Optimal Constant-Weight Codes under $l_1$-Metric

Tingting Chen, Yiming Ma, and Xiande Zhang

Abstract

Motivated by the duplication-correcting problem for data storage in live DNA, we study the construction of constant-weight codes in $l_1$-metric. By combinatorial methods, we give constructions of optimal ternary codes with $l_1$-weight $w \leq 4$ for all possible distances. In general, we determine the maximum size of ternary codes with constant weight $w$ and distance $2w - 2$ for sufficiently large length under certain conditions.

Index Terms

DNA storage, tandem duplication error, constant-weight code, $l_1$-metric, packing, group divisible design.

I. INTRODUCTION

CONSTANT-WEIGHT codes (CWCs) with Hamming distance constraint have attracted a lot of attention in recent years due to their vast applications, such as in coding for bandwidth-efficient channels [1] and the design of oligonucleotide sequences for DNA computing [2], [3]. One of the central problems in their study is to determine the maximum size of CWCs due to their close relations to combinatorial design theory, see for example [4]–[12]. Although there are several different metrics which have been considered in coding theory, to the best of our knowledge, there is little known for CWCs in the literature besides Hamming distance.

In this paper, we initiate the study of CWCs with $l_1$-metric, which is motivated from the error correcting problem of data storage in live DNA [13]. To prove the reliability of information stored in live DNA, codes which can correct errors such as tandem duplication, point mutations, insertions, and deletions arising from various mutations, must be considered. Among these, duplication-correcting codes have been studied by a number of recent works, see [13]–[17]. In [13], the authors studied tandem duplication, which is a process of inserting a copy of a segment of the DNA adjacent to its original position. For example, for a sequence $AGCTCT$, $CTCT$ is a tandem duplication error of length two on $CT$. Tandem duplications constitute about 3% of the human genome [18] and may cause important phenomena such as chromosome fragility, expansion diseases, silencing genes [19], and rapid morphological variation [20]. Jain et al. [13] proposed a coding scheme to combat tandem duplications, which is based on CWCs in $l_1$-metric over integers. More specifically, there is a code correcting tandem duplications if and only if there exist CWCs in $l_1$-metric with certain weight and distance.

Codes in $l_1$-metric distance have been widely studied because of their applications in rank-modulation scheme for flash memory [21]–[26]. However, most works focus on permutation codes or multi-permutation codes. Kovačević and Tan [27] gave some properties and constructions of multiset code, based on Sidon sets and lattices, and derived bounds on the size of such codes. The multiset codes they studied are exactly the constant weight codes in $l_1$ metric over integers. Jinushi and Sakaniwa [28] proposed a construction for error-correcting codes in $l_1$-metric which relies upon the properties of generalized Hadamard matrices [29]. They used the term absolute summation distance that we believe is constant weight in $l_1$-metric (Unfortunately, we could not find a copy of their paper to confirm it).

In this work, we construct CWCs in $l_1$-metric by combinatorial methods, and determine the maximum number of codewords in CWCs of weight $w \leq 4$ for all possible distance $d$. For $w = 1$ or 2, codes are trivial. Our main contributions are listed below.

1) For CWCs of weight $w \in \{3, 4\}$ over non-negative integers, we determine the maximum number of codewords completely for all distance $d$. Some codes are constructed by optimal packings of triples by pairs [30], or by optimal packings of quadruples by triples [31].

2) For CWCs over $\{0, 1, 2\}$, that is, ternary CWCs, we construct non-trivial codes with maximum size by using Steiner triple systems [32] and packings with special leave graphs [33], [34], and solve the case completely for $w = 3$. For $w = 4$, the maximum sizes of codes are determined based on group divisible designs [35], but with very few cases unsolved, for which we provide upper and lower bounds.

3) For ternary CWCs with $d = 2w - 2$, we give a general construction using a result on graph packings of Alon et al. [36], and determine the maximum size under certain conditions.

Our paper is organized as follows. In Section II we give necessary definitions, notations and results in combinatorial design theory. In Section III we construct optimal CWCs over integers for weight three and four. In Sections IV and V we consider...
ternary CWCs, and give combinatorial constructions for optimal codes for weight three and four, respectively. In Section VI we deal with ternary codes for general weight $w$ and distance $2w - 2$ by graph packings. Finally, we conclude our results in Section VII.

II. Preliminaries

Let $\mathbb{Z}_{\geq 0}$ denote the set of non-negative integers and $\mathbb{Z}_q$ denote the ring of integers modulo $q$, for any integer $q \geq 2$. Suppose that $X$ and $Y$ are two sets with $|X| = n$, then $Y^X$ means the set of all vectors of length $n$, where a vector $c \in Y^X$ is denoted by $(c_x)_{x \in X}$ with $c_x \in Y$. For any integers $a < b$, let $[a, b]$ denote the set of integers $\{a, a+1, \ldots, b\}$. We further abbreviate $[1, b]$ as $[b]$.

A. CWCs with $l_1$-metric

A $q$-ary code $C$ of length $n$ is a set of vectors in $I_q^n$, where $I_q := \{0, 1, \ldots, q-1\} \subseteq \mathbb{Z}_{\geq 0}$ and $X$ is a set of size $n$. The elements of $C$ are called codewords. For any two codewords $u = (u_x)_{x \in X}, v = (v_x)_{x \in X} \in C$, the support of $u$ is defined as $\text{supp}(u) = \{x \in X \mid u_x \neq 0\}$, and the $l_1$-distance between $u$ and $v$ is defined as $d_{l_1}(u, v) = \sum_{x \in X} |u_x - v_x|$ (computations are over the ring of integers). The $l_1$-weight of $u$ is defined as the $l_1$-distance of $u$ and the zero vector, i.e. $w_{l_1}(u) = \sum_{x \in X} |u_x|$.

A code $C$ is said to be of constant-weight $w$ if $w_{l_1}(u) = w$ for every codeword $u \in C$, and of minimum $l_1$-distance $d$ if $d_{l_1}(u, v) \geq d$ for any two distinct codewords $u, v \in C$. If both properties are satisfied, then a code is called a constant-weight code in $l_1$-metric and denoted by $(n, d, w)_q$ code if it is $q$-ary, and an $(n, d, w)$ code if it is over $\mathbb{Z}_{\geq 0}$.

In [13], Jain et al. established a connection between codes capable of correcting tandem duplications and constant-weight codes with $l_1$-metric. Consider a string $x = 0^{m_0}w_10^{m_1}w_2 \cdots w_{l_1}0^{m_{l_1}} \in \mathbb{Z}_q^{n}$, where $0^{m_i}$ denotes $m_i$ consecutive zeros, and $w_i \in \mathbb{Z}_q$ is a single non-zero symbol. Clearly, $n = \sum m_i + t$. Given a non-negative integer $k$, define the zero signature of $x$ by $\sigma_k(x) = \left(\left\lfloor \frac{m_1}{k}\right\rfloor, \left\lfloor \frac{m_2}{k}\right\rfloor, \ldots, \left\lfloor \frac{m_l}{k}\right\rfloor\right)$ and let $\mu_k(x) = 0^{m_0 \mod k}w_10^{m_1 \mod k}w_2 \cdots w_{l_1}0^{m_{l_1} \mod k}$. Further, define a mapping of $x$ into a string of the same length by $\phi_k(x) = (y, z)$, where $y$ is the prefix of $x$ of length $k$, and $z$ equals the subtraction of the suffix of $x$ of length $n-k$ and the prefix of $x$ of length of $n-k$. Given two strings $x, x'$ with $\phi_k(x) = (y, z)$ and $\phi_k(x') = (y', z')$, we say $x$ and $x'$ are $k$-congruent if $y = y'$ and $\mu_k(z) = \mu_k(z')$. It was shown that a code is able to correct tandem duplications of length $k$ if and only if the zero signatures of the $z$-part of all $k$-congruent codewords form a constant-weight code in $l_1$-metric over integers [13] Theorem 20. More importantly, a choice of optimal $l_1$-metric constant-weight codes will result in optimal tandem duplication correcting codes [13 Construction B]. Further if we assume that each segment of length $k$ is duplicated at most $q - 1$ times, then the zero signatures are $q$-ary vectors, and hence we need $q$-ary constant-weight codes in $l_1$-metric.

Motivated by this connection, we consider constant-weight codes in $l_1$-metric with maximum possible size. Since we only consider $l_1$-metric in this paper, we omit the subscript $l_1$ or the term $l_1$-metric unless otherwise specified. The maximum number of codewords in an $(n, d, w)_q$ code is denoted by $A_q(n, d, w)$, and the $(n, d, w)_q$ code is called optimal if it has $A_q(n, d, w)$ codewords. Similarly for codes over $\mathbb{Z}_{\geq 0}$, we use $A(n, d, w)$ to denote the largest possible number of codewords.

In the remaining of this paper, we focus on determining the values of $A_q(n, d, w)$ and $A(n, d, w)$ by constructing optimal CWCs. The followings are some trivial cases.

**Theorem II.1.** (a) $A_q(n, 2\delta - 1, w) = A_q(n, 2\delta, w); A(n, 2\delta - 1, w) = A(n, 2\delta, w)$.
(b) $A_q(n, 2\delta, w) = A(n, 2\delta, w) = 1$ if $w < \delta$.
(c) $A_q(n, 2w, w) = \left\lfloor \frac{n}{2w-1}\right\rfloor \cdot \frac{n}{w}; A(n, 2w, w) = n$.
(d) $A_q(n, 2, w) = A(n, 2, w) = \binom{n+w-1}{w}$ if $w \leq q - 1; A_q(n, 2, w) = \sum_{t=0}^{\left\lfloor \frac{w}{q}\right\rfloor}(-1)^t\binom{n}{w}^{(t)}\binom{n-w-q}{w-t}$ if $w > q - 1$, where $t = \left\lfloor \frac{w}{q}\right\rfloor$.

*Proof.* The equalities in (a) follow because the $l_1$-distance between any two codewords of constant weight is even. The results in (b) are obvious, and results in (c) follow because the codewords must have disjoint supports and the minimum supports of codewords over $I_q$ and $\mathbb{Z}_{\geq 0}$ have sizes $\left\lfloor \frac{n}{2w-1}\right\rfloor$ and one, respectively.

For (d), if $w \leq q - 1$, any entry of the codeword is at most $w$, then $A_q(n, 2, w) = A(n, 2, w)$. Consider the following generating function

$$(1 + x + \cdots + x^k)^n = \sum_{j=0}^{kn} a_j x^j.$$

If $w \leq q - 1$, let $k = w$, then $A(n, 2, w) = a_w = \binom{n+w-1}{w}$; if $w > q - 1$, let $k = q - 1$, then $A_q(n, 2, w) = a_w = \sum_{j=0}^{t}(-1)^t\binom{n}{w}^{(t)}\binom{n-w-q}{w-t}$, where $t$ satisfies $w = qt + r$ and $0 \leq r < q$. 

By Theorem II.1 we only need to consider the even distance between 4 and $2w - 2$ for any code of constant weight $w$. 


B. Designs

A set system is a pair $S = (X, B)$, where $X$ is a finite set of points and $B$ is a set of subsets of $X$, called blocks. The order of $S$ is the number of points $|X|$, and the size of $S$ is the number of blocks $|B|$.

A graph $G$ is a set system $(V, E)$ with all blocks in $E$ being 2-subsets of $V$, in which a point of $V$ is called a vertex and a block of $E$ is called an edge. The degree of a vertex $v$ is the number of edges containing $v$. The minimum vertex degree of $G$ is the smallest vertex degree of $G$, denoted by $\delta(G)$. And the number of edges of $G$ is denoted by $e(G)$. A graph is called a complete graph if each pair of vertices is connected by an edge, and denoted by $K_n$ if $|V| = n$. We call a sequence $(v_1, v_2, \ldots, v_m)$ of distinct vertices a cycle of length $m$ if $(v_i, v_{i+1}) \in E$ for all $i \in [m-1]$ and $\{v_m, v_1\} \in E$.

There is a canonical one-to-one correspondence between all vectors $u$ in $I_2^n$ and all subsets $\supp(u)$ of $X$, so a binary code $C \subset I_2^n$ corresponds to a set system $(X, \{\supp(u) : u \in C\})$. For a codeword $u = (u_x)_{x \in X}$ which is $q$-ary, we associate it with a subset $\phi(u) := \{x_i : x_i \in \supp(u) \text{ and } u_x = i\} \subset X \times \{q-1\}$ to indicate different nonzero entries. By abuse of notation we sometimes do not distinguish between $u$ and $\phi(u)$.

Example II.1. Let $X = \mathbb{Z}_4$, we have a $(4, 4, 3)_3$ code $C \subset I_3^X$ with four codewords $1200, 0120, 0012$ and $2001$. Equivalently, we can describe them as $\{01, 12\}, \{11, 22\}, \{21, 32\}, \{31, 02\}$, which are subsets of $X \times \{2\}$. It is easy to see that $C$ is obtained by a base codeword $\{1, 1\}_2$ by a group action $\mathbb{Z}_4$ on $X$.

Let $K$ be a set of positive integers. A $t$-$(n, K, 1)$ packing is a set system $(X, B)$ of order $n$, such that $|B| \in K$ for each $B \in B$, and every $t$-subset of points occurs in at most one block of $B$. When $K = \{k\}$, we just write $k$ instead of $\{k\}$. The packing number $D(n, k, t)$ is the maximum number of blocks in any $t$-$(n, k, 1)$ packing. A $t$-$(n, k, 1)$ packing $(X, B)$ is optimal if $|B| = D(n, k, t)$. If every $t$-subset of points occurs in exactly one block, we call it a $t$-$(n, k, 1)$ design, or a $t$-design in short. The leave graph of a $t$-$(n, k, 1)$ packing is the set system $(X, E)$ where $E$ consists of all $t$-subsets of $X$ that do not appear in any blocks. For $t = 2$ and $k = 3, 4$, or $t = 3$ and $k = 4$, the packing numbers have been completely determined, see [30], and the corresponding leave graphs are also characterized. We list them below for later use.

Lemma II.1. For any positive $n \not\equiv 5 \pmod 6$, $D(n, 3, 2) = \left\lfloor \frac{n+1}{3} \right\rfloor$; when $n \equiv 5 \pmod 6$, $D(n, 3, 2) = \left\lfloor \frac{n-1}{3} \right\rfloor - 1$, and the leave graph is a cycle of length four.

Lemma II.2. for positive integers $n \not\in \{8, 9, 10, 11, 17, 19\}$, we have

$$D(n, 4, 2) = \begin{cases} \left\lfloor \frac{n}{4} \left\lfloor \frac{n+1}{3} \right\rfloor \right\rfloor, & n \equiv 7, 10 \pmod {12}, \\ \left\lfloor \frac{n}{4} \left\lfloor \frac{n+2}{3} \right\rfloor \right\rfloor, & \text{otherwise}. \end{cases}$$

When $n = 8, 9, 10, 11, 17, 19$, the values of $D(n, 4, 2)$ are equal to $2, 3, 5, 6, 20, 25$, respectively.

Lemma II.3. For any positive $n$, it holds that

$$D(n, 4, 3) = \begin{cases} \left\lfloor \frac{n}{4} \left\lfloor \frac{n-1}{3} \right\rfloor \right\rfloor, & \text{if } n \not\equiv 0 \pmod 6, \\ \left\lfloor \frac{n}{4} \left(\left\lfloor \frac{n+2}{3} \right\rfloor - 1\right) \right\rfloor, & \text{if } n \equiv 0 \pmod 6. \end{cases}$$

A group divisible design (K-GDD) is a triple $(X, \mathcal{G}, B)$, where $X$ is a finite set of size $n$, $\mathcal{G}$ is a partition of $X$ into subsets, called groups, and $B$ is a set of blocks of $X$ which satisfies (1) if $B \in B$ then $|B| \in K$, (2) every pair of $X$ not contained in a group appears in exactly one block, and (3) every pair contained in a group does not appear in any block. If $\mathcal{G} = \{G_1, G_2, \ldots, G_t\}$, then the type of the GDD is the multiset $\{\#G_i : i = 1, 2, \ldots, t\}$, and the exponential notation for the type is $g_1^{a_1}g_2^{a_2} \cdots g_s^{a_s}$ if there are $a_i$ groups of size $g_i, i = 1, 2, \ldots, s$. If each group is of size one, it is a K-GDD of type 1$^t$, then we call it a pairwise balanced design (a K-PBD of order $n$ or an $(n, K)$-PBD).

Lemma II.4. The necessary and sufficient conditions for the existence of a 3-GDD of type $3^u$ with $u \geq 3$ is $u \equiv 1 \pmod 2$.

Lemma II.5. Let $u \geq 4$ and $m \geq 0$. For each $g \in \{2, 3, 6, 7, 9, 12, 15, 16, 24, 27, 36\}$, there exists a 4-GDD of type $g^u m^1$ if and only if $m \leq g/(u-1)/2, gu \equiv 0 \pmod 3, g(u-1)+m \equiv 0 \pmod 3$ and $(9u+m)2 - u(9^u - 1)^2 \equiv 0 \pmod 6$ except possibly for $(g, u, m) \in \{(2, 33, 23), (2, 33, 29), (2, 39, 35), (6, 13, 27), (6, 13, 33), (6, 17, 39), (6, 19, 45), (6, 19, 51), (6, 23, 63)\}$.

III. CWCS over Integers

In this section, we consider the code with constant weight $w \leq 4$ over $\mathbb{Z}_{\geq 0}$ and determine the value of $A(n, d, w)$ for $w = 3, 4$ and all distances $d$ between 4 and $2w - 2$. For convenience, we call a codeword is of type $X$ if all its non-zero elements are 1, type $Y$, if it has exactly one position with entry $t$ and the rest non-zero elements are 1, and type $Z$, if all its non-zero elements are $t$. 
For weight three, there are three types of codewords, \( \{a_1, b_1, c_1\} \) of type \( X \), \( \{d_1, e_2\} \) of type \( Y_2 \), and \( \{f_3\} \) of type \( Z_3 \), respectively. Since \( w = 3 \), we only need to consider \( d = 4 \). It is easy to check that a code \( C \subset \mathbb{Z}_n \times \mathbb{Z}_{\geq 0} \) consisting of codewords of types \( X \), \( Y_2 \), and \( Z_3 \) is an \( (n, 4, 3) \) code if and only if the following conditions are satisfied:

1. The collection of subsets \( \text{supp}(u) \subset \mathbb{Z}_n \), for all codewords \( u \) of types \( X \) or \( Y_2 \), forms a 2-\( (n, \{2, 3\}, 1) \) packing.
2. For any two codewords \( u, v \) of types \( Y_2 \) or \( Z_3 \), if \( x_i \in u \) and \( x_j \in v \), then \( 1 \in \{i, j\} \).

**Theorem III.1.** \( A(n, 4, 3) = D(n, 3, 2) + n \).

**Proof.** Let \( x, y \), and \( z \) be the number of codewords of types \( X \), \( Y_2 \) and \( Z_3 \), respectively. By property (1), we have

\[
x \leq D(n, 3, 2),
\]

since codewords of type \( X \) form an 2-\( (n, 3, 1) \) packing. By property (2), we have

\[
y + z \leq n
\]

by counting the occurrences of symbols 2 and 3 in all codewords. So the upper bound \( A(n, 4, 3) = x + y + z \leq D(n, 3, 2) + n \) follows. The lower bound is achieved by the code consisting of all binary codewords of type \( X \) obtained from an optimal 2-\( (n, 3, 1) \) packing, and \( n \) codewords of type \( Z_3 \) with disjoint supports. \( \square \)

For weight four, there are five types of codewords, \( \{a_1, b_1, c_1, d_1\} \) of type \( X \), \( \{e_1, f_1, g_2\} \) of type \( Y_2 \), \( \{h_1, i_3\} \) of type \( Y_3 \), \( \{j_2, k_2\} \) of type \( Z_2 \), and \( \{l_i\} \) of type \( Z_4 \), respectively.

We first consider distance \( d = 4 \). A code \( C \subset \mathbb{Z}_n \times \mathbb{Z}_{\geq 0} \) consisting of codewords of constant weight four is an \( (n, 4, 4) \) code if and only if the following conditions are satisfied:

1. The collection of subsets \( \text{supp}(u) \subset \mathbb{Z}_n \), for all codewords \( u \) of types \( X \) or \( Y_2 \), forms a 3-\( (n, \{3, 4\}, 1) \) packing.
2. All ordered pairs \((i, j)\) satisfying \( \{i, j\} \subset u \) for some codeword \( u \in C \) and \( s' \geq 2 \) are different.
3. For any two codewords \( u, v \), if \( x_i \in u \) and \( x_j \in v \), then \( \{1, 2\} \cap \{i, j\} \neq \emptyset \).

**Theorem III.2.** \( A(n, 4, 4) = D(n, 4, 3) + \frac{n(n-1)}{2} + n \).

**Proof.** Let \( x, y, z, a \), and \( b \) be the number of codewords of types \( X \), \( Y_2 \), \( Z_2 \), \( Y_3 \) and \( Z_4 \), respectively. By properties (i) and (iii), we have

\[
x \leq D(n, 4, 3), \text{ and } a + b \leq n.
\]

From property (ii), we have

\[
2y + 2z + a \leq n(n-1).
\]

Combining the above inequalities, we have that \( A(n, 4, 4) = x + y + z + a + b \leq D(n, 4, 3) + \frac{n(n-1)}{2} + n \). The lower bound is achieved by the code consisting of all binary codewords of type \( X \) obtained from an optimal 3-\( (n, 4, 1) \) packing over \( \mathbb{Z}_n \), \( \frac{n(n-1)}{2} \) different codewords of type \( Z_2 \) and \( n \) codewords of type \( Z_4 \). \( \square \)

For distance \( d = 6 \), a code \( C \subset \mathbb{Z}_n \times \mathbb{Z}_{\geq 0} \) is an \( (n, 6, 4) \) code if and only if the following conditions are satisfied:

1. The collection of subsets \( \text{supp}(u) \subset \mathbb{Z}_n \), for all codewords \( u \) of types \( X \), \( Y_2 \), \( Y_3 \), or \( Z_2 \), forms a 2-\( (n, \{2, 3\}, 1) \) packing.
2. For any two codewords \( u, v \), if \( x_i \in u \) and \( x_j \in v \), then \( 1 \in \{i, j\} \).

**Theorem III.3.** \( A(n, 6, 4) = D(n, 4, 2) + n \).

**Proof.** Using the same notation as in the proof of Theorem III.2, we have

\[
x \leq D(n, 4, 2), \text{ and } y + 2y + a + b \leq n
\]

by conditions (a) and (b). Then \( A(n, 6, 4) = x + y + z + a + b \leq D(n, 4, 2) + n \). The lower bound is achieved by the code consisting of all binary codewords of type \( X \) obtained from an optimal 2-\( (n, 4, 1) \) packing over \( \mathbb{Z}_n \) and \( n \) codewords of type \( Z_4 \). \( \square \)

### IV. Ternary CWCS of Weight Three

In this section, we consider ternary codes with constant weight three in \( I^3_n \) and determine the value of \( A_3(n, d, w) \) for \( d = 4 \). Since the code is ternary, there are only two types of codewords, \( \{a_1, b_1, c_1\} \) of type \( X \), and \( \{d_1, e_2\} \) of type \( Y_2 \), respectively. Then a code \( C \subset \mathbb{Z}_n \times [2] \) is an \( (n, 4, 3) \) code if and only if the following conditions are satisfied:

1. The collection of subsets \( \text{supp}(u) \subset \mathbb{Z}_n \), for all codewords \( u \) of types \( X \) or \( Y_2 \), forms a 2-\( (n, \{2, 3\}, 1) \) packing.
2. For any two codewords \( u, v \) of type \( Y_2 \), if \( x_i \in u \) and \( x_j \in v \), then \( 1 \in \{i, j\} \).

**Lemma IV.1.** \( A_3(n, 4, 3) \leq \left\lfloor \frac{n^2 + 3n}{6} \right\rfloor \).
Proof. Let $x$ and $y$ be the number of codewords of types $\mathcal{X}$ and $\mathcal{Y}_2$ respectively. By properties (1') and (2'), we have
\[ 3x + y \leq \binom{n}{2}, \quad \text{and} \quad y \leq n. \]
Then $A_3(n, 4, 3) = x + y \leq \left\lfloor \frac{n^2 + 3n}{6} \right\rfloor$.

By the proof of Lemma [IV.1] it is possible for a code to achieve the upper bound when $y = n$ or $n - 1$. Assume we have already found $n$ or $n - 1$ codewords of type $\mathcal{Y}_2$, then we need to find a $2-(n, 3, 1)$ packing such that property (1) is satisfied. If the size of this $2-(n, 3, 1)$ packing is $\left\lfloor \frac{n^2 + 3n}{6} \right\rfloor - n$ or $\left\lfloor \frac{n^2 + 3n}{6} \right\rfloor - n + 1$, respectively, then the upper bound in Lemma [IV.1] can be achieved.

Theorem IV.1. $A_3(n, 4, 3) = \left\lfloor \frac{n^2 + 3n}{6} \right\rfloor$.

Proof. The upper bound is from Lemma [IV.1]. For the lower bound, the case $n \leq 3$ is easy to check; for all other integers $n$, we construct an $(n, 4, 3)$ code $\mathcal{C}$ of size achieving the upper bound as follows.

For each $n = 1, 5 \ (\text{mod} \ 6)$, there exists a $2-(n, 3, 1)$ packing $(X, \mathcal{B})$ of size $\frac{n^2 - 3n + 2}{6}$ whose leave graph consists of a cycle of length $n - 1$ and one isolated vertex $\left\lfloor \frac{n}{6} \right\rfloor$. We assume that $X = \mathbb{Z}_{n-1} \cup \{\infty\}$, and the cycle is $(0, 1, 2, \ldots, n - 2)$. Then the desired $\mathcal{C}$ consists of all type $\mathcal{X}$ codewords $\{a_1, b_1, c_1\}$ for each $\{a, b, c\} \in \mathcal{B}$, and $(n - 1)$ type $\mathcal{Y}_2$ codewords $\{0_1, 1_2\}, \{1_1, 2_2\}, \ldots, \{(n - 3)_1, (n - 2)_2\}, \{(n - 2)_1, 0_2\}$.

For each $n = 2, 4 \ (\text{mod} \ 6)$, there exists a $2-(n - 1, 3, 1)$ packing $(Z_{n-1}, \mathcal{B})$ of size $\frac{(n - 1)(n - 2)}{6}$, see Lemma [II.1], which is indeed a Steiner triple system of order $n - 1$, or $\text{STS}(n - 1)$. The desired code is constructed over $\mathbb{Z}_{n-1} \cup \{\infty\}$, which consists of all type $\mathcal{X}$ codewords $\{a_1, b_1, c_1\}$ for each $\{a, b, c\} \in \mathcal{B}$, and $(n - 1)$ type $\mathcal{Y}_2$ codewords $\{0_1, 1_2\}, \{1_1, 2_2\}, \ldots, \{(n - 2)_1, 0_2\}$.

When $n = 3 \ (\text{mod} \ 6)$, Colbourn and Rosa [33] (and Colbourn and Ling [34]) showed that there exists a $2-(3, 3, 1)$ packing $(Z_{n-1}, \mathcal{B})$ of size $\frac{n^2 - 3n}{6}$ whose leave graph consists of all pairs $\{i, j\}$ with $i - j \equiv 1 \ (\text{mod} \ n)$. Then the desired $\mathcal{C}$ consists of all type $\mathcal{X}$ codewords $\{a_1, b_1, c_1\}$ for each $\{a, b, c\} \in \mathcal{B}$, and $n$ type $\mathcal{Y}_2$ codewords $\{0_1, 1_2\}, \{1_1, 2_2\}, \ldots, \{(n - 1)_1, 0_2\}$.

When $n = 0 \ (\text{mod} \ 6)$, there is a $2-(n - 1, 3, 1)$ packing $(Z_{n-1}, \mathcal{B})$ whose leave graph is a cycle of length four by Lemma [II.1]. Assume the cycle is $(0, 1, 2, 3)$. The desired code $\mathcal{C}$ is constructed over $\mathbb{Z}_{n-1} \cup \{\infty\}$, which consists of all type $\mathcal{X}$ codewords $\{a_1, b_1, c_1\}$ for each $\{a, b, c\} \in \mathcal{B}$, an additional codeword $\{1_1, 2_1, 3_1\}$ of type $\mathcal{X}$, five codewords of type $\mathcal{Y}_2$: $\{0_2, 3_1\}, \{0_1, 2_2\}, \{2_2, 3_1\}, \{0_1, 3_2\}, \{3_1, 3_2\}$, and $(n - 5)$ type $\mathcal{Y}_2$ codewords $\{4_2, 3_1\}, \{5_2, 3_1\}, \ldots, \{(n - 3)_2, 3_1\}, \{(n - 2)_2, 3_1\}$.

It is routine to check that all codes constructed above are $(n, 4, 3)$ codes of the required sizes.

Next, we give examples of constructions in Theorem [IV.1].

Example IV.1. For $n = 6$, $A_3(6, 4, 3) = 9$. There exists a $2-(5, 3, 1)$ packing over $\mathbb{Z}_5$ with blocks $024, 134$ whose leave graph is a cycle $(0, 1, 2, 3)$. We adjoin an infinity point $\infty$, and construct an optimal code with the following codewords:

\[
\begin{align*}
\{0_1, 2_1, 4_1\} &\quad \{1_1, 3_1, 4_1\} &\quad \{1_1, 2_1, 3_1\} \\
\{0_2, 2_1, \infty\} &\quad \{0_1, 1_2\} &\quad \{2_1, 3_1\} \\
\{0_1, 3_2\} &\quad \{3_1, \infty\} &\quad \{4_2, \infty\}.
\end{align*}
\]

For $n = 7$, we construct an optimal $(7, 4, 3)_3$ code of size eleven as follows. Since there is a $2-(7, 3, 1)$ packing with five blocks $14\infty, 25\infty, 03\infty, 135, 024$ over $\mathbb{Z}_6 \cup \{\infty\}$ whose leave graph consists of a cycle $(0, 1, 2, 3, 4, 5)$, then the codewords are as follows:

\[
\begin{align*}
\{1_1, 4_1, \infty\} &\quad \{2_1, 5_1, \infty\} &\quad \{0_1, 3_1, \infty\} &\quad \{1_1, 3_1, 5_1\} \\
\{0_1, 2_1, 4_1\} &\quad \{0_1, 1_2\} &\quad \{1_1, 2_1\} &\quad \{2_1, 3_2\} \\
\{3_1, 4_2\} &\quad \{0_1, 5_2\} &\quad \{0_1, \infty\}.
\end{align*}
\]

For $n = 8$, we construct an optimal $(8, 4, 3)_3$ code of size fourteen as follows. From an $\text{STS}(7)$ over $\mathbb{Z}_7$ with blocks $124, 235, 346, 045, 156, 026, 013$, adjoining an infinite point $\infty$, we obtain the codewords as follows:

\[
\begin{align*}
\{1_1, 2_1, 4_1\} &\quad \{2_1, 3_1, 5_1\} &\quad \{3_1, 4_1, 6_1\} &\quad \{0_1, 4_1, 5_1\} \\
\{1_1, 5_1, 6_1\} &\quad \{0_1, 2_1, 6_1\} &\quad \{0_1, 1_2, 3_1\} &\quad \{0_2, \infty\} \\
\{1_2, \infty\} &\quad \{2_2, \infty\} &\quad \{3_2, \infty\} &\quad \{4_2, \infty\} \\
\{5_2, \infty\} &\quad \{6_2, \infty\}.
\end{align*}
\]

For $n = 9$, $A_3(9, 4, 3) = 18$. The blocks generated by 035 under $\mathbb{Z}_3$ is a 2-(9, 3, 1) packing, whose leave graph is a cycle of length nine. Then the codewords are as follows:

\[
\begin{align*}
\{0, 1, 3, 5\} & \quad \{1, 1, 4, 6\} & \quad \{2, 1, 5, 7\} & \quad \{3, 1, 6, 8\} \\
\{4, 1, 7, 0\} & \quad \{5, 1, 8, 1\} & \quad \{6, 1, 0, 2\} & \quad \{7, 1, 1, 3\} \\
\{8, 1, 2, 4\} & \quad \{0, 1, 2\} & \quad \{1, 1, 2\} & \quad \{2, 1, 3\} \\
\{3, 4\} & \quad \{4, 5\} & \quad \{5, 6\} & \quad \{6, 7\} \\
\{7, 8\} & \quad \{8, 0\}.
\end{align*}
\]

**Remark IV.1.** For $n \equiv 3 \pmod{6}$, we give another construction of optimal codes as follows. Let $u = n/3$, then there exists a 3-GDD of type $3^u$ by Lemma [14] say $(X, G, B)$. For each group $G = \{a, b, c\} \in G$, we obtain three type $\mathcal{Y}_2$ codewords: \{a_1, b_2\}, \{b_1, c_2\} and \{c_1, a_2\}. It is easy to check that all these $n$ codewords of type $\mathcal{Y}_2$, combining all type $\mathcal{X}$ codewords obtained from $B$, form an optimal $(n, 4, 3)_3$ code.

V. TERNARY CWCS OF WEIGHT FOUR

In this section, we consider ternary codes with constant weight four and determine the value of $A_3(n, d, 4)$ for $d = 4$ and 6. Since the code is ternary, there are three types of codewords, \{a_1, b_1, c_1, c_d\} of type $\mathcal{X}$, \{a_1, f_1, g_2\} of type $\mathcal{Y}_2$, and \{b_2, i_2\} of type $\mathcal{Z}_2$.

When $d = 4$, the only difference between the CWC of weight four over $\mathbb{Z}_{\geq 0}$ and $I_3$ is that the latter one does not have codewords of types $\mathcal{Y}_3$ and $\mathcal{Z}_4$. Then by similar arguments as in the proof of Theorem [III.2] we can obtain the following result.

**Theorem V.1.** $A_3(n, 4, 4) = D(n, 4, 3) + \frac{n(n-1)}{2}$.

Now we consider $d = 6$. A ternary constant-weight code $C$ is an $(n, 6, 4)_3$ code if and only if $C$ satisfies the following properties.

(a') The collection of subsets $\text{supp}(u) \subset \mathbb{Z}_n$, for all codewords $u$ of types $\mathcal{X}$, $\mathcal{Y}_2$ or $\mathcal{Z}_2$, forms a 2-(n, {2, 3, 4}, 1) packing.

(b') For any two codewords $u, v \in C$, if $x_i \in u$ and $x_j \in v$, then $1 \in \{i, j\}$.

**Lemma V.1.** $A_3(n, 6, 4) \leq \left\lfloor \frac{n(n+5)}{12} \right\rfloor$.

**Proof.** Let $x, y, z$ be the number of codewords of types $\mathcal{X}$, $\mathcal{Y}_2$ and $\mathcal{Z}_2$, respectively. By properties (a') and (b'), we have

\[6x + 3y + z \leq \binom{n}{2}, \quad y + 2z \leq n.\]

Then we get $A_3(n, 6, 4) = x + y + z \leq \left\lfloor \frac{n(n+5)}{12} \right\rfloor$.

For convenience, let

\[U(n) :=\left\lfloor \frac{n(n+5)}{12} \right\rfloor.\]

Then it is easy to see that $A_3(1, 6, 4) = 0 = U(1)$, $A_3(2, 6, 4) = 1 = U(2)$, $A_3(3, 6, 4) = 1 = U(3) - 1$ and $A_3(4, 6, 4) = 2 = U(4) - 1$. For $n = 5$, there is no $(5, 6, 4)_3$ code of size four by exhaustive search. Then $A_3(5, 6, 4) = 3 = U(5) - 1$ since 21100, 10012, 02020 form a $(5, 6, 4)_3$ code.

**Lemma V.2.** For all $n \in [6, 11]$, we have $A_3(n, 6, 4) = U(n)$.

**Proof.** For each $n \in [6, 11]$, we construct an optimal code as follows.

For $n = 6$, $A_3(6, 6, 4) \leq 5$. From an STS(7) over $\mathbb{Z}_7$ with blocks 124, 235, 346, 045, 156, 026, 013, delete the point 6 and all blocks containing it. Use the remaining blocks to get four codewords of type $\mathcal{Y}_2$ by choosing different positions \{0, 1, 2, 5\} for the symbol 2. Then add a type $\mathcal{Z}_2$ codeword by using the pair \{3, 4\}. The final code is listed as follows.

\[
\begin{align*}
\{0, 1, 2, 3\} & \quad \{1, 2, 3, 4\} & \quad \{2, 3, 4, 5\} \\
\{0, 1, 4, 5\} & \quad \{3, 2, 4\}.
\end{align*}
\]

For $n = 7$, $A_3(7, 6, 4) \leq 7$. We use the same STS(7) which indeed has a base block 013 under the group action $\mathbb{Z}_7$. Then the codewords of our code $C$ are generated by \{0, 1, 2, 3\} under the group action $\mathbb{Z}_7$ on the set of coordinates.

For $n = 8$, $A_3(8, 6, 4) \leq 8$. The optimal code $C$ is obtained by developing the codeword \{0, 2, 1, 3\} under the group action $\mathbb{Z}_8$.

For $n = 9$, $A_3(9, 6, 4) \leq 10$. The code $C$ is listed as follows.

\[
\begin{align*}
\{0, 1, 2, 3\} & \quad \{0, 1, 4, 5, 6\} & \quad \{7, 1, 0, 2\} & \quad \{3, 1, 6, 7\} \\
\{5, 1, 7, 2\} & \quad \{4, 2, 1\} & \quad \{1, 7, 6, 2\} & \quad \{3, 8, 1\} \\
\{2, 4, 1\} & \quad \{3, 2, 4\}.
\end{align*}
\]
For \( n = 10 \), \( A_3(10, 6, 4) \leq 12 \). This code \( C \) is given by a \((10, \{3, 4\})\)-PBD by assigning the symbol 2 appropriately.

\[
\begin{align*}
\{0_1, 1_1, 2_1, 3_1\} & \quad \{0_1, 4_1, 5_1, 6_1\} & \quad \{0_1, 7_1, 8_1, 9_1\} & \quad \{2_1, 4_2, 9_1\} \\
\{3_1, 5_1, 9_2\} & \quad \{1_1, 6_2, 9_1\} & \quad \{1_2, 4_1, 7_1\} & \quad \{2_2, 5_1, 7_1\} \\
\{2_1, 6_1, 8_2\} & \quad \{3_2, 4_1, 8_1\} & \quad \{1_1, 5_2, 8_1\} & \quad \{5_1, 6_1, 7_2\}.
\end{align*}
\]

For \( n = 11 \), \( A_3(11, 6, 4) \leq 14 \). The code \( C \) is obtained by computer search.

\[
\begin{align*}
\{0_1, 2_1, 4_1, 6_1\} & \quad \{2_1, 3_1, 7_1, 9_1\} & \quad \{3_1, 6_1, 8_1, 10_1\} & \quad \{0_2, 8_1, 9_1\} \\
\{1_1, 9_2, 10_1\} & \quad \{2_2, 6_1, 10_1\} & \quad \{4_1, 7_1, 10_2\} & \quad \{4_2, 5_1, 9_1\} \\
\{0_1, 3_1, 9_2\} & \quad \{1_1, 5_1, 6_2\} & \quad \{0_1, 1_1, 7_2\} & \quad \{5_1, 7_1, 8_2\} \\
\{1_2, 2_1, 8_1\} & \quad \{1_1, 3_2, 4_1\}.
\end{align*}
\]

\[\framebox{Lemma V.3.} \quad A_3(12, 6, 4) = 16.\]

\[\framebox{Proof.} \quad \text{When } n = 12, U(n) = 17. \text{ By the proof of Lemma V.1 a } (12, 6, 4)_3 \text{ code achieves the size 17 only when } z = 0, x = 5, y = 12, \text{ and } \text{supp}(u) \subseteq \mathbb{Z}_{12}, \text{ for all codewords } u \text{ of types } \chi' \text{ or } \chi_2, \text{ form a } 2\cdot(12, \{3, 4\}, 1) \text{ packing with an empty leave graph. Suppose there exists such a code } C \subseteq \mathbb{Z}_{12} \times \{2\}. \text{ For each } i \in \mathbb{Z}_{12}, \text{ let } x_i \text{ be the number of codewords of type } \chi' \text{ containing } i, \text{ that is having symbol 1 in position } i. \text{ By property } (a'), \text{ we have } x_i \leq \left\lfloor \frac{11}{3} \right\rfloor = 3 \text{ for each } i. \text{ Since we have five codewords of type } \chi', \text{ then } x_0 + x_1 + \cdots + x_{11} = 20.\]

\[\text{Since } \text{supp}(u) \subseteq \mathbb{Z}_{12}, u \in C \text{ form a } 2\cdot(12, \{3, 4\}, 1) \text{ packing with an empty leave graph, the number of pairs containing } i \text{ is } 11 = 3x_i + 2y_i, \text{ where } y_i \text{ is the number of codewords of type } \chi_2 \text{ having a nonzero entry in position } i. \text{ This forces } x_i \text{ to be an odd integer, which might be 1 or 3. Let } d_j \text{ be the number of positions } i \text{ such that } x_i = j \text{ for } j = 1, 3, \text{ then }\]

\[\begin{cases} d_1 + 3d_3 = 20, \\ d_1 + d_3 = 12. \end{cases}\]

Therefore, \( d_1 = 8 \) and \( d_3 = 4 \). Without loss of generality, assume that \( x_0 = 3 \), and we have three codewords \( \{0_1, 1_1, 2_1, 3_1\}, \{0_1, 4_1, 5_1, 6_1\}, \{0_1, 7_1, 8_1, 9_1\} \). Since \( d_3 = 4 \), we can further assume that \( x_1 = 3 \), and we have two more codewords \( \{1_1, 4_1, 7_1, 10_1\}, \{1_1, 5_1, 8_1, 11_1\} \). Since each \( x_3 \) is odd, we need to construct more codewords of type \( \chi' \), but we already have five codewords of type \( \chi \), a contradiction. So we prove that \( A_3(12, 6, 4) \leq 16 \).

In fact, we can construct a \((12, 6, 4)_3 \) code of size 16 as follows.

\[
\begin{align*}
\{3_1, 7_1, 8_1, 11_1\} & \quad \{1_1, 3_1, 4_1, 6_1\} & \quad \{2_1, 6_1, 8_1, 10_1\} & \quad \{0_1, 2_1, 7_1, 9_1\} \\
\{4_1, 11_1, 0_2\} & \quad \{1_1, 2_1, 11_2\} & \quad \{10_1, 11_1, 5_2\} & \quad \{0_1, 10_1, 3_2\} \\
\{7_1, 10_1, 4_2\} & \quad \{6_1, 11_1, 9_2\} & \quad \{1_1, 9_1, 10_2\} & \quad \{5_1, 9_1, 8_2\} \\
\{3_1, 5_1, 2_2\} & \quad \{1_1, 5_1, 7_2\} & \quad \{0_1, 8_1, 12\} & \quad \{0_1, 5_1, 6_2\}.
\end{align*}
\]

\[\framebox{A. Length } n \equiv 0 \pmod{3}\]

Inspired by Remark V.1, we will construct optimal \((n, 6, 4)_3 \) codes by using 4-GDDs and optimal codes of short length: given a 4-GDD, take all its blocks as codewords of type \( \chi' \) in a natural way; for each group of the GDD, say of length \( g \), take an optimal \((g, 6, 4)_3 \) code and then extend it to a code of length \( n \) by assigning zeros to the remaining coordinates. GDDs with appropriate group type and short optimal codes with special structures can give optimal codes of a long length.

\[\framebox{Theorem V.2.} \quad \text{Suppose there exists a 4-GDD of type } g^u m^1, \text{ where } g \equiv 0, 3, 4, 7 \pmod{12}. \text{ If } A_3(g, 6, 4) = U(g) \text{ and } A_3(m, 6, 4) = U(m), \text{ then } A_3(gu + m, 6, 4) = U(gu + m).\]

\[\framebox{Proof.} \quad \text{Let } n = gu + m. \text{ Given a 4-GDD } (X, \mathcal{G}, \mathcal{B}) \text{ of type } g^u m^1, \text{ with } |X| = gu + m, \text{ we construct an } (n, 6, 4)_3 \text{ code } C \subseteq X \times [2] \text{ as follows. For each group } G \in \mathcal{G}, \text{ construct an optimal } (|G|, 6, 4)_3 \text{ code } C_G \subseteq G \times [2], \text{ which exists by assumption. Note that we can view } C_G \text{ as a subset of } X \times [2], \text{ i.e., an } (n, 6, 4)_3 \text{ code by assigning zeros to the remaining coordinates. Let } \mathcal{C}_0 \text{ be the set of codewords of type } \chi' \text{ obtained from all blocks of } \mathcal{B}, \text{ that is } \mathcal{C}_0 = \{\{a_1, b_1, c_1, d_1\} : \{a, b, c, d\} \in \mathcal{B}\}. \text{ Then it is easy to} \]
Lemma V.6. \(A_3(n, 6, 4) = U(n)\) for \(n \equiv \pm 2, 6 (mod\ 12)\) and \(n \equiv 0 (mod\ 3)\).

Proof. Consider the cases \(n \equiv 2, 6 (mod\ 12)\). We will use Theorem V.2 with a 4-GDD of type \(36^m\). Let \(\mathcal{C} = C_{27} \cup \bigcup_{G \in G_2} C_G\) be a \((n, 6, 4)\)-code. For \(n \equiv 2, 6 (mod\ 12)\), \(\mathcal{C}\) is an \((n, 6, 4)\)-code.

Lemma V.7. \(A_3(n, 6, 4) = U(n)\) for \(n = 18, 24, 72, 84, 96\).

Proof. We will apply Theorem V.2 frequently with a 4-GDD of this type in the following proofs.
For each $n \in \{54, 66, 114, 126, 138\}$, optimal codes are obtained by Theorem V.2 with the corresponding 4-GDDs of types $15^9, 15^6, 15^9, 15^6$ and $15^53^1$, respectively and with the required short codes. Here, an optimal $(33, 6, 4)_3$ code of size $U(33) = 104$ is generated from the following codewords under the automorphism $(0 4 8 \cdots 28)(1 5 9 \cdots 29)(2 6 10 \cdots 30)(3 7 11 \cdots 31)(32)$.

\[
\begin{array}{c|c}
 n & \text{Codewords} \\
 \hline
39 & \{23, 27, 30, 34\}, \{15, 28, 30, 36\}, \{13, 14, 23, 31\}, \{9, 12, 26, 31\}, \{10, 24, 33, 35\}, \{8, 11, 13, 33\} \\
51 & \{23, 27, 48, 30\}, \{32, 34, 44, 47\}, \{61, 39, 44, 50\}, \{26, 27, 36, 42\}, \{9, 26, 31, 48\}, \{8, 22, 29, 37\}, \{20, 46, 11\} \\
63 & \{14, 22, 48, 39\}, \{8, 19, 27, 50\}, \{11, 13, 44, 60\}, \{8, 37, 47, 57\}, \{10, 29, 32\}, \{81, 12, 18, 33\} \\
87 & \{50, 69, 76, 82\}, \{31, 49, 73, 74\}, \{41, 43, 74, 83\}, \{15, 38, 44, 65\}, \{60, 10, 14, 38\}, \{91, 15, 36, 80\}, \{40, 66, 13\}, \{27, 30, 86\}, \{29, 30, 46, 69\}, \{23, 43, 59, 84\}, \{61, 81, 91, 81\} \\
99 & \{30, 69, 68, 98\}, \{31, 41, 74, 89\}, \{14, 15, 25, 46\}, \{27, 79, 81, 92\}, \{91, 38, 42, 90\}, \{121, 23, 54, 79\}, \{20, 25, 36, 77\}, \{40, 52, 72\}, \{68, 74, 29\} \\
111 & \{121, 35, 59, 88\}, \{46, 64, 91, 104\}, \{55, 70, 86, 103\}, \{21, 16, 91, 108\}, \{42, 79, 90, 99\}, \{29, 31, 50, 70\}, \{91, 101, 42\}, \{33, 72, 75, 100\}, \{34, 62, 68, 101\} \\
123 & \{32, 43, 110, 118\}, \{40, 69, 83, 111\}, \{52, 59, 111, 115\}, \{17, 26, 93, 107\}, \{13, 22, 46, 56\}, \{10, 65, 35\}, \{24, 88, 49\}, \{51, 28, 74, 75\}, \{26, 27, 36, 103\}, \{14, 15, 25, 46\}, \{27, 79, 81, 92\}, \{91, 38, 42, 90\}, \{121, 23, 54, 79\}, \{20, 25, 36, 77\}, \{40, 52, 72\}, \{68, 74, 29\} \\
\end{array}
\]

For $n = 36u + m$ with $u \geq 4$ and $m = 6, 54, 30$, an optimal code is constructed from a 4-GDD of type $36^u m^1$ and the required short codes, except when $n = 162$, for which a 4-GDD of type $39^6 3^1$ is used [7].

\[\square\]

Lemma V.7. $A_3(n, 6, 4) = U(n)$ for $n = 9$ (mod 12).

Proof. The case $n = 9$ is done in Lemma V.2 and $n = 33$ is given in the proof of Lemma V.6.

For $n = 21, 45$, the base codewords are given in Table III, which are developed by the automorphism $(0 4 8 \cdots n - 5)(1 5 9 \cdots n - 4)(2 6 10 \cdots n - 3)(3 7 11 \cdots n - 2)(n - 1)$ repeatedly.

For $n = 57, 93$, the base codewords are also listed in Table III but with a different automorphism $(0 2 4 \cdots n - 3)(1 3 5 \cdots n - 2)(n - 1)$.

For $n = 36u + m$ with $u \geq 4$ and $m = 9, 21, 33$, an optimal code is from a 4-GDD of type $36^u m^1$ by Theorem V.2.

\[\square\]

Lemma V.8. The values of $A_3(n, 6, 4)$ are at least $33, 55, 161, 461, 538, 616, 705, 802, 901$ for $n = 18, 24, 42, 72, 78, 84, 90, 96, 102$, respectively.

Proof. For $n = 18$ and 24, we can show that $A_3(n, 6, 4) \geq 33$ and 55, respectively, by computer search. The corresponding codes are available upon request. For $n = 42, 72, 78, 84, 90, 96, 102$, the codes are constructed from 4-GDDs of types $6^7, 15^4 12^1, 15^5 18^3, 12^2, 6^5 30^1, 6^{12} 30^1$ combined with the required short codes, respectively.

\[\square\]

Combining Lemmas V.3[V.5] we obtain the following theorem.

Theorem V.3. $A_3(n, 6, 4) = U(n)$ for all $n \equiv 0$ (mod 3) and $n > 3$, except for $n = 12$ which has value $U(n) - 1$, and possibly except for $n \in \{18, 24, 42, 72, 78, 84, 90, 102\}$. 

TABLE III: Base codewords of small CWCs in Lemma \[V.7\]

| n  | Codewords |
|----|------------|
| 21 | \{7_1, 9_1, 14_1, 19_1\}, \{3_1, 4_1, 6_1, 7_1\}, \{5_1, 14_1, 17_1, 18_1\}, \{9_1, 7_1, 12_1, 13_1\}, \{9_1, 5_1, 9_1, 11_1\}, \{8_1, 18_1, 4_2\} |
| 45 | \{10_1, 21_1, 22_1, 33_1\}, \{5_1, 15_1, 25_1, 41_1\}, \{10_1, 20_1, 19_2\}, \{8_1, 14_1, 62\} |
| 57 | \{19_1, 30_1, 41_1, 52_1\}, \{9_1, 17_1, 21_1, 37_1\}, \{9_1, 22_1, 31_1, 33_1\}, \{10_1, 20_1, 50_1, 55_1\} |
| 93 | \{1_1, 49_1, 64_1, 85_1\}, \{4_1, 7_1, 31_1, 57_1\}, \{1_1, 20_1, 44_1, 91_1\}, \{2_1, 38_1, 51_1, 78_1\} |

B. Length \(n \equiv 1 \pmod{3}\)

In this subsection, we determine the value of \(A_3(n, 6, 4)\) for \(n \equiv 1 \pmod{3}\) completely. Unlike the previous section, we use a 4-GDD and adjoin an infinite point. For convenience, we say an optimal \((n, 6, 4)\)-code \(\mathcal{C}\) has Property (A) if \(|\mathcal{C}| = U(n)\) and \(\mathcal{C}\) contains exactly \(n - 1\) codewords of type \(\mathcal{Y}_2\) and no codewords of type \(\mathcal{Z}_2\).

**Theorem V.4.** Suppose there exists a 4-GDD of type \(g^n(m - 1)^{1}\) where \(g \equiv 0, 5, 8, 9 \pmod{12}\). If there exists an optimal \((g + 1, 6, 4)_3\)-code with Property (A) and \(A_3(m, 6, 4) = U(m)\), then \(A_3(gu + m, 6, 4) = U(gu + m)\).

**Proof.** Suppose that \((X', G, B)\) is a 4-GDD of type \(g^n(m - 1)^{1}\). Let \(X = X' \cup \{\infty\}\). We will construct an optimal code \(\mathcal{C}\) of length \(gu + m\) in \(X \times [2]\) as follows. For each \(G' \in G\) of size \(g\), construct an optimal \((g + 1, 6, 4)_3\)-code \(C_{0} \subset (G' \cup \{\infty\}) \times [2]\) with Property (A), such that the codewords of type \(\mathcal{Z}_2\) have symbol 2 in the \(g\) positions from \(G\), and never in \(\infty\). For the group \(G\) of size \(m - 1\), let \(C_{0}\) be the collection of \((m, 6, 4)\)-codes in \((G \cup \{\infty\}) \times [2]\). Next, we view \(C_{0}\) as a code in \(X \times [2]\) in a natural way for each \(G' \in G\). Finally, let \(\mathcal{C} = C_{0} \cup (\bigcup_{G' \in G} C_{G'})\) be a \((g + m, 6, 4)\)-code in \(X \times [2]\) of size

\[
|\mathcal{C}| = |\mathcal{B}| + u \cdot U(g + 1) + U(m) = \frac{(g(u - 1) + m - 1)gu + gu(m - 1)}{12} + \frac{(g + 1)(g + 6) - 6}{12} + \frac{m(m + 5)}{12}
\]

In the second equality, we use the fact that \(U(g + 1) = \frac{(g + 1)(g + 6) - 6}{12}\) for each \(g \equiv 0, 5, 8, 9 \pmod{12}\).

When \(g = 12\), by Lemma \[II.5\] there exists a 4-GDD of type \(12^{m}1^{1}\) if and only if \(u \geq 4, m \leq 6(u - 1)\) and \(m \equiv 0 \pmod{3}\).

**Theorem V.5.** \(A_3(n, 6, 4) = U(n)\) for \(n \equiv 1 \pmod{3}\) and \(n \geq 7\).

**Proof.** For \(n = 7, 10\), it has been proved in Lemma \[V.2\] Note that the code of length 10 is of Property (A). For \(n = 13\), an optimal code with Property (A) is listed in Table \[IV\].

For each \(n \in \{16, 19, 22, 25, 31, 34, 40\}\), the base codewords are listed in Table \[V\] but with different group actions. For \(n = 16, 40\), the automorphism is \((0 \ 4 \ 8 \ \cdots \ n - 4)(1 \ 5 \ 9 \ \cdots \ n - 3)(2 \ 6 \ 10 \ \cdots \ n - 2)(3 \ 7 \ 11 \ \cdots \ n - 1)\); for \(n = 19, 31\), the automorphism is \((0 \ 1 \ 2 \ \cdots \ n - 1)\); for \(n = 22, 34\), the automorphism is \((0 \ 3 \ 6 \ \cdots \ n - 4)(1 \ 4 \ 7 \ \cdots \ n - 3)(2 \ 5 \ 8 \ \cdots \ n - 2)(n - 1)\); and for \(n = 25\), the automorphism is \((0 \ 6 \ 12 \ 18)(1 \ 7 \ 13 \ 19)(2 \ 8 \ 14 \ 20)(3 \ 9 \ 15 \ 21)(4 \ 10 \ 16 \ 22)(5 \ 11 \ 17 \ 23)(24)\). For \(n = 28\), the code is from a 4-GDD of type \(7^8\) by Theorem \[V.2\] and an optimal \((7, 6, 4)_3\) code.

For \(n = 37, 43, 46, 52\), optimal codes are obtained by Theorem \[V.4\] with 4-GDDs of types \(9^4, 9^6^4, 9^6, 9^6^3\), respectively and an optimal \((10, 6, 4)_3\) code with Property (A) and an optimal \((7, 6, 4)_3\) code.

For all other integers \(n\), write \(n = 12u + m\) with \(u \geq 4\) and \(m = 1, 7, 10, 16\). An optimal code is obtained by Theorem \[V.4\] with a 4-GDD of type \(12^{m}(m - 1)^{1}\), a \((13, 6, 4)_3\) code with Property (A), and an optimal \((m, 6, 4)_3\) code. Here, when \(m = 1\), the \((m, 6, 4)_3\) code is empty.

C. Length \(n \equiv 2 \pmod{3}\)

In this subsection, we deal with the case \(n \equiv 2 \pmod{3}\). Similar to the case \(n \equiv 1 \pmod{3}\), we use 4-GDDs but adjoin two infinite points. We call an optimal \((n, 6, 4)_3\)-code \(\mathcal{C}\) has Property (B) if \(|\mathcal{C}| = U(n)\) and \(\mathcal{C}\) contains exactly \(n - 2\) codewords of type \(\mathcal{Y}_2\), exactly one codeword of type \(\mathcal{Z}_2\), and the rest of type \(X'\).
Theorem V.6. Suppose there exists a 4-GDD of type $g^4(m - 2)^1$ where $g \equiv 0 \pmod{3}$. If there exists an optimal $(g + 2,6,4)_3$ code with Property (B) and $A_3(m,6,4) = U(m)$, then $A_3(gu + m,6,4) = U(gu + m)$.

Proof. Suppose that $(X',G,B)$ is a 4-GDD of type $g^4(m - 2)^1$, where the specific group $G_0 \in G$ is of size $m - 2$. Let $X = X' \cup \{1,2\}$, where $1,2 \notin X'$. We will construct an optimal code $C$ of length $gu + m$ in $X \times [2]$ as follows. For each $g' \in G$ of size $g$, construct an optimal $(g + 2,6,4)_3$ code $C_{g'} \subset (G \cup \{1,2\}) \times [2]$ with Property (B), such that the $g'$ codewords of type $\mathcal{Y}_2$ have symbol 2 in the $g'$ positions from $G$, and the type $\mathcal{Z}_2$ codeword is $\{r_2,j_2\}$. Let $C_{g'} = C_{g'} \setminus \{r_2,j_2\}$ for each $g' \in G$ of size $g$. For the group $G_0$ of size $m - 2$, let $C_{G_0}$ be an optimal $(m,6,4)_3$ code in $(G_0 \cup \{1,2\}) \times [2]$. Next, we view $C_{G_0}$ as a code in $X \times [2]$ in a natural way for each $G_0 \in G$. Finally, let $C_0$ be the collection of codewords of type $\mathcal{X}$ obtained from $B$. Then $C = C_0 \cup \cup_{C \in C_0} C_{g'}$ is an $(gu + m,6,4)_3$ code in $X \times [2]$ of size

$$|C| = |\mathcal{B}| + u \cdot (U(g + 2) - 1) + U(m)$$

$$= \frac{(g(u - 1) + m - 2)gu + gu(m - 2)}{12} + u \cdot \left(\frac{(g + 2)(g + 7) - 2}{12} - 1\right) + \left\lfloor \frac{m(m + 5)}{12} \right\rfloor.$$

In the second line, we use the fact that $U(g + 2) = \frac{(g+2)(g+7)-2}{12}$ for each $g \equiv 0 \pmod{3}$. □

When $g = 24$, by Lemma V.5 there exists a 4-GDD of type $24^4m^1$ if and only if $u \geq 4, m \leq 12(u - 1)$ and $m \equiv 0 \pmod{3}$.

Lemma V.9. $A_3(n,6,4) = U(n)$ for $n \equiv 2 \pmod{12}$ and $n \geq 26$.

Proof. For $n = 26$, the base codewords are listed in Table V, which are developed under the automorphism $(0 6 12 18)(1 7 13 19)(2 8 14 20)(3 9 15 21)(4 10 16 22)(5 11 17 23)(24 25)$ repeatedly.

For $n = 38, 50, 62, 74, 86, 98$, the base codewords are also given in Table V, but with a different automorphism $(0 2 4 \cdots n - 4)(1 3 5 \cdots n - 3)(n - 2 n - 1)$.

For $n = 110$, the code is obtained by Theorem V.6 with a 4-GDD of type $27^4$ and an optimal $(29,6,4)_3$ code with Property (B). Here, the base codewords of the optimal $(29,6,4)_3$ code with Property (B) are listed in Table V, which are developed under the automorphism $(0 3 6 \cdots 24)(1 4 7 \cdots 25)(2 5 8 \cdots 26)(27)(28)$.

For $n = 24u + m$ with $u \geq 4$ and $m = 26, 38$. An optimal code is obtained by Theorem V.6 with a 4-GDD of type $24^4(m - 2)^1$, a $(26,6,4)_3$ code with Property (B), and an optimal $(m,6,4)_3$ code.

Lemma V.10. $A_3(n,6,4) = U(n)$ for $n \equiv 5 \pmod{12}$ and $n \geq 29$.

Proof. The case $n = 29$ is done in Lemma V.9. For $n = 41, 53, 65, 77, 89, 101, 113$, the base codewords are listed in Table VI, which are developed under the automorphism $(0 3 6 \cdots n - 5)(1 4 7 \cdots n - 4)(2 5 8 \cdots n - 3)(n - 2)(n - 1)$ repeatedly.
For \( n = 24u + m \) with \( u \geq 4 \) and \( m = 29, 41 \). An optimal code is obtained by Theorem V.6 with a 4-GDD of type \( 24^u(m - 2)^u \), a \((26, 6, 4)_3\) code with Property (B), and an optimal \((m, 6, 4)_3\) code. There is one exception \( n = 137 \), for which an optimal code is obtained from a 4-GDD of type \( 27^u \) and a \((29, 6, 4)_3\) code with Property (B) by Theorem V.6.

### Table V: Base codewords of small CWCs in Lemma V.9

| \( n \) | Codewords |
|-----|---------|
| 26  | \{3, 9, 15, 21\} |
|     | \{4, 10, 16, 22\} |
|     | \{5, 11, 17, 23\} |
|     | \{6, 12, 18, 24\} |
|     | \{7, 13, 19, 25\} |
| 29  | \{6, 14, 20, 26\} |
|     | \{7, 15, 21, 27\} |
| 32  | \{12, 16, 18, 20\} |
| 38  | \{10, 16, 22, 28\} |
| 50  | \{1, 21, 22, 31\} |
| 62  | \{5, 15, 23, 33\} |
| 74  | \{1, 14, 15, 26\} |
| 86  | \{0, 52, 63, 73\} |
| 98  | \{4, 61, 12, 75\} |
|     | \{68, 97, 61\} |
|     | \{902, 972\} |

### Table VI: Base codewords of small CWCs in Lemma V.10

| \( n \) | Codewords |
|-----|---------|
| 41  | \{2, 27, 29, 51\} |
|     | \{3, 14, 24, 75\} |
|     | \{4, 10, 16, 39\} |
|     | \{5, 11, 17, 45\} |
| 53  | \{0, 16, 21, 24, 30\} |
|     | \{17, 9, 32, 50\} |
|     | \{21, 34, 47\} |
| 65  | \{1, 21, 22, 62\} |
|     | \{9, 20, 31, 41\} |
|     | \{10, 11, 23, 130\} |
|     | \{632, 642\} |
| 77  | \{33, 31, 51, 72\} |
|     | \{12, 62, 67, 75\} |
|     | \{10, 19, 42, 57\} |
|     | \{37, 54, 45\} |
| 89  | \{53, 65, 64, 81\} |
|     | \{31, 1, 2, 84\} |
|     | \{16, 26, 75, 78\} |
|     | \{14, 39, 46, 59\} |
|     | \{872, 88\} |
| 101 | \{24, 38, 45, 90\} |
|     | \{17, 24, 61, 73\} |
|     | \{6, 13, 41, 60\} |
|     | \{44, 47, 62\} |
| 113 | \{58, 91, 95, 101\} |
|     | \{18, 42, 65, 101\} |
|     | \{13, 27, 47, 94\} |
|     | \{12, 23, 31, 94\} |
|     | \{17, 27, 30, 65\} |
|     | \{111, 112\} |

#### Lemma V.11

\( A_4(n, 6, 4) = U(n) \) for \( n \equiv 8 \) (mod 12) and \( n \notin \{44, 56, 68, 80, 92\} \).

**Proof.** The case \( n = 8 \) is done in Lemma V.2 For \( n = 20 \), this code is generated from the base codewords by the
TABLE VII: The lower and upper bounds of $A_3(n,6,4)$ for small $n$.

| $n$  | 14 | 17 | 18 | 24 | 35 | 42 | 44 | 47 | 56 | 59 | 68 |
|-----|----|----|----|----|----|----|----|----|----|----|----|
| lower bound | 21 | 30 | 33 | 55 | 114 | 161 | 176 | 200 | 280 | 310 | 409 |
| upper bound | 22 | 31 | 34 | 58 | 116 | 164 | 179 | 203 | 284 | 314 | 413 |

| $n$  | 71 | 72 | 78 | 80 | 83 | 84 | 90 | 92 | 95 | 96 | 102 |
|-----|----|----|----|----|----|----|----|----|----|----|-----|
| lower bound | 445 | 461 | 538 | 562 | 603 | 616 | 705 | 738 | 786 | 803 | 901 |
| upper bound | 449 | 462 | 539 | 566 | 608 | 623 | 712 | 743 | 791 | 808 | 909 |

For $n = 32$, this code is generated from the following codewords under the automorphism $(0 3 6 \cdots 21)(1 4 7 \cdots 22)(2 5 8 \cdots 23)(24 26 28 30 25 27 29 31).

For $n = 24u + m$ with $u \geq 4$ and $m = 8, 20$. The optimal code is constructed from a 4-GDD of type $24^u(m-2)^1$ and the required short codes by Theorem V.9.

Lemma V.12. $A_3(n, 6, 4) = U(n)$ for $n \equiv 11 \pmod{12}$ and $n \not\in \{35, 47, 59, 71, 83, 95\}$.

Proof. The case $n = 11$ is done in Lemma V.2. For $n = 23$, this code is generated from the following codewords under the automorphism $(0 5 10)(1 6 11)(2 7 12)(3 8 13)(4 9 14)(15 17 19)(16 18 20)(21)(22).

For $n = 24u + m$ with $u \geq 4$ and $m = 11, 23$. The optimal code is constructed from a 4-GDD of type $24^nu(m-2)^1$ and the required short codes by Theorem V.9.

Lemma V.13. The values of $A_3(n, 6, 4)$ are at least 21, 30, 114, 176, 200, 280, 301, 409, 445, 562, 603, 738, 786, 803, for $n = 14, 17, 35, 44, 47, 56, 59, 68, 71, 80, 83, 92, 95, 96$, respectively.

Proof. For $n = 14, 35, 44, 47, 56, 59, 68, 71, 80, 83, 92, 95$, the codes are constructed from 4-GDDS of types $2^7$, $2^{12}11^1$, $2^{22}$, $2^{18}11^1$, $2^{24}8^1$, $2^{24}11^1$, $2^{24}20^1$, $2^{24}23^1$, $2^{27}3^1$, $2^{30}23^1$, $2^{33}26^1$, $2^{33}29^1$ combined with the required short codes, respectively.

When $n = 17$, $A_3(17, 6, 4) \geq 30$ and the corresponding code is found by computer.

When $n = 96$, we can improve the lower bound 802 given in Lemma V.8 by using a 4-GDD of type $7^{12}10^1$ and a $(9, 6, 4)_3$ code with Property (B) by Theorem V.6.

Combining Lemmas V.9, V.13, we obtain the following theorem.

Theorem V.7. $A_3(n, 6, 4) = U(n)$ for all $n \equiv 2 \pmod{3}$ and $n \geq 8$, and possibly except for $n \in \{14, 17, 35, 44, 47, 56, 59, 68, 71, 80, 83, 92, 95\}$.

By Subsections V-A, V-B and Lemma V.2 we summarize the results as follows.

Theorem V.8. Let $M = \{14, 17, 18, 24, 35, 42, 44, 47, 56, 59, 68, 71, 72, 78, 80, 83, 84, 90, 92, 95, 96, 102\}$. For any positive integer $n$,

$$A_3(n, 6, 4) = \begin{cases} U(n) - 1, & \text{if } n = 3, 4, 5, 12 \\ U(n), & \text{if } n \not\in M \cup \{3, 4, 5, 12\}. \end{cases}$$

For $n \in M$, the lower and upper bounds for $A_3(n, 6, 4)$ are given in the Table VII.
VI. Ternary CWCs of Distance 2w − 2

In this section, we consider ternary CWCs with weight \( w \) and distance \( 2w - 2 \) for general \( w \), and determine the value of \( A_3(n, 2w - 2, w) \) when \( n \) is sufficiently large under certain conditions based on graph packings.

For a graph \( H \) without isolated vertices, \( \gcd(H) \) denotes the greatest common divisor of the degrees of all vertices of \( H \). A graph \( G \) is called \( d \)-divisible if \( \gcd(G) \) is divisible by \( d \), while \( G \) is called nowhere \( d \)-divisible if no vertex of \( G \) has degree divisible by \( d \). An \( H \)-packing of a graph \( G \) is a set \( \{G_1, \ldots, G_s\} \) of edge-disjoint subgraphs of \( G \) where each subgraph is isomorphic to \( H \). Further, if \( G \) is a union of \( G_i \), \( i = 1, 2, \ldots, s \), then we call it an \( H \)-decomposition. The \( H \)-packing number of \( G \), denoted by \( P(H, G) \), is the maximum cardinality of an \( H \)-packing of \( G \). In particular, if \( H = K_k \) and \( G = K_e \) are complete graphs, then \( P(H, G) \) is equal to the packing number \( D(v, k, 2) \) introduced in Section II-B. Our main tool is the following result of Alon et al. [36].

**Theorem VI.1.** Let \( H \) be a graph with \( h \) edges, and let \( \gcd(H) = e. \) Then there exist \( N = \frac{N(H)}{e} \), and \( \varepsilon = \varepsilon(H) \) such that for any \( e \)-divisible or nowhere \( e \)-divisible graph \( G = (V, E) \) with \( n > N(H) \) vertices and \( \delta(G) > (1 - \varepsilon(H))n \),

\[
P(H, G) = \left\lfloor \frac{\sum_{v \in V} \alpha_v}{2h} \right\rfloor,
\]

unless when \( G \) is \( e \)-divisible and \( 0 < |E| \pmod{h} \leq \frac{e^2}{2} \), in which case

\[
P(H, G) = \left\lfloor \frac{\sum_{v \in V} \alpha_v}{2h} \right\rfloor - 1.
\]

Here, \( \alpha_v \) is the degree of vertex \( v \), rounded down to the closest multiple of \( e \).

Consider a ternary code \( C \subseteq \mathbb{T}_3^n \) with constant weight \( w \). We say a codeword in \( C \) has type \( 1^x2^y \) if it has \( x \) entries of symbol 1 and \( y \) entries of symbol 2, where \( x + 2y = w \). Note that \( y \) could be 0, 1, \ldots, \( \lfloor \frac{w}{2} \rfloor \). Then \( C \) is an \( (n, 2w - 2, w) \) code if and only if \( C \) satisfies the following properties:

(1’) The collection of subsets \( \text{supp}(u) \subset \mathbb{Z}_n \), for all codewords \( u \in C \), forms a \( 2-(n, \left\lfloor \frac{w}{2} \right\rfloor, \left\lceil \frac{w}{2} \right\rceil, 1) \) packing.

(2’) For any two codewords \( u, v \in C \), if \( x_i \in u \) and \( x_j \in v \), then \( i \in \{i, j\} \).

**Lemma VI.1.** \( A_3(n, 2w - 2, w) \leq \left\lfloor \frac{n(n-1-(w-1)(w-2))}{w(w-1)} \right\rfloor + n. \)

**Proof.** Let \( x_y \) be the number of codewords of type \( 1^x2^y \), \( y = 0, 1, \ldots, \left\lfloor \frac{w}{2} \right\rfloor \). By properties (1’) and (2’), we have

\[
\left( \begin{array}{c} w \vspace{1mm} \end{array} \right)_2 x_0 + \left( \begin{array}{c} w-1 \vspace{1mm} \end{array} \right)_2 x_1 + \cdots + \left( \begin{array}{c} w-2 \vspace{1mm} \end{array} \right)_2 x_{\left\lfloor \frac{w}{2} \right\rfloor} \leq \left\lfloor \frac{n}{2} \right\rfloor,
\]

and

\[
x_1 + 2x_2 + \cdots + \left( \begin{array}{c} w \vspace{1mm} \end{array} \right)_2 x_{\left\lfloor \frac{w}{2} \right\rfloor} \leq n.
\]

Note that \( \left( \begin{array}{c} w-t \vspace{1mm} \end{array} \right)_2 + t(w-1) = \left( \begin{array}{c} w \vspace{1mm} \end{array} \right)_2 + \left( \begin{array}{c} 1 \vspace{1mm} \end{array} \right)_2 \). Computing \( \left( \begin{array}{c} 1 \vspace{1mm} \end{array} \right)_2 + \left( \begin{array}{c} w \vspace{1mm} \end{array} \right)_2(2) \), we obtain

\[
x_0 + x_1 + \cdots + x_{\left\lfloor \frac{w}{2} \right\rfloor} \leq \left\lfloor \frac{n(n-1-(w-1)(w-2))}{w(w-1)} \right\rfloor + n,
\]

which complete the proof.

For convenience, let

\[
B(n) := \left\lfloor \frac{n(n-1-(w-1)(w-2))}{w(w-1)} \right\rfloor.
\]

We will prove that the upper bound in Lemma VI.1 can be achieved for certain values of \( n \) by Theorem VI.1. The desired code \( C \) has only two types of codewords, \( 1^w \) and \( 1^{w-2}2^1 \). Consider the complete graph \( K_n \) with vertex set \( \mathbb{Z}_n \). For each \( u \in C \), view it as a complete subgraph on the vertex set \( \text{supp}(u) \). Then property (1’) tells that all these subgraphs are pairwise edge-disjoint, thus form a packing of \( K_n \). We need to be careful about codewords containing symbol 2, for which the corresponding positions should be different by property (2’). We construct such codes by Golomb rulers [38].

An \( (n, w) \) modular Golomb ruler is a set of \( w \) integers \( \{a_1, a_2, \ldots, a_w\} \), such that all of the differences, \( \{a_i - a_j|1 \leq i \neq j \leq w\} \), are distinct and nonzero modulo \( n \). Suppose that we have an \( (n, w-1) \) modular Golomb ruler \( \{a_1, a_2, \ldots, a_{w-1}\} \). Then the \( n \) codewords \( \{(a_1+i)2, (a_2+i)1, \ldots, (a_{w-1}+i)1\} \) of type \( 1^{w-2}2^1 \), \( i \in \mathbb{Z}_n \), have pairwise distance at least \( 2w-2 \). Associate these \( n \) codewords with \( n \) complete graphs \( K_{w-1} \) with vertex set \( \{a_1+i, a_2+i, \ldots, a_{w-1}+i\} \), \( i \in \mathbb{Z}_n \), which are edge-disjoint subgraphs of \( K_n \) with vertex set \( \mathbb{Z}_n \). Let \( S \) be the union of these \( n \) subgraphs \( K_{w-1} \). It is easy to show that \( S \) is a regular subgraph of \( K_n \) with degree \( (w-1)(w-2) \). Denote \( G = K_n \setminus S \). We will apply Theorem VI.1 to obtain a \( K_n \)-packing of \( G \), which yields the remaining codewords of type \( 1^w \).

**Theorem VI.2.** Let \( w \geq 3 \) be any fixed integer. Then \( A_3(n, 2w - 2, w) \geq B(n) + n - 1 \) for any sufficiently large integer \( n \equiv 1 \pmod{w - 1} \). Further if \( n \equiv w - 2w + 3, 1 \) or \( -w + 2 \pmod{w(w - 1)} \), then \( A_3(n, 2w - 2, w) = B(n) + n \).
Proof. By [38], there exists an \((n, 2w-1)\) modular Golomb ruler for any \(n = \Omega(w^2)\). Thus, by above discussion, we obtain \(n\) codewords of type \(1^{w-2}2^1\), and a regular graph \(G = (V, E)\) with vertex set \(V = \mathbb{Z}_n\) and degree \(d = n - (w-1)(w-2)\). Let \(H = K_w\) in Theorem VI.1, then \(e = \gcd(H) = w - 1\) and \(h = \frac{w(w-1)}{2}\). Since \(n\) is sufficiently large, we have \(d \geq (1 - \epsilon)n\), where \(\epsilon = \epsilon(K_w)\) is defined in Theorem VI.1. Further \(n \equiv 1 \pmod{w-1}\) implies that \(d \equiv 0 \pmod{w-1}\), i.e., \(G\) is \((w-1)\)-divisible, so \(\alpha_v = d = n - (w-1)(w-2)\) for each \(v \in V\). By Theorem VI.1, we have a \(K_w\)-packing of \(G\) with packing number

\[ P(K_w, G) \geq \left\lceil \frac{\sum_{v \in V} \alpha_v}{2^{\left\lfloor \frac{w}{2} \right\rfloor}} \right\rceil - 1 = \left\lceil \frac{n(n-1-(w-1)(w-2))}{w(w-1)} \right\rceil - 1 = B(n) - 1. \]

Each \(K_w\) of this packing gives a codeword of type \(1^w\) in a natural way. Thus we obtain at least \(B(n)-1\) codewords. Combining the \(n\) codewords of type \(1^{w-2}2^1\), we have an \((n, 2w-2, w)\) code of size at least \(B(n) + n - 1\).

When \(n \equiv -2w + 3, w \pmod{w(w-1)}\), it is easy to check that \(|E| \equiv nd/2 \equiv 0 \pmod{\frac{w(w-1)}{2}}\). In fact, one can show that these are the only two congruent classes of \(n\) that satisfies \(|E| \pmod{\frac{w(w-1)}{2}} 
eq [1, \frac{(w-1)^2}{2}]\). By Theorem VI.1, we have \(P(K_w, G) = B(n)\), hence \(A_3(n, 2w-2, w) = B(n) + n\).

For \(n \equiv -w + 2, 1 \pmod{w(w-1)}\), we consider a slightly different model. Let \(S^G\) be a regular graph on \(\mathbb{Z}_{n-1}\) of degree \((w-1)(w-2)\), which is a union of \(n-1\) edge-disjoint complete subgraphs \(K_{w-1}\) obtained by an \((n-1, w-1)\) modular Golomb ruler. Note that \(S^G\) corresponds to \(n-1\) codewords of type \(1^{w-2}2^1\). Let \(G' = K_n \setminus S^G\), which has vertex set \(V' = \mathbb{Z}_{n-1} \cup \infty\) and edge set \(E'\). Then \(G'\) has degree \(n-1 - (w-1)(w-2)\) for each \(v \in \mathbb{Z}_{n-1}\), and degree \(n-1\) for the vertex \(\infty\). So \(G'\) is \((w-1)\)-divisible. Further, it is easy to check that

\[ |E'| = \frac{(n-1)(n-1-(w-1)(w-2)) + n-1}{2} \equiv 0 \pmod{\frac{w(w-1)}{2}}. \]

By Theorem VI.1, we have

\[ P(K_w, G') = \left\lceil \frac{\sum_{v \in V} \alpha_v}{2^{\left\lfloor \frac{w}{2} \right\rfloor}} \right\rceil = \frac{n(n-1-(w-1)(w-2)) + (w-1)(w-2)}{w(w-1)} = B(n) + 1. \]

The last equality is due to the fact that \(B(n) = \frac{n(n-1-(w-1)(w-2)) + 2(w-1)}{w(w-1)}\) in these cases. Hence we get \(A_3(n, 2w-2, w) = n-1 + P(K_w, G') = B(n) + n\) for \(n \equiv -w + 2, 1 \pmod{w(w-1)}\).

Theorem VI.1.2 tells us when \(n\) is sufficiently large and \(n \equiv w, -2w + 3, 1, -w + 2 \pmod{w(w-1)}\), \(A_3(n, 2w-2, w) = B(n) + n\). This is consistent with the previous sections when \(w \leq 4\) and \(d = 2w - 2\). In particular, we have \(A_3(n, 6, 4) = B(n) + n\) if \(n \equiv 1 \pmod{w-1}\) for \(w \geq 4\).

VII. Conclusion

In this paper, we determine the maximum size of constant-weight codes in \(l_1\)-metric over integers or over \(I_3 = \{0, 1, 2\}\), for weight three and four. We also provide an asymptotic result for the maximum size of ternary codes with general weight \(w\) and distance \(2w - 2\). It is plausible that we could extend the method in Section VI by looking for irregular graphs \(S\), such that by deleting edges in \(S\), the resultant has a clique decomposition. We leave this for future study. Further, constructing optimal constant-weight codes over integers is a more challenging problem.

REFERENCES

[1] D. J. Costello and G. D. Forney, “Channel coding: The road to channel capacity,” Proceedings of the IEEE, vol. 95, no. 6, pp. 1150–1177, 2007.
[2] O. D. King, “Bounds for DNA codes with constant GC-content,” The electronic journal of combinatorics, vol. 10, no. 1, p. 33, 2003.
[3] O. Milenkovic and N. Kashyap, “On the design of codes for DNA computing,” in International Workshop on Coding and Cryptography. Springer, 2005, pp. 100–119.
[4] R. Graham and N. Sloane, “Lower bounds for constant weight codes,” IEEE Transactions on Information Theory, vol. 26, no. 1, pp. 37–43, 1980.
[5] A. Brouwer, J. Shearer, N. Sloane, and W. Smith, “A new table of constant weight codes,” IEEE Transactions on Information Theory, vol. 36, no. 6, pp. 1334–1380, 1990.
[6] E. Agrell, A. Vardy, and K. Zeger, “Upper bounds for constant-weight codes,” IEEE Transactions on Information Theory, vol. 46, no. 7, pp. 2373–2395, 2000.
[7] H. Zhang, X. Zhang, and G. Ge, “Optimal ternary constant-weight codes with weight 4 and distance 5,” IEEE Transactions on Information Theory, vol. 58, no. 5, pp. 2706–2718, 2012.
[8] Y. M. Chee, H. Zhang, and X. Zhang, “Complexity of dependences in bounded domains, armstrong codes, and generalizations,” IEEE Transactions on Information Theory, vol. 61, no. 2, pp. 812–819, 2014.
[9] Y. M. Chee, G. Ge, H. Zhang, and X. Zhang, “Hanani triple packings and optimal q-ary codes of constant weight three,” Designs, Codes and Cryptography, vol. 75, no. 3, pp. 387–403, 2015.
[10] Y. M. Chee and X. Zhang, “Linear size constant-composition codes meeting the Johnson bound,” IEEE Transactions on Information Theory, vol. 64, no. 2, pp. 909–917, 2017.
[11] Y. M. Chee, H. M. Kiah, H. Zhang, and X. Zhang, “Constructions of optimal and near-optimal multiply constant-weight codes,” IEEE Transactions on Information Theory, vol. 63, no. 6, pp. 3621–3629, 2017.
[12] Y. M. Chee, F. Gao, H. M. Kiah, A. C. Hung Ling, H. Zhang, and X. Zhang, “Decompositions of edge-colored digraphs: A new technique in the construction of constant-weight codes and related families,” SIAM Journal on Discrete Mathematics, vol. 33, no. 1, pp. 209–229, 2019.

[13] S. Jain, F. F. Hassanzadeh, M. Schwartz, and J. Bruck, “Duplication-correcting codes for data storage in the DNA of living organisms,” IEEE Transactions on Information Theory, vol. 63, no. 8, pp. 4996–5010, 2017.

[14] M. Kovačević and V. Y. F. Tan, “Asymptotically optimal codes correcting fixed-length duplication errors in DNA storage systems,” IEEE Communications Letters, vol. 22, no. 11, pp. 2194–2197, 2018.

[15] A. Lenz, N. Jünger, and A. Wachter-Zeh, “Bounds and constructions for multi-symbol duplication error correcting codes,” arXiv preprint arXiv:1807.02874, 2018.

[16] Y. Tang, Y. Yehezkeally, M. Schwartz, and F. F. Hassanzadeh, “Single-error detection and correction for duplication and substitution channels,” in 2019 IEEE International Symposium on Information Theory (ISIT). IEEE, 2019, pp. 300–304.

[17] Y. Yehezkeally and M. Schwartz, “Reconstruction codes for DNA sequences with uniform tandem-duplication errors,” IEEE Communications Letters, vol. 22, no. 11, pp. 2194–2197, 2018.

[18] E. Lander, L. Linton, B. Birren, C. Nusbaum, M. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh et al., “Initial sequencing and analysis of the human genome.” Nature, vol. 409, no. 6822, pp. 860–921, 2001.

[19] K. Usdin, “The biological effects of simple tandem repeats: lessons from the repeat expansion diseases,” Genome research, vol. 18, no. 7, pp. 1011–1019, 2008.

[20] J. W. Fondon and H. R. Garner, “Molecular origins of rapid and continuous morphological evolution,” Proceedings of the National Academy of Sciences, vol. 101, no. 52, pp. 18 058–18 063, 2004.

[21] A. Barg and A. Mazumdar, “Codes in permutations and error correction for rank modulation,” IEEE Transactions on Information Theory, vol. 56, no. 7, pp. 3158–3165, 2010.

[22] A. Jiang, H. Li, and Y. Wang, “Error scrubbing codes for flash memories,” in 2009 11th Canadian Workshop on Information Theory. IEEE, 2009, pp. 32–35.

[23] L. G. Tallini and B. Bose, “On l₁-distance error control codes,” in 2011 IEEE International Symposium on Information Theory Proceedings. IEEE, 2011, pp. 1061–1065.

[24] H. Zhou, M. Schwartz, A. A. Jiang, and J. Bruck, “Systematic error-correcting codes for rank modulation,” IEEE Transactions on Information Theory, vol. 61, no. 1, pp. 17–32, 2014.

[25] F. Farnoud, Y. Skachek, and O. Milenkovic, “Error-correction in flash memories via codes in the ulam metric,” IEEE Transactions on Information Theory, vol. 59, no. 5, pp. 3003–3020, 2013.

[26] G. Kabatiansky and S. Kruglik, “On codes correcting constant number of errors in l₁ metric,” in Сборник трудов 39-й международной конференции ИППИ РАН “Информационные технологии и системы 2015”, 2015, pp. 152–157.

[27] M. Kovačević and V. Y. F. Tan, “Codes in the space of multisets-coding for permutation channels with impairments,” IEEE Transactions on Information Theory, vol. 64, no. 7, pp. 5156–5169, 2018.

[28] H. Jinushi and K. Sakaniwa, “A construction method for multilevel error-correcting codes based on absolute summation weight,” in Abst. 1990 IEEE Int. Symp. Inform. Theory, 1990, p. 87.

[29] A. Butson, “Generalized hadamard matrices,” Proceedings of the American Mathematical Society, vol. 13, no. 6, pp. 894–898, 1962.

[30] D. R. Stinson, R. Wei, and J. Yin, “Packings,” in Handbook of Combinatorial Designs. Chapman and Hall/CRC, 2006, pp. 576–582.

[31] J. Bao and L. Ji, “The completion determination of optimal (3,4)-packings,” Designs, Codes and Cryptography, vol. 77, no. 1, pp. 217–229, 2015.

[32] C. J. Colbourn and J. H. Dinitz, Handbook of combinatorial designs. CRC press, 2006.

[33] C. J. Colbourn and A. Rosa, “Quadratic leaves of maximal partial triple systems,” Graphs and Combinatorics, vol. 2, no. 1, pp. 317–337, 1986.

[34] C. J. Colbourn and A. C. Ling, “A class of partial triple systems with applications in survey sampling,” Communications in Statistics-Theory and Methods, vol. 27, no. 4, pp. 1099–1108, 1998.

[35] H. Wei and G. Ge, “Group divisible designs with block sizes from K₁(3) and Kirkman frames of type h"m1,” Discrete Mathematics, vol. 329, pp. 42–68, 2014.

[36] N. Alon, Y. Caro, and R. Yuster, “Packing and covering dense graphs,” Journal of Combinatorial Designs, vol. 6, no. 6, pp. 451–472, 1998.

[37] G. Ge, “Group divisible designs, handbook of combinatorial designs,” 2007.

[38] J. B. Shearer, “Difference triangle sets,” in Handbook of Combinatorial Designs. Chapman and Hall/CRC, 2006, pp. 436–440.