Cascade Encryption Revisited

Peter Gaži[1,2] and Ueli Maurer[1]

[1] ETH Zürich, Switzerland
Department of Computer Science
{gazipete,maurer}@inf.ethz.ch
[2] Comenius University, Bratislava, Slovakia
Department of Computer Science
gazi@dcs.fmph.uniba.sk

Abstract. The security of cascade blockcipher encryption is an important and well-studied problem in theoretical cryptography with practical implications. It is well-known that double encryption improves the security only marginally, leaving triple encryption as the shortest reasonable cascade. In a recent paper, Bellare and Rogaway showed that in the ideal cipher model, triple encryption is significantly more secure than single and double encryption, stating the security of longer cascades as an open question.

In this paper, we propose a new lemma on the indistinguishability of systems extending Maurer's theory of random systems. In addition to being of independent interest, it allows us to compactly rephrase Bellare and Rogaway's proof strategy in this framework, thus making the argument more abstract and hence easy to follow. As a result, this allows us to address the security of longer cascades as well as some errors in their paper. Our result implies that for blockciphers with smaller key space than message space (e.g. DES), longer cascades improve the security of the encryption up to a certain limit. This partially answers the open question mentioned above.

Keywords: cascade encryption, ideal cipher model, random system, indistinguishability.

1 Introduction

The cascade encryption is a simple and practical construction used to enlarge the key space of a blockcipher without the need to switch to a new algorithm. Instead of applying the blockcipher only once, it is applied l times with l independently chosen keys. A prominent and widely used example of this construction is the Triple DES encryption [2, 13, 14].

Many results investigating the power of the cascade construction have been published. It is well-known that double encryption does not significantly improve the security over single encryption due to the meet-in-the-middle attack [7]. The marginal security gain achieved by double encryption was described in [1]. Even and Goldreich [8] show that a cascade of ciphers is at least as strong as the strongest of the ciphers against attacks that are restricted to operating on full blocks. In contrast, Maurer and Massey [11] show that for the most general attack model, where it is for example possible that an attacker might obtain only half the ciphertext block for a chosen message block, the cascade is only at least as strong as the *first* cipher of the cascade.

In a recent paper [4], Bellare and Rogaway have claimed a lower bound on the security of triple encryption in the ideal cipher model. Their bound implies that for a blockcipher with key length k and block length n, triple encryption is indistinguishable from a random permutation as long as the distinguisher is allowed to make not more than roughly $2^{k+\frac{1}{2}\min\{n,k\}}$ queries. This bound is significantly higher than the known upper bound on the security of single and double encryption, proving that triple encryption is the shortest cascade that provides a reasonable security improvement over single encryption. Since a longer cascade is at least as secure as a shorter one, their bound applies also to longer cascades. They formulate as an interesting open problem to determine whether the security improves with the length of the cascade also for lengths $l > 3$. However, the proof in [4] contains a few bugs, which we describe in the appendix of this paper. The first part of our contribution is to fix these errors and to reestablish the lower bound on the security of triple encryption up to a constant factor.

Second, we have rephrased the proof into the random systems framework introduced in [10]. Our goal here is to simplify the proof and express it on the most abstract level possible, thus making the main line of reasoning easy to follow and clearly separated from the two technical arguments required. To achieve this, we extend the random systems framework by a new lemma. This lemma is a generalization of both Lemma 7 from [10] and hence also of its special case for the game-playing scenario, the Fundamental lemma of game-playing. This was introduced in [4] and subsequently used as an important tool in the game-playing proofs (see for example [15, 3, 5]). We illustrate the use of this new lemma in our proof of the security of cascade encryption. Apart from the simplification, this also gives us an improvement of the result by a constant factor.

Finally, our reformulation makes it natural to consider also the security of longer cascades. The lower bound we prove improves with the length of the cascade for all blockciphers where $k < n$. With increasing cascade length, the bound approaches very roughly the value $2^{k+\min\{n/2,k\}}$ (the exact formula can be found in Theorem 1). The condition $k < n$ is satisfied for example for the DES blockcipher, where the length of the key is 56 bits and the length of one block is 64 bits. For these parameters, the result from [4] that we reestablish proves that the triple encryption is secure up to 2^{78} queries, but our result shows that a cascade of length 5 is secure up to 2^{83} queries. The larger the difference $n - k$, the more a longer cascade can help. This partially answers the open question from [4].

2 Preliminaries

2.1 Basic Notation

Throughout the paper, we denote sets by calligraphic letters (e.g. $\mathcal{S}$). For a finite set $\mathcal{S}$, we denote by $|\mathcal{S}|$ the number of its elements. A k-tuple is denoted as $u^k = (u_1, \ldots, u_k)$, and the set of all k-tuples of elements of $\mathcal{U}$ is denoted as $\mathcal{U}^k$. The composition of mappings is interpreted from left to right, i.e., $f \circ g$ denotes the mapping $g(f(\cdot))$. The set of all permutations of $\{0,1\}^n$ is denoted by $\mathrm{Perm}(n)$ and id represents the identity mapping, if the domain is implicitly given. The notation $x^{\underline{n}}$ represents the falling factorial power, i.e., $x^{\underline{n}} = x(x-1)\cdots(x-n+1)$. The symbol $p_{coll}(n,k)$ denotes the probability that k independent random variables with uniform distribution over a set of size n contain a collision, i.e., that they are not all distinct. It is well-known that $p_{coll}(n,k) < k^2/2n$. By $\mathsf{CS}(\cdot)$ we shall denote the set of all *cyclic shifts* of a given tuple, in other words, $\mathsf{CS}(\pi_1, \pi_2, \ldots, \pi_r) = \{(\pi_1, \pi_2, \ldots, \pi_r), (\pi_2, \pi_3, \ldots, \pi_r, \pi_1), \ldots, (\pi_r, \pi_1, \ldots, \pi_{r-1})\}$.

We usually denote random variables and concrete values they can take on by capital and small letters, respectively. For events A and B and random variables U and V with ranges $\mathcal{U}$ and $\mathcal{V}$, respectively, we denote by $\mathsf{P}_{UA|VB}$ the corresponding conditional probability distribution, seen as a function $\mathcal{U} \times \mathcal{V} \to \langle 0,1 \rangle$. Here the value $\mathsf{P}_{UA|VB}(u,v)$ is well-defined for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$ such that $\mathsf{P}_{VB}(v) > 0$ and undefined otherwise. Two probability distributions P_U and $\mathsf{P}_{U'}$ on the same set $\mathcal{U}$ are equal, denoted $\mathsf{P}_U = \mathsf{P}_{U'}$, if $\mathsf{P}_U(u) = \mathsf{P}_{U'}(u)$ for all $u \in \mathcal{U}$. Conditional probability distributions are equal if the equality holds for all arguments for which both of them are defined. To emphasize the random experiment $\mathcal{E}$ in consideration, we sometimes write it in the superscript, e.g. $\mathsf{P}^{\mathcal{E}}_{U|V}(u,v)$. The expected value of the random variable X is denoted by $\mathsf{E}[X] = \sum_{x \in \mathcal{X}}(x \cdot \mathsf{P}[X = x])$. The complement of an event A is denoted by $\overline{A}$.

2.2 Random Systems

In this subsection, we present the basic notions of the random systems framework, as introduced in [10], along with some new extensions of the framework. The input-output behavior of any discrete system can be described by a *random system* in the spirit of the following definition.

Definition 1. *An $(\mathcal{X}, \mathcal{Y})$-random system $\mathbf{F}$ is a (generally infinite) sequence of conditional probability distributions $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}}$ for all $i \geq 1$.*

The behavior of the random system is specified by the sequence of conditional probabilities $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1})$ (for $i \geq 1$) of obtaining the output $y_i \in \mathcal{Y}$ on query $x_i \in \mathcal{X}$ given the previous $i-1$ queries $x^{i-1} = (x_1, \ldots, x_{i-1}) \in \mathcal{X}^{i-1}$ and their corresponding outputs $y^{i-1} = (y_1, \ldots, y_{i-1}) \in \mathcal{Y}^{i-1}$. A random system can also be defined by a sequence of conditional probability distributions $\mathsf{P}^{\mathbf{F}}_{Y^i|X^i}$ for $i \geq 1$. This description is often convenient, but is not minimal.

We shall use boldface letters (e.g. $\mathbf{F}$) to denote both a discrete system and a random system corresponding to it. This should cause no confusion. We emphasize that although the results of this paper are stated for random systems, they hold for arbitrary systems, since the only property of a system that is relevant here is its input-output behavior. It is reasonable to consider two discrete systems equivalent if their input-output behaviors are the same, even if their internal structure differs.

Definition 2. *Two systems* $\mathbf{F}$ *and* $\mathbf{G}$ *are* equivalent*, denoted* $\mathbf{F} \equiv \mathbf{G}$*, if they correspond to the same random system, i.e., if* $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}} = \mathsf{P}^{\mathbf{G}}_{Y_i|X^iY^{i-1}}$ *for all* $i \geq 1$*.*

We shall usually define a system (and hence also the corresponding random system) by a description of its internal working, as long as the transition to the probability distributions is straightforward. Examples of random systems that we consider in the following are a *uniform random permutation* $\mathbf{P} : \{0,1\}^n \to \{0,1\}^n$, which realizes a function randomly chosen from $\mathrm{Perm}(n)$; and an *ideal blockcipher* $\mathbf{E} : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$, which realizes an independent random permutation for each key $K \in \{0,1\}^k$. In this paper we assume that both $\mathbf{P}$ and $\mathbf{E}$ can be queried in both directions.

We can define a *distinguisher* $\mathbf{D}$ for an $(\mathcal{X}, \mathcal{Y})$-random system as a $(\mathcal{Y}, \mathcal{X})$-random system which is one query ahead, i.e., it is defined by the conditional probability distributions $\mathsf{P}^{\mathbf{D}}_{X_i|X^{i-1}Y^{i-1}}$ for all $i \geq 1$. In particular, the first query of $\mathbf{D}$ is determined by $\mathsf{P}^{\mathbf{D}}_{X_1}$. After a certain number of queries (say q), the distinguisher outputs a bit W_q depending on the transcript (X^q, Y^q). For a random system $\mathbf{F}$ and a distinguisher $\mathbf{D}$, let $\mathbf{DF}$ be the random experiment where $\mathbf{D}$ interacts with $\mathbf{F}$. Then for two $(\mathcal{X}, \mathcal{Y})$-random systems $\mathbf{F}$ and $\mathbf{G}$, the *distinguishing advantage* of $\mathbf{D}$ in distinguishing systems $\mathbf{F}$ and $\mathbf{G}$ by q queries is defined as $\Delta^{\mathbf{D}}_q(\mathbf{F}, \mathbf{G}) = \left|\mathsf{P}^{\mathbf{DF}}(W_q = 1) - \mathsf{P}^{\mathbf{DG}}(W_q = 1)\right|$. We are usually interested in the maximal distinguishing advantage over all such distinguishers, which we denote by $\Delta_q(\mathbf{F}, \mathbf{G}) = \max_{\mathbf{D}} \Delta^{\mathbf{D}}_q(\mathbf{F}, \mathbf{G})$.

For a random system $\mathbf{F}$, we often consider an internal *monotone condition* defined on it. Such a condition is initially satisfied (true), but once it gets violated, it cannot become true again. We characterize such a condition by a sequence of events $\mathcal{A} = A_0, A_1, \ldots$ such that A_0 always holds, and A_i holds if the condition holds after query i. The probability that a distinguisher $\mathbf{D}$ issuing q queries makes a monotone condition $\mathcal{A}$ fail in the random experiment $\mathbf{DF}$ is denoted by $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q}) = \mathsf{P}^{\mathbf{DF}}(\overline{A_q})$ and we are again interested in the maximum over all distinguishers, denoted by $\nu(\mathbf{F}, \overline{A_q}) = \max_{\mathbf{D}} \nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$. For a random system $\mathbf{F}$ with a monotone condition $\mathcal{A} = A_0, A_1, \ldots$ and a random system $\mathbf{G}$, we say that $\mathbf{F}$ *conditioned on* $\mathcal{A}$ *is equivalent to* $\mathbf{G}$, denoted $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}$, if $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}A_i} = \mathsf{P}^{\mathbf{G}}_{Y_i|X^iY^{i-1}}$ for $i \geq 1$, for all arguments for which $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}A_i}$ is defined. The following claim was proved in [10].

Lemma 1. *If* $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}$ *then* $\Delta_q(\mathbf{F}, \mathbf{G}) \leq \nu(\mathbf{F}, \overline{A_q})$*.*

Let $\mathbf{F}$ be a random system with a monotone condition $\mathcal{A}$. Following [12], we define $\mathbf{F}$ *blocked by* $\mathcal{A}$ to be a new random system that behaves exactly like $\mathbf{F}$ while the condition $\mathcal{A}$ is satisfied. Once $\mathcal{A}$ is violated, it only outputs a special blocking symbol $\perp$ not contained in the output alphabet of $\mathbf{F}$. More formally, the following mapping is applied to the i^{th} output of $\mathbf{F}$:

$$y_i \mapsto \begin{cases} y_i & \text{if } A_i \text{ holds} \\ \perp & \text{otherwise.} \end{cases}$$

The following new lemma relates the optimal advantage in distinguishing two random systems to the optimal advantage in distinguishing their blocked counterparts.

Lemma 2. *Let* $\mathbf{F}$ *and* $\mathbf{G}$ *be two random systems with monotone conditions* $\mathcal{A}$ *and* $\mathcal{B}$ *defined on them, respectively. Let* $\mathbf{F}^\perp$ *denote the random system* $\mathbf{F}$ *blocked by* $\mathcal{A}$ *and let* $\mathbf{G}^\perp$ *denote* $\mathbf{G}$ *blocked by* $\mathcal{B}$. *Then for every distinguisher* $\mathbf{D}$ *we have* $\Delta_q^{\mathbf{D}}(\mathbf{F},\mathbf{G}) \leq \Delta_q(\mathbf{F}^\perp,\mathbf{G}^\perp) + \nu^{\mathbf{D}}(\mathbf{F},\overline{A_q})$.

Proof. Let $\mathbf{D}$ be an arbitrary distinguisher for $\mathbf{F}$ and $\mathbf{G}$. Let $\mathbf{D}'$ be a distinguisher that works as follows: it simulates $\mathbf{D}$, but whenever it receives an answer $\perp$ to its query, it aborts and outputs 1. Then we have $\mathsf{P}^{\mathbf{DG}}[W_q = 1] \leq \mathsf{P}^{\mathbf{D}'\mathbf{G}^\perp}[W_q = 1]$ and $\mathsf{P}^{\mathbf{D}'\mathbf{F}^\perp}[W_q = 1] \leq \mathsf{P}^{\mathbf{DF}}[W_q = 1] + \nu^{\mathbf{D}}(\mathbf{F},\overline{A_q})$.

First, let us assume that $\mathsf{P}^{\mathbf{DG}}[W_q = 1] \geq \mathsf{P}^{\mathbf{DF}}[W_q = 1]$. Then, using the definition of advantage and the above inequalities, we get

$$\begin{aligned}
\Delta_q^{\mathbf{D}}(\mathbf{F},\mathbf{G}) &= \left|\mathsf{P}^{\mathbf{DG}}[W_q = 1] - \mathsf{P}^{\mathbf{DF}}[W_q = 1]\right| \\
&= \mathsf{P}^{\mathbf{DG}}[W_q = 1] - \mathsf{P}^{\mathbf{DF}}[W_q = 1] \\
&\leq \mathsf{P}^{\mathbf{D}'\mathbf{G}^\perp}[W_q = 1] - (\mathsf{P}^{\mathbf{D}'\mathbf{F}^\perp}[W_q = 1] - \nu^{\mathbf{D}}(\mathbf{F},\overline{A_q})) \\
&\leq \Delta_q(\mathbf{F}^\perp,\mathbf{G}^\perp) + \nu^{\mathbf{D}}(\mathbf{F},\overline{A_q}),
\end{aligned}$$

which proves the lemma in this case. On the other hand, if $\mathsf{P}^{\mathbf{DG}}[W_q = 1] < \mathsf{P}^{\mathbf{DF}}[W_q = 1]$, we can easily construct another distinguisher $\mathbf{D}^*$ with the same behavior as $\mathbf{D}$ and the opposite final answer bit. Then we can proceed with the argument as before and since $\Delta_q^{\mathbf{D}}(\mathbf{F},\mathbf{G}) = \Delta_q^{\mathbf{D}^*}(\mathbf{F},\mathbf{G})$ and $\nu^{\mathbf{D}}(\mathbf{F},\overline{A_q}) = \nu^{\mathbf{D}^*}(\mathbf{F},\overline{A_q})$, the conclusion is valid also for the distinguisher $\mathbf{D}$. □

Lemma 2 is a generalization of both Lemma 7 from [10] and of its special case, the Fundamental lemma of game-playing from [4]. Both these lemmas describe the special case when $\Delta_q(\mathbf{F}^\perp,\mathbf{G}^\perp) = 0$, i.e., when the distinguished systems behave identically until some conditions are violated. Our lemma is useful in the situations where the systems are not identical even while the conditions are satisfied, but their behavior is very similar. A good example of such a situation is presented in the proof of Theorem 1.

A random system $\mathbf{F}$ can be used as a component of a larger system: in particular, we shall consider *constructions* $\mathbf{C}(\cdot)$ such that the resulting random system $\mathbf{C}(\mathbf{F})$ invokes $\mathbf{F}$ as a subsystem. We state the following two observations about the composition of systems.

Lemma 3. *Let* $\mathbf{C}(\cdot)$ *and* $\mathbf{C}'(\cdot)$ *be two constructions invoking an internal random system, and let* $\mathbf{F}$ *and* $\mathbf{G}$ *be random systems. Then*

(i) $\Delta_q(\mathbf{C}(\mathbf{F}), \mathbf{C}(\mathbf{G})) \le \Delta_{q'}(\mathbf{F}, \mathbf{G})$*, where* q' *is the maximum number of invocations of any internal system* $\mathbf{H}$ *for any sequence of* q *queries to* $\mathbf{C}(\mathbf{H})$*, if such a value is defined.*

(ii) There exists a fixed permutation $S \in \mathrm{Perm}(n)$ *(represented by a deterministic stateless system) such that* $\Delta_q(\mathbf{C}(\mathbf{P}), \mathbf{C}'(\mathbf{P})) \le \Delta_q(\mathbf{C}(S), \mathbf{C}'(S))$*.*

Proof. The first claim comes from [10], so here we only prove the second one. Since the random system $\mathbf{P}$ can be seen as a system that picks a permutation uniformly at random from $\mathrm{Perm}(n)$ and then realizes this permutation, we have:

$$\Delta_q(\mathbf{C}(\mathbf{P}), \mathbf{C}'(\mathbf{P})) \le \frac{1}{(2^n)!} \sum_{S \in \mathrm{Perm}(n)} \Delta_q(\mathbf{C}(S), \mathbf{C}'(S)).$$

If all the values $\Delta_q(\mathbf{C}(S), \mathbf{C}'(S))$ were smaller than $\Delta_q(\mathbf{C}(\mathbf{P}), \mathbf{C}'(\mathbf{P}))$ it would contradict the inequality above, hence there exists a permutation $S \in \mathrm{Perm}(n)$ such that $\Delta_q(\mathbf{C}(\mathbf{P}), \mathbf{C}'(\mathbf{P})) \le \Delta_q(\mathbf{C}(S), \mathbf{C}'(S))$. □

2.3 Ideal Blockciphers and Chains

We introduce some specific notions related to the cascade encryption setting. Our terminology follows and extends that in [4].

A blockcipher with keyspace $\{0,1\}^k$ and message space $\{0,1\}^n$ is a mapping $E\colon \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ such that for each $K \in \{0,1\}^k$, $E(K, \cdot)$ is a permutation on the set $\{0,1\}^n$. Typically $E_K(x)$ is written instead of $E(K, x)$ and $E_K^{-1}(\cdot)$ refers to the inverse of the permutation $E_K(\cdot)$.

Throughout the paper, we shall work in the ideal blockcipher model, which was recently shown to be equivalent to the random oracle model [6]. The ideal blockcipher model is widely used to analyze blockcipher constructions (e.g. [1, 4, 9]) and consists of the assumption that for each key, the blockcipher realizes an independent random permutation.

A blockcipher can be seen as a directed graph consisting of 2^n vertices representing the message space and 2^{n+k} edges. Each vertex x has 2^k outgoing edges pointing to the encryptions of the message x using all possible keys. Each of the edges is labeled by the respective key. For a fixed blockcipher E, we denote by[3]

$$w(E) = \max_{x,y} |\{K \mid E_K(x) = y\}|$$

the maximal number of distinct keys mapping the plaintext x onto the ciphertext y, the maximum taken over all pairs of blocks (x, y). Intuitively, $w(E)$ is the weight of the heaviest edge in the graph corresponding to E. This also naturally defines a random variable $w(\mathbf{E})$ for the random system $\mathbf{E}$ realizing an ideal blockcipher.

[3] $w(E)$ was denoted as Keys^E in [4]

If a distinguisher makes queries to a blockcipher E, let $x \overset{K}{\to} y$ denote the fact that it either made a query $E_K(x)$ and received the encryption y or made a query $E_K^{-1}(y)$ and received the decryption x. An *r-chain for keys* $(K_1, \ldots, K_r)$ is an $(r+1)$-tuple $(x_0, K_1, \ldots, K_r)$ for which there exist $x_1, \ldots, x_r$ such that $x_0 \overset{K_1}{\to} x_1 \overset{K_2}{\to} \cdots \overset{K_r}{\to} x_r$ holds. Similarly, if a fixed permutation S is given and $1 \leq i < r$, then an *i-disconnected r-chain for keys* $(K_1, \ldots, K_r)$ is an $(r+1)$-tuple $(x_0, K_1, \ldots, K_r)$ for which there exist $x_1, \ldots, x_r$ such that we have both $x_0 \overset{K_{r-i+1}}{\longrightarrow} x_1 \overset{K_{r-i+2}}{\longrightarrow} \cdots \overset{K_r}{\to} x_i$ and $S^{-1}(x_i) \overset{K_1}{\to} x_{i+1} \overset{K_2}{\to} \cdots \overset{K_{r-i}}{\longrightarrow} x_r$. When describing chains, we sometimes explicitly refer to the permutations instead of the keys that define them. The purpose of the following definition will be clear from the proof of Theorem 1.

Definition 3. *A distinguisher* examines *the key tuple* $(K_1, K_2, \ldots, K_r)$ *if it creates either an r-chain or an i-disconnected r-chain for* $(K_1, K_2, \ldots, K_r)$ *for any* $i \in \{1, \ldots, r-1\}$.

3 The Security of Cascade Encryption

In this section we reestablish the lower bound on the security of triple encryption from [4] in a more general setting. Our goal here is to simplify the proof and make it more comprehensible thanks to the level of abstraction provided by the random systems framework. Using Lemma 2 we also gain an improvement by a constant factor of 2. However, in order to fix the problem of the proof in [4], a new factor l appears in the security bound.

Although Theorem 1 only explicitly states the security of cascades with odd length, we point out that a simple reduction argument proves that longer cascades cannot be less secure than shorter ones, except for a negligible term $l/2^k$. Therefore, our result also implicitly proves any even cascade to be at least as secure as a one step shorter odd-length cascade.

3.1 Proof of the Main Result

Since this subsection aims to address the overall structure of the proof, we shall use two technical lemmas without proof (Lemmas 4 and 5). These lemmas correspond to Lemmas 7 and 9 from [4], which they improve and generalize. We shall prove them in later subsections.

Let $l \geq 3$ be an odd integer. Let $\mathbf{C}_1(\cdot, \cdot)$ denote a construction which expects two subsystems: a blockcipher E and a permutation P. It chooses in advance l random distinct keys $K_1, \ldots, K_l$. These are not used by the system, their purpose is to make $\mathbf{C}_1(\cdot, \cdot)$ comparable to the other constructions. $\mathbf{C}_1(\cdot, \cdot)$ provides an interface to make forward and backward queries both to the blockcipher E and to the permutation P.

On the other hand, let $\mathbf{C}_2(\cdot)$ denote a construction which expects a blockcipher E as the only subsystem. It chooses in advance l random keys $K_1, \ldots, K_l$.

It provides an interface to make forward and backward queries both to the blockcipher E and to a permutation P, which it realizes as $E_{K_1} \circ \cdots \circ E_{K_l}$. To achieve this, $\mathbf{C}_2(\cdot)$ queries its subsystem for all necessary values. Let $\mathbf{C}_2^{\mathrm{d}}(\cdot)$ be the same construction as $\mathbf{C}_2(\cdot)$ except that it chooses the keys $K_1, \ldots, K_l$ to be distinct.

Finally, let $\mathbf{C}_3(\cdot, \cdot)$ denote a construction which again expects two subsystems: a blockcipher E and a permutation P. It chooses in advance l random distinct keys $K_1, \ldots, K_l$. It provides an interface to make forward and backward queries both to the blockcipher E and to the permutation P. However, answers to the blockcipher queries involving the key K_l are modified to satisfy the equation $E_{K_1} \circ \cdots \circ E_{K_l} = P$, for example forward queries are realized as $E_{K_l}(x) = P(E_{K_1}^{-1}(\cdots E_{K_{l-1}}^{-1}(x) \cdots))$. To achieve this, $\mathbf{C}_3(\cdot, \cdot)$ queries its subsystems for all necessary values.

The following theorem bounds $\Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_2(\mathbf{E}))$, the advantage in distinguishing cascade encryption of length l from a random permutation, given access to the underlying blockcipher.

Theorem 1. *For the constructions* $\mathbf{C}_1(\cdot, \cdot)$, $\mathbf{C}_2(\cdot)$ *and random systems* $\mathbf{E}$, $\mathbf{P}$ *defined as above we have*

$$\Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_2(\mathbf{E})) \leq 2l\alpha^{\lfloor l/2 \rfloor} \frac{q^{\lceil l/2 \rceil}}{(2^k)^{\underline{l}}} + 1.9 \left(\frac{lq}{2^{k+n/2}} \right)^{2/3} + \frac{l^2}{2^{k+1}},$$

where $\alpha = \max\{2e2^{k-n}, 2n + k\lfloor l/2 \rfloor\}$.

Proof. First, it is easy to see that $\Delta_q(\mathbf{C}_2(\mathbf{E}), \mathbf{C}_2^{\mathrm{d}}(\mathbf{E})) \leq p_{coll}(2^k, l) < l^2/2^{k+1}$ and hence we have $\Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_2(\mathbf{E})) \leq \Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_2^{\mathrm{d}}(\mathbf{E})) + l^2/2^{k+1}$. However, note that $\mathbf{C}_2^{\mathrm{d}}(\mathbf{E}) \equiv \mathbf{C}_3(\mathbf{E}, \mathbf{P})$; in both systems the permutations $E_{K_1}, \ldots, E_{K_l}, P$ are chosen randomly with the only restriction that $E_{K_1} \circ \cdots \circ E_{K_l} = P$ is satisfied. Now we can use Lemma 3 to substitute the random permutation $\mathbf{P}$ in both $\mathbf{C}_1(\mathbf{E}, \mathbf{P})$ and $\mathbf{C}_3(\mathbf{E}, \mathbf{P})$ for a fixed one. Let S denote the permutation guaranteed by Lemma 3. Then we have

$$\Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_2^{\mathrm{d}}(\mathbf{E})) = \Delta_q(\mathbf{C}_1(\mathbf{E}, \mathbf{P}), \mathbf{C}_3(\mathbf{E}, \mathbf{P})) \leq \Delta_q(\mathbf{C}_1(\mathbf{E}, S), \mathbf{C}_3(\mathbf{E}, S)).$$

Since the permutation S is fixed, it makes now no sense for the distinguisher to query this permutation; it can have the permutation S hardwired.

From now on, we shall denote all queries to a blockcipher that involve one of the keys K_1, K_2, $\ldots$, K_l as *relevant queries*. Let us now consider a monotone condition $\mathcal{A}^h$ ($h \in \mathbb{N}$ is a parameter) defined on the random system $\mathbf{C}_1(\mathbf{E}, S)$. The condition A_q^h is satisfied if the keys $(K_1, K_2, \ldots, K_l)$ were not examined (in the sense of Definition 3) by the first q queries and at most h of these q queries were relevant. Let $\mathcal{B}^h$ be an analogous condition defined on $\mathbf{C}_3(\mathbf{E}, S)$: B_q^h is satisfied if the first q queries did not form a chain for the tuple $(K_1, K_2, \ldots, K_l)$ and at most h of these queries were relevant. Let $\mathbf{G}$ and $\mathbf{H}$ denote the random systems $\mathbf{C}_1(\mathbf{E}, S)$ and $\mathbf{C}_3(\mathbf{E}, S)$ blocked by $\mathcal{A}^h$ and $\mathcal{B}^h$, respectively. Then by Lemma 2,

$$\Delta_q(\mathbf{C}_1(\mathbf{E}, S), \mathbf{C}_3(\mathbf{E}, S)) \leq \Delta_q(\mathbf{G}, \mathbf{H}) + \nu(\mathbf{C}_1(\mathbf{E}, S), \overline{A_q^h}).$$

Let us first bound the quantity $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{A_q^h})$. We can write A_q^h as $U_q \wedge V_q^h$, where U_q is satisfied if the first q queries did not examine the tuple of keys $(K_1, K_2, \ldots, K_l)$ and V_q^h is satisfied if at most h of the first q queries were relevant. Since $\overline{A_q^h} \Leftrightarrow \overline{U_q} \vee \overline{V_q^h}$, the union bound gives us

$$\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{A_q^h}) \leq \nu(\mathbf{C}_1(\mathbf{E}, S), \overline{U_q}) + \nu(\mathbf{C}_1(\mathbf{E}, S), \overline{V_q^h}).$$

We prove in Lemma 4 that $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{U_q}) \leq 2l\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil} / (2^k)^{\underline{l}}$. Since the keys $K_1, \ldots, K_l$ do not affect the outputs of $\mathbf{C}_1(\mathbf{E}, S)$, adaptivity does not help when trying to violate the condition V_q^h, therefore we can restrict our analysis to nonadaptive strategies for provoking $\overline{V_q^h}$. The probability that a given query is relevant is $l/2^k$, hence the expected number of relevant queries among the first q queries is $lq/2^k$ and by Markov's inequality we have $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{V_q^h}) \leq lq/h2^k$. All put together, $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{A_q^h}) \leq 2l\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil} / (2^k)^{\underline{l}} + lq/h2^k$.

It remains to bound $\Delta_q(\mathbf{G}, \mathbf{H})$. These systems only differ in their behavior for the first h relevant queries, so let us make this difference explicit. Let $\mathbf{G}_{\mathrm{r}}$ be a random system that allows queries to l independent random permutations $\pi_1, \pi_2, \ldots, \pi_l$, but returns $\perp$ once the queries create an l-chain for any tuple in $\mathsf{CS}(\pi_1, \pi_2, \ldots, \pi_l)$. Let $\mathbf{H}_{\mathrm{r}}$ be a random system that allows queries to l random permutations $\pi_1, \pi_2, \ldots, \pi_l$ such that $\pi_1 \circ \pi_2 \circ \ldots \circ \pi_l = id$, but returns $\perp$ once the queries create an l-chain for the tuple $(\pi_1, \pi_2, \ldots, \pi_l)$. Let $\mathbf{C}_{h,S}(\cdot)$ be a construction that allows queries to a blockcipher, let us denote it by E. In advance, it picks l random distinct keys K_1,K_2,…,K_l. Then it realizes the queries to E_{K_1},E_{K_2},…,E_{K_l} as π_1, π_2,…,π_{l-1} and $\pi_l \circ S$ respectively, where the permutations π_i for $i \in \{1, \ldots, l\}$ are provided by a subsystem. E_K for all other keys K are realized by $\mathbf{C}_{h,S}(\cdot)$ as random permutations. However, $\mathbf{C}_{h,S}(\cdot)$ only redirects the first h relevant queries to the subsystem, after this number is exceeded, it responds to all queries by $\perp$. Intuitively, the subsystem used is responsible for the answers to the first h relevant queries (hence the subscript "r"). Since the disconnected chains in $\mathbf{C}_{h,S}(\mathbf{G}_{\mathrm{r}})$ correspond exactly to the ordinary chains in $\mathbf{G}_{\mathrm{r}}$, we have $\mathbf{C}_{h,S}(\mathbf{G}_{\mathrm{r}}) \equiv \mathbf{G}$ and $\mathbf{C}_{h,S}(\mathbf{H}_{\mathrm{r}}) \equiv \mathbf{H}$. According to Lemma 3 and Lemma 5 below, we have $\Delta_q(\mathbf{G}, \mathbf{H}) \leq \Delta_h(\mathbf{G}_{\mathrm{r}}, \mathbf{H}_{\mathrm{r}}) \leq h^2/2^n$.

Now we can optimize the choice of the constant h. The part of the advantage that depends on h is $f(h) = lq/h2^k + h^2/2^n$. This term is minimal for $h^* = (lq2^{n-k-1})^{1/3}$ and we get $f(h^*) < 1.9 \left(\frac{lq}{2^{k+n/2}} \right)^{2/3}$. This completes the proof. □

3.2 Examining the Relevant Keys

Here we analyze the probability that the adversary examines the relevant keys $(K_1, \ldots, K_l)$ during its interaction with the random system $\mathbf{C}_1(\mathbf{E}, S)$. This is a generalization of Lemma 7 from [4] to longer cascades, also taking disconnected chains into account.

Lemma 4. *Let the random system* $\mathbf{C}_1(\mathbf{E}, S)$ *and the condition* U_q *be defined as in the proof of Theorem 1, with the number of keys* l *being odd. Then we have* $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{U_q}) \leq 2l\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil}/(2^k)^{\underline{l}}$, *where* $\alpha = \max\{2e2^{k-n}, 2n + k\lfloor l/2 \rfloor\}$.

Proof. The relevant keys $K_1, \ldots, K_l$ are examined by the distinguisher if it creates either an l-chain or an i-disconnected l-chain for the tuple $(K_1, K_2, \ldots, K_l)$ for any $i \in \{1, \ldots, l-1\}$.

Let $i \in \{1, \ldots, l-1\}$ be fixed. We first bound the probability that the distinguisher creates an i-disconnected l-chain. Since the relevant keys do not affect the behavior of the system $\mathbf{C}_1(\mathbf{E}, S)$, this probability is equal to the number of l-tuples of distinct keys for which an i-disconnected l-chain was created, divided by the number of all l-tuples of distinct keys, which is $(2^k)^{\underline{l}}$. The numerator can be upper bounded by the number of all i-disconnected l-chains that were created (here we also count those created for non-distinct key tuples). Hence, let $\mathsf{Ch}^E_{i,l,q}$ denote the maximum number of i-disconnected l-chains any distinguisher can create by issuing q queries to a fixed blockcipher E and let $\mathsf{Ch}^{\mathbf{E}}_{i,l,q}$ denote the expected value of $\mathsf{Ch}^E_{i,l,q}$ with respect to the choice of E by $\mathbf{E}$.

Let G be a directed graph corresponding to a blockcipher E, as described in Subsection 2.3. Let H be the spanning subgraph of G containing only the edges that were queried by the distinguisher. Any i-disconnected l-chain consists of l edges in H, let us denote them as $e_1, e_2, \ldots, e_l$, following the order in which they appear in the chain. Then for each of the odd edges $e_1, e_3, \ldots, e_l$ there are q possibilities to choose which of the queries corresponds to this edge. Once the odd edges are fixed, they uniquely determine the vertices $x_0, x_1, \ldots, x_l$ such that e_j is $x_{j-1} \rightarrow x_j$ for $j \in \{1, 3, \ldots, l\} \setminus \{i+1\}$ and e_{i+1} is $S^{-1}(x_i) \rightarrow x_{i+1}$ if i is even. Since there are at most $w(E)$ possible edges to connect any pair of vertices in G, there are now at most $w(E)$ possibilities to choose each of the even edges $e_2, e_4, \ldots, e_{l-1}$ so that e_j is $x_{j-1} \rightarrow x_j$ for $j \in \{2, 4, \ldots, l-1\} \setminus \{i+1\}$ and e_{i+1} is $S^{-1}(x_i) \rightarrow x_{i+1}$ if i is odd. Hence, $\mathsf{Ch}^E_{i,l,q} \leq w(E)^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil}$ and $\mathsf{Ch}^{\mathbf{E}}_{i,l,q} \leq w(\mathbf{E})^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil}$.

It remains to bound the value $w(\mathbf{E})$. For this, we use the bound from [4], where the inequality $\mathsf{P}[w(\mathbf{E}) \geq \beta] < 2^{2n+1-\beta}$ is proved for any $\beta \geq 2e2^{k-n}$. Using this inequality gives us

$$\begin{aligned}\mathsf{Ch}^{\mathbf{E}}_{i,l,q} &\leq \mathsf{E}[\mathsf{Ch}^E_{i,l,q} \mid w(E) < \alpha] + \mathsf{E}[\mathsf{Ch}^E_{i,l,q} \mid w(E) \geq \alpha] \cdot 2^{2n+1-\alpha} \\ &\leq \alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil} + 2^{k\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil} 2^{2n+1-\alpha} \leq 2\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil},\end{aligned}$$

where the last two inequalities hold since $w(E) \leq 2^k$ and $\alpha \geq 2n + k\lfloor l/2 \rfloor \geq 2$.

Putting all together, we get that the probability of forming an i-disconnected l-chain for the keys $(K_1, K_2, \ldots, K_l)$ can be upper bounded by $2\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil}/(2^k)^{\underline{l}}$. Since this holds for each $i \in \{1, 2, \ldots, l-1\}$ and the probability of creating an l-chain for the keys $(K_1, \ldots, K_l)$ can be bounded in the same way, by the union bound we get $\nu(\mathbf{C}_1(\mathbf{E}, S), \overline{U_q}) \leq 2l\alpha^{\lfloor l/2 \rfloor} q^{\lceil l/2 \rceil}/(2^k)^{\underline{l}}$. □

3.3 Distinguishing Independent and Correlated Permutations

Now we shall improve the bound on $\Delta_h(\mathbf{G}_\mathrm{r}, \mathbf{H}_\mathrm{r})$ stated by Lemma 9 in [4]. Using the concept of conditional equivalence from [10], our result is better by a constant factor and is applicable for the general case of l-cascade encryption.

Recall that $\mathbf{G}_\mathrm{r}$ is a random system that provides an interface to query l random independent permutations[4] $\pi_1, \ldots, \pi_l$ in both directions. However, if the queries of the distinguisher form an l-chain for any tuple of permutations in $\mathsf{CS}(\pi_1, \ldots, \pi_l)$, the system $\mathbf{G}_\mathrm{r}$ becomes blocked and answers all subsequent queries (including the one that formed the chain) with the symbol $\perp$. On the other hand, $\mathbf{H}_\mathrm{r}$ is a random system that provides an interface to query l random permutations $\pi_1, \ldots, \pi_l$ such that $\pi_1 \circ \cdots \circ \pi_l = id$, again in both directions. Similarly, if an l-chain is created for any tuple in $\mathsf{CS}(\pi_1, \ldots, \pi_l)$ (which is in this case equivalent to creating an l-chain for $(\pi_1, \ldots, \pi_l)$), $\mathbf{H}_\mathrm{r}$ answers all subsequent queries with the symbol $\perp$. Therefore, the value $\Delta_h(\mathbf{G}_\mathrm{r}, \mathbf{H}_\mathrm{r})$ denotes the best possible advantage in distinguishing l independent random permutations from l random permutations correlated in the described way, without forming an l-chain.

Lemma 5. *Let $\mathbf{G}_\mathrm{r}$ and $\mathbf{H}_\mathrm{r}$ be the random systems defined in the proof of Theorem 1. Then $\Delta_h(\mathbf{G}_\mathrm{r}, \mathbf{H}_\mathrm{r}) \leq h^2/2^n$.*

Proof. First, let us introduce some notation. In any experiment where the permutations $\pi_1, \ldots, \pi_l$ are queried, let $\mathrm{dom}_j(\pi_i)$ denote the set of all $x \in \{0,1\}^n$ such that among the first j queries, the query $\pi_i(x)$ was already answered or some query $\pi_i^{-1}(y)$ was answered by x. Similarly, let $\mathrm{range}_j(\pi_i)$ be the set of all $y \in \{0,1\}^n$ such that among the first j queries, the query $\pi_i^{-1}(y)$ was already answered or some query $\pi_i(x)$ was answered by y. In other words, $\mathrm{dom}_j(\pi_i)$ and $\mathrm{range}_j(\pi_i)$ denote the domain and range of the partial function π_i defined by the first j answers. For each pair of consecutive permutations[5] π_i and π_{i+1}, let $\mathcal{X}_i^{(j)}$ denote the set $\{0,1\}^n \setminus (\mathrm{range}_j(\pi_i) \cup \mathrm{dom}_j(\pi_{i+1}))$ of fresh, unused values. If $x \xrightarrow{\pi_i} y$ then we call the queries $\pi_i(x)$ and $\pi_i^{-1}(y)$ *trivial* and the queries $\pi_{i+1}(y)$ and $\pi_{i-1}^{-1}(x)$ are said to *extend* a chain if they are not trivial too.

Now we introduce an intermediate random system $\mathbf{S}$ and show how both $\mathbf{G}_\mathrm{r}$ and $\mathbf{H}_\mathrm{r}$ are conditionally equivalent to $\mathbf{S}$. This allows us to use Lemma 1 to bound the advantage in distinguishing $\mathbf{G}_\mathrm{r}$ and $\mathbf{H}_\mathrm{r}$. The system $\mathbf{S}$ also provides an interface to query l permutations $\pi_1, \ldots, \pi_l$. It works as follows: it answers any non-trivial forward query $\pi_i(x)$ with a value chosen uniformly from the set $\mathcal{X}_i^{(j-1)}$ and any non-trivial backward query $\pi_i^{-1}(x)$ with a value chosen uniformly from the set $\mathcal{X}_{i-1}^{(j-1)}$ (assuming it is the j^{th} query). Any trivial queries are answered consistently with previous answers. Moreover, if the queries form an l-chain for any tuple in $\mathsf{CS}(\pi_1, \ldots, \pi_l)$, $\mathbf{S}$ also gets blocked and responds with $\perp$ to any further queries. Note that $\mathbf{S}$ is only defined as long as $|\mathcal{X}_i^{(j-1)}| \geq 0$, but if this is not true, we have $h \geq 2^n$ and the lemma holds trivially.

[4] All permutations considered here are defined on the set $\{0,1\}^n$.

[5] The indexing of permutations is cyclic, e.g. π_{l+1} denotes the permutation π_1.

Let us now consider the j^{th} query that does not extend an $(l-1)$-chain (otherwise both $\mathbf{G}_{\text{r}}$ and $\mathbf{S}$ get blocked). Then the system $\mathbf{G}_{\text{r}}$ answers any non-trivial forward query $\pi_i(x)$ by a random element uniformly chosen from $\{0,1\}^n \setminus \text{range}_{j-1}(\pi_i)$ or gets blocked if this answer would create an l-chain by connecting two shorter chains. On the other hand, the system $\mathbf{S}$ answers with a random element uniformly chosen from $\mathcal{X}_i^{(j-1)}$, which is a subset of $\{0,1\}^n \setminus \text{range}_{j-1}(\pi_i)$. The situation for backward queries is analogous. Therefore, let us define a monotone condition $\mathcal{K}$ on $\mathbf{G}_{\text{r}}$: the event K_j is satisfied if K_{j-1} was satisfied and the answer to the j^{th} query was picked from the set $\mathcal{X}_i^{(j-1)}$ if it was a non-trivial forward query $\pi_i(x)$ or from the set $\mathcal{X}_{i-1}^{(j-1)}$ if it was a non-trivial backward query $\pi_i^{-1}(y)$. Note that as long as $\mathcal{K}$ is satisfied, no l-chain can emerge by connecting two shorter chains. By the previous observations and the definition of $\mathcal{K}$, we have $\mathbf{G}_{\text{r}}|\mathcal{K} \equiv \mathbf{S}$ which by Lemma 1 implies $\Delta_h(\mathbf{G}_{\text{r}}, \mathbf{S}) \leq \nu(\mathbf{G}_{\text{r}}, \overline{K_h})$. The probability that $\mathcal{K}$ is violated by the j^{th} answer is

$$\frac{|\text{dom}_{j-1}(\pi_{i+1}) \setminus \text{range}_{j-1}(\pi_i)|}{|\{0,1\}^n \setminus \text{range}(\pi_i)|} \leq \frac{|\{0,1\}^n \setminus \mathcal{X}_i^{(j-1)}|}{|\{0,1\}^n|} \leq \frac{j-1}{2^n},$$

which gives us $\nu(\mathbf{G}_{\text{r}}, \overline{K_h}) \leq \sum_{j=1}^{h} (j-1)/2^n \leq h^2/2^{n+1}$.

In the system $\mathbf{H}_{\text{r}}$, the permutations $\pi_1, \ldots, \pi_l$ can be seen as 2^n cycles of length l, each of which is formed by the edges connecting the vertices $x, \pi_1(x), \ldots, \pi_{l-1}(\cdots \pi_1(x) \cdots), x$ for some $x \in \{0,1\}^n$ and labeled by the respective permutations. We shall call such a cycle *used* if at least one of its edges was queried in either direction[6], otherwise we call it *unused*. Let us now define a monotone condition $\mathcal{L}$ on $\mathbf{H}_{\text{r}}$: the event L_j is satisfied if during the first j queries, any non-trivial query which did not extend an existing chain queried an unused cycle.

We claim that $\mathbf{H}_{\text{r}}|\mathcal{L} \equiv \mathbf{S}$. To see this, let us consider all possible types of queries. If the j^{th} query $\pi_i(x)$ is trivial or it extends an $(l-1)$-chain, both systems behave identically. Otherwise, the system $\mathbf{H}_{\text{r}}$ answers with a value y, where $y \notin \text{range}_{j-1}(\pi_i)$ (because π_i is a permutation) and $y \notin \text{dom}_{j-1}(\pi_{i+1})$, since that would mean that $\mathcal{L}$ was violated either earlier (if this query extends an existing chain) or now (if it starts a new chain). All values from $\mathcal{X}_i^{(j-1)}$ have the same probability of being y, because for any $y_1, y_2 \in \mathcal{X}_i^{(j-1)}$, there exists a natural bijective mapping between the arrangement of the cycles consistent with $\pi_i(x) = y_1$ or $\pi_i(x) = y_2$ (and all previous answers). Therefore, $\mathbf{H}_{\text{r}}$ answers with an uniformly chosen element from $\mathcal{X}_i^{(j-1)}$ and so does $\mathbf{S}$. For backward queries, the situation is analogous. By Lemma 1 this gives us $\Delta_h(\mathbf{S}, \mathbf{H}_{\text{r}}) \leq \nu(\mathbf{H}_{\text{r}}, \overline{L_h})$.

Let the j^{th} query be a non-trivial forward query $\pi_i(x)$ that does not extend a chain, i.e., $x \in \mathcal{X}_{i-1}^{(j-1)}$. Let u denote the number of elements in $\mathcal{X}_{i-1}^{(j-1)}$ that are in a used cycle on the position between π_{i-1} and π_i. Then since every element in $\mathcal{X}_{i-1}^{(j-1)}$ has the same probability of having this property (for the same reason as above), this query violates the condition $\mathcal{L}$ with probability $u/|\mathcal{X}_{i-1}^{(j-1)}| \leq (u +$

[6] We consider a separate edge connecting two vertices for each cycle in which they follow each other, hence each query creates at most one used cycle.

$|\mathrm{range}_{j-1}(\pi_{i-1}) \cup \mathrm{dom}_{j-1}(\pi_i)|)/2^n \leq (j-1)/2^n$. Hence $\nu(\mathbf{H}_\mathrm{r}, \overline{L_h}) \leq \sum_{j=1}^{h}(j-1)/2^n \leq h^2/2^{n+1}$.

Putting everything together, we have $\varDelta_h(\mathbf{G}_\mathrm{r}, \mathbf{H}_\mathrm{r}) \leq \varDelta_h(\mathbf{G}_\mathrm{r}, \mathbf{S}) + \varDelta_h(\mathbf{S}, \mathbf{H}_\mathrm{r}) \leq h^2/2^n$, which completes the proof. □

4 Conclusions

In this paper, we have studied the security of the cascade encryption. The most important recent result on this topic [4] contained a few mistakes, which we pointed out and corrected. We have formulated the proof from [4] in the random systems framework, which allows us to describe it on a more abstract level and thus in a more compact argument. This abstraction leads to a minor improvement for the case of triple encryption, as well as a generalization for the case of longer cascades. We prove that for the wide class of blockciphers with smaller key space than message space, a reasonable increase in the length of the cascade improves the encryption security. Our intention here was also to demonstrate the power of the random systems framework as a tool for modelling the behavior and interactions of discrete systems, with a focus towards analyzing their indistinguishability.

References

1. W. Aiello, M. Bellare, G. Di Crescenzo and R. Venkatesan: *Security Amplification by Composition: The case of Doubly-Iterated, Ideal Ciphers*, Advances in Cryptology - CRYPTO '98, LNCS vol. 1462, pp. 499–558, Springer-Verlag, 1998.
2. ANSI X9.52, *Triple Data Encryption Algorithm Modes of Operation*, 1998.
3. M. Bellare and Ch. Namprempre: *Authenticated Encryption: Relations among Notions and Analysis of the Generic Composition Paradigm*, full version, Cryptology ePrint Archive, Report 2000/025, 2007.
4. M. Bellare and P. Rogaway: *Code-Based Game-Playing Proofs and the Security of Triple Encryption*, Eurocrypt 2006, LNCS vol. 4004, pp. 409–426, Springer-Verlag, 2006. Full version at http://eprint.iacr.org/2004/331.
5. M. Bellare and T. Ristenpart: *Hash Functions in the Dedicated-Key Setting: Design Choices and MPP Transforms*, ICALP 2007, LNCS vol. 4596, pp. 399–410, Springer-Verlag, 2007.
6. J. S. Coron, J. Patarin and Y. Seurin: *The Random Oracle Model and the Ideal Cipher Model are Equivalent*, Advances in Cryptology - CRYPTO 2008, LNCS vol. 5157, pp. 1–20, Springer-Verlag, 2008.
7. W. Diffie and M. Hellman: *Exhaustive Cryptanalysis of the Data Encryption Standard*, Computer, vol. 10, pp. 74–84, 1977.
8. S. Even and O. Goldreich: *On the Power of Cascade Ciphers*, ACM Transactions on Computer Systems, vol. 3, no. 2, pp. 108–116, 1985.
9. S. Even and Y. Mansour: *A Construction of a Cipher from a Pseudorandom Permutation*, Asiacrypt '91, LNCS vol. 739, pp. 210–224, Springer-Verlag, 1992.
10. U. Maurer: *Indistinguishability of Random Systems*, Eurocrypt 2002, LNCS vol. 2332, pp. 110–132, Springer-Verlag, 2002.

11. U. Maurer and J. Massey: *Cascade Ciphers: the Importance of Being First*, J. of Cryptology, vol. 6, no. 1, pp. 55–61, 1993.
12. U. Maurer, K. Pietrzak and R. Renner: *Indistinguishability Amplification*, Advances in Cryptology - CRYPTO 2007, LNCS vol. 4622, pp. 130–149, Springer-Verlag, 2007.
13. National Institute of Standards and Technology: *FIPS PUB 46-3: Data Encryption Standard (DES)*, 1999.
14. National Institute of Standards and Technology: *Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher*, NIST Special Publication 800-67, 2004.
15. P. Rogaway and T. Shrimpton: *Deterministic Autenticated-Encryption*, Eurocrypt 2006, LNCS vol. 4004, pp. 373–390, Springer-Verlag, 2006.

A Problems with the Proof in [4]

The proof of a lower bound for the security of triple encryption presented in [4] contains some errors. We describe briefly where these errors come from, assuming the reader is familiar with the terminology and the proof from [4]. We shall be referring to the version 2.3 of the paper published at the online ePrint archive. The proof eventually comes down to bounding the advantage in distinguishing independent random permutations π_0, π_1, π_2 from random permutations π_0, π_1, π_2 such that $\pi_0 \circ \pi_1 \circ \pi_2 = id$ (distinguishing games G and H). This can be done easily if the distinguisher is allowed to extend a 2-chain by his queries, therefore the adversary is not allowed to do that in games G and H. To justify this, before proceeding to this part of the proof, the authors have to argue in a more complex setting (games D_S and R_3) that the probability of extending a 2-chain for the relevant keys is negligible. However, due to the construction of the adversary $B_{S,b}$ from the adversary B, extending a 2-chain by $B_{S,b}$ in the experiment $H^{B_{S,b}}$ does not correspond to extending a 2-chain by B in D_S^B, but to something we call a disconnected chain. The same can be said about the experiments R_3^B and $G^{B_{S,b}}$. Therefore, by bounding the probability of extending a 2-chain for the relevant keys in the experiment R_3^B, the authors do not bound the probability of extending a 2-chain in the experiment $G^{B_{S,b}}$, which they later need.

The second problem of the proof in [4] lies in bounding the probability of creating a chain using the game L. This is done by the equation $\mathsf{P}[R_3^B \text{ sets } x2ch] \leq 3 \cdot 2^{-k} + \mathsf{P}[B^L \text{ sets } bad]$ on page 19, which is also invalid. To see this, note that the game L only considers chains using subsequently the keys (K_0, K_1, K_2), while the flag $x2ch$ in the experiment R_3^B can also be set by a chain for any cyclic shift of this triple, e.g. (K_2, K_0, K_1). This is why a new multiplicative factor l appears in the security bound we have proved.

In the version 3.0 of the paper [4], the second bug mentioned here was fixed, while the first is still present in a different form. Now the games G and H_S can be easily distinguished by forming a disconnected chain, for example by the following trivial adversary B:

```
Adversary $B$
  $x_1 \overset{\$}{\leftarrow} \{0,1\}^n$;
  $x_2 \leftarrow \mathit{\Pi}(1, x_1)$; $x_3 \leftarrow \mathit{\Pi}(2, x_2)$; $x_0 \leftarrow S^{-1}(x_3)$; $x'_1 \leftarrow \mathit{\Pi}(0, x_0)$;
  if $x_1 = x'_1$ return 1 else return 0;
```

This problem can be fixed by introducing the concept of disconnected chains and bounding the probability of them being constructed by the adversary, as we do for the general case of l-cascades in Lemma 4.