Forking Sums of Permutations for Optimally Secure and Highly Efficient PRFs

Avijit Dutta[1], Jian Guo[2] and Eik List[2]

[1] Institute for Advancing Intelligence, TCG-CREST, India
avirocks.dutta13(at)gmail.com
[2] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
{guojian, eik.list}(at)ntu.edu.sg

Abstract. The desirable encryption scheme possesses high PRF security, high efficiency, and the ability to produce variable-length outputs. Since designing dedicated secure PRFs is difficult, a series of works was devoted to building optimally secure PRFs from the sum of independent permutations (SoP), Encrypted Davies-Meyer (EDM), its Dual (EDMD), and the Summation-Truncation Hybrid (STH) for variable output lengths, which can be easily instantiated from existing permutations. For increased efficiency, reducing the number of operations in established primitives has been gaining traction: Mennink and Neves pruned EDMD to FastPRF, and Andreeva et al. introduced ForkCiphers, which take an n-bit input, process it through a reduced-round permutation, fork it into two states, and feed each of them into another reduced-round permutation to produce a $2n$-bit output. The constructions above can be used in secure variable-length modes or generalizations such as MultiForkCiphers.

In this paper, we suggest a framework of those constructions in terms of the three desiderata: we span the spectrum of (1) output length vs. PRF security, (2) full vs. round-reduced primitives, and (3) fixed- vs. variable-length outputs. From this point of view, we identify remaining gaps in the spectrum and fill them with the proposal of several highly secure and efficient fixed- and variable-output-length PRFs.

We fork SoP and STH to ForkPRF and ForkSTH, extend STH to the variable-output-length construction STHCENC, which bridges the gap between CTR mode and CENC,and propose ForkCENC, ForkSTHCENC, ForkEDMD, as well as ForkEDM-CTR as the variable-output-length and round-reduced versions of CENC, STH, FastPRF, and FastPRF's dual, respectively.

Using recent results on Patarin's general Mirror Theory, we have proven that almost all our proposed PRFs are optimally secure under the assumption that the permutations are pairwise independent and random and STH achieves the optimal security depending on the output length. Our constructions can be highly efficient in practice. We propose efficient instantiations from round-reduced AES and back it with the cryptanalysis lessons learned from existing earlier analysis of AES-based primitives.

Keywords: Provable security · H-coefficient technique · sum of permutations · Encrypted Davies-Meyer · Summation and Truncation Hybrid · AES · Forkcipher

1 Introduction

Pseudorandom functions (PRFs) are important cryptographic primitives used in various cryptographic algorithms for encryption and authentication. However, designing dedicated pseudorandom functions is hard as one of the biggest challenges in designing PRFs is to design a secure *non-invertible* length-preserving round function that can be iterated

multiple times to produce a secure PRF, as collision probabilities are amplified with each iteration [20, 78].

On the other hand, designing a pseudorandom permutation (PRP) (a.k.a. a keyed block cipher) is relatively easy. The general strategy to design a block cipher is to design a weak *invertible* round function and to iterate it multiple times to produce a secure pseudorandom permutation. There are primarily two paradigms for designing a block cipher [69], namely the Luby-Rackoff paradigm [51, 74] and substitution-permutation networks (SPN) [51, 88]. Informally, we call a collection of functions a PRF if, for any computationally bounded adversary, it is infeasible to distinguish the collection from a collection of random functions. Similarly, we call a collection of permutations a PRP if, for any computationally bounded adversary, it is infeasible to distinguish the collection from a collection of random permutations.

While myriads of block ciphers are available and thus many more than there are practical candidates of pseudorandom functions, the usual security notion for encryption and authentication modes of operation searches for the PRF rather than the PRP security. Due to the PRP-PRF switching lemma [15, 17, 32], a block cipher is considered to be a good PRF up to the birthday bound, i.e., if the block size of a block cipher E_k is n-bits, then it is a secure PRF when the number of messages it processes is at most $2^{n/2}$. As a result, one can instantiate the mode of operation with block ciphers which makes the resultant bound of the construction up to the birthday limit. While the birthday bound is acceptable for practical values of n such as 128 bits, the bound may not be useful for small values of n, such as $n = 64$ bits. However, due to the ongoing trend of lightweight cryptographic primitives, small-state block ciphers (e.g., PRINCE [23], PRESENT [22] etc.) are frequently used in practice. Therefore, it is of utmost importance to disregard the idea of considering a block cipher to be a secure PRF and focus on designing modes of operations that are provably secure beyond the birthday bound (BBB) when instantiated with small-state block ciphers. Hereafter, we use the terms permutation and block cipher interchangeably. Before we can consider this long goal of designing highly secure and efficient variable-output-length (VOL-)PRFs, we have to briefly recall and understand the existing fixed-output-length (FOL-)PRFs first.

1.1 Designing PRFs with Beyond-birthday-bound Security

Designing PRFs with beyond-birthday security started from the proposal of Hall et al. [60], who proposed to truncate the output from an n-bit permutation to s bits. This construction was later proven secure for up to $2^{n-s/2}$ queries [14, 52], i.e., $n - s/2$-bit security. Bellare et al. [16] have proposed the Sum of Permutations (SoP) which returns the XOR of the outputs of two n-bit independent permutations Π_1, Π_2:

$$\mathsf{SoP}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_1(x) \oplus \Pi_2(x).$$

This construction was proven secure first for up to $2^{2n/3}$ queries [75] and recently for up to 2^n queries [41, 49]. Guo et al. [59] proposed SUMPIP, a contender of SoP:

$$\mathsf{SUMPIP}_{\Pi}(x) \triangleq \Pi(x) \oplus \Pi^{-1}(x)\,.$$

In contrast to the single-permutation variant of SoP which takes $(n-1)$-bit inputs, SUMPIP is the first single-permutation-based PRF that takes and returns n-bit values. In the same paper, the authors also showed that a single-permutation-variant of EDM and EDMD achieves $O(2n/3)$-bit security. Compared to just returning both outputs from $\Pi_1(x)$ and $\Pi_2(x)$ (call it PRP2), the sum of permutation is a trade-off: On the one hand, it reduces the efficiency of two permutation calls by generating only an n-bit output. At CRYPTO'20, Gunsing and Mennink [58] proposed the *Summation-Truncation Hybrid* (STH) that filled

the range between those extremes. STH outputs an a-bit part of each permutation call and sums the $n-a$ outputs of both permutations. More precisely, STH takes an $(n-1)$-bit input x, truncates the leftmost s bits of $\Pi(x\|0)$ and $\Pi(x\|1)$, and sums the discarded $n-s$ bits of $\Pi(x\|0)$ and $\Pi(x\|1)$ to produce an $(n+s)$-bit output. They showed that STH provides security roughly up to $2^{n-s/2}$ queries. This trade-off is shown in the top of the leftmost column of Figure 1.1.

While the constructions above are parallelizable PRFs, Cogliati and Seurin [37] initiated the research direction on alternative candidates of beyond-birthday-bound secure PRFs that use a sequential execution of permutations. They proposed the *Encrypted Davis Meyer* (EDM) construction and have shown that EDM achieves $2n/3$-bit security. EDM orders the permutation calls to Π_1 and Π_2 in sequence and XORs the input to Π_1 to the output of Π_1 before the XOR sum is processed by Π_2:

$$\mathsf{EDM}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_2(\Pi_1(x) \oplus x)\,.$$

Later in [77], Mennink and Neves showed the optimal security of the construction. In the same paper, the authors also proposed a dual variant of EDM called EDMD.

$$\mathsf{EDMD}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_2(\Pi_1(x)) \oplus \Pi_1(x)\,,$$

and showed its optimal PRF security. However, the proofs for the optimal bound of both EDM and EDMD are inherently based on a debated result of Mirror Theory for general $\xi_{\max}$ [79, 84].[1] While both (i.e., EDM and EDMD) are based on two independent n-bit permutations, in [38], Cogliati and Seurin have shown $2n/3$-bit security for EDM with a single permutation. Concurrent to this work, Guo et al. [59] have also shown $2n/3$-bit PRF security for the single-permutation-based EDM and the EDMD constructions. While those PRFs are useful when small size fixed in- and output lengths are needed, encryption requires PRFs with variable output lengths in general. Recently, Chen et al. [34] have shown that to design an n-to-n-bit PRF from the XOR of permutations with optimal security, one needs to resort to the constructions whose structure is inherently based on SoP, EDM, or EDMD. Thus, it makes perfect sense to consider possible extensions of these constructions for VOL PRFs.

1.2 From Fixed- to Variable-output-length PRFs

A simple instantiation of PRF-based modes is to substitute each block-cipher call with SoP or STH. For example, Iwata and Minematsu [64] replaced every block-cipher call for encryption with a sum of permutations in GCM-SIV-2 and its generalization GCM-SIV-r. Iwata [62] extended the SoP construction for a variable-output-length PRF, called XORP, that takes an m-bit input and produces a sequence of w n-bit outputs as:

$$\mathsf{XORP}(x) \triangleq \|_{i=1}^{w} \Pi(x\|\langle 0\rangle_s) \oplus \Pi(x\|\langle j\rangle_s)\,,$$

where $s = \lceil \log_2(w+1) \rceil$, $\langle j\rangle_s$ denotes the s bit binary represenation of the integer j and $m+s=n$. In [63], Iwata et al. proved optimal security of XORP and a nonce-based encryption mode CENC [62] around it.

From EDM, Menning and Neves [78] defined a counter-mode PRF with optimal security. Choi et al. derived per-message masks with a four-block chunk of CENC and encrypts in an OCB-like manner in their proposal of the SCM AE scheme [35]. If the message length is limited to $\ll 2^{n/2}$ blocks, their scheme can provide up to optimal security. However, the message lengths might be larger than that. In this work, we will strive for highly secure PRFs whose security does not limit the message length to the birthday bound.

[1] Dutta et al. [39] have given a correct and verifiable proof of the Mirror-Theory result for general $\xi_{\max}$.

For permutation-based constructions, Farfalle [18] and Farasha [1] are variable-length PRFs, although, with their security limited to the birthday bound of the primitive. In a preprint of their Megafono and Hydra PRF constructions, Grassi et al. [55] employed a variable-output-length variant of STH that summed the b-bit part of two consecutive permutations each. Given w permutation calls for even w, this approach produced $wn - (w/2)b$ bits of output. In comparison, STHCENC outputs $wn - b$ bits with almost the same PRF security (reduced by a logarithmic factor in w). In [56], they replaced the STH construction by a "feed-forward operation both in order to avoid wasting encryption material and in order to increase the security with respect to guessing attacks".

1.3 Round-reduced Primitives

While CENC and its generalization NEMO [72] constitute highly efficient modes with close-to-optimal provably security, practical efficiency demands fostered additional research. Motivated by the increased depth of understanding of established primitives (such as the AES or Skinny), there is an ongoing trend to make schemes more efficient by using round-reduced variants of them. For AEZ, Hoang et al. [61] proposed the prove-then-prune approach, where a scheme is proven secure under assumptions on its primitive, and the construction is instantiated with a downscaled primitive. For some variants of AEZ, later cryptanalysis [26, 89] showed that the assumptions are violated with the used primitives.

ForkCipher. While one can simply reduce the number of rounds in AES instances inside to eight or nine rounds (cf. [6]), this approach still needs 16 or 18 AES rounds per block. Andreeva et al. [4] suggested a new kind of primitive that they coined ForkCipher with the goal of higher efficiency for encryption or authentication. A ForkCipher maps an n-bit message to a $2n$-bit output. Instead of applying a single (full-size) permutation to each block, it uses three permutations that can be more efficient in sum. More precisely, a ForkCipher processes the input with a first permutation, before the resulting state is forked (used simultaneously as input to both) to two further independent permutations at the bottom whose outputs are returned. To instantiate a ForkCipher from a primitive's round function, the authors of [4] proposed the iterate-fork-iterate (IFI) paradigm for some r_1 and r_2, where a plaintext is encrypted through an r_1-round of the cipher at the top and further processed through key-independent r_2 rounds of the cipher in the bottom permutations. Both outputs can be used, e.g. as a ciphertext block and its authentication tag for small messages of at most a single block, or as a ciphertext block and a chaining value in modes, respectively. Andreeva et al.'s [4] security notion of MultiForkCiphers (MFCs) represented a natural extension of ForkCiphers. Similarly, the authors proposed their instantiation from the extension of the IFI principle to iterate-fork-iterate-many (IFIM). IFIM uses the forked state for more than two and potentially many more independent permutation calls. As such, it can represent a more efficient variant of counter mode.

FastPRF. Menning and Neves [78] applied the prove-then-prune paradigm to EDM and EDMD and the AES. Their proposal FastPRF was an n-bit-secure encryption scheme that used a variant of EDM with round-reduced permutations in counter mode. While FastPRF features a generic proof, it can also be analyzed with respect to the standard PRF notion. As a concrete instance, they proposed AES-PRF, which is the AES-128 reduced to five rounds in each permutation call. Thus, their instantiated encryption scheme had also only 10 AES rounds per encrypted message block, which inspired further research on pushing the AES and the understanding of attacks on its round-reduced variants further.

1.4 Filling the Gaps Towards Secure Highly Efficient VOL-PRFs

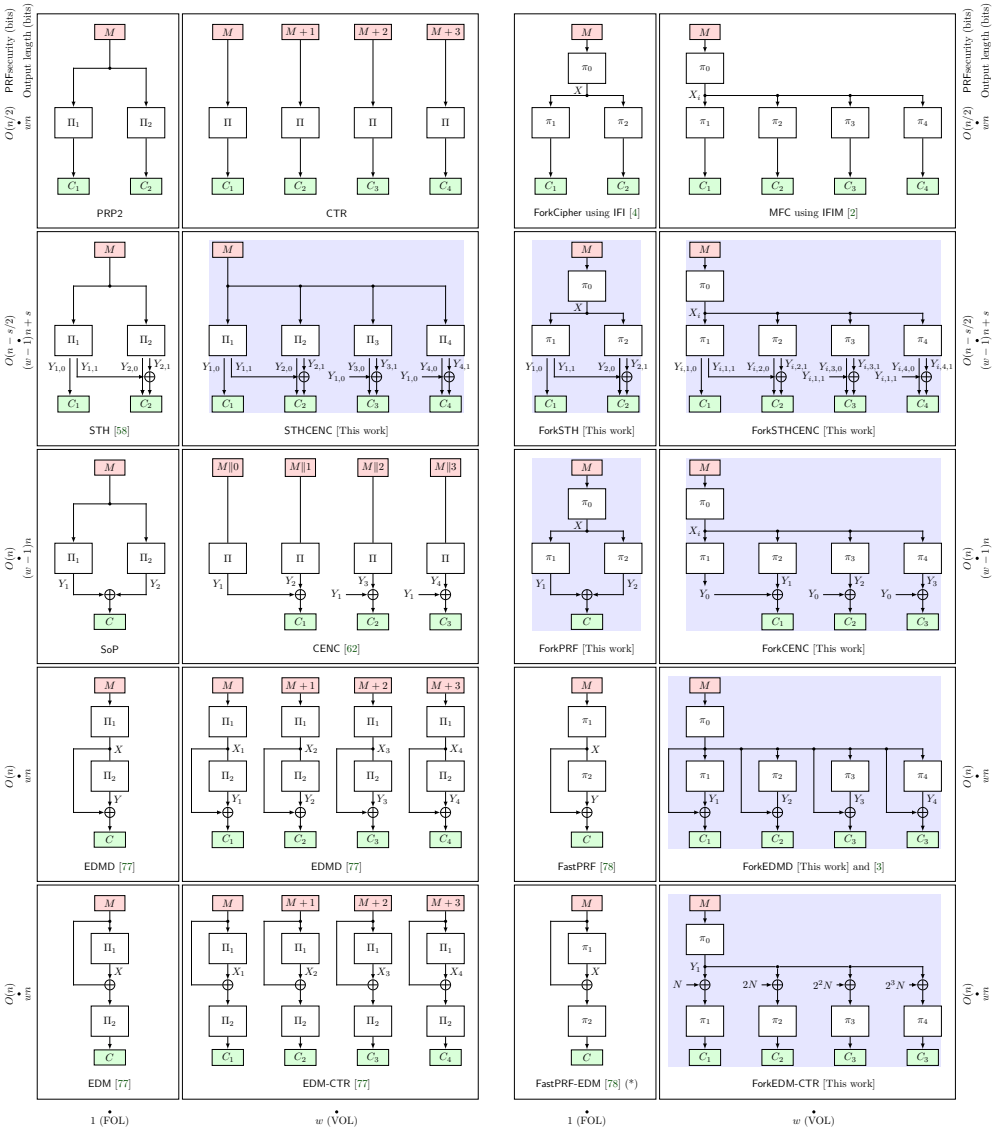
We ask if one could reduce the number of rounds of a cipher and prove a similar level of security in a design of higher efficiency. The ForkCipher tries to increase efficiency compared to two calls to a PRP, and the MultiForkCipher generalizes it to variable-output-length constructions. Though, when limiting the number of independent permutations and not imposing restrictions on the message length, they offer $O(n/2)$-bit PRFsecurity since collisions are prohibited between $2^{n/2}$ short messages.
Thus, (1) the number of primitive calls per output block, (2) the number of output blocks, and (3) the blockwise efficiency from forking span a spectrum that is visualized in Figure 1.1. After locating the 20 existing constructions therein, we identify seven gaps. One can easily observe that the round-reducing approach of ForkCiphers can also be applied to SoP and STH to obtain more efficient variants ForkPRF and ForkSTH. Still, those constructions are an intermediate step given their inferior efficiency compared to FastPRF and its dual. In the second step, we observe that we can derive VOL extensions. The equivalent of round-reduced counter mode is the MFC, i.e., the VOL extension ForkCipher.
This work proposes the additional VOL extensions ForkCENC, ForkSTHCENC, ForkEDMD, and ForkEDM-CTR for ForkPRF, ForkSTH, FastPRF, and the dual of FastPRF, respectively. Our constructions fork the output of a top permutation call in the middle for many bottom-permutation calls and differ in their outputs. As a result, this work proposes a framework of close-to-optimally secure and efficient VOL-PRFs built on reduced-round block ciphers. Assuming r, r_t, and r_b rounds in a full, the top, and the bottom permutations, respectively, ForkCENC can encrypt at a rate approaching $r/r_b > 1$.We propose an instantiation based on the AES that applies the knowledge from existing attacks and countermeasures. Figure 1.1 provides an overview of existing constructions and our proposals (the latter highlighted).

Instantiation. In the line of AES-PRF and Fork-AES, we propose an instantiation of ForkCENC, called ForkCENC-AES, where we instantiate the permutations with reduced-round tweaked AES that we believe to be more secure than AES-PRF or ForkAES. We adopt the ElasticTweak approach from ESTATE [28, 29] to separate the individual permutations, but introduce further refinements to increase the diffusion of the tweaks. As a consequence, ForkCENC is more efficient than CENC and AES-PRF while it provides the security of CENC of n bits of security, and benefits from the corpus of existing cryptanalysis.

Outline. The remainder of this work is structured as follows. Section 2 lists the notations and defines the security notions for the rest of the paper. Section 3 provides formal definitions for all of our proposed constructions. We provide a formal security argument for all of the proposed constructions in section 4. Section 6 defines an instantiation of our proposed scheme which is analyzed in depth in Section 7. Section 8 describes software-implementation results.

Comparison to [3]. In a parallel work [3], Andreeva et al. proposed a variant of ForkEDMD [4]. Theirs and our work share the idea that ForkEDMD is an excellent candidate for reduced-round encryption. Further similarly to our work, they proposed a highly efficient instantiation of it, called ButterKnife, from a round-reduced tweakable variant of an AES-round-based TBC. More precisely, they proposed a highly efficient instantiation 7+8-round Deoxys-BC with 256-bit tweakey before and after the forking point, respectively. They further demonstrate its efficiency and usefulness in highly secure deterministic authenticated encryption schemes.

Our work differs from [3] in several points: most notably, our work considers the spectrum of schemes and not only one extreme. The ForkCipher series of coins new paradigms for the individual variants: Iterate-Fork-Iterate [4], Iterate-MultiFork-Iterate [4], and Masked-Iterate-Fork-Iterate (mIFI) [3]. In contrast, we map them to fixed- or



(a) Full-round primitives. (b) Round-reduced primitives.

Figure 1.1: The spectrum of encryption schemes from PRP2 to SoP and their variable-output-length and round-reduced constructions. The schemes in the colored boxes are introduced in this work. (*) FastPRF-EDM was not proposed but considered in [78].

Table 1: Comparison of the considered constructions. $\widetilde{\pi}^t$ and $\widetilde{\pi}^T$ = reduced tweakable permutation;$\widetilde{\pi}^t$ uses small tweaks that can cover only few domains, whereas $\widetilde{\pi}^T$ allows nonce/IV to be used in the tweak. Security is given for n-bit tweaks T. (*) The dual variant of FastPRF was considered but not proposed by [78].

	Calls/block		Output	PRF Sec.	
Construction	$\widetilde{\pi}^t$	$\widetilde{\pi}^T$	(bits)	(bits)	Reference
FIL, FOL, reduced permutations					
ForkCipher	3	–	$2n$	$n/2$	[4]
ForkSTH	3	–	$n+s$	$n-s/2$	[**This work**]
ForkPRF	3	–	n	n	[**This work**]
FastPRF	2	–	n	n	[78]
FastPRF-EDM	2	–	n	n	[78] (*)
FIL, VOL, reduced permutations					
MFC$[w]$	$w+1$	–	wn	$n/2$	[2]
$\widetilde{\text{MFC}}[w]$	–	$w+1$	wn	n	[2]
ForkSTHCENC$[w]$	$w+1$	–	$(w-1)n+s$	$n-s/2$	[**This work**]
ForkCENC$[w]$	$w+1$	–	$(w-1)n$	n	[**This work**]
ButterKnife	–	$w+1$	wn	n	[3]
mIFI, ForkEDMD$[w]$	$w+1$	–	wn	n	[3], [**This work**]
ForkEDM-CTR$[w]$	$w+1$	–	wn	n	[**This work**]

variable-output-length PRFs. Viewing constructions as part of a spectrum of FOL- VOL-PRFs allows us to identify the trade-off between security and output length from ForkCENC and MultiForkCipher and intermediate constructions with variable output lengths, covered by ForkSTHCENC. Extending the round-reduced variants of EDMD and EDM-CTR, we can further identify ForkEDMD and ForkEDM-CTR as the most efficient VOL-PRFs with respect to the number of calls to the internal primitives. Then, given the full classification of two-permutation-based constructions from [34], we can finally identify ForkCENC, ForkEDMD, and ForkEDM-CTR to be the set of all n-bit secure VOL-PRFs variants of the spectrum.

Our instantiation allows us to address a slightly different use case than does ButterKnife. This is a property of the instantiation in mind and not of the construction. ForkEDMD and also mIFI can consider a variant of ElasticTweak as a tweak schedule for very small tweaks explicitly. Small tweaks suffice for separating the individual permutation calls and could support more efficient tweak scheduling. Our motivation was to offer an advantage over a tweakable MultiForkCipher [2]. The latter can be used as a variant of forked Counter-in-Tweak [87] (as proposed e.g. as Variant 3 of [2]) and therefore already allowed to build an n-bit-secure VOL-PRF. Though, it needs n-bit tweaks for n-bit security.

ButterKnife targets 256-bit tweakeys [3]. Thus, it would be interesting to compare ButterKnife – which adds the final XOR to each block – to a tweakable MultiForkCipher that uses CTRT mode also with a similar primitive as ButterKnife. We are aware that practical x64 platforms such as 6-th or 12-th generation Intel i5 processors seem able to execute the operations for the tweakey schedule in Deoxys-BC in parallel to the AES rounds. Nevertheless, platforms vary over time and so may their efficiency.

2 Preliminaries

General Notation. We use uppercase characters for functions and variables, lowercase characters for indices and lengths, and calligraphic uppercase characters for sets and distributions. We indicate lists, vectors, and matrices, but also distinguishers and adversaries in general by boldface characters. For non-negative integers x and y, we write $[x] = \{1, \ldots, x\}$, $[0..x] = \{0, 1, \ldots, x\}$, and $[x..y] = \{x, x+1, \ldots, y\}$. We write

$\{0,1\}^n$ for n-bit strings, and $X\|Y$ for the concatenation of two bitstrings X and Y. For integers x, n and bitstring $X \in \{0,1\}^n$, we use $X_1, \ldots, X_m \xleftarrow{x} X$ for the unique splitting of X into segments of $\leq x$ bit length, such that $|X_1| = \cdots = |X_{m-1}| = x$ and $|X_m| \leq x$. Similarly, we use $(X_1, X_2) \xleftarrow{x,n-x} X$ to indicate that $|X_1| = x$, $|X_2| = n - x$ and $X_1\|X_2 = X$. We write $X_1, X_2, \ldots \twoheadleftarrow \mathcal{X}$ for the uniform and pairwise independent sampling with replacement of $X_1, X_2, \ldots$ from $\mathcal{X}$. Thus, $X_i \twoheadleftarrow \mathcal{X}$, independent of the values X_j for $i \neq j$. We use $X_1, X_2, \ldots \twoheadleftarrow_{\mathsf{wor}} \mathcal{X}$ for the uniform and pairwise independent sampling with replacement in the order of $X_1 \twoheadleftarrow \mathcal{X}$, then $X_2 \twoheadleftarrow \mathcal{X} \setminus \{X_1\}$, and so on. In general, $X_i \twoheadleftarrow \mathcal{X} \setminus \{X_1, X_2, \ldots, X_{i-1}\}$. For non-empty set or spaces $\mathcal{T}$, $\mathcal{X}$, and $\mathcal{Y}$, we write $\mathsf{Perm}(\mathcal{X})$ for the set of permutations over $\mathcal{X}$ and $\widetilde{\mathsf{Perm}}(\mathcal{T}, \mathcal{X})$ for the set of tweakable permutations over $\mathcal{X}$ with tweak space $\mathcal{T}$.

Distinguishers. An adversary is an algorithm that interacts with a challenger. A distinguisher $\mathbf{D}$ is an adversary that interacts with one of several worlds that it shall distinguish between. Prior to the interaction, the challenger samples a random bit $b \leftarrow_{\$} \{0,1\}^n$ and presents $\mathbf{D}$ with one of two sets of oracles depending on the value of b. Moreover, the challenger uses internal secrets such as keys. $\mathbf{D}$ can interact with the individual oracles and collect the corresponding responses. At the end, $\mathbf{D}$ outputs a guess b' to the challenger; $\mathbf{D}$ wins if and only if (iff) $b = b'$. We write

$$\Delta_{\mathbf{D}}(\mathcal{R}_K; \mathcal{I}) = \mathbf{Adv}_{\mathcal{R}_K}(\mathbf{D}) \triangleq \left| \Pr_K \left[\mathbf{D}^{\mathcal{R}_K} = 1\right] - \Pr\left[\mathbf{D}^{\mathcal{I}} = 1\right] \right|$$

for the advantage of $\mathbf{D}$ in distinguishing a real keyed construction $\mathcal{R}_K$ from an ideal construction $\mathcal{I}$, where the probability is over the key K, the randomness of $\mathcal{I}$, the coins of $\mathbf{D}$ and that of the challenger, if any. We use the convention of $b = 1$ for the real world. W.l.o.g., we consider deterministic distinguishers and consider information-theoretic advantages that are restricted under the assumption that all queries to the construction and primitives are made through limited numbers of oracle queries. For two sets of oracles $\mathcal{I}$ and $\mathcal{R}_K$, where $\mathcal{I}$ represents an ideal and $\mathcal{R}_K$ a real world (usually a keyed construction), we write the distinguishing advantage of $\mathbf{D}$ as $\Delta_{\mathbf{D}}(\mathcal{R}_K; \mathcal{I})$.

PRF Security. PRF security refers to the maximal advantage of distinguishing the outputs of a scheme from random bits of the expected length. For primitives and schemes in general, we will often use the set $\mathcal{K} = \mathbb{F}_2^n$ for keys, $\mathcal{B} = \mathbb{F}_2^n$ for message blocks, $\mathcal{N} = \mathbb{F}_2^\nu$ for nonces, and $\mathbf{D} = \mathbb{F}_2^d$ for counters, where n, ν, d are small integers. Given two non-empty sets or spaces $\mathcal{X}, \mathcal{Y}$, let $F : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a function, $\rho \leftarrow_{\$} \mathsf{Func}(\mathcal{X}, \mathcal{Y})$ and $K \leftarrow_{\$} \mathcal{K}$ be a secret key. Then, the PRF advantage of $\mathbf{D}$ is defined as

$$\mathbf{Adv}^{\mathsf{PRF}}_{F_K}(\mathbf{D}) \triangleq \Delta_{\mathbf{D}}(F_K; \rho).$$

Nonce-based Encryption. A nonce-based encryption scheme $\mathcal{E} = (\mathcal{E}, \mathbf{D})$ is a tuple of algorithms for encryption and decryption with signatures $\mathcal{E} : \mathcal{K} \times \mathcal{N} \times \mathbb{F}_2^* \to \mathbb{F}_2^*$ and $\mathbf{D} : \mathcal{K} \times \mathcal{N} \times \mathbb{F}_2^* \to \mathbb{F}_2^*$, where $\mathcal{N}$ denotes a nonce space. The nonce $N \in \mathcal{N}$ must not repeat over all encryption queries. The ciphertext expansion function may depend on only the length of M (and also that of associated data for nonce-based authenticated encryption). Distinguishers that obey this requirement are called nonce-respecting. We assume that $\mathcal{E}$ is correct, i.e., for all $K, N, M \in \mathcal{K} \times \mathcal{N} \times \mathbb{F}_2^*$, it holds that $\mathbf{D}_K(N, \mathcal{E}_K(N, M)) = M$. Let $K \twoheadleftarrow \mathcal{K}$ and $\rho : \mathcal{N} \times \mathbb{F}_2^* \to \mathbb{F}_2^*$ be a function that, on input (N, M), computes $C \leftarrow \mathcal{E}_K(N, M)$ for random $K \twoheadleftarrow \mathcal{K}$ and outputs $C' \twoheadleftarrow \mathbb{F}_2^{|C|}$. The nE-security of a nonce-respecting distinguisher $\mathbf{D}$ is defined as

$$\mathbf{Adv}^{\mathsf{nE}}_{\mathcal{E}_K}(\mathbf{D}) \triangleq \Delta_{\mathbf{D}}(\mathcal{E}_K; \rho)\,.$$

3 Definitions of The Forking Zoo

This section describes our proposals and their evolution. We start from the existing constructions with full-round keyed permutations, from PRP2 via STH to SoP and EDMD, move to the slightly more efficient two-branch forked constructions and thereupon, show how we can extend these. Throughout the following, we assume that n, s, t, w are (small) positive integers. We write $\Pi \in \mathsf{Perm}(\{0,1\}^n)$ to denote the secret (that is, keyed) full-round n-bit permutation and $\pi \in \mathsf{Perm}(\{0,1\}^n)$ to denote secret n-bit permutation but assume that it is a reduced-round variant of the full-round permutation Π. In a natural manner, we denote by $\widetilde{\Pi} \in \widetilde{\mathsf{Perm}}(\{0,1\}^t, \{0,1\}^n)$ tweakable permutation and by $\widetilde{\pi} \in \widetilde{\mathsf{Perm}}(\{0,1\}^t, \{0,1\}^n)$ reduced-round variant thereof. Note that this does not imply a certain tweak length.

3.1 The Baseline: From PRP2 over STH to SoP, EDM, and EDMD

PRP2. The concatenation of the outputs of two independent full-round n-bit secret permutations Π_1 and Π_2, namely $\mathsf{PRP2}_{\Pi_1,\Pi_2}$, is a trivial extension to generate a $2n$ output from an n-bit input as follows:

$$\mathsf{PRP2}_{\Pi_1,\Pi_2}(x) = \Pi_1(x) \| \Pi_2(x)\,.$$

Since a PRP provides only $n/2$-bit PRF security from the switching lemma [15, 17, 32], their sum had seen a vast amount of study, e.g. [75, 77, 82, 84].

SoP. The *Sum of Permutations* (SoP) construction is defined as follows. Let Π_1 and Π_2 be two independent full round n-bit secret permutations. Then,

$$\mathsf{SoP}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_1(x) \oplus \Pi_2(x)\,,$$

where the input x to the construction is an n-bit binary string and the construction produces an n-bit output with roughly $O(n)$-bit PRF security – an old result that had been finally proven by [36, 49, 77].

STH. The Summation-Truncation-Hybrid (STH) by Gunsing and Mennink [58] generalizes both PRP2 and SoP. It feeds the n-bit input x into two independent full round n-bit secret permutations Π_1 and Π_2, and splits each of their outputs Y_i, where $Y_i = \Pi_i(x)$, into an s-bit part $Y_{i,0}$ and an $(n-s)$-bit part $Y_{i,1}$, for $i \in \{1,2\}$, i.e., $(Y_{i,0}, Y_{i,1}) \xleftarrow{s,n-s} Y_i$, for $i \in \{1,2\}$. The (s)-bit parts $Y_{1,0}$ and $Y_{2,0}$ are output in plain; the $(n-s)$-bit parts are summed and output as $Y_{1,1} \oplus Y_{2,1}$:

$$\mathsf{STH}_{\Pi_1,\Pi_2}(x) \triangleq Y_{1,0} \| Y_{1,1} \oplus Y_{2,1} \| Y_{2,0}\,.$$

As a result, Gunsing and Mennink could show $O(n - s/2)$-bit security. Thus, it provides a continuous trade-off between output length and security from two independent PRPs and SoP, depending on s.

EDM. While the constructions above consider two parallelizable calls to two independent PRPs, [37, 77] initiated a research direction with their proposal of Encrypted Davies Meyer (EDM) and its dual EDMD construction. For the EDM construction, the n-bit input x is fed into a full-round n-bit secret permutation Π_1 whose output is masked with the

input and the resulting value is finally fed into another independent full round n-bit secret permutation Π_2. Formally, the construction is defined as follows:

$$\mathsf{EDM}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_2(\Pi_1(x) \oplus x)\,.$$

EDMD. For the dual construction, i.e., EDMD, the n-bit input x is fed to a full round n-bit secret permutation Π_1 which is again fed to another independent full round n-bit secret permutation Π_2 whose output is masked with the output of Π_1. Formally, the construction is defined as follows:

$$\mathsf{EDMD}_{\Pi_1,\Pi_2}(x) \triangleq \Pi_2(\Pi_1(x)) \oplus \Pi_1(x)\,.$$

[77] showed optimal PRFsecurity for both EDM and EDMD.

3.2 Forking: From ForkCipher to ForkPRF and ForkSTH

ForkCipher. Instead of using two full round n-bit independent secret permutations, Andreeva et al. proposed ForkCipher [4] that uses three reduced-round n-bit secret permutations. It forks an *intermediate state*, i.e. the output of a first reduced-round permutation π_0. This means it feeds the intermediate state into the other two reduced-round permutations π_1 and π_2 for generating two n-bit outputs. Formally, a ForkCipher is defined as

$$\mathsf{ForkCipher}_{\pi_0,\pi_1}(x) \triangleq (\pi_0(x), \pi_1(x))\,,$$

where x is an n-bit input string. However, in practice, the IFI instantiation uses three smaller permutations

$$\mathsf{IFI}_{\pi_0,\pi_1,\pi_2}(x) \triangleq (\pi_1(\pi_0(x)), \pi_2(\pi_0(x))\,,$$

Thus, it can be more efficient than the PRP2 construction with the help of three reduced-round permutations. The IFI construction can be naturally viewed as the "**forked and reduced**" derivative of PRP2 [2]. While elegant, it aims at a variant of PRP security and provides "only" $O(n/2)$-bit PRF security.

ForkPRF. ForkPRF is derived from the SoP construction similarly as a ForkCipher has been derived from PRP2. Instead of using two full n-bit secret permutations Π_1 and Π_2, one can use three reduced-round permutations π_0, π_1, and π_2. Given an n-bit input x, the construction produces an n-bit output as follows:

$$\mathsf{ForkPRF}_{\pi_0,\pi_1,\pi_2}(x) \triangleq \pi_1(\pi_0(x)) \oplus \pi_2(\pi_0(x))\,.$$

In contrast to a ForkCipher, a ForkPRF sums $Y_1 \oplus Y_2$ and returns only a single output C, where $Y_1 = \pi_1(\pi_0(x)), Y_2 = \pi_2(\pi_0(x))$. Thus, it tries to approximate SoP with smaller permutations while maintaining $O(n)$-bit security by producing n-bits output under the assumption that all the three n-bit secret permutations π_0, π_1 and π_2 are pairwise independent.

ForkSTH. The *Summation-Truncation-Hybrid* (STH) by Gunsing and Mennink [58] generalized PRP2 and SoP. As our second proposal, we introduce a forked variant of the STH construction, called ForkSTH, which generalizes the ForkCipher construction in the sense of providing a continuous trade-off between security and output length.

[2] We call it a forked and reduced derivative because the construction is *forked* and it uses the permutations whose number of rounds is *reduced*

ForkSTH takes an n-bit input, and produces Y_1 and Y_2 exactly as ForkCipher, where $Y_1 = \pi_1(\pi_0(x)), Y_2 = \pi_2(\pi_0(x))$ and π_0, π_1 and π_2 are three reduced-round n-bit secret permutations. As STH, ForkSTH splits Y_1 and Y_2 into $Y_{1,0}, Y_{1,1}, Y_{2,0}, Y_{2,1}$, where $(Y_{1,0}, Y_{1,1}) \xleftarrow{s,n-s} Y_1$ and $(Y_{2,0}, Y_{2,1}) \xleftarrow{s,n-s} Y_2$, and produces $Y_{1,1}\|(Y_{1,0} \oplus Y_{2,0})\|Y_{2,1}$ as an $(n+s)$-bit output. More formally, let π_0, π_1 and π_2 are three reduced-round n-bit secret permutations. Then

$$\mathsf{ForkSTH}_{\pi_0,\pi_1,\pi_2}(x) \stackrel{\Delta}{=} Y_{1,1}\|(Y_{1,0} \oplus Y_{2,0})\|Y_{2,1}\,,$$

where $Y_1 = \pi_1(\pi_0(x))$ and $Y_2 = \pi_2(\pi_0(x))$ such that $(Y_{1,0}, Y_{1,1}) \xleftarrow{s,n-s} Y_1$ and $(Y_{2,0}, Y_{2,1}) \xleftarrow{s,n-s} Y_2$. We have shown that ForkSTH achieves $O(n-s/2)$-bit PRF security under the assumption that π_0, π_1 and π_2 are pairwise independent n-bit secret permutations.

3.3 Reducing Numbers of Rounds: From EDMD and EDM to FastPRF and FastPRF-EDM

FastPRF. In [78], Mennink and Neves proposed a more efficient derivative of EDMD, called FastPRF. FastPRF replaces the full-round n-bit secret permutations of EDMD with reduced-round variants. Formally, let π_1, π_2 be two reduced-round n-bit secret permutations. On an n-bit input x, it returns the n-bit output as follows:

$$\mathsf{FastPRF}_{\pi_1,\pi_2}(x) \stackrel{\Delta}{=} \pi_2(\pi_1(x)) \oplus \pi_1(x)\,.$$

Under the assumption that π_1 and π_2 are two independent n-bit permutations, FastPRF achieves optimal PRFsecurity.

FastPRF-EDM. Similar to FastPRF, Mennink and Neves considered also a round-reduced variant of EDM [37] that we call FastPRF-EDM. They preferred FastPRF in the sequel of their work since it uses the hard-to-control intermediate value as a feed-forward value whereas FastPRF-EDM XORs the (user-controlled) plaintext to the intermediate state. Again, the full round n-bit secret permutations are replaced by their reduced-round variants. Formally, let π_1 and π_2 be two reduced-round n-bit secret permutations. On an n-bit input x, it returns the n-bit output as follows:

$$\mathsf{FastPRF\text{-}EDM}_{\pi_1,\pi_2}(x) \stackrel{\Delta}{=} \pi_2(\pi_1(x) \oplus x)\,.$$

Again, FastPRF-EDM achieves the optimal PRF security under the assumption that π_1 and π_2 are two independent n-bit permutations.

3.4 Multiple Forks: From MFC to ForkCENC and ForkSTHCENC

The Multi-Fork-Cipher MFC[w] by Andreeva et al. [2] extended the ForkCipher from two to w outputs. For a given parameter $w \geq 2$, $(w+1)$ permutations $\pi_0, \pi_1, \ldots, \pi_w$ and a given n-bit input x, the multi-forkcipher produces w many n-bits output as follows:

$$\mathsf{MFC}[w]_{\pi_0,\ldots,\pi_w}(x) \stackrel{\Delta}{=} \|_{i=1}^{w} \pi_i(x)\,.$$

The IFIM principle uses it to

$$\mathsf{IFIM}[w]_{\pi_0,\ldots,\pi_w}(x) \stackrel{\Delta}{=} \|_{i=1}^{w} \pi_i(\pi_0(x))\,.$$

Note that it computes $X \leftarrow \pi_0(x)$ as the original IFI paradigm but produces w outputs $C_i \leftarrow \pi_i(X)$, for all $i \in [w]$. Thus, it needs $w+1$ calls to permutations for a wn-bit output with $O(n/2)$ bits of PRF security under the assumption that all $(w+1)$ permutations are pairwise independent. Note that, MFC[2] is actually the ForkCipher construction.

ForkCENC. Similarly to the extension that MFC represents for the ForkCipher, we can extend ForkPRF by forking more blocks in the middle and adding the first output to each of the other bottom-permutation outputs. We call the resulting construction ForkCENC since it is a variant of CENC, where the full-round primitives are replaced by reduced-round variants. We formally define the construction as follows: let $w \geq 2$ be a parameter and $\pi_0, \pi_1, \ldots, \pi_{w+1}$ be $(w+2)$ many n-bit permutations. Then, for an n-bit input x, it produces wn-bits output as follows:

$$\mathsf{ForkCENC}[w]_{\pi_0,\ldots,\pi_{w+1}}(x) \triangleq \|_{i=2}^{w+1} \pi_i(\pi_0(x)) \oplus \pi_1(\pi_0(x)) \,.$$

In other words, it computes $X \leftarrow \pi_0(x)$, and $Y_i \leftarrow \pi_i(X)$ for all $i \in [w+1]$, and $C_i \leftarrow Y_1 \oplus Y_{i+1}$ for all $i \in [w]$. We have shown that the resulting scheme provides $O(n)$-bit security under the assumption that all $(w+2)$ permutations are pairwise independent. Note that it does not inherit a slight security degradation in terms of w as CENC, i.e. $O(n - \log_2(w^2))$, since the inputs in a w-block chunk in CENC differed pairwise and targeted the same permutation. In contrast, the permutations in ForkCENC differ.

ForkSTHCENC. STH and ForkSTH generalize the spectrum between PRP2 and SoP and between ForkCipher and ForkPRF, respectively. In a similar line, we introduce ForkSTHCENC that covers the spectrum between MFC$[w]$ and ForkCENC$[w]$. We formally define the construction as follows. Let $w \geq 2$ and $t \geq 1$ be two given parameters. Then, for a given sequence of $(w+1)$ many reduced-round n-bit permutations $\pi_0, \pi_1, \ldots, \pi_w$ and for a $n-t$ bit string x, we define the construction as

$$\mathsf{ForkSTHCENC}[w]_{\pi_0,\pi_1,\ldots,\pi_w}(x) \triangleq Y_{1,1} \| \left(\|_{i=2}^{w} ((Y_{1,0} \oplus Y_{i,0}) \| Y_{i,1}) \right),$$

where $Y_i = \pi_i(\pi_0(x \| \langle 0 \rangle_t))$ and $(Y_{i,0}, Y_{i,1}) \xleftarrow{s,n-s} Y_i$. In particular, the construction computes the values Y_i, for all $i \in [w]$, as by MFC$[w]$. Then, it splits them as STH and ForkSTH did into $(Y_{i,0}, Y_{i,1}) \xleftarrow{s,n-s} Y_i$. Next, it computes $C_i \leftarrow (Y_{1,0} \oplus Y_{i,0}) \| Y_{i,1}$ for all $i \in [2, w]$ and finally output $Y_{1,1} \|_{i=2}^{w} C_i$. Note that, this construction outputs s-bits in addition to $(w-1)n$ bits, and still provides $O(n-s)$-bit PRFsecurity under the assumption that all $(w+1)$ permutations are pairwise independent.

3.5 Multiple Forks: From MFC to ForkEDMD and ForkEDM-CTR

ForkEDMD. Similar to ForkCENC and ForkSTHCENC, which are VOL extensions of ForkPRF and ForkSTH, respectively, we introduce an extension of FastPRF by using the intermediate value in the middle for more Davies-Meyer constructions in parallel with pairwise independent permutations. We call this construction ForkEDMD$[w]$, which is formally defined as follows: let $w \geq 2$ and $s \geq 1$ be two given integer parameters. Then, for a given sequence of $(w+1)$ many reduced-round n-bit permutations $\pi_0, \pi_1, \ldots, \pi_w$ and for an $n-s$-bit binary string x, we define the construction as

$$\mathsf{ForkEDMD}[w]_{\pi_0,\pi_1,\ldots,\pi_w}(x) \triangleq \|_{i=1}^{w} \pi_i(\pi_0(x \| \langle 0 \rangle_s)) \oplus \pi_0(x \| \langle 0 \rangle_s) \,.$$

Like FastPRF, it computes $X \leftarrow \pi_0(x \| \langle 0 \rangle_s)$; thereupon, it derives and outputs $C_i \leftarrow \pi_i(X) \oplus X$, for all $i \in [w]$. We have shown that the construction achieves $O(n)$-bit PRF security under the assumption that all $w+1$ permutations are pairwise independent. Thus, it combines the efficiency of MFC$[w]$ with the PRF security of ForkPRF.

ForkEDM-CTR. We introduce an extension of FastPRF-EDM similarly as ForkEDMD extends ForkEDMD. Again, we use the intermediate value in the middle for more parallel

Davies-Meyer constructions with pairwise independent permutations. Though, we have to multiply the feed-forward value by the power of a generator in the field. We call the resulting construction ForkEDM-CTR[w]: let $w \geq 2$ and $s \geq 1$ be two given integer parameters. Then, for $(w+1)$ reduced-round n-bit permutations $\pi_0, \pi_1, \ldots, \pi_w$, and for a given $n-s$ bit string x, the construction is defined as

$$\mathsf{ForkEDM\text{-}CTR}[w]_{\pi_0,\pi_1,\ldots,\pi_w}(x) \stackrel{\Delta}{=} \|_{i=1}^{w} \pi_i(\pi_0(x\|\langle 0\rangle_s) \oplus 2^{i-1}(x\|\langle 0\rangle_s)) \,.$$

In other words, it computes $X = \pi_0(x\|\langle 0\rangle_s)$. Then, the inputs to the i-th primitive call in the bottom row is given by $Y_i = X \oplus 2^{i-1}(x\|\langle 0\rangle_s)$, for all $i \in [w]$. The results are computed as $C_i = \pi_i(Y_i)$, for all $i \in [w]$. We emphasize that Mennink and Neves explicitly proposed the variant of FastPRF based on EDMD and not on EDM [78] as a heuristic. The latter variant – we name it FastPRF-EDM – might provide the adversary with too much freedom over differentials up to the point after π_0. Thus, one may have to choose a stronger permutation for π_0 here.

Remark 1. We emphasize that the field multiplications in the middle layer of ForkEDM-CTR is crucial for its security. Let ForkEDM[w] be the construction without the multiplications, which is defined analogously as:

$$\mathsf{ForkEDM}[w]_{\pi_0,\pi_1,\ldots,\pi_w}(x\|\langle 0\rangle_s) \stackrel{\Delta}{=} \|_{i=1}^{w} \pi_i(\pi_0(x\|\langle 0\rangle_s) \oplus (x\|\langle 0\rangle_s)) \,.$$

ForkEDM[w] provides only birthday-bound security, which can be illustrated easily. Let $X = \pi_0(x\|\langle 0\rangle_s) \oplus (x\|\langle 0\rangle_s)$ and $Y_i = \pi_i(X)$ for $i \in [w]$. Now, whenever two queries of $X^k = X^\ell$ for $x^k \neq x^\ell$ collide, which happens at the birthday bound, all outputs $Y_i^k = Y_i^\ell$ will collide. Thus, whenever $w \geq 2$, this will yield a distinguisher that is absent in the original EDM construction (which coincides with ForkEDM[1]).

3.6 Comparison to $\widetilde{\mathsf{MFC}}$

In [2], Andreeva et al. used MFC as a primitive to define and analyze many variants of CTR modes, instantiated with tweakable secret permutations. We denote tweaked MultiForkCiphers as $\widetilde{\mathsf{MFC}}$. Therein, an instance can use a random IV or a nonce as its tweak. The resulting construction can also provide optimal PRF security. However, it also needs a tweakable keyed primitive with a sufficiently large tweak space that can absorb the IV or nonce of the mode.

In contrast, this work considers designs that remain secure when instantiated with few *untweaked* secret independent permutations, or – to address practice – with a single tweakable permutation with a *very small* tweak space for domain separation. As a result, our constructions avoid the requirement of [2] of larger tweaks, which can save hardware area or computational effort since it can imply a simpler tweak schedule or more efficient precomputation and cache management of round tweaks. Thus, our proposal uses a slightly different but potentially more efficient primitive, where the precise efficiency impacts depend on the concrete primitive and platform. Note that our constructions could also be instantiated from tweakable permutations with larger tweaks when the tweak could be used for even higher security (if it contains a nonce or random IV) or for authenticating associated data during encryption.

3.7 Are Those All Optimally Secure Constructions?

In [34], Chen et al. conducted a systematic study of PRFs and MACs of the secure variants from the sum of independent permutations. They showed that among schemes with two permutation calls, only six constructions provided optimal PRF security: SoP, EDM, and EDMD, as well as their variants with the input summed to the output. We can safely

disregard the latter variants since they solely add a redundant and reversible operation. Since a highly secure MAC should be based upon a highly secure PRF, the former three constructions comprise the set of constructions that are relevant for our studies of how to reduce, fork, and extend them. Thus, we propose ForkCENC as a forked extension of SoP, ForkEDMD as the extension of EDMD, and ForkEDM-CTR as the extension of EDM to cover all constructions with n-bit security.

4 Security Proofs

In this section, we prove the security of our proposed constructions. Note that in all the security proofs we assume that the underlying primitives are ideal permutations, i.e., the block ciphers used in the construction are replaced by their ideal counterparts, i.e., random permutations, at the cost of the PRP advantage of the block cipher. As a result, we study the indistinguishability advantage of the resulting constructions with respect to a computationally unbounded (a.k.a information-theoretic) distinguisher.

4.1 H-coefficient Technique

Prior, we set up the general framework for proving the security of the constructions in this work. We consider an information-theoretic deterministic distinguisher $\mathbf{D}$ that interacts with oracles in either a real or an ideal world: in the former, it interacts with the construction oracle $\mathcal{O}_{\text{real}}$ of our concern, and in the ideal world, with an ideal oracle $\mathcal{O}_{\text{ideal}}$. For nonce-encryption security, we consider an ideal oracle to be a random function over an appropriate domain and space. We summarize the interaction of the distinguisher with the oracle in a transcript $\tau = \{(x_1, y_1), \ldots, (x_q, y_q)\}$, where q is the total number of queries that $\mathbf{D}$ can make to the oracle and (x_i, y_i) represent the i-th query of $\mathcal{D}$ and the corresponding response, respectively. We assume that $\mathbf{D}$ never makes any pointless queries and the transcript does not contain any duplicate elements. To simplify the proofs, we modify the experiment by releasing internal information $\mathbf{S}$ to the distinguisher after $\mathbf{D}$ has finished its interaction with the oracle, but before it has output its decision bit. In the real world, the actual internal state generated in the construction is revealed as the additional information, whereas in the ideal world, dummy states $\mathbf{S}$ are sampled closely following the distribution of $\mathbf{S}$ generated in the real world and revealed to the distinguisher. In the following, the complete transcript is $\tau = \{(x_1, y_1), (x_2, y_2), \ldots, (x_q, y_q), \mathbf{S}\}$. Note that the modified experiment only makes the distinguisher more powerful and hence the distinguishing advantage of $\mathbf{D}$ in this experiment is at least that of in the former one. Let X_{real} be a random variable for transcripts τ in the real world and $\mathsf{X}_{\text{ideal}}$ be a random variable for transcripts τ in the ideal world. The probability of realizing a transcript τ in the ideal (resp. real) world is called *ideal* (*resp. real*) *interpolation probability*. A transcript τ is said to be attainable with respect to $\mathbf{D}$ if its ideal interpolation probability is non-zero. Let Θ denote the set of all attainable transcripts. Following these notations, we now state a combinatorial result, called the H-Coefficient technique by Patarin [83], which is used to establish an upper bound on the distinguishing advantage of two random systems.

Theorem 1 (H-Coefficient Technique)**.** Let $\Theta = \mathsf{GoodT} \sqcup \mathsf{BadT}$ be some partition of the set of attainable transcripts. Suppose there exists $\epsilon_{\text{ratio}} \geq 0$ such that for any $\tau \in \mathsf{GoodT}$,

$$\frac{\mathsf{p}_{\text{re}}(\tau)}{\mathsf{p}_{\text{id}}(\tau)} := \frac{\Pr[\mathsf{X}_{\text{real}} = \tau]}{\Pr[\mathsf{X}_{\text{ideal}} = \tau]} \geq 1 - \epsilon_{\text{ratio}}\,,$$

and there exists $\epsilon_{\text{bad}} \geq 0$ such that $\Pr[\mathsf{X}_{\text{ideal}} \in \mathsf{BadT}] \leq \epsilon_{\text{bad}}$. Then,

$$\Delta_{\mathbf{D}}(\mathcal{O}_{\text{real}}; \mathcal{O}_{\text{ideal}}) \leq \epsilon_{\text{ratio}} + \epsilon_{\text{bad}}\,. \tag{1}$$

In this paper, we mainly prove the PRF security of all the proposed constructions using the H-Coefficient technique, where the ideal oracle will be simply a uniform random function. Due to the application of the H-Coefficient technique, we need to properly identify the bad events, upper bounding their probabilities in the ideal world, and finally lower bounding the ratio of the real-to-ideal interpolation probability for good transcripts. Moreover, in this paper, the additional information $\mathbf{S}$, which will be revealed to the distinguisher after the interaction, will be some internal states generated in the construction. We once again remind the reader that we will carry out the proofs in the information-theoretic setting, where all the block ciphers of the construction will be replaced by n-bit independent permutations at the cost of the PRP advantage of the underlying block ciphers.

4.2 Security Proof for ForkPRF

Theorem 2. Let $\pi_0, \pi_1, \pi_2 \twoheadleftarrow \mathsf{Perm}(\{0,1\}^n)$ be independent random permutations and $n \geq 7$ and $q \leq 2^n/17$ be positive integers. Let $\mathbf{D}$ be a PRF distinguisher on $\mathsf{ForkPRF}_{\pi_0,\pi_1,\pi_2}$. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkPRF}}(\mathbf{D}) \leq \frac{19q^2}{2^{2n}} + \frac{8n^3}{2^{2n}}\,.$$

Proof. Before we begin the proof of the construction, let $X := \pi_0(M)$ be the intermediate variable of the construction which is released to the distinguisher as additional information, i.e., when the distinguisher $\mathbf{D}$ has interacted with the oracle in the real world, we release $X_1, X_2, \ldots, X_q$ to $\mathbf{D}$ after the interaction is over, but before it has output its decision bit. In contrast, in the ideal world, we sample $X_1, X_2, \ldots, X_q \twoheadleftarrow_{\mathsf{wor}} \{0,1\}^n$ and release it to $\mathbf{D}$ after the interaction is over, but before it has output its decision bit. Thus, we represent the overall transcript of the distinguisher $\mathbf{D}$ as

$$\tau = \{(M_1, C_1, X_1), (M_2, C_2, X_2), \ldots, (M_q, C_q, X_q)\}.$$

In this proof, we identify the set of bad transcripts as the empty set; hence $\epsilon_{\text{bad}} = 0$. Therefore, it remains now to lower bound the ratio of real to ideal interpolation probability. As each C_i is uniformly and independently distributed over $\{0,1\}^n$, each X_i is uniformly distributed over $\{0,1\}^n \setminus \{X_1, X_2, \ldots, X_{i-1}\}$, and each X_i is independently distributed over all C_i, the ideal interpolation probability becomes:

$$\Pr[\mathsf{X}_{\text{ideal}} = \tau] = \frac{1}{2^{nq}} \cdot \frac{1}{(2^n)_q}.$$

To compute the real interpolation probability, we need to count the number of permutations π_0 such that $\pi_0(M_i) = X_i$ for all $i \in [q]$ and count the number of permutations (π_1, π_2) such that they satisfy the following system of equations:

$$\mathcal{E} = \begin{cases} \pi_1(X_1) \oplus \pi_2(X_1) & = C_1 \\ \pi_1(X_2) \oplus \pi_2(X_2) & = C_2 \\ \vdots & \vdots \ \vdots \\ \pi_1(X_q) \oplus \pi_2(X_q) & = C_q\,. \end{cases}$$

It is easy to see that the number of permutations π_0 that map M_i to X_i for all $i \in [q]$ is $(2^n - q)!$. However, from the result of Mirror Theory by Dutta et al. [49], one can see that the number of permutation tuples (π_1, π_2) satisfying $\mathcal{E}$ is at least

$$\frac{(2^n)_q \cdot (2^n)_q}{2^{nq}} \cdot \left(1 - \frac{19q^2}{2^{2n}} - \frac{8n^3}{2^{2n}}\right),$$

provided $n \geq 7$ and $q \leq 2^n/17$. Therefore, the real interpolation probability becomes:

$$\Pr[\mathsf{X}_{\mathrm{real}} = \tau] = \frac{1}{2^{nq}} \cdot \frac{1}{(2^n)_q} \cdot \left(1 - \frac{19q^2}{2^{2n}} - \frac{8n^3}{2^{2n}}\right).$$

The result follows by taking the ratio of the real-to-ideal interpolation probability, where $\epsilon_{\mathrm{ratio}} = 19q^2/2^{2n} + 8n^3/2^{2n}$.

4.3 Security Proof for ForkCENC

Theorem 3. Let w, n, and q be positive integers with $n \geq 30$ and $q \leq 2^n/12(w+1)^2$, and let π_0, π_1, …, $\pi_{w+1} \leftarrow \mathsf{Perm}(\{0,1\}^n)$ be independent random permutations. Let $\mathbf{D}$ be a PRF distinguisher on the construction $\mathsf{ForkCENC}_{\pi_0,\pi_1,\ldots,\pi_{w+1}}$. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkCENC}}(\mathbf{D}) = 0\,.$$

Proof. Let $X_i := \pi_0(N)$ be the intermediate variable of the construction corresponding to the i-th query which is released to the distinguisher as additional information, i.e. when the distinguisher $\mathbf{D}$ has interacted with the oracle in the real world, we release $X_1, X_2, \ldots, X_q$ to $\mathbf{D}$ after the interaction is over, but before it has output its decision bit. In contrast, in the ideal world, we sample $X_1, X_2, \ldots, X_q \xleftarrow{\mathsf{wor}} \{0,1\}^n$ and release it to $\mathbf{D}$ after the interaction is over, but before it has output its decision bit. Thus, we represent the overall transcript of the distinguisher $\mathbf{D}$ as

$$\tau = \{(N_1, (C_1[1], C_1[2], \ldots, C_1[w]), X_1), \ldots, (N_q, (C_q[1], C_q[2], \ldots, C_q[w]), X_q)\}\,.$$

In this proof, we identify the set of bad transcripts as an empty set, and hence $\epsilon_{\mathrm{bad}} = 0$. Therefore, it remains to lower bound the ratio of the real-to-ideal interpolation probability. As each C_i is uniformly and independently distributed over $\{0,1\}^n$, each X_i is uniformly distributed over $\{0,1\}^n \setminus \{X_1, X_2, \ldots, X_{i-1}\}$, and each X_i is independently distributed over all C_i, the ideal interpolation probability becomes:

$$\Pr[\mathsf{X}_{\mathrm{ideal}} = \tau] = \frac{1}{2^{wnq}} \cdot \frac{1}{(2^n)_q}\,.$$

To compute the real interpolation probability, we need to count the number of permutations π_0 such that $\pi_0(N_i) = X_i$ for all $i \in [q]$ and count the number of permutations $(\pi_1, \pi_2, \ldots, \pi_w)$ such that they satisfy the following system of equations for each $i \in [q]$:

$$\mathcal{E}_i = \begin{cases} \pi_1(X_i) \oplus \pi_2(X_i) & = C_i[1] \\ \pi_1(X_i) \oplus \pi_3(X_i) & = C_i[2] \\ \vdots & \vdots \ \vdots \\ \pi_1(X_i) \oplus \pi_w(X_i) & = C_i[w]\,. \end{cases}$$

Note that $\mathcal{E}_i$ is a system of bivariate affine equations over $w+1$ variables with block maximality $w+1$. Therefore, to lower bound the number of permutation tuples $(\pi_1, \pi_2, \ldots, \pi_w)$ that satisfy $\mathcal{E}_i$ for all $i \in [q]$, we require the result of *Mirror theory for general* $\xi_{\max}$. Therefore, from the Mirror theory result for general $\xi_{\max}$ [39] with $\xi_{\max} = w+1$, one can see that the number of permutation tuples $(\pi_1, \pi_2, \ldots, \pi_w)$ satisfying $\mathcal{E}_i$ for all $i \in [q]$ is at least

$$\prod_{i=1}^{w} \frac{(2^n)_q}{2^{nq}}\,,$$

provided $n \geq 30$ and $q \leq 2^n/12(w+1)^2$. Moreover, it is easy to see that the number of permutation π_0 that maps N_i to X_i for all $i \in [q]$ is $(2^n - q)!$. Therefore, the real interpolation probability becomes:

$$\Pr[\mathsf{X}_{\text{real}} = \tau] = \frac{1}{2^{wnq}} \cdot \frac{1}{(2^n)_q}.$$

The result follows by taking the ratio of the real-to-ideal interpolation probability. □

4.4 Security Proof for FastPRF and FastPRF-EDM

Theorem 4. Let π_1 and $\pi_2 \twoheadleftarrow \mathsf{Perm}(\{0,1\}^n)$ be two independent random permutations and $n \geq 7$ and $q \leq 2^n/17$ be positive integers. Let $\mathbf{D}$ be a PRF distinguisher on the construction FastPRF. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{FastPRF}}(\mathbf{D}) \leq \frac{19q^2}{2^{2n}} + \frac{8n^3}{2^{2n}}\,. \tag{2}$$

Moreover, let $\xi_{\max} \geq 1$, and $q \leq 2^n/12\xi^2_{\max}$, as long as $n \geq 30$, be two given integer parameters. Let $\mathbf{D}$ be a PRF distinguisher on the construction FastPRF-EDM. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{FastPRF\text{-}EDM}}(\mathbf{D}) \leq \frac{\binom{q}{\xi_{\max}+1}}{2^{n\xi_{\max}}}\,. \tag{3}$$

Note that the constructions FastPRF and FastPRF-EDM actually get boils down to the EDMD and EDM construction respectively under the assumption that π_1 and π_2 are two independent n-bit permutations. As a result, the security proof of FastPRF and FastPRF-EDM under the standard model is exactly the same as that of EDMD and EDM respectively. Therefore, by following the proof of Theorem 6 of [77], we obtain the security bound of FastPRF and by following the proof of Theorem 4 of [77], we obtain the security bound of FastPRF-EDM. Note that, the bound for EDMD as shown in [77] differs from Eqn. (2) as the earlier version of Mirror theory result for $\xi_{\max} = 2$ [85, 84] was used in [77] to derive the bound of EDMD, whereas we have used the correct bound of Mirror theory from [49]. In particular, the ratio of real to ideal interpolation probability in the proof of EDM uses the earlier bound of Mirror theory result for $\xi_{\max} = 2$ which is $(1-q/2^n)$ [85, 84], whereas we are using Theorem 3 of [49] that yields $(1 - 19q^2/2^{2n} - 8n^3/2^{2n})$ bound to the ratio of real to ideal interpolation probability of FastPRF. Similarly, the bound for EDM as shown in [77] differs from Eqn. (3) as the earlier version of Mirror theory result for general $\xi_{\max}$ [84, 79] was used in [77] to derive the bound of EDM, whereas we have used the correct bound of Mirror theory result for general $\xi_{\max}$ from [39]. In particular, Mennink and Neves have used the earlier bound of Mirror theory result for general $\xi_{\max}$ [84, 79] that yields $(1 - q/2^n)$ bound to the ratio of real to ideal interpolation probability of EDM, whereas we are using Theorem 1 of [39] that yields that the real interpolation probability is close to the ideal interpolation probability of FastPRF-EDM.

4.5 Security Proof for ForkEDMD

Theorem 5. Let w, n, and q be positive integers with $n \geq 30$ and $q \leq 2^n/12(w+1)^2$, and let $\pi_0, \pi_1, \ldots, \pi_w \twoheadleftarrow \mathsf{Perm}(\{0,1\}^n)$ be independent random permutations. Let $\mathbf{D}$ be a PRF distinguisher on the construction $\mathsf{ForkEDMD}_{\pi_0,\pi_1,\ldots,\pi_w}$. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkEDMD}}(\mathbf{D}) = 0\,.$$

Proof. Let $\pi'_1, \pi'_2, \ldots \pi'_w \leftarrow\!\!\$\ \mathsf{Perm}(\{0,1\}^n)$ such that each π'_1 is independent from π_0 and each π'_i is independent from each π_j. Let $\mathbf{D}$ be a distinguisher that distinguishes the

real oracle $\mathcal{O}_{\mathrm{re}} = (\pi_1 \circ \pi_0 \oplus \pi_0, \pi_2 \circ \pi_0 \oplus \pi_0, \ldots, \pi_w \circ \pi_0 \oplus \pi_0)$ from the ideal oracle $\mathcal{O}_{\mathrm{id}} = (\mathsf{RF}_1, \mathsf{RF}_2, \ldots, \mathsf{RF}_w)$, where each RF_i is a uniform random function from n-bits to n-bits that are independently sampled over all RF_j. Now, we consider an another real oracle $\mathcal{O}'_{\mathrm{re}} = (\pi_0 \oplus \pi'_1, \pi_0 \oplus \pi'_2, \ldots, \pi_0 \oplus \pi'_w)$. Now, we can rewrite

$$\begin{aligned}\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkEDMD}}(\mathbf{D}) &= |\Pr[\mathbf{D}^{\mathcal{O}_{\mathrm{re}}} \Rightarrow 1] - \Pr[\mathbf{D}^{\mathcal{O}_{\mathrm{id}}} \Rightarrow 1]| \\ &\leq |\Pr[\mathbf{D}^{\mathcal{O}_{\mathrm{re}}} \Rightarrow 1] - \Pr[\mathbf{D}^{\mathcal{O}'_{\mathrm{re}}} \Rightarrow 1]| + |\Pr[\mathbf{D}^{\mathcal{O}'_{\mathrm{re}}} \Rightarrow 1] - \Pr[\mathbf{D}^{\mathcal{O}_{\mathrm{id}}} \Rightarrow 1]| \\ &\overset{(1)}{=} |\Pr[\mathbf{D}^{\mathcal{O}'_{\mathrm{re}}} \Rightarrow 1] - \Pr[\mathbf{D}^{\mathcal{O}_{\mathrm{id}}} \Rightarrow 1]|\,,\end{aligned}$$

where (1) holds due to the fact that the distinguishing advantage of $\mathbf{D}$ in distinguishing the oracle $\mathcal{O}_{\mathrm{re}}$ from $\mathcal{O}'_{\mathrm{re}}$ equals to 0, as one can easily see that revealing the permutation π_0 to the distinguisher prior to the experiment effectively boils down to distinguish permutation π_i from π'_i for all $i \in [w]$. Therefore, it boils down to upper bound the distinguishing advantage of $\mathbf{D}$ in distinguishing the output of $\mathcal{O}'_{\mathrm{re}}$ from $\mathcal{O}_{\mathrm{id}}$. For this purpose, we first note that the q evaluations of the construction $\mathcal{O}'_{\mathrm{re}}$ can be translated to an equivalent system as follows:

$$\mathcal{E}_i = \begin{cases} \pi_0(N_i) \oplus \pi'_1(N_i) & = C_i[1] \\ \pi_0(N_i) \oplus \pi'_2(N_i) & = C_i[2] \\ \vdots & \vdots \ \vdots \\ \pi_0(N_i) \oplus \pi'_w(N_i) & = C_i[w]\,. \end{cases}$$

for all $i \in [q]$. Now, to upper bound the distinguishing advantage of $\mathbf{D}$ in distinguishing the output of $\mathcal{O}'_{\mathrm{re}}$ from $\mathcal{O}_{\mathrm{id}}$ using the H-coefficient technique, we identify the set of bad transcripts to be an empty set and hence $\epsilon_{\mathrm{bad}} = 0$. Therefore, it remains to lower bound the ratio of real to ideal interpolation probability. As each $C_i[\alpha]$ is uniformly and independently distributed over $\{0,1\}^n$, the ideal interpolation probability becomes:

$$\Pr[\mathsf{X}_{\mathrm{ideal}} = \tau] = \frac{1}{2^{wnq}}\,.$$

To compute the real interpolation probability, we need to count the number of permutation tuples $(\pi_0, \pi'_1, \ldots, \pi'_w)$ that satisfy the above system of equations $\mathcal{E}_i$ for each $i \in [q]$. Note that $\mathcal{E}_i$ is a system of bivariate affine equations over $w+1$ variables with block maximality $w+1$. Therefore, to lower bound the number of permutation tuples $(\pi_0, \pi'_1, \pi'_2, \ldots, \pi'_w)$ that satisfy $\mathcal{E}_i$ for all $i \in [q]$, we require the result of *Mirror theory for general* $\xi_{\max}$. Therefore, from the Mirror theory result for general $\xi_{\max}$ [39], one can see that the number of permutations $(\pi_0, \pi'_1, \pi'_2, \ldots, \pi'_w)$ satisfying $\mathcal{E}_i$ for all $i \in [q]$ is at least

$$\prod_{i=1}^{w} \frac{(2^n)_q}{2^{nq}},$$

provided $n \geq 30$ and $q \leq 2^n/12(w+1)^2$. Therefore, the real interpolation probability becomes:

$$\Pr[\mathsf{X}_{\mathrm{real}} = \tau] = \frac{1}{2^{wnq}} \cdot \frac{1}{(2^n)_q}\,.$$

The result follows by taking the ratio of the real to ideal interpolation probability. □

4.6 Security Proof for ForkEDM-CTR

Theorem 6. Let w, n, and q be positive integers with $n \geq 30$ and $q \leq 2^n/12(w+1)^2$ and let $\pi_0, \pi_1, \ldots, \pi_w \leftarrow \mathsf{Perm}(\{0,1\}^n)$ be independent random permutations. Let $\mathbf{D}$ be a PRF distinguisher on the construction $\mathsf{ForkEDM\text{-}CTR}_{\pi_0,\pi_1,\ldots,\pi_w}$. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkEDM\text{-}CTR}}(\mathbf{D}) \leq \frac{qw^2}{2^n}\,.$$

Proof. To prove its security, we consider a slightly different construction, where each π_i is replaced by its inverse π_i^{-1} for each $i \in [w]$. As $\pi_0, \pi_1, \ldots, \pi_w$ are all mutually independent, these two constructions are provably equally secure. However, it is more convenient to establish the security proof for the latter construction as one can view an evaluation

$$C_i[\alpha] = \pi_\alpha^{-1}(\pi_0(N_i) \oplus 2^{\alpha-1}N_i)$$

as the xor of two permutations in the middle of the function

$$\pi_0(N_i) \oplus \pi_\alpha(C_i[\alpha]) = 2^{\alpha-1}N_i,$$

for all $1 \leq \alpha \leq w$. Therefore, q evaluations of the latter construction can be translated to an equivalent system as follows:

$$\mathcal{E}_i = \begin{cases} \pi_0(N_i) \oplus \pi_1(C_i[1]) & = N \\ \pi_0(N_i) \oplus \pi_2(C_i[2]) & = 2N \\ \vdots & \vdots\ \vdots \\ \pi_0(N_i) \oplus \pi_w(C_i[w]) & = 2^{w-1}N\,. \end{cases}$$

Let τ denote the summary of the interaction between the distinguisher $\mathbf{D}$ and the oracle, where τ is represented as

$$\tau = \{(N_1, (C_1[1], C_1[2], \ldots, C_1[w])), \ldots, (N_q, (C_q[1], C_q[2], \ldots, C_q[w]))\}\,.$$

We call a transcript τ bad if it satisfies either of the following events:

- $\exists i \in [q], \alpha \neq \beta \in [w]$ such that $C_i[\alpha] = C_i[\beta]$.
- $\exists i, j \in [q]$, $i \neq j$, $\alpha, \beta \in [w]$ such that $C_i[\alpha] = C_j[\beta], 2^{\alpha-1}N_i = 2^{\beta-1}N_j$.

As each $C_i[\alpha]$ is uniformly and independently distributed over all $C_j[\beta]$, we upper bound the probability of the first bad event to $q\binom{w}{2}/2^n$. To upper bound the probability of the second bad event, we would first like to note that for a fixed choice of index i, there is at most one choice of j such that $N_j = 2^{\alpha-\beta}N_i$. Therefore, for a fixed choice of indices, the probability of the event is upper bounded to at most 2^{-n} using the randomness of $C_i[\alpha]$. Note that the choice of i is at most q, the choice of j is at most 1, and the choice of α, β is at most $\binom{w}{2}$. Therefore, the probability of the last bad event is upper bounded to at most $q\binom{w}{2}/2^n$. Hence, we have

$$\epsilon_{\text{bad}} = \frac{qw^2}{2^n}\,.$$

Now, it remains now to lower bound the ratio of real to ideal interpolation probability. As each C_i is uniformly and independently distributed over $\{0,1\}^n$, the ideal interpolation probability becomes:

$$\Pr[\mathsf{X}_{\text{ideal}} = \tau] = \frac{1}{2^{wnq}}\,.$$

To compute the real interpolation probability, we need to count the number of permutation tuples $(\pi_0, \pi_1, \pi_2, \ldots, \pi_w)$ such that they satisfy the above system of equations $\mathcal{E}_i$ for each $i \in [q]$. Note that $\mathcal{E}_i$ is a system of bivariate affine equations over $w+1$ variables with block maximality $w+1$. Therefore, to lower bound the number of permutations $(\pi_0, \pi_1, \pi_2, \ldots, \pi_w)$ that satisfies $\mathcal{E}_i$ for all $i \in [q]$, we require the result of *Mirror theory for general* $\xi_{\max}$. Therefore, from the Mirror theory result for general $\xi_{\max}$ [39], one can see that the number of permutation tuples $(\pi_0, \pi_1, \pi_2, \ldots, \pi_w)$ satisfying $\mathcal{E}_i$ for all $i \in [q]$ is at least

$$\prod_{i=1}^{w} \frac{(2^n)_q}{2^{nq}},$$

provided $n \geq 30$ and $q \leq 2^n/12(w+1)^2$. Therefore, the real interpolation probability becomes:

$$\Pr[\mathsf{X}_{\text{real}} = \tau] = \frac{1}{2^{wnq}} \cdot \frac{1}{(2^n)_q} .$$

The result follows by taking the ratio of the real to ideal interpolation probability. □

4.7 Security Proof for ForkSTH

Theorem 7. Let r, n, a, b and q be positive integers with $r \geq 3$, $a + b = n$, and $q < 2^{b-2}$ and $q \leq 2^n/(3r)$. Let $\pi_0, \pi_1, \ldots, \pi_r \twoheadleftarrow \mathsf{Perm}(\{0,1\}^n)$ be independent random permutations. Let $\mathbf{D}$ be a PRF distinguisher on the construction $\mathsf{ForkSTH}_a[\pi_0, \pi_1, \ldots, \pi_r]$. Then

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkSTH}_a[r]}(\mathbf{D}) \leq \left(\frac{4}{3}\right)^r \left(\frac{rq}{2^{n-a/3}}\right)^{3/2} + 2^{a-1} \cdot \left(\frac{16rq}{2^n}\right)^{2^{b-2}} + \mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{trunc}_a}(rq) .$$

Proof. The general proof strategy will follow that by [58]. Let $\pi_0, \pi_1, \ldots \pi_r \twoheadleftarrow \mathsf{Perm}(\mathbb{F}_2^n)$ such that all permutations π_j are pairwise independent. We consider two oracles, $\mathcal{O}_{\text{ideal}}$ and $\mathcal{O}_{\text{real}}$. Let $\mathbf{D}$ be a distinguisher that is given access to one of them, chosen uniformly at random. $\mathbf{D}$ shall distinguish between both worlds, given the transcript τ of queries of $\mathbf{D}$ to the oracle, the corresponding responses, and intermediate variables. We define by $\mathbf{I}_n$ the identity permutation over $\mathbb{F}_2^n$. For integers $n = a + b$ and $X \in \mathbb{F}_2^n$ with $X = V \| Y$ and $V \in \mathbb{F}_2^a$, $Y \in \mathbb{F}_2^b$, we define $\mathsf{msb}_a(X) = V$ to always return the leftmost a bits of X and $\mathsf{lsb}_b(X) = Y$ to return the b least significant b bits of X, and $(V, Y) \xleftarrow{a,n-a} X$ denotes the splitting of X into an a-bit part V and an $n-a$-bit part Y.

On message input M^i, the real world $\mathcal{O}_{\text{real}}$ uses $\mathsf{ForkSTH}_a[\pi_0, \ldots, \pi_r](M^i)$ and produces and outputs $V_1^i, V_2^i, W_2^i, \ldots, V_r^i, W_r^i$, where $W_j^i = Y_1^i \oplus Y_j^i$ for all $j \in [2..r]$. The values are collected in vectors $\mathbf{V} = (\mathbf{V}^1, \ldots, \mathbf{V}^q)$ and $\mathbf{Y} = (\mathbf{Y}^1, \ldots, \mathbf{Y}^q)$ with $\mathbf{V}^i = (V_1^i, \ldots, V_r^i)$ and $\mathbf{Y}^i = (Y_1^i, \ldots, Y_r^i)$, for all $i \in [q]$. Let $\tau = (\mathbf{V}, \mathbf{W})$ be the transcript. Over all queries, we define the short-hand notation $\mathbf{V}_j = (V_j^1, \ldots, V_j^q)$ for some $j \in [r]$.

The ideal world $\mathcal{O}_{\text{ideal}}$ samples all outputs $V_j^i \twoheadleftarrow \mathbb{F}_2^a$, for all $i \in [q]$ and $j \in [r]$ and samples $W_2^i, \ldots, W_r^i \twoheadleftarrow \mathbb{F}_2^b$, for all $i \in [q]$. We denote $\mathbf{W}^i = (W_2^i, \ldots, W_r^i)$ and $\mathbf{W} = (\mathbf{W}^1, \ldots, \mathbf{W}^q)$. We denote the real-world oracle as $\mathcal{O}_1$ since we will modify it stepwise in the following. It holds that

$$\mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{ForkSTH}_a}(\mathcal{A}) \leq \| \Pr[\mathcal{O}_{\text{ideal}}] - \Pr[\mathcal{O}_1] \| .$$

Next, we separate the a-bit values, $(V_1^i, \ldots, V_r^i)$, given out in clear from the results of the sums, $(W_2^i, \ldots, W_r^i)$. This yields the modified real world $\mathcal{O}_2$. Internally, $\mathcal{O}_2$ uses a function $\mathsf{PTrunc}[r]$ that samples the values $\mathbf{V} = (V_1, \ldots, V_r)$ as a-bit values sampled independently uniformly at random from $\mathbb{F}_2^a$ each. This is given in Algorithm 1. Moreover, we define $\mathsf{PSoP}[r]$, which takes $(V_1, \ldots, V_r)$ and samples $r-1$ permutations compatible to it (if they exist) and computes the vector of sum values, $\mathbf{W} = (W_2^i, \ldots, W_r^i)$, from it. A bad event will be defined when no such compatible permutation exists. We say a transcript τ is bad if bad occurs in τ. We partition the set of all attainable transcripts into a set BadT that consists of exactly all bad transcripts and GoodT of all attainable transcripts that are not bad. For all $j \in [r]$ and given vectors of a-bit strings $\mathbf{V}_j = (V_j^1, \ldots, V_j^q) \in (\mathbb{F}_2^a)^q$, we define $\mathsf{Perm}_{\mathsf{comp}}(\mathbf{V}_j) \subseteq \mathsf{Perm}(\mathbb{F}_2^{n-a})$ as the set of all n-bit permutations that would produce $\mathbf{V}_j$ in their most significant a-bit outputs for the inputs in $\mathbf{V}_j$. The difference between both worlds is upper bounded by

$$\|\Pr[\mathcal{O}_1] - \Pr[\mathcal{O}_2]\| \leq \Pr_{\mathcal{O}_2}[\tau \in \mathsf{bad}] + \mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{trunc}_a}(rq)$$

Algorithm 1 Real-world oracles used in the proof of $\mathsf{ForkSTH}_a[\pi_0, \pi_1, \ldots, \pi_r]$.

```
11: function O_1(M)
12:     π_0, π_1, ... π_r ←$ Perm(F_2^n)
13:     M^1, ..., M^q ← M
14:     for i ← 1 to q do
15:         X^i ← π_0(M^i)
16:         for j ← 1 to r do
17:             (V_j^i, Y_j^i) ←^{a,b} π_j(X^i)
18:             W_j^i ← Y_1^i ⊕ Y_j^i
19:         V^i ← (V_2^i, ..., V_j^i)
20:         W^i ← (W_2^i, ..., W_j^i)
21:     V ← (V^1, ..., V^q)
22:     W ← (W^1, ..., W^q)
23:     τ ← (V, W)
24:     return τ

31: function O_2(M)
32:     V ← PTrunc[r](M)
33:     W ← PSoP[r](M, V)
34:     return τ = (V, W)

41: function PTrunc[r](M)
42:     for i ← 1 to q do
43:         for j ← 1 to r do
44:             V_j^i ←$ F_2^a
45:         V^i ← (V_1^i, ..., V_r^i)
46:     return V = (V^1, ..., V^q)

51: function PSoP[r](M, V)
52:     for j ← 1 to r do
53:         if Perm_comp(V_j) = ∅ then
54:             bad ← true
55:             π_j ←$ I_n
56:         else
57:             π_j ←$ Perm_comp(V_j)
58:     for i ← 1 to q do
59:         for j ← 1 to r do
60:             Y_j^i ← lsb_b(π_j(⟨i⟩))
61:             W_j^i ← Y_1^i ⊕ Y_j^i
62:         W^i ← (W_2^i, ..., W_r^i)
63:     return W = (W^1, ..., W^q)
```

From the triangle inequality, the difference in the setting is at most

$$\|\Pr[\mathcal{O}_{\text{ideal}}] - \Pr[\mathcal{O}_1]\| \leq \|\Pr[\mathcal{O}_{\text{ideal}}] - \Pr[\mathcal{O}_2]\| + \Pr_{\mathcal{O}_2}[\tau \in \mathsf{bad}] + \mathbf{Adv}^{\mathsf{PRF}}_{\mathsf{trunc}_a}(rq)\,.$$

We want to upper bound the distance between the multi-sum of pairwise independent permutations and the function that produces random bits. For the values V_1, V_2, …, V_r, we define counters

$$C_{\mathbf{V},j}(i) \stackrel{\text{def}}{=} \left|\left\{V_j^{i'} : V_j^{i'} = V_j^i\right\}\right|\,,\ \text{for all } j \in [r]\,.$$

Those counters will later have to remain below 2^{b-2}. For the case that one of them exceeds this amount, we define a set bad of vectors $\mathbf{V}$ such that there exists $k \in [r]$ with $C_{\mathbf{V},k}(i) \geq 2^{b-2}$. Given a transcript τ that contains $\mathbf{V}$, we see that

$$\begin{aligned}&\mathbb{E}_\tau[\Pr[\mathcal{O}_{\text{ideal}} = \tau] - \Pr[\mathcal{O}_2 = \tau]]\\ \leq\ &\mathbb{E}_\tau[\Pr[\mathcal{O}_{\text{ideal}} = \tau] - \Pr[\mathcal{O}_2 = \tau | \tau \notin \mathsf{BadT}]] + \Pr[\tau \in \mathsf{BadT}]\,.\end{aligned}$$

Multi-Collision. We can upper bound $\Pr[\tau \in \mathsf{BadT}]$ first, which requires a (2^{b-2})-collision of values $V_j^{i_1} = \cdots = V_j^{i_{2^{b-2}}}$ inside any one of r vectors $\mathbf{V}_j$ in $\mathbf{V}$. Since the values V_j^i are chosen independently and uniformly at random each, the probability for a t-collision is upper bounded by

$$\frac{(rq)^t}{2^{a(t-1)} \cdot t!}$$

where Stirling's approximation can be used for

$$t! \geq \sqrt{2\pi} \cdot \left(\frac{1}{2^{3/2} \cdot t}\right)^t\,.$$

We can rewrite it and substitute $t = 2^{b-2}$

$$\begin{aligned}\Pr[\mathcal{O}_{\text{ideal}} \in \mathsf{bad}] &\leq \frac{1}{\sqrt{2\pi}} \cdot \frac{(rq)^t}{2^{a(t-1)}} \cdot \left(\frac{1}{2^{3/2} \cdot t}\right)^t \\ &\leq \frac{2^a}{\sqrt{2\pi}} \cdot \left(\frac{rq}{2^{a-3/2} \cdot t}\right)^t \\ &\leq \frac{2^a}{\sqrt{2\pi}} \cdot \left(\frac{rq}{2^{a-2} \cdot 2^{b-2}}\right)^{2^{b-2}} \\ &\leq 2^{a-1} \cdot \left(\frac{16rq}{2^n}\right)^{2^{b-2}} .\end{aligned}$$

It remains to upper bound the good transcripts. Since for good transcripts, the vectors $\mathbf{V}$ are sampled equally in both worlds, we can concentrate on the vectors $\mathbf{W}$.

Theorem 8. Let a, b, q, r be positive integers and $\tau = (\mathbf{V}, \mathbf{W})$ be a good transcript such that $C_{\mathbf{V},j}(i) < 2^{b-2}$ holds for all $i \in [q]$ and $j \in [r]$ and $q \leq 2^n/(3r)$. Then, for $r \geq 3$, it holds that

$$\mathbb{E}_\tau \left[\left|\Pr\left[\mathcal{O}_2 = \tau\right] - \Pr\left[\mathcal{O}_{\text{ideal}} = \tau\right]\right|\right] \leq \left(\frac{4}{3}\right)^r \cdot \left(\frac{rq}{2^{n-a/3}}\right)^{3/2} .$$

For the sake of space limitation, we defer the proof of Theorem 8 to Appendix B.

5 From VOL-PRF To Nonce-based Mode

We define a simple conversion of ForkSTHCENC, ForkCENC, ForkEDMD, and ForkEDM-CTR into nonce-based encryption schemes, similar to the conversion from XORP into CENC. Let n, ℓ be the block length in bits and the maximal number of blocks per message. We define a nonce space $\mathcal{N} = \{0,1\}^\nu$, an index space $\mathcal{I} = \{0,1\}^\iota$, such that $\nu, \iota > 0$, $\iota \geq \log_2(\ell)$, and $\nu + \iota = n$. Then, we Let $\Pi[w]_{\boldsymbol{\pi}}$ be a VOL-PRF instantiated with a set of secret permutations $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{w+1})$. We write Π as a short form in the following. Let $\omega = F(w, n)$ be the number of message bits that Π can output at most. Then, we define the conversion of $\Pi[w]_{\boldsymbol{\pi}}$ into a nonce-based encryption scheme $\widehat{\Pi}[\Pi[w]_{\boldsymbol{\pi}}]$ as follows. $\widehat{\Pi}[\Pi[w]_{\boldsymbol{\pi}}]$ takes a nonce $N \in \{N\}$, and the input message $M \in \{0,1\}^*$ and to encrypt M to the ciphertext $C = (C_1 \| \ldots \| Cm)$ as

$$\begin{aligned}(M_1, \ldots, M_m) &\xleftarrow{\omega} M \\ S_i &= \Pi[w]_{\boldsymbol{\pi}}(N \| \langle i-1 \rangle_\iota) \text{ for } i \in [1..m] \\ C_i &= M_i \oplus \mathsf{trunc}_{|M_i|}(S_i) \text{ for } i \in [1..m] .\end{aligned}$$

That is, it splits the message into chunks of at most ω bits each, where the final chunk M_m may be smaller and XORs it with the output of $\Pi[w]_{\boldsymbol{\pi}}(N \| \langle i-1 \rangle_\iota)$.

Theorem 9. Let $\widehat{\Pi}[\Pi[w]_{\boldsymbol{\pi}}]$ be a VOL-PRF instantiated with a set of pairwise independent secret permutations $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{w+1})$, i.e., $\pi_1, \ldots, \pi_{w+1} \leftarrow\!\!\$ \, (\mathsf{Perm}(\{0,1\}^n)^{w+1}$. Let $\mathbf{D}'$ be an adversary on the PRFsecurity of $\Pi[w]_{\boldsymbol{\pi}}$. Then, for any distinguisher $\mathbf{D}$ on the nE security of

$$\mathbf{Adv}^{\mathsf{nE}}_{\Pi[w]_{\boldsymbol{\pi}}}(\mathbf{D}) \leq \left\lceil \frac{q}{w} \right\rceil \cdot \mathbf{Adv}^{\mathsf{PRF}}_{\Pi[w]_{\boldsymbol{\pi}}}(\mathbf{D}') .$$

The result follows from replacing $\Pi[w]_{\boldsymbol{\pi}}$ by a random function and taking the gap between both as the upper bound.

Algorithm 2 Definition of TweAES'.

```
11: function TweAES'_K[w](M)
12:     (K^0, ..., K^{r_t+r_b}) ← KeySchedule(K)
13:     T_0 ← ExpandTweak(0)
14:     S^0 ← M
15:     for i ← 1..r_t do
16:         S^i ← R[K^i ⊕ T_0](S^{i-1})
17:     for t ← 1..w do
18:         T^t ← ExpandTweak(t)
19:         S^{t,r_t} ← S_{t,r_t} ⊕ BC^t
20:         for i ← r_t + 1..r_t + r_b − 1 do
21:             S^{t,i} ← R[K^i ⊕ T^t](S^{t,i-1})
22:         S^{t,r} ← R[0](S_{t,r-1})
23:     return S^{1,r}, ··· , S^{w,r}

31: function R[K](S)
32:     return MC(SR(SB(S))) ⊕ K

41: function KeySchedule(K)
42:     K^0[0..15] ← K
43:     for i ← 1..r_t + r_b do
44:         K^i[0,1] ← Sbox(K^{i-1}[13]), Sbox(K^{i-1}[14])
45:         K^i[2,3] ← Sbox(K^{i-1}[15]), Sbox(K^{i-1}[12])
46:         K^i[4..7] ← K^i[0..3] ⊕ K^{i-1}[4..7]
47:         K^i[8..11] ← K^i[0..3] ⊕ K^{i-1}[8..11]
48:         K^i[12..15] ← K^i[0..3] ⊕ K^{i-1}[12..15]

51: function ExpandTweak(T)
52:     (t_0, t_1, t_2, t_3) ← T
53:     (t_4, t_5, t_6, t_7) ← J · (t_0, t_1, t_2, t_3)^⊤
54:     for i ← 0..7 do
55:         R[i] ← (0^7 ‖ t_i)
        return (0^7 ‖ t_0, ..., 0^7 ‖ t_7, 0, 0, 0, 0, 0, 0, 0, 0)
```

6 Instantiation

6.1 Requirements

Compared to SoP and STH, we want to design more efficient PRFs by using round-reduced instead of full primitive calls and forking from an intermediate state. Thus, we need a primitive that renders our constructions (1) highly efficient, (2) single-key, (3) single-primitive, and (4) sufficiently secure against fixed-key standard attacks. We consider differential, linear, boomerang, impossible-differential, integral, and meet-in-the-middle attacks but will detail what types are particularly relevant in the context of our constructions. By sufficiently secure, we target a *security margin of at least two rounds* compared to the best known attacks.

Since our VOL-PRFs aim at optimal security, constructions based on public permutations are not fully fit for instantiations since the security of sums of two public permutations is usually capped by $O(2n/3)$ bits. Moreover, they demand multiple keys (cf. [19, 31, 33, 50]). Given classical block ciphers, the use of multiple independent primitives as in our constructions would demand either multiple keys or sacrificing parts of the input for encoding the domain.

Tweakable block ciphers (TBCs) allow efficient and effective domain separation of the individual primitive calls without expanding the key material excessively. As a disadvantage, tweak inputs represent additional degrees of freedom to adversaries: the usual tweak(ey) sizes of n, or $2n$ bit, for ciphers with n-bit block size, as e.g. in Skinny [13] or Deoxys-BC [67], exceed what we need for domain separation. The $n/2$-bit tweak in Kiasu-BC [66] suffices, but its diffusion was improved recently by the dedicated small-tweak constructions of the ElasticTweak framework [28, 30]. We employ the latter with the AES round function as a natural choice for instantiations efficient on off-the-shelf processors, as in the ElasticTweak instance TweAES.

The ElasticTweak Framework and TweAES. The ElasticTweak framework [28, 30] can produce large diffusion from very small tweaks by expanding it to larger round tweakeys with a simple code. The authors proposed two concrete instances, TweAES and TweGIFT, which are tweakable variants of AES-128 and GIFT-64, respectively. Those were employed in the NIST LwC second-round candidate EState [27, 29]. The main strategy takes a small four-bit tweak and expands it to affect wide parts of the state using a code $\mathbf{M} = [\mathbf{I}|\mathbf{J}]$ with the identity $\mathbf{I}$ and a binary matrix $\mathbf{J}$ that is the element-wise sum of $\mathbf{J} = \mathbf{I} + \mathbf{1}$, with $\mathbf{1}$ being the all-one matrix. More precisely, the four-bit tweak $\mathbf{T} = (t_0, t_1, t_2, t_3)$ is expanded

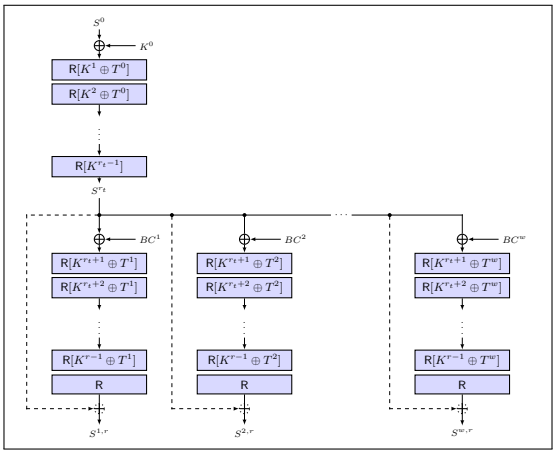


Figure 6.1: Schematic illustration of our AES-based instance TweAES'. The dashed lines and XORs are the feed-forwards that ForkEDMD adds.

to eight bits as $(t_4, t_5, t_6, t_7) = \mathbf{J} \cdot \mathbf{T}^\top$:

$$\begin{bmatrix} t_4 \\ t_5 \\ t_6 \\ t_7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} .$$

In TweAES, the bits are XORed one at a time to the least significant bit of the bytes in the two top rows, i.e., the first bytes in the topmost row are XORed with t_0, t_1, t_2, t_3, respectively. The bytes in the second row are XORed with t_4, t_5, t_6, and t_7, respectively:

$$\begin{bmatrix} t_0 & t_1 & t_2 & t_3 \\ t_4 & t_5 & t_6 & t_7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

J ensures at least four active bits for any nonzero input tweak difference. In combination with the XOR into the second row and the ShiftRows operation after a tweak addition, a non-zero tweak difference will affect at least three pairwise distinct columns of the state. Moreover, TweAES injects the tweak after Round 2, 4, 6, and 8. Our instance will differ by using the tweak injection after every round except the final one.

6.2 Definition of TweAES'

Let $\mathcal{K} = \mathcal{B} = \mathbb{F}_{2^8}^{4\times 4}$ be key and block space and $\mathcal{T} = \{0,1\}^4$ the tweak space, respectively. Both state and key are arranged in a 4×4-byte matrix in the AES, indexed as

$$\begin{bmatrix} 0 & 4 & 8 & 12 \\ 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \end{bmatrix} .$$

We employ the AES round function $R[K^i]$ which is composed of SubBytes, ShiftRows, MixColumns, and AddRoundKey$[K^i]$. SubBytes applies the AES S-box to each byte in the state. ShiftRows rotates Row i by i byte positions to the left, for $i \in \{0,1,2,3\}$. MixColumns multiplies each column by the AES MDS matrix. AddRoundKey$[K^i]$ XORs

the round key K^i to the state. Additionally, after almost every call to AddRoundKey$[K^i]$, and AddRoundTweak$[T]$ adds an expanded tweak T is added to the state. The rounds i are indexed as 1 through r.
We denote the states as S^0 through S^r, where S^0 is the plaintext and S^i the state after the i-th round, i.e., S^r is identical to the ciphertext. Before the first round, AddRoundKey$[K^0]$ and AddRoundTweak$[T^0]$ are performed, i.e., K^0 and T^0 are XORed to the S^0. We write S^i_{SB}, S^i_{SR}, S^i_{MC}, S^i_{AK}, S^i for the state directly after the application of SubBytes, ShiftRows, MixColumns, AddRoundKey and AddRoundTweak of the i-th round. The cells are indexed as usually in the AES. Compared to the AES, the final round of TweAES' employs the MixColumns operation. TweAES' adopts the key schedule of AES-128 and iterates it further (as did ForkAES) when more round keys are necessary. TweAES' is a variant of TweAES$[4, 8, 8, 2]$ that takes and expands a four-bit tweak $T = (t_0, t_1, t_2, t_3)$ to $T = (t_0, \ldots, t_7)$ and adds the bits to the least significant bits of the topmost two rows before the first and after every round except the final round, as in TweAES.

6.3 Aspects of TweAES' in ForkCENC-AES and ForkEDM-AES

We need to consider two aspects when using our instantiation in the constructions: numbers of rounds, derivation of round keys, and branch constants. We define r_t and r_b for the number of rounds in the top and bottom permutations, respectively, in our forked constructions. We fix $r_t = 5$ and $r_b = 7$ for TweAES' as a result of cryptanalysis and our requirements. To obtain more round keys, we use the keys K^0, ..., $K^{r_t+r_b-1}$ from the (further-iterated) AES key schedule, with K^0, ..., K^{r_t} in the top permutation and K^{r_t+1}, ..., $K^{r_t+r_b-1}$ in the bottom-permutation calls. We deviate from the simplest composition of two independent instances of TweAES' in our constructions as follows:

- We can omit the tweaks in the top permutation since it does not increase security. Since we use an all-zero tweak in the top permutation, we spare any operations there.

- We avoid iterating the key schedule further on than r rounds compared to the ForkCipher proposal [4]. The overhead of the schedule does not scale well beyond two branches. Instead, we use the same round keys in the bottom-permutation calls. At the start, each bottom-permutation call associated with Tweak T^i XORs a branch-dependent constant BC^i to the state, for $i \in \{1, \ldots, w\}$. The branch constants provide efficient differential diffusion through the bottom-call permutations; thus, we do not need additional tweaks at the beginning of them.

- Since K^{r_t} protects the state after the top-permutation call, we avoid a key addition before the first round of each bottom-permutation call.

- Since ForkCENC-AES employs a sum of branches at the bottom, we can omit the final round-key addition since the keys would cancel.

- Similarly, we can omit the tweak addition at the end of all bottom permutations since it does not affect the security.

We define the branch constants that employ the sequential digits of π encoded as integers as nothing-up-my-sleeve branch constants. For completeness, the 15 constants are listed in Table 2. Note that we limit our interest to constructions with at most 15 bottom branches. In use cases, where more branches are required, the tweak space can be easily extended to take longer tweaks, e.g., 16-bit tweak inputs as in AES$[16, 32, 8, 2]$ [28, 30]. Though, we do not propose such instances.

Table 2: Branch constants for Branches i.

i	Constant	i	Constant
0	0x9d7b8175 f0fec5b2 0ac020e6 4c708406	8	0xee65d4b9 ca8fdbec e97f86e6 f1634dab
1	0x17f7082f a46b0f64 6ba0f388 e1b4668b	9	0x337e03ad 4f402a5b 64cdb7d4 84bf301c
2	0x1491029f 609d02cf 9884f253 2dde0234	10	0x0098f68d 2e8b0269 bf231794 b90bccb2
3	0x794f5bfd afbcf3bb 084f7b2e e6ead60e	11	0x8a2d9d5c c89eaa4a 72556fde a67804fa
4	0x447039be 1ccdee79 8b447248 cbb0cfcb	12	0xd49f1229 2e4ffa0e 122a776b 2b9fb4df
5	0x7b058a2b ed35538d b732906e eecdea7e	13	0xee126abb ae11d632 36a249f4 4403a11e
6	0x1bef4fda 612741e2 d07c2e5e 438fc267	14	0xa6eca89c c900965f 8400054b 884904af
7	0x3b0bc71f e2fd5f67 07cccaaf b0d92429	15	0xec93e527 e3c7a278 4f9c199d d85e0221

7 Cryptanalysis of TweAES' in ForkCENC and ForkEDM

In this section, we provide a preliminary discussion of the security of TweAES' in ForkCENC. We call those instances ForkCENC-AES and ForkEDM-AES, respectively.

7.1 Rationale

In Appendix D, we summarize lessons from the community's cryptanalysis efforts on earlier AES-round-based primitives. We further summarize a collection of the best existing key-recovery attacks and distinguishers in Table 5 there. To conclude, our instantiation

- uses five rounds in each top-permutation call against differential attacks.
- uses seven rounds in each bottom-permutation call against rectangle, mixture, and impossible-differential attacks.
- adopts the branch-constant approach of ForkAES as an effective means to make inter-branch differentials harder to exploit.
- uses the tweak injection after each round to ensure sufficiently many active S-boxes for two- to four-round differentials against rectangles and to destroy mixture properties.
- injects the tweak directly at the start of the bottom-permutation calls to increase security.

In the following, we give brief details on the individual types of attacks we consider. Throughout this section, we consider three settings for attacks with tuples, such as pairs or integral sets, from

- **Setting (1):** different bottom-permutation branches (distinct branch indices i and j, with $i, j \in \{1..15\}$ of the same chunk.
- **Setting (2):** equal branches i from different chunks, i.e. different input messages.
- **Setting (3):** different branches i and j from different chunks.

7.2 Differential Bounds

Among top-permutation calls, it is well-known that four rounds of AES activate at least 25 S-boxes. Thus, any differential characteristic has a probability of at most 2^{-150} on average. Without concerning tweaks, the full top and bottom permutation are at least as secure as the full-round AES. For the bottom-permutation calls, we have to lower bound the number of active S-boxes in differentials.
For Setting (1), the differences result from the branch constant differences and the tweak differences. The branch constants ensure almost fully active differences and hence at least 28 active S-boxes in five rounds and 25 without branch constants, as given in Table 3.

Table 3: Lower bounds on the number of active S-boxes in small-tweak AES-based TBCs with difference only in the tweak.

Constr.	1	2	3	4	5	6	7	8	9	10
Active plaintext or tweak										
TweAES'	**0**	**4**	**8**	**14**	**18**	**22**	**26**	**30**	**34**	**38**
TweAES [30]	0	0	4	15	19	20	27	30	34	40
Kiasu-BC [66]	0	1	4	8	18	22	25	28	33	38
Active tweak										
TweAES'	**4**	**11**	**18**	**21**	**25**	**29**	**34**	**38**	**42**	**46**
TweAES [30]	4	15	20	20	27	30	34	40	44	50
Kiasu-BC [66]	1	4	17	23	25	26	29	37	44	50
From branch constants										
TweAES'	**14**	**15**	**19**	**23**	**28**	**32**	**36**	**40**	**44**	**48**
TweAES [30]	14	20	21	21	25	35	39	45	48	51
Kiasu-BC [66]	14	15	18	20	24	32	36	40	43	48

For Setting (2), between the i-th branch of two distinct chunks, we can assume that a certain non-zero difference to happen with probability roughly at most 2^{-128} after the top permutation. This setting can be reduced to tracing differential trails through the sequence of the top- and a bottom-permutation branch of the cipher, which implies tracing them through $(r_t + r_b)$-round AES. Our construction is likely to provide more security since partial decryptions are unavailable.
Moreover, between different branches from different chunks, we can assume any non-zero difference to happen with probability roughly at most 2^{-128} after the top permutation. Then, the bottom-permutation calls represent a conditional differential. We have at least 36 active S-boxes through seven rounds of the bottom permutation, which should prevent such differential characteristics.

7.3 Linear Attacks

Concerning standard linear attacks, similar results as for the differential analysis can be applied against linear attacks. Since the tweak schedule is linear, the tweak does not introduce additional linear trails compared to a non-tweaked cipher [71]. For the AES, four rounds are known to activate at least 25 S-boxes. Thus, we expect that resistance against linear attacks for the sequence of top- and any bottom-permutation branches should suffice to thwart them. Moreover, we expect the sum in ForkEDM and ForkCENC to render attacks even harder than for the plain AES since makes key recovery harder at the ciphertext side.

7.4 Integral Cryptanalysis

Regarding integral attacks, we have to consider again the three settings. We believe that Cases (2) and (3) are hard to exploit and thwarted by the presence of the top permutation: there are no integral attacks over five-round AES [91] without related tweaks.
Related-tweak differences can generate a balanced property over only two rounds of TweAES [29] before it is destroyed. We can formulate a similar statement for the bottom permutations when we inject tweaks in every round. If an attacker wants to exploit a blank round with zero difference, it has to hit the difference induced by different branch constants, which renders such attacks harder. Moreover, key guessing at the ciphertext side is thwarted or made harder due to the sum. Such attacks seem to be more of a threat to TweAES.

7.5 Impossible-differential and Zero-correlation Distinguishers

Impossible-differential attacks exploit differentials with probability zero. Zero-correlation distinguishers represent a corresponding attack in the linear setting, i.e., they are linear

attacks with correlation zero. As shown by Derbez et al. on AES-PRF [45], both types can be strong on EDM and AES-PRF and cover up to four rounds. There, an attacker can exploit a zero difference or a zero-correlation mask cannot occur in the ciphertext output. The upper bounds on the length of impossible differentials for the AES and TweAES are also upper bounds for the top and bottom parts of our constructions. For the AES, there exist impossible-differential distinguishers on up to four rounds and no impossible differentials for five-round AES even when taking the key schedule into account [25]. Hence, longer distinguishers need to exploit tweak differences. For TweAES, impossible differentials cover at most six rounds but only when the former or latter two rounds are inactive. In Setting (1), we expect the branch numbers to thwart long distinguishers. For ForkCENC-AES and ForkEDM-AES, we can derive that impossible-differential distinguishers cover at most five rounds in Setting (1). Moreover, Setting (2) and (3) correspond to an attack on 12-round AES; Setting (3) with an additional fixed tweak in the bottom permutation, which should not provide the adversary with better distinguishers.
Zero-correlation distinguishers are expected to hold also for only up to four rounds of the AES. Here, the presence of tweaks should not allow longer distinguishers [71]. Derbez et al. [45] exploited a four-round zero-correlation distinguisher on EDM. Given a zero-correlation trail with mask $\alpha \to \alpha$ through the bottom permutation, the adversary can evaluate α for all ciphertexts. We expect that there exist no zero-correlation distinguishers that map onto themselves through more than four rounds of the bottom permutation in ForkCENC-AES and ForkEDM-AES.

7.6 Meet-in-the-Middle (MitM) Distinguishers

Demirci-Selçuk-(DS)-MitM attacks [42, 43] trace an input set of partial sets to their corresponding output sets through parts of a cipher. The adversary guesses parts of those internal states and builds a table of all possible transitions from a partial start state to a partial end state through the cipher. If the number of computations of this offline step is significantly smaller than that of an exhaustive search, it can then use the table in an attack. For keyed ciphers, the adversary guesses the keys to obtain knowledge about the partial start and end states. It traces related texts and looks up if the sequence obtained for the current key guess is among the possible transitions. If it is not, the adversary can discard the current key guess.

In Settings (2) and (3), we can employ the same heuristic argument from Derbez et al. [45] that EDM and its dual seem at least as resistant as the sequence of top and bottom permutation. For both cases, an attacker would have to either work its way through 12 rounds of AES or have to predict a certain truncated difference in the middle. In those cases, we assume that 12-round TweAES' offers sufficient resistance against DS-MitM distinguishers. In particular, Sun showed in [90], that the length of such distinguishers for ciphers with n-bit state and key is limited by at most twice the number of operations necessary for full diffusion, taking the maximum number of operations in for- or backward direction. Given that the AES achieves full diffusion after two rounds and designs with short tweak achieves full diffusion after three rounds, the length of DS-MitM distinguishers is limited to at most six rounds in our construction. We emphasize that this is already a conservative estimate. Works on Kiasu-BC, whose larger and simpler tweak injection provides the adversary with strictly more freedom than our proposal, managed five-round distinguishers and eight-round attacks yet [73, 93]).
Differentials between different branches of the same chunk in Setting (1) may pose a considerable threat to constructions such as ours. The state-of-the-art DS-MitM attacks exploit the differential-enumeration technique [48]: they wait for a pair of texts that follow a certain differential through the distinguishing part of the cipher. The advantage over value-based guessing is that a pair of texts allows deriving the state in a round without the need for guessing the value inside the middle of the differential trail.

In contrast to DS-MitM attacks on block ciphers, tweak-induced differentials in Setting (1) for our construction prohibit guessing the key until the forking point and derive further texts from new plaintexts. However, the other branches allow us to derive a few further texts automatically without the need to guess keys from the start. What remains is to consider the composition of a distinguisher and key guessing from the end. We consider distinguishers to cover at most six rounds. We can further adopt an argument from TweAES [29]: The large weight of differences induced by expanded related tweaks and the branch constants in our constructions prohibits sparse trails and limits the lengths of distinguishers to at most five rounds. The additional sum at the end further strengthens the resistance since it blinds the output values and renders key recovery impossible or highly expensive. Therefore, we expect seven rounds in the bottom-permutation calls to thwart such attacks.

7.7 Differential-linear and Rectangle Distinguishers

Differential-linear distinguishers combine a short differential with a short linear hull and a middle phase. We consider them inapplicable to TweAES' in Settings (2) and (3) due to the strong diffusion properties and the high complexity of $p^2 r \epsilon^4$ – given a differential with probability p, a middle-phase transition with probability r, and a linear approximation with correlation ϵ – no such attacks are known for even a few rounds of the AES.
Regarding related tweaks from the same chunks in Setting (1), we see that key guessing at the ciphertext side is made substantially harder by the sum at the end and attacks will have to exploit at least six-round distinguishers. Since any combination of $r = 4$ rounds activate at least 22 and $r = 5$ rounds activate at least 25 active S-boxes, we assume that no distinguishers over six or seven rounds (plus a middle round for the transition phase) exist in the bottom part.
A similar argument as for differential-linear distinguishers follows for chosen-plaintext related-tweak rectangle attacks. Related-tweak differences in TweAES' and TweAES activate more S-boxes than for Kiasu-BC or TNT-AES [12]. Distinguishers over seven rounds seem unlikely. We note that Chakraborti et al. showed a longer distinguisher on up to seven rounds of the TweAES [29], which exploited two consecutive inactive rounds in both top and bottom difference. Since TweAES injects the tweak difference in every second round, this distinguisher is not directly applicable to our construction. Furthermore, the tool by Yang et al. refined the study of multiple consecutive S-boxes in boomerangs and found that the seven-round distinguisher on TweAES had probability zero [94].

7.8 Mixtures

Mixture-differential attacks [53, 54] are a variant of conditional differentials. Given a pair with a certain input difference and at least two different components (e.g. bytes) therein, further pairs with the same input difference can be constructed by mixing the components from the first pair. If the first pair follows a differential, a distinguisher exploits that the further pairs follow the differential with higher probability than random. The attack represents a potential threat when given AES-based permutations with adversary-controllable inputs or when tweak differences can be used for creating mixed pairs.
We found mixture-differential distinguishers on ForkCENC[w] if it would use TweAES-4, that is, four-round TweAES with injecting tweaks after every two rounds. Preventing such attacks was also a factor for using tweak injections in every round. We argue that ForkAES [5] as well as our proposals ForkCENC-AES, and ForkEDM-AES thwart them due to branch constants. ForkAES also employs the effective but costly countermeasure of using independent round keys for each branch. In our proposals, we choose the number of bottom rounds more conservatively and expect that the sum in our proposals thwarts key recovery at the ciphertext side.

Table 4: Performance in cycles per byte for our instantiations with selected number of branches w and up to 16 chunks with AES-NI, SSE4.1, and AVX2.

(a) ForkCENC-AES-5-7[w].

	#Chunks of $16w$ bytes						
w	1	2	3	4	8	12	16
i5-6300U							
4	1.15	0.89	0.93	0.78	0.76	0.75	0.74
5	0.96	0.80	0.81	0.71	0.69	0.68	0.68
8	0.79	0.70	0.68	0.59	0.58	0.58	0.58
15	0.67	0.59	0.58	0.54	0.53	0.53	0.52
Intel i5-1240P							
4	0.91	0.83	0.88	0.81	0.64	0.62	0.62
5	0.81	0.81	0.80	0.70	0.55	0.54	0.54
8	0.82	0.68	0.60	0.49	0.45	0.45	0.45
15	0.64	0.49	0.45	0.43	0.41	0.40	0.40

(b) ForkEDM-AES-5-7[w].

	#Chunks of $16w$ bytes						
w	1	2	3	4	8	12	16
Intel i5-1240P							
4	0.71	0.75	0.66	0.62	0.49	0.47	0.47
5	0.62	0.67	0.63	0.54	0.45	0.44	0.43
8	0.73	0.65	0.54	0.46	0.44	0.43	0.43
15	0.58	0.47	0.42	0.42	0.41	0.41	0.39

7.9 Reflection, Yoyo, and Boomerang Attacks

The feed-forward in ForkCENC-AES renders backward queries difficult or impossible. As a result, boomerang or yoyo key-recovery attacks that need both plain- and ciphertext queries seem unlikely to be applicable – at least, they seem costlier than exhaustive search. Compared to ForkCiphers, the feed-forward would also allow discarding chosen-ciphertext reflection differentials, which were the most effective attacks on ForkAES.

7.10 Others

We assume that techniques like slide attacks [21], rotational attacks [70], or internal-difference attacks [65, 86] are prevented by adopting the AES key schedule and its constants. Multiple-of-n attacks [57] exploit probability-1 conditional differences of related pairs after a few rounds. Grassi et al. showed that for a set of all texts in a diagonal, if a pair has a specific set of inactive anti-diagonals after almost five-round AES (without the final MixColumns operation), then, the number of pairs with this property in the set of texts is a multiple of eight. Though, no such properties are known over more than five rounds [24] AES. Plus, we see that the sum operation in our constructions can thwart key-recovery attacks such as the six-round attack by Bar-On et al. [9] on six-round AES-128.

8 Implementation

We implemented ForkCENC-AES in C using SSE4.2 and AES-Native instructions and benchmarked it on an Intel Skylake i5-6300U (6th-generation) and an Intel Alder Lake i5-1240P (12-th generation) both with adaptive power policy, TurboBoost and HyperThreading disabled. Table 4 lists the benchmark results for selected numbers of branches per chunk, where the message lengths are given in multiples of $16w$-byte chunks for varying numbers of words w. The measurements are the medians of 1 024 runs each after 1 024 warm-up runs.

The senior (pre-AVX-512) Skylake CPU allows for simpler comparison, where the usual optimum for 10-round AES-128 was simply one cycle per round, i.e. $10/16 = 0.625$ cycles per bytes (c/b) without input-specific optimizations. There remains a tiny gap to the theoretical optimum of about 0.49 c/b for ForkCENC-AES-5-7[15] on Skylake that could be further closed. Though, we note that our implementation is still conservative. We did not include the counter-specific optimizations by Park and Lee yet [81] that exploit unchanged parts of the input in counter mode and could spare effectively the equivalent of three AES

rounds. We could employ a similar optimization for the nonce in our proposals, but, to be fair, we would save that amount only in the top permutation calls, which are computed only for $1/w$ of the blocks.
We note that Intel removed its performance-boosting AVX-512 instructions from most of its off-the-shelf platforms of the 12th-generation CPUs. We provide the results for ForkCENC and ForkEDM both instantiated with TweAES'-5-7 on this platform as an example of a recent platform at the time of writing the paper for a broader overview to the reader.

9 Conclusion

This work combines the knowledge from designing optimally secure fixed-output-length PRFs and the generalized Mirror Theory. We proposed a spectrum view of constructions from forked constructions that cover (1) output-length vs. PRF security, (2) full vs reduced primitives, and (3) fixed- vs. variable-length outputs. We forked and reduced the sum of permutations and the Summation-Truncation Hybrid, and extended them to variable-output-length constructions. Given the insights about attacks on the growing corpus of primitives from reduced-round AES, e.g. AES-PRF, ForkAES, TweAES, we could propose efficient instantiations. Our instantiation based on $5 + 7$ AES rounds serves as an initial proposal of what is possible at least. We can envision the search for even more secure and more efficient instantiations as interesting future works. We also motivate third-party cryptanalysis to further increase the understanding of such settings. Moreover, there probably exist possible more lightweight instantiations from GIFT or SKINNY. We acknowledge the parallel and independent work by Andreeva et al. [3] who had focused on ForkEDMD, proved its security, derived a highly performant instantiation and highly secure deterministic AE schemes.

Acknowledgments. We are highly thankful to the reviewers and editors of ToSC 2022(4) for their very fruitful comments and suggestions.

References

[1] Najwa Aaraj, Emanuele Bellini, Ravindra Jejurikar, Marc Manzano, Raghvendra Rohit, and Eugenio Salazar. Farasha: A Provable Permutation-based Parallelizable PRF. *IACR Cryptol. ePrint Arch.*, page 1150, 2022.

[2] Elena Andreeva, Amit Singh Bhati, Bart Preneel, and Damian Vizár. 1, 2, 3, Fork: Counter Mode Variants based on a Generalized Forkcipher. *IACR Trans. Symmetric Cryptol.*, 2021(3):1–35, 2021.

[3] Elena Andreeva, Benoit Cogliati, Virginie Lallemand, Marine Minier, Antoon Purnal, and Arnab Roy. Masked Iterate-Fork-Iterate: A new Design Paradigm for Tweakable Expanding Pseudorandom Function. Cryptology ePrint Archive, Paper 2022/1534, 2022.

[4] Elena Andreeva, Virginie Lallemand, Antoon Purnal, Reza Reyhanitabar, Arnab Roy, and Damian Vizár. Forkcipher: A New Primitive for Authenticated Encryption of Very Short Messages. In Steven D. Galbraith and Shiho Moriai, editors, *ASIACRYPT II*, volume 11922 of *LNCS*, pages 153–182. Springer, 2019.

[5] Elena Andreeva, Reza Reyhanitabar, Kerem Varici, and Damian Vizár. ForkAES: a Tweakable Forkcipher. Cryptology ePrint Archive, Report 2018/916, 2018.

[6] Jean-Philippe Aumasson. Too Much Crypto. *IACR Cryptol. ePrint Arch.*, 2019:1492, 2019.

[7] Subhadeep Banik, Jannis Bossert, Amit Jana, Eik List, Stefan Lucks, Willi Meier, Mostafizar Rahman, Dhiman Saha, and Yu Sasaki. Cryptanalysis of ForkAES. In Robert H. Deng, Valérie Gauthier-Umaña, Martín Ochoa, and Moti Yung, editors, *ACNS*, volume 11464 of *LNCS*, pages 43–63. Springer, 2019.

[8] Zhenzhen Bao, Jian Guo, and Eik List. Extended Truncated-differential Distinguishers on Round-reduced AES. *IACR Trans. Symmetric Cryptol.*, 2020(3):197–261, 2020.

[9] Achiya Bar-On, Orr Dunkelman, Nathan Keller, Eyal Ronen, and Adi Shamir. Improved Key Recovery Attacks on Reduced-Round AES with Practical Data and Memory Complexities. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO II*, volume 10992 of *LNCS*, pages 185–212. Springer, 2018.

[10] Navid Ghaedi Bardeh and Sondre Rønjom. The Exchange Attack: How to Distinguish Six Rounds of AES with $2^{88.2}$ Chosen Plaintexts. In Steven D. Galbraith and Shiho Moriai, editors, *ASIACRYPT III*, volume 11923 of *LNCS*, pages 347–370. Springer, 2019.

[11] Augustin Bariant, Nicolas David, and Gaëtan Leurent. Cryptanalysis of Forkciphers. *IACR Trans. Symmetric Cryptol.*, 2020(1):233–265, 2020.

[12] Augustin Bariant and Gaëtan Leurent. Truncated Boomerang Attacks and Application to AES-based Ciphers. Cryptology ePrint Archive, Paper 2022/701, 2022.

[13] Christof Beierle, Jérémy Jean, Stefan Kölbl, Gregor Leander, Amir Moradi, Thomas Peyrin, Yu Sasaki, Pascal Sasdrich, and Siang Meng Sim. The SKINNY Family of Block Ciphers and Its Low-Latency Variant MANTIS. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO II*, volume 9815 of *LNCS*, pages 123–153. Springer, 2016. Full version at https://eprint.iacr.org/2016/660.

[14] M. Bellare and R. Impagliazzo. A tool for obtaining tighter security analyses of pseudorandom function based constructions, with applications to prp to prf conversion. Cryptology ePrint Archive, Report 1999/024, 1999. http://eprint.iacr.org/1999/024.

[15] Mihir Bellare, Joe Kilian, and Phillip Rogaway. The Security of Cipher Block Chaining. In Yvo Desmedt, editor, *CRYPTO*, volume 839 of *LNCS*, pages 341–358. Springer, 1994.

[16] Mihir Bellare, Ted Krovetz, and Phillip Rogaway. Luby-Rackoff Backwards: Increasing Security by Making Block Ciphers Non-invertible. In Kaisa Nyberg, editor, *EUROCRYPT*, volume 1403 of *LNCS*, pages 266–280. Springer, 1998.

[17] Mihir Bellare and Phillip Rogaway. The Security of Triple Encryption and a Framework for Code-Based Game-Playing Proofs. In Serge Vaudenay, editor, *EUROCRYPT*, volume 4004 of *LNCS*, pages 409–426. Springer, 2006.

[18] Guido Bertoni, Joan Daemen, Seth Hoffert, Michaël Peeters, Gilles Van Assche, and Ronny Van Keer. Farfalle: parallel permutation-based cryptography. *IACR Trans. Symmetric Cryptol.*, 2017(4):1–38, 2017.

[19] Arghya Bhattacharjee, Avijit Dutta, Eik List, and Mridul Nandi. CENCPP*: beyond-birthday-secure encryption from public permutations. *Des. Codes Cryptogr.*, 90(6):1381–1425, 2022.

[20] Ritam Bhaumik, Nilanjan Datta, Avijit Dutta, Nicky Mouha, and Mridul Nandi. The Iterated Random Function Problem. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT II*, volume 10625 of *LNCS*, pages 667–697. Springer, 2017.

[21] Alex Biryukov and David A. Wagner. Slide Attacks. In Lars R. Knudsen, editor, *FSE*, volume 1636 of *LNCS*, pages 245–259. Springer, 1999.

[22] Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *CHES*, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.

[23] Julia Borghoff, Anne Canteaut, Tim Güneysu, Elif Bilge Kavun, Miroslav Knezevic, Lars R. Knudsen, Gregor Leander, Ventzislav Nikov, Christof Paar, Christian Rechberger, Peter Rombouts, Søren S. Thomsen, and Tolga Yalçin. PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT*, volume 7658 of *LNCS*, pages 208–225. Springer, 2012.

[24] Christina Boura, Anne Canteaut, and Daniel Coggia. A General Proof Framework for Recent AES Distinguishers. *IACR Trans. Symmetric Cryptol.*, 2019(1):170–191, 2019.

[25] Christina Boura and Daniel Coggia. Efficient MILP Modelings for Sboxes and Linear Layers of SPN ciphers. *IACR Trans. Symmetric Cryptol.*, 2020(3):327–361, 2020.

[26] Colin Chaigneau and Henri Gilbert. Is AEZ v4.1 Sufficiently Resilient Against Key-Recovery Attacks? *IACR Trans. Symmetric Cryptol.*, 2016(1):114–133, 2016.

[27] Avik Chakraborti, Nilanjan Datta, Ashwin Jha, Cuauhtemoc Mancillas Lopez, Mridul Nandi, and Yu Sasaki. ESTATE Authenticated Encryption Mode: Hardware Benchmarking and Security Analysis. *National Institute of Standards and Technology (NIST)*, 2019.

[28] Avik Chakraborti, Nilanjan Datta, Ashwin Jha, Cuauhtemoc Mancillas-López, Mridul Nandi, and Yu Sasaki. Elastic-Tweak: A Framework for Short Tweak Tweakable Block Cipher. *IACR Cryptol. ePrint Arch.*, 2019:440, 2019.

[29] Avik Chakraborti, Nilanjan Datta, Ashwin Jha, Cuauhtemoc Mancillas-López, Mridul Nandi, and Yu Sasaki. ESTATE: A Lightweight and Low Energy Authenticated Encryption Mode. *IACR Trans. Symmetric Cryptol.*, 2020(S1):350–389, 2020.

[30] Avik Chakraborti, Nilanjan Datta, Ashwin Jha, Cuauhtemoc Mancillas-López, Mridul Nandi, and Yu Sasaki. Elastic-Tweak: A Framework for Short Tweak Tweakable Block Cipher. In Avishek Adhikari, Ralf Küsters, and Bart Preneel, editors, *INDOCRYPT*, volume 13143 of *LNCS*, pages 114–137. Springer, 2021.

[31] Avik Chakraborti, Mridul Nandi, Suprita Talnikar, and Kan Yasuda. On the Composition of Single-Keyed Tweakable Even-Mansour for Achieving BBB Security. *IACR Trans. Symmetric Cryptol.*, 2020(2):1–39, 2020.

[32] Donghoon Chang and Mridul Nandi. A short proof of the PRP/PRF switching lemma. *IACR Cryptol. ePrint Arch.*, page 78, 2008.

[33] Yu Long Chen, Eran Lambooij, and Bart Mennink. How to Build Pseudorandom Functions from Public Random Permutations. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO I*, volume 11692 of *LNCS*, pages 266–293. Springer, 2019.

[34] Yu Long Chen, Bart Mennink, and Bart Preneel. Categorization of Faulty Nonce Misuse Resistant Message Authentication. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT III*, volume 13092 of *LNCS*, pages 520–550. Springer, 2021.

[35] Wonseok Choi, ByeongHak Lee, Jooyoung Lee, and Yeongmin Lee. Toward a Fully Secure Authenticated Encryption Scheme from a Pseudorandom Permutation. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT III*, volume 13092 of *LNCS*, pages 407–434. Springer, 2021.

[36] Benoît Cogliati and Jacques Patarin. Mirror theory: A simple proof of the $p_i + p_j$ theorem with $xi\_max = 2$. *IACR Cryptol. ePrint Arch.*, page 734, 2020.

[37] Benoît Cogliati and Yannick Seurin. EWCDM: An Efficient, Beyond-Birthday Secure, Nonce-Misuse Resistant MAC. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO I*, volume 9814 of *LNCS*, pages 121–149. Springer, 2016.

[38] Benoît Cogliati and Yannick Seurin. Analysis of the single-permutation encrypted Davies-Meyer construction. *Des. Codes Cryptogr.*, 86(12):2703–2723, 2018.

[39] Benoît Cogliati, Avijit Dutta, Mridul Nandi, Jacques Patarin, and Abishanka Saha. Proof of mirror theory for any $\xi_{\max}$. Cryptology ePrint Archive, Paper 2022/686, 2022.

[40] Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The Block Cipher Square. In Eli Biham, editor, *FSE*, volume 1267 of *LNCS*, pages 149–165. Springer, 1997.

[41] Wei Dai, Viet Tung Hoang, and Stefano Tessaro. Information-Theoretic Indistinguishability via the Chi-Squared Method. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO Part III*, volume 10403 of *LNCS*, pages 497–523. Springer, 2017. Full version at http://eprint.iacr.org/2017/537, version 20170616:190106.

[42] Hüseyin Demirci and Ali Aydin Selçuk. A Meet-in-the-Middle Attack on 8-Round AES. In Kaisa Nyberg, editor, *FSE*, volume 5086 of *LNCS*, pages 116–126. Springer, 2008.

[43] Patrick Derbez and Pierre-Alain Fouque. Exhausting Demirci-Selçuk Meet-in-the-Middle Attacks Against Reduced-Round AES. In Shiho Moriai, editor, *FSE*, volume 8424 of *LNCS*, pages 541–560. Springer, 2013.

[44] Patrick Derbez, Pierre-Alain Fouque, and Jérémy Jean. Improved Key Recovery Attacks on Reduced-Round AES in the Single-Key Setting. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT*, volume 7881 of *LNCS*, pages 371–387. Springer, 2013.

[45] Patrick Derbez, Tetsu Iwata, Ling Sun, Siwei Sun, Yosuke Todo, Haoyang Wang, and Meiqin Wang. Cryptanalysis of AES-PRF and Its Dual. *IACR Trans. Symmetric Cryptol.*, 2018(2):161–191, 2018.

[46] Christoph Dobraunig, Maria Eichlseder, and Florian Mendel. Square Attack on 7-Round Kiasu-BC. In Mark Manulis, Ahmad-Reza Sadeghi, and Steve A. Schneider, editors, *ACNS*, volume 9696 of *LNCS*, pages 500–517. Springer, 2016.

[47] Christoph Dobraunig and Eik List. Impossible-Differential and Boomerang Cryptanalysis of Round-Reduced Kiasu-BC. In Helena Handschuh, editor, *CT-RSA*, volume 10159 of *LNCS*, pages 207–222. Springer, 2017.

[48] Orr Dunkelman, Nathan Keller, and Adi Shamir. Improved Single-Key Attacks on 8-Round AES-192 and AES-256. In Masayuki Abe, editor, *ASIACRYPT*, volume 6477 of *Lecture Notes in Computer Science*, pages 158–176. Springer, 2010.

[49] Avijit Dutta, Mridul Nandi, and Abishanka Saha. Proof of mirror theory for $\xi_{\max} = 2$. *IACR Cryptol. ePrint Arch.*, page 669, 2020.

[50] Avijit Dutta, Mridul Nandi, and Suprita Talnikar. Permutation Based EDM: An Inverse Free BBB Secure PRF. *IACR Trans. Symmetric Cryptol.*, 2021(2):31–70, 2021.

[51] Horst Feistel. Cryptography and computer privacy. *Scientific American*, 228(5):15–23, 1973.

[52] Shoni Gilboa and Shay Gueron. The Advantage of Truncated Permutations. *CoRR*, abs/1610.02518, 2016.

[53] Lorenzo Grassi. MixColumns Properties and Attacks on (Round-Reduced) AES with a Single Secret S-Box. In Nigel P. Smart, editor, *CT-RSA*, volume 10808 of *LNCS*, pages 243–263. Springer, 2018.

[54] Lorenzo Grassi. Mixture Differential Cryptanalysis: a New Approach to Distinguishers and Attacks on round-reduced AES. *IACR Transactions on Symmetric Cryptology*, 2018(2):133–160, 2018.

[55] Lorenzo Grassi, Morten Øygarden, Markus Schofnegger, and Roman Walch. From Farfalle to Megafono via Ciminion: The PRF Hydra for MPC Applications. *IACR Cryptol. ePrint Arch.*, page 342, March 14, 11:54:47 2022. version 20220314:115447.

[56] Lorenzo Grassi, Morten Øygarden, Markus Schofnegger, and Roman Walch. From Farfalle to Megafono via Ciminion: The PRF Hydra for MPC Applications. *IACR Cryptol. ePrint Arch.*, page 342, 2022.

[57] Lorenzo Grassi, Christian Rechberger, and Sondre Rønjom. A New Structural-Differential Property of 5-Round AES. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *EUROCRYPT II*, volume 10211 of *LNCS*, pages 289–317, 2017.

[58] Aldo Gunsing and Bart Mennink. The Summation-Truncation Hybrid: Reusing Discarded Bits for Free. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO I*, volume 12170 of *LNCS*, pages 187–217. Springer, 2020.

[59] Chun Guo, Yaobin Shen, Lei Wang, and Dawu Gu. Beyond-birthday secure domain-preserving PRFs from a single permutation. *Des. Codes Cryptogr.*, 87(6):1297–1322, 2019.

[60] Chris Hall, David A. Wagner, John Kelsey, and Bruce Schneier. Building prfs from prps. In *CRYPTO 1998, Proceedings*, pages 370–389, 1998.

[61] Viet Tung Hoang, Ted Krovetz, and Phillip Rogaway. Robust Authenticated-Encryption AEZ and the Problem That It Solves. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT (1)*, volume 9056 of *LNCS*, pages 15–44. Springer, 2015.

[62] Tetsu Iwata. New Blockcipher Modes of Operation with Beyond the Birthday Bound Security. In Matthew J. B. Robshaw, editor, *FSE*, volume 4047 of *LNCS*, pages 310–327. Springer, 2006.

[63] Tetsu Iwata, Bart Mennink, and Damian Vizár. CENC is Optimally Secure. *IACR Cryptol. ePrint Arch.*, 2016:1087, 2016.

[64] Tetsu Iwata and Kazuhiko Minematsu. Stronger Security Variants of GCM-SIV. *IACR Trans. Symmetric Cryptol.*, 2016(1):134–157, 2016.

[65] Jérémy Jean and Ivica Nikolic. Internal Differential Boomerangs: Practical Analysis of the Round-Reduced Keccak-f Permutation. In Gregor Leander, editor, *FSE*, volume 9054 of *LNCS*, pages 537–556. Springer, 2015.

[66] Jérémy Jean, Ivica Nikolić, and Thomas Peyrin. Kiasu v1.1. http://competitions.cr.yp.to/caesar-submissions.html, Mar 16 2014. First-round submission to the CAESAR competition.

[67] Jérémy Jean, Ivica Nikolic, and Thomas Peyrin. Tweaks and Keys for Block Ciphers: The TWEAKEY Framework. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT II*, volume 8874 of *LNCS*, pages 274–288. Springer, 2014.

[68] Zilong Jiang and Chenhui Jin. Multiple Impossible Differential Attacks for ForkAES. *Security and Communication Networks*, 2022:1–11, 2022.

[69] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography, Second Edition*. CRC Press, 2014.

[70] Dmitry Khovratovich and Ivica Nikolic. Rotational Cryptanalysis of ARX. In Seokhie Hong and Tetsu Iwata, editors, *FSE*, volume 6147 of *LNCS*, pages 333–346. Springer, 2010.

[71] Thorsten Kranz, Gregor Leander, and Friedrich Wiemer. Linear Cryptanalysis: Key Schedules and Tweakable Block Ciphers. *IACR Trans. Symmetric Cryptol.*, 2017(1):474–505, 2017.

[72] David Lefranc, Philippe Painchault, Valérie Rouat, and Emmanuel Mayer. A Generic Method to Design Modes of Operation Beyond the Birthday Bound. In Carlisle M. Adams, Ali Miri, and Michael J. Wiener, editors, *SAC*, volume 4876 of *LNCS*, pages 328–343. Springer, 2007.

[73] Ya Liu, Yifan Shi, Dawu Gu, Zhiqiang Zeng, Fengyu Zhao, Wei Li, Zhiqiang Liu, and Yang Bao. Improved Meet-in-the-Middle Attacks on Reduced-Round Kiasu-BC and Joltik-BC. *Comput. J.*, 62(12):1761–1776, 2019.

[74] Michael Luby and Charles Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM J. Comput.*, 17(2):373–386, 1988.

[75] Stefan Lucks. The Sum of PRPs Is a Secure PRF. In Bart Preneel, editor, *EUROCRYPT*, volume 1807 of *LNCS*, pages 470–484. Springer, 2000.

[76] Hamid Mala, Mohammad Dakhilalian, Vincent Rijmen, and Mahmoud Modarres-Hashemi. Improved Impossible Differential Cryptanalysis of 7-Round AES-128. In Guang Gong and Kishan Chand Gupta, editors, *INDOCRYPT*, volume 6498 of *LNCS*, pages 282–291. Springer, 2010.

[77] Bart Mennink and Samuel Neves. Encrypted Davies-Meyer and Its Dual: Towards Optimal Security Using Mirror Theory. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO III*, volume 10403 of *LNCS*, pages 556–583. Springer, 2017.

[78] Bart Mennink and Samuel Neves. Optimal PRFs from Blockcipher Designs. *IACR Trans. Symmetric Cryptol.*, 2017(3):228–252, 2017.

[79] Valérie Nachef, Jacques Patarin, and Emmanuel Volte. *Feistel Ciphers - Security Proofs and Cryptanalysis*. Springer, 2017.

[80] Chao Niu, Muzhou Li, Meiqin Wang, Qingju Wang, and Siu-Ming Yiu. Related-Tweak Impossible Differential Cryptanalysis of Reduced-Round TweAES. In Riham AlTawy and Andreas Hülsing, editors, *SAC*, volume 13203 of *LNCS*, pages 223–245. Springer, 2021.

[81] Jin Hyung Park and Dong Hoon Lee. FACE: Fast AES CTR mode Encryption Techniques based on the Reuse of Repetitive Data. *IACR Trans. Cryptogr. Hardw. Embed. Syst.*, 2018(3):469–499, 2018.

[82] Jacques Patarin. A Proof of Security in $O(2^n)$ for the Xor of Two Random Permutations. In Reihaneh Safavi-Naini, editor, *ICITS*, volume 5155 of *LNCS*, pages 232–248. Springer, 2008. Full version at https://eprint.iacr.org/2008/010.

[83] Jacques Patarin. The "Coefficients H" Technique. In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *SAC*, volume 5381 of *LNCS*, pages 328–345. Springer, 2008.

[84] Jacques Patarin. Introduction to Mirror Theory: Analysis of Systems of Linear Equalities and Linear Non Equalities for Cryptography. *IACR Cryptology ePrint Archive*, 2010:287, 2010.

[85] Jacques Patarin. Security in $\mathrm{O}(2^n)$ for the Xor of Two Random Permutations: Proof with the standard H technique. *IACR Cryptology ePrint Archive*, 2013:368, 2013.

[86] Thomas Peyrin. Improved Differential Attacks for ECHO and Grøstl. In Tal Rabin, editor, *CRYPTO*, volume 6223 of *LNCS*, pages 370–392. Springer, 2010.

[87] Thomas Peyrin and Yannick Seurin. Counter-in-Tweak: Authenticated Encryption Modes for Tweakable Block Ciphers. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO I*, volume 9814 of *LNCS*, pages 33–63. Springer, 2016.

[88] Claude E. Shannon. Communication theory of secrecy systems. *Bell Syst. Tech. J.*, 28(4):656–715, 1949.

[89] Tairong Shi, Chenhui Jin, and Jie Guan. Collision attacks against aez-prf for authenticated encryption aez. *China Communications*, 15(2):46–53, 2018.

[90] Bing Sun. Provable Security Evaluation of Block Ciphers Against Demirci-Selçuk's Meet-in-the-Middle Attack. *IEEE Trans. Inf. Theory*, 67(7):4838–4844, 2021.

[91] Yosuke Todo. Structural Evaluation by Generalized Integral Property. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT I*, volume 9056 of *LNCS*, pages 287–314. Springer, 2015.

[92] Yosuke Todo and Kazumaro Aoki. FFT Key Recovery for Integral Attack. In Dimitris Gritzalis, Aggelos Kiayias, and Ioannis G. Askoxylakis, editors, *CANS*, volume 8813 of *LNCS*, pages 64–81. Springer, 2014.

[93] Mohamed Tolba, Ahmed Abdelkhalek, and Amr M. Youssef. A Meet in the Middle Attack on Reduced Round Kiasu-BC. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 99-A(10):1888–1890, 2016.

[94] Qianqian Yang, Ling Song, Siwei Sun, Danping Shi, and Lei Hu. New Properties of Double Boomerang Connectivity Table. *IACR Cryptol. ePrint Arch.*, page 1579, 2022.

A The χ^2 Method

For each $i \in [q]$ and each vector $\mathbf{W}^{i-1} = (W_2^{i-1}, \ldots, W_r^{i-1})$ with $\mathbf{W}_j^{i-1} = (W_j^1, W_j^2, \ldots, W_j^{i-1})$, define

$$\chi^2(\mathbf{W}^{i-1}) \stackrel{\text{def}}{=} \sum_{W \in (\mathbb{F}_2^b)^{r-1}} \frac{\left(\Pr_{\mathcal{O}_{\text{real}}}[\mathbf{W}^i = W | \mathbf{W}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[\mathbf{W}^i = W | \mathbf{W}^{i-1}]\right)^2}{\Pr_{\mathcal{O}_{\text{ideal}}}[\mathbf{W}^i = W | \mathbf{W}^{i-1}]}.$$

Theorem 10 (χ^2 Method [41])**.** Consider two systems $\mathcal{O}_{\text{real}}$ and $\mathcal{O}_{\text{ideal}}$. Suppose that for any vector $\mathbf{W}$, it holds that $\Pr_{\mathcal{O}_{\text{ideal}}}[\mathbf{W}^i] > 0$ whenever $\Pr_{\mathcal{O}_{\text{real}}}[\mathbf{W}^i] > 0$. Then

$$\left\| \Pr_{\mathcal{O}_{\text{real}}}\left[\mathbf{W}^i\right] - \Pr_{\mathcal{O}_{\text{ideal}}}\left[\mathbf{W}^i\right] \right\| \le \sqrt{\frac{1}{2}\sum_{i=1}^{q} \mathbb{E}_{\mathcal{O}_{\text{real}}}[\chi^2(\mathbf{W}^{i-1})]}\,.$$

B Proof for ForkSTH

We recall the theorem to aid the reader.

Theorem 8. Let a, b, q, r be positive integers and $\tau = (\mathbf{V}, \mathbf{W})$ be a good transcript such that $C_{\mathbf{V},j}(i) < 2^{b-2}$ holds for all $i \in [q]$ and $j \in [r]$ and $q \le 2^n/(3r)$. Then, for $r \ge 3$, it holds that

$$\mathbb{E}_\tau\left[\left|\Pr\left[\mathcal{O}_2 = \tau\right] - \Pr\left[\mathcal{O}_{\text{ideal}} = \tau\right]\right|\right] \le \left(\frac{4}{3}\right)^r \cdot \left(\frac{rq}{2^{n-a/3}}\right)^{3/2}\,.$$

The proof will use the χ^2 approach by Dai et al. [41]. Gunsing and Mennink have shown in [58] that we can condition on an auxiliary variable $\mathbf{Y}^{i-1}$ instead of $\mathbf{W}^{i-1}$, even if the former exists in only one of the worlds, as long as the former allows to derive $\mathbf{W}^{i-1}$ uniquely.

Theorem 11 ([58])**.** Let $\mathbf{Y}^{i-1}$ be a random variable existing in world $\mathcal{O}_{\text{real}}$ but not necessarily in $\mathcal{O}_{\text{ideal}}$. Then,

$$\begin{aligned}&\mathbb{E}_{\mathcal{O}_{\text{real}}}[\chi^2(\mathbf{W}^{i-1})]\\ &\le \sum_{W \in (\mathbb{F}_2^b)^{r-1}} \mathbb{E}_{\mathcal{O}_{\text{real}}}\left[\frac{\left(\Pr_{\mathcal{O}_{\text{real}}}[\mathbf{W}^i = W|\mathbf{W}^{i-1}, \mathbf{Y}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[\mathbf{W}^i = W|\mathbf{W}^{i-1}]\right)^2}{\Pr_{\mathcal{O}_{\text{ideal}}}[\mathbf{W}^i = W|\mathbf{W}^{i-1}]}\right].\end{aligned}$$

The proof for $r = 2$ is provided in [58] and works in the same manner as there for general r.

Proof of Theorem 8. We can easily see that $\Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}] = 2^{-(r-1)b}$. Though, it remains to determine the probability in the real world. We denote the outputs $(Y_1^i, Y_2^i, \ldots, Y_r^i)$ also as $(y_1^i, y_2^i, \ldots, y_r^i)$ and the fixed sum values at the i-th step $(W_2^i, \ldots, W_r^i)$ also as $(w_2^i, \ldots, w_r^i)$. We consider r independent permutations $\pi_1, \ldots \pi_r$. We have to determine the probability

$$\Pr_{\mathcal{O}_{\text{real}}}\left[\mathbf{W}^i = (w_2^i, \ldots, w_r^i)|\mathbf{Y}^{i-1}\right],$$

where $\mathbf{Y}^{i-1} = (Y_1^1, \ldots, Y_r^1, \ldots, Y_1^{i-1}, \ldots, Y_r^{i-1})$.

Fix a tuple $\mathbf{W}^i = (w_2^i, \ldots, w_r^i) \in (\mathbb{F}_2^b)^{r-1}$. We define $q \times r$ sets $\mathcal{S}_j^i = \{y_j^1, \ldots, y_j^{i-1}\}$ for all $i \in [q]$ and $j \in [r]$. Furthermore, we propose sets of translated values $\mathcal{S}_{y_j \to w_j}^i = \mathcal{S}_j^i \oplus w_j =^{\text{def}} \{Y_j \in \mathcal{S}_j^i : Y_j \oplus w_j\}$ to denote the elementwise translation of $\mathcal{S}_j^i$ for the fixed scalar $w_j \in \mathbb{F}_2^b$ for all $j \in \{2, \ldots, r\}$. For consistency, we introduce $w_1^i = 0^b$ for all $i \in [q]$ so we can define $\mathcal{S}_{y_1 \to w_1}^i = \mathcal{S}_1^i$. We define cardinalities $s_j^i = |\mathcal{S}_{y_j \to w_j}^i| = |\mathcal{S}_j^i|$ for all $j \in [r]$.

We have to find the number of possible solutions $Y^i = (Y_1^i, \ldots, Y_r^i)$ for the next fixed tuple $W^i = (w_2^i, \ldots, w_r^i)$. For $Y_1^i \oplus Y_2^i = w_2^i$, $Y_1^i \oplus Y_3^i = w_3^i$, $\ldots$, it must hold that

$$Y_1^i \in \mathbb{F}_2^b \setminus \left(\mathcal{S}_1^i \cup \bigcup_{j=2}^{r} (\mathcal{S}_{y_j \to w_j}^i)\right)\,.$$

From the inclusion-exclusion principle, the number of choices for Y_1^i, that we denote by n^i, is

$$\begin{aligned}
n^i &= 2^b - \left(|\mathcal{S}^i_{y_1\to w_1}| + |\mathcal{S}^i_{y_2\to w_2}| + \cdots + |\mathcal{S}^i_{y_r\to w_r}|\right) + \\
&\quad \left(|\mathcal{S}^i_{y_1\to w_1} \cap \mathcal{S}^i_{y_2\to w_2}| + |\mathcal{S}^i_{y_1\to w_1} \cap \mathcal{S}^i_{y_3\to w_3}| + \cdots + |\mathcal{S}^i_{y_{r-1}\to w_{r-1}} \cap \mathcal{S}^i_{y_r\to w_r}|\right) - \\
&\quad \left(|\mathcal{S}^i_{y_1\to w_1} \cap \mathcal{S}^i_{y_2\to w_2} \cap \mathcal{S}^i_{y_3\to w_3}|) + \cdots\right) + \cdots \\
&= 2^b - \left(\sum_{j=1}^{r} |\mathcal{S}^i_{y_j\to w_j}|\right) + \left(\sum_{j_1<j_2} |\mathcal{S}^i_{y_{j_1}\to w_{j_1}} \cap \mathcal{S}^i_{y_{j_2}\to w_{j_2}}|\right) - \\
&\quad \left(\sum_{1\le j_1<j_2<j_3\le r} |\mathcal{S}^i_{y_{j_1}\to w_{j_1}} \cap \mathcal{S}^i_{y_{j_2}\to w_{j_2}} \cap \mathcal{S}^i_{y_{j_3}\to w_{j_3}}|\right) + \cdots \\
&= 2^b - \left(\sum_{j=1}^{r} s^i_j\right) + \left(\sum_{1\le j_1<j_2\le r} s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}\right) - \\
&\quad \left(\sum_{1\le j_1<j_2<j_3\le r} s^{i,w_{j_1},w_{j_2},w_{j_3}}_{j_1,j_2,j_3}\right) + \cdots + \\
&\quad (-1)^r \left(\sum_{1\le j_1<\cdots<j_r\le r} s^{i,w_{j_1},w_{j_2},\ldots,w_{j_r}}_{j_1,\ldots,j_r}\right), \qquad (4)
\end{aligned}$$

where we define $s^{i,w_1,w_2}_{1,2}$, $s^{i,w_1,w_2,w_3}_{1,2,3}$, ... for the cardinalities of the corresponding intersection sets in a natural manner. We call the terms $s^{i,w_1,w_2}_{1,2}$ 2-tuple-related, $s^{i,w_1,w_2,w_3}_{1,2,3}$ 3-tuple-related, and so on. For each, we have to upper bound its expectation and variance.

Expectation and Variance of 2-tuple-related Terms. We can use the knowledge about $s^{i,w_1,w_2}_{1,2} = s^{i,0,w_2}_{1,2} = D_{i,w}$ from [41, 58]. Thus, the expectation and variance of all cardinalities of two-component intersections can be taken from Equations (34) and (35) in [58] as

$$\mathbb{E}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}] = \frac{s^i_{j_1} s^i_{j_2}}{2^b} \qquad \mathbf{Var}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}] \le \frac{2 s^i_{j_1} s^i_{j_2}}{2^b}\,. \qquad (5)$$

For independent permutations $\pi_1, \ldots, \pi_r$, and independent Binomial variables, we can derive them more precisely.

Lemma 1. For distinct $j_1, j_2 \in [r]$, it holds that

$$\mathbb{E}[s^{i,w_{j_1},w_{j_2}}_{y_{j_1},y_{j_2}}] = \frac{s^i_{j_1} s^i_{j_2}}{2^b} \qquad \text{and}$$
$$\mathbf{Var}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}] = \frac{s^i_{j_1} s^i_{j_2}}{2^b} - \frac{(s^i_{j_1} s^i_{j_2})^2}{2^{3b}}\,.$$

Expectation and Variance of 3-tuple-related Terms. Next, we consider the expectation and variance of $s^{i,w_1,w_2,w_3}_{y_1,y_2,y_3}$.

Lemma 2. For distinct $j_1, j_2, j_3 \in [r]$, it holds that

$$\mathbb{E}[s^{i,w_{j_1},w_{j_2},w_{j_3}}_{y_{j_1},y_{j_2},y_{j_3}}] = \frac{s^i_{j_1} s^i_{j_2} s^i_{j_3}}{2^{2b}} \qquad \text{and}$$
$$\mathbf{Var}[s^{i,w_{j_1},w_{j_2},w_{j_3}}_{j_1,j_2,j_3}] = \frac{s^i_{j_1} s^i_{j_2} s^i_{j_3}}{2^{2b}} - \frac{(s^i_{j_1} s^i_{j_2} s^i_{j_3})^2}{2^{5b}}\,.$$

Expectation and Variance of Terms for General Tuples.

Lemma 3. Let $t \leq r$ and $\{I\} = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, r\}$. Then, it holds for the expectation and variance that

$$\mathbb{E}[s^{i,w_{j_1},w_{j_2},\ldots,w_{j_t}}_{j_1,j_2,\cdots,j_t}] = \frac{\prod_{j\in\{I\}} s^i_j}{2^{(t-1)b}}$$
$$\mathbf{Var}[s^{i,w_{j_1},w_{j_2},\ldots,w_{j_t}}_{j_1,j_2,\cdots,j_t}] = \frac{\prod_{j\in\{I\}} s^i_j}{2^{(t-1)b}} - \frac{\left(\prod_{j\in\{I\}} s^i_j\right)^2}{2^{(2t-1)b}}\,.$$

Determining the Ratio. In the real and ideal worlds, it holds that

$$\Pr_{\mathcal{O}_{\text{real}}}[W^i = (w^i_2, \ldots, w^i_r)|\mathbf{Y}^{i-1}] = \mathbb{E}[\frac{n^i}{d^i}] \qquad \text{and}$$
$$\Pr_{\mathcal{O}_{\text{ideal}}}[W^i = (w^i_2, \ldots, w^i_r)|\mathbf{W}^{i-1}] = \frac{1}{2^{(r-1)b}}\,,$$

respectively, with n^i given in Equation (4). The number of all choices of Y^i, that represents the denominator d^i, is

$$\begin{aligned} d^i &= (2^b - s^i_1)\cdot(2^b - s^i_2)\cdot\cdots\cdot(2^b - s^i_r) = \prod_{j=1}^{r}(2^b - s^i_j) \\ &= 2^{rb} - 2^{(r-1)b}\left(\sum_{j=1}^{r} s^i_j\right) + 2^{(r-2)b}\left(\sum_{1\leq j_1<j_2\leq r} s^i_{j_1}s^i_{j_2}\right) - \\ &\quad 2^{(r-3)b}\left(\sum_{1\leq j_1<j_2<j_3\leq r} s^i_{j_1}s^i_{j_2}s^i_{j_3}\right) + \cdots + (-1)^r\left(\sum_{1\leq j_1<\cdots<j_r\leq r} s^i_{j_1}\cdots s^i_{j_r}\right)\,, \end{aligned} \tag{6}$$

which yields

$$\begin{aligned} &\mathbb{E}\left[\left(\Pr_{\mathcal{O}_{\text{real}}}[W^i = (w^i_2, \ldots, w^i_r)|\mathbf{Y}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[W^i = (w^i_2, \ldots, w^i_r)|\mathbf{W}^{i-1}]\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{n^i}{d^i} - \frac{1}{2^{(r-1)b}}\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{2^{(r-1)b}\cdot n^i - d^i}{2^{(r-1)b}\cdot d^i}\right)^2\right] \\ &\leq \left(\frac{4}{3}\right)^{2r}\cdot\frac{1}{2^{(4r-2)b}}\cdot\mathbb{E}\left[\left(2^{(r-1)b}\cdot n^i - d^i\right)^2\right]\,, \end{aligned} \tag{7}$$

where we used the assumption of $s^i_j < 2^{b-2}$, for all $j \in [r]$, to upper bound $d^i \geq \left(\frac{3}{4}\cdot 2^b\right)^r$. In the following, we focus on the rightmost term of Equation (7), i.e., the expectation of the squared difference. We can observe that the leftmost two terms of $2^{(r-1)b}\cdot n^i$, that we call $\underline{n}^i$ for short,

$$\underline{n}^i \stackrel{\text{def}}{=} 2^{(r-1)b}\cdot\left(2^b - \sum_{j=1}^{r} s^i_j\right) = 2^{rb} - 2^{(r-1)b}\left(\sum_{j=1}^{r} s^i_j\right)\,,$$

are identical to the leftmost two terms in d^i as in Equation (6). Therefore, they cancel in the difference. We define

$$\begin{aligned}\overline{n}^i &\stackrel{\text{def}}{=} n^i - \left(2^b - \sum_{j=1}^{r} s_j^i\right) \qquad (8)\\ &= \left(\sum_{1\le j_1<j_2\le r} s_{j_1,j_2}^{i,w_{j_1},w_{j_2}}\right) - \left(\sum_{1\le j_1<j_2<j_3\le r} s_{j_1,j_2,j_3}^{i,w_{j_1},w_{j_2},w_{j_3}}\right) + \cdots + \\ &\quad (-1)^r \left(s_{1,\ldots,r}^{i,w_1,w_2,\ldots,w_r}\right)\\ \overline{d}^i &\stackrel{\text{def}}{=} d^i - \underline{n}^i\,. \qquad (9)\end{aligned}$$

We substitute the extended formulation of d^i from Equation (6) into Equation (8) and factor out $(2^{(r-1)b})^2$:

$$\begin{aligned}\mathbb{E}\left[\left(2^{(r-1)b}\cdot n^i - d^i\right)^2\right] &= \mathbb{E}\left[2^{2(r-1)b}\cdot\left(n^i - \frac{d^i}{2^{(r-1)b}}\right)^2\right]\\ &= 2^{2(r-1)b}\cdot\mathbb{E}\left[\left(\overline{n}^i - \frac{\overline{d}^i}{2^{(r-1)b}}\right)^2\right]\,. \qquad (10)\end{aligned}$$

We can write the rightmost term as

$$\begin{aligned}\frac{\overline{d}^i}{2^{(r-1)b}} &= \left(\sum_{1\le j_1<j_2\le r} \frac{s_{j_1}^i s_{j_2}^i}{2^b}\right) - \left(\sum_{1\le j_1<j_2<j_3\le r} \frac{s_{j_1}^i s_{j_2}^i s_{j_3}^i}{2^{2b}}\right) + \cdots\\ &\quad + (-1)^r\cdot\frac{s_1^i\cdots s_r^i}{2^{(r-1)b}}\,. \qquad (11)\end{aligned}$$

From Equation (4) for n^i, we can observe that for the sum of terms x in $\overline{n}^i$, Equation (11) consists of exactly the sum of terms $\mathbb{E}[\mathsf{x}]$.

$$\begin{aligned}(10) &= 2^{(2r-2)b}\cdot\mathbb{E}[\left(\overline{n}^i - \mathbb{E}[\overline{n}^i]\right)^2]\\ &= 2^{(2r-2)b}\cdot\mathbf{Var}[\overline{n}^i]\,.\end{aligned}$$

Inserting it into Equation (7) yields

$$\left(\frac{4}{3}\right)^{2r}\cdot\frac{1}{2^{(4r-2)b}}\cdot\mathbb{E}\left[\left(2^{(r-1)b}\cdot n^i - d^i\right)^2\right] \le \left(\frac{4}{3}\right)^{2r}\cdot\frac{1}{2^{2rb}}\cdot\mathbf{Var}[\overline{n}^i]\,.$$

For the sum of random variables x_i, it holds that

$$\mathbf{Var}[\overline{n}^i] = \sum_i\sum_j \mathbf{Cov}[\mathsf{x}_i,\mathsf{x}_j] = c^i\,,$$

where c^i is the sum of the pairwise covariances of all combinations of two addends in $\mathbf{Var}[\overline{n}^i]$, which includes the (always positive) variance terms:

$$\begin{aligned} c^i = \Bigg[& \left(\sum_{1 \le j_1 < j_2 \le r} \sum_{1 \le j'_1 < j'_2 \le r} \mathbf{Cov}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}, s^{i,w_{j'_1},w_{j'_2}}_{j'_1,j'_2}] \right) \\ & - \left(\sum_{1 \le j_1 < j_2 \le r} \sum_{1 \le j'_1 < j'_2 < j'_3 \le r} \mathbf{Cov}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}, s^{i,w_{j'_1},w_{j'_2},w_{j'_3}}_{j'_1,j'_2,j'_3}] \right) + \cdots \\ & + (-1)^r \left(\sum_{j_1,j_2} \sum_{j'_1,\ldots j'_r} \mathbf{Cov}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}, s^{i,w_{j'_1},w_{j'_2},\ldots,w_{j'_r}}_{j'_1,\ldots,j'_r}] \right) \Bigg] \\ & - \Bigg[\left(\sum_{1 \le j_1 < j_2 < j_3 \le r} \sum_{1 \le j'_1 < j'_2 < j'_3 \le r} \mathbf{Cov}[s^{i,w_{j_1},w_{j_2},w_{j_3}}_{j_1,j_2,j_3}, s^{i,w_{j'_1},w_{j'_2},w_{j'_3}}_{j'_1,j'_2,j'_3}] \right) - \cdots \\ & + (-1)^r \left(\sum_{j_1,j_2,j_3} \sum_{j'_1,\ldots j'_r} \mathbf{Cov}[s^{i,w_{j_1},w_{j_2},w_{j_3}}_{j_1,j_2,j_3}, s^{i,w_{j'_1},w_{j'_2},\ldots,w_{j'_r}}_{j'_1,\ldots,j'_r}] \right) \Bigg] + \cdots . \end{aligned}$$

Recall that the covariance of a term with itself equals its variance and is always positive: $\mathbf{Cov}[\mathsf{x}_i, \mathsf{x}_i] = \mathbf{Var}[\mathsf{x}_i]$.

Covariance. In the definition of covariance,

$$\mathbf{Cov}[\mathsf{x}_i, \mathsf{x}_j] = \mathbb{E}[\mathsf{x}_i \cdot \mathsf{x}_j] - \mathbb{E}[\mathsf{x}_i]\mathbb{E}[\mathsf{x}_j] \,, \tag{12}$$

we can compute the products of expectations, but have to find the expectations of the products $\mathbb{E}[\mathsf{x}_i \cdot \mathsf{x}_j]$, with dependent variables x_i and x_j.

Lemma 4 considers the expectation of products. For all $i \in [\ell]$, let $\mathcal{S}_i$ be a list of n independent Bernoulli trials represented by indicator variables $I_{i,j}$ for $j \in [n]$, where $\Pr[I_{i,j} = 1] = p_i$ for all i, j. For multiple pairwise distinct index combinations $i_1, \ldots, i_r \in [\ell]$, let $\mathcal{S}_{i_1,\ldots,i_r} = \{j : I_{i_1,j} = \cdots = I_{i_r,j} = 1\}$ for $j \in [n]$. Let $s_i = |\mathcal{S}_i|$ and $s_{i_1,\ldots,i_r} = |\mathcal{S}_{i_1,\ldots,i_r}|$ for all indices and all pairwise distinct index combinations.

We use $\mathcal{I}, \mathcal{J} \subseteq \{i_1, \ldots, i_r\}$ as distinct index sets and overload the notations so that for each $\mathcal{I} = \{j_1, \ldots, j_s\} \subseteq \{i_1, \ldots, i_r\}$, we define $s_{\mathcal{I}} = s_{j_1,\ldots,j_s}$. Moreover, we define $p_{\mathcal{I}} = \prod_{i \in \mathcal{I}} p_i$. Note that

$$\begin{aligned} \mathbb{E}[s_{\mathcal{I}}] \cdot \mathbb{E}[s_{\mathcal{J}}] &= np_{\mathcal{I}} \cdot np_{\mathcal{J}} \\ \mathbb{E}[s_{\mathcal{I}} \cdot s_{\mathcal{J}}] &= \mathbb{E}[s_{\mathcal{I}}] \cdot \mathbb{E}[s_{\mathcal{J}}] + \mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}] \,. \end{aligned}$$

If $\mathcal{I} \cap \mathcal{J} = \emptyset$, it follows that $p_{\mathcal{I} \cup \mathcal{J}} = p_{\mathcal{I}} \cdot p_{\mathcal{J}}$; thus, $\mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}] = 0$ and

$$\mathbb{E}[s_{\mathcal{I}} \cdot s_{\mathcal{J}}] = \mathbb{E}[s_{\mathcal{I}}] \cdot \mathbb{E}[s_{\mathcal{J}}] \,.$$

Though, for the cases when $\mathcal{I} \cap \mathcal{J} \neq \emptyset$, we have to find $\mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}]$ in Lemma 4.

Lemma 4. It holds that

$$\mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}] = np_{\mathcal{I} \cup \mathcal{J}} - np_{\mathcal{I}} \cdot p_{\mathcal{J}} \,.$$

We show that we are allowed to apply Lemma 4. Since the permutations are independent from each other and the values are sampled independently at random, we can say that

each value in $\mathcal{S}_u$, $\mathcal{S}_v$, $\mathcal{S}_w$ is chosen independently from the others. The size of all three lists is $n = 2^b$; moreover, we can instantiate the probabilities p_j, for $j \in [r]$ as

$$p_j \stackrel{\text{def}}{=} \frac{s_j^i}{2^b}\,.$$

In our case, this means

$$\mathbb{E}[s_{\mathcal{I}} \cdot s_{\mathcal{J}}] = 2^{2b} \cdot \prod_{i\in\mathcal{I}} p_j \cdot \prod_{j\in\mathcal{J}} p_j + \mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}]$$
$$\mathbf{Cov}[s_{\mathcal{I}}, s_{\mathcal{J}}] = 2^b \cdot \prod_{i\in\mathcal{I}\cup\mathcal{J}} p_i - 2^b \cdot \prod_{i\in\mathcal{I}} p_i \cdot \prod_{j\in\mathcal{J}} p_j\,.$$

For example, let $\mathcal{I} = \{1, 2\}$ and $\mathcal{J} = \{1, 3, 4\}$. Then,

$$\mathbf{Cov}[s^i_{1,2}, s^i_{1,3,4}] = 2^b \cdot \left(\frac{s_1^i s_2^i s_3^i s_4^i}{2^{4b}} - \frac{(s_1^i)^2 s_2^i s_3^i s_4^i}{2^{5b}}\right)\,.$$

Decomposing c^i. Given the covariance, we can rewrite c^i. We define $\mathcal{C}_{t,r}$ for the set of t-out-of-r element combinations, e.g. $\mathcal{C}_{2,3} = \{(1,2), (1,3), (2,3)\}$.

$$c^i = \sum_{t_1=2}^{r}\sum_{t_2=2}^{r}(-1)^{t_1+t_2} \cdot c^i_{t_1,t_2,r}\,, \quad \text{where} \tag{13}$$
$$c^i_{t_1,t_2,r} = \sum_{\mathcal{I}\in\mathcal{C}_{t_1,r}}\sum_{\mathcal{J}\in\mathcal{C}_{t_2,r}} \mathbf{Cov}[s^{i,w_{\mathcal{I}}}_{\mathcal{I}}, s^{i,w_{\mathcal{J}}}_{\mathcal{J}}]\,.$$

Lemma 4 allows us to write

$$c^i_{t_1,t_2,r} = \sum_{\mathcal{I}\in\mathcal{C}_{t_1,r}}\sum_{\mathcal{J}\in\mathcal{C}_{t_2,r}} 2^b \cdot (p_{\mathcal{I}\cup\mathcal{J}} - p_{\mathcal{I}}p_{\mathcal{J}}) \tag{14}$$
$$= 2^b \cdot \underbrace{\left(\sum_{\mathcal{I}\in\mathcal{C}_{t_1,r}}\sum_{\mathcal{J}\in\mathcal{C}_{t_2,r}}\prod_{j\in\mathcal{I}\cup\mathcal{J}} p_j\right)}_{\overline{c}^i_{t_1,t_2,r}} - 2^b \cdot \underbrace{\left(\sum_{\mathcal{I}\in\mathcal{C}_{t_1,r}}\sum_{\mathcal{J}\in\mathcal{C}_{t_2,r}}\prod_{i\in\mathcal{I}} p_i \prod_{j\in\mathcal{J}} p_j\right)}_{\underline{c}^i_{t_1,t_2,r}}\,. \tag{15}$$

Later, we will consider the case that $p_1 = p_2 = \cdots = p_r = p$. Then, we can write $c^i_{t_1,t_2,r}$ as

$$c^i_{t_1,t_2,r} = 2^b \cdot \left(\overline{c}^i_{t_1,t_2,r} - \underline{c}^i_{t_1,t_2,r}\right)$$
$$= 2^b \cdot \left(\sum_{j=0}^{u}\left(\overline{k}^i_{t_1,t_2,r,j} \cdot p^{\overline{\ell}^i_{t_1,t_2,r,j}}\right) - \underline{k}^i_{t_1,t_2,r} \cdot p^{\underline{\ell}^i_{t_1,t_2,r}}\right)$$

with $u \stackrel{\text{def}}{=} \min(r - t_2, t_1)$ and j denotes the number of elements in $\mathcal{I}$ that are not contained in $\mathcal{J}$. Thus, we can reduce the task to that of finding the multiples

$$\overline{k}^i_{t_1,t_2,r,j} = |\{(\mathcal{I},\mathcal{J}) \in \mathcal{C}_{t_1,r}, \times\mathcal{C}_{t_2,r} : |\mathcal{I}\cup\mathcal{J}| = t_2 + j\}| \qquad \text{and} \tag{16}$$
$$\overline{\ell}^i_{t_1,t_2,r,j} = |\mathcal{I}\cup\mathcal{J}| \tag{17}$$

and

$$\underline{k}^i_{t_1,t_2,r} = |\{(\mathcal{I},\mathcal{J}) \in \mathcal{C}_{t_1,r}, \times\mathcal{C}_{t_2,r}\}| = |\mathcal{C}_{t_1,r}| \cdot |\mathcal{C}_{t_2,r}| = \binom{r}{t_1} \cdot \binom{r}{t_2} \qquad \text{and} \tag{18}$$
$$\underline{\ell}^i_{t_1,t_2,r} = |\mathcal{I}| + |\mathcal{J}| = t_1 + t_2\,. \tag{19}$$

The exponent $\overline{\ell}^i_{t_1,t_2,r,j}$ is derived from the size of the union set $\mathcal{I} \cup \mathcal{J}$ when j elements of $\mathcal{I}$ are not in $\mathcal{J}$. Thus

$$\overline{\ell}^i_{t_1,t_2,r,j} = \max(t_1, t_2) + j$$

for all $j \in [0..u]$ where $u =^{\text{def}} \min(r - t_2, t_1)$. It remains to determine $\overline{k}^i_{t_1,t_2,r,j}$. For this purpose, we can use the simple combinatorial Lemma 5.

Lemma 5. Let t_1, t_2, r, j be fixed integers with $t_1 \le t_2 \le r$ and $j \in [t_2..r]$. Let $\mathcal{I}, \mathcal{J} \subseteq [r]$ be non-identical subsets of $[r]$ with $|\mathcal{I}| = t_1$ and $|\mathcal{J}| = t_2$. Then, the number of combinations of distributing $\mathcal{I}$ and $\mathcal{J}$ so that

$$|\{(\mathcal{I}, \mathcal{J}) \in \mathcal{C}_{t_1,r}, \times \mathcal{C}_{t_2,r} : |\mathcal{I} \cup \mathcal{J}| = t_2 + j\}| = \binom{r}{t_2} \cdot \binom{t_2}{t_1 - j} \cdot \binom{r - t_2}{j}.$$

Proof. W.l.o.g., we had fixed that $|\mathcal{I}| \le |\mathcal{J}|$ and therefore $t_1 \le t_2$. There are $\binom{r}{t_2}$ sets $\mathcal{J}$ among r elements. We defined that j elements of $\mathcal{I}$ are not in $\mathcal{J}$. For a fixed $\mathcal{J}$ and fixed j, there are $\binom{t_2}{t_1-j}$ combinations of the $t_1 - j$ values in $\mathcal{I} \cap \mathcal{J}$ and $\binom{r-t_2}{j}$ combinations how the j values from $\mathcal{I} \setminus \mathcal{J}$ are distributed outside of $\mathcal{J}$. The lemma follows. □

We can rewrite Lemma 5 as Lemma 6, which will serve useful.

Lemma 6. Let t_1, t_2, r, ℓ be fixed integers with $t_1, t_2 \le r$. Let $\mathcal{I}, \mathcal{J} \subseteq [r]$ such that $|\mathcal{I}| = t_1$, $|\mathcal{J}| = t_2$, and $j = \ell - t_1$. Then, the number of combinations of distributing $\mathcal{I}$ and $\mathcal{J}$ so that

$$\begin{aligned} &|\{(\mathcal{I}, \mathcal{J}) \in \mathcal{C}_{t_1,r}, \times \mathcal{C}_{t_2,r} : |\mathcal{I} \cup \mathcal{J}| = \ell\}| \\ &= \binom{r}{t_1}\binom{t_1}{t_1 + t_2 - \ell}\binom{r - t_1}{\ell - t_1}(-1)^{t_1+t_2}. \end{aligned}$$

Proof. There are $\binom{r}{t_1}$ sets $\mathcal{I}$ among r elements. The overlap, i.e., the number of shared elements in the intersection $|\mathcal{I} \cap \mathcal{J}| = t_1 + t_2 - \ell$. Among the t_1 elements of $\mathcal{I}$, there are $\binom{t_1}{t_1+t_2-\ell}$ combinations what elements of $\mathcal{I}$ and $\mathcal{J}$ could be in the intersection. Then, the remaining $\ell - t_1$ elements in $\mathcal{J} \setminus \mathcal{I}$ can be distributed by $\binom{r-t_1}{\ell-t_1}$ combinations over the remaining $r - t_1$ elements not in $\mathcal{I}$. The lemma follows. □

Upper Bounding c^i for General r. We aim at having a simplified upper bound for c^i for general r. The terms in c^i consist of multiples of powers of p from exponents 2 to $2r$. Now, we can find non-negative integer coefficients k_j, for all $j \in [2..r]$, such that we can write

$$c^i = k^i_2 \cdot p^2 + k^i_3 \cdot p^3 + \sum_{j=2}^{r} \left((-1)^{2j-1} \cdot k^i_{2j} \cdot p^{2j}\right). \tag{20}$$

We show that there the indices $j \in [2..2r]$ are the only potential non-zero non-negative coefficients k^i_j. For $k_\ell \cdot p^\ell$ with $k_\ell < 2$, there must exist $\overline{\ell}^i_{t_1,t_2,r,j} < 2$ or $\underline{\ell}^i_{t_1,t_2,r} < 2$ for some $t_1, t_2 \in [2..r]$ and $j \le r$. Though, our sets that always have $|\mathcal{I}|, |\mathcal{J}| \in [2..r]$. Hence,

$$\begin{aligned} \overline{\ell}^i_{t_1,t_2,r,j} &= |\mathcal{I} \cup \mathcal{J}| \in [2..2r] \\ \underline{\ell}^i_{t_1,t_2,r} &= |\mathcal{I}| + |\mathcal{J}| \in [4..2r]. \end{aligned}$$

Thus, $k_\ell = 0$ for all $\ell \notin [2..2r]$. We want to reduce the bound to the terms with the few lowest exponents and show that we can upper bound the tail since the positive and

negative terms will compensate each other. In particular, we want a bound so that we can reduce Equation (20) to

$$c^i \leq 2^b \cdot (k_2 \cdot p^2 + k_3 \cdot p^3)\,.$$

Later, we show two aspects: first, that $p \leq \frac{1}{3r}$ always holds, and second, the following lemma.

Lemma 7. Let $r \geq 3$ be integer. It holds for all $\ell = 2j$ for some $j \in [2..r-1]$ that

$$\frac{|k^i_{\ell+1}|}{|k^i_\ell|} \leq 3r\,, \quad k^i_{\ell+1} \geq 0\,, \quad k^i_\ell \leq 0 \quad \text{and} \quad k^i_{2r} \leq 0\,.$$

We defer the proof of Lemma 7 to Appendix C.4. Combined with our assumption that $p \leq \frac{1}{3r}$, it follows for all $\ell = 2j$ for some $j \in [2..r-1]$, that

$$\begin{aligned} k_\ell \cdot p^\ell &\geq k_{\ell+1} \cdot p^{\ell+1} \\ k_\ell \cdot p^\ell &\geq 3r \cdot k_\ell \cdot p^\ell \cdot \frac{1}{3r}\,, \end{aligned}$$

and therefore

$$\begin{aligned} c^i &= 2^b \cdot \left(k_2 \cdot p^2 + k_3 \cdot p^3 + \sum_{j=2}^{r-1} \underbrace{\left(-k_{2j} \cdot p^{2j} + k_{2j+1} \cdot p^{2j+1}\right)}_{\leq 0} - k_{2r} \cdot p^{2r} \right) \\ &\leq 2^b \cdot (k_2 \cdot p^2 + k_3 \cdot p^3)\,. \end{aligned}$$

The factors k_2 and k_3 result from only few terms in c^i. In particular, they stem from $c^i_{2,2,r}$, $c^i_{2,3,r} = c^i_{3,2,r}$, and $c^i_{3,3,r}$. Given $r \geq 3$, they result from

$$\begin{aligned} k_2 &= \overline{k}^i_{2,2,r,0} = \binom{r}{2}\binom{2}{2}\binom{2}{0} = \binom{r}{2} \\ k_3 &= \overline{k}^i_{2,2,r,1} - \overline{k}^i_{2,3,r,0} - \overline{k}^i_{3,2,r,0} + \overline{k}^i_{3,3,r,0} \\ &= \binom{r}{2}\binom{2}{1}\binom{r-2}{1} - 2\binom{r}{3}\binom{3}{2}\binom{r-3}{0} + \binom{r}{3}\binom{3}{3}\binom{r-3}{0} = \binom{r}{3}\,. \end{aligned}$$

We obtain

$$c^i \leq 2^b \cdot \left(\binom{r}{2} \cdot p^2 + \binom{r}{3} \cdot p^3 \right)\,. \tag{21}$$

Equal Probabilities p_i. It remains to show that $p_1 = \cdots = p_r$. The values of the a most significant bits of the permutation outputs, $\mathbf{V}^i_j = V^1_j, \ldots, V^i_j$, for all $j \in [r]$, are sampled uniformly and independently at random, also in the modified real world $\mathcal{O}_{\text{real}}$ since we replace their sampling with that from a truncated permutation. Thus, every V^i_j has probability 2^{-a} to be equal to a specific a-bit value. Therefore

$$\mathbb{E}_{\mathbf{V}^{i-1}}[s^i_1] = \cdots = \mathbb{E}_{\mathbf{V}^{i-1}}[s^i_r] = \frac{i-1}{2^a}\,.$$

Thus, for all $j \in [r]$, we can use

$$p_j = \mathbb{E}[\frac{s^i_j}{2^b}] = \frac{\mathbb{E}[s^i_j]}{2^b} = \frac{i-1}{2^n}\,.$$

We have to show that the expectations of the quantities $s_1^i, \ldots, s_r^i$ are independent. We can adopt the argument from [58] here: it holds since they stem from pairwise independent permutations and hence

$$\mathbb{E}_{\mathbf{V}^{i-1}}[s_2^i | s_1^i] = \mathbb{E}_{\mathbf{V}^{i-1}}[s_2^i]$$

and similarly for all other combinations. We can use

$$\mathbb{E}_{\mathbf{V}^{i-1}}[s_1^i s_2^i] = \mathbb{E}_{\mathbf{V}^{i-1}}[s_1^i] \cdot \mathbb{E}_{\mathbf{V}^{i-1}}[s_2^i]$$

and the other product combinations can be decomposed similarly.

Finalizing with the χ^2 Approach. We have that

$$\mathbb{E}\left[\left(\Pr_{\mathcal{O}_{\text{real}}}[W^i = W|\mathbf{W}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}]\right)^2\right] \leq \left(\frac{4}{3}\right)^{2r} \cdot \frac{1}{2^{2rb}} \cdot \mathbf{Var}[\overline{n}^i]\,.$$

Using the χ^2 approach and inserting $\Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}] = 2^{-(r-1)b}$, we obtain

$$\begin{aligned}
&(\|\Pr[\mathcal{O}_{\text{real}} = \tau] - \Pr[\mathcal{O}_{\text{ideal}} = \tau]\|)^2 \\
&\leq \frac{1}{2}\sum_{i=1}^{q} \mathbb{E}_{\mathcal{O}_{\text{real}}}[\chi^2(\mathbf{W}^{i-1})] \\
&\leq \frac{1}{2}\sum_{i=1}^{q} \sum_{W \in (\mathbb{F}_2^b)^{r-1}} \mathbb{E}_{\mathcal{O}_{\text{real}}}\left[\frac{\left(\Pr_{\mathcal{O}_{\text{real}}}[W^i = W|\mathbf{W}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}]\right)^2}{\Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}]}\right] \\
&\leq \frac{1}{2} \cdot 2^{(r-1)b} \cdot \sum_{i=1}^{q} \sum_{W \in (\mathbb{F}_2^b)^{r-1}} \mathbb{E}\left[\left(\Pr_{\mathcal{O}_{\text{real}}}[W^i = W|\mathbf{W}^{i-1}] - \Pr_{\mathcal{O}_{\text{ideal}}}[W^i = W|\mathbf{W}^{i-1}]\right)^2\right] \\
&\leq \frac{1}{2} \cdot 2^{(r-1)b} \cdot \sum_{i=1}^{q} \sum_{W \in (\mathbb{F}_2^b)^{r-1}} \left(\left(\frac{4}{3}\right)^{2r} \cdot \frac{1}{2^{2rb}} \cdot c^i\right) \\
&\leq \frac{1}{2^{2b+1}} \cdot \left(\frac{4}{3}\right)^{2r} \cdot \sum_{i=1}^{q} c^i\,. \qquad (22)
\end{aligned}$$

From Equation (21)

$$c^i \leq 2^b \left(\binom{r}{2} p^2 + \binom{r}{3} p^3\right)$$

and $p = (i-1)/2^n$, we obtain that

$$\begin{aligned}
(22) &= \sqrt{\frac{1}{2^{2b+1}} \cdot \left(\frac{4}{3}\right)^{2r} \cdot \sum_{i=1}^{q} 2^b \cdot \left(\binom{r}{2}\frac{(i-1)^2}{2^{2n}} + \binom{r}{3}\frac{(i-1)^3}{2^{3n}}\right)} \\
&= \sqrt{\frac{1}{2^{2b+1}} \cdot \left(\frac{4}{3}\right)^{2r} \cdot \frac{1}{2^a} \cdot \sum_{i=1}^{q} \left(\binom{r}{2}\frac{(i-1)^2}{2^n} + \binom{r}{3}\frac{(i-1)^3}{2^{2n}}\right)} \\
&\leq \sqrt{\frac{1}{2^{2n-a}} \cdot \left(\frac{4}{3}\right)^{2r} \cdot \frac{1}{2} \cdot \left(\frac{r^2q^3}{2^n} + \frac{r^3q^4}{2^{2n}}\right)} \\
&\leq \left(\frac{4}{3}\right)^{r} \cdot \frac{1}{2} \cdot \sqrt{\frac{r^2q^3}{2^{3n-a}} + \frac{r^3q^4}{2^{4n-a}}} \\
&\leq \left(\frac{4}{3}\right)^{r} \cdot \left(\frac{rq}{2^{n-a/3}}\right)^{3/2},
\end{aligned}$$

which yields the bound in Theorem 8. We used the assumption that $q \leq 2^n/3r$ to upper bound

$$\frac{r^2q^3}{2^{3n-a}} + \frac{r^3q^4}{2^{4n-a}} \leq \frac{2r^3q^3}{2^{3n-a}}\,. \quad \square$$

Though, we can obtain tighter constant factors. We give the results for $r = 3, 4$ in Corollaries 1 and 2 to aid the reader.

Corollary 1. Let a, b, q be positive integers and $\tau = (\mathbf{V}, \mathbf{W})$ be a good transcript such that $C_{\mathbf{V},j}(i) < 2^{b-2}$ holds for all $i \in [q]$ and $j \in [r]$ and $q \leq 2^n/9$. Then, for $r = 3$, it holds that

$$\mathbb{E}_\tau[\Pr[\mathcal{O}_2 = \tau] - \Pr[\mathcal{O}_{\text{ideal}} = \tau]] \leq 4 \cdot \left(\frac{q}{2^{n-a/3}}\right)^{3/2}\,.$$

Corollary 2. Let a, b, q be positive integers and $\tau = (\mathbf{V}, \mathbf{W})$ be a good transcript such that $C_{\mathbf{V},j}(i) < 2^{b-2}$ holds for all $i \in [q]$ and $j \in [r]$ and $q \leq 2^n/12$. Then, for $r = 4$, it holds that

$$\mathbb{E}_\tau[\Pr[\mathcal{O}_2 = \tau] - \Pr[\mathcal{O}_{\text{ideal}} = \tau]] \leq 8 \cdot \left(\frac{q}{2^{n-a/3}}\right)^{3/2}\,.$$

C Proof of Lemmas for ForkSTH

C.1 Proof of Lemma 1

Lemma 1. For distinct $j_1, j_2 \in [r]$, it holds that

$$\mathbb{E}[s^{i,w_{j_1},w_{j_2}}_{y_{j_1},y_{j_2}}] = \frac{s^i_{j_1}s^i_{j_2}}{2^b} \qquad \text{and}$$
$$\mathbf{Var}[s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}] = \frac{s^i_{j_1}s^i_{j_2}}{2^b} - \frac{(s^i_{j_1}s^i_{j_2})^2}{2^{3b}}\,.$$

Proof. Let us focus on $s^{i,w_1,w_2}_{1,2}$; the remaining 2-tuple-related terms $s^{i,w_{j_1},w_{j_2}}_{j_1,j_2}$ behave similarly, for all $j_1 \neq j_2$, $j_1, j_2 \in [r]$. Given fixed $w_2 \in \mathbb{F}_2^b$, for each $y_1 \in \mathbb{F}_2^b$, we define Bernoulli variables I_{y_1} as

$$I_{y_1} \stackrel{\text{def}}{=} \begin{cases} 1 & y_1 \in \mathcal{S}^i_1 \wedge y_1 \oplus w_2 \in \mathcal{S}^i_2 \\ 0 & \text{otherwise.} \end{cases}$$

Then, we derive

$$\mathbb{E}\left[s^{i,w_1,w_2}_{1,2}\right] = \sum_{y_1 \in \mathbb{F}_2^b} \Pr[I_{y_1}]\,.$$

To obtain

$$\mathbf{Var}\left[\mathsf{x}\right] = \mathbb{E}\left[\mathsf{x}^2\right] - \left(\mathbb{E}[\mathsf{x}]\right)^2\,,$$

we have to determine $\mathbb{E}\left[\mathsf{x}^2\right]$. For a sum of n independent Bernoulli variables I_{y_1}, with $\Pr[I_{y_1} = 1] = p$ for all y_1,

$$\mathsf{x} = \sum_{y_1} \Pr[I_{y_1} = 1]\,,$$

it holds that

$$\mathbb{E}\left[\mathsf{x}^2\right] = \mathbb{E}\left[\left(\sum_{j=1}^{n} I_j\right)^2\right] = n(n-1)p^2 + np\,.$$

In our case, $n = 2^b$ and $p = s_1^i s_2^i \cdot 2^{-2b}$, for all $y_1 \in \mathbb{F}_2^b$. Given that $(\mathbb{E}[\mathsf{x}])^2 = (2^b p)^2$, we obtain

$$\begin{aligned}\mathbf{Var}\left[s_{1,2}^{i,0,w_2}\right] &\leq \frac{s_1^i s_2^i}{2^b} - \frac{(s_1^i s_2^i)^2}{2^{3b}} \quad \text{and in general}\\ \mathbf{Var}\left[s_{j_1,j_2}^{i,w_{j_1},w_{j_2}}\right] &\leq \frac{s_{j_1}^i s_{j_2}^i}{2^b} - \frac{(s_{j_1}^i s_{j_2}^i)^2}{2^{3b}}\,.\end{aligned}$$

C.2 Proof of Lemma 2

Lemma 2. For distinct $j_1, j_2, j_3 \in [r]$, it holds that

$$\begin{aligned}\mathbb{E}[s_{y_{j_1},y_{j_2},y_{j_3}}^{i,w_{j_1},w_{j_2},w_{j_3}}] &= \frac{s_{j_1}^i s_{j_2}^i s_{j_3}^i}{2^{2b}} \qquad \text{and}\\ \mathbf{Var}[s_{j_1,j_2,j_3}^{i,w_{j_1},w_{j_2},w_{j_3}}] &= \frac{s_{j_1}^i s_{j_2}^i s_{j_3}^i}{2^{2b}} - \frac{(s_{j_1}^i s_{j_2}^i s_{j_3}^i)^2}{2^{5b}}\,.\end{aligned}$$

Proof. Again, the remaining 3-tuple-related terms $s_{j_1,j_2,j_3}^{i,w_{j_1},w_{j_2},w_{j_3}}$ behave similarly, for all distinct $j_1, j_2, j_3 \in [r]$. Given fixed $w_1 = 0^b$ and $w_2, w_3 \in \mathbb{F}_2^b$, for each $y_1 \in \mathbb{F}_2^b$, we define Bernoulli variables I_{y_1} as

$$I_{y_1} \stackrel{\text{def}}{=} \begin{cases} 1 & y_1 \in \mathcal{S}_1^i \wedge y_1 \oplus w_2 \in \mathcal{S}_2^i \wedge y_1 \oplus w_3 \in \mathcal{S}_3^i \\ 0 & \text{otherwise.}\end{cases}$$

Then, it holds that

$$\mathbb{E}\left[s_{1,2,3}^{i,w_1,w_2,w_3}\right] = \mathbb{E}\left[\sum_{y_1 \in \mathbb{F}_2^b} I_{y_1}\right] = \sum_{y_1 \in \mathbb{F}_2^b} \mathbb{E}[I_{y_1}]\,.$$

Since the expectations for a fixed value $y_1 \in \mathbb{F}_2^b$ and its translations to be in the list of all three permutations are mutually independent, the probability is 2^{-3b}. Over all elements of the sets $|\mathcal{S}_{y_1 \to w_1}^i| = |\mathcal{S}_1^i|$, $|\mathcal{S}_{y_2 \to w_2}^i|$, and $|\mathcal{S}_{y_3 \to w_3}^i|$, it holds that

$$\mathbb{E}[I_u] = \frac{s_1^i s_2^i s_3^i}{2^{3b}} \quad \text{and therefore} \quad \mathbb{E}\left[s_{y_1,y_2,y_3}^{i,w_1,w_2,w_3}\right] = \frac{s_1^i s_2^i s_3^i}{2^{2b}}\,. \tag{23}$$

It remains to determine its variance

$$\begin{aligned}\mathbf{Var}\left[s_{y_1,y_2,y_3}^{i,w_1,w_2,w_3}\right] &= \mathbb{E}\left[\left(s_{y_1,y_2,y_3}^{i,w_1,w_2,w_3}\right)^2\right] - \left(\mathbb{E}\left[s_{y_1,y_2,y_3}^{i,w_1,w_2,w_3}\right]\right)^2\\ &= \mathbf{Var}\left[\sum_{y_1 \in \mathbb{F}_2^b} I_{y_1}\right] = \sum_{y_1 \in \mathbb{F}_2^b} \mathbf{Var}\left[I_{y_1}\right] + \sum_{y_1 \neq y_1'} \mathbf{Cov}\left[I_{y_1}, I_{y_1'}\right]\,,\end{aligned}$$

with the covariance

$$\begin{aligned}\mathbf{Cov}\left[I_{y_1}, I_{y_1'}\right] &= \mathbb{E}\left[I_{y_1}, I_{y_1'}\right] - \mathbb{E}[I_{y_1}]\,\mathbb{E}\left[I_{y_1'}\right]\\ &= \mathbb{E}[I_{y_1}]\,\Pr[I_{y_1'} = 1 | I_{y_1} = 1] - \mathbb{E}[I_{y_1}]\,\mathbb{E}\left[I_{y_1'}\right]\,.\end{aligned}$$

For the variance of the Bernoulli variables, it holds that

$$\mathbf{Var}\left[I_{y_1}\right] = \mathbb{E}\left[(I_{y_1})^2\right] - (\mathbb{E}[I_{y_1}])^2 = \mathbb{E}[I_{y_1}] - (\mathbb{E}[I_{y_1}])^2 = \frac{s_u^i s_v^i s_w^i}{2^{3b}} - \left(\frac{s_u^i s_v^i s_w^i}{2^{3b}}\right)^2 .$$

For their covariance, we need to determine the conditional probability. We consider the case that $y_1' \notin \{y_1 \oplus w_2, y_1 \oplus w_3\}$. Since $y_1' \neq y_1$, it holds that all values differ mutually

$$\begin{aligned}\Pr[I_{y_1'} = 1 | I_{y_1} = 1] = \Pr[&(y_1' \in \mathcal{S}_1^i) \wedge (y_1' \oplus w_2 \in \mathcal{S}_2^i) \wedge (y_1' \oplus w_3 \in \mathcal{S}_3^i) | \\ &(y_1 \in \mathcal{S}_1^i) \wedge (y_1 \oplus w_2 \in \mathcal{S}_2^i) \wedge (y_1 \oplus w_3 \in \mathcal{S}_3^i)] \\ &\leq \frac{(s_1^i - 1)(s_2^i - 1)(s_3^i - 1)}{(2^b - 1)^3} .\end{aligned}$$

We conduct it for $y_1' = y_1 \oplus w_2$ exemplarily. From the requirement of the covariance that $y_1' \neq y_1$, we must exclude $w_2 = 0$.

$$\begin{aligned}&\Pr[I_{y_1 \oplus w_2} = 1 | I_{y_1} = 1] \\ &\leq \Pr[(y_1 \oplus w_2 \in \mathcal{S}_1^i) \wedge (y_1 \in \mathcal{S}_2^i) \wedge (y_1 \oplus w_2 \oplus w_3 \in \mathcal{S}_3^i) | I_{y_1} = 1] \\ &\leq \frac{(s_u^i - 1)(s_v^i - 1)(s_w^i - 1)}{(2^b - 1)^3} .\end{aligned}$$

From $s_1^i, s_2^i, s_3^i < 2^b$, it follows that

$$\Pr[I_{y_1 \oplus w_2} = 1 | I_{y_1} = 1] \leq \mathbb{E}[I_{y_1 \oplus w_2}] ,$$

and therefore $\mathbf{Cov}\left[I_{y_1}, I_{y_1 \oplus w_2}\right] \leq 0$ in this case. A similar argument holds for $y_1' = y_1 \oplus w_3$, $w_2 \neq w_3$. It remains to consider $y_1' = y_1 \oplus w_2$ with $w_2 = w_3$.

$$\begin{aligned}&\Pr[I_{y_1 \oplus w_2} = 1 | I_{y_1} = 1, w_2 = w_3] \\ &\leq \Pr[(y_1 \oplus w_2 \in \mathcal{S}_1^i) \wedge (y_1 \in \mathcal{S}_2^i) \wedge (y_1 \in \mathcal{S}_3^i) | I_{y_1} = 1] \\ &\leq \frac{(s_1^i - 1)(s_2^i - 1)(s_3^i - 1)}{(2^b - 1)^3} .\end{aligned}$$

Again, $s_1^i, s_2^i, s_3^i < 2^b$ implies

$$\Pr[I_{y_1 \oplus w_2} = 1 | I_{y_1} = 1] \leq \mathbb{E}[I_{y_1 + w_2}] ,$$

and therefore, $\mathbf{Cov}\left[I_{y_1}, I_{y_1 \oplus w_2}\right] \leq 0$. Thus, it holds that $\mathbf{Cov}\left[I_{y_1}, I_{y_1'}\right] \leq 0$ over all cases of y_1', and it follows that

$$\begin{aligned}\mathbf{Var}\left[s_{1,2,3}^{i,w_1,w_2,w_3}\right] &\leq \sum_{y_1 \in \mathbb{F}_2^b} \mathbf{Var}\left[I_{y_1}\right] \\ &= 2^b \cdot \left(\frac{s_1^i s_2^i s_3^i}{2^{3b}} - \left(\frac{s_1^i s_2^i s_3^i}{2^{3b}}\right)^2\right) = \frac{s_1^i s_2^i s_3^i}{2^{2b}} - \frac{(s_1^i s_2^i s_3^i)^2}{2^{5b}} .\end{aligned}$$

C.3 Proof of Lemma 3

Lemma 3. Let $t \leq r$ and $\{I\} = \{j_1, \ldots, j_t\} \subseteq \{1, \ldots, r\}$. Then, it holds for the expectation and variance that

$$\mathbb{E}[s_{j_1, j_2, \cdots, j_t}^{i, w_{j_1}, w_{j_2}, \ldots, w_{j_t}}] = \frac{\prod_{j \in \{I\}} s_j^i}{2^{(t-1)b}}$$

$$\mathbf{Var}[s_{j_1, j_2, \cdots, j_t}^{i, w_{j_1}, w_{j_2}, \ldots, w_{j_t}}] = \frac{\prod_{j \in \{I\}} s_j^i}{2^{(t-1)b}} - \frac{\left(\prod_{j \in \{I\}} s_j^i\right)^2}{2^{(2t-1)b}} .$$

Proof. Given fixed $w_{j_2}, \dots w_{j_t} \in \mathbb{F}_2^b$, for each $y_1 \in \mathbb{F}_2^b$, we define Bernoulli variables I_{y_1} as

$$I_{y_1} \stackrel{\text{def}}{=} \begin{cases} 1 & y_1 \in \mathcal{S}_1^i \wedge y_1 \oplus w_{j_2} \in \mathcal{S}_{j_2}^i \wedge \dots \wedge y_1 \oplus w_{j_t} \in \mathcal{S}_{j_t}^i \\ 0 & \text{otherwise.} \end{cases}$$

Then, it holds that

$$\mathbb{E}\left[s^{i,w_{j_1},w_{j_2},\dots,w_{j_t}}_{j_1,j_2,\cdots,j_t}\right] = \mathbb{E}\left[\sum_{y_1 \in \mathbb{F}_2^b} I_{y_1}\right] = \sum_{y_1 \in \mathbb{F}_2^b} \mathbb{E}[I_{y_1}]\,.$$

Since the expectations for a fixed value $y_1 \in \mathbb{F}_2^b$ and its translations to be in the list of all three permutations are mutually independent, the probability is 2^{-tb}. Over all elements of the sets, it holds that

$$\mathbb{E}[I_u] = \frac{\prod_{j\in\mathcal{I}} s_j^i}{2^{tb}} \quad \text{and therefore} \quad \mathbb{E}\left[s^{i,w_{j_1},w_{j_2},\dots,w_{j_t}}_{j_1,j_2,\cdots,j_t}\right] = \frac{\prod_{j\in\mathcal{I}} s_j^i}{2^{(t-1)b}}\,.$$

For $\mathsf{x} = \sum_{y_1} \Pr[I_{y_1} = 1]$, as a sum of n independent Bernoulli variables I_{y_1}, with $\Pr[I_{y_1} = 1] = p$ for all y_1, it holds that

$$\mathbb{E}\left[\mathsf{x}^2\right] = \mathbb{E}\left[\left(\sum_{j=1}^{n} I_j\right)^2\right] = n(n-1)p^2 + np\,.$$

In our case, $n = 2^b$ and $p = \prod_{j\in\mathcal{I}} s_j^i \cdot 2^{-tb}$, for all $y_1 \in \mathbb{F}_2^b$. Given that $(\mathbb{E}[\mathsf{x}])^2 = (2^b p)^2$, we obtain

$$\mathbf{Var}\left[s^{i,w_{j_1},w_{j_2},\dots,w_{j_t}}_{j_1,j_2,\cdots,j_t}\right] \le \frac{\prod_{j\in\mathcal{I}} s_j^i}{2^{(t-1)b}} - \frac{\left(\prod_{j\in\mathcal{I}} s_j^i\right)^2}{2^{(2t-1)b}}\,.$$

C.4 Proof of Lemma 7

Lemma 7. Let $r \ge 3$ be integer. It holds for all $\ell = 2j$ for some $j \in [2..r-1]$ that

$$\frac{|k^i_{\ell+1}|}{|k^i_\ell|} \le 3r\,, \quad k^i_{\ell+1} \ge 0\,, \quad k^i_\ell \le 0 \quad \text{and} \quad k^i_{2r} \le 0\,.$$

Proof. First, we note that k^i_ℓ will be negative whereas $k^i_{\ell+1}$ will be positive, given that $\ell \ge 4$. We can write

$$k^i_\ell = \underline{k}^i_\ell + \overline{k}^i_\ell$$

We consider $\underline{k}^i_\ell$ first. To isolate those terms that contribute to the fixed ℓ, we can see from Equation 19 and 15 that $t_1 + t_2 = \ell$ must hold. Since we consider an even exponent $\ell = t_1 + t_2 = 2j$, those summands add to

$$\begin{aligned} \underline{k}^i_\ell &= -(-1)^{t_1+t_2}\binom{r}{2}\binom{r}{\ell-2} - \binom{r}{3}\binom{r}{\ell-3} - \dots - \binom{r}{\ell-2}\binom{r}{2} \\ &= -\left(\sum_{t_1=2}^{\ell-2}\binom{r}{t_1}\binom{r}{\ell-t_1}\right)\,. \end{aligned}$$

For odd $\ell+1$, the inverse holds, i.e., all terms in $\underline{k}^i_{\ell+1}$ will be positive:

$$\underline{k}^i_\ell = \left(\sum_{t_1=2}^{\ell-1} \binom{r}{t_1}\binom{r}{\ell+1-t_2}\right).$$

Next, we consider the summands that contribute to $\overline{k}^i_\ell$. From Lemma 6, we know that the factors for fixed t_1, t_2, r, and ℓ for $\overline{k}^i_\ell$ are

$$\overline{k}^i_{t_1,t_2,r} = \binom{r}{t_1}\binom{r}{t_1+t_2-\ell}\binom{r-t_1}{\ell-t_1}(-1)^{t_1+t_2}.$$

Over all $t_1, t_2 \in \{2,\ldots,\ell\}$ in Equation 13 and considering the correct signs, we obtain

$$\overline{k}^i_\ell = \sum_{t_1=2}^{\ell}\sum_{t_2=2}^{\ell}\left(\binom{r}{t_1}\binom{t_1}{t_1+t_2-\ell}\binom{r-t_1}{\ell-t_1}(-1)^{t_1+t_2}\right).$$

Note that values of $t_1, t_2 > \ell$ do not contribute since they have no terms in c^i that produce powers p^ℓ. We observe that $\overline{k}^i_\ell$ consists of summands of different sign. To gain clarity, we decompose and group those first according to $\binom{r}{t_1}$ and second to their sign.

$$\begin{aligned}\overline{k}^i_\ell = \sum_{t_1=2}^{\ell}\binom{r}{t_1}\binom{r-t_1}{\ell-t_1}(-1)^{t_1+t_2} \\ \left(\sum_{t_2=2,4,\ldots,\ell}\binom{t_1}{t_1+t_2-\ell} - \sum_{t_2=3,5,\ldots,\underline{\ell+1}}\binom{t_1}{t_1+t_2-\ell}\right) \qquad (24)\\ \overline{k}^i_{\ell+1} = \sum_{t_1=2}^{\ell+1}\binom{r}{t_1}\binom{r-t_1}{\ell+1-t_1}(-1)^{t_1+t_2+1} \\ \left(\sum_{t_2=2,4,\ldots,\ell+1}\binom{t_1}{t_1+t_2-(\ell+1)} - \sum_{t_2=3,5,\ldots,\ell+1}\binom{t_1}{t_1+t_2-(\ell+1)}\right). \qquad (25)\end{aligned}$$

Note that for odd $\ell+1$, the cardinalities of $|\{2,4,\ldots,\ell+1\}| = |\{3,5,\ldots,\ell+1\}| = \ell/2$. Though, while $|\{2,4,\ldots,\ell\}| = \ell/2$, $|\{3,5,\ldots,\ell\}| = \ell/2-1$. Since the term $\binom{t_1}{t_1+t_2-\ell} = \binom{t_1}{-1} = 0$ for $t_2 = \ell+1$, we were allowed to extend the underlined index in the rightmost sum in Equation (24) from ℓ to $\ell+1$ without changing the result. Then, we have $\ell/2$ terms in each difference and will be able to use another helping lemma.

For some set $\mathcal{I} \subseteq \mathbf{N}_0$, let $\mathcal{I}_e = \{i \in \mathcal{I} : i \text{ is even}\}$ and $\mathcal{I}_o = \{i \in \mathcal{I} : i \text{ is odd}\}$ denote the sets of even and odd non-negative numbers in $\mathcal{I}$. The following result is well-known.

Lemma 8. Let n be a non-negative integer. Then

$$\sum_{k\in[0..n]_e}\binom{n}{k} = \sum_{k\in[0..n]_o}\binom{n}{k} = \frac{2^n}{2}.$$

It follows that

$$\begin{aligned}\sum_{k\in[0..n]_o}\binom{n}{k} - \sum_{k\in[0..n]_e}\binom{n}{k} &= 0\\ \sum_{k\in[1..n]_o}\binom{n}{k} - \sum_{k\in[1..n]_e}\binom{n}{k} &= \binom{n}{0} = 1\\ \sum_{k\in[2..n]_o}\binom{n}{k} - \sum_{k\in[2..n]_e}\binom{n}{k} &= \binom{n}{0} - \binom{n}{1} = 1-n.\end{aligned}$$

First, we consider $\overline{k}^i_{\ell+1}$ with three cases.

Case $t_1 \leq \ell - 1$: From Equation (25), we see that for all even $t_1 \leq \ell - 1$, it holds that

$$\sum_{t_2=2,4,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_o} \binom{t_1}{k}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_e} \binom{t_1}{k}$$

and from Lemma 8

$$\sum_{t_2=2,4,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} - \sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = 0\,.$$

A similar statement can be derived for all odd $t_1 \leq \ell - 1$.

Case $t_1 = \ell$: For $t_1 = \ell$, we have

$$\sum_{t_2=2,4,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_o} \binom{t_1}{k}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_e} \binom{t_1}{k} - \binom{t_1}{0}\,.$$

Case $t_1 = \ell$: For $t_1 = \ell + 1$, it holds that

$$\sum_{t_2=2,4,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_e} \binom{t_1}{k} - \binom{t_1}{0}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-(\ell+1)} = \sum_{k\in[0..\ell+1]_o} \binom{t_1}{k} - \binom{t_1}{1}$$

We obtain that

$$\begin{aligned}
\overline{k}^i_{\ell+1} &= \binom{r}{\ell}\binom{r-\ell}{\ell+1-\ell}\binom{\ell}{0} - \binom{r}{\ell+1}\binom{r-(\ell+1)}{\ell+1-(\ell+1)}\left(\binom{\ell+1}{1} - \binom{\ell+1}{0}\right) \\
&= \binom{r}{\ell}(r-\ell) - \binom{r}{\ell+1}\ell \\
&= (\ell+1)\binom{r}{\ell+1} - \ell\binom{r}{\ell+1} \\
&= \binom{r}{\ell+1}\,.
\end{aligned}$$

Next, we consider $\overline{k}^i_{\ell}$ with three similar cases.

Case $t_1 \leq \ell - 2$: From Equation (25), we see that for all even $t_1 \leq \ell - 2$, it holds that

$$\sum_{t_2=2,4,\ldots,\ell} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell+1]_e} \binom{t_1}{k}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell+1]_o} \binom{t_1}{k}$$

and from Lemma 8

$$\sum_{t_2=2,4,\ldots,\ell} \binom{t_1}{t_1+t_2-\ell} - \sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-\ell} = 0\,.$$

A similar statement can be derived for all odd $t_1 \leq \ell - 2$.

Case $t_1 = \ell - 1$: For $t_1 = \ell - 1$, we have

$$\sum_{t_2=2,4,\ldots,\ell} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell]_o} \binom{t_1}{k}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell]_e} \binom{t_1}{k} - \binom{t_1}{0}\,.$$

Case $t_1 = \ell$: For $t_1 = \ell$, it holds that

$$\sum_{t_2=2,4,\ldots,\ell} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell]_e} \binom{t_1}{k} - \binom{t_1}{0}$$
$$\sum_{t_2=3,5,\ldots,\ell+1} \binom{t_1}{t_1+t_2-\ell} = \sum_{k\in[0..\ell]_o} \binom{t_1}{k} - \binom{t_1}{1}$$

We obtain that

$$\begin{aligned}
\overline{k}^i_\ell &= \binom{r}{\ell-1}\binom{r-(\ell-1)}{\ell-(\ell-1)}\binom{\ell-1}{0} - \binom{r}{\ell}\binom{r-\ell}{\ell-\ell}\left(\binom{\ell}{1} - \binom{\ell}{0}\right) \\
&= \binom{r}{\ell-1}(r-(\ell-1)) - \binom{r}{\ell}(\ell-1) \\
&= \ell\binom{r}{\ell} - (\ell-1)\binom{r}{\ell} \\
&= \binom{r}{\ell}\,.
\end{aligned}$$

Now, we can insert our terms to bound our desired ratio

$$\begin{aligned}
\frac{k^i_{\ell+1}}{k^i_\ell} &= \frac{\underline{k}^i_{\ell+1} + \overline{k}^i_{\ell+1}}{\underline{k}^i_\ell + \overline{k}^i_\ell} \\
&= \frac{\left(\sum_{t_1=2}^{\ell-1}\binom{r}{t_1}\binom{r}{\ell+1-t_1}\right) + \binom{r}{\ell+1}}{-\left(\sum_{t_1=2}^{\ell-2}\binom{r}{t_1}\binom{r}{\ell-t_1}\right) - \binom{r}{\ell}} \\
&= \frac{\overbrace{\left(\sum_{t_1=2}^{\ell-2}\binom{r}{t_1}\binom{r}{\ell+1-t_1}\right)}^{a} + \overbrace{\binom{r}{\ell-1}\binom{r}{2}}^{b} + \overbrace{\binom{r}{\ell+1}}^{c}}{-\underbrace{\left(\sum_{t_1=2}^{\ell-2}\binom{r}{t_1}\binom{r}{\ell-t_1}\right)}_{d} - \underbrace{\binom{r}{\ell}}_{e}}\,.
\end{aligned}$$

Thus, we have shown the positivity and negativity statements from Lemma 7:

$$\begin{aligned}
k^i_{\ell+1} &\geq 0 \\
k^i_\ell &\leq 0 \\
k^i_{2r} &\leq 0\,.
\end{aligned}$$

It remains to obtain upper bound the ratio of their absolutes. We can use

$$\frac{a+b+c}{d+e} \leq \frac{a}{d+e} + \frac{b}{d+e} + \frac{c}{d+e} \leq \frac{a}{d} + \frac{b}{d} + \frac{c}{e}\,.$$

We can see that

$$\frac{a}{d} = \frac{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell+1-t_1}}{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}} = \frac{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}\frac{r-(\ell-t_1)}{\ell-t_1+1}}{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}} \leq \frac{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}\frac{r-2}{3}}{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}} \leq \frac{r}{3}\,.$$

Similarly, for all $\ell \geq 4$:

$$\frac{b}{d} = \frac{\binom{r}{\ell-1}\binom{r}{2}}{\sum_{t_1=2}^{\ell-2} \binom{r}{t_1}\binom{r}{\ell-t_1}} \leq \frac{\binom{r}{2}\binom{r}{\ell-1}}{\binom{r}{2}\binom{r}{\ell-2}} \leq \frac{\binom{r}{2}\binom{r}{\ell-2}\frac{r-(\ell-2)}{\ell-1}}{\binom{r}{2}\binom{r}{\ell-2}} \leq \frac{r}{3}\,.$$

Finally, for all $\ell \geq 4$, it holds

$$\frac{c}{e} = \frac{\binom{r}{\ell+1}}{\binom{r}{\ell}} \leq \frac{\binom{r}{\ell}\frac{r-\ell}{\ell+1}}{\binom{r}{\ell}} \leq \frac{r}{\ell+1} \leq \frac{r}{5}\,.$$

From the sum of the three bounds, we obtain our claim for all $\ell = 2j$ and $j \in [2..r-1]$:

$$\frac{|k^i_{\ell+1}|}{|k^i_\ell|} \leq 3r\,.$$

D Lessons from Related AES-round-based Block Ciphers

Lessons from AES-PRF. Derbez et al. studied the security of generalized variants of AES-PRF and its dual. The original instantiations have five rounds in each of their respective permutations. For variants AES-PRF-r_1-r_2, Derbez et al.'s works could attack up to AES-PRF-$*$-4 rounds with zero-correlation cryptanalysis and AES-PRF-dual-4-$*$. Thus, the security margin seems to be one round for those variants. Note that AES-PRF and its dual are untweaked and a tweak may introduce even further attack angles. As a takeaway, *we need at least six rounds in each bottom-permutation call* to achieve our desired security margins.

Lessons from ForkCipher Instances. ForkAES has seen third-party cryptanalysis [7, 11] after its proposal. The former showed rectangle and impossible-differential attacks using only the encryption direction on ForkAES-$*$-4-4, i.e., with one round less on each of the bottom permutations; The cipher was broken with reflection queries by [11]; Jiang and Jin [68] proposed another attack on ForkAES-$*$-5-4 with the help of multiple impossible differentials. We note that our applications prohibit reflection queries, which makes attacks with chosen reflection queries inapplicable. Moreover, our construction renders partial decryption infeasible or at least considerably harder since the adversary would have to guess parts of the internal states. We conclude from the chosen-plaintext attacks on ForkAES-$*$-4-4 that *we need at least six rounds in each bottom permutation to have at least two rounds of margin.* Moreover, we see *the branch constants of ForkAES as an effective means to make inter-branch differentials harder to exploit.*

Lessons from Kiasu-BC. Kiasu-BC [66] provided a baseline of adding tweaks to the AES. Third-party cryptanalysis showed that integral [46], MitM [73, 93], and differential-based attacks [47] can exploit the degrees of freedom from tweak inputs to extend several attacks by one round – i.e., key-recovery attacks can cover eight rounds, compared to the best

Table 5: Existing key-recovery attacks (no related keys) on round-reduced AES-like block ciphers. r = #rounds, t = #used tweaks, Mem. = memory complexity, ref. = reference, int. = integral, diff. = differential, BD/ID = biased/impossible differential, MitM = meet-in-the-middle, ZC = zero-correlation linear, rect. = rectangle, boom. = boomerang, CP/CC = chosen plaintexts/chosen ciphertexts. (*) Distinguisher found to have probability zero in [94].

(a) On AES-128.

r	Type	Time	Data	Mem.	Ref.
AES, single-key key-recovery					
6	Int.	$2^{51.7}$	$2^{34.6}$ CP	2^{32}	[92]
7	ID	$2^{110.2}$	$2^{106.2}$ CP	$2^{90.2}$	[76]
7	MitM	2^{99}	2^{97} CP	2^{98}	[44]
AES, single-key distinguishers					
4	Int.	2^{37}	2^{32} CP	2^{32}	[40]
4	ID	$2^{110.2}$	$2^{106.2}$ CP	$2^{90.2}$	[76]
4	MitM	2^{99}	2^{97} CP	2^{98}	[44]
6	Diff	$2^{96.5}$	$2^{89.5}$ CP	2^{36}	[8]
6	Mixture	$2^{88.2}$	$2^{88.2}$ CP	$2^{88.2}$	[10]

(b) On AES-PRF and its dual.

r	Type	Time	Data	Mem.	Ref.
AES-PRF					
2-$*$	ID	2^{94}	2^{94} CP	2^{88}	[45]
$*$-4	ZC	$2^{96.95}$	$2^{96.95}$ KP	2^{64}	[45]
s-$(7-s)$	MitM	2^{107}	2^{107} CP	2^{104}	[45]
Dual-AES-PRF					
$*$-2	ID	2^{104}	2^{104} CP	2^{72}	[45]
$*$-2	ZC	$2^{115.14}$	$2^{115.06}$ KP	2^{65}	[45]
3-$*$	Diff.	2^{97}	2^{97} CP	2^{32}	[45]
4-$*$	Diff.	2^{121}	2^{121} CP	2^{8}	[45]

(c) On Kiasu-BC (related tweaks).

r	Type	Time	Data	Mem.	Ref.
7	Int.	$2^{48.5}$	$2^{43.6}$ CP	–	[46]
8	ID	2^{118}	2^{118} CP	–	[47]
8	MitM	2^{116}	2^{116} CP	2^{86}	[93]
8	MitM	$2^{112.8}$	2^{109} CP	$2^{92.9}$	[73]
8	Boom.	2^{83}	2^{83} CC	–	[12]

(d) On TweAES (related tweaks).

r	t	Type	Time	Data	Mem.	Ref.
5	2	Diff.	2^{26}	2^{5} CP	$2^{28.6}$	[29]
6	16	Int.	2^{45}	2^{5} KP	negl.	[29]
7	16	Boom.	2^{125}	2^{125} CP	negl.	[29](*)
8	2	ID	$2^{124.4}$	$2^{124.3}$ CP	$2^{118.8}$	[80]

(e) On ForkAES.

Type	r	Type	Time	Data	Mem.	Ref.
ForkAES						
$*$-4-4	8	ID	2^{47}	$2^{39.5}$ CP	2^{35}	[7]
$*$-4-4	8	RD	2^{35}	2^{35} CP	2^{33}	[7]
$*$-5-4	8	ID	$2^{118.2}$	$2^{111.4}$ CP	$2^{92.7}$	[68]
$*$-5-5	10	Diff.	2^{125}	2^{119} CR	2^{83}	[11]

single-key attacks that cover seven out of ten rounds on AES-128. We observe that Kiasu-BC has only one, four, and eight active S-boxes over two, three, and four rounds, respectively, which may open up the gates for rectangle distinguishers. We consider this a crucial attack vector that we have to close for our construction and aim for more active S-boxes over a few rounds. The recent further improvements of boomerang attacks by Bariant and Leurent on Kiasu-BC and TNT-AES [12] emphasize that boomerang attacks could also pose a similar threat to our constructions. The attack on TNT-AES is not directly applicable to our setting since the tweak differences could be chosen arbitrarily in the tweak state. Though, we observe that the cipher must remain secure also against combinations of trails with few rounds for the top and bottom trails. To conclude, our proposal *must ensure sufficiently many active S-boxes for two to four rounds.*

Lessons from TweAES. TweAES applies a lightweight linear code to expand the tweak; thus, a tweak difference activates at least three cells from any non-zero tweak difference and prevents cancellations of a tweak-injected difference in the subsequent tweak addition. The design strategy injects the tweak after only every second round, which ensures at least 15 active S-boxes in four subsequent rounds with a tweak addition in the middle. Thus, it activates more S-boxes than Kiasu-BC (see Table 3). For the construction, there exist longer distinguishers than for the AES, e.g. impossible differentials can cover up to six rounds. While the designers proposed impossible-differentials attacks on up to six and possibly boomerangs on up to $1+3+3+1$ rounds, Niu et al. [80] showed that the latter attack was invalid; though, they proposed an alternative impossible-differential-based key

recovery on eight rounds in the same work. Thus, it seems that TweAES is on par with Kiasu-BC in terms of covered rounds in attacks.

TweAES opens several interesting questions. The linear code seems to be one instance out of various variants. Since we do not target being lightweight for an AES-based instantiation, little overhead compared to TweAES is less relevant. *The linear code ensures many active S-boxes effectively.* We see that the tweak injection only after every second round yields too few – only 0, 4, and 15 active S-boxes over two, three, and four rounds, respectively. Thus, while four rounds offer sufficient protection to thwart boomerang-based attacks, attacks that can combine two trails of less than four rounds may become a threat. Though, since we consider the bottom-permutation calls, we can *add the tweak addition before its first round to increase security.* Moreover, we can investigate what strategy of tweak injections – each round vs. after every second round – will be more effective against attacks.