Interference Neutralization for Separated Multiuser Uplink-Downlink with Distributed Relays

Jinyuan Chen, Arun Singh, Petros Elia and Raymond Knopp
Mobile Communications Department
EURECOM, Sophia Antipolis, France
{chenji, singhak, elia, knopp}@eurecom.fr

Abstract— **We consider the role of independent single antenna relays in allowing for interference cancellation and for multiplexing of uplink and downlink communications, in the presence of multiple interfering users. For a specific practical scenario of interest, we construct interference neutralization schemes that are simple, distributed and which encode over both space and time, to achieve performance improvements over basic TDMA schemes. These improvements are quantified in terms of the achievable degrees of freedom (DOF), where asymptotic analysis reveals that in specific cases, these simple techniques achieve optimal DOF performance.**

I. Introduction

In the setting of interference limited multiuser communications, different interference management methods have been developed that range from the hard to implement but powerful interference alignment techniques [1], to simpler precoding and network coding methods [2]. The latter include different interference neutralization (IN) methods [3], [4], which seek to properly combine signals arriving from various paths in such a way that the interfering signals are canceled while the desired signals are preserved.

The task of these methods is made harder when the interference is due to other users, i.e., when it is not self-interference, when some of the precoding nodes have single antennas thus forcing precoding to be performed in a distributed manner, and when there is a half-duplex constraint especially in two-way communications where a single node might both transmit and receive information (cf. [5]–[8]). We here explore the above and provide linear distributed solutions for a specific half-duplex constrained, two-way, interference limited multiuser scenario of practical importance. In this setting, we propose partially distributed IN schemes that properly combine signals across both time and space, and then proceed to analyze their performance by exploring their signal attenuation and noise accentuation effects, to show that in some settings, the proposed schemes are DOF optimal, whereas in other settings, the schemes provide modest improvements.

A. Notation

Throughout this paper, $(\bullet)^{-1}$, $(\bullet)^T$, $(\bullet)^\dagger$, $||\bullet||_F$ and $\text{tr}(\bullet)$ respectively denote the inverse, transpose, conjugate transpose, Frobenius norm and trace of matrices. $(\bullet)^*$ denotes the complex conjugate, $||\bullet||$ denotes Euclidean norm, $[\bullet]_j$ denotes j-th row of the matrix or column vector in the argument, and $[\bullet]_{i,j}$ denotes the matrix element in the ith row and jth column. $\text{diag}(\bullet)$ denotes a diagonal matrix. $\lambda_{max}(\bullet)$ and $\lambda_{min}(\bullet)$ respectively denote the maximum and minimum eigenvalues of the matrix. $|\bullet|$ denotes either the magnitude of a scalar or the cardinality of a set. $\Re$ and $\Re_+$ respectively denote the sets of real and positive real numbers. We use $\doteq$ to denote exponential equality, i.e., we write $f(\rho) \doteq \rho^d$ to denote $\lim_{\rho\to\infty} \frac{\log f(\rho)}{\log \rho} = d$ and $\dot{\leq}$, $\dot{\geq}$ are similarly defined.

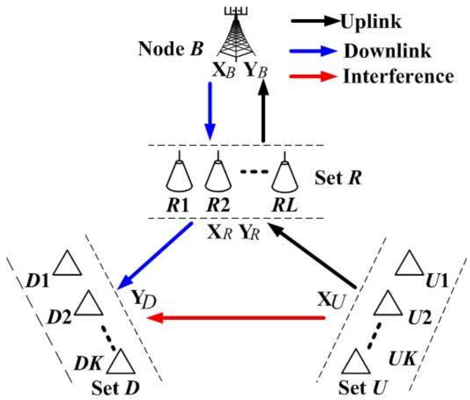

Fig. 1. K-user Uplink-Downlink System Model

II. Signal and Channel Model

We consider a setting where a node B with K antennas wants to receive from a set U of K single-antenna users, and transmit to a set D of K single antenna users, with the help of a set R of L single-antenna relays, in the presence of the half-duplex constraint, and in the presence of unlimited channel state information at all transmitters and receivers (CSITR). In the context of cellular systems, B may play the role of a base station, and U and D the role of the sets of uplink and downlink users respectively. Let $\mathbf{y}_B^{(t)}, \mathbf{y}_R^{(t)}, \mathbf{y}_D^{(t)}, \mathbf{z}_B^{(t)}, \mathbf{z}_R^{(t)}, \mathbf{z}_D^{(t)}$ respectively denote the received signal and noise vectors at B, and at all in R and in D. In the following we will ignore the time index if no ambiguity is caused. In the scale of interest we consider a uniform power constraint where the total power transmitted by a node does not exceed ρ which also, in the scale of interest, takes the role of the signal-to-noise ratio (SNR). Furthermore we let $h_{i,j}$ denote the channel fading coefficient between the ith user in U and the receiver of the jth user in D, let $\mathbf{h}_{U,j}$ denote the vector of fading coefficients between the entire set U and the receiver of the jth user in D, and let $\mathbf{H}_{U,D}, \mathbf{H}_{U,R}, \mathbf{H}_{R,D}, \mathbf{H}_{R,B}, \mathbf{H}_{B,R}$ denote the U-to-D, U-to-R, R-to-D, R-to-B and B-to-R collective channel fading matrices. Fading and additive noise coefficients are considered to be i.i.d. complex Gaussian $\mathbb{CN}(0,1)$, and the fading is assumed to remain constant during the coherence period. Here $\mathbf{H}_{B,D} = \mathbf{H}_{U,B} = \mathbf{0}$.

We are interested in analyzing the degrees of freedom, so a rate of $\mathcal{R}$ bits per channel use (bpcu), corresponds to $d = \lim_{\rho\to\infty} \frac{\mathcal{R}}{\log\rho}$ degrees of freedom.

A. Summary

Lemmas 1 and 2 provide basic results that are used in the DOF analysis, and Lemma 3 describes the DOF outer bound. Section III-A describes the IN schemes for the case of $|U| = |D| = K = L$, and Section III-B describes the IN schemes and their performance for the case of $L > |U| = |D| = K$. The same sections provide part of the proof of the performance of these schemes, which is completed in Section III-C which analyzes the effects of precoding on the signal and noise power. The Appendix holds some of the proofs.

III. Proposed IN schemes and DOF analysis

Before proceeding to establish the DOF limits for the system shown in Figure 1, we provide the following necessary lemmas whose proofs are found in the Appendix.

Lemma 1: Let $\mathbf{\Phi} = \prod_{i=1}^{k} \mathbf{H}_i^{\tau_i}$, where $\tau_i \in \{-1, 1\}$ and where $\mathbf{H}_i$ are $N \times N$ random matrices with i.i.d. $\mathbb{CN}(0,1)$ entries. Then

$$P(||\mathbf{\Phi}^{-1}||_F^2 \geq \rho^{\epsilon}) \dot{\leq} \rho^{-\epsilon/d},$$

for $d = |\{j|\tau_j = 1\}|$, for any $\epsilon > 0$.

Lemma 2: Let $\mathbf{E}$ be a rank-M, $M \times N$ matrix with entries that are either zero or (positive exponent) polynomial functions of independent $\mathbb{CN}(0,1)$ random variables. For any unit-norm vector $\mathbf{v}$ in the (right) null space of $\mathbf{E}$, and any $\epsilon > 0$, there exists a finite constant $d' > 0$ such that

$$P(|v_i|^2 \leq \rho^{-\epsilon}) \dot{\leq} \rho^{-\epsilon/d'},$$

for any element v_i of $\mathbf{v}$.

We henceforth adopt a partial rate uniformity assumption where all users in set U have the same rate $\mathcal{R}_U$, and all users in D have rate $\mathcal{R}_D$, but where $\mathcal{R}_U$ and $\mathcal{R}_D$ are not necessarily equal. We also focus on the case where $L \geq K$, and where all the relays transmit at the same time. When ergodicity is required, transmission is typically considered to take place over M blocks equaling M coherence times, each spanning n channel uses.

Under the above assumptions, we have the following straightforward upper bound on the performance.

Lemma 3: In the setting of interest, each user in U and in D can achieve at most $\frac{1}{2}$ DOF, i.e., $d_U \leq 1/2, \quad d_D \leq 1/2$.

The proof is relegated to the Appendix. We note that the above holds irrespective of the number of phases, the duration of each phase, and irrespective of the power allocation methods.

We proceed to describe the IN schemes for different settings. What is common in all settings, is that each block is divided into two phases of duration Δ_1 and $1 - \Delta_1$, where in the first phase the relays listen while in the second phase they transmit. We furthermore here consider that each phase has duration equal to one channel use.

We begin with the case of having K relays, and propose four schemes, with the first two schemes $\mathcal{X}_1, \mathcal{X}_2$ being optimal for the case $K = 1, 2$ respectively, and where the other two schemes $\mathcal{X}_3, \mathcal{X}_4$ are for the $K \geq 3$ case, and which are not proven to be optimal. A summary of the schemes' performance and basic characteristics is provided in Table I.

We will afterwards address the case where $L > K$, and design schemes $\mathcal{X}_3^{'}, \mathcal{X}_4^{'}$ that utilize the increase in the number of relays to achieve higher, and in some cases optimal DOF performance. A summary of the schemes' performance is provided in Table II.

A. IN schemes for the case $|U| = |D| = K = L$

1) Scheme $\mathcal{X}_1$ for $K = 1$: In the first phase, $U1$ transmits a_1 intended for B, while B transmits b_1 intended for $D1$. The relay and $D1$ then respectively receive

$$\begin{aligned} y_{R1}^{(1)} &= h_{B,R1}^{(1)} b_1 + h_{U1,R1}^{(1)} a_1 + z_{R1}^{(1)}, & (1)\\ y_{D1}^{(1)} &= h_{U1,D1}^{(1)} a_1 + z_{D1}^{(1)}. & (2) \end{aligned}$$

In the second phase $R1$ scales $y_{R1}^{(1)}$ and forwards $v_{R1} y_{R1}^{(1)}$, where v_{R1} will be designed later on. Hence $D1$ and B respectively receive

$$\begin{aligned} y_{D1}^{(2)} =& h_{R1,D1}^{(2)} v_{R1} y_{R1}^{(1)} + z_{D1}^{(2)} \\ =& h_{R1,D1}^{(2)} v_{R1} h_{U1,R1}^{(1)} a_1 + h_{R1,D1}^{(2)} v_{R1} h_{B,R1}^{(1)} b_1 \\ &+ h_{R1,D1}^{(2)} v_{R1} z_{R1}^{(1)} + z_{D1}^{(2)}, \qquad (3)\\ y_B^{(2)} =& h_{R1,B}^{(2)} v_{R1} y_{R1}^{(1)} + z_B^{(2)} \\ =& h_{R1,B}^{(2)} v_{R1} h_{U1,R1}^{(1)} a_1 + h_{R1,B}^{(2)} v_{R1} h_{B,R1}^{(1)} b_1 \\ &+ h_{R1,B}^{(2)} v_{R1} z_{R1}^{(1)} + z_B^{(2)}. \qquad (4) \end{aligned}$$

Towards neutralizing the received interference, $D1$ adds up the signal $y_{D1}^{(1)}$ stored during the first phase and the received signal $y_{D1}^{(2)}$, to get a new observation

$$\begin{aligned} \tilde{y}_{D1} &= y_{D1}^{(1)} + y_{D1}^{(2)} \\ &= (h_{U1,D1}^{(1)} + h_{R1,D1}^{(2)} v_{R1} h_{U1,R1}^{(1)}) a_1 + h_{R1,D1}^{(2)} v_{R1} h_{B,R1}^{(1)} b_1 \\ &\quad + h_{R1,D1}^{(2)} v_{R1} z_{R1}^{(1)} + z_{D1}^{(2)} + z_{D1}^{(1)}. \qquad (5) \end{aligned}$$

Setting

$$v_{R1} = -h_{U1,D1}^{(1)} / (h_{R1,D1}^{(2)} h_{U1,R1}^{(1)}), \qquad (6)$$

we get an interference-free signal

$$\tilde{y}_{D1} = h_{R1,D1}^{(2)} v_{R1} h_{B,R1}^{(1)} b_1 + h_{R1,D1}^{(2)} v_{R1} z_{R1}^{(1)} + z_{D1}^{(2)} + z_{D1}^{(1)}. \qquad (7)$$

Consequently $D1$ can decode one independent symbol in two channel uses, albeit with an attenuated signal $h_{R1,D1}^{(2)} v_{R1} h_{B,R1}^{(1)} b_1$ and accentuated noise $h_{R1,D1}^{(2)} v_{R1} z_{R1}^{(1)} +$

$z_{D1}^{(2)} + z_{D1}^{(1)}$. Section III-C will show that the noise accentuation and signal attenuation from precoding do not affect the DOF performance, and we can thus conclude that the scheme achieves the optimal $d_D = 1/2$. Similarly for the uplink, where the interference is self-interference (cf. (4)), we can also conclude that the scheme achieves the optimal $d_U = 1/2$.

We now proceed to describe the schemes when $K \geq 2$. We will henceforth denote as b_i the information symbol at B intended for D_i, and we will denote as $\mathbf{b}$ the vector of all such symbols. B will then be transmitting $\mathbf{V}_B\mathbf{b}$ where $\mathbf{V}_B$ is the precoder at B which will be designed for each of the cases later on. Symbols a_j will carry the information from Uj intended for B, and $\mathbf{a}$ will denote the vector of all such symbols. Furthermore $\{v_{Uj}\}, \{v_{R\ell}\}, \{v_{Di}\}$ will denote the sets of distributed precoding/weighting coefficients, respectively used at the nodes in U, the relays, and the nodes in D, and which will be designed for the different settings later on. The dimensionality of the corresponding vectors and matrices in the different settings, will be clarified when necessary.

2) Scheme $\mathcal{X}_2$ for $K = 2$: In the first phase, each U_j transmits a_j, $j = 1, 2$ and B transmits $\mathbf{V}_B\mathbf{b}$, hence the relays and each Di respectively receive

$$\mathbf{y}_R^{(1)} = \mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} + \mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{z}_R^{(1)}, \quad (8)$$

$$y_{Di}^{(1)} = \sum_{j=1}^{2} h_{Uj,Di}^{(1)} a_j + z_{Di}^{(1)}, \quad i \in \{1, 2\}. \quad (9)$$

The second phase takes place during a different fading realization. During this second phase, all in U remain in the transmit mode, and specifically each Uj transmits $v_{Uj}a_j$, whereas each relay $R\ell$ transmits $v_{R\ell}y_{R\ell}^{(1)}$, $j, \ell \in \{1, 2\}$. For $\mathbf{V}_R = \mathrm{diag}(v_{R1}, v_{R2})$, then we design

$$\mathbf{V}_B = (\mathbf{H}_{R,D}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)})^{-1}\mathrm{diag}(\beta_1, \beta_2),$$

where β_1, β_2 are scalars that ensure the power constraint at B. Consequently Di and B respectively receive

$$y_{Di}^{(2)} = \sum_{j=1}^{2}(h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j + \beta_i b_i + (\sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)}, \quad i \in \{1, 2\}, \quad (10)$$

$$\mathbf{y}_B^{(2)} = \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} + \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{z}_R^{(1)} + \mathbf{z}_B^{(2)}. \quad (11)$$

Towards neutralizing the interference from $U1$ and $U2$, each Di adds up a specifically scaled version of $y_{Di}^{(1)}$, to the received signal $y_{Di}^{(2)}$, to get

$$\begin{aligned}\tilde{y}_{Di} &= v_{Di}y_{Di}^{(1)} + y_{Di}^{(2)} \\ &= \sum_{j=1}^{2}(v_{Di}h_{Uj,Di}^{(1)} + h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j \\ &\quad + \beta_i b_i + (\sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di}z_{Di}^{(1)}. \quad (12)\end{aligned}$$

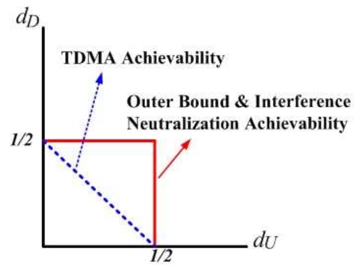

Fig. 2. DOF regions for $K = 1, 2$.

We now seek to design a unit-norm distributed precoding vector $\mathbf{v}_2 = [v_{R1} \quad v_{R2} \quad v_{U1} \quad v_{U2} \quad v_{D1} \quad v_{D2}]^T$, that will remove the uplink interference at D. Looking at (12), $\mathbf{v}_2$ is designed such that for any $j, i \in \{1, 2\}$ then

$$v_{Di}h_{Uj,Di}^{(1)} + h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)} = 0.$$

Specifically, considering the 4×6 matrix $\mathbf{G}_2$ having the mth-row nth-column element $g_{m,n}$ equal to

$$\begin{aligned} g_{2(i-1)+j,\ell} &\triangleq h_{R\ell,Di}^{(2)}h_{Uj,R\ell}^{(1)}, \\ g_{2(i-1)+j,2+j} &\triangleq h_{Uj,Di}^{(2)}, \\ g_{2(i-1)+j,4+i} &\triangleq h_{Uj,Di}^{(1)}, \quad \forall i, j, \ell \in \{1, 2\}, \quad (13)\end{aligned}$$

else $g_{m,n} \triangleq 0$, we see that $\mathbf{v}_2$ must satisfy

$$\mathbf{G}_2\mathbf{v}_2 = \begin{pmatrix} g_{1,1} & g_{1,2} & g_{1,3} & 0 & g_{1,5} & 0 \\ g_{2,1} & g_{2,2} & 0 & g_{2,4} & g_{2,5} & 0 \\ g_{3,1} & g_{3,2} & g_{3,3} & 0 & 0 & g_{3,6} \\ g_{4,1} & g_{4,2} & 0 & g_{4,4} & 0 & g_{4,6} \end{pmatrix}\mathbf{v}_2 = 0, \quad (14)$$

and consequently the unit norm $\mathbf{v}_2$ is designed simply by picking an arbitrary unit-norm vector from the non-empty null-space of $\mathbf{G}_2$. We here note from (13) that as the nonzero elements of $\mathbf{G}_2$ are independent polynomials in i.i.d. complex Gaussian random variables, the rank of $\mathbf{G}_2$ is almost surely 4, which explains the need for encoding over the second fading realization (and the reason that uplink stays in transmit mode during the second phase); i.e., so that $\mathbf{G}_2$ has a non-empty null-space, inside of which we arbitrarily draw the unit norm $\mathbf{v}_2$.

Consequently in two channel uses each user in D can decode one independent symbol without interference. In conjunction with the signal attenuation and noise accentuation analysis in Section III-C which again proves that the signal and noise effects of the proposed precoding can be ignored in terms of DOF, we can conclude that the scheme achieves the optimal $d_D = 1/2$. Similarly for the uplink where all interference is self-interference (cf. (11)), we can again conclude that the scheme achieves the optimal $d_U = 1/2$.

3) Scheme $\mathcal{X}_3$ for $K \geq 3$: The protocol builds on $\mathcal{X}_2$ and asks the users in D to split in pairs[1] and, two at a

[1]Note that the same result holds when K is odd.

time, take turns in receiving information. Thus, without loss of generality, we reformulate the problem and rewrite $D = \{D1, D2\}$, in which case the downlink information symbols are $\mathbf{b} = [b_1 \quad b_2]^T$, and the precoded downlink transmitted vector at B during the first phase is $\mathbf{V}_B\mathbf{b}$, for some $K \times 2$ precoder $\mathbf{V}_B$ to be designed. During the same first phase, U_j transmits a_j $j \in \{1, \cdots, K\}$. For $\mathbf{a} = [a_1 \cdots a_K]^T$, the relays and the different Di then respectively receive

$$\mathbf{y}_R^{(1)} = \mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} + \mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{z}_R^{(1)}, \quad (15)$$

$$y_{Di}^{(1)} = \sum_{j=1}^{K} h_{Uj,Di}^{(1)} a_j + z_{Di}^{(1)}, \quad i \in \{1, 2\}. \quad (16)$$

The second phase takes place during a different fading realization, during which, each Uj transmits $v_{Uj}a_j$ and the ℓth relay transmits $v_{R\ell}y_{R\ell}^{(1)}$, $j, \ell \in \{1, \cdots, K\}$. For $\mathbf{V}_R = \text{diag}(v_{R1}, \cdots, v_{RK})$, we design the precoder at B to be[2]

$$\mathbf{V}_B = (\mathbf{H}_{R,D}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)})^{-1}\text{diag}(\beta_1, \beta_2),$$

where $\mathbf{H}_{R,D}^{(2)}$ is a $2 \times K$ channel coefficient matrix between R and $\{D1, D2\}$, and where scalars β_1 and β_2 ensure that the power constraint at B is satisfied. Consequently each Di and B respectively receive

$$\begin{aligned} y_{Di}^{(2)} =& \sum_{j=1}^{K}(h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j + \beta_i b_i \\ &+ (\sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)}, \quad i \in \{1, 2\}, \quad (17) \\ \mathbf{y}_B^{(2)} =& \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} \\ &+ \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{z}_R^{(1)} + \mathbf{z}_B^{(2)}. \quad (18) \end{aligned}$$

Using distributed weighting coefficients v_{Di} to be derived later on, each downlink user Di then adds up the properly weighted received signals $v_{Di}y_{Di}^{(1)}$ stored during the first phase and the received signal $y_{Di}^{(2)}$ of the second phase, to get

$$\begin{aligned} \tilde{y}_{Di} &= v_{Di}y_{Di}^{(1)} + y_{Di}^{(2)} \\ &= \sum_{j=1}^{K}(v_{Di}h_{Uj,Di}^{(1)} + h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j \\ &\quad + \beta_i b_i + (\sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di}z_{Di}^{(1)}. \quad (19) \end{aligned}$$

Towards deriving the above mentioned weighting coefficients, we construct

$$\mathbf{v}_3 = [v_{R1} \cdots v_{RK} \quad v_{U1} \cdots v_{UK} \quad v_{D1} \quad v_{D2}]^T,$$

by picking an arbitrary unit norm vector from the non-empty null space of the $2K \times (2K + 2)$ matrix $\mathbf{G}_3$ with elements $g_{m,n}$

$$\begin{aligned} g_{K(i-1)+j,\ell} &\triangleq h_{R\ell,Di}^{(2)}h_{Uj,R\ell}^{(1)}, \\ g_{K(i-1)+j,K+j} &\triangleq h_{Uj,Di}^{(2)}, \\ g_{K(i-1)+j,2K+i} &\triangleq h_{Uj,Di}^{(1)}, \quad (20) \end{aligned}$$

$\forall j, \ell \in \{1, \cdots, K\}, \forall i \in \{1, 2\}$, else $g_{m,n} \triangleq 0$. The same arguments as in the case of $\mathcal{X}_2$ guarantee that all the uplink interference is removed at D, and also reveal the role of encoding over two different fading realization, i.e., so that the null-space of $\mathbf{G}_3$ is not empty. Hence both $D1$ and $D2$ can decode one independent symbol without interference in two channel uses. Signal and noise analysis in Section III-C, and the fact that there are K downlink users, allows us to conclude that each downlink user can achieve $d_D = 1/K$ DOF. Furthermore, for the uplink, the analysis in Section III-C and the fact that interference is self-interference (cf. (18)) allows us to conclude that each uplink user can achieve the optimal $d_U = 1/2$ DOF.

4) Scheme $\mathcal{X}_4$ for $K \geq 3$: Here the users in U take turns, two at a time, in transmitting information, and thus without loss of generality we reformulate the problem and rewrite $U = \{U1, U2\}$. In the first phase, the jth user in U transmits a_j intended for B, while B transmits $\mathbf{V}_B\mathbf{b}$ for $\mathbf{b} = [b_1 \cdots b_K]^T$. For $\mathbf{a} = [a_1 \quad a_2]^T$, the relay nodes and the Di respectively receive

$$\mathbf{y}_R^{(1)} = \mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} + \mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{z}_R^{(1)}, \quad (21)$$

$$y_{Di}^{(1)} = \sum_{j=1}^{K} h_{Uj,Di}^{(1)} a_j + z_{Di}^{(1)}, \quad i \in \{1, \cdots, K\}, \quad (22)$$

where $\mathbf{H}_{U,R}^{(1)}$ is of dimension $K \times 2$.

The second phase takes place during a different fading realization. During the second phase Uj transmits $v_{Uj}a_j$, and the ℓth relay transmits $v_{R\ell}y_{R\ell}^{(1)}$, $j, \ell \in \{1, \cdots, K\}$. For $\mathbf{V}_R = \text{diag}(v_{R1}, \cdots, v_{RK})$, then

$$\mathbf{V}_B = (\mathbf{H}_{R,D}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)})^{-1}\text{diag}(\beta_1, \cdots, \beta_K),$$

and consequently each Di and B respectively receive

$$\begin{aligned} y_{Di}^{(2)} =& \sum_{j=1}^{2}(h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j + \beta_i b_i \\ &+ (\sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)}, \quad i \in \{1, \cdots, K\} \quad (23) \\ \mathbf{y}_B^{(2)} =& \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} \\ &+ \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{z}_R^{(1)} + \mathbf{z}_B^{(2)}. \quad (24) \end{aligned}$$

Towards neutralizing the interference, each Di linearly combines the stored received signals of the first and second phase,

[2] At this point the problem of non-causality becomes evident, which can be handled either by channel prediction, or as we will show in the next section, by adding extra relays.

to get

$$\tilde{y}_{Di} = v_{Di}y_{Di}^{(1)} + y_{Di}^{(2)}$$
$$= \sum_{j=1}^{2}(v_{Di}h_{Uj,Di}^{(1)} + h_{Uj,Di}^{(2)}v_{Uj} + \sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j$$
$$+\beta_i b_i + (\sum_{\ell=1}^{K} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di}z_{Di}^{(1)}. \quad (25)$$

Similar to before, we construct the distributed precoding vector $\mathbf{v}_4 = [v_{R1}\cdots v_{RK} \quad v_{U1} \quad v_{U2} \quad v_{D1}\cdots v_{DK}]^T$, by picking an arbitrary unit norm vector from the non-empty null space of the $2K \times (2K+2)$ rank-$2K$ matrix $\mathbf{G}_4$ with elements $g_{m,n}$: $g_{2(i-1)+j,\ell} \triangleq h_{R\ell,Di}^{(2)}h_{Uj,R\ell}^{(1)}$, $g_{2(i-1)+j,K+j} \triangleq h_{Uj,Di}^{(2)}$, $g_{2(i-1)+j,K+2+i} \triangleq h_{Uj,Di}^{(1)}, \forall i,\ell \in \{1,\cdots,K\}, \forall j \in \{1,2\}$, else $g_{m,n} \triangleq 0$ Similar arguments as before guarantee that all the uplink interference is removed at D, and allow us to conclude that the users in D each decode, in two channel uses, one independent symbol without interference, which in conjunction with the signal and noise power analysis in Section III-C, tells us that $\mathcal{X}_4$ allows for each downlink user to achieve the optimal $d_D = 1/2$ DOF, while the uplink users can achieve $d_U = 1/K$ DOF.

TABLE I

SCHEMES FOR THE CASE OF $|U| = |D| = K = L$. B TRANSMITS DURING THE FIRST PHASE WHILE THE RELAYS TRANSMIT DURING THE SECOND PHASE. THE SECOND COLUMN DESCRIBES HOW MANY FADING REALIZATIONS ENCODING TAKES PLACE OVER, AND THE THIRD COLUMN DESCRIBES THE PHASES OVER WHICH THE UPLINK TRANSMITS.

	fading	U tx	K	achievable DOF
$\mathcal{X}_1$	1	1	$K=1$	$\underbrace{(d_U{=}\frac{1}{2}, d_D{=}\frac{1}{2})}_{\text{optimal}}$
$\mathcal{X}_2$	2	1 & 2	$K=2$	$\underbrace{(d_U{=}\frac{1}{2}, d_D{=}\frac{1}{2})}_{\text{optimal}}$
$\mathcal{X}_3$	2	1 & 2	$K\geq 3$	$(\underbrace{d_U{=}\frac{1}{2}}_{\text{optimal}}, d_D{=}\frac{1}{K})$
$\mathcal{X}_4$	2	1 & 2	$K\geq 3$	$(d_U{=}\frac{1}{K}, \underbrace{d_D{=}\frac{1}{2}}_{\text{optimal}})$
TDMA	1	–	$\forall K$	$(d_U{=}\frac{1}{4}, d_D{=}\frac{1}{4})$

B. IN schemes for the case $L > |U| = |D| = K$

In what follows we consider the case $L > |U| = |D| = K$. Naturally for $K = 1,2$, the previously described schemes will achieve, as before, the optimal performance $(1/2, 1/2)$. It is worth noting that, as it turns out, the existence of extra relays allows for $\mathcal{X}_2$ to achieve the optimal performance in just one fading realization, rather than two as it was before the case. This takes care of all the causality issues. In what follows we provide two new variations $\mathcal{X}_3^{'}, \mathcal{X}_4^{'}$, and analyze their performance for $K \geq 2$. It is worth noting that as the number of relays increases, the two schemes will achieve the DOF outer bound.

1) Scheme $\mathcal{X}_3^{'}$ for $K \geq 2$ and $(J+1)K - J > L > J(K-1)$, $\forall J \in \{2,\cdots,K\}$: The protocol asks the users in D to take turns, J at a time, in receiving information. Thus, without loss of generality, we reformulate the problem and rewrite $D = \{D1,\cdots,DJ\}$, in which case the downlink information symbols are $\mathbf{b} = [b_1\cdots b_J]^T$, and B transmits $\mathbf{V}_B\mathbf{b}$ where $\mathbf{V}_B$ is of dimension $K \times J$. Furthermore each U_j transmits a_j $j \in \{1,\cdots,K\}$, and consequently the relays and the different Di respectively receive

$$\mathbf{y}_R^{(1)} = \mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b} + \mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{z}_R^{(1)}, \quad (26)$$
$$y_{Di}^{(1)} = \sum_{j=1}^{K} h_{Uj,Di}^{(1)}a_j + z_{Di}^{(1)}, \quad i \in \{1,\cdots,J\}. \quad (27)$$

During the second phase the ℓth relay transmits $v_{R\ell}y_{R\ell}^{(1)}$, $\ell \in \{1,\cdots,L\}$, while this time around all the users in U are silent. For $\mathbf{V}_R = \text{diag}(v_{R1},\cdots,v_{RL})$, then $\mathbf{V}_B = (\mathbf{H}_{R,D}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)})^{-1}\text{diag}(\beta_1,\cdots,\beta_J)$, where $\mathbf{H}_{R,D}^{(2)}$ is a $J \times L$ channel matrix between R and $\{D1,\cdots,DJ\}$. Di and B then respectively receive

$$y_{Di}^{(2)} = \sum_{j=1}^{K}(\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j + \beta_i b_i$$
$$+(\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)}, \quad i \in \{1,\cdots,J\} \quad (28)$$
$$\mathbf{y}_B^{(2)} = \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{U,R}^{(1)}\mathbf{a} + \mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{H}_{B,R}^{(1)}\mathbf{V}_B\mathbf{b}$$
$$+\mathbf{H}_{R,B}^{(2)}\mathbf{V}_R\mathbf{z}_R^{(1)} + \mathbf{z}_B^{(2)}. \quad (29)$$

Each Di properly weighs past stored and current signals to get

$$\tilde{y}_{Di} = v_{Di}y_{Di}^{(1)} + y_{Di}^{(2)}$$
$$= \sum_{j=1}^{K}(v_{Di}h_{Uj,Di}^{(1)} + \sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)}v_{R\ell}h_{Uj,R\ell}^{(1)})a_j$$
$$+\beta_i b_i + (\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)}v_{R\ell}z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di}z_{Di}^{(1)} \quad (30)$$

and we are left having to design the distributed precoding unit-norm vector $\mathbf{v}_5 = [v_{R1}\cdots v_{RL} \quad v_{D1}\cdots v_{DJ}]^T$, simply by picking it from the non-empty null space of the $JK \times (J+L)$ full row rank matrix $\mathbf{G}_5$ with elements $g_{m,n}$ $g_{K(i-1)+j,\ell} \triangleq h_{R\ell,Di}^{(2)}h_{Uj,R\ell}^{(1)}$, $g_{K(i-1)+j,L+i} \triangleq h_{Uj,Di}^{(1)}$, $\forall j \in \{1,\cdots,K\}, \forall \ell \in \{1,\cdots,L\}, \forall i \in \{1,\cdots,J\}$, else $g_{m,n} \triangleq 0$. As before we can conclude that $D1,\cdots,DJ$ can each decode, in two channel uses, one independent symbol without interference, and as before to conclude that $d_D = J/2K$ DOF. For the uplink where interference is self-interference (cf. (29)), we get the optimal $d_U = 1/2$ DOF. We now see that for $J = K$ (and naturally for $J > K$), the scheme achieves the optimal performance $(1/2, 1/2)$.

2) Scheme $\mathcal{X}_4'$ for $K \geq 2$ and $JK > L > K(J-1)$, $\forall J \in \{2, \cdots, K\}$: Here it's the uplink users that take turns, J at a time, in transmitting information, and as before we rewrite $U = \{U1, \cdots, UJ\}$. In the first phase, the jth user in U transmits a_j, while B transmits $\mathbf{V}_B \mathbf{b} = [b_1 \cdots b_K]^T$. For $\mathbf{a} = [a_1 \cdots a_J]^T$, we have

$$\mathbf{y}_R^{(1)} = \mathbf{H}_{B,R}^{(1)} \mathbf{V}_B \mathbf{b} + \mathbf{H}_{U,R}^{(1)} \mathbf{a} + \mathbf{z}_R^{(1)}, \tag{31}$$

$$y_{Di}^{(1)} = \sum_{j=1}^{K} h_{Uj,Di}^{(1)} a_j + z_{Di}^{(1)}, \quad i \in \{1, \cdots, K\}, \tag{32}$$

where now $\mathbf{H}_{U,R}^{(1)}$ is of dimension $L \times J$.

During the second phase, the relays precode in a distributed manner, each transmitting $v_{R\ell} y_{R\ell}^{(1)}$, $\ell \in \{1, \cdots, L\}$ while all the users in U are silent. For $\mathbf{V}_R = \mathrm{diag}(v_{R1}, \cdots, v_{RL})$ then $\mathbf{V}_B = (\mathbf{H}_{R,D}^{(2)} \mathbf{V}_R \mathbf{H}_{B,R}^{(1)})^{-1} \mathrm{diag}(\beta_1, \cdots, \beta_K)$, and

$$\begin{aligned} y_{Di}^{(2)} =& \sum_{j=1}^{J} (\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)} v_{R\ell} h_{Uj,R\ell}^{(1)}) a_j + \beta_i b_i \\ &+ (\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)} v_{R\ell} z_{R\ell}^{(1)}) + z_{Di}^{(2)}, \quad i \in \{1, \cdots, K\} \end{aligned} \tag{33}$$

$$\begin{aligned} \mathbf{y}_B^{(2)} =& \mathbf{H}_{R,B}^{(2)} \mathbf{V}_R \mathbf{H}_{U,R}^{(1)} \mathbf{a} + \mathbf{H}_{R,B}^{(2)} \mathbf{V}_R \mathbf{H}_{B,R}^{(1)} \mathbf{V}_B \mathbf{b} \\ &+ \mathbf{H}_{R,B}^{(2)} \mathbf{V}_R \mathbf{z}_R^{(1)} + \mathbf{z}_B^{(2)}. \end{aligned} \tag{34}$$

After proper linear combinations of the above received signals, we get

$$\begin{aligned} \tilde{y}_{Di} &= v_{Di} y_{Di}^{(1)} + y_{Di}^{(2)} \\ &= \sum_{j=1}^{J} (v_{Di} h_{Uj,Di}^{(1)} + \sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)} v_{R\ell} h_{Uj,R\ell}^{(1)}) a_j \\ &\quad + \beta_i b_i + (\sum_{\ell=1}^{L} h_{R\ell,Di}^{(2)} v_{R\ell} z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di} z_{Di}^{(1)} \end{aligned} \tag{35}$$

and the precoding vector $\mathbf{v}_6 = [v_{R1} \cdots v_{RL} \quad v_{D1} \cdots v_{DK}]^T$, is drawn from the non-empty null-space of the $JK \times (L+K)$ rank-JK matrix $\mathbf{G}_6$ with non-zero elements $g_{m,n}$ $g_{J(i-1)+j,\ell} \triangleq h_{R\ell,Di}^{(2)} h_{Uj,R\ell}^{(1)}$, $g_{J(i-1)+j,L+i} \triangleq h_{Uj,Di}^{(1)}$. Arguing as before we can conclude that $\mathcal{X}_4'$ allows for each downlink user to achieve the optimal $d_D = 1/2$ DOF, while the uplink users can achieve $d_U = J/2K$ DOF.

C. Signal and noise power analysis

In what follows we analyze the signal attenuation and noise accentuation effects of precoding on the achieved DOF. The analysis, provided here for the case of $\mathcal{X}_2$, readily applies with minor modifications to the rest of the schemes.

We begin with the downlink, corresponding to (12), and note that the unit norm precoder $\mathbf{v}_2$ satisfying (14), removes the uplink interference a_j at all Di, and (12) can then be written as

$$\begin{aligned} \tilde{y}_{Di} &= \beta_i b_i + (\sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)} v_{R\ell} z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di} z_{Di}^{(1)} \\ &= \beta_i b_i + \tilde{z}_{Di}, \quad i \in \{1, 2\}, \end{aligned} \tag{36}$$

TABLE II

ACHIEVABLE DOF WHEN $K \geq 2$ AND $L > K, \forall J \in \{2, \cdots, K\}$. FOR $J = K$ (AND $J > K$), THE SCHEMES ACHIEVE THE OPTIMAL PERFORMANCE $(d_U = 1/2, d_D = 1/2)$.

	L	achievable DOF
$\mathcal{X}_3'$	$(J+1)K - J > L > J(K-1)$	$(\underbrace{d_U = \frac{1}{2}}_{\text{optimal}}, d_D = \frac{J}{2K})$
$\mathcal{X}_4'$	$JK > L > K(J-1)$	$(d_U = \frac{J}{2K}, \underbrace{d_D = \frac{1}{2}}_{\text{optimal}})$

where $\tilde{z}_{Di} = (\sum_{\ell=1}^{2} h_{R\ell,Di}^{(2)} v_{R\ell} z_{R\ell}^{(1)}) + z_{Di}^{(2)} + v_{Di} z_{Di}^{(1)}$ denotes the individual noise terms. The noise power in $\tilde{z}_{Di}$ naturally depends on the channel realization and is bounded by

$$\begin{aligned} \mathbb{E}[\tilde{z}_{Di} \tilde{z}_{Di}^*] &= |h_{R1,Di}^{(2)} v_{R1}|^2 + |h_{R2,Di}^{(2)} v_{R2}|^2 + |v_{Di}|^2 + 1 \\ &\leq |h_{R1,Di}^{(2)}|^2 + |h_{R2,Di}^{(2)}|^2 + 2, \quad i \in \{1, 2\} \end{aligned} \tag{37}$$

For an arbitrarily small positive constant ϵ, we define a region $\mathcal{E}_1 = \{h_{R1,Di}^{(2)}, h_{R2,Di}^{(2)} : \mathbb{E}[\tilde{z}_{Di} \tilde{z}_{Di}^*] \dot{>} \rho^0\}$ where the power of the equivalent noise term $\tilde{z}_{Di}$ is high enough to potentially affect the achievable DOF. We now note that

$$\begin{aligned} P(\mathcal{E}_1) &\overset{(a)}{\leq} P(\cup_{\ell=1}^{2} \{|h_{R\ell,Di}^{(2)}|^2 \dot{>} \frac{1}{2}(\rho^0 - 2)\}) \\ &\overset{(b)}{\dot{\leq}} \sum_{\ell=1}^{2} P(|h_{R\ell,Di}^{(2)}|^2 \dot{>} \rho^0) \overset{(c)}{\doteq} 0, \end{aligned} \tag{38}$$

where (a) results from (37), (b) results from the union bound, and (c) from the straightforward fact that $P(|f(\{x\})| \dot{>} \rho^0) \doteq 0$ where $f(\{x\})$ is a polynomial in a finite set of i.i.d. complex Gaussian random variables $\{x\}$.

With respect to (36), β_1 and β_2 maintain the power constraint at node B, i.e., $||\mathbf{V}_B||_F^2 \doteq ||(\mathbf{H}_{R,D}^{(2)} \mathbf{V}_R \mathbf{H}_{B,R}^{(1)})^{-1} \mathrm{diag}(\beta_1, \beta_2)||_F^2 = 1$[3]. Here we set

$$\beta_1 = \beta_2 = \beta = (||(\mathbf{H}_{R,D}^{(2)} \mathbf{V}_R \mathbf{H}_{B,R}^{(1)})^{-1}||_F^2)^{-1/2}$$

for simplicity and define $\mathcal{E}_2 = \{\beta^2 < \rho^{-\epsilon}\}$ as the event where β reduces the signal power in (36) enough to potentially affect the achievable DOF. We evaluate $P(\mathcal{E}_2)$ as

$$\begin{aligned} P(\mathcal{E}_2) &= P(||(\mathbf{H}_{R,D}^{(2)} \mathbf{V}_R \mathbf{H}_{B,R}^{(1)})^{-1}||_F^2 > \rho^{\epsilon}) \\ &\leq P(||(\mathbf{H}_{R,D}^{(2)})^{-1}||_F^2 ||(\mathbf{H}_{B,R}^{(1)})^{-1}||_F^2 ||(\mathbf{V}_R)^{-1}||_F^2 > \rho^{\epsilon}) \\ &\leq P(\{||(\mathbf{H}_{R,D}^{(2)})^{-1}||_F^2 > \rho^{\epsilon/3}\} \cup \{||(\mathbf{H}_{B,R}^{(1)})^{-1}||_F^2 > \rho^{\epsilon/3}\} \\ &\qquad \cup \{||(\mathbf{V}_R)^{-1}||_F^2 > \rho^{\epsilon/3}\}) \\ &\overset{(a)}{\dot{\leq}} \rho^{-\epsilon/3} + \rho^{-\epsilon/3} + P(\sum_{\ell=1}^{2} \frac{1}{||v_{R\ell}||^2} > \rho^{\epsilon/3}) \\ &\overset{(b)}{\dot{\leq}} 2\rho^{-\epsilon/3} + P(||v_{R1}||^2 < \rho^{-\epsilon/3}) + P(||v_{R2}||^2 < \rho^{-\epsilon/3}) \\ &\overset{(c)}{\dot{\leq}} 2\rho^{-\epsilon/3} + \rho^{-\epsilon/d_1} + \rho^{-\epsilon/d_2} \doteq \rho^{-\epsilon/d_0}, \end{aligned} \tag{39}$$

[3]In the scale of interest, power constraints are also satisfied on the relays with $||\mathbf{V}_B||_F^2 = 1$ and $||\mathbf{v}_2||^2 = 1$.

where (a) results from the union bound and Lemma 1, (b) results from the union bound, (c) results from Lemma 2, and where $d_1 > 3$, $d_2 > 3$ and $d_0 = \max\{d_1, d_2\}$ are finite constants.

Now we show that a downlink rate of $\mathcal{R}_D = (\frac{1}{2} - \frac{1}{2}\epsilon - \delta)\log\rho$, for any positive δ, is achievable with arbitrarily small probability of error (in the presence of ergodicity). We first note that the mutual information between b_i and $\tilde{y}_{Di}$ is

$$I_D \triangleq \frac{1}{2}\log(1 + \frac{\beta^2\rho/K}{\mathbb{E}[\tilde{z}_{Di}\tilde{z}^*_{Di}]}),$$

and then for $\mathcal{E} \triangleq \{\mathcal{E}_1 \cup \mathcal{E}_2\}$, the corresponding outage probability is bounded as

$$\begin{aligned}
P[I_D < \mathcal{R}_D] &= P[I_D < \mathcal{R}_D \mid \mathcal{E}]P[\mathcal{E}] + P[I_D < \mathcal{R}_D|\mathcal{E}^c]P[\mathcal{E}^c] \\
&\overset{(a)}{\leq} \sum_{i=1}^{2} P[\mathcal{E}_i] + P[\frac{1}{2}\log(1+\frac{\beta^2\rho/K}{\mathbb{E}[\tilde{z}_{Di}\tilde{z}^*_{Di}]}) < \mathcal{R}_D|\mathcal{E}^c] \\
&\overset{(b)}{\dot{\leq}} \rho^{-\epsilon/d_0} + P[\frac{1}{2}\text{logdet}(1 + \frac{\rho^{-\epsilon}\rho}{\rho^0}) < \mathcal{R}_D] \qquad (40) \\
&= \rho^{-\epsilon/d_0} + P[\frac{(1-\epsilon)^+}{2}\log\rho < (\frac{1-\epsilon}{2} - \delta)\log\rho] = \rho^{-\epsilon/d_0}
\end{aligned}$$

where (a) follows from the union bound, where (b) follows from (38) and (39), and where $\mathcal{E}^c$ is the complement of $\mathcal{E}$. With positive constants ϵ and δ becoming arbitrarily small, we can conclude that $d_D = 1/2$ is achievable.

Similarly for the uplink, we note that after removing the self-interference, (11) takes the form

$$\tilde{\mathbf{y}}_B = \mathbf{H}^{(2)}_{R,B}\mathbf{V}_R\mathbf{H}^{(1)}_{U,R}\mathbf{a} + \mathbf{H}^{(2)}_{R,B}\mathbf{V}_R\mathbf{z}^{(1)}_R + \mathbf{z}^{(2)}_B = \mathbf{A}_U\mathbf{a} + \tilde{\mathbf{z}}_B,$$

where $\mathbf{A}_U = \mathbf{H}^{(2)}_{R,B}\mathbf{V}_R\mathbf{H}^{(1)}_{U,R}$ and $\tilde{\mathbf{z}}_B = \mathbf{H}^{(2)}_{R,B}\mathbf{V}_R\mathbf{z}^{(1)}_R + \mathbf{z}^{(2)}_B$. We note that the covariance matrix $\mathbf{\Sigma}$ of the equivalent noise vector $\tilde{\mathbf{z}}_B$ naturally depends on the channel realization and is given by $\mathbf{\Sigma} = \mathbb{E}[\tilde{\mathbf{z}}_B\tilde{\mathbf{z}}^\dagger_B] = \mathbf{H}^{(2)}_{R,B}\mathbf{V}_R(\mathbf{H}^{(2)}_{R,B}\mathbf{V}_R)^\dagger + I$. Now let $\mathcal{E}_3 = \{\mathbf{H}^{(2)}_{R,B}, \mathbf{V}_R : \lambda_{max}(\mathbf{\Sigma}) > \rho^\epsilon\}$ be the channel region where the power of the equivalent noise term is high enough to potentially affect the achievable DOF, and observe that

$$\begin{aligned}
P(\mathcal{E}_3) &= P(\lambda_{max}(\mathbf{H}^{(2)}_{R,B}\mathbf{V}_R(\mathbf{H}^{(2)}_{R,B}\mathbf{V}_R)^\dagger) + 1 > \rho^\epsilon) \\
&\leq P(||\mathbf{H}^{(2)}_{R,B}\mathbf{V}_R||^2_F > \rho^\epsilon - 1) \\
&\leq P(||\mathbf{H}^{(2)}_{R,B}||^2_F||\mathbf{V}_R||^2_F > \rho^\epsilon - 1) \\
&\overset{(a)}{\leq} P(||\mathbf{H}^{(2)}_{R,B}||^2_F K > \rho^\epsilon - 1) \\
&\doteq P(||\mathbf{H}^{(2)}_{R,B}||^2_F > \rho^\epsilon) \overset{(b)}{\doteq} 0, \qquad (41)
\end{aligned}$$

where (a) results from the fact that $||\mathbf{v}||^2 = 1$ and (b) results from Lemma 1.

Corresponding to the equivalent channel matrix $\mathbf{A}_U$ in (41), we define $\mathcal{E}_4 = \{\mathbf{A}_U : \lambda_{min}(\mathbf{A}_U\mathbf{A}^\dagger_U) < \rho^{-\epsilon}\}$ to be the channel region where the received signal power is reduced to levels that can potentially affect the achieved DOF, and note that

$$\begin{aligned}
P(\mathcal{E}_4) &= P(\lambda_{min}(\mathbf{A}_U\mathbf{A}^\dagger_U) < \rho^{-\epsilon}) \\
&\leq P(||\mathbf{A}^{-1}_U||^2_F > \rho^\epsilon) \\
&\leq P(||(\mathbf{H}^{(2)}_{R,B})^{-1}||^2_F||(\mathbf{H}^{(1)}_{U,R})^{-1}||^2_F||(\mathbf{V}_R)^{-1}||^2_F > \rho^\epsilon) \\
&\overset{(a)}{\leq} P(||(\mathbf{H}^{(2)}_{R,B})^{-1}||^2_F|| > \rho^{\epsilon/3}) + P(||(\mathbf{H}^{(1)}_{U,R})^{-1}||^2_F > \rho^{\epsilon/3}) \\
&\quad + P(||(\mathbf{V}_R)^{-1}||^2_F > \rho^{\epsilon/3}) \\
&\overset{(b)}{\dot{\leq}} \rho^{-\epsilon/3} + \rho^{-\epsilon/3} + \sum_{\ell=1}^{2} P(|v_{R\ell}|^2 < \rho^{-\epsilon/3}) \\
&\overset{(c)}{\dot{\leq}} 2\rho^{-\epsilon/3} + \rho^{-\epsilon/d_3} + \rho^{-\epsilon/d_4} \overset{(c)}{\doteq} \rho^{-\epsilon/d_5}, \qquad (42)
\end{aligned}$$

where (a) results from the union bound, (b) results from Lemma 1 and the union bound, (c) results from Lemma 2, and where $d_3 > 3$, $d_4 > 3$ and $d_5 = \max\{d_3, d_4\}$ are finite constants.

What remains is to show that for any $\delta > 0$, an uplink sum rate $\mathcal{R} = (\frac{K}{2} - K\epsilon - \delta)\log\rho$ is achievable with arbitrarily small probability of error. Towards this we note that the mutual information between $\mathbf{a}$ and $\tilde{\mathbf{y}}_B$, is given by

$$I_U \triangleq \frac{1}{2}\text{logdet}\left(I + \rho\mathbf{A}_U\mathbf{A}^\dagger_U\mathbf{\Sigma}^{-1}\right),$$

and that the corresponding outage probability is bounded as

$$\begin{aligned}
P[I_U < \mathcal{R}] &= P[I_U < \mathcal{R} \mid \mathcal{E}_5]P[\mathcal{E}_5] + P[I_U < \mathcal{R}|\mathcal{E}^c_5]P[\mathcal{E}^c_5] \\
&\overset{(a)}{\leq} \sum_{i=3}^{4} P[\mathcal{E}_i] + P[\frac{1}{2}\log\det(I + \frac{\rho\lambda_{min}(\mathbf{A}_U\mathbf{A}^\dagger_U)I}{\lambda_{max}(\mathbf{\Sigma})}) < \mathcal{R}|\mathcal{E}^c_5] \\
&\overset{(b)}{\dot{\leq}} \rho^{-\epsilon/d_5} + 0 + P[\frac{1}{2}\text{logdet}(1 + \frac{\rho\rho^{-\epsilon}I}{\rho^\epsilon}) < \mathcal{R}] \\
&= \rho^{-\epsilon/d_5} + P[\frac{K(1-2\epsilon)^+}{2}\log\rho < (\frac{K(1-2\epsilon)}{2} - \delta)\log\rho] \\
&= \rho^{-\epsilon/d_5} \qquad (43)
\end{aligned}$$

where $\mathcal{E}_5 \triangleq \{\mathcal{E}_3 \cup \mathcal{E}_4\}$, where (a) follows from the union bound and the fact that $I_U \geq \frac{1}{2}\log\det(I + \frac{\rho\lambda_{min}(\mathbf{A}_U\mathbf{A}^\dagger_U)I}{\lambda_{max}(\mathbf{\Sigma})})$, where (b) follows from (41) and (42), and where $\mathcal{E}^c_5$ is the complement of $\mathcal{E}_5$. Setting ϵ and δ to be very small, allows us to conclude that each node in U can achieve $d_U = 1/2$ DOF.

IV. Discussion and Conclusion

We notice that the proposed schemes perform optimally in the case of having few nodes $(\mathcal{X}_1, \mathcal{X}_2)$, or in the case of having an abundance of relays $(\mathcal{X}'_3, \mathcal{X}'_4,\ J \approx K)$. Notable advantages of the schemes over possibly better performing interference alignment techniques include the computational simplicity of the linear solutions, and the much reduced time delay. Drawbacks of the proposed methods include their modest gains when not optimal, issues of causality when encoding is over two coherence times, as well as, when compared to DF schemes, the large amount of necessary CSIT. The issue of causality can be handled either with channel prediction or, as we have seen, by adding relays. Future work might consider combining the proposed schemes with Han-Kobayashi techniques.

V. Acknowledgments

The research leading to these results was partially supported by the Mitsubishi RD project Home-eNodeBS, and from the European Commissions 7th framework programme under grant agreement FP7-257616 also referred to as CONECT.

VI. Appendix

A. Proof of Lemma 1

Defining two sets of indices $S_1 = \{j|\tau_j = 1\}$ and $S_{-1} = \{j|\tau_j = -1\}$, we have

$$\begin{aligned} ||\mathbf{\Phi}^{-1}||_F^2 &= \mathrm{tr}[(\mathbf{\Phi}^\dagger\mathbf{\Phi})^{-1}] \\ &\le \theta^2 (\prod_{j\in S_1} \mathrm{tr}[(\mathbf{H}_j^\dagger \mathbf{H}_j)^{-1}]) \cdot (\prod_{j\in S_{-1}} \mathrm{tr}[\mathbf{H}_j^\dagger \mathbf{H}_j]) \\ &= \theta^2 (\prod_{j\in S_1} \sum_{i=1}^N \lambda_{j,i}^{-1}) \cdot (\prod_{j\in S_{-1}} \sum_{i=1}^N \lambda_{j,i}), \end{aligned} \quad (44)$$

with $\lambda_{j,i}$ denoting the ith eigenvalue of $\mathbf{H}_j^\dagger \mathbf{H}_j$. Thus we have

$$P(||\mathbf{\Phi}^{-1}||_F^2 \ge \rho^\epsilon) \dot{\le} P((\prod_{j\in S_1}\sum_{i=1}^N \lambda_{j,i}^{-1})(\prod_{j\in S_{-1}}\sum_{i=1}^N \lambda_{j,i}) \ge \rho^\epsilon). \quad (45)$$

Let $S_{-1}(\ell)$ denote the ℓth element of set S_{-1}, and define channel regions $\mathcal{A}_\ell \triangleq \{\mathbf{H}_{S_{-1}(\ell)} : \sum_{i=1}^N \lambda_{S_{-1}(\ell),i} \dot{>} \rho^0\}$, $\forall \ell \in \{1, \cdots, |S_{-1}|\}$, $\mathcal{B}_1 \triangleq \{\cup_{\ell=1}^{|S_{-1}|} \mathcal{A}_\ell\}$ and $\mathcal{B}_0 \triangleq \{\mathbf{H}_1, \cdots, \mathbf{H}_k : (\prod_{j\in S_1}\sum_{i=1}^N \lambda_{j,i}^{-1})(\prod_{j\in S_{-1}}\sum_{i=1}^N \lambda_{j,i}) \ge \rho^\epsilon\}$.

We know from [11] that for any eigenvalue λ of the $\mathbf{H}_j^\dagger \mathbf{H}_j$, then $P(\lambda \dot{>} \rho^0) \doteq 0$. Thus $\forall \ell$, we have that

$$P(\mathcal{A}_\ell) \doteq P(\lambda_{S_{-1}(\ell),max} \dot{>} \rho^0) \doteq 0, \quad (46)$$

where $\lambda_{S_{-1}(\ell),max}$ is the maximum eigenvalue of $\mathbf{H}_{S_{-1}(\ell)}^\dagger \mathbf{H}_{S_{-1}(\ell)}$. Thus (45) takes the form

$$\begin{aligned} &P(||\mathbf{\Phi}^{-1}||_F^2 \ge \rho^\epsilon) \\ &\dot{\le} P(\mathcal{B}_0) = P(\mathcal{B}_0|\mathcal{B}_1)P(\mathcal{B}_1) + P(\mathcal{B}_0|\mathcal{B}_1^c)P(\mathcal{B}_1^c) \\ &\overset{(a)}{\doteq} 0 + P(\mathcal{B}_0|\mathcal{B}_1^c)P(\mathcal{B}_1^c) \\ &\overset{(b)}{\le} P(\mathcal{B}_0|\bigcap_{\ell=1}^{|S_{-1}|}(\sum_{i=1}^N \lambda_{S_{-1}(\ell),i} \doteq \rho^0)) \doteq P(\prod_{j\in S_1}\sum_{i=1}^N \lambda_{j,i}^{-1} \ge \rho^\epsilon) \end{aligned} \quad (47)$$

where (a) follows from (46) and the union bound, (b) from the fact that $P(\mathcal{B}_0|\{\sum \lambda_{S_{-1}(\ell),i} \dot{\le} \rho^0\}) \le P(\mathcal{B}_0|\sum \lambda_{S_{-1}(\ell),i} \doteq \rho^0)$, and where $\mathcal{B}_1^c$ is the complement of $\mathcal{B}_1$.

We note that for the case $d = |S_1| = 0$, (47) is simply expressed as $P(||\mathbf{\Phi}^{-1}||_F^2 \ge \rho^\epsilon) \dot{\le} P(1 \ge \rho^\epsilon) = 0$. For the case $d = |S_1| \ge 1$, for $\mathcal{B}_2 \triangleq \{\mathbf{H}_1, ..., \mathbf{H}_k : \prod_{j\in S_1}\sum_{i=1}^N \lambda_{j,i}^{-1} \ge \rho^\epsilon\}$, and $\mathcal{B}_3 \triangleq \{\mathbf{H}_1, ..., \mathbf{H}_k : \cup_{j\in S_1}\{\sum_{i=1}^N \lambda_{j,i}^{-1} \ge \rho^{\epsilon/d}\}\}$, then (47) takes the form

$$\begin{aligned} P(||\mathbf{\Phi}^{-1}||_F^2 \ge \rho^\epsilon) &\dot{\le} P(\mathcal{B}_2) \\ &= P(\mathcal{B}_2|\mathcal{B}_3)P(\mathcal{B}_3) + P(\mathcal{B}_2|\mathcal{B}_3^c)P(\mathcal{B}_3^c) \\ &= P(\mathcal{B}_2|\mathcal{B}_3)P(\mathcal{B}_3) + 0 \\ &\le P(\mathcal{B}_3) \overset{(a)}{\le} \sum_{j=1}^{|S_1|} P(\sum_{i=1}^N \lambda_{j,i}^{-1} \ge \rho^{\epsilon/d}), \end{aligned} \quad (48)$$

where (a) results from the union bound. From [13] we know that for $\lambda_{j,min}$ being the smallest eigenvalue of the $\mathbf{H}_j^\dagger \mathbf{H}_j$, then $P(\lambda_{j,min} \le \rho^{-\epsilon}) \dot{\le} \rho^{-\epsilon}$. Thus we have

$$P(\sum_{i=1}^N \lambda_{j,i}^{-1} \ge \rho^{\epsilon/d}) \doteq P(\lambda_{j,min} \le \rho^{-\epsilon/d}) \dot{\le} \rho^{-\epsilon/d}, \quad (49)$$

and (48) takes the form

$$P(||\mathbf{\Phi}^{-1}||_F^2 \ge \rho^\epsilon) \overset{(a)}{\dot{\le}} |S_1|\rho^{-\epsilon/d} \doteq \rho^{-\epsilon/d}, \quad (50)$$

where (a) is from (49). $\square$

B. Proof of Lemma 2

Let $\mathbf{v} \triangleq [v_1 \cdots v_N]^T$, let $[\mathbf{E}_1 \quad \mathbf{E}_2] \triangleq \mathbf{E}$ where $\mathbf{E}_1$ and $\mathbf{E}_2$ are respectively $M \times M$ and $M \times (N-M)$ matrices, and note that

$$\mathbf{E}_1^{-1}\mathbf{E}\mathbf{v} = [I \quad \mathbf{E}_1^{-1}\mathbf{E}_2]\mathbf{v} = 0. \quad (51)$$

Now recalling that $\mathbf{E}_1^{-1} = \frac{\mathrm{adj}(\mathbf{E}_1)}{\det(\mathbf{E}_1)}$ where $\mathrm{adj}(\mathbf{E}_1)$ denotes the matrix of cofactors of $\mathbf{E}_1$, we can rewrite (51) as

$$[I \quad \frac{\mathrm{adj}(\mathbf{E}_1)}{\det(\mathbf{E}_1)}\mathbf{E}_2]\mathbf{v} = 0. \quad (52)$$

Let $\mathbf{E}_0 \triangleq \mathrm{adj}(\mathbf{E}_1)\mathbf{E}_2$ and let $\mathbf{v}' \triangleq [v_1' \cdots v_N']^T$ be a solution of (52) of arbitrary magnitude. Let v_{n+M}' be such that $|v_{n+M}'|^2 = 1/N$, $n \in \{1, \cdots, N-M\}$, and note that from (52) we have

$$v_i' = \frac{-1}{\det(\mathbf{E}_1)} \sum_{n=1}^{N-M} [\mathbf{E}_0]_{i,n} v_{n+M}, \quad i \in \{1, \cdots, M\}. \quad (53)$$

Now normalize $\mathbf{v}'$ by a factor $\gamma = (\sum_{i=1}^M |v_i'|^2 + \frac{N-M}{N})^{-1/2}$, to get $\mathbf{v} = \gamma \mathbf{v}'$ with $||\mathbf{v}||^2 = 1$. Let

$$i_1 = \arg \min_{i\in\{1,\cdots,M\}} |v_i|, \quad i_2 = \arg \max_{i\in\{1,\cdots,M\}} |v_i|,$$

and note that

$$\begin{aligned} &P(|v_i|^2 \le \rho^{-\epsilon}) \\ &\le P(|\gamma|^2|v_{i_1}'|^2 \le \rho^{-\epsilon}) \le P(\frac{|v_{i_1}'|^2}{M|v_{i_2}'|^2 + \frac{N-M}{N}} \le \rho^{-\epsilon}) \\ &= P(\frac{|\sum_{n=1}^{N-M}[\mathbf{E}_0]_{i_1,n}v_{n+M}'|^2}{M|\sum_{n=1}^{N-M}[\mathbf{E}_0]_{i_2,n}v_{n+M}'|^2 + \frac{N-M}{N}|\det(\mathbf{E}_1)|^2} \le \rho^{-\epsilon}). \end{aligned} \quad (54)$$

Defining $f_1(\{x\}) \triangleq |\sum_{n=1}^{N-M}[\mathbf{E}_0]_{i_1,n}v_{n+M}'|^2$, $f_2(\{x\}) \triangleq |\sum_{n=1}^{N-M}[\mathbf{E}_0]_{i_2,n}v_{n+M}'|^2$, $f_3(\{x\}) \triangleq |\det(\mathbf{E}_1)|^2$

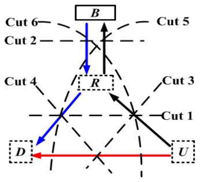

Fig. 3. Cut set model for the system.

and $f_4(\{x\}) \triangleq M f_2(\{x\}) + \frac{N-M}{N} f_3(\{x\})$, we note that all the entries of $\mathbf{E}_0$, and consequently $f_1(\{x\})$, $f_2(\{x\})$, $f_3(\{x\})$ and $f_4(\{x\})$ are, with probability 1, positive exponent polynomials (without constant terms) in the set of all pertinent random fading coefficients $\{x\}$.

It is then straightforward that $P(f_2(\{x\}) \dot{>} \rho^0) \doteq 0$ and $P(f_3(\{x\}) \dot{>} \rho^0) \doteq 0$, which consequently indicates that

$$P(f_4(\{x\}) \dot{>} \rho^0) \doteq 0. \tag{55}$$

At this point, the result in [10, Lemma 2.10] applies to tell us that there exists a finite constant d' such that

$$P(f_1(\{x\}) \leq \rho^{-\epsilon}) \quad \dot{\leq} \quad \rho^{-\epsilon/d'}. \tag{56}$$

Now define an event $\mathcal{A} = \{\{x\} : f_4(\{x\}) \dot{>} \rho^0\}$, and rewrite (54) as

$$\begin{aligned} P(|v_i|^2 \leq \rho^{-\epsilon}) &\leq P(\frac{f_1}{f_4} \leq \rho^{-\epsilon}) \\ &= P(\frac{f_1}{f_4} \leq \rho^{-\epsilon} | \mathcal{A}) P(\mathcal{A}) + P(\frac{f_1}{f_4} \leq \rho^{-\epsilon} | \mathcal{A}^c) P(\mathcal{A}^c) \\ &\overset{(a)}{\doteq} 0 + P(\frac{f_1}{f_4} \leq \rho^{-\epsilon} | \mathcal{A}^c) P(\mathcal{A}^c) \overset{(b)}{\dot{\leq}} P(\frac{f_1}{\rho^0} \leq \rho^{-\epsilon}) \\ &\doteq P(f_1 \leq \rho^{-\epsilon}) \overset{(c)}{\dot{\leq}} \rho^{-\epsilon/d'}, \quad i \in \{1, \cdots, N\}, \end{aligned} \tag{57}$$

where (a) results from (55), (b) results from the fact that $P(\frac{f_1}{f_4} \leq \rho^{-\epsilon} | \mathcal{A}^c) \leq P(\frac{f_1}{f_4} \leq \rho^{-\epsilon} | f_1 \doteq \rho^0)$, and (c) results from (56). $\square$

C. Proof of Lemma 3

We begin by noting that it is not difficult to see that the optimal DOF (not necessarily the optimal capacity though) can be achieved under a static phase duration setting, where there are two phases, one of fractional duration Δ_1 and the other of $1 - \Delta_1$. To derive the DOF outer bound we apply the Cut-Set Theorem (cf. [12]), in the ergodic setting ($M \to \infty$). For the uplink, by just considering 'cut 1' and 'cut 2' as illustrated in Fig. 3, the sum rate $K\mathcal{R}_U$ is upper bounded as

$$\begin{aligned} K\mathcal{R}_U \leq \max_{\Delta_1} \min\{ &\frac{\Delta_1}{M} \sum_{i=1}^{M} \max_{P(\mathbf{x})} \underbrace{I(\mathbf{x}_U; \mathbf{y}_R | \mathbf{x}_B, \mathbf{x}_R, \{\mathbf{H}_i\})}_{cut1}, \\ &\frac{1-\Delta_1}{M} \sum_{i=1}^{M} \max_{P(\mathbf{x})} \underbrace{I(\mathbf{x}_R; \mathbf{y}_B | \mathbf{x}_B, \{\mathbf{H}_i\})}_{cut2} \}, \end{aligned} \tag{58}$$

where $P(\mathbf{x})$ denotes the joint probability distribution of all the input signals in the network, and where $\{\mathbf{H}_i\}$ denotes the set of all channel coefficients associated with the ith block.

For sufficiently large number of blocks M, the first term ('cut 1') in (58) relates to the maximum rates in the $K \times L$ MIMO channel corresponding to K DOF. Similarly the second term ('cut 2') relates to the maximum rates of the $L \times K$ channel corresponding to K DOF. Hence, the corresponding number of DOF Kd_U is upper bounded as

$$\begin{aligned} Kd_U &\leq \max_{\Delta_1} \min\{\Delta_1 K, (1-\Delta_1)K\} \\ &= K/2, \quad \text{for} \quad \Delta_1 = 0.5, \end{aligned} \tag{59}$$

which gives that $d_U \leq 1/2$.

Similarly for the downlink, the sum rate $K\mathcal{R}_D$ is upper bounded as

$$\begin{aligned} K\mathcal{R}_D \leq \max_{\Delta_1} \min\{ &\frac{1-\Delta_1}{M} \sum_{i=1}^{M} \max_{P(\mathbf{x})} \underbrace{I(\mathbf{x}_R; \mathbf{y}_D | \mathbf{x}_U, \{\mathbf{H}_i\})}_{cut1}, \\ &\frac{\Delta_1}{M} \sum_{i=1}^{M} \max_{P(\mathbf{x})} \underbrace{I(\mathbf{x}_B; \mathbf{y}_R | \mathbf{x}_R, \mathbf{x}_U, \{\mathbf{H}_i\})}_{cut2} \}, \end{aligned} \tag{60}$$

and the corresponding number of DOF Kd_D as

$$\begin{aligned} Kd_D &\leq \max_{\Delta_1} \min\{(1-\Delta_1)K, \Delta_1 K\} \\ &= K/2, \quad \text{for} \quad \Delta_1 = 0.5, \end{aligned} \tag{61}$$

resulting in $d_D \leq 1/2$. $\square$

References

[1] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425-3441, Aug. 2008.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46: pp. 1204-1216, 2000.

[3] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse, "Transmission techniques for relay-interference networks," in *Proc. Allerton Conf. on Communication, Control, and Computing*, Illinois, Sep. 2008.

[4] Brian Smith and Sriram Vishwanath, "Network-coding in interference networks," in *Proc. Conf. on Information Sciences and Systems*, 2005.

[5] S.J. Kim, B. Smida and N. Devroye, "Capacity bounds on multi-pair two-way communication with a base-station aided by a relay," in *Proc. IEEE Int. Symp. Information Theory*, Austin, Jun. 2010.

[6] S. J. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5235-5241, Nov. 2008.

[7] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.

[8] R. Knopp, "Two-way wireless communication via a relay station," *GDRISIS meeting*, Mar. 2007.

[9] J. Chen, P. Elia and R. Knopp, "Relay-aided interference neutralization for the multiuser uplink-downlink asymmetric setting" submitted to *Proc. IEEE Int. Symp. Information Theory*, 2011.

[10] K. Sreeram, S. Birenjith and P. Vijay Kumar, "DMT of multi-hop cooperative networks-part I: basic results," submitted to *IEEE Trans. Inf. Theory*. Available Online: http://arxiv.org/abs/0808.0234.

[11] 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.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[13] Alan Edelman, *Eigenvalues and condition numbers of random matrices*, PhD thesis, Massachusetts Institute of Technology, 1989.