A coding view of network recovery and management for single-receiver communications

Tracey Ho and Muriel Médard
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139
e-mail: {trace, medard}@mit.edu

Ralf Koetter
Coordinated Science Laboratory
University of Illinois
Urbana, IL 61801
e-mail: koetter@csl.uiuc.edu

Abstract — **Communication on a network can be described by codes specifying the behavior of the network nodes, i.e the relationships between the signals on each node's incoming and outgoing links. Link failure recovery may require some of these relationships to change. The network management information needed to effect the appropriate changes in network behavior for recovery can be quantified by the number of different codes needed. In this paper, we focus on linear codes for delay-free acyclic networks, with a single receiver node and one or more source nodes. We consider two formulations for quantifying network management. The first is a centralized formulation where network behavior is described by an overall code determining the behavior of every node, and the management requirement is taken as the log of the number of such codes that the network may switch among. For this formulation, we give tight upper and lower bounds on network management for receiver-based and network-wide recovery from all single link failures, assuming they are recoverable. The second is a node-based formulation where the management requirement is taken as the sum over all nodes of the log of the number of different behaviors for each node. We show that the minimum node-based management requirement for terminal link failures and the no-failure scenario is achieved with receiver-based schemes.**

I. Introduction

Information on networks is traditionally transmitted by routing and replication. Recent work on network capacity, considering multicast connections, has shown that achieving multicast capacity on certain network topologies requires coding, i.e. combination of different signals [1]. A set of source-receiver connections is feasible if and only if the connection rate satisfies the max-flow min-cut bound [1], and, if so, can be achieved using linear coding [2] (i.e. where outputs from a node are linear combinations of the inputs to that node). Reference [3] gives an algebraic framework for checking the feasibility of a set of network connections in polynomial time. This algebraic approach allows the derivation of the results in [1] and [2], and provides powerful tools for considering multicast and non-multicast network connections.

It is not yet clear to what extent codes are useful from a capacity point of view. Apart from the example in [2] and related constructions, our searches over several thousand randomly generated graphs have not yielded other examples in which coding is necessary to achieve multicast capacity [6].

However, coding can be very useful for robust recovery from link failures. Traditionally, link failure recovery is achieved by re-routing, which can be considered a special case of coding. A surprising result is that with coding, a linear multicast network has, for all recoverable failures, a solution in which only the receiver nodes need to be informed of the failure pattern [3]. Thus, there exists a coding strategy which is static for all non-receiver nodes. However, the use of such a static approach may entail some additional complexity. For example, the network in Figure 1, with one source node multicasting two processes to two receivers, is recoverable for all single link failures in an $\mathbb{F}_2$ field with network-wide codes but not with receiver-based codes. Bounds on the required size of the code alphabet are given in [3].

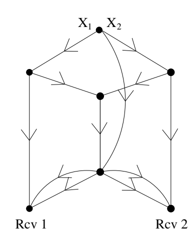

Figure 1: An example network in which recovery is achievable in an $\mathbb{F}_2$ field with network-wide codes but not with receiver-based codes.

Network coding naturally leads to consideration of network management. Indeed, network management, in the form of network state information, is needed to modify the behavior of the receiver nodes, or any other nodes involved, to achieve recovery under different failure scenarios. Apart from Gallager's 1976 paper [4] on the protocol information needed to keep track of source and receiver addresses and of the starting and stopping of messages, not much theoretical work has been done in the area of network management [5].

We describe network behavior by a code specifying the relationships between the signals on the input and output links of each node, which may have to change in order to recover from different link failures. The network management information needed to effect the appropriate changes in network behavior for recovery can be quantified by the number of different codes needed.

In this paper we consider delay-free acyclic networks, with a single receiver node and one or more source nodes. This includes the point-to-point and incast cases. We also focus on linear codes. If non-linear processing is permitted, the network management requirements may be different in some cases [6].

We consider two formulations for quantifying network management. The first is a centralized formulation where network behavior is described by an overall code determining the behavior of every node, and the management requirement is taken as the log of the number of such codes that the network may switch among. We give, in Section C, tight upper and lower bounds on network management for receiver-based and network-wide recovery from all single link failures, assuming they are recoverable.

An alternative formulation for quantifying network management, discussed in Section D, takes into account the number of nodes which change behavior, and the number of different behaviors for each node. Taking the node-based management requirement to be the sum over all nodes of the log of the number of different behaviors for each node, we show that the minimum node-based management requirement for terminal link failures and the no-failure scenario is achieved with receiver-based schemes.

II. Linear coding model

A Problem formulation

As in [3], we represent a network by a directed graph $\mathcal{G} = (V, E)$ where V is the vertex set and $E \subseteq V \times V$ the edge set. There are one or more source nodes, at which one or more discrete independent random processes X_i are observable. Processes originating at different source nodes are assumed to be independent. In this paper we consider the single-receiver case, and call the receiver node β. The network connection problem is to transmit all the source processes to the receiver through the network.

An edge e is an *incident outgoing link* of node v if $v = \text{tail}(e)$, and an *incident incoming link* of v if $v = \text{head}(e)$. We call an incident incoming link of the receiver a *terminal link*, and denote the set of terminal links by $\mathcal{T}$. Links which are not terminal links are called *interior links*. Edges e_1 and e_2 are *incident* if $\text{head}(e_1) = \text{tail}(e_2)$ or $\text{head}(e_2) = \text{tail}(e_1)$. A path is a sequence of distinct nodes that are connected by links. If there is a directed path from a link or node to another, the former is said to be *upstream* of the latter, and the latter *downstream* of the former. Edge e carries the random process $Y(e)$. Output processes at the receiver node are denoted Z_i.

We choose the time unit such that the capacity of each link is 1 bit per unit time, and the random processes X_i have a constant entropy rate of one bit per unit time. Edges with larger capacities are modeled as parallel edges, and sources of larger entropy rate are modeled as multiple sources at the same node.

The processes X_i, $Y(e)$, Z_i are binary sequences. We assume that information is transmitted as vectors of bits which are of equal length m, and can be represented as elements in the finite field $\mathbb{F}_{2^m}$. The length of the vectors is equal in all transmissions and all links are assumed to be synchronized with respect to the symbol timing.

B Transfer matrices

A linear code for a given acyclic network can be specified by transfer matrices with network-constrained zero positions, defined as follows [3], where the matrix coefficients are elements in $\mathbb{F}_{2^m}$:

1. $r \times \nu$ matrix A specifies how the source processes X_i, $i = 1, \ldots, r$ are represented on the source nodes' incident outgoing links, where r is the total number of source processes and ν the number of network links. The signal on source link j is
$$Y(j) = \sum_{i \in \mathcal{A}_{\text{tail}(j)}} A(i,j) X_i$$
where X_i, $i \in \mathcal{A}_{\text{tail}(j)}$ are processes generated at node $\text{tail}(j)$ and transmitted on j.
2. $\nu \times \nu$ matrix F specifies how signals are transmitted between incident links. $F(i,j)$ is nonzero only if $\text{head}(i) = \text{tail}(j)$, so the signal on link j is
$$Y(j) = \sum_{i\,:\,\text{head}(i) = \text{tail}(j)} F(i,j) Y(i)$$
3. $\nu \times \nu$ matrix $G = I + F + F^2 + \ldots$ sums the gains along all paths between each pair of links, and equals $(I - F)^{-1}$, since matrix F is nilpotent. Link i is said to *feed into* link j if $G(i,j)$ is nonzero.
4. $r \times \nu$ matrix B specifies how the receiver outputs Z_j are obtained, where the number of output processes r is equal to the number of input processes. Then
$$Z_j = \sum_{i \in \mathcal{T}} B(j,i) Y(i)$$
where $Y(i)$, $i \in \mathcal{T}$ are signals on the receiver node's terminal links i.

We call a triple (A, F, B) a code. Note that a code (A, F, B) is equivalently specified by the triple (A, G, B), where $G = (I - F)^{-1}$. A pair (A, F), or (A, G), is called an *interior code*.

We will use the following notation in this paper:

$\underline{c}_j$ is column j of AG. It can be viewed as a *signal map* specifying the signal $Y(j) = [X_1\ X_2\ \ldots\ X_r]\,\underline{c}_j$ given the vector $[X_1\ X_2\ \ldots\ X_r]$ of source processes.

$\underline{b}_j$ is column j of B.

$G_{\mathcal{K}}$ is the submatrix of G consisting of columns that correspond to links in a set $\mathcal{K}$.

$B_{\mathcal{K}}$ is the submatrix of B consisting of columns that correspond to links in a set $\mathcal{K}$.

G^h, $G^h_{\mathcal{K}}$ and $\underline{c}^h_j$ are the altered values of G, $G_{\mathcal{K}}$ and $\underline{c}_j$, respectively, resulting from failure of link h.

$G^{\mathcal{H}}$, $G^{\mathcal{H}}_{\mathcal{K}}$ and $\underline{c}^{\mathcal{H}}_j$ are the altered values of G, $G_{\mathcal{K}}$ and $\underline{c}_j$, respectively, under the combined failure of links in set $\mathcal{H}$.

An example illustrating the structure of the transfer matrices is given in Figure 2.

C Validity of codes for the network connection problem

For the source processes to be reconstructable at the receiver, the transfer matrix AGB^T must be nonsingular [3]. We call a code (A, G, B) *valid* for the network connection problem if $AGB^T = I$, where the specification of a particular value (chosen here to be I) for AGB^T imposes the requirement that the receiver correctly identifies which signal corresponds to which input process and outputs them in specific order (in this case, the same order as the input processes).

An interior code (A, G) is called *valid* for the network connection problem if there exists some B for which (A, G, B) is a valid code for the problem. Note that $AGB^T = AG_{\mathcal{T}}B^T_{\mathcal{T}}$, where the entries of the $r \times |\mathcal{T}|$ matrix $B_{\mathcal{T}}$ are not constrained by the network. The rank of $AG_{\mathcal{T}}B^T_{\mathcal{T}}$ is thus constrained by the rank of $AG_{\mathcal{T}}$. Thus, (A, G) is a valid interior code if and only if $AG_{\mathcal{T}}$ has rank r.

D Link failure model

Assuming that a constant zero signal is observed on failed links, failure of link h corresponds to setting to zero the h^{th} column of matrices A, B and F, and the h^{th} row of F. The overall transfer matrix after failure of link h is $AI^hG^h(BI^h)^T = AG^hB^T$, where $I^h = I - \delta_{hh}$ is the identity matrix with a zero in the $(h,h)^{th}$ position, $F^h = I^hFI^h$, and $G^h = I^h + F^h + (F^h)^2 + \ldots = I^h\left(I - FI^h\right)^{-1} = \left(I - I^hF\right)^{-1} I^h$.

If failure of link h is recoverable, there exists some (A', G', B') such that $A'{G'}^h{B'}^T = I$. We say that a code (A', G', B') *covers* (failure of) link h if $A'{G'}^h{B'}^T = I$.

Recovery codes can be classified as receiver-based (only B changes), transmitter-based (only A changes) or network-wide (any combination of A, F and B may change). An interior code (A, G) in a receiver-based scheme is called *static*.

III. Link failure recovery and management

Nodes which need to change behavior for recovery from different failures must be appropriately directed by management signals. We seek to characterize the network management requirements based on parameters of the network connection problem, as well as the type of recovery codes used, whether receiver-based or network-wide.

In this paper we consider two formulations for quantifying the essential management information. The first is a centralized formulation where the behavior of the network is described by a common code determining the behavior of all nodes. Network management can be viewed as a mapping from the code indices to the state of network behavior. If n_c different codes are needed to achieve recovery from different failure scenarios, then the network management requirement is $\lceil \log_2 n_c \rceil$ bits. An alternative node-based formulation for quantifying network management takes into account the number of nodes which change behavior, and the number of different behaviors for each node.

A Codes for multiple failure scenarios

We first characterize codes which can be used for multiple single link failures. The results in this section, from [6], provide concepts and tools which are useful in proving the results in later sections.

Lemma 1 *Let $\mathcal{T}^h \subseteq \mathcal{T}$ be the set of terminal links of receiver β that are downstream of link h.*
1. If code (A, G, B) covers the no-failure scenario and failure of link h, then $\underline{c}_h \sum_{j\in\mathcal{T}^h} G(h,j)\underline{b}_j^T = \mathbf{0}$.
2. If code (A, G, B) covers failures of links h and k, then either

(a) $\underline{c}_h \sum_{j\in\mathcal{T}^h} G(h,j)\underline{b}_j^T = \mathbf{0}$
and $\underline{c}_k \sum_{j\in\mathcal{T}^k} G(k,j)\underline{b}_j^T = \mathbf{0}$
or
(b) $\gamma_{h,k} \sum_{j\in\mathcal{T}^h} G(h,j)\underline{b}_j = \sum_{j\in\mathcal{T}^k} G(k,j)\underline{b}_j \neq \underline{0}$
and $\underline{c}_h = \gamma_{h,k}\underline{c}_k \neq \underline{0}$
where $\gamma_{h,k} \in \mathbb{F}_{2^m}$ is a constant for given h, k

Proof outline The results follow from writing $AG^h_{\mathcal{T}}B^T_{\mathcal{T}}$ in the form $\sum_{j\in\mathcal{T}} \underline{c}^h_j \underline{b}^T_j$ and noting that $\Delta\underline{c}^h_j = \underline{c}_j - \underline{c}^h_j = G(h,j)\underline{c}_h$. □

These expressions simplify considerably for terminal links as follows:

Corollary 1 *1. If code (A, G, B) covers the no-failure scenario and failure of terminal link h, then $\underline{c}_h\underline{b}_h^T = \mathbf{0}$.*
2. If (A, G, B) covers failures of terminal links h and k of β, then either

(a) $\underline{c}_h\underline{b}_h^T = \mathbf{0}$ *and* $\underline{c}_k\underline{b}_k^T = \mathbf{0}$
or
(b) $\gamma_{h,k}\underline{b}_h = \underline{b}_k \neq \underline{0}$ *and* $\underline{c}_h = \gamma_{h,k}\underline{c}_k \neq \underline{0}$
where $\gamma_{h,k} \in \mathbb{F}_{2^m}$ is a constant for given h, k

These results lead to the notion of active and non-active recovery codes. A recovery code which is *active* in a failed link h is one in which $AG^h_{\mathcal{T}}B^T_{\mathcal{T}}$ is affected by the value on link h, i.e. $\underline{c}_h \sum_{j\in\mathcal{T}^h} G(h,j)\underline{b}_j^T \neq \mathbf{0}$. Otherwise, the code is *non-active* in h.

In a code which is non-active in a failed link, the value on that link is set to zero (by upstream links ceasing to transmit on the link), cancelled out, or disregarded by the receivers. Part 1 of Lemma 1 states that a code which covers the no-failure scenario as well as one or more single link failures must be non-active in those links. Part 2 of Lemma 1 states

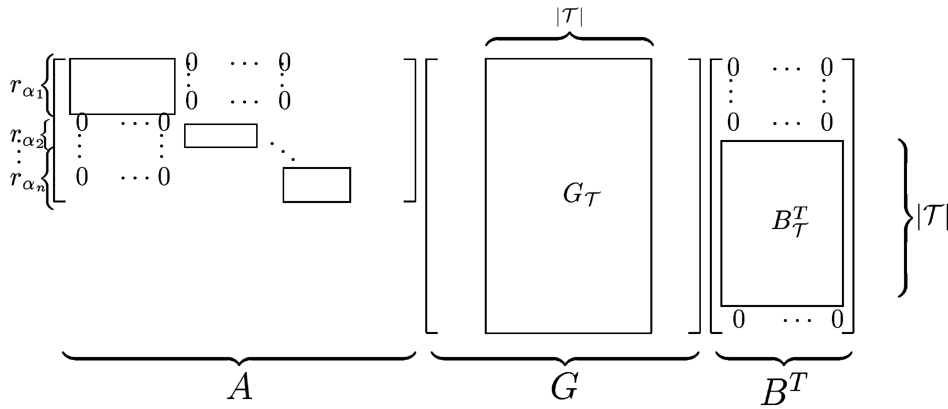

Figure 2: An example illustrating the structure of transfer matrices.

that a code which covers failures of two or more single links is either non-active in all of them (case a) or active in all of them (case b). In the latter case, those links carry signals that are multiples of each other. We term a code *active* if it is active in those links whose failures it covers, and *non-active* otherwise. Active codes cannot be used if signals on failed links are undetermined.

This classification is very useful in later proofs, as the following results from [6] show that a set of codes with a common AG matrix, covering terminal link failures, corresponds to a set of non-active codes each covering the same terminal link failures, and vice versa. The analysis for terminal link failures provides tools and results useful in the analysis of network management for all link failures.

Lemma 2 *In a given network, for any set of non-active codes* $\{(A_1, G_1, B_1),\ (A_2, G_2, B_2), \ldots, (A_n, G_n, B_n)\}$, *there exists a set of receiver-based codes* $\{(A, G, B_1'),\ (A, G, B_2'), \ldots, (A, G, B_n')\}$, *such that* (A, G, B_i') *covers the same terminal link failures as* (A_i, G_i, B_i), *for all* $i = 1, \ldots, n$.

Corollary 2 *Let* $n_\mathcal{T}$ *be the number of codes needed to cover the no-failure scenario and all single terminal link failures for a given network connection problem in which all such failures are recoverable. If only non-active codes are used, then* $n_\mathcal{T}$ *is the same for network-wide and receiver-based codes.*

Corollary 3 *The terminal link failures covered by each code in a network-wide scheme can be covered by one or two codes in a receiver-based scheme.*

Lemma 3 *For any set of* $n \geq 2$ *codes with a common* (A, G) *covering failures from a set* $\mathcal{T}_1 \subseteq \mathcal{T}$ *of terminal links, there exists a set of* n *or fewer non-active codes that cover failures in set* $\mathcal{T}_1$.

Corollary 4 *For receiver-based recovery, the minimum number of codes for terminal link failures can be achieved with non-active codes.*

B Need for network management

We call a link h *integral* if it satisfies the property that there exists some subgraph of the network containing h, on which the set of source-receiver connections is feasible if and only if h has not failed. The following result shows that there is a need for network management for recovery from all recoverable link failures, if there is at least one integral link whose failure is recoverable.

Theorem 1 *Consider any network connection problem with at least one integral link whose failure is recoverable. Then there is no single linear code (A, G, B) that can cover the no-failure scenario and all recoverable failures for this problem.*[1].

Proof outline Consider an integral link h whose failure is recoverable, and a subgraph $\mathcal{G}'$ on which the set of source-receiver connections is feasible if and only if h has not failed. $\mathcal{G}'$ does not include all links, otherwise failure of h would not be recoverable. Then the set of links not in $\mathcal{G}'$, together with h, forms a set $\mathcal{H}$ of two or more links whose individual failures are recoverable but whose combined failures are not. By Lemma 1, a code which covers the no-failure scenario and failure of a link k is non-active in k. However, a code which is non-active in all the links in $\mathcal{H}$ is not valid. $\square$

Corollary 5 *For a network connection problem with one or more nonzero source-receiver connections, there is no single linear code (A, G, B) that can cover the no-failure scenario and all single link failures, assuming they are recoverable.*

C Centralized network management for all link failures

In this section we establish bounds on centralized network management for receiver-based and network-wide recovery from all single link failures, which are presented in Theorem 3.

Our approach is to first consider the number of codes needed for failures among a set $\mathcal{M}$ of links across a minimum capacity cut between the sources and the receiver[2]. We note that $AG_\mathcal{M}$ has rank r for any valid interior code (A, G), since the signal maps on the terminal links are linear combinations of those on links in $\mathcal{M}$.

[1] A solution with static A and F matrices always exists for any recoverable set of failures in a multicast scenario [3], but the receiver code B must change.

[2] a partition of the network nodes into a set containing the sources, and another set containing the receiver, such that the minimum number of links cross from one set to the other

In the development leading up to Theorem 3, we establish Lemmas 5 to 8, making use of the following definitions. We first define transfer matrices for the set $\mathcal{J}$ of links upstream of and including links in $\mathcal{M}$. (Q, J) is a partial interior code defining the behavior of links in $\mathcal{J}$, where

1. $r \times |\mathcal{J}|$ matrix Q specifies how the source processes X_i, $i = 1, \ldots, r$ are represented on the source nodes' incident outgoing links[3]. The signal on source link j is

$$Y(j) = \sum_{i \in \mathcal{A}_{\mathrm{tail}(j)}} Q(i,j) X_i$$

where X_i, $i \in \mathcal{A}_{\mathrm{tail}(j)}$ are processes generated at tail(j) and transmitted on j.

2. $|\mathcal{J}| \times |\mathcal{J}|$ matrix D specifies how signals are transmitted between incident links in $\mathcal{J}$. $D(i,j)$ is nonzero only if head(i) = tail(j). The signal on link j is

$$Y(j) = \sum_{i\,:\,\mathrm{head}(i)=\mathrm{tail}(j)} D(i,j) Y(i)$$

3. $|\mathcal{J}| \times |\mathcal{J}|$ matrix $J = I + D + D^2 + \ldots$ sums the gains along all paths between each pair of links, and equals $(I - D)^{-1}$ since matrix D is nilpotent.

We also define the following:

$J_{\mathcal{M}}$ is the submatrix of J consisting of columns that correspond to links in $\mathcal{M}$.

$A_{\mathcal{J}}$ is the submatrix of A consisting of columns that correspond to links in $\mathcal{J}$.

$F_{\mathcal{J}\times\mathcal{J}}$ is the $|\mathcal{J}| \times |\mathcal{J}|$ submatrix of F consisting of rows and columns corresponding to links in $\mathcal{J}$.

$G_{\mathcal{J}\times\mathcal{J}}$ is the $|\mathcal{J}| \times |\mathcal{J}|$ submatrix of G consisting of rows and columns corresponding to links in $\mathcal{J}$.

Π' is a related network connection problem in which all nodes upstream of $\mathcal{M}$, and the links between them, are the same as in the original problem, but each link h in $\mathcal{M}$ is replaced by a link h' such that tail(h') = tail(h), and head$(h') = \beta'$, a new receiver node that is the head of all links h'.

Note that Q, D and J are defined analogously to A, F and G respectively, except that they are limited to specifying the behavior of links in $\mathcal{J}$. $A_{\mathcal{J}}$ is the value of Q that specifies the same behavior for links in $\mathcal{J}$ as does A. Similarly, $F_{\mathcal{J}\times\mathcal{J}}$ is the value for D that corresponds to F, and $G_{\mathcal{J}\times\mathcal{J}}$ is the value for J that corresponds to G. Note that $G_{\mathcal{J}\times\mathcal{J}} = (I - F_{\mathcal{J}\times\mathcal{J}})^{-1}$ if $G = (I - F)^{-1}$.

We also use the following theorem, from [6], which gives bounds on $n_{\mathcal{T}}$ in terms of the number of transmitted processes r and the number of terminal links $|\mathcal{T}|$ of the receiver:

Theorem 2 *The following are tight bounds on the number of receiver-based codes needed for the no-failure scenario and all terminal link failures, assuming they are recoverable (i.e. for any values of r and d, there are examples for which the bounds are met with equality):*

[3]all of which are in $\mathcal{J}$, since $\mathcal{M}$ is a cut between the source nodes and the receiver

lower bound	*upper bound*
$\left\lceil \frac{\lvert\mathcal{T}\rvert}{\lvert\mathcal{T}\rvert - r} \right\rceil$	$\max\left(\left\lceil \frac{\lvert\mathcal{T}\rvert}{\lvert\mathcal{T}\rvert - r} \right\rceil, r\right) = \begin{cases} r+1 & \textit{for} \quad r = 1 \textit{ or } \lvert\mathcal{T}\rvert - 1 \\ r & \textit{for} \quad 2 \le r \le \lvert\mathcal{T}\rvert - 2 \end{cases}$

Proof outline

Lower bound: Corollary 4 allows us to limit consideration to non-active codes, which can each cover at most $|\mathcal{T}| - r$ terminal link failures (Lemma 1).

Upper bound: $\left\lceil \frac{|\mathcal{T}|}{|\mathcal{T}| - r} \right\rceil > r$ only for $r = 1$ or $r = |\mathcal{T}| - 1$, for which cases it is easy to verify that $r + 1$ codes are needed. For $2 \le r \le |\mathcal{T}| - 2$, consider any valid static code (A, G). Let $\underline{v}_1, \ldots, \underline{v}_r$ be r columns of $AG_{\mathcal{T}}$ that form a basis, and $\underline{w}_1, \ldots, \underline{w}_{|\mathcal{T}|-r}$ the remaining columns. Assuming single link failures are recoverable, we can find two pairs $(\underline{v}_i, \underline{w}_{i'})$ and $(\underline{v}_j, \underline{w}_{j'})$ such that $\underline{w}_{i'}$ can replace $\underline{v}_i$ in the basis, and $\underline{w}_{j'}$ can replace $\underline{v}_j$ in the basis. Then the links corresponding to $\underline{v}_i$ and $\underline{w}_{j'}$ can be covered by one code, the links corresponding to $\underline{v}_j$, $\underline{w}_{i'}$ and $\{\underline{w}_k \mid k = 1, \ldots, |\mathcal{T}| - r,\ k \ne i', j'\}$ by another code, and the links corresponding to $\{\underline{v}_k \mid k = 1, \ldots, r,\ k \ne i, j\}$ by a separate code each. □

The following lemma allows us to relate the number of codes needed for terminal link failures in problem Π', bounds on which are given in Theorem 2, to the number of codes needed for failures of links in $\mathcal{M}$.

Lemma 4 *If failure of some link in $\mathcal{J}$ is recoverable, recovery can be achieved with a code in which no link in $\mathcal{M}$ feeds into another.*

Proof outline Having one link in $\mathcal{M}$ feed into another only adds a multiple of one column of $AG_{\mathcal{M}}$ to another, which does not increase its rank. □

Lemma 5 below characterizes, for a code which covers some links in $\mathcal{M}$, the behavior of links in $\mathcal{J}$, in terms of matrices Q and J.

Lemma 5 *Let (Q, J) be a partial interior code in which no link in $\mathcal{M}$ feeds into another. If there exists an $r \times |\mathcal{M}|$ matrix L such that $QJ^h_{\mathcal{M}} L^T = I$ for $h \in \mathcal{M}_1 \subseteq \mathcal{M}$, then there exists a code (A, G, B) covering failure of links in $\mathcal{M}_1$ such that $A_{\mathcal{J}} = Q$ and $G_{\mathcal{J}\times\mathcal{J}} = J$. Conversely, if (A, G, B) covers links in $\mathcal{M}_1 \subseteq \mathcal{M}$, then there exists some $r \times |\mathcal{M}|$ matrix L such that $Q = A_{\mathcal{J}}$ and $J = G_{\mathcal{J}\times\mathcal{J}}$ satisfy $QJ^h_{\mathcal{M}} L^T = I$ for $h \in \mathcal{M}_1$.*

Proof outline There exists a set of link-disjoint paths $\{P_k \mid k \in \mathcal{M}\}$ where P_k connects link k to the receiver. (Q, J) can be *extended* to a valid interior code (A, G), where $A_{\mathcal{J}} = Q$ and $G_{\mathcal{J}\times\mathcal{J}} = J$, by having each link $k \in \mathcal{M}$ transmit without coding along the path P_k to the receiver. For the converse, we can construct a matrix L which satisfies the required property as follows:

$$L^T = \begin{bmatrix} \frac{\sum_{j\in\mathcal{T}} G(e_1, j)\, \underline{b}_j^T}{} \\ \vdots \\ \frac{\sum_{j\in\mathcal{T}} G(e_{|\mathcal{M}|}, j)\, \underline{b}_j^T}{} \end{bmatrix}$$

where $e_1, \ldots, e_{|\mathcal{M}|}$ are the links of $\mathcal{M}$ in the order they appear in $J_{\mathcal{M}}$. □

Lemma 6 *Tight bounds on the number of receiver-based codes needed to cover the no-failure scenario and failures of links in $\mathcal{M}$, assuming they are recoverable, are given in the following table. These bounds are the same in the case where only non-active codes are used.*

lower bound	upper bound
$\left\lceil \frac{\lvert\mathcal{M}\rvert}{\lvert\mathcal{M}\rvert-r} \right\rceil$	$\max\left(\left\lceil \frac{\lvert\mathcal{M}\rvert}{\lvert\mathcal{M}\rvert-r} \right\rceil, r\right) = \begin{cases} r+1 & \text{for} \quad r=1 \text{ or } \lvert\mathcal{M}\rvert-1 \\ r & \text{for} \quad 2 \le r \le \lvert\mathcal{M}\rvert-2 \end{cases}$

Proof outline It follows from Lemma 4 that if failure of some link in $\mathcal{J}$ is recoverable, it is recoverable for the related problem Π'. Any code (Q', J') covering failure of terminal links $h \in \mathcal{M}_1$ in problem Π' can be extended to obtain a code (A, G, B) covering links $h \in \mathcal{M}_1$ in the original problem (Lemma 5). The upper bound from Theorem 2 thus applies, with $|\mathcal{M}|$ in place of $|\mathcal{T}|$.

For the lower bound, from Lemma 1, a single code in a valid receiver-based scheme can cover at most $|\mathcal{M}| - r$ of the links in $\mathcal{M}$.

By Corollary 4, restricting consideration to non-active codes does not increase the receiver-based lower bound for the related terminal link problem Π', which is also $\left\lceil \frac{|\mathcal{M}|}{|\mathcal{M}|-r} \right\rceil$, and so does not increase the receiver-based lower bound here. $\square$

Lemma 7 *A tight lower bound on the number of network-wide codes needed to cover the no-failure scenario and failures of links in $\mathcal{M}$, assuming they are recoverable, is given by* $\left\lceil \frac{|\mathcal{M}|+1}{|\mathcal{M}|-r+1} \right\rceil$.

Proof outline It follows from Lemma 1 that a single non-active code covers the no-failure scenario and at most $|\mathcal{M}| - r$ single link failures among links in $\mathcal{M}$, while a single active code covers at most $|\mathcal{M}| - r + 1$ links in $\mathcal{M}$. Each code therefore covers at most $|\mathcal{M}| - r + 1$ out of $|\mathcal{M}| + 1$ scenarios of no failures and failures of links in $\mathcal{M}$. $\square$

Lemma 8 *For receiver-based recovery with a single receiver and a set of connections recoverable under all single link failures, there exists a valid static interior code (A, G) such that no link feeds into more than one link in $\mathcal{M}$.*

Proof outline From Lemma 4, there exist valid codes for failures of links in $\mathcal{J}$ in problem Π'. Thus, a static interior code (Q', J') covering these failures exists for Π' [3]. This can be extended (Lemma 5) to a static interior code (A, G) in which no link in $\mathcal{M}$ feeds into another. For any such code (A, G), consider any link h which feeds into more than one link in $\mathcal{M}$. Let the set of these links be $\mathcal{M}^h = \{h_1, \ldots, h_x\}$, and let the set of remaining links in $\mathcal{M}$ be $\overline{\mathcal{M}^h}$.

Case 1: h feeds into some link h_i in $\mathcal{M}$ via some path P without further coding with other signals. We can construct a partial code (Q, J) in which h feeds only into $h_i \in \mathcal{M}^h$, whose extension is a valid static code.

Case 2: Coding occurs between h and each $h_i \in \mathcal{M}^h$. We show that there exists a proper subset $\mathcal{L} \subset \mathcal{M}$ such that $AG_{\mathcal{L}}^h$ has full rank and which does not include all links in $\mathcal{M}^h$. Let h_j be some link in $\mathcal{M}^h \cap \mathcal{M}/\mathcal{L}$.

Case 2a: There exists a set R of links forming a single path from h to h_j, excluding h and h_j, such that none of the links $h' \in R$ feeding into some other link h_i, $i = 1, \ldots, x$, $i \neq j$ has a signal map other than $G(h, h')\underline{c}_h$. We can then construct a partial code (Q', D') which is the same as $(A_{\mathcal{J}}, F_{\mathcal{J}\times\mathcal{J}})$ except that h feeds only into links in R, whose extension is a valid static code.

Case 2b: Every path from h to h_j contains some link that feeds into one or more links $h_i \in \mathcal{M}^h$ besides h_j, and has a signal map which is a linear combination of $\underline{c}_h$ and some other signal map. Consider any path R' from h to h_j and let $\tilde{h}$ be the furthest upstream of these links.

We apply the entire argument described from paragraph 1 onwards with (A, G) and $\tilde{h}$. If case 1 or case 2a applies, then we have a modified code (A', G') in which $\tilde{h}$ feeds into only one link in $\mathcal{M}$. We then apply the same argument once again, this time to (A', G') and h, with h feeding into strictly fewer links in $\mathcal{M}$ than before. If on the other hand case 2b applies, we proceed recursively, with $\tilde{h}$ replaced by one of its downstream links. If we come to a link that is incident to a link in $\mathcal{M}$, then case 1 or case 2a will apply, allowing us to eliminate a nonzero number of links in $\mathcal{M}$ from consideration. Thus, the procedure terminates with a valid static interior code in which h feeds into only one link in $\mathcal{M}$. $\square$

Theorem 3 *For a network with r source processes, a single receiver and a minimum cut capacity of $|\mathcal{M}|$, tight bounds on the number of codes needed for the no-failure scenario and all single link failures, assuming they are recoverable, are given below (i.e. for any values of r and $|\mathcal{M}|$, there are examples for which the bounds are met with equality).*

recovery scheme	lower bound	upper bound
receiver-based	$\left\lceil \frac{\lvert\mathcal{M}\rvert}{\lvert\mathcal{M}\rvert-r} \right\rceil$	$\max\left(\left\lceil \frac{\lvert\mathcal{M}\rvert}{\lvert\mathcal{M}\rvert-r} \right\rceil, r\right) = \begin{cases} r+1 & \text{for } r=1 \text{ or } \lvert\mathcal{M}\rvert-1 \\ r & \text{for } 2 \le r \le \lvert\mathcal{M}\rvert-2 \end{cases}$
network-wide	$\left\lceil \frac{\lvert\mathcal{M}\rvert+1}{\lvert\mathcal{M}\rvert-r+1} \right\rceil$	$\max(2, r) = \begin{cases} 2 & \text{for } r=1 \\ r & \text{for } 2 \le r \le \lvert\mathcal{M}\rvert-1 \end{cases}$

Proof outline We can find a valid static interior code such that the subgraphs S_k of links which feed into each $k \in \mathcal{M}$ are link disjoint with each other, and the paths P_k along which k transmits to the receiver are also link disjoint (Lemmas 5 and 8). Any non-active code which covers failure of link k also covers failure of all links in S_k and P_k. Thus the receiver-based bounds here are the same as those in Lemma 6.

The network-wide lower bound is the same as that in Lemma 7.

For $1 \le r \le |\mathcal{M}| - 2$, the receiver-based upper bound of $\max(2, r)$ is also a tight upper bound for network-wide recovery, which includes the former as a special case. For $r = |\mathcal{M}| - 1$, we start with a static interior code (A, G) in which no link feeds into more than one link in $\mathcal{M}$, and use this to construct an active code which covers two of the links in $\mathcal{M}$ with their upstream links. The remaining $r-1$ links in $\mathcal{M}$, and their upstream links, can be covered by their corresponding receiver-based codes. $\square$

D Node-based management requirement

In this section we consider an alternative measure of network management, the *node-based management requirement*, which is defined as the sum over all nodes of the log of the number of different behaviors for each node. This formulation imputes higher management overhead to recovery schemes which require multiple nodes to change their behavior.

We show in Theorem 4 that in the single-receiver case, receiver-based codes are sufficient to achieve the minimum node-based management requirement for terminal link failures and the no-failure scenario. This is in contrast to the centralized formulation of the preceding sections, in which network-wide schemes may in some cases require fewer codes than receiver-based schemes for terminal link failures [6][4]. Thus, although network-wide schemes may require fewer overall codes than receiver-based schemes for terminal link failure recovery, more nodes are involved, which offsets the effect of having fewer codes where node-based management requirement is concerned. We first establish a lemma needed in the proof of Theorem 4:

Lemma 9 *If the no failure scenario and all single terminal link failures are covered by a set of n codes $\{(A_1,G_1,B), (A_2,G_2,B),\ldots,(A_n,G_n,B)\}$ having a common B matrix, then they can be covered by a set of n codes $\{(A,G,B_1), (A,G,B_2),\ldots,(A,G,B_n)\}$ with a common AG matrix.*

Proof outline Since an active code cannot cover the no-failure scenario (Lemma 1), there is at least one non-active code. If codes $\{(A_1,G_1,B),(A_2,G_2,B),\ldots,(A_n,G_n,B)\}$ are all non-active, there is a set of n codes with common (A,G) that cover the same terminal link failures (Lemma 2). Otherwise, there is at least one active code among them. We can show that one of links covered by each active code can be covered together with some non-active code, and the rest by a separate non-active code. $\square$

Theorem 4 *For the single-receiver case, the minimum node-based management requirement for terminal link failures and the no-failure scenario is achieved with receiver-based schemes.*

Proof outline If interior nodes $i=1,\ldots,x$ each switch among m_i codes and the receiver switches among n codes, the node-based management requirement is $\sum_{i=1}^{x}\log_2 m_i+\log_2 n=\log_2(\Pi_{i=1}^{x}m_i)\,n\geq\log_2 mn$, where m is the number of different values for AG among all the codes. $m\geq\Pi_{i=1}^{x}m_i$ because between two distinct values of AG, there is at least one interior node which switches code.

Let a set of codes covering the no-failure scenario and all terminal link failures be called *complete*. We show that for any complete set of network-wide codes with m values for AG and n values for B, there exists a complete set of $\leq mn$ receiver-based codes. Then the receiver-based management requirement is $\leq\log_2 mn$, which is less than or equal to the network-wide requirement.

Case 1: $m=1$. There exists a complete set of $n=mn$ codes with a static AG matrix, which are receiver-based codes.

Case 2: $n=1$. There exists a complete set of m codes with a static B matrix. By Lemma 9, there exists a complete set of $m=mn$ receiver-based codes with a static B matrix.

Case 3: $m\geq 2, n\geq 2$. If any set of $n_1\geq 2$ codes $\{(A,G,B_1), (A,G,B_2),\ldots,(A,G,B_{n_1})\}$ has a common AG matrix, there is a corresponding set of $\leq n_1$ non-active codes covering the same terminal links (Lemma 3). Each of the remaining codes can be covered by one or two non-active codes (Corollary 3). The maximum number of non-active codes resulting from this procedure is mn. Thus, there exists a complete set of $\leq mn$ receiver-based codes (Lemma 2). $\square$

[4]unless attention is restricted to non-active codes (Corollary 2)

IV. Conclusions and further work

We have analyzed two formulations for quantifying network management for link failure recovery. The first is a centralized formulation where network behavior is described by an overall code determining the behavior of every node, and the management requirement is taken as the log of the number of codes the network may switch among. For this formulation, we have given tight upper and lower bounds on network management for receiver-based and network-wide recovery from all single link failures, assuming they are recoverable. In this paper we consider the single-receiver case, but the approach appears to be amenable to extension to the multi-receiver case.

The second is a node-based formulation where the management requirement is taken as the sum over all nodes of the log of the number of different behaviors for each node. We have shown that minimum node-based management requirement for terminal link failures and the no-failure scenario is achieved with receiver-based schemes in the single-receiver case. This does not however hold for the multi-receiver case. An example of a multi-receiver scenario for which this does not hold is shown in Figure 3. Here, the source multicasts one process to two receivers. Receiver-based recovery requires each of the two receivers to switch between two codes, whereas network-wide recovery allows for recovery with only the source node switching between two codes.

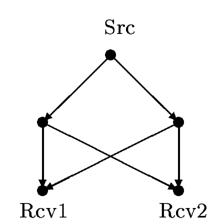

Figure 3: Counter-example showing that Theorem 4 does not hold for the multi-receiver case.

Further work includes studying the capacity requirements for decentralized network management, where management information is transmitted from a failed link to nodes requiring the information for code selection. For a given network and set of connections, there may be a number of possible codes for each recoverable scenario. As different choices of coding scheme have different network management requirements, the coding scheme may be chosen to meet constraints on, or some objective criterion for, the management network. If, for instance, we want to find the minimum capacity management network, we will look for localized recovery schemes involving nodes close to the failed link. Alternatively, given a management network, we would look for coding schemes which the management network can support.

References

[1] R. Ahlswede, N. Cai, S.-Y.R. Li and R.W. Yeung, "Network Information Flow", IEEE-IT, vol. 46, pp. 1204-1216, 2000.

[2] S.-Y.R. Li and R.W. Yeung, "Linear Network Coding", preprint, 1999.

[3] R. Koetter and M. Médard, "An Algebraic Approach to Network Coding", Proceedings of the 2001 IEEE International Symposium on Information Theory, pp. 104, 2001.

[4] R. G. Gallager, "Basic limits on protocol information in data communication networks", IEEE Transactions on Information Theory, Vol. 22, pp. 385-399, 1976.

[5] A. Ephremides and B. Hajek, " Information Theory and Communication Networks: An Unconsummated Union", IEEE Transactions on Information Theory, vol. 44, 1998.

[6] T. Ho, M. Médard and R. Koetter, "A coding view of network capacity, recovery and management", Submitted to the 2002 IEEE International Symposium on Information Theory, 2001.