The Dissertation Committee for Brian Matthew Smith
certifies that this is the approved version of the following dissertation:

Capacities of Erasure Networks

Committee:

Sriram Vishwanath, Supervisor

Ari Arapostathis

Piyush Gupta

Robert Heath

Lili Qiu

Sanjay Shakkottai

Capacities of Erasure Networks

by

Brian Matthew Smith, S.B., M.Eng.

Dissertation

Presented to the Faculty of the Graduate School of
the University of Texas at Austin
in Partital Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

The University of Texas at Austin
May 2008

Capacities of Erasure Networks

Publication No. _____

Brian Matthew Smith, Ph.D.

The University of Texas at Austin, 2008

Supervisor: Sriram Vishwanath

We have investigated, in various multiple senses, the "capacity" of several models of erasure networks. The defining characteristic of a memoryless erasure network is that each channel between any two nodes is an independent erasure channel. The models that we explore differ in the absence or presence of interference at either the transmitters, the receivers, or both; and in the availability of feedback at the transmitters. The crux of this work involves the investigation and analysis of several different performance measures for these networks: traditional information capacity (including multicast capacity and feeback capacity), secrecy capacity, and transport capacity.

Contents

List of Figures

Chapter 1

Introduction

Network information theory is a subset of information theory which deals with the transfer of information over a networked combination of channels, rather than across a single channel. There may be either a single, or more than one, source or destination, which require the transfer of the same, or perhaps unique, pieces of data. Large wireless systems, those with many users which often lack the ability for coordinated, universal control, are particularly valuable to study, both for practical and theoretical reasons. These kinds of systems are becoming ubiquitous: cellular networks continually must support increasing numbers of customers; more people expect higher-speed broadband access over their laptops, personal organizers, and phones; military applications which keep track of the course of battle can help save lives.

While both theoretical limits and many practical methods are known for single-user point-to-point communication models, basic understanding of these complex multiple-user systems is still lacking. There are several very natural, easy to formulate, of practical use, and still open problems in the field of multi-user information theory. For example, the precise capacity regions of the broadcast channel (where one transmitter wishes to communicate two independent data sources to two different receivers) and the interference channel (where two transmitters send independent data to two different receivers) are, in general, unknown. In these cases, it is important to do as much as possible to characterize the behavior of these multi-user channels. We can do so in several ways - by providing upper and lower bounds on the general case, by choosing simple but representative models that are more amenable to analytic description, and by studying the asymptotic behavior of such systems. My research focuses primarily on the last two methods.

We choose the erasure channel as a simple, yet wide-reaching, model of wireless packet communication, and then study both traditional information capacity and asymptotic transport and throughput capacities.

If we take a directed graph as the basis of our network model, then the distinguishing feature of an erasure network is that the edges each represent erasure channels: Either the symbol transmitted across an edge is received correctly at the endpoint, or a distinct "Error" symbol is received. The model results from viewing a system from the *network layer*, rather than the physical layer. We assume that some mode of error detection coding has already been performed, so that when a packet is received, we can be assured either that its contents are correct, or that we have no information about what the contents might actually be. Pioneering work on networks of erasure channels was done in [1] and [2]. We expand this work by examining more general models of erasure networks, involving different types of transmitter and receiver interference and feedback availability. My primary research aim is to understand the fundamental mathematical limits which govern communication in different models of erasure networks, in determining the information theoretic capacity when possible, and in other cases determining the properties that exact solutions to problems involving both systems that increase in size and density must have.

1.1 Notions of Capacity: Information Capacity, Transport Capacity, Secrecy Capacity, and Feedback

The information capacity of a single channel is the maximum rate at which reliable communication is possible; it allows us to answer the question "In n uses, how many unique inputs can the output distinguish amongst with high probability?" (The number of distinguishable signals grows exponentially in n, with the capacity as the exponent. [3]) Information capacity is a theoretical limit on reliable communication; it ignores practical consideration by allowing arbitrarily long delays and unlimited computational abilities as the transmitter and receivers.

Nonetheless, it is important to understand and be aware of the fundamental limits of communication, in an attempt to approach them with real-world algorithms and approximations.

When dealing with more than a single source and a single destination, the information capacity of a system becomes a capacity region, rather than the maximum of a scalar rate. The convex region is a set of all rate vectors $(R_1, R_2, ...)$ where the R_i are rates corresponding to the different source-destination pairs. Even the simple act of describing such a region, let along computing it, becomes exponentially more difficult as the number of possible sources and destinations increases.

One example of a multiple-source multiple-destination network problem is "multicast," which refers to the case when a number of different destinations all require the information generated by a single source. Alternatively, "multiple unicast" is the situation when several source-destination pairs all communication unique data.

Even with a single source-destination pair, networks composed of more than two nodes and more than a single channel also become problematic: these networks are generally called relay networks (since the intermediate nodes have no data of their own to transmit, and have the purpose of aiding the source in relaying its data to the destination), and the general capacity of such networks is unknown. There are several achievable strategies for the three-terminal discrete memoryless relay channel [4] dating back twenty-five years or more, but these only correspond to best known upper-bound in a particular subset of cases (the physically degraded and deterministic relay channels, for example).

Because there are only a relatively few cases in which the information capacity region of multiple-source and/or multiple-destination networks is known, different approaches must be considered. In order to still obtain a descriptive measure of the communication capability of networks, other tools are available. Even though information capacity is seen by some as the apex of information theory, it does not always tell the whole story. In addition to the difficulty of actually computing the capacity region, it reflects nothing about the ease of implementation or the

delay requirement of a particular coding scheme. Random coding, the stock tool for demonstrating the achievability of a particular capacity value, is largely not considered for actual use in any real (or even simulated) system - the computation required and delay incurred would be astronomically huge (doubly exponential growth, with respect to desired probability of error, is hardly unknown).

All of this makes looking at other measures of the capabilities of any network a valuable exercise. One popular quantity is the *transport capacity*. Transport capacity is the distance-weighted sum of rates for a network, and provides a convenient scalar description of the amount of traffic that a network can support. The notion was introduced by Gupta and Kumar [5] to study the capabilities of wireless networks with additive Gaussian noise. Later, Xie and Kumar [6] provided a information-theoretic scaling law that shows, under certain high-attenuation conditions, the transport capacity of the additive Gaussian interference network can grow no faster than linearly in the number of nodes in the network. Franceschetti et al. [7] demonstrate that linear growth in the number of nodes is achievable in a random network with an additive Gaussian noise model using routing alone (and no network coding). The work is often done in two different network settings: First, the case of *dense* scaling, where more and more nodes are placed in a fixed area. This is the sense of Gupta and Kumar's original paper. [5] Alternatively, the case of an *expansive* network is studied, where the physical size of the network grows and the density of nodes remains a constant as the total number of nodes increases. These views are unified for the Gaussian interference network in [8], where it is shown that (arbitrarily close to) linear growth can be achieved in a dense network using a cooperative, multi-layered MIMO technique based on the single-layer work of [9].

We use the descriptive tool of transport capacity to investigate the asymptotic capabilities of several different models of erasure networks.

Another interesting question to ask about the capacity of a network is, "What happens in the presence of a malicious eavesdropper?" The user would like to

communicate information as efficiently as possible to a destination, but realizes that there may exist an eavesdropper who has knowledge of all the codebooks in the network, and access to the symbols received by nodes along one or more links in the network. Therefore, the user wants to prevent the eavesdropper from gaining any information about his secret message: Not only should the eavesdropper be unable to determine exactly which message the transmitter is sending, but he should gain no information about that message, at all. That is, we want the entropy of the message, given the data symbols which the eavesdropper sees, to be arbitrarily close to the entropy of the message, given no additional information. The rate at which the user can communication under such a rate is the secrecy capacity, otherwise known as the equivocation rate. We try to answer this question in this work, in the context of wireless erasure networks.

Finally, we know that in point-to-point channels, feedback cannot increase the channel capacity, but still may have valuable benefits: Increasing the probability of error decay exponent, for example, or simplifying the coding scheme. We demonstrate the benefits of feedback - primarily, a much simplified coding and transmission scheme - for wireless erasure networks in this work, as well.

1.2 Erasure Network Models

There is a growing interest in the study of the capacity of erasure networks with constraints that reflect the underlying physical layer [2]. One of the primary techniques used to study such networks is network coding. Network coding was first used to achieve the multicast capacity of deterministic wireline networks [10]. It has since been put to a variety of purposes, including wireless erasure networks with broadcast constraints (but no interference constraints) which were studied in [2]. This work studied systems with independent erasures between nodes using both random coding and random linear encoding techniques.

Erasure networks generate interest for two main reasons: First, they can be a reasonable model for a packetized network which uses error correction and de-

tection coding. In fact, some practical in-use communication protocols, such as ethernet, use a check to decide whether to accept or reject a packet. Secondly, erasure networks are one of the few multiple-terminal networks which offer themselves up to analysis. By comparison, much recent work has been concentrated on network coding models in zero-error, wired networks without any kind of interference whatsoever [10]. Work involving network coding with interference is now beginning to be more common, for example [11–13].

We look at several different models of erasure networks, in order to cover multiple physical phenomenon, and to gain as much understanding as to the underlying fundamental dynamics of networks.

The simplest of these is the non-interference network model introduced in [1]. It represents a network by a directed acyclic graph, where each edge is an independent erasure channel, with a dedicated input and output for each edge. When the erasure probability is a constant ϵ on each edge, then it was shown that the single-source single destination capacity of the network is the traditional min-cut value (the sum of the number of edges crossing the cut) multiplied by $\bar{\epsilon}$.

The broadcast nature of wireless networks is accounted for by a new erasure network model in [2]. In this model, all of the edges that depart any given node are required to carry an identical symbol in any given timeslot, just as a wireless antenna transmits just one signal to all the antennas which may be receiving at that time. This *wireless erasure network* model assumes some time-sharing or other interference mitigation scheme, so that the symbols along all incoming edges to a node are received without interference, as in a vector. It has been demonstrated that, if the final destinations all know the positions of any erasures in the network, then the information theoretic cut-set upper-bound (a modification of the traditional min-cut bound for flow networks) is indeed achievable for multicasting.

This result capacity-achieving result inspires the question, what about a network with no broadcast constraint, but a multiple-access interference constraint instead? Indeed, for any wireless erasure network with only a broadcast constraint,

one can create a "dual" multiple-access constraint network: Swap the source and destination nodes, and reverse the direction of every edge. Eliminate the broadcast constraint, and institute an additive finite-field multiple access channel at every receiver node. The cut-set upper bound for this network is then identical to that of the wireless broadcast erasure network.

Further, we can generalize and create a model which takes into account all of the cases so far described: In this model, each node is allowed multiple outputs, and each output is connected to one or more nodes' receivers by erasure channels. Each node has at least one receiver, which obtains the finite-field sum of all the unerased inputs to that receiver. The cut-set upper bound is easy to derive for this global model, but whether that bound is achievable is still unknown. Further, if we wish to investigate the transport capacity of any of the above networks, it is necessary to have a model which translates the geographic distance between any two nodes and the probability of erasure along the channel between them. We investigate three separate models: a threshold model (where perfect reception is assumed for distances smaller than a certain value, and no reception for greater distances), an exponential decay model, and a polynomial decay model.

1.3 Summary of Results

Our research on erasure networks can be categorized under four main headings: Work in the field of information capacity for networks with receiver interference, work in secrecy capacity, work in networks with feedback, and work in transport capacity.

In Chapter 2, for a single-source single-destination or multicast erasure network, with additive-finite field receiver interference and with or without broadcast requirements at the transmitter, we have shown that the min-cut max-flow capacity is in fact achievable when side information detailing the erasure locations along all the links is available to the final destination.

In Chapter 3, we demonstrate a randomized routing scheme in a unicast wireless

erasure network which allows for a throughput-optimal, capacity achieving network operation. The routing scheme is novel in its simplicity, requiring no link-level feedback and no knowledge of the network topology for success, and only a very simple acknowledgment feedback from the final destination. The proof mechanism is also notable because of the intuitive connection between the model of the system state, and the individual cut-set bounds on throughput which necessarily must follow.

In Chapter 4, we prove upper and lower bounds on the secrecy capacity of wireless erasure networks. We give sufficient conditions for when those bounds meet, and provide counter-examples which demonstrate cases when the intuitive (both upper and lower) bounds are not tight.

Finally, in Chapter 5, we show that with or without receiver interference, the wireless erasure network with a broadcast transmit constraint has a transport capacity which grows no faster than linearly in the number of nodes when a minimum node separation constraint is enforced. This result holds for both one- and two-dimensional networks both when the probability of a successful packet transmission decays exponentially with increasing distance between two nodes and when the probability decays polynomially as long as the decay exponent β is greater than 3. Further, it has been shown that linear growth in the number of nodes is achievable by routing only, without any need for network coding at the intermediate nodes. Routing, therefore, is an order-optimal strategy in wireless erasure networks.

We hope that, taken together, this work provides a thorough survey of several different applications of wireless networks and a variety of the characteristics of these networks.

1.4 General Notation

This section contains some notation which is commonly used throughout the document. Specialized notation required for individual sections will be defined and introduced where appropriate.

In general, we use a directed graph model as the basis for our network topology. A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ has vertex set $\mathcal{V} = \{1, 2, ..., |\mathcal{V}|\}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The number of nodes/vertices $|\mathcal{V}|$ in a network is usually referred to as n, and nodes are most often indexed by the indices i and j.

For a vertex $i \in \mathcal{V}$, let $\mathcal{N}_I(i)$ and $\mathcal{N}_O(i)$ be the sets of edges entering and leaving the vertex v, respectively. That is,

$$\mathcal{N}_I(j) = \{(i, j) : (i, j) \in \mathcal{E}\}$$
$$\mathcal{N}_O(i) = \{(i, j) : (i, j) \in \mathcal{E}\}$$

An s-d cut for $s, d \in \mathcal{V}$ is a partition of $\mathcal{V}$ into two subsets $S \subset \mathcal{V}$ and $S^C \subset \mathcal{V}$ such that $s \in S$ and $d \in S^C$. Let $\mathcal{S}$ be the set of all s-d cuts. Further, for any s-d cut S, define S^* and S^{C*} as

$$S^* = \{i | \exists (i, j) \in \mathcal{E} \text{ s.t. } i \in S, j \in S^C\}$$
$$S^{C*} = \{j | \exists (i, j) \in \mathcal{E} \text{ s.t. } i \in S, j \in S^C\}$$

so that S^* and S^{C*} are the sets of nodes with edges that cross the cut-set partition.

Random variables are represented by uppercase letters, specific instantiations of a random variable by lowercase. For example, channel inputs (i.e. node outputs) are represented by X or x, and channel outputs by Y or y, which may be indexed as X_{ij} or x_{ij} to denote the specific edge to which the input or output refers.

In this work we use a binary symmetric erasure channel model, where a transmitted symbol is either correctly received at the destination, or remains completely unknown. Specifically, the input/output relationship is a conditional probability distribution

$$p(y|x) = \begin{cases} 1 - \epsilon & y = x \\ \epsilon & y = E \end{cases}$$

with input alphabet chosen from some finite field F_q and output alphabet $F_q \bigcup E$, the input alphabet plus a symbol to represent an erasure. The character ϵ thus refers to the erasure probability, and can also be indexed ϵ_{ij} to denote the probability of erasure along a specific edge. If an edge (i, j) does not exist in a network,

we write $\epsilon_{ij} = 1$ to denote the fact that the input is always erased and never transmitted. The symbols γ_{ij} refer to realizations of Bernoulli binary random variables with a probability ϵ_{ij} of being zero.

The uppercase letters H, G, and A usually refer to transfer matrices, which in this work are usually $0-1$ binary random matricies with entries γ_{ij}. (The uppercase H is also used to represent entropy, but the context should be clear.) The notation lg is used to represent the logarithm in base 2.

Chapter 2

Information Capacity: Receiver Interference Models

Here we look at the capacity of erasure networks in the traditional information theory sense. We are interested in both single-source, single-destination capacity and multicast capacity. The distinguishing feature of erasure networks is that each edge in the directed graph describing the network topology represents an independent erasure channel. The innovation of the work in this chapter is the inclusion of interference in the network model, in particular finite-field addition. For simplicity, in this section we will consider networks whose transmit alphabet is limited to the two symbols $\{0, 1\}$, but all results hold for larger alphabets(say, of size q), with the trivial modification of multiplying the rate values by $\lg q$. This work was performed simultaneously with, but independently from, the work in [14], which has very similar things to say about deterministic networks with both broadcast and multiple-access constraints (with an emphasis on the finite-field sum), and very interestingly, a parallel proof mechanism.

2.1 Prior Work

David Julian help pioneer the concept of the erasure network in [1]. This original model is loosely based on the concept of flow networks [15]: each link in the network is an independent erasure channel with an identical erasure probability. Further, each node in the network sees all of its incoming and outgoing edges as distinct: it is a completely interference-free network, as if it were a wireline network.

Julian proves, among other statements concerning the erasure channel, that the capacity of such a network is equal to the capacity of the equivalent flow network,

multiplied by $\bar{\epsilon}$, where ϵ is the uniform erasure probability.

The concept of an erasure network was expanded to provide applicability to wireless networks in [2]. It has been known for several years that the multicast cut-set bound for *wireline* networks (with no interference constraints at either transmitters or receivers, the cut-set bound is simply the sum of the individual capacities of all the links across the cut) was achievable [10]. In the work of [2] a tractable model which accounts for the broadcast nature of the wireless medium was proposed: specifically, a *wireless erasure network* is still represented by a directed graph, but each node is required to transmit an identical symbol down each of its outgoing links. Incorporating the broadcast nature of the wireless medium was a significant step forward. While routing (a simple forwarding of information) is sufficient to achieve the capacity of a unicast wireline network, the more powerful tool of network coding was required to achieve the modified cut-set rate-bound for the wireless erasure network. We continue to use network coding to demonstrate the capacity of alternate classes of erasure networks.

This modification requires a change in the evaluation of the cut-set bound. While in flow networks and in the network of [1], the cut-capacity is simply the sum of the capacities of all edges that cross the cut, the broadcast requirement of [2] introduces a new formula. We take ϵ_{ij} as the erasure probability across the edge connecting node i to node j; S is a partition of the nodes such that source s and destination d satisfy $s \in S$ and $d \in S^C$. Then the communication rate is upper-bounded by

$$R \leq \min_{S:S\in\mathcal{S}} \sum_{i\in S} \left(1 - \prod_{j\in S^C} \epsilon_{ij}\right). \tag{2.1}$$

Intuitively, we can understand the rate bound (2.1) as follows: Given a cut, every node which can possibly transmit a symbol across the cut, contributes to the sum a quantity equal to the probability that the symbol transmitted along *at least one of the outgoing edges* is successfully received across the cut. Intuitively, as long as any of the symbols crossing the cut from that node are not erased, the

transmission is accounted successful.

The main result of [2] is that the cut-set rate of Equation (2.1) is indeed achievable, for multicast as well as single-destination networks, under one additional assumption: The destinations are all aware, as side-information, of the positions of all the erasures on each link of the network. This assumption is reasonable, as packets usually carry headers, and the extra amount of information required does not increase with the packet size.

To demonstrate the achievability of (2.1), the authors of [2] make use of an extremely interesting proof technique. Random coding is employed over B blocks, each of size n, at all nodes in the network. Given the locations of all erasures, the operation of the network is a *deterministic function* of any input. The destinations simply simulate the operation of the network, and an error only occurs if two different input sequences (corresponding to two different messages) produce identical output sequences at the destination.

Because there is no interference between incoming edges at any node, there is no ambiguity as to which time block any symbol corresponds. (Each node waits until all symbols corresponding to a given message block arrive before calculating its output function for that particular message block). It is therefore possible to follow the evolution of a message block as it traverses the network. As noted above, an error occurs when there exists at least one input sequence (corresponding to a message other than the message actually sent) which produces exactly the same output at the destination node as the actual message. The critical insight of the proof is in representing this error event (call it E) as the union of error events E_s. Let w_0 be the actual message transmitted, and w_1 be an alternative message. If S is a cut, then E_s is the event that, for all nodes in S, the inputs w_0 and w_1 produce distinct outputs, but for all nodes in S^C, the received symbols for that block are identical. Note that the E_s (corresponding to the different cut-sets) are disjoint, and their union is the event E that the destination has an identical block of symbols for both w_o and w_1.

Using the union bound, the probability of E is less than or equal to the sum of the probabilities of E_s, each of which decays exponentially in n, the size of a block. It is therefore the event E_s with the minimum exponent (corresponding to the min-cut) which dominates the sum and governs the achievable rate.

The key to this proof is the fact the we can follow the progression of a single message through the network - that there is no "mixing" between symbols corresponding to different blocks. As soon as receiver interference is introduced to the network, however, this simplification may no longer be available. For example, in the simple relay network of Figure 2.1, the symbols transmitted by the relay in the second time block (corresponding to the message of the first time block) become entwined with the symbols which are simultaneously being transmitted by the source (corresponding to the message of the second time block).

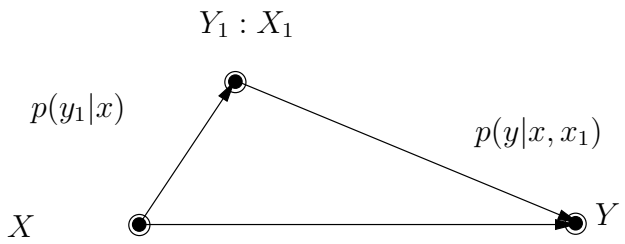

Figure 2.1: Relay Network with Interference

Other prior work has focused on the capacity of other various network models with broadcast constraints [11], multiple-access constraints [12], or both [13].

In this chapter, we prove that nonetheless, a max-flow min-cut bound, which incorporates the interference properties of the model, is achievable using random coding arguments. We look at two different cases; first, a network which is in many sense a dual of the broadcast wireless erasure network of [2], where a multiple-access finite-field sum is used to model receiver interference. We present two distinct proofs of the min-cut bound achievability for this multiple-access constraint network: a simple proof using the technique of pipelining, which was suggested to us by Gerhard Kramer; and our original proof of [16] which is a appropriate lead-in

to the more general case. The second case that we consider is a more general, in a sense "'multiple-antenna,"' model which subsumes both broadcast-constraint, only, and receiver-constraint, only, networks. The proof for this case will follow from our original MAC constraint case, and utilizes some of the results of [14].

2.2 Additive Finite-Field MAC Constraints on Erasure Networks

The two network models that we consider are described in detail in this section. Although the first model, that of the multiple-access finite-field sum erasure networks, are technically simply a special case of the second, the erasure network with arbitrary combinations of broadcast and finite-field sum interference, we present them separately because of the different proof techniques which are applicable and valid for each.

2.2.1 Erasure Networks with Only an Additive Finite-Field MAC Constraint

For this model, we consider a network very similar to that of [2], but with a dual multiple-access channel constraint, instead of the broadcast channel constraint. That is, each node now can send a distinct symbol down all outgoing edges. Instead of a vector of incoming symbols, however, each node receives the finite-field sum of the unerased symbols along all incoming edges (erasures are treated as zeros in the sum). Specifically,

- Each edge (i, j) in the network acts as an independent erasure channel with a specified erasure probability ϵ_{ij}.
- Each node i may transmit unique symbols across each outgoing edge (i, j).
- Each node j receives the finite-field sum of all the non-erased symbols along incoming edges.

- The final destination node has access to side information concerning the erasure locations within the entire network.

In addition, we show that the cut-set upperbound for a given MAC erasure network is identical to the upperbound obtained for the wireless erasure network in which

- The source and destination nodes of the MAC network are interchanged, and
- the direction of every edge in the network is reversed, but
- the erasure probabilities along each edge remain unchanged.

Further, we can show that this cut-set bound is achievable, and therefore the capacity of this network model.

2.3 System Model

We consider a single-source (denoted by s), single-destination (denoted by d) network, modeled by an acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each node i has $|\mathcal{N}_O(i)|$ outputs and exactly one input (i.e. the output of a finite-field additive MAC) and at time t transmits the $|\mathcal{N}_O(i)|$-length vector of symbols $X_{ij}(t)$, where each symbol $X_{ij}(t)$, $(i,j) \in \mathcal{N}_O(i)$, is chosen from the alphabet $\{0,1\}$. The symbol $X_{ij}(t)$ can depend on inputs to the node i from times 1 to $t-1$ and, if i is the source node, the current message.

At each time t, the node j will receive the single symbol $Y_j(t)$, where

$$Y_j(t) = \sum_{(i,j)\in\mathcal{N}_I(j)} \gamma_{ij}(t) X_{ij}(t) \tag{2.2}$$

and the $\gamma_{ij}(t)$ are independent (over both time and edge indices) Bernoulli random variables which take the value 0 with probability ϵ_{ij}. We assign $\epsilon_{ij} = 1$ if the edge (i,j) does not exist in $\mathcal{E}$.

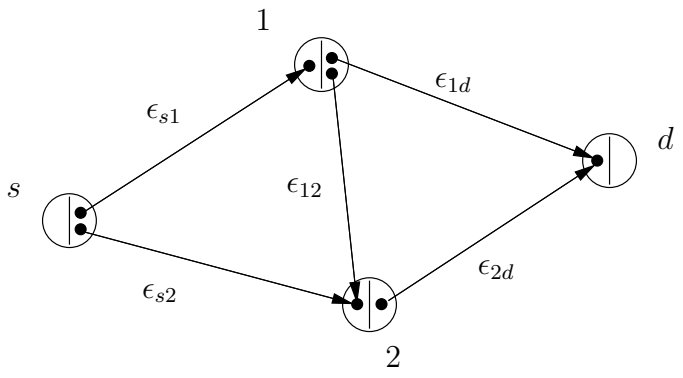

Figure 2.2: Erasure Network With MAC Constraint

In summary, each node has a single input and multiple outputs, which are connected to other nodes' inputs via erasure channels. At each timestep, every node receives an input $Y_j(t)$, the value of which is the finite-field sum of all the non-erased symbols transmitted by exactly the edges which are connected that node. This is illustrated in Figure 2.2. The probability that a symbol transmitted from an output of node i is successfully received (and therefore added to the sum) of the input of node j is then $1-\epsilon_{ij}$. Equation (2.3) represents the relationships between the input and outputs of each node in Figure 2.2.

$$\begin{aligned} y_1 =& \gamma_{s1} x_{s1} \\ y_2 =& \gamma_{s2} x_{s2} + \gamma_{12} x_{12} \\ y_d =& \gamma_{1d} x_{1d} + \gamma_{2d} x_{2d} \end{aligned} \tag{2.3}$$

The destination node d is provided with side information which gives it the locations of all erasure events throughout the network. Intermediate nodes are provided no knowledge of the erasure locations.

2.3.1 Erasure Networks with Generalized Broadcast and Finite-Field Sum Multiple-Access Constraints

The two erasure network models described so far each only consider a single type of interference - at the transmitter, only, or alternatively, at the receiver, only. The

erasure network with arbitrary interference is a fairly general network description which encompasses both of the above models, incorporating both broadcast and receiver interference.

In this model, each node is allowed multiple outputs, and each output is connected to one or more nodes' receivers by erasure channels. Each node has at least one receiver, which obtains the finite-field sum of all the unerased inputs to that receiver. Specifically,

- Each node i is allowed to have multiple inputs (receivers) and multiple outputs (transmitters).
- Each edge in the network is a connection between an output i^m of one node and an input j_n of a different node. Each edge acts as an independent erasure channel.
- Each output i^m of a node is constrained to send the same symbol along all outgoing edges, but different outputs of a the same node may transmit different symbols.
- Each input j_n to a node receives the finite-field sum of all the non-erased symbols along incoming edges.

We must begin with a description of notation used to represent such a model.

2.3.2 Notation and Preliminaries

We modify the notation for directed graphs in order to allow for a broad variety of access and broadcast constraints on a network. A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ has vertex set $\mathcal{V} = \{1, 2, ..., |\mathcal{V}|\}$ and edge set $\mathcal{E} \subseteq (\mathcal{V} \times \mathbb{Z}^+) \times (\mathcal{V} \times \mathbb{Z}^+)$. An edge $((i, m), (j, n)) \in \mathcal{E}$ will be abbreviated as (i^m, j_n). Let M_i and N_i denote the number of outputs and inputs a node i has (which is a distinct concept from the total number of edges entering or leaving the vertex - each edge connects one

particular output i^m of a vertex i to one particular input j_n of a different vertex j).

For a node $i \in \mathcal{V}$, let $\mathcal{N}_I(i)$ and $\mathcal{N}_O(j)$ be the sets of edges entering and leaving the vertex v, respectively. That is,

$$\mathcal{N}_I(j) = \{(i^m, j_n) : (i^m, j_n) \in \mathcal{E}\}$$
$$\mathcal{N}_O(i) = \{(i^m, j_n) : (i^m, j_n) \in \mathcal{E}\}$$

We define similar notation for the set of edges entering and leaving any particular input or output of the vertex i, as well: $\mathcal{N}_I(j_n)$ and $\mathcal{N}_O(i^m)$.

An s-d cut for $s, d \in \mathcal{V}$ is defined just as in Section 1.4 a partition of $\mathcal{V}$ into two subsets S and S^C) such that $s \in S$ and $d \in S^C$.

2.3.3 System Model

We consider a single-source (denoted by s), single-destination (denoted by d) network, modeled by by an acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Each node i has M_i outputs and N_i inputs, and at time t transmits the M_i-length vector of symbols $X_i(t)$, where each symbol $X_i^m(t)$, $m \in 1..M_i$, is chosen from the alphabet $\{0, 1\}$.

At each time t, the node j will receive the vector of symbols $Y_j(t)$ (of length N_j), where

$$Y_j^n(t) = \sum_{(i^m, j_n) \in \mathcal{N}_I(j^n)} \gamma_{j,n}^{i,m}(t) X_i^m(t) \tag{2.4}$$

and the $\gamma_{j,n}^{i,m}(t)$ are all independent (over both time and edge indices) Bernoulli random variables which take the value 0 with probability $\epsilon_{j,n}^{i,m}$. We assume that each node knows the state of each of the channels incoming to that node, as well.

In summary, each node has multiple inputs and outputs, which are connected to other nodes' inputs and outputs via erasure channels. At each timestep, every node receives a vector of symbols, where each element of the vector corresponds to one of that node's inputs. The value of each input's symbol is the finite-field sum of all the non-erased symbols transmitted by exactly the outputs which are

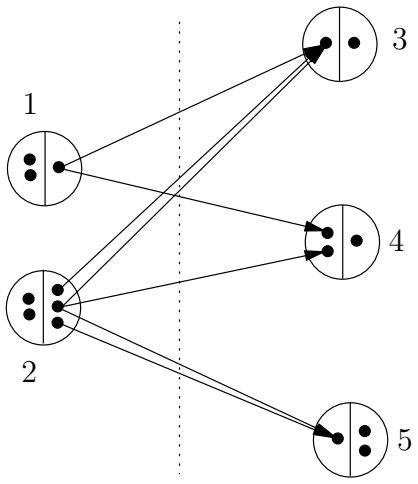

Figure 2.3: Detail of a General Erasure Network With Interference

connected that input. This is illustrated in Figure 2.3. The probability that a symbol transmitted from the m^{th} output of node i is successfully received (and therefore added to the sum) of the n^{th} input of node j is then $1 - \epsilon_{j,n}^{i,m}$.

2.4 Results

For both of our network models, we have been able to demonstrate that the capacity is indeed given by a generalized min-cut max-flow bound, by providing both an upper-bound converse and by providing multiple achievable strategies. This section explicitly gives those cut-set bounds for the networks.

2.4.1 Cut-Set Bound for Generalized Network Model

For any $s-d$ cut, define the matrix A_s to be a random matrix with $0-1$ entries of size

$$\sum_{j \in S^{C*}} N_j \times \sum_{i \in S^*} M_i.$$

Each column represents an output i^m of a node in S^*, (i.e. an input into an erasure channel in the cut-set) and each row represents an input j_n of a node in S^{C*} (i.e. the output of a channel, on the right side of the cut). For every edge (i^m, j_n) in

the cut-set, there is an entry in (the appropriate row and column of) A_s which takes the value 0 with probability $\epsilon_{j,n}^{i,m}$ and 1 with the probability $1-\epsilon_{j,n}^{i,m}$. Every entry in A_s which does have a corresponding edge in $\mathcal{G}$ will be zero with probability one. The matrix A_s then acts as a transfer matrix between the outputs of $\mathcal{V}_s$ and the inputs of $\mathcal{V}_d$: if all nodes on each of the two sides of the cut could cooperate perfectly, then we could collect all the outputs of the nodes in S^* into the vector X^* and the part of the inputs of the nodes in S^{C*} that depends only on outputs from the s side of the $s-d$ cut into a vector Y^*. The relationship $Y^*(t) = A_s(t)X^*(t)$ describes the transfer of information across the cut.

Theorem 2.4.1. *The rate of reliable communication between the source s and the destination d in an erasure network, as defined in Section 2.3, is upperbounded as*

$$R \leq \min_{S:S\in\mathcal{S}} E[rank(A_s)]. \tag{2.5}$$

For example, the matrix A_s for the s-d cut $S = \{1, 2\}$ illustrated in Figure 2.3 is demonstrated as

$$\begin{bmatrix} y_3^1 \\ y_4^1 \\ y_4^2 \\ y_5^1 \end{bmatrix} = \begin{bmatrix} \gamma_{3,1}^{1,1} & \gamma_{3,1}^{2,1} & \gamma_{3,1}^{2,2} & 0 \\ \gamma_{4,1}^{1,1} & 0 & 0 & 0 \\ 0 & 0 & \gamma_{4,2}^{2,2} & 0 \\ 0 & 0 & \gamma_{5,1}^{2,2} & \gamma_{5,1}^{2,3} \end{bmatrix} \begin{bmatrix} x_1^1 \\ x_2^1 \\ x_2^2 \\ x_2^3 \end{bmatrix}.$$

2.4.2 Cut-set Bound for the Multiple Access, Only, Constrained Network

The single-source single-destination rate must satisfy the upperbound

$$R \leq \min_{S:S\in\mathcal{S}} \sum_{j\in S^C} \left(1 - \prod_{i\in S} \epsilon_{ij}\right). \tag{2.6}$$

Further, the rate of Equation (2.6) is achievable, when side information on the locations of all erasures is available to the destination node.

Note that Equation (2.6) is merely a simplification of Equation (2.5) for the special case of a network with multiple-access constraints, only. For example, consider the network of Figure 2.4, where the expression

$$\begin{bmatrix} Y_d^* \\ Y_2^* \end{bmatrix} = \begin{bmatrix} 0 & 0 & \gamma_{1d} \\ \gamma_{s2} & \gamma_{12} & 0 \end{bmatrix} \begin{bmatrix} X_{s2} \\ X_{12} \\ X_{1d} \end{bmatrix}$$

defines A_s for the cut S illustrated.

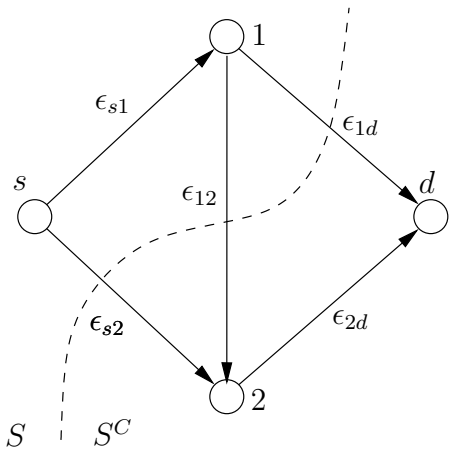

Figure 2.4: Example Network

2.4.3 Cut-Set Bound for Other Specific Network Instances

We have previously remarked that both the wireless erasure network (with broadcast only constraint) of [2] and the no-interference model are both specific instances of our generalized network.

For a wireline network with no interference at either transmitters or receivers, we could similarly define another matrix $A_s^{wireline}$, a random diagonal matrix. Each entry on the diagonal, representing a particular edge that crosses the cut, say, from the i^{th} node to the j^{th} node, will be 1 with probability $1 - \epsilon_{ij}$. The rank of this random diagonal matrix is the sum of all the diagonal entries; the expected value of the rank is then the sum of the probabilities that the diagonal entries are 1.

This sum is precisely the sum of the capacities of each edge crossing the cut, which is the definition of the cut-capacity.

For the wireless erasure network of [2], the transfer matrix A_s will have $|S^*|$ columns (one for each node which has a departing edge across the cut) and one row for each edge in the cut-set. Since only one entry in each row of the matrix can possibly be non-zero, the rank of the transfer matrix is equal to the number of column which contain at least one zero. The probability that a column contains at least one zero is the probability that, of all the edges departing some node in $\mathcal{V}_s$, at least one of these is not erased. The expected value of the rank of the matrix A_S is thus given by the cut-capacity given in [2].

2.5 Upper Bounds

We will use the cut-set bound, as described in [3], to demonstrate an upper bound on the rate of reliable communication from source node s to destination node d in both of our network models.

Recall the definition of the random transfer matrix A_s with 0-1 entries for any $s-d$ cut, from Section 2.4.1. We have the following theorem:

Theorem 2.5.1. *The rate of reliable communication between the source s and the destination d in an erasure network with generalized broadcast and finite-field sum multiple-access constraints, as defined in Section 2.3.3, when the destination d is provided side-information concerning the locations of all erasures in the network, is bounded above by*

$$R \leq \min_{S \in \mathbb{S}} E[rank(A_s)]. \tag{2.7}$$

Further, when A_s is the matrix obtained from the cut S in a multiple access erasure network, the expression's right-hand side evaluates as

$$E[rank(A_s)] = C(S), \tag{2.8}$$

where we define

$$C(S) = \sum_{j \in S^C} \left(1 - \prod_{i \in S} \epsilon_{ij} \right). \tag{2.9}$$

The bound $C(S)$ for the cut $S = \{s, 1\}$ illustrated in Figure 2.4 evaluates as

$$1 - \epsilon_{1d} + 1 - \epsilon_{12}\epsilon_{s2}.$$

Proof. If we assemble all of the outputs of nodes in S of the s-d cut into a vector X_s, the outputs of all nodes in S^* into the vector X^*, the outputs of nodes in S^C into the vector X_d, the inputs of all nodes in S^C into the vector Y_d, and the parts of the inputs of all nodes in S^{C*} that depend only on outputs of nodes in S into the vector Y^*, we can write

$$Y_d = \tilde{A}_s X_s + A_d X_d \tag{2.10}$$

where $\tilde{A}_s$ is a matrix which contains the entries of A_s (at the appropriate input-output positions) and zeros elsewhere, and A_d is a random matrix defined similarly as A_s, but as the transfer matrix between outputs and inputs on the destination side of the s-d cut.

Assuming all the nodes on each side of the cut can mutually cooperate (i.e., all the nodes in S can jointly decide their output values, nodes in S^C can do the same, and nodes in S^C are given each others input values and the erasure locations), an upper bound on rate is given by

$$\begin{aligned} R &\leq \sup_{p(x_{ij} \forall (i,j) \in \mathcal{E})} I(Y_d, A_s; X_s | X_d) \\ &= \sup_{p(x_{ij} \forall (i,j) \text{s.t.} i \in S, j \in S^C)} I(Y^*, A_s; X^*). \end{aligned} \tag{2.11}$$

But,

$$\begin{aligned} I(X^*; Y^*, A_s) &= H(Y^*, A_s) - H(Y^*, A_s | X^*) \\ &= H(A_s) + H(Y^* | A_s) - H(A_s | X^*) - H(Y^* | X^*, A_s) \\ &= H(Y^* | A_s) \end{aligned} \tag{2.12}$$

where the final equality in Equation (2.12) comes the facts that X^* and A_s are independent and Y^* is a deterministic function of X^* and A_s.

Now, for any given transfer matrix $A_s(t)$, the maximum entropy in the vector Y^* is the number of elements in Y^* which are linearly independent of each other (since each element can have maximum entropy 1, but some elements are functions of other elements). Choosing the entries of X^* as i.i.d. Bernoulli(1/2) 0-1 random variables maximizes the entropy of Y^*; this follows because the sum of i.i.d Bernoulli random variables in F_2 is also distributed Bernoulli(1/2), and the entropy $H(Y^*|A_s = A_s(t))$ is then $rank(A_s(t))$. Therefore, the rate over all realizations of A_s is bounded by $E[rank(A_s)]$.

Since for every s-d cut, the rate is upper-bounded by the expected value of the rank of the transfer matrix A_s, the rate is then upper-bounded by the minimum, over all possible s-d cuts, of $E[rank(A_s)]$.

For the multiple-access erasure network, A_s will have one row for each node j in S^{C*}, and there will be $|\mathcal{N}_I(j) \cap S|$ entries (one for each edge connected across the cut-set to node j) which are possibly non-zero. The rank of A_s will be the number of rows which contain at least one 1 entry. Since the probability that the row corresponding to node j has all zeros is

$$\prod_{i:(i,j) \text{ s.t } i \in S^*, j \in S^{C*}} \epsilon_{ij},$$

the expected value of the rank of A_s evaluates to the expression in Equation (2.9).

□

2.6 Achievability in Multiple-Access Erasure Networks

In this section, we show that rates arbitrarily close to the cut-set upper bound given in Section 2.5 are achievable for the multiple-access constraint erasure networks by using random coding arguments. Our proofs will follow the same general outline as that of Theorem 1 of [2]. Therefore, we shall summarize the arguments which are similar and be more precise in the exposition of those steps which are

particular to our proof. The following section will contain the more recent proof for the networks with generalized interference, which builds upon the work required for these proofs.

First, we state the main result:

Theorem 2.6.1. *Consider a single-source single-destination multiple-access erasure network given by the directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The capacity of the network, under the assumptions of Section 2.3.3 (including the assumption of side information on erasure locations known to the destination node), is given by*

$$R \leq \min_{S \in \mathcal{S}} C(S). \tag{2.13}$$

We proceed to demonstrate the achievability of this rate via two different methods.

2.6.1 Achievability by Pipelining

Recall from our description of the achievability argument of [2] in Section 2.1 that the achievability proof depended on the no-receiver interference condition: it was necessary to keep symbols from different blocks independent of each other. Because each node receives a vector of the incoming data symbols, this was straightforward to achieve: before transmitting its codeword relating to any specific block, the node waits until all data for that block (which will in general occur at different times) is received. For example, in a network with the topology of Figure 2.1, the receiver Y will receive from the transmitter X the symbols related to a block b in timeblock b of the network's operation. However, in timeblock b, the receiver Y will simultaneously be receiving symbols related to block $b-1$ from the transmitter X_1. Since there is no interference, it is straightforward to keep these sets of symbols separate, and then decode the data from block b only after timeblock $b+1$ has been completely received.

In the multiple-access constraint network, pipelining allows us to keep a similar separation of data from different blocks. In this case, in Figure 2.1, in block b

the transmitter X will send symbols for the codeword of block b along the link to receiver Y, while simultaneously sending the codeword for block $b+1$ along the edge to receiver Y_1. This allows, in block b, the transmitter X_1 to send the codeword for the message of block b to the receiver Y in timeblock b, so data from timeblock b only interferes with itself - there is no mixing between blocks. Now that there is no mixing, the simulation technique and probability of error analysis of [2] follows explicitly.

Note that the scheduling of such blocks for any directed acyclic graph completely parallels the scheduling that would be necessary for the dual wireless erasure network with broadcast constraint, only (as described in Section 2.8). Thanks goes to Gerhard Kramer, who suggested the pipelining technique. This technique, however, is unable to achieve the capacity of a network with both broadcast and receiver interference, because there is no way to differentiate timeblocks at either the transmitter or receiver. This case is studied in Section 2.7.

2.6.2 Achievability by Conditioning on Previous Blocks

This section contains the proof mechanism which we demonstrated in [16].

Let $\mathcal{W} = \{1, 2, ..., 2^{nRB}\}$ be a set of message indices, and let $w \in \mathcal{W}$ be the message, chosen uniformly and independently, that we desire to communicate between the source and destination node. The network will operate over $(B+L)n$ timeslots in blocks of size n, where L is the length of the longest path between source and destination. As B becomes large the achieved rate of $\frac{B}{B+L}R$ approaches R.

2.6.3 Codebook Generation

Each node $i \in \mathcal{V}$ generates $(B+L) \times |\mathcal{N}_O(i)|$ codebooks which are encoding functions over all the possible inputs to that node. Specifically, the source node generates the $(B+L) \times |\mathcal{N}_O(s)|$ functions

$$f^b_{s^m} : \mathcal{W} \to \{0,1\}^n$$

and each node i other than the destination generates the $(B+L)\times|\mathcal{N}_O(i)|$ different functions

$$f_{i^m}^b : \{0,1\}^n \rightarrow \{0,1\}^n.$$

That is, for each output labeled i^m, $m \in [1, ..., |\mathcal{N}_O(i)|]$, of the node i, the codebook will contain a different function for each combination of the n inputs from the previous length-n block.

All codewords are drawn i.i.d. from a binary Bernoulli distribution with parameter $1/2$. All nodes share their codebooks with the destination node d.

2.6.4 Encoding

After each block of n time slots, each node chooses the appropriate set of codewords, which are based on the inputs it received in that block and the block number, and transmits those codewords from its outputs in the next block.

2.6.5 Decoding

Decoding follows a similar, but not identical, procedure to that of [2]. Because of the lack of interference in the wireless erasure network model, there is no ambiguity about the information encoded in a block – the codeword in each time-block was a function of exactly one message. In the multiple-access erasure network, however, the input received at any particular node in each block is possibly a function of the codewords sent by the source node over multiple different blocks.

Therefore, the destination node, with knowledge of all the erasures throughout the network, decodes by simulating all 2^{nRB} messages' transit of the network. At least one of these codewords (the true codeword) will yield a result at the destination node identical to the sequence that was actually received (since the behavior of the network is a deterministic function of the message and the erasure pattern); an error occurs either if there are one or more additional codewords that result in an identical output, or if the pattern of erasures throughout the network in any time-block is not strongly typical.

2.6.6 Probability of Error Calculation

Intuitively, an error event occurs either if

- The pattern of erasures in any time-block is atypical, or
- For every time-block, the simulation of the codewords for the two messages w and w_1 produces identical inputs to the destination node.

We write this error event E as

$$E = E_{typ} \bigcup \left(\bigcup_{k=1}^{2^{nRB}-1} E_k \right) \tag{2.14}$$

where E_{typ} is the event that the pattern of erasures in not strongly typical and E_k is the event that the messages w and w_k produce identical inputs to the destination node d for all time-blocks.

Formally, define E^b to the event that pattern of erasures in time-block b is not strongly typical, i.e.

$$E^b = \{\gamma_{ij}(t) | (i,j) \in \mathcal{E}, t \in [n(b-1), ..., nb]\} \notin \mathcal{A}_\delta^{(n)}. \tag{2.15}$$

Then, define E_{typ} as

$$E_{typ} = \bigcup_{b \in (1..B+L)} E^b.$$

We choose n sufficiently large so that the probability $P(E_{typ}) < \delta_1$, i.e. we condition the remainder of the probability of error analysis on the event E_{typ}^C that the erasures are strongly typical.

For any cut $\mathcal{V}_s$, define B_s^b as the event that, after the b^{th} block is simulated, the inputs to all of the nodes in $\mathcal{V}_d$ in the b^{th} block are identical for the true message w and the incorrect message $w_1 \neq w$ and the inputs to all of the nodes in $\mathcal{V}_s$ contain some difference for the two codewords. We will average the probability of error over all codebooks and codewords, as per standard coding arguments, so the message w_1 is arbitrarily chosen.

When w and w_1 produce identical inputs to the destination node in a time-block b, exactly one of the events B_s^b for one s-d cut occurs. We define E_1 as

$$E_1 = \bigcap_{b \in (1..B+L)} \left(\bigcup_{S \in \mathcal{S}} B_s^b \right). \tag{2.16}$$

Letting $\mathcal{S}_{seq}$ be the set of all $(B+L)$-length sequences of cut-sets $(S_1, S_2, ..., S_{B+L})$, the expression for E_1 in Equation (2.16) can be rewritten as

$$E_1 = \bigcup_{(S_1,S_2,...,S_{B+L}) \in \mathcal{S}_{seq}} \left(B_{S_1}^1 \bigcap B_{S_2}^2 \bigcap ... \bigcap B_{S_{B+L}}^{B+L} \right) \tag{2.17}$$

Note that the events $B_{s_l}^b$ and $B_{s_m}^{b+1}$ are clearly not independent of each other. However, given $B_{s_l}^b$, the events $B_{s_k}^{b-1}$ and $B_{s_m}^{b+1}$ are conditionally independent, so we study the probabilities $P(B_{s_m}^{b+1} | B_{s_l}^b, E_{typ}^C)$ where l and m are indices of cut-sets.

Because of startup transients, the probabilities of the B_s^b events for $b \in (1..L)$ (as the first codeword actually dependent on the message w spreads through the network) are especially difficult to define. We will show that we can ignore them in the error analysis, however.

We imagine the events B_s^b as states, each sequence $S_{seq} \in \mathcal{S}_{seq}$ as a path over time through the different states, and the conditional probabilities $P(B_{s_m}^{b+1} | B_{s_l}^b, E_{typ}^C)$ as transition probabilities. The goal then is to find the sequence with the maximum product of transition probabilities, because we desire to upper bound the probability of error. As long as

$$2^{nRB} \times \max_{S_{seq} \in \mathcal{S}_{seq}} P(B_{s_1}^1 | E_{typ}^C) \prod_{b \in (2..B+L)} P(B_{s_b}^b | B_{s_{b-1}}^{b-1}, E_{typ}^C) \tag{2.18}$$

can be made arbitrarily small by increasing n appropriately, then the output at the destination is uniquely decodable. Since we are upper-bounding the probabilities, we can set the first L product terms in the right side of Equation (2.18) to unity.

First, examine terms of the form $P(B_{s_l}^{b+1} | B_{s_l}^b, E_{typ}^C)$, which are transitions from one state back to the same state. Precisely, this means the probability, for the

different inputs w and w_1, given that they produce the same input for exactly the nodes in S^C in time-block b, that they will also produce the same input for exactly those same nodes in time-block $b+1$. All nodes in $S^C \cap (S^{C*})^C$, i.e. those nodes on the destination side of the cut that have no input edges in the cut-set, will necessarily still have identical inputs for time-block $b+1$ – there are no non-identical inputs to those nodes.

For every node $j \in S^{C*}$, there is at least one edge with independent codewords for the two messages w and w_1. Say that for node j there is exactly one edge (i,j) in $\mathcal{N}_I(j)$ in the cut-set. Then, since the codewords generated at i for the two messages w and w_1 will be independent, the input to j will only be identical for the two messages if the two codewords *differ only in the erased locations*. The probability that the codewords are identical in all unerased locations is no more than

$$2^{-n(1-\epsilon_{ij}-\delta)}.$$

Now say node j has the set $\{m_j\} = \mathcal{N}_I(j) \bigcap \left\{(i,j)|i \in S_l, j \in S_l^C\right\}$ of incoming edges in the cut-set. For any node j in S_l^C, the sum of the unerased values along incoming edges for the message w_1 must be equal to the sum of the unerased values for the message w for all n time slots in the block $b+1$ for the event $B_{s_l}^{b+1}$ to occur. If the bits along all edges in $\{m_j\}$ are erased, then the probability of the incoming sum being identical is 1. If any of the incoming edges has incoming bits which are unerased, then the probability that the sum of these bits for messages w and w_1 are equal is 1/2 for each time slot. Since the erasures are assumed to be typical, the probability that all timeslots in the block have identical bits for the two messages is less than

$$2^{-n(1-\prod_{(i,j)\in\{m_j\}}\epsilon_{ij}-\delta)} \tag{2.19}$$

when all the codewords being transmitted by nodes in S_l are independent. (Recall, that all the codewords transmitted by nodes in S_l during time-block $b+1$ will be independent, given the event $B_{s_l}^b$.) For the event $B_{s_l}^{b+1}$ to occur, all nodes j in S_l^C

must receive identical inputs, which will occur with probability less than

$$P(B_{s_l}^{b+1}|B_{s_l}^{b}, E_{typ}^{C}) \leq \prod_{j \in S_l^{C*}} 2^{-n(1-\prod_{(i,j)\in\{m_j\}} \epsilon_{ij} - \delta)}$$
$$\leq 2^{-n(C(S_l)-\delta\left|S_l^{C*}\right|)}. \tag{2.20}$$

Recall that we wish to show that Equation (2.18) can be made arbitrarily small for any sequence of error cut-sets S_{seq} as n and B grow large. Assume that $\hat{S}$ is the min-cut, i.e. the argument which minimizes the right-hand side of Equation (2.8). To prove that the capacity of the network is given by Equation (2.13), we must show that no sequence of error cut-sets S_{seq} has probability with error exponent asymptotically smaller than that of the sequence of min-cuts $\hat{S}_{seq} = \{\hat{S}, \hat{S}, \hat{S}, ...\}$.

There exist transitions which have probabilities much greater than the min-cut self-transition probability $P(B_{\hat{s}}^{b+1}|B_{\hat{s}}^{b}, E_{typ}^{C})$: Consider the sets $\hat{S}$ and $S_2 = \hat{S} \bigcup \hat{S}^{C*}$, illustrated in Figure 2.5. The set S_2^C has the property that there are no edges originating from $\hat{S}$ with endpoints in S_2^C. All nodes in $S_2^* \bigcup S_2^C$ will then generate the same codeword for messages w and w_1 for block $b+1$, since they have had exactly the same inputs in block b. Therefore, $P(B_{s_2}^{b+1}|B_{\hat{s}}^{b}, E_{typ}^{C})$ is nearly 1. (The probability is slightly less than unity because when $S_l \subset \hat{S}$, $P(B_{s_l}^{b+1}|B_{\hat{s}}^{b}, E_{typ}^{C})$ is small, but still positive.)

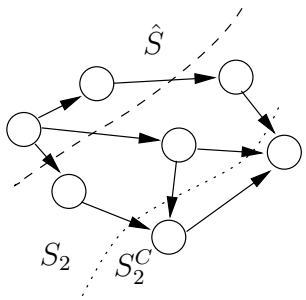

Figure 2.5: The probability $P(B_{s_2}^{b+1}|B_{\hat{s}}^{b}, E_{typ}^{C})$ is nearly 1.

Define $C(S_m|S_l)$ to be the exponent of the error transition probability $P(B_{s_m}^{b+1}|B_{s_l}^{b}, E_{typ}^{C})$, that is

$$P(B_{s_m}^{b+1}|B_{s_l}^{b}, E_{typ}^{C}) \leq 2^{-nC(S_m|S_l)+n\delta\left|S_m^{C*}\right|}.$$

We have already shown that $C(S_l|S_l)$ is given by $C(S_l)$, defined in Equation (2.9). By the same argument,

$$C(S_m|S_l) = \sum_{j \in S_m^C} (1 - \prod_{i \in S_l} \epsilon_{ij}). \tag{2.21}$$

The expression is the sum over every node j in the new error cut S_m^C, the probability that at least one edge which originates in the source-side of the old error cut S_l and terminates at that node j remains unerased.

Now we use the expression of Equation (2.21) to demonstrate that the probability of transitioning from error events $B_{s_l}^b$ to $B_{s_m}^{b+1}$ and back to $B_{s_l}^{b+2}$ always has an asymptotically smaller probability that the error event corresponding to staying in the min-cut error events $B_{\hat{s}}$ over the two time-block transitions. This is expressed in Lemma 2.6.1.

Lemma 2.6.1. *The expression*

$$C(S_m|S_l) + C(S_l|S_m) \geq C(\hat{S}|\hat{S}) + C(\hat{S}|\hat{S})$$

is valid for any cuts S_m and S_l when $\hat{S}$ is the min-cut.

Proof. Observe that $C(S_m \bigcup S_l) \geq C(\hat{S})$ and $C(S_m \bigcap S_l) \geq C(\hat{S})$, since $\hat{S}$ is defined to be the min-cut. Therefore, demonstrating that

$$C(S_m|S_l) + C(S_l|S_m) - C(S_m \bigcap S_l) - C(S_m \bigcup S_l)$$

is non-negative will complete the proof. We decompose the expressions resulting from the application of Equation (2.21) into summations over the sets $\{j \in$

$S_m^C \bigcap S_l^C\}$, $\{j \in S_m^C \bigcap S_l\}$, and $\{j \in S_l^C \bigcap S_m\}$.

$$
\begin{aligned}
& C(S_m|S_l) + C(S_l|S_m) - C(S_m \bigcap S_l) - C(S_m \bigcup S_l) \\
&= \sum_{j \in S_m^C} (1 - \prod_{i \in S_l} \epsilon_{ij}) + \sum_{j \in S_l^C} (1 - \prod_{i \in S_m} \epsilon_{ij}) \\
&\quad - \sum_{j \in S_m^C \bigcap S_l} (1 - \prod_{i \in S_m \bigcap S_l} \epsilon_{ij}) - \sum_{j \in S_m^C \bigcup S_l} (1 - \prod_{i \in S_m \bigcup S_l} \epsilon_{ij}) \qquad (2.22) \\
&= - \sum_{j \in S_m^C \bigcap S_l^C} \left(\prod_{i \in S_l} \epsilon_{ij} + \prod_{i \in S_m} \epsilon_{ij} \right) \\
&\quad - \sum_{j \in S_m^C \bigcap S_l} \left(\prod_{i \in S_l} \epsilon_{ij} \right) - \sum_{j \in S_l^C \bigcap S_m} \left(\prod_{i \in S_m} \epsilon_{ij} \right) \\
&\quad + \sum_{j \in S_m^C \bigcap S_l^C} \left(\prod_{i \in S_l \bigcup S_m} \epsilon_{ij} + \prod_{i \in S_l \bigcap S_m} \epsilon_{ij} \right) \\
&\quad + \sum_{j \in S_m^C \bigcap S_l} \left(\prod_{i \in S_l \bigcap S_m} \epsilon_{ij} \right) + \sum_{j \in S_l^C \bigcap S_m} \left(\prod_{i \in S_l \bigcap S_m} \epsilon_{ij} \right) \qquad (2.23)
\end{aligned}
$$

By combining the common sums and factoring out the common products of ϵ_{ij}'s within those sums, we obtain

$$
\begin{aligned}
& \sum_{j \in S_m^C \bigcap S_l^C} \left(\prod_{i \in S_l \bigcap S_l^C} \epsilon_{ij} \right) \left(1 + \prod_{i \in S_l \bigcap S_m^C} \epsilon_{ij} \prod_{i \in S_m \bigcap S_l^C} \epsilon_{ij} - \prod_{i \in S_l \bigcap S_m^C} \epsilon_{ij} - \prod_{i \in S_m \bigcap S_l^C} \epsilon_{ij} \right) \\
& + \sum_{j \in S_m^C \bigcap S_l} \left(\prod_{i \in S_l \bigcap S_l^C} \epsilon_{ij} \right) \left(1 - \prod_{i \in S_l \bigcap S_m^C} \epsilon_{ij} \right) \\
& + \sum_{j \in S_l^C \bigcap S_m} \left(\prod_{i \in S_l \bigcap S_l^C} \epsilon_{ij} \right) \left(1 - \prod_{i \in S_m \bigcap S_l^C} \epsilon_{ij} \right) \qquad (2.24)
\end{aligned}
$$

where each of the three terms is clearly non-negative. □

Now, we generalize Lemma 2.6.1:

Lemma 2.6.2. *For any sequence of K error-state transitions $(s_b, s_{b+1}, ..., s_{b+K} = s_b)$ which begins and ends with the same error cut-set, the error exponent for the*

K transitions is greater than the error exponent for the error event of staying in the min-cut error cut-set for those K time-blocks. Formally,

$$\sum_{k=1}^{K} C(S_{b+k}|S_{b+k-1}) \geq \sum_{k=1}^{K} C(S_k) \geq KC(\hat{S}), \tag{2.25}$$

where new self-transition error-cuts S_k are defined for $k \in (1, ..., K)$ as

$$S_k = \{i : i \text{ is in exactly } k \text{ of the sets } (S_{b+1}, ..., S_{b+K})\}.$$

The proof proceeds along the same lines as that of Lemma 2.6.1.

Thus, the probability of the error event E is bounded by

$$\begin{aligned} P(E) &\leq P(E_{typ}) + \sum_{k=1}^{2^{nRB}} P(E_k|E_{typ}^C) \\ &\leq \delta_1 + 2^{nRB} \max_{S\in\mathcal{S}} P(B_{s_1}^1|E_{typ}^C) \prod_{b\in(2..B+L)} P(B_{s_b}^b|B_{s_{b-1}}^{b-1}, E_{typ}^C) \\ &\leq \delta_1 + 2^{nRB} \prod_{b\in(1+L..B+L)} P(B_{\hat{s}}^b|B_{\hat{s}}^{b-1}, E_{typ}^C) \\ &\leq \delta_1 + 2^{nRB - nBC(\hat{S}) + nB\delta\left|\hat{S}^{C*}\right|}. \end{aligned}$$

Thus, if $R < C(\hat{S})$, a code with rate R and arbitrarily small error probability exists to communicate information from the source node to the destination node when the destination node has side information describing the locations of all the erasures in the network. We have thus demonstrated that the unicast capacity of the multiple-access erasure network is given by a modified min-cut max-flow cut-set rate bound.

2.7 Achievability in Erasure Networks with Generalized Broadcast and Finite-Field Sum Interference

In Section 2.6.2, we explicitly compute the conditional entropy $H(Y^*|A_s)$, as defined in Section 2.5 and use that entropy to compute $C(S_m|S_l)$, the exponent

of the error transition probabilities $P(B_{s_m}^{b+1}|B_{s_l}^{b}, E_{typ}^{C})$ in Equation (2.21). We are able to prove Lemmas 2.6.1 and 2.6.2 through manipulations of these particular expressions, as in Equations (2.22, 2.23, and 2.24).

The work of [14] makes the same claims in the context of deterministic networks, but contains a more general statement. That is, assuming a product distribution on the transmitted symbols at each of the nodes (which is the distribution that maximizes the cutset bound on our linear finite-field sum multiple-access interference model),

$$H(Y_{S_2^C}|X_{S_1^C}) + H(Y_{S_3^C}|X_{S_2^C}) + ... + H(Y_{S_1^C}|X_{S_K^C}) \geq KH(Y_{\hat{S}^C}|X_{\hat{S}^C}) \qquad (2.26)$$

where the terms are defined for any non-repeating sequence of cuts. (Lemma 6.4 of [14]). We then apply this lemma directly to our probability of error calculations, and the achievability follows with the same proof mechanism of Section 2.6.2.

2.8 Duality Observations

This work was inspired by studying the wireless erasure network of [2] and considering what form a receiver interference constraint might take. The finite-field sum is a first-cut choice to model a multiple-access channel, and in calculating the cut-set bound of such networks (with and without a broadcast channel constraint) the following observation was made:

Lemma 2.8.1. *The capacity of any multiple-access erasure network is equal to the capacity a wireless erasure network (of [2]) where the source and destination nodes from the multiple-access erasure network are interchanged, where the direction of every edge is reversed, and where the erasure probabilities associated with each edge remains unchanged.*

For example, the multiple-access erasure network in the top of Figure 2.6 has the same capacity of the wireless erasure network at the bottom of the figure. The relationship between the node inputs and outputs in the multiple-access network

is given by the matrix equation

$$\begin{bmatrix} Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} \gamma_{12} & 0 & 0 & 0 & 0 \\ 0 & \gamma_{13} & \gamma_{23} & 0 & 0 \\ 0 & 0 & 0 & \gamma_{24} & \gamma_{34} \end{bmatrix} \begin{bmatrix} X_{12} \\ X_{13} \\ X_{23} \\ X_{24} \\ X_{34} \end{bmatrix}$$

while the relationship in the dual wireless erasure network is given by

$$\begin{bmatrix} Y_{12} \\ Y_{13} \\ Y_{23} \\ Y_{24} \\ Y_{34} \end{bmatrix} = \begin{bmatrix} \gamma_{12} & 0 & 0 \\ 0 & \gamma_{13} & 0 \\ 0 & \gamma_{23} & 0 \\ 0 & 0 & \gamma_{24} \\ 0 & 0 & \gamma_{34} \end{bmatrix} \begin{bmatrix} X_2 \\ X_3 \\ X_4 \end{bmatrix}.$$

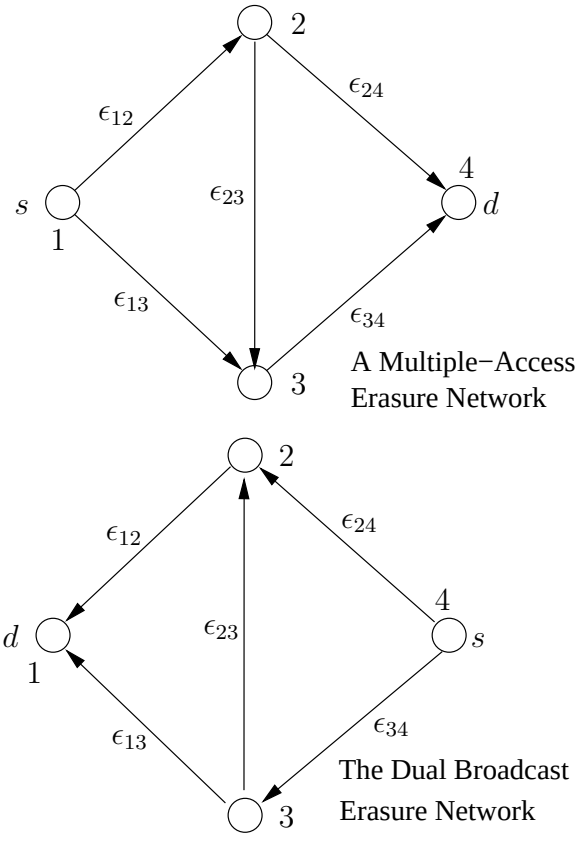

Figure 2.6: Dual multiple-access and wireless broadcast erasure networks

Is is interesting to note that the same capacity duality exists for the generalized interference setting, as well, again simply by swapping the source and destination

nodes, reversing the direction of each edge, and interchanging broadcast and receiver interference constraints at each appropriate antenna. It is easy to see that the capacities are identical, because the corresponding transfer matrices A_s for the dual networks are just transposes of each other. The rank of A^T, obviously, is equal to the rank of A.

Chapter 3

Benefits of Feedback

This chapter is joint work with Professor Babak Hassibi, at the California Institute of Technology.

3.1 Introduction

This chapter presents a throughput-optimal transmission strategy for a unicast wireless erasure network when feedback is available, which has the following advantages: The algorithm requires a very limited form of acknowledgment feedback. It is completely distributed, and independent of the network topology. Finally, communication at the information theoretic cut-set rate requires no network coding and no rateless coding on the packets. This simple strategy consists of each node randomly choosing a packet from its buffer to transmit at each opportunity. However, the packet is only deleted from a node's buffer once it has been successfully received by the final destination.

It is well know that in the point-to-point channel model, feedback can never increase the value of the information theoretic capacity[3]. However, there several significant advantages to having feedback. Feedback allows coding strategies which can significantly increase the probability of error exponent, for example the Schalkwijk-Kailath scheme for additive Gaussian noise channels[17]. Feedback can also allow transmission strategies with extremely simple coding algorithms. Specifically, consider the binary symmetric erasure channel. When feedback is available, the transmitter can simply repeat each bit until successfully received. Capacity is achieved, and in some sense, no coding whatsoever is required.

In this chapter, a unicast model of a lossy wireless network of queues is con-

sidered, similar in spirit to the wireless erasure network[2] that we have been discussing, but from more of a networking perspective. This network model still contains the crucial characteristics of independent erasure channels/loss probabilities on a directed graph, a wireless broadcast requirement, and unicast for a single source-destination pair, but now also includes asynchronous transmission timing. With transmit opportunities occurring as a unit rate Poisson process, a transmission by one node will be received independently with some fixed probability by each other node in the network. The network model will allow general feedback, but it will be shown that only a very limited form of acknowledgment feedback is required to achieve the throughput-optimal cut-set capacity. The primary differences between the model of this chapter and that of [2] are first, the availability of feedback, and second, an asynchronous, memoryless arrival process (rather than a slotted-time model). We do not believe the difference in timing model to be critical, and conjecture that the given transmission algorithm will be throughput-optimal in a slotted-time model. Additionally, the addition of feedback eliminates the requirement for any side-information concerning the location of erasures throughout the network to achieve capacity, in contrast with the decoding strategies of [2] and of Chapter 2.

A similar asynchronous network model was studied in [18]. The authors' model demonstrates the usefulness of network coding: with no feedback, but allowing network coding and additionally, a packet header describing the linear combinations of data packets included in the transmission, they demonstrate the achievability of the cut-set bound. This work highlights somewhat of a dual statement: without any sort of coding, but with feedback, the same cut-set packet rate is achievable.

The paper [19] also is concerned with a similar wireless lossy packet network model. With a backpressure algorithm, throughput-optimality in a multicommodity sense is also achieved in a multiple-source multiple-destination network. This algorithm requires link-level feedback, and for each node to maintain knowledge of the queue state of, in worst case, every other node in the network. It provides a decision process, when multiple nodes in the network receive copies of the same

packet, to determine which (if any) of those nodes should keep that packet and attempt to forward it onward.

In contrast, the routing algorithm described in this chapter is completely decentralized and requires no conferencing among nodes to decide who should "keep" a packet that it has received. Instead, there will in general be multiple copies of each packet throughout the network.

Specifically, the algorithm is as follows: Whenever a node has an opportunity to transmit a packet, it will randomly choose one packet from its buffer. Every time that a packet successfully reaches the final destination node, that node will (errorlessly) broadcast an acknowledgment to every node in the system stating that this particular packet has successfully completed its transit of the network. Only after receiving this acknowledgment from the *final destination node* will any node remove the packet from its buffer. Indeed, the entire network will then flush that packet from all the buffers. This paper shows via Foster's Theorem and an application of an appropriate and novel Lyapunov function the stability of all network queues under this operation as long as the input data rate is less than the minimum-cut of the network. The authors are unaware of previous uses of an exponential Lyapunov function of the form we consider in showing stability results.

The advantages of this throughput optimal strategy include

- It requires no coding, particularly no network coding at intermediate nodes.
- The only information that a packet header must contain is an identifier - no additional information is required.
- It is completely decentralized. No coordination or conferencing, other than the acknowledgment feedback, is required.
- It is topology independent. No node other than the source needs any information about the layout of the network. The source must only be given the value of the min-cut, which could even be adaptively estimated, if desired.

- The only feedback required, a simple acknowledgment from the destination, is practically already implemented in real systems.

The main thrust of this chapter: A demonstration that, in this lossy unicast wireless network, feedback obviates the need for coding, network coding in particular.

3.2 Network Model and Notation

Consider a directed (possibly cyclic) graph $G(V, E)$ with $n + 2$ nodes: a source node, a destination node, and n intermediate nodes. Label the source node s, the destination d, and index the other nodes as $i \in 1, .., n$. To each edge pair $(i, j) \in V \times V$ assign an erasure probability $0 \leq \epsilon_{ij} \leq 1$. If the directed edge (i, j) does not exist in the graph, then assign $\epsilon_{ij} = 1$. Define $\mu_{ij} = 1 - \epsilon_{ij}$.

Because of the wireless nature of the model, when a node i transmits a packet, each other node in the system j has the probability μ_{ij} of successfully receiving that packet. The events that packets are dropped are independent, that is i.i.d. across time for any fixed edge (i, j), and independent between every pair of edges. We will consider the case where the events corresponding to combinations of packet drops from a single transmitter at a fixed time can be correlated in a later section.

Allow an infinite buffer to exist at each node in the network. Packets will exogenously arrive at the source node s according to a Poisson process with arrival rate λ. At average rate 1 exponentially distributed intervals, each node in the network (other than the destination node) receives an opportunity to transmit a packet.

Each packet has a unique identifier in its header. Therefore, if a node already has a copy of a particular packet and it receives that packet again, the contents of that node's buffer remain unchanged.

A feedback mechanism exists such that when the destination node receives a packet, it instantaneously, via a delay-free feedback, notifies all of the other nodes in the system of that fact. All nodes in the system can then immediately remove that particular packet from their buffer.

Finally, this asynchronous model does not consider any receiver interference or the possibility of simultaneous arrivals.

3.3 Cut-set Upper Bound and Transmission Strategy

Under any transmission strategy, the cut-set upper-bound remains valid. Intuitively, the cut-set upper-bound is obtained by dividing the network into two parts S and S^C and creating two super-nodes. That is, by allowing free, unlimited communication among the nodes in S and among the nodes in S^C, we can only increase the capacity of the system.

With that in mind, let S be a subset of the $n+2$ nodes such that $s \in S$ and $d \in S^C$. There are 2^n such subsets. Let $\mathcal{S}$ be the set of all such subsets. The super-node created by joining all nodes in S together will still have opportunities to transmit at exponentially distributed intervals, but now the sum rate will be $|S|$ – a rate of 1 for each node in S. For each node $i \in S$, because of the unlimited free communication on the right side of the cut in S^C, as long as one of the nodes $j \in S^C$ successfully receives the packet, we can count it in the total communication throughput. Therefore, define

$$C(S) = \sum_{i \in S} \left(1 - \prod_{j \in S^C} \epsilon_{ij} \right) \tag{3.1}$$

as the cut-set capacity for the subset S, i.e. an upperbound on the rate of packets that can be transmitted across the $S - S^C$ cut, exactly as per [2].

The total throughput $T < C(S)$ then, for every subset S, and

$$R < \min_{S \in \mathcal{S}} C(S).$$

The authors would like to emphasize the key role that the subsets S will play in the proof and the derivation of the stability results. The minimum of $C(S)$ over all $S - S^C$ cuts must emerge from any stability equations; therefore it is reasonable that each cut-set represented by S must play a role. As will be further explained,

the sets S will become essential as indices to the variables m_S which describe the state of our Markov chain model. It will become clear that as the state variable m_S corresponding to the subset S becomes large, the requirement $\lambda < C(S)$ becomes a dominant constraint.

The network operates in the following manner: At every transmission opportunity for a node, that node *randomly* chooses one of the packets in its buffer to transmit. If the buffer is empty, then that transmission opportunity is lost. Only when acknowledgment from the final destination d is received will a node remove a packet from its buffer; therefore in general there are multiple copies of each packet in the network.

Theorem 3.3.1. *Under this randomized transmission strategy, all queues in a wireless erasure network with feedback are stable as long as $\lambda < C(S)$ for all $S \in \mathcal{S}$.*

At first glance, this randomized strategy seems unnecessarily wasteful. Consider a network which is a simple serial line of queues. In this case, it is obvious that an optimal strategy, when link-level feedback is available, is to stop attempting to transmit a packet (and remove it from one's queue) as soon as it is successfully received at the next queue down the line. Leaving a successfully transmitted packet in the queue could result in the retransmission of that packet, possibly wasting a transmission opportunity that could be put to better use sending a new packet.

However, the randomization is crucially important in achieving the minimum-cut value for this network and for a general network. To achieve the min-cut, it is essential that all transmitters on the min-cut boundary transmit packets at almost every channel use and that these packets be almost always distinct. As the input rate λ increases, the min-cut slowly becomes the bottleneck of the network and the queues on its boundary will grow large. This will ensure that each transmitter always has a packet to transmit with high probability. The randomization in packet transmission guarantees that for such long queues the probability that two transmitters along the min-cut transmit the same packet is very low. Deterministic strategies, such as FIFO for example, cannot guarantee this without coordination,

and so the randomized strategy is essential to achieving the optimal throughput in a completely decentralized manner.

In the line network in particular, edges which are not the minimum cut can afford to retransmit a certain number of packets, since they have extra capacity. In fact, edges which lie downstream of the minimum cut edge will have relatively short queue lengths (compared to the queues upstream of the minimum cut edge) since they can remove packets from their queue at a faster rate than those packets can arrive across the minimum cut edge.

All queues upstream of the minimum cut, however, will have a relatively large number of packets. If many packets are transmitted multiple times across the minimum-cut edge, then the queue length at that edge will grow large. However, as the queue length grows large (with new arrivals), the probability of picking a "useless" packet will *decrease* as most of the packets in the queue have not yet been successfully sent. This unwanted probability will be made as arbitrarily small as required (depending on the ratio between λ and the minimum μ) as the queue length grows.

Note that strategies such as the one in [19] implement an algorithm to assure that there is only one copy of each packet in the network at a time. Such strategies necessarily require some amount of link-level feedback and inter-node communication to guarantee the single copy, under the broadcast nature of the wireless medium. The strategy of this work eliminates the need for any additional intra-network communication, other than the single feedback acknowledgment.

3.4 Proof Preliminaries

3.4.1 Notation and Description of Markov Chain Model

Before formally beginning the proof of Theorem 3.3.1, some additional notation must be defined.

The subset S has already been defined to be an element of $\mathcal{S}$, which is essentially the power-set of n. Precisely, $\mathcal{S}$ differs the power-set of n only in that all $S \in \mathcal{S}$

always include the source node s and never include destination node d. Equivalently, each element S can represent an index in the set $\{0, 1, 2, ..., 2^n - 1\}$. With this notion, the length-n binary expansion of S indicates which of the n nodes are contained within the subset S. This yields a one-to-one correspondence between subsets, cut-sets, and indices, all represented by the overloaded notation S.

A continuous time Markov chain model is used to describe the state of the queuing network. Transitions between states will occur when one of three different types of events happen in the network:

- A new packet is received (at rate λ) by the source node s.
- A packet is successfully transmitted from some node i in the system to some subset of the receivers.
- A packet is successfully received by the the destination node d and therefore exits the network.

By the asynchronous, continuous time model of the network, no two of these events can occur simultaneously.

In the $n = 1$ three node network, the size of the buffers at the source node s and the intermediate node 1 are sufficient to describe any state of the system. There must be more packets at the source node s than at the intermediate node 1 at any point in time. By the given network operation protocol, no packet is deleted from a queue until it reaches the final destination, so that if a packet is present anywhere within the system, it must be present at the source node s.

One option for the state variable of the system is to use $\underline{q} = (q(s), q(1))$, representing the lengths of the queues at source and relay nodes respectively. This notation has the disadvantage that there would exist constraints such as "the number of packets at 1 must be smaller than the number of packets at s", i.e. $q(s) \geq q(1)$, on the state space. To eliminate the need for such constraints, consider an alternate notation to describe the system state.

Let m_1 be the number of packets which appear at both the nodes s and 1. Let m_0 be the number of packets which appear at the source node uniquely. Then the source node has a total of $m_0 + m_1$ packets, while the relay node has exactly m_1 packets in its buffer.

This state description can be generalized to an $n+2$ node network. The Markov chain describing the system state is a vector $\underline{m}$ with 2^n dimensions:

$$\underline{m} = (m_0, m_1, ..., m_S, ..., m_{2^n-1}) \tag{3.2}$$

The dimensions of the state vector $\underline{m}$ are indexed by the subsets $S \in \mathcal{S}$. The value m_S is the number of packets which appear at every node $i \in S$ and at no node $j \in S^C$. Therefore, the number of packets $q(i)$ which appear any node $i \neq s, d$ in the network is a function of $\underline{m}$. Let

$$\mathcal{S}_i = \left\{ S \in \mathcal{S} | \text{the } i^{th}\text{-least significant bit in the binary expansion of } S \text{ is a } 1 \right\}.$$

Then

$$q(i) = \sum_{S \in \mathcal{S}_i} m_S,$$

while the destination node d retains no buffer, and the source node s has

$$q(s) = \sum_{S \in \mathcal{S}} m_S$$

packets in its buffer.

Figure 3.1 illustrates the queue lengths for a network with $n = 2$, using the binary expansion of the S indices.

3.4.2 Markov Chain Evolution - Transition Model

To understand the evolution of the Markov chain model describing the state $\underline{m}$ of the queuing system, first take an example of the network where $n = 1$.

Successful transmission events can cause three different kinds of transitions to the state vector $\underline{m} = (m_0, m_1)$.

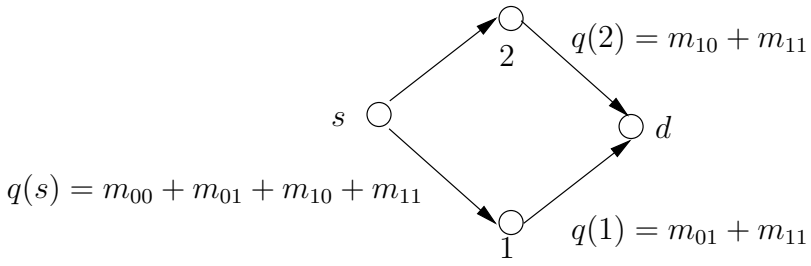

Figure 3.1: Relationship between queue lengths and state for the case with two relay nodes

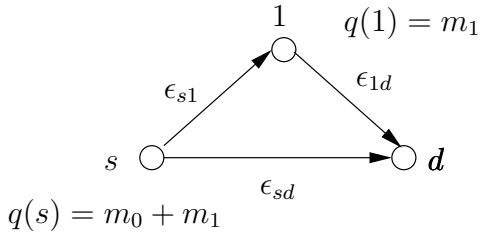

Figure 3.2: A general $n = 1$ wireless erasure network.

- There is an exogenous arrival to the system. In this case, the source node receives a new packet; the source is therefore the only node in the system which has that particular packet in its buffer. Thus, the value of m_0 is increased by 1.

- A packet at the source node s or the relay node 1 can be successfully received by the destination d, and therefore flushed from the network. If this packet was transmitted by the source, it may have come from either the set of m_0 packets (with probability $m_0/(m_0 + m_1)$) or from the set of m_1 packets (with probability $m_1/(m_0 + m_1)$). In this case, the appropriate variable m_0 or m_1 would decrease by 1. If this packet was transmitter by the relay node 1, then by definition it must have been one of the m_1 packets at both the source and the relay, so m_1 would decrease by 1.

- If the transmitter selects a packet from the set of m_0 packets, and that packet is successfully received by node 1, but not by the receiver, then that

particular packet would now be in both nodes' queues. In that case, we have a transition in which m_0 decreases by 1 (there is one less packet which is unique to node 1) and m_1 increases by 1 (there is one additional packet which is located at both the source node s and the relay node 1.)

Each of these possible transitions and their individual rates are illustrated in Figure 3.3.

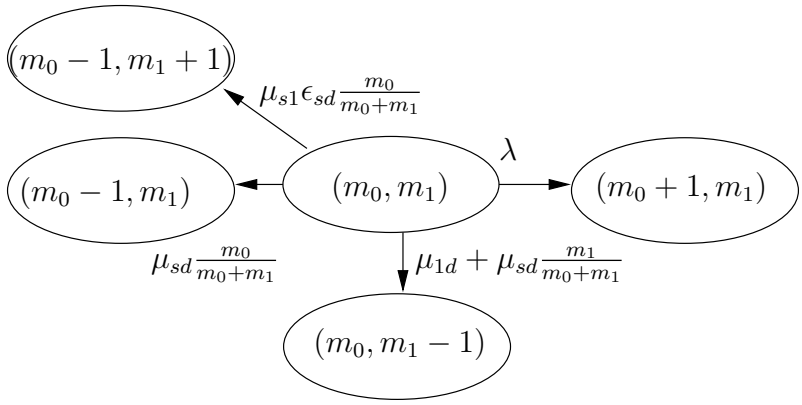

Figure 3.3: Possible transitions and transition rates from a state (m_0, m_1) in the $n = 1$ wireless erasure network.

In general, a network with n relay nodes has these same three kinds of transitions:

- A packet arrives at the source node, with rate λ. In this case, m_0 increases by 1.

- A packet (which exists in the subset S_1 of nodes) exits the system from some node i, with rate

$$\sum_{i \in S_1} \mu_{id} \frac{m_{S_1}}{q(i)}.$$

  Here, m_{S_1} decreases by 1.

- A packet (which exists in the subset S_1) transmitted at some node i is successfully received at some subset of possible receiver nodes, at least one of which did not previously have that particular packet in its buffer. In this case,

let S_2 be the new subset of nodes which have this packet. This constrains $S_1 \subset S_2$, and this occurs with rate

$$\sum_{i \in S_1} \left(\prod_{j \in S_2 / S_1} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) \frac{m_{S_1}}{q(i)}.$$

Here, m_{S_1} decreases by 1 while m_{S_2} increases by 1. It is important to note that in this kind of transition, the subset S_2 whose variable m_{S_2} increases must always be a superset of the subset S_1 whose variable m_{S_1} decreases.

3.4.3 Queue Stability and Foster's Theorem

We desire to show that, for any arrival rate $\lambda < \min_{S \in \mathcal{S}} C(S)$, all the queues in the network are stable. We first present a review of Foster's Theorem, which is the main proof mechanism[20].

Theorem 3.4.1. *Foster's Theorem. Let the transition matrix $\boldsymbol{P}$ on the countable state space M be irreducible and suppose there exists a function $V : M \to \mathbb{R}$ such that $\inf_m V(m) > -\infty$ and*

$$\sum_{k \in M} p_{mk} V(k) < \infty \qquad \text{for all } m \in F$$

$$\sum_{k \in M} p_{mk} V(k) < V(m) - \delta \qquad \text{for all } m \notin F$$

for some finite set F and some $\delta > 0$. Then the corresponding homogeneous Markov chain is positive recurrent.

Intuitively, the theorem states that as long as there is a Lyapunov function which is on average decreasing, then the value of that function cannot go to infinity with increasing time.

3.5 Proof for the Case $n = 1$

This section contains a demonstration of the stability proof for the simplest network, the case where $n = 1$, illustrated in Figure 3.2. Note that for this

particular network, the cut-set bound evaluates to

$$\min\left(1-\epsilon_{s1}\epsilon_{sd}, 1-\epsilon_{1d}+1-\epsilon_{sd}\right)$$

Lemma 3.5.1. *The network illustrated in Figure 3.2 is stable for*

$$\lambda < \frac{N}{N+1}\frac{1}{1+\delta}\min\left(1-\epsilon_{s1}\epsilon_{sd}, 1-\epsilon_{1d}+1-\epsilon_{sd}\right)$$

for any fixed $N > 0$ *and* $\delta > 0$.

By choosing $N >> 1$ and $\delta << 1$ appropriately, for any λ less than the cut-set bound, the randomized transmission policy with feedback stabilizes all the network queues.

Proof. Consider the Lyapunov function

$$V\left(m_0, m_1\right) = N\left(1+\delta\right)^{m_0} + \left(1+\delta\right)^{m_0+m_1}. \tag{3.3}$$

This Lyapunov function is "rewarded" (i.e. decreases) when m_0 decreases and penalized when m_0 increases. When a packet is received at the relay node, the function is rewarded (while m_1 increases, m_0 simultaneously decreases) and when a packet leaves the system (i.e. m_1 decreases) the function is also rewarded.

We identify three different cases to study. These cases arise first, because several state transitions in the Markov chain of Figure 3.3 become unavailable in certain states (for example, when $m_0 = 0$ a packet cannot transition from the subset m_0 to the subset m_1). Secondly, some of the cases individually give rise to the required cut-set constraints on λ that the cut-set bound requires.

- Case 1: When $m_0 = 0$ and $m_1 > 0$.
- Case 2: When $m_0 > 0$ and $m_1 = 0$.
- Case 3: When both $m_0 > 0$ and $m_1 > 0$.

As previously stated, when one of the variables in the state description $\underline{m}$ is equal to zero, one or more transitions from the state transition Figure 3.3 become unavailable.

3.5.1 Case 1 : $m_0 = 0$

Evaluate the expected change in the value of the Lyapunov function $V(0, m_1)$ to determine when it is bounded away from zero:

$$\begin{aligned}
&\lambda \left(V\left(1, m_1\right) - V(0, m_1)\right) + \left(\mu_{1d} + \mu_{sd}\right)\left(V(0, m_1 - 1) - V(0, m_1)\right) < 0 \\
&\lambda \left(N(1+\delta)^1 + (1+\delta)^{1+m_1} - N(1+\delta)^0 - (1+\delta)^{m_1}\right) \\
&\quad + \left(\mu_{1d} + \mu_{sd}\right)\left(N(1+\delta)^0 + (1+\delta)^{m_1 - 1} - N(1+\delta)^0 - (1+\delta)^{m_1}\right) < 0 \\
&\lambda \left(N\delta + (1+\delta)^{m_1}(1+\delta-1)\right) + \left(\mu_{1d} + \mu_{sd}\right)(1+\delta)^{m_1-1}\left(1 - (1+\delta)\right) < 0 \\
&\lambda \left(N + (1+\delta)^{m_1}\right) < \left(\mu_{1d} + \mu_{sd}\right)(1+\delta)^{m_1-1} \\
&\lambda < \left(\mu_{1d} + \mu_{sd}\right)\frac{(1+\delta)^{m_1-1}}{(1+\delta)^{m_1} + N}
\end{aligned}$$

The first line represents the change in Lyapunov function for all possible state transitions from the state $\underline{m} = (0, m_1)$, weighted by the appropriate rates to calculate the expectation. The right hand side of the final inequality approaches as arbitrarily close to $(\mu_{1d} + \mu_{sd})\frac{1}{1+\delta}$ as desired for sufficiently large m_1. That is, for any given $\lambda < (\mu_{1d} + \mu_{sd})\frac{1}{1+\delta}$, there exists a finite $\tilde{m}_1$ such that the expected value of the Lyapunov function decreases for states $\underline{m} = (0, m_1)$ for all $m_1 > \tilde{m}_1$.

Note that for this case, one of the two cut-set bounds on λ is obtained.

3.5.2 Case 2 : $m_1 = 0$

For $\underline{m} = (m_0, 0)$, Figure 3.3 indicates that we desire

$$\begin{aligned}
&\lambda \left(V\left(m_0 + 1, 0\right) - V(m_0, 0)\right) \\
&\quad + \mu_{sd}\left(V(m_0 - 1, 0) - V(m_0, 0)\right) + \mu_{s1}\epsilon_{sd}\left(V(m_0 - 1, 1) - V(m_0, 0)\right) < 0 \\
&\lambda \left(N(1+\delta)^{m_0} + (1+\delta)^{m_0}\right) < \mu_{sd}\left(N(1+\delta)^{m_0-1} + (1+\delta)^{m_0-1}\right) + \mu_{s1}\epsilon_{sd}N(1+\delta)^{m_0-1} \\
&\lambda(N+1)(1+\delta) < \mu_{sd}\epsilon_{sd}(N+1) + \mu_{s1}\epsilon_{sd}N \\
&\lambda < \mu_{sd}\frac{1}{1+\delta} + \mu_{s1}\frac{N}{N+1}\frac{1}{1+\delta}
\end{aligned}$$

As long as

$$\lambda < \left(1 - \epsilon_{sd} + \epsilon_{sd}(1 - \epsilon_{s1})\right)\frac{N}{N+1}\frac{1}{1+\delta} = \left(1 - \epsilon_{sd}\epsilon_{s1}\right)\frac{N}{N+1}\frac{1}{1+\delta}$$

then the expected value of the Lyapunov function decreases for all states of the form $\underline{m} = (m_0, 0)$.

3.5.3 Case 3 : $m_0, m_1 > 0$

All of the transitions in Figure 3.3 are possible from the state $\underline{m} = (m_0, m_1)$.

$$
\begin{aligned}
&\lambda(V(m_0+1,m_1) - V(m_0,m_1)) \\
&\quad + \left(\mu_{1d} + \mu_{sd}\frac{m_1}{m_0+m_1}\right)(V(m_0,m_1-1) - V(m_0,m_1)) \\
&\quad + \left(\mu_{sd}\frac{m_0}{m_0+m_1}\right)(V(m_0-1,m_1) - V(m_0,m_1)) \\
&\quad + \left(\mu_{s1}\epsilon_{sd}\frac{m_0}{m_0+m_1}\right)(V(m_0-1,m_1+1) - V(m_0,m_1)) < 0 \\
&\lambda N(1+\delta)^{m_0} + (1+\delta)^{m_0+m_1} \\
&\quad < (\mu_{1d}+\mu_{sd})(1+\delta)^{m_0+m_1-1} \\
&\quad + \mu_{sd}\frac{m_0}{m_0+m_1}N(1+\delta)^{m_0-1} + (1+\delta)^{m_0+m_1-1} \\
&\quad + \mu_{s1}\epsilon_{sd}\frac{m_0}{m_0+m_1}N(1+\delta)^{m_0-1} \\
&\lambda\left(N(1+\delta) + (1+\delta)^{m_1+1}\right) \\
&\quad < (\mu_{1d}+\mu_{sd})(1+\delta)^{m_1} + (1-\epsilon_{s1}\epsilon_{sd})\frac{m_0}{m_0+m_1}N
\end{aligned}
$$

By inspection, note that regardless of the value of m_0, an m_1^* can be chosen sufficiently large such that if $\lambda < (\mu_{s1}+\mu_{sd})\frac{1}{1+\delta}$, the expected value of the Lyapunov function is decreasing for all states with $m_1 > m_1^*$. Likewise, for any fixed m_1, choose $m_0 > Nm_1$, and if $\lambda < \min(1-\epsilon_{s1}\epsilon_{sd}, \mu_{1d}+\mu_{sd})$, then the expected value of the Lyapunov function is decreasing. Thus, there are only a finite number of states where the expected value of Lyapunov function is increasing, and the requirements of Foster's Theorem are fulfilled. □

3.6 Proof for General Network

Recall Theorem 3.3.1, which we desire to prove:

Theorem 1: Under the given randomized transmission strategy, all queues in a wireless erasure network with feedback are stable as long as $\lambda < C(S)$ for all $S \in \mathbb{S}$.

For a general wireless erasure network with $n+2$ nodes, recall the Markov chain describing the system evolution described in Section 3.4.1. Foster's Theorem is utilized to demonstrate the stability of this Markov chain for a general $n+2$ node network.

3.6.1 General Lyapunov Function

For a $n+2$ node network, define the Lyapunov function $V(\underline{m})$ as

$$V(\underline{m}) = \sum_{S \in \mathbb{S}} N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'}} \tag{3.4}$$

where the $N_{|S|}$ and δ are fixed constants. The $N_{|S|}$ should jointly satisfy

$$N_{|S|} > N \sum_{S' \supset S} N_{|S'|}. \tag{3.5}$$

To form some intuition on the particular choice of Equation (3.4), consider the three kinds of transitions that can occur in our Markov Chain. When a packet arrives in the system, m_0 increases by 1, and the value of the Lyapunov function increases. Whenever any other transition occurs, the system is, in some sense, advancing a packet toward the final destination, and we would like the value of the Lyapunov function to decrease. This can happen in two ways:

- A packet which appears in the subset S_1 of nodes can exit the system. Then, m_{S_1} will decrease, and all of the terms in the summation corresponding to $S \supseteq S_1$ will decrease in value. (i.e., those terms which contain the factor $(1+\delta)^{m_{S_1}}$). The Lyapunov function therefore decreases in value.

- A packet which appears in the subset S_1 of nodes will arrive at some other nodes, and then will appear in the subset $S_2 \supset S_1$. Then, all the terms in the summation corresponding to S such that $S_1 \subseteq S$, but $S_2 \nsubseteq S$, will decrease

in value (i.e. those which contain the factor $(1+\delta)^{m_{S_1}}$ but not $(1+\delta)^{m_{S_2}}$). However, those which contain both the factor $(1+\delta)^{m_{S_1}}$ and $(1+\delta)^{m_{S_2}}$ will remain unchanged (since m_{S_1} decreases by 1 and m_{S_2} increases by 1).

3.6.2 Proof

The proof of Theorem 3.3.1 follows directly from the following lemma:

Lemma 3.6.1. *The expected value of the function $V(\underline{m})$, defined in Equation (3.4), is increasing only on a finite number of states whenever $\lambda < \frac{N}{N+1}\frac{1}{(1+\delta)^2}\min_{S\in\mathcal{S}} C(S)$.*

Thus, for any $\lambda < \min_{S\in\mathcal{S}} C(S)$, we can find an appropriate Lyapunov function to show the system's stability by choosing N sufficiently large and δ sufficiently small.

Proof. First fix $S \in \mathcal{S}$, and examine the term in the main summation of Equation (3.4) corresponding that S. Then, determine which transitions of the Markov chain effect the value of that term.

Let

$$V_S(\underline{m}) = N_{|S|}\,(1+\delta)^{\sum_{S'\subseteq S} m_{S'}}\,. \tag{3.6}$$

An arrival to the system effects every V_S, since every term V_S contains m_0. Thus $\forall S \in \mathcal{S}$,

$$\begin{aligned} &V_S(m_0+1, m_1, ...) - V_S(m_0, m_1, ...) \\ &\quad = N_{|S|}\,(1+\delta)^{\sum_{S'\subseteq S} m_{S'}+1} - N_{|S|}\,(1+\delta)^{\sum_{S'\subseteq S} m_{S'}} \\ &\quad = \delta N_{|S|}\,(1+\delta)^{\sum_{S'\subseteq S} m_{S'}} \end{aligned} \tag{3.7}$$

These events occur at rate λ.

If a packet appearing in cut-set S_1 departs the system, precisely the terms $V_S(\underline{m})$ when $S \supseteq S_1$ will decrease, since they are the only terms in the Lyapunov function

Equation (3.4) which contain m_{S_1}. For $S \supseteq S_1$,

$$\begin{aligned} &V_S(m_0, m_1, ..., m_{S_1} - 1, ...) - V_S(m_0, m_1, ...) \\ &\quad = N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1} - N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'}} \\ &\quad = -\delta N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1}. \end{aligned} \tag{3.8}$$

These events will occur when any node $i \in S_1$ transmits a packet in S_1 which is successfully received by the destination node d. Given an opportunity to transmit, the node i chooses a packet in S_1 with probability $\frac{m_{S_1}}{q(i)}$, and the packet is successfully received at the destination with probability μ_{id}. Thus, packets from S_1 will leave the system with rate

$$\sum_{i \in S_1} \frac{m_{S_1}}{q(i)} \mu_{id} \tag{3.9}$$

The final possible transition type occurs when a packet located at the nodes in subset S_1 is successfully received at some set of nodes which did not previously have that packet, but not the destination d, resulting in that packet being finally in the subset $S_2 \supset S_1$. Thus m_{S_1} will decrease by 1, and m_{S_2} will increase by 1. The only terms $V_S(\underline{m})$ that will change are those containing m_{S_1} but not m_{S_2}. Thus, for S such that $S \supseteq S_1$ and $S \not\supseteq S_2$,

$$\begin{aligned} &V_S(m_0, m_1, ..., m_{S_1} - 1, ..., m_{S_2} + 1, ...) - V_S(m_0, m_1, ...) \\ &\quad = N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1} - N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'}} \\ &\quad = -\delta N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1}. \end{aligned} \tag{3.10}$$

These events occur when any node $i \in S_1$ transmits a packet in S_1, and that packet is successfully received by all the nodes $j \in S_2/S_1$, and not successful in reaching nodes $\{j | j \notin S_2\}$, including the destination node d. The total rate of such events is

$$\sum_{i \in S_1} \frac{m_{S_1}}{q(i)} \prod_{j \in S_2/S_1} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij}. \tag{3.11}$$

The expected increase in the total Lyapunov function due to arrivals should be less than the expected decrease due to departures and transitions on all but a finite

number of state $\underline{m}$. The sum of changes over all of the terms must therefore satisfy

$$\lambda \sum_{S \in \mathcal{S}} N_{|S|} \left(1+\delta\right)^{\sum_{S' \subseteq S} m_{S'}}$$
$$< \sum_{S \in \mathcal{S}} \sum_{\{(S_1,S_2)|S_1 \subset S_2, S_1 \subseteq S, S_2 \nsubseteq S\}} \left(\sum_{i \in S_1} \frac{m_{S_1}}{q(i)} \prod_{j \in S_2/S_1} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) N_{|S|} \left(1+\delta\right)^{\sum_{S' \subseteq S} m_{S'}-1}$$
$$+ \sum_{S \in \mathcal{S}} \sum_{S_1 \subseteq S} \left(\sum_{i \in S_1} \frac{m_{S_1}}{q(i)} \mu_{id} \right) N_{|S|} \left(1+\delta\right)^{\sum_{S' \subseteq S} m_{S'}-1} \tag{3.12}$$

In the second line of Equation (3.12), the first summation is over terms in the Lyapunov function. The second summation is over transitions of the possible pairs of S_1 and S_2 which will effect that particular term, and the third summation is over nodes which could possibly transmit and create that transition. The final terms of the second line represent the value of the change in that term $V_S(\underline{m})$.

Similarly, in the third line of Equation (3.12), the first summation is over the terms of the Lyapunov function, and the second is over the possible departures from the system which can effect the value of each term. Within the parentheses is the rate of those departures, and the final terms again represent the value of the change in the term $V_S(\underline{m})$.

Note that if $q(i) = 0$, that is, no packets are currently in the queue at node i, then for any S such that $i \in S$, $m_S = 0$. In this case, take

$$\frac{m_S}{q(i)} = \frac{0}{0} = 0$$

since this node cannot transmit any packets.

We must show that Equation (3.12) holds for all but a finite number of states $\underline{m}$. To begin, consider the states of the form $\underline{m} = (0, 0, ..., 0, m_{S''}, 0, ..., 0)$, where all but a single one of the 2^n variables $m_S = 0$. As in Section 3.5, each of these states will provide the individual cut-set bounds on λ required for stability by the

theorem. For $\underline{m}$ of this form, Equation (3.12) reduces to

$$\begin{aligned}
&\lambda \sum_{S \supseteq S''} N_{|S|} \left(1+\delta\right)^{m_{S''}} + \lambda \sum_{S \not\supseteq S''} N_{|S|} \\
&< \sum_{S \supseteq S''} \sum_{\{S_2 | S_2 \supset S'', S_2 \not\subseteq S\}} \left(\sum_{i \in S''} \prod_{j \in S_2 / S''} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) N_{|S|} \left(1+\delta\right)^{m_{S''}-1} \\
&+ \sum_{S \supseteq S''} \sum_{i \in S''} \mu_{id} N_{|S|} \left(1+\delta\right)^{m_{S''}-1}
\end{aligned} \tag{3.13}$$

As long as $N_{|S''|}$ is chosen such that

$$N_{|S''|} + \sum_{S \supset S''} N_{|S|} < N_{|S''|} \frac{N+1}{N}$$

which is equivalent to the requirement of Equation (3.5).

We can replace Equation (3.13) with

$$\begin{aligned}
&\lambda N_{|S''|} \frac{N+1}{N} (1+\delta) + \lambda \sum_{S \not\supseteq S''} N_{|S|} \left(1+\delta\right)^{-m_{S''}+1} \\
&< \sum_{S_2 \supset S''} \left(\sum_{i \in S''} \prod_{j \in S_2 / S''} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) N_{|S''|} \\
&+ \sum_{i \in S''} \mu_{id} N_{|S''|}.
\end{aligned} \tag{3.14}$$

Because the left-hand side has been increased and the right-hand side has been decreased, satisfying Equation (3.14) assures that Equation (3.13) holds.

Rearrange Equation (3.14) as

$$\begin{aligned}
&\lambda \frac{N+1}{N} (1+\delta) + \lambda \frac{1}{N_{|S''|}} \sum_{S \not\supseteq S''} N_{|S|} \left(1+\delta\right)^{-m_{S''}+1} \\
&< \sum_{i \in S''} \left(\mu_{id} + \sum_{S_2 \supset S''} \left(\prod_{j \in S_2 / S''} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) \right) \\
&= \sum_{i \in S''} \left(1 - \epsilon_{id} + \epsilon_{id} \sum_{S_2 \supset S''} \left(\prod_{j \in S_2 / S''} \mu_{ij} \prod_{\{j | j \notin S_2, j \neq d\}} \epsilon_{ij} \right) \right)
\end{aligned} \tag{3.15}$$

and recognize that for any particular S_2, the product in the inner parentheses represents the probability that a packet transmitted from a particular node i remains unerased at exactly the subset S_2/S'' of possible receiver nodes. As S_2 runs over all possible supersets of S'' that contain at least one node, the cut-set bound $C(S'')$ is recovered by definition. Observe, for example, that if $A = \{1, 2, ..., n\}$ and $0 \leq a_j \leq 1$, then

$$\sum_{A_1 \subseteq A} \left(\prod_{j \in A_1} a_j \prod_{j \notin A_1} (1 - a_j) \right) = 1 \tag{3.16}$$

Choose $m_{S''}$ such that

$$\sum_{S \not\supseteq S''} N_{|S|} (1 + \delta)^{-m_{S''}+1} < \frac{N+1}{N}(1+\delta)^2 - \frac{N+1}{N}(1+\delta)$$

and Equation (3.15) reduces to $\lambda < \frac{N}{N+1} \frac{1}{(1+\delta)^2} C(S'')$.

Thus, it has been shown that for states of the form $\underline{m} = (0, 0, ..., 0, m_{S''}, 0, ..., 0)$ whenever $\lambda < \frac{N}{N+1} \frac{1}{(1+\delta)^2} C(S'')$, there exists a $\tilde{m}_{S''}$ sufficiently large such that the expected value of the Lyapunov function is decreasing for $m_{S''} > \tilde{m}_{S''}$. Each of the required cut-set bounds on λ for all of the different cuts $S \in \mathcal{S}$ are obtained in this manner.

It remains to show that there are only a finite number of general states $\underline{m} = (m_0, m_1, ..., , m_S, , ..., m_{2^n-1})$ where the expected value of the Lyapunov function is increasing.

To do so, first examine the state variable m_{2^n-1}; that is, the variable counting the number of packets which appear at every node in the system other than the destination. We will show that there exists a finite $m^*_{2^n-1}$ for which, as long as $m_{2^n-1} > m^*_{2^n-1}$, regardless of the value of m_0, m_1, and every other state variable up to m_{2^n-2}, the expected value of the Lyapunov function Equation (3.12) will be decreasing.

Let $\hat{S}$ be the subset $\hat{S} \in \mathcal{S}$ which contains all n relay nodes and the source node s, i.e. the largest subset of the nodes. Also, let $N_{|\hat{S}|} = 1$. Equation (3.12) can be

rewritten as

$$\lambda\left(1+\delta\right)^{m_{\hat{S}}+\sum_{S'\subset\hat{S}}m_{S'}}+\lambda\sum_{S\subset\hat{S}}N_{|S|}\left(1+\delta\right)^{\sum_{S'\subseteq S}m_{S'}}$$
$$<\sum_{S_1\subseteq\hat{S}}\left(\sum_{i\in S_1}\frac{m_{S_1}}{q(i)}\mu_{id}\right)\left(1+\delta\right)^{m_{\hat{S}}+\sum_{S'\subset\hat{S}}m_{S'}-1}$$
$$+\sum_{S\subset\hat{S}}\sum_{S_1\subseteq S}\left(\sum_{i\in S_1}\frac{m_{S_1}}{q(i)}\mu_{id}\right)N_{|S|}\left(1+\delta\right)^{\sum_{S'\subseteq S}m_{S'}-1} \tag{3.17}$$

(by ignoring the second line of Equation (3.12) and breaking the third line into two parts) since there can be no transitions from $\hat{S}$ to a larger subset. By dividing Equation (3.17) through by

$$\left(1+\delta\right)^{m_{\hat{S}}+\sum_{S'\subset\hat{S}}m_{S'}-1},$$

increasing the second term on the left-hand side, and decreasing the right-hand side by dropping the final term, the constraint

$$\lambda(1+\delta)+\lambda(1+\delta)^{1-m_{\hat{S}}}\sum_{S\subset\hat{S}}N_{|S|} \tag{3.18}$$
$$<\sum_{S_1\subseteq\hat{S}}\sum_{i\in S_1}\frac{m_{S_1}}{q(i)}\mu_{id}$$
$$=\sum_{i\in\hat{S}}\sum_{\{S_1|i\in S_1\}}\frac{m_{S_1}}{q(i)}\mu_{id}=\sum_{i\in\hat{S}}\mu_{id}=C(\hat{S}) \tag{3.19}$$

is obtained.

Therefore there exists a $m^*_{\hat{S}}$ for which, for all states $\underline{m}$ with $m_{\hat{S}}>m^*_{\hat{S}}$, regardless of the values of the other state variables, the expected value of the Lyapunov function will be decreasing when $\lambda<C(\hat{S})\frac{1}{(1+\delta)^2}$.

Next, consider any set $\hat{\hat{S}}$ which contains all but one of the n relay nodes. Define

$$C_i(S)=1-\prod_{j\notin S}\epsilon_{ij}$$

for pairs (i,S) such that $i\in S$. $C_i(S)$ represents the contribution to the cut-set bound for the cut S from the node $i\in S$. Use the same logic as in Equations (3.15

and 3.16) to note that for a fixed set S and fixed node $i \in S$,

$$\sum_{\{(S_1,S_2)|S_1 \subset S_2, S_1 \subseteq S, S_2 \not\subseteq S\}} \left(\sum_{i \in S_1} \frac{m_{S_1}}{q(i)} \prod_{j \in S_2/S_1} \mu_{ij} \prod_{j \notin S_2} \epsilon_{ij} \right) + \sum_{S_1 \subseteq S} \frac{m_{S_1}}{q(i)} \mu_{id}$$
$$= \sum_{\{S_1|S_1 \subseteq S, i \in S_1\}} \frac{m_{S_1}}{q(i)} C_i(S).$$

Use this simplification to rewrite Equation (3.12) as

$$\lambda \sum_{S \in \mathcal{S}} N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'}}$$
$$< \sum_{S \in \mathcal{S}} \left(\sum_{i \in S} \frac{\sum_{\{S_1|S_1 \subseteq S, i \in S_1\}} m_{S_1} C_i(S_1)}{q(i)} \right) N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1} \tag{3.20}$$

and since whenever $S_1 \subseteq S$, $C_i(S_1) > C_i(S)$, satisfying

$$\lambda \sum_{S \in \mathcal{S}} N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'}}$$
$$< \sum_{S \in \mathcal{S}} \left(\sum_{i \in S} \frac{\sum_{\{S_1|S_1 \subseteq S, i \in S_1\}} m_{S_1}}{q(i)} C_i(S) \right) N_{|S|} (1+\delta)^{\sum_{S' \subseteq S} m_{S'} - 1} \tag{3.21}$$

will assure that Equation (3.20) holds.

Assume that $m_{\hat{S}} < m^*_{\hat{S}}$. If $m_{\hat{S}} > m^*_{\hat{S}} \frac{2^n}{\delta}$, it will be shown that Equation (3.21) holds for any $\lambda < \frac{N}{N+1} \frac{1}{(1+\delta)^2} \min_{S \in \mathcal{S}} C(S)$.

There are two cases to consider:

- None of the m_S other than $m_{\hat{\hat{S}}}$ are greater than $m_{\hat{\hat{S}}} \frac{\delta}{2^n}$.
- At least one of the m_S is greater than $m_{\hat{\hat{S}}} \frac{\delta}{2^n}$.

In the first case, divide all the sets S into two classes: $\mathcal{S}_1 = \{S | \hat{\hat{S}} \subseteq S\}$ and $\mathcal{S}_1^C$. For $S \in \mathcal{S}_1$,

$$\frac{\sum_{\{S_1|S_1 \subseteq S, i \in S_1\}} m_{S_1}}{q(i)} > \frac{m_{S''}}{m_{S''} + \delta m_{S''}} = \frac{1}{1+\delta},$$

so term by term of the outer summation of Equation (3.21), when $\lambda < \frac{N}{N+1} \frac{1}{(1+\delta)^2} C(S)$, the terms in the class $\mathcal{S}_1$ are satisfied.

Now consider any of the terms in the class $S \in \mathcal{S}_1^C$. Divide both sides of Equation (3.21) by

$$m_{sum} = (1+\delta)^{\sum_{S' \subseteq \hat{S}} m_{S'} - 1}$$

and note that m_{sum} is greater than the sum in the exponent of the $(1+\delta)$ for any the terms $S \in \mathcal{S}_1^C$ by more than $m^*_{\hat{S}}$. The contribution to the left hand side from these terms becomes arbitrarily small, just as in Equation (3.18). Equation (3.21) is thus satisfied when none of the m_S other than $m_{\hat{S}}$ are greater than $m_{\hat{S}} \frac{1}{2^n}$.

In the case where at least one of the other m_S is greater than $m_{\hat{S}} \frac{\delta}{2^n}$, define

$$S_U = \bigcup_{\{S | m_S > m_{\hat{S}} \frac{\delta}{2^n}\}} S$$

as the union of all these sets S. The same analysis holds from the above case: For each term $S \in \mathcal{S}$ of Equation (3.21), either $S \supseteq S_U$ and the right hand side is greater than a $\frac{1}{1+\delta}$ fraction of $C(S)$, or the exponent of that term is more than $m^*_{\hat{S}}$ less than the exponent of the S_U term, and is thus inconsequential.

Define m^*_S so that whenever $|S_1| = |S_2| - 1$, $m^*_{S_1} > m^*_{S_2} \frac{2^n}{\delta}$. The same arguments already made are used inductively to show that as long as $m_S > m^*_S$ for at least one S, then Equation (3.21) is satisfied. □

3.7 Spatial Correlation of Dropped Packets

It is also possible to consider a model where the dropping of packets transmitted from each transmitter $i \in V$ are correlated events. That is, if the node i transmits a packet, the probability that exactly the set $W \subseteq V$ successfully receives that packet can be considered to be $p(i, W)$. In the independent model used in the majority of this chapter,

$$p(i, W) = \prod_{j \in W} \mu_{ij} \prod_{j \in W^C, j \neq i} \epsilon_{ij},$$

and

$$\sum_{W \in \mathcal{S}} p(i, W) = 1$$

just as observed in Equation (3.16).

The cut-set bound for the $S - S^C$ cut can still be interpreted as the sum of rates for which nodes in S can transmit and at least one node in S^C will successfully receive the packet:

$$C(S) = \sum_{i \in S} \sum_{\{W | W \cap S^C \neq \emptyset\}} p(i, W). \tag{3.22}$$

Replacing the transition probabilities in the Markov chain model with these $p(i, W)$ requires no substantive changes in the proof technique - the values of the cutset bounds $C(S)$ and the probabilities $C_i(S)$ change accordingly, and the proof of queue stability follows.

Allowing correlations across time for a single or multiple edges is a much different problem. An entire new set layers of the Markov chain would be required, and the whether the cut-set bound is achievable is still unknown in even the feedback-free model of [2]. Because of the asynchronous nature of our model, this same difficulty is encountered if it is desired to correlate erasures of packets from different transmitters. Such events would be simultaneous in the slotted-time model, and therefore are dealt with in [2], but would induce correlations over time in our model.

3.8 Conclusion

In this chapter, we have demonstrated a parallel between the erasure channel and a network of such channels: When acknowledgment feedback is available, there exists a simple transmission strategy by which the information-theoretic capacity (calculated by the cut-set bound) can be achieved for a unicast network without any need for a coding scheme. We have described a novel randomized and decentralized

strategy which requires only a surprisingly small amount of information about the network (specifically, no knowledge whatsoever about the network topology) to succeed in stabilizing the queues and achieving throughput optimality.

Chapter 4

Secrecy Capacity

This chapter is joint work with Emina Soljanin, of Alcatel-Bell Labs, T. Charles Clancy of the University of Maryland, and Andrew Mills, of the Computer Science Department of the University of Texas at Austin.

4.1 Introduction

In this chapter, we are concerned with the secrecy capacity of unicast communications over broadcast erasure networks with no interference in the presence of a wiretapper that has access to a certain number of network links of his choice. The key difference in this section from much prior work on secrecy capacity is that the model incorporates the broadcast nature of the wireless medium - the work in [21] studies the secrecy capacity of a broadcast channel, rather than that of a network of such channels. The wireless erasure network model captures the physical nature of the medium to a large extent, provides key insights into viable network protocols, and is much more tractable than the more general network information-theoretic setting that also includes interference. As we have pointed out before, the capacity of wireless erasure networks has been studied in significant depth, and for multiple settings it has been characterized in closed form [2, 18, 22, 23].

Just as the capacity of networks has been analyzed in different domains using different assumptions and notions of network throughput, secure communication over networks has been studied using multiple distinct assumptions and notions of secrecy. In particular, perfect secrecy capacity for the general class of information-theoretic channels has seen a resurgence of interest in recent years [21, 24, 25]. For a single source multicast networks implementing network coding, the problem

of making a linear network code information-theoretically secure in the presence of a wiretap adversary that can look at a bounded number, say k, of network edges was first studied by Cai and Yeung in [26]. They considered directed (V, E) graphs and demonstrated the existence of a code over an alphabet with at least $\binom{|E|}{k}$ elements which can support a secure multicast rate of up to $n - k$. They also showed that such codes can be designed in $\mathcal{O}(\binom{|E|}{k})$ steps. The required edge bandwidth and the secure code design complexity are main drawbacks of this pioneering work. Feldman *et al.* derived trade-offs between security, code alphabet size, and multicast rate of secure linear network coding schemes in [27], by using ideas from secret sharing and abstracting network topology. El Rouayheb and Soljanin showed that network security can be achieved by using the Ozarow-Wyner approach of coset coding at the source on top of the implemented network code [28]. Weakly secure network coding (which insures that only useless information rather than none is revealed to the adversary) was studied by Bhattad and Narayanan in [29], and practical schemes are missing in this case as well. Another approach was taken by Jain in [30] who obtained security by merely exploiting the topology of the network in question.

In this chapter, we deal with wireless erasure networks in an attempt to find a network generalization of the information-theoretic perfect secrecy analysis for the Wyner wiretap channel model [24] to the capacitated network secrecy capacity for unicast networks. In particular, we determine achievable strategies and upper bounds on secrecy capacity, which match for a class of broadcast-constrained erasure networks. We will also present example networks which fall outside of this class, demonstrating that the intuitive upper and lower bounds we describe are not sufficient to characterize the secrecy capacity of wireless erasure networks in general. Finding a general description of the secrecy capacity of wireless erasure networks may indeed be a problem with an elusive solution.

4.2 System Model

The network under investigation is a single-source, single-destination, lossy, wireless packet network, modeled as a directed acyclic graph (V, E) with nodes V and communication links E. An example is illustrated in Figure 4.2. The "wireless" component of the network is manifested in a broadcast constraint; that is, each transmitter must send a single, identical packet on every communication link exiting that node in any given time-slot. These broadcasted packets are all taken to be symbols of a finite field transmit alphabet $GF(q)$ for some (large) q. The network is "lossy" because each edge in the graph experiences packet drops, or equivalently, symbol erasures. These erasures are independent across both time and space, and associated with each directed edge (i, j) between nodes i and j is the value of that erasure probability, denoted by ϵ_{ij}. We assume that each receiver obtains all of the symbols along its incoming edges without interference. The capacity of this non wire-tapped network model was first given in [2] (and its achievablilty alternatively demonstrated by a random linear network coding scheme in in [18]). We again provide the expression for convenience:

$$C = \min_{S} \lg q \sum_{i \in S} \left(1 - \prod_{j \in S^C} \epsilon_{ij} \right) \tag{4.1}$$

where S is a vertex-cut. Specifically, S is any subset of the nodes which contains the source and not the destination. We now extend the work on this notion of a lossy wireless network to demonstrate its secrecy capacity.

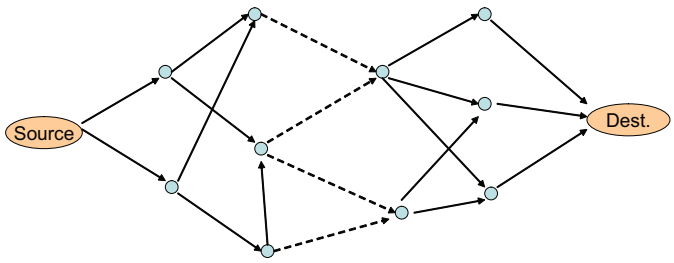

Figure 4.1: System Model - Dotted Edges Represent an Example Edge Cut-set.

Along the lines of the model in [26], the eavesdropper has access to *any* k edges of this network, in the sense that it can observe the outputs of each edge in the network. Specifically, if the packet transmitted along any wire-tapped edge is erased (or dropped), the wire-tapper as well as the receiving node fails to receive it. Our objective is to ensure perfect secrecy of the message in the network from the wire-tapper, defined in the sense of [24]. Specifically, we wish to determine the highest rate possible such that the wire-tapper gains no information about the message being communicated, and that the mutual information between the wire-tapper's information and the source's intended message is zero.

Let $s \in V$ denote the source and $d \in V$ the destination. We define an edge cut as any set $T \subset E$ such that there does not exist a path from s to d in $E \setminus T$. We assume a path from s to d exists in E, so a cut must be nonempty. Also, note that E itself defines a cut in this network. The definition of an edge cut in this chapter is distinct from that used in [2] and the rest of this document, which is a vertex-based definition of a "cut" in the network.

Letting $A \subseteq E$ be a set of edges of the graph, we define a size $m \times n$ incidence matrix. The number of columns n will be equal to the number of distinct parent vertexes of the edges in A. That is, $n = |I_A|$, where $I_A = \{i | (i,j) \in A\}$. The number of rows m in a network with no interference will be equal to the number of edges in A, that is $m = |A|$. For each time slot t, the incidence matrix $G_A(t)$ will contain a 1 in each row corresponding to an edge whose symbol was not dropped, in the column corresponding to the corresponding parent vertex. Let X_i denote the symbol transmitted from node i, $Y_{j,i}$ denote the symbol received at node j which had been transmitted from node i, and γ_{ij} be independent Bernoulli random variables with $P[\gamma_{ij} = 0] = \epsilon_{ij}$. Then, for the network displayed in Figure 4.2, the matrix

$$\begin{bmatrix} Y_{2,s} \\ Y_{2,1} \\ Y_{d,1} \end{bmatrix} = \begin{bmatrix} \gamma_{s2} & 0 \\ 0 & \gamma_{12} \\ 0 & \gamma_{1d} \end{bmatrix} \begin{bmatrix} X_s \\ X_1 \end{bmatrix}$$

defines G_A for the cut $A = \{(s,2),(1,2),(1,d)\}$. Note that when A is chosen to be the set of edges crossed by the vertex min-cut, then the expected value of G_A

corresponds precisely to the value of the min-cut capacity given in Equation (4.1).

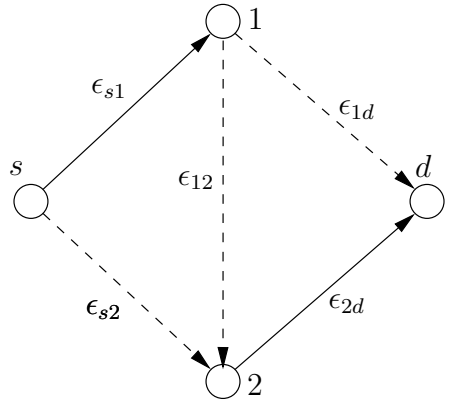

Figure 4.2: Example Network and Edge Cut Set

4.3 Upper Bound on Secrecy Capacity

Example: Consider for a moment a simple unicast scenario shown in Fig. 4.4. Assume that the source is directly connected to the destination through n network

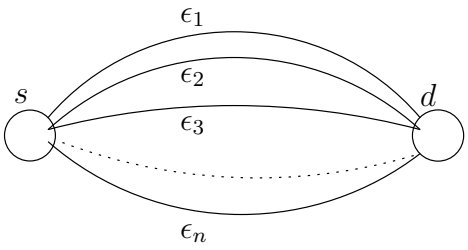

Figure 4.3: Example Network and Edge Cut Set

Figure 4.4: At the source node, k information symbols are encoded into h coded symbols, which are then simultaneously transmitted to the destination.

links where link i, $1 \leq i \leq n$ has erasure rate ϵ_i, and that any k of these edges can be accessed by a wiretapper. If we assume that all erasure rates are equal, that is $\epsilon \triangleq \epsilon_1 = \ldots \epsilon_n$, then this network unicast is equivalent to the Wyner wiretap channel model in which the intended user observes the output of an erasure channel

with the erasure rate ϵ^n, and the wiretapper observes the output of a degraded channel with the erasure rate ϵ^k.

Outer Bound: Let S be any set of k edges in the network. Let $T : S \subseteq T \subseteq E$ be a cut in the network. Then, the secrecy capacity of the network is bounded by:

$$C \leq \lg q \left(\mathbb{E}[rank(H_T)] - \mathbb{E}[rank(G_S)] \right) \tag{4.2}$$

where H_T is the incidence matrix of the cut T and G_S the incidence matrix of S.

Specifically, secrecy capacity for the network is upper bounded by:

$$C \leq \lg q \min_{T:\ T \text{ is a cut}} \Big[\mathbb{E}[rank(H_T)] - \max_S \mathbb{E}[rank(G_S)] \Big]$$

In the notation of [2], this expression is equivalent to

$$C \leq \min_A \lg q \sum_{i \in A} \left(\prod_{j: j \in A^C, (i,j) \in S} \epsilon_{ij} \right) \left(1 - \prod_{j: j \in A^C, (i,j) \notin S} \epsilon_{ij} \right) \tag{4.3}$$

where, to be clear, A refers to a vertex cut and S refers to the set of edges to which the wire-tapper has access.

Proof: Wire-tappers are placed on any k edges of the graph (V, E), forming the subset $S \subseteq E$. Consider any edge cut, and if necessary, supplement it with the edges in S to form the edge cut $T \supseteq S$ with inputs $X^n(T)$ and outputs $Y^n(T)$.

Intuitively, in this setting, our goal is to upper bound the secrecy capacity of the network by a cutset bound: The amount of information that the receiver can get while the wire-tapper still has no information about the source should be no more than the rate that the network can get across any cut, minus the amount that the wire-tapper can see on that same cut.

$$
\begin{aligned}
nR_e &\overset{(a)}{\leq} H(W|Y(S)^n) \\
&\overset{(b)}{\leq} H(W|Y(S)^n) - H(W|Y(T)^n) + n\epsilon_n \\
&= I(W;Y(T)^n) - I(W;Y(S)^n) + n\epsilon_n \\
&= H(Y(T)^n) - H(Y(T)^n|W) - H(Y(S)^n) \\
&\qquad + H(Y(S)^n|W) + n\epsilon_n \\
&\overset{(c)}{=} H(Y(R)^n|Y(S)^n) - H(Y(R)^n|W, Y(S)^n) + n\epsilon_n \\
&\leq H(Y(R)^n|Y(S)^n) \\
&\qquad - H(Y(R)^n|X(T)^n, W, Y(S)^n) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} (H(Y(R)_i|Y(S)_i) \\
&\qquad - H(Y(R)_i|X(T)^n, W, Y(S)^n, Y(R)^{i-1}))
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(d)}{=} \sum_{i=1}^{n} (H(Y(R)_i|Y(S)_i) - H(Y(R)_i|X(T)_i, Y(S)_i)) \\
&= \sum_{i=1}^{n} (I(X(T)_i; Y(T)_i) - I(X(T)_i; Y(S)_i)) \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} (I(X(T)_i; Y(T)_i) - I(X(S)_i; Y(S)_i)) \\
&\leq \sum_{i=1}^{n} \max_{p(X(T)_i)} (I(X(T)_i; Y(T)_i) - I(X(S)_i; Y(S)_i)) \\
&= n \max_{p(X(T))} (I(X(T); Y(T)) - I(X(S); Y(S)))
\end{aligned}
$$

where

(a) follows from the definition of secrecy rate

(b) follows from Fano's inequality [31]. All the information must be retrieved from the final destination node with high probability, and the Data Processing Inequality applies since $W \rightarrow T \rightarrow\ Y_d$ is a Markov Chain.

In (c), R is defined as $T \setminus S$ so that $X(T) = (X(R), X(S))$ (d) follows from the

fact that, given $X(T)_i$, $Y(S)_i$ is conditionally independent of W and $Y(S)_{i-1}$. (e) follows because $X(S)_i$ is a degraded version of $X(T)_i$.

Intuitively, for every cut $T \supseteq S$, The network behaves as an information-theoretic wiretap channel with a (physically) degraded wire-tapper. Note that in the (symmetric) erasure network setting, it is easy to show that the optimum input distribution is uniform (straightforward extension of result in [2]), and thus the upper bound reduces to the difference between the max-rate across the cut T and the max-rate across the subset of nodes S. From [16], the rate across any subset of nodes is given by the expected rank of its incidence matrix. This gives us the result. This converse is general, and holds for some additional models of erasure networks. Specifically, in a network with additive finite-field interference at the receivers (for example, [13] or [16], or an extension of [32]) then the cut-capacity still evaluates to the expected rank of the incidence matrix.

4.4 Achievability for a Special Class

In this section, we show that the outer bound given by (4.2) can be achieved under certain conditions. Specifically, let T^* be the cut that minimizes $\lg q\mathbb{E}(rank(H_{T^*}))$, i.e, T^* is the cut that minimizes the (non-wiretap) capacity of the network. Let S^* be a $k-$edge subset of T^* such that the upper bound given by (4.2) is minimized for T^*. [1] Define

$$C^* \triangleq \lg q\mathbb{E}(rank(H_{T^*})) - \mathbb{E}(rank(G_{S^*}))$$

Our goal in this section is to show that in a broadcast-constrained erasure network, if T^{**} is the minimizing cut for (4.2) and $T^{**} = T^*$, then C^* is the secrecy capacity of the network.

Achievability: Given that the system has k wire-tappers placed arbitrarily on the cut T^* as defined above, a (linear) encoding scheme exists such that C^* is

[1] Note that T^* may not be the minimizing cut for (4.2) of this graph.

achievable with a perfect secrecy constraint.

Proof: The key idea, similar to that in [27] and [33], is to send information packets and "noise" packets such that the legitimate receiver can decode both the information and noise packets, while the wire-tapper can only decode the "noise" packets. The noise packets are independent of information packets thus guaranteeing perfect secrecy.

Note that two coding schemes exist that achieve the cut-set upper bound on capacity of broadcast-constrained erasure networks (in a non-wiretap setting). The first is based on random coding arguments in [2], while the second is based on random linear network coding in [2, 13, 18]. Note that in [2], but not necessarily in [13, 18], it is assumed that the exact erasure locations at the intermediate nodes is known to the receiver. In the rest of this document, we will utilize the linear network coding framework for coding at the intermediate nodes in [2]. Note that our proof for secrecy capacity can be modified to the case when random network coding is used [2].

Let n be a positive integer. Define $m \triangleq n\mathbb{E}(rank(H_{T^*}))$ and $l \triangleq n\mathbb{E}(rank(G_{S^*}))$. Then $C^* = \frac{m-l}{n} \lg q$. Intuitively, we show that if m un-erased packets reach the legitimate destination and l un-erased packets reach the wire-tapper, then there exists a coding scheme that achieves C^*.

The encoding scheme is as follows: A transmit vector of length m packets for some $\epsilon > 0$ is constructed consisting of $(m - l - n\epsilon)$ information (or data) packets and $(l + n\epsilon)$ noise packets that are independent of the data source and chosen uniformly from $GF(q)$. We refer to this as the vector x^m, with $a^{(m-l-n\epsilon)}$ denoting data packets, $b^{l+n\epsilon}$ denoting "noise" packets, and $x^m = [a^{(m-l-n\epsilon)} \;\; b^{(l+n\epsilon)}]$.

A *linear* encoding scheme is used at each node, which results in $n \times m$ transfer matrices M_n and F_n for the legitimate destination and the wire-tapper respectively. By the weak law of large numbers (which results in the notion of strong typicality as in [31]), the received vector at the legitimate destination is at least an $(m - n\epsilon)$-sized subset of the n-length vector $M_n x^m$ with probability greater than $1 - \delta$.

Similarly, the wire-tapper observes at most an $(l + n\epsilon)$-ary subset of the vector $F_n x^m$ with high probability.

In [2, Section 7-A], the authors show that averaging over all matrices, the average probability of error in decoding can be made less than ϵ. Therefore, the information vector $a^{m-l-n\epsilon}$ is received successfully with high probability at the destination allowing for a rate arbitrarily close to $C^* - \epsilon$.

Let the $(l + n\epsilon)$-ary subset of the vector $F_n x^m$ received at the wire-tapper with high probability be denoted as:

$$z^{l+n\epsilon} \triangleq F_{l+n\epsilon} x^m$$

Define

$$\hat{z}^{(l+n\epsilon)} \triangleq \hat{F}_{(l+n\epsilon)} b^{(l+n\epsilon)}$$

where $\hat{F}_{(l+n\epsilon)}$ is a $(l+n\epsilon) \times (l+n\epsilon)$ matrix of the last $(l+n\epsilon)$ columns from $F_{l+n\epsilon}$.

A key step is for us to choose coefficients for the network (i.e., one particular M_n and F_n) such that:

- $a^{(m-l-n\epsilon)}$ is decodable at the legitimate destination with probability of error arbitrarily small, and
- $\hat{F}_{l+n\epsilon}$ is invertible.

Note that a random choice in coefficients ensures both of these simultaneously with high probability [2, 34], and therefore a particular set of M_n and F_n exist that satisfy these requirements.

By Fano's inequality [31], the information recovered by the wire-tapper is upper bounded by:

$$
\begin{aligned}
I(a^{m-l-n\epsilon}; z^{l+n\epsilon}) &= H(z^{l+n\epsilon}) - H(z^{l+n\epsilon}|a^{m-l-n\epsilon}) \\
&\overset{(a)}{=} H(z^{l+n\epsilon}) - H(\hat{z}^{l+n\epsilon}|a^{m-l-n\epsilon}) \\
&\overset{(b)}{=} H(z^{l+n\epsilon}) - H(\hat{z}^{l+n\epsilon}) \\
&\overset{(c)}{=} H(z^{l+n\epsilon}) - (l+n\epsilon)\lg q \\
&\overset{(d)}{\leq} (l+n\epsilon)\lg q - (l+n\epsilon)\lg q \\
&= 0
\end{aligned}
$$

where:

(a) follows from the definition of $\hat{z}^{(l+n\epsilon)}$ and the property of entropy.

(b) follows from the independence between $a^{(m-l-n\epsilon)}$ and $b^{(l+n\epsilon)}$.

(c) follows from the invertibility of $\hat{F}_{(l+n\epsilon)}$, and that $b^{(l+n\epsilon)}$ are i.i.d. and uniform.

Or in essence, no information is recovered by the wire-tapper. This concludes the achievability proof.

4.5 Counter-Example to the Upper Bound

Counter to the authors' intuitive expectations, the upper bound of Section 4.3 is not tight in general. This section provides a counter example demonstrating this fact.

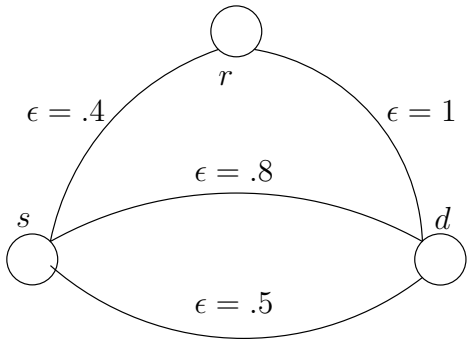

Figure 4.5: Upper Bound Counter Example

Consider the three node network shown in Figure 4.5. This network has two independent edges which connect the source to the destination. The wiretapper may choose any one edge to wiretap, but it is clear that the set consisting of the

edge $S = \{sr\}$ is the most favorable. Note that this set S does not lie in the minimum cut edge set of the original network $T_{min} = \{rd, sd_1, sd_2\}$, whose cut value is $1 - (.8)(.5) = .6$.

We evaluate along the cut $T = \{sr, sd_1, sd_2\}$,

$$\mathbb{E}[rank(H_T)] - \mathbb{E}[rank(G_S)] = (1 - (.4)(.5)(.8)) - .6 = .24$$

to find the minimum of the outerbounds on capacity from Equation 2.13. However, the actual secrecy capacity of this network is equal to zero.

The wiretapper receives the data transmitted by the source, with a loss of 40% of all the symbols. The destination will also receive the data transmitter by the source, again with a loss of 40% of the symbols. Therefore, whatever the destination d can decode, the wiretapper will also be able to decode, and there is no secrecy capacity.

4.6 An Alternate Achievability and Counter-Example

We had conjectured that the capacity described as in Equation (4.3), was in all cases achievable, through a strategy more standard to the information-theoretic community, that of binning [21]. Intuitively, the strategy is equivalent to that described in Section 4.4: Provide more random, noisy information to the destination, in an amount equivalent to the amount of information able to be decoded by any wiretapper, so that the desired message remains completely unknown. Specifically, operate the network as follows: Follow the coding strategy of [2], creating codebooks for the $2^{nR_{tot}}$ messages using n channel uses, for any sufficiently large n and R_{tot} less than the untapped network's capacity. Bin the $2^{nR_{tot}}$ messages randomly into $2^{nR_{sec}}$ bins, and the encoding strategy to communicate at a rate of R_{sec} is to randomly choose any message in the appropriate bin, and transmit using the original codebook. The destination node will be able to decode the expanded message, and therefore be able to determine in which bin it lies.

We then need to compute at what rate R_{sec} we can bin the messages so that a

wiretapper cannot gain any information about which bin the true expanded message was chosen from. The wiretapper can only attempt to decode the message using the same strategy as the final destination does it [2]: Assuming the wiretapper knows all of the erasure locations in the network and all of the relay nodes' codebooks, simulate the progress each of the possible messages through the network. The sender must guarantee that the wiretapper finds, to within a small factor, an equivalent number of possible messages in every bin. By the same arguments of distinguishable and indistinguishable messages used in both [2] and Chapter 2, this will be the case as long as $R_{sec} < R_{tot} - R_{sec-cut}$, where $R_{sec-cut}$ is the cutset for the minimum cut *separating the source and all of the wiretapped edges.*

We had conjectured that the difference $R_{tot} - R_{sec-cut}$ was equal to the expression in Equation (4.3), and would therefore be the capacity of the network. However, we show in this second counterexample of Figure 4.6 that only binning at the source is not a good enough strategy to achieve the network secrecy capacity.

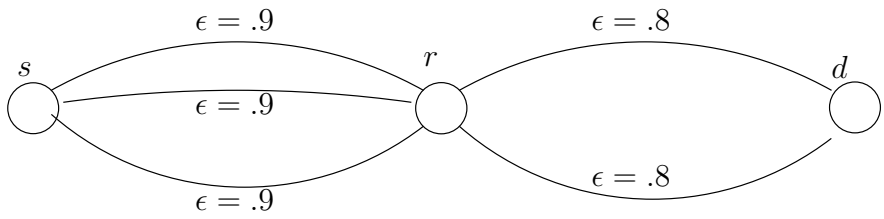

Figure 4.6: Counter Example: Source-side Binning is Insufficient to Achieve Secrecy Capacity

The untapped capacity of the wireless erasure network of Figure 4.6 is equal to the minimum of $(1 - .9^3) = .271$ and $(1 - .8^2) = .36$, or .271. We will assume that the wiretapper can only choose one edge to tap. If he chooses one of the edges between the source and the relay, then $R_{sec-cut} = .1$; however, if he chooses one of the edges between the relay and the destination, his minimum cut (and therefore $R_{sec-cut}$) is equal to .2. By the strategy of source-side binning, a rate of .071 is thus achieved.

We will demonstrate, however, that there exists a strategy which can achieve a secrecy capacity of .16, which is the value obtained by the application of Equation (4.3). This strategy consists of re-binning the messages at the relay node.

The source should encode messages at the source at a rate of $R_{tot} = .26$ and bin them to achieve $R_{sec} = .16$. Since a wiretapper at the intermediate node can only see a rate of $R_{sec-cut}$ of .1, perfect secrecy is assured to the relay node.

The relay node will have a codebook designed for the rate of $R_{tot} = .36$, and again bin these messages with $R_{sec-cut} = .2$ redundancy so that again, a rate of $R_{sec} = .16$ can be transmitted from relay to destination with perfect secrecy, as a wiretapper on one of these edges can only see a maximum rate of .2. So a secrecy rate of .16, which is the maximum, by our upper bounds, is indeed achievable on the network of Figure 4.6.

We have demonstrated in this section that source-side binning alone is therefore unable to achieve the capacity of every wireless erasure network.

4.7 Conclusions

This work investigates the secrecy capacity, or specifically the equivocation rate, of a wire-tapped wireless erasure network. Secrecy capacity for this network takes the same intuitive form as it does in many other network models: we desire to maximize the difference in the amount of data that the receiver can interpret and the amount of information from the source the wire-tapper could receive. In a simple, error and interference-free network, this result has the straightforward interpretation of subtracting the number of wire-tapped edges from the number of edges in the minimum cut. For the wireless erasure network, the subtraction is equivalent, given the modified cut-set bound, for a few special cases only. We have shown that the upper-bound is achievable when the wire-tappers are chosen from a specific subset of the nodes, that is, along the minimum cut of the original network. The strategy for this network class is interesting because of its history in the secrecy context, and its application to a non-traditional secrecy model.

Because there are networks where the intuitive min-cut max-flow upperbound is not achievable, and networks where the intuitive source-side binning strategy is sub-optimal, the general problem of secrecy capacity in wireless networks seems unlikely to be easily solvable.

Chapter 5

Transport Capacity

In general, finding the capacity region for multi-terminal networks is an open problem. Even simple cases are difficult to solve, and the exact capacity region is known only in a few special cases, such as the multiple-access channel and degraded broadcast channel. Wireless erasure networks, first introduced by Dana, Gowaikar, and Hassibi [2], as we have seen in Chapters 2 and 3, constitute another class of networks where a precise characterization of capacity is possible. The authors of [2], and our additions to their model, determine the capacity for any single pair of source and destination nodes, as well as the multicast capacity. Wireless erasure networks are also practical: most ARQ systems with packet errors act like erasure channels, with either successful or completely unsuccessful delivery of each packet.

While the exact capacity of the single-source wireless erasure network is known in the unicast and multicast scenarios, current research can only bound the capacity region for more general networks with many nodes and multiple sources and destinations. The alternate notion of transport capacity provides a descriptive and convenient scalar quantity that captures much of the overall network's behavior.

We assume the erasure probability over each channel is a function of the physical distance between the transmitting and receiving nodes. Intuitively, nodes spaced far apart from each other should have a difficult time communicating, thus an increase in the erasure probability should occur with increasing distance. This concept parallels the models used in [5] and [6], where the received power from a wireless source decays according to a polynomial power-law or exponentially with increasing distance. The authors of [6] use minimum-cut maximum-flow arguments to upper-bound the transport capacity of a network with Gaussian channels, a

technique we will also follow.

Adding in receive-side as well as transmit constraints allows us to capture both the broadcast nature of the wireless medium and the influence of interference in a wireless network.

In our setting, we allow each transmitting node to broadcast one symbol across all of its outgoing edges (so-called broadcast constraint on each transmitting node). We consider two disparate settings on the receive-side: One with no interference among nodes, and the other with finite-field additive interference at the receivers. The reasons these two models are of great interest are as follows:

- An upper bound on throughput under the "no-interference model" provides an upper bound on throughput for a class of wireless channels with interference. If X_i, $1 \leq i \leq m$ represent the input to the channel, settings where the interference can be decomposed into *separate* erasures on each link producing Y_i, $1 \leq i \leq m$ followed by an interference mapping $Z = g(Y_1, \ldots, Y_m)$ with Z being the channel output, the no-interference case provides an outer bound on throughput. This is because the receiver in the no-interference case can "mimic" the interference function g, thus making the maximum throughput for the interference case less than that of the no-interference case.

- The finite field additive interference is intuitively a pessimistic interference model. This is because interference is traditionally thought to increase the total received power (or equivalently, for finite field inputs $X_i \in F$, the channel output Z belongs to a field with alphabet size larger than that of F). Thus, restricting the output to belong to the same (finite) field as each of the inputs represents a "stringent" interference requirement.

Thus, the no-interference and finite-field interference represent an optimistic and a pessimistic extreme respectively, with many other interference settings lying in between these two settings.

5.1 Prior Work

Transport capacity, introduced in [5], was first used to describe the asymptotic capabilities of a Gaussian interference network on a fixed area, as the number of nodes in the network n (and therefore the node density of the network) scales larger and larger. Transport capacity is defined as the maximum, over all possible sets of feasible rate vectors, of the distance-weighted rate-sum, i.e.

$$T = \sup_{feasible\{R_1,\ldots,R_L\}} \sum_{l=1}^{L} R_l d(l) \tag{5.1}$$

where the supremum is taken over all possible sets of source-destination pairs, there are L source-destination pairs, $d(l)$ be the distance between the source node s_l and destination node t_l, and R_l is the rate associated with the l^{th} pair.

Making transport capacity meaningful requires that increasing physical distance separating two nodes to have a degrading effect on the communication ability between those nodes. Many authors [5],[6],[7],[8] invoke a polynomially or exponentially decaying power law to accomplish this.

Under the outage capacity and routing-only constraints of the model, [5] showed that the total transport capacity of the Gaussian interference network is upper-bounded as $O(\sqrt{n})$, and that at least $\Omega(\sqrt{n}/\lg n)$ growth is achievable by routing alone. Franceschetti et. al [7] sharpen the lower bound and demonstrate that linear growth in the number of nodes (removing the $\lg n$ term) is achievable in a random network with an additive Gaussian noise model using routing alone.

In an alternative to the fixed area, dense network scaling of [5], some authors [6] [7], [8] study a network whose geographic expanse increases with the number of nodes, so that the density remains constant. Xie and Kumar [6] provide an information theoretic upperbound of linear transport capacity growth for a sufficiently high-attenuation power law. Finally, Tse et. al demonstrate in [8] a hierarchical cooperative MIMO scheme which achieves linear transport capacity growth for both the dense and expansive network cases, building upon a cooperative scheme introduced by Aeron and Saligrama [9].

We have investigated the capabilities of erasure networks from a transport capacity viewpoint, and present our results in the following sections. Recall that for the wireless erasure network without any interference, the cut-set bound on capacity is given by

$$R \leq \min_{S \in \mathcal{S}} \sum_{i \in S} \left(1 - \prod_{j \in S^C} \epsilon_{ij} \right). \tag{5.2}$$

or for the general erasure network with additive interference, by the minimum of the expected value of the rank of the transfer matrices for each source-destination cut.

5.2 Transport Capacity of Erasure Networks

In this chapter, a wide-variety of models for erasure networks are analyzed. The differences between the various models fall into the following different categories:

- **Expansive** networks (those with constant density as the number of nodes n increases versus **dense** networks (those with constant area as n increases).
- **Random** networks versus **arbitrary** networks. We shall prove, in general, a single identical upper bound on the rate of transport capacity growth for both networks, but demonstrating achievability in a network designed for maximizing transport capacity growth versus one where the sources and destinations are randomly assigned require different approaches.
- The **threshold**, **exponential**, and **polynomial** decay models are three different cases we study in relating the probability of an erasure to the physical distance between the transmitting and receiving nodes.
- While it is simple to show that the upper-bound on any wireless erasure network **without receiver interference** is also valid in the same wireless network **with additive finite-field interference**, achievability in the interference model requires substantial additional proof.

This work provides a linear upper-bound (in terms of the number of nodes n) to the transport capacity in both the wireless erasure network (broadcast constraint, but no receiver interference) and the broadcast erasure network with interference (a single transmit and single receive antenna at each node), under each of the three decay models for expansive and dense networks.

Further, the work demonstrates that in expansive random wireless erasure network with and without interference, routing alone (no network coding) can achieve this order-wise optimal linear bound.

It is shown that only in the dense wireless network with interference (with and without random source destination pairs) are the achievable lower and obtainable upper bounds not tight. We conjecture that the cut-set upperbound, used in various methods in all of these cases, is not strong enough to show the true capacity for the dense model with interference.

5.2.1 Models

In demonstrating the upper bound, we use three different models to describe how the probability of an unerased transmission decays with distance:

In the threshold model, over distances less than or equal to d^*, there is never an erasure, but for distances greater than d^*, an erasure will always occur. This model is motivated by straightforward signal-to-noise ratio considerations.

In the exponential model, the probability of successful transmission decays exponentially with increasing distance:

$$\epsilon\left(d\right) = 1 - \alpha^d = 1 - e^{-d/d^*},$$

where the decay parameter α satisfies $0 < \alpha < 1$. The exponential model is motivated by the error rates for coding over long blocks with a signal-to-noise ratio which decays with increasing distance according to a power-law.

In the polynomial model, the probability of successful transmission follows a

power-law decay:

$$\epsilon\left(d\right) = 1 - \frac{1}{1+d^{\beta}},$$

where $\beta > 0$.

Note that for all of the models, there is no erasure when $d = 0$ and a zero probability of successful transmission when $d = \infty$.

In all models, transmit output symbols (i.e. channel input symbols) are chosen from the alphabet F_q.

Because of the broadcasting requirement of the wireless medium, we can consider the network to be a complete graph: There exists some probability (non-zero, in all but the threshold model) that any node will successfully receive the symbol transmitted by any node in the network. Therefore, we no longer have a directed *acyclic* graph, as per the system models in Chapter 2. We do prove, however, that back-edges cannot increase the cut-set bound on capacity.

In the expansive network model, we make use of one of two assumptions in order to prove a linear upperbound. Either, a random placement is assumed, or we impose a minimum separation distance constraint on the node placement: For all nodes i and j, $d_{ij} \geq d_{min}$ for some d_{min} which is independent of n. A random placement, intuitively, assures that "not too many" nodes are within this minimum distance of each other.

In the wireless erasure network model, we make the same assumption of no receiver interference as [2]. For the broadcast erasure network with interference, we assume that in each timeslot, the node j receives the sum $Y_j(t)$, where

$$Y_j(t) = \sum_{i \neq j, i=1}^{n+1} h_{ij}(t) X_i(t) \tag{5.3}$$

and $h_{ij}(t)$ are independent (over both indices and timeslots) zero-one binary random variables that take the value 0 with probability ϵ_{ij}. When the receiver node j has knowledge of the states h_{ij}, this formulation provides the same information to a receiver in the broadcast erasure network with no interference as a model with

the additional output symbol "Erased" in the output alphabet: if $h_{ij} = 0$, then the receiver knows that the corresponding symbol was erased.

The network contains L source-destination nodes pairs (s_l, t_l). We desire to reliably decode each of the L independent information sources, each available to one of the source nodes s_l, at the corresponding destination node t_l at the rate R_l.

Let $d(l)$ be the distance between the source node s_l and destination node t_l. The transport capacity of a network is the supremum of the distance-weighted sum of reliable rates

$$T = \sup_{feasible\{R_1,\ldots,R_L\}} \sum_{l=1}^{L} R_l d(l) \tag{5.4}$$

where the supremum is taken over all possible sets of source-destination pairs.

5.3 Summary of Results

The following proofs will show

- For expansive networks with a geometric threshold, a polynomial decay ($\beta > 3$), or an exponential decay, a linear growth in transport capacity is both the upperbound and achievable (to a polylog factor) in arbitrary and random networks.
- For dense networks with a geometric threshold, a polynomial decay ($\beta > 3$), or an exponential decay exponential decay, a linear growth in transport capacity is an upperbound.
- For dense networks with no receiver interference, a linear growth in transport capacity is achievable in both arbitrary and random networks.

The sole case where upperbound and lowerbounds do not meet, in a asymptotic growth rate sense, are dense networks with interference.

All of the achievability strategies considered in this paper are routing-only strategies. That is, they do not require network-coding, that is, any combination of the

information within packets at any of the relay nodes. Forwarding of packets is sufficient. Thus for the settings under consideration ($\beta > 3$) only routing is required to provide the correct order of maximum throughput. That is, network coding in the case $\alpha > 3$ for an erasure network with broadcast constraints with or without receive side constraints can provide no more than a polylog factor improvement in performance.

5.4 Converse Techniques

In this section, the various different converse techniques that can be used to upperbound the transport capacity of these networks are described.

5.4.1 Threshold Model

The proof of linear growth in the geometric model is similar in spirit to that of the proofs in [6]. We allow n nodes to be distributed arbitrarily on the plane, subject only to the $d \geq d_{min}$ minimum distance constraint. Let d^* be the maximum distance over which transmissions are successful. We will examine an infinite number of cuts, vertical and horizontal, corresponding to the lines $y = id_{min}/2$ and $x = jd_{min}/2$. Any source-destination pair of distance d must cross at least d/d_{min} of these cuts. The cut-set bound on rates across each of the horizontal and vertical cuts (in either direction, i.e. from S to S^C and from S^C to S) will be denoted $R_v(i)$ and $R_h(j)$, respectively. We are now interested in the expression

$$\frac{\sqrt{5}}{2} d_{min} \sum_{i=-\infty}^{\infty} R_v(i) + \frac{\sqrt{5}}{2} d_{min} \sum_{j=-\infty}^{\infty} R_h(j) \tag{5.5}$$

which bounds the total transport capacity. The value of each cut is bounded by the total number of nodes which are within a distance d^* of that cut, so each node in the network can contribute $\lg q$ bits of rate to a maximum of $8d^*/d_{min}$ cuts (four cardinal directions, and the transmission can reach to at most $2d^*/d_{min}$ cuts in any of those directions). Any source-destination pair which crosses a cut can be separated by a distance of at most $\sqrt{5}/2$, so Equation (5.5) and the transport

capacity are thus bounded by $4\sqrt{5}nd^* \lg q$.

5.4.2 Information Theoretic Background for Converses

Each of the remaining converse proofs uses the information-theoretic notion of cut-set bounds to show the upperbound. Before the main converse proofs, we will prove and evaluate a general cut-set bound for the sum-rate of reliable information flow across a partition of the nodes.

Since we assume that each receiver j knows not only the symbol, but also the channel state for all the channels which terminate at node j, define $Y_j^*(t)$ to be the vector consisting of both $Y_j(t)$ (the received symbol or symbols at j)and all channel states $h_{ij}(t)$, for $i \in 1..n+1$.

Assemble the random variables into vector notation as follows: For any subset S of nodes, consider $X_S(t)$, $X_{S^C}(t)$, $Y_{S^C}(t)$ which are defined intuitively. Define matrices $H_1(t)$ and $H_2(t)$ such that

$$Y_{S^C}(t) = H_1(t)X_S(t) + H_2(t)X_{S^C}(t). \tag{5.6}$$

Then $Y^*(t)$ is the collection $(Y(t), H_1(t), H_2(t))$. As we have discussed in Chapter 2, the form of the various H transfer matrices depends whether we are discussing networks with or without interference, but for all general H the arguments follow identically.

Define $V(t)$ as the received vector $Y_{S^C}(t)$, under the situation that the nodes in S^C did not have any transmitters:

$$V(t) = H_1(t)X_S(t) \tag{5.7}$$

Also, define $V^*(t)$ and the collection of $(V(t), H_1(t))$.

Define X^T as the combined vector of all $X(t)$ vectors, $t \in 1..T$. Similarly define X_S^T, $X_{S^C}^T$, V^T, V^{*T}, $Y_{S^C}^T$, and $Y_{S^C}^{*T}$.

Over T timeslots, we wish to transmit messages from some set of L source-destination node pairs at rates $R(l)$, $l \in 1..L$. The l^{th} source-destination pair has

source node $s(l)$ and destination node $d(l)$. Let W_{cut} be the vector of messages whose source nodes are in S and whose destination nodes are in S^C. W_{S^C} will be the messages whose source nodes are in S^C, regardless of their destinations. All messages $W(l)$ are independent of each other and uniformly chosen from the set $(1, 2, .., 2^{TR(l)})$.

Lemma 5.4.1. *If a set of rates $R(l)$ are achievable, then*

$$\sum_{s(l)\in S, d(l)\in S^C} R(l) \leq I(X_S; V|H_1) \tag{5.8}$$

for some joint distribution $p(x_1, x_2, .., x_{n+1})$. Further, this cut-set bound evaluates to the expected value of the rank (in F_q) of the random matrix H_1 multiplied by $\lg q$.

Proof. The proof is based on the work of Xie and Kumar [6] and the text by Cover and Thomas.[3]

Starting from the fact that the rate across the cut is equal to the entropy of the messages that go across the cut:

$$\begin{aligned} T \sum_{s(l)\in S, d(l)\in S^C} R(l) &= H(W_{cut}) && (5.9)\\ &= I(W_{cut}; V^T, H_1^T, H_2^T, W_{S^C}) \\ &\quad + H(W_{cut}|V^T, H_1^T, H_2^T, W_{S^C}) && (5.10)\\ &\leq I(W_{cut}; V^T, H_1^T, H_2^T, W_{S^C}) + T\epsilon_T && (5.11) \end{aligned}$$

where Equation (5.11) comes from Fano's inequality and the fact that

$$W_{cut} \rightarrow (V^T, H_1^T, H_2^T, W_{S^C}) \rightarrow (Y_{S^C}^{*T}, W_{S^C}) \tag{5.12}$$

forms a Markov chain. The messages W_{cut} are decoded from knowledge of Y_{S^C}, W_{S^C}, H_1^T, and H_2^T.

Returning to Equation (5.11), the steps continue

$$= I(W_{cut}; W_{S^C}) + I(W_{cut}; V^{*T}, H_2^T | W_{S^C}) + T\epsilon_T \tag{5.13}$$

$$= 0 + H(V^{*T}|H_2^T, W_{S^C}) + H(H_2^T|W_{S^C})$$

$$- H(H_2^T|W_{S^C}, W_{cut}) - H(V^{*T}|H_2^T, W_{cut}, W_{S^C}) + T\epsilon_T \tag{5.14}$$

$$= H(V^{*T}|H_2^T, W_{S^C}) - H(V^{*T}|H_2^T, W_{cut}, W_{S^C}) + T\epsilon_T \tag{5.15}$$

$$\leq H(V^{*T}) - H(V^{*T}|H_2^T, W_{cut}, W_{S^C}) + T\epsilon_T \tag{5.16}$$

Equations (5.14 and 5.15) follow because messages W_l and channel states h_{ij} are independent.

We'll now examine the second term in Equation (5.16).

$$H(V^{*T}|H_2^T, W_{cut}, W_{S^C})$$

$$= \sum_{t=1}^{T} H(V^*(t)|V^{*t-1}, W_{cut}, W_{S^C}, H_2^T) \tag{5.17}$$

$$\geq \sum_{t=1}^{T} H(V^*(t)|X_S(t), W_{cut}, W_{S^C}, H_2^T) \tag{5.18}$$

$$= \sum_{t=1}^{T} H(V^*(t)|X_S(t)) \tag{5.19}$$

Equation (5.19) follows because

$$(V^{*t-1}, W_{S^C}, W_{cut}, H_2^T) \rightarrow X_S(t) \rightarrow V^*(t) \tag{5.20}$$

is a Markov chain, since $V(t) = H_1(t)X_S(t)$.

Thus,

$$T \sum_{s(l)\in S, d(l)\in S^C} R(l) = H(W_{cut})$$

$$\leq \sum_{t=1}^{T} I(X_S(t); V^*(t)) + T\epsilon_T$$

$$= \sum_{t=1}^{T} I(X_S(t); V(t)|H_1(t)) + T\epsilon_T \tag{5.21}$$

since X_S and H_1 are independent, which shows the first part of the lemma. Now, examine each mutual information term $I(X_S(t); V(t)|H_1(t))$ which equals $H(V(t)|H_1(t))$ since $V(t)$ is a deterministic function of $H_1(t)$ and $X_S(t)$.

Maximizing the entropy in the vector V is achieved by making all of the elements in $X_S(t)$ independent random variables uniformly distributed over the field F_q and observing that

$$\begin{aligned} & H(V(t)|H_1(t)) \\ & \quad = H(V_1(t)|H_1(t)) + H(V_2(t)|H_1(t), V_1(t)) + ... \\ & \qquad + H(V_m(t)|V^{m-1}(t), H_1(t)) \end{aligned} \tag{5.22}$$

where $m = |S_C|$, $V_k(t)$ is the received symbol at the k^{th} node in S_C, and $V^k(t)$ is the collected vector $(V_1(t), ..., V_k(t))$. Each term in Equation (5.22) has a maximum value of $\lg q$, and, given a particular instance of the transfer matrix $H_1(t)$, if a term can be written as a linear combination of terms with smaller indices, then the conditional entropy of that term is zero. Thus, the value of $H(V(t)|H_1(t))$ is $\lg q$ times the number of linearly independent elements of the vector $V(t)$, or in other words, the rank of the matrix $H_1(t)$ in the finite field F_q times $\lg q$. Taken over all possible instances of the matrix $H_1(t)$, the cutset bound on rates across the cut becomes the expected value of the rank of $H_1(t)$ times $\lg q$. □

5.4.3 One-Dimensional Exponential and Polynomial Decay Models

This section contains the proofs of the first techniques that we discovered for upper-bounding the transport capacity. They are of mathematical interest because the specifics and uniqueness of the proof mechanism, despite having been superseded by the more general and simpler proofs of Section 5.4.5.

Our network consists of $n + 1$ nodes placed along a line, subject to a minimum separation of d_{min}. Index them from 1 to $n + 1$, in geographic order. Let d_i be the distance between the i^{th} and $(i + 1)^{th}$ nodes in this linear network.

We will examine n cuts, where the m^{th} cut separates the nodes into the two sets

S, consisting of nodes with indexes from 1 to m, and S^C, consisting of the nodes with indexes from $m+1$ to $n+1$. Let R_m be the cut-set bound on sum-rate from S to S^C across the m^{th} cut, and R'_m be the sum rate bound on the information from S^C to S.

The total distance weighted sum of reliable rates is therefore bounded by

$$T_{tot} \leq \sum_{m=1}^{n} d_m \left(R_m + R'_m\right) \tag{5.23}$$

In Section 5.4.1 and in [6] the proofs are formed by examining an infinite number of cuts, each placed a constant distance apart from each other. In contrast, the proof in this section examines exactly n cuts, with unequal spacing between cuts. The basic strategy of the proof is to examine the maximum rate across a single cut. In the exponential decay model, we will show that the rate across any cut is bounded by α^{d_i} times some constant K. This is accomplished by assuming that all other inter-node distances are reduced to d_{min}, making them as small as possible, an act which only increases the rate bound across the cut under examination. Thus, transport capacity across this is bounded by $Kd_i\alpha^{d_i}$, which is bounded by a constant T_{e1}, and total transport capacity is bounded by nT_{e1}.

In the polynomial model, we will construct a similar bound (in the form of a summation) for a single cut and again show that the summation converges to a constant for the case when $\beta > 3$.

Theorem 5.4.1. *In a one-dimensional wireless erasure network with an exponential decay law and $n+1$ nodes, the total transport capacity is upperbounded by nT_{e1}, where*

$$T_{e1} = 2(e \ln \frac{1}{\alpha})^{-1}(\frac{1}{1-\alpha^{d_{min}}})^2. \tag{5.24}$$

Proof. We will examine the rate R_m across the m^{th} cut. By Equation (5.2), this rate is

$$R_m = \sum_{j=1}^{m} \left(1 - \prod_{k=m}^{n} \left(1 - \alpha^{\sum_{i=j}^{k} d_i}\right)\right) \tag{5.25}$$

Using the fact that

$$\prod_{i=1}^{n}(1-a_i) \geq 1 - \sum_{i=1}^{n} a_i \tag{5.26}$$

for $0 < a_i < 1$, we can rewrite Equation 5.25 as

$$R_m \leq \sum_{j=1}^{m}\sum_{k=m}^{n} \alpha^{\sum_{i=j}^{k} d_i}. \tag{5.27}$$

Reducing all of the d_i, $i \neq m$, to d_{min} only increases the value of this expression, so

$$R_m \leq \sum_{j=1}^{m}\sum_{k=m}^{n} \alpha^{d_{min}(k-j)}\alpha^{d_m} \tag{5.28}$$

$$= \alpha^{d_m} \sum_{k=m}^{n} \alpha^{kd_{min}} \sum_{j=1}^{m} \alpha^{-jd_{min}} \tag{5.29}$$

$$= \alpha^{d_m}\alpha^{md_{min}} \left(\frac{1-\alpha^{(n-m+1)d_{min}}}{1-\alpha^{d_{min}}}\right) \alpha^{-d_{min}} \left(\frac{\alpha^{-md_{min}}-1}{\alpha^{-d_min}-1}\right) \tag{5.30}$$

$$\leq \alpha^{d_m} \frac{1}{(1-\alpha^{d_{min}})^2}. \tag{5.31}$$

R'_m is bounded similarly, and therefore, the transport capacity across this link only is bounded by

$$d_m\alpha^{d_m} \frac{2}{(1-\alpha^{d_{min}})^2}. \tag{5.32}$$

Finally, because

$$d_m\alpha^{d_m} \leq \left(e \ln \frac{1}{\alpha}\right)^{-1} \tag{5.33}$$

the per-node transport capacity is bounded by T_{e1} for all node distributions which obey the minimum-distance constraint.

□

In a result which closely parallels that of [6], we find that different asymptotic growth rates are possible under various values for the power-law decay parameter β. We will restrict attention to the high-attenuation regime, specifically, when $\beta > 3$ for a linear, one-dimensional network.

Theorem 5.4.2. *In a one-dimensional wireless erasure network with high-attenuation polynomial decay ($\beta > 3$), and $n+1$ nodes, the total transport capacity is upper-bounded by $n2T_{p1}$, where T_{p1} is a constant independent of n.*

Proof. Again we examine the sum-rate bounds R_m and R'_m across the m^{th} cut:

$$R_m \leq \sum_{j=1}^{m} \left(1 - \prod_{k=m}^{n} \frac{1}{1 + (\sum_{i=j}^{k} d_i)^{\beta}} \right). \tag{5.34}$$

As in Section 5.4.3, bound the product term in 5.34 by a summation,

$$\leq \sum_{j=1}^{m} \sum_{k=m}^{n} \frac{1}{1 + \left(\sum_{i=j}^{k} d_i \right)^{\beta}}, \tag{5.35}$$

and upper-bound by decreasing the denominator by dropping the '1+' and changing all $d_i \neq d_m$ to d_{min}.

$$\leq \sum_{j=1}^{m} \sum_{k=1}^{n} \frac{1}{(d_m + (k-j)d_{min})^{\beta}} \tag{5.36}$$

Bound the summation with integrals

$$\leq \int_{0}^{m} \int_{m}^{\infty} (d_m - x d_{min} + y d_{min})^{-\beta} \tag{5.37}$$

which evaluate to

$$= \frac{1}{d_{min}^2} \frac{1}{(\beta-1)(\beta-2)} \left[\frac{1}{d_m^{\beta-2}} - \frac{1}{(d_m + m d_{min})^{\beta-2}} \right] \tag{5.38}$$

Equation 5.38 is upper bounded by

$$= \frac{1}{d_{min}^2} \frac{1}{(\beta-1)(\beta-2)} \frac{1}{d_m^{\beta-2}} \tag{5.39}$$

Recall that Equation (5.39) is an upper-bound R_m on the rate across the m^{th} cut. Multiply (5.39 by d_m to bound the transport capacity across this cut only, and the result is a decreasing function of d_m when $\beta > 3$. Since $d_m \geq d_{min}$, the transport capacity across each and every cut is upper-bounded by the constant

$$T_{p1} = \frac{1}{(\beta-1)(\beta-2)} \frac{d_{min}}{d_{min}^{\beta}}. \tag{5.40}$$

The same analysis holds for R'_m, thus the value of $d_m\,(R_m + R'_m)$ is upper-bounded by the constant $2T_{p1}$. Therefore, the total transport capacity of the network is linearly bounded by $2nT_{p1}$. □

5.4.4 The Squish Technique

Again, this is an early upper-bound technique for a two-dimension model. It was successful for providing the upperbound for a network with the exponential decay model, but failed for a polynomial decay model, with even large values of the polynomial decay exponent β, representing high attenuations.

In Section 5.4.3, we were able to give an upper bound on the rate across the m^{th} cut (specifically, $\alpha^{d_{min}}/(1-\alpha^{d_{min}})^2$) so that the transport capacity across any single cut is bounded by a constant. In a planar two-dimension model, this is no longer the case: Consider a network of $2n$ nodes, placed at (x, y) coordinates $(1, k)$ and $(-1, k)$ for $k = 1..n$. The value of the cut along the y-axis grows linearly with n, so a different proof strategy is required.

Theorem 5.4.3. *In a two-dimensional wireless erasure network with an exponential decay law and $n+1$ nodes, the total transport capacity is upperbounded by nT_{e2}, where T_{e2} is a constant independent of n.*

Before beginning this proof, we will introduce some new notation. As in Section 5.4.1, we will use two sets of cuts, both vertical and horizontal. Unlike Section 5.4.1, we will not have equal spacing between the cuts, but rather proceed as in Section 5.4.3. Label the nodes from 1 to $n+1$ in order of increasing y-coordinate. If any set of nodes share the same y-coordinate, label those in order of increasing x-coordinate. Let d_{ij} be the distance between the i^{th} and j^{th} nodes, and let d_i^v be the *vertical projection* of the distance between nodes i and $i+1$. Similarly define d_i^h, for when the nodes are re-labeled in order of increasing x-coordinate. R_i^v and R_i^h are defined as the rates across the i^{th} vertical and horizontal cuts. $R_i^{v\prime}$ and $R_i^{h\prime}$ are the sum rates for the opposite direction across the cuts.

The transport capacity T_{tot} is therefore bounded by

$$T_{tot} \leq \sum_{m=1}^{n} d_m^v \left(R_m^v + R_m^{v\,\prime}\right) + \sum_{m=1}^{n} d_m^h \left(R_m^h + R_m^{h\,\prime}\right) \tag{5.41}$$

We only need to follow the expressions for one set of cuts, say the horizontal cuts, and for one direction of information flow across the cuts, because the analysis of the other cuts and flow will be identical.

Proof. The expression for the vertical portion of the transport capacity, T_v, is

$$T_v \leq \sum_{m=1}^{n} d_m^v \sum_{j=1}^{m} \left(1 - \prod_{k=m+1}^{n+1} \left(1 - \alpha^{d_{jk}}\right)\right) \tag{5.42}$$

$$\leq \sum_{m=1}^{n} d_m^v \sum_{j=1}^{m} \sum_{k=m+1}^{n+1} \alpha^{d_{jk}} \tag{5.43}$$

We use geometry to lowerbound the value of d_{jk}. As long as $j \leq m < k$,

$$d_{jk} \geq \sqrt{\left(\sum_{i=j}^{m} d_i^v\right)^2 + \frac{d_{min}^2}{3}(k-m)} - \frac{d_{min}}{2} \tag{5.44}$$

where the bound comes from packing $(k-m)$ circles of radius $d_{min}/2$ into a semi-circle whose near side is at least $\sum_{i=j}^{m} d_i^v - d_{min}/2$ away from node j.

Thus, Equation(5.43) is upperbounded by

$$\leq \alpha^{-d_{min}/2} \sum_{m=1}^{n} \sum_{j=1}^{m} d_m^v \sum_{k=1}^{\infty} \alpha^{\sqrt{\left(\sum_{i=j}^{m} d_i^v\right)^2 + k d_{min}^2/2}} \tag{5.45}$$

which can be bounded again by replacing the innermost summation with an integral to obtain

$$\leq \frac{6\alpha^{-d_{min}/2}}{d_{min}^2} \sum_{m=1}^{n} \sum_{j=1}^{m} d_m^v \alpha^{\sum_{i=j}^{m} d_i^v} \left(\sum_{i=j}^{m} d_i^v \frac{1}{-\ln\alpha} + \frac{1}{(\ln\alpha)^2}\right) \tag{5.46}$$

Equation (5.46) can be split into two parts. By redefining $a = -\ln\alpha$, $K_1 = \frac{6\alpha^{-d_{min}/2}}{a^2 d_{min}^2}$, and $K_2 = aK_1$ we can write

$$\begin{aligned} = \quad & K_1 \sum_{m=1}^{n} \sum_{j=1}^{m} d_m^v e^{-a\sum_{i=j}^{m} d_i^v} \\ & + K_2 \sum_{m=1}^{n} \sum_{j=1}^{m} d_m^v \left(\sum_{i=j}^{m} d_i^v\right) e^{-a\sum_{i=j}^{m} d_i^v} \end{aligned} \tag{5.47}$$

and examine the two parts separately.

By defining $x_i = d^v_{n-i}$ for $i = 0, 1, ..., n-1$, the first summation in Equation (5.47) can be written as $K_1 \sum_{k=0}^{n-1} S_k$ where $S_0 = x_0 e^{-ax_0}$ and the S_k are defined recursively as

$$S_k = e^{-ax_k}(S_{k-1} + x_k). \tag{5.48}$$

The S_k are therefore upperbounded by a constant, since all $x_k \geq 0$. Similarly, the second summation in Equation (5.47) can be written as $K_2 \sum_{k=0}^{n-1} V_k$, where $V_0 = S_0$ and

$$V_k = e^{-ax_k}(V_{k-1} + x_k S_{k-1} + x_k^2). \tag{5.49}$$

Each V_k is also upper-bounded by a constant, so the summation T_v, and likewise T_h and T_{tot}, can grow no faster than linearly in the number of nodes n.

□

5.4.5 General Technique for Two Dimensional Networks

As we have stated before, the previous proof techniques were devised and developed in order to solve specific cases of the one- and two-dimensional expansive networks, and build on the work done previously by other authors [5], [6], for example. The most general proof technique, presented in this section, makes use of ideas, lemmas, and the experience gained in the previous sections to create a more elegant exposition.

Theorem 5.4.4. *The transport capacity of an arbitrary expansive broadcast erasure network, with or without interference, grows no faster than linearly in the number of nodes when there exists a constant minimum node separation d_{min}, under the geometric threshold, exponential, and polynomial (with decay $\beta > 3$) decay models.*

Proof. To begin, we will examine the cut-set bound across and infinite number of equally spaced cuts, both horizontal and vertical, as in Section 5.4.1 and in [6]. Space these vertical and horizontal cuts a distance of d_min apart from each

other, and index them by k. Similarly to Equation (5.5), let $R(m)$ denote the cutset bounds across each particular cut. The n nodes of the network will be indexed using i and j, and let $\mathcal{L}_k$ be the set of nodes on the left side of the k^{th} cut (similarly define $\mathcal{R}_k$. Define ϵ_{ij} as ϵd_{ij} and similarly, μ_{ij} as $1 - \epsilon ij$. We bound the total transport capacity of the network by

$$\begin{aligned} &\sum_{m=1}^{\infty} 2d_{min}R(m) \\ =&\sum_{m=1}^{\infty} 2d_{min} \lg q \sum_{i \in \mathcal{L}_m} \left(1 - \prod_{j \in \mathcal{R}_m \epsilon(ij} \right) \end{aligned} \tag{5.50}$$

We can write the rate as we do in Equation (5.50) because of the derivation in Section 5.4.2 showing that back edges do not add anything to the value of the cutset bound. Applying Equation(5.26) yields

$$\sum_{m=1}^{\infty} 2d_{min} \lg q \sum_{i \in \mathcal{L}_m} \sum_{j \in \mathcal{R}_m} \mu_{ij}. \tag{5.51}$$

Define

$$\mathcal{C}_{ij} = \{\text{cuts } m \text{ which separate node } i \text{ from node } j\}$$

in order to rearrange Equation (5.51) as

$$2d_{min} \lg q \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \sum_{m \in \mathcal{C}_{ij}} \mu_{ij}. \tag{5.52}$$

The value of μ_{ij} is fixed for every cut m which it crosses. Defining a normalized internodal distance $\tilde{d_{ij}} = d_{ij}/d_{min}$, then there are at most $2\tilde{d_{ij}}$ cuts (both horizontal and vertical) that separate node i from node j. The transport capacity is thus bounded by

$$4d_{min} \lg q \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \tilde{d_{ij}} \mu_{ij}.$$

Now, take and apply the main idea from the Squish technique of Section 5.4.4: Because of the d_min minimum node-separation constraint, the k^{th} closest node to

any node i must be a distance at least $d_{min}\sqrt{k/6}$ away from node i, by a packing argument. The transport capacity is then no greater than

$$\frac{4d_{min}\lg q}{\sqrt{6}}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\sqrt{k}\mu_{ik} \tag{5.53}$$

where we have reindexed the nodes j in order of increasing distance from the node i under the index k.

At this point, we can apply the different decay models to Equation (5.53) to bound the sum

$$\sum_{k=1}^{\infty}\sqrt{k}\mu_{ik} \tag{5.54}$$

for each node i. In the threshold model, no more than a constant number of nodes can lie within the distance given by the threshold d^*, and Equation (5.54) is less than $\sqrt{d^*/d_{min}}$. In the exponential decay model, Equation (5.54) becomes

$$\sum_{k=1}^{\infty}\sqrt{k}\exp(-\alpha d_{min}\sqrt{k})$$

which is again bounded by a constant (as was shown in Equation (5.46)). Finally, in the polynomial decay model, we again (as in Equation (5.36) bound μ_{ij} as $1/d_{ij}^{\beta}$ to bound Equation (5.54) by

$$\sum_{k=1}^{\infty}\frac{\sqrt{k}}{\left(\sqrt{k}d_{min}\right)^{\beta}} = d_{min}^{-\beta}\sum_{k=1}^{\infty}k^{-(\beta-1)/2}$$

which is a constant for all $\beta > 3$.

The final summation is over each node i, and therefore covers all n nodes in the network once each. Therefore, the transport capacity of such networks are always upperbounded by a linear function of n, the total number of nodes.

□

We have thus provided linear upper-bounds on all arbitrary wireless erasure networks with a minimum distance constraint.

5.4.6 Dense Network Converses

Consider a network with n nodes. Each source (of which there are no more than n of) can only generate symbols with a maximum entropy of $\lg q$. The rate of information transfer between any source and any destination is therefore bounded above by a constant. Also, the maximum distance between any source and any associated destination is bounded by a constant - the network does not grow in size as the number of nodes increases. Therefore, an upperbound on the transport capacity is the maximum rate per node, multiplied by the maximum distance that the information travels, multiplied by the total number of source destination pairs. Two of these quantities are constants, and the third (the number of pairs) by definition grows no faster than linearly in the number of nodes. The transport capacity of a dense wireless erasure network thus can grow no faster than linearly in the number of nodes, whether there be interference or no interference. It might be argued that, while this is clearly an upperbound on the transport capacity of a broadcast erasure network with additive interference, it may not be a "good" upperbound. Here, we demonstrate that the best cut-set upperbound does, in fact, grow linearly in the number of nodes. We will lower-bound the value of the cut-set upperbound and show that our lower-bound grows linearly.

From the more general work of [35], we derive the following lemma:

Lemma 5.4.2. *Consider a random* $0-1$ $n \times n$ *matrix in* GF_2 *where the probability of the entry in the* i^{th} *column and* j^{th} *row being a* 1 *is* p_{ij}*. As long as* $1 - \log n/n > p_{ij} > /logn$*, then with probability* $1 - O(n)$ *the matrix is non-singular as* n *goes to infinity.*

We consider one specific node configuration for a dense network. This is sufficient, since the upperbound for all networks must be at least as large as the upperbound for a specific network. Place each of n nodes along a line spaced at intervals of $1/n$. The cut-set upperbound for this network will be computed exactly as done in, for example, Subsection 5.4.3: We will sum, over all cuts, the width of that cut multiplied by the cut-set bound on rate across that cut.

The width of each cut is $1/n$. The cutset rate for the k^{th} is $E[H_k]$, that is the expected value of the k^{th} transfer matrix. The cutset bound for transport capacity is thus equal to

$$\sum_{k=1}^{n} \frac{1}{n} E[rank(H_k)]$$
$$\geq \frac{1}{n} \sum_{k=n/5}^{4n/5} E[rank(H_k)].$$

Now, the transfer matrix H_k will have k rows and $n-k$ columns. As n grows large, the probability that the entries in all but at most $O(\log n)$ of those rows will be 1 will be between the required bounds on p_{ij} of $\log n$ and $1 - \log n$. We can therefore choose a sub-matrix of H_k of size $2k/3$, all of whose entries will be within the required bounds. Thus, $E[rank(H_k)] > 2k/3$ with high probability.

$$\frac{1}{n} \sum_{k=n/5}^{4n/5} E[rank(H_k)]$$
$$\geq \frac{1}{n} \sum_{k=n/5}^{4n/5} \frac{2k}{3}$$
$$= \Theta(n)$$

Therefore, the best upper-bound on the transport capacity of a broadcast erasure network with interference that uses the technique of the cut-set upperbound must be at least linear in the number of nodes in the network.

We know of no linear achievability strategy for a dense network with interference, however.

5.4.7 Relating Upperbounds for networks with and without additive interference

Precisely computing the expected value of the rank of a random matrix is a difficult problem [35]. However, to linearly bound the transport capacity of the broadcast erasure network with interference, only a simple bound is required:

Lemma 5.4.3. *The expected value of the rank of any random matrix is upper-bounded by the expected value of the number of columns with non-zero entries in that random matrix.*

The proof is obvious.

The important relationship here is that the expected value of the rank of a random matrix with the form required by a broadcast erasure network with no interference is exactly equal to the number of columns with non-zero entries. The interpretation is the usual one - for every transmitting node on the left hand side of the cut, (represented by a column of the transfer matrix), is there at least one successful reception (or, is there at least one 1 in that column of the transfer matrix)? For example, we again present the appropriate transfer matrices A_s for an example network in Figure 5.1.

If Figure 5.1 represents a broadcast erasure network with no interference, then the expression

$$\begin{bmatrix} Y_{d1} \\ Y_{d2} \\ Y_{21} \end{bmatrix} = \begin{bmatrix} 0 & \gamma_{1d} \\ \gamma_{s2} & 0 \\ 0 & \gamma_{12} \end{bmatrix} \begin{bmatrix} X_s \\ X_1 \end{bmatrix}$$

defines A_s for the cut S.

However, if Figure 5.1 represents a broadcast erasure network with additive finite-field interference, then the expression

$$\begin{bmatrix} Y_d \\ Y_2 \end{bmatrix} = \begin{bmatrix} 0 & \gamma_{1d} \\ \gamma_{s2} & \gamma_{12} \end{bmatrix} \begin{bmatrix} X_s \\ X_1 \end{bmatrix}$$

defines A_s for the same cut.

The conclusion: The upperbound on a wireless erasure network without interference is always also an upperbound on a wireless erasure network with additive finite field interference.

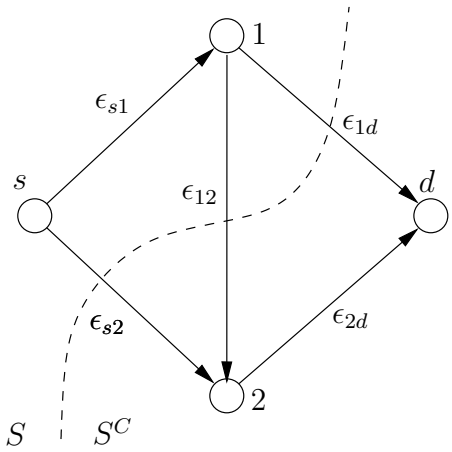

Figure 5.1: Example Network

5.5 Achievability Proofs

5.5.1 Dense Networks Without Any Interference

In a network with a constant expanse - that is, the *maximum* distance between any two nodes is upperbounded by a constant d_{max} - and no receiver interference, achieving a linear growth rate is a straightforward process for any node placement when the polynomial or exponential decay models are used. In each of these models, there is a non-zero constant rate between any source and destination, regardless of what any other nodes in the network are transmitting. Each source and destination pair therefore can communicate at this rate, and ignore all the other traffic in the network. Simultaneous single-hop routing is therefore optimal.

The situation is only slightly different in the geometric threshold model. If the threshold d^* is greater than the expanse of the network, then there is no change in the procedure and a rate of $\lg q$ is available to every single source-destination pair. If the threshold is less than the expanse, then multi-hop routing, with a TDMA scheme dividing the network into $T = (d_{max}/d^*)^2$ squares with side-length d^* each, will again achieve the constant rate arbitrarily close to $\lg q/T$ for each pair in a situation where the nodes are uniformly and randomly placed in the network. If the user is allowed to arbitrarily place the nodes, simply place each destination

within the threshold distance d^* of its source to achieve the rate of $\lg q$.

5.5.2 Expansive Networks - Achievability for Interference Networks Implies Achievability for Non-Interference Networks

In a dual statement to that of Section 5.4.7, it is easy to see that if a set of rates R_l can be achieved in a network with the additive finite field interference, that same set of rates can be achieved in a network without interference, since each receiver can simulate the input in the interference network (by summing up all of its received symbols) with the data that it is given. In the following, then, we shall only consider achievability of a linear growth rate in the expansive network with additive finite field interference.

5.5.3 Random Expansive Networks

Our proof of the achievability of $\Omega(n)$ transport capacity follows from the work of M. Franceschetti et al [7]. We shall summarize their prior results of which we will make use and be more precise with the exposition of those changes that are particular to our model.

The authors of [7] examine a Gaussian interference network of size $\sqrt{n} \times \sqrt{n}$, and show that for a random matching of source-destination pairs on a uniformly randomly distributed set of nodes, a per-node throughput capacity of $\Omega(1/\sqrt{n})$ bit/sec is achievable. The main idea of the proof is to divide the network into horizontal rectangles and show that there exit "highways," or disjoint paths which can be used to carry data from the left edge to the right edge of the network. Similarly, a known density of vertical highways also exist, and this mesh of paths transports the information from sources to destinations. The network operation protocol has three parts: A draining phase, where data is routed from source nodes to the highways, a transport phase, where data travels along the highways, and a distribution phase, where the data is routed from the highway to its final destination node.

Theorem 5.5.1. *A random broadcast erasure network with interference can achieve*

a per-node throughput capacity of $\Omega(1/\sqrt{n})$ *bit/sec, and hence a total transport capacity of* $\Omega(n)$.

The construction of the highways and the routing protocol are identical to the procedure in [7]. Our contribution is determine necessary conditions for the required rate to be achievable in our erasure model, and to show that those conditions are met.

Nodes are distributed according to a Poisson process of unit intensity on square of area n. The total area is divided into smaller squares of side-length c, and a horizontal "highway" consists of a collection these smaller squares, each containing at least one node, which form a continuous path across the total area.

Our proof proceeds as follows: First, we will show the analog of Theorem 3 in [7], namely that there exists a rate $R(d) > 0$ such that a node in our network can transmit w.h.p. at rate $R(d)$ to any destination within a distance d using a TDMA scheme on squares of side length c. Further, as d tends to infinity,

$$R(d) = \Omega\left(d^{-2}e^{-d/d^*}\right). \tag{5.55}$$

Next, the network is divided in horizontal slabs of constant width $c\kappa \ln \sqrt{n/c^2}$, where the parameter κ is to be later determined, and consecutively assign highways to slabs. From [7], Theorem 5 we see that w.h.p., each node is at most a distance $d \leq \sqrt{2}c\kappa \ln \frac{\sqrt{n}}{\sqrt{2}c}$ from its highway. Substituting this d into $R(d)$ in Equation 5.55 and appropriately choosing κ, we find the rate that one node in a square of side length c can transmit data to the highway. There may be as many as $\ln \frac{\sqrt{n}}{c}$ nodes in a square which will have to share this rate. Transmitting data from the highway back to a destination node (the distribution phase) is the dual problem, and therefore is possible to perform at the same rate.

In the highway phase, we note that each node is no more than a distance $2\sqrt{2}c$ from the next node along the highway, so each highway can carry data at a constant rate. There are no more than $O(\sqrt{n})$ nodes in any slab, so the highway can devote a $1/\sqrt{n}$ fraction of its (constant) throughput to each node. As long as

the $R(d)$ achievable in the draining and distribution phases is $\Omega(1/\sqrt{n})$, then a throughput of rate $\Omega(1/\sqrt{n})$ is achievable. In summary, there are $n/2$ random transmit-receive pairs, each providing a rate $\Omega(1/\sqrt{n})$ over a distance $\Omega(\sqrt{n})$, for a transport capacity of $\Omega(n)$.

Proof. Assume that nodes in the network wish to transmit data to destination nodes at most a distance d away. We will say that the transmission was successful only if both 1) the symbol sent by the transmitting node was not erased at the receiver and 2) any symbols sent by simultaneously transmitting nodes are all erased. We will operate the network in a TDMA scheme with $T = (kd/c)^2$ timeslots, where k is to be determined.

If the intended receiver is at most a distance d from the transmitter, then it is at most d/c squares away from the transmitter. Under the TDMA scheme, the nearest simultaneously operating transmitter is at least $(k-1)\frac{d}{c} - 1$ squares away, and the 8 closest transmitters are all (at least) this distance from the receiver. The next 16 operating transmitters are all at least $(2k-1)\frac{d}{c} - 1$ squares away, and so on so that there are $8i$ transmitters at least distance of $(ik-1)d - c$ from the intended receiver, for all positive integers i.

The union bound on the probability that the symbol from at least one of these transmitter is not erased, P_{int}, is

$$\begin{aligned} P_{int} &\leq \sum_{i=1}^{\infty} 8ie^{-((ik-1)d-c)/d^*} \\ &= 8e^{(d+c)/d^*} \frac{e^{-kd/d^*}}{\left(1-e^{-kd/d^*}\right)^2}. \end{aligned} \tag{5.56}$$

We can show that by choosing

$$k > 1 + \left(d^* \ln 32 + c\right)/d \tag{5.57}$$

so that the TDMA scheme operates in

$$T = \lceil (d/c + d^* \ln 32/c + 1)^2 \rceil \tag{5.58}$$

timeslots, $P_{int} < 1$ and thus

$$R(d) \geq e^{-d/d^*}(1 - P_{int})T^{-1}. \tag{5.59}$$

Using that facts that in the draining and distribution phases, $d \leq \sqrt{2}c\kappa \ln \frac{\sqrt{n}}{\sqrt{2}c}$, that the number of timeslots $T = \Theta(d^2)$, and that there may be $O(\ln n)$ nodes in each square, we see that the achievable rate in these phases is

$$\Omega\left(e^{-d/d^*}(1 - P_{int})d^{-2}(\ln n)^{-1}\right) \tag{5.60}$$

$$= \Omega\left(e^{-\sqrt{2}c\kappa \ln \frac{\sqrt{n}}{\sqrt{2}c}/d^*}(\ln n)^{-3}\right) \tag{5.61}$$

$$= \Omega\left(n^{-\sqrt{2}c\kappa/2d^*}(\ln n)^{-3}\right) \tag{5.62}$$

As long as we choose

$$\sqrt{2}c\kappa/2d^* < 1/2 \tag{5.63}$$

while keeping

$$c^2 > \ln 6 + 2/\kappa \tag{5.64}$$

to fulfill the requirements of [7], Theorem 5, a per-node throughput of $\Omega(1/\sqrt{n})$ in our random broadcast erasure network with interference is achievable. □

When the network uses a polynomial decay, rather than the exponential decay, the analysis is similar.

Replace the union bound on the probability that the symbol from at least one of these transmitter is not erased, P_{int}, is thus

$$P_{int} \leq \sum_{k=1}^{\infty} 8k\left((ck-1)\rho\right)^{-\alpha} \tag{5.65}$$

The sum converges for $\alpha > 3$, the range we are interested in, and by choosing an appropriate ρ (independent of n) the upperbound on the probability P_{int} can be made less than 1. The probability of a successful transmission between two nodes in adjacent cells (located no further apart than 2ρ, with no interfering symbols simultaneously received, is then better than

$$R_{neighbor} = \frac{1}{1 + (2\rho)^{\alpha}}\left(1 - P_{int}\right). \tag{5.66}$$

Each node in the cell gets at least a $1/\rho^2 \log n$ fraction of this rate; and the TDMA scheme allows the cell to operate at $1/c^2$ fraction of the time.

5.6 Discussion of Results

We have studied the transport capacity of non-interference wireless erasure networks, where the probability of a erasure increases with distance according to three different models. Under the geometric model, the exponential model, and the high-attenuation polynomial decay model, we have shown that the transport capacity can grow no faster than linearly in the number of nodes. These results nicely parallel the theorems of [6], despite the fact that we are studying the same phenomenon under two different network models.

Chapter 6

Summary

6.1 Summary

The erasure channel is an appropriate model to describe the action of channel-coded communication from the network-layer point of view. It has the additional advantage of analytic tractability. Investigation of the operation of networks of such channels has allowed us to ask and answer a number of compelling and intellectually interesting system level questions. For example, we were able to investigate the capacity of a more generalized class of erasure networks, incorporating both broadcast and receiver interference. We demonstrated the tremendous benefits of feedback, particularly in simplifying coding and routing, in erasure networks. We explored bounds on the secrecy capacity of wireless erasure networks, and showed that our intuitive guesses on such capacity do not hold for all possible networks. Finally, we demonstrated that, very similarly to Gaussian interference networks, the asymptotic transport capacity of wireless erasure networks is linear in a variety of physical layer models. In addition, unlike for Gaussian networks, we have shown that routing is an order-optimal strategy - network coding can yield only constant, or at most, logarithmic gains.

Bibliography

[1] D. Julian, "Erasure networks," in *Proc. IEEE ISIT 2002*, Lausanne, Switzerland, Jul 2002.

[2] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 789-804, Mar 2006.

[3] T. Cover and J. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[4] T. M. Cover and A. A. El Gamal,"Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, No.5, pp. 572-584, Sep 1979.

[5] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 388–404, Mar 2000.

[6] L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Inform. Theory,* vol. 50, no. 5, pp. 748-767, May 2004.

[7] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "On the throughput capacity of random wireless networks," *IEEE Trans. Inform. Theory*, submitted for publication.

[8] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," in *Proc. 44th Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, Oct 2006.

[9] S. Aeron, and V. Saligrama, "'Wireless ad hoc networks: strategies and scaling laws for the fixed SNR regime,"' *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2044-2059, Jun 2007.

[10] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204-1216, Jul 2000.

[11] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2425-2432, June 2006.

[12] B. A. Nazer and M. Gastpar, "Computing over multiple-access channels with connections to wireless network coding," in *Proc. IEEE ISIT 2006*, Seattle, WA, Jul 2006.

[13] S. Bhadra, P. Gupta, and S. Shakkottai, "On network coding for interference networks," in *Proc. IEEE ISIT 2006*, Seattle, WA, Jul 2006.

[14] A. S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow," in *Proc. 45th Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, Oct 2007.

[15] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms, Second Edition.* Cambridge: MIT Press, 2001.

[16] B. Smith and S. Vishwanath, "Unicast transmission over multiple access erasure networks: capacity and duality", IEEE Information Theory Workshop, Lake Tahoe, CA, Sep 2007.

[17] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback-I: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, pp. 172-182, Apr. 1966.

[18] D. S. Lun, M. Médard, and M. Effros, "On coding for reliable communication over packet network", *Proc. 42nd Annual Allerton Conf. on Commun. Control, and Computing*, September 2004.

[19] M. J. Neely and R. Urgaonkar, "Optimal Backpressure Routing in Wireless Networks with Multi-Receiver Diversity," in *Proc. Conference on Information Sciences and Systems*, Mar. 2006.

[20] P. Bremaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*: Springer Science, 2001

[21] I. Csiszár and J. Körner, "Broadcast channels with confidential messages", *IEEE Trans. Inform. Theory*, 24(3):339348, May 1978.

[22] D. Julian, "Dependent and Gaussian erasure networks," in *Proc. IEEE ISIT 2003*, Yokohama, Japan, Jul 2003.

[23] M. Xiao, M. Médard, and T. Aulin,"A Binary Coding Approach for Combination Networks and General Erasure Networks," *Proc. IEEE ISIT*, 2007.

[24] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, 54(8):13551387, October 1975.

[25] Y. Liang and H. V. Poor, "Generalized multiple access channels with confidential messages", *IEEE Trans. Inform. Theory*, 2006.

[26] N. Cai and R. W. Yeung, "Secure Network Coding," ISIT 2002.

[27] J. Feldman, T. Malkin, C. Stein, R. A. Servedio "On the Capacity of Secure Network Coding", Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing, September 2004.

[28] S. E. Rouayheb and E. Soljanin, "On wiretap networks II," in *Proc. 2007 International Symposium on Information Theory, (ISIT'07)*, Nice, France, June 2007.

[29] K. Bhattad and K. R. Narayanan, "Weakly secure network coding", Netcod 2005, Italy, April 2005.

[30] K. Jain, "Security based on network topology against the wiretapping attack," *IEEE Wireless Communications*, pp. 68–71, Feb. 2004.

[31] R. W. Yeung, "A First Course in Information Theory," Kluwer Academic Press.

[32] A. Avestimehr, S. Diggavi, D. Tse, "Wireless network information flow", Proc. 45th Annual Allerton Conference on Communication, Control, and Computing, September 2007.

[33] A. Shamir, "How to Share a Secret", Communications, 1979.

[34] T. Ho, R. Koetter, M. Medard, D. R. Karger and M. Effros, "The benefits of coding over routing in a randomized setting," IEEE International Symposium on Information Theory (ISIT), 2003.

[35] C. Cooper, "On the distribution of rank of a random matrix over a finite field," *Random Structures and Algorithms*, vol. 17, no. 3-4, pp. 197-212, Nov 2000.

Vita

Brian Matthew Smith was born in Youngstown, Ohio on August 15th, 1975, the son of Nancy Smith and Gregory Smith. After completing his work at Webster High School, Webster, New York, he entered the Massachusetts Institute of Technology. He received the degree Bachelor of Science in Electrical Engineering degree from the Massachusetts Institute of Technology in June 2000. He received the degree of Master of Engineering in Electrical Engineering and Computer Science degree from the Massachusetts Institute of Technology in September 2000. In August 2003 he entered the Graduate School of the University of Texas at Austin.

Permanent Address: 3203 Helms, Austin, Texas 78705

This dissertation was typed by the author.