Construction of q-ary Constant Weight Sequences using a Knuth-like Approach

Elie N. Mambou and Theo G. Swart
Dept. of Electrical and Electronic Engineering Science, University of Johannesburg
P. O. Box 524, Auckland Park, 2006, South Africa
Email: {emambou, tgswart}@uj.ac.za

***Abstract*—We present an encoding and decoding scheme for constant weight sequences, that is, given an information sequence, the construction results in a sequence of specific weight within a certain range. The scheme uses a prefix design that is based on Gray codes. Furthermore, by adding redundant symbols we extend the range of weight values for output sequences, which is useful for some applications.**

I. Introduction

Constant weight (CW) sequences have found applications in the fields of computer science, information security and communications. They play an important role in communication system where high security and confidentiality are needed, because of various properties such as low correlations, balanced value distributions and strong linear complexity, amongst others.

Practically, CW sequences are used in several modern applications including frequency hopping in GSM networks, detection of unidirectional errors and threshold setting in barcode implementations. However, our target application is the domain of visible light communication (VLC) where CW sequences can be designed to perform light dimming which decreases flickering issues.

In the coding theory field, a CW code can be viewed as an error detection and correction code such that all codewords in that code have the same Hamming weight. Binary CW codes are also called m of n codes as each codeword has a length n and m instances of 1s (weight equals m). A special case of m of n code is the 1 of n code, it encodes $\log_2 n$ bits in a codeword of length n. For instance, the 1 of 2 code has inputs 0 and 1 and generates codewords 01 and 10. A 1 of 4 code generates codewords 0001, 0010, 0100 and 1000 given inputs 00, 01, 10 and 11. In this case the Hamming distance is $d = 2$ and each sequence has a weight $w = 1$.

There are many algorithms to generate binary CW codes. In [1], a construction of CW codes from a given code length is presented. The obtained codes are usually referred to as optical orthogonal codes. This scheme has an efficient algorithm for error code correction and performs the encoding and the decoding of CW codes. However, the construction is limited to specific constant dimension codes. In [2], a construction of CW codes based on Knuth's balancing approach [3] is presented. The flexibility on the tail bits is used to generate CW codes including balanced codes. As in Knuth's algorithm, the information about the changes in the word is carried in the prefix.

In [4], a construction for a set of non-binary constant-weight sequences was proposed with finite period from cyclic difference sets, this is based on the generalization of the binary case proposed in [5]. Other existing works on binary and non-binary constant-weight sequences generated from q-ary sequences and cyclic difference sets include [6]–[8].

The rest of this paper is structured as follows: in Section II, we present preliminary work for our construction. Section III shows the encoding of the q-ary CW sequences based on Gray code prefixes. Section IV presents the encoding of CW sequences with higher weights, then Section V describes the decoding method for this algorithm. Finally, in Section VI we analyze the redundancy for q-ary CW sequences.

II. Preliminaries

A. Balancing of q-ary information sequences

Let $\boldsymbol{x} = x_1x_2 \dots x_k$ be a q-ary information sequence of length k, with $x_i \in \{0, 1, \dots, q-1\}$, and $w(\boldsymbol{x})$ be the weight of $\boldsymbol{x}$, that is $w(\boldsymbol{x}) = \sum_{i=1}^{k} x_i$. When $w(\boldsymbol{x}) = k(q-1)/2$, $\boldsymbol{x}$ is said to be balanced, and $\beta_{k,q}$ represents this balancing value.

A construction was presented in [9] for the balancing of q-ary sequences, that stipulates that for any q-ary information sequence of k symbols, there is always a way to achieve balancing by adding (modulo q) one sequence from a set of weighting sequences to $\boldsymbol{x}$. The weighting sequence, $\boldsymbol{b}(s,p) = b_1b_2 \dots b_k$ is defined as follows:

$$b_i = \begin{cases} s, & i-1 \geq p, \\ s+1 \pmod{q}, & i-1 < p, \end{cases}$$

with s and p positive integers such that $0 \leq s \leq q-1$ and $0 \leq p \leq k-1$. Let z be the counter of all possible weighting sequences, $z = sk+p$ and $0 \leq z \leq kq-1$. We use $\boldsymbol{b}(s,p)$ and $\boldsymbol{b}(z)$ interchangeably to denote the z-th weighting sequence.

We define $\boldsymbol{y} = \boldsymbol{x} \oplus_q \boldsymbol{b}(z)$, as the addition (modulo q) of the information sequence $\boldsymbol{x}$ to the weighting sequence $\boldsymbol{b}(z)$.

Example 1: Consider $q = 3$ and $k = 4$, where we want to balance the ternary information sequence of length four, 2102.

The process (illustrated on the next page) leads to the construction of at least one balanced sequence. Bold weights indicate that the desired weight has been obtained. Fig. 1 presents the weight progression $w(\boldsymbol{y})$ vs. z for this example. The red line represents the balancing value, $\beta_{4,3} = 4$.

z	$\boldsymbol{x} \oplus_q \boldsymbol{b}(z) = \boldsymbol{y}$	$w(\boldsymbol{y})$
0	$2102 \oplus_3 0000 = 2102$	5
1	$2102 \oplus_3 1000 = 0102$	3
2	$2102 \oplus_3 1100 = 0202$	**4**
3	$2102 \oplus_3 1110 = 0212$	5
4	$2102 \oplus_3 1111 = 0210$	3
5	$2102 \oplus_3 2111 = 1210$	**4**
6	$2102 \oplus_3 2211 = 1010$	2
7	$2102 \oplus_3 2221 = 1020$	3
8	$2102 \oplus_3 2222 = 1021$	**4**
9	$2102 \oplus_3 0222 = 2021$	5
10	$2102 \oplus_3 0022 = 2121$	6
11	$2102 \oplus_3 0002 = 2101$	**4**

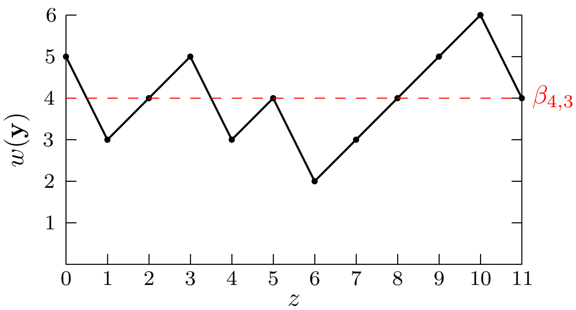

Fig. 1. Weight progression $w(\boldsymbol{y})$ vs. z for Example 1.

From this we can see that weight values other than $\beta_{4,3}$ are easily achievable, and this will form the basis for our CW algorithm in Section III.

B. q-ary Gray codes

Gray codes were invented by Gray [10] for solving problems in pulse code communication, and have been extended to various other applications.

We define an (r', q)-Gray code as the set of q-ary Gray sequences of length r'. This set presents the property that any two consecutive codewords (provided that the sequences being mapped from are listed in the normal lexicographic order) differ only in one symbol position and the difference between any two consecutive sequences' weights is ± 1.

Let $\boldsymbol{d} = d_1 d_2 \ldots d_{r'}$ be any sequence within the set of q-ary sequences of length r', listed in the normal lexicographic order. These sequences are mapped to (r', q)-Gray code sequences, $\boldsymbol{g} = g_1 g_2 \ldots g_{r'}$, by making use of the following encoding and decoding algorithms [11] for Gray codes.

Encoding algorithm for (r', q)-Gray code: Let S_i be the sum of the first $i-1$ symbols of $\boldsymbol{g}$, with $2 \le i \le r'$ and $g_1 = d_1$. Then

$$S_i = \sum_{j=1}^{i-1} g_j, \quad \text{and} \quad g_i = \begin{cases} d_i, & \text{if } S_i \text{ is even,} \\ q - 1 - d_i, & \text{if } S_i \text{ is odd.} \end{cases}$$

The parity of S_i determines symbols of $\boldsymbol{g}$ from $\boldsymbol{d}$. If S_i is even then the symbol stays the same, otherwise the q-ary complement of the symbol is taken.

Decoding algorithm for (r', q)-Gray code: As before, S_i is the sum of the first $i-1$ symbols of $\boldsymbol{g}$, with $2 \le i \le r'$ and $d_1 = g_1$. Then

$$S_i = \sum_{j=1}^{i-1} g_j, \quad \text{and} \quad d_i = \begin{cases} g_i, & \text{if } S_i \text{ is even,} \\ q - 1 - g_i, & \text{if } S_i \text{ is odd.} \end{cases}$$

C. Encoding of balanced q-ary information sequences based on Gray code prefixes

A method for the encoding and decoding of balanced sequences based on Gray code prefixes was presented in [12]. It was proven that any q-ary information sequence of length k can be balanced by adding (modulo q) an appropriate weighting sequence $\boldsymbol{b}(z)$, and prefixing a redundant symbol u with a Gray code sequence. This method consists of generating $\boldsymbol{y} = \boldsymbol{x} \oplus \boldsymbol{b}(z)$ outputs as presented in Section II-A. However, q-ary Gray code prefixes of length $r' = \log_q k + 1$, are appended to $\boldsymbol{y}$, this only works for information sequences where $k = q^t$, with t being a positive integer. For each output sequence, the q-ary representation of the index z is converted into its corresponding Gray code representation, as described in Section II-B. Then an extra digit u, is appended to the Gray code prefix to enforce balancing as necessary, where $u = \beta_{n,q} - w(\boldsymbol{c}')$, $n = k + r' + 1$ being the overall length of output sequence $\boldsymbol{c} = [u|\boldsymbol{g}|\boldsymbol{y}]$ and $\boldsymbol{c}' = [\boldsymbol{g}|\boldsymbol{y}]$, where $|$ represents concatenation of sequences.

Example 2: We consider encoding the ternary information sequence of length three, 102.

The process below shows the encoding, where the underlined part represents the prefix and the bold symbol is the symbol u. The Gray code length is $r' = 2$, the total length of the transmitted sequence is $n = 6$ and $\beta_{6,3} = 6$.

z	$\boldsymbol{x} \oplus_q \boldsymbol{b}(z) = \boldsymbol{y}$	$\boldsymbol{c} = [u \vert \boldsymbol{g} \vert \boldsymbol{y}]$	$w(\boldsymbol{c})$
0	$102 \oplus_3 000 = 102$	$\mathbf{0}\underline{00}102$	3
1	$102 \oplus_3 100 = 202$	$\mathbf{1}\underline{01}202$	**6**
2	$102 \oplus_3 110 = 212$	$\mathbf{0}\underline{02}212$	7
3	$102 \oplus_3 111 = 210$	$\mathbf{0}\underline{12}210$	**6**
4	$102 \oplus_3 211 = 010$	$\mathbf{0}\underline{11}010$	3
5	$102 \oplus_3 221 = 020$	$\mathbf{0}\underline{10}020$	3
6	$102 \oplus_3 222 = 021$	$\mathbf{1}\underline{20}021$	**6**
7	$102 \oplus_3 022 = 121$	$\mathbf{0}\underline{21}121$	7
8	$102 \oplus_3 002 = 101$	$\mathbf{0}\underline{22}101$	**6**

We have four occurrences of balanced outputs: $\mathbf{1}\underline{01}202$, $\mathbf{0}\underline{12}210$, $\mathbf{1}\underline{20}021$ and $\mathbf{0}\underline{22}101$. The encoding of $(2, 3)$-Gray code prefixes can be followed from Table I. Fig. 2 presents the weight progression for this example.

TABLE I
ENCODING OF $(2, 3)$-GRAY CODE

z	s, p	$\boldsymbol{b}(z)$	Sequence ($\boldsymbol{d}$)	Gray code ($\boldsymbol{g}$)
0	$0, 0$	000	00	00
1	$0, 1$	100	01	01
2	$0, 2$	110	02	02
3	$1, 0$	111	10	12
4	$1, 1$	211	11	11
5	$1, 2$	221	12	10
6	$2, 0$	222	20	20
7	$2, 1$	022	21	21
8	$2, 2$	002	22	22

The proposed algorithm for the encoding of q-ary CW sequences makes use of this process.

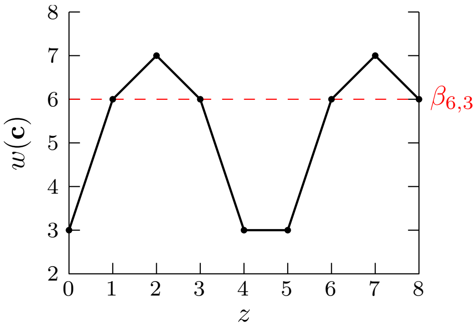

Fig. 2. Weight progression $w(\boldsymbol{c})$ vs. z for Example 2.

III. Encoding of q-ary CW sequences based on Gray code prefixes

We now extend this work on encoding balanced sequences, to encode CW sequences. Let an (n, k, W, q) CW sequence be the q-ary CW sequence of length n, weight W with k information symbols.

The encoding process consists of using the (r', q)-Gray code to encode the index z as prefix. The Gray code sequence, $\boldsymbol{g} = g_1 g_2 \dots g'_r$ is prefixed to all kq outputs as $\boldsymbol{c}' = [\boldsymbol{g}|\boldsymbol{y}]$. We have to find the appropriate length of (r', q)-Gray code prefixes that can uniquely match the kq weighting sequences.

We impose a condition that $k = q^t$, where t is a positive integer, therefore

$$r' = \log_q(kq) = \log_q(k) + 1, \tag{1}$$

such that the cardinalities of the Gray code set and that of the weighting sequences are equal. As in the case for balanced q-ary sequences, an extra digit u is also added to $\boldsymbol{c}'$ to control the overall weight of the CW sequence.

For the construction of (n, k, W, q) CW sequences, let $\boldsymbol{c} = [u|\boldsymbol{g}|\boldsymbol{y}]$ be the concatenation of u, $\boldsymbol{g}$ and $\boldsymbol{y}$. For a specific z, if $W \geq w(\boldsymbol{c}')$, provided that $u \in \{0, 1, \dots, q-1\}$, $u = W - w(\boldsymbol{c}')$, else if $W < w(\boldsymbol{c}')$ then $u = 0$. The overall encoded sequence has a length n and the prefix has a length of $r = r' + 1 = \log_q k + 2$.

Encoding algorithm for q-ary CW sequence:

1) The length of the (r', q)-Gray code prefix is calculated as in (1), $r' = \log_q k + 1$.
2) Incrementing through z, determine weighting sequences, $\boldsymbol{b}(z)$ and add them to $\boldsymbol{x}$, $\boldsymbol{y} = \boldsymbol{x} \oplus \boldsymbol{b}(z)$.
3) For each increment of z, determine the corresponding Gray code sequence, $\boldsymbol{g}$, using the Gray code encoding algorithm presented in Section II-B, append it to sequence $\boldsymbol{y}$ and obtain sequence $\boldsymbol{c}'$.
4) Finally, append the extra digit, u, to sequence $\boldsymbol{c}'$, where $u = W - w(\boldsymbol{c}')$ provided that $u \in \{0, 1, \dots, q-1\}$, otherwise $u = 0$.

Lemma 1: For any q-ary information sequence $\boldsymbol{x}$ of length k, where parameters k and q are not coprime, we can find a $\boldsymbol{b}(z)$ such that the weight of $\boldsymbol{y} = \boldsymbol{x} \oplus_q \boldsymbol{b}(z)$ is $\omega_1 \leq w(\boldsymbol{y}) \leq \omega_2$, where $\omega_1 = \beta_{k,q} - (q-1) = \frac{(k-2)(q-1)}{2}$ and $\omega_2 = \beta_{k,q} + (q-1) = \frac{(k+2)(q-1)}{2}$.

Proof: It was proven in [9] that the weight progression graph of $\boldsymbol{y} = \boldsymbol{x} \oplus_q \boldsymbol{b}(z)$ represents a path with increases of 1 and decreases of $q-1$, and it is such that $\min\{w(\boldsymbol{y})\} \leq \beta_{k,q}$ and $\max\{w(\boldsymbol{y})\} \geq \beta_{k,q}$. The constraint on the information sequences $k = q^t$, implies that k and q are not coprime. Lemma 1 is true if $\min\{w(\boldsymbol{y})\} \geq t_1$ and $\max\{w(\boldsymbol{y})\} \leq t_2$. The weighting sequence $\boldsymbol{b}(z)$ is such that $0 \leq w(\boldsymbol{b}(z)) \leq k(q-1)$, so the range of weighting sequence is greater than the range of $[\omega_2 - \omega_1] = 2(q-1)$. Therefore $\min\{w(\boldsymbol{y})\} \geq \omega_1$ and $\max\{w(\boldsymbol{y})\} \leq \omega_2$. ■

Theorem 1: An (n, k, W, q) CW sequence can be constructed from any q-ary information sequence $\boldsymbol{x}$ of length k where

$$\frac{(k-2)(q-1)}{2} \leq W \leq \frac{(k+2r'+4)(q-1)}{2}. \tag{2}$$

Proof: According to Lemma 1, $\frac{(k-2)(q-1)}{2} \leq W \leq \frac{(k+2)(q-1)}{2}$, however because of the flexibility of u and $\boldsymbol{g}$, the upper bound increases to $\frac{(k+2r'+4)(q-1)}{2}$. ■

Example 3: We want to encode the ternary sequence $\boldsymbol{x} = 212$ into a CW sequence of weight $W = 8$. The condition $k = q^t \Rightarrow 3$, is fulfilled and the Gray code length is $r' = \log_3 3 + 1 = 2$, according to (1).

The cardinality of the $(2, 3)$-Gray code equals that of the weighting sequences, with $kq = 9$. The overall length of the CW sequence is $n = 3 + 2 + 1 = 6$. The weight range as presented in (2) is such that $2 \leq W \leq 10$, therefore it is possible to construct $(6, 3, 8, 3)$ CW sequences. The process below presents the encoding, with bold weight values indicating that the desired $W = 8$ is attained, the underline part represents the prefix and the bold symbol is the digit u.

z	$\boldsymbol{x} \oplus_q \boldsymbol{b}(z) = \boldsymbol{y}$	$\boldsymbol{c} = [u\|\boldsymbol{g}\|\boldsymbol{y}]$	$w(\boldsymbol{c})$
0	$212 \oplus_3 000 = 212$	**0**00212	5
1	$212 \oplus_3 100 = 012$	**0**01012	4
2	$212 \oplus_3 110 = 022$	**2**02022	**8**
3	$212 \oplus_3 111 = 020$	**0**12020	5
4	$212 \oplus_3 211 = 120$	**0**11120	5
5	$212 \oplus_3 221 = 100$	**0**10100	2
6	$212 \oplus_3 222 = 101$	**0**20101	4
7	$212 \oplus_3 022 = 201$	**2**21201	**8**
8	$212 \oplus_3 002 = 211$	**0**22211	**8**

In this case, there are three occurrences of a $(6, 3, 8, 3)$ CW sequence: **2**02022, **2**21201 and **0**22211. Fig. 3 presents the weight progression for this example.

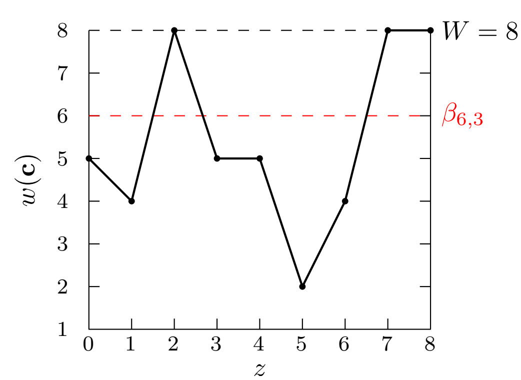

Fig. 3. Weight progression of $(6, 3, 8, 3)$ CW sequence of Example 3.

IV. CONSTRUCTION OF q-ARY CONSTANT WEIGHT SEQUENCES WITH EXTENDED WEIGHT RANGE

In the previous section, the construction of q-ary CW sequences was achieved with a weight range shown in (2). However, because of the limited interval, we will present an approach to extend this range.

The method consists of appending a redundant vector $\boldsymbol{u}$ of length e to $\boldsymbol{c}' = [\boldsymbol{g}|\boldsymbol{y}]$, then the output sequence becomes $\boldsymbol{c} = [\boldsymbol{u}|\boldsymbol{g}|\boldsymbol{y}]$. This leads to (n,k,W,q) CW sequences where $n = k + r' + e$.

In general, generating (n,k,W,q) CW sequence from a q-ary information sequence of length k with the redundant vector $\boldsymbol{u}$ of length e will lead to an increase of weight range. Combining (2) and $w(\boldsymbol{u}) \in [0, e(q-1)]$ results in a CW sequence $\boldsymbol{c} = [\boldsymbol{u}|\boldsymbol{g}|\boldsymbol{y}]$ of weight W is such that

$$\frac{(k-2)(q-1)}{2} \leq W \leq \frac{(k+2r'+2e+2)(q-1)}{2}. \qquad (3)$$

Example 3 can then be viewed as a special case with $e = 1$.

Theorem 2: Any q-ary information sequence of length k can generate an (n,k,W,q) CW sequence where $W \in \left[\frac{(k-2)(q-1)}{2}, \frac{(k+2r'+2e+1)(q-1)}{2}\right]$.

Theorem 2 can be proved using a similar argument as in Theorem 1. However, the condition $k = q^t$, where t is a positive integer must still be fulfilled in order to observe equality between the cardinality of (r',q)-Gray code prefixes and that of the weighting sequences.

The redundant vector $\boldsymbol{u} = u_1u_2\ldots u_e$ is such that $u_i \in \{0,1,\ldots,q-1\}$ and $w(\boldsymbol{u}) = W - w(\boldsymbol{c}')$ if and only if $W \geq w(\boldsymbol{c}')$, otherwise $\boldsymbol{u} = \mathbf{0}$. Table II presents some parameters evaluation, that is values of total length n and the achievable range of weights W, given an alphabet size q and an information sequence length k.

Example 4: Consider the same ternary information sequence $\boldsymbol{x} = 212$ of length 3 as in Example 3. We would like to generate a $(7,3,12,3)$ CW sequence of weight $W = 12$ and $n = 7$ as described in Table II.

We observe from (2) that the information sequence 212 can only generate (n,k,W,q) CW sequences with $W \in [2,10]$. In order to extend this range of weight, we append a ternary redundant vector $\boldsymbol{u}$ of length e to $\boldsymbol{c}'$. For $e = 2$, (3) stipulates that $W \in [2,12]$, therefore the weight $W = 12$ can be obtained.

- There are nine possible vectors, $\boldsymbol{u}$ of length two that can be appended to $\boldsymbol{c}'$ as follow: 00, 01, 02, 10, 11, 12, 20, 21, 22.

TABLE II
PARAMETERS EVALUATION

	t	$k=q^t$	W	n	r'	e
	2	4	$[4,9]$	10	3	3
$q=2$	3	8	$[7,13]$	16	4	4
	4	16	$[11,17]$	24	5	3
	1	3	$[5,12]$	7	2	2
$q=3$	2	9	$[13,22]$	15	3	3
	3	27	$[32,42]$	34	4	3
	1	4	$[12,20]$	8	2	2
$q=4$	2	16	$[30,43]$	22	3	3
	3	64	$[105,121]$	72	4	4

- The redundant vector weight is such that $w(\boldsymbol{u}) = 12 - w(\boldsymbol{c}')$ for $W \geq w(\boldsymbol{c}')$.
- The length of Gray code prefixes is $r' = \log_3 3 + 1 = 2$.
- Below we repeat the process defined in Section III:

z	$\boldsymbol{x} \oplus_q \boldsymbol{b}(z) = \boldsymbol{y}$	$\boldsymbol{c} = [\boldsymbol{u}\vert\boldsymbol{g}\vert\boldsymbol{y}]$	$w(\boldsymbol{c})$
0	$212 \oplus_3 000 = 212$	$\underline{\mathbf{00}00}212$	5
1	$212 \oplus_3 100 = 012$	$\underline{\mathbf{00}01}012$	4
2	$212 \oplus_3 110 = 022$	$\underline{\mathbf{00}02}022$	6
3	$212 \oplus_3 111 = 020$	$\underline{\mathbf{00}12}020$	5
4	$212 \oplus_3 211 = 120$	$\underline{\mathbf{00}11}120$	5
5	$212 \oplus_3 221 = 100$	$\underline{\mathbf{00}10}100$	2
6	$212 \oplus_3 222 = 101$	$\underline{\mathbf{00}20}101$	4
7	$212 \oplus_3 022 = 201$	$\underline{\mathbf{00}21}201$	6
8	$212 \oplus_3 002 = 211$	$\underline{\mathbf{22}22}211$	**12**

The underline sequence represents the overall prefix where the bold part is the redundant vector $\boldsymbol{u}$ and the rest is the Gray code prefix. We observe that by adding a redundant vector of length $e = 2$ in the process, there is one occurrence of a $(7,3,12,3)$ CW sequence which is $\underline{\mathbf{22}22}211$. Fig. 4 presents the weight progression for this example.

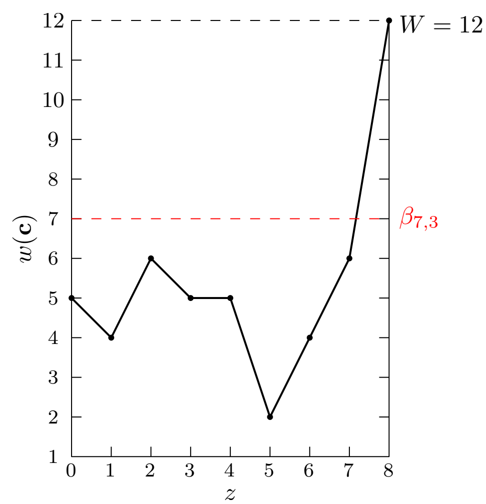

Fig. 4. Weight progression for Example 4.

We observe that by adding a redundant vector of length $e = 2$, we increased the weight range from $[2,10]$ to $[2,12]$, that is the weight range increases proportionally with the length of the redundant symbols.

V. DECODING OF q-ARY CONSTANT WEIGHT SEQUENCES

The decoding of (n,k,W,q) CW sequences follows the same process as the decoding of balanced sequences presented in [12]. This process consists of the following steps:

1) The redundant vector $\boldsymbol{u}$ is dropped, then the r' symbols are extracted as the Gray code prefix and converted to z as presented in Section II-B.
2) z is used to determine the parameters s and p, then $\boldsymbol{b}(s,p)$ can be derived.
3) Finally, the original sequence is recovered through $\boldsymbol{x} = \boldsymbol{y} \ominus_q \boldsymbol{b}(s,p)$.

Example 5: We want to decode the $(7,4,14,4)$ CW sequence, $\underline{\mathbf{2}31}3113$.

The illustration of the decoding process for the $(2,4)$-Gray code prefixes is presented in Table III, the bold row represents the decoding of the extracted Gray code prefix, 31.

- The redundant vector $\boldsymbol{u} = 2$ is dropped. Then the Gray code sequence of length 2, is extracted as 31.
- The Gray code $\boldsymbol{g} = 31$ corresponds to $\boldsymbol{d} = 32$, and index $z = 14$ according to Table III. This implies that $s = 3$ and $p = 2$, therefore $\boldsymbol{b}(3,2) = 0033$.
- Finally, the information sequence is recovered as
$$\boldsymbol{x} = \boldsymbol{y} \ominus_q \boldsymbol{b}(s,p) = 3113 \ominus_3 0033 = 3120.$$

VI. Redundancy and Complexity Analysis

The redundancy of the presented method comes from the Gray code prefix and the redundant vector. The overall redundancy is $r = \log_q k + e + 1 \Rightarrow k = q^{r-1-e}$. Therefore, the total length of encoded CW sequence is $n = k + \log_q k + e + 1$. An encoding and decoding algorithm for balanced q-ary sequences based on Gray code prefixes was presented in [12], with a redundancy of $r = \log_q k + 1$, which was compared against existing constructions. However, the two above redundancies differ only with the parameter e which gets larger for high values of W and equals zero for $W = \beta_{n,q}$.

Table IV presents the comparison of cardinalities for the full set of CW versus the ones for information sequences. Here $\mathcal{N}_1$ represents the cardinality of q-ary CW sequences for specific W of length n while $\mathcal{N}_2$ represents the cardinality of q-ary information sequences of length k. To have a q-ary CW sequence of weight W and length k, one clearly requires enough parity bits r such that $\mathcal{N}_1 \geq \mathcal{N}_2 = q^k$, where $n = k + r$.

Our method requires $\mathcal{O}(qk \log_q k)$ digit operations for the encoding and $\mathcal{O}(k)$ digit operations for the decoding process; this complexity is similar to the one in construction [12].

TABLE III
Decoding of $(2,4)$-Gray code

Gray code ($\boldsymbol{g}$)	Sequence ($\boldsymbol{d}$)	z	s, p	$\boldsymbol{b}(s,p)$
00	00	0	$0,0$	0000
01	01	1	$0,1$	1000
02	02	2	$0,2$	1100
03	03	3	$0,3$	1110
13	10	4	$1,0$	1111
12	11	5	$1,1$	2111
11	12	6	$1,2$	2211
10	13	7	$1,3$	2221
20	20	8	$2,0$	2222
21	21	9	$2,1$	3222
22	22	10	$2,2$	3322
23	23	11	$2,3$	3332
33	30	12	$3,0$	3333
32	31	13	$3,1$	0333
31	**32**	**14**	$\mathbf{3,2}$	**0033**
30	33	15	$3,3$	0003

TABLE IV
Comparison of our construction against the full set of CW and balanced sequences

W		q	n	k	$\mathcal{N}_1$	$\mathcal{N}_2$
$\beta_{n,q} - q + 1$	3	2	7	4	35	16
	5	2	12	8	792	256
	10	2	21	16	352716	65536
	3	3	5	3	30	27
	10	3	12	9	58278	19683
	6	4	6	4	336	256
$\beta_{n,q}$	4	2	7	4	35	16
	6	2	12	8	924	256
	11	2	21	16	352716	65536
	5	3	5	3	51	27
	12	3	12	9	737789	19683
	9	4	4	6	580	256
$\beta_{n,q} + q$	6	2	8	4	28	16
	9	2	13	8	715	256
	13	2	22	16	497420	65536
	9	3	6	3	50	27
	16	3	13	9	129844	19683
	15	4	7	4	728	256

VII. Conclusion

An efficient algorithm was proposed for encoding and decoding (n,k,W,q) CW sequences based on Gray code prefixes, with a simple method to extend the achievable weight range, if necessary. The construction does not make use of memory-consuming lookup tables, and only simple operations such as addition and subtraction are needed, and most of the decoding process can be performed in parallel. Seeing as the proposed method is only applicable to information sequences of length k where $k = q^t$, the most obvious improvement would be to extend this algorithm to the case where $k \neq q^t$.

References

[1] T. Etzion and A. Vardy, “A new construction for constant weight codes”, *Int. Symp. Inform. Theory Applic.*, Melbourne, Australia, Oct. 26–29, 2014, pp. 338–342.

[2] V. Skachek and K. A. S. Immink, “Constant weight codes: an approach based on Knuth’s balancing method,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 5, pp. 908–918, Apr. 2014.

[3] D. E. Knuth, “Efficient balanced codes,” *IEEE Trans. Inform. Theory*, vol. 32, no. 1, pp. 51–53, Jan. 1986.

[4] T. Kaida and J. Zheng, “A note on constant weight sequences over q-ary from cyclic difference sets,” in *Proc. Int. Workshop Signal Design Applic. Commun.*, Fukuoka City, Japan, Oct. 19–23, 2009, pp. 112–113.

[5] N. Li, X. Zheng and L. Hu, “Binary constant weight codes based on cyclic difference sets,” in *IEICE Trans. Fundamentals Electron. Commun. Comput. Sci.*, vol. E91-A, no. 5, pp. 1288–1292, Jan. 2010.

[6] T. Kaida and J. Zheng, “On q-ary constant weight sequences from cyclic difference sets,” in *Proc. Int. Workshop Signal Design Applic. Commun.*, Guilin, China, Oct. 10–14, 2011, pp. 126–129.

[7] Y. M. Chee and S. Ling, “Constructions for q-ary constant-weight codes,” *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 135–146, Jan. 2007.

[8] F. Zheng, “Properties of m-sequence and construction of constant weight codes,” *IEICE Trans. Fundamentals Electron. Commun. Comput. Sci.*, vol. E88-A, no. 12, pp. 3675–3676, Dec. 2005.

[9] T. G. Swart and J. H. Weber, “Efficient balancing of q-ary sequences with parallel decoding,” in *Proc. IEEE Int. Symp. Inform. Theory*, Seoul, Korea, Jun. 28–Jul. 3, 2009, pp. 1564–1568.

[10] F. Gray, “Pulse code communication,” *U. S. Patent 2632058*, Mar. 1953.

[11] D.-J. Guan, “Generalized Gray codes with applications,” in *Proc. Nat. Sci. Council, Rep. of China, Part A*, vol. 22, no. 6, pp. 841–848, 1998.

[12] E. N. Mambou and T. G. Swart, “Encoding and decoding of balanced q-ary sequences using a Gray code prefix,” in *Proc. IEEE Int. Symp. Inform. Theory*, Barcelona, Spain, Jul. 10–15, 2016, pp. 380–384.