

Improving upon HCTR and matching attacks for Hash-Counter-Hash approach

Mridul Nandi

Indian Statistical Institute, Kolkata
mridul.nandi@gmail.com

Abstract. McGrew and Fluhrer first proposed hash-counter-hash approach to encrypt arbitrary length messages. By its nature, counter can handle incomplete message blocks as well as complete message blocks in the same manner. HCTR is the till date best (in terms of efficiency) strong pseudo random permutation or SPRP among all known counter based SPRPs. But as of now, a cubic bound for HCTR is known. Moreover, all invocations of underlying block ciphers can not be made in parallel. Our new proposal (we call it HMC or Hash Modified Counter) provides a quadratic security bound and all block cipher invocations are parallel in nature even if we have an incomplete message block. We also present a prp-distinguishing attack on a generic counter based encryption, which makes q non-adaptive encryption queries consisting of $(\ell + 1)$ n -bit blocks and has success probability roughly $\ell^2 q^2 / 2^n$. Loosely speaking, the success probability matches with the upper bound of distinguishing probability. As a result, we prove that the known quadratic bounds for XCB, HCH and HMC are tight.

Keywords: Strong pseudo random permutation, modes of operation, HCTR, distinguishing attack.

1 Introduction

A mode of operation is a method of constructing an encryption algorithm which can encrypt arbitrary length messages. It uses a cryptographic object called block cipher, as an underlying object and possibly some algebraic operations such as finite field multiplication. (Strong) Pseudo Random Permutation or (S)PRP [12], authenticity and privacy [8] are some of the desired security notions for symmetric key encryptions. Later, Liskov et al. [10] followed by Halevi-Rogaway [7] considered tweakable version of length-preserving SPRP, which allows us to process associated data or tweak as a part of the messages. Disk-encryption is one of the important application for the length-preserving tweakable SPRP as mentioned in [7]. Motivated by disc-encryption algorithms, there are several tweakable-SPRP proposals. We list some of these important constructions based on three categories.

1. Hash-Encrypt-Hash: First introduced by Naor-Reingold [17, 18], consists of ECB layer between two layers of invertible hash. Similar approach is considered in TET [5] and HEH [20], where latter is an improvement over TET.
2. Encrypt-Mix-Encrypt: Halevi-Rogaway [7] introduced ECB, mixing and ECB approach. Some of the constructions in this types are CMC [7], EME [6] and EME* [4] (modification of EME which can encrypt arbitrary size messages).
3. Hash-Ctr-Hash: This approach is first observed in the original proposal of XCB [15]. Later, HCTR [22], HCH [20] and a new version of XCB with security bound [16] are of this type. The first hash function layer is to generate counter. Based on the counter, we obtain ciphertext except one block which is computed by using the second hash layer. In this paper, we are mainly interested in this type of modes of operations.

1.1 Brief outline of the paper

HMC: a new hash-counter-hash encryption proposal This paper is broadly classified into two parts. In the first part we propose an improvement over HCTR which is itself one of the most efficient encryption algorithms. The improvement is made in efficiency as well as security. We call this new construction as HMC or Hash Modified Counter. To be precise we modify the definition of counter in HCTR. This is why we name it as Hash Modified Counter. We show that encryption algorithm of HMC is more efficient than HCTR in both hardware and software (all block cipher invocations can be made parallel for any type of messages). We also prove that HMC has quadratic security bound, roughly $\ell^2 q^2 / 2^n$ where ℓ is the number of n -bit blocks in the longest query among all q queries. Whereas, as of now a cubic security bound $\sigma^3 / 2^n$ for HCTR is known where $\sigma = O(\ell q)$ is the number of message blocks in all q queries.

The new construction is having similar performance as HEH when we have messages of size multiple of n . But HMC is more efficient and suitable than HEH when we need to process messages of size not multiple of n . We consider several issues related to modes of operations like hardware and software efficiency, key-storage, internal storage, simplicity and security bounds. We choose one of the efficient candidates from each class. In particular, HEH from the Hash-Encrypt-Hash type, EME from the Encrypt-Mix-Encrypt class and HCTR from the hash-counter-hash type. A detailed comparison among all these constructions along with the new construction HCM is provided and comparison is mainly based on those issues.

Distinguishing attack on hash-counter-hash encryption : In the second part, we consider the opposite direction. We first define a wide generic class of counter based encryption and then provide a distinguishing attack which can distinguish a counter-based encryption from a random permutation with advantage roughly $\ell^2 q^2 / 2^n$ where q is the number of queries and ℓ is the number of blocks of each query. We already know similar bound for XCB [16], HCH [2] and HMC (in this paper) as a maximum distinguishing advantage. Thus we prove that this quadratic bound is in fact tight. In other words, the known security bounds for these constructions can not be further improved. As of now, this is the first result showing distinguishing attack as well as the tightness of the bound for strong pseudorandom permutations of arbitrary length (i.e. advantage includes the length of queries).

Organization of the paper. In section 2, we provide a brief discussion on the two underlying function namely, hash function and counter function which have been used in hash-counter-hash approach. In section 3, we provide our new proposal called HMC and provide a detailed comparison with some of the best constructions known so far. In the following section, we provide the security analysis of this new construction. We define a large class of counter based encryption algorithm and state a quadratic distinguishing attack in section 5. Finally we conclude with possible future research works.

2 Hash functions and Counter functions

poly hash function

We first define poly-hash [21] which is an useful algebraic object for cryptography. Let n be the block-size of an underlying block cipher. In this paper we identify the Galois field \mathbb{F}_{2^n} and $\{0, 1\}^n$

and we use \oplus for bit-wise xor and $*$ for the field multiplication¹. We represent $\mathbf{0} = 0^n$ for the additive identity and $\mathbf{1}$ for the multiplicative identity. Let $h, X_i \in \mathbb{F}_{2^n}$, $1 \leq i \leq m$. We define

$$\mathcal{H}_h^{\text{poly}}(X_1, \dots, X_m) = X_m \oplus X_{m-1} * h \oplus \dots \oplus X_1 * h^{m-1}.$$

Trivially for any $i \geq 1$, $\mathcal{H}_h^{\text{poly}}(X) = \mathcal{H}_h^{\text{poly}}(\mathbf{0}^i, X)$ for all $h \in \mathbb{F}_{2^n}$. In the following result, we state that this is the only possibility to get same hash values with probability one. The proof is a direct application of fundamental theorem of algebra. In this paper, we denote \mathbf{h} to represent the uniform random variable taking values on $\mathbb{F}_{2^n}^* := \mathbb{F}_{2^n} \setminus \{\mathbf{0}\}$.

Lemma 1. *Let $X = (X_1, \dots, X_m) \in \mathbb{F}_{2^n}^m$ and $X' = (X'_1, \dots, X'_{m'}) \in \mathbb{F}_{2^n}^{m'}$, $X \neq X'$ and $m' \leq m$. If $X_1 \neq \mathbf{0}$ then $\Pr_{\mathbf{h}}[\mathcal{H}_{\mathbf{h}}^{\text{poly}}(X) = \mathcal{H}_{\mathbf{h}}^{\text{poly}}(X')] \leq \frac{m-1}{2^n-1}$.*

Corollary 1. *Let $X = (X_1, \dots, X_m) \in \mathbb{F}_{2^n}^m$ and $X' = (X'_1, \dots, X'_{m'}) \in \mathbb{F}_{2^n}^{m'}$ such that $m' \leq m$ and $X \neq X'$. Moreover, $\tau_k : \mathbb{F}_{2^n}^k \rightarrow \mathbb{F}_{2^n}^k$ is a permutation for all $k \geq 1$. Then, $\Pr_{\mathbf{h}}[\mathcal{H}_{\mathbf{h}}^{\text{poly}}(\mathbf{1}, \tau_m(X)) = \mathcal{H}_{\mathbf{h}}^{\text{poly}}(\mathbf{1}, \tau_{m'}(X'))] \leq \frac{m-1}{2^n-1}$.*

Horner's rule To compute $X_m \oplus X_{m-1} * h \oplus \dots \oplus X_1 * h^{m-1}$, one can use Horner's rule.

- $Y \leftarrow X_1$;
- for $i = 1$ to $m - 1$
 - $Y \leftarrow (Y * h) \oplus X_{i+1}$;
- end for
- return Y ;

We need $m - 1$ multiplication to compute the hash value. Note that, when $X_1 = \mathbf{1}$ (corresponding to monic polynomial) we do not need first multiplication in the loop. In this case we need, only $m - 2$ multiplications for $m \geq 2$.

Counter-function

Notations For $a, b \in \{0, 1\}^n$, $a + b$ is defined to be $(a + b) \bmod 2^n$. We denote $\overline{X}_m = X_m \mathbf{0}^i$ where $i = n - |X_m|$. If $Z \in \{0, 1\}^n$ then we denote $Z \oplus X_m$ for the first $|X_m|$ bits of $Z \oplus \overline{X}_m$. The n -bit standard binary representation of an integer $0 \leq i < 2^n$ is given by $\text{bin}_n(i)$. We also denote $[a, b] = \{a, a + 1, \dots, b\}$ for two integers $a \leq b$ and $\text{Bin}_{[a, b]} = \{\text{bin}_n(i) : i \in [a, b]\}$. We simply denote Bin_x for $\text{Bin}_{[0, x]}$. Moreover, given a set $A \subseteq \{0, 1\}^n$, and $S \in \{0, 1\}^n$ we define $S \oplus A = \{S \oplus a : a \in A\}$ and $S + A = \{S + a : a \in A\}$.

Let $E : \{0, 1\}^\kappa \times \{0, 1\}^n \rightarrow \{0, 1\}^n$ be a block-cipher (in other words, it is a keyed family of permutation which are efficiently computable). The function $E_K(\cdot) := E(K, \cdot)$ is a permutation on $\{0, 1\}^n$. Now we state two examples of *counter functions* which have been used in many counter modes of encryptions such as HCTR, HCH and XCB etc.

Example 1. (Modulo Counter) $\text{Ctr}_{K, S}(X_2, \dots, X_m) = (C_2, \dots, C_m)$, where $C_i = E_K(S + i - 2) \oplus X_i$, $|X_2| = \dots = |X_{m-1}| = n$, $|X_m| \leq n$. The set of inputs of E_K for computing counter function is $S + [0, m - 2]$.

¹ Once we fix an irreducible polynomial of degree n and the representation of \mathbb{F}_{2^n} as $\{0, 1\}^n$, field multiplication operation on $\{0, 1\}^n$ are determined.

Example 2. (XOR Counter) $\text{XCtr}_{K,S}(X_2, \dots, X_m) = (C_2, \dots, C_m)$, where $C_i = E_K(S \oplus \text{bin}_n(i-1)) \oplus X_i$, $|X_2| = \dots = |X_{m-1}| = n$, $|X_m| \leq n$. Here, the sets of inputs of E_K is $S \oplus \text{Bin}_{[1, m-1]}$. One can similarly define $C_i = E_K(S \oplus \text{bin}_n(i-2)) \oplus X_i$, $2 \leq i \leq m$, and hence the set of inputs is $S \oplus \text{Bin}_{m-2}$. We denote this variant of xor-counter function by $\text{XCtr}'_{K,S}$.

In these above examples, S is also called as the *counter* for corresponding counter-function. All inputs of E_K are derived from this counter either by adding or xor-ing some distinct constants. We state a simple and useful observation on these derived elements.

Lemma 2.

1. $(S + [0, k]) \cap (S' + [0, k']) \neq \emptyset \implies (S - S') \in [-k, k']$.
2. $(S \oplus \text{Bin}_k) \cap (S' \oplus \text{Bin}_{k'}) \neq \emptyset \implies (S \oplus S') \in \text{Bin}_x$ where $x = 2 \times \max\{k, k'\} - 1$.

Proof. The first part is straightforward from the definition of $S + [0, k]$ and $S' + [0, k']$. We prove the second part here. We identify n -bit binary sequences and integers in a standard way. Without loss of generality, let us assume that $k' \leq k$. Let b denote the minimum number of bits needed to represent k . In other words, $b = \lfloor \log_2 k \rfloor + 1$. If $z = \text{bin}_n(a) \oplus \text{bin}_n(b)$ where $0 \leq a \leq k$ and $0 \leq b \leq k'$ then z can also be represented by b bits. Hence $z \leq 2^b - 1 \leq 2k - 1 = x$. Thus, $S \oplus a = S' \oplus a'$ implies $S \oplus S' = a \oplus a' \in \text{Bin}_x$. ■

Definition 1. Let S^i be a random variable taking values on $\{0, 1\}^n$, $1 \leq i \leq q$. We say a q -tuple of random variables (S^1, \dots, S^q) is ϵ -AXU if

$$\Pr[S^i \oplus S^j = \alpha] \leq \epsilon \quad \forall i \neq j, \text{ and } \forall \alpha \in \{0, 1\}^n.$$

Now we prove an important result based on the above observation (Lemma 2).

Proposition 1. Let ℓ_1, \dots, ℓ_q be q positive integers and $\ell = \max_i \ell_i \leq 2^{n/2}$. If (S^1, \dots, S^q) is $\frac{\ell+1}{2^n-1}$ -AXU then the q sets, $S^i \oplus \text{Bin}_{m_i-1}$, $1 \leq i \leq q$ are disjoint with probability at least $1 - \frac{\ell^2 q^2}{2^n}$.

Proof. We show that the complement probability is at most $\frac{\ell^2 q^2}{2^n}$. For any pair (i, i') with $i \neq i'$, we define the following collision event

$$\text{Coll}_{i,i'} : (S^i \oplus \text{Bin}_{\ell_i-1}) \cap (S^{i'} \oplus \text{Bin}_{\ell_{i'}-1}) \neq \emptyset.$$

Thus, $\text{Coll}_{i,i'}$ implies $S^i \oplus S^{i'} \in \text{Bin}_{2\ell-3}$ (from Lemma 2). Since $|\text{Bin}_{2\ell-3}| = 2(\ell-1)$, $\Pr[\text{Coll}_{i,i'}] \leq 2(\ell-1) \times \frac{(\ell+1)}{2^n-1} \leq \frac{2\ell^2}{2^n}$. Since there are less than $q^2/2$ pairs of (i, i') ,

$$\Pr\left[\bigcup_{1 \leq i \neq i' \leq q} \text{Coll}_{i,i'} \right] \leq \frac{\ell^2 q^2}{2^n}.$$

The complement of $\bigcup_{1 \leq i \neq i' \leq q} \text{Coll}_{i,i'}$ corresponds to the event where $S^i \oplus \text{Bin}_{\ell_i-1}$, $1 \leq i \leq q$ are disjoint. Hence proved. ■

step-1 $V = \mathcal{H}_h^{\text{poly}}(\mathbf{1}, C_2, \dots, \overline{C}_m, C_1)$;
step-2 $S = E_K^{-1}(V) \oplus \text{bin}_n(m-1)$;
step-3 $X_i = E_K(S \oplus \text{bin}_n(i-2)) \oplus C_i, 2 \leq i \leq m$;
step-4 $X_1 = (h^{-1} * S) \oplus \mathcal{H}_h^{\text{poly}}(\mathbf{1}, X_2, \dots, \overline{X}_m, \mathbf{0})$;

Remark 1. The set of all inputs for E_K for $\text{HMC}_{E_K, h}$ is $\{S, S \oplus \text{bin}_n(1), \dots, \text{bin}_n(m-1)\} = S \oplus \text{Bin}_{m-1}$. Thus, all m computation of E_K can be made in parallel. We also have only one type of input for E_K (unlike HCTR, where we need two different types of inputs, U and the elements derived from the counter). This helps us to get much simpler hardware design for HMC.

MCTR_{K,h}(S, X₂, ..., X_m)

```

001 Sum ← h;
002 C2 ← EK(S) ⊕ X2;
003 for i = 3 to m
004   Ci ← EK(S ⊕ binn(i-2)) ⊕ Xi;
005   Sum ← (Sum ⊕ Ci-1) * h;
006 end for
007 C1 ← EK(S ⊕ binn(m-1)) ⊕ Sum;
008 return (C1, ..., Cm);

```

Fig. 1. Modified xor-Counter function. In AES pipeline implementation for E_K , the line 004 and the line 005 can be done parallel in each clock-cycle

Tweakable mode

For a tweakable mode we can compute the counter $S = \mathcal{H}_h^{\text{poly}}(\mathbf{1}, T, X_2, \dots, \overline{X}_m, X_1, \mathbf{0})$ where T is one block tweak. One can similarly extend for arbitrary length and in that case we need to provide length of the tweak as an input. One can find a standard literature how to consider tweak in [10, 7, 2, 22]. As in HCH [2], we can also encrypt tweak by block cipher (when tweak has size at most n -bits) and then we simply xor with hash value. A similar steps should be done for the second hash layer. Security analysis for tweakable case can be done similarly. To have a simple and clear proof to the readers we will only consider non-tweak case.

3.3 HMC is an improvement over HCTR

Table 1 is providing a comparison among EME*, HEH, HCTR and HMC based on several efficiency and design issues. The improvement is made in the following points.

1. **all encryptions are parallel.** We make all invocation of the underlying block cipher parallel even if we need to encrypt the partial block messages. This is not present in HCTR and in

HEH (in the case of partial block messages). In fact, HMC is the first SPRP-construction which has complete parallel invocations of block ciphers while encrypting incomplete message blocks. Thus, a pipeline implementation of AES can make it most efficient construction so far.

2. **quadratic security bound.** It has quadratic security bound. In particular, we show that any distinguisher \mathcal{A} has advantage of distinguishing at HMC from an uniform random permutation at most $(5\ell^2 + 1)^2 q^2 / 2^n$, where \mathcal{A} can make at most q queries with ℓ as the number of blocks of the longest query. On the other hand, HCTR has cubic security bound. Thus, our construction guarantees more security than HCTR.
3. **simpler design.** The design is simpler than HCTR (HCTR has simpler design also) and it is broadly broken into two main steps. In first step, we compute the counter (by using only field multiplication) and then based on the counter we compute the complete cipher text. In HCTR, there are two types of AES inputs whereas in new construction we have only one type of AES input. This helps us to have an efficient hardware design and fewer amounts of multiplexers.
4. **a better hash function.** We choose hash function different from that of HCTR. By choosing the new hash function we do not need any information of length of the message (which we need for HCTR and some other constructions). This helps us in two ways. Firstly, we have again a simpler hardware design for the hash function. Secondly, we also save one multiplication which is indeed an advantage for software implementation (for hardware, we can save one clock-cycle as Karatsuba-Ofman [9] implementation of field multiplication needs one clock cycle.). Moreover, we choose h as a part of the key and hence we need constant multiplier. A constant multiplier is always having advantage in hardware implementation than variable multiplier (which is used in one variant of HEH to save one key).

Table 1. Comparison among EME*, HEH, HCTR and the new construction HMC. Here, [BC] denotes block cipher invocations and [M] denotes finite field multiplication. Constant multiplier has one input as constant.

Parameter	EME*[4]	HEH [20]	HCTR [22]	HMC
Block cipher key	1	1	1	1
Auxiliary key	2	0	1	1
computational cost	$2m + \frac{m}{n}[\text{BC}]$	$(m + 1)[\text{BC}] + 2(m - 1)[\text{M}]$	$m[\text{BC}] + 2m[\text{M}]$	$m[\text{BC}] + (2m - 1)[\text{M}]$
clock-cycle (32 blocks)	107	75	89	76
Block-cipher layer (enc)	≥ 3	1 (or 2 for incomplete block)	2	1
Block-cipher layer (dec)	≥ 3	1 (or 2 for incomplete block)	2	2
multiplier	NO	YES	YES but const.	YES but const.
length as input	NO	YES	YES	NO

4 Security Analysis of HMC

4.1 Security notions for length-preserving enciphering schemes

Length-preserving enciphering scheme Let $A, B \subseteq \{0, 1\}^*$. A function $f : A \rightarrow B$ is called length-preserving if for all $a \in A$, $|a| = |f(a)|$. Formally, a tweakable enciphering scheme is a function $\mathbf{E} : \mathcal{K} \times \mathcal{T} \times \mathcal{M} \rightarrow \mathcal{M}$, where $\mathcal{K} \neq \emptyset$ and $\mathcal{T} \neq \emptyset$ are the key space and the tweak space

respectively such that for each key $K \in \mathcal{K}$ and each tweak $T \in \mathcal{T}$, $\mathbf{E}_K^T(\cdot) := \mathbf{E}(K, T, \cdot) : \mathcal{M} \rightarrow \mathcal{M}$ is a length-preserving permutation. The message and the cipher spaces are \mathcal{M} . The inverse of an enciphering scheme is $\mathbf{D} = \mathbf{E}^{-1}$ where $X = \mathbf{D}_K^T(Y)$ if and only if $\mathbf{E}_K^T(X) = Y$.

Uniform random permutation Let $\text{Perm}(n)$ denote the set of all permutations on $\{0, 1\}^n$ and Π corresponds to the uniform distribution on the set $\text{Perm}(n)$. It is easy to see that for any distinct $x_1, \dots, x_\sigma \in \{0, 1\}^n$ and distinct $y_1, \dots, y_\sigma \in \{0, 1\}^n$, we have $\Pr[\Pi(x_1) = y_1, \dots, \Pi(x_\sigma) = y_\sigma] = \frac{1}{N(N-1)\dots(N-\sigma+1)}$ where $N = 2^n$. The uniform random permutation Π is the ideal candidate for block cipher.

Ideal enciphering schemes In what follows, by the notation $X \stackrel{\$}{\leftarrow} \mathcal{S}$, we will denote the event of choosing X uniformly at random from the set \mathcal{S} when $|\mathcal{S}|$ is finite. Let $\text{Perm}^{\mathcal{T}}(\mathcal{M})$ denote the set of all functions $\pi : \mathcal{T} \times \mathcal{M} \rightarrow \mathcal{M}$ where $\pi(\mathcal{T}, \cdot)$ is a length preserving permutation. Such a $\pi \in \text{Perm}^{\mathcal{T}}(\mathcal{M})$ is called a tweak indexed permutation. Here, $\pi \stackrel{\$}{\leftarrow} \text{Perm}^{\mathcal{T}}(\mathcal{M})$ means that for each ℓ such that $\{0, 1\}^\ell \subseteq \mathcal{M}$ and $T \in \mathcal{T}$ we choose a tweakable random permutation π^T from $\text{Perm}(\ell)$ uniformly and independently.

Advantage of a distinguisher An adversary \mathcal{A} is a probabilistic algorithm which has access to some oracles and which outputs either 0 or 1. Oracles are written as superscripts. The notation $\mathcal{A}^{\mathcal{O}_1, \mathcal{O}_2} \Rightarrow 1$ denotes the event that the adversary \mathcal{A} , interacts with the oracles $\mathcal{O}_1, \mathcal{O}_2$, and finally outputs the bit 1.

We require a block cipher $E(\cdot)$ to be a strong pseudorandom permutation. The advantage of an adversary in breaking the strong pseudorandomness of $E(\cdot)$ is defined in the following manner.

$$\mathbf{Adv}_E^{\pm\text{prp}}(\mathcal{A}) = \left| \Pr \left[K \stackrel{\$}{\leftarrow} \mathcal{K} : \mathcal{A}^{E_K(\cdot), E_K^{-1}(\cdot)} \Rightarrow 1 \right] - \Pr \left[\pi \stackrel{\$}{\leftarrow} \text{Perm}(n) : \mathcal{A}^{\pi(\cdot), \pi^{-1}(\cdot)} \Rightarrow 1 \right] \right|.$$

For a tweakable enciphering scheme $\mathbf{E} : \mathcal{K} \times \mathcal{T} \times \mathcal{M} \rightarrow \mathcal{M}$, we define the advantage of an adversary \mathcal{A} at distinguishing \mathbf{E} and its inverse \mathbf{E}^{-1} from a random tweak indexed permutation and its inverse in the following manner.

$$\mathbf{Adv}_{\mathbf{E}}^{\pm\widetilde{\text{prp}}}(\mathcal{A}) = \left| \Pr \left[K \stackrel{\$}{\leftarrow} \mathcal{K} : \mathcal{A}^{\mathbf{E}_K(\cdot, \cdot), \mathbf{E}_K^{-1}(\cdot, \cdot)} \Rightarrow 1 \right] - \Pr \left[\pi \stackrel{\$}{\leftarrow} \text{Perm}^{\mathcal{T}}(\mathcal{M}) : \mathcal{A}^{\pi(\cdot, \cdot), \pi^{-1}(\cdot, \cdot)} \Rightarrow 1 \right] \right|. \quad (1)$$

We define $\mathbf{Adv}_{\mathbf{E}}^{\pm\widetilde{\text{prp}}}(q, \sigma, \ell)$ by $\max_{\mathcal{A}} \mathbf{Adv}_{\mathbf{E}}^{\pm\widetilde{\text{prp}}}(\mathcal{A})$ where maximum is taken over all distinguishers which makes at most q queries having at most σ many blocks and maximum ℓ blocks in a query.

For a computational advantage we define $\mathbf{Adv}_{\mathbf{E}}^{\pm\widetilde{\text{prp}}}(q, \sigma, \ell, t)$ by $\max_{\mathcal{A}} \mathbf{Adv}_{\mathbf{E}}^{\pm\widetilde{\text{prp}}}(\mathcal{A})$. In addition to the previous restrictions on \mathcal{A} , he can run in time at most t .

Pointless queries: We assume that an adversary never repeats a query, i.e., it does not ask the encryption oracle with a particular value of (T, P) more than once and neither does it ask the decryption oracle with a particular value of (T, C) more than once. Furthermore, an adversary never queries its deciphering oracle with (T, C) if it got C in response to an encipher query (T, P) for some P . Similarly, the adversary never queries its enciphering oracle with (T, P) if it got P as a response to a decipher query of (T, C) for some C . These queries are called *pointless* as the adversary knows what it would get as responses for such queries.

4.2 Security analysis for complete message blocks

We first make our analysis for complete message blocks without tweak. Later we state how same security analysis will work when a distinguisher can ask incomplete message blocks queries.

Definition 2. A tuple $v = (M^1, \dots, M^q, C^1, \dots, C^q)$ is called view if $M^1, \dots, M^q, C^1, \dots, C^q \in \mathbb{F}_{2^n}^+$ such that M^i 's are distinct and C^i 's are distinct and $\|M^i\| = \|C^i\|$ ($= \ell_i$, say). We write $M^i = (M_1^i, \dots, M_{\ell_i}^i)$ where $|M_1^i| = \dots = |M_{\ell_i}^i| = n$. Similarly we denote C_j^i 's. We say that a view is **good** if the collection of the following blocks are distinct :

$$A = \{M_j^i \oplus C_j^i : 2 \leq j \leq \ell_i, 1 \leq i \leq q\}$$

Now we define two class of blocks known as input blocks and output blocks. Intuitively, these are inputs and outputs for the block ciphers when we compute the encryption values of the messages M^1, \dots, M^q and obtain responses (C^1, \dots, C^q) . Define, $S^i = \mathcal{H}_h^{\text{poly}}(\mathbf{1}, M_2^i, \dots, M_{\ell_i}^i, M_1^i, \mathbf{0})$ is the counter for the message M^i . Similarly, we define $V^i = \mathcal{H}_h^{\text{poly}}(\mathbf{1}, C_2^i, \dots, C_{\ell_i}^i, C_1^i)$ when C^i is the ciphertext for M^i (it is same as $\Pi(S^i \oplus \text{bin}_n(\ell_i - 1))$).

- $\text{In}_h = \{S^i \oplus \text{bin}_n(j) : 0 \leq j \leq \ell_i - 1 \text{ and } 1 \leq i \leq q\} = \cup_{i=1}^q (S^i \oplus \text{Bin}_{\ell_i-1})$. It denotes the set of all inputs for Π while computing $\text{HMC}_{\Pi, h}(M^1), \dots, \text{HMC}_{\Pi, h}(M^q)$.
- $\text{Out}_h = A \cup \{V^i : 1 \leq i \leq q\}$. This is the set of all corresponding outputs of Π .

Denote two bad events Bad_{in} and Bad_{out} as the collision in the set In_h and Out_h respectively. From corollary 1, we know that (S^1, \dots, S^q) is $\frac{\ell+1}{2^n-1}$ -universal where $\ell = \max_i \ell_i$. Hence, by proposition 1,

$$\Pr[\text{Bad}_{in}] \leq \frac{\ell+1}{2^n-1} \times (2\ell-2) \times \binom{q}{2} \leq \frac{\ell^2 q^2}{2^n}.$$

Similarly, V^i is a non-constant monic polynomial of degree ℓ . Moreover, for $i \neq j$, either $V^i \oplus V^j$ is non-zero constant or it is the polynomial of degree at most ℓ . Thus, $\Pr_h[V^i = V^j] \leq \frac{\ell}{2^n-1}$ and $\Pr[V^i = \alpha] \leq \frac{\ell}{2^n-1}$.

$$\Pr[\text{Bad}_{out}] \leq \frac{\ell}{2^n-1} \times q(\sigma-1) \leq \frac{\ell^2 q^2}{2^n}.$$

Here we assume that A is derived from a good view. Let $\text{Bad} = \text{Bad}_{in} \cup \text{Bad}_{out}$. Thus, $\overline{\text{Bad}}$ implies that all inputs and outputs of Π (we assume that the underlying permutation is Π , an uniform random permutation) are distinct. Thus,

$$\Pr[\text{HMC}_{\Pi, h}(M^1) = C^1, \dots, \text{HMC}_{\Pi, h}(M^q) = C^q \mid \overline{\text{Bad}}] = \frac{1}{\mathbf{P}(2^n, \sigma)} \geq \frac{1}{2^{n\sigma}}$$

where $\sigma = \sum_i \ell_i$ and $\mathbf{P}(a, b)$ denotes the value $a(a-1)\dots(a-b+1)$. We have already observed that $\Pr_h[\text{Bad}] \leq \frac{2\ell^2 q^2}{2^n}$. Hence we have proved that

$$\Pr[\text{HMC}_{\Pi, h}(M^1) = C^1, \dots, \text{HMC}_{\Pi, h}(M^q) = C^q] \geq \frac{1 - 2\ell^2 q^2 / 2^n}{2^{n\sigma}}. \quad (2)$$

Now we prove the following result which says that for any distinguisher interacting with two independent length preserving uniform random function, obtains a good view with probability at least $1 - \frac{\sigma(\sigma-1)}{2^{n+1}}$.

Lemma 3. $\mathcal{A}^{\mathbf{F}_1, \mathbf{F}_2}$ is a distinguisher interacting with two independent length-preserving uniform random function. Suppose \mathcal{A} is not making any pointless query. Let $v = (M^1, \dots, M^q, C^1, \dots, C^q)$ be the view of \mathcal{A} . Then the probability that v is good is at least $1 - \frac{\sigma(\sigma-1)}{2^{n+1}}$

Proof. Since \mathcal{A} is making all distinct queries, the output distribution of any query is uniformly distributed. Hence $M_j^i \oplus C_j^i$ are uniformly and independently distributed over $\{0, 1\}^n$. Since there are σ many blocks there is a collision with probability at most $\sigma(\sigma - 1)/2^{n+1}$. ■

Let us denote \mathbf{E} for the HMC based on an uniform random permutation Π and \mathbf{h} chosen uniformly from $\{0, 1\}^n \setminus \{0\}$. Let \mathbf{E}^{-1} be its inverse function. We denote \mathbf{F}_1 and \mathbf{F}_2 for two independently distributed length-preserving uniform random functions.

Lemma 4. Let \mathcal{A} be an adversary which do not make any pointless query. \mathcal{A} is either interacting with \mathbf{E} and \mathbf{E}^{-1} or \mathbf{F}_1 and \mathbf{F}_2 . Now the distinguishing advantage of \mathcal{A} is at most $5\ell^2 q^2 / 2^{n+1}$.

Proof. Let V_1 be the set of all views on which \mathcal{A} outputs one. Let $v_{\mathbf{E}, \mathbf{E}^{-1}}$ denote the random variable view which is obtained when \mathcal{A} is interacting with \mathbf{E} and \mathbf{E}^{-1} . Similarly we denote $v_{\mathbf{F}_1, \mathbf{F}_2}$. Thus, advantage of \mathcal{A} is $\Pr[v_{\mathbf{F}_1, \mathbf{F}_2} \in V_1] - \Pr[v_{\mathbf{E}, \mathbf{E}^{-1}} \in V_1]$. Let V_1^G denote the set of all views from V_1 which are good. Now, for all $v \in V_1^G$, we have $\Pr[v_{\mathbf{F}_1, \mathbf{F}_2} = v] - \Pr[v_{\mathbf{E}, \mathbf{E}^{-1}} = v] \leq \frac{2\ell^2 q^2}{2^n} \times \Pr[v_{\mathbf{F}_1, \mathbf{F}_2} = v]$ (from equation 2). By using lemma 3,

$$\Pr[v_{\mathbf{F}_1, \mathbf{F}_2} \in V_1] - \Pr[v_{\mathbf{E}, \mathbf{E}^{-1}} \in V_1] \leq \frac{2\ell^2 q^2}{2^n} + \Pr[v_{\mathbf{F}_1, \mathbf{F}_2} \text{ is not good}] \leq \frac{5\ell^2 q^2}{2^{n+1}}. \quad \blacksquare$$

Theorem 1. For any q, ℓ , we have $\text{Adv}_{\mathbf{E}}^{\pm \text{prp}}(q, \sigma, \ell) \leq \frac{5(\ell^2+1)q^2}{2^{n+1}}$.

Proof. This is followed from above Lemma 4 and the fact that two independent length-preserving random functions can be distinguished from an uniform length-preserving permutation and its inverse by at most $q(q-1)/2^{n+1}$ probability (a statement with proof can be found in [7]). Thus, $\text{Adv}_{\mathbf{E}}^{\pm \text{prp}}(q, \sigma, \ell) \leq \frac{5\ell^2 q^2}{2^{n+1}} + q(q-1)/2^{n+1} \leq \frac{5(\ell^2+1)q^2}{2^{n+1}}$. ■

4.3 Security analysis in presence of incomplete message blocks

Here, we provide how one can make similar security analysis when distinguisher is allowed to have queries of length not multiple of n . Given $X = X_1 \parallel \dots \parallel X_{m-1} \parallel X_m$ where $|X_1| = \dots = |X_{m-1}| = n$ and $|X_m| \leq n$. Now we define modified encryption/decryption algorithm as $\text{HMC}'(X) = \text{HMC}(X) \parallel R$, where R is an independently distributed string of size $n - |X_m|$ and HMC denotes both encryption and decryption algorithm for HMC. In other words, we add remaining $n - |X_m|$ bits which is uniformly and independently distributed. Similarly, we define a pair of two independent uniform random functions $\mathbf{F}' = (\mathbf{F}'_1, \mathbf{F}'_2)$ which returns nm bits for the query X . Distinguishing, HMC' from \mathbf{F}' is equivalent in distinguishing HMC from \mathbf{F} . By previous section, we can prove that all inputs and outputs of the underlying uniform random permutation are distinct with probability at least $(1 - 2\ell^2 q^2 / N) / N$. Hence, we have the following similar bound for variable length version.

Theorem 2. Let \mathcal{A} be an adversary which do not make any pointless query. \mathcal{A} is either interacting with \mathbf{E} and \mathbf{E}^{-1} (with input space $\{0, 1\}^{\geq n}$) or two independent length-preserving random functions. Now the distinguishing advantage of \mathcal{A} is at most $5\ell^2 q^2 / 2^{n+1}$. Hence, for any q, ℓ, σ , we have $\text{Adv}_{\mathbf{E}}^{\pm \text{prp}}(q, \sigma, \ell) \leq \frac{5(\ell^2+1)q^2}{2^{n+1}}$.

5 A distinguishing attack on a wide class of counter-based encryption

5.1 \mathcal{C} : A class of counter-based encryptions

Now we define a wide class of counter based encryption algorithms. It uses three functions $H(\cdot)$, $S(\cdot)$ and $B(\cdot)$ based on secret keys. These functions satisfy some conditions (stated in the encryption algorithm) so that decryption is also possible. In this paper, we are mainly interested on the behavior of the function $S : \{0, 1\}^{2n} \rightarrow \{0, 1\}^n$. We denote both function and functional value by S . All probability calculation are based on probability distribution of the secret keys of an enciphering scheme.

- We say that S is ϵ - Δ in first coordinate if for all $M_0 \neq M'_0 \in \{0, 1\}^n$ and $\alpha, N \in \{0, 1\}^n$ we have $\Pr[S(M_0, N) - S(M'_0, N) = \alpha] \leq \epsilon$ where “ $-$ ” is the inverse operation of 2^n -modular addition.
- Similarly, we say S is ϵ -AXU in first coordinate if for all $M_0 \neq M'_0 \in \{0, 1\}^n$ and $\alpha, N \in \{0, 1\}^n$ we have $\Pr[S(M_0, N) \oplus S(M'_0, N) = \alpha] \leq \epsilon$.

Encryption \mathbf{E} :

input. A message $M = M_0 \parallel M_1 \cdots \parallel M_\ell$ with $|M_0| = \cdots = |M_{\ell-1}| = n, |M_\ell| \leq n$ and $\ell \geq 0$. If $\ell = 0$, $M = M_0 \in \{0, 1\}^n$.

step-1 We compute $S = S(M_0, N)$ (called as counter) where $N = H(M_1, \dots, M_\ell) \in \{0, 1\}^n$.

step-2 If $\ell = 0$, then we skip this step. Otherwise, we compute the last ℓ blocks of ciphertexts² as

$$(C_1, \dots, C_\ell) = (E_K(S_1) \oplus M_1, \dots, E_K(S_\ell) \oplus M_\ell).$$

1. If $S_i = S \oplus a_i$ we say \mathbf{E} is XOR-based encryption.
2. If $S_i = S + a_i$ we say \mathbf{E} is addition-based encryption.

step-3 The first block ciphertext C_0 is computed as $C_0 = B(M, C_1, \dots, C_\ell)$. The function B and S are defined in such a way so that S is also computable only from the ciphertext $C = (C_0, \dots, C_\ell)$. Moreover, M_0 should be computable from the ciphertext C , (M_1, \dots, M_ℓ) . B can use some intermediate computation in step-1 to make it more efficient.

Decryption \mathbf{E}^{-1} : We first compute the counter S from the ciphertext C (it is possible as described in step-3 of the encryption algorithm), and based on S we compute M_1, \dots, M_ℓ where $M_i = C_i \oplus E_K(S_i)$. Finally, we can compute M_0 from C and (M_1, \dots, M_ℓ) (again by requirements of the functions S and B as mentioned in step-3). This completes the decryption. A specific method to decrypt depends on the specification of the functions $S(\cdot)$ and $B(\cdot)$. Since we are interested in prp-distinguishing attack which needs only encryption query, we do not go to the detail of the decryption algorithm.

The class \mathcal{C}_1 . Let \mathcal{C}_1 be the the set of all counter based encryption algorithm \mathbf{E} defined as above where $S(\cdot)$ is $c/2^n$ - Δ (or $c/2^n$ -AXU) in first coordinate if \mathbf{E} is addition-based (or XOR-based respectively) and c is some constant.

Fig. 2. Encryption using HCH. K is the block-cipher key and h the hash key.

Algorithm $\mathbf{E}_{K,h}(M_0, M_1, \dots, M_\ell)$
1. $MM \leftarrow M_1 \oplus H_h(M_1 \dots M_\ell);$
2. $CC \leftarrow E_K(MM);$
3. $U \leftarrow MM \oplus CC$
4. $S = E_K(U)$
5. $(C_1, \dots, C_{\ell-1}, C_\ell)$ $\quad \leftarrow \mathbf{Ctr}_{K,S}(M_1, \dots, M_\ell);$
6. $C_0 \leftarrow CC \oplus H_h(C_1 C_2 \dots C_\ell);$
7. return $(C_0, C_1, \dots, C_\ell).$

Fig. 3. Encryption using XCB. $K = (K_1, K_2, K_3)$ is a tuple of block-cipher keys and h the hash key.

Algorithm $\mathbf{E}_{K,h}(M_0, M_1, \dots, M_\ell)$
1. $S \leftarrow E_{K_1}(M_1) \oplus H_h(M_1 \dots M_\ell);$
2. $(C_1, \dots, C_{\ell-1}, C_\ell)$ $\quad \leftarrow \mathbf{Ctr}_{K_2,S}(M_1, \dots, M_\ell);$
3. $C_0 \leftarrow E_{K_3}^{-1}(S \oplus H_h(C_1 C_2 \dots C_\ell));$
4. return $(C_0, C_1, \dots, C_\ell).$

Examples : HCTR, HCH, XCB, HMC

We define HCH, XCB encryption algorithms in the figures 2 and 3 respectively. We define it for without tweak. For a complete definition with tweak one can refer [2, 16]. Now we see that these encryptions along with HCTR [22] and HMC (stated in section 3) are included in the class defined above. The function H is nothing but different variants of poly-hash function for all four encryption schemes. The counter S is defined as

$$\text{XCB} : S(M_0, N) = E_K(M_0) \oplus N,$$

$$\text{HMC} : S(M_0, N) = h * M_0 \oplus N,$$

$$\text{HCTR} : S(M_0, N) = E_K(M_0 \oplus N) \oplus M_0 \oplus N,$$

$$\text{HCH} : S(M_0, N) = E_K(E_K(M_0 \oplus N) \oplus M_0 \oplus N).$$

S_i 's are computed from the counter S as follows :

- In case of HCTR and HCH, $S_i = S \oplus \text{bin}_n(i)$ and in case of HMC $S_i = S \oplus \text{bin}_n(i-1)$, $1 \leq i \leq \ell$. We call this type of encryption as XOR-based encryption.
- In case of XCB, $S_i \equiv (S + i - 1) \pmod{2^n}$ (more precisely, $\pmod{2^{32}}$ for the last 32-bits). We call this type of encryption as addition based encryption.

Now we show that the function $S(\cdot)$ for the above encryptions are $2/2^n$ - Δ or $5/2^n$ -AXU in first coordinate. More precisely, we have the following lemma. We denote $N = 2^n$.

² Recall that, $E_K(S_\ell) \oplus M_\ell = E_K(S_\ell)_{|M_\ell|} \oplus M_\ell$ where $E_K(S_\ell)_{|M_\ell|}$ denotes the first $|M_\ell|$ bits of $E_K(S_\ell)$.

Lemma 5. *We assume that the underlying block cipher E_K is an uniform random permutation and hash key \mathbf{h} is chosen uniformly from $\mathbb{F}_{2^n}^*$. In case of HMC, HCTR and HCH, $S(\cdot)$ is $5/2^n$ -AXU in first coordinate, i.e., for all $M_0^i \neq M_0^{i'} \in \{0, 1\}^n$ and $\alpha \in \{0, 1\}^n$,*

$$\Pr[S(M_0^i, N_0) \oplus S(M_0^{i'}, N_0) = \alpha] \leq 5/2^n.$$

For XCB, $S(\cdot)$ is $2/2^n$ - Δ in first coordinate. That is, for all $M_0^i \neq M_0^{i'} \in \{0, 1\}^n$ and $\alpha \in \{0, 1\}^n$,

$$\Pr[S(M_0^i, N_0) - S(M_0^{i'}, N_0) = \alpha] \leq 2/2^n.$$

Proof.

- XCB : One can see that there are N possible pair of values for $S(M_0^i, N_0)$, $S(M_0^{i'}, N_0)$ such that $S(M_0^i, N_0) - S(M_0^{i'}, N_0) = \alpha$. Note that, $S(M_0^i, N_0) = E_K(M_0^i) \oplus N_0$ and $S(M_0^{i'}, N_0) = E_K(M_0^{i'}) \oplus N_0$ and $M_0^i \neq M_0^{i'}$. Thus, for each such choice the probability is $1/(N(N-1))$ (here we assume that the underlying permutation is an uniform random permutation). So, for any given $\alpha \neq 0$ and $i \neq i'$, $\Pr[S(M_0^i, N_0) - S(M_0^{i'}, N_0) = \alpha] = 1/(N-1) \leq 2/N$.
- HMC : For any given α and $i \neq i'$, $\Pr[S(M_0^i, N_0) \oplus S(M_0^{i'}, N_0) = \alpha] = \Pr[\mathbf{h} * (M_0^i \oplus M_0^{i'}) = \alpha] \leq 1/2^{n-1} \leq 2/2^n$.
- HCTR : For any given α and $i \neq i'$, $\Pr[S(M_0^i, N_0) \oplus S(M_0^{i'}, N_0) = \alpha] = \Pr[E_K(M_0^i \oplus N_0) \oplus E_K(M_0^{i'} \oplus N_0) = M_0^i \oplus M_0^{i'} \oplus \alpha]$. By similar argument as in XCB, one can show that $\Pr[E_K(M_0^i \oplus N_0) \oplus E_K(M_0^{i'} \oplus N_0) = M_0^i \oplus M_0^{i'} \oplus \alpha] \leq 1/(N-1)$. Hence c for HCTR is 2.
- HCH : Recall that, $S(M_0, N_0) = E_K(M_0 \oplus N_0 \oplus E_K(M_0 \oplus N_0))$. Let $v^i = E_K(M_0^i \oplus N_0)$ and $v^{i'} = E_K(M_0^{i'} \oplus N_0)$. Define an event E such that all inputs of E_K are distinct, i.e.,

$$E : M_0^i \oplus N_0, M_0^{i'} \oplus N_0, v^i \oplus M_0^i \oplus N_0, v^{i'} \oplus M_0^{i'} \oplus N_0 \text{ are distinct}$$

Given that, E is true, one can similarly prove that the conditional probability of $X^{i,i'} := E_K(v^i \oplus M_0^i \oplus N_0) \oplus E_K(v^{i'} \oplus M_0^{i'} \oplus N_0) = \alpha$ is $1/(N-1)$ provided $\alpha \neq \mathbf{0}$. Now, $\Pr[E] \geq (N-2)(N-3)/N(N-1)$. Thus,

$$\Pr[X^{i,i'} = \alpha] \leq 1 - \sum_{\beta \neq \alpha, \mathbf{0}} \Pr[X^{i,i'} = \beta] \leq 1 - \frac{(N-2)^2(N-3)}{N(N-1)^2} \leq \frac{5}{N}. \quad \blacksquare$$

5.2 Some known related results on PRP-advantages

In [2], it has been shown that for any PRP-distinguisher \mathcal{A} which makes q queries of $(\ell+1)$ blocks, the PRP-advantage for HCH based on an uniform random permutation is at most $7(\ell+1)^2 q^2 / 2^n = O(\ell^2 q^2 / 2^n)$. Similar results have been shown for XCB [16] and HMC (in the section 4.2). In case of XCB, the maximum PRP-advantage is bounded by $8q^2(\ell+3)^2 / 2^n$ and in case of HMC the maximum advantage is bounded by $5(\ell+2)^2 q^2 / 2^n$. For HCTR, a cubic bound is known so far. In the next section, we state a distinguisher which has the PRP-advantage $\Omega(\ell^2 q^2 / 2^n)$. By this result we prove that $\mathbf{Adv}_{\mathbf{E}}^{\pm \text{PRP}}(q, \sigma, \ell) = \Theta(\ell^2 q^2 / 2^n)$ where \mathbf{E} is either HMC or XCB or HCH.

5.3 Distinguishing Attacks

Distinguisher \mathcal{A} :

- step-1 It first chooses distinct $M_0^i \in \{0, 1\}^n$ and define $M_j^i = \mathbf{0}$ for $1 \leq i \leq q$, $1 \leq j \leq \ell$. Denote $M^i = (M_0^i, M_1^i, \dots, M_\ell^i)$.
- step-2 It makes encryption queries M^i for $1 \leq i \leq q$. Let $C^i = (D^i, Y^i) = (D^i, Y_1^i, \dots, Y_\ell^i)$ be the response corresponding to the message M^i where $|D^i| = |Y_j^i| = n$.
- step-3 Return 1 if all Y_j^i 's are distinct for $1 \leq i \leq q$, $1 \leq j \leq \ell$. Otherwise it returns 0.

The event DIST : It corresponds to the event where Y_j^i 's are distinct for $1 \leq i \leq q$, $1 \leq j \leq \ell$ when the above distinguisher runs.

Let $\mathbf{E} \in \mathcal{C}$. Let $p_1 := \Pr[\text{DIST is true}]$ when \mathcal{A}_1 is interacting with \mathbf{E} and p_2 denote the probability for same event when \mathcal{A}_1 is interacting with an uniform length-preserving permutation. By definition of advantage (see section 4), $\text{Adv}_{\mathbf{E}}^{\pm\text{PRP}}(\mathcal{A}_1) \geq |p_2 - p_1|$. Now we compute p_1 and p_2 and obtains an estimate of the advantage of the given distinguisher. Let $N = 2^n$.

An observation : When \mathcal{A} interacts with any XOR-based encryption for the above choices of messages we have that $Y_j^i = E_K(S^i \oplus \text{bin}_n(j)) \oplus M_j^i = E_K(S^i \oplus \text{bin}_n(j))$. Thus, DIST is true if and only if $S^i \oplus \text{bin}_n(j)$'s are distinct. Even if there are ℓq many possible values we actually need to compare $2\ell \times \binom{q}{2}$ times. This is because of the following main observation.

$$S^i \oplus \text{bin}_n(j) = S^{i'} \oplus \text{bin}_n(j') \Rightarrow (S^i \oplus S^{i'}) = \text{bin}_n(j) \oplus \text{bin}_n(j') \in \{\text{bin}_n(k) : 0 \leq k \leq 2\ell - 1\}.$$

One can similarly study for the modular-addition based encryption. Let $N_0 = H_h(\mathbf{0}, \dots, \mathbf{0})$. Thus, $S^i = S(M_0^i, N_0)$ where M_0^i 's are distinct. Since $S(\cdot)$ is $c/2^n$ -AXU, probability that DIST is not true is at most $2\ell \times \binom{q}{2} \times \frac{c}{2^n}$. Thus we have proved the following result.

$$p_2 := \Pr[\text{DIST is true}] \geq \left(1 - \frac{c\ell q(q-1)}{N}\right) \quad (3)$$

Computation of probability of DIST in case of an uniform random permutation

Let \mathbf{P} be an uniform random permutation on $\{0, 1\}^{n(\ell+1)}$. Thus, for any distinct $M^1, \dots, M^q \in \{0, 1\}^{n(\ell+1)}$, the joint probability distribution of $(\mathbf{P}(M^1), \dots, \mathbf{P}(M^q))$ is uniform over all q distinct $n(\ell+1)$ bits. More precisely, for any distinct C^1, \dots, C^q we have

$$\Pr[\mathbf{P}(M^1) = C^1, \dots, \mathbf{P}(M^q) = C^q] = \frac{1}{\mathbf{P}(N^{\ell+1}, q)}$$

where $N = 2^n$. We write $C^i = D^i \parallel Y^i$ where $|D^i| = n$. Let \mathcal{D} be the set of all q -tuples (C^1, \dots, C^q) such that the q -tuple (Y^1, \dots, Y^q) has no block-wise collision. Note that,

$$(Y^1, \dots, Y^q) \text{ has no block-wise collision} \implies C^i\text{'s are always distinct.}$$

Now we see that $|\mathcal{D}| = \mathbf{P}(N, \ell q) \times N^q$. This is true since we can choose all ℓq blocks of Y^i 's in $\mathbf{P}(N, \ell q)$ ways and then we can choose D^i 's all possible way i.e., in N^q ways.

Proposition 2. $\Pr[\text{DIST is true}] = \Pr[\mathbf{P}(M^1, \dots, M^q) \in \mathcal{D}] \leq (1 - \frac{(\ell q - 1)\ell q}{4N})$ where M^1, \dots, M^q are q distinct elements from $\{0, 1\}^{n(\ell+1)}$ and \mathcal{D} is defined as above. Here, we assume that $\frac{(\ell q - 1)\ell q}{4N} \leq 1$.

Proof. Roughly speaking, the complement is the collision event for randomly chosen ℓq elements. Thus, the complement probability is the order of $O(\ell^2 q^2 / 2^n)$. We know from the above discussion that

$$p_1 := \Pr[\mathbf{P}(M^1, \dots, M^q) \in \mathcal{D}] = \frac{\mathbf{P}(N, \ell q) \times N^q}{\mathbf{P}(N^{\ell+1}, q)} = \frac{\prod_{i=1}^{\ell q - 1} (1 - i/N)}{\prod_{i=1}^{q-1} (1 - i/N^{\ell+1})}.$$

Now the term $(1 - i/N)/(1 - i/N^{\ell+1})$ is less than $(1 - i/N + i/N^{\ell+1})$ for $1 \leq i \leq q - 1$. Thus,

$$p_1 \leq \prod_{i=1}^{q-1} (1 - i/N + i/N^{\ell+1}) \times \prod_{i=q}^{\ell q - 1} (1 - i/N) \leq \prod_{i=1}^{\ell q - 1} (1 - i/2N).$$

The last inequality is true provided $\ell \geq 1$. Now, $\prod_i (1 - a_i) \leq \exp(-\sum_i a_i) \leq 1 - \frac{1}{2} \times \sum_i a_i$ where $0 \leq a_i \leq 1$ and $\sum_i a_i \leq 1$. Since we assume that $\frac{(\ell q - 1)\ell q}{4N} \leq 1$, $p_1 \leq 1 - \frac{(\ell q - 1)\ell q}{4N}$. Hence proved. ■

Combining equation 3 and proposition 2 the distinguishing PRP-advantage of \mathcal{A}_1 for a counter-based $\mathbf{E} \in \mathcal{C}_1$ based on $c/2^n$ - Δ or $c/2^n$ -AXU function, is at least

$$p_2 - p_1 \geq \frac{(\ell q - 1)\ell q}{4N} - \frac{c\ell q(q - 1)}{N} \geq \frac{q^2 \ell (\ell - 4c)}{4N} \geq \frac{\ell^2 q^2}{8N}$$

provided $8c \leq \ell$. Since c is constant, we can choose ℓ more than $8c$.

Theorem 3. Let $\mathbf{E} \in \mathcal{C}$ with $c/2^n$ -function $S(\cdot)$ as described in the generic definition of counter based encryption where c is a constant. Then, $\mathbf{Adv}_{\mathbf{E}}(q, \sigma, \ell + 1) \geq \frac{q^2 \ell^2}{2^{n+3}}$ where $\ell \geq 8c$. Thus, $\mathbf{Adv}_{\mathbf{E}}(q, \sigma, \ell) = \Theta(\frac{q^2 \ell^2}{2^n})$ in case of $\mathbf{E} = \text{HMC}$ or XCB or HCTR .

6 Conclusion

The new counter based encryption HMC is having improved performance than HCTR. Moreover, it has quadratic security, whereas a cubic security bound for HCTR is known. Since the modified construction has similar nature as HCTR, an important question would be to find a quadratic security bound for HCTR. We also present a wide class of counter based encryption algorithms and present a distinguishing attack which has quadratic advantage on this class. Thus, we can not expect more than quadratic security from a counter based encryption. It looks possible to characterize the counter based subclass which has exactly quadratic security bound. Outside counter based enciphering schemes there are several good constructions with quadratic security bound. It would be nice to find a quadratic advantage for those constructions also.

References

1. D. Chakraborty and P. Sarkar. A new mode of encryption providing a strong tweakable pseudo-random permutation. Fast Software Encryption FSE 2006, LNCS vol. 4047, Springer, pp. 293309, 2006.

2. D. Chakraborty and P. Sarkar. HCH: A new tweakable enciphering scheme using the Hash- Encrypt-Hash approach. *Advances in Cryptology INDOCRYPT 2006*, LNCS vol. 4329, Springer, pp. 287302, 2006.
3. Joan Daemen and Vincent Rijmen. The Design of Rijndael: AES The Advanced Encryption Standard. Springer 2002. <http://csrc.nist.gov/CryptoToolkit/aes/rijndael/Rijndael-ammended.pdf>
4. S. Halevi. EME*: Extending EME to handle arbitrary-length messages with associated data. *Advances in Cryptology INDOCRYPT 2004*, LNCS vol. 3348, Springer, pp. 315327, 2004.
5. S. Halevi. TET: A wide-block tweakable mode based on Naor-Reingold. *Advances in Cryptology 2007*. Cryptology ePrint archive, report 20007/14, 2007.
6. S. Halevi and P. Rogaway. A parallelizable enciphering mode. *Topics in Cryptology CT-RSA 2004*, LNCS vol. 2964, Springer, pp. 292304, 2004.
7. S. Halevi and P. Rogaway. A tweakable enciphering mode. *Advances in Cryptology CRYPTO 2003*, LNCS vol. 2729, Springer, pp. 482499, 2003.
8. C. Jutla. Encryption modes with almost free message integrity. *Advances in Cryptology EUROCRYPT 2001*. Lecture Notes in Computer Science, vol. 2045, Springer-Verlag, 2001.
9. A. Karatsuba and Y. Ofman. Multiplication of multidigit numbers by automata. *Soviet Physics-Doklady*, 7:595596, 1963.
10. M. Liskov, R. Rivest, and D. Wagner. Tweakable block ciphers. *Advances in Cryptology CRYPTO 2002*, LNCS vol. 2442, Springer, pp. 3146, 2002.
11. C. M. Lopez, D. Chakraborty and F. R. Henriquez. Efficient Implementations of Some Tweakable Enciphering Schemes in Reconfigurable Hardware. To appear in *Indocrypt-2007*.
12. M. Luby and C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal of Computing*, vol. 17, no. 2, pp. 373386, 1988.
13. S. Lucks. Faster Luby-Rackoff ciphers. *Fast Software Encryption 1996*, LNCS vol. 1039, Springer, pp. 189203, 1996.
14. D. McGrew and S. Fluhrer. The extended codebook (XCB) mode of operation. Cryptology ePrint archive, report 2007/298. To appear in *Proceedings in Selected Areas in Cryptography*.
15. McGrew and S. Fluhrer, The Extended Codebook (XCB) Mode of Operation, Cryptology ePrint Archive: Report 2004/278, October 25, 2004. <http://eprint.iacr.org/2004/278>
16. D. McGrew and S. Fluhrer. The extended codebook (XCB) mode of operation. Cryptology ePrint archive, report 2007/298. To appear in *Proceedings in Selected Areas in Cryptography*.
17. Moni Naor and Omer Reingold. A pseudo-random encryption mode. Manuscript available from www.wisdom.weizmann.ac.il/naor.
18. M. Naor and O. Reingold. On the construction of pseudorandom permutations: Luby-Rackoff revisited. *Journal of Cryptology*, vol. 12, no. 1, pp. 2966, 1999.
19. P. Rogaway. Authenticated-encryption with associated-data. Ninth ACM Conference on Computer and Communications Security (CCS-9). ACM Press, 2002.
20. Palash Sarkar. Improving Upon the TET Mode of Operation. Cryptology ePrint archive, report 2007/317.
21. Victor Shoup. On fast and provably secure message authentication based on universal hashing. CRYPTO-1996, volume 1109, Lecture Notes in Computer Science, pages 313328. Springer, 1996.
22. P. Wang, D. Feng, and W. Wu. HCTR: a variable-input-length enciphering mode. *Information Security and Cryptography, CISC 2005*, LNCS vol. 3822, Springer, pp. 175188, 2005.