One-Key Compression Function Based MAC with BBB Security

Avijit Dutta, Mridul Nandi, Goutam Paul

Indian Statistical Institute, Kolkata 700 108, India.
avirocks.dutta13@gmail.com, mridul.nandi@gmail.com,
goutam.paul@isical.ac.in

Abstract. Gaži et al. [CRYPTO 2014] analyzed the NI-MAC construction proposed by An and Bellare [CRYPTO 1999] and gave a tight birthday-bound of $O(\ell q^2/2^n)$, as an improvement over the previous bound of $O(\ell^2 q^2/2^n)$. In this paper, we design a simple extension of NI-MAC, called NI$^+$-MAC, and prove that it has security bound beyond birthday (BBB) of order $O(q^2\ell^2/2^{2n})$ provided $\ell \leq 2^{n/4}$. Our construction not only lifts the security of NI-MAC beyond birthday, it also reduces the number of keys from 2 (NI uses 2 independent keys) to 1. Before this work, Yasuda had proposed [FSE 2008] a single fixed-keyed compression function based BBB-secure MAC with security bound $O(\ell q^2/2^{2n})$ that uses an extra mask, requires a storage space to store the mask. However, our proposed construction NI$^+$ does not require any extra mask and thereby has reduced the state size compared to Yasuda's proposal [FSE 2008] with providing the same order of security bound for light-weight applications

Keywords: Beyond Birthday, MAC, NI, Structure-Graph.

1 Introduction

In symmetric key paradigm, MAC (Message Authentication Code) is used for preserving message integrity and message origin authentication. The design of a MAC should not only consider achieving security, but also target attaining efficiency. In the literature, three different approaches of designing a MAC exists: (a) universal hash function based MAC, a popular example of which is UMAC [8], (b) a compression function based MAC, like NMAC [2], HMAC [2], NI [1] etc. (c) Block cipher based MAC, such as CBC MAC [4], PMAC [9], OMAC [17]. etc.

Most of the popular MACs are block cipher based MACs, but each one of them suffers from the same problem - security is guaranteed up to the *birthday bound*. When the block length of the underlying block cipher is 128-bit, then birthday bound does not seem to be a problem, as we are guaranteed to have 64 bits of security which is well acceptable for many practical applications. But when we deal with 64-bit block cipher (e.g HIGHT [16], PRESENT [10]) as used in many light weight crypto devices (e.g RFID, smartcard) then birthday bound problem becomes the main bottleneck.

Related Work on NMAC and HMAC. NMAC and its variant HMAC [2] is the first re-keying compression function based MAC where a key is appended to a message and then the appended message is hashed using Merkle-Damgård technique. It has been standardized in [23]. and has become popular and widely used in many network protocols like SSH, IPSec, TLS etc. Bellare et al. in [2] proves that NMAC is a secure PRF based on the assumption (i) $\mathtt{f}$ is a secure PRF and (ii) $\mathtt{Casc}^{\mathtt{f}}$ is a WCR (weakly collision resistant). HMAC when instantiated with MD4 or SHA-1, both of them play the role of $\mathtt{Casc}^{\mathtt{f}}$ then they have been found not to satisfy the WCR property [37, 38] and hence the security of HMAC [2] stands void. To restore the PRF security of NMAC, Bellare in [6] investigates the proof and drops assumption (ii). Koblitz and Menezes in [22] criticizes the way [6] discusses the practical implication of their result against uniform and non-uniform reductions used in the proof.

Dodis et.al in [12] investigates the indifferentiable property of HMAC from a keyed random oracle. In a recent line of researches, generic attack against iterated hash based MAC are being investigated [31, 32, 30, 25]. More recently, Gaži et. al in [14] showed a tight bound on NMAC. There is also a recent result [15] on the generic security analysis of NMAC and HMAC with input whitening.

Yasuda in [40] had proposed a novel way of iterating a compression function dedicated for the use of MAC which is more efficient than standard HMAC to process data much faster. In [42] Yasuda has showed that classical sandwiched construction with Merkle-Damgård iteration based hashing provides a secure MAC which is an alternative for HMAC, useful in situation where the message size is small and high performance is required. A new secret-prefix MAC based on hash functions is presented in [45] which is similar to HMAC but does not require the second key.

U.Maurer et. al in [27] has presented a MAC construction namely PDI, that tranforms any Fixed-Input Length (FIL) MAC to Alternative Input Length (AIL) MAC and investigated the tradeoff between the efficiency of MAC and the tightness of its security reduction. In [28] construction of AIL MAC from a FIL MAC with a single key was presented which is better than NI [1].

Related Work on Beyond Birthday Secure MAC.

• **Block Cipher Based Beyond Birthday Secure MAC.** Recently, many MAC constructions have been proposed with security beyond the birthday barrier without degrading the performance. The first attempt was made in ISO 9797-1 [3] without security proof. But Algorithm 4 of ISO 9797-1 was attacked by Joux et al. [20] that falsified the security bound. Algorithm 6 of ISO 9797-1 was proven to be secure against $O(2^{2n/3})$ queries with restrictions on the message length [46]. In [46] Yasuda also presented SUM-ECBC, a 4-key rate-1/2 construction with beyond birthday bound security. In 2011, Yasuda improved the number of keys and rate over SUM-ECBC and proposed a 3-key rate-1 PMAC_Plus construction [47] with beyond birthday security. In 2012, Zhang et al. [50] proposed a 3key version of f9 MAC (3kf9) that achieves BBB security.

There is also another deterministic MAC mode provides security beyond the birthday bound. Given an n-bit to n-bit fixed-key blockcipher with MAC

security ϵ against q queries, Dodis et al. [13] have designed a variable-length MAC achieving $O(\epsilon q \text{poly}(n))$ MAC security. However, this design requires even longer keys and more block cipher invocations. By parity method, Bellare et al. present MACRX [3] with BBB security, conditioned on the input parameters are random and distinct. In [18], Jaulmes et al. proposed a randomized MAC that provides BBB security based on the ideal model (or possibly based on tweakable block cipher). Another BBB secure randomized construction called generic enhanced hash then MAC has been proposed in [29] by Minematsu. In [24], the authors propose a tweakable block-cipher based two-key rate-2 BBB-secure MAC with security margin of $O(q^2\ell^2/2^n)$. Recently Datta et al. in [11] unify PMAC_Plus and 3kf9 in one key setting with beyond birthday security.

• **Compression Function Based Beyond Birthday Secure MAC.** Besides the block cipher based BBB MAC constructions, Yasuda in [41] proposed a compression function based MAC construction - Multi-lane HMAC, that achieves BBB security. In [44] Yasuda presented a double pipe mode operation (Lucks Construction [26]) for constructing AIL MAC from a FIL MAC that acheives BBB security. This work is further extended to provide full security in [48]. In [43] Yasuda has proposed a fixed single keyed compression function based cascaded MAC in a tweakable setting where the tweaks are some distinct masking keys of b bits. Thus for a l blocks message, one needs to compute l many different masks where the masks are generated from a single mask Δ_0 using the field multiplication. The security of the scheme has been proved to be $O(\ell q^2/2^{2n})$. Further improvement on [43] is followed in [49].

Related Work on Fixed-Key MAC. An et al.in [1] proposed a fixed-keyed compression function based MAC called NI-MAC. The construction of NI-MAC is similar to that of NMAC [2], the only difference is that NI-MAC uses two independent keyed compression functions f_1, f_2. The motivation of designing NI was to avoid constant re-keying on multi-block messages in NMAC and to allow for a security proof starting by the standard switch from a PRF to a random function, followed by information-theoretic analysis.

We mention here that the security proof technique for re-keying compression function based MAC is completely different from that of fixed-keyed compression function based MAC. The security of the former scheme is proved using reduction argument, whereas that of the latter is proved by replacing the fixed-keyed compression function with a random function.

Gaži et al.in [14] revisited the proof of NI-MAC and gave a tight birthday bound of $O(\frac{\ell q^2}{2^n})$, a better bound than earlier $O(\frac{\ell^2 q^2}{2^n})$.

Our Contributions.

• In this paper, we propose a rate-$b/(b+n)$[1], fixed key compression function based MAC NI$^+$, which is an extension of existing NI-MAC, that achieves beyond-birthday security of security bound $O(q^2\ell^2/2^{2n})$, where b is the block length

[1] Rate $\triangleq \frac{b}{rs}$, where b-size of message block, s-total input size of the function without the key part and r is the total number of function calls to process a single message block.

and n is the number of output bits. Our proposed construction not only lifts the security of NI beyond birthday (Sect. 4), but also reduces the number of required keys from two (NI uses two independent keys) to one.

• Yasuda in [43] proposed a one pass mode BBB secure MAC with a security bound of $O(\ell q^2/2^{2n})$. The construction uses an extra b-bit mask Δ_0 for which all the keyed compression functions are eventually turned out to be the tweaked compression functions. Moreover, state size of the proposed construction is $2(b+n)$[2] as one needs to store the b-bit masking value and the b-bit checksum value along with two n bits partial outputs. In this regard, our construction NI^+ is a non-tweakable single-keyed compression function based MAC in which the keyed-compression function takes a $(b+n)$-bit input and produces an n-bit output. Our construction does not use any mask and therefore the state size of NI^+ is $(b+2n)$ as one needs to store b-bit checksum value along with two n bit partial outputs.

However, to compare the state-size of Yasuda's construction with our design, one needs to consider the compression functions with the same input-width, i.e., one needs to replace input width (b) bits of the compression function used in the construction proposed in [43] by $b+n$ bits which gives the state size of Yasuda's scheme to $2(b+n)+2n = 2(b+2n)$, which is twice of our state size.

Though reducing this state-size to $2n$ bits was placed as an open problem in [43, Section 7], our construction has slightly improved the state size, albeit with the cost of an extra factor of ℓ in the security bound. However, we note that this bound is comparable to that of [43] for light-weight applications.

It is to be noted that using tweaked functions, which can be easily replaced by independent uniform random functions, security proof of Yasuda's construction is more or less straight-forward. Whereas due to the absence of tweaked functions in our construction, such replacement cannot be done and therefore the security proof of our construction uses a completely different technique which exploits the structure graph analysis of [5], where we consider more bad cases than that of [43].

In the following table we compare different parameters and the security bound of known BBB secure MACs. We write BC to denote block cipher based MAC in which the underlying primitive is a block cipher and CF_{rk} denotes re-keying compression function based MAC in which the underlying primitive is a compression function (e.g HMAC), CF_{fk} denotes fixed-keyed compression function based MAC (e.g NI).

[2] In [43] author has mistakenly stated the state size for the construction is $b+2n$ bits, without considering the state size required for storing the b-bit mask, thus eventually state size becomes $2(b+n)$.

Construction	Type	# Keys	Rate	Security Bound	State size (#bits)
SUM-ECBC [46]	BC	4	1/2	$O(\ell^3 q^3/2^{2n})$	$2n$
PMAC_Plus [47]	BC	3	1	$O(\ell^3 q^3/2^{2n})$	$4n$
3kf9 [50]	BC	3	1	$O(\ell^3 q^3/2^{2n})$	$2n$
1kf9 [11]	BC	1	1	$O(q^3 \ell^4/2^{2n})$	$2n$
1k_PMAC+ [11]	BC	1	1	$O(q^3 \ell^4/2^{2n})$	$4n$
L-Lane ($L = 2$) HMAC [41]	CF_{rk}	3	1/2	$O(\ell^2 q^2/2^{2n})$	$2n$
1-pass mode [43]	CF_{fk}	1	1	$O(\ell q^2/2^{2n})$	$(2b + 4n)$
NI^+[This paper]	CF_{fk}	1	$b/(b + n)$	$O(\ell^2 q^2/2^{2n})$	$(b + 2n)$

2 Preliminaries

In this section, we briefly discuss the notations and definitions used in this paper. We also state some existing basic results.

We denote $|S|$ as the cardinality of set S. Let $x \xleftarrow{\$} S$ denote that x is chosen uniformly at random from S. $[n]$ denotes the set of integers $\{1, 2, \ldots, n\}$. $(s)_{|n}$ denotes the last n bit substring of b bit string s.

Let M be a binary string over $\{0, 1\}$. Length of M in bits is denoted by $|M|$. When $|M| \mod b \neq 0$, we pad 10^d to M to make $|M| \mod b = 0$ where $d = n - 1 - |M| \mod b$ and b denotes the block length of M. $M_1||M_2||\ldots||M_l$ denotes the partition of message M after M is being padded, where each $M_i \in \{0, 1\}^b$ and l denotes the number of blocks of M. ℓ denotes the maximum number of blocks in a message. By a q-set or a q-tuple $x := (x_i : i \in I)$ for an index set I, we mean a set or a tuple of size q. When all elements $x_i's$ are distinct we write $x \in \mathsf{dist}_q$.

Random Functions. Let $Func(A, B)$ denote the set of all functions from A to B. A **random function** $\mathtt{F}$ is a function which is chosen from $Func(A, B)$ following some distribution, not necessarily uniform. In particular, a function ρ_n is said to be a uniform random function, if ρ_n is chosen uniformly at random from the set of all functions from a specified finite domain $\mathcal{D}$ to $\{0, 1\}^n$. Throughout the paper we fix a positive integer n.

We will specify a uniform random function by performing *lazy sampling*. In lazy sampling, initially the function ρ is undefined at every point of its domain. We maintain a set $\mathrm{Dom}(\rho)$ that grows dynamically to keep the record of already defined domain points of ρ. $\mathrm{Dom}(\rho)$ is initialized to be empty. If $x \notin \mathrm{Dom}(\rho)$ then we will choose $y \xleftarrow{\$} \{0, 1\}^n$ and add x in $\mathrm{Dom}(\rho)$. In this regard, x is said to be *fresh*. On the other hand, if $x \in \mathrm{Dom}(\rho)$ (i.e $x = x'$) then $y \leftarrow f(x')$. In this regard x is said to be *covered*.

2.1 Security Definitions

We consider that an adversary $\mathcal{A}$ is an oracle algorithm with access to its oracle $\mathcal{O}(\cdot)$ and outputs either 1 or 0. Accordingly, we write $\mathcal{A}^{\mathcal{O}(\cdot)} = 1$ or 0. The resource of $\mathcal{A}$ is measured in terms of the time complexity t which takes into account the

time it takes to interacts with its oracle $\mathcal{O}(\cdot)$ and the time for its internal computations, query complexity q takes into account the number of queries asked to the oracle by the adversary, data complexity ℓ takes into account the maximum number of blocks in each query.

Pseudo-Random Function. We define **distinguishing advantage** of an oracle algorithm $\mathcal{A}$ for distinguishing two random functions $\mathtt{F}$ from $\mathtt{G}$ as

$$\mathbf{Adv}_{\mathcal{A}}(\mathtt{F}\ ;\ \mathtt{G}) := \mathsf{Pr}[\mathcal{A}^{\mathtt{F}} = 1] - \mathsf{Pr}[\mathcal{A}^{\mathtt{G}} = 1].$$

We define prf-advantage of $\mathcal{A}$ for an n-bit construction $\mathtt{F}$ by

$$\mathbf{Adv}^{\mathrm{prf}}_{\mathtt{F}}(\mathcal{A}) := \mathbf{Adv}_{\mathcal{A}}(\mathtt{F}\ ;\ \rho_n).$$

We call $\mathcal{A}$ a (q, ℓ, t)-distinguisher if it makes at most q queries with at most ℓ-blocks in each query and runs in time at most t. We write $\mathbf{Adv}^{\mathrm{prf}}_{\mathtt{F}}(q, \ell, t) = \max_{\mathcal{A}} \mathbf{Adv}^{\mathrm{prf}}_{\mathtt{F}}(\mathcal{A})$ where maximum is taken over all (q, ℓ, t)-distinguisher $\mathcal{A}$. In an information theoretic situation we also ignore the time parameter t. We call a keyed construction $\mathtt{F}$ is (q, ℓ, ϵ)-prf if $\mathbf{Adv}^{\mathrm{prf}}_{\mathtt{F}}(q, \ell) \leq \epsilon$. Informally, if ϵ is negligible then $\mathtt{F}$ is said to be a secure PRF.

Collision-Free and Cover-Free. Now we define some other information-theoretic security advantages (in which there is no presence of an adversary). Let $\mathtt{H}$ be a random function which outputs two n bit blocks, denoted by $(\Sigma, \Theta) \in (\{0,1\}^n)^2$. For a q-tuple of distinct messages $\mathcal{M} = (M^1, \ldots, M^q)$, we write $\mathtt{H}(M^i) = (\Sigma^i, \Theta^i)$. For a q-tuple of pairs $(\Sigma^i, \Theta^i)_i$, we say that

1. A tuple $(\Sigma^i, \Theta^i)_i$ is **collided** if $\exists i, j \in [q]$ such that $\Sigma^i = \Sigma^j$ and $\Theta^i = \Theta^j$ for some $j \neq i$. Otherwise the tuple is said to be **collision-free**.
2. A tuple $(\Sigma^i, \Theta^i)_i$ is **covered** if $\exists i, j \in [q]$ such that $\Sigma^i = (M^j_\alpha)_{|n}$ and $\Theta^i = Y^j_{\alpha-1}$ where $\alpha \in [l_i]$ or $\alpha \in [l_j]$ and j could be equal to i, M^j_α denotes the α^{th} block of j^{th} message M^j and $Y^j_{\alpha-1}$ is a n bit binary string that denotes the output of $(\alpha-1)^{th}$ block corresponding to j^{th} message M^j. Otherwise the tuple is said to be **cover-free**.

Definition 1. *We define* (q, ℓ)-collision advantage *and* (q, ℓ)-cover-free advantage *as*

$$\mathbf{Adv}^{\mathrm{coll}}_{\mathtt{F}}(q, \ell) = \max_{M \in \mathsf{dist}_q} \mathsf{Pr}[(\Sigma_i, \Theta_i)_i \textit{ is not collision-free}].$$
$$\mathbf{Adv}^{\mathrm{cf}}_{\mathtt{F}}(q, \ell) = \max_{M \in \mathsf{dist}_q} \mathsf{Pr}[(\Sigma_i, \Theta_i)_i \textit{ is not cover-free}].$$

Clearly, $\mathbf{Adv}^{\mathrm{coll}}_{\mathtt{F}}(q, \ell) \leq \frac{q^2}{2}\mathbf{Adv}^{\mathrm{coll}}_{\mathtt{F}}(2, \ell)$. Similarly, $\mathbf{Adv}^{\mathrm{cf}}_{\mathtt{F}}(q, \ell) \leq \frac{q^2}{2}\mathbf{Adv}^{\mathrm{cf}}_{\mathtt{F}}(2, \ell)$. So it would be sufficient to concentrate on a pair of messages while bounding collision free or cover-free advantages. We say that a construction F is (q, ℓ, ϵ)-xxx if $\mathbf{Adv}^{\mathrm{xxx}}_{\mathtt{F}}(q, \ell) \leq \epsilon$ where xxx denotes either **collision-free** or **cover-free**.

2.2 Structure Graphs

In this section, we briefly revisit the structure graph analysis [5, 14].

Consider a cascaded construction with a function f, where f is a uniform random function, that works on a message $M = M_1||M_2||\ldots||M_l$ of length l blocks as follows:

$Y_0 = \mathbf{0}$, and $Y_i = f(Y_{i-1}, M_i)$ for $i = 1, \ldots, l$ where M_i is the i^{th} block of message M.
Informally, for a set of any two fixed distinct messages $\mathcal{M} = \{M^1, M^2\}$ and a uniformly chosen random function f, we construct the structure graph $\mathcal{G}^f(\mathcal{M})$ with $\{0,1\}^n$ as the set of nodes as follows. We follow the computations for M^1 followed by those of M^2 by creating nodes labelled by the values y_i of the intermediate chaining variables Y_i with the edge (y_i, y_{i+1}) labelled by the block M_{i+1}. In this process, if we arrive at a vertex already labelled, while not following an existing edge, we call this event an f-collision.[3] The sequence of alternating vertices and edges corresponding to the computations for a message M^j is called an M^j-walk or more generally a message walk, denoted by W^j. A more formal discussion on structure graph appears in Appendix A.

Let $\mathcal{G}(\mathcal{M})$ denote the set of all structure graphs corresponding to the set of messages $\mathcal{M}$ (by varying f over a function family). For a fixed graph $G \in \mathcal{G}(\mathcal{M})$, let $fColl(G)$ denote the set of all f-collisions in G. We state the following results.

Proposition 1. *[14, Lemma 2] For a fixed graph G, $\Pr_f[\mathcal{G}^f(\mathcal{M}) = G] \leq 2^{-n|fColl(G)|}$.*

Proposition 2. $\Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : |fColl(G)| \geq 3] \leq \frac{27\ell^6}{2^{2n}}$*, where ℓ is the total number of blocks of the messages in $\mathcal{M}$.*

Proof of the Prosposition can be found in Appendix B.

It is to be noted that for CBC-MAC analysis [5], $f(\alpha, \beta)$ is taken as $\pi(\alpha \oplus \beta)$ and for the NI-MAC analysis [14], $f(\alpha, \beta)$ is taken as $\rho(\alpha || \beta)$, where π is a random permutation over n bits and ρ is a random function from $b + n$ bits to n bits, where b is the message block-length and n is the length of the chaining variable as well as the tag.

3 Proposed Construction of NI$^+$ for Beyond-Birthday Secure MAC

We present the schematic diagram of NI$^+$ in Fig. 3.1 followed by the description in Algorithm 1. Let $f_k : \{0,1\}^{b+n} \to \{0,1\}^n$ be a keyed function from $b + n$

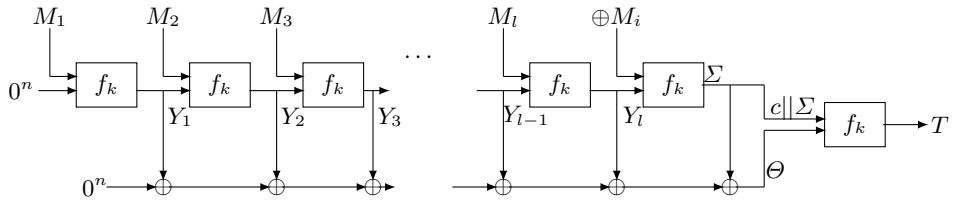

Fig. 3.1. Construction of NI$^+$ MAC

[3] We use the term collision and accident interchangably.

Input: $f_k : k \xleftarrow{\$} \mathcal{K}$, $M \leftarrow \{0,1\}^*, c \leftarrow 10^{b-n-1}$
Output: $T \in \{0,1\}^n$
$M_1||M_2||\dots M_l \leftarrow M||10^*$; //$l$ is the number of message blocks in M
$Z \leftarrow 0^n$; $Y \leftarrow 0^n$;
for $i = 1$ *to* l **do**
  $Y \leftarrow f_k(M_i, Y)$; $Z \leftarrow Z \oplus Y$;
end
$CS \leftarrow \oplus_{i=1}^{l} M_i$;
$Y \leftarrow f_k(CS, Y)$; $Z \leftarrow Z \oplus Y$;
$\Sigma \leftarrow Y$; $\Theta \leftarrow Z$;
$T \leftarrow f_k(c||\Sigma, \Theta)$;
Return T;

Algorithm 1: Algorithm for NI$^+$ MAC

bits to n bits where $b > n$ where b refers to the block length of a message block and n refers to the output length in bits. Let $M \in \{0,1\}^{bl}$. So we can write $M = (M_1, M_2, \dots, M_l)$ where each $M_i \in \{0,1\}^b$. We define a checksum block $CS = \oplus_{i=1}^{l} M_i$. We denote $\mathbf{Casc}^{\mathbf{f_k}}(M) := f_{k_1}(\dots(f_k(f_k(0, M_1), M_2), \dots, M_l)$. Output of $\mathbf{Casc}^{\mathbf{f_k}}(M)$ and the checksum block CS is passed through the same function f_k and the output is denoted as Σ. We obtain Θ by xoring all the intermediate chaining values (i.e $\oplus_{i=1}^{l} Y_i \oplus \Sigma$). We concatenate a fixed $b-n$ bit string $c = 10^{b-n-1}$ with the $2n$ bit string $\Sigma||\Theta$ to match the input size of f_k and then the entire concatenated b bit string (i.e $c||\Sigma||\Theta$) is passed through f_k and finally outputs the tag T. We sometimes denote CS by M_{l+1}.
Note that, NI$^+$ is similar to that of NI upto $\mathbf{Casc}^{\mathbf{f_k}}(M)$ except the following differences. Schematic diagram of NI is given in Appendix C.

(a) In NI construction, b-bit encoding of $|M|$ and the last message block output Y_ℓ is passed through a different keyed compression function f_{k_2}. In NI$^+$, we substitute the b-bit length encoding by the checksum block CS. Moreover, CS and Y_ℓ is passed through the same keyed compression function.

(b) NI is a two fixed-keyed compression function based MAC. NI$^+$ is a single fixed-keyed compression function based MAC.

(c) NI provides only birthday bound ($\ell q^2/2^n$) security. NI$^+$ provides beyond birthday bound security ($q^2\ell^2/2^{2n}$) when $\ell \leq 2^{n/4}$.

Remark 1 *We note that the beyond birthday security is not possible to achieve if we just keep the original structure of NI-MAC and output Σ as the last block output (i.e $\Sigma = f_{K_2}(|M|, Y_l)$) and Θ as the sum of all intermediate chaining variables (i.e $\Theta = \oplus_{i=1}^{l} Y_i \oplus \Sigma$) as the birthday bound attack is followed from Prennnel and Oorschot's attack [33].*

4 Security Analysis of NI$^+$-MAC

Gaži et. al in [14] have shown that the advantage of NI-MAC is bounded above by $\frac{q^2}{2^n}\left(l + \frac{64l^4}{2^n}\right)$. In this section we analyse the advantage of our construction NI$^+$-MAC and show that the advantage of NI$^+$-MAC achieves beyond birthday bound security; better than that of NI-MAC. Thus we have the following theorem.

Theorem 1. *Let* $f : \{0,1\}^k \times \{0,1\}^b \times \{0,1\}^n \to \{0,1\}^n$ *be a* (ϵ, t, q) *secure PRF. Then NI*$^+$ *be a* $\left(\epsilon^{'}, t^{'}, q, l\right)$ *secure PRF, where*

$$\epsilon^{'} \leq \epsilon + \frac{q}{2^n} + \frac{2q^2}{2^{2n}} + \frac{2q^2\ell^2}{2^{2n}} + \frac{2q^2\ell^4}{2^{3n}} + \frac{54q^2\ell^6}{2^{3n}},$$

such that $t = t^{'} + \tilde{O}(lq)$. *Moreover, if* $\ell \leq 2^{n/4}$ *then,*

$$\epsilon^{'} \leq \epsilon + \frac{q}{2^n} + \frac{2q^2\ell^2}{2^{2n}}.$$

Proof. Let $\mathcal{A}$ be a adaptive PRF-adversary against NI$^+$ running in time t and asking at most q queries, each of length at most ℓ blocks. NI$^+$ uses a single keyed function f. Now if we replace f by a uniformly distributed random function r such that $r \xleftarrow{\$} Func(\{0,1\}^b \times \{0,1\}^n, \{0,1\}^n)$ and call the resulting construction NI$_r^+$, then using the standard reduction from information theoretic setting to complexity theoretic setting we have,

$$\mathbf{Adv}^{\mathbf{prf}}_{\mathrm{NI}^+} \leq \epsilon + \mathbf{Adv}^{\mathbf{prf}}_{\mathrm{NI}^+_r}.$$

Therefore to prove Theorem 1, we only need to prove

$$\mathbf{Adv}^{\mathbf{prf}}_{\mathrm{NI}^+_r} \leq \tfrac{q}{2^n} + \tfrac{2q^2}{2^{2n}} + \tfrac{2q^2\ell^2}{2^{2n}} + \tfrac{2q^2\ell^4}{2^{3n}} + \tfrac{54q^2\ell^6}{2^{3n}}.$$

Consider the following Game as shown in Algorithm 2 where the adversary $\mathcal{A}$ queries to oracle O with distinct messages M^i and obtains the response T^i. Note that Game G_0 truly simulates a uniform random function and G_1 simulates the actual construction NI$_r^+$. Therefore using the fundamental lemma of game-playing technique [7], we have the following:

$$\begin{aligned}
\mathbf{Adv}^{\mathbf{prf}}_{\mathrm{NI}^+_r} &= \left|\Pr[\mathcal{A}^{G_1} = 1] - \Pr[\mathcal{A}^{G_0} = 1]\right| \\
&\leq \Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{badsigma} \vee \mathcal{A}^{G_1} \text{ sets } \mathbf{bad}] \\
&\leq \Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{badsigma}] + \Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{bad}]. \qquad (1)
\end{aligned}$$

Therefore, we evaluate now the probability $\Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{bad}]$. To evaluate this, let us define a double block function $\mathcal{H}_f(M) := (\Sigma, \Theta)$ with respect to a uniform random function f. Recall that the tuple $\mathcal{H}_f(M^i) := (\Sigma^i, \Theta^i)_i, \forall i \in [q]$ is said to be collision-free if $\forall i$, either $\Sigma^i \neq \Sigma^j$ or $\Theta^i \neq \Theta^j$ or both $\forall j \in [i-1]$. Similarly, the tuple $(\Sigma^i, \Theta^i)_i$ is said to be cover-free if $\forall i$, either $\Sigma^i \neq (M^j_\alpha)_{|n}$ or $\Theta^i \neq Y^j_{\alpha-1}$ or both $\forall j \in [i]$. Therefore, it is then easy to see that,

$$\begin{aligned}
\Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{bad}] &\leq \mathbf{Adv}^{\mathrm{coll}}_{\mathtt{H}}(q,\ell) + \mathbf{Adv}^{\mathrm{cf}}_{\mathtt{H}}(q,\ell) \\
&\leq \frac{q^2}{2}(\mathbf{Adv}^{\mathrm{coll}}_{\mathtt{H}}(2,\ell) + \mathbf{Adv}^{\mathrm{cf}}_{\mathtt{H}}(2,\ell)). \qquad (2)
\end{aligned}$$

$initialize:$ **badsigma, bad** $\leftarrow false$;
On the j^{th} query M^j;
$M_1^j||M_2^j||\ldots M_l^j \leftarrow M^j||10^* \leftarrow$ **Partition**(M^j), $Y_0 = 0$;
for $i = 1$ to l ;
  if $((M_i^j, Y_{i-1}^j) \in Dom(f))$ $Y_i^j \leftarrow f(M_i^j, Y_{i-1}^j)$;
  Else $Y_i^j \leftarrow \{0,1\}^n$;
  $f(M_i^j, Y_{i-1}^j) \leftarrow Y_i^j$;
  $Dom(f) \leftarrow Dom(f) \cup (M_i^j, Y_{i-1}^j)$;
if $((\oplus_{i=1}^l M_i^j, Y_l^j) \in Dom(f))$ $Y_{l+1}^j \leftarrow f(\oplus_{i=1}^l M_i^j, Y_l^j)$;
Else $Y_{l+1}^j \leftarrow \{0,1\}^n$;
$f(\oplus_{i=1}^l M_i^j, Y_l^j) \leftarrow Y_{l+1}^j$;
$Dom(f) \leftarrow Dom(f) \cup (\oplus_{i=1}^l M_i^j, Y_l^j)$;
$\Sigma^j \leftarrow Y_{l+1}^j$, $\Theta^j \leftarrow \oplus_{i=1}^{l+1} Y_i^j$;
if $(\Sigma^j = 0)$ **badsigma** $\leftarrow true$;
$T^j \xleftarrow{\$} \{0,1\}^n$;
if $((\Sigma^j, \Theta^j) = (\Sigma^i, \Theta^i)$ for some $i \in \{1, 2, \ldots, j-1\}$, or $(c||\Sigma^j, \Theta^j) = (M_s^*, Y_{s-1}^*)$ such that $s \in [l_i + 1]$ or $s \in [l_j + 1], * \in \{i, j\}$);
  if $(\neg$**bad**$)$;
    Coll$(i, j) \leftarrow true$, **bad** $\leftarrow true$;
      if $((\Sigma^j, \Theta^j) = (\Sigma^i, \Theta^i))$ $\boxed{T^j \leftarrow f(\Sigma^i, \Theta^i)}$;
      Else $\boxed{T^j \leftarrow f(M_s^*, Y_{s-1}^*)}$;
Return T^j;

Algorithm 2: Game G_0 is without boxed statement and G_1 is with boxed statement.

Now we state the following lemma, proof of which is deferred until next section. The first three cases of the lemma bound the collision-free advantage and the last three cases bound the cover-free advantage of function $H_f(\cdot)$.

Notation: Let E_{coll} denotes the collision event (i.e. $\Sigma^i = \Sigma^j \wedge \Theta^i = \Theta^j$) and E_{cf} denotes the covered event (i.e. $\Sigma^i = x \wedge \Theta^i = Y_t^s$) for some n bit constant x. W^i denotes the walk graph corresponding to message M^i. $\boldsymbol{Y}$ denotes the vector of intermdiate computations (i.e $(Y_1, Y_2, \ldots, Y_l)$). l_i and l_j denote the message length in number of blocks of M^i and M^j respectively. When M^i is not a prefix of M^j or M^j is not a prefix of M^i, p denotes longest common prefix (LCP) of M^i and M^j. That means $M_{p+1}^i \neq M_{p+1}^j$ and $M_\alpha^i = M_\alpha^j$ where $1 \leq \alpha \leq p$. Let $\mathcal{G}(\mathcal{M}^i, \mathcal{M}^j)$ denotes the set of all structure graphs corresponding to two fixed messages $\mathcal{M}^i$ and $\mathcal{M}^j$. $\mathcal{G}^a \subset \mathcal{G}(\mathcal{M}^i, \mathcal{M}^j)$ be the set of all structure graphs with accident a where, in this paper, we consider $a = 0, 1, 2$. Moreover, when $a = 1$ or 2 we denote $\mathcal{G}_{nl}^a \subset \mathcal{G}_a$ be the set of all structure graphs such that none of the two message walks W^i, W^j contains a loop. $\mathcal{G}_l^a$ denotes the set of all remaining sturcture graphs. Moreover, $\mathcal{G}^a = \mathcal{G}_{nl}^a \sqcup \mathcal{G}_l^a$ for $a = 1, 2$.

Lemma 1. *Let us consider* $G \xleftarrow{\$} \mathcal{G}(\mathcal{M}^i, \mathcal{M}^j)$*, where* M^i *and* M^j *are any two distinct messages, each of length at most* ℓ *blocks and a particualr* n *bit constant* x*, we have the followings:*

Case (A) : $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$.
Case (B) : $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 1] \leq \frac{\ell^2}{2^{2n}}$.
Case (C) : $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 2] \leq \frac{\ell^4}{2^{3n}}$.
Case (D) : $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$.
Case (E) : $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 1] \leq \frac{\ell^2}{2^{3n}}$.
Case (F) : $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 2] \leq \frac{\ell^4}{2^{3n}}$.

Resume the proof of Theorem 1: Now we have all the materials to prove Theorem 1 which is given in the following.

It is easy to see the followings:

$$\mathbf{Adv}^{\text{coll}}_{\mathtt{H}}(2, \ell) \leq \sum_{k=0}^{2} \Pr[E_{\text{coll}} \wedge |fColl(G)| = k] + \Pr[|fColl(G)| \geq 3].$$

$$\mathbf{Adv}^{\text{cf}}_{\mathtt{H}}(2, \ell) \leq \sum_{k=0}^{2} \Pr[E_{\text{cf}} \wedge |fColl(G)| = k] + \Pr[|fColl(G)| \geq 3].$$

Therefore, we have the following results,

$$\mathbf{Adv}^{\text{coll}}_{\mathtt{H}}(2, \ell) \leq \frac{1}{2^{2n}} + \frac{\ell^2}{2^{2n}} + \frac{\ell^4}{2^{3n}} + \frac{27\ell^6}{2^{3n}}. \tag{3}$$

$$\mathbf{Adv}^{\text{cf}}_{\mathtt{H}}(2, \ell) \leq \frac{1}{2^{2n}} + \frac{\ell^2}{2^{2n}} + \frac{\ell^4}{2^{3n}} + \frac{27\ell^6}{2^{3n}}. \tag{4}$$

Equation (3) follows from Case (A),(B) and (C) of Lemma 1. Similarly, Equation (4) follows from Case (D),(E) and (F) of Lemma 1.
Substituting Equation (3) and (4) into Equation (2) we obtain

$$\Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{bad}] \leq \frac{2q^2}{2^{2n}} + \frac{2q^2\ell^2}{2^{2n}} + \frac{2q^2\ell^4}{2^{3n}} + \frac{54q^2\ell^6}{2^{3n}}.$$

Moreover it is easy to see that $\Pr[\mathcal{A}^{G_1} \text{ sets } \mathbf{badsigma}] \leq \frac{q}{2^n}$. Therefore, substituting these two probability expressions back to Equation (1) will give

$$\mathbf{Adv}^{\mathbf{prf}}_{\text{NI}^+_r} \leq \frac{q}{2^n} + \frac{2q^2}{2^{2n}} + \frac{2q^2\ell^2}{2^{2n}} + \frac{2q^2\ell^4}{2^{3n}} + \frac{54q^2\ell^6}{2^{3n}}. \qquad \square$$

4.1 Proof of Lemma 1

We prove all the following cases using structure graph analysis. After fixing two distinct messages we choose a structure graph uniformly at random from the set of all structure graphs. Then we analyse mainly two events E_{coll} and E_{cf} in view

of the number of collisions occured in the randomly chosen structure graph G. Therefore, we have,

$$\Pr[E_{\text{coll}} \wedge |fColl(G)| = a] = \sum_{H \in \mathcal{G}^a} \Pr[E_{\text{coll}} \wedge G = H]$$
$$\Pr[E_{\text{cf}} \wedge |fColl(G)| = a] = \sum_{H \in \mathcal{G}^a} \Pr[E_{\text{cf}} \wedge G = H]$$

It is easy to see that $|\mathcal{G}^a| \leq \ell^{2a}$ as structure graph is uniquely determined by the number of accidents occured in the graph when the two messages are fixed. Therefore, we only need to bound (i) $\Pr[E_{\text{coll}} \wedge G = H]$ and (ii) $\Pr[E_{\text{cf}} \wedge G = H]$ for some fixed structure graph H having accident a where we consider $a = 0, 1$ or 2.

Case (A) : Proof of $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$. We fix a structure graph $H \in \mathcal{G}^0$ and then analyse the probability of the event E_{coll} with respect to H in a case-by-case basis.

Case (i) When M^i or M^j is not a prefix of each other, we recall that p be the LCP of M^i and M^j. Therefore, all Y^i_α and Y^j_β are distinct where $p+1 \leq \alpha \leq l_i, p+1 \leq \beta \leq l_j$. Moreover, $Y^i_\alpha \neq Y^j_\alpha$, $p+1 \leq \alpha \leq \min\{l_i, l_j\}$ as the number of collisions in H is 0. Therefore, we have,

$$\Pr[E_{\text{coll}} \wedge G = H] = \Pr[\Theta^i = \Theta^j \wedge G = H | \Sigma^i = \Sigma^j] \cdot \Pr[\Sigma^i = \Sigma^j]$$

It is obvious that $\Pr[\Sigma^i = \Sigma^j] \leq \frac{1}{2^n - 2\ell}$ and the event $\Theta^i = \Theta^j \wedge G = H$ conditioned on the event $\Sigma^i = \Sigma^j$ implies a non trivial equation on $\boldsymbol{Y}$ as we will obtain Y^i_{p+1} and Y^j_{p+1} for which $\Theta^i \oplus \Theta^j = 0$ would become non-trivial. Thus, $\Pr[\Theta^i = \Theta^j \wedge G = H | \Sigma^i = \Sigma^j] \leq \frac{1}{2^n - 2\ell}$. Therefore,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}. \quad \text{Assuming } \ell \leq 2^{n-1}$$

Case (ii) Consider either of the two messages is a prefix of other (w.l.o.g M^j is a prefix of M^i). Since $l_i > l_j$ therefore, $p = l_j$. Since the number of collision in H is 0, $Y^i_{p+1}, \ldots Y^i_{l_i}$ are all distinct with each other and with $Y^i_1, \ldots, Y^i_{l_j}$. This implies that $Y^i_{l_i} \neq Y^j_{l_j}$ as depicted in Fig. 4.1. Therefore, the probability of $\Theta^i = \Theta^j \wedge G = H$ conditioned on the event $\Sigma^i = \Sigma^j$ will be $O(1/2^n)$ as we will obtain two random variables $Y^i_{l_i}$ and $Y^j_{l_j}$ for which $\Theta^i \oplus \Theta^j = 0$ would become non-trivial. Moreover, $\Pr[\Sigma^i = \Sigma^j] \leq \frac{1}{2^n}$. Therefore again,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}.$$

Since, $\mathcal{G}^0 = 1$, we have, $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$. □

Fig. 4.1. Structure graph with 0 accident

Case (B) : Proof of $\mathbf{Pr}[\boldsymbol{E}_{\mathbf{coll}} \wedge |\boldsymbol{fColl}(\boldsymbol{G})| = \mathbf{1}] \leq \frac{\boldsymbol{\ell^2}}{\boldsymbol{2^{2n}}}$. Like the earlier case, we fix a structure graph $H \in \mathcal{G}^1$ and then analyse the probability of the event E_{coll} with respect to H in a case-by-case basis. Since $\mathcal{G}^1 = \mathcal{G}^1_{nl} \sqcup \mathcal{G}^1_l$, it follows that $H \in \mathcal{G}^1_{nl}$ or $H \in \mathcal{G}^1_l$. We analyse each case separately as follows:

Case (B.1) When $H \in \mathcal{G}^1_{nl}$. It essentially implies that H is the union of two walk graphs W^i, W^j such that W_i and W_j are path. Without loss of generality, we consider $l_i \geq l_j$.

Case (B.2) When $H \in \mathcal{G}^2_l$. It implies that either of the walks W^i or W^j contains a loop.

(B.1) Analysis of $\boldsymbol{\mathcal{G}^1_{nl}}$. Let us consider $H \in \mathcal{G}^1_{nl}$. First of all we would like to note that if M^j is a proper prefix of M^i then $|\mathcal{G}^1_{nl}| = 0$, as in that case number of accidents of H will be 0. So, without loss of generality, lets assume that M^j is not a prefix of M^i and p be the LCP of M^i and M^j. Therefore, $Y^i_\alpha = Y^j_\alpha$, $1 \leq \alpha \leq p$. As number of collision is 1 therefore, let the colliding pair is $(Y^i_{\beta_i}, Y^j_{\beta_j})$, where $p+1 \leq \beta_i \leq l_i, p+1 \leq \beta_j \leq l_j$.

Case (i) Let $\beta_i = \beta_j = p+1$ and $l_i = l_j$ and after the collision $Y^i_\beta = Y^j_\beta$, for $p+2 \leq \beta \leq l_i$. In this case, it is clear that checksum block of i^{th} message CS^i and checksum block of j^{th} message CS^j would not be equal and therefore even if $Y^i_{l_i} = Y^j_{l_j}$, the event $\Sigma^i = \Sigma^j$ would not be trivial. So, even though $\Pr[\Theta^i = \Theta^j | \Sigma^i = \Sigma^j \wedge G = H] = 1$, but the required randomness will be obtained from the following two equations : (i) $Y^i_{p+1} \oplus Y^j_{p+1} = 0$, (ii) $\Sigma^i \oplus \Sigma^j = 0$ such that the rank of the system of equations is 2. Therefore,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}.$$

Case (ii) Let $\beta_i = \beta_j = p+1$ and $l_i = l_j$ and after the collision $Y^i_\beta \neq Y^j_\beta$, for $p+2 \leq \beta \leq l_i$. Then we will always obtain Y^i_k and $Y^j_{k'}$ such that $\Theta^i = \Theta^j$ is non-trivial for some k, k'. Therefore again in this case we have,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}.$$

Case (iii) Let $\beta_i = \beta_j = p+2$ and $l_i = l_j$ and $Y^i_\beta = Y^j_\beta$, for $p+3 \leq \beta \leq l_i$, then $\Theta^i = \Theta^j$ would imply $Y^i_{p+1} = Y^j_{p+1}$; creates one more collision which violates the condition that the structure graph has only one collision.

Therefore, in general, we assume that the colliding pair is $(Y^i_{\beta_i}, Y^j_{\beta_j})$, where $p+1 \leq \beta_i \leq l_i, p+1 \leq \beta_j \leq l_j$. Since the number of collision allowed in H is 1,

after the collision point either W^i and W^j follow the same path or they will get bifurcated right from the collision point and will never meet again. If W^i and W^j follows the same path, then for Case (i) we have shown that we can ensure to get the probability $O(1/2^{2n})$. If not, then except Case (iii) where $\beta_i = \beta_j = p + 2$, we will obtain two random variables Y^i_k and $Y^j_{k'}$ such that equation $\Theta^i \oplus \Theta^j = 0$ becomes non-trivial. If W^i and W^j gets bifurcated right after the collision point, then the equality of Θ becomes non-trivial for two random variables Y^i_{p+1} and Y^j_{p+1} as depicted in (a) and (b) of Fig. 4.2. Note that it is easy to follow that we will always obtain two such random variables.

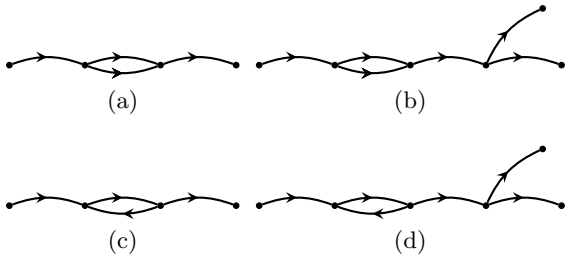

Fig. 4.2. Structure graphs with 1 accident. (a) and (b): no loop, (c) and (d): one loop.

Case (iv) Finally, if $\beta_i = l_i$ and $\beta_j = l_j$ then one can easily find out two random variables from the set $\{Y^i_{p+1}, \ldots, Y^i_{l_i-1}\} \cup \{Y^j_{p+1}, \ldots, Y^j_{l_j-1}\}$ such that the equation on Θ becomes non-trivial.

Therefore, in each of the above cases we have obtained

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}.$$

Since $|\mathcal{G}^1_{nl}| \leq |\mathcal{G}^1| \leq \ell^2$, we have, $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 1] \leq \frac{\ell^2}{2^{2n}}$.

(B.2) Analysis of $\mathcal{G}^1_l$. Let us fix a structure graph $H \in \mathcal{G}^1_l$. Without loss of generality we assume that W^i contains a loop. That means α is a smallest integer such that $Y^i_\alpha = Y^i_{\alpha+c}$ for $c \geq 1$. Here c denotes the loop size. Note that, the loop actually creates a collision and therefore, neither (i) W^j or W^i makes another different loop, nor (ii) W^j collides with W^i as in both of the cases number of collisions will increase to 2. Thus, the only possibilities are either

(i) W^j completely lies on W^i
(ii) W^j could follow W^i but after a point W^j and W^i gets bifurcated and never meets.

We will analyze the probability of the event $E_{\text{coll}} \wedge G = H$ separately for each of the above cases.

Case (i) : W^j completely lies on W^i.
Let us assume $W^i = Y_1^i||\ldots||Y_{\alpha-1}^i||(Y_\alpha^i||\ldots||Y_{\alpha+c-1}^i)^k||Y_{\alpha+c+1}^i||\ldots||Y_{l_i}^i$ and $W^j = Y_1^j||\ldots||Y_{\alpha-1}^j||(Y_\alpha^j||\ldots||Y_{\alpha+c-1}^j)^{k'}||Y_{\alpha+c+1}^j||\ldots||Y_{l_j}^j$ where $k' \geq 0$. Now we have the following cases:

• As W^j lies on W^i, it is easy to see that if $k' = 0$ then W^j be a subsequence of $Y_1^i||\ldots||Y_{\alpha-1}^i$ and therefore one can ensures the non-triviality of equation $\Theta^i \oplus \Theta^j = 0$ which holds with probability $\frac{1}{2^n}$. Moreover, $Y_{l_i}^i \neq Y_{l_j}^j$ and thus $\Sigma^i = \Sigma^j$ also holds with probability $\frac{1}{2^n}$ and therefore $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

• If $k' \geq 1$, then it is obvious that $Y_1^j||\ldots||Y_{\alpha-1}^j = Y_1^i||\ldots||Y_{\alpha-1}^i$. Now, if we assume that the length of the tail of W^i (i.e $Y_{\alpha+c+1}^i||\ldots||Y_{l_i}^i$) is same as that of W^j then it must have been the case that $k \neq k'$ and without loss of generality we can assume that $k > k'$. Since $Y_{l_i}^i = Y_{l_j}^j$, depending on the equality of CS^i and CS^j we have $\Pr[\Sigma^i = \Sigma^j |\ \ |fColl(G)| = 1] = 1$. Therefore,

$$\begin{aligned}\Pr[E_{\text{coll}} \wedge G = H] = \Pr[\Theta^i = \Theta^j | \Sigma^i = \Sigma^j \wedge G = H] \\ \cdot\ \Pr[\Sigma^i = \Sigma^j |\ \ G = H] \cdot \Pr[G = H]\end{aligned}$$

As $k > k'$ therefore, it is obvious to see that there must be at least two random variables Y_s^i and $Y_{s'}^i$ for which $\Theta^i = \Theta^j$ would become non-trivial as depicted in (c) of Fig. 4.2.

Thus in the above equation, $\Pr[\Theta^i = \Theta^j | \Sigma^i = \Sigma^j \wedge G = H] \leq \frac{1}{2^n}$ and $\Pr[G = H] \leq \frac{1}{2^n}$. Therefore, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}$. Moreover, if we assume that the tail length of W^i and W^j are not same (w.l.o.g $tail(W^i) > tail(W^j)$) then we have either $k = k'$ or $k \neq k'$. The case of $k = k'$ has already been taken care of. If $k \neq k'$ then $Y_{l_i}^i \neq Y_{l_j}^j$ and therefore, $\Theta^i \oplus \Theta^j = 0$ would become non-trivial for the random variable $Y_{l_i}^i$ and $Y_{l_j}^j$. Moreover, $\Pr[\Sigma^i = \Sigma^j] \leq \frac{1}{2^n}$. Thus,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}.$$

Case (ii) : W^j follows W^i but after they get bifurcated and never meets. In this case W^j bifurcates from W^i right after some point X. This condition necessarily implies that $Y_{l_i}^i \neq Y_{l_j}^j$. Now it is to be noted that if W^j completely lies on W^i (as in $head(W^i) = head(W^j)$ and $k = k'$) and bifurcates right from the point $X = Y_{l_i-1}^i$, then $\Theta^i = \Theta^j$ would imply $Y_{l_i}^i = Y_{l_j}^j$, introduces one more collision and hence the number of collision would increase. Therefore, even if $head(W^i) = head(W^j)$ either $k \neq k'$ or W^j must get bifurcated from W^i from some earlier point of $Y_{l_i-1}^i$. In both of these cases one should obtain at least two random variables (*either from portion of loop or from portion of tail*) Y_s^i and $Y_{s'}^i$ for some s and s' that ensures the non-triviality of equation on Θ as depicted in (d) of Fig. 4.2.

Moreover as $Y_{l_i}^i \neq Y_{l_j}^j$ this ensures that $\Pr[\Sigma^i = \Sigma^j] \leq \frac{1}{2^n}$. Hence, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

Therefore, in all of the above cases we have obtained $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{2n}}$. Moreover, $|\mathcal{G}_l^1| \leq |\mathcal{G}^1| \leq \ell^2$. So, $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 1] \leq \frac{\ell^2}{2^{2n}}$.

Case (C) : Proof of $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 2] \leq \frac{x\ell^4}{2^{3n}}$ Likewise the analysis of Case (B), we first fix a graph $H \in \mathcal{G}^2$ and analyze the probability of E_{coll} with respect to H in a case-by-case basis. With the same argument, either $H \in \mathcal{G}_{nl}^2$ or $H \in \mathcal{G}_l^2$.

Case (C.1) Let us consider $H \in \mathcal{G}_{nl}^2$ which implies that none of the message walks W^i or W^j contains a loop.

Case (C.2) Let us consider $H \in \mathcal{G}_l^2$ which implies that either of the message walks W^i or W^j contains a loop.

(C.1) Analysis of $\mathcal{G}_{nl}^2$. Let p be the LCP of M^i and M^j. Since number of accident of H is 2, we denote the collision pairs are : $(Y_{\alpha_i}^i, Y_{\alpha_j}^j)$ and $(Y_{\beta_i}^i, Y_{\beta_j}^j)$ where $p+1 \leq \alpha_i, \beta_i \leq l_i$ and $p+1 \leq \alpha_j, \beta_j \leq l_j$.

Case (i) Let $l_i = l_j$, $\alpha_i = \alpha_j = p+1$ and $\beta_i = \beta_j = p+2$ and after collision $Y_s^i = Y_s^j$, for $p+3 \leq s \leq l_i$. This case is boiled down to the analysis of subcase (i) under Case (B.1). Therefore, even though $\Pr[\Theta_i = \Theta_j | \Sigma_i = \Sigma_j \wedge G = H] = 1$, we obtain the required randomness from the following three linearly independent equations : (i) $Y_{p+1}^i \oplus Y_{p+1}^j = 0$, (ii) $Y_{p+2}^i \oplus Y_{p+2}^j = 0$ and (iii) $\Sigma^i \oplus \Sigma^j = 0$ such that the rank of the system of equations become 3. Therefore, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{3n}}$.

Case (ii) This case is similar to Case (i) except that after the collision $Y_s^i \neq Y_s^j$. Again this case is boiled down to the analysis of subcase (ii) under Case (B.1). Therefore it is easy to see that the obtained rank of the linear system of equations will be at least 3. Therefore in this case also, we obtain, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{3n}}$.

Case (iii) Let $l_i = l_j$ and two collision points are not consecutive like Case (i). We can also assume that after the final collision point (i.e $Y_{\beta_i}^i = Y_{\beta_j}^j$) $Y_s^i = Y_s^j$ for $s \leq l_i$. So, we can obtain a system of linear equations of rank 3 such that $\Theta^i \oplus \Theta^j = 0$ along with two collisions give three linearly independent equations. Therefore, in this case we obtain, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{3n}}$.

In general, we assume that the colliding pair is $(Y_{\alpha_i}^i, Y_{\alpha_j}^j)$ and $(Y_{\beta_i}^i, Y_{\beta_j}^j)$, where $p+1 \leq \alpha_i, \beta_i \leq l_i, p+1 \leq \alpha_j, \beta_j \leq l_j$. Since the number of collision allowed in H is 2, after the first collision point $(Y_{\alpha_i}^i, Y_{\alpha_i}^j)$, W^i and W^j must bifurcate and then meets with each other to form the second collision point $(Y_{\beta_i}^i, Y_{\beta_j}^j)$ and then W^i, W^j follow the same path or they will get bifurcated from the second collision point and will never meet again. If W^i and W^j follows the same path, then for Case (i) we have shown that we can ensure to get the probability $O(1/2^{2n})$. If not, then we will obtain two random variables Y_k^i and $Y_{k'}^j$ such that equation $\Theta^i \oplus \Theta^j = 0$ becomes non-trivial. If W^i and W^j gets bifurcated after the second collision point, then the equality of Θ becomes non-trivial for two

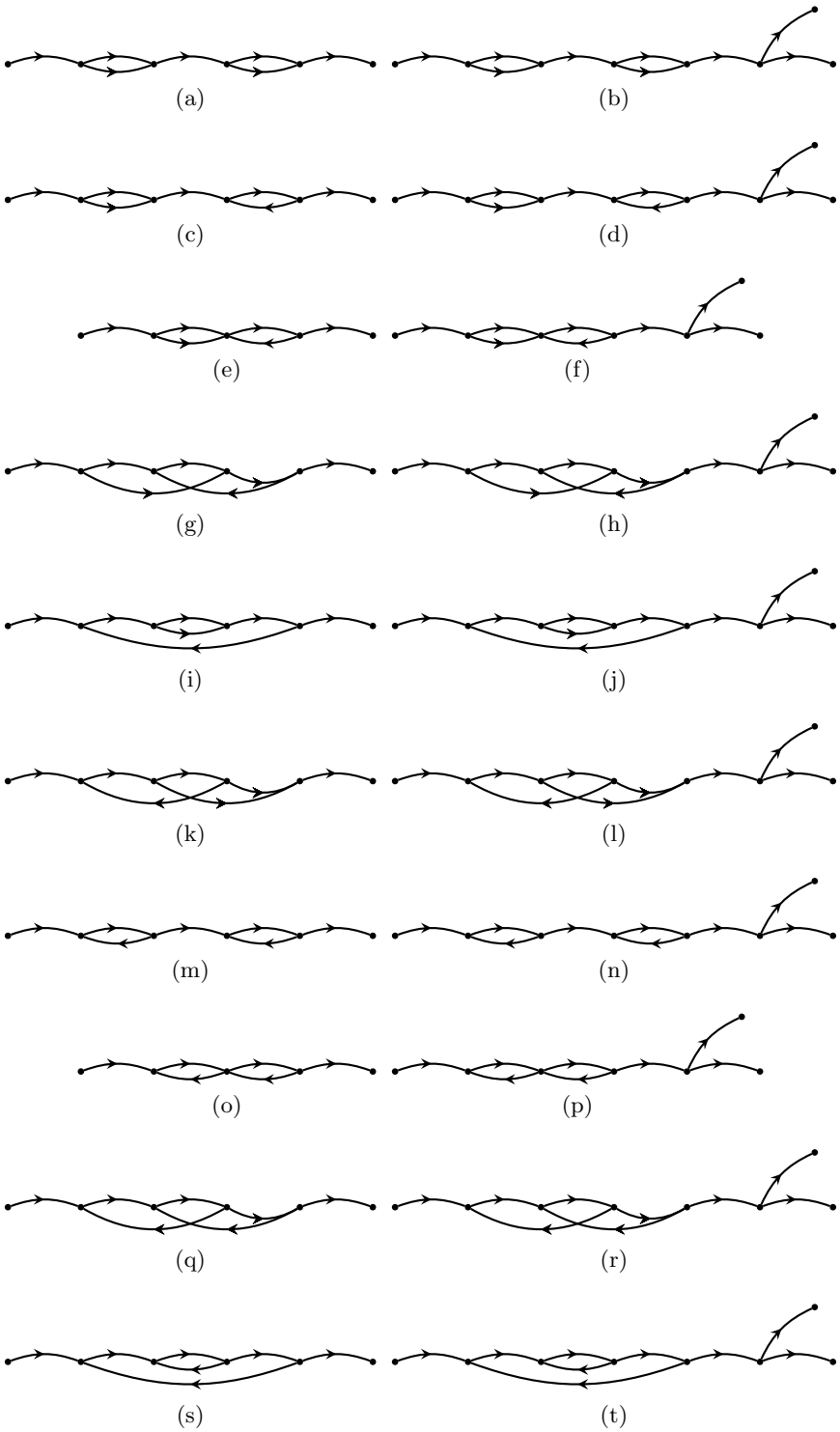

Fig. 4.3. Structure graphs with 2 accidents. (a) and (b) : No loop, (c) to (l) : one loop, (m) to (t) : two loops.

random variables Y_k^i and $Y_{k'}^j$ as depicted in (a) and (b) of Fig. 4.3. Note that it is easy to follow that we will always obtain two such random variables. Therefore, the obtained rank of the linear system of equations comprising of equations (i) $\Sigma^i \oplus \Sigma^j = 0, \Theta^i \oplus \Theta^j = 0, Y_{\alpha_i}^i \oplus Y_{\alpha_j}^j = 0, Y_{\beta_i}^i = Y_{\beta_j}^j = 0$ will be at least 3. Therefore, $\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{3n}}$.

Case (iv) Finally, $\beta_i = l_i$ and $\beta_j = l_j$ where $\alpha_i < \beta_i, \alpha_j < \beta_j$, then one can easily find out two random variables from the set $\{Y_{p+1}^i, \ldots, Y_{l_i-1}^i\} \cup \{Y_{p+1}^j, \ldots, Y_{l_j-1}^j\}$ such that the equation on Θ becomes non-trivial.

Therefore, from the all of the above cases we have the following,

$$\Pr[E_{\text{coll}} \wedge G = H] \leq \frac{1}{2^{3n}}.$$

Moreover, $|\mathcal{G}_{nl}^2| \leq |\mathcal{G}^2| \leq \ell^4$. Therefore, $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 2] \leq \frac{\ell^4}{2^{3n}}$.

(C.2) Analysis of $\mathcal{G}_l^2$. We characterize the all possible graphs in following two ways :

(i) When both the accident comes from a single message walk.
(ii) When two message walks are involved to yield two accidents.
  (ii.i) When each message walk contributes a single accident.
  (ii.ii) When two message walk jointly contributes two accidents.

Let p be the LCP of M^i and M^j. Since number of accident of H is 2, here the collision pairs will be one of the followings based on the three cases listed above: (a) $(Y_{\alpha_i}^i, Y_{\alpha_i'}^i)$ and $(Y_{\beta_i}^i, Y_{\beta_i'}^i)$ (single message walk) where $\alpha_i < \alpha_i^{'} < \beta_i < \beta_i^{'}$, (b) $(Y_{\alpha_i}^i, Y_{\alpha_i'}^i)$ and $(Y_{\beta_j}^j, Y_{\beta_j'}^j)$ (each message walk contributes single accident) where $\alpha_i < \alpha_i^{'}$ and $\beta_i < \beta_i^{'}$, (c) $(Y_{\alpha_i}^i, Y_{\alpha_j}^j)$ and $(Y_{\beta_i}^i, Y_{\beta_j}^j)$ (two message walks jointly contribute two accidents) where $p+1 \leq \alpha_i, \beta_i \leq l_i$ and $p+1 \leq \alpha_j, \beta_j \leq l_j$.

Case (i) : Both accidents come from a single message walk. To analyze this case, note that, only a single message walk (e.g W^i) yields two accidents; that means, the accident pair is $(Y_{\alpha_i}^i, Y_{\alpha_i'}^i)$ and $(Y_{\beta_i}^i, Y_{\beta_i'}^i)$, thus W^i contains two distinct loops, whereas W^j does not contain any loop. In this regard, it is to be noted that W^j either lies on W^i or W^j eventually bifurcates from W^i and never meets again. Now we have two possibilities under this case. (a) When $l_i = l_j$, then it has to be the case that W^j must bifurcates from W^j from some fixed certain point node X in H. Note that, it may also happen that X does not exist in some H and in that specific cases we will obtain two parallel walks. Now one can easily see that two distinct accidents yields two linearly independent equations. That is

$$Y_{\alpha_i}^i \oplus Y_{\alpha_i'}^i = 0$$
$$Y_{\beta_i}^i \oplus Y_{\beta_i'}^i = 0.$$

Moreover, the following two equations $\Sigma^i \oplus \Sigma^j = 0$ and $\Theta^i \oplus \Theta^j = 0$ is not implied from the previous two linearly independent equations coming from accidents as one can easily see that $Y^i_{l_i} \neq Y^j_{l_j}$ and thus, $\Sigma^i \oplus \Sigma^j = 0$ is not a trivial equation. Thus one can ensure that the rank of this system of linear equations is at least 3. (b) When $l_i \neq l_j$, then without loss of generality we assume that $l_i > l_j$. Therefore, either W^j bifurcates from W^i or W^j completely lies on W^i. Former case has already been treated. So, when W^j completely lies on W^i where $|W^j| < |W^i|$[4], then again $Y^i_{l_i} \neq Y^j_{l_j}$, making the equation $\Sigma^i \oplus \Sigma^j = 0$ non-trivial. Moreover, two accidents imply two linearly independent equations. Altogether, the rank of the system of equations become at least 3. Therefore, in this case, we obtain

$$Pr[E_{coll} \wedge G = H] \leq \frac{1}{2^{3n}}. \tag{5}$$

Case (ii.i) : Each message walk contributes a single accident. When two message walk W^i, W^J individually contributes a single accident, that is the accident pair is $(Y^i_{\alpha_i}, Y^i_{\alpha'_i})$ and $(Y^j_{\beta_j}, Y^j_{\beta'_j})$. Note that the last collision point, say, $(Y^j_{\beta_j}, Y^j_{\beta'_j})$ must be after the `LCP` point. Therefore, each of W^i and W^j contains a single loop and they never meet again, otherwise that will contribute to one more accident. Therefore, the structure of the graph is simple as depicted in (m) and (n) of Fig 4.3. It is very straight-forward to see that $Y^i_{l_i} \neq Y^j_{l_j}$. Moreover, two distinct accident gives two linearly independent equations and therefore, one can see that the rank becomes at least 3. Thus, Equation (5) holds in this case.

Case (ii.ii) : Two message walks jointly contribute two accidents. Former two cases were easy to handle as those cases contain simple structure graphs. This case is little involved to handle as it contains many kind of structure graphs as depicted in (c) to (t) of Fig, 4.3.

Let d denotes the gap of two colliding nodes[5]. Note that for (e), (f), (o) and (p) of Fig. 4.3, value of d is 0. For the rest of the cases, $d > 0$.
To keep our discussion simple, we give the details proof of (c) of Fig. 4.3 and then one can use the similar analysis for the proof of the rest of the cases.

Details analysis for Case (c) of Fig. 4.3. Let the fist collision point is $(Y^i_{\alpha_i}, Y^j_{\alpha_j})$. This accident is contributed by two message walks W^i and W^j. After this first accident point, the second message walk may or may not take part in forming the second collision. (a) If W^j takes part in forming the second collision then after the first collision point W^i and W^j will move in unison and after forming the second collision W^j and W^i may bifurcates or again they move in unison depending on the message blocks of M^j. (b) On the other hand, if W^j does not take part then either (i) W^j bifurcates from a node X where $X \in \{Y^i\alpha_i, Y^i_{\alpha_i+1}, \ldots, Y^i_{\alpha_i+d}\}$ and never meets again or (ii) W^j completely lies on W^i and $|W^j| < \beta_i$. Note that in both of the cases (a) and (b), two collision

[4] Length of a walk W is denoted as $|W|$.

[5] gap of two colliding nodes means the number of edges in the structure graphs between two vertices which are collided.

give rises to two linearly independent equations

$$Y^i_{\alpha_i} \oplus Y^j_{\alpha_j} = 0$$
$$Y^i_{\beta_i} \oplus Y^j_{\beta_j} = 0.$$

(a) We consider W^j takes part in forming the second collision. If $l_i = l_j$, then we will find Y^i_{p+1} for which $\Theta^i \oplus \Theta^j = 0$ becomes non-trivial and hence the rank of the above two equations along with $\Theta^i \oplus \Theta^j = 0$ becomes 3. If $l_i \neq l_j$ then again one can ensure to obtain $Y^j_{l_i}$ such that the variable is fresh in the equation $\Theta^i \oplus \Theta^j = 0$ which makes the rank of the above three equations to 3.

(b) We consider w^j does not take part in forming the second collision. Therefore (i) When W^j bifurcates from W^i then again $Y^j_{l_j}$ will be the fresh random variable in the equation $\Theta^i \oplus \Theta^j = 0$; making the rank of the system of equations to at least 3. (ii) If W^j completely lies on W^i, which essentially implies $l_j < l_i$, and therefore, one can obtain $Y^i_{l_i}$ which will be fresh in the equation $\Theta^i \oplus \Theta^j = 0$; making the rank at least 3.

Therefore, in all of the above cases, we have observed that the rank of the following system of equations is at least 3.

$$Y^i_{\alpha_i} \oplus Y^j_{\alpha_j} = 0$$
$$Y^i_{\beta_i} \oplus Y^j_{\beta_j} = 0$$
$$\Sigma^i \oplus \Sigma^j = 0$$
$$\Theta^i \oplus \Theta^j = 0.$$

Therefore, we have,

$$\Pr[E_{coll} \wedge G = H] \leq \frac{1}{2^{3n}}.$$

All of the remaining cases can be analyzed similarly and one can show the rank to be at least 3. Since, $|\mathcal{G}^2_l| \leq |\mathcal{G}^2| \leq \ell^4$. Therefore, $\Pr[E_{\text{coll}} \wedge |fColl(G)| = 2] \leq \frac{\ell^4}{2^{3n}}$.

Case (D) : Proof of $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$. We fix a structure graph $H \in \mathcal{G}^0$ and then analyse the probability of the event E_{cf} with respect to H in a case-by-case basis.

Case (i) Let p be the LCP of M^i and M^j. Therefore, $Y^i_\alpha = Y^j_\alpha$ where $1 \leq \alpha \leq p$ and $Y^i_\beta \neq Y^j_\beta$ where $p+1 \leq \beta \leq \min\{l_i, l_j\}$ as the number of accident in H is 0. Moreover, if $l_i > l_j$ then all Y^i_β would have been distinct as $|fColl(G)| = 0$ where $l_j + 1 \leq \beta \leq l_i$. Note that, it is also true that $Y^i_{l_i} \neq Y^j_{l_j}$. Therefore, we have the following set of equations:

$$Y^i_{l_i+1} = x, \tag{6}$$
$$Y^i_1 \oplus Y^i_2 \oplus \ldots \oplus Y^i_{l_i+1} + Y^s_t = 0, \tag{7}$$

where s could be either i or j and $t \in [l_i+1]$ or $t \in [l_j+1]$. For each of these cases one can easily check that the above system of equation has rank 2. Therefore, $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

Case (ii). Without loss of generality let us consider that M^j is a prefix of M^i. Since $l_i > l_j$ therefore, $p = l_j$. Since, number of collisions in H is 0, $Y^i_{p+1}, \ldots Y^i_{l_i}$ are all distinct with each other and with $Y^j_1, \ldots, Y^j_{l_j}$. This implies that $Y^i_{l_i} \neq Y^j_{l_j}$ as depicted in Fig. 4.1. Therefore, the set of equations (Equation (6) and (7)) has the full rank. Therefore, again we have, $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

Therefore from the above two cases we have, $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$ for any non-zero n bit constant x. Moreover $|\mathcal{G}^0| \leq 1$. So $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 0] \leq \frac{1}{2^{2n}}$.

Case (E) : Proof of $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 1] \leq \frac{l^2}{2^{2n}}$. Again, we fix a structure graph $H \in \mathcal{G}^1$ and then analyse the probability of the event E_{cf} with respect to H in a case-by-case basis. Therefore, $H \in \mathcal{G}^1_{nl}$ or $H \in \mathcal{G}^1_l$. We analyse each case separately as follows.

Case (E.1) Let us consider $H \in \mathcal{G}^1_{nl}$ which implies that none of the message walks W^i or W^j contains a loop.

Case (E.2) Let us consider $H \in \mathcal{G}^1_l$ which implies that either of the message walks W^i or W^j contains a loop

(E.1) Analysis of $\mathcal{G}^1_{nl}$. As before M^i or M^j could not be a prefix of each other. Let p be the LCP of M^i and M^j and let the colliding pair is $(Y^i_{\beta_i}, Y^j_{\beta_j})$, where $p+1 \leq \beta_i \leq l_i, p+1 \leq \beta_j \leq l_j$. In this case, it is easy to check that the following system of equations will have rank 2.

$$Y^i_{l_i+1} = x,$$
$$Y^i_1 \oplus Y^i_2 \oplus \ldots \oplus Y^i_{l_i+1} + Y^s_t = 0.$$

Therefore, we have $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

Note that, $|\mathcal{G}^1_{nl}| \leq |\mathcal{G}^1| \leq \ell^2$. Therefore $\Pr[E_{\text{cf}} \wedge |fColl(G) = 1|] \leq \frac{l^2}{2^{2n}}$.

(E.2) Analysis of $\mathcal{G}^2_l$. As before let us assume that W^i contains a loop of size c such that $Y^i_\alpha = Y^i_{\alpha+c}$ for $c \geq 1$. Since the loop creates a collision, neither (i) W^j or W^i makes another different loop, nor (ii) W^j collides with W^i as in both of the cases the number of collisions will increase to 2. Thus we have the following two possibilities.

(1) W^j coincides with W^i
(2) W^j could follow W^i but after a point W^i and W^j departs and never meets again.

We analyze the probability of the event $E_{\text{cf}} \wedge G = H$ separately for each of the two above cases. In particular, in each of the following analysis our main concern will be to show the rank of the set of equations as defined earlier (i.e Equation (6) and (7)) to be 2, that is it achieves full rank in each of the following subcases.

Case (i) : W^j coincides with W^i.
Let k denotes the number of iterations in the loop of W^i and k' be the number of iterations in the loop of W^j. Now irrespective of the value of k and k', the system of equations (Equation (6) and (7)) will have rank 2 and therefore, we can upper bound the probability of our desired event to $\frac{1}{2^{2n}}$.

Case (ii) : W^j could follow W^i but after a point W^i and W^j departs and never meets again.
The analysis for this case would be similar to Case (i). Here W^i and W^j bifurcates from a certain point say X and $l_i - X, l_j - X \neq 0$. Therefore, it is trivial to see that the set of equations (i.e Equation (6) and (7)) will have full rank. Again, as we have shown in the previous case that $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$.

Therefore, for the above two cases $\Pr[E_{\text{cf}} \wedge G = H] \leq \frac{1}{2^{2n}}$. Moreover, $|\mathcal{G}_l^1| \leq |\mathcal{G}^1| \leq \ell^2$. Therefore, $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 1] \leq \frac{\ell^2}{2^{2n}}$.

Case (F) : Proof of $\Pr[E_{\text{cf}} \wedge |fColl(G)| = 2] \leq \frac{xl^4}{2^{3n}}$ Proof of this bound is similar to Case (C) and thus we skip the proof of the bound.

5 Conclusion

In this paper, we have proposed a non-tweaked single fixed-key compression function based MAC NI$^+$, a variant of NI-MAC that achieves BBB security and efficient than NI-MAC in terms of number of keys. Moreover, our construction is better than Yasuda's proposed single-fixed key compression function based MAC construction that uses an extra mask of b bits which needs a storage space. Moreover, we have been able to *slightly* reduce the state size from $2(b + 2n)$ bits to $(b + 2n)$ bits which was an open problem in [43] to reduce the state size to $2n$ bits. Thus we are leaving the problem still open which does not require now a extra mask.

References

1. Jee Hea An and Mihir Bellare. Constructing vil-macsfrom fil-macs: Message authentication under weakened assumptions. In Wiener [39], pages 252–269.
2. Mihir Bellare, Ran Canetti, and Hugo Krawczyk. Keying hash functions for message authentication. In Neal Koblitz, editor, CRYPTO '96, volume 1109 of *LNCS*, pages 1–15. Springer, 1996.
3. Mihir Bellare, Oded Goldreich, and Hugo Krawczyk. Stateless evaluation of pseudorandom functions: Security beyond the birthday barrier. In Wiener [39], pages 270–287.
4. Mihir Bellare, Joe Kilian, and Phillip Rogaway. The security of cipher block chaining. In Yvo Desmedt, editor, CRYPTO '94, volume 839 of *LNCS*, pages 341–358. Springer, 1994.

5. Mihir Bellare, Krzysztof Pietrzak, and Phillip Rogaway. Improved security analyses for CBC macs. In Shoup [34], pages 527–545.
6. Mihir Bellare. New proofs for NMAC and HMAC: security without collision-resistance. In Cynthia Dwork, editor, CRYPTO 2006, volume 4117 of *LNCS*, pages 602–619. Springer, 2006.
7. Mihir Bellare and Phillip Rogaway. The security of triple encryption and a framework for code-based game-playing proofs. In Vaudenay [35], pages 409–426.
8. John Black, Shai Halevi, Hugo Krawczyk, Ted Krovetz, and Phillip Rogaway. UMAC: fast and secure message authentication. In Wiener [39], pages 216–233.
9. John Black and Phillip Rogaway. A block-cipher mode of operation for parallelizable message authentication. In Knudsen [21], pages 384–397.
10. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: an ultra-lightweight block cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, CHES 2007, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.
11. Nilanjan Datta, Avijit Dutta, Mridul Nandi, Goutam Paul, and Liting Zhang. One-key double-sum mac with beyond-birthday security. Cryptology ePrint Archive, Report 2015/958, 2015. http://eprint.iacr.org/.
12. Yevgeniy Dodis, Thomas Ristenpart, John P. Steinberger, and Stefano Tessaro. To hash or not to hash again? (in)differentiability results for H 2 and HMAC. In Reihaneh Safavi-Naini and Ran Canetti, editors, CRYPTO 2012, volume 7417 of *LNCS*, pages 348–366. Springer, 2012.
13. Yevgeniy Dodis and John P. Steinberger. Domain extension for macs beyond the birthday barrier. In Kenneth G. Paterson, editor, EUROCRYPT 2011, volume 6632 of *LNCS*, pages 323–342. Springer, 2011.
14. Peter Gazi, Krzysztof Pietrzak, and Michal Rybár. The exact prf-security of NMAC and HMAC. In Juan A. Garay and Rosario Gennaro, editors, CRYPTO 2014, volume 8616 of *LNCS*, pages 113–130. Springer, 2014.
15. Peter Gaži, Krzysztof Pietrzak, and Stefano Tessaro. Generic security of nmac and hmac with input whitening. Cryptology ePrint Archive, Report 2015/881, 2015. http://eprint.iacr.org/.
16. Deukjo Hong, Jaechul Sung, Seokhie Hong, Jongin Lim, Sangjin Lee, Bonseok Koo, Changhoon Lee, Donghoon Chang, Jaesang Lee, Kitae Jeong, Hyun Kim, Jongsung Kim, and Seongtaek Chee. HIGHT: A new block cipher suitable for low-resource device. In Louis Goubin and Mitsuru Matsui, editors, CHES 2006, volume 4249 of *LNCS*, pages 46–59. Springer, 2006.
17. Tetsu Iwata and Kaoru Kurosawa. OMAC: one-key CBC MAC. In Johansson [19], pages 129–153.
18. Éliane Jaulmes, Antoine Joux, and Frédéric Valette. On the security of randomized CBC-MAC beyond the birthday paradox limit: A new construction. In FSE, 2002, volume 2365 of *LNCS*, pages 237–251. Springer, 2002.
19. Thomas Johansson, editor. In FSE, 2003, volume 2887 of *LNCS*. Springer, 2003.
20. Antoine Joux, Guillaume Poupard, and Jacques Stern. New attacks against standardized macs. In Johansson [19], pages 170–181.
21. Lars R. Knudsen, editor. EUROCRYPT 2002, volume 2332 of *LNCS*. Springer, 2002.
22. Neal Koblitz and Alfred Menezes. Another look at hmac. *J. Mathematical Cryptology*, 7(3):225–251, 2013.
23. H. Krawczyk, M. Bellare, and R. Canetti. HMAC: Keyed-Hashing for Message Authentication. RFC 2104 (Informational), February 1997.

24. Will Landecker, Thomas Shrimpton, and R. Seth Terashima. Tweakable block-ciphers with beyond birthday-bound security. In Reihaneh Safavi-Naini and Ran Canetti, editors, CRYPTO 2012, volume 7417 of *LNCS*, pages 14–30. Springer, 2012.
25. Gaëtan Leurent, Thomas Peyrin, and Lei Wang. New generic attacks against hash-based macs. In Kazue Sako and Palash Sarkar, editors, ASIACRYPT 2013, volume 8270 of *LNCS*, pages 1–20. Springer, 2013.
26. Stefan Lucks. A failure-friendly design principle for hash functions. In Bimal K. Roy, editor, ASIACRYPT 2005, volume 3788 of *LNCS*, pages 474–494. Springer, 2005.
27. Ueli M. Maurer and Johan Sjödin. Domain expansion of macs: Alternative uses of the FIL-MAC. In Nigel P. Smart, editor, *Cryptography and Coding, 2005*, volume 3796 of *LNCS*, pages 168–185. Springer, 2005.
28. Ueli M. Maurer and Johan Sjödin. Single-key ail-macs from any FIL-MAC. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, ICALP 2005, volume 3580 of *LNCS*, pages 472–484. Springer, 2005.
29. Kazuhiko Minematsu. How to thwart birthday attacks against macs via small randomness. In Seokhie Hong and Tetsu Iwata, editors, FSE, 2010, volume 6147 of *LNCS*, pages 230–249. Springer, 2010.
30. Yusuke Naito, Yu Sasaki, Lei Wang, and Kan Yasuda. Generic state-recovery and forgery attacks on chopmd-mac and on NMAC/HMAC. In Kazuo Sakiyama and Masayuki Terada, editors, IWSEC 2013, volume 8231 of *LNCS*, pages 83–98. Springer, 2013.
31. Thomas Peyrin, Yu Sasaki, and Lei Wang. Generic related-key attacks for HMAC. In Wang and Sako [36], pages 580–597.
32. Thomas Peyrin and Lei Wang. Generic universal forgery attack on iterative hash-based macs. In Phong Q. Nguyen and Elisabeth Oswald, editors, EUROCRYPT 2014, volume 8441 of *LNCS*, pages 147–164. Springer, 2014.
33. Bart Preneel and Paul C. van Oorschot. Mdx-mac and building fast macs from hash functions. In Don Coppersmith, editor, CRYPTO 1995, volume 963 of *Lecture Notes in Computer Science*, pages 1–14. Springer, 1995.
34. Victor Shoup, editor. CRYPTO 2005, volume 3621 of *LNCS*. Springer, 2005.
35. Serge Vaudenay, editor. EUROCRYPT 2006, volume 4004 of *LNCS*. Springer, 2006.
36. Xiaoyun Wang and Kazue Sako, editors. ASIACRYPT 2012, volume 7658 of *LNCS*. Springer, 2012.
37. Xiaoyun Wang, Yiqun Lisa Yin, and Hongbo Yu. Finding collisions in the full SHA-1. In Shoup [34], pages 17–36.
38. Xiaoyun Wang and Hongbo Yu. How to break MD5 and other hash functions. In Ronald Cramer, editor, EUROCRYPT 2005, volume 3494 of *LNCS*, pages 19–35. Springer, 2005.
39. Michael J. Wiener, editor. CRYPTO '99, volume 1666 of *LNCS*. Springer, 1999.
40. Kan Yasuda. Boosting merkle-damgård hashing for message authentication. In Kaoru Kurosawa, editor, ASIACRYPT 2007, volume 4833 of *LNCS*, pages 216–231. Springer, 2007.
41. Kan Yasuda. Multilane HMAC - security beyond the birthday limit. In K. Srinathan, C. Pandu Rangan, and Moti Yung, editors, INDOCRYPT 2007, volume 4859 of *LNCS*, pages 18–32. Springer, 2007.
42. Kan Yasuda. "sandwich" is indeed secure: How to authenticate a message with just one hashing. In Josef Pieprzyk, Hossein Ghodosi, and Ed Dawson, editors, ACISP 2007, volume 4586 of *LNCS*, pages 355–369. Springer, 2007.

43. Kan Yasuda. A one-pass mode of operation for deterministic message authentication- security beyond the birthday barrier. In Kaisa Nyberg, editor, FSE, 2008, volume 5086 of *LNCS*, pages 316–333. Springer, 2008.
44. Kan Yasuda. A double-piped mode of operation for macs, prfs and pros: Security beyond the birthday barrier. In Antoine Joux, editor, EUROCRYPT 2009, volume 5479 of *LNCS*, pages 242–259. Springer, 2009.
45. Kan Yasuda. HMAC without the "second" key. In Pierangela Samarati, Moti Yung, Fabio Martinelli, and Claudio Agostino Ardagna, editors, ISC 2009, volume 5735 of *LNCS*, pages 443–458. Springer, 2009.
46. Kan Yasuda. The sum of CBC macs is a secure PRF. In Josef Pieprzyk, editor, CT-RSA 2010, volume 5985 of *LNCS*, pages 366–381. Springer, 2010.
47. Kan Yasuda. A new variant of PMAC: beyond the birthday bound. In Phillip Rogaway, editor, CRYPTO 2011, volume 6841 of *LNCS*, pages 596–609. Springer, 2011.
48. Kan Yasuda. On the full MAC security of a double-piped mode of operation. *IEICE Transactions*, 94-A(1):84–91, 2011.
49. Kan Yasuda. A parallelizable prf-based MAC algorithm: Well beyond the birthday bound. *IEICE Transactions*, 96-A(1):237–241, 2013.
50. Liting Zhang, Wenling Wu, Han Sui, and Peng Wang. 3kf9: Enhancing 3gpp-mac beyond the birthday bound. In Wang and Sako [36], pages 296–312.

A Formal Discussion on Structure Graph

Let for two distinct messages M^1 and M^2 of l_1 and l_2 blocks respectively, where

$$M^1 = M^1_1||M^1_2||\dots||M^1_{l_1} \text{ and } M^2 = M^2_1||M^2_2||\dots||M^2_{l_2},$$

and the corresponding Y-values be given by

$$y^1_0, y^1_1, y^1_2, \dots, y^1_{l_1} \text{ and } y^2_0, y^2_1, y^2_2, \dots, y^2_{l_2}$$

respectively. Let $\tau = l_1 + l_2$. We use the notation M_i where $1 \le i \le \tau$ to refer to the block M^1_i, when $i \le l_1$, otherwise refer to the block $M^2_{i-l_1}$. Similarly, let Y_i to refer to $\mathbf{0}$ when $i = 0$; Y^1_i, when $1 \le i \le l_1$; and $Y^2_{i-l_1}$, when $l_1 + 1 \le i \le \tau$. Now, we give a few definitions.

Definition 2. *We define two mappings* $[[\cdot]]$ *and* $[[\cdot]]'$ *on* $\{0, \dots, \tau\}$ *as follows:*
(1) $[[i]] \triangleq \min\{j : Y_i = Y_j\}$*, and*
(2) $[[i']] \triangleq [[i]]$ *for* $i \neq l_1$ *except that* $[[l_1]]' = 0$.

Definition 3. *For any fixed* f *and any two distinct messages* $\mathcal{M} = \{M^1, M^2\}$*, we define the structure graph* $\mathcal{G}^f(\mathcal{M})$ *as follows:*
$\mathcal{G}^f(\mathcal{M}) \triangleq (V, E, L)$*, where* $V = \{[[i]] : 0 \le i \le \tau\}$, $E = \{([[i-1]]', [[i]]) : 1 \le i \le \tau\}$, *and* $L((u, v)) = \{M_i : [[i-1]]' = u \text{ and } [[i]] = v\}$ *is an edge-labeling.*

Definition 4. *For the computation of* M^1*, the sequence* 0, $([[0]]', [[1]])$, $[[1]]$, $([[1]]', [[2]])$, …, $[[l_1]]$ *of alternating vertices and edges is called an* M^1*-walk. (An* M^2*-walk is defined analogously).*

Let (V_i, E_i, L_i) be the graph obtained after processing only the first i out of τ blocks of $\mathcal{M}$. We define a collision event as follows.

Definition 5. $(i, [[i]])$ *is an* f*-collision if* $[[i]] < i$ *and* $M_i \notin L_{i-1}([[i-1]]', [[i]])$.

Note that the last condition on M_i implies that collision occurred due to parallel edges with the same message label is not considered.

B Proof of Proposition 2

Propostion 2 $\Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : |fColl(G)| \geq 3] \leq \frac{27\ell^6}{2^{3n}}$, where ℓ is the total number of blocks of the messages in $\mathcal{M}$ where $\ell \leq 2^{n/2}$.

$$\begin{aligned}\Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : |fColl(G)| \geq 3] &= \sum_{i=3}^{\infty} \Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : |fColl(G)| = i] \\ &\leq \sum_{i=3}^{\infty} \sum_{H \in \mathcal{G}^i(\mathcal{M})} \Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : G = H] \\ &\leq \sum_{i=3}^{\infty} \frac{|\mathcal{G}^i(\mathcal{M})|}{2^{in}}\end{aligned}$$

Now, note that the a graph G is uniquely determined by its number of collisions. Therefore, $|\mathcal{G}^i(\mathcal{M})| \leq (\frac{2\ell(2\ell+1)}{2})^i \leq (3\ell^2)^i$. Now let a denotes $\frac{3\ell^2}{2^n}$. Assuming $\ell \leq 2^{n/2}$ we can write,

$$\begin{aligned}\Pr[G \xleftarrow{\$} \mathcal{G}(\mathcal{M}) : |fColl(G)| \geq 3] &\leq \frac{a^3}{(1-a)} \\ &\leq \frac{27\ell^6}{2^{3n}}.\end{aligned}$$

□

C Diagram of NI

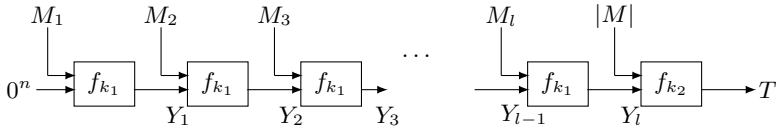

Fig. C.1. Construction of NI MAC