Throughput and Stability for Relay-Assisted Wireless Broadcast with Network Coding

Yalin E. Sagduyu, Randall A. Berry, Dongning Guo

***Abstract*—The throughput and stability properties of wireless network coding are evaluated for an arbitrary number of terminals exchanging broadcast traffic with the aid of a relay. First, coding and scheduling schemes are derived that minimize the number of transmissions needed for each node to broadcast one packet. For stochastically varying traffic, the stable throughput is then compared under both digital and analog network coding schemes. The initial analysis focuses on a network with a single relay. Extensions to arbitrary terminal-relay configurations are then outlined for a general multihop network. Backpressure-like algorithms for jointly achieving throughput optimal scheduling and network coding are given for each network coding scheme.**

***Index Terms*—Digital Network Coding; Analog Network Coding; Throughput; Delay; Stability; Network Control.**

I. Introduction

Network coding (NC) generalizes store-and-forward-based routing by providing the general capability to process traffic at relay nodes. For single-source multicast in wired networks, NC achieves the maximum (min-cut) capacity [1]. Multi-source NC [2] expands the achievable rate region of multiple (multicast and unicast) flows; however, the characterization of the general capacity region remains an open problem [3]. The extension to wireless networks involves the joint design of NC with medium access control under half-duplex broadcast communication and interference effects [4]–[7]. Different from packet-level *digital network coding* (DNC), a relay can also amplify and forward the superpositions of signals received in wireless access (without decoding) thereby performing *analog network coding* (ANC) in the physical layer [8], [9].

Much of the work on multi-source wireless NC has focused on *pairwise* NC of unicast flows [10]–[13]; the canonical example being the exchange channel, in which two terminals transmit packets to each other over a third relay node [14]–[19]. To gain insights into multi-source inter-session NC, we consider a star network, which generalizes the exchange channel by allowing an arbitrary number of sources to exchange packets with relay assistance. For star networks with backlogged unicast traffic, the coding gain of NC has been illustrated in [7], the information-theoretic capacity region has been derived in [20], and the effects of MIMO switching have been studied with physical-layer NC in [21]. Here, our focus is on maximizing the *broadcast* throughput achieved by DNC and ANC while stabilizing packet queues under *stochastic* traffic. This requires the joint design of network codes and dynamic transmission scheduling.

Our approach for handling stochastic traffic is motivated by the *backpressure* algorithm [22]. This approach applies to multiple unicast flows and extends to multicast flows under joint scheduling and routing [23]. We consider generalizations of such policies for broadcast traffic in wireless networks that employ DNC and ANC. Backpressure algorithms have been formulated for intra-session NC of packets from the same source [24] and for the special case of pairwise NC of unicast flows [25], [26]. The queue stability with multi-source NC has been studied for other network models, including exchange channels with one relay [27] or two relays [28], line networks [6], and butterfly network structures [29].

We consider random broadcast traffic among multiple source terminals with two possible paths, either by overhearing of direct transmissions or by two-hop transmissions over a relay that uses *local* NC. Such broadcast traffic may be the dominant form of traffic in some wireless networks such as one supporting first responders, and may also be a good model for certain control traffic, video conferencing, and file synchronization through a relay/access point. The star network topology we consider in this paper can be used to interconnect gateway nodes via relays with applications in satellite backhaul, peer to peer, sensor, and cellular networks [20], [30]. Our objectives are: 1) to develop network codes and schedules that minimize the number of transmissions to exchange broadcast traffic, 2) to characterize the maximum throughput region, and 3) to derive throughput optimal control to stabilize the packet queues for rates in this region.

Meeting these objectives is not a straightforward extension of the plain routing case in [22] because NC serves packet queues *jointly*. Hence, the service rates of the relay and terminal queues depend on each other's queue backlogs, preventing us from applying Lyapunov techniques over instantaneous queue sizes to show stability as in [22]. Instead, we consider Lyapunov stability arguments (with *state-dependent step-sizes* in Lyapunov drift) over the entire duration of any NC session to decouple the packet queues. This leads to different formulations of backpressure algorithms, which are shown to achieve the maximum stable throughput. They result in *crosslayer* queue-based NC and dynamic scheduling without *a priori* information on packet arrival statistics.

We consider both DNC and ANC at the relay nodes. DNC can be viewed as index coding [31], where the relay finds

Y. E. Sagduyu is with Intelligent Automation Inc. Rockville, MD 20855, USA, and with the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. E-mail: yalinsagduyu@ieee.org. R. A. Berry and D. Guo are with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208, USA. E-mail: rberry@eecs.northwestern.edu, dguo@northwestern.edu. This work was presented in part at CISS 2008, WCNC 2008, and WICON 2008. This material is based on work supported by DARPA under grant W911NF-07-1-0028 and by NSF under grant CCF-0905224.

the minimum number of coded packets needed to exchange the traffic between terminals with side information. Here, both traffic demand and side information (from packet overhearing) vary stochastically, requiring DNC to adapt. In ANC, transmissions of terminals are scheduled to ensure that they receive sufficient degrees of freedom to decode packets after the relay amplifies and forwards superpositions of the received signals. Hence, side information is again dynamically built along with transmission scheduling. We quantify the improvement of stable throughput rates over plain routing for each NC scheme. Our initial focus is on a single relay model. We then outline the extension to multiple relays.

The paper is organized as follows. The single relay model with DNC and ANC is introduced in Section II. We derive the maximum throughput region and give a throughput optimal control for DNC and ANC in Sections III and IV, respectively. The generalization of the model to multiple relays is given in Section V. The paper concludes in Section VI.

II. Broadcast via a Single Relay

Consider $K \geq 2$ terminals communicating with the aid of a *single* relay R, as in Figure 1. *Half-duplex* communication is assumed in a synchronous slotted system, in which each transmission takes one time-slot. Each terminal has broadcast traffic that is destined to all others. Any terminal i can overhear n neighbor terminals on each side, collectively denoted as $\mathcal{N}_i$, where $i \in \mathcal{N}_i$ so that $|\mathcal{N}_i| = 2n+1$. Terminals do not forward packets received from the other terminals. We assume $2n < K-1$; otherwise, all terminals can overhear each other and exchange packets over K time-slots without using the relay.

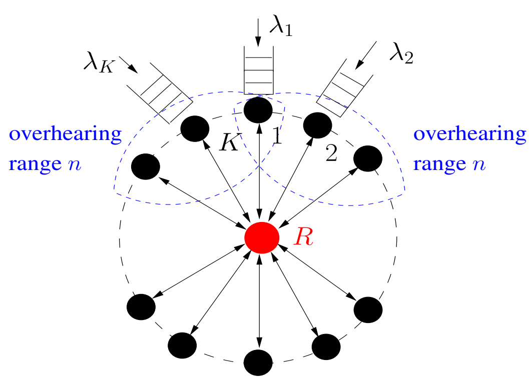

Fig. 1. Star topology with $K \geq 2$ terminals (with overhearing range n) and a single relay R.

Two network coding schemes are considered:

(a) *Digital Network Coding (DNC)*: Terminals take turns transmitting their packets to the relay. The relay decodes, re-encodes and broadcasts the incoming packets, performing linear network coding in finite field F_q with field size q.

(b) *Analog Network Coding (ANC)*: Terminals transmit first to the relay (possibly over concurrent slots). The relay amplifies and broadcasts the superpositions of the signals received from different terminals, and sends them back to all terminals, without decoding the packets.

We consider a *two-phase* operation separated in time. In phase 1 (multiaccess), terminals transmit packets according to a *scheduling matrix* $\tilde{G}$, where $\tilde{G}_{t,j} = 1$, if terminal j transmits in time-slot t, and $\tilde{G}_{t,j} = 0$, otherwise. In plain routing and DNC, $\tilde{G}$ is I_K, the $K \times K$ identity matrix. In phase 2 (broadcast), the relay transmits to the terminals. In DNC, the relay decodes and re-encodes the packets according to the *coding matrix* G, where $G_{i,j}$ is the linear coding coefficient for terminal j's packet in the ith coded packet. In plain routing, $G = I_K$. In ANC, the relay amplifies and forwards the received signals without decoding (equivalently, $G = I_K$). Here, our focus is on the network-layer aspects of relaying. The physical-layer rates achieved by DNC and ANC have been extensively studied for backlogged traffic in exchange channels (e.g., in [32], [33]) and we can readily apply them here.

A. Delay-Optimal Coding and Scheduling with the Minimum Number of Transmissions

To begin we consider minimizing the total delay when all nodes have one packet to broadcast and can reliably send one packet in each time-slot. Let $T(K,n)$ denote the number of successful transmissions to exchange K packets, one packet from each of K terminals with the overhearing range n. The minimal value of $T(K,n)$ under different coding assumption is given in Lemma 1, whose proof is in the Appendix.[1]

***Lemma** 1:* For $K > 2n+1$, the minimum value of $T(K,n)$ for plain routing, DNC, and ANC, respectively, is given by

$$\begin{aligned} T_R(K,n) &= 2K, & (1)\\ T_{DNC}(K,n) &= 2K-1-2n, & (2)\\ T_{ANC}(K,n) &= \max\left(2\left\lceil K/2\right\rceil, 2(K-1-2n)\right). & (3) \end{aligned}$$

For DNC, the minimum value of $T(K,n)$ can be achieved by deterministic NC, as shown in the proof of Lemma 1, or by random NC for sufficiently large coding field size. Since $K-1$ packets are broadcast from each terminal over $T(K,n)$ slots, the broadcast throughput per terminal is $\lambda(K,n) = \frac{K-1}{T(K,n)}$, assuming no error. For finite n, $\lim_{K\to\infty} \lambda(K,n) = 0.5$ for all the three schemes. Next, we consider a geometric model in which n increases with K. Let d be the distance over which a terminal can overhear another. If nodes are equally spaced on a circle of radius r around the relay, we have $n = \left\lfloor \frac{K}{\pi} \sin^{-1}\left(\frac{d}{2r}\right)\right\rfloor$, where we assume that r is small enough so that terminals can always reach the relay. The gain of DNC and ANC over plain routing (in terms of reducing the transmission number) is shown in Figure 2, as K grows to infinity. For $d = r$, we have $n = \left\lfloor \frac{K}{6}\right\rfloor$ and the gain is given by $\lim_{K\to\infty} 1 - \frac{T_{DNC}(K,n)}{T_R(K,n)} = 16.67\%$ and $\lim_{K\to\infty} 1 - \frac{T_{ANC}(K,n)}{T_R(K,n)} = 33.33\%$ for DNC and ANC, respectively. This gain increases with $\frac{d}{r}$ to the maximum value of 50% that is attained by DNC at $d = 2r$ and by ANC at $d = \sqrt{2}r$.

B. Stochastic Traffic Model

For the remainder of the paper, we assume that the broadcast traffic at each terminal is generated according to an inde-

[1] This result first appeared in [33].

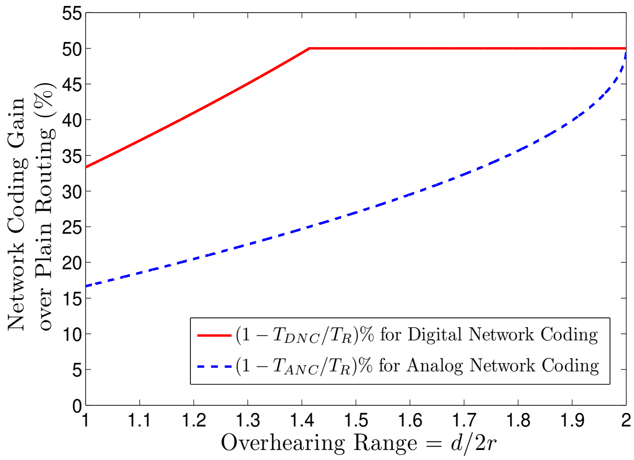

Fig. 2. The gain of DNC and ANC over plain routing, as $K \to \infty$.

pendent stationary random process. For terminal i, let λ_i be the average arrival rate and $A_i(t)$ be the number of packets arriving in time-slot t. We assume that $A_i(t) \le A_{i,\max}$ and the second moment is bounded. Each terminal keeps a packet queue of infinite capacity. In plain routing and DNC, the relay maintains *different* queues for packets of different terminals. In ANC, the relay amplifies and forwards packets immediately without queueing at the relay. Transmissions of terminal i and the relay are allocated over separate time fractions f_i and f_r, respectively. Let f_t be the total time fraction for terminal transmissions. The maximum throughput region Λ is the convex hull of rates $\{\lambda_i\}_{i=1}^K$ that are achieved by optimizing over $\{f_i\}_{i=1}^K$, f_t, and f_r. Any arrival rate within Λ can be supported by a fixed schedule, given *a priori* knowledge of the arrival rates.

Our goal is to find a throughput optimal policy to stabilize any set of rates in Λ, when the arrival statistics are not known. We still assume $K > 2n+1$ as otherwise, the optimal policy reduces to scheduling the terminal with the *longest* (non-empty) queue to transmit without using the relay.[2] We generalize the previous example by allowing a different (constant) channel rate on each link.[3] Let $C_{i,r}$ denote the capacity of the channel from terminal i to the relay and $C_{r,i}$ the capacity of the channel from the relay to terminal i. Any terminal i can decode transmissions of any terminal $j \in \mathcal{N}_i$. Otherwise, $C_{i,j} = 0$. In ANC, the terminal and relay transmissions are coupled and the end-to-end rates will be considered in Section IV.

III. Stochastic Network Control with Digital Network Coding

Let Q_i and Q_i^r be the queues for terminal i's packets at the terminal and relay, respectively. The corresponding backlogs in time-slot t are $U_i(t)$ and $U_i^r(t)$. We define $\mu_i(t)$ as the transmission rate from queue Q_i and $\mu_i^r(t)$ as the transmission rate from queue Q_i^r.[4] Plain routing timeshares the terminal and relay transmissions, achieving the rate region

$$\Lambda_R = \left\{\lambda > 0 : \sum_{i=1}^{K} \lambda_i \left(\frac{1}{C_i^r} + \frac{1}{C_{i,r}}\right) \le 1\right\}, \qquad (4)$$

where $C_i^r = \min_{j: i \notin \mathcal{N}_j} C_{r,j}$ is the broadcast rate from the relay for packets from terminal i. In DNC when all nodes are backlogged, packets in the relay queues are *linearly coded* as $G\mathbf{y}^r$, where $\mathbf{y}^r = \{y_i^r\}_{i=1}^K$ and y_i^r is the first packet in Q_i^r. With stochastic traffic, the service rates of relay queues are coupled in DNC and in any slot t they depend on $q_{br}(t) = \sum_{i=1}^K 1_{\{U_i^r(t)>0\}}$, the number of backlogged relay queues with at least one packet to transmit (where $(1_{\{.\}}$ is the indicator function). Here, we consider a relay which operates in one of two broadcast modes. One broadcast mode corresponds to using DNC with a coding matrix G corresponding to the minimum number of $m = K - 1 - 2n$ coded transmissions given in Lemma 1. The relay includes dummy packets from all empty queues and serves all backlogged queues at a rate of $\frac{C^r}{m}$, where $C^r = \min_{i=1,\ldots K} C_i^r$ is the capacity of the weakest channel (used for relay transmissions of packets from every terminal). The second broadcast mode is to simply employ plain routing. If $q_{br}(t) > m$, either DNC or plain routing may achieve a higher rate depending on the channel capacities. If $q_{br}(t) \le m$, plain routing is optimal. We define $[U]^+ = \max(U, 0)$ and $\mu_i^r\left([\tau]_{\tau=t}^{t+m-1}\right)$ as the service rate of relay queue Q_i^r over the m slots, $[\tau]_{\tau=t}^{t+m-1}$.

In any time-slot, the scheduler chooses one from *three* possible transmission strategies: 1) only terminal i transmits to the relay, 2) the relay forwards plain packets from terminal i, 3) the relay transmits network-coded packets over m time-slots. Here, DNC refers to joint digital NC and routing such that when the relay does not have enough packets to code with, it proceeds with routing only[5]. For each of the three cases, the queue lengths evolve over time as follows:

1) $U_i(t+1) = [U_i(t) - \mu_i(t)]^+ + A_i(t)$,
   $U_i^r(t+1) = U_i^r(t) + \min(\mu_i(t), U_i(t))$.
2) $U_i(t+1) = U_i(t) + A_i(t)$,
   $U_i^r(t+1) = [U_i^r(t) - \mu_i^r(\tau)]^+$.
3) $U_i(t+m) = U_i(t) + \sum_{\tau=t}^{t+m-1} A_i(\tau)$,
   $U_i^r(t+m) = \left[U_i^r(t) - \mu_i^r\left([\tau]_{\tau=t}^{t+m-1}\right)\right]^+$.

A. Maximum Throughput Region

Let Λ_{DNC} denote the stable throughput region with DNC, which is characterized next.

Theorem 1: A rate vector $\{\lambda_i\}_{i=1}^K$ lies in Λ_{DNC} if and only if there exist non-negative constants f_t, f_{rr}, f_{rn}, and ϕ_i for $i = 1, \ldots, K$ with $f_t + f_{rr} + f_{rn} = 1$ and $\phi_i \in [0, 1]$ for

[2] This is similar to the longest connected queue policy for multiple terminals transmitting to a single receiver [34].

[3] The analysis also extends to time-varying *dynamic* channel conditions. We skip this extension for brevity.

[4] Note that $\mu_i(t) = 0$, if $U_i(t) = 0$, and $\mu_i^r(t) = 0$, if $U_i^r(t) = 0$.

[5] Strategy 2) with plain routing is necessary for DNC to stabilize relay queues. It was shown in [35] that the two-terminal relay system that operates only with strategies 1) and 3) cannot sustain positive throughput with finite delay for both sources.

all i, that satisfy the following equations

$$\sum_{i=1}^{K} \frac{\lambda_i}{C_{i,r}} \leq f_t, \tag{5}$$

$$\sum_{i=1}^{K} \frac{\phi_i \lambda_i}{C_i^r} \leq f_{rr}, \tag{6}$$

$$\frac{(1-\phi_i)\lambda_i m}{C^r} \leq f_{rn},\ \forall i = 1, ..., K. \tag{7}$$

Proof: This proof follows from the relationship between the stability region and a multi-commodity flow problem as in [37]. The key difficulty here is that the capacity constraints from the relay to the terminals depend on if DNC or plain routing is used. This can be addressed by first viewing the choice between DNC and routing as a choice between two possible "routes" and noting that given the arrival rates, we can *a priori* decide what fraction of traffic is served via each route. In the theorem, ϕ_i denotes this fraction for relay transmissions of packets from terminal i. The flows are then subject to the "conservation equations" in (5)-(7), where f_t denotes the fraction of time the system operates in phase 1, f_{rr} denotes the fraction of time that the system operates in phase 2 using plain routing, and f_{rn} denotes the fraction of time the system operates in phase 2 using DNC. Note that (7) follows since, with DNC a flow can always be served at rate $\frac{C^r}{m}$ regardless of the number of other backlogged flows. ■

In the broadcast phase it may be possible to achieve the rate C_i^r individually for any terminal i by advanced channel coding and physical layer NC techniques [16], [18], [32] beyond DNC. In that case the rate region increases by replacing C_i in (7) with the node dependent rate C_i^r. There may be also a trade-off between C_i^r, $i = 1, ..., K$, i.e., the broadcast capacities form a region of their own. We do not consider such generalizations and continue assuming a common broadcast rate C^r for all terminals.

B. The Throughput Optimal Control

Next we give a backpressure-like control policy which dynamically makes scheduling and DNC decisions based on the queue backlogs.

DNC Backpressure Policy: For each time-slot t for which there is no on-going DNC session, compare $\{C_{i,r}\ [U_i(t) - U_i^r(t)]^+\}_{i=1}^K$, $\{C_i^r\, U_i^r(t)\}_{i=1}^K$ and $\frac{1}{m}\sum_{i=1}^{K} C^r\, U_i^r(t)$, and schedule

1) terminal i to transmit a packet from queue Q_i, if $C_{i,r}\ [U_i(t) - U_i^r(t)]^+$ is the largest,
2) the relay to transmit a plain packet from queue Q_i^r, if $C_i^r\, U_i^r(t)$ is the largest, or
3) the relay to transmit linearly network-coded packets from queues Q_i^r, $i = 1, \ldots, K$, over m time-slots, if $\frac{1}{m}\sum_{i=1}^{K} C^r\, U_i^r(t)$ is the largest.

***Theorem** 2:* The *DNC backpressure policy* is throughput optimal, i.e., it can stabilize the system for any arrival rates in the stability region Λ_{DNC} characterized in Theorem 1.

Proof: The standard for proof for such results relies on defining a Lyapunov function over the joint buffer-states and then showing that for sufficiently large backlogs, the expected one-step drift of this function is negative. The key difficulty in applying such an argument here is that whenever the relay decides to use DNC, there cannot be another control decision for m time-slots; however, when either a terminal transmits or the relay employs plain routing, one can make another control decision in the next time-slot. Furthermore, when DNC is chosen at time t, the queue backlog (and thus the one-step Lyapunov drift) is not well-defined at times $t+1, \ldots, t+m-1$. To circumvent this, we use a generalization that applies when the drift is evaluated using *state-dependent step-sizes* [38]. Specifically, we define the quadratic Lyapunov function $L(\mathbf{U}(t)) = \sum_{i=1}^{K} \left((U_i(t))^2 + (U_i^r(t))^2\right)$ for the set of queue backlogs $\mathbf{U}(t)$ in time-slot t. Let $g(\mathbf{U}(t))$ be the state-dependent step-size, such that $g(\mathbf{U}(\mathbf{t})) = m$ for any time t where the relay chooses DNC, and $g(\mathbf{U}(\mathbf{t})) = 1$ for any time t when either the relay employs plain routing or a terminal transmits. Since $g(\mathbf{U}(t))$ is bounded, we can show that the system is stable under the *DNC backpressure policy* provided that $\Delta(\mathbf{U}(t)) = \mathbb{E}\left[L(\mathbf{U}(t + g(\mathbf{U}(t))) - L(\mathbf{U}(t)) \mid \mathbf{U}(t)\right]$ is negative for sufficiently large backlogs [38].

The backlogs in terminal and relay queues Q_i and Q_i^r, $i = 1, \ldots, K$, over the $g(\mathbf{U}(t))$ time-slots into the future can be bounded in terms of the current backlogs as

$$\begin{aligned} & U_i(t + g(\mathbf{U}(t))) \\ & \leq \left[U_i(t) - \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1} \mu_i(\tau)\right]^+ + \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1} A_i(\tau), \end{aligned} \tag{8}$$

$$\begin{aligned} & U_i^r(t + g(\mathbf{U}(t))) \\ & \leq \left[U_i^r(t) - \mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right)\right]^+ + \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1} \mu_i(\tau), \end{aligned} \tag{9}$$

where the service rates satisfy $\mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right) \leq mC_i^r$ or $\mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right) = C^r$, when the relay proceeds with plain routing or DNC, respectively. The conditions (8) and (9) are upper bounds, because new arrivals may depart before the interval of $g(\mathbf{U}(t))$ time-slots is finished and $\mu_i(\tau)$ contributes to (9) only if $U_i(\tau) > 0$.

Note that for any non-negative real numbers V, U, μ and A, $V^2 \leq U^2 + \mu^2 + A^2 - 2U(\mu - A)$, if $V \leq [U - \mu]^+ + A$, as given in [37]. Also note that $\mu_i(t) \leq 1$, $\mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right) \leq m$, $A_i(t) \leq A_i^{max}$, $\mathbb{E}[g(\mathbf{U}(t))] \leq m$, $\mathbb{E}[g^2(\mathbf{U}(t))] \leq m^2$, and terminal and relay queues are separately served. Then, from (8)-(9) the Lyapunov drift is bounded as (10), where $B = m^2(2 + \sum_{i=1}^{K}(A_i^{max})^2)$ and $g(\mathbf{U}(t))$ and $h(\mathbf{U}(t))$ denote the remaining terms in (10). If the terminals are scheduled to transmit for a given realization of $\mathbf{U}(t)$, we have

$$h(\mathbf{U}(t)) \leq \max_{i=1,\ldots,K} \left(C_{i,r}\ [U_i(t) - U_i^r(t)]^+\right). \tag{12}$$

$$\Delta(\mathbf{U}(t)) \leq m^2\Big(2+\sum_{i=1}^{K}(A_i^{max})^2\Big) - 2\sum_{i=1}^{K}U_i(t)\mathbb{E}\Bigg[\sum_{\tau=t}^{t+g(\mathbf{U}(t))-1}\mu_i(\tau) - \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1}A_i(\tau)\Big|\mathbf{U}(t)\Bigg]$$
$$-2\sum_{i=1}^{K}U_i^r(t)\mathbb{E}\Bigg[\mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right) - \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1}\mu_i(\tau)\Big|\mathbf{U}(t)\Bigg] \tag{10}$$
$$:= B - 2g(\mathbf{U}(t))\Big(h(\mathbf{U}(t)) - \sum_{i=1}^{K}U_i(t)\lambda_i\Big) \tag{11}$$

On the other hand, if the relay is scheduled, we have

$$h(\mathbf{U}(t)) \leq \max\Bigg(\max_{i=1,\ldots,K}\left(C_i^r\,U_i^r(t)\right), \frac{1}{m}\sum_{i=1}^{K}C^rU_i^r(t)\Bigg). \tag{13}$$

From (12)-(13), it can be seen that the DNC backpressure policy minimizes $h(\mathbf{U}(t))$ at all times t. Now consider any rate $\{\lambda_i\}_{i=1}^K$ in the interior of the maximum throughput region Λ_{DNC}. For such a rate, there must exist a stationary control policy independent of the backlogs that, for an arbitrary small ϵ, satisfies the following flow constraints [37] for $i = 1, \ldots, K$:

$$\mathbb{E}\left[\mu_i(t) \mid \mathbf{U}(t)\right] = \epsilon + \lambda_i,$$
$$\mathbb{E}\Bigg[\mu_i^r\left([\tau]_{\tau=t}^{t+g(\mathbf{U}(t))-1}\right) - \sum_{\tau=t}^{t+g(\mathbf{U}(t))-1}\mu_i(\tau)\Big|\mathbf{U}(t)\Bigg] = m\epsilon.$$

Substituting these into (10) and (11), the Lyapunov drift $\Delta_s(\mathbf{U}(t))$ is bounded by

$$B - g_s(\mathbf{U}(t))\Big(h_s(\mathbf{U}(t)) - \sum_{i=1}^{K}U_i(t)\lambda_i\Big)$$
$$= B - 2m\epsilon\sum_{i=1}^{K}\left(U_i(t) + U_i^r(t)\right), \tag{14}$$

where the subscript s denotes this stationary policy.

Since the DNC backpressure policy maximizes $h(\mathbf{U}(t))$, it follows that

$$g_s(\mathbf{U}(t))\Big(h_s(\mathbf{U}(t)) - \sum_{i=1}^{K}U_i(t)\lambda_i\Big)$$
$$\leq mg_b(\mathbf{U}(t))\Big(h_b(\mathbf{U}(t)) - \sum_{i=1}^{K}U_i(t)\lambda_i\Big),$$

where the subscript b denotes the backpressure policy. Hence,

$$\Delta_b(\mathbf{U}(t)) \leq B - 2\epsilon\sum_{i=1}^{K}\left(U_i(t) + U_i^r(t)\right)$$

showing that when $\sum_{i=1}^{K}\left(U_i(t) + U_i^r(t)\right)$ is large enough, the Lyapunov drift under the backpressure policy must be negative and so the system is stable. ■

If random NC is used with finite field size, the coding block length, m, becomes a random variable and then we need to bound the expected Lyapunov drift, where the expectation is taken over m. As m has bounded second moment, the previous Lyapunov stability arguments follow and the DNC backpressure policy stabilizes the rate region using instantaneous values of m.

IV. Stochastic Network Control with Analog Network Coding

We now consider two transmission modes for the terminals. One mode corresponds to using ANC with a *scheduling matrix* $\tilde{G}$ corresponding to the minimum number of $m = \max(\lceil K/2\rceil, K-1-2n)$ transmissions, given in Lemma 1, in each of the multiaccess and broadcast phases. The relay includes dummy transmissions from all empty terminal queues. The second transmission mode is to simply schedule one terminal with amplify-and-forward (AF) relaying. We consider immediate forwarding at the relay without queueing so that every slot of terminal transmissions in multiaccess phase is followed immediately by a slot of relay transmission in broadcast phase.

The channel rates in the multiaccess and broadcast phases are fully *coupled* for each transmission mode. Therefore, we express the *end-to-end* rate for any terminal pair and partition it between the two phases to satisfy the Max-Flow Min-Cut condition. Given a scheduling matrix $\tilde{G}$, we define $C_{i\to j}(\tilde{G})$ as the end-to-end rate from terminal i to terminal j over two hops (via AF scheme) for $i \notin \mathcal{N}_j$.[6] The rate $C_{i\to j}$ strongly depends on the physical layer properties (e.g., studied in [32], [33] for AWGN channels). Then, the broadcast rate for packets from terminal i to the rest of terminals is given by $\tilde{C}_i(\tilde{G}) = \min_{j:j\notin\mathcal{N}_i} C_{i\to j}(\tilde{G})$, with overhearing taken into account.

The service rates of terminal queues are coupled in ANC and in any slot t they depend on $q_{bt}(t) = \sum_{i=1}^{K} 1_{\{U_i(t)>0\}}$, the number of backlogged terminal queues with at least one packet to transmit. If $q_{bt}(t) > m$, ANC or plain routing may achieve a higher rate depending on the channel capacities. If $q_{bt}(t) \leq m$, plain routing is optimal. We define $\mu_i\left([\tau]_{\tau=t}^{t+m-1}\right)$ as the service rate of terminal queue Q_i achievable over the interval $[\tau]_{\tau=t}^{t+m-1}$ of m time-slots.

In any time-slot, the scheduler chooses one from *two* possible transmission strategies: 1) terminal i transmits to the

[6] We can further improve the end-to-end rates by excluding terminals with empty queues from ANC and replacing the scheduling matrix with $\tilde{G}D$, where the diagonal matrix D is given by $D_{i,i} = 1$ for backlogged terminal i.

relay that amplifies and forwards the received signal over the next time-slot, 2) terminals transmit with scheduling matrix $\tilde{G}$ over m time-slots to the relay that amplifies and forwards the received signals over the next m time-slots. Here, ANC refers to joint analog NC and routing such that when there are not enough terminals with packets to transmit, backlogged terminals transmit separately and the relay proceeds with routing only. For each of the two cases, the queue lengths evolve over time as follows:

1) $U_i(t+2) = [U_i(t) - \mu_i(t)]^+ + \sum_{\tau=t}^{t+1} A_i(t)$,
2) $U_i(t+2m)$
$= \left[U_i(t) - \mu_i\left([\tau]_{\tau=t}^{t+2m-1}\right)\right]^+ + \sum_{\tau=t}^{t+2m-1} A_i(\tau)$.

Following any terminal transmission, no terminal can transmit (since the relay is) over the next time-slot, if plain routing is used, or over the next m time-slots, if ANC is used. Hence, in these times the queue dynamics do not have any departures.

A. The Maximum Throughput Region

Let Λ_{ANC} denote the stable throughput region with ANC, which is characterized next.

Theorem 3: A rate vector $\{\lambda_i\}_{i=1}^K$ lies in Λ_{ANC} if and only if there exist non-negative constants f_{tr} and f_{tn}, and ϕ_i for $i = 1, \dots, K$ with $f_{tr} + f_{tn} = 1$ and $\phi_i \in [0, 1]$ for all i, that satisfy the following equations

$$\sum_{i=1}^{K} \frac{\phi_i \lambda_i}{\tilde{C}_i([I_K]_i)} \le \frac{1}{2} f_{tr}, \tag{15}$$

$$\frac{(1-\phi_i)\lambda_i m}{\tilde{C}_i(\tilde{G})} \le \frac{1}{2} f_{tn}, \ \forall i = 1, \dots, K. \tag{16}$$

Proof: The proof of this theorem follows again from the relationship between the stability region and a multi-commodity flow problem as in [37]. In this case the key difficulty is that the end-to-end achievable rates depend on if ANC or simple AF is used. We again view this as a choice between two possible "routes". In the theorem, ϕ_i denotes the fraction of terminal i's traffic for each route. The resulting flows are then subject to the "conservation equations" in (15)-(16), where f_{tr} denotes the fraction of time that the system operates using plain routing and f_{tn} denotes the fraction of time the system operates using ANC. Note that (16) follows since, with ANC a flow from terminal i can always be served at rate $\frac{\tilde{C}_i(\tilde{G})}{2m}$ regardless of the number of other backlogged flows. ■

B. The Throughput Optimal Control

Next we give a backpressure-like control policy for our model which dynamically makes scheduling and ANC decisions based on the queue backlogs.

ANC Backpressure Policy: For each time-slot t for which there is no on-going ANC session, compare $\tilde{C}_i([I_K]_i)U_i(t)$ and $\frac{1}{m}\sum_{i=1}^{K} \tilde{C}_i(\tilde{G})\, U_i(t)$, and schedule

1) terminal i to transmit a packet from queue Q_i, if $\tilde{C}_i([I_K]_i)U_i(t)$ is the largest, or
2) terminals $i = 1, \dots, K$ to transmit cooperatively (i.e., they start ANC session) over m time-slots, if $\frac{1}{m}\sum_{i=1}^{K} \tilde{C}_i(\tilde{G})\, U_i(t)$ is the largest.

Theorem 4: The *ANC backpressure policy* is throughput optimal, i.e., it can stabilize the system for any arrival rates in the stability region Λ_{ANC} characterized in Theorem 3.

Proof: As with DNC, we use a state-dependent drift now over $2m$ time-slots for ANC and 2 time-slots for the AF case in plain routing, since this is the time before another control choice can be made. Let $2g(\mathbf{U}(t))$ be the state-dependent step-size, such that $g(\mathbf{U}(t)) = m$ for any time t when terminals transmit in ANC , and $g(\mathbf{U}(t)) = 1$ for any time t when one terminal transmits in plain routing. The backlogs in terminal queues Q_i, $i = 1, \dots, K$, over $2(g(\mathbf{U}(t))$ time-slots into the future can be bounded in terms of the current backlogs as

$$U_i(t + 2g(\mathbf{U}(t)))$$
$$\le \left[U_i(t) - \mu_i\left([\tau]_{\tau=t}^{t+2g(\mathbf{U}(t))-1}\right)\right]^+ + \sum_{\tau=t}^{t+2g(\mathbf{U}(t))-1} A_i(\tau),$$

where $\mu_i\left([\tau]_{\tau=t}^{t+2g(\mathbf{U}(t))-1}\right) \le m\tilde{C}_i([I_K]_i)$ or $\mu_i\left([\tau]_{\tau=t}^{t+2g(\mathbf{U}(t))-1}\right) = \tilde{C}_i(\tilde{G})$, when the terminal i transmits separately in plain routing or participates in an ANC session, respectively.

Consider the quadratic Lyapunov function $L(\mathbf{U}(t)) = \sum_{i=1}^{K}(U_i(t))^2$. Since $g(\mathbf{U}(t))$ is bounded, we can show that the system is stable under the *ANC backpressure policy* provided that $\Delta(\mathbf{U}(t)) = \mathbb{E}\left[L(\mathbf{U}(t + 2g(\mathbf{U}(t))) - L(\mathbf{U}(t)) \mid \mathbf{U}(t)\right]$ is negative for sufficiently large backlogs [38]. The rest of the proof is similar to the proof of Theorem 2 and omitted here for brevity. ■

C. Stable Throughput Comparison

Consider symmetric channels with $C_{i,r} = C_{r,i} = c_i$ for plain routing and DNC. For $K = 2$, Figure 3 shows the throughput region for DNC, ANC and plain routing. DNC expands the throughput region of plain routing and the potential gain of ANC depends on the AF-based achievable channel rates.

Next, we increase K and evaluate the physical channel effects in the symmetric case, where $\lambda_i = \lambda$ and $c_i = c$ for all i. We consider AWGN channels with channel rate $\mathcal{C}(\gamma) = \log(1+\gamma)$ for SNR γ. In DNC and plain routing, the channel rate between any terminal and the relay is $\mathcal{C}(\gamma)$. In ANC, the end-to-end rate $\tilde{C}_i$ is a function of scheduled transmissions. In particular, if one or two terminals transmit at a time for $K = 2$, the end-to end rate is $\mathcal{C}(\frac{\gamma^2}{2\gamma+1})$ or $\mathcal{C}(\frac{\gamma^2}{3\gamma+1})$, respectively. Figures 4 and 5 show how the achievable rates change with SNR and K, respectively. The corresponding sum rates are shown in Figures 6 and 7 for different values of SNR γ and K. As the number of terminals increases, new trade-offs emerge since more terminals contribute to the sum rate but each terminal is allocated less time for transmission. As shown in Figure 6 for high SNR, the sum rate of routing is

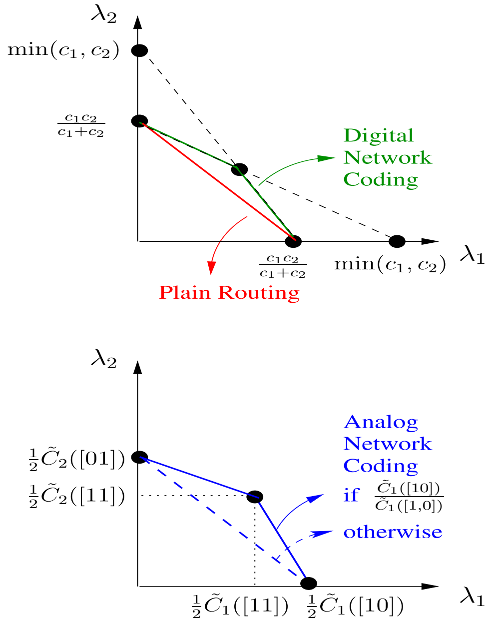

Fig. 3. Stable throughput region for $K = 2$ and $n = 0$.

smaller than the sum rate of ANC or DNC for any K and n. However, Figure 4 shows that the individual rates may be better with routing compared to ANC or DNC depending on the values of K and n.

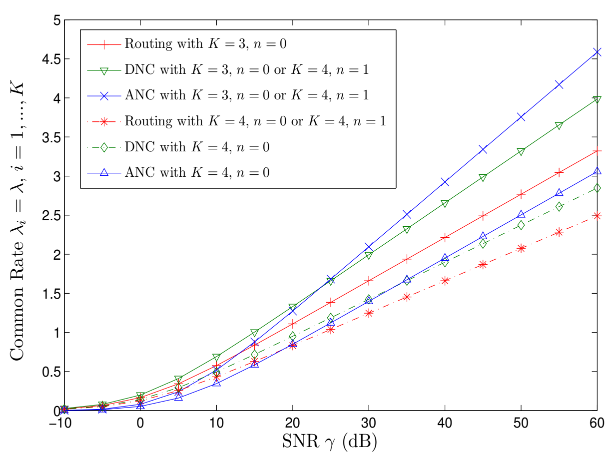

Fig. 4. Common rate for ANC, DNC and plain routing over AWGN channels as a function of SNR γ for different values of K and n.

In general, ANC amplifies the noise when relaying so it performs better as SNR increases. The individual rate achieved by ANC for each terminal exceeds the rate of plain routing at SNR γ_1^* and the rate of DNC at SNR $\gamma_2^* \geq \gamma_1^*$. Figure 8 illustrates γ_1^* and γ_2^* as a function of K and n. Both γ_1^* and γ_2^* increase with K, i.e., ANC needs higher SNR to mitigate the throughput decrease as the interference increases. In addition, both γ_1^* and γ_2^* decrease with n, i.e., ANC makes better use of overheard terminal transmissions in general.

We also evaluated the effects of routing and NC strategies on queueing dynamics as a function of packet arrival rate.

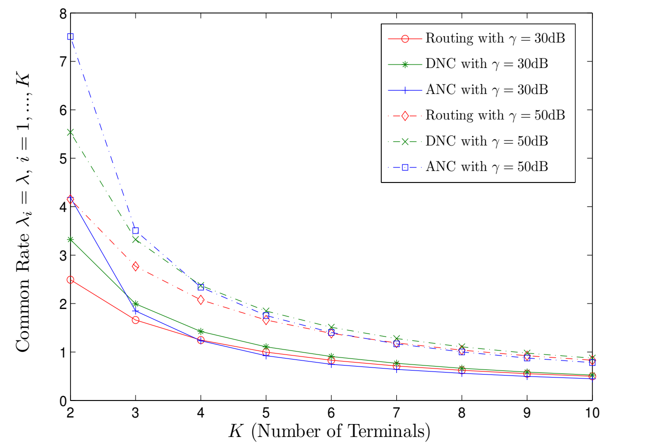

Fig. 5. Common rate for ANC, DNC and plain routing over AWGN channels as a function of K for different values of SNR γ, where $n = 0$.

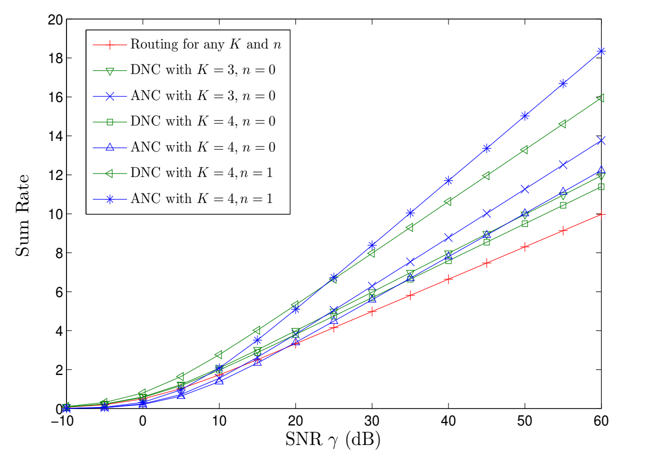

Fig. 6. um rate for ANC, DNC and plain routing over AWGN channels as a function of SNR γ for different values of K and n.

For different values of SNR, Figures 9 and 10 show the average packet delay for $K = 2$ and $K = 3$, respectively, under Poisson packet arrivals. For smaller values of SNR, DNC extends the maximum stable throughput and reduces the average packet delay, whereas ANC improves both stable throughput and packet delay for larger values of SNR.

V. Extension to Multiple Relays

We next consider *multihop* operation with *multiple* relays as in Figure 11. Let the network be connected and the set of nodes be denoted by $\mathcal{N}$. Nodes with one neighbor are called the *end terminals*. Nodes with more than one neighbors are called *relays*, the set of which is denoted by $\mathcal{R}$. Each node (relay or end-terminal) is a source with broadcast packets for all others. We focus on DNC here. A similar approach applies to ANC, where relays do not decode packets and end-terminals are the destinations. We define $\mathcal{S}_r$ as the set of one-hop neighbors of relay $r \in \mathcal{R}$ (that have packets to be relayed by relay r). $\mathcal{S}_r$ is predetermined by a given set of paths $\mathcal{P}$ without cycles. Packets of nodes in $\mathcal{S}_r$ are exchanged over relay $r \in \mathcal{R}$. For

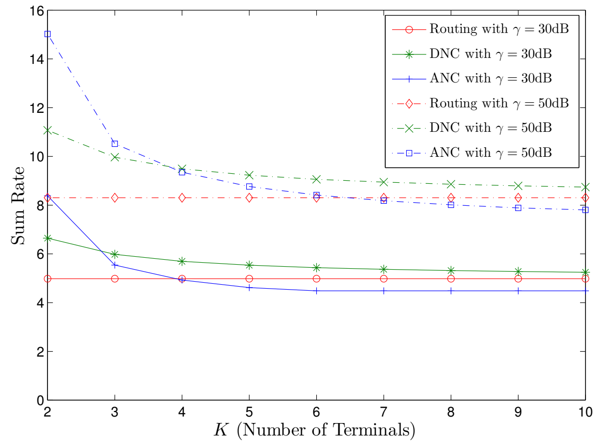

Fig. 7. Sum rate for ANC, DNC and plain routing over AWGN channels as a function of K for different values of SNR γ, where $n = 0$.

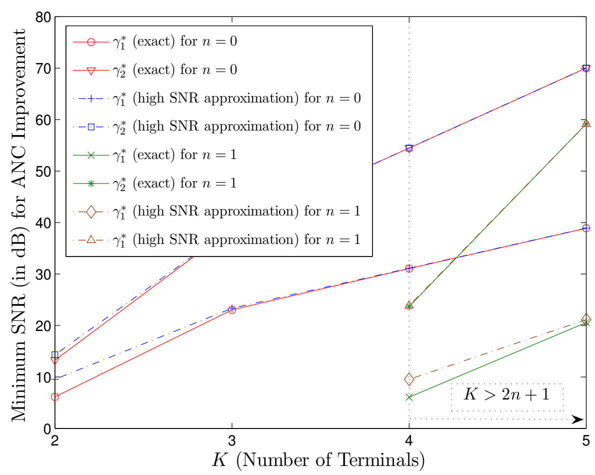

Fig. 8. Minimum SNR values for ANC to improve the rate over DNC and plain routing for different values of K and n.

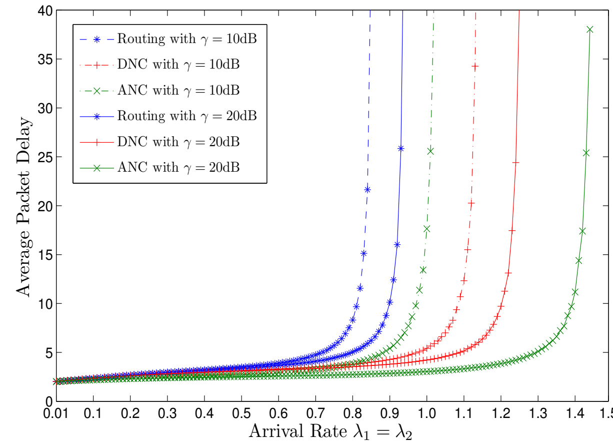

Fig. 9. Average packet delay as a function of arrival rate for different values of SNR (where $K = 2$).

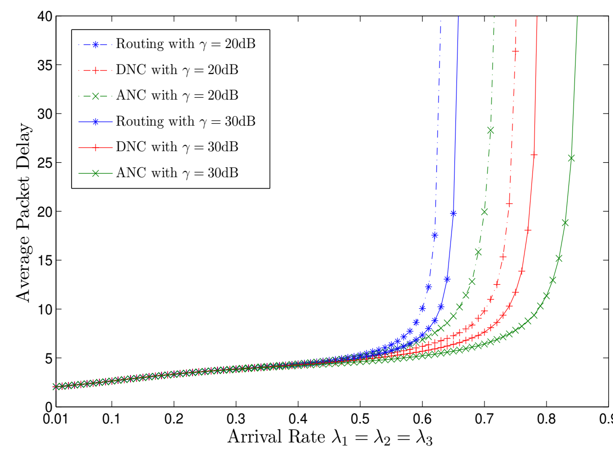

Fig. 10. Average packet delay as a function of arrival rate for different values of SNR (where $K = 3$).

any relay r, we define $\tilde{\lambda}_{i,r}$ as the total rate of traffic from $i \in \mathcal{S}_r$ that is relayed by r. This rate includes self-generated traffic and relay traffic incoming from neighbors such that $\tilde{\lambda}_{i,r} = \lambda_i$, $i \notin \mathcal{R}$, $i \in \mathcal{S}_r$, $r \in \mathcal{R}$, and $\tilde{\lambda}_{i,r} = \lambda_i + \sum_{l \in \mathcal{S}_i \setminus \{r\}} \tilde{\lambda}_{l,i}$, $i \in \mathcal{R}$, $i \in \mathcal{S}_r$, $r \in \mathcal{R}$.

The *source* and *relay* traffic can be transmitted separately (without the need to code across the source and relay traffic). Therefore, we define Q_i^i as the queue at (relay or end terminal) node i for the self-generated packets and Q_j^r as the queue at relay r for the relay packets incoming from node $j \in \mathcal{S}_r$. The backlogs of queues Q_i^i and Q_j^r in time-slot t are $U_i^i(t)$ and $U_j^r(t)$, respectively. We define $\mathcal{N}_k^r$ as the set of nodes $\mathcal{S}_r$, $r \in \mathcal{R}$, that a node $k \in \mathcal{S}_r$ can overhear, and we define $m_r = |\mathcal{S}_r| - 1 - 2n_r$, where n_r is the number of hops any node $i \in \mathcal{S}_r$ can overhear on each side from the neighbors of relay $r \in \mathcal{R}$.

The rate regions (5)-(7) for the single relay generalize to multiple relays as follows. For each relay r, there are four types of transmissions that need to be separated over time to satisfy the half-duplex condition: 1) relay r transmits its own packets to $\mathcal{S}_r$ with rate $C_r = \min_{i \in \mathcal{S}_r} C_{r,i}$ over time fraction f_1^r, 2) end terminal $i \in \mathcal{S}_r$ transmits to relay r with rate $C_{i,r}$ over time fraction $f_{2,i}^r$,[7] 3) relay r transmits uncoded

[7]This model can also be extended to simultaneous transmissions from end terminals in $\mathcal{S}_r$ to relay r.

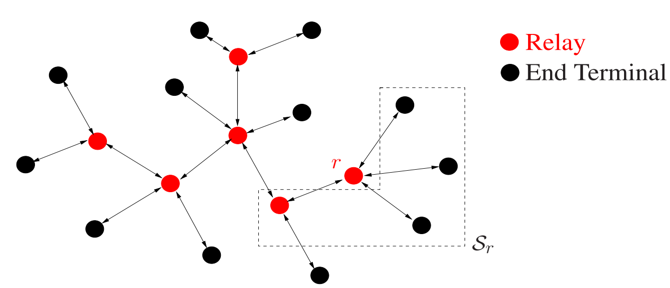

Fig. 11. General topology with multiple relays $\mathcal{R}$, where $\mathcal{S}_r$ is the set of one-hop neighbors of relay r.

packet from node $i \in \mathcal{S}_r$ with rate $C_i^r = \min_{k: k \in \mathcal{S}_r \setminus \{i\}, i \in N_k^r} C_{r,k}$ over time fraction $f_{3,i}^r$, and 4) relay r transmits network-coded packets with rate $C^r = \min_{i \in \mathcal{S}_r} C_i^r$ over time fraction f_4^r. The throughput region Λ_{DNC} of DNC is achieved by time sharing the rate regions $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$ and $\mathcal{R}_4$ for transmissions 1), 2), 3) and 4), respectively. Let $\phi_i^r \in [0,1]$ denote the fraction of time where node i's packets are relayed by relay r via plain routing. Then, the rate regions are given by $\mathcal{R}_1 = \left\{\lambda_r \geq 0 : \frac{\lambda_r}{C_r} \leq f_1^r\right\}$, $\mathcal{R}_2 = \left\{\lambda \geq 0 : \frac{\lambda_i}{C_{i,r}} \leq f_{2,i}^r, i \in \mathcal{S}_r\right\}$, $\mathcal{R}_3 = \left\{\lambda \geq 0 : \frac{\phi_i^r \tilde{\lambda}_{i,r}}{C_i^r} \leq f_{3,i}^r, i \in \mathcal{S}_r\right\}$, and $\mathcal{R}_4 = \left\{\lambda \geq 0 : \frac{(1-\phi_i^r)\tilde{\lambda}_{i,r} m_r}{C^r} \leq f_4^r, i \in \mathcal{S}_r\right\}$.

The throughput optimal control requires a new formulation of differential backlogs:

1) For transmissions from $r \in \mathcal{R}$ to $\mathcal{S}_r$, we define, $B_1^r(t) = \left[U_r^r(t) - \sum_{k: r \in \mathcal{S}_k, k \in \mathcal{R} \setminus \{r\}} U_k^r(t)\right]^+$,
2) For transmissions from i (where $i \in \mathcal{S}_r$ and $i \notin \mathcal{R}$) to $r \in \mathcal{R}$, we define $B_{2,i}^r(t) = \left[U_i^i(t) - U_i^r(t)\right]^+$,
3) For uncoded transmissions of packets of $i \in \mathcal{S}_r$ from $r \in \mathcal{R}$ to $\mathcal{S}_r \setminus \{i\}$, we define $B_{3,i}^r(t) = \left[U_i^r(t) - \sum_{k: r \in \mathcal{S}_k \setminus \{i\}, k \in \mathcal{R} \setminus \{i,r\}} U_k^r(t)\right]^+$, and
4) For coded transmissions from $r \in \mathcal{R}$, we define $B_4^r = \frac{1}{m_r} \sum_{i \in \mathcal{S}_r} \left[U_i^r(t) - \sum_{k: r \in \mathcal{S}_k \setminus \{i\}, k \in \mathcal{R} \setminus \{i,r\}} U_k^r(t)\right]^+$.

In time-slot t, throughput optimal control follows from choosing the set of transmitting nodes to bound the time-average value of total queue length $\sum_{i \in \mathcal{N}} U_i^i(t) + \sum_{r \in \mathcal{R}} \sum_{i \in \mathcal{S}_r} U_i^r(t)$ by maximizing

$$\begin{aligned} W(t) &= \sum_{r \in \mathcal{R}} W_1^r(t) B_1^r(t) + \sum_{r \in \mathcal{R}, i \in \mathcal{S}_r, i \notin \mathcal{R}} W_{2,i}^r(t) B_{2,i}^r(t) \\ &\quad + \sum_{r \in \mathcal{R}, i \in \mathcal{S}_r} W_{3,i}^r(t) B_{3,i}^r(t) + \sum_{r \in \mathcal{R}} W_4^r(t) B_4^r(t), \end{aligned}$$

where we set the weights $W_1^r(t) = C_r$, $W_{2,i}^r(t) = C_{i,r}$, $W_{3,i}^r(t) = C_i^r$ and $W_4^r(t) = C^r$ if the corresponding transmissions are successful; otherwise, they are set to zero.

For illustration purposes, we consider a ring network of $N > 3$ relays, each attached to K different end terminals. Each node generates traffic broadcast to the rest at the common rate of λ. For each relay, only K end terminals and two relays on both sides may overhear each other depending on the overhearing range n (where $m = K + 1 - 2n > 0$). Because of half-duplex model, nodes are scheduled to transmit over four time fractions of equal length, three for relays (without interference at neighbors on both sides) and one for end terminals. Using the shortest paths, every node receives traffic with rate $\left(\frac{N-3}{2}\right)(K+1)\lambda$ from each neighbor. In plain routing, the time for each relay is separated to (*i*) transmissions of end-terminals to relay with rate $K\lambda$, (*ii*) transmissions of self-generated traffic of relay with rate λ, and (*iii*) relaying of traffic incoming from two relays on both sides with total rate $(N-3)(K+1)\lambda$ and from K terminals with total rate $K\lambda$. Then, the maximum throughput achieved by plain routing is $\lambda = \left(4((N-2)(K+1)+K)\right)^{-1}$.

In DNC, the set of neighbors with m largest rates includes two neighbor relays and $K - 1 - 2n$ end terminals, if $K \geq 2n + 1$, and each relay needs to carry with DNC the traffic incoming at rate $\big((N-3)(K+1) + K - 1 - 2n\big)\lambda$. Then, the maximum throughput achieved by DNC is $\lambda = \left(4((N-2)(K+1) + K - 1 - 2n)\right)^{-1}$. DNC improves the rate of plain routing and this gain increases with overhearing range n. The maximum gain is achieved for $K = 2n$ when DNC carries the relay traffic with rate $\left(\frac{N-3}{2}\right)(K+1)\lambda$ and DNC maximizes the throughput to $\lambda = \left(4(\frac{N-1}{2})(K+1)\right)^{-1}$.

VI. Conclusions

We have derived the maximum throughput region for relay-assisted wireless broadcast with digital and analog network coding. For stochastic packet traffic, variations of the maximum differential backlog policy are shown to be throughput optimal. We started with the case of a single relay and extended the analysis to the simultaneous operation of multiple relays. The results quantify the stable throughput gain of digital and analog network coding over plain routing. Throughput optimal control relies on centralized scheduling with full queue backlog information. Future work should look at *distributed* control with limited or delayed information on queue and channel properties. In addition, it would be interesting to analyze the trade-offs of stable throughput with *delay* and *energy* measures with joint network coding and scheduling.

Appendix: Proof of Lemma 1

A. Plain Routing

Plain routing requires K packets to be separately delivered to the relay in phase 1 and then forwarded to terminals over additional K transmissions in phase 2. Therefore, the minimum number of transmissions $T_R(K, n)$ for plain routing is given by (1).

B. Digital Network Coding

Phase 1 requires K transmissions. In Phase 2, each transmission represents a linear equation in terms of the K packets. If $n = 0$, each user requires $K-1$ independent equations to be able to solve for the $K-1$ needed packets. This requires that coding matrix G has at least $K-1$ rows. $G = [I_{K-1}, \mathbf{1}_{K-1}]$ meets this bound as an example, where $\mathbf{1}_{K-1}$ is the $(K-1)$-dimensional column vector of all 1's. This coding matrix G has the minimum number of $K-1$ rows needed to have the rank of $K-1$. If terminals overhear each other's transmissions over n neighbor hops in phase 1, $2n$ degrees of freedom are delivered to each terminal in phase 1. Here, a degree of freedom received at node i represents the change in the rank of coding matrix (one row for the coding coefficients of each received packet). Then, at least $K-1-2n$ transmissions are necessary in phase 2 to deliver $K-1$ linearly independent coded packets to each terminal i.

Next, we show that $K - 1 - 2n$ coded transmissions are sufficient to form the decoding matrix of rank K at any terminal. Define g_i as the ith column of G, and define matrix $\hat{G}_i$ by removing the consecutive columns g_i and g_j, for all $j \in N_i$, from G. If $\hat{G}_i$ has full column rank, terminal i can decode $K-2n-1$ packets missing from phase 1. The optimal coding matrix is not unique. We can use the generator matrix of $[K, K-1-2n]$ Maximum Distance Separable (MDS) Code for a sufficiently large field size, since any $K-1-2n$ columns of this matrix are independent [39]. Here, we can also use random NC to construct coding matrix G with entries uniformly and randomly chosen from a finite field. As the field size grows to infinity, any m columns out of matrix G become independent from each other and $\hat{G}_i$ becomes full-rank.

To construct G for a general field size, we define the $r_i \times s_i$ matrix $U_i = [\, I_{r_i}, \ldots, I_{r_i}, V_i \,]$, where $r_{i+1} = \text{mod}(r_i, s_{i+1})$ and $s_{i+1} = \text{mod}(s_i, r_i)$ follows extended Euclidean algorithm with the initial conditions of $r_0 = K-1-2n$ and $s_0 = K$. The $r_i \times s_{i+1}$ matrix V_i is defined as $V_i = [I_{s_{i+1}}, \ldots, I_{s_{i+1}}, U_{i+1}^T]^T$, where $\{\cdot\}^T$ is the matrix transpose. The $(K-1-2n) \times K$ coding matrix is given by $G = U_0$. As an example, for $K = 4$ and $n = 1$, we have $m = 1$, $r_0 = 1$, $s_0 = 4$, $s_1 = 0$ (V_0 does not exist), and the corresponding coding matrix is $G = [1\ 1\ 1\ 1]$, and for $K = 5$ and $n = 1$, we have $m = 2$, $r_0 = 2$, $s_0 = 5$, $s_1 = 1$, $r_1 = 0$, and the corresponding coding matrix is $G = [I_2, I_2, V_0]$, where $V_0 = [1, 1]^T$.

Any $K-2n-1$ adjacent columns of matrix G are independent, i.e., the matrix $\hat{G}_i$ is full-rank, and any terminal i can decode $K-1-2n$ missing packets. Therefore, the minimum number of transmissions $T_{DNC}(K, n)$ for DNC is given by (2).

C. Analog Network Coding

For $n = 0$, the relay needs to deliver $K-1$ degrees of freedom in phase 2 to each terminal over at least $K-1$ transmissions, i.e., $T_{ANC}(K, n) \geq 2(K-1)$. If the scheduling matrix $\tilde{G}$ is chosen as $[I_{K-1}, \mathbf{1}_{K-1}]$, each terminal i receives $K-1$ linearly independent combinations of K signals and can decode $K-1$ missing signals by using its own signal. Then, $T_{ANC}(K, n)$ is $2(K-1)$. For $n > 0$, we define $\mathbf{Y}_i = H_i \mathbf{X}$ as the received signals, where $\mathbf{X}$ denotes the transmitted signals, one signal from each terminal, and H_i is the end-to-end transfer matrix at terminal i. For random channel gains, the rank of H_i is $K-1$ with probability one, if all terminals transmit over $K-1$ slots in phase 1. However, continuous transmissions do not allow packet overhearing because of half-duplex constraints on the terminals. At most $2n$ degrees of freedom can be delivered to each terminal by overhearing in phase 1 and we need at least $K-1-2n$ transmissions in phase 2 such that $T_{ANC}(K, n) \geq 2(K-1-2n)$. Since at least K transmissions are needed to exchange K packets and each transmission of phase 1 is repeated in phase 2, $T_{ANC}(K, n) \geq 2\lceil K/2 \rceil$. The optimal scheduling matrix $\tilde{G}$ is not unique.

Consider the $m \times K$ scheduling matrix $\tilde{G} = [U, V]$, where $m = \max(\lceil K/2 \rceil, K-1-2n)$, and the $m \times m$ matrix U and the $m \times r$ matrix V are given by $U_{i,j} = 1$, if $i = j$ or $i = m-s+1, \ldots, m$ and $j = 1, \ldots, n+1$, else $U_{i,j} = 0$, and $V = [\, I_r, \ldots, I_r, W^T \,]^T$, where $r = K - m$, $s = \text{mod}(m, r)$, and $W_{i,j} = 1$ if $i = 1, \ldots, s$ and $j = r-n, \ldots, r$, else $W_{i,j} = 0$. As an example, for $K = 4$ and $n = 1$, we have $m = 2$, $r = 2$, $s = 0$, (W does not exist), and the corresponding scheduling matrix is $\tilde{G} = [I_2, I_2]$, and for $K = 5$ and $n = 1$, we have $m = 3$, $r = 2$, $s = 1$ ($W = [1, 1]$), and the rows of the corresponding scheduling matrix $\tilde{G}$ are $\tilde{G}_1 = [1, 0, 0, 1, 0]$, $\tilde{G}_2 = [0, 1, 0, 0, 1]$, and $\tilde{G}_3 = [1, 1, 1, 1, 1]$.

Note that $\tilde{G}$ ensures that all terminals transmit once over a period of m time-slots. After the first $2(m-s)$ slots (i.e., after $m-s$ terminal transmissions, each followed by forwarding at the relay), terminals $1, \ldots n+1$ and $K-n, \ldots, K$ can decode $K-s$ packets. Then, these terminals continuously transmit for the rest of s time-slots (since they do not need to overhear transmissions of other terminals) and decode s more packets from relay transmissions. The rest of terminals (i.e., terminals $n+2, ..., K-n-1$) miss at most m degrees of freedom from the packets transmitted in the first $2(m-s)$ slots and they also need to decode $s-1$ of remaining packets (since every terminal knows its own packet). So, at most $2s-1$ packets are missing at these terminals. Each of them transmits once in the last $2s$ slots and receives $2s-1$ degrees of freedom to decode the remaining (at most $2s-1$) packets. Then, $s-1$ of these packets are decoded from the terminal transmissions that are overheard (no overhearing is possible in the time-slot when the terminal transmits) and s of these packets are decoded from the relay transmissions.

References

[1] 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, July 2000.

[2] X. Yan, R. W. Yeung, and Z. Zhang, "The Capacity Region for Multi-source Multi-sink Network Coding, *Proc. IEEE International Symposium on Information Theory*, 2007.

[3] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of Linear Coding in Network Information Flow," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2745-2759, Aug. 2005.

[4] Y. Wu, P. Chou, Q. Zhang, K. Jain, W. Zhu, and S. Kung, "Network Planning in Wireless Ad Hoc Networks: A Cross-layer Approach," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 136-150, Jan. 2005.

[5] D. S. Lun, N. Ratnakar, M. Medard, R. Koetter, D. R. Karger, T. Ho, E. Ahmed, and F. Zhao, "Minimum-cost Multicast over Coded Packet Networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2608-2623, June 2006.

[6] Y. E. Sagduyu and A. Ephremides, "Cross-Layer Optimization of MAC and Network Coding in Wireless Queueing Tandem Networks," *IEEE Trans. Inform. Theory*, vol. 54, no. 2, pp. 554-571, Feb. 2008.

[7] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the Air: Practical Wireless Network Coding," *IEEE/ACM Trans. Networking*, vol. 16, no. 3, June 2008.

[8] S. Katti, S. Gollakota, and D. Katabi, "Embracing Wireless Interference: Analog Network Coding," *Proc. Sigcomm*, 2007.

[9] S. Zhang, S. C. Liew, and P. P. Lam, "Hot Topic: Physical-Layer Network Coding," *Proc. MobiCom*, 2006.

[10] D. Traskov, N. Ratnakar, D. S. Lun, R. Koetter, and M. Medard, "Network Coding for Multiple Unicasts: An Approach based on Linear Optimization," *Proc. IEEE International Symposium on Information Theory*, 2006.

[11] T. Ho, Y. Chang and K. J. Han, "On Constructive Network Coding for Multiple Unicasts", *Proc. Allerton Conference on Communication, Control and Computing*, 2006.

[12] C.-C Wang and N. B. Shroff, "Pairwise Intersession Network Coding on Directed Networks," *IEEE Trans. Information Theory*, vol. 56, no. pp. 3879-3900, Aug. 2010.

[13] S. Sengupta, S. Rayanchu, and S. Banerjee, "Network Coding-Aware Routing in Wireless Networks," *IEEE/ACM Trans. Networking*, pp. 1028-1036, Apr. 2010.

[14] P. Popovski and H. Yomo, "Wireless Network Coding by Amplify-and-Forward for Bi-Directional Traffic Flows," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 16-18, Jan. 2007.

[15] Y. Hao , D. Goeckel, Z. Ding, D. Towsley, and K. K. Leung, "Achievable Rates for Network Coding on the Exchange Channel," *Proc. IEEE Military Communications Conference*, 2007.

[16] F. Xue, C.-H. Liu, and S. Sandhu, "MAC-Layer and PHY-Layer Network Coding for Two-Way Relaying: Achievable Rate Regions and Opportunistic Scheduling," *Proc. Allerton Conference on Communication, Control and Computing*, 2007.

[17] S. Katti, I. Maric, A. Goldsmith, D. Katabi, and M. Medard, "Joint Relaying and Network Coding in Wireless Networks, *Proc. IEEE International Symposium on Information Theory*, 2007.

[18] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast Capacity Region of Two-Phase Bidirectional Relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454-458, 2008.

[19] M. P. Wilson, K. Narayanan, H. Pfister, and A. Sprintson."Joint Physical Layer Coding and Network Coding for Bi-Directional Relaying," *IEEE Trans. Information Theory*, vol. 56, no. 11, pp. 5641-5654, Nov. 2010.

[20] S. M. S. Tabatabaei Yazdi, S. A. Savari, and G. Kramer, "Network Coding in Node-constrained Line and Star Networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4452-4468, July 2011.

[21] F. Wang, S. C. Liew and D. Guo, "Wireless MIMO Switching with Zero-forcing Relaying and Network-coded Relaying," *IEEE Journal on Selected Areas in Communications*, this issue.

[22] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling for Maximum Throughput in Multihop Radio Networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936-1949, Dec. 1992.

[23] S. Sarkar and L. Tassiulas, "A Framework for Routing and Congestion Control for Multicast Information Flows," *IEEE Trans. Inform. Theory*, vol. 48, no. 10, pp. 2690-2708, Oct. 2002.

[24] T. Ho and H. Viswanathan, "Dynamic Algorithms for Multicast with Intra-session Network Coding," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 797-815, Feb. 2009.

[25] A. Proutiere and P. Chaporkar, "Adaptive Network Coding and Scheduling for Maximizing Throughput in Wireless Networks," *Proc. ACM Mobicom*, 2007.

[26] A. Khreishah, C.-C. Wang, and N. B. Shroff, "Rate Control with Pairwise Inter-session Network Coding," *IEEE Trans. Networking,* vol. 18, no. 3, pp. 816-829, June 2010.

[27] T. Oechtering and H. Boche, "Stability Region of an Optimized Bidirectional Regenerative Half-duplex Relaying Protocol," *IEEE Trans. Communications*, vol. 56, no. 9, pp. 1519-1529, Sept. 2008.

[28] E. N. Ciftcioglu, A. Yener, and R. A. Berry, "Stability of Bi-Directional Cooperative Relay Networks," *Proc. IEEE ITW*, 2008.

[29] A. Eryilmaz, D. S. Lun, and B. T. Swapna, "Control of Multi-hop Communication Networks for Inter-session Network Coding," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 1092-1110, Feb. 2011.

[30] R. Knopp, "Two-way radio networks with a star topology," *Proc. IEEE International Zurich Seminar on Communications,* Zurich, Switzerland, Feb. 2006, pp. 154157.

[31] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index Coding with Side Information," *Proc. IEEE Symposium on Foundations of Computer Science*, 2006.

[32] P. Popovski and T. Koike-Akino, "Coded Bidirectional Relaying in Wireless Networks," *New Directions in Wireless Communications*, Springer, 2009.

[33] Y. E. Sagduyu, D. Guo, and R. A. Berry, "On the Delay and Throughput of Digital and Analog Network Coding for Wireless Broadcast," *Proc. Conference on Information Sciences and Systems, CISS'08,* Princeton, NJ, March 2008.

[34] L. Tassiulas and A. Ephremides, "Dynamic Server Allocation to Parallel Queues with Randomly Varying Connectivity," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 466-478, Mar. 1993.

[35] X. He and A. Yener, "On the Energy Delay Trade-off of a Two-Way Relay Network," *Proc. Conference on Information Sciences and Systems, CISS'08,* Princeton, NJ, March 2008.

[36] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ, 1992.

[37] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource Allocation and Cross-Layer Control in Wireless Networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-144, 2006.

[38] S. Foss and T. Konstantopoulos, "Lecture Notes on Stochastic Stability by Lyapunov Methods," Heriot-Watt University, Edinburgh.

[39] F. J. MacWilliams and N. J. Sloane, The Theory of Error-Correcting Codes, North-Holland, 1977.

PLACE PHOTO HERE

Yalin E. Sagduyu (S'02-M'08) received the B.S. degree from Bogazici University, Turkey, in electrical and electronics engineering and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park.

Dr. Sagduyu is currently a Program Manager in Intelligent Automation Inc., Rockville MD, and a Visiting Assistant Research Scientist in the Institute for Systems Research at the University of Maryland, College Park. He has been a Graduate Research Assistant with the Institute for Systems Research at the University of Maryland and a Postdoctoral Fellow in the Department of Electrical Engineering and Computer Science at Northwestern University. His research interests are in the areas of design and optimization of wireless networks, multi-user communications, network coding, information theory, network security, game theory, and socio-technological networks.

PLACE PHOTO HERE

Randall A. Berry (S'93-M'00) received the B.S. degree in Electrical Engineering from the University of Missouri-Rolla in 1993 and the M.S. and PhD degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1996 and 2000, respectively.

In 2000, he joined Northwestern University, where he is currently a Professor in the Department of Electrical Engineering and Computer Science. In 1998 he was on the technical staff at MIT Lincoln Laboratory in the Advanced Networks Group. His primary research interests include wireless communication, data networks, and information theory.

Dr. Berry is the recipient of a 2003 NSF CAREER award. He has served as an Editor for the IEEE Transactions on Wireless Communications from 2006 to 2009, and an Associate Editor for the IEEE Transactions on Information Theory from 2009 to 2011, in the area of communication networks. He was a co-chair of the 2012 IEEE Communication Theory Workshop and a technical co-chair of 2010 IEEE ICC Wireless Netowrking Symposium. He has also been a guest editor for the IEEE Journal on Selected Topics in Signal Processing special issue on Dynamic Spectrum Access and the IEEE Transactions on Information Theory special issue on Relaying and Cooperation.

PLACE PHOTO HERE

Dongning Guo (S'97-M'05-SM'11) joined the faculty of Northwestern University, Evanston, IL, in 2004, where he is currently an Associate Professor in the Department of Electrical Engineering and Computer Science. He received the B.Eng. degree from the University of Science & Technology of China, the M.Eng. degree from the National University of Singapore, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ. He was an R&D Engineer in the Center for Wireless Communications (now the Institute for Infocom Research), Singapore, from 1998 to 1999. He has held visiting positions at Norwegian University of Science and Technology in summer 2006 and in the Institute of Network Coding at the Chinese University of Hong Kong in 2010–2011. He is an Associate Editor of the IEEE Transactions on Information Theory and an Editor of Foundations and Trends in Communications and Information Theory.

Dongning Guo received the Huber and Suhner Best Student Paper Award in the International Zurich Seminar on Broadband Communications in 2000 and is a co-recipient of the IEEE Marconi Prize Paper Award in Wireless Communications in 2010 (with Y. Zhu and M. L. Honig). He is also a recipient of the National Science Foundation Faculty Early Career Development (CAREER) Award in 2007. His research interests are in information theory, communications, and networking.