Tight Security Bounds for Multiple Encryption

Yuanxi Dai[1] and John Steinberger[1*]

Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing.
shustdc@gmail.com, jpsteinb@gmail.com

Abstract. Multiple encryption—the practice of composing a blockcipher several times with itself under independent keys—has received considerable attention of late from the standpoint of provable security. Despite these efforts proving definitive security bounds (i.e., with matching attacks) has remained elusive even for the special case of triple encryption. In this paper we close the gap by improving both the best known attacks and best known provable security, so that both bounds match. Our results apply for arbitrary number of rounds and show that the security of ℓ-round multiple encryption is precisely $\exp(\kappa + \min\{\kappa(\ell' - 2)/2), n(\ell' - 2)/\ell'\})$ where $\exp(t) = 2^t$ and where $\ell' = 2\lceil \ell/2 \rceil$ is the even integer closest to ℓ and greater than or equal to ℓ, for all $\ell \geq 1$. Our technique is based on Patarin's H-coefficient method and reuses a combinatorial result of Chen and Steinberger originally required in the context of key-alternating ciphers.

1 Introduction

Let $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ be a blockcipher with key space $\{0,1\}^\kappa$ and message/ciphertext space $\{0,1\}^n$. The *ℓ-cascade of E*, denoted $E^{(\ell)}$, is the blockcipher of key space $\{0,1\}^{\ell\kappa}$ and of message space $\{0,1\}^n$ obtained by composing E ℓ times with itself under independent keys. Thus

$$E_k^{(\ell)}(x) = E_{k_\ell}(E_{k_{\ell-1}}(\ldots(E_{k_1}(x))\ldots)) \tag{1}$$

where $k = k_1\|\ldots\|k_\ell \in \{0,1\}^{\ell\kappa}$. (The inverse of $E^{(\ell)}$ is computed the obvious way.) In particular $E^{(1)} = E$.

Since $E^{(\ell)}$ has longer keys than E for $\ell \geq 2$, the ℓ-cascade can be viewed as a natural mechanism for increasing the key space of a blockcipher and, hence, potentially, enhancing the security level. Security does not necessarily increase linearly with the key length, however. For example there exist meet-in-the-middle (key-recovery) attacks against cascades of length 2 that cost no more[1] than generic (key-recovery) attacks against cascades of length 1 [9]. Indeed, when a variant of DES with longer keys was needed, designers eschewed double encryption (cascades of length 2) in favor of triple encryption [9,24]. The standard which eventually resulted, so-called Triple DES [2,13,27], is still widely deployed.

Even while generic attacks have guided the considerations of designers since the beginning, finding nontrivial provable security results for multiple encryption in idealized models remained an open problem for a very long time. In the ideal model which we and most previous authors envisage [1,4,14,15,20] the security of the ℓ-cascade is quantified by the information-theoretic indistinguishability of two worlds, "real" and "ideal". In the "real" world the adversary A is given oracle access to an ideal[2] cipher E, to its inverse E^{-1}, and to a randomly keyed ℓ-cascade instance

[*] Supported by National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61033001, 61361136003, and by the China Ministry of Education grant number 20121088050.

[1] This should be qualified: the memory costs are much larger and the query complexity is *slightly* greater [1].

[2] I.e., $E(k,\cdot) : \{0,1\}^n \to \{0,1\}^n$ is a random permutation for each key $k \in \{0,1\}^\kappa$.

$E_k^{(\ell)}$ of E (for hidden k) as well as to the inverse $(E_k^{(\ell)})^{-1}$ of the ℓ-cascade; in the "ideal" world the ℓ-cascade instance $E_k^{(\ell)}$ is replaced by a random independent permutation π and its inverse π^{-1}. We emphasize that A has no computational restrictions and that A knows the value of ℓ in question.

We note this experiment already makes sense for $\ell = 1$. The case $\ell = 1$, while quite simple, is already instructive to analyze. In that case the adversary must distinguish between $E_k^{(1)} = E_k$ and a random permutation π, while being given oracle access to E. Since E is ideal, it is easy to argue that the adversary has no advantage as long as it has not queried its oracle E on key k. With k being uniform at random, and with other queries to $E/\pi/E_k$ giving no clue as to the value of k, the adversary's distinguishing advantage is thus upper bounded by—and in fact basically equal to—$q/2^\kappa$, where q is the number of queries made. (We note this bound holds even if n is very small compared to κ, e.g., $n = 1, 2$. For the sake of completeness, we formalize the argument just sketched in Appendix C.) An easy reduction[3] argument, moreover, shows that $E^{(\ell)}$ is at least as secure as $E^{(r)}$ for all $r \leq \ell$. Hence $E^{(\ell)}$ achieves *at least* κ bits of security for all $\ell \geq 1$, and the basic question is to determine how security grows with ℓ.

The first nontrivial results obtained pertaining to this question were by Aiello et al. [1] who show that $E_k^{(2)}$ is *slightly* harder to distinguish from a random π than $E_k^{(1)} = E_k$. More precisely, Aiello et al. show that A's distinguishing advantage for $E^{(2)}$ is upper bounded by an expression of the form $q^2/2^{2\kappa}$, as opposed to $q/2^\kappa$ for $E^{(1)}$, where q is the number of queries made by A. In either event, thus, $E^{(1)}$ and $E^{(2)}$ both essentially offer κ bits of security, given the meet-in-the-middle attack for length two cascades of cost $q = 2^\kappa$ [9]. (See also Appendix C, where we revisit Aiello et al.'s result.)

Subsequently we will write $\exp(\kappa)$ for 2^κ, somewhat in line with the computer science convention of writing $\log(t)$ for $\log_2(t)$. We thus say, e.g., that $E^{(1)}$ and $E^{(2)}$ "achieve security $\exp(\kappa)$", in the sense that it requires about $\exp(\kappa) = 2^\kappa$ queries to achieve constant distinguishing advantage between the real and ideal worlds for those cascade lengths.

After Aiello et al., substantial progress had to wait for Bellare and Rogaway [4], who showed that $E^{(3)}$ achieves security (up to lower-order terms, which we disregard to the remainder of the introduction)

$$\exp(\kappa + \min\{\kappa/2, n/2\}). \tag{2}$$

Bellare and Rogaway did not present matching attacks, but Lucks [22] had previously presented a key-recovery attack on $E^{(3)}$ of cost $\exp(\kappa + n/2)$, indicating that (2) is at least tight for $n \leq \kappa$.

Further progress was made by Gaži and Maurer [15] who, besides finding and correcting some mistakes in the proof of [4], generalized the approach to ℓ rounds. They prove that $E^{(\ell)}$ achieves security at least

$$\exp(\kappa + \min\{\kappa(\ell' - 2)/\ell', n/2\}) \tag{3}$$

where ℓ' (here and later) is the smallest even integer greater than or equal to ℓ. We note this bound also makes sense for $\ell = 1, 2$ and is equivalent to (2) for $\ell = 3$. We also note that (3) approaches

$$\exp(\kappa + \min\{\kappa, n/2\})$$

[3] Since the adversaries considered are information-theoretic, we note that we don't even have to consider the reduction's running time lossiness.

as ℓ grows large. In some sense, thus, Gaži and Maurer could claim to have shown that security increases with the number of rounds ℓ. (At least, say, for $\kappa \leq n/2$.) On the other hand, Gaži and Maurer never proved matching *upper* bounds; thus the security they *proved* for, say, 15 rounds, might already be *achieved* (but without proof) at 3 rounds! (And potentially, by the same token, the "true" security for 3 rounds *is never exceeded* by using a higher number of rounds.) This criticism of the Gaži-Maurer bound (or, more precisely, of the conclusions one might derive therefrom) is partly justified in the sense, and as we will see in this paper, that security

$$\exp(\kappa + \min\{\kappa, n/2\})$$

is indeed already achieved at $\ell = 3$ rounds.

We note that while Lucks's attack caps the security of $E^{(3)}$ at $\exp(\kappa + n/2)$ nothing precludes, a priori, $E^{(\ell)}$ from reaching larger security than this (and hence larger security than (2), (3)) for larger values of ℓ. Indeed, Lee [20], using coupling techniques, provided a partial answer to this question by showing that $E^{(\ell)}$ achieves security at least

$$\exp(\kappa + \min\{\kappa, n\} - 8n/\ell). \tag{4}$$

We note this bound is void for $\ell \leq 8$ and that it consistently beats Gaži and Maurer's bound (i.e., for all values of κ) only for $\ell > 16$. Qualitatively, however, Lee's bound shows that security can approach

$$\exp(\kappa + \min\{\kappa, n\})$$

for large numbers of rounds. Lee does not prove any new upper bounds on the security of $E^{(\ell)}$ but makes the basic observation that security cannot exceed

$$\exp(\kappa + n)$$

since with that many queries A can completely learn E, at which point a few queries to $E_k^{(\ell)}$ suffice for the information-theoretic A to recover the key k (a key-recovery attack obviously implies a distinguishing attack).

Shortly thereafter Gaži [14] presented attacks on ℓ-cascades of cost

$$\exp(\kappa + n(\ell' - 2)/\ell'). \tag{5}$$

Plugging in $\ell = 1, 2, 3$ one recovers the cost of the previous best-known attacks for those cascade lengths. Indeed, Gaži's attacks are both a generalization of the meet-in-the-middle attack for length two cascades and of Lucks's attack for length three cascades (in fact Gaži is the first to rigorously analyze Luck's attack).

These results constitute the current state-of-the-art for cascade encryption in the idealized information-theoretic model we have described. In a nutshell, (3) and (4) constitute the best-known lower bounds, while (5) is the best-known upper bound. These bounds, however, still leave much room for wiggle. For example, the exact security of $E^{(3)}$ in the regime $\kappa \leq n$ is still an open question, as all we know is that $E^{(3)}$'s security lies somewhere in the interval

$$[\exp(\kappa + \min\{\kappa/2, n/2\}), \exp(\kappa + n/2)]$$

by (3) and (5).

Our results. In this paper we close all remaining gaps between upper and lower bounds, up to customary lower-order terms. More precisely, we show that $E^{(\ell)}$ has security

$$\exp(\kappa + \min\{\kappa(\ell' - 2)/2, n(\ell' - 2)/\ell'\}) \tag{6}$$

by exhibiting matching attacks and security proofs, for all $\ell \geq 1$. (Note by the form of (6) that new attacks are only needed when $\kappa(\ell' - 2)/2 < n(\ell' - 2)/\ell'$; otherwise the attacks of Gaži, cf. (5), suffice.) One can observe from (6) that $\ell = 2r$ rounds buy the same amount of security as $\ell = 2r - 1$ rounds. In fact, we expect the curve describing the adversary's advantage to be slightly more advantageous for $2r - 1$ rounds than for $2r$ rounds, as observed by Aiello et al. for $r = 1$, but our analysis is not fine-grained enough to verify this.

Techniques. Tightening the security bounds for triple encryption is already an interesting problem in itself. Besides devising a new rather easy attack of cost $\exp(2\kappa)$, it turns out that the bound directly follows from tightening a key combinatorial lemma in Bellare and Rogaway's original proof (Lemma 10 in [5]). Improving the lemma is not particularly difficult. Hence, the fact that triple encryption has gone without a tight security analysis all this time seems, in retrospect, something of an oversight.

We found the case of larger number of rounds (in particular, $\ell \geq 5$) to be more challenging. While we copied the basic approach of Bellare and Rogaway [4] and of Gaži and Maurer [15] some major structural changes were required in order to achieve tightness. In particular, we had to rebundle a key two-step game transition from [15] into a single-step transition. Moreover we found that the best way to handle this (now rather delicate) single-step transition was by Patarin's H-coefficient technique [29]. Here we drew inspiration from Chen and Steinberger [8] and, indeed, reused the key combinatorial lemma of that paper. Roughly speaking, this lemma gives an explicit expression for the probability that

$$(P_\ell \circ \cdots \circ P_1)(a) = b$$

where each P_i is a *partially defined* random permutation of $\{0, 1\}^n$, where $\circ$ denotes function composition, where $a, b \in \{0, 1\}^n$ are two values such that $P_1(a)$ and $P_\ell^{-1}(b)$ are undefined. Here the probability is expressed (in particular, lower-bounded) as a function of the number of edges[4] already defined in the P_i's as well as of the number of "chains" of various lengths[5] formed by those edges in the composition $P_1 \circ \cdots \circ P_\ell$. (In our case $P_i = E_{k_i}$ where $k = k_1\| \ldots \|k_\ell$ is the secret key.) It is noteworthy that the security proofs for three different classes of composed ciphers (key-alternating ciphers [8], cascade ciphers (this paper), and XOR-cascade ciphers [8, 14, 16]) now rely on this lemma, and in each case tight bounds are achieved.

In order to successfully apply the H-coefficient technique and Chen and Steinberger's lemma a crucial step is to upper bound the probability of the adversary obtaining (too many) long chains in $P_\ell \circ \cdots \circ P_1 = E_{k_\ell} \circ \cdots \circ E_{k_1}$. Like Bellare and Rogaway [4] and like Gaži and Maurer [15] before us, we do this by upper bounding the *total* number of query chains of a given length formed by *all* of the adversary's queries to E, regardless of the underlying key, and then by applying a Markov inequality—but in our case *tight* bounds on the total number of query chains are needed. At

[4] If $x \in \{0, 1\}^n$ is a value such that $y = P_i(x)$ is defined, then the pair (x, y) is also called an *edge* of P_i, equating P_i with a bipartite graph (more precisely, a partial matching) from $\{0, 1\}^n$ to $\{0, 1\}^n$. The composition $P_\ell \circ \cdots \circ P_1$ is visualized by "gluing" these bipartite graphs sequentially next to one another.

[5] See the previous footnote.

first glance the combinatorial question is nonobvious (especially given the presence of an adaptive adversary) but we observe that on any path of queries at least half the queries are "backwards" (meaning contrary to the path's direction, in this instance) *for at least one of the two possible ways of orienting the path* (as a given path can be traversed right-to-left or left-to-right). Together with some classical balls-in-bins occupancy results, this simple symmetry-breaking observation gives an easy means of upper bounding the total number of query chains formed, and the bounds we find are also tight. We refer to Proposition 1 for more details.

Other related work. We have already briefly mentioned related work on key-alternating ciphers [7,8,12,19,33] as well as on XOR cascades [14,16,20], to which the beautiful work of Rogaway and Kilian on DESX (a special case of an XOR-cascade) should be added [17].

Coming back to cascade ciphers Merkle and Hellman [24] show an attack on two-key triple encryption, which attack is revisited by Oorschot and Wiener [26]. (See also [25].) Even and Goldreich [11] present a medley of observations on multiple encryption in various models, including some conclusions which are disputed by Maurer and Massey [23]. Finally, the best paper award at CRYPTO 2012, by Dinur et al. [10], concerns, in large part, non-information-theoretic key-recovery attacks on cascade ciphers.

Open questions. As will be seen, our results actually hold even if the adversary is always allowed to make 2^n queries to its permutation oracle (which is $E_k^{(\ell)}$ or π) for free, i.e., to entirely learn its permutation oracle for free. It would be interesting to know if better bounds can be achieved by exploiting possible restrictions on the number of permutation queries. Such restrictions might correspond to natural limitations on the number of encryptions/decryptions available to the adversary.

2 Definitions

Blockciphers and Cascades. A blockcipher is a function $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ such that $E(k,\cdot) : \{0,1\}^n \to \{0,1\}^n$ is a permutation for each key $k \in \{0,1\}^\kappa$. We also write $E_k(x)$ for $E(k,x)$. By the "inverse" E^{-1} of E we mean the blockcipher $E^{-1} : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ such that E_k^{-1} is the inverse permutation of E_k for each $k \in \{0,1\}^\kappa$. We note that $\{0,1\}^\kappa$ is also called the *key space* of the blockcipher whereas $\{0,1\}^n$ is both the *message space* and *ciphertext space*.

For a blockcipher E and an integer $\ell \geq 1$ we define the *ℓ-cascade* of E, written $E^{(\ell)}$, by equation (1). We note that $E^{(\ell)}$ is a blockcipher of key space $\{0,1\}^{\ell\kappa}$ and of message space $\{0,1\}^n$.

Ideal Ciphers. A blockcipher $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ which is sampled uniformly at random from the space of all blockciphers of key space $\{0,1\}^\kappa$ and of message space $\{0,1\}^n$ is called an *ideal cipher*. In this case E_k is a random independent permutation of $\{0,1\}^n$ for each $k \in \{0,1\}^\kappa$.

Security Game. Let ℓ, κ and n be given. Let A be an information-theoretic adversary (or "distinguisher") with oracle access to, among others, an ideal cipher $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$, which we write A^E but by which we mean that A can query *both* E and E^{-1}. (Along the same lines writing A^π indicates that A has access to both π and π^{-1} when π is a permutation.) Then A's *distinguishing advantage* against ℓ-cascades, written $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(A)$ is defined as

$$\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(A) = \Pr[k = k_1^* \| \ldots \| k_\ell^* \xleftarrow{\$} \{0,1\}^{\ell\kappa}; A^{E,E^{(\ell)}_{k^*}} = 1] - \Pr[\pi \xleftarrow{\$} \mathcal{P}; A^{E,\pi} = 1]$$

where the notation

$$k^* = k_1^* \| \ldots \| k_\ell^* \xleftarrow{\$} \{0,1\}^{\ell\kappa}; A^{E,E_k^{(\ell)}} = 1$$

indicates the event that A outputs 1 after interacting with oracles E/E^{-1} and $E_k^{(\ell)}/(E_k^{(\ell)})^{-1}$ where k is sampled uniformly at random from the key space of $E^{(\ell)}$, and hidden from A; whereas the notation

$$\pi \xleftarrow{\$} \mathcal{P}; A^{E,\pi} = 1$$

indicates the event that A outputs 1 after interacting with oracles E/E^{-1} and π/π^{-1} where π is a permutation of $\{0,1\}^n$ sampled uniformly at random from the set of all permutations of $\{0,1\}^n$, here denoted $\mathcal{P}$; and where in either case the sampling of the ideal cipher E at the start of the experiment is kept implicit for the sake of succinctness.

We write

$$\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(q)$$

for the supremum of $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(A)$ taken over all q-query information-theoretic adversaries A. (The notation $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}$ is thus overloaded.)

What Constitutes Security? For our purposes, and following a long tradition of ideal model analyses (such as [1,6,14,15,17,20] and dozens more) the "security" of $E^{(\ell)}$ is the value of q at which $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(q)$ reaches $\Omega(1)$ or, to be more concrete, the smallest value q_0 such that $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(q_0) \geq 0.5$. This single-number measure has revealed itself of practical significance because the adversary's advantage usually falls to zero very quickly as the number of queries q falls beneath q_0. Typically, for example, the advantage is upper bounded by $(q/q_0)^c$ for some $c > 0$, plus a small constant. The value of c roughly determines the quality of the "security curve": the larger the c, the sharper the security threshold (i.e., the faster the drop in the adversary's advantage as q falls below q_0). E.g., single encryption has $c = 1$ whereas double encryption achieves $c = 2$, by Aiello et al. [1]. Our own security exponents are constants between 0 and 1 and which depend on ℓ, as we discuss in more detail in the next section.

We will refer to c as the *security exponent*, when it can be identified. A security exponent thus gives further qualitative information about a security bound, with higher security exponents being better. Unfortunately, comparing the quality of our bounds with previous multiple-round result such as Gaži and Maurer [15] and Lee [20] is made more difficult by the fact that those authors do not explicitly identify their security exponents (which are sometimes, indeed, hard to extract from the midst of complicated security bounds).

3 Statement of Results

Lower bounds. Our paper's main result is the following theorem (as always, $\ell' = 2\lceil \ell/2 \rceil$; we also write $(\ell+1)'$ for $2\lceil(\ell+1)/2\rceil$, etc):

Theorem 1. *If $q \geq 2^n$ then*

$$\mathbf{Adv}^{\mathsf{casc}}_{\ell}(q) \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{2^n} + \alpha(\ell/2+2)\ell^{1/2}\left(\frac{8q}{2^{\kappa+n(\ell'-2)/\ell'}}\right)^{\ell'/(\ell+3)'}$$

where $\alpha = \ell^2 2^\ell (7n)^{\ell'/2}$. Moreover if $q \leq 2^n$ and $2^\ell(6n+2)^{\ell'/2} \leq 2^n$ then

$$\mathbf{Adv}^{\mathsf{casc}}_{\ell}(q) \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{2^n} + \beta(\ell/2+2)\left(\frac{\ell 3^{\ell'} q}{2^{\kappa\ell'/2}}\right)^{2/(\ell+3)'}$$

where $\beta = \ell^2 2^\ell (6\log q + 2)^{\ell'/2-1}$*. Moreover these results also hold if the adversary is allowed to ask, for free, all possible* 2^n *queries to its second oracle.*

We note the constraint $2^\ell(6n+2)^{\ell'/2} \leq 2^n$ that appears in the second part of Theorem 1 is almost always satisfied by practical parameters and is always asymptotically verified as $n \to \infty$. (Indeed, we imagine ℓ as fixed whereas $n, \kappa \to \infty$ according to some fixed ratio.)

We actually prove something a bit stronger and more general, given by Theorem 3 in Appendix A, but the statement of this more general result is also less digestible. Appendix A discusses how Theorem 1 can be obtained as a corollary of Theorem 3.

It directly follows from Theorem 1 that $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(q)$ is small if

$$q \ll \exp(\kappa + \min\{\kappa(\ell'-2)/2, n(\ell'-2)/\ell'\})$$

(note $\kappa + \kappa(\ell'-2)/2 = \kappa\ell'/2$) or, a little more precisely, if

$$q \ll (2^{-\ell/2}(7n)^{-\ell'/4}\ell^{-1})^{(\ell+3)'} \cdot \exp(\kappa + \min\{\kappa(\ell'-2)/2 - 2\ell, n(\ell'-2)/\ell' - 3\}). \tag{7}$$

Indeed, if q is a factor r smaller than the expression on the right of (7), then it is easy to see from Theorem 1 that the adversary's advantage is upper bounded by either $r^{\ell'/(\ell+3)'}$ or $r^{2/(\ell+3)'}$, depending on whether $q \geq 2^n$ or $q \leq 2^n$, disregarding the negligible terms $\ell^2/2^{\kappa+1}$ and $4\ell^2/2^\kappa$. (Note that $\log q \leq n$ in the second part of Theorem 1, so $\beta \leq \alpha$.)

Our security exponents are thus "$\ell'/(\ell+3)'$ if $q \geq 2^n$, $2/(\ell+3)'$ if $q \leq 2^n$" but this is not very informative since what is really of interest is the security exponent that governs *near the threshold*. Here one can note that

$$\kappa(\ell'-2)/2 \geq n(\ell'-2)/\ell' \iff \kappa/2 \geq n/\ell'$$

which leads us to examine the cases $\kappa/2 \geq n/\ell'$ and $\kappa/2 \leq n/\ell'$ separately. If $\kappa/2 \geq n/\ell'$ then

$$\kappa + n(\ell'-2)/\ell' \geq 2n/\ell' + n(\ell'-2)/\ell' = n$$

so that security is given by the first part of Theorem 1 for q's near the security threshold (which, in this case, is $\exp(\kappa + n(\ell'-2)/\ell')$), and the security exponent is $\ell'/(\ell+3)'$. On the other hand if $\kappa/2 \leq n/\ell'$ then

$$\kappa\ell'/2 \leq n$$

so that security is given by the second part of Theorem 1 for *all* q's beneath the security threshold (which is now $\exp(\kappa\ell'/2)$), and the security exponent is $2/(\ell+3)'$.

By means of comparison we point out that Gaži and Maurer, whose security threshold is at $q_0 = \exp(\kappa + \min\{\kappa(\ell'-2)/\ell', n/2\})$, roughly achieve security exponent

$$c = \begin{cases} 2/3 & \text{if } q_0 = \exp(\kappa + n/2) \\ \ell'/2 & \text{if } q_0 = \exp(\kappa + \kappa(\ell'-2)/\ell'). \end{cases}$$

For $\ell = 3$, for example, their security exponents of 2/3 and 2 for respectively $q_0 = \exp(\kappa + n/2)$ and $q_0 = \exp(3\kappa/2)$ (these are also achieved by Bellare and Rogaway for $\ell = 2$). The former security exponent of 2/3 for security threshold $q_0 = \exp(\kappa + n/2)$ matches our own security exponent for this threshold. The latter security exponent of 2 cannot be compared with any of our security exponents, since it belongs to a different (suboptimal) security threshold. Along the same lines, it doesn't make sense to compare our security exponents for $\ell > 4$ with those of Gaži and Maurer, since these belong to different security thresholds[6]. (Lee's security bound [20] has a rather complicated form, and we did not attempt to identify a security exponent.)

Upper bounds. In Section 4 we present a simple attack of query complexity

$$9 \cdot \exp(\kappa\ell'/2)$$

that succeeds in distinguishing $(E, E_k^{(\ell)})$ from (E, π) with overwhelming advantage. This complements the previously quoted attack by Gaži, of query complexity

$$\ell \cdot \exp(\kappa + n(\ell' - 2)/\ell')$$

and which also succeeds with overwhelming advantage. Hence the gap left between lower and upper bounds is essentially the gap left between

$$\min\{9 \cdot \exp(\kappa\ell'/2), \ell \cdot \exp(\kappa + n(\ell' - 2)/\ell')\}$$

and the right-hand side of (7).

4 An attack of cost $\exp(\kappa\ell'/2)$

In this section we describe a new "start-in-the-middle-attack" on $E^{(\ell)}$ of complexity $\exp(\kappa\ell'/2)$, which complements Gaži's attack of query complexity $\exp(\kappa + n(\ell' - 2)/\ell')$. A precise statement is given in the following theorem.

Theorem 2. *For any $n \geq 2$ and any ℓ, κ there exists an adversary A making at most $6\exp(\kappa\ell'/2)$ queries to E and at most $3\exp(\kappa(\ell - \ell'/2))$ queries to $E_k^{(\ell)}/\pi$, such that*

$$\mathbf{Adv}_\ell^{\mathsf{casc}}(A) \geq 1 - \frac{2^{\kappa\ell}}{2^n(2^n - 1)(2^n - 2)}.$$

Comments. Note that $\kappa\ell'/2 \leq n(\ell' - 2)/\ell'$ implies $\kappa\ell \leq 2n$, so that $2^{\kappa\ell}/2^{3n}$ is indeed negligible for $\kappa\ell'/2 \leq n(\ell' - 2)/\ell'$. We also note that

$$\ell - \ell'/2 = \begin{cases} \ell'/2 & \text{if } \ell \text{ is even,} \\ \ell'/2 - 1 & \text{if } \ell \text{ is odd.} \end{cases}$$

In particular $\ell - \ell'/2 \leq \ell'/2$.

Proof. The adversary A, which implements a "start-in-the-middle" attack, is given by the pseudocode of Figure 4. A iterates a main loop $R = 3$ times. In each loop it tries to discard potential candidates for the secret key k. If no candidate keys remain after the final loop it outputs 0, otherwise 1.

A's second oracle, which is either $E_k^{(\ell)}$ or π depending on the world, is written P. We also adopt the convention, not reflected in the pseudocode, that A avoids redundant queries to P.

[6] For $\ell = 4$ and $q_0 = \exp(\kappa + n/2)$, the Gaži-Maurer exponent of 2/3 "momentarily" beats our own exponent of $\ell'/(\ell+3)' = 4/8 = 1/2$ for the same q_0. Since security for $\ell = 3$ transfers (via lossless reduction) to $\ell = 4$, however, one could equally well argue a security exponent of 2/3 based on our own bounds.

```
$K \leftarrow \{0,1\}^{\ell\kappa}$
$R \leftarrow 3$
for $r = 0$ to $R - 1$ do
    for $i = 0$ to $\ell$ let $S_i \leftarrow \emptyset$
    $h \leftarrow \ell - \ell'/2$
    $S_h \leftarrow \{1^r 0^{n-r}\}$
    for $i = h$ to $\ell - 1$
        forall $x \in S_i$ do
            forall $k \in \{0,1\}^\kappa$ do
                query $e_k(x) := E(k, x)$
                let $S_{i+1} \leftarrow S_{i+1} \cup \{e_k(x)\}$
    end for // (for $i = h$ to $\ell - 1$)
    for $i = h$ to $1$
        forall $x \in S_i$ do
            forall $k \in \{0,1\}^\kappa$ do
                query $e_k^{-1}(x) := E^{-1}(k, x)$
                let $S_{i-1} \leftarrow S_{i-1} \cup \{e_k^{-1}(x)\}$
    end for // (for $i = h$ to $1$)
    forall $x \in S_0$, $k_1 \| \ldots \| k_\ell \in K$ do
        if $(e_{k_\ell}(\cdots e_{k_1}(x) \cdots)$ is known and not equal to $P(x))$ then
            $K \leftarrow K \backslash \{k_1 \| \ldots \| k_\ell\}$
end for // (for $r = 0$ to $R - 1$)
If $K = \emptyset$ then return $0$
return $1$
```

Fig. 1. The adversary A for Theorem 2.

It is easy to see that A makes at most $\exp(\kappa\ell'/2) + \exp(\kappa(\ell - \ell'/2)) \leq 2\exp(\kappa\ell'/2)$ queries to E at each iteration of the main loop. The fact that A makes at most $\exp(\kappa(\ell - \ell'/2))$ queries to P at each iteration of the main loop can be seen from the fact that $|S_0| \leq (2^\kappa)^{\ell - \ell'/2}$ at each iteration, and by the convention that A never makes redundant queries to P.

A obviously outputs 1 in the real world, since the real key always remains in K. To upper bound the probability that A outputs 1 in the ideal world we consider the probability that a given value $k = k_1 \| \ldots \| k_\ell$ survives all three iterations. Since $P = \pi$ is a random permutation, this probability is easily[7] seen to be $(2^{-n})(2^n - 1)^{-1}(2^n - 2)^{-1}$. The theorem statement then follows by a union bound over all keys, and by definition of $\mathbf{Adv}^{\mathsf{casc}}_{\ell,\kappa,n}(A)$. □

5 Preliminary reductions and proof overview

In this section we lay some basic groundwork for the proof of Theorem 3, stated in Appendix A. This "groundwork" partly consists of some innocuous transformations made to the adversary and/or to the security game, and partly of a quick introduction to the H-coefficient technique, which we will use.

Modifications of Bellare and Rogaway [4]. We start by modifying the game in the following way. At the very start of the experiment we send a symbol $\star \in \{\bot, \top\}$ to the adversary. In the ideal world we send $\star = \top$, and in the real world we also send $\star = \top$ unless $k^*_\ell = k^*_i$ for some $i < \ell$,

[7] If this causes confusion, it should be noted that "extraneous" queries to P—those made in the process of checking keys that are not equal to k^*—do not affect this probability. And if *this* causes doubt, one should consider that all such "extraneous" queries can be made *after* the queries which are relevant to k^* are made, without changing the probability that k^* survives. One should also note that $E_{k_1} \circ \cdots \circ E_{k_h}$ is a permutation.

where $k^* = k_1^* \| \ldots \| k_\ell^*$ is the secret key, in which case we send $\star = \bot$. Since the adversary is free to disregard $\star$, this modification is without loss of generality.

Next, we make a second modification, namely that if $\star = \bot$ then we forbid the adversary from making any queries. Since $\star$ can only be $\bot$ in the real world this is without loss of generality either (as the adversary already knows which world it is in anyway).

Now we make yet another modification to the real world, by generating a random permutation π like in the ideal world at the beginning of the experiment. If $\star = \top$ we answer queries to $E_{k^*}^{(\ell)}$ by π instead and, to compensate, we define $E_{k_\ell^*} = \pi \circ E_{k_1^*}^{-1} \circ \cdots \circ E_{k_{\ell-1}^*}^{-1}$ (thus "overwriting" $E_{k_\ell^*}$). Since this simply trades the randomness of $E_{k_\ell^*}$ for the randomness in π, it is easy to see that this is an equivalent way of defining the real world.

Note that both worlds now involve an independent[8] random permutation π. For each fixed permutation S one can also consider the distinguishing experiment where π is set to S in each world. A simple averaging argument over π (see, e.g., Appendix A of [8] for something very similar) shows, moreover, that there must exist some S for which the adversary's distinguishing advantage is at least as great when π is fixed to S as when π is random. We can thus assume without loss of generality that π is not sampled at random, but set to the same fixed permutation S in both worlds. Since S is fixed, now, and since we are quantifying over all information-theoretic adversaries A, we can assume that A knows S and, hence, makes no queries to its second oracle[9].

To summarize, modifications so far amount to this: in the real world, we abort the experiment if $k_\ell^* = k_i^*$ for some $i < \ell$, whereas in the contrary (generic) case there is some fixed permutation S, known to the adversary, such that $E_{k_\ell^*} = S \circ E_{k_1^*}^{-1} \circ \cdots \circ E_{k_{\ell-1}^*}^{-1}$. The ideal world never aborts.

Further Normalizations. Since A is information-theoretic we can assume without loss of generality that A is *deterministic.*

As customary, we also assume that A never makes a query to which it already knows the answer. such as querying $E(k, x)$, obtaining answer y, and later querying $E^{-1}(k, y)$. For the remainder of the proof we assume the presence of a fixed adversary A confirming to these conventions.

As in [8] we will also modify the experiment by *giving the secret key to A after it has finished making all its queries.* More precisely, in the real world we give the "real" key k^* used to key the second oracle $E_{k^*}^{(\ell)}$ whereas in the ideal world (where no such key exists) we sample a "dummy" key $k^* \in \{0,1\}^{\kappa\ell}$ uniformly at random and give this dummy key to A. Since A is free to disregard this extra information this is also without loss of generality. For consistency, we also give this key in the real world if the real world aborts.

For linguistic convenience we will also view the strings $\star$ and k^* which A learns as being the result of "queries" made by A to its "oracles" (as, indeed, would be easy to formalize).

Transcripts. The interaction of A with its oracles is encoded by a *transcript* which, basically, a list of questions asked and answers received, together also with the key value received at the end of the experiment.

More precisely, a transcript can be encoded by a triple of the form $(\star, Q_E, k^*)$ where $\star \in \{\bot, \top\}$, where $k^* \in \{0,1\}^{\kappa\ell}$ is the final key value received, and where Q_E is an *unordered* set of triples of

[8] The real world now has three "random tapes": one for k^*, one for π, and one for the ideal cipher E. Every query made by the adversary is deterministically answered as a function of these three random tapes, and these random tapes are independently sampled. This is the sense in which π is "independent" from other randomness in the real world.

[9] It is this modification which allows the proof to support 2^n second oracle queries since, in the end, we "give" the second oracle to A anyway.

the form $(k, x, y) \in \{0,1\}^\kappa \times \{0,1\}^n \times \{0,1\}^n$ with each such tuple indicating that either $E(k, x)$ was queried with answer y or that $E^{-1}(k, y)$ was queried with answer x. Indeed, A's interaction with its oracles can be unambiguously reconstructed from such an "unordered and undirected" set Q_E by using the fact that A is deterministic, cf. [8].

We write $\mathcal{T}$ for the set of all possible transcripts.

Probability space of Oracles. Let $\mathcal{P}$ be the set of all permutations from $\{0,1\}^n$ to $\{0,1\}^n$. Then a blockcipher of key space $\{0,1\}^\kappa$ and message space $\{0,1\}^n$ can be viewed as an element of $\mathcal{P}^{\exp(\kappa)}$ (2^κ-fold direct product). Thus, an ordered pair

$$(E', k^*) \in \mathcal{P}^{\exp(\kappa)} \times \{0,1\}^{\kappa\ell}$$

uniquely determines a real-world environment for A. More precisely, unless $\star = \perp$ in which case A receives no further information except for k^*, A's ideal cipher oracle E is defined by

$$E_k = \begin{cases} E'_k & \text{if } k \neq k^*_\ell \\ S \circ E'^{-1}_{k^*_1} \circ \cdots \circ E'^{-1}_{k^*_{\ell-1}} & \text{if } k = k^*_\ell \end{cases}$$

where $k^* = k^*_1 \| \ldots \| k^*_\ell$. We thus identify elements of

$$\Omega_X := \mathcal{P}^{\exp(\kappa)} \times \{0,1\}^{\kappa\ell}$$

with real-world oracles. We view Ω_X as a probability space with uniform measure (indeed, the definition of the real-world experiment induces uniform measure on Ω_X).

We similarly define

$$\Omega_Y := \mathcal{P}^{\exp(\kappa)} \times \{0,1\}^{\kappa\ell}$$

to be identified with the set of all ideal-world oracles, and which we also view as a probability space with uniform measure. Here the last coordinate corresponds to the "dummy key" given to the adversary at the end of the experiment. We emphasize that, for $(E, k^*) \in \Omega_Y$, the ideal cipher oracle to which A has access is precisely E, i.e., with no key being overwritten as a function of k^* and S; this is precisely the difference between the real and ideal worlds in the (generic) case when $k^*_\ell \notin \{k^*_1, \ldots, k^*_{\ell-1}\}$.

We can view the transcript produced by A in the real world as a random variable defined over Ω_X. Formally, let $X : \Omega_X \to \mathcal{T}$ be the function defined by letting $X(\omega)$ be the transcript obtained by running A on oracle ω. Thus X is a random variable of range $\mathcal{T}$, and the distribution of X is exactly the distribution of transcripts in the real world. We similarly define $Y : \Omega_Y \to \mathcal{T}$, so that Y is the transcript distribution in the ideal world.

The H-coefficient technique [28, 29], in its simplest form, states that if we can divide $\mathcal{T}$ into a set of (so-called) "good" transcripts $\mathcal{T}_1$ and (so-called) "bad" transcripts $\mathcal{T}_2$, such that[10]

$$\frac{\Pr[X = \tau]}{\Pr[Y = \tau]} \geq 1 - \varepsilon_1 \tag{8}$$

for some $\varepsilon_1 > 0$ and for all $\tau \in \mathcal{T}_1$, then the adversary's distinguishing advantage is upper bounded by

$$\Pr[Y \in \mathcal{T}_2] + \varepsilon_1.$$

[10] By convention, the ratio $\Pr[X = \tau] / \Pr[Y = \tau]$ is considered to be ∞ if $\Pr[Y = \tau] = 0$.

I.e., the distinguishing advantage is upper bounded by the probability of obtaining a bad transcript in the ideal world, plus

$$\max_{\tau\in\mathcal{T}_1}(1-\Pr[X=\tau]/\Pr[Y=\tau]).$$

We refer to [8] for more details. (One could reverse the roles of the real and ideal worlds, since the method is completely general, but $\Pr[Y\in\mathcal{T}_2]$ is typically much easier to compute than $\Pr[X\in\mathcal{T}_2]$ due to the ideal world's nice structure.)

Computing transcript probabilities. Another key insight of the H-coefficient technique is that the probability of obtaining a transcript in either world can be computed via the formulas

$$\Pr[X=\tau]=\frac{|\mathsf{comp}_X(\tau)|}{|\Omega_X|},\qquad \Pr[Y=\tau]=\frac{|\mathsf{comp}_Y(\tau)|}{|\Omega_Y|}\tag{9}$$

as long as $\Pr[Y=\tau]>0$, and where $\mathsf{comp}_X(\tau)\subseteq\Omega_X$ (resp. $\mathsf{comp}_Y(\tau)\subseteq\Omega_Y$) is the set of real-world (resp. ideal-world) oracles that are compatible with a transcript τ, where "compatibility" is defined the obvious[11] way: an oracle ω is compatible with a transcript τ if each individual query in τ is compatible with ω, i.e., if asking that query to ω would result in the answer seen in τ (in particular τ's key value should match ω's, since the key appears in the transcript). We note that query order and query direction (E versus E^{-1}) have no bearing on which transcripts are compatible with which oracles (even should the transcript contain such information). See [8] and Appendix D for further discussion of these identities.

Terminology: Chains. Let $\tau=(\star,Q_E,k^*)$ be a transcript, where $k^*=k_1^*\|\ldots\|k_\ell^*$. Loosely following [15], a tuple $(h,x_h,k_{h+1},x_{h+1},k_{h+2},\ldots,k_{h+r},x_{h+r})$ where $0\le h\le\ell-1$ is an integer, where $1\le r\le\ell$, and where

$$\begin{cases}(k_i,x_{i-1},x_i)\in Q_E & \text{if } i-1\neq\ell\\ (k_i,S^{-1}(x_{i-1}),x_i)\in Q_E & \text{if } i-1=\ell\end{cases}$$

for $h+1\le i\le h+r$ (in particular, $x_i\in\{0,1\}^n$ and $k_i\in\{0,1\}^\kappa$ for each x_i, k_i) is called an *r-chain of τ starting at index h* or simply an *r-chain of τ*. Moreover, an r-chain is said to *fit* τ if $k_{h+i}=k_{h+i}^*$ for $1\le i\le r$, indices taken mod ℓ and in the range $\{1,\ldots,\ell\}$. We sometimes commit a slight abuse of language by saying that a chain "fits k^*" instead of "fits τ" when it is clear which transcript τ is intended.

By means of emphasis, a chain which doesn't (necessarily) fit the key of τ is said to be *generic*; thus all r-chains of τ are by definition generic.

The rest of the proof in a nutshell. Broadly, our "bad transcripts" are transcripts that either have a bad key (i.e., $k_i^*=k_j^*$ for some $i\neq j$) or transcripts with too many (long) fitting chains, where "too many" depends geometrically on the chain length r, as might be expected. When there are not too many long chains that fit the transcript's key, indeed, we are in a position to apply the lemma of Chen and Steinberger [8] to show that the probability of obtaining the given transcript in the real world is not far off from the probability of obtaining the same transcript in the ideal world, as required by (8).

The main technical challenge that arises is that of upper bounding the probability of obtaining too many length r chains that fit the key. Here one must emphasize that this probability (which is

[11] Slightly more formally—but less intuitively—an oracle (or "environment") ω is compatible with a transcript τ if there exists *some* (wlog, deterministic) adversary A' that produces τ as transcript when given ω as oracle.

the probability of obtaining a "bad" transcript) is being computed in the ideal world. In the ideal world, the key value $k^* \in \{0,1\}^{\kappa\ell}$ is chosen at random *after* all queries are completed. Hence, by a Markov bound, it suffices to show that, with high probability, not too many *generic* r-chains are created by the adversary's queries. While [4, 15] are up against the same challenge, we deliver a tight bound on the number of generic chains by using a fairly simple argument, as already discussed in the paper's introduction (see in particular Proposition 1 in Section 6).

The details of all this are implemented in Section 6.

6 Proof of Theorem 3

For the remainder of the proof of Theorem 3 we will assume that $n \geq 2$ and also, if $q \geq 2^n$, that

$$4Cq \leq 2^{\kappa+n} \qquad \text{and} \qquad 2^n \left(\frac{q}{2^{\kappa+n}}\right)^{\ell'/2} < 1. \tag{10}$$

These assumptions are without loss of generality because Theorem 3 is void otherwise, as can easily be checked. We also let $N = 2^n$.

We start by making a few more definitions that will be useful for the definition of bad transcripts and thereafter. Firstly, for a transcript $\tau = (\star, Q_E, k^*)$ we let Q_E^+, Q_E^- be the sets of queries in Q_E obtained respectively by *forward* and *backward* queries to E by the adversary. (To wit, a query to E is forward, a query to E^{-1} is backward.) We note that while Q_E does not explicitly encode forward/backward information by design, such information can be uniquely reconstructed from Q_E given the fact that A is deterministic; hence, this information is implicitly contained in Q_E. Moreover, we note that $Q_E^+ \cap Q_E^- = \emptyset$ by the fact that A never makes redundant queries, so Q_E is the disjoint union of Q_E^+ and Q_E^-.

The *maximum forward query occupancy* of τ, denoted $\mathsf{fwd}(\tau)$, is given by

$$\mathsf{fwd}(\tau) := \max_{y_0 \in \{0,1\}^n} |\{(k,x,y) \in Q_E^+ : y = y_0\}| \tag{11}$$

and $\mathsf{bwd}(\tau)$, the *maximum backward query occupancy*, is similarly given by

$$\mathsf{bwd}(\tau) = \max_{x_0 \in \{0,1\}^n} |\{(k,x,y) \in Q_E^- : x = x_0\}|.$$

We also define

$$\mathsf{fitkey}(\tau, r, h)$$

as the number of r-chains in τ that fit k^* and that start at position h.

Note that back-of-the-envelope computations suggest that $\mathsf{fwd}(\tau)$, $\mathsf{bwd}(\tau)$ should be around q/N for $q \geq N = 2^n$ and should be around $\log(q) \leq n$ for $q \leq N$. This motivates the definition of the following threshold $\zeta(q)$:

$$\zeta(q) := \begin{cases} 7nq/N & \text{if } q \geq N, \\ 6\log(q) + 2 & \text{if } q \leq N. \end{cases}$$

For now, the factors $7n$, $6\log(q) + 2$ that appear in the definition of $\zeta(q)$ should be more or less ignored; these coefficients are necessary to make bad transcripts, as defined next, unlikely. In fact, we find it convenient to factor $\zeta(q)$ into "essential" an "non-essential" parts $\zeta'(q)$ and $\zeta''(q)$:

$$\zeta''(q) = \begin{cases} 7n & \text{if } q \geq N, \\ 6\log(q) + 2 & \text{if } q \leq N. \end{cases} \qquad \zeta'(q) = \begin{cases} q/N & \text{if } q \geq N, \\ 1 & \text{if } q \leq N. \end{cases} \tag{12}$$

Thus $\zeta(q) = \zeta''(q)\zeta'(q)$. Note also that $\zeta(q) \leq 2^\kappa$ by the wlog assumptions made in (10).

Bad transcripts. We say that a transcript $\tau = (\star, Q_E, k^*)$ is *bad* if either (i) $k_i^* = k_j^*$ for some $i \neq j$, or (ii) $\mathsf{fwd}(\tau) \geq \zeta(q)$ or $\mathsf{bwd}(\tau) \geq \zeta(q)$, or (iii) there exists some r, $1 \leq r \leq \ell$ and some h, $0 \leq h \leq \ell - 1$ such that

$$\mathsf{fitkey}(\tau, r, h) \geq C z_r.$$

where

$$z_r := \min\{q, N\} \cdot \left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil r/2 \rceil}. \tag{13}$$

We let $\mathcal{T}_2$ be the set of bad transcripts, and let $\mathcal{T}_1 = \mathcal{T} \backslash \mathcal{T}_2$. One can note that every transcript with $\star = \bot$ is a bad transcript, since in that case $k_\ell^* = k_i^*$ for some $i \neq \ell$.

Bounding the probability of bad transcripts. Here we attach ourselves to upper bounding $\Pr[Y \in \mathcal{T}_2]$, as required by the H-coefficient technique. This is the probability of obtaining a bad transcript in the ideal world.

The probability that two subkeys of k^* are equal is obvioulsy at most $\binom{\ell}{2} 2^{-\kappa} \leq \ell^2/2^{\kappa+1}$. For the other two events we need the help of the following lemmas:

Lemma 1. *One has*

$$\Pr_{\tau \sim Y}[\mathsf{fwd}(\tau) \geq \zeta(q)] \leq \frac{2}{N} \qquad \textit{and} \qquad \Pr_{\tau \sim Y}[\mathsf{bwd}(\tau) \geq \zeta(q)] \leq \frac{2}{N}$$

for all q, n.

(Here $\Pr_{\tau \sim Y}$ indicates that τ is sampled according to the ideal world distribution on transcripts. The same probabilities could equivalently be written $\Pr[\mathsf{fwd}(Y) \geq \zeta(q)]$, $\Pr[\mathsf{bwd}(Y) \geq \zeta(q)]$.)

Lemma 2. *One has*

$$\Pr_{\tau \sim Y}[\mathsf{fitkey}(\tau, r, h) \geq C z_r \wedge \mathsf{fwd}(\tau) \leq \zeta(q) \wedge \mathsf{bwd}(\tau) \leq \zeta(q)] \leq \frac{2^r \zeta''(q)^{\lceil r/2 \rceil}}{C}$$

for each $1 \leq r \leq \ell$, $0 \leq h \leq \ell - 1$ with z_r as defined in (13)*.*

Combining lemmas 1 and 2 we directly obtain by a union bound that

$$\Pr[Y \in \mathcal{T}_2] \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{N} + \frac{\ell^2 2^\ell \zeta''(q)^{\lceil \ell/2 \rceil}}{C} \tag{14}$$

since there are ℓ^2 choices for the pair (r, h).

The proof of Lemma 1 is found in Appendix B. Even while the proof of this lemma ultimately relies on basic balls-in-bins statistics, we note that subtleties arise due to the fact that A is querying *permutations* on keys of its choice, and that a permutation of $\{0,1\}^n$ "loses randomness" after $\approx 2^n$ queries. In particular, the proof of Lemma 2 requires the "super-query" technique developed for beyond-message-length-security [3, 18, 21] and, to achieve something fully formal, some combinatorial results related to partitions (see Appendix B).

The proof of Lemma 2 is, despite appearances, significantly simpler than the proof of Lemma 1. The main component is the following proposition (which happens to be a key part of our proof and which significantly sharpens similar bounds found in [4, 15]):

Proposition 1. *Assume $\tau = (\star, Q_E, k^*)$ is a q-query transcript such that $\mathsf{fwd}(\tau) \le \zeta(q)$, $\mathsf{bwd}(\tau) \le \zeta(q)$. Then the total number of r-chains of τ starting at position h is at most*

$$2^r \cdot \min\{q, N\} \cdot \zeta(q)^{\lceil r/2 \rceil} 2^{\kappa \lfloor r/2 \rfloor}.$$

Proof. Let $\nu = (h, x_h, k_{h+1}, x_{h+1}, \ldots, k_{h+r}, x_{h+r})$ be an r-chain of τ. Thus either $(k_i, x_{i-1}, x_i) \in Q_E^+$ or $(k_i, x_{i-1}, x_i) \in Q_E^-$ for $h+1 \le i \le h+r$. Let ν's *signature* be the string $sig^\nu \in \{+,-\}^r$ such that $(k_i, x_{i-1}, x_i) \in Q_E^{sig_i^\nu}$ for $h+1 \le i \le h+r$.

We start by fixing a signature $sig^0 \in \{+,-\}^r$ and by upper bounding the number of r-chains ν of τ starting at position h such that $sig^\nu = sig^0$. We can assume without loss of generality that sig^0 contains at least as many $-$'s as $+$'s, i.e., that the number of $-$'s is at least $\lceil r/2 \rceil$.

If $\nu = (h, x_h, k_{h+1}, x_{h+1}, \ldots, k_{h+r}, x_{h+r})$ is a ν-chain with signature sig^0 then there are, firstly, at most

$$\min\{q, N\}$$

choices for x_h given that $(k_{h+1}, x_h, x_{h+1}) \in Q_E$. Then, presuming x_h fixed, there are at most 2^κ choices for x_{h+1} if $sig_1^0 = +$ and at most $\zeta(q)$ choices for x_{h+1} if $sig_1^0 = -$, given that τ is a transcript such that $\mathsf{bwd}(\tau) \le \zeta(q)$. Similarly, each subsequent step introduces a factor of either 2^κ or $\zeta(q)$ depending on the sign of that step in sig^0. Hence (and since $2^\kappa \ge \zeta(q)$) the total number of choices for $x_h, k_{h+1}, \ldots, x_{h+r}$ is at most

$$\min\{q, N\} \cdot \zeta(q)^{\lceil r/2 \rceil} 2^{\kappa \lfloor r/2 \rfloor}.$$

Multiplying by 2^r to account for all possible signatures concludes the proof. □

Since $\Pr[A \wedge B] \le \Pr[A|B]$ we have

$$\begin{aligned} &\Pr_{\tau \sim Y}[\mathsf{fitkey}(r, \tau) \ge C z_r \wedge \mathsf{fwd}(\tau) \le \zeta(q) \wedge \mathsf{bwd}(\tau) \le \zeta(q)] \\ \le &\Pr_{\tau \sim Y}[\mathsf{fitkey}(r, \tau) \ge C z_r \mid \mathsf{fwd}(\tau) \le \zeta(q) \wedge \mathsf{bwd}(\tau) \le \zeta(q)] \end{aligned}$$

where z_r is the quantity defined in Lemma 2. When we condition on $\mathsf{fwd}(\tau) \le \zeta(q) \wedge \mathsf{bwd}(\tau) \le \zeta(q)$, however, k^* is still independent uniformly at random (being entirely independent from Q_E in the ideal world), and so the expected number of r-chains that fit τ at position h is upper bounded by

$$2^r \cdot \min\{q, N\} \cdot \zeta(q)^{\lceil r/2 \rceil} 2^{\kappa \lfloor r/2 \rfloor} \frac{1}{2^{\kappa r}} \tag{15}$$

by Proposition 1. (Each r-chain of Q_E, indeed, has probability of exactly $1/2^{\kappa r}$ of being "hit" by k^*.) Since $r - \lfloor r/2 \rfloor = \lceil r/2 \rceil$, (15) can be written

$$2^r \zeta''(q)^{\lceil r/2 \rceil} \min\{q, N\} \left(\frac{\zeta'(q)}{2^\kappa} \right)^{\lceil r/2 \rceil} = 2^r \zeta''(q)^{\lceil r/2 \rceil} z_r$$

with z_r as defined in (13). It thus follows by Markov's inequality that

$$\Pr_{\tau \sim Y}[\mathsf{fitkey}(r, \tau) \ge C z_r \mid \mathsf{fwd}(\tau) \le \zeta(q) \wedge \mathsf{bwd}(\tau) \le \zeta(q)] \le \frac{2^r \zeta''(q)^{\lceil r/2 \rceil}}{C}$$

which proves Lemma 2 and inequality (14).

Good transcripts. The rest of the proof consists in lower-bounding the ratio

$$\frac{\Pr[X = \tau]}{\Pr[Y = \tau]}$$

for all $\tau \in \mathcal{T}_1$ such that $\Pr[Y = \tau] > 0$. We recall this is equivalent to lower bounding the ratio

$$\frac{|\mathsf{comp}_X(\tau)|}{|\Omega_X|} \Big/ \frac{|\mathsf{comp}_Y(\tau)|}{|\Omega_Y|} \tag{16}$$

by (9).

Fix $\tau = (\star, Q_E, k^*) \in \mathcal{T}_1$ such that $\Pr[Y = \tau] > 0$. (Note $\star = \top$ because $\tau \in \mathcal{T}_1$.) Our approach for lower-bounding (16) is very similar to that of [8].

Let $\mathsf{comp}'_X(\tau) \subseteq \Omega_X$ be the set of real-world oracles that are compatible with τ on everything except for the queries in Q_E with key k^*_ℓ. (In more detail, $\omega = (E'', k'') \in \Omega_X$ is in $\mathsf{comp}'_X(\tau)$ if and only if $k'' = k^*$ and if $E''_k(x) = y$ for all $(k, x, y) \in Q_E$ such that $k \neq k^*_\ell$.) We similarly define $\mathsf{comp}'_Y(\tau)$ to be the set of all ideal-world oracles that are compatible with τ on everything except for the queries to E with key k^*_ℓ. It is easy to see that

$$\frac{|\mathsf{comp}'_X(\tau)|}{|\Omega_X|} = \frac{|\mathsf{comp}'_Y(\tau)|}{|\Omega_Y|}$$

from the fact that k^* and the set of random permutations $\{E_k : k \neq k^*_\ell\}$ are independent in both the real and ideal worlds. (Alternatively, one can show by direct counting that $|\mathsf{comp}'_X(\tau)| = |\mathsf{comp}'_Y(\tau)|$. Also note that $|\Omega_X| = |\Omega_Y|$.) It thus suffices to lower bound

$$\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|} \Big/ \frac{|\mathsf{comp}_Y(\tau)|}{|\mathsf{comp}'_Y(\tau)|} \tag{17}$$

in order to lower bound (16).

Let $q_\ell = |p_\ell|$ be the number of queries with key k^*_ℓ that appear in Q_E. For the ideal world, it is easy to see that

$$\frac{|\mathsf{comp}_Y(\tau)|}{|\mathsf{comp}'_Y(\tau)|} = \prod_{i=0}^{q_\ell - 1} \frac{1}{2^n - i}. \tag{18}$$

For the real world the situation is more complicated, but the basic idea is to view

$$\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|} \tag{19}$$

as the probability that when a random $\omega \in \mathsf{comp}'_X(\tau)$ is sampled, this ω happens to also be compatible with the q_ℓ queries in τ of key k^*_ℓ. (This probability is to be compared with (18).) Note that sampling $\omega \in \mathsf{comp}'_X(\tau)$ can be viewed as randomly extending a set of already partially-defined random permutations $\{E_k : k \neq k^*_\ell\}$ (partially defined, that is, by the queries that appear in Q_E). Note moreover that whether such a random extension is "successful" in the sense of producing an oracle compatible with τ depends only on the extension of the permutations $\{E_{k^*_i} : 1 \leq i \leq \ell - 1\}$, since queries with key k^*_ℓ are (in the real world) answered as a function of those permutations. Details follow.

Graph view. To enable us to reason combinatorially we introduce a "graph view" of the problem.

For $1 \leq i \leq \ell$ let p_i be the set of queries in Q_E with key k_i^*; more precisely, $p_i = \{(x, y) : (k_i^*, x, y) \in Q_E\}$. We associate a bipartite graph G_i to p_i, where G_i has shores $\{0,1\}^n$ and $\{0,1\}^n$; an edge connects $u \in \{0,1\}^n$ to $v \in \{0,1\}^n$ in G_i if and only if $(u, v) \in p_i$. (Thus G_i is a partial matching from $\{0,1\}^n$ to $\{0,1\}^n$, by the permutation structure of $E_{k_i^*}$.) Moreover, define

$$\tilde{p}_\ell = \{(S^{-1}(y), x) : (x, y) \in p_\ell\}.$$

Thus $|\tilde{p}_\ell| = |p_\ell| = q_\ell$.

We introduce a new graph $G(\tau)$ that, essentially, "glues" the graphs $G_1, \ldots, G_{\ell-1}$ together. $G(\tau)$ has $2^n\ell$ vertices grouped into ℓ "shores", which each shore being identified with $\{0,1\}^n$. We number the ℓ shores as 0, 1, …, $\ell - 1$. We place a copy of G_i between shores $i - 1$ and i for $1 \leq i \leq \ell - 1$; more precisely, vertices x in shore $i - 1$ and y in shore i are joined if and only if $(x, y) \in p_i$. We moreover equate G_i with its "copy" in $G(\tau)$ for $1 \leq i \leq \ell - 1$.

A *path from shore i to shore j* in $G(\tau)$, $0 \leq i < j \leq \ell - 1$, is just what it sounds like: a sequence of vertices $(x_i, x_{i+1}, \ldots, x_j)$ such that x_i is in shore i and x_j is in shore j and such that (x_h, x_{h+1}) is an edge of $G(\tau)$ for $i \leq h \leq j - 1$. (This automatically implies that x_h is in shore h.) The *length* of the path is $j - i$. We emphasize that we do not require a path to be maximal, i.e., one path can be a subset of another, longer path.

Let $(\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_{q_\ell}, \tilde{y}_{q_\ell})$ be the q_ℓ elements of $\tilde{p}_\ell$. We view each $\tilde{x}_i$ as a vertex in shore 0 of $G(\tau)$ and each $\tilde{y}_i$ as a vertex in shore $\ell - 1$ of $G(\tau)$. We emphasize that we do not place an edge between $\tilde{x}_i$ and $\tilde{y}_i$. Also let $\mathsf{R}(\tilde{x}_i)$ be the rightmost vertex in $G(\tau)$ reachable by a path of edges from $\tilde{x}_i$ in $G(\tau)$ (shores with higher indices being "to the right" of shores with lower indices), and, similarly, let $\mathsf{L}(\tilde{y}_i)$ be the leftmost vertex in $G(\tau)$ reachable from $\tilde{y}_i$. For example, $\mathsf{R}(\tilde{x}_i) = \tilde{x}_i$ if $\tilde{x}_i$ isn't adjacent to any edges in $G(\tau)$. Finally, let $\mathrm{Sh}(u) \in \{0, 1, \ldots, \ell - 1\}$ be the index of the shore of a vertex u in $G(\tau)$.

Note that every path of length r in $G(\tau)$ corresponds to a distinct r-chain in τ that fits k^*. Moreover, for each $1 \leq i \leq q_\ell$, there corresponds a distinct r-chain of length

$$\mathrm{Sh}(\mathsf{R}(\tilde{x}_i)) + \ell - \mathrm{Sh}(\mathsf{L}(\tilde{y}_i))$$

that fits k^* as well.

Incidentally, note that

$$Cz_\ell = C \min\{q, N\} \left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil \ell/2 \rceil} = C \cdot \begin{cases} N \left(\frac{q}{2^\kappa N}\right)^{\ell'/2} & \text{if } q \geq N, \\ q \left(\frac{1}{2^\kappa}\right)^{\ell'/2} & \text{if } q \leq N \end{cases} \tag{20}$$

is less than 1; for $q \geq N$ this follows from our initial observation (cf. (10)) that Theorem 3 is void otherwise whereas for $q \leq N$ we recall that $Cq < 2^{\kappa\ell'/2}$ is one of the theorem's assumptions. It follows there are no r-chains of length[12] ℓ that fit τ, by virtue of τ being a good transcript. In particular $G(\tau)$ contains no paths of length ℓ and, also,

$$\mathrm{Sh}(\mathsf{R}(\tilde{x}_i)) + \ell - \mathrm{Sh}(\mathsf{L}(\tilde{y}_i)) \leq \ell - 1$$

[12] In the extreme case $\ell = 1$ one can observe that $G(\tau)$ has only one shore, that $q_\ell = 0$ and that $|\mathsf{comp}_X(\tau)|/|\mathsf{comp}'_X(\tau)| = |\mathsf{comp}_Y(\tau)|/|\mathsf{comp}'_Y(\tau)| = 1$. The adversary's advantage is then upper bounded by (14), but still subject to the condition that $Cq/2^\kappa < 1$ in order that (20) remain upper bounded by 1. Of course, tighter analyses can be made for the special case $\ell = 1$.

for each $1 \leq i \leq q_\ell$, i.e.,

$$\mathrm{Sh}(\mathsf{L}(\tilde{y}_i)) - \mathrm{Sh}(\mathsf{R}(\tilde{x}_i)) \geq 1$$

for $1 \leq i \leq q_\ell$.

More definitions: a vertex in shore i of $G(\tau)$ is *left-free* if it is not adjacent to a vertex in shore $i-1$. It is *right-free* if it is not adjacent to a vertex in shore $i+1$.

We can evaluate (19) via the following experiment in q_ℓ stages. Each stage modifies the graph $G(\tau)$, and we write $G(\tau)^i$ for the graph as it stands after the i-th stage, with $G(\tau)^0 = G(\tau)$. At the i-th stage, we randomly extend the path ending at $\mathsf{R}(\tilde{x}_i)$ by choosing a vertex uniformly at random among the left-free vertices of shore $\mathrm{Sh}(\mathsf{R}(\tilde{x}_i)) + 1$, and connecting this vertex to $\mathsf{R}(\tilde{x}_i)$; this new edge changes the value of $\mathsf{R}(\tilde{x}_i)$, and we repeat this step until $\mathrm{Sh}(\mathsf{R}(\tilde{x}_i)) \geq \mathrm{Sh}(\mathsf{L}(\tilde{y}_i))$; this defines the graph $G(\tau)^i$ from $G(\tau)^{i-1}$; if $\tilde{x}_i$ and $\tilde{y}_i$ are connected by a path (i.e., $\mathsf{R}(\tilde{x}_i) = \tilde{y}_i$, $\mathsf{L}(\tilde{y}_i) = \tilde{x}_i$) in $G(\tau)^i$ then we proceed to stage $i+1$; if not, we abort the experiment and declare failure.

We write

$$G(\tau)^i \downarrow \tilde{p}_\ell$$

for the event that stages 1 through i complete successfully in the above experiment, where $0 \leq i \leq q_\ell$, and we write

$$\Pr[G(\tau)^{i+1} \downarrow \tilde{p}_\ell \,|\, G(\tau)^i \downarrow \tilde{p}_\ell]$$

for the probability that the $(i+1)$-th stage of the experiment succeeds, given that the first i stages have succeeded. It is easy to see that

$$\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|} = \Pr[G(\tau)^{q_\ell} \downarrow \tilde{p}_\ell]$$

and, thus,

$$\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|} = \prod_{i=0}^{q_\ell - 1} \Pr[G(\tau)^{i+1} \downarrow \tilde{p}_\ell \,|\, G(\tau)^i \downarrow \tilde{p}_\ell].$$

Also note that $\mathsf{R}(\tilde{x}_{i+1})$, $\mathsf{L}(\tilde{y}_{i+1})$ remain unchanged for stages 0 through i as long as $G(\tau)^i \downarrow \tilde{p}_\ell$. In fact, assuming $G(\tau)^i \downarrow \tilde{p}_\ell$, the first i successfully completed paths constitute "dead weight" in $G(\tau)^i$ in the sense that these paths can be entirely removed from $G(\tau)^i$ without affecting future probabilities of path completions. It will indeed prove convenient to remove successfully completed paths as we go along, which means that $G(\tau)^i$ (still assuming $G(\tau)^i \downarrow \tilde{p}_\ell$) becomes a graph with $N-i$ vertices in each shore. Moreover, this means there remains an injective correspondence from the set of paths of length r and starting in shore h of $G(\tau)^i$ to the set of r-chains of τ fitting k^* and starting at position h, since we have removed from $G(\tau)^i$ all edges added during previous path completions.

With these explanations in place we can state Chen and Steinberger's lemma [8], based on an inclusion-exclusion principle:

Lemma 3. (Chen and Steinberger [8]) *One has*

$$\Pr[G(\tau)^{j+1} \downarrow \tilde{p}_\ell \,|\, G(\tau)^j \downarrow \tilde{p}_\ell] = \frac{1}{N-j} - \frac{1}{N-j} \sum_{\sigma} (-1)^{|\sigma|} \prod_{h=1}^{|\sigma|} \frac{|U_{i_h i_{h-1}}|}{N-j-|E_{i_h}|} \tag{21}$$

where the sum is taken over all sequences $\sigma = (i_0, \ldots, i_s)$ with $\mathsf{R}(\tilde{x}_{j+1}) = i_0 < i_1 < \ldots < i_s = \mathsf{L}(\tilde{y}_{j+1})$, where $|\sigma| = s$, where U_{uv} denotes the set of paths in $G(\tau)^j$ from shore u to shore v that *start at a left-free vertex in shore u and where E_i is the set of edges of $G(\tau)^j$ from shore $i-1$ to shore i of $G(\tau)^j$.*

Computations. The object, now, is to upper bound the sum which appears in (21), in order to lower bound (21). As a preliminary, note that

$$|U_{ij}| \leq Cz_{j-i} = C\min\{q,N\}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil (j-i)/2\rceil}$$

with z_r as defined in (13) because $|U_{ij}|$ is at most the number of paths of length $j-i$ in $G(\tau)$ starting in shore i, and because τ is a good transcript. Moreover

$$|E_i| \leq Cz_1 = C\min\{q,N\}\frac{\zeta'(q)}{2^\kappa} = \frac{Cq}{2^\kappa} \leq N/4$$

by (10) for $q \geq N$ and by the assumption $Cq \leq 2^{\kappa+n-2}$ (explicit in Theorem 3) for $q \leq N$. Let $\mathsf{Odd} : \mathbb{Z} \to \{0,1\}$ be the indicator function for odd integers: $\mathsf{Odd}(x) = (x \bmod 2)$. Fixing a value of j, $0 \leq j \leq q_\ell - 1$ and letting $t = \mathsf{L}(\tilde{y}_{j+1}) - \mathsf{R}(\tilde{x}_{j+1})$ then, with notations as in Lemma 3, and since also $q_\ell \leq Cq/2^\kappa \leq N/4$,

$$\begin{aligned}
&\sum_\sigma (-1)^{|\sigma|}\prod_{h=1}^{|\sigma|}\frac{|U_{i_h i_{h-1}}|}{N-j-|E_{i_h}|}\\
&\leq \sum_\sigma \prod_{h=1}^{|\sigma|}\frac{|U_{i_h i_{h-1}}|}{N-N/4-N/4}\\
&\leq \sum_\sigma \prod_{h=1}^{|\sigma|}\frac{C\min\{q,N\}\,(\zeta'(q)/2^\kappa)^{\lceil (i_h-i_{h-1})/2\rceil}}{N/2}\\
&= \sum_\sigma \left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\prod_{h=1}^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{(i_h-i_{h-1})/2+\mathsf{Odd}(i_h-i_{h-1})/2}\\
&= \sum_\sigma \left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{t/2+\sum_{h=1}^{|\sigma|}\mathsf{Odd}(i_h-i_{h-1})/2}
\end{aligned}$$

If $|\sigma| \geq \lceil t/2\rceil$ then there are at least $2|\sigma|-t$ different indices h such that $i_h - i_{h-1} = 1$; in this case, therefore, and because $\zeta'(q)/2^\kappa \leq 1$,

$$\begin{aligned}
&\left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{t/2+\sum_{h=1}^{|\sigma|}\mathsf{Odd}(i_h-i_{h-1})/2}\\
&\leq \left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{t/2+(2|\sigma|-t)/2}\\
&= \left(\frac{2C\min\{q,N\}\zeta'(q)}{2^\kappa N}\right)^{|\sigma|}\\
&= \left(\frac{2Cq}{2^\kappa N}\right)^{|\sigma|}
\end{aligned}$$

$$\leq \left(\frac{2Cq}{2^\kappa N}\right)^{\lceil t/2\rceil}. \tag{22}$$

On the other hand if $|\sigma| \leq \lfloor t/2\rfloor$ then there exists an h such that $i_h - i_{h-1}$ is odd if t is odd; thus

$$\begin{aligned}
&\left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{t/2+\sum_{h=1}^{|\sigma|}\mathsf{Odd}(i_h-i_{h-1})/2}\\
&\leq \left(\frac{2C\min\{q,N\}}{N}\right)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil t/2\rceil}\\
&\leq \frac{\min\{q,N\}}{N}(2C)^{|\sigma|}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil t/2\rceil}\\
&\leq \frac{\min\{q,N\}}{N}(2C)^{\lfloor t/2\rfloor}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil t/2\rceil}\\
&\leq \frac{\min\{q,N\}}{N}\left(\frac{2C\zeta'(q)}{2^\kappa}\right)^{\lceil t/2\rceil}. \qquad (23)
\end{aligned}$$

If $q \geq N$ then (23) equals $(2Cq/2^\kappa N)^{\lceil t/2\rceil}$, which yields

$$\sum_\sigma (-1)^{|\sigma|}\prod_{h=1}^{|\sigma|}\frac{|U_{i_h i_{h-1}}|}{N-j-|E_{i_h}|} \leq \sum_\sigma \left(\frac{2Cq}{2^\kappa N}\right)^{\lceil t/2\rceil} \leq 2^t\left(\frac{2Cq}{2^\kappa N}\right)^{\lceil t/2\rceil}.$$

On the other hand if $q \leq N$ then (22), (23) are both upper bounded by $(q/N)(2C/2^\kappa)^{\lceil t/2\rceil}$ which gives us

$$\sum_\sigma (-1)^{|\sigma|}\prod_{h=1}^{|\sigma|}\frac{|U_{i_h i_{h-1}}|}{N-j-|E_{i_h}|} \leq \sum_\sigma \frac{q}{N}\left(\frac{2C}{2^\kappa}\right)^{\lceil t/2\rceil} \leq \frac{2^t q}{N}\left(\frac{2C}{2^\kappa}\right)^{\lceil t/2\rceil}. \tag{24}$$

Let $\mathcal{L}_t \subseteq \{1,\ldots,q_\ell\}$ be the set of indices j such that $\mathsf{L}(\tilde{y}_j) - \mathsf{R}(\tilde{x}_j) = t$, $1 \leq t \leq \ell - 1$. Note that because τ is good,

$$|\mathcal{L}_t| \leq \ell C z_{\ell-t} = \ell C \min\{q,N\}\left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil(\ell-t)/2\rceil}.$$

Also write $\mathfrak{C}_j$ for the set of sequences σ over which summation takes place in Lemma 3, and similarly let U^j_{uv}, E^j_i stand for U_{uv}, E_i in $G(\tau)^j$. Then for $q \geq N$ we find

$$\begin{aligned}
\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|}\Big/\frac{|\mathsf{comp}_Y(\tau)|}{|\mathsf{comp}'_Y(\tau)|} &= \frac{\prod_{j=0}^{q_\ell-1}\Pr[G(\tau)^{j+1}\downarrow\tilde{p}_\ell \mid G(\tau)^j\downarrow\tilde{p}_\ell]}{\prod_{j=0}^{q_\ell-1}1/(N-j)}\\
&= \prod_{j=0}^{q_\ell-1}\left(1-\sum_{\sigma\in\mathfrak{C}_j}(-1)^{|\sigma|}\prod_{h=1}^{|\sigma|}\frac{|U^j_{i_h i_{h-1}}|}{N-j-|E^j_{i_h}|}\right)\\
&= \prod_{t=1}^{\ell-1}\prod_{j\in\mathcal{L}_t}\left(1-\sum_{\sigma\in\mathfrak{C}_j}(-1)^{|\sigma|}\prod_{h=1}^{|\sigma|}\frac{|U^j_{i_h i_{h-1}}|}{N-j-|E^j_{i_h}|}\right)
\end{aligned}$$

$$
\begin{aligned}
&\geq \prod_{t=1}^{\ell-1} \prod_{j \in \mathcal{L}_t} \left(1 - 2^t \left(\frac{2Cq}{2^\kappa N}\right)^{\lceil t/2 \rceil}\right) \\
&= \prod_{t=1}^{\ell-1} \left(1 - 2^t \left(\frac{2Cq}{2^\kappa N}\right)^{\lceil t/2 \rceil}\right)^{\ell C N \left(\frac{q}{2^\kappa N}\right)^{\lceil (\ell-t)/2 \rceil}} \\
&\geq 1 - \sum_{t=1}^{\ell-1} 2^t \ell C N (2C)^{\lceil t/2 \rceil} \left(\frac{q}{2^\kappa N}\right)^{\lceil t/2 \rceil + \lceil (\ell-t)/2 \rceil} \\
&\geq 1 - \sum_{t=1}^{\ell-1} 2^{t + \lceil t/2 \rceil} \ell C N C^{\lceil (\ell-1)/2 \rceil} \left(\frac{q}{2^\kappa N}\right)^{\lceil \ell/2 \rceil} \\
&\geq 1 - 2^{\ell + \lceil (\ell-1)/2 \rceil} \ell C^{\lceil (\ell+1)/2 \rceil} N \left(\frac{q}{2^\kappa N}\right)^{\lceil \ell/2 \rceil} \\
&\geq 1 - \ell C^{\lceil (\ell+1)/2 \rceil} N \left(\frac{8q}{2^{\kappa+n}}\right)^{\lceil \ell/2 \rceil}
\end{aligned}
$$

Which, since $\lceil \ell/2 \rceil = \ell'/2$, and in conjunction with the upper bound (14) on the probability of a bad transcript, gives us the first part of Theorem 3.

On the other hand if $q \leq N$ then

$$
|\mathcal{L}_t| \leq \ell C \min\{q, N\} \left(\frac{\zeta'(q)}{2^\kappa}\right)^{\lceil (\ell-t)/2 \rceil} \leq \ell C q \left(\frac{1}{2^\kappa}\right)^{\lceil (\ell-t)/2 \rceil}
$$

using (24) which gives us

$$
\begin{aligned}
\frac{|\mathsf{comp}_X(\tau)|}{|\mathsf{comp}'_X(\tau)|} \Bigg/ \frac{|\mathsf{comp}_Y(\tau)|}{|\mathsf{comp}'_Y(\tau)|} &= \prod_{t=1}^{\ell-1} \prod_{j \in \mathcal{L}_t} \left(1 - \sum_{\sigma \in \mathfrak{C}_{j+1}} (-1)^{|\sigma|} \prod_{h=1}^{|\sigma|} \frac{|U_{i_h i_{h-1}}|}{N - j - |E_{i_h}|}\right) \\
&\geq \prod_{t=1}^{\ell-1} \left(1 - \frac{2^t q}{N} \left(\frac{2C}{2^\kappa}\right)^{\lceil t/2 \rceil}\right)^{|\mathcal{L}_t|} \\
&\geq \prod_{t=1}^{\ell-1} \left(1 - \frac{2^t q}{N} \left(\frac{2C}{2^\kappa}\right)^{\lceil t/2 \rceil}\right)^{\ell C q \left(\frac{1}{2^\kappa}\right)^{\lceil (\ell-t)/2 \rceil}} \\
&\geq 1 - \frac{q}{N} \sum_{t=1}^{\ell-1} 2^t \ell C q \, (2C)^{\lceil t/2 \rceil} \left(\frac{1}{2^\kappa}\right)^{\lceil t/2 \rceil + \lceil (\ell-t)/2 \rceil} \\
&\geq 1 - \frac{q^2}{N} \sum_{t=1}^{\ell-1} 2^{t + \lceil t/2 \rceil} \ell C^{1 + \lceil (\ell-1)/2 \rceil} \left(\frac{1}{2^\kappa}\right)^{\lceil \ell/2 \rceil} \\
&\geq 1 - \frac{q^2 \ell}{N} 2^{\ell + \lceil (\ell-1)/2 \rceil} C^{\lceil (\ell+1)/2 \rceil} \left(\frac{1}{2^\kappa}\right)^{\lceil \ell/2 \rceil} \\
&\geq 1 - \frac{q^2 \ell}{N} C^{\lceil (\ell+1)/2 \rceil} \left(\frac{8}{2^\kappa}\right)^{\lceil \ell/2 \rceil}
\end{aligned}
$$

which gives us the second part of Theorem 3, together with (14).

References

1. William Aiello, Mihir Bellare, Giovanni Di Crescenzo, and Ramarathnam Venkatesan. Security amplification by composition: the case of doubly-iterated, ideal ciphers, CRYPTO 1998, LNCS 1462, pp. 390–407.
2. ANSI X9.52: Triple Data Encryption Algorithm Modes of Operation (withdrawn), 1998.
3. Frederik Armknecht, Ewan Fleischmann, Matthias Krause, Jooyoung Lee, Martijn Stam and John Steinberger, The preimage security of double-block length compression functions. Asiacrypt 2011, LNCS 7073, Springer, 233–251.
4. Mihir Bellare and Phillip Rogaway, The security of triple encryption and a framework for code-based game-playing proofs. Eurocrypt 2006, LNCS 4004 pp409–426.
5. Mihir Bellare and Phillip Rogaway, Code-based game-playing proofs and the security of triple encryption. IACR eprint report. `eprint.iacr.org/2004/331`
6. John Black, Phillip Rogaway, Thomas Shrimpton, Black-Box Analysis of the Block Cipher-Based Hash-Function Constructions from PGV. CRYPTO 2002, LNCS XXXX, pages 320–335.
7. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Francois-Xavier Standaert, John Steinberger and Elmar Tischhauser, Key-Alternating Ciphers in a Provable Setting: Encryption Using a Small Number of Public Permutations. EUROCRYPT 2012, LNCS 7237, pp. , Springer-Verlag, 2012.
8. Shan Chen and John Steinberger, Tight security bounds for key-alternating ciphers, IACR Cryptology ePrint Archive, 2013/222, `http://eprint.iacr.org/2013/222.pdf`.
9. Whitfield Diffie and Martin Hellman, Exhaustive cryptanalysis of the NBS data encryption standard. Computer 10 (6), 74–84, 1997.
10. Itai Dinur, Orr Dunkelman, Nathan Keller and Adi Shamir, Efficient Dissection of Composite Problems, with Applications to Cryptanalysis, Knapsacks, and Combinatorial Search Problems. CRYPTO 2012, LNCS 7417, pp. 719–740.
11. Shimon Even and Oded Goldreich, On the power of cascade ciphers. ACM Transactions on Computer Systems, vol. 3, no. 2, pp. 108–116, 1985.
12. Shimon Even and Yishay Mansour, A Construction of a Cipher From a Single Pseudorandom Permutation. ASIACRYPT 1991, LNCS 739, pp. 210–224, Springer-Verlag, 1993.
13. FIPS46-3: Data Encryption Standard. National Institute of Standards and Technology (withdrawn), 1999.
14. Peter Gaži, Plain versus Randomized Cascading-Based Key-Length Extension for Block Ciphers, CRYPTO 2013, LNCS 8042, pp551–570.
15. Peter Gaži and Ueli Maurer, Cascade encryption revisited, Asiacrypt 2009, LNCS 5912, pp37–51.
16. Peter Gaži and Stefano Tessaro, Efficient and Optimally Secure Key-Length Extension for Block Ciphers via Randomized Cascading. Eurocrypt 2012, LNCS 7237, pp. 63–80, Springer, Heidelberg (2012).
17. Joe Kilian and Phillip Rogaway, How to protect DES against exhaustive key search (an analysis of DESX). Journal of Cryptology 14 (1), 17-35 (2001).
18. Matthias Krause, Frederik Armknecht and Ewan Fleischmann, Preimage resistance beyond the birthday bound: Double-length hashing revisited. IACR eprint report, `http://eprint.iacr.org/2010/519.pdf`.
19. Rudolphe Lampe, Jacques Patarin and Yannick Seurin, An Asymptotically Tight Security Analysis of the Iterated Even-Mansour Cipher, Asiacrypt 2012, Lecture Notes in Computer Science Volume 7658, pp 278-295, 2012.
20. Jooyoung Lee, Towards Key-Length Extension with Optimal Security: Cascade Encryption and Xor-cascade Encryption, Eurocrypt 2013, LNCS 7881, pp405–425.
21. Jooyoung Lee, John Steinberger and Martijn Stam, The preimage security of double-block-length compression functions. IACR eprint report, `http://eprint.iacr.org/2011/210.pdf`.
22. Stefan Lucks, Attacking triple encryption. *Fast Software Encryption* 1998, LNCS 1372, pp. 239–253.
23. Ueli M. Maurer and James L. Massey, Cascade ciphers: The importance of being first. Journal of Cryptology 6(1), pp. 55–61, 1993.
24. Ralph Merkle and Martin Hellman, On the Security of Multiple Encryption, Communications of the ACM, vol. 24, no. 7, pp. 465–467, ,July 1981. See also: Communicutionr of the ACM, vol. 24, no. 11, p. 776, November 1981.
25. Paul C. van Oorschot and Michael Wiener, Improving implementable meet-in-the-middle attacks by orders of magnitude, CRYPTO 1996, LNCS 1109 pp. 229–236.
26. Paul C. van Oorschot and Michael Wiener, A Known-Plaintext Attack on Two-Key Triple Encryption, Eurocrypt 1990, LNCS 473 pp. 318–325.
27. NIST SP 800-67, Revision 1: Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher. National Institute of Standards and Technology, 2012.
28. Jacques Patarin, Etude de Génerateurs de Permutations Bases sur les Schemas du DES. In Ph.D. Thesis. Inria, Domaine de Voluceau, France, 1991.

29. Jacques Patarin, The “Coefficients H” Technique, Selected Areas in Cryptography, LNCS 5381, 2009, pp. 328-345.
30. Martin Raab and Angelika Steger, “Balls into Bins”-A Simple and Tight Analysis, RANDOM 1998, LNCS 1528, pp159–170.
31. Richard P. Stanley, *Enumerative Combinatorics.* Wadsworth & Brooks/Cole, 1986.
32. John Steinberger, The collision intractability of MDC-2 in the ideal-cipher model. Eurocrypt 2007, LNCS 4515, pp. 34–51.
33. John Steinberger, Improved Security Bounds for Key-Alternating Ciphers via Hellinger Distance, `http://eprint.iacr.org/2012/481.pdf`.

A Main Theorem Statement

Our main theorem is as follows:

Theorem 3. *If $q \geq 2^n$ then, for every real number $C \geq 1$,*

$$\mathbf{Adv}_{\ell}^{\mathsf{casc}}(q) \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{2^n} + \frac{\alpha}{C} + 2^n \ell C^{(\ell+1)'/2} \left(\frac{8q}{2^{\kappa+n}}\right)^{\ell'/2}$$

where $\alpha = \ell^2 2^\ell (7n)^{\ell'/2}$. Moreover if $q \leq 2^n$ then, for every $C \geq 1$ such that $Cq < \min\{2^{\kappa\ell'/2}, 2^{\kappa+n-2}\}$,

$$\mathbf{Adv}_{\ell}^{\mathsf{casc}}(q) \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{2^n} + \frac{\beta}{C} + \frac{q^2\ell}{2^n} C^{(\ell+1)'/2} \left(\frac{8}{2^{\kappa}}\right)^{\ell'/2}$$

where $\beta = \ell^2 2^\ell (6 \log q + 2)^{\ell'/2}$. Moreover these results also hold if the adversary is allowed to ask, for free, all possible 2^n queries to its second oracle.

The presence of the adjustable constant C is typical of security proof that involve a “bad event” whose definition is based on an adjustable threshold (this threshold being, namely, controlled by C). We note that for every value of q, κ, n and ℓ there is an *optimal* C, i.e., a C which gives the best upper bound on $\mathbf{Adv}_{\ell,\kappa,n}^{\mathsf{casc}}(q)$ in either part of Theorem 3. (The constraint $C \geq 1$ is included mostly for readability, since the bounds are vacuously true anyway for $0 < C < 1$.)

Nominally, at a glance, the second part of Theorem 3 indicates security of

$$q \approx \sqrt{2^n \cdot 2^{\kappa\ell'/2}} = 2^{\kappa\ell'/4+n/2}$$

which is *better* (higher) than $2^{\kappa\ell'/2}$ if $\kappa\ell'/2 < n$, and which thus appears to contradict the attack of Section 4. Upon closer inspection, however, the constraint $Cq < 2^{\kappa\ell'/2}$ actually prevents any security claim for $q \geq \exp(\kappa\ell'/2)$, so there is no contradiction. In fact, it is worth noting that the second part of Theorem 3 admits the following (maybe less mysterious) corollary:

Corollary 1. *If $q \leq 2^n$ then, for every $C \geq 1$ such that $Cq \leq 2^{\kappa+n-2}$,*

$$\mathbf{Adv}_{\ell}^{\mathsf{casc}}(q) \leq \frac{\ell^2}{2^{\kappa+1}} + \frac{4}{2^n} + \frac{\beta}{C} + q\ell C^{(\ell+1)'/2} \left(\frac{8}{2^{\kappa}}\right)^{\ell'/2}$$

where $\beta = \ell^2 2^\ell (6 \log q + 2)^{\ell'/2}$. Moreover this also holds if the adversary is allowed to ask, for free, all possible 2^n queries to its second oracle.

In this corollary we have removed the constraint $Cq < \exp(\kappa\ell'/2)$, but the last term of the security bound is worse by a factor $2^n/q$. Corollary 1 follows easily from Theorem 3: if $Cq \geq 2^{\kappa\ell'/2}$ then Corollary 1 is vacuously true because the last term of the security bound is greater than 1, whereas if $Cq < 2^{\kappa\ell'/2}$ then we can fall back on Theorem 3, since the bound in Corollary 1 is strictly weaker than the bound in the second part of Theorem 3.

As we prove Theorem 1 from Theorem 3 we will, in fact, rely on Corollary 1 for the case $q \leq 2^n$.

Theorem 1 is obtained, essentially, by finding the optimum C for $q \geq 2^n$ and $q \leq 2^n$ (in the latter case, referring to Corollary 1) and by substituting this value into the bounds of Theorem 3. Details follow.

In the first part of Theorem 3 the terms containing C are

$$\frac{\alpha}{C} + \ell C 2^n \left(\frac{8Cq}{2^{\kappa+n}}\right)^{\ell'/2}.$$

This expression has the form

$$\alpha C^{-1} + C^{\gamma} B \tag{25}$$

for

$$B = \ell 2^n \left(\frac{8q}{2^{\kappa+n}}\right)^{\ell'/2}$$

and

$$\gamma = (\ell+1)'/2.$$

Differentiating (25) with respect to C, setting to 0, and solving, we find

$$\begin{aligned} & -\alpha C^{-2} + \gamma C^{\gamma-1} B = 0 \\ \iff\ & \gamma C^{\gamma-1} B = \alpha C^{-2} \\ \iff\ & C^{\gamma+1} = \alpha/\gamma B \\ \iff\ & C = (\alpha/\gamma B)^{1/(\gamma+1)}. \end{aligned}$$

Substituting this expression for C back into (25), we find

$$\begin{aligned} & \alpha(\alpha/\gamma B)^{-1/(\gamma+1)} + (\alpha/\gamma B)^{\gamma/(\gamma+1)} B \\ & = \alpha^{\gamma/(\gamma+1)} B^{1/(\gamma+1)} (\gamma^{1/(\gamma+1)} + \gamma^{-\gamma/(\gamma+1)}) \\ & \leq \alpha^{\gamma/(\gamma+1)} B^{1/(\gamma+1)} (\gamma^{1/(\gamma+1)} + 1) \\ & \leq \alpha(\gamma+1) B^{1/(\gamma+1)} \end{aligned}$$

using $\gamma \geq 1$, $\alpha^{\gamma/(\gamma+1)} \leq \alpha$, $\gamma^{1/(\gamma+1)} \leq \gamma$.

One can observe that

$$2^n \left(\frac{2^{\kappa+n(\ell'-2)/\ell'}}{2^{\kappa+n}}\right)^{\ell'/2} = 2^n \left(2^{n[(\ell'-2)/\ell' - \ell'/\ell']}\right)^{\ell'/2} = 1$$

and

$$1/(\gamma+1) = 2/(\ell+3)'$$

since $\gamma + 1 = (\ell+3)'/2$, so that if we let

$$q = r2^{\kappa+n(\ell'-2)/\ell'}$$

then

$$\begin{aligned}
B^{1/(\gamma+1)} &= \left(\ell 2^n \left(\frac{8q}{2^{\kappa+n}}\right)^{\ell'/2}\right)^{2/(\ell+3)'} \\
&= \left(\ell 2^n \left(\frac{8r2^{\kappa+n(\ell'-2)/\ell'}}{2^{\kappa+n}}\right)^{\ell'/2}\right)^{2/(\ell+3)'} \\
&= \left((8r)^{\ell'/2}\ell 2^n \left(\frac{2^{\kappa+n(\ell'-2)/\ell'}}{2^{\kappa+n}}\right)^{\ell'/2}\right)^{2/(\ell+3)'} \\
&= (8r)^{\ell'/(\ell+3)'}\ell^{2/(\ell+3)'} \\
&= \left(\frac{8q}{2^{\kappa+n(\ell'-2)/\ell'}}\right)^{\ell'/(\ell+3)'}\ell^{2/(\ell+3)'} \\
&\leq \ell^{1/2}\left(\frac{8q}{2^{\kappa+n(\ell'-2)/\ell'}}\right)^{\ell'/(\ell+3)'}
\end{aligned}$$

and, altogether,

$$\alpha(\gamma+1)B^{1/(\gamma+1)} = \alpha((\ell+3)'/2)\ell^{1/2}\left(\frac{8q}{2^{\kappa+n(\ell'-2)/\ell'}}\right)^{\ell'/(\ell+3)'}$$

wherefrom the first part of Theorem 1, given that $(\ell+3)'/2 \leq (\ell+4)/2 = \ell/2+2$.

For the second part of Theorem 1 we apply a similar minimization to the bound found in Corollary 1; this time the expression containing C is of the form

$$\beta C^{-1} + C^{\gamma}B \tag{26}$$

with

$$B = q\ell\left(\frac{8}{2^{\kappa}}\right)^{\ell'/2}$$

and

$$\gamma = (\ell+1)'/2$$

again. The value C which minimizes (26) is

$$\begin{aligned}
C &= (\beta/\gamma B)^{1/(\gamma+1)} \\
&= \left(\frac{2\beta}{(\ell+1)'q\ell}\left(\frac{2^{\kappa}}{8}\right)^{\ell'/2}\right)^{2/(\ell+3)'} \\
&= \left(\frac{2\beta}{(\ell+1)'q\ell}\right)^{2/(\ell+3)'}\left(\frac{2^{\kappa}}{8}\right)^{\ell'/(\ell+3)'}
\end{aligned}$$

so that

$$\begin{aligned} Cq &= q\left(\frac{2\beta}{(\ell+1)'q\ell}\right)^{2/(\ell+3)'}\left(\frac{2^\kappa}{8}\right)^{\ell'/(\ell+3)'} \\ &\le q\left(\frac{2\beta}{(\ell+1)'q\ell}\right)\left(\frac{2^\kappa}{8}\right) \\ &= \frac{2\ell^2 2^\ell(6\log q+2)^{\ell'/2}}{(\ell+1)'\ell}\left(\frac{2^\kappa}{8}\right) \\ &\le 2^{\ell-2}(6n+2)^{\ell'/2}2^\kappa \end{aligned}$$

which is at most $2^{\kappa+n-2}$ if $2^\ell(6n+2)^{\ell'/2} \le 2^n$, as is assumed in the second part of Theorem 1. We can therefore use this C in Corollary 1 and, as found above, the optimal value of (26) using this C is

$$\beta C^{-1} + C^\gamma B \le \beta(\gamma+1)B^{1/(\gamma+1)} = \beta((\ell+3)'/2)\left(\frac{\ell 8^{\ell'/2}q}{2^{\kappa\ell'/2}}\right)^{2/(\ell+3)'}.$$

Moreover, $8^{\ell'/2} \le 3^{\ell'}$ and $(\ell+3)'/2 \le \ell/2+2$. This completes the proof of Theorem 1 from Theorem 3 and Corollary 1.

B Proof of Lemma 1

To prove Lemma 1 we can focus on $\mathsf{fwd}(\tau)$. Adopting the "super-query" technique of [3] we imagine the following modified game: when the adversary has already made $N/2$ queries to $E(k,\cdot)$ under the same key k, we give the remaining $N/2$ queries to E under that key to A for free. In this case we say that a "super query" occurs, and the $N/2$ free queries are said to be *part of* that super query. We can keep the assumption that A never makes redundant queries in this modified game. We note that Q_E can reach length $2q$ in this modified game, since A can obtain half its queries (but no more) for free. We emphasize that this modified game is a "mental experiment" that we consider *only* for the purpose of proving Lemma 1, and which otherwise bears no connection to the proof of Theorem 3. In particular, the set of queries Q_E which we obtain as part of the modified experiment is *not* the same as the original set of queries Q_E in the real transcript τ. In fact we will eschew mention of τ, to avoid confusion, and focus instead on

$$\mathsf{fwd}(Q_E^+) := \max_{y_0\in\{0,1\}^n} |\{(k,x,y) \in Q_E^+ : y = y_0\}| \tag{27}$$

where Q_E^+ is (for the purpose of this section) the set of forward queries made by A *plus* the set of queries obtained as part of super queries, i.e., plus the set of freely obtained queries. (Though making backward queries has no obvious benefit, the adversary is still allowed to make backward queries in the modified game.)

Write $Q_E^{+\mathsf{N}}$, $Q_E^{+\mathsf{S}}$ for the queries in Q_E^+ that A obtains via normal queries and via super queries respectively. Then Q_E^+ is the disjoint union of $Q_E^{+\mathsf{N}}$ and $Q_E^{+\mathsf{S}}$, and

$$\mathsf{fwd}(Q_E) \le \mathsf{fwd}(Q_E^{+\mathsf{N}}) + \mathsf{fwd}(Q_E^{+\mathsf{S}}) \tag{28}$$

with $\mathsf{fwd}(Q_E^{+\mathsf{N}})$, $\mathsf{fwd}(Q_E^{+\mathsf{S}})$ defined by analogy with (27). Moreover

$$\mathsf{fwd}(Q_E^{+\mathsf{S}}) \le \frac{q}{N/2} = \frac{2q}{N} \le \begin{cases} \frac{qn}{N} & \text{if } q \ge N, \\ 2 & \text{if } q \le N \end{cases} \tag{29}$$

because each super-query consists of $N/2$ queries with distinct y coordinates (by the fact that $E_k(\cdot)$ is a permutation) so that each super-query can only contribute 1 to $\mathsf{fwd}(Q_E^{+\mathsf{S}})$; moreover, there are at most $q/(N/2)$ super-queries; and $n \geq 2$ follows from the observations initially made at the beginning of Section 6.

It thus remains to upper bound $\mathsf{fwd}(Q_E^{+\mathsf{N}})$. The intuition, here, is that each forward query made by the adversary (thus, not obtained for free as part of a super query) is answered at random from a set of size at least $N/2$. Thus we expect that $\mathsf{fwd}(Q_E^{+\mathsf{N}})$ will be governed by the balls-in-bins statistics of q balls thrown independently at random into $N/2$ bins, which are well understood. Of course, a subtlety occurs because the adversary has (some) control over *which* $N/2$ bins the next ball will appear in. While this extra power quite intuitively gives no advantage in increasing the size of the maximally occupied bin, providing a formal proof of this intuitive fact can reveal itself an annoying and even nontrivial task. Hence, we give such a proof for completeness (and also, partly, for the sake of amusement).

Formally, let $\mathsf{BinGame}(q, p, \alpha)$ be the following game: an adversary A has an infinite set of bins, initially empty, that we identify with the natural numbers $\mathbb{N}$; A has q rounds to play; at the i-th round A selects a set $\mathcal{S} \subseteq \mathbb{N}$ such that $|\mathcal{S}| \geq p$, and a "ball" is thrown uniformly at random into one of the bins in $\mathcal{S}$; the adversary wins if, at the end of the game after the q-th round, there is at least one bin with α or more balls. For example, if $\alpha > q$ then $\mathsf{BinGame}(q, p, \alpha)$ is impossible to win.

Let B denote the "dumbest" adversary for this game: at each round, B selects $\mathcal{S} = \{1, \dots, p\}$ regardless of prior history. Then we can prove:

Lemma 4. *Let A be an arbitrary adversary for $\mathsf{BinGame}(q, p, \alpha)$ and let B be the nonadaptive adversary just sketched. Then B has chance at least as great as A of winning $\mathsf{BinGame}(q, p, \alpha)$.*

Proof. We will couple the executions of A and B such that B wins whenever A wins. As the game proceeds we need to argue that B is always doing at least as well as A in a sense explained below.

Let a_i^ℓ be the number of balls in A's i-th bin right *before* the ℓ-th ball is thrown, and define b_i^ℓ likewise for B. We can assume wlog that bins are rearranged after each ball is thrown so that $a_1^\ell \geq a_2^\ell \geq a_3^\ell \geq \dots$ and $b_1^\ell \geq b_2^\ell \geq b_3^\ell \geq \dots$. For B this causes no change, and A can obviously adopt its strategy to accomodate for such bin rearrangement. We note that

$$\sum_{i=1}^{\infty} a_i^\ell = \sum_{i=1}^{\infty} b_i^\ell = \ell - 1.$$

Thus $(a_i^\ell)_{i=1}^\infty$ and $(b_i^\ell)_{i=1}^\infty$ are *partitions* of $\ell - 1$, in the parlance of algebraic combinatorics [31]. Put $a^\ell = (a_i^\ell)_{i=1}^\infty$ and $b^\ell = (b_i^\ell)_{i=1}^\infty$ for $1 \leq \ell \leq q+1$, with a^{q+1}, b^{q+1} being the final ball distributions.

Let $\mathcal{S}_A^\ell$ be the set of bins selected by A for the ℓ-th round of the game, $1 \leq \ell \leq q$. We note that $\mathcal{S}_A^\ell$ is a random variable. We also write $\mathcal{S}_B^\ell$ for $\{1, \dots, p\}$, which is a constant set.

When the ℓ-th ball is thrown, we couple the randomness for A and B as follows. Let x be a number uniformly at random in $[0, 1)$. For A, we place the ℓ-th ball in the i-th bin of $\mathcal{S}_A^\ell$ if

$$\frac{i-1}{|\mathcal{S}_A^\ell|} \leq x < \frac{i}{|\mathcal{S}_A^\ell|}$$

and we place the ℓ-th ball in the j-th bin of $\mathcal{S}_B^\ell = \{1, \dots, p\}$ if

$$\frac{j-1}{p} \leq x < \frac{j}{p}.$$

Obviously, then, A and B are both playing fair versions of $\mathsf{BinGame}(q, p, \alpha)$, albeit with correlated randomness.

Note that

$$\Pr[a_1^{q+1} \geq \alpha] \qquad \text{and} \qquad \Pr[b_1^{q+1} \geq \alpha]$$

are respectively the probability that A and B win $\mathsf{BinGame}(q, p, \alpha)$. It will thus suffice if we can show that $b_1^{q+1} \geq a_1^{q+1}$ at the end of every execution. In fact we will show something stronger, namely that

$$\sum_{i=1}^{m} b_i^{\ell} \geq \sum_{i=1}^{m} a_i^{\ell}$$

for all $m \in \mathbb{N}$ and all $1 \leq \ell \leq q+1$. In other words, the partition b^ℓ *dominates* the partition a^ℓ, traditionally written $b^\ell \succeq a^\ell$, for all ℓ.

We show that $b^\ell \succeq a^\ell$ by induction on ℓ. For this, let $\lambda = (\lambda_i)_{i=1}^{\infty}$, $(\mu)_{i=1}^{\infty}$ be two partitions[13] such that $\lambda \succeq \mu$; let $i_\lambda \leq i_\mu$ be positive indices; and let λ', μ' be obtained by increasing each of λ_{i_λ}, μ_{i_μ} by 1 and by rearranging coordinates of λ, μ to keep the sequences in nonincreasing order. It's enough to show that $\lambda' \succeq \mu'$, since, obviously, the index of the bin which receives a ball in A's game is always at least as large as the index of the bin which receives a ball in B's game.

To prove $\lambda' \succeq \mu'$, let $i_\mu^* \leq i_\mu$ be the smallest index $i \geq 1$ such that $\mu_i = \mu_{i_\mu}$ and likewise let $i_\lambda^* \leq i_\lambda$ be the smallest index $i \geq 1$ such that $\lambda_i = \lambda_{i_\lambda}$. Then note that λ', μ' can be directly obtained from λ, μ by adding 1 to $\lambda_{i_\lambda^*}$, $\mu_{i_\mu^*}$, respectively, without further rearrangement of coordinates. If $i_\lambda^* \leq i_\mu^*$ then $\lambda' \succeq \mu'$ obviously follows from $\lambda \succeq \mu$, so we can assume $i_\mu^* < i_\lambda^*$.

Now to show $\lambda' \succeq \mu'$ it is sufficient and necessary to show that

$$\sum_{j \leq i} \lambda_j \geq 1 + \sum_{j \leq i} \mu_j$$

for all i such that $i_\mu^* \leq i \leq i_\lambda^* - 1$. Arguing by contradiction, say that

$$\sum_{j \leq i_0} \lambda_j \leq \sum_{j \leq i_0} \mu_j.$$

for some i_0 such that $i_\mu^* \leq i_0 \leq i_\lambda^* - 1$. Then: (i) $\lambda_{i_0} \leq \mu_{i_0}$ because $\lambda \succeq \mu$, (ii) μ_j cannot decrease for $i_0 \leq j \leq i_\mu$ by definition of i_μ^*, and (iii) λ_j cannot decrease for $i_0 \leq j \leq i_\mu$ by $\lambda \succeq \mu$ and by (ii). But (iii) and $i_\lambda \leq i_\mu$ contradicts $i_0 \leq i_\lambda^* - 1$. We conclude that $\lambda' \succeq \mu'$, as desired. □

Now we can apply the following proposition, which constitutes a classical maximum occupancy result:

Proposition 2. *[30] Let q balls be thrown independently into p bins, $p \geq 16$. If $q \geq p$, there is probability less than $1/p$ that the maximum number of balls per bin exceeds $3\log(p)(q/p)$. If $q \leq p$, there is probability less than $1/p$ that the maximum number of balls per bin exceeds $3\log(q)$.*

Since we wish to apply Proposition 2 with $p = N/2$ (by Lemma 4 and the fact that the answer to each of the adversary's forward queries comes at random from a set of size at least $N/2$) but our main theorem divides cases according to whether $q \geq N$, $q \leq N$ instead of according to whether $q \geq N/2$, $q \leq N/2$. Thus we are more interested in the following corollary to Proposition 2:

[13] Technically, $\lambda = (\lambda_i)_{i=1}^{\infty}$ is a partition if $\lambda_1 \geq \lambda_2 \geq \ldots$, if the λ_i's are nonnegative integers, and if $\sum_i \lambda_i < \infty$.

Corollary 2. *Let q balls be thrown independently into p bins, $p \geq 16$. If $q \geq 2p$, there is probability less than $1/p$ that the maximum number of balls per bin exceeds $3\log(p)(q/p)$. If $q \leq 2p$, there is probability less than $1/p$ that the maximum number of balls per bin exceeds $6\log(q)$.*

Corollary 2 is immediate from Proposition 2.

It follows from Corollary 2 applied with $p = N/2$ and from Lemma 4 that

$$\Pr[\mathsf{fwd}(Q_E^{+\mathsf{N}}) \geq 6nq/N] \leq \frac{2}{N}$$

for $q \geq N$, and that

$$\Pr[\mathsf{fwd}(Q_E^{+\mathsf{N}}) \geq 6\log q] \leq \frac{2}{N}$$

for $q \leq N$. Lemma 1 then follows from (28) and from (29).

C Double (and single) encryption

In this appendix we give a simple proof of the main result of [1] (slightly corrected—see below) concerning double encryption. In fact Bellare and Rogaway [4] already hint at the existence of a simple game-based proof of [1]'s main result, as the original proof of [1] is quite involved. Instead of game-playing we use the H-coefficient technique, which is more practical for us partly because we have already set the stage for the technique in the main proof. Moreover this should help convince that the H-coefficient technique can also be quite effective in settings that are combinatorially simple, and is not only a tool of "last resort" for hairy, quasi-intractable situations.

For the sake of completeness we follow up with a formal analysis (again H-coefficient-based) of single encryption. (We emphasize, however, that the results of previous sections are valid for $\ell = 1, 2$ as well; the point here is just to give tighter bounds with cleaner proofs.)

Double encryption. The result we prove on double encryption is the following:

Theorem 4. *One has*

$$\mathbf{Adv}^{\mathsf{casc}}_{2,\kappa,n}(q) \leq \frac{q^2}{2^{2\kappa}} + \frac{1}{2^\kappa}$$

for all q, κ, n. This bound also holds if the adversary is allowed 2^n free queries to its second oracle.

Aiello et al. actually claim a stronger bound of $q^2/2^{2\kappa}$, but this bound is incorrect as can be seen by considering the case $q = 0$. In more detail, one might have $k_1^* = k_2^*$, while the composition of a random permutation with itself is no longer random. Hence the adversary already has a nonzero distinguishing advantage by making only queries to $E^{(2)}_{k^*}/\pi$, and not making any queries to E/E^{-1}, and which contradicts a security bound of $q^2/2^{2\kappa}$. (If further convincing is needed one should consider the case $n = 1$.) The error can be found in the proof of Lemma 3.6 in [1], when it is claimed that "the composition of two permutations with one random, is random" (overlooking that the composition of two random *but equal* permutations is not random). We emphasize, however, that the proof of [1] is easily repairable by adding $k_1^* = k_2^*$ to their list of "bad events".

In the next few bullets we give a proof of Theorem 4, assuming familiarity with the material and notations of Section 5.

PRELIMINARY REDUCTIONS. We again follow Bellare and Rogaway [4] and replace $E^{(2)}/\pi$ by a fixed permutations S. More precisely, and like in the general case, we start by sending the adversary a symbol $\star \in \{\bot, \top\}$ where $\star = \top$ unless we are in the real world and $k_1^* = k_2^*$ where $k_1^* \| k_2^*$ is the secret key, in which case we send $\star = \bot$ and the real world aborts. If the real world doesn't abort we overwrite $E_{k_2^*} = S \circ E_{k_1^*}^{-1}$ in the real world. As before the adversary no longer requires access to its second oracle, since it knows S. We refer to the description of the same step in the general case for more details.

In terms of the underlying probability space, we recall that the space of all real world (or, for that matter, ideal world) oracles is the set of all pairs

$$(E', k^*) \in \mathcal{P}^{\exp(\kappa)} \times \{0,1\}^{2\kappa}.$$

In the real world, the adversary's oracle E is built from E' by setting

$$E_k = \begin{cases} E'_k & \text{if } k \neq k_2^* \\ S \circ E'^{-1}_{k_1^*} & \text{if } k = k_2^* \end{cases}$$

(presuming $\star \neq \bot$) whereas $E = E'$ in the ideal world.

We also assume that the adversary is deterministic and makes no redundant queries.

Transcripts. Here we adopt all the same conventions as in our main proof. In particular the key k^* is included at the end of the transcript, where k^* is a dummy key sampled uniformly at random in the ideal world. As before, $\mathcal{T}$ denotes the set of all transcripts.

Bad transcripts. A transcript $\tau = (\star, Q_E, k^*)$ with $k^* = k_1^* \| k_2^*$ is *bad* if $k_1^* = k_2^*$ or if Q_E contains *both* a query of the form (k_1^*, x, y) *and* a query of the form (k_2^*, x, y). As usual $\mathcal{T}_2$ denotes the set of bad transcripts, and $\mathcal{T}_1 := \mathcal{T} \backslash \mathcal{T}_2$.

Probability of bad transcript. We recall the probability of bad transcript is computed in the ideal world, and in the ideal world we can think of k^* as being sampled at random *after* all the queries in Q_E have been made. Since Q_E contains q queries, k_i^* has probability at most $q/2^\kappa$ of being a key in the transcript, and this probability is independent for $i = 1, 2$; thus

$$\Pr[Y \in \mathcal{T}_2] \leq \frac{q^2}{2^{2\kappa}} + \frac{1}{2^\kappa}$$

where the second term accounts for the probability that $k_1^* = k_2^*$ and where where Y is the probability distribution over transcripts in the ideal world.

Probability ratio for good transcripts. Let $\tau = (\star, Q_E, k^*) \in \mathcal{T}_1$ be a good transcript. Let q_k be the number of queries in Q_E appearing in Q_E with key k, so that either $q_{k_1^*} = 0$ or $q_{k_2^*} = 0$ by definition of good transcripts. If $q_{k_2^*} = 0$ then it is easy to see that

$$\mathsf{comp}_X(\tau) = \prod_{k \in \{0,1\}^\kappa} (2^n - q_k)! \tag{30}$$

$$\mathsf{comp}_Y(\tau) = \prod_{k \in \{0,1\}^\kappa} (2^n - q_k)! \tag{31}$$

with the set of all possible real-world/ideal-world oracles Ω_X, Ω_Y and with the sets of compatible oracles $\mathsf{comp}_X(\tau) \subseteq \Omega_X$, $\mathsf{comp}_Y(\tau) \subseteq \Omega_Y$ defined as in Section 5. (Note that in (30) the product term with $k = k_2^*$ accounts for the $2^n!$ possibilities for $E'_{k_2^*}$. Indeed, in the real world the transcript never constrains $E'_{k_2^*}$.) On the other hand if $q_{k_2^*} > 0$ then each query under key k_2^* in τ induces, in the real world, a unique constraint on $E_{k_1^*} = E'_{k_1^*}$; thus (30) also holds if $q_{k_2^*} > 0$ (though now the product term with $k = k_2^*$ is counting the number of possibilities for $E'_{k_1^*}$, whereas the term with $k = k_1^*$ is counting the number of possibilities for $E'_{k_2^*}$, which is still $2^n!$). Since (31) also obviously holds if $q_{k_2^*} > 0$, we have $|\mathsf{comp}_X(\tau)| = |\mathsf{comp}_Y(\tau)|$ whether $q_{k_1^*} > 0$ or whether $q_{k_2^*} > 0$. Thus, by equations (9), we have

$$\frac{\Pr[X = \tau]}{\Pr[Y = \tau]} = 1$$

for all $\tau \in \mathcal{T}_1$ such that $\Pr[Y = \tau] > 0$. It follows that the adversary's advantage is at most the probability of obtaining a bad transcript, which, as shown above, is upper bounded by $q^2/2^{2\kappa}+1/2^\kappa$. This completes the proof of Theorem 4.

Single encryption. For single encryption we can prove the following theorem:

Theorem 5. *One has*

$$\mathbf{Adv}^{\mathsf{casc}}_{1,\kappa,n}(q) \leq \frac{q}{2^\kappa}$$

for all q, κ, n. This bound also holds if the adversary is allowed 2^n free queries to its second oracle.

By now, the proof should be transparent. We apply the customary reduction of Bellare and Rogaway [4], except that this time we don't even need the initial message $\star$. Thus we replace the second oracle by a fixed permutation S, and set

$$E_k = \begin{cases} E'_k & \text{if } k \neq k^* \\ S & \text{if } k = k^* \end{cases}$$

in the real world and by setting $E = E'$ in the ideal world, and where $(E', k^*) \in \mathcal{P}^{\exp(\kappa)} \times \{0,1\}^\kappa$ is the underlying oracle (a.k.a. random tape) in either Ω_X or Ω_Y. Transcripts are defined as before, with the key $k^* \in \{0,1\}^\kappa$ included. A transcript $\tau = (Q_E, k^*)$ is bad if a query with key k^* appears in Q_E. The probability of a bad transcript in the ideal world is at most $q/2^\kappa$ since k^* is independent from Q_E in the ideal world. Moreover, $\Pr[X = \tau] = \Pr[Y = \tau]$ for a good transcript τ, as is easy to see that (30), (31) hold as well here. Theorem 5 follows.

D Some leftover formalities

In this appendix we make a few more brief comments on the identities

$$\Pr[X = \tau] = \frac{|\mathsf{comp}_X(\tau)|}{|\Omega_X|}, \qquad \Pr[Y = \tau] = \frac{|\mathsf{comp}_Y(\tau)|}{|\Omega_Y|} \tag{32}$$

mentioned in Section 5. Our comments have a significant intersection with similar comments made in [8] but are not a proper subset thereof (nor vice-versa).

The fact that $\Pr[Y = \tau] > 0$, mentioned as a sufficient condition for (32) to hold, can be replaced, in a general[14] context, by the requirement that τ be *attainable* [8], in the sense that there

[14] For the purpose of applying the H-coefficient technique, however, the difference between an attainable a transcript and a transcript for which $\Pr[Y = \tau] > 0$ is immaterial, since in any case we are only interested in applying (32) for transcripts τ such that $\Pr[Y = \tau] > 0$ (cf. footnote 5).

exists *some* oracle ω' (which could be neither real nor ideal, but something else entirely) such that A produces transcript τ on oracle ω', notated $A^{\omega'} \to \tau$. Without loss of generality ω' is deterministic (just as any $\omega \in \mathit{\Omega}_X \cup \mathit{\Omega}_Y$ is deterministic). Moreover, as indicated in footnote 5, an oracle ω is compatible with a transcript τ if and only if there exists some (wlog deterministic) adversary A' that produces transcript τ on oracle ω, notated $A'^{\omega} \to \tau$. Thus, once the the definitions of "attainability" (necessary for (32)) and "compatibility" are unfolded, both equalities in (32) become equivalent to the fact that

$$(A^{\omega} \to \tau) \iff ((\exists A' \text{ s.t. } A'^{\omega} \to \tau) \wedge (\exists \omega' \text{ s.t. } A^{\omega'} \to \tau)) \tag{33}$$

for every τ and A, where the existential quantifications are taken over deterministic adversaries A' and over deterministic oracles ω'. Here the forward implication is obvious. The reverse is equally straightforward: if B, C, B', C' are four deterministic interactive Turing machines such that both of the interactions $B \leftrightarrow C$ and $B' \leftrightarrow C'$ produce the same transcript τ, then, obviously, $B \leftrightarrow C'$ and $B' \leftrightarrow C$ also produce τ as transcript.