Classification of Linear Codes by Extending Their Residuals

Stefka Bouyuklieva[1](✉) and Iliya Bouyukliev[2]

[1] St. Cyril and St. Methodius University of Veliko Tarnovo, Veliko Tarnovo, Bulgaria
stefka@ts.uni-vt.bg

[2] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
P.O. Box 323, Veliko Tarnovo, Bulgaria
iliyab@math.bas.bg

Abstract. An approach for classification of linear codes with given parameters starting from their proper residual codes or subcodes is presented. The base of the algorithm is the concept of canonical augmentation which is important for parallel implementations. The algorithms are implemented in the programs LengthExtension and DimExtension of the package QextNewEdition. As an application, the nonexistence of binary $[41, 14, 14]$ codes is proved.

Keywords: Linear code · Classification · Residual code

1 Introduction

The paper is a contribution to the problem of classifying linear codes with given parameters over finite fields with q elements. Many authors have considered this problem before [2,3,5,10], and it is known to be very hard. The structure of the codes for classification is very important in the generation process. We discuss an algorithm that solves the following problem: Find all inequivalent codes with given parameters if the set of all residual codes with respect to a codeword with a given weight is given. The extension of the generator matrix of a given residual code can be done row by row or column by column. We consider in more details the problem how to generate only inequivalent codes and obtain all of needed codes. To do this, we use the concept of canonical augmentation [10,12]. This concept is very important for parallel implementations. We also mention the dual problem namely the classification of linear codes by extending their proper subcodes.

The algorithms presented in this paper are implemented in the programs LengthExtension and DimExtension of the package QextNewEdition.

Supported by Grant DN 02/2/13.12.2016 of the Bulgarian National Science Fund.
I. Bouyukliev—The research of the second author was supported, in part, by the Bulgarian Ministry of Education and Science by Grant No. DO1-221/03.12.2018 for NCHDC, a part of the Bulgarian National Roadmap on RIs.

A. M. Bigatti et al. (Eds.): ICMS 2020, LNCS 12097, pp. 173–180, 2020.
https://doi.org/10.1007/978-3-030-52200-1_17

Restrictions on the dual distance, minimum distance, etc. can be applied. The program will be available on the webpage
http://www.moi.math.bas.bg/moiuser/~data/Software/QextNewEdition

2 Preliminaries

Let q be a prime power and $\mathbb{F}_q$ the finite field with q elements, $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. A linear code of length n, dimension k, and minimum distance d over $\mathbb{F}_q$ is called an $[n,k,d]_q$ code. Two linear codes of the same length and dimension are equivalent if one can be obtained from the other by a sequence of the following transformations: (1) a permutation of the coordinate positions of all codewords; (2) a multiplication of a coordinate of all codewords with a nonzero element from $\mathbb{F}_q$; (3) a field automorphism. A sequence of the transformations given above that maps a code C to itself is called an automorphism of C. The set of all automorphisms of C forms a group, called the automophism group of the code and denoted by $\mathrm{Aut}(C)$. The action of $\mathrm{Aut}(C)$ on the code partitions the set of its codewords into orbits.

The defined equivalence relation in the set of all linear $[n,k,d]_q$ codes partitions this set into equivalence classes. We choose a canonical representative of each equivalence class. If C is a linear $[n,k,d]_q$ code, we call the canonical representative of its equivalence class the canonical form of C and denote it by $\rho(C)$. If two codes C_1 and C_2 are equivalent they have the same canonical form, or $\rho(C_1) = \rho(C_2)$.

Let C be an $[n,k,d]_q$ code and let c be a codeword of weight w. Then the residual code of C with respect to c, denoted $Res(C;c)$, is the code of length $n-w$ punctured on the set of coordinates on which c is nonzero. If only the weight w of c is of importance, we will denote it by $Res(C;w)$. The next result gives a lower bound for the minimum distance of residual codes.

Theorem 1. [8] *Let C be an $[n,k,d]$ code over $\mathbb{F}_q$ and let c be a codeword of weight $w < qd/(q-1)$. Then $Res(C;c)$ is an $[n-w, k-1, d']$ code, where $d' \geq d - w + \lceil w/q \rceil$.*

We need also the following theorem

Theorem 2. *Let C be an $[n,k,d]$ code over $\mathbb{F}_q$ and $x, y \in C$ be codewords of the same weight $w < qd/(q-1)$ such that $y = \phi(x)$ for an automorphism $\phi \in Aut(C)$. Then the residual codes $Res(C;x)$ and $Res(C;y)$ are equivalent.*

Proof. Let $\phi = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)\pi$, where $\gamma_i \in \mathbb{F}_q^*$, $\pi \in S_n$. Then for any $v = (v_1, v_2, \ldots, v_n) \in C$ we have

$$\phi(v) = (\gamma_1 v_1, \ldots, \gamma_n v_n)\pi = (\gamma_{1\pi^{-1}} v_{1\pi^{-1}}, \ldots, \gamma_{n\pi^{-1}} v_{n\pi^{-1}}) \in C.$$

Without loss of generality we can take $x = (00\cdots 0\underbrace{11\cdots 1}_{w})$. Then the support of $y = \phi(x)$ will be $\{(n-w+1)\pi^{-1}, \ldots, n\pi^{-1}\}$. If v is a codeword in C

then $(v_1, \ldots, v_{n-w}) \in Res(C; x)$ and $(\gamma_{1\pi^{-1}} v_{1\pi^{-1}}, \ldots, \gamma_{(n-w)\pi^{-1}} v_{(n-w)\pi^{-1}}) \in Res(C; y)$. Hence the restriction of ϕ on the first $n - w$ coordinates maps $Res(C; x)$ to $Res(C; y)$.

To see the connection to the dual code, we use a theorem that gives the relation between a punctured of a code C and a shortened of its dual code $C^\perp$. A code C can be punctured on a coordinate set T of size t. We denote the resulting code by C^T. Consider the set $C(T)$ of codewords whose i-th coordinate is 0 if $i \in T$. $C(T)$ is a subcode of C. Shortening $C(T)$ on T gives a code of length $n - t$ called shortened code of C on T and denoted by C_T. If we take T to be the support of the codeword $c \in C$ of weight w, then C^T is the residual code of $Res(C; c)$ with respect to c.

Theorem 3 ([9, Theorem 1.5.7]). *Let C be an $[n, k, d]$ code and T be a set of t coordinates. Then:*

(i) $(C^\perp)_T = (C^T)^\perp$ *and* $(C^\perp)^T = (C_T)^\perp$*;*
(ii) if $t < d$, then C^T and $(C^\perp)_T$ have dimensions k and $n-t-k$, respectively;
(iii) if $t = d$ and T is the set of coordinates where a minimum weight codeword is nonzero, then C^T and $(C^\perp)_T$ have dimensions $k - 1$ and $n - d - k + 1$, respectively.

As a corollary we obtain

Corollary 1. *Let C be an $[n, k, d]$ code over $\mathbb{F}_q$ with dual distance $d^\perp$ and let c be a codeword of weight $w < qd/(q-1)$. If T is the support of c then $Res(C; c) = C^T$ is a linear $[n - w, k - 1, d']$ code and $Res(C; c)^\perp = (C^\perp)_T$ is a linear $[n - w, n - w - k + 1, \geq d^\perp]$ code.*

Since $Res(C; c)^\perp$ is a shortened code of $C^\perp$, its minimum distance is at least $d^\perp$. Therefore we consider all $[n - w, k - 1, d' \geq d - w + \lceil w/q \rceil]_q$ codes with dual distance $\geq d^\perp$ as residual codes and then extend them to the linear $[n, k, d]_q$ codes with dual distance $\geq d^\perp$.

We developed a second algorithm which extends all possible $[n-w, k-w+1, \geq d]$ shortened codes to the $[n, k, d]$ codes provided that their dual codes contain codewords of weight w, $w < qd^\perp/(q - 1)$. The theoretical base of this algorithm is the following corollary.

Corollary 2. *If C is a linear $[n, k, d]_q$ code whose dual code $C^\perp$ contains a codeword of weight w, $w < qd^\perp/(q - 1)$, then C has a shortened code with parameters $[n - w, k - w + 1, \geq d]_q$ and dual distance $d' \geq d^\perp - w + \lceil w/q \rceil$.*

Proof. Let $x \in C^\perp$ be a vector of weight w. According to Theorem 2, its residual code $Res(C^\perp; x)$ has parameters $[n - w, n - k - 1, d']$ where $d' \geq d^\perp - w + \lceil w/q \rceil$. Then $Res(C^\perp; x)^\perp$ is a shortened code of C with parameters $[n-w, k-w+1, \geq d]$ (see Theorem 3 and Corollary 1).

Corollary 3. *Let C be a linear $[n, k, d]_q$ code with dual distance $d^\perp$. If no linear $[n - i, k - i + 1, \geq d]_q$ codes exist for $1 \leq i \leq w - 1$ then $d^\perp \geq w$.*

Proof. Suppose that $d^{\perp} = i < w$ and $x \in C^{\perp}$ is a vector of weight $d^{\perp}$. Then $Res(C^{\perp}; x)^{\perp}$ is a shortened code of C with parameters $[n-i, k-i+1, \geq d]_q$ which is not possible. Hence $d^{\perp} \geq w$.

3 The Construction

We are looking for all inequivalent linear codes with length n, dimension k, minimum distance d and dual distance at least $d^{\perp} \geq 2$. We propose two algorithms depending on the input codes.

The input in the first algorithm is a set of all inequivalent linear $[n-w, k-1, \geq d']_q$ codes with dual distance $\geq d^{\perp}$ where $d' > d - w + \lceil w/q \rceil$. These codes are all possible residual codes of $[n, k, d]_q$ linear codes with dual distance at least $d^{\perp}$ with respect to a codeword of weight w.

Without loss of generality, we can consider the generator matrices in the form

$$\left(\begin{array}{c|c} 00\cdots0 & 11\cdots1 \\ \hline G_{res} & G_1 \end{array}\right)$$

where G_{res} is a $(k-1)\times(n-w)$ matrix that generates the residual code $Res(C; x)$, $x = (00\cdots0, 11\cdots1) \in C$, $wt(x) = w$. We construct the matrix G_1 row by row in the same way as it is in the program QEXT_L of the package Q-EXTENSION [3]. The main question is which of the constructed in this way codes to take in our set of representatives of the equivalence classes. To do this, we use canonical augmentation [10,12]. The presentation that follows differs from the original McKay's paper [12] but the idea is the same.

First, we find the canonical form and the automorphism group of the constructed $[n, k, d]$ code C. The orbits are ordered in the way described in [1] and this ordering depends on the canonical form $\rho(C)$ and the automorphism group $\mathrm{Aut}(C)$. Then we check if the vector x is in the first orbit in the set of all codewords of weight w in C. If not, we reject it (it can be obtained by another residual code), if yes we say that this code passes the parent test. Finally, we check for equivalence the codes obtained from the same residual code that have passed the parent test. A pseudocode is presented in Algorithm 1.

Theorem 4. *The set M, obtained by Algorithm 1, consists of all inequivalent $[n, k, d]_q$ codes with dual distance $\geq d^{\perp}$ that have codewords of weight w.*

Proof. We have to prove that (1) any $[n, k, d]_q$ code with the needed dual distance is equivalent to a code in the set M, and (2) the codes in M are not equivalent.

(1) Let C be an $[n, k, d]_q$ code with dual distance $\geq d^{\perp}$. The set of all codewords of weight w is partitioned into orbits under the action of $\mathrm{Aut}(C)$. These orbits are ordered depending on the canonical form $\rho(C)$ (see [1] for details). Take a codeword x in the first orbit and the residual code $Res(C; x)$. There is a code $B \cong Res(C; x)$ in the set R. If ϕ maps $Res(C; x)$ into B, we can extend the map ϕ to $\overline{\phi} : C \to C'$, $C' = \overline{\phi}(C)$. If $x' = \overline{\phi}(x)$, then

$B = Res(C', x')$ and the code C' passes the parent test (the codeword $x' \in C'$ belongs to the first orbit in the partition of the set of all codewords of weight w in C' since $\rho(C) = \rho(C')$). Hence there is a code that is equivalent to C, has a residual code in the set R and passes the parent test.

(2) If $C_1 \cong C_2$ are two codes with the needed parameters, $x_i \in C_i$, $i = 1, 2$ are vectors of weight w, and both codes pass the parent test, then their residuals $Res(C_1, x_1)$ and $Res(C_2, x_2)$ are also equivalent (see Theorem 2).

Algorithm 1: Extension of a residual code.

```
Input:  The set $R$ of all inequivalent linear $[n-w, k-1, \geq d']_q$ codes with dual
        distance at least $d^\perp$
Output: A set $M$ of all inequivalent linear $[n, k, d]_q$ codes with dual distance
        $\geq d^\perp$
begin   $M = \emptyset$
      | for all codes $B \in R$ do
      |    $M_B = \emptyset$;
      |    for all constructed codes C with a residual code $B$ do:
      |        Obtain $\rho(C)$ and $Aut(C)$;
      |        if $x \in O_1$ then $M_B = M_B \cup C$
      |    end for;
      |    Remove equivalent codes from the set $M_B$
      |    $M = M \cup M_B$;
      | end for;
end.
```

The second algorithm extends all $[n-w, k-w+1, \geq d]$ codes to the $[n, k, \geq d]$ codes with dual distance $d^\perp$ whose dual codes contain codewords of weight w. The generator matrices of the considered codes have the form

$$\left(\begin{array}{cc|c} & 1 & \\ I_{w-1} & \vdots & A \\ & 1 & \\ \hline \multicolumn{2}{c|}{O} & G_0 \end{array}\right)$$

where I_{w-1} is the identity matrix, O is the $(k-w+1) \times w$ zero matrix, A and G_0 are $(w-1) \times (n-w)$ and $(k-w+1) \times (n-w)$ matrices, respectively. We fill out the matrix A row by row in a similar way as it is done in [4]. The dual code $C^\perp$ has a generation matrix $\left(\begin{array}{c|c} 11\cdots1 & 00\cdots0 \\ \hline G_1 & G_2 \end{array}\right)$ where G_2 generates the residual code of $C^\perp$ with respect to the codewords $(11\cdots100\cdots0)$ of weight w and it is the dual code of C_0. To take only inequivalent codes, we apply Algorithm 1 to the dual codes.

4 Examples

We use the presented algorithms implemented in the programs LengthExtension and DimExtension to obtain a systematic classification of linear codes with specific properties and parameters over fields with 2, 3 and 4 elements. Besides specifying the parameters such as length (n), dimension (k) and minimum distance (d), many other constraints can be considered. We give two examples, both over the filed $\mathbb{F}_2$, but the first one uses the program LengthExtension and the second one DimExtension. All calculations have been done on 2 × Intel Xeon E5-2620 V4, 32 thread computer.

Example 1. We construct all inequivalent $[45, 8, 20]_2$ codes from their residual $[25, 7, 10]_2$ codes with respect to a codeword of minimum weight 20. Since no $[44, 8, 20]_2$ code exists, the dual distance $d^{\perp}$ must be at least 2. Using the program Generation, we obtain 188572 inequivalent $[25, 7, 10]_2$ codes. Six of these codes have dual distance 1 (these codes have a zero coordinate) and therefore we cannot use them as residual codes. The other 188566 have dual distances 2 (30522 codes), 3 (158036 codes), and 4 (only 8 codes). Considering these codes as residual codes, the program LengthExtension constructs 424208 inequivalent $[45, 8, 20]_2$ codes. The calculations took 459 min. All doubly-even $[45, 8, 20]_2$ codes are classified in [11] and their number is 424207. There is only one code (up to equivalence) with these parameters which is not doubly-even. This code has a generator matrix

$$\begin{pmatrix}
111111111111111111110000000000000000000000000\\
000000000001111111100001111111111111111000000\\
000000011110011111101110000000001111110100000\\
000111100010100011100110000111110001110010000\\
011001100100101100100110011001110110010001000\\
100110101101000101100010101110011010010000100\\
101011010010110101000101100010110110100000010\\
101010101011010110001001010110101000110000001
\end{pmatrix}$$

and weight enumerator $W(y) = 1 + 99y^{20} + 90y^{22} + 15y^{24} + 45y^{28} + 6y^{30}$. Its automorphism group is isomorphic to $(C_{15} : C_4) \times S_3$, where $C_{15} : C_4$ is the semidirect product of the cyclic groups of orders 15 and 4, and S_3 is the symmetric group (calculated by GAP Computer Algebra System [6]). The group acts transitively on the coordinates and has order 360. The code is not self-orthogonal.

The following proposition allows one to reduce the number of cases that need to be considered for an exhaustive search for a certain class of codes.

Proposition 1. *If binary linear $[n, k, 2d]$ codes exist then at least one of these codes is even.*

Proof. Let C be a binary linear $[n, k, 2d]$ code. Suppose that C contains codewords of odd weight. If C^* is the punctured code of C on the right-most coordinate then C^* is an $[n-1, k, d^*]$ code where $d^* = 2d - 1$ or $2d$. Then we extend

C^* with one coordinate by adding an overall parity check. The resulting code $\widehat{C}^*$ is even and its parameters are $[n, k, 2d]$.

Proposition 2. *Binary linear* $[41, 14, 14]$ *codes do not exist.*

Proof. According to Proposition 1, it is enough to prove the nonexistence of even codes with these parameters. Feulner proved in [5] that binary $[35, 10, 13]$ code does not exist. We prove that binary $[36, 11, 13]$ and $[37, 12, 13]$ codes do not exist. The nonexistence of codes with these parameters proves that binary linear $[36, 10, 14]$, $[37, 11, 14]$ and $[38, 12, 14]$ codes do not exist. This gives us that no linear $[41 - i, 15 - i, 14]_2$ codes exist for $1 \leq i \leq 5$. According to Corollary 3, the dual distance of a binary $[41, 14, 14]$ must be at least 6. Since no $[41, 27, \geq 7]_2$ codes exist [7], $d^{\perp} = 6$. Therefore we are looking for binary even $[41, 14, 14]$ codes with dual distance 6 and we try to construct them by extending all possible even $[35, 9, 14]_2$ codes with dual distance ≥ 3. The program Generation shows that there are exactly 209 inequivalent even $[35, 9, 14]_2$ codes with needed dual distance. Then we try to extend them using the program DimExtension. The result is '`RES 0, Elapsed time: 432m`' which means that these codes cannot be extended to $[41, 14, 14]$ codes and this result is obtained in 432 min.

Remark 1. The table of optimal codes [7] indicates that the existence of $[40, 13, 14]$ binary codes is also unknown. If a code with these parameters exists, its dual distance can be 5 or 6. If C is a $[40, 13, 14]$ binary even code with dual distance 5, it contains an even $[35, 9, 14]$ shortened code with dual distance ≥ 3. By the program DimExtension, we obtain that these codes cannot be extended to $[40, 13, 14]$ binary codes. This means that if a $[40, 13, 14]$ binary even code exists, its dual distance is 6. Then this code contains a shortened code with parameters $[34, 8, 14]$ and dual distance ≥ 3. There are 10 607 917 inequivalent $[34, 8, 14]$ codes with needed dual distance. We were not able to extend all these codes for a reasonable time and therefore we have no result for the codes with parameters $[40, 13, 14]$.

Acknowledgements. We are greatly indebted to the unknown referees for their useful suggestions.

References

1. Bouyukliev, I.: About the code equivalence. In: Shaska, T., Huffman, W., Joyner, D., Ustimenko, V. (eds.) Advances in Coding Theory and Cryptology, pp. 126–151 (2007)
2. Bouyukliev, I., Bouyuklieva, S., Kurz, S.: Computer classification of linear codes. arXiv:2002.07826 [cs.IT] (2020)
3. Bouyukliev, I., Simonis, J.: Some new results for optimal ternary linear codes. IEEE Trans. Inf. Theory **48**(4), 981–985 (2002)
4. Bouyuklieva, S., Bouyukliev, I.: Classification of the extremal formally self-dual even codes of length 30. Adv. Math. Commun. **4**(3), 433–439 (2010)
5. Feulner, T.: Classification and nonexistence results for linear codes with prescribed minimum distances. Des. Codes Cryptogr. **70**, 127–138 (2014)

6. The GAP Group: GAP - Groups, Algorithms, and Programming, Version 4.11.0 (2020). https://www.gap-system.org
7. Grassl, M.: Bounds on the minimum distance of linear codes and quantum codes. http://www.codetables.de. Accessed 10 Mar 2020
8. Hill, R., Newton, D.E.: Optimal ternary linear codes. Des. Codes Crypt. **2**, 137–157 (1992)
9. Huffman, W.C., Pless, V.: Fundamentals of Error-Correcting Codes. Cambridge University Press, Cambridge (2003)
10. Kaski, P., Östergård, P.R.: Classification Algorithms for Codes and Designs. Springer, Heidelberg (2006)
11. Kurz, S.: The $[46, 9, 20]_2$ code is unique. arXiv:1906.02621v2 (2020)
12. McKay, B.: Isomorph-free exhaustive generation. J. Algorithms **26**, 306–324 (1998)