♯Pencil: A Domain-Extended Committing BBB PRF for Strengthening GCM*

Ritam Bhaumik Jean Paul Degabriele Chandranan Dhar

TII, Abu Dhabi, UAE

bhaumik.ritam@gmail.com jeanpaul.degabriele@tii.ae chandranandhar@gmail.com

Abstract

We consider the problem of constructing efficient committing pseudorandom functions with Beyond-Birthday-Bound (BBB) security from blockciphers. More specifically, we are interested in expanding pseudorandom functions (PRF) whose domain is roughly twice that of the underlying blockcipher. The main motivation behind our work is to construct an AES-based key derivation function (KDF) that can be combined with GCM in order to improve its security bound, accommodate larger nonces that can be generated randomly without risking collisions, and make it a committing AEAD scheme. NIST has recently announced a pre-draft call for comments to standardise AEAD schemes that can encrypt larger amounts of data and admit larger nonces. The call lists two approaches. The first is to define an analogue of GCM using a 256-bit blockcipher, and the second is based on a recent proposal by Gueron, to extend GCM with a key derivation function (KDF) called DNDK to increase its security. The latter has clear benefits in terms of backwards compatibility and is likely to be the preferred interim solution. Moreover, DNDK-GCM is already deployed in production at Meta. DNDK is essentially a BBB-secure expanding weak pseudorandom function with a domain size of 192 bits realised from AES. We here propose an alternative AES-based KDF called ♯Pencil with comparable efficiency but stronger provable security guarantees. Specifically, ♯Pencil is a full PRF, whereas DNDK is only a weak PRF, which allows us to prove the ♯Pencil-GCM composition secure as an AEAD scheme. Our most technically challenging result is to show that ♯Pencil is committing, whereas DNDK claims this property without proof. Finally, in contrast to DNDK and other alternatives, ♯Pencil can be safely used with arbitrary (non-random) nonces and remains committing even when the nonce is not included as part of the ciphertext.

*An extended abstract of this paper will appear in the proceedings of CRYPTO 2026.

Contents

1 Introduction

Authenticated Encryption with Associated Data (AEAD) is the understated pillar of cryptography—often considered uninteresting or easy, yet essential for virtually all of its applications. While it was already deployed in major cryptographic protocols like SSL (TLS), SSH, and IPsec, its foundations can be traced back to the seminal works of Bellare and Namprempre [BN00], Jutla [Jut01], and Rogaway [Rog02]. Yet, some 25 years later, it remains an important area of research, as new practical challenges continue to emerge, and we continue to deepen our understanding of this fundamental primitive. Besides security, a crucial property of this primitive is its performance, which is why it is found in so many applications in the first place. Today, the most popular AEAD scheme is Galois Counter Mode (GCM), which typically offers the best performance [MV04]. Its superior performance derives in large part from the native AES and carry-less multiplication instructions now supported in most CPUs. The ubiquity of this hardware support, together with its standardisation by NIST, has cemented GCM as the go-to AEAD scheme and is now found practically everywhere.

Of course, GCM also has its limitations, which have become more noticeable in recent years. Notably, in cloud environments, it must be rekeyed every couple of seconds due to the high volumes of data being processed [KCC+23]. Moreover, many practical applications require AEAD schemes that can admit a randomly-generated nonce, because maintaining state is costly, cumbersome, and even unsafe in distributed systems. However, in GCM, nonces are effectively limited to 96 bits[1]. Thus, for randomly generated nonces and a 32-bit security level, GCM can encrypt at most 2^{32} messages with a single key. It should also be noted that GCM breaks down completely as soon as a single nonce value repeats. In addition, GCM provides only birthday-bound security [IOM12], which, in turn, limits the total number of data blocks that can be encrypted under a single key. Finally, commitment security has recently been identified as an important security property for AEAD, which is necessary in several applications [GLR17, LGR21, ADG+22]. Unfortunately, GCM also lacks this property.

NIST has recently initiated multiple standardisation efforts to address GCM's limitations. One is to standardise an AES-based Accordion mode, which would automatically yield AEAD with superior security, such as nonce-misuse resistance [RS06] and security against the release of unverified plaintext [ABL+14]. However, an Accordion mode is inherently a two-pass primitive and thus cannot process its inputs in an online manner. For most Internet and cloud applications, the ability to process data online is crucial, and as such, an Accordion mode is unlikely to replace GCM in these settings. Accordingly, NIST has also published a pre-draft call for comments, recognising the need to standardise efficient AEAD schemes that support larger nonces and encrypt larger amounts of data under the same key [Nat25]. This call lists two possible approaches to address this, the first is to define an analogue of GCM using a 256-bit blockcipher, and the second is to augment GCM with an AES-based key derivation function (KDF) as proposed earlier by Gueron in DNDK-GCM [Gue24b, Gue24a]. In the latter approach, the KDF (DNDK) effectively acts as a wrapper for GCM, transforming it into an AEAD scheme (DNDK-GCM) with a larger nonce size and enhanced quantitative and qualitative security. The latter approach has the benefit of being backwards compatible with existing AES-GCM deployments without needing to introduce a new blockcipher, which is a more pragmatic and timely solution for existing deployments. Indeed, DNDK-GCM is already deployed in production at Meta [Gue24b]. In this work, we propose ♯Pencil-GCM as an improved alternative to DNDK-GCM, offering stronger provable security.

Augmenting AEAD with a KDF to improve its security is not new, and was in fact already used in the design of AES-GCM-SIV to attain a better security bound [GL17, GLL17]. In DNDK-GCM, however, the construction of the KDF (DNDK) is more elaborate in order to: (a) admit a larger nonce to accommodate randomly generated ones, and (b) provide Beyond-Birthday-Bound (BBB) security so as to avoid

[1] Technically, GCM admits nonces larger than 96 bits, but they are hashed internally to a 96-bit value, thereby resulting in the same collision probability as a 96-bit nonce.

introducing KDF-related birthday terms in the security bound. However, its BBB security only holds when the nonce is sampled uniformly at random [Gue24a]. In essence, DNDK can be understood as an expanding weak PRF construction consisting of two instances of the XORP [Iwa06] construction evaluated over independently sampled random inputs and combined in an XOR fashion using the result of Bellare, Goldreich, and Krawczyk [BGK99]. Combining the results of Iwata, Mennink, and Vizar [IMV16] and Bellare, Goldreich, and Krawczyk [BGK99] it follows that DNDK achieves $\mathcal{O}(q/2^n)$ security as a weak PRF. In terms of efficiency, DNDK requires 10 AES calls to produce 4 blocks of output, where the AES calls are parallelizable. However, the security of DNDK-GCM as a composite AEAD scheme is not supported by a security proof, and in fact, it seems rather unlikely that such a proof is even possible. DNDK is only known to be secure as a weak PRF, but in the AEAD security game, the adversary can make decryption queries on any (non-random) nonce of its choice, and accordingly, the derived GCM key for decryption is not necessarily random, thus precluding a reduction to the security of GCM and DNDK. In addition, DNDK is claimed to be key committing, but no proof or security bound is provided.

1.1 Contribution

Our work is primarily motivated by the use case of DNDK, where we propose an alternative AES-based KDF construction to strengthen GCM and obtain a stateless AEAD with BBB security and key-commitment security. From a more theoretical perspective, it also addresses a rather understudied problem: that of constructing pseudorandom functions from blockciphers with extended domains and BBB security. Our proposed construction, ♯Pencil, improves over DNDK in that it is a full-fledged PRF and thus results in a provably-secure AEAD scheme when combined with GCM. It also means that it can be operated with non-random nonces, whereas DNDK is insecure in such a setting. It requires two fewer AES calls than DNDK in order to generate a four-block output, and it provides a stronger form of key-commitment security than DNDK, i.e., CMT-1 instead of FOR-KC. More specifically, our contribution is as follows.

A New Domain-Extended BBB PRF. Our first contribution is ♯Pencil, a new PRF construction from a an n-bit blockcipher taking $(2n-16)$-bit inputs yielding full n-bit security. It is realised as the composition of two components, Sharp (Fig. 1, left) and Pencil (Fig. 1, right). Intuitively, Pencil can be understood as a BBB expanding weak PRF construction on extended inputs. Specifically, $\mathsf{Pencil}_{K_2}(R_1, R_2, 0^{n-8}, 0^{n-8})$ is an expanding weak PRF over the $(2n-16)$-bit input (R_1,R_2). Compared to DNDK, another weak PRF with comparable parameters, Pencil requires only half the number of blockcipher calls, while retaining the same optimal (full n-bit) security bound. Pencil adapts the XORP construction to take two random blocks of input instead of one block. For this to work, we introduce a preprocessing step on the PRF input to generate the values X_1 to X_5. This preprocessing requires minimal overhead: five binary-field multiplications by 2, which can be realised very efficiently using one shift and one XOR operation. To obtain a full-fledged BBB PRF, we augment Pencil with a $(2n-16)$-to-$(4n-32)$-bit component called Sharp consisting of three additional blockcipher calls with a separate key to preprocess the PRF input. We prove the PRF security of ♯Pencil in Sec. 3, attaining a bound of $q/2^{n-11} + nq/2^{n-4}$.

Composition with CAU/GCM. We next consider using ♯Pencil as a KDF in combination with the CAU AEAD scheme (a generalisation of GCM [BT16]) to yield a composite AEAD scheme ♯Pencil-CAU analogous to DNDK-GCM. However, in contrast to DNDK-GCM, which lacks a security proof as an AEAD scheme, we show that ♯Pencil-CAU is provably secure as an AEAD scheme with a significantly better security bound than plain CAU/GCM. To prove its security, we utilise the multi-user security analysis of CAU/GCM by Hoang, Tessaro, and Thiruvengadam [HTT18]. In ♯Pencil-CAU, the key derivation adds only 8 additional blockcipher calls per encryption while attaining a similar bound to that claimed (without proof) by Gueron for DNDK-GCM [Gue24a]. In addition, note that ♯Pencil-CAU attains this bound for

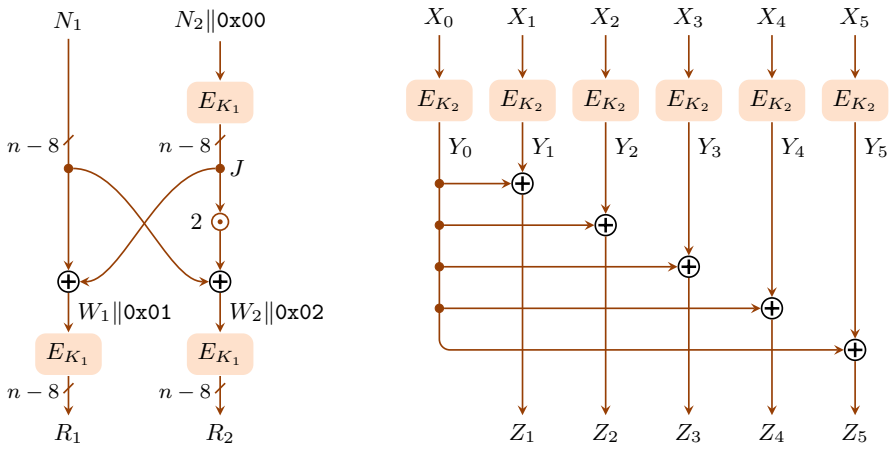

Figure 1: Illustration of the blockcipher-based constructions $\mathsf{Sharp}_{K_1}(N_1, N_2)$ (**left**) and $\mathsf{Pencil}_{K_2}(R_1, R_2, W_1, W_2)$ (**right**), together they yield $\sharp\mathsf{Pencil}_K(N_1, N_2)$. Here $X_j = (R_1 \oplus 2^j R_2 \oplus \mathbf{1}_{j\geq 4} W_{j-3}) \| \langle j \rangle_8$ where $\mathbf{1}$ is the indicator function, E_K is a blockcipher of blocksize n and keysize $2n$, and K_1 and K_2 are derived from a master key K as $K_1 \leftarrow E_K(\langle 0 \rangle_n) \| E_K(\langle 1 \rangle_n)$ and $K_2 \leftarrow E_K(\langle 2 \rangle_n) \| E_K(\langle 3 \rangle_n)$.

arbitrary nonces and not just random nonces as in the case of DNDK-GCM and other proposals. Finally, our KDF-AEAD composition diverges from DNDK-GCM in that we derive a random nonce-key pair for each message, whereas the latter fixes the GCM nonce to a constant.

Commitment Security of ♯Pencil. Our final contribution is to show that ♯Pencil is committing, and consequently ♯Pencil-CAU is also committing. This is arguably our most novel and technically challenging result. Note that the values (W_1, W_2) in the Pencil construction can be zeroed out, and Pencil and ♯Pencil would still be a BBB weak PRF and a BBB PRF, respectively. Indeed, their sole purpose in the construction is to make it a committing PRF. We consider two security notions, FOR-KC [FOR17], where the adversary is restricted to collisions resulting from the same nonce, and the more general CMT-1 [BH22] notion where the adversary is allowed to produce collisions over distinct nonces. While all other proposals focus on the former notion, the latter property is required in applications where the nonce is not included in the ciphertext. Such applications include scenarios where the nonce is generated in a stateful manner (e.g., a counter) to reduce bandwidth. We prove full n-bit FOR-KC security and $3n/4$-bit CMT-1 security, where these bounds transfer directly to the AEAD commitment security of ♯Pencil-CAU.

1.2 Other Related Work and Proposals

Since the publication of DNDK, two other concurrent proposals have emerged, which we now discuss and compare. Kampanakis et al. [KHEC25] have proposed KC-XAES, which augments Valsorda's earlier design XAES [Val24] with key commitment security. Both schemes use the FIPS-compliant construction CMAC as a KDF. Gueron and Ristenpart [GR25] have proposed a significantly different version of DNDK, named DNDKv2, that uses only a single instance of XORP and a different KDF-GCM composition approach together with an accompanying provable security analysis. Both proposals combine the KDF with GCM by feeding half of the double-block nonce into the KDF to generate the GCM key and then using the other half as the GCM nonce directly. A comparison of the security of all proposals is shown in Figure 2, which we elaborate on below.

In contrast to its predecessor, DNDKv2 only requires 5 blockcipher calls and is accompanied by a security analysis on the composition with GCM. However, we note that ♯Pencil provides better security

Scheme	$AEAD	AEAD	FOR-KC	CMT-1
AES-GCM	$q^2/2^{96} + \sigma^2/2^{128}$	$\sigma^2/2^{128}$	Insecure	Insecure
DNDK-GCM	Unknown	DNDK Insecure	Unknown	Unknown
DNDKv2-GCM	$q^3/2^{240} + qL^2/2^{127}$	$\sigma^2/2^{128}$	$q/2^n$	Unknown
XAES	$qL^2/2^{126}$	$\sigma^2/2^{128}$	Insecure	Insecure
KC-XAES	$\geq q^2/2^{128}$	$\geq q^2/2^{128}$	$q/2^n$	Insecure
♯Pencil-GCM	$q/2^{113} + \sigma L/2^{127}$	$q/2^{113} + \sigma L/2^{127}$	$q/2^n$	$q^4/2^{3n}$

Figure 2: Security comparison of the different KDF-GCM proposals when instantiated with AES-256, displaying only dominant terms in the security bounds. AEAD and $AEAD denote AEAD security in the standard nonce-respecting and restricted random-nonce settings, respectively. Here q denotes the number of encrypted messages, σ the total number of encrypted message blocks, and L the maximum message size in blocks.

when encrypting messages of varying lengths or a large number of short messages. Moreover, whereas ♯Pencil imposes no restriction on how nonces are generated, DNDKv2 provides no security advantage over plain GCM when nonces are not random. Finally, DNDKv2 is only committing with respect to fixed nonces (FOR-KC), and is insecure when nonces are under adversarial control (CMT-1).

Moving on to the security of XAES and KC-XAES, a standard analysis would introduce birthday terms due to the KDF (CMAC) producing non-uniform key blocks. However, in [BHKK25], the authors present a more elaborate analysis yielding a security bound of roughly $qL^2/2^{126}$ for the multiuser security GCM when taking into account the non-uniformity of the key blocks. This analysis assumes stronger security on the part of GCM, which the authors argue to be a mild assumption. It is not immediately clear how this multiuser security bound translates to a security bound for XAES and KC-XAES as AEAD schemes with random nonces. In [KHEC25] the designers of KC-XAES quote this result to argue the AEAD security of XAES, and seem to implicitly assume that it also holds for KC-XAES. However, in Appendix F we show an attack that contradicts this claim for the case of KC-XAES. Thus, in comparison to ♯Pencil-GCM, KC-XAES provides weak AEAD security, while XAES provides no committing security. Similarly to DNDKv2, the AEAD security of XAES and KC-XAES reverts back to that of plain GCM when nonces are not generated randomly, and KC-XAES is only committing with respect to fixed nonces (FOR-KC).

Since the first version of this paper [BD25] was uploaded to ePrint, two parallel works have been published which proposed designs similar to our weak PRF Pencil, albeit with different goals. Chen et al. [CDJN25] have proposed a weak PRF SnowFlake and a nonce-randomisation function F^* which they combine to build a deterministic encryption scheme. Chung et al. [CHK+25] have proposed a weak PRF eCTR and a strong PRF HteC as a BBB alternative to the counter-mode component used in AES-GCM and AES-GCM-SIV.

1.3 The Need for a Blockcipher-Based KDF

One may naturally wonder why a blockcipher-based KDF is needed—why can't we simply use a 'more standard' hash-based KDF instead? Firstly, the NIST call is looking to augment NIST SP 800-38D (GCM), which is a standard for AES *modes of operation*. Thus, the inclusion of the KDF must result in a composition that retains an AEAD interface and is based solely on AES. In addition, it is important to note that for many industry applications, adhering to FIPS (NIST) compliance is of paramount importance. Secondly, having a solution that is purely AES-based means that it can benefit from the AES-NI instruction set, which is now widely supported on a variety of hardware platforms. On the other hand, SHA-2

instructions are not as widely supported on legacy hardware, and there is no hardware support for SHA3. Thirdly, it is cleaner both conceptually and practically to have an AEAD scheme that is based on a single hardness assumption and a single primitive.

2 Preliminaries

For $m \leq n$, we write $[m..n]$ to denote the range $\{m, \ldots, n\}$. We use the Pochhammer falling factorial power notation $(n)_m := n(n-1)\ldots(n-m+1)$. For a finite set $\mathcal{X}$ we write $R \leftarrow\!\$\ \mathcal{X}$ to denote R being uniformly sampled from $\mathcal{S}$. For sets $\mathcal{X}$ and $\mathcal{Y}$, $\mathsf{Func}(\mathcal{X}, \mathcal{Y})$ will denote the set of all functions $\mathcal{X} \to \mathcal{Y}$, and $\mathsf{Perm}(\mathcal{X})$ will denote the set of all permutations over $\mathcal{X}$.

For an $x \leq 2^d$, $\langle x \rangle_d$ will denote the d-bit encoding of x. $x \| y$ will denote the concatenation of two bit-strings x and y. $|x|_b$ will denote the length of x in bits, $|x|_{\mathrm{B}}$ will denote the length of x in bytes, and $|x|_n$ will denote its length in n-bit blocks. $\mathsf{lsb}_d(x)$ and $\mathsf{msb}_d(x)$ will respectively denote the right-most and left-most d-bit substring of x. $x \ll d$ denotes the left-shift of x by d bits.

Bold face capital letters will usually denote vectors or matrices over $\mathbb{GF}(2^d)$ for some bit-length d. Unless there is scope for confusion, we'll not make a distinction between a field element $x \in \mathbb{GF}(2^d)$ and its d-bit encoding $\langle x \rangle_d \in \{0,1\}^d$, and we'll treat these two sets interchangeably.

Field Arithmetic. We will work with $\mathbb{GF}(2^{128})$, with the field polynomial $f(t) = t^{128} + t^7 + t^2 + t + 1$. This field has the convenient property that for any $x \in \{0,1\}^{120}$, $\mathsf{msb}_{120}(2 \cdot (x \| \mathtt{0x00})) = x \ll 1$. We will use this for the shorthand notation $2^j \cdot x := x \ll j$ for $x \in \{0,1\}^{120}$ and any positive integer j (though we will use it for only up to $j \leq 5$). By extension, $3 \cdot x$ will denote $2 \cdot x \oplus x$.

Blockciphers and Ideal Cipher. A blockcipher is a function $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ such that for any $K \in \mathcal{K}$, $E_K(\cdot) := E(K, \cdot)$ is a permutation on $\{0,1\}^n$. K is called the key of E. Let $\mathcal{BC}$ be the set of all blockciphers from $\mathcal{K} \times \{0,1\}^n$ to $\{0,1\}^n$. An ideal cipher from $\mathcal{K} \times \{0,1\}^n$ to $\{0,1\}^n$ is an element from $\mathcal{BC}$ chosen uniformly at random. In security games involving a mode built on top of an ideal cipher E, we allow the adversary make oracle queries to E and E^{-1}.

Adversarial Advantage. For any game G defined with respect to some primitive F and any adversary $\mathcal{A}$ we denote the adversary's corresponding advantage by $\mathbf{Adv}^{\mathsf{G}}_F(\mathcal{A})$. We use $\mathcal{A}^{\mathsf{G}} \Rightarrow b'$ to denote the event where the adversary interacts with game G and returns the bit b' upon terminating the game.

PRF Security. For a keyed function $F : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$, the PRF security is defined via the game in Fig. 3 (right). The adversary $\mathcal{A}$ has access to a single oracle FUNC. The PRF advantage of $\mathcal{A}$ against F is defined as

$$\mathbf{Adv}^{\mathsf{PRF}}_F(\mathcal{A}) := \left| \Pr[\mathcal{A}^{\mathsf{PRF}_F(1)} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathsf{PRF}_F(0)} \Rightarrow 1] \right|,$$

2.1 Authenticated Encryption with Associated Data (AEAD)

AEAD syntax. A nonce-based AEAD scheme $\Pi = (\Pi.\mathsf{Enc}, \Pi.\mathsf{Dec})$ consists of the following pair of algorithms:

- A deterministic encryption algorithm $\Pi.\mathsf{Enc} : \mathcal{K} \times \mathcal{N} \times \mathcal{AD} \times \mathcal{M} \to \mathcal{C}$ takes as input a secret key $K \in \mathcal{K}$, a nonce $N \in \mathcal{N}$, associated data $AD \in \mathcal{AD}$, and a message $M \in \mathcal{M}$ and returns a ciphertext $C \in \mathcal{C}$. We require $\Pi.\mathsf{Enc}$ to have constant expansion, i.e. for any $(K, N, AD, M) \in (\mathcal{K}, \mathcal{N}, \mathcal{AD}, \mathcal{M})$, the expansion $\tau = |\Pi.\mathsf{Enc}_K(N, AD, M)|_b - |M|_b$ is constant.

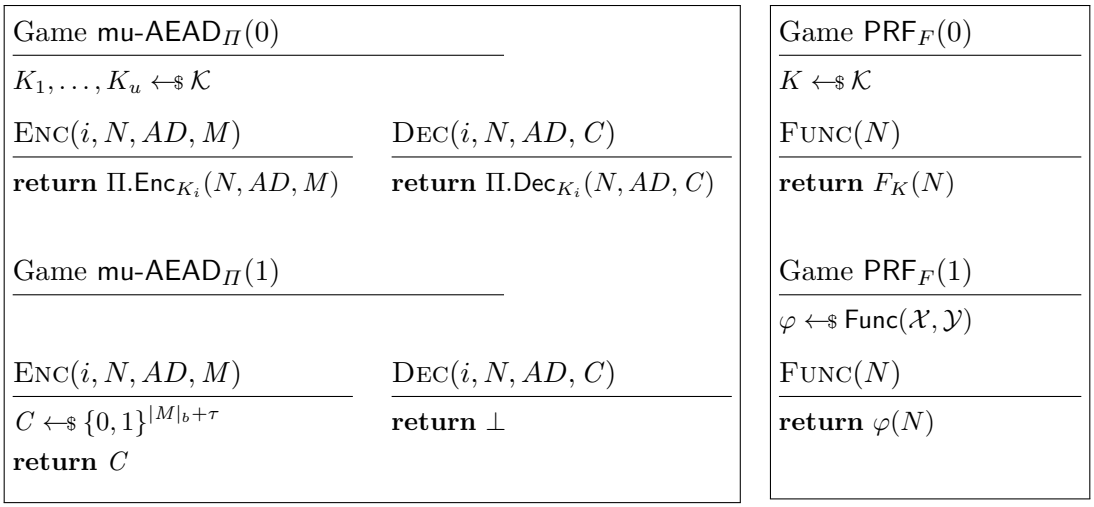

Figure 3: Games defining mu-AEAD security (left) and PRF security (right).

- A deterministic decryption algorithm $\Pi.\mathsf{Dec} : \mathcal{K} \times \mathcal{N} \times \mathcal{AD} \times \mathcal{C} \to \mathcal{M} \cup \{\perp\}$ takes as input a secret key $K \in \mathcal{K}$, a nonce $N \in \mathcal{N}$, associated data $AD \in \mathcal{AD}$, and a ciphertext $C \in \mathcal{C}$ and returns either a message $M \in \mathcal{M}$ or the symbol $\perp$ to indicate an invalid ciphertext.

Multi-User AEAD security (mu-AEAD). Multi-user AEAD security is defined via the game described in Fig. 3 (left). The adversary $\mathcal{A}$ is given access to multiple keyed instances of an AEAD scheme. Each instance consists of an encryption oracle $\textsc{Enc}(i, N, AD, M)$ and a decryption oracle $\textsc{Dec}(i, N, AD, C)$. An adversary is considered valid if it never makes a decryption query $\textsc{Dec}(i, N, AD, C)$ with C being the output of a previous encryption query $\textsc{Enc}(i, N, AD, M)$. Moreover an adversary is said to be nonce-respecting if it never repeats the pair (i, N) across encryption queries. The mu-AEAD advantage of $\mathcal{A}$ against Π is defined as

$$\mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\Pi}(\mathcal{A}) := \left| \Pr[\mathcal{A}^{\mathsf{mu\text{-}AEAD}_{\Pi}(1)} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathsf{mu\text{-}AEAD}_{\Pi}(0)} \Rightarrow 1] \right|,$$

The single user setting corresponds to the special case when $u = 1$, and in this setting, we denote the AEAD advantage of $\mathcal{A}$ to be $\mathbf{Adv}^{\mathsf{AEAD}}_{\Pi}(\mathcal{A})$.

2.2 Committing Security

Here we define two different notions of key-committing security of an AEAD which are considered of practical relevance; in addition, we define notions to capture the suffix-committing security of a PRF.

FOR-KC security. In the FOR-KC game against an AEAD scheme Π, an adversary $\mathcal{A}$ outputs (K, N, AD, M) and (K', N', AD', M'); $\mathcal{A}$ wins if $K \neq K'$, $N = N'$, and $\Pi.\mathsf{Enc}_K(N, AD, M) = \Pi.\mathsf{Enc}_K(N', AD', M')$. We write $\mathbf{Adv}^{\mathsf{FOR\text{-}KC}}_{\Pi}(\mathcal{A})$ to denote the probability that $\mathcal{A}$ wins.

CMT-1 security. In the CMT-1 game against an AEAD scheme Π, an adversary $\mathcal{A}$ outputs (K, N, AD, M) and (K', N', AD', M'); $\mathcal{A}$ wins if $K \neq K'$, and $\Pi.\mathsf{Enc}_K(N, AD, M) = \Pi.\mathsf{Enc}_K(N', AD', M')$. We write $\mathbf{Adv}^{\mathsf{CMT\text{-}1}}_{\Pi}(\mathcal{A})$ to denote the probability that $\mathcal{A}$ wins.

Suffix-Committing security of a function. For a keyed function $F : \mathcal{K} \times \mathcal{X} \to \{0,1\}^m$ with an output length $m \geq t$, we define the t-suffix-committing security games CMT-1$\lfloor t \rfloor$ and FOR-KC$\lfloor t \rfloor$ as follows: an adversary $\mathcal{A}$ playing against F outputs (K, X) and (K', X') and

- wins the CMT-1$\lfloor t \rfloor$ game if $K \neq K'$ and $\mathsf{lsb}_t(F_K(X)) = \mathsf{lsb}_t(F_{K'}(X'))$;
- wins the FOR-KC$\lfloor t \rfloor$ game if $K \neq K'$, $\mathsf{lsb}_t(F_K(X)) = \mathsf{lsb}_t(F_{K'}(X'))$, and $X = X'$.

2.3 Description of the CAU Mode

The AEAD mode CAU [BT16] is a generalisation of GCM. Let $E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ be a blockcipher with a κ-bit key and an n-bit output, and let $\mathcal{H} : \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^n$ be a c-AXU hash function with an n-bit hash key and an n-bit output. CAU uses a single κ-bit key $\tilde{K}$, and generates the n-bit hash key as $K_h \leftarrow E_{\tilde{K}}(\langle 0 \rangle_n)$. It accepts a ν-bit nonce $\tilde{N}$ for some $\nu < n$, and sets the n-bit IV as $IV \leftarrow \langle \tilde{N} \rangle_\nu \| \langle 1 \rangle_{n-\nu}$.

Encryption of an L_M-byte message M is done in counter mode, with the i-th block of the keystream being generated as $E_{\tilde{K}}(\langle IV + i \rangle_n)$ (where $+$ denotes integer addition), and the final keystream block being truncated if needed to bring the total number of keystream bytes to L_M. The authentication tag is generated from the L_{AD}-byte associated data AD and L_M-byte ciphertext as $T \leftarrow \mathcal{H}(K_h, AD \| C \| \langle L_{AD} \rangle_{64} \| \langle L_M \rangle_{64}) \oplus E_{\tilde{K}}(IV)$, and truncated down to the desired length.

$\mathsf{CAU.Enc}_{\tilde{K}}(\tilde{N}, AD, M)$

1 : $IV \leftarrow \langle \tilde{N} \rangle_\nu \| \langle 1 \rangle_{n-\nu}$
2 : $L_M \leftarrow |M|_{\mathrm{B}}$
3 : $\mu \leftarrow \lceil L_M / 16 \rceil$
4 : **for** $j = 1$ to μ **do**
5 : $\quad Z_i \leftarrow E_{\tilde{K}}(\langle IV + i \rangle_n)$
6 : $Z \leftarrow \mathsf{msb}_{8L_M}(Z_1 \| \ldots \| Z_\mu)$
7 : $C \leftarrow M \oplus Z$
8 : $L_{AD} \leftarrow |AD|_{\mathrm{B}}$
9 : $\mathsf{data} \leftarrow AD \| C \| \langle L_{AD} \rangle_{64} \| \langle L_M \rangle_{64}$
10 : $K_h \leftarrow E_{\tilde{K}}(\langle 0 \rangle_n)$
11 : $T \leftarrow \mathcal{H}(K_h, \mathsf{data}) \oplus E_{\tilde{K}}(IV)$
12 : **return** (C, T)

$\mathsf{CAU.Dec}_{\tilde{K}}(\tilde{N}, AD, C, T)$

1 : $IV \leftarrow \langle \tilde{N} \rangle_\nu \| \langle 1 \rangle_{n-\nu}$
2 : $L_{AD} \leftarrow |AD|_{\mathrm{B}}$
3 : $L_M \leftarrow |C|_{\mathrm{B}}$
4 : $\mathsf{data} \leftarrow AD \| C \| \langle L_{AD} \rangle_{64} \| \langle L_M \rangle_{64}$
5 : $K_h \leftarrow E_{\tilde{K}}(\langle 0 \rangle_n)$
6 : $T' \leftarrow \mathcal{H}(K_h, \mathsf{data}) \oplus E_{\tilde{K}}(IV)$
7 : **if** $T' \neq T$
8 : $\quad$ **return** $\perp$
9 : $\mu \leftarrow \lceil L_M / 16 \rceil$
10 : **for** $j = 1$ to μ **do**
11 : $\quad Z_i \leftarrow E_{\tilde{K}}(\langle IV + i \rangle_n)$
12 : $Z \leftarrow \mathsf{msb}_{8L_M}(Z_1 \| \ldots \| Z_\mu)$
13 : $M \leftarrow C \oplus Z$
14 : **return** M

Figure 4: Algorithm of CAU.

Fig. 4 gives a formal description of the encryption and decryption algorithms of CAU. AES-GCM-256 is an instantiation of CAU, with $E = \mathsf{AES}$, $\mathcal{H} = \mathsf{GHASH}$, $n = 128$, $\nu = 96$, and $\kappa = 256$.

Multi-user security of CAU. Hoang et al. [HTT18] showed the following multi-user security bound for CAU in the ideal cipher model.

Theorem 1 ([HTT18], Theorem 3.1)**.** *For an adversary $\mathcal{A}$ playing an mu-AEAD game against CAU, making $p \leq 2^{n-2}$ ideal cipher queries to E and q total construction queries with the total number of*

queried blocks not exceeding σ and the number of queried blocks for each user not exceeding B, and subject to the restriction that each nonce is repeated by at most d users, the advantage of $\mathcal{A}$ satisfies the bound

$$\mathbf{Adv}^{\textit{mu-AEAD}}_{\textit{CAU}}(\mathcal{A}) \leq \frac{d(p+q)+n(p+q+\sigma)}{2^{\kappa}} + \frac{\sigma(2B+cn+3)}{2^{n}} + \frac{2q+1}{2^{2n}} + \frac{\sigma(\sigma+cnd)+2pq}{2^{\kappa+n}}. \tag{1}$$

2.4 H-Coefficient Technique

The H-coefficient technique is a proof technique devised by Patarin [Pat09] and popularised by the exposition due to Chen and Steinberger [CS14]. We outline the main idea below; for a more detailed exposition, the reader is referred to [CS14, JN22].

A distinguisher $\mathcal{A}$ interacts with one of two oracles, possibly providing an interface to multiple functionalities: $\mathcal{O}_1$ in the real world and $\mathcal{O}_0$ in the ideal world. A record of its interaction with oracle $\mathcal{O}$ is collected in a transcript $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}})$. The transcript may optionally include additional information that is revealed to the adversary at the end of the interactive phase. We assume, without loss of generality, that the oracles sample all their random coins, typically a key or an ideal primitive, before the experiment starts. A transcript tr is called *attainable* if $\mathcal{A}$ can observe tr with non-zero probability in the ideal world, i.e., if $\Pr[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) = \mathsf{tr}] > 0$. The Fundamental Theorem of the H-coefficient technique states the following:

Theorem 2 (H-Coefficient Technique [Pat09])**.** *Suppose we can partition the set $\mathcal{T}$ of all attainable transcripts as $\mathcal{T} = \mathcal{T}_{bad} \sqcup \mathcal{T}_{good}$, and find $\epsilon_1, \epsilon_2 \geq 0$ such that*

$$\Pr\left[\mathit{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{bad}\right] \leq \epsilon_1, \tag{2}$$

and for any $\mathsf{tr} \in \mathcal{T}_{good}$,

$$\frac{\Pr\left[\mathit{Trs}\,(\mathcal{A}^{\mathcal{O}_1}) = \mathit{tr}\right]}{\Pr\left[\mathit{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) = \mathit{tr}\right]} \geq 1-\epsilon_2. \tag{3}$$

Then the distinguishing advantage of $\mathcal{A}$ obeys the bound

$$\left|\Pr[\mathcal{A}^{\mathcal{O}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{O}_0} \Rightarrow 1]\right| \leq \epsilon_1 + \epsilon_2. \tag{4}$$

2.5 Mirror Theory

The mirror theory of Patarin [Pat06, Pat10b, Pat03] gives a lower bound on the number of solutions to a system of bivariate equations. We give a formal description of mirror theory here. The motivation of mirror theory and a more detailed overview can be found in Appendix E.

Let $\mathbf{Y} := (Y_1, \ldots, Y_r)^{\intercal}$ be a vector of variables, $\mathbf{Z} := (Z_1, \ldots, Z_m)^{\intercal}$ be a vector of values in $\mathbb{GF}(2^n)$ and $\mathbf{A}$ be a coefficient matrix whose (i,j)-th entry is A_{ij}. Consider the system of equations

$$\mathbf{A}_{m\times r} \cdot \mathbf{Y} = \mathbf{Z}. \tag{5}$$

We write $\Phi := \Phi[\mathbf{A}, \mathbf{Z}]$ to denote the system of equations described by Eqn. (5), parametrised by $\mathbf{A}$ and $\mathbf{Z}$. We call $\Phi[\mathbf{A}, \mathbf{Z}]$ a *bivariate mirror system* if $\mathbf{A}$ has the special form where each row of $\mathbf{A}$ has *exactly two 1's and* $(r-2)$ *0's*. A solution to Φ which satisfies the constraint that $Y_i \neq Y_j$ for all $i, j \in [1..r]$ with $i \neq j$ is called a *mirror solution*; let $\Gamma(\Phi)$ denote the set of all miror solutions to the system Φ.

We associate a weighted undirected graph $\mathsf{Gr} := \mathsf{Gr}(\Phi)$ with a bivariate mirror system $\Phi = \Phi[\mathbf{A}_{m\times r}, \mathbf{Z}]$ as follows: Gr has r vertices labelled $Y_1, \ldots, Y_r$, and m edges with weights $Z_1, \ldots, Z_m$, subject to the

following rule: there is an edge between Y_i and Y_j with weight Z_t if and only if $A_{ti} = A_{tj} = 1$. We call $\Phi[\mathbf{A}, \mathbf{Z}]$ *redundancy-free* if $m \leq r - 1$ and $\mathbf{A}$ is of full row-rank; in which case it is easy to see that Φ is redundancy-free if and only if $\mathsf{Gr}(\Phi)$ is acyclic. Assuming that $\Phi[\mathbf{A}, \mathbf{Z}]$ is redundancy-free and following graph theoretic terminology, we call Gr a *forest* and each of its connected components a *tree*. Then a basic result from graph theory tells us that Gr has exactly $c := r - m$ trees.

As observed in [CDN+23], we can assume without loss of generality that each tree in Gr is in fact a *star*, meaning that all of its nodes except one are *leaf nodes*. We call the solitary non-leaf node the *center* of the star. The ξ variables involved in a particular star component can be reordered to begin with the center, so that the corresponding coefficient matrix becomes $(\mathbf{1}_{\xi-1}, \mathbf{I}_{\xi-1})$, where $\mathbf{1}$ denotes a vector of 1's, and $\mathbf{I}$ denotes the identity matrix. Then Φ can be decomposed into c systems $\Phi^1, \ldots, \Phi^c$, where for each $p \in [1..c]$, $\Phi^p := \Phi^p[\mathbf{A}^p := (\mathbf{1}_{\xi_p-1}, \mathbf{I}_{\xi_p-1}), \mathbf{Z}^p := (Z^p_1, \ldots, Z^p_{\xi_p})^\intercal]$, with ξ_p denoting the size of the p-th component.

With the above setup in mind, we call Φ a *consistent mirror system* if for each $p \in [1..c]$, the p-th component satisfies the following conditions:

- $Z^p_i \neq 0$ for each $i \in [1..\xi_p]$;
- $Z^p_i \neq Z^p_j$ for each $i, j \in [1..\xi_p]$ with $i \neq j$.

With these definitions laid out, we are now ready to restate the main mirror theory result [CDN+23, Theorem 1] which we will employ later in this paper.

Theorem 3 (Mirror Theorem [CDN+23])**.** *Let Φ be a consistent mirror system of m equations in r variables, and let $\xi_{\max}$ denote the size of the largest component in $\mathsf{Gr}(\Phi)$. Then, if either $m \leq 2^{n/2}$, or $\xi_{\max}$ satisfies the conditions $n\xi^2_{\max} + \xi_{\max} \leq 2^{n/2}$ and $12m\xi_{\max} \leq 2^n$, we have the bound*

$$|\Gamma(\Phi)| \geq \frac{(2^n)_r}{2^{nm}}. \tag{6}$$

Note that Φ being a consistent mirror system is a necessary and sufficient condition to ensure $|\Gamma(\Phi)| > 0$. We note that we define consistency only for redundancy-free mirror systems; this does not lead to loss of generality since for every mirror system Φ we can find a redundancy-free mirror system Φ' such that Φ and Φ' have an identical set of solutions. Moreover, it may be interesting to observe that the characterisation of a system of equations $\Phi[\mathbf{A}, \mathbf{Z}]$ as a bivariate mirror system or as redundancy-free depends solely on the coefficient matrix $\mathbf{A}$, while its consistency depends solely on the difference vector $\mathbf{Z}$.

3 ♯Pencil: A BBB PRF

We begin this section by describing the design rationale of our main construction ♯Pencil, followed by its specifications and security results. Our goal was to come up with a construction that: (1) is based on a blockcipher with $2n$-bit key and n-bit output (so we can instantiate it with AES-256); (2) accepts a nonce that is almost $2n$ bits in length (less some bits reserved for domain separation); (3) generates a $2n$-bit key and an $(n - 32)$-bit nonce for AES-GCM as well as a $2n$-bit commitment tag; (4) achieves a PRF security close to n bits, even when the input nonce is not random but adversarially chosen; and (5) provides a key-commitment security that is BBB in n.

Design Rationale. As an intermediate goal, we design a weak PRF that takes a $2n$-bit input and produces a variable-length output. An XORP-like structure [Iwa06] is a suitable choice, as long as the two (random) input blocks are appropriately combined to feed the XORP. We generate the j-th input block as $X_j \leftarrow (R_1 \oplus 2^j \cdot R_2) \| \langle j \rangle_\mu$, where (R_1, R_2) is the random input. When fed into the XORP layer, this

$\sharp\mathsf{Pencil}_K(\mathbf{N})$	$\mathsf{Sharp}_{K_1}(\mathbf{N})$	$\mathsf{Pencil}_{K_2}(\mathbf{W},\mathbf{R})$
for $i = 0$ to 3 **do**	$(N_1, N_2) \leftarrow \mathbf{N}$	$(W_1, W_2) \leftarrow \mathbf{W}$
$\quad L_i \leftarrow E_K(\langle i \rangle_{128})$	$J \leftarrow \mathsf{msb}_{120}(E_{K_1}(N_2 \| \mathtt{0x00}))$	$(R_1, R_2) \leftarrow \mathbf{R}$
$K_1 \leftarrow L_0 \| L_1$	$W_1 \leftarrow N_1 \oplus J$	**for** $j = 0$ to 5 **do**
$K_2 \leftarrow L_2 \| L_3$	$W_2 \leftarrow N_1 \oplus 2 \cdot J$	$\quad X_j \leftarrow (R_1 \oplus 2^j \cdot R_2) \| \langle j \rangle_8$
$(\mathbf{W}, \mathbf{R}) \leftarrow \mathsf{Sharp}_{K_1}(\mathbf{N})$	$R_1 \leftarrow \mathsf{msb}_{120}(E_{K_1}(W_1 \| \mathtt{0x01}))$	$X_4 \leftarrow X_4 \oplus (W_1 \| \mathtt{0x00})$
$\mathbf{Z} \leftarrow \mathsf{Pencil}_{K_2}(\mathbf{W}, \mathbf{R})$	$R_2 \leftarrow \mathsf{msb}_{120}(E_{K_1}(W_2 \| \mathtt{0x02}))$	$X_5 \leftarrow X_5 \oplus (W_2 \| \mathtt{0x00})$
return $\mathbf{Z}$	$\mathbf{W} \leftarrow (W_1, W_2)$	**for** $j = 0$ to 5 **do**
	$\mathbf{R} \leftarrow (R_1, R_2)$	$\quad Y_j \leftarrow E_{K_2}(X_j)$
	return $(\mathbf{W}, \mathbf{R})$	$\quad$ **if** $j \geq 1$ **then**
		$\quad\quad Z_j \leftarrow Y_0 \oplus Y_j$
		$\mathbf{Z} \leftarrow (Z_1, Z_2, Z_3, Z_4, Z_5)$
		return $\mathbf{Z}$

Figure 5: Algorithms for $\sharp$Pencil, Sharp, and Pencil for $n = 128$.

construction achieves a weak PRF security close to n bits. For more details on how we arrived at the present design, see Appendix B.1.

Once we have the weak PRF, the next step is to convert it into a strong PRF, using a suitable preprocessing function. While an obvious choice is to apply a $2n$-bit-to-$2n$-bit PRF to the nonce to randomize it, this PRF would require at least six blockcipher calls to achieve n-bit security, going by the state of the art, which when combined with the augmented XORP layer would be more expensive than what we are aiming for. We observe that our weak PRF proof does not require its inputs to be fully random—it suffices that they appear random enough so as to satisfy the conditions required by Mirror theory for the proof to go through. Accordingly, we design a three-call preprocessing layer Sharp, which is weaker than a PRF, but achieves sufficient randomization so that its composition with the weak PRF construction gives a strong PRF with approximately n-bit security.

For the final goal of commitment, we discovered that generating a commitment tag by simply extending the output of the augmented XORP layer would not suffice to achieve BBB commitment security. This is evidenced by a birthday-bound commitment attack on this construction that is described in Appendix B.2. To overcome this, we expose two intermediate values from the Sharp layer as part of its output and mix them into the augmented XORP layer when generating the last two blocks of the final PRF output in order to effectively obtain a commitment tag. This approach preserves the (strong) PRF security of the earlier construction, while also achieving BBB key-commitment security. We call this weak PRF construction extending the augmented XORP layer Pencil.

The $\sharp$Pencil Construction. We now present our main construction $\sharp$Pencil and establish its security as a PRF. $\sharp$Pencil is realised from a blockcipher E with an n-bit input and a $2n$-bit key and consists of two subcomponents: Sharp and Pencil. $\mathsf{Sharp} : \{0,1\}^{2n} \times \{0,1\}^{2n-16} \to \{0,1\}^{4n-32}$ uses a $2n$-bit key and accepts a $(2n-16)$-bit user-chosen input, and returns four blocks of $n-8$ bits each (Fig. 1, left; Fig. 5, middle). $\mathsf{Pencil} : \{0,1\}^{2n} \times \{0,1\}^{4n-32} \to \{0,1\}^{5n}$, uses a $2n$-bit key and accepts a $(4n-32)$-bit input, and outputs five n-bit blocks (Fig. 1, right; Fig. 5, right). Combining the two yields $\sharp\mathsf{Pencil} : \{0,1\}^{2n} \times \{0,1\}^{2n-16} \to \{0,1\}^{5n}$, a PRF taking a $2n$-bit key and a $(2n-16)$-bit input and returning a $5n$-bit output (Fig. 5, left). Looking ahead, we will split this $5n$-bit output into a $2n$-bit key and a nonce for GCM, and

use the last $2n$-bits as a commitment tag that we append to the ciphertext. We now show that $\sharp$Pencil is a BBB PRF with a nearly-optimal bound. The security of $\sharp$Pencil as a $2n$-suffix-committing PRF is established in Sec. 5.

3.1 PRF Security of $\sharp$Pencil

Theorem 4 (PRF security)**.** *Suppose $25n^3 + 5n \leq 2^{n/2}$. Let E be an ideal cipher with a $2n$-bit key and an n-bit output. For any q-query PRF-adversary $\mathcal{A}$ also making p ideal cipher queries to E, we have*

$$\mathbf{Adv}^{PRF}_{\sharp Pencil}(\mathcal{A}) \leq \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}} + \frac{3p}{2^{2n}} + \frac{6}{2^n}, \tag{7}$$

as long as $nq \leq 2^{n-11}$.

For $n = 128$, the condition $25n^3 + 5n \leq 2^{n/2}$ holds, and we can simplify the above bound to say that as long as $q \leq 2^{110}$ and $p \ll 2^{254}$, $\mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}}(\mathcal{A})$ is bounded by $q/2^{113}$.

Proof of Theorem 4. As an intermediate we consider the unkeyed hybrid construction $\sharp\mathsf{Pencil}[\tilde{\pi} := (\pi_1, \pi_2)]$, which replaces E_{K_i} with a uniform random permutation $\pi_i : \{0,1\}^n \to \{0,1\}^n$ for $i = 1, 2$. We first establish the following lemma (proof to follow).

Lemma 1. *For any q-query PRF-adversary $\mathcal{A}$ also making p ideal cipher queries to E, we have*

$$\mathbf{Adv}^{PRF}_{\sharp Pencil}(\mathcal{A}) \leq \mathbf{Adv}^{PRF}_{\sharp Pencil[\tilde{\pi}]}(\mathcal{A}) + \frac{3p}{2^{2n}} + \frac{6}{2^n}. \tag{8}$$

Next, in Sec. 3.2, we establish a bound on $\mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}[\tilde{\pi}]}(\mathcal{A})$ for any q-query adversary $\mathcal{A}$. Specifically, we prove the following lemma.

Lemma 2. *Suppose $25n^3 + 5n \leq 2^{n/2}$. Then for any q-query PRF-adversary $\mathcal{A}$,*

$$\mathbf{Adv}^{PRF}_{\sharp Pencil[\tilde{\pi}]}(\mathcal{A}) \leq \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}}, \tag{9}$$

as long as $nq \leq 2^{n-11}$.

Theorem 4 follows from Lemmas 1 and 2. □

Proof of Lemma 1. We show this bound through a series of game hops, as illustrated in Fig. 6, where $\mathcal{G}_1$ represents $\sharp$Pencil and $\mathcal{G}_4$ represents $\sharp\mathsf{Pencil}[\tilde{\pi}]$:

- $\mathcal{G}_1$ and $\mathcal{G}_2$ are identical unless one of the adversary's ideal cipher queries is keyed with K in $\mathcal{G}_1$, and the probability of this is $p/2^{2n}$, since K is sampled uniformly at random from $\{0,1\}^{2n}$, so

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1]\right| \leq \frac{p}{2^{2n}}; \tag{10}$$

- $\mathcal{G}_2$ and $\mathcal{G}_3$ are identical unless there is a collision in the set $\{L_0, L_1, L_2, L_3\}$ in $\mathcal{G}_3$, and the probability of such a collision is upper bounded by $6/2^n$ by union bound, so

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1]\right| \leq \frac{6}{2^n}; \tag{11}$$

Game $\mathcal{G}_1$	Game $\mathcal{G}_2$	Game $\mathcal{G}_3$	Game $\mathcal{G}_4$
$K \leftarrow_\$ \{0,1\}^{256}$	$\pi \leftarrow_\$ \mathsf{Perm}(\{0,1\}^{128})$ **for** $i = 0$ to 3 **do** $\quad L_i \leftarrow \pi(\langle i \rangle_{128})$ $K_1 \leftarrow L_0 \| L_1$ $K_2 \leftarrow L_2 \| L_3$	**for** $i = 0$ to 3 **do** $\quad L_i \leftarrow_\$ \{0,1\}^{128}$ $K_1 \leftarrow L_0 \| L_1$ $K_2 \leftarrow L_2 \| L_3$	$\pi_1 \leftarrow_\$ \mathsf{Perm}(\{0,1\}^{128})$ $\pi_2 \leftarrow_\$ \mathsf{Perm}(\{0,1\}^{128})$
FUNC($\mathbf{N}$)	FUNC($\mathbf{N}$)	FUNC($\mathbf{N}$)	FUNC($\mathbf{N}$)
return $\sharp\mathsf{Pencil}_K(\mathbf{N})$	$(\mathbf{W}, \mathbf{R}) \leftarrow \mathsf{Sharp}_{K_1}(\mathbf{N})$ **return** $\mathsf{Pencil}_{K_2}(\mathbf{W}, \mathbf{R})$	$(\mathbf{W}, \mathbf{R}) \leftarrow \mathsf{Sharp}_{K_1}(\mathbf{N})$ **return** $\mathsf{Pencil}_{K_2}(\mathbf{W}, \mathbf{R})$	$(\mathbf{W}, \mathbf{R}) \leftarrow \mathsf{Sharp}[\pi_1](\mathbf{N})$ **return** $\mathsf{Pencil}[\pi_2](\mathbf{W}, \mathbf{R})$

Figure 6: Games from the proof of Lemma 1.

- $\mathcal{G}_3$ and $\mathcal{G}_4$ are identical unless one of the adversary's ideal cipher queries is keyed with K_1 or K_2 in $\mathcal{G}_3$, and the probability of this is upper bounded by $2p/2^{2n}$, since K_1 and K_2 are sampled uniformly at random from $\{0,1\}^{2n}$, so

$$\left| \Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1] \right| \leq \frac{2p}{2^{2n}}. \tag{12}$$

We can write

$$\begin{aligned} \mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}}(\mathcal{A}) &\leq \mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}[\tilde{\pi}]}(\mathcal{A}) + \left| \Pr[\mathcal{A}^{\mathcal{G}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1] \right| \\ &\leq \mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}[\tilde{\pi}]}(\mathcal{A}) \qquad (13) \\ &\qquad + \sum_{i=1}^{3} \left| \Pr[\mathcal{A}^{\mathcal{G}_i} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_{i+1}} \Rightarrow 1] \right| \\ &\leq \mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}[\tilde{\pi}]}(\mathcal{A}) + \frac{3p}{2^{2n}} + \frac{6}{2^n}, \qquad (14) \end{aligned}$$

thus completing the proof of the lemma. □

3.2 Proof of Lemma 2

For proving Lemma 2, we need to bound the advantage of a q-query PRF-adversary $\mathcal{A}$ against $\sharp\mathsf{Pencil}[\tilde{\pi}]$. We first describe how all the intermediate and final outputs are computed in the two worlds. In the real world, for each input $\mathbf{N}^{(i)} := (N_1^{(i)}, N_2^{(i)})$, the values $J^{(i)}, \mathbf{W}^{(i)} := (W_1^{(i)}, W_2^{(i)}), \mathbf{R}^{(i)} := (R_1^{(i)}, R_2^{(i)})$ are first computed by running $\mathsf{Sharp}[\pi_1](N_1^{(i)}, N_2^{(i)})$ (see Fig. 7, left). (Note that $J^{(i)}$ is not an output of Sharp, but is computed and stored as an intermediate value while running Sharp.) Next $\mathbf{Z}^{(i)} := (Z_1^{(i)}, \ldots, Z_5^{(i)})$ is computed by running $\mathsf{Pencil}[\pi_2](W_1^{(i)}, W_2^{(i)}, R_1^{(i)}, R_2^{(i)})$ (see Fig. 7, right), and in the process $\mathbf{X}^{(i)} := (X_0^{(i)}, \ldots, X_5^{(i)})$ and $\mathbf{Y}^{(i)} := (Y_0^{(i)}, \ldots, Y_5^{(i)})$ are computed as intermediate values and stored. At the end of the query phase, $\mathbf{J} := (J^{(1)}, \ldots, J^{(q)})$, $\mathbf{R} := (\mathbf{R}^{(1)}, \ldots, \mathbf{R}^{(q)})$, and $\mathbf{Y} := (\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(q)})$ are also released to $\mathcal{A}$; $\mathbf{W} := (\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(q)})$ and $\mathbf{X} := (\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(q)})$ can be computed by $\mathcal{A}$.

In the ideal world, on receiving a query $\mathbf{N}^{(i)}$, the oracle $\mathcal{O}_0$ computes and returns $\mathbf{Z}^{(i)} \leftarrow \varphi(\mathbf{N}^{(i)})$; at the end of the query phase $\mathcal{O}_0$ computes $\mathbf{J}$, $\mathbf{W}$, and $\mathbf{R}$ exactly as in the real world, by running $\mathsf{Sharp}[\pi_1]$; it then samples $\mathbf{Y} := (\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(q)})$ as described below, and returns $\mathbf{J}$, $\mathbf{W}$, $\mathbf{R}$, and $\mathbf{Y}$ to $\mathcal{A}$.

$\sharp\mathsf{Pencil}[\tilde{\pi}](\mathbf{N})$	$\mathsf{Sharp}[\pi_1](\mathbf{N})$	$\mathsf{Pencil}[\pi_2](\mathbf{W},\mathbf{R})$
$(\pi_1,\pi_2) \leftarrow \tilde{\pi}$	$(N_1,N_2) \leftarrow \mathbf{N}$	$(W_1,W_2) \leftarrow \mathbf{W}$
$(\mathbf{W},\mathbf{R}) \leftarrow \mathsf{Sharp}[\pi_1](\mathbf{N})$	$J \leftarrow \mathsf{msb}_{120}(\pi_1(N_2 \Vert \mathtt{0x00}))$	$(R_1,R_2) \leftarrow \mathbf{R}$
$\mathbf{Z} \leftarrow \mathsf{Pencil}[\pi_2](\mathbf{W},\mathbf{R})$	$W_1 \leftarrow N_1 \oplus J$	**for** $j = 0$ to 5 **do**
return $\mathbf{Z}$	$W_2 \leftarrow N_1 \oplus 2 \cdot J$	$\quad X_j \leftarrow (R_1 \oplus 2^j \cdot R_2) \Vert \langle j \rangle_8$
	$R_1 \leftarrow \mathsf{msb}_{120}(\pi_1(W_1 \Vert \mathtt{0x01}))$	$X_4 \leftarrow X_4 \oplus (W_1 \Vert \mathtt{0x00})$
	$R_2 \leftarrow \mathsf{msb}_{120}(\pi_1(W_2 \Vert \mathtt{0x02}))$	$X_5 \leftarrow X_5 \oplus (W_2 \Vert \mathtt{0x00})$
	$\mathbf{W} \leftarrow (W_1,W_2)$	**for** $j = 0$ to 5 **do**
	$\mathbf{R} \leftarrow (R_1,R_2)$	$\quad Y_j \leftarrow \pi_2(X_j)$
	return $(\mathbf{W},\mathbf{R})$	$\quad$ **if** $j \geq 1$ **then**
		$\quad\quad Z_j \leftarrow Y_0 \oplus Y_j$
		$\mathbf{Z} \leftarrow (Z_1,Z_2,Z_3,Z_4,Z_5)$
		return $\mathbf{Z}$

Figure 7: Algorithms for $\sharp\mathsf{Pencil}[\tilde{\pi}]$, $\mathsf{Sharp}[\pi_1]$, and $\mathsf{Pencil}[\pi_2]$ for $n = 128$.

In order to sample $\mathbf{Y}$, $\mathcal{O}_0$ first sets $X_j^{(i)} \leftarrow (R_1^{(i)} \oplus 2^j \cdot R_2^{(i)}) \Vert \langle j \rangle_8$ for $0 \leq j \leq 3$ and $X_j^{(i)} \leftarrow (R_1^{(i)} \oplus 2^j \cdot R_2^{(i)}) \Vert \langle j \rangle_8 \oplus W_{j-3}^{(i)} \Vert \mathtt{0x00}$ for $j = 4,5$ for each $i \in [1..q]$. It then looks at the set of distinct values $\{X_j^{(i)} \mid i \in [1..q], j \in [0..5]\}$, which are relabelled as $X_1,\ldots,X_r$, where r denotes the number of distinct values taken by $X_j^{(i)}$. Let $h : [1..q] \times [0..5] \to [1..r]$ denote the relabelling function, such that for each $i \in [1..q], j \in [0..5]$, $X_{h(i,j)} = X_j^{(i)}$. Next, $\mathcal{O}_0$ constructs a mirror system of $5q$ equations Φ on r variables $Y_1,\ldots,Y_r$ as follows: for each $i \in [1..q], j \in [1..5]$, $\mathcal{O}_0$ adds to Φ the equation $Y_{h(i,0)} \oplus Y_{h(i,j)} = Z_j^{(i)}$. Finally, if Φ is not a consistent mirror system, $\mathcal{O}_0$ samples $\mathbf{Y}^{(i)} \leftarrow_{\$} \{0,1\}^{6n}$ for each $i \in [1..q]$; otherwise, it samples $(Y_1,\ldots,Y_r) \leftarrow_{\$} \Gamma(\Phi)$, and assigns $Y_j^{(i)} = Y_{h(i,j)}$ for each $i \in [1..q], j \in [0..5]$.

Partitioning the Transcripts. Each transcript tr is represented as a tuple $(\mathbf{N},\mathbf{Z};\mathbf{J},\mathbf{W},\mathbf{R},\mathbf{X},\mathbf{Y})$. We will classify the set $\mathcal{T}$ of all attainable transcripts into $\mathcal{T}_{\mathsf{bad}}$ and $\mathcal{T}_{\mathsf{good}}$. Let Φ_{tr} be the system of equations associated with a transcript tr in the manner described above (where Φ is completely determined by $\mathbf{W}$, $\mathbf{R}$, and $\mathbf{Z}$), and, when Φ_{tr} is a mirror system, let $\xi_{\max}(\mathsf{tr})$ be the size of the largest component in $\mathsf{Gr}(\Phi_{\mathsf{tr}})$. We recall that for a positive integer L we say Φ_{tr} is *L-scattered* if $\xi_{\max}(\mathsf{tr}) \leq L$. We first define the following two sets:

$$\mathcal{T}_{\mathsf{bad},1} := \{\mathsf{tr} \mid \Phi_{\mathsf{tr}} \text{ is not a redundancy-free, } 5n\text{-scattered mirror system}\},$$
$$\mathcal{T}_{\mathsf{bad},2} := \{\mathsf{tr} \mid \Phi_{\mathsf{tr}} \text{ is not a consistent mirror system}\}.$$

Then we define $\mathcal{T}_{\mathsf{bad}} := \mathcal{T}_{\mathsf{bad},1} \cup \mathcal{T}_{\mathsf{bad},2}$, and $\mathcal{T}_{\mathsf{good}} := \mathcal{T} \setminus \mathcal{T}_{\mathsf{bad}}$. We state below two lemmas, which we prove in Sec. 3.3 and Appendix C.1 respectively.

Lemma 3. *For any q-query adversary $\mathcal{A}$ with $nq \leq 2^{n-11}$,*

$$\Pr\left[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right) \in \mathcal{T}_{\mathsf{bad}}\right] \leq \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}}. \tag{15}$$

Lemma 4. *For any q-query adversary $\mathcal{A}$ and for any $\mathsf{tr} \in \mathcal{T}_{\mathit{good}}$,*

$$\frac{\Pr\left[\mathit{Trs}\left(\mathcal{A}^{\mathcal{O}_1}\right) = \mathit{tr}\right]}{\Pr\left[\mathit{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right) = \mathit{tr}\right]} \geq 1. \tag{16}$$

under the assumptions that $25n^3 + 5n \leq 2^{n/2}$, and $nq \leq 2^{n-9}$.

Lemma 2 follows from Theorem 2 and Lemmas 3 and 4, by choosing $\epsilon_1 = q/2^{n-15} + nq/2^{n-4}$ and $\epsilon_2 = 0$. □

3.3 Proof of Lemma 3

In this subsection, we'll denote $\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right)$ as $\mathsf{tr} = (\mathbf{N}, \mathbf{Z}; \mathbf{J}, \mathbf{W}, \mathbf{R}, \mathbf{X}, \mathbf{Y})$; of these, $\mathbf{N}$ is under direct adversarial control, and the rest we treat as random variables sampled in the ideal world as described earlier. First we observe that whether or not $\mathsf{tr} \in \mathcal{T}_{\mathsf{bad},1}$ is determined entirely by $(\mathbf{R}, \mathbf{W})$. Since $\mathbf{N}$ and $\mathbf{J}$ together determine $\mathbf{W}$, we begin by bounding various kinds of degeneracy events on the sampling of $(\mathbf{R}, \mathbf{J})$.

Bounding Arms and Elbows. For distinct $i_1, i_2 \in [1..q]$, the pair (i_1, i_2) is called an *arm* if either $W_1^{(i_1)} = W_1^{(i_2)}$ or $W_2^{(i_1)} = W_2^{(i_2)}$. Let $\mathcal{S}_{\mathsf{ar}}$ denote the set of all arms. Let $\mathcal{E}_{\mathsf{ar}}$ denote the event that $|\mathcal{S}_{\mathsf{ar}}| > q$. We will show the following bound on $\Pr[\mathcal{E}_{\mathsf{ar}}]$ in Appendix C.2.

Lemma 5.

$$\Pr[\mathcal{E}_{\mathit{ar}}] \leq \frac{q}{2^{n-9}}. \tag{17}$$

For distinct $i_1, i_2, i_3 \in [1..q]$, the triple (i_1, i_2, i_3) is called an *elbow* if $W_1^{(i_1)} = W_1^{(i_2)}$ and $W_2^{(i_2)} = W_2^{(i_3)}$. Let $\mathcal{S}_{\mathsf{el}}$ denote the set of all elbows. Let $\mathcal{E}_{\mathsf{el}}$ denote the event that $|\mathcal{S}_{\mathsf{el}}| > q$. We will show the following bound on $\Pr[\mathcal{E}_{\mathsf{el}}]$ in Appendix C.3.

Lemma 6.

$$\Pr[\mathcal{E}_{\mathit{el}}] \leq \frac{q}{2^{n-8}} + \frac{q^2}{2^{2n-16}}. \tag{18}$$

Bounding Stale Nodes on Chains. For some $k \geq 2$ and distinct $i_1, \ldots, i_k \in [1..q]$, we call $\zeta := (i_1, \ldots, i_k)$ a *k-chain* if for some $j_1, \ldots, j_{k-1} \in [0..5]$,

$$(X_{j_1}^{(i_2)}, \ldots, X_{j_{k-1}}^{(i_k)}) = (X_{j_1}^{(i_1)}, \ldots, X_{j_{k-1}}^{(i_{k-1})}). \tag{19}$$

Each query index i_d for $d \in [1..k]$ in a k-chain ζ is called a *node* of ζ. A node i_d for $d \geq 2$ is called *fresh* if either $W_1^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \leq a < d\}$ or $W_2^{(i_d)} \notin \{W_2^{(i_a)} \mid 1 \leq a < d\}$ (or both), and *stale* otherwise.

Let $\mathcal{E}_{\mathsf{ch},k}$ denote the event that there exists a k-chain $\zeta = (i_1, \ldots, i_k)$ such that i_d is stale for some $d \in [2..k]$. Let $\mathcal{E}_{\mathsf{ch}}$ denote the event $\mathcal{E}_{\mathsf{ch},3} \vee \cdots \vee \mathcal{E}_{\mathsf{ch},n+1}$. We will show the following bound on $\Pr[\mathcal{E}_{\mathsf{ch}}]$ in Appendix C.4.

Lemma 7. *As long as $q \leq 2^{n-11}$,*

$$\Pr[\mathcal{E}_{\mathit{ch}}] \leq \frac{q}{2^{n-13}}. \tag{20}$$

Bounding Stale Nodes on Trees. For some $k \geq 2$ and distinct $i_1, \ldots, i_k \in [1..q]$, we call $\vartheta := (i_1, \ldots, i_k)$ a *k-tree* if for some $i_1^\star, \ldots, i_{k-1}^\star$ satisfying the constraint that for each $d \in [1..k-1]$, $i_d^\star \in \{i_1, \ldots, i_d\}$, and some $j_1, \ldots, j_{k-1} \in [0..5]$,

$$(X_{j_1}^{(i_2)}, \ldots, X_{j_{k-1}}^{(i_k)}) = (X_{j_1}^{(i_1^\star)}, \ldots, X_{j_{k-1}}^{(i_{k-1}^\star)}). \tag{21}$$

As for a chain, each query index i_d for $d \in [1..k]$ in a k-tree ϑ is called a node of ϑ. A node i_d for $d \geq 2$ is called *fresh* if either $W_1^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \leq a < d\}$ or $W_2^{(i_d)} \notin \{W_2^{(i_a)} \mid 1 \leq a < d\}$ (or both), and *stale* otherwise.

Let $\mathcal{E}_{\mathsf{t},k}$ denote the event that there exists a k-tree $\vartheta = (i_1, \ldots, i_k)$ such that i_d is stale for some $d \in [2..k]$. Let $\mathcal{E}_{\mathsf{t}}$ denote the event $\mathcal{E}_{\mathsf{t},3} \vee \cdots \vee \mathcal{E}_{\mathsf{t},n+1}$. We will show the following bound on $\Pr[\mathcal{E}_{\mathsf{t}}]$ in Appendix C.5.

Lemma 8. *As long as* $nq \leq 2^{n-11}$,

$$\Pr[\mathcal{E}_t] \leq \frac{q}{2^{n-13}}. \tag{22}$$

Bounding two-query cycles. Let $\mathcal{E}_2$ denote the event that for some $i_1, i_2 \in [1..q]$ with $i_1 \neq i_2$, and some $j_1, j_2 \in [0..5]$ with $j_1 < j_2$,

$$(X_{j_1}^{(i_2)}, X_{j_2}^{(i_2)}) = (X_{j_1}^{(i_1)}, X_{j_2}^{(i_1)}). \tag{23}$$

We will show the following bound on $\Pr[\mathcal{E}_2 \mid \neg\mathcal{E}_{\mathsf{ar}}]$ in Appendix C.6.

Lemma 9. *As long as* $q \leq 2^{n-11}$,

$$\Pr[\mathcal{E}_2 \mid \neg\mathcal{E}_{ar}] \leq \frac{q}{2^{n-13}}. \tag{24}$$

Bounding long paths. Let $\mathcal{E}_3$ denote the event that for some $i_1, \ldots, i_{n+1} \in [1..q]$ with $i_1 < \cdots < i_{n+1}$, and some $j_1, \ldots, j_n \in [0..5]$,

$$(X_{j_1}^{(i_2)}, \ldots, X_{j_n}^{(i_{n+1})}) = (X_{j_1}^{(i_1)}, \ldots, X_{j_n}^{(i_n)}). \tag{25}$$

We will show the following bound on $\Pr[\mathcal{E}_3 \mid \neg\mathcal{E}_{\mathsf{ch}}]$ in Appendix C.7.

Lemma 10. *As long as* $q \leq 2^{n-11}$,

$$\Pr[\mathcal{E}_3 \mid \neg\mathcal{E}_{ch}] \leq \frac{q}{2^n}. \tag{26}$$

Bounding multi-query cycles. For each $k \in [3..n]$, let $\mathcal{E}_{4,k}$ denote the event that for some $i_1, \ldots, i_k \in [1..q]$ with $i_1 < \cdots < i_k$, and some $j_1, \ldots, j_k \in [0..5]$,

$$(X_{j_1}^{(i_2)}, \ldots, X_{j_{k-1}}^{(i_k)}, X_{j_k}^{(i_1)}) = (X_{j_1}^{(i_1)}, \ldots, X_{j_{k-1}}^{(i_{k-1})}, X_{j_k}^{(i_k)}). \tag{27}$$

Let $\mathcal{E}_4$ denote the event $\mathcal{E}_{4,3} \vee \cdots \vee \mathcal{E}_{4,n}$. We will show the following bound on $\Pr[\mathcal{E}_4 \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}]$ in Appendix C.8.

Lemma 11. *As long as* $q \leq 2^{n-11}$,

$$\Pr[\mathcal{E}_4 \mid \neg\mathcal{E}_{ch}, \neg\mathcal{E}_{ar}] \leq \frac{q^2}{2^{2n-26}}. \tag{28}$$

Bounding large components. Let $\mathcal{E}_5$ denote the event that for some $i_1, \ldots, i_{n+1} \in [1..q]$ with $i_1 < \cdots < i_{n+1}$, some $i_1^\star, \ldots, i_n^\star$ satisfying the constraint that for each $k \in [1..n]$, $i_k^\star \in \{i_1, \ldots, i_k\}$, and some $j_1, \ldots, j_n \in [0..5]$,

$$(X_{j_1}^{(i_2)}, \ldots, X_{j_n}^{(i_{n+1})}) = (X_{j_1}^{(i_1^\star)}, \ldots, X_{j_n}^{(i_n^\star)}). \tag{29}$$

This event is a generalisation of $\mathcal{E}_3$: in $\mathcal{E}_3$ for each query i_{k+1} we require $\mathbf{X}^{(i_{k+1})}$ to collide somewhere with $\mathbf{X}^{(i_k)}$ for the immediately preceding query i_k in the sequence $(i_1, \ldots, i_{n+1})$, whereas in $\mathcal{E}_5$ we relax this requirement to colliding somewhere with $\mathbf{X}^{(i_k^\star)}$ for any query $i_k^\star$ appearing before i_{k+1} in the sequence. Thus, $\mathcal{E}_5$ captures a more general component instead of just a path. We will show the following bound on $\Pr[\mathcal{E}_5 \mid \neg\mathcal{E}_\mathsf{t}, \neg\mathcal{E}_2, \neg\mathcal{E}_3, \neg\mathcal{E}_4]$ in Appendix C.9.

Lemma 12. *As long as* $nq \le 2^{n-11}$,

$$\Pr[\mathcal{E}_5 \mid \neg\mathcal{E}_t, \neg\mathcal{E}_2, \neg\mathcal{E}_3, \neg\mathcal{E}_4] \le \frac{q}{2^n}. \tag{30}$$

Getting a bound for $\Pr[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad},1}]$. Let $\mathcal{E}_{\mathsf{bad},1}$ be the event $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad},1}$. We will show the following bound on $\Pr[\mathcal{E}_{\mathsf{bad},1}]$ in Appendix C.10.

Lemma 13. *As long as* $nq \le 2^{n-11}$,

$$\Pr[\mathcal{E}_{bad,1}] \le \frac{q}{2^{n-15}}. \tag{31}$$

Bounding zero blocks. For the rest of this proof, we condition on $\neg\mathcal{E}_{\mathsf{bad},1}$, so we will only consider tr such that Φ_tr is redundancy-free and $5n$-scattered. As discussed in Sec. 2.5, we can assume without loss of generality that each component of $\mathsf{Gr}(\Phi_\mathsf{tr}) = \Phi[\mathbf{A}, \mathbf{Z}]$ is a star graph. Let $\mathcal{E}_6$ be the event that $Z_j = 0$ for some $j \in [1..5q]$. Then

$$\Pr[\mathcal{E}_6 \mid \neg\mathcal{E}_{\mathsf{bad},1}] \le \frac{5q}{2^n}. \tag{32}$$

Bounding output collisions. Let $\mathcal{E}_7$ be the event that $Z_i^p = Z_j^p$ for some $p \in [1..c]$ and $i, j \in [1..\xi_p]$ with $i \ne j$. Then

$$\Pr[\mathcal{E}_7 \mid \neg\mathcal{E}_{\mathsf{bad},1}] \le \sum_{p=1}^{c} \binom{\xi_p}{2} \frac{1}{2^n} \le \frac{\xi_{\max}}{2^{n+1}} \sum_{p=1}^{c} \xi_p \le \frac{25nq}{2^{n+1}}. \tag{33}$$

Getting a bound for $\Pr[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad},2} \mid \neg\mathcal{E}_{\mathsf{bad},1}]$. Let $\mathcal{E}_{\mathsf{bad},2}$ be the event $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad},2}$. We observe that

$$\mathcal{E}_{\mathsf{bad},2} \wedge \neg\mathcal{E}_{\mathsf{bad},1} \subseteq (\mathcal{E}_6 \wedge \neg\mathcal{E}_{\mathsf{bad},1}) \vee (\mathcal{E}_7 \wedge \neg\mathcal{E}_{\mathsf{bad},1}).$$

Thus, applying union-bound to Eqns. (32) and (33), we get

$$\Pr[\mathcal{E}_{\mathsf{bad},2} \mid \neg\mathcal{E}_{\mathsf{bad},1}] \le \Pr[\mathcal{E}_5 \mid \neg\mathcal{E}_{\mathsf{bad},1}] + \Pr[\mathcal{E}_6 \mid \neg\mathcal{E}_{\mathsf{bad},1}] \tag{34}$$

$$\le \frac{5q}{2^n} + \frac{25nq}{2^{n+1}} \le \frac{nq}{2^{n-4}}. \tag{35}$$

Getting a bound for $\Pr[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad}}]$. From the definitions of $\mathcal{E}_{\mathsf{bad},1}$, $\mathcal{E}_{\mathsf{bad},2}$, and $\mathcal{T}_{\mathsf{bad}}$, and from Eqns. (31) and (35), we get

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}) \in \mathcal{T}_{\mathsf{bad}}\right] = \Pr[\mathcal{E}_{\mathsf{bad},1} \vee \mathcal{E}_{\mathsf{bad},2}]$$

$$= \Pr[\mathcal{E}_{\mathsf{bad},1}] + \Pr[\mathcal{E}_{\mathsf{bad},2} \mid \neg\mathcal{E}_{\mathsf{bad},1}] \le \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}}, \tag{36}$$

thus establishing Eqn. (15), and completing the proof of Lemma 3. □

$\sharp$Pencil-CAU.Enc$_K(\mathbf{N}, AD, M)$

1 : $(\tilde{K}, \tilde{N}, \mathsf{KC}) \leftarrow \mathtt{parse}(\sharp\mathsf{Pencil}_K(\mathbf{N}))$
2 : **return** $(\mathsf{CAU.Enc}_{\tilde{K}}(\tilde{N}, AD, M), \mathsf{KC})$

$\sharp$Pencil-CAU.Dec$_K(\mathbf{N}, AD, C, T, \mathsf{KC})$

1 : $(\tilde{K}, \tilde{N}, \mathsf{KC}') \leftarrow \mathtt{parse}(\sharp\mathsf{Pencil}_K(\mathbf{N}))$
2 : **if** $\mathsf{KC} \neq \mathsf{KC}'$
3 : **return** $\perp$
4 : **return** $\mathsf{CAU.Dec}_{\tilde{K}}(\tilde{N}, AD, C, T)$

$\mathtt{parse}(\mathbf{Z})$

1 : $\tilde{K} \leftarrow \mathsf{msb}_\kappa(\mathbf{Z})$
2 : $\tilde{N} \leftarrow \mathsf{msb}_\nu(\mathsf{lsb}_{5n-\kappa}(\mathbf{Z}))$
3 : $\mathsf{KC} \leftarrow \mathsf{lsb}_{2n}(\mathbf{Z})$
4 : **return** $(\tilde{K}, \tilde{N}, \mathsf{KC})$

Figure 8: The $\sharp$Pencil-CAU composition, where $\mathbf{Z}$ is the $5n$-bit output of $\sharp$Pencil.

4 $\sharp$Pencil-CAU: A Strengthened AEAD Mode

In this section, we consider the use of $\sharp$Pencil as a KDF in combination with CAU (a generalisation of GCM) to yield a composite AEAD scheme that supports longer nonces and provides better quantitative security. Specifically, we now establish the security of the resulting composition $\sharp$Pencil-CAU as an AEAD scheme. In contrast to DNDK-GCM, we use $\sharp$Pencil to derive both the key and the nonce of the underlying CAU instance. Furthermore, because $\sharp$Pencil is a full-fledged PRF and we feed it the full nonce, this composition yields a standard nonce-based AEAD scheme, thus imposing no condition on the nonce to be sampled uniformly at random like DNDK-GCM, DNDKv2-GCM, XAES, and KC-XAES. When the CAU in $\sharp$Pencil-CAU is instantiated with AES-GCM-256, we refer to the resulting scheme as $\sharp$Pencil-GCM.

$\sharp$Pencil-CAU works as follows: it uses a κ-bit root key K (where κ is the key-length of E and hence of CAU), and accepts a $(2n-16)$-bit nonce $\mathbf{N}$; for each fresh nonce $\mathbf{N}$, it makes one call to $\sharp$Pencil with key K and input $\mathbf{N}$, and receives a $5n$-bit output $\mathbf{Z}$. It parses this output as $(\tilde{K}, \tilde{N}, \_s, \mathsf{KC})$, where $\tilde{K}$ is a κ-bit encryption key, $\tilde{N}$ is an ν-bit nonce, and KC is a $2n$-bit key-commitment tag; $\_s$ contains the remaining $3n - \kappa - \nu$ bits which are discarded. $\tilde{K}$ and $\tilde{N}$ go on to be used in the ensuing call to the appropriate interface of CAU, with the rest of the inputs to $\sharp$Pencil-CAU passed on untouched. KC does not interact with CAU, but is appended to the output of CAU to give the final output of $\sharp$Pencil-CAU. Fig. 8 gives a concise formal definition of the encryption and decryption functions of $\sharp$Pencil-CAU. The security of $\sharp$Pencil-CAU is formally stated in the following theorem, which relates its single-user AEAD security to the multiuser AEAD security of CAU.

Theorem 5. *Suppose $25n^3 + 5n \leq 2^{n/2}$. For an AEAD adversary $\mathcal{A}$ playing against $\sharp$Pencil-CAU, making $p \leq 2^{n-2}$ ideal cipher queries to E and $q \leq 2^{\nu-1}$ total construction queries with the total number of queried blocks not exceeding σ and the number of queried blocks for each nonce not exceeding B, we can find an mu-AEAD adversary $\mathcal{A}'$ playing against CAU, making p ideal cipher queries to E and q total construction queries with the total number of queried blocks not exceeding σ and the number of queried blocks for each user not exceeding B, and subject to the restriction that each nonce is repeated by at most n users, such that the advantage of $\mathcal{A}$ satisfies the bound*

$$\mathbf{Adv}^{\mathsf{AEAD}}_{\sharp\mathsf{Pencil\text{-}CAU}}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\mathsf{CAU}}(\mathcal{A}') + \frac{3q}{2^\kappa} + \frac{3p}{2^{2n}} + \frac{q+6}{2^n} + \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}}, \tag{37}$$

as long as $nq \leq 2^{n-11}$.

The following Corollary to Theorem 5 immediately follows from the bounds in Theorems 1 and 4.

Corollary 1. *Suppose $25n^3 + 5n \le 2^{n/2}$. For an AEAD adversary $\mathcal{A}$ playing against $\sharp$Pencil-CAU, making $p \le 2^{n-2}$ ideal cipher queries to E and $q \le 2^{\nu-1}$ total construction queries with the total number of queried blocks not exceeding σ and the number of queried blocks for each nonce not exceeding B, we have*

$$\begin{aligned}\mathbf{Adv}^{\mathsf{AEAD}}_{\sharp\mathsf{Pencil\text{-}CAU}}(\mathcal{A}) \le{}& \frac{n(p+q)+n(p+q+\sigma)+3q}{2^{\kappa}} \\ &+ \frac{\sigma(2B+cn+3)+q+6}{2^n} + \frac{2q+1+3p}{2^{2n}} \\ &+ \frac{\sigma(\sigma+cn^2)+2pq}{2^{\kappa+n}} + \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}},\end{aligned} \tag{38}$$

as long as $nq \le 2^{n-11}$.

Game $\mathcal{G}_1$

$K \leftarrow_{\$} \{0,1\}^{256}$

ENC($\mathbf{N}, AD, M$)

return $\sharp$Pencil-CAU.Enc$_K(\tilde{N}, AD, M)$

DEC($\mathbf{N}, AD, C, T, \mathsf{KC}$)

return $\sharp$Pencil-CAU.Dec$_K(\mathbf{N}, AD, C, T, \mathsf{KC})$

Game $\mathcal{G}_2$

$\pi_1 \leftarrow_{\$} \mathsf{Perm}(\{0,1\}^{128})$
$\pi_2 \leftarrow_{\$} \mathsf{Perm}(\{0,1\}^{128})$
$\tilde{\pi} \leftarrow (\pi_1, \pi_2)$

ENC($\mathbf{N}, AD, M$)

$(\tilde{K}, \tilde{N}, \mathsf{KC}) \leftarrow \mathtt{parse}(\sharp\mathsf{Pencil}[\tilde{\pi}](\mathbf{N}))$
return $(\mathsf{CAU.Enc}_{\tilde{K}}(\tilde{N}, AD, M), \mathsf{KC})$

DEC($\mathbf{N}, AD, C, T, \mathsf{KC}$)

$(\tilde{K}, \tilde{N}, \mathsf{KC}') \leftarrow \mathtt{parse}(\sharp\mathsf{Pencil}[\tilde{\pi}](\mathbf{N}))$
if $\mathsf{KC} \neq \mathsf{KC}'$
  return $\perp$
return $\mathsf{CAU.Dec}_{\tilde{K}}(\tilde{N}, AD, C, T)$

Game $\mathcal{G}_3$

$\varphi \leftarrow_{\$} \mathsf{Func}(\{0,1\}^{2n-16}, \{0,1\}^{5n})$

ENC($\mathbf{N}, AD, M$)

$(\tilde{K}, \tilde{N}, \mathsf{KC}) \leftarrow \mathtt{parse}(\varphi(\mathbf{N}))$
return $(\mathsf{CAU.Enc}_{\tilde{K}}(\tilde{N}, AD, M), \mathsf{KC})$

DEC($\mathbf{N}, AD, C, T, \mathsf{KC}$)

$(\tilde{K}, \tilde{N}, \mathsf{KC}') \leftarrow \mathtt{parse}(\varphi(\mathbf{N}))$
if $\mathsf{KC} \neq \mathsf{KC}'$
  return $\perp$
return $\mathsf{CAU.Dec}_{\tilde{K}}(\mathbf{N}, AD, C, T)$

Game $\mathcal{G}_4$

ENC($\mathbf{N}, AD, M$)

$C \leftarrow_{\$} \{0,1\}^{|M|}$
$T \leftarrow_{\$} \{0,1\}^{n}$
$\mathsf{KC} \leftarrow_{\$} \{0,1\}^{2n}$
return (C, T, KC)

DEC($\mathbf{N}, AD, C, T, \mathsf{KC}$)

return $\perp$

Figure 9: Games from the proof of Theorem 5.

Proof of Theorem 5. We show this bound through a series of game hops, described in Fig. 9. In game $\mathcal{G}_1$, the adversary interacts with $\sharp$Pencil-CAU. In game $\mathcal{G}_2$, $\sharp$Pencil is replaced with $\sharp$Pencil$[\tilde{\pi} := (\pi_1, \pi_2)]$ in $\sharp$Pencil-CAU, for uniform random permutations π_1 and π_2 independent of E. We bound the distance

between these two games by arguing like in the proof of Lemma 1, except we now also consider collisions with the derived keys fed into CAU in addition to the keys in $\mathcal{A}$'s ideal cipher queries. Thus we have

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1]\right| \leq \frac{3q}{2^\kappa} + \frac{3p}{2^{2n}} + \frac{6}{2^n}. \tag{39}$$

In game $\mathcal{G}_3$, the $\sharp\mathsf{Pencil}$ layer is replaced by a $(2n-16)$-bit-to-$5n$-bit uniform random function φ. Consider a PRF adversary $\mathcal{B}$ against $\sharp\mathsf{Pencil}[\tilde{\pi}]$, which simulates $\mathcal{A}$, defined as follows: on receiving an encryption query $(\mathbf{N}, AD, M)$ from $\mathcal{A}$, $\mathcal{B}$ queries its own oracle with $\mathbf{N}$, and receives a response $\mathbf{Z}$; it then sets $(\tilde{K}, \tilde{N}, \mathsf{KC}) \leftarrow \mathtt{parse}(\mathbf{Z})$, and returns $(\mathsf{CAU.Enc}_{\tilde{K}}(\tilde{N}, AD, M), \mathsf{KC})$ to $\mathcal{A}$; on receiving an decryption query $(\mathbf{N}, AD, C, T, \mathsf{KC})$ from $\mathcal{A}$, $\mathcal{B}$ again queries its own oracle with $\mathbf{N}$, and receives a response $\mathbf{Z}$; it then sets $(\tilde{K}, \tilde{N}, \mathsf{KC}') \leftarrow \mathtt{parse}(\mathbf{Z})$, checks if $\mathsf{KC} = \mathsf{KC}'$ and (if yes) returns $\mathsf{CAU.Dec}_{\tilde{K}}(\tilde{N}, AD, C, T)$ to $\mathcal{A}$. At the end of the game $\mathcal{B}$ copies the output bit of $\mathcal{A}$. Then

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1]\right| \leq \mathbf{Adv}^{\mathsf{PRF}}_{\sharp\mathsf{Pencil}[\tilde{\pi}]}(\mathcal{B}). \tag{40}$$

Substituting the bound from Lemma 2 in Eq. (40) gives

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1]\right| \leq \frac{q}{2^{n-15}} + \frac{nq}{2^{n-4}}. \tag{41}$$

In game $\mathcal{G}_3$, the CAU layer is dropped; the encryption queries are answered uniform random strings of the required length ($\ell + 3n$ for an ℓ-bit plaintext), and decryption queries are always answered with $\perp$. Consider an $\mathsf{mu\text{-}AEAD}$ adversary $\mathcal{A}''$ against CAU. $\mathcal{A}''$ simulates $\mathcal{A}$, and associates a unique user number $i(\mathbf{N})$ with each unique nonce $\mathbf{N}$ queried by $\mathcal{A}$ (e.g., by storing the nonces in a table and assigning the corresponding row number to i). Further, for each user number i, $\mathcal{A}''$ samples a uniform random ν-bit nonce $\tilde{N}_i \leftarrow\!\!\$\ \{0,1\}^\nu$. On receiving an encryption query $(\mathbf{N}, AD, M)$ from $\mathcal{A}$, $\mathcal{A}''$ queries its own encryption oracle with $(i(\mathbf{N}), \tilde{N}_{i(\mathbf{N})}, AD, M)$, and passes on the response to $\mathcal{A}$; on receiving a decryption query $(\mathbf{N}, AD, C, T)$ from $\mathcal{A}$, $\mathcal{A}$ queries its own decryption oracle with $(i(\mathbf{N}), \tilde{N}_{i(\mathbf{N})}, AD, C, T)$, and passes on the response to $\mathcal{A}$. It is clear that as long as $\mathcal{A}$'s encryption queries are nonce-respecting, $\mathcal{A}''$'s encryption queries are also nonce-respecting. At the end of the game, $\mathcal{A}''$ copies the final output bit of $\mathcal{A}$. Then we have

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1]\right| \leq \mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\mathsf{CAU}}(\mathcal{A}''). \tag{42}$$

Next consider an adversary $\mathcal{A}'$ which behaves identically to the adversary $\mathcal{A}''$, except after sampling each $\tilde{N}_i$, it counts the number $\eta(\tilde{N}_i)$ of previous users i' with $\tilde{N}_{i'} = \tilde{N}_i$; if $\eta(\tilde{N}_i) \leq n$, it sets $\tilde{N}_i$ to the lexicographically smallest $\tilde{N} \in \{0,1\}^r$ such that $\eta(\tilde{N}) \leq n$, and continues as $\mathcal{A}''$ does. Let $\mathcal{E}$ be the event that there is an $(n+1)$-multicollision on the ν-bit nonces. As long as $q \leq 2^{\nu-1}$, we have

$$\Pr[\mathcal{E}] \leq \binom{q}{n+1}\left(\frac{1}{2^r}\right)^n \leq q\left(\frac{q}{2^r}\right)^n \leq \frac{q}{2^n}. \tag{43}$$

We observe that $\mathcal{A}'$ and $\mathcal{A}''$ are identical unless $\mathcal{E}$ happens. Thus, from Eqn. (43), we have

$$\mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\mathsf{CAU}}(\mathcal{A}'') \leq \mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\mathsf{CAU}}(\mathcal{A}') + \frac{q}{2^n}. \tag{44}$$

Plugging Eqn. (44) in Eqn. (42) yields

$$\left|\Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1]\right| \leq \mathbf{Adv}^{\mathsf{mu\text{-}AEAD}}_{\mathsf{CAU}}(\mathcal{A}') + \frac{q}{2^n}. \tag{45}$$

The game $\mathcal{G}_4$ is identical to the ideal world of the AEAD game. Thus, we have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{AEAD}}_{\sharp\mathsf{Pencil\text{-}CAU}}(\mathcal{A}) &= \Big|\Pr[\mathcal{A}^{\mathcal{G}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1]\Big| \\ &\le \Big|\Pr[\mathcal{A}^{\mathcal{G}_1} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1]\Big| \\ &\quad + \Big|\Pr[\mathcal{A}^{\mathcal{G}_2} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1]\Big| \\ &\quad + \Big|\Pr[\mathcal{A}^{\mathcal{G}_3} \Rightarrow 1] - \Pr[\mathcal{A}^{\mathcal{G}_4} \Rightarrow 1]\Big|. \end{aligned} \tag{46}$$

Plugging in Eqns. (39), (41), and (45) in Eqn. (46) completes the proof. □

5 Committing Security of ♯Pencil-CAU

In this section we establish the committing security of ♯Pencil-CAU. We proceed in two steps. First, we observe that both committing-security notions of ♯Pencil-CAU reduce to the $2n$-suffix-committing security of its ♯Pencil key derivation layer: this reduction is immediate from the construction, since the key commitment tag is simply the $2n$-bit suffix of the ♯Pencil output. The main technical contribution of the section is the second step: bounding the suffix-committing security of ♯Pencil as summarised in Theorems 6 and 7.

Concretely, a CMT-1 adversary against ♯Pencil-CAU is successful if it outputs two tuples $(K, \mathbf{N}, AD, M)$ and $(K', \mathbf{N}', AD', M')$ with $K \neq K'$ that result in the same ciphertext. Since the ciphertext contains the key-commitment tag, colliding ciphertexts implies that $\mathsf{lsb}_{2n}(\sharp\mathsf{Pencil}_K(\mathbf{N})) = \mathsf{lsb}_{2n}(\sharp\mathsf{Pencil}_{K'}(\mathbf{N}'))$, which constitutes a winning pair for the CMT-1$\lfloor 2n \rfloor$ game against ♯Pencil. A similar argument holds for the FOR-KC case, with the additional restriction that the nonces be equal ($\mathbf{N} = \mathbf{N}'$) as required by FOR-KC$\lfloor 2n \rfloor$. This reduction is thus advantage-preserving and requires the same number of queries, i.e., ♯Pencil-CAU is CMT-1- and FOR-KC-committing with exactly the same advantage bounds established for ♯Pencil in Theorems 6 and 7 below.

Theorem 6 (CMT-1$\lfloor 2n \rfloor$ security of ♯Pencil)**.** *For any CMT-1$\lfloor 2n \rfloor$ adversary $\mathcal{A}$ making at most q blockcipher queries (in either direction) in the ideal cipher model,*

$$\mathbf{Adv}^{\textit{CMT-1}\lfloor 2n \rfloor}_{\sharp\textit{Pencil}}(\mathcal{A}) \le \frac{q^8}{2^{6(n-9)}} + \frac{q^4}{2^{3(n-9)}}$$

for $q \ge 2^{2(n-8)/3}$.

For $n = 128$, $\mathbf{Adv}^{\mathsf{CMT\text{-}1}\lfloor 256 \rfloor}_{\sharp\mathsf{Pencil}}(\mathcal{A}) \le (q/2^{87})^4$ for $q \ge 2^{80}$. Since the advantage is non-decreasing in q, for all $q < 2^{80}$, the advantage is at most 2^{-39}.

Theorem 7 (FOR-KC$\lfloor 2n \rfloor$ security of ♯Pencil)**.** *For any FOR-KC$\lfloor 2n \rfloor$ adversary $\mathcal{A}$ making at most q blockcipher queries (in either direction) in the ideal cipher model,*

$$\mathbf{Adv}^{\textit{FOR-KC}\lfloor 2n \rfloor}_{\sharp\textit{Pencil}}(\mathcal{A}) \le \frac{q^4}{2^{4(n-7)}}$$

for $q \ge 2^{2(n-8)/3}$.

The proof of Theorem 6 occupies the rest of this section. Theorem 7 is obtained from the same analysis specialised to $(N_1, N_2) = (N_1', N_2')$, and is proved in Appendix D.1.

Proof of Theorem 6. Only six query–response pairs feed the $2n$-bit suffix (Z_4, Z_5): the pairs $(N_2\|\texttt{0x00}, J\|*)$, $(W_1\|\texttt{0x01}, R_1\|*)$, $(W_2\|\texttt{0x02}, R_2\|*)$ to E_{K_1}, and (X_0, Y_0), (X_4, Y_4), (X_5, Y_5) to E_{K_2}. We collect them in tables $T_1, \ldots, T_6$, in this order. Since committing security places the blockcipher key under adversarial control, each entry of T_i is a triple (K, u, v), and by domain separation (the fixed last byte of each E_{K_i}-query) every query lands in at most one table. A forward query can be issued in the required shape directly, whereas an inverse query—e.g. $E_{K_1}^{-1}(R_1\|*)$—cannot control its last byte and must use rejection sampling, so the inverse tables are typically small; to be safe we work in a stronger model where every table has size q. The adversary's goal is two tuples (K, N_1, N_2) and (K', N_1', N_2') with $K \neq K'$ such that

$$\mathsf{lsb}_{2n}(\sharp\mathsf{Pencil}_K(N_1, N_2)) = (Z_4, Z_5) = \mathsf{lsb}_{2n}(\sharp\mathsf{Pencil}_{K'}(N_1', N_2')).$$

Fix a root key K with derived keys (K_1, K_2) (Fig. 5). Only entries under these keys matter for K, so we set

$$T_i^K := \begin{cases} \{(u, v) : (K_1, u, v) \in T_i\}, & i \in \{1, 2, 3\}, \\ \{(u, v) : (K_2, u, v) \in T_i\}, & i \in \{4, 5, 6\}, \end{cases}$$

and $T^K := \bigcup_i T_i^K$. We write $q_{K,i} := |T_i^K|$, $q_K := |T^K|$, and let $T_i^{K,+}$, $T_i^{K,-}$ be the forward and backward entries of T_i^K.

A tuple $\bar{x} = (x_1, \ldots, x_6)$ with $x_i = (u_i, v_i) \in T_i^K$ is *computable* if its six queries arise from a single valid (K, N_1, N_2, Z_4, Z_5) evaluation. As some variables are n-bit and others $(n-8)$-bit, the queries relate to the construction variables by

$$\begin{aligned} &u_1 = N_2\|\texttt{0x00}, \quad \mathsf{msb}_{n-8}(v_1) = J; \qquad u_4 = X_0, \quad v_4 = Y_0; \\ &u_2 = W_1\|\texttt{0x01}, \quad \mathsf{msb}_{n-8}(v_2) = R_1; \qquad u_5 = X_4, \quad v_5 = Y_4; \\ &u_3 = W_2\|\texttt{0x02}, \quad \mathsf{msb}_{n-8}(v_3) = R_2; \qquad u_6 = X_5, \quad v_6 = Y_5; \end{aligned}$$

and we write $\tilde{X}_i := \mathsf{msb}_{n-8}(X_i)$ for $i \in \{0, 4, 5\}$.

Fix a key K and an output pair (Z_4, Z_5), and let $\mathcal{E}_{(K,Z_4,Z_5)}$ be the event that some computable $\bar{x} \in T^K$ evaluates to (Z_4, Z_5); such an $\bar{x}$ is a *witness*. A computable $\bar{x}$ evaluates to (Z_4, Z_5) exactly when

$$R_1 \oplus R_2 = \tilde{X}_0, \tag{47}$$
$$R_1 \oplus 2^4 \cdot R_2 \oplus W_1 = \tilde{X}_4, \tag{48}$$
$$R_1 \oplus 2^5 \cdot R_2 \oplus W_2 = \tilde{X}_5, \tag{49}$$
$$W_1 \oplus W_2 = 3 \cdot J, \tag{50}$$
$$Y_0 \oplus Y_4 = Z_4, \tag{51}$$
$$Y_0 \oplus Y_5 = Z_5, \tag{52}$$

so that $\Pr[\mathcal{E}_{(K,Z_4,Z_5)}] = \Pr[\textit{Eqns.}\ (47)\text{-}(52)]$. For fixed (K, Z_4, Z_5), let the number of backward queries at indices 4, 5 and 6 be bounded above by $q_{(K,Z_4,Z_5)}$. For ease of calculations, we will assume $q_{K,i} = q_K$ for all $i \leq 6$ since we are only bothered with upper bounds and each of these is upper bounded by the total number of queries q_K. We classify a computable $\bar{x}$ by the directions in which the Pencil queries x_4, x_5, x_6 are generated:

- We call a computable tuple $\bar{x}$ as $\mathsf{Type\ 1}$ if the last two queries (in the order of execution) from x_4, x_5, x_6 are forward. For a $\mathsf{Type\ 1}$ tuple, equations (51) and (52) are probabilistic, and there is no advantage in fixing (Z_4, Z_5) beforehand. Define $\mathcal{E}^1_{(K,Z_4,Z_5)}$ to be the event that $\mathcal{E}_{(K,Z_4,Z_5)}$ happens with a $\mathsf{Type\ 1}$ witness.

- We call a computable tuple $\bar{x}$ as Type 2 if the last two queries from x_4, x_5, x_6 are backward. In this case, the adversary can set equations (51) and (52) to be true by fixing (Z_4, Z_5) beforehand. Define $\mathcal{E}^2_{(K,Z_4,Z_5)}$ to be the event that $\mathcal{E}_{(K,Z_4,Z_5)}$ happens with a Type 2 witness.
- We call a computable tuple $\bar{x}$ a Type 3 tuple if it is neither Type 1 nor Type 2; that is, exactly one of the last two queries (in order of execution) among x_4, x_5, x_6 is forward and the other backward. The adversary can then force any single one of (51), (52), or

$$Y_4 \oplus Y_5 = Z_4 \oplus Z_5 \tag{53}$$

to hold, by fixing Z_4, Z_5, or $Z_4 \oplus Z_5$ in advance. Since Z_4, Z_5, and $Z_4 \oplus Z_5$ are linearly dependent, any two of them determine the third, fixing one is equivalent to fixing a single coordinate of (Z_4, Z_5), and exactly one of the three equations remains probabilistic. Define $\mathcal{E}^3_{(K,Z_4,Z_5)}$ to be the event that $\mathcal{E}_{(K,Z_4,Z_5)}$ happens with a Type 3 witness.

For each of these cases, we will bound $\Pr[\textit{Eqns.}\ (47)\text{-}(52)]$. First we define some bad events.

- $\mathcal{E}_{\mathsf{bad},1}$: $|\{(x_2,x_3,x_5) \in T_2^K \times T_3^K \times T_5^{K,+} \mid (x_2,x_3,x_5) \text{ satisfies } \textit{Eqn.}\ (48)\}| > q_K^3/2^{n-8} + q_K\sqrt{3nq_K}$.
- $\mathcal{E}_{\mathsf{bad},2}$: $|\{(x_2,x_3,x_5) \in T_2^K \times T_3^K \times T_5^{K,-} \mid (x_2,x_3,x_5) \text{ satisfies } \textit{Eqn.}\ (48)\}| > q_{(K,Z_4,Z_5)}q_K^2/2^{n-8} + q_{(K,Z_4,Z_5)}\sqrt{3nq_K}$.
- $\mathcal{E}_{\mathsf{bad},3}$: $|\{(x_2,x_3,x_6) \in T_2^K \times T_3^K \times T_6^{K,+} \mid (x_2,x_3,x_6) \text{ satisfies } \textit{Eqn.}\ (49)\}| > q_K^3/2^{n-8} + q_K\sqrt{3nq_K}$.
- $\mathcal{E}_{\mathsf{bad},4}$: $|\{(x_2,x_3,x_6) \in T_2^K \times T_3^K \times T_6^{K,-} \mid (x_2,x_3,x_6) \text{ satisfies } \textit{Eqn.}\ (49)\}| > q_{(K,Z_4,Z_5)}q_K^2/2^{n-8} + q_{(K,Z_4,Z_5)}\sqrt{3nq_K}$.
- $\mathcal{E}_{\mathsf{bad},5}$: $|\{(x_2,x_3,x_4) \in T_2^K \times T_3^K \times T_4^{K,-} \mid (x_2,x_3,x_4) \text{ satisfies } \textit{Eqn.}\ (47)\}| > q_{(K,Z_4,Z_5)}q_K^2/2^{n-8} + q_{(K,Z_4,Z_5)}\sqrt{3nq_K}$.

Define $\mathcal{E}_{\mathsf{bad}} := \cup_{i=1}^5 \mathcal{E}_{\mathsf{bad},i}$ to be the union of all bad events defined above. Then, we have the following lemma which we prove in Appendix D.2.

Lemma 14. $\Pr[\mathcal{E}_{bad}] \le 2^{12}/2^n$.

Conditioned on the event $\neg\mathcal{E}_{\mathsf{bad}}$, we proceed to bound $\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}]$. We prove the following lemma in Appendix D.3.

Lemma 15. *For $q_K \ge 2^{2(n-8)/3}$, we have the following bounds:*

$$\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{bad}] \le \frac{2^{25}q_K^4}{2^{4n}},$$
$$\Pr[\mathcal{E}^2_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{bad}] \le \frac{2^{21}q^2_{(K,Z_4,Z_5)}}{2^{2n}},$$
$$\Pr[\mathcal{E}^3_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{bad}] \le \frac{2^{13}q_{(K,Z_4,Z_5)}q_K}{2^{2n}}.$$

We have, for $i = 1, 2, 3$,

$$\begin{aligned}\Pr[\mathcal{E}^i_{(K,Z_4,Z_5)}] &= \Pr[\mathcal{E}^i_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] + \Pr[\mathcal{E}^i_{(K,Z_4,Z_5)} \wedge \mathcal{E}_{\mathsf{bad}}] \\ &\le \Pr[\mathcal{E}^i_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] + \Pr[\mathcal{E}_{\mathsf{bad}}],\end{aligned}$$

and since $\Pr[\mathcal{E}_{\mathsf{bad}}]$ is negligible, it contributes only an additive negligible term to the final bound, which we suppress for clarity.

Type	last two of x_4, x_5, x_6	fixed in advance	$\Pr[\mathcal{E}^i_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \le$
Type 1	both forward	nothing	$2^{25} q_K^4 / 2^{4n}$
Type 2	both backward	Z_4 and Z_5	$2^{21} q^2_{(K,Z_4,Z_5)} / 2^{2n}$
Type 3	one of each	one of $Z_4,\ Z_5,\ Z_4 \oplus Z_5$	$2^{13} q_{(K,Z_4,Z_5)}\, q_K / 2^{2n}$

Figure 10: The three tuple types in the proof of Theorem 6, classified by the directions of the last two K_2-queries. The third column lists the part of (Z_4, Z_5) that a witness of each type can pin down in advance (cf. Eqns. (51)–(53)); the last column restates the bounds of Lemma 15.

Combining two witnesses. It remains to bound

$$\sum_{K \neq K'} \sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}_{(K,Z_4,Z_5)}]\ \Pr[\mathcal{E}_{(K',Z_4,Z_5)}].$$

Expanding each $\mathcal{E}_{(K,Z_4,Z_5)}$ as the union of its three types turns the inner sum

$$\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}_{(K,Z_4,Z_5)}]\ \Pr[\mathcal{E}_{(K',Z_4,Z_5)}] \tag{54}$$

into nine contributions, one for each pair of types for $\bar{x}$ and $\bar{x}'$. Two observations keep this manageable. First, the summand is symmetric under exchanging $(K, \bar{x})$ with $(K', \bar{x}')$, so only six of the nine type-pairs are distinct; the remaining three are obtained by swapping $K \leftrightarrow K'$. Second, by Table 10 the number of (Z_4, Z_5) a pair can jointly realise is governed only by how many coordinates are fixed in advance: a Type 1 factor leaves all 2^{2n} values free, a Type 2 factor fixes both coordinates, and a Type 3 factor fixes exactly one of $Z_4, Z_5, Z_4 \oplus Z_5$. We treat the six pairs in turn, using the per-type bounds of Table 10 throughout.

Both Type 1. Neither factor fixes (Z_4, Z_5), so summing over all 2^{2n} values,

$$\sum_{(Z_4,Z_5)} \Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}]\ \Pr[\mathcal{E}^1_{(K',Z_4,Z_5)}] \ \le\ 2^{2n} \cdot \frac{2^{25} q_K^4}{2^{4n}} \cdot \frac{2^{25} q_{K'}^4}{2^{4n}} \ =\ \frac{2^{50} q_K^4 q_{K'}^4}{2^{6n}}. \tag{55}$$

Both Type 2. Both coordinates are fixed, so the outer sum collapses through $\sum_{(Z_4,Z_5)} q^2_{(K,Z_4,Z_5)} \le q_K^2$:

$$\sum_{(Z_4,Z_5)} \Pr[\mathcal{E}^2_{(K,Z_4,Z_5)}]\ \Pr[\mathcal{E}^2_{(K',Z_4,Z_5)}] \ \le\ \frac{2^{21} q_K^2}{2^{2n}} \cdot \frac{2^{21} q_{K'}^2}{2^{2n}} \ =\ \frac{2^{42} q_K^2 q_{K'}^2}{2^{4n}}. \tag{56}$$

Type 1 with Type 2. The Type 2 factor already fixes (Z_4, Z_5), so the freedom of the Type 1 factor over (Z_4, Z_5) is irrelevant and no 2^{2n} factor appears:

$$\sum_{(Z_4,Z_5)} \Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}]\ \Pr[\mathcal{E}^2_{(K',Z_4,Z_5)}] \ \le\ \frac{2^{25} q_K^4}{2^{4n}} \cdot \frac{2^{21} q_{K'}^2}{2^{2n}} \ =\ \frac{2^{46} q_K^4 q_{K'}^2}{2^{6n}}. \tag{57}$$

By symmetry, (Type 2, Type 1) contributes $2^{46} q_K^2 q_{K'}^4 / 2^{6n}$.

Type 2 with Type 3. Both coordinates are fixed by the Type 2 factor, and bounding $q_{(K',Z_4,Z_5)} \le q_{K'}$ in the Type 3 factor,

$$\sum_{(Z_4,Z_5)} \Pr[\mathcal{E}^2_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}^3_{(K',Z_4,Z_5)}] \;\le\; \frac{2^{21}q_K^2}{2^{2n}} \cdot \frac{2^{13}q_{K'}^2}{2^{2n}} \;=\; \frac{2^{34}q_K^2 q_{K'}^2}{2^{4n}}. \tag{58}$$

The pair (Type 3, Type 2) is identical by symmetry.

Type 1 with Type 3. The Type 3 factor fixes one coordinate, say Z_4; summing the free coordinate against $q_{(K',Z_4)} := \sum_{Z_5} q_{(K',Z_4,Z_5)}$ (so that $\sum_{Z_4} q_{(K',Z_4)} = q_{K'}$),

$$\begin{aligned}\sum_{(Z_4,Z_5)} &\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}^3_{(K',Z_4,Z_5)}] \\ &\le 2^n \sum_{Z_4} \frac{2^{25}q_K^4}{2^{4n}} \cdot \frac{2^{13}q_{(K',Z_4)}q_{K'}}{2^{2n}} \;\le\; \frac{2^{38}q_K^4 q_{K'}^2}{2^{5n}}.\end{aligned} \tag{59}$$

By symmetry, (Type 3, Type 1) contributes $2^{38}q_K^2 q_{K'}^4/2^{5n}$.

Both Type 3. Each factor fixes one of $Z_4, Z_5, Z_4 \oplus Z_5$. If both fix the *same* coordinate, say Z_4,

$$\begin{aligned}\sum_{(Z_4,Z_5)} &\Pr[\mathcal{E}^3_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}^3_{(K',Z_4,Z_5)}] \\ &\le 2^n \sum_{Z_4} \frac{2^{13}q_{(K,Z_4)}q_K}{2^{2n}} \cdot \frac{2^{13}q_{(K',Z_4)}q_{K'}}{2^{2n}} \;\le\; \frac{2^{26}q_K^2 q_{K'}^2}{2^{3n}}.\end{aligned} \tag{60}$$

If they fix *different* coordinates, then between them both coordinates are fixed, and (e.g. $\bar{x}$ fixing Z_4 and $\bar{x}'$ fixing Z_5) the two sums separate to give $2^{26}q_K^2 q_{K'}^2/2^{4n}$; the case where one factor fixes $Z_4 \oplus Z_5$ instead is identical after summing over the pair $(Z_4,\, Z_4 \oplus Z_5)$. Combining the two sub-cases,

$$\sum_{(Z_4,Z_5)} \Pr[\mathcal{E}^3_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}^3_{(K',Z_4,Z_5)}] \;\le\; \frac{2^{26}q_K^2 q_{K'}^2}{2^{3n}} + \frac{2^{26}q_K^2 q_{K'}^2}{2^{4n}}. \tag{61}$$

Summing the six pairs in Eqns. (55)–(61) together with their three symmetric partners, the inner sum (54) is bounded, for sufficiently large q, by

$$\frac{2^{54}q_K^4 q_{K'}^4}{2^{6n}} + \frac{2^{27}q_K^2 q_{K'}^2}{2^{3n}}.$$

Taking the outer sum over $K \ne K'$ with $\sum_K q_K = q$ and applying the bounds $\sum_K q_K^4 \le q^4$ and $\sum_K q_K^2 \le q^2$ yields the bound in Theorem 6. □

Acknowledgements

This work was originally motivated by Shay Gueron's talk on DNDK at RWC 2024. We are grateful to Shay Gueron and Tom Ristenpart for helpful discussions on the design of DNDK and the security proof of DNDKv2. We also thank Ashwin Jha for useful suggestions in relation to the security proof of ♯Pencil.

References

[ABL+14] Elena Andreeva, Andrey Bogdanov, Atul Luykx, Bart Mennink, Nicky Mouha, and Kan Yasuda. How to securely release unverified plaintext in authenticated encryption. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT 2014, Part I*, volume 8873 of *LNCS*, pages 105–125. Springer, Berlin, Heidelberg, December 2014. *(Cited on p. 3.)*

[ADG+22] Ange Albertini, Thai Duong, Shay Gueron, Stefan Kölbl, Atul Luykx, and Sophie Schmieg. How to abuse and fix authenticated encryption without key commitment. In Kevin R. B. Butler and Kurt Thomas, editors, *USENIX Security 2022*, pages 3291–3308. USENIX Association, August 2022. *(Cited on p. 3.)*

[BBN22] Arghya Bhattacharjee, Ritam Bhaumik, and Mridul Nandi. Offset-based BBB-secure tweakable block-ciphers with updatable caches. In Takanori Isobe and Santanu Sarkar, editors, *INDOCRYPT 2022*, volume 13774 of *LNCS*, pages 171–194. Springer, Cham, December 2022. *(Cited on p. 47.)*

[BCF+24] Ritam Bhaumik, André Chailloux, Paul Frixons, Bart Mennink, and Mar4a Naya-Plasencia. Block cipher doubling for a post-quantum world. *IACR Communications in Cryptology*, 1(3), 2024. *(Cited on p. 47.)*

[BD25] Ritam Bhaumik and Jean Paul Degabriele. Pencil: A Domain-Extended PRF with Full n-bit Security for Strengthening GCM and More. Cryptology ePrint Archive, Paper 2025/383, version 20250228:113934, 2025. *(Cited on p. 6.)*

[BGK99] Mihir Bellare, Oded Goldreich, and Hugo Krawczyk. Stateless evaluation of pseudorandom functions: Security beyond the birthday barrier. In Michael J. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 270–287. Springer, Berlin, Heidelberg, August 1999. *(Cited on p. 4.)*

[BH22] Mihir Bellare and Viet Tung Hoang. Efficient schemes for committing authenticated encryption. In Orr Dunkelman and Stefan Dziembowski, editors, *EUROCRYPT 2022, Part II*, volume 13276 of *LNCS*, pages 845–875. Springer, Cham, May / June 2022. *(Cited on p. 5.)*

[BHKK25] Fabrice Benhamouda, Shai Halevi, Panos Kampanakis, and Hugo Krawczyk. Blockcipher-based key derivation without PRP/PRF switching. Cryptology ePrint Archive, Report 2025/878, 2025. *(Cited on p. 6.)*

[BN00] Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In Tatsuaki Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 531–545. Springer, Berlin, Heidelberg, December 2000. *(Cited on p. 3.)*

[BT16] Mihir Bellare and Björn Tackmann. The multi-user security of authenticated encryption: AES-GCM in TLS 1.3. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 247–276. Springer, Berlin, Heidelberg, August 2016. *(Cited on pp. 4 and 9.)*

[CDJN25] Yu Long Chen, Avijit Dutta, Ashwin Jha, and Mridul Nandi. Towards Optimally Secure Deterministic Authenticated Encryption Schemes. In Serge Fehr and Pierre-Alain Fouque, editors, *EUROCRYPT 2025*, volume 15601 of *LNCS*, pages 3–32. Springer, 2025. *(Cited on p. 6.)*

[CDN+23] Beno4t Cogliati, Avijit Dutta, Mridul Nandi, Jacques Patarin, and Abishanka Saha. Proof of mirror theory for a wide range of $\xi_{\max}$. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023, Part IV*, volume 14007 of *LNCS*, pages 470–501. Springer, Cham, April 2023. *(Cited on pp. 11 and 47.)*

[CHK+25] Woohyuk Chung, Seongha Hwang, Seongkwang Kim, Byeonghak Lee, and Jooyoung Lee. Making GCM Great Again: Toward Full Security and Longer Nonces. In Serge Fehr and Pierre-Alain Fouque, editors, *EUROCRYPT 2025*, volume 15601 of *LNCS*, pages 33–61. Springer, 2025. *(Cited on p. 6.)*

[CJN20] Bishwajit Chakraborty, Ashwin Jha, and Mridul Nandi. On the security of sponge-type authenticated encryption modes. *IACR Trans. Symmetric Cryptol.*, 2020(2):93–119, 2020. *(Cited on p. 31.)*

[CLL24] Wonseok Choi, Jooyoung Lee, and Yeongmin Lee. Toward full n-bit security and nonce misuse resistance of block cipher-based macs. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024*, volume 15492 of *LNCS*, pages 251–279. Springer, 2024. *(Cited on p. 47.)*

[CLP15] Benoit Cogliati, Rodolphe Lampe, and Jacques Patarin. The indistinguishability of the XOR of k permutations. In Carlos Cid and Christian Rechberger, editors, *FSE 2014*, volume 8540 of *LNCS*, pages 285–302. Springer, Berlin, Heidelberg, March 2015. *(Cited on p. 47.)*

[CS14] Shan Chen and John P. Steinberger. Tight security bounds for key-alternating ciphers. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 327–350. Springer, Berlin, Heidelberg, May 2014. *(Cited on p. 10.)*

[CS18] Beno4t Cogliati and Yannick Seurin. Analysis of the single-permutation encrypted davies-meyer construction. *Des. Codes Cryptogr.*, 86(12):2703–2723, 2018. *(Cited on p. 31.)*

[DNS22] Avijit Dutta, Mridul Nandi, and Abishanka Saha. Proof of Mirror Theory for $\xi_{\max} = 2$. *IEEE Trans. Inf. Theory*, 68(9):6218–6232, 2022. *(Cited on p. 47.)*

[FOR17] Pooya Farshim, Claudio Orlandi, and Răzvan Roşie. Security of symmetric primitives under incorrect usage of keys. *IACR Trans. Symm. Cryptol.*, 2017(1):449–473, 2017. *(Cited on p. 5.)*

[GL17] Shay Gueron and Yehuda Lindell. Better bounds for block cipher modes of operation via nonce-based key derivation. In Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *ACM CCS 2017*, pages 1019–1036. ACM Press, October / November 2017. *(Cited on p. 3.)*

[GLL17] Shay Gueron, Adam Langley, and Yehuda Lindell. AES-GCM-SIV: Specification and analysis. Cryptology ePrint Archive, Report 2017/168, 2017. *(Cited on p. 3.)*

[GLR17] Paul Grubbs, Jiahui Lu, and Thomas Ristenpart. Message franking via committing authenticated encryption. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part III*, volume 10403 of *LNCS*, pages 66–97. Springer, Cham, August 2017. *(Cited on p. 3.)*

[GR25] Shay Gueron and Thomas Ristenpart. DNDK: Combining nonce and key derivation for fast and scalable AEAD. Cryptology ePrint Archive, Paper 2025/785, 2025. *(Cited on pp. 5 and 32.)*

[Gue24a] Shay Gueron. Double Nonce Derive Key AES-GCM (DNDK-GCM). Internet-Draft draft-gueron-cfrg-dndkgcm-00, Internet Engineering Task Force, April 2024. Work in Progress. *(Cited on pp. 3, 4, and 32.)*

[Gue24b] Shay Gueron. Double-nonce-derive-key-gcm (dndk-gcm), 2024. Presented at the IACR Real World Crypto Symposium 2024. *(Cited on p. 3.)*

[HTT18] Viet Tung Hoang, Stefano Tessaro, and Aishwarya Thiruvengadam. The multi-user security of GCM, revisited: Tight bounds for nonce randomization. In David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang, editors, *ACM CCS 2018*, pages 1429–1440. ACM Press, October 2018. *(Cited on pp. 4 and 9.)*

[IMV16] Tetsu Iwata, Bart Mennink, and Damian Vizár. CENC is optimally secure. Cryptology ePrint Archive, Report 2016/1087, 2016. *(Cited on pp. 4 and 47.)*

[IOM12] Tetsu Iwata, Keisuke Ohashi, and Kazuhiko Minematsu. Breaking and repairing GCM security proofs. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 31–49. Springer, Berlin, Heidelberg, August 2012. *(Cited on p. 3.)*

[Iwa06] Tetsu Iwata. New blockcipher modes of operation with beyond the birthday bound security. In Matthew J. B. Robshaw, editor, *FSE 2006*, volume 4047 of *LNCS*, pages 310–327. Springer, Berlin, Heidelberg, March 2006. *(Cited on pp. 4 and 11.)*

[JN22] Ashwin Jha and Mridul Nandi. A Survey on Applications of H-Technique: Revisiting Security Analysis of PRP and PRF. *Entropy*, 24(4):462, 2022. *(Cited on p. 10.)*

[Jut01] Charanjit S. Jutla. Encryption modes with almost free message integrity. In Birgit Pfitzmann, editor, *EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 529–544. Springer, Berlin, Heidelberg, May 2001. *(Cited on p. 3.)*

[KCC+23] Panos Kampanakis, Matt Campagna, Eric Crocket, Adam Petcher, and Shay Gueron. Practical challenges with aes-gcm and the need for a new cipher. In *Third NIST Workshop on Block Cipher Modes of Operation*, 2023. *(Cited on p. 3.)*

[KHEC25] Panos Kampanakis, Shai Halevi, Nevine Ebeid, and Matt Campagna. Blockcipher-based key commitment for nonce-derived schemes. Cryptology ePrint Archive, Paper 2025/758, 2025. *(Cited on pp. 5 and 6.)*

[LGR21] Julia Len, Paul Grubbs, and Thomas Ristenpart. Partitioning oracle attacks. In Michael Bailey and Rachel Greenstadt, editors, *USENIX Security 2021*, pages 195–212. USENIX Association, August 2021. *(Cited on p. 3.)*

[MN17] 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 2017, Part III*, volume 10403 of *LNCS*, pages 556–583. Springer, Cham, August 2017. *(Cited on p. 47.)*

[MV04] David A. McGrew and John Viega. The security and performance of the Galois/counter mode (GCM) of operation. In Anne Canteaut and Kapalee Viswanathan, editors, *INDOCRYPT 2004*, volume 3348 of *LNCS*, pages 343–355. Springer, Berlin, Heidelberg, December 2004. *(Cited on p. 3.)*

[Nat25] National Institute of Standards and Technology. Pre-draft call for comments: GCM and GMAC block cipher modes of operation, Jan 2025. *(Cited on p. 3.)*

[Pat03] Jacques Patarin. Luby-Rackoff: 7 rounds are enough for 2n(1-epsilon)security. In Dan Boneh, editor, *CRYPTO 2003*, volume 2729 of *LNCS*, pages 513–529. Springer, Berlin, Heidelberg, August 2003. *(Cited on pp. 10 and 47.)*

[Pat06] Jacques Patarin. On linear systems of equations with distinct variables and small block size. In Dongho Won and Seungjoo Kim, editors, *ICISC 05*, volume 3935 of *LNCS*, pages 299–321. Springer, Berlin, Heidelberg, December 2006. *(Cited on pp. 10 and 47.)*

[Pat08] Jacques Patarin. A proof of security in O(2n) for the xor of two random permutations. In Reihaneh Safavi-Naini, editor, *ICITS 08*, volume 5155 of *LNCS*, pages 232–248. Springer, Berlin, Heidelberg, August 2008. *(Cited on p. 47.)*

[Pat09] Jacques Patarin. The "coefficients H" technique (invited talk). In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *SAC 2008*, volume 5381 of *LNCS*, pages 328–345. Springer, Berlin, Heidelberg, August 2009. *(Cited on p. 10.)*

[Pat10a] Jacques Patarin. Introduction to mirror theory: Analysis of systems of linear equalities and linear non equalities for cryptography. Cryptology ePrint Archive, Report 2010/287, 2010. *(Cited on p. 47.)*

[Pat10b] Jacques Patarin. Transfinite cryptography. Cryptology ePrint Archive, Report 2010/001, 2010. *(Cited on pp. 10 and 47.)*

[Rog02] Phillip Rogaway. Authenticated-encryption with associated-data. In Vijayalakshmi Atluri, editor, *ACM CCS 2002*, pages 98–107. ACM Press, November 2002. *(Cited on p. 3.)*

[RS06] Phillip Rogaway and Thomas Shrimpton. A provable-security treatment of the key-wrap problem. In Serge Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 373–390. Springer, Berlin, Heidelberg, May / June 2006. *(Cited on p. 3.)*

[Val24] Filippo Valsorda. The xaes-256-gcm extended-nonce aead. https://github.com/C2SP/C2SP/blob/main/XAES-256-GCM.md, 2024. *(Cited on p. 5.)*

[ZHY18] Ping Zhang, Honggang Hu, and Qian Yuan. Close to Optimally Secure Variants of GCM. *Secur. Commun. Networks*, 2018:9715947:1–9715947:12, 2018. *(Cited on p. 47.)*

Appendix

A Some Useful Results in Probability

Markov's Inequality. For a non-negative random variable X and some $a > 0$, *Markov's Inequality* states that

$$\Pr[X \geq a] \leq \frac{\mathbb{E}(X)}{a}.$$

Sum-Capture Lemma. Let $\mathcal{S}$ be a sequence of uniformly random and independently chosen elements of $\{0,1\}^n$ with $|\mathcal{S}| = p$, and define

$$\delta_q(\mathcal{S}) = \max_{\substack{\mathcal{X},\mathcal{Y}\subset\{0,1\}^n \\ |\mathcal{X}|=|\mathcal{Y}|=q}} |\{(a,b,c) \in \mathcal{S}\times\mathcal{X}\times\mathcal{Y} : a = b \oplus c\}|.$$

Lemma 16 (a variant of sum-capture lemma [CS18])**.**

$$\Pr\left[\delta_q(\mathcal{S}) \geq \frac{pq^2}{2^n} + p\sqrt{3nq}\right] \leq \frac{2}{2^n}.$$

Multicollision bound. Let $\mathcal{S}$ be a sequence with elements from a set $\mathcal{X}$. For any $x \in \mathcal{X}$, we define its multiplicity in $\mathcal{S}$ as $\mathsf{mul}_{\mathcal{S}}(x) := |\{y \in \mathcal{S} : y = x\}|$, i.e., the number of times x appears in the multiset. Finally, we define multicollision of $\mathcal{S}$ as $\mathsf{mc}(\mathcal{S}) := \max_{x \in \mathcal{X}} \mathsf{mul}_{\mathcal{S}}(x)$. For $n \geq 2$, define

$$\mathsf{mcoll}(p, 2^n) = \begin{cases} \frac{4 \log_2 p}{\log_2 \log_2 p} & \text{if } 4 < p \leq 2^n \\ 5n \left\lceil \frac{p}{n \cdot 2^n} \right\rceil & \text{if } p > 2^n. \end{cases}$$

Lemma 17 (multicollision bound on random samples [CJN20])**.** *Let $\mathcal{X}$ be the set $\{0,1\}^n$. Given random variables $X_1, \ldots, X_p \leftarrow\!\$\, \mathcal{X}$, we have*

$$\mathbb{E}[mc(X_1, \ldots, X_p)] \leq \mathsf{mcoll}(p, 2^n).$$

B Design Rationale of Pencil

B.1 Designing a weak PRF

To motivate the final design adopted for Pencil, we present first a more naive approach towards designing a weak PRF that accepts a 2-block random input, why it doesn't work, and how we can circumvent this problem.

Naive Approach. We wish to design a weak PRF mode F based on a block cipher E of width n bits which accepts a random nonce of length $2n - 2\mu$ bits, and outputs a keystream of ℓn bits. As a first approach, we may try to mimic XORP as follows: we split the nonce in half as $R_1 \| R_2$, and reserve the last μ bits of each call to E_K for domain separation (we assume $\mu \ll n$); then we take each output block to be of the form $F_{i,j}(R_1, R_2) := E_K(R_1 \| \langle i \rangle_\mu) \oplus E_K(R_2 \| \langle j \rangle_\mu)$ for some $i, j \in [0..2^\mu - 1]$. Let $\mathcal{I} \subseteq [0..2^\mu - 1]^2$ be the index set of size ℓ such that $F(R_1, R_2)$ is made up of the blocks $\{F_{i,j}(R_1, R_2) \mid (i,j) \in \mathcal{I}\}$ in some order.

Consider the bipartite graph $\mathcal{G}$ where the vertex set is $\{1,2\} \times [0..2^\mu - 1]$, and the edge set is $\{\{(1,i), (2,j)\} \mid (i,j) \in \mathcal{I}\}$, i.e., there is an edge between $(1,i)$ and $(2,j)$ if one of the output blocks is $F_{i,j}(R_1, R_2)$. Then it is clear that we cannot allow a cycle in $\mathcal{G}$ as it leads to a trivial distinguisher—for instance, if $(i,j), (i,j'), (i',j'), (i',j) \in \mathcal{I}$ with $i \neq i, j \neq j$ (which implies that $(1,i), (2,j), (1,i'), (2,j')$ lie on a cycle in $\mathcal{G}$), the distinguishing event can be $F_{i,j}(R_1, R_2) \oplus F_{i,j'}(R_1, R_2) \oplus F_{i',j'}(R_1, R_2) \oplus F_{i',j}(R_1, R_2) = 0$, which will always be true for F but only has a probability $\approx 1/2^n$ for an ideal random function; a similar attack works when longer cycles exist in $\mathcal{G}$. Thus, we can get at most $2^\mu - 1$ output blocks from 2^μ calls to E_K, and this happens when $\mathcal{G}$ is a tree.

Why this doesn't work. This approach fails if we want BBB security from the weak PRF. This is because if at least one vertex $(1,i)$ has two neighbours $(2,j)$ and $(2,j')$ in $\mathcal{G}$, then we have two output blocks $Z := F_{i,j}(R_1, R_2)$ and $Z' := F_{i,j'}(R_1, R_2)$. Then $Z \oplus Z' = E_K(R_2 \| \langle j \rangle_\mu) \oplus E_K(R_2 \| \langle j' \rangle_\mu)$ is a function of R_2 alone, and any collision on R_2 (which should happen within $2^{n/2}$ samplings of (R_1, R_2)) leads to a collision on $Z \oplus Z'$, resulting in a distinguishing attack.

Using a Linear Transformation. We observe that the above problem is caused by the fact that each E_K call is function of exactly one of R_1 and R_2, so that a collision in either leads to a collision in the outputs of *all* calls involving it. We can get around this by generating 2^μ independent linear combinations of R_1 and R_2, to be used with the 2^μ domain separators. The idea is that if two of them are to collide simultaneously (which is needed for the above attack), it should have about the same probability as a collision over the whole $2n - 2\mu$ bits. A simple choice is to use the following inputs: $(R_1 \oplus R_2)\|\langle 0\rangle_\mu$, $(R_1 \oplus 2 \cdot R_2)\|\langle 1\rangle_\mu$, $\ldots$, $(R_1 \oplus 2^{2^\mu-1} \cdot R_2)\|\langle 2^\mu - 1\rangle_\mu$. This has the advantage that all the inputs can be computed using $2^\mu - 1$ multiplications by 2, and no other field multiplications are needed.

Resistance to attack. For the above attack to work, we need to find a pair of nonces $R_1\|R_2$ and $R'_1\|R'_2$ and indices i and j such that the inputs to i for the two nonces collide with each other, and the inputs to j collide with each other. This is equivalent to finding $R_1\|R_2$ and $R'_1\|R'_2$ such that

$$R_1 \oplus 2^i \cdot R_2 = R'_1 \oplus 2^i \cdot R'_2,$$
$$R_1 \oplus 2^j \cdot R_2 = R'_1 \oplus 2^j \cdot R'_2.$$

Together they imply that $(2^i \oplus 2^j) \cdot (R_2 \oplus R'_2) = 0$, from which we can see that $R_2 = R'_2$ and hence $R_1 = R'_1$. Thus, this needs a collision on the entire $(2n - 2\mu)$-bit nonce.

Computing PRF outputs. The 2^μ outputs of E_K for a nonce $R_1\|R_2$ can be arranged in a tree to produce $2^\mu - 1$ output blocks. The specific choice of the tree does not affect security, since all pairwise sums of the E_K outputs can be computed from any tree. For the design we propose, we choose the same tree as in XORP, fixing the first call as a 'root' and treating the other calls as 'leaves', so that the output of each leaf is XOR-ed to the output of the root to get one block of final output.

B.2 Adding Suffix-Commitment Security

Since AES-GCM is not inherently key-committing, we would like to follow the approach of [Gue24a, GR25] and have our PRF output a separate commitment tag, with the goal that it should be difficult to generate the same tag from two different keys. Since a collision attack is always possible on the tag, commitment security cannot be more than birthday-bound in the length of the tag. Since we are aiming for a commitment security close to n bits, we opt for a tag length of $2n$ bits.

Straightforward Approach. For generating a $2n$-bit commitment tag, a straightforward approach would be to just extend the PRF output by two blocks and output the final two blocks at the tag. Following the scheme in the previous subsection, suppose (R_1, R_2) are obtained from the input (non-random) nonce through an invertible preprocessing step. Denoting $\tilde{X}_j := R_1 \oplus 2^j \cdot R_2$, $X_j := \tilde{X}_j\|\langle j\rangle_\mu$, and $Y_j := E_K(X_j)$, we get the j-th output block as $Z_j := Y_0 \oplus Y_j$, so the commitment tag is $(Z_{\ell+1}, Z_{\ell+2})$ if the PRF was earlier generating an ℓn-bit keystream.

Commitment Attack. Consider the following algorithm:

- Take as input K and Z;
- Set $Z_{\ell+1} := Z$;
- Repeat the following until $\mathsf{lsb}_\mu(X_{\ell+1}) = \langle \ell + 1\rangle_\mu$:
  - Sample $\tilde{X}_0 \xleftarrow{\$} \{0,1\}^{n-\mu}$ and set $X_0 := \tilde{X}_0\|\langle 0\rangle_\mu$;

– Compute $Y_0 := E_K(X_0)$, $Y_{\ell+1} := Y_0 \oplus Z_1$, and $X_{\ell+1} := E_K^{-1}(Y_{\ell+1})$;

- Set $\tilde{X}_{\ell+1} := \mathsf{msb}_{n-\mu}(X_{\ell+1})$;
- Compute $R_1 := (1 \oplus 2^{\ell+1})^{-1} \cdot (2^{\ell+1} \cdot \tilde{X}_0 \oplus \tilde{X}_{\ell+1})$ and $R_2 := (1 \oplus 2^{\ell+1})^{-1} \cdot (\tilde{X}_0 \oplus \tilde{X}_{\ell+1})$;
- Set $X_{\ell+2} := \tilde{X}_{\ell+2} \| \langle \ell + 2 \rangle_\mu$;
- Compute $Y_{\ell+2} := E_K(X_{\ell+2})$ and $Z_{\ell+2} := Y_0 \oplus Y_{\ell+2}$;
- Return $(Z_{\ell+1}, Z_{\ell+2})$.

On an average the loop should terminate by 2^μ iterations, and for any input (K, Z) the algorithm returns a tag $(Z_{\ell+1}, Z_{\ell+2})$ generated from K where $Z_{\ell+1} = Z$ and $Z_{\ell+2}$ is randomly generated. By executing the algorithm around $2^{n/2}$ times each for two different keys K and K' and the same fixed value Z, we can expect to find a pair of executions, one from K and one from K', that return the same $Z_{\ell+2}$). Since the preprocessing layer can be inverted, this leads to a commitment attack in about $2^{n/2+\mu}$ total iterations. Since μ is typically small ($\mu = 8$ in our final construction), this straightforward approach thus fails to give any significant commitment security beyond $n/2$.

C Deferred Proofs from Sec. 3

C.1 Proof of Lemma 4

For each oracle $\mathcal{O}$, let the stochastic component of $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}})$ be split into $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}}; 1)$, $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}}; 2)$, and $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}}; 3)$), to denote the intermediate values $(\mathbf{J}, \mathbf{R})$, final output $\mathbf{Z}$, and sampled block cipher outputs $\mathbf{Y}$ respectively. Note that $\mathbf{W}$ and $\mathbf{X}$ can be computed given $\mathbf{R}$ and the stochastic outputs, and can be dropped from the transcript for the probability calculations.

Fix a $\mathsf{tr} = (\mathbf{N}, \mathbf{Z}; \mathbf{J}, \mathbf{R}, \mathbf{Y}) \in \mathcal{T}_{\mathsf{good}}$. By definition of $\mathcal{T}_{\mathsf{bad}}$, we know that $\Phi_{\mathsf{tr}} := \Phi[\mathbf{A}_{r \times 5q}, \mathbf{Z}]$ is a consistent mirror system and $\xi_{\max}(\mathsf{tr}) \leq n\ell$. First we observe that since the preprocessing step is identical for both oracles, we have

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 1) = (\mathbf{J}, \mathbf{R})\right] = \Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}; 1) = (\mathbf{J}, \mathbf{R})\right] = \rho(\mathbf{J}, \mathbf{R}) \text{ (say)}. \tag{62}$$

For the rest of the transcript, we examine the real world and the ideal world separately. Since $\mathbf{A}$ has r rows, there must be r distinct values of $X_j^{(i)}$ in $\mathbf{X}$. Since the corresponding outputs are sampled uniformly without replacement by π_2 in the real world, we have

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 3) = \mathbf{Y} \mid \mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 1) = (\mathbf{J}, \mathbf{R})\right] = \frac{1}{(2^n)_r}. \tag{63}$$

Finally, in the real world, $\mathbf{Z}$ is determined from $\mathbf{Y}$, so we have

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 2) = \mathbf{Z} \mid \mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 1) = (\mathbf{J}, \mathbf{R}), \mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}; 3) = \mathbf{Y}\right] = 1. \tag{64}$$

From Eqns. (62), (63), and (64) we have

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_1}) = \mathsf{tr}\right] = \frac{\rho(\mathbf{J}, \mathbf{R})}{(2^n)_r}. \tag{65}$$

Next we turn our attention to the ideal world. Here, $(\mathcal{A}^{\mathcal{O}_0}; 2)$ is sampled first, and independently of $(\mathcal{A}^{\mathcal{O}_0}; 1)$, so we have

$$\Pr\left[\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}; 2) = \mathbf{Y} \mid \mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0}; 1) = (\mathbf{J}, \mathbf{R})\right] = \frac{1}{2^{5nq}}. \tag{66}$$

In the final step, sampling $(\mathcal{A}^{\mathcal{O}_0}; 3)$ is equivalent to a uniform sampling from $\Gamma(\Phi_{\mathsf{tr}})$, so we have

$$\Pr\left[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}; 3\right) = \mathbf{Y} \mid \mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}; 1\right) = (\mathbf{J}, \mathbf{R}), \mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}; 2\right) = \mathbf{Y}\right] = \frac{1}{|\Gamma(\Phi_{\mathsf{tr}})|}. \tag{67}$$

From Eqns. (62), (66), and (67) we have

$$\Pr\left[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right) = \mathsf{tr}\right] = \frac{\rho(\mathbf{J}, \mathbf{R})}{2^{5nq}|\Gamma(\Phi_{\mathsf{tr}})|}. \tag{68}$$

To complete the proof, we need to find an appropriate bound for the ratio of $\Pr[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_1}\right) = \mathsf{tr}]$ and $\Pr[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right) = \mathsf{tr}]$. From Eqns. (65) and (68) we have

$$\frac{\Pr\left[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_1}\right) = \mathsf{tr}\right]}{\Pr\left[\mathsf{Trs}\left(\mathcal{A}^{\mathcal{O}_0}\right) = \mathsf{tr}\right]} = \frac{2^{5nq}|\Gamma(\Phi_{\mathsf{tr}})|}{(2^n)_r}. \tag{69}$$

As the final step in the proof, we want to bound $|\Gamma(\Phi_{\mathsf{tr}})|$ by applying Theorem 3. We already know that Φ_{tr} is a consistent mirror system, so we just need to verify the conditions on $\xi_{\max}(\mathsf{tr})$. Using the bound $\xi_{\max}(\mathsf{tr}) \leq 5n$, we have

$$n\xi_{\max}(\mathsf{tr})^2 + \xi_{\max}(\mathsf{tr}) \leq 25n^3 + 5n \leq 2^{n/2}, \tag{70}$$

and

$$12 \cdot 5q\xi_{\max}(\mathsf{tr}) \leq 300nq \leq 2^n. \tag{71}$$

Thus, all the conditions for Theorem 3 are satisfied, and applying this theorem gives

$$|\Gamma(\Phi_{\mathsf{tr}})| \geq \frac{(2^n)_r}{2^{5nq}}. \tag{72}$$

Substituting Eqn. (72) in Eqn. (69) completes the proof of Lemma 4. □

C.2 Proof of Lemma 5

For each pair $(i_1, i_2) \in [1..q]^2$, let $\mathcal{I}_{i_1,i_2}$ be the indicator random variable for membership in $\mathcal{S}_{\mathsf{ar}}$, defined as

$$\begin{aligned} \mathcal{I}_{i_1,i_2} &= 1, \quad \text{if } (i_1, i_2) \in \mathcal{S}_{\mathsf{ar}}, \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

If $N_2^{(i_1)} = N_2^{(i_2)}$, then $W_1^{(i_1)} \neq W_1^{(i_2)}$ and $W_2^{(i_1)} \neq W_2^{(i_2)}$, so (i_1, i_2) can only form an arm if $N_2^{(i_1)} \neq N_2^{(i_2)}$, so that $J^{(i_1)}$ and $J^{(i_2)}$ are outputs of π_1 on two different inputs. We observe that

$$\begin{aligned} \Pr\left[W_1^{(i_1)} = W_1^{(i_2)}\right] &= \Pr\left[N_1^{(i_1)} \oplus J^{(i_1)} = N_1^{(i_2)} \oplus J^{(i_2)}\right] = \frac{2^8 - 1}{2^n - 1} \leq \frac{1}{2^{n-8}}, \\ \Pr\left[W_2^{(i_1)} = W_2^{(i_2)}\right] &= \Pr\left[N_1^{(i_1)} \oplus 2 \cdot J^{(i_1)} = N_1^{(i_2)} \oplus 2 \cdot J^{(i_2)}\right] \leq \frac{1}{2^{n-8}}. \end{aligned}$$

By union-bound, we have

$$\begin{aligned} \mathbb{E}\left[\mathcal{I}_{i_1,i_2}\right] &= \Pr\left[(i_1, i_2) \in \mathcal{S}_{\mathsf{ar}}\right] \\ &\leq \Pr\left[W_1^{(i_1)} = W_1^{(i_2)}\right] + \Pr\left[W_2^{(i_1)} = W_2^{(i_2)}\right] \leq \frac{1}{2^{n-9}}. \end{aligned}$$

Since $|\mathcal{S}_{\mathsf{ar}}| = \sum_{(i_1,i_2)\in[1..q]^2} I_{i_1,i_2}$, we have

$$\mathbb{E}[|\mathcal{S}_{\mathsf{ar}}|] = \sum_{(i_1,i_2)\in[1..q]^2} \mathbb{E}[I_{i_1,i_2}] \leq \frac{q^2}{2^{n-9}}.$$

By applying Markov's Inequality, we get

$$\Pr[\mathcal{E}_{\mathsf{ar}}] = \Pr[|\mathcal{S}_{\mathsf{ar}}| > q] \leq \frac{\mathbb{E}[|\mathcal{S}_{\mathsf{ar}}|]}{q} \leq \frac{q}{2^{n-9}},$$

which establishes the bound in Eqn. (17). □

C.3 Proof of Lemma 6

For each triple $(i_1, i_2, i_3) \in [1..q]^3$, let $\mathcal{I}_{i_1,i_2,i_3}$ be the indicator random variable for membership in $\mathcal{S}_{\mathsf{el}}$, defined as

$$\begin{aligned} \mathcal{I}_{i_1,i_2,i_3} &= 1, \quad \text{if } (i_1, i_2, i_3) \in \mathcal{S}_{\mathsf{el}}, \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

If $N_2^{(i_1)} = N_2^{(i_2)}$, then $W_1^{(i_1)} \neq W_1^{(i_2)}$, and if $N_2^{(i_2)} = N_2^{(i_3)}$, then $W_2^{(i_2)} \neq W_2^{(i_3)}$. Thus, the event $W_1^{(i_1)} = W_1^{(i_2)}, W_2^{(i_2)} = W_2^{(i_3)}$ can have non-zero probability only when $N_2^{(i_2)} \notin \{N_2^{(i_1)}, N_2^{(i_3)}\}$. In this case, we can write this event as

$$N_1^{(i_1)} \oplus J^{(i_1)} = N_1^{(i_2)} \oplus J^{(i_2)}, \tag{73}$$

$$N_1^{(i_2)} \oplus 2 \cdot J^{(i_2)} = N_1^{(i_3)} \oplus 2 \cdot J^{(i_3)}. \tag{74}$$

Since $N_2^{(i_1)} \neq N_2^{(i_2)}$, $J^{(i_1)}$ and $J^{(i_2)}$ are outputs of π_1 on two different inputs. Thus,

$$\Pr\left[Eqn.\ (73)\right] \leq \frac{1}{2^{n-8}}. \tag{75}$$

Now we consider two cases. If in addition $N_2^{(i_1)} \neq N_2^{(i_3)}$, then $J^{(i_3)}$ is the output of π_1 on yet another different input, so that

$$\Pr\left[Eqn.\ (74) \mid Eqn.\ (73)\right] = \frac{2^8 - 1}{2^n - 2} \leq \frac{1}{2^{n-8}}. \tag{76}$$

From Eqns. (75) and (76), we get

$$\begin{aligned} \Pr\left[W_1^{(i_1)} = W_1^{(i_2)}, W_2^{(i_2)} = W_2^{(i_3)}\right] &= \Pr\left[Eqns.\ (73),\ (74)\right] \\ &= \Pr\left[Eqn.\ (73)\right] \Pr\left[Eqn.\ (74) \mid Eqn.\ (73)\right] \\ &\leq \frac{1}{2^{2n-16}}. \end{aligned} \tag{77}$$

This leaves the case $N_2^{(i_1)} = N_2^{(i_3)} \implies J^{(i_1)} = J^{(i_3)}$. Multiplying Eqn. (73) by 2, adding to Eqn. (74), setting $J^{(i_1)} = J^{(i_3)}$, and rearranging the terms gives

$$N_1^{(i_3)} = 3 \cdot N_1^{(i_2)} \oplus 2 \cdot N_1^{(i_1)}. \tag{78}$$

Define the set $\mathcal{S}^* \subseteq [1..q]^3$ as

$$\mathcal{S}^* := \left\{(i_1, i_2, i_3) \mid N_1^{(i_3)} = 3 \cdot N_1^{(i_2)} \oplus 2 \cdot N_1^{(i_1)}, N_2^{(i_3)} = N_2^{(i_1)}\right\}.$$

For $(i_1, i_2, i_3) \in \mathcal{S}^*$, Eqn. (73) implies Eqn. (74), so we have

$$\Pr\left[W_1^{(i_1)} = W_1^{(i_2)}, W_2^{(i_2)} = W_2^{(i_3)}\right] = \Pr\left[Eqn.\ (73)\right] = \frac{1}{2^{n-8}}. \tag{79}$$

If $(i_1, i_2, i_3) \notin \mathcal{S}^*$ and yet $N_2^{(i_3)} = N_2^{(i_1)}$, $\Pr[W_1^{(i_1)} = W_1^{(i_2)}, W_2^{(i_2)} = W_2^{(i_3)}] = 0$, because Eqn. (78) does not hold. From Eqns. (77) and (79), we can conclude that

$$\mathbb{E}[\mathcal{I}_{i_1,i_2,i_3}] = \Pr[(i_1, i_2, i_3) \in \mathcal{S}_{\mathsf{el}}] \leq \begin{cases} \frac{1}{2^{n-8}}, & \text{if } (i_1, i_2, i_3) \in \mathcal{S}^*, \\ \frac{1}{2^{2n-16}}, & \text{otherwise.} \end{cases}$$

Since $|\mathcal{S}_{\mathsf{el}}| = \sum_{(i_1,i_2,i_3)\in[1..q]^3} I_{i_1,i_2,i_3}$, we have

$$\begin{aligned} \mathbb{E}[|\mathcal{S}_{\mathsf{el}}|] &= \sum_{(i_1,i_2,i_3)\in[1..q]^3} \mathbb{E}[I_{i_1,i_2,i_3}] \\ &= \sum_{(i_1,i_2,i_3)\in\mathcal{S}^*} \mathbb{E}[I_{i_1,i_2,i_3}] + \sum_{(i_1,i_2,i_3)\notin\mathcal{S}^*} \mathbb{E}[I_{i_1,i_2,i_3}] \\ &\leq \sum_{(i_1,i_2,i_3)\in\mathcal{S}^*} \frac{1}{2^{n-8}} + \sum_{(i_1,i_2,i_3)\notin\mathcal{S}^*} \frac{1}{2^{2n-16}} \\ &\leq \frac{|\mathcal{S}^*|}{2^{n-8}} + \frac{q^3 - |\mathcal{S}^*|}{2^{2n-16}} \leq \frac{|\mathcal{S}^*|}{2^{n-8}} + \frac{q^3}{2^{2n-16}}. \end{aligned} \tag{80}$$

To bound $|\mathcal{S}^*|$, we observe that for each $(i_1, i_2, i_3) \in \mathcal{S}^*$, fixing i_1 and i_2 completely determines i_3. Thus, $|\mathcal{S}^*| \leq q^2$. Substituting this bound in Eqn. (80) gives

$$\mathbb{E}[|\mathcal{S}_{\mathsf{el}}|] \leq \frac{q^2}{2^{n-8}} + \frac{q^3}{2^{2n-16}}.$$

By applying Markov's Inequality, we get

$$\Pr[\mathcal{E}_{\mathsf{el}}] = \Pr[|\mathcal{S}_{\mathsf{el}}| > q] \leq \frac{\mathbb{E}[|\mathcal{S}_{\mathsf{el}}|]}{q} \leq \frac{q}{2^{n-8}} + \frac{q^2}{2^{2n-16}},$$

which establishes the bound in Eqn. (18). □

C.4 Proof of Lemma 7

A chain ζ can be treated interchangeably with the event of the chain forming, i.e., the conjunction of all the collisions necessary for ζ to satisfy the definition of a chain. In this way we can make sense of an expression like $\Pr[\zeta]$. For $2 < d \leq k$ and a k-chain ζ, let $\zeta_{<d}$ denote the $(d-1)$-chain $(i_1, \ldots, i_{d-1})$ (or, equivalently, the conjunction of the $d-2$ collisions required for this chain), and let ζ_d denote the collision required to link i_{d-1} to i_d, i.e., the event that for some $j_{d-1} \in [0..5]$, $X_{j_{d-1}}^{(i_d)} = X_{j_{d-1}}^{(i_{d-1})}$. Note that by this notation,

$$\zeta = \bigwedge_{d=2}^{k} \zeta_d.$$

We observe that if $W_1^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \leq a < d\}$, $R_1^{(i_d)}$ will be the truncated output of $\pi_1(\cdot \| \texttt{0x01})$ on an input that did not appear earlier in the chain, and similarly, if $W_2^{(i_d)} \notin \{W_2^{(i_a)} \mid 1 \leq a < d\}$, $R_2^{(i_d)}$ will

be the truncated output of $\pi_1(\cdot\|\texttt{0x02})$ on an input that did not appear earlier in the chain. Thus, if i_d is fresh, at least one of $R_1^{(i_d)}$ and $R_2^{(i_d)}$ will be a source of fresh randomness after conditioning on $\zeta_{<d}$. This is a key feature of fresh nodes that we will repeatedly invoke.

With the above in mind, we can start analysing the probability of $\mathcal{E}_{\mathsf{ch}}$. We first look at the probability of $\mathcal{E}_{\mathsf{ch},2}$. In a 2-chain (i_1, i_2), the only candidate for a stale node is i_2, and the only way for i_2 to be stale if to have both $W_1^{(i_2)} = W_1^{(i_1)}$ and $W_2^{(i_2)} = W_2^{(i_1)}$, which is impossible because $\mathbf{N}^{(i_1)} \neq \mathbf{N}^{(i_2)}$, and the mapping $\mathbf{N}^{(i)} \mapsto \mathbf{W}^{(i)}$ is injective. Thus,

$$\Pr[\mathcal{E}_{\mathsf{ch},2}] = 0.$$

Next, for $k > 2$, define

$$\mathcal{E}_{\mathsf{ch},<k} := \bigvee_{2 \leq k' < k} \mathcal{E}_{\mathsf{ch},k'}.$$

We want to examine the probability of $\mathcal{E}_{\mathsf{ch},k}$ conditioned on $\neg\mathcal{E}_{\mathsf{ch},<k}$. So suppose $\mathcal{E}_{\mathsf{ch},<k}$ has not occurred. Consider a k-chain $\zeta = (i_1, \ldots, i_k)$ with one or more stale nodes, and let i_d be its first stale node from the left. We observe that if $d < k$, $(i_1, \ldots, i_d)$ is a d-chain with i_d as a stale node, which contradicts the assumption that $\mathcal{E}_{\mathsf{ch},<k}$ (and thus in particular $\mathcal{E}_{\mathsf{ch},d}$) has not occurred. Thus, the only candidate for a stale node in ζ is i_k.

Now, since i_2 is fresh (it cannot be otherwise, as we saw earlier), for any fixed $j_1 \in [0..5]$ we have

$$\Pr\left[X_{j_1}^{(i_2)} = X_{j_1}^{(i_1)}\right] \leq \frac{2^8}{2^n - 1},$$

by the randomness of $R_1^{(i_2)}$ or $R_2^{(i_2)}$, whichever lends freshness to i_2. Since there are 6 choices each for j_1, we have

$$\Pr[\zeta_2] \leq \frac{6 \cdot 2^8}{2^n - 1} \leq \frac{1}{2^{n-11}}. \tag{81}$$

Further, if $k \geq 4$, for any d with $2 < d < k$, since i_d is fresh (by the above discussion), for any fixed $j_{d-1} \in [0..\ell]$ we have

$$\Pr\left[X_{j_{d-1}}^{(i_d)} = X_{j_{d-1}}^{(i_{d-1})} \mid \zeta_{<d}\right] \leq \frac{2^8}{2^n - d + 1},$$

and counting the choices for j_{d-1} gives

$$\Pr\left[\zeta_d \mid \zeta_{<d}\right] \leq \frac{6 \cdot 2^\mu}{2^n - d + 1} \leq \frac{1}{2^{n-11}}. \tag{82}$$

By induction, we can show that

$$\Pr[\zeta_{<d+1}] \leq \frac{1}{(2^{n-11})^{d-1}}. \tag{83}$$

For the proof of this bound, Eqn. (81) serves as base case; for the induction step we use the identity $\Pr[\zeta_{<d+1}] = \Pr[\zeta_d \wedge \zeta_{<d}] = \Pr[\zeta_d \mid \zeta_{<d}] \Pr[\zeta_{<d}]$ and plug in the bound from Eqn. (82).

Taking $d = k - 1$ in Eqn. (83) gives

$$\Pr[\zeta_{<k}] \leq \frac{1}{(2^{n-11})^{k-2}}.$$

Since $\zeta_k \implies \zeta_{<k}$, we have

$$\Pr[\zeta_k \wedge i_k \text{ is stale}] \leq \Pr[\zeta_k] \leq \Pr[\zeta_{<k}] \leq \frac{1}{(2^{n-11})^{k-2}}.$$

Since i_k is stale and $\mathcal{E}_{\mathsf{ch},<k}$ did not occur, i_k must be fresh in the chain $(i_2, \ldots, i_k)$. Thus, either $W_1^{(i_k)} \notin \{W_1^{(i_a)} \mid 2 \le a < k\}$ or $W_2^{(i_k)} \notin \{W_2^{(i_a)} \mid 2 \le a < k\}$. First suppose $W_1^{(i_k)} \notin \{W_1^{(i_a)} \mid 2 \le a < k\}$. Since i_k is stale in ζ, we must have $W_1^{(i_k)} = W_1^{(i_1)}$. This in turn implies that $W_2^{(i_k)} \neq W_2^{(i_1)}$, so again by staleness of i_k, we must have $W_2^{(i_k)} = W_2^{(i_a)}$ for some $a \in [2..k-1]$. Thus, (i_1, i_k, i_a) must be an elbow. Similarly, if instead $W_2^{(i_k)} \notin \{W_2^{(i_a)} \mid 2 \le a < k\}$, we can show that (i_a, i_k, i_1) must be an elbow.

In either case, the number of choices for i_1, i_a, i_k is bounded by $|\mathcal{S}_{\mathsf{el}}|$, which is bounded by q if $\mathcal{E}_{\mathsf{el}}$ has not occurred. Then the other $k-3$ indices in ζ can be picked in $(q-3)_{k-3} \le q^{k-3}$ ways. Thus, by union-bound, we have

$$\Pr[\mathcal{E}_{\mathsf{ch},k} \mid \neg\mathcal{E}_{\mathsf{ch},<k}, \neg\mathcal{E}_{\mathsf{el}}] \le \frac{q(q-3)_{k-3}}{(2^{n-11})^{k-2}} \le \left(\frac{q}{2^{n-11}}\right)^{k-2}. \tag{84}$$

Recall that $\mathcal{E}_{\mathsf{ch}}$ denotes the event $\mathcal{E}_{\mathsf{ch},3} \vee \cdots \vee \mathcal{E}_{\mathsf{ch},n+1}$. Then by union-bound we have

$$\Pr[\mathcal{E}_{\mathsf{ch}} \mid \neg\mathcal{E}_{\mathsf{el}}] \le \sum_{k=3}^{n+1} \left(\frac{q}{2^{n-11}}\right)^{k-2} = \frac{q}{2^{n-11}} \cdot \sum_{k=0}^{n-2} \left(\frac{q}{2^{n-11}}\right)^{k}. \tag{85}$$

Using the assumption that $q \le 2^{n-11}$, we can bound Eqn. (85) as

$$\Pr[\mathcal{E}_{\mathsf{ch}} \mid \neg\mathcal{E}_{\mathsf{el}}] \le \frac{q}{2^{n-11}} \cdot \sum_{k=0}^{n-2} \left(\frac{1}{2}\right)^{k} \le \frac{q}{2^{n-12}}. \tag{86}$$

From Eqns. (18) (Lemma 6) and (86) we get

$$\begin{aligned}\Pr[\mathcal{E}_{\mathsf{ch}}] &= \Pr[\mathcal{E}_{\mathsf{ch}} \wedge \mathcal{E}_{\mathsf{el}}] + \Pr[\mathcal{E}_{\mathsf{ch}} \wedge \neg\mathcal{E}_{\mathsf{el}}] \\ &\le \Pr[\mathcal{E}_{\mathsf{el}}] + \Pr[\mathcal{E}_{\mathsf{ch}} \mid \neg\mathcal{E}_{\mathsf{el}}] \le \frac{q}{2^{n-8}} + \frac{q^2}{2^{2n-16}} + \frac{q}{2^{n-12}} \le \frac{q}{2^{n-13}}\end{aligned}$$

(again by $q \le 2^{n-11}$), which establishes the bound in Eqn. (20). □

C.5 Proof of Lemma 8

This proof closely follows the proof of Lemma 7 in Appendix C.4. A tree ϑ can be treated interchangeably with the event of the tree forming, i.e., the conjunction of all the collisions necessary for ϑ to satisfy the definition of a tree, thus letting us make sense of an expression like $\Pr[\vartheta]$. For $2 < d \le k$ and a k-chain ϑ, let $\vartheta_{<d}$ denote the $(d-1)$-tree $(i_1, \ldots, i_{d-1})$ (or, equivalently, the conjunction of the $d-2$ collisions required for this tree), and let ϑ_d denote the collision required to link $i^\star_{d-1}$ to i_d, i.e., the event that for some $j_{d-1} \in [0..5]$, $X^{(i_d)}_{j_{d-1}} = X^{(i^\star_{d-1})}_{j_{d-1}}$. Note that by this notation,

$$\vartheta = \bigwedge_{d=2}^{k} \vartheta_d.$$

We observe that if $W_1^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \le a < d\}$, $R_1^{(i_d)}$ will be the truncated output of $\pi_1(\cdot\|\mathtt{0x01})$ on an input that did not appear earlier in the tree, and similarly, if $W_2^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \le a < d\}$, $R_2^{(i_d)}$ will be the truncated output of $\pi_1(\cdot\|\mathtt{0x02})$ on an input that did not appear earlier in the tree. Thus, if i_d is fresh, at least one of $R_1^{(i_d)}$ and $R_2^{(i_d)}$ will be a source of fresh randomness after conditioning on $\vartheta_{<d}$.

Now we analyse the probability of $\mathcal{E}_{\mathsf{t}}$. We first observe that a 2-tree is also a 2-chain, so

$$\Pr[\mathcal{E}_{\mathsf{t},2}] = \Pr[\mathcal{E}_{\mathsf{ch},2}] = 0.$$

Next, for $k > 2$, define

$$\mathcal{E}_{\mathsf{t},<k} := \bigvee_{2 \le k' < k} \mathcal{E}_{\mathsf{t},k'}.$$

We want to examine the probability of $\mathcal{E}_{\mathsf{t},k}$ conditioned on $\neg\mathcal{E}_{\mathsf{t},<k}$. So suppose $\mathcal{E}_{\mathsf{t},<k}$ has not occurred. Consider a k-tree $\vartheta = (i_1, \ldots, i_k)$ with one or more stale nodes, and let i_d be its first stale node from the left. We observe that if $d < k$, $(i_1, \ldots, i_d)$ is a d-tree with i_d as a stale node, which contradicts the assumption that $\mathcal{E}_{\mathsf{t},<k}$ (and thus in particular $\mathcal{E}_{\mathsf{t},d}$) has not occurred. Thus, the only candidate for a stale node in ϑ is i_k.

Now, since i_2 is fresh (it cannot be otherwise, as we saw earlier), for any fixed $j_1 \in [0..5]$ we have

$$\Pr\left[X_{j_1}^{(i_2)} = X_{j_1}^{(i_1^\star)}\right] \le \frac{2^8}{2^n - 1},$$

by the randomness of $R_1^{(i_2)}$ or $R_2^{(i_2)}$, whichever lends freshness to i_2. Since there are 6 choices each for j_1, we have

$$\Pr[\vartheta_2] \le \frac{6 \cdot 2^8}{2^n - 1} \le \frac{1}{2^{n-11}}. \tag{87}$$

Further, if $k \ge 4$, for any d with $2 < d < k$, since i_d is fresh (by the above discussion), for any fixed $j_{d-1} \in [0..\ell]$ we have

$$\Pr\left[X_{j_{d-1}}^{(i_d)} = X_{j_{d-1}}^{(i_{d-1}^\star)} \mid \vartheta_{<d}\right] \le \frac{2^8}{2^n - d + 1},$$

and counting the choices for $i_{d-1}^\star$ and j_{d-1} gives

$$\Pr\left[\vartheta_d \mid \vartheta_{<d}\right] \le \frac{6(d-1)2^8}{2^n - d + 1} \le \frac{d-1}{2^{n-11}}. \tag{88}$$

By induction, we can show that

$$\Pr[\vartheta_{<d+1}] \le \frac{(d-1)!}{(2^{n-11})^{d-1}}. \tag{89}$$

For the proof of this bound, Eqn. (87) serves as base case; for the induction step we use the identity $\Pr[\vartheta_{<d+1}] = \Pr[\vartheta_d \wedge \vartheta_{<d}] = \Pr[\vartheta_d \mid \vartheta_{<d}] \Pr[\vartheta_{<d}]$ and plug in the bound from Eqn. (88).

Taking $d = k - 1$ in Eqn. (89) gives

$$\Pr[\vartheta_{<k}] \le \frac{(k-2)!}{(2^{n-11})^{k-2}}.$$

Since $\vartheta_k \implies \vartheta_{<k}$, we have

$$\Pr[\vartheta_k \wedge i_k \text{ is stale}] \le \Pr[\vartheta_k] \le \Pr[\vartheta_{<k}] \le \frac{(k-2)!}{(2^{n-11})^{k-2}}.$$

Since i_k is stale and $\mathcal{E}_{\mathsf{t},<k}$ did not occur, i_k must be fresh in the tree $(i_2, \ldots, i_k)$. Thus, either $W_1^{(i_k)} \notin \{W_1^{(i_a)} \mid 2 \le a < k\}$ or $W_2^{(i_k)} \notin \{W_2^{(i_a)} \mid 2 \le a < k\}$. First suppose $W_1^{(i_k)} \notin \{W_1^{(i_a)} \mid 2 \le a < k\}$. Since i_k is stale in ϑ, we must have $W_1^{(i_k)} = W_1^{(i_1)}$. This in turn implies that $W_2^{(i_k)} \ne W_2^{(i_1)}$, so again by staleness of i_k, we must have $W_2^{(i_k)} = W_2^{(i_a)}$ for some $a \in [2..k-1]$. Thus, (i_1, i_k, i_a) must be an elbow. Similarly, if instead $W_2^{(i_k)} \notin \{W_2^{(i_a)} \mid 2 \le a < k\}$, we can show that (i_a, i_k, i_1) must be an elbow.

In either case, the number of choices for i_1, i_a, i_k is bounded by $|\mathcal{S}_{\mathsf{el}}|$, which is bounded by q if $\mathcal{E}_{\mathsf{el}}$ has not occurred. Then the other $k-3$ indices in ϑ can be picked in $(q-3)_{k-3}$ ways. Thus, by union-bound, we have

$$\Pr[\mathcal{E}_{\mathsf{t},k} \mid \neg\mathcal{E}_{\mathsf{t},<k}, \neg\mathcal{E}_{\mathsf{el}}] \leq \frac{(k-2)!q(q-3)_{k-3}}{(2^{n-11})^{k-2}} \leq \left(\frac{(k-2)q}{2^{n-11}}\right)^{k-2}. \tag{90}$$

Recall that $\mathcal{E}_{\mathsf{t}}$ denotes the event $\mathcal{E}_{\mathsf{t},3} \vee \cdots \vee \mathcal{E}_{\mathsf{t},n+1}$. Then by union-bound we have

$$\Pr[\mathcal{E}_{\mathsf{t}} \mid \neg\mathcal{E}_{\mathsf{el}}] \leq \sum_{k=3}^{n+1} \left(\frac{(k-2)q}{2^{n-11}}\right)^{k-2} \leq \frac{q}{2^{n-11}} \cdot \sum_{k=0}^{n-2} \left(\frac{nq}{2^{n-11}}\right)^{k}. \tag{91}$$

Using the assumption that $nq \leq 2^{n-11}$, we can bound Eqn. (91) as

$$\Pr[\mathcal{E}_{\mathsf{t}} \mid \neg\mathcal{E}_{\mathsf{el}}] \leq \frac{q}{2^{n-11}} \cdot \sum_{k=0}^{n-2} \left(\frac{1}{2}\right)^{k} \leq \frac{q}{2^{n-12}}. \tag{92}$$

From Eqns. (18) (Lemma 6) and (92) we get

$$\begin{aligned}\Pr[\mathcal{E}_{\mathsf{t}}] &= \Pr[\mathcal{E}_{\mathsf{t}} \wedge \mathcal{E}_{\mathsf{el}}] + \Pr[\mathcal{E}_{\mathsf{t}} \wedge \neg\mathcal{E}_{\mathsf{el}}] \\ &\leq \Pr[\mathcal{E}_{\mathsf{el}}] + \Pr[\mathcal{E}_{\mathsf{t}} \mid \neg\mathcal{E}_{\mathsf{el}}] \leq \frac{q}{2^{n-8}} + \frac{q^2}{2^{2n-16}} + \frac{q}{2^{n-12}} \leq \frac{q}{2^{n-13}},\end{aligned}$$

which establishes the bound in Eqn. (22). □

C.6 Proof of Lemma 9

We observe that for any $i \in [1..q], j \in [0..5]$,

$$X_j^{(i)} = (R_1^{(i)} \oplus 2^j \cdot R_2^{(i)} \oplus \mathcal{I}_{[j\geq 4]} W_{j-3}^{(i)}) \| \langle j \rangle_8,$$

where $\mathcal{I}$ denotes the indicator function. Then we can rewrite Eqn. (23) as the two equations

$$R_1^{(i_2)} \oplus 2^{j_1} \cdot R_2^{(i_2)} = R_1^{(i_1)} \oplus 2^{j_1} \cdot R_2^{(i_1)} \oplus \mathcal{I}_{[j_1\geq 4]}(W_{j_1-3}^{(i_1)} \oplus W_{j_1-3}^{(i_2)}), \tag{93}$$

$$R_1^{(i_2)} \oplus 2^{j_2} \cdot R_2^{(i_2)} = R_1^{(i_1)} \oplus 2^{j_2} \cdot R_2^{(i_1)} \oplus \mathcal{I}_{[j_2\geq 4]}(W_{j_2-3}^{(i_1)} \oplus W_{j_2-3}^{(i_2)}). \tag{94}$$

If $(i_1, i_2) \notin \mathcal{S}_{\mathsf{ar}}$, $W_1^{(i_1)} \neq W_1^{(i_2)}$ and $W_2^{(i_1)} \neq W_2^{(i_2)}$, so $R_1^{(i_1)}$, $R_1^{(i_2)}$, $R_2^{(i_1)}$, and $R_2^{(i_2)}$ are truncated outputs of π_1 on different inputs. Since $j_1 \neq j_2$, Eqns. 93 and 94 are independent, so we can write

$$\Pr\left[\textit{Eqns. } (93),\ (94)\right] = \frac{2^{16}}{(2^n-2)(2^n-3)}.$$

If $(i_1, i_2) \in \mathcal{S}_{\mathsf{ar}}$, Eqns. 93 and 94 may become dependent. But one of $W_1^{(i_1)} \neq W_1^{(i_2)}$ and $W_2^{(i_1)} \neq W_2^{(i_2)}$ must still hold, so we must still have

$$\Pr\left[\textit{Eqns. } (93),\ (94)\right] \leq \frac{2^8}{2^n-1}. \tag{95}$$

There will be $\binom{6}{2} = 15$ choices for j_1 and j_2. Let $\mathcal{E}_2^{i_1,i_2}$ denote the event that for some $j_1, j_2 \in [0..5]$ with $j_1 < j_2$, $(X_{j_1}^{(i_2)}, X_{j_2}^{(i_2)}) = (X_{j_1}^{(i_1)}, X_{j_2}^{(i_1)})$. Then we have

$$\Pr[\mathcal{E}_2^{i_1,i_2} \mid \neg\mathcal{E}_{\mathsf{ar}}] \le \begin{cases} \dfrac{15 \cdot 2^{16}}{(2^n-2)(2^n-3)} & \le \dfrac{1}{2^{2n-20}}, \quad \text{if } (i_1,i_2) \in \mathcal{S}_{\mathsf{ar}}, \\ \dfrac{15 \cdot 2^8}{2^n-1} & \le \dfrac{1}{2^{n-12}}, \quad \text{otherwise.} \end{cases}$$

Recalling that $|\mathcal{S}_{\mathsf{ar}}| \le q$ as long as $\mathcal{E}_{\mathsf{ar}}$ has not occurred, we get

$$\begin{aligned} \Pr[\mathcal{E}_2 \mid \neg\mathcal{E}_{\mathsf{ar}}] &= \sum_{(i_1,i_2)\in[1..q]^2} \Pr[\mathcal{E}_2^{i_1,i_2} \mid \neg\mathcal{E}_{\mathsf{ar}}] \\ &= \sum_{(i_1,i_2)\in\mathcal{S}_{\mathsf{ar}}} \Pr[\mathcal{E}_2^{i_1,i_2} \mid \neg\mathcal{E}_{\mathsf{ar}}] + \sum_{(i_1,i_2)\notin\mathcal{S}_{\mathsf{ar}}} \Pr[\mathcal{E}_2^{i_1,i_2} \mid \neg\mathcal{E}_{\mathsf{ar}}] \\ &\le \frac{|\mathcal{S}_{\mathsf{ar}}|}{2^{n-12}} + \frac{q(q-1)-|\mathcal{S}_{\mathsf{ar}}|}{2^{2n-20}} \le \frac{q}{2^{n-12}} + \frac{q^2}{2^{2n-20}} \le \frac{q}{2^{n-13}} \end{aligned}$$

(by the assumption $q \le 2^{n-11}$), which establishes the bound in Eqn. (24). □

C.7 Proof of Lemma 10

Eqn. (25) is equivalent to an $(n+1)$-chain ζ. If $\mathcal{E}_{\mathsf{ch}}$ has not happened, each node in this chain must be fresh, so arguing as in the bounding of $\Pr[\zeta_{<d+1}]$ in Appendix C.4, we have

$$\Pr[\zeta \mid \neg\mathcal{E}_{\mathsf{ch}}] \le \frac{1}{(2^{n-11})^n}. \tag{96}$$

Since there are $(q)_{n+1}$ choices for the indices in ζ, by union-bound we have

$$\Pr[\mathcal{E}_3 \mid \neg\mathcal{E}_{\mathsf{ch}}] \le \frac{(q)_{n+1}}{(2^{n-11})^n} \le q\left(\frac{q}{2^{n-11}}\right)^n \le \frac{q}{2^n},$$

using the assumption that $q \le 2^{n-11}$. This establishes the bound in Eqn. (26). □

C.8 Proof of Lemma 11

Eqn. (27) is equivalent to $(i_1, \ldots, i_{k-1})$ forming a $(k-1)$-chain $\zeta_{<k}$, and i_k satisfying the condition

$$(X_{j_{k-1}}^{(i_k)}, X_{j_k}^{(i_k)}) = (X_{j_{k-1}}^{(i_{k-1})}, X_{j_k}^{(i_1)}). \tag{97}$$

As in Supplementary Material C.6, we can rewrite Eqn. (97) as

$$R_1^{(i_k)} \oplus 2^{j_{k-1}} \cdot R_2^{(i_k)} = R_1^{(i_{k-1})} \oplus 2^{j_{k-1}} \cdot R_2^{(i_{k-1})} \oplus \mathcal{I}_{[j_{k-1}\ge 4]}(W_{j_{k-1}-3}^{(i_{k-1})} \oplus W_{j_{k-1}-3}^{(i_k)}), \tag{98}$$

$$R_1^{(i_k)} \oplus 2^{j_k} \cdot R_2^{(i_k)} = R_1^{(i_1)} \oplus 2^{j_k} \cdot R_2^{(i_1)} \oplus \mathcal{I}_{[j_k\ge 4]}(W_{j_k-3}^{(i_1)} \oplus W_{j_k-3}^{(i_k)}). \tag{99}$$

Since $\zeta_k := (i_1, \ldots, i_k)$ is also a chain, i_k will be fresh in ζ_k unless $\mathcal{E}_{\mathsf{ch}}$ has occurred. Now we consider two cases. First, if both $W_1^{(i_d)} \notin \{W_1^{(i_a)} \mid 1 \le a < d\}$ and $W_2^{(i_d)} \notin \{W_2^{(i_a)} \mid 1 \le a < d\}$, then we can use the randomness of both $R_1^{(i_k)}$ and $R_2^{(i_k)}$ after conditioning on $\zeta_{<k}$, to get

$$\Pr\left[\textit{Eqns. (98), (99)} \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \le \frac{2^{16}}{(2^n-2k+4)(2^n-2k+3)}. \tag{100}$$

Like in Eqn. (96), taking $k-2$ instead of n, we get

$$\Pr[\zeta_{<k} \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}] \leq \frac{1}{(2^{n-11})^{k-2}}. \tag{101}$$

Let $\zeta_k^{\circlearrowleft}$ denote the *cyclic closure* of ζ_k, i.e., the chain $\zeta_{<k}$ together with the event that we can find j_{k-1}, $j_k \in [0..\ell]$ such that Eqn. (97) holds. Then we have

$$\begin{aligned}
&\Pr\left[\zeta_k^{\circlearrowleft} \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \\
&= \Pr\left[\bigvee_{j_{k-1},j_k} (X_{j_{k-1}}^{(i_k)}, X_{j_k}^{(i_k)}) = (X_{j_{k-1}}^{(i_{k-1})}, X_{j_k}^{(i_1)}) \;\middle|\; \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \\
&\leq \sum_{j_{k-1},j_k} \Pr\left[(X_{j_{k-1}}^{(i_k)}, X_{j_k}^{(i_k)}) = (X_{j_{k-1}}^{(i_{k-1})}, X_{j_k}^{(i_1)}) \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \\
&\leq \frac{36 \cdot 2^{16}}{(2^n - 2k + 4)(2^n - 2k + 3)} \leq \frac{1}{2^{2n-22}},
\end{aligned} \tag{102}$$

by applying union-bound to Eqn. (100). Recalling that $\zeta^{\circlearrowleft} \wedge \zeta_{<k} = \zeta^{\circlearrowleft}$, from Eqns. (101) and (102) we get

$$\begin{aligned}
\Pr\left[\zeta^{\circlearrowleft} \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] &\leq \Pr\left[\zeta^{\circlearrowleft} \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \Pr\left[\zeta_{<k} \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \\
&\leq \frac{1}{2^{2n-22+(n-11)(k-2)}} = \frac{1}{(2^{n-11})^k}.
\end{aligned} \tag{103}$$

The other case is when for some $d \in [1..k-1]$ either $W_1^{(i_d)} \in \{W_1^{(i_a)} \mid 1 \leq a < d\}$ or $W_2^{(i_d)} \in \{W_2^{(i_a)} \mid 1 \leq a < d\}$ holds. Then $(i_d, i_k) \in \mathcal{S}_{\mathsf{ar}}$, and we can argue as before Eqn. (95) that

$$\Pr\left[\textit{Eqns. (98), (99)} \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \leq \frac{2^8}{2^n - 2k + 2},$$

so that

$$\Pr\left[\zeta^{\circlearrowleft} \mid \zeta_{<k}, \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \leq \frac{36 \cdot 2^8}{2^n - 2k + 2} \leq \frac{1}{2^{n-14}},$$

and hence

$$\Pr\left[\zeta^{\circlearrowleft} \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}\right] \leq \frac{1}{2^{(n-11)(k-1)-3}}. \tag{104}$$

For this case, since $\mathcal{E}_{\mathsf{ar}}$ did not occur, (i_d, i_k) can be chosen in at most q ways, and the rest of the queries in at most $(q-2)_{k-2}$ ways. Thus, combining the two cases from Eqns. (103) and (104), we have

$$\Pr[\mathcal{E}_{4,k} \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}] \leq \frac{(q)_k}{2^{(n-11)k}} + \frac{q(q-2)_{k-2}}{2^{(n-11)(k-1)-3}} \leq 8\left(\frac{q}{2^{n-11}}\right)^{k-1}, \tag{105}$$

where in the last inequality we use the assumption that $q \leq 2^{n-11}$. Recall that $\mathcal{E}_4$ denotes the event $\mathcal{E}_{4,3} \vee \cdots \vee \mathcal{E}_{4,n}$. Then by Eqn. (105) and union-bound we have

$$\begin{aligned}
\Pr[\mathcal{E}_4 \mid \neg\mathcal{E}_{\mathsf{ch}}, \neg\mathcal{E}_{\mathsf{ar}}] &\leq \sum_{k=3}^{n} \left[8\left(\frac{q}{2^{n-11}}\right)^{k-1}\right] \\
&= 8\left(\frac{q}{2^{n-11}}\right)^2 \cdot \sum_{k=0}^{n-3} \left(\frac{q}{2^{n-11}}\right)^k \\
&\leq 8\left(\frac{q}{2^{n-11}}\right)^2 \cdot \sum_{k=0}^{n-3} \left(\frac{1}{2}\right)^k \leq \frac{q^2}{2^{2n-26}},
\end{aligned}$$

which establishes the bound in Eqn. (28). □

C.9 Proof of Lemma 12

Consider the graph $\hat{\mathsf{Gr}}$, whose vertices are $i_1, \ldots, i_{n+1}$, and which contains a path between i_k and $i_{k'}$ if and only if $X_{j_k}^{(i_k)} = X_{j_k}^{(i_{k'})}$ for some $j_k \in [0..5]$. If Eqn. (29) holds and none of $\mathcal{E}_2$, $\mathcal{E}_3$, or $\mathcal{E}_4$ has happened, we claim that for $k' < k$ there can only be an edge between i_k and $i_{k'}$ if $i_{k'} = i^\star_{k-1}$. Since $\mathcal{E}_3$ has not happened, $\hat{\mathsf{Gr}}$ cannot contain a path of length $n+1$, and hence it cannot be a cycle of length $n+1$; and since neither $\mathcal{E}_2$ nor $\mathcal{E}_4$ has happened, $\hat{\mathsf{Gr}}$ cannot contain a cycle of length n or less; thus $\hat{\mathsf{Gr}}$ must by acyclic, and can have at most n edges. Eqn. (29) says that for each $k \in [1..n]$ there is an edge between $i^\star_k$ and i_{k+1}, and these edges are all distinct, so they must be the only n edges in $\hat{\mathsf{Gr}}$. Thus for the rest of the proof we assume $\hat{\mathsf{Gr}}$ is a tree, i.e., there are no other collisions over the queries $i_1, \ldots, i_{n+1}$ except those mentioned in Eqn. (29). Now, Eqn. (29) is equivalent to an $(n+1)$-tree ϑ. If $\mathcal{E}_{\mathsf{t}}$ has not happened, each node in this tree must be fresh, so arguing as in the bounding of $\Pr[\vartheta_{<d+1}]$ in Appendix C.5, we have

$$\Pr[\vartheta \mid \neg\mathcal{E}_{\mathsf{t}}, \neg\mathcal{E}_2, \neg\mathcal{E}_3, \neg\mathcal{E}_4] \le \frac{n!}{(2^{n-11})^n}. \tag{106}$$

Since there are $(q)_{n+1}$ choices for the indices in ϑ, by union-bound we have

$$\Pr[\mathcal{E}_5 \mid \neg\mathcal{E}_{\mathsf{t}}, \neg\mathcal{E}_2, \neg\mathcal{E}_3, \neg\mathcal{E}_4] \le \frac{n!(q)_{n+1}}{(2^{n-11})^n} \le q\left(\frac{nq}{2^{n-11}}\right)^n \le \frac{q}{2^n},$$

using the assumption that $nq \le 2^{n-11}$. This establishes the bound in Eqn. (30). □

C.10 Proof of Lemma 13

We observe that $\mathsf{Trs}\,(\mathcal{A}^{\mathcal{O}_0})$ cannot have redundancy except by triggering $\mathcal{E}_2$, or $\mathcal{E}_4$, and cannot fail to be $5n$-scattered except by triggering $\mathcal{E}_5$. Thus,

$$\mathcal{E}_{\mathsf{bad},1} \subseteq \mathcal{E}_2 \vee \mathcal{E}_4 \vee \mathcal{E}_5. \tag{107}$$

Since $\mathcal{E}_{\mathsf{ch}} \subseteq \mathcal{E}_{\mathsf{t}}$, we have $\neg\mathcal{E}_{\mathsf{t}} \subseteq \neg\mathcal{E}_{\mathsf{ch}}$. Define $\mathcal{E}_{\mathsf{aux}} := \mathcal{E}_{\mathsf{ar}} \vee \mathcal{E}_{\mathsf{t}}$. Then we can write

$$\begin{aligned}
\Pr[\mathcal{E}_{\mathsf{bad},1}] &= \Pr[\mathcal{E}_{\mathsf{bad},1} \wedge \mathcal{E}_{\mathsf{aux}}] + \Pr[\mathcal{E}_{\mathsf{bad},1} \wedge \neg\mathcal{E}_{\mathsf{aux}}] \\
&\le \Pr[\mathcal{E}_{\mathsf{aux}}] + \sum_{i\in\{2,4,5\}} \Pr[\mathcal{E}_i \wedge \neg\mathcal{E}_{\mathsf{aux}}] \\
&\le \Pr[\mathcal{E}_{\mathsf{ar}}] + \Pr[\mathcal{E}_{\mathsf{t}}] + \Pr[\mathcal{E}_2 \wedge \neg\mathcal{E}_{\mathsf{ar}}] \\
&\qquad + \Pr[\mathcal{E}_4 \wedge \neg\mathcal{E}_{\mathsf{ar}} \wedge \neg\mathcal{E}_{\mathsf{ch}}] + \Pr[\mathcal{E}_5 \wedge \neg\mathcal{E}_{\mathsf{t}}] \\
&\le \Pr[\mathcal{E}_{\mathsf{ar}}] + \Pr[\mathcal{E}_{\mathsf{t}}] + \Pr[\mathcal{E}_2 \mid \neg\mathcal{E}_{\mathsf{ar}}] + \Pr[\mathcal{E}_3 \mid \neg\mathcal{E}_{\mathsf{ch}}] \\
&\qquad + \Pr[\mathcal{E}_4 \mid \neg\mathcal{E}_{\mathsf{ar}}, \neg\mathcal{E}_{\mathsf{ch}}] + \Pr[\mathcal{E}_5 \mid \neg\mathcal{E}_{\mathsf{t}}, \neg\mathcal{E}_2, \neg\mathcal{E}_3, \neg\mathcal{E}_4].
\end{aligned} \tag{108}$$

Substituting the bounds from Eqns. (17), (22), (24), (26), (28), and (30) (Lemmas 5, 8, 9, 10, 11, and 12) in Eqn. (108) gives

$$\Pr[\mathcal{E}_{\mathsf{bad},1}] \le \frac{q}{2^{n-9}} + \frac{q}{2^{n-13}} + \frac{q}{2^{n-13}} + \frac{q}{2^n} + \frac{q^2}{2^{2n-26}} + \frac{q}{2^n} \le \frac{q}{2^{n-15}}, \tag{109}$$

using the assumption $nq \le 2^{n-11}$. □

D Deferred Proofs from Sec. 5

D.1 Proof of Theorem 7

We prove this as a special case of Theorem 6 when $(N_1, N_2) = (N_1', N_2')$. Let us fix a key K, an input pair (N_1, N_2) and an output pair (Z_4, Z_5). Let $\mathcal{E}_{(K,N_1,N_2,Z_4,Z_5)}$ be the event that given (K, N_1, N_2, Z_4, Z_5), there exists a computable tuple $\bar{x} \in T^K$ that matches both the input (N_1, N_2) and output (Z_4, Z_5). Informally, $\bar{x}$ can be seen as a collection of primitive queries that map (N_1, N_2) to (Z_4, Z_5). It is easy to see that $\Pr[\mathcal{E}_{(K,N_1,N_2,Z_4,Z_5)}] = 1/2^{2n}$. Hence, when $(N_1, N_2) = (N_1', N_2')$,

$$\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}_{(K',Z_4,Z_5)}]$$
$$\leq \sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}_{(K,Z_4,Z_5)}] \Pr[\mathcal{E}_{(K',N_1,N_2,Z_4,Z_5)}] \leq \frac{1}{2^{2n}} \sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}_{(K,Z_4,Z_5)}]. \tag{110}$$

So, it is enough to bound $\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}^i_{(K,Z_4,Z_5)}]$ for $i = 1, 2, 3$. Again using the fact that when $\bar{x}$ is Type 1, neither Z_4 nor Z_5 is fixed, when $\bar{x}$ is Type 2, both Z_4 and Z_5 are fixed, and $\bar{x}$ is Type 3, one of Z_4, Z_5 or $Z_4 \oplus Z_5$ is fixed (say Z_4), we have the following:

$$\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}] \leq 2^{2n} \Pr[\mathcal{E}^1_{(K,Z_4,Z_5)}] \leq \frac{2^{25} q_K^4}{2^{2n}},$$
$$\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}^2_{(K,Z_4,Z_5)}] \leq \sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \frac{2^{21} q^2_{(K,Z_4,Z_5)}}{2^{2n}} \leq \frac{2^{21} q_K^2}{2^{2n}},$$
$$\sum_{(Z_4,Z_5)\in\{0,1\}^{2n}} \Pr[\mathcal{E}^3_{(K,Z_4,Z_5)}] \leq 2^n \sum_{Z_4\in\{0,1\}^n} \Pr[\mathcal{E}^3_{(K,Z_4,Z_5)}] \leq \frac{2^{13} q_K^2}{2^n}.$$

Combining the above bounds in equation (110) yields the desired result. □

D.2 Proof of Lemma 14

We bound the probability of only $\mathcal{E}_{\mathsf{bad},1}$ and $\mathcal{E}_{\mathsf{bad},2}$. The bound of $\mathcal{E}_{\mathsf{bad},3}$ follows from the bound of $\mathcal{E}_{\mathsf{bad},1}$, and the bound for $\mathcal{E}_{\mathsf{bad},4}$ and $\mathcal{E}_{\mathsf{bad},5}$ follows from the bound of $\mathcal{E}_{\mathsf{bad},2}$.

Bounding $\mathcal{E}_{\mathsf{bad},1}$. Let $\beta = W_1 \oplus R_1$. Then equation (48) becomes

$$\beta \oplus 2^4 R_2 = \tilde{X}_4.$$

β is uniformly random, irrespective of the direction in which x_2 is queried, since it is the XOR of the input and output of a random permutation. Let $\mathcal{S} := \{W_1 \oplus R_1 \mid x_2 \in T_2^K\}$, $\mathcal{X} := \{\tilde{X}_4 \mid x_5 \in T_5^{K,+}\}$, and $\mathcal{Y} = \{2^4 \cdot R_2 \mid x_3 \in T_3^K\}$. Then, by Lemma 16, we have $\Pr[\mathcal{E}_{\mathsf{bad},1}] \leq 2/2^{n-8}$.

Bounding $\mathcal{E}_{\mathsf{bad},2}$. Here, the sizes of the sets corresponding to $W_1 \oplus R_1$ and $2^4 \cdot R_2$ are q_K each, but the size of the set corresponding to $\tilde{X}_4$ is $q_{(K,Z_4,Z_5)}$. However, since $\tilde{X}_4$ is random, we can assign $\mathcal{S} := \{\tilde{X}_4 \mid x_5 \in T_5^{K,-}\}$, $\mathcal{X} := \{W_1 \oplus R_1 \mid x_2 \in T_2^K\}$, and $\mathcal{Y} = \{2^4 \cdot R_2 \mid x_3 \in T_3^K\}$, and by applying Lemma 16, we have $\Pr[\mathcal{E}_{\mathsf{bad},2}] \leq 2/2^{n-8}$. □

D.3 Proof of Lemma 15

Bounding $\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}]$. There are two ways in which a computable tuple $\bar{x}$ can be Type 1:

- For $i = 4, 5, 6$, $x_i \in T_i^{K,+}$, i.e. all three queries are forward.
- $x_i \in T_i^{K,-}$ for exactly one $i \in \{4, 5, 6\}$, and for all $j \in \{4, 5, 6\} \setminus \{i\}$, x_j is queried after x_i, i.e., exactly one query is backward, and the backward query precedes the two forward queries.

First, let us consider the case when all three of x_4, x_5 and x_6 are forward. For any such fixed tuple, since all of Y_0, Y_4 and Y_5 are random, the probability that equations (51) and (52) are satisfied is bounded by $1/2^{2n}$. Hence, it is enough to bound the number of tuples that satisfy equations (47)-(50).

We first consider the case when x_4 is queried before both x_5 and x_6. This is because by choosing R_1 and R_2 first, the adversary can choose $\tilde{X}_0$ such that all q_K^2 pairs of (R_1, R_2) satisfy equation (47). Then, considering equation (48), $\alpha := R_1 \oplus 2^4 \cdot R_2 \oplus W_1$ is random, since one of R_1 and W_1 is always random. Let $\mathcal{L}_1$ be the list of all such α with $|\mathcal{L}_1| = q_K^2$. By Lemma 17, an element α occurs in the list at most $\mathsf{mcoll}(q_K^2, 2^{n-8})$ times. Hence, an $\tilde{X}_4$ can match with at most $\mathsf{mcoll}(q_K^2, 2^{n-8})$ such α, and hence there can be at most $q_K \cdot \mathsf{mcoll}(q_K^2, 2^{n-8})$ tuples of (x_2, x_3, x_4, x_5) satisfying equations (47) and (48). We are interested in the case when $q_K^2 > 2^{n-8}$. Hence, $\mathsf{mcoll}(q_K^2, 2^{n-8}) \leq 5q_K^2/2^{n-8}$, and therefore the number of (x_2, x_3, x_4, x_5) tuples is at most $5q_K^3/2^{n-8}$.

Now, considering equation (49), $\alpha' := R_1 \oplus 2^5 \cdot R_2 \oplus W_2$ is random, since one of R_2 and W_2 is always random. Let $\mathcal{L}_2$ be the list of all such α' with $|\mathcal{L}_2| = q_K^3/2^{n-8}$. By Lemma 17, an element α' occurs in the list at most $\mathsf{mcoll}(q_K^3/2^{n-8}, 2^{n-8})$ times. Hence, an $\tilde{X}_5$ can match with at most $\mathsf{mcoll}(5q_K^3/2^{n-8}, 2^{n-8})$ such α', and hence there can be at most $q_K \cdot \mathsf{mcoll}(5q_K^3/2^{n-8}, 2^{n-8})$ tuples of $(x_2, x_3, x_4, x_5, x_6)$ satisfying equations (47), (48) and (49). Since the adversary can choose J after it fixes W_1 and W_2, we assume that the adversary can always force equation (50) to be true. (In fact, this will be true for all the cases, and therefore, we will not consider equation (50) in our analysis hereafter). Therefore in this case,

$$\Pr[\textit{Eqns. } (47)\text{-}(52)] \leq \frac{q_K \cdot \mathsf{mcoll}(5q_K^3/2^{n-8}, 2^{n-8})}{2^{2n}}.$$

As we are interested in the case when $q_K^3 > 2^{2(n-8)}$, we can write

$$\Pr[\textit{Eqns. } (47)\text{-}(52)] \leq \frac{25 \cdot 2^{16} \cdot q_K^4}{2^{4n}} \leq \frac{2^{21} q_K^4}{2^{4n}}.$$

When x_5 or x_6 is queried first instead (say x_5), if $\mathcal{E}_{\mathsf{bad}}$ does not occur, there can be at most $q_K^3/2^{n-8} + q_K\sqrt{3nq_K}$ tuples of (x_2, x_3, x_5) satisfying equation (48) instead of the q_K^2 tuples of (x_2, x_3, x_4) we were getting when x_4 is queried first. Thus, $\Pr[\textit{Eqns. } (47)\text{-}(52)]$ becomes worse.

Next, let us consider the case when $x_4 \in T_4^{K,-}$ and x_4 is queried before both x_5 and x_6. Since x_4 is backward here, $\tilde{X}_0$ is random. If $\mathcal{E}_{\mathsf{bad}}$ does not occur, there can be at most $q_K^3/2^{n-8} + q_K\sqrt{3nq_K}$ tuples of (x_2, x_3, x_4) satisfying equation (47) instead of the q_K^2 tuples of (x_2, x_3, x_4) we were getting when x_4 is queried in forward direction. Since x_5 and x_6 are forward queries, the remaining analysis remains the same, and thus $\Pr[\textit{Eqns. } (47)\text{-}(52)]$ becomes worse.

The analysis is similar when x_5 or x_6 is the backward query. Hence,

$$\Pr[\textit{Eqns. } (47)\text{-}(52)] \leq \frac{2^{21} q_K^4}{2^{4n}}$$

in every case. Summing over the 12 possible execution orders of x_4, x_5, x_6, and over the choice of which query is the backward one, we have

$$\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \leq \frac{12 \cdot 2^{21} q_K^4}{2^{4n}} \leq \frac{2^{25} q_K^4}{2^{4n}}.$$

Bounding $\Pr[\mathcal{E}^2_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}]$**.** There are two ways in which a computable tuple $\bar{x}$ can be Type 2:

- For $i = 4, 5, 6$, $x_i \in T_i^{K,-}$, i.e. all three queries are backward.
- $x_i \in T_i^{K,+}$ for exactly one $i \in \{4, 5, 6\}$, and for all $j \in \{4, 5, 6\} \setminus \{i\}$, x_i is queried before x_j i.e., exactly one query is forward, and the forward query precedes the backward queries.

First, let us consider the case when all the three queries are backward. Then, all of $\tilde{X}_0, \tilde{X}_4$ and $\tilde{X}_5$ are generated randomly. Since $\tilde{X}_4$ is generated randomly, given $\mathcal{E}_{\mathsf{bad}}$ does not occur, there can be at most $q' := q_{(K,Z_4,Z_5)}q_K^2/2^{n-8} + q_{(K,Z_4,Z_5)}\sqrt{3nq_K}$ tuples of (x_2, x_3, x_5) satisfying equation (48). Now, since both x_4 and x_6 are inverse queries, then $Y_0 = Y_4 \oplus Z_4$ and $Y_5 = Y_4 \oplus Z_4 \oplus Z_5$ are fixed, and for the corresponding fixed values of $\tilde{X}_0$ and $\tilde{X}_5$, the probability that these satisfy equations (47) and (49) is $q'/2^{2(n-8)}$.

Now, let us consider the case that x_4 is a forward query, and precedes both the backward queries x_5 and x_6. As before, $\tilde{X}_0$ can actually match with q_K pairs of (R_1, R_2), and hence there can be at most q_K^2 tuples of (x_2, x_3, x_4) satisfying equation (47). Now, since both x_5 and x_6 are backward queries, $Y_4 = Y_0 \oplus Z_4$ and $Y_5 = Y_0 \oplus Z_5$ are fixed, and for the corresponding fixed values of $\tilde{X}_4$ and $\tilde{X}_5$, the probability that these satisfy equations (48) and (49) is $q^2_{(K,Z_4,Z_5)}/2^{2(n-8)}$. Note that even though each $\tilde{X}_4$ and $\tilde{X}_5$ corresponds to a unique $\tilde{X}_0$, not all of them can be queried, since the number of inverse queries at x_5 and x_6 are both $q_{(K,Z_4,Z_5)}$. The analysis is similar when either x_5 or x_6 are the only forward queries. The probability is actually worse, as argued while bounding $\Pr[\mathcal{E}^1_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}]$. Hence,

$$\Pr[\mathcal{E}^2_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \leq \frac{6q'}{2^{2(n-8)}} + \frac{6q^2_{(K,Z_4,Z_5)}}{2^{2(n-8)}}.$$

Note that we have multiplied the probability of each term by 6 accounting for the six possible execution orders of x_4, x_5, x_6. Expanding out the right hand side, we have

$$\Pr[\mathcal{E}^2_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \leq \frac{2^{27}q_{(K,Z_4,Z_5)}q_K^2}{2^{3n}} + \frac{2^{20}n^{1/2}q_{(K,Z_4,Z_5)}q_K^{1/2}}{2^{2n}} + \frac{2^{19}q^2_{(K,Z_4,Z_5)}}{2^{2n}}.$$

When $q_{(K,Z_4,Z_5)}$ and q_K are large, we have,

$$\Pr[\mathcal{E}^2_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \leq \frac{2^{21}q^2_{(K,Z_4,Z_5)}}{2^{2n}}.$$

Bounding $\Pr[\mathcal{E}^3_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}]$**.** We first consider the case when exactly two out of x_4, x_5 and x_6 are forward queries, and at least one of them is queried before the backward query. As our first instance, let $x_4 \in T_4^{K,+}, x_5 \in T_5^{K,+}, x_6 \in T_6^{K,-}$ and x_4 is queried first. Since x_4 is a forward query, $\tilde{X}_0$ can actually match with q_K pairs of (R_1, R_2). But then, for a fixed $\tilde{X}_0$, all of Y_0, $Y_5 = Y_0 \oplus Z_5$ and hence $\tilde{X}_5$ are unique, and $\tilde{X}_5$ needs to match with $R_1 \oplus 2^5 \cdot R_2 \oplus W_2$ as well. Summing over all choices of $\tilde{X}_5$, equations (47) and (49) are satisfied by at most $\tilde{q} = q_{(K,Z_4,Z_5)}q_K/2^{n-8}$ tuples of (x_2, x_3, x_4, x_6). Note that even though each $\tilde{X}_0$ corresponds to a unique $\tilde{X}_5$, not all of them can be queried, since the number of backward queries at x_6 is $q_{(K,Z_4,Z_5)}$. Since x_5 is forward, the adversary can choose $\tilde{X}_4$. Now, since $q_{(K,Z_4,Z_5)}q_K/2^{n-8} < q_K$, the adversary can make forward queries with all $q_{(K,Z_4,Z_5)}q_K/2^{n-8}$ choices of $\tilde{X}_4$, and for a fixed $\tilde{X}_4$, the probability that it satisfies equation (51) is $1/2^n$. Hence, in this case,

$$\Pr[\textit{Eqns. (47)-(52)}] \leq q_{(K,Z_4,Z_5)}q_K/2^{2n-8}.$$

A very similar analysis holds for all other instances satisfying the condition that exactly two out of x_4, x_5 and x_6 are forward queries, and at least one of them is queried before the backward query. In fact, if

the first forward query is not to E_4, the bound becomes worse as neither $\tilde{X}_4$ nor $\tilde{X}_5$ can match with q_K pairs of (R_1, R_2).

Next, we consider the case when exactly two out of x_4, x_5 and x_6 are backward queries, and at least one of them is queried before the forward query. Let x_4 be the first query. As x_4 is backward, $\tilde{X}_0$ is random. Given $\mathcal{E}_{\mathsf{bad}}$ does not occur, the number of (x_2, x_3, x_4) tuples satisfying equation (47) is at most $q'' := q_{(K,Z_4,Z_5)}q_K^2/2^{n-8} + q_{(K,Z_4,Z_5)}\sqrt{3nq_K}$. Since $q'' < q_K^2$ and Cases 1 and 2 only differ in the direction of the first query, we have $\Pr[Eqns.\ (47)\text{-}(52)] \leq q_{(K,Z_4,Z_5)}q_K/2^{2n-8}$ in this case too. Hence,

$$\Pr[\mathcal{E}^3_{(K,Z_4,Z_5)} \wedge \neg\mathcal{E}_{\mathsf{bad}}] \leq \frac{24q_{(K,Z_4,Z_5)}q_K}{2^{2n-8}} \leq \frac{2^{13}q_{(K,Z_4,Z_5)}q_K}{2^{2n}}.$$

This completes the proof of the lemma. □

E A Brief Overview of Mirror Theory

The mirror theory of Patarin [Pat06, Pat10b, Pat03] gives a lower bound on the number of solutions to a systems of bivariate equations. Consider a system of m bivariate equations over $\mathbb{GF}(2^n)$ on r variables $(Y_1, \ldots, Y_r)$, that can be written as

$$Y_i \oplus Y_j = Z_{i,j}, \tag{111}$$

for some $i \neq j$. Informally, mirror theory says the following: if $Z_{i,j} \neq 0$ in all of the equations and the graph corresponding to these equations (where the variables are nodes and the equations are edges) does not have a cycle or a 0-sum path, and provided that the size $\xi_{\max}$ of the largest component obeys certain bounds, the number of solutions to this system of equations where $Y_i \neq Y_j$ for all i, j is at least

$$\frac{(2^n)_r}{2^{nm}}. \tag{112}$$

The intuition behind this lower bound is that the numerator in the above expression is the total number of solutions satisfying just the distinctness constraint, and any randomly chosen solution has a probability of around $1/2^{nm}$ of satisfying all m bivariate equations.

Mirror theory has a somewhat contentious history. It was originally proposed by Patarin and then refined in a series of works [Pat08, Pat10a, CLP15]. After that, several works used it in their security analysis to obtain better security bounds [IMV16, MN17, ZHY18, BBN22, BCF+24, CLL24]. At the same time, consensus was growing among a number of cryptographers that the original proof of mirror theory contained a number of gaps that could not be readily justified, though these gaps do not seem to have been documented anywhere. Subsequently, Dutta et al. [DNS22] provided a more complete proof of the bound for $\xi_{\max} = 2$, and Cogliati et al. [CDN+23] followed it up with a proof for a wider range of $\xi_{\max}$. With these two works, the technique has gained wider acceptance in the community. In this work, we use one of the results from [CDN+23], which we formally restate as Theorem 3.

F A Birthday Bound Privacy Attack on KC-XAES

The value KC is included in the ciphertext and therefore must also satisfy privacy guarantees. Since KC is a $2n$-bit value, one would expect that finding a collision in this part alone should require about 2^n queries. However, this is not the case for KC-XAES. Recall that

$$\mathsf{KC} = E_K(W_1) \| E_K(W_2),$$

where W_1 and W_2 are as defined in Fig. 11.

We describe the following attack. For $i = 1, 2, \ldots, q$, the adversary:

$\mathsf{KC\text{-}XAES.Enc}_K(\mathbf{N}, AD, P)$

$N_1 \| N_2 \leftarrow \mathbf{N}$
$K_1 \leftarrow X \cdot E_K(\langle 0 \rangle_{128})$
$M_1 \leftarrow (\mathsf{const}_1 \| N_1) \oplus K_1$
$M_2 \leftarrow (\mathsf{const}_2 \| N_1) \oplus K_1$
$\tilde{K} \leftarrow E_K(M_1) \| E_K(M_2)$
$C \leftarrow \mathsf{AES\text{-}GCM.Enc}_{\tilde{K}}(N_2, AD, P)$
$X_1 \leftarrow E_K(\mathsf{const}_3 \| N_1)$
$W_1 \leftarrow X_1 \oplus (N_2 \| \mathsf{const}_4) \oplus K_1$
$W_2 \leftarrow X_1 \oplus (N_2 \| \mathsf{const}_5) \oplus K_1$
$\mathsf{KC} \leftarrow E_K(W_1) \| E_K(W_2)$
return $C \| \mathsf{KC}$

Figure 11: Encryption algorithm of KC-XAES.

1. Chooses arbitrary triples $(\mathbf{N}^i, AD^i, P^i)$.
2. Queries the encryption oracle to obtain

$$(C^i, \mathsf{KC}^i) = \mathsf{KC\text{-}XAES.Enc}_K(\mathbf{N}^i, AD^i, P^i).$$

The value X_1^i is computed as

$$X_1^i = E_K(\mathsf{const}_3 \| N_1^i).$$

Since E behaves as a random permutation, the outputs X_1^i are uniformly distributed over $\{0,1\}^n$. By the birthday bound, after about $2^{n/2}$ queries, the adversary finds indices $i \neq j$ such that

$$X_1^i \oplus (N_2^i \| \langle 0 \rangle_{32}) = X_1^j \oplus (N_2^j \| \langle 0 \rangle_{32}).$$

From the structure of the construction, this implies

$$W_1^i = W_1^j \quad \text{and} \quad W_2^i = W_2^j$$

and therefore

$$\mathsf{KC}^i = \mathsf{KC}^j.$$