A Modular Approach to the Security Analysis of Two-Permutation Constructions

Yu Long Chen

imec-COSIC, KU Leuven, Belgium
yulong.chen@kuleuven.be

Abstract. Constructions based on two public permutation calls are very common in today's cryptographic community. However, each time a new construction is introduced, a dedicated proof must be carried out to study the security of the construction. In this work, we propose a new tool to analyze the security of these constructions in a modular way. This tool is built on the idea of the classical mirror theory for block cipher based constructions, such that it can be used for security proofs in the ideal permutation model. We present different variants of this public permutation mirror theory such that it is suitable for different security notions.

We also present a framework to use the new techniques, which provides the bad events that need to be excluded in order to apply the public permutation mirror theory. Furthermore, we showcase the new technique on three examples: the Tweakable Even-Mansour cipher by Cogliati et al. (CRYPTO '15), the two permutation variant of the pEDM PRF by Dutta et al. (ToSC '21(2)), and the two permutation variant of the nEHtM$_p$ MAC algorithm by Dutta and Nandi (AFRICACRYPT '20). With this new tool we prove the *multi-user* security of these constructions in a considerably simplified way.

Keywords: mirror theory, two permutation calls constructions, multi-user security, modular framework

1 Introduction

Permutation-Based Crypto. Following the selection of Keccak as the winner of the SHA-3 competition [2], cryptographic schemes based on public permutations gained a lot of traction in the research community. Nowadays, permutation-based constructions have become a trend in cryptography, and form a successful and full-fledged alternative to block-cipher based designs. Recently, in the first round of the ongoing NIST lightweight competition [1], 24 out of 57 submissions are based on public permutations, and 16 out of 24 permutation-based designs have been selected for the second round. These statistics show without a doubt the wide acceptance of permutation based designs in the community. The long line of research on the design of secret key constructions using public permutations originates with Even and Mansour [23], who designed a secret random permutation using a public permutation by xoring random keys to the input and output of this permutation. Later, their work was generalized to the Iterated Even-Mansour construction or the Key Alternating Cipher by [6, 9, 17, 26], which is the backbone of today's block ciphers.

Constructions Based on Two Permutation Calls. In recent years, beyond birthday bound security has become a very popular topic in the field of symmetric key cryptography due to the rise of lightweight primitives. Admittedly, the state size of a permutation is typically very large: for example the SHA-3 permutation is of size 1600 bits, and a simple birthday bound secure construction built on SHA-3 would be secure up to an attack complexity of 2^{800}. However, this example permutation is on the extreme end: a big drawback of these big permutations is that they were not designed with lightweight applications in mind, the state of lightweight permutations such as SPONGENT [5] and PHOTON [25] can be as small as 88 and 100 bits, respectively. Hence, birthday bound secure constructions using these types of permutations are inadequate.

Due to the above-mentioned reason, beyond birthday-bound constructions based on two permutations are interesting to investigate. Indeed, it is possible to break any single-permutation construction by finding a collision between the input of the underlying permutation of the given construction and an input to the oracle of the public primitive, which happens with probability $\Omega(qp/2^n)$, where q is the number of queries to the construction and p is the number of queries to the underlying permutation. Constructions using more than two permutation calls can achieve even better security, however these are less efficient and difficult to analyze. On the other hand, constructions based on two permutation calls can achieve a resulting security bound of the form $O(qp^2/2^{2n})$, which is usually sufficient for most practical applications. In the last years, several types of constructions based on two permutation calls were proposed and analyzed. The most notable examples are the 2-round Even-Mansour cipher by Bogdanov et al. [6] and Chen et al. [8], the tweakable block cipher TEM by Cogliati et al. [15] and Dutta [19], the pseudorandom function SoEM by Chen et al. [11] and pEDM by Dutta et al. [22], the FPTP hash function by Chen and Tessaro [12], and the nonce-based MAC algorithm nEHtM$_p$ by Dutta and Nandi [20]. Due to the similarity in the structures, the security proofs of these construction all share some relevance.

Single vs Multi-User Security. The security of most of the above constructions has been proven in the single-user setting. In practice, however, commonly used cryptographic constructions are usually deployed in contexts with a large number of users. An obvious question is to what extent the number of users will affect the security bound of these permutation-based constructions, or more specifically, can these constructions still have a security bound of the form $O(qp^2/2^{2n})$ in the multi-user setting? The concept of multi-user security was first introduced by Bellare, Boldyreva and Micali [3] in the context of public key cryptography, and was later extended by Biham [4] to symmetric key cryptanalysis. In the multi-user setting, attackers can adaptively distribute their q construction queries across multiple users with independent keys, and the attackers succeed as long as they can compromise one user key. Unfortunately, research on provable multi-user security for permutation-based constructions has been missing until now. The notable exceptions are the work of key-alternating ciphers by Mouha and Luykx for a single round [31], and Hoang and Tessaro for multiple rounds [26]. These works show that evaluating how security degrades as the number of users grows is a challenging technical problem, even when the security is known in the single-user setting. The generic reduction [7], however, does not help the constructions to maintain beyond birthday-bound security in the multi-user setting. For example, suppose the number of users is u, then simply applying the generic reduction to obtain multi-user security from single-user security introduces an extra factor u in the security bound. If the attacker only asks one query per user, then the security bound becomes

$$\frac{uqp^2}{2^{2n}} \leq \frac{q^2p^2}{2^{2n}} ,$$

which is only comparable to the $O(qp/2^n)$ security of one-call constructions.

Therefore, it appears that a dedicated analysis of the multi-user security is needed. Most security proofs in symmetric key cryptography today are based on the H-coefficients technique [9,33]. The idea behind this technique is that only a smaller number of transcripts are significantly more likely to appear in the ideal world than in the real world, namely: the bad transcripts. Usually, such proofs are performed as follows: (1) we first define a set of bad transcripts, (2) then the probability of observing bad transcripts in the ideal world is upper bounded, (3) and finally the ratio of observing good transcripts in the real and the ideal world is lower bounded. Note that points (1) and (3) are completely different problems than point (2). Since upper bounding the probability of the bad transcripts is a purely combinatorial problem and has little to do with cryptography, it relies heavily on the randomness of the generated keys. Defining the bad transcripts and lower bounding the ratio of the good transcripts depend, however, strongly on the way a particular construction is built. Unlike the case of block cipher-based constructions, where single-user security is usually proven in the standard model and multi-user security in the ideal-cipher model. For permutation-based constructions, ideal-permutation model analysis is used for both the single and multi-user settings. This raises the question whether or not it is possible to derive a modular approach that can be applied to constructions based on two permutation calls, which generically find the set of bad transcripts and a tight lower bound for the ratio of the corresponding good transcripts in both the *single* and *multi*-user settings, avoiding the long and involved dedicated analysis.

Patarin's Mirror Theory. Before we give an answer to this question, we recall Patarin's mirror theory [34], which is a very powerful but currently still unverified technique. Mirror theory is concerned with systems of $q_m \geq 1$ equations with $r \geq q_m$ unknowns of the form $v \oplus y = \lambda$, where v and y are two unknowns, and λ is a known value. The goal is to determine a lower bound on the number of possible solutions to the unknowns such that the solution does not contain collisions. Originally, Patarin derived mirror theory in order to prove the optimal n-bits security of the Xor of two secret Permutations construction (XoP).

After the modernization by Mennink and Neves [30], who used mirror theory to prove the pseudorandom function security of EDM, EDMD and EWCDM, the applications of mirror theory seem to be increasing. For example, mirror theory was used to prove the security of the 2-round CLRW tweakable block cipher by Jha and Nandi [27], and to prove the security of the nonce-less MAC algorithms PolyMAC, SUM-ECBC, PMAC-Plus, 3kf9, and LightMAC-Plus by Kim et al. [28]. Datta et al. [18] were first to extend the mirror theory by including a system of $q_a \geq 1$ non-equations of the form $v \oplus y \neq \lambda$, and used it to prove the security of the nonce-based MAC algorithm DWCDM. Later, Dutta et al. [21], and Kim et al. [14] used it to prove the security of the nonce-based MAC algorithm nHEtM.

The New Idea. The reason why we can apply mirror theory to the above mentioned block cipher based constructions is because all these constructions can be viewed as the xor of two secret permutations. Note that when the permutations become public, the constructions have a structure that follows the Sum of Even-Mansour (SoEM) construction of Chen et al. [11]. Since the proofs of public permutation based constructions are all performed in the ideal permutation model, the attacker also gets access to the underlying permutations. Hence it is necessary to have an xor before the input and after the output of each of the permutation evaluations. Due to this important observation, almost all constructions based on two permutation calls can be viewed as the xor of two public permutations in the middle. This observation leads to the answer to the previous question, and the goal of this paper is to use the idea of mirror theory to build a modular technique that can be applied to all of the above mentioned permutation-based constructions.

1.1 Our Contribution

The goal of this paper is to derive a generic tool that can be used for the security analysis of constructions based on two public permutation calls and for different security notions. In order to do that, there are a few difficulties that we need to resolve. First of all, the traditional mirror theory is only suitable for the prf security, and we cannot simply apply it to the other settings. The second problem is that mirror theory does not consider primitive queries, and we need to include these queries in order to apply the theory for ideal permutation model proofs.

We solve the first problem using the approach proposed by Jha and Nandi [27] (see SUP Material **??**-**??**), and we derive different versions of public permutation mirror theory that are suitable for almost all popular security notions in symmetric key cryptography. On the other hand, since each primitive query defines exactly the input and the output of one permutation evaluation, we can solve the second problem by including a set of uni-variant affine equations of the form $v = \lambda$ and $y = \lambda$. Each uni-variate affine equation defines exactly one primitive query (where the input and output values of these queries are well defined).

With all these in mind, we derive two new theorems for the ideal public permutation model proofs. In Section 3.1, we explain the general setting for traditional mirror theory, and our new technique for the ideal permutation model is defined and given in Section 3.2. We provide four theorems for different type of settings. Theorem 1 (a) is suitable for security notions such as sprp and tsprp, Theorem 1 (b) is suitable for security notions such as prf, weak prf, and (t)ccr [24], and Theorem 2 (b) is suitable for notions such as mac security (prf with non-equations). Since these three theorems already cover all currently know security notions, Theorem 2 (a) (sprp or tsprp with non-equations) has been added for the sake of completeness, as we are not aware of any notions on which it can be applied and will leave this for future research. We want to emphasize that our aim here is not to fix the proofs in the traditional mirror theory, but rather to focus on a new modular technique for permutation-based constructions. Hence, our technique focuses only on $2n/3$-bit security, since this is a tight security bound (as we will see from the three examples) for constructions based on two permutations due to the presence of the primitive queries.

Another important contribution in this work is to provide a general framework to use the new techniques for the (multi-user) security analysis of constructions based on two independent permutation calls. The framework is given in Section 4. In general, to prove the security of constructions based on two permutation calls, one should first turn the query transcript into a system of bi-variate (non-)equations and uni-variate equations. This system of (non-)equations can be used to define the transcript graph (see Section 4.2). In our case, we decompose our graph into four subgraphs - the union of the components containing one "colliding vertex" (defined by an uni-variate affine equation), the union of "star" components, the set of isolated edges, and the set of isolated vertices. However, the graph may contain other types of components that prevent us from using the new technique, these components need to be excluded in our analysis. A very important part about this framework is that it provides a set of bad events that need to be considered in the security analysis to exclude these components. It seems, especially when non-equations also need to be considered, the analysis becomes very complex, which increases the chance to miss some bad events (see Section 4.3). The probability of these bad events must be upper bounded based on the randomness of the generated keys, sometimes difficult combinatorial techniques are required in the case of limited randomness. After these bad events are excluded, our new theory can be applied to the given system to determine a lower bound on the number of possible solutions to the unknowns, which in its turn defines the ratio of observing good transcripts in the real and the ideal world (see Section 4.4). This framework is useful for future designs, such that the future analysis will not miss any necessary bad events.

Applications. We illustrate the new techniques by applying them to prove the *multi-user* security of Tweakable Even Mansour (TEM), permutation-based Encrypted Davies-Mayer (pEDM), and permutation-based version of nonce-based Enhance Hash-then-Mask (nEHtM$_p$). These three constructions are chosen because they use three important security notions in symmetric-key cryptography, namely tsprp, prf and mac, allowing us to demonstrate the new technique on different notions. On the other hand, the three constructions are the permutation-based variants of important block cipher-based schemes LRW2 [27], EDM [16], nEHtM [21], which have already received much attention in the field.

Firstly, we consider the 2-round TEM construction that was proposed by Cogliati et al. [15]. They showed that 2-round TEM achieves $2n/3$-bit security in the single-user setting. In Section 5, we apply the public permutation mirror theory suitable for tsprp (Theorem 1 (a)) to the TEM construction, and show that TEM also achieves $2n/3$-bit security in the multi-user setting.

Secondly, we consider the pEDM construction that was proposed by Dutta et al. [22]. Again, pEDM was showed to achieve $2n/3$-bit security in the single-user setting. In Section 6, we apply the public permutation mirror theory suitable for prf (Theorem 1 (b)) to pEDM construction, and show that pEDM also achieves $2n/3$-bit security in the multi-user setting. In this example we can clearly see that the analysis in the multi-user setting is more complex than the one in the single-user setting.

Thirdly, we consider the nEHtM$_p$ MAC algorithm, proposed by Dutta and Nandi [20]. They showed that nEHtM$_p$ based on a single permutation (using domain separation) achieves $2n/3$-bit security. We first note that the proof of [20] is incomplete since, according to our framework for MAC designs, the authors missed some bad events in their analysis. This observation was also verified by the authors [10]. Some of these additional bad events, however, require involved arguments to bound. In order to solve this problem, we modify the nEHtM$_p$ construction by adding an extra universal hash function call. This modified construction uses more randomness, which in turn enables us to bound the additional bad events easily. We would like to note that our analysis does not imply infeasibility in fixing the proof of nEHtM$_p$. In Section 7, we will prove the *multi-user* security of this modified variant nEHtM$_p$ using our public permutation extended mirror theory (Theorem 2 (b)), and we show that it achieves $2n/3$-bit security in the multi-user setting.

We believe that the techniques have a wide range of applications in the future design of public permutation based schemes. For example, when building nonce-less MAC algorithms and (authenticated) encryption schemes with beyond birthday bound security using public permutations, as done in the case of block cipher-based mirror theory [13, 28].

2 Preliminaries

For $n \in \mathbb{N}$, we denote by $[n]$ the shorthand notation for $\{1, \ldots, n\}$, and by $\{0,1\}^n$ the set of bit strings of length n. For two bit strings $X, Y \in \{0,1\}^n$, we denote their bitwise addition as $X \oplus Y$. For a value Z, we denote by $A \leftarrow Z$ the assignment of Z to the variable A. For a finite set $\mathcal{S}$, we denote by $S \xleftarrow{\$} \mathcal{S}$ the uniformly random selection of S from $\mathcal{S}$. We denote by $\mathrm{Func}(m, n)$ the set of all functions that map $\{0,1\}^m$ to $\{0,1\}^n$, and by $\mathrm{Func}(n)$ the set of all functions that maps $\{0,1\}^n$ to $\{0,1\}^n$. We denote by $\mathrm{Perm}(n)$ the set of all permutations on $\{0,1\}^n$, and by $\widetilde{\mathrm{Perm}}(t, n)$ the set of all functions $\tilde{\pi} : \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$

such that $\tilde{\pi}(T, \cdot)$ is in $\mathrm{Perm}(n)$ for all $T \in \{0,1\}^t$. For any integers a, b such that $1 \leq b \leq a$, we have $(a)_b = a \cdot (a-1) \ldots (a-b+1)$ and $(a)_0 = 1$.

For $q \in \mathbb{N}$, we denote by x^{*q} the q-tuple $(x_1, \ldots, x_q)$, and by $\hat{x}^{*q}$ the set $\{x_i : i \in [q]\}$. By an abuse of notation we also use x^{*q} to denote the multiset $\{x_i : i \in [q]\}$, and we denote by $\mu(x^{*q}, x')$ the multiplicity of $x' \in x^{*q}$. For two disjoint sets P and Q, we denote their (disjoint) union as $P \sqcup Q$.

2.1 Tweakable Block Ciphers Based on Public Permutations

For $k, n, r, t, u \in \mathbb{N}$, consider a tweakable block cipher $\tilde{E} : \{0,1\}^k \times \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$ that is based on $\pi_1, \ldots, \pi_r \xleftarrow{\$} \mathrm{Perm}(n)$, such that for fixed key $K \in \{0,1\}^k$, the function $\tilde{E}_K(T, \cdot) = \tilde{E}(K, T, \cdot)$ is a permutation on $\{0,1\}^n$. We denote its inverse (for fixed key and tweak) by $\tilde{E}_K^{-1}(T, \cdot) = \tilde{E}^{-1}(K, T, \cdot)$, and $\tilde{E}_K^{-1}$ should behave independently for different tweaks. We will consider the multi-user tweakable strong pseudorandom permutation (mu-tsprp) security of $\tilde{E}$, where the distinguisher $\mathcal{D}$ is given two-directional access to either $(\tilde{E}^{\pm}_{K_1}, \ldots, \tilde{E}^{\pm}_{K_u}, \pi_1^{\pm}, \ldots, \pi_r^{\pm})$ for secret keys $K_1, \ldots, K_u \xleftarrow{\$} \{0,1\}^k$, or $(\tilde{\pi}_1^{\pm}, \ldots, \tilde{\pi}_u^{\pm}, \pi_1^{\pm}, \ldots, \pi_r^{\pm})$ for $\tilde{\pi}_1, \ldots, \tilde{\pi}_u \xleftarrow{\$} \widetilde{\mathrm{Perm}}(t, n)$. The goal is to determine which world it interacted with:

$$\mathbf{Adv}_{\tilde{E}}^{\text{mu-tsprp}}(\mathcal{D}) = \left| \Pr\left[\mathcal{D}^{\tilde{E}^{\pm}_{K_1}, \ldots, \tilde{E}^{\pm}_{K_u}, \pi_1^{\pm}, \ldots, \pi_r^{\pm}} = 1\right] - \Pr\left[\mathcal{D}^{\tilde{\pi}_1^{\pm}, \ldots, \tilde{\pi}_u^{\pm} \pi_1^{\pm}, \ldots, \pi_r^{\pm}} = 1\right] \right| .$$

Here the superscript ± indicates that the distinguisher has bi-directional access. When $u = 1$, we consider the single-user security of $\tilde{E}$, and we simply denote $\mathcal{D}$'s advantage in distinguishing the real world from random by $\mathbf{Adv}_{\tilde{E}}^{\text{tsprp}}(\mathcal{D})$.

2.2 Pseudorandom Functions Based on Public Permutations

For $k, m, n, r, u \in \mathbb{N}$, consider a pseudorandom function $F \colon \{0,1\}^k \times \{0,1\}^m \to \{0,1\}^n$ that is based on $\pi_1, \ldots, \pi_r \xleftarrow{\$} \mathrm{Perm}(n)$, such that for fixed key $K \in \{0,1\}^k$, $F_K(\cdot) = F(K, \cdot)$ is a function that maps $\{0,1\}^m$ to $\{0,1\}^n$. We will consider the multi-user pseudorandom function (mu-prf) security of F, where the distinguisher $\mathcal{D}$ is given access to either $(F_{K_1}, \ldots, F_{K_u}, \pi_1^{\pm}, \ldots, \pi_r^{\pm})$ for secret keys $K_1, \ldots, K_u \xleftarrow{\$} \{0,1\}^k$, or $(\varphi_1, \ldots, \varphi_u, \pi_1^{\pm}, \ldots, \pi_r^{\pm})$ for $\varphi_1, \ldots, \varphi_u \xleftarrow{\$} \mathrm{Func}(n)$. The goal is to determine which world it interacted with:

$$\mathbf{Adv}_{F}^{\text{mu-prf}}(\mathcal{D}) = \left| \Pr\left[\mathcal{D}^{F_{K_1}, \ldots, F_{K_u}, \pi_1^{\pm}, \ldots, \pi_r^{\pm}} = 1\right] - \Pr\left[\mathcal{D}^{\varphi_1, \ldots, \varphi_u, \pi_1^{\pm}, \ldots, \pi_r^{\pm}} = 1\right] \right| .$$

Here the superscript ± for π's indicates that the distinguisher has bi-directional access. When $u = 1$, we consider the single-user security of F, and we simply denote $\mathcal{D}$'s advantage in distinguishing the real world from random by $\mathbf{Adv}_{F}^{\text{prf}}(\mathcal{D})$.

2.3 Nonce-Based MAC Algorithms Based on Public Permutations

For $k, n, r, t \in \mathbb{N}$, consider a nonce-based message authentication code (MAC) algorithm $F \colon \{0,1\}^k \times \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^t$ that is based on $\pi_1, \ldots, \pi_r \xleftarrow{\$}$

$\mathrm{Perm}(n)$. For any fixed key $K \in \{0,1\}^k$, we write $F_K(\cdot,\cdot) = F(K,\cdot,\cdot)$. We denote by $\mathsf{Ver}\colon \{0,1\}^k \times \{0,1\}^n \times \{0,1\}^* \times \{0,1\}^t \to 1/0$ the verification oracle that is based on $\pi_1,\ldots,\pi_r$, which takes as input a key $K \in \{0,1\}^k$, a nonce $N \in \{0,1\}^n$, a message $M \in \{0,1\}^*$, and tag $T \in \{0,1\}^t$, and outputs 1 if the tag T is correct and 0 otherwise.

For $u \in \mathbb{N}$, the multi-user message authentication code (mu-mac) security of F is measured by considering a distinguisher $\mathcal{D}$ that is given access to $(F_{K_1}, \mathsf{Ver}_{K_1}),\ldots,(F_{K_u}, \mathsf{Ver}_{K_u})$ for secret keys $K_1,\ldots,K_u \xleftarrow{\$} \{0,1\}^k$, and the primitive oracles $\pi_1,\ldots,\pi_r$. For any $j = 1,\ldots,u$, the goal of $\mathcal{D}$ is to fool the verification oracle with a valid but new (j,N,M,T), and its advantage with respect to this task is defined as

$$\mathbf{Adv}_F^{\text{mu-mac}}(\mathcal{D}) = \Pr\Big[K_1,\ldots,K_u \xleftarrow{\$} \{0,1\}^k : \\ \mathcal{D}^{(F_{K_1},\mathsf{Ver}_{K_1}),\ldots,(F_{K_u},\mathsf{Ver}_{K_u}),\pi_1^{\pm},\ldots,\pi_r^{\pm}} \text{ forges}\Big]\,,$$

where "forges" means that the distinguisher enters a tuple (j,N,M,T) such that $\mathsf{Ver}_{K_j}(N,M,T)$ returns 1 and $F_{K_j}(N,M)$ has never been queried.

We call a MAC query to the j-th user (j,N,M) a faulty query if the distinguisher $\mathcal{D}$ has already queried F_{K_j} with the same nonce N and a different message M. The distinguisher $\mathcal{D}$ is allowed to make at most μ faulty MAC queries over u users. We call $\mathcal{D}$ a nonce-respecting adversary if $\mu = 0$, and nonce-misusing if $\mu \geq 1$. We stress that $\mathcal{D}$ may always repeat nonces in its verification queries.

It will be more convenient to express $\mathbf{Adv}_F^{\text{mu-mac}}(\mathcal{D})$ as a distinguisher's advantage. For $j = 1,\ldots,u$, we define perfectly random oracles $\mathsf{Rand}_j\colon \{0,1\}^n \times \{0,1\}^* \to \{0,1\}^t$, and rejection oracles $\mathsf{Rej}_j\colon \{0,1\}^n \times \{0,1\}^* \times \{0,1\}^t \to 0$.

To obtain an upper bound for the forging advantage of a message authentication code F with respect to the distinguisher $\mathcal{D}$, we consider another distinguisher $\mathcal{D}'$, that is given access to either the real world oracles $\mathcal{O},\pi_1^{\pm},\ldots,\pi_r^{\pm}$, or the ideal world oracles $\mathcal{P},\pi_1^{\pm},\ldots,\pi_r^{\pm}$. Then, $\mathcal{D}'$'s advantage is upper bounded by:

$$\mathbf{Adv}_F^{\text{mu-mac}}(\mathcal{D}') \leq \left|\Pr\left[\mathcal{D}'^{\mathcal{O},\pi_1^{\pm},\ldots,\pi_r^{\pm}} = 1\right] - \Pr\left[\mathcal{D}'^{\mathcal{P},\pi_1^{\pm},\ldots,\pi_r^{\pm}} = 1\right]\right|\,,$$

with $\mathcal{O} = \big((F_{K_1},\mathsf{Ver}_{K_1}),\ldots,(F_{K_u},\mathsf{Ver}_{K_u})\big)$ for secret keys $K_1,\ldots,K_u \xleftarrow{\$} \{0,1\}^k$, and $\mathcal{P} = \big((\mathsf{Rand}_1,\mathsf{Rej}_1),\ldots,(\mathsf{Rand}_u,\mathsf{Rej}_u)\big)$.

Here the superscript ± for the π_i's indicates that the distinguisher has bidirectional access. We call a distinguisher $\mathcal{D}'$ non-trivial if it never makes a query (j,N,M,T) to its j-th verification oracle when a previous query (j,N,M) to its j-th MAC oracle returned T. When $u = 1$, we consider the single-user security of F, and we simply denote $\mathcal{D}'$'s advantage in distinguishing the real world from random by $\mathbf{Adv}_F^{\text{mac}}(\mathcal{D}')$.

2.4 Universal Hash Functions

For $n \in \mathbb{N}$, let $H \colon \mathcal{K}_h \times \mathcal{M} \to \{0,1\}^n$ such that for $K_h \in \mathcal{K}_h$, $H_{K_h}(\cdot) = H(K_h, \cdot)$ is called an ϵ-almost XOR universal (ϵ-AXU) hash function [29] if for all distinct $M, M' \in \mathcal{M}$ and all $C \in \{0,1\}^n$, we have

$$\Pr\left[K_h \xleftarrow{\$} \mathcal{K}_h \colon H_{K_h}(M) \oplus H_{K_h}(M') = C\right] \leq \epsilon\,.$$

For $q \in \mathbb{N}$, fix $M_1, \ldots, M_q \in \mathcal{M}$. For $K_h \in \mathcal{K}_h$, let $X_i = H_{K_h}(M_i)$ for $i = 1, \ldots, q$. We define an equivalence relation $\sim$ on $[q]$ as: $\alpha \sim \beta$ if and only if $X_\alpha = X_\beta$, for $\alpha, \beta \in [q]$. For $r \in \mathbb{N}$, we denote by $\mathcal{P}_1, \ldots, \mathcal{P}_r$ the non-trivial equivalence classes of $[q]$ with respect to $\sim$, and we define $\nu_i = |\mathcal{P}_i| \geq 2$ for $i = 1, \ldots, r$. Jha and Nandi [27] proved the following lemmas, that will be useful in our security proof.

Lemma 1. *Let $\nu_i, i = 1, \ldots, r$, be the random variables as defined above. Then, we have*

$$\boldsymbol{E}\Big[\sum_{i=1}^{r} \nu_i\Big] \leq q^2\epsilon/2\,, \qquad \boldsymbol{E}\Big[\sum_{i=1}^{r} \nu_i^2\Big] \leq 2q^2\epsilon\,.$$

Lemma 2. *Let $\nu_{\max} = \max\{\nu_i : i \in [r]\}$. Then, for some $a \geq 2$, we have*

$$\Pr[\nu_{\max} \geq a] \leq \frac{2q^2\epsilon}{a^2}\,.$$

2.5 Expectation Method

In this work, we use the expectation method by Hoang and Tessaro [26], a generalization of Patarin's H-coefficient technique [9, 33].

Consider two oracles $\mathcal{O}$ and $\mathcal{P}$, and a deterministic distinguisher $\mathcal{D}$ that has query access to either of these oracles. The distinguisher's goal is to distinguish both worlds, and we denote by

$$\mathbf{Adv}(\mathcal{D}) = \left|\Pr\left[\mathcal{D}^{\mathcal{O}} = 1\right] - \Pr\left[\mathcal{D}^{\mathcal{P}} = 1\right]\right|$$

its advantage. We define a transcript τ which summarizes all query-response tuples learned by $\mathcal{D}$ during its interaction with its oracle $\mathcal{O}$ or $\mathcal{P}$. We denote by $X_{\mathcal{O}}$ and $X_{\mathcal{P}}$ the random variables equal to transcript produced when interacting with $\mathcal{O}$ and $\mathcal{P}$, respectively. We call a transcript $\tau \in \mathcal{T}$ attainable if $\Pr[X_{\mathcal{P}} = \tau] > 0$, or in other words if the transcript τ can be obtained from an interaction with $\mathcal{P}$.

Lemma 3 (expectation method [26]). *Consider a deterministic distinguisher $\mathcal{D}$. Define a partition $\mathcal{T} = \mathcal{T}_{\text{good}} \sqcup \mathcal{T}_{\text{bad}}$, where $\mathcal{T}_{\text{good}}$ is the subset of $\mathcal{T}$ which contains all the "good" transcripts and $\mathcal{T}_{\text{bad}}$ is the subset with all the "bad" transcripts. Let $\phi \colon \mathcal{T} \to [0, \infty)$ be a non-negative function mapping any attainable transcript to a non-negative real value, such that for all $\tau \in \mathcal{T}_{\text{good}}$:*

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} \geq 1 - \phi(\tau)\,. \qquad (1)$$

Then, we have $\mathbf{Adv}(\mathcal{D}) \leq \boldsymbol{E}[\phi(X_{\mathcal{P}})] + \Pr[X_{\mathcal{P}} \in \mathcal{T}_{\mathrm{bad}}]$.

The H-coefficients technique can be seen as a simple corollary of the expectation method when ϕ is equal to a constant function.

Preliminary Observations. For $\pi \xleftarrow{\$} \mathrm{Perm}(n)$ and a permutation queries transcript τ_π, we say that π extends τ_π, denoted $\pi \vdash \tau_\pi$, if $\pi(u) = v$ for all $(u, v) \in \tau_\pi$. By extension, for $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_r) \xleftarrow{\$} \Big(\mathrm{Perm}(n)\Big)^r$ and a tuple of permutation queries transcript $\tau_{\boldsymbol{\pi}} = (\tau_{\pi_1}, \ldots, \tau_{\pi_r})$, we say that $\boldsymbol{\pi}$ extends $\tau_{\boldsymbol{\pi}}$, denoted $\boldsymbol{\pi} \vdash \tau_{\boldsymbol{\pi}}$, if $\pi_i \vdash \tau_{\pi_i}$ for $i = 1, \ldots, r$.

Consider an attainable transcript $\tau \in \mathcal{T}_{\mathrm{good}}$, and let $\mathcal{P}$ be an uniformly chosen random oracle. For permutation based constructions, let $\tau = (\tau_0, \tau_{\boldsymbol{\pi}})$, where τ_0 contains queries to the construction oracle $\mathcal{O}$ or $\mathcal{P}$, and $\tau_{\boldsymbol{\pi}}$ contains the queries to the primitive oracles $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_r)$. To compute $\Pr[X_{\mathcal{O}} = \tau]$ and $\Pr[X_{\mathcal{P}} = \tau]$, it suffices to compute the probability of oracles that could result in view τ. We first consider the ideal world oracle $\mathcal{P}$, and obtain

$$\Pr[X_{\mathcal{P}} = \tau] = \frac{1}{|\mathcal{K}|^r} \cdot \left(\frac{1}{(2^n)_p}\right)^r \cdot \Pr[\mathcal{P} \colon \mathcal{P} \vdash \tau_0]\,.$$

The first term corresponds to the number of dummy keys that are drawn uniformly at random; the second term is the probability that $\boldsymbol{\pi}$ extends $\tau_{\boldsymbol{\pi}}$; and the last term is the probability that $\mathcal{P}$ extends τ_0.

Similarly we say that a real world oracle $\mathcal{O}$ extends τ if it extends τ_0 and $\tau_{\boldsymbol{\pi}}$. For $K \xleftarrow{\$} \mathcal{K}^r$, we have

$$\Pr[X_{\mathcal{O}} = \tau] = \frac{1}{|\mathcal{K}|^r} \cdot \left(\frac{1}{(2^n)_p}\right)^r \cdot \Pr\left[\boldsymbol{\pi} \xleftarrow{\$} \Big(\mathrm{Perm}(n)\Big)^r : \mathcal{O}_K^{\boldsymbol{\pi}} \vdash \tau_0 \mid \boldsymbol{\pi} \vdash \tau_{\boldsymbol{\pi}}\right].$$

The first term corresponds to the number of randomly drawn keys that are used in the construction; the second term is the probability that $\boldsymbol{\pi}$ extends $\tau_{\boldsymbol{\pi}}$; and the last term is the probability that $\mathcal{O}_K^{\boldsymbol{\pi}}$ extends τ_0, given that $\boldsymbol{\pi}$ extends $\tau_{\boldsymbol{\pi}}$.

Let $\rho(\tau) = \Pr\left[\boldsymbol{\pi} \xleftarrow{\$} \Big(\mathrm{Perm}(n)\Big)^r : \mathcal{O}_K^{\boldsymbol{\pi}} \vdash \tau_0 \mid \boldsymbol{\pi} \vdash \tau_{\boldsymbol{\pi}}\right]$. Take for instance $r = 2$, and assume that each primitive query transcript contains p queries to the given permutation. Suppose we sample distinct outputs of π_1 (resp., π_2) over for example q_V (resp., q_Y) distinct inputs. Then, it is easy to see that $\rho(\tau) = h_q/(2^n - p)_{q_V - p}(2^n - p)_{q_Y - p}$, where h_q is the number of solutions of distinct outputs of π_1 and π_2. Then we have

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} = \rho(\tau) / \Pr[\mathcal{P} \colon \mathcal{P} \vdash \tau_0] \geq 1 - \varepsilon_1\,. \tag{2}$$

3 (Extended) Mirror Theory in Ideal Permutation Model

We explain the general settings behind the traditional (extended) mirror theory in Section 3.1, which involves only bi-variate affine equations and possible non-equations. In Section 3.2, we introduce the new public permutation mirror theory, that takes primitive queries into account by including uni-variate affine equations.

3.1 System of Bi-Variate Affine Equations and Non-Equations

Let $q_m, q_a, q_V, q_Y \geq 1$. Let $\mathcal{V} = \{v_1, \ldots, v_{q_V}\}$ be a set of q_V unknowns and $\mathcal{Y} = \{y_1, \ldots, y_{q_Y}\}$ be a set of q_Y unknowns. We consider a system $\mathcal{E}_m$ of q_m bi-variate affine equations

$$\mathcal{E}_m = \{v_{I_1} \oplus y_{I_1} = \lambda_1, \ldots, v_{I_{q_m}} \oplus y_{I_{q_m}} = \lambda_{q_m}\}\,.$$

In some cases (for example mac security), we also need to consider a system $\mathcal{E}_a$ of q_a bi-variate affine non-equations

$$\mathcal{E}_a = \{v'_{J_1} \oplus y'_{J_1} \neq \lambda'_1, \ldots, v'_{J_{q_a}} \oplus y'_{J_{q_a}} \neq \lambda'_{q_a}\}\,,$$

where v_{I_i}'s, y_{I_i}'s, v'_{J_j}'s, and y'_{J_j}'s are unknowns, and λ_i's and λ'_j's are knowns, for $i = 1, \ldots, q_m$ and $j = 1, \ldots, q_a$. We want to state that the sets $\mathcal{V}$ and $\mathcal{Y}$ are disjoint.

We define two surjective index mappings:

$$\varphi_V \colon \{I_1, \ldots, I_{q_m}, J_1, \ldots, J_{q_a}\} \to \{1, \ldots, q_V\}\,,$$
$$\varphi_Y \colon \{I_1, \ldots, I_{q_m}, J_1, \ldots, J_{q_a}\} \to \{1, \ldots, q_Y\}\,.$$

such that $q_V, q_Y \leq q_m + q_a$. Note that I_i and J_j are respectively the indices of the unknowns in $\mathcal{E}_m$ and $\mathcal{E}_a$. However, multiple unknowns with different indices can be the same. In that case, these unknown are all mapped to the same value using φ_V or φ_Y. The system $\mathcal{E} = \mathcal{E}_m \sqcup \mathcal{E}_a$ is uniquely determined by $(\varphi'_V, \varphi'_Y, \lambda^{*q_m}, \lambda'^{*q_a})$.

Consider a graph $\mathcal{G}(\mathcal{E}) = (\mathcal{V}, \mathcal{Y}, \mathcal{S} \sqcup \mathcal{S}')$, where the edge set is partitioned into two disjoint sets $\mathcal{S}$ and $\mathcal{S}'$. Here $\mathcal{S}$ and $\mathcal{S}'$ denote the set of λ-labeled edges and the set of λ'-labeled edges, respectively. The graph $\mathcal{G}(\mathcal{E})$ can be seen as a superposition of two subgraphs $\mathcal{G}(\mathcal{E}_m) = (\mathcal{V}, \mathcal{Y}, \mathcal{S})$ and $\mathcal{G}(\mathcal{E}_a) = (\mathcal{V}, \mathcal{Y}, \mathcal{S}')$. Let $\overline{v_s y_t} \in \mathcal{S}$ be an edge for $v_s \in \mathcal{V}$ and $y_t \in \mathcal{Y}$, then $\overline{v_s y_t}$ is labeled with an element in λ^{*q_m}. If the given edge is labeled with λ_i (for $i = 1, \ldots, q_m$), then this edge and the connected vertices v_s and y_t represent the equation $v_s \oplus y_t = \lambda_i$, where $s = \varphi_V(I_i)$ and $t = \varphi_Y(I_i)$. Similarly, let $\overline{v_s y_t} \in \mathcal{S}'$ be an edge for $v_s \in \mathcal{V}$ and $y_t \in \mathcal{Y}$, then $\overline{v_s y_t}$ is labeled with an element in λ'^{*q_a}. If the given edge is labeled with λ'_j (for $j = 1, \ldots, q_a$), then this edge and the connected vertices v_s and y_t represent the non-equation $v_s \oplus y_t \neq \lambda'_j$, where $s = \varphi_V(J_j)$ and $t = \varphi_Y(J_j)$. Here, each equation in $\mathcal{E}_m$ corresponds to a unique λ-labeled edge in $\mathcal{G}(\mathcal{E}_m)$, and each non-equation in $\mathcal{E}_a$ corresponds to a unique λ'-labeled edge in $\mathcal{G}(\mathcal{E}_a)$. Note that when the system of non-equations $\mathcal{E}_a$ is empty, then the graph $\mathcal{G}(\mathcal{E})$ does not contain isolated vertices, every vertex is incident with at least one λ-labeled edge. Otherwise, the subgraph $\mathcal{G}(\mathcal{E}_m)$ may contain isolated vertices, and these vertices are connected with a λ'-labeled edge in $\mathcal{G}(\mathcal{E}_a)$.

We say two distinct equations in $\mathcal{E}_m$ are in the same component if and only if the corresponding edges (or vertices) in $\mathcal{G}(\mathcal{E}_m)$ are in the same component. Let $\ell > 0$ and a path

$$\mathcal{L} : a_0 \overset{\lambda_1}{-} a_1 \overset{\lambda_2}{-} \ldots \overset{\lambda_\ell}{-} a_\ell$$

in $\mathcal{G}(\mathcal{E}_m)$, for some vertices $a_0, a_1, \ldots, a_\ell \in \mathcal{V} \sqcup \mathcal{Y}$ that are in the same component. The label of $\mathcal{L}$ is defined as

$$\lambda(\mathcal{L}) = \lambda_1 \oplus \lambda_2 \oplus \ldots \oplus \lambda_\ell \, .$$

In case $\mathcal{E}_a$ is empty, a graph $\mathcal{G}(\mathcal{E})$ is called a *good* graph if its subgraph $\mathcal{G}(\mathcal{E}_m)$ satisfies the following properties.

Definition 1 (acyclic). *There is an unique path $\mathcal{L}$ in the subgraph $\mathcal{G}(\mathcal{E}_m)$ between any two vertices a and b in the same connected component, for $a, b \in \mathcal{V} \sqcup \mathcal{Y}$.*

Definition 2 (non-degeneracy). *For all paths $\mathcal{L}$ of even length at least 2 in the subgraph $\mathcal{G}(\mathcal{E}_m)$, we have $\lambda(\mathcal{L}) \neq 0$.*

Definition 3 (ξ-block-maximality). *For a component $\mathcal{I}$, we denote its size by $\xi(\mathcal{I})$, which is the number of vertices in $\mathcal{I}$. We denote the maximum component size by $\xi_{\max}$ such that $\xi(\mathcal{I}) \leq \xi_{\max}$ for all $\mathcal{I}$ in $\mathcal{G}(\mathcal{E}_m)$.*

In case the system $\mathcal{E}_a$ contains at least one non-equation, a graph $\mathcal{G}(\mathcal{E})$ is called a *good* graph if its subgraph $\mathcal{G}(\mathcal{E}_m)$ satisfies the above three properties, and if $\mathcal{G}(\mathcal{E})$ also satisfies the following property.

Definition 4 (non-zero cycle label (NCL)). *If vertices v and y are connected with a λ'-labeled edge, then they are not connected by a $\lambda(\mathcal{L})$-labeled path in $\mathcal{G}(\mathcal{E}_m)$ such that $\lambda(\mathcal{L}) = \lambda'$, for $v \in \mathcal{V}'$ and $y \in \mathcal{Y}'$.*

In an edge-labeled bipartite graph $\mathcal{G}$, we call a component $\mathcal{I}$ of $\mathcal{G}$ an isolated component if $\mathcal{I}$ only contains a path of length one. So $\mathcal{I}$ only contains an edge (v, y, λ) where both v and y have degree 1. We call a component $\mathcal{I}$ of $\mathcal{G}$ a star component if $\xi(\mathcal{I}) \geq 3$, and if there is an unique vertex in $\mathcal{I}$ with degree $\xi(\mathcal{I}) - 1$. We call this vertex the center of $\mathcal{I}$. Further, we call $\mathcal{I}$ a v-$\star$ (resp., y-$\star$) component if its center lies in v (resp., y).

3.2 System of Bi-Variate and Uni-Variate Affine Equations and Bi-Variate Affine Non-Equations

In order to handle the primitive queries, we extend the system of bi-variate affine equations and non-equations with $2p$ uni-variate affine equations. Each uni-variate affine equation defines one primitive queries. Let $q_m, q_a, q_V^p, q_Y^p, p \geq 1$. Let $\mathcal{V}^p = \{v_1, \ldots, v_{q_V^p}\}$ be a set of q_V^p unknowns and $\mathcal{Y}^p = \{y_1, \ldots, y_{q_Y^p}\}$ be a set of q_Y^p unknowns. The new systems are $\mathcal{E}_m^p$, that contains q_m bi-variate affine equations and $2p$ uni-variate affine equations

$$\begin{aligned} \mathcal{E}_m^p = \{ & v_{I_1} \oplus y_{I_1} = \lambda_1, \ldots, v_{I_{q_m}} \oplus y_{I_{q_m}} = \lambda_{q_m} \, , \\ & v_{I_{q_m+1}} = \lambda_{q_m+1}, \ldots, v_{I_{q_m+p}} = \lambda_{q_m+p}, \\ & y_{I_{q_m+1}} = \lambda_{q_m+p+1}, \ldots, y_{I_{q_m+p}} = \lambda_{q_m+2p}\} \, . \end{aligned}$$

In some cases (for example mac security), we also need to consider a system $\mathcal{E}_a$ of q_a bi-variate affine non-equations

$$\mathcal{E}_a = \{v'_{J_1} \oplus y'_{J_1} \neq \lambda'_1, \ldots, v'_{J_{q_a}} \oplus y'_{J_{q_a}} \neq \lambda'_{q_a}\}\,,$$

where v_{I_i}'s, y_{I_i}'s, v'_{J_j}'s, and y'_{J_j}'s are unknowns, and λ_k's and λ'_j's are knowns, for $i = 1, \ldots, q_m + p$, $j = 1, \ldots, q_a$, and $k = 1, \ldots, q_m + 2p$, such that $\lambda_{q_m+1} \neq \ldots \neq \lambda_{q_m+p}$ and $\lambda_{q_m+p+1} \neq \ldots \neq \lambda_{q_m+2p}$. We want to state that the sets $\mathcal{V}^p$ and $\mathcal{Y}^p$ are disjoint.

We define two surjective index mappings:

$$\begin{aligned}\varphi^p_V &\colon \{I_1, \ldots, I_{q_m+p}, J_1, \ldots, J_{q_a}\} \to \{1, \ldots, q^p_V\}\,,\\ \varphi^p_Y &\colon \{I_1, \ldots, I_{q_m+p}, J_1, \ldots, J_{q_a}\} \to \{1, \ldots, q^p_Y\}\,,\end{aligned}$$

such that $q_V, q_Y \leq q_m + q_a + p$. The system $\mathcal{E}^p = \mathcal{E}^p_m \sqcup \mathcal{E}_a$ is uniquely determined by $(\varphi^p_V, \varphi^p_Y, \lambda^{*q_m+2p}, \lambda'^{*q_a})$.

Since the last $2p$ uni-variate affine equations define the values of the $2p$ unknowns v_{I_i}'s and y_{I_i}'s exactly, for $i = q_m + 1, \ldots, q_m + p$. Hence, we know that exactly p unknowns in $\mathcal{V}^p$ and p unknowns in $\mathcal{Y}^p$ are already well defined by the system $\mathcal{E}^p_m$. We define

$$V_0 = \{v_{\varphi^p_V(I_{q_m+1})}, \ldots, v_{\varphi^p_V(I_{q_m+p})}\}\,, \qquad Y_0 = \{y_{\varphi^p_Y(I_{q_m+1})}, \ldots, y_{\varphi^p_Y(I_{q_m+p})}\}\,,$$

as the sets that contain these $2p$ unknowns such that $|V_0| = p$ and $|Y_0| = p$. We are particularly interested in the unknowns from the sets $\mathcal{V}^p \setminus V_0$ and $\mathcal{Y}^p \setminus Y_0$, since these are the unknowns that appear in the q_m bi-variate affine equations and q_a bi-variate affine non-equations.

Consider a bipartite edge-labeled graph $\mathcal{G}(\mathcal{E}^p) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S} \sqcup \mathcal{S}')$, the edge set is partitioned into two disjoint sets $\mathcal{S}$ and $\mathcal{S}'$ as before. The graph $\mathcal{G}(\mathcal{E}^p)$ can be seen as a superposition of two subgraphs $\mathcal{G}(\mathcal{E}^p_m) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S})$ and $\mathcal{G}(\mathcal{E}_a) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S}')$. Here, each of the q_m bi-variate affine equations in $\mathcal{E}^p_m$ corresponds to a unique λ-labeled edge in $\mathcal{G}(\mathcal{E}^p_m)$, each non-equation in $\mathcal{E}_a$ corresponds to a unique λ'-labeled edge in $\mathcal{G}(\mathcal{E}_a)$, and each of the $2p$ uni-variate affine equations in $\mathcal{E}^p_m$ corresponds to a vertex with well defined value in $\mathcal{G}(\mathcal{E}^p)$. Note that the subgraph $\mathcal{G}(\mathcal{E}^p_m)$ may contain isolated vertices, and these vertices are either connected with a λ'-labeled edge in $\mathcal{G}(\mathcal{E}_a)$ or they are isolated colliding vertices in $\mathcal{G}(\mathcal{E}^p)$ with a well-defined value. The subgraph $\mathcal{G}(\mathcal{E}^p_m)$ may also contain components that contain vertices with well defined value. We call these components the "colliding components", and the vertices with well defined values the "colliding vertices".

We distinguish two different cases. In the first case, assume that $\mathcal{E}_a$ is empty, hence we will focus on a graph $\mathcal{G}(\mathcal{E}^p)$ where its subgraph $\mathcal{G}(\mathcal{E}^p_m)$ satisfies the (i) acyclic, (ii) non-degeneracy, and (iii) ξ-block-maximality properties (Definition 1-3). In addition, $\mathcal{G}(\mathcal{E}^p_m)$ also needs to satisfy the following property.

Definition 5 (single colliding vertex (SCV)). *Each component in the graph $\mathcal{G}(\mathcal{E}^p_m)$ contains at most one colliding vertex.*

Note that Definition 5 is necessary in order to give a unique assignment to every vertex in the graph, since if a vertex is assigned with any value, then the labeled edges determine the values of all other vertices in the component containing this vertex. We call any graph that satisfies these four properties a *good* graph. Given a good graph $\mathcal{G}(\mathcal{E}_m^p) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S})$, a solution to $\mathcal{G}(\mathcal{E}_m^p)$ is an assignment of distinct values to the v vertices in $\mathcal{V}^p$ and distinct values to the y vertices in $\mathcal{Y}^p$ satisfying all λ-labeled equations.

We consider a system of bi-variate and uni-variate affine equations $\mathcal{E}_m^p$, such that each component in $\mathcal{G}(\mathcal{E}_m^p)$ is either an isolated edge, a star, or isolated colliding vertex. In order to find the number of solutions to $\mathcal{G}(\mathcal{E}_m^p)$, we first decompose the graph $\mathcal{G}(\mathcal{E}_m^p)$ into its connected components such that $\mathcal{G}(\mathcal{E}_m^p) = \mathcal{I} \sqcup \mathcal{A} \sqcup \mathcal{B} \sqcup \mathcal{C}$, where

$$\begin{aligned}
\mathcal{A} &= \mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_{c_1} \sqcup \mathcal{A}_{c_1+1} \sqcup \cdots \sqcup \mathcal{A}_{c_1+c_2}\ , \\
\mathcal{B} &= \mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_{c_3} \sqcup \mathcal{B}_{c_3+1} \sqcup \cdots \sqcup \mathcal{B}_{c_3+c_4}\ , \\
\mathcal{C} &= \mathcal{C}_1 \sqcup \cdots \sqcup \mathcal{C}_{c_5}\ ,
\end{aligned}$$

for some $c_1, c_2, c_3, c_4, c_5 \geq 0$. Here $\mathcal{I}$ is the union of isolated colliding vertices. $\mathcal{A}$ is the union of colliding components, where $\mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_{c_1}$ is the union of colliding components with a colliding v vertex; and $\mathcal{A}_{c_1+1} \sqcup \cdots \sqcup \mathcal{A}_{c_1+c_2}$ is the union of colliding components with a colliding y vertex. $\mathcal{B}$ is the union of the remaining star components (that are not colliding components), where $\mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_{c_3}$ is the union of v-$\star$ components, and $\mathcal{B}_{c_3+1} \sqcup \cdots \sqcup \mathcal{B}_{c_3+c_4}$ is the union of y-$\star$ components. $\mathcal{C}$ is the union of the remaining isolated components (that are not colliding components).

Let q_1, q_2, q_3, q_4, and q_5 denote the number of equations (edges) in $\mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_{c_1}$, $\mathcal{A}_{c_1+1} \sqcup \cdots \sqcup \mathcal{A}_{c_1+c_2}$, $\mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_{c_3}$, $\mathcal{B}_{c_3+1} \sqcup \cdots \sqcup \mathcal{B}_{c_3+c_4}$, and $\mathcal{C}$, respectively. Therefore, we have $c_5 = q_5$. Note that the equations in $\mathcal{E}_m^p$ can be arranged in any arbitrary order without affecting the number of solutions. For the sake of simplicity, we fix the ordering in such a way that the union $\mathcal{I}$ comes first, followed by $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. Now, our goal is to give a lower bound on the number of solutions of $\mathcal{E}_m^p$.

Theorem 1. *For positive integers q_m and p, let $\mathcal{G}(\mathcal{E}_m^p) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S})$ be a good graph as described above such that $|\mathcal{S}| = q_m$. Assume that $p + q_m \leq 2^{n-2}$ and $\xi_{\max} \cdot (p + q_m) \leq 2^{n-1}$, let $h(\mathcal{G}(\mathcal{E}_m^p))$ denote the number of solutions to $\mathcal{G}(\mathcal{E}_m^p)$.*

(a) For settings such as (t)sprp, we have

$$\frac{h(\mathcal{G}(\mathcal{E}_m^p)) \prod_{\lambda' \in \hat{\lambda}^{q_m}} (2^n)_{\mu(\lambda^{*q_m}, \lambda')}}{(2^n - p)_{q_2+c_3+q_4+q_5} (2^n - p)_{q_1+q_3+c_4+q_5}} \geq 1 - \frac{\sum_{i=1}^{c_1+c_2} q_m}{2^n} - \frac{3q_m^3}{2^{2n}} - \frac{2(p+q_m)^2}{2^{2n}} \Big(\eta + q_m \Big)\,.$$

(b) For settings such as prf, weak prf, (t)ccr, we have

$$\frac{h(\mathcal{G}(\mathcal{E}_m^p)) 2^{nq_m}}{(2^n - p)_{q_2+c_3+q_4+q_5} (2^n - p)_{q_1+q_3+c_4+q_5}} \geq 1 - \frac{2(p+q_m)^2}{2^{2n}} \Big(\eta + q_m \Big)\,.$$

where $\eta = \sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1}(\eta_{i+1}^2 + \eta_{i+1})$*,* $\eta_j = \xi_j - 1$ *and* ξ_j *denotes the size (number of vertices) of the* j*-th component, for all* $j \in [c_1 + c_2 + c_3 + c_4 + c_5]$*.*

Proof. The proof is given in the full version of the paper. □

We will illustrate Theorem 1 (a) to tweakable block ciphers in Section 5, and Theorem 1 (b) to PRFs in Section 6. Looking back at the discussion given towards the end of Section 2.5, one can see the motivation behind the difference between the expressions given in Theorem 1 (a) and Theorem 1 (b).

For the second case, the system $\mathcal{E}_a$ contains at least one non-equation. Here we will focus on a graph $\mathcal{G}(\mathcal{E}^p)$ such that its subgraph $\mathcal{G}(\mathcal{E}_m^p)$ satisfies the (i) acyclic, (ii) non-degeneracy, (iii) ξ-block-maximality, (v) NCL, and (iv) SCV properties (Definition 1-5). In addition, $\mathcal{G}(\mathcal{E})$ also need to satisfy the following property.

Definition 6 (non-zero distance label (NDL)). *There are no* λ'*-labeled edges that connect two vertices* v *and* y *from two colliding components* $\mathcal{I}_1$ *and* $\mathcal{I}_2$*, where the distance between* v *and* y *(defined as* $v \oplus y$*) is* λ'*, for* $v \in \mathcal{V}'^p$ *and* $y \in \mathcal{Y}'^p$*.*

Note that if there is a λ'-labeled non-equations between vertices v and y of two different colliding components, then it means $v \oplus y \neq \lambda'$. However, for any colliding component, the values of all vertices in this component are uniquely defined. If the distance $v \oplus y$ is equal to λ', then we will have a contradiction with the non-equation. Definition 6 actually excludes this situation. We call any graph that satisfies these six properties a *good* graph. Given a good graph $\mathcal{G}(\mathcal{E}^p) = (\mathcal{V}^p, \mathcal{Y}^p, \mathcal{S} \sqcup \mathcal{S}')$, a solution to $\mathcal{G}(\mathcal{E}^p)$ is an assignment of distinct values to the v vertices in $\mathcal{V}^p$ and distinct values to the y vertices in $\mathcal{Y}^p$ satisfying all λ-labeled equations and λ'-labeled non-equations.

We consider a system $\mathcal{E}^p$ with its corresponding graph $\mathcal{G}(\mathcal{E}^p)$ such that each component in the subgraph $\mathcal{G}(\mathcal{E}_m^p)$ is a star, an isolated edge, or an isolated vertex. In order to find the number of solutions to $\mathcal{G}(\mathcal{E}^p)$, we first decompose the subgraph $\mathcal{G}(\mathcal{E}_m^p)$ into its connected components such that $\mathcal{G}(\mathcal{E}_m^p) = \mathcal{I} \sqcup \mathcal{A} \sqcup \mathcal{B} \sqcup \mathcal{C} \sqcup \mathcal{D}$, with $\mathcal{I}, \mathcal{A}, \mathcal{B}$, and $\mathcal{C}$ the union of components defined before. Here, $\mathcal{D}$ is the union of isolated vertices in the subgraph $\mathcal{G}(\mathcal{E}_m^p)$ that are not colliding vertices, note that these vertices are connected with λ'-labeled edges in the subgraph $\mathcal{G}(\mathcal{E}_a)$. Let c_6 be the number of such isolated v vertices, and c_7 be the number of such isolated y vertices. Again, we fix the ordering in such a way that the union $\mathcal{I}$ comes first, followed by $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$. Now, our goal is to give a lower bound on the number of solutions of $\mathcal{E}^p$.

Theorem 2. *For positive integers* q_m*,* q_a *and* p*, let* $\mathcal{G}(\mathcal{E}^p) = (\mathcal{V}'^p, \mathcal{Y}'^p, \mathcal{S} \sqcup \mathcal{S}')$ *be a good graph as described above such that* $|\mathcal{S}| = q_m$*,* $|\mathcal{S}'| = q_a$*. Assume that* $p + q_m \leq 2^{n-2}$ *and* $\xi_{\max} \cdot (p + q_m) \leq 2^{n-1}$*, let* $h(\mathcal{G}(\mathcal{E}^p))$ *denote the number of solutions to* $\mathcal{G}(\mathcal{E}^p)$*.*

(a) For settings such as (t)sprp with non-equations, we have

$$\frac{h(\mathcal{G}(\mathcal{E}^p))\prod_{\lambda'\in\hat{\lambda}^{q_m}}(2^n)_{\mu(\lambda^{*q_m},\lambda')}}{(2^n-p)_{q_2+c_3+q_4+q_5+c_6}(2^n-p)_{q_1+q_3+c_4+q_5+c_7}} \geq 1-\frac{\sum_{i=1}^{c_1+c_2} q_m}{2^n}-\frac{3q_m^3}{2^{2n}} \\ -\frac{2(p+q_m)^2}{2^{2n}}\Big(\eta+q_m\Big)-\frac{2q_a}{2^n}\,.$$

(b) For settings such as mac (prf with non-equations), we have

$$\frac{h(\mathcal{G}(\mathcal{E}^p))2^{nq_m}}{(2^n-p)_{q_2+c_3+q_4+q_5+c_6}(2^n-p)_{q_1+q_3+c_4+q_5+c_7}} \geq 1-\frac{2(p+q_m)^2}{2^{2n}}\Big(\eta+q_m\Big)-\frac{2q_a}{2^n}\,.$$

where $\eta = \sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1}(\eta_{i+1}^2+\eta_{i+1})$, $\eta_j = \xi_j - 1$ and ξ_j denotes the size (number of vertices) of the j-th component, for all $j \in [c_1+c_2+c_3+c_4+c_5]$.

Proof. The proof is given in the full version of the paper. □

We will illustrate Theorem 2 (b) to nonce-based MAC algorithms in Section 7.

4 A Framework for Security Proof Using Public Permutation Mirror Theory

The goal of this section is to establish a general framework for (multi-user) security proof using Theorem 1-2. Note that a framework for specific security notions such as sprp, tsprp, prf, mac, ... can be derived directly from this framework. We consider an algorithm F which is built on two independent public permutations with the following special structure.

Let $n, s, t \in \mathbb{N}$, and let $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$. One can consider the generic construction $F^{\pi_1,\pi_2}\colon \mathcal{K}\times\mathcal{I}_1\times\cdots\times\mathcal{I}_s \to \mathcal{R}_1\times\cdots\times\mathcal{R}_t$ based on π_1 and π_2, where $\mathcal{K}$ is the key space, $\mathcal{I}_1\times\cdots\times\mathcal{I}_s$ are the input spaces, and $\mathcal{R}_1\times\cdots\times\mathcal{R}_t$ are the output spaces. Note that here F can be a tweakable block cipher, a PRF, a MAC algorithm, etc. In this work, we will focus on algorithms that can be viewed as the xor of the public permutations

$$Z = \pi_1(A) \oplus \pi_2(B)\,,$$

for $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$. Here A, B, and Z are functions of the secret key K, the inputs $I_1, \ldots, I_s$, and the outputs $R_1, \ldots, R_t$. Although there are no strict restrictions for Z, we do require that the equality patterns of A and B satisfy certain conditions. More precisely, equality pattern of B should not depend on the value of $\pi_1(A)$ and vice versa. This is the condition to use the mirror theory based lower bound as formalized in [32].

4.1 General Setting and Transcripts.

Let $u \in \mathbb{N}$, $K_1, \ldots, K_u \xleftarrow{\$} \mathcal{K}$, and $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$. Consider any distinguisher $\mathcal{D}$ that has access to the oracles: $(\mathcal{O}_m, \mathcal{O}_a, \pi_1^{\pm}, \pi_2^{\pm})$ in the real world or $(\mathcal{P}_m, \mathcal{P}_a, \pi_1^{\pm}, \pi_2^{\pm})$ in the ideal world. Here we have $\mathcal{O}_m = (F_{K_1}^{\pi_1,\pi_2}, \ldots, F_{K_u}^{\pi_1,\pi_2})$, and $\mathcal{O}_a$ is the possible set of verification oracles (for example the case of mac notion in Section 2.3). The oracle $\mathcal{P}_m$ is the idealized version of $(F_{K_1}^{\pi_1,\pi_2}, \ldots, F_{K_u}^{\pi_1,\pi_2})$, which these idealized oracles are depend on the considered security notion (see Section 2.1-2.3 for more details), and $\mathcal{P}_a$ is the possible set of rejection oracles. We require that $\mathcal{D}$ is computationally unbounded and deterministic. For user index $j \in \{1, \ldots, u\}$, $\mathcal{D}$ makes q_m queries to $\mathcal{O}_m$ or $\mathcal{P}_m$, and these are summarized in a transcript

$$\tau_m = \{(j^{(1)}, I_1^{(1)}, \ldots, I_s^{(1)}, R_1^{(1)}, \ldots, R_t^{(1)}), \ldots, \\ (j^{(q_m)}, I_1^{(q_m)}, \ldots, I_s^{(q_m)}, R_1^{(q_m)}, \ldots, R_t^{(q_m)})\},$$

and q_a queries to $\mathcal{O}_a$ or $\mathcal{P}_a$, these are summarized in a transcript

$$\tau_a = \{(j'^{(1)}, I_1'^{(1)}, \ldots, I_s'^{(1)}, R_1'^{(1)}, \ldots, R_t'^{(1)}, b'^{(1)}), \ldots, \\ (j'^{(q_a)}, I_1'^{(q_a)}, \ldots, I_s'^{(q_a)}, R_1'^{(q_a)}, \ldots, R_t'^{(q_a)}, b'^{(q_a)})\}.$$

Note that τ_a is empty for notions where no verification oracles are considered (such as sprp, tsprp, prf, tccr, etc). $\mathcal{D}$ also makes p primitive queries to $\pi_1^{\pm}$ and p primitive queries to $\pi_2^{\pm}$, and like before, these are respectively summarized in transcripts τ_1 and τ_2. We assume that τ_m, τ_a, τ_1, and τ_2 do not contain duplicate elements. After $\mathcal{D}$'s interaction with the oracles, but before it outputs its decision, we disclose the keys $K_1, \ldots, K_u$ to the distinguisher. In the real world, these are the keys used in the construction. In the ideal world, $K_1, \ldots, K_u$ are dummy keys that are drawn uniformly at random. The complete view is denoted $\tau = (\tau_m, \tau_a, \tau_1, \tau_2, K_1, \ldots, K_u)$.

4.2 Attainable Index Mappings.

In the real world, each query $(j^{(i)}, I_1^{(i)}, \ldots, I_s^{(i)}, R_1^{(i)}, \ldots, R_t^{(i)}) \in \tau_m$ corresponds to an evaluation of the $j^{(i)}$-th oracle in $\mathcal{O}_m$, each query $(j'^{(a)}, I_1'^{(a)}, \ldots, I_s'^{(a)}, R_1'^{(a)}, \ldots, R_t'^{(a)}, b'^{(a)}) \in \tau_a$ corresponds to an evaluation of the $j'^{(a)}$-th oracle in $\mathcal{O}_a$, and each primitive query $(u, v) \in \tau_1$ (resp., $(x, y) \in \tau_2$) corresponds to an evaluation of the primitive oracle $\pi_1^{\pm}$ (resp., $\pi_2^{\pm}$). Note that each algorithm F consists of an evaluation of π_1 and an evaluation of π_2. For the queries in τ_m, these are of the form $A^{(i)} \mapsto \pi_1(A^{(i)})$ and $B^{(i)} \mapsto \pi_2(B^{(i)})$ such that $\pi_1(A^{(i)}) \oplus \pi_2(B^{(i)}) = Z^{(i)}$. Likewise, for the queries in τ_a, there are evaluations $A'^{(a)} \mapsto \pi_1(A'^{(a)})$ and $B'^{(a)} \mapsto \pi_2(B'^{(a)})$, such that $\pi_1(A'^{(a)}) \oplus \pi_2(B'^{(a)}) \neq Z'^{(a)}$. The values of $A^{(i)}, B^{(i)}, Z^{(i)}$ and $A'^{(a)}, B'^{(a)}, Z'^{(a)}$ are specific for the particular construction, and can be deduced from τ. Without loss of generality, we assume that all primitive queries are made in the forward direction, then these are of the form $u \mapsto \pi_1(u)$ or $x \mapsto \pi_2(x)$ such that $\pi_1(u) = v$ and $\pi_2(x) = y$. The transcript τ defines $q_m + 2p$ equations and q_a non-equations on the unknowns, and these (non)-equations are

$$\mathcal{E}_m^p = \begin{cases} \pi_1(A^{(1)}) \oplus \pi_2(B^{(1)}) = Z^{(1)}, \\ \vdots \\ \pi_1(A^{(q_m)}) \oplus \pi_2(B^{(q_m)}) = Z^{(q_m)}, \\ \pi_1(u) = v \quad \text{for } (u,v) \in \tau_1, \\ \pi_2(x) = y \quad \text{for } (x,y) \in \tau_2, \end{cases} \mathcal{E}_a = \begin{cases} \pi_1(A'^{(1)}) \oplus \pi_2(B'^{(1)}) \neq Z'^{(1)}, \\ \vdots \\ \pi_1(A'^{(q_a)}) \oplus \pi_2(B'^{(q_a)}) \neq Z'^{(q_a)}. \end{cases} \tag{3}$$

In the above q_m+2p equations, some of the unknowns may be equal to each other. We have that $\pi_1(A^{(i)}) \neq \pi_1(A^{(k)})$ if and only if $A^{(i)} \neq A^{(k)}$, and $\pi_2(B^{(i)}) \neq \pi_2(B^{(k)})$ if and only if $B^{(i)} \neq B^{(k)}$. No condition holds for $\pi_1(A^{(i)})$ versus $\pi_2(B^{(i)})$, as these are defined by independent permutations. The same holds for verification queries and primitive queries. However, no a priori condition holds for (non-)equality between values $\pi_1(A^{(i)})$ versus $\pi_1(A'^{(a)})$ versus $\pi_1(u)$, and $\pi_2(B^{(i)})$ versus $\pi_2(B'^{(a)})$ versus $\pi_2(x)$.

Thus,

$$\{\pi_1(A^{(i)})\}_{1 \le i \le q_m} \cup \{\pi_1(A'^{(a)})\}_{1 \le a \le q_a} \cup \{\pi_1(u)\}_{(u,v) \in \tau_1} ,$$
$$\{\pi_2(B^{(i)})\}_{1 \le i \le q_m} \cup \{\pi_2(B'^{(a)})\}_{1 \le a \le q_a} \cup \{\pi_2(x)\}_{(x,y) \in \tau_2} ,$$

are identified with two sets of unknowns $\mathcal{V}'^p = \{v_1, \ldots, v_{q^p_{V'}}\}$ and $\mathcal{Y}'^p = \{y_1, \ldots, y_{q^p_{Y'}}\}$, with $q^p_{V'}, q^p_{Y'} \le q_m + q_a + p$. Since $\mathcal{V}'^p$ and $\mathcal{Y}'^p$ are defined by independent permutations, we know that $\mathcal{V}'^p$ and $\mathcal{Y}'^p$ are disjoint. We also know that

$$V_0 = \{\pi_1(u) \mid (u,v) \in \tau_1\} , \qquad Y_0 = \{\pi_2(x) \mid (x,y) \in \tau_2\} .$$

are already well defined by the system. Hence the only unknowns that are left are in the sets $\mathcal{V}'^p \setminus V_0$ and $\mathcal{Y}'^p \setminus Y_0$. For $v_s \in \mathcal{V}'^p$ and $y_t \in \mathcal{Y}'^p$, we connect v_s and y_t with a λ-labeled edge of label $Z^{(i)}$ if $\pi_1(A^{(i)}) = v_s$ and $\pi_2(B^{(i)}) = y_t$ for some $i \in [q_m]$. Similarly, we connect v_s and y_t with a λ'-labeled edge of label $Z'^{(a)}$ if $\pi_1(A'^{(a)}) = v_s$ and $\pi_2(B'^{(a)}) = y_t$ for some $a \in [q_a]$. Finally, v_s (resp., y_t) represents an isolated colliding vertex if it is not connected with an edge, for these vertices we have $\pi_1(u) = v_s$ (resp., $\pi_2(v) = y_t$) for $(u, v_s) \in \tau_1$ and $(x, y_t) \in \tau_2$. In this way, we obtain a graph on $\mathcal{V}'^p \sqcup \mathcal{Y}'^p$, called the transcript graph of τ, and we denote it by $\mathcal{G}_\tau(\mathcal{E}^p)$.

4.3 Bad Transcripts.

Informally, bad events are the properties which would make the public permutation extended mirror theory inapplicable. One can only apply the mirror theory if $\mathcal{G}_\tau(\mathcal{E}^p)$ is (1). acyclic, (2). satisfies the non-degeneracy condition, (3). satisfies the NCL condition, (4). satisfies the SCV condition, (5). satisfies the NDL condition, and (6). is $(\xi+1)$-block-maximal. For some parameter ξ that will be defined later on, we say a system of equations is $(\xi+1)$-block-maximal if it does not contain a $(\xi+1)$-block-collision, which means that neither of the two permutations evaluates the same input more than ξ times. As our security analysis will cap on $2n/3$-bit security only, we can keep it simple by introducing an event that excludes all alternating paths of length 3, and events that exclude all v-$\star$ component with a y-colliding vertex and y-$\star$ component with a v-colliding vertex. Below, we will give a formal description of the bad events.

For simplicity, we denote by $A^{(i)}$ the i-th input to π_1, $B^{(i)}$ the i-th input to π_2, and $Z^{(i)} = \pi_1(A^{(i)}) \oplus \pi_2(B^{(i)})$ for the MAC queries. Similarly we have $A'^{(a)}$, $B'^{(a)}$ and $Z'^{(a)}$ for the verification queries. Given a parameter $\xi \in \mathbb{N}$, we say that $\tau \in \mathcal{T}_{\text{bad}}$ if and only if one of the following conditions holds:

(i) A component with two colliding vertices.

$$\begin{aligned}
&\exists i \in [q_m], (u,v) \in \tau_1, (x,y) \in \tau_2 \,\text{such that}\, A^{(i)} = u \,\wedge\, B^{(i)} = x\,,\\
&\exists i \in [q_m], (u,v) \in \tau_1, (x,y) \in \tau_2 \,\text{such that}\, A^{(i)} = u \,\wedge\, Z^{(i)} = v \oplus y\,,\\
&\exists i \in [q_m], (u,v) \in \tau_1, (x,y) \in \tau_2 \,\text{such that}\, Z^{(i)} = v \oplus y \,\wedge\, B^{(i)} = x\,.
\end{aligned}$$

(ii) An alternating path of length 3.

$$\exists i \neq k, k \neq l \in [q_m] \,\text{such that}\, A^{(i)} = A^{(k)} \,\wedge\, B^{(k)} = B^{(l)}\,.$$

(iii) An alternating path of length 2 such that $\lambda(\mathcal{L}) = 0$.

$$\begin{aligned}
&\exists i \neq k \in [q_m] \,\text{such that}\, A^{(i)} = A^{(k)} \,\wedge\, Z^{(i)} = Z^{(k)}\,,\\
&\exists i \neq k \in [q_m] \,\text{such that}\, Z^{(i)} = Z^{(k)} \,\wedge\, B^{(i)} = B^{(k)}\,.
\end{aligned}$$

(iv) A v-$\star$ colliding component with y-colliding vertices, or a y-$\star$ colliding component with v-colliding vertices.

$$\begin{aligned}
&\exists i \neq k \in [q_m], (u,v) \in \tau_1 \,\text{such that}\, A^{(i)} = u \,\wedge\, B^{(i)} = B^{(k)}\,,\\
&\exists i \neq k \in [q_m], (x,y) \in \tau_2 \,\text{such that}\, B^{(i)} = x \,\wedge\, A^{(i)} = A^{(k)}\,,\\
&\exists i \neq k \in [q_m], (u,v), (u',v') \in \tau_1 \,\text{such that}\\
&\qquad A^{(i)} = u \,\wedge\, A^{(k)} = u' \,\wedge\, v \oplus Z^{(i)} = v' \oplus Z^{(k)}\,,\\
&\exists i \neq k \in [q_m], (x,y), (x',y') \in \tau_2 \,\text{such that}\\
&\qquad B^{(i)} = x \,\wedge\, B^{(k)} = x' \,\wedge\, y \oplus Z^{(i)} = y' \oplus Z^{(k)}\,.
\end{aligned}$$

(v) A $(\xi+1)$-block-collision.

$$\begin{aligned}
&\exists i_1, \ldots, i_{\xi+1} \in \{1, \ldots, q_m\} \,\text{such that}\, A^{(1)} = \cdots = A^{(\xi+1)}\,,\\
&\exists i_1, \ldots, i_{\xi+1} \in \{1, \ldots, q_m\} \,\text{such that}\, B^{(1)} = \cdots = B^{(\xi+1)}\,.
\end{aligned}$$

(vi) An alternating circle of length 2 with a λ'-labeled edge.

$$\exists i \in [q_m], a \in [q_a] \,\text{such that}\, A^{(i)} = A'^{(a)} \,\wedge\, B^{(i)} = B'^{(a)} \,\wedge\, Z^{(i)} = Z'^{(a)}\,.$$

(vii) A λ'-labeled edge between two vertices with distance λ'.

$$\exists a \in [q_a], (u,v) \in \tau_1, (x,y) \in \tau_2 \text{ such that}$$
$$A'^{(a)} = u \wedge B'^{(a)} = x \wedge Z'^{(a)} = v \oplus y\,,$$
$$\exists i \in [q_m], a \in [q_a], (u,v), (u'v') \in \tau_1 \text{ such that}$$
$$A^{(i)} = u \wedge B^{(i)} = B'^{(a)} \wedge A'^{(a)} = u' \wedge Z'^{(a)} = v \oplus Z^{(i)} \oplus v'\,,$$
$$\exists i \in [q_m], a \in [q_a], (x,y), (x',y') \in \tau_2 \text{ such that}$$
$$B^{(i)} = x \wedge A^{(i)} = A'^{(a)} \wedge B'^{(a)} = x' \wedge Z'^{(a)} = y \oplus Z^{(i)} \oplus y'\,,$$
$$\exists i \neq k \in [q_m], a \in [q_a], (u,v) \in \tau_1, (x,y) \in \tau_2 \text{ such that } A^{(i)} = u$$
$$\wedge\, B^{(k)} = x \wedge B^{(i)} = B'^{(a)} \wedge A^{(k)} = A'^{(a)} \wedge Z'^{(a)} = v \oplus Z^{(i)} \oplus y \oplus Z^{(k)}\,.$$

Note that by (ii) and (iv), we will end up with a graph that contains only isolated and v-$\star$ colliding components with a v-colliding vertex, isolated and y-$\star$ colliding components with a y-colliding vertex, v-$\star$ components, y-$\star$ components, isolated components, and isolated vertices. The resulting graph is good since it

1. satisfies the SCV condition by conditions (i), (ii), and (iv),
2. acyclic by conditions (ii),
3. satisfies the non-degeneracy condition by conditions (ii) and (iii),
4. is $(\xi+1)$-block-maximal by conditions (ii) and (v),
5. satisfies the NCL condition by conditions (ii) and (vi),
6. satisfies the NDL condition by conditions (ii), (iv), and (vii).

The probability that $\tau \in \mathcal{T}_{\text{bad}}$ happens, is given by the sum of the probabilities that each of the above mentioned bad events happens. When the above mentioned events can be excluded in the transcript, then $\mathcal{G}_\tau$ forms a good transcript graph for $\tau \in \mathcal{T}_{\text{good}}$.

4.4 Ratio for Good Transcripts.

Once bad transcripts have been defined, we will show that

$$\Pr[X_{\mathcal{P}} \in \mathcal{T}_{\text{bad}}] \leq \varepsilon_{\text{bad}}\,,$$

for a small $\varepsilon_{\text{bad}} > 0$. Next, we fix a good transcript τ. According to (2), we only have to consider $\rho(\tau)/\Pr[\mathcal{P}_m \colon \mathcal{P}_m \vdash \tau_m \wedge \mathcal{P}_a \vdash \tau_a]$, with

$$\rho(\tau) = \Pr\left[\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n) \colon \mathcal{O}_m \vdash \tau_m \wedge \mathcal{O}_a \vdash \tau_a \mid \pi_1 \vdash \tau_1 \wedge \pi_2 \vdash \tau_2\right].$$

This is exactly the ratio given by Theorem 1 and 2. From (2), we obtain

$$\frac{\Pr[X_{\mathcal{O}} = \tau]}{\Pr[X_{\mathcal{P}} = \tau]} \geq 1 - \varepsilon_{\text{ratio}}\,,$$

and by Lemma 3, we have

$$\mathbf{Adv}^{\text{mac}}_{\text{MAC}}(\mathcal{D}) \leq \boldsymbol{E}[\varepsilon_{\text{ratio}}] + \varepsilon_{\text{bad}}\,.$$

5 Multi-User Security of Tweakable Even-Mansour Cipher

In this section we consider the 2-round Tweakable Even-Mansour (TEM) construction that was proposed by Cogliati et al. [15]. They showed that 2-round TEM achieves $2n/3$-bit security in the single-user setting. Here we show that same level of security can be achieved in the *multi-user* setting using the technique proposed in this work.

Let $n \in \mathbb{N}$, let $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$, and let $\mathcal{H}$ be an ϵ-AXU function family. One can consider a generic construction $\mathrm{TEM}\colon \mathcal{H}^2 \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ as

$$\mathrm{TEM}(h_1, h_2, T, M) = \pi_2(\pi_1(M \oplus h_1(T)) \oplus h_1(T) \oplus h_2(T)) \oplus h_2(T)\,. \qquad (4)$$

The security of TEM based on π_1 and π_2 is given in the following Theorem.

Theorem 3. *Let $n \in \mathbb{N}$ and let $\mathcal{H}$ be a uniform ϵ-AXU family of functions from $\mathcal{T}$ to $\{0,1\}^n$. We consider $\mathrm{TEM}\colon \mathcal{H}^2 \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ based on two permutations $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$ and u pairs of uniform user hash keys $(h_1^1, h_2^1), \ldots, (h_1^u, h_2^u) \xleftarrow{\$} \mathcal{H}^2$. For any distinguisher $\mathcal{D}$ making at most q construction queries distributed over its u construction oracles, at most p primitive queries to $\pi_1^{\pm}$ and p primitive queries to $\pi_2^{\pm}$, we have*

$$\mathbf{Adv}^{\text{mu-tsprp}}_{\mathrm{TEM}}(\mathcal{D}) \le 3q^3\epsilon^2 + q^2p\epsilon^2 + 6\sqrt{q}p\epsilon + \frac{6q^{3/2}}{2^n} + \frac{2q(p+q)^2}{2^{2n}}\Big(1 + 13q\epsilon\Big)\,.$$

Proof. Let $(h_1^1, h_2^1), \ldots, (h_1^u, h_2^u) \xleftarrow{\$} \mathcal{H}^2$, $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$, and $\tilde{\pi}_1, \ldots, \tilde{\pi}_u \xleftarrow{\$} \widetilde{\mathrm{Perm}}(t, n)$. Here, we consider any distinguisher $\mathcal{D}$ that has access to either $(\mathrm{TEM}^{\pi_1,\pi_2^{-1}}_{h_1^1,h_2^1}, \ldots, \mathrm{TEM}^{\pi_1,\pi_2^{-1}}_{h_1^u,h_2^u}, \pi_1, \pi_2)$ in the real world, or $(\tilde{\pi}_1, \ldots, \tilde{\pi}_u, \pi_1, \pi_2)$ in the ideal world. The security proof relies on Theorem 1 (a), although this application is not straightforward. Most importantly, we consider $(\mathrm{TEM}^{\pi_1,\pi_2^{-1}}_{h_1^j,h_2^j})_{j=1}^u$ instead of $(\mathrm{TEM}^{\pi_1,\pi_2}_{h_1^j,h_2^j})_{j=1}^u$. As π_1, π_2 are drawn independently, these two constructions are provably equally secure. We can view an evaluation $C = \mathrm{TEM}^{\pi_1,\pi_2^{-1}}_{h_1^j,h_2^j}(T, M)$ as the xor of two public permutations in the middle of the function, $\pi_1(M \oplus h_1^j(T)) \oplus \pi_2(M \oplus h_2^j(T)) = h_1^j(T) \oplus h_2^j(T)$. Therefore, q evaluations of $\mathrm{TEM}^{\pi_1,\pi_2^{-1}}_{h_1^j,h_2^j}$ can be translated to a system of q bi-variate affine equations. Including $2p$ univariate affine equations that are defined by the primitive queries, these equations can be written in the form (3).

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}} \in \boldsymbol{\mathcal{T}_{\mathrm{bad}}}]$. Following the framework given in Section 4, we first perform the bad transcripts analysis. By replacing $A = M \oplus h_1^j(T)$, $B = M \oplus h_2^j(T)$, and $Z = h_1^j(T) \oplus h_2^j(T)$ in the framework of Section 4.3, we get the following bad events. Given a parameter $\xi \in \mathbb{N}$, we say that $\tau \in \mathcal{T}_{\mathrm{bad}}$ if and only if there exist construction queries $(j, T, M, C), (j', T', M', C'), (j'', T'', M'', C'') \in \tau_m$, primitive queries $(u, v), (u', v') \in \tau_1$ and $(x, y), (x', y') \in \tau_2$ such that one of the following conditions holds:

(i) A component with two colliding vertices.

$$\begin{aligned}
&\mathrm{bad}_1\colon M \oplus u = h_1^j(T) \,\wedge\, C \oplus x = h_2^j(T)\,,\\
&\mathrm{bad}_2\colon M \oplus u = h_1^j(T) \,\wedge\, v \oplus y = h_1^j(T) \oplus h_2^j(T)\,,\\
&\mathrm{bad}_3\colon v \oplus y = h_1^j(T) \oplus h_2^j(T) \,\wedge\, C \oplus x = h_2^j(T)\,.
\end{aligned}$$

(ii) An alternating path of length 3.

$$\mathrm{bad}_4\colon M \oplus h_1^j(T) = M' \oplus h_1^{j'}(T') \,\wedge\, C' \oplus h_2^{j'}(T') = C'' \oplus h_2^{j''}(T'')\,.$$

(iii) An alternating path of length 2 such that $\lambda(\mathcal{L}) = 0$.

$$\begin{aligned}
&\mathrm{bad}_5\colon M \oplus h_1^j(T) = M' \oplus h_1^{j'}(T') \,\wedge\, h_1^j(T) \oplus h_2^j(T) = h_1^{j'}(T') \oplus h_2^{j'}(T')\,,\\
&\mathrm{bad}_6\colon h_1^j(T) \oplus h_2^j(T) = h_1^{j'}(T') \oplus h_2^{j'}(T') \,\wedge\, C \oplus h_2^j(T) = C' \oplus h_2^{j'}(T')\,.
\end{aligned}$$

(iv) A v-$\star$ colliding component with y-colliding vertices, or a y-$\star$ colliding component with v-colliding vertices .

$$\begin{aligned}
\mathrm{bad}_7\colon & M \oplus u = h_1^j(T) \,\wedge\, C \oplus h_2^j(T) = C' \oplus h_2^{j'}(T)\,,\\
\mathrm{bad}_8\colon & C \oplus x = h_2^j(T) \,\wedge\, M \oplus h_1^j(T) = M' \oplus h_1^{j'}(T')\,,\\
\mathrm{bad}_9\colon & M \oplus u = h_1^j(T) \,\wedge\, M' \oplus u' = h_1^{j'}(T')\\
& \wedge\, v \oplus h_1^j(T) \oplus h_2^j(T) = v' \oplus h_1^{j'}(T') \oplus h_2^{j'}(T')\,,\\
\mathrm{bad}_{10}\colon & C \oplus x = h_2^j(T) \,\wedge\, C' \oplus x' = h_2^{j'}(T)\\
& \wedge\, y \oplus h_1^j(T) \oplus h_2^j(T) = y' \oplus h_1^{j'}(T') \oplus h_2^{j'}(T')\,.
\end{aligned}$$

(v) A $(\xi+1)$-block-collision.

$$\begin{aligned}
&\mathrm{bad}_{11}\colon \{i_1,\ldots,i_{\xi+1}\} \in [q]\,\text{such that}\, M_{i_1} \oplus h_1^{j_{i_1}}(T_{i_1}) = \cdots = M_{i_{\xi+1}} \oplus h_1^{j_{\xi+1}}(T_{i_{\xi+1}})\,,\\
&\mathrm{bad}_{12}\colon \{i_1,\ldots,i_{\xi+1}\} \in [q]\,\text{such that}\, C_{i_1} \oplus h_2^{j_{i_1}}(T_{i_1}) = \cdots = C_{i_{\xi+1}} \oplus h_2^{j_{\xi+1}}(T_{i_{\xi+1}})\,.
\end{aligned}$$

Since there is a $\sum_{i=1}^{c_1+c_2} q/2^n$ term in Theorem 1 (a), and we want to get $2n/3$-bits security, we also need the following two bad events

$$\begin{aligned}
&\mathrm{bad}_{c_1}: c_1 = |(j,T,M,C) \in \tau_m \colon M \oplus h_1^j(T) \in \tau_1| \geq \sqrt{q}\,,\\
&\mathrm{bad}_{c_2}: c_2 = |(j,T,M,C) \in \tau_m \colon C \oplus h_2^j(T) \in \tau_2| \geq \sqrt{q}\,.
\end{aligned}$$

Lemma 4. *For any integers q and p, one has*

$$\Pr[\tau \in \mathcal{T}_{\mathrm{bad}}] \leq 4qp^2\epsilon^2 + 3q^3\epsilon^2 + q^2p\epsilon^2 + \frac{q^3}{2^{2n}} + 2\sqrt{q}p\epsilon + \frac{16q^2(p+q)^2\epsilon}{2^{2n}}\,.$$

The proof of the lemma is given in the full version of the paper.

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{O}}} = \boldsymbol{\tau}] / \mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}} = \boldsymbol{\tau}]$. The next step is the calculate the ratio for good transcripts. Note that by $\neg\text{bad}_{c_1}$ and $\neg\text{bad}_{c_2}$, we have $\sum_{i=1}^{c_1+c_2} q \leq 2q^{3/2}$. We use Theorem 1 (a) to get

$$\epsilon_{\text{ratio}} \leq \frac{2q^{3/2}}{2^n} + \frac{3q^3}{2^{2n}} + \frac{2(p+q)^2 \sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1} (\eta_{i+1}^2 + \eta_{i+1})}{2^{2n}} + \frac{2q(p+q)^2}{2^{2n}} .$$

Let $\sim_1$ (resp., $\sim_2$) be an equivalence relation on $[q]$ as $\alpha \sim_1 \beta$ (resp., $\alpha \sim_2 \beta$) if and only if $A_\alpha = A_\beta$ (resp. $B_\alpha = B_\beta$). Now, each η_i random variable denotes the cardinality of some non-singleton equivalence class of $[q]$ with respect to either $\sim_1$ or $\sim_2$. For $r, s \in \mathbb{N}$, we denote by $\mathcal{P}_1^1, \ldots, \mathcal{P}_r^1$ and $\mathcal{P}_1^2, \ldots, \mathcal{P}_s^2$ the non-singleton equivalence classes of $[q]$ with respect to $\sim_1$ and $\sim_2$, respectively. Further, for $k \in [r]$ and $l \in [s]$, let $\nu_k = |\mathcal{P}_k^1|$ and $\nu'_l = |\mathcal{P}_l^2|$. Then, we have

$$\begin{aligned} \boldsymbol{E}\left[\sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1} (\eta_{i+1}^2 + \eta_{i+1})\right] &\leq \boldsymbol{E}\left[\sum_{k=1}^{r} \nu_k^2 + \nu_k\right] + \boldsymbol{E}\left[\sum_{l=1}^{s} \nu_l'^2 + \nu'_l\right] \\ &\leq 5q^2\epsilon\,, \end{aligned}$$

using Lemma 1 and the fact that $(h_1^1, h_2^1), \ldots, (h_1^u, h_2^u) \xleftarrow{\$} \mathcal{H}^2$.

Finally, Theorem 3 is proven by combining Lemma 4 and ϵ_{ratio} with Lemma 3. □

6 Multi-User Security of pEDM PRF

In this section we consider the permutation based version of Encrypted Davies-Mayer (pEDM) construction, that was proposed by Dutta et al. [22]. They showed that pEDM based on a single permutation achieves $2n/3$-bit security. Here we will prove the *multi-user* security of pEDM based on two independent permutations, and we show that same level of security can be achieved using the technique proposed in this work. In this case, the multi-user security analysis is more complex than the single-user analysis, since the inputs to π_1 do not need to be fresh, this leads to more bad events and a more complex good transcripts ratio analysis when a dedicated proof need to be performed.

Let $n \in \mathbb{N}$, let $\pi_1, \pi_2 \xleftarrow{\$} \text{Perm}(n)$. One can consider a generic construction pEDM: $\{0,1\}^{2n} \times \{0,1\}^n \to \{0,1\}^n$ as

$$\text{pEDM}(K_1, K_2, M) = \pi_2(\pi_1(M \oplus K_1) \oplus M \oplus K_1 \oplus K_2) \oplus K_1\,. \tag{5}$$

The security of pEDM based on π_1 and π_2 is given in the following Theorem.

Theorem 4. *Let $n \in \mathbb{N}$ and $1 \leq \xi \leq 2^{n-1}/(p+q)$. We consider* pEDM: $\{0,1\}^{2n} \times \{0,1\}^n \to \{0,1\}^n$ *based on two permutations $\pi_1, \pi_2 \xleftarrow{\$}$* Perm$(n)$*, and u pairs of uniform user keys $(K_1^1, K_2^1), \ldots, (K_1^u, K_2^u) \xleftarrow{\$} \{0,1\}^{2n}$. For any distinguisher $\mathcal{D}$ making at most q construction queries distributed over its u construction oracles,* *at most p primitive queries to $\pi_1^{\pm}$ and p primitive queries to $\pi_2^{\pm}$, we have*

$$\mathbf{Adv}^{\text{mu-prf}}_{\text{pEDM}}(\mathcal{D}) \le \frac{2}{2^n} + \frac{4qp^2}{2^{2n}} + \frac{3q^2p}{2^{2n}} + \frac{3p\sqrt{nq}}{2^n} + \frac{2q^3}{2^{2n}} + \frac{p^{3/2}}{2^n} + \frac{(p+q)^2}{2^{2n}}\Big(7q + \frac{5u(u-1)}{2^n}\Big) + \frac{\binom{q}{\xi+1}}{2^{n\xi}}\,.$$

Proof. Let $(K_1^1, K_2^1), \ldots, (K_1^u, K_2^u) \xleftarrow{\$} \{0,1\}^{2n}$, $\pi_1, \pi_2 \xleftarrow{\$} \text{Perm}(n)$, and $\varphi_1, \ldots, \varphi_u \xleftarrow{\$} \text{Func}(n)$. Here, we consider any distinguisher $\mathcal{D}$ that has access to either $(\text{pEDM}^{\pi_1,\pi_2^{-1}}_{K_1^1,K_2^1}, \ldots, \text{pEDM}^{\pi_1,\pi_2^{-1}}_{K_1^u,K_2^u}, \pi_1, \pi_2)$ in the real world, or $(\varphi_1, \ldots, \varphi_u, \pi_1, \pi_2)$ in the ideal world. The security proof relies on Theorem 1 (b). As before, we consider $(\text{pEDM}^{\pi_1,\pi_2^{-1}}_{K_1^j,K_2^j})_{j=1}^u$ instead of $(\text{pEDM}^{\pi_1,\pi_2}_{K_1^j,K_2^j})_{j=1}^u$. We can view an evaluation $C = \text{pEDM}^{\pi_1,\pi_2^{-1}}_{K_1^j,K_2^j}(M)$ as the xor of two public permutations in the middle of the function, $\pi_1(M \oplus K_1^j) \oplus \pi_2(C \oplus K_1^j) = M \oplus K_1^j \oplus K_2^j$. Therefore, q evaluations of $\text{pEDM}^{\pi_1,\pi_2^{-1}}_{K_1^j,K_2^j}$ can be translated to a system of q bi-variate affine equations. Including $2p$ uni-variate affine equations that are defined by the primitive queries.

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}} \in \boldsymbol{\mathcal{T}_{\text{bad}}}]$. Following the framework given in Section 4, we first perform the bad transcripts analysis. By replacing $A = M \oplus K_1^j$, $B = C \oplus K_1^j$, and $Z = M \oplus K_1^j \oplus K_2^j$ in the framework of Section 4.3, we get the following bad events. Given a parameter $\xi \in \mathbb{N}$, we say that $\tau \in \mathcal{T}_{\text{bad}}$ if and only if there exist construction queries $(j, M, C), (j', M', C'), (j'', M'', C'') \in \tau_m$, and primitive queries $(u, v), (u', v') \in \tau_1$ and $(x, y), (x', y') \in \tau_2$ such that one of the following conditions holds:

(i) A component with two colliding vertices.

$$\begin{aligned} &\text{bad}_1\colon M \oplus u = K_1^j \,\wedge\, C \oplus x = K_1^j\,,\\ &\text{bad}_2\colon M \oplus u = K_1^j \,\wedge\, v \oplus y = M \oplus K_1^j \oplus K_2^j\,,\\ &\text{bad}_3\colon v \oplus y = M \oplus K_1^j \oplus K_2^j \,\wedge\, C \oplus x = K_1^j\,.\end{aligned}$$

(ii) Alternating paths of length 3 *across different users.*

$$\text{bad}_4\colon M \oplus K_1^j = M' \oplus K_1^{j'} \,\wedge\, C' \oplus K_1^{j'} = C'' \oplus K_1^{j''}\,.$$

(iii) Alternating paths of length 2 such that $\lambda(\mathcal{L}) = 0$ *across different users.*

$$\begin{aligned} &\text{bad}_5\colon M \oplus K_1^j = M' \oplus K_1^{j'} \,\wedge\, M \oplus K_1^j \oplus K_2^j = M' \oplus K_1^{j'} \oplus K_2^{j'}\,,\\ &\text{bad}_6\colon M \oplus K_1^j \oplus K_2^j = M' \oplus K_1^{j'} \oplus K_2^{j'} \,\wedge\, C \oplus K_1^j = C' \oplus K_1^{j'}\,.\end{aligned}$$

(iv) A v-$\star$ colliding component with y-colliding vertices, or a y-$\star$ colliding component with v-colliding vertices.

$$\begin{aligned}
\mathrm{bad}_7 &: M \oplus u = K_1^j \wedge C \oplus K_1^j = C' \oplus K_1^{j'}\,,\\
\mathrm{bad}_8 &: C \oplus x = K_1^j \wedge M \oplus K_1^j = M' \oplus K_1^{j'}\,,\\
\mathrm{bad}_9 &: M \oplus u = K_1^j \wedge M' \oplus u' = K_1^{j'}\\
&\quad \wedge v \oplus M \oplus K_1^j \oplus K_2^j = v' \oplus M' \oplus K_1^{j'} \oplus K_2^{j'}\,,\\
\mathrm{bad}_{10} &: C \oplus x = K_1^j \wedge C' \oplus x' = K_1^{j'}\\
&\quad \wedge y \oplus M \oplus K_1^j \oplus K_2^{j'} = y' \oplus M' \oplus K_1^j \oplus K_2^{j'}\,.
\end{aligned}$$

(v) A $(\xi+1)$-block-collision.

$$\mathrm{bad}_{11} : \{i_1, \ldots, i_{\xi+1}\} \in [q] \,\text{such that}\, C_{i_1} \oplus K_1^{j_{i_1}} = \cdots = C_{i_{\xi+1}} \oplus K_1^{j_{\xi+1}}\,.$$

Note that the events bad_4-bad_6 and bad_8 will not appear when the *single user* setting is considered, since in that case the distinguisher is not allow to query the same M to the construction oracle.

Lemma 5. *Let* $1 \le \xi \le 2^{n-1}/(p+q)$*. For any integers* q *and* p*, one has*

$$\Pr[\tau \in \mathcal{T}_{\mathrm{bad}}] \le \frac{2}{2^n} + \frac{4qp^2}{2^{2n}} + \frac{3q^2p}{2^{2n}} + \frac{3p\sqrt{nq}}{2^n} + \frac{2q^3}{2^{2n}} + \frac{p^{3/2}}{2^n} + \frac{\binom{q}{\xi+1}}{2^{n\xi}}\,.$$

The proof of the lemma is given in the full version of the paper.

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{O}}} = \boldsymbol{\tau}]/\,\mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}} = \boldsymbol{\tau}]$. The next step is the calculate the ratio for good transcripts. We use Theorem 1 (b) to get

$$\epsilon_{\mathrm{ratio}} \le \frac{2(p+q)^2 \sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1} (\eta_{i+1}^2 + \eta_{i+1})}{2^{2n}} + \frac{2q(p+q)^2}{2^{2n}}\,.$$

As before, for $r, s \in \mathbb{N}$, we denote by $\mathcal{P}_1^1, \ldots, \mathcal{P}_r^1$ and $\mathcal{P}_1^2, \ldots, \mathcal{P}_s^2$ the non-singleton equivalence classes of $[q]$ with respect to $\sim_1$ and $\sim_2$, respectively. Further, for $k \in [r]$ and $l \in [s]$, let $\nu_k = |\mathcal{P}_k^1|$ and $\nu_l' = |\mathcal{P}_l^2|$. Then, we have

$$\begin{aligned}
\boldsymbol{E}\left[\sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1} (\eta_{i+1}^2 + \eta_{i+1})\right] &\le \boldsymbol{E}\left[\sum_{k=1}^{r} \nu_k^2 + \nu_k\right] + \boldsymbol{E}\left[\sum_{l=1}^{s} \nu_l'^2 + \nu_l'\right]\\
&\le \frac{5u(u-1)}{2^{n-1}} + \frac{5q^2}{2^{n-1}}\,.
\end{aligned}$$

The non-freshness of π_1 in the multi-user setting leads to the existence of v-$\star$ components ($c_3 \neq 0$). The difficulty introduced by this can easily be handled by our new technique without performing a long and complicated analysis. Note that when $u = 1$, we are back to the single user setting, then there are no v-$\star$ components ($c_3 = 0$), since M is always different. Finally, Theorem 4 is proven by combining Lemma 5 and $\epsilon_{\mathrm{ratio}}$ with Lemma 3. □

7 Multi-User Security of nEHtM$_p$ MAC Algorithm

In this section we consider the public permutation based nonce-based Enhance Hash-then-Mask (nEHtM$_p$) MAC algorithm, that was proposed by Dutta and Nandi [20]. They showed that nEHtM$_p$ based on a single permutation (using domain separation) achieves $2n/3$-bit security when the number of faulty nonces μ is sufficiently smaller than $2^{n/3}$. However, according to the framework given in Section 4.3, the authors missed some bad events in their analysis, namely the two last events of (iv) and the three last events of (vii) of Section 4.3. Taking into account these missing bad events, the extended mirror theory used in [20] is *not* sufficient for the good transcripts ratio analysis of nEHtM$_p$. As a result, their ratio analysis of the construction is also incomplete. More precisely, non-equations between a colliding component and a normal component were not considered in the ratio analysis of [20]. This observation was also verified by the authors [10].

In this section, we will fix this problem using the techniques proposed in this work without performing a new complicated analysis, since these non-equations are already covered in our public permutation extended mirror theory (Theorem 2). Some of these additional bad events, however, require involved arguments to bound. Since the goal of this work is to illustrate the power of our new modular proof approaches, rather than presenting strong combinatorial results to bound these events. We will modify the design of nEHtM$_p$ by xoring an universal hash evaluation of the input message M using an extra hash key h^* to the output tag. This modified m-nEHtM$_p$ construction uses more randomness, which in turn enables us to bound the additional bad events easily. We would like to note that our analysis of m-nEHtM$_p$ does not imply infeasibility in fixing the proof of nEHtM$_p$. In fact, we believe that the security of the original nEHtM$_p$ construction can also be proven with our new approaches in combination with some strong techniques to bound these bad events. Here we will prove that this m-nEHtM$_p$ construction based on two independent permutations achieves $2n/3$-bit security in the *multi-user* setting using the technique proposed in this work.

Let $n \in \mathbb{N}$, let $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$, and let $\mathcal{H}$ be an ϵ-AXU function family. One can consider a generic construction $\text{m-nEHtM}_\text{p} \colon \{0,1\}^n \times \mathcal{H}^2 \times \{0,1\}^n \times \mathcal{M} \to \{0,1\}^n$ as

$$\text{m-nEHtM}_\text{p}(K, h, h^* N, M) = \pi_1(N \oplus K) \oplus \pi_2(N \oplus h(M)) \oplus h^*(M)\,. \qquad (6)$$

The security of m-nEHtM$_p$ based on π_1 and π_2 is given in the following Theorem.

Theorem 5. *Let $n \in \mathbb{N}$, and let $\mathcal{H}$ be a uniform ϵ-AXU family of functions from $\mathcal{M}$ to $\{0,1\}^n$. We consider* $\text{m-nEHtM}_p \colon \{0,1\}^n \times \mathcal{H}^2 \times \{0,1\}^n \times \mathcal{M} \to \{0,1\}^n$ *based on two permutations $\pi_1, \pi_2 \xleftarrow{\$} \mathrm{Perm}(n)$, u uniform user keys $K_1, \ldots, K_u \xleftarrow{\$} \{0,1\}^n$ and u pairs of uniform user hash keys $(h_1, h_1^*), \ldots, (h_u, h_u^*) \xleftarrow{\$} \mathcal{H}^2$. Let μ be a fixed parameter. For any distinguisher $\mathcal{D}$ making at most q_m queries with at most μ faulty nonces distributed over its u construction MAC oracles, q_a queries* *distributed over its u construction verification oracles, at most p primitive queries to $\pi_1^{\pm}$ and p primitive queries to $\pi_2^{\pm}$, we have*

$$\mathbf{Adv}^{\text{mu-mac}}_{\text{m-nEHtM}_p}(\mathcal{D}) \le 7\sqrt{q_m}p\epsilon + 2\mu^2\epsilon + 4q_m^3\epsilon^2 + \frac{q_m^2 p\epsilon}{2^n} + 2\mu p\epsilon + q_m^2 q_a\epsilon^2 + \frac{q_a p^2\epsilon}{2^n}$$
$$+ \frac{3q_m q_a p\epsilon}{2^n} + p\sqrt{q_m q_a}\epsilon^{\frac{3}{2}} + \frac{(p+q_m)^2}{2^{2n}}\Big(5\mu^2 + 7q_m + \frac{5u(u-1)}{2^n}\Big) + \frac{q_a}{2^n}\,.$$

Proof. Let $K_1,\ldots,K_u \xleftarrow{\$} \{0,1\}^n$, $(h_1,h_1^*),\ldots,(h_u,h_u^*) \xleftarrow{\$} \mathcal{H}^2$, and $\pi_1,\pi_2 \xleftarrow{\$} \text{Perm}(n)$. Here, we consider any distinguisher $\mathcal{D}$ that has access to either $(\mathcal{O},\pi_1,\pi_2)$ in the real world or $(\mathcal{P},\pi_1,\pi_2)$ in the ideal world, with $\mathcal{O} = \big((\text{m-nEHtM}^{\pi_1,\pi_2}_{p(K_1,h_1,h_1^*)}, \mathsf{Ver}^{\pi_1,\pi_2}_{(K_1,h_1,h_1^*)}),\ldots,(\text{m-nEHtM}^{\pi_1,\pi_2}_{p(K_u,h_u,h_u^*)}, \mathsf{Ver}^{\pi_1,\pi_2}_{(K_u,h_u,h_u^*)})\big)$ and $\mathcal{P} = \big((\mathsf{Rand}_1,\mathsf{Rej}_1),\ldots,(\mathsf{Rand}_u,\mathsf{Rej}_u)\big)$. The security proof relies on Theorem 2 (b).

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}} \in \boldsymbol{\mathcal{T}_{\text{bad}}}]$. Following the framework given in Section 4, we first perform the bad transcripts analysis. For the notational simplicity, we denote $H_j = h_j(M)$, and By replacing $A = N \oplus K_j$, $B = N \oplus h_j(M)$, and $Z = T \oplus h_j^*(M)$ in Section 4.3, we get the following bad events. Given a parameter $\xi \in \mathbb{N}$, we say that $\tau \in \mathcal{T}_{\text{bad}}$ if and only if there exist construction MAC queries $(j,N,M,T),(j',N',M',T'),(j'',N'',M'',T'') \in \tau_m$, a construction verification query $(j^{(a)},N^{(a)},M^{(a)},T^{(a)},b^{(a)}) \in \tau_a$ and primitive queries $(u,v),(u',v') \in \tau_1$ and $(x,y),(x',y') \in \tau_2$ such that one of the following conditions holds:

(i) A component with two colliding vertices.

$$\begin{aligned}
\text{bad}_1\colon & N \oplus u = K_j \ \wedge\ N \oplus x = H_j\,,\\
\text{bad}_2\colon & N \oplus u = K_j \ \wedge\ v \oplus y = T \oplus H_j^*\,,\\
\text{bad}_3\colon & v \oplus y = T \oplus H_j^* \ \wedge\ N \oplus x = H_j\,.
\end{aligned}$$

(ii) An alternating path of length 3.

$$\text{bad}_4\colon N \oplus K_j = N' \oplus K_{j'} \ \wedge\ N' \oplus H_{j'} = N'' \oplus H_{j''}\,.$$

(iii) An alternating path of length 2 such that $\lambda(\mathcal{L}) = 0$.

$$\begin{aligned}
\text{bad}_5\colon & N \oplus K_j = N' \oplus K_{j'} \ \wedge\ T \oplus H_j^* = T' \oplus H_{j'}^*\,,\\
\text{bad}_6\colon & T \oplus H_j^* = T' \oplus H_{j'}^* \ \wedge\ N \oplus H_{j'} = N' \oplus H_{j'}\,.
\end{aligned}$$

(iv) A v-$\star$ colliding component with y-colliding vertices, or a y-$\star$ colliding component with v-colliding vertices.

$$\begin{aligned}
\text{bad}_7\colon & N \oplus u = K_j \ \wedge\ N \oplus H_j = N' \oplus H_{j'}\,,\\
\text{bad}_8\colon & N \oplus x = H_j \ \wedge\ N \oplus K_j = N' \oplus K_{j'}\,,\\
\text{bad}_9\colon & N \oplus u = K_j \ \wedge\ N' \oplus u' = K_{j'} \ \wedge\ v \oplus T \oplus H_j^* = v' \oplus T' \oplus H_{j'}^*\,,\\
\text{bad}_{10}\colon & N \oplus x = H_j \ \wedge\ N' \oplus x' = H_{j'} \ \wedge\ y \oplus T \oplus H_j^* = y' \oplus T' \oplus H_{j'}^*\,.
\end{aligned}$$

(v) A $(\xi+1)$-block-collision.

$$\text{bad}_{11}\colon \{i_1,\ldots,i_{\xi+1}\}\in[q_m] \text{ such that } N_{i_1}\oplus H_{j_{i_1}}=\cdots=N_{i_{\xi+1}}\oplus H_{j_{i_{\xi+1}}}\,.$$

(vi) An alternating circle of length 2 with a λ'-labeled edge.

$$\begin{aligned}\text{bad}_{12}\colon\, & N\oplus K_j=N^{(a)}\oplus K_{j^{(a)}}\ \wedge\ N\oplus H_j=N^{(a)}\oplus H_{j^{(a)}}\\ & \wedge\ T\oplus H^*_j=T^{(a)}\oplus H^*_{j^{(a)}}\,.\end{aligned}$$

(vii) A λ'-labeled edge between two vertices with distance λ'.

$$\begin{aligned}&\text{bad}_{13}\colon N^{(a)}\oplus u=K_{j^{(a)}}\ \wedge\ N^{(a)}\oplus x=H_{j^{(a)}}\ \wedge\ T^{(a)}\oplus H^*_{j^{(a)}}=v\oplus y\,,\\ &\text{bad}_{14}\colon N\oplus u=K_j\ \wedge\ N\oplus H_j=N^{(a)}\oplus H_{j^{(a)}}\ \wedge\ N^{(a)}\oplus u'=K_{j^{(a)}}\\ &\qquad\wedge\ T^{(a)}\oplus H^*_{j^{(a)}}=v\oplus T\oplus H^*_j\oplus v'\,,\\ &\text{bad}_{15}\colon N\oplus x=H_j\ \wedge\ N\oplus K_j=N^{(a)}\oplus K_{j^{(a)}}\ \wedge\ N^{(a)}\oplus x'=H_{j^{(a)}}\\ &\qquad\wedge\ T^{(a)}\oplus H^*_{j^{(a)}}=y\oplus T\oplus H^*_j\oplus y'\,,\\ &\text{bad}_{16}\colon N\oplus u=K_j\ \wedge\ N'\oplus x=H_{j'}\ \wedge\ N\oplus H_j=N^{(a)}\oplus H_{j^{(a)}}\ \wedge\\ &\quad N'\oplus K_{j'}=N^{(a)}\oplus K_{j^{(a)}}\ \wedge\ T^{(a)}\oplus H^*_{j^{(a)}}=v\oplus T\oplus H^*_j\oplus y\oplus T'\oplus H^*_{j'}\,.\end{aligned}$$

Note that the events bad_9-bad_{10} and bad_{14}-bad_{16} are the missing events that were not considered by the authors of [20].

Lemma 6. *For any integers q_m, q_a and p, then one has*

$$\begin{aligned}\Pr[\tau\in\mathcal{T}_{\text{bad}}]\le{}& 7\sqrt{q_m}p\epsilon+2\mu^2\epsilon+4q_m^3\epsilon^2+\frac{q_m^2p\epsilon}{2^n}+2\mu p\epsilon\\ &+\frac{8q_m^2(p+q_m)^2\epsilon}{2^{2n}}+q_m^2q_a\epsilon^2+\frac{q_ap^2\epsilon}{2^n}+\frac{3q_mq_ap\epsilon}{2^n}+p\sqrt{q_mq_a}\epsilon^{3/2}\,.\end{aligned}$$

The proof of this Lemma is given in the full version of the paper.

$\mathbf{Pr}[\boldsymbol{X_{\mathcal{O}}}=\boldsymbol{\tau}]/\mathbf{Pr}[\boldsymbol{X_{\mathcal{P}}}=\boldsymbol{\tau}]$. The next step is the calculate the ratio for good transcripts. We use Theorem 2 (b) to get

$$\epsilon_{\text{ratio}}\le\frac{2(p+q_m)^2\sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1}(\eta_{i+1}^2+\eta_{i+1})}{2^{2n}}+\frac{2q_m(p+q_m)^2}{2^{2n}}+\frac{2q_a}{2^n}\,.$$

As before, for $r,s\in\mathbb{N}$, we denote by $\mathcal{P}^1_1,\ldots,\mathcal{P}^1_r$ and $\mathcal{P}^2_1,\ldots,\mathcal{P}^2_s$ the non-singleton equivalence classes of $[q_m]$ with respect to $\sim_1$ and $\sim_2$, respectively. For $k\in[r]$ and $l\in[s]$, let $\nu_k=|\mathcal{P}^1_k|$ and $\nu'_l=|\mathcal{P}^2_l|$. Then, we have

$$\begin{aligned}\boldsymbol{E}\left[\sum_{i=c_1+c_2}^{c_1+c_2+c_3+c_4-1}(\eta_{i+1}^2+\eta_{i+1})\right]&\le\boldsymbol{E}\left[\sum_{k=1}^{r}\nu_k^2+\nu_k\right]+\boldsymbol{E}\left[\sum_{l=1}^{s}\nu_l'^2+\nu'_l\right]\\ &\le\frac{5\mu^2}{2}+\frac{5u(u-1)}{2^{n-1}}+\frac{5q_m^2\epsilon}{2}\,,\end{aligned}$$

using Lemma 1 and the fact that $K_1, \ldots, K_u \xleftarrow{\$} \{0,1\}^n$ and $(h_1, h_1^*), \ldots, (h_u, h_u^*) \xleftarrow{\$} \mathcal{H}^2$. Note that when $u = 1$, we are back to the single-user setting, in this case the v-$\star$ components (collision in the inputs of π_1) can only be formed by queries with repeated nonces, hereby the $5\mu^2/2$ term in the bound. Finally, Theorem 5 is proven by combining Lemma 6 and ϵ_{ratio} with Lemma 3. □

Acknowledgments. This work was supported in part by the Research Council KU Leuven: GOA TENSE (C16/15/058). The author was supported by a Ph.D. Fellowship from the Research Foundation - Flanders (FWO). I want to thank Tim Beyne and the reviewers for their valuable comments and suggestions.

References

1. NIST Lightweight Cryptography https://csrc.nist.gov/Projects/Lightweight-Cryptography
2. NIST SHA-3 Project https://csrc.nist.gov/projects/hash-functions/sha-3-project
3. Bellare, M., Boldyreva, A., Micali, S.: Public-key encryption in a multi-user setting: Security proofs and improvements. In: EUROCRYPT 2000. LNCS, vol. 1807, pp. 259–274
4. Biham, E.: How to decrypt or even substitute des-encrypted messages in 2^{28} steps. Inf. Process. Lett. **84**(3), 117–124
5. Bogdanov, A., Knežević, M., Leander, G., Toz, D., Varici, K., Verbauwhede, I.: Spongent: A lightweight hash function. In: CHES 2011. LNCS, vol. 6917, pp. 312–325
6. Bogdanov, A., Knudsen, L.R., Leander, G., Standaert, F.X., Steinberger, J.P., Tischhauser, E.: Key-alternating ciphers in a provable setting: Encryption using a small number of public permutations - (extended abstract). In: EUROCRYPT 2012. LNCS, vol. 7237, pp. 45–62
7. Chatterjee, S., Menezes, A., Sarkar, P.: Another look at tightness. In: SAC 2011. LNCS, vol. 7118, pp. 293–319
8. Chen, S., Lampe, R., Lee, J., Seurin, Y., Steinberger, J.P.: Minimizing the two-round Even-Mansour cipher. In: CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 39–56
9. Chen, S., Steinberger, J.P.: Tight security bounds for key-alternating ciphers. In: EUROCRYPT 2014. LNCS, vol. 8441, pp. 327–350
10. Chen, Y.L., Dutta, A., Nandi, M.: Multi-user BBB security of public permutations based MAC. Cryptogr. Commun. **14**(5), 1145–1177
11. Chen, Y.L., Lambooij, E., Mennink, B.: How to build pseudorandom functions from public random permutations. In: CRYPTO 2019, Part I. LNCS, vol. 11692, pp. 266–293
12. Chen, Y.L., Tessaro, S.: Better security-efficiency trade-offs in permutation-based two-party computation. In: ASIACRYPT 2021, Part II. LNCS, vol. 13091, pp. 275–304
13. Choi, W., Lee, B., Lee, J., Lee, Y.: Toward a fully secure authenticated encryption scheme from a pseudorandom permutation. In: ASIACRYPT 2021, Part III. LNCS, vol. 13092, pp. 407–434

14. Choi, W., Lee, B., Lee, Y., Lee, J.: Improved security analysis for nonce-based enhanced hash-then-mask macs. In: ASIACRYPT 2020, Part I. LNCS, vol. 12491, pp. 697–723
15. Cogliati, B., Lampe, R., Seurin, Y.: Tweaking Even-Mansour ciphers. In: CRYPTO 2015, Part I. LNCS, vol. 9215, pp. 189–208
16. Cogliati, B., Seurin, Y.: EWCDM: An efficient, beyond-birthday secure, nonce-misuse resistant MAC. In: CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 121–149
17. Daemen, J., Rijmen, V.: The Design of Rijndael - The Advanced Encryption Standard (AES), Second Edition. Information Security and Cryptography, Springer (2020)
18. Datta, N., Dutta, A., Nandi, M., Yasuda, K.: Encrypt or decrypt? to make a single-key beyond birthday secure nonce-based MAC. In: CRYPTO 2018, Part I. LNCS, vol. 10991, pp. 631–661
19. Dutta, A.: Minimizing the two-round tweakable even-mansour cipher. In: ASIACRYPT 2020, Part I. LNCS, vol. 12491, pp. 601–629
20. Dutta, A., Nandi, M.: BBB secure nonce based MAC using public permutations. In: AFRICACRYPT 2020. LNCS, vol. 12174, pp. 172–191
21. Dutta, A., Nandi, M., Talnikar, S.: Beyond birthday bound secure MAC in faulty nonce model. In: EUROCRYPT 2019, Part I. LNCS, vol. 11476, pp. 437–466
22. Dutta, A., Nandi, M., Talnikar, S.: Permutation based EDM: an inverse free BBB secure PRF. IACR Trans. Symmetric Cryptol. **2021**(2), 31–70
23. Even, S., Mansour, Y.: A construction of a cipher from a single pseudorandom permutation. In: ASIACRYPT'91. LNCS, vol. 739, pp. 210–224
24. Guo, C., Katz, J., Wang, X., Yu, Y.: Efficient and secure multiparty computation from fixed-key block ciphers. In: 2020 IEEE Symposium on Security and Privacy. pp. 825–841
25. Guo, J., Peyrin, T., Poschmann, A.: The PHOTON family of lightweight hash functions. In: CRYPTO 2011. LNCS, vol. 6841, pp. 222–239
26. Hoang, V.T., Tessaro, S.: Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In: CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 3–32
27. Jha, A., Nandi, M.: Tight security of cascaded LRW2. Journal of Cryptology **33**(3), 1272–1317
28. Kim, S., Lee, B., Lee, J.: Tight security bounds for double-block hash-then-sum MACs. In: EUROCRYPT 2020, Part I. LNCS, vol. 12105, pp. 435–465
29. Krawczyk, H.: LFSR-based hashing and authentication. In: CRYPTO'94. LNCS, vol. 839, pp. 129–139
30. Mennink, B., Neves, S.: Encrypted Davies-Meyer and its dual: Towards optimal security using mirror theory. In: CRYPTO 2017, Part III. LNCS, vol. 10403, pp. 556–583
31. Mouha, N., Luykx, A.: Multi-key security: The Even-Mansour construction revisited. In: CRYPTO 2015, Part I. LNCS, vol. 9215, pp. 209–223
32. Nandi, M.: Mind the composition: Birthday bound attacks on EWCDMD and SoKAC21. In: EUROCRYPT 2020, Part I. LNCS, vol. 12105, pp. 203–220
33. Patarin, J.: The "coefficients H" technique (invited talk). In: SAC 2008. LNCS, vol. 5381, pp. 328–345
34. Patarin, J.: Mirror theory and cryptography. Appl. Algebra Eng. Commun. Comput. **28**(4), 321–338