Short Variable Length Domain Extenders With Beyond Birthday Bound Security

Yu Long Chen[1], Bart Mennink[2], and Mridul Nandi[3]

[1] imec-COSIC, KU Leuven, Belgium
yulong.chen@kuleuven.be
[2] Digital Security Group, Radboud University, Nijmegen, The Netherlands
b.mennink@cs.ru.nl
[3] Indian Statistical Institute, Kolkata, India
mridul.nandi@gmail.com

Abstract. Length doublers are cryptographic functions that transform an n-bit cryptographic primitive into an efficient and secure cipher that length-preservingly encrypts strings of length in $[n, 2n-1]$. All currently known constructions are only proven secure up to the birthday bound, and for all but one construction this bound is known to be tight. We consider the remaining candidate, LDT by Chen et al. (ToSC 2017(3)), and prove that it achieves beyond the birthday bound security for the domain $[n, 3n/2)$. We generalize the construction to multiple rounds and demonstrate that by adding one more encryption layer to LDT, beyond the birthday bound security can be achieved for all strings of length in $[n, 2n-1]$: security up to around $2^{2n/3}$ for the encryption of strings close to n and security up to around 2^n for strings of length close to $2n$. The security analysis of both schemes is performed in a modular manner through the introduction and analysis of a new concept called "harmonic permutation primitives."

Keywords: length doublers, LDT, beyond birthday bound, harmonic primitives, chi-squared

1 Introduction

Block ciphers are keyed deterministic functions that encrypt bit strings of a fixed size n bits to ciphertext blocks of the same size. They play a predominant role in cryptography, and yet, most cryptographic applications deal with arbitrary-length messages. To achieve this, the applications evaluate a block cipher in a certain mode of operation.

A simple example of this is counter mode encryption. Given block cipher E_K on n bits, counter mode encrypts a message M of arbitrary length as follows. First, the message is partitioned into blocks $M_1, \ldots, M_\ell$, where the first $\ell - 1$ are of size n bits, and the last one may be smaller. Second, the message is encrypted as

$$C_i = E_k(\mathrm{ctr} + i) \oplus M_i \text{ for } i = 1, \ldots, \ell\,,$$

where the ℓ-th ciphertext block is truncated to have the same size as M_ℓ, and where ctr is a carefully specified counter.

Counter mode is unique in the sense that it allows for easy length-preservation due to its "streaming" property. Whereas this property is fine in some use cases, in many others it is lacking. For example, stream cipher encryption is inapplicable to disk sector encryption for security reasons. Alternative encryption modes like CBC [47], OCB [26,38,39], XTS [15], and TC3 [41], however, feed the message to the block cipher and there is no easy way of keeping length preservation. One often pads input to size a multiple of n-blocks and takes ciphertext expansion for granted [1,2,26,28]. Ciphertext expansion is, in many cases, not desirable: it creates overhead, making it unsuitable for disk encryption and low-bandwidth network protocols.

A generic method for length-preserving variable-length encryption is ciphertext stealing [13,40]. Informally, it encrypts the first $\ell - 1$ blocks as is, but to encrypt the non-integral ℓ-th block, it is first expanded to n bits by scraping sufficiently many ciphertext bits from the $(\ell - 1)$-th block and gluing these to M_ℓ. The approach is appealing, but it only works on modes of use for which ciphertext blocks can be decrypted independently of each other: otherwise one cannot recover the ciphertext bits scraped off of $C_{\ell-1}$.

Besides these two generic solutions, many dedicated designs that support variable-length encryption have appeared, e.g., EME [20], TET [21], HEH [43], HCTR [46], HCH [10], and XCB [27], but a golden method for generically transforming an existing block cipher mode of operation for integral data to one for arbitrary-length data was long due.

1.1 Length Doublers

In 2007, Ristenpart and Rogaway [37] introduced *length doublers* as an elegant way of achieving variable-length encryption. A length doubler is a length-preserving encryption mode on the set of bit strings of size between n and $2n-1$ bits, where n is the state size of the underlying primitive.

By allowing flexibility of the size of the second block, length doublers suit well as modular building blocks for variable-length encryption and authenticated encryption. For example, whereas the possibility to apply ciphertext stealing depends on the mode in consideration, length doubling can be used generically for black-box authenticated encryption schemes as demonstrated by Chen et al. [11]. We discuss further applications of length doublers in Section 1.4.

Alongside the formalization, Ristenpart and Rogaway introduced the XLS length doubler, based on three block cipher calls and two evaluations of a so-called *ϵ-good mixing function*. It found application in first-round CAESAR submission AES-COPA [2,3]. Only 7 years after its introduction, Nandi found an attack on XLS [32], an attack that also rendered the solution in the COPA mode insecure [34]. Nandi further proved that a secure length doubler must make at least four block cipher calls [33]. Other length doublers introduced after XLS are DE by Nandi [31] and HEM by Zhang [48], both of which make four block cipher calls and match the lower bound of [33].

Chen et al. considered the design of length doublers from tweakable block ciphers and introduced LDT [11]. It makes two calls to a tweakable block cipher and uses a *pure mixing function*, noting that an ϵ-good mixing function is pure but not necessarily vice versa. The transition to using tweakable block ciphers is a natural one: 18 initial submissions to the CAESAR competition were based on tweakable block ciphers, various novel cryptographic authentication and/or encryption modes use a tweakable block cipher as black box [23, 35, 41, 45], and dedicated tweakable block ciphers like TWEAKEY [24] and SKINNY [5] are gaining traction. The recently announced ARMv8.3 [36] uses an implementation of the lightweight tweakable blockcipher QARMA [4]. The approach allows for more modular (and thus simpler) security proofs.

1.2 Towards Beyond Birthday Bound Security

All of the length doublers mentioned so far, barring XLS, are proven secure up to $2^{n/2}$. For DE and HEM this bound is tight as there is an attack matching this complexity. For LDT, Chen et al. [11] derived an attack in approximately $2^{n-s/2}$ queries, as long as all queries are of size at least $n+s$. The bound suggests tightness for $s = n - 1$, but it leaves the possibility of proving beyond birthday bound security for $s \ll n - 1$ open.

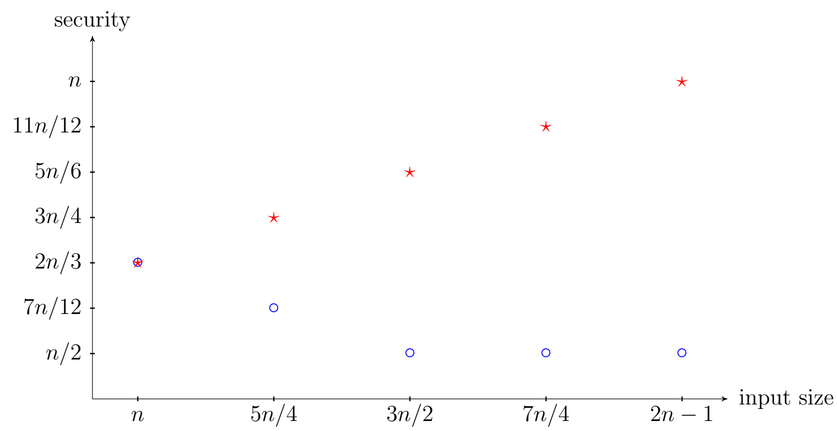
Although all length doublers known to date have only birthday bound proven security, beyond birthday bound secure length doublers are relevant for various scenarios. First, consider the case of a cryptographic mode that uses a length doubler in a black-box manner and achieves beyond birthday bound security. If it is instantiated with any off-the-shelf solution (DE, HEM, LDT) the provable security guarantee degrades to birthday bound security. Second, considering the case of format-preserving encryption and electronic product code tag encryption (see Section 1.4), using a birthday bound secure length doubler with a lightweight 64-bit block cipher yields 32-bit security at best. A beyond birthday bound secure length doubler would guarantee security up to well beyond 32 bits.

1.3 Our Contribution

We challenge the problem of proving beyond birthday bound security of length doublers. The starting point of our work is Chen et al.'s LDT: it is simple, modular, and so far the only existing candidate that may offer beyond birthday bound security.

As first contribution, we prove in Section 5 that the original LDT achieves beyond the birthday security for queries of size in $[n, 3n/2)$: if only evaluations of size around n are permitted, $2n/3$-bit security is achieved, but if evaluations of size around $3n/2$ are permitted, the proven security bound degrades to $n/2$. The bound is not tight, but we recall that Chen et al. [11] already demonstrated a birthday bound attack if $s = n - 1$, testifying of the fact that the security *decreases* with s. As second and main contribution, we generalize the mode to r-round LDT, recalling that the original construction consists of 2 rounds, and prove in Section 6 that 3-round LDT achieves beyond the birthday bound

Fig. 1: Security bound of 2-LDT and 3-LDT for various choices of input size, where ∘ stands for 2-LDT and ⋆ stands for 3-LDT.



security for the entire domain $[n, 2n-1]$. As proven so far, the security of 3-round LDT *increases* with s: for evaluations of size around n we achieve $2n/3$-bit security, and for evaluations of size around $2n$ we get optimal n-bit security. Figure 1 plots the simplified security results of 2-LDT and 3-LDT for bit strings of length in $\{n, 5n/4, 3n/2, 7n/4, 2n-1\}$ (these data are taken from the discussion in Sections 5 and 6). In aforementioned example of an 80-bit cipher using a 64-bit primitive, 3-LDT achieves $2^{3n/4} = 2^{48}$ security.

Central to our proofs is the introduction and usage of a new concept: "harmonic permutation primitives." These can be seen as lazily-sampled permutations where one part of the state is always sampled uniformly at random and the other part in such a way that permutation consistency is maintained. We describe two harmonic primitives: a harmonic tweakable permutation in Section 4.1 and a harmonic variable-length pseudorandom permutation in Section 4.2.

These harmonic permutation primitives allow for compact, neat, and modular security proofs of both 2-round and 3-round LDT. Both proofs use the two harmonic permutation primitives of Section 4 in a different setting, but using the chi-squared method by Dai et al. [14] and properties of the hypergeometric distribution, security of both LDT modes is reduced to the security of the harmonic permutation primitives. What then remains is an analysis of these primitives in Sections 7 and 8.

Inspired by the proof approach in this work, one may likewise use the two harmonic permutation primitives to prove security of r-round LDT for $r \geq 4$. However, it would only render marginal improvement of the bound, with a large efficiency penalty. It nevertheless appears that the idea of harmonic permutation primitives and our proof technique may be broadly applicable beyond LDT, for example in the direction of sponge functions [7].

1.4 Application

An example use case of length doublers is format-preserving encryption, a field that got significant attention recently in light of the standardization [16] of FF1 [6] and FF3 [8]. Format-preserving encryption considers the problem of encrypting data from a small domain that does not fit the parameters of standardized block ciphers. For example, there is no practical way to length-preservingly encrypt 80-bit strings using AES-128 (other than streaming-based). Whereas the standardized FF1 and FF3 are made to facilitate arbitrary types of domains, for certain cases this can equally well be resolved using a length doubler. Above example of 80-bit strings can be resolved with a birthday bound secure length doubler on top of a lightweight 64-bit tweakable block cipher, but that would only give 32-bit security. As shown in Figure 1, for this scenario 2-LDT would achieve around 37-bit security and 3-LDT even 48-bit security.

A more concrete example is that of electronic product code tag encryption, which is considered as a replacement for bar codes using low-cost passive RFID-tags. The standard EPC Class 1 Gen 2 RFID tag [18] proposes to use a unique 96-bit identifier for any physical item [19]. As for above generic case, a birthday bound secure length doubler on top of a 64-bit block cipher would give 32-bit security at best. Our bound of 2-LDT does not improve for this regime (see Figure 1), but 3-LDT does achieve beyond birthday bound security: instantiated with a 64-bit tweakable block cipher, it reaches around 53-bit security.

It is straightforward to transform r-LDT into a *tweakable* length doubler, where the tweak is fed as additional tweak input to the underlying tweakable block ciphers (this requires extending the tweak space of the underlying primitive). This observation has two implications. First, one can obtain multi-user security of r-LDT by considering user IDs as tweak inputs and feeding those to the underlying tweakable block cipher. Second, r-LDT is an interesting and non-obvious generalization of the tweakable block cipher based domain extender of Coron et al. [12]. Stated simply, Coron et al. considered the problem of transforming a tweakable block cipher with $2n$-bit tweaks and n-bit blocks into a domain extender with n-bit tweaks and $2n$-bit blocks. They presented a 2-round scheme (achieving birthday bound security) and a 3-round scheme (achieving optimal n-bit security). Our tweakable length doublers, instead, transform that tweakable block cipher into a length doubler with n-bit tweaks and $[n, 2n-1]$-bit blocks, therewith enabling support for variable length input. For the specific case of $s \approx n$, our schemes achieve the same level of security as those of [12].

Finally, we remark that if one considers 2-LDT for fixed s, and sandwiches it by two universal hash functions in a specific way, the resulting construction is identical to the Small-Block Cipher (SBC) construction proposed by Minematsu and Iwata [30] (an extension of ENR [29]). As SBC is designed to achieve beyond birthday bound security quite efficiently, it makes sense to compare it with 3-LDT. It turns out that 3-LDT compares favorably in various aspects. First, Minematsu and Iwata showed that SBC achieves $(n+s)/2$-bit security, whereas 3-LDT achieves $(2n+s)/3$-bit security for any fixed s (see also the last column of Table 1 in Section 6). Second, SBC uses two tweakable block ciphers and two universal hash functions, whereas 3-LDT uses three tweakable block ciphers. The latter could be beneficial for implementation on constrained devices. Finally, SBC is ultimately still a fixed input length cipher, whereas 3-LDT allows for inputs of size $[n, 2n-1]$.

2 Preliminaries

For $n \in \mathbb{N}$, we denote the set of all bit strings of length n as $\{0,1\}^n$, and the set of all bit strings of arbitrary length as $\{0,1\}^*$. For $m \in \mathbb{N}$ and $m \leq n$ we define $\{0,1\}^{[m,n]} = \bigcup_{m \leq i \leq n} \{0,1\}^i$. Given two bit strings $X, Y \in \{0,1\}^*$, we use both $X \| Y$ and XY interchangeably to denote their concatenation. The length of X is denoted $|X|$, and if X and Y satisfy $|X| = |Y|$, we denote their bitwise addition as $X \oplus Y$. For $X \in \{0,1\}^n$, we denote $\mathsf{left}_m(X)$ the m leftmost bits of X and $\mathsf{right}_m(X)$ the m rightmost bits of X, in such a way that $X = \mathsf{left}_{n-m}(X) \| \mathsf{right}_m(X)$.

For $n \in \mathbb{N}$ and $X \in \{0,1\}^{[0,n-1]}$, we define a padding function $\mathsf{pad}(X) = X \| 10^{n-|X|-1}$. We denote its inverse unpad that on input of a string of length n removes the rightmost string 10^* and returns the resulting string. Note that unpad is an injective mapping.

The expression $S \leftarrow T$ denotes the assignment of the value T to variable S, $\mathcal{L} \xleftarrow{\cup} S$ the addition of S to list $\mathcal{L}$, and $S \xleftarrow{\$} \mathcal{S}$ for finite set $\mathcal{S}$ the uniformly random sampling of S from $\mathcal{S}$. For an algorithm $\mathcal{D}$ and a function/oracle $\mathcal{O}$, $\mathcal{D}^{\mathcal{O}}$ represents the evaluation of $\mathcal{D}$ with oracle interaction to $\mathcal{O}$, and $\Delta_{\mathcal{D}}\left(\mathcal{O}\,;\mathcal{P}\right)$ represents the advantage of $\mathcal{D}$ in distinguishing $\mathcal{O}$ from an oracle $\mathcal{P}$.

2.1 (Tweakable) Block Ciphers

For arbitrary finite key space $\mathcal{K}$ and $n \in \mathbb{N}$, a block cipher is a function $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ such that for every fixed key $K \in \mathcal{K}$, $E_K(\cdot) = E(K, \cdot)$ is a permutation on $\{0,1\}^n$. We denote its inverse for fixed key K by $E_K^{-1}(\cdot) = E^{-1}(K, \cdot)$. Denote by $\mathrm{Perm}(n)$ the set of all permutations on $\{0,1\}^n$. Tweakable block ciphers generalize over ordinary block ciphers by input of a t-bit tweak, for $t \in \mathbb{N}$. More detailed, a tweakable block cipher is a function $\tilde{E} : \mathcal{K} \times \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$ such that for every fixed key $K \in \mathcal{K}$ and tweak $T \in \{0,1\}^t$, $\tilde{E}_K(T, \cdot) = \tilde{E}(K, T, \cdot)$ is a permutation on $\{0,1\}^n$. Its inverse for fixed key K and tweak T is denoted by $\tilde{E}_K^{-1}(T, \cdot) = \tilde{E}^{-1}(K, T, \cdot)$. Denote by $\widetilde{\mathrm{Perm}}(t, n)$ the set of all families of permutations $\tilde{\pi} : \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$ indexed by tweak $T \in \{0,1\}^t$.

The security of a tweakable block cipher $\tilde{E}$ is measured by considering a distinguisher $\mathcal{D}$ that has two-sided query access to either $\tilde{E}_K$ for a randomly drawn key $K \xleftarrow{\$} \mathcal{K}$, or a random tweakable permutation $\tilde{\pi} \xleftarrow{\$} \widetilde{\mathrm{Perm}}(t, n)$, and its goal is try to distinguish the real construction from the ideal one:

$$\mathbf{Adv}_{\tilde{E}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}) = \left| \Pr\left[K \xleftarrow{\$} \mathcal{K} : \mathcal{D}^{\tilde{E}_K^{\pm}} = 1\right] - \Pr\left[\tilde{\pi} \xleftarrow{\$} \widetilde{\mathrm{Perm}}(t, n) : \mathcal{D}^{\tilde{\pi}^{\pm}} = 1\right] \right| .$$

For the left probability, the key space $\mathcal{K}$ is a fair representation of the randomness of $\tilde{E}^{\pm}_K$. Often, it is the set of k-bit strings, where k is the key size. In proofs, specifically in hybrid arguments within proofs, one regularly considers tweakable block ciphers with idealized primitives. For example, one may consider the construction of a tweakable block cipher $\tilde{E}$ from a secret permutation (that could, in turn, be instantiated using a block cipher with secret key). In this case, the key space of $\tilde{E}$ is $\mathcal{K} = \mathrm{Perm}(n)$. More involved examples appear if the construction internally consists of lazy sampling, as will for instance be the case with our harmonic tweakable SPRP in Section 4.1.

2.2 Chi-Squared Method

Our proof will rely on the chi-squared method by Dai et al. [14].

Consider two stateless systems $\mathcal{O}_0, \mathcal{O}_1$ and any computationally unbounded deterministic distinguisher $\mathcal{D}$ that has query access to either of these systems. The distinguisher's goal is to distinguish both systems. If we denote the maximum amount of queries by q, we can define a transcript $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(q)})$ and let $\boldsymbol{\tau}^{(i)} = (\tau^{(1)}, \ldots, \tau^{(i)})$ for every $i \le q$. Distinguisher $\mathcal{D}$ can make its queries adaptively, but as it makes them in a deterministic manner, the $(i+1)$-th query input is determined by the first i query-responses $\boldsymbol{\tau}^{(i)}$.

For system $\mathcal{O} \in \{\mathcal{O}_0, \mathcal{O}_1\}$ and fixed tuple $\boldsymbol{\tau}^{(i)}$, we denote by $p_{\mathcal{O},\mathcal{D}}(\boldsymbol{\tau}^{(i)})$ the probability that distinguisher $\mathcal{D}$ interacting with $\mathcal{O}$ obtains transcript $\boldsymbol{\tau}^{(i)}$ for its first i queries. If $p_{\mathcal{O},\mathcal{D}}(\boldsymbol{\tau}^{(i)}) > 0$, then we denote by $p_{\mathcal{O},\mathcal{D}}(Y^{(i+1)} \mid \boldsymbol{\tau}^{(i)})$ the conditional probability that $\mathcal{D}$ receives response $Y^{(i+1)}$ upon its $(i+1)$-th query, given transcript $\boldsymbol{\tau}^{(i)}$ of the first i queries (that deterministically fixes the input to the $(i+1)$-th query). Define for any $i \in \{1, \ldots, q\}$ and any query-response tuple $\boldsymbol{\tau}^{(i)}$:

$$\chi^2(\boldsymbol{\tau}^{(i-1)}) = \sum_{Y^{(i)}} \frac{\left(p_{\mathcal{O}_1,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)}) - p_{\mathcal{O}_0,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)})\right)^2}{p_{\mathcal{O}_0,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)})}\,, \tag{1}$$

where the sum is taken over all $Y^{(i)}$ in the support of the distribution $p_{\mathcal{O}_0,\mathcal{D}}(\cdot \mid \boldsymbol{\tau}^{(i-1)})$. The chi-squared method states the following:

Lemma 1 (Chi-squared method [14]). *Consider a fixed deterministic distinguisher $\mathcal{D}$ and two systems $\mathcal{O}_0, \mathcal{O}_1$. Suppose that for any $i \in \{1, \ldots, q\}$ and any query-response tuple $\boldsymbol{\tau}^{(i)}$, $p_{\mathcal{O}_0,\mathcal{D}}(\boldsymbol{\tau}^{(i)}) > 0$ whenever $p_{\mathcal{O}_1,\mathcal{D}}(\boldsymbol{\tau}^{(i)}) > 0$. Then:*

$$\Delta_{\mathcal{D}}\Big(\mathcal{O}_0\,;\mathcal{O}_1\Big) = \|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| \le \left(\frac{1}{2}\sum_{i=1}^{q} \mathrm{Exp}[\chi^2(\boldsymbol{\tau}^{(i-1)})]\right)^{1/2}\,, \tag{2}$$

where the expectation is taken over $\boldsymbol{\tau}^{(i-1)}$ of the $i-1$ first answers sampled according to interaction with $\mathcal{O}_1$.

2.3 Hypergeometric Distribution

The hypergeometric distribution $\mathrm{HG}(N, K, n)$ considers the case of n draws without replacement from a set of size N elements, denote by K the total number of successes out of N and h the number of successes present in a sample of size n. It is well-known that for $h \sim \mathrm{HG}(N, K, n)$,

$$\begin{aligned}\mathrm{Exp}[h] &= n \cdot \frac{K}{N}\,, \\ \mathrm{Var}[h] &= n \cdot \frac{K}{N} \cdot \frac{(N-K)}{N} \cdot \frac{N-n}{N-1}\,.\end{aligned}$$

3 Length Doublers and LDT

Following Chen et al. [11], we recall the formalization of length doublers in Section 3.1, and present generalized LDT in Section 3.2.

3.1 Length Doublers

For arbitrary finite key space $\mathcal{K}$ and $n \in \mathbb{N}$, a length doubler is a function $\mathcal{E} : \mathcal{K} \times \{0,1\}^{[n,2n-1]} \to \{0,1\}^{[n,2n-1]}$ such that for every fixed key $K \in \mathcal{K}$, $\mathcal{E}_K(\cdot) = \mathcal{E}(K, \cdot)$ is a length preserving invertible function on $\{0,1\}^{[n,2n-1]}$. We denote its inverse for fixed key K by $\mathcal{E}_K^{-1}(\cdot) = \mathcal{E}^{-1}(K, \cdot)$. Note that $\mathcal{E}$ should behave like a random permutation for every length input in $[n, 2n-1]$. Denote by $\mathrm{VPerm}([n, 2n-1])$ the set of all length-preserving and invertible functions on $\{0,1\}^{[n,2n-1]}$. The security of $\mathcal{E}$ is measured by considering a distinguisher $\mathcal{D}$ that has two-sided query access to either $\mathcal{E}_K$ for a randomly drawn key $K \xleftarrow{\$} \mathcal{K}$, or a random length-preserving permutation $\rho \xleftarrow{\$} \mathrm{VPerm}([n, 2n-1])$, and its goal is to try to distinguish the real construction from the ideal one:

$$\begin{aligned}\mathbf{Adv}^{\mathrm{vsprp}}_{\mathcal{E}}(\mathcal{D}) = \Big| &\Pr\left[K \xleftarrow{\$} \mathcal{K} : \mathcal{D}^{\mathcal{E}_K^{\pm}} = 1\right] - \\ &\qquad \Pr\left[\rho \xleftarrow{\$} \mathrm{VPerm}([n, 2n-1]) : \mathcal{D}^{\rho^{\pm}} = 1\right]\Big|\,.\end{aligned}$$

As in Section 2.1, the key space $\mathcal{K}$ corresponds to the source of randomness of the construction $\mathcal{E}_K^{\pm}$. It may take various shapes, but it will always be clear from the context.

3.2 Generalized LDT

Chen et al. [11] introduced length doubler LDT that internally makes two calls to an underlying tweakable block cipher, separated by an evaluation of a "pure mixing function" (a weaker variant of a multipermutation [44]) on part of the state. In this work, we will consider a generalization of LDT to multiple rounds, but we simplify it by discarding the pure mixing function and replacing it by the simplest possible option: $\mathsf{mix}(A, B) = (B, A)$, i.e., a function that swaps the two halves of its input. This simplification is without loss of generality: all results in this work generalize to arbitrary pure mixing functions with some notational overhead. For completeness, we describe pure mixing functions as defined by Chen et al. [11] in Appendix A.

Algorithm 1 Round function $F[\tilde{E}_K]$

Input: $K \in \mathcal{K}$, $M \in \{0,1\}^{[n,2n-1]}$
Output: $C \in \{0,1\}^{|M|}$

$s \leftarrow |M| - n$
$M_1 \leftarrow \mathsf{left}_n(M)$, $M_2 \leftarrow \mathsf{right}_s(M)$
$Y \leftarrow \tilde{E}_K(\mathsf{pad}(M_2), M_1)$
$C \leftarrow \mathsf{left}_{n-s}(Y) \| M_2 \| \mathsf{right}_s(Y)$
return C

Algorithm 2 Round function $F^{-1}[\tilde{E}_K]$

Input: $K \in \mathcal{K}$, $C \in \{0,1\}^{[n,2n-1]}$
Output: $M \in \{0,1\}^{|C|}$

$s \leftarrow |C| - n$
$C_1 \leftarrow \mathsf{left}_n(C)$, $C_2 \leftarrow \mathsf{right}_s(C)$
$Y \leftarrow \mathsf{left}_{n-s}(C_1) \| C_2$
$M \leftarrow \tilde{E}_K^{-1}(\mathsf{pad}(\mathsf{right}_s(C_1)), Y) \| \mathsf{right}_s(C_1)$
return M

Consider finite key space $\mathcal{K}$ and let $n \in \mathbb{N}$. Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider the round function F (and its inverse) that uses $\tilde{E}_K$ for secret key $K \in \mathcal{K}$, and length-preservingly transforms a plaintext $M \in \{0,1\}^{[n,2n-1]}$ (resp. a ciphertext $C \in \{0,1\}^{[n,2n-1]}$) into a ciphertext C (resp. a plaintext M) as in Algorithm 1 (resp. Algorithm 2). For $r \geq 2$, the r-round length doubler r-LDT is defined as

$$r\text{-LDT}_{\boldsymbol{K}}(M) = F_{K_r} \circ \cdots \circ F_{K_1}(M)\,, \tag{3}$$

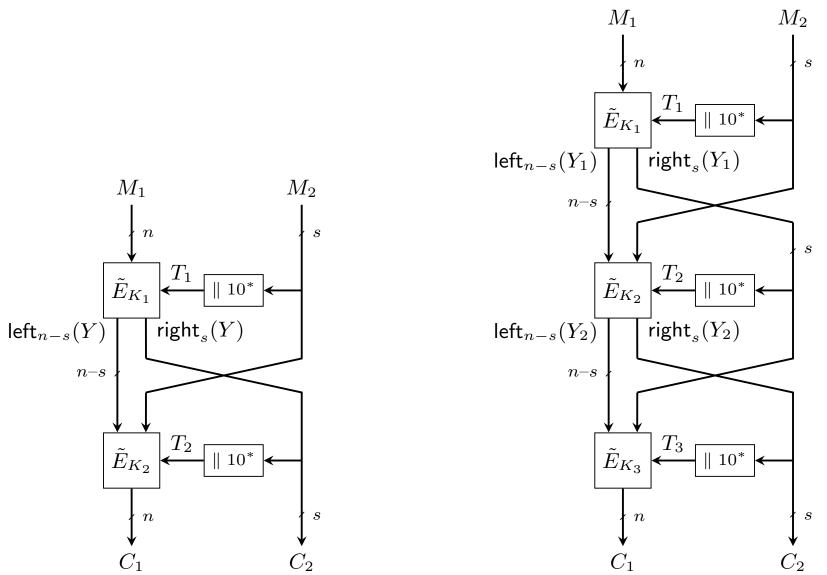
where $\boldsymbol{K} = (K_1, \ldots, K_r) \in \mathcal{K}^r$ and $M \in \{0,1\}^{[n,2n-1]}$. In this evaluation, the mixing of the last round function evaluation is irrelevant for the scheme's security and therefore ignored. For $r = 2$ and $r = 3$, the doubler r-LDT is depicted in Figure 2.

Chen et al. proved that two rounds of LDT (with arbitrary pure mixing) is secure against any adversary making around $2^{n/2}$ queries.

Proposition 1 (Chen et al. [11]). *Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider two-round* 2-LDT. *For any distinguisher $\mathcal{D}$ making at most q queries, there exist distinguishers $\mathcal{D}'_1$ and $\mathcal{D}'_2$ with the same query complexity such that*

$$\mathbf{Adv}^{\mathrm{vsprp}}_{2\text{-LDT}}(\mathcal{D}) \leq \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_1) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_2) + \frac{q(q-1)}{2^n}\,. \tag{4}$$

Chen et al. also presented a distinguisher against 2-LDT that succeeds in approximately $2^{n-s/2}$ queries, where the distinguisher makes queries of size $n+s$ bits. The analysis of this attack supports on earlier proofs and attacks by Hall et al. [22] and Gilboa and Gueron [17] on the truncated permutation construction. The attack only works if the distinguisher takes large enough $s \gg 0$ [11]. In addition, it shows that the birthday bound security analysis is tight for $s \approx n-1$, and that we may only be able to prove beyond birthday bound security for $s \ll n-1$. Based on these observations, in future analyses we will explicitly limit s to a certain range by using lower and upper bounds $s_{\min}$ and $s_{\max}$.



(a) 2-Round LDT. (b) 3-Round LDT.

Fig. 2: Depiction of 2-round and 3-round LDT. Here, $s = |M| - n$.

4 Harmonic Permutation Primitives

In this section we introduce two harmonic permutation primitives: a tweakable SPRP in Section 4.1 and a variable SPRP in Section 4.2.

4.1 Harmonic Tweakable SPRP $G_{a,b}$

We introduce a tweakable pseudorandom permutation $G_{a,b}$ parameterized by $a, b \in \{0,1\}$. The primitive will be used as intermediate in the analysis of 2-LDT (for $(a,b) = (1,0)$ and $(a,b) = (0,1)$) and in the analysis of 3-LDT (for $(a,b) = (1,1)$).

$G_{a,b}$ is a tweakable permutation with n-bit tweaks and data blocks (so $G_{a,b} \in \widetilde{\mathrm{Perm}}(n,n)$). It maintains an initially empty list $\mathcal{L}$ to store all query-response tuples (T, X, Y). For $T \in \{0,1\}^n$, write $\mathsf{dom}(\mathcal{L}_T) = \{X \mid (T, X, \cdot) \in \mathcal{L}\}$ and $\mathsf{rng}(\mathcal{L}_T) = \{Y \mid (T, \cdot, Y) \in \mathcal{L}\}$. The tweakable pseudorandom permutation $G_{a,b}$ on input of a new query is described in Algorithm 3 (forward) and Algorithm 4 (inverse).

Algorithm 3 Harmonic $G_{a,b}$

Input: $T \in \{0,1\}^n \backslash \{0^n\}, X \in \{0,1\}^n$
Output: $Y \in \{0,1\}^n$

$s \leftarrow |\mathsf{unpad}(T)|$
if $a = 0$ **then**
  $Y \xleftarrow{\$} \{0,1\}^n \backslash \mathsf{rng}(\mathcal{L}_T)$
if $a = 1$ **then**
  $Z \xleftarrow{\$} \{0,1\}^s$
  $Y \xleftarrow{\$} \{\{0,1\}^{n-s} \| Z\} \backslash \mathsf{rng}(\mathcal{L}_T)$
$\mathcal{L} \xleftarrow{\cup} (T, X, Y)$
return Y

Algorithm 4 Harmonic $G_{a,b}^{-1}$

Input: $T \in \{0,1\}^n \backslash \{0^n\}, Y \in \{0,1\}^n$
Output: $X \in \{0,1\}^n$

$s \leftarrow |\mathsf{unpad}(T)|$
if $b = 0$ **then**
  $X \xleftarrow{\$} \{0,1\}^n \backslash \mathsf{dom}(\mathcal{L}_T)$
if $b = 1$ **then**
  $Z \xleftarrow{\$} \{0,1\}^s$
  $X \xleftarrow{\$} \{\{0,1\}^{n-s} \| Z\} \backslash \mathsf{dom}(\mathcal{L}_T)$
$\mathcal{L} \xleftarrow{\cup} (T, X, Y)$
return X

In our work, $G_{a,b}$ will never be called for tweak $T = 0^n$, hence the assignment $s \leftarrow |\mathsf{unpad}(T)|$ is sound. If $a = b = 0$, $G_{0,0}$ describes a randomly drawn tweakable permutation from $\widetilde{\mathrm{Perm}}(n, n)$ (lazily sampled). We will use $G_{a,b}$ for the case where a or b is 1.

Lemma 2. *Let $a, b \in \{0,1\}$, and consider $G_{a,b}$. Let $s_{min}, s_{max} \in [0, n-1]$ such that $s_{min} \leq s_{max}$. Let $1 \leq \theta \leq 2^{n-s_{max}-2}$ be an integral threshold. For any distinguisher $\mathcal{D}$ making at most $q \leq 2^{n-1}$ queries, all with tweaks satisfying $|\mathsf{unpad}(T)| \in [s_{min}, s_{max}]$, and restricted to making at most θ inverse queries per tweak (if $a = 1$) and at most θ forward queries per tweak (if $b = 1$), we have*

$$\mathbf{Adv}_{G_{a,b}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}) \leq \begin{cases} 0 & \textit{if } (a,b) = (0,0)\,, \\ \left(\dfrac{2q^3}{2^{2n-s_{max}}}\right)^{1/2} + \dbinom{q}{\theta}\dfrac{1}{2^{(\theta-1)s_{min}}}\,, & \textit{if } (a,b) \in \{(1,0),(0,1)\}\,, \\ \left(\dfrac{4\left(\theta+\theta^2\right)q}{2^{2n-s_{max}}}\right)^{1/2}\,, & \textit{if } (a,b) = (1,1)\,. \end{cases}$$

Note that no limitation is put on the number of times a single tweak is queried in forward direction in case $a = 0$ or in inverse direction in case $b = 0$. The proof will be given in Section 7.

4.2 Harmonic VSPRP Permutation $\boldsymbol{H}$

We introduce a variable pseudorandom permutation H, that operates similarly as $G_{1,1}$, but on domain $\{0,1\}^{[n,2n-1]}$ and without tweak input. H likewise maintains an initially empty list $\mathcal{L}$ to store all query-response tuples (X, Y). For $s \in [0, n-1]$, write $\mathsf{dom}(\mathcal{L}_s) = \{X \in \{0,1\}^{n+s} \mid (X, \cdot) \in \mathcal{L}\}$ and $\mathsf{rng}(\mathcal{L}_s) = \{Y \in \{0,1\}^{n+s} \mid (\cdot, Y) \in \mathcal{L}\}$. The variable pseudorandom permutation H on input of a new query is described in Algorithm 5 (forward) and Algorithm 6 (inverse).

Algorithm 5 Harmonic H

Input: $X \in \{0,1\}^{[n,2n-1]}$
Output: $Y \in \{0,1\}^{|X|}$

$s \leftarrow |X| - n$
$Z \xleftarrow{\$} \{0,1\}^s$
$Y \xleftarrow{\$} \{\{0,1\}^n \| Z\} \backslash \mathsf{rng}(\mathcal{L}_s)$
$\mathcal{L} \xleftarrow{\cup} (X, Y)$
return Y

Algorithm 6 Harmonic H^{-1}

Input: $Y \in \{0,1\}^{[n,2n-1]}$
Output: $X \in \{0,1\}^{|Y|}$

$s \leftarrow |Y| - n$
$Z \xleftarrow{\$} \{0,1\}^s$
$X \xleftarrow{\$} \{\{0,1\}^n \| Z\} \backslash \mathsf{dom}(\mathcal{L}_s)$
$\mathcal{L} \xleftarrow{\cup} (X, Y)$
return X

Lemma 3. *Consider H. Let $s_{min} \in [0, n-1]$. For any distinguisher $\mathcal{D}$ making at most $q \leq 2^{n-1}$ queries, all of length in $[n + s_{min}, 2n - 1]$ bits, we have*

$$\mathbf{Adv}_H^{\mathrm{vsprp}}(\mathcal{D}) \leq \left(\frac{2q^3}{2^{2n+s_{min}}}\right)^{1/2} .$$

The proof will be given in Section 8.

5 2-Round LDT

As main result on 2-LDT, we derive the following reduction to harmonic primitives $G_{a,b}$ and H.

Theorem 1. *Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider two-round 2-LDT. Let $s_{min}, s_{max} \in [0, n-1]$ such that $s_{min} \leq s_{max}$. Let $1 \leq \theta \leq 2^{n-s_{max}-2}$ be an integral threshold. For any distinguisher $\mathcal{D}$ making at most q queries, all of length in $[n+s_{min}, n+s_{max}]$ bits, there exist distinguishers $\mathcal{D}'_1, \ldots, \mathcal{D}'_5$ with the same query complexity such that*

$$\mathbf{Adv}_{\text{2-LDT}}^{\mathrm{vsprp}}(\mathcal{D}) \leq \mathbf{Adv}_{\tilde{E}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_1) + \mathbf{Adv}_{\tilde{E}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_2) + \mathbf{Adv}_H^{\mathrm{vsprp}}(\mathcal{D}'_3) \tag{5a}$$

$$+ \mathbf{Adv}_{G_{1,0}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_4) + \mathbf{Adv}_{G_{0,1}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_5) + \binom{q}{\theta} \frac{1}{2^{(\theta-1)s_{min}}} , \tag{5b}$$

where $\mathcal{D}'_4$ may make at most θ inverse queries per tweak and $\mathcal{D}'_5$ at most θ forward queries per tweak.

We will prove Theorem 1 in Section 5.1. Plugging the bounds of Lemmas 2 and 3 into the equation yields the following corollary.

Corollary 1. *Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider two-round 2-LDT. Let $s_{min}, s_{max} \in [0, n-1]$ such that $s_{min} \leq s_{max}$. Let $1 \leq \theta \leq 2^{n-s_{max}-2}$ be an integral threshold. For any distinguisher $\mathcal{D}$ making at most q queries, all of length in $[n+s_{min}, n+s_{max}]$ bits, there exist distinguishers $\mathcal{D}'_1, \mathcal{D}'_2$ with the same query complexity such that*

$$\mathbf{Adv}_{\text{2-LDT}}^{\mathrm{vsprp}}(\mathcal{D}) \leq \mathbf{Adv}_{\tilde{E}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_1) + \mathbf{Adv}_{\tilde{E}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_2) + \left(\frac{2q^3}{2^{2n+s_{min}}}\right)^{1/2} + 2\left(\frac{2q^3}{2^{2n-s_{max}}}\right)^{1/2} + 3\binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{min}}} .$$

The first two advantages represent the security of the underlying tweakable block cipher $\tilde{E}$. By Stirling's approximation, if $s_{\min} \leq \theta$, the last term satisfies

$$3\binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{\min}}} \leq 3 \cdot 2^{s_{\min}} \cdot \left(\frac{qe}{\theta 2^{s_{\min}}}\right)^{\theta} \leq 3\left(\frac{2qe}{\theta 2^{s_{\min}}}\right)^{\theta} .$$

As the term decreases with θ but θ is limited by side condition $\theta \leq 2^{n-s_{\max}-2}$, it makes sense to choose $\theta = 2^{n-s_{\max}-2}$, and this term equals

$$3\left(\frac{8qe}{2^{n+s_{\min}-s_{\max}}}\right)^{\theta} .$$

We obtain security up to approximately $\min\left\{\frac{2n+s_{\min}}{3}, \frac{2n-s_{\max}}{3}, n+s_{\min}-s_{\max}\right\}$ bits, provided that $s_{\min} \leq 2^{n-s_{\max}-2}$. For $s_{\max} \geq n/2$, the middle term dominates and we achieve $n/2$-bit security at most. In this case, the bound of Chen et al. [11] is better. For $s_{\max} < n/2$, our bound guarantees up to at most $2n/3$ bits of security, depending of the choice of $s_{\max}$, where $s_{\min}$ is adapted to $s_{\min} \leq 2^{n-s_{\max}-2}$.

5.1 Proof of Theorem 1

Consider any distinguisher $\mathcal{D}$ making at most q queries, all of length in $[n+s_{\min}, n+s_{\max}]$ bits. It has access to either 2-LDT$_{\boldsymbol{K}}$ for $\boldsymbol{K} = (K_1, K_2) \xleftarrow{\$} \mathcal{K}^2$ or a random length-preserving invertible permutation $\rho \xleftarrow{\$} \mathrm{VPerm}([n \ldots 2n-1])$. For ease of discussion, write

$$\text{2-LDT}_{\boldsymbol{K}} = \mathcal{E}[\tilde{E}_{K_1}, \tilde{E}_{K_2}] .$$

Let $\tilde{\pi}_1, \tilde{\pi}_2 \xleftarrow{\$} \widetilde{\mathrm{Perm}}(n, n)$. We have

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{vsprp}}_{\text{2-LDT}}(\mathcal{D}) &= \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{E}_{K_1}, \tilde{E}_{K_2}]^{\pm} \ ; \ \rho^{\pm}\Big) \\
&\leq \Delta_{\mathcal{D}'_1}\Big(\tilde{E}^{\pm}_{K_1} \ ; \ \tilde{\pi}^{\pm}_1\Big) + \Delta_{\mathcal{D}'_2}\Big(\tilde{E}^{\pm}_{K_2} \ ; \ \tilde{\pi}^{\pm}_2\Big) + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \rho^{\pm}\Big) \\
&= \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_1) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_2) + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \rho^{\pm}\Big) , \qquad (6)
\end{aligned}$$

for some distinguishers $\mathcal{D}'_1$ and $\mathcal{D}'_2$ with the same query complexity as $\mathcal{D}$.

We will focus on the remaining distance in (6). Without loss of generality, we will consider computationally unbounded and deterministic distinguishers. Consider three harmonic primitives, $G_{1,0}$ and $G_{0,1}$ of Section 4.1 and H of Section 4.2. We obtain via the triangle inequality:

$$\begin{aligned}
\Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \rho^{\pm}\Big) &\leq \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}\Big) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[G_{1,0}, G_{0,1}]^{\pm} \ ; \ H^{\pm}\Big) + \Delta_{\mathcal{D}}\Big(H^{\pm} \ ; \ \rho^{\pm}\Big) \\
&= \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}\Big) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[G_{1,0}, G_{0,1}]^{\pm} \ ; \ H^{\pm}\Big) + \mathbf{Adv}^{\mathrm{vsprp}}_{H}(\mathcal{D}'_3) , \qquad (7)
\end{aligned}$$

for some distinguisher $\mathcal{D}'_3$ with the same query complexity as $\mathcal{D}$ (in fact, $\mathcal{D}'_3 = \mathcal{D}$).

Below two claims bound the remaining distances in (7) and complete the proof.

Claim. We have $\Delta_{\mathcal{D}}\Big(\mathcal{E}[G_{1,0}, G_{0,1}]^{\pm} \ ; \ H^{\pm}\Big) = 0$.

Proof (of claim). For any query to H of length $n+s$ bits (either forward or inverse), the last s bits of the response are drawn uniformly at random from $\{0,1\}^s$ and the first n bits are drawn uniformly at random in such a way that the permutativity of H is retained (see Algorithms 5 and 6). Consider any query to $\mathcal{E}[G_{1,0}, G_{0,1}]$, without loss of generality a forward query of length $n+s$ bits. The s rightmost bits of the output equal the s rightmost bits of $G_{1,0}$, and are generated uniformly at random (see Algorithm 3). Denote this s-bit block by C_2. The remaining n bits of the response, say C_1, come from the evaluation of $G_{0,1}$ for tweak C_2, on input of a data block that never appeared for this tweak before. As can be deduced from Algorithm 3, $G_{0,1}$ behaves like a tweakable permutation: for every tweak input, it behaves like a permutation. Therefore, C_1 is generated uniformly at random in such a way that $C_1 \| C_2$ has never appeared before. Concluding, $\mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}$ and $H^{\pm}$ follow identical distributions. □

Claim. We have

$$\Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm} \ ; \ \mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}\Big) \leq \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{1,0}}(\mathcal{D}'_4) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{0,1}}(\mathcal{D}'_5) + \binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{\min}}}\,,$$

for some distinguishers $\mathcal{D}'_4$ and $\mathcal{D}'_5$ with the same query complexity as $\mathcal{D}$, where $\mathcal{D}'_4$ may make at most θ inverse queries per tweak and $\mathcal{D}'_5$ at most θ forward queries per tweak.

Proof (of claim). Consider a computationally unbounded and deterministic distinguisher $\mathcal{D}$ making at most q queries. It has access to either $\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm}$ or $\mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}$. Summarize the queries in a transcript $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(q)})$, where the i-th tuple $\tau^{(i)} = (\ell^{(i)}, X^{(i)}, Y^{(i)})$ is comprised of a bit $\ell^{(i)} \in \{-1,1\}$ denoting the direction of the query, $X^{(i)}$ is the query input and $Y^{(i)}$ the query output, in such a way $Y^{(i)} = \mathcal{O}^{\ell^{(i)}}(X^{(i)})$. Write $s^{(i)} = |X^{(i)}| - n$. We assume that the distinguisher $\mathcal{D}$ does not repeat any query, which means that $\boldsymbol{\tau^{(i)}}$ does not contain duplicate elements.

For the threshold θ of the theorem statement, define the following bad event:

$$\mathsf{BAD}\ :\ \max_{\ell\in\{-1,1\}}\ \max_{s\in[s_{\min},s_{\max}]}\ \max_{Z\in\{0,1\}^s}\Big|\{i \mid \ell^{(i)} = \ell \wedge s^{(i)} = s \wedge \mathsf{right}_s(Y^{(i)}) = Z\}\Big| > \theta\,.$$

Clearly, writing $\mathcal{O}_{\tilde{\pi}} = \mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2]^{\pm}$ and $\mathcal{O}_G = \mathcal{E}[G_{1,0}, G_{0,1}]^{\pm}$ for brevity,

$$\begin{aligned}
\Delta_{\mathcal{D}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big) &= \left|\Pr\left[\mathcal{D}^{\mathcal{O}_{\tilde{\pi}}} = 1\right] - \Pr\left[\mathcal{D}^{\mathcal{O}_G} = 1\right]\right| \\
&\leq \left|\Pr\left[\mathcal{D}^{\mathcal{O}_{\tilde{\pi}}} = 1 \wedge \neg\mathsf{BAD}\right] - \Pr\left[\mathcal{D}^{\mathcal{O}_G} = 1 \wedge \neg\mathsf{BAD}\right]\right| \\
&\quad + \left|\Pr\left[\mathcal{D}^{\mathcal{O}_{\tilde{\pi}}} = 1 \wedge \mathsf{BAD}\right] - \Pr\left[\mathcal{D}^{\mathcal{O}_G} = 1 \wedge \mathsf{BAD}\right]\right| \\
&\leq \left|\Pr\left[\mathcal{D}^{\mathcal{O}_{\tilde{\pi}}} = 1 \wedge \neg\mathsf{BAD}\right] - \Pr\left[\mathcal{D}^{\mathcal{O}_G} = 1 \wedge \neg\mathsf{BAD}\right]\right| \\
&\quad + \max\Big\{\Pr\left[\mathcal{O}_{\tilde{\pi}} \text{ sets } \mathsf{BAD}\right], \Pr\left[\mathcal{O}_G \text{ sets } \mathsf{BAD}\right]\Big\}. \qquad (8)
\end{aligned}$$

Denoting the distance in (8) by $\Delta_{\mathcal{D}}^{\neg\mathsf{BAD}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big)$ for brevity, a straightforward triangle argument shows that

$$\Delta_{\mathcal{D}}^{\neg\mathsf{BAD}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big) \leq \mathbf{Adv}_{G_{1,0}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_4) + \mathbf{Adv}_{G_{0,1}}^{\widetilde{\mathrm{sprp}}}(\mathcal{D}'_5), \qquad (9)$$

for some distinguishers $\mathcal{D}'_4$ and $\mathcal{D}'_5$ with the same query complexity as $\mathcal{D}$, where $\mathcal{D}'_4$ may make at most θ inverse queries per tweak and $\mathcal{D}'_5$ at most θ forward queries per tweak. These two restrictions follow from the way $\mathcal{E}$ evaluates its primitives ($\tilde{\pi}_1, \tilde{\pi}_2$ in the left oracle and $G_{1,0}, G_{0,1}$ in the right oracle) and from the conditioning of the bad event.

Consider the max-term in (8). Consider any $\ell \in \{-1, 1\}$ and $s \in [s_{\min}, s_{\max}]$, and denote the number of queries with $\ell^{(i)} = \ell$ and $s^{(i)} = s$ by $q_{\ell,s}$. For $\mathcal{O}_G$, in forward queries the $\mathsf{right}_s(Y^{(i)})$-values come from the evaluation of $G_{1,0}$ and are always uniformly randomly drawn (see Algorithm 3), whereas in inverse queries they come from evaluations of $G_{0,1}^{-1}$ and are also uniformly randomly drawn (see Algorithm 4). Therefore,

$$\Pr\left[\mathcal{O}_G \text{ sets } \mathsf{BAD} \text{ for } (\ell, s)\right] \leq \binom{q_{\ell,s}}{\theta} \frac{1}{2^{(\theta-1)s}}.$$

On the other hand, for $\mathcal{O}_{\tilde{\pi}}$, the $\mathsf{right}_s(Y^{(i)})$-values come from a truncated permutation evaluation, and we find

$$\Pr\left[\mathcal{O}_{\tilde{\pi}} \text{ sets } \mathsf{BAD} \text{ for } (\ell, s)\right] = \binom{q_{\ell,s}}{\theta} \cdot 2^s \cdot \prod_{i=0}^{\theta-1} \frac{2^{n-s} - i}{2^n - i} \leq \binom{q_{\ell,s}}{\theta} \frac{1}{2^{(\theta-1)s}}.$$

We thus obtain

$$\max\Big\{\Pr\left[\mathcal{O}_{\tilde{\pi}} \text{ sets } \mathsf{BAD}\right], \Pr\left[\mathcal{O}_G \text{ sets } \mathsf{BAD}\right]\Big\} \leq \binom{q}{\theta} \frac{1}{2^{(\theta-1)s_{\min}}}, \qquad (10)$$

using that $\binom{q_a}{\theta} + \binom{q_b}{\theta} \leq \binom{q_a+q_b}{\theta}$ and the distinguisher maximizes its probability for $s = s_{\min}$. The proof is concluded by combining (8), (9), and (10). $\square$

6 3-Round LDT

We derive the following reduction from the security of 3-LDT to harmonic primitives $G_{a,b}$ and H.

Theorem 2. *Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider three-round* 3-LDT. *Let $s_{min}, s_{max} \in [0, n-1]$ such that $s_{min} \le s_{max}$. Let $1 \le \theta \le 2^{n-s_{max}-2}$ be an integral threshold. For any distinguisher $\mathcal{D}$ making at most q queries, all of length in $[n{+}s_{min}, n{+}s_{max}]$ bits, there exist distinguishers $\mathcal{D}'_1, \dots, \mathcal{D}'_5$ with the same query complexity such that*

$$\mathbf{Adv}^{\mathrm{vsprp}}_{\text{3-LDT}}(\mathcal{D}) \le \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_1) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_2) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_3) \tag{11a}$$

$$+ \mathbf{Adv}^{\mathrm{vsprp}}_{H}(\mathcal{D}'_4) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{1,1}}(\mathcal{D}'_5) + \binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{min}}}\,, \tag{11b}$$

where $\mathcal{D}'_5$ may make at most θ forward and θ inverse queries per tweak.

We will prove Theorem 2 in Section 6.1.

The improvement of the bound of 3-LDT over that of 2-LDT of Theorem 1 is readily visible: $\mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{1,0}} + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{0,1}}$ has been replaced with $\mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{1,1}}$, which by Lemma 2 achieves a better bound. Plugging the bounds of Lemmas 2 and 3 into the equation yields the following corollary.

Corollary 2. *Let $\tilde{E} : \mathcal{K} \times \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a tweakable block cipher. Consider three-round* 3-LDT. *Let $s_{min}, s_{max} \in [0, n-1]$ such that $s_{min} \le s_{max}$. Let $1 \le \theta \le 2^{n-s_{max}-2}$ be an integral threshold. For any distinguisher $\mathcal{D}$ making at most q queries, all of length in $[n{+}s_{min}, n{+}s_{max}]$ bits, there exist distinguishers $\mathcal{D}'_1, \mathcal{D}'_2, \mathcal{D}'_3$ with the same query complexity such that*

$$\mathbf{Adv}^{\mathrm{vsprp}}_{\text{3-LDT}}(\mathcal{D}) \le \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_1) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_2) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_3)$$
$$+ \left(\frac{2q^3}{2^{2n+s_{min}}}\right)^{1/2} + \left(\frac{4\left(\theta+\theta^2\right)q}{2^{2n-s_{max}}}\right)^{1/2} + \binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{min}}}\,.$$

The first three advantages represent the security of the underlying tweakable block cipher $\tilde{E}$. Two of the terms in the remaining portion of the bound depend on θ: the first one increases with θ whereas the latter decreases with θ. Recalling from Section 5 that, for $s_{\min} \le \theta$,

$$\binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{\min}}} \le \left(\frac{2qe}{\theta 2^{s_{\min}}}\right)^{\theta}\,,$$

equating the two θ-dependent fractions in the corollary gives $\theta \approx 2^{(2n-s_{\max}-s_{\min})/3}$. This threshold value still has to obey to the condition $s_{\min} \le \theta \le 2^{n-s_{\max}-2}$, or stated differently,

$$s_{\max} \le \min\left\{(n-6+s_{\min})/2\,,\, 2n-s_{\min}-3\log_2(s_{\min})\right\}. \tag{12}$$

The minimum is achieved for its left element as long as $s_{\min} \le n - 2\log_2(n)$. In Table 1, we list the simplified security bound of Corollary 2 (omitting constants) for $s_{\min} \in \{\mathit{const}, n/4, n/2, 3n/4, n-2\log_2(n)\}$ and three possible choices of $s_{\max}$: arbitrary, $s_{\max} \approx (n+s_{\min})/2$ of (12), and $s_{\max} \approx s_{\min}$. For $s_{\min}$ approaching n, n-bit security is achieved.

Table 1: Interpretation of the bound of Corollary 2 for various choices of $s_{\min}$, where *const* is a constant to make the bound meaningful. Small constants are omitted in the security upper bound.

$s_{\min}$	security up to		
	arbitrary $s_{\max}$	$s_{\max} \approx \frac{n+s_{\min}}{2}$ of (12)	$s_{\max} \approx s_{\min}$
const	$\min\left\{\frac{8n}{12}, \frac{4n}{6} - \frac{s_{\max}}{3}\right\}$	$\frac{n}{2}$	$\frac{2n}{3}$
$\frac{n}{4}$	$\min\left\{\frac{9n}{12}, \frac{5n}{6} - \frac{s_{\max}}{3}\right\}$	$\frac{5n}{8}$	$\frac{3n}{4}$
$\frac{n}{2}$	$\min\left\{\frac{10n}{12}, \frac{6n}{6} - \frac{s_{\max}}{3}\right\}$	$\frac{3n}{4}$	$\frac{5n}{6}$
$\frac{3n}{4}$	$\min\left\{\frac{11n}{12}, \frac{7n}{6} - \frac{s_{\max}}{3}\right\}$	$\frac{7n}{8}$	$\frac{11n}{12}$
$n - 2\log_2(n)$	$\min\left\{\frac{12n}{12}, \frac{8n}{6} - \frac{s_{\max}}{3}\right\}$	n	n

Note that these two choices of $s_{\max}$ set its two extremes: for given $s_{\min}$, we require that $s_{\min} \le s_{\max} \le (n + s_{\min})/2$. The security bounds for the two extremes are plotted in Figure 3: the level of security given by Corollary 2 is in the shaded area of Figure 3 and depends on $s_{\min}$ and $s_{\max}$. For example, fixing $s_{\min} = n/2$, the security bound of Corollary 2 lies between $3n/4$ (for $s_{\max} \approx (n+s_{\min})/2$) and $5n/6$ (for $s_{\max} \approx s_{\min}$), using that $s_{\min} \le s_{\max} \le (n+s_{\min})/2$ by condition.

6.1 Proof of Theorem 2

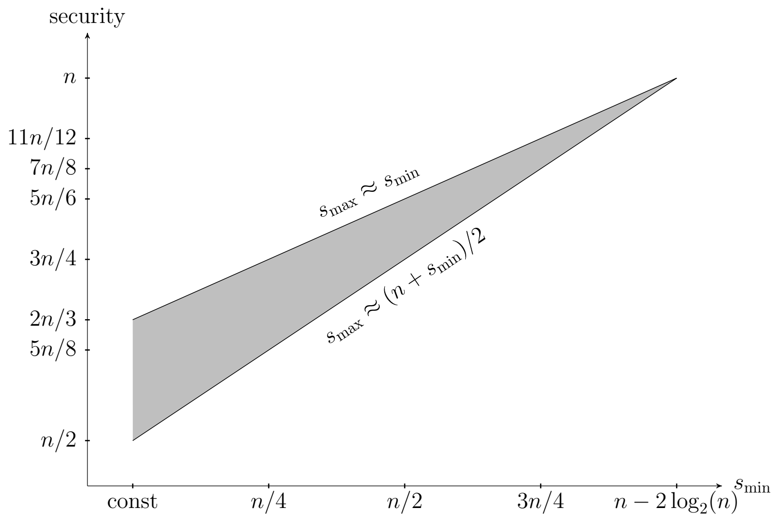
The first steps of the proof resemble those of Section 5.1. Consider any distinguisher $\mathcal{D}$ making at most q queries. It has access to either 3-LDT$_{\boldsymbol{K}}$ for $\boldsymbol{K} = (K_1, K_2, K_3) \xleftarrow{\$} \mathcal{K}^3$ or a random length-preserving invertible permutation $\rho \xleftarrow{\$} \mathrm{VPerm}([n \dots 2n-1])$. For ease of discussion, write

$$\text{3-LDT}_{\boldsymbol{K}} = \mathcal{E}[\tilde{E}_{K_1}, \tilde{E}_{K_2}, \tilde{E}_{K_3}]\,.$$

Let $\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3 \xleftarrow{\$} \widetilde{\mathrm{Perm}}(n, n)$. We have

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{vsprp}}_{\text{3-LDT}}(\mathcal{D}) &= \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{E}_{K_1}, \tilde{E}_{K_2}, \tilde{E}_{K_3}]^{\pm}\ ;\ \rho^{\pm}\Big) \\
&\le \Delta_{\mathcal{D}'_1}\Big(\tilde{E}^{\pm}_{K_1}\ ;\ \tilde{\pi}^{\pm}_1\Big) + \Delta_{\mathcal{D}'_2}\Big(\tilde{E}^{\pm}_{K_2}\ ;\ \tilde{\pi}^{\pm}_2\Big) + \Delta_{\mathcal{D}'_3}\Big(\tilde{E}^{\pm}_{K_3}\ ;\ \tilde{\pi}^{\pm}_3\Big) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm}\ ;\ \rho^{\pm}\Big) \\
&= \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_1) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_2) + \mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{\tilde{E}}(\mathcal{D}'_3) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm}\ ;\ \rho^{\pm}\Big)\,, \qquad (13)
\end{aligned}$$

Fig. 3: Simplified security bound of 3-LDT for various choices of $s_{\min}$. Lower line is for $s_{\max} \approx (n + s_{\min})/2$, upper line for $s_{\max} \approx s_{\min}$. Security of 3-LDT is indicated by the shaded area and depends on $s_{\min}$ and $s_{\max}$, where $s_{\min} \leq s_{\max} \leq (n + s_{\min})/2$.



for some distinguishers $\mathcal{D}_1'$, $\mathcal{D}_2'$, and $\mathcal{D}_2'$ with the same query complexity as $\mathcal{D}$.

We will focus on the remaining distance in (13). Without loss of generality, we will consider computationally unbounded and deterministic distinguishers. Consider two harmonic primitives, $G_{1,1}$ of Section 4.1 and H of Section 4.2. We obtain via the triangle inequality:

$$\begin{aligned}
\Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm} \;;\; \rho^{\pm}\Big) &\leq \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm} \;;\; \mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}\Big) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm} \;;\; H^{\pm}\Big) + \Delta_{\mathcal{D}}\Big(H^{\pm} \;;\; \rho^{\pm}\Big) \\
&= \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm} \;;\; \mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}\Big) \\
&\quad + \Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm} \;;\; H^{\pm}\Big) + \mathbf{Adv}_{H}^{\mathrm{vsprp}}(\mathcal{D}_4')\,,
\end{aligned} \tag{14}$$

for some distinguisher $\mathcal{D}_4'$ with the same query complexity as $\mathcal{D}$ (in fact, $\mathcal{D}_4' = \mathcal{D}$).

Below two claims bound the remaining distances in (14) and complete the proof.

Claim. We have $\Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm} \;;\; H^{\pm}\Big) = 0$.

Proof (of claim). For any query to H of length $n + s$ bits (either forward or inverse), the last s bits of the response are drawn uniformly at random from $\{0,1\}^s$ and the first n bits are drawn uniformly at random in such a way that the permutativity of H is retained (see Algorithms 5 and 6). Consider any query to $\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]$, without loss of generality a forward query of length $n+s$ bits. The s rightmost bits of the output equal the s rightmost bits of $G_{1,1}$, and are generated uniformly at random (see Algorithm 3; here, we make explicit use of the fact that $G_{1,1}$ never receives the same input twice). Denote this s-bit block by C_2. The remaining n bits of the response, say C_1, come from the evaluation of $\tilde{\pi}_3$ for tweak C_2 on input of a data block that never appeared for this tweak before. For every tweak input, the tweakable permutation $\tilde{\pi}_3$ behaves like a permutation. Therefore, C_1 is generated uniformly at random in such a way that $C_1 \| C_2$ has never appeared before. Concluding, $\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}$ and $H^{\pm}$ follow identical distributions. □

Claim. We have

$$\Delta_{\mathcal{D}}\Big(\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm} \ ; \ \mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}\Big) \leq \mathbf{Adv}^{\widetilde{\text{sprp}}}_{G_{1,1}}(\mathcal{D}'_5) + \binom{q}{\theta} \frac{1}{2^{(\theta-1)s_{\min}}} ,$$

for some distinguisher $\mathcal{D}'_5$ with the same query complexity as $\mathcal{D}$, but that may make at most θ forward and θ inverse queries per tweak.

Proof (of claim). The first part of the proof resembles that of the corresponding claim in Section 5.1.

Consider a computationally unbounded and deterministic distinguisher $\mathcal{D}$ making at most q queries. It has access to either $\mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm}$ or $\mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}$. Summarize the queries in a transcript $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(q)})$, where the i-th tuple $\tau^{(i)} = (\ell^{(i)}, X^{(i)}, Y^{(i)})$ is comprised of a bit $\ell^{(i)} \in \{-1, 1\}$ denoting the direction of the query, $X^{(i)}$ is the query input and $Y^{(i)}$ the query output, in such a way $Y^{(i)} = \mathcal{O}^{\ell^{(i)}}(X^{(i)})$. Write $s^{(i)} = |X^{(i)}| - n$. We further denote by $Z^{(i)}$ the last s bits of the output of $\tilde{\pi}_1$ (which is also the last s bits of the input of $\tilde{\pi}_3$) in forward queries, and the last s bits of the output of $\tilde{\pi}_3^{-1}$ (which is also the last s bits of the input of $\tilde{\pi}_1^{-1}$) in inverse queries. We assume that the distinguisher $\mathcal{D}$ does not repeat any query, which means that $\boldsymbol{\tau^{(i)}}$ does not contain duplicate elements.

For the threshold θ of the theorem statement, define the following bad event:

$$\mathsf{BAD} \ : \ \max_{\ell \in \{-1,1\}} \ \max_{s \in [s_{\min}, s_{\max}]} \ \max_{Z \in \{0,1\}^s} \Big| \{ i \mid \ell^{(i)} = \ell \wedge s^{(i)} = s \wedge Z^{(i)} = Z \} \Big| > \theta \, .$$

As before, writing $\mathcal{O}_{\tilde{\pi}} = \mathcal{E}[\tilde{\pi}_1, \tilde{\pi}_2, \tilde{\pi}_3]^{\pm}$ and $\mathcal{O}_G = \mathcal{E}[\tilde{\pi}_1, G_{1,1}, \tilde{\pi}_3]^{\pm}$,

$$\begin{aligned} \Delta_{\mathcal{D}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big) \leq & \left| \Pr\left[\mathcal{D}^{\mathcal{O}_{\tilde{\pi}}} = 1 \wedge \neg\mathsf{BAD}\right] - \Pr\left[\mathcal{D}^{\mathcal{O}_G} = 1 \wedge \neg\mathsf{BAD}\right] \right| \\ & + \max\Big\{ \Pr\left[\mathcal{O}_{\tilde{\pi}} \text{ sets } \mathsf{BAD}\right] , \Pr\left[\mathcal{O}_G \text{ sets } \mathsf{BAD}\right] \Big\} . \end{aligned} \tag{15}$$

Denoting the distance in (15) by $\Delta^{\neg\mathsf{BAD}}_{\mathcal{D}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big)$ for brevity, a straightforward triangle argument shows that

$$\Delta^{\neg\mathsf{BAD}}_{\mathcal{D}}\Big(\mathcal{O}_{\tilde{\pi}} \ ; \ \mathcal{O}_G\Big) \leq \mathbf{Adv}^{\widetilde{\text{sprp}}}_{G_{1,1}}(\mathcal{D}'_5) , \tag{16}$$

for some distinguisher $\mathcal{D}_5'$ with the same query complexity as $\mathcal{D}$, but that may make at most θ forward and θ inverse queries per tweak. These two restrictions follow from the way $\mathcal{E}$ evaluates its primitives ($\tilde{\pi}_2$ in the left oracle and $G_{1,1}$ in the right oracle) and from the conditioning of the bad event.

Consider the max-term in (15). Consider any $\ell \in \{-1, 1\}$ and $s \in [s_{\min}, s_{\max}]$, and denote the number of queries with $\ell^{(i)} = \ell$ and $s^{(i)} = s$ by $q_{\ell,s}$. For both $\mathcal{O}_{\tilde{\pi}}$ and $\mathcal{O}_G$, the $Z^{(i)}$-values come from a truncated permutation evaluation, and we find for $\mathcal{O} \in \{\mathcal{O}_{\tilde{\pi}}, \mathcal{O}_G\}$:

$$\Pr[\mathcal{O} \text{ sets } \mathsf{BAD} \text{ for } (\ell, s)] = \binom{q_{\ell,s}}{\theta} \cdot 2^s \cdot \prod_{i=0}^{\theta-1} \frac{2^{n-s} - i}{2^n - i} \leq \binom{q_{\ell,s}}{\theta} \frac{1}{2^{(\theta-1)s}} \,.$$

We thus obtain

$$\max\Big\{ \Pr[\mathcal{O}_{\tilde{\pi}} \text{ sets } \mathsf{BAD}] \,,\, \Pr[\mathcal{O}_G \text{ sets } \mathsf{BAD}] \Big\} \leq \binom{q}{\theta} \frac{1}{2^{(\theta-1)s_{\min}}} \,, \tag{17}$$

as before. The proof is concluded by combining (15), (16), and (17). □

7 Proof of Lemma 2 on $G_{a,b}$

For $a = b = 0$, the lemma is trivial. Let $a, b \in \{0, 1\}$ such that $a + b \geq 1$, and consider any distinguisher $\mathcal{D}$ making at most q queries, all with tweaks satisfying $|\mathsf{unpad}(T)| \in [s_{\min}, s_{\max}]$, and it makes at most θ inverse queries per tweak (if $a = 1$) and at most θ forward queries per tweak (if $b = 1$). The distinguisher has access to either random system $G_{a,b}$ or $\tilde{\pi} \xleftarrow{\$} \widetilde{\mathrm{Perm}}(n, n)$, and without loss of generality, $\mathcal{D}$ is computationally unbounded and deterministic.

We will use the chi-squared method of Section 2.2, with $\mathcal{O}_2 = G_{a,b}$ being the real system and $\mathcal{O}_0 = \tilde{\pi}$ the ideal system. Define an intermediate world $\mathcal{O}_1$ that implements $\mathcal{O}_2$, unless some event "BAD" (defined below) happens, from which point it implements $\mathcal{O}_0$. Summarize the communication of $\mathcal{D}$ with its oracle in a transcript $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(q)})$, where $\tau^{(i)} = (\ell^{(i)}, T^{(i)}, X^{(i)}, Y^{(i)})$ consists of a bit $\ell^{(i)} \in \{-1, 1\}$ indicating the direction of the query (1 means forward, and -1 means inverse), a tweak value $T^{(i)}$, an input value $X^{(i)}$, and a response $Y^{(i)}$, in such a way that $O^{\ell^{(i)}}(T^{(i)}, X^{(i)}) = Y^{(i)}$ for $\mathcal{O} \in \{\mathcal{O}_0, \mathcal{O}_2\}$.

By a triangle inequality,

$$\begin{aligned}\mathbf{Adv}^{\widetilde{\mathrm{sprp}}}_{G_{a,b}}(\mathcal{D}) &= \|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_2,\mathcal{D}}(\cdot)\| \\ &\leq \|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| + \|p_{\mathcal{O}_1,\mathcal{D}}(\cdot) - p_{\mathcal{O}_2,\mathcal{D}}(\cdot)\| \,.\end{aligned} \tag{18}$$

Let $\mathcal{T}$ denote the set of all possible transcripts, and $\mathcal{T}_{\mathrm{bad}}$ the set of all transcripts that satisfy BAD. We have that $p_{\mathcal{O}_1,\mathcal{D}}(\boldsymbol{\tau}) = p_{\mathcal{O}_2,\mathcal{D}}(\boldsymbol{\tau})$ for any $\boldsymbol{\tau} \in \mathcal{T} \backslash \mathcal{T}_{\mathrm{bad}}$, and hence,

$$\begin{aligned}\|p_{\mathcal{O}_1,\mathcal{D}}(\cdot) - p_{\mathcal{O}_2,\mathcal{D}}(\cdot)\| &= \sum_{\boldsymbol{\tau}\in\mathcal{T}} \max\{0, p_{\mathcal{O}_1,\mathcal{D}}(\boldsymbol{\tau}) - p_{\mathcal{O}_2,\mathcal{D}}(\boldsymbol{\tau})\} \\ &= \sum_{\boldsymbol{\tau}\in\mathcal{T}_{\rm bad}} \max\{0, p_{\mathcal{O}_1,\mathcal{D}}(\boldsymbol{\tau}) - p_{\mathcal{O}_2,\mathcal{D}}(\boldsymbol{\tau})\} \\ &\leq \sum_{\boldsymbol{\tau}\in\mathcal{T}_{\rm bad}} p_{\mathcal{O}_1,\mathcal{D}}(\boldsymbol{\tau}) = \Pr\left[\mathcal{O}_1 \text{ sets } \mathsf{BAD}\right] .\end{aligned}$$

We obtain from (18):

$$\mathbf{Adv}^{\widetilde{\rm sprp}}_{G_{a,b}}(\mathcal{D}) \leq \|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| + \Pr\left[\mathcal{O}_1 \text{ sets } \mathsf{BAD}\right] . \tag{19}$$

We will formalize and analyze BAD in Section 7.1 and analyze the remaining distance using the chi-squared technique in Section 7.2. These will immediately conclude the proof by (19).

7.1 Bad Transcripts

For the threshold θ of the theorem statement, define the following bad events:

$$\mathsf{BAD}_1 : \max_{s\in[s_{\min},s_{\max}]} \max_{Z,Z'\in\{0,1\}^s} \left|\{i \mid a=1 \wedge \ell^{(i)}=1 \wedge \mathsf{unpad}(T^{(i)})=Z \wedge \mathsf{right}_s(Y^{(i)})=Z'\}\right| > \theta\,,$$

$$\mathsf{BAD}_2 : \max_{s\in[s_{\min},s_{\max}]} \max_{Z,Z'\in\{0,1\}^s} \left|\{i \mid a=1 \wedge \ell^{(i)}=-1 \wedge \mathsf{unpad}(T^{(i)})=Z \wedge \mathsf{right}_s(X^{(i)})=Z'\}\right| > \theta\,,$$

$$\mathsf{BAD}_3 : \max_{s\in[s_{\min},s_{\max}]} \max_{Z,Z'\in\{0,1\}^s} \left|\{i \mid b=1 \wedge \ell^{(i)}=1 \wedge \mathsf{unpad}(T^{(i)})=Z \wedge \mathsf{right}_s(X^{(i)})=Z'\}\right| > \theta\,,$$

$$\mathsf{BAD}_4 : \max_{s\in[s_{\min},s_{\max}]} \max_{Z,Z'\in\{0,1\}^s} \left|\{i \mid b=1 \wedge \ell^{(i)}=-1 \wedge \mathsf{unpad}(T^{(i)})=Z \wedge \mathsf{right}_s(Y^{(i)})=Z'\}\right| > \theta\,.$$

Define $\mathsf{BAD} = \mathsf{BAD}_1 \vee \mathsf{BAD}_2 \vee \mathsf{BAD}_3 \vee \mathsf{BAD}_4$.

The bad events look complicated, but in fact they are not. If $(a,b)=(0,0)$, none of the four bad events are satisfied, and BAD does not hold by construction. On the other hand, if $(a,b)=(1,1)$, the distinguisher makes at most θ forward queries per tweak and at most θ inverse queries per tweak, and also in this case BAD does not hold by construction. The cases $(a,b)=(1,0),(0,1)$ are symmetric, and we treat the former only. If $(a,b)=(1,0)$, $\mathsf{BAD}_3,\mathsf{BAD}_4$ do not hold as $b=0$, and BAD_2 does not hold as the distinguisher makes at most θ inverse queries. We are left with BAD_1. Consider any $s\in[s_{\min},s_{\max}]$ and any $Z\in\{0,1\}^s$, and denote the number of queries with $\ell^{(i)}=1$ and $\mathsf{unpad}(T^{(i)})=Z$ by $q_{s,Z}$. The $\mathsf{right}_s(Y^{(i)})$-values come from the evaluation of $G_{1,0}$ and are always uniformly randomly drawn (see Algorithm 3). Therefore,

$$\Pr\left[\mathcal{O}_1 \text{ sets } \mathsf{BAD} \text{ for } (s,Z)\right] \leq \binom{q_{s,Z}}{\theta}\frac{1}{2^{(\theta-1)s}}\,.$$

We thus obtain

$$\Pr\left[\mathcal{O}_1 \text{ sets } \mathsf{BAD}\right] \leq \binom{q}{\theta}\frac{1}{2^{(\theta-1)s_{\min}}}\,, \tag{20}$$

using that $\binom{q_a}{\theta} + \binom{q_b}{\theta} \leq \binom{q_a+q_b}{\theta}$ and the distinguisher maximizes its probability for $s = s_{\min}$. Recalling that the case of $(a,b) = (0,1)$ is symmetric and that BAD is not set for $(a,b) = (0,0),(1,1)$, we obtain

$$\Pr[\mathcal{O}_1 \text{ sets } \mathsf{BAD}] \leq |a-b| \cdot \binom{q}{\theta} \frac{1}{2^{(\theta-1)s_{\min}}}\,.$$

7.2 Distance Between $\mathcal{O}_0$ and $\mathcal{O}_1$

Our aim is to bound the term of (1). Consider a given transcript $\boldsymbol{\tau}^{(i-1)}$, which in turn determines the values $\ell^{(i)}$, $T^{(i)}$, and $X^{(i)}$. Let $s = |\mathsf{unpad}(T^{(i)})|$, and consider any value $Y^{(i)}$. As both oracles behave independently for different tweaks, it suffices to focus on earlier queries of the same tweak. We additionally refine into the number of queries with the same or opposite query direction. Let

$$\begin{aligned} i_{\text{pos}} &= \left|\left\{ j \in \{1,\dots,i-1\} \mid T^{(j)} = T^{(i)} \wedge \ell^{(j)} = \ell^{(i)} \right\}\right|\,, \\ i_{\text{neg}} &= \left|\left\{ j \in \{1,\dots,i-1\} \mid T^{(j)} = T^{(i)} \wedge \ell^{(j)} = -\ell^{(i)} \right\}\right|\,, \end{aligned}$$

and write $i' = i_{\text{pos}} + i_{\text{neg}}$ for brevity. Let

$$\begin{aligned} h_{\text{pos}}(Y^{(i)}) &= \left|\left\{ j \in \{1,\dots,i-1\} \mid T^{(j)} = T^{(i)} \wedge \ell^{(j)} = \ell^{(i)} \wedge \mathsf{right}_s(Y^{(j)}) = \mathsf{right}_s(Y^{(i)}) \right\}\right|\,, \\ h_{\text{neg}}(Y^{(i)}) &= \left|\left\{ j \in \{1,\dots,i-1\} \mid T^{(j)} = T^{(i)} \wedge \ell^{(j)} = -\ell^{(i)} \wedge \mathsf{right}_s(X^{(j)}) = \mathsf{right}_s(Y^{(i)}) \right\}\right|\,, \end{aligned}$$

where $h_{\text{pos}}(Y^{(i)}) \leq i_{\text{pos}}$ and $h_{\text{neg}}(Y^{(i)}) \leq i_{\text{neg}}$, and write $h(Y^{(i)}) = h_{\text{pos}}(Y^{(i)}) + h_{\text{neg}}(Y^{(i)})$. We can distinct the following cases.

(1) $\ell^{(i)} = 1$ (forward query) and $Y^{(i)} \in \mathsf{rng}(\mathcal{L}_T)$. This case is excluded as $Y^{(i)}$ is not in the support of both probabilities;

(2) $\ell^{(i)} = 1$ (forward query) and $Y^{(i)} \notin \mathsf{rng}(\mathcal{L}_T)$.
   - (a) $a = 0$. We have $p_{\mathcal{O}_0,\mathcal{D}}(Y^{(i)}|\boldsymbol{\tau}^{(i-1)}) = p_{\mathcal{O}_1,\mathcal{D}}(Y^{(i)}|\boldsymbol{\tau}^{(i-1)}) = \frac{1}{2^n - i'}$, as the response is drawn uniformly at random from a set of size 2^n minus the amount of earlier queries for the same tweak;
   - (b) $a = 1$. We have $p_{\mathcal{O}_0,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)}) = \frac{1}{2^n - i'}$ as before, and $p_{\mathcal{O}_1,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)}) = \frac{1}{2^s(2^{n-s} - h(Y^{(i)}))}$ as $\mathsf{right}_s(Y^{(i)})$ is generated uniformly at random, and $\mathsf{left}_{n-s}(Y^{(i)})$ from a set of size 2^{n-s} minus $h(Y^{(i)})$. For later usage, note that this case fixes $(\ell^{(i)}, T^{(i)}, Y^{(i)})$, and we have $h_{\text{pos}}(Y^{(i)}) \leq \theta$ by $\neg\mathsf{BAD}_1$ and $h_{\text{neg}}(Y^{(i)}) \leq \theta$ by $\neg\mathsf{BAD}_2$. Therefore, in this case, we have $h(Y^{(i)}) \leq 2\theta$.

(3) $\ell^{(i)} = -1$ (inverse query) and $Y^{(i)} \in \mathsf{dom}(\mathcal{L}_T)$. The case is symmetric to (1).

(4) $\ell^{(i)} = -1$ (inverse query) and $Y^{(i)} \notin \mathsf{dom}(\mathcal{L}_T)$.
   - (a) $b = 0$. The case is symmetric to (2a).
   - (b) $b = 1$. The case is symmetric to (2b), where now we rely on the fact that by $\neg(\mathsf{BAD}_3 \vee \mathsf{BAD}_4)$, $h(Y^{(i)}) \leq 2\theta$.

Cases (2b) and (4b) dominate the chi-squared technique, and we obtain for $\chi^2(\boldsymbol{\tau}^{(i-1)})$ of (1):

$$\begin{aligned}
\chi^2(\boldsymbol{\tau}^{(i-1)}) &= \sum_{Y^{(i)}} \frac{\left(\frac{1}{2^n-i'} - \frac{1}{2^s(2^{n-s}-h(Y^{(i)}))}\right)^2}{\frac{1}{2^n-i'}} \\
&= \sum_{Y^{(i)}} (2^n - i') \cdot \left(\frac{1}{2^n-i'} - \frac{1}{2^s(2^{n-s}-h(Y^{(i)}))}\right)^2 \\
&= \sum_{Y^{(i)}} \frac{1}{(2^n-i')(2^{n-s}-h(Y^{(i)}))^2} \cdot \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 \\
&\leq \frac{1}{(2^n-i')(2^{n-s}-2\theta)^2} \cdot \sum_{Y^{(i)}} \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 \\
&\leq \frac{8}{2^{3n-2s}} \cdot \sum_{Y^{(i)}} \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 ,
\end{aligned}$$

using that $h(Y^{(i)}) \leq 2\theta$ by $\neg\mathsf{BAD}$ (see above), and $i' \leq 2^{n-1}$ and $\theta \leq 2^{n-s_{\max}-2}$. We find for its expectation:

$$\mathrm{Exp}[\chi^2(\boldsymbol{\tau}^{(i-1)})] \leq \frac{8}{2^{3n-2s}} \cdot \sum_{Y^{(i)}} \mathrm{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] . \tag{21}$$

Recalling that $i' = i_{\mathrm{pos}} + i_{\mathrm{neg}}$ and $h(Y^{(i)}) = h_{\mathrm{pos}}(Y^{(i)}) + h_{\mathrm{neg}}(Y^{(i)})$, the remaining expectation satisfies:

$$\begin{aligned}
\mathrm{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] &= \mathrm{Exp}\left[\left(h_{\mathrm{pos}}(Y^{(i)}) + h_{\mathrm{neg}}(Y^{(i)}) - \frac{i_{\mathrm{pos}} + i_{\mathrm{neg}}}{2^s}\right)^2\right] \\
&= \mathrm{Exp}\left[\left(h_{\mathrm{pos}}(Y^{(i)}) - \frac{i_{\mathrm{pos}}}{2^s}\right)^2\right] + \mathrm{Exp}\left[\left(h_{\mathrm{neg}}(Y^{(i)}) - \frac{i_{\mathrm{neg}}}{2^s}\right)^2\right] \\
&\quad + 2 \cdot \mathrm{Exp}\left[\left(h_{\mathrm{pos}}(Y^{(i)}) - \frac{i_{\mathrm{pos}}}{2^s}\right)\left(h_{\mathrm{neg}}(Y^{(i)}) - \frac{i_{\mathrm{neg}}}{2^s}\right)\right] \\
&= \mathrm{Exp}\left[\left(h_{\mathrm{pos}}(Y^{(i)}) - \frac{i_{\mathrm{pos}}}{2^s}\right)^2\right] + \left(h_{\mathrm{neg}}(Y^{(i)}) - \frac{i_{\mathrm{neg}}}{2^s}\right)^2 \\
&\quad + 2 \cdot \left(\mathrm{Exp}\left[h_{\mathrm{pos}}(Y^{(i)})\right] - \frac{i_{\mathrm{pos}}}{2^s}\right)\left(h_{\mathrm{neg}}(Y^{(i)}) - \frac{i_{\mathrm{neg}}}{2^s}\right) .
\end{aligned} \tag{22}$$

As $h_{\mathrm{pos}}(Y^{(i)}) \sim \mathrm{HG}(2^n, 2^{n-s}, i_{\mathrm{pos}})$, by Section 2.3 it satisfies

$$\begin{aligned}
\mathrm{Exp}[h_{\mathrm{pos}}(Y^{(i)})] &= \frac{i_{\mathrm{pos}}}{2^s} , \\
\mathrm{Var}[h_{\mathrm{pos}}(Y^{(i)})] &= \frac{i_{\mathrm{pos}}}{2^s} \cdot \left(1 - \frac{1}{2^s}\right) \cdot \frac{2^n - i_{\mathrm{pos}}}{2^n - 1} ,
\end{aligned}$$

and we obtain that

$$\operatorname{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] = \frac{i_{\text{pos}}}{2^s} \cdot \left(1 - \frac{1}{2^s}\right) \cdot \frac{2^n - i_{\text{pos}}}{2^n - 1} + \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2$$

$$\leq \frac{i_{\text{pos}}}{2^s} + \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 . \tag{23}$$

We furthermore claim the following.

Claim. We have $\sum_{Y^{(i)}} \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 \leq i_{\text{neg}}^2 2^{n-s}$.

Proof (of claim). We have

$$\sum_{Y^{(i)}} \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 = \sum_{Z^{(i)}} \sum_{Y^{(i)} = * \| Z^{(i)}} \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 .$$

As $h_{\text{neg}}(Y^{(i)}) = h_{\text{neg}}(Y^{(i)\prime})$ for $\mathsf{right}_s(Y^{(i)}) = \mathsf{right}_s(Y^{(i)\prime})$, we subsequently have

$$\begin{aligned}
\sum_{Y^{(i)}} \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 &= \sum_{Z^{(i)}} 2^{n-s} \left(h_{\text{neg}}(0^{n-s} \| Z^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 \\
&= 2^{n-s} \sum_{Z^{(i)}} \left(\left(h_{\text{neg}}(0^{n-s} \| Z^{(i)})\right)^2\right. \\
&\qquad \left. -2h_{\text{neg}}(0^{n-s} \| Z^{(i)}) \frac{i_{\text{neg}}}{2^s} + \left(\frac{i_{\text{neg}}}{2^s}\right)^2\right) \\
&\leq 2^{n-s} \left(i_{\text{neg}}^2 - \frac{2 i_{\text{neg}}^2}{2^s} + \frac{2^s i_{\text{neg}}^2}{2^{2s}}\right) \\
&= 2^{n-s} \left(i_{\text{neg}}^2 - \frac{i_{\text{neg}}^2}{2^s}\right) \leq i_{\text{neg}}^2 2^{n-s} . \qquad \square
\end{aligned}$$

Equations (21) and (23), alongside above claim, constitute to

$$\begin{aligned}
\operatorname{Exp}[\chi^2(\boldsymbol{\tau}^{(i-1)})] &\leq \frac{8}{2^{3n-2s}} \cdot \left(i_{\text{pos}} 2^{n-s} + i_{\text{neg}}^2 2^{n-s}\right) \\
&= \frac{8\left(i_{\text{pos}} + i_{\text{neg}}^2\right)}{2^{2n-s}} .
\end{aligned} \tag{24}$$

If $(a, b) \in \{(1,0), (0,1)\}$, we have $i_{\text{pos}}, i_{\text{neg}} \leq (i-1)$ and

$$(24) \leq \frac{8\left((i-1) + (i-1)^2\right)}{2^{2n-s}} .$$

This bound is independent of the direction of the i-th query ($\ell^{(i)} \in \{-1, 1\}$), but the parameter s depends on the query, as $s = |\mathsf{unpad}(T^{(i)})|$. The adversary maximizes its chances by sticking to the maximal s. This, finally, gives by

Lemma 1:

$$\|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| \leq \left(\frac{1}{2}\sum_{i=1}^{q} \frac{8\left((i-1)+(i-1)^2\right)}{2^{2n-s_{\max}}}\right)^{1/2}$$
$$= \left(\frac{4}{3}\frac{q^3-q}{2^{2n-s_{\max}}}\right)^{1/2} \leq \left(\frac{2q^3}{2^{2n-s_{\max}}}\right)^{1/2} .$$

On the other hand, if $(a,b) = (1,1)$, we have $i_{\text{pos}}, i_{\text{neg}} \leq \theta$ and

$$(24) \leq \frac{8\left(\theta+\theta^2\right)}{2^{2n-s}} .$$

Again sticking to the maximal s, this gives by Lemma 1:

$$\|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| \leq \left(\frac{1}{2}\sum_{i=1}^{q} \frac{8\left(\theta+\theta^2\right)}{2^{2n-s_{\max}}}\right)^{1/2} = \left(\frac{4\left(\theta+\theta^2\right)q}{2^{2n-s_{\max}}}\right)^{1/2} .$$

8 Proof of Lemma 3 on H

Consider any distinguisher $\mathcal{D}$ making at most $q \leq 2^{n-1}$ queries, all of length in $[n+s_{\min}, 2n-1]$ bits. The distinguisher has access to either random system H or $\rho \xleftarrow{\$} \text{VPerm}([n, 2n-1])$, and without loss of generality, $\mathcal{D}$ is computationally unbounded and deterministic.

We will use the chi-squared method of Section 2.2, with $\mathcal{O}_0 = H$ being the real system and $\mathcal{O}_1 = \rho$ the ideal system. Summarize the communication of $\mathcal{D}$ with its oracle in a transcript $\boldsymbol{\tau} = (\tau^{(1)}, \ldots, \tau^{(q)})$, where $\tau^{(i)} = (\ell^{(i)}, X^{(i)}, Y^{(i)})$ consists of a bit $\ell^{(i)} \in \{-1, 1\}$ indicating the direction of the query (1 means forward, and -1 means inverse), an input value $(X^{(i)})$, and a response $Y^{(i)}$, in such a way that $O^{\ell^{(i)}}(X^{(i)}) = Y^{(i)}$ for $\mathcal{O} \in \{\mathcal{O}_0, \mathcal{O}_1\}$.

Unlike the proof of Section 7, we will not rely on additional bad events, and there is no need to perform a hybrid argument and to upper bound the probability of bad events. We immediately move to bounding the term of (1), and the proof is very similar to that in Section 7.2. Consider a given transcript $\boldsymbol{\tau}^{(i-1)}$, which in turn determines the values $\ell^{(i)}$ and $X^{(i)}$. Let $s = |X^{(i)}| - n$, and consider any value $Y^{(i)}$. As both oracles behave independently for different input lengths, it suffices to focus on earlier queries of the same size. We additionally refine into the number of queries with the same or opposite query direction. Let

$$i_{\text{pos}} = \left|\left\{j \in \{1, \ldots, i-1\} \mid |X^{(j)}| = |X^{(i)}| \wedge \ell^{(j)} = \ell^{(i)}\right\}\right| ,$$
$$i_{\text{neg}} = \left|\left\{j \in \{1, \ldots, i-1\} \mid |X^{(j)}| = |X^{(i)}| \wedge \ell^{(j)} = -\ell^{(i)}\right\}\right| ,$$

and write $i' = i_{\text{pos}} + i_{\text{neg}}$. Let

$$h_{\text{pos}}(Y^{(i)}) = \left|\left\{j \in \{1, \ldots, i-1\} \mid |X^{(j)}| = |X^{(i)}| \wedge \ell^{(j)} = \ell^{(i)} \wedge \mathsf{right}_s(Y^{(j)}) = \mathsf{right}_s(Y^{(i)})\right\}\right| ,$$
$$h_{\text{neg}}(Y^{(i)}) = \left|\left\{j \in \{1, \ldots, i-1\} \mid |X^{(j)}| = |X^{(i)}| \wedge \ell^{(j)} = -\ell^{(i)} \wedge \mathsf{right}_s(X^{(j)}) = \mathsf{right}_s(Y^{(i)})\right\}\right| ,$$

where $h_{\rm pos}(Y^{(i)}) \leq i_{\rm pos}$ and $h_{\rm neg}(Y^{(i)}) \leq i_{\rm neg}$, and write $h(Y^{(i)}) = h_{\rm pos}(Y^{(i)}) + h_{\rm neg}(Y^{(i)})$. We can distinct the following cases.

(1) $\ell^{(i)} = 1$ (forward query) and $Y^{(i)} \in \mathsf{rng}(\mathcal{L}_s)$. This case is excluded as $Y^{(i)}$ is not in the support of both probabilities;

(2) $\ell^{(i)} = 1$ (forward query) and $Y^{(i)} \notin \mathsf{rng}(\mathcal{L}_s)$. We have $p_{\mathcal{O}_0,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)}) = \frac{1}{2^{n+s}-i'}$ as the response is drawn uniformly at random from a set of size 2^{n+s} minus the amount of earlier queries for the same tweak, and $p_{\mathcal{O}_1,\mathcal{D}}(Y^{(i)} \mid \boldsymbol{\tau}^{(i-1)}) = \frac{1}{2^s(2^n-h(Y^{(i)}))}$ as $\mathsf{right}_s(Y^{(i)})$ is generated uniformly at random, and $\mathsf{left}_n(Y^{(i)})$ from a set of size 2^n minus $h(Y^{(i)})$.

(3) $\ell^{(i)} = -1$ (inverse query) and $Y^{(i)} \in \mathsf{dom}(\mathcal{L}_s)$. The case is symmetric to (1).

(4) $\ell^{(i)} = -1$ (inverse query) and $Y^{(i)} \notin \mathsf{dom}(\mathcal{L}_s)$. The case is symmetric to (2).

Cases (2) and (4) dominate the chi-squared technique, and we obtain for $\chi^2(\boldsymbol{\tau}^{(i-1)})$ of (1):

$$\begin{aligned}
\chi^2(\boldsymbol{\tau}^{(i-1)}) &= \sum_{Y^{(i)}} \frac{\left(\frac{1}{2^{n+s}-i'} - \frac{1}{2^s(2^n-h(Y^{(i)}))}\right)^2}{\frac{1}{2^{n+s}-i'}} \\
&= \sum_{Y^{(i)}} \left(2^{n+s}-i'\right) \cdot \left(\frac{1}{2^{n+s}-i'} - \frac{1}{2^s(2^n-h(Y^{(i)}))}\right)^2 \\
&= \sum_{Y^{(i)}} \frac{1}{(2^{n+s}-i')(2^n-h(Y^{(i)}))^2} \cdot \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 \\
&\leq \frac{1}{(2^{n+s}-i')(2^n-i')^2} \cdot \sum_{Y^{(i)}} \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 \\
&\leq \frac{8}{2^{3n+s}} \cdot \sum_{Y^{(i)}} \left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2 ,
\end{aligned}$$

using that $h(Y^{(i)}) \leq i'$, and $i' \leq 2^{n-1}$. We find for its expectation:

$$\mathrm{Exp}[\chi^2(\boldsymbol{\tau}^{(i-1)})] \leq \frac{8}{2^{3n+s}} \cdot \sum_{Y^{(i)}} \mathrm{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] . \tag{25}$$

Recalling that $i' = i_{\rm pos}+i_{\rm neg}$ and $h(Y^{(i)}) = h_{\rm pos}(Y^{(i)})+h_{\rm neg}(Y^{(i)})$, the remaining expectation satisfies (identically to (22)):

$$\begin{aligned}
\mathrm{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] &= \mathrm{Exp}\left[\left(h_{\rm pos}(Y^{(i)}) - \frac{i_{\rm pos}}{2^s}\right)^2\right] + \left(h_{\rm neg}(Y^{(i)}) - \frac{i_{\rm neg}}{2^s}\right)^2 \\
&\quad + 2 \cdot \left(\mathrm{Exp}\left[h_{\rm pos}(Y^{(i)})\right] - \frac{i_{\rm pos}}{2^s}\right)\left(h_{\rm neg}(Y^{(i)}) - \frac{i_{\rm neg}}{2^s}\right) .
\end{aligned}$$

As $h_{\text{pos}}(Y^{(i)}) \sim \text{HG}(2^{n+s}, 2^n, i_{\text{pos}})$, by Section 2.3 it satisfies

$$\begin{aligned}
\text{Exp}[h_{\text{pos}}(Y^{(i)})] &= \frac{i_{\text{pos}}}{2^s}\,, \\
\text{Var}[h_{\text{pos}}(Y^{(i)})] &= \frac{i_{\text{pos}}}{2^s} \cdot \left(1 - \frac{1}{2^s}\right) \cdot \frac{2^{n+s} - i_{\text{pos}}}{2^{n+s} - 1}\,,
\end{aligned}$$

and we obtain that

$$\begin{aligned}
\text{Exp}\left[\left(h(Y^{(i)}) - \frac{i'}{2^s}\right)^2\right] &= \frac{i_{\text{pos}}}{2^s} \cdot \left(1 - \frac{1}{2^s}\right) \cdot \frac{2^{n+s} - i_{\text{pos}}}{2^{n+s} - 1} + \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 \\
&\leq \frac{i_{\text{pos}}}{2^s} + \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 . \qquad (26)
\end{aligned}$$

We furthermore claim the following.

Claim. We have $\sum_{Y^{(i)}} \left(h_{\text{neg}}(Y^{(i)}) - \frac{i_{\text{neg}}}{2^s}\right)^2 \leq i_{\text{neg}}^2 2^n$.

Proof (of claim). The proof is identical to that of the claim in Section 7.2, except that now $(n + s)$-bit values $Y^{(i)}$ are involved. □

Equations (25) and (26), alongside above claim, constitute to

$$\begin{aligned}
\text{Exp}[\chi^2(\boldsymbol{\tau}^{(i-1)})] &\leq \frac{8}{2^{3n+s}} \cdot \left(i_{\text{pos}} 2^n + i_{\text{neg}}^2 2^n\right) \\
&= \frac{8\left(i_{\text{pos}} + i_{\text{neg}}^2\right)}{2^{2n+s}} \\
&\leq \frac{8\left((i-1) + (i-1)^2\right)}{2^{2n+s}} .
\end{aligned}$$

This bound is independent of the direction of the i-th query ($\ell^{(i)} \in \{-1, 1\}$), but the parameter s depends on the query, as $s = |X^{(i)}| - n$. The adversary maximizes its chances by sticking to the minimal s. This, finally, gives by Lemma 1:

$$\begin{aligned}
\mathbf{Adv}_H^{\text{vsprp}}(\mathcal{D}) = \|p_{\mathcal{O}_0,\mathcal{D}}(\cdot) - p_{\mathcal{O}_1,\mathcal{D}}(\cdot)\| &\leq \left(\frac{1}{2} \sum_{i=1}^{q} \frac{8\left((i-1) + (i-1)^2\right)}{2^{2n+s_{\min}}}\right)^{1/2} \\
&= \left(\frac{4}{3} \frac{q^3 - q}{2^{2n+s_{\min}}}\right)^{1/2} \leq \left(\frac{2q^3}{2^{2n+s_{\min}}}\right)^{1/2} .
\end{aligned}$$

A Example Mixing Functions

Chen et al. [11] defined pure mixing functions as follows.

Definition 1. *Let $m, n \in \mathbb{N}$ such that $m \leq n$, and let* $\mathsf{mix} : \cup_{s=m}^{n}(\{0,1\}^s)^2 \to \cup_{s=m}^{n}(\{0,1\}^s)^2$ *be a length-preserving permutation. Define* mix_L *as the left half of its evaluation and* mix_R *as its right half. The mixing function is called* pure *if for all $s \in [m, n]$:*

– $\mathsf{mix}_L(A, \cdot)$ *is a permutation for all* $A \in \{0,1\}^s$*, and*
– $\mathsf{mix}_R(\cdot, B)$ *is a permutation for all* $B \in \{0,1\}^s$*.*

Chen et al. already pointed out that the simplest possible pure mixing function, $\mathsf{mix}(A, B) = (B, A)$, is sufficient for LDT.

Acknowledgments. This work was supported in part by the Research Council KU Leuven: GOA TENSE (C16/15/058). Yu Long Chen is supported by a Ph.D. Fellowship from the Research Foundation - Flanders (FWO). Bart Mennink is supported by a postdoctoral fellowship from the Netherlands Organisation for Scientific Research (NWO) under Veni grant 016.Veni.173.017. Mridul Nandi is supported by the Wisekey Project in the R.C.Bose Centre of Cryptology and Security. The authors would like to thank the anonymous reviewers for their comments and suggestions.

References

1. Andreeva, E., Bogdanov, A., Datta, N., Luykx, A., Mennink, B., Nandi, M., Tischhauser, E., Yasuda, K.: COLM v1 (2016), submission to CAESAR competition
2. Andreeva, E., Bogdanov, A., Luykx, A., Mennink, B., Tischhauser, E., Yasuda, K.: Parallelizable and Authenticated Online Ciphers. In: Sako and Sarkar [42], pp. 424–443
3. Andreeva, E., Bogdanov, A., Luykx, A., Mennink, B., Tischhauser, E., Yasuda, K.: AES-COPA v.1 (2015), submission to CAESAR competition
4. Avanzi, R.: The QARMA block cipher family. almost MDS matrices over rings with zero divisors, nearly symmetric even-mansour constructions with non-involutory central rounds, and search heuristics for low-latency s-boxes. IACR Trans. Symmetric Cryptol. 2017(1), 4–44 (2017), https://doi.org/10.13154/tosc.v2017.i1.4-44
5. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY Family of Block Ciphers and Its Low-Latency Variant MANTIS. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016, Part II. LNCS, vol. 9815, pp. 123–153. Springer (2016)
6. Bellare, M., Rogaway, P., Spies, T.: The FFX Mode of Operation for Format-Preserving Encryption (2010), submission to NIST
7. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Sponge Functions. ECRYPT Hash Workshop 2007 (May 2007)
8. Brier, E., Peyrin, T., Stern, J.: BPS: A Format-Preserving Encryption Proposal (2010), submission to NIST
9. Canteaut, A., Viswanathan, K. (eds.): INDOCRYPT 2004, LNCS, vol. 3348. Springer (2004)
10. Chakraborty, D., Sarkar, P.: HCH: A New Tweakable Enciphering Scheme Using the Hash-Encrypt-Hash Approach. In: Barua, R., Lange, T. (eds.) INDOCRYPT 2006. LNCS, vol. 4329, pp. 287–302. Springer (2006)
11. Chen, Y., Luykx, A., Mennink, B., Preneel, B.: Efficient Length Doubling From Tweakable Block Ciphers. IACR Trans. Symmetric Cryptol. 2017(3), 253–270 (2017)

12. Coron, J., Dodis, Y., Mandal, A., Seurin, Y.: A Domain Extender for the Ideal Cipher. In: Micciancio, D. (ed.) TCC 2010. LNCS, vol. 5978, pp. 273–289. Springer (2010)
13. Daemen, J.: Hash Function and Cipher Design: Strategies Based on Linear and Differential Cryptanalysis. Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium (1995)
14. Dai, W., Hoang, V.T., Tessaro, S.: Information-Theoretic Indistinguishability via the Chi-Squared Method. In: Katz and Shacham [25], pp. 497–523
15. Dworkin, M.: NIST SP 800-38E: Recommendation for Block Cipher Modes of Operation: The XTS-AES Mode for Confidentiality on Storage Devices (2010)
16. Dworkin, M.: NIST SP 800-38G: Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption (2016)
17. Gilboa, S., Gueron, S.: Distinguishing a truncated random permutation from a random function. Cryptology ePrint Archive, Report 2015/773 (2015)
18. GS1: EPC$^{\mathrm{TM}}$ Radio-Frequency Identity Protocols Generation-2 UHF RFID, Version 2.0.1 (2015), https://www.gs1.org/sites/default/files/docs/epc/Gen2_Protocol_Standard.pdf
19. GS1: EPC Tag Data Standard, Version 1.11 (2017), https://www.gs1.org/sites/default/files/docs/epc/GS1_EPC_TDS_i1_11.pdf
20. Halevi, S.: EME*: Extending EME to Handle Arbitrary-Length Messages with Associated Data. In: Canteaut and Viswanathan [9], pp. 315–327
21. Halevi, S.: Invertible Universal Hashing and the TET Encryption Mode. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 412–429. Springer (2007)
22. Hall, C., Wagner, D.A., Kelsey, J., Schneier, B.: Building PRFs from PRPs. In: Krawczyk, H. (ed.) CRYPTO '98. LNCS, vol. 1462, pp. 370–389. Springer (1998)
23. Iwata, T., Minematsu, K., Peyrin, T., Seurin, Y.: ZMAC: A Fast Tweakable Block Cipher Mode for Highly Secure Message Authentication. In: Katz and Shacham [25], pp. 34–65
24. Jean, J., Nikolic, I., Peyrin, T.: Tweaks and Keys for Block Ciphers: The TWEAKEY Framework. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014, Part II. LNCS, vol. 8874, pp. 274–288. Springer (2014)
25. Katz, J., Shacham, H. (eds.): CRYPTO 2017, Part III, LNCS, vol. 10403. Springer (2017)
26. Krovetz, T., Rogaway, P.: The Software Performance of Authenticated-Encryption Modes. In: Joux, A. (ed.) FSE 2011. LNCS, vol. 6733, pp. 306–327. Springer (2011)
27. McGrew, D.A., Fluhrer, S.R.: The Security of the Extended Codebook (XCB) Mode of Operation. In: Adams, C.M., Miri, A., Wiener, M.J. (eds.) SAC 2007. LNCS, vol. 4876, pp. 311–327. Springer (2007)
28. McGrew, D.A., Viega, J.: The Security and Performance of the Galois/Counter Mode (GCM) of Operation. In: Canteaut and Viswanathan [9], pp. 343–355
29. Minematsu, K.: Beyond-Birthday-Bound Security Based on Tweakable Block Cipher. In: Dunkelman, O. (ed.) FSE 2009. LNCS, vol. 5665, pp. 308–326. Springer (2009)
30. Minematsu, K., Iwata, T.: Building Blockcipher from Tweakable Blockcipher: Extending FSE 2009 Proposal. In: Chen, L. (ed.) IMACC 2011. LNCS, vol. 7089, pp. 391–412. Springer (2011)
31. Nandi, M.: A Generic Method to Extend Message Space of a Strong Pseudorandom Permutation. Computación y Sistemas 12(3) (2009)
32. Nandi, M.: XLS is Not a Strong Pseudorandom Permutation. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014, Part I. LNCS, vol. 8873, pp. 478–490. Springer (2014)

33. Nandi, M.: On the optimality of non-linear computations of length-preserving encryption schemes. In: Iwata, T., Cheon, J.H. (eds.) ASIACRYPT 2015, Part II. LNCS, vol. 9453, pp. 113–133. Springer (2015)
34. Nandi, M.: Revisiting security claims of XLS and COPA. Cryptology ePrint Archive, Report 2015/444 (2015)
35. Peyrin, T., Seurin, Y.: Counter-in-Tweak: Authenticated Encryption Modes for Tweakable Block Ciphers. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 33–63. Springer (2016)
36. Qualcomm: Pointer Authentication on ARMv8.3 — Design and Analysis of the New Software Security Instructions (2017), https://www.qualcomm.com/media/documents/files/whitepaper-pointer-authentication-on-armv8-3.pdf
37. Ristenpart, T., Rogaway, P.: How to Enrich the Message Space of a Cipher. In: Biryukov, A. (ed.) FSE 2007. LNCS, vol. 4593, pp. 101–118. Springer (2007)
38. Rogaway, P.: Efficient Instantiations of Tweakable Blockciphers and Refinements to Modes OCB and PMAC. In: Lee, P.J. (ed.) ASIACRYPT 2004. LNCS, vol. 3329, pp. 16–31. Springer (2004)
39. Rogaway, P., Bellare, M., Black, J., Krovetz, T.: OCB: a block-cipher mode of operation for efficient authenticated encryption. In: Reiter, M.K., Samarati, P. (eds.) ACM CCS 2001. pp. 196–205. ACM (2001)
40. Rogaway, P., Wooding, M., Zhang, H.: The Security of Ciphertext Stealing. In: Canteaut, A. (ed.) FSE 2012. LNCS, vol. 7549, pp. 180–195. Springer (2012)
41. Rogaway, P., Zhang, H.: Online Ciphers from Tweakable Blockciphers. In: Kiayias, A. (ed.) CT-RSA 2011. LNCS, vol. 6558, pp. 237–249. Springer (2011)
42. Sako, K., Sarkar, P. (eds.): ASIACRYPT 2013, Part I, LNCS, vol. 8269. Springer (2013)
43. Sarkar, P.: Improving Upon the TET Mode of Operation. In: Nam, K., Rhee, G. (eds.) ICISC 2007. LNCS, vol. 4817, pp. 180–192. Springer (2007)
44. Schnorr, C., Vaudenay, S.: Parallel FFT-Hashing. In: Anderson, R.J. (ed.) FSE 1993. LNCS, vol. 809, pp. 149–156. Springer (1993)
45. Shrimpton, T., Terashima, R.S.: A Modular Framework for Building Variable-Input-Length Tweakable Ciphers. In: Sako and Sarkar [42], pp. 405–423
46. Wang, P., Feng, D., Wu, W.: HCTR: A Variable-Input-Length Enciphering Mode. In: Feng, D., Lin, D., Yung, M. (eds.) CISC 2005. LNCS, vol. 3822, pp. 175–188. Springer (2005)
47. William F. Ehrsam, Carl H. W. Meyer, John L. Smith, Walter L. Tuchman: Message verification and transmission error detection by block chaining (1976), US Patent 4074066
48. Zhang, H.: Length-Doubling Ciphers and Tweakable Ciphers. In: Bao, F., Samarati, P., Zhou, J. (eds.) ACNS 2012. LNCS, vol. 7341, pp. 100–116. Springer (2012)