A Matroid Theory Approach to Multicast Network Coding

Xunrui Yin[†], Zongpeng Li[†], and Xin Wang[‡]

[†]Department of Computer Science, University of Calgary, Canada
[‡]School of Computer Science, Fudan University, Shanghai, China

***Abstract*—Network coding encourages the mixing of information flows at intermediate nodes of a network for enhanced network capacity, especially for one-to-many multicast applications. A fundamental problem in multicast network coding is to construct a feasible solution such that encoding and decoding are performed over a finite field of size as small as possible. Coding operations over very small finite fields (*e.g.*, $\mathbb{F}_2$) enable low computational complexity in theory and ease of implementation in practice. In this work, we propose a new approach based on matroid theory to study multicast network coding and its minimum field size requirements. Applying this new approach that translates multicast networks into matroids, we derive the first upper-bounds on the field size requirement based on the number of relay nodes in the network, and make new progresses along the direction of proving that coding over very small fields ($\mathbb{F}_2$ and $\mathbb{F}_3$) suffices for multicast network coding in planar networks.**

I. Introduction

Departing from the store-and-forward data networking practice, network coding [1] encourages coding at intermediate nodes of a network to improve the end-to-end communication throughput. Two classes of problems are typically studied in the field of network coding: single source problems and multiple-source problems. For a single source multicast session, linear network coding over a finite field is proven to achieve the maximum throughput [2]. However, such throughput improvement does come with a computation overhead associated with the encoding and decoding operations conducted across the network. A fundamental problem in network coding research is to find a feasible coding solution over a finite field as small as possible. The advantages of performing coding operations over a very small field are three-fold. First, it leads to lower computational complexity for encoding and decoding the same amount of information. Second, it is amenable for hardware and software implementations in practice. Third, perhaps less obviously, it often comes with very efficient (*e.g.* linear time) algorithms for constructing the network coding solution itself [3].

A large series of research in network coding has been devoted to the study of sufficient field sizes for multicast network coding, using different approaches and through different perspectives, relating the field size to the number of receivers [3], the size of the max-flow [4], and characteristic of the underlying network topology [5]. For general multicast networks, Sanders *et al.* [3] proved that there exists a coding solution if the field size is no smaller than $|T|$, where T is the set of receivers. However, in all known types of multicast networks, coding over a field of size $O(\sqrt{|T|})$ always suffices [6]. The gap between the upper-bound and the lower-bound has remained open for a decade. The case of multicasting two information flows then attracted considerable interest, because it is relatively easier to analyze yet fundamental. Fragouli *et al.* proved an upper-bound of $\sqrt{2|T| - 7/4} + 1/2$ on the required field size in this case. Subsequent efforts devoted to extending the results to the case of multicasting more than 2 flows, even for just 3 flows, have proven sterile. Such an extension is believed to be hard [7].

Matroid theory is a branch of mathematics developed from linear algebra and graph theory, generalizing the structure of independence relations. It has been applied to the multiple source network coding setting, usually for constructing multi-source network coding problem instances from specific matroids, resulting in desired properties in the former. For example, Dougherty *et al.* used the Fano matroid and non-Fano matroid [8] to construct a network whose coding capacity is unachievable [9], and used the Vámos matroid to construct a network to show that non-Shannon type inequality is necessary in characterizing the coding capacity [10].

Coding in a single source multicast network is fundamentally different from coding for multiple sources. Linear network coding is sufficient to achieve the maximum throughput in the former [2] but not the latter [11]. The rich body of matroid theory concepts and techniques, focusing on a central notion of "independence" as does network coding (linearly independent information flows are required for decoding), has rarely been explored for the case of single session multicast network coding.

Towards this direction, we propose a new *network matroid* approach, and apply it to study the minimum field size requirement in multicast network coding. In contrast to existing literature that applies matroid theory to network coding, which transforms matroids into networks (the *matroid network* approach) [9][10][12], our *network matroid* approach transforms multicast networks into matroids, for study as a matroid theory problem.

The network matroid paradigm connects matroid theory to multicast network coding by translating the problem of finding

This work is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the National Science Foundation of China under Grant 61171074.

a coding solution in a multicast network into a problem of finding a representable matroid satisfying properties specified by the network. Capitalizing on the mature tools and methods from matroid theory, we use this approach to unify existing proofs, to prove new upper-bounds for the minimum field size in multicast networks, and to identify classes of multicast problems for which coding over small fields suffices.

The majority of this paper focuses on *homogeneous* multicast networks, where the in-degree of each non-source node falls into two cases: 1 or $h > 1$. Homogeneous networks generalize the basic case of multicasting 2 information flows, because in the network induced by two information flows, the in-degree of each node is either 1 or 2. In addition, homogeneous networks are closely related to a practical paradigm of network coding, *terminal coding*. With terminal coding, the coding operations are allowed only at the source and receivers. Relays merely forward and replicate received messages. Terminal coding is of practical relevance for two reasons. First, in a real-world network, coding requires extra hardware capabilities and resources than replication or forwarding does. The intermediate nodes, such as routers and switches, are usually not capable of information encoding. Second, empirical studies show that terminal coding can achieve almost the same throughput as general network coding [13]. We show that to find a coding solution for terminal coding is equivalent to find a network coding solution in a homogeneous network.

For the first time in the literature of network coding, this work presents results that relate the required field size to the number of relay nodes and to the total network size, respectively. We conjecture that coding over a field of size no more than $|R| - 1$ is sufficient, where R is the set of relays. We show that this conjecture implies an important open problem on maximum distance separable (MDS) codes [14], and prove it for networks with up to 6 relays. Through careful examination of sub-network structures of the multicast network, we further prove that in many cases, the filed size can be even smaller. Employing these results, we prove that coding over $\mathbb{F}_2$ suffices for a multicast network with up to 8 nodes, and coding over $\mathbb{F}_3$ suffices for networks with up to 10 nodes. Such bounds are independent from existing bounds based the number of receivers [3][15], and cannot be derived from the latter.

We also consider network coding in planar networks, which are closely related to practical networks and practical network algorithm design [5][16]. It was conjectured that coding over small fields suffices in planar networks [16]. We use the network matroid approach to recover the existing result of multicasting 2 information flows in a planar network, and then extend it to the case of multicasting 3 information flows. For general cases, we discuss a potential proof to this conjecture using the network matroid approach.

The rest of this paper is organized as follows. We overview existing literature in Sec. II. In Sec. III, we introduce the preliminaries of multicast network coding and matroid theory. We introduce the concept of network matroid and overview the proposed approach in Sec. IV. In Sec. V, we generalize 2-minimal networks into homogeneous networks and discuss their relation to terminal coding. We prove a series of bounds on the minimum field size with regard to the number of relays and the network size in Sec. VI, and study network coding in planar networks in Sec. VII. Sec. VIII concludes the paper.

II. Related Work

Existing research on the theory of network coding can be categorized into two classes: single multicast session versus multiple unicast/multicast sessions. For a single multicast session, Li *et al.* [2] proved that linear network coding over a sufficiently large field suffices to achieve the maximum multicast throughput. The size of the finite field over which coding operations are performed directly affects the computation overhead in practical network coding applications.

Koetter and Médard [4] used an algebraic approach to upper-bound the field size by $h|T|$, where h is the minimum max-flow from the multicast source to each receiver and $|T|$ is the number of multicast receivers. This bound was then improved to $|T|$ by Sanders *et al.* [3], which is essentially the best bound known for general cases, with little improvement in a decade. For the case of $h = 2$, Fragouli *et al.* [15] proved an upper-bound of $\sqrt{2|T| - 7/4} + 1/2$ on the minimum required field size.

We earlier proved upper-bounds on the required field size by exploiting the characteristics of the underlying network topology [5]. We conjectured that the required field size is bounded by the size of the maximum clique minor of the multicast network, and proved the conjecture for the case of $h = 2$. Xiahou *et al.* [16] considered how the positions of relays in the network affect the minimum field size, and proposed efficient algorithms to compute a coding solution in planar multicast networks.

Li *et al.* [13] studied a more restricted type of terminal coding, where relay nodes are not allowed to replicate messages. Note that, terminal coding is different from the classic paradigm of source coding [17], which abstracts a network as a simple end-to-end communication medium, without catering to details in the network topology.

Previous work that combine network coding and matroid theory are usually for the case of multiple communication sessions. For example, Dougherty *et al.* utilized specific matroids to construct a network to show the insufficiency of linear network coding for multiple sessions [11], and a network whose coding capacity is unachievable [9]. El. Rouayheb *et al.* [18] and He *et al.* [12] presented approaches to construct networks from matroids, which can be applied to construct networks with interesting properties in terms of network coding solutions. Sun *et al.* [19] resorted to matroid theory to construct linear coding solutions for cyclic networks.

III. Preliminaries

A. Multicast Network Coding

We denote a network with a directed acyclic multigraph $D(V, A)$ where each link can transmit one symbol per unit time. In order to model link capacities larger than one unit, parallel links are allowed. Let $s \in V$ be the source node, and $T \subset V$ be the set of multicast receivers. A multicast rate r is achieved if each receiver can recover the message transmitted from the source at rate r. Let $\lambda(t)$ be the volume of max-flow from the source to receiver t. A celebrated result [1][2] in network coding shows that with linear network coding, we can achieve the max multicast throughput $h = \min\{\lambda(t) \mid t \in T\}$, as if there is no competition for network resources among receivers. We will use $D(V, A, s, T, h)$ to denote a multicast network.

With linear network coding, the intermediate nodes send coded symbols as linear combinations of their received symbols. Therefore, each transmitted symbol can be represented as a linear combination of the source symbols. The combination coefficient vectors are called *coding vectors*. To disseminate h source symbols to all receivers, each receiver must receive h coded symbols whose coding vectors are linearly independent.

Without loss of generality, we may assume the network is link minimal, *i.e.* removing any link will make some receiver unable to recover all h source messages. Such networks are called h-minimal networks [5][7][16].

B. Basics of Matroid Theory

A matroid can be viewed as an abstraction of the linear independence relation in vector spaces. Specially, a matroid is a pair $(E, \mathcal{I})$, where E is a finite set (called the *ground set*), and $\mathcal{I}$ is a collection of subsets of E (called *the independence family*) satisfying:

1) $\emptyset \in \mathcal{I}$, *i.e.*, the empty set is independent;
2) if $A \subset B \in \mathcal{I}$, then $A \in \mathcal{I}$, *i.e.*, every subset of an independent set is independent;
3) for two independent sets $I_1, I_2 \in \mathcal{I}$, if $|I_1| > |I_2|$, we can find an element $e \in I_1 \backslash I_2$ such that $I_2 \cup \{e\} \in \mathcal{I}$.

For simplicity, we use the element itself to represent a singleton set. For a matroid $M(E, \mathcal{I})$, a maximal independent set $B \subset E$ is a *base*, and a minimal dependent set $C \subset E$ is a *circuit*. We use $\mathcal{B}, \mathcal{C}$ to denote the base family and circuit family, respectively. Note that, we can derive the independence family $\mathcal{I}$ through $\mathcal{B}$ or $\mathcal{C}$ as $\mathcal{I} = \{I \subset E \mid \exists B \in \mathcal{B} : I \subset B\}$ or $\mathcal{I} = \{I \subset E \mid \forall C \in \mathcal{C} : C \subsetneq I\}$. Therefore, we can use the notation $M(E, \mathcal{B})$ or $M(E, \mathcal{C})$ to denote a matroid. A matroid has a number of equivalent definitions. We will often refer to the following characterization by circuit family:

Proposition 1. *Let $\mathcal{C}$ be a family of subsets of a ground set E, then there exists a unique matroid with $\mathcal{C}$ as its circuit family if and only if $\mathcal{C}$ satisfies:*
(C1): $\emptyset \notin \mathcal{C}$;
(C2): If $C_1, C_2 \in \mathcal{C}$ and $C_1 \subset C_2$, $C_1 = C_2$;
(C3): For distinct $C_1, C_2 \in \mathcal{C}$ and $x \in C_1 \cap C_2$, there exists another member of $\mathcal{C}$ contained in $(C_1 \cup C_2) - x$.

The rank function of a set of vectors can also be generalized to a matroid $M(E, \mathcal{I})$. Specifically, the rank function $r(X), X \subset E$ is defined as:

$$r(X) = \max\{|I| \mid I \subset X, I \in \mathcal{I}\}$$

The rank of the ground set $r(E)$ is also the rank of the matroid.

As matroids generalize the linear independence relation in vector spaces, a set of vectors E over a finite field $\mathbb{F}_q$ naturally defines a matroid $M(E, \mathcal{I})$ with $\mathcal{I} = \{I \subset E \mid \text{vectors in } I \text{ are linearly independent}\}$, known as a *vector matroid*. A matroid is *representable* over a finite field $\mathbb{F}_q$ if it is isomorphic to a vector matroid over a finite field $\mathbb{F}_q$. A representable matroid can be represented by a matrix whose column vectors correspond to the elements of the ground set. The matrix is called a *representation*. Note that we can shuffle basis vectors to derive different matrix representations of the same matroid.

A central problem in matroid theory is to characterize the representability of matroids. An important concept in such characterization is that of a *matroid minor*.

For a matroid $M(E, \mathcal{I})$ and a subset $X \subset E$, two types of operations can be applied to derive a new matroid over $E - X$: *deletion* and *contraction*. Let $M \backslash X$ and M/X denote the matroid derived from deleting and contracting X, respectively. They are defined as :

$$\mathcal{I}(M \backslash X) = \{I \subset E - X \mid I \in \mathcal{I}(M)\}$$

$$\mathcal{I}(M/X) = \{I \subset E - X \mid I \cup B_X \in \mathcal{I}(M)\}$$

where B_X is a maximal independent set contained in X.

For two disjoint subsets $X, Y \subset E$, it can be verified that $M/X/Y = M/(X \cup Y)$, $M \backslash X \backslash Y = M \backslash (X \cup Y)$ and that the sequence of the operations does not matter, *i.e.*, $M/X \backslash Y = M \backslash Y/X$. A matroid N is a *minor* of matroid M if we can obtain N from M by a series of contractions and deletions.

For a vector matroid, deleting an element is equivalent to removing the corresponding column of its representation matrix. Contracting an element is equivalent to first representing the matroid with the matrix where the contracted element is one of the basis vectors and then removing the row containing the unique '1' of the contracted vector. Therefore, a representable matroid remains representable after deletions and contractions. Consequently, if a matroid has a minor that is not representable over $\mathbb{F}_q$, then the matroid itself is not representable over $\mathbb{F}_q$ either. This observation provides a way to characterize the class of $\mathbb{F}_q$-representable matroids by listing their set of forbidden minors. The following lemmas are for the cases of $q = 2, 3, 4$ [8].

Lemma 1. *A matroid is representable over $\mathbb{F}_2$ if and only if it does not contain a $U_{2,4}$ minor.*

Lemma 2. *A matroid is representable over $\mathbb{F}_3$ if and only if it does not contain any of the $U_{2,5}$, $U_{3,5}$, F_7 and F_7^* minors.*

Lemma 3. *A matroid is representable over $\mathbb{F}_4$ if and only if it does not contain any of the $U_{2,6}$, $U_{4,6}$, P_6, F_7^-,$(F_7^-)^*$, P_8 and $P_8^{=}$ minors.*

Here $U_{k,n}$ denotes a uniform matroid, which is a simple but important class of matroids [12]. Specifically, the uniform matroid $U_{k,n}$ has n elements in the ground set and any subset of up to k elements are independent. For the other types of matroid minors in these lemmas, their definitions are provided in the Appendix, and can also be found in a standard textbook on matroid theory (*e.g.*, [8]). We simply note that their subscriptions represent the number of elements contained in their ground sets.

IV. Overview of the Network Matroid Approach

To connect multicast networks with matroids, we introduce the concept of a *network matroid*.

Definition 1. *For a multicast network $D(V, A, s, T, h)$, a matroid on the link set A is a* network matroid *of D if it satisfies the following properties:*
(P1) For each link e leaving a non-source node u, there is a circuit consisting of e and a subset of u's incoming links.
(P2) For each receiver $t \in T$, the set of incoming links contains a base of rank h.

Note that these properties only depend on the structure of the multicast network and are independent from its specific coding solutions. Given the concept of a network matroid, we can translate the problem of finding a linear network coding solution to a multicast network into the problem of finding a representable matroid satisfying the desired properties.

Theorem 1. *A multicast network $D(V, A, s, T, h)$ has a linear network coding solution over $\mathbb{F}_q$ if and only if there exists a network matroid for D, which satisfies conditions $P1$ and $P2$ and is representable over $\mathbb{F}_q$.*

Proof: '$\Rightarrow$' (the "only if" part): If the multicast network has a coding solution over $\mathbb{F}_q$, we construct the matroid $M(E = A, \mathcal{I})$ according to the coding solution: $\mathcal{I}(M) = \{I \subset A \mid \text{coding vectors on links } I \text{ are independent}\}$. For a link e leaving a non-source node u, if the coding vector of e is a zero vector, e itself is a circuit. Otherwise, the coding vector of e is generated as a non-zero linear combination of the coding vectors on some of the incoming links, which form a circuit with e. Thus, we can see that (P1) is satisfied. As each receiver is able to recover the source messages, there must be a set of links whose coding vectors are of full rank, which form a base, *i.e.*, (P2) is satisfied and M is a network matroid of D.

'$\Leftarrow$' (the "if" part): Let $M(E = A, \mathcal{I})$ be a matroid representable over $\mathbb{F}_q$ and satisfying (P1) and (P2). As $r(M) = h$, we assume the dimension of vectors in the representation equals h, since otherwise we may represent each vector as the coordinates with respect to a base of h vectors. Let the vectors in the representation be the coding vectors on each link. Since M satisfies (P1), each coding vector can be generated as a linear combination of the incoming links' coding vectors. Each receiver can recover the source messages from the set of coding vectors forming a base. ∎

We note that similar forms of Theorem 1 have been proved by Dougherty *et al.* [20] and Kim *et al.* [21], which mainly focus on the general case of multiple sources. Theorem 1 refines their results for the case of single source multicast.

As matroid representability over small finite fields has been explicitly characterized by forbidden minors, we propose the network matroid approach for proving the sufficiency of coding over small fields in different types of multicast networks. Generally speaking, such a proof under the network matroid approach consists of four steps, which are shown in Table 1.

Table 1: The General Network Matroid Approach

1: *Given that the multicast network is solvable over a sufficiently large field, there exists a representable network matroid M;*
2: *If M contains a forbidden minor, eliminate the minor through techniques such as introducing new circuits to the matroid, while maintaining the invariant that M satisfies properties (P1) and (P2);*
3: *Repeat step 2 until M does not contain any forbidden minors;*
4: *Apply Lemma 1∼3 to prove that M is representable over a small finite field, and the coding solution can be derived from M's representations according to Theorem 1.*

V. Generalization of 2-minimal Networks

A fundamental and extensively studied scenario in multicast network coding is the case of 2-minimal networks, *i.e.*, the case of multicasting 2 information flows. A number of interesting results that are hard to prove for general networks are proven in the 2-minimal case. For example, Fragouli *et al.* used the technique of information flow decomposition to derive a tight bound of $\sqrt{2|T| - 7/4} + 1/2$ on the required field size for 2-minimal multicast networks, and Xiahou *et al.* proved that coding over $\mathbb{F}_3$ is sufficient for 2-minimal multicast networks that are planar [16].

In this section, we generalize the study of 2-minimal networks to homogeneous networks where h may be larger than 2. Specifically, a homogeneous network is defined as follows:

Definition 2. *An h-minimal multicast network is* homogeneous *if the in-degrees of non-source nodes are either 1 or h. The nodes of in-degree 1 are called relays and the nodes of in-degree h are called receivers.*

All 2-minimal networks are special cases of homogeneous networks with $h = 2$. This is because in an h-minimal network, every node's in-degree is no more than h [22]. Thus, the in-degree of a non-source node is either 1 or 2 in 2-minimal networks.

The fundamental class of combination networks [23] in the network coding literature form another special class of homogeneous networks. In a combination network $C(n, k)$, the source is connected to n relays, and for each set of k relays, there is a unique receiver connecting to them.

Besides generalizing existing classes of networks, homogeneous networks are also related to a practical paradigm of network coding called *terminal coding*, where source and receivers are allowed to encoding and decode, and relays are allowed to replicate and forward messages only. Terminal coding is of important practical relevance for two reasons. First, in a realworld network, coding requires much more hardware complexity and computation resource than replication and forwarding. Thus, the intermediate nodes, such as routers and switches, are usually not capable of encoding operations. Second, empirical studies show that terminal coding can achieve almost the same throughput as general network coding

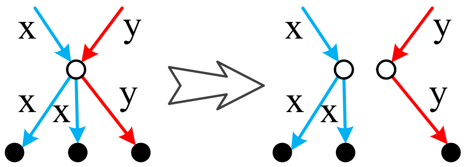


Fig. 1. Converting a multicast network in to a homogeneous network with respect to the flow solution for terminal coding.

does, in most practical situations where the network topology is not highly contrived [13]. A terminal coding solution has two components: 1) a flow transmission scheme that specifies how the relays replicate and forward the messages; and 2) a coding solution that specifies how to assign coding vectors to each flow. The relation between homogeneous network and terminal coding is that, once the flow solution is determined, the network can be modified into a homogeneous network by splitting each relay into several relays of in-degree 1 according to the flow solution (Fig. 1). There is a terminal coding solution over $\mathbb{F}_q$ for the terminal coding problem if and only if there is a coding solution over $\mathbb{F}_q$ in the converted homogeneous network.

In a homogeneous network, we can assume that there are no adjacent relays, because if there is a link from a relay u to another relay v, we may contract this link to make all outgoing links of v directly leaving u. Such contractions do not affect the solution of this network, since each relay has only one incoming link, the coding vectors on v's outgoing links must be the same as the coding vector of v's incoming link, which is the same as u's incoming link for the same reason.

The key feature of homogeneous networks is that we have less restrictions on determining the coding vectors. Each coding terminal should now be able to decode and recover all source messages, and consequently should be able to re-encode and generate any new coding vector desired.

Meanwhile, Property (P1) translates to that each outgoing link of a relay node and its incoming link must form a circuit. Two elements of a matroid are *parallel* if they form a circuit. The two parallel elements have the same role in the matroid, because if $\{e_1, e_2\} \in \mathcal{C}$, for any circuit C that contains e_2, $C \backslash \{e_2\} \cup \{e_1\}$ is also a circuit according to property (C3).

This motivates us to simplify properties (P1) and (P2) for homogeneous networks.

Theorem 2. *A homogeneous network $D(V, A)$ has a coding solution over $\mathbb{F}_q$ if and only if there is an $\mathbb{F}_q$-representable matroid $M(R, \mathcal{B})$ with the set of relays R as its ground set and satisfying:*

(S1) $\mathcal{B} \supset \mathcal{B}_0$, where $\mathcal{B}_0 = \{B_t \subset R \mid B_t$ is the set of relays that have outgoing links to receiver $t \in T\}$, i.e. *$\mathcal{B}_0$ is the collection of subsets of R that are connected to the same receiver.*

Proof: Given Theorem 1, it suffices to prove that there is a matroid defined over the link set A and satisfying (P1) and (P2) if and only if there is a network defined over the set of relays R and satisfying (S1).

For a network matroid $M(A, \mathcal{B})$, construct a matroid $M'(R, \mathcal{B}')$ by mapping each relay to one of its adjacent links. Properties (P1) and (P2) imply that each set $B_t \in \mathcal{B}_0$ is independent in M'.

For a matroid $M'(R, \mathcal{B}')$, construct the network matroid $M(A, \mathcal{B})$ as $\mathcal{C}(M) = \{C \subset A|$ links in C are adjacent to distinct relays that form a circuit in $M'\} \cup \{\{e_1, e_2\}|e_1 e_2$ are adjacent to the same relay$\}$. It can be verified that M satisfies (P1) and (P2). ■

We also refer to the matroid defined on the set of relays R and satisfying (S1) as the network matroid.

VI. Bounding Field Size by the Number of Relays

In this section, we upper-bound the field size required for multicast network coding from a new perspective: the number of relays in the multicast network. Specifically, we first propose the following conjecture.

Conjecture 1. *For a homogeneous multicast network $D(V, A)$, there exists a network coding solution over $\mathbb{F}_q$ if the number of relays in D is no more than $q + 1$.*

We point out that this conjecture partially implies a major conjecture on maximum distance separable (MDS) codes. A (n, k)-MDS code $(n \geq k)$ is a linear code that encodes k symbols into n symbols such that the source symbols can be recovered from any k coded symbols. The well-known MDS conjecture can be stated as

The MDS Conjecture [14] A (n, k)-MDS code exists over $\mathbb{F}_q$ if $q \geq n - 1$. Furthermore, if q is even and $q \geq n - 2$, there exists a (n, k)-MDS code for $k = 3$ and $n - k = 3$.

Conjecture 1 implies the first part of the MDS Conjecture, due to the following argument. If Conjecture 1 is true, then there is a coding solution over $\mathbb{F}_q$ for combination networks $C(n, k), q \geq n - 1$, which have n relays. This implies we can encode k symbols in to n symbols as the symbols sent to the n relays in the coding solution. The MDS Conjecture has remained open for a long time and is believed to be hard. Consequently, Conjecture 1 in its general form may be hard to prove too. We next prove a number of special cases of Conjecture 1, which are exactly the cases related to the smallest finite fields.

Theorem 3. *For a homogeneous multicast network $D(V, A)$ with up to 3,4,5 relays, coding over $\mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_4$ suffices, respectively.*

Proof: As each relay has a single incoming link, the coding vectors on its outgoing links must be identical to the coding vector on the incoming link. To find a coding solution is equivalent to assign to each relay a coding vector such that every set of relays accessed by the same receiver have linearly independent coding vectors. Let $\mathcal{B}_0$ denote the set of subsets of relays connected to the same receiver.

We regard the set of relays as the ground set of the network matroid. As the multicast network is solvable over a sufficiently large field, we first create a network matroid $M(E, \mathcal{B})$ as the vector matroid of the coding solution. As each receiver is able to recover the source message, all sets in $\mathcal{B}_0$ must be bases of M, *i.e.*, $\mathcal{B}_0 \subset \mathcal{B}$.

If there are no more than 5 relays, matroid M has 5 elements at most, which excludes the forbidden minors for $\mathbb{F}_4$ representability, since they have 6 elements at least. According to lemma 3, M is representable over $\mathbb{F}_4$. Therefore, we may assign coding vectors to relays according to the $\mathbb{F}_4$ representation of M. As $\mathcal{B}$ is not changed in switching representations, we conclude that all receivers can recover the source message, which implies that the multicast network has a coding solution over $\mathbb{F}_4$.

Applying similar techniques to the case of 4 relays, we conclude that coding over $\mathbb{F}_3$ suffices for homogeneous networks with 4 relays, since all forbidden minors of $\mathbb{F}_3$ representability have 5 elements at least.

Furthermore, coding over $\mathbb{F}_2$ suffices for homogeneous networks with 3 relays, since the forbidden minor of $\mathbb{F}_2$ representability have 4 elements. ■

The bounds given in Theorem 3 are tight. The following example (Fig. 2) shows an homogeneous network with 4 relays that requires coding over $\mathbb{F}_3$.

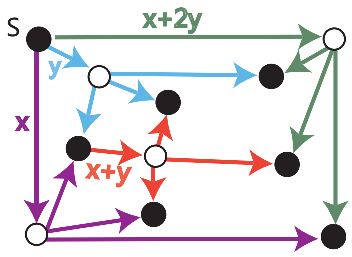


Fig. 2. An example homogeneous network with 4 relays (white nodes), requiring coding over $\mathbb{F}_3$. The five black nodes include a multicast source and four receivers. Two unit information flows x and y are to be multicast from the source to the four receivers.

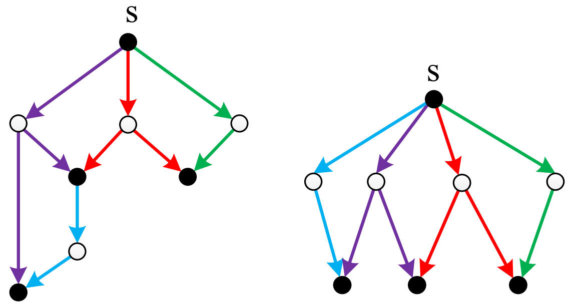
According to the relationship between homogeneous networks and terminal coding, we derive the following corollary.

Corollary 1. *In a multicast network D, if the sum of relays' in-degrees does not exceed 3, 4, 5, coding over $\mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_4$ is sufficient for terminal coding, respectively.*

Proof: Assume there is a flow solution indicating how each relay forwards messages under terminal coding, we construct the corresponding homogeneous network H for the coding solution as discussed in section V. As in the homogeneous network, the number of relays generated from splitting a relay u in D does not exceed u's in-degree, the number of relays in the homogeneous network does not exceed the sum of all relays' in-degrees in D. According to Theorem 3, we conclude that there is a coding solution over the corresponding finite fields. ■

Theorem 3 ignores the inner structure of the homogeneous networks. We next derive more refined results by considering the sub-network structure of a homogeneous network, leading to new bounds on the required field size through the number of receivers.

For a given homogeneous network, we may turn it into a sub-network of $C(n, k)$ by replacing each link leaving a receiver node with a node v, with a link from the source to v, as illustrated in Fig. 3. We refer to such a sub-network of a combination network as a *quasi-combination* network. It can be verified that a homogeneous network is solvable over $\mathbb{F}_q$ if and only if the corresponding quasi-combination network is solvable over $\mathbb{F}_q$.



(a) A homogeneous network (b) A sub-network of a combination network

Fig. 3. Translating a homogeneous network into a sub-network of a combination network.

Theorem 4. *For a homogeneous multicast network $D(V, A)$ with up to 5 relays whose quasi-combination network excludes a $C(5, 2)$ sub-network, coding over $\mathbb{F}_3$ suffices.*

Proof: Because the network $D(V, A)$ is solvable over a sufficiently large field, there is a network matroid M for D. As there are 5 elements at most, M has a $U_{2,5}$ or $U_{3,5}$ minor only if $M = U_{2,5}$ or $M = U_{3,5}$.

If $M = U_{2,5}$, the network must be 2-minimal since $r(M) = 2$. As it does not contain a $C(5, 2)$ sub-network, there are two relays e_1, e_2 with no receiver connecting to both of them. We can modify M into $U_{2,4}$ with a pair of parallel elements corresponding to e_1 and e_2. The circuit family is $\mathcal{C} = \{C \subset E \mid |C| = 3 \wedge |C \cap \{e_1, e_2\}| \leq 1\} + \{e_1, e_2\}$. $\mathcal{C}$ satisfies property (C1–C3). As any set $B \in \mathcal{B}_0$ contains 2 elements, B does not contain any circuits. Therefore, this new matroid contains all sets in $\mathcal{B}_0$ as bases.

If $M = U_{3,5}$, the network must be 3-minimal and each receiver connects to 3 relays. We claim that there are three relay nodes with no receiver connecting to them, since otherwise, we can find a $C(5, 3)$ sub-network, which contains $C(5, 2)$ as a sub-network. Let e_1, e_2, e_3 be the three elements in E corresponding to the three relays. Circuits in $U_{3,5}$ include all possible subsets of 4 elements. We replace the two circuits containing e_1, e_2, e_3 with a single circuit (e_1, e_2, e_3). Specifically, $\mathcal{C} = \{C \subset E \mid |C| = 4 \wedge |C \cap \{e_1, e_2, e_3\}| \leq 2\} + \{e_1, e_2, e_3\}$. It can be verified that $\mathcal{C}$ satisfies property (C1–C3). As any set $B \in \mathcal{B}_0$ contains 3 elements, B does not contain any circuits. Therefore, this new circuit collection defines a matroid which contains all sets in $\mathcal{B}_0$ as bases.

In summary, if M contains a $U_{2,5}$ or $U_{3,5}$ minor, we can find a new matroid M' excluding these two minors but containing all sets in $\mathcal{B}_0$. Further, M' does not contain F_7 or F_7^* minors, since M' has 5 elements at most. Thus, M' is representable over $\mathbb{F}_3$. ■

Theorem 5. *For a homogeneous multicast network with up to 4 relays, if its corresponding quasi-combination network excludes a $C(4, 2)$ sub-network, coding over $\mathbb{F}_2$ suffices.*

Proof: The proof is similar to the proof of theorem 4, except that we only need to consider the $U_{2,4}$ minor instead of $U_{2,5}$ and $U_{3,5}$ minors. ■

Theorem 6. *For a homogeneous multicast network with up to 3 relays, if its corresponding quasi-combination network excludes a $C(3,2)$ sub-network, routing suffices.*

Proof: Network coding is unnecessary if and only if there is a network matroid that can be decomposed into the direct-sum of rank 1 uniform matroids. Let M be the network matroid. As there are 3 elements at most, $r(M) \leq 3$.

If $r(M) = 2$, there must be two relays e_1, e_2 with no common receivers, since the quasi-combination network does not contain a $C(3,2)$ sub-network. Let e_1, e_2 be the two elements of $U_{1,2}$ and the other element (if any) be the element of $U_{1,1}$, we can see that $U_{1,2} \oplus U_{1,1}$ is a network matroid for this multicast network.

If $r(M) = 3$, every receiver must connect to the 3 relays. Thus, network coding is unnecessary. ■

Corollary 2. *For a homogeneous network with up to 5 relays and 9 receivers, coding over $\mathbb{F}_3$ suffices. For a homogeneous network with up to 4 relays and 5 receivers, coding over $\mathbb{F}_2$ suffices.*

Proof: Theses two statements can be proved by combining Theorem 4, Theorem 5 and the fact that the $C(5,2)$ and $C(4,2)$ sub-networks contain $\binom{5}{2} = 10$ receivers and $\binom{4}{2} = 6$ receivers, respectively. ■

We note that these bounds can not be derived from the hitherto best bounds due to Fragouli *et al.* of $\sqrt{2|T| - 7/4} + 1/2$. First, their bound only applies to the case of multicasting two information flows, while our bounds apply to networks multicasting an arbitrary number of flows. Second, for the case of 9 receivers and 5 receivers, their characterization ensures the existence of coding solution over $\mathbb{F}_4$ and $\mathbb{F}_3$ respectively, which are larger (looser) than our bounds of $\mathbb{F}_3$ and $\mathbb{F}_2$.

Corollary 3. *For a homogeneous network with up to 10 nodes, coding over $\mathbb{F}_3$ suffices. For a homogeneous network with up to 8 nodes, coding over $\mathbb{F}_2$ suffices.*

Proof: Recall that $|T|$ is the number of receivers in the network, and let $|R|$ denote the number of relays.

For the sufficiency of $\mathbb{F}_3$: Since the network is composed of a source node, relays and receivers, $1 + |T| + |R| \leq 10$. As coding over a field of size larger than the number of receivers is always sufficient, we only need to consider the case of $|T| \geq 4$ receivers, which implies $|R| \leq 5$. According to Theorem 3, coding over $\mathbb{F}_3$ suffices for $|R| \leq 4$, leaving the only case of $|R| = 5, |T| \leq 4$. According to Corollary 2, coding over $GR(3)$ suffices for this case.

The case of $\mathbb{F}_2$ is similar. ■

Parallel to Conjecture 1, we conjecture that the following general case proposition holds.

Conjecture 2. *For a prime power q and a homogeneous network with up to $q+2$ relays, if its corresponding quasi-combination network excludes a $C(q+2,2)$ sub-network, coding over $\mathbb{F}_q$ suffices.*

The network matroid approach can be further applied to derive the following characterization for multicast networks with up to 5 relays.

Theorem 7. *For a homogeneous network of up to 5 relays and 4 receivers, coding over $\mathbb{F}_2$ suffices.*

Proof: As the network is solvable over a sufficiently large field, there exists a network matroid $M(R, \mathcal{B})$ for this homogeneous network. If M contains a $U_{2,4}$ minor, we aim to find a new matroid M' excluding a $U_{2,4}$ minor but contains all sets in $\mathcal{B}_0$ as bases. As there are 5 elements in M at most, $r(M) \leq 3$, since we can contract or delete at most one element to obtain the $U_{2,4}$ minor, whose rank is 2. As there are up to 4 receivers, the corresponding quasi-combination network does not contain $C(4,2)$ sub-network. According to Theorem 5, we only need to consider the case of 5 relay nodes.

If $r(M) = 2$, the $U_{2,4}$ minor must be obtained by deleting an element. Denote the element as x and the corresponding relay node as v. As the quasi-combination network does not contain the $C(4,2)$ sub-network, there are at least two pairs of relay nodes in $V \backslash \{v\}$ with no receivers connecting to both of them. Let e_1, e_2 be one of the pairs in R. We add a circuit (e_1, e_2) to M to obtain a new matroid M'. Specifically, its circuit set $\mathcal{C}' = \{\{e_1, e_2\}\} \cup \{C \in \mathcal{C}(M) | e_1, e_2 \notin C\} \cup \{C - e_1, C - e_2 | e_1, e_2 \in C \in \mathcal{C}(M)\} \cup \{C, C - e_1 + e_2 | C \in \mathcal{C}(M), e_1 \in C, e_2 \notin C\} \cup \{C, C - e_2 + e_1 | C \in \mathcal{C}(M), e_2 \in C, e_1 \notin C\}$. As $r(M) = 2$, any set of 3 elements or more is dependent and any two elements in $U_{2,4}$ are independent, there is at most one new circuit of length 2 in $\mathcal{C}' \backslash \mathcal{C}$ besides $\{e_1, e_2\}$, which must include the deleted element x. Let $\{e_1, x\}$ be the new circuit, it must be introduced to $\mathcal{C}'$ by the circuit $\{e_2, x\} \in \mathcal{C}$. If $\{e_1, x\} \in \mathcal{B}_0$, we can use the other pair of relays in $U_{2,4}$ that have no common receivers as circuits and repeat the operation to obtain a matroid $M'(R, \mathcal{B}')$ satisfying $\mathcal{B}_0 \subset \mathcal{B}'$.

If $r(M) = 3$, the network is 3-minimal. According to the following theorem,

Schum Theorem [8] If N is a minor of $M(R, \mathcal{B})$, there exist disjoint sets $X, Y \subset R$ such that $N = M/X \backslash Y$ and $r(M/X) = r(N)$.

the $U_{2,4}$ minor can be obtained by contracting an element of M. Denote the contracted element as x and the set of elements in $U_{2,4}$ as E. As any two elements in $U_{2,4}$ are independent, any circuit in M contains at least 3 elements. Since there are up to 3 receivers, each of which is adjacent to 3 relays, there exist two elements $e_1, e_2 \in R$ such that there is no receiver adjacent to e_1, e_2, x simultaneously. We add a new circuit $\{e_1, e_2, x\}$ to the circuit family $\mathcal{C}$ and remove the non-minimal dependent set to obtain $\mathcal{C}'$. According to the construction and the fact that $\mathcal{C}$ is a circuit family, $\mathcal{C}'$ satisfies (C1) and (C2). To show that $\mathcal{C}'$ also satisfies (C3), consider a circuit $C \in \mathcal{C}' - \{e_1, e_2, x\}$. If $e \in C \cap \{e_1, e_2, x\}$, $|(C \cup \{e_1, e_2, x\}) - e| \geq |C| + 3 - 1 \geq 5$. Because $r(M) = 3$, $(C \cup \{e_1, e_2, x\}) - e$ contains a circuit. Thus, $\mathcal{C}'$ satisfies (C3) and determines a matroid M'. As $\{e_1, e_2, x\}$ is dependent in M', we can see that the original $U_{2,4}$ is minor is eliminated. If M' still contains $U_{2,4}$ minor that is derived by contracting another element, we can repeat adding circuits to the matroid. Since the number of length 3 circuits increases in each step, we will eventually obtain a network matroid that excludes a $U_{2,4}$ minor. ■

VII. Network Coding in Planar networks

Planar networks are networks whose underlying topology is a planar graph, *i.e.*, a graph that can be drawn in a 2D plane without intersecting edges. The following theorem, which considers multicasting 2 information flows in planar networks, was proved in recent network coding literature [5][16]. We show that the same result can be recovered and then generalized using the network matroid approach.

For a 2-minimal network, the interference graph is an undirected graph defined on the set of relays: two relays are adjacent if they connect to a common receiver.

Theorem 8. *Coding over $\mathbb{F}_3$ is sufficient for planar 2-minimal networks.*

Proof: It is sufficient to construct a network matroid representable over $\mathbb{F}_3$. We construct such a matroid M by constructing its circuit family $\mathcal{C}$.

First, with a given planar embedding, we may divide the plane into regions such that each region contains a relay. Note that each regions's boundary nodes are either the source or receivers. As each receiver receives from 2 relays, we may require that the two relays' regions are adjacent.

Then we color these regions such that adjacent regions have different colors. According to the Four Color Theorem [24], these regions can be properly colored with 4 colors only. We construct $\mathcal{C}$ with respect to the colors: Add a circuit $\{e_1, e_2\}$ to $\mathcal{C}$ for any two relays e_1, e_2 having the same color, and for any 3 relays e_1, e_2, e_3 with different colors, add the circuit $\{e_1, e_2, e_3\}$ to $\mathcal{C}$. $\mathcal{C}$ satisfies (C1∼C3), and therefore, determines a matroid M on the ground set of relays.

For any $B \in \mathcal{B}_0$, the two relays contained in B have different colors. Therefore, B does not contain any circuits, and M is a network matroid for the planar network. As parallel elements in M can be regarded as one element, M contains only 4 distinct elements, preventing the $U_{2,5}, U_{3,5}, F_7, (F_7)^*$ minors. Consequently, Lemma 2 assures that M is representable over $\mathbb{F}_3$. ■

We next proceed to prove the case of $h = 3$, which is unknown in existing literature. Indeed, it has been believed that generating from the case $h = 2$ to $h = 3$ is hard [7].

Theorem 9. *For the case of $h = 3$ (i.e., 3-minimal networks), coding over $\mathbb{F}_2$ is sufficient for planar homogeneous multicast networks.*

Proof: It is sufficient to construct a network matroid representable over $\mathbb{F}_2$. We construct such a matroid M by constructing its circuit family $\mathcal{C}$.

First, with a given planar embedding of the network, we may divide the plane into regions such that each region contains a relay. Note that each regions's boundary nodes are either the source or receivers. As each receiver receives from 3 relays, we stretch the receiver into a "Y" pattern (as shown in Fig. 4), so that the three relay regions are adjacent to each other.

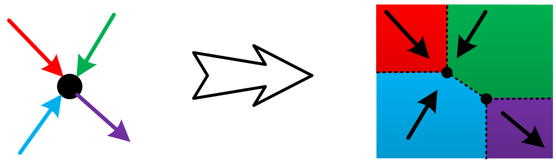


Fig. 4. Dividing the plane into regions such that each region contains a relay and the 3 regions with relays connected to the same receiver are adjacent to each other.

Then we color these regions such that adjacent regions have different colors. According to the Four Color Theorem, these regions can be colored with 4 colors. We construct $\mathcal{C}$ with respect to the colors: Add a circuit $\{e_1, e_2\}$ to $\mathcal{C}$ for any two relays e_1, e_2 having the same color, and for any 4 relays e_1, e_2, e_3, e_4 with different colors, add the circuit $\{e_1, e_2, e_3, e_4\}$ to $\mathcal{C}$. It can be verified that $\mathcal{C}$ satisfies (C1∼C3), and therefore, determines a matroid M on the ground set of relays.

For any $B \in \mathcal{B}_0$, the three relays contained in B have different colors. Therefore, B does not contain any circuits, and M is a network matroid for the planar network. As parallel elements in M can be regarded as one element, we can see that M is identical to $U_{3,4}$ matroid with parallel elements. As $U_{3,4}$ excludes the $U_{2,4}$ minor, M is representable over $\mathbb{F}_2$ according to Lemma 1.

■

Conjecture 3. *For a planar homogeneous network, coding over $\mathbb{F}_3$ suffices.*

A possible approach to prove this conjecture is to use induction on the number of flows h. We have proved the basic case of $h = 2$ and $h = 3$. For the induction step, we may assume coding over $\mathbb{F}_3$ suffices for h-minimal networks to prove the case of $h + 1$ by way of contradiction. For a $(h + 1)$-minimal network, assume all of its h-minimal subnetworks are solvable over $\mathbb{F}_3$, which implies the existence of $\mathbb{F}_3$ representable network matroids. Then try to show that there is a network matroid for the $(h+1)$-minimal network, whose every minor derived by contacting or deleting one element corresponds to a sub-network's matroid which is representable over $\mathbb{F}_3$. This would contradict the fact that the network matroid is not representable over $\mathbb{F}_3$, leading to a proof to Conjecture 3.

VIII. Conclusion

In this paper, we proposed the network matroid approach for studying multicast network coding, including in particular the minimum field size requirement for a multicast code. The network matroid approach was shown to be powerful for identifying networks where coding over small fields suffices. Applying this new approach that translates multicast networks into matroids for studying with mature tools in matroid theory, we prove new results on the sufficiency of coding over small fields in networks with a bounded number of relays and in planar networks. In particular, for the first time in the literature of network coding, we derive new upper-bounds on the field size of a multicast coding solution based on the number of relays in the multicast network. Further progress was also made along the direction of proving that coding over $\mathbb{F}_3$ suffices for all planar multicast networks.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[2] S.-Y. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[3] P. Sanders, S. Egner, and L. Tolhuizen, "Polynomial time algorithms for network information flow," in *Proceedings of the fifteenth annual ACM symposium on Parallel algorithms and architectures*, 2003.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE Transactions on Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[5] X. Yin, Y. Wang, X. Wang, X. Xue, and Z. Li, "A graph minor perspective to network coding: Connecting algebraic coding with network topologies," in *Proceedings of IEEE INFOCOM*, 2013.

[6] A. R. Lehman and E. Lehman, "Complexity classification of network information flow problems," in *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, ser. SODA '04, 2004, pp. 142–150.

[7] S. el Rouayheb, C. Georghiades, and A. Sprintson, "Network coding in minimal multicast networks," in *Proceedings of Information Theory Workshop*, Mar. 2006.

[8] J. Oxley, *Matroid Theory*. New York: Oxford University Press, 1992.

[9] R. Dougherty, C. Freiling, and K. Zeger, "Unachievability of network coding capacity," *IEEE Transactions on Information Theroy*, vol. 52, no. 6, pp. 2365–2372, 2006.

[10] ——, "Networks, matroids, and non-shannon information inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, 2007.

[11] ——, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745–2759, 2005.

[12] M. He, Z. Li, C. Wu, and X. Yin, "On uniform matroidal networks," in *IEEE International Symposium on Information Theory*, 2013.

[13] Z. Li, B. Li, D. Jiang, and L. C. Lau, "On achieving optimal throughput with network coding," in *Proceedings of IEEE INFOCOM*, 2005.

[14] J. Hirschfeld, "The main conjecture for mds codes," in *Cryptography and Coding*, ser. Lecture Notes in Computer Science, C. Boyd, Ed. Springer Berlin Heidelberg, 1995, vol. 1025, pp. 44–52.

[15] C. Fragouli and E. Soljanin, "Information flow decomposition for network coding," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 829–848, Mar. 2006.

[16] T. Xiahou, Z. Li, and C. Wu, "Information multicast in (pseudo-)planar networks: Efficient network coding over small finite fields," in *Proceedings of the 2013 IEEE International Symposium on Network Coding*, 2013.

[17] R. M. Gray, *Source coding theory*. Springer, 1990, vol. 83.

[18] S. El Rouayheb, A. Sprintson, and C. Georghiades, "On the index coding problem and its relation to network coding and matroid theory," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3187–3195, 2010.

[19] Q. Sun, S. T. Ho, and S.-Y. Li, "On network matroids and linear network codes," in *IEEE International Symposium on Information Theory*, 2008.

[20] R. Dougherty, C. Freiling, and K. Zeger, "Matroidal networks," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2007.

[21] A. Kim and M. Medard, "Scalar-linear solvability of matroidal networks associated with representable matroids," in *Proceedings of the 6th IEEE International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*, 2010.

[22] X. Yin, Y. Wang, Z. Li, X. Wang, and X. Xue, "A graph minor perspective to multicast network coding," submitted to IEEE Transactions on Information Theory, 2013.

[23] S. Maheshwar, Z. Li, and B. Li, "Bounding the coding advantage of combination network coding in undirected networks," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 570 – 584, Feb. 2012.

[24] R. Diestel, *Graph Theory*. Springer-Verlag, 1997.

APPENDIX

Several specific matroids appear in Lemmas $1 \sim 3$ as forbidden minors of $\mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_4$ representability. Here we present the definitions of these matroids except the uniform matroids $U_{k,n}$, which have been explained.

Definition of the F_7 matroid. F_7 is representable over $\mathbb{F}_2$ with the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Definition of the F_7^* matroid. F_7^* is representable over $\mathbb{F}_2$ with the following matrix representation:

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Definition of the F_7^- matroid. F_7^- is representable over $\mathbb{F}_3$ with the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Definition of the $(F_7^-)^*$ matroid. $(F_7^-)^*$ is representable over $\mathbb{F}_3$ with the following matrix representation:

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Definition of the P_6 matroid. P_6 is representable over $\mathbb{F}_5$ with the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 0 & 1 & 3 \end{bmatrix}$$

Definition of the P_8 matroid. P_8 is representable over $\mathbb{F}_3$ with the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 & 1 & 0 \end{bmatrix}$$

Definition of the $P_8^=$ matroid. $P_8^=$ is representable over $\mathbb{F}_5$ with the following matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 & 4 & 0 & 4 \\ 0 & 0 & 0 & 1 & 1 & 2 & 1 & 0 \end{bmatrix}$$