Insufficiency of Linear Coding in Network Information Flow

Randall Dougherty, Christopher Freiling, and Kenneth Zeger, *Fellow, IEEE*

***Abstract*—It is known that every solvable multicast network has a scalar linear solution over a sufficiently large finite-field alphabet. It is also known that this result does not generalize to arbitrary networks. There are several examples in the literature of solvable networks with no scalar linear solution over any finite field. However, each example has a linear solution for some vector dimension greater than one. It has been conjectured that every solvable network has a linear solution over some finite-field alphabet and some vector dimension. We provide a counterexample to this conjecture. We also show that if a network has no linear solution over any finite field, then it has no linear solution over any finite commutative ring with identity. Our counterexample network has no linear solution even in the more general algebraic context of modules, which includes as special cases all finite rings and Abelian groups. Furthermore, we show that the network coding capacity of this network is strictly greater than the maximum linear coding capacity over any finite field (exactly 10% greater), so the network is not even asymptotically linearly solvable. It follows that, even for more general versions of linearity such as convolutional coding, filter-bank coding, or linear time sharing, the network has no linear solution.**

***Index Terms*—Asymptotics, flows, linear coding, network information theory, routing.**

I. Introduction

In the context of network information theory [1], [18], a *network* is a directed acyclic multigraph, some of whose nodes are sources or sinks. Associated with the sources are *messages* and associated with the sinks are *demands*.[1] The demands at each sink are a subset of all the messages of all the sources. Each directed edge (u, v) in a network carries information from node u to node v. The goal is for each sink to deduce its demanded messages from its in-edges by having information propagate from the sources through the network. A *multicast network* is a network with exactly one source and such that each sink demands all of the source's messages.

A network's messages are assumed to be arbitrary elements of a fixed finite alphabet. At any node in the network, each out-edge carries an alphabet symbol which is a function (called an *edge function*) of the symbols carried on the in-edges to the node, or a function of the node's messages if it is a source. Also, each sink has *demand functions* for each of its demands, which attempt to deduce the node's demands from its inputs. A network *code* is a collection of edge functions, one for each edge in the network, and demand functions, one for each demand of each node in the network. A *solution* is a code which results in every sink being able to deduce its demands from its demand functions, and a network that has a solution is called *solvable*. It was noted by Ahlswede, Cai, Li, and Yeung [1] that for some networks, coding can achieve solutions that are otherwise unachievable using only routing or switching.

One way of modeling multiple uses of a network is to view each network edge as carrying a vector of alphabet symbols. For a network code using vector transmission, the out-edge of each node carries a vector of alphabet symbols which is a function of the vectors carried on the in-edges to the node, or a function of the node's message vectors if it is a source. Also, each source has a vector of messages and each sink demands a subset of all the source vector messages. All edge vectors are assumed to have the same dimension n and all message vectors are assumed to have the same dimension k. Note that the definition of a solution is with respect to the case when $k = n$. If there is a solution with $k = n = 1$, the solution is said to be *scalar*. For general k and n, a code that allows the sink nodes to deduce their demands is called a (k, n) *fractional coding solution*.

For a network alphabet with an algebraic structure (such as a ring or field), a fractional coding solution is said to be *linear* if all edge functions and all demand functions are linear combinations of their vector inputs, where the coefficients are matrices over the alphabet. That is, in a linear solution, if a node has in-edges and/or source messages carrying vectors $x_1, \ldots, x_r \in \mathcal{A}^k \cup \mathcal{A}^n$, then an out-edge of the node carries a vector

$$y = \sum_{i=1}^{r} M_i x_i$$

where each matrix M_i has elements in the alphabet $\mathcal{A}$, and is of dimension $n \times k$ when x_i is a source message and is of dimension $n \times n$ when x_i is an in-edge. A demand function is linear if it has an identical form as the equation for y, but with the number of rows in each matrix equal to k instead of n.

Manuscript received March 2, 2004; revised January 7, 2005. This work was supported by the Institute for Defense Analyses, the National Science Foundation, and Ericsson.

R. Dougherty is with the Center for Communications Research, San Diego, CA 92121-1969 USA (e-mail: rdough@ccrwest.org).

C. Freiling is with the Department of Mathematics, California State University, San Bernardino, San Bernardino, CA 92407-2397 USA (e-mail: cfreilin@csusb.edu).

K. Zeger is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA (e-mail: zeger@ucsd.edu).

Communicated by M. Médard, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2005.851744

[1] Here we use the terms "source" and "sink" in the graph-theoretic sense, namely, nodes with no in-edges and no out-edges, respectively. More generally, a network can have messages and demands associated with non-sources and non-sinks, respectively, but such a network can always be converted into a network satisfying our given definition, without altering the network solvability properties. Also, some definitions of a network allow directed cycles.

The coding capacity of a network with respect to an alphabet $\mathcal{A}$ and a class $\mathcal{C}$ of network codes (e.g., see [2] and a related definition in [18, p. 339]) is

$$\sup\{k/n : \exists\ (k,n) \text{ fractional coding solution in } \mathcal{C} \text{ over } \mathcal{A}\}.$$

If $\mathcal{C}$ consists of all network codes, then we simply refer to the above quantity as the *coding capacity* of the network with respect to $\mathcal{A}$. The following result was recently shown in [2].

Lemma I.1: The coding capacity of a network is independent of the alphabet size.

The linear coding capacity is the coding capacity when $\mathcal{C}$ consists of all fractional linear codes. Whereas the coding capacity of a network is known to be independent of the alphabet size [2], the linear coding capacity of a network does in general depend on the alphabet size chosen (e.g., see Theorems IV.3 and IV.4). We say that a class of network codes is *sufficient* over a class of alphabets if every solvable network has a solution in the class of codes over some member of the alphabet class. A network is *asymptotically solvable* with respect to an alphabet and a class of codes if its coding capacity is at least 1. We say that a class of network codes is *asymptotically sufficient* over a class of alphabets if every solvable network is asymptotically solvable in the class of codes over some member of the alphabet class.

In this paper, we first show that network linear codes are insufficient over finite field alphabets (Theorem II.4), and then over commutative ring alphabets (Corollary III.2), and even over the general class of alphabets consisting of R-modules[2] (Theorem III.4). Finally, we show that network linear codes are asymptotically insufficient over finite field alphabets (Corollary IV.6). Interestingly, a single network is used to establish all four of these counterexamples. Also, we compute the exact network coding capacity and the linear network coding capacity of this network for any finite-field alphabet (Corollary IV.5). The method used to obtain the network exploits techniques from the theory of matroids, which we will discuss in a future publication.

Li, Yeung, and Cai [12] showed that any solvable multicast network has a scalar linear solution over a sufficiently large finite-field alphabet. Riis [15] noted in particular that every solvable multicast network has a linear solution over GF (2) in some vector dimension. For multicast networks, there have been various studies of algorithms for constructing scalar linear codes as well as the alphabet sizes needed for obtaining scalar linear solutions [3]–[6], [9]–[12].

For nonmulticast networks, various results have been given. Riis [15] constructed a network which is solvable over a binary alphabet, but which has no scalar linear solution over the finite field GF (2), and yet does have a linear solution over GF (2) in three dimensions. He also demonstrated in [15] solvable networks which can achieve linear solutions over GF (2) only if the vector dimension grows at least linearly with the number of nodes in the network.

Rasala Lehman and Lehman [11] gave a collection of networks which are solvable, but which have no scalar linear solution over any finite-field alphabet. Médard, Effros, Ho, and Karger [13] pointed out that the networks in [11] have linear solutions (based purely on routing) over every finite field in two dimensions. Similarly, it was noted in [13] that a certain network given by Koetter has no scalar linear solution but does have a linear (routing) solution in two dimensions.

It is clear that linear codes in dimensions two and higher are more powerful than scalar linear codes. Riis stated in [14]: "Maybe the most important question is whether any flow problem can be solved using linear coding." In fact, Médard, Effros, Ho, and Karger stated in [13]: "We conjecture that linear coding under its most general definition is sufficient for network coding in systems with arbitrary demands." The "most general definition" of linear coding is not specified in [13], but some clarification is given by Jaggi, Effros, Ho, and Médard [8] who state that the "most general possible linear codes" are filter-bank network codes, a generalization of convolutional codes. It is also stated in [8] that in [13] "it is conjectured that (linear codes) are asymptotically optimal."

We prove that vector linear coding is insufficient over the general class of R-modules, which includes as special cases finite fields, commutative rings with identity, and Abelian groups. Thus, the result is not restricted to alphabet cardinalities which are powers of primes, nor to linearity with respect to only a finite field. In addition, we show that linear coding (over finite fields) is not sufficient even asymptotically using fractional coding, as the ratio of message dimensions to edge dimensions approaches one. (In fact, we show that, in our example network, nonlinear network coding gives exactly 10% more capacity than the maximum capacity achievable using linear coding over finite fields.) From this, we deduce that even convolutional or filter-bank linear coding is not sufficient for network coding.[3]

Another form of "linearity" that one might consider (as suggested by R. Yeung) consists of time sharing between linear codes on different finite field alphabets. We note at the end of Section IV that this form of linearity is not sufficient for our example network either.

In what follows, the insufficiency of linear network codes is shown for finite fields in Section II, for rings and modules in Section III, and asymptotically for finite fields in Section IV.

We will often need to handle separately the cases of finite fields with even cardinality (i.e., characteristic two) and odd cardinality (i.e., odd characteristic).

II. Insufficiency of Network Linear Codes Over Finite Fields

In this section, we establish the existence of a solvable network that has no linear solution over any finite field and any vector dimension.

First we give a useful lemma (an alternative proof follows from the max-flow bound, e.g., see [18, p. 328]).

Lemma II.1: Suppose a network has a message m which is demanded by a node y and is produced by exactly one source node x. If there is a unique directed path from x to y, then the coding capacity of the network is at most 1.

[2] An R-module is like a vector space, but the coefficients of its vectors come from a ring R rather than from a field.

[3] Our proofs that linear codes are neither optimal for R-modules nor asymptotically optimal did not appear in the present paper until we submitted our revision in January 2005. (The quest for a proof based on R-modules was motivated by a comment from one referee and for asymptotics by a question from another referee.)

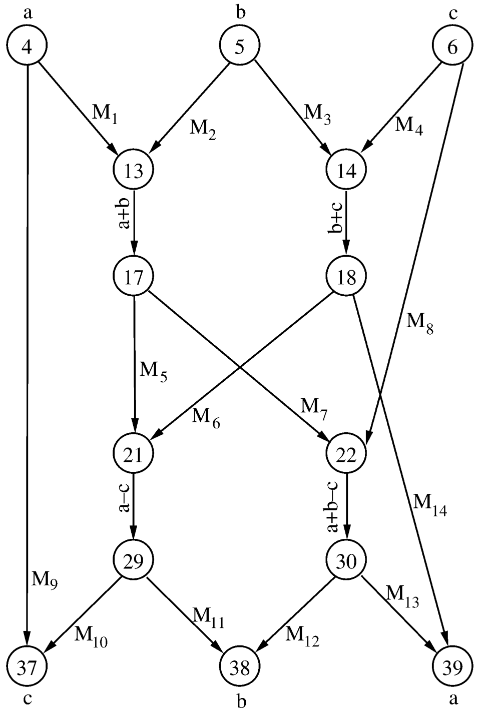

Fig. 1. The network $\mathcal{N}_1$ has sources n_4, n_5, n_6 emitting messages a, b, c, respectively, and sinks n_{37}, n_{38}, n_{39} with demands c, b, a, respectively. Some edges are labeled to their left to illustrate a scalar linear solution over any ring alphabet with characteristic two (used in Lemma II.2), and edges are labeled to their right with matrix coefficients M_i of an assumed solution used in Lemma II.2.

Proof: Suppose there exists a (k, n) fractional coding solution over alphabet $\mathcal{A}$ with $n < k$. If all messages other than m are fixed, then each edge of the path from x to y can take on at most $|\mathcal{A}|^n$ different values. So y can only decode at most $|\mathcal{A}|^n$ different values. But, $|\mathcal{A}|^n < |\mathcal{A}|^k$ so not every possible message m at x can be decoded at y, a contradiction. □

Suppose we impose on a network code the constraint that for every node with in-degree one, the out-edges must carry the same symbol as the lone in-edge, and for every source with exactly one message, the out-edges must carry the source's lone message. Then, it is easy to see that the network has a linear solution under this constraint for a given vector dimension over a given finite field if and only if the network has an unconstrained linear solution for the same vector dimension and over the same finite field. This fact is used implicitly in the proofs of Lemmas II.2 and II.3 and Theorem II.4 by assuming the described code constraint.

Denote by $\mathcal{N}_1$ the network shown in Fig. 1.

Lemma II.2: The network $\mathcal{N}_1$ has a scalar linear solution over any ring with characteristic two, but has no linear solution for any vector dimension over a finite field with odd characteristic. Also, the coding capacity of $\mathcal{N}_1$ is 1.

Proof: A scalar linear solution (as illustrated in Fig. 1) to the network $\mathcal{N}_1$ over any ring of characteristic two is given by the following edge functions and sink decoding functions:

$$\begin{aligned}
e_{13,17} &= e_{4,13} + e_{5,13} = a + b\\
e_{14,18} &= e_{5,14} + e_{6,14} = b + c\\
e_{21,29} &= e_{17,21} - e_{18,21} = (a + b) - (b + c) = a - c\\
e_{22,30} &= e_{17,22} - e_{6,22} = (a + b) - c\\
n_{37} : c &= e_{4,37} - e_{29,37} = a - (a - c)\\
n_{38} : b &= e_{30,38} - e_{29,38} = (a + b - c) - (a - c)\\
n_{39} : a &= e_{30,39} + e_{18,39} = (a + b - c) + (b + c) = a + 2b
\end{aligned}$$

(any edge function not shown is assumed to be an identity mapping). Note that the fact that the alphabet is a ring with characteristic two is used only in decoding the message a at node n_{39} where $2b = 0$.

Now, suppose the network has a linear solution over a finite field F with odd characteristic and some vector dimension k. Let I be the $k \times k$ identity matrix and for each i and j, let $e_{i,j}$ be the vector carried on the edge from n_i to n_j. Then there exist $k \times k$ matrices $M_1, \ldots, M_{14}$ with entries in F (as illustrated in Fig. 1), such that

$$\begin{aligned}
e_{13,17} &= M_1 a + M_2 b && (1)\\
e_{14,18} &= M_3 b + M_4 c && (2)\\
e_{21,29} &= M_5 e_{13,17} + M_6 e_{14,18} && (3)\\
e_{22,30} &= M_7 e_{13,17} + M_8 c && (4)\\
c &= M_9 a + M_{10} e_{21,29} && (5)\\
b &= M_{11} e_{21,29} + M_{12} e_{22,30} &&\\
a &= M_{13} e_{22,30} + M_{14} e_{14,18}. && (6)
\end{aligned}$$

Equating coefficients of a, b, c in (5) and (6) gives

$$\begin{aligned}
M_9 + M_{10} M_5 M_1 &= 0 && (7)\\
M_{10}(M_5 M_2 + M_6 M_3) &= 0 && (8)\\
M_{10} M_6 M_4 &= I && (9)\\
M_{11} M_5 M_1 + M_{12} M_7 M_1 &= 0 && (10)\\
M_{11} M_5 M_2 + M_{11} M_6 M_3 + M_{12} M_7 M_2 &= I && (11)\\
M_{11} M_6 M_4 + M_{12} M_8 &= 0 && (12)\\
M_{13} M_7 M_1 &= I && (13)\\
M_{13} M_7 M_2 + M_{14} M_3 &= 0 && (14)\\
M_{13} M_8 + M_{14} M_4 &= 0. && (15)
\end{aligned}$$

By (9) and (13), the matrices M_1, M_4, M_6, M_7, M_{10}, M_{13} are invertible. Therefore, $M_5 M_2 + M_6 M_3 = 0$ by (8), and hence,

$$M_{12} M_7 M_2 = I \qquad (16)$$

by (11). This implies that the matrices M_2 and M_{12} are invertible. We have $M_{11} M_5 = -M_{12} M_7$ by (10), so matrices M_5 and M_{11} are invertible. Since $M_{14} M_3 = -M_{13} M_7 M_2$ by (14), the matrices M_3 and M_{14} are invertible. Since $M_9 = -M_{10} M_5 M_1$ by (7), the matrix M_9 is invertible. Thus, since $M_8 = -M_{13}^{-1} M_{14} M_4$ by (15), the matrix M_8 is invertible. So, M_i is invertible for all i.

Now, we have

$$\begin{aligned}
0 &= M_{11} M_5 + M_{12} M_7 && \text{[from (10)]}\\
&= M_{11} M_5 + (M_{12} M_7 M_2) M_2^{-1} &&\\
&= M_{11} M_5 + M_2^{-1} && \text{[from (16)]}\\
-I &= M_{11} M_5 M_2 && (17)
\end{aligned}$$

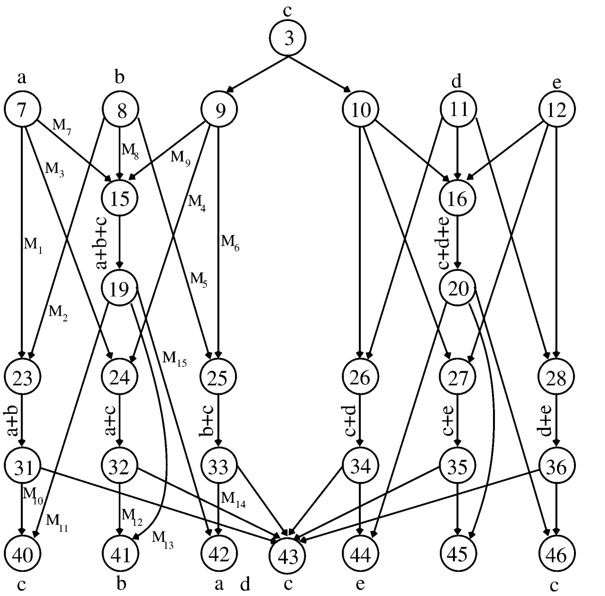

Fig. 2. The network $\mathcal{N}_2$ has sources $n_7, n_8, n_3, n_{11}, n_{12}$ with messages a, b, c, d, e, respectively. Sinks n_{40} through n_{46} each demand one of the messages, as indicated. Some edges are labeled to illustrate a scalar linear solution over any ring alphabet where 2 is invertible (used in Lemma II.3), and edges are labeled to their right with matrix coefficients M_i of an assumed solution used in Lemma II.3.

and

$$\begin{aligned}0 &= M_{11}M_6M_4 + M_{12}M_8 && \text{[from (12)]}\\ &= M_{11}(M_6M_3)M_3^{-1}M_4\\ &\quad - M_{12}M_{13}^{-1}M_{14}M_4 && \text{[from (15)]}\\ &= -M_{11}(M_5M_2)M_3^{-1}M_4\\ &\quad + (M_{12}M_7M_2)M_3^{-1}M_4 && \text{[from (8), (14)]}\\ &= (-M_{11}M_5M_2 + I)M_3^{-1}M_4 && \text{[from (16)]}\\ I &= M_{11}M_5M_2. && (18)\end{aligned}$$

But (17) and (18) imply that $I = -I$ which is impossible in a field with odd characteristic.

Finally, since $\mathcal{N}_1$ has a scalar linear solution over GF(2), its coding capacity (independent of alphabet size by Lemma I.1) is at least 1. Since n_6 is the only node that produces message c and node n_{37} demands message c, and since there is a unique directed path from n_6 to n_{37}, the coding capacity of $\mathcal{N}_1$ is at most 1 by Lemma II.1. Hence, the coding capacity of $\mathcal{N}_1$ is exactly 1. □

Denote by $\mathcal{N}_2$ the network shown in Fig. 2.

Lemma II.3: The network $\mathcal{N}_2$ has a scalar linear solution over any ring where 2 is a unit, but has no linear solution for any vector dimension over a finite field with characteristic two. Also, the coding capacity of $\mathcal{N}_2$ is 1.

Proof: A scalar linear solution (as illustrated in Fig. 2) to the network $\mathcal{N}_2$ over any ring where 2 is invertible is given by the following edge functions and sink decoding functions (any edge function not shown is assumed to be an identity mapping):

$$\begin{aligned}e_{15,19} &= e_{7,15} + e_{8,15} + e_{9,15} = a + b + c\\ e_{23,31} &= e_{7,23} + e_{8,23} = a + b\\ e_{24,32} &= e_{7,24} + e_{9,24} = a + c\\ e_{25,33} &= e_{8,25} + e_{9,25} = b + c\\ e_{16,20} &= e_{10,16} + e_{11,16} + e_{12,16} = c + d + e\\ e_{26,34} &= e_{10,26} + e_{11,26} = c + d\\ e_{27,35} &= e_{10,27} + e_{12,27} = c + e\\ e_{28,36} &= e_{11,28} + e_{12,28} = d + e\\ n_{40} : c &= e_{19,40} - e_{31,40} = (a + b + c) - (a + b)\\ n_{41} : b &= e_{19,41} - e_{32,41} = (a + b + c) - (a + c)\\ n_{42} : a &= e_{19,42} - e_{33,42} = (a + b + c) - (b + c)\\ n_{43} : c &= 2^{-1} \cdot (e_{32,43} + e_{33,43} - e_{31,43})\\ &= 2^{-1} \cdot ((a + c) + (b + c) - (a + b))\\ n_{44} : e &= e_{20,44} - e_{34,44} = (c + d + e) - (c + d)\\ n_{45} : d &= e_{20,45} - e_{35,45} = (c + d + e) - (c + e)\\ n_{46} : c &= e_{20,46} - e_{36,46} = (c + d + e) - (d + e).\end{aligned}$$

Now, suppose the $\mathcal{N}_2$ network has a linear solution over some finite field F of characteristic two and with some finite vector dimension k. Henceforth, let $+$ denote addition in F. We can write

$$\begin{aligned}e_{23,31} &= M_1a + M_2b\\ e_{24,32} &= M_3a + M_4c && (19)\\ e_{25,33} &= M_5b + M_6c && (20)\\ e_{15,19} &= M_7a + M_8b + M_9c\\ n_{40} : c &= M_{10}(M_1a + M_2b)\\ &\quad + M_{11}(M_7a + M_8b + M_9c) && (21)\\ n_{41} : b &= M_{12}(M_3a + M_4c)\\ &\quad + M_{13}(M_7a + M_8b + M_9c) && (22)\\ n_{42} : a &= M_{14}(M_5b + M_6c)\\ &\quad + M_{15}(M_7a + M_8b + M_9c) && (23)\end{aligned}$$

where each M_i is a $k \times k$ matrix with elements in F, and the messages a, b, c, d, e are k-dimensional vectors. Equations (21)–(23) come from the demands at sinks n_{40}, n_{41}, n_{42}. Let I be the $k \times k$ identity matrix over F. Equating coefficients of a, b, c in (21)–(23) gives

$$\begin{aligned}I &= M_{11}M_9 = M_{13}M_8 = M_{15}M_7 && (24)\\ M_{10}M_1 &= M_{11}M_7 && (25)\\ M_{10}M_2 &= M_{11}M_8 && (26)\\ M_{12}M_3 &= M_{13}M_7 && (27)\\ M_{12}M_4 &= M_{13}M_9 && (28)\\ M_{14}M_5 &= M_{15}M_8 && (29)\\ M_{14}M_6 &= M_{15}M_9 && (30)\end{aligned}$$

where minus signs have been omitted since the finite-field alphabet has characteristic two. Equation (24) implies that

$$M_7, M_8, M_9, M_{11}, M_{13}, M_{15}$$

are invertible. Since the right-hand sides of (25)–(30) are invertible, the left-hand-side matrices

$$M_1, M_2, M_3, M_4, M_5, M_6, M_{10}, M_{12}, M_{14}$$

must also be invertible. So M_i is invertible for all i. Thus,

$$M_2 = M_1M_7^{-1}M_8 \qquad \text{[from (25), (26)]}$$

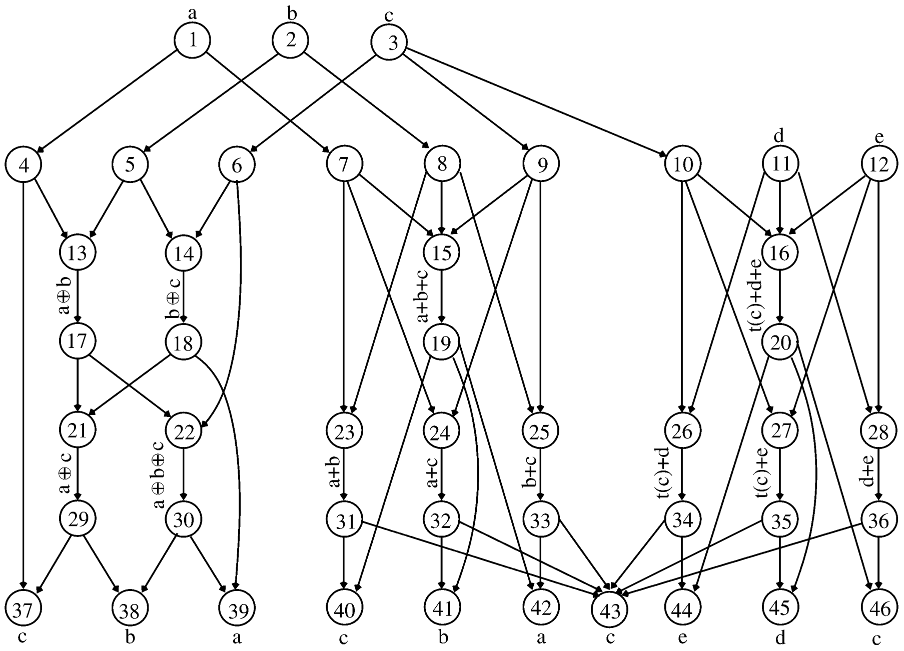

Fig. 3. The network $\mathcal{N}_3$ has sources $n_1, n_2, n_3, n_{11}, n_{12}$ with messages a, b, c, d, e, respectively. Sinks n_{37} through n_{46} each demand one of the messages, as indicated. Some edges are labeled to illustrate a nonlinear solution over an alphabet of size 4 (used in Theorem II.4).

$$M_4 = M_3 M_7^{-1} M_9 \qquad \text{[from (27), (28)]}$$
$$M_6 = M_5 M_8^{-1} M_9 \qquad \text{[from (29),(30)]}$$

and therefore,

$$\begin{aligned} M_1 a + M_2 b &= M_1 \left(a + M_7^{-1} M_8 b\right) \\ M_3 a + M_4 c &= M_3 \left(a + M_7^{-1} M_9 c\right) \\ M_5 b + M_6 c &= M_5 \left(b + M_8^{-1} M_9 c\right). \end{aligned}$$

Finally

$$\begin{aligned} &M_1^{-1} e_{23,31} + M_3^{-1} e_{24,32} + M_1^{-1} M_2 M_5^{-1} e_{25,33} \\ &\quad = M_1^{-1}(M_1 a + M_2 b) + M_3^{-1}(M_3 a + M_4 c) \\ &\qquad + M_1^{-1} M_2 M_5^{-1}(M_5 b + M_6 c) \\ &\quad = (a + M_7^{-1} M_8 b) + (a + M_7^{-1} M_9 c) \\ &\qquad + M_1^{-1}(M_1 M_7^{-1} M_8)(b + M_8^{-1} M_9 c) \\ &\quad = 0 \end{aligned}$$

so

$$e_{23,31} = M_1 M_3^{-1} e_{24,32} + M_2 M_5^{-1} e_{25,33}. \tag{31}$$

Hence, for any message assigned to c, if the messages $M_3^{-1} M_4 c$ and $M_5^{-1} M_6 c$ are assigned to a and b, respectively, then $e_{24,32} = e_{25,33} = 0$, by (19) and (20), and therefore, $e_{23,31} = 0$ by (31). A similar argument shows that, for any message assigned to c, there exist messages that can be assigned to d and e that result in $e_{26,34} = e_{27,35} = e_{28,36} = 0$.

Thus, for every message vector assigned to c, there exist assignments of messages to a, b, d, e such that all six inputs

$$e_{24,32}, e_{25,33}, e_{23,31}, e_{26,34}, e_{27,35}, e_{28,36}$$

to node n_{43} are zero. This contradicts the assumption that the demand c at node n_{43} can be recovered, since c is not uniquely determined by the node's inputs.

Finally, since $\mathcal{N}_2$ has a scalar linear solution over GF (3), its coding capacity (independent of alphabet size by Lemma I.1) is at least 1. Since n_3 is the only node that produces message c and node n_{40} demands message c, and since there is a unique directed path from n_3 to n_{40}, the coding capacity of $\mathcal{N}_2$ is at most 1 by Lemma II.1. Hence, the coding capacity of $\mathcal{N}_2$ is exactly 1. □

Denote by $\mathcal{N}_3$ the network shown in Fig. 3, with nodes $n_1, \dots, n_{46}$. In the $\mathcal{N}_3$ network, the left-most part is the $\mathcal{N}_1$ network (with sinks n_{37}, n_{38}, n_{39}) and the rest of $\mathcal{N}_3$ is the $\mathcal{N}_2$ network.

Theorem II.4 shows that linear network codes are insufficient over finite-field alphabets.

Theorem II.4: There exists a solvable network that has no linear solution over any finite field and any vector dimension.

Proof: The proof is achieved with $\mathcal{N}_3$, which combines networks $\mathcal{N}_1$ and $\mathcal{N}_2$. Lemmas II.2 and II.3 show that network $\mathcal{N}_3$ does not have a vector linear solution over any finite-field alphabet.

We now demonstrate a solution to the network over an alphabet of cardinality 4, as indicated in Fig. 3. The symbols $+$ and $-$ indicate addition and subtraction in the ring $\boldsymbol{Z}_4$ of integers modulo 4, the symbol $\oplus$ indicates addition in the ring $\boldsymbol{Z}_2 \oplus \boldsymbol{Z}_2$ (i.e., bitwise XOR), and $t(x)$ indicates the result of exchanging the order of the bits in a 2-bit binary word x. We represent the elements of the alphabet either as members of $\boldsymbol{Z}_4$ when using $+$ or $-$, or as elements of $\boldsymbol{Z}_2 \oplus \boldsymbol{Z}_2$ (i.e., 2-bit binary words) when using $\oplus$ or $t(\cdot)$. Note that the functions $+$ and $-$ are linear over $\boldsymbol{Z}_4$ but not over GF (4) or $\boldsymbol{Z}_2 \oplus \boldsymbol{Z}_2$, the function $\oplus$ is linear over $\boldsymbol{Z}_2 \oplus \boldsymbol{Z}_2$ and GF (4) but not over $\boldsymbol{Z}_4$, and the function $t(\cdot)$ is not linear over any of these. The demands are met as follows:

$$\begin{aligned} n_{37} &: c = e_{4,37} \oplus e_{29,37} = a \oplus (a \oplus c) \\ n_{38} &: b = e_{29,38} \oplus e_{30,38} = (a \oplus c) \oplus (a \oplus b \oplus c) \\ n_{39} &: a = e_{30,39} \oplus e_{18,39} = (a \oplus b \oplus c) \oplus (b \oplus c) \end{aligned}$$

$$\begin{aligned}
n_{40} : c &= e_{19,40} - e_{31,40} = (a+b+c) - (a+b)\\
n_{41} : b &= e_{19,41} - e_{32,41} = (a+b+c) - (a+c)\\
n_{42} : a &= e_{19,42} - e_{33,42} = (a+b+c) - (b+c)\\
n_{43} : c &= t(e_{32,43} + e_{33,43} - e_{31,43})\\
&\quad + (e_{34,43} + e_{35,43} - e_{36,43})\\
&= t((a+c) + (b+c) - (a+b))\\
&\quad + (t(c)+d) + (t(c)+e) - (d+e)\\
&= t(2c) + 2t(c) \quad \text{(see below)}\\
n_{44} : e &= e_{20,44} - e_{34,44} = (t(c)+d+e) - (t(c)+d)\\
n_{45} : d &= e_{20,45} - e_{35,45} = (t(c)+d+e) - (t(c)+e)\\
n_{46} : c &= t(e_{20,46} - e_{36,46}) = t((t(c)+d+e) - (d+e)).
\end{aligned}$$

In decoding c at node n_{43}, we used the fact that if the 2-bit binary representation of c is (x, y), then the following binary representations also hold:

$$\begin{aligned}
2c &= (y, 0)\\
t(c) &= (y, x)\\
t(2c) &= (0, y)\\
2t(c) &= (x, 0)\\
2t(c) + t(2c) &= (x, y) = c.
\end{aligned}$$

□

Corollary II.5: The coding capacity of the network $\mathcal{N}_3$ is 1.

Proof: By Theorem II.4, the network $\mathcal{N}_3$ is solvable and therefore the coding capacity is at least 1 (independent of the alphabet size by Lemma I.1). Since n_3 is the only node that produces message c and node n_{37} demands message c, and since there is a unique directed path from n_3 to n_{37}, the coding capacity of $\mathcal{N}_3$ is at most 1 by Lemma II.1. Hence, the coding capacity of $\mathcal{N}_3$ is exactly 1. □

III. Insufficiency of Network Linear Codes Over Rings and Modules

In this section, we start by showing how to extend nonsolvability over finite fields to nonsolvability over finite commutative rings with identity. Whereas finite fields are uniquely characterized up to isomorphism by their cardinality, the same is not true of rings. Such rings exist for every cardinality and often in many different forms. For a given finite alphabet, the linearity of a network code can be considered with respect to any commutative ring with identity whose cardinality is the same as that of the alphabet. The lack of inverses in a ring does not prevent the use of linear network codes. In fact, consideration of rings increases the variety of codes to choose from and also allows linear codes over arbitrary alphabet sizes, instead of only powers of primes.

Theorem III.1: If a network does not have a linear solution over any finite field in k dimensions, then it does not have a linear solution over any finite commutative ring with identity in k dimensions.

Proof: Let R be a finite commutative ring with identity and J a maximal ideal in R. Consider an arbitrary network with a linear solution over the ring R in dimension k. We will show that the linear ring solution induces a linear solution over the quotient field R/J in dimension k.

To do this, we need to redefine the edge functions and demand functions for the new alphabet R/J. Let

$$h : R \to R/J$$

be the ring homomorphism defined by

$$h(x) = x + J.$$

Since the solution is presumed to be linear, all edge functions and demand functions are linear combinations of their input vectors, where the linear combinations' constant coefficients are matrices over R. We replace each entry of each such matrix and of each node's input vectors by its image under h. This clearly gives us linear edge functions and linear demand functions under the new alphabet R/J. We need only show that the demands are met.

Since the map h is surjective, each message vector x in the new alphabet R/J has a corresponding message vector $\bar{x}$ in the original alphabet R, satisfying $h(\bar{x}) = x$. It follows by induction that along every edge, if y is the vector carried by the new coding and $\bar{y}$ is the vector carried by the old coding, then $h(\bar{y}) = y$ (component-wise). This is because addition and multiplication are preserved by ring homomorphisms.

In particular, each sink will recover its demands in R/J, and the new code is linear over the field R/J. Thus, the new network has a linear solution.

Thus, since a linear solution over any finite commutative ring with identity induces a linear solution over some finite field, if no linear solution exists over any finite field, then there cannot possibly be any linear solution over any commutative ring with identity. □

The next corollary follows immediately from Theorem II.4 and Theorem III.1. The corollary establishes that linear network codes are insufficient over a class of rings that includes finite fields.

Corollary III.2: There exists a solvable network such that for every vector dimension there is no linear solution over any finite commutative ring with identity.

In [11], solvable networks were given, whose minimum alphabet size required for a solution could be made arbitrarily large. By combining such a network with the network $\mathcal{N}_2$ used in the proof of Theorem II.4 (i.e., taking the disjoint union of them) one obtains a solvable network with no linear solution for any vector dimension and an arbitrarily large minimum alphabet size for a solution. From this fact and Corollary III.2, we immediately obtain the following corollary.

Corollary III.3: For each $n > 0$, there exists a solvable network which has no scalar solution for any alphabet of cardinality smaller than n, and such that for every vector dimension there is no linear solution over any finite commutative ring with identity.

We can talk about linearity in even more generality than the above, if we are willing to separate the set of coefficients allowed in linear functions from the set of inputs to the linear functions (the set of messages). For example, it makes sense to talk about linear functions over any Abelian group G if we restrict the coefficients of those functions to be integers, because ng makes sense for any integer n and any element g of G. Or we can let the set of coefficients be any field F and let the message set be any vector space over F.

If we generalize the definition of vector space to use a ring R instead of a field F, we get what are called R-modules. For any ring R, an *R-module* (or, more specifically, a *left R-module*) is an Abelian group G together with an action of R on G (i.e., a mapping from $R \times G$ to G), denoted here by concatenation: rg is the result of ring element r acting on group element g. This action must satisfy the analogues of the usual vector space laws: for any $r, s \in R$ and $g, h \in G$, we have

$$\begin{aligned}(rs)g &= r(sg)\\ (r+s)g &= rg + sg\\ r(g+h) &= rg + rh\\ 0g &= 0\end{aligned}$$

(the first 0 here is the ring zero and the second 0 is the group zero). If R is a ring with identity I, then we also require that $Ig = g$ for all $g \in G$.

This generalizes the two previous examples; any Abelian group is a $\boldsymbol{Z}$-module under the obvious action of the integers $\boldsymbol{Z}$ on the group by repeated addition, and any vector space over a field F is an F-module.

The notions of scalar linear solution and (vector) linear solution for a network now easily generalize to the context of an R-module G. For a scalar R-linear solution over G, the set of messages to select from is G, and each edge function or decoding function must be an R-linear function (i.e., one of the form

$$f(x_1, \dots, x_j) = r_1x_1 + \cdots + r_jx_j$$

where $r_1, \dots, r_j$ are fixed elements of R). For an R-linear solution of vector dimension k, the set of messages is G^k and the edge and decoding functions are such that each component of the output vector is a fixed R-linear combination of the components of the input vectors.

Any ring R is itself an R-module acting on itself by left multiplication, so module linearity includes ring linearity as a special case.

Note that if R is a ring and G is an R-module, then $M_k(R)$, the set of $k \times k$ matrices over R with matrix addition and multiplication defined in the usual way, is a ring (with identity if R has an identity) and G^k is an $M_k(R)$-module, and any k-dimensional R-linear solution over G becomes a scalar $M_k(R)$-linear solution over G^k. So, in this very general context, (vector) linear solvability gives no more generality than scalar linear solvability (on a larger module).

Theorem III.4: There exists a solvable network that does not have an R-linear solution over G for any ring R, any finite R-module G with more than one element, and any vector dimension.

Proof: The network $\mathcal{N}_3$ will again prove the assertion. First, by the remark preceding the theorem, it will be enough to show that, for any ring R and any finite R-module G with more than one element, there is no scalar R-linear solution over G.

Next, we may restrict ourselves to the case where R acts faithfully on G; that is, if $r_1, r_2 \in R$ are such that $r_1g = r_2g$ for all $g \in G$, then $r_1 = r_2$. For suppose we have R and G as in the hypotheses of the theorem so that there is a scalar R-linear solution over G for the network $\mathcal{N}_3$. Let J be the set of all $z \in R$ such that $zg = 0$ for all $g \in G$. It is easy to see that J contains 0 and is closed under addition; it is also true that, if $z \in J$ and $r \in R$, then zr and rz are in J, because $(zr)g = z(rg) = 0$ and $(rz)g = r(zg) = r0 = 0$ for all $g \in G$. So J is a two-sided ideal in R, and we can form the quotient ring $R' = R/J$. We define an action R' on G by the formula

$$(r + J)g = rg$$

for all $r \in R$ and $g \in G$; this is well defined, since by the definition of J, it does not matter which member of the coset $r+J$ is chosen. It is easy to see that R' acts faithfully on G. And, just as in the proof of Theorem III.1, a scalar R-linear solution to the network yields a scalar R'-linear solution to the network; we simply have to replace each coefficient r in the edge and decoding functions with $r + J$. (The message group G does not change.) Hence, if there is a scalar linear solution, then there is a scalar linear solution in a module where the action is faithful. So assume faithfulness from now on.

If we have a scalar R-linear solution over G, then for each demand at a sink node we get an equation of the form

$$a_i = f(a_1, a_2, \dots, a_m)$$

where $a_1, a_2, \dots, a_m$ are the source messages and f is the composition of decoding and edge functions given by the specified solution; this equation must hold for all choices of $a_1, a_2, \dots, a_m$ from G. Now f will be R-linear, so we can write

$$f(x_1, \dots, x_m) = r_1x_1 + r_2x_2 + \cdots + r_mx_m$$

with coefficients $r_1, r_2, \dots, r_m$ from R. Since the decoding must in particular be correct when all messages other than a_i are zero, we have $r_ia_i = a_i$ for all $a_i \in G$. So R has an element I such that $Ig = g$ for all $g \in G$; because of faithfulness, this element is unique. For any $r \in R$ and $g \in G$, we have $(Ir)g = I(rg) = rg$ and $(rI)g = r(Ig) = rg$; hence, $Ir = rI = r$ by faithfulness. Also, if g is a nonzero element of G, then $Ig = g \neq 0 = 0g$, so $I \neq 0$. Therefore, R is in fact a ring with identity.

Faithfulness states that different elements r of R yield different functions $g \mapsto rg$ from G to G. Since G is finite, the number of such functions is finite, so R must be finite.

We now make use of the following fact: in a finite ring R with identity I, if $ab = I$, then $ba = I$. This result and much more general versions can be found in the literature (see, e.g., Jacobson [7]), but the simple version here can be proved quickly as follows. If $ab = I$, then the map f from R to R defined by $f(x) = xa$ is one-to-one, because $f(x) = f(x')$ means $xa = x'a$, which implies $xab = x'ab$ and hence $x = x'$. Hence, since R is finite, f must be onto, so there exists x_0 in R such that $f(x_0) = I$, or $x_0a = I$. Now we have $x_0 = x_0ab = b$, so $ba = I$. (The authors thank Daniel Goldstein for this clean proof of the result and Lance Small for the reference.)

So, in R, any left inverse is a two-sided inverse and so is any right inverse. If it exists, the two-sided inverse of a is unique (because $ba = I$ and $ac = I$ imply $b = bac = c$)—call it a^{-1}, and say that a is invertible. If $abc = I$, then a^{-1} and c^{-1} exist by the above, but b^{-1} also exists, because we have $bca = cab = I$.

We can now follow the proofs in the preceding section almost verbatim, changing "matrix" to "member of R" and so on. (The faithfulness of the R-action allows us to conclude from "$rg = g$ for all $g \in G$" that $r = I$, or from "$rg = 0$ for all $g \in G$" that $r = 0$.) The division into cases will be on whether $I + I = 0$ in the final ring R (or, equivalently, whether $g + g = 0$ for all g in G). If $I + I \neq 0$, then the proof of Lemma II.2 shows that $\mathcal{N}_1$ does not have a scalar R-linear solution over G. If $I + I = 0$, then the proof of Lemma II.3 shows that $\mathcal{N}_2$ does not have a scalar R-linear solution over G. Therefore, in all cases, $\mathcal{N}_3$ does not have a scalar R-linear solution over G, so we are done. □

Since any ring R is itself an R-module, we get the following.

Corollary III.5: Corollaries III.2 and III.3 remain true if "finite commutative ring with identity" is replaced with "finite ring with more than one element."

IV. Asymptotic Insufficiency of Network Linear Codes Over Finite Fields

Throughout this section, F is a finite field, all matrices have entries in F, I_j denotes the $j \times j$ identity matrix for each j, and $e_{i,j}$ denotes the vector carried on the edge from a node n_i to a node n_j, where n_i and n_j are two adjacent nodes in some given network. Without loss of generality, we will assume that the first k components of each out-edge of a source consist of the k components of the corresponding source message. Also, we can assume that the out-edges of any node with in-degree 1 are copies of the in-edge to the node (see the discussion before Lemma II.2). The following notation will be used in proofs in this section.

Notation: Let the network messages be denoted by $m_1, \ldots, m_r \in F^k$. For $1 \leq i \leq i_0$ and $1 \leq j \leq j_0$, let $f_i : F^{kr} \to F^{s_i}$ and $g_j : F^{kr} \to F^{t_j}$ be linear functions where $s_i, t_j > 0$. We use the notation

$$\begin{aligned} & f_1(m_1, \ldots, m_r), \ldots, f_{i_0}(m_1, \ldots, m_r) \\ & \qquad \longrightarrow g_1(m_1, \ldots, m_r), \ldots, g_{j_0}(m_1, \ldots, m_r) \end{aligned}$$

to indicate that there exist $t_j \times s_i$ matrices $A_{i,j}$ such that for all $m_1, \ldots, m_r \in F^k$

$$g_j(m_1, \ldots, m_r) = \sum_i A_{i,j} f_i(m_1, \ldots, m_r).$$

Intuitively, the definition says that if a network node can linearly compute $f_1, \ldots, f_{i_o}$ then it is also able to linearly compute $g_1, \ldots, g_{j_o}$. (We will often arrange the terms on the left of $\longrightarrow$ vertically for ease of readability.) Note that $\longrightarrow$ is transitive.

Lemma IV.1 (e.g., see [16, p. 124]): If $A : F^m \to F^n$ and $B : F^k \to F^m$ are linear maps, then

$$\begin{aligned} \mathsf{rank}(A) + \mathsf{rank}(B) - m &\leq \mathsf{rank}(AB) \\ &\leq \min(\mathsf{rank}(A), \mathsf{rank}(B)). \end{aligned}$$

The next lemma follows immediately from Gaussian elimination.

Lemma IV.2: If A is an $n \times k$ matrix of rank k, then there exists an $n \times n$ invertible matrix B such that

$$BA = \begin{bmatrix} I_k \\ 0 \end{bmatrix}.$$

Theorem IV.3: The linear coding capacity of network $\mathcal{N}_1$ is $4/5$ over any odd-characteristic finite field and is 1 over any even-characteristic finite field.

Proof: If the alphabet is a finite field F of characteristic two, then a scalar linear solution (i.e., with $k = n = 1$) is guaranteed by Lemma II.2. Thus, the linear coding capacity of $\mathcal{N}_1$ is at least 1 for even $|F|$. Also, the linear coding capacity is upper-bounded by the coding capacity, which equals 1 by Lemma II.2. Thus, the linear coding capacity of $\mathcal{N}_1$ is 1 if $|F|$ is even.

Henceforth in the proof, assume the alphabet is a finite field F with odd characteristic. Suppose there exists a (k, n) fractional linear solution for $\mathcal{N}_1$ over F. First we show in general that k/n cannot exceed $4/5$ and then we demonstrate a specific linear code which achieves $k/n = 4/5$.

Since the coding capacity of $\mathcal{N}_1$ is 1 by Lemma II.2, we may assume $k \leq n$, for otherwise the linear coding capacity would exceed the coding capacity.

Let

$$\delta = n - k.$$

By supposition, there exist $n \times k$ matrices M_1, M_2, M_3, M_4, M_8; $n \times n$ matrices M_5, M_6, M_7; $k \times n$ matrices M_{10}, M_{11}, M_{12}, M_{13}, M_{14}; and a $k \times k$ matrix M_9 with entries in F, such that (1)–(15) are satisfied. Henceforth, the only properties we will use will be (1)–(15).

We will write each matrix M_i in terms of a $k \times k$ matrix R_i, a $k \times \delta$ matrix S_i, a $\delta \times k$ matrix T_i, and a $\delta \times \delta$ matrix U_i, as

$$M_i = \begin{bmatrix} R_i & S_i \\ T_i & U_i \end{bmatrix}.$$

If M_i is $n \times k$, then S_i and U_i are omitted; if M_i is $k \times n$, then T_i and U_i are omitted.

Claim: We may assume without loss of generality that

$$\begin{aligned} R_1 &= R_4 = R_6 = R_7 = I_k \\ T_1 &= T_4 = T_6 = T_7 = 0_{\delta \times k}. \end{aligned}$$

Proof of Claim: Since we have $M_{13} M_7 M_1 = I_k$ and $M_{10} M_6 M_4 = I_k$ from (13) and (9), Lemma IV.1 gives $\mathsf{rank}(M_1) \geq k$ and $\mathsf{rank}(M_4) \geq k$. Hence, by Lemma IV.2 we can find invertible $n \times n$ matrices B_1 and B_4 such that

$$B_1 M_1 = B_4 M_4 = \begin{bmatrix} I_k \\ 0_{\delta \times k} \end{bmatrix}.$$

Define the following matrices:

$$\begin{aligned} M_1' &= B_1 M_1 & M_5' &= M_5 B_1^{-1} \\ M_2' &= B_1 M_2 & M_6' &= M_6 B_4^{-1} \\ M_3' &= B_4 M_3 & M_7' &= M_7 B_1^{-1} \\ M_4' &= B_4 M_4 & M_{14}' &= M_{14} B_4^{-1} \end{aligned}$$

and suppose a new (k, n) fractional linear code is formed by replacing each matrix $M_1, M_2, M_3, M_4, M_5, M_6, M_7$, and M_{14}

in the assumed fractional linear solution by the corresponding matrix with a prime notation. It is easy to see from (1)–(6) that the new code is also a solution, and the new M_1 and M_4 satisfy

$$\begin{aligned} R_1 &= R_4 = I_k \\ T_1 &= T_4 = 0_{\delta\times k} \end{aligned}$$

(which we will henceforth assume).

Now we can apply Lemma IV.1 to (13) and (9) again to get

$$\begin{aligned} \mathsf{rank}(M_7M_1) &\geq k \\ \mathsf{rank}(M_6M_4) &\geq k. \end{aligned}$$

Hence, we can find invertible $n\times n$ matrices B_7 and B_6 such that

$$B_7M_7M_1 = B_6M_6M_4 = \begin{bmatrix} I_k \\ 0_{\delta\times k} \end{bmatrix}.$$

Define the following matrices

$$\begin{array}{ll} M_5' = B_6M_5 & M_{10}' = M_{10}B_6^{-1} \\ M_6' = B_6M_6 & M_{11}' = M_{11}B_6^{-1} \\ M_7' = B_7M_7 & M_{12}' = M_{12}B_7^{-1} \\ M_8' = B_7M_8 & M_{13}' = M_{13}B_7^{-1} \end{array}$$

and replace M_5, M_6, M_7, M_8, M_{10}, M_{11}, M_{12}, and M_{13} by the corresponding matrices with prime notation. Again, the new code is also a solution, and we now have

$$M_7M_1 = M_6M_4 = \begin{bmatrix} I_k \\ 0_{\delta\times k} \end{bmatrix}.$$

But since $R_1 = R_4 = I_k$ and $T_1 = T_4 = 0_{\delta\times k}$, we have

$$\begin{aligned} M_7M_1 &= \begin{bmatrix} R_7 \\ T_7 \end{bmatrix} \\ M_6M_4 &= \begin{bmatrix} R_6 \\ T_6 \end{bmatrix} \end{aligned}$$

so

$$\begin{aligned} R_6 &= R_7 = I_k \\ T_6 &= T_7 = 0_{\delta\times k}. \end{aligned}$$

Thus, in general, if there is a (k,n) fractional linear solution to $\mathcal{N}_1$, then there is also a (k,n) fractional linear solution with the claimed constraints on M_1, M_4, M_6, and M_7. □

We have

$$\begin{aligned} e_{13,17} &= \begin{bmatrix} a + R_2b \\ T_2b \end{bmatrix} && \text{[from (1)]} \quad (32) \\ e_{14,18} &= \begin{bmatrix} R_3b + c \\ T_3b \end{bmatrix} && \text{[from (2)]} \quad (33) \\ c &= M_9a + [R_{10} \quad S_{10}]\, e_{21,29} && \text{[from (5)]} \\ &= M_9a + [R_{10} \quad S_{10}] \begin{bmatrix} x + c \\ y \end{bmatrix} && \text{[from (3), (32), (33)]} \\ & && (34) \end{aligned}$$

where

$$\begin{aligned} x &= R_5a + R_5R_2b + S_5T_2b + R_3b + S_6T_3b \\ y &= T_5a + T_5R_2b + U_5T_2b + U_6T_3b. \end{aligned}$$

Claim: We may assume without loss of generality that $R_{10} = I_k$ and $S_{10} = 0$.

Proof of Claim: Define the following matrices:

$$\begin{aligned} M_5' &= \begin{bmatrix} I_k & S_{10} \\ 0 & I_\delta \end{bmatrix} M_5 \\ M_6' &= \begin{bmatrix} I_k & S_{10} \\ 0 & I_\delta \end{bmatrix} M_6 \\ M_{10}' &= M_{10} \begin{bmatrix} I_k & -S_{10} \\ 0 & I_\delta \end{bmatrix} \\ M_{11}' &= M_{11} \begin{bmatrix} I_k & -S_{10} \\ 0 & I_\delta \end{bmatrix} \end{aligned} \qquad (35)$$

and note that

$$\begin{aligned} M_{10}'M_5' &= M_{10}M_5 \\ M_{10}'M_6' &= M_{10}M_6. \end{aligned} \qquad (36)$$

Since c appears only once on the right-hand side of (34), it must be the case that $R_{10} = I_k$. So from (35) we have $R_{10}' = I_k$ and $S_{10}' = 0$.

Now suppose we replace M_5, M_6, M_{10}, M_{11} by M_5', M_6', M_{10}', M_{11}', respectively, and for each i and j let $e_{i,j}'$ be the resulting vector carried on edge $e_{i,j}$. Then we have

$$\begin{aligned} e_{13,17}' &= e_{13,17} \\ e_{14,18}' &= e_{14,18} \end{aligned} \qquad (37)$$

and thus,

$$\begin{aligned} M_{10}'e_{21,29}' &= M_{10}'(M_5'e_{13,17}' + M_6'e_{14,18}') && \text{[from (3)]} \\ &= M_{10}(M_5e_{13,17} + M_6e_{14,18}) && \text{[from (36)–(37)]} \\ &= M_{10}e_{21,29} && \text{[from (3)].} \end{aligned}$$

A similar argument shows

$$M_{11}'e_{21,29}' = M_{11}e_{21,29}.$$

These facts imply that we still have a linear solution to the network. Also note that the assumptions from the first claim remain true. So, if there exists a (k,n) fractional linear solution to the network, then there exists a (k,n) fractional linear solution satisfying $R_{10} = I_k$ and $S_{10} = 0$. □

From (34) we obtain

$$c = M_9a + x + c$$

which upon equating the terms containing b yields

$$R_5R_2 + S_5T_2 + R_3 + S_6T_3 = 0. \qquad (38)$$

We have

$$\begin{aligned} & a + R_2b + S_7T_2b + R_8c, \\ & U_7T_2b + T_8c, \\ & c + R_3b, \\ & T_3b \\ & \longrightarrow \begin{bmatrix} I_k & S_7 \\ 0 & U_7 \end{bmatrix} \begin{bmatrix} a + R_2b \\ T_2b \end{bmatrix} + \begin{bmatrix} R_8 \\ T_8 \end{bmatrix} c, \\ & \qquad \begin{bmatrix} R_3 \\ T_3 \end{bmatrix} b + \begin{bmatrix} I_k \\ 0 \end{bmatrix} c \\ & = M_7e_{13,17} + M_8c, \\ & \qquad M_3b + M_4c && \text{[from (32)]} \\ & = e_{22,30}, e_{14,18} && \text{[from (2), (4)]} \\ & \longrightarrow a && \text{[from } n_{39} \text{ demand]} \quad (39) \end{aligned}$$

which implies a is a linear combination of the terms on the left-hand side of (39). The only such term containing an a is $a + R_2b + S_7T_2b + R_8c$, so it must be the case that the linear combination's coefficient multiplying $a + R_2b + S_7T_2b + R_8c$ is the $k \times k$ identity matrix, and thus, we can conclude

$$\begin{aligned} &U_7T_2b + T_8c,\\ &c + R_3b,\\ &T_3b\\ &\quad\longrightarrow R_2b + S_7T_2b + R_8c. \end{aligned} \tag{40}$$

Therefore, using (40) and the identity

$$U_7T_2b + T_8c = T_8(R_3b + c) + (U_7T_2b - T_8R_3b)$$

gives

$$\begin{aligned} &U_7T_2b - T_8R_3b,\\ &c + R_3b,\\ &T_3b\\ &\quad\longrightarrow R_2b + S_7T_2b + R_8c \end{aligned}$$

which, in turn, implies

$$\begin{aligned} &U_7T_2b - T_8R_3b,\\ &c + R_3b,\\ &T_3b\\ &\quad\longrightarrow (R_2b + S_7T_2b + R_8c) - R_8(c + R_3b)\\ &\quad= R_2b + S_7T_2b - R_8R_3b. \end{aligned} \tag{41}$$

Since the right-hand side of (41) has no c terms and must be written as a linear combination of the terms on the left-hand side of (41), the term $c + R_3b$ on the left-hand side of (41) must have a zero $k \times k$ matrix coefficient in the linear combination. This implies

$$\begin{aligned} &U_7T_2b - T_8R_3b,\\ &T_3b\\ &\quad\longrightarrow R_2b + S_7T_2b - R_8R_3b. \end{aligned} \tag{42}$$

Now, we have

$$\begin{aligned} &(U_7T_2 - T_8R_3)b\\ &\quad= U_7T_2b - T_8(R_3 + S_6T_3)b + T_8S_6(T_3b)\\ &\quad= (U_7T_2b + T_8R_5R_2b + T_8S_5T_2b) + T_8S_6(T_3b)\\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{[from (38)]}. \end{aligned} \tag{43}$$

So

$$\begin{aligned} &U_7T_2b + T_8R_5R_2b + T_8S_5T_2b,\\ &T_3b\\ &\longrightarrow (U_7T_2 - T_8R_3)b,\\ &\qquad T_3b \qquad\qquad\qquad\qquad\qquad \text{[from (43)]}\\ &\longrightarrow (R_2 + S_7T_2 - R_8R_3)b,\\ &\qquad T_3b \qquad\qquad\qquad\qquad\qquad \text{[from (42)]}\\ &\longrightarrow (R_2 + S_7T_2 - R_8R_3)b - R_8S_6T_3b\\ &= (R_2 + S_7T_2 - R_8R_3)b\\ &\qquad + R_8(R_3 + R_5R_2 + S_5T_2)b \qquad \text{[from (38)]}\\ &= (I + R_8R_5)R_2b + (S_7 + R_8S_5)T_2b \end{aligned}$$

which implies

$$\begin{aligned} &T_2b,\\ &T_3b,\\ &T_8R_5R_2b,\\ &\quad\longrightarrow (I + R_8R_5)R_2b. \end{aligned} \tag{44}$$

In what follows, we use the fact that 2 is invertible in an odd-characteristic finite field (the inverse in GF (p^n) is $(p+1)/2$). The following relations hold:

$$\begin{aligned} &a + 2^{-1}R_2b,\\ &T_2b,\\ &T_3b,\\ &T_8R_5R_2b,\\ &T_5R_2b\\ &\longrightarrow T_5(a + 2^{-1}R_2b) + (2^{-1}T_5R_2 + U_5T_2 + U_6T_3)b,\\ &\qquad (I - R_8R_5)(a + 2^{-1}R_2b)\\ &\qquad\quad + (2^{-1}(I + R_8R_5)R_2 + S_7T_2)b, \qquad \text{[from (44)]}\\ &\qquad (U_7T_2 + 2^{-1}T_8R_5R_2)b - T_8R_5(a + 2^{-1}R_2b)\\ &= T_5a + T_5R_2b + U_5T_2b + U_6T_3b,\\ &\qquad a + R_2b + S_7T_2b - R_8R_5a,\\ &\qquad U_7T_2b - T_8R_5a. \end{aligned} \tag{45}$$

We have

$$\begin{aligned} &R_5a + c,\\ &T_5a + T_5R_2b + U_5T_2b + U_6T_3b,\\ &a + R_2b + S_7T_2b - R_8R_5a,\\ &U_7T_2b - T_8R_5a\\ &\longrightarrow R_5a + c,\\ &\qquad T_5a + T_5R_2b + U_5T_2b + U_6T_3b,\\ &\qquad (a + R_2b + S_7T_2b - R_8R_5a) + R_8(R_5a + c),\\ &\qquad (U_7T_2b - T_8R_5a) + T_8(R_5a + c)\\ &= R_5a + c,\\ &\qquad T_5a + T_5R_2b + U_5T_2b + U_6T_3b,\\ &\qquad a + R_2b + S_7T_2b + R_8c,\\ &\qquad U_7T_2b + T_8c,\\ &\longrightarrow \begin{bmatrix} (R_5a + c) + (R_5R_2 + S_5T_2 + R_3 + S_6T_3)b \\ T_5a + T_5R_2b + U_5T_2b + U_6T_3b \end{bmatrix},\\ &\qquad \begin{bmatrix} a + R_2b + S_7T_2b + R_8c \\ U_7T_2b + T_8c \end{bmatrix} \qquad \text{[from (38)]}\\ &= \begin{bmatrix} R_5 & S_5 \\ T_5 & U_5 \end{bmatrix} \begin{bmatrix} a + R_2b \\ T_2b \end{bmatrix}\\ &\qquad + \begin{bmatrix} I_k & S_6 \\ 0 & U_6 \end{bmatrix} \begin{bmatrix} R_3b + c \\ T_3b \end{bmatrix},\\ &\qquad \begin{bmatrix} I_k & S_7 \\ 0 & U_7 \end{bmatrix} \begin{bmatrix} a + R_2b \\ T_2b \end{bmatrix} + \begin{bmatrix} R_8 \\ T_8 \end{bmatrix} c\\ &= M_5e_{13,17} + M_6e_{14,18},\\ &\qquad M_7e_{13,17} + M_8c \qquad\qquad \text{[from (32), (33)]}\\ &= e_{21,29}, e_{22,30}\\ &\longrightarrow b \qquad\qquad\qquad\qquad \text{[from } n_{38} \text{ demand]}. \end{aligned} \tag{46}$$

Since there are no c terms in b and b must be written as a linear combination of the terms on the left-hand side of (46), the term $R_5a + c$ on the left-hand side of (46) must have a zero $k \times k$ matrix coefficient in the linear combination, since $R_5a + c$ is the only term on the left-hand side of (46) containing a c. This implies

$$\begin{aligned}&T_5a + T_5R_2b + U_5T_2b + U_6T_3b,\\ &a + R_2b + S_7T_2b - R_8R_5a,\\ &U_7T_2b - T_8R_5a\\ &\longrightarrow b. \qquad (47)\end{aligned}$$

Thus, (45) and (47) together imply

$$\begin{aligned}&a + 2^{-1}R_2b,\\ &T_2b,\\ &T_3b,\\ &T_8R_5R_2b,\\ &T_5R_2b\\ &\longrightarrow b\end{aligned}$$

which when combined with the fact that

$$a + 2^{-1}R_2b,\ \ b \longrightarrow a$$

gives

$$\begin{aligned}&a + 2^{-1}R_2b,\\ &T_2b,\\ &T_3b,\\ &T_8R_5R_2b,\\ &T_5R_2b\\ &\longrightarrow a, b. \qquad (48)\end{aligned}$$

Thus, the two (independent) k-dimensional vectors a and b are a linear combination of the five terms on the left-hand side of (48). Since the dimension of $a + 2^{-1}R_2b$ is k and the dimension of each of the four vectors

$$T_2b, T_3b, T_8R_5R_2b, T_5R_2b$$

is $n - k$, we must have

$$\begin{aligned}k + 4(n-k) &\geq 2k\\ k/n &\leq 4/5.\end{aligned}$$

The notion of "dimension" here is in terms of the space of values that can be taken on as the messages vary independently over F^k. Another way of phrasing this step is as follows. From the left-hand side of (48), one can compute the right-hand side. Therefore, the number of possible values for the right-hand side is no larger than the number of possible values for the left-hand side. There are $|F|^{2k}$ possible values for the right-hand side and at most $|F|^{k+4(n-k)}$ values for the left-hand side, so $k + 4(n - k) \geq 2k$.

Finally, we show that (again with $|F|$ odd) there exists a linear solution to $\mathcal{N}_1$ which achieves $k/n = 4/5$. Specifically, a $(4,5)$ fractional linear coding solution (see Fig. 4) for any alphabet which is a finite commutative ring with identity is given by (where hyphenated subscripts indicate ranges of component indices)

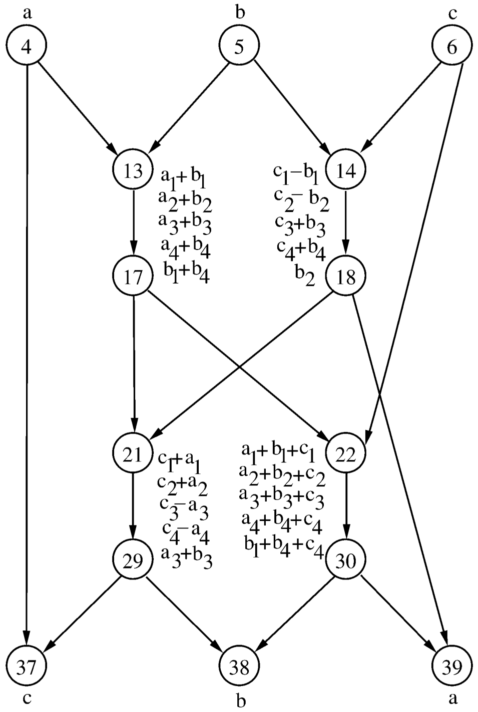

Fig. 4. Some edges are labeled to illustrate a $(4, 5)$ fractional linear solution over any ring alphabet (used in Theorem IV.3).

$$\begin{aligned}e_{13,17} &= (a+b)_{1\text{-}4},\\ &\quad\ b_1 + b_4\\ e_{14,18} &= (-b+c)_{1\text{-}2},\\ &\quad\ (b+c)_{3\text{-}4},\\ &\quad\ b_2\\ e_{21,29} &= (a+b)_{1\text{-}2} + (-b+c)_{1\text{-}2},\\ &\quad\ -(a+b)_{3\text{-}4} + (b+c)_{3\text{-}4},\\ &\quad\ a_3 + b_3\\ &= (a+c)_{1\text{-}2},\\ &\quad\ (-a+c)_{3\text{-}4},\\ &\quad\ a_3 + b_3\\ e_{22,30} &= (a+b)_{1\text{-}4} + c_{1\text{-}4},\\ &\quad\ (b_1 + b_4) + c_4\\ &= (a+b+c)_{1\text{-}4},\\ &\quad\ b_1 + b_4 + c_4\\ n_{37} : c_{1\text{-}2} &= (a+c)_{1\text{-}2} - a_{1\text{-}2}\\ c_{3\text{-}4} &= (-a+c)_{3\text{-}4} + a_{3\text{-}4}\\ n_{38} : b_{1\text{-}2} &= (a+b+c)_{1\text{-}2} - (a+c)_{1\text{-}2}\\ b_3 &= -(a_3+b_3+c_3) + (-a_3+c_3) + 2(a_3+b_3)\\ b_4 &= -2(a_1+b_1+c_1) - (a_4+b_4+c_4)\\ &\quad + 2(b_1+b_4+c_4) + 2(a_1+c_1) - (-a_4+c_4)\\ n_{39} : a_1 &= (a_1+b_1+c_1) - 2(b_1+b_4+c_4)\end{aligned}$$

$$\begin{aligned} &-(-b_1+c_1)+2(b_4+c_4)\\ a_2 &= (a_2+b_2+c_2)-(-b_2+c_2)-2b_2\\ a_{3\text{-}4} &= (a+b+c)_{3\text{-}4}-(b+c)_{3\text{-}4}. \end{aligned} \qquad \square$$

Theorem IV.4: The linear coding capacity of the network $\mathcal{N}_2$ is $10/11$ over any even-characteristic finite field and is 1 over any odd-characteristic finite field.

Proof: If the alphabet is an odd-characteristic finite field F, then a scalar linear solution (i.e., with $k = n = 1$) is guaranteed by Lemma II.3. Thus, the linear coding capacity of $\mathcal{N}_2$ is at least 1 for odd $|F|$. Also, the linear coding capacity is bounded above by the coding capacity, which equals 1 by Lemma II.3. Thus, the linear coding capacity of $\mathcal{N}_2$ is 1 if $|F|$ is odd.

Henceforth in the proof, assume the alphabet is a finite field F with characteristic two. Suppose there exists a (k, n) fractional linear solution for $\mathcal{N}_2$ over F. First we show in general that k/n cannot exceed $10/11$ and then we demonstrate a specific linear code which achieves $k/n = 10/11$.

Since the coding capacity of $\mathcal{N}_2$ is 1 by Lemma II.3, we may assume $k \le n$, for otherwise the linear coding capacity would exceed the coding capacity. Let

$$\delta = n - k$$

and assume a (k, n) fractional linear solution for $\mathcal{N}_2$ with the same labeling of M_i's assumed in Lemma II.3.

Given any $m \times n$ matrix M_i over F of rank r, we can find an $(n-r) \times n$ matrix Q_i over F such that

$$\mathsf{rank}\left(\begin{bmatrix} M_i \\ Q_i \end{bmatrix}\right) = n.$$

(Choose r independent rows of M_i, find $n - r$ more members of F^n which together with the r rows form a basis of F^n, and let the rows of Q_i be these $n - r$ members of F^n.) So we have

$$\mathsf{rank}(Q_i) = n - r$$

and

$$Q_i x, M_i x \longrightarrow x \tag{49}$$

(since by Lemma IV.2, there exists an $n \times (m + n - r)$ matrix B such that $B\begin{bmatrix} M_i \\ Q_i \end{bmatrix} = I_n$). There is some flexibility in the choice of Q_i that we will use later.

We have

$$\begin{aligned} 0 = {} & M_{10}(M_1a+M_2b)\\ & + M_{11}(M_7a+M_8b) \qquad \text{[from (21)]} \end{aligned} \tag{50}$$

$$\begin{aligned} 0 = {} & M_{12}(M_3a+M_4c)\\ & + M_{13}(M_7a+M_9c) \qquad \text{[from (22)]} \end{aligned} \tag{51}$$

$$\begin{aligned} 0 = {} & M_{14}(M_5b+M_6c)\\ & + M_{15}(M_8b+M_9c) \qquad \text{[from (23)]} \end{aligned} \tag{52}$$

and so

$$M_7a+M_8b \longrightarrow M_{10}(M_1a+M_2b) \qquad \text{[from (50)]} \tag{53}$$

$$M_3a+M_4c \longrightarrow M_{13}(M_7a+M_9c) \qquad \text{[from (51)]} \tag{54}$$

$$M_5b+M_6c \longrightarrow M_{15}(M_8b+M_9c). \qquad \text{[from (52)].} \tag{55}$$

Also, since (24) gives $M_{15}M_7 = M_{13}M_8 = I_k$, we have

$$M_7a \longrightarrow a \tag{56}$$

$$M_8b \longrightarrow b. \tag{57}$$

Now, let Q_{10}, Q_{13}, Q_{15} correspond to M_{10}, M_{13}, M_{15} as above. Let L be the list consisting of

$$\begin{aligned} &M_3a+M_4c\\ &M_5b+M_6c\\ &Q_{13}(M_7a+M_9c)\\ &Q_{15}(M_8b+M_9c)\\ &Q_{10}(M_1a+M_2b) \end{aligned}$$

and the corresponding five objects from the c, d, e side of $\mathcal{N}_2$.

Claim: $L \longrightarrow a, b, c, d, e.$

Proof of Claim:

$$L \longrightarrow M_{13}(M_7a+M_9c) \qquad \text{[from (54)]} \tag{58}$$

$$L \longrightarrow M_7a+M_9c \qquad \text{[from (49), (58)]} \tag{59}$$

$$L \longrightarrow M_{15}(M_8b+M_9c) \qquad \text{[from (55)]} \tag{60}$$

$$L \longrightarrow M_8b+M_9c \qquad \text{[from (49), (60)].} \tag{61}$$

Note that since F has characteristic two

$$(M_7a+M_9c)+(M_8b+M_9c) = M_7a+M_8b. \tag{62}$$

We get

$$L \longrightarrow M_7a+M_8b \qquad \text{[from (59), (61), (62)]} \tag{63}$$

$$L \longrightarrow M_{10}(M_1a+M_2b) \qquad \text{[from (53), (63)]}$$

$$L \longrightarrow M_1a+M_2b \qquad \text{[from (49), (64)]} \tag{64}$$

and so

$$L \longrightarrow e_{23,31}, e_{24,32}, e_{25,33}. \tag{65}$$

The same reasoning on the other side of the network gives

$$L \longrightarrow e_{26,34}, e_{27,35}, e_{28,36}. \tag{66}$$

Now, the decoding function at node n_{43} gives

$$L \longrightarrow c \qquad \text{[from (65), (66)]} \tag{67}$$

$$L \longrightarrow M_7a \qquad \text{[from (59), (67)]} \tag{68}$$

$$L \longrightarrow a \qquad \text{[from (56), (68)]}$$

$$L \longrightarrow M_8b \qquad \text{[from (61), (67)]} \tag{69}$$

$$L \longrightarrow b \qquad \text{[from (57), (69)]}$$

$$L \longrightarrow d, e \ \text{[from same reasoning on other side of } \mathcal{N}_2\text{].} \tag{70}$$

$\square$

As in the argument for $\mathcal{N}_1$, we can think of each n-length edge vector as one part of length k followed by one part of length $n - k$, and break up the matrices M_i accordingly, so

$$\begin{aligned} M_i &= \begin{bmatrix} R_i \\ T_i \end{bmatrix}, && \text{for } 1 \le i \le 9\\ M_i &= \begin{bmatrix} R_i & S_i \end{bmatrix}, && \text{for } 10 \le i \le 15. \end{aligned}$$

Claim: We may assume without loss of generality that $R_9 = I_k$ and $T_9 = 0$.

Proof of Claim: Since $M_{11}M_9 = I_k$ by (24), using Lemma IV.1 we have

$$\begin{aligned} k = \min(k,n) &\geq \mathsf{rank}(M_9) \\ &\geq \mathsf{rank}(M_{11}M_9) = \mathsf{rank}(I_k) = k \end{aligned}$$

so $\mathsf{rank}(M_9) = k$. By Lemma IV.2, there exists an $n \times n$ invertible matrix W such that

$$WM_9 = \begin{bmatrix} I_k \\ 0 \end{bmatrix}.$$

Replace

$$M_7, M_8, M_9, M_{11}, M_{13}, M_{15}$$

with

$$WM_7, WM_8, WM_9, M_{11}W^{-1}, M_{13}W^{-1}, M_{15}W^{-1}$$

respectively; this will yield a new (k,n) fractional linear solution, because the W^{-1}'s will cancel the W's, and the new M_9 will be $\begin{bmatrix} I_k \\ 0 \end{bmatrix}$. □

Now $M_{11}M_9 = I_k$ (from (24)) gives $R_{11} = I_k$.

Claim: We may assume without loss of generality that $S_{11} = 0$.

Proof of Claim: Let W_2 be the $n \times n$ matrix

$$\begin{bmatrix} I_k & S_{11} \\ 0 & I_{n-k} \end{bmatrix}$$

and replace

$$M_7, M_8, M_9, M_{11}, M_{13}, M_{15}$$

with

$$W_2M_7, W_2M_8, W_2M_9, M_{11}W_2^{-1}, M_{13}W_2^{-1}, M_{15}W_2^{-1}$$

respectively. This is still a (k,n) fractional linear solution, M_9 is still $\begin{bmatrix} I_k \\ 0 \end{bmatrix}$, and M_{11} is now $[\, I_k \quad 0 \,]$. □

Now let

$$\alpha = \mathsf{rank}\left([\, R_7 \quad R_8 \,]\right). \tag{71}$$

From (50) and the fact that F has characteristic two, we have

$$M_{10}(M_1a + M_2b) = M_{11}(M_7a + M_8b) = R_7a + R_8b$$

and thus

$$M_{10}[\, M_1 \quad M_2 \,] = [\, R_7 \quad R_8 \,]$$

which implies $\mathsf{rank}(M_{10}) \geq \alpha$. Therefore,

$$\mathsf{rank}(Q_{10}) \leq n - \alpha. \tag{72}$$

By (71), there exists a $k \times k$ permutation matrix P such that the top α rows of $P[\, R_7 \quad R_8 \,]$ are independent. Therefore, each of the bottom $k - \alpha$ rows of $P[\, R_7 \quad R_8 \,]$ can be written as a linear combination of the first α rows of the same matrix. Thus, there exists a $(k-\alpha) \times k$ matrix X' whose right-most $k - \alpha$ columns form the matrix $-I_{k-\alpha}$ (and, therefore, $\mathsf{rank}(X') = k - \alpha$) such that

$$X'P[\, R_7 \quad R_8 \,] = 0.$$

Let

$$X = X'P$$

and note that $\mathsf{rank}(X) = k - \alpha$ since

$$\begin{aligned} k - \alpha &= \min(k, k-\alpha) \\ &\geq \mathsf{rank}(X) \\ &= \mathsf{rank}(X'P) \\ &\geq \mathsf{rank}(X') + \mathsf{rank}(P) - k \quad \text{[from Lemma IV.1]} \\ &= (k - \alpha) + k - k \\ &= k - \alpha. \end{aligned}$$

Define a $(k - \alpha) \times n$ matrix

$$Y = [\, X \quad 0 \,]$$

and note that

$$YM_7 = YM_8 = 0.$$

Since $M_{13}M_8 = I_k$, the rows of M_{13} and the rows of Y are jointly linearly independent. (If v is a nontrivial linear combination of rows of M_{13}, then $vM_8 \neq 0$; if v' is a linear combination of rows of Y, then $v'M_8 = 0$, so $v \neq v'$.) Therefore, we may choose Q_{13} so that its first $k - \alpha$ rows are Y. Similarly, since $M_{15}M_7 = I_k$ and $YM_7 = 0$, we may choose Q_{15} so that its first $k - \alpha$ rows are Y. Now, in the left half of the list L, the $M_3a + M_4c$ and the $M_5b + M_6c$ give n entries each. The $Q_{13}(M_7a + M_9c)$ and $Q_{15}(M_8b + M_9c)$ each give $\delta = n - k$ entries, but the first $k - \alpha$ entries from each of them are

$$\begin{aligned} Y(M_7a + M_9c) &= YM_9c \quad \text{and} \\ Y(M_8b + M_9c) &= YM_9c \end{aligned}$$

which are the same; hence, these two together give only $2\delta - (k-\alpha)$ new entries. Finally, by (72), the $Q_{10}(M_1a + M_2b)$ gives only $n - \alpha$ independent entries. Therefore, the left half of L has only

$$n + n + [2\delta - (k - \alpha)] + [n - \alpha] = 2n + 3\delta$$

independent entries. The same applies to the right half, so L has at most $4n + 6\delta$ independent entries. Therefore,

$$\begin{aligned} 4n + 6(n - k) &\geq 5k \\ k/n &\leq 10/11. \end{aligned}$$

Finally, we show that (over any ring) there exists a (k,n) fractional linear solution to $\mathcal{N}_2$ which achieves $k/n = 10/11$. Specifically, a $(10, 11)$ fractional linear coding solution is given by (where hyphenated subscripts indicate ranges of component indices)

$$\begin{aligned} e_{15,19} &= ((a + b + c)_{3\text{-}10}, a_1 + b_1, c_1 + a_2, b_2 + c_2) \\ e_{23,31} &= ((a + b)_{3\text{-}10}, a_2, b_2, b_3) \\ e_{24,32} &= ((a + c)_{3\text{-}10}, a_1, c_2, c_4) \end{aligned}$$

$$\begin{aligned}
e_{25,33} &= ((b+c)_{3\text{-}10}, b_1, c_1, c_5)\\
e_{16,20} &= ((c+d+e)_{1\text{-}8}, c_9+d_9, e_9+c_{10}, d_{10}+e_{10})\\
e_{26,34} &= ((c+d)_{1-8}, c_{10}, d_{10}, c_6)\\
e_{27,35} &= ((c+e)_{1\text{-}8}, c_9, e_{10}, c_7)\\
e_{28,36} &= ((d+e)_{1\text{-}8}, d_9, e_9, d_8)\\
n_{40}: c_{3\text{-}10} &= (a+b+c)_{3\text{-}10} - (a+b)_{3\text{-}10}\\
c_1 &= (c_1+a_2) - a_2\\
c_2 &= (b_2+c_2) - b_2\\
n_{41}: b_{3-10} &= (a+b+c)_{3-10} - (a+c)_{3-10}\\
b_1 &= (a_1+b_1) - a_1\\
b_2 &= (b_2+c_2) - c_2\\
n_{42}: a_{3\text{-}10} &= (a+b+c)_{3\text{-}10} - (b+c)_{3\text{-}10}\\
a_1 &= (a_1+b_1) - b_1\\
a_2 &= (c_1+a_2) - c_1\\
n_{43}: c_3 &= (b+c)_3 - b_3\\
c_8 &= (c+d)_8 - d_8\\
&\text{Already known}: c_{1\text{-}2}, c_{4\text{-}7}, c_{9\text{-}10}\\
n_{44}: e_{1\text{-}8} &= (c+d+e)_{1-8} - (c+d)_{1-8}\\
e_9 &= (e_9+c_{10}) - c_{10}\\
e_{10} &= (d+e)_{10} - d_{10}\\
n_{45}: d_{1\text{-}8} &= (c+d+e)_{1\text{-}8} - (c+e)_{1\text{-}8}\\
d_9 &= (c_9+d_9) - c_9\\
d_{10} &= (d+e)_{10} - e_{10}\\
n_{46}: c_{1\text{-}8} &= (c+d+e)_{1\text{-}8} - (d+e)_{1-8}\\
c_9 &= (c_9+d_9) - d_9\\
c_{10} &= (e_9+c_{10}) - e_9. \qquad \square
\end{aligned}$$

The next corollary follows immediately from Theorems IV.3 and IV.4, and together with Corollary II.5 shows that the coding capacity of $\mathcal{N}_3$ (i.e., 1) is exactly 10% greater than the maximum linear coding capacity (i.e., $10/11$) over any finite field.

Corollary IV.5: The linear coding capacity of the network $\mathcal{N}_3$ is $10/11$ over any even-characteristic finite field and is $4/5$ over any odd-characteristic finite field.

From this and Corollary II.5, we get the following.

Corollary IV.6: There exists a solvable network which is not asymptotically linearly solvable. In other words, linear network codes are asymptotically insufficient over finite fields.

Asymptotic insufficiency allows us to deduce results about the extended linear coding methods known as convolutional coding and filter-bank network coding (see [8] for the definitions). This is because of the following simple result which appears to be well known (it is just a variant of Lemma 8 from [8]):

Proposition IV.7: If a network is solvable by means of convolutional coding or filter-bank coding, then it is asymptotically linearly solvable.

Proof: Both convolutional coding and filter-bank coding use inputs and outputs that are (potentially) infinite sequences of members of a finite field F. They both have the feature that there is a fixed nonnegative integer d (the delay) such that, for any $k \geq 0$, the coding will produce (in a linear way) the first k components of each output given only the first $k+d$ components of each source message. Hence, for any k, we can obtain a linear $(k, k+d)$ fractional coding solution for the network as follows. Each interior node will take $k+d$ inputs and produce $k+d$ outputs on each output edge by the same procedure as would have been applied in the first $k+d$ steps of the convolutional or filter-bank code. A source node will be supplied with k inputs for each source message; it will append d 0's and then simulate what the convolutional or filter-bank code would do in the first $k+d$ steps. Finally, the decoding operations will simulate the first $k+d$ steps of the convolutional or filter-bank decoding operations and then output only the k relevant entries. Each of these node operations is linear by the definition of convolutional or filter-bank coding, and they form a solution because the given coding scheme was a solution.

Since $k/(k+d)$ becomes arbitrarily close to 1 as k increases, the network is asymptotically linearly solvable. □

Therefore we get the following.

Corollary IV.8: There exists a solvable network which is not solvable by means of convolutional coding or filter-bank coding.

A more general network coding model allows different rates for different source messages. This corresponds to a collection of source dimensions $(k_1, k_2, \ldots, k_m)$, where m is the number of messages. The linear rate region consists of the set of all points in $\mathbb{R}^m$ of the form

$$\left(\frac{k_1}{n}, \frac{k_2}{n}, \ldots, \frac{k_m}{n}\right)$$

for which there exists a fractional $(k_1, k_2, \ldots, k_m, n)$ linear coding solution. See [18] for an alternate definition of such a region. One might consider a form of "linear coding" where linear codes on different finite-field alphabets are time shared. This corresponds to taking the convex hull of the linear rate region of a network.

It turns out that even this form of linearity is not sufficient for our example network. To see this, note that for the network in Fig. 3, each of the source messages a, b, c, d, e is demanded by at least one network node which can be reached by exactly one directed path from the corresponding message. It follows easily that any point $(r_1, r_2, r_3, r_4, r_5)$ in the linear rate region must satisfy $r_i \leq 1$ for all i. But we must also have $r_i \leq 10/11$ for at least one i, by Corollary IV.5. Hence, the sum $r_1 + \cdots + r_5$ must be at most $54/11 < 5$. This is true for any point in the linear rate region, and hence for any point in the convex hull of the region. Therefore, the point $(1, 1, 1, 1, 1)$ is not in this convex hull or even in its closure. So time sharing different linear codes will not allow us to achieve or even approach capacity 1 for this network.

Acknowledgment

The authors thank Kris Popat for providing a helpful suggestion.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1204–1216, Jul. 2000.

[2] J. Cannons, R. Dougherty, C. Freiling, and K. Zeger, "Network routing capacity," *IEEE/ACM Trans. Networking*, submitted for publication.

[3] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2003.

[4] R. Dougherty, C. Freiling, and K. Zeger, "Linearity and solvability in multicast networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2243–2256, Oct. 2004.

[5] M. Feder, D. Ron, and A. Tavory, "Bounds on linear codes for network multicast," in *Proc. Electronic Colloquium on Computational Complexity (ECCC)*, 2003, Rep.33, pp. 1–9.

[6] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "Toward a random operation of networks," *IEEE Trans. Inf. Theory*. [Online]. Available: http://web.mit.edu/trace/www/, submitted for publication.

[7] N. Jacobson, "Some remarks on one-sided inverses," *Proc. Amer. Math. Soc.*, vol. 1, no. 3, pp. 352–355, Jun. 1950.

[8] S. Jaggi, M. Effros, T. Ho, and M. Médard, "On linear network coding," in *Proc. 42st Annu. Allerton Conf. Communication Control and Computing*, Monticello, IL, Oct. 2004.

[9] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973–1982, Jun. 2005.

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

[11] A. Rasala Lehman and E. Lehman, "Complexity classification of network information flow problems," in *Proc. 41st Annu. Allerton Conf. Communication Control and Computing*, Monticello, IL, Oct. 2003.

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

[13] M. Médard, M. Effros, T. Ho, and D. Karger, "On coding for nonmulticast networks," in *Proc. 41st Annu. Allerton Conf. Communication Control and Computing*, Monticello, IL, Oct. 2003.

[14] S. Riis. (2003) Linear versus non-linear Boolean functions in network flow. Tech. Rep. [Online]. Available: http://nick.dcs.qmul.ac.uk/~smriis

[15] ——, "Linear versus nonlinear boolean functions in network flow," in *Proc. 38th Annu. Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2004.

[16] I. Satake, *Linear Algebra*. New York: Marcel Dekker, 1975.

[17] S. R. Searle, *Matrix Algebra Useful for Statistics*. New York: Wiley, 1982.

[18] R. W. Yeung, *A First Course in Information Theory*. Norwell, MA: Kluwer, 2002.