On the Combinatorics of Locally Repairable Codes via Matroid Theory

Thomas Westerbäck, Ragnar Freij, Toni Ernvall, and Camilla Hollanti, *Member, IEEE*

Abstract

This paper provides a link between matroid theory and locally repairable codes (LRCs) that are either linear or more generally almost affine. Using this link, new results on both LRCs and matroid theory are derived. The parameters (n, k, d, r, δ) of LRCs are generalized to matroids, and the matroid analogue of the generalized Singleton bound in [P. Gopalan *et al.*, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*] for linear LRCs is given for the parameters (n, k, d, r, δ) of a matroid. It is shown that the given bound is not tight for certain classes of parameters, implying a nonexistence result for the corresponding optimal locally repairable almost affine codes.

Constructions of classes of matroids with a large span of the parameters (n, k, d, r, δ) and the corresponding local repair sets are given. With the aid of these matroid constructions, new methods for constructing linear LRCs are provided. The constructions are based on a collection of subsets of a finite set with an associated function on these subsets, or a graph with associated functions on the vertices and edges, and finally result in several classes of optimal linear LRCs. The existence results on linear LRCs and the nonexistence results on almost affine LRCs given in this paper strengthen the nonexistence and existence results on optimal linear LRCs given in [W. Song *et al.*, "Optimal locally repairable codes," *IEEE Sel. Areas Commun.*].

T. Westerbäck and C. Hollanti are with the Department of Mathematics and Systems Analysis, Aalto University, Finland. E-mails: {thomas.westerback,camilla.hollanti}@aalto.fi.

R. Freij is with the Department of Communications and Networking, Aalto University, Finland. E-mail: ragnar.freij@aalto.fi.

T. Ernvall is with the Department of Mathematics and Statistics, University of Turku, Finland. E-mail: toni.ernvall@utu.fi.

The research of R. Freij is partially supported by the Finnish Academy of Science and Letters. The research of C. Hollanti is supported by the Academy of Finland grants #276031, #282938, and #283262, and by Magnus Ehrnrooth Foundation, Finland. The support from the European Science Foundation under the ESF COST Action IC1104 is also gratefully acknowledged.

Preliminary and partial results of this paper were presented at the 2014 IEEE Information Theory Workshop (ITW) in Hobart, Tasmania [1].

I. INTRODUCTION

Due to the ever-growing need for more efficient and scalable systems for cloud storage and data storage in general, distributed storage has become an increasingly important ingredient in many data systems. In their seminal paper [2], Dimakis *et al.* introduced network coding techniques for large-scaled distributed storage systems such as data centers, cloud storage, peer-to-peer storage systems and storage in wireless networks. These techniques can, for example, considerably improve the storage efficiency compared to traditional storage techniques such as replication and erasure coding.

Failing devices are not uncommon in large-scale distributed storage systems [3]. A central problem for this type of storage is therefore to design codes that have good distributed repair properties. Several cost metrics and related tradeoffs [2], [4], [5], [6], [7], [8] are studied in the literature, for example *repair bandwidth* [2], [4], *disk-I/O* [9], and *repair locality* [10], [11], [12]. In this paper repair locality is the subject of interest.

The notion of a *locally repairable code* (LRC) was introduced in [13], and such repair-efficient codes are already used in existing distributed storage systems, *e.g.*, in *RAID* used by Facebook and Windows Azure Storage [14].

It is well-known that nonlinear codes often achieve better performance than linear ones, *e.g.*, in the context of coding rates for error-correcting codes and maximal throughput for network codes. Almost affine codes were introduced in [15] as a generalization of linear codes. This class of codes contains codes over any alphabet size, not only of prime power size as is the case for vector codes. Many naturally occurring vector-linear codes, but not all, are also almost affine. In this paper we are especially interested in almost affine LRCs.

We will consider five key invariants (n, k, d, r, δ) of locally repairable codes. The technical definitions are given in Section II-A, but in short, a good code should have large rate k/n as well as high global and local failure tolerance d and δ, respectively. In addition, it is desirable to have small r, which will determine the maximum size of a "local" repair set. In coding-theoretic terms, d will be the minimum distance of the underlying code.

There are two notions of *symbol locality* considered in the literature, namely information locality and all-symbol locality. A linear code has *information locality* (r, δ) if all the information symbols have (r, δ)-locality. Further, a linear code has *all-symbol locality* (r, δ) if all the code symbols have (r, δ)-locality. The subject of interest in this paper is the all-symbol locality.

A. *Related work*

One of the most classical theorems in coding theory is the Singleton bound, discussed in Section II-B [16]. Its classical version upper bounds the minimum distance d of a code in terms of the length n and dimension k. Recent work sharpens the bound in terms of the local parameters (r, δ) [10], [17], as well as in terms of other parameters [13], [18], [19].

There are different constructions of LRCs that are optimal in the sense that they achieve a generalized Singleton bound, *e.g.* [14], [17], [20], [21], [22]. Song *et al.* [21] investigate for which parameters (n, k, r, δ) there exists a linear LRC with all-symbol locality and minimum distance d achieving the generalized Singleton bound from [17]. The parameter set (n, k, r, δ) is divided into eight different classes. In four of these classes it is proven that there are linear LRCs achieving the bound, in two of these classes it is proven that there are no linear LRCs achieving the bound, and the existence of linear LRCs achieving the bound in the remaining two cases is an open question.

A matroid representable by a code can in general be represented by many different codes. Further, there are matroids which can be represented by almost affine codes but not by any linear code [15]. It has been proven for linear codes that many of their properties are *matroid-invariant*, that is, only depend on their associated matroid. Examples of matroid invariants for linear codes are the dimension, the length, and the distributions of supports, weights, higher supports and higher weights [23], [24]. Other properties, such as the covering radius, are not matroid-invariant for linear codes [25]. It was shown in [14], that the r-locality of a linear LRC is a matroid invariant. This was used in [14] to prove that the distance of a class of linear LRCs achieves a generalized Singleton bound. Moreover, there are several instances of results in the theory of linear codes that have been generalized to all matroids. Examples on how these results can be interpreted for other objects that can represent a matroid, such as graphs, tranversals and certain designs can be found in [26].

Recently, the present authors have studied locally repairable codes with all-symbol locality [27]. Methods to modify already existing codes were presented and it was shown that with high probability, a certain random matrix will be a generator matrix for a locally repairable code with a good minimum distance. Constructions were given for three infinite classes of optimal vector-linear locally repairable codes over an alphabet of small size. The present paper extends and deviates from this work by studying the combinatorics of LRCs in general and relating LRCs to matroid theory. This allows for the derivation of fundamental bounds for matroids and linear and almost affine LRCs, as well as for the characterization of the matroids achieving this bound. Next, we will describe our contributions in more detail.

B. Contributions and organization

In this paper we investigate the combinatorics of LRCs. The main focus is on the connections between matroid theory and LRCs with all-symbol locality (r, δ). In particular, we are interested in locally repairable codes that are almost affine. This class contains all linear LRCs over finite fields.

Our first contribution is to extend the definition of linear LRCs with parameters (n, k, d, r, δ) given in [17] to a much bigger class of LRCs, and to show that the parameters (n, k, d, r, δ) are matroid invariant for all almost affine LRCs. We then proceed to prove our main results in this paper, which can be summarized as follows:

(i) A matroid analogue of the generalized Singleton bound in [17] is given for (n, k, d, r, δ)-matroids, and in particular to all almost affine codes in Theorem III.2.

(ii) In Theorem III.3, some necessary structural properties are given for an (n, k, d, r, δ)-matroid meeting the generalized Singleton bound.

(iii) In Theorem III.4, a class of matroids is given with different values of the parameters (n, k, d, r, δ). Simple and explicit constructions of matroids in this class are given in Corollary III.1, Theorem III.5, Theorem III.6, and Corollary III.2. The matroids are constructed via certain ranked subsets of the storage nodes, or via weighted graphs.

(iv) In Section IV-B, for the matroids mentioned above (cf. (iii)), we construct a directed graph that supports a gammoid isomorphic to the given matroid.

(v) Using the results in (iii), (iv) and results on representability of gammoids in [28], we obtain four explicit constructions of linear LRCs with given parameters in Section IV-D.

(vi) Theorem III.7 characterizes values of (n, k, r, δ) for which there exist (n, k, d, r, δ)-matroids meeting the bound (i). In particular, the nonexistence results for linear LRC in [21] are extended to the nonexistence of almost affine codes and matroids. Moreover, in Theorem IV.4 and Theorem IV.5, we settle the existence in one of the regimes left open in [21], leaving open only a minor subregime of $b > a \geq \lceil \frac{k}{r} \rceil - 1$, where $a = r \lceil \frac{k}{r} \rceil - k$ and $b = (r + \delta - 1)\lceil \frac{n}{r+\delta-1} \rceil - n$.

The proofs of some of the longer theorems and the explicit constructions of matroids with certain parameters (n, k, d, r, δ) are given in the Appendix.

II. Locally Repairable Codes, Almost Affine Codes and Matroids

A. *Parameters* (n, k, d, r, δ) *of locally repairable codes*

In this subsection, we introduce the parameters (n, k, d, r, δ) defined in [17] for linear locally repairable codes. We extend this definition to a much wider class of codes, including all almost affine codes, to be introduced in II-F. Figure 1 serves as a visual aid for the technical definitions. The information symbols (a, b, c, d, e, f) are stored on twelve nodes (code symbols) as in the figure. Within each of the local clouds (or locality sets), three symbols are enough to determine the other two. Globally, three lost code symbols can be restored using the remaining nine. Thus, Figure 1 depicts a $(12, 6, 4, 3, 3)$-LRC, according to the following definitions.

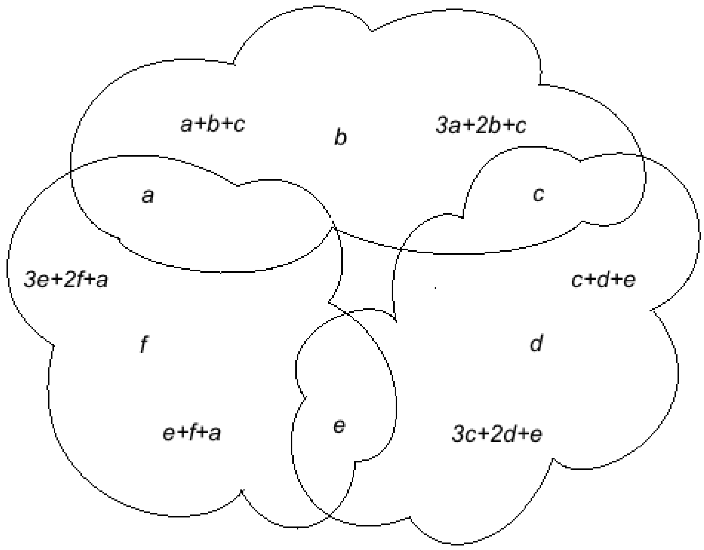

Fig. 1. A storage system from a $(12, 6, 4, 3, 3)$-LRC.

The ith coded symbol c_i in an $[n, k]$ linear code is said to have *repair locality* r, $(1 \leq r \leq k)$, if the code symbol c_i can be recovered by accessing at most r other symbols in the code. By a *linear* code we mean more precisely a linear code *over a finite field* $\mathbb{F}_q$, with length n, dimension k, and field size q. The concept of r-locality for linear codes was generalized to (r, δ)-locality in [17].

Let $C \subseteq \mathbb{A}^n$ be a code such that $|C| = |\mathbb{A}|^k$, where $\mathbb{A}$ is a finite set, also referred to as the *alphabet*. Moreover, for any subset $X = \{i_1, \ldots, i_m\} \subseteq [n] = \{1, 2, \ldots, n\}$, let C_X denote the *projection* of the code into $\mathbb{A}^{|X|}$, that is

$$C_X = \{(c_{i_1}, \ldots, c_{i_m}) : \boldsymbol{c} = (c_1, \ldots, c_n) \in C\}. \tag{1}$$

For $1 \leq r \leq k$ and $\delta \geq 2$, an (r, δ)-*locality set* of C is a subset $S \subseteq [n]$ such that

$$\begin{aligned} &(i) \quad |S| \leq r + \delta - 1, \\ &(ii) \quad l \in S,\, L = \{i_1, \ldots, i_{|L|}\} \subseteq S \setminus \{l\} \text{ and } |L| = |S| - (\delta - 1) \quad \Rightarrow \\ &\qquad\quad \exists f : C_L \to \mathbb{A} \text{ such that } f((c_{i_1}, \ldots, c_{i_{|L|}})) = c_l \text{ for all } \boldsymbol{c} \in C. \end{aligned} \tag{2}$$

We say that C is a *locally repairable code (LCR)* with *all-symbol locality* (r, δ) if all the n symbols of the code are contained in an (r, δ)-locality set. The locality sets can be also referred to as the local repair sets.

We remark that the symbols in a locality set S can be used to recover up to $\delta - 1$ lost symbols in the same locality set. Further, we note that the following conditions are equivalent to the condition (iii) above:

$$\begin{aligned} &(iii)' \quad l \in S,\, L = \{i_1, \ldots, i_{|L|}\} \subseteq S \setminus \{l\} \text{ and } |L| = |S| - (\delta - 1) \quad \Rightarrow \\ &\qquad\quad |C_{L \cup \{l\}}| = |C_L|, \\ &(iii)'' \quad d(C_S) \geq \delta, \text{ where } d(C_S) \text{ is the minimum (Hamming) distance of } C_S. \end{aligned}$$

An LRC with parameters (n, k), minimum (Hamming) distance d and all-symbol locality (r, δ) is an (n, k, d, r, δ)-LRC. Since we focus only on all-symbol locality in this paper, we will henceforth use the term LRC to mean a locally repairable code with all-symbol locality.

B. The Singleton bound

For any $[n, k]$-linear code with minimum distance d, the Singleton bound is given by

$$d \leq n - k + 1. \tag{3}$$

This bound was generalized in [10] to locally repairable codes with information locality. For any linear LRC with parameters (n, k, d, r), we have

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{4}$$

In [17], the bound was further generalized to codes with information locality (r, δ). A linear LRC with parameters (n, k, d, r, δ) satisfies

$$d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1). \tag{5}$$

Note that the bound (4) is derived from (5) by setting $\delta = 2$, and that the bound (3) is derived from (4) by setting $r = k$. Both the bounds (4) and (5) are stated assuming only information locality, but since a code with all-symbol locality (r, δ) has in particular the same information locality, the bounds are valid for all codes considered in this paper. Other generalizations of the Singleton bound for linear and nonlinear LRCs can be found in [13], [18], [19].

C. Graphs, $G = (V, E)$

Let us fix some standard graph-theoretic notation that will be used at two stages in the constructions. A (finite) directed graph $G = (V, E)$ is a pair of a finite *vertex set* V, whose elements are called nodes or vertices, and an *edge set* $E \subseteq V \times V$ of pairs called arcs or egdes. Graphs are often drawn with the vertices as points and arcs (v, u) as arrows $v \to u$. We call v the tail of (v, u), and u the head. A path from $S \subseteq V$ to $T \subseteq V$ is a sequence $v_0, v_1, \ldots, v_n$, where $v_0 \in S$, $v_n \in T$, and $(v_i, v_{i+1}) \in E$ for each $e = 0, \ldots, n-1$. If $v_0 = v_n$, then the path is called a (directed) cycle.

An important case of graphs is when E is *symmetric*, *i.e.*, $(u, v) \in E$ if and only if $(v, u) \in E$. In such case, it is customary to identify the two pairs (u, v) and (v, u) with the set $\{u, v\}$, and erase all the heads of the arrows in the drawing. When talking about a graph without specifying that it is directed, the symmetric situation is assumed. Observe that this definition allows for loops edges (where the tail and the head is the same), but not multiple edges. In this paper, we will assume that all graphs, both symmetric and directed, are without multiple edges and loops.

D. Posets and lattices, $(\mathcal{P}, \subseteq)$

Before studying matroids, we need a minumum of background on poset and lattice theory. We refer the reader to [29] for more information on posets and lattices.

A collection of sets $\mathcal{P} \subseteq 2^E$ ordered by inclusion $\subseteq$ defines a (finite) poset $(\mathcal{P}, \subseteq)$. A *chain* C of $(\mathcal{P}, \subseteq)$ is a set of elements $X_0, \ldots, X_m \in \mathcal{P}$ such that $X_0 \subsetneq X_1 \subsetneq \ldots \subsetneq X_m$. The *length* of a chain C is defined as the integer $l(C) = |C| - 1 = m$. For $X, Y \in \mathcal{P}$, let

$$L_{X,Y} = \{Z \in \mathcal{P} : Z \subseteq X \text{ and } Z \subseteq Y\},$$
$$U_{X,Y} = \{Z \in \mathcal{P} : X \subseteq Z \text{ and } Y \subseteq Z\}.$$

An element $Z \in L_{X,Y}$ is the *meet* of X and Y, denoted by $X \wedge Y$, if it contains every $V \in L_{X,Y}$. Dually, $Z \in U_{X,Y}$ is the *join* of X and Y, denoted by $X \vee Y$, if it is contained in every $V \in U_{X,Y}$. A poset $(\mathcal{P}, \subseteq)$ is a *lattice* if every pair of elements of $\mathcal{P}$ has a meet and a join. If $(\mathcal{P}, \subseteq)$ is a (finite) lattice, then there are two elements $0_{\mathcal{P}}, 1_{\mathcal{P}} \in \mathcal{P}$ such that $0_{\mathcal{P}} \subseteq X$ and $X \subseteq 1_{\mathcal{P}}$ for all $X \in \mathcal{P}$. Furthermore, the *atoms* and *coatoms* of a lattice $(\mathcal{L}, \subseteq)$ are defined as

$$\begin{aligned} A_{\mathcal{L}} &= \{X \in \mathcal{L} \setminus 0_{\mathcal{L}} : \nexists Y \in \mathcal{L} \text{ such that } 0_{\mathcal{L}} \subsetneq Y \subsetneq X\}, \\ coA_{\mathcal{L}} &= \{X \in \mathcal{L} \setminus 1_{\mathcal{L}} : \nexists Y \in \mathcal{L} \text{ such that } X \subsetneq Y \subsetneq 1_{\mathcal{L}}\}, \end{aligned}$$

respectively.

E. *Matroids,* $M = (\rho, E)$

In this paper, our main tools for analysing LRCs come from matroid theory. This is a branch of algebraic combinatorics with natural links to a great number of different topics, *e.g.*, to coding theory, graph theory, matching theory and combinatorial optimization. Matroids were introduced in [30] in order to abstractly capture the similar properties of linear independence in vector spaces and independence in graphs. Since its introduction, matroid theory have been successfully used to solve problems in many areas of mathematics and computer sciences, and as an example matroids play a fundamental role in combinatorial optimization. Today, matroid theory also plays important roles in information and coding theory, for example in the areas of network coding, secret sharing, index coding, information inequalities and more [31], [32], [33]. Any almost affine code C, so in particular any linear codes, are associated with a matroid M_C in a specific way [15]. However, there are many matroids which cannot be represented by any almost affine code [34]. A matroid is called *representable* over a finite field $\mathbb{F}_q$ if it can be represented by a linear code over $\mathbb{F}_q$. In this paper, while our main goal is investigating almost affine LRCs by matroid theory, ideas from the theory of LRCs will also be used to give new results on matroids in general.

Matroids can be defined in many equivalent ways, for example by their rank function, nullity function, independent sets, circuits and more [35]. For our purpose, the following definition will be the most useful. Let 2^E denote the set of all subsets of E. A *matroid* M on a finite set E is defined by a *rank function* $\rho : 2^E \to \mathbb{Z}$ satisfying the following conditions:

$$\begin{array}{ll} (R1) & 0 \leq \rho(X) \leq |X| \text{ for } X \subseteq E, \\ (R2) & X \subseteq Y \subseteq E \Rightarrow \rho(X) \leq \rho(Y), \\ (R3) & X, Y \subseteq E \Rightarrow \rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y). \end{array} \tag{6}$$

The *nullity function* $\eta : 2^E \to \mathbb{Z}$ of the matroid $M = (E, \rho)$ is defined by

$$\eta(X) = |X| - \rho(X), \text{ for } X \subseteq E.$$

Let X be any subset of E. The subset X is *independent* if $\rho(X) = |X|$, otherwise it is *dependent*. A dependent set X is a *circuit* if all proper subsets of X are independent, *i.e.*, $\rho(X) = |X| - 1$ and $\rho(Y) = |Y|$ for all subsets $Y \subsetneq X$. The *closure* of X is defined as

$$\mathrm{cl}(X) = \{x \in E : \rho(X \cup x) = \rho(X)\}.$$

The subset X is a *flat* if $\mathrm{cl}(X) = X$. It is *cyclic* if it is a (possible empty) union of circuits. Equivalently, X is cyclic if and only if there is no $x \in X$ such that $\rho(X \setminus \{x\}) = \rho(X) - 1$. A flat that is cyclic is called a *cyclic flat*. The sets of circuits, independent sets, cyclic sets and cyclic flats of a matroid M is denoted by $\mathcal{C}(M)$, $\mathcal{I}(M)$, $\mathcal{U}(M)$ and $\mathcal{Z}(M)$, respectively. We omit the subscript M when the matroid is clear and write $\mathcal{C}$, $\mathcal{I}$, $\mathcal{U}$ and $\mathcal{Z}$, respectively. The set of cyclic flats together with inclusion defines the *lattice of cyclic flats* $(\mathcal{Z}, \subseteq)$ of the matroid. The *restriction* of M to X is the matroid $M|X = (\rho_{|X}, X)$ where

$$\rho_{|X}(Y) = \rho(Y), \text{ for all subsets } Y \subseteq X. \tag{7}$$

The *dual* of the matroid $M = (\rho, E)$ is the matroid $M^* = (\rho^*, E)$, where

$$\rho^*(X) = \rho(E \setminus X) + |X| - \rho(E), \ \text{ for } X \subseteq E. \tag{8}$$

If we omit the argument M^*, then we denote the sets of circuits and cyclic flats of the dual matroid by $\mathcal{C}^*$ and $\mathcal{Z}^*$, respectively.

Example II.1. *An example of a matroid $M_G = (\rho, E)$ is defined by the matrix*

$$G = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{cccccccccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} & \mathbf{9} & \mathbf{10} & \mathbf{11} & \mathbf{12} \end{array} \\ \begin{pmatrix} 1 & & & & & & 1 & & 1 & 3 & & 1 \\ & 1 & & & & & 1 & & & 2 & & \\ & & 1 & & & & 1 & 1 & & 1 & 3 & \\ & & & 1 & & & & 1 & & & 2 & \\ & & & & 1 & & & 1 & 1 & & 1 & 3 \\ & & & & & 1 & & & 1 & & & 2 \end{pmatrix} \end{array}, \tag{9}$$

which we think of as a generator matrix of a linear code C over the field $\mathbb{F}_5$. $E = \{\mathbf{1}, \dots, \mathbf{12}\}$ is the set of columns, and the rank of a subset of E is the rank of the corresponding submatrix, i.e.,

$$\rho(I) = rank(G_I) \textit{ for } I \subseteq E,$$

where G_I is the submatrix of G whose columns are the columns indexed by I. Below are some independent sets, circuits, cyclic flats and rank functions of some subsets of E for the matroid M.

$\mathcal{I} = \{\emptyset, \{2,3,7\}, \{3,4,5\}, \{7,8,9\}, [6], \dots\}$,
$\mathcal{C} = \{\{1,2,3,7\}, \{4,5,8,11\}, \dots\}$,
$\mathcal{Z} = \{\{1,2,3,7,10\}, \{3,4,5,8,11\}, \{1,2,3,4,5,7,8,10,11\}, [12], ...\}$,
$\rho(\emptyset) = 0, \ \rho(\{3,4,5\}) = \rho(\{4,5,8,11\}) = \rho(\{3,4,5,8,11\}) = 3, \ \rho([6]) = \rho([12]) = 6.$

The reader can verify that the code generated by this matrix corresponds to the storage system in Figure 1, when the rows are the information symbols.

F. Almost affine codes and their associated matroids

A code $C \subseteq \mathbb{A}^n$, where $\mathbb{A}$ is a finite set of size $s \geq 2$, is *almost affine* if

$$\log_s(|C_X|) \in \mathbb{Z}$$

for each $X \subseteq [n]$. Note that if C is an almost affine code, then all projections C_X of C are also almost affine.

Two special cases of almost affine codes are linear codes and m-multilinear codes over finite fields. A code $C \subseteq \mathbb{F}_q^{mn}$ is an *m-multilinear code* over $\mathbb{F}_q$ if C is a linear code over $\mathbb{F}_q$ and, when considering C as a code in $(\mathbb{F}_q^m)^n$, the dimension of the vector space C_X over $\mathbb{F}_q$ is divisible by m for all $X \subseteq [n]$. We remark that a linear code is a 1-multilinear code and that an m-multilinear code is a special case of vector-linear codes (linear codes where each code symbol is a vector). Not all vector-linear codes are almost affine and vice versa.

In [15] it is proven that every almost affine code $\mathbb{A} \subseteq S^n$ induces a matroid $M_C = (\rho_C, [n])$, where

$$\rho_C(X) = \log_s(|C_X|). \tag{10}$$

Examples of matroids M which cannot be represented by any almost affine code are given in [34]. Moreover, an example of a matroid (the so called non-Pappus matroid) which can be represented by a 2-multilinear code over $\mathbb{F}_3$ but not by any linear code is given in [15].

G. Basic properties of matroids and the lattice of cyclic flats

For the applications in this paper, the most important matroid attribute is its lattice of cyclic flats. We will need a few basic properties of this lattice (for instance the nontrivial property that it is indeed a lattice), that can be found in [36].

Proposition II.1. *Let $M = (\rho, E)$ be a matroid. Then*

(i) $X \in \mathcal{Z} \iff (E \setminus X) \in \mathcal{Z}^*$,

(ii) $\rho(X) = \min\{\rho(F) + |X \setminus F| : F \in \mathcal{Z}\}$, *for* $X \subseteq E$,

(iii) Define $\mathcal{D} = \{X : \textit{there is } F \in \mathcal{Z} \textit{ with } X \subseteq F \textit{ and } |X| = \rho(F) + 1\}$.

Then $\mathcal{C} = \{X \in \mathcal{D} : \nexists Y \in \mathcal{D} \textit{ such that } Y \subsetneq X\}$, *(iv)* $(\mathcal{Z},\subseteq)$ *is a lattice with the following meet and join for* $X,Y\in\mathcal{Z}$*,*

$X\wedge Y=\bigcup_{\{C\in\mathcal{C}:C\subseteq X\cap Y\}}C$ *and* $X\vee Y=\mathrm{cl}(X\cup Y)$*.*

The condition (ii) in Proposition II.1 shows that a matroid is determined by its cyclic flats and their ranks. Conversely, the following theorem gives an axiomatic scheme for a collection of subsets on E and a function on these sets to define the cyclic flats of a matroid and their ranks.

Theorem II.1 (see [36] Th. 3.2)**.** *Let* $\mathcal{Z}\subseteq 2^E$ *and let* ρ *be a function* $\rho:\mathcal{Z}\to\mathbb{Z}$*. There is a matroid* M *on* E *for which* $\mathcal{Z}$ *is the set of cyclic flats and* ρ *is the rank function restricted to the sets in* $\mathcal{Z}$ *if and only if*

$$
\begin{aligned}
(Z0)\quad & \mathcal{Z}\ \textit{is a lattice under inclusion},\\
(Z1)\quad & \rho(0_{\mathcal{Z}})=0,\\
(Z2)\quad & X,Y\in\mathcal{Z}\ \textit{and}\ X\subsetneq Y\Rightarrow\\
& 0<\rho(Y)-\rho(X)<|Y|-|X|,\\
(Z3)\quad & X,Y\in\mathcal{Z}\Rightarrow\rho(X)+\rho(Y)\geq\\
& \rho(X\vee Y)+\rho(X\wedge Y)+|(X\cap Y)\setminus(X\wedge Y)|.
\end{aligned}
$$

Many of the results in Proposition II.2 below have already been proved many times in different places in the rich literature of matroids. We give a proof for these results that we have not been able to found in the literature, for the other results we give a reference.

Proposition II.2. *Let* $M=(\rho,E)$ *be a matroid and let* X,Y *be subsets of* E*, then*

(i) *If* $X\subseteq Y$*, then* $\eta(X)\leq\eta(Y)$*,*

(ii) $\eta(X\cup Y)\geq\eta(X)+\eta(Y)-\eta(X\cap Y)$*,*

(iii) *If* $\rho(X)<\rho(E)$ *and* $1_{\mathcal{Z}}=E$*, then* $\eta(X)\leq\max\{\eta(Z):Z\in coA_{\mathcal{Z}}\}$*,*

(iv) $\mathrm{cl}(U)\in\mathcal{Z}(M)$ *for* $U\in\mathcal{U}(M)$*,*

(v) $\mathcal{U}(M|X)=\{U\subseteq X:U\in\mathcal{U}(M)\}$*,*

(vi) $\mathcal{C}(M|X)=\{C\subseteq X:C\in\mathcal{C}(M)\}$*,*

(vii) $\mathcal{Z}(M|X)=\{Z\in\mathcal{Z}(M):Z\subseteq X\}$ *if* $X\in\mathcal{F}(M)$*,*

(viii) $X\notin\mathcal{U}(M)$ *if and only if* $\min\{|D|:D\in\mathcal{C}((M|X)^*)\}=1$*,*

(ix) $\rho(\mathrm{cl}(X))=\rho(X)$*,*

(x) *If* $X\subseteq Y$*, then* $\mathrm{cl}(X)\subseteq\mathrm{cl}(Y)$*.*

Proof. Property (i), (ii), (v) and (vii) can be found in [37, Lemma 2.2.4, Lemma 2.3.1]. Property (iv) is a consequence of [35, Proposition 1.4.10 (ii)]. For (iii), assume that $\rho(X)<\rho(E)$ and $1_{\mathcal{Z}}=E$.

Thus, $\mathrm{cl}(X) \neq E$ and $\eta(X) \leq \eta(\mathrm{cl}(X))$. Let U be the largest cyclic set such that $U \subseteq \mathrm{cl}(X)$. From [37, Lemma 2.4.8, Lemma 2.5.2], we have that $\eta(\mathrm{cl}(X)) = \eta(U)$ and that U is a cyclic flat. Property (iv) now follows from the fact that

$$\rho(U) \leq \rho(\mathrm{cl}(X)) < \rho(E) = \rho(1_{\mathcal{Z}}).$$

Property (vi) is a direct consequence of (v). For (viii), by using (i) and (iii) in Proposition II.1, (Z1) in Theorem II.1 and (v) in this proposition, we obtain that

$$\begin{aligned} X \notin \mathcal{U}(M) \quad &\iff \quad 1_{\mathcal{Z}(M|X)} \neq X \\ &\iff \quad 0_{\mathcal{Z}((M|X)^*)} \neq \emptyset \\ &\iff \quad \min\{|D| : D \in \mathcal{C}((M|X)^*)\} = 1. \end{aligned}$$

Property (ix) is a consequence of property (x) which can be found in [35, Lemma 1.4.2] □

Example II.2. *Continuing the example from Example II.1, and remembering that the elements of M_G are the columns of G, we see that the cyclic flats of M_G are the submatrices in Figure 2. The atomic cyclic flats are thus the submatrices corresponding to column sets $\{1, 2, 3, 7, 10\}$, $\{3, 4, 5, 8, 11\}$ and $\{1, 5, 6, 9, 12\}$. Remembering from (9) that the rows are indexed by the information symbols (a, b, c, d, e, f), these atomic cyclic flats agree exactly with the local clouds in Figure 1.*

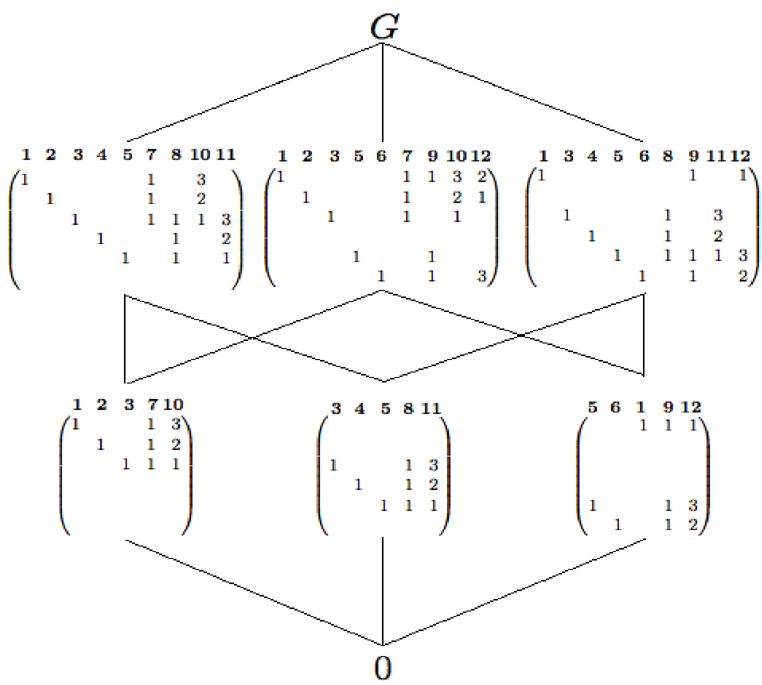

Fig. 2. The lattice $\mathcal{Z}(M_G)$ of cyclic flats.

III. Locally Repairable Matroids

A. The parameters (n, k, d, r, δ) for matroids

In this subsection we will first show that the parameters (n, k, d, r, δ) are matroid invariants for an almost affine LRC. This will allow us to extend the definition of these parameters to matroids in general.

Let C be an almost affine (n, k, d, r, δ)-LRC over some finite alphabet $\mathbb{A}$. The parameters (n, k, d, r, δ) of C can be detected by its matroid representation $M_C = (\rho_C, [n])$ as follows:

In [15] it is proved that d equals the minimum size of a circuit of the dual matroid M_C^*, and that $M_{C_X} = M|X$ for $X \subseteq [n]$. Since the projection C_X also is almost affine, it now follows that

$$d(C_X) = \min\{|X| : X \in \mathcal{C}((M|X)^*)\},$$

where $d(C_X)$ denotes the minimum distance of C_X. Furthermore, from (10) it follows that

$$|C_X| = |\mathbb{A}|^{\rho_C(X)} \text{ for } X \subseteq [n].$$

Hence $|C| = |\mathbb{A}|^{\rho_C([n])}$, which implies that $k = \rho_C([n])$.

By using the observations above and the definition of an (n, k, d, r, δ)-LRC given in Section II-A, we can state the following theorem.

Theorem III.1. *Let C be an almost affine LRC with the associated matroid $M_C = (\rho_C, [n])$. Then, the parameters (n, k, d, r, δ) of C are matroid invariants, where*

(i) $k = \rho_C([n])$,

(ii) $d = \min\{|X| : X \in \mathcal{C}(M_C^*)\}$,

(iii) *C has all-symbol locality (r, δ) if and only if, for every $j \in [n]$ there exists a subset $S_j \subseteq [n]$ such that*

- *a)* $j \in S_j$,
- *b)* $|S_j| \leq r + \delta - 1$,
- *c)* $d(C_{S_j}) = \min\{|X| : X \in \mathcal{C}((M_C|S_j)^*)\} \geq \delta$,

or the equivalent condition to (c) that

- *c′)* *For all $L \subseteq S_j$ with $|L| = |S_j| - (\delta - 1)$, and all $l \in S_j \setminus L$, we have $\rho_C(L \cup l) = \rho_C(L)$.*

These results can now be taken as the definition of the parameters (n, k, d, r, δ) for an arbitrary matroid.

Definition III.1. *Let $M = (\rho, E)$ be a matroid. Then, for $0 < r \leq \rho(E)$ and $\delta \geq 2$,*

(i) $n = |E|$,

(ii) $k = \rho(E)$,

(iii) $d = \min\{|X| : X \in \mathcal{C}(M^*)\}$,

(iv) *M has all-symbol locality (r, δ) if and only if for all $x \in E$ there exists a subset $S_x \subseteq E$ such that*

a) $x \in S_x$,

b) $|S_x| \leq r + \delta - 1$,

c) $d(M|S_x) = \min\{|X| : X \in \mathcal{C}((M|S_x)^*)\} \geq \delta$.

A subset $S \subseteq E$ is called a *locality set* of M if the conditions (b)–(c) above is satisfied by S. For any matroid $N = (\rho_N, Y)$ and $X \subseteq Y$, X is a circuit of N^* if and only if $Y \setminus X$ is a flat of N with $\rho_N(X) = \rho_N(Y) - 1$, [35]. So two equivalent statements of (c) in Lemma III.1 (iv) are that

$$\begin{aligned} d(M|S_x) \geq \delta \quad &\iff \quad (X \subseteq S_x,\, |X| \geq |S_x| - (\delta - 1) \Rightarrow \rho(X) = \rho(S_x)) \\ &\iff \quad \min\{|X| : X \subseteq S_x,\, \rho(S_x \setminus X) < \rho(S_x)\} \geq \delta. \end{aligned} \tag{11}$$

The parameters n and k are obviously defined for all matroids. We note that the parameter d is finite if and only if the dual matroid M^* contains a circuit. Thus, by Proposition II.1, d is infinite if and only if

$$\mathcal{C}(M^*) = \emptyset \iff 0_{\mathcal{Z}^*} = 1_{\mathcal{Z}^*} = \emptyset \iff 0_Z = 1_Z = E \iff k = \rho(E) = 0.$$

Furthermore, we notice that every element $x \in E$ is contained in some cyclic set S_x if and only if $1_{\mathcal{Z}} = E$. If this is the case, and $r = \max\{|S_x| - 1 : x \in X\}$, then M has $(r, 2)$-locality. As a consequence of the observations above, we get the following proposition.

Proposition III.1. *Let $M = (\rho, E)$ be a matroid. All the parameters (n, k, d, r, δ) are defined and finite for M if and only if $0 < \rho(E)$ and $1_{\mathcal{Z}} = E$.*

We observe that, if M has (r, δ)-locality then, by definition III.1 (iv) and (11), M has (r', δ')-locality for $r \leq r' \leq k$ and $2 \leq \delta' \leq \delta$ with $r' + \delta' \geq r + \delta$. So neither the values of (r, δ) nor the locality sets S_x are in general uniquely determined in the matrix M.

B. A generalized Singleton bound for (n, k, d, r, δ)-matroids

In this subsection we relate the parameters (n, k, d, r, δ) of a matroid to its lattice of cyclic flats in Lemmas III.1–III.3. Using this, in Theorem III.2 we prove a generalized Singleton bound for matroids.

Lemma III.1. *Let $M = (\rho, E)$ be a matroid with $0 < \rho(E)$ and $1_{\mathcal{Z}} = E$. Then*

(i) $n = |1_{\mathcal{Z}}|$,

(ii) $k = \rho(1_{\mathcal{Z}})$,

(iii) $d = n - k + 1 - \max\{\eta(Z) : Z \in coA_{\mathcal{Z}}\}$,

(iv) M has locality (r, δ), if and only if for each $x \in E$, there is a cyclic set $S_x \in \mathcal{U}(M)$ such that

a) $x \in S_x$,

b) $|S_x| \leq r + \delta - 1$,

c) $d(M|S_x) = \eta(S_x) + 1 - \max\{\eta(Z) : Z \in coA_{\mathcal{Z}(M|S_x)}\} \geq \delta$.

Proof. The proof is given in the Appendix. □

We remark that

$$\rho(S_x) \leq r \tag{12}$$

for any S_x given in Lemma III.1 (iv). This is a consequence of (b) and (c) in Lemma III.1 which implies that $\delta + \rho(S_x) - 1 \leq |S_x| \leq r + \delta - 1$. Moreover, we observe that for any atom S in a lattice of cyclic flats with $0_{\mathcal{Z}} = \emptyset$, we can use any subset $S' \subseteq S$ as a locality set when $|S'| > \rho(S)$. However, different choices of locality sets may give different values on the parameters (r, δ).

Example III.1. *Representing the cyclic flats associated to the matroid M_G from Example II.2 just by their corresponding sets and ranks in Figure 3, we use Lemma III.1 to get the parameters (n, k, d, r, δ) of the linear LRC that is generated by the matrix G given in Example II.1.*

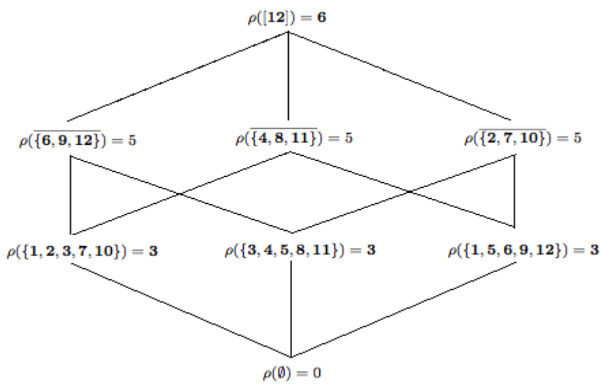

Fig. 3. The lattice $\mathcal{Z}(M_G)$ of cyclic flats, without reference to the matrix G.

The values for (n,k,d) *are*

$$n = 12,$$
$$k = 6,$$
$$d = 12 - 6 + 1 - 4 = 3.$$

Using $S_1 = \{1,2,3,7,10\}$*,* $S_2 = \{3,4,5,8,11\}$ *and* $S_3 = \{1,5,6,9,12\}$ *as the locality sets, we get the parameters* $(r,\delta) = (3,3)$*. If we choose some other locality sets, e.g.* $S_1' = \{2,3,7,10\}$*,* $S_2' = \{4,5,8,11\}$*,* $S_3' = \{1,5,6,9\}$ *and* $S_4' = \{5,6,9,12\}$*, we get the parameters* $(r,\delta) = (3,2)$*.*

From Lemma III.1 above, we can derive the following:

Lemma III.2. *Let* $M = (\rho, E)$ *be a matroid with parameters* (n,k,d,r,δ)*, and let* $\{S_x\}_{x\in E}$ *be a collection of cyclic sets of* M *for which the conditions* $(a)-(c)$ *in Lemma III.1 are satisfied. Then there is a subset of cyclic sets* $\{S_j\}_{j\in[m]}$ *of* $\{S_x\}_{x\in E}$ *such that*

(i) $C : 0_{\mathcal{Z}} = Y_0 \subsetneq Y_1 \subseteq \ldots \subsetneq Y_m = E$ *is a chain in* $(\mathcal{Z}(M), \subseteq)$*,*

(ii) $\rho(Y_j) \leq \rho(Y_{j-1}) + \rho(S_j) - \rho(Y_{j-1} \cap S_j) \leq \rho(Y_{j-1}) + r$*,*

(iii) $\eta(Y_j) \geq \eta(Y_{j-1}) + \eta(S_j) - \eta(Y_{i-1} \cap S_j) \geq \eta(Y_{j-1}) + (\delta - 1)$

for $Y_j = \mathrm{cl}(Y_{j-1} \cup S_j) = Y_{j-1} \vee \mathrm{cl}(S_j)$*, where* $j = 1, \ldots, m$*.*

Proof. The proof is given in the Appendix. □

The following lemma is now a consequence of Lemma III.1 and Lemma III.2.

Lemma III.3. *Let* $M = (\rho, E)$ *be a matroid with parameters* (n,k,d,r,δ)*, and let*

$$C : 0_{\mathcal{Z}} = Y_0 \subsetneq Y_1 \subseteq \ldots \subsetneq Y_m = E$$

be any chain of $(\mathcal{Z}, \subseteq)$ *given in Lemma III.2. Then*

$$d \leq n - k + 1 - \eta(Y_{m-1}) \quad \text{and} \quad m \geq \left\lceil \frac{k}{r} \right\rceil.$$

Proof. The proof is given in the Appendix. □

The following generalized Singleton bound for matroids, now follows immediately from Lemmas III.1–III.3.

Theorem III.2. *Let* $M = (\rho, E)$ *be an* (n,k,d,r,δ)*-matroid. Then*

$$d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1).$$

Proof. Let

$$C : 0_{\mathcal{Z}} = Y_0 \subsetneq Y_1 \subsetneq \ldots \subsetneq Y_m = E$$

be a chain of $(\mathcal{Z}, \subseteq)$ given in Lemma III.2 (i). Observe that $\eta(Y_{m-1}) \geq (m-1)(\delta-1)$, by Lemma III.2 (iii). Hence, it follows from Lemma III.3 that

$$d \leq n - k + 1 - \eta(Y_{m-1}) \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1).$$

□

We can now give the following bound on δ, k and the rate $\frac{k}{n}$.

Proposition III.2. *Let $M = (\rho, E)$ be an (n, k, d, r, δ)-matroid. Then*

(i) $\delta \leq d$,

(ii) $k \leq n - \left\lceil \frac{k}{r} \right\rceil (\delta - 1)$,

(iii) $\frac{k}{n} \leq \frac{r}{r+\delta-1}$.

Proof. For (i), let Y be any subset of E with $|Y| < \delta$. From Definition III.1 and (11), we conclude for every $x \in E$ that there is a subset $S_x \subseteq E$ with $x \in S_x$ and $\rho(S_x \setminus Y) = \rho(S_x)$. Hence, by Proposition II.2 (ix) and (x), $\mathrm{cl}(S_x \setminus Y) = \mathrm{cl}(S_x)$. This implies that $E = \bigcup_{x \in E} S_x \subseteq \mathrm{cl}(E \setminus Y)$, and consequently $\rho(E \setminus Y) = \rho(E)$. Since

$$d = \min\{|X| : X \subseteq E,\, \rho(E \setminus X) < \rho(E)\}$$

by (11), we now obtain that $\delta \leq d$.

For (ii), by (i) and Lemma III.2,

$$\delta \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1).$$

Therefore

$$k \leq n - \left\lceil \frac{k}{r} \right\rceil (\delta - 1).$$

For (iii), by (ii), we have

$$\frac{n}{k} \geq 1 + \left\lceil \frac{k}{r} \right\rceil \frac{(\delta - 1)}{k} \geq 1 + \frac{\delta - 1}{r} = \frac{r + \delta - 1}{r}.$$

□

C. A structure theorem for matroids achieving the generalized Singleton bound

The following theorem gives some necessary structural properties for (n, k, d, r, δ)-matroids, when d meets the generalized Singleton bound for matroids in Theorem III.2, and $r < k$. The case when $r = k$ is easier, and is considered in Section III-D3.

A collection of sets $X_1 \ldots, X_j$ has a *non trivial union* if

$$X_l \nsubseteq \bigcup_{i \in [j] \setminus \{l\}} X_i, \text{ for } l = 1, \ldots, j.$$

Theorem III.3. *Let $M = (\rho, E)$ be an (n, k, d, r, δ)-matroid with $r < k$ and*

$$d = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1).$$

Also, let $\{S_x : x \in E\} \subseteq \mathcal{U}(M)$, be a collection of cyclic sets for which the conditions $(a) - (c)$ in Lemma III.1 are satisfied. Then

(i) $0_{\mathcal{Z}} = \emptyset$,

(ii) for each $x \in E$,

a) $\eta(S_x) = (\delta - 1)$,

b) S_x is a cyclic flat and $\mathcal{Z}(M|S_x) = \{X \in \mathcal{Z}(M) : X \subseteq S_x\} = \{\emptyset, S_x\}$,

(iii) For each collection $F_1, \ldots F_j$ of cyclic flats in $\{S_x : x \in E\}$ that has a non trivial union,

c) $\eta(\bigvee_{i=1}^{j} F_i) = \begin{cases} j(\delta - 1) & \text{if} \quad j < \lceil \frac{k}{r} \rceil, \\ n - k \geq \lceil \frac{k}{r} \rceil (\delta - 1) & \text{if} \quad j \geq \lceil \frac{k}{r} \rceil, \end{cases}$

d) $\bigvee_{i=1}^{j} F_i = \begin{cases} \bigcup_{i=1}^{j} F_i & \text{if} \quad j < \lceil \frac{k}{r} \rceil, \\ E & \text{if} \quad j \geq \lceil \frac{k}{r} \rceil, \end{cases}$

e) $\rho(\bigvee_{i=1}^{j} F_i) = \begin{cases} |\bigcup_{i=1}^{j} F_i| - j(\delta - 1) & \text{if} \quad j < \lceil \frac{k}{r} \rceil, \\ k & \text{if} \quad j \geq \lceil \frac{k}{r} \rceil. \end{cases}$

f) $|F_j \cap (\bigcup_{i=1}^{j-1} F_i)| \leq |F_j| - \delta$ *if* $j \leq \lceil \frac{k}{r} \rceil$.

Proof. The proof is given in the Appendix. □

Note that even if the conditions (i)–(iii) are satisfied by a subset of cyclic flats of $\mathcal{Z}(M)$ of a matroid M, $\mathcal{Z}(M)$ can still contain cyclic flats such that the parameter d of the matroid does not meet the generalized Singleton bound.

D. *The maximal values of d for (n,k,r,δ)-matroids*

The generalized Singleton bound theorem for matroids gives an upper bound for the value of d in terms of the parameters (n,k,r,δ) for a matroid. In this section we will investigate, given different classes of the parameters (n,k,r,δ), whether or not this bound can be achieved. We will divide the parameters (n,k,r,δ) into two main classes, namely when $r=k$ and $r<k$. The parameters (n,k,r,δ) will be divided into four subclasses in the case when $r<k$.

1) A construction of (n,k,d,r,δ)-matroids: Let $F_1,\dots,F_m$ be a collection of subsets of a finite set E, and let $F_I=\bigcup_{i\in I}F_i$ for $I\subseteq[m]$. Let k be a positive integer and let ρ be a function $\rho:\{F_i\}_{i\in[m]}\to\mathbb{Z}$ such that

(i) $\{F_i\}_{i\in[m]}$ is a nontrivial union, with $F_{[m]}=E$,

(ii) $0<\rho(F_i)<|F_i|$ for every $i\in[m]$,

(iii) There exists $I\subseteq[m]$ such that $|F_I|-\sum_{i\in I}\eta(F_i)\geq k$,

(iv) If $F_I\in\mathcal{Z}_{<k}$ and $j\in[m]\setminus I$, then $|F_I\cap F_j|<\rho(F_j)$,

(v) If $F_I,F_J\in\mathcal{Z}_{<k}$ and $F_{I\cup J}\notin\mathcal{Z}_{<k}$, then $|F_{I\cup J}|-\sum_{t\in I\cup J}\eta(F_t)\geq k$,

where we define

$$\eta(F_i)=|F_i|-\rho(F_i)\text{ for }i\in[m]$$

and

$$\mathcal{Z}_{<k}=\{F_J:J\subseteq[m]\text{ and }|F_I|-\sum_{i\in I}\eta(F_i)<k\text{ for all }I\subseteq J\}.$$

Now, to a collection of subsets $F_1,\dots,F_m$ of E, integer k and function ρ that satisfy the conditions (i)–(v), we extend $\mathcal{Z}_{<k}$ to $\mathcal{Z}$ and extend the function ρ to a function on $\mathcal{Z}$, by $\mathcal{Z}=\mathcal{Z}_{<k}\cup E$ and

$$\rho(X)=\begin{cases}|F_I|-\sum_{i\in I}\eta(F_i) & \text{if } X=F_I\in\mathcal{Z}_{<k},\\ k & \text{if } X=E.\end{cases}\tag{13}$$

Note that the extension of ρ given in (13) is well defined, as by property (iv) the set E is not in $\mathcal{Z}_{<k}$. Also note that $F_\emptyset=\emptyset$ and $\rho(F_\emptyset)=0$.

Theorem III.4. *Let $F_1,\dots,F_m$ be a collection of subsets of a finite set E, k a positive integer and $\rho:\{F_i\}_{i\in[m]}\to\mathbb{Z}$ a function such that the conditions (i)–(v) in Section III-D1 are satisfied. Then $\mathcal{Z}$ and $\rho:\mathcal{Z}\to\mathbb{Z}$, defined in* (13)*, define an (n,k,d,r,δ)-matroid $M(F_1,\dots,F_m;k;\rho)$ on E for which $\mathcal{Z}$ is the collection of cyclic flats, ρ is the rank function restricted to the cyclic flats and*

(i) $n=|E|$,

(ii) $k = \rho(E)$,

(iii) $d = n - k + 1 - \max\{\sum_{i\in I} \eta(F_i) : F_I \in \mathcal{Z}_{<k}\}$,

(iv) $\delta - 1 = \min_{i\in[m]}\{\eta(F_i)\}$,

(v) $r = \max_{i\in[m]}\{\rho(F_i)\}$.

For each $i \in [m]$, *any subset* $S \subseteq F_i$ *with* $|S| = \rho(F_i) + \delta - 1$ *is a locality set of the matroid.*

Proof. The proof is given in the Appendix □

A subclass of the matroids that we can get from Theorem III.4 is given in the following corollary.

Corollary III.1. *Let* $F_1, \dots, F_m$ *be a collection of subsets of a finite set* E, k *a positive integer and* $\rho : \{F_i\}_{i\in[m]} \to \mathbb{Z}$ *a function such that*

(i) $0 < \rho(F_i) < |F_i|$ *for* $i \in [m]$,

(ii) $F_{[m]} = E$,

(iii) $k \leq |F_{[m]}| - \sum_{i\in[m]} \eta(F_i)$,

(iv) $I \subseteq [m], j \in [m] \setminus I \Rightarrow |F_I \cap F_j| < \rho(F_j)$.

Then $F_1, \dots, F_m$, k *and* ρ *define a matroid* $M(F_1, \dots, F_m; k; \rho)$ *given in Theorem III.4.*

Proof. The proof is given in the Appendix □

Proposition III.3. *Let* $F_1, \dots, F_m$ *be a collection of subsets of a finite set* E, k *a positive integer and* $\rho : \{F_i\}_{i\in[m]} \to \mathbb{Z}$ *a function such that the statements* $(i) - (iv)$ *in Corollary III.1 hold. Let* $\mathcal{I}$ *denote the collection of independent sets in the matroid* $M(F_1, \dots, F_m; k; \rho)$ *that we obtain in Corollary III.1. Then*

(i) $|F_J| - \sum_{j\in J} \eta(F_j) > |F_I| - \sum_{i\in I} \eta(F_i)$, *for* $I \subsetneq J \subseteq [m]$,

(ii) $\mathcal{Z} = \{F_I : |F_I| - \sum_{i\in I} \eta(F_i) < k\} \cup E$,

(iii) $\rho(F_I) = \min\{|F_I| - \sum_{i\in I} \eta(F_i), k\}$, *for* $I \subseteq [m]$,

(iv) $\mathcal{I} = \{X \subseteq E : |F_I \cap X| \leq \min\{|F_I| - \sum_{i\in I} \eta(F_i), k\}$ *for all* $I \subseteq [m]\}$,

(v) For $X \subseteq F_I$, $X \in \mathcal{I} \iff |F_{I'} \cap X| \leq \min\{|F_{I'}| - \sum_{i\in I'} \eta(F_i), k\}$ *for all* $I' \subseteq I\}$.

Proof. The proof is given in the Appendix □

2) Constructions of (n, k, d, r, δ)*-matroids via graphs:* In Section III-D4 we will use two subclasses of the $M(F_1, \dots, F_m; k; \rho)$ matroids given in Corollary III.1 in order to construct (n, k, d, r, δ)-matroids with large d. These two subclasses can be represented by graphs together with some functions on the vertices and edges.

The graph construction 1: Let $G = G(\alpha, \beta, \gamma; k, r, \delta)$ be a graph with vertices $[m]$ and edges W, where (α, β) are two functions $[m] \to \mathbb{Z}$, $\gamma : W \to \mathbb{Z}$ and (k, r, δ) are three integers with $0 < r < k$ and $\delta \geq 2$, such that

(14)

(i) G is a graph with no 3-cycles,
(ii) $0 \leq \alpha(i) \leq r - 1$ for $i \in [m]$,
(iii) $\beta(i) \geq 0$ for $i \in [m]$,
(iv) $\gamma(w) \geq 1$ for $w \in W$,
(v) $k \leq rm - \sum_{i\in[m]} \alpha(i) - \sum_{w\in W} \gamma(w)$,
(vi) $r - \alpha(i) > \sum_{w=\{i,j\}\in W} \gamma(w)$ for $i \in [m]$.

Theorem III.5. *Let $G(\alpha, \beta, \gamma; k, r, \delta)$ be a graph on $[m]$ such that the conditions (i)–(vi) given in* (14) *are satisfied. Then there is an (n, k, d, r, δ)-matroid $M(F_1, \ldots, F_m; k; \rho)$ given in Corollary III.1 with*

(i) $n = (r + \delta - 1)m - \sum_{i\in[m]} \alpha(i) + \sum_{i\in[m]} \beta(i) - \sum_{w\in W} \gamma(w)$,
(ii) $d = n - k + 1 - \max_{I\in V_{<k}} \{(\delta - 1)|I| + \sum_{i\in I} \beta(i)\}$,

where

$$V_{<k} = \{I \subseteq [m] : r|I| - \sum_{i\in I} \alpha(i) - \sum_{i,j\in I,\ w=\{i,j\}\in W} \gamma(w) < k\}.$$

Proof. The proof is given in the Appendix. □

The construction of the sets $F_1, \ldots, F_m$ and the rank function ρ proceeds by first assigning the folowing:

(15)

(i) $\rho(F_i) = r - \alpha(i)$ for $i \in [m]$,
(ii) $|F_i| = r + \delta - 1 - \alpha(i) + \beta(i)$ for $i \in [m]$,
(iii) $|F_i \cap F_j| = \gamma(\{i, j\})$ for $\{i, j\} \in W$,
(iv) $|F_h \cap F_i \cap F_j| = 0$ for all three distinct elements $h, i, j \in [m]$.

Example III.2. *Let $G(\alpha, \beta, \gamma; 7, 3, 3)$ denote the graph with vertex set* [3] *and empty edge set, with $(k, r, \delta) = (7, 3, 3)$, $\alpha(3) = 1$ and all other values of α and β constantly zero.*

By Theorem III.5 and (15)*, this corresponds to a* $(14, 7, 4, 3, 3)$*-matroid* $M(F_1, F_2, F_3; k; \rho)$*, where* F_1*,* F_2 *and* F_3 *are pairwise disjoint sets with* $|F_1| = |F_2| = 5$*,* $|F_3| = 4$*,* $\rho(F_1) = \rho(F_2) = 3$ *and* $\rho(F_3) = 2$*.*

The graph construction 2: Let $G = G(\gamma; k, r, \delta, a, b)$ be a graph with vertices $[m]$ and edges W, $\gamma : [m] \to \mathbb{Z}$, and let (k, r, δ, a, b) be five integers with $0 < r < k$, $\delta \geq 2$, $0 \leq a < r$ and $0 \leq b < r + \delta - 1$ such that

(16)

(i) G is a graph with no l-cycles, for $l \leq \max\{3, \lceil \frac{k}{r} \rceil\}$,
(ii) $\gamma(w) \geq 1$ for $w \in W$,
(iii) $\sum_{w \in W} \gamma(w) = b$,
(iv) $k \leq rm - b$,
(v) For $I \subseteq [m]$ with $|I| = \lceil \frac{k}{r} \rceil$, $\sum_{w \in W \cap I \times I} \gamma(w) \leq a$,
(vi) $r > \sum_j \gamma(\{i, j\})$ for $i \in [m]$.

Theorem III.6. *Let* $G(\gamma; k, r, \delta, a, b)$ *be a graph on* $[m]$ *such that the conditions (i)–(vi) given in* (16) *are satisfied. Then there is an* (n, k, d, r, δ)*-matroid* $M(F_1, \ldots, F_m; k; \rho)$ *given in Corollary III.1 with*

(i) $n = (r + \delta - 1)m - b$,
(ii) $k = \lceil \frac{k}{r} \rceil r - a$
(iii) $d = n - k + 1 - (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$.

Proof. The proof is given in the Appendix. □

The construction of the sets $F_1, \ldots, F_m$ and the rank function ρ proceeds by first assigning the folowing:

(17)

(i) $\rho(F_i) = r$ for $i \in [m]$,
(ii) $|F_i| = r + \delta - 1$ for $i \in [m]$,
(iii) $|F_i \cap F_j| = \gamma(\{i, j\})$ for $\{i, j\} \in W$,
(iv) $|F_h \cap F_i \cap F_j| = 0$ for all three distinct elements $h, i, j \in [m]$.

Corollary III.2. *Let* $G = G(\gamma; k, r, \delta, a, b)$ *be a graph with* $b > a$ *such that the conditions (i), (iii) and (vi) in* (16) *are satisfied and* $1 \leq \gamma(w) \leq \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$ *for* $w \in W$*. Then* G *defines a matroid* $M(F_i, \ldots, F_m; k; \rho)$ *given in Theorem III.6.*

Proof. To prove this corollary we will prove that the conditions (ii), (iv) and (v) in (16) are satisfied by G. Condition (ii) follows directly. Since G has no l-cycles for $l \leq \lceil \frac{k}{r} \rceil$ and $\gamma(w) \leq \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$ we obtain for $I \subseteq [m]$ with $|I| = \lceil \frac{k}{r} \rceil$ that

$$\sum_{w= \in W \cap I \times I} \gamma(w) \leq \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) \leq a.$$

Hence condition (v) holds. Now, properties (iii) and (v) implies that (iv) holds for G as $b > a$. □

We remark that in order to find as small n as possible for a chosen (k, r, δ, a, b) in Corollary III.2, we want to find a good graph with as few nodes as possible. To find such a graph, preferable properties for the graph are: many small cycles of length $\max\{4, \lceil \frac{k}{r} \rceil + 1\}$, large values of γ on every edge, *i.e.* $\gamma(w) = \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$ for $w \in W$, and that the sum of γ-values incident to each node is large, *i.e.* $\sum_j \gamma(\{i, j\}) = r - 1$ for all nodes $i \in [m]$).

Example III.3. *Let* $G = G(\gamma; 14, 4, 2, 2, 3)$ *denote the graph below on the vertex set* $[6]$*, where the values of* γ *are written above the edges in the graph.*

Fig. 4. The graph $G(\gamma; 14, 4, 2, 2, 3)$

By Corollary III.2 and (17)*, this graph corresponds to an* (n, k, d, r, δ)*-matroid* $M(F_1, \ldots, F_6; k; \rho)$ *with*

$$(n, k, d, r, \delta) = (27, 14, 11, 4, 2),$$
$$F_1 = \{1, \ldots, 5\},\ F_2 = \{5, \ldots, 9\},\ F_3 = \{9, \ldots, 13\},$$
$$F_4 = \{14, \ldots, 18\},\ F_4 = \{18, \ldots, 22\} \text{ and } F_5 = \{23, \ldots, 27\},$$
$$\rho(F_1) = \rho(F_2) = \rho(F_3) = \rho(F_4) = \rho(F_5) = \rho(F_6) = 5.$$

Example III.4. *Let* $G = G(\gamma; k, r, \delta, a, b)$ *denote the graph below on the vertex set* $[11]$*, where* G *consists of one subgraph* B*, one subgraph* $B'(5)$ *and a single vertex with no edges. The* γ*-values for the edges are written in the graph and* $(k, r, \delta, a, b) = (19, 9, 5, 8, 21)$.

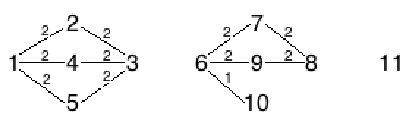

Fig. 5. The graph $G(\gamma; 19, 9, 5, 8, 21)$

By Corollary III.2, this graph corresponds to an (n, k, d, r, δ)-matroid $M(F_1, \ldots, F_6; k; \rho)$ with

$$(n, k, d, r, \delta) = (122, 19, 96, 9, 5).$$

3) The maximal value of d when $r = k$: A well known class of matroids is the class of uniform matroids [35]. A uniform matroid U_n^k can be characterized as a matroid on E, where $|E| = n$ and a subset $X \subseteq E$ is a circuit if and only if $|X| = k + 1$. This implies that the cyclic sets of U_n^k is

$$\mathcal{U}(U_n^k) = \{\emptyset\} \cup \{X \subseteq E : |X| \geq k + 1\}.$$

In terms of cyclic flats, when $0 < k < n$, the uniform matroid U_n^k is defined by

$$\mathcal{Z} = \{0_{\mathcal{Z}}, 1_{\mathcal{Z}}\}, \text{ where } 0_{\mathcal{Z}} = \emptyset,\ 1_{\mathcal{Z}} = E,\ \rho(0_{\mathcal{Z}}) = 0 \text{ and } \rho(1_{\mathcal{Z}}) = k.$$

Assuming that $M = (\rho, E)$ is a (n, k, d, r, δ)-matroid for which $r = k$ and d achieves the generalized Singleton bound given in Theorem III.2, *i.e.*, $d = n - k + 1$. Then using Lemma III.1 (iii), we get that $\mathcal{Z} = \{\emptyset, E\}$, so M is the uniform matroid U_n^k. For (r, δ)-locality, let $S_x = U_n^k$ for each $x \in E$ and $\delta = d = n - k + 1$. Then $|S_x| = r + (\delta - 1)$ and $(S_x) = \delta$. Consequently, U_n^k is a matroid with parameters $(n, k, d, r, \delta) = (n, k, n - k + 1, r, n - k + 1)$.

4) The maximal value of d when $r < k$: Theorem III.7 gives results on the maximal value of d, for which there exists an (n, k, r, δ)-matroid, for every (n, k, r, δ) with $0 < r < k \leq n - (\delta - 1)\lceil \frac{k}{r} \rceil$ and $\delta \geq 2$. We will denote this maximal value $d_{\max} = d_{\max}(n, k, r, \delta)$. Note that, by Theorem III.2, we always have $d_{\max}(n, k, r, \delta) \leq n - k - (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$. We will use the graph constructions given in Theorem III.5 and Theorem III.6, in order to construct matroids with large d. In the cases when $d_{\max} < n - k - (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$, we will use Theorem III.3 to prove this.

We will let a and b be defined by $a = \lceil \frac{k}{r} \rceil r - k$ and $b = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r + \delta - 1) - n$. Before stating the technical theorem on $d_{\max}$, we need the following qualitative result.

Proposition III.4. *Let M be an (n, k, d, r, δ)-matroid and let $a = \lceil \frac{k}{r} \rceil r - k$ and $b = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r + \delta - 1) - n$. Then the following hold,*

$$\left\lceil \frac{n}{r + \delta - 1} \right\rceil \geq \begin{cases} \lceil \frac{k}{r} \rceil & \text{if} \quad b \leq a, \\ \lceil \frac{k}{r} \rceil + 1 & \text{if} \quad b > a, \end{cases}$$

Proof. Let $\left\lceil \frac{n}{r+\delta-1} \right\rceil = \lceil \frac{k}{r} \rceil + t$. Note that $n - k \geq \lceil \frac{k}{r} \rceil (\delta - 1)$ by Proposition III.2. Hence,

$$\begin{aligned} n - k &= (\lceil \tfrac{k}{r} \rceil + t)(r + \delta - 1) - b - (\lceil \tfrac{k}{r} \rceil r - a) \\ &= \lceil \tfrac{k}{r} \rceil (\delta - 1) + t(r + \delta - 1) - (b - a) \\ &\geq \lceil \tfrac{k}{r} \rceil (\delta - 1). \end{aligned}$$

This implies that $t \geq 0$ if $b \leq a$ and $t \geq 1$ if $b > a$. □

Theorem III.7. *Let (n, k, r, δ) be integers such that $0 < r < k \leq n - \left\lceil \frac{k}{r} \right\rceil (\delta - 1)$, $k = \left\lceil \frac{k}{r} \right\rceil r - a$ and $n = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r + \delta - 1) - b$. Let $d_{\max} = d_{\max}(n, k, r, \delta)$ be the largest d such that there exists an (n, k, d, r, δ)-matroid. Then the following hold.*

(i) If $a \geq b$, then $d_{\max} = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta - 1)$;

(ii) If $b > a$, then

$$d_{\max} \geq \begin{cases} n - k + 1 - \left\lceil \frac{k}{r} \right\rceil (\delta - 1) & \text{if} \quad b \leq r - 1, \\ n - k + 1 - \left\lceil \frac{k}{r} \right\rceil (\delta - 1) + (b - r) & \text{if} \quad b \geq r; \end{cases}$$

(iii) If $b > a$ and $a < \left\lceil \frac{k}{r} \right\rceil - 1$, then
$d_{\max} = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta - 1)$ if and only if $\left\lfloor \left\lceil \frac{k}{r} \right\rceil / 2 \right\rfloor \leq a$ and

$$\left\lceil \frac{n}{r + \delta - 1} \right\rceil \geq \left\lceil \frac{k}{r} \right\rceil - 1 + (b - a)\left(1 + \frac{1}{t}\right),$$

where $t = \left\lfloor a / \left(\left\lceil \frac{k}{r} \right\rceil - 1 - a\right) \right\rfloor$;

(iv) If $b > a \geq \left\lceil \frac{k}{r} \right\rceil - 1$, $\left\lceil \frac{k}{r} \right\rceil \geq 3$ and

$$\left\lceil \frac{n}{r + \delta - 1} \right\rceil \geq \left\lfloor \frac{b}{stu} \right\rfloor (t(u - 1) + 2) + y,$$

where $s = \left\lfloor \frac{a}{\left\lceil \frac{k}{r} \right\rceil - 1} \right\rfloor$, $t = \left\lfloor \frac{r-1}{s} \right\rfloor$, $u = \left\lceil \frac{\left\lceil \frac{k}{r} \right\rceil + 1}{2} \right\rceil$, $x = \left\lceil \frac{b - \left\lfloor \frac{b}{stu} \right\rfloor stu}{s} \right\rceil$, and

$$y = \begin{cases} 0 & \text{if} \quad stu \mid b, \\ x - \left\lfloor \frac{x}{u} \right\rfloor + 1 + \min\{\left\lfloor \frac{x}{u} \right\rfloor, 1\} & \text{if} \quad stu \nmid b, \end{cases}$$

then $d_{\max} = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta - 1)$;

(v) If $b > a \geq \left\lceil \frac{k}{r} \right\rceil - 1$, $\left\lceil \frac{k}{r} \right\rceil = 2$, and

$$\left\lceil \frac{n}{r + \delta - 1} \right\rceil \geq \begin{cases} \left\lceil \frac{b}{a} \right\rceil + 1 & \text{if} \quad 2a \leq r - 1, \\ \left\lceil \frac{b}{\left\lfloor \frac{r-1}{2} \right\rfloor} \right\rceil + 1 & \text{if} \quad 2a > r - 1, \end{cases}$$

then $d_{\max} = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta - 1)$.

Proof. The proof is given in the Appendix. □

In the proof of Theorem III.7(iv), we will notice that a simplier bound, but in general not as good, is $\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \left\lceil \frac{b}{stu} \right\rceil (t(u - 1) + 2)$.

Example III.5. *Examples of constructions of matroids in Theorem III.7(i), (iii) and (iv) given by the proofs of the theorem are given in Example III.2, III.3 and III.4 respectively.*

IV. Applications of (n, k, d, r, δ)-Matroids to (n, k, d, r, δ)-LRCs

In this section we will use the previous results on (n, k, d, r, δ)-matroids to get new results on linear and almost affine (n, k, d, r, δ)-LRCs. In order to do so we will first show that the class of matroids given in Corollary III.1 is a subclass of a class of matroids called gammoids. Gammoids have the property that they are representable over large enough fields.

A. Transversal matroids and gammoids

In this section we will give a short introduction to gammoids and transversal matroids. For more information on these classes of matroids we refer the reader to [35], [38].

Let $G = (V, D, S, T)$ be a directed graph where V is the set of vertices in G, D is the directed edges in G, $S \subseteq V$ and $T \subseteq V$. A *gammoid* is a matroid $M(G)$ on S where the independent sets of $M(G)$ equals

$$\mathcal{I}(M(G)) = \{X \subseteq S : \exists \text{ a set of } |X| \text{ vertex-disjoint paths from } X \text{ to } T\}.$$

Our interest in gammoids in this paper stems from the following result.

Theorem IV.1 ([28])**.** *Every gammoid over a finite set E is representable over every finite field of size greater than or equal to $2^{|E|}$.*

A *transversal matroid* is a gammoid $M(G')$ over a directed graph $G' = (V', D', S', T')$, where S' and T' are disjoint sets, $V' = S' \cup T'$ and the set of directed edges D' is a subset of $\{(\overrightarrow{s, t}) : s \in S' \text{ and } t \in T'\}$. By the Hall's Theorem [35, Thm. 12.2.1] we have for a transversal matroid $M(G')$ that

$$\mathcal{I}(M(G')) = \{X \subseteq S' : |X'| \leq |A(X')| \text{ for each } X' \subseteq X\}, \tag{18}$$

where $A(X') = \{t \in T : \exists x' \in X' \text{ such that } (\overrightarrow{x', t}) \in D'\}$.

B. Constructions of linear (n, k, d, r, δ)-LRCs $C(F_1, \dots, F_m; k; \rho)$

The main result in this subsection is Theorem IV.2.

Theorem IV.2. *Let $M(F_1, \dots, F_m; k; \rho)$ be an (n, k, d, r, δ)-matroid that we get in Corollary III.1. Then for every large enough finite field there is a linear LRC over the field with parameters (n, k, d, r, δ).*

This theorem follows immediately from Lemma IV.2 and Theorem IV.1. The key element is the construction of a directed graph whose associated gammoid is the matroid from Corollary III.1. This construction is detailed in Algorithms 1 and 2.

Algorithm 1 From $(F_1, \ldots, F_m; \rho)$ to $G(F_1, \ldots, F_m; \rho) = (V, D, S, T = H)$

$S = E$, $H = \emptyset$, $D = \emptyset$
s is a function from S to $2^{[m]}$
h is a function from H to $2^{[m]}$
for all $e \in E$ **do**
    $s(e) = \{i : e \in F_i\}$
for all $e \in E$ **do**
    if $|s(e)| \geq 2$ **then**
        $H \leftarrow H \cup u_e$
        $h(u_e) = s(e)$
for all $i \in [m]$ **do**
    $l_i = \rho(F_i) - |\{u \in H : i \in h(u)\}|$
    $H \leftarrow H \cup \{v_1^i, \ldots, v_{l_i}^i\}$
    $h(v_1^i) = \ldots = h(v_{l_i}^i) = \{i\}$
for all $(e, u) \in S \times H$ **do**
    if $s(e) \subseteq h(u)$ **then**
        $D \leftarrow D \cup (\overrightarrow{e, u})$
$V = S \cup H$
$T = H$
$G(F_1, \ldots, F_m; \rho) = (V, D, S, T)$

Lemma IV.1. *Let $F_1, \ldots, F_m$ be a collection of subsets of a finite set E and $\rho : \{F_i\}_{i \in [m]} \to \mathbb{Z}$ a function such that the statements (i), (ii) and (iv) in Corollary III.1 are satisfied. Then the transversal matroid $M(G(F_1, \ldots, F_m; \rho))$, that we get from Algorithm 1 is equal to the matroid $M(F_1, \ldots, F_m; k; \rho)$ with $k = |F_{[m]}| - \sum_{i \in [m]} \eta(F_i)$ that we get in Corollary III.1.*

Proof. The proof is given in the Appendix □

Algorithm 2 From $(F_1, \dots, F_m; k; \rho)$ to $G(F_1, \dots, F_m; k; \rho) = (V, D, S, T)$

Get the graph $G'(F_1, \dots, F_m; \rho) = (V', D', S', T')$ *from Algorithm 1*

$H = T'$, $S = S'$, $D = D'$

$T = \{t_1, \dots, t_k\}$

$V = V' \cup T$

for all $(u, t) \in H \times T$ **do**

    $D \leftarrow D \cup (\overrightarrow{u, t})$

$G(F_1, \dots, F_m; k; \rho) = (V, D, S, T)$

Lemma IV.2. *Let* $F_1, \dots, F_m$ *be a collection of subsets of a finite set* E, k *a positive integer and* $\rho : \{F_i\}_{i \in [m]} \to \mathbb{Z}$ *a function such that the statements (i)–(iv) in Corollary III.1 are satisfied. Then the gammoid* $M(G(F_1, \dots, F_m; k; \rho))$*, that we get from Algorithm 2, is equal to the matroid* $M(F_1, \dots, F_m; k; \rho)$ *that we get in Corollary III.1.*

Proof. The proof is given in the Appendix □

Example IV.1. *Consider the matroid* M_G*, associated to the storage system in Figure 1 and the code in Example II.1, and whose lattice of cyclic flats is written out in Example II.2. By Lemma IV.2, this is the gammoid associated to the following graph, with* $|T| = k = 6$ *and* $|S| = n = 12$*. Note that in this particular setting, Algorithm 2 is superfluous, since* H *already has only* 6 *nodes. Indeed, the inclusion of the bipartite graph* (H, T) *corresponds to truncating the gammoid at rank* k*.*

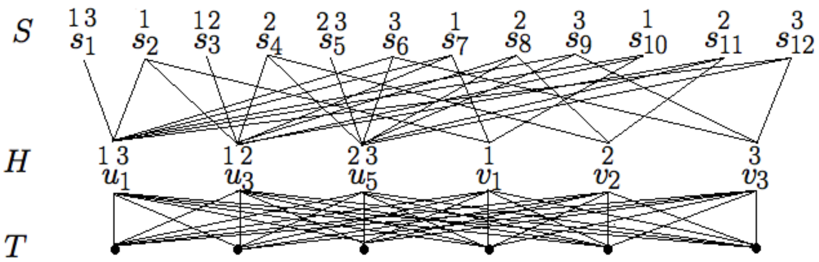

Fig. 6. A storage system from a $(12, 6, 4, 3, 3)$-LRC

C. Bounds on the parameters (n, k, d, r, δ) *for LRCs*

In this section we will give results on the parameters (n, k, d, r, δ) for linear, and more generally almost affine LRCs. The results are direct consequences from the fact that the parameters (n, k, d, r, δ)

are matroid invariants as in Theorem III.1, the representability results from Theorem IV.2), and previously proven results on (n, k, d, r, δ)-matroids. Before we give the results we recall that all linear LRCs also are almost affine LRCs. Hence, Theorem IV.3-IV.4 and Proposition IV.1 also hold for linear LRCs, and Theorem IV.5 also holds for almost affine LRCs.

The following theorem is the generalized Singleton bound for almost affine codes.

Theorem IV.3. *If C is an almost affine LRCs with the parameters (n, k, d, r, δ), then*

$$d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1).$$

Proof. This is an immediate consequence of Theorem III.1 and Theorem III.2. □

Next, we present some bounds on δ, k and the rate $\frac{k}{n}$ of almost affine LRCs.

Proposition IV.1. *Let C be an almost affine LRC with parameters (n, k, d, r, δ). Then*

(i) $\delta \leq d$,
(ii) $k \leq n - \left\lceil \frac{k}{r} \right\rceil (\delta - 1)$,
(iii) $\frac{k}{n} \leq \frac{r}{r+\delta-1}$.

Proof. These statements are immediate consequences of Theorem III.1 and Proposition III.2. □

The next theorem gives a nonexistence result for almost affine LRCs attending the generalized Singleton bound for a certain class of parameters (n, k, r, δ).

Theorem IV.4. *Let C be an almost affine LRC with parameters (n, k, r, d, δ), and let $a = \left\lceil \frac{k}{r} \right\rceil r - k$ and $b = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r + \delta - 1) - n$. Then the following hold.*

(i) If $b > a$ and $a < \left\lfloor \left\lceil \frac{k}{r} \right\rceil /2 \right\rfloor$, then $d < n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1)$;

(ii) If $b > a$ and $\left\lfloor \left\lceil \frac{k}{r} \right\rceil /2 \right\rfloor \leq a \leq \left\lceil \frac{k}{r} \right\rceil - 1$, then $d < n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1)$.

Proof. The theorem follows from Theorem III.1 and Theorem III.7(iii). □

Theorem IV.5. *Let (n, k, r, δ) be integers such that $0 < r < k \leq n - \left\lceil \frac{k}{r} \right\rceil (\delta - 1)$, $a = \left\lceil \frac{k}{r} \right\rceil r - k$ and $b = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r + \delta - 1) - n$. Let $d_{\max} = d_{\max}(n, k, r, \delta)$ be the largest d such that there exists a linear LRC with parameters (n, k, d, r, δ). Then the following hold.*

(i) If $a \geq b$, then $d_{\max} = n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1)$;

(ii) If $b > a$*, then*

$$d_{\max} \geq \begin{cases} n-k+1-\left\lceil \frac{k}{r} \right\rceil (\delta-1) & \text{if} \quad b \leq r-1, \\ n-k+1-\left\lceil \frac{k}{r} \right\rceil (\delta-1)+(b-r) & \text{if} \quad b \geq r; \end{cases}$$

(iii) If $b > a$, $\left\lfloor \left\lceil \frac{k}{r} \right\rceil /2 \right\rfloor \leq a < \left\lceil \frac{k}{r} \right\rceil - 1$ *and* $\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \left\lceil \frac{k}{r} \right\rceil - 1 + (b-a)\left(1+\frac{1}{t}\right)$*, where* $t = \left\lfloor a/\left(\left\lceil \frac{k}{r} \right\rceil - 1 - a\right) \right\rfloor$*, then*

$$d_{\max} = n-k+1-\left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta-1);$$

(iv) If $b > a \geq \left\lceil \frac{k}{r} \right\rceil - 1$, $\left\lceil \frac{k}{r} \right\rceil \geq 3$ *and* $\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \left\lfloor \frac{b}{stu} \right\rfloor (t(u-1)+2)+y$*,*
where $s = \left\lfloor \frac{a}{\left\lceil \frac{k}{r} \right\rceil - 1} \right\rfloor$, $t = \left\lfloor \frac{r-1}{s} \right\rfloor$, $u = \left\lceil \frac{\left\lceil \frac{k}{r} \right\rceil + 1}{2} \right\rceil$, $x = \left\lceil \frac{b - \left\lfloor \frac{b}{stu} \right\rfloor stu}{s} \right\rceil$ *and*

$$y = \begin{cases} 0 & \text{if} \quad stu \mid b, \\ x - \left\lfloor \frac{x}{u} \right\rfloor + 1 + \min\{\left\lfloor \frac{x}{u} \right\rfloor, 1\} & \text{if} \quad stu \nmid b, \end{cases}$$

then

$$d_{\max} = n-k+1-\left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta-1);$$

(v) If $b > a \geq \left\lceil \frac{k}{r} \right\rceil - 1$, $\left\lceil \frac{k}{r} \right\rceil = 2$ *and*

$$\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \begin{cases} \left\lceil \frac{b}{a} \right\rceil + 1 & \text{if} \quad 2a \leq r-1, \\ \left\lceil \frac{b}{\left\lfloor \frac{r-1}{2} \right\rfloor} \right\rceil + 1 & \text{if} \quad 2a > r-1, \end{cases}$$

then

$$d_{\max} = n-k+1-\left(\left\lceil \frac{k}{r} \right\rceil - 1\right)(\delta-1).$$

Proof. The corresponding (n,k,d,r,δ)-matroids in Theorem III.7 are constructed using Theorem III.5 and Corollary III.2. Hence these (n,k,d,r,δ)-matroids is contained in the class of matroids given in Corollary III.1. The theorem now follows from Theorem IV.2. □

We observe from the the remark below Theorem III.7 that a simplier bound, but in general not as good, in Theorem IV.5(iv) is $\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \left\lceil \frac{b}{stu} \right\rceil (t(u-1)+2)$.

D. Constructions of linear LRCs

We will now explicitly construct linear LRCs using the results in this paper. We will give four different constructions, of decreasing generality, increasingly specialized to give good values of d. By Constructions 3 and 4 we are only able to construct linear LRCs for parameter values where the generalized Singleton bound can be achieved. By Constructions 1 and 2 we are also able to construct linear LRCs for other parameter values.

Construction 1: (A linear LRC from Corollary III.1)

The following construction will give a linear LRCs with the same parameters (n, k, d, r, δ) as the matroid constructed in Corollary III.1:

(i) Let $F_1, \ldots, F_m$ be a collection of subsets of a finite set E, k a positive integer and $\rho : \{F_i\}_{i \in [m]} \to \mathbb{Z}$ a function, such that the conditions (i)–(iv) in Corollary III.1 hold. This gives a matroid $M(F_1, \ldots, F_m; k; \rho)$.

(ii) Use Theorem III.4 to get the values on the parameters (n, d, r, δ).

(iii) Use $(F_1, \ldots, F_m; k; \rho)$ in Algorithm 2 to get the graph $G(F_1, \ldots, F_m; k; \rho)$.

(iv) Use the methods in [28] to get a linear code C from the gammoid $M(G(F_1, \ldots, F_m; k; \rho))$. The linear code C is a linear (n, k, d, r, δ)-LRCs with repair sets $F_1, \ldots, F_m$.

Construction 2: (A linear LRC from Theorem III.5)

The following construction will give a linear LRCs with the same parameters (n, k, d, r, δ) as the graph constructed in Theorem III.5:

(i) Let G be a graph with vertices $[m]$ and edges W, α and β be two functions $[m] \to \mathbb{Z}$, γ be a function $W \to \mathbb{Z}$ and (k, r, δ) be three integers with $0 < r < k$ and $\delta \geq 2$, such that the conditions (i)–(vi) in (14) hold. This gives a graph $G = G(\alpha, \beta, \gamma; k, r, \delta)$.

(ii) Use Theorem III.5 to get the values on the parameters (n, d).

(iii) Use (15) to get subsets $(F_1, \ldots, F_m)$ and a function $\rho : [m] \to \mathbb{Z}$.

(iv) Use $(F_1, \ldots, F_m; k; \rho)$ in Algorithm 2 to get the graph $G(F_1, \ldots, F_m; k; \rho)$.

(v) Use the methods in [28] to get a linear code C from the gammoid $M(G(F_1, \ldots, F_m; k; \rho))$. The linear code C is a linear (n, k, d, r, δ)-LRCs with repair sets $F_1, \ldots, F_m$.

Construction 3: (A linear LRC with $d = n - k + 1 - (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$ from Theorem III.6)

The following construction will give a linear LRCs with the same parameters (n, k, d, r, δ) as the graph constructed in Theorem III.6:

(i) Let $G = G(\gamma; k, r, \delta, a, b)$ be a graph with vertices $[m]$ and edges W, γ be a function $W \to \mathbb{Z}$, and (k, r, δ, a, b) be integers with $0 < r < k$, $\delta \geq 2$, $0 \leq a < r$ and $0 \leq b < r + \delta - 1$, such that the conditions (i)–(vi) in (16) hold.

(ii) Use (17) in order to get subsets $(F_1, \ldots, F_m)$ and a function $\rho : [m] \to \mathbb{Z}$.

(iii) Use $(F_1, \ldots, F_m; k; \rho)$ in Algorithm 2 to get the graph $G(F_1, \ldots, F_m; k; \rho)$.

(iv) Use the methods in [28] to get a linear code C from the gammoid $M(G(F_1,\dots,F_m;k;\rho))$. The linear code C is a linear (n,k,d,r,δ)-LRCs with $d=n-k+1-(\lceil\frac{k}{r}\rceil-1)(\delta-1)$, $k=\lceil\frac{k}{r}\rceil r-a$ and $n=(r+\delta-1)m-b$ and repair sets $F_1,\dots,F_m$.

Construction 4: (A linear LRC with $d=n-k+1-(\lceil\frac{k}{r}\rceil-1)(\delta-1)$ from Corollary III.2)

The following construction will give a linear LRCs with the same parameters (n,k,d,r,δ) as the graph constructed in Corollary III.2:

(i) Let $G=G(\gamma;k,r,\delta,a,b)$ be a graph with vertices $[m]$ and edges W, γ be a function $W\to\mathbb{Z}$, and let (k,r,δ,a,b) be integers with $0<r<k$, $\delta\geq 2$, $0\leq a<r$, $0\leq b<r+\delta-1$ and $b>a$, such that the conditions in Corollary (III.2) hold.

(ii) Use (17) in order to get subsets $(F_1,\dots,F_m)$ and a function $\rho:[m]\to\mathbb{Z}$.

(iii) Use $(F_1,\dots,F_m;k;\rho)$ in Algorithm 2 to get the graph $G(F_1,\dots,F_m;k;\rho)$.

(iv) Use the methods in [28] to get a linear code C from the gammoid $M(G(F_1,\dots,F_m;k;\rho))$. The linear code C is a linear (n,k,d,r,δ)-LRCs with $d=n-k+1-(\lceil\frac{k}{r}\rceil-1)(\delta-1)$, $k=\lceil\frac{k}{r}\rceil r-a$ and $n=(r+\delta-1)m-b$, and repair sets $F_1,\dots,F_m$.

V. CONCLUSIONS

Recent progress in coding theory has proven matroid theory a valuable tool in many different contexts. This trend carries over to locally repairable codes too as we have seen in this paper. Especially the lattice of cyclic flats turned out useful in our quest for optimal locally repairable codes and related bounds.

We have thoroughly studied linear and more generally almost affine LRCs with all-symbol locality, as well as the connections of these codes to matroid theory. By generalizing the concepts related to the parameters (n,k,d,r,δ) of almost affine LRCs to matroids, we have derived a generalized Singleton bound for matroids and nonexistence results for certain classes of (n,k,d,r,δ)-matroids. These results can then be directly translated to nonexistence results for almost affine LRCs.

Further, we have given several constructions of matroids with certain values of the parameters (n,k,d,r,δ). Using these matroid constructions, novel constructions of linear LRCs are given based on either some finite sets and a function on these sets, or a graph with functions on the vertices and the edges. Several classes of optimal linear LRCs then arise from these constructions.

As future work, (non)existence results for matroids and linear and almost affine LRCs achieving the generalized Singleton bound remain open for certain classes of parameters (n,k,r,δ), when $\lceil\frac{k}{r}\rceil-1\leq$

$a < b$. In addition, the size of the underlying finite field that our linear (n, k, d, r, δ)-LRCs can be constructed over is left for future research. We expect that the upper bound 2^n arising from the related bound for all gammoids is loose for our class of matroids. We conjecture that the bound for the linear (n, k, d, r, δ)-LRCs constructed by any of the four methods given in Section IV-D is approximately n.

References

[1] T. Westerbäck, T. Ernvall, and C. Hollanti, "Almost affine locally repairable codes and matroid theory," *2014 IEEE Inf. Theory Workshop (ITW)*, 2014.

[2] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, 56(9), pp. 4539–4551, September 2010.

[3] S. Ghemawat, H. Gobioff, and S. T. Leung, "The Google file system,“ In *SOSP03, Proceedings of the nineteenth ACM symposium on Operating systems principles*, 2003.

[4] T. Ernvall, S. El Rouayheb, C. Hollanti, and H. V. Poor, "Capacity and security of heterogeneous distributed storage systems," *IEEE J. Sel. Areas Comm.*, 31(12), pp. 2701–2709, Dec. 2013.

[5] B. Sasidharan, and P. V. Kumar, "High-rate regenerating codes through layering," *2013 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1611–1615, 2013.

[6] C. Tian, V. Aggarwal, and V. A. Vaishampayan, "Exact-repair regenerating codes via layered erasure correction and block designs," *2013 IEEE Int. Symp. Inf. Theory (ISIT)* , pp. 1431–1435, 2013.

[7] T. Ernvall, "Codes between MBR and MSR points with exact repair property," *IEEE Trans. Inf. Theory*, 60(11), pp. 6993–7005, Nov. 2014.

[8] S. Goparaju, S. El Rouayheb, and R. Calderbank, "New codes and inner bounds for exact repair in distributed storage systems," *2014 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1036–1040, 2014.

[9] I. Tamo, Z. Wang, and J. Bruck, "MDS array codes with optimal rebuilding," *2011 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1240–1244, 2011.

[10] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, 58(11), pp. 6925–6934, 2011.

[11] F. Oggier and A. Datta, "Self-repairing homomorphic codes for distributed storage systems," *2011 IEEE INFOCOM*, pp. 1215–1223.

[12] D. S. Papailiopoulos, J. Luo, A. G. Dimakis, and C. Huang, J. Li "Simple regenerating codes: Network coding for cloud storage," *2012 IEEE INFOCOM*, pp. 2801–2805.

[13] D. S. Papailiopoulos, and A. G. Dimakis, "Locally repairable codes," *2012 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2771–2775.

[14] I. Tamo, D. S. Papailiopoulos, and A. G. Dimakis, "Optimal locally repairable codes and connections to matroid theory," *2013 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1814–1818.

[15] J. Simonis and A. Ashikhmin, "Almost affine codes", *Design, codes and cryptography*, 14, pp. 179–197, 1998.

[16] R. C. Singleton, "Maximum distance q-nary codes", *IEEE Trans. Inf. Theory*, 10, pp. 116–118, 1964.

[17] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, “Optimal linear codes with a local-error-correction property,” *2012 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2776 – 2780, 2012.

[18] V. Cadambe and A. Mazumdar, “An upper bound on the size of locally recoverable codes”, In *Proc. IEEE Symp. Netw. Coding*, pp. 1–5, Jun. 2013.

[19] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, “Optimal locally repairable and secure codes for distributed storage systems”, *IEEE Trans. Inf. Theory*, 60(1), pp. 212–236, 2014.

[20] N. Silberstein, A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, “Optimal locally repairable codes via rank-metric codes,” *2013 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1819–1823.

[21] W. Song, S. H. Dau, C. Yuen, and T. J. Li, “Optimal locally repairable linear codes”,*IEEE Sel. Areas Commun.,* 32(5), pp. 1019–1036, 2014.

[22] I. Tamo and A. Barg, “A family of optimal locally recoverable codes” *IEEE Trans. Inf. Theory*, 60(8), pp. 4661–4676, 2014.

[23] A. Barg, “The matroid supports of a linear code”, *Appl. Algebra Engrg. Comm. Comput.*, 8, pp. 165–172, 1997.

[24] T. Britz, “Code enumerators and Tutte polynomials”, *IEEE Trans. Inf. Theory*, 56, pp. 4350–4358, 2010.

[25] T. Britz and C. G. Rutherford, “Covering radii are not matroid invariants”, *Discrete Math.*, 296, pp. 117–120, 2005.

[26] T. Britz, T. Johnsen, D. Mayhew, and K. Shiromoto, “Wei-type duality theorems for matroids”, *Designs, Codes and Cryptography*, 62, pp. 331–341, 2012.

[27] T. Ernvall, T. Westerbäck, C. Hollanti, and R. Freij, “Constructions and properties of linear locally repairable codes,” submitted, arXiv:1410.6339.

[28] B. Lindström, “On the vector representations of induced matroids” *Bull. London Math. Soc.*, 5, pp. 85–90, 1973.

[29] R. Stanley “Enumerative combinatorics, vol 1” *2:ed Cambridge University Press*, 2011.

[30] H. Whitney, ‘On the abstract properties of linear dependence”, *Amer. J. Math*, 57, pp. 509–533, 1935.

[31] R. Dougherty, C. Freiling, and K. Zeger, “Networks, matroids, and non-Shannon information inequalities”, *IEEE Trans. Inf. Theory*, 53(6), pp. 1949–1969, 2007.

[32] J. Martí-Farré and C. Padró, “On secret sharing schemes, matroids and polymatroids”, In S. Vadhan edt., *4th Theory of Crypt. Conf. TCC 2007, Lecture Notes in Computer Science*, vol. 4392, pp. 253–272, 2007.

[33] S. El Rouayheb, A.Sprintson, and C. Georghiades, ‘On the index coding problem and its relation to network coding and matroid theory”, *IEEE Trans. Inf. Theory*, 56(7), pp. 3187–3195, 2010.

[34] F. Matúŝ, “Matroid representation by partitions”, *Discrete Math.*, 203, pp. 169–194, 1999.

[35] J. Oxley, “Matroid Theory” *Oxford Graduate Texts in Mathematics*, 3rd ed., Oxford University Press, 1992.

[36] J. E. Bonin and A. de Mier, “The lattice of cyclic flats of a matroid”, *Annals of Combinatorics*, 12, pp. 155–170, 2008.

[37] K. Shoda, “Large families of matroids with the same Tutte polynomial”, *PhD thesis, The Geroge Washington University*, August 2012.

[38] A. Schrijver, “Combinatorial Optimization: Polyhedra and Efficiency. Vol B: Matroids, Trees, Stable Sets” *Algorithms and Combinatorics 24*, Springer-Verlag, 2003.

[39] T. H. Brylawski, “Constructions”, *Theory of Matroids* (edited by N. White), Cambridge University Press, pp. 127–223, 1986.

APPENDIX

Proof of Lemma III.1. Statements (i) and (ii) are obvious from the definitions of n and k in Definition III.1.

For statement (iii), note that $\rho^*(0_{\mathcal{Z}^*}) = 0$, and $\rho^*(X) < \rho^*(Y)$ if $X \subsetneq Y$ and $X, Y \in \mathcal{Z}^*$, by Theorem II.1. Hence, by Proposition II.1 (iii) and the fact that $0_{\mathcal{Z}^*} = E \setminus 1_{\mathcal{Z}} = \emptyset$, it follows that

$$d = \min\{\rho^*(X) + 1 : X \in A_{\mathcal{Z}^*}\}. \tag{19}$$

Equation (8) implies that

$$\begin{aligned} \rho^*(X) + 1 &= \rho(E \setminus X) + |X| - \rho(E) + 1 \\ &= |X| + |E \setminus X| - \rho(E) - |E \setminus X| + \\ &\quad\ \rho(E \setminus X) + 1 \\ &= n - k + 1 - (|E \setminus X| - \rho(E \setminus X)), \end{aligned} \tag{20}$$

for $X \in A_{\mathcal{Z}^*}$. By proposition II.1 (i), we observe that

$$coA_{\mathcal{Z}} = \{E \setminus X : X \in A_{\mathcal{Z}^*}\}. \tag{21}$$

Condition (iii) now follows from (19), (20) and (21).

For statement (iv), assume that we have a collection of sets $\{S_x \subseteq E\}_{x \in E}$ for which (a)–(c) in Definition III.1 (iv) are satisfied. Then it follows immediately that statements (a) and (b) in this lemma are satisfied. Since $d(M|S_x) \geq \delta \geq 2$, it follows that S_x is a cyclic set by Proposition II.2 (viii). Statement (c) in this lemma now follows from statement (iii) in this lemma and Definition III.1 (iv)(c).

Now, assume that the statements (a)–(c) in this lemma are satisfied by a collection of cyclic sets $\{S_x\}_{x \in E}$. Statement (a) and (b) in Definition III.1 (iv) follow directly from statement (a) and (b) in this lemma. Condition (c) in Definition III.1 (iv) follows from the fact that $1_{M|S_x} = S_x$ for each $x \in E$ and statement (iii) in this lemma. $\square$

Proof of Lemma III.2. First assume that $\{S_j\}_{j \in [m]}$ is a collection of cyclic sets from $\{S_x\}_{x \in E}$. Let j be any integer in$[m]$. We first observe that $\mathrm{cl}(S_j)$ is a cyclic flat, by Proposition II.2 (iv). Hence, by Proposition II.1 (iv), if Y_{j-1} is a cyclic flat then Y_j is a cyclic flat with

$$Y_j = \mathrm{cl}(Y_{j-1} \cup S_j) = \mathrm{cl}(Y_{j-1} \cup \mathrm{cl}(S_j)) = Y_{j-1} \vee \mathrm{cl}(S_j). \tag{22}$$

For $j = 1, \ldots, m$ it follows that Y_j is a cyclic flat and that (22) holds, as Y_0 is a cyclic flat.

For (i), let S_1 be a cyclic flat of $\{S_x\}_{x\in E}$. Note that $\rho(S_1) > 0$, because if $\rho(S_1) = 0$, then the remark in the paragraph above Proposition III.1 implies that $d(M|S_1)$ is not defined. Thus, as $\mathrm{cl}(S_1) \in \mathcal{Z}(M)$ and $\rho(0_{\mathcal{Z}}) = 0 < \rho(\mathrm{cl}(S_1))$, we obtain that

$$Y_0 = 0_{\mathcal{Z}} \subsetneq \mathrm{cl}(S_1).$$

Now, using that $\bigcup_{x\in E} S_x = E$, we obtain that there is a subset $\{S_1, \ldots, S_m\}$ of $\{S_x : x \in E\}$ such that $Y_m = E$ and $S_j \nsubseteq Y_{j-1}$ for $j = 1, \ldots, m$. Consequently,

$$Y_{j-1} \subsetneq \mathrm{cl}(Y_{j-1} \cup S_j) = Y_j.$$

For (ii), we remark as in (12) that $\rho(S_j) \le r$ for any $j \in [m]$. Hence, by axiom (R3) in (6),

$$\rho(Y_j) = \rho(\mathrm{cl}(Y_{j-1} \cup S_j)) = \rho(Y_{j-1} \cup S_j) \le \rho(Y_{j-1}) + \rho(S_j) - \rho(Y_{j-1} \cap S_j) \le \rho(Y_{j-1}) + r.$$

For (iii), the fact that $Y_{j-1} \subsetneq Y_j$ implies that $\rho(Y_{j-1}) < \rho(Y_j)$, as Y_{j-1} and Y_j are flats. Then, by Axiom (R3) in (6),

$$\rho(S_j \cap Y_{j-1}) \le \rho(S_j) + \rho(Y_{j-1}) - \rho(S_j \cup Y_{j-1}) = \rho(S_j) + \rho(Y_{j-1}) - \rho(Y_j) < \rho(S_j).$$

Hence, by the statements (i) - (iii) in Section II.2 and LemmaIII.1 (c),

$$\begin{aligned}
\eta(Y_j) &= \eta(\mathrm{cl}(Y_{j-1} \cup S_j)) \\
&\ge \eta(Y_{j-1} \cup S_j) \\
&\ge \eta(Y_{j-1}) + \eta(S_j) - \eta(Y_{j-1} \cap S_j) \\
&\ge \eta(Y_{j-1}) + \eta(S_j) - \max\{\eta(X) : X \in coA_{\mathcal{Z}(M|S_j)}\} \\
&= \eta(Y_{j-1}) + d(M|S_j) - 1 \\
&\ge \eta(Y_{j-1}) + \delta - 1.
\end{aligned}$$

□

Proof of Lemma III.3. As $(\lceil \frac{k}{r} \rceil - 1)r < k = \rho(Y_m)$, it follows from Lemma III.2 (ii) that

$$m - 1 \ge \left\lceil \frac{k}{r} \right\rceil - 1.$$

Moreover, since $\rho(Y_{m-1}) < \rho(Y_m) = \rho(E)$, we have that

$$\eta(Y_{m-1}) \le \max\{\eta(Z) : Z \in coA_{\mathcal{Z}}\},$$

by Proposition II.2 (iii). Hence, by Lemma III.1 (iii),

$$d = n - k + 1 - \max\{\eta(Z) : \eta(Z) \in coA_{\mathcal{Z}}\} \le n - k + 1 - \eta(Y_{m-1}).$$

$\square$

Proof of Theorem III.3. Let

$$C : 0_{\mathcal{Z}} = Y_0 \subsetneq Y_1 \subsetneq \ldots \subsetneq Y_m = E, \tag{23}$$

be a chain of $(\mathcal{Z}(M), \subseteq)$ as given in Lemma III.2 (i), from a subset $\{S_j\}_{j\in[m]}$ of $\{S_x\}_{x\in E}$. Since d achieves the generalized Singleton bound in Theorem III.2, we get that $m = \lceil \frac{k}{r} \rceil$ and $\eta(Y_{m-1}) \leq (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$. Hence,

$$\eta(Y_j) = j(\delta - 1) \text{ for } j = 0, 1, \ldots, m-1, \quad \text{where} \quad m = \left\lceil \frac{k}{r} \right\rceil \geq 2, \tag{24}$$

by Lemma III.2 (iii) and Lemma III.3.

For (i), observe that $\eta(0_{\mathcal{Z}}) = |0_{\mathcal{Z}}|$. Hence, if $0_{\mathcal{Z}} \neq \emptyset$, then $\eta(Y_0) = \eta(0_{\mathcal{Z}}) > 0$. This is a contradiction by (24).

To prove (ii), first observe that for any S_x we can get a chain, as in (23), with $S_1 = S_x$.

For (a), choose $x \in E$ and let $S_1 = S_x$ in (23). By (24) and the property that $\eta(X) \leq \eta(\mathrm{cl}(X))$ for any $X \subseteq E$, we obtain that

$$\delta - 1 = \eta(Y_1) \geq \eta(S_x).$$

Moreover, as we know from (11),

$$d(M|S_x) \leq \delta \iff \min\{|X| : X \subseteq S_x, \, \rho(S_x \setminus X) < \rho(S_x)\} \geq \delta,$$

which implies that $\eta(S_x) \geq \delta - 1$.

For (b), choose $x \in E$ and let $S_1 = S_x$ in (23). If S_x is not a cyclic flat then

$$\eta(Y_1) = \eta(\mathrm{cl}(S_x)) > \eta(S_x) = \delta - 1.$$

This is a contradiction by (24). Thus,

$$0_{\mathcal{Z}(M|S_x)} = \emptyset, \, 1_{\mathcal{Z}(M|S_x)} = S_x \text{ and } \mathcal{Z}(M|S_x) = \{X \in \mathcal{Z}(M) : X \subseteq S_x\}$$

by Proposition II.2 (vii). Suppose that there is a cyclic flat $Z \in \mathcal{Z}(M)$ such that $\emptyset \subsetneq Z \subsetneq S_x$. Then $\rho(Z) < \rho(S_x)$ and $\eta(Z) > \eta(\emptyset) = 0$ by axiom (Z2) in Theorem II.1. Consequently, by Proposition II.2 (iii) and Lemma III.1 (iv) (c),

$$d(M|S_x) = \eta(S_x) + 1 - \max\{\eta(X) : X \in coA_{\mathcal{Z}(M|S_x)}\} \leq \eta(S_x) + 1 - \eta(Z) \leq \delta - 1.$$

This is a contradiction.

For (iii), we will first prove that any non trivial union $F_1, \ldots, F_m$ of cyclic sets from $\{S_x : x \in E\}$ with $m = \lceil \frac{k}{r} \rceil$ constitutes a chain as given in (23), with $Y_j = Y_{j-1} \vee F_j$ for $j = 1, \ldots, m$. Note by statement (b) in this theorem that F_i is a cyclic flat for $j = 1, \ldots, m$. Also, observe that

$$\emptyset = 0_{\mathcal{Z}} = Y_0 \subsetneq F_1 = Y_0 \vee F_1 = Y_1. \tag{25}$$

Assume that for some $j \leq m-1$ we had

$$Y_l = F_1 \cup \ldots \cup F_l \text{ and } Y_{l-1} \subsetneq Y_l \text{ for } l = 1, \ldots, j. \tag{26}$$

Since $F_1, \ldots, F_m$ is a non trivial union of subsets it would follow that $F_{j+1} \nsubseteq Y_j$. If $|Y_j \cap F_{j+1}| \geq \rho(F_{j+1})$, then we obtain that $\mathrm{cl}(Y_j \cap F_{j+1}) = F_{j+1} \subseteq \mathrm{cl}(Y_j) = Y_j$ by Proposition II.2 (x). This is a contradiction, and consequently

$$|Y_j \cap F_{j+1}| < \rho(F_{j+1}). \tag{27}$$

Now, by (b) and Proposition II.1 (iii), any subset $X \subseteq S_x$ contains a circuit if and only if $|X| > \rho(S_x)$, *i.e.*, $\rho(X) = |X|$ if and only if $|X| \leq \rho(S_x)$. Consequently, $\eta(Y_j \cap F_{j+1}) = 0$. This implies, using Proposition II.2 (ii), (24) and statement (ii), that

$$\eta(Y_j \cup F_{j+1}) \geq \eta(Y_j) + \eta(F_{j+1}) - \eta(Y_j \cap F_{j+1}) = (j+1)(\delta - 1). \tag{28}$$

Furthermore by (24), if $j+1 \leq m-1$, then

$$\eta(Y_j \cup F_{j+1}) \leq \eta(\mathrm{cl}(Y_j \cup F_{j+1})) = \eta(Y_{j+1}) = (j+1)(\delta - 1).$$

Hence, $Y_{j+1} = F_1 \cup \ldots \cup F_{j+1}$ if $j+1 \leq m-1$. If $Y_m \neq E$, then $\rho(Y_m) < \rho(E)$ and $\eta(Y_m) > (\lceil \frac{k}{r} \rceil - 1)(\delta - 1)$. Then it follows, by Proposition II.2 (iii) and Lemma III.1 (iii), that

$$d < n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right)(\delta - 1).$$

This is a contradiction. Consequently, $F_1, \ldots, F_m$ constitutes a chain as given in (23), with

$$Y_j = \bigvee_{i=1}^{j} F_i = \begin{cases} \bigcup_{i=1}^{j} F_i & \text{if} \quad j < \lceil \frac{k}{r} \rceil, \\ E & \text{if} \quad j = m = \lceil \frac{k}{r} \rceil. \end{cases} \tag{29}$$

For statement (c), we first notice that the statement follows directly from (24) when $j < \lceil \frac{k}{r} \rceil$. When $j \geq m = \lceil \frac{k}{r} \rceil$ we conclude, using (29), that $\bigvee_{i=1}^{j} F_i = E$. Also, by (24),

$$n - k \geq \eta(F_1 \cup \ldots \cup F_m) \geq \left\lceil \frac{k}{r} \right\rceil (\delta - 1).$$

Statement (d) follows directly from (29), and statement (e) is a immediately consequence of (c) and (d). Statement (f) follows from (27) and (a). □

Proof of Theorem III.4. We will show that $\mathcal{Z}$ and ρ define a matroid, by proving that the axioms $(Z0)--(Z3)$ in Theorem II.1 are satisfied by $\mathcal{Z}$ and ρ. We let I, J be two subsets of $[m]$.

(Z0) Since the collection of sets $F_1, \dots, F_m$ has a non trivial union by property (i) in III-D1, it follows that $F_I \subsetneq F_J$ if and only if $I \subsetneq J$. Hence, we immediately get that $\mathcal{Z}$ is a lattice under inclusion with

$$F_I \wedge F_J = F_{I \cap J} \quad \text{and} \quad F_I \vee F_J = \begin{cases} F_{I\cup J} & \text{if} \quad F_{I \cup J} \in \mathcal{Z}_{<k}, \\ E & \text{if} \quad F_{I\cup J} \notin \mathcal{Z}_{<k}, \end{cases}$$

for $F_I, F_J \in \mathcal{Z}_{<k}$. Also, the bottom element in the lattice $0_{\mathcal{Z}}$ equals $\emptyset$ and by property (iv) in III-D1 the top element $1_{\mathcal{Z}}$ equals E.

(Z1) Since $0_{\mathcal{Z}} = F_\emptyset$, we obtain that

$$\rho(0_{\mathcal{Z}}) = \rho(F_\emptyset) = 0.$$

(Z2) Assume that $F_I, F_J \in \mathcal{Z}_{<k}$ and that $F_I \subsetneq F_J$. Then $I \subsetneq J$, and by property (v) in III-D1 we obtain that

$$|F_{J\setminus I} \setminus F_I| > |F_{J\setminus I}| - \sum_{j \in J\setminus I} \rho(F_j) = \sum_{j\in J\setminus I} \eta(F_j). \tag{30}$$

Hence,

$$\begin{aligned} \rho(F_J) - \rho(F_I) &= (|F_J| - \textstyle\sum_{j\in J} F_j) - (|F_i| - \sum_{i \in I} \eta(F_i)) \\ &= |F_J \setminus F_I| - \textstyle\sum_{j\in J\setminus I} \eta(F_j) \\ &= |F_{J\setminus I} \setminus F_I| - \textstyle\sum_{j \in J\setminus I} \eta(F_j) \\ &> 0. \end{aligned}$$

Moreover, by property (ii) in III-D1,

$$\begin{aligned} &(|F_J| - \rho(F_J)) - (|F_I| - \rho(F_I)) = \\ &|F_J| - (|F_J| - \textstyle\sum_{j\in J} F_j) - (|F_I| - (|F_I| - \sum_{i\in I} F_i)) = \\ &\textstyle\sum_{j\in J\setminus I} \eta(F_j) > 0. \end{aligned}$$

Since $\mathcal{Z}_{<k}$ contains all cyclic flats in $\mathcal{Z}$ of rank less than k, and $\rho(E) = k$ we immediately get that $\rho(E) - \rho(F_I) > 0$. By properties (iii) and (iv) in III-D1, we notice that there is a set $J \supsetneq I$ and an element $j \in J \setminus I$ such that

$$\rho(I) \leq \rho(J \setminus \{j\}) < \rho(J) = k.$$

Hence, by property (v) in III-D1,

$$(|E| - \rho(E)) - (|F_I| - \rho(F_I)) \geq$$
$$(|F_J| - \rho(F_J)) - (|F_I| - \rho(F_I)) =$$
$$|F_j \setminus F_{J\setminus\{j\}}| - \rho(j) + (|F_{J-\{j\}}| - \rho(F_{J-\{j\}})) - (|F_I| - \rho(F_I)) > 0.$$

(Z3) Suppose that $F_I, F_J \in \mathcal{Z}_{<k}$. Then

$$\rho(F_I) + \rho(F_J) - (\rho(F_I \vee F_J) + \rho(F_I \wedge F_J) + |(F_I \cap F_J) \setminus (F_I \wedge F_J)|) =$$
$$\rho(F_I) + \rho(F_J) - \rho(F_I \vee F_J) - \rho(F_{I\cap J}) - |F_I \cap F_J| + |F_{I\cap J}| =$$
$$|F_I| - \textstyle\sum_{i\in I} \eta(F_i) + |F_J| - \sum_{j\in J} \eta(F_j) -$$
$$|F_{I\cap J}| + \textstyle\sum_{j\in I\cap J} \eta(F_j) - |F_{I\cap J}| + |F_I \cap F_J| - \rho(F_I \vee F_J) =$$
$$|F_I \cup F_J| - \textstyle\sum_{j\in I\cup J} \eta(F_j) - \rho(F_I \vee F_j) =$$
$$|F_{I\cup J}| - \textstyle\sum_{j\in I\cup J} \eta(F_j) - \rho(F_I \vee F_j).$$

If $F_{I\cup J} \in \mathcal{Z}_{<k}$, then

$$|F_{I\cup J}| - \sum_{j\in I\cup J} \eta(F_j) - \rho(F_I \vee F_j) = |F_{I\cup J}| - \sum_{j\in I\cup J} \eta(F_j) - |F_{I\cup J}| - \sum_{j\in I\cup J} \eta(F_j) = 0.$$

If $F_{I\cup J} \notin \mathcal{Z}_{<k}$, then

$$|F_{I\cup J}| - \sum_{j\in I\cup J} \eta(F_j) - \rho(F_I \vee F_j) = |F_{I\cup J}| - \sum_{j\in I\cup J} \eta(F_j) - k \geq 0$$

by property (iii) in III-D1. Moreover, for E and F_I we have that

$$\rho(F_I) + \rho(E) - (\rho(F_I \vee E) + \rho(F_I \wedge E) + |(F_I \cap E) \setminus (F_I \wedge E)|) =$$
$$\rho(F_I) + \rho(E) - \rho(E) - \rho(F_I) + |F_I| - |F_I| = 0.$$

We have now proven that the axioms (Z0)–(Z3) in Theorem II.1 are satisfied by $\mathcal{Z}$ and ρ. Hence, $\mathcal{Z}$ and ρ define a matroid $M = M(F_1, \ldots, F_m; k; \rho)$ over E.

The parameters (n, k, d, r, δ) will now be investigated using Lemma III.1. Firstly, the parameters (n, k, d, r, δ) are defined for M with $n = |1_{\mathcal{Z}}| = |E|$ and $k = \rho(1_{\mathcal{Z}}) = \rho(E)$, since $E \in \mathcal{Z}$ and $\rho(E) = k > 0$. By Axiom $(Z2)$ in Theorem II.1, $\eta(Y) > \eta(X)$ for all $X, Y \in \mathcal{Z}$ when $X \subsetneq Y$. Hence, by Lemma III.1 (iii),

$$\begin{aligned} d &= n - k + 1 - \max\{\eta(F_I) : F_I \in \mathcal{Z}_{<k}\} \\ &= n - k + 1 - \max\{|F_I| - (|F_I| - \textstyle\sum_{i\in I} \eta(F_i)) : F_I \in \mathcal{Z}_{<k}\} \\ &= n - k + 1 - \max\{\textstyle\sum_{i\in I} \eta(F_i) : F_I \in \mathcal{Z}_{<k}\}. \end{aligned}$$

Let $\delta - 1 = \min_{i\in[m]}\{\eta(F_i)\}$ and S be a subset of F_i such that $|S| = \rho(F_i) + \delta - 1$. By Proposition II.2(vii) and property (v) in III-D1,

$$\mathcal{Z}(M|F_i) = \{Z \in \mathcal{Z}(M) : Z \subseteq F_i\} = \{\emptyset, F_i\}.$$

Hence, from Proposition II.1 (ii) and (iii),

$$\rho(X) = \{|X|, \rho(F_i)\} \text{ for } X \subseteq F_i$$

and

$$\mathcal{C}(M) \cap F_i = \{X \subseteq F_i : |X| = \rho(F_i) + 1\}.$$

This implies that S is a cyclic set and that

$$d(M|S) = |S| - \rho(S) + 1 = \rho(F_i) + \delta - 1 - \rho(F_i) + 1 = \delta.$$

by (11). Therefore, with $r = \max_{i\in[m]}\{\rho(F_i)\}$ and as $F_{[m]} = E$, statements (a) - (c) in Lemma III.1 are satisfied. Consequently, M has (r,δ)-locality and S is a locality set. $\square$

Proof of Corollary III.1. That assumptions (ii)–(v) in Theorem III.4 are satisfied follow directly from the assumptions (i)–(iv) in this corollary. Assumption (i) in Theorem III.4 is satisfied since (i) and (iv) in this corollary imply that

$$|F_{[m]\setminus\{j\}} \cap F_j| < \rho(F_j) < |F_j|$$

for every $j \in [m]$. To prove that assumption (vi) in Theorem III.4 is satisfied, it is enough to prove that

$$|F_J| - \sum_{j\in J}\eta(F_j) \geq |F_I| - \sum_{i\in I}\eta(F_i) \tag{31}$$

when $I \subsetneq J$. Assume that $|J \setminus I| = 1$. Then $J = I \cup \{j\}$, and consequently from (iv) in this corollary,

$$\begin{aligned} &|F_J| - \textstyle\sum_{j\in J}\eta(F_j) - (|F_I| - \sum_{i\in I}\eta(F_i)) = \\ &|F_j| - |F_I \cap F_j| - \eta(F_j) > \\ &|F_j| - \rho(F_j) - \eta(F_j) = 0. \end{aligned}$$

Now, assume for some $|J \setminus I| > 1$ and that (31) holds for $I \subsetneq I' \subsetneq J$. Then, from (iv) in this corollary, for $t \in J \setminus I$,

$$\begin{aligned} &|F_J| - \textstyle\sum_{j\in J}\eta(F_j) - (|F_{J\setminus\{t\}}| - \sum_{i\in J\setminus\{t\}}\eta(F_i)) = \\ &|F_t| - |F_{J\setminus\{t\}} \cap F_t| - \eta(F_t) > \\ &|F_t| - \rho(F_t) - \eta(F_t) = 0. \end{aligned}$$

Hence,

$$|F_J| - \sum_{j\in J}\eta(F_j) - (|F_I| - \sum_{i\in I}\eta(F_i)) =$$
$$|F_J| - \sum_{j\in J}\eta(F_j) - (|F_{J\setminus\{t\}}| - \sum_{i\in J\setminus\{t\}}\eta(F_i)) +$$
$$(|F_{J\setminus\{t\}}| - \sum_{i\in J\setminus\{t\}}\eta(F_i)) - (|F_I| - \sum_{i\in I}\eta(F_i)) > 0.$$

□

Proof of Proposition III.3. Statement (i) is given in (31). Statement (ii) follows from (i), which implies that

$$\mathcal{Z} = \mathcal{Z}_{<k} \cup E = \{F_I : |F_I| - \sum_{i\in I}\eta(F_i) < k\} \cup E.$$

Statement (iii) also follows directly from (i). For Statement (iv) we first remark that a set $X \subseteq E$ is independent if and only if X does not contain any circuits. Hence, from (i) and Proposition II.1(iii),

$$\begin{aligned}\mathcal{I} &= \{X \subseteq E : |X| \leq \rho(Y) \text{ for all } Y \in \mathcal{Z}\}\\ &= \{X \subseteq E : |F_I \cap X| \leq \min\{|F_I| - \textstyle\sum_{i\in I}\eta(F_i), k\} \text{ for all } I \subseteq [m]\}.\end{aligned}$$

The right implication in Statement (v) follows directly from (iv). For the proof of the left implication, assume that for some $I \subseteq [m]$ and $X \subseteq F_I$ that

$$|F_{I'} \cap X| \leq \min\{|F_{I'}| - \sum_{i\in I'}\eta(F_i), k\} \text{ for all } I' \subseteq I\}.$$

Observe that, by Statement (iii), $\min\{|F_{I'}| - \sum_{i\in I'}\eta(F_i), k\} = \rho(F_{I'})$. Now, choose an element $j \in [m] \setminus I$ and a subset $I' \subseteq I$. If $\rho(F_{I'\cup\{j\}}) = k$, then

$$|X \cap F_{I'\cup\{j\}}| \leq |X| = |X \cap F_I| \leq \rho(F_I) \leq k = \rho(F_{I'\cup\{j\}}).$$

Now suppose that $\rho(F_{I'\cup\{j\}}) < k$. Then, using Statement (iv) in Corollary III.1,

$$\begin{aligned}|X \cap F_{I'\cup\{j\}}| &\leq |X \cap F_{I'}| + |F_I \cap F_j| - |F_{I'} \cap F_j|\\ &\leq \rho(F_{I'}) + \rho(F_j) - |F_{I'} \cap F_j|\\ &= |F_{I'}| - \textstyle\sum_{i\in I'}\eta(F_i) + |F_j| - \eta(F_j) - |F_{I'} \cap F_j|\\ &= |F_{I'\cup\{j\}}| - \textstyle\sum_{i\in(I\cup\{j\})}\eta(F_i)\\ &= \rho(F_{I'\cup\{j\}}).\end{aligned}$$

The left implication of (v) now follows from (iv). □

Proof of Theorem III.5. To prove the theorem, we will first show that the assumptions (i)–(iv) in Corollary III.1 are satisfied by $(F_1, \ldots, F_m; k; \rho)$, obtained from the graph $G(\alpha, \beta, \gamma; k, r, \delta)$ in (15). We will then show that the values of the parameters (n, k, d, r, δ) of $M(F_1, \ldots, F_m; k; \rho)$ are the ones requested in Theorem III.5.

Corollary III.1 (i) follows directly from (14) (ii) and (iii), and (15) (i) and (ii). Moreover, Corollary III.1 (ii) is obvious. For Corollary III.1 (iii), we first notice that by (14) (iv) and (vi), we can define γ as the size of a nonempty intesection of two sets F_i and F_j, as in (15) (iii). Now, the nonexistence of 3-cycles in G by (14) (i) implies (15) (iv). Hence, we get

$$|F_{[m]}| = \sum_{i\in[m]} |F_i| - \sum_{w\in W} \gamma(w). \tag{32}$$

Moreover, for $i \in [m]$, we have

$$\eta(F_i) = |F_i| - \rho(F_i) = \delta - 1 + \beta(i).$$

Consequently,

$$\begin{aligned} |F_{[m]}| - \textstyle\sum_{i\in[m]} \eta(F_i) &= \textstyle\sum_{i\in[m]} |F_i| - m(\delta-1) - \sum_{i\in[m]}(\beta(i) + \gamma(i)) \\ &= mr - \textstyle\sum_{i\in[m]} \alpha(i) - \sum_{i\in[m]} \gamma(i). \end{aligned}$$

Therefore, by (14) (v), Corollary III.1 (iii) holds. For Corollary III.1 (iv), we first remark that

$$F_{[m]\setminus i} \cap F_i = \sum_{w=\{i,j\}\in W} \gamma(w)$$

and $\rho(F_i) = r - \alpha(i)$ for $i \in [m]$. Hence Corollary III.1 (iv) holds, by (14)(vi).

We will now determine the parameters (n, k, d, r, δ), proving that they agree for the graph and the matroid. The given parameters (r, δ) for the graphs also give (r, δ)-locality of the matroid as $\rho(F_i) \le r$ and $\eta(F_i) \ge \delta - 1$ by (14) (ii) and (iii), and (15) (i) and (ii). We have already proven that the parameter k of the graph is the rank of the matroid. Moreover, by (32),

$$n = |F_{[m]}| = |\sum_{i\in[m]} |F_i| - \sum_{w\in W} \gamma(w) = (r+\delta-1)m - \sum_{i\in[m]} \alpha(i) + \sum_{i\in[m]} \beta(i) - \sum_{w\in W} \gamma(w).$$

The statement about d in Theorem III.5 (ii) holds as a consequence of Proposition III.3, Theorem III.4(iii) and the properties that

$$\sum_{i\in I} \eta(F_i) = |I|(\delta-1) + \sum_{i\in I} \beta(i) \text{ and } |F_I| - \sum_{i\in I} \eta(F_i) = r|I| - \sum_{i\in I} \alpha(i) - \sum_{w\subseteq I, w\in W\in I} \gamma(w).$$

This concludes the proof. □

Proof of Theorem III.6. We will first prove that the graph $G(\gamma; k, r, \delta, a, b)$ in this theorem is an instance of the graph $G(\alpha, \beta, \gamma; k, r, \delta)$ from Theorem III.5, with $\alpha(i) = \beta(i) = 0$ for $i \in [m]$.

The statements (14) (i)–(iv) and (vi) follow directly from (16). For (14) (v), by (16) (v), and assigning $k = \lceil \frac{k}{r} \rceil r - a$, we get

$$k \leq |I|r - \sum_{w=\{i,j\}\in W, i,j\in I} \gamma(i), \tag{33}$$

when $I \subseteq [m]$ and $|I| = \lceil \frac{k}{r} \rceil$. Thus, by (16)(vi),

$$k \leq |I|r - \sum_{w=\{i,j\}\in W, i,j\in I} \gamma(i),$$

for any $I \subseteq [m]$ with $|I| \geq \lceil \frac{k}{r} \rceil$. Hence, statement (14) (v) also holds. Moreover, (17) also follows. Theorem III.6 (ii) has already been proved. Theorem III.6 (i) and (iii) follows from III.5 (i) and (ii), (33) and (14). □

Proof of Theorem III.7: We will divide the proof of Theorem III.7 into the the parts (i)–(v). First, we recall that a and b are the integers where

$$k = \left\lceil \frac{k}{r} \right\rceil r - a \text{ and } n = \left\lceil \frac{n}{r+\delta-1} \right\rceil (r+\delta-1) - b.$$

Construction 1: In order to prove Theorem III.7 (i), we will define graphs $G = G(\alpha, \beta, \gamma; k, r, \delta)$ that satisfy the conditions given in (14) and then use Theorem III.5. Let

$$\begin{array}{ll} (a) & m = \left\lceil \frac{n}{r+\delta-1} \right\rceil = \lceil \frac{k}{r} \rceil + t \text{ where } t \geq 0, \\ (b) & G \text{ be the graph on } [m] \text{ with no edges}, \\ (c) & \alpha(i) = \begin{cases} 0 & \text{if } \ i \in [m-1], \\ b & \text{if } \ i = m, \end{cases} \\ (d) & \beta(i) = 0 \text{ for } i \in [m]. \end{array} \tag{34}$$

Proof of Theorem III.7 (i). The conditions (14) (i), (iii) and (iv) follow directly from (b) and (d). Condition (14)(ii) follows from the property that $0 \leq b \leq a \leq r+1$. For condition (14)(v) we obtain from (a) that

$$k = \left\lceil \frac{k}{r} \right\rceil r - a \leq \left\lceil \frac{k}{r} \right\rceil r - b \leq (\left\lceil \frac{k}{r} \right\rceil + t)r - b = mr - \sum_{i\in[m]} \alpha(i) - \sum_{w\in W} \gamma(w).$$

Condition (14) (vi) follows from the facts that $r - \alpha(i) \geq 1$ for $i \in [m]$ and that G has no edges. Now, by Theorem III.5, there is an (n, k, d, r, δ)-matroid with

$$n = (r+\delta-1)m - b = (r+\delta-1) \left\lceil \frac{n}{r+\delta-1} \right\rceil - b$$

and

$$d = n - k - 1 - (\left\lceil \frac{k}{r} \right\rceil - 1)(\delta - 1),$$

as

$$r|I| - \sum_{i \in I} \alpha(i) - \sum_{i \in I} \gamma(i) \geq r|I| - b \geq r|I| - a$$

for $I \subseteq [m]$. □

Construction 2 when $b \leq r - 1$: To prove Theorem III.7 (ii) when $b \leq r - 1$, we will define graphs $G = G(\alpha, \beta, \gamma; k, r, \delta)$ that satisfy the conditions in (14), and then use Theorem III.5. Let

$$\begin{aligned}
&(a) \quad m = \left\lceil \tfrac{n}{r+\delta-1} \right\rceil = \left\lceil \tfrac{k}{r} \right\rceil + t \text{ where } t \geq 1, \\
&(b) \quad G \text{ be the graph on } [m] \text{ with no edges}, \\
&(c) \quad \alpha(i) = \begin{cases} 0 & \text{if} \quad i \in [m-1], \\ b & \text{if} \quad i = m, \end{cases} \\
&(d) \quad \beta(i) = 0 \text{ for } i \in [m],
\end{aligned} \tag{35}$$

Proof of Theorem III.7 (ii) when $\mathbf{b \leq r - 1}$. The conditions (14) (i)–(iv) follow directly from (b)–(d). For condition (14) (v), we obtain from (a) that

$$k = \left\lceil \frac{k}{r} \right\rceil r - a \leq (\left\lceil \frac{k}{r} \right\rceil + 1)r - b \leq (\left\lceil \frac{k}{r} \right\rceil + t)r - b = mr - \sum_{i \in [m]} \alpha(i) - \sum_{w \in W} \gamma(W).$$

Condition (14) (vi) follows as $r - \alpha(i) \geq 1$ for $i \in [m]$, and as G has no edges. Now, by Theorem III.5, there is an (n, k, d, r, δ)-matroid with

$$n = (r + \delta - 1)m - b = (r + \delta - 1) \left\lceil \frac{n}{r + \delta - 1} \right\rceil - b.$$

Moreover, by Theorem III.5, this matroid satisfies

$$d = n - k - 1 - \left\lceil \frac{k}{r} \right\rceil (\delta - 1),$$

as

$$\max_{I \in V_{<k}} \{(\delta - 1)|I| + \sum_{i \in I} \beta(i)\} = \left\lceil \frac{k}{r} \right\rceil$$

for

$$V_{<k} = \{I \subseteq [m] : r|I| - \sum_{i \in I} \alpha(i) - \sum_{i,j \in I,\ w = \{i,j\} \in W} \gamma(w) < k\}.$$

□

Construction 2 when $b \geq r$: To prove Theorem III.7 (ii) when $b \geq r$, we will define graphs $G = G(\alpha, \beta, \gamma; k, r, \delta)$ that satisfy the conditions in (14), and then use Theorem III.5. Let

$$\begin{aligned} &(a) \quad m = \left\lceil \tfrac{n}{r+\delta-1} \right\rceil - 1 = \lceil \tfrac{k}{r} \rceil + t, \text{ where } t \geq 0, \\ &(b) \quad G \text{ be the graph on } [m] \text{ with no edges}, \\ &(c) \quad \alpha(i) = 0 \text{ for } i \in [m], \\ &(d) \quad \beta(i) = \begin{cases} 0 & \text{if} \quad i \in [m-1], \\ r+\delta-1-b & \text{if} \quad i = m. \end{cases} \end{aligned} \tag{36}$$

Proof of Theorem III.7 (ii) when $\mathbf{b \geq r}$. The conditions (14) (i)–(iv) follow directly from (b)–(d). For condition (14) (v), we obtain from (a) that

$$k = \left\lceil \frac{k}{r} \right\rceil r - a \leq \left\lceil \frac{k}{r} \right\rceil r = mr - \sum_{i \in [m]} \alpha(i) - \sum_{w \in W} \gamma(w).$$

Condition (14) (vi) follows as $r - \alpha(i) \geq 1$ for $i \in [m]$ and as G has no edges. Now, by Theorem III.5, there is an (n, k, d, r, δ)-matroid with

$$n = (r+\delta-1)m + (r+\delta-1-b) = (r+\delta-1) \left\lceil \frac{n}{r+\delta-1} \right\rceil - b.$$

Moreover, by Theorem III.5,

$$d = n - k - 1 - \left(\left\lceil \frac{k}{r} \right\rceil (\delta - 1) - (b - r) \right),$$

as

$$\max_{I \in V_{<k}} \{(\delta-1)|I| + \sum_{i \in I} \beta(i)\} = \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) + (r + \delta - 1 - b)$$

for

$$V_{<k} = \{I \subseteq [m] : r|I| - \sum_{i \in I} \alpha(i) - \sum_{i,j \in I,\ w = \{i,j\} \in W} \gamma(w) < k\}.$$

This concludes the proof. □

Proof of Theorem III.7 (iii), right implication. By the structure theorem III.3, we see that the existence of an (n, k, d, r, δ)-matroid with $d = n - k + 1 - \left(\lceil \frac{k}{r} \rceil - 1 \right) (\delta - 1)$ implies the existence of subsets $F_1, \dots, F_m$ of $[n]$ such that:

(i) $F_j \nsubseteq \bigcup_{i \in [m] \setminus \{j\}} F_i$ for $j = 1, \dots, m$,

(ii) $|F_i| \leq r + \delta - 1$ for $i = 1, \dots, m$,

(iii) $|\bigcup_{i \in [m]} F_i| = n$,

(iv) $|F \cap (\bigcup_{i \in I} F_i)| \leq |F| - \delta$ for every $F \in \{F_i\}_{i \in [m] \setminus I}$ with $I \subseteq [m]$ and $|I| < \lceil \frac{k}{r} \rceil$,

(v) $|\bigcup_{i\in I} F_i| - |I|(\delta-1) \geq k = \lceil \frac{k}{r} \rceil r - a$ for every $I \subseteq [m]$ with $|I| \geq \lceil \frac{k}{r} \rceil$.

For simplicity, denote $\lceil \frac{k}{r} \rceil = h$. For any set system $F_1, \ldots, F_m$ where $|F_i| \leq r+\delta-1$ for every i, construct a graph $\mathcal{G}$ on vertex set $[m]$, with an edge between i and j if and only if $F_i \cap F_j \neq \emptyset$. Note that, when $I \subseteq [m]$ is such that the induced subgraph $\mathcal{G}[I]$ on I is connected, then

$$|F_I| \leq (r+\delta-2)|I| + 1.$$

If $\mathcal{G}[I]$ is connected and equality holds in the above inequality, then I is said to be *full*. Note that for every full component I in $\mathcal{G}$ and integer $1 \leq u \leq |I|$, there is a subset $I' \subseteq I$ such that $|I'| = t$ and I' is full. Denoting by $c(\mathcal{G}[I])$ the number of full components of $\mathcal{G}[I]$, we get

$$|F_I| \leq (r+\delta-2)|I| + c(\mathcal{G}[I]).$$

Let J be the union of the $h-a-1$ largest full components of $\mathcal{G}$ together with all non-full components of $\mathcal{G}$. If $|J| \geq h$, then we have a subset of nodes $J' \subseteq J$ with $|J'| = h$, such that $c(\mathcal{G}[J']) \leq h-a-1$. However, assuming

$$|F_I| \geq h(r+\delta-1) - a = h(r+\delta-2) + h - a$$

for every subset $I \subseteq [m]$ with $|I| = h$, then $c(\mathcal{G}[J']) \geq h-a$. Hence, $|J| \leq h-1$ and $\mathcal{G}[[m]\setminus J]$ is a union of full components $I_1, \ldots, I_s$ of $\mathcal{G}$, and these full components contain at most $\left\lfloor \frac{h-1}{h-1-a} \right\rfloor$ nodes each.

When bounding $\left\lceil \frac{n}{r+\delta-1} \right\rceil$, we first notice that

(i) $|I|(r+\delta-1) - |F_I| = |I| - 1$ if I is connected and full,

(ii) $|I|(r+\delta-1) - |F_I| \geq |I|$ if I is connected and not full,

(iii) $h(r+\delta-1) - |F_I| \leq a$ if $|I| \leq h$,

(iv) $|I_i|(r+\delta-1) - |F_{I_i}| \leq \left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1$ for $1 \leq i \leq s$.

Hence,

$$\begin{aligned} b &= m(r+\delta-1) - |F_{[m]}| \\ &= |J|(r+\delta-1) - |F_J| + \textstyle\sum_{i=1}^{s} |I_i|(r+\delta-1) - |F_{I_i}|. \end{aligned}$$

Also, as $|J| < h$, we get

$$|J|(r+\delta-1) - |F_J| + \sum_{i=1}^{s} |I_i|(r+\delta-1) - |F_{I_i}| \leq a + s\left(\left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1\right).$$

Hence, as $b > a$, we have $\left\lfloor \frac{h-1}{h-1-a} \right\rfloor \geq 2$, or equivalently $a \geq \lceil \frac{h}{2} \rceil$. Now, assume that $a \geq \lceil \frac{h}{2} \rceil$. For the cardinality of $F_{[m]}$ we have that

$$|F_{[m]}| = (|J| + \sum_{i=1}^{s} |I_i|)(r+\delta-1) - b.$$

By using (iii), (iv) and the property that $|J| < h$, we now obtain that

$$\left\lceil \frac{|F_{[m]}|}{r+\delta-1} \right\rceil \geq h-1+\left\lfloor \frac{b-a}{\left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1} \right\rfloor \left\lfloor \frac{h-1}{h-1-a} \right\rfloor + t, \tag{37}$$

where

$$t = \begin{cases} 0 \text{ if } (\left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1)|(b-a) \\ (b-a) - \left\lfloor \frac{b-a}{\left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1} \right\rfloor \left(\left\lfloor \frac{h-1}{h-1-a} \right\rfloor - 1 \right) + 1 \text{ otherwise.} \end{cases}$$

Rearranging equation (37), we find the bound

$$\left\lceil \frac{n}{r+\delta-1} \right\rceil \geq \left\lceil \frac{k}{r} \right\rceil - 1 + (b-a)\left(1+\frac{1}{t}\right), \tag{38}$$

where

$$t = \lfloor a/\left(h-1-a\right) \rfloor = \left\lfloor a/\left(\left\lceil \frac{k}{r} \right\rceil - 1 - a \right) \right\rfloor.$$

□

Construction 3: To prove Theorem III.7 (iii), we will construct graphs $G = G(\gamma; k, r, \delta, a, b)$ that satisfy the conditions given in (16) and then use Theorem III.6. For simplicity, denote

$$s = \left\lfloor \frac{\lceil \frac{k}{r} \rceil - 1}{\lceil \frac{k}{r} \rceil - 1 - a} \right\rfloor, \ u = \left\lceil \frac{k}{r} \right\rceil - 1 - a + \left\lceil \frac{b-a}{s-1} \right\rceil \text{ and } x = \left\lceil \frac{k}{r} \right\rceil - 1 - s\left(\left\lceil \frac{k}{r} \right\rceil - 1 - a \right)$$

In order to prove Theorem III.7 (iii), we will define graphs $G = G(\gamma; k, r, \delta, a, b)$ that satisfy the conditions given in (16) and then use Theorem III.6. Let

(39)

(i) $m \geq \lceil \frac{k}{r} \rceil - 1 + (b-a)\left(1+\frac{1}{t}\right)$, where $t = \left\lfloor a/\left(\lceil \frac{k}{r} \rceil - 1 - a\right) \right\rfloor$,

(ii) G be the graph consisting of vertex-disjoint paths $P_1, \ldots, P_u$ with

$$|P_i| = \begin{cases} s+1 & \text{if } 1 \leq i \leq x, \\ s & \text{if } x+1 \leq i \leq u-1, \\ s & \text{if } i = u \text{ and } s-1 \mid b-a, \\ b-a-\lfloor \frac{b-a}{s-1} \rfloor (s-1)+1 & \text{if } i = u \text{ and } s-1 \nmid b-a, \end{cases}$$

(iii) $\gamma(w) = 1$ for each $w \in W$.

Proof of Theorem III.7 (iii), left implication. : We first note that condition (16) (i) follows directly as G has no cycles. Condition (16) (ii) is a consequence of (c). For condition (16) (iii), we first remark that by (37) and (38), we get

$$\sum_{w \in W} \gamma(w) = |W| = (\sum_{i \in u} |P_i|) - u = (s+1)x + (u-1-x)s + |P_u| - u = b.$$

Condition (16) (iv) follows directly from (a). Condition (16)(v) follows from the fact that

$$\sum_{i=1}^{\lceil \frac{k}{r} \rceil - 1 - a} |P_i| = \left\lceil \frac{k}{r} \right\rceil - 1$$

and

$$\sum_{i,j \in P, w=\{i,j\} \in W} \gamma(w) = \left\lceil \frac{k}{r} \right\rceil - 1 - (\left\lceil \frac{k}{r} \right\rceil - 1 - a) = a,$$

where $P = \bigcup_{1 \le i \le \lceil \frac{k}{r} \rceil - 1 - a} P_i$. Moreover, condition (16) (vi) follows from the property that $\gamma(w) = 1$ for all $w \in W$. The result now follows using (37) and (38), which imply that

$$|\bigcup_{i=1}^{u} P_i| = \left\lceil \frac{k}{r} \right\rceil - 1 + (b - a) \left(1 + \frac{1}{t}\right),$$

where $t = \left\lfloor a / \left(\lceil \frac{k}{r} \rceil - 1 - a\right) \right\rfloor$. □

Construction 4: To prove Theorem III.7(iv), we will construct graphs $G = G(\gamma; k, r, \delta, a, b)$ that satisfy the conditions in Corollary III.2, and then use Theorem III.6. For simplicity, denote

$$s = \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor, \; t = \left\lfloor \frac{r-1}{s} \right\rfloor, \; u = \left\lceil \frac{\lceil \frac{k}{r} \rceil + 1}{2} \right\rceil \text{ and } x = \left\lceil \frac{b - \lfloor \frac{b}{stu} \rfloor stu}{s} \right\rceil.$$

Before we are ready to construct G, we need some subgraphs that will be the building blocks of G. For $1 \le i \le t$, let P_i denote a path containing $u+1$ vertices, with p_i as start vertex and q_i as end vertex. Now, let B denote the graph obtained from $\sqcup_i P_i$ by identifying all p_i to the same vertex $p \in B$, all all the end vertices q_i the same vertex $q \in B$

We will now define a subgraph of $B'(h)$ of B, where h denotes the number of edges that the subgraph should have. First we remark that the number of edges in B equals tu. Now, order the edges in B from 1 to tu by starting from the start vertex p and ending in the end vertex q for each path, ordering the edges path by path from P_1 to P_t. This is

- the path P_i is the sequence of vertices $p = v_1^{(i)}, v_2^{(i)}, \ldots v_u^{(i)}, v_{u+1}^{(i)} = q$, then edge $\{v_j^{(i)}, v_{j+1}^{(i)}\}$ is ordered as edge number $(i-1)u + j$.

The subgraph $B'(h)$ is now defined as the subgraph of B that consists of the edges numbered from 1 to x and the vertices associated to these edges. By $B'(0)$ we mean the graph with no vertices.

The number of vertices of B equals the number of internal nodes in paths $P_1, \ldots, P_t$ plus 2, *i.e.*,

$$t(u-1) + 2.$$

Moreover, the number of vertices in $B'(h)$, when $h \neq 0$, equals

$$\left\lfloor \frac{h}{u} \right\rfloor (u-1) + (h - \left\lfloor \frac{h}{u} \right\rfloor u) + (1 + \min\{\left\lfloor \frac{h}{u} \right\rfloor, 1\}) = h - \left\lfloor \frac{h}{u} \right\rfloor + 1 + \min\{\left\lfloor \frac{h}{u} \right\rfloor, 1\}.$$

Now, for the construction of G, let

(40)

(i) $m \geq \lfloor \frac{b}{stu} \rfloor (t(u-1)+2)+y$, where

$$y = \begin{cases} 0 & \text{if} \quad stu \mid b, \\ x - \lfloor \frac{x}{u} \rfloor + 1 + \min\{\lfloor \frac{x}{u} \rfloor, 1\} & \text{if} \quad stu \nmid b; \end{cases}$$

(ii) G be the graph with vertices $[m]$ and edges W, where G consists of $\lfloor \frac{b}{stu} \rfloor$ copies of B, one copy of $B'(x)$ and possibly some additional isolated vertices;

(iii)
- If $s \mid b$ then $\gamma(w) = s$ for all $w \in W$,
- If $s \nmid b$ then

$$\gamma(w) = \begin{cases} s & \text{if} \quad w \text{ is not the vertex number x in } B'(x), \\ b - \lfloor \frac{b}{s} \rfloor s & \text{if} \quad w \text{ is the vertex number x in } B'(x); \end{cases}$$

Proof of Theorem III.7 (iv). As $\lceil \frac{k}{r} \rceil \geq 3$, and the smallest size of a cycle in the graph is $2u \geq \lceil \frac{k}{r} \rceil + 1$, it follows that G has no l-cycles for $l \leq \max\{3, \lceil \frac{k}{r} \rceil\}$. Also, the property that $1 \leq \gamma(w) \leq \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$ for all edges w in G follows from (c) as $s = \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$. For condition (16) (iii), we remark that for a copy of B in G, the total sum of $\gamma(w)$ for all edges w in B equals stu. Moreover, the total sum of $\gamma(w)$ for all edges w in $B'(x)$ equals sx if $s|b$, and $s(x-1) + b - \lfloor \frac{b}{s} \rfloor s$ if $s \nmid b$. Hence

$$\sum_{w \in W} \gamma(w) = \left\lfloor \frac{b}{stu} \right\rfloor \sum_{edges \ w \in B} \gamma(w) + \sum_{edges \ w \in B'(x)} \gamma(w) = b.$$

Condition (16)(vi) follows as $ts \leq r-1$ and $2s \leq 2 \lfloor \frac{a}{2} \rfloor \leq a \leq r-1$. Hence, by Corollary III.2 and Theorem III.6, the theorem is now proven. □

Construction 5 when $2a \leq r-1$*:* We will construct graphs $G = G(\gamma; k, r, \delta, a, b)$ that satisfy the conditions in (16) and then use Theorem III.6. To construct G, let

(a) $m \geq \lceil \frac{b}{a} \rceil + 1$;

(b) G be the graph with vertices $[m]$ and edges $W = \{\{i, i+1\} : 1 \leq i \leq \lceil \frac{b}{a} \rceil\}$;

(c) For $\{i, i+1\} \in W$ let,

$$\gamma(\{i, i+1\}) = \begin{cases} a & \text{if} \quad i < \lceil \frac{b}{a} \rceil, \\ a & \text{if} \quad i = \lceil \frac{b}{a} \rceil \text{ and } a|b, \\ b - \lfloor \frac{b}{a} \rfloor a & \text{if} \quad i = \lceil \frac{b}{a} \rceil \text{ and } a \nmid b. \end{cases} \tag{41}$$

Proof of Theorem III.7(v) when $\mathbf{2a \le r - 1}$. That G has no l-cycles for $l \le \max\{3, \lceil \frac{k}{r} \rceil\}$ follows directly as G has no cycles. Also, that $1 \le \gamma(w) \le \left\lfloor \frac{a}{\lceil \frac{k}{r} \rceil - 1} \right\rfloor$ for all edges w in G follows directly from (c). For condition (16) (iii), we obtain from (c) that

$$\sum_{w \in W} \gamma(w) = \left(\frac{b}{a} - 1\right) a + a = b \text{ if } a|b,$$

and

$$\sum_{w \in W} \gamma(w) = \left\lfloor \frac{b}{a} \right\rfloor a + b - \left\lfloor \frac{b}{a} \right\rfloor a = b \text{ if } a \nmid b.$$

As the maximal number of neighbours of a vertex in G is 2, we get that for any $i \in [m]$,

$$\sum_{w = \{i,j\} \in W} \gamma(w) \le 2a \le r - 1.$$

Hence, by Corollary III.2 and Theorem III.6, the theorem is now proved. □

Construction 5 when $2a > r - 1$*:* In order to prove Theorem III.7(v), we will construct graphs $G = G(\gamma; k, r, \delta, a, b)$ that satisfy the conditions in Corollary III.2, and then use Theorem III.6. For simplicity, denote $h = \lfloor \frac{r-1}{2} \rfloor$. Now, to construct G, let

(42)

(a) $m \ge \lceil \frac{b}{h} \rceil + 1$;
(b) G be the graph with vertices $[m]$ and edges $W = \{\{i, i+1\} : 1 \le i \le \lceil \frac{b}{h} \rceil\}$;
(c) For $\{i, i+1\} \in W$, let

$$\gamma(\{i, i+1\}) = \begin{cases} h & \text{if } \ i < \lceil \frac{b}{h} \rceil, \\ h & \text{if } \ i = \lceil \frac{b}{h} \rceil \text{ and } h|b, \\ b - \lfloor \frac{b}{h} \rfloor h & \text{if } \ i = \lceil \frac{b}{h} \rceil \text{ and } h \nmid b. \end{cases}$$

Proof of Theorem III.7(v) when $\mathbf{2a > r - 1}$. The proof is completely analogous to the proof Theorem III.7(v) when $2a \le r - 1$, replacing a by h. □

Proof of Lemma IV.1. The independent sets of the transversal associated to $G = G(F_1, \ldots, F_m; \rho)$ are

$$\mathcal{I}(M(G)) = \{X \subseteq E : \exists \text{ a set of } \ |X| \text{ vertex-disjoint paths from } X \text{ to } H\}.$$

Let $k = |F_{[m]}| - \sum_{i \in [m]} \eta(F_i)$. Then, by Proposition III.3(i), $|F_I| - \sum_{i \in I} \eta(F_i) \le k$ for all $I \subseteq [m]$. Then, by Proposition III.3 (iv), the independent sets of the matroid $M = M(F_1, \ldots, F_m; k; \rho)$ are

$$\mathcal{I}(M) = \{X \subseteq E : |X \cap F_I| \le |F_I| - \sum_{i \in I} \eta(F_i) \text{ for each } I \subseteq [m]\}.$$

We will prove that the matroid $M(G)$ is equal to M by proving that the independent sets $\mathcal{I}(M)$ and $\mathcal{I}(M(G))$ are equal.

Let

$$H_I = \{u \in H : h(u) \cap I \neq \emptyset\} = \{u \in H : \exists e \in E \text{ with } e \in F_I \text{ and } (\overrightarrow{e,u}) \in D\}.$$

To prove the inclusion $\mathcal{I}(M(G)) \subseteq \mathcal{I}(M)$, we will prove, by induction on $|I|$, that

$$|H_I| = |F_I| - \sum_{i \in I} \eta(F_i) \tag{43}$$

for each $I \subseteq [m]$. For $I = \emptyset$, equation (43) is trivially true. For $i \in [m]$, as $|F_{[m]\setminus\{i\}} \cup F_i| < \rho(F_i)$, we get from $(6-12)$ in Algorithm 1 that $|H_i| = \rho(F_i) = |F_i| - \eta(F_i)$. Now, for some $l \geq 1$, assume that (43) is true for each $I \subseteq [m]$ with $|I| \leq l$. Let I be a subset of $[m]$ with $|I| = l$. Take any $j \in [m] \setminus I$, and let $J = I \cup \{j\}$. Now,

$$\begin{aligned}
|H_J| &= |H_I| + |H_j| - |H_I \cap H_j| \\
&= |F_I| - \textstyle\sum_{i \in I} \eta(F_i) + |F_j| - \eta(F_j) - |\{u \in H : h(u) \cap I \neq \emptyset \text{ and } j \in h(u)\}| \\
&= |F_I| - \textstyle\sum_{i \in I} \eta(F_i) + |F_j| - \eta(F_j) - |F_I \cap F_j| \\
&= |F_J| - \textstyle\sum_{i \in J} \eta(F_i).
\end{aligned}$$

We will now prove the reverse inclusion $\mathcal{I}(M) \subseteq \mathcal{I}(M(G))$. By Proposition III.3(v), for any $X \subseteq F_I$,

$$X \in \mathcal{I}(M) \iff |X \cap F_{I'}| \leq |F_{I'}| - \sum_{i \in I'} \eta(F_i) \text{ for all } I' \subseteq I. \tag{44}$$

For $X \subseteq E$ and G, let

$$A(X) = \{u \in H : \exists x \in X \text{ such that } (\overrightarrow{x,u}) \in D\}.$$

Let $X = X' \cup X''$ be a subset of F_I and an independent set in M, where $X' = \{x \in X : |s(x)| = 1\}$ and $X'' = \{x \in X : |s(x)| \geq 2\}$ respectively. For $x \in X''$, by (7–9) in Algorithm1, there is a node $u_x \in H$ for which $(\overrightarrow{x,u_x}) \in D$. Consequently, for $H'' = \{u_x \in H : x \in X''\}$, we have

$$|H''| = |\{u_x \in H : x \in X''\}| = |X''|. \tag{45}$$

Moreover, for $x \in X$, let $H_x = \{u \in H : (\overrightarrow{x,u}) \in D\}$. As a consequence of (13–15) in Algorithm 1 and (43),

$$|H_x| = |H_{s(x)}| = |F_{s(x)}| - \eta(F_{s(x)}).$$

Hence, for $H' = \{u : \exists x \in X' \text{ such that } (\overrightarrow{x,u}) \in D\}$ and $I' = \{s(x) : x \in X'\}$, we get

$$|H'| = |H_{I'}| = |F_{I'}| - \sum_{i \in I'} \eta(F_i). \tag{46}$$

From (44), (45) and (46), we now obtain that $|X| \leq |H' \cup H''| \leq |A(X)|$. Hence for any subset X of an independent set $Y \in \mathcal{I}(M)$, we see that

$$|X| \leq |A(X)|,$$

as all subsets of an independent set is independent. But by (18) this implies that there is a set of vertex-disjoint paths from Y to H in G, and that $Y \in \mathcal{I}(M(G))$. Consequently, we have now proven that $\mathcal{I}(M) \subseteq \mathcal{I}(M(G))$. □

Proof of Lemma IV.2. The independent sets of the gammoid associated to $G = G(F_1, \ldots, F_m; k, \rho)$ are

$$\mathcal{I}(M(G)) = \{X \subseteq E : \exists \text{ a set of } |X| \text{ vertex-disjoint paths from } X \text{ to } T\}.$$

The independent sets of the matroid $M = M(F_1, \ldots, F_m; k; \rho)$ are

$$\mathcal{I}(M) = \{X \subseteq E : |X \cap F_I| \leq \min\{|F_I| - \sum_{i \in I} \eta(F_i), k\} \text{ for each } I \subseteq [m]\},$$

by Proposition III.3 (iii). From Lemma IV.1 we know that there is a set of vertex-disjoint paths from a subset $X \subseteq E$ to H if and only if $|X \cap F_I| \leq |F_I| - \sum_{i \in I} \eta(F_i)$ for all $I \subseteq [m]$. From (5–6) in Algorithm 2, this implies that

$$\begin{aligned}
\mathcal{I}(M(G)) &= \{X \subseteq E : |X| \leq k \text{ and } |X \cap F_I| \leq |F_I| - \textstyle\sum_{i \in I} \eta(F_i) \text{ for each } I \subseteq [m]\} \\
&= \{X \subseteq E : |X \cap F_I| \leq \min\{|F_I| - \textstyle\sum_{i \in I} \eta(F_i), k\} \text{ for each } I \subseteq [m]\} \\
&= \mathcal{I}(M).
\end{aligned}$$

Consequently, the matroids $M(G)$ and M are equal, since their independent sets are the same. □