Generalized Singleton Type Upper Bounds

Hao Chen, Longjiang Qu, Chengju Li, Shanxiang Lyu, Member, IEEE, Liqing Xu, and Mingshuo Zhou

Abstract—In this paper, we give many new Singleton type upper bounds on the sizes of codes with given minimum Hamming distances. These upper bounds are stronger than the Griesmer bound when the lengths of codes are large. Some upper bounds on the lengths of general Singleton type codes are presented. Our generalized Singleton type upper bounds have wide applications to symbol-pair codes, insertion-deletion codes and locally recoverable codes. The generalized Singleton type upper bounds on symbol-pair codes and insertion-deletion codes are much stronger than the direct Singleton bounds on symbol-pair codes and insertion-deletion codes when the lengths are large and the Hamming minimum distances are small. Upper bounds on the lengths of small dimension optimal locally recoverable codes and small dimension optimal $(r, \delta)$ locally recoverable codes with any given minimum distance are also presented.

Index Terms—Covering code, generalized singleton type upper bound, symbol-pair code, insertion-deletion code, singleton-optimal locally recoverable code.

I. INTRODUCTION

Let $\mathbb{F}_q$ be the finite field of size $q$, where $q$ is a prime power. When $q$ is prime, this finite field is called a prime field. For a vector $a \in \mathbb{F}_q^n$, the Hamming weight $wt(a)$ is the number of non-zero coordinate positions. The Hamming distance $d(a, b)$ between two vectors $a$ and $b$ is defined to be the Hamming weight of $a - b$. The minimum Hamming distance $d(C)$ of a code $C \subseteq \mathbb{F}_q^n$ is the minimum of Hamming distances $d(a, b)$ between any two different codewords $a$ and $b$ in $C$. It is well-known that the minimum Hamming distance of a linear code $C$ is the minimum Hamming weight of its non-zero codewords. Let $A_q(n, d)$ be the maximal size of a code $C \subseteq \mathbb{F}_q^n$ with the minimum Hamming distance $d$. One of the main problems in the error-correcting coding theory is to determine $A_q(n, d)$. There are many upper bounds on $A_q(n, d)$, see Chapter 2 of [41]. For example it was known that $A_q(n, d) \leq qA_q(n - 1, d)$, see [64].

A. Singleton Bound, Griesmer Bound and MDS Codes

The Singleton bound $A_q(n, d) \leq q^{n - d + 1}$ in [69] for a general code with the length $n$ and the given minimum Hamming distance $d$ is the basic upper bound in coding theory, see also [44]. The number $n + 1 - d - \log_q |C|$ is called the Singleton defect of a code $C \subseteq \mathbb{F}_q^n$. A code attaining this bound or with the zero Singleton defect is called a maximal distance separable (MDS) code. A code with the Singleton defect $1$ is called almost MDS. A linear almost MDS code $C$, which satisfies that the dual code $C^\perp$ is also almost MDS, is called a near MDS code. Near and almost MDS codes have received a lot of attentions, see e.g. [5], [52], [62], [66] and references therein.

To determine the maximal lengths of general MDS codes, almost MDS codes or near MDS codes is a classical problem in coding theory and finite geometries, see [1], [51], [62], [68]. The Bush upper bound

$$A_q(q^2 + 2, q) \leq q^{3 - 2}$$

for odd $q$, was given in [7] and improved in [64]. From this bound it follows that no MDS code of the minimum Hamming distance $q$ over $\mathbb{F}_q$, $q$ is odd, and of the length $n \geq q + 2$ exists. It was proved that the length of an MDS $(n, q^k, n - k + 1)_q$ code satisfies $n \leq q + k - 1$, see [1], [68]. For nontrivial linear MDS codes, the dimension $k$ has to satisfy $k \leq q - 1$, then $n \leq 2q - 2$, see [41], page 264. The main conjecture of linear MDS codes proposed in [65] claims that the length of a linear MDS code over $\mathbb{F}_q$ is at most $q + 1$, except some exceptional cases. In [2] the main conjecture was proved for linear MDS codes over prime fields. The main conjecture for linear MDS codes has been proved for small dimensions, small $q$’s and some geometric codes, see [1], [6], [14], and [15] and references therein. For the recent progress in the main conjecture of linear MDS codes, we refer to [3]. The main conjecture of near MDS codes was proposed in [52] and has been studied in many papers, see [62] and references therein. The maximal lengths of linear almost MDS codes were studied in [5] and [40]. Some upper bounds were given for small $q$’s and the special case such as $d > q$. Some classification
results about non-linear MDS codes over small fields were given in [51]. We refer to [34] and [49] for the counting of the number of MDS linear codes. New MDS codes which are not equivalent to Reed-Solomon codes and twisted Reed-Solomon codes were constructed from algebraic curves in [18]. In this paper, our upper bounds on the lengths of MDS codes and small Singleton defect codes are for both linear and non-linear codes.

The Griesmer bound for a linear \([n, k, d]_q\) code proposed in [39] asserts that

\[ n \geq \sum_{i=0}^{k-1} \left\lfloor \frac{d}{q^i} \right\rfloor, \]

see Section 2.7 of [41]. This bound is stronger than the Singleton bound for linear codes, since \(n \geq d + \sum_{i=0}^{k-1} \left\lfloor \frac{d}{q^i} \right\rfloor \geq d + k - 1\) and serves as the best upper bound for many presently known best linear codes in [38].

Let \(F_q\) be an arbitrary finite field, \(P_1, \ldots, P_n\) be \(n \leq q\) elements in \(F_q\). The Reed-Solomon codes \(R S(n, k)\) is defined by

\[ RS(n, k) = \{(f(P_1), \ldots, f(P_n)) : f \in F_q[x], \deg(f) \leq k - 1\}. \]

This is an \([n, k, n-k+1]_q\) linear MDS codes based on the fact that the polynomial \(f\) has at most \(k - 1\) roots, since \(\deg(f) \leq k - 1\).

B. Covering Codes

For a code \(C \subset F_q^n\), we define its covering radius by

\[ R_{c}(C) = \max_{c \in F_q^n} \min_{x \in C} \{wt(x - c)\}. \]

Hence the Hamming balls \(B(x, R_{c}(C))\) centered at all codewords \(x \in C\), with the radius \(R_{c}(C)\) cover the whole space \(F_q^n\). Then \(C\) is called a covering code with the radius \(R_{c}(C)\). We refer to the excellent book [20] on this classical topic of coding theory. Actually the covering radius \(R_{c}(C)\) of a linear \([n, k]_q\) code \(C \subset F_q^n\) can be determined as follows. If \(H\) is any \((n - k) \times n\) parity check matrix of \(C\), \(R_{c}(C)\) is the least integer such that every vector in \(F_q^{n-k}\) can be represented as \(F_q\) linear combinations of \(R_{c}(C)\) or fewer columns of \(H\), see [24], [43]. The redundancy upper bound

\[ R_{c}(C) \leq n - k \]

for a linear \([n, k]_q\) code, follows immediately, see [20]. Let \(n\) be a fixed positive integer and \(q\) be a fixed prime power, for a given positive integer \(R < n\), we denote \(K_q(n, R)\) the minimal size of a code \(C \subset F_q^n\) with the covering radius smaller than or equal to \(R\). Set \(\frac{\log K_q(n, R)}{n}\) as \(k_q(n, R)\). The following asymptotic bound is well-known,

\[ 1 - H_q(\rho) \leq k_q(n, \rho) \leq 1 - H_q(\rho) + O\left(\frac{\log n}{n}\right), \]

where \(H_q(\rho) = r\log_q(q - 1) - r\log_q(r) - (1 - r)\log_q(1 - r)\) is the \(q\)-ary entropy function, see [20, Chapter 12]. In particular when each vector in \(F_q^n\) is in exactly one Hamming ball centered in a codeword in \(C\) with the radius of \(R\), we call this code a perfect codes. The existence and determination of perfect codes in various metrics are a fascinating topic in coding theory, related to many other topics of mathematics. Hamming code and Golay code are basic examples of perfect codes, see [20, Chapter 11]. It is well-known that the covering radius of the Reed-Solomon \([n, k]_q\) code is \(n - k\) if \(n \leq q\), see [20]. There have been extensive research on covering radii of codes in the Hamming metric. For example, in the paper [31], covering radii of more than six thousand binary cyclic codes were calculated and determined. We refer to [32] for the recent work on the generalized covering radii of linear codes.

C. Generalized Singleton Bound for List-Decodable Codes

A code \(C\) in the Hamming metric space \(F_q^n\) is called (combinatorial) \((d_{list}, L)\) list-decodable if the Hamming ball of the radius \(d_{list}\) centered at any \(x \in F_q^n\) contains at most \(L\) codewords of \(C\). We refer to [46] and [70] and [41, Chapter 5] for the list-decoding algorithm and list-decodable codes in the Hamming metric space \(F_q^n\).

The classical Singleton bound

\[ |C| \leq q^{n-2d_{list}} \]

for the \((d_{list}, 1)\) list-decodable codes in [44] and [69] was generalized to

\[ |C| \leq Lq^{n - \frac{(L+1)d_{list}}{L}} \]

for \((d_{list}, L)\) list-decodable codes in [67]. When \(L = 2\) and \(L = 3\) the existence of list decodable Reed-Solomon codes attaining this bound was proved in [67]. An improvement on this bound was recently presented in [36]. The main focus of the bounds in [36] and [67] is for list-decodable codes over large fields. The main focus of this paper is for list-decodable codes over small fields.

D. Our Contributions

In this paper, we show that any small covering code with the covering radius smaller than or equal to \(R\) gives a good upper bound on the sizes of \((R, L)\) list-decodable codes. When \(L = 1\), from classical results on covering radius of various codes, good upper bounds on the sizes of codes with the given minimum Hamming distances \(\geq 2R + 1\) are obtained. This is the first paper which translates classical results about covering codes to upper bound sizes of general codes of given minimum distances. When the lengths are large these upper bounds are stronger than the Griesmer bound for linear codes. One of our Singleton type upper bounds is as follows, see Section IV.

Theorem 1: [see Section IV] Let \(p\) be an odd prime and \(m\) be a positive integer. Let \(R\) be an even positive integer satisfying \(R \leq 2(p^m + 1)\). An \((n, M, 2R + 1)\) code of the length \(n \geq p^{2m} - 1\) has to satisfy

\[ M \leq p^{n - Rm}. \]

Let \(R\) be an odd positive integer satisfying \(R \leq 2(p^n + 1)\). An \((n, M, 2R + 1)\) code of the length \(n \geq p^{2m} - 1\) has to satisfy

\[ M \leq p^{n - (R-1)m}. \]
All our generalized Singleton type upper bounds depend on small covering codes. Theorem 1 is based on covering codes in [73]. We refer to Corollary 6 for the binary case.

When lengths are larger, the positive integers $m$’s in the Singleton type upper bounds
\[ A_q(n, d) \leq q^{n-md} \]
are larger. These upper bounds are suitable to upper bound the sizes of subfield subcodes. From these upper bounds, we give some upper bounds on the lengths of general (both linear and non-linear) MDS codes and small Singleton defect codes. An upper bound on the lengths of list-decodable codes attaining the generalized Singleton bound in [67] is also presented. Some good covering codes as dual codes of linear codes with few nonzero weights in [28], [29], [30], and [73] can be obtained. Then some generalized Singleton type upper bounds can be proved by the using of these covering codes.

The generalized Singleton type upper bounds presented in this paper have wide applications to symbol-pair codes, insertion-deletion codes and locally recoverable codes. We show that the direct Singleton upper bounds on symbol-pair codes and insertion-deletion codes are far away from the tightness when the lengths of codes are large. Sizes of a $q$-ary symbol-pair code with a given minimum pair distance $d_{pair}$ or a $q$-ary insertion-deletion code with a given minimum insdel distance $d_{insdel}$ is much smaller than the direct Singleton bound
\[ |C| \leq q^{n-d_{pair}+2}, \]
or
\[ |C| \leq q^{n-d_{insdel}+1} \]
when the length is large.

An upper bound on the lengths of small dimension Singleton-optimal locally recoverable codes with any fixed minimum distance is given. We also give an upper bound on the lengths of small dimension optimal $(r, \delta)$ locally recoverable codes with any fixed given minimum distance. In Section VIII we prove the following upper bound on the maximal length of Singleton-optimal locally recoverable codes of small dimensions.

**Theorem 2:** [see Section VIII] For a Singleton-optimal locally recoverable code $C$ of the minimum Hamming distance $2R + 1$ or $2R + 2$ and the dimension $k \leq Rr$, the length $n$ of $C$ has to satisfy
\[ n \leq cq^{(2R+1)n} \cdot (lnq)^{1/R}, \]
where $c$ is a constant independent of $q$ and only depending on $R$.

II. COVERING CODE UPPER BOUNDS

The following covering code upper bound is general, since we can take any code $C'$ with the covering radius $R_{covering}(C') \leq d$. Actually any good upper bound for the covering radius implies a good upper bound on the sizes of $(d_{list}, L)$ list-decodable codes from our covering code upper bound.

**Theorem 3:** Let $C \subset \mathbb{F}_q^n$ be a $(d_{list}, L)$ list-decodable code. Suppose that $C' \subset \mathbb{F}_q^n$ is a code with the covering radius $R_{covering} \leq d_{list}$, then we have
\[ |C| \leq L|C'|. \]
Hence
\[ |C| \leq LK_q(n, d_{list}). \]

**Proof:** The balls of the radius $d_{list} \geq R_{covering}(C')$ centered at the codewords of $C'$ cover the whole space $\mathbb{F}_q^n$. Then in each such ball there are at most $L$ codewords of $C$. We have $|C| \leq L \cdot |C'|$. The other conclusions follow directly.

The covering upper bounds for a $(d_{list}, L)$ list-decodable code $C \subset \mathbb{F}_q^n$ are weaker than the sphere-packing upper bound
\[ |C| \leq L \cdot \frac{q^n}{\sum_{j=0}^{d_{list}} \binom{n}{j} (q-1)^j}. \]
When $n$ and $d_{list}$ are large, it is computationally infeasible to get clear and explicit upper bounds from the sphere-packing bound. Hence by applying various results in covering codes, explicit upper bounds on sizes of list-decodable codes in the Hamming metric space can be obtained. In this section some such upper bounds on length $n$ list-decodable codes are obtained for concrete length $n$.

From Theorem 3 and the asymptotic bound of covering codes, it follows immediately that for a sequence of $[n_i, k_i]_q$ codes $C_i \subset \mathbb{F}_q^{n_i}$ which is $(rn_i, L_i)$ list-decodable, where $n_i$ goes to the infinity, and the rate $\lim \frac{k_i}{n_i} \geq 1 - H_q(r) + \epsilon$, then the list sizes $L_i$ are exponentials of $n_i$. Then this sequence of codes is non list-decodable.

There are a lot of classical results about the binary covering codes, we refer to [4], [20], [31], and [56]. From the table [56] of Litsyn, we have the following upper bound on the binary $(3, L)$ list decodable code $C \subset \mathbb{F}_2^{16}$,
\[ |C| \leq 192L, \]
since $K_2(16, 3) \leq 192$. Thus if the linear $[16, 9, 4]_2$ code is $(3, L)$ list-decodable, the list size has to be at least $3$.

The almost trivial redundancy upper bound $R_{covering} \leq n - k$ for the covering radius of the linear $[n, k]_q$ code, see [20, page 217], and our covering code upper bounds imply an upper bound which is close to the generalized Singleton upper bound in [67], when the list size is large.

**Corollary 1:** A $(d_{list}, L)$ list-decodable code $C \subset \mathbb{F}_q^n$ have to satisfy
\[ |C| \leq Lq^{n-d_{list}}. \]

**Proof:** Let $C'$ be a linear $[n, k]_q$ code in $\mathbb{F}_q^n$, then $R_{covering}(C') \leq n - k$. This follows from the fact that for any $k$ information set coordinate positions $1 \leq i_1 < i_2 \cdots < i_k \leq n$, the coordinates at these positions of codewords in $C'$ can be arbitrary vectors in $\mathbb{F}_q$. For the second conclusion, we take a linear $[n, n - d_{list}]_q$ code $C'$, then $R_{covering}(C') \leq d_{list}$. The conclusion follows immediately.
This is weaker than the generalized Singleton upper bound
\[ |C| \leq Lq^{n+\frac{1}{2}(n+1)\log_{q}(n+1)}. \]
in [67]. However when the list size \( L \geq d_{\text{list}} \), this almost trivial upper bound is equivalent to the generalized Singleton upper bound in [67].

From the known linear perfect codes such as linear \( \frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m, q \) perfect Hamming code, binary perfect Golay [23, 12, 7] code and ternary perfect Golay [11, 6, 3] code, see, Chapter 11, [20], we have the following result.

**Corollary 2:** 1) Let \( n = \frac{q^m-1}{q-1} \), where \( q \) is a fixed prime power and \( m \geq 2 \) is an integer. Let \( k > n - m \) be a positive integer. Then if a linear \([n, k, q]_q\) code is \((1, L)\) list-decodable, then the list size has to satisfy
\[ L \geq q^{k-n+m}. \]

2) If a binary linear \([2^m, 2^m+1, 2^m-1]_2\) Reed-Muller code is \((3, L)\) list-decodable, then the list size has to satisfy
\[ L \geq 2^{k-12}. \]

3) If a ternary linear \([11, 6, 3]_3\) code is \((2, L)\) list-decodable, then the list size has to satisfy
\[ L \geq 3^{k-6}. \]

It is well-known that the covering radius of the first order \([2^m, m+1, 2^m-1]_2\) Reed-Muller code is \(2^m-1-2^{\frac{m+1}{2}}\) when \( m \) is even, see page 243 of [20]. Then the following upper bound for binary \((rn, L)\) list-decodable codes, where \( \frac{1}{2} < \frac{1}{2^{2n+r}} \leq r < \frac{1}{2} \), follows from our covering code bound. The second conclusion follows from Theorem 11.5.3 in page 450 of [41].

**Corollary 3:** 1) Let \( n = 2^m \), \( m \) is even. Let \( r \) be a positive real number satisfying \( \frac{1}{2} < r \leq \frac{1}{2^{2n+r}} \). If a binary length \( n = 2^m \) code \( C \) is \((rn, L)\) list-decodable, then
\[ |C| \leq L2^{m+1}. \]

Hence a length \( 2^m \) binary linear code with the minimum Hamming weight \( 2^m - 2^{\frac{m}{2}} + 1 \) has the dimension at most \( m+1 \).

2) Moreover for any binary \((s-1, L)\) list-decodable code \( C \) in \( F_{2^{s-1}} \) where \( s \) is any positive integer, we have \( |C| \leq 64L \).

For the positive list size \( L \geq 2 \) we have \( \frac{L-1}{2} \geq \frac{3}{2} \), hence the generalized Singleton upper bound in [67] for this case is at least
\[ L2^{2m-3}2^{m-2} = L2^{m-2}. \]

Our upper bound is an exponential \( 2^{\frac{1}{2} - \log_{2} n + 1} \) improvement of the bound in [67] for binary codes.

Let \( r \) and \( R \) be two fixed positive integers satisfying \( r \geq R \). Let \( l_q(r, R) \) be the smallest length of a linear \( q \)-ary code with the covering radius \( R \) and the redundancy \( r = n - k \). It is clear that there are \([n, n-r]_q\) code with the covering radius \( R \) for each integer \( n \geq l_q(r, R) \), see [41, Section 11.5]. Actually this can be proved directly from the characterization of the covering radius of linear codes from its parity check matrix. There are a lot of results about this topic, we refer to [4], [24], and [25] and references therein. From our covering code upper bounds the size of an \((R, L)\) list-decodable code \( C \) of length \( n \) satisfying \( n \geq l_q(r, R) \) has to satisfy
\[ |C| \leq Lq^{n-r}. \]

When the list size is not one, and \( r \geq \frac{3}{2} R \), the upper bound is better than the generalized Singleton bound in [67]. For example for \( r = 5 \geq \frac{3}{2} \cdot 3 \), the size of a length \( n \geq l_q(5, 3) \) \((3, L)\) list-decodable code satisfies \( |C| \leq Lq^{n-5} \). In this case, the conclusion \( l_q(5, 3) < 2.88q^{\frac{1}{2}}(1+q)/3 \) was proved in [4]. More generally the following result follows from the bound in [24].

**Corollary 4:** 1) Let \( q \geq 8 \) be an even prime power. Let \( R \geq 4 \) be a positive integer. Set \( m = [\log_q (R + 1)] + 1 \). For any given positive integer \( t \geq 3m + 2 \), and the length \( n \geq \frac{R}{q(t-1)}2^t - 2^{m+2} + \sum_{j=0}^{m+2} q^{m+j} \), the size of a \((R, L)\) list-decodable code \( C \subset F_q^n \) has to satisfy
\[ |C| \leq Lq^{n-tR}. \]

**Proof:** The conclusion follows from Theorem 8 in [24].

We observe that when \( t \geq 3 \), the \((R, L)\) list-decodable codes in Corollary 4 cannot attain the generalized Singleton bound in [67], since \( \frac{(t+1)R}{t} < 3R \). Hence this gives an upper bound on the lengths of codes attaining the generalized Singleton bound in [67].

III. COVERING CODES FROM LINEAR CODES WITH FEW NONZERO WEIGHTS AND THE RELATED UPPER BOUNDS

One interesting point of the classical theory of covering codes is the upper bounds on the covering radius of a code from its dual code. This was originated from the work of [26]. We refer to [26] and [42] for some upper bounds on the covering radii from dual codes. The Delsarte upper bound on the covering radius of a linear code \( C \) asserts
\[ R_{\text{covering}}(C) \leq s, \]
where \( s \) is the number of nonzero weights of its dual code, see [41] page 440 Theorem 11.3.3. Hence we have the following result.

**Corollary 5 (Delsarte Type Bound):** Let \( C \subset F^n_2 \) be a \((d_{\text{list}}, L)\) list-decodable code. Suppose that \( C_1 \subset F^n_q \) is a linear code such that the total number of nonzero weights of \( C_1^\perp \) is smaller than or equal to \( d_{\text{list}} \), then we have
\[ |C| \leq L|C_1|. \]

The construction and the combinatorial structure of linear codes with few nonzero weights are the classical topic in coding theory, Boolean functions and sequences, we refer to [9], [10], [27], [28], [29], [30], [50], [54], [55], [63], [72], and [73] for many such linear code constructions and the study of their combinatorial structures. The binary simplex code and the Kasami code are typical such linear codes. From the Delsarte upper bound on the covering radius, we have a lot of good covering codes over small fields as in the following examples. The length function \( l_q(r, R) \) for covering codes has been studied in [21], [22], [24], and [25]. Actually these construction of linear codes with few nonzero weights can be thought as upper bounds on this length function \( l_q(r, R) \) for covering codes.

**Example 1:** The Kasami codes proposed in [50] are binary linear \([2^{2m+1} - 1, 3m, 2^{2m+1} - 2^{m+1}]_2 \) codes with three nonzero weights, \( w_1 = 2^{2m-1} - 2^{m-1}, w_2 = 2^{2m-1} \) and \( w_3 = \)
2^{m-1} + 2^{m-1}, m = 1, 2, \ldots. Its dual is a binary linear 
[2^{2m} - 1, 2^{2m} - 3m]_2 code with the covering radius at most 3. Therefore \( l_2(3m, 3) \leq 2^{2m} - 1 \). These dual codes of Kasami codes as covering codes can be compared with linear covering codes with the covering radius 3 in [21] and [22].

Example 2: Some binary linear \([2^{m-1} - 1, m]_2\) codes with three nonzero weights were constructed in [27], \( m = 2, 3, \ldots \). Their dual codes are linear codes with the covering radius at most 3 and the codimension \( m \). Then \( l_2(m, 3) \leq 2^{m-1} - 1 \).

Let \( m \) be an odd positive integer. Then a \([2^{m-1}, m+1]_2\) code with 4 nonzero weights was constructed in [30]. The dual code is a binary linear \([2^{2m-1}, 2^{2m-1} - m - 1]_2\) code with the covering radius at most 4. Then \( l_2(m+1, 4) \leq 2^{m-1} \).

Example 3: Let \( m \) be an even positive integer and \( k/m \) be a positive divisor of \( m \) satisfying \( k \neq m \) and \( k \neq \frac{m}{2} \). A binary linear \([2^{m-1}, m+k]_2\) code with 3 nonzero weights was constructed in Theorem 2 of [72, page 5143]. The dual code is a binary linear \([2^{m-1}, 2^{m-1} - m - k]_2\) code with the covering radius at most 3. Then \( l_2(m+k, 3) \leq 2^{m-1} \). This is weaker than the bound in Example 2.

Let \( m \) be an even positive integer of the form \( 2(2k + 1) \) and \( t \) be a positive divisor satisfying \( t \neq 2^{m/2} + 1 \), \( lcm(t, 3)(2^{m/2} + 1), 3|t \). Then a binary linear \([\frac{2^{m-1}}{t}, m+1]_2\) code with 5 nonzero weights was constructed in Theorem 24 of [72, page 5144]. The dual code is a binary linear \([\frac{2^{m-1}}{t}, 2^{m-1} - m - 1]_2\) code with the covering radius at most 5. Then \( l_2(m+1, 5) \leq \frac{2^{m-1}}{t} \).

Let \( m \) be an even positive integer and \( t \) be a positive divisor satisfying \( t \neq 2^{m/2} + 1 \), \( lcm(t, 3)(2^{m/2} + 1), t \) is not a multiple of 3. Then a binary linear \([\frac{2^{m-1}}{t}, m+2]_2\) code with 6 nonzero weights was constructed in Theorem 24 of [72, page 5144]. The dual code is a binary linear \([\frac{2^{m-1}}{t}, 2^{m-1} - m - 2]_2\) code with the covering radius at most 6. Then \( l_2(m+2, 6) \leq \frac{2^{m-1}}{t} \).

Example 4: Let \( m \) be a positive integer satisfying \( m \equiv 0 \mod 6 \). Then a binary cyclic \([2^{m-1}, 5m]_2\) code with 7 nonzero weights was constructed in [71]. From the Delsarte upper bound, the dual code is a binary linear \([2^{2m-1}, 2^{2m-1} - 1 - \frac{5m}{2}]_2\) code with the covering radius at most 7. Then \( l_2(\frac{2m}{7} - 7) \leq 2^{m-1} - 1 \) for \( m = 6, 12, \ldots \).

Example 5: Let \( q \) be an odd prime power and \( m \) be an odd positive integer. Then a linear \([q^m - 1, 2m]_q\) code with 3 nonzero weights was constructed in Theorem 1 of [54]. The dual code is a linear \([q^m - 1, q^m - 1 - 2m]_q\) code with the covering radius at most 3. Then \( l_2(2m, 3) \leq q^m - 1 \).

Let \( q \) be an odd prime power and \( m \) be an even positive integer. Then a linear \([q^m - 1, 2m]_q\) code with 4 nonzero weights was constructed in Theorem 1 of [54]. The dual code is a linear \([q^m - 1, q^m - 1 - 2m]_q\) code with the covering radius at most 4. Then \( l_2(2m, 4) \leq q^m - 1 \).

Let \( q \) be an odd prime power and \( m \) be an odd positive integer. Then a linear \([q^m, 2m+1]_q\) code with 4 nonzero weights was constructed in Theorem 2 of [54]. The dual code is a linear \([q^m, q^m - 1 - 2m]_q\) code with the covering radius at most 4. Then \( l_2(2m+1, 4) \leq q^m \).

Let \( q \) be an odd prime power and \( m \) be an even positive integer. Then a linear \([q^m, 2m+1]_q\) code with 6 nonzero weights was constructed in Theorem 2 of [54]. The dual code is a linear \([q^m, q^m - 1 - 2m]_q\) code with the covering radius at most 6. Then \( l_2(2m+1, 6) \leq q^m \).

Let \( q = p^m \) be an odd prime power with the even exponent \( m \), then a linear \([\frac{q^m - 1}{2}, m, (p+1)(q^m - \sqrt{q} + 1)]_p\) code with two nonzero weights is explicitly given in Corollary 4 of [28]. Hence for any \((2, L)\) list-decodable code \( C \subset F_p^2 \), we have

\[ |C| \leq Lp^{\frac{2m}{2} - m}, \]

from the above Delsarte type upper bound.

From Theorem 1 of [73] we get the following upper bound from the above Delsarte type upper bound on length \( q^2 - 1 \) code over \( F_p \), \( q = p^m \). Let \( h \) be a fixed positive integer satisfying \( h \equiv 0 \mod q + 1, e = gcd(h, q+1), t \leq \frac{q+1}{2e} \). Then when \( p = 2 \) the cardinality of an \((2t + 1, L)\) list-decodable code \( C \subset F_p^{2^{m-1}} \) has to satisfy

\[ |C| \leq L \cdot 2^{2^{m-1} - 2(2t+1)m}. \]

When \( p \) is an odd prime, then the cardinality of a \((2t, L)\) list-decodable code \( C \subset F_p^{2^{m-1}} \) has to satisfy

\[ |C| \leq L \cdot p^{2^{m-1} - 2tm}. \]

IV. Generalized Singleton Type Upper Bounds

When \( L = 1 \), we can obtain many upper bounds on sizes of codes with the given minimum Hamming distance \( d \), from good covering codes applied in Theorem 3. It is verified in this section that these upper bounds are much better than the Griesmer bounds, when lengths are large. By combining the best known covering codes in [56], we show that our generalized Singleton type upper bounds are strong for presently best known codes in [38]. For codes of the given minimum Hamming distance over a small field, the classical Singleton bound is far away from the tightness. In many cases, the Griesmer bound was used to get the upper bound on its minimum distances, see the table [38] of presently known best codes.

A. Generalized Singleton Type Upper Bound I

If there is a covering code \( C' \subset F_q^n \) with the covering radius \( R \). Then the size of a code \( C \subset F_q^n \) with the minimum Hamming distance \( 2R + 1 \) satisfies

\[ |C| \leq |C'|. \]

We can compare the above upper bound for small length best known binary codes in [38]. For example from Livingston's table [56], \( K_2(15,3) = 112 \). Then from the general Singleton bound I the cardinality we have \( M \leq 112 \) for the \((15, M, 7)_2\) code. The presently best known code is a linear \([15, 5, 7]_2\) code, see [38]. The ratio is \( K_2(15,3) \leq \frac{K_2(15,3)}{2^d} \leq 4 \). From Livingstone's table [56], for a binary \((33, M, 9)_2\) code, \( M \leq K_2(33,4) \leq 2^{17} \cdot 3 \) from the generalized Singleton bound I. The presently best known code is a linear \([33, 14, 9]_2\) code. The ratio is \( K_2(33,4) \leq 24 \). Similarly for a
binary \((33, M, 11)\) code, \(M \leq 2^{13} \cdot 11\) from the best known covering code in [56]. The presently best known binary code is a linear \([33, 12, 11]_2\) code. The ratio is \(k_2(33,5) = 22\). From the generalized Singleton bound I, for a \([32, M, 7]_2\) code \(M \leq 2^{20}\). The presently best known code is a binary \([33, 17, 8]_2\) code, see [38]. The ratio is \(k_2(32,7) = 8\). This illustrates that the above generalized Singleton bound I is strong even for small length binary codes, since the ratio \(k_2(n,M) = \frac{2^m}{2^r}\) is small, where \(k\) is the dimension of the best known \([n, k, 2R + 1]_2\) code.

The following result follows from the binary linear codes with few nonzero weights constructed in [73].

Corollary 6: Let \(R\) be an odd positive integer satisfying \(R \leq 2(2^m + 1) + 1\). An \((n, M, 2R + 1)_2\) code of the length \(n \geq 2^{2m} - 1\) has to satisfy
\[
M \leq 2^{n - Rm}.
\]

Let \(R\) be an even positive integer satisfying \(R \leq 2(2^m + 1) + 1\). An \((n, M, 2R + 1)_2\) code of the length \(n \geq 2^{2m} - 1\) has to satisfy
\[
M \leq 2^{n - (R-1)m}.
\]

Actually any good upper bound on the length function \(l_q(r, R)\) studied in [22], [24], and [25] gives a generalized Singleton type upper bound as follows.

B. Generalized Singleton Type Upper Bound II

Let \(l_q(r, R)\) be the length function for covering codes. Then an \((n, M, 2R + 1)_2\) code of the length \(n \geq l_q(r, R)\) has to satisfy \(M \leq q^{n-r}\).

Proof: Let \(C\) be a linear covering code of the length \(l_q(r, R)\), covering radius \(R\) and the codimension \(r\). For any nonnegative integer \(s \geq 0\), the linear covering code \(C \times F_q^{s}\) of length \(l_q(r, R) + s\) has its covering radius \(R\) and codimension \(r\). The conclusion follows immediately.

There are some upper bounds on the length function \(l_q(t, R)\) and \(l_q(t + 1, R)\), see [24], [25]. Then if \(t \geq 4 + \epsilon\), where \(\epsilon > 0\) is a small positive real number, the generalized Singleton bound for codes of the length at least \(l_q(t, R)\) is stronger than the Griesmer bound. The following corollary follows from the upper bound on the length function for covering codes in [25].

Corollary 7: Let \(q\) be an arbitrary prime power. Let \(t\) and \(R\) be two positive integers satisfying \(R \geq 3\) and \(t \geq 1\). Then an \((n, M, 2t + 1)\) code of the length \(n \geq c q \left( \frac{m}{t(m-1)} - 1\right)\) has to satisfy
\[
M \leq q^{n-tR-1}.
\]

Here \(c\) is a constant independent of \(q\).

C. Generalized Singleton Type Upper Bound III

Let \(p\) be an odd prime and \(m\) be a positive integer. Let \(R\) be an even positive integer satisfying \(R \leq 2(p^m + 1)\). An \((n, M, 2R + 1)_p\) code of the length \(n \geq p^{2m} - 1\) has to satisfy
\[
M \leq p^{n - Rm}.
\]

Let \(R\) be an odd positive integer satisfying \(R \leq 2(p^m + 1)\). An \((n, M, 2R + 1)_p\) code of the length \(n \geq p^{2m} - 1\) has to satisfy
\[
M \leq p^{n - (R-1)m}.
\]

This upper bound follows from the linear codes with \(R\) nonzero weights constructed in [73]. An upper bound
\[
k \leq \left\lfloor \frac{n - d - m + 3 - \log_q((n - d - m)(q - 1) + 1)}{m} \right\rfloor
\]
for a linear \([n, k, d]_q\) code satisfying \(d > mq, m \geq 3, n - d - m + 2 \geq 0\), was proved in [60]. When \(q = 2\), we have
\[
k \leq \left\lfloor \frac{n - d - m + 3 - \log_2((n - d - m)(2 - 1) + 1)}{m} \right\rfloor.
\]

It is clear our bounds Corollary 6 and Generalized Singleton type bound III are much stronger when \(n\) is large.

Let \(q\) be a prime power and \(m\) be a positive integer. Let \(e \geq 2\) be a divisor of \(q^m - 1\) and \(a\) be a positive integer satisfying \((q^m - 1)\) is not a divisor of \(a\). Let \(\Delta_1, \ldots, \Delta_e\) be \(e\) nonnegative integers satisfying \(\Delta_i \neq \Delta_j \mod e\), for \(i \neq j\), and \(\gcd(\Delta_2 - \Delta_1, \ldots, \Delta_e - \Delta_1) = 1\). For example we can take \(\Delta_i = i\). Set \(a_i = a + \frac{q^m - 1}{\Delta_i}\), \(\delta = \gcd(q^m - 1, a_1, \ldots, a_e)\), and \(n_1 = \frac{q^m - 1}{\delta}\). Let \(N = \gcd(\frac{q^m - 1}{\delta}, d)\). We assume the condition \(N = 1\) is satisfied. Then a linear cyclic \([n_1, em]_q\) code with \(e\) nonzero weights was constructed in [74] Theorem 7.

D. Generalized Singleton Type Upper Bound IV

An \((n, M, 2e + 1)_q\) code of the length \(n \geq n_1\) has to satisfy
\[
M \leq q^{n-em}.
\]

We can compare this bound with the Griesmer bound for a linear \([n, k, d]_q\) code. The Griesmer bound claims
\[
n \geq \sum_{i=0}^{k-1} \left\lfloor \frac{d}{q^i} \right\rfloor.
\]

From the Griesmer bound the right side is smaller than or equal to \(2d + k - 1\). That is, we can get at most \(k \leq n - 2d + 1\) from the Griesmer bound for a linear \([n, k, d]_q\) code. However when \(t \geq 5\), our above generalized Singleton type bounds is much stronger than the Griesmer bound when the length is large. The above generalized Singleton type bounds are suitable to upper bound the sizes of various subfield subcodes.

Example 6: We consider the subfield subcode \(C_{\mathbb{F}_2}\) of the generalized Reed-Solomon \([2^{2m}, 2k, 2^{2m} - 2k + 1]_{2^{2m}}\) code \(C \subset \mathbb{F}_{2^{2m}}\), where \(k\) is a positive integer satisfying \(k \geq 2^{2m-1} - 2^m\). This is a binary linear \([2^{2m}, k_1, 2^{2m-1} - 2^m + 1]_2\) code with the unknown dimension \(k_1\). It follows from Corollary 6,
\[
k_1 \leq 2^{2m} - m(2^{2m-1} - k).
\]

E. Generalized Singleton Type Upper Bound V

Let \(m\) and \(u\) be two positive integers satisfying \(m \geq 3\). Then the cardinality of an \((n, M, 2u + 1)_q\) code of the length \(n \geq u \cdot \frac{q^m - 1}{q - 1}\) satisfies \(M \leq q^{n - mu}\).
Proof: Let \( H_m \) be an \( m \times \frac{q^m-1}{q-1} \) matrix over \( F_q \), with \( \frac{q^m-1}{q-1} \) columns taking all nonzero vectors in \( F_m^q \). Let \( H \) be a \( um \times \frac{u(q^m-1)}{q-1} \) matrix of the following form.

\[
\begin{pmatrix}
H_m & 0 & 0 & \cdots & 0 \\
0 & H_m & 0 & \cdots & 0 \\
0 & 0 & H_m & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & H_m
\end{pmatrix}
\]

Then \( H \) is a full-rank matrix such that each vector in \( F_m^{um} \) can be represented as the sum of at most \( u \) columns of \( H \). Let \( C \) be the linear \( [\frac{u(q^m-1)}{q-1}, \frac{u(q^m-1)}{q-1} - um]_q \) code defined by the parity-check matrix \( H \). Then the covering radius of \( C \) is at most \( u \). On the other hand if \( n = \frac{u(q^m-1)}{q-1} + n_1 \) satisfying \( n_1 < \frac{q^m-1}{q-1} \), we take the covering code \( C \times F_m^{n_1} \subseteq F_m^n \) with the covering radius \( u \). The conclusion follows from Theorem 3 immediately.

Actually from examples in the previous section, we can list more generalized Singleton upper bounds on the sizes of codes with a given minimum Hamming distance.

V. LENGTHS OF GENERAL MDS CODES AND SMALL SINGLETON DEFECT CODES

The construction of MDS codes is a fundamental problem. For example, there have been a lot of efforts in the construction of MDS quantum codes, MDS entanglement-assisted quantum codes, the construction of MDS codes in symbol-pair metric and the construction of Singleton-optimal locally recoverable codes, see [8], [13], [17], [48] and references therein. Thus it is always interesting to give an upper bound on lengths of various MDS codes containing various Singleton bounds. In this section, we give some upper bounds on the length of general MDS codes and almost MDS codes (then near MDS codes) in the Hamming metric. Actually our method is not restricted to general MDS or almost MDS codes satisfying that \( n+1-(k+d) = 1 \), our upper bounds work for general small Singleton defect codes.

Theorem 4: If there exists a linear \([n, n-d]_q\) code with the covering radius \( \lceil \frac{d-1}{2} \rceil \), then the length of a general MDS code can not be bigger than \( n \).

Proof: It is clear that we have a linear \([n+s, n-d+s]_q\) covering code \( C \times F_s^q \subseteq F_{s+d}^q \), \( s \geq 0 \), with the covering radius \( \lceil \frac{d-1}{2} \rceil \). From the generalized Singleton bound I, an \((n+s, M, d)_q\) code has to satisfy

\[
M \leq q^{n+s-d}.
\]

This is not an MDS code.

From Theorem 4, if there could exist a covering \([q+2, q+2-d]_q\) code \( C \subseteq F_{q+2}^q \) with the covering radius \( \lceil \frac{d-1}{2} \rceil \) for \( d = 1, 2, \ldots, \frac{q+2}{2} \), then the main conjecture follows immediately. However this claim seems too strong for covering codes.

From the generalized Singleton bound III, an \((n, M, 2R+1)_q\) code with the length \( n \geq l_q(2R+1, R) \) has to satisfy \( M \leq q^{n-2R-1} \), that is, the code is not MDS. Hence an MDS \((n, q^{n-d+1}, d)_q\) code, \( d = 2R + 1, 2R + 2 \), can not have length \( n \geq l_q(2R+1, R) \). From the bound in Corollary 7, an \((n, M, 2R+1)_q\) MDS code or an \((n, M, 2R+2)_q\) MDS code can not have the length \( n \geq cq^{\frac{n+1}{2}} \cdot (lnq)^{\frac{k}{4}} \), where \( c \) is a constant independent of \( q \). In these upper bounds on lengths of general MDS codes, the upper bounds depend on the minimum Hamming distances.

From generalized Singleton bounds III and IV in the case \( m \geq 3 \), we can have upper bounds on the lengths of an \((n, M, d)_q\) MDS code with the minimum distance \( d \leq (p^m + 1) \) or \( d+1 \) is a divisor of \( q^{m-1} - 1 \).

Theorem 5: Let \( q \) be a prime power, \( m \) be a positive integer satisfying \( m \geq 3 \) and \( e \geq 2 \) be a divisor of \( q^{m-1} - 1 \). We assume that the condition \( N = 1 \) in Section V is satisfied. Then an \((n, M, 2e+1)_q\) MDS code or an \((n, M, 2e+2)_q\) MDS code can not have the length \( n \geq q^{m-1} - 1 \).

Corollary 8: Let \( e \geq 2 \) be an odd divisor of \( q-1 \) satisfying \( \gcd(e, 3) = 1 \). Then an \((n, M, 2e+2)_q\) MDS code or almost MDS code can not have length \( n \geq q^{m-1} - 1 \).

Proof: Set \( m = 2 \). We have \( \gcd(q^2 + q + 1, q-1) \leq 3 \). Then \( \gcd(q^2-1, ae) = 1 \) if we set \( a = 1 \). The conclusion follows from Theorem 5 immediately.

In [5], it was proved that for a linear \([n, k, d]_q\) code with the Singleton defect \( n + 1 - k - d = s \), and satisfying \( d > q \) the length \( n \) has to satisfy

\[
n \leq d - 2 + \frac{2(q^{s+1} - 1)}{q - 1}.
\]

Corollary 8 can be thought as an upper bound on the lengths of MDS codes and almost MDS codes with small Hamming distances. Moreover our bounds are for both linear and non-linear general MDS codes and almost MDS codes.

When the list size is 2 or 3, Reed-Solomon codes with \( n << q \) attaining the generalized Singleton bound \(|C| \leq Lq^n - \left(\frac{l_q(n)}{2}\right)\) were constructed in [67]. Since for any list size \( L \), \( \left(\frac{l_q(n)}{2}\right) < 2R + 1 \), we have the following upper bound on the lengths of \((R, L)\) list-decodable codes attaining the generalized bound in [67], from Corollary 7.

Theorem 6: Let \( q \) be a prime power and \( R \) be a fixed positive integer satisfying \( R \geq 3 \). If \( C \) is a \((R, L)\) list-decodable code attaining the generalized Singleton bound \(|C| = L \cdot q^n - \left(\frac{l_q(n)}{2}\right)\). Then the length \( n \) has to satisfy

\[
n \leq l_q(2R + 1, R) \leq cq^{\frac{R+1}{2}} \cdot (lnq)^{\frac{k}{4}}.
\]

VI. GENERALIZED SINGLETON TYPE UPPER BOUNDS ON SYMBOL-PAIR CODES

Symbol-pair codes were introduced for high density data storage, we refer to [11] and [33]. Set \( (F_q^2)^n = \{(x_1, x_2), (x_2, x_3), \ldots, (x_{n-1}, x_n), (x_n, x_1) : x = (x_1, \ldots, x_n) \in F_q^n\} \). The pair metric on \( F_q^n \) is defined as follows. For \( x = (x_1, \ldots, x_n) \in F_q^n \), the mapping \( \pi_2 : F_q^n \rightarrow (F_q^2)^n \) is defined by \( \pi_2(x) = ((x_1, x_2), \ldots, (x_n, x_1)) \in (F_q^2)^n \). The weight of \( x \in F_q^n \) is

\[
wt_{pair}(x) = wt(\pi_2(x)) = |\{i : (x_i, x_{i+1}) \neq 0\}|.
\]
The pair distance \(d_{\text{pair}}\) is defined by
\[
d_{\text{pair}}(x, y) = wt_{\text{pair}}(\pi_2(x) - \pi_2(y)),
\]
see [11]. It follows from the definition that
\[
\min\{d(x, y) + 1, n\} \leq d_{\text{pair}}(x, y) \leq 2d(x, y).
\]
For a code \(C \subset \mathbb{F}_q^n\), the minimum pair distance is
\[
d_{\text{pair}}(C) = \min\{d_{\text{pair}}(x, y) : x \neq y, x \in C, y \in C\}.
\]
It is obvious that
\[
d_{\text{pair}}(C) \leq 2d(C).
\]

The direct Singleton bound
\[
|C| \leq q^{n-d_{\text{pair}}+2}
\]
on symbol-pair codes was proved in [12] and some MDS symbol-pair codes have been constructed in [12], [48], and [59]. For some new bounds on symbol-pair codes, we refer to the recent paper [33]. From the inequality
\[
d_{\text{pair}}(C) \leq 2d(C),
\]
we have the following result directly from our generalized Singleton type upper bounds in Section IV.

**Corollary 9:** If the length of a symbol-pair code \(C\) is bigger than the length \(n_1\) in Corollary 6, the generalized Singleton upper bounds II, III, V in Section IV, and the minimum Hamming distance of \(C\) satisfies the requirements in Corollary 6, the generalized Singleton upper bounds II, III, V in Section IV, then the size of this symbol-pair code has to satisfy
\[
|C| \leq q^{n-2d_{\text{pair}}}
\]
or
\[
|C| \leq q^{n-4d_{\text{pair}}},
\]
When \(m\) or \(t\) is bigger than 7, these upper bounds are much stronger than the above direct Singleton bound in [12] for symbol-pair codes if the lengths are large. The above bound can be used to give some upper bounds on the maximal lengths of MDS symbol-pair codes, if their minimum pair distance satisfies the condition in Corollary 9.

**VII. Generalized Singleton Type Upper Bounds on Insertion-Deletion Codes**

Insertion-deletion codes have been extensively studied in recent years, see [16], [19], [35], [45] and references therein. The insdel distance \(d_{\text{insdel}}(a, b)\) between two vectors \(a\) and \(b\) in \(\mathbb{F}_q^n\) is the minimal number of insertions and deletions which are needed to transform \(a\) into \(b\). Actually it was well-known, see [19], that
\[
d_{\text{insdel}}(a, b) = 2(n - l),
\]
where \(l\) is the length of a longest common subsequence of \(a\) and \(b\). This insdel distance \(d_{\text{insdel}}\) is indeed a metric on \(\mathbb{F}_q^n\).

It is clear \(d_{\text{insdel}}(a, b) \leq 2d(a, b)\) since \(l \geq n - d(a, b)\) is valid for arbitrary two different vectors \(a\) and \(b\) in \(\mathbb{F}_q^n\). The insdel distance \(d_{\text{insdel}}(C)\) of a code \(C \subset \mathbb{F}_q^n\) is the minimum of the insdel distances of all different two codewords in this code. Hence we have the direct upper bound \(d_{\text{insdel}}(C) \leq 2d(C)\), and the direct Singleton bound on the insdel distance of a linear \([n, k]\) code
\[
d_{\text{insdel}} \leq 2(n - k + 1),
\]
or
\[
|C| \leq q^{n-\left[\frac{d_{\text{insdel}}}{2}\right]+1},
\]
see [16], [19]. For linear insertion-deletion codes, the half-Singleton bound
\[
d_{\text{insdel}} \leq 2(n - k + 2),
\]
was proposed and proved in [19], also see [16]. This half-Singleton bound was improved in [45] to
\[
d_{\text{insdel}} \leq 2(n - k + 1),
\]
for linear insertion-deletion codes without the all 1 codeword. In this section, we prove that for both linear and nonlinear insertion-deletion codes, the above upper bounds are far from the tightness when the length are large compared with \(q\). This is a direct application of the above Singleton type upper bounds in Section IV, since \(d_{\text{insdel}}(C) \leq 2d(C)\).

**Corollary 10:** If the length of a general insertion-deletion code \(C\) is bigger than the length \(n_1\) in Corollary 6, the generalized Singleton upper bounds II, III, V in Section IV, and the minimum Hamming distance of \(C\) satisfies the requirements in Corollary 6, the generalized Singleton upper bounds II, III, V in Section IV, then the size of this insertion-deletion code has to satisfy
\[
|C| \leq q^{n-\frac{2}{3}d_{\text{insdel}}}
\]
or
\[
|C| \leq q^{n-\frac{4}{3}d_{\text{insdel}}}.
\]

When \(m\) or \(t\) is bigger than 6, these upper bounds are much stronger than the direct Singleton bound for general insertion-deletion codes if the lengths are large. In a recent paper [57], it was also showed that the direct Singleton bound on insertion-deletion codes are not tight. Our result is stronger than their bounds when the minimum insdel distance \(d_{\text{insdel}}\) is small.

**VIII. Lengths of Optimal Locally Recoverable Codes**

For a linear code \(C\) over \(\mathbb{F}_q\) with length \(n\), dimension \(k\) and minimum Hamming distance \(d\), its locality is defined as follows. Given \(a \in \mathbb{F}_q\), set \(C(i, a) = \{x \in C : x_i = a\}\), where \(i \in \{1, \ldots, n\}\) is an arbitrary coordinate position. The set \(C_A(i, a)\) is the restriction of \(C(i, a)\) to the coordinate positions in \(A \subset \{1, \ldots, n\}\). The linear code \(C\) is a locally recoverable code with locality \(r\), if for each \(i \in \{1, \ldots, n\}\), there exists a subset \(A_i \subseteq \{1, \ldots, n\} - \{i\}\) of cardinality at most \(r\) such that \(C_{A_i}(i, a) \cap C_{A_i}(i, a') = \emptyset\) for any given \(a \neq a'\). A Singleton-like bound for LRC codes
\[
d \leq n - k + 2 - \left\lceil \frac{k}{r} \right\rceil
\]
was proved in [37]. It is clear that this upper bound is just the Singleton bound \( d \leq n - k + 1 \) for linear codes when \( r \geq k \). A linear code attaining this upper bound is called a Singleton-optimal locally recoverable code, see [58]. The maximal length of a Singleton-optimal LRC code was studied in [13] and [58]. In this section we show that our generalized Singleton upper bounds for Hamming error-correcting codes give a natural upper bound on the lengths of Singleton-optimal LRC codes with a given minimum distance \( 2R + 1 \) or \( 2R + 2 \).

In [13] and [58] and references therein, the maximum lengths of Singleton-optimal LRC codes were upper bounded for the given minimum distances \( d = 5, 6, 7 \) with some additional conditions on the locality \( r \). For example in [58], it was proved that for a Singleton-optimal LRC code with the minimum Hamming distance \( d = 7 \), if the locality satisfies \( r \geq 6 \) and \( r = o(n) \), then \( n = O(q^{r/2}) \).

In [13] about the length of Singleton-optimal LRC codes of the minimum Hamming distances 5 and 6, the strong condition that the recovery sets are disjoint is needed. Our following upper bound is valid for any given minimum Hamming distance \( d = 2R + 1 \) or \( 2R + 2 \), with an additional condition \( k \leq R \).

**Corollary 11:** For a Singleton-optimal locally recoverable code \( C \) of the minimum Hamming distance \( 2R + 1 \) or \( 2R + 2 \) and the dimension \( k \leq Rr \), the length \( n \) of \( C \) has to satisfy

\[
  n \leq cq^{\frac{(2R+1)}{R}} \cdot (lnq)^{1/R},
\]

where \( c \) is a constant independent of \( q \) and only depending on \( R \).

**Proof:** This is direct from the generalized Singleton type upper bound in Corollary 7.

Similarly from generalized Singleton type upper bound VI, we have the following result.

**Corollary 12:** For a Singleton-optimal locally recoverable code \( C \) of the minimum Hamming distance \( 2R + 1 \) or \( 2R + 2 \) and the dimension \( k \leq Rr \), the length \( n \) of \( C \) has to satisfy

\[
  n \leq R \cdot \frac{q^k - 1}{q - 1}.
\]

A linear \([n, k, d]_q \) code \( C \) is called \((r, \delta)\) locally recoverable if for each coordinate position \( j \in \{1, \ldots, n\} \) there is a subset \( S_j \subset \{1, \ldots, n\} \) satisfying, \( j \in S_j, |S_j| \leq r + \delta - 1 \), and the minimum Hamming distance of the punctured code \( C|_{S^c} \) is at least \( \delta \). The above locally recoverable code is just the case of \( \delta = 2 \). The Singleton-like bound for an \((r, \delta)\) locally recoverable code is

\[
  d \leq n - k + 1 - \left(\left\lfloor \frac{k}{r} \right\rfloor - 1\right)(\delta - 1).
\]

The code attaining this bound is called an optimal \((r, \delta)\) locally recoverable code. The maximal lengths of optimal \((r, \delta)\) locally recoverable code was studied in [8] and an upper bound was given. We also give an upper bound on the lengths of small dimension \((r, \delta)\) locally recoverable codes using our generalized Singleton type upper bound.

**Corollary 13:** For an optimal \((r, \delta)\) locally recoverable code \( C \) of the minimum Hamming distance \( 2R + 1 \) or \( 2R + 2 \) and the dimension \( k \leq \left(\frac{R+1}{R} + 1\right)r \), the length \( n \) of \( C \) has to satisfy

\[
  n \leq cq^{\frac{(2R+1)}{R}} \cdot (lnq)^{1/R},
\]

where \( c \) is a constant independent of \( q \) and only depending on \( R \). Or we have

\[
  n \leq R(q^l - 1) \frac{q^l - 1}{q - 1}
\]

from the generalized Singleton type upper bound VI.

**Proof:** This upper bound follows from the generalized Singleton type upper bound in Corollary 7 and the generalized Singleton upper bound V.

This upper bound on the lengths of optimal \((r, \delta)\) locally recoverable codes is stronger than the general upper bound in [8] when the dimensions are small.

**IX. Conclusion**

In this paper, we showed that any small covering code \( C' \) in a general finite metric space with the covering radius \( R(C') \leq d_{\text{list}} \) gives a good upper bound on the sizes of \((d, L)\) list-decodable codes. When the list size is \( L = 1 \), many generalized Singleton type upper bounds on the sizes of codes with a given minimum Hamming distance \( d \) can be obtained from small covering codes with the covering radius \( \leq \frac{d-1}{2} \). These upper bounds are stronger than the Griesmer bound and some previous upper bounds when the lengths are large.

From our generalized Singleton type upper bounds, some upper bounds on the length of general small Singleton defect codes were presented. We also showed that many previous constructions of linear codes with few nonzero weights give natural covering codes from the Delsarte bound on covering radii. As applications, generalized Singleton type upper bounds for long symbol-pair codes and long insertion-deletion codes were given. Upper bounds on the lengths of small dimension optimal locally recoverable codes and small dimension \((r, \delta)\) locally recoverable codes with any given minimum distance were also presented.

**Acknowledgment**

The authors would like to thank an Associate Editor, Prof. Sudhir Ghorpade, and two reviewers for their helpful comments and suggestions that improved the presentation of this paper.

**References**

[1] T. L. Alderson, “On the weights of general MDS codes,” IEEE Trans. Inf. Theory, vol. 66, no. 9, pp. 5414–5418, Sep. 2020.

[2] S. Ball, “On sets of vectors of a finite vector space in which every subset of basis size is a basis,” J. Eur. Math. Soc., vol. 14, no. 3, pp. 733–748, Mar. 2012.

[3] S. Ball and M. Lavrauw, “Planar arcs,” J. Comb. Theory Ser. A, vol. 160, pp. 261–287, 2018.

[4] D. Bartoli, A. A. Davydov, S. Marcugini, and F. Pambianco, “Tables, bounds and graphics of short linear codes with covering radius 3 and codimension 4 and 5,” 2020, arXiv:1712.07078v6.

[5] M. A. de Boer, “Almost MDS codes,” Designs, Codes Cryptography, vol. 9, no. 2, pp. 143–155, Oct. 1996.

[6] A. A. Bruen, J. A. Thas, and A. Blokhuis, “On MDS codes, arcs in PG(n, q) with q even, and a solution of three fundamental problems of B. Segre,” Invent. Math., vol. 92, no. 3, pp. 441–459, 1988.
[66] R. M. Roth and V. Skachek, “Improved nearly-MDS expander codes,” IEEE Trans. Inf. Theory, vol. 52, no. 8, pp. 3650–3661, Aug. 2006.

[67] C. Shangguan and I. Tamo, “Combinatorial list-decoding of Reed–Solomon codes beyond the Johnson radius,” in Proc. 52nd Annu. ACM SIGACT Symp. Theory Comput., Jun. 2020, pp. 538–551.

[68] R. Silverman, “A metrization for power-sets with applications to combinatorial analysis,” Can. J. Math., vol. 12, pp. 158–176, Mar. 1960.

[69] R. Singleton, “Maximum distance q-ary codes,” IEEE Trans. Inf. Theory, vol. IT-10, pp. 116–118, 1964.

[70] M. Sudan, “Decoding of Reed Solomon codes beyond the error-correction bound,” J. Complex., vol. 13, no. 1, pp. 180–193, Mar. 1997.

[71] X. Wang, D. Zheng, L. Hu, and X. Zeng, “The weight distributions of two classes of binary cyclic codes,” Finite Fields Appl., vol. 34, pp. 192–207, Jul. 2015.

[72] X. Wang, D. Zheng, and C. Ding, “Some punctured codes of several families of binary linear codes,” IEEE Trans. Inf. Theory, vol. 67, no. 8, pp. 5133–5148, Aug. 2021.

[73] M. Xiong, N. Li, Z. Zhou, and C. Ding, “Weight distribution of cyclic codes with arbitrary number of generalized Niho type zeroes,” Des., Codes Cryptogr., vol. 78, no. 3, pp. 713–730, Mar. 2016.

[74] J. Yang, M. Xiong, C. Ding, and J. Luo, “Weight distribution of a class of cyclic codes with arbitrary number of zeros,” IEEE Trans. Inf. Theory, vol. 59, no. 9, pp. 5985–5993, Sep. 2013.

Hao Chen received the Ph.D. degree in mathematics from the Institute of Mathematics, Fudan University, in 1991. He is currently a Professor with the College of Information Science and Technology/Cyber Security, Jinan University. His research interests include coding theory and cryptography, quantum information and computation, lattices, and algebraic geometry.

Chengju Li received the Ph.D. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2014. From March 2015 to February 2016, he was a Post-Doctoral Researcher with the Department of Mathematics, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. From March 2016 to August 2016, he held a post-doctoral position with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong. From July 2018 to September 2018, he visited the University of Paris VIII, Paris, France. He is currently a Professor with the Software Engineering Institute, East China Normal University, Shanghai, China. His research interests include exponential sums, coding theory, and cryptography.

Longjiang Qu received the B.A. degree in 2002 and the Ph.D. degree in mathematics from the National University of Defense Technology, Changsha, China, in 2007. He is currently a Professor with the College of Science, National University of Defense Technology. His research interests include cryptography and coding theory.

Shanxiang Lyu (Member, IEEE) received the B.S. and M.S. degrees in electronic and information engineering from the South China University of Technology, Guangzhou, China, in 2011 and 2014, respectively, and the Ph.D. degree from the Electrical and Electronic Engineering Department, Imperial College London, U.K., in 2018. He is currently an Associate Professor with the College of Cyber Security, Jinan University, Guangzhou. His research interests include lattice codes and information security. He was a recipient of the 2021 CIE Information Theory Society Yong-Star Award and the 2020 Superstar Supervisor Award of the National Crypto-Math Challenge of China. He was in the Organizing Committee of Inscripy 2020.

Liqing Xu received the master’s degree in mathematics from the Institute of Mathematics, Fudan University, in 1991. She is currently an Associate Professor with the College of Information Science and Technology/Cyber Security, Jinan University. Her research interests include coding and cryptography.

Mingshuo Zhou received the Ph.D. degree in mathematics from the Institute of Mathematics, Chinese Academy of Sciences, in 2014. He is currently an Associate Professor with the Center of Applied Mathematics, School of Mathematics, Tianjin University. His research interests include coding and algebraic geometry.