Spectrum of Sizes for Perfect Burst Deletion-Correcting Codes

Yeow Meng Chee, Yang Li, and Xiande Zhang
School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
email: {ymchee, li_yang, xiandezhang}@ntu.edu.sg

***Abstract*—Perfect deletion-correcting codes of the same length over the same alphabet can have different sizes. The interesting problem of determining the possible sizes of perfect deletion-correcting codes has previously been studied. In this paper, we study the corresponding problem for burst deletion-correcting codes. We completely determine the spectrum of sizes for perfect burst deletion-correcting codes for certain classes of parameters and also construct new classes of perfect deletion-correcting codes.**

I. Introduction

Deletion errors are very common in channels such as magnetic and optical recording [1], packet-switched communication [2], and DNA replication [3]. Codes that are capable of correcting deletions were first introduced by Levenshteĭn [4] in the 1960s, and have been systematically studied since [5]–[16].

Most of the work on deletion-correcting codes have focused on the channels where deletions are randomly scattered in a codeword. However, in certain memory systems, such as spacecraft memories subject to soft upsets [17], and digital video recording [18], [19], the effect of burst deletions (that is, a run of deletions that occur in consecutive components of a codeword) is significant. It is therefore important to consider codes for combating burst deletion errors, which we call *burst deletion-correcting codes*. Levenshteĭn [4] presented asymptotic upper bounds on the size of such codes and constructed asymptotically optimally codes capable of correcting burst of up to two deletions. Codes capable of correcting bursts of higher number of deletions are considered recently by Cheng et al. [20].

Sloane [11] observed that unlike error-correcting codes, perfect deletion-correcting codes of the same length over the same alphabet can surprisingly have different sizes. The same holds for perfect burst deletion-correcting codes. It is interesting, therefore, to ask for the possible sizes a perfect deletion-correcting code and a perfect burst deletion-correcting code can have. The spectrum of possible sizes for perfect q-ary deletion-correcting codes of length n, capable of correcting up to t deletions, was determined when $(t, n) \in \{(1, 3), (2, 4)\}$ for almost all q by Chee et al. [14].

In this paper, we initiate the determination of the spectrum of possible sizes for perfect burst deletion-correcting codes. In particular, for perfect q-ary burst deletion-correcting code of length n, capable of correcting a burst deletion of length t, we completely determine the spectrum of their sizes when

(i) $n \geq 4$ and $t = n - 2$.
(ii) $n \geq 6$, $t = n - 3$, and q even.

II. Preliminaries and Notation

For $a \in \mathbb{Z}$ and $S \subseteq \mathbb{Z}$, aS denotes the set $\{a \cdot s : s \in S\}$.

Let X be a finite set of q elements and n be a positive integer. A set $\mathcal{C} \subseteq X^n$ is called a *q-ary code of length n*. The set X is called the *alphabet* of the code $\mathcal{C}$ and the elements of $\mathcal{C}$ are called *codewords*. The *size* of $\mathcal{C}$ is $|\mathcal{C}|$, the number of codewords it contains.

For $x \in X^n$ and $0 \leq t \leq n$, let $D_t(x)$ denote the set of *t-th order descendants*, that is, the set of $y \in X^{n-t}$ that are obtained if any t components are deleted from x. The q-ary code $\mathcal{C} \subset X^n$ is said to be *t-deletion-correcting* if $D_t(x) \cap D_t(y) = \emptyset$ for all distinct $x, y \in \mathcal{C}$. We call such a code an *$(n, t)_q$-deletion-correcting code* and denote it by $(n, t)_q$-DCC. An $(n, t)_q$-DCC is called *optimal* if it has maximum size among all the codes with same parameters.

An $(n, t)_q$-DCC $\mathcal{C} \subseteq X^n$ is *perfect* if the balls $D_t(x)$, $x \in \mathcal{C}$, partition X^{n-t}. As observed by Sloane [11], perfect $(n, t)_q$-DCCs can have different sizes because the balls $D_t(x)$ have different sizes depending on x.

Define the spectrum of sizes of a perfect $(n, t)_q$-DCC to be

$$\mathsf{Spec}(q, n, t) = \{|\mathcal{C}| : \mathcal{C} \text{ is a perfect } (n, t)_q\text{-DCC}\}.$$

Chee et al. [14] determined that

(i) $\mathsf{Spec}(q, n, n) = \{1\}$;
(ii) $\mathsf{Spec}(q, n, n - 1) = [\lceil q/n \rceil, q]$;
(iii) $\mathsf{Spec}(q, 3, 1) = [\lceil q^2/3 \rceil, \lfloor q(q + 2)/3 \rfloor]$;
(iv) $\mathsf{Spec}(q, 4, 2) = [\lceil \frac{q}{4} \lceil \frac{2q}{3} \rceil \rceil, \lfloor \frac{q}{4} \lfloor \frac{2(q-1)}{3} \rfloor \rfloor + q]$,

where $[a, b]$ denotes the interval of integers $\{a, a + 1, \ldots, b\}$, for positive integers $a < b$.

Recently, it was shown that there exists a perfect $(4, 1)_q$-DCC for all q [13], [15], [21]. For even q, the perfect $(4, 1)_q$-DCC constructed by Kim et al. [15] is also optimal. However, it seems difficult to determine $\mathsf{Spec}(q, 4, 1)$ completely.

We now introduce notations for codes that can correct burst deletion errors. For $x \in X^n$ and $0 \leq t \leq n$, let $\bar{D}_t(x)$ denote the set of all $y \in X^{n-t}$ that are obtained by deleting t consecutive components from x. A q-ary code $\mathcal{C} \subseteq X^n$ is said to be *capable of correcting a burst deletion of length t* if $\bar{D}_t(x) \cap \bar{D}_t(y) = \emptyset$ for all distinct $x, y \in \mathcal{C}$. We call such a code an *$(n, t)_q$-burst deletion-correcting code* and denote it by $(n, t)_q$-BDCC. Similar to deletion-correcting codes, an

978-1-4673-7704-1/15/$31.00 ©2015 IEEE

$(n,t)_q$-BDCC is *optimal* if it has maximum size among all $(n,t)_q$-BDCCs. An $(n,t)_q$-BDCC is *perfect* if the balls $\bar{D}_t(x)$, $x \in \mathcal{C}$, partition X^{n-t}. Since $\max_{x\in X^n} |\bar{D}_t(x)| = n-t+1$, a lower bound on the size of a perfect $(n,t)_q$-BDCC is $q^{n-t}/(n-t+1)$.

We define the spectrum of sizes of a perfect $(n,t)_q$-BDCC as

$$\mathsf{Spec}^{\mathsf{B}}(q,n,t) = \{|\mathcal{C}| : \mathcal{C} \text{ is a perfect } (n,t)_q\text{-BDCC}\}.$$

Then $\mathsf{Spec}^{\mathsf{B}}(q,n,1) = \mathsf{Spec}(q,n,1)$. We also have $\mathsf{Spec}^{\mathsf{B}}(q,n,n) = \{1\}$ and $\mathsf{Spec}^{\mathsf{B}}(q,n,n-1) = [\lceil q/2\rceil, q]$.

Our results in this paper can be summarized as follows:

(i) $\mathsf{Spec}^{\mathsf{B}}(q,n,n-2) = [\lceil q^2/3\rceil, q^2]$ for all $n \geq 4$.

(ii) $\mathsf{Spec}^{\mathsf{B}}(q,5,2) \supseteq q[\lceil q^2/3\rceil, \lfloor q(q+2)/3\rfloor]$ for all q, and $\inf \mathsf{Spec}^{\mathsf{B}}(q,5,2) = q^3/4$ for even q.

(iii) Let $n \geq 6$. Then $\mathsf{Spec}^{\mathsf{B}}(q,n,n-3) \supseteq [q\lceil q^2/3\rceil, q^3]$ for odd q, and $\mathsf{Spec}^{\mathsf{B}}(q,n,n-3) = [q^3/4, q^3]$ for even q.

(iv) There exist perfect $(4,1)_q$-DCCs of size $q(q^2+q+4)/4$ and $\frac{q^3+5q^2/4}{4}$ for all $q \equiv 0 \bmod 8$, and of size $\frac{q^3+3q^2/2}{4}$ for all $q \equiv 0 \bmod 4$.

III. Constructions from Shorter Deletion-Correcting Codes

This section is devoted to the study of $\mathsf{Spec}^{\mathsf{B}}(q,n,t)$, with $t \in \{n-3, n-2\}$, by constructing burst deletion-correcting codes from shorter deletion-correcting codes and burst deletion-correcting codes.

A. Determination of $\mathsf{Spec}^{\mathsf{B}}(q,n,n-2)$

Recall that when $n=3$, $\mathsf{Spec}^{\mathsf{B}}(q,3,1) = \mathsf{Spec}(q,3,1) = [\lceil \frac{q^2}{3}\rceil, \lfloor \frac{q^2+2q}{3}\rfloor]$. Thus, we assume $n \geq 4$.

Let $x = (x_1, x_2, \ldots, x_{n-1}, x_n) \in X^n$. Then $\bar{D}_{n-2}(x) = \{(x_1,x_2),(x_1,x_n),(x_{n-1},x_n)\}$, which is independent of the symbols $x_3,\ldots,x_{n-2}$. So any $(n,n-2)_q$-BDCC of length $n>4$ can be obtained from a $(4,2)_q$-BDCC by inserting any $n-4$ symbols into the center of each codeword. Further, since $1 \leq |\bar{D}_{n-2}(x)| \leq 3$, we have $\mathsf{Spec}^{\mathsf{B}}(q,n,n-2) \subseteq [\lceil \frac{q^2}{3}\rceil, q^2]$ for all $n \geq 4$.

By the analysis above, we only need to consider the case $n=4$.

First, we construct a perfect $(4,2)_q$-BDCC from a perfect $(3,1)_q$-DCC as follows. Suppose $\mathcal{D} \subseteq X^3$ is a perfect $(3,1)_q$-DCC. For each $x = (a,b,c) \in \mathcal{D}$, let $\epsilon_x = (a,b,b,c)$. Note that $\bar{D}_2(\epsilon_x) = D_1(x)$. Then $\mathcal{C} = \{\epsilon_x : x \in \mathcal{D}\}$ is a perfect $(4,2)_q$-BDCC. Since $\lceil \frac{q^2}{3}\rceil \in \mathsf{Spec}(q,3,1)$, we have a perfect $(4,2)_q$-BDCC of size $\lceil \frac{q^2}{3}\rceil$.

Next, we construct perfect $(4,2)_q$-BDCCs with bigger sizes as follows. For each word $x = (a,b,c,d)$ with $(a,b) \neq (c,d)$, let

$$A_x = \begin{cases} \{(a,d,c,d),(a,b,a,b)\}, & \text{if } b \neq d \\ \{(a,b,a,d),(c,d,c,d)\}, & \text{otherwise.} \end{cases}$$

Note that $\bar{D}_2(y)$, $y \in A_x$ partition $\bar{D}_2(x)$. Then $(\mathcal{C}\setminus\{x\}) \cup A_x$ is also a perfect $(4,2)_q$-BDCC but with size increased by one. Continuing this procedure until all the codewords have the form (a,b,a,b), we can obtain a perfect $(4,2)_q$-BDCC for any size in $[\lceil \frac{q^2}{3}\rceil, q^2]$. Hence, $\mathsf{Spec}^{\mathsf{B}}(q,4,2) = [\lceil \frac{q^2}{3}\rceil, q^2]$ and consequently $\mathsf{Spec}^{\mathsf{B}}(q,n,n-2) = [\lceil \frac{q^2}{3}\rceil, q^2]$ for all $n \geq 4$.

B. $\mathsf{Spec}^{\mathsf{B}}(q,5,2)$

We show that $\mathsf{Spec}^{\mathsf{B}}(q,5,2) \supseteq q\mathsf{Spec}(q,3,1)$ in this subsection.

Let $x = (x_1,x_2,x_3,x_4,x_5) \in X^5$, then $\bar{D}_2(x) = \{(x_1, x_2,x_3),(x_1,x_2,x_5),(x_1,x_4,x_5),(x_3,x_4,x_5)\}$. Note that x_2 and x_4 are not related in $\bar{D}_2(x)$. Hence, we consider a special case that $x_2 = x_4$, where $\bar{D}_2(x) = \{(x_1,x_2,x_3),(x_1, x_2,x_5),(x_3,x_2,x_5)\}$.

Now we construct a perfect $(5,2)_q$-BDCC from a perfect $(3,1)_q$-DCC as follows. Suppose $\mathcal{D} \subseteq X^3$ is a perfect $(3,1)_q$-DCC. For each $d \in X$, let $E_d = \{(a,d,b,d,c) : (a,b,c) \in \mathcal{D}\}$. Denote $T_d = X \times \{d\} \times X$. Note that T_d, $d \in X$ partition X^3, while $\bar{D}_2(x)$, $x \in E_d$ partition T_d. Hence, $\mathcal{C} = \cup_{d\in X} E_d$ is a perfect $(5,2)_q$-BDCC of size $q|\mathcal{D}|$. Thus, we have $\mathsf{Spec}^{\mathsf{B}}(q,5,2) \supseteq q\mathsf{Spec}(q,3,1) = q[\lceil \frac{q^2}{3}\rceil, \lfloor \frac{q^2+2q}{3}\rfloor]$.

C. $\mathsf{Spec}^{\mathsf{B}}(q,n,n-3)$ *for* $n \geq 6$

Let $x = (x_1,x_2,\ldots,x_{n-1},x_n) \in X^n$. Then $\bar{D}_{n-3}(x) = \{(x_1,x_2,x_3),(x_1,x_2,x_n),(x_1,x_{n-1},x_n),(x_{n-2},x_{n-1},x_n)\}$, which is independent of the symbols $x_4,\ldots,x_{n-3}$. So any $(n,n-3)_q$-BDCC of length $n>6$ can be obtained from a $(6,3)_q$-BDCC by inserting any $n-6$ symbols into the center of each codeword. So we only consider $n=6$ here. Let $x = (x_1,x_2,\ldots,x_5,x_6) \in X^6$. Note that x_2 and x_5 are not related in $\bar{D}_3(x)$. As in Section III-B, we consider a special case when $x_2 = x_5$, and $\bar{D}_3(x)$ is of size three.

First, we construct a perfect $(6,3)_q$-BDCC from a perfect $(4,2)_q$-BDCC as follows. Suppose $\mathcal{D} \subseteq X^4$ is a perfect $(4,2)_q$-BDCC. For each $e \in X$, let $E_e = \{(a,e,b,c,e,d) : (a,b,c,d) \in \mathcal{D}\}$. Denote $T_e = X \times \{e\} \times X$. Note that T_e, $e \in X$ partition X^3, while $\bar{D}_3(x)$, $x \in E_e$ partition T_e. Hence, $\mathcal{C} = \cup_{e\in X} E_e$ is a perfect $(6,3)_q$-BDCC of size $q|\mathcal{D}|$. By $\mathsf{Spec}(q,4,2)$, we have a perfect $(6,3)_q$-BDCC of size $q\lceil \frac{q^2}{3}\rceil$.

Second, we construct perfect $(6,3)_q$-BDCCs with bigger sizes as follows. For each word $x = (a,b,c,d,e,f)$ with $(a,b,c) \neq (d,e,f)$, let

$$A_x = \begin{cases} \{(a,b,c,a,e,f),(d,e,f,d,e,f)\}, & \text{if } a \neq d \\ \{(a,b,c,a,b,f),(a,e,f,a,e,f)\}, & \text{if } a = d \text{ and } b \neq e \\ \{(a,b,c,a,b,c),(a,b,f,a,b,f)\}, & \text{otherwise.} \end{cases}$$

Note that $\bar{D}_3(y)$, $y \in A_x$ partition $\bar{D}_3(x)$. Then $(\mathcal{C}\setminus\{x\}) \cup A_x$ is also a perfect $(6,3)_q$-BDCC but with code size being increased by one. Continuing this procedure until all the codewords have the form (a,b,c,a,b,c), we can obtain a perfect $(6,3)_q$-BDCC for any size in $[q\lceil \frac{q^2}{3}\rceil, q^3]$. Hence, $\mathsf{Spec}^{\mathsf{B}}(q,6,3) \supseteq [q\lceil \frac{q^2}{3}\rceil, q^3]$ and consequently $\mathsf{Spec}^{\mathsf{B}}(q,n,n-3) \supseteq [q\lceil \frac{q^2}{3}\rceil, q^3]$ for all $n \geq 6$.

IV. PERFECT $(n, n-3)_q$-BDCCS OF MINIMUM SIZE

Note that the smallest possible size of a perfect $(n, n-3)_q$-BDCC is $\lceil q^3/4 \rceil$. In this section, we give recursive constructions of perfect $(n, n-3)_q$-BDCCs from codes over smaller alphabets, which yield the existence of perfect $(n, n-3)_q$-BDCCs of size $\frac{q^3}{4}$ for all $n \geq 5$ and even q.

Suppose X, Y are two finite sets. For brevity, we write a_b for an element $(a, b) \in X \times Y$. If $x = (x_1, \ldots, x_n) \in X^n$ and $y = (y_1, \ldots, y_n) \in Y^n$, define $x \circ y = (x_{1y_1}, \ldots, x_{ny_n})$.

Theorem 1. *If $\mathcal{C}$ is a perfect $(5, 2)_q$-BDCC of size s such that $i \neq k$ and $k \neq m$ for each $(i, j, k, l, m) \in \mathcal{C}$, then there exits a perfect $(5, 2)_{vq}$-BDCC of size $v^3 s$ for any positive integer v.*

Proof: Let $\mathcal{C} \subseteq X^5$ be a perfect $(5, 2)_q$-BDCC of size s such that $i \neq k$ and $k \neq m$ for each $(i, j, k, l, m) \in \mathcal{C}$.

Let Y be a cyclic additive group of order v, and define $\mathcal{I} \subseteq Y^5$ so that

$$\mathcal{I} = \{(a, b, c, b, a + c) : a, b, c \in Y\}.$$

Let $X' = Y \times X$. For each $x \in \mathcal{C}$, define

$$\mathcal{I}_x = \{y \circ x : y \in \mathcal{I}\}.$$

Let $\mathcal{D} = \cup_{x \in \mathcal{C}} \mathcal{I}_x$. It is easy to see that $\mathcal{D}$ is of size $v^3 s$. We claim that $\mathcal{D}$ is a perfect $(5, 2)_q$-BDCC.

First, we prove $\mathcal{D}$ is a $(5, 2)_q$-BDCC, that is, $\bar{D}_2(\mathsf{u}) \cap \bar{D}_2(\mathsf{u}') = \emptyset$ for any two different codewords $\mathsf{u} = y \circ x$ and $\mathsf{u}' = y' \circ x'$. If $x \neq x'$, then $\bar{D}_2(\mathsf{u}) \cap \bar{D}_2(\mathsf{u}') = \emptyset$ since $\bar{D}_2(x) \cap \bar{D}_2(x') = \emptyset$. If $x = x'$, then $y \neq y'$. Suppose that $x = (i, j, k, l, m)$, $y = (a, b, c, b, a + c)$ and $y' = (a', b', c', b', a' + c')$. Then

$$\begin{aligned}\bar{D}_2(\mathsf{u}) = &\{(a_i, b_j, c_k), (a_i, b_j, (a + c)_m), \\ &(a_i, b_l, (a + c)_m), (c_k, b_l, (a + c)_m)\}\end{aligned}$$

and

$$\begin{aligned}\bar{D}_2(\mathsf{u}') = &\{(a'_i, b'_j, c'_k), (a'_i, b'_j, (a' + c')_m), \\ &(a'_i, b'_l, (a' + c')_m), (c'_k, b'_l, (a' + c')_m)\}.\end{aligned}$$

By the assumption that $i \neq k$, $k \neq m$ and $(a, b, c) \neq (a', b', c')$, it is easy to verify that $\bar{D}_2(\mathsf{u}) \cap \bar{D}_2(\mathsf{u}') = \emptyset$.

Next, we prove $\mathcal{D}$ is a perfect code, that is, any element of X'^3 belongs to $\bar{D}_2(\mathsf{u})$ for some $\mathsf{u} \in \mathcal{D}$. This is true since $\mathcal{C}$ is a code and after any deletion of two consecutive components of elements of $\mathcal{I}$, the derived 3-tuples cover exactly all the elements of Y^3. ■

Let $\mathsf{u} = (0, 1, 1, 0, 0)$ and $\mathsf{u}' = (1, 1, 0, 0, 1)$. Then $\{\mathsf{u}, \mathsf{u}'\}$ is a perfect $(5, 2)_2$-BDCC of size two. Applying Theorem 1 gives the following.

Corollary 1. *There exists a perfect $(5, 2)_q$-BDCC of size $q^3/4$, for all even q.*

Theorem 2. *Let $n \geq 6$. If there exits a perfect $(n, n-3)_q$-BDCC of size s, then there exists a perfect $(n, n-3)_{vq}$-BDCC of size $v^3 s$, for any positive integer v.*

Proof: It suffices to prove the case when $n = 6$. Let $\mathcal{C} \subseteq X^6$ be a perfect $(6, 3)_q$-BDCC of size s, and Y be an alphabet of size v. Define $\mathcal{I} \subseteq Y^6$ so that

$$\mathcal{I} = \{(a, b, c, a, b, c) : a, b, c \in Y\}.$$

Let $X' = Y \times X$. For each $x \in \mathcal{C}$, define

$$\mathcal{I}_x = \{y \circ x : y \in \mathcal{I}\}.$$

Let $\mathcal{D} = \cup_{x \in \mathcal{C}} \mathcal{I}_x$. The proof that $\mathcal{D}$ is a perfect $(6, 3)_q$-BDCC of size $v^3 s$ is similar to that in Theorem 1, and is omitted. ■

Let $\mathsf{u}_1 = (0, 1, 1, 1, 0, 0)$ and $\mathsf{u}_2 = (1, 1, 0, 0, 0, 1)$. Then $\{\mathsf{u}, \mathsf{u}'\}$ is a perfect $(6, 3)_2$-BDCC of size two. Applying Theorem 2, we obtain a perfect $(6, 3)_q$-BDCC of size $q^3/4$ for all even q. Further, we can determine the spectrum of possible sizes of perfect $(6, 3)_q$-BDCCs by using the same technique as in Section III-C.

Corollary 2. $\mathsf{Spec}^{\mathsf{B}}(q, n, n-3) = [q^3/4, q^3]$ *for all $n \geq 6$ and even q.*

V. A CONSTRUCTION FOR $(4, 1)_q$-DCCS

In this section, we apply similar ideas as in Section IV to construct single-deletion-correcting codes.

An *orthogonal array* with k constraints, v levels, and strength three, denoted $\mathrm{OA}(3, k, v)$, is a $k \times v^3$ array with entries from a set of $v \geq 2$ symbols, having the property that in every $3 \times v^3$ submatrix, every 3×1 column vector appears exactly once. It is known that an $\mathrm{OA}(3, 4, v)$ exists for all v [22].

Proposition 1. *Suppose $\mathcal{C}$ is a $(4, 1)_q$-DCC of size s. If each codeword $(i, j, k, l) \in \mathcal{C}$ satisfies that $i \neq j$, $j \neq l$ and $l \neq k$, then there exits a $(4, 1)_{vq}$-DCC of size $v^3 s$.*

Proof: Let $\mathcal{A}$ be an $\mathrm{OA}(3, 4, v)$ with entries from X, and let $\mathcal{C}$ be a $(4, 1)_q$-DCC of size s over Y. Note that $|X| = v$ and $|Y| = q$. Now construct a code over $X \times Y$ as follows. For each column $B = (a, b, c, d)^T$ in $\mathcal{A}$, define $\mathcal{D}_B \subseteq (X \times Y)^4$ so that

$$\mathcal{D}_B = \{B \circ x : x \in \mathcal{C}\}.$$

Let $\mathcal{D} = \cup_{B \in \mathcal{A}} \mathcal{D}_B$, which has size $v^3 s$. It is easy to verify that $\mathcal{D}$ is a $(4, 1)_{vq}$-DCC. ■

Note that we do not use perfect codes in Proposition 1 since any perfect $(4, 1)_q$-DCC does not satisfy the hypothesis. In fact, if a code is perfect, then for any symbol a, there exists a codeword x such that $(a, a, a) \in D_1(x)$. This codeword x must have consecutive equal entries. However, we can still produce perfect DCCs based on the construction in Proposition 1.

For any finite set Y, the set of all ordered k-tuples of Y with distinct components is denoted $\overline{\binom{Y}{k}}$. Levenshteĭn [21] shows that a $(4, 1)_q$-DCC $\mathcal{C}$ over Y, with the property that $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{C}$ partition $\overline{\binom{Y}{3}}$, exists if and only if q is even. Such a $(4, 1)_q$-DCC satisfies the hypothesis of Proposition 1 and has size $s = q(q-1)(q-2)/4$. Let us denote this $(4, 1)_q$-DCC as $L(q)$.

Theorem 3. *Let $v, q > 0$ be even integers. Suppose there exits a perfect $(4,1)_v$-DCC of size s. Then there exits a perfect $(4,1)_{vq}$-DCC of size $\frac{1}{4}q(q-1)(q-2)v^3 + qs + q(q-1)v^2 + \frac{3}{4}q(q-1)v^2(v-1)$.*

Proof: Let X be a set of size v and Y be a set of size q. We construct a perfect $(4,1)_{vq}$-DCC $\mathcal{C} \subseteq (X \times Y)^4$, which consists of four subcodes as follows.

Let $\mathcal{D} \subseteq (X \times Y)^4$ be the $(4,1)_{vq}$-DCC that is obtained by applying Proposition 1 with $L(q)$. Note that $|\mathcal{D}| = \frac{1}{4}q(q-1)(q-2)v^3$ and $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{D}$ partition

$$T_1 = \left\{ x \circ y : x \in X^3, y \in \overline{\binom{Y}{3}} \right\}.$$

Let $\mathcal{A} = \cup_{d \in Y} \mathcal{C}_d$ be the second $(4,1)_{vq}$-DCC, where $\mathcal{C}_d$ is a perfect $(4,1)_v$-DCC of size s over $X \times \{d\}$ for any $d \in Y$. Here, $\mathcal{C}_d$ exists by assumption. Then $|\mathcal{A}| = qs$ and $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{A}$ partition

$$T_2 = \{x \circ (d,d,d) : x \in X^3, d \in Y\}.$$

Let $\mathcal{E} = \{(i_a, j_b, i_a, j_b), (j_b, j_b, i_a, i_a) : i_a, j_b \in X \times Y, a \prec b\}$ be the third $(4,1)_{vq}$-DCC, where $\prec$ is an order defined over Y. Then $|\mathcal{E}| = q(q-1)v^2$ and the sets $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{E}$ partition

$$T_3 = \left\{ (i_a, j_b, j_b), (j_b, i_a, j_b), (j_b, j_b, i_a) : (i,j) \in X^2, (a,b) \in \overline{\binom{Y}{2}} \right\}.$$

Finally, since v is even, for any $d \in Y$, there is a one-factorization $\mathcal{F}^d = \{F_1^d, \ldots, F_{v-1}^d\}$ of the complete graph on vertex set $X \times \{d\}$. Let

$$\mathcal{B} = \{\{i,j,k,l\} : \{i,j\} \in F_n^a, \{k,l\} \in F_n^b, n \in [1, v-1], a, b \in Y, a \neq b\}.$$

Note that quadruples in $\mathcal{B}$ cover all triples of the form $\{i_a, j_b, k_b\} \subseteq X \times Y$ with $j \neq k$ and $a \neq b$ exactly once. For each $B \in \mathcal{B}$, let $\mathcal{L}_B$ be an $L(3,4,4)$ over B. Then $\mathcal{G} = \cup_{B \in \mathcal{B}} \mathcal{L}_B$ is our fourth $(4,1)_{vq}$-DCC. Note that $|\mathcal{G}| = \frac{3}{4}q(q-1)v^2(v-1)$ and $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{G}$ partition

$$T_4 = \{(i_a, j_b, k_b), (j_b, i_a, k_b), (j_b, k_b, i_a) : i_a, j_b, k_b \in X \times Y, a \neq b, j \neq k\}.$$

Let $\mathcal{C} = \mathcal{D} \cup \mathcal{A} \cup \mathcal{E} \cup \mathcal{G}$. It is easy to check that $T_i, 1 \leq i \leq 4$ are disjoint and partition $(X \times Y)^3$. Hence, $D_1(\mathsf{u}), \mathsf{u} \in \mathcal{C}$ partition $(X \times Y)^3$ and $\mathcal{C}$ is a perfect $(4,1)_{vq}$-DCC of size $\frac{1}{4}q(q-1)(q-2)v^3 + qs + q(q-1)v^2 + \frac{3}{4}q(q-1)v^2(v-1)$. ∎

Corollary 3. *There exits a perfect $(4,1)_q$-DCC of size $\frac{q^3+q^2+4q}{4}$ and $\frac{q^3+5q^2/4}{4}$ for all $q \equiv 0 \bmod 8$, and of size $\frac{q^3+3q^2/2}{4}$ for all $q \equiv 0 \bmod 4$.*

Proof: Kim et al. [15] gives a perfect $(4,1)_4$-DCC of size 24. Then applying Theorem 3 with $v = 4$, a perfect $(4,1)_q$-DCC of size $\frac{q^3+q^2+4q}{4}$ exists for any $q \equiv 0 \bmod 8$.

Kim et al. [15] also gives a perfect $(4,1)_v$-DCC of size $\frac{v^3+2v^2}{4}$ for all even v. Then applying Theorem 3 with $v = q/4$, a perfect $(4,1)_q$-DCC of size $\frac{q^3+5q^2/4}{4}$ exists for any $q \equiv 0 \bmod 8$. If we apply Theorem 3 with $v = q/2$, then a perfect $(4,1)_q$-DCC of size $\frac{q^3+3q^2/2}{4}$ exists for all $q \equiv 0 \bmod 4$. ∎

There are only two sizes of perfect $(4,1)_q$-DCCs known for all even q:

(i) $(q^3 + q^2 + 2q)/4$ by Levenshteĭn [21], and
(ii) $(q^3 + 2q^2)/4$ by Kim et al. [15].

Thus, the sizes of the codes obtained in Corollary 3 are new. Our construction is also different from those of Wang and Ji [13]. Similar to Corollary 3, it is also possible to obtain perfect $(4,1)_q$-DCCs of further new sizes by applying Theorem 3. We do not explore them in this paper due to lack of space.

VI. Conclusion

We initiate the investigation into the possible sizes of q-ary perfect codes of length n that are capable of correcting burst deletions of length t. The spectrum of sizes for such codes is completely determined when

(i) $n \geq 4$ and $t = n - 2$, and
(ii) $n \geq 6$, $t = n - 3$, and q even.

We also show that the smallest possible size for such codes with $t = 2$, $n = 5$, and even q is $q^3/4$. Finally, new classes of perfect $(4,1)_q$-DCCs are obtained.

Acknowledgments

Yang Li was supported in part by NSFC under Grant No.11201328.

References

[1] R. M. Roth and P. H. Siegel, "Lee-metric BCH codes and their application to constrained and partial-response channels," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1083–1096, 1994.

[2] L. J. Schulman and D. Zuckerman, "Asymptotically good codes correcting insertions, deletions, and transpositions," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2552–2557, 1999.

[3] M. Marron, K. M. Swenson, and B. M. E. Moret, "Genomic distances under deletions and insertions," *Theoret. Comput. Sci.*, vol. 325, no. 3, pp. 347–360, 2004.

[4] V. I. Levenshteĭn, "Binary codes capable of correcting deletions, insertions, and reversals," *Soviet Physics Dokl.*, vol. 10, pp. 707–710, 1965.

[5] F. F. Sellers, Jr., "Bit loss and gain correction code," *IRE Trans.*, vol. IT-8, pp. 35–38, 1962.

[6] R. R. Varshamov and G. M. Tenengolts, "Codes which correct single asymmetric errors," *Autom. Remote Control*, vol. 26, no. 2, pp. 286–290, 1965.

[7] L. Calabi and W. E. Hartnett, "A family of codes for the correction of substitution and synchronization errors," *IEEE Trans. Information Theory*, vol. IT-15, pp. 102–106, 1969.

[8] R. Fuji-Hara, Y. Miao, J. Wang, and J. Yin, "Directed $B(K, 1; v)$ with $K = \{4, 5\}$ and $\{4, 6\}$ related to deletion/insertion-correcting codes," *J. Combin. Des.*, vol. 9, no. 2, pp. 147–156, 2001.

[9] J. Yin, "A combinatorial construction for perfect deletion-correcting codes," *Des. Codes Cryptogr.*, vol. 23, no. 1, pp. 99–110, 2001.

[10] N. Shalaby, J. Wang, and J. Yin, "Existence of perfect 4-deletion-correcting codes with length six," *Des. Codes Cryptogr.*, vol. 27, no. 1-2, pp. 145–156, 2002.

[11] N. J. A. Sloane, "On single-deletion-correcting codes," in *Codes and Designs (Columbus, OH, 2000)*, ser. Ohio State Univ. Math. Res. Inst. Publ. Berlin: de Gruyter, 2002, vol. 10, pp. 273–291.

[12] A. Klein, "On perfect deletion-correcting codes," *J. Combin. Des.*, vol. 12, no. 1, pp. 72–77, 2004.

[13] J. Wang and L. Ji, "Existence of $T^*(3,4,v)$-codes," *J. Combin. Des.*, vol. 13, no. 1, pp. 42–52, 2005.

[14] Y. M. Chee, G. Ge, and A. C. H. Ling, "Spectrum of sizes for perfect deletion-correcting codes," *SIAM J. Discrete Math.*, vol. 24, no. 1, pp. 33–55, 2010.

[15] H. K. Kim, J. Y. Lee, and D. Y. Oh, "Optimal single deletion correcting code of length four over an alphabet of even size," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3217–3220, 2010.

[16] D. Cullina and N. Kiyavash, "An improvement of Levenshtein's upper bound on the cardinality of deletion correcting codes," *IEEE Trans. Inform. Theory*, vol. 60, no. 7, pp. 3862–3870, 2014.

[17] J. Schwartz and J. Wolf, "A systematic (12,8) code for correcting single errors and detecting adjacent errors," *IEEE Trans. Comput.*, vol. 39, no. 11, pp. 1403–1404, 1990.

[18] D. C. Bossen, "b-adjacent error correction," *IBM J. Res. Develop.*, vol. 14, no. 4, pp. 402–408, 1970.

[19] T. Fukunaga and S. Tazaki, "Assessment on characteristics of b-adjacent error-correcting code," *Electron. Comm. Japan Part I Comm.*, vol. 68, no. 2, pp. 50–58, 1985.

[20] L. Cheng, T. G. Swart, H. C. Ferreira, and K. Abdel-Ghaffar, "Codes for correcting three or more adjacent deletions or insertions," in *ISIT 2014 – Proceedings of the 2014 IEEE International Symposium on Information Theory*, 2014, pp. 1246–1250.

[21] V. I. Levenshteĭn, "Perfect codes in the metric of deletions and insertions," *Diskret. Mat.*, vol. 3, no. 1, pp. 3–20, 1991.

[22] C. J. Colbourn and J. H. Dinitz, Eds., *The CRC Handbook of Combinatorial Designs*, 2nd ed. Boca Raton: CRC Press, 2007.