Towards Key-Length Extension with Optimal Security: Cascade Encryption and Xor-cascade Encryption*

Jooyoung Lee**

Faculty of Mathematics and Statistics
Sejong University, Seoul, Korea 143-747
jlee05@sejong.ac.kr

Abstract. This paper discusses provable security of two types of cascade encryptions. The first construction CE^l, called l-cascade encryption, is obtained by sequentially composing l blockcipher calls with independent keys. The security of CE^l has been a longstanding open problem until Gaži and Maurer [9] proved its security up to $2^{\kappa+\min\{\frac{n}{2},\kappa\}}$ query complexity for large cascading length, where κ and n denote the key size and the block size of the underlying blockcipher, respectively. We improve this limit by proving the security of CE^l up to $2^{\kappa+\min\{\kappa,n\}-\frac{16}{l}\left(\frac{n}{2}+2\right)}$ query complexity: this bound approaches $2^{\kappa+\min\{\kappa,n\}}$ with increasing cascade length l.
The second construction XCE^l is a natural cascade version of the DESX scheme with intermediate keys xored between blockcipher calls. This can also be viewed as an extension of double XOR-cascade proposed by Gaži and Tessaro [10]. We prove that XCE^l is secure up to $2^{\kappa+n-\frac{8}{l}\left(\frac{n}{2}+2\right)}$ query complexity. As cascade length l increases, this bound approaches $2^{\kappa+n}$.
In the ideal cipher model, one can obtain all the evaluations of the underlying blockcipher by making $2^{\kappa+n}$ queries, so the $(\kappa+n)$-bit security becomes the maximum that key-length extension based on a single κ-bit key n-bit blockcipher is able to achieve. Cascade encryptions CE^l (with $n \leq \kappa$) and XCE^l provide almost optimal security with large cascade length.

1 Introduction

The key length of a blockcipher, say κ, is a crucial factor that limits its achievable security level: no matter how carefully designed, one can recover its secret key simply by trying all possible 2^κ keys. For example, the Data Encryption Standard (DES) [1] using 56-bit keys was one of the most predominant algorithms for encryption of data. No feasible attacks faster than a brute-force attack have been proposed (as most of them require a huge amount of data), while advances in computational power made a brute-force attack itself practical. As a result, DES was replaced by a new standard algorithm AES [4]. On the other hand, in order to protect legacy applications based on DES, there have been considerable research on constructing DES-based encryption schemes which employ longer keys. This approach is called *key-length extension*, for which Triple-DES [2, 3, 5] and DESX (due to Rivest) are the most popular constructions.

The Triple-DES approach transforms a κ-bit key n-bit blockcipher E into an encryption scheme that accepts three κ-bit keys $k_1, k_2, k_3 \in \{0,1\}^\kappa$ and encrypts an n-bit message block u as $v = E_{k_3}(E_{k_2}(E_{k_1}(u)))$. Bellare and Rogaway [6] proved its security up to $2^{\kappa+\frac{\min\{n,\kappa\}}{2}}$

** The work of J. Lee was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2012-0003157).

query complexity assuming E is an ideal blockcipher. The triple encryption can be naturally extended to sequentially composing more than three blockcipher calls with independent keys. It has been a longstanding open problem if the security of cascade encryption improves with increasing cascade length. Recently, Gaži and Maurer [9] partially answered this question by showing the security bound (in terms of the threshold number of queries) approaches roughly the value $2^{\kappa+\min\{\frac{n}{2},\kappa\}}$. In this paper, we will revisit this question.

The DESX approach transforms a κ-bit key n-bit blockcipher E into an encryption scheme that accepts a κ-bit key $k \in \{0,1\}^\kappa$ and additional n-bit whitening keys $k_i, k_o \in \{0,1\}^n$ and encrypts an n-bit message block u as $v = k_o \oplus E_k(k_i \oplus u)$. Killan and Rogaway [13] proved its security up to $2^{\frac{\kappa+n}{2}}$ query complexity. As an efficient key-length extension, Gaži and Tessaro [10] proposed a cascade of two DESX schemes with some refinement, and proved its security up to $2^{\kappa+\frac{n}{2}}$ query complexity.

Our Contribution. Since one can obtain all the evaluations of a κ-bit key n-bit blockcipher by making $2^{\kappa+n}$ queries, the $(\kappa+n)$-bit security becomes the maximum that key-length extension based on a single κ-bit key n-bit blockcipher is able to achieve: a standard brute-force attack of $2^{\kappa+n}$ query complexity is given in Appendix A.

Therefore it is natural to ask if there is key-length extension with the optimal $(\kappa+n)$-bit security. In order to answer this question, we consider two types of cascade encryptions. The first construction is a regular cascade encryption. Formally, *l-cascade encryption* CE^l accepts an $l\kappa$-bit key $\mathbf{k} = (k_1, \ldots, k_l) \in (\{0,1\}^\kappa)^l$ and encrypts a plaintext $u \in \{0,1\}^n$ by computing

$$v = \mathsf{CE}^l_{\mathbf{k}}[E](u) = E_{k_l} \circ E_{k_{l-1}} \circ \cdots \circ E_{k_2} \circ E_{k_1}(u).$$

In this paper, we prove that CE^l is pseudorandom up to $2^{\kappa+\min\{\kappa,n\}-\frac{16}{l}\left(\frac{n}{2}+2\right)}$ query complexity (ignoring log factor). As cascade length l increases, this bound approaches $2^{\kappa+\min\{\kappa,n\}}$, improving the limit $2^{\kappa+\min\{\frac{n}{2},\kappa\}}$ given by Gaži and Maurer when $\frac{n}{2} < \kappa$.

The second construction can be viewed as a cascade of DESX: *l-xor-cascade encryption* XCE^l accepts an $(l\kappa+(l+1)n)$-bit key $(\mathbf{k},\mathbf{z}) \in (\{0,1\}^\kappa)^l \times (\{0,1\}^n)^{l+1}$ and for

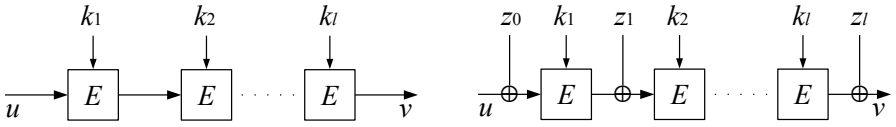
$$\mathbf{k} = (k_1, \ldots, k_l) \in (\{0,1\}^\kappa)^l \text{ and } \mathbf{z} = (z_0, \ldots, z_l) \in (\{0,1\}^n)^{l+1},$$

encrypts a plaintext $u \in \{0,1\}^n$ by computing

$$v = \mathsf{XCE}^l_{\mathbf{k},\mathbf{z}}[E](u) = \oplus_{z_l} \circ E_{k_l} \circ \oplus_{z_{l-1}} \circ \cdots \circ \oplus_{z_1} \circ E_{k_1} \circ \oplus_{z_0}(u),$$

where for $z \in \{0,1\}^n$, $\oplus_z$ denotes the mapping $x \mapsto x \oplus z$ from $\{0,1\}^n$ to itself. We prove the security of XCE^l up to $2^{\kappa+n-\frac{8}{l}\left(\frac{n}{2}+2\right)}$ query complexity. With increasing cascade length, this bound approaches $2^{\kappa+n}$. So XCE^l asymptotically provides optimal security with large cascade length, and this observation also applies to cascade encryption CE^l if $n \le \kappa$ (as in the case of DES and AES). See Figure 1 for pictorial representation of CE^l and XCE^l.

Proof Techniques. We will use a combinatorial framework that lifts the NCPA-security of $l/2$-cascade construction to the CPA security of l-cascade construction. Maurer, Pietrzak, and Renner [15] proved that if two independent encryption schemes F and G are NCPA-secure, then $F \circ G^{-1}$ is CPA-secure. Combinatorial interpretation of this property, based on Lemma 2, was first introduced in [14], where the key-alternating cipher of t rounds is viewed as a composition of two independent key-alternating cipher of $t/2$ rounds, and the NCPA-security of each component is analyzed. A similar approach can be applied to our



(a) l-cascade encryption CE^l (b) l-xor-cascade encryption XCE^l

Fig. 1. Two types of cascade encryptions

constructions, while a difficulty comes from the fact that the two components are commonly based on a single blockcipher. We address this problem by using *random key space separation*: randomly partition a key space into two subspaces of the same size and make the first $l/2$ blockcipher calls use keys from one key subspace and the last $l/2$ calls from the other. The modified key sampling process is shown to be indistinguishable from the original one, while by having the two components use their keys from separate key subspaces, we can view a cascade encryption as a composition of two independent ones.

The NCPA-security of each component is proved by coupling technique. Since first introduced by Mironov [16] in a cryptographic context and recently revisited by Morris, Rogaway and Stegers [17] to analyze the security of maximally unbalanced Feistel networks, it became a powerful tool used for the security proof of various types of iterated constructions including generalized Feistel networks, shuffling-based encryption schemes and key-alternating ciphers [14, 11, 12]. Careful definition and analysis of a coupling, given in the proof of Lemma 5, is the core of our security proof.

Other Related Work. Recently, Gaži [8] presented a distinguishing attack on cascade encryption of odd (resp. even) length l using roughly $2^{\kappa+\frac{l-1}{l+1}n}$ (resp. $2^{\kappa+\frac{l-2}{l}n}$) queries. For xor-cascade encryption of length l (and its generalization), a distinguishing attack of $2^{\kappa+\frac{l-1}{l}n}$ query complexity is presented. In the random system framework, the security of xor-cascade encryption of odd (resp. even) length l is proved up to $2^{\kappa+\frac{l-1}{l+1}n}$ (resp. $2^{\kappa+\frac{l-2}{l}n}$) query complexity, and especially up to $2^{\kappa+\frac{l-1}{l}n}$ query complexity for $l \in \{3, 4\}$. These lower bounds are tighter than ours.

2 Preliminaries

2.1 General Notation

For an integer $n \geq 1$, let $I_n = \{0,1\}^n$ be the set of binary strings of length n. The set of all permutations on I_n will be denoted $\mathcal{P}_n$. We will usually write $N = 2^n$.

For a set T and an integer $s \geq 1$, T^{*s} denotes the set of all sequences that consists of s pairwise distinct elements of T. For integers $1 \leq s \leq t$, we will write $(t)_s = t(t-1)\cdots(t-s+1)$. If $|T| = t$, then $(t)_s$ becomes the size of T^{*s}.

2.2 The Ideal Cipher Model

A blockcipher is a function family $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ such that for all $k \in \mathcal{K}$ the mapping $E(k,\cdot)$ is a permutation on I_n. We write $BC(\mathcal{K}, n)$ to mean the set of all such blockciphers, shortening to $BC(\kappa, n)$ when $\mathcal{K} = \{0,1\}^\kappa$. In the ideal cipher model, a blockcipher E is chosen from $BC(\mathcal{K}, n)$ uniformly at random. It allows for two types of oracle queries $E(k, x)$

and $E^{-1}(k,y)$ for $x,y \in \{0,1\}^n$ and $k \in \mathcal{K}$.[1] The response to an inverse query $E^{-1}(k,y)$ is $x \in \{0,1\}^n$ such that $E(k,x) = y$.

2.3 Indistinguishability

Let $\mathsf{C} \in \{\mathsf{CE}^l, \mathsf{XCE}^l\}$ be an n-bit encryption scheme that employs λ-bit keys and makes oracle queries to a blockcipher $E \in BC(\kappa, n)$. So each key $\mathbf{k} \in \{0,1\}^\lambda$ and a blockcipher $E \in BC(\kappa, n)$ define a permutation $\mathsf{C}_\mathbf{k}[E]$ on I_n. In the *indistinguishability* framework (in the ideal cipher model), $\mathsf{C}_\mathbf{k}[E]$ uses a random secret key $\mathbf{k}$ and makes oracle queries to an ideal blockcipher E, while a permutation P is chosen uniformly at random from $\mathcal{P}_n$. A distinguisher $\mathcal{A}$ would like to tell apart two worlds $(\mathsf{C}_\mathbf{k}[E], E)$ and (P, E) by adaptively making forward and backward queries to the permutation and the blockcipher. Formally, $\mathcal{A}$'s distinguishing advantage is defined by

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{C}}(\mathcal{A}) = \mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa,n) : \mathcal{A}[P,E] = 1\right] - \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} \{0,1\}^\lambda, E \xleftarrow{\$} BC(\kappa,n) : \mathcal{A}[\mathsf{C}_\mathbf{k}[E], E] = 1\right].$$

For q_1, $q_2 > 0$, we define

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{C}}(q_1, q_2) = \max_{\mathcal{A}} \mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{C}}(\mathcal{A}),$$

where the maximum is taken over all adversaries $\mathcal{A}$ making at most q_1 queries to the outer permutation and at most q_2 queries to the underlying blockcipher.

Combinatorial Framework. We assume that a distinguisher $\mathcal{A}$ making q_1 forward and/or backward queries to the permutation oracle records a query history

$$\mathcal{Q}_1 = (u^i, v^i)_{1 \le i \le q_1},$$

where (u^i, v^i) represents the evaluation obtained by the i-th query to the permutation oracle. So according to the instantiation, it implies either $\mathsf{C}_\mathbf{k}[E](u^i) = v^i$ or $P(u^i) = v^i$. By making q_2 queries to the underlying blockcipher E, $\mathcal{A}$ also records the second query history

$$\mathcal{Q}_2 = (x^i, k^i, y^i)_{1 \le i \le q_2},$$

where (x^i, k^i, y^i) represents the evaluation $E(k^i, x^i) = y^i$ obtained by the i-th query to the blockcipher. The pair of the query histories

$$\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$$

is called the *transcript* of the attack; it contains all the information that $\mathcal{A}$ has obtained at the end of the attack. In this work, we will only consider information theoretic distinguishers. Therefore we can assume that a distinguisher is deterministic without making any redundant queries, and hence the output of $\mathcal{A}$ can be regarded as a function of $\mathcal{T}$, denoted $\mathcal{A}(\mathcal{T})$ or $\mathcal{A}(\mathcal{Q}_1, \mathcal{Q}_2)$.

If a permutation $\mathsf{C}_\mathbf{k}[E]$(resp. P) is consistent with $\mathcal{Q}_1$, i.e., $\mathsf{C}_\mathbf{k}[E](u^i) = v^i$(resp. $P(u^i) = v^i$) for every $i = 1, \ldots, q_1$, then we will write $\mathsf{C}_\mathbf{k}[E] \vdash \mathcal{Q}_1$(resp. $P \vdash \mathcal{Q}_1$). Similarly, if a

[1] We interchangeably use both representations $E(k,x)$ and $E_k(x)$, and similarly $E^{-1}(k,y)$ and $E_k^{-1}(y)$.

blockcipher $E \in BC(\kappa, n)$ is consistent with $\mathcal{Q}_2$ (i.e., $E(k^i, x^i) = y^i$ for $i = 1, \ldots, q_2$), then we will write $E \vdash \mathcal{Q}_2$. Using these notations, we have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{C}}(\mathcal{A}) = &\sum_{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1} \mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa, n) : P \vdash \mathcal{Q}_1 \wedge E \vdash \mathcal{Q}_2\right] \\ &- \sum_{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1} \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} \{0,1\}^\lambda, E \xleftarrow{\$} BC(\kappa, n) : \mathsf{C}_{\mathbf{k}}[E] \vdash \mathcal{Q}_1 \wedge E \vdash \mathcal{Q}_2\right], \quad (1)\end{aligned}$$

where the sum is taken over all the possible transcripts $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$ such that $\mathcal{A}(\mathcal{Q}_1, \mathcal{Q}_2) = 1$.[2]

2.4 Coupling Technique

Given a finite event space Ω and two probability distributions μ and ν defined on Ω, the *total variation distance* between μ and ν, denoted $\|\mu - \nu\|$, is defined as

$$\|\mu - \nu\| = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

The following definitions are also all equivalent.

$$\|\mu - \nu\| = \max_{S \subset \Omega}\{\mu(S) - \nu(S)\} = \max_{S \subset \Omega}\{\nu(S) - \mu(S)\} = \max_{S \subset \Omega}\{|\mu(S) - \nu(S)|\}.$$

A *coupling* of μ and ν is a distribution τ on $\Omega \times \Omega$ such that for all $x \in \Omega$, $\sum_{y \in \Omega} \tau(x, y) = \mu(x)$ and for all $y \in \Omega$, $\sum_{x \in \Omega} \tau(x, y) = \nu(x)$. In other words, τ is a joint distribution whose marginal distributions are respectively μ and ν. We will use the following two lemmas in subsequent security proofs.

Lemma 1. *Let μ and ν be probability distributions on a finite event space Ω, let τ be a coupling of μ and ν, and let (X, Y) be a random variable sampled according to distribution τ. Then $\|\mu - \nu\| \leq \mathbf{Pr}[X \neq Y]$.*

Lemma 2. *Let Ω be some finite event space and ν be the uniform probability distribution on Ω. Let μ be a probability distribution on Ω such that $\|\mu - \nu\| \leq \epsilon$. Then there is a set $S \subset \Omega$ such that*

1. $|S| \geq (1 - \sqrt{\epsilon})|\Omega|$,
2. $\mu(x) \geq (1 - \sqrt{\epsilon})\nu(x)$ *for every* $x \in S$.

The proof of the above lemmas is given in [14]. For completeness, we include the same proof in Appendix B.

[2] Here we only consider "valid" transcripts that $\mathcal{A}$ might produce by communicating with a permutation $P \in \mathcal{P}_n$ and a blockcpher $E \in BC(\kappa, n)$. For example, in a valid transcript $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$, (x, y) and (x', y) with $x \neq x'$ could not be both contained in $\mathcal{Q}_1$.

3 Security Proofs

In the security proof of cascade encryption CE^l, we will assume that for any x, $y \in I_n$, there are at most β keys k such that $(x, k, y) \in \mathcal{Q}_2$. Define the weight of $\mathcal{Q}_2$ by

$$\omega(\mathcal{Q}_2) = \max_{x,y \in I_n} |\{k : (x, k, y) \in \mathcal{Q}_2\}| .$$

Then we have

$$\mathbf{Pr}\left[E \xleftarrow{\$} BC(\kappa, n) : \omega(\mathcal{Q}_2) > \beta\right] \leq 2^{2n-\beta} \tag{2}$$

for any $\beta \geq e2^{\kappa-n+1}$. Note that a distinguisher $\mathcal{A}$ is deterministic, so once E is chosen then $\mathcal{Q}_2$, and hence $\omega(\mathcal{Q}_2)$ is uniquely determined. This bound has already been used in [6, 9], while for completeness we give a proof in Appendix C. With this probabilistic restriction, the security proof of cascade encryption CE^l will use the following lemma.

Lemma 3. *Let $\delta > 0$ and $\beta \geq e2^{\kappa-n+1}$. Assume that for any transcript $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$ such that $|\mathcal{Q}_1| = q_1$, $|\mathcal{Q}_2| = q_2$ and $\omega(\mathcal{Q}_2) \leq \beta$, we have*

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) \geq (1 - \delta)\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2),$$

where

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) = \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} I_\kappa^l, E \xleftarrow{\$} BC(\kappa, n) : \mathsf{CE}_{\mathbf{k}}^l[E] \vdash \mathcal{Q}_1 \;\middle|\; E \vdash \mathcal{Q}_2\right],$$
$$\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2) = \mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa, n) : P \vdash \mathcal{Q}_1 \;\middle|\; E \vdash \mathcal{Q}_2\right] = 1/(N)_{q_1}.$$

Then we have

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CE}^l}(\mathcal{A}) \leq \delta + 2^{2n-\beta}.$$

Proof. For a transcript $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$, define

$$\begin{aligned}
\mathsf{p}(\mathcal{Q}_2) &= \mathbf{Pr}\left[E \xleftarrow{\$} BC(\kappa, n) : E \vdash \mathcal{Q}_2\right], \\
\mathsf{p}_1(\mathcal{Q}_1, \mathcal{Q}_2) &= \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} I_\kappa^l, E \xleftarrow{\$} BC(\kappa, n) : \mathsf{CE}_{\mathbf{k}}^l[E] \vdash \mathcal{Q}_1 \wedge E \vdash \mathcal{Q}_2\right] \\
&= \mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2), \\
\mathsf{p}_2(\mathcal{Q}_1, \mathcal{Q}_2) &= \mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa, n) : P \vdash \mathcal{Q}_1 \wedge E \vdash \mathcal{Q}_2\right] \\
&= \mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2).
\end{aligned}$$

Then by (1) and (2), we have

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CE}^l}(\mathcal{A}) &= \sum_{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1} \mathsf{p}_2(\mathcal{Q}_1,\mathcal{Q}_2) - \sum_{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1} \mathsf{p}_1(\mathcal{Q}_1,\mathcal{Q}_2) \\
&= \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) - \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) \\
&+ \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)>\beta}} \mathsf{p}_2(\mathcal{Q}_1,\mathcal{Q}_2) - \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)>\beta}} \mathsf{p}_1(\mathcal{Q}_1,\mathcal{Q}_2) \\
&\le \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) - \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) + \sum_{\omega(\mathcal{Q}_2)>\beta} \mathsf{p}_2(\mathcal{Q}_2) \\
&\le \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) - (1-\delta) \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2)\mathsf{p}(\mathcal{Q}_2) + 2^{2n-\beta} \\
&\le \delta \sum_{\substack{\mathcal{A}(\mathcal{Q}_1,\mathcal{Q}_2)=1 \\ \omega(\mathcal{Q}_2)\le\beta}} \mathsf{p}_2(\mathcal{Q}_1,\mathcal{Q}_2) + 2^{2n-\beta} \le \delta + 2^{2n-\beta}. \qquad \square
\end{aligned}$$

In the security proof of xor-cascade encryption XCE^l, we put no restriction on the weight of $\mathcal{Q}_2$. In this case, we can use the following lemma whose proof is similar as Lemma 3. (We might simply apply $\beta = \infty$ to Lemma 3.)

Lemma 4. *Let $\delta > 0$. Assume that for any transcript $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$ such that $|\mathcal{Q}_1| = q_1$ and $|\mathcal{Q}_2| = q_2$, we have*

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) \ge (1-\delta)\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2),$$

where

$$\begin{aligned}
\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) &= \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} I_\kappa^l, \mathbf{z} \xleftarrow{\$} I_n^{l+1}, E \xleftarrow{\$} BC(\kappa,n) : \mathsf{XCE}^l_{\mathbf{k},\mathbf{z}}[E] \vdash \mathcal{Q}_1 \;\middle|\; E \vdash \mathcal{Q}_2\right], \\
\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2) &= \mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa,n) : P \vdash \mathcal{Q}_1 \;\middle|\; E \vdash \mathcal{Q}_2\right] = 1/(N)_{q_1}.
\end{aligned}$$

Then we have

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{XCE}^l}(q_1,q_2) \le \delta.$$

3.1 Security of Cascade Encryption

In this section, we analyze the security of cascade encryption CE^l for even length $l = 2d$. We begin with slightly modifying the key sampling process of CE^l. Consider the following three key sampling processes.

A: Choose $\mathbf{k} \in I_\kappa^l$ uniformly at random.
B: Choose $\mathbf{k} \in (I_\kappa)^{*l}$ uniformly at random.
C: Randomly partition $T_1 \cup T_2 = I_\kappa$ so that $|T_1| = |T_2|$, choose $\mathbf{k}' \in (T_1)^{*d}$ and $\mathbf{k}'' \in (T_2)^{*d}$ uniformly at random, and then define $\mathbf{k} = (\mathbf{k}', \mathbf{k}'')$.

One can distinguish sampling processes **A** and **B** with advantage at most

$$\binom{l}{2}\frac{1}{2^{\kappa}} \leq \frac{l^2}{2^{\kappa+1}}. \tag{3}$$

On the other hand, sampling processes **B** and **C** have exactly the same probability distribution. (See Appendix D for the proof.) Taking into account (3), we will analyze the security of

$$\mathsf{CE}^{l}_{\mathbf{k}}[E] = \mathsf{CE}^{d}_{\mathbf{k}''}[E] \circ \mathsf{CE}^{d}_{\mathbf{k}'}[E],$$

where $\mathbf{k}$, $\mathbf{k}'$ and $\mathbf{k}''$ are defined by key sampling process **C** instead of the original process **A**.

If $\mathsf{CE}^{l}_{\mathbf{k}}[E] \vdash \mathcal{Q}_1$ for a query history $\mathcal{Q}_1 = (u^i, v^i)_{1\leq i\leq q_1}$, then it follows that

$$\mathsf{CE}^{d}_{\mathbf{k}'}[E] \vdash (u^i, w^i)_{1\leq i\leq q_1} \text{ and } \mathsf{CE}^{d}_{\mathbf{k}''}[E] \vdash (w^i, v^i)_{1\leq i\leq q_1},$$

for some $\mathbf{w} = (w^i)_{1\leq i\leq q_1} \in (I_n)^{*q_1}$. Therefore for a transcript $\mathcal{T} = (\mathcal{Q}_1, \mathcal{Q}_2)$, we have

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) = \sum_{\mathbf{w}\in\Omega} \mathbf{Pr}\Big[T_1 \xleftarrow{\$} \mathcal{P}_{2^{\kappa-1}}(I_\kappa), \mathbf{k}' \xleftarrow{\$} (T_1)^{*d}, \mathbf{k}'' \xleftarrow{\$} (T_2)^{*d},$$
$$E \xleftarrow{\$} BC(\kappa, n) : \mathsf{CE}^{d}_{\mathbf{k}'}[E] \vdash (u^i, w^i) \wedge \mathsf{CE}^{d}_{\mathbf{k}''}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}_2\Big],$$

where $\Omega = (I_n)^{*q_1}$, $\mathcal{P}_{2^{\kappa-1}}(I_\kappa)$ is the set of all subsets of I_κ of size $2^{\kappa-1}$, and $T_2 = I_\kappa \backslash T_1$.

Given a partition (T_1, T_2) of I_κ, a blockcipher $E \in BC(\kappa, n)$ is naturally partitioned into two blockciphers $E' \in BC(T_1, n)$ and $E'' \in BC(T_2, n)$, and vice versa. Given a query history $\mathcal{Q}_2$ for E, then this partition also induces two query histories $\mathcal{Q}'_2$ for E' and $\mathcal{Q}''_2$ for E''. Namely, for $\mathcal{Q}_2 = (x^i, k^i, y^i)_{1\leq i\leq q_2}$, $\mathcal{Q}'_2 = (x^i, k^i, y^i)_{1\leq i\leq q_2, k^i\in T_1}$ and $\mathcal{Q}''_2 = (x^i, k^i, y^i)_{1\leq i\leq q_2, k^i\in T_2}$. With these notations, we have

$$\mathbf{Pr}\Big[T_1 \xleftarrow{\$} \mathcal{P}_{2^{\kappa-1}}(I_\kappa), \mathbf{k}' \xleftarrow{\$} (T_1)^{*d}, \mathbf{k}'' \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(\kappa, n) :$$
$$\mathsf{CE}^{d}_{\mathbf{k}'}[E] \vdash (u^i, w^i) \wedge \mathsf{CE}^{d}_{\mathbf{k}''}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}_2\Big]$$
$$= \frac{1}{\binom{2^\kappa}{2^{\kappa-1}}} \sum_{\substack{T_1\cup T_2 = I_\kappa \\ |T_1|=|T_2|=2^{\kappa-1}}} \mathbf{Pr}\Big[\mathbf{k}' \xleftarrow{\$} (T_1)^{*d}, \mathbf{k}'' \xleftarrow{\$} (T_2)^{*d}, E' \xleftarrow{\$} BC(T_1, n),$$
$$E'' \xleftarrow{\$} BC(T_2, n) : \mathsf{CE}^{d}_{\mathbf{k}'}[E'] \vdash (u^i, w^i) \wedge \mathsf{CE}^{d}_{\mathbf{k}''}[E''] \vdash (w^i, v^i) \ \Big| E' \vdash \mathcal{Q}'_2 \wedge E'' \vdash \mathcal{Q}''_2\Big],$$

and hence

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) = \frac{1}{\binom{2^\kappa}{2^{\kappa-1}}} \sum_{\substack{T_1\cup T_2 = I_\kappa \\ |T_1|=|T_2|=2^{\kappa-1}}} \sum_{\mathbf{w}\in\Omega} \mathbf{Pr}\Big[\mathbf{k}' \xleftarrow{\$} (T_1)^{*d}, \mathbf{k}'' \xleftarrow{\$} (T_2)^{*d}, E' \xleftarrow{\$} BC(T_1, n),$$
$$E'' \xleftarrow{\$} BC(T_2, n) : \mathsf{CE}^{d}_{\mathbf{k}'}[E'] \vdash (u^i, w^i) \wedge \mathsf{CE}^{d}_{\mathbf{k}''}[E''] \vdash (w^i, v^i) \ \Big| E' \vdash \mathcal{Q}'_2 \wedge E'' \vdash \mathcal{Q}''_2\Big],$$

where

$$\sum_{\mathbf{w}\in\Omega} \mathbf{Pr}\Big[\mathbf{k}' \xleftarrow{\$} (T_1)^{*d}, \mathbf{k}'' \xleftarrow{\$} (T_2)^{*d}, E' \xleftarrow{\$} BC(T_1,n), E'' \xleftarrow{\$} BC(T_2,n) :$$

$$\mathsf{CE}^d_{\mathbf{k}'}[E'] \vdash (u^i, w^i) \wedge \mathsf{CE}^d_{\mathbf{k}''}[E''] \vdash (w^i, v^i) \ \Big| E' \vdash \mathcal{Q}'_2 \wedge E'' \vdash \mathcal{Q}''_2\Big]$$

$$= \sum_{\mathbf{w}\in\Omega} \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_1)^{*d}, E \xleftarrow{\$} BC(T_1,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i) \ \Big| E \vdash \mathcal{Q}'_2\Big]$$

$$\times \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(T_2,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}''_2\Big]. \quad (4)$$

In order to upper bound each factor of the products appearing in (4), we fix a query history $\mathcal{Q}_2 = (x^i, k^i, y^i)_{1\le i\le q'}$ such that $q' \le q_2$ and $\omega(\mathcal{Q}_2) \le \beta$, and define a probability distribution $\mu_{\mathbf{s}}$ for each $\mathbf{s} = (s^i)_{1\le i\le q_1} \in \Omega$, where for each $\mathbf{w} = (w^i)_{1\le i\le q_1} \in \Omega$,

$$\mu_{\mathbf{s}}(\mathbf{w}) = \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} (I_\kappa)^{*d}, E \xleftarrow{\$} BC(\kappa,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (s^i, w^i)_{1\le i\le q_1} \ \Big| E \vdash \mathcal{Q}_2\right].$$

Using the coupling technique, we can upper bound the statistical distance between $\mu_{\mathbf{s}}$ and the uniform probability distribution. The proof will be given at the end of this section.

Lemma 5. *Let d be even, let $\mu_{\mathbf{s}}$ be the probability distribution defined as above, and let ν be the uniform probability distribution on Ω. Then for $M > 0$, we have $\|\mu_{\mathbf{s}} - \nu\| \le \epsilon$, where*

$$\epsilon = q_1 \left(\frac{2q_2}{M(2^\kappa - d)} + \frac{2M\beta}{2^\kappa - d} + \frac{2M}{N-M}\right)^{\frac{d}{2}}.$$

Applying Lemma 5 with $\mathbf{s} = \mathbf{u} = (u^i)_{1\le i\le q_1}$, $\mathcal{Q}_2 = \mathcal{Q}'_2$ and

$$\mu_{\mathbf{u}}(\mathbf{w}) = \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_1)^{*d}, E \xleftarrow{\$} BC(T_1,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i) \ \Big| E \vdash \mathcal{Q}'_2\Big],$$

and using Lemma 2, we have a subset $S_1 \subset \Omega$ such that $|S_1| \ge (1-\sqrt{\epsilon})|\Omega|$ and

$$\mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_1)^{*d}, E \xleftarrow{\$} BC(T_1,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i) \ \Big| E \vdash \mathcal{Q}'_2\Big]$$

$$\ge (1-\sqrt{\epsilon})\nu(\mathbf{w}) = \frac{1-\sqrt{\epsilon}}{(N)_{q_1}}$$

for every $\mathbf{w} \in S_1$, where

$$\epsilon = q_1 \left(\frac{2q_2}{M(2^{\kappa-1} - d)} + \frac{2M\beta}{2^{\kappa-1} - d} + \frac{2M}{N-M}\right)^{\frac{d}{2}}.$$

Here $BC(T_1, n)$ is viewed as equivalent to $BC(\kappa - 1, n)$.

For $\mathcal{Q}''_2$, define $\mathcal{Q}'''_2$ where $(x,k,y) \in \mathcal{Q}''_2$ if and only if $(y,k,x) \in \mathcal{Q}'''_2$. Again, applying Lemma 5 with $\mathbf{s} = \mathbf{v} = (v^i)_{1\le i\le q_1}$, $\mathcal{Q}_2 = \mathcal{Q}'''_2$ and

$$\mu_{\mathbf{v}}(\mathbf{w}) = \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(T_2,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (v^i, w^i) \ \Big| E \vdash \mathcal{Q}'''_2\Big]$$

$$= \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(T_2,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}''_2\Big],$$

we have a subset $S_2 \subset \Omega$ such that $|S_2| \geq (1-\sqrt{\epsilon})|\Omega|$ and

$$\mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(T_2,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}_2''\Big] \geq (1-\sqrt{\epsilon})\nu(\mathbf{w}) = \frac{1-\sqrt{\epsilon}}{(N)_{q_1}}$$

for every $w \in S_2$. Let $S = S_1 \cap S_2$. Since $|S| \geq (1-2\sqrt{\epsilon})|\Omega|$, it follows that

$$\begin{aligned}\sum_{\mathbf{w}\in\Omega} \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_1)^{*d}, E \xleftarrow{\$} BC(T_1,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i) \ \Big| E \vdash \mathcal{Q}_2'\Big] \\ \times \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, E \xleftarrow{\$} BC(T_2,n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (w^i, v^i) \ \Big| E \vdash \mathcal{Q}_2''\Big] \\ \geq (1-2\sqrt{\epsilon})|\Omega| \cdot \left(\frac{1-\sqrt{\epsilon}}{(N)_{q_1}}\right)^2 \geq (1-4\sqrt{\epsilon})\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2).\end{aligned}$$

Therefore we have

$$\begin{aligned}\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) &\geq \frac{1}{\binom{2^\kappa}{2^{\kappa-1}}} \sum_{\substack{T_1 \cup T_2 = I_\kappa \\ |T_1|=|T_2|=2^{\kappa-1}}} (1-4\sqrt{\epsilon})\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2) \\ &= (1-4\sqrt{\epsilon})\mathsf{p}_2(\mathcal{Q}_1|\mathcal{Q}_2). \qquad (5)\end{aligned}$$

By (3), (5) and Lemma 3, we have the following theorem.

Theorem 1. *Let CE^l be an l-cascade encryption scheme using a κ-bit key n-bit blockcipher. If $l = 2d$ and d is even, then for $M > 0$ and $\beta \geq e2^{\kappa-n}$,*

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CE}^l}(q_1,q_2) \leq \frac{l^2}{2^{\kappa+1}} + 4q_1^{\frac{1}{2}}\left(\frac{2q_2}{M(2^{\kappa-1}-d)} + \frac{2M\beta}{2^{\kappa-1}-d} + \frac{2M}{N-M}\right)^{\frac{d}{4}} + 2^{2n-\beta}.$$

Optimizing parameters. Let $\beta \geq \max\{3n, e2^{\kappa-n}\}$. Then $3n \leq \beta$, and hence $2^{2n-\beta} \leq 1/2^n$. Let $M = \sqrt{\frac{q_2}{\beta}}$ by solving $\frac{2q_2}{M(2^{\kappa-1}-d)} = \frac{2M\beta}{2^{\kappa-1}-d}$. Then for $q_2 \leq 2^{\kappa+n}$, $M \leq \sqrt{\frac{2^{\kappa+n}}{e2^{\kappa-n}}} \leq \frac{N}{\sqrt{e}}$ and hence

$$\frac{1}{N-M} \leq \left(1-\frac{1}{\sqrt{e}}\right)^{-1}\frac{1}{N} \leq \frac{e}{N}.$$

This implies

$$\frac{2M}{N-M} \leq \frac{2M \cdot e2^{\kappa-n}}{2^\kappa} \leq \frac{2M\beta}{2^{\kappa-1}-d}\left(= \frac{4M\beta}{2^\kappa - l}\right).$$

Using this inequality, the upper bound of Theorem 1 is simplified as follows.

Corollary 1. *Let CE^l be an l-cascade encryption scheme using a κ-bit key n-bit blockcipher. If l is a multiple of 4, then for $\beta \geq \max\{3n, e2^{\kappa-n}\}$,*

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CE}^l}(q_1,q_2) \leq \frac{l^2}{2^{\kappa+1}} + 4q_1^{\frac{1}{2}}\left(\frac{12\sqrt{\beta q_2}}{2^\kappa - l}\right)^{\frac{l}{8}} + \frac{1}{2^n}.$$

INTERPRETATION. Assuming that $l^2/2^{\kappa+1}$ and $1/2^n$ are negligible, focus on the second term of the above upper bound. If we set $q_1 = 2^n$ to the maximum number of queries to the outer permutation and approximate $2^\kappa - l \approx 2^\kappa$, then the distinguishing advantage becomes negligible when

$$q_2 \ll \frac{2^{2\kappa - \frac{16}{l}(\frac{n}{2}+2)}}{144\beta} \leq \min\left\{ \frac{2^{2\kappa - \frac{16}{l}(\frac{n}{2}+2)}}{432n}, \frac{2^{\kappa+n - \frac{16}{l}(\frac{n}{2}+2)}}{144e} \right\}.$$

Alternatively, let $q_2 = \min\left\{\frac{2^{2\kappa}}{432n}, \frac{2^{\kappa+n}}{144e}\right\}$. Then the second term is upper bounded by $2^{\frac{n}{2}+2-\frac{l}{8}}$, approaching zero as the length l increases.

PROOF OF LEMMA 5. Fix $\mathbf{s} = (s^i)_{1\leq i\leq q_1}$ and for $m = 0, \ldots, q_1$, define probability distributions π_m where for each $\mathbf{w} = (w^1, \ldots, w^{q_1}) \in \Omega$,

$$\pi_m(\mathbf{w}) = \mathbf{Pr}\Big[(u^{m+1}, \ldots, u^{q_1}) \xleftarrow{\$} (I_n \backslash \{s^1, \ldots, s^m\})^{*(q_1-m)}, \mathbf{k} \xleftarrow{\$} (I_\kappa)^{*d},$$
$$E \xleftarrow{\$} BC(\kappa, n) : \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (s^i, w^i)_{1\leq i\leq m} \wedge \mathsf{CE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i)_{m+1\leq i\leq q_1} \ \Big| E \vdash \mathcal{Q}_2\Big].$$

Then we can check that $\pi_0 = \nu$ and $\pi_{q_1} = \mu_{\mathbf{s}}$. Since

$$\|\mu_{\mathbf{s}} - \nu\| \leq \sum_{m=0}^{q_1-1} \|\pi_{m+1} - \pi_m\|, \tag{6}$$

we will focus on upper bounding $\|\pi_{m+1} - \pi_m\|$ for each $m = 0, \ldots, q_1 - 1$. In order to couple π_{m+1} and π_m, we will define a random variable (T, V) on $\Omega \times \Omega$ by the sampling process described in Figure 2. In this description,

$$\mathsf{D}(k) = \{x \in I_n : (x, k, y) \in \mathcal{Q}_2 \text{ for some } y\},$$
$$\mathsf{R}(k) = \{y \in I_n : (x, k, y) \in \mathcal{Q}_2 \text{ for some } x\},$$

for each key $k \in I_\kappa$. So they denote the domain points and the range points of the evaluations of $E(k, \cdot)$ determined by $\mathcal{Q}_2$, respectively.

In lines 1 to 4, the first $m+1$ elements are initialized. They are updated in lines 5 to 23 along cascade encryption. Specifically, the first m elements are faithfully updated in lines 7 to 11, while the $(m+1)$-th element is updated in lines 12 to 23 according to four conditions. The last $q_1 - m - 1$ elements of the output are determined in lines 24 to 29 without any update process.

As for this random variable, we point out some noteworthy properties.

1. In any case, the first m elements of T and V are equal.
2. If $t[d] = v[d]$, then $T = V$ at the end of the experiment.
3. By ignoring the steps used to sample V, we obtain the process for sampling T as described in Figure 3(a). Similarly, we obtain the process for sampling V as described in Figure 3(b). We can check that T and V follow probability distributions π_{m+1} and π_m, respectively.

Therefore by Lemma 1, we have

$$\|\pi_{m+1} - \pi_m\| \leq \mathbf{Pr}\left[T \neq V\right] = \mathbf{Pr}\left[t[d] \neq v[d]\right]. \tag{7}$$

1: **for** $i \leftarrow 1$ to m **do**
2: $\quad w^i[0] \leftarrow s^i$
3: $t[0] \leftarrow s^{m+1}$
4: $v[0] \xleftarrow{\$} I_n \backslash \{s^1, \ldots, s^m\}$
5: **for** $j \leftarrow 1$ to d **do**
6: $\quad k[j] \xleftarrow{\$} I_\kappa \backslash \{k[1], \ldots, k[j-1]\}$
7: $\quad$ **for** $i \leftarrow 1$ to m **do**
8: $\quad\quad$ **if** $w^i[j-1] \in \mathsf{D}(k[j])$ **then**
9: $\quad\quad\quad w^i[j] \leftarrow E(k[j], w^i[j-1])$
10: $\quad\quad$ **else if** $w^i[j-1] \notin \mathsf{D}(k[j])$ **then**
11: $\quad\quad\quad w^i[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^{i-1}[j]\} \cup \mathsf{R}(k[j]))$
12: $\quad$ **if** $t[j-1] \in \mathsf{D}(k[j])$ and $v[j-1] \in \mathsf{D}(k[j])$ **then**
13: $\quad\quad t[j] \leftarrow E(k[j], t[j-1])$
14: $\quad\quad v[j] \leftarrow E(k[j], v[j-1])$
15: $\quad$ **else if** $t[j-1] \in \mathsf{D}(k[j])$ and $v[j-1] \notin \mathsf{D}(k[j])$ **then**
16: $\quad\quad t[j] \leftarrow E(k[j], t[j-1])$
17: $\quad\quad v[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
18: $\quad$ **else if** $t[j-1] \notin \mathsf{D}(k[j])$ and $v[j-1] \in \mathsf{D}(k[j])$ **then**
19: $\quad\quad t[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
20: $\quad\quad v[j] \leftarrow E(k[j], v[j-1])$
21: $\quad$ **else if** $t[j-1] \notin \mathsf{D}(k[j])$ and $v[j-1] \notin \mathsf{D}(k[j])$ **then**
22: $\quad\quad t[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
23: $\quad\quad v[j] \leftarrow t[j]$
24: **if** $t[d] = v[d]$ **then**
25: $\quad (v^{m+2}, \ldots, v^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], v[d]\})^{*(q_1-m-1)}$
26: $\quad (t^{m+2}, \ldots, t^{q_1}) \leftarrow (v^{m+2}, \ldots, v^{q_1})$
27: **else**
28: $\quad (v^{m+2}, \ldots, v^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], v[d]\})^{*(q_1-m-1)}$
29: $\quad (t^{m+2}, \ldots, t^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], t[d]\})^{*(q_1-m-1)}$
30: $T \leftarrow (w^1[d], \ldots, w^m[d], t[d], t^{m+2}, \ldots, t^{q_1})$
31: $V \leftarrow (w^1[d], \ldots, w^m[d], v[d], v^{m+2}, \ldots, v^{q_1})$
32: **return** (T, V)

Fig. 2. Sampling process for random variable (T, V)

for $i \leftarrow 1$ to m **do**
    $w^i[0] \leftarrow s^i$
$t[0] \leftarrow s^{m+1}$
for $j \leftarrow 1$ to d **do**
    $k[j] \xleftarrow{\$} I_\kappa \backslash \{k[1], \ldots, k[j-1]\}$
    for $i \leftarrow 1$ to m **do**
        if $w^i[j-1] \in \mathsf{D}(k[j])$ **then**
            $w^i[j] \leftarrow E(k[j], w^i[j-1])$
        else if $w^i[j-1] \notin \mathsf{D}(k[j])$ **then**
            $w^i[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^{i-1}[j]\} \cup \mathsf{R}(k[j]))$
    if $t[j-1] \in \mathsf{D}(k[j])$ **then**
        $t[j] \leftarrow E(k[j], t[j-1])$
    else if $t[j-1] \notin \mathsf{D}(k[j])$ **then**
        $t[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
$(t^{m+2}, \ldots, t^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], t[d]\})^{*(q_1-m-1)}$
return $T = (w^1[d], \ldots, w^m[d], t[d], t^{m+2}, \ldots, t^{q_1})$

(a) Sampling T

for $i \leftarrow 1$ to m **do**
    $w^i[0] \leftarrow s^i$
$v[0] \xleftarrow{\$} I_n \backslash \{s^1, \ldots, s^m\}$
for $j \leftarrow 1$ to d **do**
    $k[j] \xleftarrow{\$} I_\kappa \backslash \{k[1], \ldots, k[j-1]\}$
    for $i \leftarrow 1$ to m **do**
        if $w^i[j-1] \in \mathsf{D}(k[j])$ **then**
            $w^i[j] \leftarrow E(k[j], w^i[j-1])$
        else if $w^i[j-1] \notin \mathsf{D}(k[j])$ **then**
            $w^i[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^{i-1}[j]\} \cup \mathsf{R}(k[j]))$
    if $v[j-1] \in \mathsf{D}(k[j])$ **then**
        $v[j] \leftarrow E(k[j], v[j-1])$
    else if $v[j-1] \notin \mathsf{D}(k[j])$ **then**
        $v[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
$(v^{m+2}, \ldots, v^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], v[d]\})^{*(q_1-m-1)}$
return $V = (w^1[d], \ldots, w^m[d], v[d], v^{m+2}, \ldots, v^{q_1})$

(b) Sampling V

Fig. 3. Sampling T and V separately

Since $t[j] = v[j]$ implies $t[j+2] = v[j+2]$ for $j = 0, \ldots, d-2$ (actually, $t[j'] = v[j']$ for every $j' > j$), we have

$$\mathbf{Pr}\left[t[d] \neq v[d]\right] \leq \prod_{h=1}^{\frac{d}{2}} \mathbf{Pr}\left[t[2h] \neq v[2h] \;\middle|\; t[2h-2] \neq v[2h-2]\right]. \tag{8}$$

For a fixed $h = 1, \ldots, \frac{d}{2}$, assume that $t[2h-2] \neq v[2h-2]$, and on this condition, consider the probability that $t[2h]$ and $v[2h]$ are different. In order for this event to happen, either $t[2h-2]$ or $v[2h-2]$ should map to a point within $\mathsf{D}(k[2h])$ since otherwise $t[2h-1]$ and $v[2h-1]$ both outside $\mathsf{D}(k[2h])$ would map to an identical point $t[2h] = v[2h]$. We divide this event into three subcases. In the following description, we fix a parameter $M > 0$, and call a key k *heavy* if $|\mathsf{R}(k)| = |\mathsf{D}(k)| > M$.

Case 1: Either $k[2h-1]$ or $k[2h]$ is heavy. Since there are at most q_2/M heavy keys and $k[2h-1]$ and $k[2h]$ are chosen from the set of size at least $2^\kappa - d$, the probability of this case is at most

$$\frac{2q_2}{M(2^\kappa - d)}. \tag{9}$$

Case 2: $k[2h]$ is not heavy, and either $(t[2h-2], k[2h-1], y) \in \mathcal{Q}_2$ or $(v[2h-2], k[2h-1], y) \in \mathcal{Q}_2$ for some $y \in \mathsf{D}(k[2h])$. First, assume that $k[2h]$ is not heavy. Since $|\mathsf{D}(k[2h])| \leq M$ and $\omega(\mathcal{Q}_2) \leq \beta$, the number of keys k such that either $(t[2h-2], k, y) \in \mathcal{Q}_2$ or $(v[2h-2], k, y) \in \mathcal{Q}_2$ for some $y \in \mathsf{D}(k[2h])$ is at most $2M\beta$. Therefore the probability that one of such keys is chosen as $k[2h-1]$ is at most

$$\frac{2M\beta}{2^\kappa - d}. \tag{10}$$

Case 3: The remaining case. Here we assume that any of $k[2h-1]$ and $k[2h]$ is not heavy. Furthermore, $k[2h-1]$ and $\mathcal{Q}_2$ do not determine a mapping from one of $t[2h-2]$ and $v[2h-2]$ to any point within $\mathsf{D}(k[2h])$. However either $t[2h-2]$ or $v[2h-2]$ might still go into $\mathsf{D}(k[2h])$ by probabilistic sampling. Since $|\mathsf{D}(k[2h])| \leq M$ and $|\mathsf{R}(k[2h-1])| \leq M$, this case occurs with probability at most

$$\frac{2M}{N-M}. \tag{11}$$

We notice that the update of $w^i[2h-2]$, $i = 1, \ldots, m$, does not affect this upper bounding. By (6), (7), (8), (9), (10) and (11), we obtain

$$\|\mu_{\mathbf{s}} - \nu\| \leq q_1 \left(\frac{2q_2}{M(2^\kappa - d)} + \frac{2M\beta}{2^\kappa - d} + \frac{2M}{N-M}\right)^{\frac{d}{2}}.$$

3.2 Security of Xor-cascade Encryption

In this section, we analyze the security of xor-cascade encryption XCE^l for even length $l = 2d$. The argument is very similar to the security proof of the original cascade encryption except modifying key sampling process and applying Lemma 6. First, the following original key sampling process **A** is modified into **B**:

A: Choose $\mathbf{k} \in I_\kappa^l$ and $\mathbf{z} \in I_n^{l+1}$ uniformly at random.

B: Randomly partition $T_1 \cup T_2 = I_\kappa$ so that $|T_1| = |T_2|$, choose $\mathbf{k}' \in (T_1)^{*d}$ and $\mathbf{k}'' \in (T_2)^{*d}$ uniformly at random, and then define $\mathbf{k} = (\mathbf{k}', \mathbf{k}'')$. Next, choose $\mathbf{z}' = (z'_0, \ldots, z'_d) \in I_n^{d+1}$ and $\mathbf{z}'' = (z''_0, \ldots, z''_d) \in I_n^{d+1}$ uniformly at random, and then define

$$\mathbf{z} = (z'_0, \ldots, z'_d \oplus z''_0, \ldots, z''_d) \in I_n^{l+1}.$$

One can distinguish sampling processes **A** and **B** with advantage at most

$$\binom{l}{2}\frac{1}{2^\kappa} \leq \frac{l^2}{2^{\kappa+1}}. \tag{12}$$

Taking into account (12), we analyze the security of

$$\mathsf{XCE}^l_{\mathbf{k},\mathbf{z}}[E] = \mathsf{XCE}^d_{\mathbf{k}'',\mathbf{z}''}[E] \circ \mathsf{XCE}^d_{\mathbf{k}',\mathbf{z}'}[E],$$

where $(\mathbf{k}, \mathbf{z})$, $(\mathbf{k}', \mathbf{z}')$ and $(\mathbf{k}'', \mathbf{z}'')$ are defined by key sampling process **B**.

For $\mathcal{Q}_1 = (u^i, v^i)_{1\leq i\leq q_1}$ and $\mathcal{Q}_2 = (x^i, k^i, y^i)_{1\leq i\leq q_2}$, we can prove

$$\mathsf{p}_1(\mathcal{Q}_1|\mathcal{Q}_2) = \frac{1}{\binom{2^\kappa}{2^{\kappa-1}}} \times$$

$$\sum_{\substack{T_1\cup T_2 = I_\kappa \\ |T_1|=|T_2|=2^{\kappa-1}}} \sum_{\mathbf{w}\in\Omega} \left(\mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} (T_1)^{*d}, \mathbf{z} \xleftarrow{\$} I_n^{d+1}, E \xleftarrow{\$} BC(T_1, n) : \mathsf{XCE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i) \;\middle|\; E \vdash \mathcal{Q}'_2\right] \right.$$

$$\left. \times \mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} (T_2)^{*d}, \mathbf{z} \xleftarrow{\$} I_n^{d+1}, E \xleftarrow{\$} BC(T_2, n) : \mathsf{XCE}^d_{\mathbf{k}}[E] \vdash (w^i, v^i) \;\middle|\; E \vdash \mathcal{Q}''_2\right] \right),$$

with the same notations as the previous section. In order to estimate the probabilities appearing as the summands, we fix a query history $\mathcal{Q}_2 = (x^i, k^i, y^i)_{1\leq i\leq q'}$ such that $q' \leq q_2$, and for each $\mathbf{s} \in \Omega$ define a probability distribution $\mu_\mathbf{s}$ such that for each $\mathbf{w} = (w^i)_{1\leq i\leq q_1} \in \Omega$,

$$\mu_\mathbf{s}(\mathbf{w}) = \mathbf{Pr}\Big[\mathbf{k} \xleftarrow{\$} (I_\kappa)^{*d}, \mathbf{z} \xleftarrow{\$} I_n^{d+1}, E \xleftarrow{\$} BC(\kappa, n) :$$

$$\mathsf{XCE}^d_{\mathbf{k}}[E] \vdash (s^i, w^i)_{1\leq i\leq q_1} \;\Big|\; E \vdash \mathcal{Q}_2\Big].$$

Then we have the following lemma.

Lemma 6. *Let $d > 0$, let $\mu_\mathbf{s}$ be the probability distribution defined as above, and let ν be the uniform probability distribution on Ω. Then for $M > 0$, we have $\|\mu_\mathbf{s} - \nu\| \leq \epsilon$, where*

$$\epsilon = q_1 \left(\frac{q_2}{M(2^\kappa - d)} + \frac{2M}{N}\right)^d.$$

The proof will be given at Appendix E. Using this lemma and exactly the same argument for the original cascade encryption, we can also prove the following theorem.

Theorem 2. *Let XCE^l be an l-xor-cascade encryption scheme using a κ-bit key n-bit block-cipher. If $l = 2d$, then for $M > 0$,*

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{XCE}^l}(q_1, q_2) \leq \frac{l^2}{2^{\kappa+1}} + 4q_1^{\frac{1}{2}} \left(\frac{q_2}{M(2^{\kappa-1} - d)} + \frac{2M}{N}\right)^{\frac{d}{2}}.$$

Optimizing Parameters. By solving $\frac{q_2}{M(2^{\kappa-1}-d)} = \frac{2M}{N}$, we set $M = \sqrt{\frac{Nq_2}{2^\kappa - l}}$, obtaining the following corollary.

Corollary 2. *Let* XCE^l *be an* l*-cascade encryption scheme using a* κ*-bit key* n*-bit blockcipher. If* l *is even, then*

$$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{XCE}^l}(q_1, q_2) \leq \frac{l^2}{2^{\kappa+1}} + 4q_1^{\frac{1}{2}} \left(\frac{16q_2}{N(2^\kappa - l)} \right)^{\frac{l}{8}}.$$

Interpretation. Assuming that $l^2/2^{\kappa+1}$ is negligible, set $q_1 = 2^n$ and approximate $2^\kappa - l \approx 2^\kappa$. Then the distinguishing advantage becomes negligible when

$$q_2 \ll 2^{\kappa+n-4-\frac{8}{l}(\frac{n}{2}+2)}.$$

Alternatively, let $q_2 = 2^{\kappa+n-5}$. Then we can check that $\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{CE}^l}(2^n, 2^{\kappa+n-5})$ approaches zero as the length l increases (up to the condition that $l^2/2^{\kappa+1}$ is negligible).

References

1. FIPS PUB 46: Data Encryption Standard (DES). National Institute of Standards and Technology (1977)
2. ANSI X9.52: Triple Data Encryption Algorithm Modes of Operation (1998)
3. FIPS PUB 46-3: Data Encryption Standard (DES). National Institute of Standards and Technology (1999)
4. FIPS PUB 197: Advanced Encryption Standard (AES). National Institute of Standards and Technology (2001)
5. NIST ST 800-67: Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher. National Institute of Standards and Technology (2004)
6. M. Bellare and P. Rogaway: The Security of Triple Encryption and a Framework for Code-Based Game-Playing Proofs. Eurocrypt 2006, LNCS 4004, pp. 409–426, Springer, Heidelberg (2006)
7. A. Bogdanov, L. R. Knudsen, G. Leander, F. Standaert, J. Steinberger, and E. Tischhauser: Key-Alternating Ciphers in a Provable Setting: Encryption Using a Small Number of Public Permutations. Eurocrypt 2012, LNCS 7237, pp. 45–62, Springer, Heidelberg (2012)
8. P. Gaži: Plain versus Randomized Cascading-Based Key-Length Extension for Block Ciphers. Cryptology ePrint Archive, Report 2013/019, 2013. Available at http://eprint.iacr.org/2013/019
9. P. Gaži and U. Maurer: Cascade Encryption Revisited. Asiacrypt 2009, LNCS 5912, pp. 37–51, Springer, Heidelberg (2009)
10. P. Gaži and S. Tessaro: Efficient and Optimally Secure Key-Length Extension for Block Ciphers via Randomized Cascading. Eurocrypt 2012, LNCS 7237, pp. 63–80, Springer, Heidelberg (2012)
11. V. T. Hoang, B. Morris and P. Rogaway: An Enciphering Scheme Based on a Card Shuffle. CRYPTO 2012, LNCS 7417, pp. 7–13, Springer, Heidelberg (2012)
12. V. T. Hoang and P. Rogaway: On Generalized Feistel Networks. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 613630. Springer, Heidelberg (2010)
13. J. Kilian and P. Rogaway: How to Protect DES Against Exhaustive Key Search (an Analysis of DESX). Journal of Cryptology 14, pp. 1735. Springer, Heidelberg (2001)
14. R. Lampe, J. Patarin and Y. Seurin: An Asymptotically Tight Security Analysis of the Interated Even-Mansour Cipher. Asiacrypt 2012, To appear.
15. U. Maurer, K. Pietrzak and R. Renner: Indistinguishability Amplification. CRYPTO 2007, LNCS 4622, pp. 130-149. Springer, Heidelberg (2007)
16. I. Mironov: (Not So) Random Shuffles of RC4. CRYPTO 2002, LNCS 2442, pp. 304-319. Springer, Heidelberg (2002)
17. B. Morris, P. Rogaway and T. Stegers: How to Encipher Messages on a Small Domain: Deterministic Encryption and the Thorp Shuffle. CRYPTO 2009. LNCS 5677, pp. 286–302, Springer, Heidelberg (2009)

A A Brute-force Attack of $2^{\kappa+n}$ Query Complexity

In this section, we describe a standard information theoretic brute-force attack against a λ-bit key m-bit encryption scheme C that makes a certain number of calls to a κ-bit key n-bit blockcipher E.[3] An adversary $\mathcal{A}$ executes the following steps.

1. $\mathcal{A}$ makes all possible $2^{\kappa+n}$ queries to the underlying blockcipher E.
2. $\mathcal{A}$ makes t nonadaptive forward queries to the outer permutation, recording query history $\mathcal{Q} = (u^i, v^i)_{1 \le i \le t}$.
3. If there is a λ-bit key $\mathbf{k}$ such that $\mathsf{C}_{\mathbf{k}}[E](u^i) = v^i$ for every $i = 1, \ldots, t$, then $\mathcal{A}$ outputs 0. Otherwise, $\mathcal{A}$ outputs 1.

Since we have

$$\mathbf{Pr}\left[\mathbf{k} \xleftarrow{\$} \{0,1\}^\lambda, E \xleftarrow{\$} BC(\kappa, n) : \mathcal{A}[\mathsf{C}_{\mathbf{k}}[E], E] = 0\right] = 1,$$
$$\mathbf{Pr}\left[P \xleftarrow{\$} \mathcal{P}_n, E \xleftarrow{\$} BC(\kappa, n) : \mathcal{A}[P, E] = 0\right] \le \frac{2^\lambda}{(2^m)_t},$$

$\mathbf{Adv}^{\mathsf{PRP}}_{\mathsf{C}}(\mathcal{A})$ gets close to 1 as $t \gg \frac{\lambda}{m}$.

B Proof of Lemma 1 and Lemma 2

Proof of Lemma 1. Let λ be a coupling of μ and ν and let $(X, Y) \sim \lambda$. By definition, for any $z \in \Omega$, $\lambda(z, z) \le \min\{\mu(z), \nu(z)\}$. Since $\mathbf{Pr}[X = Y] = \sum_{z \in \Omega} \lambda(z, z)$, we have

$$\mathbf{Pr}[X = Y] \le \sum_{z \in \Omega} \min\{\mu(z), \nu(z)\}.$$

Therefore we have

$$\begin{aligned}\mathbf{Pr}[X \neq Y] &\ge 1 - \sum_{z \in \Omega} \min\{\mu(z), \nu(z)\} = \sum_{z \in \Omega} (\mu(z) - \min\{\mu(z), \nu(z)\}) \\ &= \sum_{\substack{z \in \Omega \\ \mu(z) \ge \nu(z)}} (\mu(z) - \nu(z)) = \max_{S \subset \Omega} \{\mu(S) - \nu(S)\} = \|\mu - \nu\|.\end{aligned}$$

Proof of Lemma 2. Let $S = \{x \in \Omega : \mu(x) \ge (1 - \sqrt{\epsilon})\nu(x)\}$. By definition, any element of S satisfies the second condition. Contary to the first condition, suppose that $|S| < (1 - \sqrt{\epsilon})|\Omega|$. This implies $\nu(\Omega \backslash S) = 1 - |S|/|\Omega| > \sqrt{\epsilon}$, and hence

$$\nu(\Omega \backslash S) - \mu(\Omega \backslash S) \ge \nu(\Omega \backslash S) - \left(1 - \sqrt{\epsilon}\right) \nu(\Omega \backslash S) = \sqrt{\epsilon}\nu(\Omega \backslash S) > \left(\sqrt{\epsilon}\right)^2 = \epsilon.$$

This is a contradiction to $\|\mu - \nu\| \le \epsilon$.

[3] The output size m of C might be different from the block size n of E.

C Proof of Inequality (2)

Fix x, $y \in I_n$. For any $\beta > 0$,

$$\mathbf{Pr}\left[E \xleftarrow{\$} BC(\kappa, n) : |\{k : E(k, x) = y\}| \geq \beta\right] \leq \binom{2^\kappa}{\beta}\left(\frac{1}{2^n}\right)^\beta \leq \left(\frac{e2^\kappa}{\beta 2^n}\right)^\beta.$$

Therefore for any $\mathcal{Q}_2$ (which might be the maximum query history of size $2^{\kappa+n}$ including all the evaluations of E), $\omega(\mathcal{Q}_2)$ is smaller than β except with probability

$$\mathsf{P} = 2^{2n}\left(\frac{e2^\kappa}{\beta 2^n}\right)^\beta,$$

where $\mathsf{P} \leq 2^{2n-\beta}$ if $\beta \geq e2^{\kappa-n+1}$.

D Equivalence of Key Sampling Processes B and C for CE^l

Fix a key $\mathbf{k} = (\mathbf{k}', \mathbf{k}'') \in (I_\kappa)^{*l}$, where $\mathbf{k}' = (k'_1, \ldots, k'_d)$ and $\mathbf{k}'' = (k''_1, \ldots, k''_d)$. Then the number of partitions (T_1, T_2) such that $\{k'_1, \ldots, k'_d\} \subset T_1$ and $\{k''_1, \ldots, k''_d\} \subset T_2$ is $\binom{2K-2d}{K-d}$, where $K = 2^{\kappa-1}$. For each (T_1, T_2), key sampling process $\mathbf{C}$ chooses $\mathbf{k}'$ and $\mathbf{k}''$ from T_1 and T_2, respectively, with probability $(1/(K)_d)^2$. So the probability that $\mathbf{C}$ chooses $\mathbf{k}'$ and $\mathbf{k}''$ is

$$\frac{\binom{2K-2d}{K-d}}{\binom{2K}{K}}\left(\frac{1}{(K)_d}\right)^2 = \frac{(2K-2d)!(K!)^2}{(2K)!((K-d)!)^2}\cdot\left(\frac{(K-d)!}{K!}\right)^2 = \frac{(2K-2d)!}{(2K)!} = \frac{1}{(2^\kappa)_{2d}},$$

which is the same as the probability that key sampling process $\mathbf{B}$ chooses $\mathbf{k} = (\mathbf{k}', \mathbf{k}'')$.

E Proof of Lemma 6

Fix $\mathbf{s} = (s^i)_{1\leq i\leq q_1}$ and for $m = 0, \ldots, q_1$, define probability distributions π_m where for each $\mathbf{w} = (w^1, \ldots, w^{q_1}) \in \Omega$,

$$\pi_m(\mathbf{w}) = \mathbf{Pr}\Big[(u^{m+1}, \ldots, u^{q_1}) \xleftarrow{\$} (I_n \backslash \{s^1, \ldots, s^m\})^{*(q_1-m)}, \mathbf{k} \xleftarrow{\$} (I_\kappa)^{*d}, \mathbf{z} \xleftarrow{\$} I_n^{d+1},$$
$$E \xleftarrow{\$} BC(\kappa, n) : \mathsf{XCE}^d_{\mathbf{k}}[E] \vdash (s^i, w^i)_{1\leq i\leq m} \wedge \mathsf{XCE}^d_{\mathbf{k}}[E] \vdash (u^i, w^i)_{m+1\leq i\leq q_1} \;\Big|\; E \vdash \mathcal{Q}_2\Big].$$

Then we can check that $\pi_0 = \nu$ and $\pi_{q_1} = \mu_{\mathbf{s}}$. Since

$$\|\mu_{\mathbf{s}} - \nu\| \leq \sum_{m=0}^{q_1-1} \|\pi_{m+1} - \pi_m\|, \tag{13}$$

we focus on upper bounding $\|\pi_{m+1} - \pi_m\|$ for each $m = 0, \ldots, q_1 - 1$. In order to couple π_{m+1} and π_m, we will define a random variable (T, V) on $\Omega \times \Omega$ by the sampling process described in Figure 4. Then we can check that their marginal distributions are π_{m+1} and π_m, and

$$\|\pi_{m+1} - \pi_m\| \leq \mathbf{Pr}\left[T \neq V\right] = \mathbf{Pr}\left[t[d] \neq v[d]\right]. \tag{14}$$

for $i \leftarrow 1$ to m **do**
  $w^i[0] \leftarrow s^i$
$t[0] \leftarrow s^{m+1}$
$v[0] \xleftarrow{\$} I_n \backslash \{s^1, \ldots, s^m\}$
for $j \leftarrow 1$ to d **do**
  $k[j] \xleftarrow{\$} I_\kappa \backslash \{k[1], \ldots, k[j-1]\}$
  $z[j-1] \xleftarrow{\$} I_n$
  for $i \leftarrow 1$ to m **do**
    $w^i[j-1] \leftarrow w^i[j-1] \oplus z[j-1]$
    if $w^i[j-1] \in \mathsf{D}(k[j])$ **then**
      $w^i[j] \leftarrow E(k[j], w^i[j-1])$
    else if $w^i[j-1] \notin \mathsf{D}(k[j])$ **then**
      $w^i[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^{i-1}[j]\} \cup \mathsf{R}(k[j]))$
  if $t[j-1] \oplus z[j-1] \in \mathsf{D}(k[j])$ and $v[j-1] \oplus z[j-1] \in \mathsf{D}(k[j])$ **then**
    $t[j] \leftarrow E(k[j], t[j-1] \oplus z[j-1])$
    $v[j] \leftarrow E(k[j], v[j-1] \oplus z[j-1])$
  else if $t[j-1] \oplus z[j-1] \in \mathsf{D}(k[j])$ and $v[j-1] \oplus z[j-1] \notin \mathsf{D}(k[j])$ **then**
    $t[j] \leftarrow E(k[j], t[j-1] \oplus z[j-1])$
    $v[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
  else if $t[j-1] \oplus z[j-1] \notin \mathsf{D}(k[j])$ and $v[j-1] \oplus z[j-1] \in \mathsf{D}(k[j])$ **then**
    $t[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
    $v[j] \leftarrow E(k[j], v[j-1] \oplus z[j-1])$
  else if $t[j-1] \oplus z[j-1] \notin \mathsf{D}(k[j])$ and $v[j-1] \oplus z[j-1] \notin \mathsf{D}(k[j])$ **then**
    $t[j] \xleftarrow{\$} I_n \backslash (\{w^1[j], \ldots, w^m[j]\} \cup \mathsf{R}(k[j]))$
    $v[j] \leftarrow t[j]$
$z[d] \xleftarrow{\$} I_n$
if $t[d] = v[d]$ **then**
  $(v^{m+2}, \ldots, v^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], v[d]\})^{*(q_1-m-1)}$
  $(t^{m+2}, \ldots, t^{q_1}) \leftarrow (v^{m+2}, \ldots, v^{q_1})$
else
  $(v^{m+2}, \ldots, v^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], v[d]\})^{*(q_1-m-1)}$
  $(t^{m+2}, \ldots, t^{q_1}) \xleftarrow{\$} (I_n \backslash \{w^1[d], \ldots, w^m[d], t[d]\})^{*(q_1-m-1)}$
$T \leftarrow (w^1[d] \oplus z[d], \ldots, w^m[d] \oplus z[d], t[d] \oplus z[d], t^{m+2}, \ldots, t^{q_1})$
$V \leftarrow (w^1[d] \oplus z[d], \ldots, w^m[d] \oplus z[d], v[d] \oplus z[d], v^{m+2}, \ldots, v^{q_1})$
return (T, V)

Fig. 4. Sampling process for random variable (T, V)

Since $t[j] = v[j]$ implies $t[j+1] = v[j+1]$ for $j = 0, \ldots, l-1$, we have

$$\mathbf{Pr}\left[t[d] \neq v[d]\right] \leq \prod_{j=1}^{d} \mathbf{Pr}\left[t[j] \neq v[j] \;\middle|\; t[j-1] \neq v[j-1]\right]. \tag{15}$$

In order to upper bound $\mathbf{Pr}\left[t[j] \neq v[j] \;\middle|\; t[j-1] \neq v[j-1]\right]$ for each j, we first choose $k[j]$ from the set of size at least $2^\kappa - d$. For a parameter $M > 0$, there are at most q_2/M heavy keys k such that

$$|\mathsf{R}(k)| = |\mathsf{D}(k)| > M.$$

Therefore the probability that $k[j]$ is heavy is at most

$$\frac{q_2}{M(2^\kappa - d)}. \tag{16}$$

Conditioned on the case that $k[j]$ is not heavy, either $t[j-1] \oplus z[j-1]$ or $v[j-1] \oplus z[j-1]$ should map to a point within $\mathsf{D}(k[j])$ since otherwise $t[j] \oplus z[j-1]$ and $v[j] \oplus z[j-1]$ both outside $\mathsf{D}(k[j])$ would map to an identical point $t[j+1] = v[j+1]$. The probability of this event over the random choice of $z[j-1]$ is at most

$$\frac{2M}{N}. \tag{17}$$

Then by (13), (14), (15), (16) and (17), we obtain

$$\|\mu_{\mathbf{s}} - \nu\| \leq q_1 \left(\frac{q_2}{M(2^\kappa - d)} + \frac{2M}{N}\right)^d.$$