Информационные процессы, Том 9, № 4, 2009, стр. 301–331

Low density parity check codes on bipartite graphs with Reed-Solomon constituent codes[1]

V.B. Afanassiev, A.A. Davydov, V.V. Zyablov

Institute for Information Transmission Problems of Russian Academy of Sciences (Kharkevich Institute), Bol'shoi Karetnyi per. 19, GSP-4, Moscow, 127994, Russia
Email: afanv@iitp.ru, adav@iitp.ru, zyablov@iitp.ru

Received December 3, 2009

Abstract—The following two important problems are considered in the paper: constructing a low density parity check code on a bipartite graph and rapid encoding of this code. For a given constituent code, the first problem solving is reduced to constructing and investigation of parameters of the matrix describing connections of two vertex subsets of a regular bipartite graph (biadjacency matrix). It is convenient to treat the such matrix as a support-matrix of a code word. We propose a number of constructions that essentially extend the region of accessible parameters of the such matrices including these providing graphs without 4-cycles. Biadjacency matrices of regular bipartite graphs without 4-cycles are treated also as the incidence matrices of symmetric combinatorial configurations. This contributes to understanding and solving of the first problem. The second problem solving leads to search of such support-matrix transformations that maximize the encoding speed and allow us to find non trivial complexity estimates.

1. INTRODUCTION

Binary low density parity check codes (LDPC codes) are proposed by Gallager [1]. Further advances in this area are connected with the works of Zyablov and Pinsker [2], Margulis [3], Tanner [4], Sipser and Spielman [5]. By definition, a *regular* LDPC code is given by a sparse $r \times n$ parity check matrix, every row of which contains exactly $k < n$ units and every column exactly $j < r$ units [6–15]. In the general case, the number of units in the parity check matrix is of order $O(n)$ where n is the code length. In Tanner's paper [4], a construction of LDPC codes including a short binary linear code (in further "*constituent code*") is introduced. In the last decade, interest in LDPC codes and, accordingly, the number of scientific publications essentially increased, see e.g. the works [6–24] and the references therein. The contemporary interest in these codes is conditioned on the fact that probabilistic iterative decoding algorithms correct a great part of errors beyond one-half distance bound. Moreover, the modern digital technic provides the implementation of encoding and decoding of very long LDPC codes.

In the paper [5], the construction "*expander codes*" connected with the corresponding graphs is proposed. The term "*graph codes*" is introduced in the work [14] for LDPC codes with non binary constituent codes (for example, Reed-Solomon codes, Hamming codes). A similar construction with binary constituent codes is considered in [20] and called "*bipartite graph codes*".

The "graph codes" develop some approaches of the construction "product code". A support of an $[N, K]$ product code with $N = n^2, K = k^2$, is an $n \times n$ matrix, every row and every column of which contains a word of a constituent $[n, k]$ code. If we juxtapose n matrix rows to n left vertices of a graph and n columns to n its right vertices then we obtain a bipartite graph such that every its vertex from one subset is connected by one edge with every vertex from another subset. If one

[1] This work was supported in part by the State Contract no. 02.514.11.4025 of May 1, 2007.

marks every graph edge by a code symbol then n edges connected with any graph vertex contain a word of a constituent $[n, k]$ code. The adjacency matrix of the graph considered has the form

$$\begin{pmatrix} \mathbf{0} & \mathbf{J}_n \\ \mathbf{J}_n^T & \mathbf{0} \end{pmatrix}.$$

Here and in further, $\mathbf{0}$ is the zero matrix with a convenient size, T is the sign of transposition, $\mathbf{J}_n$ is the square matrix of order n all elements of which are units. As a code word support of the product code one may treat both the graph and the matrix $\mathbf{J}_n$.

An obvious development of the approach considered is use of the following adjacency matrix of an n-regular bipartite graph with $m > n$ vertex in every subset

$$\begin{pmatrix} \mathbf{0} & \mathbf{M}_{m,n} \\ \mathbf{M}_{m,n}^T & \mathbf{0} \end{pmatrix} \tag{1.1}$$

where $\mathbf{M}_{m,n}$ is an $m \times m$ matrix describing connections of two vertex subsets of the graph (biadjacency matrix). If the graph does not contain multiple edges then n units and $m - n$ zeroes are written in every row and every column of the matrix $\mathbf{M}_{m,n}$. The matrix $\mathbf{M}_{m,n}$ can be used as a support of a code word of an $[mn, K]$ LDPC code with a constituent $[n, k]$ code. At that, units of the support-matrix are changed by code symbols so that every row and every column of the matrix contains a code word of the corresponding constituent code. The graph also can be treated as a code word support.

A generalization for the case with distinct constituent codes in rows and columns is obvious. It will be considered in Section 5.

By constructing manners, LDPC codes can be partitioned to *random-like* codes [1, 6, 23] and codes *structured* on the base of algebraic and combinatoric methods [6–19, 24]. In this paper the both approaches are considered.

For implementation of LDPC codes, probabilistic iterative decoding algorithms are used usually. The efficiency of these algorithms grows if girth (the length of the minimal cycle) of the code support-graph increases. Many works are devoted to the problem of removal of short cycles, see e.g. the papers [6–17, 24] and the references there. Use of the incidence matrices of 2-designs (including Steiner systems) and the incidence matrices of projective and Euclidian planes and spaces turned out effective. For introduction in these combinatorics areas, see [25, Chapter II], [26, 27].

In a bipartite graph without multiple edges, the cycles have an even length. The support-matrix $\mathbf{M}_{m,n}$ provides a support-graph without 4-cycles if and only if any pair of rows (columns) does not contain units pairs on the same positions. In other words, the matrix $\mathbf{M}_{m,n}$ has no submatrices $\mathbf{J}_2$, where $\mathbf{J}_2$ is the 2×2 matrix all elements of which are units. In further, the such matrix $\mathbf{M}_{m,n}$ is called $\mathbf{J}_2$-*free*. It can be treated also as the incidence matrix of a *symmetric combinatorial configuration* m_n [29–36].

Definition 1. [29]

(i) A *(combinatorial) configuration* (m_r, b_n) is an incidence structure of m points and b lines. Every line contains n points, every point lies on r lines, and *at most one* line passes through two distinct points.
(ii) If $m = b$ and, hence, $n = r$, the configuration is called *symmetric* and is denoted by m_n.
(iii) A configuration is called *cyclic* if its incidence matrix is circulant.

Cyclic combinatorial configurations are considered in [17, 24, 32, 34, 35]. Their incidence matrices can be constructed, in particular, on the base *Golomb rulers* [32, 34, 35, 37–40].

Definition 2. [37]

(i) A *Golomb ruler* of *order* n is an ordered set of n integer $(a_1, a_2, \ldots, a_n)$ such that $0 \le a_1 < a_2 < \ldots < a_n$ and all the differences, $\{a_i - a_j \,|\, 1 \le j < i \le n\}$, are distinct. The *length* $L_{\mathrm{G}}(n)$ of the ruler is equal to $a_n - a_1$.

(ii) A Golomb ruler is an *optimal ruler* of length $L_{\mathrm{OG}}(n)$ if no shorter Golomb ruler of the same order n exists.

(iii) A Golomb ruler is called an (m, n) *modular Golomb ruler* if all the differences, $\{a_i - a_j \,|\, 1 \le i, j \le n,\ i \ne j\}$ are distinct and nonzero modulo m.

In other words, a Golomb ruler of order n is a set of n non negative integers placed as marks on a ruler so that distances between any two marks are distinct. The length of a ruler is the greatest distance between two marks. For any value $\delta \ge 0$, rulers $(a_1, a_2, \ldots, a_n)$ and $(a_1+\delta, a_2+\delta, \ldots, a_n+\delta)$ have the same properties.

Denote by $\mathbf{M}_{m,n}(a_1, a_2, \ldots, a_n)$ the circulant matrix $\mathbf{M}_{m,n}$, units on the first row of which are disposed in the columns with numbers $a_1, a_2, \ldots, a_n$.

Proposition 1. [32, Section 4] *Let* $a_1 = 1$.

(i) *Let* $(a_1, a_2, \ldots, a_n)$ *be a Golomb ruler of order* n *and length* $L_{\mathrm{G}}(n) = a_n - a_1$. *Then the circulant matrix* $\mathbf{M}_{m,n}(a_1, a_2, \ldots, a_n)$ *is* $\mathbf{J}_2$*-free for all values*

$$m \ge 2L_{\mathrm{G}}(n) + 1. \tag{1.2}$$

(ii) *A circulant matrix* $\mathbf{M}_{m,n}(a_1, a_2, \ldots, a_n)$ *is* $\mathbf{J}_2$*-free if and only if the set* $(a_1, a_2, \ldots, a_n)$ *is an* (m, n) *modular Golomb ruler of order* n.

In Proposition 1 the condition $a_1 = 1$ is written for exposition simplicity. Note that the point (ii) of this proposition is not written directly in [32, Section 4], but in fact it follows from the context of the work [32]. This point is given also in [35, Theorem 4].

In the present time, the optimal lengths $L_{\mathrm{OG}}(n)$ are known only for orders $n \le 25$ [37–40]. The proof of the optimality of a Golomb ruler of order $n \ge 20$ is an extremely hard problem. For example [40], in the framework of the project "distributed.net", 124387 researchers, executing distributed computing, participated in the proof of the equality $L_{\mathrm{OG}}(25) = 480$ announced in 2008. The corresponding ruler was obtained in 1984. In the other hand, for sufficiently great orders n, relatively short rulers are constructed and are available "online", see internet-resources [38–40] and the references therein. In [39] it is shown that

$$L_{\mathrm{OG}}(n) > n^2 - 2n\sqrt{n} + \sqrt{n} - 2 \text{ for all } n; \quad L_{\mathrm{OG}}(n) < n^2 \text{ for } n < 65000. \tag{1.3}$$

For practically interesting $n \le 150$, the lengths $L_{\mathrm{G}}(n)$ of the known Golomb rulers are of order $\sim (0.7 - 0.9)n^2$ [32, 34, 35, 37–40]. Accordingly to (1.2), it means the existence of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ with $m \ge (1.4 - 1.8)n^2$. In the other hand, for $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ it holds that

$$n^2 - n + 1 \le m, \tag{1.4}$$

where the equality is possible if and only if there exists a projective plane of order $n - 1$ [29, 32]. In the region

$$n^2 - n + 1 < m < 2L_{\mathrm{G}}(n) + 1, \tag{1.5}$$

for $n \ge 7$, the problem of the existence and constructing of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ is open, see the work [35] and the references therein. The known constructions [6–15, 17, 24, 29–36], including

the modular Golomb rulers [37] and constructions on the graph theory language [28, 41], give a relatively wide spectrum of possible parameters m, n. However, this spectrum consists of *non connected values* with sufficiently big gaps. The survey of the known parameters and constructions of symmetric combinatorial configurations (and, hence, of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$) is given in [35], see also the references in [24].

In this paper a new construction of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ is proposed. The construction is called "*Cancellation+Enlargement*" (*Construction CE*). Using modifications of the incidence matrices of Euclidian planes and spaces, Construction CE allows us to obtain $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ in *connected regions* of values m for the fixed n. This essentially increases the area of accessible parameters of the matrices.

As a rule, for the fixed parameters m, n, Construction CE allows us to obtain a matrix $\mathbf{M}_{m,n}$ by many ways that is convenient for the practice. Moreover, this increases the cardinality of the code ensemble drawing it to a random ensemble and thereby facilitating estimates production.

Decrease of the encoding complexity is a very important problem for implementation of LDPC codes. The complexity of encoding of an $[N, K = NR]$ code with the help of its parity check matrix has order $O(N^2(1-R)^2)$ for $N \to \infty$ and $R = const$. There are encoding algorithms with linear (by the code length) complexity, see e.g. the papers [22],[23]. In these works, attaining of the asymptotic linear complexity is connected either with structure restrictions on the parity check matrix or with the fact that the code distance of an $[N, K]$ LDPC code grows as $\sqrt{N}$. In the given work, a new encoding algorithm of an LDPC code is proposed. The asymptotic complexity of this algorithm has order $O((N^2(1-R)^4)/(1+R)^2)$.

The main results of the work

1. *Constructions of* $\mathbf{J}_2$*-free square matrices and symmetric combinatorial configurations*

The constructions proposed essentially extend the region of accessible parameters of $\mathbf{J}_2$-free square matrices and symmetric combinatorial configurations. The parameters, that can be obtained with the help of the constructions proposed, are described by Theorem 1. Theorem 1 combines results of Theorems 3-5 of Section 3.

Theorem 1. *Let q be a power of prime. Then the constructions of Section 3 of the given work allows us to obtain $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ (and, hence, the incidence matrices of symmetric combinatorial configurations m_n) with the following parameters m and n:*

(i)

$$m = q^2 - tq + \theta, \quad n = q - t - \Delta, \quad t = 0, 1, 2, ..., q-1, \quad \Delta = 0, 1, 2, ..., q-t-1, \\ \theta = 0, 1, 2, ..., q-t+1. \tag{1.6}$$

(ii)

$$m = q^s - tq^{s-1} + \theta, \quad n = q - t - \Delta, \quad s \geq 2, \quad t = 0, 1, 2, ..., q-1, \quad \Delta = 0, 1, 2, ..., q-t-1, \\ \theta = 0, 1, 2, ..., f_1 + D_s(n, q, t), \quad f_1 = (q-t)\left\lfloor \frac{q^{s-1}}{n-1} \right\rfloor \geq q - t, \tag{1.7}$$

where

$$D_s(n, q, t) = \sum_{i=1,\, f_i \geq n-1} \left\lfloor \frac{f_i}{n-1} \right\rfloor \geq 1, \quad f_{i>1} = f_{i-1} - \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor (n-2).$$

The parameters of (1.6) is a particular (and the most important) case of these from (1.7) when $s = 2$, $\lfloor q/(n-1) \rfloor = 1$, $f_1 = q - t$, $\lfloor (q-t)/(n-1) \rfloor = 1$, $D_s(n, q, t) = 1$. In the region (1.5), the parameters of (1.6) effectively fill wide connected areas in which there are no the known parameters, see Section 3. One can obtain the connected areas of values m changing q, t, and Δ for a fixed n.

2. *An encoding algorithm of LDPC codes on bipartite graphs with Reed-Solomon constituent codes*

A special transformation of the parity check matrix of an LDPC code on bipartite graph is proposed. For the transformation executing, so called a "*trajectory of rapid encoding*" is being formed. The trajectory includes the most part of the constituent codes. The codes included to the trajectory admit independent encoding. This decrease essentially the encoding complexity of LDPC code as a whole, because constituent codes are short. The rest of the constituent codes is encoded in common, and this part of the encoding procedure determines mainly its complexity. For an $[N = mn, K = NR]$ LDPC code, an approximated estimate of the maximal length of the trajectory of rapid encoding, i.e. the number of $[n, k]$ constituent codes included to it, has the form

$$L_{apr} \approx \frac{4NR}{n\,(1+R)} \left(1 + \frac{1}{k-1}\right), \tag{1.8}$$

see Section 4. The asymptotic encoding complexity with the help of the algorithm proposed has order

$$O\left(\frac{N^2(1-R)^4}{(1+R)^2}\right). \tag{1.9}$$

3. *An estimate of the code distance of LDPC codes on bipartite graphs with Reed-Solomon constituent codes*

For an $[N = mn, K = NR]$ LDPC code with $[n, k]$ Reed-Solomon constituent codes, the code distance estimate has the form

$$D_{apr} \approx \frac{N}{2} \frac{(1-R)^2}{1+R}, \tag{1.10}$$

see Section 4.

Some results of the given work were briefly represented [16] on XI International Symposium on Problems of Redundancy in Information and Control Systems (2007).

The work is organized as follows. In Section 2 random procedures constructing support-matrixs are proposed. In Section 3 constructing $\mathbf{J}_2$-free support-matrixs of a code on the base of finite geometries is considered. The survey of the known results, founded on the work [35], and the comparison of the new parameters with the known ones are given. In Section 4 a new encoding algorithm of LDPC codes on bipartite graphs is described. An upper estimate of the code distance of LDPC codes is given. Finally, in Section 5 a possible development of the approaches considered is noted.

2. RANDOM METHODS CONSTRUCTING THE BIPARTITE GRAPH ADJACENCY MATRIX

The adjacency matrix of an n-regular bipartite graph with m vertices in every subset has the form (1.1). In (1.1), we consider submatrices $\mathbf{M}_{m,n}$ with the following properties:

(A) Every row and every column contain n units and $m - n$ zeroes.

(B) The matrix $\mathbf{M}_{m,n}$ is $\mathbf{J}_2$-free, i.e. any pair of rows (columns) does not contain unit pairs on the same positions.

Property A is sufficient for a graph without 2-cycles, while Property B is sufficient for a graph without 4-cycles.

2.1. Constructing a random matrix $\mathbf{M}_{m,n}$

We consider constructing matrices $\mathbf{M}_{m,n}$ having only Property A . Let $m = sn$ and let n be fixed. Take an $n \times n$ matrix $\mathbf{J}_n$ filled by units. Change every unit of $\mathbf{J}_n$ by a random $s \times s$ permutation matrix, every row and every column of which contains exactly one unit. As a result, we obtain an $n \times n$ block matrix with the blocks of size $s \times s$. Every row and every column of this matrix contains n units and $m - n$ zeroes.

Example 1. s =2, $n = 4$, 2×2 permutation matrices are separated.

1	0	0	1	1	0	0	1
0	1	1	0	0	1	1	0
1	0	1	0	0	1	0	1
0	1	0	1	1	0	1	0
0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0
0	1	1	0	0	1	1	0
1	0	0	1	1	0	0	1

If only the identity $s \times s$ matrix (the identical permutation) is used then the graph obtained is an interleaving of s the complete $n \times n$ graphs. In all, there exist $s!$ distinct $s \times s$ permutation matrices. The complete set of variants of the matrices $\mathbf{M}_{m,n}$ having the Property A without extra restrictions is equal to $(s!)^{n^2}$ for any $s, n \geq 2$.

An arbitrary permutation matrix can be reduced to the diagonal (identity) form by permutations only rows or only columns. Therefore, as a *standard* form of the block matrices $\mathbf{M}_{m,n}$ one can use the matrix with the first block row and the first block column consisting only of identity matrices. So, the number of non equivalent block matrices does not exceed $(s!)^{(n-1)^2}$.

2.2. Constructing $\mathbf{J}_2$*-free matrices* $\mathbf{M}_{m,n}$

The submatrix $\mathbf{M}_{m,n}$ of the adjacency matrix (1.1) is a convenient object for estimates of conditions and probability of the existence of t-cycles. A cycle of length t is a closed way on a graph passing through t vertices. The cycles of the minimal length (graph girth) are at most interesting. Bipartite graph cycles have an even length. A graph has 4-cycle if in $\mathbf{M}_{m,n}$ there is a pair of rows and a pair of columns, on the intersection of which units are placed (in Example 1, see the intersections of the first and fifth columns with the first and sixth rows).

Consider an algorithm of random constructing and necessary and sufficient conditions of the $\mathbf{J}_2$-free matrices existence. In concept, this algorithm is close to Gallager's approach to constructing a parity check matrix of an LDPC code [1].

An algorithm constructing a $\mathbf{J}_2$-free matrix

Parameters: the block size is equal to s, the row weight is equal to n.

Goal: constructing a $\mathbf{J}_2$-free $sn \times sn$ matrix $\mathbf{M}$ with the fixed weight n of every row and every column using blocks, every of which is an $s \times s$ permutation matrix.

We denote $\mathbf{M} = [\mathbf{b}_{ij}]$ where $\mathbf{b}_{ij}$ is a permutation matrix, $i = 1, 2, ..., n$, $j = 1, 2, ..., n$.

1. In a block matrix of size $n \times n$, fill the top row $\mathbf{b}_{1j}$ and the left column $\mathbf{b}_{i1}$ by $s \times s$ blocks chosen randomly.
2. Put $i = j = 2$.
3. While $j < n$, in a block $\mathbf{b}_{ij}$, mark as banned all positions completing the submatrix $\mathbf{J}_2$ in common with all blocks filled before. If at least one position in every row and every column is empty, fill them by any convenient permutation matrix, otherwise, **Surrender**.
   Put $j = j + 1$.
4. Put $i = i + 1$, $j = 2$. While $i < n$, execute the point 3, otherwise, **Stop**.

Proposition 2. *If $s < n$ there is no any $\mathbf{J}_2$-free matrix $\mathbf{M}$ with the parameters s, n.*

Proof. If on the first step of the algorithm only the identity matrix is used (the standard form) then in the all rest of blocks the main diagonal positions are banned. For filling the block $\mathbf{b}_{22}$, any of $(s-1)!$ permutation matrices with the banned diagonal may be used. In $\mathbf{b}_{23}$, in addition to the main diagonal, the banned positions are these connected with the block $\mathbf{b}_{22}$ by conditions constructing the submatrix $\mathbf{J}_2$. So, in every column (row) of $\mathbf{b}_{23}$ two positions are banned, while free these may be filled by any permutation matrix from $(s-2)!$ permissible ones. It is easy to see that in a block $\mathbf{b}_{2j}$, the number of banns in a row (column) is equal to $(j-1)$. It may not be smaller as coincidence of the banns (i.e. *multiple banns on the same position*) means presence at least one submatrix $\mathbf{J}_2$ in the blocks filled before. From the analyze given, the **necessary condition** $s \geq n$ follows. Under this condition, in the block $\mathbf{b}_{2n}$ at least one free position in every row is present. □

We consider a sufficient condition of the existence of $\mathbf{J}_2$-free matrices.

Proposition 3. *Let $s \geq (n-1)^2 + 1$. Then a $\mathbf{J}_2$-free matrix $\mathbf{M}$ with the parameters s, n can be obtained.*

Proof. We continue infill of blocks described in the proof of Proposition 2. The situation with filling of the block $\mathbf{b}_{32}$ is similar to $\mathbf{b}_{23}$. Starting with the block $\mathbf{b}_{33}$ the accumulation of banns happens by more complicate rules. Assume that banns from distinct combinations of previous blocks do not put over each other. In other words, there are no multiple banns. Then the maximal number of banns in a row (column) of blocks $\mathbf{b}_{3j}$ does not exceed $2\,(j-1)$. So, in a block $\mathbf{b}_{ij}$ the total number of the banns does not exceed $(i-1)\,(j-1)$, and the **sufficient condition** has the form $s \geq (n-1)^2 + 1$. □

The sufficient condition obtained is relatively weak. The relations (1.2),(1.3) and distinct combinatorial methods constructing $\mathbf{J}_2$-free matrices, considered in Section 3, give better bounds.

2.3. An enlargement of matrices

The block structure of the matrix $\mathbf{M}_{m,n}$ is convenient for constructing, but it limits the set of accesible parameters by the values multiple to the component code length. Two variants of an enlargement of support-matrices will be considered below.

Enlargement 1: A square matrix $\mathbf{U}$ of the order u with the fixed number n of units in every row and column is given.

1. Enlarge the matrix $\mathbf{U}$ to the size $u + 1$ bordering it by the zero row from the bottom and the zero column from the right. Write the unit into "corner" element of the bordering.

2. Choose at random a row of **U** unused before; choose at random a unit in this row on the intersection of it with a column unused before.
3. "Clone" the chosen unit writing its "projections" to the bordering row and column; then change this unit by zero.
4. Perform Steps 2-3 by $n-1$ times and finish the process.

It is evident that starting from the $n \times n$ matrix filled by units and executing the procedure by $m-n$ times, we get a random matrix $\mathbf{M}_{m,n}$. It is simple to calculate that the total number of the matrices is not smaller than

$$n \prod_{i=0}^{m-n-1} \binom{n+i}{n-1}.$$

The following procedure aims at increase of a $\mathbf{J}_2$-free matrix size with conservation of the property "$\mathbf{J}_2$-free".

Enlargement 2: A square $\mathbf{J}_2$-free matrix $\mathbf{B}$ of the order b with the fixed number n of units in every row and column is given.

1. Enlarge the matrix $\mathbf{B}$ to size $b+1$ bordering it by the zero row from the bottom and the zero column from the right. Write the unit into "corner" element of the bordering.
2. Choose at random a row of $\mathbf{B}$ unmarked; choose at random a unit in this row on the intersection of it with a column unmarked. Interrupt the procedure, if there is no any row unmarked or the units needed are absent.
3. "Clone" the chosen unit writing its "projections" to the bordering row and column; then change this unit by zero.
4. In the bordering, mark elements that could complete new submatrices $\mathbf{J}_2$ on the next steps (i.e. mark rows and columns banned for further using).
5. Perform Steps 2-4 by $n-1$ times and finish the procedure.

Note that "Enlargement 2" procedure not necessary can be applied to an arbitrary starting matrix $\mathbf{B}$ as forced interruptions on Step 2 are possible. It is possible also that we will not able to use the procedure by the necessary number of times. In Section 3 we discuss a geometric interpretation of this procedure with conditions for its application. Note also that the both variants of the enlargement do not save the block structure of the initial matrix (if the structure was that).

3. CONSTRUCTING $\mathbf{J}_2$-FREE MATRICES $\mathbf{M}_{m,n}$ ON THE BASE OF FINITE GEOMETRIES

The matrix $\mathbf{M}_{m,n}$ is a constructional part of the adjacency matrix (1.1) of a bipartite n-regular graph with m vertices in every subset.

For constructing $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ providing graphs without 4-cycles, it is convenient to use the incidence matrices of 2-$(v,k,1)$-designs called also Steiner systems $S(2,k,v)$ [25, Chapter II]. Here, as usually, v is the total number elements of a design, k is the cardinality of the block. Every column of the such incidence matrix corresponds to a design element and every row corresponds to a block of the design. Every subset of two design elements is contained in **exactly** one block. Clearly that Property B of Section 2 holds. Moreover, for a support of this property it is enough if every subset of two elements is contained in **at most** one block. So, for constructing matrices $\mathbf{M}_{m,n}$ one may use submatrices of the incidence matrix of a 2-design (Steiner system) without some blocks. Some elements of the initial v-set (columns of the incidence matrix) may be excluded too.

The *resolvable* 2-designs are the most convenient for implementation of the such approach. The number of blocks of a 2-$(v,k,1)$-design is equal to $\frac{v(v-1)}{k(k-1)}$. Every element appears in exactly r blocks,

$r = \frac{v-1}{k-1}$. The blocks of a resolvable 2-$(v, k, 1)$-design can be partitioned into r resolution classes called also "parallel classes". Blocks of the every class do not intersect each other. Every design element is contained in exactly one block of every class [9–11],[25, Section II.7]. Every resolution class contains $\frac{v}{k}$ blocks. The incidence matrices of projective and Euclidian spaces are a fruitful source of 2-$(v, k, 1)$-designs [6, 9, 11, 14, 25–27]. In this case points of a space are elements of the v-set and space lines are blocks. The spaces pointed have the following important properties:

- two lines can intersect each other in at most one point;
- one and only one line passes through two points.

Property B (see Section 2) holds in any submatrix of the incidence matrix of a geometrical space.

Projective planes are not resolvable 2-designs. But they allow us to obtain a number useful parameters including the lower bound (1.4) on the value m for a given n. A projective plane provides equality $m = n^2 - n + 1$. Use of a projective plane in graph codes is considered in [14]. In [6, 7] a projective space, a particular case of which is a projective plane, is applied.

We consider constructing $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ on the base Euclidian planes and spaces. Euclidian planes and spaces are used for constructing parity check matrices LDPC codes in [7, 8]. Our approach distinguishes by the fact that the incidence matrix of an Euclidian geometry is a starting point for distinct constructions transforming it for needed parameters obtaining.

It should be emphasized that the basis of a field used for component codes is not connected with that of a finite geometry field. Nevertheless, we use the same symbol q as it is clear, by the context, what field is considered now.

3.1. $\mathbf{J}_2$*-free matrices* $\mathbf{M}_{m,n}$ *on the base of Euclidean plane*

Euclidian plane $EG(2, q)$ over the field $GF(q)$ contains q^2 points and $q^2 + q$ lines. It is called also an affine plane $AG(2, q)$. In the $(q^2 + q) \times q^2$ incidence matrix $\mathbf{M}_{EG}$ of Euclidian plane, every row corresponds to a line, every column corresponds to a point. A line contains q points, a point belongs to $q + 1$ lines. A point is given by coordinates (x_1, x_2), $x_i \in GF(q)$.

The incidence matrix $\mathbf{M}_{EG}$ gives the resolvable 2-$(q^2, q, 1)$-design with $q^2 + q$ blocks and $q + 1$ resolution classes. Every class contains q *parallel* lines. The first q classes are lines of equation $x_2 = wx_1 + u$ where w is constant for the given class and u runs over the whole field $GF(q)$. Once more class contains q lines $x_1 = c$. We do not use it. The incidence matrix truncated by this way is denoted $\overline{\mathbf{M}}_{EG}$. It provides the parameters $m = q^2$, $n = q$.

We represent $q^2 \times q^2$ matrix $\overline{\mathbf{M}}_{EG}$ so that it consists of q^2 permutation matrices of size $q \times q$. Columns of $\overline{\mathbf{M}}_{EG}$ (points of the plane) are placed in the lexicographical order. The points are enumerated by groups of the form (c, x_2) where c is constant for every group. Rows of the matrix (lines on the plane) are enumerated by the parallel classes. The matrix has the form (3.1) where all indexes are elements of the field $GF(q)$ with a primitive element α.

$$\overline{\mathbf{M}}_{EG} = \begin{bmatrix} \mathbf{G}_{0,0} & \mathbf{G}_{0,1} & \mathbf{G}_{0,\alpha} & \mathbf{G}_{0,\alpha^2} & \dots & \mathbf{G}_{0,\alpha^{q-2}} \\ \mathbf{G}_{1,0} & \mathbf{G}_{1,1} & \mathbf{G}_{1,\alpha} & \mathbf{G}_{1,\alpha^2} & \dots & \mathbf{G}_{1,\alpha^{q-2}} \\ \mathbf{G}_{\alpha,0} & \mathbf{G}_{\alpha,1} & \mathbf{G}_{\alpha,\alpha} & \mathbf{G}_{\alpha,\alpha^2} & \dots & \mathbf{G}_{\alpha,\alpha^{q-2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_{\alpha^{q-2},0} & \mathbf{G}_{\alpha^{q-2},1} & \mathbf{G}_{\alpha^{q-2},\alpha} & \mathbf{G}_{\alpha^{q-2},\alpha^2} & \dots & \mathbf{G}_{\alpha^{q-2},\alpha^{q-2}} \end{bmatrix}. \tag{3.1}$$

A *permutation matrix* $\mathbf{G}_{w,c}$ of size $q \times q$ corresponds to q points (c, x_2) and q lines $x_2 = wx_1 + u$. The matrices $\mathbf{G}_{0,c}$ and $\mathbf{G}_{w,0}$ are the identity matrices of order q.

Example 2. An example of the representation of the truncated Euclidian plane $\overline{\mathbf{M}}_{EG}$ for $q = 4$ is given in Table 1 where α is a primitive element of the field $GF(4)$, $\beta = \alpha^2$.

Table 1. The incidence matrix of the truncated Euclidian plane $\overline{\mathbf{M}}_{EG}$ for $q = 4$

x_1	0	0	0	0	1	1	1	1	α	α	α	α	β	β	β	β
x_2	0	1	α	β	0	1	α	β	0	1	α	β	0	1	α	β
$x_2 = 0$	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0
$x_2 = 1$	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0
$x_2 = \alpha$	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0
$x_2 = \beta$	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**
$x_2 = x_1$	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	**1**
$x_2 = x_1 + 1$	0	**1**	0	0	**1**	0	0	0	0	0	0	**1**	0	0	**1**	0
$x_2 = x_1 + \alpha$	0	0	**1**	0	0	0	0	**1**	**1**	0	0	0	0	**1**	0	0
$x_2 = x_1 + \beta$	0	0	0	**1**	0	0	**1**	0	0	**1**	0	0	**1**	0	0	0
$x_2 = \alpha x_1$	**1**	0	0	0	0	0	**1**	0	0	0	0	**1**	0	**1**	0	0
$x_2 = \alpha x_1 + 1$	0	**1**	0	0	0	0	0	**1**	0	0	**1**	0	**1**	0	0	0
$x_2 = \alpha x_1 + \alpha$	0	0	**1**	0	**1**	0	0	0	0	**1**	0	0	0	0	0	**1**
$x_2 = \alpha x_1 + \beta$	0	0	0	**1**	0	**1**	0	0	**1**	0	0	0	0	0	**1**	0
$x_2 = \beta x_1$	**1**	0	0	0	0	0	0	**1**	0	**1**	0	0	0	0	**1**	0
$x_2 = \beta x_1 + 1$	0	**1**	0	0	0	0	**1**	0	**1**	0	0	0	0	0	0	**1**
$x_2 = \beta x_1 + \alpha$	0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	**1**	0	0	0
$x_2 = \beta x_1 + \beta$	0	0	0	**1**	**1**	0	0	0	0	0	**1**	0	0	**1**	0	0

Note that for q power prime, $q^2 \times q^2$ matrices formed from q^2 permutation matrices of size $q \times q$ can be obtained by distinct methods, see e.g. [8, 15, 27] and the references therein. The form (3.1) is convenient for transformations used in our constructions. Removing units from the matrix $\overline{\mathbf{M}}_{EG}$ in (3.1) and excluding from the matrix some resolution classes in whole, one can obtain matrices $\mathbf{M}_{m,n}$ with other parameters.

We introduce operations *q-cancellation* and Δ-*cancellation.*

q-cancellation

From the matrix $\overline{\mathbf{M}}_{EG}$ in (3.1), one removes t block rows (resolution classes) and t block columns. Arbitrary block rows and columns may be removed. A matrix $\mathbf{M}_{m,n}$ is obtained with the parameters

$$m = q^2 - tq, \quad n = q - t, \quad q > t \geq 0. \tag{3.2}$$

Δ-cancellation

A zeroing binary $(q-t) \times (q-t)$ matrix $\mathbf{S}_0$ containing Δ units in every row and column is given. In the matrix resulting q-cancellation or in the initial matrix $\overline{\mathbf{M}}_{EG}$ (if $t = 0$), the square submatrices $\mathbf{G}_{w,c}$ noted by units of $\mathbf{S}_0$ are zeroed. We obtain a matrix $\mathbf{M}_{m,n}$ with the parameters

$$m = q^2 - tq, \quad n = q - t - \Delta, \quad q > t \geq 0, \quad q - t > \Delta \geq 0. \tag{3.3}$$

3.2. An enlargement of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ obtained from Euclidian plane

We give a variant of the procedure Enlargement 2 of Section 2.3.

Definition 3. Let $\mathbf{B}$ be a $\mathbf{J}_2$-free square matrix with the fixed number n of units in every row and column. In the matrix $\mathbf{B}$, we consider an aggregate $\mathcal{A}$ of $n-1$ rows without units in common and $n-1$ columns without mutual units. The $(n-1) \times (n-1)$ submatrix $\mathbf{C}(\mathcal{A})$ formed by the intersection of the rows and columns of $\mathcal{A}$ is called a *critical submatrix* of $\mathcal{A}$. The aggregate $\mathcal{A}$ is called an *enlarging aggregate* if its critical submatrix $\mathbf{C}(\mathcal{A})$ is a permutation matrix. The matrix $\mathbf{B}$ *admits an enlargement* if it contains at least one enlarging aggregate. The matrix $\mathbf{B}$ *admits ϕ enlargements* if it contains ϕ enlarging aggregates that do not intersect each other.

Enlargement 2a. Let $\mathbf{B}$ be a $\mathbf{J}_2$-free $b \times b$ matrix with the fixed number n of units in every row and column. Assume that $\mathbf{B}$ admits an enlargement.

1. Enlarge the matrix $\mathbf{B}$ to size $(b+1)\times(b+1)$ bordering it by the zero row from the bottom and the zero column from the right. Write the unit into "corner" element of the bordering.
2. In an arbitrary way take one of the enlarging aggregates, say $\mathcal{A}$.
3. 'Clone" all $n-1$ units of the critical submatrix $\mathbf{C}(\mathcal{A})$ writing their "projections" to the new row and column; then change the units cloned by zeroes.

Remark 1. We can treat rows of the matrix $\mathbf{B}$ as "lines" and its columns as "points". Then one may say that an *enlarging aggregate* contains $n-1$ *parallel lines* $\ell_1,\ell_2,\ldots,\ell_{n-1}$ (the corresponding rows have no units in common) and $n-1$ *pairwise non collinear points* $P_1,P_2,\ldots,P_{n-1}$ (the corresponding columns have no units in common). We number the lines and the points mentioned so that, before the enlargement, it holds that $P_i\in\ell_i$, $i=1,2,\ldots,n-1$; $P_i\notin\ell_j$ if $i\neq j$, $\{i,j\}\subset\{1,2,\ldots,n-1\}$.

The procedure Enlargement 2a can be interpreted as follows. The addition to $\mathbf{B}$ of a row and a column can be treated as the addition of a new line ℓ_{new} and a new point P_{new}. The "corner" unit provides that $P_{\text{new}}\in\ell_{\text{new}}$.

The "cloning" of $n-1$ units of the critical submatrix $\mathbf{C}(\mathcal{A})$ means that all points corresponding to the aggregate $\mathcal{A}$ are included into the line ℓ_{new}. Also, the point P_{new} is included into all lines of $\mathcal{A}$. In other words, after the cloning we have $\{P_1,P_2,\ldots,P_{n-1}\}\subset\ell_{\text{new}}$ and $P_{\text{new}}\in\ell_i$, $i=1,2,\ldots,n-1$.

The change of the units cloned by zeroes means that all points of $\mathcal{A}$ are removed from the lines of $\mathcal{A}$. In other words, after this zeroing we have $P_i\notin\ell_i$, $i=1,2,\ldots,n-1$.

Note also, that after the enlargement all lines $\ell_1,\ell_2,\ldots,\ell_{n-1}$ of $\mathcal{A}$ and the new line ℓ_{new} are intersecting in the new point P_{new}.

It is interestingly that in 1887, in the work [30] V. Martinetti proposed a construction of an enlargement of the incidence matrices of a symmetric configuration m_3. In this construction quoted in [33, Introduction], two parallel lines a, b and two non collinear points A_0, B_0 are chosen so that $A_0\in a$, $B_0\in b$. Then a line z and a point Z are added. The points A_0, B_0 are removed from the lines a and b and are included into the new line z. The new point Z is included into all lines a, b, and z. Therefore all the lines are intersecting now.

It is easy to see that the procedure Enlargement 2a can be treated as a generalization of Martinetti's construction to symmetric configurations m_n, $n>3$.

Theorem 2. *From a* $\mathbf{J}_2$*-free matrix* $\mathbf{M}_{m,n}$ *the procedure Enlargement 2a obtains a* $\mathbf{J}_2$*-free matrix* $\mathbf{M}_{m+1,n}$.

Proof. The number n of units in every row and column is saved as the critical $(n-1)\times(n-1)$ submatrix $\mathbf{C}(\mathcal{A})$ is a permutation matrix. Also, the "corner" unit should be taken into account. Consider the n-set consisting of all rows of the enlarging aggregate $\mathcal{A}$ and the new row. After the enlargement, every two rows of the set have exactly one common unit disposed in the new column. These two rows are not able to have the second common unit as before the enlargement the rows of $\mathcal{A}$ had no mutual units, see Definition 3. □

Remark 2. By Remark 1, the proof of Theorem 2 can be written on the geometrical language. After the enlargement, as before every line contains n points and every point lies on n lines, see the relations $P_{\text{new}}\in\ell_{\text{new}}$, $\{P_1,P_2,\ldots,P_{n-1}\}\subset\ell_{\text{new}}$, $P_{\text{new}}\in\ell_i, i=1,2,\ldots,n-1$, and $P_i\notin\ell_i$, $i=1,2,\ldots,n-1$. Moreover, exactly one line ℓ_{new} passes through any two points of the set $\{P_1,P_2,\ldots,P_{n-1},P_{\text{new}}\}$, as before the enlargement the points $\{P_1,P_2,\ldots,P_{n-1}\}$ were pairwise non collinear.

Now we show that the procedure Enlargement 2a can be applied to a $\mathbf{J}_2$-free matrix $\mathbf{M}_{m,n}$ obtained by the operations q-cancellation and Δ-cancellation.

Lemma 1. *Any $n-1$ new rows and $n-1$ new columns obtained as a result of multiple applying of the procedure Enlargement 2a form an enlarging aggregate.*

Proof. In this case the critical submatrix is the identity $(n-1)\times(n-1)$ matrix from the "corner" units. □

Denote by $\overline{\mathbf{M}}_{q\Delta}$ a matrix obtained as a result of multiple applying of q-cancellation and Δ-cancellation to the matrix $\overline{\mathbf{M}}_{EG}$. Remind that some square submatrices $\mathbf{G}_{w,c}$ of $\overline{\mathbf{M}}_{q\Delta}$ can be zeroed by Δ-cancellation. We call a submatrix $\mathbf{G}_{w,c}$ "square".

Lemma 2. **(i)** *Any aggregate of $n-1$ rows and $n-1$ columns passing through a non zeroed square $\mathbf{G}_{w,c}$ of a matrix $\overline{\mathbf{M}}_{q\Delta}$ is an enlarging aggregate.*
(ii) *Every non zeroed square $\mathbf{G}_{w,c}$ of $\overline{\mathbf{M}}_{q\Delta}$ is connected with $\lfloor q/(n-1)\rfloor$ non intersecting enlarging aggregates.*
(iii) *A matrix $\overline{\mathbf{M}}_{q\Delta}$ admits at least $(q-t)\lfloor q/(n-1)\rfloor$ enlargements.*

Proof. **(i)** Lines $x_2 = wx_1 + u$ passing through a square $\mathbf{G}_{w,c}$ are parallel as they do not contain points in common. Similarly, q points (c, x_2) associated with this square are pairwise non collinear as their columns do not have common units. Remind also that $\mathbf{G}_{w,c}$ is a $q \times q$ permutation matrix.
(ii) In a matrix $\overline{\mathbf{M}}_{q\Delta}$, the inequality $n \le q$ holds.
(iii) As enlarging aggregates one can take the aggregates connected with $q-t$ non zeroed squares $\mathbf{G}_{w,c}$ of $\overline{\mathbf{M}}_{q\Delta}$ so that the "configuration" of these squares form some binary $(q-t)\times(q-t)$ permutation matrix. □

Construction CE - "Cancellation+Enlargement"

1. Let t and Δ be such that $q > t \ge 0$ and $q - t > \Delta \ge 0$. By q-cancellation and Δ-cancellation, a matrix $\overline{\mathbf{M}}_{q\Delta}$ is formed with parameters $m = q^2 - tq$, $n = q - t - \Delta$.
2. A binary $(q-t)\times(q-t)$ permutation matrix $\mathbf{S}_t$ is given so that every its unit corresponds to a non zeroed square $\mathbf{G}_{w,c}$ of the matrix $\overline{\mathbf{M}}_{q\Delta}$.
3. One executes $(q-t)\lfloor q/(n-1)\rfloor$ enlargements. For this, the enlarging aggregates connected with $q-t$ non zeroed squares $\mathbf{G}_{w,c}$ are used. The squares are pointed by units of the matrix $\mathbf{S}_t$.
4. In accordingly to Lemma 1, new rows and columns are used iteratively to form enlarging aggregates and to do enlargements. The iterative process continues while the number of non used new rows (columns) is not smaller than $n-1$.

Remark 3. On the i-th stage of the iterative process of Step 4, $\lfloor f_i/(n-1)\rfloor$ enlargements are done where $f_1 = (q-t)\lfloor q/(n-1)\rfloor$ and $f_{i>1}$ is the number of new rows (columns) non used in enlargements on the stages $1, 2, \ldots, i-1$ of the process. *At least one enlargement is executed on Step 4* as $f_1 \ge q - t \ge n$.

Example 3. In Tables 2 and 3, an example of Construction CE is given with $q = 4$, $t = \Delta = 0$, $m = 16$, $n = 4$. The sign ◯ notes units "cloned" and changed by zeroes. In Table 2 execution of $(4-0)\lfloor 4/(4-1)\rfloor = 4$ enlargements is illustrated accordingly to Step 3 of Construction CE. The matrix $\mathbf{S}_t$ has the form

$$\mathbf{S}_t = \begin{bmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{bmatrix}.$$

In Table 3 one enlargement is done accordingly to Step 4 of Construction CE.

Table 2. Step 3 of Construction CE, $q = 4$, $t = \Delta = 0$, $m = 16$, $n = 4$

1	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	0	0	0
0	◯	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	0	**1**
0	0	◯	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	**1**
0	0	0	◯	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**
1	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0
0	**1**	0	0	◯	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0	**1**	0
0	0	**1**	0	0	0	0	◯	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0
0	0	0	**1**	0	0	◯	0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	0
1	0	0	0	0	0	**1**	0	0	0	0	**1**	0	**1**	0	0	0	0	0	0
0	**1**	0	0	0	0	0	**1**	0	0	◯	0	**1**	0	0	0	0	**1**	0	0
0	0	**1**	0	**1**	0	0	0	0	◯	0	0	0	0	0	**1**	0	**1**	0	0
0	0	0	**1**	0	**1**	0	0	◯	0	0	0	0	0	**1**	0	0	**1**	0	0
1	0	0	0	0	0	0	**1**	0	**1**	0	0	0	0	**1**	0	0	0	0	0
0	**1**	0	0	0	0	**1**	0	**1**	0	0	0	0	0	0	◯	**1**	0	0	0
0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	◯	0	0	0	**1**	0	0	0
0	0	0	**1**	**1**	0	0	0	0	0	**1**	0	0	◯	0	0	**1**	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	0	**1**	**1**	0	0	0
0	0	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0	0	0	0	**1**	0	0
0	0	0	0	**1**	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	**1**	0
0	**1**	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**

Table 3. Step 4 of Construction CE, $q = 4$, $t = \Delta = 0$, $m = 16$, $n = 4$

1	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	0	0	0	0
0	0	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	0	**1**	0
0	0	0	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	0	**1**	0
0	0	0	0	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0	0	**1**	0
1	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	0
0	**1**	0	0	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0	**1**	0	0
0	0	**1**	0	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0	0
0	0	0	**1**	0	0	0	0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	0	0
1	0	0	0	0	0	**1**	0	0	0	0	**1**	0	**1**	0	0	0	0	0	0	0
0	**1**	0	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0
0	0	**1**	0	**1**	0	0	0	0	0	0	0	0	0	0	**1**	0	**1**	0	0	0
0	0	0	**1**	0	**1**	0	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0
1	0	0	0	0	0	0	**1**	0	**1**	0	0	0	0	**1**	0	0	0	0	0	0
0	**1**	0	0	0	0	**1**	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0
0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0
0	0	0	**1**	**1**	0	0	0	0	0	**1**	0	0	0	0	0	**1**	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	0	**1**	◯	0	0	0	**1**
0	0	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0	0	0	0	◯	0	0	**1**
0	0	0	0	**1**	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	◯	0	**1**
0	**1**	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	**1**	0	**1**

Theorem 3. *By Construction CE, from the truncated incidence matrix of Euclidian plane* $\overline{\mathbf{M}}_{EG}$*, see (3.1),* $\mathbf{J}_2$*-free matrices* $\mathbf{M}_{m,n}$ *(and, hence, the incidence matrices of symmetric combinatorial configurations* m_n*) can be obtained with the following parameters* m *and* n*:*

$$m = q^2 - tq + \theta, \quad n = q - t - \Delta, \quad t = 0, 1, 2, ..., q-1, \quad \Delta = 0, 1, 2, ..., q-t-1,$$
$$\theta = 0, 1, 2, ..., q-t+1, \tag{3.4}$$

where θ *is the total number of enlargements done.*

Proof. The assertion follows from Lemmas 1 and 2 and the description of Construction CE. As $n \le q-t \le q$ it holds that $n-1 < q$ and $\lfloor q/(n-1) \rfloor \ge 1$. So, on Step 3 of the construction at least $q-t$ enlargements can be done. It gives at least $q-t$ new rows and columns. By above, $n-1 < q-t$ and $\lfloor (q-t)/(n-1) \rfloor \ge 1$. So, on Step 4 at least one enlargement can be done. It means, that at all, at least $q-t+1$ enlargements can be done. □

The following theorem describes parameters connected with the iterative process of Step 4 of Construction CE in more detail.

Theorem 4. *By Construction CE, from the truncated incidence matrix of Euclidian plane* $\overline{\mathbf{M}}_{EG}$*, see (3.1),* $\mathbf{J}_2$*-free matrices* $\mathbf{M}_{m,n}$ *(and, hence, the incidence matrices of symmetric combinatorial configurations* m_n*) can be obtained with the following parameters* m *and* n*:*

$$m = q^2 - tq + \theta, \quad n = q - t - \Delta, \quad t = 0, 1, 2, ..., q-1, \quad \Delta = 0, 1, 2, ..., q-t-1,$$
$$\theta = 0, 1, 2, ..., (q-t)\left\lfloor \frac{q}{n-1} \right\rfloor + D_2(n, q, t), \tag{3.5}$$

where θ *is the total number of enlargements done,* $D_2(n, q, t)$ *is the number of enlargements done in the iterative process of Step 4 of Construction CE,*

$$D_2(n, q, t) = \sum_{i=1,\, f_i \ge n-1} \left\lfloor \frac{f_i}{n-1} \right\rfloor \ge 1, \tag{3.6}$$

$\lfloor f_i/(n-1) \rfloor$ *is the number of enlargements done on the* i*-th stage of the iterative process,*

$$f_1 = (q-t)\left\lfloor \frac{q}{n-1} \right\rfloor \ge q - t \ge n, \tag{3.7}$$

$f_{i>1}$ *is the number of new rows (columns) non used in enlargements on the stages* $1, 2, \ldots, i-1$ *of the iterative process,*

$$f_{i>1} = f_{i-1} - \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor (n-2). \tag{3.8}$$

Proof. Similarly to Theorem 3, the assertion follows from Lemmas 1 and 2 and the description of Construction CE. Step 4 of the construction is executed iteratively by a few stages. In the last term of the θ values list and in the equality for f_1 in (3.7), the summand $(q-t)\lfloor q/(n-1) \rfloor$ follows from the point(iii) of Lemma 2. It corresponds to Step 3 of Construction CE. On the $(i-1)$-th iterative stage we use $\lfloor f_{i-1}/(n-1) \rfloor\,(n-1)$ new rows and columns forming $\lfloor f_{i-1}/(n-1) \rfloor$ enlarging aggregates. Accordingly, we add $\lfloor f_{i-1}/(n-1) \rfloor$ new rows and columns. So,

$$f_{i>1} = f_{i-1} - \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor (n-1) + \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor = f_{i-1} - \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor (n-2).$$

The iterative process finishes when $f_i < n-1$. □

It should be noted that as a result of the iterative process we obtain an uninterrupted series of m values for fixed n and t, see (3.4)-(3.8). In particular, as $q - t \ge n$ and $\lfloor q/(n-1) \rfloor \ge 1$, it holds that $(q-t)\lfloor q/(n-1) \rfloor + D_2(n, q, t) \ge q - t + 1$. The last relation is noted in Theorem 3.

3.3. The matrices $\mathbf{M}_{m,n}$ on the base of the Euclidean space.

Euclidian space $EG(s,q)$ of order $s \geq 2$ over the field $GF(q)$ contains $q^{s-1}(q^s-1)/(q-1)$ lines and q^s points [6,8,26]. It is called also an *affine* space $AG(s,q)$. In the incidence $q^{s-1}(q^s-1)/(q-1) \times q^s$ matrix $\mathbf{M}_{EG}^{(s)}$ of the Euclidian space, every row corresponds to a line and every column corresponds to a point. Every line contains q points. A point is given by an s-positional vector $(x_1, x_2, ..., x_s)$, $x_i \in GF(q)$. The incidence matrix $\mathbf{M}_{EG}^{(s)}$ gives a *resolvable* 2-$(q^s, q, 1)$-design with $q^{s-1}(q^s-1)/(q-1)$ blocks and $(q^s-1)/(q-1)$ resolution classes. Every class contains q^{s-1} *parallel* lines. A line is either an one-dimensional subspace of the space of s-positional vectors or a coset of this subspace. We use $q^{2(s-1)}$ lines with equations of the form

$$x_i = w_i x_1 + u_i, \quad i = 2, 3, ..., s, \quad w_i, u_i \in GF(q). \tag{3.9}$$

Points of the such line have the form $(x_1, w_2x_1 + u_2, w_3x_1 + u_3, ..., w_sx_s + u_s)$.

A set of q^{s-1} lines with the same vector $\bar{\mathbf{w}} = (w_2, w_3, ..., w_s)$ forms a *resolution class* called a *bundle of parallel lines* [8]. In the class with the given vector $\bar{\mathbf{w}}$, a line is defined by the vector $\bar{\mathbf{u}} = (u_2, u_3, ..., u_s)$, see (3.9). If $(u_2, u_3, ..., u_s) = (0, 0, ..., 0)$ then the line contains "the coordintates beginning", i.e. the point $(0, 0, ..., 0)$. This line is a subspace. Other lines of the class are cosets of this subspace. All used lines of the space $EG(s,q)$ form q^{s-1} resolution classes. Every line of a class intersects q lines of another class and is parallel to the rest of its lines.

Columns of the matrix $\mathbf{M}_{EG}^{(s)}$ (i.e. points of the space) are placed in the lexicographical order. Points $(x_1, x_2, ..., x_s)$ are numerated by groups of q points of the form $(c_1, c_2, ..., c_{s-1}, x_s)$ where c_i are constants and x_s runs over all the field. Similarly, the aggregate of constants $c_1, c_2, ..., c_{s-1}$ is changed as follows: the constant c_{s-1} runs over all the field, while the rest of constants is saved, and so on.

Rows of $\mathbf{M}_{EG}^{(s)}$ (i.e. lines of the space) are numerated by groups of q^{s-1} rows chosen so that the $q^{s-1} \times q^s$ submatrix corresponding to the group consists of q *permutation* $q^{s-1} \times q^{s-1}$ *matrices*. All lines in the group are parallel. The groups noted can be chosen by distinct manners [8]. We consider the situations when every group is a bundle of lines with the same vector $\bar{\mathbf{w}}$. The vectors $\bar{\mathbf{u}}$ are numerated in the lexicographical order. In this case a permutation $q^{s-1} \times q^{s-1}$ matrix can be represented by a multilevel cartesian product of permutation matrices the least of which has order $q \times q$. The such representation can be used for obtaining matrices $\mathbf{M}_{m,n}$ with distinct parameters. But in the given work we use only the fact that a row group corresponds to q permutation $q^{s-1} \times q^{s-1}$ matrices.

In order to obtain a matrix $\mathbf{M}_{m,n}$ we use q bundles of parallel lines chosen arbitrarily. The incidence matrix truncated in this manner is denoted by $\overline{\mathbf{M}}_{EG}^{(s)}$ and provides parameters $m = q^s$, $n = q$. Let $\mathbf{G}_{\bar{w},c}$ be a *permutation* $q^{s-1} \times q^{s-1}$ submatrix placed on the intersection on the bundle of parallel lines with the vector $\bar{w}$ and columns corresponding to the points $(c, x_2, ..., x_{s-1}, x_s)$. Denote by $\bar{0}$ the vector $\bar{w} = (0, 0, ..., 0)$. The squares $\mathbf{G}_{\bar{w},0}$ and $\mathbf{G}_{\bar{0},c}$ are the identity matrices of the order q^{s-1}. Let $\bar{w}_i$ be a vector $\bar{w}$ for the i-th chosen bundle of parallel lines. Put $\bar{w}_1 = \bar{0}$. The matrix $\overline{\mathbf{M}}_{EG}^{(s)}$ has the form

$$\overline{\mathbf{M}}_{EG}^{(s)} = \begin{bmatrix} \mathbf{G}_{\bar{0},0} & \mathbf{G}_{\bar{0},1} & \mathbf{G}_{\bar{0},\alpha} & \mathbf{G}_{\bar{0},\alpha^2} & ... & \mathbf{G}_{\bar{0},\alpha^{q-2}} \\ \mathbf{G}_{\bar{w}_2,0} & \mathbf{G}_{\bar{w}_2,1} & \mathbf{G}_{\bar{w}_2,\alpha} & \mathbf{G}_{\bar{w}_2,\alpha^2} & ... & \mathbf{G}_{\bar{w}_2,\alpha^{q-2}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{G}_{\bar{w}_q,0} & \mathbf{G}_{\bar{w}_q,1} & \mathbf{G}_{\bar{w}_q,\alpha} & \mathbf{G}_{\bar{w}_q,\alpha^2} & ... & \mathbf{G}_{\bar{w}_q,\alpha^{q-2}} \end{bmatrix}. \tag{3.10}$$

In concept and technically, constructions of matrices for Euclidian space are similar to those for Euclidian plane. Relations for matrix parameters and other formulas for the space can be

obtained from these for the plane by substitute q^s instead of q^2 and q^{s-1} instead of q. Therefore, in this subsection we only reformulate in the manner mentioned the corresponding constructions and results. Also, Theorem 5 is given without proof.

We introduce operations $q^{(s)}$-*cancellation* and $\Delta^{(s)}$-*cancellation.*

$q^{(s)}$-cancellation

From the matrix $\overline{\mathbf{M}}_{EG}^{(s)}$ in (3.10), one removes t block rows and t block columns. Arbitrary block rows and columns may be removed. A matrix $\mathbf{M}_{m,n}$ is obtained with parameters

$$m = q^s - tq^{s-1}, \quad n = q - t, \quad s \geq 2, \quad q > t \geq 0.$$

$\Delta^{(s)}$-cancellation

A zeroing binary $(q-t) \times (q-t)$ matrix $\mathbf{S}_0$ containing Δ units in every row and column is given. In the matrix resulting q-cancellation or in the initial matrix $\overline{\mathbf{M}}_{EG}^{(s)}$ (if $t = 0$), the square submatrices $\mathbf{G}_{\bar{w},c}$ noted by units of $\mathbf{S}_0$ are zeroed. We obtain a matrix $\mathbf{M}_{m,n}$ with parameters

$$m = q^s - tq^{s-1}, \quad n = q - t - \Delta, \quad s \geq 2, \quad q > t \geq 0, \quad q - t > \Delta \geq 0.$$

For the next construction description, note that q^{s-1} lines with the same vector $\bar{w}$ passing through a square $\mathbf{G}_{\bar{w},c}$ are parallel. Similarly, q^{s-1} points of the form $(c, x_2, ..., x_{s-1}, x_s)$ associated with this square are pairwise non collinear. A square $\mathbf{G}_{\bar{w},c}$ is a permutation matrix. So, any aggregate from $n-1$ rows and $n-1$ columns, passing through some square $\mathbf{G}_{\bar{w},c}$, is an enlarging aggregate. Evidently, $\lfloor q^{s-1}/(n-1) \rfloor$ non intersecting enlarging aggregates are connected with a one square. Therefore a matrix obtained by $q^{(s)}$-cancellation and $\Delta^{(s)}$-cancellation admits $(q-t)\lfloor q^{s-1}/(n-1) \rfloor$ enlargements. As enlarging aggregates one can take aggregates connected with $q-t$ non zeroing squares $\mathbf{G}_{\bar{w},c}$ forming a $(q-t) \times (q-t)$ permutation matrix.

Construction CE$^{(s)}$ - "Cancellation+Enlargement" in a space

1. Let t and Δ be such that $q > t \geq 0$ and $q-t > \Delta \geq 0$. By $q^{(s)}$-cancellation and $\Delta^{(s)}$-cancellation, a matrix $\overline{\mathbf{M}}_{q\Delta}^{(s)}$ is formed with parameters $m = q^s - tq^{s-1}$, $n = q - t - \Delta$, $s \geq 2$.
2. A binary $(q-t) \times (q-t)$ permutation matrix $\mathbf{S}_t$ is given so that every its unit corresponds to a non zeroed square $\mathbf{G}_{\bar{w},c}$ of the matrix $\overline{\mathbf{M}}_{q\Delta}^{(s)}$.
3. One executes $(q-t)\lfloor q^{s-1}/(n-1) \rfloor$ enlargements. For this, the enlarging aggregates connected with $q-t$ non zeroed squares $\mathbf{G}_{\bar{w},c}$ are used. The squares are pointed by units of the matrix $\mathbf{S}_t$.
4. In accordingly to Lemma 1, new rows and columns are used iteratively to form enlarging aggregates and to do enlargements. The iterative process continues while the number of non used new rows (columns) is not smaller than $n-1$.

Theorem 5. *By Construction CE$^{(s)}$, from the truncated incidence matrix of Euclidian space $\overline{\mathbf{M}}_{EG}^{(s)}$, see (3.10), $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ (and, hence, the incidence matrices of symmetric combinatorial configurations m_n) can be obtained with the following parameters m and n:*

$$m = q^s - tq^{s-1} + \theta, \quad n = q - t - \Delta, \quad s \geq 2, \quad t = 0, 1, 2, ..., q-1, \quad \Delta = 0, 1, 2, ..., q-t-1,$$
$$\theta = 0, 1, 2, ..., (q-t)\left\lfloor \frac{q^{s-1}}{n-1} \right\rfloor + D_s(n, q, t), \tag{3.11}$$

where

$$D_s(n,q,t) = \sum_{i=1,\, f_i \geq n-1} \left\lfloor \frac{f_i}{n-1} \right\rfloor \geq 1, \;\; f_1 = (q-t)\left\lfloor \frac{q^{s-1}}{n-1} \right\rfloor \geq n, \;\; f_{i>1} = f_{i-1} - \left\lfloor \frac{f_{i-1}}{n-1} \right\rfloor (n-2). \tag{3.12}$$

The notations and the proof are similar to Theorems 3 and 4.

3.4. Obtaining of connected regions of parameters

Using the relations (1.6),(1.7),(3.4)-(3.8),(3.11),(3.12), one can obtain the connected areas of values m changing q, t, and Δ for a fixed n

Example 4. We consider $n = 16$ and $n = 22$. The lengths of the optimal Golomb rulers are as follows $L_{\mathrm{OG}}(16) = 177$, $L_{\mathrm{OG}}(22) = 356$ [37]. Accordingly to (1.6) and (3.4), taking convenient q, t, and Δ, we obtain the intervals $(q^2 - tq)...(q^2 - tq + q - t + 1)$ of m values in the region $n^2 - n + 1 < m < 2L_{\mathrm{OG}}(n) + 1$, see (1.5). The intervals are written in Table 4. Some intervals slightly exceed the right bound $2L_{\mathrm{OG}}(n) + 1$. Combining the intervals of Table 4 we obtain $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ with the following parameters

$$n = 16,\ 256 \le m \le 321,\ 323 \le m \le 361;$$
$$n = 22,\ 506 \le m \le 573,\ 575 \le m \le 729.$$

Table 4. Connected areas of values m of Construction CE

n	n^2-n+1	q	t	Δ	q^2-tq	$q-t+1$	$(q^2-tq)...(q^2-tq+q-t+1)$	$2L_{\mathrm{OG}}(n)+1$
16	241	16	0	0	256	17	256...273	355
16	241	17	1	0	272	17	272...289	355
16	241	17	0	1	289	18	289...307	355
16	241	19	3	0	304	17	304...321	355
16	241	19	2	1	323	18	323...341	355
16	241	19	1	2	342	19	342...361	355
22	463	23	1	0	506	23	506...529	713
22	463	23	0	1	529	24	529...553	713
22	463	25	3	0	550	23	550...573	713
22	463	25	2	1	575	24	575...599	713
22	463	27	5	0	594	23	594...617	713
22	463	25	1	2	600	25	600...625	713
22	463	27	4	1	621	24	621...645	713
22	463	25	0	3	625	26	625...651	713
22	463	29	7	0	638	23	638...661	713
22	463	27	3	2	648	25	648...673	713
22	463	29	6	1	667	24	667...691	713
22	463	27	2	3	675	26	675...701	713
22	463	31	9	0	682	23	682...705	713
22	463	29	5	2	696	25	696...721	713
22	463	27	1	4	702	27	702...729	713
22	463	32	10	0	704	23	704...727	713

3.5. Comparison with the known parameters

A survey of the known parameters of symmetric configurations m_n (and, hence, of $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$) is given in [35]. In the paper [35] it is noted that in the works [17,24],[28, Constructions (i),(ii), Conjecture 4.4, Remark 4.5, Example 4.6],[29, 31, 32, 34, 36, 37],[41, Constructions 3.2, 3.3, 3.7, Remark 3.5, Theorem 3.8] infinite families of symmetric configurations m_n are obtained with

the following parameters m and n.

$$\begin{gathered} m = q^2 + q + 1,\ n = q + 1 - \delta,\ q + 1 > \delta \geq 0; \\ m = q^2 - 1, \quad n = q - \delta, \quad q > \delta \geq 0; \\ m = p^2 - p,\ n = p - 1 - \delta,\ p - 1 > \delta \geq 0; \\ m = q^2 - qt,\ n = q - t - \delta,\ q > t \geq 0,\ q - t > \delta \geq 0; \\ m = q^2 - (q-1)t - 1,\ n = q - t - \delta,\ q > t \geq 0,\ q - t > \delta \geq 0; \\ m = c(q^2 + q + 1),\ n = q + c - \delta,\ c = 2, 3, \ldots, q^2 - q,\ q + c > \delta \geq 0; \\ m = 2p^2,\ n = p + t - \delta,\ 0 < t \leq q + 1,\ q^2 + q + 1 \leq p,\ p + t > \delta \geq 0. \end{gathered} \tag{3.13}$$

In (3.13), p is a prime, q is a power prime. The first three families consist of cyclic configurations.

By (3.13), it holds that the known spectrum of possible parameters m, n is a set of *non connected values* with reltaively big gaps. At the same time, Construction CE proposed in the given paper allows us to obtain $\mathbf{J}_2$-free matrices $\mathbf{M}_{m,n}$ in *connected regions* of m values for a fixed n, see above.

Example 5. For comparison, in Table 5 for $n = 12, 16, 17, 22, 32$ in the region (1.5), the known m values from (3.13) and connected areas of values obtained by the given work methods are written. For $n = 32$ the length of the optimal Golomb ruler is not known in the present time. The least known length $L_{\mathrm{G}}(32) = 784$ [38] is used. In Table 5, Σ_{known} and Σ_{new} is, respectively, the total number of the known and new values of m for a fixed n.

Table 5. Comparison of known and new values of m for fixed n

n	$n^2 - n + 1$	The known values of m from (3.13) [17, 24, 28, 29, 31, 32, 34, 36, 37, 41]	the areas of m values obtained by Construction CE	$2L_{\mathrm{OG}}(n) + 1$	Σ_{known}	Σ_{new}
12	133	133, 135, 156, 168	**156 – 170**	171	4	13
16	241	252, 255, 256, 272, 273, 288, 304, 306, 307, 338, 341, 342	**256 – 321, 323 – 354**	355	12	88
17	273	273, 288, 289, 307, 323, 324, 338, 342, 360, 372, 381, 391, 396	**289 – 307, 323 – 381, 391 – 398**	399	13	75
22	463	506, 527, 528, 550, 552, 553, 594, 598, 624, 638, 644, 651, 682, 690, 704	**506 – 573, 575 – 712**	713	15	191
32	993	993, 1023, 1024, 1057, 1184, 1188, 1312, 1320, 1332, 1368, 1376, 1386, 1407, 1425, 1504, 1518, 1568	**1024 – 1057, 1184 – 1217, 1221 – 1255, 1258 – 1293, 1295 – 1568**	1569	17	398

4. ENCODING OF LDPC CODES

Denote by C an $[N, K, D]_q$ LDPC code of length $N = mn$ over the field $GF(q)$. The constituent generalized Reed-Solomon $[n, k, d]_q$ code (GRS) has $r = n - k$ parity check symbols and distance $d = r + 1$. A support of a code word of the code C is a square $m \times m$ matrix $\mathbf{M}_{m,n}$, in every row and every column of which n units and $m - n$ zeroes are written. To every unit of $\mathbf{M}_{m,n}$ we assign a number from 1 to N (in an arbitrary order) and the corresponding code symbol of the code C. Every row and every column of the support-matrix $\mathbf{M}_{m,n}$ (i.e. the support-graph vertex) is juxtaposed to an unique GRS. In other words, a code word of the code C consists of $2m$ code words of constituent codes GRS placed on all rows and columns of the support-matrix. The rate of GRS code is equal to $\rho = k/n$. The rate R of the LDPC code C is bounded as $R \geq 2\rho - 1$. The number of information symbols is equal to $K = m\,(n - 2r)$ or greater than this value if in a parity check matrix of C there are linearly dependent rows. It is known [14] that the lower bound on the

minimum distance D of the LDPC code C has the form $D \geq d\,(d\,(d-1)+1)$ and that D exceeds the lower bound if values of n and m are relatively great.

4.1. Encoding method

Let an $[N, K, D]_q$ LDPC code C be given on a random support $\mathbf{M}_{m,n}$. We introduce a parameter x and consider the following encoding procedure:

- $m - x$ rows are encoded by constituent GRS codes;
- $m - x$ columns are encoded by constituent GRS codes;
- the rest of symbols placed in $x \times x$ submatrix (in the right lower corner of $\mathbf{M}_{m,n}$) either are calculated or are given.

The meaning of the procedure proposed is as follows. The encoding process is partitioned by two stages. The first stage is "independent" encoding of L codes GRS (i.e. independent calculation of Lr check symbols). The second stage is calculation of the rest of $(2m - L)\,r$ check symbols by solving of a truncated system of linear equations. The parameter x should be chosen so that rows and columns independently encoded contain all information symbols of the code C. It means that the lower right $x \times x$ submatrix contains only check symbols. The number of codes GRS independently encoded is equal to $L = 2(m - x)$. The number of check equations non used on the first stage is equal to $2xr$. In further, we often call constituent codes as "blocks".

Proposition 4. *For a random support-matrix $\mathbf{M}_{m,n}$ an estimate of the number of codes GRS independently encoded is as follows.*

$$L \geq L_{ran} = 2m\left(1 - \frac{2r}{n}\right) = 2mR. \tag{4.1}$$

Proof. The necessary condition of the successful completion of the encoding procedure is as follows: the number of check equations non used on the first stage should be equal to the number of check symbols calculated in $x \times x$ submatrix.

Let $L = 2(m - x)$ where x is some parameter. Then accordingly to the encoding procedure, check symbols non calculated on the first its stage are placed in a square $x \times x$ submatrix $\mathbf{M}^*_{m,n}$ allocated in the right lower corner of the matrix $\mathbf{M}_{m,n}$. The value L is random as matrices $\mathbf{M}_{m,n}$ are random. Assume that units in a random matrix $\mathbf{M}_{m,n}$ are uniformly distributed. Then the average density of units in the all matrix $\mathbf{M}_{m,n}$ and in the submatrix $\mathbf{M}^*_{m,n}$ is equal to $\delta = n/m$ and the number of check symbols non calculated equals to δx^2. There is the equation $2xr = \delta x^2$ whence $x = 2r/\delta = 2rm/n$. From this we obtain an estimate on average L_{ran} of the number of codes GRS (rows and columns) for a random matrix $\mathbf{M}_{m,n}$. The assertion of (4.1) follows from the properties of average of a random value.

In the other hand, encoding of $m - x$ rows and columns needs an assignment of $k\,(m - x)$ information symbols, $2\,(m - x)^2\,\delta$ of which are assigned repeatedly. Assume that the number of the information symbols assigned is equal to the code C dimension. We obtain equation $2k\,(m - x) - (m - x)^2\,\delta = m\,(n - 2r)$ whence $x = 2rm/n$ follows again. □

From the estimate above it follows also that the both row and column of the submatrix $\mathbf{M}^*_{m,n}$ contains $2r$ units in average.

The encoding variant above does not use the fact that every code symbol of C belongs to exactly two codes GRS. This fact can be used on a preliminary stage in order to find such permutations of rows and columns of the initial support-matrix that do these rows and columns dependent of each

other to a maximal extent. Remind the following important property of GRS code: *every set of k positions of a code word is information.* This property is used below.

The initial state of all code word positions, corresponding to units of $\mathbf{M}_{m,n}$, is called **free**. For executing of an encoding procedure, either information or check symbols will be written on the free positions. After this the positions is called **occupied**. Let some permutation of rows and columns of the initial support-matrix be done so that $\pi_1\mathbf{M}_{m,n}\pi_2 \Rightarrow \mathbf{M}_{m,n}$. The number of free positions in a block GRS denote by n_i for the i-th row and $n_i^{''}$ for the i-th column of the support-matrix.

Encoding procedure:

1. Put $i = 1, K = m\,(n - 2r)\,, L = 0$.
2. On the i-th row of the matrix $\mathbf{M}_{m,n}$, fill by information symbols any $k_i = n_i - r$ of $n_i \le n$ free positions, then calculate and write values of r check positions of the corresponding GRS code. Put $K = K - k_i, L = L + 1$.
3. If $K > 0$, in the i-th column of $\mathbf{M}_{m,n}$, fill by information symbols any $k_i^{''} = n_i^{''} - r$ of $n_i^{''} \le n$ free positions, then calculate and write values of r check positions of the corresponding GRS code. Put $K = K - k_i^{''}, L = L + 1$.
4. While $K > 0$, put $i = i + 1$ and execute Steps 2-4.
5. Calculate the rest of $(2m - L)\,r$ check symbols.

In fact, the procedure described not always comes to Step 5. We need a preliminary stage the goal of which is to find convenient permutations π_1, π_2 providing both the successful completion of the encoding procedure and the maximal number of Steps 2-4.

4.2. A preliminary stage

In order to creat a rapid encoding procedure, it is necessary to solve on the preliminary stage the following two problems: to maximize L by picking up permutations and to minimize the calculation complexity on Step 5 of the basis procedure. The main idea leading to a relatively large L is to find on the preliminary stage a pair of permutations π_1, π_2 such that, for every i, a row (column) with the maximal number of positions occupied (i.e. given or calculated) is chosen. Below a non formal algorithm of permutations search is described.

1. Choose the identical permutations as the initial state of π_1, π_2.
2. Choose some row in the initial matrix $\mathbf{M}_{m,n}$. If the choice satisfies to Rule A, then it is fixed in the permutation π_1 and all units of the row chosen are remarked as occupied.
3. Choose some column in the initial matrix $\mathbf{M}_{m,n}$. If the choice satisfies to Rule A, then it is fixed in the permutation π_2 and all units of the column chosen are remarked as occupied.
4. If on some step of permutations search there are no rows or columns of $\mathbf{M}_{m,n}$ containing r or greater free positions, then the search procedure comes back to the arbitrary number of steps and continues on another branch of the conceptual tree of posibilities.
5. If on some step t the equality $K = \sum\limits_{i=1}^{t}\left(k_i + k_i^{''}\right)$ holds, then the permutation forming finishes by movement (to the left and to the top) of the rest of rows and columns containing exactly r free positions. This increases L and decreases size of the submatrix $\mathbf{M}_{m,n}^{*}$.

Rule A: *A row (column) should have the minimal number (but not smaller than r) free positions. Also, every block connected with it and non chosen yet should have at least $r + 1$ free positions.*

In further, the permutations π_1, π_2 maximizing the number of independently encoded codes GRS are called a *rapid encoding trajectory.*

Note the following facts:

- The movements (to the left and to the top) of rows and columns containing exactly r free positions do not upset the balance between the check equations non used and the check symbols non calculated.
- The equality $K = m(n - 2r)$ is a lower bound of dimension of the code C on a graph. Use of unique codes GRS in every row and column increases the probability of coincidence the real dimension with the lower bound. Also, the presence of additional information symbols will be found on the stage of solving of the linear equation system corresponding to the submatrix $\mathbf{M}^*_{m,n}$.
- If a support-graph contains minimal cycles of length ℓ, then starting from the ℓ-th block, blocks GRS with two or greater occupied positions can appear in the rapid encoding trajectory. In other words, a zone of appearance of blocks with one occupied position is bounded, mainly, by the trajectory beginning.

4.3. An estimate of the maximal trajectory length

Denote by Q_w the number of blocks GRS having w intersections with blocks preceding to them in the rapid encoding trajectory. Then, evidently

$$L = \sum\nolimits_{w=0}^{k} Q_w, \quad K = \sum\nolimits_{w=0}^{k} (k - w)Q_w.$$

Proposition 5. *For the above-stated procedure finding the rapid encoding trajectory of an* $[N, NR]$ *LDPC code on a bipartite graph with constituent* $[n, k]$ *codes GRS, the following estimate of the maximal trajectory length holds:*

$$L_{apr} \approx \frac{4NR}{n\,(1 + R)} \left(1 + \frac{1}{k - 1}\right). \tag{4.2}$$

Proof. The proof consists of two parts. In the first part, an heuristic basis of the constructing process for the rapid encoding trajectory is given. In the second one, an estimate of the trajectory length is obtained under condition that the heuristic hypothesis holds.

Part 1. Let the graph given by a support-matrix contain ℓ-cycles. Accordingly by the procedure finding the rapid encoding trajectory, the first block has zero intersections with the previous blocks, i.e. $Q_0 = 1$. The next blocks have only one intersection. Only on the ℓ-th step finding the maximal trajectory, a block with two intersections will taken. So, for the maximal trajectory it holds that $Q_1 \geq \ell - 2$. The further development of the process finding the trajectory depends on a density of cycles on the graph and of the corresponding groups rows and columns in the support-matrix.

It is easy to see that an ℓ-cycle on a graph corresponds to an aggregate of $\ell/2$ rows and $\ell/2$ columns in a support-matrix such that $\ell/2 \times \ell/2$ submatrix obtained by their intersection contains exactly two units in every row and column. The submatrix can be reduced to the circulant form with the first row $110\ldots0$. When $\ell = 4$ this submatrix is the matrix $\mathbf{J}_2$. Say that the aggregate of rows and columns noted *contains an ℓ-cycle*.

We define a *density* of ℓ-cycles as *low*, if in a support-matrix for any group of $\ell/2$ rows (or columns) there is at most one group of $\ell/2$ columns (rows) containing an ℓ-cycle. Then we define a density of ℓ-cycles as *middle*, if for any group of $\ell/2$ rows (or columns) there is at most c group of $\ell/2$ columns (rows) containing an ℓ-cycle. Here c is some constant. Finally, define a density of ℓ-cycles as *maximum*, if any group of $\ell/2$ rows and $\ell/2$ columns contains an ℓ-cycle. An evident example of an object with maximum density of any ℓ-cycles is the complete graph.

Taking into account the definitions introduced, we can assume that the $(\ell + 1)$-th block of the maximal trajectory has one intersection with the before chosen blocks, if the cycle density is low,

and two intersections, if the density is middle or maximum. Then the $(\ell+2)$-th block of the maximal trajectory can have two intersections for the low density, two or three those for the middle density, and three those for the maximum density of the cycles. To this moment, the total number of blocks intersecting the blocks of the initial part of the trajectory is approximately $(\ell+2)\,n$. Therefore, it can be assumed that blocks with the only intersection will not appear now, as a block completing a cycle will be found. One may assume also that the series of blocks with two intersections cannot be longer than m/n, as at this time the total number of blocks intersecting the chosen trajectory would be of order m and a block with three or more intersection would be found. In future, the possibility including (in the maximal trajectory) a block having the great number of intersections should increase with simultaneous depletion of blocks with the smaller number of intersections. The growth rate of the distribution of Q_w on the initial part should decrease with increase of the length of minmal cycle (graph girth).

It can be assumed also that in the region $w \approx n/2$ the distribution Q_w is similar to uniform. In the region $w \approx k$ the distribution Q_w should quickly terminate because of Rule A and the fact that blocks not included to the trajectory have $2r$ free positions in average.

The above heuristic reasoning allows us to assume permissibility of the uniform approximation of the distribution of values $Q_w = const, 2 \le w \le k-2$. We assume also a week symmetry $Q_0 = Q_k = 1,\ Q_1 = \ell - 1$.

Part 2. A $(k-2)$-set of blocks from the maximal trajectory with the number of intersections $w = 2, 3, 4, ..., k-1$ (by one block for every value of w) is called a **package**. The length of the package (i.e the number of blocks in it) is equal to $(k-2)$. The number of information symbols in one package is equal to

$$v_k = \sum_{i=1}^{k-2} i = \frac{(k-2)\,(k-1)}{2}.$$

The estimate of length of the maximal trajectory is obtained from the following equation (under the condition that the above-mentioned hypothesis holds) :

$$\frac{K - k - (\ell-1)\,(k-1)}{v_k} \approx \frac{L - (\ell-1) - 1}{k-2}, \tag{4.3}$$

where the numerator of the left part is the total number of the information symbols containing in the packages of the volume v_k; the numerator of the right part is the total number of the blocks of the maximal trajectory without blocks non included into the packages. From (4.3), by elementary transformations we obtain

$$L \approx 2\frac{K - \ell\,(k-1) - 1}{k-1} + \ell,$$

whence

$$L \approx \frac{2K}{k-1} - \frac{1}{k-1}$$

or

$$L \approx \frac{2K}{k-1}.$$

Introduce relative values $R = \frac{K}{N},\ \rho = \frac{k}{n}$. Taking into account $R = 2\rho - 1$, we transform the estimate of L to the form (4.2). □

From (4.2) in the asymptotic form for $N \to \infty$, we obtain

$$L_{apr} \approx \begin{cases} \frac{4NR}{n(1+R)} \left(1 + \frac{1}{k-1}\right), \ n = const. \\ \frac{4NR}{n(1+R)}, \ n \to \infty. \end{cases} \tag{4.4}$$

Evidently that the estimate L_{apr} of the maximal trajectory length in (4.2) is better by $2/(1+R)$ times than the estimate (4.1) of Proposition 4.

The verisimilitude of the estimate proposed is checked by computer simulation. Here and future, for the simulation, we consider $[N_m, K_r, D_r]_{16}$ LDPC codes C_r over the field $GF(16)$ with $r = 2-6$. The $[16, 16-r, r+1]_{16}$ codes GRS are taken as constituent these. The matrices $\mathbf{M}_{m,n}$ are constructed by a random manner with the parameters $m = 256, 800, 1600, 3200$, $n = 16$. The code length is $N_m = mn$. The code dimension is $K = m(n-2r)$. The simulation results are given in Table 6 where t_w is the average number of blocks in the trajectory having w intersections with blocks included into the trajectory before.

Table 6. Average values of Q_w by an ensemble of trajectories, $m = 256$

w	t_w code C_2	t_w code C_3	t_w code C_4
0	1.0	1.0	1.0
1	3.2	3.2	3.2
2	14.2	14.2	14.2
3	28.2	28.2	28.2
4	38.3	38.4	38.4
5	44.9	44.9	44.9
6	49.1	49.0	49.1
7	51.3	51.4	51.3
8	52.0	52.0	52.0
9	51.2	51.2	51.2
10	48.8	48.9	34.2
11	44.6	43.5	3.0
12	38.0	16.1	0.1
13	22.8	0.5	
14	2.9		

The comparison of estimates with the experiment can be done by Table 7, where $L_{\min}$, L_{avr} and $L_{\max}$ denote, respectively, the minimal, average, and maximal trajectory length, obtained by the simulation. The table gives also upper and low estimates of the minimum distance D of the code C described below.

From comparison with the simulation one can see that the estimate L_{apr} really gives an order of the maximal trajectory length. At that end, L_{apr} is the lower estimate for $r \leq 4$, while it is an upper estimate if $r \geq 5$. The evident reason is roughness of the uniform approximation of the distribution of Q_w increasing with decrease of a constituent code rate.

4.4. The final calculation of check symbols

We determine a correspondence between m rows and m columns and $2m$ distinct parity check matrices of constituent $[n, k, r+1]_q$ codes GRS. Every unit of the matrix $\mathbf{M}_{m,n}$ is juxtaposed to a code symbol and, hence, to a column of a parity check matrix of a code GRS. Code symbols are numerated in the order corresponding to the *rapid encoding trajectory*. A matrix $\mathbf{M}_{m,n}$ in the clear unique manner is unfolded to a rectangular binary $2m \times N$ matrix $\mathbf{P}$ that is a **skeleton** of a parity check matrix of an $[N, K, D]_q$ LDPC code C. In the matrix $\mathbf{P}$ every row contains exactly n units,

Table 7. Estimates of trajectory length and minimum distance

n =16 $r : K$	*Simulation* $L_{\min} : L_{avr} : L_{\max}$	Estimate L_{apr}	*Simulation* $D_{\max} : D_{avr} : D_{\min}$	Estimate D_{apr}
$m = 256$, $N_m = 4096$				
2:3072	499: 500: 503	473	38: 28: 21	59
3:2560	457: 459: 468	427	160: 151: 125	171
4:2048	383: 387: 399	372	480: 461: 418	349
5:1536	292: 295: 303	307		614
6:1024	195: 197: 202	228		996
$m = 800$, $N_m = 12800$				
2:9600	1564:1568:1583	1477	69: 61: 46	185
3:8000	1417:1421:1440	1333	515: 501: 450	533
4:6400	1188:1193:1213	1164	1528:1498:1431	1090
5:4800	913: 917: 929	960		1920
6:3200	608: 610: 620	711		3111
$m = 1600$, $N_m = 25600$				
2:19200	3124:3128:3157	2954		369
3:16000	2825:2833:2878	2667		1067
4:12800	2368:2373:2412	2327		2182
5: 9600	1821:1825:1851	1920		3840
6: 6400	1213:1217:1236	1422		6222
$m = 3200$, $N_m = 51200$				
2:38400	6240:6245:6292	5908		738
3:32000	5645:5651:5730	5333		2133
4:25600	4728:4735:4804	4654		4364

Also, the i-th row of $\mathbf{P}$ corresponds to the i-th block GRS in a rapid encoding trajectory. Every column of $\mathbf{P}$ contains exactly two units. This constructing method leads to the following structure of the skeleton

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 & \mathbf{0} \\ \mathbf{P}_2 & \mathbf{P}_3 \end{pmatrix}$$

where submatrices have the following properties:

- The submatrix $\mathbf{P}_1$ corresponds only blocks GRS included in the rapid encoding trajectory. It contains n units in every row and one or two units in every column.
- The submatrix $\mathbf{P}_2$ contains smaller that k units in every row and at most one unit in every column.
- The submatrix $\mathbf{P}_3$ contains exactly two units in every column and greater than r units in every row. Its size is equal to $(2m - L) \times (2m - L)$. The submatrices $\mathbf{P}_2, \mathbf{P}_3$ correspond only blocks GRS non included in the rapid encoding trajectory.

A complete parity check $2mr \times N$ matrix $\mathbf{H}$ of the code C is formed by change of every unit of the skeleton matrix $\mathbf{P}$ to the corresponding column of a parity check matrix of a code GRS. In other words, parity check matrices of $2m$ distinct codes GRS are "hanged" up the skeleton $\mathbf{P}$. The structure of $\mathbf{H}$ has the form

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_1 & 0 \\ \mathbf{H}_2 & \mathbf{H}_3 \end{pmatrix}.$$

The final calculation of check symbols consists in solving of the equation system

$$\mathbf{H}_3 \mathbf{c}^T = \mathbf{s}^T \tag{4.5}$$

where $\mathbf{H}_3$ is a $t \times t$ matrix, $\mathbf{c}$ is the vector consisting of

$$t = (2m - L)\, r \tag{4.6}$$

check symbols non calculated in blocks included to the rapid encoding trajectory, $\mathbf{s}$ is a vector depending on the matrix $\mathbf{H}_2$ and on results of encoding by trajectory, T is the sign of transposition.

Let the matrix $\mathbf{H}_3$ be non singular. The system of (4.5) solving can be represented as

$$\mathbf{c}^T = \mathbf{H}_3^{-1}\mathbf{s}^T.$$

Though the matrix $\mathbf{H}_3$ is sparse, its inverse matrix is dense. Consider a block approach to decrease of the calculation complexity [42]. The "skeleton" $\mathbf{P}_3$ contains exactly two units in every column. Also, a parity check matrix of a constituent code GRS contains the identity $r \times r$ submatrix. Therefore, the matrix $\mathbf{H}_3$ always can be transformed to the following block form

$$\mathbf{H}_3 = \begin{pmatrix} \mathbf{I}_u & \mathbf{A} \\ \mathbf{C} & \mathbf{B} \end{pmatrix} \tag{4.7}$$

where $\mathbf{I}_u$ is the identity $u \times u$ matrix, $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are sparse matrices of the sizes $u \times (t-u)$, $(t-u) \times (t-u)$ and $(t-u) \times u$, respectively. We denote $\mathbf{s} = (\mathbf{s}_0; \mathbf{s}_1)$, $\mathbf{c} = (\mathbf{c}_0; \mathbf{c}_1)$, where $\mathbf{s}_0$ and $\mathbf{c}_0$ are u dimensional vectors. Now, from (4.5),(4.7) it follows that

$$\begin{pmatrix} \mathbf{I_u} & \mathbf{0} \\ \mathbf{C} & -\mathbf{I_{t-u}} \end{pmatrix} \mathbf{H_3 c}^T = \begin{pmatrix} \mathbf{I_u} & \mathbf{A} \\ \mathbf{0} & \mathbf{CA - B} \end{pmatrix} [\mathbf{c_0}; \mathbf{c_1}]^T = \begin{pmatrix} \mathbf{I_u} & \mathbf{0} \\ \mathbf{C} & -\mathbf{I_{t-u}} \end{pmatrix} [\mathbf{s_0}; \mathbf{s_1}]^T = \begin{pmatrix} \mathbf{s}_0^T \\ \mathbf{Cs_0^T - s_1^T} \end{pmatrix}.$$

We obtain the equation system

$$\begin{cases} \mathbf{c_0^T + Ac_1^T = s_0^T} \\ \mathbf{(CA - B)c_1^T = (Cs_0^T - s_1^T)} \end{cases}.$$

The system solving has the form

$$\mathbf{c_1^T = (CA - B)^{-1}(Cs_0^T - s_1^T)}, \quad \mathbf{c_0^T = s_0^T - Ac_1^T}.$$

As the matrix $(\mathbf{CA} - \mathbf{B})$ has the size $(t-u) \times (t-u)$, the complexity of the system solving is defined by the value $(t-u)^2$. The interliving of rows and columns included to the maximal trajectory gives $u \geq t/2$. The simulations shows that by simple heuristic algorithms one can obtain $u \geq 0,7t$.

Proposition 6. *For the above-stated procedure finding the rapid encoding trajectory of an $[N, NR]$ LDPC code on a bipartite graph with constituent $[n, k]$ codes GRS, the encoding complexity is*

$$T_{enc} = T_{traj} + T_{fin}, \tag{4.8}$$

where the volume T_{traj} of calculations by the trajectory can be estimated with the help of the relation

$$T_{traj} = 2rNR\frac{3-R}{1+R}, \quad r = n - k, \tag{4.9}$$

and the complexity T_{fin} of the final checks calculation can be estimated as

$$T_{fin} \leq \frac{3}{4}N^2\frac{(1-R)^4}{(1+R)^2}. \tag{4.10}$$

Proof. The volume of calculations by the rapid encoding trajectory can be estimated as the complexity of multiplication of the information vector by the part of a parity check matrix

$$\mathbf{H}_{traj} = \begin{pmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{pmatrix}.$$

The number of columns of this matrix is equal to $K+Lr$. Every column contains at most $2r$ nonzero elements of a finite field. As a result of these calculations, code words of constituent codes included to the rapid encoding trajectory and the auxiliary syndrome $\mathbf{s}$ in equation (4.5) are calculated. The complexity of this stage has the form

$$T_{traj} = 2r\left(K + Lr\right) = 2r\left(NR + \frac{4NRr}{n\left(1+R\right)}\right) = 2rNR\left(1 + \frac{2\left(1-R\right)}{1+R}\right) = 2rNR\frac{3-R}{1+R}.$$

The complexity of the final checks calculation can be estimated as the complexity of multiplication of the vector $\mathbf{s} = (\mathbf{s}_0; \mathbf{s}_1)$ by the inverse matrix $\mathbf{H}_3^{-1}$ taking into account possible accelerations. The size of this matrix is equal to

$$u = N - K - Lr = N\left(1 - R\frac{3-R}{1+R}\right) = N\frac{(1-R)^2}{1+R}.$$

Using the accelerations proposed above, the complexity of the final stage can be estimated as

$$T_{fin} \le 3\left(\frac{u}{2}\right)^2 = \frac{3}{4}N^2\left(\frac{(1-R)^2}{1+R}\right)^2 = \frac{3}{4}N^2\frac{(1-R)^4}{(1+R)^2}.$$

□

It is easy to check that for the code rate $R > 0.75$ in a practice region code lengths (3000 symbols and more), for a constituent code with distance 3 we obtain $T_{traj} \ge T_{fin}$. For a constituent code with distance 3, the bound is moved to $R > 0.7$ for the same lengths of a LDPC code.

4.5. Estimates of the code distance

The encoding method proposed allows us to find upper estimates of minimum distance by direct constructing code words of a small weight. After permutations of rows and columns, corresponding to rapid encoding trajectory, the code word support-matrix $\mathbf{M}_{m,n}$ has the structure

$$\mathbf{M}_{m,n} = \begin{pmatrix} \mathbf{M_1}\ \mathbf{M_2} \\ \mathbf{M_3}\ \mathbf{M_4} \end{pmatrix}.$$

The submatrix $\mathbf{M}_1$ of size $(L/2) \times (L/2)$ is a support of only information symbols. Let $x = (2m - L)/2$. The $x \times x$ submatrix $\mathbf{M}_4$ (an analogue of $\mathbf{M}^*_{m,n}$ in Section 4.1) is a support of only check symbols calculated on the final encoding stage. The submatrices $\mathbf{M}_2, \mathbf{M}_3$ of size $x \times L/2$ (taking into account the transposition) contain a part of information symbols and all check these calculated by the rapid encoding trajectory. The submatrix $\mathbf{M}_1$ corresponds to the block $\mathbf{H}_1$ of a parity check matrix. The submatrices $\mathbf{M}_2, \mathbf{M}_3$ correspond to $\mathbf{H_2}$. Finally, $\mathbf{M_4}$ corresponds to $\mathbf{H}_3$.

An evident way constructing words of a small weight is as follows. Fill by zeroes the supports $\mathbf{M_1}, \mathbf{M_2}$. In the support-submatrix $\mathbf{M_3}$, choose any column with the maximal number of support-positions. Denote by z weight of the column chosen. Remind that an average value of z is equal to $2r$. Place an arbitrary word $\mathbf{v}$ of a shortened $[z, z-r, r+1]$ code GRS in the column chosen. Accordingly to the final part of the encoding procedure, a word of a shortened code GRS gives

the vector $\mathbf{s}$ and the right part of the equation (4.5). Evidently, the complete weight of the code word is equal to $wt(\mathbf{v}) + wt(\mathbf{c}_0, \mathbf{c}_1)$. An upper estimate of minimum distance can be obtained as $D \le \min\limits_{\mathbf{v}} (wt(\mathbf{v}) + wt(\mathbf{c}_0, \mathbf{c}_1))$ by running over all words of a shortened code GRS. Experimental results of the estimate of code distance $D_{\min}$, D_{avr} and $D_{\max}$ are given in Table 7.

Accordingly to the above proposed procedure finding code word of minimal weight, a simple estimate of code distance can be obtained with the help of an estimate of the maximal trajectory length. The size of the part $\mathbf{M}_4$ of the support-matrix could be considered as equal to $(2m - L_{apr})/2$. By the procedure, a code word of the minimal weight placed in the submatrix $\mathbf{M}_4$ should contain a constituent code word in every nonzero row. Similarly situation is for columns of this matrix. If we assume that all columns of the submatrix contain constituent code words of the minimal weight, then, with great probability, there is an arrangement of code symbols such that rows (maybe not all) of this submatrix contain constituent code words. From this simple argumentation, a simple estimate follows

$$D_{apr} \approx \frac{(2m - L_{apr})}{2}(r+1). \tag{4.11}$$

This estimate is given in Table 7.

For $N \to \infty$, the estimate of (4.11) has the following asymptotic form

$$D_{apr} \approx \frac{N}{2}\frac{(1-R)^2}{1+R}.$$

5. ON DEVELOPMENT OF THE APPROACHES CONSIDERED. CONCLUSION

Above we consider LDPC codes with constituent $[n, k, d]$ codes having the same parameters n, k, and d. The natural development of approaches considered is investigation of situations when these parameters are distinct.

For constituent codes with distinct parameters, the main ideas of algorithms and estimates of the encoding procedure, in particular, the rapid encoding trajectory, can be used, at whole. Of course, some modification is needed. For example, in Rule A of Section 4.2, the value $r = n - k$, equal for all codes, should be changed by the value $r_i = n_i - k_i$, where n_i, k_i are parameters of a constituent code chosen on Step 3 of the algorithm founding permutations.

The case when all constituent codes have the same length n, but the transfer rate k/n may be distinct in distinct codes (see e.g. [20]), does not need new methods for creating matrices $\mathbf{M}_{m,n}$.

Assume that $[n_1, k_1, d_1]$ constituent codes are used on rows and $[n_2, k_2, d_2]$ these are applied in columns. Then in the adjacency matrix (1.1) of a bipartite graph, the $m \times m$ submatrix $\mathbf{M}_{m,n}$ giving connections of two vertex subsets should be changed by $m_1 \times m_2$ matrix $\mathbf{M}_{m_1,m_2,n_1,n_2}$. In every row of $\mathbf{M}_{m_1,m_2,n_1,n_2}$, n_1 units and $m_1 - n_1$ zeroes are written, while in every its column n_2 units and $m_2 - n_2$ zeroes are placed. See, for example, the work [24] and the references therein. In this case, the graph becomes biregular: all vertices of the first subset have degree n_1, while all vertices of the second one have degree n_2. The adjacency matrix has the form

$$\begin{pmatrix} \mathbf{0} & \mathbf{M}_{m_1,m_2,n_1,n_2} \\ \mathbf{M}^T_{m_1,m_2,n_1,n_2} & \mathbf{0} \end{pmatrix}. \tag{5.1}$$

To avoid 4-cycles in the graph, the matrix $\mathbf{M}_{m_1,m_2,n_1,n_2}$ should be $\mathbf{J}_2$-free. It can be treated as the incidence matrix of a *non symmetric* combinatorial configuration (m_r, b_n) with parameters $m = m_2, r = n_2, b = m_1, n = n_1$, see Definition 1.

The algorithms and conditions of Sections 2.1 and 2.2 can be applied also in the case (5.1), but convenient rectangular matrices should be used. For example, in Section 2.1 one should start

with the $n_1 \times n_2$ matrix filled by units. In matrices $\mathbf{M}_{m_1,m_2,n_1,n_2}$ obtained, parameters m_1, m_2 are multiplied to order s of the permutation matrices used in the algorithms of Sections 2.1 and 2.2.

The enlargement algorithms of Section 2.3 should be modified. In the case (5.1), it is not possible to increase the order matrix by one on every step. For an enlargement of the matrix $\mathbf{M}_{m_1,m_2,n_1,n_2}$, rows and columns are added by groups so that

$$\Delta_1 = \frac{n_2}{t},\ \Delta_2 = \frac{n_1}{t}, \tag{5.2}$$

where Δ_1 and Δ_2 are, respectively, the number of rows and columns added, $t = (n_1, n_2)$ is the greatest common divisor of n_1 and n_2.

Note also, that the block matrix $\mathbf{M}_{m_1,m_2,n_1,n_2}$ of (5.1) with sizes m_1, m_2, n_1, n_2, multiplied to q and q^s, can be easy constructed from the matrices (3.1) and (3.10), doing similarly to the operations q-cancellation and q^s-cancellation of Sections 3.1 and 3.3.

The algorithm finding the rapid encoding trajectory of Section 4.2 uses the fact that in an $[n, k]$ code GRS, any k symbols can be treated as an information set. An arbitrary MDS code has the such property. Therefore, if constituent codes are codes MDS, the algorithm does not need changes. Otherwise, the corresponding additional conditions should be included into Rule A.

Finally, it should be noted that in (5.2) it is assumed that the "corner" units are not used for the enlargement. Moreover, all enlargement algorithms of Sections 2 and 3 can be modified so that the "corner" units are not filled.

In this case, for example, the enlarging aggregate $\mathcal{A}$ of Section 3 should contain n row and n columns and the corresponding critical submatrix should have size $n \times n$, cf. Definition 3. In the procedure Enlargement 2a, an unit is not written to the bordering corner element. On the geometrical language (see Remark 1) it means that the enlarging aggregate $\mathcal{A}$ contains n parallel lines $\ell_1, \ldots, \ell_n$ and n pairwise non collinear points $P_1, \ldots, P_n$. After the enlargement the following holds (as before, but with the naturally change $n-1$ by n): $\{P_1, \ldots, P_n\} \subset \ell_{\text{new}}$; $P_{\text{new}} \in \ell_i$, $i = 1, \ldots, n$; $P_i \notin \ell_i$, $i = 1, \ldots, n$; all the lines $\ell_1, \ldots, \ell_n$ of $\mathcal{A}$ intersect in the new point P_{new}. However now, because of absence of the corner unit, we have $P_{\text{new}} \notin \ell_{\text{new}}$, i.e. the new line ℓ_{new} is parallel to the lines $\ell_1, \ldots, \ell_n$. Step 4 of Construction CE is not executed as it is based on the corner units.

Elimination of the corner units does not give new parameters of $\mathbf{J}_2$-free *square* support-matrices $\mathbf{M}_{m,n}$. Moreover, the list of parameters that can be obtained is reduced (non essentially). The reason is that an enlarging aggregate can be found with slightly more efforts before of increase of its sizes. Also, Step 4 of Construction CE is not executed. In the other hand, the such approach allows us to obtain new matrices $\mathbf{M}_{m,n}$ with the same parameters m, n. It increases the ensemble of codes and the set of non equivalent combinatorial configurations that can be obtained by the methods considered.

For *non square* support-matrices of the type (5.1), elimination of the corner units is a perspective tool as it facilitates essentially constructing matrices with parameters needed, see e.g. (5.2).

In Sections 2 and 3, methods constructing support-matrices of words of an LDPC code with generalized Reed-Solomon constituent codes are considered. Conditions of the existence of such matrices are founded. Sets of possible parameters of regular structured support-matrices providing absence of 4-cycles in the corresponding bipartite graphs are obtained in Section 3. These sets extend essentially the region of possible parameters of codes. The parameters are a collection of series of consequent values. In a number of cases these series fill a region completely. However the such filling is executed not always. Therefore, the problem constructing codes with an uninterrupted parameters set is open. Methods of the given work can be used directly to create a parity check matrix of a binary or non binary LDPC code.

The encoding method of Section 4 and estimates connected with it are investigated for LDPC codes with random constructing manners. We plan to do similar researches for structured codes, including these based on finite geometries and other 2-designs.

We note a sufficient well "coincidence" of simple estimates of the rapid encoding trajectory length and of code distance with experiment results for high code rate.

The matrices considered can be treated as the incidence matrices of symmetric configurations in combinatorics. Therefore, the results obtained are useful for studying and solving of the corresponding problems. In particular, they extend essentially our knowledge on possible parameters of symmetric combinatorial configurations.

The applying area of the matrices considered is not restricted by LDPC codes constructing. For example, they are widely used in CDMA systems and in other tasks of the independent division of a common resource.

ACKNOWLEDGMENTS

The authors would like to thank M. Giulietti, S. Marcugini, and F. Pambianco for useful discussions of problems connected with symmetric combinatorial configurations and Golomb rulers.

REFERENCES

1. Gallager R.G. *Low-Density Parity-Check Codes.* Cambridge, MA: MIT Press, 1963.
2. Zyablov V.V., Pinsker M.S. Estimation of the error-correction complexity for Gallager low-density codes. *Problems of Information Transmission*, 1975, vol. 11, no. 1, pp. 23-26.
3. Margulis G.A. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problems of Information Transmission*, 1988, vol. 24, no. 1, pp. 39-46.
4. Tanner R.M. A Recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, 1981, vol. 27, no. 5, pp. 533-547.
5. Sipser M., Spielman D. Expander codes. *IEEE Transactions on Information Theory*, 1996, vol. 42, no. 6, pp. 1710-1722.
6. Kou Y., Lin S., Fossorier M.P.C. Low-density parity-check codes based on finite geometries: a rediscovery and new results. *IEEE Transactions on Information Theory*, 2001, vol. 47, no. 7, pp. 2711-2736.
7. Liva G., Song S., Lan L., Zhang Y., Lin S., Ryan W. E. Design LDPC codes: a survey and new results. *Journal of Communications Software and Systems (JCOMSS)*, 2006, vol. 2, no. 3, pp. 191-211.
8. Xu J., Chen L., Djurdjevic I., Abdel-Ghaffar K. Construction of regular and irregular LDPC codes: geometry decomposition and masking. *IEEE Transactions on Information Theory*, 2007. vol. 53, no. 1, pp. 121-134.
9. Johnson S.J., Weller S.R. Construction of low-density parity-check codes from Kirkman triple systems. *Proc. IEEE Globecom Conference*, San Antonio, TX, USA. 2001, pp. 970-974.
10. Weller S.R., Johnson S.J. Regular low-density parity-check codes from oval designs. *European Transactions on Telecommunications*, 2003, vol. 14, no. 5, pp. 399-409.
11. Johnson S.J., Weller S.R. Resolvable 2-designs for regular low-density parity-check codes. *IEEE Transactions on Communications,* 2003, vol. 51, no. 9, pp. 1413-1419.
12. Johnson S.J., Weller S.R. High-rate LDPC codes from unital designs. *Proc. IEEE Globecom 2003*, San Francisco, CA, USA. 2003.
13. Johnson S.J., Weller S.R. Codes for iterative decoding from partial geometries. *IEEE Transactions on Communications*, 2004, vol. 52, no. 2, pp. 236-243.

14. Hoholdt T., Justesen J. Graph codes with Reed-Solomon component codes. *Proc. ISIT 2006.* Seattle, USA: 2006.

15. Gabidulin E., Moinian A., Honary B., Generalized construction of quasi-cyclic regular LDPC codes based on permutation matrices. *Proc. ISIT 2006.* Seattle, USA: 2006, pp. 679-683.

16. Afanassiev V.B., Davydov A.A., Zyablov V.V. Low density concatenated codes with Reed-Solomon component codes. *Proc. XI International Symposium on Problems of Redundancy in Information and Control Systems.* St.-Petersburg, Russia: 2007, pp. 47-51. Available at http://k36.org/redundancy2007

17. Davydov A.A., Giulietti M., Marcugini S., Pambianco F. Symmetric configurations for bipartite-graph codes. *Proc. XIth International Workshop on Algebraic and Combinatorial Coding Theory, ACCT2008.* Pamporovo, Bulgaria: 2008, pp. 63-69. Available at http://www.moi.math.bas.bg/acct2008/b11.pdf

18. Lentmaier M., Zigangirov K.Sh. On generalized low-density parity-check codes based on Hamming component codes. *IEEE Communications Letters,* 1999, vol. 3, no. 8, pp. 248-260.

19. Boutros J., Pothier O., Zemor G. Generalized low density (Tanner) codes. *Proc. IEEE International Conference on Communications* (*ICC*). Vancouver, BC, Canada: 1999, vol. 1, pp. 441-445.

20. Barg A., Zemor G. Distances properties of expander codes. *IEEE Transactions on Information Theory,* 2006, vol. 52, no. 1, pp. 78-90.

21. Stiglmayr S., Zyablov V.V. Asymptotically good low-density codes based on Hamming codes. *Proc. XI International Symposium on Problems of Redundancy in Information and Control Systems.* St.-Petersburg, Russia: 2007, pp. 98-103. Available at http://k36.org/redundancy2007

22. Richardson T., Urbanke R. Efficient encoding of low-density parity-check codes. *IEEE Transactions on Information Theory*, 2001. vol. 47, no. 2, pp. 638-656.

23. Freundlich S., Burshtein D., Litsyn S. Approximately lower triangular ensembles of LDPC codes with linear encoding complexity. *IEEE Transactions on Information Theory*, 2007, vol. 53, no. 4, pp. 1484-1494.

24. Davydov A.A., Giulietti M., Marcugini S., Pambianco F. Some combinatorial aspects of constructing bipartite-graph codes. Submitted. Available at http://arxiv.org/abs/0909.5669

25. Colbourn C.J. and Dinitz J., editors. *The CRC Handbook of Combinatorial Designs.* Boca Raton, FL: CRC Press, 2006, 2-nd edition.

26. Storme L. Finite geometry. In: *The CRC Handbook of Combinatorial Designs*, 2-nd edition, C.J. Colbourn and J. Dinitz, editors, 2006, 2-nd edition, Boca Raton, FL: CRC Press, Section VII.2, pp. 702-728.

27. Kártesi F. *Introduction to finite geometries.* Budapest: Akadémiai Kiadó, 1976.

28. Abreu M., Funk M., Labbate D., Napolitano V. On (minimal) regular graphs of girth 6. *Australasian Journal of Combinatorics*, 2006, vol. 35, pp. 119-132.

29. Gropp H. Configurations. In: *The CRC Handbook of Combinatorial Designs*, 2-nd edition, C.J. Colbourn and J. Dinitz, editors, 2006, 2-nd edition, Boca Raton, FL: CRC Press, Section VI.7, pp. 353-355.

30. Martinetti V. Sulle configurazioni piane μ_3. *Annali di matematica pura ed applicata (2).* 1887-88, vol. 15, pp. 1-26.

31. Mendelsohn N.S., Padmanabhan R., Wolk B., Planar projective configurations I. *Note di Matematica*, 1987, vol. 7, pp. 91-112. Available at http://siba2.unile.it/ese/issues/1/13/Notematv7n1p91.pdf

32. Gropp H. On the existence and non-existence of configurations n_k. *Journal of Combinatorics, Information & System Sciences,* 1990, vol. 15, pp. 34-48.

33. Carstens H.G., Dinski T., Steffen E. Reduction of symmetric configurations n_3. *Discrete Applied Mathematics,* 2000, vol. 99, no. 1-3, pp. 401-411.

34. Funk M. Cyclic Difference Sets of Positive Deficiency. *Bulletin of Institute of Combinatorics and its Applications*, 2008, vol. 53, pp. 47-56.

35. Davydov A.A., Giulietti M., Marcugini S., Pambianco F. On the spectrum of possible parameters of symmetric configurations. *Proc. XII International Symposium on Problems of Redundancy in Information and Control Systems.* St.-Petersburg, Russia: 2009, pp. 59-64. Available at http://k36.org/redundancy2009

36. Funk M., Labbate D., Napolitano V. Tactical (de-)compositions of symmetric configurations. *Discrete Mathematics,* 2009, vol. 309, pp. 741-747.

37. Shearer J. Difference triangle sets. In: *The CRC Handbook of Combinatorial Designs*, C.J. Colbourn and J. Dinitz, editors, 2006, 2-nd edition, Boca Raton, FL: CRC Press, Section VI.19, pp. 436-440.

38. Shearer J. Golomb ruler tables. Available at http://www.research.ibm.com/people/s/shearer/grtab.html

39. Dimitromanolakis A. *Analysis of the Golomb Ruler and the Sidon Set problems, and determination of large, near-optimal Golomb Rulers.* Depart. Electronic Comput. Eng. Techn. University of Crete, 2002. Available at http://www.cs.toronto.edu/∼apostol/golomb/main.pdf

40. Golomb Ruler. In: *Wolfram MathWorld*. Available at http://mathworld.wolfram.com/GolombRuler.html

41. Gács A., Héger T. On geometric constructions of (k, g)-graphs. *Contributions to Discrete Mathematics*, 2008, vol. 3, pp. 63-80.

42. Gantmaher F. R. *Matrix Theory.* Moscow: Nauka, Fizmatgiz, 1988.