Blockcipher-Based MACs: Beyond the Birthday Bound Without Message Length

Yusuke Naito(✉)

Mitsubishi Electric Corporation, Kanagawa, Japan
Naito.Yusuke@ce.MitsubishiElectric.co.jp

Abstract. We present blockcipher-based MACs (Message Authentication Codes) that have beyond the birthday bound security without message length in the sense of PRF (Pseudo-Random Function) security. Achieving such security is important in constructing MACs using blockciphers with short block sizes (e.g., 64 bit).

Luykx *et al.* (FSE 2016) proposed LightMAC, the first blockcipher-based MAC with such security and a variant of PMAC, where for each n-bit blockcipher call, an m-bit counter and an $(n-m)$-bit message block are input. By the presence of counters, LightMAC becomes a secure PRF up to $O(2^{n/2})$ tagging queries. Iwata and Minematsu (TOSC 2016, Issue 1) proposed F_t, a keyed hash function-based MAC, where a message is input to t keyed hash functions (the hash function is performed t times) and the t outputs are input to the xor of t keyed blockciphers. Using the LightMAC's hash function, F_t becomes a secure PRF up to $O(2^{tn/(t+1)})$ tagging queries. However, for each message block of $(n-m)$ bits, it requires t blockcipher calls.

In this paper, we improve F_t so that a blockcipher is performed only once for each message block of $(n-m)$ bits. We prove that our MACs with $t \leq 7$ are secure PRFs up to $O(2^{tn/(t+1)})$ tagging queries. Hence, our MACs with $t \leq 7$ are more efficient than F_t while keeping the same level of PRF-security.

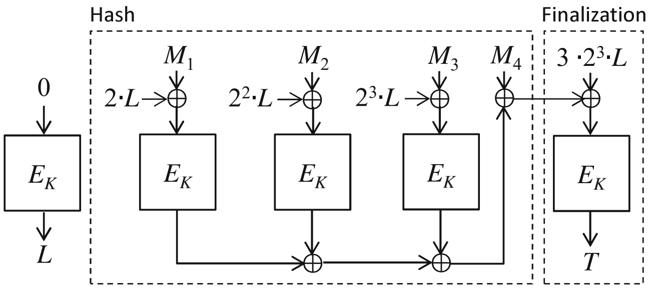
Keywords: MAC · Blockcipher · PRF · PRP · Beyond the birthday bound · Message length · Counter

1 Introduction

A MAC (Message Authentication Code) is a fundamental symmetric-key primitive that produces a tag to authenticate a message. MACs are often realized by using a blockcipher so that these become secure PRFs (Pseudo-Random Functions) under the standard assumption that the underlying keyed blockciphers are pseudo-random permutations. Hence, in security proofs, these are replaced with random permutations. The advantage of PRF-security is commonly measured by using the parameters: n the block length, q the total number of tagging queries, ℓ the maximum message length (in blocks) of each query and σ the total message length (in blocks) of all queries. Many blockcipher-based MACs are provided with the so-called birthday security. The basic birthday bound looks like $O(\ell^2 q^2/2^n)$ or $O(\sigma^2/2^n)$.

T. Takagi and T. Peyrin (Eds.): ASIACRYPT 2017, Part III, LNCS 10626, pp. 446–470, 2017.
https://doi.org/10.1007/978-3-319-70700-6_16

Blockcipher-based MACs are mainly categorized into CBC-type MACs and PMAC-type ones. These MACs are constructed from two functions: hash and finalization functions, where a hash function produces a fixed length hash value from an arbitrary length message; a finalization function produces a tag from a hash value. CBC-type MACs [2,8,15,20,30,31] use hash functions that iterate a keyed blockcipher. The PRF-security bound becomes the birthday one due to the collision in the chaining values. PMAC-type MACs [9,33] use hash functions using a keyed blockcipher parallelly. The following figure shows the structure of PMAC1, where E_K is a keyed blockcipher (K is a secret key), M_1, M_2, M_3 and M_4 are n-bit message blocks and multiplications are performed over the multiplication subgroup of $GF(2^n)$. For collision inputs to the keyed blockcipher,the outputs are canceled out before the finalization function. Hence, the collision might trigger a distinguishing attack. By the birthday analysis for the input collision, the PRF-security bound becomes the birthday one.



MACs with Beyond the Birthday Bound Security. The birthday bound security may not be enough for blockciphers with short block sizes such as Triple-DES and lightweight blockciphers, as mentioned in [7]. Hence, designing a MAC with *beyond the birthday bound* (BBB) security is an important research of MAC design. Such MACs contribute not only to blockciphers with short block sizes but also to the longevity of 128-bit blockciphers.

Yasuda proposed a CBC-type MAC, called SUM-ECBC [36], and a PMAC-type one, called `PMAC_Plus` [37]. He proved that the PRF-security bounds become $O(\ell^3 q^3/2^{2n})$. Later, Zhang *et al.* proposed a CBC-type MAC, called 3kf9 [40] that is more efficient than SUM-ECBC. These hash functions have a double length ($2n$ bit) internal state and produce a $2n$-bit value. These finalization functions have the xor of two keyed blockciphers that generates a tag from a $2n$-bit hash value. By the double length internal state, the influences of ℓ and q on the bounds are weakened.

Yasuda designed a PMAC-type MAC, called PMAC with Parity [38], with the aim of weakening the influence of ℓ. He proved that the PRF-security bound becomes $O(q^2/2^n + \ell q\sigma/2^{2n})$. Later, Zhang proposed a PMAC-type MAC with better efficiency, called PMACX [41]. Luykx *et al.* proposed a PMAC-type MAC, called LightMAC [25]. LightMAC is the counter-based construction that is used in the XOR MAC [1] and the protected counter sum [6]. LightMAC can be seen as a counter-based PMAC in which $(i)_m \| M_i$ is input to the i-th keyed blockcipher call, where $(i)_m$ is the m-bit binary representation of i and M_i is the i-th message block of $n-m$ bits. By the presence of counters, the input collision can be avoided, thereby the influence ℓ can completely be removed. They proved that the PRF-security bound becomes $O(q^2/2^n)$, namely, LightMAC is a secure PRF up to $O(2^{n/2})$ tagging queries.

Recently, Iwata and Minematsu proposed MACs with beyond the $O(2^{n/2})$-security, called F_t [16]. F_t is based on t keyed hash functions $H_{L_1}, \ldots, H_{L_t}$ and t keyed blockciphers $E_{K_1}, \ldots, E_{K_t}$, where $L_1, \ldots, L_t$ are hash keys. For a message M, the tag is defined as $\mathsf{F}_t(M) = \bigoplus_{i=1}^{t} E_{K_i}(S_i)$ where $S_i = H_{L_i}(M)$. They proved that the PRF-security bound becomes $O(q^{t+1} \cdot \epsilon^t)$ as long as the keyed hash functions are ϵ-almost universal. They pointed out that the hash function of LightMAC is a $O(1/2^n)$-almost universal hash function, and adopting it as these hash functions, the PRF-security bound becomes $O(q^{t+1}/2^{tn})$. Namely, it is a secure PRF up to $O(2^{tn/(t+1)})$ tagging queries.

Why BBB-Security Without Message Length? We explain the importance of achieving BBB-security without message length. Here we consider the following example: the block length $n = 64$, the message length 2^{15} bits (4 Kbytes), and the threshold $1/2^{20}$ (a key is changed when the security bound equals the threshold). The message length is the case of HTTPS connection given in [7] and the threshold is given in [25]. We define the counter size as $m = n/3$ (rounded to the nearest multiple of 8) (in this case, $n = 64$ and $m = 24$). Putting these parameters into security bounds of `PMAC_Plus` ($O(\ell^3 q^3/2^{2n})$), LightMAC ($O(q^2/2^n)$), and F_t using LightMAC ($O(q^{t+1}/2^{tn})$), a key is changed after the tagging queries given in Table 1 (Line with "Queries"). Then, we consider the case that 2900 tagging queries of message length 4 Kbytes per second can be made. This example is the case of HTTPS connection given in [7]. In this case, a key is changed after the times given in Table 1 (Line with "Times"). Note that the security bound of `PMAC_Plus` depends on the message length, thereby increasing the length decreases the time. As shown Table 1, `PMAC_Plus` and LightMAC require a rekeying within a day, whereas F_t does not require such frequent rekeyings.

Table 1. The numbers of tagging queries of changing a key and the times.

	`PMAC_Plus`	LightMAC	F_2 $(t=2)$	F_3 $(t=3)$	F_4 $(t=4)$	$\cdots$
Queries	2^{29}	2^{22}	2^{36}	2^{43}	2^{47}	$\cdots$
Times	13 hrs	12 min	274 days	96 years	1539 years	$\cdots$

Question. As mentioned above, achieving BBB-security without message length is important for blockciphers with short block sizes, and F_t using LightMAC achieves such security. However, it is inefficient because for each input block $(i)_m \| M_i$ it requires t blockcipher calls. It is roughly t times slower than LightMAC. Therefore, the main question of this paper is: *can we design more efficient MACs than* F_t *while keeping* $O(2^{tn/(t+1)})$*-security?*

Our Results. Firstly, we focus to design a MAC that is more efficient than F_2 and achieves the $O(2^{2n/3})$-security. As the research direction from PMAC to LightMAC, it is natural to consider a counter-based `PMAC_Plus`. We call the resultant scheme "`LightMAC_Plus`". Regarding the efficiency, `LightMAC_Plus` requires roughly one blockcipher call for each input block $(i)_m \| M_i$, while F_2 requires two blockcipher calls. Hence, `LightMAC_Plus` is more efficient than F_2. Regarding the PRF-security, by the presence of counters, the influence of ℓ can be removed. We prove that the PRF-security bound becomes $O(q^3/2^{2n})$, namely, `LightMAC_Plus` is a secure PRF up to $O(2^{2n/3})$ queries.

Next, we focus to design a MAC that is more efficient than F_t and achieves $O(2^{tn/(t+1)})$-security, where $t \geq 3$. Regarding the hash function, we also use that of `LightMAC_Plus`. Hence, this hash function is roughly t times faster than that of F_t. In order to ensure randomnesses of tags, we use the xor of t keyed blockciphers. However, there is a gap between the output length of the hash function ($2n$ bit) and the input length of the xor function (tn bit). Therefore, we propose a new construction that links between a $2n$-bit output and a tn-bit input. We call the resultant scheme "`LightMAC_Plus2`", and prove that if $t \leq 7$, then the PRF-security bound becomes $O(q^{t+1}/2^{tn} + q^2/2^{2n})$, namely, it is a secure PRF up to $O(2^{tn/(t+1)})$ tagging queries. In the proof of `LightMAC_Plus2`, we generalize the hash function by an ϵ-almost universal one, and prove that if $t \leq 7$, then the PRF-security bound is $O(q^{t+1}/2^{tn} + \epsilon)$. We prove that the counter-based hash function is $O(q^2/2^{2n})$-almost universal, which offers the PRF-security bound: $O(q^{t+1}/2^{tn} + q^2/2^{2n})$.

Table 2. Comparison of our MACs and other BBB-secure MACs. Column "# bits/BCs" refers to the number of bits of input message processed per blockcipher call. Column "# BCs in FF" refers to the number of blockcipher calls in a finalization function. F_t uses the hash function of LightMAC. `LightMAC_Plus2` has the condition $t \leq 7$.

Scheme	# keys	# bits/BC	# BCs in FF	Security	Ref.
`PMAC_Plus`	3	n	2	$O(\ell^3 q^3/2^{2n})$	[37]
LightMAC	2	$n-m$	1	$O(q^2/2^n)$	[25]
F_t	$2t$	$(n-m)/t$	t	$O(q^{t+1}/2^{tn})$	[16]
`LightMAC_Plus`	3	$n-m$	2	$O(q^3/2^{2n})$	This paper
`LightMAC_Plus2`	$t+3$	$n-m$	$t+2$	$O(q^{t+1}/2^{tn} + \epsilon)$	This paper

Finally, in Table 2, we compare our MACs with BBB-secure MACs `PMAC_Plus`, LightMAC, and F_t. These MACs are PMAC-type ones, and thus parallelizable. We note that the PRF-security bound of `LightMAC_Plus2` is satisfied when $t \leq 7$. Proving the PRF-security with $t > 7$ is left as an open problem.

Related Works. The PRF-security bounds of CBC-type MACs and PMAC-type MACs were improved to $O(\ell q^2/2^n)$ [3,27] and $O(\sigma q/2^n)$ [29]. Luykx *et al.*

studied the influence of ℓ in the PMAC's bound [24]. They showed that PMAC with Gray code [9] may not achieve the PRF-security bound of $O(q^2/2^n)$. Gaži *et al.* [14] showed that there exists an attack to PMAC with Gray code with the probability of $\Omega(\ell q^2/2^n)$, and instead proved that PMAC with 4-wise independent masks achieves the PRF-security bound of $O(q^2/2^n)$, where the input masks are defined by using 4 random values. Dodis and Steinberger [12] proposed a secure MAC from unpredicable keyed blockciphers with beyond the birthday bound security. Note that the security bound of their MAC includes the message length. Several randomized MACs achieve beyond the birthday bound security [18,19,26]. These require a random value for each query, while our MACs are deterministic, namely, a random value is not required.

Several compression function-based MACs achieve BBB security e.g., [13,21,35,39]. Naito [28], List and Nandi [22], and Iwata *et al.* [17] proposed tweakable blockcipher-based MACs with BBB security. These MACs also employ the counter-based `PMAC_Plus`-style construction, where a counter is input as tweak. Namely, in the security proofs, the power of a tweakable blockcipher is used (distinct tweaks offer distinct random permutations). On the other hand, our MACs do not change the permutation in the hash function for each message block and the permutations in the finalization function. Peyrin and Seurin proposed a nonce-based and tweakable blockcipher-based MAC with BBB security [32]. Several Wegman-Carter-type MACs with BBB security were proposed e.g., [10,11,34]. These MACs use a random value or a nonce, whereas our MACs do not require either of them.

Organization. In Sect. 2, we give notations and the definition of PRF-security. In Sect. 3, we give the description of `LightMAC_Plus` and the PRF-security bound. In Sect. 4, we give the proof of the PRF-security. In Sect. 5, we give the description of `LightMAC_Plus2` and the PRF-security bound. In Sect. 6, we give the proof of the PRF-security. Finally, in Sect. 7, we improve the efficiency of the hash function of `LightMAC_Plus2`.

2 Preliminaries

Notation. Let $\{0,1\}^*$ be the set of all bit strings. For a non-negative integer n, let $\{0,1\}^n$ be the set of all n-bit strings, and 0^n the bit string of n-bit zeroes. For a positive integer i, $[i] := \{1,2,\ldots,i\}$. For non-negative integers i,m with $i < 2^m$, $(i)_m$ denotes the m-bit binary representation of i. For a finite set X, $x \xleftarrow{\$} X$ means that an element is randomly drawn from X and is assigned to x. For a positive integer n, $\mathsf{Perm}(n)$ denotes the set of all permutations: $\{0,1\}^n \to \{0,1\}^n$ and $\mathsf{Func}(n)$ denotes the set of all functions: $\{0,1\}^* \to \{0,1\}^n$. For sets X and Y, $X \leftarrow Y$ means that Y is assigned to X. For a bit string x and a set X, $|x|$ and $|X|$ denote the bit length of x and the number of elements in X, respectively. X^s denotes the s-array cartesian power of X for a set X and a positive integer s.

Let $GF(2^n)$ be the field with 2^n points and $GF(2^n)^*$ the multiplication subgroup of $GF(2^n)$ which contains $2^n - 1$ points. We interchangeably think of a point a in $GF(2^n)$ in any of the following ways: as an n-bit string $a_{n-1} \cdots a_1 a_0 \in \{0,1\}^n$ and as a formal polynomial $a_{n-1}\mathtt{x}^{n-1} + \cdots + a_1\mathtt{x} + a_0 \in GF(2^n)$. Hence we need to fix an irreducible polynomial $a(\mathtt{x}) = \mathtt{x}^n + a_{n-1}\mathtt{x}^{n-1} + \cdots + a_1\mathtt{x} + a_0$. This paper uses an irreducible polynomial with the property that the element $2 = \mathtt{x}$ generates the entire multiplication group $GF(2^n)^*$ of order $2^n - 1$. Examples of irreducible polynomial for $n = 64$ and $n = 128$ are given in [33]: $a(\mathtt{x}) = \mathtt{x}^{64} + \mathtt{x}^4 + \mathtt{x}^3 + \mathtt{x} + 1$ and $a(\mathtt{x}) = \mathtt{x}^{128} + \mathtt{x}^7 + \mathtt{x}^2 + \mathtt{x} + 1$, respectively.

PRF-Security. We focus on the information-theoretic model, namely, all keyed blockciphers are assumed to be random permutations, where a random permutation is defined as $P \xleftarrow{\$} \mathsf{Perm}(n)$. Through this paper, a distinguisher $\mathcal{D}$ is a computationally unbounded algorithm. It is given query access to an oracle $\mathcal{O}$, denoted by $\mathcal{D}^{\mathcal{O}}$. Its complexity is solely measured by the number of queries made to its oracles. Let $F[\mathbf{P}]$ be a function using s permutations $\mathbf{P} = (P^{(1)}, \ldots, P^{(s)})$.

The PRF-security of $F[\mathbf{P}]$ is defined in terms of indistinguishability between the real and ideal worlds. In the real world, $\mathcal{D}$ has query access to $F[\mathbf{P}]$ for $\mathbf{P} \xleftarrow{\$} \mathsf{Perm}(n)^s$. In the ideal world, it has query access to a random function $\mathcal{R}$, where a random function is defined as $\mathcal{R} \xleftarrow{\$} \mathsf{Func}(n)$. After interacting with an oracle $\mathcal{O}$, $\mathcal{D}$ outputs $y \in \{0,1\}$. This event is denoted by $\mathcal{D}^{\mathcal{O}} \Rightarrow y$. The advantage function is defined as

$$\mathbf{Adv}^{\mathsf{prf}}_{F[\mathbf{P}]}(\mathcal{D}) = \Pr\left[\mathbf{P} \xleftarrow{\$} \mathsf{Perm}(n)^s; \mathcal{D}^{F[\mathbf{P}]} \Rightarrow 1\right] - \Pr\left[\mathcal{R} \xleftarrow{\$} \mathsf{Func}(n); \mathcal{D}^{\mathcal{R}} \Rightarrow 1\right] .$$

Note that the probabilities are taken over $\mathbf{P}, \mathcal{R}$ and $\mathcal{D}$.

3 LightMAC_Plus

3.1 Construction

Let $\{E_K\}_{K \in \mathcal{K}} \subseteq \mathsf{Perm}(n)$ be a family of n-bit permutations (or a blockcipher) indexed by the key space $\mathcal{K}$, where $k > 0$ is the key length. Let m be the counter size with $m < n$. Let $K, K_1, K_2 \in \mathcal{K}$ be three keys for E. For a message M, the response of $\mathtt{LightMAC\_Plus}[E_K, E_{K_1}, E_{K_2}]$ is defined by Algorithm 1. Figure 1 illustrates the subroutine $\mathsf{Hash}[E_K]$. Here, $M \| 10^*$ means that first 1 is appended to M, and if the bit length of $M \| 1$ is not a multiple of $n - m$ bits, then a sequence of the minimum number of zeros is appended to $M \| 1$ so that the bit length becomes a multiple of $n - m$ bits. Note that $M \| 10^* = M_1 \| M_2 \| \cdots \| M_l$ and $\forall i \in [l] : |M_i| = n - m$. By the counter size m and the padding value 10^*, the maximum message length in bits is at most $(2^m - 1) \times (n - m) - 1$ bit.

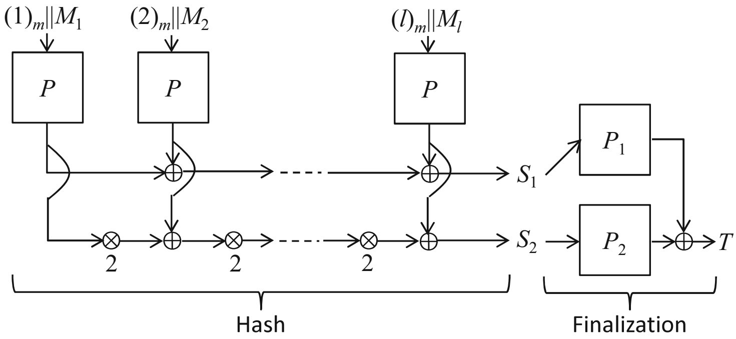


Fig. 1. `LightMAC_Plus` where $P := E_K$, $P_1 := E_{K_1}$ and $P_2 := E_{K_2}$.

Algorithm 1. `LightMAC_Plus`

▶ Main Procedure `LightMAC_Plus`$[E_K, E_{K_1}, E_{K_2}](M)$

$(S_1, S_2) \leftarrow \mathsf{Hash}[E_K](M)$

$T_1 \leftarrow E_{K_1}(S_1)$; $T_2 \leftarrow E_{K_2}(S_2)$; $T \leftarrow T_1 \oplus T_2$

return T

▶ Subroutine $\mathsf{Hash}[E_K](M)$

Partition $M \| 10^*$ into $n - m$-bit blocks $M_1, \ldots, M_l$; $S_1 \leftarrow 0^n$; $S_2 \leftarrow 0^n$

for $i = 1, \ldots, l$ **do**

  $B_i \leftarrow (i)_m \| M_i$; $C_i \leftarrow E_K(B_i)$; $S_1 \leftarrow S_1 \oplus C_i$; $S_2 \leftarrow S_2 \oplus 2^{l-i} \cdot C_i$

end for

return (S_1, S_2)

3.2 Security

We prove the PRF-security of `LightMAC_Plus` in the information-theoretic model, namely, E_K, E_{K_1} and E_{K_2} are replaced with random permutations P, P_1 and P_2, respectively. The upper-bound of the PRF-security advantage is given below, and the security proof is given in Sect. 4.

Theorem 1. *Let $\mathcal{D}$ be a distinguisher making q tagging queries. Then we have*

$$\mathbf{Adv}^{\mathsf{prf}}_{\texttt{LightMAC\_Plus}[P,P_1,P_2]}(\mathcal{D}) \leq \frac{2q^2}{2^{2n}} + \frac{4q^3}{2^{2n}} \ .$$

4 Proof of Theorem 1

Let $F =$ `LightMAC_Plus`. In this section, we upper-bound the PRF-advantage

$$\mathbf{Adv}^{\mathsf{prf}}_{F[P,P_1,P_2]}(\mathcal{D}) = \Pr[(P, P_1, P_2) \xleftarrow{\$} \mathsf{Perm}(n)^3; \mathcal{D}^{F[P,P_1,P_2]} \Rightarrow 1] - \Pr[\mathcal{R} \xleftarrow{\$} \mathsf{Func}(n); \mathcal{D}^{\mathcal{R}} \Rightarrow 1] \ .$$

Without loss of generality, we assume that $\mathcal{D}$ is deterministic and makes no repeated query.

In this proof, we use the following notations. For $\alpha \in [q]$, values defined at the α-th query are denoted by using the superscript character of α such as $B_i^\alpha, C_i^\alpha, S_i^\alpha$, etc., and the message length l at the α-th query is denoted by l_α. For $\alpha \in [q]$ and $j \in [2]$, $\mathsf{Dom}P_j^\alpha := \bigcup_{\delta=1}^{\alpha}\{S_j^\delta\}$, $\mathsf{Rng}P_j^\alpha := \bigcup_{\delta=1}^{\alpha}\{T_j^\delta\}$ and $\overline{\mathsf{Rng}P_j^\alpha} := \{0,1\}^n \backslash \mathsf{Rng}P_j^\alpha$.

4.1 Proof Strategy

This proof largely depends on the so-called game-playing technique [4,5]. In this proof, a random permutation P used in Hash is defined before starting the game, whereas other random permutations P_1 and P_2 are realized by lazy sampling. Before starting the game, for $i \in [2]$, all responses of P_i are not defined, that is, $\forall S_i \in \{0,1\}^n : P_i(S_i) = \perp$. When $P_i(S_i^\alpha)$ becomes necessary, if $P_i(S_i^\alpha) = \perp$ (or $S_i^\alpha \notin \mathsf{Dom}P_i^{\alpha-1}$), then it is defined as $P_i(S_i^\alpha) \xleftarrow{\$} \overline{\mathsf{Rng}P_i^{\alpha-1}}$, and otherwise, $P_i(S_i^\alpha)$ is not updated.

The main game is given in Fig. 2, where there are three sub-cases (See lines 2–4 in Fig. 2) and these procedures are defined in Fig. 3. The analysis of Case C is based on the proofs of $\textsc{sum}^2$ construction by Lucks [23] and SUM-ECBC by Yasuda [36]. We say a set $\mathsf{Fair}^\alpha \subseteq (\{0,1\}^n)^2$ is fair if for each $T \in \{0,1\}^n$,

$$|\{(T_1, T_2) \in \mathsf{Fair}^\alpha \mid T_1 \oplus T_2 = T\}| = \frac{|\mathsf{Fair}^\alpha|}{2^n}.$$

Let $L^\alpha = \overline{\mathsf{Rng}P_1^{\alpha-1}} \times \overline{\mathsf{Rng}P_2^{\alpha-1}}$. Lucks pointed out that at the α-th query, there exists a set $W \subset L^\alpha$ of size at most $(\alpha-1)^2$ such that $L^\alpha \backslash W$ is fair. In Case C, the fair set is defined as $\mathsf{Fair}^\alpha := L^\alpha \backslash W$. Hence, the α-th output ($T^\alpha = T_1^\alpha \oplus T_2^\alpha$) is uniformly random over $\{0,1\}^n$ as long as $(T_1^\alpha, T_2^\alpha) \in \mathsf{Fair}^\alpha$. See Lemma 2 of [23] or [36] for explicit constructions of fair sets.

Initialization

1: $P \xleftarrow{\$} \mathsf{Perm}(n)$
2: $\forall i \in [2], S_i \in \{0,1\}^n : P_i(S_i) \leftarrow \perp$

Main Game: **Upon** the α-th query M^α **do**

1: $(S_1^\alpha, S_2^\alpha) \leftarrow \mathsf{Hash}[P](M^\alpha)$
2: **If** $S_1^\alpha \in \mathsf{Dom}P_1^{\alpha-1}$ and $S_2^\alpha \in \mathsf{Dom}P_2^{\alpha-1}$ **then** goto **Case A**
3: **If** ($S_1^\alpha \in \mathsf{Dom}P_1^{\alpha-1}$ and $S_2^\alpha \notin \mathsf{Dom}P_2^{\alpha-1}$) or ($S_1^\alpha \notin \mathsf{Dom}P_1^{\alpha-1}$ and $S_2^\alpha \in \mathsf{Dom}P_2^{\alpha-1}$) **then** goto **Case B**
4: **If** $S_1^\alpha \notin \mathsf{Dom}P_1^{\alpha-1}$ and $S_2^\alpha \notin \mathsf{Dom}P_2^{\alpha-1}$ **then** goto **Case C**
5: **return** T^α

Fig. 2. Main game.

Case A:

If $\neg\mathsf{bad}$ **then** $\mathsf{bad_A} \leftarrow \mathsf{true}$
$T^\alpha \xleftarrow{\$} \{0,1\}^n$
$\boxed{T_1^\alpha \leftarrow P_1(S_1^\alpha);\ T_2^\alpha \leftarrow P_2(S_2^\alpha);\ T^\alpha \leftarrow T_1^\alpha \oplus T_2^\alpha}$ ▷ Removed in the ideal world

Case B: In the following procedure, $S_j^\alpha \in \mathsf{Dom}P_j^{\alpha-1}$ and $S_{j+1}^\alpha \notin \mathsf{Dom}P_{j+1}^{\alpha-1}$, where $j \in [2]$ and if $j = 2$ then $j+1$ is regarded as 1.

$T_{j+1}^\alpha \xleftarrow{\$} \{0,1\}^n$
if $T_{j+1}^\alpha \in \mathsf{Rng}P_{j+1}^{\alpha-1}$ **then**
  if $\neg\mathsf{bad}$ **then** $\mathsf{bad_B} \leftarrow \mathsf{true}$
  $\boxed{T_{j+1}^\alpha \xleftarrow{\$} \overline{\mathsf{Rng}P_{j+1}^{\alpha-1}}}$ ▷ Removed in the ideal world
end if
$P_{j+1}(S_{j+1}^\alpha) \leftarrow T_{j+1}^\alpha;\ T_j^\alpha \leftarrow P_j(S_j^\alpha);\ T^\alpha \leftarrow T_1^\alpha \oplus T_2^\alpha$

Case C:

Choose a fair set $\mathsf{Fair}^\alpha \subseteq \overline{\mathsf{Rng}P_1^{\alpha-1}} \times \overline{\mathsf{Rng}P_2^{\alpha-1}}$
$(T_1^\alpha, T_2^\alpha) \xleftarrow{\$} \overline{\mathsf{Rng}P_1^{\alpha-1}} \times \overline{\mathsf{Rng}P_2^{\alpha-1}};\ T^\alpha \leftarrow T_1^\alpha \oplus T_2^\alpha$
if $(T_1^\alpha, T_2^\alpha) \notin \mathsf{Fair}^\alpha$ **then**
  if $\neg\mathsf{bad}$ **then** $\mathsf{bad_C} \leftarrow \mathsf{true}$
  $\boxed{\boxed{(T_1^\alpha, T_2^\alpha) \xleftarrow{\$} \mathsf{Fair}^\alpha;\ T^\alpha \leftarrow T_1^\alpha \oplus T_2^\alpha}}$ ▷ Removed in the real world
end if
$P_1(S_1^\alpha) \leftarrow T_1^\alpha;\ P_2(S_2^\alpha) \leftarrow T_2^\alpha$

Fig. 3. Case A, **Case B** and **Case C**.

Let $\mathsf{bad} = \mathsf{bad_A} \vee \mathsf{bad_B} \vee \mathsf{bad_C}$. By the fundamental lemma of game-playing [4,5], we have

$$\mathbf{Adv}^{\mathsf{prf}}_{F[P,P_1,P_2]}(\mathcal{D}) \leq \Pr[\mathsf{bad}] \leq \Pr[\mathsf{bad_A}] + \Pr[\mathsf{bad_B}] + \Pr[\mathsf{bad_C}]. \tag{1}$$

Hereafter, we upper-bound $\Pr[\mathsf{bad_A}]$, $\Pr[\mathsf{bad_B}]$ and $\Pr[\mathsf{bad_C}]$.

4.2 Upper-Bound of $\Pr[\mathsf{bad_A}]$

First we define the following event:

$$\mathsf{coll} \Leftrightarrow \exists \alpha, \beta \in [q] \text{ with } \alpha \neq \beta \text{ s.t. } (S_1^\alpha, S_2^\alpha) = (S_1^\beta, S_2^\beta).$$

Then we have

$$\Pr[\mathsf{bad_A}] \leq \Pr[\mathsf{coll}] + \Pr[\mathsf{bad_A} | \neg\mathsf{coll}] \ .$$

By Propositions 1 and 2, we have

$$\Pr[\mathsf{bad_A}] \leq \frac{2q^2}{2^{2n}} + \frac{\frac{4}{3}q^3}{2^{2n}} \ . \tag{2}$$

Proposition 1. $\Pr[\mathsf{coll}] \leq \frac{2q^2}{2^{2n}}$.

Proof. Lemma 1 shows the upper-bound of the probability that for distinct two messages $M^\alpha, M^\beta \in \{0,1\}^*$, $\mathsf{Hash}[P](M^\alpha) = \mathsf{Hash}[P](M^\beta)$, which is at most $4/2^{2n}$. The sum of the upper-bounds for all combinations of message pairs gives

$$\Pr[\mathsf{coll}] \leq \binom{q}{2} \cdot \frac{4}{2^{2n}} \leq \frac{2q^2}{2^{2n}} \ .$$

□

Lemma 1. *For distinct two messages $M^\alpha, M^\beta \in \{0,1\}^*$, the probability that $\mathsf{Hash}[P](M^\alpha) = \mathsf{Hash}[P](M^\beta)$ is at most $4/2^{2n}$.*

Proof. Without loss of generality, we assume that $l_\alpha \leq l_\beta$. $\mathsf{Hash}[P](M^\alpha) = \mathsf{Hash}[P](M^\beta)$ implies that

$$S_1^\alpha = S_1^\beta \text{ and } S_2^\alpha = S_2^\beta \Leftrightarrow$$

$$\underbrace{\bigoplus_{i=1}^{l_\alpha} C_i^\alpha \oplus \bigoplus_{i=1}^{l_\beta} C_i^\beta}_{A_{3,1}} = 0^n \quad \text{and} \quad \underbrace{\bigoplus_{i=1}^{l_\alpha} 2^{l_\alpha - i} \cdot C_i^\alpha \oplus \bigoplus_{i=1}^{l_\beta} 2^{l_\beta - i} \cdot C_i^\beta}_{A_{3,2}} = 0^n. \qquad (3)$$

We consider the following three cases.

1. $\left(l_\alpha = l_\beta\right) \wedge \left(\exists a \in [l_\alpha] \text{ s.t. } B_a^\alpha \neq B_a^\beta\right) \wedge \left(\forall i \in [l_\alpha] \backslash \{a\} : B_i^\alpha = B_i^\beta\right)$.
2. $\left(l_\alpha = l_\beta\right) \wedge \left(\exists a_1, a_2 \in [l_\alpha] \text{ s.t. } B_{a_1}^\alpha \neq B_{a_1}^\beta \wedge B_{a_2}^\alpha \neq B_{a_2}^\beta\right)$
3. $\left(l_\alpha \neq l_\beta\right)$

The first case is that there is just one position a where the inputs are distinct, whereas the second case is that there are at least two positions a_1, a_2 where the inputs are distinct. For each case, we upper-bound the probability that (3) is satisfied.

– Consider the first case: $\exists a \in [l_\alpha]$ s.t. $B_a^\alpha \neq B_a^\beta$ and $\forall i \in [l_\alpha] \backslash \{a\} : B_i^\alpha = B_i^\beta$. Since $B_a^\alpha \neq B_a^\beta \Rightarrow C_a^\alpha \neq C_a^\beta$ and $B_i^\alpha = B_i^\beta \Rightarrow C_i^\alpha = C_i^\beta$, $A_{3,1} \neq 0^n$ and $A_{3,2} \neq 0^n$. Hence, the probability that (3) is satisfied is 0.
– Consider the second case: $\exists a_1, a_2, \ldots, a_j \in [l_\alpha]$ with $j \geq 2$ s.t. $\forall i \in [j]$: $B_{a_i}^\alpha \neq B_{a_i}^\beta$. Note that $B_{a_i}^\alpha \neq B_{a_i}^\beta \Rightarrow C_{a_i}^\alpha \neq C_{a_i}^\beta$. Eliminating the same outputs between $\{C_i^\alpha : 1 \leq i \leq l_\alpha\}$ and $\{C_i^\beta : 1 \leq i \leq l_\beta\}$, we have

$$A_{3,1} = \bigoplus_{i=1}^{j} \left(C_{a_i}^\alpha \oplus C_{a_i}^\beta\right) \text{ and } A_{3,2} = \bigoplus_{i=1}^{j} 2^{l_\alpha - a_i} \cdot \left(C_{a_i}^\alpha \oplus C_{a_i}^\beta\right) \ .$$

  Since in $A_{3,1}$ and $A_{3,2}$ there are at most $l_\alpha + l_\beta$ outputs, the numbers of possibilities for $C_{a_1}^\alpha$ and $C_{a_2}^\alpha$ are at least $2^n - (l_\alpha + l_\beta - 2)$ and $2^n - (l_\alpha + l_\beta - 1)$, respectively. Fixing other outputs, the equations in (3) provide a unique solution for $C_{a_1}^\alpha$ and $C_{a_2}^\alpha$. As a result, the probability that (3) is satisfied is at most $1/(2^n - (l_\alpha + l_\beta - 2))(2^n - (l_\alpha + l_\beta - 1))$.

– Consider the third case. Without loss of generality, assume that $l_\alpha < l_\beta$. Eliminating the same outputs between $\{C_i^\alpha : 1 \le i \le l_\alpha\}$ and $\{C_i^\beta : 1 \le i \le l_\beta\}$, we have

$$A_{3,1} = \bigoplus_{i=1}^{u} C_{a_i}^\alpha \oplus \bigoplus_{i=1}^{v} C_{b_i}^\beta \ ,$$

where $a_1, \dots, a_u \in [l_\alpha]$ and $b_1, \dots, b_v \in [l_\beta]$. By $l_\alpha < l_\beta$, $l_\beta \in \{b_1, \dots, b_v\}$ and $l_\beta \neq 1$. Since in $A_{3,1}$ and $A_{3,2}$ there are at most $l_\alpha + l_\beta$ outputs, the numbers of possibilities for C_1^β and $C_{l_\beta}^\beta$ are at least $2^n - (l_\alpha + l_\beta - 2)$ and $2^n - (l_\alpha + l_\beta - 1)$, respectively. Fixing other outputs, the equations in (3) provide a unique solution for C_1^β and $C_{l_\beta}^\beta$. As a result, the probability that (3) is satisfied is at most $1/(2^n - (l_\alpha + l_\beta - 2))(2^n - (l_\alpha + l_\beta - 1))$.

The above upper-bounds give

$$\Pr\left[\mathsf{Hash}[P](M^\alpha) = \mathsf{Hash}[P](M^\beta)\right] \le \frac{1}{(2^n - (l_\alpha + l_\beta))^2} \le \frac{4}{2^{2n}} \ ,$$

assuming $l_\alpha + l_\beta \le 2^{n-1}$.

$\square$

Proposition 2. $\Pr[\mathsf{bad_A} | \neg\mathsf{coll}] \le \frac{\frac{4}{3}q^3}{2^{2n}}$.

Proof. First, fix $\alpha \in [q]$ and $\beta, \gamma \in [\alpha - 1]$ with $\beta \neq \gamma$ (from the condition $\neg\mathsf{coll}$), and upper-bound the probability that $S_1^\alpha = S_1^\beta \wedge S_2^\alpha = S_2^\gamma$, which implies

$$\underbrace{\bigoplus_{i=1}^{l_\alpha - 1} C_i^\alpha \oplus \bigoplus_{i=1}^{l_\beta - 1} C_i^\beta}_{A_{4,1}} = 0^n \text{ and } \underbrace{\bigoplus_{i=1}^{l_\alpha - 1} 2^{l_\alpha - i} \cdot C_i^\alpha \oplus \bigoplus_{i=1}^{l_\gamma - 1} 2^{l_\gamma - i} \cdot C_i^\gamma}_{A_{4,2}} = 0^n. \quad (4)$$

Since M^α, M^β and M^γ are distinct, there are at least two distinct outputs $C^{\alpha,\beta}$ and $C^{\alpha,\gamma}$ where $C^{\alpha,\beta}$ appears in $A_{4,1}$ and $C^{\alpha,\gamma}$ appears in $A_{4,2}$. Fixing other outputs in $A_{4,1}$ and $A_{4,2}$, the equations in (4) provide a unique solution for $C^{\alpha,\beta}$ and $C^{\alpha,\gamma}$. Since there are at most $l_\alpha + l_\beta$ outputs in $A_{4,1}$, the number of possibilities for $C^{\alpha,\beta}$ is at least $2^n - (l_\alpha + l_\beta - 1)$. Since there are at most $l_\alpha + l_\gamma$ outputs in $A_{4,2}$, the number of possibilities for $C^{\alpha,\gamma}$ is at least $2^n - (l_\alpha + l_\gamma - 1)$. Hence, the probability that (4) is satisfied is at most

$$\frac{1}{(2^n - (l_\alpha + l_\beta - 1))(2^n - (l_\alpha + l_\gamma - 1))} \le \frac{4}{2^{2n}} \ ,$$

assuming $l_\alpha + l_\beta - 1 \le 2^{n-1}$ and $l_\alpha + l_\gamma - 1 \le 2^{n-1}$.

Finally, we just run induces α, β, and γ to get

$$\begin{aligned}\Pr[\mathsf{bad_A} | \neg\mathsf{coll}] &\le \sum_{\alpha=1}^{q} \left(\sum_{\beta,\gamma \in [1,\alpha-1] \text{ s.t. } \beta \neq \gamma} \frac{4}{2^{2n}} \right) \le \sum_{\alpha=1}^{q} \frac{4(\alpha-1)^2}{2^{2n}} = \sum_{\alpha=1}^{q-1} \frac{4\alpha^2}{2^{2n}} \\ &\le \frac{4}{2^{2n}} \times \frac{q(q-1)(2q-1)}{6} \le \frac{\frac{4}{3}q^3}{2^{2n}} \ .\end{aligned}$$

$\square$

4.3 Upper-Bound of $\Pr[\mathsf{bad_B}]$

First, fix $\alpha \in [q]$ and $j \in [2]$, and upper-bound the probability that $\mathcal{D}$ sets $\mathsf{bad_B}$ at the α-th query, namely, $S_j^\alpha \in \mathsf{Dom}P_j^{\alpha-1}$, $S_{j+1}^\alpha \notin \mathsf{Dom}P_{j+1}^{\alpha-1}$, and $T_{j+1}^\alpha \in \mathsf{Rng}P_{j+1}^{\alpha-1}$. Note that if $j = 2$ then $j+1$ is regarded as 1.

– Regarding $S_j^\alpha \in \mathsf{Rng}P_j^{\alpha-1}$, fix $\beta \in [\alpha-1]$ and consider the case that $S_j^\alpha = S_j^\beta$. Since $M^\alpha \neq M^\beta$, there is an output $C^{\alpha,\beta}$ in $\{C_1^\alpha, \ldots, C_{l_\alpha}^\alpha, C_1^\beta, \ldots, C_{l_\beta}^\beta\}$ that is distinct from other outputs. Fixing other outputs, $S_j^\alpha = S_j^\beta$ provides a unique solution for $C^{\alpha,\beta}$. There are at most $2^n - (l_\alpha + l_\beta - 1)$ possibilities for $C^{\alpha,\beta}$. Hence, the probability that $S_j^\alpha \in \mathsf{Dom}P_j^{\alpha-1}$ is at most $|\mathsf{Dom}P_j^{\alpha-1}| \times 1/(2^n - (l_\alpha + l_\beta - 1)) \leq 2(\alpha-1)/2^n$, assuming $l_\alpha + l_\beta - 1 \leq 2^{n-1}$.
– Regarding $T_{j+1}^\alpha \in \mathsf{Rng}P_{j+1}^{\alpha-1}$, T_{j+1}^α is randomly drawn from $\{0,1\}^n$ after $S_j^\alpha \in \mathsf{Rng}P_j^{\alpha-1}$ and $S_{j+1}^\alpha \notin \mathsf{Dom}P_{j+1}^{\alpha-1}$ are satisfied. In this case, T_{j+1}^α is defined independently from S_j^α and S_{j+1}^α. Since $|\mathsf{Rng}P_{j+1}^{\alpha-1}| \leq \alpha - 1$, this probability that $T_{j+1}^\alpha \in \mathsf{Rng}P_{j+1}^{\alpha-1}$ is at most $(\alpha-1)/2^n$.

Hence, the probability that $\mathcal{D}$ sets $\mathsf{bad_B}$ at the α-th query is upper-bounded by the multiplication of the above probabilities, which is $\frac{2(\alpha-1)^2}{2^{2n}}$.

Finally, we just run induces α and j to get

$$\Pr[\mathsf{nosol}] \leq \sum_{\alpha=1}^{q} \sum_{j=1}^{2} \frac{2(\alpha-1)^2}{2^{2n}} \leq \frac{\frac{4}{3}q^3}{2^{2n}} \ . \tag{5}$$

4.4 Upper-Bound of $\Pr[\mathsf{bad_C}]$

For each $\alpha \in [q]$, since $\left|\overline{\mathsf{Rng}P_1^{\alpha-1}} \times \overline{\mathsf{Rng}P_2^{\alpha-1}} \backslash \mathsf{Fair}^\alpha\right| \leq (\alpha-1)^2$, the probability that $(T_1^\alpha, T_2^\alpha) \notin \mathsf{Fair}^\alpha$ is at most

$$\frac{(\alpha-1)^2}{(2^n - (\alpha-1))^2} \leq \frac{4(\alpha-1)^2}{2^{2n}} \ ,$$

assuming $\alpha - 1 \leq 2^{n-1}$. Hence, we have

$$\Pr[\mathsf{bad_C}] \leq \sum_{\alpha=1}^{q} \frac{4(\alpha-1)^2}{2^{2n}} = \sum_{\alpha=1}^{q-1} \frac{4(\alpha-2)^2}{2^{2n}} \leq \frac{\frac{4}{3}q^3}{2^{2n}} \ . \tag{6}$$

4.5 Conclusion of Proof

Putting (2), (5) and (6) into (1) gives

$$\mathbf{Adv}^{\mathsf{prf}}_{F[P,P_1,P_2]}(\mathcal{D}) \leq \frac{2q^2}{2^{2n}} + \frac{\frac{4}{3} \cdot q^3}{2^{2n}} + \frac{\frac{4}{3}q^3}{2^{2n}} + \frac{\frac{4}{3}q^3}{2^{2n}} \leq \frac{2q^2}{2^{2n}} + \frac{4q^3}{2^{2n}} \ .$$

Algorithm 2. $\texttt{LightMAC\_Plus2}[H_{K_H}, E_{K_{0,1}}, E_{K_{0,2}}, E_{K_1}, \ldots, E_{K_t}]$

▶ Main Procedure $\texttt{LightMAC\_Plus2}[H_{K_H}, E_{K_{0,1}}, E_{K_{0,2}}, E_{K_1}, \ldots, E_{K_t}](M)$

$(S_1, S_2) \leftarrow H_{K_H}(M)$

$R_1 \leftarrow E_{K_{0,1}}(S_1)$; $R_2 \leftarrow E_{K_{0,2}}(S_2)$; $T \leftarrow 0^n$

for $i = 1, \ldots, t$ **do**

  $X_i \leftarrow R_1 \oplus 2^{i-1} \cdot R_2$; $Y_i \leftarrow E_{K_i}(X_i)$; $T \leftarrow T \oplus Y_i$

end for

return T

5 LightMAC_Plus2

5.1 Construction

Let $\mathcal{K}$, $\mathcal{K}_H$ and $\mathsf{Dom}H$ be three non-empty sets. Let $\{E_K\}_{K \in \mathcal{K}} \subset \mathsf{Perm}(n)$ be a family of n-bit permutations (or a blockcipher) indexed by key space $\mathcal{K}$. Let $\{H_{K_H}\}_{K_H \in \mathcal{K}_H}$ be a family of hash functions: $\mathsf{Dom}H \to \{0,1\}^{2n}$ indexed by key space $\mathcal{K}_H$. Let m be the counter size with $m < n$. Let $K_{0,1}, K_{0,2}, K_1, \ldots, K_t \in \mathcal{K}$ be the E's keys and $K_H \in \mathcal{K}_H$ the hash key. For a message M, the response of $\texttt{LightMAC\_Plus2}[H_{K_H}, E_{K_{0,1}}, E_{K_{0,2}}, E_{K_1}, \ldots, E_{K_t}]$ is defined by Algorithm 2, where $|S_1| = n$ and $|S_2| = n$. The finalization function is illustrated in Fig. 4.

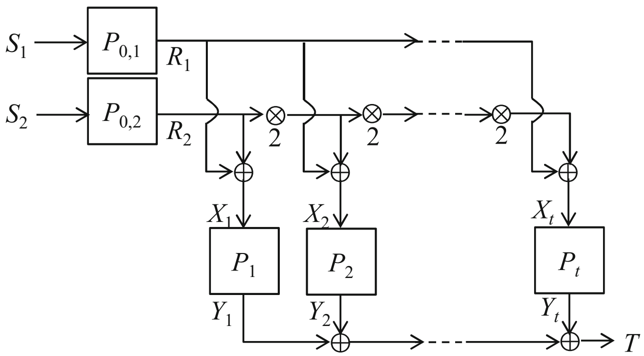


Fig. 4. Finalization function of $\texttt{LightMAC\_Plus2}$, where $P_{0,1} := E_{K_{0,1}}, P_{0,2} := E_{K_{0,2}}, P_1 := E_{K_1}, \ldots, P_t := E_{K_t}$.

5.2 Almost Universal Hash Function

In the security proof, we assume that the hash function H is an almost universal (AU) hash function. The definition is given below.

Definition 1. *Let $\epsilon > 0$. H is an ϵ-AU hash function if for any two distinct messages $M, M' \in \mathsf{Dom}H$, $\Pr[K_H \xleftarrow{\$} \mathcal{K}_H; H_{K_H}(M) = H_{K_H}(M')] \leq \epsilon$.*

5.3 Security

We prove the PRF-security of LightMAC_Plus2 in the information-theoretic model, where permutations $E_{K_{0,1}}, E_{K_{0,2}}, E_{K_1}, \ldots, E_{K_{t-1}}$ and E_{K_t} are replaced with random permutations $P_{0,1}, P_{0,2}, P_1, \ldots, P_{t-1}$ and P_t, respectively, and H is assumed to be an ϵ-AU hash function, where a key is drawn as $K_H \xleftarrow{\$} \mathcal{K}_H$. The upper-bound of the PRF-security advantage is given below, and the security proof is given in Sect. 6.

Theorem 2. *Assume that $t \leq 7$. Let H is an ϵ-AU hash function. Let $\mathcal{D}$ be a distinguisher making q tagging queries. Then we have*

$$\mathbf{Adv}^{\mathsf{prf}}_{\mathtt{LightMAC\_Plus2}[H_{K_H}, P_{0,1}, P_{0,2}, P_1, \ldots, P_{t-1}, P_t]}(\mathcal{D}) \leq 0.5q^2\epsilon + \frac{2^t q^{t+1}}{(2^n - q)^t} \enspace .$$

Define the hash function as $H_{K_H} := \mathsf{Hash}[P]$ (given in Algorithm 1). By Lemma 1, Hash is a $4/2^{2n}$-AU hash function, where $\mathcal{K}_H = \mathsf{Perm}(n)$ and $K_H = P$. Hence, combining Lemma 1 and Theorem 2, the following corollary is obtained.

Corollary 1. *Let $H_{K_H} := \mathsf{Hash}[P]$. Then we have*

$$\mathbf{Adv}^{\mathsf{prf}}_{\mathtt{LightMAC\_Plus2}[H_{K_H}, P_{0,1}, P_{0,2}, P_1, \ldots, P_{t-1}, P_t]}(\mathcal{D}) \leq \frac{2q^2}{2^{2n}} + \frac{2^t q^{t+1}}{(2^n - q)^t} \enspace .$$

6 Proof of Theorem 2

Assume that $t \leq 7$. Let $F = \mathtt{LightMAC\_Plus2}$ and $\mathbf{P} = (P_{0,1}, P_{0,2}, P_1, \ldots, P_t)$. In this section, we upper-bound the PRF-advantage

$$\begin{aligned}\mathbf{Adv}^{\mathsf{prf}}_{F[H_{K_H}, \mathbf{P}]}(\mathcal{D}) =& \Pr[\mathbf{P} \xleftarrow{\$} \mathsf{Perm}(n)^{t+2}; K_H \xleftarrow{\$} \mathcal{K}_H; \mathcal{D}^{F[H_{K_H}, \mathbf{P}]} \Rightarrow 1] \\ &- \Pr[\mathcal{R} \xleftarrow{\$} \mathsf{Func}(n); \mathcal{D}^{\mathcal{R}} \Rightarrow 1] \enspace .\end{aligned}$$

Without loss of generality, we assume that $\mathcal{D}$ is deterministic and makes no repeated query.

In this proof, we use the following notations. For $\alpha \in [q]$, values defined at the α-th query are denoted by using the superscript of α such as $B_i^\alpha, C_i^\alpha, S_i^\alpha$, etc., and the message length l at the α-th query is denoted by l_α. For $\alpha \in [q]$ and $j \in [t]$, $\mathsf{Dom}P_j^\alpha := \bigcup_{\delta=1}^{\alpha}\{X_j^\delta\}$, $\mathsf{Rng}P_j^\alpha := \bigcup_{\delta=1}^{\alpha}\{Y_j^\delta\}$ and $\overline{\mathsf{Rng}P_j^\alpha} := \{0,1\}^n \backslash \mathsf{Rng}P_j^\alpha$.

6.1 Proof Strategy

This proof uses the same strategy as in the proof of Theorem 1 (given in Subsect. 4.1). In this proof, random permutations $P_{0,1}$ and $P_{0,2}$ are defined before starting the game, whereas other random permutations are realized by lazy sampling. The main game is given in Fig. 5, where there are three sub-cases defined by inputs to random permutations $X_1^\alpha, \dots, X_t^\alpha$ (See lines 4–6 in Fig. 5). The sub-cases are given in Fig. 6. Note that for $i \in [t]$, "X_i^α is new" means that $X_i^\alpha \notin \mathsf{Dom}P_i^{\alpha-1}$, and "$X_i^\alpha$ is not new" means that $X_i^\alpha \in \mathsf{Dom}P_i^{\alpha-1}$.

As is the case with the proof of Theorem 1, Case C uses a fair set for the xor of s random permutations with $s \geq 2$. For s random permutations $P_{a_1}, \dots, P_{a_s}$ at the α-th query, we say a set $\mathsf{Fair}^\alpha \subseteq (\{0,1\}^n)^s$ is fair if for each $T \in \{0,1\}^n$,

$$\left| \left\{ (Y_{a_1}, Y_{a_2}, \dots, Y_{a_s}) \in \mathsf{Fair}^\alpha \,\middle|\, \bigoplus_{i \in [s]} Y_{a_i} = T \right\} \right| = \frac{|\mathsf{Fair}^\alpha|}{2^n}.$$

Let $L^\alpha := \overline{\mathsf{Rng}P_{a_1}^{\alpha-1}} \times \overline{\mathsf{Rng}P_{a_2}^{\alpha-1}} \times \cdots \times \overline{\mathsf{Rng}P_{a_s}^{\alpha-1}}$. Lucks [23] pointed out that when s is even, there exists a set $W \subset L^\alpha$ of size at most $(\alpha-1)^s$ such that $L^\alpha \backslash W$ is fair, and when s is odd, there exists a set $W' \subset (\{0,1\}^n)^s$ of size at most $(\alpha-1)^s$ with $W' \cap L^\alpha = \emptyset$ such that $W' \cup L^\alpha$ is fair. See Lemma 2 of [23] or [36] for explicit constructions of fair sets. In Case C, the fair set is defined as $\mathsf{Fair}^\alpha := L^\alpha \backslash W$ when s is even; $\mathsf{Fair}^\alpha := L^\alpha \cup W'$ when s is odd.

Let $\mathsf{bad} = \mathsf{bad_A} \vee \mathsf{bad_B} \vee \mathsf{bad_C}$. Then by the fundamental lemma of game-playing [4,5], we have

$$\mathbf{Adv}_{F[\mathbf{P}]}^{\mathsf{prf}}(\mathcal{D}) \leq \Pr[\mathsf{bad}] \leq \Pr[\mathsf{bad_A}] + \Pr[\mathsf{bad_B}] + \Pr[\mathsf{bad_C}]. \tag{7}$$

Hereafter, we upper-bound $\Pr[\mathsf{bad_A}]$, $\Pr[\mathsf{bad_B}]$ and $\Pr[\mathsf{bad_C}]$.

Initialization

1: $K_H \xleftarrow{\$} \mathcal{K}_H$; $(P_{0,1}, P_{0,2}) \xleftarrow{\$} \mathsf{Perm}(n)^2$
2: $\forall i \in [t], X_i \in \{0,1\}^n : P_i(X_i) \leftarrow \perp$

Main Game: **Upon** the α-th query M^α **do**

1: $(S_1^\alpha, S_2^\alpha) \leftarrow H_{K_H}(M^\alpha)$
2: $R_1^\alpha \leftarrow P_{0,1}(S_1^\alpha)$; $R_2 \leftarrow P_{0,2}(S_2^\alpha)$;
3: **for** $i \in [t]$ **do** $X_i^\alpha = R_1^\alpha \oplus \left(2^{i-1} \cdot R_2^\alpha\right)$
4: **if** all of $X_1^\alpha, \dots, X_t^\alpha$ are not new **then** goto **Case A**
5: **if** one of $X_1^\alpha, \dots, X_t^\alpha$ is new **then** goto **Case B**
6: **if** two ore more of $X_1^\alpha, \dots, X_t^\alpha$ are new **then** goto **Case C**
7: **return** T^α

Fig. 5. Main Game.

Case A:

if $\neg\mathsf{bad}$ **then** $\mathsf{bad_A} \leftarrow \mathsf{true}$
$T^\alpha \xleftarrow{\$} \{0,1\}^n$
for $i \in [t]$ **do** $Y_i^\alpha \leftarrow P_i(X_i^\alpha)$
$\boxed{T^\alpha \leftarrow \bigoplus_{i=1}^{t} Y_i^\alpha}$ ▷ Removed in the ideal world

Case B: In the following procedure, X_a^α is new, and for all $i \in [t]\backslash\{a\}$ X_i^α is not new.

$Y_a^\alpha \xleftarrow{\$} \{0,1\}^n$
if $Y_a^\alpha \in \mathsf{Rng}P_a^{\alpha-1}$ **then**
    if $\neg\mathsf{bad}$ **then** $\mathsf{bad_B} \leftarrow \mathsf{true}$
    $\boxed{Y_a^\alpha \xleftarrow{\$} \overline{\mathsf{Rng}P_a^{\alpha-1}}}$ ▷ Removed in the ideal world
end if
$P_i(X_a^\alpha) \leftarrow Y_a^\alpha$
for $i \in [t]\backslash\{a\}$ **do** $Y_i^\alpha \leftarrow P_i(X_i^\alpha)$
$T^\alpha \leftarrow \bigoplus_{i=1}^{t} Y_i^\alpha$

Case C: In the following procedure, $X_{a_1}^\alpha, \ldots, X_{a_s}^\alpha$ are new with $a_1, \ldots a_s \in [t]$ and other inputs are not new where $s \geq 2$.

$L^\alpha \leftarrow \overline{\mathsf{Rng}P_{a_1}^{\alpha-1}} \times \overline{\mathsf{Rng}P_{a_2}^{\alpha-1}} \times \cdots \times \overline{\mathsf{Rng}P_{a_s}^{\alpha-1}}$
if s is even **then**
    Choose a fair set $\mathsf{Fair}^\alpha \subseteq L^\alpha$; $(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \xleftarrow{\$} L^\alpha$
    if $(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \notin \mathsf{Fair}^\alpha$ **then**
        if $\neg\mathsf{bad}$ **then** $\mathsf{bad_C} \leftarrow \mathsf{true}$
        $\boxed{\boxed{(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \xleftarrow{\$} \mathsf{Fair}^\alpha}}$ ▷ Removed in the real world
    end if
end if
if s is odd **then**
    Choose a fair set $\mathsf{Fair}^\alpha \supseteq L^\alpha$; $(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \xleftarrow{\$} \mathsf{Fair}^\alpha$
    if $(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \notin L^\alpha$ **then**
        if $\neg\mathsf{bad}$ **then** $\mathsf{bad_C} \leftarrow \mathsf{true}$
        $\boxed{(Y_{a_1}^\alpha, Y_{a_2}^\alpha, \ldots, Y_{a_s}^\alpha) \xleftarrow{\$} L^\alpha}$ ▷ Removed in the ideal world
    end if
end if
for $i \in [s]$ **do** $P_i(X_{a_i}^\alpha) \leftarrow Y_{a_i}^\alpha$
for $i \in [t]\backslash\{a_1, \ldots, a_s\}$ **do** $Y_i^\alpha \leftarrow P_i(X_i^\alpha)$
$T^\alpha \leftarrow \bigoplus_{i=1}^{t} Y_i^\alpha$

Fig. 6. Case A, **Case B** and **Case C**.

6.2 Upper-Bound of $\Pr[\mathsf{bad_A}]$

First we define the following event:

$$\mathsf{coll} \Leftrightarrow \exists \alpha, \beta \in [q] \text{ with } \alpha \neq \beta \text{ s.t. } (S_1^\alpha, S_2^\alpha) = (S_1^\beta, S_2^\beta).$$

Then we have

$$\Pr[\mathsf{bad_A}] \le \Pr[\mathsf{coll}] + \Pr[\mathsf{bad_A}|\neg\mathsf{coll}] \ ,$$

Regarding $\Pr[\mathsf{coll}]$, since H is an ϵ-AU hash function, the sum of ϵ for all combinations of message pairs gives

$$\Pr[\mathsf{coll}] \le \binom{q}{2} \cdot \epsilon \le 0.5q^2\epsilon \ .$$

Regarding $\Pr[\mathsf{bad_A}|\neg\mathsf{coll}]$, for $\alpha \in [q]$, Lemma 2 gives the upper-bound of the probability that all of $X_1^\alpha, \dots, X_t^\alpha$ are not new, which is $\left(\frac{\alpha-1}{2^n-q}\right)^t$. Then, we run the index α to get

$$\Pr[\mathsf{bad_A}|\neg\mathsf{coll}] \le \sum_{\alpha=1}^{q} \left(\frac{\alpha-1}{2^n-q}\right)^t = \sum_{\alpha=1}^{q-1} \left(\frac{\alpha}{2^n-q}\right)^t \ .$$

Finally we have

$$\Pr[\mathsf{bad_A}] \le 0.5q^2\epsilon + \sum_{\alpha=1}^{q-1} \left(\frac{\alpha}{2^n-q}\right)^t \ . \tag{8}$$

Lemma 2. *Assume that* coll *does not occur. Fix* $\alpha \in [q]$, $s \le t$ *and* $a_1, a_2, \dots, a_s \in [t]$ *such that* $a_1, a_2, \dots, a_s$ *are distinct. Then the probability that* $\forall i \in [s]$*:* $X_{a_i}^\alpha$ *is not new, that is,* $\exists \beta_i \in [\alpha-1]$ *s.t.* $X_{a_i}^\alpha = X_{a_i}^{\beta_i}$ *is at most* $\left(\frac{\alpha-1}{2^n-q}\right)^s$.

Proof. First, fix $\beta_1, \dots, \beta_s \in [\alpha-1]$, and upper-bound the probability that

$$\underbrace{\forall i \in [s] : X_{a_i}^\alpha \oplus X_{a_i}^{\beta_i}}_{A_9} = 0^n. \tag{9}$$

By Lemma 3, we have only to consider the case where $\beta_1, \dots, \beta_s$ are distinct. Thus if $\alpha \le s$, then this probability is 0. In the following, we consider the case: $\alpha > s$. Note that A_9 is defined as

$$\begin{aligned} X_{a_i}^\alpha \oplus X_{a_i}^{\beta_i} &= \left(R_1^\alpha \oplus 2^{a_i-1} \cdot R_2^\alpha\right) \oplus \left(R_1^{\beta_i} \oplus 2^{a_i-1} \cdot R_2^{\beta_i}\right) \\ &= \left(R_1^\alpha \oplus R_1^{\beta_i}\right) \oplus 2^{a_i-1} \cdot \left(R_2^\alpha \oplus R_2^{\beta_i}\right) \ , \end{aligned}$$

where $R_1^\alpha = P_{0,1}(S_1^\alpha)$, $R_2^\alpha = P_{0,2}(S_2^\alpha)$, $R_1^{\beta_i} = P_{0,1}(S_1^{\beta_i})$ and $R_2^{\beta_i} = P_{0,2}(S_2^{\beta_i})$. Then, the number of independent random variables in $\{R_1^\alpha, R_1^{\beta_1}, \dots, R_1^{\beta_s}, R_2^\alpha, R_2^{\beta_1}, \dots, R_2^{\beta_s}\}$ that appear in A_9 is counted. Note that $\{R_1^\alpha, R_1^{\beta_1}, \dots, R_1^{\beta_s}\}$ are independently defined from $\{R_2^\alpha, R_2^{\beta_1}, \dots, R_2^{\beta_s}\}$.

First, the number of independent random variables in $\{R_1^{\beta_1}, \dots, R_1^{\beta_s}\}$ and $\{R_2^{\beta_1}, \dots, R_2^{\beta_s}\}$ is counted. By $\neg\mathsf{coll}$, for all $i, j \in [s]$ with $i \ne j$,

$(S_1^{\beta_i}, S_2^{\beta_i}) \neq (S_1^{\beta_j}, S_2^{\beta_j})$, that is, $(R_1^{\beta_i}, R_2^{\beta_i}) \neq (R_1^{\beta_j}, R_2^{\beta_j})$. Note that if there are z_1 (resp., z_2) independent random variables in $\{R_1^{\beta_1}, \ldots, R_1^{\beta_s}\}$ (resp., $\{R_2^{\beta_1}, \ldots, R_2^{\beta_s}\}$), then the number of distinct pairs for (R_1, R_2) is $z_1 \cdot z_2$ and the number of distinct random variables is $z_1 + z_2$. If $(z_1 \leq 2 \wedge z_2 \leq 2)$ or $(z_1 = 1 \wedge z_2 \leq 4)$, then $z_1 \cdot z_2 \leq z_1 + z_2$, and if $z_1 = 2$ and $z_2 = 3$, then $z_1 + z_2 = 5 < z_1 \cdot z_2 = 6$. Since $s \leq z_1 \cdot z_2$, the sum of the numbers of independent random variables in $\{R_1^{\beta_1}, \ldots, R_1^{\beta_s}\}$ and in $\{R_2^{\beta_1}, \ldots, R_2^{\beta_s}\}$ is at least $\min\{5, s\}$.

By Lemma 4, we have only to consider the case that $\forall i \in [s] : R_1^{\alpha} \neq R_1^{\beta_i}$ and $R_2^{\alpha} \neq R_2^{\beta_i}$. Hence, the number of independent random variables in $\{R_1^{\beta_1}, \ldots, R_1^{\beta_s}\}$ and $\{R_2^{\beta_1}, \ldots, R_2^{\beta_s}\}$ is at least $s \leq \min\{5, s\} + 2$. By $s \leq t \leq 7$, there are at least s independent random variables in A_9.

Fixing other outputs in A_9 except for the s outputs, the equations in (9) provide a unique solution for the s outputs. The number of possibilities for the s outputs are at least $2^n - s$. Hence, the probability that (9) is satisfied is at most $(1/(2^n - s))^s$.

Finally, the probability that $\forall i \in [s] : \exists \beta_i \in [\alpha - 1]$ s.t. $X_{a_i}^{\alpha} = X_{a_i}^{\beta_i}$ is at most

$$(\alpha - 1)^s \cdot \left(\frac{1}{2^n - s}\right)^s \leq \left(\frac{\alpha - 1}{2^n - q}\right)^s .$$

□

Lemma 3. *Assume that* coll *does not occur. For $\alpha, \beta \in [q]$ with $\alpha \neq \beta$, if there exists $j \in [t]$ such that $X_j^{\alpha} = X_j^{\beta}$, then for all $i \in [t] \backslash \{j\}$, $X_i^{\alpha} \neq X_i^{\beta}$.*

Proof. Assume that $X_j^{\alpha} = X_j^{\beta}$, which implies

$$X_j^{\alpha} = X_j^{\beta} \Leftrightarrow R_1^{\alpha} \oplus R_1^{\beta} = 2^{j-1} \cdot \left(R_2^{\alpha} \oplus R_2^{\beta}\right) .$$

By $\neg$coll, $R_1^{\alpha} \oplus R_1^{\beta} \neq 0^n$ and $R_2^{\alpha} \oplus R_2^{\beta} \neq 0^n$. Then, for any $i \in [t] \backslash \{j\}$

$$\begin{aligned} X_i^{\alpha} \oplus X_i^{\beta} &= \left(R_1^{\alpha} \oplus R_1^{\beta}\right) \oplus 2^{i-1} \cdot \left(R_2^{\alpha} \oplus R_2^{\beta}\right) \\ &= \left(2^{j-1} \oplus 2^{i-1}\right) \cdot \left(R_2^{\alpha} \oplus R_2^{\beta}\right) \neq 0^n , \end{aligned}$$

namely, $X_i^{\alpha} \neq X_i^{\beta}$.

□

Lemma 4. *For $\alpha, \beta \in [q]$ with $\alpha \neq \beta$, if $(R_1^{\alpha} \neq R_1^{\beta} \wedge R_2^{\beta} = R_2^{\beta})$ or $(R_1^{\alpha} = R_1^{\beta} \wedge R_2^{\beta} \neq R_2^{\beta})$, then for all $i \in [t]$ $X_i^{\alpha} \neq X_i^{\beta}$.*

Proof. Let $\alpha, \beta \in [q]$ with $\alpha \neq \beta$. If $R_1^{\alpha} \neq R_1^{\beta} \wedge R_2^{\beta} = R_2^{\beta}$, then for any $i \in [t]$,

$$X_i^{\alpha} \oplus X_i^{\beta} = \left(R_1^{\alpha} \oplus 2^{i-1} \cdot R_2^{\alpha}\right) \oplus \left(R_1^{\beta} \oplus 2^{i-1} \cdot R_2^{\beta}\right) = R_1^{\alpha} \oplus R_1^{\beta} \neq 0^n.$$

If $R_1^{\alpha} = R_1^{\beta} \wedge R_2^{\beta} \neq R_2^{\beta}$, then for any $i \in [t]$,

$$X_i^{\alpha} \oplus X_i^{\beta} = \left(R_1^{\alpha} \oplus 2^{i-1} \cdot R_2^{\alpha}\right) \oplus \left(R_1^{\beta} \oplus 2^{i-1} \cdot R_2^{\beta}\right) = 2^{i-1} \cdot \left(R_2^{\alpha} \oplus \cdot R_2^{\beta}\right) \neq 0^n.$$

□

6.3 Upper-Bound of $\Pr[\mathsf{bad_B}]$

First, fix $\alpha \in [q]$ and $a \in [t]$, and upper-bound the probability that

$$X_a^\alpha \text{ is new}, \ \underbrace{\forall i \in [t]\backslash\{a\} : X_i^\alpha \text{ is not new}}_{A_{10,2}}, \text{ and } \underbrace{Y_a^\alpha \in \mathsf{Rng}P_a^{\alpha-1}}_{A_{10,3}}. \quad (10)$$

Regarding $A_{10,2}$, by Lemma 2, the probability that $A_{10,2}$ is satisfied is at most $\left(\frac{\alpha-1}{2^n-q}\right)^{t-1}$. Regarding $A_{10,3}$, since Y_a^α is randomly drawn and $|\mathsf{Rng}P_a^{\alpha-1}| \leq \alpha-1$, the probability that $A_{10,3}$ is satisfied is at most $\frac{\alpha-1}{2^n}$. Hence the probability that (10) is satisfied is at most

$$\left(\frac{\alpha-1}{2^n-q}\right)^{t-1} \cdot \frac{\alpha-1}{2^n} \leq \left(\frac{\alpha-1}{2^n-q}\right)^{t} .$$

Finally, we run induces α and a to get

$$\Pr[\mathsf{bad_B}] \leq \sum_{\alpha=1}^{q}\sum_{a=1}^{t}\left(\frac{\alpha-1}{2^n-q}\right)^{t} \leq \sum_{\alpha=1}^{q-1} t \cdot \left(\frac{\alpha}{2^n-q}\right)^{t} . \quad (11)$$

6.4 Upper-Bound of $\Pr[\mathsf{bad_C}]$

First, fix $\alpha \in [q]$, $s \in \{2, \ldots, t\}$ and $a_1, \ldots, a_s \in [t]$ such that $a_1, \ldots, a_s$ are distinct, and consider the case that

$$X_{a_1}^\alpha, \ldots, X_{a_{s-1}}^\alpha \text{ and } X_{a_s}^\alpha \text{ are new}, \ \underbrace{\forall i \in [t]\backslash\{a_1, \ldots, a_s\} : X_i^\alpha \text{ is not new}}_{A_{12,2}}, \text{ and}$$

$$\underbrace{(Y_{a_1}^\alpha, \ldots, Y_{a_{s-1}}^\alpha, Y_{a_s}^\alpha) \notin \mathsf{Fair}^\alpha \text{ if } s \text{ is even}; (Y_{a_1}^\alpha, \ldots, Y_{a_{s-1}}^\alpha, Y_{a_s}^\alpha) \notin L^\alpha \text{ if } s \text{ is odd}}_{A_{12,3}}. \quad (12)$$

Regarding $A_{12,2}$, by Lemma 2, the probability that $A_{12,2}$ is satisfied is at most $\left(\frac{\alpha-1}{2^n-q}\right)^{t-s}$. Regarding $A_{12,3}$, if s is even, then since $|L^\alpha\backslash\mathsf{Fair}^\alpha| \leq (\alpha-1)^s$, the probability that $A_{12,3}$ is satisfied is at most $\left(\frac{\alpha-1}{2^n-q}\right)^{s}$; if s is odd, then since $|\mathsf{Fair}^\alpha\backslash L^\alpha| \leq (\alpha-1)^s$, the probability that $A_{12,3}$ is satisfied is at most $\left(\frac{\alpha-1}{2^n-q}\right)^{s}$. Hence, the probability that the conditions in (12) are satisfied is at most

$$\left(\frac{\alpha-1}{2^n-q}\right)^{t-s} \cdot \left(\frac{\alpha-1}{2^n-q}\right)^{s} = \left(\frac{\alpha-1}{2^n-q}\right)^{t} .$$

Finally, we run induces α and s to get

$$\Pr[\mathsf{bad_C}] \leq \sum_{\alpha=1}^{q}\sum_{s=2}^{t}\left(\binom{t}{s} \cdot \left(\frac{\alpha-1}{2^n-q}\right)^{t}\right) = \sum_{s=2}^{t}\binom{t}{s} \cdot \left(\sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^{t}\right) . \quad (13)$$

6.5 Conclusion of Proof

Putting (8), (11) and (13) into (7) gives

$$\begin{aligned}
&\mathbf{Adv}^{\mathsf{prf}}_{F[H_{K_H},\mathbf{P}]}(\mathcal{D})\\
&\le 0.5q^2\epsilon + \sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^t + t\cdot\sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^t + \sum_{s=2}^{t}\binom{t}{s}\left(\sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^t\right)\\
&\le 0.5q^2\epsilon + \sum_{s=0}^{t}\binom{t}{s}\cdot\left(\sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^t\right) = 0.5q^2\epsilon + 2^t\cdot\left(\sum_{\alpha=1}^{q-1}\left(\frac{\alpha}{2^n-q}\right)^t\right)\\
&= 0.5q^2\epsilon + \sum_{\alpha=1}^{q-1}\left(\frac{2\alpha}{2^n-q}\right)^t \le 0.5q^2\epsilon + \frac{2^t q^{t+1}}{(2^n-q)^t}\ ,
\end{aligned}$$

where the last term uses the fact that $\sum_{\alpha=1}^{x}\alpha^t \le x^{t+1}$ for $x \ge 1$ and $t \ge 1$.

7 Improving the Efficiency of Hash

In this section, we consider a hash function Hash^* with better efficiency than Hash. Hash^* is defined in Algorithm 3 and is illustrated in Fig. 7. Here, $M\|10^*$ means that first 1 is appended to M, and if $|M\|1| \le n$, then a sequence of the minimum number of zeros is appended to $M\|1$ so that the length in bits becomes n bit; if $|M\|1| > n$, then a sequence of the minimum number of zeros is appended to $M\|1$ so that the total length minus n becomes a multiple of $n-m$.

The difference between Hash and Hash^* is that in Hash the last block message M_l is input to E_K, while in Hash^* it is not input. Therefore, replacing Hash with Hash^*, the efficiency of `LightMAC_Plus2` is improved.

In Lemma 5, the collision probability of Hash^* is given, where E_K is replaced with a random permutation P. Combining Theorem 2 and Lemma 5 offers the following corollary.

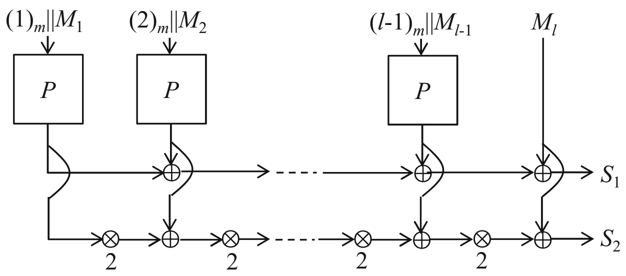


Fig. 7. Hash^*.

Algorithm 3. $\mathsf{Hash}^*[E_K](M) = (S_1, S_2)$

Partition $M\|10^*$ into $n-m$-bit blocks $M_1, \ldots, M_{l-1}$ and n-bit block M_l
$S_1 \leftarrow 0^n$; $S_2 \leftarrow 0^n$
for $i = 1, \ldots, l-1$ **do**
    $B_i \leftarrow (1)_m\|M_i$; $C_i \leftarrow E_K(B_i)$; $S_1 \leftarrow S_1 \oplus C_i$; $S_2 \leftarrow S_2 \oplus 2^{l-i} \cdot C_i$
end for
$S_1 \leftarrow S_1 \oplus M_l$; $S_2 \leftarrow S_2 \oplus M_l$
return (S_1, S_2)

Corollary 2. *Assume that $t \leq 7$. Then we have*

$$\mathbf{Adv}^{\mathsf{prf}}_{\mathtt{LightMAC\_Plus2}[\mathsf{Hash}^*[P], P_{0,1}, P_{0,2}, P_1, \ldots, P_{t-1}, P_t]}(\mathcal{D}) \leq \frac{2q^2}{2^{2n}} + \frac{2^t q^{t+1}}{(2^n - q)^t} \ .$$

Lemma 5. *Let $P \xleftarrow{\$} \mathsf{Perm}(n)$ be a random permutation. For distinct two messages $M^\alpha, M^\beta \in \{0,1\}^*$, the probability that $\mathsf{Hash}^*[P](M^\alpha) = \mathsf{Hash}^*[P](M^\beta)$ is at most $4/2^{2n}$.*

Proof. In this proof, values defined from M^α (resp., M^β) are denoted by using the superscript of α (resp., β), length l of M^α (resp., M^β) is denoted by l_α (resp., l_β). Without loss of generality, we assume that $l_\alpha \leq l_\beta$. $H[P](M^\alpha) = H[P](M^\beta)$ implies that

$$S_1^\alpha = S_1^\beta \text{ and } S_2^\alpha = S_2^\beta \Leftrightarrow$$
$$\underbrace{\bigoplus_{i=1}^{l_\alpha-1} C_i^\alpha \oplus \bigoplus_{i=1}^{l_\beta-1} C_i^\beta = Z^{\alpha,\beta}}_{A_{14,1}} \quad \text{and} \quad \underbrace{\bigoplus_{i=1}^{l_\alpha-1} 2^{l_\alpha-i} \cdot C_i^\alpha \oplus \bigoplus_{i=1}^{l_\beta-1} 2^{l_\beta-i} \cdot C_i^\beta = Z^{\alpha,\beta}}_{A_{14,2}} \quad (14)$$

where $Z^{\alpha,\beta} = M^\alpha_{l_\alpha} \oplus M^\beta_{l_\beta}$. We consider the following six cases.

1. $\left(l_\alpha = l_\beta = 1\right)$
2. $\left(l_\alpha = l_\beta \neq 1\right) \wedge \left(\forall a \in [l_\alpha - 1] \text{ s.t. } B_a^\alpha = B_a^\beta\right) \wedge \left(M_{l_\alpha} \neq M_{l_\beta}\right)$
3. $\left(l_\alpha = l_\beta \neq 1\right) \wedge \left(\exists a \in [l_\alpha - 1] \text{ s.t. } B_a^\alpha \neq B_a^\beta\right) \wedge$ $\left(\forall i \in [l_\alpha - 1] \backslash \{a\} : B_i^\alpha = B_i^\beta\right)$.
4 $\left(l_\alpha = l_\beta \neq 1\right) \wedge \left(\exists a_1, a_2 \in [l_\alpha - 1] \text{ s.t. } B_{a_1}^\alpha \neq B_{a_1}^\beta \wedge B_{a_2}^\alpha \neq B_{a_2}^\beta\right)$
5. $\left(l_\alpha \neq l_\beta\right) \wedge \left(l_\beta = 2\right)$
6. $\left(l_\alpha \neq l_\beta\right) \wedge \left(l_\beta \geq 3\right)$

Note that by $l_\alpha \leq l_\beta$, when $l_\alpha \neq l_\beta$, $l_\beta \neq 1$, thereby we do not have to consider the case of $\left(l_\alpha \neq l_\beta\right) \wedge \left(l_\beta = 1\right)$. The third case is that there is just one position

a where the inputs are distinct, whereas the fourth case is that there are at least two positions a_1, a_2 where the inputs are distinct. For each case we evaluate the probability that the equalities in (14) hold.

- Consider the first and second cases. In these cases, $A_{14,1} = A_{14,2} = 0^n$ and $Z^{\alpha,\beta} \neq 0^n$. Hence (14) is not satisfied.
- Consider the third case. In this case, $A_{14,1} = (C_a^\alpha \oplus C_a^\beta) \neq 2^{l_\alpha - a} \cdot (C_a^\alpha \oplus C_a^\beta) = A_{14,2}$. Hence, in (14) is not satisfied.
- Consider the fourth case. First we eliminate the same outputs between $\{C_i^\alpha, 1 \leq i \leq l_\alpha - 1\}$ and $\{C_i^\beta, 1 \leq i \leq l_\beta - 1\}$ from $A_{14,1}$ and $A_{14,2}$, and then we have

$$A_{14,1} = \bigoplus_{i=1}^{j} \left(C_{a_i}^\alpha \oplus C_{a_i}^\beta\right) \text{ and } A_{14,2} = \bigoplus_{i=1}^{j} 2^{l_\alpha - a_i} \cdot \left(C_{a_i}^\alpha \oplus C_{a_i}^\beta\right) ,$$

  where $a_1, \ldots, a_j \in [l_\alpha - 1]$ with $j \geq 2$. Since in $A_{14,1}$ and $A_{14,2}$ there are at most $l_\alpha + l_\beta - 2$ outputs, the numbers of possibilities for $C_{a_1}^\alpha$ and $C_{a_2}^\alpha$ are at least $2^n - (l_\alpha + l_\beta - 3)$ and $2^n - (l_\alpha + l_\beta - 4)$, respectively. Fixing other outputs, the equations in (14) provide a unique solution for $C_{a_1}^\alpha$ and $C_{a_2}^\alpha$. Thus, the probability that (14) is satisfied is at most $1/(2^n - (l_\alpha + l_\beta - 2))(2^n - (l_\alpha + l_\beta - 3))$.
- Consider the fifth case. In this case, $l_\alpha = 1$ and $A_{14,1} = C_1^\beta \neq 2 \cdot C_1^\beta = A_{14,2}$. Hence (14) is not satisfied.
- Consider the sixth case. We eliminate the same outputs between $\{C_i^\alpha : 1 \leq i \leq l_\alpha - 1\}$ and $\{C_i^\beta : 1 \leq i \leq l_\beta - 1\}$ from $A_{14,1}$. By $l_\alpha < l_\beta$, $C_{l_\beta}^\beta$ remains in $A_{14,1}$. Since in $A_{14,1}$ and $A_{14,2}$ there are at most $l_\alpha + l_\beta - 2$ outputs, the numbers of possibilities for $C_{l_\beta}^\beta$ and C_1^β are at least $2^n - (l_\alpha + l_\beta - 3)$ and $2^n - (l_\alpha + l_\beta - 4)$, respectively. Fixing other outputs, the equations in (14) provide a unique solution for $C_{l_\beta}^\beta$ and C_1^β. As a result, the probability of (14) is at most $1/(2^n - (l_\alpha + l_\beta - 3))(2^n - (l_\alpha + l_\beta - 4))$.

Thus, we have

$$\Pr\left[\mathsf{Hash}^*[P](M^\alpha) = \mathsf{Hash}^*[P](M^\beta)\right] \leq \frac{1}{(2^n - (l_\alpha + l_\beta))^2} \leq \frac{4}{2^{2n}} ,$$

assuming $l_\alpha + l_\beta \leq 2^{n-1}$.

□

References

1. Bellare, M., Guérin, R., Rogaway, P.: XOR MACs: new methods for message authentication using finite pseudorandom functions. In: Coppersmith, D. (ed.) CRYPTO 1995. LNCS, vol. 963, pp. 15–28. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-44750-4_2

2. Bellare, M., Kilian, J., Rogaway, P.: The security of cipher block chaining. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 341–358. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48658-5_32
3. Bellare, M., Pietrzak, K., Rogaway, P.: Improved security analyses for CBC MACs. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 527–545. Springer, Heidelberg (2005). https://doi.org/10.1007/11535218_32
4. Bellare, M., Rogaway, P.: Code-based game-playing proofs and the security of triple encryption. Cryptology ePrint Archive, Report 2004/331 (2004). http://eprint.iacr.org/2004/331
5. Bellare, M., Rogaway, P.: The security of triple encryption and a framework for code-based game-playing proofs. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 409–426. Springer, Heidelberg (2006). https://doi.org/10.1007/11761679_25
6. Bernstein, D.J.: How to stretch random functions: the security of protected counter sums. J. Cryptol. **12**(3), 185–192 (1999)
7. Bhargavan, K., Leurent, G.: On the practical (in-)security of 64-bit block ciphers: collision attacks on HTTP over TLS and OpenVPN. In: Weippl, E.R., Katzenbeisser, S., Kruegel, C., Myers, A.C., Halevi, S. (eds.) Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24–28 October 2016, pp. 456–467. ACM (2016)
8. Black, J., Rogaway, P.: CBC MACs for arbitrary-length messages: the three-key constructions. In: Bellare, M. (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 197–215. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44598-6_12
9. Black, J., Rogaway, P.: A block-cipher mode of operation for parallelizable message authentication. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 384–397. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-46035-7_25
10. Brassard, G.: On computationally secure authentication tags requiring short secret shared keys. In: Chaum, D., Rivest, R.L., Sherman, A.T. (eds.) Advances in Cryptology, pp. 79–86. Plenum Press, New York (1982)
11. Cogliati, B., Seurin, Y.: EWCDM: an efficient, beyond-birthday secure, nonce-misuse resistant MAC. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9814, pp. 121–149. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53018-4_5
12. Dodis, Y., Steinberger, J.: Domain extension for MACs beyond the birthday barrier. In: Paterson, K.G. (ed.) EUROCRYPT 2011. LNCS, vol. 6632, pp. 323–342. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-20465-4_19
13. Dutta, A., Nandi, M., Paul, G.: One-key compression function based MAC with security beyond birthday bound. In: Liu, J.K.K., Steinfeld, R. (eds.) ACISP 2016. LNCS, vol. 9722, pp. 343–358. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40253-6_21
14. Gaži, P., Pietrzak, K., Rybar, M.: The exact security of PMAC. Cryptology ePrint Archive, Report 2017/069 (2017). http://eprint.iacr.org/2017/069
15. Iwata, T., Kurosawa, K.: OMAC: one-key CBC MAC. In: Johansson, T. (ed.) FSE 2003. LNCS, vol. 2887, pp. 129–153. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-39887-5_11
16. Iwata, T., Minematsu, K.: Stronger security variants of GCM-SIV. Cryptology ePrint Archive, Report 2016/853, to appear at IACR Transactions on Symmetric Cryptology. http://eprint.iacr.org/2016/853
17. Iwata, T., Minematsu, K., Peyrin, T., Seurin, Y.: ZMAC: a fast tweakable block cipher mode for highly secure message authentication. In: Katz, J., Shacham,

H. (eds.) CRYPTO 2017. LNCS, vol. 10403, pp. 34–65. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63697-9_2

18. Jaulmes, É., Joux, A., Valette, F.: On the security of randomized CBC-MAC beyond the birthday paradox limit a new construction. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 237–251. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45661-9_19
19. Jaulmes, E., Lercier, R.: FRMAC, a fast randomized message authentication code. Cryptology ePrint Archive, Report 2004/166 (2004). http://eprint.iacr.org/2004/166
20. Kurosawa, K., Iwata, T.: TMAC: two-key CBC MAC. In: Joye, M. (ed.) CT-RSA 2003. LNCS, vol. 2612, pp. 33–49. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-36563-X_3
21. Lee, J., Steinberger, J.: Multi-property-preserving domain extension using polynomial-based modes of operation. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 573–596. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13190-5_29
22. List, E., Nandi, M.: Revisiting Full-PRF-secure PMAC and using it for beyond-birthday authenticated encryption. IACR Cryptology ePrint Archive 2016, 1174 (2016)
23. Lucks, S.: The sum of PRPs is a secure PRF. In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 470–484. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-45539-6_34
24. Luykx, A., Preneel, B., Szepieniec, A., Yasuda, K.: On the influence of message length in PMAC's security bounds. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9665, pp. 596–621. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49890-3_23
25. Luykx, A., Preneel, B., Tischhauser, E., Yasuda, K.: A MAC mode for lightweight block ciphers. In: Peyrin, T. (ed.) FSE 2016. LNCS, vol. 9783, pp. 43–59. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-52993-5_3
26. Minematsu, K.: How to thwart birthday attacks against MACs via small randomness. In: Hong, S., Iwata, T. (eds.) FSE 2010. LNCS, vol. 6147, pp. 230–249. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13858-4_13
27. Minematsu, K., Matsushima, T.: New bounds for PMAC, TMAC, and XCBC. In: Biryukov, A. (ed.) FSE 2007. LNCS, vol. 4593, pp. 434–451. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74619-5_27
28. Naito, Y.: Full PRF-secure message authentication code based on tweakable block cipher. In: Au, M.-H., Miyaji, A. (eds.) ProvSec 2015. LNCS, vol. 9451, pp. 167–182. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-26059-4_9
29. Nandi, M.: A unified method for improving PRF bounds for a class of blockcipher based MACs. In: Hong, S., Iwata, T. (eds.) FSE 2010. LNCS, vol. 6147, pp. 212–229. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13858-4_12
30. NIST: Recommendation for Block Cipher Modes of Operation: the CMAC Mode for Authentication. Spp. 800–38B (2005)
31. Petrank, E., Rackoff, C.: CBC MAC for real-time data sources. J. Cryptol. **13**(3), 315–338 (2000)
32. Peyrin, T., Seurin, Y.: Counter-in-tweak: authenticated encryption modes for tweakable block ciphers. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9814, pp. 33–63. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53018-4_2

33. Rogaway, P.: Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In: Lee, P.J. (ed.) ASIACRYPT 2004. LNCS, vol. 3329, pp. 16–31. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30539-2_2
34. Wegman, M.N., Carter, L.: New hash functions and their use in authentication and set equality. J. Comput. Syst. Sci. **22**(3), 265–279 (1981)
35. Yasuda, K.: A double-piped mode of operation for MACs, PRFs and PROs: security beyond the birthday barrier. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 242–259. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01001-9_14
36. Yasuda, K.: The sum of CBC MACs is a secure PRF. In: Pieprzyk, J. (ed.) CT-RSA 2010. LNCS, vol. 5985, pp. 366–381. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-11925-5_25
37. Yasuda, K.: A new variant of PMAC: beyond the birthday bound. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 596–609. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22792-9_34
38. Yasuda, K.: PMAC with parity: minimizing the query-length influence. In: Dunkelman, O. (ed.) CT-RSA 2012. LNCS, vol. 7178, pp. 203–214. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-27954-6_13
39. Yasuda, K.: A parallelizable PRF-based MAC algorithm: well beyond the birthday bound. IEICE Trans. **96–A**(1), 237–241 (2013)
40. Zhang, L., Wu, W., Sui, H., Wang, P.: 3kf9: enhancing 3GPP-MAC beyond the birthday bound. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 296–312. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34961-4_19
41. Zhang, Y.: Using an error-correction code for fast, beyond-birthday-bound authentication. In: Nyberg, K. (ed.) CT-RSA 2015. LNCS, vol. 9048, pp. 291–307. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16715-2_16