Cryptogr. Commun. (2018) 10:555–565
DOI 10.1007/s12095-017-0235-8

Bounds and constructions for $\overline{3}$-strongly separable codes with length 3

Xuli Zhang[1] · Jing Jiang[2] · Minquan Cheng[2]

Received: 17 November 2016 / Accepted: 15 June 2017 / Published online: 4 July 2017

Abstract Separable code (SC, Cheng and Miao IEEE Trans. Inf. Theory **57**, 4843–4851, 2011), frameproof code (FPC, Boneh and Shaw IEEE Trans. Inf. Theory **44**, 1897–1905, 1998) and strongly separable code (SSC, Jiang et al. Des. Codes Cryptogr. 79:303–318, 2016) are used to construct anti-collusion codes. SSC is better than FPC and SC in the applications for multimedia fingerprinting since SSC has lower identifying complexity than that of SC (the same complexity as FPC) and weaker structure than that of FPC. In this paper, we first derive several upper bounds on the number of codewords of a $\overline{t}$-SSC. Then we focus on $\overline{3}$-SSCs with codeword length 3 and obtain the following two main results: (1) An equivalence between an SSC and an SC is derived; (2) An improved lower bound $\Omega(q^{5/3}+q^{4/3}-q)$ on the size of a q-ary SSC when $q=q_1^6$ for any prime power $q_1\equiv 1 \pmod 6$, which is better than the previously known bound $\lfloor\sqrt{q}\rfloor^3$, is obtained by means of a difference matrix and a known result on the subsets of $\mathbb{F}_q^n$ containing no three points on a line.

Keywords Multimedia fingerprinting · Separable code · Strongly separable code · Forbidden configuration · Difference matrix

Mathematics Subject Classification (2010) 94B25 · 68P30

✉ Minquan Cheng
chengqinshi@hotmail.com

Xuli Zhang
xulizhang@hotmail.com

Jing Jiang
jjiang2008@hotmail.com

[1] School of Mathematics and Statistics, Guangxi Normal University, Guilin, 541004, China

[2] Guangxi Key Lab of Multi-source Information Mining & Security, Guangxi Normal University, Guilin, China

1 Introduction

With the advancement of multimedia technologies, coupled with the development of an infrastructure of ubiquitous broadband communication networks, a large amount of multimedia content, such as image, video, audio and speech, is available in the digital marketplace. However, pirate copies are an increasingly serious problems in copyright protection of multimedia contents.

In order to against pirate copies, fingerprinting system, in which each multimedia content has a unique embedded fingerprint, has recently become quite popular in the field of copyright protection for multimedia content. Clearly each customer obtains an embedded version which consists of a unique fingerprint and the same content. Consequently attacks mounted by individuals are no longer a main security issue in digital rights management. However, multiuser could carry out attacks against the embedded fingerprints by comparing their different embedded versions collectively. There are a variety of embedding techniques such as [2, 12, 19]. Spread spectrum embedding technique which can be used to restrict multiuser's attacking strategy as largely as possible [22], is widely used in fingerprinting systems [6, 12, 18, 22]. Such a system is called multimedia fingerprinting. Averaging attack is one of the most feasible approaches to perform a collusion attack in multimedia fingerprinting [17].

Cheng et al. [11] proposed a concept of a logical anti-collusion code (LACC), which can be used to against the averaging attack. They also showed that separable codes and frampeproof codes can be used to construct LACCs. There are several researches on separable codes and frameproof codes, for instances, [5, 7, 9–11, 14] and so on. The LACCs constructed by $\bar{t}$-separable codes can identify all colluders with computation complexity exponential in the number of authorized users, and those constructed by frameproof codes can identify all the colluders with computational complexity linear in the number of authorized users (by Theorem 5.5 and tracing algorithm `LACCIdenAlg` in [11]). This is in contrast to the fact that frameproof codes have no traceability properties under the embedding technique in [2]. In fact, the number of codewords in a frampeproof code is too small to be of practical use in most cases. Jiang et al. introduced a new concept of a strongly separable code which is weaker than a frameproof code but can also be used to identify all colluders with the same complexity (by Algorithm `SSCTraceAlg(R)` in [15]) as that of a frameproof code. Usually, strongly $\bar{t}$-separable codes have much more codewords than t-frameproof codes could have. So compared with frameproof codes, strongly separable codes have an advantage in copyright protection. In this paper, we will pay our attention to strongly separable codes.

When $t = 2$, some strongly $\bar{t}$-separable codes were studied in [15]. Especially the cases of codeword length 2 and 3 were discussed in detail. When $t \geq 3$, the structure of a strongly $\bar{t}$-separable code becomes more complex so that little is known about strongly $\bar{t}$-separable codes. In this paper, we will focus on strongly $\bar{t}$-separable codes with $t \geq 3$. First several upper bounds and a lower bound on the size of a strongly separable code are derived. Then we further improve the above lower bound when t and codeword length equal 3.

The remainder of the paper is organized as follows. Section 2 introduces preliminaries about separable codes, strongly separable codes and frameproof codes. In Section 3, several upper bounds on the size of strongly separable codes are derived by discussing the relationships among separable codes, strongly separable codes and frameproof codes. In Section 4, an improved lower bound $\Omega(q^{5/3} + q^{4/3} - q)$ on the size of a q-ary SSC when $q = q_1^6$ for any prime power $q_1 \equiv 1 \pmod 6$, which is better than the previously known bound $\lfloor \sqrt{q} \rfloor^3$, is obtained. Finally, conclusions are drawn in Section 5.

2 Preliminaries

Let n, M and q be positive integers, and Q an alphabet with $|Q| = q$. A set $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\} \subseteq Q^n$ is called an (n, M, q) code and each $\mathbf{c}_i$ is called a codeword. Without loss of generality, we may assume $Q = \{0, 1, \ldots, q-1\}$. When $Q = \{0, 1\}$, we also use the word “binary”.

For any code $\mathcal{C} \subseteq Q^n$, we define the set of ith coordinates of $\mathcal{C}$ as

$$\mathcal{C}(i) = \{\mathbf{c}(i) \in Q \mid \mathbf{c} = (\mathbf{c}(1), \mathbf{c}(2), \ldots, \mathbf{c}(n))^T \in \mathcal{C}\}$$

for any $1 \leq i \leq n$. For any subset of codewords $\mathcal{C}_0 \subseteq \mathcal{C}$, we define the descendant code of $\mathcal{C}_0$ by

$$\mathsf{desc}(\mathcal{C}_0) = \{(\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(n))^T \in Q^n \mid \mathbf{x}(i) \in \mathcal{C}_0(i), 1 \leq i \leq n\},$$

that is,

$$\mathsf{desc}(\mathcal{C}_0) = \mathcal{C}_0(1) \times \mathcal{C}_0(2) \times \ldots \times \mathcal{C}_0(n).$$

Clearly the set $\mathsf{desc}(\mathcal{C}_0)$ consists of the n-tuples that could be produced by a coalition holding the codewords in $\mathcal{C}_0$.

Definition 1 ([11, 15]) Let $\mathcal{C}$ be an (n, M, q) code and $t \geq 2$ be an integer.

- $\mathcal{C}$ is a $\overline{t}$-separable code, or $\overline{t}$-SC(n, M, q), if for any $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{C}$ such that $1 \leq |\mathcal{C}_1| \leq t$, $1 \leq |\mathcal{C}_2| \leq t$, and $\mathcal{C}_1 \neq \mathcal{C}_2$, we have $\mathsf{desc}(\mathcal{C}_1) \neq \mathsf{desc}(\mathcal{C}_2)$, that is there is at least one coordinate i, $1 \leq i \leq n$, such that $\mathcal{C}_1(i) \neq \mathcal{C}_2(i)$.
- $\mathcal{C}$ is a strongly $\overline{t}$-separable code, or $\overline{t}$-SSC(n, M, q), if for any $\mathcal{C}_0 \subseteq \mathcal{C}$ such that $1 \leq |\mathcal{C}_0| \leq t$, we have $\bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$, where $S(\mathcal{C}_0) = \{\mathcal{C}' \subseteq \mathcal{C} | \mathsf{desc}(\mathcal{C}') = \mathsf{desc}(\mathcal{C}_0)\}$.
- $\mathcal{C}$ is a t-frameproof code, or t-FPC(n, M, q), if for any $\mathcal{C}' \subseteq \mathcal{C}$ such that $|\mathcal{C}'| \leq t$, it holds that $\mathsf{desc}(\mathcal{C}') \bigcap \mathcal{C} = \mathcal{C}'$, that is, for any $\mathbf{c} = (\mathbf{c}(1), \ldots, \mathbf{c}(n))^T \in \mathcal{C} \setminus \mathcal{C}'$, there is at least one coordinate i, $1 \leq i \leq n$, such that $\mathbf{c}(i) \notin \mathcal{C}'(i)$.

Since the parameter M of a $\overline{t}$-SSC(n, M, q) corresponds to the number of fingerprints assigned to authorized users who purchased the right to access the copyrighted multimedia data, we should try to construct strongly separable codes with M as large as possible, given length n. Let $M(\overline{t}, n, q) = \max\{M \mid$ there exists a $\overline{t}$-SSC $(n, M, q)\}$. A $\overline{t}$-SSC(n, M, q) is said to be optimal if $M = M(\overline{t}, n, q)$. Similarly, a $\overline{t}$-SC(n, M, q) (or a t-FPC(n, M, q)) is optimal if M is the largest possible value given n, q and t.

3 Upper bounds

In this section, we first investigate the relationships among SC, SSC and FPC, and then derive the upper bounds on $M(\overline{t}, n, q)$ according to these relationships.

3.1 SC, SSC and FPC

The relationship between SC and FPC was described in [11].

Lemma 1 ([11]) *Any t-FPC(n, M, q) is a $\overline{t}$-SC(n, M, q), $t \geq 1$. Conversely any $\overline{t}$-SC(n, M, q) is a $(t-1)$-FPC(n, M, q), $t \geq 2$.*

Jiang et al. [15] established the following relationships among SC, SSC and FPC.

Lemma 2 ([15]) *Any t-FPC(n, M, q) is a $\overline{t}$-SSC(n, M, q).*

The following example shows that the converse of Lemma 2 does not always hold.

Example 1 ([15]) The following $(3, 4, 2)$ code $\mathcal{C}$ is a $\overline{2}$-SSC$(3, 4, 2)$, but is not a 2-FPC$(3, 4, 2)$.

$$\begin{array}{cc} & \begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \mathbf{c}_4 \end{array} \\ \mathcal{C} = & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Lemma 3 ([15]) *Any $\overline{t}$-SSC(n, M, q) is a $\overline{t}$-SC(n, M, q).*

Although, the converse of Lemma 3 does not always hold, when $t = n = 2$, Jiang et al. proved that the converse of Lemma 3 is also true.

Example 2 ([15]) Let $\mathbf{c}_i$, $1 \le i \le 5$, be the ith codeword of the following code $\mathcal{C}$, then $\mathcal{C}$ is a $\overline{2}$-SC$(3, 5, 2)$, but not a $\overline{2}$-SSC$(3, 5, 2)$, because $\mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_5) = \mathsf{desc}(\{\mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\})$.

$$\mathcal{C} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Lemma 4 ([15]) *A $(2, M, q)$ code $\mathcal{C}$ is a $\overline{2}$-SSC$(2, M, q)$ if and only if $\mathcal{C}$ is a $\overline{2}$-SC$(2, M, q)$.*

Example 3 The following code $\mathcal{C}$ is an optimal $\overline{3}$-SC$(3, 3, 2)$, and we can check that it is also a $\overline{3}$-SSC$(3, 3, 2)$.

$$\mathcal{C} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Furthermore, it is very interesting that the converse of Lemma 3 also holds for $t = n = 3$ and $q \ge 3$. We first state the two useful results. From Lemmas 1 and 3, the following statement holds.

Corollary 1 *Any $\overline{t}$-SSC(n, M, q) is a $(t-1)$-FPC(n, M, q) where $t \ge 2$.*

Lemma 5 *Suppose $\mathcal{C}$ is a $\overline{3}$-SC$(3, M, q)$. Then for any $\mathcal{C}_0 \subseteq \mathcal{C}$ with $|\mathcal{C}_0| \le 3$, and any $\mathbf{c} \in \mathcal{C}_0$, the Hamming distance $d(\mathbf{c}, \mathbf{c}') \ge 2$ holds for any $\mathbf{c}' \in \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} \setminus \mathcal{C}_0$.*

Proof By Lemma 1, $\mathcal{C}$ is a 2-FPC. By the definition of an FPC, we have $\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} = \mathcal{C}_0$ when $|\mathcal{C}_0| = 1, 2$. This implies $\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} \setminus \mathcal{C}_0 = \emptyset$. Clearly the statement holds. So we only need to consider the case $|\mathcal{C}_0| = 3$. For any $\mathcal{C}_0 = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$, where $\mathbf{c}_i = (a_i, b_i, e_i)^T$, $1 \le i \le 3$, suppose that there exits one codeword $\mathbf{c}' = (a', b', e')^T \in \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} \setminus \mathcal{C}_0$, such that $d(\mathbf{c}_1, \mathbf{c}') = 1$. Without loss of generality, assume $a_1 = a'$, $b_1 = b'$, $e_1 \ne e'$. This implies that e' equals e_2 or e_3 since $\mathbf{c}' \in \mathsf{desc}(\mathcal{C}_0)$. If $e' = e_2$ (or $e' = e_3$), we have

$\mathbf{c}' \in \mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2\})$ (or $\mathbf{c}' \in \mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_3\})$), a contradiction to the definition of a 2-FPC. So the statement also holds when $|\mathcal{C}_0| = 3$. □

Theorem 1 *For any $q \geq 3$, a $(3, M, q)$ code $\mathcal{C}$ is a $\overline{3}$-SSC$(3, M, q)$ if and only if $\mathcal{C}$ is a $\overline{3}$-SC$(3, M, q)$.*

Proof The necessity of the condition directly follows from Lemma 3. We now show that any $\overline{3}$-SC$(3, M, q)$ $\mathcal{C}$ over Q is also a $\overline{3}$-SSC$(3, M, q)$. That is, for any $\mathcal{C}_0 \subseteq \mathcal{C}$, $|\mathcal{C}_0| \leq 3$, we should show $\bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$ from the definition of an SSC.

By Lemma 1, $\mathcal{C}$ is a 2-FPC. From Lemma 2, we have $\mathcal{C}$ is a $\overline{2}$-SSC. So when $|\mathcal{C}_0| = 1$, 2, $\bigcap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$ holds. Now we consider the case $|\mathcal{C}_0| = 3$. For any $\mathcal{C}_0 = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$, $\mathbf{c}_i = (a_i, b_i, e_i)$, we have $\mathsf{desc}(\mathcal{C}_0)$:

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_1 & a_2 & a_1 & a_2 & a_3 & a_3 & a_1 & a_1 & a_1 & a_1 & a_1 & a_1 & a_2 & a_2 & a_2 & a_2 & a_2 & a_2 & a_3 & a_3 & a_3 & a_3 & a_3 & a_3 \\ b_1 & b_2 & b_3 & b_2 & b_1 & b_3 & b_3 & b_1 & b_2 & b_1 & b_1 & b_2 & b_2 & b_3 & b_3 & b_2 & b_3 & b_3 & b_2 & b_1 & b_1 & b_3 & b_3 & b_1 & b_1 & b_2 & b_2 \\ e_1 & e_2 & e_3 & e_3 & e_3 & e_2 & e_1 & e_2 & e_1 & e_2 & e_3 & e_1 & e_2 & e_1 & e_3 & e_3 & e_2 & e_3 & e_1 & e_2 & e_1 & e_1 & e_2 & e_3 & e_1 & e_3 & e_2 \end{pmatrix} \quad (1)$$

Let $\mathbf{c}_i$, $1 \leq i \leq 27$, be the ith codeword of $\mathsf{desc}(\mathcal{C}_0)$ in (1).

According to Lemma 5, $\mathbf{c}_i \notin \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$, $10 \leq i \leq 27$. Hence we have

$$\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} \subseteq \begin{pmatrix} a_1 & a_2 & a_3 & a_1 & a_2 & a_1 & a_2 & a_3 & a_3 \\ b_1 & b_2 & b_3 & b_2 & b_1 & b_3 & b_3 & b_1 & b_2 \\ e_1 & e_2 & e_3 & e_3 & e_3 & e_2 & e_1 & e_2 & e_1 \end{pmatrix} \quad (2)$$

Now we consider formula (2) by discussing cardinalities of sets $\mathcal{C}_0(i)$, $1 \leq i \leq 3$.

- If there exists an integer $1 \leq i \leq 3$ such that $|\mathcal{C}_0(i)| \leq 2$, without loss of generality, assume $|\mathcal{C}_0(1)| \leq 2$ and $a_1 = a_2$. According to Lemma 5, we have $\mathbf{c}_i \notin \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$, $4 \leq i \leq 7$. So we only need to consider $\mathbf{c}_8$ and $\mathbf{c}_9$.
  - If $a_1 = a_3$, then $\mathbf{c}_8, \mathbf{c}_9 \notin \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$ from Lemma 5. So $\cap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$.
  - If $a_1 \neq a_3$, $|\{b_1, b_2, b_3\}| < 3$ or $|\{e_1, e_2, e_3\}| < 3$ holds, then $\mathbf{c}_8, \mathbf{c}_9 \notin \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$ from Lemma 5. So $\cap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$.
  - If $a_1 \neq a_3$ and $|\{b_1, b_2, b_3\}| = |\{e_1, e_2, e_3\}| = 3$, then $\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$ contains at most one of $\mathbf{c}_8$ and $\mathbf{c}_9$. Otherwise, we have $\mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_8\}) = \mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_9\})$, a contradiction to the definition of a $\overline{3}$-SC. Without loss of generality, suppose that $\mathbf{c}_8 \in \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$. We have $\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_8\}$. Clearly $\mathrm{S}(\mathcal{C}_0) = \{\mathcal{C}_0, \mathcal{C}_1\}$ where $\mathcal{C}_1 = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_8\}$. It is easy to check that $\cap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$.
- If $|\mathcal{C}_0(i)| = 3$ for any $1 \leq i \leq 3$, it is easy to check that $d(\mathbf{c}_{j_1}, \mathbf{c}_{j_2}) = 2$ for all $j_1 = 1, 2, 3$ and $j_2 = 4, 5, \ldots, 9$. If $\mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C} = \mathcal{C}_0$, clearly it is a $\overline{3}$-SSC. Now we consider the case that there is at least one codeword $\mathbf{c}_i \in \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$, $4 \leq i \leq 9$, without loss of generality, we assume $\mathbf{c}_4 \in \mathsf{desc}(\mathcal{C}_0) \bigcap \mathcal{C}$. By the distance, $\mathbf{c}_i$, $5 \leq i \leq 9$, can be divided into two subsets $\mathcal{C}_1 = \{\mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_9\}$ and $\mathcal{C}_2 = \{\mathbf{c}_7, \mathbf{c}_8\}$ such that $d(\mathbf{c}_4, \mathbf{c}) = 2$ if $\mathbf{c} \in \mathcal{C}_1$ and $d(\mathbf{c}_4, \mathbf{c}) = 3$ if $\mathbf{c} \in \mathcal{C}_2$.
  - If $\mathsf{desc}(\mathcal{C}_0) \cap \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}$, then $\mathrm{S}(\mathcal{C}_0) = \{\mathcal{C}_0, \mathcal{C}'\}$, where $\mathcal{C}' = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}$. Clearly $\cap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$.
  - If $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}\} \subseteq \mathsf{desc}(\mathcal{C}_0) \cap \mathcal{C}$, $\mathbf{c} \in \mathcal{C}_1$, we claim that this case does not happen. We take $\mathbf{c} = \mathbf{c}_5$ as an example. Then we have $\mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_4\}) = \mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_5\})$, a contradiction to the definition of a $\overline{3}$-SC.

- If $\mathsf{desc}(\mathcal{C}_0) \cap \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}\}$, $\mathbf{c} \in \mathcal{C}_2$, we claim that this case satisfies the conditions of $\overline{3}$-SSC. We take $\mathbf{c} = \mathbf{c}_7$ as an example. Let $\mathcal{C}' = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}, \mathcal{C}'' = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_7\}$ and $\mathcal{C}''' = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_7\}$. It is easy to check that $S(\mathcal{C}_0) = \{\mathcal{C}_0, \mathcal{C}', \mathcal{C}'', \mathcal{C}'''\}$ and $\cap_{\mathcal{C}' \in S(\mathcal{C}_0)} \mathcal{C}' = \mathcal{C}_0$.
- If $\mathsf{desc}(\mathcal{C}_0) \cap \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_7, \mathbf{c}_8\}$, then $\mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}) = \mathsf{desc}(\{\mathbf{c}_4, \mathbf{c}_7, \mathbf{c}_8\})$, a contradiction to the definition of a $\overline{3}$-SC.

From the above discussions, we know that $\mathcal{C}$ is a $\overline{3}$-SSC. Then the proof is complete. □

3.2 Upper bounds on $M(\bar{t}, n, q)$

As an important class of anti-collusion codes in multimedia copyright protection, separable codes and frameproof codes were widely studied, e.g., [5, 7, 9–11, 14].

Lemma 6 ([8]) *Given a $\bar{t}$-SC(n, M, q) with $t \geq 3$ and $n \geq 2$, let $r \in \{0, 1, \ldots, t-2\}$ be the remainder of n on division by $t-1$. If $M > q$, then*

$$M \leq \max\left\{q^{\lceil n/(t-1)\rceil}, r(q^{\lceil n/(t-1)\rceil} - 1) + (t-1-r)(q^{\lfloor n/(t-1)\rfloor} - 1)\right\}.$$

Lemma 7 ([9]) *For any positive integers t and $q \geq 2$,*

- *when $2 \leq n < t$, there always exists an optimal $\bar{t}$-SC$(n, n(q-1), q)$ and an optimal t-FPC$(n, n(q-1), q)$;*
- *when $n = t$, for any $\bar{t}$-SC(n, M, q) we have $M \leq q^2$ if $n \leq q$, otherwise $M \leq nq$.*

When $t = 3, n = 3$, Cheng et al. improved the above result.

Lemma 8 ([9]) *The maximum value of M in a $\overline{3}$-SC$(3, M, q)$ must be between $\lfloor\sqrt{q}\rfloor^3$ and $\lfloor\frac{3q^2}{4}\rfloor$, where $q \geq 4$.*

From Lemmas 3 and 6, the following upper bounds on $M(\bar{t}, n, q)$ can be obtained.

Corollary 2 *Let n, q and t be positive integers such that $t \geq 3$ and $n \geq 2$, and $r \in \{0, 1, \ldots, t-2\}$ the remainder of n on division by $t-1$. If $M(\bar{t}, n, q) > q$, then*

$$M(\bar{t}, n, q) \leq \max\left\{q^{\lceil n/(t-1)\rceil}, r(q^{\lceil n/(t-1)\rceil} - 1) + (t-1-r)(q^{\lfloor n/(t-1)\rfloor} - 1)\right\}.$$

From Lemmas 2, 3 and 7, the following statement holds.

Corollary 3 *For any positive integers t, n and $q \geq 2$,*

- *when $2 \leq n < t$, $M(\bar{t}, n, q) = n(q-1)$;*
- *when $n = t$, $M(\bar{t}, n, q) \leq q^2$ if $n \leq q$, and otherwise $M(\bar{t}, n, q) \leq nq$.*

From Theorem 1 and Lemma 8, we have the following result.

Corollary 4 $\lfloor\sqrt{q}\rfloor^3 \leq M(\bar{3}, 3, q) \leq \lfloor\frac{3q^2}{4}\rfloor$ *holds for $q \geq 4$.*

To our best knowledge, the lower bound in [9] is the best known result on $\overline{3}$-SC$(3, M, q)$. In the following section, we will improve the lower bound in Corollary 4 to $\Omega(q^{5/3}+q^{4/3}-q)$ for some prime powers q.

4 Construction

From Theorem 1, it is sufficient to consider $\overline{3}$-SC$(3, M, q)$ for studying $\overline{3}$-SSC$(3, M, q)$. First the following notations are necessary.

For any $(3, M, q)$ code $\mathcal{C}$ defined on $Q = \{0, 1, \cdots, q-1\}$, we define the column vector sets $\mathcal{A}_i^{(1)}$ for $i \in Q$ as follows:

$$\mathcal{A}_i^{(1)} = \{(x_2, x_3)^T \mid (x_1, x_2, x_3)^T \in \mathcal{C},\ x_1 = i\}.$$

Obviously, $\mathcal{A}_i^{(1)} \subseteq Q^2$ for any $i \in Q$ and $|\mathcal{A}_0^{(1)}| + \cdots + |\mathcal{A}_{q-1}^{(1)}| = M$ hold. Similar to the above notation, vector sets $\mathcal{A}_i^{(j)}$ for $j = 2, 3$ can be also defined.

Lemma 9 ([9]) *A* $(3, M, q)$ *code is a 2-FPC*$(3, M, q)$ *if and only if* $|\mathcal{A}_i^{(j)} \bigcap \mathcal{A}_{i'}^{(j)}| \le 1$ *holds for any* $j \in \{1, 2, 3\}$ *and distinct* $i, i' \in Q$*, where if* $|\mathcal{A}_i^{(j)} \bigcap \mathcal{A}_{i'}^{(j)}| = 1$*, then* $|\mathcal{A}_i^{(j)}| = |\mathcal{A}_{i'}^{(j)}| = 1$.

Cheng et al. showed that for any $\overline{3}$-SC$(3, M, q)$, $\mathcal{C}$, there is no subcode $\triangle_i \subseteq \mathcal{C}$ described in (3), $a \neq b$, $c \neq d$, $e \notin \{f, g\}$, and there is no subcode $\nabla \subseteq \mathcal{C}$ described in (4), $|\{a_i, b_i, c_i\}| = 3$, $i = 1, 2, 3$.

$$\triangle_1 = \begin{pmatrix} a & a & b & b \\ e & f & g & e \\ c & d & c & d \end{pmatrix} \quad \triangle_2 = \begin{pmatrix} a & a & b & b \\ c & d & c & d \\ e & f & g & e \end{pmatrix} \quad \triangle_3 = \begin{pmatrix} e & f & g & e \\ a & a & b & b \\ c & d & c & d \end{pmatrix} \tag{3}$$

$$\nabla = \begin{pmatrix} a_1 & b_1 & c_1 & a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 & b_2 & c_2 & a_2 \\ a_3 & b_3 & c_3 & c_3 & a_3 & b_3 \end{pmatrix} \tag{4}$$

We call such $\triangle_i$ and $\triangle$ *forbidden configurations* of $\mathcal{C}$.

Theorem 2 ([9]) *A* $(3, M, q)$ *code* $\mathcal{C}$ *is a* $\overline{3}$-*SC*$(3, M, q)$ *if and only if it satisfies the following conditions:*

(i) *$\mathcal{C}$ is a 2-FPC$(3, M, q)$;*
(ii) *Configurations in* (3) *and* (4) *are all the forbidden configurations of* $\mathcal{C}$.

In the following, for any prime power q, we will take advantage of difference matrices to construct $\overline{3}$-SC$(3, M, q)$s.

Definition 2 ([3]) For any prime power q, a *difference matrix* $(q, 3, 1)$DM is a $3 \times q$ matrix $D = (d_{j,i})$ with $d_{j,i} \in \mathbb{F}_q$ such that for any $1 \le j_1 \neq j_2 \le 3$, the differences $d_{j_1,i} - d_{j_2,i}$ over $\mathbb{F}_q$, $i \in \mathbb{F}_q$, comprise all the elements of $\mathbb{F}_q$.

Given a $3 \times s$ matrix N with entries from $\mathbb{F}_q$ and s distinct columns $\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_s$, we can define a $(3, qs, q)$ code $\mathcal{C}$ on $\mathbb{F}_q$ as

$$\mathcal{C} = \{N + g \mid g \in \mathbb{F}_q\} = \{\mathbf{n}_i + g \mid g \in \mathbb{F}_q,\ 1 \le i \le s\}. \tag{5}$$

We say N is a base of $\mathcal{C}$, or $\mathcal{C}$ is constructed by N.

For any given $(q, 3, 1)$DM, D, we can obtain a $(3, q^2, q)$ code $\mathcal{C} = \{D + g \mid g \in \mathbb{F}_q\}$. By the definition of a DM, we know that $|\mathcal{A}_{i_1}^{(j)} \cap \mathcal{A}_{i_2}^{(j)}| = 0$ holds for any $1 \le j \le 3$ and for any distinct $i_1, i_2 \in \mathbb{F}_q$, which implies that $\mathcal{C}$ is a 2-FPC$(3, q^2, q)$ by Lemma 9. Unfortunately, this code is not always a $\overline{3}$-SC.

Example 4 The following code $\mathcal{C}$ is constructed by $(3, 3, 1)$DM in (5). Let $\mathbf{c}_i$ denote the ith codewode, $1 \le i \le 9$. From the above discussion, $\mathcal{C}$ is a 2-FPC$(3, 9, 3)$, but is not a $\overline{3}$-SC since $\mathsf{desc}(\{\mathbf{c}_1, \mathbf{c}_4, \mathbf{c}_7\}) = \mathsf{desc}(\{\mathbf{c}_2, \mathbf{c}_5, \mathbf{c}_8\})$.

$$\mathcal{C} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 2 & 1 & 1 & 0 & 2 & 2 & 1 & 0 \end{pmatrix}$$

In fact, we can obtain the base of a $\overline{3}-$SC$(3, M, q)$ by deleting some codewords of $\mathcal{C}$ constructed by $(q, 3, 1)$DM in (5). For any prime power q, if $q \ge 3$, the following array D is a $(q, 3, 1)$DM

$$D = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & \varepsilon^{q-2} \\ 0 & \alpha & \ldots & \alpha\varepsilon^{q-2} \end{pmatrix}, \tag{6}$$

where ε is a primitive element of $\mathbb{F}_q$ and α is an element of $\mathbb{F}_q \setminus \{0, 1\}$ [13]. For any subset $S \subseteq \mathbb{F}_q$, let sub-matrix $N = D|_S$ obtained by deleting the columns $i \in \mathbb{F}_q \setminus S$. Clearly the code $\mathcal{C}$ constructed by N in (5) is a 2-FPC$(3, q|S|, q)$. From Theorem 2, in order that $\mathcal{C}$ may be a $\overline{3}$-SC$(3, M, q)$, we only need to consider the forbidden configurations in (3) and (4). Suppose $C \subseteq \mathcal{C}$,

(I) When $C \in \{\triangle_1, \triangle_2, \triangle_3\}$, we may assume

$$C = \begin{pmatrix} k_1 & k_2 & k_3 & k_4 \\ x + k_1 & y + k_2 & z + k_3 & w + k_4 \\ \alpha x + k_1 & \alpha y + k_2 & \alpha z + k_3 & \alpha w + k_4 \end{pmatrix},$$

where $x, y, z, w, k_1, k_2, k_3, k_4 \in \mathbb{F}_q$.

- When $C = \triangle_1$, we have

$$k_1 = k_2,\ x + k_1 = w + k_4,\ \alpha y + k_2 = \alpha w + k_4.$$

This means

$$x + (\alpha - 1)w = y\alpha \tag{7}$$

with $|\{x, y, w\}| = 3$. In fact, if $x = y$, we have that the first codeword equals the second codeword of $\triangle_1$, a contradiction to the assumption. This implies $x \neq y$. Similarly we can check that $x \neq w$ and $y \neq w$.

- When $C = \triangle_2$, we have

$$k_1 = k_2,\ y + k_2 = w + k_4,\ \alpha x + k_1 = \alpha w + k_4.$$

This means

$$y + (\alpha - 1)w = x\alpha. \tag{8}$$

It is easy to check that $|\{x, y, w\}| = 3$ holds in (8).

- When $C = \triangle_3$, we have

$$k_1 = k_4,\ x + k_1 = y + k_2,\ \alpha y + k_2 = \alpha w + k_4.$$

This means

$$x + (\alpha - 1)y = w\alpha. \tag{9}$$

It is easy to check that $|\{x, y, w\}| = 3$ holds in (9).

(II) When $C = \nabla$, we may assume

$$C = \begin{pmatrix} k_1 & k_2 & k_3 & k_1 & k_2 & k_3 \\ x + k_1 & y + k_2 & z + k_3 & u + k_1 & v + k_2 & w + k_3 \\ \alpha x + k_1 & \alpha y + k_2 & \alpha z + k_3 & \alpha u + k_1 & \alpha v + k_2 & \alpha w + k_3 \end{pmatrix}.$$

$$\begin{cases} x + k_1 = w + k_3 \\ y + k_2 = u + k_1 \\ z + k_3 = v + k_2 \\ \alpha x + k_1 = \alpha v + k_2 \\ \alpha y + k_2 = \alpha w + k_3 \\ \alpha z + k_3 = \alpha u + k_1 \end{cases} \Longrightarrow \begin{cases} k_3 - k_1 = x - w \\ k_2 - k_1 = u - y \\ k_3 - k_2 = v - z \\ k_2 - k_1 = \alpha x - \alpha v \\ k_3 - k_2 = \alpha y - \alpha w \\ k_3 - k_1 = \alpha u - \alpha z \end{cases}. \tag{10}$$

This means

$$\begin{cases} \alpha x + \alpha(\alpha - 1)z = (\alpha - 1)y + (\alpha^2 - \alpha + 1)u \\ \alpha w + \alpha(\alpha - 1)u = (\alpha - 1)v + (\alpha^2 - \alpha + 1)z \end{cases} \tag{11}$$

Then we know $\{x, y, z\} \cap \{u, v, w\} = \emptyset$ always holds.

- $x \notin \{u, v, w\}$ always holds. If $x = u$, we have the first codeword equals the forth codeword of ∇, a contradiction to the assumption. Similarly, we can prove that $x \neq w, v$ always holds.
- $y \notin \{u, v, w\}$ always holds. If $y = u$, we have $k_1 = k_2$ from $y + k_2 = u + k_1$. This implies that the second codeword equals the forth codeword of ∇, a contradiction to the assumption. Similarly, we can prove that $y \neq w, v$.
- $z \notin \{u, v, w\}$ always holds. If $z = u$, we have $k_1 = k_3$ from $\alpha z + k_3 = \alpha u + k_1$. This implies that the third codeword equals the forth codeword of ∇, a contradiction to the assumption. Similarly, we have $z \neq w, v$.

For any prime power q_1 and positive integer n, let $\mathbb{F}_{q_1}^n$ be the n-dimensional vector space over $\mathbb{F}_{q_1}^n$. When $q = q_1^n$, it is well known that the element of $\mathbb{F}_q$ can be represented by the n-dimensional vector over $\mathbb{F}_{q_1}$. Suppose that S is a subset of $\mathbb{F}_{q_1}^n$, of which no three distinct elements are collinear. Then (7), (8) and (9) have no solution in S. This implies that $\mathcal{C}$ does not contain $\triangle_1$, $\triangle_2$ and $\triangle_3$. Together with (11), we have the following result.

Theorem 3 *For any subset $S \subseteq \mathbb{F}_q$, of which no three distinct elements are collinear, if there is no solution of (11) in S, then the code constructed by $N = D|_S$ in (5) is a $\overline{3}$-$SC(3, q|S|, q)$.*

Combining forbidden configurations with subsets containing no nontrivial solution to certain equations is an efficient method. This similar idea was also used in [1] and [20]. Now, we focus on the formula (11). Let $q_1 = 6t + 1$ be a prime power, and α be a primitive 6th root of unity in $\mathbb{F}_{q_1}$, where $t \geq 1$. Clearly α is a root of $f(x) = x^2 - x + 1$. Then (11) can be written as

$$\begin{cases} x + (\alpha - 1)z = \alpha y \\ w + (\alpha - 1)u = \alpha v \end{cases} \tag{12}$$

From (12), if $|\{x, y, z\}| < 3$ (or $|\{u, v, w\}| < 3$), then $|\{x, y, z\}| = 1$ (or $|\{u, v, w\}| = 1$) always holds. Furthermore, from (10) we claim if $x = y = z$ (or $u = v = w$), then $|\{u, v, w\}| = 3$ (or $|\{x, y, z\}| = 3$) always holds in (11). If $x = y = z$ and $u = v = w$, then $x + k_1 = w + k_3$ and $\alpha x + k_3 = \alpha w + k_1$ hold by (10). We have $(\alpha + 1)x = (\alpha + 1)w$. This implies $x = w$, $k_1 = k_3$ since $\alpha \neq -1$. That is, the first codeword equals the sixth codeword in C, a contradiction. So we have

$$\begin{aligned} &|\{x, y, z, u, v, w\}| = 6; \textit{ or} \\ &|\{x, y, z\}| = 3 \text{ and } |\{u, v, w\}| = 1; \textit{ or} \\ &|\{x, y, z\}| = 1 \text{ and } |\{u, v, w\}| = 3. \end{aligned} \tag{13}$$

According to (12) and (13), Theorem 3 can be written as follows.

Corollary 5 *Let $q = q_1^n$, where $q_1 = 6t + 1$ is a prime power, $t \geq 1$. For any subset $S \subseteq \mathbb{F}_q$, of which no three distinct elements are collinear, the code constructed by $N = D|_S$ in (5) is a $\overline{3}$-SC$(3, q|S|, q)$.*

Denoting by $r(q_1^n)$ the maximum size of a subset of $\mathbb{F}_{q_1}^n$ that contains no three points on a line. There are many studies on the value of $r(q_1^n)$ over $\mathbb{F}_{q_1}^n$. The interested reader is referred to [4, 16, 21].

Lemma 10 ([16]) *For any prime power $q_1 \geq 3$, we have $r(\mathbb{F}_{q_1}^6) = \Omega(q_1^4 + q_1^2 - 1)$.*

From Corollary 5 and Lemma 10, the following lower bound can be obtained.

Theorem 4 *For any prime power $q = q_1^6$, where $q_1 = 6t + 1$ is a prime power, there exists a $\overline{3}$-SSC$(3, M, q)$, where $M = \Omega(q^{5/3} + q^{4/3} - q)$.*

5 Conclusion

In this paper, we first derived several upper bounds on the number of codewords of $\bar{t}$-SSC. Then we focused on $\overline{3}$-SSCs with codeword length 3, and obtained the following two main results: (1) An equivalence between an SSC and an SC was derived. (2) An improved lower bound $\Omega(q^{5/3} + q^{4/3} - q)$ on the size of a q-ary SSC when $q = q_1^6$ for any prime power $q_1 \equiv 1 \pmod 6$, which is better than the previously known bound $\lfloor \sqrt{q} \rfloor^3$, was obtained by means of a difference matrix and a known result on the subsets of $\mathbb{F}_q^n$ containing no three points on a line.

It would be of interest if we could improve the upper bounds $\lfloor \frac{3q^2}{4} \rfloor$ or the lower bound $\Omega(q^{5/3} + q^{4/3} - q)$. It would be also interesting if we could get more properties and constructions of strongly separable codes.

Acknowledgements The authors express their sincere thanks to the two anonymous reviewers for their critical comments and suggestions which made this paper significantly improved, and to the editor for his/her excellent editorial job. This work is in part supported by NSFC (No.11301098, 11601096 and 61672176), 2016GXNSF(No. FA380009 and CA380021), 2014GXNSFDA118001, Guangxi Higher Institutions' Program of Introducing 100 High-Level Overseas Talents, Guangxi Collaborative Innovation Center of Multi-source Information Integration and Intelligent Processing, and Guangxi Bagui Scholar Teams for Innovation and Research.

References

1. Alon, N., Fischer, E., Szegedy, M.: Parent-identifying codes. J. Combin. Theory Ser. A **95**, 349–359 (2001)
2. Boneh, D., Shaw, J.: Collusion-secure fingerprinting for digital data. IEEE Tran. Inf. Theory **44**, 1897–1905 (1998)
3. Bose, R.C., Bush, K.A.: Orthogonal arrays of strength two and three. Ann. Math. Statist. **23**, 508–524 (1952)
4. Bierbrauer, J.: Large caps. J. Geom. **76**, 16–51 (2003)
5. Blackburn, S.R.: Frameproof codes. SIAM J. Discrete Math. **16**, 499–510 (2003)
6. Chen, B., Wornell, G.W.: Quantization index modulation: a class of provably good methods for digital watermarking and information embedding. IEEE Trans. Inf. Theory **47**, 1423–1443 (2001)
7. Cheng, M., Fu, H.-L., Jiang, J., Lo, Y.-H., Miao, Y.: New bounds on $\overline{2}$-separable codes of length 2. Des. Codes Cryptogr. **74**, 31–40 (2015)
8. Cheng, M., Ji, L., Miao, Y.: Separable codes. IEEE Trans. Inf. Theory **58**, 1791–1803 (2012)
9. Cheng, M., Jiang, J., Li, H., Miao, Y., Tang, X.: Bounds and construction for $\overline{3}$-separable codes with short length 3. Des. Codes Cryptogr. doi:10.1007/s10623-015-0160-9
10. Cheng, M., Jiang, J., Tang, X.: Asymptotically optimal $\overline{2}$-separable codes with length 4. Cryptogr. Commun. doi:10.1007/s12095-016-0182-9
11. Cheng, M., Miao, Y.: On anti-collusion codes and detection algorithms for multimedia fingerprinting. IEEE Trans. Inf. Theory **57**, 4843–4851 (2011)
12. Cox, I.J., Kilian, J., Leighton, T., Shamoon, T.: Secure spread spectrum watermarking for multimedia. IEEE Trans. Image Process. **6**, 1673–1687 (1997)
13. Drake, D.A.: Partial λ-geometries and generalized Hadamard matrices over groups. Canad. J. Math. **31**, 617–627 (1979)
14. Gao, F., Ge, G.: New bounds on separable codes for multimedia fingerprinting. IEEE Trans. Inf. Theory **60**, 5257–5262 (2014)
15. Jiang, J., Cheng, M., Miao, Y.: Strongly separable codes. Des. Codes Cryptogr. **79**, 303–318 (2016)
16. Lin, Y., Wolf, J.: On subset of $\mathbb{F}_q^n$ containing no k-term progressions. Eur. J. Comb. **31**, 1398–1403 (2010)
17. Liu, K.J.R., Trappe, W., Wang, Z.J., Wu, M., Zhao, H.: Multimedia Fingerprinting Forensics for Traitor Tracing. Hindawi, New York (2005)
18. Moulin, P., O'Sullivan J.A.: Information-theoretic analysis of information hiding. IEEE Trans. Inf. Theory **49**, 563–593 (2003)
19. Podilchuk, C.I., Zeng, W.: Image-adaptive watermarking using visual models. IEEE J. Select. Areas Commun. **16**, 525–539 (1998)
20. Shangguan, C., Ge, G.: Separating hash families: a Johnson-type bound and new constructions. SIAM J. Discrete. Math. **30**, 2243–2264 (2016)
21. Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. Acta Arith. **27**, 199–245 (1975)
22. Trappe, W., Wu, M., Wang, Z.J., Liu, K.J.R.: Anti-collusion fingerprinting for multimedia. IEEE Trans. Signal Process. **51**, 1069–1087 (2003)