Provably Secure Reflection Ciphers

Tim Beyne and Yu Long Chen

imec-COSIC, KU Leuven, Belgium
name.lastname@esat.kuleuven.be

Abstract. This paper provides the first analysis of reflection ciphers such as PRINCE from a provable security viewpoint.
As a first contribution, we initiate the study of key-alternating reflection ciphers in the ideal permutation model. Specifically, we prove the security of the two-round case and give matching attacks. The resulting security bound takes form $\mathcal{O}(qp^2/2^{2n} + q^2/2^n)$, where q is the number of construction evaluations and p is the number of direct adversarial queries to the underlying permutation. Since the two-round construction already achieves an interesting security lower bound, this result can also be of interest for the construction of reflection ciphers based on a single public permutation.
Our second contribution is a generic key-length extension method for reflection ciphers. It provides an attractive alternative to the FX construction, which is used by PRINCE and other concrete key-alternating reflection ciphers. We show that our construction leads to better security with minimal changes to existing designs. The security proof is in the ideal cipher model and relies on a reduction to the two-round Even-Mansour cipher with a single round key. In order to obtain the desired result, we sharpen the bad-transcript analysis and consequently improve the best-known bounds for the single-key Even-Mansour cipher with two rounds. This improvement is enabled by a new sum-capture theorem that is of independent interest.

Keywords: Reflection ciphers · Public random permutations · Ideal cipher model · Sum capture theorem · PRINCE

1 Introduction

Cryptographers have long been fascinated by self-inverse, or almost self-inverse, encryption schemes. For example, the Enigma rotor machine has the surprising property that its encryption and decryption operations are identical. This feature, enabled by the middle reflector or *Umkehrwalze*, made the encryption device considerably more compact.

Although the reflector ultimately contributed to the demise of Enigma, the use of self-inverse structures was not abandoned and persists in modern cryptography. Feistel ciphers such as the DES, for instance, are equal to their own inverse up to a reordering of the round keys. Despite this property, it was later shown by Luby and Rackoff [28] and follow-up work that the generic Feistel construction is indeed sound.

Many traditional key-alternating ciphers also use involutions, i.e. self-inverse functions, as their components in order to keep the hardware implementation costs for encryption and decryption similar and to save area. The block ciphers ANUBIS [3], KHAZAD [4] and NOEKEON [16] are early examples of this strategy. Key-alternating ciphers have been extensively analyzed from the perspective of provable security [7, 13, 20, 21, 25], with results demonstrating their resistance against generic attacks. The provable security of key-alternating ciphers based on an involution instead of permutations has been studied by Lee [26].

At ASIACRYPT 2012, Borghoff et al. [8] introduced the block cipher PRINCE as an alternative approach to minimizing the overhead of supporting both efficient encryption and decryption. PRINCE has the following *reflection property*: decryption is the same as encryption using a related key. This feature is achieved by using the structure shown in Figure 1, which we will call the *key-alternating reflection cipher*. Although the use of both permutations and their inverse risks increasing area requirements, this is not a concern for the low-latency use-case that PRINCE aims for. Indeed, PRINCE targets fully unrolled hardware implementations that encrypt a plaintext in a single cycle.

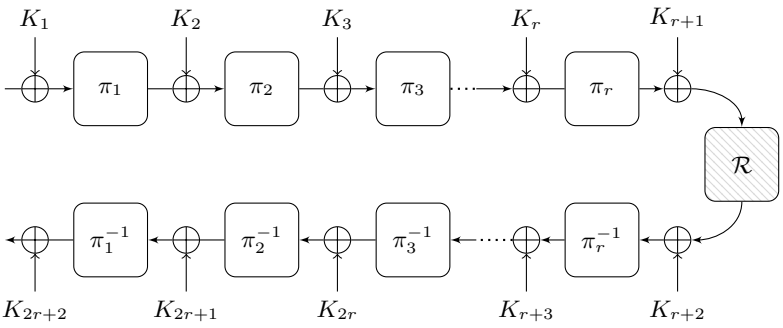

Fig. 1. A $2r$-round key-alternating reflection cipher based on r public permutations $\pi_1, \ldots, \pi_r$ and $2r+2$ keys $K_1, \ldots, K_{2r+2}$. Various key-schedules are possible. In the PRINCE core cipher $K_1 = \ldots = K_{r+1}$ and $K_{r+2} = \ldots = K_{2r+2} \oplus \alpha$ for some constant $\alpha \neq 0$. The reflector $\mathcal{R}$ is an involution.

Following increased interest in lightweight cryptography, and low-latency encryption in particular, several other key-alternating reflection ciphers were subsequently proposed. For example, PRINCESS [10] and PRINCE V2 [11] are variants of PRINCE. The tweakable block ciphers MANTIS [5] and QARMA [1] combine the key-alternating reflection cipher structure with involutive components and target applications such as memory encryption.

Despite their widespread use, the generic security of key-alternating reflection ciphers has not been analyzed from a provable security viewpoint. This stands in sharp contrast to Feistel ciphers and traditional key-alternating ciphers, which have been a regular subject of study in symmetric-key provably security. This is

remarkable, since it is natural to wonder whether or not the additional structure of reflection ciphers leads to generic flaws.

Related work. The block cipher PRINCE has been extensively analyzed from a cryptanalytic point of view, see for instance the results of the PRINCE cryptanalysis challenge which ran between 2014 and 2016 [9]. Boura et al. [10] discuss the choice of the reflector $\mathcal{R}$ and the key-schedule of general key-alternating reflection ciphers.

No results, for any number of rounds or any kind of key-schedule, are known about the provable security of key-alternating reflection ciphers. The study of traditional key-alternating ciphers, in contrast, goes back to Even and Mansour [20] for one round. The analysis of multiple rounds was initiated by Bogdanov et al. [7] and continued in [13,21,25]. Their results consider the case with independent round keys. For the two-round case, the security with three equal keys was shown by Chen et al. [12] at CRYPTO 2014.

Despite the lack of results about the provable security of key-alternating reflection ciphers, the design of PRINCE does rely on results from provable security for the purpose of key-length extension. Specifically, PRINCE uses a variant of the FX construction [24] to extend the key-length of its 64-bit core reflection cipher from 64 to 128 bits. This construction is shown in Figure 2. The designers of PRINCE prove that, under the strong assumption that E^* is an *ideal reflection cipher*, the resulting construction is secure up to the tradeoff curve $pq = 2^{128}$ with p the number of queries to E^* and q the number of construction queries. MANTIS uses the same approach to key-length extension.

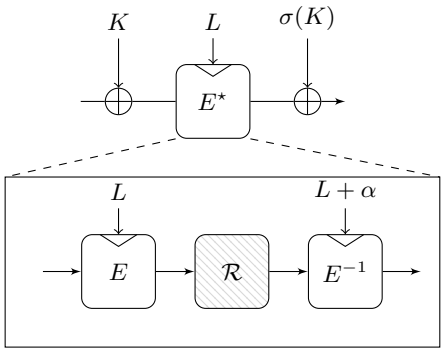

Fig. 2. The structure of PRINCE and MANTIS, with secret keys K and L, and E a block cipher. The map σ is an invertible linear map and $\mathcal{R}$ is a linear involution.

Although the construction in Figure 2 can offer reasonable security when the number of construction queries q is limited, it has been observed that the security margin offered by the $pq = 2^{128}$ tradeoff may be less comfortable than expected. In particular, at EUROCRYPT 2015, Dinur [18] proposed new time-memory-data tradeoff attacks against PRINCE. Recently, PRINCE V2 [11] was proposed

with the explicit goal of obtaining improved security with minimal changes to the original design. The approach taken by PRINCE v2 is to use alternating round keys, i.e. $K_{2i-1} = K_1$ and $K_{2i} = K_2$ for $i = 1, \ldots, r$ in Figure 1. They also slightly modify the reflector $\mathcal{R}$.

Contribution. The contribution of this paper is twofold. First, we initiate the study of the provable security of key-alternating reflection ciphers. Second, we provide a simple and generic key-extension method for reflection ciphers that achieves much better security than the FX construction.

For the first contribution, we analyze the security of the two-round variant of the general construction from Figure 1 in the ideal permutation model. Specifically, our results focus on the case with a linear reflector $\mathcal{R}$ and two alternating round keys (i.e. $K_3 = K_1$, $K_4 = K_2$), similar to the PRINCE v2 construction. Decryption is then the same as encryption up to swapping of the keys K_1 and K_2. We denote this construction by KARC2. Our Theorem 1 shows that any adaptive distinguisher making p primitive queries and q construction queries to KARC2 achieves an advantage of at most $\mathcal{O}(p^2q/2^{2n} + q^2/2^n)$. In Section 3.2, alternative key-schedules are discussed, and we show that reducing the number of round keys is nontrivial and even results in insecure constructions for many natural choices of the key-schedule.

The KARC2 construction is the first generic reflection cipher construction with a security proof. This resolves the first case of a problem of intrinsic theoretical interest, similar to the study of key-alternating ciphers. From a more practical perspective, the result limits the power of generic attacks and motivates the general soundness of a widely used construction.

Although KARC2 achieves only birthday-bound security with respect to the number of construction queries q, the best tradeoff between primitive and construction queries satisfies $p^2q = 2^{2n}$. Since the amount of data q is often limited in practice, the latter tradeoff is usually dominant. Hence, we believe the KARC2 construction could also be instantiated directly with concrete reduced-round permutations to build an attractive reflection cipher. Although many permutations are only designed to be efficient in the forward direction, there are exceptions such as Friet [32].

In Section 4, we show that Theorem 1 is tight for general choices of the reflector $\mathcal{R}$, by providing two matching generic attacks. The first attack is information-theoretic and shows that the tradeoff curve $p^2q = 2^{2n}$ cannot be improved. The second attack is a variant of the mirror slide attack of Dunkelman, Keller and Shamir [19]. It uses $\mathcal{O}(2^{n/2})$ construction queries and has a similar time-complexity. The advantage achieved by the attack is lower bounded in Theorem 2, thereby showing that the $q^2/2^n$ term in Theorem 1 can not be avoided in general. Although this may suggest that the reflector $\mathcal{R}$ is not that important from a generic viewpoint, it is important from the viewpoint of dedicated cryptanalysis (when all permutations are instantiated). Another reason for considering $\mathcal{R}$ is simply that all practical reflection ciphers have such a layer, and we want our results to say something about their generic security.

The proof of Theorem 1 is given in Section 5. It relies on Patarin's H-coefficient technique [13, 29]. The good transcript analysis resembles ideas of the first iteration of Patarin's mirror theory [30, 31], but additional difficulties appear due to the fact that the underlying permutation can be queried by the distinguisher. Note that the framework of Chen et al. [14] relies on mirror theory for two independent permutations, so it cannot be applied to KARC2, which requires the single permutation variant of mirror theory. For the secret permutation case, different techniques can be used in order to obtain domain separation [17,30]. In our proof, the domain separation is covered by a bad event, which leads to the $q^2/2^n$ term in the final security bound. The proof, like many proofs in provable security, is in an idealized model. The assumption that the primitive is ideal will never be satisfied in practice. For this reason, it is good practice to complement the provable security analysis (which rules out generic attacks) with dedicated cryptanalysis when all components are instantiated.

Our second contribution is a general method to extend the key-length of reflections ciphers, similar to the FX construction shown in Figure 2, but achieving much better security. Specifically, our proposal is to add the keys K and $\sigma(K)$ again before and after the reflector $\mathcal{R}$ respectively. For this construction, we model the block cipher E as an ideal cipher. Our Theorem 7 shows that any distinguisher making adaptively chosen plaintext and ciphertext queries to this construction achieves an advantage of at most $\widetilde{\mathcal{O}}(p\sqrt{q}/2^{n+k})$, with n the block size and k the key-length of the ideal cipher.

The proof of Theorem 7 is by a reduction to the security of the two-round Even-Mansour cipher with a single key. However, in order to be able to prove that $p^2q = 2^{2(n+k)}$ is the optimal tradeoff for our ideal cipher construction, we had to sharpen the analysis of two-round Even-Mansour by Chen et al. [12]. Hence, as a side-result that is of independent interest, we improve the best known bounds for the two-round Even-Mansour cipher with identical round keys. Figure 3 shows the difference between our new bound and the bound of Chen et al.. This result is presented in Theorem 3.

The proof of Theorem 3 is given in Section 6. Our improvement over the result of Chen et al. [12] is due to a sharpening of their bad-transcript analysis. This sharpening is made possible by an improved sum-capture theorem, which we present in Theorem 5 and prove in Section 6.1. Our sharpened sum-capture theorem is also of independent interest, as it is applicable to all other proofs relying on this result. In a nutshell, the new result removes the unnecessary discrepancy between the best-known sum-capture theorems for random functions and random permutations. Hence, we are able to avoid a term of order $p^2\sqrt{pq}/2^{2n}$ in the security bound. A detailed discussion of this result is given in Section 6.

Section 7 presents our ideal cipher construction and the proof of Theorem 7. When applied to PRINCE or MANTIS, we obtain a reflection cipher with an optimal tradeoff of $p^2q = 2^{256}$. This should be compared to the tradeoff curve $pq = 2^{128}$ for the FX construction. Hence, our construction can tolerate far more construction queries before becoming insecure. Compared to the dedicated

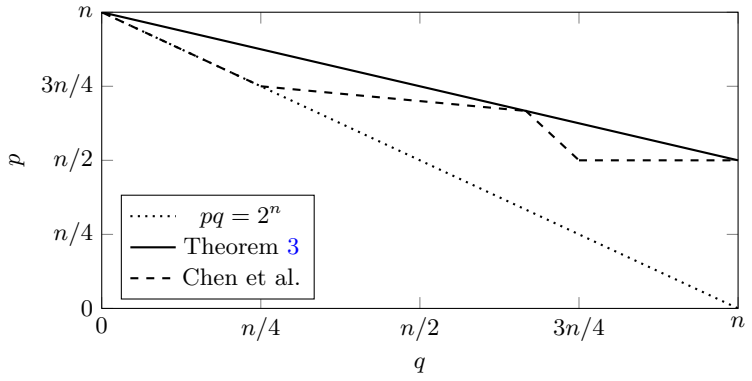

Fig. 3. Comparison between the result of Chen et al. [12] for 2-round Even-Mansour and Theorem 3. The lines correspond to an advantage upper bound equal to one.

construction PRINCE v2, it has the advantage of introducing a more minimalist change. In addition, PRINCE v2 does not completely preserve the reflection property of PRINCE due to the changes it introduces in the reflector $\mathcal{R}$.

Future work. Our work opens up several directions for interesting future research. Currently, our results only apply when two independent keys are used. Several difficulties in using a single key are discussed in Section 3.2, but we believe that using a nonlinear involution σ could resolve these issues. However, this seems to require novel proof techniques, as the sum-capture theorem requires linear mappings. Likewise, it is an open question to categorize all strong linear key schedules using two independent master keys.

Another challenging problem is that the mirror slide attack from Section 4.2 suggests that a good choice of the reflector may improve the security of KARC2, in the sense that the birthday bound term $q^2/2^n$ can be avoided. However, proving this seems difficult with state-of-the-art techniques.

A third tantalizing open problem is to generalize our results to a larger number of rounds. Namely, for $r > 1$, can we find sufficient conditions on the key-schedule such that the $2r$-round key-alternating reflection cipher achieves tight security?

It would also be interesting to reduce the time complexity of attacks against the KARC2 construction (potentially down to $\widetilde{\mathcal{O}}(2^{2n/3})$). Note that the analogous problem for two-round Even-Mansour cipher is also open, with the best attack due to Leurent and Sibleyras [27] having a time-complexity of $\mathcal{O}(2^n/\sqrt{n})$.

Another possible future research direction is to design tweakable reflection ciphers from public random permutations. Finally, it could be interesting to study the related key security of KARC2 – apart from the intentional reflection relation, and to perform cryptanalysis of concrete instances of the KARC2 construction.

2 Preliminaries

For a non-negative integer n, the set of bitstrings of length n will be denoted by $\{0,1\}^n$. For any two bitstrings $X, Y \in \{0,1\}^n$, we denote their bitwise exclusive-or as the bitstring $X \oplus Y \in \{0,1\}^n$.

For any finite set $\mathcal{S}$, the notation $S \xleftarrow{\$} \mathcal{S}$ indicates that S is a random variable uniformly distributed on $\mathcal{S}$. In particular, $\mathrm{Perm}(n)$ denotes the set of all permutations on $\{0,1\}^n$ and $\pi \xleftarrow{\$} \mathrm{Perm}(n)$ defines π as a uniform random permutation. For a list of input-output tuples $\mathcal{Q}_\pi = \{(x_1, y_1), \dots\}$, we denote by $\pi \vdash \mathcal{Q}_\pi$ the event that the permutation π is consistent with the queries-response tuples in $\mathcal{Q}_\pi$, i.e. that $\pi(x) = y$ for all $(x, y) \in \mathcal{Q}_\pi$.

Finally, for any non-negative integers $b \leq a$, the falling factorial of a with respect to b will be denoted by $(a)_b$. The value $(a)_b$ is equal to the number of injections from a set of size b to a set of size a. In particular,

$$(a)_b = \begin{cases} 1 & \text{if } b = 0, \\ a(a-1)\dots(a-b+1) & \text{otherwise.} \end{cases}$$

2.1 Block Ciphers

For non-negative integers k and n, a block cipher is a function $F\colon \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$, such that for every fixed key $K \in \{0,1\}^k$, the function $F_K(\cdot) = F(K, \cdot)$ is a permutation on $\{0,1\}^n$. The inverse of F_K will be denoted by $F_K^{-1}(\cdot) = F^{-1}(K, \cdot)$.

We will consider block ciphers F based on r public random permutations $\pi_1, \dots, \pi_r \xleftarrow{\$} \mathrm{Perm}(n)$. Our analysis of such constructions will use the strong pseudorandom permutation (sprp) security notion. Specifically, let $\mathcal{D}$ be a distinguisher with bi-directional access to either $(F_K[\pi_1, \dots, \pi_r], \pi_1, \dots, \pi_r)$ for secret key $K \xleftarrow{\$} \{0,1\}^k$, or $(\pi, \pi_1, \dots, \pi_r)$ for $\pi \xleftarrow{\$} \mathrm{Perm}(n)$. The goal of $\mathcal{D}$ is to determine which oracle it was given access to and its advantage with respect to this task is defined as

$$\mathbf{Adv}_F^{\mathrm{sprp}}(\mathcal{D}) = \left| \Pr\left[\mathcal{D}^{F_K^{\pm}[\pi_1,\dots,\pi_r],\pi_1^{\pm},\dots,\pi_r^{\pm}} = 1\right] - \Pr\left[\mathcal{D}^{\pi^{\pm},\pi_1^{\pm},\dots,\pi_r^{\pm}} = 1\right] \right| .$$

It is possible to build a new block cipher F from an ideal cipher E. The sprp security notion carries over to this case, but the distinguisher $\mathcal{D}$ is given access to the ideal cipher E rather than to r random permutations. This means that $\mathcal{D}$ can query the random permutations $F(K, \cdot)$ or its inverse for any chosen key K. Formally, let $\mathcal{D}$ be a distinguisher with bi-directional access to either $(F_K[E], E)$ for a secret key $K \xleftarrow{\$} \{0,1\}^n$, or (π, E) with $\pi \xleftarrow{\$} \mathrm{Perm}(n)$. The sprp-advantage of $\mathcal{D}$ against F is defined as

$$\mathbf{Adv}_F^{\mathrm{sprp}}(\mathcal{D}) = \left| \Pr\left[\mathcal{D}^{F_K^{\pm}[E],E^{\pm}} = 1\right] - \Pr\left[\mathcal{D}^{\pi^{\pm},E^{\pm}} = 1\right] \right| .$$

Here $\mathcal{D}^{\mathcal{O}}$ denotes the value returned by $\mathcal{D}$ when interacting with the oracle $\mathcal{O}$ and the superscript $\pm$ indicates that the distinguisher has bi-directional access.

2.2 Patarin's H-Coefficient Technique

We use the H-coefficient technique of Patarin [29], and our description of it follows the modernization of Chen and Steinberger [13].

Consider a deterministic distinguisher $\mathcal{D}$ that is given access to either a real world oracle $\mathcal{O}$ or an ideal world oracle $\mathcal{P}$. The distinguisher's goal is to determine which oracle it is given access to and we denote its advantage by

$$\mathbf{Adv}(\mathcal{D}) = \left|\Pr\left[\mathcal{D}^{\mathcal{O}} = 1\right] - \Pr\left[\mathcal{D}^{\mathcal{P}} = 1\right]\right| .$$

The query-response tuples learned by $\mathcal{D}$ during its interaction with the oracle $\mathcal{O}$ or $\mathcal{P}$ can be summarized in a transcript τ. Let $X_{\mathcal{O}}$ (respectively $X_{\mathcal{P}}$) be a random variable equal to transcript produced by the interaction between $\mathcal{D}$ and $\mathcal{O}$ (respectively $\mathcal{P}$). A particular transcript τ is called attainable if $\Pr[X_{\mathcal{P}} = \tau] > 0$ and the set of all attainable transcripts is denoted by $\mathcal{T}$.

Lemma 1 (H-coefficient technique). *Let $\mathcal{D}$ be any deterministic distinguisher. Define a partition $\mathcal{T} = \mathcal{T}_{\text{good}} \cup \mathcal{T}_{\text{bad}}$, where $\mathcal{T}_{\text{good}}$ is the subset of attainable transcripts $\mathcal{T}$ which contains all the "good" transcripts and $\mathcal{T}_{\text{bad}}$ is the subset with all the "bad" transcripts. If there exists an $\epsilon \geq 0$ such that for all attainable $\tau \in \mathcal{T}_{\text{good}}$,*

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} \geq 1 - \epsilon ,$$

then $\mathbf{Adv}(\mathcal{D}) \leq \epsilon + \Pr[X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}]$.

3 Construction Based on a Public Permutation

In this section, we consider the two-round variant of the general construction shown in Figure 1. In particular, as shown in Figure 4, we consider the case with $K_3 = K_1$ and $K_4 = K_2$ and a linear reflector $\mathcal{R}$. This case is of particular interest because it is both a natural choice for the key-schedule, and one which is used by concrete reflection ciphers such as PRINCE v2 [11]. A few alternative choices of the key-schedule are discussed in Section 3.2 below.

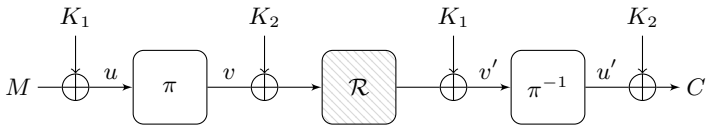

Fig. 4. The KARC2 construction based on a public permutation π and with secret keys K_1 and K_2.

The construction shown in Figure 4 will be referred to as KARC2, for *key-alternating reflection cipher* with two rounds. Formally, let n be a positive integer, $\pi \in \text{Perm}(n)$, and $\mathcal{R}\colon \{0,1\}^n \to \{0,1\}^n$ a linear involution. The generic

construction $\mathsf{KARC2}\colon \{0,1\}^{2n} \times \{0,1\}^n \to \{0,1\}^n$ is defined as

$$\mathsf{KARC2}_{K_1,K_2}[\pi](M) = \pi^{-1}(\mathcal{R}(\pi(M \oplus K_1) \oplus K_2) \oplus K_1) \oplus K_2\,.$$

The KARC2 construction has the following reflection property:

$$\left(\mathsf{KARC2}_{K_1,K_2}[\pi]\right)^{-1} = \mathsf{KARC2}_{K_2,K_1}[\pi]$$

The security of KARC2 is discussed in Section 3.1.

3.1 Security Lower Bound

In Section 5, we prove the following security bound for KARC2. As will be shown in Section 4, it is also the case that this bound is tight for general choices of the reflector $\mathcal{R}$, i.e., there are specific $\mathcal{R}$ (such as the identity) with a matching attack.

Theorem 1. *Let n be a positive integer, $\pi \xleftarrow{\$} \mathrm{Perm}(n)$ and $K_1, K_2 \xleftarrow{\$} \{0,1\}^n$. Let $\mathcal{R}$ be a linear involution on $\{0,1\}^n$. For any distinguisher $\mathcal{D}$ for $\mathsf{KARC2}_{K_1,K_2}[\pi]$ making at most q construction queries, and at most p primitive queries to $\pi^{\pm}$ such that $p + 2q < 2^{n-1}$, we have*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC2}}(\mathcal{D}) \leq \frac{3qp^2}{2^{2n}} + \frac{q^2}{2^n} + \frac{4q^{3/2}}{2^n} + \frac{4q(p+2q)(p+2q+1)}{2^{2n}}\,.$$

On the one hand, Theorem 1 'only' shows that KARC2 achieves birthday-bound security with respect to the number of construction queries q. On the other hand, it also shows that the best possible tradeoff curve between construction and primitive queries is $p^2q = 2^{2n}$ up to a small constant. This is much better than the typical birthday-bound tradeoff $pq = 2^n$. This result is especially important since in practice the number of construction queries is usually limited by the application. The number of primitive queries, however, is only limited by the computational power of the adversary.

The attacks that will be presented in Section 4 show that the term $q^2/2^n$ cannot be avoided unless the reflector $\mathcal{R}$ is carefully chosen. However, for any *linear* involution $\mathcal{R}$, there is an attack with advantage approximately $2^{-n/2}$ using $q = 2^{n/2}$ construction queries and no primitive queries. Hence, some terms independent of p cannot be avoided. It will also be shown that the term $p^2q/2^{2n}$ is tight from an information-theoretic point of view, but we are not aware of any attacks achieving the $p^2q = 2^{2n}$ tradeoff with reasonable time complexity.

3.2 Variants

The choice of the key-schedule in Figure 4 is not the only possibility. One tempting option is to further reduce the number keys by setting $K_2 = \sigma(K_1)$ for some involution σ. However, when σ is linear, this construction would not even be secure up to $q^2/2^n$ for general choices of $\mathcal{R}$. The reason is that $K_1 \oplus K_2 = K_1 \oplus \sigma(K_1)$ can then no longer be uniform random, and this significantly facilitates the attack presented in Section 4.2 below. Indeed, one has the following result.

Lemma 2. *Let n be a positive integer and $\sigma : \{0,1\}^n \to \{0,1\}^n$ a linear involution. Then σ has at least $2^{n/2}$ fixed points and the image of $\sigma \oplus \mathsf{id}$, where id is the identity function, contains at most $2^{n/2}$ distinct values.*

Proof. Since $f = \sigma \oplus \mathsf{id}$ is linear, the cardinality of its image is $2^{\dim(\operatorname{im} f)}$. Furthermore, $f^2 = 0$, so $\operatorname{im}(f) \subseteq \ker(f)$ and

$$\dim(\operatorname{im} f) \leq \dim(\ker f) = n - \dim(\operatorname{im} f)\,.$$

It follows that $\dim(\operatorname{im} f) \leq n/2$. The claim about the number of fixed points follows from the observation that the fixed points of σ are precisely the elements of $\ker f$. □

Due to the above issue, we focus on constructions with two keys. The case of one key, which necessarily requires either a special choice of $\mathcal{R}$ or a nonlinear σ, will be left as interesting (but likely challenging) future work. However, even with two keys, several constructions are possible. For example, Boura et al. [10] propose general key-schedules in which the third and fourth key-addition in Figure 4 (counting from the left) are replaced by $F_2(K_1, K_2)$ and $F_1(K_1, K_2)$ respectively, where F_1 and F_2 are (possibly nonlinear) functions. The construction we analyze is arguably the simplest secure case: $F_1(K_1, K_2) = K_1$ and $F_2(K_1, K_2) = K_2$.

4 Attacks on the Public Permutation Construction

This section shows that the security bound in Theorem 1 is essentially tight by providing two matching generic attacks. The first attack is only information theoretic and has no practical significance: it shows that the tradeoff curve $p^2q = 2^{2n}$ between the number of construction queries q and the number of primitive queries p can be achieved with a time-complexity of $\mathcal{O}(2^{2n})$. The second attack only uses construction queries and corresponds to the $q^2/2^n$ term in Theorem 1. Contrary to the first attack, the time-complexity of the second attack is limited to $\widetilde{\mathcal{O}}(2^{n/2})$ operations.

4.1 Information Theoretic Attack

Suppose the attacker makes $2q$ construction queries and p primitive queries with inputs-output pairs denoted by $(u_1, v_1), \ldots, (u_p, v_p)$. If $p^2q = 2^{2n}$, then the expected number of plaintext-ciphertext pairs (M, C) and primitive query indices (i, j) such that

$$\begin{aligned} M \oplus K_1 &= u_i \\ C \oplus K_2 &= u_j\,, \end{aligned} \tag{1}$$

is equal to two. Whenever the above conditions hold, one also has $\mathcal{R}(v_i) \oplus v_j = K_1 \oplus \mathcal{R}(K_2)$. This suggests the following method for obtaining the keys K_1 and K_2. For each possible choice of K_1 and K_2, the adversary proceeds as follows:

(i) Identify the pairs (M, C) and (i, j) for which a collision of type (1) occurs.
(ii) For each of the cases identified in step i, check that $\mathcal{R}(v_i) \oplus v_j = K_1 \oplus \mathcal{R}(K_2)$. If this relation holds for all pairs that were identified, add (K_1, K_2) to a list of candidate keys.

Since the expected number of pairs satisfying (1) is equal to two, each incorrect key (K_1, K_2) has an average probability of $1/2^{2n}$ of being accepted. Hence, the adversary obtains a list of a constant number of candidate keys. These candidate keys can be checked using a few additional queries.

The attack sketched above is purely information theoretic and does not account for the computational cost of the procedure. Since the attack uses $\mathcal{O}(2^{2n})$ table lookups, it indeed has no practical significance. Nevertheless, it shows that the $p^2q/2^{2n}$ term in Theorem 1 cannot be avoided.

Finding attacks with lower computational cost is left for future work and we believe this is an interesting problem, as the situation for the two-round Even-Mansour cipher is similar. In that case, the best known attack is due to Leurent and Sibleyras [27] and has a time-complexity of $\mathcal{O}(2^n/\sqrt{n})$ [27]. Their attack is based on a reduction to the 3-XOR problem. However, since the KARC2 construction has two keys, this approach does not help to reduce the time-complexity below $\mathcal{O}(2^n)$.

4.2 Mirror Slide Attack

The second attack is a variant of the *mirror slide* attack of Dunkelman, Keller and Shamir [19]. The attack is applicable whenever $\mathcal{R}$ has many fixed points and recovers the value of $K_1 \oplus K_2$.

The original mirror slide attack is applicable to the one-round Even-Mansour cipher with an involutive permutation. To apply a similar technique to KARC2, we let

$$\mathcal{I}(x) = \pi^{-1}(K_1 \oplus \mathcal{R}(K_2) \oplus \mathcal{R}(\pi(x)))\,.$$

The KARC2 construction can then be written as $M \mapsto \mathcal{I}(x \oplus K_1) \oplus K_2$. In general, $\mathcal{I}$ is not an involution since

$$\mathcal{I}^{-1}(x) = \pi^{-1}(K_2 \oplus \mathcal{R}(K_1) \oplus \mathcal{R}(\pi(x)))\,.$$

Nevertheless, the equation above shows that $\mathcal{I}$ is an involution iff $K_1 \oplus K_2$ is a fixed point of the reflector $\mathcal{R}$. Since by Lemma 2 any linear involution has at least $2^{n/2}$ fixed points, the mirror slide attack is applicable for a fraction of at least $2^{-n/2}$ weak keys. However, if $\mathcal{R}$ is chosen as the identity map, then all keys are weak.

The attack is based on the following observation. Let (M, C) and (M^*, C^*) be two input-output pairs for the construction such that $M \oplus C^* = K_1 \oplus K_2$ with $K_1 \oplus K_2$ a fixed point of $\mathcal{R}$. Since $M \oplus K_1 = C^* \oplus K_2$, it then follows that

$$M^* = K_1 \oplus \mathcal{I}^{-1}(C^* \oplus K_2) = K_1 \oplus \mathcal{I}(M \oplus K_1) = K_1 \oplus K_2 \oplus C\,.$$

The attack itself is then simple: choose $\Theta(2^{n/2})$ distinct values $M_1, M_2, \ldots$ and $C_1, C_2, \ldots$. With high probability, there exist indices $i \neq j$ such that $M_i \oplus C_j = K_1 \oplus K_2 = M_j \oplus C_i$. Furthermore, since the expected number of collisions is small, one obtains a short list of candidates for $K_1 \oplus K_2$.

Theorem 2 gives a lower bound on the advantage of a distinguisher based on the same principle. Hence, the security lower bound in Theorem 1 is tight in the sense that the $\mathcal{O}(q^2/2^n)$ term cannot be avoided for some choices of $\mathcal{R}$. Finding matching attacks when $\mathcal{R}$ has only $2^{n/2}$ fixed points, or improving the security lower bound in this case, will be left as future work.

Theorem 2 (Mirror slide attack). *Let $n \geq 2$ be an even integer, $\pi \xleftarrow{\$} \mathrm{Perm}(n)$, and $K_1, K_2 \xleftarrow{\$} \{0,1\}^n$. Let $\mathcal{R}$ be a linear involution on $\{0,1\}^n$ with $\ell \geq 4$ fixed points. There exists a distinguisher $\mathcal{D}$ for $\mathsf{KARC2}_{K_1,K_2}[\pi]$ making $3 \cdot 2^{n/2} + 1$ construction queries such that*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC2}}(\mathcal{D}) \geq \frac{\ell}{2^n} - \frac{4}{2^n}\,.$$

Proof. Let Δ be an arbitrary constant which is zero on the first $n/2$ bits, such as $\Delta = 0^{n-1}\|1$. The distinguisher $\mathcal{D}$ follows the approach described above, but using a slightly different approach to make the attack deterministic in the real world (assuming $K_1 \oplus K_2$ is a fixed point of $\mathcal{R}$). Specifically, $\mathcal{D}$ operates as follows:

(i) For $i = 1, \ldots, 2^{n/2}$, query $M_i = \langle i \rangle_{n/2} \,\|\, 0^{n/2}$ to obtain its encryption C_i. Likewise, query $\widetilde{M}_i = M_i \oplus \Delta$ to obtain its encryption $\widetilde{C}_i$.
(ii) For $i = 1, \ldots, 2^{n/2}$, query $C^*_i = 0^{n/2} \,\|\, \langle i \rangle_{n/2}$ to obtain M^*_i. Likewise, define $\widetilde{C}^*_i = C_i \oplus \Delta$ and denote the corresponding plaintext by $\widetilde{M}^*_i$.
(iii) If there exists a pair of indices (i, j) such that $M_i \oplus C^*_j = M^*_i \oplus C_j$ and $\widetilde{M}_i \oplus \widetilde{C}^*_j = \widetilde{M}^*_i \oplus \widetilde{C}_j$, then output 1. Otherwise, output 0.

Since in step ii only $2^{n/2} + 1$ new queries are made, the total number of queries made is $3 \cdot 2^{n/2} + 1$. The distinguisher's advantage satisfies

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC2}}(\mathcal{D}) = \left| \Pr\left[\mathcal{D}^{\mathsf{KARC2}^{\pm}_{K_1,K_2}[\mathcal{R},\pi],\pi^{\pm}} = 1\right] - \Pr\left[\mathcal{D}^{\pi^{\pm}_I,\pi^{\pm}} = 1\right] \right| .$$

Suppose that $K_1 \oplus K_2$ is a fixed point of $\mathcal{R}$. In the real world, there is a unique pair (i, j) such that $M_i \oplus C^*_j = K_1 \oplus K_2$. It then also holds that $(M_i \oplus \Delta) \oplus (C^*_j \oplus \Delta) = K_1 \oplus K_2$. Hence, as detailed in the explanation of the mirror slide attack above, the following two events then necessarily hold:

$$\begin{aligned} A_{i,j} &: \quad M_i \oplus C^*_j = M^*_j \oplus C_i \\ B_{i,j} &: \quad \widetilde{M}_i \oplus \widetilde{C}^*_j = \widetilde{M}^*_j \oplus \widetilde{C}_i\,. \end{aligned}$$

Thus, since the number of fixed points of $\mathcal{R}$ is ℓ,

$$\Pr\left[\mathcal{D}^{\mathsf{KARC2}^{\pm}_{K_1,K_2}[\mathcal{R},\pi],\pi^{\pm}} = 1\right] \geq \ell/2^n\,.$$

For the ideal world, we have

$$\Pr\left[\mathcal{D}^{\pi_I^{\pm},\pi^{\pm}}=1\right]=\Pr\left[\bigvee_{i,j} A_{i,j}\wedge B_{i,j}\right]\leq 2^n\Pr\left[A_{1,1}\wedge B_{1,1}\right]\leq\frac{4}{2^n}\,.$$

Hence, the result follows provided that $\ell\geq 4$.

5 Security Proof for the Public Permutation Construction

In this section we prove Theorem 1. Let $K_1, K_2 \xleftarrow{\$} \{0,1\}^n$ and $\pi_I, \pi \xleftarrow{\$} \mathrm{Perm}(n)$. Consider any computationally unbounded and deterministic distinguisher $\mathcal{D}$ with access to the oracles $(\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi], \pi^{\pm})$ in the real world and $(\pi_I^{\pm}, \pi^{\pm})$ in the ideal world.

The distinguisher makes q construction queries to $\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi]$ or $\pi_I^{\pm}$, and these are summarized in a transcript of the form $\tau_0 = \{(M_1, C_1), \ldots, (M_q, C_q)\}$. It also makes p primitive queries to $\pi^{\pm}$, and these are summarized in the transcript $\tau_1 = \{(u_1, v_1), \ldots, (u_p, v_p)\}$. Without loss of generality, it can be assumed that the distinguisher does not make duplicate construction or primitive queries.

After $\mathcal{D}$'s interaction with the oracles, but before it outputs its decision, we disclose the keys K_1 and K_2 to the distinguisher. This can only increase its advantage. In the real world, these are the keys used in the construction. In the ideal world, K_1 and K_2 are dummy keys drawn uniformly at random. The complete view is denoted by $\tau = (\tau_0, \tau_1, K_1, K_2)$.

5.1 Bad Events

Throughout the proof, let $U = \{u \mid (u,v) \in \tau_1\}$ and $V = \{v \mid (u,v) \in \tau_1\}$. Recall that $\mathcal{R}\colon \{0,1\}^n \to \{0,1\}^n$ is an involution, i.e. $\mathcal{R}^{-1} = \mathcal{R}$. We say that $\tau \in \mathcal{T}_{\mathrm{bad}}$ if and only if there exist construction queries $(M_i, C_i), (M_j, C_j) \in \tau_0$ and primitive queries $(u,v), (u',v') \in \tau_1$ such that one of the following conditions holds:

$$\mathrm{bad}_1\colon M_j \oplus C_i = K_1 \oplus K_2\,, \tag{2}$$
$$\mathrm{bad}_2\colon M_j \oplus u = K_1 \text{ and } C_j \oplus u' = K_2\,, \tag{3}$$
$$\mathrm{bad}_3\colon M_j \oplus u = K_1 \text{ and } \mathcal{R}(v) \oplus v' = K_1 \oplus \mathcal{R}(K_2)\,, \tag{4}$$
$$\mathrm{bad}_4\colon C_j \oplus u' = K_2 \text{ and } v \oplus \mathcal{R}(v') = \mathcal{R}(K_1) \oplus K_2\,, \tag{5}$$

When $p < q$, we also need the following two bad events for our good transcripts analysis:

$$\mathrm{bad}_5\colon \alpha_1 = |\{(M_j, C_j) \in \tau_0 \mid M_j \oplus K_1 \in U\}| \geq \sqrt{q}\,, \tag{6}$$
$$\mathrm{bad}_6\colon \alpha_2 = |\{(M_j, C_j) \in \tau_0 \mid C_j \oplus K_2 \in U\}| \geq \sqrt{q}\,. \tag{7}$$

Any attainable transcript τ for which $\tau \notin \mathcal{T}_{\mathrm{bad}}$ will be called a good transcript.

We give an informal explanation of the definition of the first four bad events. The first bad event is necessary to exclude the mirror slide attack that was

described in Section 4.2. The second bad event is exploited by the information-theoretic attack from Section 4.1. The motivation behind bad_3 and bad_4 is similar. In fact, note that $\mathcal{R}(v) \oplus v' = K_1 \oplus \mathcal{R}(K_2)$ in bad_3 and $v \oplus \mathcal{R}(v') = \mathcal{R}(K_1) \oplus K_2$ in bad_4 express the same equation. In the real world, if bad_1 does not hold, then every construction query j induces *exactly two* evaluations $(u, v), (u', v')$ of the underlying public permutation π, and these two pairs satisfy

$$\begin{aligned} M_j \oplus u &= K_1 \,, \\ C_j \ \oplus u' &= K_2 \,, \\ \mathcal{R}(v) \oplus v' &= K_1 \oplus \mathcal{R}(K_2) \,. \end{aligned}$$

Clearly, u and u' are fixed by M_j (if in the forward direction) or C_j (if in the inverse direction) and K_1, K_2, but there is "freedom" in the value $\mathcal{R}(v) \oplus v'$. If it happens to be that the distinguisher queried u, i.e., that $(u, v) \in \tau_1$, the construction query also fixes the input-output tuple (u', v'). However, in the ideal world, there is no such dependency. This means that if the adversary queries $u = M_j \oplus K_1$ and $u' = C_j \oplus K_2$ to π, with high probability the third equation would not hold. An identical reasoning applies for the case where the distinguisher happened to have set any other two out of three equations.

5.2 Probability of Bad Events in the Ideal World

We want to bound the probability $\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\mathrm{bad}}]$ that an ideal world transcript τ satisfies either of (2)-(7). Therefore, by the union bound, the probability that $X_{\mathcal{P}} \in \mathcal{T}_{\mathrm{bad}}$ can be bounded as

$$\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\mathrm{bad}}] \leq \sum_{i=1}^{6} \Pr[\mathrm{bad}_i] \,.$$

1st bad event. We first consider the bad event bad_1. Here, we rely on the randomness of $K_1 \oplus K_2$. Since K_1 and K_2 are dummy keys generated independently of τ_0 and τ_1, the probability that (2) holds for fixed i and j is $1/2^n$. Summing over q^2 possible choices of the pair (i, j), we have

$$\Pr[\mathrm{bad}_1] \leq \frac{q^2}{2^n} \,.$$

2nd bad event. We now consider the event bad_2. For any construction query $(M_j, C_j) \in \tau_0$ and any primitive queries (u, v) and (u', v'), the only randomness in the first equation of (3) is K_1 and the only randomness in the second equation is K_2. This means that the event that one of the equations defining bad_2 holds is independent of the event that the other one holds. Since the keys $K_1, K_2 \xleftarrow{\$} \{0,1\}^n$ are dummy keys generated independently of τ_0 and τ_1, the probability that bad_2 holds for a fixed choice of j, (u, v), and (u', v') is $1/2^{2n}$. Summing over the q possible construction queries and p^2 possible pairs of primitive queries, we get

$$\Pr[\mathrm{bad}_2] \leq \frac{qp^2}{2^{2n}} \,.$$

3rd bad event. Next, we consider the bad event bad_3. Note that in the second equation of (4), we can replace K_1 by $M_j \oplus u$. Hence, the only randomness in the first equation is K_1 and the only randomness in the second equation (conditional on the first) is K_2. The events that one of the equations defining bad_2 holds is therefore independent of the other. Summed over q possible construction queries and p^2 possible pairs of primitive queries, we get

$$\Pr[\text{bad}_3] \leq \frac{qp^2}{2^{2n}}\,.$$

4th bad event. The same reasoning as in the case of bad_3 applies to bad_4. Hence, it also holds that $\Pr[\text{bad}_4] \leq qp^2/2^{2n}$.

5th bad event. Finally, if $p < q$, we also consider the bad event bad_5. Note that α_1 is a random variable over the random choice of K_1, and it is independent of K_2. Furthermore, by the uniformity of K_1,

$$\mathbb{E}[\alpha_1] = \sum_{j=1}^{q} \sum_{u \in U} \Pr[M_j \oplus K_1 = u] = \frac{qp}{2^n},$$

Hence, by Markov's inequality and because we only consider this event for $p < q$,

$$\Pr[\text{bad}_5] \leq \frac{\sqrt{q}p}{2^n} \leq \frac{q^{3/2}}{2^n}\,.$$

6th bad event. The analysis of the last bad event is similarly to that of bad_5. Hence, we also have $\Pr[\text{bad}_6] \leq q^{3/2}/2^n$.

Conclusion. Summing the probabilities of the bad events, we get

$$\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}] \leq \frac{3qp^2}{2^{2n}} + \frac{q^2}{2^n} + \frac{2q^{3/2}}{2^n}\,. \tag{8}$$

This concludes the analysis of the bad transcripts in the ideal world.

5.3 Ratio for Good Transcripts

Before we continue with the proof, we present the following lemma, which will be useful in the good transcript analysis.

Lemma 3. *Let $a, b, c \geq 0$ and $N \geq 1$ be integers such that $2a + b \leq N/2$ and $2a + c + 1 \leq N/2$. Then*

$$\prod_{i=1}^{a} \frac{(N-i)(N-b-c-3i)}{(N-b-2i)(N-c-2i-1)} \geq 1 - \frac{4a(2a+b)(2a+c+1)}{N^2}$$

Proof. One has

$$\begin{aligned}
&\prod_{i=1}^{a} \frac{(N-i)(N-b-c-3i)}{(N-b-2i)(N-c-2i-1)} \\
&\quad \geq \prod_{i=1}^{a} \frac{N^2 - N(b+c+4i) - N}{N^2 - N(b+c+4i+1) + (b+2i)(c+2i+1)} \\
&\quad = \prod_{i=1}^{a} \Big(1 - \frac{(b+2i)(c+2i+1)}{N^2 - N(b+c+4i+1) + (b+2i)(c+2i+1)}\Big) \\
&\quad = \prod_{i=1}^{a} \Big(1 - \frac{(b+2i)(c+2i+1)}{(N-b-2i)(N-c-2i-1)}\Big) \\
&\quad \geq 1 - \frac{a(2a+b)(2a+c+1)}{(N-b-2a)(N-c-2a-1)} \\
&\quad \geq 1 - \frac{4a(2a+b)(2a+c+1)}{N^2}\,,
\end{aligned}$$

where for the last inequality we used $2a+b \leq N/2$ and $2a+c+1 \leq N/2$. □

Consider an attainable transcript $\tau \in \mathcal{T}_{\text{good}}$. We now lower bound $\Pr[X_{\mathcal{O}} = \tau]$ and compute $\Pr[X_{\mathcal{P}} = \tau]$ in order to obtain a lower bound for the ratio of these probabilities. For the ideal world oracle $\mathcal{P}$, the probability of any good transcript τ is equal to

$$\begin{aligned}
\Pr[X_{\mathcal{P}} = \tau] &= \frac{1}{2^{2n}} \cdot \frac{(2^n - p)!}{2^n!} \cdot \frac{(2^n - q)!}{2^n!} \\
&= \frac{1}{2^{2n}} \cdot \frac{1}{(2^n)_p} \cdot \frac{1}{(2^n)_q}\,.
\end{aligned}$$

The first factor is due to the number of possible keys K_1 and K_2. The second and third factors correspond to the probability that the uniform random permutations π and π_I are consistent with the transcripts τ_1 and τ_0 respectively.

Similarly, the real world oracle $\mathcal{O}$ is compatible with a good transcript τ if and only if it is compatible with τ_0 and τ_1. Hence,

$$\Pr[X_{\mathcal{O}} = \tau] = \frac{1}{2^{2n}} \cdot \frac{1}{(2^n)_p} \cdot \Pr[\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi] \vdash \tau_0 \mid \pi \vdash \tau_1]\,,$$

where the probability is taken with respect to $\pi \xleftarrow{\$} \mathrm{Perm}(n)$ and conditional on the keys. As before, the first factor corresponds to the number of possible keys K_1 and K_2. The second factor is the probability that π is consistent with τ_1. The third factor is the probability that the construction $\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi]$ is consistent with τ_0, given the keys K_1, K_2, and given that π is compliant with τ_1.

If we let $\rho(\tau) = \Pr[\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi] \vdash \tau_0 \mid \pi \vdash \tau_1]$, then from the above we obtain that

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} = (2^n)_q\, \rho(\tau)\,. \tag{9}$$

In order to bound $\rho(\tau)$, we re-group the construction queries in τ_0 according to their collisions with the primitive queries:

$$\begin{aligned} Q_{U_1} &= \{(M_j, C_j) \in \tau_0 \mid M_j \oplus K_1 \in U\}\,, \\ Q_{U_2} &= \{(M_j, C_j) \in \tau_0 \mid C_j \oplus K_2 \in U\}\,, \\ Q_0 &= \{(M_j, C_j) \in \tau_0 \mid M_j \oplus K_1, C_j \oplus K_2 \notin U\}\,. \end{aligned}$$

By definition, $\alpha_1 = |Q_{U_1}|$ and $\alpha_2 = |Q_{U_2}|$. Also note that $Q_{U_1} \cap Q_{U_2} = \emptyset$ by $\neg\mathrm{bad}_2$, $Q_{U_1} \cap Q_0 = \emptyset$ and $Q_{U_2} \cap Q_0 = \emptyset$ by the definition of Q_{U_1}, Q_{U_2}, and Q_0. Denote respectively by E_1, E_2, and E_0 the events that $\mathsf{KARC2}^{\pm}_{K_1,K_2}[\pi] \vdash Q_{U_1}$, Q_{U_2}, and Q_0 such that

$$\rho(\tau) = \Pr[E_1 \wedge E_2 \mid \pi \vdash \tau_1] \Pr[E_0 \mid E_1 \wedge E_2 \wedge \pi \vdash \tau_1]\,. \tag{10}$$

Lower bounding $\Pr[E_1 \wedge E_2 \mid \pi \vdash \tau_1]$. The consistency condition $\pi \vdash \tau_1$ already defines *exactly* p distinct input-output relations for π. We know that for each $(M_j, C_j) \in Q_{U_1}$, there is an unique $(u, v) \in \tau_1$ such that $M_j \oplus K_1 = u$, and $\pi(M_j \oplus K_1) = v$. We define

$$\begin{aligned} \tilde{V}_2 &= \{\mathcal{R}(\pi(M_j \oplus K_1) \oplus K_2) \oplus K_1 \colon (M_j, C_j) \in Q_{U_1}\}\,, \\ \tilde{U}_2 &= \{C_j \oplus K_2 \colon (M_j, C_j) \in Q_{U_1}\}\,. \end{aligned}$$

Similarly, for each $(M_j, C_j) \in Q_{U_2}$, there is a unique $(u, v) \in \tau_1$ such that $C_j \oplus K_2 = u$, and $\pi(C_j \oplus K_2) = v$. Again, define

$$\begin{aligned} \tilde{V}_1 &= \{\mathcal{R}(\pi(C_j \oplus K_2) \oplus K_1) \oplus K_2 \mid (M_j, C_j) \in Q_{U_2}\}\,, \\ \tilde{U}_1 &= \{M_j \oplus K_1 \mid (M_j, C_j) \in Q_{U_2}\}\,. \end{aligned}$$

Note that all values in $\tilde{U}_1$ and all values in $\tilde{V}_2$ are distinct since the M_j's are distinct, and all values in $\tilde{U}_2$ and all values in $\tilde{V}_1$ are distinct since the C_j's are distinct. We also have $\tilde{U}_1 \cap \tilde{U}_2 = \tilde{V}_1 \cap \tilde{V}_2 = \emptyset$ by $\neg\mathrm{bad}_1$, $U \cap \tilde{U}_1 = U \cap \tilde{U}_2 = \emptyset$ by $\neg\mathrm{bad}_2$, $V \cap \tilde{V}_2 = \emptyset$ by $\neg\mathrm{bad}_3$, and $V \cap \tilde{V}_1 = \emptyset$ by $\neg\mathrm{bad}_4$. Hence, the events E_1 and E_2 define *exactly* $\alpha = |Q_{U_1}| + |Q_{U_2}|$ new and distinct input-output pairs of π and it follows that

$$\Pr[E_1 \wedge E_2 \mid \pi \vdash \tau_1] = \frac{1}{(2^n - p)_\alpha}\,. \tag{11}$$

Lower bounding $\Pr[E_0 \mid E_1 \wedge E_2 \wedge \pi \vdash \tau_1]$. The conditions $\pi \vdash \tau_1$, E_1 and E_2 now define *exactly* $p' = |U \cup \tilde{U}_1 \cup \tilde{U}_2| = |V \cup \tilde{V}_1 \cup \tilde{V}_2| = p + \alpha$ distinct input-output pairs of π. Our goal now is to count the number of new distinct input-output relations for π induced by the event E_0. Recall that the event E_0 holds if and only if the reflection cipher is consistent with the construction queries in Q_0, i.e. $\mathsf{KARC2}_{K_1,K_2}[\pi] \vdash Q_0$. The queries in Q_0 can be labeled as

$$Q_0 = \{(M_{l_1}, C_{l_1}), \ldots, (M_{l_{q'}}, C_{l_{q'}})\}\,,$$

where $q' = |Q_0| = q - \alpha$ is the total number of these queries.

The event E_0 defines exactly $2q'$ relations for π of the form $\pi(\bar{u}_{2i-1}) = \bar{v}_{2i-1}$ and $\pi(\bar{u}_{2i}) = \bar{v}_{2i}$, where $\bar{u}_{2i-1} = M_{l_i} \oplus K_1$ and $\bar{u}_{2i} = C_{l_i} \oplus K_2$ for $i = 1, \ldots, q'$. By the definition of Q_0 and because bad_1 does not hold for good transcripts, it follows that

$$\{\bar{u}_1, \ldots, \bar{u}_{2q'}\} \not\subseteq U \cup \tilde{U}_1 \cup \tilde{U}_2 \,.$$

Hence, taking into account that π is a permutation, the values $\bar{v}_1, \ldots, \bar{v}_{2q'}$ must satisfy the following conditions (for $i = 1, \ldots, q'$) in the real world:

(1) $\mathcal{R}(\bar{v}_{2i-1}) \oplus \bar{v}_{2i} = K_1 \oplus \mathcal{R}(K_2)$.
(2) The variables $\bar{v}_{2i-1}$ additionally satisfy:
  (a) $\bar{v}_{2i-1} \notin V \cup \tilde{V}_1 \cup \tilde{V}_2$,
  (b) $\bar{v}_{2i-1} \notin \{\bar{v}_1, \ldots, \bar{v}_{2i-2}\}$ if $i > 1$.
(3) The variables $\bar{v}_{2i}$ additionally satisfy:
  (a) $\bar{v}_{2i} \notin V \cup \tilde{V}_1 \cup \tilde{V}_2$,
  (b) $\bar{v}_{2i} \notin \{\bar{v}_1, \bar{v}_3, \ldots, \bar{v}_{2i-1}\}$ if $i > 1$.

Observe that whenever conditions (1) and (2b) are satisfied, then it also holds that $\bar{v}_{2i} \notin \{\bar{v}_2, \bar{v}_4, \ldots, \bar{v}_{2i-2}\}$, since $K_1 \oplus \mathcal{R}(K_2)$ is a fixed value. It follows that conditions (1), (2b) and (3b) ensure that the values $\bar{v}_1, \ldots, \bar{v}_{2q'}$ are distinct.

For any positive integer $m \leq q'$, let N_m denote the number of distinct tuples $(\bar{v}_1, \ldots, \bar{v}_{2m})$ satisfying the conditions above for $i = 1, \ldots, m$. In particular, for each of the $N_{q'}$ possible consistent choices of $(\bar{v}_1, \ldots, \bar{v}_{2q'})$, the event E_0 is equivalent to exactly $2q'$ new input-output relations for π. Hence,

$$\Pr[E_0 \mid E_1 \wedge E_2 \wedge \pi \vdash \tau_1] = \frac{N_{q'}}{(2^n - p')_{2q'}} \,. \tag{12}$$

Below, a recursive formula for N_m in terms of N_{m-1} will be determined. This formula leads to a lower bound for N_m/N_{m-1}. Finally, in order to lower bound $N_{q'}$, the following telescoping product will be used ($N_0 = 1$):

$$N_{q'} = \prod_{m=1}^{q'} \frac{N_m}{N_{m-1}} \,. \tag{13}$$

Define R_m as the set of all tuples $(\bar{v}_1, \ldots, \bar{v}_{2m})$ that satisfy all conditions above for $i = 1, \ldots, m-1$ and satisfy condition (1) for $i = m$, but not (2) and (3). It is easy to see that $|R_m| = 2^n N_{m-1}$.

Furthermore, let S_m be the set of values $(\bar{v}_1, \ldots, \bar{v}_{2m})$ also satisfying all conditions for $i = 1, \ldots, m-1$, and additionally satisfying (1) and (2) but *not* (3) for $i = m$. Define T_m analogously but with values satisfying (1) and (3) but *not* (2) for $i = m$. The set of complete solutions can then be written as $R_m \setminus (S_m \cup T_m)$. Hence, by the union bound,

$$N_{2m+2} = |R_m \setminus (S_m \cup T_m)| = |R_m| - |S_m \cup T_m| \geq |R_m| - |S_m| - |T_m| \,. \tag{14}$$

Since any $(\bar{v}_1, \ldots, \bar{v}_{2m}) \in S_m$ satisfies $\bar{v}_{2m-1} \in \{\bar{v}_1, \ldots, \bar{v}_{2m-2}\} \cup V_1 \cup \tilde{V}_1 \cup \tilde{V}_2$, one has that $|S_m| \leq (p' + 2m - 2)N_{m-1}$. Similarly, $|T_m| \leq (p' + m - 1)N_{m-1}$.

Hence, substituting these inequalities and $|R_m| = 2^n N_{m-1}$ in (14) and dividing out N_{m-1} yields

$$\frac{N_m}{N_{m-1}} \geq 2^n - (p' + 2m - 2) - (p' + m - 1) = 2^n - 2p' - 3m + 3\,.$$

Using the telescoping product (13), it follows that

$$N_{q'} \geq \prod_{m=1}^{q'} (2^n - 2p' - 3m + 3) \geq \prod_{i=0}^{q'-1} (2^n - 2p' - 3i)\,.$$

Combining (10), (11) and (12), we obtain

$$\begin{aligned} \frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} &\geq N_{q'} \, \frac{(2^n)_q}{(2^n - p')_{2q'} \, (2^n - p)_\alpha} \\ &\geq \underbrace{N_{q'} \frac{(2^n)_{q'}}{(2^n - p')_{2q'}}}_{A} \cdot \underbrace{\frac{(2^n - q')_\alpha}{(2^n - p)_\alpha}}_{B}\,. \end{aligned} \tag{15}$$

Plugging in the lower bound for $N_{q'}$ in A yields

$$\begin{aligned} A &\geq \frac{\prod_{i=0}^{q'-1} (2^n - i)(2^n - 2p' - 3i)}{(2^n - p')_{2q'}} \\ &\geq \prod_{i=0}^{q'-1} \frac{(2^n - i)(2^n - 2p' - 3i)}{(2^n - p' - 2i)(2^n - p' - 2i - 1)} \\ &\geq 1 - \frac{4q'(p' + 2q')(p' + 2q' + 1)}{2^{2n}} \\ &\geq 1 - \frac{4q(p + 2q)(p + 2q + 1)}{2^{2n}}\,, \end{aligned} \tag{16}$$

where we used Lemma 3 with $a = q'$ and $b = c = p'$, and the fact that $q' \leq q$ and $p' + 2q' + 1 \leq p + 2q + 1 \leq 2^n/2$.

Next, we consider the factor B in (15). Note that for $p \geq q \geq q'$ and using the fact that $q = q' + \alpha$, we have $B \geq 1$. For $p < q$, we have

$$B \geq \frac{(2^n - q')_\alpha}{2^{\alpha n}} \geq \left(\frac{2^n - q}{2^n}\right)^\alpha \geq 1 - \frac{2q^{3/2}}{2^n}\,, \tag{17}$$

where we used $\alpha = \alpha_1 + \alpha_2 \leq 2\sqrt{q}$, which is due to $\neg\mathrm{bad}_5$, and $\neg\mathrm{bad}_6$.

Conclusion. From (15), (16), and (17) we conclude that

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} \geq 1 - \frac{4q(p + 2q)(p + 2q + 1)}{2^{2n}} - \frac{2q^{3/2}}{2^n} =: 1 - \epsilon\,,$$

using $(1 - x)(1 - y) \geq 1 - x - y$.

5.4 Conclusion

Using Patarin's H-Coefficient technique (Lemma 1), we obtain

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC2}}(\mathcal{D}) \leq \frac{3qp^2}{2^{2n+1}} + \frac{q^2}{2^n} + \frac{4q(p+2q)(p+2q+1)}{2^{2n}} + \frac{4q^{3/2}}{2^n}\,.$$

6 Sharpened Analysis of Two-Round Even-Mansour

As an intermediate result that will be used to prove the security of our ideal cipher construction, we consider the following single-key variant of the 2-round Even-Mansour cipher. For any positive integer n, let $\pi_1, \pi_2 \in \mathrm{Perm}(n)$, and let $\gamma_1, \gamma_2 \colon \{0,1\}^n \to \{0,1\}^n$ be arbitrary invertible linear maps on $\{0,1\}^n$ with respect to $\oplus$. Define the generic construction $\mathsf{EMIP2} \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ as

$$\mathsf{EMIP2}_K[\pi_1, \pi_2](M) = \pi_2(\pi_1(M \oplus K) \oplus \gamma_1(K)) \oplus \gamma_2(K)\,.$$

Chen et al. [12] showed that for $\gamma_1 = \gamma_2 = \mathsf{id}$, $\mathsf{EMIP2}$ is secure up to $\widetilde{\mathcal{O}}(2^{2n/3})$ queries. In this section, the following sharpened result will be shown. The result is sharper because, as explained below, our proof avoids the term $p^2\sqrt{qp}/2^{2n}$ in the bad transcript analysis. The latter term can play an important role when p is large. The difference between Theorem 3 and the result of Chen et al. is illustrated in Figure 3 in the introduction.

Theorem 3. *Let $n \geq 4$ be an integer, let $K \xleftarrow{\$} \{0,1\}^n$ and $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$ independent and uniform random permutations. Let $\mathcal{D}$ be any distinguisher for $\mathsf{EMIP2}_K[\pi_1, \pi_2]$ making at most $q > 1$ construction queries, at most p primitive queries to $\pi_1^{\pm}$ and at most p primitive queries to $\pi_2^{\pm}$. For all $q < 2^{n-1}$ or $q = 2^n$, we have*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{EMIP2}}(\mathcal{D}) \leq \frac{12}{2^{2c-n}} + \frac{7qp^2}{2^{2n}} + \frac{6\sqrt{3cqp^2}}{2^n}\,,$$

with $c > 0$ an arbitrary real number.

We prove Theorem 3 in Section 6.2. The bad transcript analysis of Chen et al. [12] relies on a sum-capture theorem. The sharpened bound in Theorem 3 is due to a sharpening of this result. Several variants of the sum-capture theorem exist for different situations [12,15]. These results build on the work of Babai [2] and Steinberger [33]. Typically, a sum-capture theorem states that for a random subset Z of $\{0,1\}^n$ of size q, the quantity

$$\mu(Z, A, B) = |\{(z, a, b) \in Z \times A \times B \colon z = a \oplus b\}|$$

is not much larger than $q\,|A|\,|B|\,/2^n$ for any possible choice of A and B, except with negligible probability. In our setting, Z will consist of query-response tuples from a permutation, i.e. Z consists of values $u_i \oplus v_i$ where $\{(u_1, v_1), \ldots, (u_q, v_q)\}$ is a permutation transcript. For this case, Chen et al. [12] proved the following result.

Theorem 4 (Chen et al. [12]). *Let Γ be an invertible linear map on the $\mathbb{F}_2$-vector space $\{0,1\}^n$. Let $\pi \xleftarrow{\$} \mathrm{Perm}(n)$, let $\mathcal{D}$ be some probabilistic algorithm making exactly q distinct two-sided adaptive queries to π. Let $Z = \{(u_1, v_1), \ldots, (u_q, v_q)\}$ be the transcript of the interaction of $\mathcal{D}$ with π, which consists of $q \geq 1$ pairs such that either $v_i = \pi(u_i)$ or $u_i = \pi(v_i)$ for all $i = 1, \ldots, q$. For any two subsets $A, B \subseteq \{0,1\}^n$, let*

$$\mu(Z, A, B) = \left|\{((u, v), a, b) \in Z \times A \times B \colon u \oplus a = \Gamma(v \oplus b)\}\right| .$$

Then, for $9n \leq q \leq 2^{n-1}$, we have

$$\Pr\left[\mu(Z, A, B) \geq \frac{q\,|A|\,|B|}{2^n} + \frac{2q^2\sqrt{|A|\,|B|}}{2^n} + 3\sqrt{nq\,|A|\,|B|}\right] \leq \frac{2}{2^n} .$$

In Section 6.1, we prove the following sharpened and simplified version of their result. For $c = n$, the bound in the theorem below is essentially identical to the one given in the sum-capture theorem of Cogliati et al. [15, Lemma 1] for the case where Z results from the interaction with a random *function*. Hence, our result removes the unnecessary discrepancy between the sum-capture theorems for random functions and random permutations.

Theorem 5 (Sum-capture theorem). *Let Γ be an invertible linear map on the $\mathbb{F}_2$-vector space $\{0,1\}^n$. Let $\pi \xleftarrow{\$} \mathrm{Perm}(n)$, and let $\mathcal{D}$ be some probabilistic algorithm making exactly q distinct two-sided adaptive queries to π. Let $Z = \{(u_1, v_1), \ldots, (u_q, v_q)\}$ be the transcript of the interaction of $\mathcal{D}$ with π, which consists of $q \geq 1$ pairs such that either $v_i = \pi(u_i)$ or $u_i = \pi(v_i)$ for all $i = 1, \ldots, q$. For any two subsets $A, B \subseteq \{0,1\}^n$, let*

$$\mu(Z, A, B) = \left|\{((u, v), a, b) \in Z \times A \times B \colon u \oplus a = \Gamma(v \oplus b)\}\right| .$$

For any real number $c > 0$, it then holds that

$$\Pr\left[\mu(Z, A, B) \geq \frac{q\,|A|\,|B|}{2^n} + 2\sqrt{3cq\,|A|\,|B|}\right] \leq \frac{4}{2^{2c-n}} .$$

As can be seen by comparing Theorem 4 and Theorem 5, our version of the sum-capture theorem does not contain the term $2q^2\sqrt{|A|\,|B|}/2^n$ and avoids the condition $9n \leq q \leq 2^{n-1}$. This eliminates the terms $2q^2p/2^{2n}$ and $4p^2\sqrt{qp}/2^{2n}$ in our bad transcript analysis. The latter term can play an important role when p is large.

6.1 Proof of the Sharpened Sum-Capture Theorem

For a subset Z of $\{0,1\}^n \times \{0,1\}^n$ and an invertible linear map Γ of the $\mathbb{F}_2$-vector space $\{0,1\}^n$, we define the quantity

$$\Phi_\Gamma(Z) = \max_{\substack{\alpha \in \{0,1\}^n \\ \alpha \neq 0}} \left| \sum_{(x,y) \in Z} (-1)^{\langle \alpha, x \rangle \oplus \langle \alpha, \Gamma(y) \rangle} \right| .$$

In the expression above, $\langle \alpha, x \rangle = \oplus_{i=1}^{n} \alpha_i x_i$ denotes the standard dot product between bitstrings of length n. The following lemma was proven by Chen et al. [12], but in the statement of their result they replaced the smaller quantity $\Phi_\Gamma(Z)$ by the quantity

$$\Phi(Z) = \max_{\substack{\alpha,\beta \in \{0,1\}^n \\ \alpha,\beta \neq 0}} \left| \sum_{(x,y) \in Z} (-1)^{\langle \alpha, x \rangle \oplus \langle \beta, y \rangle} \right| \geq \Phi_\Gamma(Z) \,.$$

However, their proof carries over essentially completely.

Lemma 4 (Chen et al. [12]). *Let Γ be an automorphism of the $\mathbb{F}_2$-vector space $\{0,1\}^n$. For all sets $Z \subseteq \{0,1\}^n \times \{0,1\}^n$ and $A, B \subseteq \{0,1\}^n$, define*

$$\mu(Z, A, B) = |\{((u, v), a, b) \in Z \times A \times B \colon u \oplus a = \Gamma(v \oplus b)\}| \,.$$

Then it holds that

$$\mu(Z, A, B) \leq \frac{|Z|\,|A|\,|B|}{2^n} + \Phi_\Gamma(Z)\sqrt{|A|\,|B|} \,.$$

In order to obtain the simplified sum-capture theorem, it suffices to compute a tail bound for the quantity $\Phi_\Gamma(Z)$. Our improvement over the result of Chen et al. is enabled by the following theorem of Hoeffding [22], which is stated for the special case of zero-mean uniformly bounded populations below.

Theorem 6 (Hoeffding [22]). *If $x_1, x_2, \ldots, x_q$ is a random sample without replacement from a finite population (multiset) $\{\{c_1, c_2, \ldots, c_N\}\}$ such that $a \leq c_i \leq b$ for all $i = 1, \ldots, N$ and $\sum_{i=1}^{N} c_i = 0$, then for all $\delta > 0$, it holds that*

$$\Pr\left[\sum_{i=1}^{q} x_i \geq \sqrt{q}\delta\right] \leq \exp\left(\frac{-2\delta^2}{(b-a)^2}\right) \,.$$

Theorem 6 is precisely the same bound as the classical Hoeffding inequality for sampling *with* replacement [22, Theorem 2]. It is not surprising that the same result should be true for sampling without replacement, since the latter tends to decrease variability. To prove Theorem 6, Hoeffding first showed that the average of any continuous convex function of $\sum_{i=1}^{q} x_i$ is less than the same function of an equivalent sum involving random variables sampled with replacement. The result then follows by applying this argument for the exponential function (which is clearly convex) and by using Markov's inequality.

Lemma 5. *Let $\pi \xleftarrow{\$} \mathrm{Perm}(n)$ and let $\mathcal{D}$ be some probabilistic algorithm making exactly q distinct two-sided adaptive queries to π. Let $Z = \{(u_1, v_1), \ldots, (u_q, v_q)\}$ be the transcript of the interaction of $\mathcal{D}$ with π, which consists of $q \geq 1$ pairs such that $v_i = \pi(u_i)$ or $u_i = \pi(v_i)$. For any real number $c > 0$, the tail of $\Phi_\Gamma(Z)$ can be bounded as*

$$\Pr[\Phi_\Gamma(Z) \geq 2\sqrt{3cq}] \leq \frac{4}{2^{2c-n}} \,.$$

Proof. By swapping inputs and outputs where necessary for $i = 1, \ldots, q$, there exist pairs (x_i, y_i) such that $y_i = \pi(x_i)$ and

$$\Phi_\Gamma(Z) = \max_{\substack{\alpha \in \{0,1\}^n \\ \alpha \neq 0}} \left| \sum_{i=1}^{q} (-1)^{\langle \alpha, x_i \rangle \oplus \langle \alpha, \Gamma(y_i) \rangle} \right| .$$

For any $\alpha \neq 0$ the values $z_i = \langle \alpha, \Gamma(y_i) \rangle$ with $i = 1, \ldots, q$ are random samples *without replacement* from a population consisting of 2^{n-1} values 0 and 2^{n-1} values 1. Indeed, any nonzero linear combination of the output bits of a uniform random permutation is a uniform random balanced Boolean function and no queries to π can be repeated. Furthermore, due to the fact that π is a uniform random permutation, $z_1, \ldots, z_q$ are independent of $x_1, \ldots, x_q$. Hence, consider the sum

$$S_\alpha = \sum_{i=1}^{q} (-1)^{\langle \alpha, x_i \rangle} (-1)^{z_i} .$$

Note that S_α is a symmetric random variable and $\mathbb{E}[S_\alpha] = 0$. Applying the union bound[1] and Theorem 6 to the terms with positive and negative coefficients separately gives the tail bound

$$\Pr\left[|S_\alpha| \geq \delta \sqrt{q} \mid x_1, \ldots, x_q \right] \leq 4\, e^{-\delta^2/8} .$$

The law of total probability then directly yields the upper bound $\Pr\left[|S_\alpha| \geq \delta\sqrt{q}\right] \leq 4\, e^{-\delta^2/8}$. By the union bound,

$$\Pr\left[\Phi_\Gamma(Z) \geq \delta\sqrt{q}\right] = \Pr\left[\max_{\alpha \neq 0} |S_\alpha| \geq \delta\sqrt{q} \right] \leq 2^{n+2}\, e^{-\delta^2/8} .$$

Let $\delta = 2\sqrt{3c} > 4\sqrt{\ln 2^c}$ for $c > 0$, then

$$\Pr\left[\Phi_\Gamma(Z) \geq 2\sqrt{3cq}\right] \leq 2^{n+2}\, e^{-2 \ln 2^c} = \frac{4}{2^{2c-n}} .$$

This concludes the proof. □

6.2 Proof of Theorem 3

In this section we prove Theorem 3. Let $K \xleftarrow{\$} \{0,1\}^n$ and $\pi_I, \pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$. Consider any computationally unbounded and deterministic distinguisher $\mathcal{D}$ with access to the oracles $(\mathsf{EMIP2}_K^{\pm}[\pi_1, \pi_2], \pi_1^{\pm}, \pi_2^{\pm})$ in the real world and $(\pi_I^{\pm}, \pi_1^{\pm}, \pi_2^{\pm})$ in the ideal world.

The distinguisher makes q construction queries to $\mathsf{EMIP2}_K^{\pm}[\pi_1, \pi_2]$ or $\pi_I^{\pm}$, and these are summarized in a transcript of the form $\tau_0 = \{(M_1, C_1), \ldots, (M_q, C_q)\}$. It also makes p primitive queries to $\pi_1^{\pm}$, and p primitive queries to $\pi_2^{\pm}$, these are respectively summarized in the transcript $\tau_1 = \{(u_1, v_1), \ldots, (u_p, v_p)\}$ and

[1] In the form $\Pr[X + Y \geq t] \leq \Pr[X \geq t/2] + \Pr[Y \geq t/2]$.

$\tau_2 = \{(x_1, y_1), \ldots, (x_p, y_p)\}$. Without loss of generality, it can be assumed that the distinguisher does not make duplicate construction or primitive queries.

After $\mathcal{D}$'s interaction with the oracles, but before it outputs its decision, we disclose the key K to the distinguisher. In the real world, this is the key used in the construction. In the ideal world, K is a dummy key that is drawn uniformly at random. The complete view is denoted by $\tau = (\tau_0, \tau_1, \tau_2, K)$.

Bad events. We say that $\tau \in \mathcal{T}_{\text{bad}}$ if and only if there exist a construction query $(M_j, C_j) \in \tau_0$ and primitive queries $(u, v) \in \tau_1$ and $(x, y) \in \tau_2$ such that one of the following conditions holds:

$$\text{bad}_1 \colon M_j \oplus u = K \text{ and } C_j \oplus y = \gamma_2(K)\,, \tag{18}$$
$$\text{bad}_2 \colon M_j \oplus u = K \text{ and } v \oplus x = \gamma_1(K)\,, \tag{19}$$
$$\text{bad}_3 \colon C_j \oplus y = \gamma_2(K) \text{ and } v \oplus x = \gamma_1(K)\,. \tag{20}$$

Any attainable transcript τ for which $\tau \notin \mathcal{T}_{\text{bad}}$ will be called a good transcript.

Probability of bad events in the ideal world. We want to bound the probability $\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}]$ that an ideal world transcript τ satisfies either of (18)-(20). Therefore, the probability that $X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}$ is given by

$$\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}] \le \Pr[\text{bad}_1] + \Pr[\text{bad}_2] + \Pr[\text{bad}_3]\,.$$

Throughout the proof, let $U = \{u \mid (u, v) \in \tau_1\}$, $V = \{v \mid (u, v) \in \tau_1\}$, $X = \{x \mid (x, y) \in \tau_2\}$ and $Y = \{y \mid (x, y) \in \tau_2\}$. In addition, denote

$$\begin{aligned}
\Omega_1 &= \left|\left\{\big(j, (u, v), (x, y)\big) \mid M_j \oplus u = \gamma_2^{-1}(C_j \oplus y)\right\}\right|\,,\\
\Omega_2 &= \left|\left\{\big(j, (u, v), (x, y)\big) \mid M_j \oplus u = \gamma_1^{-1}(v \oplus x)\right\}\right|\,,\\
\Omega_3 &= \left|\left\{\big(j, (u, v), (x, y)\big) \mid C_j \oplus y = \gamma_2 \circ \gamma_1^{-1}(v \oplus x)\right\}\right|\,.
\end{aligned}$$

In the ideal world, Ω_1, Ω_2, and Ω_3 only depend on π_1, π_2 and π, and not on the key $K \xleftarrow{\$} \{0,1\}^n$, which is drawn uniformly at random at the end of the interaction. For any $i \in \{1, 2, 3\}$ and $\lambda_i > 0$ a real constant, we have

$$\Pr[\text{bad}_i] \le \Pr[\Omega_i \ge \lambda_i] + \frac{\lambda_i}{2^n}\,.$$

To upper bound the first term above, the sharpened sum-capture theorem (Theorem 5) will be used. This application of the sum-capture theorem will also rely on the linearity of γ_1 and γ_2.

1st bad event. The first bad event can be rewritten as $M_j \oplus u = \gamma_2^{-1}(C_j) \oplus \gamma_2^{-1}(y) = K$. To apply the sum-capture lemma, define

$$\begin{aligned}
Z_1 &= \{M_j \oplus \gamma_2^{-1}(C_j) \mid (M_j, C_j) \in \tau_0\}\,,\\
A_1 &= U\,,\\
B_1 &= \{\gamma_2^{-1}(y) \mid y \in Y\}\,.
\end{aligned}$$

Since γ_2^{-1} is a permutation, Lemma 4 can be applied with $\Omega_1 = \mu(Z_1, A_1, B_1)$,

$$\Pr\left[\mu(Z_1, A_1, B_1) \geq \frac{qp^2}{2^n} + 2\sqrt{3cqp^2}\right] \leq \frac{4}{2^{2c-n}}\,.$$

We thus set $\lambda_1 = qp^2/2^n + 2\sqrt{3cqp^2}$ and obtain

$$\Pr[\mathrm{bad}_1] \leq \frac{4}{2^{2c-n}} + \frac{qp^2}{2^{2n}} + \frac{2\sqrt{3cqp^2}}{2^n}\,.$$

2^{nd} bad event. For $i = 2$, we rewrite bad_2 as $M_j \oplus u = \gamma_1^{-1}(v) \oplus \gamma_1^{-1}(x) = K$, and we define

$$\begin{aligned} Z_2 &= \{u \oplus \gamma_1^{-1}(v) \mid (u, v) \in \tau_1\}\,, \\ A_2 &= \{M_j \mid (M_j, C_j) \in \tau_0\}\,, \\ B_2 &= \{\gamma_1^{-1}(x) \mid x \in X\}\,. \end{aligned}$$

Then, since γ_1^{-1} is a permutation, we can apply Lemma 4 with $\Omega_2 = \mu(Z_2, A_2, B_2)$,

$$\Pr\left[\mu(Z_2, A_2, B_2) \geq \frac{qp^2}{2^n} + 2\sqrt{3cqp^2}\right] \leq \frac{4}{2^{2c-n}}\,.$$

We thus set $\lambda_2 = qp^2/2^n + 2\sqrt{3cqp^2}$ and obtain

$$\Pr[\mathrm{bad}_2] \leq \frac{4}{2^{c-2n}} + \frac{qp^2}{2^{2n}} + \frac{2\sqrt{3cqp^2}}{2^n}\,.$$

3^{rd} bad event. For $i = 3$, we rewrite bad_3 as $C_j \oplus y = \gamma_2 \circ \gamma_1^{-1}(v) \oplus \gamma_2 \circ \gamma_1^{-1}(x) = \gamma_2(K)$ and we define

$$\begin{aligned} Z_3 &= \{\gamma_2 \circ \gamma_1^{-1}(x) \oplus y \mid (x, y) \in \tau_2\}\,, \\ A_3 &= \{C_j \mid (M_j, C_j) \in \tau_1\}\,, \\ B_3 &= \{\gamma_2 \circ \gamma_1^{-1}(v) \mid v \in V\}\,. \end{aligned}$$

Then, since $\gamma_2 \circ \gamma_1^{-1}$ is a permutation, we can apply Lemma 4 with $\Omega_3 = \mu(Z_3, A_3, B_3)$,

$$\Pr\left[\mu(Z_3, A_3, B_3) \geq \frac{qp^2}{2^n} + 2\sqrt{3cqp^2}\right] \leq \frac{4}{2^{2c-n}}\,.$$

We thus set $\lambda_3 = qp^2/2^n + \sqrt{5nqp^2}$ and obtain

$$\Pr[\mathrm{bad}_3] \leq \frac{4}{2^{2c-n}} + \frac{qp^2}{2^{2n}} + \frac{2\sqrt{3cqp^2}}{2^n}\,.$$

Conclusion. Summing the probabilities of the three bad events, we get

$$\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\mathrm{bad}}] \leq \frac{12}{2^{2c-n}} + \frac{3qp^2}{2^{2n}} + \frac{6\sqrt{3cqp^2}}{2^n}\,. \tag{21}$$

Probability ratio for good transcripts. Since our bad events are the same as in the analysis of Chen et al. [12], their analysis of the good transcript ratio can be recycled. In particular, their Lemma 8 (i) implies that for any good transcript τ and any integers q and p such that $2q + 2p \leq 2^n$,

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} \in \tau]} \geq 1 - \frac{4qp^2}{2^{2n}} .$$

However, the above bound is trivial whenever $p \geq 2^{n-1}/\sqrt{q}$. Hence, $2q + 2p \leq 2^n/\sqrt{q} + 2q$ and for $n \geq 4$ this is lower than 2^n whenever $q > 1$ and $q < 2^{n-1}$. Furthermore, by [12, Lemma 8 (ii)], the result also holds for $q = 2^n$.

Conclusion. Using Patarin's H-Coefficient technique (Lemma 1), we obtain

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{EMIP2}}(\mathcal{D}) \leq \frac{12}{2^{2c-n}} + \frac{3qp^2}{2^{2n}} + \frac{6\sqrt{3cqp^2}}{2^n} + \frac{4qp^2}{2^{2n}} .$$

7 Construction Based on an Ideal Cipher

We now turn to our second reflection cipher construction, which is illustrated in Figure 5 below. Theorem 7 will show that, for an n-bit ideal block cipher with a k-bit key, this construction achieves a $\widetilde{\mathcal{O}}(p\sqrt{q}/2^{n+k})$ security bound. The proof of this result is based on a reduction to our sharpened security bound for the two-round Even-Mansour cipher from Theorem 3.

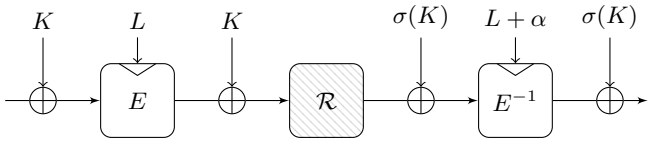

Fig. 5. The KARC-IC construction uses two secret keys K and L, and a block cipher E. The reflector $\mathcal{R}$ is a fixed linear involution and σ is an invertible linear map. To obtain a pure reflection property with respect to both keys, σ should be an involution.

Although the construction in Figure 5 is based on the more powerful ideal cipher model, it is of considerable practical interest. Indeed, block-ciphers such as PRINCE [8], MANTIS [5] and QARMA [1] are designed to support a 64 bit block size with 128 bit keys (internally split into two 64 bit keys), and claim a security tradeoff of $pq = 2^{128}$.

In the case of PRINCE and MANTIS, this is achieved by instantiating the XEX-construction [23] with an ideal reflection cipher. Their construction is shown in Figure 2 (in the introduction). Importantly, although this achieves the desired tradeoff, the construction of the ideal reflection cipher E^* in PRINCE and MANTIS closely follows our proposed construction: the only difference is the presence

of key-additions in the middle layer of our construction. Hence, by minimally modifying PRINCE and MANTIS, our results show that an improved security tradeoff of $pq^2 = 2^{256}$ can be achieved. However, it should be stressed that our results only establish security against *generic* attacks. Careful analysis by cryptanalysts remains necessary, even for minor changes such as the one proposed by our construction. For instance, in the case of MANTIS, reduced-round nonlinear invariant attacks have been discovered [6]. The presence of key additions in the middle could provide additional flexibility to propagate the invariant property over more rounds. We believe a detailed analysis of this case would make for interesting future work.

The design of QARMA follows a very similar approach to our construction. In fact, Avanzi [1] remarks that the true security of the QARMA construction is likely to exceed the claimed $pq = 2^n$ trade-off. Our results corroborate this to some extent. However, our Theorem 7 is not directly applicable because QARMA uses a nonlinear reflector $\mathcal{R}$ between the middle key-additions. Analyzing the security of such construction would be possible if the sum-capture theorem could be extended to allow for nonlinearity. This is an interesting problem by itself.

Before giving Theorem 7 and its proof, we formalize our second construction. For any positive integers n and k, let E be a block cipher with key $L \in \{0,1\}^k$, and let $K \in \{0,1\}^n$ be a second construction key. Furthermore let $\mathcal{R}$ be a linear involution and σ an invertible linear map on $\{0,1\}^n$ such that $\mathrm{id} + \mathcal{R} \circ \sigma$ is invertible. The generic construction $\mathsf{KARC\text{-}IC2}\colon \{0,1\}^{n+k} \times \{0,1\}^n \to \{0,1\}^n$ is defined by

$$\mathsf{KARC\text{-}IC2}_{K,L}[E](M) = E^{-1}_{L+\alpha}(\mathcal{R}(E_L(M \oplus K) \oplus K) \oplus \sigma(K)) \oplus \sigma(K)\,, \qquad (22)$$

with $\alpha \in \{0,1\}^k$ a nonzero constant. The condition that $\mathrm{id} + \mathcal{R} \circ \sigma$ is invertible is an important one, since Theorem 3 requires that γ_2 is invertible. Note that this condition is equivalent to the requirement that $\mathcal{R} \circ \sigma$ does not have any fixed points. The security of $\mathsf{KARC\text{-}IC2}$ is given in Theorem 7, which can be proven by a reduction to the security of $\mathsf{EMIP2}$.

Theorem 7. *For any positive integers $n \geq 2$ and k, let $K \xleftarrow{\$} \{0,1\}^n$ and $L \xleftarrow{\$} \{0,1\}^k$ be uniform random keys and E an ideal cipher. If $\mathcal{D}$ is any distinguisher for $\mathsf{KARC\text{-}IC2}_{K,L}[E]$ making at most $q > 1$ construction queries, and at most p primitive queries to $E^{\pm}$, then for all $q < 2^{n-1}$ or $q = 2^n$ it holds that:*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC\text{-}IC2}}(\mathcal{D}) \leq \frac{12}{2^{n+k}} + 9\sqrt{2n+k}\,\frac{\sqrt{q}\,p}{2^{n+k}}\,.$$

Proof. Enumerate all $\ell = 2^k$ possible ideal cipher keys as $L_1, \ldots, L_\ell$. Suppose the distinguisher $\mathcal{D}$ makes $p_{1,i}$ queries to $E^{\pm 1}$ with key L_i. Likewise, let $p_{2,i}$ denote the number of queries to $E^{\pm 1}$ with key $L_i \oplus \alpha$. For convenience, let $p_i = \max\{p_{1,i}, p_{2,i}\}$ be the maximum number of queries made for either L_i or $L_i \oplus \alpha$. Since the total number of queries is equal to p, we have

$$\sum_{i=1}^{\ell} p_i \leq \sum_{i=1}^{\ell} p_{1,i} + p_{2,i} = 2p\,.$$

It follows from the law of total probability and the triangle inequality that

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC\text{-}IC2}}(\mathcal{D}) \le \sum_{i=1}^{\ell} \frac{1}{\ell}\, \mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC\text{-}IC2}_{K,L_i}[E]}(\mathcal{D})\,.$$

Let $\mathcal{D}_i$ be a distinguisher running $\mathcal{D}$ to play the indistinguishability game against the $\mathsf{EMIP2}_K[\pi_1,\pi_2]$ construction with $\pi_1 = E_{L_i}$ and $\pi_2 = E^{-1}_{L_i\oplus\alpha}$ using $p_{1,i}$ primitive queries to π_1, $p_{2,i}$ primitive queries to π_2 and q construction queries. In order to do this, $\mathcal{D}_i$ simulates $\mathcal{D}$'s queries to E whenever the key is different from L_i or $L_i \oplus \alpha$. A standard hybrid argument then shows that

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC\text{-}IC2}_{K,L_i}[E]}(\mathcal{D}) \le \mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{EMIP2}_K[E_{L_i},E^{-1}_{L_i+\alpha}]}(\mathcal{D}_i)\,.$$

Since $L_i \neq L_i \oplus \alpha$, the permutations π_1 and π_2 are indeed independent and uniform random. Hence, Theorem 3 (with $c = n + k/2$) yields the upper bound

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{EMIP2}_{K_1}[E_{L_i},E^{-1}_{L_i+\alpha}]}(\mathcal{D}_i) &\le \frac{12}{2^{n+k}} + \frac{7qp_i^2}{2^{2n}} + 6\sqrt{3(n+k/2)}\frac{\sqrt{q}\,p_i}{2^n} \\
&\le \frac{12}{2^{n+k}} + (6\sqrt{3(n+k/2)} + \sqrt{7})\frac{\sqrt{q}\,p_i}{2^n} \\
&\le \frac{12}{2^{n+k}} + 9\sqrt{2n+k}\,\frac{\sqrt{q}\,p_i}{2^n}\,,
\end{aligned}$$

where the second inequality follows from $x^2 \le x$ for all $x \in [0,1]$. Hence, it follows that

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{KARC\text{-}IC2}}(\mathcal{D}) \le \frac{12}{2^{n+k}} + 9\sqrt{2n+k}\,\frac{\sqrt{q}\,p}{2^{n+k}}\,.$$

This concludes the proof. □

To apply Theorem 7 to PRINCE, it remains to show that the linear map $\mathcal{R}\circ\sigma$ does not have any fixed points when $\mathcal{R}$ is the linear reflector and σ the whitening-key orthomorphism[2] of PRINCE. Specifically, $\sigma : \{0,1\}^n \to \{0,1\}^n$ is defined by

$$\sigma(x) = (x \ggg 1) \oplus (x \gg 63)\,. \tag{23}$$

One can verify that $\mathrm{rank}(\mathrm{id} + \mathcal{R}\circ\sigma) = 64$. That is, $\mathcal{R}\circ\sigma$ does not have any fixed points.

Observe that the σ defined by (23) is not an involution. Hence, the PRINCE decryption algorithm is not the exactly same as the encryption algorithm: K and $\sigma(K)$ must also be swapped. Our construction preserves the same property, but we note that it is also possible to choose an involution σ such that $\mathcal{R}\circ\sigma$ does not have any fixed points. In this case, decryption and encryption are purely related by the coupling map $(K, L) \mapsto (\sigma(K), L \oplus \alpha)$.

However, since the block cipher E used in PRINCE starts by xoring L to the state, using an involution σ has the potential downside that $(K + L, \sigma(K) + L)$

[2] An orthomorphism such as σ is a linear map such that both σ and $\sigma\oplus\mathsf{id}$ are invertible.

is no longer jointly uniform for uniform random keys K and L. Indeed, for any linear involution σ, it holds that $\mathrm{rank}(\mathrm{id} + \sigma) \leq n/2$. This may facilitate partial key guessing. Again, this illustrates the importance of performing additional cryptanalysis when instantiating our (or, more generally, any) generic construction.

Acknowledgments. This work was supported in part by the Research Council KU Leuven: GOA TENSE (C16/15/058). Tim Beyne and Yu Long Chen are supported by a Ph.D. Fellowship from the Research Foundation - Flanders (FWO). The authors thank the reviewers for their valuable comments and suggestions.

References

1. Avanzi, R.: The QARMA block cipher family. IACR Trans. Symm. Cryptol. **2017**(1), 4–44
2. Babai, L.: The Fourier transform and equations over finite Abelian groups: an introduction to the method of trigonometric sums, Lecture notes
3. Barreto, P., Rijmen, V.: The Anubis block cipher. Primitive submitted to NESSIE
4. Barreto, P., Rijmen, V.: The Khazad legacy-level block cipher. Primitive submitted to NESSIE
5. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY family of block ciphers and its low-latency variant MANTIS. In: CRYPTO 2016, Part II. LNCS, vol. 9815, pp. 123–153
6. Beyne, T.: Block cipher invariants as eigenvectors of correlation matrices. In: ASIACRYPT 2018, Part I. LNCS, vol. 11272, pp. 3–31
7. Bogdanov, A., Knudsen, L.R., Leander, G., Standaert, F.X., Steinberger, J.P., Tischhauser, E.: Key-alternating ciphers in a provable setting: Encryption using a small number of public permutations - (extended abstract). In: EUROCRYPT 2012. LNCS, vol. 7237, pp. 45–62
8. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knežević, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In: ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225
9. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knežević, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: The PRINCE challenge (2014-2016), https://www.emsec.ruhr-uni-bochum.de/research/research_startseite/prince-challenge/
10. Boura, C., Canteaut, A., Knudsen, L.R., Leander, G.: Reflection ciphers. Des. Codes Cryptogr. **82**(1-2), 3–25
11. Bozilov, D., Eichlseder, M., Knezevic, M., Lambin, B., Leander, G., Moos, T., Nikov, V., Rasoolzadeh, S., Todo, Y., Wiemer, F.: PRINCEv2 - More security for (almost) no overhead. IACR Cryptol. ePrint Arch. **2020**, 1269
12. Chen, S., Lampe, R., Lee, J., Seurin, Y., Steinberger, J.P.: Minimizing the two-round Even-Mansour cipher. In: CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 39–56
13. Chen, S., Steinberger, J.P.: Tight security bounds for key-alternating ciphers. In: EUROCRYPT 2014. LNCS, vol. 8441, pp. 327–350

14. Chen, Y.L., Lambooij, E., Mennink, B.: How to build pseudorandom functions from public random permutations. In: CRYPTO 2019, Part I. LNCS, vol. 11692, pp. 266–293
15. Cogliati, B., Seurin, Y.: Analysis of the single-permutation encrypted davies-meyer construction. Des. Codes Cryptogr. **86**(12), 2703–2723
16. Daemen, J., Van Assche, G., Peeters, M., Rijmen, V.: Noekeon. Primitive submitted to NESSIE
17. Datta, N., Dutta, A., Nandi, M., Yasuda, K.: Encrypt or decrypt? To make a single-key beyond birthday secure nonce-based MAC. In: CRYPTO 2018, Part I. LNCS, vol. 10991, pp. 631–661
18. Dinur, I.: Cryptanalytic time-memory-data tradeoffs for FX-constructions with applications to PRINCE and PRIDE. In: EUROCRYPT 2015, Part I. LNCS, vol. 9056, pp. 231–253
19. Dunkelman, O., Keller, N., Shamir, A.: Minimalism in cryptography: The Even-Mansour scheme revisited. In: EUROCRYPT 2012. LNCS, vol. 7237, pp. 336–354
20. Even, S., Mansour, Y.: A construction of a cipher from a single pseudorandom permutation. In: ASIACRYPT'91. LNCS, vol. 739, pp. 210–224
21. Hoang, V.T., Tessaro, S.: Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In: CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 3–32
22. Hoeffding, W.: Probability inequalities for sums of bounded random variables. In: The Collected Works of Wassily Hoeffding, pp. 409–426. Springer (1994)
23. Kilian, J., Rogaway, P.: How to protect DES against exhaustive key search. In: CRYPTO'96. LNCS, vol. 1109, pp. 252–267
24. Kilian, J., Rogaway, P.: How to protect DES against exhaustive key search (an analysis of DESX). Journal of Cryptology **14**(1), 17–35
25. Lampe, R., Patarin, J., Seurin, Y.: An asymptotically tight security analysis of the iterated Even-Mansour cipher. In: ASIACRYPT 2012. LNCS, vol. 7658, pp. 278–295
26. Lee, J.: Key alternating ciphers based on involutions. Des. Codes Cryptogr. **86**(5), 955–988
27. Leurent, G., Sibleyras, F.: Low-memory attacks against two-round even-mansour using the 3-XOR problem. In: CRYPTO 2019, Part II. LNCS, vol. 11693, pp. 210–235
28. Luby, M., Rackoff, C.: How to construct pseudo-random permutations from pseudo-random functions (abstract). In: CRYPTO'85. LNCS, vol. 218, p. 447
29. Patarin, J.: The "coefficients H" technique (invited talk). In: SAC 2008. LNCS, vol. 5381, pp. 328–345
30. Patarin, J.: Introduction to mirror theory: Analysis of systems of linear equalities and linear non equalities for cryptography. IACR Cryptol. ePrint Arch. **2010**, 287
31. Patarin, J.: Mirror theory and cryptography. IACR Cryptol. ePrint Arch. **2016**, 702
32. Simon, T., Batina, L., Daemen, J., Grosso, V., Massolino, P.M.C., Papagiannopoulos, K., Regazzoni, F., Samwel, N.: Friet: An authenticated encryption scheme with built-in fault detection. In: EUROCRYPT 2020, Part I. LNCS, vol. 12105, pp. 581–611
33. Steinberger, J.P.: The sum-capture problem for abelian groups. arXiv preprint arXiv:1309.5582