DMT of Multihop Networks: End Points and Computational Tools

K. Sreeram, S. Birenjith, and P. Vijay Kumar, *Fellow, IEEE*

***Abstract*—In this paper, the diversity-multiplexing gain tradeoff (DMT) of single-source, single-sink (ss-ss), multihop relay networks having slow-fading links is studied. In particular, the two end-points of the DMT of ss-ss full-duplex networks are determined, by showing that the maximum achievable diversity gain is equal to the min-cut and that the maximum multiplexing gain is equal to the min-cut rank, the latter by using an operational connection to a deterministic network. Also included in the paper, are several results that aid in the computation of the DMT of networks operating under amplify-and-forward (AF) protocols. In particular, it is shown that the colored noise encountered in amplify-and-forward protocols can be treated as white for the purpose of DMT computation, lower bounds on the DMT of lower-triangular channel matrices are derived and the DMT of parallel MIMO channels is computed. All protocols appearing in the paper are explicit and rely only upon AF relaying. Half-duplex networks and explicit coding schemes are studied in a companion paper.**

***Index Terms*—Amplify-and-forward protocols, cooperative diversity, degrees of freedom, deterministic network, diversity-multiplexing gain tradeoff, explicit codes, multihop networks, parallel channel, relay networks.**

I. Introduction

In fading relay networks, cooperative diversity provides an efficient means of network operation. While much of the work in the literature on cooperative diversity is based on two-hop networks, the attention here is on multihop networks.

A. Prior Work

The concept of user cooperative diversity was introduced in [3]. Cooperative diversity protocols were first discussed in [5] for the two-hop, single-relay network (see Fig. 1 for example of a general multirelay network). Zheng and Tse [6] proposed the diversity-multiplexing gain tradeoff (DMT) as a means of evaluating point-to-point, multiple-antenna schemes in the context of slow-fading channels.

Manuscript received July 22, 2008; revised October 13, 2010; accepted May 02, 2011. Date of current version February 08, 2012. This work was supported in part by the NSF-ITR Grant CCR-0326628, in part by the DRDO-IISc Program on Advanced Mathematical Engineering, and in part by Motorola's University Research Partnership Program with IISc. The material in this paper was presented in part at the 10th International Symposium on Wireless Personal Multimedia Communications, Jaipur, India, December 2007, at the Information Theory and Applications Workshop, San Diego, CA, January 2008, and at the 2008 IEEE International Symposium on Information Theory.

K. Sreeram is with the Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, IL 61820 USA (e-mail: kannan1@illinois.edu).

S. Birenjith is with the Department of ECE, Indian Institute of Science, Bangalore, 560012 India (e-mail: birenjith@gmail.com).

P. V. Kumar is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India (e-mail: vijayk@usc.edu).

Communicated by M. C. Gastpar, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2011.2173243

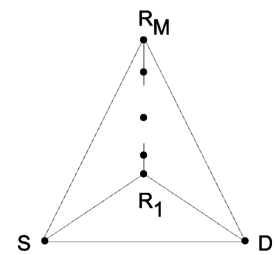

Fig. 1. Two-hop cooperative relay network.

1) Two-Hop Networks: Several papers [7], [8], [11], [12] have studied the DMT of two-hop relay networks (Fig. 1). The proposed schemes for two-hop relay networks include the nonorthogonal amplify-and-forward protocol [4], [8], selection decode-and-forward [7], dynamic-decode-and-forward [8], [9], compress-and-forward [2], [11] and slotted amplify-and-forward protocols [16]. However, the DMT of even the simplest of these networks was undetermined until recently, when the DMT of the single relay network was solved by using a quantize and re-encode protocol [24].

Other papers have considered two-hop relay networks in the absence of a direct-link between the source and destination [10], [26]. In a parallel and independent work [28], [29], the DMT of the two-hop network without direct link is proved to be equal to the cut-set bound.

2) Multihop Networks: Yang and Belfiore in [15] consider AF protocols for a family of MIMO multihop networks. They derive the DMT of the Rayleigh-product channel which they prove is equal to the DMT of the AF protocol applied to this channel. They also propose AF protocols to achieve the maximum diversity gain of these multiantenna layered networks.

Oggier and Hassibi [37] propose distributed space time codes for multiantenna layered networks that achieve diversity gain equal to the minimum number of relay nodes among the hops. Recently, Vaze and Heath [39] have constructed distributed space time codes based on orthogonal designs, that achieve the maximum diversity gain of the multiantenna layered network with low decoding complexity. In [40], the same authors study the circumstances under which full diversity gain can be achieved without coding in a layered network in the presence of partial CSIT.

Borade, Zheng, and Gallager in [25] consider AF schemes for a class of multihop layered networks where each layer has the same number of relays. They show that AF strategies are optimal in terms of multiplexing gain. They also compute lower bounds on the DMT of the product Rayleigh channel.

3) Capacity: There has been recent interest in determining approximations to the capacity of wireless networks. The pre-log coefficient of the capacity, termed as the degrees-of-freedom (DOF) of wireless multiantenna networks is studied in [42] and [43].[1] The DOF of the N-user interference channel was derived in [32], of the MIMO X network in [33], [34] and the DOF of single-source, single-sink (ss-ss) layered networks was determined in [25].

In [31], given a wireline network and an associated network code, a compute-and-forward scheme for the corresponding wireless Gaussian relay channel is derived, wherein each relay computes linear transformations of its input signals, and the achievable rate region for this scheme is determined.

In a different direction, deterministic wireless networks, which model the effect of interaction between various nodes have been studied [30]. The capacity of ss-ss and multicast deterministic wireless networks has been characterized there. Intuition drawn from the deterministic wireless networks is then used to identify capacity to within an additive constant for Gaussian relay networks in the same work. It is also pointed out there that this result extends to give the compound-channel capacity to within a constant number of bits for full-duplex networks.

4) Other Work: Cooperative networks with asynchronous transmissions have also been studied in the literature [53]–[55]. Codes for two-hop cooperative networks having low decoding complexity and full diversity are studied in [56], [55] and [57].

B. Setting and Channel Model

1) Notation: In this paper, we use uppercase letters to denote matrices and lowercase letters to denote either a vector or a scalar. Thus vectors and scalars are differentiated only through context. Boldface letters are used to identify a random entity, regardless of whether the particular random entity is a scalar, vector or matrix.

2) Network Representation: Any wireless network can be associated with a directed graph, with vertices representing nodes in the network and edges representing connectivity between nodes. A bidirectional edge should be represented by two edges, one pointing in each direction; however for simplicity we will draw one undirected edge to denote the same. In the wireless networks considered here, relay nodes operate in either half or full-duplex mode. Under half-duplex operation, a node cannot listen and transmit simultaneously.

Between any two adjacent nodes v_x, v_y of a wireless network, we assume the following channel model:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{1}$$

where $\mathbf{y}$ corresponds to the received vector at node v_y, $\mathbf{w}$ is the noise vector, $\mathbf{H}$ is a matrix and $\mathbf{x}$ is the vector transmitted by the node v_x.

3) Assumptions: We follow the literature in making the assumptions listed below. Our description is in terms of the equivalent, complex-baseband, discrete-time channel.

1) All channels are assumed to be quasi-static and to experience Rayleigh fading, i.e., all fade coefficients are i.i.d., circularly-symmetric, complex Gaussian $\mathbb{C}\mathcal{N}(0,1)$ random variables.
2) The additive noise at each receiver is also modeled as possessing an i.i.d., circularly-symmetric, complex Gaussian $\mathbb{C}\mathcal{N}(0,1)$ distribution.
3) Each receiver (but none of the transmitters) is assumed to have perfect channel state information of all the upstream channels in the network.[2]
4) We impose the following energy constraint on the vector $\mathbf{x}$ (of length n) transmitted by the source
$$\mathrm{Tr}\,(\Sigma_x) := \mathrm{Tr}\,(\mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\}) \leq n\rho$$
where Tr denotes the trace operator and Σ_x is the covariance matrix of $\mathbf{x}$.
5) We will impose the constraint that the relays operate under the same power constraint ρ as the source.

C. Results

In this paper, we characterize the maximum diversity, the maximum multiplexing gain and the achievable DMT of full-duplex cooperative networks. Some of these results were presented in conference versions of this paper [17]–[20] (see also [21] and [23]). Special classes of half-duplex networks are considered in a companion paper, [1]. The design of explicit codes that achieve the DMT of all protocols proposed in the present paper, can also be found there.

The principal results established in this paper are given below (see Table I-C for a tabular listing).

1) The maximum diversity gain achieved for any specific source-destination pair in a general multiterminal network with many sources, many sinks and multiple-antenna nodes is equal to the value of the min-cut between the particular source-destination pair. Moreover the maximum diversity gain can be achieved for all source-destination pairs simultaneously.
2) The maximum multiplexing gain (MMG) of a ss-ss full-duplex, multiantenna network is equal to the minimum rank of any cut between the source and the destination.
3) A DMT which is linear between the maximum diversity gain and maximum multiplexing gain is achievable for directed-acyclic, full-duplex, single-antenna, ss-ss networks.

We also establish the following general results, that are useful in computing the DMT of cooperative networks operated under AF protocols:

1) the colored noise encountered in cooperative networks can be treated as white for the purpose of DMT computation,

[1]The degrees-of-freedom is alternately referred to as the maximum multiplexing gain in the literature, although the former is typically used for ergodic capacity characterizations while the latter is typically used in the context of outage characterization. This paper deals with the DMT, which is an outage characterization and for this reason, we use the term maximum multiplexing gain.

[2]However, in the protocols proposed in the current paper, CSIR is utilized only at the sink, as all the relay nodes are required to simply amplify and forward the received signal.

TABLE I
PRINCIPAL RESULTS SUMMARY

Network	No of source-sink pairs	No of antennas in nodes	Full-Duplex/ Half-Duplex (FD/HD)	Direct Link	Upper bound on Diversity/DMT $d_{\text{bound}}(r)$	Achievable Diversity/DMT $d_{\text{achieved}}(r)$	Is upper bound achieved?	Reference
Arbitrary	Multiple	Multiple	FD/HD	✓	$d(0) =$ Min-cut	$d(0) =$ Min-cut	✓ ($d_{\max}$ achieved)	Thm. 3.1
Arbitrary	Single	Multiple	FD	✓	$r_{\max}$ = Rank of Min-cut	$r_{\max}$ = Rank of Min-cut	✓ ($r_{\max}$ achieved)	Thm. 3.3
Arbitrary Directed Acyclic Networks	Single	Single	FD	✓	Concave in general	$d_{\max}(1-r)^+$	A linear DMT between $d_{\max}$ and $r_{\max}$ is achieved	Thm. 5.1

2) a lower bound on the DMT of channels associated with triangular matrices is provided,
3) the DMT of a parallel MIMO channel is computed in terms of the DMT of its constituent MIMO links.

D. Relation to Existing Literature

1) Proof of a Conjecture by Rao and Hassibi: The results in Example 4 in Section IV-C prove Conjecture 1 appearing in [26] and [27].
2) Lower bound on the DMT of various AF Protocols: We present a simple means of deriving in a concise and intuitive manner, lower bounds on the DMT of two-hop relay networks operating under either the NAF [8], SAF [16] or MIMO-NAF protocols [14]. The lower bound in the case of the first two protocols is tight, while the third lower bound, improves upon the best-known bound.
3) The diversity gain of arbitrary cooperative networks. The result on maximum diversity gain attainable by an arbitrary multiterminal network, is derived by making use of an argument that while obvious in retrospect, was not noted earlier. The maximum diversity gain of certain specific network families has previously been derived: in the case of the MIMO two-hop relay channel in [14], in the case of layered networks in [15] while upper bounds for general networks are derived in [52]. In parallel work, [28] have used a different approach to characterize the diversity gain of general ss-ss networks. In work that has appeared subsequent to the initial submission of the present paper, low-complexity full-diversity codes have been designed in [39] and [40].
4) DMT of single-antenna full-duplex networks The compound channel results contained in [30], can be used to establish that the DMT of full-duplex networks is equal to the cut-set bound. It must be noted though, that the schemes in [30] involve complicated schemes with random codes of large block length, in contrast with the AF protocols presented here that use short block-length, explicit coding schemes.
5) The DMT of the parallel channel in closed form is provided here in Lemma 4.1. A special case of this result is derived in [15] where the authors characterize the parallel channel DMT for the particular case when all the individual channels have the same DMT.
6) Polynomial-time algorithms for constructing zero-error codes for the layered linear deterministic channel with a block-length of 1 have been recently developed in [38]. These codes can be used in conjunction with our results, to derive low-complexity coding schemes for the fading relay channel, that are optimal in terms of achieving the MMG.

E. Outline

In Section II, we present some background on the DMT, and prove two results that aid in the computation of the DMT of a network, one showing that noise can be treated as white for the purpose of DMT computation and the other computing the MMG of a MIMO channel whose entries are polynomial functions of a common pool of i.i.d. Rayleigh random variables. In Section III, we determine the extreme points of the DMT of arbitrary ss-ss networks by drawing upon the results in Section II. In Section IV, we provide additional computational tools for evaluating the DMT of AF networks. In particular, we establish the DMT of a bank of parallel MIMO channels. Next, a lower bound on the DMT of channels characterized by lower triangular block matrix is presented. The section concludes by applying these results to determine lower-bounds on the DMT of several networks. In Section V, we present a protocol for operating an arbitrary ss-ss network with single-antenna, full-duplex relays, and establish a lower bound on the associated DMT.

A companion paper [1] to the present submission, makes use of the basic results and techniques introduced here to characterize the DMT of certain classes of half-duplex networks as well as provide DMT-achieving code designs.

Expanded versions of some of the abbreviated proofs contained in the present paper can be found in [22].

II. Basic Results for Cooperative Networks

We begin by reviewing the notion of DMT in point-to-point channels and then go on to explain how the DMT becomes a meaningful tool in the study of cooperative wireless networks.

We then go on to develop general techniques, which will prove useful in deriving results on the DMT of ss-ss networks.

A. Background

1) Diversity-Multiplexing Gain Tradeoff: Let us assume a power constraint of ρ at each node, and scale the rate of transmission with ρ as $R = r\log(\rho)$. A protocol $\wp$ is a method of operation of the network, defined for every value of power constraint ρ and data rate R, such that the power constraint is satisfied at all nodes. Let the protocol over blocks of n symbols of data, be such that the transmitted vector at the source is $\vec{x} = x^n$, and let the corresponding received vector under the protocol, at the destination be $\vec{y} = y^n$. Then the protocol essentially induces a channel $p_{\vec{y}|\vec{x}}$ between the transmitted source block and the destination block. The probability of outage under the operating protocol $\wp$ is then given as

$$P_{\text{out}}(\wp, r, \rho) = \Pr(I(\vec{\mathbf{x}}; \vec{\mathbf{y}}) < nr\log\rho | \mathbf{H}(\wp) = H(\wp))$$

where $\mathbf{H}(\wp)$ denotes the collection of all random variables associated with all the channels in the network. The outage exponent, $d_{\text{out}}(\wp, r)$, of the network under the protocol $\wp$ is a function of the multiplexing gain r and is defined by

$$d_{\text{out}}(\wp, r) = -\lim_{\rho\to\infty} \frac{\log P_{\text{out}}(\wp, r, \rho)}{\log(\rho)}$$

and we will indicate this by writing

$$\rho^{-d_{\text{out}}(\wp, r)} \doteq P_{\text{out}}(\wp, R).$$

The symbols $\dot{\geq}$, and $\dot{\leq}$ are similarly defined.

Given a protocol $\wp$ and a coding scheme for the network that achieves a probability-of-error $P_e(\wp, r, \rho)$, the diversity-multiplexing gain tradeoff $d(\wp, r)$ achieved by the coding scheme is defined by

$$d(\wp, r) = \lim_{\rho\to\infty} \frac{\log P_{\text{e}}(\wp, R)}{\log(\rho)}.$$

Random coding can be shown to achieve a probability-of-error P_e whose rate of exponential decay with ρ matches that of the outage probability, i.e.

$$P_{\text{e}}(\wp, r, \rho) \doteq P_{\text{out}}(\wp, R). \tag{2}$$

The DMT $d(\wp, r)$ achieved by the protocol is thus equal to the outage exponent, i.e., $d(\wp, r) = d_{\text{out}}(\wp, r)$. The DMT of the network, $d(r)$ is the supremum of the DMT over all protocols that satisfy the power constraints.

Definition 1: Given a random matrix $\mathbf{H}$ of size $m \times n$, we define the DMT of the matrix $\mathbf{H}$ as the optimal DMT of the associated channel $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$ where $\mathbf{y}$ is an m-length received column vector, $\mathbf{x}$ is an n-length transmitted column vector with total power constraint ρ, and $\mathbf{w}$ is a $\mathbb{C}\mathcal{N}(0, I)$ column vector. We denote the DMT of matrix $\mathbf{H}$ by $d_H(.)$.

2) Cut-Set Bound on DMT: For any network, the cut-set upper-bound on mutual information of a general multiterminal network [46], [47] translates into an upper bound on the DMT. This can be formalized as follows [11]:

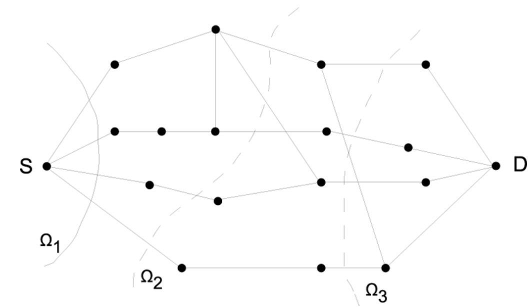

Fig. 2. Some example cuts in a network. The cut Ω_1 turns out to be an example of a min-cut.

Lemma 2.1: Let $r\log(\rho)$ be the rate of communication between the source and the sink. Given a cut ω between source and destination, let $\mathbf{H}_\omega$ denote the transfer matrix between nodes on the source-side of the cut and those on the sink-side, and let $d_\omega(r)$ be the DMT of $\mathbf{H}_\omega$. Then the DMT of the network is upper bounded as

$$d(r) \leq \min_{\omega\in\Lambda}\{d_\omega(r)\}$$

where Λ is the set of all cuts between the source and the destination.

In the example network shown in Fig. 2, it turns out that Ω_1 is the min-cut.

3) Amplify and Forward Protocols: Our attention here will be restricted to AF protocols since as we shall see, this class of protocols can often achieve the DMT of a network. Under an AF protocol $\wp$, each relay node is permitted to linearly process the incoming signals prior to its transmission. It is assumed that the node will scale its transmitted symbols so as to meet its power constraint. Such protocols induce a linear channel model between source and sink of the form

$$\mathbf{y} = \mathbf{H}(\wp)\mathbf{x} + \mathbf{w} \tag{3}$$

where $\mathbf{y} \in \mathbb{C}^m$ denotes the signal received at the sink, $\mathbf{w}$ is the noise vector, $\mathbf{H}(\wp)$ is the $(m \times n)$ induced channel matrix, and $\mathbf{x} \in \mathbb{C}^n$ is the vector transmitted by the source. The noise vector $\mathbf{w}$ is in general, colored, and we will deal with this issue in Section II-B.

It turns out that the specific value of the constant used by the nodes in order to ensure the power constraint is met, does not affect the DMT of the protocol, since the DMT is a high SNR characterization (see [8] for a detailed explanation). Without loss of accuracy therefore, we will assume this constant to be equal to 1.

B. White in the Scale of Interest

In this section, we prove a result for AF protocols, that will be extensively used in all future sections: Theorem 2.3, which states that noise, even though correlated, can be treated as white in the scale of interest and also state a known result (Lemma

2.4), which says that i.i.d. Gaussian inputs are sufficient to attain the outage exponent of any Gaussian channel of the form $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$.

If $\mathbf{h}$ is complex Gaussian, i.e., a $\mathbb{C}\mathcal{N}(0,1)$ random variable, then we have, for any given ϵ and ρ

$$Pr\{|\mathbf{h}|^2 > \rho^{\epsilon}\} = \exp(-\rho^{\epsilon}).$$

It turns out that a similar statement holds even when we replace h by a polynomial in several complex Gaussian random variables.

Lemma 2.2: Let $\{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M\}$ be a collection of i.i.d. complex Gaussian random variables. Let $f \in \mathbb{C}[X_1, X_2, \ldots, X_M]$ be a polynomial without constant term, in the variables $\{X_i\}$. Then there exist $A > 0$, $B > 0$, $d > 0$, $\delta > 0$ such that

$$Pr\{|f(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M)|^2 > k\} \leq A\exp(-Bk^{\frac{1}{d}}), \forall\, k \geq \delta$$

where the constants A, B, d, δ depend only on f.

Proof: See Appendix A. ■

We are now ready to establish that if the noise covariance matrix has a certain structure, then it can be considered as white noise for the purpose of DMT computation.

Theorem 2.3: Consider a channel of the form $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$, with $\mathbf{z} = \mathbf{z}_0 + \sum_{i=1}^{M} \mathbf{G}_i\mathbf{z}_i$, where the $\{\mathbf{z}_i\}_{i=0}^{M}$ are i.i.d. $\mathbb{C}\mathcal{N}(\underline{0}, I)$ random vectors, and where each entry of the $N \times N$ random matrix $\mathbf{G}_i$ is a polynomial function of some underlying i.i.d. complex Gaussian random variables $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_L$. Then the DMT of the channel $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$ is the same as the DMT of the channel $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$ where $\mathbf{w}$ is now a $\mathbb{C}\mathcal{N}(\underline{0}, I)$ random vector.

Proof: See Appendix B. ■

We now state a known result on input distributions that achieve the DMT:

Lemma 2.4: [6] For any channel that is of the form $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$ where $\mathbf{w}$ is white Gaussian noise, i.i.d. Gaussian inputs are sufficient to attain the DMT of the channel.

The noise that we deal with in this paper will always satisfy the conditions laid down in Theorem 2.3 and hence we can assume without loss of generality, that

- the noise is white, and
- the transmitted signal has an i.i.d. Gaussian distribution.

C. Maximum Multiplexing Gain

In this section, we derive the MMG of a MIMO channel matrix with each entry of the matrix being a polynomial function of certain complex Gaussian random variables. We begin by deriving certain properties of polynomial functions of Gaussian random variables and will subsequently use these characteristics to determine the MMG.

Lemma 2.5: Let $p \in \mathbb{R}[X]$ be a nonconstant polynomial of degree d, and let k, m be given real numbers. Consider the set $\mathcal{R}$ of all $x \in \mathbb{R}$ for which the following two conditions are satisfied:

$$|p(x)| \leq k, \tag{4}$$

$$|p'(x)| \geq m \tag{5}$$

where $p'(x)$ is the derivative of $p(x)$. Then this subset $\mathcal{R}$ of $\mathbb{R}$ can be expressed as the union

$$\mathcal{R} = \cup_{i=1}^{L} R_i \tag{6}$$

of disjoint intervals $R_i = [a_i, b_i]$. Furthermore, $L \leq 2d$.

Proof: See Appendix C. ■

Lemma 2.6: Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ be a collection of independent (real) Gaussian random variables. Let $f \in \mathbb{R}[X_1, X_2, \ldots, X_N]$ be a polynomial in the variables $\{X_i\}$. Then there exist constants $A > 0$, $d > 0$, $K > 0$ such that

$$\Pr\{|f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| < \delta\} \leq A\delta^{\frac{1}{d}}, \ \forall\, 0 \leq \delta < K$$

where the constants A, d, K depend only on f and not on δ.

Proof: See Appendix D. ■

We will now make use of this lemma to determine the MMG.

Definition 2: Given a random matrix $\mathbf{H}$, we define the structural rank, $\mathbb{R}\text{ank}(\mathbf{H})$, as the maximum rank attained by $\mathbf{H}$ over all possible realizations.

Theorem 2.7: Let $\mathbf{h}_1, \ldots, \mathbf{h}_N$ be independent complex Gaussian random variables, and let $\mathbf{H}$ be a random matrix, whose entries are polynomial functions of the random variables $\mathbf{h}_i$. Consider a channel of the form $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{w}$ are n-length column vectors representing the transmitted signal, received signal and the noise vector respectively, with the noise being white, i.e., distributed as $\mathbb{C}\mathcal{N}(0, I)$. Then the MMG of the channel, denoted by D, is given by the structural rank of $\mathbf{H}$, i.e.,

$$D = \mathbb{R}\text{ank}(\mathbf{H}).$$

Proof: Let m be the structural rank of $\mathbf{H}$. Clearly, for any given $\mathbf{H}$, the MMG is upper-bounded by the rank of $\mathbf{H}$, which does not exceed m. Next, we will show that a MMG of m is achievable. Consider transmission at a multiplexing gain of $r = m - \delta$. Since $\mathbf{H}$ is of structural rank m, there is a $m \times m$ submatrix $\mathbf{H}_m$ of structural rank m. Then $\mathbf{H}_m\mathbf{H}_m^{\dagger}$ is a principal submatrix of $\mathbf{H}\mathbf{H}^{\dagger}$.

Let $\lambda_i(A)$ denote the ith smallest eigenvalue of matrix A. Using the inclusion principle ([48, Theorem 4.3.15]), we get $\lambda_1(\mathbf{H}_m\mathbf{H}_m^{\dagger}) \leq \lambda_{n-m+1}(\mathbf{H}\mathbf{H}^{\dagger}), \ldots, \lambda_i(\mathbf{H}_m\mathbf{H}_m^{\dagger}) \leq \lambda_{n-m+i}(\mathbf{H}\mathbf{H}^{\dagger}), \ldots, \lambda_m(\mathbf{H}_m\mathbf{H}_m^{\dagger}) \leq \lambda_n(\mathbf{H}\mathbf{H}^{\dagger})$. Also we have $\lambda_1(\mathbf{H}\mathbf{H}^{\dagger}) = \ldots = \lambda_{n-m}(\mathbf{H}\mathbf{H}^{\dagger}) = 0$, since $\mathbf{H}\mathbf{H}^{\dagger}$ has at most rank m. Therefore, we obtain that

$$\log\det(I + \rho\mathbf{H}\mathbf{H}^{\dagger}) \geq \log\det(I + \rho\mathbf{H}_m\mathbf{H}_m^{\dagger}).$$

Therefore, the outage exponent for rate $r = m - \delta$ is

$$\begin{aligned}\rho^{-d_{out}(r)} &= \Pr\{\log\det(I + \rho\mathbf{H}\mathbf{H}^\dagger) < r\log\rho\}\\ &\leq \Pr\{\log\det(I + \rho\mathbf{H}_m\mathbf{H}_m^\dagger) < r\log\rho\}\\ &\leq \Pr\{\det(\rho\mathbf{H}_m\mathbf{H}_m^\dagger) < \rho^{(m-\delta)}\}\\ &= \Pr\{|\det(\mathbf{H}_m)|^2 < \rho^{-\delta}\}.\end{aligned}$$

Let $\mathbf{H}$, and thereby $\mathbf{H}_m$, be a function of the independent complex Gaussian random variables $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M$. Let us denote the real and imaginary parts of this collection of complex Gaussian random variables by $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, where $N = 2M$. Now the $\{\mathbf{x}_i\}$ are i.i.d. Gaussian random variables, i.e., they are distributed as $\mathcal{N}(0, 0.5)$. Then $|\det(\mathbf{H}_m)|^2$ is a nonzero real polynomial $p(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ in the $\{\mathbf{x}_i\}$. Since $p = |\det(\mathbf{H}_m)|^2$ is positive, $|p(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| = p(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$.

We can now use Lemma 2.6 to obtain that

$$\Pr\{|\det(\mathbf{H}_m)|^2 < \rho^{-\delta}\} \leq A\rho^{-\delta/d} \tag{7}$$

for some positive constants A, d, K with $\rho^{-\delta} < K$. Let $\rho_0^{-\delta} = K$. Then we can see that (7) is valid for all $\rho > \rho_0$.

This leads to

$$\begin{aligned}&\rho^{-d_{out}(m-\delta)} \dot{\leq} \rho^{-\delta/d}\\ \Rightarrow\ & d_{out}(m-\delta) \geq \delta/d > 0.\end{aligned}$$

Thus a MMG of m is achievable. ■

III. Characterization of Extreme Points of DMT of Arbitrary Networks

The focus in this section, is on multihop networks. We show that the min-cut is equal to the maximum diversity gain for arbitrary multiterminal networks with multiantenna nodes irrespective of whether the relays operate under the half-duplex constraint or not. We also show for ss-ss full-duplex networks that the maximum multiplexing gain is equal to the min-cut rank. These two results put together, characterize the two end-points of the DMT of full-duplex ss-ss networks.

A. Min-Cut Equals Diversity

Theorem 3.1: Consider a multiterminal fading network with channels undergoing i.i.d. complex Gaussian fading. The maximum diversity gain achievable for any source-destination (s-d) pair is equal to the min-cut between the source and the destination. Each s-d pair can achieve its maximum diversity gain simultaneously.

Proof: We first consider the case where there is only a single s-d pair. We will handle the case of single and multiple-antenna nodes separately.

Case I: Network With Single-Antenna Nodes: Let the source be S_i and sink be D_j. Let Λ_{ij} denote the set of all cuts between S_i and D_j.

From the cut-set upper bound on DMT (Lemma 2.1), $d(0) \leq m$, where $m := \min_\omega m_\omega$ is the min-cut. It suffices to prove that a diversity gain equal to m is achievable. We know from Menger's theorem in graph theory (see for eg. [50]), that the number of edges in the min-cut is equal to the maximum number of edge-disjoint paths between source and the sink. Given a collection of edge-disjoint paths, we schedule the network in such a way that each edge in a given edge-disjoint path is activated one by one. The same is repeated for all the edge-disjoint paths. This in effect creates a parallel channel between the source S_i and destination D_j. Let the number of edges along the ith edge-disjoint path be n_i. Let the fading coefficient on the jth edge in the ith edge-disjoint path be $\mathbf{h}_{ij}$. Now define $\mathbf{h}_i := \prod_{j=1}^{n_i} \mathbf{h}_{ij}$, $i = 1, 2, \ldots, m$. The induced parallel channel therefore contains m links, with the fading coefficients $\mathbf{h}_i$ on the link i. Thus, the equivalent channel seen by a symbol is

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & 0 & \ldots & & 0\\ 0 & \mathbf{h}_2 & 0 & & 0\\ \vdots & & \vdots & \ddots & \\ 0 & \ldots & \ldots & & \mathbf{h}_m \end{bmatrix}.$$

This is a parallel channel with all the channels being independent of each other and the DMT of the channels being identical. Therefore we can use Corollary 4.3 and obtain the DMT of the parallel channel as

$$d_H(r) = (m - r)^+. \tag{8}$$

This DMT can be achieved by using a DMT-optimal parallel channel code [14], [41], [13].

The protocol utilizes N time instants to induce this effective channel matrix, and therefore, the DMT of the protocol can be given in terms of the DMT of the channel matrix as

$$d(r) = d_H(Nr) = (m - Nr)^+ \tag{9}$$

with $N := \Sigma_{i=1}^m n_i$. Hence the maximum achievable diversity gain is m.

Case II: Network With Multiple-Antenna Nodes: In the multiple antenna case, we regard any link between a n_t-transmit-antenna node and an n_r-receive-antenna node as being composed of $n_t n_r$ links, with one link between each transmit and each receive antenna. Note that it is possible to selectively activate precisely one of the $n_t n_r$ transmit-antenna receive-antenna pairs by appropriately transmitting from just one antenna and listening at just one receive antenna. Now, the same strategy as in the single antenna case can then be applied to achieve a diversity gain equal to the min-cut in the network.

Thus the proof is complete for the single flow from S_i to D_j.

When there are multiple flows in the network, we simply schedule the data of all the flows in a time-division manner. This will entail a rate loss—however, since we are interested only in the diversity, we can still achieve each flow's maximum diversity gain simultaneously. ■

B. Maximum Multiplexing Gain Equals Minimum Rank

In this section, we determine the maximum multiplexing gain (MMG) for multiantenna ss-ss networks to be equal to the min-cut rank. For ss-ss networks with single-antennas, the MMG cannot exceed one in value, because the source has a single antenna and the cut with source at one side and the rest of the nodes on the other side will yield an upper bound of value one for the MMG. It is possible to attain the optimal MMG of 1 by activating one path between the source to the destination either using amplify-and-forward or a decode-and-forward strategy. However, the MMG-optimal strategy becomes unclear when the number of antennas is greater than 1.

We use results from recent work on linear deterministic networks [30] to arrive at strategies for achieving the MMG of a fading network. We use simple algebraic techniques to lift the achievability strategies for linear deterministic networks to fading networks. We will now introduce linear deterministic networks, and then move on to connect strategies from the linear deterministic network to the fading network.

1) Deterministic Networks: Linear deterministic networks, which will be called deterministic networks henceforth for brevity, were introduced in [30] as a model for wireless networks. In deterministic networks, the effect of noise is suppressed and the focus is on how wireless signals transmitted by various nodes interact. Two fundamental effects of wireless transmission are captured in the model: broadcast and interference. Broadcasting is captured by the fact that the information transmitted by any node affects the received information at all destinations to which the node is connected. Interference is captured by the addition of signals over a certain finite field. The finite capacity of each link is captured by the fact that each node can transmit a certain number of symbols from a finite field, which get linearly transformed and received at the destination.

More formally, the vector $\mathbf{y}_i$ received by a terminal i can be given in terms of the transmitted vectors $\mathbf{x}_j$ of various terminals by

$$\mathbf{y}_i = \sum_{j \in \mathrm{In}(i)} G_{ij}\mathbf{x}_j$$

where $\mathbf{y}_i$ and $\mathbf{x}_j$ are q-length column vectors over $\mathbb{F}_p$, and G_{ij} is a $q \times q$ transfer matrix between the node i and node j, whose entries are drawn from $\mathbb{F}_p$. Every cut ω in the deterministic network is associated with a channel matrix, which we will denote by G_ω.

The theorem below from [30], computes the capacity[3] of a ss-ss linear deterministic wireless network.

Theorem 3.2: [30] Given a linear deterministic ss-ss wireless network over any finite field $\mathbb{F}_p$, the capacity C of such a relay network is given by

$$C = \min_{\omega \in \Omega} \mathrm{rank}(G_\omega)$$

where Ω is the collection of all cuts between source and sink and where the capacity is specified in terms of the number of finite field symbols per unit time. A strategy utilizing only linear transformations over $\mathbb{F}_p$ at the relays is sufficient to achieve this capacity.

The capacity-achieving strategy in [30] utilizes matrix transformations of the input vector received over a period of T time slots at each relay. This process continues for L blocks, therefore the total number of time instants required for the scheme is $M := LT$. The achievability shows the existence of relay matrices A_i at each relay node $i \in |\mathcal{V}|$, where $\mathcal{V}$ is the set of vertices in the graph. The relay matrix A_i is of size $qT \times qT$, and represents the transformation between the received and transmitted vectors at the particular relay node.

[3]We use the term capacity to signify ϵ-error capacity, as is conventional.

2) MMG of SS-SS Networks: The main result of this section is given below.

Theorem 3.3: Given a ss-ss multiantenna wireless network, with i.i.d. Rayleigh fading coefficients, the MMG of the network is given by

$$D = \min_{\omega \in \Omega} \mathbb{R}\mathrm{ank}(\mathbf{H}_\omega)$$

where Ω is the collection of all cuts between source and sink.

An amplify-and-forward strategy utilizing only linear transformations at the relays (that do not depend on the channel realization) is sufficient to achieve this MMG.

Proof: (Outline): The proof proceeds as follows (see Appendix E for details.)

1) First, a converse for the MMG is provided using simple cut-set bounds.
2) Next, we convert the fading network into a deterministic network with the property that the cut-set bound on MMG for the fading network is the same as the cut-set bound on the capacity of the deterministic network. It must be noted that the conversion to deterministic network used here is different from that used in [30] and [35], where the small nodes are used to model bits transmitted or received by a node, as opposed to their being modeled in the current paper as antennas.
3) We then characterize the zero-error capacity of the linear deterministic wireless network. We can also use the results of [38] to construct short-block-length zero-error codes for the linear deterministic channel. This will ensure that our MMG achieving scheme has very low complexity.
4) Finally, we convert a capacity-achieving scheme for the deterministic network into a MMG-achieving scheme for the fading network, which matches the converse.

■

C. MMG for Multicasting

In this section, we extend the result on MMG to the multicasting scenario.

Theorem 3.4: Given a single-source, D-sink, multicast Gaussian wireless network with i.i.d. Rayleigh fading coefficients, the MMG of the network is given by

$$D = \min_{\{j=1,2,\ldots,D\}} \min_{\omega \in \Omega_j} \mathbb{R}\mathrm{ank}(\mathbf{H}_\omega) \tag{10}$$

where Ω_j is the set of all cuts between the source and sink j. An amplify-and-forward strategy utilizing only linear transformations at the relays is sufficient to achieve this MMG.

Proof: The proof proceeds along lines similar to the proof of Theorem 3.3 and is omitted here for brevity. ■

IV. Tools for Computing DMT

In this section, we present further tools for computing the DMT in cooperative networks. We begin by establishing the

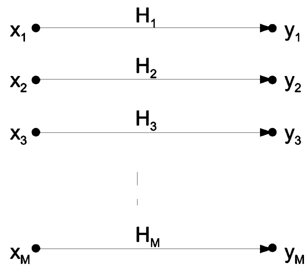

Fig. 3. The Parallel channel with M subchannels.

DMT of a bank of parallel MIMO channels. We then present a lower bound on the DMT of channels characterized by lower triangular block matrices. We conclude in the final subsection by applying these results to determine lower-bounds on the DMT of several networks. In several instances the lower bound turns out to be tight. These techniques will also be used in the companion paper [1] to compute the DMT of certain families of half-duplex networks.

A. DMT of Elementary Network Connections

1) Parallel Network: The lemma below presents an expression for the DMT of a parallel channel in terms of the DMT of the individual links under the assumption that the individual links experience independent fading.

Lemma 4.1: Consider a parallel channel with M links (see Fig. 3), with the ith link having representation $\mathbf{y_i} = \mathbf{H_i}\mathbf{x_i} + \mathbf{w_i}$, where $\{\mathbf{H}_i\}$ are independent and let $d_i(\cdot)$ denote the corresponding DMT. Then the DMT of the overall parallel channel is given by

$$d(r) = \inf_{(r_1,r_2,\ldots,r_M):\sum_{i=1}^{M} r_i = r} \sum_{i=1}^{M} d_i(r_i). \tag{11}$$

Proof: See Appendix F. ∎

The following lower and upper bounds on the outage exponent are immediate from (11):

$$\sum_{i=1}^{M} d_i(r) \le d(r) \le \sum_{i=1}^{M} d_i\left(\frac{r}{M}\right). \tag{12}$$

To determine the DMT of the parallel channel when all component channels are identically distributed and possess individual DMTs that are a convex function of the rate, we will make use of the following lemma from the theory of majorization [45]:

Lemma 4.2: [45] If $f(.)$ is a symmetric function in variables $r_1, r_2, \ldots, r_N$ and is convex in each of the variables r_i, $i = 1, 2, \ldots, N$, then

$$\inf_{(r_1,r_2,\ldots,r_N):\sum_{i=1}^{N} r_i = r} f(r_1, r_2, \ldots, r_N) = f\left(\frac{r}{N}, \frac{r}{N}, \ldots, \frac{r}{N}\right).$$

The corollary below follows as a result.

Corollary 4.3: The DMT of a parallel channel with all the M channels being i.i.d. and having a common convex DMT $d_1(r)$ is given by

$$d(r) = M d_1\left(\frac{r}{M}\right). \tag{13}$$

The result in Corollary 4.3 was also obtained in [14].

2) Parallel Channel With Repeated Coefficients:

Lemma 4.4: Consider a parallel channel with M links and repeated channel matrices. More precisely, let there be N independent channel matrices $H^{(1)}, H^{(2)}, \ldots, H^{(N)}$, with $H^{(i)}$ repeating in n_i subchannels.

Then the DMT of such a parallel channel is given by

$$d(r) = \inf_{(r_1,r_2,\ldots,r_M):\sum_{i=1}^{N} n_i r_i = r} \sum_{i=1}^{N} d_i(r_i). \tag{14}$$

Proof: The proof is along the lines of the proof of Lemma 4.1, and can be found in [22]. ∎

B. A Lower Bound on the DMT of Block-Lower-Triangular Matrices

In many situations, the matrices induced by AF protocols in a ss-ss network will turn out to posses block-lower-triangular (blt) structure. In this section, we provide a lower bound to the DMT of channels associated to such matrices.

Definition 3: Consider a collection $\{A_{ij} \mid 1 \le j \le i \le N\}$ of matrices where matrix A_{ij} is of size $(N_i \times N_j)$. Let A be the blt matrix whose (i,j)th entry is the matrix A_{ij} for $1 \le j \le i \le N$ and zeros elsewhere, i.e.

$$A = \begin{bmatrix} A_{11} & 0 & \ldots & 0 \\ A_{21} & A_{22} & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ A_{N1} & A_{N2} & \ldots & A_{NN} \end{bmatrix}.$$

We define the ℓth subdiagonal matrix, $A^{(\ell)}$ of such a blt matrix A as the matrix comprising only of the entries $A_{(\ell+1)1}, A_{(\ell+2)2}, \ldots, A_{N(N-\ell)}$ in their respective positions with zeros everywhere else, i.e.

$$(A^{(\ell)})_{ij} = \begin{cases} A_{ij} & \text{if } i-j=\ell \\ 0_{N_i \times N_j} & \text{otherwise.} \end{cases}$$

The 0th subdiagonal matrix corresponds to the diagonal part of the matrix. The *last* subdiagonal matrix of A is defined as the subdiagonal matrix $A^{(\ell)}$ of A, where ℓ is the largest integer for which $A^{(\ell)}$ is nonzero.

The theorem below establishes lower bounds on the DMT of channel matrices which have a blt structure.

Theorem 4.5: Consider a random blt matrix $\mathbf{H}$ having component matrices $\mathbf{H}_{ij}$ of size $N_i \times N_j$. Let $\mathbf{H}^{(0)}$ be the diagonal part of the matrix $\mathbf{H}$ and $\mathbf{H}^{(\ell)}$ denote the last subdiagonal matrix of $\mathbf{H}$, as given by Definition 3. Then:

1) $d_H(r) \ge d_{H^{(0)}}(r)$.
2) $d_H(r) \ge d_{H^{(\ell)}}(r)$.

3) In addition, if the entries of $H^{(\ell)}$ are independent of the entries in $H^{(0)}$, then $d_H(r) \geq d_{H^{(0)}}(r) + d_{H^{(\ell)}}(r)$.

Proof: Let $M := \sum_{i=1}^{N} N_i$ be the size of the square matrix $\mathbf{H}$. To estimate the outage exponent, we begin by identifying lower bounds on the mutual information. Since ℓ is the largest size of a nonzero subdiagonal matrix, we have

$$\mathbf{y}_i = \mathbf{H}_{ii}\mathbf{x}_i + \mathbf{H}_{i(i-1)}\mathbf{x}_{i-1} + \cdots + \mathbf{H}_{i(i-\ell)}\mathbf{x}_{i-\ell} + \mathbf{w}_i.$$

We begin with the mutual-information expression expansion,

$$I(\mathbf{x};\mathbf{y}|\mathbf{H}=H) = \sum_{i=1}^{M} I(\mathbf{x}_i;\mathbf{y}|\mathbf{H}=H,\mathbf{x}_1^{i-1}). \qquad (15)$$

Consider next, the following series of inequalities for all $i = 1,\ldots,N$:

$$\begin{aligned} I(\mathbf{x}_i;\mathbf{y}|\mathbf{H}=H,\mathbf{x}_1^{i-1}) &\geq I(\mathbf{x}_i;\mathbf{y}_i|\mathbf{H}=H,\mathbf{x}_1^{i-1}) \\ &= I(\mathbf{x}_i; H_{ii}\mathbf{x}_i + H_{i(i-1)}\mathbf{x}_{i-1} \\ &\quad + \cdots + H_{i(i-\ell)}\mathbf{x}_{i-\ell} + \mathbf{w}_i|\mathbf{x}_1^{i-1}) \\ &= I(\mathbf{x}_i; H_{ii}\mathbf{x}_i + \mathbf{w}_i|\mathbf{x}_1^{i-1}) \\ &= I(\mathbf{x}_i; H_{ii}\mathbf{x}_i + \mathbf{w}_i). \end{aligned}$$

The last step follows since $\{\mathbf{x_i}\}$ are independent under the assumed i.i.d input distribution. We thus have

$$\begin{aligned} I(\mathbf{x};\mathbf{y}|\mathbf{H}=H) &\geq \sum_{i=1}^{M} I(\mathbf{x}_i; H_{ii}\mathbf{x}_i + \mathbf{w}_i) \\ &= I(\mathbf{x}; \mathbf{H^{(0)}}\mathbf{x} + \mathbf{w}|\mathbf{H^{(0)}} = H^{(0)}). \quad (16) \end{aligned}$$

From (16), it follows that

$$\begin{aligned} \rho^{-d_H(r)} &\doteq \Pr\{I(\mathbf{x};\mathbf{y}|\mathbf{H}=H) \leq r\log\rho\} \\ &\leq \Pr\{I(\mathbf{x};\mathbf{H}^{(0)}\mathbf{x}+\mathbf{w}|\mathbf{H}^{(0)}=H^{(0)}) \leq r\log\rho\} \\ &\doteq \rho^{-d_{H^{(0)}}(r)} \end{aligned} \qquad (17)$$

i.e.

$$d_H(r) \geq d_{H^{(0)}}(r). \qquad (18)$$

Similarly, we can derive

$$I(\mathbf{x};\mathbf{y}|\mathbf{H}=H) \geq I(\mathbf{x};\mathbf{H}^{(\ell)}x + \mathbf{w}|\mathbf{H}^{(\ell)} = H^{(\ell)}). \qquad (19)$$

This corresponds effectively to replacing the matrix H by the last subdiagonal matrix $H^{(\ell)}$. This will lead us to $d_H(r) \geq d_{H^{(\ell)}}(r)$.

It follows therefore, from (16) and (19) that

$$\begin{aligned} &I(\mathbf{x};\mathbf{y}|\mathbf{H}=H) \\ &\quad \geq \max\{I(\mathbf{x};\mathbf{H^{(0)}}\mathbf{x}+\mathbf{w}|\mathbf{H^{(0)}}=H^{(0)}) \\ &\qquad I(\mathbf{x};\mathbf{H}^{(\ell)}\mathbf{x}+\mathbf{w}|\mathbf{H}^{(\ell)}=H^{(\ell)})\}. \end{aligned} \qquad (20)$$

Using the above together with the independence of the entries in $\mathbf{H^{(0)}}$ and $\mathbf{H}^{(\ell)}$, we obtain

$$\rho^{-d_H(r)} \dot{\leq} \rho^{-d_{H^{(0)}}(r)}\rho^{-d_{H^{(\ell)}}(r)}$$

$$\Rightarrow d_H(r) \geq d_{H^{(0)}}(r) + d_{H^{(\ell)}}(r). \qquad (21)$$

■

Remark 1: The two matrix inequalities below can be deduced from the proof of Theorem 4.5, with $\mathbf{H}^{(0)}$ and $\mathbf{H}^{(\ell)}$ defined as in the theorem

$$\det(I+\rho\mathbf{H}\mathbf{H}^\dagger) \geq \det(I+\rho\mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger}) \qquad (22)$$
$$\det(I+\rho\mathbf{H}\mathbf{H}^\dagger) \geq \det(I+\rho\mathbf{H}^{(\ell)}\mathbf{H}^{(\ell)\dagger}). \qquad (23)$$

Remark 2: Although the result is derived for lower triangular matrices, it extends easily to banded matrices as well. The only change is that the first nonzero diagonal replaces the diagonal in the statement of the theorem.

C. Example Applications of the DMT Lower Bound

Here, we derive lower bounds based on the results in the previous two subsections to the DMT of two-hop networks under the operation of various existing AF protocols. One lower bound proves a conjecture by Rao and Hassibi [26], while a second is tighter than lower bounds known earlier. In the remaining instances, two other examples provide lower bounds for the NAF and SAF protocol which from the literature are known to be tight. Additionally, the derivations presented here are surprisingly simple and provide some intuitive explanation as to how these protocols achieve the DMT.

Example 1: Single Relay, NAF Protocol: Consider the relay network in Fig. 1 operating under the NAF protocol. Let $\mathbf{g}_d$, $\mathbf{g}_1$, $\mathbf{h}_1$ denote the channel coefficients associated with the source-sink, source-relay, and relay-sink links, respectively. The induced channel under the NAF protocol is given by

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{g}_d & 0\\ \mathbf{g}_1\mathbf{h}_1 & \mathbf{g}_d\end{bmatrix}\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix} + \begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_2 + h_1\mathbf{v}\end{bmatrix}.$$

The noise in this channel can be seen to correlated in general with the channel. However, since the noise satisfies conditions of Theorem 2.3, the theorem implies that for DMT computations, we can treat the noise as white. Applying Theorem 4.5 to this channel, the DMT $d(r)$ of the protocol can be lower bounded as

$$d(r) = d_H(2r) \geq (1-r)^+ + (1-2r)^+$$

where the factor of 2 in $d_H(2r)$ carries out rate compensation.

From [8] we know that this bound is indeed tight.

Remark 3: For the case of NAF protocol used with N relays, it can be similarly shown that $d(r) \geq (1-r)^+ + N(1-2r)^+$, now also using the parallel channel result in Lemma 4.1. This lower bound is proved to be tight for the N-relay case as well in [8].

Example 2: Multiple Relays, SAF Protocol: Consider the network in Fig. 1 with N relays operating under the Slotted Amplify-and-Forward (SAF) protocol with M slots, introduced in [16]. We assume that the relays are isolated from each others' transmissions (see [16]). For the example case when $M = 5$ and $N = 2$, the induced channel matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{g}_d & 0 & 0 & 0 & 0 \\ \mathbf{g}_1 & \mathbf{g}_d & 0 & 0 & 0 \\ 0 & \mathbf{g}_2 & \mathbf{g}_d & 0 & 0 \\ 0 & 0 & \mathbf{g}_1 & \mathbf{g}_d & 0 \\ 0 & 0 & 0 & \mathbf{g}_2 & \mathbf{g}_d \end{bmatrix}.$$

Now the DMT of this protocol can be bounded using Theorem 4.5 as

$$d(r) = d_H(Mr) \geq (1-r)^+ + N\left(1 - \frac{M}{M-1}r\right)^+.$$

The right-hand side (RHS) is in fact shown to be equal to the DMT in [16]. For large M, the RHS of this bound approaches the cutset bound.

Example 3: Multiple Antenna, Multiple Relays, NAF Protocol: We consider a N-relay network with each node in the network having multiple antennas. Let $\mathbf{H}_d$ be the channel matrix for the direct link, $\mathbf{H}_{si}$ and $\mathbf{H}_{id}$ be the matrices corresponding to the channel between source to relay i, and relay i to destination respectively. Let the DMT of the product channel $\mathbf{H}_i = \mathbf{H}_{si}\mathbf{H}_{id}$ through relay i be given by $d_i(r)$. The NAF protocol for N relays can be viewed as if one were operating by applying the NAF protocol to each relay separately and then cycling through all the relays. When each user has a different number of antennas, then it is in general, advantageous to use each relay for different fractions f_i of time. Under this protocol with asymmetric activation, we obtain the DMT using Theorem 4.5 and Lemma 4.4 as (see [22] for details)

$$d(r) \geq d_{H_d}(r) + \sup_{\substack{\{(f_1,f_2,\ldots,f_N):\\ \sum_{i=1}^N f_i=1\}}} \inf_{\substack{\{(r_1,r_2,\ldots,r_N):\\ \sum_{i=1}^N f_i r_i=2r\}}} \sum_{i=1}^N d_i(r_i). \qquad (24)$$

The scheme in [14] is now a special case of this protocol where all the relays are used for a equal duration of time, i.e., $f_i = 1/N$ for all i.

Example 4: Multiple-Antenna, Multiple-Relay, Generalized NAF Protocol: Let us now consider a N-relay network with the source and destination having n_s and n_d antennas and the relays having a single antenna each. For this network, a generalized NAF protocol was proposed in [26], where during the first T time instants, the source transmits to the N relays. Over the next T time slots, the relays transmit a *linear transformation* of the vector received over the prior T time slots. This induces an effective channel matrix between the source and the destination of the form:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_D & 0 \\ \mathbf{H}_R & \mathbf{H}_D \end{bmatrix} \qquad (25)$$

where $\mathbf{H}_D = I_T \otimes \mathbf{H}_d$, with I_T denoting the identity matrix of size T, $\otimes$ denoting the tensor product and $\mathbf{H}_d$ denoting the $n_s \times n_d$ fading matrix corresponding to the direct link between the source and the destination. $\mathbf{H}_R$ is a $Tn_d \times Tn_s$ matrix corresponding to the effective channel through the relays, which depends on the source-relay and relay-destination channels as well as on the linear transformations employed at the relays.

Now, $\mathbf{H}$ is blt and therefore, we invoke Theorem 4.5 to get, $d_H(r) \geq d_{H^{(0)}}(r) + d_{H^{(\ell)}}(r)$. Next we note that the matrix $\mathbf{H}^{(0)}$ corresponds to a block-diagonal matrix with $\mathbf{H}_D$ repeated twice along the diagonal or effectively, $\mathbf{H}_d$ repeated $2T$ times along the diagonal and clearly $\mathbf{H}^{(\ell)} = \mathbf{H}_R$. It follows that $d_{H^{(0)}}(r) = d_{H_d}(\frac{r}{2T})$.

The protocol utilizes $2T$ time instants to induce the effective channel matrix $\mathbf{H}$ and therefore the DMT of the protocol $d(r)$ can be given in terms of the DMT of the matrix $\mathbf{H}$ as $d(r) = d_H(2Tr)$. Thus

$$\begin{aligned} d(r) &= d_H(2Tr) \\ &\geq d_{H^{(0)}}(2Tr) + d_{H^{(\ell)}}(2Tr) \\ &= d_{H_D}(2Tr) + d_{H_R}(2Tr) \\ &= d_{H_d}(r) + d_{H_R}(2Tr). \end{aligned} \qquad (26)$$

This proves [26, Conjecture 1].

V. DMT Bounds for Single Antenna Relay Networks

In this section, we consider ss-ss networks equipped with full-duplex single-antenna nodes. We provide a lower bound to the DMT of such a network by exploiting Menger's theorem.

Definition 4: Consider a network $\mathcal{N}$ and a path P from source to sink. This path P is said to have a *shortcut* if there is a single edge in $\mathcal{N}$ connecting two nonconsecutive nodes in P.

Theorem 5.1: Consider a ss-ss full-duplex network with single-antenna nodes. Let the min-cut of the network be M. Let the network satisfy *either* of the two conditions below:

1) The network has no directed cycles, or
2) There exists a set of M edge-disjoint paths between source and sink such that *none* of the M paths have shortcuts.

Then, a linear DMT $d(r) = M(1-r)^+$ between a MMG of 1 and maximum diversity gain M is achievable.

Proof: The proof is along the lines of the proof of Theorem 3.1, and is skipped here for brevity. Details can be found in [22]. ■

Appendix A
Proof of Lemma 2.2

Let $f \in \mathbb{C}[X_1, X_2, \ldots, X_M]$ be written as $f(X_1, X_2, \ldots, X_M) := \sum_{i=1}^{S} c_i f_i(X_1, X_2, \ldots, X_M)$, where $c_i \in \mathbb{C}$ are constants and f_i are monomials. Then for every assignment, $X_i = h_i$

$$|f(h_1, h_2, \ldots, h_M)| \leq \sum_i |c_i||f_i(h_1, h_2, \ldots, h_M)|.$$

Now we have

$$\begin{aligned} &\Pr\{|f(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M)|^2 > k\} \\ &\quad \leq \Pr\{\sum_i |c_i||f_i(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M)| > k^{\frac{1}{2}}\} \end{aligned} \qquad (27)$$

$$\le \Pr\left\{\bigcup_i \left(|c_i||f_i(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_M)| > \frac{k^{\frac{1}{2}}}{S}\right)\right\}$$
$$\le \sum_i \Pr\left\{|f_i(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_M)|^2 > \frac{k}{c_{\max}S^2}\right\} \quad (28)$$

where $c_{\max}$ is the maximum over all $\{|c_i|^2\}$. We know $\mathbf{u}_j{:=}|\mathbf{h}_j|^2$ has an exponential distribution. Since f_i is a monomial

$$|f_i(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_M)|^2 = \prod_{j=1}^{M} \mathbf{u}_j{}^{a_{ij}}$$

where $0 \le a_{ij} \le D$ is an integer, where D is the maximum degree of any of the monomials f_i in any of the variables $\mathbf{h}_i$.

Consider a single term in the RHS of (28), we have

$$\begin{aligned}
&\Pr\left\{|f_i(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_M)|^2 > \frac{k}{c_{\max}S^2}\right\}\\
&\le \Pr\left\{\bigcup_j \left(\mathbf{u}_j{}^{a_{ij}} > \left(\frac{k}{c_{\max}S^2}\right)^{\frac{1}{M}}\right)\right\}\\
&\le \sum_j \Pr\left\{\mathbf{u}_j{}^{a_{ij}} > \left(\frac{k}{c_{\max}S^2}\right)^{\frac{1}{M}}\right\}\\
&\le \sum_j \Pr\left\{\mathbf{u}_j > \left(\frac{k}{c_{\max}S^2}\right)^{\frac{1}{MD}}\right\} \quad (29)\\
&= M\ \exp\left(-\left(\frac{k}{c_{\max}S^2}\right)^{\frac{1}{MD}}\right). \quad (30)
\end{aligned}$$

Equation (29) follows if $k \ge c_{\max}S^2$. We get this condition by setting $\delta := c_{\max}S^2 > 0$, since by the hypothesis of the lemma, we have $k \ge \delta$.

Combining (28) and (30), we get, for $k \ge \delta$

$$\begin{aligned}
&\Pr\{|f(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_M)|^2 > k\}\\
&\le MS\ \exp\left(-\left(\frac{k}{c_{\max}S^2}\right)^{\frac{1}{MD}}\right)\\
&= A\ \exp\left(-Bk^{\frac{1}{d}}\right)
\end{aligned}$$

where $A{:=}MS$, $B{:=}\left(\frac{1}{c_{\max}S^2}\right)^{\frac{1}{MD}}$, $\delta = c_{\max}S^2$ and $d := MD$.

APPENDIX B
PROOF OF THEOREM 2.3

We need to show

$$\begin{aligned}
&Pr(\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\mathbf{\Sigma}^{-1}) < r\log\rho)\\
&\doteq Pr(\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger) < r\log\rho)
\end{aligned}$$

where $\mathbf{\Sigma}$ is the correlation matrix of noise, given by

$$\mathbf{\Sigma} = \mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] = I + \sum_{i=1}^{M}\mathbf{G}_i\mathbf{G}_i^\dagger. \quad (31)$$

Let $\lambda_i(A)$, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote the ith largest, maximum and minimum eigenvalues of a positive semi-definite matrix A. By the Amir-Moez bound [51], for any two positive definite $n \times n$ Hermitian matrices A, B

$$\lambda_i(A)\lambda_{\min}(B) \le \lambda_i(AB) \le \lambda_i(A)\lambda_{\max}(B).$$

So we get

$$\begin{aligned}
\det(I+\rho AB) &= \prod_i (1+\rho\lambda_i(AB))\\
&\le \prod_i (1+\rho\lambda_i(A)\lambda_{\max}(B))\\
&= \det(I+\rho\lambda_{\max}(B)A). \quad (32)
\end{aligned}$$

Similarly

$$\det(I+\rho AB) \ge \det(I+\rho\lambda_{\min}(B)A). \quad (33)$$

Applying (32) and (33) to $A = \mathbf{H}\mathbf{H}^\dagger$ and $B = \mathbf{\Sigma}^{-1}$

$$\begin{aligned}
&\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\lambda_{\min}(\mathbf{\Sigma}^{-1}))\\
&\quad\le \det(I+\rho\mathbf{H}\mathbf{H}^\dagger\mathbf{\Sigma}^{-1}) \quad (34)\\
&\quad\le \det(I+\rho\mathbf{H}\mathbf{H}^\dagger\lambda_{\max}(\mathbf{\Sigma}^{-1})). \quad (35)
\end{aligned}$$

Continuing from (34) and (35), we have

$$\begin{aligned}
&\Pr\{\log(\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\lambda_{\min}(\mathbf{\Sigma}^{-1}))) < r\log\rho\}\\
&\quad\ge \Pr\{\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\mathbf{\Sigma}^{-1}) < r\log\rho\}\\
&\quad\ge \Pr\{\log(\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\lambda_{\max}(\mathbf{\Sigma}^{-1}))) < r\log\rho\}. \quad (36)
\end{aligned}$$

We will show that both the bounds coincide as $\rho \to \infty$. We begin with bounding $\lambda_{\min}(\mathbf{\Sigma})$ and $\lambda_{\max}(\mathbf{\Sigma})$. Let e_i be the eigen vector corresponding to $\lambda_i(\Sigma)$ for a realization Σ of $\mathbf{\Sigma}$. Then

$$\begin{aligned}
\lambda_i(\Sigma) \parallel e_i \parallel^2 &= e_i{}^\dagger\Sigma e_i\\
&= e_i{}^\dagger\left(I+\sum_{i=1}^{M}G_iG_i^\dagger\right)e_i\\
&\ge \parallel e_i \parallel^2\\
\Rightarrow \lambda_i(\Sigma) &\ge 1 \quad \forall i\\
\Rightarrow \lambda_{\max}(\Sigma^{-1}) &\le 1. (37)
\end{aligned}$$

Since this is true for every realization we can write $\lambda_{\max}(\mathbf{\Sigma}^{-1}) \le 1$

$$\begin{aligned}
\text{Hence, }\ &\Pr\{\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\mathbf{\Sigma}^{-1}) < r\log\rho\}\\
&\ge \Pr\{\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger\lambda_{\max}(\mathbf{\Sigma}^{-1})) < r\log\rho\}\\
&\ge \Pr\{\log\det(I+\rho\mathbf{H}\mathbf{H}^\dagger) < r\log\rho\} \quad (38)\\
&\doteq \rho^{-d_{out}(r)}. \quad (39)
\end{aligned}$$

Now we proceed to get an upper bound on $\lambda_{\max}(\mathbf{\Sigma})$

$$\lambda_{\max}(\mathbf{\Sigma}) = \lambda_{\max}(I + \sum_{i=1}^{M}\mathbf{G}_i\mathbf{G}_i^\dagger)$$

$$\leq 1 + \text{Tr}\,(\sum_{i=1}^{M} \mathbf{G}_i \mathbf{G}_i^\dagger)$$
$$= 1 + \sum_{i=1}^{M} \|\mathbf{G}_i\|^2{}_F = 1 + \sum_{i=1}^{M}\sum_{j=1}^{N^2} |f_{ij}|^2 \quad (40)$$

where f_{ij} represents a polynomial entry of the matrix G_i. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2L}$ denote in some order, the real and imaginary parts of $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_L$. Then, the RHS in (40) is a real polynomial in the variables $\mathbf{x}_i$, $i = 1, 2, \ldots, 2L$.

This leads to

$$\lambda_{\max}(\mathbf{\Sigma}) \leq g(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2L}) + 1 \quad (41)$$

where $g(x_1, x_2, \ldots, x_{2L})$ is a polynomial without constant term in the variables $\{x_i\}$. Let us invoke Lemma 2.2 for the polynomial g which does not possess any constant term, with $k = \rho^\epsilon - 1$. Then, $\forall \rho^\epsilon - 1 \geq \delta$, we get

$$\Pr\{\lambda_{\max}(\mathbf{\Sigma}) > \rho^\epsilon\} \leq \Pr\{g(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_S) > \rho^\epsilon - 1\}$$
$$\leq A \exp(-B(\rho^\epsilon - 1)^{\frac{1}{d}}) \quad (42)$$

for some constants A, B, $d > 0$. Also $\rho^\epsilon - 1 \geq \delta$ is equivalent to $\rho \geq \rho_0$, if we choose ρ_0 such that $\rho_0^\epsilon - 1 = \delta$. So (42) holds for sufficiently large ρ, which is our regime of interest.

Let $\hat{\mathbf{h}}$ denote the set of all the random fading coefficients in the network, and let $\hat{h}$ denote a realization of the fading coefficients. Given a $\hat{h}$, H and G_i are deterministic.

Let $U = \{\hat{\mathbf{h}} \mid \log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger \mathbf{\Sigma}^{-1}) < r \log \rho\}$ and $V = \{\hat{\mathbf{h}} \mid \lambda_{\max}(\mathbf{\Sigma}) > \rho^\epsilon\}$ be two events. Then,

$$\Pr(U) \leq \Pr(U \cap V^c) + \Pr(V)$$
$$\dot{\leq} \Pr(U \cap V^c) + A \exp(-B(\rho^\epsilon - 1)^{\frac{1}{d}}) \quad (43)$$
$$\text{Now, } U = \{\hat{\mathbf{h}} \mid \log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger \mathbf{\Sigma}^{-1}) < r \log \rho\}$$
$$\subset \{\hat{\mathbf{h}} \mid \log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger \lambda_{\min}(\mathbf{\Sigma}^{-1})) < r \log \rho\}$$
$$= \{\hat{\mathbf{h}} \mid \log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger \lambda_{\max}(\mathbf{\Sigma})^{-1}) < r \log \rho\}$$
$$U \cap V^c \subset \{\hat{\mathbf{h}} \mid \log\det(I + \rho^{1-\epsilon} \mathbf{H}\mathbf{H}^\dagger) < r \log \rho\}. \quad (44)$$

Substituting (44) and (42) in (43), we have

$$\Pr(U) \leq \Pr(U \cap V^c) + \Pr(U)$$
$$\dot{\leq}\ \Pr\{\hat{\mathbf{h}} \mid \log\det(I + \rho^{1-\epsilon} \mathbf{H}\mathbf{H}^\dagger) < r \log \rho\}$$
$$+ A \exp(-B(\rho^\epsilon - 1)^{\frac{1}{d}}).$$
$$\doteq \Pr\{\hat{\mathbf{h}} \mid \log\det(I + \rho^{1-\epsilon} \mathbf{H}\mathbf{H}^\dagger) < r \log \rho\}.$$
$$\doteq (1 - \epsilon)\rho^{d_{out}\left(\frac{r}{1-\epsilon}\right)}. \quad (45)$$

In (45), ϵ is arbitrary, and we let it tend to zero. Hence, by (45) and (39), the exponents for both the bounds in (36) coincide and we obtain

$$\Pr\{\log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger \Sigma^{-1}) < r \log \rho\}$$
$$\doteq \Pr\{\log\det(I + \rho \mathbf{H}\mathbf{H}^\dagger) < r \log \rho\}.$$

This proves the assertion of the theorem.

Appendix C
Proof of Lemma 2.5

Consider any two intervals $R_1, R_2 \in \mathcal{R}$ such that the interval in between is not contained in $\mathcal{R}$. If we are not able to find two such intervals, then clearly $L \leq 1$, and we are done. Let $R_1 = [a, b]$ and $R_2 = [c, d]$, and without loss of generality assume that $a \leq b < c \leq d$. First, we claim that there exists a point $b \leq x_0 \leq c$, such that either $p'(x_0) = 0$, or $p''(x_0) = 0$. We now proceed to prove this claim.

Clearly, either of the two conditions (4) or (5) is violated just to the right of the point $x = b$, else, the interval would extend beyond b. We can show that in the first case, $p'(x_0) = 0$, i.e., the first derivative of p vanishes, and in the second case, the second derivative of p vanishes, i.e., $p''(x_0) = 0$ for some $b \leq x_0 \leq c$ (see [22] for a detailed explanation). Thus there exists a real root of $p'(x)$ or $p''(x)$ between those two intervals. Since the number of roots of a polynomial is bounded by its degree, there will be only finitely many such intervals. In particular, the number of intervals L is bounded by $2d$, which is an upper bound on the total number of zeros of $p'(x)$ and $p''(x)$.

Appendix D
Proof of Lemma 2.6

For any polynomial f in independent Gaussian random variables, we have that

$$\Pr\{f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \neq 0\} = 1. \quad (46)$$

Let $\mathbf{x} := \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. Let us define an indicator function $I_\delta(\mathbf{x})$ as follows:

$$I_\delta(\mathbf{x}) := \left\{ \begin{array}{ll} 1, & |f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| < \delta \\ 0, & \text{else} \end{array} \right\}.$$
$$\text{Then } \Pr\{|f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| < \delta\}$$
$$= \mathbb{E}_{\mathbf{x}}\ I_k(\mathbf{x})$$
$$= \mathbb{E}_{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N-1}}\ \mathbb{E}_{\mathbf{x}_N}\{I_k(\mathbf{x}) \mid \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}\}$$
$$= \mathbb{E}_{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N-1}}\ \Pr\{|f(x_1, x_2, \ldots, x_{N-1}, \mathbf{x}_N)| < \delta\}. \quad (47)$$

Let $f(\mathbf{x_N}) := f(x_1, x_2, \ldots, x_{N-1}, \mathbf{x}_N)$, where the dependence of f on the first $N - 1$ variables is made implicit. Let

$$f(\mathbf{x}_N) = \sum_{k=0}^{d_N} b_k \mathbf{x}_N^k$$

where d_N is the degree of the polynomial f in the variable x_N. Since b_{d_N} is a polynomial in the variables $x_1, \ldots, x_{N-1}$, it follows from (46) that with probability one, $b_{d_N} \neq 0$. Let

$$g(x_N) = \frac{\partial f(x_N)}{\partial x_N}$$

be the partial derivative of $f(x_N)$ with respect to x_N. Then

$$\Pr\{|f(x_1, x_2, \ldots, x_{N-1}, \mathbf{x}_N)| < \delta\}$$
$$= \Pr\{|f(\mathbf{x}_N)| < \delta, |g(\mathbf{x}_N)| \geq \delta^{1/2}\}$$

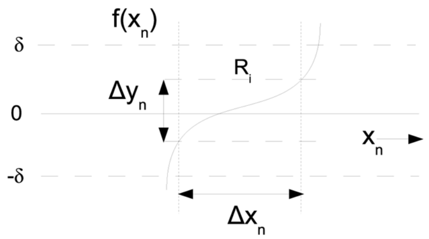

Fig. 4. f(x) in a region R_i.

$$\begin{aligned} &+ \Pr\{|f(\mathbf{x}_N)| < \delta, |g(\mathbf{x}_N)| < \delta^{1/2}\} \\ \leq &\Pr\{|f(\mathbf{x}_N)| < \delta, |g(\mathbf{x}_N)| \geq \delta^{1/2}\} \\ &+ \Pr\{|g(\mathbf{x}_N)| < \delta^{1/2}\}. \end{aligned} \tag{48}$$

Let us consider the first term on the RHS. The region $\mathcal{R} := \{|f(x_N)| < \delta, |g(x_N)| \geq \delta^{1/2}\}$ is described by two conditions $|f(x_N)| < \delta$ and $|g(x_N)| \geq \delta^{1/2}$. It is shown in Lemma 2.5 that the set of all values of x_N satisfying both conditions can be expressed as the union of L pairwise-disjoint intervals R_i, $i = 1, 2, \ldots, L$ with $L < 2d_N$. Now $\Pr(x_N \in R) = \sum_{i=1}^{L} \Pr(\mathbf{x}_N \in R_i)$. We will now proceed to upper-bound the probability $\mathbf{x}_N \in R_i$. To do so, consider Fig. 4. Let Δx_n be the width of the interval R_i and $\Delta f(x_n)$ be the height. Since the slope of the curve $g(x)$ is greater than $\delta^{1/2}$ throughout R_i, we have that

$$\left|\frac{\Delta f(x_n)}{\Delta x_n}\right| \geq \delta^{1/2}.$$

We can assume without loss of generality that

$$\frac{\Delta f(x_n)}{\Delta x_n} \geq \delta^{1/2}.$$

Also in any contiguous region, $\Delta f(x_n) \leq 2\delta$. This implies that

$$\Delta x_n \leq \frac{2\delta}{\delta^{1/2}} = 2\delta^{1/2}.$$

Since $\mathbf{x}_n$ is a $\mathcal{N}(0,1)$ random variable, the probability that x_n lies in a range of $2\delta^{1/2}$ is less than $2c\delta^{1/2}$ where c is the maximum of the Gaussian pdf.

Therefore

$$\begin{aligned} \Pr\{\mathbf{x} \in R_i\} &\leq 2c\delta^{1/2}. \\ \text{Now } \Pr\{\mathbf{x} \in R\} &= \sum_{i=1}^{L} \Pr\{\mathbf{x} \in R_i\} \\ &\leq L2c\delta^{1/2} = C\delta^{1/2}. \end{aligned} \tag{49}$$

Plugging (49) into (48) yields

$$\Pr\{|f(\mathbf{x}_N)| < \delta\} \leq C\delta^{1/2} + \Pr\{|g(\mathbf{x}_N)| < \delta^{1/2}\}. \tag{50}$$

Since $g(x)$ is of lower degree than $f(x)$, the process can be continued to yield

$$\begin{aligned} \Pr\{|f(\mathbf{x}_N)| < \delta\} \leq &C\delta^{1/2} + C\delta^{1/4} + \cdots + C\delta^{1/2^{d_N - 1}} \\ &+ \Pr\{b_{d_N} \leq \delta^{1/2^{d_N}}\}. \end{aligned} \tag{51}$$

Only the last term involving b_{d_N} is a function of the remaining variables $x_1, x_2, \ldots, x_{N-1}$. The last term is identical to that in the RHS of (47), except that the polynomial b_{d_N} involves $(N-1)$ or fewer variables and hence this recursion is finite. Eventually, we will be left with the probability that a constant coefficient J is greater than $\delta^{1/s}$ for some integer s. Choosing the constant K appearing in the statement of the lemma to equal J^s, we obtain that this probability is equal to the probability that $K \leq \delta$. But by hypotheses, $K > \delta$ and hence this probability is equal to zero. This allows us to rewrite the bound on probability appearing in (51) as

$$\Pr\{|f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| < \delta\} \leq C_1(\delta^{1/2} + \delta^{1/4} + \cdots + \delta^{1/2^e})$$

for a suitable constant C_1 and some integer e.

Choosing $K \leq 1$ forces $\delta < 1$ since by hypotheses, $\delta < K$. In this case

$$\delta^{1/2} + \delta^{1/4} + \cdots + \delta^{1/2^e} \leq e\delta^{1/2^e}.$$

With $A := eC_1$ and $d := 2^e$, we get

$$\Pr\{|f(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)| < \delta\} \leq A\delta^{1/d} \tag{52}$$

as desired.

Appendix E
Proof of Theorem 3.3

Since the converse is clear, we proceed to the achievability part of the proof. First, we convert the wireless fading network into a derived linear deterministic network. The construction of the derived deterministic network is described below. We will show that the zero-error capacity of this derived deterministic network is lower bounded by the upper-bound on the MMG of the fading network.

Let the number of edges in the fading network be N. Fading coefficients associated with edges of the network are denoted by $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_N$. To construct the derived deterministic network, consider a deterministic network with the same topology as that of the original fading network. We take q, the vector length in the deterministic network to be equal to the maximum number of antennas of any terminal in the fading network. For terminals with number of antennas less than q, we assign zero as the finite-field fading coefficient between that antenna and all other terminals. We still need to decide the finite field size, p, and finite field coefficients on the edges to completely characterize the equivalent finite-field deterministic network. We shall denote these finite-field coefficients by ξ_i, $i = 1, 2, \ldots, N$. We shall consider $\{\xi_i\}$ as indeterminates, before values are assigned to them.

For determining the field size p and $\{\xi_i,\ i = 1, 2, \ldots, N\}$, we will impose further conditions. In particular, we will ensure that the deterministic network will have at least the same capacity as the upper bound on MMG for the fading network. Due to the similarity between the expression for capacity in Theorem 3.2 and MMG terms in Theorem 3.3, the above condition can be met by making sure that, cut-by-cut, the rank of the transfer matrix (G_ω) in the deterministic network is at least as large as the structural rank of the transfer matrix $\mathbf{H}_\omega$, i.e., $\text{rank}(G_\omega) \geq \mathbb{R}\text{ank}(\mathbf{H}_\omega)$.

Let us fix a cut ω, and let $r_\omega := \mathbb{R}\text{ank}(\mathbf{H}_\omega)$ be the structural rank of the transfer matrix of the cut in the fading network. Then, there exists a $r_\omega \times r_\omega$ submatrix (say $\mathbf{H}'_\omega$) of $\mathbf{H}_\omega$, which has structural rank r_ω. Consider the same cut on the deterministic network and find the corresponding $r_\omega \times r_\omega$ submatrix G'_ω of the transfer matrix G_ω. Now consider the determinant of the matrix G'_ω. The determinant is a polynomial in several variables ξ_i, $i = 1, 2, .., N$ with rational integer coefficients. Let us call this polynomial as $f_\omega(\xi_1, \xi_2, .., \xi_N)$. This polynomial is not identically a zero polynomial over $\mathbb{C}$ and hence not over $\mathbb{Q}$. This is because if it had been, then the substitution of $\xi_i = h_i$ will also yield the value zero irrespective of the choice of h_i, making the determinant zero even for the Gaussian case, leading to a contradiction. Therefore, f_ω is a nonzero polynomial. We also observe that the degree of f_ω in each of the variable ξ_i is at most one. The lemma below, easily proved using elementary algebra, shows that it is possible to identify a finite field $\mathbb{F}_p$ and an allocation to $\{\xi_i\}$ with numbers from $\mathbb{F}_p$ such that f_ω does not vanish (see [49] for instance).

Lemma E.1: Given a polynomial $f(\xi_1, \xi_2, \ldots, \xi_N)$ with integer coefficients, which is not identically zero, there exists a prime field $\mathbb{F}_p$ with p large enough, such that the polynomial evaluates to a nonzero value at least for one assignment of field values to the formal variables.

However, we wish to ensure that the above condition is met for every cut in the network. To do so, consider the polynomial

$$f(\xi_1, \xi_2, .., \xi_N) := \prod_{\omega \in \Omega} f_\omega(\xi_1, \xi_2, .., \xi_N). \tag{53}$$

Now, the polynomial f is nonzero since it is a product of nonzero polynomials f_ω and the degree of f in any of the variables is at-most $|\Omega|$. We want a field $\mathbb{F}_p$ and an assignment for ξ_i from the field such that f is nonzero. Using Lemma E.1, such an assignment exists. Let us choose that p and the assignment that makes f nonzero. Thus we have a deterministic wireless network whose capacity is guaranteed to be greater than or equal to the upper bound on MMG.

Next, we prove that the zero error capacity, C_{ZE}, of a linear deterministic network, is equal to its ϵ-error capacity.

Definition 5: [36] The zero error capacity of a channel is defined as the supremum of all achievable rates across the channels such that the probability of error is exactly zero.

Theorem E.2: The zero error capacity of a ss-ss deterministic wireless network is equal to

$$C_{ZE} = \min_{\omega \in \Omega} \text{rank}(G_\omega).$$

This capacity can be achieved using a linear code and linear transformations in all relays.

Proof: We will prove this theorem using the ϵ-error capacity result from Theorem 3.2. Given an ϵ-error achieving scheme from Theorem 3.2, we have a codebook $\mathcal{C}$ of rate r and linear transformations A_j used by the relays with T being the block length and L the number of blocks for which the network is operated. The matrix of transformation from the source to the destination is $y = Gx$. A given codeword either always results in error at the destination or is always decoded correctly due to the deterministic nature of the channel. The fraction of erroneous codewords is less than ϵ. If we expurgate the codewords that do cause error, then the new codebook of size $\bar{r} = r - \frac{\log\{(1-\epsilon)^{-1}\}}{LT}$ can be decoded error free, which approaches r as LT becomes large. The zero-error capacity is therefore the same as the ϵ-error capacity, which equals the min-cut rank r. Now, there are at least $p^{\bar{r}LT}$ vectors in the rangespace $\mathcal{R}(G)$ in order to have a zero error rate of $\bar{r}$ over LT time instants. Since $\mathcal{R}(G) = p^{\text{rank}(G)}$, the transfer matrix G is at least of rank $\bar{r}LT$. Therefore a linear codebook can be used at the source in order to achieve the rate $\bar{r}$. ■

Finally, we lift the achievability strategy of zero-error capacity in the equivalent deterministic networks to arrive at an achievable strategy for MMG in corresponding fading network.

In the reduced deterministic network of a fading network, to achieve the zero-error capacity, the relays perform matrix operations A_i on received vectors for T time durations and for L blocks with $M = LT$. Since each received vector is of size q, the matrix A_i is of size $qT \times qT$. Now we use the same strategy for the fading network, i.e., all relays use the same matrices A_i, that are obtained via the zero-error strategy in the reduced deterministic network. Though the entries of A_i belong to $\mathbb{F}_p$, they can be treated as integers by identifying the elements of $\mathbb{F}_p$ with the integers $0, 1, \ldots, (p-1)$. Therefore the matrices A_i can also be interpreted as matrices over $\mathbb{C}$. By using linear maps A_i at relays in the fading network, we get an induced channel matrix $\mathbf{H}$, and effective channel would be of the form, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$. As is shown in Theorem 2.7, MMG offered by this channel is equal to $\mathbb{R}\text{ank}(\mathbf{H})$. We shall prove that the MMG offered by this induced channel is greater than or equal to $\bar{r}M$, i.e., we show that $\mathbb{R}\text{ank}(\mathbf{H}) \geq \bar{r}M$. That is equivalent to showing that there exists an assignment of $\mathbf{h}_i = h_i$ in the fading network such that $\text{rank}(H) \geq \bar{r}M$.

In the proof of Theorem E.2, we created a transfer matrix G of size $qM \times qM$ with rank greater than or equal to $\bar{r}M$. Now we have a similar transfer matrix $\mathbf{H}$ in the fading network. If we assign the underlying random variables $\mathbf{h}_i$ to be equal to ξ_i, again by identifying the elements of $\mathbb{F}_p$ with the integers $0, 1, , \ldots, (p-1)$, we have an assignment of $\mathbf{H}$ that has rank at least $\bar{r}M$ (since the rank of a matrix over $\mathbb{C}$ is greater than or equal to its rank over $\mathbb{F}_p$). Since the structural rank is the maximum possible rank under any assignment, we get that,

$$\mathbb{R}\text{ank}(\mathbf{H}) \geq \text{rank}(G) \geq \bar{r}M.$$

The induced channel therefore has a MMG equal to $\bar{r}$ by Theorem 2.7. Thus the cut-set bound is achieved, and hence the MMG of the ss-ss fading network is given by

$$D = \min_{\{\omega \in \Lambda\}} \mathbb{R}\text{ank}(\mathbf{H}_\Lambda).$$

APPENDIX F
PROOF OF LEMMA 4.1

With $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M^T]$, $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_M]^T$ and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \ddots & \\ & & & \mathbf{H}_M \end{bmatrix} \quad (54)$$

the parallel channel is given by $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$. We know independent inputs are optimal for this parallel channel, so we will choose the $\mathbf{x}_i$ to be independent. We have

$$\begin{aligned} &\Pr\{I(\mathbf{x}; \mathbf{y} \mid \mathbf{H} = H) \leq r \log \rho\} \\ &= \Pr\{\sum_{i=1}^{M} I(\mathbf{x}_i; \mathbf{y}_i | \mathbf{H}_i = H_i) \leq r \log \rho\}. \end{aligned}$$

Define $\mathbf{Z}_i := I(\mathbf{x}_i; \mathbf{y}_i | \mathbf{H}_i = H_i)$. Thus $\mathbf{Z}_i$ is a function of the channel realization H_i. Since $\{\mathbf{H}_i\}$ are independent by the hypothesis of the lemma and $\mathbf{x}_i$ are independent by the argument above, $\{\mathbf{Z}_i\}$ are also independent.

Let $R = r\ \log(\rho)$ and $R_i = r_i\ \log(\rho)$ for $i = 1, 2, \ldots, M$. Our next goal is to evaluate $\Pr\{\ \sum_{i=1}^{M} \mathbf{Z}_i \leq r \log(\rho)\ \}$. To do this, we first consider the case when $M = 2$ and we evaluate $\Pr\{\ \mathbf{Z}_1 + \mathbf{Z}_2 \leq r \log(\rho)\ \}$. Then we extend this to general M by induction. We define

$$\begin{aligned} F_{Z_i}(R_i) &:= \Pr\{\mathbf{Z}_i < R_i\}, f_{Z_i}(R_i) := \frac{d}{dR_i} F_{Z_i}(R_i) \\ \text{Let } F_{Z_i}(R_i) &\doteq \rho^{-d_i(r_i)} \\ \text{Then } f_{Z_i}(R_i) &\doteq \frac{d}{dr_i\ \log(\rho)} \rho^{-d_i(r_i)} \\ &\doteq \rho^{-d_i(r_i)} (-1) \frac{d}{dr_i} d_i(r_i) \\ &\doteq \rho^{-d_i(r_i)}. \end{aligned}$$

The last equation follows since $d_i(r_i)$ is a decreasing function making the derivative negative. Now

$$\begin{aligned} \Pr(\mathbf{Z}_1 + \mathbf{Z}_2 \leq R) &= \int_0^\infty f_{Z_1}(R_1) F_{Z_2}(R - R_1) dR_1 \\ &\doteq \int_0^\infty \rho^{-d_1(r_1)}\ \rho^{-d_2(r-r_1)} \log(\rho) d(r_1). \end{aligned}$$

By Varadhan's Lemma [44], the SNR exponent of the integral is given by

$$\begin{aligned} d(r) &= \inf_{r_1 \geq 0} d_1(r_1) + d_2(r - r_1) \\ &= \inf_{(r_1, r_2): r_1 + r_2 = r} \sum_{i=1}^{2} d_i(r_i). \end{aligned}$$

Proceeding by induction, we get for the general case with M parallel channels that

$$d(r) = \inf_{(r_1, r_2, \ldots, r_M): \sum_{i=1}^{M} r_i = r} \sum_{i=1}^{M} d_i(r_i).$$

ACKNOWLEDGMENT

The authors thank K. Vinodh and M. Anand for several useful discussions. The authors would also like to thank the associate editor M. Gastpar and the anonymous reviewers for suggestions that led to significant improvement in the presentation.

REFERENCES

[1] K. Sreeram, S. Birenjith, and P. V. Kumar, "DMT of parallel-path and layered networks under the half-duplex constraint," *IEEE Trans. Inf. Theory*, submitted for publication.

[2] T. M. Cover and A. A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[4] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. IEEE Infocom*, New York, Jun. 2002.

[5] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[6] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[7] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[8] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[9] P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communications," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2041–2057, Jun. 2005.

[10] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[11] M. Yuksel and E. Erkip, "Multiple-antenna cooperative wireless systems: A diversity–multiplexing tradeoff perspective," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3371–3393, Oct. 2007.

[12] P. Elia, K. Vinodh, M. Anand, and P. V. Kumar, "D-MG tradeoff and optimal codes for a class of AF and DF cooperative communication protocols," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3161–3185, Jul. 2009.

[13] P. Elia and P. V. Kumar, "Space-time codes that are approximately universal for the parallel, multiblock and cooperative DDF channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Jun.–Jul. 28–3, 2009.

[14] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 647–663, Feb. 2007.

[15] S. Yang and J.-C. Belfiore, "Diversity of MIMO multihop relay channels," *IEEE Trans. Inf. Theory*, submitted for publication.

[16] S. Yang and J.-C. Belfiore, "Towards the optimal amplify-and-forward cooperative diversity scheme," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3114–3126, Sep. 2007.

[17] K. Sreeram, S. Birenjith, and P. V. Kumar, "Multihop cooperative wireless networks: Diversity multiplexing tradeoff and optimal code design," in *Proc. Inf. Theory and Appl. Workshop, UCSD*, Feb. 2008.

[18] K. Sreeram, S. Birenjith, K. Vinod, M. Anand, and P. V. Kumar, "On the throughput, DMT and optimal code construction of the K-parallel-path cooperative wireless fading network," in *Proc. 10th Int. Symp. Wireless Pers. Multimedia Commun.*, Dec. 2007.

[19] K. Sreeram, S. Birenjith, and P. V. Kumar, "DMT of multihop cooperative networks-Part I: K-parallel-path-networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, Canada, Jul. 2008.

[20] K. Sreeram, S. Birenjith, and P. V. Kumar, "Diversity and degrees of freedom of cooperative wireless networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, Canada, Jul. 2008.

[21] K. Sreeram, S. Birenjith, and P. V. Kumar, Multihop Cooperative Wireless Networks: Diversity Multiplexing Tradeoff and Optimal Code Design [Online]. Available: http://arxiv.org/pdf/0802.1888

[22] K. Sreeram, S. Birenjith, and P. V. Kumar, DMT of Multi-Hop Cooperative Networks—Part I: Basic Results [Online]. Available: http://arxiv.org/pdf/0802.0234

[23] K. Sreeram, S. Birenjith, and P. V. Kumar, On the Throughput, DMT and Optimal Code Construction of the K-Parallel-Path Cooperative Wireless Fading Network 2007, USC CSI Technical Report, CSI-2007-06-07.
[24] S. Pawar, S. Avestimehr, and D. Tse, "Diversity multiplexing tradeoff of the half-duplex relay channel," in *Proc. 45th Allerton Conf. Commun., Contr., Comput.*, Sep. 2007.
[25] S. Borade, L. Zheng, and R. Gallager, "Amplify and forward in wireless relay networks: Rate, diversity and network size," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3302–3318, Oct. 2007.
[26] C. Rao and B. Hassibi, "Diversity-multiplexing gain trade-off of a MIMO system with relays," in *Proc. IEEE Inf. Theory Workshop*, Norway, Jul. 2007.
[27] C. Rao, "Asymptotic analysis of wireless systems with Rayleigh fading," Ph.D., Calif. Inst. Technol., Pasadena, 2007.
[28] S. O. Gharan, A. Bayesteh, and A. K. Khandani, "On the diversity-multiplexing tradeoff in multiple-relay network," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5423–5444, Dec. 2009.
[29] S. O. Gharan, A. Bayesteh, and A. K. Khandani, "Diversity-multiplexing tradeoff in multiple-relay network-Part I: Proposed scheme and single-antenna networks," in *Proc. 46th Allerton Conf. Commun., Contr., Comput.*, Sep. 2008.
[30] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory.*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
[31] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
[32] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the K user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
[33] S. A. Jafar and S. Shamai, "Degrees of freedom region for the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.
[34] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
[35] V. R. Cadambe, S. A. Jafar, and S. Shamai, "Interference alignment on the deterministic channel and application to fully connected AWGN interference networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 274–296, Jan. 2009.
[36] C. E. Shannon, "The zero-error capacity of a noisy channel," *IEEE Trans. Inf. Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.
[37] F. Oggier and B. Hassibi, "Code design for multihop wireless relay networks," *EURASIP J. Adv. Signal Process.*, Nov. 2007, to be published.
[38] S. M. S. T. Yazdi and S. A. Savari, "A combinatorial study of linear deterministic relay networks," in *Proc. Inf. Theory Workshop*, Cairo, Egypt, Jan. 2010.
[39] R. Vaze and R. W. Heath Jr., "Maximizing reliability in multihop wireless networks with cascaded space-time codes," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, Feb. 2008.
[40] R. Vaze and R. W. Heath Jr., "To code in space and time or not in multihop relay channels," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2736–2747, Jul. 2009.
[41] H.-F. Lu, "Explicit construction of multiblock space-time codes that achieve the diversity-multiplexing gain tradeoff," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, 2006.
[42] M. Godavarti and A. O. Hero, III, "Diversity and degrees of freedom in wireless communications," in *Proc. IEEE ICASSP*, May 2002, vol. 3, pp. 2854–2861.
[43] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Adeliade, 2005.
[44] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer-Verlag, 1998.
[45] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*. New York: Academic , 1979.
[46] A. E. Gamal, "On information flow in relay networks," in *Proc. IEEE Nat. Telecomm. Conf.*, Nov. 1981, vol. 2, pp. D4.1.1–D4.1.4.
[47] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley , 2006.
[48] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.
[49] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
[50] D. B. West, *Introduction to Graph Theory*. Englewood Cliffs, NJ: Prentice-Hall, 2000.
[51] A. R. Amir-Moez, "Extreme properties of eigenvalues of a hermitian transformation and singular values of the sum and product of linear transformations," *Duke Math. J.*, vol. 23, no. 3, pp. 463–476, 1956.
[52] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Diversity order bounds for wireless relay networks," in *Proc. Wireless Commun. Netw. Conf.*, Mar. 2007.
[53] S. Wei, "Diversity multiplexing tradeoff of asynchronous cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4150–4172, Nov. 2007.
[54] Z. Li and X.-G. Xia, "A simple alamouti space-time transmission scheme for asynchronous cooperative systems," *IEEE Signal Process. Lett.*, vol. 15, no. 1, pp. 804–807, Jan. 2008.
[55] G. S. Rajan and B. S. Rajan, "Multigroup ML decodable collocated and distributed space time block codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3221–3247, Jul. 2010.
[56] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," in *Proc. IEEE Globecom*, Dec. 2006.
[57] Z. Yi and I.-M. Kim, "Single-symbol ML decodable distributed STBCs for cooperative networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2977–2985, Aug. 2007.

K. Sreeram received the B.E. degree in electronics and communication engineering from the College of Engineering, Guindy, Anna University, and the M.E. degree from the Indian Institute of Science, Bangalore, in 2008.

Since 2008, he has been a Ph.D. student at the University of Illinois Urbana-Champaign. His research interests include information theory and wireless networks.

Mr. Sreeram is a recipient of the S.V.C. Aiya medal from the Indian Institute of Science (2008), a recipient of a Qualcomm Roberto Padovani Scholarship (2010), a corecipient of Intel India Student Research Contest first prize (2006), and a corecipient of the Qualcomm Cognitive Radio Contest prize (2010).

S. Birenjith received the B.Tech. degree in electronics and communication engineering from the College of Engineering, Trivandrum, and the M.Sc.(Engg.) from the Indian Institute of Science, Bangalore, in 2008.

During 2008–2011, he was on the faculty of Electronics and Communication Engineering, Govt Engineering Colleges, Kerala. He is currently pursuing the Ph.D. degree at the Department of Electrical Communication, Indian Institute of Science.

P. Vijay Kumar (S'80–M'82–SM'01–F'02) received the B.Tech. and M.Tech. degrees from the Indian Institutes of Technology (Kharagpur and Kanpur), and the Ph.D. degree from the University of Southern California (USC), Los Angeles, in 1983, all in electrical engineering.

From 1983 to 2003, he was on the faculty of the EE-Systems Department, USC. Since 2003, he has been on the faculty of the Indian Institute of Science, Bangalore. He also holds the position of Adjunct Research Professor at USC. His current research interests include codes for distributed storage, distributed function computation, sensor networks, and space-time codes for MIMO and cooperative communication networks.

Dr. Kumar is an ISI highly cited author. He is also a corecipient of the 1995 IEEE Information Theory Society Prize Paper Award, as well as of a Best Paper Award at the DCOSS 2008 conference on sensor networks.