Triple and Quadruple Encryption: Bridging the Gaps

Bart Mennink and Bart Preneel

Dept. Electrical Engineering, ESAT/COSIC, KU Leuven, and iMinds, Belgium
bart.mennink@esat.kuleuven.be, bart.preneel@esat.kuleuven.be

Abstract. Triple encryption is a cascade of three block cipher evaluations with independent keys, in order to enlarge its key size. This design is proven secure up to approximately $2^{\kappa+\min\{\kappa/2,n/2\}}$ queries (by Bellare and Rogaway, EUROCRYPT 2006, and Gaži and Maurer, ASIACRYPT 2009), where κ denotes the key size and n the block length of the underlying block cipher. On the other hand, the best known attack requires about $2^{\kappa+n/2}$ queries (by Lucks, FSE 1998, and Gaži, CRYPTO 2013). These bounds are non-tight for $\kappa \leq n$.
In this work, we close this gap. By strengthening the best known attack as well as tightening the security bound, we prove that triple encryption is tightly secure up to $2^{\kappa+\min\{\kappa,n/2\}}$ queries. Additionally, we prove that the same tight security bound holds for quadruple encryption (which consists of four sequentially evaluated block ciphers), and derive improved security and attack bounds for cascades consisting of five or more rounds. This work particularly solves the longstanding open problem of proving tight security of the well-known Triple-DES construction in the ideal model.

Keywords. Cascade encryption; Indistinguishability; Tight; Triple-DES.

1 Introduction

From 1977 to the late 1990s, the Data Encryption Standard DES : $\{0,1\}^{56} \times \{0,1\}^{64} \rightarrow \{0,1\}^{64}$ [14] was the dominant algorithm for encryption. During its lifetime DES managed to withstand a wide range of attacks, such as the differential cryptanalysis of Biham and Shamir [4–6], the linear cryptanalysis of Matsui [26, 27], and others [8, 23], but in the end it was the simple brute force attack DES fell victim to: due to advances in computer technology, keys of 56 bits had become too small to guarantee security. Already since the introduction of DES, concerns about the short key size existed and research had been done to "artificially" extend the key of DES. This approach is called key length extension, and we can identify two popular approaches in this direction: *cascade encryption* (such as the Triple-DES construction [2, 15, 20, 32]) and *XOR-cascade encryption* (DESX being the most prominent example). In this work, we will focus on cascade encryption (the state of the art on XOR-cascading is discussed at the end of this section).

Cascade Encryption

Let $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher with a κ-bit key and an n-bit state. Informally, the idea of cascade encryption is to evaluate r block cipher calls sequentially, for $r \geq 1$, with usually different keys. It is well-known that a cascade of length two only offers a marginal security increase over E, due to the meet-in-the-middle attack of Diffie and Hellman [10] (although it is shown that in a different model a security increase can be achieved [1]). As such, the shortest length of a "meaningful" cascade is three. Triple encryption $\mathcal{E}^3 : \{0,1\}^{3\kappa} \times \{0,1\}^n \to \{0,1\}^n$ takes as input a key $(k_1, k_2, k_3) \in \{0,1\}^{3\kappa}$ and maps an n-bit message m to

$$\mathcal{E}^3_{k_1,k_2,k_3}(m) = E_{k_3} \circ E_{k_2} \circ E_{k_1}(m)\,. \tag{1}$$

This function has gained popularity and found widespread usage as the *Triple-DES* construction [2,15,20,32],[1] and even despite the presence of better block ciphers such as AES [7], Triple-DES still remains popular, particularly due to its short block size of 64 bits. For instance, Triple-DES is used in ATMs, the EMV standard [12], TLS 1.0 [9], and in Microsoft Outlook 2007. In general, the triple data encryption algorithm finds over 1600 by NIST validated implementations worldwide [33].

Assuming ideality of the underlying block cipher E (we discuss the state of the art for different security models later in the intro), Bellare and Rogaway [3] proved that triple encryption is secure up to $2^{\kappa+\min\{\kappa/2,n/2\}}$. This bound was later confirmed by Gaži and Maurer [17] who spotted a couple of bugs in the proof. On the other hand, Lucks [25] presented an attack on Triple-DES in approximately 2^{90} queries, an attack which got generalized by Gaži [16] to an attack on triple encryption in approximately $2^{\kappa+n/2}$ queries.

Generalizing triple encryption, cascade encryption is a construction $\mathcal{E}^r : \{0,1\}^{r\kappa} \times \{0,1\}^n \to \{0,1\}^n$ that consists of evaluating r block cipher calls sequentially, for $r \geq 1$:

$$\mathcal{E}^r_{k_1,\ldots,k_r}(m) = E_{k_r} \circ \cdots \circ E_{k_1}(m)\,. \tag{2}$$

The function $\mathcal{E}^r$ is depicted in Figure 1. Throughout, we denote $r' = \lceil r/2 \rceil$. Gaži and Maurer [17] proved that $\mathcal{E}^r$ is secure up to at least $2^{\kappa+\min\left\{\frac{(r'-1)}{r'}\kappa,n/2\right\}}$ queries, approaching $2^{\kappa+\min\{\kappa,n/2\}}$ for increasing r. Recently, Lee [24] proved security up to $2^{\kappa+\min\{\kappa,n\}-\frac{16}{r}(\frac{n}{2}+2)}$ queries for r a multiple of 4. This bound implies that the security approaches

[1] For compatibility reasons, Triple-DES is originally defined as $E_{k_3} \circ E^{-1}_{k_2} \circ E_{k_1}(m)$, where E is the block cipher DES. All findings in this work, however, apply to both cases.

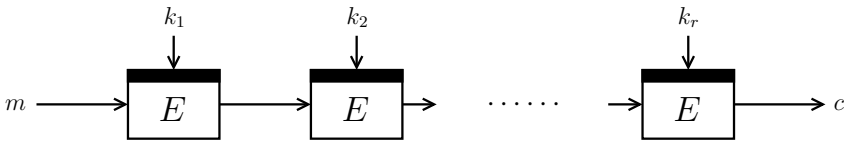

Fig. 1: Cascade encryption $\mathcal{E}^r$, for $r \geq 1$.

$2^{\kappa+\min\{\kappa,n\}}$ for increasing size of the cascade, but it only improves over [17] for relatively large values of r. On the other hand, Gaži [16] generalized the attack of Lucks to arbitrarily-sized cascades, leading to a security upper bound of $2^{\kappa+\frac{r'-1}{r'}n}$. These results are summarized in Table 1.

Our Contributions

For triple and quadruple encryption (for which the current security bounds are the same), the known bounds are only tight for $\kappa \geq n$ and a gap remains for $\kappa \leq n$. Remarkably, as DES has a key length of 56 bits and a message space of 64 bits, this means there is a non-trivial gap in the security bounds for the famous Triple-DES construction. The primary goal of this work is to close this gap. The first question to investigate is which of the bounds is "the tight one:" the security bound $2^{\kappa+\min\{\kappa/2,n/2\}}$ or the attack bound $2^{\kappa+n/2}$? As it turns out, both of them can be improved.

Improving Best Known Attacks. Consider the following pathological example. Let $\kappa = 0$, hence E_k is one and the same permutation π for all k, and $\mathcal{E}^3_{k_1,k_2,k_3} = \pi \circ \pi \circ \pi(m)$. On the one hand, $\mathcal{E}^3$ is trivially distinguishable from a random permutation, but on the other hand the best known attack claims the distinguisher needs to make at most $2^{n/2}$ queries. This demonstrates the presence of an attack independent of the block length n. In Section 3, we generalize and formalize this attack and find that $\mathcal{E}^r$, for $r \geq 1$, can be distinguished from an ideal permutation in $2^{r'\kappa}$ queries. Together with the attacks of Lucks [25] and Gaži [16], this result implies that an attack on $\mathcal{E}^r$ can be mounted in at most $2^{\kappa+\frac{r'-1}{r'}\min\{r'\kappa,n\}}$ queries, where we recall $r' = \lceil r/2 \rceil$. For $r = 3, 4$, these bounds read $\kappa + \min\{\kappa, n/2\}$. The attack can be considered as an extension of the meet-in-the-middle attack [10]. It borrows some ideas from Dinur et al. [11], who derive an attack with complexity about $2^{(r-\sqrt{2r})\kappa}$, but it is in an entirely different model and thus incomparable.

Tightening Security Bounds. For $r = 3, 4$, the newly obtained attack bound does not yet meet the security bound $2^{\kappa+\min\{\kappa/2,n/2\}}$ from [3,17].

Table 1: Security lower and upper bounds for cascaded encryption (in $\log_2$). Here, $r' = \lceil r/2 \rceil$. All results in **bold** are derived in this work.

$\mathcal{E}^r$	security	attack	tight
$r = 1, 2$	κ	κ [10]	✓
$r = 3, 4$	$\kappa + \min\{\kappa/2, n/2\}$ [3, 17]	$\kappa + n/2$ [16, 25]	✗
	$\boldsymbol{\kappa + \min\{\kappa, n/2\}}$	$\boldsymbol{\kappa + \min\{\kappa, n/2\}}$	✓
$r \geq 5^\star$	$\kappa + \min\left\{\frac{(r'-1)}{r'}\kappa, n/2\right\}$ [17]	$\kappa + \frac{r'-1}{r'}n$ [16]	✗
	$\boldsymbol{\kappa + \min\{\kappa, n/2\}}$	$\boldsymbol{\kappa + \frac{r'-1}{r'}\min\{r'\kappa, n\}}$	✗

* Starting from $r \geq 16$ and depending on the choice of κ, n, Lee [24] proved that if r is a multiple of 4, $\mathcal{E}^r$ is secure up to $2^{\kappa+\min\{\kappa,n\}-\frac{16}{r}(\frac{n}{2}+2)}$ queries.

As second contribution, we strengthen this security bound to achieve sharp security for $r = 3$ and $r = 4$. This security result is presented in Section 4. Informally and at a very high level, the improvement we make is as follows.

It turns out that one isolated lemma in [3, 17] is responsible for the current gap (Lemma 4 in both works). This lemma bounds the number of evaluations $E_{k''} \circ E_{k'} \circ E_k(m)$ for arbitrary m, k, k', k'' a distinguisher finds when making q queries to its ideal cipher E. While both [3, 17] bound this number of 3-paths by first fixing the first and last block cipher evaluation, and then smartly bounding the number of middle evaluations, we follow a different approach: we start from the middle, making a distinction between forward and inverse queries, and compute the number of 3-paths accordingly. (The formal argument is more delicate.)

Discussion. The freshly obtained security lower bound $2^{\kappa+\min\{\kappa,n/2\}}$ for $\mathcal{E}^3$ naturally carries over to $\mathcal{E}^r$ for $r \geq 3$, and this particularly implies also tight security for quadruple encryption. These findings are included in Table 1. Our bound improves over the previously known bounds of [3,17,24] for relatively small numbers of r. Particularly, Gaži and Maurer [17] claimed that $\mathcal{E}^r$ will achieve asymptotic security up to $2^{\kappa+\min\{\kappa,n/2\}}$ for increasing r. Our proof shows that this bound is already achieved for $r = 3$. For $r \geq 16$ and specific values of κ, n, the bound of Lee [24] is still better. A further discussion of the results is given in the conclusions in Section 5.

Other models

In line with the works of [3, 16, 17, 24], the focus of this work is on the security of cascade encryption in the ideal model, where the underlying block cipher E is considered perfect. In this setting, the distinguisher can make block cipher queries and queries to the construction: $\mathcal{E}^r$ or a random permutation $\mathcal{P}$.

Cascade encryption has also been analyzed in different settings, beyond the ideal cipher model. For instance, Even and Goldreich [13] demonstrate that cascade encryption is at least as strong as the strongest of the underlying ciphers, and Massey and Maurer [28] prove that, in a generalized attack model, the cascading is at least as strong as the first one. Interesting results on amplified security in the information-theoretic model via composition are by Gaži, Maurer, Pietrzak, and Renner [18, 29, 30], and Vaudenay [36]. Maurer and Tessaro [31, 35] consider security of cascades based on the PRP security of the underlying cipher. In a computational setting, Dinur et al. [11] generalized [10, 34] to derive a time memory tradeoff for cascade encryption.

These works, however, employ different models than ours, and hence the results are incomparable. Nevertheless, in [11], Dinur et al. report in their CRYPTO 2012 best paper award winning paper: "the exact security of double and triple encryption are well understood and we cannot push their analysis any further." In this work, we prove that this is not entirely accurate for the security in the ideal cipher model.

Further Related Work

Related to the plain cascading discussed in this paper is the idea of XOR-cascade encryption. For $r \geq 1$, XOR-cascade encryption $\mathcal{XE}^r : \left(\{0,1\}^{r\kappa} \times \{0,1\}^{(r+1)n}\right) \times \{0,1\}^n \to \{0,1\}^n$ takes as input keys $k = (k_1, \ldots, k_r) \in \{0,1\}^{r\kappa}$ and $z = (z_0, \ldots, z_r) \in \{0,1\}^{(r+1)n}$ and maps an n-bit message m to

$$\mathcal{XE}^r_{k,z}(m) = \oplus_{z_r} \circ E_{k_r} \circ \cdots \circ E_{k_1} \circ \oplus_{z_0}(m)\,,$$

where $\oplus_y(x) = x \oplus y$. For $r = 1$, XOR-cascading boils down to the DESX construction (accredited to Rivest), which has been proven secure up to $2^{\kappa/2+n/2}$ queries by Kilian and Rogaway [22]. An extension of DESX, with some refinements, has been analyzed by Gaži and Tessaro [19], proving tight security up to $2^{\kappa+n/2}$ queries. For general even $r \geq 2$, Lee [24] proved $\mathcal{XE}^r$ secure up to $2^{\kappa+n-\frac{8}{r}(\frac{n}{2}+2)}$, a bound approaching the optimal $2^{\kappa+n}$ for increasing r. Gaži [16] subsequently improved these results, showing that $\mathcal{XE}^r$ achieves security up to $2^{\kappa+\frac{r-1}{r}n}$ queries for $r = 3, 4$ and up to $2^{\kappa+\frac{r'-1}{r'}n}$ queries for $r \geq 5$. Next, Gaži also describes a distinguishing attack on $\mathcal{XE}^r$ in $2^{\kappa+\frac{r-1}{r}n}$ queries.

2 Security Model

For an integer $n \in \mathbb{N}$, $\{0,1\}^n$ denotes the set of bit strings of length n. If $\mathcal{X}$ is a set, we denote by $x \xleftarrow{\$} \mathcal{X}$ the uniformly random drawing of x from $\mathcal{X}$. For $x, y \in \mathbb{N}$, $x^{\underline{y}}$ denotes the falling factorial power $x(x-1)\cdots(x-y+1)$.

Block Ciphers. For integral $\kappa, n \geq 1$, a block cipher $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ is a mapping such that for every key $k \in \{0,1\}^\kappa$, $E_k(\cdot) = E(k,\cdot)$ is a permutation on $\{0,1\}^n$. We denote by $\mathsf{Block}(\kappa, n)$ the set of all such block ciphers, and by $\mathsf{Perm}(n)$ the set of all n-bit permutations.

Cascade Encryption. Let $\kappa, n, r \geq 1$ be integral, $E \in \mathsf{Block}(\kappa, n)$, and consider $\mathcal{E}^r : \{0,1\}^{r\kappa} \times \{0,1\}^n \to \{0,1\}^n$ based on E (see eqn. (2)). We consider a distinguisher that tries to tell apart $\mathcal{E}^r$ from a perfectly random permutation $\mathcal{P} \xleftarrow{\$} \mathsf{Perm}(n)$. We consider the ideal cipher model, where we assume that E is randomly drawn from $\mathsf{Block}(\kappa, n)$, and give the distinguisher query access to both E and its inverse E^{-1}. Formally, the advantage of a distinguisher $\mathcal{D}$ in attacking $\mathcal{E}^r$ is defined as

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^r}(\mathcal{D}) = \mathbf{Pr}\left(\mathcal{P} \xleftarrow{\$} \mathsf{Perm}(n), E \xleftarrow{\$} \mathsf{Block}(\kappa, n) \ : \ \mathcal{D}^{\pm\mathcal{P},\pm E} = 1\right) - \mathbf{Pr}\left(k \xleftarrow{\$} \{0,1\}^{r\kappa}, E \xleftarrow{\$} \mathsf{Block}(\kappa, n) \ : \ \mathcal{D}^{\pm\mathcal{E}^r_k,\pm E} = 1\right).$$

For $q \in \mathbb{N}$, we define by $\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^r}(q)$ the maximum advantage over all distinguishers making at most q queries to its oracles (the outer permutation $\mathcal{E}^r_k$ or $\mathcal{P}$, and the block cipher E).

3 Improved Distinguishing Attack on $\mathcal{E}^r$

In [25], Lucks presented an attack on triple-DES in approximately 2^{90} queries. Gaži [16] generalized this result to $\mathcal{E}^r$ in the ideal cipher model, for any $r \geq 1$. Paraphrased,

Lemma 1 (Gaži [16]). *Let $\kappa, n, r \geq 1$ be integral. Put $r' = \lceil r/2 \rceil$. Then, for any $0 < \tau < 2^{2n/r-1}$,*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^r}(2\tau 2^{\frac{r'-1}{r'}n} + r2^{\kappa+\frac{r'-1}{r'}n}) \geq 1 - 2/\tau - 2^{r\kappa-\tau(n-1)}.$$

Here, $\mathcal{D}$ makes $2\tau 2^{\frac{r'-1}{r'}n}$ queries to its oracle $\mathcal{E}^r_k/\mathcal{P}$ and $r2^{\kappa+\frac{r'-1}{r'}n}$ queries to E.

Intuitively, this result means that $\mathcal{E}^r$ can be successfully attacked in approximately $2^{\kappa+\frac{r'-1}{r'}n}$ queries. In the following proposition, we extend this result, describing a simple attack that works in approximately $2^{r'\kappa}$ queries, hence being faster if $\kappa < n/r'$.

Proposition 1. *Let $\kappa, n, r \geq 1$ be integral. Put $r' = \lceil r/2 \rceil$. Then, for any $0 < \tau < 2^{n-1}/r$,*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^r}(\tau + 2\tau 2^{r'\kappa}) \geq 1 - 2^{r\kappa - \tau(n-1)}\,.$$

Here, $\mathcal{D}$ makes τ queries to its oracle $\mathcal{E}^r_k/\mathcal{P}$ and $r\tau 2^{r'\kappa}$ queries to E.

Proof. We describe a distinguisher $\mathcal{D}$ that tries all keys via the meet-in-the-middle attack. Intuitively, it chooses τ messages $m_1, \ldots, m_\tau$ and their corresponding ciphertexts $c_1, \ldots, c_\tau$. Then, for each of these message/ciphertext pairs (m_i, c_i), it evaluates the first $\lceil r/2 \rceil$ block ciphers for all possible keys starting from m_i, and the last $\lfloor r/2 \rfloor$ block ciphers in inverse direction starting from c_i. One possible key $k = (k_1, \ldots, k_r)$ must "stand out;" if not, $\mathcal{D}$ likely converses with the ideal world. Formally, $\mathcal{D}$ operates as follows:

fix distinct $m_1, \ldots, m_\tau \in \{0,1\}^n$
for $i = 1, \ldots, \tau$ **do**
  $c_i \leftarrow \mathcal{E}^r_k(m_i)$
  for all $k^*_L = (k^*_1, \ldots, k^*_{\lceil r/2 \rceil}) \in \{0,1\}^{\lceil r/2 \rceil \kappa}$ **do**
    $a_{i,k^*_L} \leftarrow E_{k^*_{\lceil r/2 \rceil}} \circ \cdots \circ E_{k^*_1}(m_i)$
  for all $k^*_R = (k^*_{\lceil r/2 \rceil + 1}, \ldots, k^*_r) \in \{0,1\}^{\lfloor r/2 \rfloor \kappa}$ **do**
    $b_{i,k^*_R} \leftarrow E^{-1}_{k^*_{\lceil r/2 \rceil + 1}} \circ \cdots \circ E^{-1}_{k^*_r}(c_i)$
// identifying correct key:
for all $(k^*_L, k^*_R) \in \{0,1\}^{\lceil r/2 \rceil \kappa} \times \{0,1\}^{\lfloor r/2 \rfloor \kappa}$ **do**
  if $a_{i,k^*_L} = b_{i,k^*_R}$ **for all** $i = 1, \ldots, \tau$ **then**
    return 1
return 0

Clearly, if $\mathcal{D}$ is in the real world $(\mathcal{E}^r_k, E)$, then for $(k^*_L, k^*_R) = k$ we have $a_{i,k^*_L} = b_{i,k^*_R}$ for all $i = 1, \ldots, \tau$. We consider the probability $\mathcal{D}$ returns 1 when in the random world $(\mathcal{P}, E)$. Using that E is evaluated at most $r\tau < 2^{n-1}$ times for every key, $a_{i,k^*_L} = b_{i,k^*_R}$ holds for all i with probability at most $(2/2^n)^\tau$. Summing over all possible choices of (k^*_L, k^*_R), we observe that $\mathcal{D}$ returns 1 with probability at most $2^{r\kappa - \tau(n-1)}$. This completes the proof. □

Proposition 1 and Lemma 1 together imply that the security of $\mathcal{E}^r$ is at most up to $2^{\kappa+\frac{r'-1}{r'}\min\{r'\kappa, n\}}$ queries. Particularly, $\mathcal{E}^3$ and $\mathcal{E}^4$ can be attacked in approximately $2^{\kappa+\min\{\kappa, n/2\}}$ queries.

4 Improved Security Bound on $\mathcal{E}^r$

The main result of this section is to prove that $\mathcal{E}^3$ is secure up to approximately $2^{\kappa+\min\{\kappa,n/2\}}$.

Theorem 1. *Let $\kappa, n \geq 1$ be integral. Then,*

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^3}(q) \leq 12\frac{\alpha 2^\kappa q}{(2^\kappa)^{\underline{3}}} + 4\left(\frac{q}{2^{\kappa+n/2}}\right)^{2/3} + \frac{3}{2^\kappa}\,,$$

where $\alpha = \max\{2eq/2^n, \kappa + n\}$.

Theorem 1 can be interpreted as saying that a distinguisher must make at least approximately $\min\{2^{2\kappa}, 2^{\kappa+n/2}\}$ queries in order to distinguish $\mathcal{E}^3$ from $\mathcal{P}$. As $\mathcal{E}^4$ cannot be less secure than $\mathcal{E}^3$ up to a negligible term (cf. [17]), the bound carries over. Concretely this means that also $\mathcal{E}^4$ is tightly secure up to this bound. For $r \geq 5$, Theorem 1 similarly improves over the previous bounds on $\mathcal{E}^r$.

In order to prove Theorem 1, we can rely on the previous result from Bellare and Rogaway [3] and its generalization from Gaži and Maurer [17], and we only need to improve one rather isolated lemma (Lemma 4 in both works). In Section 4.1 we formally state and prove our lemma, and compare it with [3, 17]. Then, in Section 4.2, we prove Theorem 1. This proof is in essence based on the proofs of [3,17], but then using the new result from Section 4.1.

4.1 Improved Lemma: Finding a Chain

Consider the following definition, derived from [17]:

Definition 1. *Let P be a given permutation. Let $\mathcal{D}$ be a distinguisher with query access to E. Let $r \geq 1$ be integral, and let $k = (k_1, \ldots, k_r)$ be a key. We say that $\mathcal{D}$ finds an* i-disconnected chain for k, *where $i \in \{1, \ldots, r\}$, if for some $m \in \{0,1\}^n$ it makes all necessary queries to E for the evaluation of*

$$E_{k_{r-i}} \circ \cdots \circ E_{k_1} \circ P^{-1} \circ E_{k_r} \circ \cdots \circ E_{k_{r-i+1}}(m)\,.$$

We say that $\mathcal{D}$ evaluates *k if this happens for some i.*

In [17] (the derivation of [3] is fairly the same), Gaži and Maurer proved that $\mathcal{D}$ evaluates secret key $k = (k_1, \ldots, k_r)$ with probability at most $2r\alpha^{\lfloor r/2 \rfloor}\frac{q^{\lceil r/2 \rceil}}{(2^\kappa)^{\underline{r}}}$, where $\alpha = \max\{2e2^{\kappa-n}, 2n + \kappa\lfloor r/2 \rfloor\}$. We improve this bound for $r = 3$: informally, we demonstrate that this happens with a significantly smaller probability, therewith allowing to achieve tight security for $\mathcal{E}^3$ and $\mathcal{E}^4$. We note that the trick we employ is specific to $r = 3$: if we generalize this approach to $r \geq 5$ it only leads to a very small

improvement. In other words, it appears to be a non-straightforward problem to improve this bound further for $r \geq 5$. We elaborate on this in Section 5.

Lemma 2. *Let $k = (k_1, k_2, k_3) \xleftarrow{\$} \{0,1\}^{3\kappa}$ distinct. Consider $\mathcal{D}$, making at most q queries to ideal cipher E. Then,*

$$\mathbf{Pr}\left(\mathcal{D}^E \text{ evaluates } k\right) \leq 12\frac{\alpha 2^\kappa q}{(2^\kappa)^{\underline{3}}}\,,$$

where $\alpha = \max\{2eq/2^n, \kappa + n\}$.

Proof. Our goal is to bound the event that $\mathcal{D}$ finds an i-disconnected chain for k, where $i \in \{1,2,3\}$. Let $\mathsf{Ch}_{i,3}(E)$ denote the expected number of i-disconnected chains $\mathcal{D}$ finds (for arbitrary keys, hence not necessarily just (k_1, k_2, k_3)) when making q queries to the ideal cipher E. Denote by $x \xrightarrow{k} y$ the event that the distinguisher queried $y \leftarrow E_k(x)$ or $x \leftarrow E^{-1}(y)$.

The probability that $\mathcal{D}^E$ finds an i-disconnected chain for the secret key k can be upper bounded by the expected number of such chains $\mathcal{D}$ finds *for any key*, $\mathsf{Ch}_{i,3}(E)$, divided by the total number of possible keys, $(2^\kappa)^{\underline{3}}$. Formally, we find (cf. [3,17]):

$$\mathbf{Pr}\left(\mathcal{D}^E \text{ evaluates } k\right) \leq \frac{\sum_{i=1}^{3} \mathsf{Ch}_{i,3}(E)}{(2^\kappa)^{\underline{3}}}\,.$$

As P is considered to be a given permutation, the cases are the same for all i, and we focus on $\mathsf{Ch}_{3,3}(E)$. Technically, we bound the expected number of solutions to

$$m \xrightarrow{k_1'} a \xrightarrow{k_2'} b \xrightarrow{k_3'} c\,, \text{ for } m, a, b, c, k_1', k_2', k_3' \text{ arbitrary.}$$

Write $\overrightarrow{\mathsf{Ch}}_{3,3}(E)$ to be the expected number of chains where the middle block cipher call is a forward query, and $\overleftarrow{\mathsf{Ch}}_{3,3}(E)$ the expected number of chains where the middle block cipher call is an inverse query. Clearly

$$\mathsf{Ch}_{3,3}(E) = \overrightarrow{\mathsf{Ch}}_{3,3}(E) + \overleftarrow{\mathsf{Ch}}_{3,3}(E)\,,$$

and we bound these expected values separately. We focus on $\overrightarrow{\mathsf{Ch}}_{3,3}(E)$. Define

$$\overrightarrow{w}(E) = \max_b \left|\{(k,a) \mid b \leftarrow E_k(a),\ \text{forward query}\}\right|.$$

Now, fix any query $b \xrightarrow{k_3'} c$ (q choices). There are at most $\overrightarrow{w}(E)$ choices for the second query, fixing k_2' and a. Starting from a, there are at most 2^κ possible queries $m \xrightarrow{k_1'} a$. Hence $\overrightarrow{\mathsf{Ch}}_{3,3}(E) \leq \overrightarrow{w}(E)2^\kappa q$.

Claim. Let $\alpha \geq 2eq/2^n$, then $\mathbf{Pr}\left(\overrightarrow{w}(E) \geq \alpha\right) \leq 2^{n-\alpha}$.

Proof (Proof of claim). Fix any $b \in \{0,1\}^n$. Observe that for every key k there is exactly one value a that satisfies $E_k(a) = b$. Adaptivity does not help the adversary in hitting b [21], hence every guess succeeds with probability $\frac{1}{2^n}$. Then,

$$\begin{aligned}\mathbf{Pr}\left(\left|\{(k,a) \mid b \leftarrow E_k(a),\ \text{forward query}\}\right| \geq \alpha\right) &\leq \binom{q}{\alpha}\left(\frac{1}{2^n}\right)^{\alpha} \\ &\leq \left(\frac{eq}{\alpha 2^n}\right)^{\alpha}.\end{aligned}$$

Now, given $\alpha \geq 2eq/2^n$, this value is bounded by $2^{-\alpha}$. The proof is completed by summing over all choices of b. □

This claim gives us, where we use that $\overrightarrow{w}(E) \leq 2^{\kappa}$:

$$\begin{aligned}\overrightarrow{\mathsf{Ch}}_{3,3}(E) &\leq \mathbf{E}\left(\overrightarrow{\mathsf{Ch}}_{3,3}(E)\middle|\overrightarrow{w}(E) < \alpha\right) + \mathbf{E}\left(\overrightarrow{\mathsf{Ch}}_{3,3}(E)\middle|\overrightarrow{w}(E) \geq \alpha\right) 2^{n-\alpha} \\ &\leq \alpha 2^{\kappa} q + 2^{2\kappa} q 2^{n-\alpha} \leq \alpha 2^{\kappa} q + 2^{\kappa} q\,,\end{aligned}$$

using $\alpha \geq \kappa + n$. Naturally, $\alpha \geq 1$ and henceforth we find $\overrightarrow{\mathsf{Ch}}_{3,3}(E) \leq 2\alpha 2^{\kappa} q$. The analysis for $\overleftarrow{\mathsf{Ch}}^{E}_{3,3}$ is the same. This completes the proof. □

4.2 Proof of Theorem 1

The proof of Theorem 1 is similar to the proofs of [3,17]. Gaži and Maurer [17] generalized the result of Bellare and Rogaway [3], and pointed out a few small flaws in the proof. Additionally, the proof of [17] is more compact, and we use this proof as starting point. For ease of presentation, we will employ our own terminology, and stick to $r = 3$.

Let $E \xleftarrow{\$} \mathsf{Block}(\kappa, n)$ and $\mathcal{P} \xleftarrow{\$} \mathsf{Perm}(n)$. Write $\mathcal{O}_1$ as $(\pm\mathcal{P}, \pm E)$. For simplicity, we assume $\mathcal{O}_1$ randomly samples 3 distinct keys $k = (k_1, k_2, k_3)$ from $\{0,1\}^{\kappa}$ in advance (these will not be used). Next, write $\mathcal{O}_2$ as the real world $(\pm\mathcal{E}^3_k, \pm E)$ where $k = (k_1, k_2, k_3) \xleftarrow{\$} (\{0,1\}^{\kappa})^3$ is selected in advance. Let $\mathcal{D}$ be any distinguisher making q queries. Our goal is to bound

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^3}(\mathcal{D}) = \left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_2} = 1\right)\right|.$$

Let $\mathcal{O}_3$ be $\mathcal{O}_2$ with the difference that the keys $k = (k_1, k_2, k_3)$ are distinct. As these two worlds can only be distinguished in case of a key collision, which happens with probability at most $\binom{3}{2}/2^{\kappa} = 3/2^{\kappa}$, we find:

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^3}(\mathcal{D}) \leq \left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_3} = 1\right)\right| + \frac{3}{2^{\kappa}}.$$

Next, let $\mathcal{O}_4$ be the world $\mathcal{O}_1$ with the difference that for k_3, E_{k_3} is defined to satisfy the equation $\mathcal{P} = E_{k_3} \circ E_{k_2} \circ E_{k_1}$ (for all other key inputs, E operates as is).

Note that in $\mathcal{O}_3$, the oracles $E_{k_1}, E_{k_2}, E_{k_3}, \mathcal{E}_k^3$ are all random permutations with the sole restriction that $\mathcal{E}_k^3 = E_{k_3} \circ E_{k_2} \circ E_{k_1}$, and similarly for $\mathcal{O}_4$ with the role of $\mathcal{E}_k^3$ replaced by $\mathcal{P}$. As such, $\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_3} = 1\right) = \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1\right)$, and we obtain:

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^3}(\mathcal{D}) \leq \left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1\right)\right| + \frac{3}{2^\kappa}\,.$$

We simplify the analysis, giving $\mathcal{D}$ free access to the permutation $\mathcal{P}$, which means that only its queries to E count.

Say that a query to E is *relevant* if it is made for one of the keys k_1, k_2, k_3. Let $2^{n-1} > \tau > 0$ be a threshold, and denote by $\mathsf{relevant}(\tau)$ the event that $\mathcal{D}$ makes more than τ relevant queries. Define the following event lucky, where

$$\mathsf{lucky} := (\mathcal{D} \text{ sets } \mathsf{relevant}(\tau)) \vee (\mathcal{D} \text{ evaluates } k)\ .$$

Then, via the fundamental lemma of game-playing [3], or its counterpart in random systems [17],

$$\begin{aligned}\mathbf{Adv}^{\mathrm{sprp}}_{\mathcal{E}^3}(\mathcal{D}) \leq \left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1 \mid \neg\mathsf{lucky}\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1 \mid \neg\mathsf{lucky}\right)\right| + \\ \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1}\ \mathsf{lucky}\right) + \frac{3}{2^\kappa}\,,\end{aligned} \tag{3}$$

for which additionally,

$$\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1}\ \mathsf{lucky}\right) \leq \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} \text{ sets } \mathsf{relevant}(\tau)\right) + \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} \text{ evaluates } k\right)\ .$$

In Lemma 2 it is proven that

$$\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} \text{ evaluates } k\right) \leq 12\frac{\alpha 2^\kappa q}{(2^\kappa)^{\underline{3}}}\,,$$

where $\alpha = \max\{2eq/2^n, \kappa + n\}$. To bound $\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} \text{ sets } \mathsf{relevant}(\tau)\right)$, note that any query is relevant with probability $\frac{3}{2^\kappa}$. Hence, the expected number of relevant queries is at most $\frac{3q}{2^\kappa}$. By Markov's inequality, we obtain

$$\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} \text{ sets } \mathsf{relevant}(\tau)\right) \leq \frac{3q}{\tau 2^\kappa}\,. \tag{4}$$

Consider the remaining term of eqn. (3). It equals $\Delta_q(\mathbf{G}, \mathbf{H})$ of Gaži and Maurer [17], who proved the following bound (for completeness, we rewrote this proof in our terminology in Lemma 3 in Appendix A):

$$\left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1 \mid \neg\mathsf{lucky}\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1 \mid \neg\mathsf{lucky}\right)\right| \leq \frac{\tau^2}{2^n}\,. \tag{5}$$

Equations (4-5) hold for any τ, and we select it to minimize the sum of the terms, balancing between eqn. (4) and eqn. (5). The minimum is achieved for $\tau = (3q2^{n-\kappa-1})^{1/3}$, for which

$$\min_{\tau} \frac{3q}{\tau 2^{\kappa}} + \frac{\tau^2}{2^n} = \left(\frac{3q}{2^{\kappa+n/2-1/2}}\right)^{2/3} + \left(\frac{3q}{2^{\kappa+n/2+1}}\right)^{2/3} \leq 4\left(\frac{q}{2^{\kappa+n/2}}\right)^{2/3}.$$

This completes the proof of Theorem 1.

5 Conclusions

Recent security analysis has rendered a better understanding of cascaded encryption [3, 16, 17, 19, 24], but the best known security bounds for the classical triple and quadruple encryption have always been non-tight. As main contribution of this work, we close this gap, proving that both schemes are tightly secure up to approximately $2^{\kappa+\min\{\kappa,n/2\}}$ queries. Our results particularly prove tight security of the well-known and still widely employed triple-DES construction.

Also for longer cascades, $\mathcal{E}^r$ with $r \geq 5$, our results render improved attacks *and* improved bounds, with the remark that for $r \geq 16$ and for specific choices of κ, n a better bound by Lee [24] applies. These bounds, however, still leave a gap (see Table 1) and it remains an important open problem to prove tight security for $r \geq 5$. It is possible to apply the techniques of Lemma 2 to larger cascades, but this will only minimally decrease the gap between the best known security and the best known attack. For instance, for $r = 5$ a similar approach would give a dominating term $\alpha^3 q^2/(2^{\kappa})^{\underline{5}}$, and this will not readily allow proving tight security. Additionally, an improvement of this type only affects the first term in $\min\{\cdot,\cdot\}$ in the security bound. Yet, this is still insufficient: also the gap in the second term of $\min\{\cdot,\cdot\}$ needs to be tightened, and for this a structurally different approach seems to be needed. In this aspect, we note that Lee [24] proved that the security of $\mathcal{E}^r$ approaches $2^{\kappa+\min\{\kappa,n\}}$ for increasing r.

Acknowledgments. This work was supported in part by the Research Fund KU Leuven, OT/13/071, and in part by the Research Council KU Leuven: GOA TENSE (GOA/11/007).

References

[1] Aiello, W., Bellare, M., Di Crescenzo, G., Venkatesan, R.: Security amplification by composition: The case of doubly-iterated, ideal ciphers. In: Advances in Cryptology - CRYPTO '98. Lecture Notes in Computer Science, vol. 1462, pp. 390–407. Springer, Heidelberg (1998)

[2] ANSI X9.52: Triple Data Encryption Algorithm Modes of Operation (withdrawn) (1998)
[3] Bellare, M., Rogaway, P.: The security of triple encryption and a framework for code-based game-playing proofs. In: Advances in Cryptology - EUROCRYPT 2006. Lecture Notes in Computer Science, vol. 4004, pp. 409–426. Springer, Heidelberg (2006)
[4] Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. In: Advances in Cryptology - CRYPTO '90. Lecture Notes in Computer Science, vol. 537, pp. 2–21. Springer, Heidelberg (1991)
[5] Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology 4(1), 3–72 (1991)
[6] Biham, E., Shamir, A.: Differential cryptanalysis of the full 16-round DES. In: Advances in Cryptology - CRYPTO '92. Lecture Notes in Computer Science, vol. 740, pp. 487–496. Springer, Heidelberg (1993)
[7] Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Springer (2002)
[8] Davies, D.W., Murphy, S.: Pairs and triplets of DES S-Boxes. Journal of Cryptology 8(1), 1–25 (1995)
[9] Dierks, T., Allen, C.: The TLS protocol. Request for Comments (RFC) 2246 (January 1999), `http://tools.ietf.org/html/rfc2246`
[10] Diffie, W., Hellman, M.: Exhaustive cryptanalysis of the NBS data encryption standard. Computer 10(6), 74–84 (1977)
[11] Dinur, I., Dunkelman, O., Keller, N., Shamir, A.: Efficient dissection of composite problems, with applications to cryptanalysis, knapsacks, and combinatorial search problems. In: Advances in Cryptology - CRYPTO 2012. Lecture Notes in Computer Science, vol. 7417, pp. 719–740. Springer, Heidelberg (2012)
[12] EMVCo: EMV Integrated Circuit Card Specifications for Payment Systems. Book 2: Security and Key Management, version 4.2 (2008)
[13] Even, S., Goldreich, O.: On the power of cascade ciphers. ACM Trans. Comput. Syst. 3(2), 108–116 (1985)
[14] FIPS 46: Data Encryption Standard. National Institute of Standards and Technology (1977)
[15] FIPS 46-3: Data Encryption Standard. National Institute of Standards and Technology (withdrawn) (1999)
[16] Gaži, P.: Plain versus randomized cascading-based key-length extension for block ciphers. In: Advances in Cryptology - CRYPTO (I) 2013. Lecture Notes in Computer Science, vol. 8042, pp. 551–570. Springer, Heidelberg (2013)
[17] Gaži, P., Maurer, U.M.: Cascade encryption revisited. In: Advances in Cryptology - ASIACRYPT 2009. Lecture Notes in Computer Science, vol. 5912, pp. 37–51. Springer, Heidelberg (2009)
[18] Gaži, P., Maurer, U.M.: Free-start distinguishing: Combining two types of indistinguishability amplification. In: International Conference on Information Theoretic Security - ICITS 2009. Lecture Notes in Computer Science, vol. 5973, pp. 28–44. Springer, Heidelberg (2010)
[19] Gaži, P., Tessaro, S.: Efficient and optimally secure key-length extension for block ciphers via randomized cascading. In: Advances in Cryptology - EUROCRYPT 2012. Lecture Notes in Computer Science, vol. 7237, pp. 63–80. Springer, Heidelberg (2012)
[20] ISO/IEC 18033-3:2010. Information technology – Security techniques – Encryption algorithms Part 3: Block ciphers (2010)
[21] Jetchev, D., Özen, O., Stam, M.: Understanding adaptivity: Random systems revisited. In: Advances in Cryptology - ASIACRYPT 2012. Lecture Notes in Computer Science, vol. 7658, pp. 313–330. Springer, Heidelberg (2012)

[22] Kilian, J., Rogaway, P.: How to protect DES against exhaustive key search (an analysis of DESX). Journal of Cryptology 14(1), 17–35 (2001)
[23] Knudsen, L.R., Mathiassen, J.: A chosen-plaintext linear attack on DES. In: Fast Software Encryption 2000. Lecture Notes in Computer Science, vol. 1978, pp. 262–272. Springer, Heidelberg (2001)
[24] Lee, J.: Towards key-length extension with optimal security: Cascade encryption and xor-cascade encryption. In: Advances in Cryptology - EUROCRYPT 2013. Lecture Notes in Computer Science, vol. 7881, pp. 405–425. Springer, Heidelberg (2013)
[25] Lucks, S.: Attacking triple encryption. In: Fast Software Encryption '98. Lecture Notes in Computer Science, vol. 1372, pp. 239–253. Springer, Heidelberg (1998)
[26] Matsui, M.: The first experimental cryptanalysis of the Data Encryption Standard. In: Advances in Cryptology - CRYPTO '94. Lecture Notes in Computer Science, vol. 839, pp. 1–11. Springer, Heidelberg (1994)
[27] Matsui, M.: Linear cryptoanalysis method for DES cipher. In: Advances in Cryptology - EUROCRYPT '93. Lecture Notes in Computer Science, vol. 765, pp. 386–397. Springer, Heidelberg (1994)
[28] Maurer, U.M., Massey, J.L.: Cascade ciphers: The importance of being first. Journal of Cryptology 6(1), 55–61 (1993)
[29] Maurer, U.M., Pietrzak, K.: Composition of random systems: When two weak make one strong. In: Theory of Cryptography Conference 2004. Lecture Notes in Computer Science, vol. 2951, pp. 410–427. Springer, Heidelberg (2004)
[30] Maurer, U.M., Pietrzak, K., Renner, R.: Indistinguishability amplification. In: Advances in Cryptology - CRYPTO 2007. Lecture Notes in Computer Science, vol. 4622, pp. 130–149. Springer, Heidelberg (2007)
[31] Maurer, U.M., Tessaro, S.: Computational indistinguishability amplification: Tight product theorems for system composition. In: Advances in Cryptology - CRYPTO 2009. Lecture Notes in Computer Science, vol. 5677, pp. 355–373. Springer, Heidelberg (2009)
[32] NIST SP 800-67, Revision 1: Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher. National Institute of Standards and Technology (2012)
[33] NIST: Triple DES Validation List (2014), `http://csrc.nist.gov/groups/STM/cavp/documents/des/tripledesval.html`
[34] Oorschot, P., Wiener, M.: Improving implementable meet-in-the-middle attacks by orders of magnitude. In: Advances in Cryptology - CRYPTO '96. Lecture Notes in Computer Science, vol. 1109, pp. 229–236. Springer, Heidelberg (1996)
[35] Tessaro, S.: Security amplification for the cascade of arbitrarily weak PRPs: Tight bounds via the interactive hardcore lemma. In: Theory of Cryptography Conference 2011. Lecture Notes in Computer Science, vol. 6597, pp. 37–54. Springer, Heidelberg (2011)
[36] Vaudenay, S.: Adaptive-attack norm for decorrelation and super-pseudorandomness. In: Selected Areas in Cryptography '99. Lecture Notes in Computer Science, vol. 1758, pp. 49–61. Springer, Heidelberg (2000)

A Proof of Equation (5)

For completeness of this work, we rewrite the proof of Gaži and Maurer [17] of (5) in our terminology and restricted to the case of $\mathcal{E}^3$.

Lemma 3. *Let oracles $\mathcal{O}_1, \mathcal{O}_4$, distinguisher $\mathcal{D}$, and event* lucky *be as in the proof of Theorem 1. Then,*

$$\left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1 \mid \neg\mathsf{lucky}\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1 \mid \neg\mathsf{lucky}\right)\right| \leq \frac{\tau^2}{2^n}\,. \tag{6}$$

Proof. Recall that we consider $\mathcal{O}_1$ and $\mathcal{O}_4$, where we assume that lucky does not get satisfied. In other words, $\mathcal{D}$ makes at most τ relevant queries and $\mathcal{D}$ does not evaluate k. The first condition means that it suffices to focus on the τ relevant queries. More formally, define an oracle $\mathcal{O}'_1$ consisting of three independent random permutations π_1, π_2, π_3, and let $\mathcal{O}'_4$ be an oracle consisting of three random permutations such that $\pi_3 \circ \pi_2 \circ \pi_1 = id$.

Consider the following equivalent construction of $\mathcal{O}_1$: randomly sample three distinct keys $k = (k_1, k_2, k_3)$, use $\mathcal{O}'_1$ to instantiate $E_{k_1}, E_{k_2}, E_{k_3}$ the obvious way, and generate random permutations for the remaining block cipher instances. Similarly, an equivalent construction for $\mathcal{O}_4$ uses π_1, π_2 of $\mathcal{O}'_4$ to instantiate E_{k_1}, E_{k_2} and π_3 to instantiate $\mathcal{P}^{-1} \circ E_{k_3}$. The condition that $\mathcal{D}$ does not evaluate k translates to the condition that $\pi_3 \circ \pi_2 \circ \pi_1$ does not evaluated.

Let $\mathcal{C}$ be any distinguisher that makes τ queries to $\mathcal{O}'_1$ or $\mathcal{O}'_4$ and does not evaluate $\pi_3 \circ \pi_2 \circ \pi_1$. Then, via the fundamental lemma of game-playing, or its counterpart in random systems,

$$\begin{aligned} &\left|\mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_1} = 1 \mid \neg\mathsf{lucky}\right) - \mathbf{Pr}\left(\mathcal{D}^{\mathcal{O}_4} = 1 \mid \neg\mathsf{lucky}\right)\right| \\ \leq\ &\left|\mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'_1} = 1\right) - \mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'_4} = 1\right)\right|. \end{aligned} \tag{7}$$

Before proceeding, we first introduce some terminology. For a permutation π_i ($i \in \{1,2,3\}$), let $\mathsf{dom}_j(\pi_i)$ (resp. $\mathsf{rng}_j(\pi_i)$) denote the set of domain (resp. range) points that are defined for π_i by the first j queries of the distinguisher. Note that we have $|\mathsf{dom}_j(\pi_i)| = |\mathsf{rng}_j(\pi_i)|$ and $\sum_{i=1}^{3} |\mathsf{dom}_j(\pi_i)| \leq j$ by construction. Next, for $i \in \{0,1,2,3\}$, we define $\mathcal{X}_j(i) = \{0,1\}^n \backslash (\mathsf{rng}_j(\pi_i) \cup \mathsf{dom}_j(\pi_{i+1}))$, where the indexing of the permutations is considered cyclic (i.e. π_4 denotes π_1 and π_0 denotes π_3, and thus $\mathcal{X}_j(0) = \mathcal{X}_j(3)$).

Next, we define a new oracle $\mathcal{O}'$. It lazily samples three permutations π_1, π_2, π_3 as follows. Repeat queries are handled the obvious way; consider the jth query by the distinguisher and assume it is still undefined. If it is a forward query to π_i ($i \in \{1,2,3\}$) the response is uniformly randomly drawn from $\mathcal{X}_{j-1}(i)$, and if it is an inverse query to π_i^{-1} the response is likewise drawn from $\mathcal{X}_{j-1}(i-1)$. Note that, as $\tau < 2^{n-1}$, we have $\mathcal{X}_j(i) > 0$ for all i, j. Recall that the distinguisher is not allowed to evaluate $\pi_3 \circ \pi_2 \circ \pi_1$.

In the following two claims we bound the distance of $\mathcal{O}'$ to $\mathcal{O}'_1$ and $\mathcal{O}'_4$, respectively.

Claim. $\left|\mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'_1}=1\right)-\mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'}=1\right)\right| \leq \frac{\tau^2}{2^{n+1}}$.

Proof (Proof of claim). Without loss of generality, consider the jth query and assume it is a forward query to π_i. Assume it does not complete an evaluation. Oracle $\mathcal{O}'$ responds to this query with an element from $\mathcal{X}_{j-1}(i)$. Oracle $\mathcal{O}'_1$, on the other hand, answers with a random element from $\{0,1\}^n \backslash \mathsf{rng}_{j-1}(\pi_i) \supseteq \mathcal{X}_{j-1}(i)$. For inverse queries the situation is analogous. Oracles $\mathcal{O}'_1$ and $\mathcal{O}'$ are indistinguishable as long as in $\mathcal{O}'_1$ a forward query to π_i is never responded with an element from $\mathsf{dom}_{j-1}(\pi_{i+1})$ and an inverse query never with an element from $\mathsf{rng}_{j-1}(\pi_{i-1})$. For forward queries, this happens with probability at most

$$\frac{|\mathsf{dom}_{j-1}(\pi_{i+1})|}{|\{0,1\}^n \backslash \mathsf{rng}_{j-1}(\pi_i)|} \leq \frac{|\mathsf{dom}_{j-1}(\pi_{i+1})| + |\mathsf{rng}_{j-1}(\pi_i)|}{|\{0,1\}^n|} \leq \frac{j-1}{2^n}.$$

The same bound holds for inverse queries. As the distinguisher makes τ queries, we obtain our bound as $\sum_{j=1}^{\tau} \frac{j-1}{2^n} \leq \frac{\tau^2}{2^{n+1}}$. □

Claim. $\left|\mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'}=1\right)-\mathbf{Pr}\left(\mathcal{C}^{\mathcal{O}'_4}=1\right)\right| \leq \frac{\tau^2}{2^{n+1}}$.

Proof (Proof of claim). The permutations of $\mathcal{O}'_4$ satisfy $\pi_3 \circ \pi_2 \circ \pi_1 = id$. To suit the intuition, one may think of this as a set of 2^n cycles in a graph of 2^n vertices, where the edges are defined and labeled by the permutations in the following way: $x \xrightarrow{\pi_1} \pi_1(x) \xrightarrow{\pi_2} \pi_2 \circ \pi_1(x) \xrightarrow{\pi_3} x$, for $x \in \{0,1\}^n$.

We claim that as long as a query to π_i ($i \in \{1,2,3\}$) never hits an old domain point of π_{i+1} or an inverse query to π_i never hits an old range point of π_{i-1}, both oracles are identical. To see this, consider jth query and assume it is a forward query to π_i (without loss of generality). It is responded with a value randomly drawn from $\{0,1\}^n$ but not from $\mathsf{rng}_{j-1}(\pi_i)$ (as π_i is a permutation) and not from $\mathsf{dom}_{j-1}(\pi_{i+1})$ (as it would invalidate the above condition). Hence, it is virtually drawn from $\mathcal{X}_{j-1}(i)$. Here, we use that $\mathcal{C}$ does not evaluate $\pi_3 \circ \pi_2 \circ \pi_1$. Similar for inverse queries. Thus, we find that under the posited condition, $\mathcal{O}'_4$ and $\mathcal{O}'$ behave indistinguishably.

Remains to consider the event that in $\mathcal{O}'_4$ a query to π_i ($i \in \{1,2,3\}$) hits an older query at position $i+1$ (for forward queries) or $i-1$ (for inverse queries). The same analysis as for the previous claim shows that this happens with probability at most $\frac{\tau^2}{2^{n+1}}$. □

Equation (6) follows directly from (7) and the two claims via the triangle inequality. □