To Cancel or Not to Cancel: Exploiting Interference Signal Strength in the Eigenspace for Efficient MIMO DoF Utilization

Yongce Chen† Shaoran Li† Chengzhang Li† Y. Thomas Hou† Brian Jalaian‡
†Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
‡US Army Research Laboratory, Adelphi, MD, USA

***Abstract*—Degree-of-Freedom (DoF) based models have been widely used to study MIMO networks. To cancel interference, the number of DoFs used in the state-of-the-art DoF models is solely based on the number of interfering data streams. However, by decomposing an interference into the eigenspace, we find that signal strengths varies significantly in different directions for the same interference link. In this paper, we exploited the difference in interference signal strength in the eigenspace and differentiate strong and weak interference signals via their singular values. By introducing a concept of effective rank threshold, we propose to use DoFs only to cancel strong interference in the eigenspace based on this threshold while treating weak interference signals as noise in throughput calculation. We explore a fundamental trade-off between network throughput and effective rank threshold. Using simulation results on MU-MIMO networks, we show that network throughput under optimal rank threshold setting is significantly higher than that under existing DoF IC models. To ensure feasibility at the PHY layer, we present an algorithm that can find Tx and Rx weights at each node that can offer our desired DoF allocation.**

I. Introduction

Degree-of-Freedom (DoF) based models are a powerful tool to analyze MIMO's capabilities in spatial multiplexing (SM) and interference cancellation (IC) [1, 2]. By getting around the complex matrix manipulation (intractable in most cases) associated with each node's Tx/Rx weights and channel matrix, a DoF-based model can easily achieve resource allocation for SM and IC with simple "+/-" arithmetic calculations. Due to its simplicity and tractability, DoF-based models have been widely used in the research community for modeling, analysis, and optimization of MIMO networks [3–7].

Under a DoF-based model, a node's DoF resource can be used either for SM or IC, and the total number of available DoFs at the node is limited by its number of antennas. Existing DoF IC models require to consume DoFs to cancel all interference in the channel, without differentiating interference strength in different directions in the eigenspace. However, as we shall see in the following motivating example, interference signal strengths vary greatly in different directions in the eigenspace on the *same* interference link, especially as the number of antenna at a node becomes large and the channel exhibits correlation. This motivates us to reconsider the IC strategy in existing DoF models and explore a more efficient IC strategy beyond the state-of-the-art.

A Motivating Example. Considering a simple two-cell MIMO network shown in Fig. 1. There are two APs (AP1 and AP2) and two users (u_1 and u_2). Suppose each node (AP or user) is equipped with 12 antennas. AP1 transmits z_{11} data streams to user u_1 (marked with solid arrow lines) which interfere with user u_2 (marked with dashed arrow lines). Likewise, AP2 transmits concurrently z_{22} data streams to user u_2. In this motivating example, we assume the interference from AP2 to user u_1 is negligible.[1]

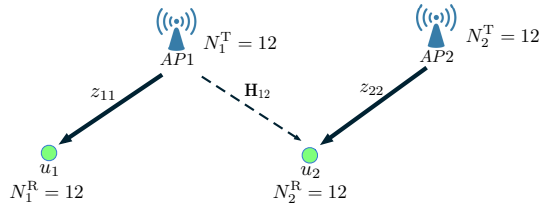


Fig. 1: A motivating example with two APs and two users.

Due to lack of rich multipath propagation and spatial separations, experimental studies showed that spatial channels within a MIMO link are often correlated [8, 9]. As a result, the transmit power from a node is generally not uniformly distributed in all directions of the channel's eigenspace. Consider the channel matrix $\mathbf{H}_{12}$ in Fig. 1. Using Kronecker channel model [9], we can write $\mathbf{H}_{12}$ as $\mathbf{H}_{12} = \mathbf{R}_{tx}^{1/2}\mathbf{H}_w\mathbf{R}_{rx}^{1/2}$, where $\mathbf{H}_w$ is an 12×12 random matrix with zero-mean i.i.d. complex Gaussian entries; $\mathbf{R}_{tx}^{1/2}$ and $\mathbf{R}_{rx}^{1/2}$ are 12×12 square root matrices of the Tx and Rx antenna correlation matrices, respectively. The (i,j)-th element in the correlation matrix $\mathbf{R}_{tx}$ and $\mathbf{R}_{rx}$ is calculated as $\rho_{tx}^{|i-j|}$ and $\rho_{rx}^{|i-j|}$, where $\rho_{tx} \in [0,1)$ and $\rho_{rx} \in [0,1)$ represent the levels of correlation between any two adjacent antennas at the respective transmitter and receiver (linear array with exponential correlation case [10]). For different values of ρ_{tx} and ρ_{rx}, we can simulate the expectations of singular values $\boldsymbol{\sigma}$ of $\mathbf{H}_{12}^{\dagger}\mathbf{H}_{12}$, which we show in Fig. 2. It is easy to see that for any given value of ρ_{tx} and ρ_{rx}, the expectations of singular values vary significantly. Here, a high singular value indicates that a large portion of AP1's power is projected into the direction of the corresponding singular vector. Likewise, a close-to-zero singular value indicates a close-to-zero portion of AP1's power is projected into the direction of the corresponding singular vector. When the values of ρ_{tx} and ρ_{rx} increases (i.e., with increased channel correlation), more and

[1] Such weak interference will be considered later.

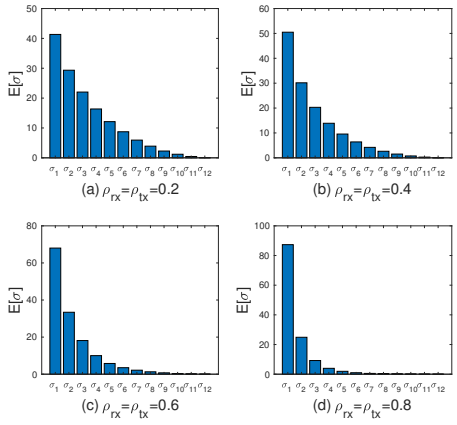
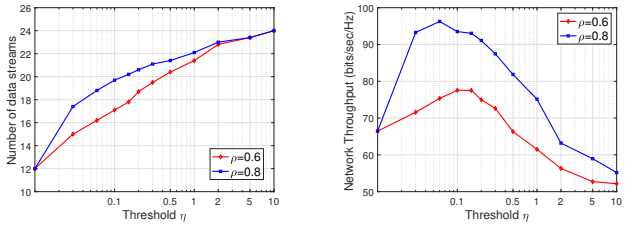


Fig. 2: Expectations of singular values $\mathbb{E}[\sigma]$ under different levels of correlation (ρ_{tx} and ρ_{rx}).



(a) (b)

Fig. 3: Total DoFs for SM and throughput performance as a function of threshold setting (used to differentiate strong and weak interferences). (a) Total number of data streams in the network. (b) Network throughput.

more expectations of singular values diminish toward zero.[2]

Figure 2 suggests that the interference strength varies significantly in different directions in its eigenspace. Under traditional IC scheme (see, e.g., [3–7]), all interference from AP1 to u_2 shall be cancelled by either AP1 (Tx side, using z_{22} DoFs) or u_2 (Rx side, using z_{11} DoFs). This approach does not differentiate strong and weak interferences in different directions and thus blindly cancels them all with precious DoFs. But if we exploit the difference of interference power strength in each direction, we could consider only cancelling the strong interference with DoFs and treat the weak ones just as noise. In other words, by exploiting the disparity in interference strength, we could conserve precious DoFs from cancelling the weaker ones. As shown in Fig. 2((c) and (d) in particular), the vast majority interference power only appears in the directions corresponding to the high singular values of $\mathbf{H}_{12}$, which can be properly cancelled by using a small number of DoFs. The remaining weak (small) interference power in these figures may be better treated as noise instead of being cancelled with precious DoFs. Although there may be some throughput loss due to un-cancelled weak interference, the DoFs savings could be used to transport more data streams (SM), which is a better direction for making a trade-off in DoF allocation. By judiciously exploiting the threshold used to differentiate strong and weak interference, one could achieve a better design objective (e.g., more data streams and/or higher throughput) than blindly cancelling all interferences (weak or strong) with DoFs, as in existing approaches [3–7].

To show the potential benefits of our idea, suppose we set $z_{11} = 12$ in the example in Fig. 1. Following traditional IC approach (i.e., no differentiation between strong and weak interferences), AP2 cannot send any data stream to user u_2 as there is no remaining DoF available at user u_2 to cancel interference from AP1. On the other hand, if u_2 treats the interference coming from AP1 in the direction corresponding to the smallest singular value of $\mathbf{H}_{12}$ as weak interference and does not use a DoF to cancel it, then it only needs to use 11 DoFs for IC from AP1 to u_2 and use the remaining one DoF to support one data stream transmission from AP2 to u_2. Following the same token, as more interferences from AP1 (corresponding to the smallest singular values) are treated as weak interferences and thus not to be cancelled with DoFs, more DoFs could be saved and be used to support SM from AP2 to u_2. As shown in Fig. 3(a), by increasing interference threshold η (more on this notation in Section II) to differentiate strong and weak interferences, more DoFs can be conserved from cancelling fewer number of weak interferences at u_2 and more data streams (SM) can be sent from AP2 to u_2. Fig. 3(b) shows the total network throughput (in bits/s/Hz) on all data streams (from AP1 to u_1 and AP2 to u_2) as a function of interference threshold η. Clearly, there is a trade-off between number of DoFs that can be transported in the network and total throughput. Under each ρ, there is an optimal knee point that offers the best trade-off between interference threshold η and total throughput. ∎

The above motivating example captures the key idea of this paper. As a major departure from existing approach for DoF IC, we exploit the differences in interference signal strength among different directions by examining singular values in the eigenspace, and we propose to expend DoFs only to cancel strong interference. Through optimal setting of interference strength threshold, we can achieve the best trade-off between the total number of DoFs for SM and network throughput. The main contributions of this paper are the following:

- This is the first paper on DoF IC models that exploits interference signal strengths in the eigenspace. Instead of performing IC with DoFs on all directions in the eigenspace as in existing DoF models, we advocate the potential benefits of performing IC with DoFs only on those directions with strong signals in the eigenspace.
- To differentiate strong and weak interference within an interference link, we introduce the concept of effective rank threshold. IC will only be performed for strong interference corresponding to large singular values in the eigenspace based on this effective rank threshold while weak interference will be treated as noise in throughput calculation.
- We study the fundamental trade-off between throughput and effective rank threshold for a MU-MIMO network. Through simulation results, we show that there exists an optimal trade-off between throughput and effective rank threshold. Further, the network throughput under optimal effective rank threshold setting is considerable higher than

[2] Apart from correlation, singular values can also be zero due to the presence of "key-hole" effect [11].

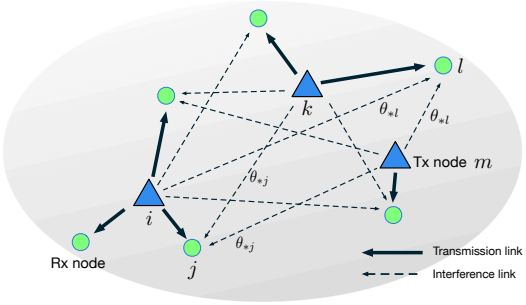


Fig. 4: A general MU-MIMO network with multiple Tx nodes and Rx nodes.

that under existing DoF IC models.

- To ensure the new DoF IC model is feasible at the PHY layer, we propose an algorithm to determine weights for all Tx and Rx nodes for a given DoF allocation. Through an iterative process, our algorithm can set the weights for all Tx and Rx nodes such that the interferences signals beyond the effective rank threshold can be suppressed nearly to zero, which is sufficient for decoding purpose.

II. Determine Effective Channel Rank of a Link

Consider a general MU-MIMO network (see Fig. 4) with a set $\mathcal{K}^{\mathrm{T}}$ of Tx nodes and a set $\mathcal{K}^{\mathrm{R}}$ of Rx nodes, respectively. Each Tx node $i \in \mathcal{K}^{\mathrm{T}}$ and Rx node $j \in \mathcal{K}^{\mathrm{R}}$ are equipped with N_i^{T} and N_j^{R} antennas, respectively. Under MU-MIMO, a Tx node is able to transmit to multiple Rx nodes concurrently while each Rx node can receive from at most one Tx node. For a Tx node $i \in \mathcal{K}^{\mathrm{T}}$, denote $\mathcal{K}_i^{\mathrm{R}}$ as the set of its Rx nodes. For a Rx node $j \in \mathcal{K}^{\mathrm{R}}$, denote $s(j)$ as its Tx node.

A. Effective Rank of A Single Interference Link

We first differentiate strong and weak interferences on a single interference link and use this differentiation to determine its effective rank. For a single interference link $k \to j$, instead of dealing directly with the fast fading channel matrix $\mathbf{H}_{kj} \in \mathbb{C}^{N_k^{\mathrm{T}} \times N_j^{\mathrm{R}}}$, we take into consideration of transmit power and path loss fading. Denote P_k as the transmit power at Tx node k and L_{kj} as the path loss from Tx node k to Rx node j. Define $\mathbf{Y}_{kj}$ as an $N_j^{\mathrm{R}} \times N_j^{\mathrm{R}}$ symmetric matrix by:

$$\mathbf{Y}_{kj} = \frac{P_k L_{kj}}{N_k^{\mathrm{T}}} \mathbf{H}_{kj}^{\dagger} \mathbf{H}_{kj}. \tag{1}$$

In matrix $\mathbf{Y}_{kj}$, each entry represents the received interference power on the corresponding channel on interference link $k \to j$. We will use $\mathbf{Y}_{kj}$ to determine the effective rank of interference link $k \to j$.

To differentiate strong and weak interferences, we employ the so-called *best rank-r approximation* of a matrix [13]. Under this approximation, $\mathbf{Y}_{kj}$ is decomposed through a SVD process and we use only the first r largest singular values and their corresponding singular vectors as an approximation.

Fact 1 *For a matrix $\mathbf{A} \in \mathbb{C}^{m \times n} (m \geq n)$, denote $\tilde{\mathbf{A}}$ as a rank-r matrix approximation of $\mathbf{A}$ with $r \in \{1, 2, \cdots, n\}$. The optimal solution to minimize approximation error*

$$\min_{\tilde{\mathbf{A}} \in \mathbb{C}^{m \times n}} \quad \left\| \mathbf{A} - \tilde{\mathbf{A}} \right\|_F, \qquad s.t. \quad \mathrm{rank}(\tilde{\mathbf{A}}) = r \tag{2}$$

where $||\cdot||_F$ denotes Frobenius norm, is

$$\tilde{\mathbf{A}} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\dagger}, \tag{3}$$

where σ_i, $\boldsymbol{u}_i$, and $\boldsymbol{v}_i$ are singular value, left and right singular vectors respectively from the SVD of $\mathbf{A}$, i.e., $\mathbf{A} = \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\dagger}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. The minimum approximation error (i.e., optimal objective value for (2)) is $\sqrt{\sum_{i=r+1}^{n} \sigma_i^2}$.

The SVD process in Fact 1 clearly shows the relative strength of interferences in different directions. The larger the singular value is, the stronger the interference in that direction. Based on the desired level of approximation error, we can approximate a rank-n matrix $\mathbf{A}$ by a rank-r matrix $\tilde{\mathbf{A}}$ with the r-strongest singular values of $\mathbf{A}$ through (3).

To apply best rank-r approximation on a single interference link $\mathbf{Y}_{kj}$, define θ as a threshold for singular values and denote r_{kj} as the effective channel rank of $\mathbf{H}_{kj}$. Then r_{kj} is given by

$$r_{kj} = \sum_{l=1}^{N_j^{\mathrm{R}}} \mathbb{1}\left\{\sigma_l(\mathbf{Y}_{kj}) \geq \theta\right\}, \tag{4}$$

where $\sigma_l(\mathbf{Y}_{kj})$ is the l-th singular value based on SVD of $\mathbf{Y}_{kj}$, and $\mathbb{1}\{\text{event}\}$ is an indicator function, which is 1 if event is true and 0 otherwise.

B. Interference Threshold at a Rx Node

Note that in a network with a set $\mathcal{K}^{\mathrm{T}}$ of Tx nodes and a set $\mathcal{K}^{\mathrm{R}}$ of Rx nodes, the interference threshold θ in (4) should be dependent upon the Rx node of this interference link. This is because the received power (from its intended transmitter) differs at each Rx node. As an example, consider Rx nodes j and l in Fig. 4. Rx node j is closer to its (intended) Tx node i than Rx node l to its (intended) Tx node k. For the same transmit power at i and k, Rx node j will receive a higher signal power (from its intended transmitter) and could thus tolerate a stronger interference. Then, for the interference links at Rx node j ($k \to j$ and $m \to j$), the threshold used to differentiate strong and weak interference should be larger than that used to differentiate stronger and weak interference on interference links ($i \to l$ and $m \to l$) for Rx node l. Based on the above discussion, for a receive node j, denote θ_{*j} as the threshold for singular values on its interference link. Then we should have $\theta_{*j} > \theta_{*l}$.

In this paper, instead of optimizing the settings of θ_{*j} for each individual Rx node j based on its (intended) received power level at Rx node j, we introduce a common scaling factor η across all Rx nodes to normalize its received power and only optimize the setting of this common scaling factor for the entire network. We define η as follows:

$$\theta_{*j} = \eta \frac{P_{s(j)} L_{s(j)j}}{N_{s(j)}^{\mathrm{T}}}. \tag{5}$$

Based on this definition of common scaling factor η, the effective rank r_{kj} of $\mathbf{H}_{kj}$ can be determined by the number

of $\mathbf{Y}_{kj}$'s singular values that are greater than or equal to the threshold $\eta \frac{P_{s(j)} L_{s(j)j}}{N^{\mathrm{T}}_{s(j)}}$. That is

$$r_{kj} = \sum_{l=1}^{N_j} \mathbb{1}\left\{\sigma_l(\mathbf{Y}_{kj}) \geq \eta \frac{P_{s(j)} L_{s(j)j}}{N^{\mathrm{T}}_{s(j)}}\right\},\ k \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}^{\mathrm{R}}_k. \tag{6}$$

Note that any negligible interference for IC will be treated as noise in the throughput calculation (see Section IV).

C. Effective Rank of An SM Link

For SM from node i to node j (intended transmission), the effective channel rank of $\mathbf{H}_{ij}$ can be determined by

$$r_{ij} = \sum_{l=1}^{N^{\mathrm{R}}_j} \mathbb{1}\left\{\sigma_l\left(\mathbf{H}^{\dagger}_{ij}\mathbf{H}_{ij}\right) \geq \theta_{\mathrm{SM}}\right\}, \quad i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}_i, \tag{7}$$

where θ_{SM} is the rank threshold for singular values on SM link $i \rightarrow j$. Note that the DoF savings by exploiting strong and weak interference can be made available for SM (more independent data streams) or diversity, both of which have the potential to increase the throughput. In this paper, for simplicity and focusing our attention on IC with DoF on interference links, we do not explore best SM-diversity trade-off. Instead, we will only pursue SM by transmitting more data streams as long as we have enough DoFs and assume θ_{SM} is a given constant throughout the paper.

III. IC Based on Effective Channel Rank

In the last section, we showed how to differentiate strong and weak interferences at a Rx node by setting a threshold for singular value and use this threshold to determine effective channel rank. In this section, we show how to perform IC (for strong interference only) in a MU-MIMO network based on this effective channel rank.

Note that DoF allocation for IC cannot be done arbitrarily and must follow certain rules to be feasible. By "feasibility", we mean that all the strong interference can be cancelled at the PHY layer. Section V will present details on PHY layer feasibility for our DoF allocation.

If DoF allocation for IC and SM is feasible at the PHY layer, then multiple data streams can be transmitted concurrently while all strong interference under best rank-r channels are cancelled. The remaining un-cancelled weak interference will be treated as noise and included in the throughput calculation in Section IV.

We employ the DoF-based IC model [12] to perform DoF allocation. In [12], the rank of a channel is assumed to be given *a priori* instead of being a function of effective rank threshold in this paper.

A. Modeling of DoF Constraints

DoF Constraints for SM For an intended transmission from Tx node i to Rx node j, denote the number of data streams on this link as z_{ij}. Denote $x_i(t)$ as a binary variable to indicate whether Tx node i is active or not at time t, i.e., $x_i(t) = 1$ if Tx node i is transmitting at time t and 0 otherwise. Likewise, denote $y_j(t)$ as a binary variable to indicate whether a Rx node j is active or not at time t, i.e., $y_j(t) = 1$ if Rx node j is receiving at time t and 0 otherwise.

If Tx node i is transmitting, then the total number of data streams transmitted to different receivers (under MU-MIMO) cannot exceed the total number of antennas at node i (i.e., N^{T}_i). We have

$$x_i(t) \leq \sum_{j \in \mathcal{K}^{\mathrm{R}}_i} z_{ij}(t) \leq N^{\mathrm{T}}_i x_i(t), \quad i \in \mathcal{K}^{\mathrm{T}}. \tag{8}$$

Similarly, if a Rx node j is active at time t, then the total number of DoFs used for reception (from only one transmitter under MU-MIMO) cannot exceed the number of antennas at node j (i.e., N^{R}_j). We have

$$y_j(t) \leq z_{ij}(t) \leq N^{\mathrm{R}}_j y_j(t), \quad i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}_i. \tag{9}$$

Taking into consideration of the effective rank of the SM link $i \rightarrow j$, the number of data streams that can be sent on this SM link cannot exceed the link's effective rank (see Section II). We have

$$z_{ij}(t) \leq r_{ij}(t), \quad i \in \mathcal{K}^{\mathrm{T}},\ j \in \mathcal{K}^{\mathrm{R}}_i. \tag{10}$$

For a Rx node l that is not Tx node i's intended receiver, i.e., $l \notin \mathcal{K}^{\mathrm{R}}_i$, the transmission at Tx node i is considered interference (instead of SM) and there is zero data streams over this link. We have

$$z_{il}(t) = 0, \quad k \in \mathcal{K}^{\mathrm{T}}, l \in \mathcal{K}^{\mathrm{R}}, l \notin \mathcal{K}^{\mathrm{R}}_i. \tag{11}$$

DoF Constraints for IC For interference from Tx node k to Rx node j, denote $d^{\mathrm{T}}_{kj}(t)$ as the number of consumed DoFs at Tx node k and $d^{\mathrm{R}}_{kj}(t)$ as the number of consumed DoFs at Rx node j that are needed to cancel this interference. Based on [12], a collaborative DoF consumption at both interfering Tx node k and Rx node j is the most efficient approach for IC when the rank of the interference channel is not full, as in our case. Denote $\mathbf{1}^{\mathrm{T}}_{kj}$ and $\mathbf{1}^{\mathrm{R}}_{kj}$ as two binary variables to indicate whether Tx node i (or Rx node j) consumes any DoFs for IC from k to j. That is, $\mathbf{1}^{\mathrm{T}}_{kj} = 1$ if Tx node k consumes DoFs for IC from k to j, $\mathbf{1}^{\mathrm{T}}_{kj} = 0$ otherwise; $\mathbf{1}^{\mathrm{R}}_{kj} = 1$ if Rx node j consumes DoFs for IC from k to j, $\mathbf{1}^{\mathrm{R}}_{kj} = 0$ otherwise.

If $x_k(t) = 1$ and $y_j(t) = 1$, then

$$d^{\mathrm{T}}_{kj}(t)\mathbf{1}^{\mathrm{T}}_{kj}(t) + d^{\mathrm{R}}_{kj}(t)\mathbf{1}^{\mathrm{R}}_{kj}(t) = \min\left\{\mathbf{1}^{\mathrm{R}}_{kj}(t)\sum_{l \in \mathcal{K}^{\mathrm{R}}_k}^{l \neq j} z_{kl}(t) + \mathbf{1}^{\mathrm{T}}_{kj}(t)\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq k} z_{ij}(t),\ r_{kj}(t)\right\}, \tag{12a}$$

$$\left(\mathbf{1}^{\mathrm{T}}_{kj}(t), \mathbf{1}^{\mathrm{R}}_{kj}(t)\right) \neq (0,0),\ k \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}} \tag{12b}$$

That is, the interference from k to j can be cancelled by consuming DoFs on Tx node k only (when $\left(\mathbf{1}^{\mathrm{T}}_{kj}(t), \mathbf{1}^{\mathrm{R}}_{kj}(t)\right) = (1,0)$), Rx node only (when $\left(\mathbf{1}^{\mathrm{T}}_{kj}(t), \mathbf{1}^{\mathrm{R}}_{kj}(t)\right) = (0,1)$), or both Tx and Rx nodes (when $\left(\mathbf{1}^{\mathrm{T}}_{kj}(t), \mathbf{1}^{\mathrm{R}}_{kj}(t)\right) = (1,1)$). Constraint (12) can be reformulated as mixed integer linear (MIL) constraints, which is omitted here to conserve space.

DoF Constraints at A Node A node can use its DoFs for both SM and IC, as long as the total number of consumed DoFs does not exceed the total available DoFs at the node. We consider DoF constraints at Tx node and Rx node separately. If node i is an active Tx node, we have

$$\text{if } x_i(t) = 1, \text{ then } \sum_{j \in \mathcal{K}_i^{\mathrm{R}}} z_{ij}(t) + \sum_{l \in \mathcal{K}^{\mathrm{R}}} d_{il}^{\mathrm{T}}(t) \mathbf{1}_{il}^{\mathrm{T}}(t) \le N_i^{\mathrm{T}}, \quad i \in \mathcal{K}^{\mathrm{T}}. \tag{13}$$

If node j is an active Rx node, we have

$$\text{if } y_j(t) = 1, \text{ then } \sum_{i \in \mathcal{K}^{\mathrm{T}}} z_{ij}(t) + \sum_{k \in \mathcal{K}^{\mathrm{T}}} d_{kj}^{\mathrm{R}}(t) \mathbf{1}_{kj}^{\mathrm{R}}(t) \le N_j^{\mathrm{R}}, \quad j \in \mathcal{K}^{\mathrm{R}}. \tag{14}$$

For constraint (13), it can be reformulated by incorporating binary variable $x_i(t)$ into the expression as follows:

$$\sum_{j \in \mathcal{K}_i^{\mathrm{R}}} z_{ij}(t) + \sum_{l \in \mathcal{K}^{\mathrm{R}}} d_{il}^{\mathrm{T}}(t) \mathbf{1}_{il}^{\mathrm{T}}(t) \le N_i^{\mathrm{T}} x_i(t) + (1 - x_i(t))B, \ i \in \mathcal{K}^{\mathrm{T}}, \tag{15}$$

where B is a large constant, which can be set as $B = \sum_{i \in \mathcal{K}^{\mathrm{T}}} N_i^{\mathrm{T}} + \sum_{j \in \mathcal{K}^{\mathrm{R}}} N_j^{\mathrm{R}}$ to ensure that B is an upper bound of $\sum_{l \in \mathcal{K}^{\mathrm{R}}} d_{il}^{\mathrm{T}}(t)$.

Similarly, constraint (14) can be reformulated as follows:

$$\sum_{i \in \mathcal{K}^{\mathrm{T}}} z_{ij}(t) + \sum_{k \in \mathcal{K}^{\mathrm{T}}} d_{kj}^{\mathrm{R}}(t) \mathbf{1}_{kj}^{\mathrm{R}}(t) \le N_j^{\mathrm{R}} y_j(t) + (1 - y_j(t))B, \ j \in \mathcal{K}^{\mathrm{R}}. \tag{16}$$

Constraints (15) and (16) can be reformulated as MIL constraints, which are omitted here to conserve space.

B. An Example with Numerical Results

As an example to illustrate the relationship between the total achievable data streams (SM) in the network and η (the common scaling factor to differentiate strong and weak interference and effective channel rank), consider the simple MU-MIMO network in Fig. 5. Suppose our objective is to maximize the sum of log of all data streams (SM) in the network. Then we have the following optimization problem:

$$\begin{aligned} \max \quad & \sum_{i \in \mathcal{K}^{\mathrm{T}}} \sum_{j \in \mathcal{K}^{\mathrm{R}}} \log(z_{ij}) \\ \text{s.t.} \quad & \text{SM constraints:(8) – (11);} \\ & \text{IC constraints:(12);} \\ & \text{Node's DoF constraints:(15), (16),} \end{aligned}$$

where $z_{ij}, d_{kj}^{\mathrm{T}}, d_{kj}^{\mathrm{R}}, \mathbf{1}_{kj}^{\mathrm{T}}$ and $\mathbf{1}_{kj}^{\mathrm{R}}$ are variables while all other symbols are constants.

As discussed earlier, the constraints in the above formulation can be reformulated into MIL constraints. However, the objective function (sum of log) remains non-linear. Fortunately, the sum of log objective can be reformulated (along with the MIL constraints) as a second order conic program (SOCP) [14]. Off-the-shelf optimization tools, such as Gurobi, can solve this SOCP (with integer variables) optimally.

Some numerical results follow. Suppose the six Tx nodes in Fig. 5 are uniformly generated in a 400 m × 400 m space. For each Tx node, there are two Rx nodes uniformly generated with a radius of 70 m of the Tx node. Each Tx node and

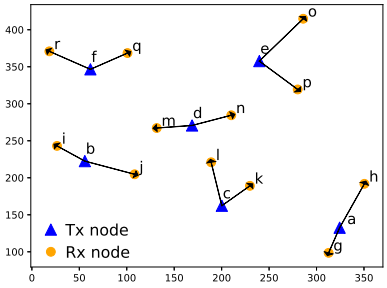
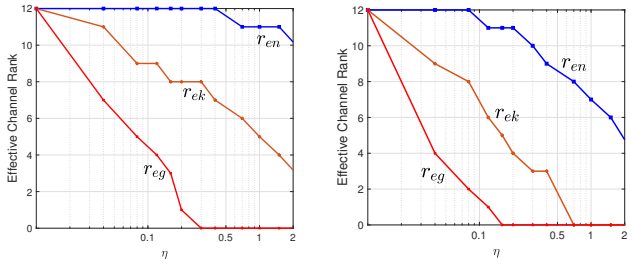


Fig. 5: An instance of MU-MIMO network topology.



(a) $\rho = 0.2$ (b) $\rho = 0.4$

Fig. 6: Effective ranks on interference links versus rank threshold scaling factor η.

Rx node are equipped with 16 and 12 antennas, respectively. Assume a fixed (constant) transmit power for each Tx node i, with SNR $P_i/n_0^2 = 80$ dB, where n_0^2 is the white noise power. Path loss is modeled as $L_{ij} = D_{ij}^{-3}$, with D_{ij} being the distance between Tx node i and Rx node j. Fast fading is modeled by Kronecker channel model, i.e., $\mathbf{H}_{ij} = \mathbf{R}_{tx}^{1/2} \mathbf{H}_w \mathbf{R}_{rx}^{1/2}$ ($i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}$), where $\mathbf{H}_w$ is an $N_i^{\mathrm{T}} \times N_j^{\mathrm{R}}$ random matrix with zero-mean i.i.d. complex Gaussian random numbers. The (k, l)-th element of the correlation matrix $\mathbf{R}_{rx}$ and $\mathbf{R}_{tx}$ is taken here as $\rho^{|k-l|}$ with $\rho \in \{0.2, 0.4, 0.6\}$. The rank threshold for SM links θ_{SM} is set to be 1.

Fig. 6 shows the effective ranks on three representative links ($e \to n, e \to k$ and $e \to g$) as a function of rank threshold scaling factor η (in log scale). We draw η in log scale since singular value distribution is more like a log-shape rather than a linear shape (see Fig. 2). As expected, all effective channel ranks are decreasing steadily. For $\rho = 0.2$ shown in Fig. 6(a), note that r_{en} remains full rank until η becomes greater than 0.4 while r_{ek} and r_{eg} starts to decrease when η starts to increase from 0. This is because Rx node n is close to the interfering Tx node e than k and g and thus experiences much stronger interference from Tx node e than k and g. On the other hand, r_{eg} drops very fast because Rx node g is further away from Tx node e than n and k. When η is greater than 0.3, $r_{eg} = 0$ and Rx node g is considered out of interference range of Tx node e. For $\rho = 0.4$ shown in Fig. 6(b), effective ranks have the similar trend but drop faster than those when $\rho = 0.2$, since the higher channel correlation causes interference strength more concentrated in fewer directions (see Fig. 2). Similar conclusion can be found for $\rho = 0.6$, and we omit the figure to conserve space. Clearly, the setting of rank threshold scaling factor η has different effect on different interference links

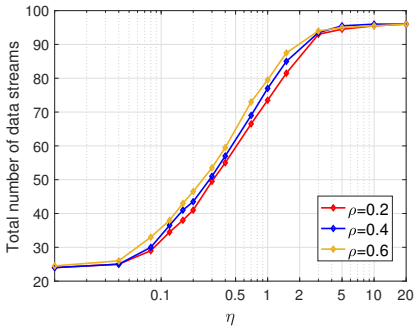
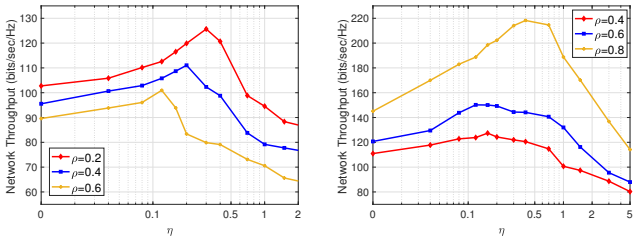


Fig. 7: Total number of data streams in the network (averaged over 10 instances).



(a) (b)

Fig. 8: Performance of network throughput under increasing threshold η. (a) Kronecker model used for both intended and interference links. (b) Kronecker model used for interference links and Rayleigh model used for intended links.

in terms of effective rank determination. Fig. 7 shows total number of data streams in the network from our optimal objective (averaged over 10 random network instances similar to Fig. 5). As shown in this figure, for a given ρ, the total number of data streams steadily increases from 24 to 96 and then flattens out. This is because the higher the rank threshold scaling factor η, the lower the effective channel ranks on interference links in the network. As a result, fewer DoFs are needed to cancel interferences and more DoFs can be allocated for SM. When η is greater than 10, the number of data streams cannot be further increased, either there is no room for further decrease of effective ranks on interference links (all effective ranks are 0), or further decrease of effective ranks on interference links will not improve objective value, due to the bounds on effective ranks on SM links. We also observe that for the same rank threshold η, larger number of data streams can be achieved for higher channel correlation level, due to lower effective ranks.

The above example shows the impact of effective rank threshold setting on the number of data streams that can be transported in the network. However, more number of data streams in the network do not necessarily mean higher throughput (in bits/s/Hz), due to un-cancelled interference (considered as noise) and channel hardening effect. In the next section, we investigate the impact of effective rank threshold setting on achievable throughput in the network.

IV. Throughput Calculation and Optimal Throughput-η trade-off

In this section, we calculate the actual throughput for a given DoF allocation for SM and IC. Then we explore the trade-off between throughput maximization and interference threshold scaling factor η directly.

A. Throughput Calculation

Assume a DoF allocation for SM and IC is feasible for a MU-MIMO network. Then the network throughput is the sum of the throughput achieved on each data stream under SM. So the key question is how to calculate throughput for each SM stream.

For each data stream, we can calculate its throughput by finding its SINR and then apply the Shannon capacity formula. The only subtlety here is that the SINR calculation should include all interferences that this data stream is suffering from, which includes all un-cancelled interference at PHY layer and white noise. To do this, we need to go to the PHY layer and work with the Tx and Rx vectors for each data stream. Denote $\mathbf{U}_i \in \mathbb{C}^{N_i^{\mathrm{T}} \times z_{i*}}$ as the weight matrix at Tx node i with z_{i*} outgoing SM data streams and $\mathbf{V}_j \in \mathbb{C}^{N_j^{\mathrm{R}} \times z_{*j}}$ as the weight matrix at Rx node j with z_{*j} incoming SM data streams. In Section V, we will show one implementation on how to derive $\mathbf{U}_i$ and $\mathbf{V}_j$ based on a DoF allocation while guaranteeing PHY layer feasibility. For now, let's assume the $\mathbf{U}_i$'s and $\mathbf{V}_j$'s are already found. Define the partition of matrix $\mathbf{U}_i$ as $\begin{bmatrix}\mathbf{U}_{i,j_1} & \mathbf{U}_{i,j_2} & \cdots & \mathbf{U}_{i,j_M}\end{bmatrix}$, where $j_1, j_2, \cdots, j_M$ are Tx node i's M recipients, i.e., $\{j_1, j_2, \cdots, j_M\} = \mathcal{K}_i^{\mathrm{R}}$, then $\mathbf{U}_{i,j_1}, \mathbf{U}_{i,j_2}, \cdots, \mathbf{U}_{i,j_M}$ are sub-weights corresponding to Rx nodes $j_1, j_2, ..., j_M$, with dimensions $N_i^{\mathrm{T}} \times z_{ij_1}, N_i^{\mathrm{T}} \times z_{ij_2}, \cdots, N_i^{\mathrm{T}} \times z_{ij_M}$ ($z_{i*} = \sum_{n=1}^{M} z_{ij_n}$), respectively.

For $\forall j \in \mathcal{K}_i^{\mathrm{R}}$, the signal-to-interference-plus-noise ratio (SINR) of f-th stream in the link $i \to j$ is then given by

$$\mathrm{SINR}_{ij}^{f} = \frac{\gamma_{ij}^{f}}{\mathbf{V}_j^{[*f]\dagger}\mathbf{Q}_j\mathbf{V}_j^{[*f]} - \gamma_{ij}^{f}}, \tag{17}$$

where $(\cdot)^{[*f]}$ is the f-th column of $(\cdot)$ and

$$\begin{aligned}\gamma_{ij}^{f} &= P_i L_{ij} \mathbf{V}_j^{[*f]\dagger} \mathbf{H}_{ij}^{\dagger} \mathbf{U}_{i,j}^{[*f]} \mathbf{U}_{i,j}^{[*f]\dagger} \mathbf{H}_{ij} \mathbf{V}_j^{[*f]},\\ \mathbf{Q}_j &= n_0^2 \mathbf{I}_{N_j} + \sum_{k \in \mathcal{K}^{\mathrm{T}}} P_k L_{kj} \mathbf{H}_{kj}^{\dagger} \mathbf{U}_k \mathbf{U}_k^{\dagger} \mathbf{H}_{kj}.\end{aligned} \tag{18}$$

Finally, the network throughput in bits/sec/Hz is given by

$$C = \sum_{i \in \mathcal{K}^{\mathrm{T}}} \sum_{j \in \mathcal{K}_i^{\mathrm{R}}} \sum_{f=1}^{z_{ij}} \log_2\left(1 + \mathrm{SINR}_{ij}^{f}\right). \tag{19}$$

B. Optimal Throughput-η Trade-off

From the network throughput expression (19), it is evident that there exists a trade-off between throughput and η. When η increases, more DoFs become available to support a larger number of SM data streams z_{ij} (as shown in Section III) and we have a larger value of z_{ij} in (19) to increase throughput. On the other hand, higher η means more weak interferences are not cancelled and left in the network. This will decrease the SINR term in (19) and decrease throughput. Thus, we have a trade-off. Unfortunately, due to the highly non-convex

nature of (19), a closed-form expression to explore optimal throughput-η trade-off remains unknown. In the rest of this section, we use simulation study to understand this throughput-η trade-off and gain insights.

We first use the same MU-MIMO network setting in Section III-B. 10 instances are randomly generated and we evaluate the average performance. Fig. 8(a) shows network throughput vs. η under different channel correlation levels ρ. Note that $\eta = 0$ stands for traditional DoF IC which uses DoFs to cancel interference indiscriminately in all directions in the eigenspace. For $\rho = 0.2$, we can see network throughput keep ascending until threshold $\eta = 0.3$, owing to more data streams are supported (see Fig. 7) while un-cancelled weak interference impacts are insignificant (Section V will show the interference level versus η). However, if we keep increasing η, throughput decreases due to un-cancelled interference. Throughput under $\eta = 0.6$ can be as good as that with traditional IC (i.e., $\eta = 0$). Nevertheless, by aggressively increasing η larger than 0.6, even though more DoFs can be made available for SM, un-cancelled interference will play a greater role and will result in worse performance than tradition IC. For $\rho = 0.4$ and 0.6, we can see a similar trade-off. For this network setting, the optimal effective rank threshold η should be set to $\eta = 0.3, 0.2$ and 0.12 for $\rho = 0.2, 0.4$ and 0.6, respectively. The peak throughput achieved at optimal η is 22.3%, 16.25%, 12.71% greater than that achieved at $\eta = 0$ for $\rho = 0.2, 0.4$ and 0.6, respectively. We can also note that with higher channel correlation level ρ, network throughput becomes lower. This is because high channel correlation also hinders SM capability, which results in low overall performance in this network setting.

In the scenarios where interference link presents high correlations (e.g., high correlation caused by key-hole effect) while intended link has low correlations, our rank-based IC can be even more beneficial. For Fig. 8(b), fast fading for interference links are modeled by $\mathbf{H}_{ij} = \mathbf{R}_{tx}^{1/2}\mathbf{H}_w\mathbf{R}_{rx}^{1/2}$ ($i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}_i^{\mathrm{R}}$), with $\rho \in \{0.4, 0.6, 0.8\}$, while fast fading for intended links are modeled by Rayleigh channel, i.e., $\mathbf{H}_{ij} = \mathbf{H}_w$ ($i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}_i^{\mathrm{R}}$). As shown in Fig. 8(b), the network throughput has a similar trend to Fig. 8(a) as we increase effective rank threshold η. However, we observed that for a higher channel correlation level ρ at interference links, we obtained much higher throughput gain by setting optimal effective rank threshold η. Specifically, the peak throughput achieved at optimal η is 15.82%, 24.44%, 50.48% greater than that achieved at $\eta = 0$ for $\rho = 0.4, 0.6$ and 0.8, respectively. This is because well-conditioned intended channels have the capability to carry more data streams for higher throughput, thus can fully take advantage of exploiting interference signal strength in the eigenspace on correlated interference channels. The trade-off shown in Fig. 8 reaffirms that cancelling interference in all directions is not efficient in terms of throughput performance.

V. Physical Layer Feasibility

In Section IV we assumed feasible weight matrices $\mathbf{U}_i$ and $\mathbf{V}_j$ at the PHY layer are given *a priori* corresponding to a particular DoF allocation. In this section, we show how to find such weight matrices at each node.

As expected, finding these feasible at the PHY layer for a MU-MIMO network is not trivial. First and foremost, the Tx weights and Rx weights are interdependent on each other. That is, the Tx weights for IC depend on the corresponding Rx weights, while the Rx weights for IC also depend on the corresponding Tx weights. There is no established guideline in the literature on how to find feasible weight matrices corresponding to a DoF allocation such that interference can be cancelled completely. Second, since we are exploring effective channel ranks in this paper and some weak interferences are not cancelled by DoFs, one cannot guarantee the existence of feasible $\mathbf{U}_i$ and $\mathbf{V}_j$ to achieve perfect (100%) interference-free transmission. This makes finding feasible weight matrices even more challenging.

In the rest of this section, we design an iterative algorithm that is able to realize the DoF allocation (based on the DoF solution for a specific objective as shown in Section III), where the strong interferences in best rank-r channels are "almost" cancelled. By "almost", we mean the remaining signal strength in the directions of strong interferences is close to zero.

A. Basic Idea and Algorithm Design

The main idea of our algorithm is as follows. For a given DoF allocation, we have the data stream allocation (i.e. z_{ij}) on each SM link in the network, which we can use to determine the dimension for each $\mathbf{U}_i$ and $\mathbf{V}_j$. Then, under the original channel matrix $\mathbf{H}_{ij}$, to cancel all the inter-stream and inter-node interference, we must have

$$\mathbf{U}_i^{\dagger}\left[\mathbf{H}_{ij_1}\mathbf{V}_{j_1} \quad \mathbf{H}_{ij_2}\mathbf{V}_{j_2}\cdots\right] = \mathbf{\Lambda}_{z_{i*}}, i \in \mathcal{K}^{\mathrm{T}}, j_1, j_2 \ldots \in \mathcal{K}_i^{\mathrm{R}}, \quad (20)$$

$$\mathbf{U}_i^{\dagger}\mathbf{H}_{ij}\mathbf{V}_j = \mathbf{0}, \quad i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}_i^{\mathrm{R}}. \quad (21)$$

Although (20) can always be satisfied for all SM links, (21), however, cannot be satisfied for all $i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}_i^{\mathrm{R}}$ if there are not enough remaining DoFs to cancel those weak interference on some links. Recognizing that not all interference can be perfectly cancelled, we focus our goal on cancelling all the strong interference, which is based on the best rank$-r$ approximate channel $\tilde{\mathbf{H}}_{ij} = \sum_{l=1}^{r_{ij}} \sigma_l \boldsymbol{u}_l \boldsymbol{v}_l^{\dagger}$ via SVD of $\mathbf{H}_{ij}$. That is, we want to have

$$\mathbf{U}_i^{\dagger}\tilde{\mathbf{H}}_{ij}\mathbf{V}_j = \mathbf{0}, \quad \text{for } i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}_i^{\mathrm{R}}. \quad (22)$$

The weak interference that is not cancelled will reduce network throughput and will be taken into account in throughput calculation (as we did in Section IV).

(20) and (22) constitute a system of bilinear equations and a general solution to bilinear equations remains unknown. Instead of finding a feasible solution to (20) and (22), we propose to minimize the LHS of (22) for all $i \in \mathcal{K}^{\mathrm{T}}, j \in \mathcal{K}^{\mathrm{R}}, j \notin \mathcal{K}_i^{\mathrm{R}}$, subject to (20). Denote Δ_{LI} as the leakage interference in the network,[3] which is defined as

$$\Delta_{\mathrm{LI}} = \sum_{i \in \mathcal{K}^{\mathrm{T}}} \sum_{j \in \mathcal{K}^{\mathrm{R}}}^{j \notin \mathcal{K}_i^{\mathrm{R}}} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger}\tilde{\mathbf{H}}_{ij}\mathbf{V}_j \right\|_F^2. \quad (23)$$

[3]Incidentally, a similar definition of leakage interference involving only channel matrix $\mathbf{H}_{ij}$ is given in [15, 16].

The problem to solve is to minimize Δ_{LI} subject to (20). To this end, we propose a simple yet effective approach to address the dependency between Tx weight matrices $\mathbf{U}_i$ and Rx weight matrices $\mathbf{V}_j$ by updating each in an alternating fashion (i.e., fixing $\mathbf{U}_i$ and update $\mathbf{V}_j$ and vice versa). Specifically, at each iteration, Tx weight matrices $\mathbf{U}_i$ are optimized first with given Rx weight matrices $\mathbf{V}_j$ and channel information. Then we optimize Rx weight matrices $\mathbf{V}_j$ with given Tx weight matrices $\mathbf{U}_i$ and channel information. For each weight matrix optimization, the weight matrix is updated by solving a minimization problem with the objective Δ_{LI} and updated set of constraints. The iteration terminates when we find no improvement following a number of consecutive iterations.

To conserve space we offer the key steps of the algorithm here. Details can be found in [17].

- **Step 1: Initialization.** Initially all the Tx and Rx weight matrices, which can be set arbitrarily but have to be full rank with dimension $N_i^{\mathrm{T}} \times z_{i*}$ and $N_j^{\mathrm{R}} \times z_{*j}$, respectively.
- **Step 2: Tx weights optimizing.** In this step, channel information $\tilde{\mathbf{H}}_{ij}$ and Rx weight matrices $\mathbf{V}_j$ are given. Denote $\Delta_{\mathrm{LI},i}^{\mathrm{T}} = \sum_{j \in \mathcal{K}^{\mathrm{R}}}^{j \notin \mathcal{K}_i^{\mathrm{R}}} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \tilde{\mathbf{H}}_{ij} \mathbf{V}_j \right\|_F^2$ as the leakage interference at Tx node i. It follows that $\min \Delta_{\mathrm{LI}}$ can be solved separately by solving $|\mathcal{K}^{\mathrm{T}}|$ independent sub-problems $\min_{\mathbf{U}_i} \Delta_{\mathrm{LI},i}^{\mathrm{T}}$, i.e., one sub-problem for each Tx node. Then the Tx weight matrix $\mathbf{U}_i$ is updated by solving the following optimization problem:

$$\min_{\mathbf{U}_i \in \mathbb{C}^{N_i^{\mathrm{T}} \times z_{i*}}} \Delta_{\mathrm{LI},i}^{\mathrm{T}} = \sum_{j \in \mathcal{K}^{\mathrm{R}}}^{j \notin \mathcal{K}_i^{\mathrm{R}}} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \tilde{\mathbf{H}}_{ij} \mathbf{V}_j \right\|_F^2. \tag{24}$$

- **Step 3: Rx weights optimizing.** Similar to optimizing Tx weight matrices, we have $|\mathcal{K}^{\mathrm{R}}|$ independent sub-problems for $|\mathcal{K}^{\mathrm{R}}|$ Rx nodes, and each Rx weight matrix $\mathbf{V}_j$ is updated by solving the following optimization problem:

$$\min_{\mathbf{U}_i \in \mathbb{C}^{N_i^{\mathrm{T}} \times z_{i*}}} \Delta_{\mathrm{LI},i}^{\mathrm{T}} = \sum_{j \in \mathcal{K}^{\mathrm{R}}}^{j \notin \mathcal{K}_i^{\mathrm{R}}} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \tilde{\mathbf{H}}_{ij} \mathbf{V}_j \right\|_F^2. \tag{25}$$

  Step 2 and Step 3 will be performed iteratively until there is no improvement for W consecutive iterations, i.e., $\Delta_{\mathrm{LI}}(t-w-1) - \Delta_{\mathrm{LI}}(t-w) < \epsilon, w = 0, 1, ..., W-1$ is met for given convergence threshold ϵ.
- **Step 4: Separating desired data streams.** To separate desired data streams within an SM link (i.e., to satisfy (20)), standard linear ZF design can be employed.
- **Step 5: Power allocation.** Equal power allocation for every data streams is applied subject to power constraints $\mathrm{Tr}(\mathbf{U}_i \mathbf{U}_i^{\dagger}) = 1, \mathrm{Tr}(\mathbf{V}_j \mathbf{V}_j^{\dagger}) = 1$.

While the algorithm minimizes leakage interference at every iteration and is guaranteed to converge, convergence to global minimum is not guaranteed due to the non-convex nature. Nevertheless, our own experience shows that the solution in each iteration is in closed form and the algorithm typically converges by less than 50 iterations. So this algorithm is computationally fast and effective for practical purpose. Numerical results for performance are presented in the following subsection.

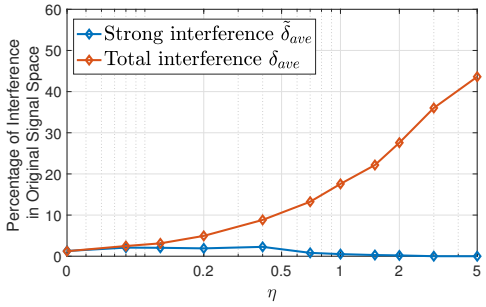


Fig. 9: Average percentage of residual strong interference and total interference in original signal space.

B. Performance

In this section, we provide numerical results to show that by applying this algorithm, the interferences which are intended to be cancelled (i.e., strong interferences under best rank-r channels) can be appropriately suppressed almost to zero. After applying the proposed algorithm, $\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \tilde{\mathbf{H}}_{ij} \mathbf{V}_j \right\|_F^2$ indicates the sum of residual strong interference power at the Rx node j after IC, which is ideally to be 0, while $\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} \frac{P_i}{N_i^{\mathrm{T}}} L_{ij} \left\| \mathbf{H}_{ij} \right\|_F^2$ indicates the interference power in the original signal space without IC performed. Then the average ratio of residual strong interference in the original signal space over all Rx nodes is defined as

$$\tilde{\delta}_{\mathrm{ave}} = \frac{1}{|\mathcal{K}^{\mathrm{R}}|} \sum_{j \in \mathcal{K}^{\mathrm{R}}} \frac{\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \tilde{\mathbf{H}}_{ij} \mathbf{V}_j \right\|_F^2}{\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} \frac{P_i}{N_i^{\mathrm{T}}} L_{ij} \left\| \mathbf{H}_{ij} \right\|_F^2}. \tag{26}$$

For analysis, we also evaluate the total interference level including un-cancelled interference, for which the original channels should be considered. The average ratio of residual total interference in the original signal space is defined as

$$\delta_{\mathrm{ave}} = \frac{1}{|\mathcal{K}^{\mathrm{R}}|} \sum_{j \in \mathcal{K}^{\mathrm{R}}} \frac{\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} P_i L_{ij} \left\| \mathbf{U}_i^{\dagger} \mathbf{H}_{ij} \mathbf{V}_j \right\|_F^2}{\sum_{i \in \mathcal{K}^{\mathrm{T}}}^{i \neq s(j)} \frac{P_i}{N_i^{\mathrm{T}}} L_{ij} \left\| \mathbf{H}_{ij} \right\|_F^2}. \tag{27}$$

Considering the same network setting as Section IV with $\rho = 0.4$, we set $\epsilon = 0.01$ and $W = 5$, and take average of simulation results over 10 instances. Fig. 9 shows our algorithm keeps the strong interferences that are to be cancelled at a very low level (within 2% in the original signal space without IC performed), while the total interference level increases as more weak interferences are not being cancelled by DoFs. More importantly, Fig. 9 suggests that when η is small (e.g., $\eta \leq 0.3$), by not cancelling weak interference, the gap between total interference level and strong interference level is small (within 6% when $\eta \leq 0.3$). Thus, leaving some weak interferences un-cancelled by DoFs will not cause much interference. However, when η is large, the different between total interference and strong interference could become large. Here, un-cancelled interference may cause considerable performance loss in network throughput. Such trade-off has been validated in Section IV.

VI. Related Work

There have been active research activities on DoF-based IC in MIMO networks. However, none of the existing DoF models differentiate strong and weak interference in different directions in the eigenspace per interference link, as we have done in this paper.

In Information Theory (IT) community, DoF characterizations are mainly based on idealized channel models, i.e., either full rank (e.g. [18, 19]) or rank-deficient with zero singular values (e.g. [20, 21]). Such idealized channel rank models do not exactly capture what happens in reality, where singular values for weak interference are not exactly zero. As a result, they cannot accurately represent channel behavior in the real world.

In networking community, most existing DoF-based models assume that channels are of full rank [3–7]. To identify interference footprint, the so-called protocol model is widely used [3–7], where a Rx node inside a disc is considered interfered and would require DoF for IC while a Rx node outside the disc is considered to experience negligible interference (i.e., no IC is needed). The main issue with this model is that, for the same Rx node (inside the interference range), it does not differentiate interference strength in different directions in the eigenspace and thus would require DoFs to cancel interference in all directions (for the same Rx node) even though the signal strength in certain directions may be very weak and does not require DoFs to cancel. The weakness of these models is further amplified when the number of antennas at Tx/Rx nodes becomes large and channels exhibit high correlation. In contrast, instead of using a disc (or interference range), we differentiate interference strength by examining singular values in the eigenspace regardless of the location of the Rx node. Strong interferences (corresponding to large singular values) are cancelled by DoFs while weak interferences (corresponding to small singular values) are treated as noise in throughput calculation. This approach combines the best of the simplicity of DoF model and the accuracy of the physical model.

VII. Conclusions

This paper exploited interference signal strengths among different directions in the eigenspace to achieve efficient DoF IC for MU-MIMO networks. By decomposing an interference into its eigenspace and introducing an effective rank threshold to differentiate strong and weak interference, we showed that precious DoFs can be conserved if we only use DoFs to cancel those strong interference signals in the eigenspace. We investigated the trade-off between network throughput and effective rank threshold and showed that network throughput under the optimal effective rank threshold is significantly higher than that under existing DoF IC models. To ensure feasibility at the PHY layer, we presented an algorithm that can find Tx and Rx weights at each node that can offer our desired DoF allocation.

Acknowledgments

This research was supported in part by NSF grants 1642873, 1617634, and 1443889.

References

[1] L. Zheng and D.N.C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] S.A. Jafar and M.J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2637–2642, June 2007.

[3] D.M. Blough, G. Resta, P. Santi, R. Srinivasan, and L. M. Cortés-Pena, "Optimal one-shot scheduling for MIMO networks," in *Proc. IEEE SECON* Salt Lake City, UT, USA, , pp. 404–412, June, 2011.

[4] Y. Shi, J. Liu, C. Jiang, C. Gao, and Y.T. Hou, "A DoF-based link layer model for multi-hop MIMO networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1395–1408, July 2014.

[5] H. Yu, O. Bejarano, and L. Zhong, "Combating inter-cell interference in 802.11ac-based multi-user MIMO networks," *In Proc. ACM MobiCom*, pp. 141–152, Maui, Hawaii, USA, Sept. 2014.

[6] B. Hamdaoui and K.G. Shin, "Characterization and analysis of multi-hop wireless MIMO network throughput," in *Proc. ACM MobiHoc*, pp. 120–129, Montreal, Quebec, Canada, Sept. 2007.

[7] R. Bhatia and L. Li, "Throughput optimization of wireless mesh networks with MIMO links," in *Proc. IEEE INFOCOM*, pp. 2326–2330, Barcelona, Spain, May 2007.

[8] J.P. Kermoal, L. Schumacher, K.I. Pedersen, P.E. Mogensen and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation", *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[9] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wide-band MIMO radio channels based on NLoS indoor measurements", *IEEE Transactions on Vehicular Technology*, vol. 53, no. 3, pp. 655–665, May 2004.

[10] J. Choi and D.J. Love, "Bounds on eigenvalues of a spatial correlation matrix", *IEEE Communications Letters*, vol. 18 no. 8, pp. 1391–1394, Aug. 2014.

[11] H. Shin and J.H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole". *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[12] Y. Chen, Y. Huang, Y. Shi, Y.T. Hou, W. Lou and S. Kompella, "A general model for DoF-based interference cancellation in MIMO networks with rank-deficient channels," in *Proc. IEEE INFOCOM*, pp. 900–907, Honolulu, Hawaii, USA, Apr. 2018.

[13] C. Eckart, and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[14] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: Analysis, algorithms, and engineering applications.* SIAM, 2001. ISBN: 9780898714913.

[15] K. Gomadam, V.R. Cadambe and S.A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," *in Proc. IEEE GLOBECOM*, pp. 1–6, New Orleans, LO, USA, Nov. 2008.

[16] S.W. Peters and R.W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 206–218, Oct. 2011.

[17] Y. Chen, S. Li, C. Li, Y.T. Hou and B. Jalaian, "To cancel or not to cancel: Exploiting interference signal strength in the eigenspace for efficient MIMO DoF utilization," *Technical Report*, Dept. of ECE, Virginia Tech, Jan. 2019. [Online]. https://sites.google.com/vt.edu/ychen-infocom2019-tr-pdf

[18] S.A. Jafar, and M.J. Fakhereddin, "Degrees of freedom for the MIMO interference channel,", *IEEE Transactions on Information Theory*,vol. 53, no. 7, pp. 2637–2642, June 2007.

[19] M. Razaviyayn, G. Lyubeznik and Z.Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO interference channel," *IEEE Transactions on Signal Processing*, vol. 60 no. 2, pp. 812–821, Feb. 2012.

[20] S.R. Krishnamurthy, A. Ramakrishnan, and S.A. Jafar, "Degrees of freedom of rank-deficient MIMO interference channels," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 341–365, Jan. 2015.

[21] Y. Zeng, X. Xu, Y.L. Guan, E. Gunawan, and C. Wang, "Degrees of freedom of the three-user rank-deficient MIMO interference channel," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4179–4192, Aug. 2014.