2011 IEEE International Symposium on Information Theory Proceedings

Worst-Case Linear Transmit Strategies with Limited Relay Information in Gaussian Relay Networks

Steven Corroy and Rudolf Mathar
Institute for Theoretical Information Technology
RWTH Aachen University
D-52074 Aachen, Germany
Email: {corroy, mathar}@ti.rwth-aachen.de

Abstract—**In practical large scale relay networks, it is often an unfeasible task to adapt the strategy of the relays to each source/sink pair communicating through the network. In this paper we consider the strategy of the relays as fixed and chosen in advance. In this context, it is a crucial problem for the source node transmitting information, to know the strategy of the relays in the network. Therefore we optimize the transmit strategy of the source node, assuming only partial *relay information* knowledge, i.e., channel state information and relay strategies. We also study the impact of getting complete relay information from certain relays only, derive the reduction in the problem uncertainty due to this knowledge and evaluate numerically the performance of the network.**

I. Introduction

Relay networks are models for communication systems where one or more sources transmit information to one or more sinks through relays. In order to reduce the complexity of relays, a lot of attention has been paid to amplify-and-forward (AF) relay networks (e.g., [1]–[5]). Less work has focused on the case where the channel state information (CSI) is unknown or incomplete at the source/relays/sink nodes. In [6], the authors developed power allocation strategies for a three-terminal AF relay network requiring only the knowledge of the mean of the channel gains. In [7], the authors derive the optimal relay precoders for a parallel relay network with K relays and partial CSI at relays and sink. In [8], the authors consider a 2-hop network and develop an opportunistic relaying strategy using distributed space-time codes requiring only statistical CSI at the relays.

The present paper differs from these previous works by two main aspects. First the present work considers general networks with no specific number of hops or topology. The only mathematical restriction is that our networks must be representable by a directed graph, which is not a limitation from the practical point of view. The second, more important difference, is that we consider the strategies of the relays to be fixed, chosen in advance and that the source node shall optimize its strategy given this fixed relaying choice. It is indeed very costly and unpractical to adapt the strategy of the relays in a network for each source/sink communication using the network. A simple method for a network owner would be to set the amplification factors of the different relays to a random value or a predefined value.

This work was supported by the UMIC research cluster of the RWTH Aachen University.

In this context, it is crucial for the source node to know the relay strategies. This knowledge is however expensive to gather since all relays participating in a communication should first transmit their strategy to the source node. To save energy and bandwidth it can be advantageous to assume only partial knowledge of relay strategy and CSI at the source node and develop a worst-case transmit strategy. Note that if the source/sink pairs are fixed it could have been an option to gather once the relay strategies at all sources. In a real network however, the source/sink pairs change constantly, using each time different sets of relays, which renders this simple approach infeasible. To tackle this problem we first define relay information (RI) as CSI and relaying strategy, in other words the RI of a relay consists in its amplification factors (i.e. its strategy) and its CSI as a receiver (CSIR). The strategy of the source node strongly depends on its RI knowledge.

The contribution of this paper is manifolds. We present a framework, based on our recent work [9], enabling to study the impact of partial RI at the source node while assuming complete RI at the sink. We develop a worst-case optimization solution for the source precoder in order to cope with missing RI. We explain how gained RI from specific relays decreases the problem uncertainty and evaluate the performance of the network with different amount of RI.

The rest of the present paper is organized as follow. In Section II, we describe our network model, in Section III we present the optimization problem, in Section IV we propose a solution to this problem, in Section V we study the influence of gained RI on the overall problem uncertainty, in Section VI we show numerical results validating our approach and finally, Section VII concludes this work.

II. Relay Network Model

We consider networks that can be represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a vertex set $\mathcal{V}$ and an edge set $\mathcal{E}$ with $|\mathcal{E}| = e$. The network has one source node, one sink node and arbitrarily many relays. The source node transmits a vector $\mathbf{x} \in \mathbb{C}^n$ to the sink node which decodes a vector $\hat{\mathbf{x}} \in \mathbb{C}^n$. We assume that $\mathbf{x}$ is zero-mean, normalized and uncorrelated such that $\mathbf{x} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_n)$. For simplicity we assume that the

978-1-4577-0594-6/11/$26.00 ©2011 IEEE

source has n outgoing edges, the sink has n incoming edges, the min-cut of the network is at least n and the source has no direct connection to the sink. The relays can linearly combine incoming signals and forward them to the next relays or sink. We illustrate an example network in Figure 1.

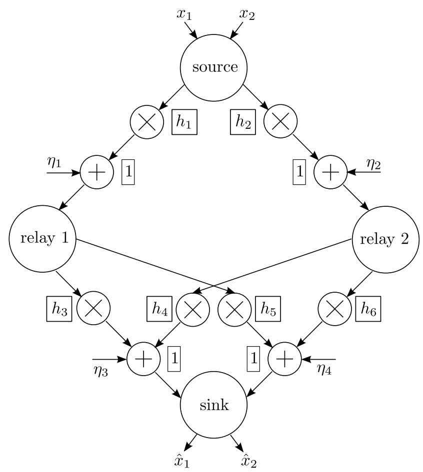

Fig. 1. Amplify and forward relay network.

We call a connection, a chain of edges, $\times$ nodes and $+$ nodes between the source or a relay and a relay or the sink. As explained in more detail in our work [9], a $\times$ node models the channel gain of a connection and a $+$ node the Gaussian noise entering a connection. The vector $\mathbf{h}$ contains the channel gains in the network (a $+$ node has always a channel gain of 1) and $\boldsymbol{\eta} \in \mathbb{C}^N$ is the noise vector applied to the N noisy connections of the network such that $\boldsymbol{\eta} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{C}_{\boldsymbol{\eta}})$.

We numerate edges in the network from 1 to e and define (similar to [10], [9]) the matrix $\mathbf{A} \in \mathbb{C}^{e\times(n+N)}$ which contains 1) the linear coefficients chosen by the source node to send information through the network and 2) a coefficient of value 1 where noise enters a $+$ node. Since we can always numerate the edges outgoing from the source first, the matrix $\mathbf{A}$ can be represented as a block matrix with the following structure

$$\mathbf{A} = \begin{bmatrix} \mathbf{A_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\boldsymbol{\eta}} \end{bmatrix}, \tag{1}$$

where $\mathbf{A_x} \in \mathbb{C}^{n\times n}$ is a precoding matrix applied by the source to the transmitted signal, $\mathbf{A}_{\boldsymbol{\eta}} \in \mathbb{C}^{(e-n)\times N}$ is a matrix composed of zeros and ones which distributes the noise on the edges of the network. Further we define $\mathbf{F} \in \mathbb{C}^{e\times e}$ which contains 1) the amplification factors $\tilde{f}_{ij} \in \mathbb{C}$ chosen by the relays to forward signals 2) the channel coefficients represented by $\times$ nodes and 3) a coefficient of value 1 for $+$ nodes. The matrix $\mathbf{F}$ has the form

$$f_{ij} = \begin{cases} \tilde{f}_{ij} & \exists \text{ direct flow from edge } j \text{ to } i \text{ through a relay} \\ h_i & \exists \text{ direct flow from edge } j \text{ to } i \text{ through a} \times \text{node} \\ 1 & \exists \text{ direct flow from edge } j \text{ to } i \text{ through a} + \text{node} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Finally we define $\mathbf{B} \in \mathbb{C}^{e\times n}$ which represents the coefficients chosen by the sink to filter information out of the network. Since we can always numerate the edges incoming at the sink last, the matrix $\mathbf{B}$ can be represented as a block matrix with the following structure

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B_x} \end{bmatrix}, \tag{3}$$

where $\mathbf{B_x} \in \mathbb{C}^{n\times n}$ is a filter applied to the received signal to get $\hat{\mathbf{x}}$. Note that $\mathbf{F}$ represents the amplification of the signal after one hop. Therefore $\mathbf{F}^k$ represents the amplification of the signal after k hops. Interestingly $\mathbf{F}$ is strictly lower diagonal and is therefore a nilpotent matrix, so there exists a power q such that $\mathbf{F}^q = \mathbf{0}$. It is possible to represent the effective amplification of the signal at each edge of the network, with only one lower diagonal matrix $\mathbf{M}$ of size $e \times e$,

$$\mathbf{M} = \mathbf{I}_e + \mathbf{F} + \mathbf{F}^2 + \cdots + \mathbf{F}^{q-1} = (\mathbf{I}_e - \mathbf{F})^{-1}. \tag{4}$$

We have shown in [9] that

$$\hat{\mathbf{x}} = \mathbf{B}^{\mathrm{H}}\mathbf{M}\mathbf{A} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\eta} \end{bmatrix} = \mathbf{B}^{\mathrm{H}}(\mathbf{I}_e - \mathbf{F})^{-1}\mathbf{A} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\eta} \end{bmatrix}. \tag{5}$$

Given that we can always numerate the outgoing edges of the source first, the incoming edges at the sink last and that there is no direct connection between the source and the sink, $\mathbf{F}$ can be written as

$$\mathbf{F} = \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{H} & \mathbf{L} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times n} & \mathbf{N} & \mathbf{0}_{n\times n} \end{bmatrix}, \tag{6}$$

with $m = e-2n$, $\mathbf{H} \in \mathbb{C}^{m\times n}$ contains the channel coefficients of connections between the source and relays, $\mathbf{N} \in \mathbb{C}^{n\times m}$ contains ones for noise components entering connections between relays and the sink and $\mathbf{L} \in \mathbb{C}^{m\times m}$ a strictly lower triangular matrix containing all other coefficients of $\mathbf{F}$ as described in (2). It is simple to calculate, for $n > 1$ (this hypothesis simply enables to write more compactly the term in the middle of the matrix and does not influence the rest of the paper), that $\mathbf{M}$ is equal to

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ (\mathbf{I}_m - \mathbf{L})^{-1}\mathbf{H} & (\mathbf{I}_m - \mathbf{L})^{-1} & \mathbf{0}_{m\times n} \\ \mathbf{N}(\mathbf{I}_m - \mathbf{L})^{-1}\mathbf{H} & \mathbf{N}(\mathbf{I}_m - \mathbf{L})^{-1} & \mathbf{I}_n \end{bmatrix}. \tag{7}$$

By putting (1), (3) and (7) together into (5) we can express $\hat{\mathbf{x}}$ as

$$\hat{\mathbf{x}} = \mathbf{B_x}^{\mathrm{H}}(\mathbf{M_x}\mathbf{A_x}\mathbf{x} + \mathbf{M}_{\boldsymbol{\eta}}\mathbf{A}_{\boldsymbol{\eta}}\boldsymbol{\eta}), \tag{8}$$

where we have defined $\mathbf{M_x} \triangleq \mathbf{N}(\mathbf{I}_m - \mathbf{L})^{-1}\mathbf{H}$ and $\mathbf{M}_{\boldsymbol{\eta}} \triangleq [\mathbf{N}(\mathbf{I}_m - \mathbf{L})^{-1} \quad \mathbf{I}_n]$. Note that this network model is very general and enables to represent different kind of multiplexing, in particular it is possible to model a multihop multiple-input multiple-output (MIMO) relay network.

III. Worst-Case Optimization Problem Formulation

As mentioned previously, we want to consider the strategy of the relays to be the same for any source/sink pair and therefore we assume that $\mathbf{F}$ is fixed and cannot be optimized. This assumption clearly raises the question, how does the network perform when the source node has only partial knowledge of the channel gains and more important of the relaying strategy. To answer this question we optimize the matrix $\mathbf{A_x}$ to maximize the mutual information between the source and the sink given partial RI. The matrix $\mathbf{B_x}$ is taken to be the minimum mean-square error (MMSE) filter.

We model uncertainty on the relay strategies and on the channel gains, i.e., the uncertainty on $\mathbf{F}$, as $\mathbf{H} = \hat{\mathbf{H}} + \mathbf{\Delta_H}$ with $\|\mathbf{\Delta_H}\|_\mathrm{F} \le \epsilon_H$ where $\hat{\mathbf{H}}$ is a known matrix, $\mathbf{\Delta_H}$ is an uncertainty component with limited norm and $\|\mathbf{\Delta_H}\|_\mathrm{F} = \sqrt{\mathrm{Tr}(\mathbf{\Delta_H}\mathbf{\Delta_H}^\mathrm{H})}$ is the Frobenius norm of $\mathbf{\Delta_H}$. Similar we have $\mathbf{L} = \hat{\mathbf{L}} + \mathbf{\Delta_L}$ with $\|\mathbf{\Delta_L}\|_\mathrm{F} \le \epsilon_L$. Since $\mathbf{N}$ is only composed of zeros and ones, depending only on the physical structure of the network, we consider $\mathbf{N}$ as certain.

We would like to express the uncertainty on $\mathbf{M_x}$ and $\mathbf{M}_{\boldsymbol{\eta}}$ caused by the uncertainty on $\mathbf{F}$. We have

$$\begin{aligned}\mathbf{M_x} &= \mathbf{N}(\mathbf{I}_m - \mathbf{L})^{-1}\mathbf{H} \\ &= \mathbf{N}(\mathbf{I}_m - \hat{\mathbf{L}} - \mathbf{\Delta_L})^{-1}(\hat{\mathbf{H}} + \mathbf{\Delta_H}).\end{aligned} \tag{9}$$

Lemma 1. *(Woodbury Identity) Let $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{n\times n}$. Assume $\mathbf{X}$ and $\mathbf{X}+\mathbf{Y}$ are invertible, it holds*

$$(\mathbf{X}+\mathbf{Y})^{-1} = \mathbf{X}^{-1} - \mathbf{X}^{-1}\mathbf{Y}(\mathbf{I}_n + \mathbf{X}^{-1}\mathbf{Y})^{-1}\mathbf{X}^{-1}.$$

Using Lemma 1 and defining $\hat{\mathbf{L}}' \triangleq (\mathbf{I}_m - \hat{\mathbf{L}})^{-1}$, we can write (9) as

$$\mathbf{M_x} = \mathbf{N}(\hat{\mathbf{L}}' + \hat{\mathbf{L}}'\mathbf{\Delta_L}(\mathbf{I}_m - \hat{\mathbf{L}}'\mathbf{\Delta_L})^{-1}\hat{\mathbf{L}}')(\hat{\mathbf{H}} + \mathbf{\Delta_H}). \tag{10}$$

Now we can write

$$\mathbf{M_x} = \hat{\mathbf{M}}_\mathbf{x} + \mathbf{\Delta}_{\mathbf{M_x}} \tag{11}$$

with

$$\hat{\mathbf{M}}_\mathbf{x} \triangleq \mathbf{N}\hat{\mathbf{L}}'\hat{\mathbf{H}} = \mathbf{N}(\mathbf{I}_m - \hat{\mathbf{L}})^{-1}\hat{\mathbf{H}} \tag{12}$$

and $\mathbf{\Delta}_{\mathbf{M_x}}$ an uncertainty matrix for which we need to bound the norm. For a given matrix $\mathbf{X}$, we define $\lambda_i(\mathbf{X})$ and $\sigma_i(\mathbf{X})$ as the i-th eigenvalue and respectively the i-th singular value of $\mathbf{X}$. $\sigma_{\max}(\mathbf{X}) = \sigma_1(\mathbf{X})$ is the largest singular value of $\mathbf{X}$.

Lemma 2. *Let $\mathbf{X} \in \mathbb{C}^{n\times n}$. Assume that $\mathbf{X}$ is singular and $\mathbf{I}_n + \mathbf{X}$ is invertible, it holds*

$$\|(\mathbf{I}_n + \mathbf{X})^{-1}\|_\mathrm{F} \le \sqrt{n}.$$

Proof:

$$\begin{aligned}\|(\mathbf{I}_n + \mathbf{X})^{-1}\|_\mathrm{F} &= \sqrt{\textstyle\sum_{i=1}^n \sigma_i((\mathbf{I}_n+\mathbf{X})^{-1})^2} \\ &= \sqrt{\textstyle\sum_{i=1}^n \sigma_i(\mathbf{I}_n+\mathbf{X})^{-2}} \\ &= \sqrt{\textstyle\sum_{i=1}^n \lambda_i((\mathbf{I}_n+\mathbf{X})^\mathrm{H}(\mathbf{I}_n+\mathbf{X}))^{-1}} \\ &\le \sqrt{\textstyle\sum_{i=1}^n (1+2\lambda_n(\mathbf{X}) + \lambda_i(\mathbf{X}^\mathrm{H}\mathbf{X}))^{-1}} \\ &\overset{(a)}{=} \sqrt{\textstyle\sum_{i=1}^n (1+\sigma_i(\mathbf{X}))^{-1}} \\ &\le \sqrt{\textstyle\sum_{i=1}^n (1+\sigma_n(\mathbf{X}))^{-1}} \\ &\overset{(b)}{=} \sqrt{n},\end{aligned}$$

where (a) and (b) comes from the fact that $\mathbf{X}$ is singular, i.e., $\lambda_n(\mathbf{X}) = 0$ and $\sigma_n(\mathbf{X}) = 0$. ■

Using Lemma 2, we can bound the norm of $\mathbf{\Delta}_{\mathbf{M_x}}$ as

$$\begin{aligned}\|\mathbf{\Delta}_{\mathbf{M_x}}\|_\mathrm{F} &\le \sigma_{\max}(\mathbf{N}\hat{\mathbf{L}}')(\epsilon_H + \sqrt{m}\epsilon_L(\sigma_{\max}(\hat{\mathbf{L}}'\hat{\mathbf{H}}) + \sigma_{\max}(\hat{\mathbf{L}}')\epsilon_H)) \\ &\triangleq \epsilon_x.\end{aligned} \tag{13}$$

Similarly we can write

$$\mathbf{M}_{\boldsymbol{\eta}} = \hat{\mathbf{M}}_{\boldsymbol{\eta}} + \mathbf{\Delta}_{\mathbf{M}_\eta} \tag{14}$$

with

$$\hat{\mathbf{M}}_{\boldsymbol{\eta}} = [\mathbf{N}(\mathbf{I}_m - \hat{\mathbf{L}})^{-1} \quad \mathbf{I}_n] \tag{15}$$

and $\mathbf{\Delta}_{\mathbf{M}_\eta}$ an uncertainty matrix with

$$\|\mathbf{\Delta}_{\mathbf{M}_\eta}\|_\mathrm{F} \le \sigma_{\max}(\mathbf{N}\hat{\mathbf{L}}')\sigma_{\max}(\hat{\mathbf{L}}')\sqrt{m}\epsilon_L \triangleq \epsilon_\eta. \tag{16}$$

We would like to maximize the mutual information between the transmitted signal and the received signal $I(\mathbf{x}, \hat{\mathbf{x}})$. We define the matrix $\mathbf{E}$ as the mean square error (MSE) matrix defined as $\mathbf{E} = \mathbb{E}[(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^\mathrm{H}]$. As shown in [9], $I(\mathbf{x}, \hat{\mathbf{x}})$ can be written as follow

$$\begin{aligned}I(\mathbf{x}, \hat{\mathbf{x}}) &= -\log\det(\mathbf{E}) \\ I(\mathbf{x}, \hat{\mathbf{x}}) &= \log\det(\mathbf{I}_n + \mathbf{A_x}^\mathrm{H}\mathbf{M_x}^\mathrm{H}(\mathbf{M}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{M}_{\boldsymbol{\eta}}^\mathrm{H})^{-1}\mathbf{M_x}\mathbf{A_x})\end{aligned} \tag{17}$$

with $\mathbf{N}_{\boldsymbol{\eta}} \triangleq \mathbf{A}_{\boldsymbol{\eta}}\mathbf{C}_{\boldsymbol{\eta}}\mathbf{A}_{\boldsymbol{\eta}}^\mathrm{H}$. Now we define $\mathbf{R}$ as

$$\mathbf{R} = \mathbf{M_x}^\mathrm{H}(\mathbf{M}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{M}_{\boldsymbol{\eta}}^\mathrm{H})^{-1}\mathbf{M_x} \tag{18}$$

and the optimization problem we want to solve is

$$\begin{aligned}&\underset{\mathbf{A_x}}{\text{maximize}} \quad \underset{\mathbf{\Delta}_{\mathbf{M_x}}, \mathbf{\Delta}_{\mathbf{M}_\eta}}{\text{infimum}} (\log\det(\mathbf{I}_n + \mathbf{A_x}^\mathrm{H}\mathbf{R}\mathbf{A_x})) \\ &\text{subject to} \quad \mathrm{Tr}(\mathbf{A_x}\mathbf{A_x}^\mathrm{H}) \le P_T,\end{aligned} \tag{19}$$

where P_T is the available power at the source node. Note that although we only express a power constraint for the source, the relays are obviously also power constrained. Simply the source node is not concerned with the power constraint of relays.

IV. Worst-Case Solution

Similar to [11, pp. 228–229], we need to find a matrix $\breve{\mathbf{R}}$ such that for all $(\mathbf{\Delta}_{\mathbf{M_x}}, \mathbf{\Delta}_{\mathbf{M}_\eta})$ with $\|\mathbf{\Delta}_{\mathbf{M_x}}\|_\mathrm{F} \le \epsilon_x$ and $\|\mathbf{\Delta}_{\mathbf{M}_\eta}\|_\mathrm{F} \le \epsilon_\eta$ we have $\breve{\mathbf{R}} \preccurlyeq \mathbf{R}$. If we find such a matrix, we have $\log\det(\mathbf{I}_n + \mathbf{A_x}^\mathrm{H}\breve{\mathbf{R}}\mathbf{A_x}) \le \log\det(\mathbf{I}_n + \mathbf{A_x}^\mathrm{H}\mathbf{R}\mathbf{A_x}), \forall(\mathbf{\Delta}_{\mathbf{M_x}}, \mathbf{\Delta}_{\mathbf{M}_\eta})$.

A. Worst-Case $\mathbf{M}_{\boldsymbol{\eta}}$ for an arbitrary $\mathbf{M_x}$

We first define $\mathbf{R}_{\boldsymbol{\eta}} = \mathbf{M}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{M}_{\boldsymbol{\eta}}^\mathrm{H}$ and decompose it as $\mathbf{R}_{\boldsymbol{\eta}} = \hat{\mathbf{R}}_{\boldsymbol{\eta}} + \mathbf{\Delta}_{\mathbf{R}_\eta}$ with $\hat{\mathbf{R}}_{\boldsymbol{\eta}} = \hat{\mathbf{M}}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\hat{\mathbf{M}}_{\boldsymbol{\eta}}^\mathrm{H}$ and $\mathbf{\Delta}_{\mathbf{R}_\eta} = \hat{\mathbf{M}}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{\Delta}_{\mathbf{M}_\eta}^\mathrm{H} + \mathbf{\Delta}_{\mathbf{M}_\eta}\mathbf{N}_{\boldsymbol{\eta}}\hat{\mathbf{M}}_{\boldsymbol{\eta}}^\mathrm{H} + \mathbf{\Delta}_{\mathbf{M}_\eta}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{\Delta}_{\mathbf{M}_\eta}^\mathrm{H}$. We now bound $\mathbf{\Delta}_{\mathbf{R}_\eta}$. We find

$$\begin{aligned}\|\mathbf{\Delta}_{\mathbf{R}_\eta}\|_\mathrm{F} &\le 2\|\hat{\mathbf{M}}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{\Delta}_{\mathbf{M}_\eta}^\mathrm{H}\|_\mathrm{F} + \|\mathbf{\Delta}_{\mathbf{M}_\eta}\mathbf{N}_{\boldsymbol{\eta}}\mathbf{\Delta}_{\mathbf{M}_\eta}^\mathrm{H}\|_\mathrm{F} \\ &\le 2\sigma_{\max}(\hat{\mathbf{M}}_{\boldsymbol{\eta}}\mathbf{N}_{\boldsymbol{\eta}})\epsilon_\eta + \sigma_{\max}(\mathbf{N}_{\boldsymbol{\eta}})\epsilon_\eta^2 \triangleq \epsilon'_\eta.\end{aligned}$$

Since the matrix $\mathbf{\Delta}_{\mathbf{R}_\eta}$ is Hermitian it follows that $\sum_i |\lambda_i(\mathbf{\Delta}_{\mathbf{R}_\eta})| \le \epsilon'_\eta$. Using the dual Weyl inequality we find that $\lambda_i(\mathbf{R}_{\boldsymbol{\eta}}) \le \lambda_i(\hat{\mathbf{R}}_{\boldsymbol{\eta}}) + \epsilon'_\eta$ which can be expressed as $\mathbf{R}_{\boldsymbol{\eta}} \preccurlyeq \mathbf{U}_{\hat{\mathbf{R}}_\eta}(\Lambda_{\hat{\mathbf{R}}_\eta} + \epsilon'_\eta\mathbf{I}_n)\mathbf{U}_{\hat{\mathbf{R}}_\eta}^\mathrm{H} \triangleq \widehat{\mathbf{R}}_{\boldsymbol{\eta}}$ by writing the eigenvalue decomposition of $\hat{\mathbf{R}}_{\boldsymbol{\eta}}$ as $\hat{\mathbf{R}}_{\boldsymbol{\eta}} \triangleq \mathbf{U}_{\hat{\mathbf{R}}_\eta}\Lambda_{\hat{\mathbf{R}}_\eta}\mathbf{U}_{\hat{\mathbf{R}}_\eta}^\mathrm{H}$, where $\mathbf{U}_{\hat{\mathbf{R}}_\eta}$ is an orthonormal matrix and $\Lambda_{\hat{\mathbf{R}}_\eta}$ a diagonal matrix with the eigenvalues of $\hat{\mathbf{R}}_{\boldsymbol{\eta}}$ on the diagonal. Finally we have that $\mathbf{R}_{\boldsymbol{\eta}}^{-1} \succcurlyeq \widehat{\mathbf{R}}_{\boldsymbol{\eta}}^{-1}$ and thus $\mathbf{M_x}^\mathrm{H}\mathbf{R}_{\boldsymbol{\eta}}^{-1}\mathbf{M_x} \succcurlyeq \mathbf{M_x}^\mathrm{H}\widehat{\mathbf{R}}_{\boldsymbol{\eta}}^{-1}\mathbf{M_x}$.

B. Worst-Case $\mathbf{M_x}$ for an arbitrary $\mathbf{M}_{\boldsymbol{\eta}}$

Similarly we decompose the matrix $\mathbf{R}$ as $\mathbf{R} = \hat{\mathbf{R}} + \mathbf{\Delta_R}$ with $\hat{\mathbf{R}} = \hat{\mathbf{M}}_{\mathbf{x}}^{\mathrm{H}}\mathbf{R}_{\boldsymbol{\eta}}^{-1}\hat{\mathbf{M}}_{\mathbf{x}}$ and $\mathbf{\Delta_R} = \hat{\mathbf{M}}_{\mathbf{x}}^{\mathrm{H}}\mathbf{R}_{\boldsymbol{\eta}}^{-1}\mathbf{\Delta}_{\mathbf{M_x}} + \mathbf{\Delta}_{\mathbf{M_x}}^{\mathrm{H}}\mathbf{R}_{\boldsymbol{\eta}}^{-1}\hat{\mathbf{M}}_{\mathbf{x}} + \mathbf{\Delta}_{\mathbf{M_x}}^{\mathrm{H}}\mathbf{R}_{\boldsymbol{\eta}}^{-1}\mathbf{\Delta}_{\mathbf{M_x}}$ and bound $\mathbf{\Delta_R}$ as follow

$$\|\mathbf{\Delta_R}\| \leq 2\sigma_{\max}(\hat{\mathbf{M}}_{\mathbf{x}}\mathbf{R}_{\boldsymbol{\eta}}^{-1})\epsilon_x + \sigma_{\max}(\mathbf{R}_{\boldsymbol{\eta}}^{-1})\epsilon_x^2 \triangleq \epsilon_x'.$$

It follows that $\sum_i |\lambda_i(\mathbf{\Delta_R})| \leq \epsilon_x'$. Since $\mathbf{R}$ is positive semidefinite we have $\lambda_i(\mathbf{R}) \geq (\lambda_i(\hat{\mathbf{R}}) - \epsilon_x')^+$ which can be written as $\mathbf{R} \succcurlyeq \mathbf{U}_{\hat{\mathbf{R}}}(\Lambda_{\hat{\mathbf{R}}} - \epsilon_x'\mathbf{I}_n)^+\mathbf{U}_{\hat{\mathbf{R}}}^{\mathrm{H}}$ by defining $\hat{\mathbf{R}} \triangleq \mathbf{U}_{\hat{\mathbf{R}}}\Lambda_{\hat{\mathbf{R}}}\mathbf{U}_{\hat{\mathbf{R}}}^{\mathrm{H}}$.

C. Joint Worst-Case Optimization

We now combine the two previous results. We use $\widehat{\mathbf{R}}_{\boldsymbol{\eta}}$ instead of $\mathbf{R}_{\boldsymbol{\eta}}$ and decompose $\hat{\mathbf{R}} = \hat{\mathbf{M}}_{\mathbf{x}}^{\mathrm{H}}\widehat{\mathbf{R}}_{\boldsymbol{\eta}}^{-1}\hat{\mathbf{M}}_{\mathbf{x}} = \widehat{\mathbf{U}}_{\hat{\mathbf{R}}}\widehat{\Lambda}_{\hat{\mathbf{R}}}\widehat{\mathbf{U}}_{\hat{\mathbf{R}}}^{\mathrm{H}}$. We find a lower bound for $\mathbf{R}$ as

$$\mathbf{R} \succcurlyeq \widehat{\mathbf{U}}_{\hat{\mathbf{R}}}(\widehat{\Lambda}_{\hat{\mathbf{R}}} - \epsilon_x''\mathbf{I}_n)^+\widehat{\mathbf{U}}_{\hat{\mathbf{R}}}^{\mathrm{H}} = \breve{\mathbf{R}}$$

where

$$\epsilon_x'' \triangleq 2\sigma_{\max}(\hat{\mathbf{M}}_{\mathbf{x}}\widehat{\mathbf{R}}_{\boldsymbol{\eta}}^{-1})\epsilon_x + \sigma_{\max}(\widehat{\mathbf{R}}_{\boldsymbol{\eta}}^{-1})\epsilon_x^2.$$

It remains to solve the problem (19) with $\breve{\mathbf{R}}$,i.e.,

$$\begin{array}{ll} \underset{\mathbf{A_x}}{\text{maximize}} & \log\det(\mathbf{I}_n + \mathbf{A_x}^{\mathrm{H}}\breve{\mathbf{R}}\mathbf{A_x}) \\ \text{subject to} & \mathrm{Tr}(\mathbf{A_x}\mathbf{A_x}^{\mathrm{H}}) \leq P_T, \end{array} \tag{20}$$

As shown in [12] the optimal precoding matrix $\mathbf{A_x}^*$ for this problem has the form

$$\mathbf{A_x}^* = \mathbf{V}_{\breve{\mathbf{R}}}\ \mathrm{diag}(\sqrt{p}), \tag{21}$$

where $\mathbf{V}_{\breve{\mathbf{R}}} \in \mathbb{C}^{n\times n}$ is the right singular matrix of $\breve{\mathbf{R}}$ and p is a power allocation vector calculated using waterfilling as

$$p_i = (\mu - \lambda_{\breve{R},i}^{-1})^+, \qquad 1 \leq i \leq n, \tag{22}$$

where $\lambda_{\breve{R},i}$ are the eigenvalues of $\breve{\mathbf{R}}$ and μ is the water level chosen such that $\sum_{i=1}^n p_i = P_T$.

V. Relay and Channel Information

We now look what happens when the source node receives information from relays. We assume that the RI of a relay contains its own strategy and the channel coefficients of its incoming connections. We can write the matrix $\mathbf{F}$ as

$$\mathbf{F} = \mathbf{R}_1 + \mathbf{R}_2 + \cdots + \mathbf{R}_K, \tag{23}$$

where $K-1$ is the number of relays in the network, $\mathbf{R}_k \in \mathbb{C}^{e\times e}$, for $k = 1,\ldots,K-1$, is a matrix containing the relay information of the relay k and $\mathbf{R}_K$ contains the channel coefficients of connexions going from relays to the sink. We can write each $\mathbf{R}_k$ as a block matrix similar to (6),i.e.,

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \\ \mathbf{R}_{\mathbf{H}k} & \mathbf{R}_{\mathbf{L}k} & \mathbf{0}_{m\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n} \end{bmatrix}, \tag{24}$$

and it follows that

$$\mathbf{H} = \sum_{k=1}^{K}\mathbf{R}_{\mathbf{H}k}, \qquad \mathbf{L} = \sum_{k=1}^{K}\mathbf{R}_{\mathbf{L}k}, \tag{25}$$

with $\mathbf{R}_{\mathbf{H}k} \in \mathbb{C}^{m\times n}$ and $\mathbf{R}_{\mathbf{L}k} \in \mathbb{C}^{m\times m}$ for $k = 1,\ldots,K$. The questions we want to answer are: what is the influence of having relay information at the source node on performance and how much relay information is necessary. To answer these questions we formulate the loss in uncertainty provided by the knowledge of relay information. We assume, to simplify notations and without loss of generality, that the K' first relays transmit their relay information to the source node. Further we define $\mathcal{T}_H$ as the set of nonzero coefficients of the matrix $\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{H}k}$ and $\overline{\mathcal{T}_H}$ as the complement of $\mathcal{T}_H$. We then can write

$$\hat{\mathbf{H}} = \hat{\mathbf{H}}(\mathcal{T}_H) + \hat{\mathbf{H}}(\overline{\mathcal{T}_H}), \tag{26}$$

and

$$\mathbf{\Delta_H} = \mathbf{\Delta_H}(\mathcal{T}_H) + \mathbf{\Delta_H}(\overline{\mathcal{T}_H}), \tag{27}$$

where, e.g, $\hat{\mathbf{H}}(\mathcal{T}_H)$ is a restriction of $\hat{\mathbf{H}}$ with zero coefficients outside of $\mathcal{T}_H$. Clearly we have

$$\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{H}k} = \hat{\mathbf{H}}(\mathcal{T}_H) + \mathbf{\Delta_H}(\mathcal{T}_H), \tag{28}$$

which describes the fact that the relay information has a part which is known already and an unknown part which decreases the problem uncertainty. By plugging (28) into (27) and using the fact that $\mathcal{T}_H$ and $\overline{\mathcal{T}_H}$ are disjoints, we can show that

$$\|\mathbf{\Delta_H}(\overline{\mathcal{T}_H})\|_{\mathrm{F}} \leq \sqrt{\epsilon_H^2 - \left\|\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{H}k} - \hat{\mathbf{H}}(\mathcal{T}_H)\right\|_{\mathrm{F}}^2} \triangleq \epsilon_H', \tag{29}$$

where $\mathbf{\Delta_H}(\overline{\mathcal{T}_H})$ is a matrix representing the rest uncertainty about $\mathbf{H}$. The known part of $\mathbf{H}$ is now $\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{H}k} + \hat{\mathbf{H}}(\overline{\mathcal{T}_H})$. Similarly it can be shown that

$$\|\mathbf{\Delta_L}(\overline{\mathcal{T}_L})\|_{\mathrm{F}} \leq \sqrt{\epsilon_L^2 - \left\|\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{L}k} - \hat{\mathbf{L}}(\mathcal{T}_L)\right\|_{\mathrm{F}}^2} \triangleq \epsilon_L'. \tag{30}$$

To see the influence of relay information on the uncertainty of $\mathbf{M_x}$ and $\mathbf{M}_{\boldsymbol{\eta}}$ it remains to replace $\hat{\mathbf{H}}$ by $\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{H}k} + \hat{\mathbf{H}}(\overline{\mathcal{T}_H})$, $\hat{\mathbf{L}}$ by $\sum_{k=1}^{K'}\mathbf{R}_{\mathbf{L}k} + \hat{\mathbf{L}}(\overline{\mathcal{T}_L})$, ϵ_H by $\hat{\epsilon}_H'$, and ϵ_L by $\hat{\epsilon}_L'$ in (13) an (16). This gives us new values for ϵ_x and ϵ_η, which are smaller than the former ones.

VI. Numerical Results

We simulate the network in Figure 1. The coefficients of the matrix $\mathbf{F}$ are generated at random $\sim \mathcal{N}_{\mathbb{C}}(0,1)$ to represent the fact that the strategy of the relays is fixed. The matrix $\hat{\mathbf{H}}$, $\hat{\mathbf{L}}$, $\mathbf{\Delta_H}$ and $\mathbf{\Delta_L}$ are also chosen at random such that $\|\mathbf{\Delta_H}\|_{\mathrm{F}} \leq \gamma\|\hat{\mathbf{H}}\|_{\mathrm{F}}$ and $\|\mathbf{\Delta_L}\|_{\mathrm{F}} \leq \gamma\|\hat{\mathbf{L}}\|_{\mathrm{F}}$, where we call γ the effective uncertainty. The source node assumes $\epsilon_H = 0.5\|\hat{\mathbf{H}}\|_{\mathrm{F}}$ and $\epsilon_L = 0.5\|\hat{\mathbf{L}}\|_{\mathrm{F}}$, where 0.5 is the expected uncertainty. In Figure 2 we plot the achieved mutual information for the case where 1) the source node has complete relay information (RI), 2) the source node has no RI and optimize $\mathbf{A_x}$ based upon $\hat{\mathbf{H}}$ and $\hat{\mathbf{L}}$ (we call it the nominal case), 3) the source node has no RI and performs a worst-case optimization as described in Section IV, 4) the source node knows the RI of

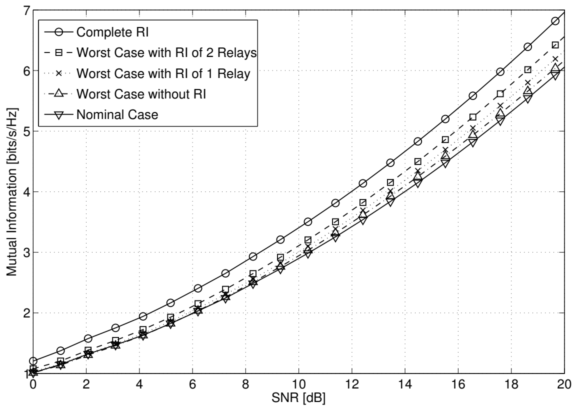

Fig. 2. Mutual information for different available RI ($\gamma = 0.5$)

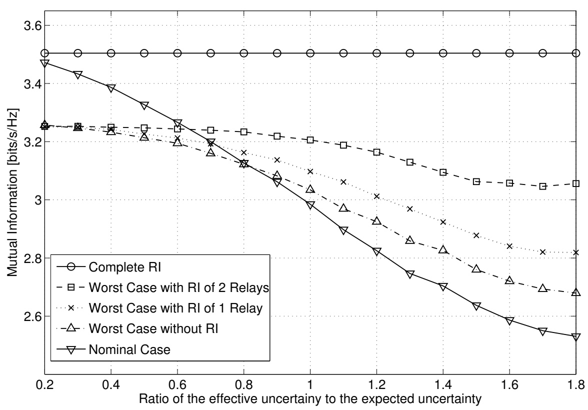

Fig. 3. Mutual information for different available RI (SNR = 10 dB)

one single relay and 5) the source node knows the RI of both relays and perform a worst-case optimization as described in Section IV. In all cases we choose $\gamma = 0.5$. As expected, the complete RI case is always best, the nominal case is the worst and each RI information gathered at the source node improves the mutual information. In Figure 3, we plot again the achieved mutual information but this time with respect to $\gamma/0.5$, for a SNR of 10 dB. In other words we want to observe the behavior of the network when the effective uncertainty is lower or higher than the expected uncertainty. The mutual information stays constant for the case of complete RI since there is no uncertainty. For a low uncertainty (< 0.6) the nominal case is better than the worst-cases (with or without RI) since the worst-case optimization overestimates the risk and designs $\mathbf{A_x}$ for an unrealistic network. For a higher uncertainty (> 0.6), the worst-case design starts becoming interesting if one has RI available and from 0.8 on, the worst-case design is always better. Finally, we plot in Figure 4 the frequency of occurence of the ratio from the achieved mutual information to the mutual information with complete RI. We want to see how often the performance is far from the maximum. We see that the case that the performance is less than 60% of the maximum occurs much more often in the nominal case. A performance between 60% and 90% from the maximum occurs more often using a worst-case optimization and the maximum mutual information occurs as often for all cases.

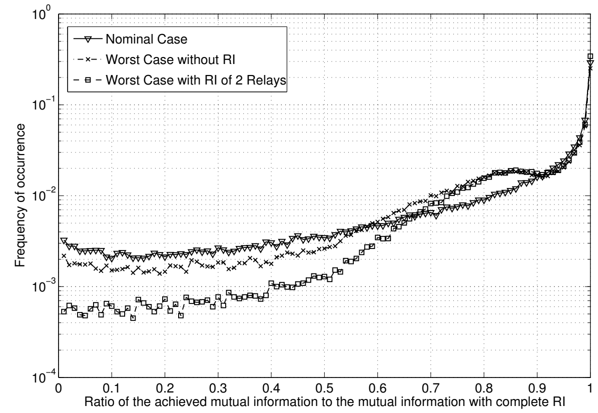

Fig. 4. Frequency of occurrence from ratio of achieved mutual information to mutual information with complete RI ($\gamma = 0.5$, SNR = 10 dB)

VII. Conclusion

In this work, we have presented a new model for studying the influence of relay information on large-scale multihop networks when the strategy of the relays is fixed. We have derived a solution to the worst-case optimization problem consisting in maximizing the mutual information with partial knowledge of relay information at the source. A future research direction is to develop algorithms to choose which relay should send RI to the source given a power budget.

References

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

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

[3] 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, 2007.

[4] Y. Zhao, R. Adve, and T. Lim, "Beamforming with limited feedback in amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5145–5149, 2008.

[5] S. Yeh and O. Lévêque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE International Symposium on Inf. Theory (ISIT 2007)*, Nice, France, Jun. 2007, pp. 1436–1440.

[6] X. Deng and A. Haimovich, "Power allocation for cooperative relaying in wireless networks," *IEEE Communications Letters*, vol. 9, no. 11, pp. 994–996, 2005.

[7] Z. Yi and I. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 447–458, 2007.

[8] B. Maham and A. Hjorungnes, "Opportunistic relaying through amplify-and-forward distributed space-time codes with partial statistical CSI at relays," in *Proc. 47th Ann. Allerton Conf. Communication, Control and Computing*, Monticello, IL, USA, 2009, pp. 1004–1008.

[9] S. Corroy and R. Mathar, "Strategies for source/sink pairs in amplify-and-forward Gaussian relay networks," in *Proc. 48th Ann. Allerton Conf. Communication, Control and Computing*, Monticello, IL, USA, Sep.-Oct. 2010.

[10] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 3, pp. 782–795, Oct. 2003.

[11] D. P. Palomar, "A unified framework for communications through MIMO channels," *Ph.D. dissertation*, May 2003.

[12] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.