Analyzing Constructions for key-alternating Pseudorandom Functions with Applications to Stream Cipher Operation Modes

Matthias Krause

University of Mannheim, Germany
krause@uni-mannheim.de
http://th.informatik.uni-mannheim.de/

Abstract. In the last years, much research work has been invested into the security analysis of key alternating ciphers in the random oracle model. These are pseudorandom permutations (PRPs), sometimes also called iterated Even-Mansour ciphers, which are defined by alternatingly adding n-bit sub-keys k_i and calling public n-bit permutations P_i. Besides the fact, that results of this kind concern the fundamental questions of understanding the nature of pseudorandomness, a practical motivation for this study is that many modern block cipher designs correspond exactly to variants of iterated Even-Mansour ciphers.

In this paper, we study similar construction for pseudorandom functions (PRFs), where additionally the access to a public n-bit (one-way) function F is allowed. In particular, we show a sharp $n/2$-security bound for the simplest possible construction $F(x \oplus k)$ and a sharp $2/3 \cdot n$-bound for the $FP(1)$-construction $F(P(x \oplus k) \oplus k)$, both in the random oracle model. The latter result contrasts with a sharp bound of the same order for $P(P(x \oplus k) \oplus \pi(k)) \oplus k$, recently proved by Chen et. al. in [6].

One practical motivation for our research is due to the fact that operation modes of key stream generator based (KSG-based) stream ciphers can be modeled in a very straightforward way by FP-constructions. Our research shows a way to save KSG inner state length by using operation modes, which yield provable security beyond the birthday bound against time-space-data tradeoff attacks. For instance, we demonstrate that a slight change in the operation mode of the Bluetooth cipher (adding the session key twice in the initialization phase) raises the security w.r.t. to generic time-space-data tradeoff attacks from $n/2$ to $2/3 \cdot n$, where n denotes the KSG inner state length.

Keywords: Pseudorandom functions, Even-Mansour Constructions, Lower Bound Proofs in the Random Oracle Model, Stream Ciphers

1 Introduction

1.1 General Motivation

In the last years, many research work has been invested into the security analysis of iterated Even-Mansour ciphers, $EM(r)$, $r \geq 1$ (see, e.g., [11],[4],[14],[7],[1] and

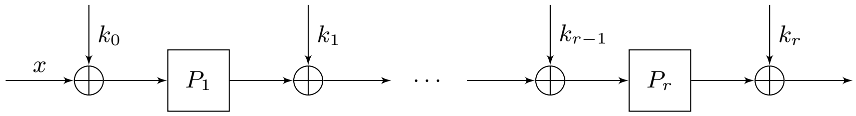

Fig. 1. The Cipher $EM(r)$.

the citations therein). These are pseudorandom permutations (PRPs) defined as described in figure 1

over a collection $P_1, \cdots, P_r$ of publicly known permutations, where $x \in \{0,1\}^n$ denotes the input and $(k_0, \cdots, k_r) \in (\{0,1\}^n)^{r+1}$ the secret key. The security is typically analyzed in the random oracle model, where the attacker tries to reach her goals (e.g., recovery of the secret key or distinguishing from a truly random permutation) on the basis of oracle queries to P_j/P_j^{-1}-oracles, $j = 1, \cdots, r$, and to an E-oracle, which answers queries $E(x) =?$-queries according to $E_{EM(r)}(x, (k_0, \cdots, k_r))$.

A first result here was a sharp $\frac{n}{2}$-bound for the original Even Mansour cipher $EM(1)$: *Daemen* found a generic attack finding the secret key with significant success probability with $O(2^{n/2})$ oracle queries [8], where *Even* and *Mansour* showed that for all $\alpha < 1/2$, attacks which make only $O(2^{\alpha \cdot n})$ oracle queries will reach this goal only with a success probability exponentially small in n [11]. After a long series of subsequent papers proving security results for several types of iterated Even-Mansour ciphers (see e.g. [4], [1] and the citations therein), *Chen* and *Steinberger* were finally able to prove for all naturals $r \geq 1$ a sharp $\frac{r}{r+1}$-security bounds of $EM(r)$, the Even-Mansour cipher of iteration depth r, w.r.t. to distinguishing from a truly random permutation [7].

There are several reasons motivating the analysis of Even-Mansour ciphers. The first is that results concerning simple generic construction of provable secure cryptographic primitives build quasi the formal foundation of cryptography, yielding justification for (or discovering weaknesses of) many designs of practically used cryptographic algorithms and systems. The concrete practical motivation for analyzing iterated Even-Mansour ciphers was to gain a better understanding of the security mechanisms underlying modern block cipher designs. Many practical block ciphers, like, e.g., AES, are key-alternating ciphers in the sense that the process of encrypting a given plaintext block x is partitioned into a number of rounds, where in each round a round function is applied to the sum of the outcome of the previous round and a round key coming from a key schedule algorithm. Note that this mechanism is exactly reflected by iterated Even-Mansour ciphers.

In this paper, we suggest to extend the focus of key alternating ciphers through the study of similar constructions for pseudorandom functions (PRFs), where, besides public permutations, a publicly known (one-way) function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$ is allowed as an possible additional component. We will use the name **FP-constructions** for this kind of generic PRF-constructions.

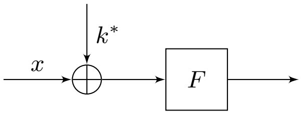

Fig. 2. The $F(0)$-construction.

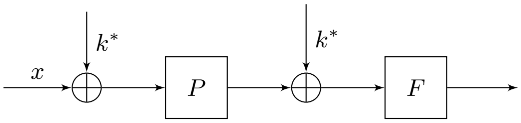

Fig. 3. The $FP(1)$-construction.

We present sharp security bounds for two concrete examples of FP-constructions in the random oracle model: We show a sharp $n/2$ security bound for the $F(0)$-construction $E_{F(0)}(x,k) = F(x \oplus k)$ (see figure 2) and a sharp $2/3 \cdot n$-security bound for the $FP(1)$-construction, where $E_{FP(1)}(x,k) = F(P(x \oplus k) \oplus k)$ (see figure 3). Here, P denotes a public permutation and k a secret n-bit key.

Our motivation for this kind of research can be summarized as follows. In general, PRFs represent a basic cryptographic primitive, which occur as building blocks in many cryptographic applications, especially in the context of authentication. Thus, results on provable security of simple PRF-constructions over simple component operations like adding the secret key or calls to (random looking) functions and permutations should be of some general interest.

The main practical motivation for our research is due to the fact that operation modes for key stream generator based (KSG-based) stream ciphers can be modeled in a straightforward way by FP-constructions. Here, P corresponds to to the iterated KSG internal state update function, and $F = F(x)$ to the KSG output function, which assigns to an inner KSG state $x \in \{0,1\}^n$ the block of the first n keystream bits generated on X. One practical consequence of or result is, that a slight change in the operation mode of the Bluetooth cipher [17], namely adding the session key twice in the initialization phase, raises the security w.r.t. to generic time-space-data tradeoff attackers from $n/2$ to $2/3 \cdot n$, where n denote the KSG inner state length (which is 128 for this cipher).

Note that the vulnerability of KSG-based stream ciphers against time-space-data tradeoff attacks represent an inherent weakness of existing practical stream ciphers, which yields the well known rule, that for achieving n-bit security by a stream cipher (in standard operation mode) one has to choose inner state length of at least $2n$ for the underlying key stream generator. This has the implication that practical stream ciphers operate with comparatively large inner state lengths (e.g., 288 for TRIVIUM [5] or 160 for GRAIN [13]).

Our research shows a way to save inner state length by using operation modes which yield provable security beyond the birthday bound against time-space-data tradeoff attacks. It is related to resent research of *Armknecht, Mikhalev* [2] which use another approach to reach this goal. The stream cipher operation modes we suggest can be seen as a compromise between the *one-stream mode* (used e.g. by TRIVIUM and GRAIN), which provides minimal initialization effort per keystream bit but minimal security w.r.t. generic collision attacks, and block cipher based modes, (used e.g. by the GSM standard A5/3), which provide maximal security w.r.t. generic collision attacks but maximal initialization effort per key stream bit. Concerning a more detailed discussion about the practical implications of our results for stream ciphers we refer to subsection 1.4 and to a parallel paper of *Hamann and Krause* [12].

In the following subsections, we discuss more related results, give a formal definition of the random oracle model underlying our security analysis, and describe our results, which are outlined in the subsequent sections, in more detail.

1.2 The Search for Minimal Constructions

Another view of our research in this paper results from the demand of minimality: What is the minimal construction over a given set of possible components which reaches a given security goal. Concerning the security bound $n/2$, for Even-Mansour ciphers this problem was solved by *Dunkelman, Keller* and *Shamir* in [10], who could show that even the one-key variant $P(x \oplus k) \oplus k$ of the original Even-Mansour cipher is $n/2$-secure w.r.t. key-recovery attacks. Note that our $F(0)$-construction solves the minimality problem for the security bound $n/2$ on the level of FP-constructions for PRFs.

Regarding the security bound $2/3 \cdot n$, the problem of finding a minimal Even-Mansour construction was completely solved very recently by *Chen, Lampe, Lee, Seurin, Steinberger* in [6]. They showed that the $EM(2)$-variant $P(P(x \oplus k) \oplus \pi(k)) \oplus k$ reaches $2/3 \cdot n$-security, where the simplest possible variant $P(P(x \oplus k) \oplus k) \oplus k$ reaches only security level $n/2$. Here, π denotes an orthomorphism, i.e., an n-bit permutation for which $x \to x \oplus \pi(x)$ is also a permutation, which even is allowed to be linear. Note that one of the technical tools underlying the lower bound proof in [6], the Sum-Capture Theorem, plays also an important role in our lower bound proof for the $FP(1)$-construction.

On the level of FP-constructions, the $2/3 \cdot n$-security can be reached already by the FP(1)-construction, i.e., by only two key-additions of the same key. However, it is not clear so far if the $FP(1)$-construction is really the minimal solution for this security level. It is an interesting open questions, if this can be reached even by the $FF(1)$-construction $F(F(x \oplus k) \oplus k)$. Note that the $O(2^{n/2})$-attack against $P(P(x \oplus k) \oplus k) \oplus k$ from [6] does not work against the $FF(1)$-construction.

1.3 Security Bounds in the Random Oracle Model

Our security bounds refer to the following game between Alice and the adversary Eve, which we explain at hand of the $FP(1)$-construction.

We suppose that Alice chooses randomly, independently and w.r.t. the uniform distribution a secret function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$, a secret permutation $P : \{0,1\}^n \longrightarrow \{0,1\}^n$ and a secret key $k^* \in \{0,1\}^n$. (In the case of the $F(0)$-construction she chooses only k^* and F)

Eve, who is supposed to be a randomized algorithm, tries to compute the secret key $k^* \in \{0,1\}^n$, on the basis of oracle queries to Alice. In particular, she is allowed to pose $P(u) =?$-queries, $P^{-1}(v) =?$-queries, $F(y) =?$-queries, and $E(x) =?$-queries for arbitrary inputs $u, v, y, x \in \{0,1\}^n$, which are immediately answered by $P(u)$, $P^{-1}(v)$, $F(y)$, and $E(x) = F(P(x \oplus k^*) \oplus k^*$, respectively, by Alice. (In the case of the $F(0)$-construction only F and E-queries are allowed, where $E(x) =?$-queries are answered by $E(x) = F(x \oplus k^*)$.

The relevant cost measures of attacks are the number of oracle queries and the success probability for finding the secret key, the costs are quantified in dependence of n. We consider an attack to be successfull if it finds the secret key with constant positive success probability.

Note that the two-way property of P is reflected by the ability of Eve to ask also P^{-1}-queries, and that the one-way property of F results from the fact, that queries of this kind are not allowed w.r.t. F. This implies that, for a given $z \in \{0,1\}$, the only way for Eve to find an input y fulfilling $F(y) = z$, is to run an exhaustive search in $\{0,1\}^n$ using F queries.

In section 2 we present first a simple successful collision attack against the $F(0)$-construction which needs $O(2^{n/2})$ oracle queries, and prove then a matching lower bound.

In section 3 we give a successful $O(2^{2/3 \cdot n})$-attack against the two-key variant of the $FP(1)$-construction (i.e., $F(P \oplus k_1) \oplus k_2)$), which uses the Slidex attack of Dunkelman, Keller, and Shamir (2012) [10] against the Even-Mansour cipher of iteration depth one as a subprogram.

Note that both attacks are *practically efficient* in the sense that the overall running time differs at most by a polynomial factor in n from the number of oracle queries. Such *practically efficient* attacks are not known for Even-Mansour ciphers of iteration depth $r \geq 2$ which use more than two different keys. Though successfull attacks which compute the secret key with $O(2^{r/(r+1) \cdot n})$ queries are known here, the overall running time of all these attacks is $\Omega(2^{(r+1)n})$, see [4].

Very recently, Dinur, Dunkelman, Keller and Shamir [9] could prove upper bounds for all variants of two key Even Mansour ciphers of iteration depth at most 4, which are practically efficient in the above sense.

Section 4 contains the main technical result of this paper, the proof of a matching lower bound for the $FP(1)$-construction. We show that for all $\alpha < 2/3$ Eve's success probability of computing the secret key k^* with $O(2^{\alpha \cdot n})$ oracle queries is exponentially small in n.

In both lower bound proofs, this for the $F(0)$-construction in section 2 and this for the $FP(1)$-construction, we follow the same line of argumentation as in the $EM(1)$-lower bound proof in [11]. We consider a sequence of queries as a process which makes an increasing set of keys $k \in \{0,1\}^n$ *bad* in the sense, that either the queries prove that k cannot be the secret key, or the queries make

k to a hopeful candidate for being the secret key. We show that, on the one hand, the probability that the secret key becomes bad during $O(2^{\alpha \cdot n})$ queries is exponentially small, and, on the other hand, that after posing $O(2^{\alpha \cdot n})$ queries which leave the secret key good, the success probability of guessing the secret key is also exponentially small.

While the lower bound proof for the $F(0)$-construction can be done quite straightforwardly, in the $FP(1)$-case we have to solve a number of new and difficult problems. One problem is the formal treatment of the probability space determining the success probabilities for making the secret key bad. In contrast to [11], the events that given good keys equal to the secret key are not equally likely, which complicates the analysis.

Both in the upper bound proof as in the lower bound proof, so-called EF-collisions play an essential role. EF-collision arises when asking an $E(x) =?$-query and an $F(y) =?$ for which $E(x) = F(y)$.

The main technical problem in the lower bound proof is to show that the probability that the number of EF-collisions exceeds $2^{1/3n}$ after $O(2^{\alpha \cdot n})$ queries is exponentially small. This proof would be easy if E and F were independent random functions. But in our case, each EF-collision increases the success probability for generating another EF-collision, which causes various difficulties. We solve this problem by applying an iterated Chernov bound argument, which could be helpful in other situation and interesting for their own (see lemma 5).

Another difficult combinatorial problem to be solved is caused by the fact that only one key k^* is used. The question is if, given a permutation P, subsets X, U, Y of $\{0,1\}^n$, $|X| \leq 2^{\alpha n}$, $|U| \leq 2^{\alpha n}$, $|Y| \leq 2^{\alpha n}$, can be constructed in such a way that for almost all keys $k \in \{0,1\}^n$ there is a pair $(x, u) \in X \times U$ such that $x \oplus u = k$ and $P(u) \oplus k \in Y$. This would imply a successful information theoretic attack with $O(2^{\alpha n})$ oracle queries against the $FP(1)$-construction.

Chen, Lampe, Lee, Seurin, Steinberger could show in [6] a so-called Sum-Capture theorem saying that the probability, that such a construction of sets X, U, Y is possible, is exponentially small, where the probability is taken w.r.t. the randomization of P. They used this theorem for proving that the $2/3 \cdot n$ security lower bound for the $P(P(x \oplus k) \oplus \pi(k)) \oplus k$-construction, which was already mentioned above.

Note that in most lower bound results on iterated Even Mansour ciphers, the bounds were proved in the stronger setting of distinguishing $E_{EM(r)}(., K)$ from a truly random permutation over $\{0,1\}^n$ ([6],[4][7]), and by using s special technique called *Patarin's H*-coefficient technique [16]. The question if our lower bound results can be also transformed into this setting remains as a topic for further research.

Due to their length, some parts of the lower bound proof of section 4 are outsourced into the appendix.

For obtaining an adequate modelling of stream cipher operation modes by FP-constructions, we have to consider their security w.r.t. a generalized random oracle model, where the attacker is able to ask a new type of oracle queries (see subsection 1.4 for explaination). In section 5 we show that the sharp $2/3 \cdot n$-

bound for the FP(1)-construction even holds in the generalized random oracle model.

1.4 The Connection to KSG-based Stream Ciphers

In this subsection, we will review the basic definitions and security requirements of keystream generator based (KSG-based) stream ciphers and then describe the connection of stream cipher operation modes and FP-constructions in a rather informal manner. For a more detailed and formal description of this topic we refer to our parallel paper [12].

Stream ciphers are symmetric encryption algorithms intended for encrypting, in an online manner, plaintext bit streams X, which have to pass an insecure channel. The encryption is done by adding bitwise a key stream $S(k)$ which is generated in dependence of a secret key k. The legal recipient, which also knows k, decrypts the encrypted bitstream $Y = X \oplus S(k)$ by generating $S(k)$ and computing $X = Y \oplus S(k)$.

In the heart of a stream cipher typically lies a key stream generator (KSG). These are clockwise working devices which can be formally specified by finite automata. A KSG is defined by an inner state length n, the set of inner states $\{0,1\}^n$, a state update function $\delta : \{0,1\}^n \longrightarrow \{0,1\}^n$ and an output function $out : \{0,1\}^n \longrightarrow \{0,1\}^*$. Starting from an initial state q_1, in each clock cycle $i \geq 1$, the KSG produces a piece of key stream $z_i = out(q_i)$ (as a rule, one bit long) and changes the inner state according to $q_{i+1} = \delta(q_i)$. The output bit stream $S(q_1)$ is defined by concatenating all the outputs $z_1 z_2 z_3 \cdots$.

The connection between stream ciphers and FP-constructions is that we associate the underlying KSG of inner state length n with a public function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$, where $F(x)$ is defined as the first n key bits generated on inner state x, and with a public permutation $P : \{0,1\}^n \longrightarrow \{0,1\}^n$, where $P(x)$ is defined to be the inner state obtained after clocking the KSG a certain number of times on state x. (Note that we suppose that the state transition function δ is bijective and efficiently invertible, which is true for many practical stream ciphers.)

The main security requirement for key stream generators is that it should be hard to distinguish a bit stream $S(q_1)$, generated by the KSG on a secret initial state $q_1 \in \{0,1\}^n$, from a truly random bitstream. This implies that it should be impossible to compute q_1 from a prefix of length n of $S(q_1)$, i.e., F should be a cryptographically hard one-way function.

In the last decades, many KSGs for practical use have been suggested and many different techniques for cryptanalyzing stream ciphers has been developed (correlation attacks, fast correlation attacks, guess-and-verify attacks, BDD-attacks, time-memory tradeoff attacks etc.).

In our context we concentrate on a very basic type of time-memory tradeoff attacks which we call generic collision attacks and which was first described in papers of *Babbage* [3]: Suppose that Eve knows a prefix of length $D + n - 1$ of $S(q_1)$. Then, with high probability, after sampling $2^n/D$-times a random state $q \in \{0,1\}^n$ and computing $F(q)$, Eve finds some inner state q such that $F(q)$ is a substring of $S(q_1)$, which allows to compute q_1 from q. Note that this yields a time-memory tradeoff of $O(2^{n/2})$. The vulnerability of KSG based stream ciphers against this generic type of attack is a main reason for the rule, that an inner state length of $2n$ has to be invested for reaching n bit security. Note that this rule influenced many practical stream ciphers, which operate with an inner state length of at least 160.

Stream cipher operation modes for encrypting a plaintext corresponding to a communication between partners Alice and Bob are usually defined by the following components:

- In a first phase, the session key generation phase, a common symmetric session key k for Alice and Bob is generated by running some key exchange protocol.
- The next phase is the state initialization phase, determined by an state initialization algorithm StateInit, which computes the initial state $q_1 = StateInit(IV, k)$ from a (public) initial value IV and the secret session key k.
- Finally, in the keystream generation phase, the keystream $S(q_1)$ will be generated for encrypting the plaintext.

In our context, we concentrate on the state initialization phase. In [12] we list the specifications of the StateInit algorithm for various practically used stream ciphers like Trivium, GRAIN, Bluetooth etc. Typically, the KSG itself is used for performing the state initialization. For instance, the way the Bluetooth systems computes an initial state q_1 from the secret session key k and an initial value IV can be modeled as

$$q_1 = F(P(IV' \oplus k')), \tag{1}$$

where IV' and k' depend linearly on IV and k, respectively [17][1].

Among stream cipher operation modes we distinguish *one-stream modes* and *packet modes*. In a one-stream mode, the whole plaintext stream of the session is encrypted by a sufficiently long prefix of the keystream $S(q_1)$, where in a packet mode, the plaintext X is partitioned into packets $X = X^1X^2X^3\cdots$ of a moderate packet length R, and separate pieces of keystream $S^i = S(q_1^i)$ are produced for each package. Hereby, the initial packet states q_1^i will be computed via

$$q_1^i = StateInit(IV^{(i)}, k),$$

in dependence of the secret session key k and a separate public initial value $IV^{(i)}$ for each packet.

Note that in the one-stream mode, the above described generic collision attack yields the initial state q_1 with $O(2^{n/2})$ time and memory on the basis of a

[1] In particular, k and IV are loaded into the KSG-registers in a $GF(2)$-linear manner, then the KSG is clocked a certain number of times without producing output, then n output bits are generated and reloaded into the KSG registers as initial state q_1, see [12].

piece of keystream $O(2^{n/2})$. In the packet-mode, we get the same time-memory-data tradeoff of $O(2^{n/2})$ for a successful recovery of the initial state of **one** packet. However, for a successful attack against all the other packets, one has to recover the secret session key k.

Consequently, we study the complexity of a session key recovery attack in dependence of the number of known key stream packets. The corresponding question is how to design the state initialization algorithm for getting provable security beyond the birthday bound against generic collision attacks w.r.t. to a recovery of the secret session key k.

Note that, according to (1), the way the Bluetooth system generates the first n keystream bits in dependence of the secret session key k and an initial value x can be modeled as

$$E(x,k) = F(F(P(x \oplus k))). \tag{2}$$

Defining $\tilde{F} = F \circ F \circ P$ one gets $E(x,k) = \tilde{F}(x \oplus k)$ which corresponds to the $F(0)$-construction. Consequently, the secret key k can be recovered with time and memory $O(2^{n/2})$ on the basis of $O(2^{n/2})$ keystream packets (see the upper bound results in section 2).

A more promising approach one gets by using a state initialization algorithm which mimics the $FP(1)$-construction:

- Load $x \oplus k$ into the inner state registers.
- Run the KSG (which is supposed to have a nonlinear state update function) a certain number of clock cycles without producing output.
- Add k to the resulting inner state.

1.5 A Generalized Random Oracle Model and Summary

Note that our lower bound argument from section 4 can still not be applied directly this scenario. The reason is that the packet length R has to be supposed to be larger than the inner state length n. Consequently, with each package, Eve gets not only $F(P(x \oplus k) \oplus k)$ (which corresponds to the first n bits of the keystream package) but even the value

$$F(\delta^j(P(x \oplus k) \oplus k)),$$

for $j = 0, \cdots, R-n$, corresponding to the key stream bits $j+1, \cdots, j+n$ of the packet.

We give a solution of this problem in section 5 by showing a lower bound of the same order for the $FP(1)$-construction in a more general setting. In particular, we show that our lower bound remains valid even against an adversary, which is allowed to ask an additional kind of oracle queries, called $E^{(j)}(x) =?$-queries, $j = 0, \cdots, R-n$, which are answered with $F(\delta^j(P(x \oplus k) \oplus k))$, and if R is polynomially bounded in n.

Summary: The resistance of KSG-based stream ciphers against generic collision attacks can be raised from $n/2$ to $2/3 \cdot n$, if it works in the packet mode, and if a state initialization algorithm is used which follows the scheme of the

Even-Mansour cipher $P(x \oplus k) \oplus k$. This is because this implies that the first n key stream bits per packet depend on the secret session key k and the initial value x in a way modeled by the $FP(1)$-construction. Consequently, 80-bit security against this type of attack can be already reached with $n = 128$ KSG inner state length.

The resistance against generic collision attacks can be raised beyond $2/3 \cdot n$ by using a state initialization algorithms corresponding to Even-Mansor ciphers of higher iteration depth. A hopeful candidate for security level $3/4 \cdot n$ would be the following state initialization algorithm

- Load $IV \oplus k$ into the inner state registers.
- Run the KSG (which is supposed to have a nonlinear state update function) a certain number C of clock cycles without producing output.
- Add $\pi(k)$ to the resulting inner state, where $\pi : \{0,1\}^n \longrightarrow \{0,1\}^n$ denotes an orthomorphism.
- Run the KSG again C clock cycles without producing output.
- Add k to the resulting inner state for obtaining the initial state for the keystream generation.

This corresponds to the construction $P(P(x \oplus k) \oplus \pi(k)) \oplus k$, for which a sharp $2/3 \cdot n$-security bound was proved in [6]. For getting a provable $3/4 \cdot n$-security lower bound, one had to show a corresponding lower bound for the minimized $FP(2)$-construction

$$E(x, k) = F(P(P(x \oplus k) \oplus \pi(k)) \oplus k)$$

which is still open.

Note that also block cipher based stream ciphers are used in practical systems, an example is the A5/3-cipher included in the GSM-standard [15]. A straightforward approach is to use the block cipher in the counting mode, i.e., the stream cipher works in the packet mode, where the packet length equals the block length length n of the block cipher, and the key stream $S_i \in \{0,1\}^n$ for the encryption of the i-th packet is computed according to

$$S_i = E_k(IV^{(i)}),$$

where $E_k : \{0,1\}^n \longrightarrow \{0,1\}^n$ denotes the encryption function of the underlying block cipher w.r.t. to the secret session key k.

Observe that this kind of stream cipher operation modes seem to be secure w.r.t. to generic collision attacks of the above described kind. However, the prize which has to be paid for this security is a much higher computational effort per key stream bit compared to KSG-based stream ciphers.

Thus, the FP-construction based operation modes suggested in this paper can be seen as a compromise between the one-stream mode, which provides minimal initialization effort per keystream bit but minimal security w.r.t. generic collision attacks, and the block cipher mode, providing maximal security w.r.t. generic collision attacks but maximal initialization effort per keystream bit.

2 A Sharp Bound on the Security of the $F(x \oplus k^*)$-construction

In this section, we present first a key-recovery attacks against the construction

$$E(x) = F(x \oplus k^*)$$

which needs time and space and number of oracle queries $O(2^{n/2})$. We show then a matching lower bound concerning the minimal number of oracle queries necessary for computing the secret key k^* with significant success probability.

We do this in the random oracle model, which was introduced in subsection 1.3.

A straightforward attack is to take subsets X, Y of $\{0,1\}^n$, $|X| = |Y| = 2^{n/2}$, such $X \oplus Y = \{0,1\}^n$, and ask the oracles for $E(x)$ for all $x \in X$, and for $F(y)$ for all $y \in Y$.[2] This yields a set of candidates for the secret key consisting of all elements $k \in \{0,1\}^n$ for which there is some $x \in X$ and some $y \in Y$ such that $k = x \oplus y$ and $E(x) = F(y)$. Note that this set of candidates has expected size less than two and contains the secret key k^*. Note further that this attack uses $O(2^{n/2})$ oracle queries but has expected running time of $\Theta(2^n)$.

For getting an attack of expected running time $O(2^{n/2})$, take an arbitrary subset X of $\{0,1\}^n$, $|X| = 2^{n/2}$, and ask the E-oracle for $E(x)$ for all $x \in X$. Then start to ask for $F(y)$ for randomly chosen elements $y \in Y$. Whenever you were successful in choosing an y fulfilling $F(y) = E(x)$ for some $x \in X$, test the hypothesis that $k^* = x \oplus y$ by posing a few further query pairs $E(x'), F(x' \oplus x \oplus y)$ and testing their equality. Note that a successful guess corresponds to the event that $y \in X \oplus k^*$ which has probability $2^{n/2} \cdot 2^{-n} = 2^{-n/2}$. Consequently, the expected running time of this attack is $O(2^{n/2})$.

Now we prove a matching lower bound

Theorem 1. *Suppose an attacker Eve who is allowed to pose at most $\mathcal{O}(2^{\alpha \cdot n})$ oracle queries for some $\alpha < 1/2$. Then there is some $\epsilon > 0$ such that Eve's success probability to compute the secret key k^* is bounded from above by $\mathcal{O}(2^{-\epsilon \cdot n})$, if n is large enough.*

Proof. We prove Theorem 1 by analyzing Eve's success probability to win the following game 1 between Alice and Eve: Alice chooses randomly, independently and w.r.t. the uniform distribution, a secret function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$, and a secret key $k^* \in \{0,1\}^n$.

In **phase A** of game 1, Eve asks a sequence of at most $\mathcal{O}(2^{\alpha \cdot n})$ oracle queries to Alice, where $E(x) =?$-queries are answered according to $E(x) = F(x \oplus k^*)$.

In **phase B**, Eve publishes a pair $(\tilde{x}, \tilde{z})$, where $\tilde{x}$ has not been asked in phase A in an E-Query. Eve wins if $\tilde{z} = E(\tilde{x})$.

[2] Let, e.g., be X and Y the set of all $x \in \{0,1\}^n$ for which the first $n/2$ components, resp. the last $n/2$ components are zero.

Clearly, if Eve succeeds in computing k^* on the basis of the queries posed in phase A, then Eve wins in phase B. Thus, Eves success probability in computing k^* is bounded from above by the probability that Eve wins game 1.

During asking a sequence of queries, say Q, Eve generates the following subsets X, Y of $\{0,1\}^n$:

– $X = \{x \in \{0,1\}^n\, ;\, Q$ contains an E-query w.r.t. $x\}$.
– $Y = \{y \in \{0,1\}^n\, ;\, Q$ contains an F-query w.r.t. $y\}$.

At some points, Eve will succeed in generating EF-collisions (x, y), fulfilling $x \in X$ and $y \in Y$, and $E(x) = F(y)$. An EF-collisions (x, y) can occur in two situations, the first is that $y = x \oplus k^*$ which directly yields the secret key $k^* = x \oplus y$. The second is that $y \neq x \oplus k^*$ but $F(y) = F(x \oplus k^*)$.

Let us call a query successful if it is a query for $F(y) =?$ for some $y \in X \oplus k^*$ or it is a query for $E(x) =?$ for some $x \in Y \oplus k^*$.

We change **game 1** to **game 2**: Whenever, in phase A, Eve succeeds to pose a successful query, Alice immediately gives up and Eve wins. Note that Eve's success probability will not be lowered by changing from game 1 to game 2.

This implies that after each sequence of unsuccessful queries Q, Eve knows that $k^* \notin X \oplus Y$.

Lemma 1. *After each sequence of unsuccessful queries Q, from Eve's point of view, for all $k \in \{0,1\}^n \setminus (X \oplus Y)$ the events that $k = k^*$ are equally likely.*

Proof. Note that the corresponding probability space is formed by the elementary events $(k, \tilde{F})$, where $k \in \{0,1\}^n \setminus (X \oplus Y)$ and $\tilde{F} : \{0,1\}^n \setminus Y \longrightarrow \{0,1\}^n$ denotes a completion of $F|_Y$ on $\{0,1\}^n \setminus Y$ which is consistent with k and Q. These elementary events correspond to Eve's hypotheses which all are equally likely to be equal to the correct hypothesis $(k^*, F|_{\{0,1\}^n \setminus Y})$.

Now observe that for all $k \in \{0,1\}^n \setminus (X \oplus Y)$, a completion $\tilde{F} : \{0,1\}^n \setminus Y \longrightarrow \{0,1\}^n$ makes $(k, \tilde{F})$ consistent with k and Q if and only if $\tilde{F}(x \oplus k) = E(x)$ for all $x \in X$ [3].

This implies that the number of all possible completions $\tilde{F} : \{0,1\}^n \setminus Y \longrightarrow \{0,1\}^n$, which make $(k, \tilde{F})$ consistent with k and Q is equal for all $k \in \{0,1\}^n \setminus (X \oplus Y)$.

Consequently, from Eve's point of view, all elements $k \in \{0,1\}^n \setminus (X \oplus Y)$ are equally likely to be the secret key k^*. □

For the rest of the proof we fix some positive constant $\alpha < 1/2$, let $M = \lceil 2^{\alpha \cdot n} \rceil$ and fix some sequence of queries $Q = (q_1, \cdots, q_M)$ (for simplicity we suppose that M is integral).

We estimate first the probability of the event E_0, that Eve wins game 3 in phase A with a sequence Q consisting of M P/P^{-1}-queries, M E-queries and M F-queries.

[3] Note that $Y \cap (X \oplus k) = \emptyset$

For all i, $1 \leq i \leq M$, let us denote by $(q_i \ succ)$ the event that query q_i is successful, and by G_i the event that $q_1, \cdots, q_i$ are unsuccessful.

Clearly, if Eve wins in phase A then there is some i, $1 \leq i \leq M$, such that q_i is successful and q_j is unsuccessful for all j, $1 \leq j \leq i-1$. Consequently, the success probability of Eve of winning in phase A is bounded from above by

$$\sum_{i=1}^{M} Pr[G_{i-1} \wedge (q_i \ succ)].$$

$$= \sum_{i=1}^{M} Pr[q_i \ succ | G_{i-1}] \cdot Pr[G_{i-1}]$$

$$\leq \sum_{i=1}^{M} Pr[q_i \ succ | G_{i-1}].$$

Here, G_0 denotes an event of probability one.

Let us estimate the probability for the event that q_i is successful under the condition G_{i-1} that this is not the case for $q_1, \cdots, q_{i-1}$. Note first that this probability equals zero for $i = 1$. Now fix some $i > 1$ and let X and Y denote the input sets generated by $q_1, \cdots, q_{i-1}$ in the way described above.

Consider first the case that query q_i asks for $F(y)$ for some $y \notin Y$. The event that this query is successful corresponds to the event that $k^* \in y \oplus X$ which, by Lemma 1, has probability $|X|/(2^n - |X \oplus Y|)$ which is not greater than $M/(2^n - 2^{2\alpha n}) \leq 2 \cdot 2^{(\alpha-1)n}$, if n is large enough.[4]

In the case that query q_i asks for $E(x)$ for some $x \notin X$, the event that this query is successful corresponds to the event that $k^* \in x \oplus Y$ which yields the same bound for the success probability.

We obtain that the probability for the event $E(A)$, that Eve wins in phase A, is bounded by $2^{\alpha n} \cdot 2^{(\alpha-1)n} = 2^{(2\alpha-1)n}$ if n is large enough.

Let us now estimate the probability for the event $E(B)$ that Eve wins in phase B. Note first that

$$Pr[E(B)] = Pr[E(B) \cap (\neg E(A))] = Pr[E(B)|\neg E(A)] \cdot Pr[\neg E(A)]$$

$$\leq Pr[E(B)|\neg E(A)] = Pr[E(B)|G_M],$$

i.e., we estimate the the probability for the event $E(B)$ that Eve wins in phase B under the condition that all queries in $Q = (q_1, \cdots, q_M)$ were unsuccessful.

Let again X and Y denote the input sets generated by $q_1, \cdots, q_M$ in the way described above.

According to game 1 and 2, Eve has to submit a pair $(\tilde{x}, \tilde{z})$. Eve wins if and only if $\tilde{z} = E(\tilde{x}) = F(\tilde{x} \oplus k^*)$.

We change again the rules of the game by allowing Eve to submit first only $\tilde{x}$. Then, Alice gives k^* to Eve, and only then Eve has to submit $\tilde{z}$.

[4] Here we used the fact that $2^{2\alpha n} \leq 2^{n-1}$ if n is large enough.

Clearly, if Eve manages to choose $\tilde{x}$ in $Y \oplus k^*$ then Eve wins, as then $\tilde{x} \oplus k^*$ belongs to Y and Eve can determine $\tilde{z} = F(\tilde{x} \oplus k^*)$ herself. The event that $\tilde{x} \in Y \oplus k^*$ corresponds to the event that $k^* \in \tilde{x} \oplus Y$ which, by Lemma 1, has probability $|Y|/(2^n - |X \oplus Y|)$, which is not greater than $M/(2^n - 2^{2\alpha n}) \leq 2 \cdot 2^{(\alpha-1)n}$ if n is large enough.

If Eve chooses $\tilde{x}$ outside of $Y \oplus k^*$ then $\tilde{x} \oplus k^*$ is outside of Y, and the probability that Eve guesses for $\tilde{z}$ the correct value $F(\tilde{x} \oplus k^*)$ without asking the F-oracle is 2^{-n}.

Consequently, the probability that Eve wins in phase B is bounded by $2 \cdot 2^{(\alpha-1)n} + 2^{-n}$, which yields that the overall winning probability is bounded by $2 \cdot 2^{(2\alpha-1)n} + 2 \cdot 2^{(\alpha-1)n} + 2^{-n}$ which proves our theorem.

3 Attacks against the $FP(1)$-Construction and the Sum-Capture Theorem of Chen et.al.

In this section, we present nontrivial key-recovery attacks against the FP(1)-construction

$$E(x) = F(P(x \oplus k^*) \oplus k^*).$$

We do this in the random oracle model, i.e., we refer to an attacker, Eve, who aims to compute the secret key $k^* \in \{0,1\}^n$ behind E by running a randomized algorithm and who has access to a P-oracle, a P^{-1}-oracle, an F-oracle and an E-oracle in a black box way.

In a first step, we present two attacks against the two-key variant of the $FP(1)$-construction

$$E(x) = F(P(x \oplus k_1^*) \oplus k_2^*)$$

which compute the secret key (k_1^*, k_2^*) with high probability by using $O(2^{2/3 \cdot n})$ oracle queries. We then discuss an approach for finding more efficient attacks for the one-key variant.

3.1 An attack against the two-key variant with $O(2^{2/3 \cdot n})$ oracle queries but $\Theta(2^{2n})$ running time

We sketch only the idea of the attack and omit any detailed analysis of the success probability. Note that Bogdanov et.al. [4] used the same idea for proving similar upper bounds on iterated Even-Mansour ciphers.

Suppose that Eve poses a sequence Q of oracle queries containing

- $|X|$ E-queries yielding a set $\{(x, E(x)) | x \in X\}$ for some $X \subseteq \{0,1\}^n$,
- $|U|$ P-queries yielding a set $\{(u, P(u)) | u \in U\}$ for some $U \subseteq \{0,1\}^n$,
- $|Y|$ F-queries yielding a set $\{(y, F(y)) | y \in Y\}$ for some $Y \subseteq \{0,1\}^n$.

We call a pair $(k_1, k_2) \in \{0,1\}^{2n}$ to be **bad w.r.t.** Q if there is some triple $(x, u, y) \in X \times U \times Y$ such that

- $k_1 = x \oplus u$,

- $k_2 = P(u) \oplus y$.

Note that if $E(x) \neq F(y)$ then (k_1, k_2) can not be the secret key as it computes a wrong value on x. If $E(x) = F(y)$ then (k_1, k_2) is a promising candidate for being the secret key, which can be tested by only a few further queries.

Note further that if X, U, Y are randomly chosen and $|X| = |Y| = |U| \in \mathcal{O}(2^{2/3 \cdot n})$ then allmost all pairs (k_1, k_2) become bad with high probability. In this case, Q contain all information for finding the secret key with high probability. However, the effort for deriving this information from Q corresponds to exhaustive key search in $\{0,1\}^{2n}$.

3.2 An attack against the two-key variant with $O(2^{2/3 \cdot n})$ oracle queries and $O(2^{2/3 \cdot n})$ running time

We now present an alternative attack which requires asymptotically the same number of queries but significantly less computation. It uses the slidex attack against the Even-Mansour cipher of iteration depth one,

$$EM(x) = P\left(x \oplus k_1^*\right) \oplus k_2^*.$$

Note that for our $FP(1)$-construction it holds $E(x) = F(EM(x))$. Let us first recall the slidex attack as it occurs in [10]. We recall also the proof, as we later need some details of it.

Theorem 2 (Slidex Attack [10]). *For any set of T pairs $(x, EM(x))$ the secret key (k_1^*, k_2^*) can be computed with constant success probability using $\mathcal{O}(2^n/T)$ P-queries.* □

Proof (Theorem 2). Suppose we have a pair x, x' and a known difference Δ such that

$$x \oplus x' = k_1^* \oplus \Delta.$$

This implies $EM(x) = P(x \oplus k_1^*) \oplus k_2^* = P(x' \oplus \Delta) \oplus k_2^*$ and $EM(x') = P(x \oplus \Delta) \oplus k_2^*$. Consequently,

$$EM(x) \oplus EM(x') = P(x \oplus \Delta) \oplus P(x' \oplus \Delta), \text{ i.e.,}$$
$$EM(x) \oplus P(x \oplus \Delta) = EM(x') \oplus P(x' \oplus \Delta).$$

Now suppose that we have a set Z of D pairs (x_i, y_i), where $y_i = EM(x)$ for $i = 1, \cdots, D$. The probability that the set of corresponding pairs

$$M(Z, \Delta) = \{(x_i, y_i \oplus P(x_i \oplus \Delta))\}$$

contains a collision is in $\Omega(D^2/2^n)$.

Let $T' = 2^n/D^2$ and fix T' differences $\Delta_1, \ldots, \Delta_{T'}$. Then, with constant probability, one of the sets $M(Z, \Delta_j)$ contains a collision. Suppose that this happens with difference Δ_t and for x, x'. Then the hypothesis $k_1^* = x \oplus x' \oplus \Delta_t$ is true with constant probability. The remaining key k_2^* can be computed via $k_2^* = P(x \oplus k_1^*) \oplus E(x)$. Note that, for mounting this attack, we have to pose $T = D \cdot T' = 2^n/D$ P-queries. □

The idea for a corresponding attack against the $FP(1-$ construction is as follows. We construct a sufficiently large set Z of pairs $(x, EM(x))$ by generating a sufficiently large number of EF-collisions (x, y), fulfilling $E(x) = F(y)$, and apply the the slidex attack to Z.

Theorem 3. *The secret key (k_1^*, k_2^*) can be computed with constant success probability using $\mathcal{O}(2^{2/3 \cdot n})$ P-queries, $\mathcal{O}(2^{2/3 \cdot n})$ E-queries, and $\mathcal{O}(2^{2/3 \cdot n})$ F-queries.*

Proof (Theorem 3). The corresponding attack proceeds as follows. First, randomly choose a subset $X \subseteq \{0,1\}^n$ and build the set $\{(x, E(x)) \mid x \in X\}$. Likewise, choose a random $Y \subseteq \{0,1\}^n$ and set up $\{(y, F(y)) \mid y \in Y\}$.

Let $Z \subseteq X \times Y$ denote the set of pairs (x, y) for which $E(x) = F(y)$.

Note that, for a random y added to Y, there are two different events that yield $(x, y) \in Z$ for some $x \in X$. The first event is that $y = P(x \oplus k_1^*) \oplus k_2^* = EM(x)$ for some $x \in X$. This occurs with probability $|X| \cdot 2^{-n}$. The second event is that $y \notin P(X \oplus k_1^*) \oplus k_2^*$ but $F(y) \in E(X)$. This occurs with probability at most $(1 - |X| \cdot 2^{-n}) \cdot |X| \cdot 2^{-n}$, which is smaller than the probability of the first event.

Consequently, the expected value of $|Z|$ is near $2 \cdot |X| \cdot |Y| \cdot 2^{-n}$ and the expected number D of pairs (x, y) in Z which satisfy $y = EM(x)$ is at least $|X| \cdot |Y| \cdot 2^{-n}$.

Now fix a set of $2^n / D^2 = 2^{3n} / (|X|^2 \cdot |Y|^2)$ differences an build the sets $M(Z, \Delta)$ for all differences Δ in this set. Then check for all collisions if they yield the correct subkey k_1^* in the manner described above. This will happen with constant probability.

Note that we have to pose $2^{2n} / (|X| \cdot |Y|)$ P-queries, which implies an overall number of oracle queries of $\Theta(2^{2/3 \cdot n})$ if $|X| = |Y| \in \Theta(2^{2/3 \cdot n})$.

3.3 An approach for a more efficient attack to the one-key variant

Let us discuss an approach to finding more efficient key recovery attacks for the case that the same key is used in both rounds of the FP-construction. We then show nontrivial limitations of this approach by applying the sum-capture theorem of Chen et.al. [6].

Consider again the situation after Eve having asked a sequence of queries Q, which yields subsets X, U, Y of $\{0,1\}^n$ as described in subsection 3.1.

Again, we call a potential key $k \in \{0,1\}^n$ to be **bad** w.r.t. Q if there are inputs $x \in X$, $u \in U$ and $y \in Y$ such that

$$k = x \oplus u = P(u) \oplus y. \tag{3}$$

Note here that $Bad(P, U, X, Y)$ depends only on U, and P and X and Y, and does not depend on F, resp. E.

As described in subsection 3.1, if Eve manages to pose the queries Q in such a way that $Bad(P, U, X, Y)$ contains almost all $k \in \{0,1\}^n$ then Eve is done, as the remaining keys can be probed one by one in a straightforward way.

This yields the following attack with $3M$ oracle queries, which we shortly will call M-strategy:

- Fix an appropriate set $U = \{u_1, \cdots, u_M\} \subseteq \{0,1\}^n$
- Determine $P(u_i)$ by posing P-queries for all u_i, $i = 1, \cdots, M$
- Determine subsets X, Y of $\{0,1\}^n$, $|X| = |Y| = M$, such that $|Bad(P, U, X, Y)|$ is maximal.

Eve can be considered to be successfull if $|Bad(P, U, X, Y)| \geq 2^n - M$, as then the secret key can be found with $O(M)$ queries.

Note that in the two-key case, for any $\alpha < 2/3$, $O(2^{\alpha n})$ queries can make only an exponentially small part of all 2^{2n} possible keys bad, as $|Bad(P, U, X, Y)| \leq |X \times U \times Y| \in O(2^{3\alpha n}$ and $3\alpha < 2$.

However, this argument does not work in the one-key case. The following example shows that, under certain circumstances, Eve can be successful even for $M = 2^{n/2}$, which can never happen in the two-key case:

Example: We choose $X = \{x_1, \cdots, x_M\}$ and $U = \{u_1, \cdots, u_M\}$ in such a way that $X \oplus U = \{0,1\}^n$.[5] Consider the case that $P(u_i) = u_i$ for all $i = 1, \cdots, M$ and let $Y = X$.

We know that for all $k \in \{0,1\}^n$ there is a pair (i, j), $1 \leq i, j \leq M$, such that $k = x_i \oplus u_j$. But this implies that $P(u_j) \oplus k = x_i \in Y$, i.e., $k \in Bad(P, U, X, Y)$.

Clearly, this example says almost nothing about Eve's success probability for $M = 2^{n/2}$, as the event that $P(u_i) = u_i$ for all $i = 1, \cdots, M$ has extremely small probability.

However, the question remains open if there is a successful M-strategy for some $M \in O(2^{\alpha n}$, $\alpha < 2/3$.

The following Theorem, which is a straightforward consequence of the Sum-Capture Theorem of Chen et.al. in [6] (it is Theorem 1 in section 3), gives a negative answer to this question.

Theorem 4. *Let P denote a uniformly random permutation over $\{0,1\}^n$, let $N = 2^n$, and fix an arbitrary number M, $9n \leq M \leq N/2$. Suppose that Eve (who is supposed to be a probabilistic algorithm) poses a sequence $U = \{u_1, \cdots, u_M\}$ of M P-queries. For any subsets $X, Y \subseteq \{0,1\}^n$ let*

$$\mu(P, U, X, Y) = |\{(u, x, y) \in U \times X \times Y, x \oplus u = y \oplus P(u)\}| .$$

Then the probability for the event that there are subsets $X, Y \subseteq \{0,1\}^n$ such that

$$\mu(P, U, X, Y) \geq \frac{M \cdot |X| \cdot |Y|}{N} + \frac{2M^2 \cdot \sqrt{|X| \cdot |Y|}}{N} + 3\sqrt{n \cdot M \cdot |X| \cdot |Y|} \tag{4}$$

is at most $\frac{2}{N}$, where the probability is taken over the random choice of P and the internal randomization of Eve. □

This has the following consequence

[5] Take, for instance X and U as the set of inputs having zero at the first resp. last $n/2$ components.

Corollary 1. *Let $9n \leq M \leq 2^{\alpha \cdot n}$ for some $\alpha < 2/3$ and let Q denote a sequence of queries posed by Eve consisting of M E-queries, M P-queries, and M F-queries, yielding subsets X, U, Y as described above. Then there is a number $\epsilon_0 > 0$ such that the probability for the event that*

$$|Bad(P, X, U, Y)| \geq 2^{(1-\epsilon_0)\cdot n}$$

is, if n is large enough, at most $\frac{2}{N}$, where the probability is taken over the random choice of P and the internal randomization of Eve.

Proof: Note that a key $k \in \{0,1\}^n$ is bad if and only if there is some tripel $(u, x, y) \in U \times X \times Y$ such that $k = x \oplus u = P(u) \oplus y$. Consequently, $|Bad(P, X, U, Y)| \leq \mu(P, U, X, Y)$. The claim of Corollary 1 follows from the observation that, under the condition that $|U| = |X| = |Y| = M \leq 2^{\alpha \cdot n}$ for $\alpha < 2/3$, all three summands of (4) belong to $O\left(2^{(1-\epsilon_0)\cdot n}\right)$ for some $\epsilon_0 > 0$, if n is large enough. □

4 A Matching Lower Bound for the $FP(1)$-Construction

In this section, we prove a matching lower bound on the security of the $FP(1)$-construction

$$E(x) = F(P(x \oplus k^*) \oplus k^*)$$

in the random oracle model.

Theorem 5. *Suppose an attacker Eve who has black-box access to a P/P^{-1}-oracle, an F-oracle and an E-oracle, and who is allowed to pose at most $\mathcal{O}(2^{\alpha \cdot n})$ oracle queries for some $\alpha < 2/3$. Then there is some $\epsilon > 0$ such that Eve's success probability to compute the secret key k^* is bounded from above by $\mathcal{O}(2^{-\epsilon \cdot n})$, if n is large enough.*

We prove Theorem 5 by analyzing Eve's success probability to win the following game 1 between Alice and Eve: Alice chooses randomly, independently and w.r.t. the uniform distribution a secret function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$, a secret permutation $P : \{0,1\}^n \longrightarrow \{0,1\}^n$ and a secret key $k^* \in \{0,1\}^n$.

In **phase A** of game 1, Eve asks a sequence of at most $\mathcal{O}(2^{\alpha \cdot n})$ oracle queries to Alice, where $E(x) =?$-queries are answered according to $F(P(x \oplus k^*) \oplus k^*)$.

In **phase B**, Eve publishes a pair $(\tilde{x}, \tilde{z})$, where $\tilde{x}$ has not been asked in phase A in an E-Query. Eve wins if $\tilde{z} = E(\tilde{x})$.

Clearly, if Eve succeeds in computing k^* on the basis of the queries posed in phase A, then Eve wins in phase B. Thus, Eves success probability in computing k^* is bounded from above by the probability that Eve wins game 1.

During asking a sequence of queries Q, Eve generates the following subsets X, Y, U, V of $\{0,1\}^n$:

- $X = \{x \in \{0,1\}^n\,;\, Q$ contains an E-query w.r.t. $x\}$.
- $Y = \{y \in \{0,1\}^n\,;\, Q$ contains an F-query w.r.t. $y\}$.

- $U = \{u \in \{0,1\}^n\,;\, Q$ contains a P-query w.r.t. u, or a P^{-1}-query w.r.t. $P(u)\}$.
- $V = P(U)$.

As the size of Q is allowed to be beyond the birthday bound, with high probability Eve will succeed in generating EF-collisions (x, y), fulfilling $x \in X$, $y \in Y$, and $E(x) = F(y)$. We have seen already in subsection 3.2 that EF-collisions (x, y) can be induced by a collision of F, or they can be *structural* in the sense that $y = P(x \oplus k^*) \oplus k^*$.

Alice simplifies Eve's life by changing **game 1** to **game 2**: Whenever, in phase A, Eve succeeds to generate an EF-collision, Alice gives Eve the information if it structural or not. If yes, then Alice gives additionally a pointer to the collision partner. In particular, if Eve succeeds to pose an F-query for some $y \in P(X \oplus k^*) \oplus k^*$ then Alice gives a pointer to $x = P^{-1}(y \oplus k^*) \oplus k^* \in X$. Correspondingly, If Eve succeeds to pose an E-query for some $x \in P^{-1}(Y \oplus k^*) \oplus k^*$ then Alice gives a pointer to $y = P(x \oplus k^*) \oplus k^* \in Y$.

Note that Eve'e success probability will not be lowered by changing from game 1 to game 2.

The EF-collisions of type 1 induce the following sets $X^* \subseteq X$ and $Y^* \subseteq Y$:

- $X^* = \{x \in X; \exists y \in Y, \text{s.t. } (x, y)$ is EF-collision of type $(1)\}$,
- $Y^* = \{y \in Y; \exists x \in X, \text{s.t. } (x, y)$ is EF-collision of type $(1)\}$.

Note that, due to the rules of game 2, it holds $|X^*| = |Y^*|$, and that Eve knows a 1:1-correspondence between X^* and Y^* saying for each $x \in X^*$ which $y \in Y^*$ fulfills $y = P(x \oplus k_1^*) \oplus k_2^*$.

How queries help Eve to find the secret key?

Let $k \in X \oplus U$, i.e., there are $x \in X$ and $u \in U$ such that $k = x \oplus u$.

Consider first the case that $x \in X^*$ and remember that Eve knows the partner $y = P(x \oplus k^*) \oplus k^*$ in Y^*. If $P(u) \oplus k \neq y$ implies that k can not be the secret key. If $P(u) \oplus k = y$ then $k = k^*$ with high probability. Consequently, for all $k \in X^* \oplus U$, Eve gains nontrivial information from Q.

Now consider the case that $x \in X \setminus X^*$ and $y := P(u) \oplus k$ belongs to Y. We know that (x, y) is not a structural collision, otherwise we had $x \in X^*$. This implies that k can not be the secret key.

A similar classification can be done for all potential keys which are differences between elements of V and Y.

This observations motivate the following definition.

Definition 1. *A key $k \in \{0,1\}^n$ is called to be **bad** w.r.t. a sequence of queries Q if*

b.1 $k = x \oplus u$ for some $x \in X^*$ and $u \in U$ or $k = v \oplus y$ for some $v \in V$ and $y \in Y^*$, or,

b.2 $k = x \oplus u$ for some $x \in X \setminus X^*$ and $u \in U$ and it holds that $P(u) \oplus k \in Y$, or, *b.3 It holds that $|(X \oplus k) \cap U| > 2^{(1/3)\cdot n}$ or $|(Y \oplus k) \cap V| > 2^{(1/3)\cdot n}$.*

A key k which does not fulfil any of the conditions b.1, b.2, b.3 is called to be ***good w.r.t.*** *Q*

Items b.1 and b.2 correspond to the observations made above. Item b.3 ensures that all good keys have, from Eve's point of view, approximately the same probability to be the secret key (see lemma 6 below).

Note that the set $Bad(P, X, U, Y)$ defined in subsection 3.3 contains exactly those keys which satisfy condition b.2.

Definition 2. *A key $k \in \{0,1\}^n$ is called to be* ***consistent with*** *Q, if, from Eve's point of view, F can be completed on $\{0,1\}^n \setminus Y$ and P on $\{0,1\}^n \setminus U$ in such a way that no contradiction with Q occurs, i.e. $E(x) = F(P(x \oplus k) \oplus k)$ for all $x \in X$.*

An important property of good keys is given in

Lemma 2. *Each $k \in \{0,1\}^n$, which is good w.r.t. Q, is consistent with Q.*

Proof. We construct completions P' of $P|_U$ on $\{0,1\}^n \setminus U$ and F' of $F|_U$ on $\{0,1\}^n \setminus Y$ which make a good key k consistent with Q.

We run with u through $\{0,1\}^n$ in a certain order and define (if still necessary) the function values $P'(u)$ and $F'(P(u) \oplus k)$ (resp. $F'(P'(u) \oplus k)$ in a manner which makes k consistent with Q. We dynamically maintain a set $Target(P')$ which is initially set to $\{0,1\}^n \setminus V$. Whenever we define $P'(u)$ for a *new* u, we delete $P'(u)$ from $Target(P')$.

- **Phase 1**, we define $P'(u)$ for all $u \in \{0,1\}^n$ for which $u \oplus k = x \in X^*$. As k is good, it holds $u \notin U$. We fix the unique $y \in Y^*$ for which $y = P(x \oplus k^*) \oplus k^*$. As k is good, it holds $y \oplus k \notin V$. Thus, we can define $P'(u) \leftarrow y \oplus k$.
- **Phase 2** concerns all $u \in U$ for which $u \oplus k \in X$ for some r, $0 \leq r \leq R-1$. Note that $u \oplus k \notin X^*$, otherwise, k would be bad via definition 4, b.1. Moreover, $P(u) \oplus k \notin Y$, otherwise k would be bad via definition 1, b.2. We set
$$F'(P(u) \oplus k) \leftarrow E(u \oplus k).$$
- **Phase 3** concerns all $u \notin U$ for which $x = u \oplus k \in X \setminus X^*$. Eve knows that for all $y \in Y$, (x, y) is not a structural collision. Consequently, $P'(u)$ has to be chosen in such a way that $P(x \oplus k) \oplus k \notin Y$. Corresponding to this we choose
$$P'(u) \in Target(P') \setminus (Y \oplus k),$$
and set $F'(P'(u) \oplus k) = E(x)$.
Note that $Target(P') \setminus Y$ has definitely enough elements for doing this, as $Target(P') \geq 2^n - O(2^{\alpha n})$ after phase 2 and 3.

For all ll remaining $u \in \{0,1\}^n$ and $y \in \{0,1\}^n$, the values of $P'(u) \in Target(P')$ and $F(y) \in \{0,1\}^n$ can be freely chosen without generating contradictions with the E-queries occuring in Q. □

Note that the proof of Lemma 2 gives some valuable information about how completions P' of $P|_U$ on $\{0,1\}^n \setminus U$ and F' of $F|_U$ on $\{0,1\}^n \setminus Y$, which are consistent with the assumption that a good key k is the secret key, are looking like. In particular,

- Q determines the values of P' on $X^* \oplus k$, a set of size $|X^*|$ (Phase 1).
- Q implies that the function values of P' on the set $((X \setminus X^*) \oplus k) \setminus U$ are forbidden to be in $Y \oplus k$ (Phase 3).
- Q determines the values of F' on a set of size $|X \setminus X^*|$, in particular on

$$Z = (P(U \cap ((X \setminus X^*) \oplus k)) \oplus k) \cup (P'(((X \setminus X^*) \oplus k) \setminus U) \oplus k)\,.$$

We need this statement later in the proof of lemma 6.

Now we consider again a slightly modified **game 3**. The difference with game 2 is that Eve wins immediately in phase A if she manages to pose a query which makes k^* bad. Clearly, the success probability of Eve for winning game 1 is bounded from above by the success probability of Eve for winning game 3.

For the rest of the proof we fix some positive constant $\alpha < 2/3$ and denote $M = 2^{\alpha \cdot n}$. For simplicity we suppose that M is an integer.

We estimate first the probability of the event E_0, that Eve wins game 3 in phase A with a sequence Q consisting of M P/P^{-1}-queries, M E-queries and M F-queries. We know from Theorem 4 and Corollary 1 that there is some $\epsilon_0 > 0$ such that the probability for the event that

$$|Bad(Q,X,U,Y)| \geq 2^{(1-\epsilon_0)n}$$

is at most $2^{-(n-1)}$ if n is large enough.

Correspondingly, we say that Q satisfies the condition E^* if $|Bad_2(Q)| < 2^{(1-\epsilon_0)n}$. Note that

$$Pr[E_0] \leq Pr[E_0|E^*] + Pr[\neg E^*] \leq Pr[E_0|E^*] + 2^{-(n-1)}.$$

In the following we estimate $Pr[E_0|E^*]$ and consider only sequence $Q = (q_1, \cdots, q_M)$ of queries which satisfy E^*.

For all i, $1 \leq i \leq M$, let us denote by $(q_i\ bad)$ the event that query q_i makes k^* bad, and by G_i the event that k^* remained good after $q_1, \cdots, q_i$.

Clearly, if Eve wins in phase A then there is some i, $1 \leq i \leq M$, such that q_i is bad and q_j is good for all j, $1 \leq j \leq i-1$. Consequently, the success probability of Eve of winning in phase A is bounded from above by

$$\sum_{i=1}^{M} Pr[G_{i-1} \wedge (q_i\ bad)].$$

$$= \sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1}] \cdot Pr[G_{i-1}]$$

$$\leq \sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1}].$$

Here, G_0 denotes an event of probability one. We will show

Lemma 3. *There is some $\epsilon_1 > 0$ such that*

$$\sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1}] \leq O\left(2^{-\epsilon_1 \cdot n}\right).$$

This implies an exponentially small upper bound for Eve's success probability for winning game 3 in phase A.

Our way to prove lemma 3 is to prove it under the condition that $|X^*|$ grows only moderately during Q and to prove that the probability that $|X^*|$ grows too fast is exponentially small. In particular, let us denote by H_i the event that $|X^*| \leq 2^{(1/3)n}$ after $q_1, \cdots, q_i$. We show

Lemma 4. *There is some $\epsilon_3 > 0$ such that*

$$\sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1} \cap H_{i-1}] \leq O\left(2^{-\epsilon_3 \cdot n}\right).$$

and

Lemma 5. *For all i, $1 \leq i \leq M$, it holds $Pr[\neg H_i|G_i] < e^{-n}$ if n is large enough.*

Lemma 3 can be derived from lemmata 4 and 5 by using that

$$Pr[A|B] \leq Pr[A|B \cap C] + Pr[\bar{C}|B]$$

for all events A, B, C. [6]

[6] Note that

$$Pr[A|B] = \frac{Pr[A \cap B}{Pr[B]} = \frac{Pr[A \cap B \cap C}{Pr[B]} + \frac{A \cap B \cap \bar{C}}{Pr[B]}$$

$$\leq \frac{Pr[A \cap B \cap C}{Pr[B \cap C]} + \frac{B \cap \bar{C}}{Pr[B]} = Pr[A|B \cap C] + Pr[\bar{C}|B]. \ \Box$$

Consequently

$$\sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1}] \leq \sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1} \cap H_{i-1}] + Pr[\neg H_{i-1}|G_{i-1}]$$

$$\leq \sum_{i=1}^{M} Pr[q_i \ bad|G_{i-1} \cap H_{i-1}] + 2^{\alpha \cdot n} \cdot e^{-n},$$

which can be easily shown to be in $O\left(2^{-\epsilon_2 \cdot n}\right)$ for some $\epsilon_2 > 0$, as $\log_2(e) > 2/3$.

We have shown

Corollary 2. *Eve's success probability of winning game 3 in phase A is bounded by $O\left(2^{-\epsilon \cdot n}\right)$ for some constant $\epsilon > 0$.* □

For proving lemma 4 we specify first the underlying probability space. Suppose that Eve has posed a sequence Q of $\leq 2^{\alpha \cdot n}$ queries, $\alpha < 2/3$, that E^* holds after Q and that the secret key k^* is still good after Q.

The relevant probability space Ω_Q, which models Eve's knowledge after Q, is formed by all triples $(k, \tilde{P}, \tilde{F})$ such that k is good w.r.t. Q, and $\tilde{P}$ and $\tilde{F}$ are completions of P and F, resp., which make k consistent with Q. From Eve's point of view, all these triples have the same probability to be equal to (k^*, P^*, F^*), where P^* and F^* are the completions of the real P and F on $\{0,1\}^n \setminus U$ and $\{0,1\}^n \setminus Y$, resp.

Consider the induced probability distribution on the set $K^g \subseteq \{0,1\}^n$ of good keys. For each $k \in K^g$ it holds that

$$Pr\left[k = k^*\right] = \frac{|\{(\tilde{P}, \tilde{F}); (k, \tilde{P}, \tilde{F}) \in \Omega_Q\}|}{|\Omega_Q|}.$$

Let us denote this value by $Pr_Q(k)$.

The following lemma is the basis for all our estimations around determining Eve's success probability.

Lemma 6. *(i) The Pr_Q-values of two arbitrarily fixed keys in K^g differ by a factor of at most $e^{2 \cdot 2^{(\alpha - \frac{2}{3})n}}$, which tends to one for increasing n.*

(ii) Suppose that E^* holds and that $|X^*| \leq 2^{(1/3)n}$ after performing a sequence Q of at most $2^{\alpha \cdot n}$ queries. Then Q yields, for some $\epsilon_3 > 0$, at most $O(2^{(1-\epsilon_3)n})$ bad keys.

Note that lemma 6 implies the following corollary

Corollary 3. *If E^* holds and $|X^*| \leq 2^{(1/3)n}$ after performing Q then $Pr_Q(k) \leq 2^{-(n-1)}$ for all $k \in K^g$ if n is large enough.*

Proof of corollary 3: Lemma 6, part (ii), implies, that after performing Q, we have that

$$|K^g| \geq 2^n - 2^{(1-\epsilon)n}$$

(under the assumptions of lemma 6, (ii)). We obtain that $|K^g| \geq \frac{2^n}{\sqrt{2}}$ if n is large enough.

On the other hand, lemma 6, part (i), yields that all Pr_Q-values of elements from K^g differ only by a factor of at most $\sqrt{2}$ if n is large enough.

This implies that $Pr_Q[k] \leq 2 \cdot 2^{-n} = 2^{-(n-1)}$ for all $k \in K^g$. □

The Proof of lemma 6: The proof of part (i) is based on deriving an exact formula for $Pr_Q(k)$ for $k \in K^g$. It will become clear that $Pr_Q(k)$ depends on the values $|Y \cap (V \oplus k)|$ and $|X \cap (U \oplus k)|$. The claim of part (i) will follow from the fact that none of these values exceeds $2^{(1/3)n}$ (see condition b.3 of definition 1). The whole proof of lemma 6, part (i), can be found in section A.

For **proving part (ii) of lemma 6** note first that it follows directly from the definition 1 that there are at most $2 \cdot |X^*| \cdot |U| \leq 2^{((1/3+\alpha)n}$ bad keys of type b.1. As $\alpha < 2/3$ it holds that $1/3 + \alpha < 1$. As E^* is true, it follows directly that the number of keys which are bad w.r.t. b.2 is bounded by $O(2^{(1-\epsilon_0)n})$.

For upper bounding the number of bad keys of type b.3 note that

$$\sum_{k \in \{0,1\}^n} |(X \oplus k) \cap U| = \sum_{k \in \{0,1\}^n} |\{(x,u) \in X \times U; x \oplus u = k\}|$$

$$= |U \times X| \leq 2^{2\alpha n}.$$

This implies that

$$\left|\{k;\ |(X \oplus k) \cap U| > 2^{(1/3)n}\}\right| \leq 2^{(2\alpha-(1/3))n}.$$

In exactly the same way one can prove that

$$\left|\{k;\ |(Y \oplus k) \cap V| > 2^{(1/3)n}\}\right| \leq 2^{(2\alpha-(1/3))n}.$$

Consequently, the number of keys k which fulfill b.3 is bounded by

$$2 \cdot 2^{(2\alpha-(1/3))n}.$$

As $2\alpha - (1/3) < 1$, this proves Lemma 6, part (ii), for bad keys of type b.3. □

From lemma 6 we obtain the following important tool for proving lemma 4.

Lemma 7. *Suppose that E^* holds and that $|X^*| \leq 2^{(1/3)n}$ after performing a sequence Q of at most $2^{\alpha \cdot n}$ queries, and that the secret key k^* is good after Q. Let $X' \subseteq X$, $Y' \subseteq Y$, $U' \subseteq U$, $V' \subseteq V$ be subsets of a fixed size t. Then the success probabilities of Eve when trying to choose*

(1) an element u from $X' \oplus k^*$,
(2) an element x from $U' \oplus k^*$,
(3) an element v from $Y' \oplus k*$,
(4) an element y from $V' \oplus k^*$

are all bounded by $2 \cdot t \cdot |K^g|^{-1}$.

The success probabilities of Eve when trying to choose

(5) an element y from $P(X' \oplus k^*) \oplus k^*$,
(6) an element x from $P^{-1}(Y' \oplus k^*) \oplus k^*$

are all bounded by $86 \cdot t \cdot |K^g|^{-1}$, respectively.

Proof (Lemma 7). We prove first the probability corresponding to item 1, the probabilities corresponding to items 2,3,4 can be proved similarly.

Note that $u \in X' \oplus k^*$ is equivalent to $k^* \in u \oplus X'$ (a set of size $|X'|$). This event has, by lemma 6, part (i), a probability at most $2 \cdot |X'| \cdot |K^g|^{-1}$ if n is large enough.

The proof corresponding to items 5 and 6 is more complicated and can be found in section B. □

Remember that corollary 2 states that the probability that Eve wins in phase A with a sequence Q of at most $2^{\alpha \cdot n}$ queries, $\alpha < \alpha^* \leq 2/3$, is bounded from above by $O(2^{(1-\epsilon)n}$ for some $\epsilon > 0$. For proving this claim we have still to prove Lemma 4 and Lemma 5.

Lemma 4 gives an exponentially small upper bound for Eve's success probability (conditioned to $|X^*| \leq 2^{1/3 \cdot n}$ and condition E^*) in making the secret key bad with during Q. It is proved by listing all possible situations, in which, after a sequence of good queries, a new query makes the secret key bad, and then upper bounding the corresponding probabilities using lemma 7. Concerning condition b.3 we have to show that under the given assumptions, the probability that $|(X \oplus k_1^*) \cap U|$ or $|(Y \oplus k_2^*) \cap V|$ exceeds $2^{1/3 \cdot n}$, is exponentially small. This will be done using a Chernov Bound argument. The complete proof of lemma 4 can be found in section C.

At the end we have to show Lemma 5 saying that the probability that $|X^*|$ exceeds $2^{1/3 \cdot n}$ during a sequence Q of $2^{\alpha \cdot n}$, $\alpha < \alpha^* \leq 2/3$, queries, is exponentially small, if condition E^* holds and the secret key k^* is good after Q. The technical problem here is that Chernov Bound arguments can not be applied in a direct way. The solution we found is to partition the sequence of queries in subintervals and to apply a Chernov Bound argument in a recursive way. The whole proof can be found in section D.

Completing the proof of Theorem 5:

If Eve did not win in phase A, then the secret key k^* is still good after phase A. We estimate now Eve's success probability to win in phase B under the additional condition, that $|X^*| \leq 2^{(1/3)n}$ and that E^* holds after phase A.

Remember that Eve has to submit a pair $(\tilde{x}, \tilde{z})$ with $\tilde{x} \notin X$ and that Eve wins if $E(\tilde{x}) = \tilde{z}$.

We again increase Eve's chance to win by a slight change of the rules of game 3, phase B:

- Eve is allowed to first submit only $\tilde{x}$.
- After this, Alice gives k^* and $P(\tilde{x} \oplus k^*)$ to Eve.
- Now, Eve has to submit $\tilde{z}$ without posing any further oracle query and wins if $E(\tilde{x}) = \tilde{z}$.

Denote $\tilde{y} = P(\tilde{x} \oplus k^*) \oplus k^*$. Note that $E(\tilde{x}) = F(\tilde{y})$.

Clearly, if $\tilde{x} \in P^{-1}(Y \oplus k_2^*) \oplus k_1^*$ (which, by lemma 7, has probability $\leq 86 \cdot 2^{-(n-1)}|Y|$), then Eve wins with probability one, as then $\tilde{y} \in Y$ and Eve can derive the correct answer $\tilde{z} = F(\tilde{y})$ from the answers of phase A.

Consider now the case that $\tilde{x} \notin P^{-1}(Y \oplus k^*) \oplus k^*$ (i.e. if $\tilde{y} \notin Y$). Here, from Eve's point of view, $F(\tilde{y})$ is uniformly distributed. Consequently, Eve's success probability in this case is 2^{-n}.

Altogether, Eve's chance to win in phase B is at most $86 \cdot 2^{(\alpha-1)n} + 2^{-n}$ if $|X^*| \leq 2^{(1/3) \cdot n}$ and E^* holds after phase A.

Let us now put things together and denote by

- E_1 the event that Eve wins game 3 with the sequence of queries Q, $|Q| \leq 2^{\alpha \cdot n}$, $\alpha < \alpha^* \leq 2/3$,
- E_2 the event that the secret key k^* is good after phase A
- E_3 the event that $|X^*| \leq 2^{(1/3) \cdot n}$ after phase A.

Then

$$\begin{aligned}
Pr\,[E_1] &\leq Pr\,[E_1\,|E_2 \cap E_3 \cap E^*\,] + Pr\,[\neg E_2 \cup \neg E_3 \cup \neg E^*] \\
&\leq Pr\,[E_1\,|E_2 \cap E_3 \cap E^*\,] + Pr\,[\neg E_2] + Pr\,[\neg E_3] + Pr\,[\neg E^*] \\
&\leq Pr\,[E_1\,|E_2 \cap E_3 \cap E^*\,] + Pr\,[\neg E_2] + Pr\,[\neg E_3\,|E_2\,] + Pr\,[\neg E_2] + Pr\,[\neg E^*] \\
&= Pr\,[E_1\,|E_2 \cap E_3 \cap E^*\,] + 2 \cdot Pr\,[\neg E_2] + Pr\,[\neg E_3\,|E_2\,] + Pr\,[\neg E^*]\,. \qquad (5)
\end{aligned}$$

Consequently, there is some $\epsilon^* > 0$ such that the probability that Eve wins game 3 with the sequence of queries Q, $|Q| \leq 2^{\alpha \cdot n}$, $\alpha < \alpha^* \leq 2/3$, is bounded by $O(2^{-\epsilon^* \cdot n}$ as this is true for all terms occuring in relation 5. □

if n is large enough. □

5 A Lower Bound for the FP-construction in a generalized Random Oracle Model

In this section we prove a stronger lower bound on the security of the $FP(1)$-construction $E(x) = F(P(x \oplus k^*) \oplus k^*)$, which has the same order as this in Theorem 5, but which refers to a more powerful attacker, who is allowed to pose additional kinds of oracle queries, so called $E^{(r)}$-oracle, where $0 \leq r \leq R-1$.

Our extension is based on an additional public permutation $\pi : \{0,1\}^n \longrightarrow \{0,1\}^n$ and a natural parameter R, which we suppose to be polynomially bounded in n. We assume that for all $z \in \{0,1\}^n$, the period of the sequence $(\pi^r(z))_{r=0}^{\infty}$ is larger than r. We suppose further that Eve can compute function values of π and π^{-1} for her own, i.e., without posing oracle queries.

Posing a query $E^{(r)}(x) =?$ yields the answer

$$E^{(r)}(x) = F\left(\pi^r(P(x \oplus k^*) \oplus k^*)\right).$$

Note that the $E^{(0)}$-oracle equals the E-oracle.

The motivation of introducing this extension of the random oracle model is that it reflects security aspects of stream ciphers which produce the keystream in a packet mode as described in subsection 1.4.

The main result of this section is that the lower bound of Theorem 5 does even hold in the extended random oracle model.

Theorem 6. *Let $\alpha < 2/3$ and suppose an attacker Eve who who tries to compute the secret key $k^* \in \{0,1\}^n$ on the basis of at most $\mathcal{O}(2^{\alpha \cdot n})$ queries to a P-oracle, a P^{-1}-oracle, an F-oracle, and to $E^{(r)}$-oracles for $0 \leq r \leq R-1$. Then there is some $\epsilon > 0$ such that Eve's success probability to compute the secret key k^* is bounded from above by $\mathcal{O}(2^{-\epsilon \cdot n})$, if n is large enough.*

Proof. The proof follows the same line as this in Theorem 5, in a few situations we have to generalize some definitions. We again analyze Eve's success probability to win the following game 1 between Alice and Eve: Alice chooses randomly, independently and w.r.t. the uniform distribution a secret function $F : \{0,1\}^n \longrightarrow \{0,1\}^n$, a secret permutation $P : \{0,1\}^n \longrightarrow \{0,1\}^n$ and a secret key $k^* \in \{0,1\}^n$.

In **phase A** of game 1, Eve asks a sequence of at most $\mathcal{O}(2^{\alpha \cdot n})$ oracle queries to Alice, where $E^{(r)}(x) =?$-queries are answered according to $E(x) = F(\pi^r(P(x \oplus k^*) \oplus k^*))$.

In **phase B**, Eve publishes a pair $(\tilde{x}, \tilde{z})$, where for all r, $0 \leq r \leq R$, $\tilde{x}$ has not been asked in phase A in an $E^{(r)}$-query. Eve wins if $\tilde{z} = E^{(0)}(\tilde{x})$.

Clearly, if Eve succeeds in computing k^* on the basis of the queries posed in phase A, then Eve wins in phase B. Thus, Eve's success probability in computing k^* is bounded from above by the probability that Eve wins game 1.

During asking a sequence of queries Q, Eve generates the following subsets X, $X^{(r)}$, $0 \leq r \leq R-1$, and Y, U, V of $\{0,1\}^n$:

- $X^{(r)} = \{x \in \{0,1\}^n \,;\, Q$ contains an $E^{(r)}$-query w.r.t. $x\}$,
- $X = \bigcup_{r=1}^{R} X^{(r)}$,
- $Y = \{y \in \{0,1\}^n \,;\, Q$ contains an F-query w.r.t. $y\}$.
- $U = \{u \in \{0,1\}^n \,;\, Q$ contains a P-query w.r.t. u, or a P^{-1}-query w.r.t. $P(u)\}$,
- $V = P(U)$.

The main difference to the basic case is that we have to consider an additional type of collisions.

Definition 3. – *A pair $(x, y) \in X \times Y$ is called to be a structural $E^{(r)}F$-collision, $0 \leq r \leq R-1$, if $x \in X^{(r)}$ and $\pi^r(P(x \oplus k^*) \oplus k^*) = y$ (which implies $E^{(r)}(x) = F(y)$).*

- *A pair $(x, x') \in X \times X$, $x \neq x'$, is called to be a structural $E^{(r)}E^{(r')}$-collision, if x occurs in a $E^{(r)}$-query, x' occurs in a $E^{(r')}$-query, where $r \neq r'$, $0 \leq r, r' \leq R-1$, and $\pi^r(P(x \oplus k^*) \oplus k^*) = \pi^{r'}(P(x' \oplus k^*) \oplus k^*)$ (which implies $E^r(x) = E^{r'}(x')$).*

Note that there may occur also collision, which are not structural and caused by collisions of F.

Again, we switch from game 1 to game 2, where in game 2, Alice helps Eve in recognizing structural collisions and gives additional information.

Whenever Eve asks an $E^{(r)}$-query or an F-query causing a structural $E^{(r)}F$-collision, then Alice informs Eve about this and gives a pointer to the collision partner.

Whenever Eve asks an $E^{(r)}$-query x causing a structural $E^{(r)}E^{(r')}$-collision (x, x') then Alice informs Eve about this and gives a pointer to the collision partner x'. Moreover, if x' is not part of an structural $E^{(r')}F$-collision, then Alice gives Eve two structural collisions for free by making $y = \pi^r(P(x \oplus k^*) \oplus k^*)$ and $y' = \pi^{r'}(P(x' \oplus k^*) \oplus k^*) = \pi^{r'-r}(y)$ public for Eve, and putting y and y' to Y.

Note that, as a result of these rules, each input x, which is part of a structural collision is automatically part of a structural $E^{(r)}F$-collision.

Again, the structural collisions occuring during Q build the following set $X^* \subseteq X$, which contain all inputs $x \in X$, which are part of a structural collision, and the set $Y^* \subseteq Y$, containing all elements $y \in \{0,1\}^n$ for which there is a structural $E^{(r)}F$-collision (x, y') for some r, $0 \leq r \leq R-1$, such that $\pi^r(y) = y'$.

Note that the function $P(x \oplus k^*) \oplus k^*$ defines again a bijective mapping from X^* to Y^*.

One difference to Theorem 5 is that it can happen that $Y^* \not\subseteq Y$. However, for all $y \in Y^*$ there is some r, $0 \leq r \leq R-1$, such that $\pi^r(y) \in Y$.

How in this scenario, queries help Eve to eliminate key candidates or to find good candidates for being the secret key?

Let $k \in X \oplus U$, i.e., there are $x \in X$ and $u \in U$ such that $k = x \oplus u$.

Consider the case that $x \in X \setminus X^*$ and that there is some r, $0 \leq r \leq R-1$, such that $y' := \pi^r(P(u) \oplus k)$ belongs to Y. As (x, y') is not a structural collision, it holds that $y' \neq \pi^r(P(x \oplus k^*) \oplus k^*)$, i.e., k cannot be the secret key.

Now consider the case that $x \in X^*$, i.e., x is part of a structural $E^{(r)}F$-collision (x, y'). Then Eve checks if $P(x \oplus k) \oplus k = \pi^{-r}(y')$. If yes, k is a good candidate, if not, k can be discarded. A similar argumentation holds for all $k \in V \oplus Y^*$.

These observations motivate the following definition.

Definition 4. *A key $k \in \{0,1\}^n$ is called to be* ***bad*** *w.r.t. a sequence of queries Q if* *b.1* $k = x \oplus u$ *for some* $x \in X^*$ *and* $u \in U$ *or* $k = v \oplus y$ *for some* $v \in V$ *and* $y \in Y^*$*, or,*

b.2 $k = x \oplus u$*, where* $x \in X^{(r)} \setminus X^*$ *for some* r*,* $0 \leq r \leq R-1$*, and* $u \in U$*, and it holds that* $\pi^r(P(u) \oplus k) \in Y$*, or,*

b.3 It holds that $|(X \oplus k) \cap U| > 2^{(1/3) \cdot n}$ *or* $|(Y \oplus k) \cap V| > 2^{(1/3) \cdot n}$*.*

A key k *which does not fulfilling any of the conditions b.1, b.2, b.3 is called to be* ***good w.r.t.*** Q

Items b.1 and b.2 correspond to the observations made above. Item b.3 ensures, as in the proof of Theorem 5, that all good keys have, from Eve's point of view, approximately the same probability to be the secret key.

Definition 5. *A key* $k \in \{0,1\}^n$ *is called to be* ***consistent with*** Q*, if,* $F|_Y$ *can be completed on* $\{0,1\}^n \setminus Y$ *and* $P|_U$ *on* $\{0,1\}^n \setminus U$ *in such a way that no contradiction with* Q *occurs, i.e.* $E^{(r)}(x) = F(\pi^r(P(x \oplus k) \oplus k))$ *for all* r*,* $0 \leq r \leq R-1$ *and* $x \in X^{(r)}$*.*

An important property of good keys is given in

Lemma 8. *Each* $k \in \{0,1\}^n$*, which is good w.r.t.* Q*, is consistent with* Q*.*

Proof. We construct completions of P' of $P|_U$ on $\{0,1\}^n \setminus U$ and F' of $F|_U$ on $\{0,1\}^n \setminus Y$ which make a good key k consistent with Q.

We run with u through $\{0,1\}^n$ in a certain order and define (if still necessary) the function values $P'(u)$ and $F'(\pi^r(P(u) \oplus k))$, resp. $F'(\pi^r(P'(u) \oplus k))$, in a manner which makes k consistent with Q. We dynamically maintain a set $Target(P')$ which is initially set to $\{0,1\}^n \setminus V$. Whenever we define $P'(u)$ for a *new* u, we delete $P'(u)$ from $Target(P')$.

- **Phase 1**, we define $P'(u)$ for all $u \in \{0,1\}^n$ for which $u \oplus k = x \in X^*$. As k is good, it holds $u \notin U$. We fix the unique $y \in Y^*$ for which $y = P(x \oplus k^*) \oplus k^*$. As k is good, it holds $y \oplus k \notin V$. Thus, we define $P'(u) \leftarrow y \oplus k$.
- **Phase 2** concerns all $u \in U$ for which $u \oplus k \in X^{(r)}$ for some r, $0 \leq r \leq R-1$. Note that $u \oplus k \notin X^*$, otherwise, k would be bad via definition 4, b.1. Moreover, $\pi^r(P(u) \oplus k) \notin Y$, otherwise k would be bad via definition 1, b.2. We set
$$F'(\pi^r(P(u) \oplus k)) \leftarrow E^{(r)}(u \oplus k).$$
- **Phase 3** concerns all $u \notin U$ for which $x = u \oplus k \in X^{(r)} \setminus X^*$ for some r, $0 \leq r \leq R-1$. Here, it is not allowed that $\pi^r(P(x \oplus k) \oplus k) \in Y$. Corresponding to this, we define a set
$$Forbidden(x) = \{y \in \{0,1\}^n \,; \exists r(0 \leq r \leq R-1) \wedge (x \in X^{(r)}) \wedge (\pi^r(y) \in Y)\},$$
and choose
$$P'(u) \in Target(P') \setminus (Forbidden(x) \oplus k).$$

Note that for all remaining $u \in \{0,1\}^n$ and $y \in \{0,1\}^n$, the values of $P'(u) \in Target(P')$ and $F'(y) \in \{0,1\}^n$ can be freely chosen without generating contradictions with the E-queries occuring in Q. □

Note that the proof gives information how completions P' of $P|_U$ on $\{0,1\}^n \setminus U$ and F' of $F|_U$ on $\{0,1\}^n \setminus Y$ look like, which are consistent with the assumption that a good key k is the secret key. In particular,

- Q determines the values of P' on $X^* \oplus k$, a set of size $|X^*|$.
- Q implies that the function values of P' on the set $((X \setminus X^*) \oplus k) \setminus U$ are forbidden to be in a set of at most $|Y|$ elements.
- Q determines the values of F' on a set of size $|X \setminus X^*|$.

We consider again a slightly modified **game 3**. The difference with game 2 is that Eve wins immediately in phase A if she manages to pose a query which makes k^* bad. Clearly, the success probability of Eve for winning game 1 is bounded from above by the success probability of Eve for winning game 3.

Starting from here, the proof can be completed by exactly the same arguments used in the proof of Theorem 5. At some points, the probabilities for making the secret key bad when asking queries or for generating a structural collision increase by a factor of R. But this does not destroy anything as R is supposed to be polynomially bounded in n. □

6 Conclusion and Acknowledgement

In this work, we studied FP-construction, which are key-alternating PRF construction. We proved sharp security bounds for two basic types of FP-constructions in the random oracle model, namely a sharp $n/2$-bound for the $F(0)$-construction and a sharp $2/3 \cdot n$-bound for the $FP(1)$-construction. We showed further, how FP-constructions can be used to model the state initialization mechanisms of keystream generator based stream ciphers, and how corresponding operation modes for KSG-based stream ciphers can be defined which provide provable security w.r.t. taime-space-data tradeoff attacks beyond the birthday bound.

One interesting problem is to find out if a lower bound of the same order can be shown for distinguishing the FP(1)-PRF from a truly random permutation. Another open problem is to show sharp bounds on the security of FP-constructions of higher iteration depth. In analogy to Even-Mansour ciphers, we conjecture that the security level of the $FP(r)$-construction of iteration depth r over r independent permutations over $\{0,1\}^n$, and $r+1$ independent n-bit keys (see figure 4) is $\frac{r+1}{r+2}n$.

Acknowledgments I would like to thank Pavel Pudlak, Stefan Lucks, Andreas Jakoby, Frederik Armknecht, and Matthias Hamann for many fruitful discussions.

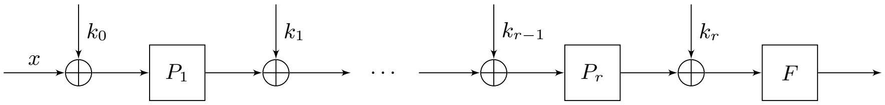

Fig. 4. The $FP(r)$-Construction.

References

1. Elena Andreeva, Andrey Bogdanov, Yevgeniy Dodis, Bart Mennink, and John P. Steinberger. On the indifferentiability of key-alternating ciphers. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 531–550. Springer, 2013.
2. Frederik Armknecht and Vasily Mikhalev. On lightweight stream ciphers with shorter internal states. FSE - 22. International Workshop on Fast Software Encryption, 2015.
3. S.H. Babbage. Improved "exhaustive search" attacks on stream ciphers. In *Security and Detection, 1995., European Convention on*, pages 161–166, May 1995.
4. Andrey Bogdanov, LarsR. Knudsen, Gregor Leander, Francois-Xavier Standaert, John Steinberger, and Elmar Tischhauser. Key-alternating ciphers in a provable setting: Encryption using a small number of public permutations. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology – EUROCRYPT 2012*, volume 7237 of *Lecture Notes in Computer Science*, pages 45–62. Springer Berlin Heidelberg, 2012.
5. Christophe De Cannière and Bart Preneel. Trivium - specifications (eSTREAM). Technical report, ECRYPT (European Network of Excellence for Cryptology), 2005. `http://www.ecrypt.eu.org/stream/p3ciphers/trivium/trivium_p3.pdf`.
6. Shan Chen, Rodolphe Lampe, Jooyoung Lee, Yannick Seurin, and John Steinberger. Minimizing the two-round even-mansour cipher. In JuanA. Garay and Rosario Gennaro, editors, *Advances in Cryptology – CRYPTO 2014*, volume 8616 of *Lecture Notes in Computer Science*, pages 39–56. Springer Berlin Heidelberg, 2014.
7. Shan Chen and John Steinberger. Tight security bounds for key-alternating ciphers. In PhongQ. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology – EUROCRYPT 2014*, volume 8441 of *Lecture Notes in Computer Science*, pages 327–350. Springer Berlin Heidelberg, 2014.
8. Joan Daemen. Limitations of the even-mansour construction. In Hideki Imai, RonaldL. Rivest, and Tsutomu Matsumoto, editors, *Advances in Cryptology — ASIACRYPT '91*, volume 739 of *Lecture Notes in Computer Science*, pages 495–498. Springer Berlin Heidelberg, 1993.
9. Itai Dinur, Orr Dunkelman, Nathan Keller, and Adi Shamir. Cryptanalysis of iterated even-mansour schemes with two keys. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014*, volume 8873 of *Lecture Notes in Computer Science*, pages 439–457. Springer Berlin Heidelberg, 2014.
10. Orr Dunkelman, Nathan Keller, and Adi Shamir. Minimalism in cryptography: The even-mansour scheme revisited. In *Proceedings of the 31st Annual Interna-*

tional Conference on Theory and Applications of Cryptographic Techniques, EUROCRYPT'12, pages 336–354, Berlin, Heidelberg, 2012. Springer-Verlag.

11. Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. In Hideki Imai, RonaldL. Rivest, and Tsutomu Matsumoto, editors, *Advances in Cryptology — ASIACRYPT '91*, volume 739 of *Lecture Notes in Computer Science*, pages 210–224. Springer Berlin Heidelberg, 1993.
12. Matthias Hamann and Matthias Krause. Stream cipher operation modes with provable security beyond the birthday bound against generic collision attacks. [To appear in Cryptology ePrint Archive], 2015.
13. Martin Hell, Thomas Johansson, and Willi Meier. Grain - a stream cipher for constrained environments (eSTREAM). Technical report, ECRYPT (European Network of Excellence for Cryptology), 2005. `http://www.ecrypt.eu.org/stream/p3ciphers/grain/Grain_p3.pdf`.
14. Rodolphe Lampe, Jacques Patarin, and Yannick Seurin. An asymptotically tight security analysis of the iterated even-mansour cipher. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology – ASIACRYPT 2012*, volume 7658 of *Lecture Notes in Computer Science*, pages 278–295. Springer Berlin Heidelberg, 2012.
15. 3GPP Organizational Partners. 3GPP TS 55.216 V6.2.0 (2003-09), 2003. `http://www.gsma.com/technicalprojects/wp-content/uploads/2012/04/a53andgea3specifications.pdf`.
16. Jacques Patarin. The "coefficients H" technique. In RobertoMaria Avanzi, Liam Keliher, and Francesco Sica, editors, *Selected Areas in Cryptography*, volume 5381 of *Lecture Notes in Computer Science*, pages 328–345. Springer Berlin Heidelberg, 2009.
17. Bluetooth SIG. Bluetooth core specification 4.2, 2014. `https://www.bluetooth.org/DocMan/handlers/DownloadDoc.ashx?doc_id=286439`.

A Proof of Lemma 6, part (i)

Proof of lemma 6, part (i):

We have to show that the Pr_Q-values of two arbitrarily fixed keys in K^g differ by a factor of at most $e^{2\cdot 2^{(\alpha-\frac{2}{3})n}}$, which tends to one for increasing n.

For all keys $k \in K^g$ let $ConsPF_Q(k)$ denote the set of all pairs $(\tilde{P}, \tilde{F})$ for which $k, \tilde{P}, \tilde{F}) \in \Omega_Q$.

Note that

$$Pr_Q(k) = \frac{|ConsPF_Q(k_1, k_2)|}{|\Omega_Q|}.$$

Let us further denote by $ConsP_Q(k)$ the set of all completions $\tilde{P}$ of P on $\{0,1\}^n \setminus U$ for which there is some completion $\tilde{F}$ of F on $\{0,1\}^n \setminus Y$ which yields $(k, \tilde{P}, \tilde{F}) \in \Omega_Q$.

Moreover, for all $\tilde{P} \in ConsP_Q(k)$ let us denote by $ConsF_Q(k, \tilde{P})$ the set of all completion $\tilde{F}$ of F on $\{0,1\}^n \setminus Y$ which yields $(k, \tilde{P}, \tilde{F}) \in \Omega_Q$. Note that

$$|ConsPF_Q(k)| = \sum_{\tilde{P} \in ConsP_Q(k)} |ConsF_Q(k, \tilde{P})|.$$

At least note that for all $k \in K^g$ and all $\tilde{P} \in ConsP_Q(k)$ it holds that $|ConsF_Q(k, \tilde{P})|$ is equal, namely

$$|ConsF_Q(k, \tilde{P})| = 2^{\mu(Q)},$$

where $\mu(Q) = 2^n - |Y| - |X \setminus X^*|$.

This is because the E-queries in Q determine the $\tilde{F}$-values in $\tilde{P}((X \setminus X^*) \oplus k) \oplus k$. Outside of this set, the $\tilde{F}$-values can be arbitrarily chosen. The fact that k is good implies that

$$\left(\tilde{P}((X \setminus X^*) \oplus k) \oplus k\right) \cap Y = \emptyset.$$

This implies that for all k and k' from K^g it holds that

$$\frac{Pr_Q(k)}{Pr_Q(k')} = \frac{|ConsP_Q(k)|}{|ConsP_Q(k')|}.$$

At next observe that for all $\tilde{P} \in ConsP(k)$ and $u \in ((X \setminus X^*) \oplus k) \setminus U$ it has to be true that

$$\tilde{P}(u) \notin (Y \oplus k) \setminus V.$$

(See phase 4 of the algorithm described in the proof of lemma 2).

Hence, $|ConsP_Q(k)|$ depends on $|(X \setminus X^*) \oplus k) \cap U|$ and $|(Y \oplus k) \cap V|$, which can vary between 0 and $\Delta := 2^{(1/3)n}$ (Here, item b.3 of the definition 1 comes in).

For $k \in K^g$, the value $|ConsP_Q(k)|$ is minimal if

$$|X \oplus k) \cap U| = |(Y \oplus k) \cap V| = 0.$$

In this case we have that

$$|ConsP_Q(k)| = T! \cdot T(T-1) \cdots (T-(t-1))$$

for $T = 2^n - |U| - |X|$ and $t = |X|$. (Let us suppose $|X| = |Y|$.)

On the other hand, $|ConsP_Q(k)|$ is maximal if

$$|(X \setminus X^*) \oplus k) \cap U| = |(Y \oplus k_2) \cap V| = \Delta,$$

which implies that

$$|ConsP_Q(k)| = (T+\Delta)! \cdot (T+\Delta) \cdots (T+\Delta-(t-\Delta-1)).$$

This implies that the Pr_Q-values of elements from K^g can differ by a factor which is at most $P_1 \cdot P_2$, where

$$P_1 = \frac{(T+\Delta)(T+\Delta-1)\cdots(T+\Delta-(t-\Delta-1))}{T(T-1)\cdots(T-(t-\Delta-1))},$$

$$P_2 = \frac{(T+1)(T+2)\cdots(T+\Delta)}{(T-t+1)(T-t+2)\cdots(T-t+\Delta)}.$$

Note that

$$P_1 = \frac{(T+1)(T+2)\cdots(T+\Delta)}{(T-(t-\Delta)+1)\cdots(T-(t-\Delta)+\Delta)}$$

and that

$$P_1 \cdot P_2 \leq \left(\frac{T}{T-(t-\Delta)}\right)^{\Delta} \cdot \left(\frac{T}{T-t}\right)^{\Delta} \leq \left(\frac{T}{T-t}\right)^{2\Delta}.$$

We derive a lower bound for the inverse of $\left(\frac{T}{T-t}\right)^{2\Delta}$.

Note that

$$\left(\frac{T-t}{T}\right)^{2\Delta} = \left(\left(1-\frac{1}{T/t}\right)^{T/t}\right)^{2\Delta\cdot(t/T)}$$

$$\approx e^{-\frac{2\Delta\cdot t}{T}} \geq e^{-4\Delta\cdot 2^{(\alpha-1)n}}.$$

Here we used that $t \leq 2^{\alpha\cdot n}$ and $T \geq \frac{1}{2}2^n$ if n large enough.

Inserting $\Delta = 2^{(1/3)n}$ completes the proof of part (i) of lemma 6. □

□

B Completing the Proof of Lemma 7, Items (5) and (6)

Let $X' \subseteq X$, $|X'| = t$. We have to estimate the success probability of Eve to choose an element y from $P(X' \oplus k^*) \oplus k^*$ with the next query after Q (which we suppose to be an F-query).

Therefore, we fix some $x \in X$ and estimate Eve's success probability for the event $Succ_x$ that Eve succeeds with choosing y such that $y = P(x \oplus k^*) \oplus k^*$.

We divide K^g into to subsets K_1^g and K_2^g, where

$$K_1^g = \{k \in K^g ; x \oplus k \notin U\},$$

$$K_2^g = \{k \in K^g ; x \oplus k \in U\}.$$

Note that

$$Pr[Succ_x] = Pr[k^* \in K_1^g] \cdot Pr[Succ_x | k^* \in K_1^g] + Pr[k^* \in K_2^g] \cdot Pr[Succ_x | k^* \in K_2^g]$$

$$\leq Pr[Succ_x | k^* \in K_1^g] + Pr[k^* \in K_2^g] \cdot Pr[Succ_x | k^* \in K_2^g].$$

Note further that for all $k \in K_1^g$ Eve does not know $P(x \oplus k)$. Consequently, for all $y \in \{0,1\}^n$ it holds the following. If $k = k^*$ then Eve succeeds with y if and only if $P(x \oplus k) = y \oplus k$. The probability for this event is at most $1/(2^n - |U|) \leq 2^{-(n-1)}$. Consequently,

$$Pr[Succ_x | k^* \in K_1^g] \leq 2^{-(n-1)}.$$

For estimating $Pr[Succ_x(y) | k^* \in K_2^g]$ note first that it may happen that $P(x \oplus k) \oplus k = P(x \oplus k') \oplus k'$ for different keys $k \neq k'$. Eve maximizes her success probability if she chooses a maximal set $K \subseteq K^g$ such that $P(x \oplus k) \oplus k$ is equal for all $k \in K$, and guesses that $k^* \in K$, i.e., that $P(x \oplus k^*) \oplus k^* = y$ where $y = P(x \oplus k) \oplus k$ for some $k \in K$. In this case, Eve's success probability is at most $2 \cdot |K| \cdot |K^g|^{-1}$.

We estimate Eve's success probability by showing that the probability that $|K| \geq 6$ is sufficiently small.

Consider the following equivalence relation on U. For all $u, u' \in U$ let us write

$$u \equiv_P u' \iff P(u) \oplus P(u') = u \oplus u'.$$

Let us further denote by $Max(P, U)$ the size of the largest equivalence class w.r.t. $\equiv_P$.

Note that if $u \equiv_P u'$ and $k = u \oplus x$ and $k' = u' \oplus x$ then $P(x \oplus k) \oplus k = P(x \oplus k') \oplus k'$.

This means, that under the condition that $k^* \in K_2^g$, Eve maximizes her success probability by doing the following:

- computing a maximal set $K \subseteq K_2^g$ such that $x \oplus k \equiv_P x \oplus k'$ for all $k, k' \in K$,
- choosing y in such a way that $y = P(x \oplus k) \oplus k$ for some (i.e., for all) all $k \in K$.

We obtain that

$$Pr[Succ_x|k^* \in K_2^g] \leq 2 \cdot \frac{|K|}{|K_2^g|},$$

if n is large enough. Taking into account that $|K| \leq Max(P,U)$ we obtain that

$$Pr[Succ_x|k^* \in K_2^g] \leq Pr[Max(P,U) < 6] \cdot Pr[Succ_x(y)|k^* \in K_2^g, Max(P,U) < 6]$$

$$+Pr[Max(P,U) \geq 6] \cdot Pr[Succ_x(y)|k^* \in K_2^g, Max(P,U) \geq 6]$$

$$\leq Pr[Succ_x(y)|k^* \in K_2^g, Max(P,U) < 6] + Pr[Max(P,U) \geq 6]$$

$$\leq \frac{10}{|K_2^g|} + Pr[Max(P,U) \geq 6].$$

Consequently,

$$Pr[k^* \in K_2^g] \cdot Pr[Succ_x|k^* \in K_2^g] \leq 2\frac{|K_2^g|}{|K^g|}\left(\frac{10}{|K_2^g|} + Pr[Max(P,U) \geq 6]\right)$$

$$\leq \frac{20}{|K^g|} + 2 \cdot Pr[Max(P,U) \geq 6].$$

Clearly, the event $Max(P,U) \geq 6$ implies the existence of some $U' \subseteq U$, $|U'| = 6$, such that $u \equiv_P u'$ for all $u, u' \in U'$. Given a subset $U' \subseteq U$ with 6 elements, the probability for the event that $u \equiv_P u'$ for all $u, u' \in U'$ equals

$$\prod_{i=1}^{5} \frac{1}{2^n - i} \leq \left(\frac{1}{1/2 \cdot 2^n}\right)^5 = 2^5 \cdot 2^{-5 \cdot n}.$$

Consequently,

$$Pr[Max(P,U) \geq 6] \leq |U|^6 \cdot 2^5 \cdot 2^{-5 \cdot n} \leq 2^5 \cdot 2^{(2/3)n} \cdot \left(2^{-(1/3)n}\right)^5.$$

Consequently, for $r = 6$, it holds

$$Pr[Max(P,U) \geq 6] \leq 32 \cdot 2^{-n}.$$

Putting all things together we obtain that

$$Pr[Succ_x(y)] \leq 2^{-(n-1)} + \frac{20}{|K^g|} + 64 \cdot 2^{-n} \leq \frac{20}{|K^g|} + 66 \cdot 2^{-n} < \frac{86}{|K^g|}.$$

The proof of item (6) can be done by using exactly the same arguments. □

C Proof of Lemma 4

Let $Q = (q_1, \cdots, q_M)$, be a sequence of queries of $M \leq 2^{\alpha \cdot n}$ queries, which fulfils E^*. For all $i = 1, \cdots, M$ let G_i denote the event that k^* remained good after $(q_1, \cdots, q_i)$ and H_i the event that $|X^*| \leq 2^{(1/3)n}$ after $(q_1, \cdots, q_i)$. We have to show that

$$\sum_{i=1}^{M} Pr[q_i bad | G_{i-1} \cap H_{i-1}] \leq O\left(2^{-\epsilon_2 \cdot n}\right)$$

for some $\epsilon_2 > 0$.

We estimate first the probabilities that during Q the secret key k^* becomes bad via item b.1 of the definition 1 of bad keys, In a last step we then estimate the probability that during Q the secret key k^* becomes bad via item b.3. Note that k^* can never fulfill item b.2 of definition 1.

We will distinguish four cases corresponding to the four possible types of queries. We fix an arbitrary i, $1 \leq i \leq M$.

Case 1: Suppose that query q_i asks for $P(u)$ for some $u \notin U$:

1.1 If $u \in X^* \oplus k^*$ then k^* becomes bad (due to corollary 2, the probability for this is at most $2^{-(n-1)} \cdot |X^*| \leq 2 \cdot 2^{-(2/3) \cdot n}$).
1.2 If $u \in P^{-1}(Y^* \oplus k^*)$ then k^* becomes also bad (probability also at most $2^{-(n-1)} \cdot |Y^*| \leq 2 \cdot 2^{-(2/3) \cdot n}$).
1.3 If $u \notin X^* \oplus k^*$ and $P(u) \notin Y^* \oplus k^*$ then k^* remains good.

The reason for the last item is that from $u \in (X \setminus X^*) \oplus k^*$ it follows $P(u) \notin Y \oplus k^*$. Otherwise we had $E(u \oplus k^*) = F(P(u) \oplus k^*)$ which would imply $u \oplus k^* \in X^*$.

Case 2: Query q_i asks for $P^{-1}(v)$ for some $v \notin V$:

2.1 If $v \in Y^* \oplus k^*$ then k^* becomes bad (probability at most $2^{-(n-1)} \cdot |X^*| \leq 2 \cdot 2^{-(2/3) \cdot n}$).
2.2 If $v \in P(X^* \oplus k^*)$ then k^* becomes also bad (probability at most $2^{-(n-1)} \cdot |X^*| \leq 2 \cdot 2^{-(2/3) \cdot n}$).
2.3 If $v \notin (Y^* \oplus k^*) \cup P(X^* \oplus k^*)$ then k^* remains good (similar argument as in case 1.3).

This implies that the overall probability that, during Q, the secret k^* will be made to fulfill item b.1 via a P- or a P^{-1}-query is bounded by $2^{\alpha \cdot n} \cdot 2^{-(2/3) \cdot n}$ which belongs to $O\left(2^{-\epsilon \cdot n}\right)$ for some $\epsilon > 0$.

Case 3: Query q_i asks for $E(x)$ for some $x \notin X$. Let $y = P(x \oplus k^*) \oplus k^*$. Then k^* becomes bad if and only if $x \oplus k^* \in U$ and $y = P(x \oplus k^*) \oplus k^* \in Y$. (In this case, Eve sees that $E(x) = F(y)$, gets the information from Alice that $y = P(x \oplus k^*) \oplus k^*$, and x will be added to X^* and y to Y^*.)

Note that this is equivalent to the event that $k^* \in K_x^g$, where

$$K_x^g = \{k \in K^g ; x \oplus k \in U \ \wedge P(x \oplus k) \oplus k \in Y\}$$

Consequently, the probability that q_i makes k^* bad via item b.1 is bounded by $2 \cdot 2^{-n}|K_x^g|$.

This implies that the overall probability that, during Q, the secret k^* will be made to fulfill item b.1 via an E-query is bounded by

$$2 \cdot 2^{-n} \cdot \sum_{x \in X} |K_x^g|.$$

Note that

$$\bigcup_{x \in X} K_x^g \subseteq Bad_(Q)$$

and remember that $|Bad_2(Q)| \leq 2^{(1-\epsilon_0)n}$ (see corollary 1), as Q fulfills E^*. Consequently, the probability that k^* becomes bad via some E-query during Q is bounded by $2 \cdot 2^{-(n-1)}$.

The overall probability of **Case 4** that, during Q, an F-query makes k^* fulfilling item b.1 can be estimated in the same way.

We fix now some number N, $1 \leq N \leq |Q|$, and estimate the probability for the event $\tilde{E}$ that k^* becomes bad via b.3 through query q_N under the condition that k^* remained good after $q_1, \cdots, q_{N-1}$. Note that $\tilde{E}$ implies that the maximum of $|(X \oplus k^*) \cap U|$ and $|(Y \oplus k^*) \cap V|$ is $\lfloor 2^{(1/3)n} \rfloor$ after $q_1, \cdots, q_{N-1}$ and $\lfloor 2^{(1/3)n} \rfloor + 1$ after $q_1, \cdots, q_N$.

For $i = 1, \cdots, N$ we denote by $C_i \in \{0,1\}$ a random $\{0,1\}$-variable for which $C_i = 1$ if q_i enlarges $|(X \oplus k^*) \cap U|$ or $|(Y \oplus k^*) \cap V|$, and denote $C = \sum_{i=1}^{N} C_i$. We obtain that

$$Pr[\tilde{E}] \leq Pr\left[C > 2^{(1/3)n}\right].$$

We complete the proof by showing that $Pr\left[C > 2^{(1/3)n}\right] < e^{-n} < 2^{-(2/3)n}$.

Note that $C_i = 1$ if and only if Eve is successful in asking with the i-th query

- for $P(u)$ for some u from $X \oplus k^*$, or
- for $P^{-1}(v)$ for some v from $Y \oplus k^*$, or
- for $E(x)$ for some x from $U \oplus k^*$, or
- for $F(y)$ for some y from $V \oplus k^*$.

Here, X, Y, U, V denote the corresponding sets after performing the first $i-1$ queries along Q.

Due, to Lemma 7 the probability of each of the four events is bounded by $2 \cdot 2^{(\alpha-1)n}$, i.e.,

$$Pr[C_i = 1] \leq 2 \cdot 2^{(\alpha-1)n}.$$

We estimate the probability that $\sum_{i=1}^{N} C_i$ exceeds $2^{(1/3)n}$ by using the following technique called **Chernov Bounds** in the literature.

Theorem 7. *Let $p \in (0, 1)$, and $A_1, \cdots, A_N$ be a set of independent random variables, where for all $i = 1, \cdots, N$ holds that $Pr[A_i = 1 - p] = p$ and $Pr[A_i = -p] = 1 - p$. Let $A = \sum_{i=1}^{N} A_i$. Then*

$$Pr[A > a] < e^{-\frac{2a^2}{N}} \tag{6}$$

for all $a > 0$.

For a **proof** see, e.g., Alon, Spencer, Erdos, The Probabilistic Method, Wiley Interscience 1992, Theorem A4 on page 235. □

We derive from Theorem 7 a corresponding result for random $\{0, 1\}$-variables.

Lemma 9. *Let $B_i = A_i + p$. Note that $B_i \in \{0, 1\}$ and $Pr[B_i = 1] = p$. Let $B = \sum_{i=1}^{N} B_i$. Then, for all $\delta > 0$, it holds*

$$Pr[B > (p + \delta)N] < e^{-2\delta^2 N}. \tag{7}$$

Proof: By definition, $B = A + N \cdot p$. The proof is completed by putting $a = \delta \cdot N$ into the relation in Theorem 7. □

We will apply Chernov Bound arguments in the following modified scenario.

Lemma 10. *Let $C_1, \cdots, C_N$ denote a collection of (not necessarily independent) random $\{0, 1\}$-variables fulfilling $Pr[C_i = 1] = p_i < p$ for all i, $1 \le i \le N$, and some p, $0 < p < 1$. Let $C = \sum_{i=1}^{N} C_i$.*

We suppose that, for $i > 1$, the probabilities p_i depend deterministically on (and can be computed from) the outcomes of the experiments $E_1, \cdots, E_{i-1}$ behind $C_1, \cdots, C_{i-1}$.

Then, for all $\delta > 0$, it holds

$$Pr[C > (p + \delta)N] < e^{-2\delta^2 N}.$$

At several places we will take $\delta = \sqrt{\Delta/(2N)}$ and obtain

$$Pr[C > (p + \delta)N] = Pr\left[C > pN + \sqrt{\frac{\Delta \cdot N}{2}}\right] < e^{-\Delta}. \tag{8}$$

Proof: We construct a collection of mutually independent binary random variables $B_1, \cdots, B_N$ satisfying

- $C_i = 1$ implies $B_i = 1$,
- $Pr[B_i = 1] = p$.

for all i, $1 \leq i \leq N$.

This proves our lemma 10, as $\sum_{i=1}^{N} C_i \leq \sum_{i=1}^{N} B_i$ with probability one, and as lemma 9 can be applied to $B = \sum_{i=1}^{N} B_i$.

The experiments $\tilde{E}_i$ behind B_i are for all i, $1 \leq i \leq N$, defined as follows:

- Compute p_i from the outcomes of the experiments $E_1, \cdots, E_{i-1}$.
- Perform E_i and output one (i.e. $B_i = 1$) if E_i is successful (i.e. if $C_i = 1$).
- If E_i is not successful (i.e. $C_i = 0$) then perform a completely independent experiment E_i' with success probability $q_i = \frac{p-p_i}{1-p_i}$ and output one (i.e. $B_i = 1$) if E_i' is successful.

Note that $Pr[B_i = 1]$ equals

$$Pr[Y_i = 1|X_i = 1] \cdot Pr[X_i = 1] + Pr[Y_i = 1|X_i = 0] \cdot Pr[X_i = 0]$$

$$= 1 \cdot p_i + q_i \cdot (1 - p_i) = p. \ \square$$

For completing the proof of lemma 4 we go into relation (8) and set $p = 2 \cdot 2^{(\alpha-1)n}$ and $\Delta = n$. We obtain that

$$Pr\left[\sum_{i=1}^{N} C_i > 2 \cdot 2^{(\alpha-1)n} \cdot 2^{\alpha \cdot n} + \sqrt{n/2} \cdot 2^{(\alpha/2)n}\right] < e^{-n}.$$

From $\alpha < 2/3$ it follows that

$$2 \cdot 2^{(2\alpha-1)n} + \sqrt{n/2} \cdot 2^{(\alpha/2)n} < 2^{(1/3)n}$$

if n is large enough. Consequently, $Pr\left[\sum_{i=1}^{N} C_i \geq 2^{(1/3)n}\right] < e^{-n}$.

This completes the proof of lemma 4. $\square$

D Proof of Lemma 5

Proof (Lemma 5).

Let $\alpha = \alpha^* - \epsilon$ for some $\epsilon > 0$, and let $Q = (q_1, \cdots, q_N)$, $N \leq 2^{\alpha \cdot n}$, denote a sequence of queries fulfilling E^*. We have to show that if the secret key k^* is good after Q then the probability that $|X^*|$ exceeds $2^{1/3 \cdot n}$ is smaller than $e^{-n} < 2^{-2/3 \cdot n}$, if n is large enough.

Note first that if F and E were independent random functions then, for each E- and F-query in Q, the probability to generate an EF-collision is at most $2^{(\alpha-1)n}$. Thus, the proof of lemma 5 could be completed with the same Chernov Bound argument (Lemma 8) used above.

But here the situation is different: Note that an query q generates an EF-collision of type 1 iff

E1 q asks for $E(x)$ for some $x \in P^{-1}(Y \oplus k^*) \oplus k^*$, or
E2 q asks for $F(y)$ for some $y \in P(X \oplus k^*) \oplus k^*$.

The probability for this is at most $86 \cdot \frac{|Y|}{|K^g|}$ in case E1 and $86 \cdot \frac{|X|}{|K^g|}$ in case E2 (see lemma 7, items 5,6). The problem is that these probability increases with $|X^*|$ and $|Y^*|$.

We denote by D_i, $1 \leq i \leq N$, a random $\{0,1\}$-variable which outputs one if, with the i-th query q_i, Eve is successfull in asking for some

- $x \in P^{-1}(Y \oplus k^*) \oplus k^*$, or some
- $y \in P(X \oplus k^*) \oplus k^*$

We denote by p_i the probability that $D_i = 1$ and by ψ_i the size of $|X^*|$ before posing query q_i. We suppose that n is large enough such that the number of keys which are bad with respect to b.2 and b.3 of definition 1 is smaller than $1/2 \cdot 2^n$ (this can be supposed due to lemma 6). It follows that

$$\begin{aligned} p_i &\leq 86 \cdot \frac{2^{\alpha \cdot n}}{1/2 \cdot 2^n - 2 \cdot 2^{\alpha \cdot n} \cdot \psi_i} \\ &\leq \frac{174}{2^{(1-\alpha)n} - 4 \cdot \psi_i}. \end{aligned}$$

We have to show that

$$Pr\left[\sum_{i=1}^{N} D_i > 2^{1/3 \cdot n}\right] < e^{-n}.$$

Note that, in the worst case, p_i can become 1 if $i \geq 1/4 \cdot 2^{(1-\alpha)n}$.

Consequently, for $N \geq 1/4 \cdot 2^{(1-\alpha)n}$, the probability that $\sum_{i=1}^{N} D_i$ exceeds $2^{1/3 \cdot n}$ cannot be estimated by a direct application of Chernov Bounds.

In the following, we have to estimate probabilities of type

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > M \,\middle|\, \sum_{i=1}^{j} \leq R\right]$$

and of type

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > M \,\middle|\, \sum_{i=1}^{j} D_i \leq R, \sum_{i=j+1}^{j+N} D_i \leq S\right].$$

Note that these probabilities does not depend on j.

We will use a recursive argument which is based on the following idea. Suppose that we know that

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > 2^{1/3 \cdot n} \,\middle|\, \sum_{i=1}^{j} D_i \leq R \cdot 2^{1/3 \cdot n}\right] < e^{-C \cdot n}$$

for some $N < 2^{\alpha \cdot n}$, $C > 0$, $R \geq 1$. Let $M < 2^{1/3 \cdot n}$. Then

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > M \middle| \sum_{i=1}^{j} D_i \leq R \cdot 2^{1/3 \cdot n}\right] \leq$$

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > M \middle| \sum_{i=1}^{j} D_i \leq R \cdot 2^{1/3 \cdot n}, \sum_{i=j+1}^{j+N} D_i \leq 2^{1/3 \cdot n}\right] + e^{-C \cdot n}.$$ [7]

If $\sum_{i=1}^{j} D_i \leq R \cdot 2^{1/3 \cdot n}$ and $\sum_{i=j+1}^{j+N} D_i \leq 2^{1/3 \cdot n}$ then we can derive the following upper bound for all p_i, $1 \leq i \leq j+N$:

$$p_i < \frac{174}{2^{(1/3+\epsilon)n} - 4 \cdot (R+1) \cdot 2^{1/3 \cdot n}} < 2^{-1/3 \cdot n}$$

if n large enough. Now, the Chernov Bound argument (Lemma 8) can be applied directly. For $M \geq (2^{-1/3 \cdot n} + \sqrt{\Delta/(2N)})N$ it holds

$$Pr\left[\sum_{i=j+1}^{j+N} D_i > M \middle| \sum_{i=1}^{j} D_i \leq R \cdot 2^{1/3 \cdot n}, \sum_{i=j+1}^{j+N} D_i \leq 2^{1/3 \cdot n}\right] < e^{-\Delta}.$$

Now suppose that $M < 2^{1/3 \cdot n}$ and let $a \leq 2^{1/3 \cdot n}/M$. We obtain

$$Pr\left[\sum_{i=j+1}^{j+a \cdot N} D_i > 2^{1/3 \cdot n} \middle| \sum_{i=1}^{j} D_i \leq (R-1) \cdot 2^{1/3 \cdot n}\right]$$

$$\leq Pr\left[\sum_{i=j+1}^{j+a \cdot N} D_i > a \cdot M \middle| \sum_{i=1}^{j} D_i \leq (R-1) \cdot 2^{1/3 \cdot n}\right]$$

$$\leq a \cdot Pr\left[\sum_{i=j'+1}^{j'+N} D_i > M \middle| \sum_{i=1}^{j'} D_i \leq R \cdot 2^{1/3 \cdot n}\right]$$

$$< a \cdot (e^{-\Delta} + e^{-C \cdot n}).$$

For seeing this, we divide the interval $j+1, \cdots, j+a \cdot N$ into a consecutive intervals of length N.

The event that $\sum_{i=j+1}^{j+a \cdot N} D_i > a \cdot M$, under the condition that $\sum_{i=1}^{j} D_i \leq (R-1) \cdot 2^{1/3 \cdot n}$, implies that the number of hits in some of these subintervals of length N has to exceeds M.

If this is not the case for all subintervals, then, additionally, for all subintervals holds, that less then $(R-1) \cdot 2^{1/3 \cdot n} + a \cdot M \leq R \cdot 2^{1/3 \cdot n}$ hits happened before starting with the subinterval.

[7] as $Pr(A) = Pr(A|B)Pr(B) + Pr(A|\neg B)Pr(\neg B) \leq Pr(A|B) + Pr(\neg B)$.

This implies a recursive argument.

For all natural $k \geq 0$ let

$$\alpha_k = \frac{1}{3}\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^k}\right).$$

Note that

$$\lim_{k \to \infty} \alpha_k = 2/3.$$

Note further that for all $k \geq 0$

$$\alpha_{k+1} = \frac{1}{3} + \frac{\alpha_k}{2}.$$

Define now for all natural $k \geq 0$ and $\delta > 0$

$$\tilde{\alpha_k} := \alpha_k - \delta$$

Note that all $\tilde{\alpha_k}$ satisfy the recursion

$$\tilde{\alpha_{k+1}} = \frac{1}{3} + \frac{\tilde{\alpha_k} - \delta}{2}.$$

It follows that

$$\tilde{\alpha_k} - \frac{1}{3} = \frac{\tilde{\alpha_{k-1}} - \delta}{2} < \frac{\tilde{\alpha_{k-1}}}{2} < \frac{\tilde{\alpha_k}}{2}. \tag{9}$$

Lemma 11. *For all naturals $K \geq 0$ and $\Delta > 0$ it holds the following. Let $\delta = \frac{\log_2(\Delta)}{n}$ and $\tilde{\alpha}_k = \alpha_k - \delta$ for all k, $0 \leq k \leq K$. Then*

$$Pr\left[\sum_{i=1}^{2^{\tilde{\alpha}_K \cdot n}} D_i > 2^{1/3 \cdot n}\right] < e^{-(\Delta - (K/3)n)}.$$

We show first that the Recursion Lemma proves the Crucial Lemma.

Let $K = \min\{k, \alpha_k > \alpha\}$. Let $\Delta = (1 + (K/3)) \cdot n$ and fix some n_0 such that $\alpha_K - \alpha > \delta = (\log_2(n) + \log_2(1 + K/3))/n$ for $n \geq n_0$.

We obtain that $\tilde{\alpha}_K = \alpha_K - \delta \geq \alpha$. Consequently,

$$Pr\left[\sum_{i=1}^{2^{\alpha \cdot n}} D_i > 2^{1/3 \cdot n}\right] \leq Pr\left[\sum_{i=1}^{2^{\tilde{\alpha}_K \cdot n}} D_i > 2^{1/3 \cdot n}\right] < e^{-n}. \square$$

We prove the Recursion Lemma by showing the following more general statement per induction over k, $0 \leq k \leq K$:

$$Pr\left[\sum_{i=j+1}^{j+2^{\tilde{\alpha}_k \cdot n}} D_i > 2^{1/3 \cdot n} \left| \sum_{i=1}^{j} D_i \leq (K - k) \cdot 2^{1/3 \cdot n}\right.\right] < e^{-(\Delta - (k/3)n)}.$$

For $k = 0$ the claim is obviously true as $\tilde{\alpha_0} < 1/3$. Thus

$$\sum_{i=j+1}^{j+2^{\tilde{\alpha}_0 \cdot n}} D_i \leq 2^{(1/3)\cdot n}$$

is true with probability one.

We fix now some k, $0 \leq k \leq K-1$, suppose that the more general statement is true for this k and show the more general statement for $k+1$. We repeat our argumentation above for $N = 2^{\tilde{\alpha}_k \cdot n}$. Note that

$$\begin{aligned}
(2^{-1/3\cdot n} + \sqrt{\mathit{\Delta}/(2N)})N &= 2^{(\tilde{\alpha}_k - 1/3)\cdot n} + 2^{(\tilde{\alpha}_k/2)\cdot n} \cdot \sqrt{\mathit{\Delta}/2} \\
&\leq \; 2^{(\tilde{\alpha}_k/2)\cdot n} + 2^{(\tilde{\alpha}_k/2)\cdot n} \cdot \frac{1}{2} \cdot \sqrt{\mathit{\Delta}} \qquad (cf.(9)) \\
&\leq \sqrt{\mathit{\Delta}} \cdot 2^{(\tilde{\alpha}_k/2)\cdot n} = 2^{\frac{1}{2}\log_2(\mathit{\Delta}) + (\tilde{\alpha}_k/2)\cdot n} \\
&= 2^{\frac{\tilde{\alpha}_k + \delta}{2}\cdot n}
\end{aligned}$$

We obtain, as argued above, by the induction hypothesis, that

$$\begin{aligned}
&Pr\left[\sum_{i=j+1}^{j+2^{\tilde{\alpha}_k \cdot n}} D_i > 2^{\frac{\tilde{\alpha}_k+\delta}{2}\cdot n} \,\middle|\, \sum_{i=1}^{j} D_i \leq (K-k)\cdot 2^{(1/3)\cdot n} \right] \\
&\qquad < e^{-\mathit{\Delta}} + e^{-(\mathit{\Delta} - (k/3)\cdot n)}.
\end{aligned}$$

Now divide (as above) the interval $1, \cdots, 2^{\tilde{\alpha}_{k+1}\cdot n}$ into $2^{(\tilde{\alpha}_{k+1} - \tilde{\alpha}_k)\cdot n}$ consecutive intervals of length $2^{\tilde{\alpha}_k \cdot n}$.

Note that

$$\tilde{\alpha}_{k+1} - \tilde{\alpha}_k = \frac{1}{3} - \frac{\tilde{\alpha}_k - \delta}{2} - \delta = \frac{1}{3} - \frac{\tilde{\alpha}_k + \delta}{2}.$$

Consequently,

$$\begin{aligned}
&Pr\left[\sum_{i=j+1}^{j+2^{\tilde{\alpha}_{k+1}\cdot n}} D_i > 2^{\frac{1}{3}\cdot n} \,\middle|\, \sum_{i=1}^{j} D_i \leq (K-(k+1))\cdot 2^{\frac{1}{3}\cdot n} \right] \\
&\leq 2^{\left(\frac{1}{3} - \frac{\tilde{\alpha}_k+\delta}{2}\right)\cdot n} \cdot Pr\left[\sum_{i=j'+1}^{j'+2^{\tilde{\alpha}_k\cdot n}} D_i > 2^{\frac{1}{3}\cdot n} \,\middle|\, \sum_{i=1}^{j'} D_i \leq (K-k)\cdot 2^{\frac{1}{3}\cdot n} \right] \\
&\leq 2^{\left(\frac{1}{3} - \frac{\tilde{\alpha}_k+\delta}{2}\right)\cdot n} \left(e^{-\mathit{\Delta}} + e^{-(\mathit{\Delta} - (k/3)\cdot n)} \right) \\
&< e^{\frac{1}{3}\cdot n - 1} \left(e^{-\mathit{\Delta}} + e^{-(\mathit{\Delta} - (k/3)\cdot n)} \right) \leq e^{\frac{1}{3}\cdot n - 1} \cdot 2 \cdot e^{-(\mathit{\Delta} - (k/3)\cdot n)} \\
&< e^{\frac{1}{3}\cdot n} \cdot e^{-(\mathit{\Delta} - (k/3)\cdot n)} = e^{-(\mathit{\Delta} - ((k+1)/3)\cdot n)}.
\end{aligned}$$

□