

Full length article

On coding for reliable communication over packet networks

Desmond S. Lun^{a,1}, Muriel Médard^{a,*}, Ralf Koetter^{b,2}, Michelle Effros^c

^a *Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, United States*

^b *Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States*

^c *Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, United States*

Abstract

We consider the use of random linear network coding in lossy packet networks. In particular, we consider the following simple strategy: nodes store the packets that they receive and, whenever they have a transmission opportunity, they send out coded packets formed from random linear combinations of stored packets. In such a strategy, intermediate nodes perform additional coding yet do not decode nor wait for a block of packets before sending out coded packets. Moreover, all coding and decoding operations have polynomial complexity.

We show that, provided packet headers can be used to carry an amount of side-information that grows arbitrarily large (but independently of payload size), random linear network coding achieves packet-level capacity for both single unicast and single multicast connections and for both wireline and wireless networks. This result holds as long as packets received on links arrive according to processes that have average rates. Thus packet losses on links may exhibit correlations in time or with losses on other links. In the special case of Poisson traffic with i.i.d. losses, we give error exponents that quantify the rate of decay of the probability of error with coding delay. Our analysis of random linear network coding shows not only that it achieves packet-level capacity, but also that the propagation of packets carrying “innovative” information follows the propagation of jobs through a queueing network, thus implying that fluid flow models yield good approximations.

© 2008 Elsevier B.V. All rights reserved.

Keywords: Network coding; Packets; Capacity; Erasure networks; Relay networks

1. Introduction

Network information theory generally focuses on applications that, in the open systems interconnection (OSI) model of network architecture, lie in the physical layer. In this context, there are some networked

systems, such as those represented by the multiple-access channel and the broadcast channel, that are well understood, but there are many that remain largely intractable. Even some very simple networked systems, such as those represented by the relay channel and the interference channel, have unknown capacities.

However, the relevance of network information theory is not limited to the physical layer. In practice, the physical layer never provides a fully-reliable bit pipe to higher layers, and reliability then falls on the data link control, network, and transport layers. These layers need to provide reliability not only because of an unreliable physical layer, but also because of packet losses resulting from causes such as congestion (which

* Corresponding author.

E-mail addresses: dslun@mit.edu (D.S. Lun), medard@mit.edu (M. Médard), koetter@uiuc.edu (R. Koetter), effros@caltech.edu (M. Effros).

¹ Current address: Broad Institute of MIT and Harvard, Cambridge, MA 02142, United States.

² Current address: Institute for Communications Engineering, Technische Universität München, D-80290 München, Germany.

leads to buffer overflows) and interference (which leads to collisions). Rather than coding over channel symbols, though, coding is applied over packets, i.e. rather than determining each node's outgoing channel symbols through arbitrary, causal mappings of their received symbols, the contents of each node's outgoing packets are determined through arbitrary, causal mappings of the contents of their received packets. Such packet-level coding offers an alternative domain for network information theory and an alternative opportunity for efficiency gains resulting from cooperation, and it is the subject of our paper.

Packet-level coding differs from symbol-level coding in three principal ways. First, in most packetized systems, packets received in error are dropped, so we need to code only for resilience against erasures and not for noise. Second, it is acceptable to append a degree of side-information to packets by including it in their headers. Third, packet transmissions are not synchronized in the way that symbol transmissions are—in particular, it is not reasonable to assume that packet transmissions occur on every link in a network at identical, regular intervals. These factors make for a different, but related, problem to symbol-level coding. Thus our work addresses a problem of importance in its own right as well as possibly having implications to network information theory in its regular, symbol-level setting.

Aside from these three principal differences, packet-level coding is simply symbol-level coding with packets as the symbols. Thus, given a specification of network use (i.e. a schedule specifying packet injection locations and times), a code specifies the causal mappings that nodes apply to packets to determine their contents; and, given a specification of the probability distribution of erasure patterns in addition to the specification of network use, we can define capacity as the maximum reliable rate (in packets per unit time) that can be achieved. Thus, when we speak of capacity, we speak of Shannon capacity as it is normally defined in network information theory except with packets as the symbols. We do not speak of the various other notions of capacity in networking literature.

The prevailing approach to packet-level coding uses a feedback code: automatic repeat request (ARQ) is used to request the retransmission of lost packets either on a link-by-link basis, an end-to-end basis, or both. This approach often works well and has a sound theoretical basis: it is well known that, given perfect feedback, retransmission of lost packets is a capacity-achieving strategy for reliability on a point-to-point link (see, for example, [1, Section

8.1.5]). Thus, if achieving a network connection meant transmitting packets over a series of uncongested point-to-point links with reliable, delay-free feedback, then retransmission is clearly optimal. This situation is approximated in lightly-congested, highly-reliable wireline networks, but it is not the case in general. First, feedback may be unreliable or too slow, which is often the case in satellite or wireless networks or when servicing real-time applications. Second, congestion can always arise in packet networks; hence the need for retransmission on an end-to-end basis. But, if the links are unreliable enough to also require retransmission on a link-by-link basis, then the two feedback loops can interact in complicated, and sometimes undesirable, ways [2,3]. Moreover, such end-to-end retransmission requests are not well-suited for multicast connections, where, because requests are sent by each terminal as packets are lost, there may be many requests, placing an unnecessary load on the network and possibly overwhelming the source; and packets that are retransmitted are often only of use to a subset of the terminals and therefore redundant to the remainder. Third, we may not be dealing with point-to-point links at all. Wireless networks are the obvious case in point. Wireless links are often treated as point-to-point links, with packets being routed hop-by-hop toward their destinations, but, if the lossiness of the medium is accounted for, this approach is sub-optimal. In general, the broadcast nature of the links should be exploited; and, in this case, a great deal of feedback would be required to achieve reliable communication using a retransmission-based scheme.

In this paper, therefore, we eschew this approach in favor of one that operates mainly in a feedforward manner. Specifically, we consider the following coding scheme: nodes store the packets they receive into their memories and, whenever they have a transmission opportunity, they form coded packets with random linear combinations of their memory contents. This strategy, which we refer to as *random linear network coding*, originates from [4–7], where it was applied to the problem of multicast over lossless wireline packet networks. We shall show that, provided the side-information in packet headers is allowed to grow arbitrarily large (but independently of payload size), random linear network coding achieves packet-level capacity for both single unicast and single multicast connections and for models of both wireline and wireless networks. Our erasure model assumes only that packets received on each link arrive according to a process that has an average rate, thus packet losses on a link may exhibit correlation in time or with losses

on other links, capturing various mechanisms for loss—including collisions.

That we allow the side-information in packet headers to grow arbitrarily large means that, in practice, the capacity we find cannot be approached arbitrarily closely for a finite payload size (since the side-information overhead eventually overwhelms the payload). Rather, a certain fraction of the capacity can be achieved, and, because larger payloads can accommodate larger amounts of side-information, this fraction grows with the size of the payload and approaches one as the size of the payload approaches infinity. Thus our capacity results only truly hold in the limit of infinitely-large payloads, which is an absurd limit. Nevertheless, we believe that our results are useful because they define a limit that can be approached with large payload sizes.

Random linear network coding has several attractive properties besides its ability to approach capacity: it is decentralized, requiring no coordination among nodes; and it can be operated ratelessly, i.e. it can be run indefinitely until successful decoding (at which stage that fact is signaled to other nodes, requiring an amount of feedback that, compared to ARQ, is small), which is a particularly useful property in packet networks, where loss rates are often time-varying and not known precisely.

Decoding can be done by matrix inversion, which is a polynomial-time procedure. Thus, though we speak of random coding, our work differs significantly from that of Shannon [8,9] and Gallager [10] in that we do not seek to demonstrate existence. Indeed, the existence of capacity-achieving linear codes for the scenarios we consider already follows from the results of [11]. Rather, we seek to show the asymptotic rate optimality of a specific scheme that we believe may be practicable and that can be considered as the prototype for a family of related, improved schemes; for example, LT codes [12], Raptor codes [13], Online codes [14], RT oblivious erasure-correcting codes [15], and the greedy random scheme proposed in [16] are related coding schemes that apply only to specific, special networks but, using varying degrees of feedback, achieve lower decoding complexity or memory usage. Our work therefore brings forth a natural code design problem, namely to find such related, improved schemes.

We begin by describing random linear network coding in the following section. In Section 3, we describe our model and illustrate it with several examples. In Section 4, we present coding theorems that prove that random linear network coding is capacity-achieving and, in Section 5, we strengthen these results

in the special case of Poisson traffic with i.i.d. losses by giving error exponents. These error exponents allow us to quantify the rate of decay of the probability of error with coding delay and to determine the parameters of importance in this decay.

2. Random linear network coding

The specific coding scheme we consider is as follows. We suppose that, at the source node, we have K message packets w_1, w_2, \dots, w_K , which are vectors of length λ over the finite field \mathbb{F}_q . (If the packet length is b bits, then we take $\lambda = \lceil b / \log_2 q \rceil$.) The message packets are initially present in the memory of the source node.

The coding operation performed by each node is simple to describe and is the same for every node: received packets are stored into the node's memory, and packets are formed for injection with random linear combinations of its memory contents whenever a packet injection occurs on an outgoing link. The coefficients of the combination are drawn uniformly from \mathbb{F}_q .

Since all coding is linear, we can write any packet x in the network as a linear combination of w_1, w_2, \dots, w_K , namely, $x = \sum_{k=1}^K \gamma_k w_k$. We call γ the *global encoding vector* of x , and we assume that it is sent along with x , as side-information in its header. The overhead this incurs (namely, $K \log_2 q$ bits) is negligible if packets are sufficiently large.

Nodes are assumed to have unlimited memory. The scheme can be modified so that received packets are stored into memory only if their global encoding vectors are linearly-independent of those already stored. This modification keeps our results unchanged while ensuring that nodes never need to store more than K packets.

A sink node collects packets and, if it has K packets with linearly-independent global encoding vectors, it is able to recover the message packets. Decoding can be done by Gaussian elimination. The scheme can be run either for a predetermined duration or, in the case of rateless operation, until successful decoding at the sink nodes. We summarize the scheme in Fig. 1.

The scheme is carried out for a single block of K message packets at the source. If the source has more packets to send, then the scheme is repeated with all nodes flushed of their memory contents.

Random linear network coding schemes similar to the one described above are examined in [4–7] for the application of multicast over lossless wireline packet networks, in [17] for data dissemination, in [18] for data storage, and in [19] for content distribution over peer-to-peer overlay networks. Other coding schemes

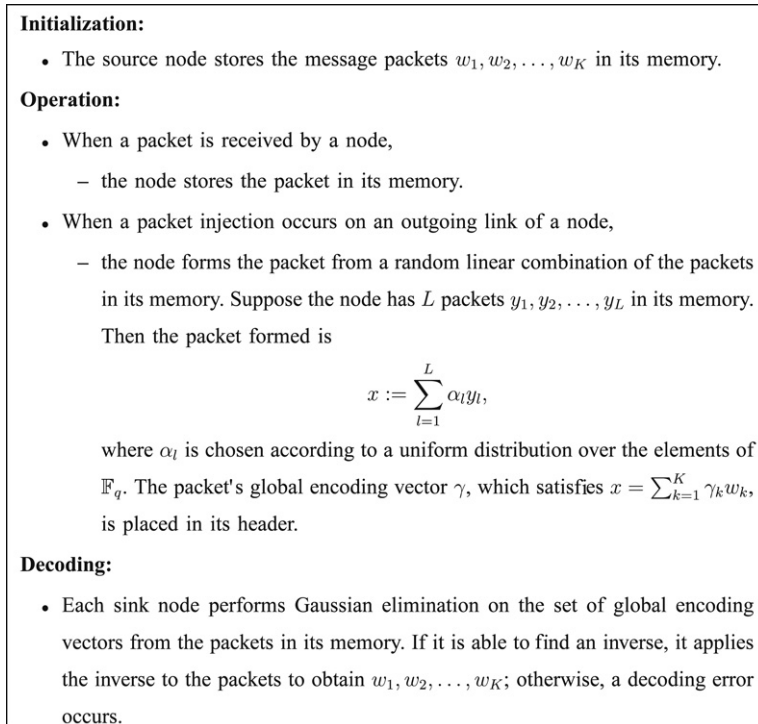


Fig. 1. Summary of the random linear network coding scheme we consider.

for lossy packet networks are described in [11,20]; the scheme described in the former requires placing in the packet headers side-information that grows with the size of the network, while that described in the latter requires no side-information at all, but achieves lower rates in general. Both of these coding schemes, moreover, operate in a block-by-block manner, where coded packets are sent by intermediate nodes only after decoding a block of received packets—a strategy that generally incurs more delay than the scheme we consider, where intermediate nodes perform additional coding yet do not decode [16].

3. Model

Existing models used in network information theory (see, for example, [1, Section 14.10]) are generally conceived for symbol-level coding and, given the peculiarities of packet-level coding, are not suitable for our purpose. One key difference, as we mentioned, is that packet transmissions are not synchronized in the way that symbol transmissions are. Thus we do not have a slotted system where packets are injected on every link at every slot, and we must therefore have a schedule that determines when (in continuous time) and where (i.e. on which link) each packets is injected. In this paper, we assume that such a schedule is given,

and we do not address the problem of determining it. This problem, of determining the schedule to use, is called *subgraph selection*, and it is a difficult problem in its own right, especially in wireless packet networks. Various instances of the problem are treated in [21–27].

Given a schedule of packet injections, the network responds with packet receptions at certain nodes. The difference between wireline and wireless packet networks, in our model, is that the reception of any particular packet may only occur at a single node in wireline packet networks while, in wireless packet networks, it may occur at more than one node.

The model, which we now formally describe, is one that we believe is an accurate abstraction of packet networks as they are viewed at the level of packets, given a schedule of packet injections. In particular, our model captures various phenomena that complicate the efficient operation of wireless packet networks, including interference (insofar as it is manifested as lost packets, i.e. as collisions), fading (again, insofar as it is manifested as lost packets), and the broadcast nature of the medium.

We begin with wireline packet networks. We model a wireline packet network (or, rather, the portion of it devoted to the connection we wish to establish) as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where \mathcal{N} is the set of nodes and \mathcal{A} is the set of arcs. Each arc (i, j) represents a lossy

point-to-point link. Some subset of the packets injected into arc (i, j) by node i are lost; the rest are received by node j without error. We denote by z_{ij} the average rate at which packets are received on arc (i, j) . More precisely, suppose that the arrival of received packets on arc (i, j) is described by the counting process A_{ij} , i.e. for $\tau \geq 0$, $A_{ij}(\tau)$ is the total number of packets received between time 0 and time τ on arc (i, j) . Then, by assumption, $\lim_{\tau \rightarrow \infty} A_{ij}(\tau)/\tau = z_{ij}$ a.s. We define a lossy wireline packet network as a pair (\mathcal{G}, z) .

We assume that links are delay-free in the sense that the arrival time of a received packet corresponds to the time that it was injected into the link. Links with delay can be transformed into delay-free links in a simple way. Suppose that arc (i, j) represents a link with delay. The counting process A_{ij} describes the arrival of received packets on arc (i, j) , and we use the counting process A'_{ij} to describe the injection of these packets. (Hence A'_{ij} counts a subset of the packets injected into arc (i, j) .) We insert a node i' into the network and transform arc (i, j) into two arcs (i, i') and (i', j) . These two arcs, (i, i') and (i', j) , represent delay-free links where the arrival of received packets are described by A'_{ij} and A_{ij} , respectively. We place the losses on arc (i, j) onto arc (i, i') , so arc (i', j) is lossless and node i' simply functions as a first-in first-out queue. It is clear that functioning as a first-in first-out queue is an optimal coding strategy for i' in terms of rate and complexity; hence, treating i' as a node implementing the coding scheme of Section 2 only deteriorates performance and is adequate for deriving achievable connection rates. Thus we can transform a link with delay and average packet reception rate z_{ij} into two delay-free links in tandem with the same average packet reception rate, and it will be evident that this transformation does not change any of our conclusions.

For wireless packet networks, we model the network as a directed hypergraph $\mathcal{H} = (\mathcal{N}, \mathcal{A})$, where \mathcal{N} is the set of nodes and \mathcal{A} is the set of hyperarcs. A hypergraph is a generalization of a graph where generalized arcs, called hyperarcs, connect two or more nodes. Thus a hyperarc is a pair (i, J) , where i , the head, is an element of \mathcal{N} , and J , the tail, is a non-empty subset of \mathcal{N} . Each hyperarc (i, J) represents a lossy broadcast link. For each $K \subset J$, some disjoint subset of the packets injected into hyperarc (i, J) by node i are received by exactly the set of nodes K without error.

We denote by z_{iJK} the average rate at which packets, injected on hyperarc (i, J) , are received by exactly the set of nodes $K \subset J$. More precisely, suppose that the arrival of packets that are injected on hyperarc (i, J) and received by all nodes in K (and no nodes in $\mathcal{N} \setminus K$)

is described by the counting process A_{iJK} . Then, by assumption, $\lim_{\tau \rightarrow \infty} A_{iJK}(\tau)/\tau = z_{iJK}$ a.s. We define a lossy wireless packet network as a pair (\mathcal{H}, z) .

3.1. Examples

3.1.1. Network of independent transmission lines with non-bursty losses

We begin with a simple example. We consider a wireline network where each transmission line experiences losses independently of all other transmission lines, and the loss process on each line is non-bursty, i.e. it is accurately described by an i.i.d. process.

Consider the link corresponding to arc (i, j) . Suppose the loss rate on this link is ε_{ij} , i.e. packets are lost independently with probability ε_{ij} . Suppose further that the injection of packets on arc (i, j) is described by the counting process B_{ij} and has average rate r_{ij} , i.e. $\lim_{\tau \rightarrow \infty} B_{ij}(\tau)/\tau = r_{ij}$ a.s. The parameters r_{ij} and ε_{ij} are not necessarily independent and may well be functions of each other.

For the arrival of received packets, we have

$$A_{ij}(\tau) = \sum_{k=1}^{B_{ij}(\tau)} X_k,$$

where $\{X_k\}$ is a sequence of i.i.d. Bernoulli random variables with $\Pr(X_k = 0) = \varepsilon_{ij}$. Therefore

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \frac{A_{ij}(\tau)}{\tau} &= \lim_{\tau \rightarrow \infty} \frac{\sum_{k=1}^{B_{ij}(\tau)} X_k}{\tau} \\ &= \lim_{\tau \rightarrow \infty} \frac{\sum_{k=1}^{B_{ij}(\tau)} X_k}{B_{ij}(\tau)} \frac{B_{ij}(\tau)}{\tau} = (1 - \varepsilon_{ij})r_{ij}, \end{aligned}$$

which implies that

$$z_{ij} = (1 - \varepsilon_{ij})r_{ij}.$$

In particular, if the injection processes for all links are identical, regular, deterministic processes with unit average rate (i.e. $B_{ij}(\tau) = 1 + \lfloor \tau \rfloor$ for all (i, j)), then we recover the model frequently used in information-theoretic analyses (for example, in [11,20]).

A particularly simple case arises when the injection processes are Poisson. In this case, $A_{ij}(\tau)$ and $B_{ij}(\tau)$ are Poisson random variables with parameters $(1 - \varepsilon_{ij})r_{ij}\tau$ and $r_{ij}\tau$, respectively. We shall revisit this case in Section 5.

3.1.2. Network of transmission lines with bursty losses

We now consider a more complicated example, which attempts to model bursty losses. Bursty losses

arise frequently in packet networks because losses often result from phenomena that are time-correlated, for example, fading and buffer overflows. (We mention fading because a point-to-point wireless link is, for our purposes, essentially equivalent to a transmission line.) In the latter case, losses are also correlated across separate links—all links coming into a node experiencing a buffer overflow will be subjected to losses.

To account for such correlations, Markov chains are often used. Fading channels, for example, are often modeled as finite-state Markov channels [28,29], such as the Gilbert–Elliot channel [30]. In these models, a Markov chain is used to model the time evolution of the channel state, which governs its quality. Thus, if the channel is in a bad state for some time, a burst of errors or losses is likely to result.

We therefore associate with arc (i, j) a continuous-time, irreducible Markov chain whose state at time τ is $E_{ij}(\tau)$. If $E_{ij}(\tau) = k$, then the probability that a packet injected into (i, j) at time τ is lost is $\varepsilon_{ij}^{(k)}$. Suppose that the steady-state probabilities of the chain are $\{\pi_{ij}^{(k)}\}_k$. Suppose further that the injection of packets on arc (i, j) is described by the counting process B_{ij} and that, conditioned on $E_{ij}(\tau) = k$, this injection has average rate $r_{ij}^{(k)}$. Then, we obtain

$$z_{ij} = \pi_{ij}' y_{ij},$$

where π_{ij} and y_{ij} denote the column vectors with components $\{\pi_{ij}^{(k)}\}_k$ and $\{(1 - \varepsilon_{ij}^{(k)})r_{ij}^{(k)}\}_k$, respectively. Our conclusions are not changed if the evolutions of the Markov chains associated with separate arcs are correlated, such as would arise from bursty losses resulting from buffer overflows.

If the injection processes are Poisson, then arrivals of received packets are described by Markov-modulated Poisson processes (see, for example, [31]).

3.1.3. Slotted Aloha wireless network

We now move from wireline packet networks to wireless packet networks or, more precisely, from networks of point-to-point links (transmission lines) to networks where links may be broadcast links.

In wireless packet networks, one of most important issues is medium access, i.e. determining how radio nodes share the wireless medium. One simple, yet popular, method for medium access control is slotted Aloha (see, for example, [32, Section 4.2]), where nodes with packets to send follow simple random rules to determine when they transmit. In this example, we consider a wireless packet network using slotted Aloha

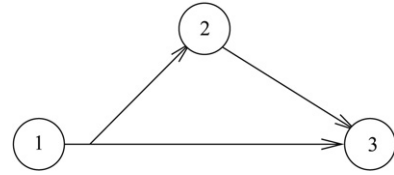


Fig. 2. The slotted Aloha relay channel. We wish to establish a unicast connection from node 1 to node 3.

for medium access control. The example illustrates how a high degree of correlation in the loss processes on separate links sometimes exists.

For the coding scheme we consider, nodes transmit whenever they are given the opportunity and thus effectively always have packets to send. So suppose that, in any given time slot, node i transmits a packet on hyperarc (i, J) with probability q_{iJ} . Let $p'_{iJK|C}$ be the probability that a packet transmitted on hyperarc (i, J) is received by exactly $K \subset J$ given that packets are transmitted on hyperarcs $C \subset \mathcal{A}$ in the same slot. The distribution of $p'_{iJK|C}$ depends on many factors: in the simplest case, if two nodes close to each other transmit in the same time slot, then their transmissions interfere destructively, resulting in a collision where neither node's packet is received. It is also possible that simultaneous transmission does not necessarily result in collision, and one or more packets are received—sometimes referred to as multipacket reception capability [33]. It may even be the case that physical-layer cooperative schemes, such as those presented in [34–36], are used, where nodes that are not transmitting packets are used to assist those that are.

Let p_{iJK} be the unconditioned probability that a packet transmitted on hyperarc (i, J) is received by exactly $K \subset J$. So

$$p_{iJK} = \sum_{C \subset \mathcal{A}} p'_{iJK|C} \left(\prod_{(j,L) \in C} q_{jL} \right) \left(\prod_{(j,L) \in \mathcal{A} \setminus C} (1 - q_{jL}) \right).$$

Hence, assuming that time slots are of unit length, we see that $A_{iJK}(\tau)$ follows a binomial distribution and

$$z_{iJK} = q_{iJ} p_{iJK}.$$

A particular network topology of interest is shown in Fig. 2. The problem of setting up a unicast connection from node 1 to node 3 in a slotted Aloha wireless network of this topology is a problem that we refer to as the slotted Aloha relay channel, in analogy to the symbol-level relay channel widely-studied in network information theory. The latter problem is a well-known open problem, while the former is, as we shall see,

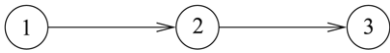


Fig. 3. A network consisting of two links in tandem.

tractable and deals with the same issues of broadcast and multiple access, albeit under different assumptions.

A case similar to that of slotted Aloha wireless networks is that of untuned radio networks, which are detailed in [37]. In such networks, nodes are designed to be low-cost and low-power by sacrificing the ability for accurate tuning of their carrier frequencies. Thus nodes transmit on random frequencies, which leads to random medium access and contention.

4. Coding theorems

In this section, we specify achievable rate regions for random linear network coding in various scenarios. That the regions we specify are the largest possible (i.e. that random linear network coding is capacity-achieving) can be seen by simply noting that the rate between any source and any sink must be limited by the rate at which distinct packets are received over any cut between that source and that sink. A formal converse can be obtained using the cut-set bound for multi-terminal networks (see [1, Section 14.10]).

4.1. Wireline networks

4.1.1. Unicast connections

We develop our general result for unicast connections by extending from some special cases. We begin with the simplest non-trivial case: that of two links in tandem (see Fig. 3).

Suppose we wish to establish a connection of rate arbitrarily close to R packets per unit time from node 1 to node 3. Suppose further that random linear network coding is run for a total time Δ , from time 0 until time Δ , and that, in this time, a total of N packets is received by node 2. We call these packets v_1, v_2, \dots, v_N .

Any received packet x in the network is a linear combination of v_1, v_2, \dots, v_N , so we can write

$$x = \sum_{n=1}^N \beta_n v_n.$$

Since v_n is formed by a random linear combination of the message packets w_1, w_2, \dots, w_K , we have

$$v_n = \sum_{k=1}^K \alpha_{nk} w_k$$

for $n = 1, 2, \dots, N$, where each α_{nk} is drawn from a uniform distribution over \mathbb{F}_q . Hence

$$x = \sum_{k=1}^K \left(\sum_{n=1}^N \beta_n \alpha_{nk} \right) w_k,$$

and it follows that the k th component of the global encoding vector of x is given by

$$\gamma_k = \sum_{n=1}^N \beta_n \alpha_{nk}.$$

We call the vector β associated with x the *auxiliary encoding vector* of x , and we see that any node that receives $\lfloor K(1 + \varepsilon) \rfloor$ or more packets with linearly-independent auxiliary encoding vectors has $\lfloor K(1 + \varepsilon) \rfloor$ packets whose global encoding vectors collectively form a random $\lfloor K(1 + \varepsilon) \rfloor \times K$ matrix over \mathbb{F}_q , with all entries chosen uniformly. If this matrix has rank K , then node 3 is able to recover the message packets. The probability that a random $\lfloor K(1 + \varepsilon) \rfloor \times K$ matrix has rank K is, by a simple counting argument, $\prod_{k=1}^{\lfloor K(1 + \varepsilon) \rfloor - K} (1 - 1/q^k)$, which can be made arbitrarily close to 1 by taking K arbitrarily large. Therefore, to determine whether node 3 can recover the message packets, we essentially need only to determine whether it receives $\lfloor K(1 + \varepsilon) \rfloor$ or more packets with linearly-independent auxiliary encoding vectors.

Our proof is based on tracking the propagation of what we call *innovative* packets. Such packets are innovative in the sense that they carry new, as yet unknown, information about v_1, v_2, \dots, v_N to a node.³ It turns out that the propagation of innovative packets through a network follows the propagation of jobs through a queueing network, for which fluid flow models give good approximations. We present the following argument in terms of this fluid analogy and defer the formal argument to Appendix A.1.

Since the packets being received by node 2 are the packets v_1, v_2, \dots, v_N themselves, it is clear that every packet being received by node 2 is innovative. Thus innovative packets arrive at node 2 at a rate of z_{12} , and this can be approximated by fluid flowing in at rate z_{12} . These innovative packets are stored in node 2's memory, so the fluid that flows in is stored in a reservoir.

Packets, now, are being received by node 3 at a rate of z_{23} , but whether these packets are innovative depends on the contents of node 2's memory. If node 2 has more information about v_1, v_2, \dots, v_N than node 3 does, then

³ Note that, although we are ultimately concerned with recovering w_1, w_2, \dots, w_K rather than v_1, v_2, \dots, v_N , we define packets to be innovative with respect to v_1, v_2, \dots, v_N . This serves to simplify our proof. In particular, it means that we do not need to be very strict in our tracking of the propagation of innovative packets since the number of innovative packets required at the sink is only a fraction of N .

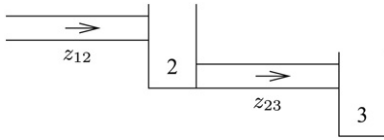


Fig. 4. Fluid flow system corresponding to two-link tandem network.



Fig. 5. A network consisting of L links in tandem.

it is highly likely that new information will be described to node 3 in the next packet that it receives. Otherwise, if node 2 and node 3 have the same degree of information about v_1, v_2, \dots, v_N , then packets received by node 3 cannot possibly be innovative. Thus the situation is as though fluid flows into node 3's reservoir at a rate of z_{23} , but the level of node 3's reservoir is restricted from ever exceeding that of node 2's reservoir. The level of node 3's reservoir, which is ultimately what we are concerned with, can equivalently be determined by fluid flowing out of node 2's reservoir at rate z_{23} .

We therefore see that the two-link tandem network in Fig. 3 maps to the fluid flow system shown in Fig. 4. It is clear that, in this system, fluid flows into node 3's reservoir at rate $\min(z_{12}, z_{23})$. This rate determines the rate at which innovative packets – packets with new information about v_1, v_2, \dots, v_N and, therefore, with linearly-independent auxiliary encoding vectors – arrive at node 3. Hence the time required for node 3 to receive $\lfloor K(1 + \varepsilon) \rfloor$ packets with linearly-independent auxiliary encoding vectors is, for large K , approximately $K(1 + \varepsilon) / \min(z_{12}, z_{23})$, which implies that a connection of rate arbitrarily close to R packets per unit time can be established provided that

$$R \leq \min(z_{12}, z_{23}). \tag{1}$$

Thus we see that rate at which innovative packets are received by the sink corresponds to an achievable rate. Moreover, the right-hand side of (1) is indeed the capacity of the two-link tandem network, and we therefore have the desired result for this case.

We extend our result to another special case before considering general unicast connections: we consider the case of a tandem network consisting of L links and $L + 1$ nodes (see Fig. 5).

This case is a straightforward extension of that of the two-link tandem network. It maps to the fluid flow system shown in Fig. 6. In this system, it is clear that fluid flows into node $(L + 1)$'s reservoir at rate $\min_{1 \leq i \leq L} \{z_{i(i+1)}\}$. Hence a connection of rate arbitrarily close to R packets per unit time from node 1

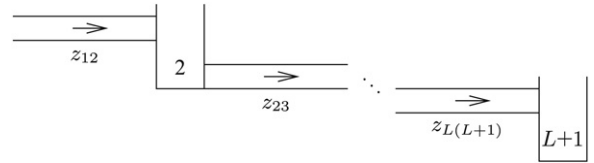


Fig. 6. Fluid flow system corresponding to L -link tandem network.

to node $L + 1$ can be established provided that

$$R \leq \min_{1 \leq i \leq L} \{z_{i(i+1)}\}. \tag{2}$$

Since the right-hand side of (2) is indeed the capacity of the L -link tandem network, we therefore have the desired result for this case. A formal argument is in Appendix A.2.

We now extend our result to general unicast connections. The strategy here is simple: a general unicast connection can be formulated as a flow, which can be decomposed into a finite number of paths. Each of these paths is a tandem network, which is the case that we have just considered.

Suppose that we wish to establish a connection of rate arbitrarily close to R packets per unit time from source node s to sink node t . Suppose further that

$$R \leq \min_{Q \in \mathcal{Q}(s,t)} \left\{ \sum_{(i,j) \in \Gamma_+(Q)} z_{ij} \right\},$$

where $\mathcal{Q}(s, t)$ is the set of all cuts between s and t , and $\Gamma_+(Q)$ denotes the set of forward arcs of the cut Q , i.e.

$$\Gamma_+(Q) := \{(i, j) \in \mathcal{A} \mid i \in Q, j \notin Q\}.$$

Therefore, by the max-flow/min-cut theorem (see, for example, [38, Section 3.1]), there exists a flow vector f satisfying

$$\sum_{\{j \mid (i,j) \in \mathcal{A}\}} f_{ij} - \sum_{\{j \mid (j,i) \in \mathcal{A}\}} f_{ji} = \begin{cases} R & \text{if } i = s, \\ -R & \text{if } i = t, \\ 0 & \text{otherwise,} \end{cases}$$

for all $i \in \mathcal{N}$, and

$$0 \leq f_{ij} \leq z_{ij}$$

for all $(i, j) \in \mathcal{A}$. We assume, without loss of generality, that f is cycle-free in the sense that the subgraph $\mathcal{G}' = (\mathcal{N}, \mathcal{A}')$, where $\mathcal{A}' := \{(i, j) \in \mathcal{A} \mid f_{ij} > 0\}$, is acyclic. (If \mathcal{G}' has a cycle, then it can be eliminated by subtracting flow from f around it.)

Using the conformal realization theorem (see, for example, [38, Section 1.1]), we decompose f into a finite set of paths $\{p_1, p_2, \dots, p_M\}$, each carrying positive flow R_m for $m = 1, 2, \dots, M$, such that

$\sum_{m=1}^M R_m = R$. We treat each path p_m as a tandem network and use it to deliver innovative packets at rate arbitrarily close to R_m , resulting in an overall rate for innovative packets arriving at node t that is arbitrarily close to R . A formal argument is in [Appendix A.3](#).

4.1.2. Multicast connections

The result for multicast connections is, in fact, a straightforward extension of that for unicast connections. In this case, rather than a single sink t , we have a set of sinks T . As in the framework of static broadcasting (see [39,40]), we allow sink nodes to operate at different rates. We suppose that sink $t \in T$ wishes to achieve rate arbitrarily close to R_t , i.e., to recover the K message packets, sink t wishes to wait for a time Δ_t that is only marginally greater than K/R_t . We further suppose that

$$R_t \leq \min_{Q \in \mathcal{Q}(s,t)} \left\{ \sum_{(i,j) \in \Gamma_+(Q)} z_{ij} \right\}$$

for all $t \in T$. Therefore, by the max-flow/min-cut theorem, there exists, for each $t \in T$, a flow vector $f^{(t)}$ satisfying

$$\sum_{\{j|(i,j) \in \mathcal{A}\}} f_{ij}^{(t)} - \sum_{\{j|(j,i) \in \mathcal{A}\}} f_{ji}^{(t)} = \begin{cases} R_t & \text{if } i = s, \\ -R_t & \text{if } i = t, \\ 0 & \text{otherwise,} \end{cases}$$

for all $i \in \mathcal{N}$, and $f_{ij}^{(t)} \leq z_{ij}$ for all $(i, j) \in \mathcal{A}$.

For each flow vector $f^{(t)}$, we go through the same argument as that for a unicast connection, and we find that the probability of error at every sink node can be made arbitrarily small by taking K sufficiently large.

We summarize our results regarding wireline networks with the following theorem statement.

Theorem 1. *Consider the lossy wireline packet network (\mathcal{G}, z) . The random linear network coding scheme described in Section 2 is capacity-achieving for multicast connections, i.e., for K sufficiently large, it can achieve, with arbitrarily small error probability, a multicast connection from source node s to sink nodes in the set T at rate arbitrarily close to R_t packets per unit time for each $t \in T$ if*

$$R_t \leq \min_{Q \in \mathcal{Q}(s,t)} \left\{ \sum_{(i,j) \in \Gamma_+(Q)} z_{ij} \right\}$$

for all $t \in T$.⁴

⁴In earlier versions of this work [41,42], we required the field size q of the coding scheme to approach infinity for Theorem 1 to hold.

Remark. The capacity region is determined solely by the average rate z_{ij} at which packets are received on each arc (i, j) . Therefore the packet injection and loss processes, which give rise to the packet reception processes, can take any distribution, exhibiting arbitrary correlations, as long as these average rates exist.

4.2. Wireless packet networks

The wireless case is actually very similar to the wireline one. The main difference is that we now deal with hypergraph flows rather than regular graph flows.

Suppose that we wish to establish a connection of rate arbitrarily close to R packets per unit time from source node s to sink node t . Suppose further that

$$R \leq \min_{Q \in \mathcal{Q}(s,t)} \left\{ \sum_{(i,J) \in \Gamma_+(Q)} \sum_{K \subset J} z_{iJK} \right\},$$

where $\mathcal{Q}(s, t)$ is the set of all cuts between s and t , and $\Gamma_+(Q)$ denotes the set of forward hyperarcs of the cut Q , i.e.

$$\Gamma_+(Q) := \{(i, J) \in \mathcal{A} \mid i \in Q, J \setminus Q \neq \emptyset\}.$$

Therefore there exists a flow vector f satisfying

$$\sum_{\{j|(i,j) \in \mathcal{A}\}} \sum_{J \in \mathcal{J}} f_{iJj} - \sum_{\{j|(j,i) \in \mathcal{A}, i \in I\}} f_{jIi} = \begin{cases} R & \text{if } i = s, \\ -R & \text{if } i = t, \\ 0 & \text{otherwise,} \end{cases}$$

for all $i \in \mathcal{N}$,

$$\sum_{j \in K} f_{iJj} \leq \sum_{\{L \subset J \mid L \cap K \neq \emptyset\}} z_{iJL} \quad (3)$$

for all $(i, J) \in \mathcal{A}$ and $K \subset J$, and $f_{iJj} \geq 0$ for all $(i, J) \in \mathcal{A}$ and $j \in J$. We again decompose f into a finite set of paths $\{p_1, p_2, \dots, p_M\}$, each carrying positive flow R_m for $m = 1, 2, \dots, M$, such that $\sum_{m=1}^M R_m = R$. Some care must be taken in the interpretation of the flow and its path decomposition because, in a wireless transmission, the same packet may be received by more than one node. The details of the interpretation are in [Appendix A.4](#) and, with it, we can use path p_m to deliver innovative packets at rate arbitrarily close to R_m , yielding the following theorem.

Theorem 2. *Consider the lossy wireless packet network (\mathcal{H}, z) . The random linear network coding scheme*

This requirement is in fact not necessary, and the formal arguments in [Appendix](#) do not require it.

described in Section 2 is capacity-achieving for multicast connections, i.e., for K sufficiently large, it can achieve, with arbitrarily small error probability, a multicast connection from source node s to sink nodes in the set T at rate arbitrarily close to R_t packets per unit time for each $t \in T$ if

$$R_t \leq \min_{Q \in \mathcal{Q}(s,t)} \left\{ \sum_{(i,J) \in \Gamma_+(Q)} \sum_{K \notin Q} z_{iJK} \right\}$$

for all $t \in T$.

5. Error exponents for Poisson traffic with i.i.d. losses

We now look at the rate of decay of the probability of error p_e in the coding delay Δ . In contrast to traditional error exponents where coding delay is measured in symbols, we measure coding delay in time units—time $\tau = \Delta$ is the time at which the sink nodes attempt to decode the message packets. The two methods of measuring delay are essentially equivalent when packets arrive in regular, deterministic intervals.

We specialize to the case of Poisson traffic with i.i.d. losses. Hence, in the wireline case, the process A_{ij} is a Poisson process with rate z_{ij} and, in the wireless case, the process A_{iJK} is a Poisson process with rate z_{iJK} . Consider the unicast case for now, and suppose we wish to establish a connection of rate R . Let C be the supremum of all asymptotically-achievable rates.

To derive exponentially-tight bounds on the probability of error, it is easiest to consider the case where the links are in fact delay-free, and the transformation, described in Section 3, for links with delay has not been applied. The results we derive do, however, apply in the latter case. We begin by deriving an upper bound on the probability of error. To this end, we take a flow vector f from s to t of size C and, following the development in Appendix, develop a queueing network from it that describes the propagation of innovative packets for a given innovation order ρ . This queueing network now becomes a Jackson network. Moreover, as a consequence of Burke's theorem (see, for example, [43, Section 2.1]) and the fact that the queueing network is acyclic, the arrival and departure processes at all stations are Poisson in steady-state.

Let $\Psi_t(m)$ be the arrival time of the m th innovative packet at t , and let $C' := (1 - q^{-\rho})C$. When the queueing network is in steady-state, the arrival of innovative packets at t is described by a Poisson process of rate C' . Hence we have

$$\lim_{m \rightarrow \infty} \frac{1}{m} \log \mathbb{E}[\exp(\theta \Psi_t(m))] = \log \frac{C'}{C' - \theta} \quad (4)$$

for $\theta < C'$ [44,45]. If an error occurs, then fewer than $\lceil R\Delta \rceil$ innovative packets are received by t by time $\tau = \Delta$, which is equivalent to saying that $\Psi_t(\lceil R\Delta \rceil) > \Delta$. Therefore

$$p_e \leq \Pr(\Psi_t(\lceil R\Delta \rceil) > \Delta),$$

and, using the Chernoff bound, we obtain

$$p_e \leq \min_{0 < \theta < C'} \exp(-\theta \Delta + \log \mathbb{E}[\exp(\theta \Psi_t(\lceil R\Delta \rceil))]).$$

Let ε be a positive real number. Then using Eq. (4) we obtain, for Δ sufficiently large,

$$\begin{aligned} p_e &\leq \min_{0 < \theta < C'} \exp\left(-\theta \Delta + R\Delta \left\{ \log \frac{C'}{C' - \theta} + \varepsilon \right\}\right) \\ &= \exp(-\Delta(C' - R - R \log(C'/R)) + R\Delta\varepsilon). \end{aligned}$$

Hence we conclude that

$$\lim_{\Delta \rightarrow \infty} \frac{-\log p_e}{\Delta} \geq C' - R - R \log(C'/R). \quad (5)$$

For the lower bound, we examine a cut whose flow capacity is C . We take one such cut and denote it by Q^* . It is clear that, if fewer than $\lceil R\Delta \rceil$ distinct packets are received across Q^* in time $\tau = \Delta$, then an error occurs. For both wireline and wireless networks, the arrival of distinct packets across Q^* is described by a Poisson process of rate C . Thus we have

$$\begin{aligned} p_e &\geq \exp(-C\Delta) \sum_{l=0}^{\lceil R\Delta \rceil - 1} \frac{(C\Delta)^l}{l!} \\ &\geq \exp(-C\Delta) \frac{(C\Delta)^{\lceil R\Delta \rceil - 1}}{\Gamma(\lceil R\Delta \rceil)}, \end{aligned}$$

and, using Stirling's formula, we obtain

$$\lim_{\Delta \rightarrow \infty} \frac{-\log p_e}{\Delta} \leq C - R - R \log(C/R). \quad (6)$$

Since (5) holds for all positive integers ρ , we conclude from (5) and (6) that

$$\lim_{\Delta \rightarrow \infty} \frac{-\log p_e}{\Delta} = C - R - R \log(C/R). \quad (7)$$

Eq. (7) defines the asymptotic rate of decay of the probability of error in the coding delay Δ . This asymptotic rate of decay is determined entirely by R and C . Thus, for a packet network with Poisson traffic and i.i.d. losses employing the coding scheme described in Section 2, the flow capacity C of the minimum cut of the network is essentially the sole figure of merit of importance in determining the effectiveness of the coding scheme for large, but finite, coding delay. Hence, in deciding how to inject packets to support the

desired connection, a sensible approach is to reduce our attention to this figure of merit, which is indeed the approach taken in [21].

Extending the result from unicast connections to multicast connections is straightforward—we simply obtain (7) for each sink.

6. Conclusion

We have considered the application of random linear network coding to lossy packet networks and shown that it achieves the capacity of single connections as long as packets received on links arrive according to processes that have average rates. In the special case of Poisson traffic with i.i.d. losses, we have given error exponents that quantify the rate of decay of the probability of error with coding delay. Our analysis took into account various peculiarities of packet-level coding that distinguish it from symbol-level coding. Thus our work intersects both with information theory and networking theory and, as such, draws upon results from the two usually-disparate fields [46]. Whether our results have implications for particular problems in either field remains to be explored.

Though we believe that random linear network coding may be practicable, we also believe that, through a greater degree of design or use of feedback, it can be improved. Indeed, feedback can be readily employed to reduce the memory requirements of intermediate nodes by getting them to clear their memories of information already known to their downstream neighbors. Aside from memory requirements, we may wish to improve coding and decoding complexity and side-information overhead. (Some work in the direction of improving coding and decoding complexity can be found in [47].) We may also wish to improve delay—a very important performance factor that we have not explicitly considered, largely owing to the difficulty of doing so. The margin for improvement is elucidated in part in [16], which analyses various packet-level coding schemes, including ARQ and the scheme of this paper, and assesses their delay, throughput, memory usage, and computational complexity for the two-link tandem network of Fig. 3. In our search for such improved schemes, we may be aided by the existing schemes that we have mentioned that apply to specific, special networks.

We should not, however, focus our attention solely on the packet-level code. The packet-level code and the symbol-level code collectively form a type of concatenated code, and an endeavor to understand the interaction of these two coding layers is worthwhile. Some work in this direction can be found in [48].

Acknowledgments

The authors would like to thank Pramod Viswanath and John Tsitsiklis for helpful discussions and suggestions. This work was supported by the National Science Foundation under grant nos. CCR-0093349, CCR-0325496, and CCR-0325673; by the Army Research Office through University of California subaward no. S0176938; and by the Office of Naval Research under grant no. N00014-05-1-0197. This paper was presented in part at the 42nd Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September–October 2004; and in part at the 2005 International Symposium on Information Theory, Adelaide, Australia, September 2005.

Appendix. Formal arguments for main result

Here, we give formal arguments for Theorems 1 and 2. Appendices A.1–A.3 give formal arguments for three special cases of Theorem 1: the two-link tandem network, the L -link tandem network, and general unicast connections, respectively. Appendix A.4 gives a formal argument for Theorem 2 in the case of general unicast connections.

A.1. Two-link tandem network

We consider all packets received by node 2, namely v_1, v_2, \dots, v_N , to be innovative. We associate with node 2 the set of vectors U , which varies with time and is initially empty, i.e. $U(0) := \emptyset$. If packet x is received by node 2 at time τ , then its auxiliary encoding vector β is added to U at time τ , i.e. $U(\tau^+) := \{\beta\} \cup U(\tau)$.

We associate with node 3 the set of vectors W , which again varies with time and is initially empty. Suppose that packet x , with auxiliary encoding vector β , is received by node 3 at time τ . Let ρ be a positive integer, which we call the *innovation order*. Then we say x is innovative if $\beta \notin \text{span}(W(\tau))$ and $|U(\tau)| > |W(\tau)| + \rho - 1$. If x is innovative, then β is added to W at time τ .

The definition of innovative is designed to satisfy two properties: first, we require that $W(\Delta)$, the set of vectors in W when the scheme terminates, is linearly-independent; second, we require that, when a packet is received by node 3 and $|U(\tau)| > |W(\tau)| + \rho - 1$, it is innovative with high probability. The innovation order ρ is an arbitrary factor that ensures that the latter property is satisfied.

Suppose that packet x , with auxiliary encoding vector β , is received by node 3 at time τ and that

$|U(\tau)| > |W(\tau)| + \rho - 1$. Since β is a random linear combination of vectors in $U(\tau)$, it follows that x is innovative with some non-trivial probability. More precisely, because β is uniformly-distributed over $q^{|U(\tau)|}$ possibilities, of which at least $q^{|U(\tau)|} - q^{|W(\tau)|}$ are not in $\text{span}(W(\tau))$, it follows that

$$\begin{aligned} \Pr(\beta \notin \text{span}(W(\tau))) &\geq \frac{q^{|U(\tau)|} - q^{|W(\tau)|}}{q^{|U(\tau)|}} \\ &= 1 - q^{|W(\tau)| - |U(\tau)|} \geq 1 - q^{-\rho}. \end{aligned}$$

Hence x is innovative with probability at least $1 - q^{-\rho}$. Since we can always discard innovative packets, we assume that the event occurs with probability exactly $1 - q^{-\rho}$. If instead $|U(\tau)| \leq |W(\tau)| + \rho - 1$, then we see that x cannot be innovative, and this remains true at least until another arrival occurs at node 2. Therefore, for an innovation order of ρ , the propagation of innovative packets through node 2 is described by the propagation of jobs through a single-server queueing station with queue size $(|U(\tau)| - |W(\tau)| - \rho + 1)^+$.

The queueing station is serviced with probability $1 - q^{-\rho}$ whenever the queue is non-empty and a received packet arrives on arc (2, 3). We can equivalently consider “candidate” packets that arrive with probability $1 - q^{-\rho}$ whenever a received packet arrives on arc (2, 3) and say that the queueing station is serviced whenever the queue is non-empty and a candidate packet arrives on arc (2, 3). We consider all packets received on arc (1, 2) to be candidate packets.

The system we wish to analyze, therefore, is the following simple queueing system: jobs arrive at node 2 according to the arrival of received packets on arc (1, 2) and, with the exception of the first $\rho - 1$ jobs, enter node 2's queue. The jobs in node 2's queue are serviced by the arrival of candidate packets on arc (2, 3) and exit after being serviced. The number of jobs exiting is a lower bound on the number of packets with linearly-independent auxiliary encoding vectors received by node 3.

We analyze the queueing system of interest using the fluid approximation for discrete-flow networks (see, for example, [49,50]). We do not explicitly account for the fact that the first $\rho - 1$ jobs arriving at node 2 do not enter its queue because this fact has no effect on job throughput. Let B_1 , B , and C be the counting processes for the arrival of received packets on arc (1, 2), of innovative packets on arc (2, 3), and of candidate packets on arc (2, 3), respectively. Let $Q(\tau)$ be the number of jobs queued for service at node 2 at time τ . Hence $Q = B_1 - B$. Let $X := B_1 - C$ and

$Y := C - B$. Then

$$Q = X + Y. \quad (8)$$

Moreover, we have

$$Q(\tau)dY(\tau) = 0, \quad (9)$$

$$dY(\tau) \geq 0, \quad (10)$$

and

$$Q(\tau) \geq 0 \quad (11)$$

for all $\tau \geq 0$, and

$$Y(0) = 0. \quad (12)$$

We observe now that Eqs. (8)–(12) give us the conditions for a Skorohod problem (see, for example, [49, Section 7.2]) and, by the oblique reflection mapping theorem, there is a well-defined, Lipschitz-continuous mapping Φ such that $Q = \Phi(X)$.

Let

$$\bar{C}^{(K)}(\tau) := \frac{C(K\tau)}{K},$$

$$\bar{X}^{(K)}(\tau) := \frac{X(K\tau)}{K},$$

and

$$\bar{Q}^{(K)}(\tau) := \frac{Q(K\tau)}{K}.$$

Recall that A_{23} is the counting process for the arrival of received packets on arc (2, 3). Therefore $C(\tau)$ is the sum of $A_{23}(\tau)$ Bernoulli-distributed random variables with parameter $1 - q^{-\rho}$. Hence

$$\begin{aligned} \bar{C}(\tau) &:= \lim_{K \rightarrow \infty} \bar{C}^{(K)}(\tau) \\ &= \lim_{K \rightarrow \infty} (1 - q^{-\rho}) \frac{A_{23}(K\tau)}{K} \quad \text{a.s.} \\ &= (1 - q^{-\rho}) z_{23} \tau \quad \text{a.s.,} \end{aligned}$$

where the last equality follows by the assumptions of the model. Therefore

$$\bar{X}(\tau) := \lim_{K \rightarrow \infty} \bar{X}^{(K)}(\tau) = (z_{12} - (1 - q^{-\rho})z_{23})\tau \quad \text{a.s.}$$

By the Lipschitz-continuity of Φ , then, it follows that $\bar{Q} := \lim_{K \rightarrow \infty} \bar{Q}^{(K)} = \Phi(\bar{X})$, i.e. \bar{Q} is, almost surely, the unique \bar{Q} that satisfies, for some \bar{Y} ,

$$\bar{Q}(\tau) = (z_{12} - (1 - q^{-\rho})z_{23})\tau + \bar{Y}, \quad (13)$$

$$\bar{Q}(\tau)d\bar{Y}(\tau) = 0, \quad (14)$$

$$d\bar{Y}(\tau) \geq 0, \quad (15)$$

and

$$\bar{Q}(\tau) \geq 0 \quad (16)$$

for all $\tau \geq 0$, and

$$\bar{Y}(0) = 0. \quad (17)$$

A pair (\bar{Q}, \bar{Y}) that satisfies (13)–(17) is

$$\bar{Q}(\tau) = (z_{12} - (1 - q^{-\rho})z_{23})^+ \tau \quad (18)$$

and

$$\bar{Y}(\tau) = (z_{12} - (1 - q^{-\rho})z_{23})^- \tau.$$

Hence \bar{Q} is given by Eq. (18).

Recall that node 3 can recover the message packets with high probability if it receives $\lfloor K(1 + \varepsilon) \rfloor$ packets with linearly-independent auxiliary encoding vectors and that the number of jobs exiting the queueing system is a lower bound on the number of packets with linearly-independent auxiliary encoding vectors received by node 3. Therefore node 3 can recover the message packets with high probability if $\lfloor K(1 + \varepsilon) \rfloor$ or more jobs exit the queueing system. Let ν be the number of jobs that have exited the queueing system by time Δ . Then

$$\nu = B_1(\Delta) - Q(\Delta).$$

Take $K = \lceil (1 - q^{-\rho})\Delta R_c R / (1 + \varepsilon) \rceil$, where $0 < R_c < 1$. Then

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{\nu}{\lfloor K(1 + \varepsilon) \rfloor} &= \lim_{K \rightarrow \infty} \frac{B_1(\Delta) - Q(\Delta)}{K(1 + \varepsilon)} \\ &= \frac{z_{12} - (z_{12} - (1 - q^{-\rho})z_{23})^+}{(1 - q^{-\rho})R_c R} \\ &= \frac{\min(z_{12}, (1 - q^{-\rho})z_{23})}{(1 - q^{-\rho})R_c R} \\ &\geq \frac{1}{R_c} \frac{\min(z_{12}, z_{23})}{R} > 1 \end{aligned}$$

provided that

$$R \leq \min(z_{12}, z_{23}). \quad (19)$$

Hence, for all R satisfying (19), $\nu \geq \lfloor K(1 + \varepsilon) \rfloor$ with probability arbitrarily close to 1 for K sufficiently large. The rate achieved is

$$\frac{K}{\Delta} \geq \frac{(1 - q^{-\rho})R_c}{1 + \varepsilon} R,$$

which can be made arbitrarily close to R by varying ρ , R_c , and ε .

A.2. L -link tandem network

For $i = 2, 3, \dots, L + 1$, we associate with node i the set of vectors V_i , which varies with time and is initially empty. We define $U := V_2$ and $W := V_{L+1}$. As in the case of the two-link tandem, all packets received

by node 2 are considered innovative and, if packet x is received by node 2 at time τ , then its auxiliary encoding vector β is added to U at time τ . For $i = 3, 4, \dots, L + 1$, if packet x , with auxiliary encoding vector β , is received by node i at time τ , then we say x is innovative if $\beta \notin \text{span}(V_i(\tau))$ and $|V_{i-1}(\tau)| > |V_i(\tau)| + \rho - 1$. If x is innovative, then β is added to V_i at time τ .

This definition of innovative is a straightforward extension of that in Appendix A.1. The first property remains the same: we continue to require that $W(\Delta)$ is a set of linearly-independent vectors. We extend the second property so that, when a packet is received by node i for any $i = 3, 4, \dots, L + 1$ and $|V_{i-1}(\tau)| > |V_i(\tau)| + \rho - 1$, it is innovative with high probability.

Take some $i \in \{3, 4, \dots, L + 1\}$. Suppose that packet x , with auxiliary encoding vector β , is received by node i at time τ and that $|V_{i-1}(\tau)| > |V_i(\tau)| + \rho - 1$. Thus the auxiliary encoding vector β is a random linear combination of vectors in some set V_0 that contains $V_{i-1}(\tau)$. Hence, because β is uniformly-distributed over $q^{|V_0|}$ possibilities, of which at least $q^{|V_0|} - q^{|V_i(\tau)|}$ are not in $\text{span}(V_i(\tau))$, it follows that

$$\begin{aligned} \Pr(\beta \notin \text{span}(V_i(\tau))) &\geq \frac{q^{|V_0|} - q^{|V_i(\tau)|}}{q^{|V_0|}} \\ &= 1 - q^{|V_i(\tau)| - |V_0|} \geq 1 - q^{|V_i(\tau)| - |V_{i-1}(\tau)|} \geq 1 - q^{-\rho}. \end{aligned}$$

Therefore x is innovative with probability at least $1 - q^{-\rho}$.

Following the argument in Appendix A.1, we see, for all $i = 2, 3, \dots, L$, that the propagation of innovative packets through node i is described by the propagation of jobs through a single-server queueing station with queue size $(|V_i(\tau)| - |V_{i+1}(\tau)| - \rho + 1)^+$ and that the queueing station is serviced with probability $1 - q^{-\rho}$ whenever the queue is non-empty and a received packet arrives on arc $(i, i + 1)$. We again consider candidate packets that arrive with probability $1 - q^{-\rho}$ whenever a received packet arrives on arc $(i, i + 1)$ and say that the queueing station is serviced whenever the queue is non-empty and a candidate packet arrives on arc $(i, i + 1)$.

The system we wish to analyze in this case is therefore the following simple queueing network: jobs arrive at node 2 according to the arrival of received packets on arc $(1, 2)$ and, with the exception of the first $\rho - 1$ jobs, enter node 2's queue. For $i = 2, 3, \dots, L - 1$, the jobs in node i 's queue are serviced by the arrival of candidate packets on arc $(i, i + 1)$ and, with the exception of the first $\rho - 1$ jobs, enter node $(i + 1)$'s queue after being serviced. The jobs in node L 's queue are serviced by the arrival of candidate packets on arc $(L, L + 1)$ and exit after being serviced. The number of

jobs exiting is a lower bound on the number of packets with linearly-independent auxiliary encoding vectors received by node $L + 1$.

We again analyze the queueing network of interest using the fluid approximation for discrete-flow networks, and we again do not explicitly account for the fact that the first $\rho - 1$ jobs arriving at a queueing node do not enter its queue. Let B_1 be the counting process for the arrival of received packets on arc $(1, 2)$. For $i = 2, 3, \dots, L$, let B_i , and C_i be the counting processes for the arrival of innovative packets and candidate packets on arc $(i, i + 1)$, respectively. Let $Q_i(\tau)$ be the number of jobs queued for service at node i at time τ . Hence, for $i = 2, 3, \dots, L$, $Q_i = B_{i-1} - B_i$. Let $X_i := C_{i-1} - C_i$ and $Y_i := C_i - B_i$, where $C_1 := B_1$. Then we obtain a Skorohod problem with the following conditions: for all $i = 2, 3, \dots, L$,

$$Q_i = X_i - Y_{i-1} + Y_i;$$

for all $\tau \geq 0$ and $i = 2, 3, \dots, L$,

$$Q_i(\tau)dY_i(\tau) = 0,$$

$$dY_i(\tau) \geq 0,$$

and

$$Q_i(\tau) \geq 0;$$

for all $i = 2, 3, \dots, L$,

$$Y_i(0) = 0.$$

Let

$$\bar{Q}_i^{(K)}(\tau) := \frac{Q_i(K\tau)}{K}$$

and $\bar{Q}_i := \lim_{K \rightarrow \infty} \bar{Q}_i^{(K)}$ for $i = 2, 3, \dots, L$. Then the vector \bar{Q} is, almost surely, the unique \bar{Q} that satisfies, for some \bar{Y} ,

$$\bar{Q}_i(\tau) = \begin{cases} (z_{12} - (1 - q^{-\rho})z_{23})\tau + \bar{Y}_2(\tau) & \text{if } i = 2, \\ (1 - q^{-\rho})(z_{(i-1)i} - z_{i(i+1)})\tau + \bar{Y}_i(\tau) - \bar{Y}_{i-1}(\tau) & \text{otherwise,} \end{cases} \quad (20)$$

$$\bar{Q}_i(\tau)d\bar{Y}_i(\tau) = 0, \quad (21)$$

$$d\bar{Y}_i(\tau) \geq 0, \quad (22)$$

and

$$\bar{Q}_i(\tau) \geq 0 \quad (23)$$

for all $\tau \geq 0$ and $i = 2, 3, \dots, L$, and

$$\bar{Y}_i(0) = 0 \quad (24)$$

for all $i = 2, 3, \dots, L$. A pair (\bar{Q}, \bar{Y}) that satisfies (20)–(24) is

$$\bar{Q}_i(\tau) = (\min(z_{12}, \min_{2 \leq j < i} \{(1 - q^{-\rho})z_{j(j+1)}\}) - (1 - q^{-\rho})z_{i(i+1)})^+ \tau \quad (25)$$

and

$$\bar{Y}_i(\tau) = (\min(z_{12}, \min_{2 \leq j < i} \{(1 - q^{-\rho})z_{j(j+1)}\}) - (1 - q^{-\rho})z_{i(i+1)})^- \tau.$$

Hence \bar{Q} is given by Eq. (25).

The number of jobs that have exited the queueing network by time Δ is given by

$$v = B_1(\Delta) - \sum_{i=2}^L Q_i(\Delta).$$

Take $K = \lceil (1 - q^{-\rho})\Delta R_c R / (1 + \varepsilon) \rceil$, where $0 < R_c < 1$. Then

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{v}{\lfloor K(1 + \varepsilon) \rfloor} &= \lim_{K \rightarrow \infty} \frac{B_1(\Delta) - \sum_{i=2}^L Q_i(\Delta)}{K(1 + \varepsilon)} \\ &= \frac{\min(z_{12}, \min_{2 \leq i \leq L} \{(1 - q^{-\rho})z_{i(i+1)}\})}{(1 - q^{-\rho})R_c R} \\ &\geq \frac{1}{R_c} \frac{\min_{1 \leq i \leq L} \{z_{i(i+1)}\}}{R} > 1 \end{aligned} \quad (26)$$

provided that

$$R \leq \min_{1 \leq i \leq L} \{z_{i(i+1)}\}. \quad (27)$$

Hence, for all R satisfying (27), $v \geq \lfloor K(1 + \varepsilon) \rfloor$ with probability arbitrarily close to 1 for K sufficiently large. The rate can again be made arbitrarily close to R by varying ρ , R_c , and ε .

A.3. General unicast connection

As described in Section 4.1.1, we decompose the flow vector f associated with a unicast connection into a finite set of paths $\{p_1, p_2, \dots, p_M\}$, each carrying positive flow R_m for $m = 1, 2, \dots, M$ such that $\sum_{m=1}^M R_m = R$. We now rigorously show how each path p_m can be treated as a separate tandem network used to deliver innovative packets at rate arbitrarily close to R_m .

Consider a single path p_m . We write $p_m = \{i_1, i_2, \dots, i_{L_m}, i_{L_m+1}\}$, where $i_1 = s$ and $i_{L_m+1} = t$. For $l = 2, 3, \dots, L_m + 1$, we associate with node i_l the set of vectors $V_l^{(p_m)}$, which varies with time and is initially empty. We define $U^{(p_m)} := V_2^{(p_m)}$ and

$W^{(p_m)} := V_{L_m+1}^{(p_m)}$. Suppose packet x , with auxiliary encoding vector β , is received by node i_2 at time τ . We associate with x the independent random variable P_x , which takes the value m with probability R_m/z_{si_2} . If $P_x = m$, then we say x is innovative on path p_m , and β is added to $U^{(p_m)}$ at time τ . Now suppose packet x , with auxiliary encoding vector β , is received by node i_l at time τ , where $l \in \{3, 4, \dots, L_m + 1\}$. We associate with x the independent random variable P_x , which takes the value m with probability $R_m/z_{i_{l-1}i_l}$. We say x is innovative on path p_m if $P_x = m$, $\beta \notin \text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})$, and $|V_{l-1}^{(p_m)}(\tau)| > |V_l^{(p_m)}(\tau)| + \rho - 1$, where $\tilde{V}_{\setminus m} := \cup_{n=1}^{m-1} W^{(p_n)}(\Delta) \cup \cup_{n=m+1}^M U^{(p_n)}(\Delta)$.

This definition of innovative is somewhat more complicated than that in [Appendices A.1 and A.2](#) because we now have M paths that we wish to analyze separately. We have again designed the definition to satisfy two properties. First, we require that $\cup_{m=1}^M W^{(p_m)}(\Delta)$ is linearly-independent. This is easily verified: vectors are added to $W^{(p_1)}(\tau)$ only if they are linearly-independent of existing ones; vectors are added to $W^{(p_2)}(\tau)$ only if they are linearly-independent of existing ones and ones in $W^{(p_1)}(\Delta)$; and so on. Second, we require that, when a packet is received by node i_l , $P_x = m$, and $|V_{l-1}^{(p_m)}(\tau)| > |V_l^{(p_m)}(\tau)| + \rho - 1$, it is innovative on path p_m with high probability.

Take $l \in \{3, 4, \dots, L_m + 1\}$. Suppose that packet x , with auxiliary encoding vector β , is received by node i_l at time τ , that $P_x = m$, and that $|V_{l-1}^{(p_m)}(\tau)| > |V_l^{(p_m)}(\tau)| + \rho - 1$. Thus the auxiliary encoding vector β is a random linear combination of vectors in some set V_0 that contains $V_{l-1}^{(p_m)}(\tau)$. Hence β is uniformly-distributed over $q^{|V_0|}$ possibilities, of which at least $q^{|V_0|} - q^d$ are not in $\text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})$, where $d := \dim(\text{span}(V_0) \cap \text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m}))$. We have

$$\begin{aligned} d &= \dim(\text{span}(V_0)) + \dim(\text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\quad - \dim(\text{span}(V_0 \cup V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\leq \dim(\text{span}(V_0 \setminus V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\quad - \dim(\text{span}(V_0 \cup V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\leq \dim(\text{span}(V_0 \setminus V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\quad - \dim(\text{span}(V_{l-1}^{(p_m)}(\tau) \cup V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})). \end{aligned}$$

Since $V_{l-1}^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m}$ and $V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m}$ both form linearly-independent sets,

$$\begin{aligned} &\dim(\text{span}(V_{l-1}^{(p_m)}(\tau))) + \dim(\text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &= \dim(\text{span}(V_{l-1}^{(p_m)}(\tau))) + \dim(\text{span}(V_l^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(\tilde{V}_{\setminus m})) \\ &= \dim(\text{span}(V_l^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_{l-1}^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})). \end{aligned}$$

Hence it follows that

$$\begin{aligned} d &\leq \dim(\text{span}(V_0 \setminus V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_l^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_{l-1}^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\quad - \dim(\text{span}(V_{l-1}^{(p_m)}(\tau) \cup V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) \\ &\leq \dim(\text{span}(V_0 \setminus V_{l-1}^{(p_m)}(\tau))) \\ &\quad + \dim(\text{span}(V_l^{(p_m)}(\tau))) \\ &\leq |V_0 \setminus V_{l-1}^{(p_m)}(\tau)| + |V_l^{(p_m)}(\tau)| \\ &= |V_0| - |V_{l-1}^{(p_m)}(\tau)| + |V_l^{(p_m)}(\tau)|, \end{aligned}$$

which yields

$$d - |V_0| \leq |V_l^{(p_m)}(\tau)| - |V_{l-1}^{(p_m)}(\tau)| \leq -\rho.$$

Therefore

$$\begin{aligned} \Pr(\beta \notin \text{span}(V_l^{(p_m)}(\tau) \cup \tilde{V}_{\setminus m})) &\geq \frac{q^{|V_0|} - q^d}{q^{|V_0|}} \\ &= 1 - q^{d-|V_0|} \geq 1 - q^{-\rho}. \end{aligned}$$

We see then that, if we consider only those packets such that $P_x = m$, the conditions that govern the propagation of innovative packets are exactly those of an L_m -link tandem network, which we dealt with in [Appendix A.2](#). By recalling the distribution of P_x , it follows that the propagation of innovative packets along path p_m behaves like an L_m -link tandem network with average arrival rate R_m on every link. Since we have assumed nothing special about m , this statement applies for all $m = 1, 2, \dots, M$.

Take $K = \lceil (1 - q^{-\rho}) \Delta R_c R / (1 + \varepsilon) \rceil$, where $0 < R_c < 1$. Then, by Eq. (26),

$$\lim_{K \rightarrow \infty} \frac{|W^{(p_m)}(\Delta)|}{[K(1 + \varepsilon)]} > \frac{R_m}{R}.$$

Hence

$$\lim_{K \rightarrow \infty} \frac{|\cup_{m=1}^M W^{(p_m)}(\Delta)|}{[K(1+\varepsilon)]} = \sum_{m=1}^M \frac{|W^{(p_m)}(\Delta)|}{[K(1+\varepsilon)]} > \sum_{m=1}^M \frac{R_m}{R} = 1.$$

As before, the rate can be made arbitrarily close to R by varying ρ , R_c , and ε .

A.4. Wireless packet networks

The constraint (3) can also be written as

$$f_{iJj} \leq \sum_{\{L \subset J | j \in L\}} \alpha_{iJL}^{(j)} z_{iJL}$$

for all $(i, J) \in \mathcal{A}$ and $j \in J$, where $\sum_{j \in L} \alpha_{iJL}^{(j)} = 1$ for all $(i, J) \in \mathcal{A}$ and $L \subset J$, and $\alpha_{iJL}^{(j)} \geq 0$ for all $(i, J) \in \mathcal{A}$, $L \subset J$, and $j \in L$. Suppose packet x is placed on hyperarc (i, J) and received by $K \subset J$ at time τ . We associate with x the independent random variable P_x , which takes the value m with probability $R_m \alpha_{iJK}^{(j)} / \sum_{\{L \subset J | j \in L\}} \alpha_{iJL}^{(j)} z_{iJL}$, where j is the outward neighbor of i on p_m . Using this definition of P_x in place of that used in Appendix A.3 in the case of wireline packet networks, we find that the two cases become identical, with the propagation of innovative packets along each path p_m behaving like a tandem network with average arrival rate R_m on every link.

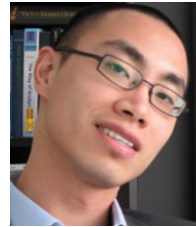
References

- [1] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley & Sons, New York, NY, 1991.
- [2] C. Liu, Cross-layer protocol interactions in heterogeneous data networks, Ph.D. dissertation, Massachusetts Institute of Technology, 2005.
- [3] C. Liu, E. Modiano, On the interaction between layered protocols: The case of window flow control and ARQ, in: Proc. 2002 Conference on Information Sciences and Systems, CISS 2002, March 2002.
- [4] T. Ho, R. Koetter, M. Médard, D.R. Karger, M. Effros, The benefits of coding over routing in a randomized setting, in: Proc. 2003 IEEE International Symposium on Information Theory, ISIT 2003, 2003.
- [5] T. Ho, Networking from a network coding perspective, Ph.D. dissertation, Massachusetts Institute of Technology, May 2004.
- [6] T. Ho, M. Médard, R. Koetter, D.R. Karger, M. Effros, J. Shi, B. Leong, A random linear network coding approach to multicast, IEEE Trans. Inform. Theory 52 (10) (2006) 4413–4430.
- [7] P.A. Chou, Y. Wu, K. Jain, Practical network coding, in: Proc. 41st Annual Allerton Conference on Communication, Control, and Computing, October 2003.
- [8] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (July) (1948) 379–423.
- [9] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (October) (1948) 623–656.
- [10] R.G. Gallager, A simple derivation of the coding theorem and some applications, IEEE Trans. Inform. Theory 11 (1) (1965) 3–18.
- [11] A.F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, M. Effros, Capacity of wireless erasure networks, IEEE Trans. Inform. Theory 52 (3) (2006) 789–804.
- [12] M. Luby, LT codes, in: Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science, November 2002, pp. 271–280.
- [13] A. Shokrollahi, Raptor codes, IEEE Trans. Inform. Theory 52 (6) (2006) 2551–2567.
- [14] P. Maymounkov, Online codes, NYU, Technical Report TR2002-833, November 2002.
- [15] A. Beimel, S. Dolev, N. Singer, RT oblivious erasure correcting, in: Proc. 2004 IEEE Information Theory Workshop, ITW 2004, October 2004, pp. 236–241.
- [16] P. Pakzad, C. Fragouli, A. Shokrollahi, Coding schemes for line networks, in: Proc. 2005 IEEE International Symposium on Information Theory, ISIT 2005, September 2005, pp. 1853–1857.
- [17] S. Deb, M. Médard, C. Choute, Algebraic gossip: A network coding approach to optimal multiple rumor mongering, IEEE Trans. Inform. Theory 52 (6) (2006) 2486–2507.
- [18] S. Acedański, S. Deb, M. Médard, R. Koetter, How good is random linear coding based distributed networked storage? in: Proc. WINMEE, RAWNET and NETCOD 2005 Workshops, April 2005.
- [19] C. Gkantsidis, P.R. Rodriguez, Network coding for large scale content distribution, in: Proc. IEEE Infocom 2005, vol. 4, March 2005, pp. 2235–2245.
- [20] R. Khalili, K. Salamatian, On the capacity of erasure relay channel: Multi-relay case, in: Proc. 2005 IEEE Information Theory Workshop, ITW 2005, August 2005.
- [21] D.S. Lun, N. Ratnakar, M. Médard, R. Koetter, D.R. Karger, T. Ho, E. Ahmed, F. Zhao, Minimum-cost multicast over coded packet networks, IEEE Trans. Inform. Theory 52 (6) (2006) 2608–2623.
- [22] Y.E. Sagduyu, A. Ephremides, Joint scheduling and wireless network coding, in: Proc. WINMEE, RAWNET and NETCOD 2005 Workshops, April 2005.
- [23] Y. Wu, M. Chiang, S.-Y. Kung, Distributed utility maximization for network coding based multicasting: A critical cut approach, in: Proc. 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, WiOpt '06, April 2006.
- [24] Y. Wu, P.A. Chou, S.-Y. Kung, Minimum-energy multicast in mobile ad hoc networks using network coding, IEEE Trans. Commun. 53 (11) (2005) 1906–1918.
- [25] Y. Wu, P.A. Chou, Q. Zhang, K. Jain, W. Zhu, S.-Y. Kung, Network planning in wireless ad hoc networks: A cross-layer approach, IEEE J. Select. Areas Commun. 23 (1) (2005) 136–150.
- [26] Y. Xi, E.M. Yeh, Distributed algorithms for minimum cost multicast with network coding, in: Proc. 43rd Annual Allerton Conference on Communication, Control, and Computing, September 2005.

- [27] Y. Xi, E.M. Yeh, Distributed algorithms for minimum cost multicast with network coding in wireless networks, in: Proc. 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, WiOpt '06, April 2006.
- [28] H.S. Wang, N. Moayeri, Finite-state Markov channel—a useful model for radio communication channels, *IEEE Trans. Veh. Technol.* 44 (1) (1995) 163–171.
- [29] A.J. Goldsmith, P.P. Varaiya, Capacity, mutual information, and coding for finite-state Markov channels, *IEEE Trans. Inform. Theory* 42 (3) (1996) 868–886.
- [30] M. Mushkin, I. Bar-David, Capacity and coding for the Gilbert–Elliott channels, *IEEE Trans. Inform. Theory* 35 (6) (1989) 1277–1290.
- [31] W. Fischer, K. Meier-Hellstern, The Markov-modulated Poisson process (MMPP) cookbook, *Perform. Eval.* 18 (1992) 149–171.
- [32] D.P. Bertsekas, R. Gallager, *Data Networks*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1992.
- [33] S. Ghez, S. Verdú, S.C. Schwartz, Stability properties of slotted Aloha with multipacket reception capability, *IEEE Trans. Automat. Control* 33 (7) (1988) 640–649.
- [34] G. Kramer, M. Gastpar, P. Gupta, Cooperative strategies and capacity theorems for relay networks, *IEEE Trans. Inform. Theory* 51 (9) (2005) 3037–3063.
- [35] J.N. Laneman, D.N.C. Tse, G.W. Wornell, Cooperative diversity in wireless networks: Efficient protocols and outage behavior, *IEEE Trans. Inform. Theory* 50 (12) (2004) 3062–3080.
- [36] K. Azarian, H. El Gamal, P. Schniter, On the achievable diversity-multiplexing tradeoff in half duplex cooperative channels, in: Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing, September–October 2004.
- [37] D. Petrović, K. Ramchandran, J. Rabaey, Overcoming untuned radios in wireless networks with network coding, in: Proc. WINMEE, RAWNET and NETCOD 2005 Workshops, April 2005.
- [38] D.P. Bertsekas, *Network Optimization: Continuous and Discrete Models*, Athena Scientific, Belmont, MA, 1998.
- [39] N. Shulman, Communication over an unknown channel via common broadcasting, Ph.D. dissertation, Tel Aviv University, July 2003.
- [40] N. Shulman, M. Feder, Static broadcasting, in: Proc. 2000 IEEE International Symposium on Information Theory, ISIT 2000, June 2000, p. 23.
- [41] D.S. Lun, M. Médard, M. Effros, On coding for reliable communication over packet networks, in: Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing, September–October 2004, invited paper.
- [42] D.S. Lun, M. Médard, R. Koetter, M. Effros, Further results on coding for reliable communication over packet networks, in: Proc. 2005 IEEE International Symposium on Information Theory, ISIT 2005, September 2005, pp. 1848–1852.
- [43] F.P. Kelly, *Reversibility and Stochastic Networks*, John Wiley & Sons, Chichester, 1979.
- [44] D. Bertsimas, I.C. Paschalidis, J. Tsitsiklis, On the large deviations behavior of acyclic networks of $G/G/1$ queues, *Ann. Appl. Probab.* 8 (4) (1998) 1027–1069.
- [45] I.C. Paschalidis, Y. Liu, Large deviations-based asymptotics for inventory control in supply chains, *Oper. Res.* 51 (3) (2003) 437–460.
- [46] A. Ephremides, B. Hajek, Information theory and communication networks: An unconsummated union, *IEEE Trans. Inform. Theory* 44 (6) (1998) 2416–2434.
- [47] P. Maymounkov, N.J.A. Harvey, D.S. Lun, Methods for efficient network coding, in: Proc. 44th Annu. Allerton Conf.

Communication, Control, and Computing, Monticello, IL, September 2006.

- [48] M. Vehkaperä, M. Médard, A throughput-delay trade-off in packetized systems with erasures, in: Proc. 2005 IEEE International Symposium on Information Theory, ISIT 2005, September 2005.
- [49] H. Chen, D.D. Yao, *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization*, in: Ser. Applications of Mathematics, vol. 46, Springer, New York, NY, 2001.
- [50] H. Chen, A. Mandelbaum, Discrete flow networks: Bottleneck analysis and fluid approximations, *Math. Oper. Res.* 16 (2) (1991) 408–446.



Desmond S. Lun is a Computational Biologist at the Broad Institute of MIT and Harvard and a Research Fellow in Genetics at Harvard Medical School. Prior to his present position, he was a Postdoctoral Research Associate in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. He received bachelor's degrees in Mathematics and Computer Engineering from the University of Melbourne, Australia in 2001, and S.M. and Ph.D. degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology (MIT), Cambridge in 2002 and 2006, respectively. His research interests are in networking and in synthetic and systems biology. He is co-author, with Tracey Ho, of *Network Coding: An Introduction*, forthcoming from Cambridge University Press.



Muriel Médard is an Associate Professor in the Electrical Engineering and Computer Science at MIT. She was previously an Assistant Professor in the Electrical and Computer Engineering Department and a member of the Coordinated Science Laboratory at the University of Illinois, Urbana-Champaign. From 1995 to 1998, she was a Staff Member at MIT Lincoln Laboratory in the Optical Communications and the Advanced Networking Groups. Professor Médard received B.S. degrees in EECS and in Mathematics in 1989, a B.S. degree in Humanities in 1990, a M.S. degree in EE 1991, and a Sc. D. degree in EE in 1995, all from the Massachusetts Institute of Technology (MIT), Cambridge. She has served as an Associate Editor for the *Optical Communications and Networking Series of the IEEE Journal on Selected Areas in Communications*, as an Associate Editor in Communications for the *IEEE Transactions on Information Theory*, as a Guest Editor for the Joint special issue of the *IEEE Transactions on Information Theory* and the *IEEE/ACM Transactions on Networking and Information Theory*, as a Guest Editor for the *IEEE Journal of Lightwave Technology* and as an Associate Editor for the *OSA Journal of Optical Networking*. She is an Associate Editor for the *IEEE Journal of Lightwave Technology* and is a guest editor for the *IEEE Transactions on Information Forensics and Security*. Professor Médard's research interests are in the areas of network coding and reliable communications, particularly for optical and wireless networks. She was awarded the IEEE Leon K. Kirchmayer Prize Paper Award 2002 for her paper, "The Effect Upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel", *IEEE Transactions on Information Theory*, Volume 46 Issue 3, May 2000, Pages: 935–946. She was co-awarded the Best Paper Award for G. Weichenberg, V. Chan, M. Médard, "Reliable Architectures for Networks Under Stress", Fourth International Workshop on the Design of Reliable Communication Networks (DRCN 2003), October 2003, Banff, Alberta, Canada. She received a NSF Career Award in 2001 and was co-winner of 2004 Harold E. Edgerton Faculty Achievement Award, established in 1982 to honor junior faculty

members, “for distinction in research, teaching and service to the MIT community”. She was named a 2007 Gilbreth Lecturer by the National Academy of Engineering. Professor Médard is a House Master at Next House and a Fellow of IEEE.



Ralf Koetter received a Diploma degree in Electrical Engineering from the Technical University Darmstadt, Darmstadt, Germany in 1990 and the Ph.D. degree from the Department of Electrical Engineering at Linköping University, Linköping, Sweden. From 1996 to 1997, he was a Visiting Scientist at the IBM Almaden Research Laboratory in San Jose, CA. He was a Visiting Assistant Professor at the University of Illinois at Urbana-Champaign, Urbana, IL, and a Visiting Scientist at CNRS in Sophia-Antipolis, France, from 1997 to 1998. In the years 1999–2006, he was a member of the faculty of the University of Illinois at Urbana-Champaign. In 2006, he joined the faculty of the Technical University of München as the Head of the Institute for Communications Engineering. His research interests include coding and information theory and their application to communication systems. From 1999 to 2001, Prof. Koetter served as an Associate Editor for Coding Theory and Techniques for the IEEE Transactions on Communications. In 2003, he concluded a term as Associate Editor for Coding Theory of the IEEE Transactions on Information Theory. He received an IBM Invention Achievement Award in 1997, an NSF Career Award in 2000, an IBM Partnership Award in 2001, and a 2006 Xerox award for faculty research. He is corecipient of the 2004 Paper Award of the IEEE Information Theory Society. Since 2003 he has been a Member of the Board of Governors of the IEEE Information Theory Society.

His research interests include coding and information theory and their application to communication systems. From 1999 to 2001, Prof. Koetter served as an Associate Editor for Coding Theory and Techniques for the IEEE Transactions on Communications. In 2003, he concluded a term as Associate Editor for Coding Theory of the IEEE Transactions on Information Theory. He received an IBM Invention Achievement Award in 1997, an NSF Career Award in 2000, an IBM Partnership Award in 2001, and a 2006 Xerox award for faculty research. He is corecipient of the 2004 Paper Award of the IEEE Information Theory Society. Since 2003 he has been a Member of the Board of Governors of the IEEE Information Theory Society.



Michelle Effros received the B.S. degree with distinction in 1989, the M.S. degree in 1990, and the Ph.D. degree in 1994, all in Electrical Engineering from Stanford University, Stanford, CA. During the summers of 1988 and 1989 she worked at Hughes Aircraft Company. She joined the faculty at the California Institute of Technology, Pasadena, CA, in 1994 and is currently a Professor of Electrical Engineering. Her research interests

include information theory, data compression, communications, pattern recognition, and image processing. Effros received Stanford’s Frederick Emmons Terman Engineering Scholastic Award (for excellence in engineering) in 1989, the Hughes Masters Full-Study Fellowship in 1989, the National Science Foundation Graduate Fellowship in 1990, the AT&T Ph.D. Scholarship in 1993, the NSF Career Award in 1995, the Charles Lee Powell Foundation Award in 1997, the Richard Feynman-Hughes Fellowship in 1997, and an Okawa Research Grant in 2000. She is a member of Tau Beta Pi, Phi Beta Kappa, Sigma Xi, and the IEEE Information Theory, Signal Processing, and Communications societies. She served as the Editor of the IEEE Information Theory Society Newsletter from 1995 to 1998 and as a Member of the Board of Governors of the IEEE Information Theory Society from 1998 to 2003. She rejoins the IEEE Information Theory Society Board of Governors for another three-year term in January 2008. She served on the IEEE Signal Processing Society Image and Multi-Dimensional Signal Processing (IMDSP) Technical Committee from 2001 to 2006. She was an Associate Editor for the joint special issue on Networking and Information Theory in the IEEE Transactions on Information Theory and the IEEE/ACM Transactions on Networking and served as Associate Editor for Source Coding for the IEEE Transactions on Information Theory from 2004 to 2007.