Noname manuscript No.
(will be inserted by the editor)

A Tight Asymptotic Bound on the Size of Constant-Weight Conflict-Avoiding Codes

Kenneth W. Shum · Wing Shing Wong

the date of receipt and acceptance should be inserted later

Abstract In the study of multiple-access in the collision channel, conflict-avoiding code is used to guarantee that each transmitting user can send at least one packet successfully in the worst case within a fixed period of time, provided that at most k users out of M potential users are active simultaneously. The number of codewords in a conflict-avoiding code determines the number of potential users that can be supported in a system. Previously, upper bound on the size of conflict-avoiding code is known only for Hamming weights three, four and five. The asymptotic upper in this paper extends the known results to all Hamming weights, and is proved to be tight by exhibiting infinite sequences of conflict-avoiding codes which meet this bound asymptotically for all Hamming weights.

Keywords Conflict-avoiding codes · Protocol sequences · Collision channel without feedback

Mathematics Subject Classification (2000) 94B65 · 94B25 · 11P99

1 Introduction

Let $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$ denote the group of residues modulo n and $\mathcal{P}_w^n$ the collection of all subsets of size w in $\mathbb{Z}_n$. For a subset A of $\mathbb{Z}_n$, we define the *set of differences* between pairs of distinct elements in A as

$$d^*(A) := \{x - y : x, y \in A, x \neq y\}.$$

A *conflict-avoiding code* (CAC) of length n and weight w is a collection of subsets in $\mathcal{P}_w^n$, satisfying the property that

$$d^*(A) \cap d^*(B) = \emptyset$$

This work was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region under Project 416906.

Dept. of Information Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
E-mail: kshum2010@gmail.com, wswong@ie.cuhk.edu.hk.

for any $A \neq B$ in this collection of subsets. Each subset in a CAC is also called a *codeword*, and n is also called the *code length*. Given n and w, we use the notation CAC(n, w) for a CAC of length n and weight w

We note that for each codeword A, we can always find an integer b such that the translation $\{x + b \in \mathbb{Z}_n : x \in A\}$ contains the zero element in $\mathbb{Z}_n$. As the set of differences $d^*(A)$ is invariant under translation, we assume without loss of generality that $A \ni 0$ for every codeword A.

A codeword A represents a binary sequence of length n, denoted by s_A, with the t-th component equal to 1 if and only if $t \in A$, i.e.,

$$s_A(t) := \begin{cases} 1 & \text{if } t \in A \\ 0 & \text{if } t \notin A \end{cases}$$

for $t = 0, 1, \ldots, n-1$. In other words, A is the characteristic set of s_A. The definition of CAC is equivalent to the requirement that, for two distinct codewords A and B, the Hamming crosscorrelation between the corresponding sequences s_A and s_B, defined as

$$H_{s_A s_B}(\tau) := \sum_{t=0}^{n-1} s_A(t) s_B(t + \tau),$$

is either 0 or 1 no matter what the relative delay offset τ is. Here, the addition $+$ is done modulo n. Indeed, if $H_{s_A s_B}(\tau) \geq 2$ for some relative delay offset τ, then there are two distinct time indices t_0 and t_1, such that

$$s_A(t_0) = s_B(t_0 + \tau) = 1 = s_A(t_1) = s_B(t_1 + \tau),$$

which implies t_0 and t_1 belong to A and $t_0 + \tau$ and $t_1 + \tau$ belong to B, and hence $0 \neq t_1 - t_0 \in d^*(A) \cap d^*(B)$, contradicting the defining property of CAC. The *Hamming weight* of a binary sequence s of length n is defined as $\sum_{t=0}^{n-1} s(t)$. By construction, the Hamming weight of $s_A(t)$ is w for each codeword A.

Conflict-avoiding codes find applications in the multiple-access collision channel without feedback, and is also called protocol sequences in this context [1–3, 13, 14, 22]. In multiple-access collision channel, there are M users who share a common transmission medium and want to send packets to a common destination node. We consider the synchronous model, in which time is divided into slots and all users are slot-synchronized, i.e., a packet sent from user must be within the duration of a time slot. If exactly one user transmits a packet in a time slot, while the others are silent, then the packet can be received successfully without error. However, if two or more users transmit in the same time slot, a collision occurs, and the collided packets are assumed unrecoverable.

Each user is assigned a protocol sequence, which is a binary sequence of period n. Suppose after a duration of being inactive, a user becomes active at time T. A user sends a packet in slot $T + i$ if the i-th component of the protocol sequence is 1, or keep silent if 0. Packets are transmitted continually by repeating the protocol sequence periodically, until there is no more data to send. At that time the user will become idle, and remain inactive for at least n time slots before he becomes active again. Because the users become active at different times, we have relative delay offsets among the protocol sequences. A set of M binary sequences is called an (M, k, n, w) *protocol sequence set* if each sequence in this set is of length n and Hamming weight w, such that in a period of n slots, each active user can successfully send at least one packet without suffering collision no matter what the relative delay offsets are, provided that there are at most k simultaneously active users. It is easy to see that in order to support k active users, each user must send at least k packets in a period. Otherwise, there is a combination of delay offsets such that all packets of a particular user are collided.

In this paper, we consider the (M, k, n, w) protocol sequence set with $w = k$, i.e., the number of active users is exactly equal to the Hamming weight of the protocol sequences. As we have mentioned in the previous paragraph, this is the smallest Hamming weight for k simultaneously active users. Given a CAC of length n and weight w, we construct a protocol sequence set by generating binary sequence s_A of length n for each codeword A. By the defining property of CAC, there is at most one collision between s_A and s_B for any delay offset τ. If there are no more than w active users, then each user suffers at most $w-1$ collisions in a period, and hence can send at least one packet successfully. The constructed protocol sequence set is thus an (M, w, n, w) protocol sequence set. Conversely, we can see that the characteristic sets of the sequences in a (M, w, n, w) sequence set form a CAC of length n and weight w.

A codeword A is called *equi-difference* if the elements in A form an arithmetic progression, i.e.,

$$A = \{a, a+\delta, a+2\delta, \dots, a+(w-1)\delta\}$$

for some a and δ in $\mathbb{Z}_n$. A CAC is said to be equi-difference if all codewords are equi-difference. We use the symbol $\mathrm{CAC}^e(n, w)$ for an equi-difference CAC.

Example 1 Let $n = 45$, $w = 5$. Consider the following equi-difference codewords $\{0, 1, 2, 3, 4\}$, $\{0, 5, 10, 15, 20\}$, $\{0, 9, 18, 27, 36\}$, $\{0, 19, 38, 12, 31\}$, $\{0, 28, 11, 39, 22\}$ and $\{0, 37, 29, 21, 13\}$. The sets of differences

$$\begin{aligned}
d^*(\{0,1,2,3,4\}) &= \{1,2,3,4,41,42,43,44\},\\
d^*(\{0,5,10,15,20\}) &= \{5,10,15,20,25,30,35,40\},\\
d^*(\{0,9,18,27,36\}) &= \{9,18,27,36\},\\
d^*(\{0,19,38,12,31\}) &= \{7,12,14,19,26,31,33,38\},\\
d^*(\{0,28,11,39,22\}) &= \{6,11,17,22,23,28,34,39\},\\
d^*(\{0,37,29,21,13\}) &= \{8,13,16,21,24,29,32,37\}.
\end{aligned}$$

are disjoint. We thus have an equi-difference CAC of length 45 and weight 5, consisting of six codewords.

More examples of CACs of weights 3, 4 and 5 are available online at [21].

In this paper, we are interested in the largest number of codewords in a CAC of length n and weight w. Denote the maximal number of codewords in the class of all CACs with length n and weight w by $M(n, w)$. A $\mathrm{CAC}(n, w)$ is said to be *optimal* if the number of codewords is equal to $M(n, w)$. We let $M^e(n, w)$ be the maximal number of codewords in the sub-class of equi-difference CACs of length n and weight w.

The main result in this paper is the following

Theorem 1 *For all $w \geq 2$, we have*

$$\limsup_{n\to\infty} \frac{M(n,w)}{n} = \limsup_{n\to\infty} \frac{M^e(n,w)}{n} = \frac{1}{2w-2}. \tag{1}$$

Remark: Recall that the limit superior of a sequence $(a_i)_{i=1}^{\infty}$ of real numbers is less than or equal to a constant c if and only if for each $\epsilon > 0$, a_i is less than $c + \epsilon$ for all but finitely many i. The limit superior of $(a_i)_{i=1}^{\infty}$ is larger than or equal to c if and only if for each $\epsilon > 0$, there are infinitely many a_i which are larger than $c - \epsilon$ [20].

Theorem 1 can be interpreted as follows. Given weight w and arbitrarily small real number $\epsilon > 0$, the number of codewords in a CAC(n, w), normalized by the code length n, is less than $(2w - 2)^{-1} + \epsilon$ for all sufficiently large n, i.e., there is an integer $N(\epsilon)$, such that

$$\frac{M(n,w)}{n} \leq \frac{1}{2w-2} + \epsilon, \quad \text{for all } n \geq N(\epsilon). \tag{2}$$

This provides an asymptotic upper bound on the number of codewords in CAC. Furthermore, this bound is *tight*, meaning that if $(2w - 2)^{-1}$ in (2) is replaced by any number strictly smaller than $(2w - 2)^{-1}$, then (2) no longer holds for all $\epsilon > 0$.

Since obviously $M(n, w) \geq M^e(n, w)$, we have

$$\limsup_{n\to\infty} \frac{M(n,w)}{n} \geq \limsup_{n\to\infty} \frac{M^e(n,w)}{n}.$$

We thus divide the proof of Theorem 1 into two parts:

1. $\limsup_{n\to\infty} \frac{M(n,w)}{n} \leq \frac{1}{2w-2}$, and
2. $\limsup_{n\to\infty} \frac{M^e(n,w)}{n} \geq \frac{1}{2w-2}$.

The first part is proved in Section 3 (Prop. 2) by establishing a general upper bound on the size of CAC, which may or may not be equi-difference. The second part is proved in Section 4 (Prop. 3), by exhibiting, for each w, infinitely many equi-difference CACs, with size larger than $(2w - 2)^{-1} - \epsilon$ times the code length. We note that in the second part we do *not* need to show that $M^e(n, w)/n \geq (2w - 2)^{-1} - \epsilon$ for *all* sufficiently large n. We only need to show that there are infinitely many such n.

This paper is organized as follows. In Section 2, we review some existing bounds on the number of codewords for $w = 3, 4, 5$. Two constructions of CAC in the literatures are also presented. A new upper bound on the size of CAC is derived in Section 3. In Section 4, we construct an explicit sequence of CACs which achieve this bound asymptotically.

2 Results Previously Reported in the Literatures

2.1 Existing Bounds on Size of CAC

A lot of works have been done for the case $w = 3$. In [10], Levenshtein derived the upper bound $M(n, 3) \leq (n + 1)/4$, especially for $n \equiv 2 \bmod 4$,

$$M(n,3) = (n-2)/4,$$

and for the sequence of all odd integers n,

$$M(n,3) \sim n/4 \tag{3}$$

as $n \to \infty$. When n is a multiple of 4, Jimbo *et al.* [7] showed that

$$\frac{n}{6} + O(\log_4(n/4)) \leq M^e(n,3) \leq \frac{3}{16}n + \delta,$$

where δ is a constant depending on the congruence of $(n/4) \bmod 12$. They also gave optimal constructions of $\mathrm{CAC}(n, 3)$ for $n \equiv 8 \bmod 16$. Some optimal constructions of $\mathrm{CAC}^e(p, 3)$ for prime p is considered in [16, 17]. For length divisible by 16, optimal constructions can be found in [15].

For $w = 4$ and 5, constructions and bounds for $\mathrm{CAC}(n, w)$ is studied in [18]. It can be derived from [18] that

$$\limsup_{n\to\infty} M(n, 4)/n = \frac{1}{6},$$

and

$$\limsup_{n\to\infty} M(n, 5)/n = \frac{1}{8}.$$

Relatively less result is known for $w \geq 6$. In this paper, we extend the above asymptotic results to $w \geq 6$.

2.2 Known Constructions

Let p be an odd prime, $n = p^r$ for some integer r larger than or equal to 2, and w be $(p + 1)/2$. For each integer c, $0 \leq c < p^r$, consider the p-ary representation $c = c_0 + c_1 p + \ldots c_{r-1} p^{r-1}$. Let S be the set of integer c, $0 \leq c < p^r$, whose first nonzero symbol in its p-ary representation is 1. Consider the collection of codewords in the form $\{0, c, 2c, \ldots, (w - 1)c\}$ with $c \in S$. Obviously S contains $(p^r - 1)/(p - 1)$ elements. The set of nonzero differences of a codeword A is of the form

$$d^*(A) = \{\pm jc \bmod p^r : j = 0, 1, \ldots, (p - 1)/2\}.$$

It can be shown that the equality $\pm jc = \pm jc' \bmod p^r$, for $j, j' = 0, 1, \ldots, (p-1)/2$ and $c, c' \in S$, $c \neq c'$ holds only when j and j' are both zero. This implies that it is a CAC with $(p^r - 1)/(p - 1)$ codewords of weight w.

Theorem 2 ([9–11])

$$M(p^r, (p + 1)/2) \geq \frac{p^r - 1}{p - 1} = \frac{n - 1}{2w - 2}.$$

Example 2 Let $p = 5$, $r = 2$. We have $n = 25$ and $w = 3$. The following 6 codewords

$$\{0, 1, 2\}, \{0, 5, 10\}, \{0, 6, 12\}, \{0, 11, 22\}, \{0, 16, 7\}, \{0, 21, 17\}$$

form a CAC of length 25 and weight 3.

The CAC in Theorem 2 is conjectured to be optimal by Levenshtein in [10]. We will prove later that Levenshtein's conjecture is true.

The next construction uses some result about quadratic residues [6]. Given an odd prime p, a nonzero element $a \in \mathbb{Z}_p$ is called a *quadratic residue* if we can find an element $x \in \mathbb{Z}_p$ such that $a = x^2 \bmod p$, otherwise, a is called a *quadratic non-residue*. The quadratic residues under multiplication form a subgroup of index 2 within the multiplicative group of nonzero elements. Hence there are precisely $(p-1)/2$ quadratic residues and $(p-1)/2$ non-residues mod p. The Legendre symbol on $\mathbb{Z}_p$ is defined as

$$\left(\frac{a}{p}\right) := \begin{cases} 0 & \text{if } a = 0 \bmod p, \\ 1 & \text{if } a \neq 0 \text{ is a quadratic residue mod } p, \\ -1 & \text{if } a \neq 0 \text{ is a quadratic non-residue mod } p. \end{cases}$$

It can be shown that the Legendre symbol is multiplicative, i.e.,

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{4}$$

The following construction is an adaptation from a construction in [18, Theorem 3.7].

Theorem 3 ([18]) *Let p be an odd prime and w be an integer such that $2 \leq w \leq p$. If -1 is a quadratic non-residue mod p and*

$$\left(\frac{i}{p}\right)\left(\frac{i-w+1}{p}\right) = -1 \tag{5}$$

for $i = 1, \dots, w-2$, then there exists a conflict-avoiding code consisting of $(p-1)/2$ codewords, with length $(w-1)p$ and weight w.

We include the proof for completeness.

Proof Since $w-1 < p$, $w-1$ and p are relatively prime. The cyclic group $\mathbb{Z}_{(w-1)p}$ can be identified with $\mathbb{Z}_{w-1} \oplus \mathbb{Z}_p$ via the Chinese remainder theorem [6]. We will construct a CAC on $\mathbb{Z}_{w-1} \oplus \mathbb{Z}_p$.

Let $\mathcal{Q}$ be the set of quadratic residues mod p. For each g in $\mathcal{Q}$, we define a codeword

$$A_g := \{(0,0),\ (1,g),\ (2,2g),\ \dots, (w-2,(w-2)g),\ (0,(w-1)g)\}.$$

This is an arithmetic progression in $\mathbb{Z}_{w-1} \oplus \mathbb{Z}_p$ with common difference $(1,g)$. The set of differences $d^*(A_g)$ can be written as

$$\{(i,ig) : i = 1,\dots,w-2\} \cup \{(i,(i-(w-1))g) : i = 1,\dots,w-2\} \cup \{(0,\pm(w-1)g)\}.$$

Suppose for the sake of contradiction that $d^*(A_g) \cap d^*(A_h)$ is non-empty for some g and $h \in \mathcal{Q}$, with $g \neq h$. Let the common element be (i,j). We consider two cases.

Case 1, $i = 0$. We have $(w-1)g = \pm(w-1)h$, and this implies that $g = \pm h$. Since $g \neq h$, we get $g = -h$, and consequently $-1 = gh^{-1}$. However, g and h are both quadratic residues by definition. This contradicts the assumption that -1 is a quadratic non-residue.

Case 2, $i = 1, 2, \dots, w-2$. There are four possibilities: (i) $j = ig = ih$, (ii) $j = ig = (i-(w-1))h$, (iii) $j = (i-(w-1))g = ih$, and (iv) $j = (i-(w-1))g = (i-(w-1))h$. In case (i) and (iv), we have $g = h$, which is false by assumption. If (ii) holds, then by taking the Legendre symbol on both side of $ig = (i-(w-1))h$, and using the multiplicative property in (4), we get $\left(\frac{i}{p}\right) = \left(\frac{i-(w-1)}{p}\right)$, which contradicts (5). Similar argument shows that (iii) cannot hold as well.

This concludes that $d^*(A_g)$ and $d^*(A_h)$ must be disjoint, and hence the $(p-1)/2$ codewords form a CAC on $\mathbb{Z}_{w-1} \oplus \mathbb{Z}_p$. Via the isomorphism between $\mathbb{Z}_{w-1} \oplus \mathbb{Z}_p$ and $\mathbb{Z}_{(w-1)p}$, we obtain a CAC on $\mathbb{Z}_{(w-1)p}$ with $(p-1)/2$ codewords of weight w. □

Example 3 Let $p = 19$ and $w = 6$. The quadratic residues mod 19 are 1, 4, 5, 6, 7, 9, 11, 16 and 17. We can check that

$$\left(\frac{-1}{p}\right), \left(\frac{1}{p}\right)\left(\frac{-4}{p}\right), \left(\frac{2}{p}\right)\left(\frac{-3}{p}\right), \left(\frac{3}{p}\right)\left(\frac{-2}{p}\right), \left(\frac{4}{p}\right)\left(\frac{-1}{p}\right)$$

are all equal to -1. The CAC obtained by the construction in Theorem 3 is of length 95, consisting of the following 9 codewords

$$\begin{array}{lll} \{0,1,2,3,4,5\}, & \{0,61,27,88,54,20\}, & \{0,81,67,53,39,25\}, \\ \{0,6,12,18,24,30\}, & \{0,26,52,78,9,35\}, & \{0,66,37,8,74,45\}, \\ \{0,11,22,33,44,55\}, & \{0,16,32,48,64,80\}, & \{0,36,72,13,49,85\}. \end{array}$$

We note the CACs in Theorem 2 and 3 are equi-difference.

3 Upper Bound on Number of Codewords

An upper bound on the size of CAC is derived in this section. The main idea of is that, despite some exceptional cases, each codeword provably contributes at least $2w-2$ distinct differences. If we can show that the number of exceptional codewords is very small, then, as the differences cannot overlap, the number of codewords is roughly speaking no larger than $n/(2w-2)$.

To formulate the upper bound, we need some result in additive number theory. We first introduce some notions for abelian group in general. Let G be an abelian group. For two subsets $A, B \subset G$, the *sum set* of A and B is defined as

$$A + B := \{a + b : a \in A, b \in B\},$$

while the *difference set* is

$$A - B := \{a - b : a \in A, b \in B\}.$$

The self difference set $A - A$ is denoted by

$$d(A) := \{a_1 - a_2 : a_1, a_2 \in A\}.$$

The set $d^*(A)$ defined in the introduction is equal to $d(A) \setminus \{0\}$.

The following simple observation will be useful.

Proposition 1 *$|d(A)| \geq |A|$ for any subset A in G.*

Proof Suppose that A contains m elements, labeled by a_1, $a_2, \ldots, a_m$. The $m-1$ difference $a_i - a_1$, for $i = 2, 3, \ldots, m$ are distinct and non-zero. Together with the zero element in G, we already have m distinct elements in $d(A)$. Hence, $|d(A)| \geq m$. □

A nonempty subset A of G is called *H-periodic*, where H is a subgroup of G, if it is a union of cosets of H. In terms of set addition, it is equivalent to saying that $A = A + H$. A subset is called *periodic* if it is H-periodic for some non-trivial subgroup H. H is said to be a period of A if A is H-periodic. (The notion of "periodic" and "H-periodic" should not be confused with the period of protocol sequences.) The following theorem due to Kneser [8] is the key to the derivation of the upper bound on the size of CAC.

Theorem 4 (Kneser [8]) *Let S and T be finite nonempty subsets of a finite abelian group G. If $|S+T| < |S|+|T|-1$, then $S+T$ is periodic for some nontrivial subgroup H of G.*

Proofs of Kneser's theorem can be found in [12] and [19].

By applying Theorem 4 with $T = \{-s : s \in S\}$, we have the following corollary.

Corollary 1 *Let S be a non-empty subset in a finite abelian group. We have $|d^*(S)| \geq 2|S|-2$ unless $d(S)$ is periodic. Moreover, if H is a period of $d(S)$, then $d(S)$ contains H as a subset.*

Proof The first statement follows directly from Theorem 4, because, if $|d^*(S)| < 2|S|-2$, then

$$|S-S| = |d(S)| = |d^*(S)|+1 \leq 2|S|-2 < |S|+|S|-1.$$

Hence, $d(S)$ must be periodic for some nontrivial subgroup H of $\mathbb{Z}_n$. Since $0 \in d(S)$, we have $H \subseteq d(S)$. □

Example 4 (Continue from Example 1) The codeword $A = \{0, 9, 18, 27, 36\}$ in $\mathbb{Z}_{45}$ has $d(A) = \{0, 9, 18, 27, 36\}$ which is periodic. In fact, $d(A)$ itself is a subgroup of $\mathbb{Z}_{45}$. For all other codeword B, $d(B)$ is not periodic and we can check that $d^*(B) = 2|B|-2 = 8$.

Example 5 Consider the codeword $A = \{0, 1, 4, 5, 8, 9\}$ in $\mathbb{Z}_{12}$. We check that

$$d^*(A) = \{1, 3, 4, 5, 7, 8, 9, 11\}$$

and $|d^*(A)| = 8 < 2|A|-2$. By Corollary 1, $d(A)$ must be periodic. Indeed, $d(A)$ is the union of subgroup $\{0, 4, 8\}$ of $\mathbb{Z}_{12}$ and two cosets $\{1, 5, 9\}$ and $\{3, 7, 11\}$.

From Kneser's theorem, we have immediately the following upper bound on number of codewords when the length is prime, or when the length is a product of large prime factors.

Theorem 5 *Suppose that the length of a CAC of weight w satisfies one of the following conditions:*
(i) n is prime,
(ii) the prime factors of n are all larger than or equal to $2w-1$.
Then

$$M(n, w) \leq \left\lfloor \frac{n-1}{2w-2} \right\rfloor.$$

Proof (i) When n is prime, $\mathbb{Z}_n$ has no nontrivial subgroup. Therefore, there is no codeword A such that $d(A)$ is periodic. By Corollary 1, we have $|d^*(A)| \geq 2w-2$ for all codeword A. Since $d^*(A)$ and $d^*(B)$ are disjoint for any pair of distinct codewords A and B, the number of codewords is no more than $(n-1)/(2w-2)$.

(ii) Suppose that $|d^*(A)| < 2w-2$ for some codeword A. Then $d(A)$ is periodic and contains a nontrivial subgroup of $\mathbb{Z}_n$ by Corollary 1. However, since the smallest divisor of n is at least $2w-1$, the smallest nontrivial subgroup of $\mathbb{Z}_n$ has cardinality at least $2w-1$. Therefore,

$$|d^*(A)| = |d(A)|-1 \geq (2w-1)-1 = 2w-2.$$

It contradicts the assumption that $d^*(A)$ has cardinality strictly less than $2w-2$. Thus, we have $d^*(A) \geq 2w-2$ for every codeword A. The proof continues as in part (i). □

We now show that the two constructions described in Section 2 are optimal.

Theorem 6 *Let p be an odd prime, $n = p^r$ for some integer $r \geq 2$ and $w = (p+1)/2$. We have*

$$M(p^r, (p+1)/2) = \frac{n-1}{2w-2}.$$

Proof As p is the only prime factor of $n = p^r$ and $p = 2w - 1$ by definition, the results follows from Theorem 5 (ii). □

This proves Levenshtein's conjecture that the construction in Theorem 2 is optimal [10].

Theorem 7 *Let p and $w-1$ are distinct odd primes such that $p > 2w-2$. Then there are at most $(p-1)/2$ codewords in $CAC((w-1)p, w)$.*

Proof Let $n = (w-1)p$. Note that $\mathbb{Z}_n$ has precisely two nontrivial subgroups. Let H_1 be the subgroup which consists of all multiples of p, and H_2 be the subgroup which consists of all multiples of $w-1$. H_1 and H_2 contain $w-1$ and p elements, respectively. If there is a codeword A with the property that $d(A)$ is H_2-periodic, then we have $|d^*(A)| = |d(A)| - 1 \geq 2w - 2$ because $d(A)$ contains the subgroup H_2 of cardinality p, which is strictly larger than $2w-2$ by assumption. This yields a contradiction to Corollary 1. Therefore, for any codeword A in a CAC$((w-1)p, w)$, $d(A)$ is either aperiodic or H_1-periodic.

We consider two cases.

Case 1, all codewords are aperiodic. We have $|d^*(A)| \geq 2w-2$ for all codeword A. The number of codewords is no more than

$$\left\lfloor \frac{n-1}{2w-2} \right\rfloor = \left\lfloor \frac{(w-1)p-1}{2w-2} \right\rfloor = \left\lfloor \frac{p}{2} - \frac{1}{2w-2} \right\rfloor \leq \frac{p-1}{2}.$$

Case 2, there is at least one codeword that is H_1-periodic. In this case, we cannot have two distinct codewords A and B that are both H_1-periodic, otherwise $d(A)$ and $d(B)$ both contain H_1 as subset and it contradicts the requirement that $d^*(A)$ and $d^*(B)$ are disjoint. So there is a unique codeword, say $\tilde{B}$, that is H_1-periodic. $\tilde{B}$ must contain at least two cosets of H_1, because $|d(\tilde{B})| \geq w$ by Prop. 1. Therefore $|d(\tilde{B})| \geq 2(w-1)$. The number of codewords is no more than

$$1 + \frac{n - 2(w-1)}{2(w-1)}.$$

After substituting n by $(w-1)p$, we can simplify the above expression to $p/2$. The number of codewords is thus less than or equal to $\lfloor p/2 \rfloor = (p-1)/2$. □

This shows that the construction in Theorem 3 is optimal when $w-1$ is prime and $p > 2w-2$. Next, we obtain a general upper bound on $M(n, w)$ which holds for all w.

Theorem 8 *Let $\omega(n)$ denote the number of distinct prime divisors of n. For $n \geq 2$ and $w \geq 2$, we have*

$$M(n, w) \leq \frac{n-1}{2w-2} + \frac{\omega(n)}{2}.$$

Proof We divide the codewords into two types by checking whether the difference set of a codeword is periodic or not. If $d(A)$ is aperiodic, we have $d^*(A) \geq 2w-2$ by Corollary 1. For a periodic $d(B)$, we have $d^*(B) \geq w-1$. Since the totality of all distinct differences cannot be larger than the number of nonzero element in $\mathbb{Z}_n$, we obtain

$$n-1 \geq \sum_{A\colon d(A) \text{ aperiodic}} |d^*(A)| + \sum_{B\colon d(B) \text{ periodic}} |d^*(B)|$$

where the first summation is over codeword A such that $d(A)$ is aperiodic and the second summation is over codeword B such that $d(B)$ is periodic.

By Prop. 1 and Corollary 1, we have

$$n-1 \geq \sum_{A\colon d(A) \text{ aperiodic}} (2w-2) + \sum_{B\colon d(B) \text{ periodic}} (w-1). \tag{6}$$

Suppose that there are M_p codewords B with $d(B)$ periodic. It follows from (6) that

$$n-1 \geq (M-M_p)(2w-2) + M_p(w-1),$$

or equivalently

$$M \leq \frac{n-1}{2w-2} + \frac{M_p}{2}. \tag{7}$$

It remains to show that M_p is no larger than $\omega(n)$. Suppose that B_1 and B_2 are two codewords such that $d(B_1)$ and $d(B_2)$ are periodic. Suppose that $d(B_1)$ is H_1-periodic and $d(B_2)$ is H_2-periodic, for two nontrivial subgroups H_1 and H_2 of $\mathbb{Z}_n$. By Corollary 1, $d(B_i)$ contains H_i, for $i=1,2$. Suppose that H_1 and H_2 are generated by two proper divisors of n, α_1 and α_2, respectively, i.e., H_i consists of the multiples of α_i for $i=1,2$. We claim that n/α_1 and n/α_2 must be relatively prime. Otherwise, if the greatest common divisor of n/α_1 and n/α_2, denoted by g, is greater than 1, then for $i=1,2$, we have $n/\alpha_i = gx_i$ for some integer x_i, and hence n/g is a multiple of both α_1 and α_2. This implies that n/g is contained in both $d^*(B_1)$ and $d^*(B_2)$, contradicting the assumption that B_1 and B_2 are codewords of a CAC.

Suppose that $B_1, B_2, \ldots, B_{M_p}$ are the codewords with $d(B_i)$ being periodic for $i=1,2,\ldots,M_p$. We can find M_p proper divisors of n, say $\alpha_1,\ldots,\alpha_{M_p}$ such that $d(B_i)$ contains the additive subgroup of $\mathbb{Z}_n$ generated by α_i, for $i=1,\ldots,M_p$. By the argument in the previous paragraph, n/α_i, for $i=1,2,\ldots,M_p$ are mutually relatively prime. On the other hand, the number of divisors of n which are mutually relatively prime is less than or equal to $\omega(n)$. Therefore $M_p \leq \omega(n)$. By replacing M_p in (7) by $\omega(n)$, we obtain the upper bound in the theorem. □

The function $\omega(n)$ grows very slowly in n. It was shown by Hardy and Ramanujan that $\omega(n)$ is close to $\log\log(n)$ [4, p.51]. For our purpose, it is sufficient to use the fact that $\omega(n)$ is upper bounded by $\log_2(n)$. Indeed, if n is factorized into $p_1^{e_1}\cdots p_{\omega(n)}^{e_{\omega(n)}}$, where $p_1,p_2,\ldots,p_{\omega(n)}$ are distinct primes, then

$$n \geq p_1p_2\cdots p_{\omega(n)} \geq 2^{\omega(n)}.$$

We are now ready to prove half of Theorem 1.

Proposition 2 *For $w \geq 2$,*

$$\limsup_{n\to\infty} \frac{M(n,w)}{n} \leq \frac{1}{2w-2}.$$

Proof We replace $\omega(n)$ by $\log_2(n)$ in Theorem 8 and divide by n,

$$\frac{M(n,w)}{n} \leq \frac{n-1}{n(2w-2)} + \frac{\log_2(n)}{2n}.$$

The result follows from taking $\limsup$ on both sides. □

4 Tightness of the Asymptotic Upper Bound

For weight w such that $2w-1$ is a prime, we have for all integer $r \geq 1$

$$\frac{M^e(p^r,w)}{p^r} \geq \frac{1}{p^r} \cdot \frac{p^r-1}{2w-2}$$

by Theorem 2, where $p = 2w-1$. Hence, for each $\epsilon > 0$, the ratio $M^e(n,w)/n$ is larger than $(2w-2)^{-1} - \epsilon$ for infinitely many n. Combining this result with Prop. 2, we get

$$\limsup_{n\to\infty} \frac{M^e(n,w)}{n} = \frac{1}{2w-2},$$

for all w such that $2w-1$ is prime.

In this section, we show that the above equality holds for all $w \geq 2$, by showing that for each w, there exists an infinite sequence of $\text{CAC}^e(n,w)$ attaining the upper bound in Prop. 2. This will complete the proof of Theorem 1

We first consider the case $w = 6$. Let $p = 2m+1$ be a prime. The conditions in Theorem 3 hold if p satisfies either one of the following set of conditions:

$$\begin{cases} \left(\frac{-1}{p}\right) = -1 \\ \left(\frac{2}{p}\right) = \left(\frac{3}{p}\right) = 1 \end{cases} \qquad \begin{cases} \left(\frac{-1}{p}\right) = -1 \\ \left(\frac{2}{p}\right) = \left(\frac{3}{p}\right) = -1. \end{cases}$$

By quadratic reciprocity, the primes that satisfy the above conditions are congruent to 19 or 23 mod 24, namely 19, 23, 43, 47, 67, 71, 91, and so on. They give rise to $\text{CAC}^e(95,6)$, $\text{CAC}^e(115,6)$, $\text{CAC}^e(215,6)$, $\text{CAC}^e(235,6)$, $\text{CAC}^e(335,6)$, $\text{CAC}^e(355,6)$, $\text{CAC}^e(455,6)$, $\ldots$, containing 9, 11, 21, 23, 33, 35, 45,... codewords respectively. We have shown in Theorem 7 that they are optimal. From Dirichlet's theorem on primes in arithmetic progression [6], we know that there are infinitely may prime p which are congruent to either 19 or 23 mod 24. With p going over all such primes, we get

$$\limsup_{p\to\infty} \frac{M^e(5p,6)}{5p} \geq \lim_{p\to\infty} \frac{(p-1)/2}{5p} = \frac{1}{10}.$$

For general w, we have the following

Proposition 3 *Let $w \geq 2$ be a fixed integer. For any arbitrarily small real number $\epsilon > 0$, there are infinitely many integer n which can be written in the form $(w-1)p$ for some prime number p, such that*

$$\frac{M^e(n,w)}{n} \geq \frac{1}{2w-2} - \epsilon,$$

i.e.,

$$\limsup_{n\to\infty} \frac{M^e(n,w)}{n} \geq \frac{1}{2w-2}.$$

Proof If we can find a prime p such that

$$\left(\frac{a}{p}\right) = 1 \qquad \text{for } a = 1,2,\dots,w-2, \text{ and} \tag{8}$$

$$\left(\frac{-1}{p}\right) = -1, \tag{9}$$

then for all $i = 1,2,\dots,w-2$, we have

$$\begin{aligned}\left(\frac{i}{p}\right)\left(\frac{i-w+1}{p}\right) &= \left(\frac{i}{p}\right)\left(\frac{(-1)(w-1-i)}{p}\right)\\ &= \left(\frac{i}{p}\right)\left(\frac{-1}{p}\right)\left(\frac{w-1-i}{p}\right) = -1.\end{aligned}$$

So, the conditions in Theorem 3 are satisfied, and we can construct a $\mathrm{CAC}^e((w-1)p, w)$ with $(p-1)/2$ codewords. Suppose that there exists infinitely many such prime p, we then have a sequence of $\mathrm{CAC}^e((w-1)p, w)$ such that the ratio of number of codewords to code length is

$$\frac{(p-1)/2}{(w-1)p} = \frac{p-1}{p} \cdot \frac{1}{2w-2}.$$

This ratio approaches $1/(2w-2)$ as p approaches infinity. Hence, given an arbitrarily small $\epsilon > 0$, there are infinitely many prime p such that

$$\frac{p-1}{p} \cdot \frac{1}{2w-2} \geq \frac{1}{2w-2} - \epsilon.$$

This implies that

$$\frac{M^e((w-1)p, w)}{(w-1)p} \geq \frac{(p-1)/2}{(w-1)p} \geq \frac{1}{2w-2} - \epsilon$$

for infinitely many prime p.

Therefore it suffices to show that there are infinity many primes which satisfy the conditions in (8) and (9). Let $\mathcal{S} = \{p_1, p_2, \dots, p_r\}$ be the set of all the distinct prime numbers less than or equal to $w-2$. For each i between 2 and $w-2$, let $\mathcal{S}_i \subseteq \mathcal{S}$ denote the set of prime factors that appear in the factorization of i with odd multiplicity. We can thereby express i as

$$i = y_i^2 \prod_{q \in \mathcal{S}_i} q$$

for some integer y_i. After applying the Legendre symbol to both sides of the above equation, we obtain

$$\left(\frac{i}{p}\right) = \prod_{q \in \mathcal{S}_i} \left(\frac{q}{p}\right).$$

This shows that if we can find a prime p such that $p_1, p_2, \ldots, p_r$ are all quadratic residues of p, then $\left(\frac{i}{p}\right) = 1$ for all $i = 2, 3, \ldots, w-2$. Our task reduces to searching for prime p which satisfies

$$\left(\frac{p_1}{p}\right) = \left(\frac{p_2}{p}\right) = \ldots = \left(\frac{p_r}{p}\right) = 1$$

and $\left(\frac{-1}{p}\right) = -1$. The infinitude of such primes is established by appealing to the following theorem, which roughly says that there are infinitely many primes with prescribed quadratic residues and non-residues. □

Theorem 9 ([5]) *Let $a_1, a_2, \ldots, a_m$ be integers such that the product of powers*

$$a_1^{u_1} a_2^{u_2} \cdots a_m^{u_m}$$

is a square only if all u_i are even. Then there are infinity many primes p which satisfy $\left(\frac{a_i}{p}\right) = c_i$, for $i = 1, 2, \ldots, m$, with c_i taken arbitrarily in $\{1, -1\}$.

A proof of Theorem 9 can be found in [5, §49]. The proof of Prop. 3 is completed by choosing $a_i = p_i$ and $c_i = 1$ for $i = 1, \ldots, r$, and $a_{r+1} = c_{r+1} = -1$.

Example 6 Consider the case $w = 11$. By Theorem 9, there are infinitely many primes such that $(2/p) = (3/p) = (5/p) = (7/p) = 1$ and $(-1/p) = -1$. Any such prime p corresponds to a $\mathrm{CAC}^e(10p, 11)$ with $(p-1)/2$ codewords. The primes that satisfy these requirements are 311, 479, 719, 839, 1151, 1319, 1511, 1559, 2351, 2399 and so on. If we take the limit over this sequence of primes, we have

$$\lim_{p \to \infty} \frac{(p-1)/2}{10p} = \frac{1}{20}.$$

Therefore, for each arbitrarily small $\epsilon > 0$, there are infinitely many prime p such that

$$\frac{M^e(10p, 11)}{10p} \geq \frac{(p-1)/2}{10p} \geq \frac{1}{20} - \epsilon.$$

This verifies Prop. 3 when $w = 11$.

5 Conclusion

We obtain an asymptotic upper bound on the size of CAC, which holds for all weights in general, and thus extend previously known upper bounds on the size of CAC. By constructing asymptotically optimal sequences of CAC with increasing length, we show that this asymptotic upper bound is tight. By the result in this paper, some existing constructions of CAC are proved to be optimal as well.

References

1. A., N.Q., Györfi, L., Massey, J.L.: Constructions of binary constant-weight cyclic codes and cyclically permutable codes. IEEE Trans. Inf. Theory **38**(3), 940–949 (1992)
2. da Rocha Jr., V.C.: Protocol sequences for collision channel without feedback. IEE Electron. Lett. **36**(24), 2010–2012 (2000)
3. Györfi, L., Vajda, I.: Construction of protocol sequences for multiple-access collision channel without feedback. IEEE Trans. Inf. Theory **39**(5), 1762–1765 (1993)
4. Hardy, G.H.: Ramanujan: Twelve Lectures on Subjects Suggested by His Life and Work, 3rd edn. Chelsea, New York (1999)
5. Hecke, E.: Lectures on the Theory of Algebraic Numbers, *Graduate Texts in Math.*, vol. 77. Springer-Verlag, New York (1981)
6. Ireland, K., Rosen, M.: A Classical Introduction to Modern Number Theory. Springer-Verlag, Yew York (1990)
7. Jimbo, M., Mishima, M., Janiszewski, S., Teymorian, A.Y., Tonchev, V.D.: On conflict-avoiding codes of length $n = 4m$ for three active users. IEEE Trans. Inf. Theory **53**(8), 2732–2742 (2007)
8. Kneser, M.: Abschätzungen der asymptotischen dichte von summenmengen. Math. Zeit. **58**, 459–484 (1953)
9. Levenshtein, V.I.: Conflict-avoiding codes with multiple active users. In: Proc. 14th Int. Conf. on Problems of Theoretical Cybernetics, p. 86. Moscow (2005)
10. Levenshtein, V.I.: Conflict-avoiding codes for three active users and cyclic triple systems. Problems of Information Transmission **43**(3), 199–212 (2007)
11. Levenshtein, V.I., Han Vinck, A.J.: Perfect (d, k)-codes capable of correcting single peak-shift. IEEE Trans. Inf. Theory **39**(2), 656–662 (1993)
12. Mann, H.B.: Addition Theorems: the Addition Theorems of Group Theory and Number Theory. No. 18 in Interscience Tracks in Pure and Applied Mathematics. Interscience Publisher, New York (1965)
13. Massey, J.L., Mathys, P.: The collision channel without feedback. IEEE Trans. Inf. Theory **31**(2), 192–204 (1985)
14. Mathys, P.: A class of codes for a T-active-users-out-of-N multiple-access communication system. IEEE Trans. Inf. Theory **36**(6), 1206–1219 (1990)
15. Mishima, M., Fu, H.L., Uruno, S.: Optimal conflict-avoiding codes of length $n \equiv 0 \pmod{16}$ and weight 3. Designs, Codes and Cryptography **52**(3), 275–291 (2009)
16. Momihara, K.: Necessary and sufficient conditions for tight equi-difference conflict-avoiding codes of weight three. Designs, Codes and Cryptography **45**, 379–390 (2007)
17. Momihara, K.: On cyclic $2(k-1)$-support $(n, k)_{k-1}$ difference families. Finite Fields and Their Applications **15**, 415–427 (2009)
18. Momihara, K., Müller, M., Satoh, J., Jimbo, M.: Constant weight conflict-avoiding codes. SIAM J. Discrete Math. **21**(4), 959–979 (2007)
19. Nathanson, M.B.: Additive Number Theory – Inverse Problems and Geometry of Sumsets. No. 165 in Graduate Texts in Mathematics. Springer-Verlag, New York (1996)
20. Rudin, W.: Principles of Mathematical Analysis. McGraw Hill (1976)
21. Tonchev, V.D.: Tables of conflict-avoiding codes (2005). Avaiable online at http://www.math.mtu.edu/˜tonchev/CAC.html
22. Wong, W.S.: New protocol sequences for random access channels without feedback. IEEE Trans. Inf. Theory **53**(6), 2060–2071 (2007)