

Linear Size Constant-Composition Codes

Yeow Meng Chee and Xiande Zhang

School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
email: {ymchee, xiandezhang}@ntu.edu.sg

Abstract—The Johnson-type upper bound on the maximum size of a code of length n , distance $d = 2w - 1$ and constant composition $\bar{w}$ is $\lfloor \frac{n}{w_1} \rfloor$, where w is the total weight and w_1 is the largest component of $\bar{w}$. Recently, Chee *et al.* proved that this upper bound can be achieved for all constant-composition codes of sufficiently large lengths. Let $N_{ccc}(\bar{w})$ be the smallest such length. The determination of $N_{ccc}(\bar{w})$ is trivial for binary codes. This paper provides a lower bound on $N_{ccc}(\bar{w})$, which is shown to be tight for all ternary and quaternary codes by giving several new combinatorial constructions. Consequently, by refining method, we determine the values of $N_{ccc}(\bar{w})$ for all q -ary constant-composition codes such that $3w_1 \geq w$ with a few possible exceptions.

I. INTRODUCTION

Constant-composition codes have attracted a lot attention [1]–[21] in recent years due to their vast applications, such as in determining the zero error decision feedback capacity of discrete memoryless channels [22], [23], multiple-access communications [24], [25], spherical codes for modulation [26], DNA codes [27], [28], powerline communications [29], [30], and frequency hopping [4].

Although constant-composition codes have been used since the early 1980s to bound error and erasure probabilities in decision feedback channels [31], their systematic study only began in late 1990s with Svanström [32]. Nowadays, the problem of determining the maximum size of a constant-composition code constitutes a central problem in their study due to their close relations to combinatorial design theory [2]–[5], [7], [9], [10], [14]–[17], [19], [20].

For integers $m \leq n$, the set of integers $\{m, m+1, \dots, n\}$ is denoted by $I_{m,n}$. When $m = 1$, the set $I_{1,n}$ is further abbreviated to I_n . If $m > n$, then $I_{m,n}$ is defined to be empty. The ring $\mathbb{Z}/n\mathbb{Z}$ is denoted by $\mathbb{Z}_n$. For finite sets R and X , R^X denotes the set of vectors of length $|X|$, where each component of a vector $\mathbf{u} \in R^X$ has value in R and is indexed by an element of X , that is, $\mathbf{u} = (u_x)_{x \in X}$, and $u_x \in R$ for each $x \in X$.

A q -ary code of length n is a set $\mathcal{C} \subseteq \mathbb{Z}_q^X$, for some X of size n . The elements of $\mathcal{C}$ are called *codewords*. The *support* of a vector $\mathbf{u} \in \mathbb{Z}_q^X$ is $\text{supp}(\mathbf{u}) = \{x \in X : u_x \neq 0\}$. The *Hamming weight* of a vector $\mathbf{u} \in \mathbb{Z}_q^X$ is defined as $\|\mathbf{u}\| = |\text{supp}(\mathbf{u})|$. The distance induced by this weight is the *Hamming distance*, denoted by $d_H(\cdot, \cdot)$, so that $d_H(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, for $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_q^X$.

A code $\mathcal{C}$ is said to have *distance* d if $d_H(\mathbf{u}, \mathbf{v}) \geq d$ for all distinct $\mathbf{u}, \mathbf{v} \in \mathcal{C}$. The *composition* of a vector $\mathbf{u} \in \mathbb{Z}_q^X$ is the tuple $\bar{w} = \llbracket w_1, \dots, w_{q-1} \rrbracket$, where $w_i = |\{x \in X : u_i = i\}|$, $i \in \mathbb{Z}_q \setminus \{0\}$. A code $\mathcal{C}$ is said to have *constant weight*

w if every codeword in $\mathcal{C}$ has weight w , and is said to have *constant composition* $\bar{w}$ if every codeword in $\mathcal{C}$ has composition $\bar{w}$. Hence, every constant-composition code is a constant-weight code. In this paper, attention is restricted to constant-composition codes. For constant-weight codes, interested readers are referred to [33].

A q -ary code of length n , distance d , and constant composition $\bar{w}$ is denoted an $(n, d, \bar{w})_q$ -code. The maximum size of an $(n, d, \bar{w})_q$ -code is denoted $A_q(n, d, \bar{w})$, and an $(n, d, \bar{w})_q$ -code attaining the maximum size is said to be *optimal*. In an $(n, d, \bar{w})_q$ -code, reordering the components of $\bar{w}$ or deleting zero components of $\bar{w}$ will not affect the distance and composition properties. Hence, through out this paper, when we talk about a composition $\bar{w} = \llbracket w_1, w_2, \dots, w_{q-1} \rrbracket$, we always assume that all components are positive and listed in non-increasing order, that is, $w_1 \geq w_2 \geq \dots \geq w_{q-1} \geq 1$. For succinctness, define the total weight $w := \sum_{i=1}^{q-1} w_i$.

The Johnson-type bound of Svanström for ternary constant-composition codes [1] could be easily extends to the following (see also [4]).

Proposition 1. (Johnson Bound):

$$A_q(n, d, \llbracket w_1, w_2, \dots, w_{q-1} \rrbracket) \leq \left\lfloor \frac{n}{w_1} A_q(n-1, d, \llbracket w_1-1, w_2, \dots, w_{q-1} \rrbracket) \right\rfloor.$$

Definition 1. A composition $\bar{w} = \llbracket w_1, \dots, w_q \rrbracket$ is a *refinement* of $\bar{v} = \llbracket v_1, \dots, v_{q'} \rrbracket$ if there exist pairwise disjoint sets $S_1, \dots, S_{q'} \subset I_q$ satisfying $\cup_{j \in I_{q'}} S_j = I_q$ such that $\sum_{i \in S_j} w_i = v_j$ for each $j \in I_{q'}$.

Chu *et al.* [4] made the following observation.

Lemma 1. If $\bar{w}$ is a refinement of $\bar{v}$, then $A_q(n, d, \bar{w}) \geq A_q(n, d, \bar{v})$.

In [2], Chee *et al.* show that $A_q(n, d, \bar{w}) = O(n)$ if and only if $d \geq 2w - 1$. For $d \geq 2w$, it is trivial to determine values of $A_q(n, d, \bar{w})$. For $d = 2w - 1$, the knowledge of $A_q(n, d, \bar{w})$ is limited. By Proposition 1,

$$A_q(n, 2w - 1, \bar{w}) \leq \left\lfloor \frac{n}{w_1} \right\rfloor$$

for all $\bar{w}$. When $q = 2$, we know that $A_2(n, 2w - 1, \llbracket w_1 \rrbracket) = \lfloor \frac{n}{w_1} \rfloor$, trivially. When $q = 3$, the values of $A_3(n, 2w - 1, \bar{w})$ has been completely determined by Svanström *et al.* [3]. Besides this, the following asymptotic statement was proved in [2].

Theorem 1 (Chee *et al.* [2]). Let $\bar{w} = \llbracket w_1, w_2, \dots, w_{q-1} \rrbracket$. Then $A_q(n, 2w-1, \bar{w}) = \lfloor \frac{n}{w_1} \rfloor$ for all sufficiently large n .

A. Problem Status and Contribution

In Theorem 1, the hypothesis that n is sufficiently large must be satisfied. But how large must n be? More precisely, for a composition $\bar{w} = \llbracket w_1, w_2, \dots, w_{q-1} \rrbracket$, define

$$N_{ccc}(\bar{w}) = \min\{n_0 \in \mathbb{N} : A_q(n, 2w-1, \bar{w}) = \lfloor \frac{n}{w_1} \rfloor \text{ for all } n \geq n_0\}.$$

For binary codes, it is trivial that $N_{ccc}(\llbracket w_1 \rrbracket) = 1$. Explicit bounds on $N_{ccc}(\bar{w})$ for general $\bar{w}$ were given in [2].

Proposition 2. For any composition $\bar{w}$, we have

$$w^2 - w_1(w-1) \leq N_{ccc}(\bar{w}) \leq 4w_1(w-1)^2 + 1.$$

The upper and lower bounds on $N_{ccc}(\bar{w})$ in Proposition 2 differ approximately by a factor of $4w_1$. Our interest in this paper is in determining the values of $N_{ccc}(\bar{w})$ for all possible compositions $\bar{w}$. In fact, a stronger lower bound of $N_{ccc}(\bar{w})$ is established in Section II, and proved to be tight for ternary constant-composition codes. Sections III and IV serve to prove our new lower bound of $N_{ccc}(\bar{w})$ is also tight for quaternary constant-composition codes. Finally, by refining and lengthening techniques, we determine the values of $N_{ccc}(\bar{w})$ for all $\bar{w}$ provided that $3w_1 \geq w$. We summarize our main results as below.

Main Theorem 1. Given a composition $\bar{w}$ with at least two components. Let $\lambda = \lceil \frac{w}{w_1} \rceil$ and $s = \lambda w_1 - w$. Then

$$N_{ccc}(\bar{w}) \geq \lambda(\lambda-1)w_1^2 - 2(\lambda-1)sw_1 + w_1 - \lfloor \frac{2s}{\lambda} \rfloor.$$

Equality holds for all $\bar{w}$ provided that $3w_1 \geq w$, and $\bar{w}$ is not a refinement of any composition in $\{\llbracket 4, 4, 2 \rrbracket, \llbracket 4, 3, 3 \rrbracket, \llbracket 5, 5, 3 \rrbracket, \llbracket 5, 4, 4 \rrbracket\}$.

II. LOWER BOUNDS

In this section, we prove the lower bound of $N_{ccc}(\bar{w})$ in Main Theorem 1. Chee *et al.* [2] showed that the following two conditions are necessary and sufficient for a q -ary code $\mathcal{C}$ of constant weight w to have distance $2w-1$:

- (C1) for any distinct $u, v \in \mathcal{C}$, $|\text{supp}(u) \cap \text{supp}(v)| \leq 1$, and
- (C2) for any distinct $u, v \in \mathcal{C}$, if $x \in \text{supp}(u) \cap \text{supp}(v)$, then $u_x \neq v_x$.

The idea of deducing our lower bound is based on the above two conditions, which have been used in [2] to obtain the lower bound in Proposition 2.

Let $\mathcal{C} = \{u^{(1)}, \dots, u^{(|\mathcal{C}|)}\}$ be an $(n, 2w-1, \bar{w})_q$ -code. Then, $\mathcal{C}$ can be regarded as an $|\mathcal{C}| \times n$ matrix C , whose i th row is $u^{(i)}$, $i \in I_{|\mathcal{C}|}$. Let N_i be the number of nonzero entries in column i of C . Then,

$$\sum_{i \in I_n} N_i = |\mathcal{C}|w. \quad (1)$$

In each column of C , we associate each pair of distinct nonzero entries with the pair of rows that contain these entries. There are $\binom{N_i}{2}$ such pairs of nonzero entries in column i of C . Therefore, there are $\sum_{i \in I_n} \binom{N_i}{2}$ such pairs in all columns of C . Since there are no pairs of distinct codewords in $\mathcal{C}$ whose support intersect in two elements, the $\sum_{i \in I_n} \binom{N_i}{2}$ pairs of rows associated with the $\sum_{i \in I_n} \binom{N_i}{2}$ pairs of distinct nonzero entries are also all distinct. Hence,

$$\sum_{i \in I_n} \binom{N_i}{2} \leq \binom{|\mathcal{C}|}{2}. \quad (2)$$

We will use Eq. (2) to obtain our lower bound on $N_{ccc}(\bar{w})$. Given a composition $\bar{w}$, let $\lambda := \lceil \frac{w}{w_1} \rceil$ and $s := \lambda w_1 - w$. Hence, $\lambda \geq 2$ (since $q \geq 3$) and $0 \leq s < w_1$.

We first deal with the case when $w_1 | n$. Let $n = Mw_1$ and $|\mathcal{C}| = M$. It is easy to show that the left hand side of Eq. (2), $\sum_{i \in I_n} \binom{N_i}{2}$ achieves the minimum value when all N_i have almost the same values, that is, N_i equals λ or $\lambda-1$, $i \in I_n$ by Eq. (1). Assume that there are x columns such that $N_i = \lambda-1$. Then by Eq. (1),

$$Mw = \lambda(Mw_1 - x) + (\lambda-1)x.$$

Hence $x = \lambda Mw_1 - Mw = Ms$. By Eq. (2), we have

$$\binom{\lambda}{2}(Mw_1 - Ms) + \binom{\lambda-1}{2}Ms \leq \binom{M}{2},$$

which yields that $M \geq \lambda(\lambda-1)w_1 - 2(\lambda-1)s + 1$. Hence,

$$n \geq \lambda(\lambda-1)w_1^2 - 2(\lambda-1)sw_1 + w_1.$$

Next, suppose that $|\mathcal{C}| = \lambda(\lambda-1)w_1 - 2(\lambda-1)s$ and $n = |\mathcal{C}|w_1 + r$, where $1 \leq r < w_1$. We need to find the smallest integer r such that Eqs. (1) and (2) both hold. By doing the same steps as the case when $w_1 | n$, we deduce that $\lambda(\lambda-1)r \geq |\mathcal{C}|$, that is,

$$r \geq w_1 - \left\lfloor \frac{2s}{\lambda} \right\rfloor.$$

Hence, we have proved the following lower bound on $N_{ccc}(\bar{w})$.

Proposition 3. Given a composition $\bar{w}$ with at least two components. Let $\lambda = \lceil \frac{w}{w_1} \rceil$ and $s = \lambda w_1 - w$. Then

$$N_{ccc}(\bar{w}) \geq L_{ccc}(\bar{w}) - \left\lfloor \frac{2s}{\lambda} \right\rfloor,$$

where $L_{ccc}(\bar{w}) := \lambda(\lambda-1)w_1^2 - 2(\lambda-1)sw_1 + w_1$.

Note that $L_{ccc}(\bar{w})$ is the smallest possible integer n such that $w_1 | n$ and $A_q(n, 2w-1, \bar{w}) = \lfloor \frac{n}{w_1} \rfloor$. Further,

$$\begin{aligned} & L_{ccc}(\bar{w}) - \left\lfloor \frac{2s}{\lambda} \right\rfloor \\ &= w^2 - w_1(w-1) + (w_1s - s^2 - \lfloor \frac{2s}{\lambda} \rfloor). \end{aligned}$$

Since $w_1s - s^2 - \lfloor \frac{2s}{\lambda} \rfloor \geq 0$ when $\lambda \geq 2$, the lower bound in Proposition 3 is stronger than that in Proposition 2.

Observe that in Proposition 3, the lower bound only depends on the total weight w and the biggest component w_1 . By Lemma 1, it is easy to prove the following fact.

Lemma 2. *Suppose that $\bar{w}$ is a refinement of $\bar{v}$ such that $w_1 = v_1$. If $N_{ccc}(\bar{v})$ achieves the lower bound in Proposition 3, so does $N_{ccc}(\bar{w})$.*

Next, we show that for ternary constant-composition codes, the lower bound in Proposition 3 is always achievable.

Proposition 4. *For all $w_1 \geq w_2 \geq 1$, we have $N_{ccc}(\llbracket w_1, w_2 \rrbracket) = 2w_1w_2 + w_2$.*

Proof: Let $\bar{w} = \llbracket w_1, w_2 \rrbracket$. Hence, $\lambda = 2$ and $s = w_1 - w_2$. By Proposition 3, $N_{ccc}(\llbracket w_1, w_2 \rrbracket) \geq 2w_1w_2 + w_2$. By [3],

$$A_3(n, 2w - 1, \llbracket w_1, w_2 \rrbracket) = \max\{M : n \geq M(w_1 + \max\{w_2 - \frac{M-1}{2}, 0\})\}.$$

So we only need to check that for all $n \geq 2w_1w_2 + w_2$, $A_3(n, 2w - 1, \llbracket w_1, w_2 \rrbracket) = \lfloor \frac{n}{w_1} \rfloor$. Let

$$F(n) := n - M(w_1 + \max\{w_2 - \frac{M-1}{2}, 0\}),$$

where $M = \lfloor \frac{n}{w_1} \rfloor$ is a function of n . Since $A_3(n, 2w - 1, \bar{w}) \leq \lfloor \frac{n}{w_1} \rfloor$ for all $\bar{w}$, it suffices to check that $F(n) \geq 0$ for all $n \geq 2w_1w_2 + w_2$. We prove it by induction. It is easy to show that $F(2w_1w_2 + w_2) = 0$. Suppose that $F(n) \geq 0$ for some $n \geq 2w_1w_2 + w_2$, we want to show that $F(n+1) \geq 0$ too. Let $n = Mw_1 + r$, where $0 \leq r < w_1$ and $M \geq 2w_2$. If $r < w_1 - 1$, then $n+1 = Mw_1 + r + 1$. Hence $F(n+1) = F(n) + 1 \geq 1$. If $r = w_1 - 1$, then $n+1 = (M+1)w_1$. Hence $F(n+1) = 0$. This completes the proof. ■

The next two sections will study linear size quaternary constant-composition codes. Given a composition $\bar{w} = \llbracket w_1, w_2, w_3 \rrbracket$, if $w_1 \geq w_2 + w_3$, then $\bar{w}$ is a refinement of $\llbracket w_1, w_2 + w_3 \rrbracket$. By Lemma 2 and Proposition 4, the value of $N_{ccc}(\bar{w})$ can be determined for this case. Hence we assume that $w_1 < w_2 + w_3$, that is, $\lambda = 3$ in the remaining of this paper.

III. CONSTRUCTIONS FOR $n \geq L_{ccc}(\llbracket w_1, w_2, w_3 \rrbracket)$

In this section, we show that $A_4(n, 2w - 1, \bar{w}) = \lfloor \frac{n}{w_1} \rfloor$ for all $n \geq L_{ccc}(\bar{w}) = 6w_1^2 - 4sw_1 + w_1$ based on the existence of difference families.

A. Difference Families

Let $B = \{b_1, \dots, b_k\}$ be a k -subset of $\mathbb{Z}_n$. The list of differences from B is the multiset $\Delta B = \langle b_i - b_j : i, j \in I_k, i \neq j \rangle$. A collection $\{B_1, \dots, B_t\}$ of k -subsets of $\mathbb{Z}_n$ forms an $(n, k; t)$ difference packing, or t -DP (n, k) , if every nonzero element of $\mathbb{Z}_n$ occurs at most once in $\Delta B_1 \cup \dots \cup \Delta B_t$. The sets B_i are base blocks. If every nonzero element of $\mathbb{Z}_n$ occurs exactly once in $\Delta B_1 \cup \dots \cup \Delta B_t$, it is well known as an (n, k) difference family, or DF (n, k) [34]. The parameter t is omitted

since it could be computed from n and k , that is $t = \frac{n-1}{k(k-1)}$. Since t must be an integer, for a DF (n, k) exists, we must have $n \equiv 1 \pmod{k(k-1)}$.

The sizes of base blocks are same in a difference packing. It is natural to generalize difference packings to a collection of subsets with the same property but with varying block sizes. If $t = e_1 + \dots + e_s$, and if there are e_i base blocks of size k_i , then the generalized difference packing is of block type $k_1^{e_1} \dots k_s^{e_s}$, and denoted by GDP $(n, k_1^{e_1} \dots k_s^{e_s})$. Without loss of generality, we assume that $k_1 \geq \dots \geq k_s \geq 2$.

We associate a generalized difference packing with a matrix as follows. Given a GDP $(n, k_1^{e_1} \dots k_s^{e_s})$ with base blocks $B_1, \dots, B_t$. Let $q = k_1 + 1$. For any base block $B_i = \{b_1, \dots, b_k\}$, and any subset $S = \{s_1, \dots, s_k\} \subset I_{q-1}$ of the same size, define $T(B_i, S)$ be the cyclic (right-shifted) $n \times n$ matrix with rows and columns indexed by $\mathbb{Z}_n$, and with the first row u having entries $u_{b_i} = s_i$, $i \in I_k$, and zero entries otherwise. Some useful properties of $T(B_i, S)$ are listed below:

- (P1) for each row and each column, each symbol from S appears exactly once, and
- (P2) for each column, rows r_1 and r_2 both have nonzero entries if and only if $r_1 - r_2 \in \Delta B_i$. Further, such pairs $\{r_1, r_2\}$ are all distinct for all columns.

Proposition 5. *If there exists a GDP $(M, 3^{e_1}2^{e_2})$, then there exists an $(n, 2w - 1, \bar{w})_4$ -code of size M , where $n = w_1M$, $\bar{w} = \llbracket w_1, w_2, w_3 \rrbracket$, and $w_1 = e_1 + e_2$, for any $w_1 \geq w_2 \geq w_3$ such that $w_2 + w_3 = 2e_1 + e_2$.*

Proof: Given a GDP $(M, 3^{e_1}2^{e_2})$. Partition the set of base blocks into three parts $\mathcal{B}_1, \mathcal{B}_2$ and $\mathcal{B}_3$, where $\mathcal{B}_1$ consists of all e_1 blocks of size three, and $\mathcal{B}_2$ contain $w_2 - e_1$ blocks of size two, and $\mathcal{B}_3$ contains the remaining $w_3 - e_1$ blocks of size two. Let $S_1 = I_3$, $S_2 = \{1, 2\}$ and $S_3 = \{1, 3\}$. Define C be the concatenation of w_1 cyclic $M \times M$ matrices: e_1 matrices $T(B, S_1)$ with $B \in \mathcal{B}_1$, followed by $w_2 - e_1$ matrices $T(C, S_2)$ with $C \in \mathcal{B}_2$, and $w_3 - e_1$ matrices $T(D, S_3)$ with $D \in \mathcal{B}_3$.

Let $\mathcal{C}$ be the collection of rows of C . It is immediate that $\mathcal{C}$ is a quaternary code of length n with constant composition $\bar{w}$ fulfilling Conditions (C1) and (C2). ■

In a DF (n, k) , if $n = tk(k-1) + 1$, then t blocks $B_i = \{b_{i,1}, \dots, b_{i,k}\}$ form a perfect (n, k) difference family over $\mathbb{Z}_n$ if the $tk(k-1)/2$ differences $b_{i,h} - b_{i,g}$ ($i \in I_t, 1 \leq g < h \leq k$) cover the set $\{1, 2, \dots, (n-1)/2\}$. If instead, they cover the set $\{1, 2, \dots, (n-3)/2\} \cup \{(n+1)/2\}$, then the difference family is quasi-perfect. We denote them by PDF (n, k) and quasi-PDF (n, k) respectively. The existences of PDF (n, k) s and quasi-PDF (n, k) s are known when $k = 3$.

Theorem 2. [34] *A PDF $(n, 3)$ exists when $n \equiv 1$ or $7 \pmod{24}$, and a quasi-PDF $(n, 3)$ exists when $n \equiv 13$ or $19 \pmod{24}$.*

Corollary 1. *Let $e_1, e_2 \geq 0$ be two integers. Then a GDP $(M, 3^{e_1}2^{e_2})$ exists for all $M \geq 6e_1 + 2e_2 + 1$ except when $e_1 \equiv 2$ or $3 \pmod{4}$ and $(M, e_2) = (6e_1 + 2, 0)$.*

Proof: For each $e_1 \equiv 0$ or $1 \pmod{4}$, let $m = 6e_1 + 1$. By Theorem 2, there exists a PDF($m, 3$) over $\mathbb{Z}_m$. Let $\mathcal{B}$ be the collection of e_1 base blocks. Given any $e_2 \geq 0$, let $P_i = \{0, \frac{m-1}{2} + i\}$, $i \in I_{e_2}$. Then $\mathcal{B} \cup \{P_i : i \in I_{e_2}\}$ is a GDP($M, 3^{e_1 2^{e_2}}$) for all $M \geq m + 2e_2$.

For each $e_1 \equiv 2$ or $3 \pmod{4}$, let $m = 6e_1 + 1$. By Theorem 2, there exists a quasi-PDF($m, 3$) over $\mathbb{Z}_m$, which is also a GDP($M, 3^{e_1}$) for all $M \geq m$ except when $M = m + 1$. Let $\mathcal{B}$ be the collection of e_1 base blocks. Given any $e_2 \geq 1$, let $P_1 = \{0, \frac{m-1}{2}\}$ and $P_i = \{0, \frac{m-1}{2} + i\}$ for all $i \in I_{2, e_2}$. Then $\mathcal{B} \cup \{P_i : i \in I_{e_2}\}$ is a GDP($M, 3^{e_1 2^{e_2}}$) for all $M \geq m + 2e_2$. ■

By the relations among all parameters in Proposition 5, it is easy to show that $6e_1 + 2e_2 + 1 = 6w_1 - 4s + 1$. Combining Corollary 1, it is immediate that the following holds.

Proposition 6. Let $\bar{w} = \llbracket w_1, w_2, w_3 \rrbracket$. Then $A_4(n, 2w - 1, \bar{w}) = \lfloor \frac{n}{w_1} \rfloor$ for all $n = Mw_1$, where $M \geq 6w_1 - 4s + 1$, except when $w_1 = w_2 = w_3 \equiv 2$ or $3 \pmod{4}$ and $n = 6w_1^2 + 2w_1$.

B. Exceptions in Proposition 6

Now we settle the exception cases in Proposition 6. That is, we need to prove that

$$A_4(6w_1^2 + 2w_1, 2w - 1, \llbracket w_1, w_1, w_1 \rrbracket) = 6w_1 + 2$$

for all $w_1 \equiv 2$ or $3 \pmod{4}$. First, we introduce several necessary terminology in combinatorial design theory.

A *set system* is a pair $(X, \mathcal{B})$ such that X is a finite set of *points* and $\mathcal{B}$ is a set of subsets of X , called *blocks*. The *order* of the set system is $|X|$, the number of points. For a nonnegative integer k , a set system $(X, \mathcal{B})$ is said to be *k-uniform* if $|B| = k$ for all $B \in \mathcal{B}$.

Let $v \geq k$. A (v, k) -*packing* is a k -uniform set system $(X, \mathcal{B})$ of order v , such that each pair of X occurs in at most one block in $\mathcal{B}$. The *packing number* $D(v, k)$ is the maximum number of blocks in any (v, k) -packing. A (v, k) -packing $(X, \mathcal{B})$ is said to be *optimal* if $|\mathcal{B}| = D(v, k)$. The values of $D(v, k)$ have been determined for all v when $k \in \{3, 4\}$ [35]. In particular, we have

$$D(v, 3) = \begin{cases} \left\lfloor \frac{v}{3} \left\lfloor \frac{v-1}{2} \right\rfloor \right\rfloor - 1, & \text{if } v \equiv 5 \pmod{6}; \\ \left\lfloor \frac{v}{3} \left\lfloor \frac{v-1}{2} \right\rfloor \right\rfloor, & \text{otherwise.} \end{cases}$$

In fact, when $v \equiv 1, 3 \pmod{6}$, an optimal $(v, 3)$ -packing is also called a *Steiner triple system of order v*, denoted by STS(v). In this case, each pair of points occurs exactly once.

Suppose that $\mathcal{C}$ is an $(n, 2w - 1, \bar{w})_4$ -code of size $6w_1 + 2$ with composition $\bar{w} = \llbracket w_1, w_1, w_1 \rrbracket$ and length $n = 6w_1^2 + 2w_1$. As in Section II, let $\mathcal{C} = \{u^{(1)}, \dots, u^{(6w_1+2)}\}$. Then, $\mathcal{C}$ can be regarded as an $(6w_1 + 2) \times (6w_1^2 + 2w_1)$ matrix C , whose i th row is $u^{(i)}$, $i \in I_{6w_1+2}$. Using the same notation in Section II, for each column $c \in I_{6w_1^2+2w_1}$, $N_c = 3$ and each nonzero entry from I_3 occurs exactly once. Thus we can define a triple $A_c = (a_1, a_2, a_3)$ for each column c , such that

the nonzero element i occurs in row a_i , $i \in I_3$. Let $\mathcal{A} = \{A_c : c \in I_{6w_1^2+2w_1}\}$, then $\mathcal{A}$ has the following properties.

- (T1) For each c , three elements in A_c are all different. Let B_c be the set with elements in A_c and $\mathcal{B} = \{B_c : c \in I_{6w_1^2+2w_1}\}$. Then $(I_{6w_1+2}, \mathcal{B})$ is an $(6w_1 + 2, 3)$ -packing of size $6w_1^2 + 2w_1$, which is optimal according to the packing number.
- (T2) For each position $i \in I_3$, each element in I_{6w_1+2} occurs in position i of w_1 triples from $\mathcal{A}$.

The converse is true. Suppose that $\mathcal{A}$ is a collection of $6w_1^2 + 2w_1$ triples in I_{6w_1+2} satisfying (T1) and (T2). We construct an $(6w_1 + 2) \times (6w_1^2 + 2w_1)$ matrix C in a natural way, then the rows of C form an $(n, 2w - 1, \bar{w})_4$ -code of size $6w_1 + 2$ with composition $\bar{w} = \llbracket w_1, w_1, w_1 \rrbracket$ and length $n = 6w_1^2 + 2w_1$.

Now we aim to construct a collection of $6w_1^2 + 2w_1$ triples in I_{6w_1+2} satisfying (T1) and (T2) for all $w_1 \equiv 2$ or $3 \pmod{4}$. Actually, they exist for all positive integers w_1 . Before stating our general construction, we give a small example first.

Example 1. Let $w_1 = 2$ and $G = \mathbb{Z}_5 \oplus \mathbb{Z}_3$. Write x_y for the pair $(x, y) \in \mathbb{Z}_u \oplus \mathbb{Z}_3$. Let $B_0 = \{0_0, 0_1, 0_2\}$, $B_1 = \{0_0, 2_0, 1_1\}$ and $B_2 = \{0_0, 4_0, 2_1\}$. For each B_i , define $\text{Orb}_G B_i := \{B_i + g : g \in G\}$. Then $\mathcal{B}' = \cup_{i \in I_{0,2}} \text{Orb}_G B_i$ is the block set of an STS(15) over G due to Skolem [36]. Let $\mathcal{B} = \mathcal{B}' \setminus \{B \in \mathcal{B}' : 4_2 \in B\}$, then $(G \setminus \{4_2\}, \mathcal{B})$ is an optimal $(14, 3)$ -packing of size 28. View blocks in $\mathcal{B}'$ as ordered triples. We show the reordering procedures in Table I.

First look at the set $\text{Orb}_G B_1 \cup \text{Orb}_G B_2$, in which each element from G occurs twice in each position. After deleting triples containing the element 4_2 , elements $2_1, 3_1, 0_2, 2_2$ occurs only once in the first position, elements $0_1, 1_1, 1_2, 3_2$ occurs only once in the second position, and elements $0_0, 1_0, 2_0, 3_0$ occurs only once in the third position. It is natural to think of reordering triples from $\text{Orb}_G B_0$ to increase the occurrences of these elements. After reordering the remaining triples of $\text{Orb}_G B_0$ as in Table I, element 1_2 occurs three times in the first position but only once in the second position, while element 2_2 occurs only once in the first position but three times in the second position. Finally, we exchange the first two elements in the triple $(1_2, 3_2, 2_0)$ from $\text{Orb}_G B_1$ and $(3_2, 2_2, 0_0)$ from $\text{Orb}_G B_2$ to balance the occurrences.

Proposition 7. For all positive integers w_1 , there exists a $(6w_1 + 2, 3)$ -packing of size $6w_1^2 + 2w_1$, such that the blocks can be directed to satisfy (T2).

Proof: Let $u = 2w_1 + 1$. We start from an STS($3u$) which is due to Skolem [36]. Let $G = \mathbb{Z}_u \oplus \mathbb{Z}_3$. Choose base blocks $B_0 = \{0_0, 0_1, 0_2\}$ and $B_x = \{0_0, (2x)_0, x_1\}$, $x \in I_{w_1}$. It is easy to see that $|\text{Orb}_G B_0| = u$ and $|\text{Orb}_G B_x| = 3u$, $x \in I_{w_1}$. Let $\mathcal{B}' = \cup_{x \in I_{0, w_1}} \text{Orb}_G B_x$. Then $(G, \mathcal{B}')$ is an STS($3u$). Let $\mathcal{B} = \mathcal{B}' \setminus \{B \in \mathcal{B}' : (2w_1)_2 \in B\}$ and $X = G \setminus \{(2w_1)_2\}$. Then $(X, \mathcal{B})$ is an optimal $(6w_1 + 2, 3)$ -packing of size $6w_1^2 + 2w_1$.

Now we direct blocks in $\mathcal{B}$ to satisfy (T2). For each $x \in I_{0, w_1}$, let A_x be the triple with entries from B_x and orders as displayed. Then define $\text{Orb}_G A_x$ accordingly by keeping the orders. To be clear, $\text{Orb}_G A_0 = \{(i_0, i_1, i_2) : i \in \mathbb{Z}_u\}$. Let

TABLE I: Reordering Procedures in Example 1

Orb _G B ₀		reordering	
(0 ₀ , 0 ₁ , 0 ₂)		→ (0 ₂ , 0 ₁ , 0 ₀)	
(1 ₀ , 1 ₁ , 1 ₂)		→ (1 ₂ , 1 ₁ , 1 ₀)	
(2 ₀ , 2 ₁ , 2 ₂)		→ (2 ₁ , 2 ₂ , 2 ₀)	
(3 ₀ , 3 ₁ , 3 ₂)		→ (3 ₁ , 3 ₂ , 3 ₀)	
~~(4₀, 4₁, 4₂)~~			
Orb _G B ₁	reordering	Orb _G B ₂	reordering
(0 ₀ , 2 ₀ , 1 ₁)	→ (3 ₂ , 1 ₂ , 2 ₀)	(0 ₀ , 4 ₀ , 2 ₁)	
(1 ₀ , 3 ₀ , 2 ₁)		(1 ₀ , 0 ₀ , 3 ₁)	
(2 ₀ , 4 ₀ , 3 ₁)		(2 ₀ , 1 ₀ , 4 ₁)	
(3 ₀ , 0 ₀ , 4 ₁)		(3 ₀ , 2 ₀ , 0 ₁)	
(4 ₀ , 1 ₀ , 0 ₁)		(4 ₀ , 3 ₀ , 1 ₁)	
(0 ₁ , 2 ₁ , 1 ₂)		(0 ₁ , 4 ₁ , 2 ₂)	
(1 ₁ , 3 ₁ , 2 ₂)		(1 ₁ , 0 ₁ , 3 ₂)	
(2 ₁ , 4 ₁ , 3 ₂)		~~(2₁, 1₁, 4₂)~~	
~~(3₁, 0₁, 4₂)~~		(3 ₁ , 2 ₁ , 0 ₂)	
(4 ₁ , 1 ₁ , 0 ₂)		(4 ₁ , 3 ₁ , 1 ₂)	
(0 ₂ , 2 ₂ , 1 ₀)		~~(0₂, 4₂, 2₀)~~	
(1 ₂ , 3 ₂ , 2 ₀)		(1 ₂ , 0 ₂ , 3 ₀)	
~~(2₂, 4₂, 3₀)~~		(2 ₂ , 1 ₂ , 4 ₀)	
(3 ₂ , 0 ₂ , 4 ₀)		(3 ₂ , 2 ₂ , 0 ₀)	
~~(4₂, 1₂, 0₀)~~		~~(4₂, 3₂, 1₀)~~	
			→ (2 ₂ , 3 ₂ , 0 ₀)

$\mathcal{A}$ be the collection of all triples from $\text{Orb}_G A_x$, $x \in I_{0,w_1}$, which do not contain the element $(2w_1)_2$. Note that $\mathcal{A}$ doesn't satisfy (T2) at this moment. The first three rows of Table II point out the sets of elements occurring $w_1 - 1$ or $w_1 + 1$ in each position, all others occur in the corresponding positions exactly w_1 times in $\mathcal{A}$. The $w_1 - 1$ occurrences happen when deleting triples from $\text{Orb}_G A_x$, $x \in I_{w_1}$, while $w_1 + 1$ happens because of triples from $\text{Orb}_G A_0$.

For convenience, denote $\iota = \lfloor \frac{w_1}{2} \rfloor$ and $\kappa = \lceil \frac{w_1}{2} \rceil$. We follow the steps below.

- (S1) In $\text{Orb}_G A_0$, change $((2i+1)_0, (2i+1)_1, (2i+1)_2)$ to $((2i+1)_1, (2i+1)_2, (2i+1)_0)$ for all $i \in I_{\iota, w_1-1}$.
- (S2) In $\text{Orb}_G A_0$, change $((2i)_0, (2i)_1, (2i)_2)$ to $((2i)_1, (2i)_2, (2i)_0)$ for all $i \in I_{\kappa, w_1-1}$.
- (S3) In $\text{Orb}_G A_0$, change (i_0, i_1, i_2) to (i_2, i_1, i_0) for all $i \in I_{0, w_1-1}$.
- (S4) Finally, in $\text{Orb}_G A_\kappa$, change $((2i+1)_2, (2\kappa+2i+1)_2, (\kappa+2i+1)_0)$ to $((2\kappa+2i+1)_2, (2i+1)_2, (\kappa+2i+1)_0)$ for all $i \in I_{0, \iota-1}$. At the same time, in $\text{Orb}_G A_{w_1}$, change $((2i+1)_2, (2i)_2, (w_1+2i+1)_0)$ to $((2i)_2, (2i+1)_2, (w_1+2i+1)_0)$ for all $i \in I_{\kappa, w_1-1}$. Note that $\{(2\kappa+2i+1)_2 : i \in I_{0, \iota-1}\} = \{(2i+1)_2 : i \in I_{\kappa, w_1-1}\}$.

After each step, all elements occur at least $w_1 - 1$ and at most $w_1 + 1$ times in each position. We list the elements occurring $w_1 - 1$ or $w_1 + 1$ in each position after each step in Table II. It is easy to check that after (S4), all elements occur w_1 times in each position. Thus triples in $\mathcal{A}$ can be reordered to satisfies (T2).

■

The following consequence is immediate.

Proposition 8. For all positive integers w_1 and $n = 6w_1^2 + 2w_1$,

$$A_4(n, 2w - 1, \llbracket w_1, w_1, w_1 \rrbracket) = 6w_1 + 2.$$

Combining Propositions 6 and 8, we have shown that $A_4(n, 2w - 1, \llbracket w_1, w_2, w_3 \rrbracket) = \lfloor \frac{n}{w_1} \rfloor$ for all $n = Mw_1$, where $M \geq 6w_1 - 4s + 1$. By the lengthening method (adding zeros in the end of codewords) used in [2], we have $A_4(n, 2w - 1, \llbracket w_1, w_2, w_3 \rrbracket) = \lfloor \frac{n}{w_1} \rfloor$ for all $n \geq L_{ccc}(\overline{w}) = 6w_1^2 - 4sw_1 + w_1$.

IV. CONSTRUCTIONS FOR $n < L_{ccc}(\llbracket w_1, w_2, w_3 \rrbracket)$

To determine values of $N_{ccc}(\overline{w})$, we still need to prove that $A_4(n, 2w - 1, \llbracket w_1, w_2, w_3 \rrbracket) = \lfloor \frac{n}{w_1} \rfloor$ for $n = L_{ccc}(\overline{w}) - \lfloor \frac{2s}{3} \rfloor$. If $s \leq 1$, then $n = L_{ccc}(\overline{w})$, which has been settled in Section III. From now on, we assume that $s \geq 2$. We generalize the method in Section III-B to deal with this case.

Let K be a set of positive integers. A (v, K) -packing is a set system $(X, \mathcal{B})$ of order v , with block sizes from K , such that each pair of X occurs in at most one block in $\mathcal{B}$.

Let $\mathcal{A}$ be a collection of n triples with distinct elements in $\mathbb{Z}_M \cup \{*\}$ satisfying the following two properties.

(T1') For each $A \in \mathcal{A}$, let B_A be the set containing all elements in A but not $*$. Let $\mathcal{B} = \{B_A : A \in \mathcal{A}\}$. $(\mathbb{Z}_M, \mathcal{B})$ is an $(M, \{2, 3\})$ -packing of size n .

(T2') for each position $i \in I_3$, each element in $\mathbb{Z}_M$ occurs in position i of w_i triples in $\mathcal{A}$.

From $\mathcal{A}$, we can construct an $(n, 2w - 1, \overline{w})_4$ -code with $\overline{w} = \llbracket w_1, w_2, w_3 \rrbracket$ of size M . The construction is similar to that in Section III-B but simply ignore the symbol $*$.

To construct $\mathcal{A}$, we use different method from that in Proposition 7. We first find a candidate of $\mathcal{A}$ satisfying (T2'), then try to modify it to satisfy (T1'). We show the idea by the following example.

Example 2. Let $M = 10$ and $\overline{w} = \llbracket 3, 2, 2 \rrbracket$. Let $A_1 = (0, 1, 6)$, $A_2 = (0, 2, *)$ and $A_3 = (0, *, 3)$. Then $\mathcal{A} = \cup_{i \in I_3} \text{Orb}_{\mathbb{Z}_{10}} A_i$ is a candidate satisfying (T2') over $\mathbb{Z}_{10}$. Note that the difference 5 occurs twice in ΔA_1 . We first do the following changes in $\text{Orb}_{\mathbb{Z}_{10}} A_1$:

$$\begin{aligned} (0, 1, 6) &\rightarrow (0, *, 6), \\ (1, 2, 7) &\rightarrow (1, 2, *), \\ (3, 4, 9) &\rightarrow (3, *, 9), \\ (4, 5, 0) &\rightarrow (4, 5, *), \\ (7, 8, 3) &\rightarrow (7, 8, *). \end{aligned}$$

Then add two more triples

$$\begin{aligned} (*, 1, 0), \\ (*, 4, 3). \end{aligned}$$

Finally, change the following triple in $\text{Orb}_{\mathbb{Z}_{10}} A_2$

$$(1, 3, *) \rightarrow (1, 3, 7).$$

Note that we do not change the positions of symbols from $\mathbb{Z}_{10}$ appearing in $\mathcal{A}$. Further, the pairs newly occurring in the last two steps are pairs deleted in the first step. Thus we have a set of 32 triples satisfying both (T1') and (T2'), which yields that $A_4(32, 13, \llbracket 3, 2, 2 \rrbracket) = 10$ and $N_{ccc}(\llbracket 3, 2, 2 \rrbracket) = 32$.

TABLE II: Abnormal elements in three positions in $\mathcal{A}$ in Proposition 7

Originally	0	1	2
$w_1 - 1$	$\{(2i)_2 : i \in I_{0,w_1-1}\} \cup \{(w_1)_1 : i \in I_{w_1,2w_1-1}\}$	$\{(2i+1)_2 : i \in I_{0,w_1-1}\}$	$\{i_0 : i \in I_{0,2w_1-1}\}$
$w_1 + 1$	$\{i_0 : i \in I_{0,2w_1-1}\}$	$\{(w_1)_1 : i \in I_{w_1,2w_1-1}\}$	$\{i_2 : i \in I_{0,2w_1-1}\}$
After (S1)	0	1	2
$w_1 - 1$	$\{(2i)_2 : i \in I_{0,w_1-1}\} \cup \{(2i)_1 : i \in I_{\kappa,w_1-1}\}$	$\{(2i+1)_2 : i \in I_{0,\ell-1}\}$	$\{i_0 : i \in I_{0,w_1-1}\} \cup \{(2i)_0 : i \in I_{\kappa,w_1-1}\}$
$w_1 + 1$	$\{i_0 : i \in I_{0,w_1-1}\} \cup \{(2i)_0 : i \in I_{\kappa,w_1-1}\}$	$\{(2i)_1 : i \in I_{\kappa,w_1-1}\}$	$\{i_2 : i \in I_{0,w_1-1}\} \cup \{(2i)_2 : i \in I_{\kappa,w_1-1}\}$
After (S2)	0	1	2
$w_1 - 1$	$\{(2i)_2 : i \in I_{0,w_1-1}\}$	$\{(2i+1)_2 : i \in I_{0,\ell-1}\}$	$\{i_0 : i \in I_{0,w_1-1}\}$
$w_1 + 1$	$\{i_0 : i \in I_{0,w_1-1}\}$	$\{(2i)_2 : i \in I_{\kappa,w_1-1}\}$	$\{i_2 : i \in I_{0,w_1-1}\}$
After (S3)	0	1	2
$w_1 - 1$	$\{(2i)_2 : i \in I_{\kappa,w_1-1}\}$	$\{(2i+1)_2 : i \in I_{0,\ell-1}\}$	$\emptyset$
$w_1 + 1$	$\{(2i+1)_2 : i \in I_{0,\ell-1}\}$	$\{(2i)_2 : i \in I_{\kappa,w_1-1}\}$	$\emptyset$

Proposition 9. Let $e_1 \geq 0$, $e_2 \geq 3$, $2|M$ and $M \geq 8$ be integers. If there exists a $\text{GDP}(M, 3^{e_1}2^{e_2})$ with three specified base blocks $\{0, 1\}$, $\{0, 2\}$ and $\{0, M/2 - 1\}$, then there exists an $(n, 2w - 1, \overline{w})_4$ -code of size M , where $n = w_1M + \lceil \frac{M}{6} \rceil$, $\overline{w} = \llbracket w_1, w_2, w_3 \rrbracket$, and $w_1 = e_1 + e_2 - 1$, for any $w_1 \geq w_2 \geq w_3$ such that $w_2 + w_3 = 2e_1 + e_2$.

Proof: Suppose that $\mathcal{B}'$ is the set of base blocks of a $\text{GDP}(M, 3^{e_1}2^{e_2})$ over $\mathbb{Z}_M$. Let $B_1 = \{0, 1, M/2 + 1\}$, $B_2 = \{0, 2\}$ and

$$\mathcal{B} = (\mathcal{B}' \setminus \{\{0, 1\}, \{0, M/2 - 1\}\}) \cup \{B_1\}.$$

Note that $\mathcal{B}$ has $e_1 + e_2 - 1$ blocks. Partition $\mathcal{B}$ into three parts $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$, where $\mathcal{B}_1$ consists of all $e_1 + 1$ blocks of size three, and $\mathcal{B}_2$ contain $w_2 - e_1 - 1$ blocks of size two including B_2 specifically, and $\mathcal{B}_3$ contains the remaining $w_3 - e_1 - 1$ blocks of size two. For each block $B \in \mathcal{B}_1$, let A_B be an ordered triple with elements from B . For each block $B = \{a, b\} \in \mathcal{B}_2$, let $A_B = (a, b, *)$. For each block $B = \{a, b\} \in \mathcal{B}_3$, let $A_B = (a, *, b)$. Especially, let $A_{B_1} = (0, 1, M/2 + 1)$ and $A_{B_2} = (0, 2, *)$. Let $\mathcal{A} = \cup_{B \in \mathcal{B}} \text{Orb}_{\mathbb{Z}_M} A_B$, then $\mathcal{A}$ is a candidate of w_1M triples over $\mathbb{Z}_M \cup \{*\}$ satisfying (T2').

Now we do modifications on triples from $\mathcal{A}$ to make it satisfy both (T1') and (T2'). The main idea is as follows. The set $\mathcal{A}$ does not satisfy (T1') since the difference $M/2$ occurs twice in ΔB_1 . Thus, we first choose $M/2$ triples from $\text{Orb}_{\mathbb{Z}_M} A_{B_1}$ to change one symbol from the repeated pairs to $*$. Besides the $M/2$ repeated pairs, there are $M/2$ other pairs also broke in this step. We let them appear somewhere else by adding $\lceil \frac{M}{6} \rceil$ triples of type $(*, \cdot, \cdot)$, and changing the $*$ to some symbol of $\mathbb{Z}_M$ in $M/2 - 2\lceil \frac{M}{6} \rceil$ triples from $\text{Orb}_{\mathbb{Z}_M} A_{B_2}$. Details for different congruent classes are listed in Table III. ■

Note that for all $s \geq 2$,

$$\begin{aligned} n &= L_{ccc}(\overline{w}) - \lfloor \frac{2s}{3} \rfloor \\ &= 6w_1^2 - 4sw_1 + w_1 - \lfloor \frac{2s}{3} \rfloor \\ &= w_1(6w_1 - 4s) + \lceil \frac{6w_1 - 4s}{6} \rceil. \end{aligned}$$

To prove $A_4(n, 2w - 1, \llbracket w_1, w_2, w_3 \rrbracket) = 6w_1 - 4s$ for $n = L_{ccc}(\overline{w}) - \lfloor \frac{2s}{3} \rfloor$, by Proposition 9, we need to construct a $\text{GDP}(M, 3^{e_1}2^{e_2})$ with three specified base blocks $\{0, 1\}$,

$\{0, 2\}$ and $\{0, M/2 - 1\}$, where $M = 6w_1 - 4s$, $e_1 = w_1 - s - 1$ and $e_2 = s + 2$, for all $w_1 > s \geq 2$.

Proposition 10. Given integers $w_1 > s \geq 2$ and $(w_1, s) \notin \{(4, 2), (5, 2)\}$. There exists a $\text{GDP}(M, 3^{e_1}2^{e_2})$ over $\mathbb{Z}_M$ with three specified base blocks $\{0, 1\}$, $\{0, 2\}$ and $\{0, M/2 - 1\}$, where $M = 6w_1 - 4s$, $e_1 = w_1 - s - 1$ and $e_2 = s + 2$.

Proof: We split it into four cases based on the values of $w_1 - s$. For all cases, the e_2 base blocks of size two are of type $\{0, d\}$, where d covers all differences that do not appear in base blocks of size three. To save space, we only list the e_1 base blocks of size three in each case. Note that the differences 1, 2, $M/2$ and $M/2 - 1$ do not appear in any base blocks of size three. Thus the GDP contains the three mentioned base blocks of size two. Further, in the first two cases, that is when $w_1 - s \equiv 0, 1 \pmod{4}$, the set of base blocks of size three are obtained by modifying some blocks in a $\text{DF}(M, 3)$ found in [37].

When $w_1 - s = 4k$, then $M = 24k + 2s$ and $e_1 = 4k - 1$. If $k \geq 2$, then the e_1 base blocks are

$$\begin{aligned} &\{0, \quad 6k - 1, \quad 18k + 2s - 1\}, \\ &\{0, \quad 4k - 1, \quad 9k - 1\}, \\ &\{0, \quad 2k, \quad 10k - 1\}, \\ &\{0, \quad 4k, \quad 10k\}, \\ &\{0, \quad 2k + 2r - 1, \quad 7k + r - 1\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2k + 2r, \quad 11k + r - 1\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2r + 1, \quad 10k + r\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2r, \quad 6k + r\}, \quad 2 \leq r \leq k - 1. \end{aligned}$$

If $k = 1$, and $s = 2$, that is $6w - 4s = 28$, then the base blocks are $\{0, 3, 11\}, \{0, 4, 9\}, \{0, 6, 16\}$. If $s \geq 3$, then the base blocks are $\{0, 3, 13\}, \{0, 4, 12\}, \{0, 5, 11\}$.

When $w_1 - s = 4k + 1$, then $M = 24k + 2s + 6$ and $e_1 = 4k$. If $k \geq 2$, then the e_1 base blocks are

$$\begin{aligned} &\{0, \quad 6k, \quad 18k + 2s + 4\}, \\ &\{0, \quad 4k - 1, \quad 9k\}, \\ &\{0, \quad 4k, \quad 10k + 1\}, \\ &\{0, \quad 4k + 1, \quad 12k + 1\}, \\ &\{0, \quad 2k, \quad 12k\}, \\ &\{0, \quad 2k + 2r - 1, \quad 7k + r\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2k + 2r, \quad 11k + r\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2r + 1, \quad 10k + r + 1\}, \quad 1 \leq r \leq k - 1, \\ &\{0, \quad 2r, \quad 6k + r + 1\}, \quad 2 \leq r \leq k - 1. \end{aligned}$$

TABLE III: Modifications made on triples of $\mathcal{A}$ in Proposition 9

$M = 6k, k \geq 2$	$\text{Orb}_{\mathbb{Z}_M} A_{B_1}$	$(3i, 3i+1, 3i+3k+1) \rightarrow (3i, 3i+1, *), i \in I_{0,k-1}$ $(3i+3k-1, 3i+3k, 3i) \rightarrow (3i+3k-1, 3i+3k, *), i \in I_{0,k-1}$ $(3i+3k+1, 3i+3k+2, 3i+2) \rightarrow (3i+3k+1, *, 3i+2), i \in I_{0,k-1}$
	Add triples	$(*, 3i+3k+2, 3i+3k+1), i \in I_{0,k-1}$
	$\text{Orb}_{\mathbb{Z}_M} A_{B_2}$	$(3i+3k-1, 3i+3k+1, *) \rightarrow (3i+3k-1, 3i+3k+1, 3i), i \in I_{0,k-1}$
$M = 6k+2, k \geq 1$	$\text{Orb}_{\mathbb{Z}_M} A_{B_1}$	$(3i+1, 3i+2, 3i+3k+3) \rightarrow (3i+1, 3i+2, *), i \in I_{0,k-1}$ $(3i+3k, 3i+3k+1, 3i) \rightarrow (3i+3k, 3i+3k+1, *), i \in I_{0,k-1}$ $(3i, 3i+1, 3i+3k+2) \rightarrow (3i, *, 3i+3k+2), i \in I_{0,k-1}$ $(6k, 6k+1, 3k) \rightarrow (6k, *, 3k)$
	Add triples	$(*, 3i+1, 3i), i \in I_{0,k-1}$ $(*, 6k+1, 6k)$
	$\text{Orb}_{\mathbb{Z}_M} A_{B_2}$	$(3i+1, 3i+3, *) \rightarrow (3i+1, 3i+3, 3i+3k+3), i \in I_{0,k-2}$
$M = 6k+4, k \geq 1$	$\text{Orb}_{\mathbb{Z}_M} A_{B_1}$	$(3i+1, 3i+2, 3i+3k+3) \rightarrow (3i+1, 3i+2, *), i \in I_{0,k-1}$ $(3i+3k+1, 3i+3k+2, 3i) \rightarrow (3i+3k+1, 3i+3k+2, *), i \in I_{0,k}$ $(3i, 3i+1, 3i+3k+3) \rightarrow (3i, *, 3i+3k+3), i \in I_{0,k}$
	Add triples	$(*, 3i+1, 3i), i \in I_{0,k-1}$
	$\text{Orb}_{\mathbb{Z}_M} A_{B_2}$	$(3i+1, 3i+3, *) \rightarrow (3i+1, 3i+3, 3i+3k+4), i \in I_{0,k-1}$

If $k = 1$, and $s \geq 2$, then the base blocks are $\{0, 3, 14\}, \{0, 4, 12\}, \{0, 5, 15\}, \{0, 6, 13\}$.

When $w_1 - s = 4k + 2$, then $M = 24k + 2s + 12$ and $e_1 = 4k + 1$. If $k \geq 1$, then the e_1 base blocks are

$$\begin{aligned} &\{0, 4k+1, 10k+6\}, \\ &\{0, 4k+3, 10k+7\}, \\ &\{0, 2r, 6k+r+5\}, \quad 2 \leq r \leq 2k+1, \\ &\{0, 2r+1, 10k+r+7\}, \quad 1 \leq r \leq k, \\ &\{0, 2k+2r+1, 11k+r+7\}, \quad 1 \leq r \leq k-1. \end{aligned}$$

If $k = 0$, and $s \geq 3$, then the only base block is $\{0, 3, 7\}$.

When $w_1 - s = 4k + 3$, then $M = 24k + 2s + 18$ and $e_1 = 4k + 2$. If $k \geq 1$, then the e_1 base blocks are

$$\begin{aligned} &\{0, 4k+1, 10k+9\}, \\ &\{0, 4k+3, 10k+10\}, \\ &\{0, 4k+4, 8k+9\}, \\ &\{0, 2r, 6k+r+7\}, \quad 2 \leq r \leq 2k+1, \\ &\{0, 2r+1, 10k+r+10\}, \quad 1 \leq r \leq k-1, \\ &\{0, 2k+2r-1, 11k+r+9\}, \quad 1 \leq r \leq k. \end{aligned}$$

If $k = 0$, and $s \geq 3$, then the base blocks are $\{0, 3, 8\}, \{0, 4, 10\}$. ■

Combining Proposition 9 and 10, we obtain the following result.

Proposition 11. *Given a composition $\bar{w} = [w_1, w_2, w_3]$ such that $2 \leq s < w_1$, where $s = 3w_1 - w$. Let $n = L_{ccc}(\bar{w}) - \lfloor \frac{2s}{3} \rfloor$. We have*

$$A_4(n, 2w-1, [w_1, w_2, w_3]) = \lfloor \frac{n}{w_1} \rfloor,$$

for all $w_1 \geq w_2 \geq w_3 \geq 1$ such that $(w_1, s) \notin \{(4, 2), (5, 2)\}$.

By Propositions 6, 8, 11 and the lengthening method, we have determined the value of $N_{ccc}(\bar{w})$ for quaternary constant-composition codes. We state it in the following proposition.

Proposition 12. *Given a composition $\bar{w} = [w_1, w_2, w_3]$ such that $w_1 < w_2 + w_3$. We have*

$$N_{ccc}(\bar{w}) = 6w_1^2 - 4sw_1 + w_1 - \lfloor \frac{2s}{3} \rfloor,$$

where $s = 3w - w_1$, except possibly when $\bar{w} \in \{\llbracket 4, 4, 2 \rrbracket, \llbracket 4, 3, 3 \rrbracket, \llbracket 5, 5, 3 \rrbracket, \llbracket 5, 4, 4 \rrbracket\}$.

V. CONCLUSION

Several new direct constructions for optimal quaternary constant-composition codes are given based on combinatorial method. Consequently, we determine the values of $N_{ccc}(\bar{w})$, the smallest length n such that $A_4(n', 2w-1, \bar{w}) = \lfloor \frac{n'}{w_1} \rfloor$ for all $n' \geq n$, with only four possible exceptions. The exact values of $N_{ccc}(\bar{w})$ show that our newly established lower bound of $N_{ccc}(\bar{w})$ is tight in these cases. Our main result, Main Theorem 1 follows from Propositions 3, 4, 12 and the refining method in Lemma 2.

REFERENCES

- [1] M. Svanstrom, "Constructions of ternary constant-composition codes with weight three," *Information Theory, IEEE Transactions on*, vol. 46, no. 7, pp. 2644–2647, Nov 2000.
- [2] Y. M. Chee, S. H. Dau, A. Ling, and S. Ling, "Linear size optimal q-ary constant-weight codes and constant-composition codes," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 140–151, Jan 2010.
- [3] M. Svanström, P. R. J. Östergård, and G. T. Bogdanova, "Bounds and constructions for ternary constant-composition codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 1, pp. 101–111, 2002.
- [4] W. Chu, C. J. Colbourn, and P. Dukes, "On constant composition codes," *Discrete Applied Mathematics*, vol. 154, no. 6, pp. 912 – 929, 2006.
- [5] C. Ding and J. Yin, "Combinatorial constructions of optimal constant-composition codes," *IEEE transactions on information theory*, vol. 51, no. 10, pp. 3671–3674, 2005.
- [6] Y. Luo, F.-W. Fu, W. Chen *et al.*, "On constant-composition codes over $\mathbb{Z}_4$," *Information Theory, IEEE Transactions on*, vol. 49, no. 11, pp. 3010–3016, 2003.
- [7] C. Ding and J. Yuan, "A family of optimal constant-composition codes," *Information Theory, IEEE Transactions on*, vol. 51, no. 10, pp. 3668–3671, 2005.
- [8] I. BOUNDS, "Algebraic constructions of constant composition codes," *IEEE transactions on information theory*, vol. 51, no. 4, p. 1585, 2005.
- [9] C. Ding and J. Yin, "A construction of optimal constant composition codes," *Des. Codes Cryptogr.*, vol. 40, no. 2, pp. 157–165, 2006.
- [10] Y. M. Chee, A. C. Ling, S. Ling, and H. Shen, "The pbd-closure of constant-composition codes," *Information Theory, IEEE Transactions on*, vol. 53, no. 8, pp. 2685–2692, 2007.
- [11] G. T. Bogdanova and S. N. Kapralov, "Enumeration of optimal ternary constant-composition codes," *Problems of Information Transmission*, vol. 39, no. 4, pp. 346–351, 2003.

- [12] C. Ding, "Optimal constant composition codes from zero-difference balanced functions," *Information Theory, IEEE Transactions on*, vol. 54, no. 12, pp. 5766–5770, 2008.
- [13] S. Huczynska, "Equidistant frequency permutation arrays and related constant composition codes," *Des. Codes Cryptogr.*, vol. 54, no. 2, pp. 109–120, 2010.
- [14] Y. M. Chee, G. Ge, and A. C. H. Ling, "Group divisible codes and their application in the construction of optimal constant-composition codes of weight three," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3552–3564, 2008.
- [15] J. Yan and J. Yin, "A class of optimal constant composition codes from gdrps," *Designs, Codes and Cryptography*, vol. 50, no. 1, pp. 61–76, 2009.
- [16] F. Gao and G. Ge, "Optimal ternary constant-composition codes of weight four and distance five," *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3742–3757, 2011.
- [17] J. Yin and Y. Tang, "A new combinatorial approach to the construction of constant composition codes," *Science in China Series A: Mathematics*, vol. 51, no. 3, pp. 416–426, 2008.
- [18] Y. Ding, "A construction for constant-composition codes," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3738–3741, 2008.
- [19] M. Zhu and G. Ge, "Quaternary constant-composition codes with weight 4 and distances 5 or 6," *Information Theory, IEEE Transactions on*, vol. 58, no. 9, pp. 6012–6022, 2012.
- [20] H. Wei, H. Zhang, M. Zhu, and G. Ge, "Optimal ternary constant-composition codes with weight four and distance six," *Discrete Mathematics*, vol. 338, no. 3, pp. 72–87, 2015.
- [21] Y. M. Chee, F. Gao, H. M. Kiah, A. C. H. Ling, H. Zhang, and X. Zhang, "Decompositions of edge-colored digraphs: A new technique in the construction of constant-weight codes and related families," in *Information Theory (ISIT), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 1436–1440.
- [22] P. Moulin, "The log-volume of optimal constant-composition codes for memoryless channels, within $o(1)$ bits," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 826–830.
- [23] J. Scarlett, A. Martinez, and A. G. i Fàbregas, "Refinements of the third-order term in the fixed error asymptotics of constant-composition codes," 2014.
- [24] A. Dyachkov, "Random constant composition codes for multiple access channels," *Probl. of Control and Inform. Theory*, vol. 13, pp. 357–369, 1984.
- [25] J. Scarlett, A. Martinez, and A. Guillen i Fabregas, "Second-order rate region of constant-composition codes for the multiple-access channel," *Information Theory, IEEE Transactions on*, vol. 61, no. 1, pp. 157–172, 2015.
- [26] T. Ericson and V. Zinoviev, "Spherical codes generated by binary partitions of symmetric pointsets," *IEEE transactions on information theory*, vol. 41, no. 1, pp. 107–129, 1995.
- [27] O. D. King, "Bounds for dna codes with constant gc-content," *Electron. J. Combin.*, vol. 10, no. 1, p. 33, 2003.
- [28] Y. M. Chee and S. Ling, "Improved lower bounds for constant gc-content dna codes," *Information Theory, IEEE Transactions on*, vol. 54, no. 1, pp. 391–394, 2008.
- [29] W. Chu, C. J. Colbourn, and P. Dukes, "Constructions for permutation codes in powerline communications," *Designs, Codes and Cryptography*, vol. 32, no. 1-3, pp. 51–64, 2004.
- [30] C. J. Colbourn, T. Klove, and A. C. Ling, "Permutation arrays for powerline communication and mutually orthogonal latin squares," *Information Theory, IEEE Transactions on*, vol. 50, no. 6, pp. 1289–1291, 2004.
- [31] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.
- [32] M. Svanström, "Ternary codes with weight constraints," Ph.D. dissertation, 1999.
- [33] D. H. Smith, L. A. Hughes, and S. Perkins, "A new table of constant weight codes of length greater than 28," *Electron. J. Combin.*, vol. 13, no. 1, pp. Article 2, 18 pp. (electronic), 2006.
- [34] R. J. R. Abel and M. Buratti, "Difference families," *CRC Handbook of Combinatorial Designs*, CRC Press, Boca Raton, FL, pp. 392–410, 2007.
- [35] D. R. Stinson, R. Wei, and J. Yin, "Packing," in *The CRC Handbook of Combinatorial Designs*, 2nd ed., C. J. Colbourn and J. H. Dinitz, Eds. Boca Raton: CRC Press, 2007, pp. 550–556.
- [36] T. Skolem, "Note 16," in *E. Netto. Kombinatorik*, 2nd ed., A. Teubner, Ed., Leipzig, 1927.
- [37] J. Dinitz and P. Rodney, "Disjoint difference families with block size 3," *UTILITAS MATHEMATICA*, pp. 153–160, 1997.