Post-Quantum Security of Tweakable Even-Mansour, and Applications

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

[1] University of Maryland
{galagic,cbai1}@umd.edu
[2] NIST
[3] Google
jkatz2@gmail.com
[4] Technical University of Denmark
chmaj@dtu.dk
[5] University of Konstanz
patrick.struck@uni.kn

Abstract. The tweakable Even-Mansour construction yields a tweakable block cipher from a public random permutation. We prove post-quantum security of tweakable Even-Mansour when attackers have *quantum* access to the random permutation but only *classical* access to the secretly-keyed construction, the relevant setting for most real-world applications. We then use our results to prove post-quantum security—in the same model—of the symmetric-key schemes Chaskey (an ISO-standardized MAC), Elephant (an AEAD finalist of NIST's lightweight cryptography standardization effort), and a variant of Minalpher (an AEAD second-round candidate of the CAESAR competition).

1 Introduction

The development of large-scale quantum computers would have a significant impact on cryptography. 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 [8] to carry out a key-recovery attack. In general, however, doubling the key length may not be sufficient [4,13,14], 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 scheme. The distinction between

⋆ Work done in part while at the University of Maryland.

Q1 and Q2 is significant: for many symmetric-key constructions, polynomial-query attacks are known in the Q2 model but not in the Q1 model [12–14]. At the same time, 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 when the honest devices implementing the construction are entirely classical). The Q1 model is thus a much better fit for realistic quantum attacks, and several recent works [1,4,11] have 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 is challenging since it requires reasoning about a combination of (related) classical and quantum oracles. Additional complications arise when reasoning about permutations (rather than functions), particularly when their inverse may also be queried, as in the random-permutation and ideal-cipher models. Indeed, most results in a "hybrid" classical-/quantum-query setting (e.g., [5,9,16]) deal with oracles for functions, and there are only a few existing results in the Q1 model that deal with random permutations. Jaeger et al. [11] gave positive results for security of the FX construction (a mechanism for key-length extension of an ideal cipher); 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] showed post-quantum security of the full Even-Mansour construction (i.e., based on a random permutation) against adaptive adversaries.

1.1 Our Results

We show post-quantum security of the *tweakable* Even-Mansour construction, a tweakable block cipher constructed from a public random permutation. We then use this result to establish post-quantum security of several symmetric-key schemes. We stress that post-quantum security of tweakable Even-Mansour does not follow from post-quantum security of Even-Mansour. Indeed, the tweak must be incorporated in a way that satisfies several technical conditions; in addition, incorporating both tweaks and possible key expansion introduces dependencies and requires significant technical work to analyze. In all of our results, adversaries can make adaptive queries to any permutations to which they have access (whether quantum or classical, as appropriate) in both the forward and inverse directions. We now summarize our results.

Tweakable Even-Mansour. Let $P : \{0,1\}^n \to \{0,1\}^n$ be a permutation. The tweakable Even-Mansour scheme $\mathsf{TEM}^{f_1,f_2}[P] : \{0,1\}^n \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ is defined as

$$\mathsf{TEM}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t,k)) \oplus f_2(t,k),$$

where the key k is of length n, the set $\mathcal{T}$ is a tweak space, and f_1, f_2 are functions satisfying some technical conditions we omit here. We also consider a variant $\mathsf{TEM\text{-}KX}^{f_1,f_2}[P] : \{0,1\}^\kappa \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ (where $\kappa \leq n$) that combines tweakable Even-Mansour with key expansion, and is defined as

$$\mathsf{TEM\text{-}KX}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t, P(k\|0^{n-\kappa}))) \oplus f_2(t, P(k\|0^{n-\kappa}))\,.$$

Our main result is that both the above are secure (post-quantum) tweakable block ciphers when P is modeled as a random permutation.

Theorem 1 (informal). *An adaptive adversary making q_C classical queries to $\mathsf{TEM\text{-}KX}_k^{f_1,f_2}[P]$ (for uniform $k \in \{0,1\}^\kappa$) and q_Q quantum queries to a random permutation P can distinguish the former from a uniform tweakable block cipher with probability at most $\mathcal{O}\big(2^{-\kappa/2} \cdot (q_C\sqrt{q_Q} + q_Q\sqrt{q_C})\big)$.*

(The above is stated formally as Theorem 3 and proved in Section 4.1.) Setting $\kappa = n$ implies security of TEM as a corollary (since $P(k)$ is uniform when $k \in \{0,1\}^n$ is uniform, for any permutation P). It follows that any post-quantum attack against TEM requires $q_C^2 \cdot q_Q + q_Q^2 \cdot q_C \approx 2^n$; hence $\Omega(2^{n/3})$ queries are necessary for constant success probability, matching known attacks [3, 10].

We also consider an alternative method of performing key expansion in which a key $k \in \{0,1\}^\kappa$ is expanded to an "effective key" of length n by computing $F_P(k) = P(k\|0^{n-\kappa}) \oplus k\|0^{n-\kappa}$. This gives rise to another variant of tweakable Even-Mansour, defined as

$$\mathsf{TEM\text{-}KX1}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t, F_P(k))) \oplus f_2(t, F_P(k))).$$

We show that the key-expansion function F_P is a pseudorandom generator (even for adversaries having quantum access to P). Using this fact, we are able to prove a tighter security bound for $\mathsf{TEM\text{-}KX1}$ than what we show for $\mathsf{TEM\text{-}KX}$ (see Theorem 5 in Section 4.2 for a formal statement):

Theorem 2 (informal). *An adaptive adversary making q_C classical queries to $\mathsf{TEM\text{-}KX1}_k^{f_1,f_2}[P]$ (for uniform $k \in \{0,1\}^\kappa$) and q_Q quantum queries to a random permutation P can distinguish the former from a uniform tweakable block cipher with probability at most $\mathcal{O}\big(2^{-\kappa/2} \cdot (q_C + q_Q) + 2^{-n/2} \cdot (q_C\sqrt{q_Q} + q_Q\sqrt{q_C})\big)$.*

A new resampling lemma. As a key technical tool used in our results, we prove a generalization of existing "resampling lemmas" [1, 7] sufficient to handle tweakable block ciphers, something we believe to be of independent interest. A resampling lemma controls the success probability of a quantum-query adversary $\mathcal{D}$ in an experiment of the following form:

1. $\mathcal{D}$ receives quantum oracle access to a random permutation P;
2. two inputs s_0, s_1 are sampled from some distribution;
3. $\mathcal{D}$ receives quantum oracle access to either P, or P with inputs s_0 and s_1 "swapped"; it succeeds if it can correctly guess which is the case.

Prior work considered only the uniform distribution on s_0, s_1. We give a new resampling lemma that handles a wider class of (adversarially influenced) distributions, and even allows the distribution to depend on information $\mathcal{D}$ learns about P during step 1 of the above experiment (cf. Lemma 3 in Section 3):

Lemma 1 (informal). *In the above experiment, for any $\mathcal{D}$ making at most q quantum queries to P in step 1, $\Pr[\mathcal{D}$ succeeds$] \le 1/2 + \mathcal{O}(\sqrt{q\varepsilon})$, where ε is the min-entropy of s_0, s_1.*

To prove the lemma, we develop a novel permutation variant of the stateful simulation technique for quantum-accessible random oracles [19] (i.e., the *superposition oracle* technique). In this context, *some* information about the input-output pairs learned by the adversary via quantum queries can be read directly from the oracle's internal quantum register. In the original superposition oracle technique [19], this useful feature is a consequence of the statistical independence of the function values of a random oracle. Existing generalizations to invertible random permutations [1] lack this feature.

Applications. In Section 5 we use our results to derive corollaries regarding the post-quantum security of various symmetric-key schemes when modeling the underlying permutations on which they are based as ideal permutations. In each case, security is established in two stages. First, we choose the tweak space $\mathcal{T}$ and the tweak functions f_1 and f_2 appropriately, and apply our theorems above to prove security for a certain block cipher construction. Then, we invoke existing results to reduce security of the overall cryptographic scheme (in the appropriate sense) to security of this cipher. Specifically:

1. We show how to specialize TEM so it captures the three pseudorandom permutations used by Chaskey [15], an ISO-standardized lightweight MAC. We can thus prove post-quantum security of Chaskey using Theorem 1.
2. We show how to specialize TEM-KX to the tweakable block cipher at the core of Elephant [2], an authenticated encryption scheme that was a finalist of NIST's lightweight standardization effort [18]. Theorem 1 then implies post-quantum security for Elephant. Using Theorem 2, we can prove a tighter security bound for a variant of Elephant that uses a slightly different key-expansion step.
3. We show how to specialize TEM-KX1 to the tweakable block cipher used by (a variant of) Minalpher [17], an authenticated encryption scheme that was a second-round candidate of the CAESAR competition. Theorem 2 then implies post-quantum security for this variant.

To our knowledge, these are the first proofs of post-quantum security for any versions of Chaskey, Elephant, or Minalpher.

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* (or random-permutation model), a uniform permutation $P \leftarrow \mathcal{P}(n)$ is sampled and then provided as an oracle (in both the forward and inverse directions) to all parties.

A block cipher $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ is a keyed permutation, i.e., $E_k(\cdot) = E(k,\cdot)$ is a permutation of $\{0,1\}^n$ for all $k \in \{0,1\}^\kappa$. We say E is a *pseudorandom permutation* if E_k (for uniform $k \in \{0,1\}^\kappa$) is indistinguishable from a uniform permutation in $\mathcal{P}(n)$ even for adversaries who may query their oracle in both the forward and inverse directions.

For a set $\mathcal{T}$, let $\mathcal{E}(\mathcal{T}, n)$ be the set of all functions $E : \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ such that $E(t, \cdot)$ is a permutation on $\{0,1\}^n$ for all $t \in \mathcal{T}$. A tweakable block cipher $\tilde{E} : \{0,1\}^\kappa \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ is a family of permutations indexed by both a key $k \in \{0,1\}^\kappa$ 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 $k \in \{0,1\}^\kappa$ and $t \in \mathcal{T}$. Tweakable block cipher $\tilde{E}_k$ is *secure* if $\tilde{E}_k$ (for uniform choice of $k \in \{0,1\}^\kappa$) is indistinguishable from a uniform $\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n)$.

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 public primitives (e.g., a random permutation), however, we provide the distinguisher with *quantum* oracle access to the public primitive. Thus, for example, a quantum distinguisher in the public-permutation model can apply the unitary operators

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

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

A reprogramming lemma. We recall here a reprogramming lemma from prior work [1] that applies to the following experiment. A distinguisher $\mathcal{D}$ chooses an arbitrary function F along with a randomized process $\mathcal{B}$ for determining a set of points B at which F should (potentially) be reprogrammed to some known value. $\mathcal{D}$ is then given quantum access to either F or a reprogrammed version of F; when it is done making its oracle queries, $\mathcal{D}$ is given B. Roughly, the lemma says that $\mathcal{D}$ cannot determine whether it was interacting with F or the reprogrammed version of F as long as no point is reprogrammed with high probability.

Formally, 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

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

The following is taken verbatim from [1, Lemma 3]:

Lemma 2. *Let $\mathcal{D}$ be a quantum 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 $\varepsilon = \max_{x \in \{0,1\}^m} \{\Pr_{B \leftarrow \mathcal{B}}[x \in B_1]\}$.*

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

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

3 A New Resampling Lemma

In this section, we describe a new resampling lemma for random permutations that generalizes earlier results [1,7]. We consider a two-phase experiment in which a distinguisher $\mathcal{D}$ is first given quantum oracle access to a uniform permutation $P : \{0,1\}^n \to \{0,1\}^n$. Then, a point $s_0 \in \{0,1\}^n$ is chosen in a manner specified by the distinguisher and a uniform point $s_1 \in \{0,1\}^n$ is also chosen; in a second phase $\mathcal{D}$ is given access either to the original permutation $P^{(0)} = P$ or a modified permutation $P^{(1)}$ that is the same as P except that the values of $P(s_0)$ and $P(s_1)$ are swapped. (See below for details.) We show, roughly speaking, that so long as the distribution of s_0 has high min-entropy and $\mathcal{D}$ makes only a bounded number of queries in the first phase of the experiment, $\mathcal{D}$ cannot distinguish those possibilities.

Compared to prior work of Alagic et al. [1], our result is more general in the following ways:

- it allows for more general distributions of s_0;
- it allows for the distribution of s_0 to be *adaptively* chosen by $\mathcal{D}$, after $\mathcal{D}$ makes queries to P in the first phase;
- it furthermore allows $\mathcal{D}$ to select a sampling algorithm for s_0 that will itself make a query to P.

In order to achieve these improvements, we use a different proof technique from that of Alagic et al. [1]. Our approach is closer in spirit to an earlier technique of Grilo et al. [7], which was previously only applied to random functions.

We now state our new resampling lemma. For $s_0, s_1 \in \{0,1\}^n$, define

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

Lemma 3. *Let $F \subset \mathcal{P}(n)$. Consider the following experiment involving a quantum distinguisher $\mathcal{D}$:*

Phase 1: *Choose uniform $P \in \mathcal{P}(n)$, and give $\mathcal{D}$ quantum access to P. $\mathcal{D}$ outputs (D, τ), where D is a distribution on $\{0,1\}^n$ and $\tau \in F$.*

Phase 2: *Sample* $\hat{s} \leftarrow D$*, set* $s_0 = \tau \circ P(\hat{s})$*, and choose* $s_1 \leftarrow \{0,1\}^n$*. Let* $P^{(0)} = P$ *and define* $P^{(1)} = P \circ \mathsf{swap}_{s_0,\, s_1}$*. A uniform bit* $b \in \{0,1\}$ *is chosen, and* $\mathcal{D}$ *is given* $\hat{s}$ *and quantum access to* $P^{(b)}$*. Then* $\mathcal{D}$ *outputs a guess* b'*.*

Let $\varepsilon = 2 \cdot \mathbb{E}_{(D,\tau)\leftarrow\mathcal{D}^P}\left[\max_{x\in\{0,1\}^n} \Pr_{x'\leftarrow D}[x' = x]\right]$*. For any* $\mathcal{D}$ *making at most* q *queries to* P *in phase 1,*

$$\begin{aligned}&|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]| \\ &\qquad \le \sqrt{\varepsilon} \cdot \left(1 + \sqrt{q + \log\left(\frac{11 \cdot |F|}{\sqrt{\varepsilon}}\right)}\right).\end{aligned}$$

The proof of Lemma 3 is given in Appendix A.

4 Post-Quantum Security of Tweakable Even-Mansour

We use the result of the previous section to prove the post-quantum security of three different variants of the tweakable Even-Mansour construction. In Section 4.1, we prove security of TEM-KX; we then prove security of TEM as a corollary. In Section 4.2, we prove security of TEM-KX1 by showing that its key-expansion function is a pseudorandom generator.

4.1 Security of TEM-KX and TEM

Let $P \in \mathcal{P}(n)$ be a permutation and $\mathcal{T}$ a finite set, and fix two functions $f_1, f_2 \colon \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$. We consider a key-expanding version of the tweakable Even-Mansour scheme $\mathsf{TEM\text{-}KX}^{f_1,f_2}[P] : \{0,1\}^\kappa \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ defined as

$$\mathsf{TEM}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t, P(k||0^{n-\kappa}))) \oplus f_2(t, P(k||0^{n-\kappa}))\,.$$

We assume the tweak functions f_1, f_2 satisfy some structural properties.

Definition 1. *A function* $f : \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ *is* ***proper*** *(with respect to* $\mathcal{T}$*) if it satisfies the following two properties:*

Uniformity: *For all* $t \in \mathcal{T}$*, the function* $f(t,\cdot)$ *is a permutation.*
XOR-universality: *For all distinct* $t, t' \in \mathcal{T}$ *and all* $y \in \{0,1\}^n$*,*

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

Theorem 3. *Let* TEM-KX *be as above, and let* $\mathcal{A}$ *be an adversary making* q_C *classical queries to its first oracle and* $q_Q \ge \max\{n, \log(11 \cdot |\mathcal{T}|)\}$ *quantum*

queries[1] *to its second oracle. If* f_1, f_2 *are proper with respect to* $\mathcal{T}$*, then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^\kappa; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{TEM\text{-}KX}_k^{f_1,f_2}[P],P} = 1 \right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T},n); \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{E},P} = 1 \right] \right| \leq 7 \cdot 2^{-\kappa/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right).$$

Proof. The high-level structure of our proof is similar to the proof of security for the Even-Mansour construction by Alagic et al. [1], though here relying heavily on our new resampling lemma. 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 a triple (t, x, y) of tweak, input, and output by making a query to its classical oracle, it never again submits a query (t, x) (resp., (t, y)) to that oracle in the forward (resp., inverse) direction. 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[2] distribute its q_Q quantum queries between these stages arbitrarily, and we let $q_{Q,j}$ be the expected number of quantum queries that $\mathcal{A}^{\tilde{E},P}$ makes in the jth stage, where the expectation is taken over $\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n)$ and $P \leftarrow \mathcal{P}(n)$ and any internal randomness/measurements of $\mathcal{A}$. Note that $\sum_{j=0}^{q_C} q_{Q,j} = q_Q$.

Fixing f_1, f_2, we write $\mathsf{TEM\text{-}KX}_k$ for $\mathsf{TEM\text{-}KX}_k^{f_1,f_2}$. In a given execution of $\mathcal{A}$, we denote its jth classical query by (t_j, x_j, y_j, b_j), where $t_j \in \mathcal{T}$ is a tweak, $(x_j, y_j) \in \{0,1\}^n \times \{0,1\}^n$ is an input/output pair, and $b_j \in \{0,1\}$ indicates the query direction, i.e., $b_j = 0$ (resp., $b_j = 1$) means that the jth 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 of the first j classical 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. We use the notation $\prod_{i=n}^{1} f_i = f_n \circ \cdots \circ f_1$ to denote the composition in reverse order. For a permutation P, a key 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,

[1] The mild assumption on q_Q can be avoided at the expense of an additive term of $\mathcal{O}(q_C \cdot 2^{-\kappa/2} \cdot (n + \log|\mathcal{T}|))$ in the bound.

[2] Alternatively, the techniques of [6] can be used to turn the adversary into one that uses a fixed query schedule; the overall bound would be unchanged.

define the operators

$$\overrightarrow{S}_{T_j,P,k} = \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{P(x_i \oplus f_1(t_i,P(k||0^{n-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))}$$

$$\overrightarrow{Q}_{T_j,P,k} = \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{x_i \oplus f_1(t_i,P(k||0^{n-\kappa})),\, P^{-1}(y_i \oplus f_2(t_i,P(k||0^{n-\kappa})))}$$

$$\overleftarrow{S}_{T_j,P,k} = \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{P(x_i \oplus f_1(t_i,P(k||0^{0-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))}$$

$$\overleftarrow{Q}_{T_j,P,k} = \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{x_i \oplus f_1(t_i,P(k||0^{n-\kappa})),\, P^{-1}(y_i \oplus f_2(t_i,P(k||0^{n-\kappa})))}$$

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

$$P^{T_j,k}(x) = \overleftarrow{S}_{T_j,P,k} \circ \overrightarrow{S}_{T_j,P,k} \circ P(x)\,.$$

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

$$\overleftarrow{S}_{j,P,k} \circ \overrightarrow{S}_{T_j,P,k} \circ P = \overleftarrow{S}_{T_j,P,k} \circ P \circ \overrightarrow{Q}_{T_j,P,k} = P \circ \overleftarrow{Q}_{T_j,P,k} \circ \overrightarrow{Q}_{T_j,P,k}\,.$$

Roughly speaking, $P^{T_j,k}$ is the minimal modification of P that is consistent with the forward ($\rightarrow$) and inverse ($\leftarrow$) queries from the transcript T_j when post-composed (S) or pre-composed (Q) with P. For compactness we occasionally write P^j in place of $P^{T_j,k}$ 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 $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $P \in \mathcal{P}(n)$, and a uniform key $k \in \{0,1\}^\kappa$. Then:

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using P, 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 $\mathsf{TEM\text{-}KX}_k[P^{T_j,k}]$ and its quantum queries using $P^{T_j,k}$.

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

$$\underbrace{P, \tilde{E}, P, \cdots, \tilde{E}, P,}_{j \text{ classical queries}}\ \underbrace{\mathsf{TEM\text{-}KX}_k[P^j], P^j, \cdots, \mathsf{TEM\text{-}KX}_k[P^j], P^j}_{q_C - j \text{ classical queries}}\,.$$

Each instance of $\tilde{E}$ or $\mathsf{TEM\text{-}KX}_k[P^j]$ represents a single classical query, while each instance of P or P^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., $\mathcal{A}^{\mathsf{TEM\text{-}KX}_k[P],P}$, and $\mathbf{H}_{q_C}$ is the execution of $\mathcal{A}$ in the ideal world, i.e., $\mathcal{A}^{\tilde{E},P}$.

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

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

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using P, 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 $\mathsf{TEM\text{-}KX}_k[P^{T_{j+1},k}]$ and its quantum queries using $P^{T_{j+1},k}$.

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

$$\underbrace{P, \tilde{E}, P, \cdots, \tilde{E}, P,}_{j \text{ classical queries}} \tilde{E}, P^{j+1}, \underbrace{\mathsf{TEM\text{-}KX}_k[P^{j+1}], P^{j+1} \cdots, \mathsf{TEM\text{-}KX}_k[P^{j+1}], P^{j+1}}_{q_C - j - 1 \text{ classical queries}}.$$

In Lemma 4 and Lemma 5, 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 $0 \le j < q_C$:

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

and

$$\begin{aligned}&\big|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\big| \\ &\le 2^{-\kappa/2} \left(1 + \sqrt{q_Q + \log(11\,|\mathcal{T}|) + n + \kappa/2}\right) + \frac{4j}{2^\kappa}.\end{aligned}$$

Using the above, we have

$$\begin{aligned}&|\Pr[\mathcal{A}(\mathbf{H}_0) = 1] - \Pr[\mathcal{A}(\mathbf{H}_{q_C}) = 1]| \\ &\le \sum_{j=0}^{q_C - 1} \left(2^{-\kappa/2} \left(1 + \sqrt{q_Q + \log(11\,|\mathcal{T}|) + n + \kappa/2} + 2q_{Q,j+1}\sqrt{2(j+1)}\right) + \frac{4j}{2^\kappa}\right) \\ &\le \frac{4q_C^2}{2^\kappa} + \sum_{j=0}^{q_C - 1} 2^{-\kappa/2} \left(1 + \sqrt{q_Q + \log(11\,|\mathcal{T}|) + n + \kappa/2} + 2 \cdot q_{Q,j+1}\sqrt{2q_C}\right) \\ &\le \frac{4q_C^2}{2^\kappa} + 2^{-\kappa/2} \left(q_C + q_C\sqrt{q_Q + \log(11\,|\mathcal{T}|) + n + \kappa/2} + 2\sqrt{2} q_Q \sqrt{q_C}\right).\end{aligned}$$

The above bound can be simplified. By assumption, $q_Q \ge \log(11 \cdot |\mathcal{T}|)$ and $q_Q \ge n \ge \kappa$. So $\sqrt{q_Q + \log(11 \cdot |\mathcal{T}|) + n + \kappa/2} \le \sqrt{7q_Q/2}$. We may also assume $q_C \le 2^{\kappa/2}$ since otherwise the bound is larger than 1. Under these assumptions,

we have $4q_C^2 \cdot 2^{-\kappa} \le 4q_C \cdot 2^{-\kappa/2} \le 4q_C\sqrt{q_Q} \cdot 2^{-\kappa/2}$ and so

$$\begin{aligned}
&\frac{4q_C^2}{2^\kappa} + 2^{-\kappa/2} \cdot \left(q_C + q_C\sqrt{q_Q + \log(11 \cdot |\mathcal{T}|) + n + \kappa/2} + 2\sqrt{2} q_Q \sqrt{q_C} \right) \\
&\le 2^{-\kappa/2} \cdot \left(5q_C + q_C\sqrt{7q_Q/2} + 2\sqrt{2} q_Q \sqrt{q_C} \right) \\
&\le 2^{-\kappa/2} \cdot \left(\left(5 + \sqrt{\frac{7}{2}} \right) q_C\sqrt{q_Q} + 2\sqrt{2} q_Q \sqrt{q_C} \right) \\
&\le 2^{-\kappa/2} \cdot \left(7q_C\sqrt{q_Q} + 2\sqrt{2} q_Q \sqrt{q_C} \right) \le 7 \cdot 2^{-\kappa/2} \cdot \left(q_C\sqrt{q_Q} + q_Q\sqrt{q_C} \right),
\end{aligned}$$

as claimed.

We now prove Lemma 4 and Lemma 5.

Lemma 4. *For $j = 0, \ldots, q_C - 1$,*

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

where $q_{Q,j+1}$ is the expected number of queries $\mathcal{A}$ makes to P 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 a distinguisher $\mathcal{D}$ for the experiment from Lemma 2:

Phase 1: $\mathcal{D}$ samples uniform $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$ and $P \in \mathcal{P}(n)$. It then runs $\mathcal{A}$, answering its quantum queries using P 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 classical queries by $\mathcal{A}$ thus far. $\mathcal{D}$ defines $F(a, x) \coloneqq P^a(x)$ for $a \in \{1, -1\}$.
It also defines the following randomized algorithm $\mathcal{B}$: sample $k \leftarrow \{0,1\}^\kappa$ and then compute the set B of input/output pairs to be reprogrammed so that $F^{(B)}(a, x) = (P^{T_{j+1},k})^a(x)$ for all a, x. Finally, $\mathcal{D}$ outputs $(F, \mathcal{B})$.

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 before $\mathcal{A}$ makes its next (i.e., $(j+2)$nd) classical query.

Phase 3: $\mathcal{D}$ is given k. It resumes running $\mathcal{A}$, answering its classical queries using $\mathsf{TEM\text{-}KX}_k[P^{T_{j+1},k}]$ and its quantum queries using $P^{T_{j+1},k}$. 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 2. 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 $P^{T_{j+1},k}$ and the fact that $F^{(B)}(a,x) = (P^{T_{j+1},k})^a(x)$. Fixing P and T_{j+1}, the probability that any particular input (a,x) is reprogrammed is at most the probability (over k) that it lies in the set

$$\left\{ \begin{array}{l} (1, x_i \oplus f_1(t_i, P(k||0^{n-\kappa}))),\ (1, P^{-1}(y_i \oplus f_2(t_i, P(k||0^{n-\kappa})))), \\ (-1, P(x_i \oplus f_1(t_i, P(k||0^{n-\kappa})))),\ (-1, y_i \oplus f_2(t_i, P(k||0^{n-\kappa}))) \end{array} \right\}_{i=1}^{j+1}.$$

We compute the probability that $(a,x) = (1, x_i \oplus f_1(t_i, P(k||0^{n-\kappa})))$ for some fixed i. P is a permutation, and so is $f_1(t_i,\cdot)$. As k is uniform,

$$\Pr_k[(a,x) = (1, x_i \oplus f_1(t_i, P(k||0^{n-\kappa})))] = \begin{cases} 2^{-\kappa} & a = 1 \\ 0 & a = -1 \end{cases}.$$

A similar bound holds for the other possibilities. By distinguishing the cases $a = 1$ and $a = -1$ and applying a union bound, we get $\varepsilon \le 2(j+1)/2^\kappa$.

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 5. *For $j = 0, \ldots, q_C$,*

$$\begin{aligned} &\left|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right| \\ &\le \frac{1}{2^{\kappa/2}} \left(1 + \sqrt{q_Q + \log(11\,|\mathcal{T}|) + n + \kappa/2}\right) + \frac{4j}{2^\kappa}. \end{aligned}$$

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

Experiment $\mathbf{H}^*_j$. Sample uniform $\tilde{E} \in \mathcal{E}(\mathcal{T}, n)$, $P \in \mathcal{P}(n)$, and $k \in \{0,1\}^\kappa$. Then

1. Run $\mathcal{A}$, answering its classical queries using $\tilde{E}$ and its quantum queries using P, 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.[3]
2. Define $s_0 = f_1(t_{j+1}, P(k||0^{n-\kappa})) \oplus x_{j+1}$ and sample uniform $s_1 \in \{0,1\}^n$. Define $P^{(1)}$ as $P^{(1)}(x) = (P \circ \mathsf{swap}_{s_0,\, s_1})(x)$. Then continue running $\mathcal{A}$, answering its remaining classical queries (including the $(j+1)$st) using $\mathsf{TEM\text{-}KX}_k[(P^{(1)})^{T_j,k}]$, and its quantum queries using $(P^{(1)})^{T_j,k}$.

Experiment $\mathbf{H}^{**}_j$ is the same as $\mathbf{H}^*_j$, except that the $(j+1)$st query is answered using $\tilde{E}$ to obtain $y_{j+1} = \tilde{E}(t_{j+1}, x_{j+1})$, and then we define $s_1 = (P^{T_j,k})^{-1}(y_{j+1} \oplus f_2(t_{j+1}, P(k||0^{n-\kappa})))$. We have

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

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

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 a distinguisher $\mathcal{D}$ for the experiment of Lemma 3, where $F = \{f_1(t,\cdot) \oplus x\}_{t\in\mathcal{T},x\in\{0,1\}^n}$.

Phase 1: $\mathcal{D}$ is given quantum access to a uniform permutation P. It samples uniform $\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n)$ and then runs $\mathcal{A}$, answering its quantum queries using P and its classical queries using $\tilde{E}$ (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 $T_j = \big((t_1, x_1, y_1, b_1), \ldots, (t_j, x_j, y_j, b_j)\big)$ of the queries $\mathcal{A}$ has made to its classical oracle thus far. $\mathcal{D}$ lets $\tau \in F$ be such that $\tau(\cdot) = f_1(t_{j+1}, \cdot) \oplus x_{j+1}$, and defines the distribution D on $\{0,1\}^n$ that chooses uniform $k \in \{0,1\}^\kappa$ and outputs $k\|0^{n-\kappa}$. Finally, $\mathcal{D}$ outputs (D, τ).

Phase 2: The challenger samples $\hat{s} \leftarrow D$ with $\hat{s} = k\|0^{n-\kappa}$. Then $\mathcal{D}$ is given $\hat{s}$ and quantum oracle access to the permutation $P^{(b)}$. It continues running $\mathcal{A}$, answering its remaining classical queries—including the $(j+1)$st—using $\mathsf{TEM\text{-}KX}_k[(P^{(b)})^{T_j,k}]$, and its remaining quantum queries using $(P^{(b)})^{T_j,k}$. $\mathcal{D}$ outputs whatever $\mathcal{A}$ does.

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 $b = 0$, $\mathcal{D}$ gets access to $P^{(0)} = P$ in phase 2. Since $\mathcal{D}$ answers all quantum queries using $(P^{(0)})^{T_j,k}$ and all classical queries using $\mathsf{TEM\text{-}KX}_k[(P^{(0)})^{T_j,k}]$, 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 $P^{(1)}$, where $P^{(1)}(x) = P \circ \mathsf{swap}_{s_0,\,s_1}(x)$. In this case $\mathcal{D}$ perfectly simulates $\mathbf{H}_j^*$ for $\mathcal{A}$. Applying Lemma 3 thus gives

$$\begin{aligned}\big|\Pr[\mathcal{A}(\mathbf{H}_j) = 1] - \Pr[\mathcal{A}(\mathbf{H}_j^*) = 1]\big| &\le \sqrt{\varepsilon}\left(1 + \sqrt{q_Q + \log\left(\frac{11\,|F|}{\sqrt{\varepsilon}}\right)}\right) \\ &= \frac{1}{2^{\kappa/2}}\left(1 + \sqrt{q_Q + \log\left(\frac{11\,|\mathcal{T}|\,2^n}{2^{-\kappa/2}}\right)}\right). \quad (1)\end{aligned}$$

Next, we bound the distinguishability of $\mathbf{H}_j^*$ and $\mathbf{H}_j^{**}$. Recall that in $\mathbf{H}_j^*$ the answer to the $(j+1)$st classical query is $y_{j+1} = \mathsf{TEM\text{-}KX}_k[(P^{(1)})^{T_j,k}](t_{j+1}, x_{j+1})$, whereas in $\mathbf{H}_j^{**}$ the response is $y_{j+1} = \tilde{E}_{t_{j+1}}(x_{j+1})$. In $\mathbf{H}_j^*$, we have

$$\begin{aligned} y_{j+1} &\stackrel{\text{def}}{=} \mathsf{TEM\text{-}KX}_k[(P^{(1)})^{T_j,k}](t_{j+1}, x_{j+1}) \\ &= (P^{(1)})^{T_j,k}(s_0) \oplus f_2(t_{j+1}, P(k||0^{n-\kappa})) \\ &= P^{T_j,k}(s_1) \oplus f_2(t_{j+1}, P(k||0^{n-\kappa}))\,. \end{aligned}$$

Since s_1 is uniform and $P^{T_j,k}(\cdot) \oplus f_2(t_{j+1}, P(k||0^{n-\kappa}))$ is a permutation, we conclude 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 $\mathcal{A}$ learned from queries with 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 that $\mathcal{A}$ can distinguish $\mathbf{H}_j^*$ and $\mathbf{H}_j^{**}$. Thus,

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

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 $(P^{(1)})^{T_j,k}$ while the latter uses $P^{T_{j+1},k}$ (both for the quantum queries of $\mathcal{A}$ and to instantiate $\mathsf{TEM\text{-}KX}$ for the classical queries of $\mathcal{A}$). Thus, the two experiments are identical if $(P^{(1)})^{T_j,k}$ and $P^{T_{j+1},k}$ are equal. In what follows we upper bound the probability that they are not equal.

Both $(P^{(1)})^{T_j,k}$ and $P^{T_{j+1},k}$ involve $j+1$ swaps: $(P^{(1)})^{T_j,k}$ involves j swaps from the first j queries plus the extra swap by the definition of $P^{(1)}$, whereas $P^{T_{j+1},k}$ involves $j+1$ swaps from the first $j+1$ queries. Since the $(j+1)$st query is a forward query, we have

$$(P^{(1)})^{T_j,k}(x) = \overleftarrow{S}_{T_j,P^{(1)},k} \circ \overrightarrow{S}_{T_j,P^{(1)},k} \circ P^{(1)}(x)$$

and

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

Let $\mathcal{X} = \{x_1 \oplus f_1(t_1, P(k||0^{n-\kappa})), \ldots, x_j \oplus f_1(t_j, P(k||0^{n-\kappa}))\}$, i.e., $\mathcal{X}$ contains the inputs to P from the first j classical queries of $\mathcal{A}$. Let Bad_0 be the event that $x_{j+1} \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \in \mathcal{X}$ and Bad_1 be the event that $s_1 \in \mathcal{X}$. We upper bound the probabilities of Bad_0, Bad_1, and then show that $(P^{(1)})^{T_j,k} = P^{T_{j+1},k}$ when neither Bad_0 nor Bad_1 occurs.

Since s_1 is $\frac{j}{2^n}$-close to uniform by Eq. (2), $\Pr[\mathsf{Bad}_1] \leq \frac{2j}{2^n}$. Bounding the probability of Bad_0 is more complex since we have to consider the tweaks from the first j queries of $\mathcal{A}$. Intuitively, for queries 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-universality of f_1, f_2. Define

$$\mathcal{X}^{=} = \{x_i \oplus f_1(t_i, P(k||0^{n-\kappa})) \mid 1 \leq i \leq j,\ t_i = t_{j+1}\}$$
$$\mathcal{X}^{\neq} = \{x_i \oplus f_1(t_i, P(k||0^{n-\kappa})) \mid 1 \leq i \leq j,\ t_i \neq t_{j+1}\}.$$

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

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

where $\mathsf{Bad}_0^{=}$ is the event that $x_{j+1} \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \in \mathcal{X}^{=}$ and $\mathsf{Bad}_0^{\neq}$ is the event that $x_{j+1} \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \in \mathcal{X}^{\neq}$. For $\mathsf{Bad}_0^{=}$, we have

$$\begin{aligned} x_{j+1} \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \in \{x_i \oplus f_1(t_i, P(k||0^{n-\kappa})) \mid t_i = t_{j+1}\} \\ \Leftrightarrow x_{j+1} \in \{x_i \mid t_i = t_{j+1}\}\,. \end{aligned}$$

Since $\mathcal{A}$ does not repeat queries, this means $\Pr[\mathsf{Bad}_0^{=}] = 0$.

For $\mathsf{Bad}_0^{\neq}$, rewriting yields

$$\begin{aligned} & x_{j+1} \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \in \{x_i \oplus f_1(t_i, P(k||0^{n-\kappa})) \,|\, t_i \neq t_{j+1}\} \\ \Leftrightarrow\ & x_{j+1} \in \{x_i \oplus f_1(t_i, P(k||0^{n-\kappa})) \oplus f_1(t_{j+1}, P(k||0^{n-\kappa})) \,|\, t_i \neq t_{j+1}\}\,. \end{aligned}$$

XOR-universality of f_1, together with the fact that $f_1(t,\cdot)$ is a permutation for all t, implies that the mapping $g_{t,t'} : x \mapsto f_1(t,x) \oplus f_1(t',x)$ is a permutation whenever $t \neq t'$. Thus $g_{t_i,t_{j+1}} \circ P$ preserves the min-entropy of $k\|0^{n-\kappa}$ and $\Pr[\mathsf{Bad}_0^{\neq}] \leq |\mathcal{X}^{\neq}|/2^{\kappa} \leq j/2^{\kappa}$. Summarizing,

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

If neither Bad_0 or Bad_1 happens, we have $P^{(1)}(x_i \oplus f_1(t_i, P(k||0^{n-\kappa}))) = P(x_i \oplus f_1(t_i, P(k||0^{n-\kappa})))$ for all $1 \leq i \leq j$; furthermore, $P^{T_j,k}(s_1) = P(s_1)$ or, in other words, $P(s_1) = y_{j+1} \oplus f_2(t_{j+1}, P(k||0^{n-\kappa}))$. Therefore,

$$\begin{aligned} \overrightarrow{S}_{T_j,P^{(1)},k} &= \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{P^{(1)}(x_i \oplus f_1(t_i,P(k||0^{n-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))} \\ &= \prod_{i=1}^{j} \mathsf{swap}^{1-b_i}_{P(x_i \oplus f_1(t_i,P(k||0^{n-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))} \;=\; \overrightarrow{S}_{T_j,P,k} \end{aligned}$$

and

$$\begin{aligned} \overleftarrow{S}_{T_j,P^{(1)},k} &= \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{P^{(1)}(x_i \oplus f_1(t_i,P(k||0^{n-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))} \\ &= \prod_{i=j}^{1} \mathsf{swap}^{b_i}_{P(x_i \oplus f_1(t_i,P(k||0^{n-\kappa}))),\, y_i \oplus f_2(t_i,P(k||0^{n-\kappa}))} \;=\; \overleftarrow{S}_{T_j,P,k}, \end{aligned}$$

and so

$$\begin{aligned} (P^{(1)})^{T_j,k}(x) &= \overleftarrow{S}_{j,P^{(1)},k} \circ \overrightarrow{S}_{j,P^{(1)},k} \circ P^{(1)}(x) \\ &= \overleftarrow{S}_{j,P,k} \circ \overrightarrow{S}_{j,P,k} \\ &\quad \circ \mathsf{swap}_{P(f_1(t_{j+1},P(k||0^{n-\kappa}))\oplus x_{j+1}),\, y_{j+1}\oplus f_2(t_{j+1},P(k||0^{n-\kappa}))} \circ P(x) \\ &= \overleftarrow{S}_{j+1,P,k} \circ \overrightarrow{S}_{j+1,P,k} \circ P(x) \;=\; P^{T_{j+1},k}. \end{aligned}$$

Putting everything together, we conclude that

$$\left|\Pr[\mathcal{A}(\mathbf{H}_j^{**}) = 1] - \Pr[\mathcal{A}(\mathbf{H}'_j) = 1]\right| \leq \Pr[\mathsf{Bad}_0] + \Pr[\mathsf{Bad}_1] \leq \frac{3j}{2^{\kappa}}\,.$$

Combining the above with Eq. (1) and Eq. (2) concludes the proof of Lemma 5, and hence the proof of Theorem 3. □

Tweakable Even-Mansour. Recall that the tweakable Even-Mansour construction TEM is defined as

$$\mathsf{TEM}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t,k)) \oplus f_2(t,k)\,.$$

Setting $\kappa = n$ and noting that $P(k)$ is uniform when k is uniform (since P is a permutation), Theorem 3 yields the following as an easy corollary:

Theorem 4. *Let $\mathcal{A}$ be an adversary making q_C classical queries to its first oracle and $q_Q \geq 1$ quantum queries to its second oracle. If f_1, f_2 are proper with respect to $\mathcal{T}$, then*

$$\left| \Pr_{\substack{k\leftarrow\{0,1\}^n;\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^{\mathsf{TEM}_k^{f_1,f_2}[P],P}=1\right] - \Pr_{\substack{\tilde{E}\leftarrow\mathcal{E}(\mathcal{T},n);\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^{\tilde{E},P}=1\right]\right| \leq 7\cdot 2^{-n/2}\cdot\left(q_C\sqrt{q_Q}+q_Q\sqrt{q_C}\right).$$

(Note: Theorem 4 is a corollary of Theorem 3 only for $q_Q \geq \max\{n, \log(11\cdot|\mathcal{T}|)\}$. While small values of q_Q are not particularly interesting, Theorem 4 can be shown to hold for $q_Q \geq 1$ by a dedicated analysis that we omit here.)

4.2 Security of TEM-KX1

We also consider an alternate method of expanding a key $k \in \{0,1\}^\kappa$ to an effective key of length n, in which we compute $F_P(k) = P(k\|0^{n-\kappa}) \oplus k\|0^{n-\kappa}$. This gives rise to TEM-KX1, a variant of tweakable Even-Mansour defined as

$$\mathsf{TEM\text{-}KX1}_k^{f_1,f_2}[P](t,x) = P(x \oplus f_1(t,F_P(k))) \oplus f_2(t,F_P(k)))\,.$$

We obtain a tighter security bound for this variant than for TEM-KX; this allows us to give a tighter bound for Elephant in Section 5.2.

We first show that F_P is a pseudorandom generator, even against adversaries with quantum oracle access to P and P^{-1}.

Lemma 6. *For any quantum algorithm $\mathcal{A}$ making q_Q quantum queries,*

$$\left| \Pr_{\substack{r\leftarrow\{0,1\}^n\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^P(r)=1\right] - \Pr_{\substack{k\leftarrow\{0,1\}^\kappa\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^P(P(k||0^{n-\kappa})\oplus k||0^{n-\kappa})=1\right]\right| \leq \frac{4\cdot q_Q}{2^{\kappa/2}}\,.$$

Proof. Given an adversary $\mathcal{A}$, we construct a distinguisher $\mathcal{D}$ for the reprogramming experiment from Lemma 2:

Phase 1: $\mathcal{D}$ samples uniform $P \in \mathcal{P}_n$ and $r \in \{0,1\}^n$, and defines a randomized algorithm $\mathcal{B}$ that proceeds as follows:

1. sample uniform $k \in \{0,1\}^\kappa$;

2. output a set of reprogramming pairs B so that P blinded with B is $P^{(B)}(x) = P \circ \mathsf{swap}_{P^{-1}((k||0^{n-\kappa})\oplus r),\, k||0^{n-\kappa}}$.

Then $\mathcal{D}$ outputs P and $\mathcal{B}$.

Phase 2: $\mathcal{B}$ is run with a uniform $k \in \{0,1\}^\kappa$ to compute B. Let $P_0 = P$ and $P_1 = P^{(B)}$. A uniform $b \in \{0,1\}$ is chosen and $\mathcal{D}$ is given access to P_b (in the forward and inverse directions). $\mathcal{D}$ runs $\mathcal{A}$ with input r and oracle P_b. This phase ends when $\mathcal{A}$ has made its last query and outputs its guess.

Phase 3: $\mathcal{D}$ outputs what $\mathcal{A}$ outputs.

Note that there are at most four reprogrammed points. By construction, it holds that $\Pr_{k\leftarrow\{0,1\}^\kappa}[x \in B_1] \le 4 \cdot 2^{-\kappa}$. By Lemma 2,

$$|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1]| \le 4q_Q \cdot 2^{-\kappa/2}\,. \tag{3}$$

When $b = 0$, $\mathcal{D}$ runs $\mathcal{A}^P(r)$ for uniform and independent P, r. When $b = 1$, $\mathcal{D}$ runs $\mathcal{A}^{P_1}(r)$ where P_1 and r are each uniform but are not independent. Indeed,

$$\begin{aligned} P_1(k||0^{n-\kappa}) \oplus k||0^{n-\kappa} &= P(P^{-1}((k||0^{n-\kappa}) \oplus r)) \oplus k||0^{n-\kappa} \\ &= k||0^{n-\kappa} \oplus r \oplus k||0^{n-\kappa} = r\,. \end{aligned}$$

We prove that P_1 is uniform subject to that constraint. Let $\ell = 2^n - 1$, and let $x_1, \ldots, x_\ell$ and $y_1, \ldots, y_\ell$ be arbitrary enumerations of $X = \{0,1\}^n \setminus \{k||0^{n-\kappa}\}$ and $Y = \{0,1\}^n \setminus \{r \oplus k||0^{n-\kappa}\}$, respectively. We show that

$$\Pr[\forall i = 1, \ldots, \ell : P_1(x_i) = y_i] = \frac{1}{(2^n - 1)!}\,.$$

Letting

$$\begin{aligned} \mathbf{A} &= \Pr[P^{-1}((k||0^{n-\kappa}) \oplus r) \notin X] \\ &\quad \cdot \Pr[\forall i = 1, \ldots, \ell : P_1(x_1) = y_i \mid P^{-1}((k||0^{n-\kappa}) \oplus r) \notin X] \\ &= 2^{-n} \cdot \frac{1}{(2^n - 1)!} \;=\; \frac{1}{2^n!} \end{aligned}$$

and

$$\begin{aligned} \mathbf{B} &= \sum_{j=1}^{\ell} \Pr[P^{-1}((k||0^{n-\kappa}) \oplus r) = x_j] \\ &\qquad \cdot \Pr[\forall i \ne j : P(k||0^{n-\kappa}) = y_j \wedge P_1(x_i) = y_i \mid P^{-1}((k||0^{n-\kappa}) \oplus r) = x_j] \\ &= \sum_{j=1}^{\ell} 2^{-n} \cdot \frac{1}{(2^n - 1)!} = \frac{\ell}{2^n!} = \frac{2^n - 1}{2^n!}\,, \end{aligned}$$

we have

$$\Pr[\forall i = 1, \ldots, \ell : P_1(x_i) = y_i] = \mathbf{A} + \mathbf{B} = \frac{1}{(2^n - 1)!}\,,$$

as desired. The claimed result thus follows from Eq. (3). □

The following is an immediate corollary of Theorem 4 and Lemma 6.

Theorem 5. *Let $\mathcal{A}$ be an adversary making q_C classical queries to its first oracle and $q_Q \geq 1$ quantum queries to its second oracle. If f_1, f_2 are proper with respect to $\mathcal{T}$, then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^\kappa; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{TEM\text{-}KX1}_k^{f_1,f_2}[P],P} = 1 \right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T},n); \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{E},P} = 1 \right] \right| \leq 4 \cdot q_Q 2^{-\kappa/2} + 7 \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right).$$

5 Applications

In this section we use our results of Section 4 to show post-quantum security of the lightweight symmetric-key schemes Chaskey [15], Elephant [2], and a variant of Minalpher [17]. Note that our proofs of security hold when some public permutation at the core of each scheme is modeled as a random permutation; we do not analyze the public permutations themselves.

5.1 Chaskey

Chaskey [15] is an ISO-standardized lightweight MAC whose construction is based on a specific permutation P that we model as a random permutation. Define $\mathsf{F}^P_{k,k'}(x) = P(x \oplus k) \oplus k'$, i.e., the Even-Mansour cipher based on P. Evaluating Chaskey using key k involves evaluating $\mathsf{F}^P_{k,k}$, $\mathsf{F}^P_{k \oplus k_1, k_1}$, and $\mathsf{F}^P_{k \oplus k_2, k_2}$, 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 [15] shows that Chaskey is a secure MAC if these three instances of F^P 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 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.

As we now show, Theorem 4 readily implies 3PRP security of F in the post-quantum setting.

Theorem 6. *Let $\mathcal{A}$ be a quantum algorithm making q_C classical queries to its first three oracles and $q_Q \geq 1$ quantum queries to its fourth oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^n, \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{F}^P_{k,k}, \mathsf{F}^P_{k \oplus k_1, k_1}, \mathsf{F}^P_{k \oplus k_2, k_2}, P} = 1 \right] - \Pr_{R_1, R_2, R_3, P \leftarrow \mathcal{P}(n)} \left[\mathcal{A}^{R_1, R_2, R_3, P} = 1 \right] \right| \leq 7 \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right),$$

where $k_1 = 2k$ and $k_2 = 4k$.

Proof. 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

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

The theorem thus follows from Theorem 4 once we verify that f_1, f_2 are proper. Uniformity of f_1 and f_2 follows 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-universality follows. This concludes the proof of the theorem. □

As discussed earlier, the above theorem in combination with prior results [15] imply post-quantum security (in the random-permutation model) of Chaskey. Below we state a simple version of the theorem, leaving out some details and parameters. We formulate MAC unforgeability in terms of a distinguishing experiment in which the adversary is equipped with the Mac_k oracle, and must distinguish the oracle implementing Ver_k from the oracle (denoted by $\perp$) that always rejects. (To exclude trivial attacks, the adversary cannot forward a message/tag pair obtained from the first oracle to the second oracle.)

Theorem 7. *Let* $(\mathsf{Mac}, \mathsf{Ver})$ *be the* Chaskey *MAC, and let* $\mathcal{A}$ *be a quantum algorithm making* q_C *classical queries to its first two oracles and* q_Q *quantum queries to its third oracle. Then*

$$\begin{aligned}
&\left| \Pr_{\substack{k\leftarrow\{0,1\}^n;\\ P\leftarrow\mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{Mac}_k,\mathsf{Ver}_k,P} = 1\right] - \Pr_{\substack{k\leftarrow\{0,1\}^n\\ P\leftarrow\mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{Mac}_k,\perp,P} = 1\right] \right| \\
&\qquad \leq \mathcal{O}(2^{-n} \cdot q_C) + 7 \cdot 2^{-n/2} \left(q_C\sqrt{q_Q} + q_Q\sqrt{q_C}\right).
\end{aligned}$$

5.2 Elephant

Elephant [2] is a lightweight authenticated encryption scheme with associated data (AEAD) that was a finalist in the NIST lightweight cryptography standardization effort [18]. It is based on a tweakable block cipher we call ELE, which is constructed from a specific permutation P. Prior work [2] proves—in the purely classical setting—that Elephant is secure if ELE is a secure tweakable block cipher, and that ELE is a secure tweakable block cipher if P is modeled as a public random permutation. As with Chaskey, it is straightforward to verify that the former result carries over to the setting of quantum adversaries with classical access to Elephant if ELE is post-quantum secure.

The tweakable block cipher $\mathsf{ELE}[P] : \{0,1\}^{n-s} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ used by Elephant is defined as

$$\mathsf{ELE}[P]_k(t,x) = P(x \oplus f(t, P(k\|0^s))) \oplus f(t, P(k\|0^s))\,, \tag{4}$$

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 f and $\mathcal{T}$ is not relevant here.) Since ELE is a special case of TEM-KX where $f_1 = f_2 = f$, post-quantum security of ELE follows directly from Theorem 3.

Theorem 8. *Let* ELE *be as above and let* $\mathcal{A}$ *be an adversary making* q_C *classical queries to its first oracle and* $q_Q \geq \max\{n, \log(11 \cdot |\mathcal{T}|)\}$ *quantum queries to its second oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^n; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{ELE}[P]_k, P} = 1\right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n); \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{E}, P} = 1\right] \right| \leq 7 \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C}\right).$$

As discussed earlier, the above theorem in combination with [2, Theorem B.3] implies post-quantum security (in the random-permutation model) of Elephant. Recall that in the authenticated encryption security experiment the adversary is tasked with distinguishing the oracles $(\mathsf{Enc}_k, \mathsf{Dec}_k)$ from the pair of oracles in which the first (denoted \$) outputs random ciphertexts and the second (denoted $\perp$) always rejects. (Typical restrictions have to be imposed on the adversary to avoid trivial attacks; we do not state these here explicitly.) A fully flexible security theorem for Elephant involves many parameters; for simplicity, we record only a simple version below.

Theorem 9. *Let* $(\mathsf{Enc}, \mathsf{Dec})$ *be the* Elephant *AEAD scheme, and let* $\mathcal{A}$ *be a quantum adversary making a total of* q_C *classical queries to its first two oracles and* $q_Q \geq \max\{n, \log(11 \cdot |\mathcal{T}|)\}$ *quantum queries to its third oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^n; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{Enc}_k, \mathsf{Dec}_k, P} = 1\right] - \Pr_{P \leftarrow \mathcal{P}(n)} \left[\mathcal{A}^{\$, \perp, P} = 1\right] \right| \leq \mathcal{O}(2^{-n} \cdot q_C) + 7 \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C}\right).$$

A variant with a tighter security bound. Next, we consider a slight variant of Elephant for which we can give a tighter security bound. Recall that ELE expands the key via $k\|0^s \mapsto P(k\|0^s)$. Here we instead expand the key via $k \mapsto k||0^s \oplus P(k||0^s)$. The tweakable block cipher then becomes

$$\mathsf{ELE\text{-}KX1}[P]_k(t,x) = P(x \oplus f(t, P(k\|0^s) \oplus k||0^s)) \oplus f(t, P(k\|0^s) \oplus k||0^s)\,. \tag{5}$$

Security of the above is then a direct consequence of Theorem 5.

Theorem 10. *Let* ELE-KX1 *be as above and let* $\mathcal{A}$ *be an adversary making* q_C *classical queries to its first oracle and* $q_Q \geq 1$ *quantum queries to its second oracle. Then*

$$\left| \Pr_{\substack{k \leftarrow \{0,1\}^{n-s}; \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\mathsf{ELE\text{-}KX1}[P]_k, P} = 1 \right] - \Pr_{\substack{\tilde{E} \leftarrow \mathcal{E}(\mathcal{T}, n); \\ P \leftarrow \mathcal{P}(n)}} \left[\mathcal{A}^{\tilde{E}, P} = 1 \right] \right|$$
$$\leq 2(q_Q + q_C) \cdot \sqrt{2/2^{n-s}} + 7 \cdot 2^{-n/2} \left(q_C \sqrt{q_Q} + q_Q \sqrt{q_C} \right).$$

The above implies post-quantum security of the variant of Elephant constructed from the cipher in Eq. (5) (in place of the cipher from Eq. (4)).

5.3 (A Variant of) Minalpher

Minalpher [17] is an AEAD scheme[4] that was a second-round candidate in the CAESAR competition. Minalpher is based on a single-round tweakable Even-Mansour cipher that we call MA, which is constructed from a specific permutation P. Prior work in the purely classical setting [17] first proves that MA is a secure tweakable block cipher when P is modeled as a random permutation and then proves, as a consequence, that Minalpher is a secure AEAD scheme. Just as with Elephant and Chaskey, the latter step easily translates to the post-quantum setting if MA is secure in that setting.

We specify MA in more detail. The tweak space $\mathcal{T}$ contains tweaks of the form (flag, N, i, j), where flag is an s-bit string that takes two possible values, $N \in \{0,1\}^{n/2-s}$, and i, j are non-negative integers with $i < 2^\ell$ giving an upper bound on the message length and $j \in \{0,1,2\}$. The tweakable block cipher $\mathsf{MA} : \{0,1\}^{n/2} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ used by Minalpher is then given by

$$\mathsf{MA}_k(t,x) = P(x \oplus L(t,k))) \oplus L(t,k),$$

where

$$L((\mathsf{flag}, N, i, j), k) = y^i (y+1)^j \cdot \big(P(k||\mathsf{flag}||N) \oplus (k||\mathsf{flag}||N) \big)$$

with y some fixed element of $GF(2^n)$. Note that Minalpher pads the key with part of the tweak (in contrast to Elephant which just pads the key with 0s), which prevents us from using Theorem 3 to analyze MA. We thus consider a variant of Minalpher based on a different tweakable block cipher MA' in which the key is padded with 0s. Specifically, we set $s = 1$ so that flag is simply a bit, encode j using two bits, and then fix the lengths of N and i so their combined length is $n - 3$ bits. We then define

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

where

$$f(t,k) = (\mathsf{flag}||N||i||j) \cdot \Big(P(k||0^{n/2}) \oplus (k||0^{n/2}) \Big).$$

Since f is proper, Theorem 5 implies:

[4] Minalpher can also be used as a MAC, but here we focus on the AEAD scheme.

Theorem 11. *Let* MA' *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*

$$\left| \Pr_{\substack{k\leftarrow\{0,1\}^{n/2};\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^{\mathsf{MA}'_k,P}=1\right] - \Pr_{\substack{\tilde{E}\leftarrow\mathcal{E}(\mathcal{T},n);\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^{\tilde{E},P}=1\right]\right| \leq 2(q_Q+q_C)\cdot\sqrt{2/2^{n/2}} + 7\cdot 2^{-n/2}\left(q_C\sqrt{q_Q}+q_Q\sqrt{q_C}\right).$$

Let $\mathsf{Minalpher}'$ be the variant of $\mathsf{Minalpher}$ constructed by using MA' in place of MA. We can combine the above with classical results about the security of $\mathsf{Minalpher}$ [17] to prove post-quantum security of $\mathsf{Minalpher}'$.

Theorem 12. *Let* $(\mathsf{Enc},\mathsf{Dec})$ *be the* $\mathsf{Minalpher}'$ *AEAD scheme, and let* $\mathcal{A}$ *be a quantum adversary making a total of* q_C *classical queries to its first two oracles and* q_Q *quantum queries to its third oracle. Then*

$$\left| \Pr_{\substack{k\leftarrow\{0,1\}^{n/2};\\ P\leftarrow\mathcal{P}(n)}}\left[\mathcal{A}^{\mathsf{Enc}_k,\mathsf{Dec}_k,P}=1\right] - \Pr_{P\leftarrow\mathcal{P}(n)}\left[\mathcal{A}^{\$,\perp,P}=1\right]\right| \leq \mathcal{O}(2^{-n/2}\cdot q_C) + 2(q_Q+q_C)\cdot\sqrt{2/2^{n/2}} + 7\cdot 2^{-n/2}\left(q_C\sqrt{q_Q}+q_Q\sqrt{q_C}\right).$$

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 an NWO VENI grant (Project No. VI.Veni.192.159) and a DFF Sapere Aude grant "IM-3PQC" (Grant Id. 10.46540/2064-00034B). Work of Patrick Struck was funded by the Bavarian State Ministry of Science and the Arts in the framework of the bidt Graduate Center for Postdocs (while working at University of Regensburg) and the Hector Foundation II.

Gorjan would like to thank Yu Sasaki for suggesting to analyze Minalpher using the results of this paper.

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, 3, 4, 5, 6, 8, 12, 24

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. 4, 18, 19, 20
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. 1, 3
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. 1, 2
5. Alexandru Cojocaru, Juan Garay, and Fang Song. Generalized hybrid search and applications, 2023. Available at https://ia.cr/2023/798. 2
6. Jelle Don, Serge Fehr, and Yu-Hsuan Huang. Adaptive versus static multi-oracle algorithms, and quantum security of a split-key PRF. In *20th Theory of Cryptography Conference—TCC 2022, Part I*, volume 13747 of *LNCS*, pages 33–51. Springer, 2022. 8
7. 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. 3, 6, 27
8. Lov K. Grover. A fast quantum mechanical algorithm for database search. In *28th Annual ACM Symp. on Theory of Computing (STOC)*, pages 212–219. ACM Press, 1996. 1
9. Yassine Hamoudi, Qipeng Liu, and Makrand Sinha. Quantum-classical tradeoffs in the random oracle model, 2022. Available at https://arxiv.org/abs/2211.12954. 2
10. Akinori Hosoyamada and Yu Sasaki. Cryptanalysis against symmetric-key schemes with online classical queries and offline quantum computations. In *Topics in Cryptology—Cryptographers' Track at the RSA Conference (CT-RSA) 2018*, volume 10808 of *LNCS*, pages 198–218. Springer, 2018. 3
11. 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. 2
12. 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. 2
13. 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. IEEE, 2010. 1, 2
14. 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, 2012. 1, 2
15. Nicky Mouha, Bart Mennink, Anthony Van Herrewege, Dai Watanabe, Bart Preneel, and Ingrid Verbauwhede. Chaskey: An efficient MAC algorithm for 32-bit microcontrollers. In *Selected Areas in Cryptography (SAC)*, volume 8781 of *LNCS*, pages 306–323. Springer, 2014. 4, 18, 19
16. Ansis Rosmanis. Hybrid quantum-classical search algorithms, 2022. Available at https://arxiv.org/abs/2202.11443. 2

17. Yu Sasaki, Yosuke Todo, Kazumaro Aoki, Yusuke Naito, Takeshi Sugawara, Yumiko Murakami, Mitsuru Matsui, and Shoichi Hirose. Minalpher v1.1, 2015. Available ay https://competitions.cr.yp.to/caesar-submissions.html. 4, 18, 21, 22
18. 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, 2021. NIST IR 8369. 4, 19
19. Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In *Advances in Cryptology—Crypto 2019, Part II*, volume 11693 of *LNCS*, pages 239–268. Springer, 2019. 4

A Proof of New Resampling Lemma

We now restate and prove Lemma 3.

Lemma 7. *Let $F \subset \mathcal{P}(n)$. Consider the following experiment involving a quantum distinguisher $\mathcal{D}$:*

Phase 1: *Choose uniform $P \in \mathcal{P}(n)$, and give $\mathcal{D}$ quantum access to P. $\mathcal{D}$ outputs (D, τ), where D is a distribution on $\{0,1\}^n$ and $\tau \in F$.*

Phase 2: *Sample $\hat{s} \leftarrow D$, set $s_0 = \tau \circ P(\hat{s})$, and choose $s_1 \leftarrow \{0,1\}^n$. Let $P^{(0)} = P$ and define $P^{(1)} = P \circ \mathsf{swap}_{s_0,\, s_1}$.*

Let $\varepsilon = 2 \cdot \mathbb{E}_{(D,\tau)\leftarrow\mathcal{D}^P}\left[\max_{x\in\{0,1\}^n} \Pr_{x'\leftarrow D}[x' = x]\right]$. For any $\mathcal{D}$ making at most q queries to P in phase 1,

$$\begin{aligned}&\left|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]\right| \\ &\leq \sqrt{\varepsilon} \cdot \left(1 + \sqrt{q + \log\left(\frac{11|F|}{\sqrt{\varepsilon}}\right)}\right).\end{aligned}$$

Proof. Note that $s_1 = s_0$ then $P^{(0)} = P^{(1)}$. Thus, the distinguishing advantage of $\mathcal{D}$ is upper bounded by its distinguishing advantage conditioned on $s_1 \neq s_0$, and this is what we analyze in the rest of the proof.

Given $s_1 \neq s_0$, let $H \subset \{0,1\}^n$ be a set of size 2^{n-1} containing s_0 but not s_1, and let M be a bijection between H and $\{0,1\}^n \setminus H$ that maps s_0 to s_1. Define

$$\langle x \rangle = \begin{cases} \{x, M(x)\} & \text{if } x \in H \\ \{x, M^{-1}(x)\} & \text{if } x \notin H \end{cases}.$$

We use the plain superposition oracle for permutations as defined, e.g., by Alagic et al. [1] to simulate the permutation P. The resampling experiment with a superposition in place of P acts on quantum registers X (query input), Y (query output), E (adversary memory), and F (the oracle simulation's internal register). The oracle register F is partitioned into 2^n registers F_x, indexed by permutation inputs x. The initial state is

$$|\eta\rangle_F = (2^n!)^{-1/2} \sum_{\pi\in\mathcal{P}(n)} |\pi\rangle_F,$$

where $|\pi\rangle_F = \bigotimes_x |\pi(x)\rangle_{F_x}$.

We begin by defining a basis B_M of $\mathbb{C}\mathcal{P}(n) = \mathrm{span}\{|\pi\rangle : \pi \in \mathcal{P}(n)\}$. Define the relation $R_M \subset \mathcal{P}(n) \times \mathcal{P}(n)$ such that

$$(\pi, \sigma) \in R_M \Leftrightarrow \{\pi(x), \pi(M(x))\} = \{\sigma(x), \sigma(M(x))\} \text{ for all } x \in H\,,$$

with the corresponding equivalence classes

$$[\pi]_M = \{\sigma \in \mathcal{P}(n) : (\pi, \sigma) \in R_M\}\,.$$

We denote the set of all equivalence classes by $\mathcal{P}(n)/R_M$. For any $x, x' \in \{0,1\}^n$ and $c \in \{0,1\}$, define the quantum state

$$|\Psi^c_{x,x'}\rangle = \frac{1}{\sqrt{2}}\left(|x\rangle|x'\rangle + (-1)^c|x'\rangle|x\rangle\right)\,.$$

Define $\Gamma_M = \mathcal{P}(n)/R_M \times \{0,1\}^H$. Although Γ_M and the equivalence classes $[\pi]_M$ depend on M, we will sometimes suppress this in the notation.

For each pair $([\pi], y) \in \Gamma$ we define a vector $|([\pi], y)\rangle_F$ as follows. Let π be such that $\pi(x) > \pi(M(x))$ for all $x \in H$, where "$<$" denotes lexicographic order; we call this π the canonical representative of $[\pi]$. Define

$$|([\pi], y)\rangle_F \coloneqq \bigotimes_{x \in H} \left|\Psi^{y_x}_{\pi(x),\pi(M(x))}\right\rangle_{F_x F_{M(x)}}\,.$$

Observe that if $[\pi] = [\sigma]$ and $y = y'$ then $\langle([\pi], y) \mid ([\sigma], y')\rangle = 1$, and otherwise $\langle([\pi], y) \mid ([\sigma], y')\rangle = 0$. The set

$$B_M = \{|([\pi], y)\rangle : ([\pi], y) \in \Gamma\}$$

is thus an orthonormal set. To see that it forms a basis of $\mathbb{C}\mathcal{P}(n)$, observe that $|B_M| = |\mathcal{P}(n)|$. It follows that any state $|\varphi\rangle_{XYEF}$ can be decomposed as

$$|\varphi\rangle_{XYEF} = \sum_{([\pi],y)\in\Gamma} |\varphi([\pi], y)\rangle_{XYE} \otimes |([\pi], y)\rangle_F\,,$$

where $|\varphi([\pi], y)\rangle$ are subnormalized such that

$$\sum_{([\pi],y)\in\Gamma} \||\varphi([\pi], y)\rangle\|^2 = 1\,.$$

Define $\Gamma_j = \{([\pi], y) \in \Gamma : |y| \leq j\}$, where $|y|$ denotes Hamming weight.

Claim. Let $|\phi_q\rangle_{XYEF}$ be the global state after the (unitary part of the) distinguisher has made q queries in phase 1 to a superposition oracle initialized in any state $|\tilde{\tau}\rangle$ such that $\langle([\pi], y) \mid \tilde{\tau}\rangle = 0$ for all $y \neq 0$. Then for all y with $|y| > q$, we have $|\,\phi_q([\pi]_M, y)\,\rangle = 0$.

Proof. We prove the claim by induction on q. The base case $q = 0$ holds by assumption. For the inductive step, say the claim holds for $q-1$, and recall that

$$|\phi_q\rangle_{XYEF} = U_{XYE} O_{XYF} |\phi_{q-1}\rangle_{XYEF} \,.$$

By the induction hypothesis we can decompose

$$|\phi_{q-1}\rangle_{XYEF} = \sum_{([\pi],y)\in\Gamma_{q-1}} |\psi_{q-1}([\pi],y)\rangle_{XYE} \otimes |([\pi],y)\rangle_F \,.$$

Using this decomposition and a linearity argument, it suffices to show that for $|y| \leq q-1$, the state $O_{XYF}|x\rangle_X|y\rangle_Y|([\pi],y)\rangle_F$ is supported on basis vectors $|([\pi'],y')\rangle_F$ with $|y'| \leq q$. This follows from the fact that

$$O_{XYF}|x\rangle_X = |x\rangle_X \otimes O^{(x)}_{YF_x} \,.$$

for some operator $O^{(x)}$. This establishes the claim. □

Next, define the projector

$$\Pi_F^{\leq q} \coloneqq \sum_{([\pi],y)\in\Gamma_q} |([\pi],y)\rangle\langle([\pi],y)|_F$$

and let $\Pi^{\pm} = \frac{1}{2}(\mathbb{1} \pm \mathsf{Swap})$ be the projectors onto the symmetric and antisymmetric subspaces of $\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n}$.

We will rely on the following claim:

Claim. For any $m \in \mathbb{N}$ we have

$$\Pr_{\sigma\leftarrow\mathcal{P}(n)}[\exists \tau \in F, S \subset \{0,1\}^n \; \forall x \in S : |S| = m \wedge \tau\circ\sigma(x) \in \langle x\rangle] \leq 11 \cdot 2^{-m} \cdot |F| \,,$$

Proof. For fixed $\tau \in F$ and $S \subset \{0,1\}^n$ of size m, the number of permutations P for which $P(x) \in \langle x\rangle$ for all $x \in S$ is at most $2^m \cdot (2^n - m)!$. Thus,

$$\Pr_{\sigma\leftarrow\mathcal{P}(n)}[\forall x \in S : \tau\circ\sigma(x) \in \langle x\rangle] \leq 2^m \frac{(2^n-m)!}{2^n!} \,.$$

A union bound over all τ and S yields

$$\Pr_{\sigma\leftarrow\mathcal{P}(n)}[\exists \tau \in F, S \subset \{0,1\}^n \text{ with } |S| = m \; \forall x \in S : \tau\circ\sigma(x) \in \langle x\rangle] \leq \frac{|F|2^m}{m!} \,.$$

Using $11m! \geq 4^m$ proves the claim. □

We now return to the proof of Lemma 3. Let $\Sigma_F^{\leq m}$ be the projector onto the subspace of $\mathbb{C}\mathcal{P}(n)$ spanned by the permutations π such that

$$|\{x \in \{0,1\}^n \mid \forall \tau \in F : \tau\circ\pi(x) \in \langle x\rangle\}| \leq m \,.$$

The claim implies

$$\left\| |\eta\rangle - \frac{1}{\sqrt{\|\Sigma_F^{\leq m}|\eta\rangle\|}} \Sigma_F^{\leq m}|\eta\rangle \right\| \leq 2 \cdot \sqrt{11 \cdot 2^{-m}|F|}\,.$$

Note that $\Pi^{\leq 0}\Sigma^{\leq m}|\eta\rangle = \Sigma^{\leq m}|\eta\rangle$. We analyze the resampling experiment where the random permutation is replaced by a superposition oracle initialized with $\frac{1}{\sqrt{\|\Sigma_F^{\leq m}|\eta\rangle\|}} \Sigma_F^{\leq m}|\eta\rangle_F$.

Let $|\psi\rangle_{XYEF}$ denote the global state after phase 1, conditioned on a particular pair (D,τ) output by the distinguisher. As in [7], we can relax the task of the distinguisher as follows: instead of merely providing access to an oracle interface acting on $|\psi\rangle_{XYEF}$ for $b=0$ and $\mathsf{Swap}_{F_{s_0}F_{s_1}}|\psi\rangle_{XYEF}$ for $b=1$, we give the distinguisher arbitrary access to all registers; the distinguisher's task is then to distinguish those quantum states.

For $x \in \{0,1\}^n$, define the projector $Q^{\langle x\rangle} = \sum_{y\in\langle x\rangle} |y\rangle\langle y|$. In the following, z is a variable that corresponds to the result of measuring $F_{\hat{s}}$, i.e., $\tau(z) = s_0$. Setting

$$\Pi_{\psi,\hat{s},z} = \frac{1}{\||z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF}\|^2} |z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle\langle\psi|_{XYEF}|z\rangle\langle z|_{F_{\hat{s}}}\,,$$

it follows that

$$\begin{aligned}
&2\Pr[b=b' \mid (D,H,M),s_0] - 1\\
&\leq \frac{1}{2}\left\| \Pi_{\psi,\hat{s},z} - \mathsf{Swap}_{F_{\langle\tau(z)\rangle}} \Pi_{\psi,\hat{s},z} \mathsf{Swap}_{F_{\langle\tau(z)\rangle}} \right\|_1\\
&= \frac{1}{2}\left\| \Pi_{\psi,\hat{s},z} (\mathbb{1}-\mathsf{Swap})_{F_{\langle\tau(z)\rangle}} + (\mathbb{1}-\mathsf{Swap})_{F_{\langle\tau(z)\rangle}} \Pi_{\psi,\hat{s},z} \mathsf{Swap}_{F_{\langle\tau(z)\rangle}} \right\|_1\\
&\leq \left\| \Pi_{\psi,\hat{s},z} \Pi^-_{F_{\langle\tau(z)\rangle}} \right\|_1 + \left\| \Pi^-_{F_{\langle\tau(z)\rangle}} \Pi_{\psi,\hat{s},z} \mathsf{Swap}_{F_{\langle\tau(z)\rangle}} \right\|_1\\
&= \frac{2}{\||z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF}\|} \left\| \Pi^-_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF} \right\|_2 .
\end{aligned}$$

(The second inequality is the triangle inequality.) Taking the expectation over $\hat{s} \leftarrow D$ and z, we get

$$\begin{aligned}
&2\Pr[b=b' \mid (D,H,M)] - 1\\
&\leq 2\mathbb{E}_{\hat{s},z} \frac{1}{\||z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF}\|} \left\| \Pi^-_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF} \right\|_2\\
&\leq 2\sqrt{\mathbb{E}_{\hat{s},z} \frac{1}{\||z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF}\|} \left\| \Pi^-_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF} \right\|^2}\\
&= 2\sqrt{\sum_{\hat{s},z} D(\hat{s}) \left\| \Pi^-_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}}|\psi\rangle_{XYEF} \right\|^2}\,, \qquad (6)
\end{aligned}$$

where the first inequality is Jensen's inequality.

It remains to prove the following claim:

Claim. For any pair (D,τ) and any normalized state $|\varphi\rangle_{XYEF}$ such that

$$\Pi_{\overline{F}}^{\leq q}|\varphi\rangle_{XYEF} = |\varphi\rangle_{XYEF} \text{ and } \Sigma_{\overline{F}}^{\leq m}|\varphi\rangle_{XYEF} = |\varphi\rangle_{XYEF}\,,$$

we have

$$\sum_{\hat{s},z} D(\hat{s}) \left\| \Pi^{-}_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}} |\psi\rangle_{XYEF} \right\|^2 \leq (m+q)\varepsilon_D\,.$$

Proof. Observe that

$$\Pi^{-}\left|\Psi^0_{\pi(x),\pi(M(x))}\right\rangle = 0 \text{ and } \Pi^{-}\left|\Psi^1_{\pi(x),\pi(M(x))}\right\rangle = \left|\Psi^1_{\pi(x),\pi(M(x))}\right\rangle$$

for all x and all canonical representatives π. It follows that

$$\Pi^{-}_{F_{s_0}F_{s_1}}|\varphi\rangle_{XYEF} = \sum_{\substack{([\pi],y)\in\Gamma_q:\\ y_{s_0}=1}} |\varphi([\pi],y)\rangle_{XYE} \otimes |([\pi],y)\rangle_F\,.$$

We can now bound

$$\begin{aligned}
&\sum_{\hat{s},z} D(\hat{s}) \left\| \Pi^{-}_{F_{\langle\tau(z)\rangle}} |z\rangle\langle z|_{F_{\hat{s}}} |\psi\rangle_{XYEF} \right\|^2 \\
&\leq \sum_{\hat{s}} \sum_{z:\hat{s}\in\langle\hat{\tau}(z)\rangle} D(\hat{s}) \left\| |z\rangle\langle z|_{F_{\hat{s}}} |\psi\rangle_{XYEF} \right\|^2 \\
&+ \sum_{\hat{s}} \sum_{z:\hat{s}\notin\langle\hat{\tau}(z)\rangle} D(\hat{s}) \left\| \left(\Pi^{-}_{F_{\langle\tau(z)\rangle}} \otimes |z\rangle\langle z|_{F_{\hat{s}}}\right) |\psi\rangle_{XYEF} \right\|^2 .
\end{aligned}$$

We bound the two terms separately, beginning with the second. We decompose

$$|\psi\rangle_{XYEF} = \sum_{([\pi],y)\in\Gamma_q} |\psi([\pi],y)\rangle_{XYE} \otimes |([\pi],y)\rangle_F$$

and denote the only element of $\langle x\rangle \cap H$ by $\tilde{x}$. We have

$$\begin{aligned}
&\sum_{\hat{s}} \sum_{z:\hat{s}\notin\langle\hat{\tau}(z)\rangle} D(\hat{s}) \left\| \left(\Pi^{-}_{F_{\langle\tau(z)\rangle}} \otimes |z\rangle\langle z|_{F_{\hat{s}}}\right) |\psi\rangle_{XYEF} \right\|^2 \\
&= \sum_{\hat{s}} \sum_{z:\hat{s}\notin\langle\hat{\tau}(z)\rangle} D(\hat{s}) \sum_{([\pi],y)\in\Gamma_q} \left\| \left(\Pi^{-}_{F_{\langle\tau(z)\rangle}} \otimes |z\rangle\langle z|_{F_{\hat{s}}}\right) |\psi([\pi],y)\rangle_{XYE} \otimes |([\pi],y)\rangle_F \right\|^2 \\
&= \sum_{([\pi],y)\in\Gamma_q} \sum_{\substack{\hat{s}\notin\langle\tau\circ\pi(x)\rangle:\\ y_{\pi(\tilde{x})}=1}} D(\hat{s}) \left\| |\psi([\pi],y)\rangle_{XYE} \right\|^2 \\
&\leq \sum_{([\pi],y)\in\Gamma_q} q\varepsilon_D \left\| |\psi([\pi],y)\rangle_{XYE} \right\|^2 \;=\; q\cdot\varepsilon_D.
\end{aligned}$$

For the first term, we have $\Sigma_{\overline{F}}^{\leq m}|\varphi\rangle_{XYEF} = |\varphi\rangle_{XYEF}$, i.e., for any permutation π in the support of this state there are at most m values x such that $\tau\circ\pi(x) \in \langle x\rangle$.

For the second term, we have $\Sigma_{\overline{F}}^{\leq m}|\varphi\rangle_{XYEF} = |\varphi\rangle_{XYEF}$, i.e., $|\varphi\rangle$ is supported on basis states $|[\pi], y\rangle$ where π has at most m fixed points. Using essentially the same chain of inequalities as for the second term, we get

$$\sum_{\hat{s}} \sum_{z:\hat{s}\in\langle\hat{\tau}(z)\rangle} D(\hat{s}) \left\| |z\rangle\langle z|_{F_{\hat{s}}} |\psi\rangle_{XYEF} \right\|^2 \leq m\varepsilon_D \,.$$

This completes the proof. □

Combining the above claim with Eq. (6), taking the expectation over (D, τ), and applying Jensen's inequality one more time results in the bound

$$|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]| \leq \sqrt{(q+m)\varepsilon}$$

for the modified resampling experiment and thus

$$|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]| \leq \sqrt{(q+m)\varepsilon} + 11 \cdot 2^{-m}|F| \,.$$

Setting $m = \log\left(\frac{11|F|}{\sqrt{\varepsilon}}\right)$ we get

$$\begin{aligned} &|\Pr[\mathcal{D} \text{ outputs } 1 \mid b = 1] - \Pr[\mathcal{D} \text{ outputs } 1 \mid b = 0]| \\ &\qquad \leq \sqrt{\varepsilon}\left(1 + \sqrt{q + \log\left(11\frac{|F|}{\sqrt{\varepsilon}}\right)}\right) \,, \end{aligned}$$

matching the lemma. □