth stage. (This probability is taken over $\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n)$ and $E \leftarrow \mathcal{E}(m,n)$ and any internal randomness/measurements of $\mathcal{A}$.) Note that $\sum_{j=0}^{q_C} q_{Q,j} = q_Q$.

We write K to stand for (k, k'). Since f_1, f_2 are fixed, we write TFX_K in place of $\mathsf{TFX}^{f_1,f_2}_{k,k'}$. In a given execution of $\mathcal{A}$, we denote its ith classical query by the tuple (t_i, x_i, y_i, b_i), where $t_i \in \mathcal{T}$ is the tweak, $(x_i, y_i) \in \{0,1\}^n \times \{0,1\}^n$ is the input/output pair, and $b_i \in \{0,1\}$ indicates the query direction, i.e., $b_i = 0$ (resp., $b_i = 1$) means that the ith classical query was in the forward (resp., inverse) direction. We let $T_j = \big((t_1, x_1, y_1, b_1), \ldots, (t_j, x_j, y_j, b_j)\big)$ be the ordered list consisting of the first j queries of $\mathcal{A}$.

Our proof involves a sequence of experiments in which $\mathcal{A}$'s oracles are modified based on the classical queries made by $\mathcal{A}$ thus far. We first establish the appropriate notation.

We use the product symbol $\prod$ to denote sequential composition of operations, i.e., $\prod_{i=1}^{n} f_i = f_1 \circ \cdots \circ f_n$. (Note that order matters, since function composition is not commutative in general.) For an ideal cipher E, a key $K = (k, k')$, and a list $T_j = \big((t_1, x_1, y_1, b_1), \ldots, (t_j, x_j, y_j, b_j)\big)$ as above, define the operators

$$\overrightarrow{S}_{T_j,E,K} = \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{E_k(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')}$$

$$\overrightarrow{Q}_{T_j,E,K} = \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{x_i \oplus f_1(t_i,k'),\, E_k^{-1}(y_i \oplus f_2(t_i,k'))}$$

$$\overleftarrow{S}_{T_j,E,K} = \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{E_k(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')}$$

$$\overleftarrow{Q}_{T_j,E,K} = \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{x_i \oplus f_1(t_i,k'),\, E_k^{-1}(y_i \oplus f_2(t_i,k'))}$$

where, as usual, f^0 is the identity map and $f^1 = f$ for any function f. We define the modified cipher $E^{T_j,K}$ as

$$E^{T_j,K}_{k^*}(x) = \begin{cases} E_{k^*}(x) & k^* \neq k \\ \overleftarrow{S}_{T_j,E,K} \circ \overrightarrow{S}_{T_j,E,K} \circ E_k(x) & k^* = k. \end{cases} \tag{1}$$

Since $E_k \circ \mathsf{swap}_{x,\,y} = \mathsf{swap}_{E_k(x),\, E_k(y)} \circ E_k$, we have

$$\overleftarrow{S}_{j,E,K} \circ \overrightarrow{S}_{T_j,E,K} \circ E_k = \overleftarrow{S}_{T_j,E,K} \circ E_k \circ \overrightarrow{Q}_{T_j,E,K} = E_k \circ \overleftarrow{Q}_{T_j,E,K} \circ \overrightarrow{Q}_{T_j,E,K}\,.$$

Roughly speaking, $E^{T_j,K}$ is the minimal modification of E that is consistent with the forward ($\rightarrow$) and backward ($\leftarrow$) queries from the transcript T_j when pre-composed (S) or post-composed (Q) with E. For compactness we occasionally write E^j in place of $E^{T_j,K}$ when T_j and K are understood from the context.

We now define a sequence of hybrid experiments $\mathbf{H}_j$, for $j = 0, \ldots, q_C$.

Experiment $\mathbf{H}_j$. Sample uniform ciphers $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $E \in \mathcal{E}(m, n)$, and a uniform key $K \in \{0,1\}^m \times \{0,1\}^\kappa$. Then:

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using E, stopping immediately *before* its $(j+1)$st classical query. Let $T_j = \big((t_1, x_1, y_1, b_1), \ldots, (t_j, x_j, y_j, b_j)\big)$ be the list of classical queries so far.
2. For the remainder of the execution of $\mathcal{A}$, answer its classical queries using $\mathsf{TFX}_K[E^{T_j,K}]$ and its quantum queries using $E^{T_j,K}$.

We can compactly represent $\mathbf{H}_j$ as the experiment in which $\mathcal{A}$'s queries are answered using the oracle sequence

$$\underbrace{E, \tilde{E}, E, \cdots, \tilde{E}, E,}_{j \text{ classical queries}} \underbrace{\mathsf{TFX}_K[E^j], E^j, \cdots, \mathsf{TFX}_K[E^j], E^j}_{q_C - j \text{ classical queries}}\,.$$

Each instance of $\tilde{E}$ or $\mathsf{TFX}_K[E^j]$ represents a single classical query, while each instance of E or E^j represents a stage during which $\mathcal{A}$ makes multiple quantum queries to that oracle but no queries to its classical oracle. Observe that $\mathbf{H}_0$ corresponds to the execution of $\mathcal{A}$ in the real world, i.e., $\mathcal{A}^{\mathsf{TFX}_K[E],E}$, and $\mathbf{H}_{q_C}$ is the execution of $\mathcal{A}$ in the ideal world, i.e., $\mathcal{A}^{\tilde{E},E}$.

For $j = 0, \ldots, q_C - 1$, we introduce additional experiments $\mathbf{H}'_j$:

Experiment $\mathbf{H}'_j$. Sample uniform ciphers $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $E \in \mathcal{E}(m, n)$, and uniform $K \in \{0,1\}^m \times \{0,1\}^\kappa$. Then:

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using E, stopping immediately *after* its $(j+1)$st classical query. Let $T_{j+1} = \big((t_1, x_1, y_1, b_1), \ldots, (t_{j+1}, x_{j+1}, y_{j+1}, b_{j+1})\big)$ be the classical queries so far.
2. For the remainder of the execution of $\mathcal{A}$, answer its classical queries using $\mathsf{TFX}_K[E^{T_{j+1},K}]$ and its quantum queries using $E^{T_{j+1},K}$.

Thus, $\mathbf{H}'_j$ corresponds to running $\mathcal{A}$ using the oracle sequence

$$\underbrace{E, \tilde{E}, E, \cdots, \tilde{E}, E,}_{j \text{ classical queries}} \tilde{E}, E^{j+1}, \underbrace{\mathsf{TFX}_K[E^{j+1}], E^{j+1} \cdots, \mathsf{TFX}_K[E^{j+1}], E^{j+1}}_{q_C - j - 1 \text{ classical queries}}.$$

In Lemmas 3 and 4, we establish the following bounds on the distinguishability of $\mathbf{H}'_j$ and $\mathbf{H}_{j+1}$, as well as $\mathbf{H}_j$ and $\mathbf{H}'_j$, for $0 \le j < q_C$:

$$\big|\Pr[\mathcal{A}(\mathbf{H}'_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}_{j+1}) = 1]\big| \le 2 \cdot q_{Q,j+1} \cdot \sqrt{\frac{2 \cdot (j+1)}{2^{m+n}}}.$$

$$\big|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\big| \le 2\sqrt{2} \cdot \sqrt{\frac{q_Q}{2^{m+n}}} + 3j \cdot 2^{-n}.$$

Using the above, we have

$$\begin{aligned}
&|\Pr[\mathcal{A}(\mathbf{H}_0) = 1] - \Pr[\mathcal{A}(\mathbf{H}_{q_C}) = 1]| \\
&\le \sum_{j=0}^{q_C-1} \left(2\sqrt{2} \cdot \sqrt{\frac{q_Q}{2^{m+n}}} + 3j \cdot 2^{-n} + 2 \cdot q_{Q,j+1}\sqrt{\frac{2 \cdot (j+1)}{2^{m+n}}}\right) \\
&\le 3q_C^2 \cdot 2^{-n} + \sum_{j=0}^{q_C-1} \left(2\sqrt{2} \cdot \sqrt{\frac{q_Q}{2^{m+n}}} + 2 \cdot q_{Q,j+1}\sqrt{\frac{2q_C}{2^{m+n}}}\right) \\
&\le 3q_C^2 \cdot 2^{-n} + 2^{-(m+n)/2} \cdot \left(2\sqrt{2} q_C \sqrt{q_Q} + 2\sqrt{2} \cdot q_Q \sqrt{q_C}\right).
\end{aligned}$$

An easy argument finishes the proof (see [1] for details). □

We now prove Lemmas 3 and 4.

Lemma 3. *For $j = 0, \ldots, q_C - 1$,*

$$\Pr[\mathcal{A}(\mathbf{H}'_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}_{j+1}) = 1]| \le 2 \cdot q_{Q,j+1}\sqrt{2 \cdot (j+1)/2^{m+n}},$$

where $q_{Q,j+1}$ is the expected number of queries $\mathcal{A}$ makes to E in the $(j+1)$st stage in the ideal world (i.e., in $\mathbf{H}_{q_C}$.)

Proof. Let $\mathcal{A}$ be a distinguisher between $\mathbf{H}'_j$ and $\mathbf{H}_{j+1}$. We construct from $\mathcal{A}$ a distinguisher $\mathcal{D}$ for the experiment from Lemma 1:

Phase 1: $\mathcal{D}$ samples uniform $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $E \in \mathcal{E}(m, n)$. It then runs $\mathcal{A}$, answering its quantum queries using E and its classical queries using $\tilde{E}$, until after it responds to $\mathcal{A}$'s $(j+1)$st classical query. Let $T_{j+1} = \big((t_1, x_1, y_1, b_1), \ldots, (t_{j+1}, x_{j+1}, y_{j+1}, b_{j+1})\big)$ be the list of input/output pairs $\mathcal{A}$ received from its classical oracle thus far. $\mathcal{D}$ defines $F(a, k^*, x) := E^a_{k*}(x)$ for $a \in \{1, -1\}$. It also defines the following randomized algorithm $\mathcal{B}$: sample $K \leftarrow \{0,1\}^m \times \{0,1\}^\kappa$ and then compute the set B of input/output pairs to be reprogrammed so that $F^{(B)}(a, k^*, x) = (E^{T_{j+1},K}_{k^*})^a(x)$ for all a, k^*, x.

Phase 2: $\mathcal{B}$ is run to generate B, and $\mathcal{D}$ is given quantum access to an oracle F_b. $\mathcal{D}$ resumes running $\mathcal{A}$, answering its quantum queries using F_b. Phase 2 ends when $\mathcal{A}$ makes its next (i.e., $(j+2)$nd) classical query.

Phase 3: $\mathcal{D}$ is given the randomness used by $\mathcal{B}$ to generate K. It resumes running $\mathcal{A}$, answering its classical queries using $\mathsf{TFX}_K[E^{T_{j+1},K}]$ and its quantum queries using $E^{T_{j+1},K}$. Finally, it outputs whatever $\mathcal{A}$ outputs.

It is immediate that if $b = 0$ (i.e., $\mathcal{D}$'s oracle in phase 2 is $F_0 = F$), then $\mathcal{A}$'s output is identically distributed to its output in $\mathbf{H}_{j+1}$, whereas if $b = 1$ (i.e., $\mathcal{D}$'s oracle in phase 2 is $F_1 = F^{(B)}$), then $\mathcal{A}$'s output is identically distributed to its output in $\mathbf{H}'_j$. It follows that $|\Pr[\mathcal{A}(\mathbf{H}'_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}_{j+1}) = 1]|$ is equal to the distinguishing advantage of $\mathcal{D}$ in the reprogramming experiment of Lemma 1. To bound this quantity, we bound the parameter ϵ and the expected number of queries made by $\mathcal{D}$ in phase 2 (when $F = F_0$.)

The value of ϵ can be bounded using the definition of $E^{T_{j+1},K}$ and the fact that $F^{(B)}(a, k^*, x) = (E^{T_{j+1},K}_{k^*})^a(x)$. Fixing E and T_{j+1}, the probability that any particular input (a, k^*, x) is reprogrammed is at most the probability (over K) that it is in the set

$$\left\{ \begin{array}{c} (1, k, x_i \oplus f_1(t_i \oplus k')),\ (1, k, E^{-1}_k(y_i \oplus f_2(t_i \oplus k'))), \\ (-1, k, E_k(x_i \oplus f_1(t_i \oplus k'))),\ (-1, k, y_i \oplus f_2(t_i \oplus k')) \end{array} \right\}_{i=1}^{j+1} .$$

Since both $f_1(t_i \oplus k')$ and $f_2(t_i \oplus k')$ are uniform (by uniformity of f_1, f_2), taking a union bound gives $\epsilon \leq 2(j+1)/2^{m+n}$.

The expected number of queries made by $\mathcal{D}$ in phase 2 when $F = F_0$ is equal to the expected number of queries made by $\mathcal{A}$ in its $(j+1)$st stage in $\mathbf{H}_{j+1}$. Since $\mathbf{H}_{j+1}$ and $\mathbf{H}_{q_E}$ are identical until after the $(j+1)$st stage is complete, this is precisely $q_{Q,j+1}$. □

Lemma 4. *For $j = 0, \ldots, q_C$,*

$$\left|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right| \leq 2\sqrt{2} \cdot \sqrt{\frac{q_Q}{2^{(m+n)}}} + \frac{3j}{2^n} .$$

Proof. We first introduce additional experiments $\mathbf{H}^*_j$ and $\mathbf{H}^{**}_j$.

Experiment $\mathbf{H}^*_j$. Sample uniform $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $E \in \mathcal{E}(m, n)$, and uniform $K = (k, k') \in \{0,1\}^m \times \{0,1\}^\kappa$. Then

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using E, until $\mathcal{A}$ makes its $(j+1)$st classical query $(t_{j+1}, x_{j+1}, b_{j+1} = 0)$, which we assume for concreteness to be in the forward direction.[1]
2. Choose uniform $s \in \{0,1\}^n$, and define $E^{(1)}$ as

$$E^{(1)}_{k^*}(x) = \begin{cases} E_{k^*}(x) & \text{if } k^* \neq k \\ \left(E_k \circ \mathsf{swap}_{f_1(t_{j+1},k') \oplus x_{j+1},\, s}\right)(x) & \text{if } k^* = k. \end{cases}$$

Continue running $\mathcal{A}$, answering its remaining classical queries (including the $(j+1)$st) using $\mathsf{TFX}_K[(E^{(1)})^{T_j,K}]$, and its quantum queries using $(E^{(1)})^{T_j,K}$.

Experiment $\mathbf{H}^{**}_j$ is the same as $\mathbf{H}^*_j$, except that the $(j+1)$st query is answered using $\tilde{E}$. Thus we can write $\mathbf{H}^*_j$ and $\mathbf{H}^{**}_j$ as the following oracle sequences:

$$\begin{array}{ll} \mathbf{H}^*_j: & E, \tilde{E}, E, \cdots, \tilde{E}, E, \quad \mathsf{TFX}_K[(E^{(1)})^j], (E^{(1)})^j, \cdots, \mathsf{TFX}_K[(E^{(1)})^j], (E^{(1)})^j \\ \mathbf{H}^{**}_j: & \underbrace{E, \tilde{E}, E, \cdots, \tilde{E}, E,}_{j \text{ classical queries}} \quad \underbrace{\qquad \tilde{E} \qquad, (E^{(1)})^j, \cdots, \mathsf{TFX}_K[(E^{(1)})^j], (E^{(1)})^j}_{q_C - j \text{ classical queries}} \end{array}$$

where, recall, we let $(E^{(1)})^j$ denote $(E^{(1)})^{T_j,K}$. We have

$$\begin{aligned} \left|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right| \leq & \left|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}^*_j) = 1]\right| \\ & + \left|\Pr[\mathcal{A}(\mathbf{H}^*_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}^{**}_j) = 1]\right| \\ & + \left|\Pr[\mathcal{A}(\mathbf{H}^{**}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right|, \end{aligned}$$

and we now bound the three differences on the right-hand side.

Let $\mathcal{A}$ be a distinguisher between $\mathbf{H}_j$ and $\mathbf{H}^*_j$. We construct from $\mathcal{A}$ a distinguisher $\mathcal{D}$ for the resampling experiment of Lemma 2. Fix D to be the uniform distribution over $\{0,1\}^{m+n}$ (so $\epsilon = 2^{-(m+n)}$ in Lemma 2). $\mathcal{D}$ does:

Phase 1: $\mathcal{D}$ is given quantum access to an ideal cipher E. It samples a uniform $\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n)$ and then runs $\mathcal{A}$, answering its quantum queries with E and its classical queries with $\tilde{E}$ (in the appropriate directions), until $\mathcal{A}$ submits its $(j+1)$st classical query $(t_{j+1}, x_{j+1}, b_{j+1} = 0)$. At that point, $\mathcal{D}$ has a list $T_j = \big((t_1, x_1, y_1, b_1), \cdots, (t_j, x_j, y_j, b_j)\big)$ of the queries/answers $\mathcal{A}$ has made to its classical oracle thus far.

Phase 2: $\mathcal{D}$ is given uniform $s_0, s_1 \in \{0,1\}^n$, $k \in \{0,1\}^m$, and quantum oracle access to a cipher $E^{(b)}$. $\mathcal{D}$ samples a uniform $k' \in \{0,1\}^\kappa$ conditioned on $f_1(t_{j+1}, k') = s_0 \oplus x_{j+1}$ (at least one such k' must exist since f_1 is surjective) and sets $K := (k, k')$. It then continues running $\mathcal{A}$, answering its remaining classical queries (including the $(j+1)$st) using $\mathsf{TFX}_K[(E^{(b)})^{T_j,K}]$, and its remaining quantum queries using $(E^{(b)})^{T_j,K}$. $\mathcal{D}$ outputs whatever $\mathcal{A}$ does.

Note that in phase 1, distinguisher $\mathcal{D}$ perfectly simulates experiments $\mathbf{H}_j$ and $\mathbf{H}^*_j$ for $\mathcal{A}$ until the point where $\mathcal{A}$ makes its $(j+1)$st classical query. If $b = 0$,

[1] As in [1], the case of an inverse query is entirely symmetric.

$\mathcal{D}$ gets access to $E^{(0)} = E$ in phase 2. Since $\mathcal{D}$ answers all quantum queries using $(E^{(0)})^{T_j,K}$ and all classical queries using $\mathsf{TFX}_K[(E^{(0)})^{T_j,K}]$, we see that $\mathcal{D}$ perfectly simulates $\mathbf{H}_j$ for $\mathcal{A}$ in that case. If, on the other hand, $b = 1$ in phase 2, then $\mathcal{D}$ gets access to $E^{(1)}$, where

$$E^{(1)}_{k^*}(x) = \begin{cases} E_{k^*}(x) & \text{if } k^* \neq k \\ E_k \circ \mathsf{swap}_{s_0,\, s_1}(x) & \text{if } k^* = k\,. \end{cases}$$

Since $f_1(t_{j+1}, k') \coloneqq s_0 \oplus x_{j+1}$, it holds that

$$E^{(1)}_{k^*}(x) = \begin{cases} E_{k^*}(x) & \text{if } k^* \neq k \\ E_k \circ \mathsf{swap}_{f_1(t_{j+1},k')\oplus x_{j+1},\, s_1}(x) & \text{if } k^* = k\,. \end{cases}$$

Moreover, the uniformity property of f_1 and the fact that s_0 (and hence $s_0 \oplus x_{j+1}$) is uniform imply that the joint distribution of k' and $s_0 \oplus x_{j+1}$ is equal to the joint distribution of $\tilde{k}$ and $f_1(t_{j+1}, \tilde{k})$ for a uniform $\tilde{k}$. Thus, in this case $\mathcal{D}$ perfectly simulates $\mathbf{H}^*_j$ for $\mathcal{A}$. Applying Lemma 2 thus gives

$$\left|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}^*_j) = 1]\right| \leq 2\sqrt{2q_Q \cdot \epsilon} \leq 2\sqrt{\frac{2q_Q}{2^{n+m}}}\,. \tag{2}$$

Next, we bound the distinguishability of $\mathbf{H}^*_j$ and $\mathbf{H}^{**}_j$. Recall they differ in that in $\mathbf{H}^*_j$ the $(j+1)$st query is answered with $\mathsf{TFX}_K[(E^{(1)})^{T_j,K}](x_{j+1})$, whereas in $\mathbf{H}^{**}_j$ that query is answered with $\tilde{E}_{t_{j+1}}(x_{j+1})$. In $\mathbf{H}^*_j$ we have

$$\begin{aligned} y_{j+1} &\stackrel{\text{def}}{=} \mathsf{TFX}_K[(E^{(1)})^{T_j,K}](t_{j+1}, x_{j+1}) \\ &= (E^{(1)}_k)^{T_j,K}(x_{j+1} \oplus f_1(t_{j+1}, k')) \oplus f_2(t_{j+1}, k') \\ &= E^{T_j,K}_k(s) \oplus f_2(t_{j+1}, k')\,; \end{aligned}$$

uniformity of f_2 implies that y_{j+1} is uniform. This is not identical to the distribution of y_{j+1} in $\mathbf{H}^{**}_j$, which is uniform subject to the constraint that $\tilde{E}_{t_{j+1}}$ is a permutation. Define the set $\mathcal{Y}_{j+1} = \{y_i \mid t_i = t_{j+1}\}$, i.e., these are the outputs of $\tilde{E}$ that $\mathcal{A}$ received for the same tweak t_{j+1} used in the $(j+1)$st query. Bounding the probability that $y_{j+1} \in \mathcal{Y}_{j+1}$ when y_{j+1} is uniform gives an upper bound on the probability with which $\mathcal{A}$ can distinguish $\mathbf{H}^*_j$ and $\mathbf{H}^{**}_j$. Thus,

$$\left|\Pr[\mathcal{A}(\mathbf{H}^*_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}^{**}_j) = 1]\right| \leq \frac{|\mathcal{Y}_{j+1}|}{2^n} \leq \frac{j}{2^n}\,. \tag{3}$$

Finally, we bound the distinguishability of $\mathbf{H}^{**}_j$ and $\mathbf{H}'_j$. Recall that the difference between these experiments is that from the $(j+1)$st query onward the former uses $(E^{(1)})^{T_j,K}$ while the latter uses $E^{T_{j+1},K}$ (both for the quantum queries of $\mathcal{A}$ and to instantiate TFX for the classical queries of $\mathcal{A}$). It follows that the two experiments are identical if $(E^{(1)})^{T_j,K}$ and $E^{T_{j+1},K}$ are equal. In what follows we bound the probability that they are not equal.

Both $(E^{(1)})^{T_j,K}$ and $E^{T_{j+1},K}$ involve $j+1$ swaps: $(E^{(1)})^{T_j,K}$ involves j swaps from the first j queries plus the extra swap by the definition of $E^{(1)}$ (i.e., $f_1(t_{j+1},k') \oplus x_{j+1}$ and s are swapped), whereas $E^{T_{j+1},K}$ induces $j+1$ swaps from the first $j+1$ queries. Since the $(j+1)$st query is a forward query, we have

$$(E^{(1)}_{k^*})^{T_j,K}(x) = \begin{cases} E_{k^*}(x) & k^* \neq k \\ \overleftarrow{S}_{T_j,E^{(1)},K} \circ \overrightarrow{S}_{T_j,E^{(1)},K} \circ E^{(1)}_k(x) & k^* = k\,. \end{cases} \tag{4}$$

Comparing Equations (1) and (4), we see that $(E^{(1)})^{T_j,K} = E^{T_{j+1},K}$ for all $k^* \neq k$. So we only need to consider $k^* = k$, in which case

$$(E^{(1)}_k)^{T_j,K}(x) = \overleftarrow{S}_{T_j,E^{(1)},K} \circ \overrightarrow{S}_{T_j,E^{(1)},K} \circ E^{(1)}_k(x)$$

and

$$(E_k)^{T_{j+1},K}(x) = \overleftarrow{S}_{T_{j+1},E,K} \circ \overrightarrow{S}_{T_{j+1},E,K} \circ E_k(x)\,.$$

Set $\mathcal{X} = \{x_1 \oplus f_1(t_1,k'), \ldots, x_j \oplus f_1(t_j,k')\}$, and let Bad_0 be the event that $x_{j+1} \oplus f_1(t_{j+1},k') \in \mathcal{X}$ and Bad_1 be the event that $s \in \mathcal{X}$. We first bound the probabilities of these events, and then show that $(E^{(1)}_k)^{T_j,K} = E^{T_{j+1},K}_k$ when neither Bad_0 nor Bad_1 occurs.

Since s is uniform and independent of everything else, it is immediate that $\Pr[\mathsf{Bad}_1] \leq j/2^n$. We continue with bounding the probability of Bad_0, which is more complex since we have to consider the tweaks from the first $j+1$ queries of $\mathcal{A}$. Intuitively, when considering a query whose tweak was the same as t_{j+1}, we rely on the assumption that $\mathcal{A}$ does not repeat queries; for queries where the tweaks are different, we use the XOR-uniformity property of f_1, f_2. We start by introducing the two sets

$$\mathcal{X}^{=} = \{x_i \oplus f_1(t_i,k') \mid 1 \leq i \leq j,\ t_i = t_{j+1}\}$$
$$\mathcal{X}^{\neq} = \{x_i \oplus f_1(t_i,k') \mid 1 \leq i \leq j,\ t_i \neq t_{j+1}\}\,.$$

These partition $\mathcal{X}$ into inputs using the same tweak as the $(j+1)$st query ($\mathcal{X}^{=}$) and those using a different tweak ($\mathcal{X}^{\neq}$). Hence,

$$\Pr[\mathsf{Bad}_0] = \Pr[\mathsf{Bad}_0^{=}] + \Pr[\mathsf{Bad}_0^{\neq}]\,,$$

where $\mathsf{Bad}_0^{=}$ is the event that $x_{j+1} \oplus f_1(t_{j+1},k') \in \mathcal{X}^{=}$ and $\mathsf{Bad}_0^{\neq}$ is the event that $x_{j+1} \oplus f_1(t_{j+1},k') \in \mathcal{X}^{\neq}$. For $\mathsf{Bad}_0^{=}$, we have

$$\begin{aligned} & x_{j+1} \oplus f_1(t_{j+1},k') \in \{x_i \oplus f_1(t_i,k') \mid t_i = t_{j+1}\} \\ \Leftrightarrow\ & x_{j+1} \in \{x_i \oplus f_1(t_i,k') \oplus f_1(t_{j+1},k') \mid t_i = t_{j+1}\} \\ \Leftrightarrow\ & x_{j+1} \in \{x_i \mid t_i = t_{j+1}\}\,, \end{aligned}$$

i.e., event $\mathsf{Bad}_0^{=}$ is equivalent to $x_{j+1} \in \{x_i \mid t_i = t_{j+1}\}$. Since $\mathcal{A}$ does not repeat queries, this means $\Pr[\mathsf{Bad}_0^{=}] = 0$. For $\mathsf{Bad}_0^{\neq}$, rewriting yields

$$\begin{aligned} & x_{j+1} \oplus f_1(t_{j+1},k') \in \{x_i \oplus f_1(t_i,k') \,|\, t_i \neq t_{j+1}\} \\ \Leftrightarrow\ & x_{j+1} \in \{x_i \oplus f_1(t_i,k') \oplus f_1(t_{j+1},k') \,|\, t_i \neq t_{j+1}\}\,. \end{aligned}$$

XOR-uniformity property of f_1 implies that every element in the set above is uniformly distributed, hence $\Pr[\mathsf{Bad}_0^{\neq}] \leq |\mathcal{X}^{\neq}|/2^n \leq j/2^n$. Summarizing,

$$\Pr[\mathsf{Bad}_0] = \Pr[\mathsf{Bad}_0^{=}] + \Pr[\mathsf{Bad}_0^{\neq}] \leq 0 + \frac{|\mathcal{X}^{\neq}|}{2^n} \leq \frac{j}{2^n}\,.$$

If neither Bad_0 or Bad_1 happens, then $E_k^{(1)}(x_i \oplus f_1(t_i, k')) = E_k(x_i \oplus f_1(t_i, k'))$ for every $1 \leq i \leq j$. Given that, we have

$$\begin{aligned}\overrightarrow{S}_{T_j,E^{(1)},K} &= \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{E_k^{(1)}(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')} \\ &= \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{E_k(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')} \;=\; \overrightarrow{S}_{T_j,E,K}\end{aligned}$$

and

$$\begin{aligned}\overleftarrow{S}_{T_j,E^{(1)},K} &= \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{E_k^{(1)}(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')} \\ &= \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{E_k(x_i \oplus f_1(t_i,k')),\, y_i \oplus f_2(t_i,k')} \;=\; \overleftarrow{S}_{T_j,E,K}\,.\end{aligned}$$

Therefore,

$$\begin{aligned}(E_k^{(1)})^{T_j,K}(x) &= \overleftarrow{S}_{j,E^{(1)},K} \circ \overrightarrow{S}_{j,E^{(1)},K} \circ E_k^{(1)}(x) \\ &= \overleftarrow{S}_{j,E,K} \circ \overrightarrow{S}_{j,E,K} \circ \mathsf{swap}_{E_k(f_1(t_{j+1},k') \oplus x_{j+1}),\, y_{j+1} \oplus f_2(t_{j+1},k')} \circ E_k(x) \\ &= \overleftarrow{S}_{j+1,E,K} \circ \overrightarrow{S}_{j+1,E,K} \circ E_k(x) \;=\; E_k^{T_{j+1},K}\,.\end{aligned}$$

Putting everything together, we conclude that

$$\left|\Pr[\mathcal{A}(\mathbf{H}_j^{**}) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right| \leq \Pr[\mathsf{Bad}_0] + \Pr[\mathsf{Bad}_1] \leq \frac{2j}{2^n}\,. \tag{5}$$

Combining Equations (2), (3), and (5) concludes the proof. $\square$

4 Applications of Our Result

In this section we show how Theorem 1 can be used to prove post-quantum security of the FX construction, a variant of the authenticated encryption scheme Elephant, and the message authentication code Chaskey.

4.1 The FX Construction

The FX construction [10] provides a mechanism for extending the key length of a cipher. Given a block cipher $E : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^n$, the FX construction yields a new block cipher $\mathsf{FX} : (\{0,1\}^m \times \{0,1\}^{2n}) \times \{0,1\}^n \to \{0,1\}^n$ via

$$\mathsf{FX}_{k,k_1,k_2}(x) = E_k(x \oplus k_1) \oplus k_2\,.$$

The FX construction is a special case of the TFX construction where $\kappa = 2n$, $\mathcal{T} = \emptyset$, $f_1(k_1, k_2) = k_1$, and $f_2(k_1, k_2) = k_2$. It is easy to verify that f_1 and f_2 are proper: they clearly satisfy uniformity, and XOR-uniformity is satisfied vacuously since $\mathcal{T} = \emptyset$. Specializing Theorem 1 to this case thus yields the following:

Theorem 2. *Let* FX *be as above and let* $\mathcal{A}$ *be an adversary making* q_C *classical queries to its first oracle and* q_Q *quantum queries to its second oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^m; k_1, k_2 \leftarrow \{0,1\}^n; \\ E \leftarrow \mathcal{E}(m,n)}} \left[\mathcal{A}^{\mathsf{FX}_{k,k_1,k_2}, E} = 1\right] - \Pr_{\substack{P \leftarrow \mathcal{P}(n); \\ E \leftarrow \mathcal{E}(m,n)}} \left[\mathcal{A}^{P,E} = 1\right] \right|$$
$$\leq (3 + 2\sqrt{2}) \cdot 2^{-(m+n)/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C}\right).$$

The above solves a problem left open by Jaeger et al. [8], who prove a similar result about security of the FX construction but only for non-adaptive attackers.

4.2 (A variant of) Elephant

$\mathsf{Elephant}$ is a lightweight authenticated encryption scheme (with associated data) under consideration for standardization by NIST [2]. It is based on a tweakable block cipher that we denote here by $\tilde{\mathsf{E}}$, which is in turn constructed from a specified public permutation P. Prior work [2] proves—in the purely classical setting—that $\mathsf{Elephant}$ is a secure authenticated encryption scheme if $\tilde{\mathsf{E}}$ is a secure tweakable block cipher, and that $\tilde{\mathsf{E}}$ is a secure tweakable block cipher if P is modeled as a public random permutation. It is straightforward to verify that this proof carries over to the setting of quantum adversaries with classical access to $\mathsf{Elephant}$, provided that $\tilde{\mathsf{E}}$ is post-quantum secure.

The tweakable block cipher $\tilde{\mathsf{E}} : \{0,1\}^m \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ (where $m \leq n$) used by $\mathsf{Elephant}$ is defined as

$$\tilde{\mathsf{E}}_k(t, x) = P(x \oplus f(t, P(k \| 0^{n-m}))) \oplus f(t, P(k \| 0^{n-m})),$$

where $f : \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ is a function that is proper with respect to $\mathcal{T}$. The particular structure of $\mathcal{T}$ is not relevant for us. We are unable to analyze $\tilde{\mathsf{E}}$ directly since it uses P both to define an Even-Mansour cipher as well as for expansion of k. Instead, we consider the simplified tweakable block cipher $\tilde{\mathsf{E}}' : \{0,1\}^n \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ defined as

$$\tilde{\mathsf{E}}'_k(t, x) = P(x \oplus f(t, k)) \oplus f(t, k).$$

This amounts to replacing $P(k \| 0^{n-m})$ with a uniform $k \in \{0,1\}^n$. Since a public random permutation is equivalent to a degenerate ideal cipher that takes no key, post-quantum security of $\tilde{\mathsf{E}}'$ follows directly from Theorem 1:

Theorem 3. *Let $\tilde{\mathsf{E}}'$ be as above and let $\mathcal{A}$ be an adversary making q_C classical queries to its first oracle and q_Q quantum queries to its second oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^n; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{\mathsf{E}}'_k, P} = 1 \right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n); \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{E}, P} = 1 \right] \right| \\ \leq (3 + 2\sqrt{2}) \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right).$$

As discussed earlier, the above theorem in combination with [2, Theorem B.3] implies post-quantum security (in the public random permutation model) of the variant of $\mathsf{Elephant}$ which uses $\tilde{\mathsf{E}}'$ in place of $\tilde{\mathsf{E}}$.

4.3 Chaskey

$\mathsf{Chaskey}$ [13], a lightweight MAC that is an ISO standard, is constructed from a specified permutation $P : \{0,1\}^n \to \{0,1\}^n$. Define $\mathsf{F}_{k,k'}(x) = P(x \oplus k) \oplus k'$; this is just an Even-Mansour cipher based on P. Evaluating $\mathsf{Chaskey}$ using key k involves evaluating $\mathsf{F}_{k,k}$, $\mathsf{F}_{k \oplus k_1, k_1}$, and $\mathsf{F}_{k \oplus k_2, k_2}$, where $k_1 = 2k$, $k_2 = 4k$, and multiplication is in the field $GF(2^n)$ with respect to a particular representation of field elements as n-bit strings. Prior work [13] shows that $\mathsf{Chaskey}$ is secure if these three instances of F are indistinguishable from three independent random permutations—a notion called *3PRP security*—and also proves 3PRP security of F when P is modeled as a public random permutation. Although this prior work considered classical adversaries only, it is not hard to verify that the proofs carry through to imply security of $\mathsf{Chaskey}$ against quantum adversaries making classical MAC queries so long as 3PRP security of F holds against adversaries making classical queries to the secretly keyed ciphers and quantum queries to P.

Theorem 1 readily implies 3PRP security of F in the post-quantum setting:

Theorem 4. *Let $\mathcal{A}$ be a quantum algorithm making q_C classical queries to its first three oracles and q_Q quantum queries to its fourth oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^n, \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{F}_{k,k}, \mathsf{F}_{k \oplus k_1, k_1}, \mathsf{F}_{k \oplus k_2, k_2}, P} = 1 \right] - \Pr_{R_1, R_2, R_3, P \leftarrow \mathcal{P}(n)} \left[\mathcal{A}^{R_1, R_2, R_3, P} = 1 \right] \right| \\ \leq (3 + 2\sqrt{2}) \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right),$$

where $k \in \{0,1\}^n$ is uniform, $k_1 = 2k$, and $k_2 = 4k$.

Proof. A public random permutation P is equivalent to a degenerate ideal cipher that takes no key (i.e., with $m = 0$). Letting $\mathcal{T} = \{0,1,2\} \subset GF(2^n)$ and defining $f_1(t,k) = k \oplus (2tk)$ and $f_2(t,k) = 2^t \cdot k$, we see that

$$\begin{aligned} \mathsf{TFX}_k^{f_1,f_2}[P](0,x) &= P(x \oplus k) \oplus k = \mathsf{F}_{k,k}(x) \\ \mathsf{TFX}_k^{f_1,f_2}[P](1,x) &= P(x \oplus k \oplus 2k) \oplus 2k = \mathsf{F}_{k \oplus k_1, k_1}(x) \\ \mathsf{TFX}_k^{f_1,f_2}[P](2,x) &= P(x \oplus k \oplus 4k) \oplus 4k = \mathsf{F}_{k \oplus k_2, k_2}(x). \end{aligned}$$

The theorem thus follows from Theorem 1 once we verify that f_1, f_2 satisfy the required properties. Uniformity of f_1 and f_2 follow readily from invertibility of non-zero elements in $GF(2^n)$. Finally, note that

$$f_1(t,k) \oplus f_1(t',k) = 2 \cdot (t \oplus t') \cdot k \text{ and } f_2(t,k) \oplus f_2(t',k) = (2^t \oplus 2^{t'}) \cdot k,$$

with $t \oplus t'$ and $2^t \oplus 2^{t'}$ non-zero for distinct t, t'; XOR-uniformity follows. This concludes the proof of the theorem. $\square$

As discussed earlier, the above theorem in combination with prior results [13, Theorem 1,2] implies post-quantum security of Chaskey (in the public random permutation model).

Acknowledgments

Work of Gorjan Alagic, Chen Bai, and Jonathan Katz was supported in part by NSF award CNS-2154705. Gorjan Alagic also acknowledges support from the U.S. Army Research Office under Grant Number W911NF-20-1-0015, the U.S. Department of Energy under Award Number DE-SC0020312, and the AFOSR under Award Number FA9550-20-1-0108. Work of Christian Majenz was funded by a NWO VENI grant (Project No. VI.Veni.192.159).

References

1. Gorjan Alagic, Chen Bai, Jonathan Katz, and Christian Majenz. Post-quantum security of the Even-Mansour cipher. In *Advances in Cryptology—Eurocrypt 2022, Part III*, volume 13277 of *LNCS*, pages 458–487. Springer, 2022.
2. Tim Beyne, Yu Long Chen, Christoph Dobraunig, and Bart Mennink. Elephant v2. Technical report, NIST, 2021. https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/finalist-round/updated-spec-doc/elephant-spec-final.pdf.
3. Xavier Bonnetain, Akinori Hosoyamada, María Naya-Plasencia, Yu Sasaki, and André Schrottenloher. Quantum attacks without superposition queries: The offline Simon's algorithm. In *Advances in Cryptology—Asiacrypt 2019, Part I*, volume 11921 of *LNCS*, pages 552–583. Springer, 2019.
4. Xavier Bonnetain, André Schrottenloher, and Ferdinand Sibleyras. Beyond quadratic speedups in quantum attacks on symmetric schemes. In *Advances in Cryptology—Eurocrypt 2022, Part III*, volume 13277 of *LNCS*, pages 315–344. Springer, 2022.
5. Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. *Journal of Cryptology*, 10(3):151–161, 1997.
6. Alex B. Grilo, Kathrin Hövelmanns, Andreas Hülsing, and Christian Majenz. Tight adaptive reprogramming in the QROM. In *Advances in Cryptology—Asiacrypt 2021, Part I*, volume 13090 of *LNCS*, pages 637–667. Springer, 2021. Available at https://eprint.iacr.org/2020/1361.
7. Lov K. Grover. A fast quantum mechanical algorithm for database search. In *STOC '96*, pages 212–219. ACM Press, 1996.

8. Joseph Jaeger, Fang Song, and Stefano Tessaro. Quantum key-length extension. In *19th Theory of Cryptography Conference—TCC 2021, Part I*, volume 13042 of *LNCS*, pages 209–239. Springer, 2021.
9. Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Breaking symmetric cryptosystems using quantum period finding. In *Advances in Cryptology—Crypto 2016, Part II*, volume 9815 of *LNCS*, pages 207–237. Springer, 2016.
10. Joe Kilian and Phil Rogaway. How to protect DES against exhaustive key search (an analysis of DESX). *J. Cryptology*, 14(1):17–35, 2001.
11. Hidenori Kuwakado and Masakatu Morii. Quantum distinguisher between the 3-round Feistel cipher and the random permutation. In *Proc. IEEE International Symposium on Information Theory*, pages 2682–2685, 2010.
12. Hidenori Kuwakado and Masakatu Morii. Security on the quantum-type Even-Mansour cipher. In *Proc. International Symposium on Information Theory and its Applications*, pages 312–316. IEEE Computer Society, 2012.
13. Nicky Mouha, Bart Mennink, Anthony Van Herrewege, Dai Watanabe, Bart Preneel, and Ingrid Verbauwhede. Chaskey: An efficient MAC algorithm for 32-bit microcontrollers. In *Intl. Conference on Selected Areas in Cryptography (SAC)*, pages 306–323. Springer, 2014.
14. Meltem Sönmez Turan, Kerry McKay, Donghoon Chang, Çağdaş Çalık, Lawrence Bassham, Jinkeon Kang, and John Kelsey. Status report on the second round of the NIST lightweight cryptography standardization process. Technical report, NIST, 2021.

A Proof of Lemma 2

Proof. The proof of this lemma is similar to the proof of the resampling lemma for random permutations. Here, we detail the parts of the proof that are different. Let F be the internal register (called "database register") of a superposition oracle for an ideal cipher, i.e., $F = F_{0^m}F_{0^{m-1}1}\dots F_{1^m}$ where each $F_k = F_{k,0^n},\dots F_{k,1^n}$ is a database register for a random permutation. Each F_k is initialized in the initial state $|\phi_0\rangle$ for a random permutation, namely,

$$|\phi_0\rangle = (2^n!)^{-1/2} \sum_{\pi\in\mathcal{P}(n)} |\pi\rangle.$$

By analogy to the proof of [1, Lemma 5], define the projectors

$$(P_{k_0 s_0 s_1})_{KX} = \begin{cases} \mathbb{1} & s_0 = s_1 \\ \mathbb{1} - |k_0\rangle\langle k_0| \otimes (|s_0\rangle\langle s_0| + |s_0\rangle\langle s_0|)_X & s_0 \neq s_2 \end{cases}$$

and

$$\begin{aligned}&\left(P^{\mathrm{inv}}_{k_0 s_0 s_1}\right)_{KYF}\\ &= \begin{cases} \mathbb{1} & s_0 = s_1 \\ |k_0\rangle\langle k_0|_K \otimes \sum_{y\in\{0,1\}^n} |y\rangle\langle y|_Y \otimes (\mathbb{1} - |y\rangle\langle y|)^{\otimes 2}_{F_{k_0,s_0}F_{k_0,s_1}} & s_0 \neq s_1. \end{cases}\end{aligned}$$

With this generalized definition of P and P^{inv}, it is straightforward to see that Equations (11) and (12) from [1] still hold, i.e.,

$$\left[\mathsf{Swap}_{F_{k,s_0}F_{k,s_1}},\ O_{KXYF}\left(P_{k,s_0s_1}\right)_{KX}\right]=0$$

and

$$\left[\mathsf{Swap}_{F_{k,s_0}F_{k,s_1}},\ O^{\mathrm{inv}}_{KXYF}\left(P^{\mathrm{inv}}_{k,s_0s_1}\right)_{KYF}\right]=0.$$

For an arbitrary state $|\psi\rangle_{KXE}$, let

$$|\psi\rangle_{KXE}=\sum_{\substack{k\in\{0,1\}^m\\x\in\{0,1\}^m}}|k\rangle_K|x\rangle_X\otimes|\psi_{kx}\rangle_E$$

be its expansion in the computational basis on X. By the definition of ϵ, we obtain generalizations of Equations (13) and (15) from [1], namely,

$$\begin{aligned}
&\mathbb{E}_{(k_0,s_0,s_1)\sim D}\left[\|\left(P_{k_0s_0s_1}\right)_{KX}|\psi\rangle_{KXE}\|_2^2\right]\\
&=\sum_{\substack{k\in\{0,1\}^m\\x\in\{0,1\}^m}}\||\psi_{kx}\rangle\|_2^2\mathbb{E}_{B\sim D}\left[\|\left(\mathit{\Pi}_B\right)_{KX}|k\rangle_K|x\rangle_X\|_2^2\right]\\
&=\sum_{\substack{k\in\{0,1\}^m\\x\in\{0,1\}^m}}\||\psi_{kx}\rangle\|_2^2\Pr_{(k_0,s_0,s_1)\sim D}[(k,x)\in\{(k_0,s_0),(k_0,s_1)\}]\\
&\leq\epsilon
\end{aligned}$$

and

$$\mathop{\mathbb{E}}_{s_0,s_1}\left[\left\|\left(\bar{P}^{\mathrm{inv}}_{k_0s_0s_1}\right)_{KYF}|\psi\rangle_{KYEF}\right\|_2^2\right]\leq\epsilon.$$

The remainder of the proof is analogous to the proof of [1, Lemma 5]. $\square$