Exploiting Coherence Time Variations for Opportunistic Blind Interference Alignment

Milad Johnny and Alireza Vahid

Abstract

The observed coherence times associated with different wireless transmitters at any given user may vary at different rates. We demonstrate how these variations can be exploited for interference management. More precisely, we propose a new opportunistic blind interference alignment (BIA) strategy in the context of K-user interference channels that exploits these variations in coherence times. We first provide a proof-of-concept setup in which the information about coherence time variations is available non-causally to the transmitters, and we demonstrate how transmitters and receivers can perform pre-coding and post-processing, respectively, to align a considerable part of the interference signal power. We note that in this non-causal scenario, no channel state information is available to the transmitters, and we make no specific assumption on the channel distributions. We take the key ideas of this scenario and consider a $K-$user interference channel in which direct links vary at a higher pace compared to cross-links. This assumption is motivated by considering mobile users, or by using our proposed transmit antenna for stationary or low-mobility users. We show how we can eliminate the need for non-causal knowledge of coherence time variations, and still provide significant capacity gains through our opportunistic BIA strategy.

Index Terms

Blind interference alignment, average achievable sum rate, channel state information, channel changing pattern, multi-layer encoding.

I. Introduction

The interference channel (IC) is a communication network with multiple transmitter-receiver (transceiver) pairs, and serves as a model for many real-world wireless networks such as cellular

Alireza Vahid is with the department of Electrical Engineering at the University of Colorado Denver, USA. Email: alireza.vahid@ucdenver.edu.

Preliminary results of this work are submitted to IEEE International Symposium on Information Theory (ISIT).

communication networks and wireless local area networks. Today, due to the increasing number of users and the higher quality of service requirements such as HD video connection, there is an increasing demand for accommodating higher data rates. This increase in demand is only going to accelerate with the emergence of 5G. One of the greatest barriers in realizing this vision is the increasing amount of interference in wireless networks which has motivated researchers to find new interference management solutions.

One approach in mitigating the impact of interference is time and/or frequency division medium access scheme, also known as orthogonal access scheme, which divides the entire transmission signal among orthogonal dimensions. Therefore, each user can use a part of signaling dimension which has no overlap with other user signaling space. Another approach to improve channel spectral efficiency and achieve higher data rate is to provide full cooperation either among the transmitters or the receivers [1], [2]. As an example, employing full cooperation among the transmitters can lead to reducing the system to a single MIMO broadcast channel with higher channel capacity. But using the full cooperation needs joint processing and data sharing over distinct nodes. It seems to be infeasible in practical scenarios to provide full cooperation among transmitters and receivers. In recent years, the concept of interference alignment (IA) [3] has attracted a lot of attention. A popular metric when considering IA techniques, is degrees-of-freedom (DoF), or capacity pre-log, which characterizes the capacity behavior of wireless networks at high SNR regime. Due to the intuitive concept of DoF, many recent results focused on characterizing the DoF region of wireless networks. In [4], for the $K-$user fast fading IC, it was shown that IA can achieve sum DoF of $K/2$. This achievement theoretically shows a major improvement compared to baseline time or frequency division multiple access schemes where total achievable sum DoF is one. Although IA has a lot of benefits, there are a number of challenges and barriers in translating this method into a practical solution. The assumption of perfect CSI, and in some cases global perfect CSI for the $K-$user IC, the fast fading assumption, and long precoder lengths at the transmitters are some of these barriers [5], [6]. These challenges cast serious doubt about the practicality of IA methods.

In terms of reliance on CSI, we have four different types of interference alignment schemes: The first one is based on perfect channel state information (CSI) [7], [8]; the second one is instantaneous but imperfect CSI [8]–[10]; the third one is delayed CSI [11]–[15]; the last one is known as blind IA (BIA) [16], [17]. For the first three cases of (perfect, imperfect, and delayed) CSI, one of the main practical challenges is attaining high-resolution CSI cannot be realized (with acceptable overhead) in real-world networks [18]. The last one, the BIA, does not rely on CSI at the transmitters but typically assumes particular channel variation structures that may not be feasible either The key in resolving these issues is to reduce the reliance of IA techniques on the CSI. However, in the case of blind channel knowledge or even finite precision CSI, it is well known that the sum DoF of many networks collapse entirely to what is achievable by time or frequency sharing methods [19], [20]. However, it is important to note that this rather discouraging observation relies on particular assumptions about the networks such as identical distribution of channel parameters and availability of single omni-directional antennas at each wireless node. On the other hand, in [16], the authors show that if the channel coherence time complies with some specific structure, then, the benefits of IA method may be attained without accessing CSI at the transmitters. To control channel coherence times, one idea is to use multi-mode switching antennas at receivers [21], [22]. In this case, every receiver is equipped with an antenna that can switch among different reception modes. The next step is to design proper precoders and switching patterns at the transmitters and at the receivers, respectively. In [23], the author analyzes the network for the 3-user IC and shows that in the case of blind CSI using a re-configurable antenna at receivers, sum DoF of $\frac{6}{5}$ is achievable.

In [22] and with much more practical insights in [24], the authors propose a method to implement IA in a multipath fading environment. They suggest a new antenna structure that can find proper channel conditions to align most of interference signal power. But hardware changes in wireless systems is a slow and daunting task. So, it would be useful to find a solution at the transmitters to use the benefits of IA without CSI and hardware modification at the receivers.

In this paper, we show that by dividing the transmission block into distinct sub-blocks, identifying transceivers with "proper" channel states, and by deploying our proposed opportunistic IA strategy, we are able to align a considerable part of the interference signal power while keeping the desired signal in a separate linearly independent subspace. First, we introduce a non-causal strategy in which transmitters access channel variation pattern before transmission to demonstrate the potential gain of exploiting coherence time variations. However, as this result relies on non-causal information, it is only of theoretical interest and serves as a proof-of-concept. We also note that it might be difficult to observe the desired channel variation pattern in some wireless networks. To remedy these challenges and shortcomings, we propose a new antenna structure with high fluctuation rate in its transmission pattern gain. This antenna pattern structure helps us generate proper channel conditions with higher probability without adversely affecting the coverage for desired users. We show that by exploiting the statistics of channel variation pattern (either through the physics of the channel or by incorporating our proposed antenna), we can achieve an average sum rate similar to what IA techniques *with perfect CSI* promise. In our approach, transmitters use a multi-layer encoding strategy with different transmission rates for different layers, and receivers perform a IA stage followed by a successive decoding strategy to recover the desired messages. We show that if cross (interfering) links vary at a lower rate compared to direct links, the chance of finding conditions to achieve higher rates increases.

In the next section, we discuss system model, channel setup, and some preliminaries. In Section III, we provide a brief overview of our contributions. We present our proof-of-concept strategy alongside an important lemma in Section IV. We analyze the achievable DoF of the proposed IA scheme of Section IV in Section V. We present a new antenna hardware design in Section VI that enables a channel variation pattern which we exploit in Section VII to devise a blind interference alignment scheme in which the transmitters not only have no access to the CSI, but also are unaware of the channel variation pattern. In other words, using the antenna design of Section VI, we present a causal BIA scheme in Section VII. Finally, Section VIII concludes the paper.

II. System Model and Definitions

We consider the $K-$user IC which consists of K transmitters TX_k, and K receivers $\mathrm{RX}_k, k \in \mathcal{K} = \{1, \ldots, K\}$, each equipped with a single antenna. We will discuss a specific antenna structure later in this paper. The received signal at RX_i can be modeled as:

$$\bar{\mathbf{Y}}^{[i]} = \sum_{j=1}^{K} \bar{\mathbf{H}}^{[ij]} \bar{\mathbf{X}}^{[j]} + \bar{\mathbf{Z}}^{[i]}, i \in \mathcal{K}, \tag{1}$$

where the $n \times 1$ column vectors of $\bar{\mathbf{Y}}^{[i]}$ and $\bar{\mathbf{X}}^{[j]}$ indicate the received and transmitted signals at RX_i and TX_j, respectively. The $n \times 1$ column vectors of $\bar{\mathbf{Z}}^{[i]}$ indicates the additive white Gaussian noise (AWGN) which is distributed according to $\mathcal{CN}(0, 1)$. The matrix $\bar{\mathbf{H}}^{[ij]}$ is an $n \times n$ diagonal matrix which represents the channel matrix coefficients from TX_j to RX_i. To make the description easy to follow, we first discuss encoding and decoding, and then, discuss the channel structure in more detail.

A. Encoding, Decoding, and Achievable Rates

In this paper, we use a multi-layer encoding and decoding strategy in which transmitter TX_i wishes to send uniformly distributed message $W^{[i]} \in \mathcal{W}^{[i]} = \mathcal{W}_1^{[i]} \times \cdots \times \mathcal{W}_M^{[i]}$ to RX_i, $i \in \mathcal{K}$, over n uses of the channel. Each set $\mathcal{W}_j^{[i]}$, $1 \leq j \leq M$, represents the message set for the j^{th} transmission layer. We further assume that the messages are independent from each other and the channel gains. Each transmitter is subject to a total average transmission power constraint of P_t. Transmitter TX_i encodes its message $W^{[i]}$ using the encoding function $\bar{\mathbf{X}}^{[i]} = e_i\left(W^{[i]}, \text{SI}_{\mathsf{TX}_i}\right)$ where $i \in \mathcal{K}$ and $\text{SI}_{\mathsf{TX}_i}$ is the available side information at TX_i. The value of $R_j^{[i]} = \log_2 |\mathcal{W}_j^{[i]}|/n$ is the transmission rate of of j^{th} layer at TX_i. Therefore, the total transmission rate at TX_i is $R^{[i]} = \sum_{j=1}^{M} R_j^{[i]}$. We assume that each receiver RX_i is aware of its channel state information and decodes its intended message $W_j^{[i]} \in \mathcal{W}_j^{[i]}$ using the decoding function $\hat{W}_j^{[i]} = \phi_{ij}\left(\bar{\mathbf{Y}}^{[i]}, \text{SI}_{\mathsf{Rx}_i}\right)$, $1 \leq j \leq M$, where $\text{SI}_{\mathsf{Rx}_i}$ is the side information available to the receiver (in this case CSI). Then, the decoding error probability at receiver RX_i for the j^{th} layer is given by

$$\lambda_j^{[i]}(n) = \sum_{W_j^{[i]} \in \mathcal{W}_j^{[i]}} P\left(W_j^{[i]}\right) \Pr\left(\hat{W}_j^{[i]} \neq W_j^{[i]}\right) = \mathbb{E}\left[\Pr\left(\hat{W}_j^{[i]} \neq W_j^{[i]}\right)\right], \tag{2}$$

where the expectation is over the random choice of messages. For $\mathcal{J}_i \subseteq \{1, \ldots, M\}$, we define $R_{\mathcal{J}_i}^{[i]} = \sum_{j \in \mathcal{J}_i} R_j^{[i]}$. Then, rate-tuple $\left(R_{\mathcal{J}_1}^{[1]}, R_{\mathcal{J}_2}^{[2]}, \ldots, R_{\mathcal{J}_K}^{[K]}\right)$, is achievable if there exist encoders and decoders such that $\lambda_j^{[i]}(n) \to 0$, when $n \to \infty$ for all $i \in \mathcal{K}$, and $j \in \mathcal{J}_i$.

To compare our work with prior results and the total transmission power of ρ, we define the degrees-of-freedom (DoF) region, $\mathcal{D}$, to be the closure of all achievable $(d_1, \ldots, d_K)$ where $d_i = \lim_{\rho \to \infty} \frac{R_{\mathcal{J}_i}^{[i]}}{\log \rho}$. The sum DoF is defined as $d_{\text{sum}} = \max_{d_1, \ldots, d_K} \sum d_i, \; i \in \mathcal{K}$.

B. Channel Setup an Assumptions

We assume all channel matrices are diagonal represented by $\bar{\mathbf{H}}^{[ij]} = \text{diag}\left(\left[h_1^{[ij]}, \ldots, h_n^{[ij]}\right]\right)$ $i, j \in \mathcal{K}$. We assume a block fading model in time, where channel states are constant for a specific time duration. Therefore, for the channel matrix of $\bar{\mathbf{H}}^{[ij]}$, we have:

$$h_{c_l^{[ij]}}^{[ij]} = \ldots = h_{c_{l+1}^{[ij]}-1}^{[ij]} \tag{3}$$

where $l \in \{1, \ldots, \eta(i, j)\}$ and $\eta(i, j)$ is the number of channel "altering points" in time between TX_j and RX_i. The value of $c_l^{[ij]}, \; l \in \{1, \ldots, \eta(i, j)\}$ represents the l^{th} point of altering state of the channel between TX_j and RX_i and as n goes to infinity $\frac{\eta(i,j)}{n}$ represents channel variation rate. It is assumed that all $h_l^{[ij]}$ are $i.i.d$ random variables with a specific distribution whose magnitude are bounded between a nonzero and a finite maximum value. Since the channel coherence time is a random variable, we can model each signaling channel similar to the state diagram which is shown in Fig. 1. At each altering point, the channel remains in its previous state with the probability of p_{ij} and changes with the probability of $1 - p_{ij}$.

We will need the following definitions in this paper.

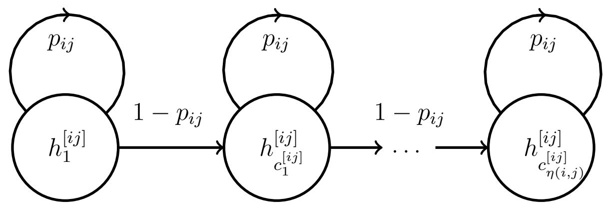


Fig. 1: The state diagram of the channel between TX_j and RX_i. This diagram is consisted of $\eta\left(i,j\right)$ states where each state has been shown by a circle.

Definition 1: We define $\mathcal{C}^{[ij]} = \left\{c_1^{[ij]}, \ldots, c_{\eta(i,j)}^{[ij]}\right\}$ *to be the changing pattern of the channel between* TX_j *and* RX_i*.*

Definition 2: We represent the collection of all sets $\mathcal{C}^{[ii]},\ i \in \mathcal{K}_r \subseteq \mathcal{K}$*, by* $\mathcal{C}_{\mathcal{K}_r}$ *as follows:*

$$\mathcal{C}_{\mathcal{K}_r} = \bigcup \mathcal{C}^{[ii]},\ i \in \mathcal{K}_r \subseteq \mathcal{K}. \tag{4}$$

The extended form of the set $\mathcal{C}_{\mathcal{K}_r}$ can be written as follows:

$$\mathcal{C}_{\mathcal{K}_r} = \left\{c_1^{\mathcal{K}_r}, c_2^{\mathcal{K}_r}, \ldots, c_{\sigma_{\mathcal{K}_r}}^{\mathcal{K}_r}\right\},\ |\mathcal{C}_{\mathcal{K}_r}| = \sigma_{\mathcal{K}_r}, \tag{5}$$

where $|\mathcal{A}|$ indicates the cardinality of set $\mathcal{A}$. During the transmission time, the value of $\sigma_{\mathcal{K}_r}$ indicates the total number of altering points of the direct channels in $\mathcal{K}_r$.

Definition 3: We define subset $\mathcal{C}'_{\mathcal{K}_r}$ *as follows:*

$$\mathcal{C}'_{\mathcal{K}_r} = \bigcup \mathcal{C}^{[ij]},\ i \neq j \in \mathcal{K}_r \subseteq \mathcal{K}, \tag{6}$$

where $\mathcal{C}'_{\mathcal{K}_r}$ indicates the collection of changing pattern sets of all cross links with indices in $\mathcal{K}_r$. Similar to (5), we can write the extended form of the set $\mathcal{C}'_{\mathcal{K}_r}$ as follows:

$$\mathcal{C}'_{\mathcal{K}_r} = \left\{c_1'^{\mathcal{K}_r}, c_2'^{\mathcal{K}_r}, \ldots, c_{\sigma'_{\mathcal{K}_r}}'^{\mathcal{K}_r}\right\},\ |\mathcal{C}'_{\mathcal{K}_r}| = \sigma'_{\mathcal{K}_r} \tag{7}$$

During the transmission time, the value of $\sigma'_{\mathcal{K}_r}$ indicates the total number of altering points of cross channels in $\mathcal{K}_r$.

Definition 4: We define the set $\mathcal{B}_t^{\mathcal{K}_r}, t \in \mathbb{N}$*, as follows:*

$$\mathcal{B}_t^{\mathcal{K}_r} = \left\{i | i \in [c_{t-1}'^{\mathcal{K}_r} : c_t'^{\mathcal{K}_r} - 1], i \in \mathbb{N}\right\}. \tag{8}$$

*This set indicates the t^{th} transmission block time snapshots in which all the cross links among the transceivers with indices in $\mathcal{K}_r$ have constant values (*i.e. *without altering state).*

Definition 5: We define the set $\mathcal{C}^{[ij]}_{\mathcal{B}^{\mathcal{K}_r}_t}, t \in \mathbb{N}$, as follows:

$$\mathcal{C}^{[ij]}_{\mathcal{B}^{\mathcal{K}_r}_t} = \mathcal{C}^{[ij]} \cap \mathcal{B}^{\mathcal{K}_r}_t. \tag{9}$$

Definition 6: We define the function $\mathcal{F}\left(\bar{\mathbf{H}}^{[ij]}\right)$ as the maximum number of the time snapshots in which the channel matrix between TX_j and RX_i is constant, or mathematically we have:

$$\mathcal{F}\left(\bar{\mathbf{H}}^{[ij]}\right) = \max_{l} c^{[ij]}_{l} - c^{[ij]}_{l-1},\ l \in \{1, \dots, \eta(i,j)+1\} \tag{10}$$

where we should note that $c^{[ij]}_0 = 1$ and $c^{[ij]}_{\eta(i,j)+1} = n+1$.

Note: $\mathcal{K}_r$ shows a subset of $\mathcal{K}$ except $\varnothing$. Therefore, the value of r is between one and $2^K - 1$.

Table 1 is an example for a 3-user IC. In this table, we illustrate how the sets we defined above, *i.e.* $\mathcal{C}^{[ij]}$, $\mathcal{C}'$, $\mathcal{B}^{\mathcal{K}}_m$ and $\mathcal{C}^{[33]}_{\mathcal{B}^{\mathcal{K}}_m}$, are created from the channel matrices $\bar{\mathbf{H}}^{[ij]},\ i,j \in \mathcal{K}$.

III. Overview of the Contributions

Before presenting the results in great details, in this section, we provide a brief overview of our contributions. This paper provides answers to the following questions:

1) **Is it possible to exploit coherence time variations for interference management?** To answer this question, we provide a proof-of-concept interference alignment strategy in Section IV that demonstrates theoretically it is possible to obtain similar gains to prior IA results. The important observation is that the only information transmitters need is the coherence time variations, and no other channel state information is needed. However, as stated, our scheme relies on *non-causal* knowledge of coherence time variations, and as such, only serves as a proof-of-concept. We address this challenge in other parts of the paper as discussed below.

2) **Can the assumed channel structure be realized in real-world?** We assume a particular channel structure in Section II. While this model is motivated by the physics of wireless systems, it is not clear how closely a typical network follows this model. To address such concerns, we propose a transmitter antenna design in Section VI that creates such coherence time fluctuations even for stationary users. In particular, using this design, we can create a scenario in which direct links vary faster than cross (interfering) links which enables us to provide a causal solution later.

3) **Is it possible to exploit coherence time variations causally for interference management?** We capitalize on our antenna design and the channel structure it creates to show how one can

TABLE I: Example of Channel Structure for 3-user IC

Channel matrices $\bar{\mathbf{H}}^{[ij]},\ i,j \in \mathcal{K}$	$\mathcal{C}^{[ij]}$	$\mathcal{C}'_{\mathcal{K}}$	$\mathcal{B}_t^{\mathcal{K}}, 1 \leq t \leq 5$	$\mathcal{C}^{[33]}_{\mathcal{B}_t^{\mathcal{K}}}$
$\bar{\mathbf{H}}^{[11]} = \text{diag}\left([1.5, 1.1, 2.3, 2.3, 0.6, 0.6, 3.0, 3.0, 1.0, 1.0]\right)$	$\mathcal{C}^{[11]} = \{2,3,5,7,9\}$			
$\bar{\mathbf{H}}^{[12]} = \text{diag}\left([1.2, 1.2, 1.2, 1.2, 1.5, 1.5, 1.5, 1.0, 1.0, 1.0]\right)$	$\mathcal{C}^{[12]} = \{5,8\}$		$\mathcal{B}_1^{\mathcal{K}} = \{1,2,3,4\}$	$\mathcal{C}^{[33]}_{\mathcal{B}_1^{\mathcal{K}}} = \{2,3\}$
$\bar{\mathbf{H}}^{[13]} = \text{diag}\left([1.0, 1.0, 1.0, 1.0, 0.8, 0.8, 1.5, 3.0, 3.0, 3.0]\right)$	$\mathcal{C}^{[13]} = \{5,7,8\}$		$\mathcal{B}_2^{\mathcal{K}} = \{5,6\}$	$\mathcal{C}^{[33]}_{\mathcal{B}_2^{\mathcal{K}}} = \{6\}$
$\bar{\mathbf{H}}^{[21]} = \text{diag}\left([2.0, 2.0, 2.0, 2.0, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3]\right)$	$\mathcal{C}^{[21]} = \{5\}$		$\mathcal{B}_3^{\mathcal{K}} = \{7\}$	$\mathcal{C}^{[33]}_{\mathcal{B}_3^{\mathcal{K}}} = \varnothing$
$\bar{\mathbf{H}}^{[22]} = \text{diag}\left([2.5, 2.1, 2.1, 2.1, 1.1, 1.1, 3.1, 1.1, 1.1, 1.1]\right)$	$\mathcal{C}^{[22]} = \{2,5,7,8\}$	$\{5,7,8,10\}$	$\mathcal{B}_4^{\mathcal{K}} = \{8,9\}$	$\mathcal{C}^{[33]}_{\mathcal{B}_4^{\mathcal{K}}} = \{9\}$
$\bar{\mathbf{H}}^{[23]} = \text{diag}\left([0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 3.0, 3.0, 3.0, 3.0]\right)$	$\mathcal{C}^{[23]} = \{7\}$		$\mathcal{B}_5^{\mathcal{K}} = \{10\}$	$\mathcal{C}^{[33]}_{\mathcal{B}_5^{\mathcal{K}}} = \varnothing$
$\bar{\mathbf{H}}^{[31]} = \text{diag}\left([0.8, 0.8, 0.8, 0.8, 0.4, 0.4, 0.4, 0.4, 0.4, 0.5]\right)$	$\mathcal{C}^{[31]} = \{5,10\}$		$\mathcal{B}_6^{\mathcal{K}} = \varnothing$	$\mathcal{C}^{[33]}_{\mathcal{B}_6^{\mathcal{K}}} = \varnothing$
$\bar{\mathbf{H}}^{[32]} = \text{diag}\left([0.2, 0.2, 0.2, 0.2, 0.3, 0.3, 1.2, 1.2, 1.2, 1.2]\right)$	$\mathcal{C}^{[32]} = \{5,7\}$			
$\bar{\mathbf{H}}^{[33]} = \text{diag}\left([0.1, 0.2, 0.2, 0.3, 0.3, 1.7, 1.7, 1.7, 1.3, 1.3]\right)$	$\mathcal{C}^{[33]} = \{2,4,6,9\}$			

align most of the interference signal power even without accessing channel coherence time variations. More precisely, we devise a new blind interference alignment strategy in Section VII and evaluate the resulting achievable rates. This result can be thought of as the generalization of the BIA technique of [16] to larger networks with further practical considerations.

IV. Non-causal Transmission Strategy: A Proof-of-Concept

In this section, exploiting coherence time variations and channel changing pattern, we propose a new methodology for data transmission. In Section II, we defined multi-layer encoding with M layers, but in this section, we assume a single layer per transmitter layer ($M = 1$), and thus, all definitions fall within the conventional settings of prior results. We note that we cannot use the method presented in this section without accessing channel changing pattern before transmitting data, but nonetheless, it demonstrates the potential benefit of exploiting channel changing patterns for interference alignment. Moreover, the results in this section provide us with intuition and ideas to design our practical causal solution in Section VII. Since we will be using linear coding, we can present an alternative representation of the transmit signal. The output of the encoding function, $\bar{\mathbf{X}}^{[j]}$, can be written as an $n \times 1$ column vector with the following relation:

$$\bar{\mathbf{X}}^{[j]} = \sum_{d=1}^{d_j} x_d^{[j]} \mathbf{v}_d^{[j]} = \bar{\mathbf{V}}^{[j]} \mathbf{x}^{[j]}, \tag{11}$$

where d_j is the number of symbols transmitted by TX_j over n channel uses, $x_d^{[j]}$ is the d^{th} transmitted symbol, and $\mathbf{v}_d{}^{[j]}$ is an $n \times 1$ transmit beamforming vector for the d^{th} symbol. Also, $\mathbf{x}^{[j]} = \left[x_1^{[j]}, \ldots, x_{d_j}^{[j]}\right]^T$ is a $d_j \times 1$ column vector that represents the transmitted data symbols, and $\bar{\mathbf{V}}^{[j]} = [\mathbf{v}_1^{[j]} \ \mathbf{v}_2^{[j]} \ ... \ \mathbf{v}_{d_j}^{[j]}]$ is the $n \times d_j$ precoder matrix. For the precoder matrix $\bar{\mathbf{V}}^{[j]}$, vector $\mathbf{v}_d^{[j]}$ indicates one of the basis vectors of the designed precoder at TX_j. So, based on presented setup, and before presenting the transmission strategy, we state the following lemma which plays an essential role in our results.

Lemma 1: Suppose $\mathbf{V}$ *is a random matrix of size* $n \times d_v$ *with nonzero elements and rank* d_v*. Let* $\bar{\mathbf{H}}$ *be a random diagonal matrix of size* $n \times n$*. Then, we have:*

$$\operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right) = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right), \tag{12}$$

where $\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]$ *is the matrix obtained by concatenating* $\mathbf{V}$ *and* $\bar{\mathbf{H}}\mathbf{V}$*.*

Proof: Consider $t_1, \ t_2 \in \mathbb{N}$ $(t_1, t_2 \leq n)$, are two time snapshots in which the random matrix $\bar{\mathbf{H}}$ from t_1-th time snapshot to t_2-th time snapshot has the constant value of α and $\mathcal{F}\left(\bar{\mathbf{H}}\right) = t_2 - t_1 + 1$. For the fast fading channel model of [3], $t_1 = t_2$ and the value of $\mathcal{F}\left(\bar{\mathbf{H}}\right)$ is one. Then, the rank of $\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]$ can be upper bounded as follows:

$$\operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right) = \operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V} - \alpha\bar{\mathbf{I}}\mathbf{V}\right]\right) \leq \underbrace{\operatorname{rank}\left(\mathbf{V}\right)}_{=d_v} + \operatorname{rank}\left(\bar{\mathbf{H}}\mathbf{V} - \alpha\bar{\mathbf{I}}\mathbf{V}\right), \tag{13}$$

where $\alpha\bar{\mathbf{I}}$ is a square diagonal matrix with all its diagonal entries equal to α. Since matrix $\bar{\mathbf{H}}\mathbf{V} - \alpha\bar{\mathbf{I}}\mathbf{V}$ has $n - \mathcal{F}\left(\bar{\mathbf{H}}\right)$ nonzero rows, we can upper bound the rank of matrix $\bar{\mathbf{H}}\mathbf{V} - \alpha\bar{\mathbf{I}}\mathbf{V}$ by $\min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right)$. The rank of $\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right)$ cannot be greater than n, therefore from (13), we get:

$$\operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right) \leq \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right). \tag{14}$$

We note that when $\bar{\mathbf{H}}$ has one state, *i.e.* $\mathcal{F}\left(\bar{\mathbf{H}}\right) = n$, and thus, $n - \mathcal{F}\left(\bar{\mathbf{H}}\right) = 0$, the rank of $\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]$ is equivalent to the rank of $\mathbf{V}$. Subsequently, from (14), we have:

$$\operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right) = d_v = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right). \tag{15}$$

So, when $\bar{\mathbf{H}}$ has one state the proof has been completed. For the case in which $\bar{\mathbf{H}}$ has two states, α_1 and α_2, the rank of $\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]$ can be simplified as follows:

$$\begin{aligned}
&\operatorname{rank}\left(\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} \mathbf{V}_1 & \alpha_1\mathbf{V}_1 \\ \mathbf{V}_2 & \alpha_2\mathbf{V}_2 \end{bmatrix}\right) \stackrel{(a)}{=} \operatorname{rank}\left(\begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{V}_2 & (\alpha_2 - \alpha_1)\mathbf{V}_2 \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} \mathbf{V}_1 & \mathbf{0} \\ \mathbf{0}' & \mathbf{V}_2 \end{bmatrix}\right)
\end{aligned} \tag{16}$$

$$\begin{aligned}
&= \min\left(\operatorname{rank}\left(\mathbf{V}_1\right) + \operatorname{rank}\left(\mathbf{V}_2\right), n\right) \stackrel{(b)}{=} \min\left(\min\left(n - \mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right) + \min\left(\mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right), n\right) \\
&\stackrel{(c)}{=} \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right),
\end{aligned} \tag{17}$$

where $\mathbf{0}$ and $\mathbf{0}'$ are all zero matrices with the same size as $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively. Equality (a) comes from subtracting α_1 times of the first d_v columns of matrix $\left[\mathbf{V} \ \bar{\mathbf{H}}\mathbf{V}\right]$ from the last d_v columns. Equality (b) comes from the fact that $\mathbf{V}_1$ is of size $\left(n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right) \times d_v$ and the size of $\mathbf{V}_2$

is $\mathcal{F}\left(\bar{\mathbf{H}}\right) \times d_v$. Equality (c) comes from the fact that if $\mathcal{F}\left(\bar{\mathbf{H}}\right) \leq d_v$, then, $\min\left(\mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right) = \mathcal{F}\left(\bar{\mathbf{H}}\right)$, and from Definition 6, since $\mathcal{F}\left(\bar{\mathbf{H}}\right) \geq n - \mathcal{F}\left(\bar{\mathbf{H}}\right)$, the value of $\min\left(n - \mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right) = n - \mathcal{F}\left(\bar{\mathbf{H}}\right)$. Subsequently, in this case, $\min\left(\min\left(n - \mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right) + \min\left(\mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right), n\right) = n = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right)$. If $\mathcal{F}\left(\bar{\mathbf{H}}\right) > d_v$, then:

$$\min\left(\min\left(\mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right) + \min\left(n - \mathcal{F}\left(\bar{\mathbf{H}}\right), d_v\right), n\right) = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right). \tag{18}$$

So, we can conclude that when the number of states in $\bar{\mathbf{H}}$ is one or two, the result of the first lemma holds. Now, we analyze three different cases as follows:

Case 1: Let $\mathcal{F}\left(\bar{\mathbf{H}}\right) \geq \frac{n}{2}$, by basic algebraic operations on $\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]$, and displacing the rows in the block with the length of $\mathcal{F}\left(\bar{\mathbf{H}}\right)$ to the first rows. We have:

$$\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) = \operatorname{rank}\left(\left[\mathbf{V}'\ \bar{\mathbf{H}}'\mathbf{V}'\right]\right). \tag{19}$$

where $\left[\mathbf{V}'\ \bar{\mathbf{H}}'\mathbf{V}'\right]$ is generated from $\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]$ through basic algebraic operations (displaced rows matrix). The column matrix $\bar{\mathbf{H}}'$ can be equivalently expressed by multiplying two matrices $\bar{\mathbf{H}}'_1$ and $\bar{\mathbf{H}}'_2$ ($\bar{\mathbf{H}}' = \bar{\mathbf{H}}'_1\bar{\mathbf{H}}'_2$) where $\bar{\mathbf{H}}'_2$ has two states and $\mathcal{F}\left(\bar{\mathbf{H}}'_2\right) = \mathcal{F}\left(\bar{\mathbf{H}}'\right) = \mathcal{F}\left(\bar{\mathbf{H}}\right)$. So:

$$\operatorname{rank}\left(\left[\mathbf{V}'\ \bar{\mathbf{H}}'\mathbf{V}'\right]\right) = \operatorname{rank}\left(\left[\mathbf{V}'\ \bar{\mathbf{H}}'_1\bar{\mathbf{H}}'_2\mathbf{V}'\right]\right) \overset{(a)}{\geq} \operatorname{rank}\left(\left[\mathbf{V}'\ \bar{\mathbf{H}}'_2\mathbf{V}'\right]\right) \tag{20}$$

$$\overset{(b)}{=} \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}'_2\right)\right), n\right) \overset{(c)}{=} \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right), \tag{21}$$

where (a) holds since $\bar{\mathbf{H}}'_1$ has random elements; similar to (16)-(17), (b) comes from the fact that $\bar{\mathbf{H}}'_2$ has two states; (c) is true since $\mathcal{F}\left(\bar{\mathbf{H}}\right) = \mathcal{F}\left(\bar{\mathbf{H}}'_2\right)$. On other hand, in (14), we show that $\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) \leq \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right)$. So, from (14), we can conclude that $\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right)$, and in this case the proof is complete.

Case 2: Let $\mathcal{F}\left(\bar{\mathbf{H}}\right) < \frac{n}{2}$ and $d_v \geq \frac{n}{2}$. Now, assume $\bar{\mathbf{H}}'$ is a matrix with $\frac{n}{2} \leq \mathcal{F}\left(\bar{\mathbf{H}}'\right) \leq \frac{n+1}{2}$ and two states. In this case, we have:

$$\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) \overset{(a)}{\geq} \operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}'\mathbf{V}\right]\right) \overset{(b)}{=} \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}'\right)\right), n\right) \overset{(c)}{=} n, \tag{22}$$

where (a) holds since $\mathcal{F}\left(\bar{\mathbf{H}}\right) \leq \mathcal{F}\left(\bar{\mathbf{H}}'\right)$; (b) is justified through (17); (c) follows from the following inequality $d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}'\right)\right) > n$. On other hand, $\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) \leq n$. Thus, we have $\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) = n$. Therefore, in this case, the proof has also been completed.

Case 3: Let $\mathcal{F}\left(\bar{\mathbf{H}}\right) < \frac{n}{2}$ and $d_v < \frac{n}{2}$. Similar to the previous case, assume $\bar{\mathbf{H}}'$ is a matrix with $\frac{n}{2} \leq \mathcal{F}\left(\bar{\mathbf{H}}'\right) \leq \frac{n+1}{2}$ and two states. In this case, we have:

$$\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) \geq \operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}'\mathbf{V}\right]\right) \overset{(a)}{=} \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}'\right)\right), n\right) \overset{(b)}{=} 2d_v. \tag{23}$$

where (a) follows from (17), and (b) holds since $n - \mathcal{F}\left(\bar{\mathbf{H}}'\right) > d_v$. Also, from (14), we have

$$\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) \leq \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right) = 2d_v. \tag{24}$$

We can conclude that $\operatorname{rank}\left(\left[\mathbf{V}\ \bar{\mathbf{H}}\mathbf{V}\right]\right) = 2d_v$. So, in this case, the proof has been completed.

Since these three cases cover all possibilities, the proof of Lemma 1 follows immediately. ■

A. Motivating Example: Transmission Strategy for a 3-user Interference Channel

Consider Table I in which we have 6 cross channels. In this case, we design precoder matrices at different transmitters according to the following steps:

Step 1: Finding all the sets of $\mathcal{C}'_{\mathcal{K}_r}$ from the third definition with $2 < |\mathcal{K}_r| \leq |\mathcal{K}|$ as follows:

$$\mathcal{C}'_{\mathcal{K}} = \bigcup_{i \neq j} \mathcal{C}^{[ij]} = \{5, 7, 8, 10\}. \; i, j \in \mathcal{K} \tag{25}$$

In this example, since since we have three users, $\mathcal{K}_r = \mathcal{K}$.

Step 2: From definition 4, finding all the sets of $\mathcal{B}_t^{\mathcal{K}_r}$ for $2 < |\mathcal{K}_r| \leq |\mathcal{K}|$ as follows:

$$\mathcal{B}_1^{\mathcal{K}} = \{1, 2, 3, 4\}, \; \mathcal{B}_2^{\mathcal{K}} = \{5, 6\}, \; \mathcal{B}_3^{\mathcal{K}} = \{7\}, \; \mathcal{B}_4^{\mathcal{K}} = \{8, 9\}, \; \mathcal{B}_5^{\mathcal{K}} = \{10\}. \tag{26}$$

Considering above sets, our transmission time is divided among 5 different transmission blocks. As an example, the first and second transmission blocks are consisted of 4 and 2 time snapshots.

Step 3: In each transmission block, we design all the precoder matrices as follows:

$$\bar{\mathbf{V}}^{[1]}(t) = \left[\mathbf{v}_1, \ldots, \mathbf{v}_{d_1(t)}\right], \; \bar{\mathbf{V}}^{[2]}(t) = \left[\mathbf{v}_1, \ldots, \mathbf{v}_{d_2(t)}\right], \; \bar{\mathbf{V}}^{[3]}(t) = \left[\mathbf{v}_1, \ldots, \mathbf{v}_{d_3(t)}\right], \tag{27}$$

where t indicates the block number. All the basis vectors $\mathbf{v}_j$ at t^{th} transmission block are nonzero random column vectors with the size of $|\mathcal{B}_t^{\mathcal{K}}| \times 1$. In this scheme, all the basis vectors at different transmitters with the same indexes are equivalent and all the basis vectors with different indexes are linearly independent.

Step 4: In each transmission block, we try to find 3-tuple $(d_1(t), d_2(t), d_3(t))\,, 1 \leq t \leq 5$, such that the sum DoF of $d_{\text{sum}} = d_1(t) + d_2(t) + d_3(t)$ at each transmission block of t, is maximized subject to the two following constraints:

$$d_i(t) + \max_{j=1:3, j \neq i} d_j(t) \overset{(a)}{\leq} |\mathcal{B}_t^{\mathcal{K}}|, \; d_i(t) \overset{(b)}{\leq} |\mathcal{B}_t^{\mathcal{K}}| - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[ii]}\right), \tag{28}$$

where $\bar{\mathbf{H}}_t^{[ii]}$ represents a diagonal matrix with elements extracted from the time snapshots of t^{th} block of the diagonal elements of the matrix $\bar{\mathbf{H}}^{[ii]}$. Inequality (a) holds since at RX_i in each transmission block of $\mathcal{B}_t^{\mathcal{K}}$ all the cross channels are diagonal matrices with the constant diagonal elements, all the interference signals are aligned and the space spanned by the interference signals has the dimension of $\max_{j \neq i}\left(\mathrm{rank}\left(\bar{\mathbf{V}}^{[j]}(t)\right)\right) = \max_{j \neq i} d_j(t)$. So, the desired and interference signal dimensions should not be greater than the number of time snapshots of each transmission block or $|\mathcal{B}_t^{\mathcal{K}}|$. Inequality (b) comes from the first lemma, which indicates that the number of desired signal at its corresponding receiver should not be greater than $|\mathcal{B}_t^{\mathcal{K}}| - \mathcal{F}\left(\bar{\mathbf{H}}^{[ii]}\right)$; otherwise, the desired signal space can not be linearly independent from interference signals.

So, for our specific problem at the first transmission block ($t = 1$) and $|\mathcal{B}_1^{\mathcal{K}}| = 4$, we can upper bound the values of $d_i(1), i \in \{1,2,3\}$ as $d_i(1) \leq |\mathcal{B}_1^{\mathcal{K}}| - \mathcal{F}\left(\bar{\mathbf{H}}_1^{[ii]}\right)$, where indicates that, $d_1(1) \leq 2$, $d_2(1) \leq 1$ and $d_3(1) \leq 2$. Setting $(d_1(1), d_2(1), d_3(1)) = (2,1,2)$, we can show that the second constrain in (28) has been satisfied.

Remark 1: For those blocks with $|\mathcal{K}_r| \leq 2$, since we cannot use the benefits of the IA scheme (we have one or two transceivers and there is no chance to align signals), we can choose one or two random transmitters TX_{i_1} and TX_{i_2} to set the 2-tuples $(d_{i_1}(t), d_{i_2}(t))$ as $d_{i_1}(t) + d_{i_2}(t) = 1$.

Step 5: Finding proper precoders to satisfy previous step constrains of (28). As an example for our specific problem, at the first transmission block we design $\bar{\mathbf{V}}^{[1]}(1)$, $\bar{\mathbf{V}}^{[2]}(1)$ and $\bar{\mathbf{V}}^{[3]}(1)$ such that:

$$\bar{\mathbf{V}}^{[1]}(1) = [\mathbf{v}_1, \mathbf{v}_2]\,,\ \bar{\mathbf{V}}^{[2]}(1) = [\mathbf{v}_1]\,,\ \bar{\mathbf{V}}^{[3]}(1) = [\mathbf{v}_1, \mathbf{v}_2]\,, \tag{29}$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are two column vectors with random elements and the size of $|\mathcal{B}_1^{\mathcal{K}}| \times 1$. Now at the first transmission block let's analyze the IA conditions at different receivers as follows:

$$\begin{aligned}
\text{at RX}_1\text{: span}\left(\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}(1)\right) &\in \text{span}\left(\bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}(1)\right)\\
\text{rank}\left(\left[\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]}(1)\ \bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}(1)\right]\right) &= 4,\\
\text{at RX}_2\text{: span}\left(\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}(1)\right) &\in \text{span}\left(\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}(1)\right)\\
\text{rank}\left(\left[\bar{\mathbf{H}}^{[22]}\bar{\mathbf{V}}^{[2]}(1)\ \bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}(1)\right]\right) &= 4,\\
\text{at RX}_3\text{: span}\left(\bar{\mathbf{H}}^{[32]}\bar{\mathbf{V}}^{[2]}(1)\right) &\in \text{span}\left(\bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[1]}(1)\right)\\
\text{rank}\left(\left[\bar{\mathbf{H}}^{[33]}\bar{\mathbf{V}}^{[3]}(1)\ \bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[1]}(1)\right]\right) &= 4.
\end{aligned} \tag{30}$$

So, at the first transmission block the first and third users get 2 resources while the second user gets 1 resource. Therefore, at the fist transmission block we can achieve 5 resources from 4 transmission snapshots. Consequently, at the first transmission block we get $\frac{5}{4}$ sum DoF. From the first remark, for other transmission blocks $|\mathcal{B}_t^{\mathcal{K}}| \leq 2,\ t > 1$, we cannot use the benefits of IA scheme. So, in this example the achievable sum DoF can be calculated as follows:

$$\frac{d_{\text{sum}}}{n} = \frac{\sum_{t=1}^{5}\sum_{i=1}^{3} d_i(t)}{n} = \frac{5+2+1+2+1}{10} = \frac{11}{10}. \tag{31}$$

Therefore, without accessing channel coefficients we can achieve $d_{\text{sum}} > 1$.

B. A search algorithm to maximize d_{sum}:

As a standard search algorithm for the problem of IA using channel changing pattern, in this subsection, we generalize the solution of previous example. For $K = 2$, the optimum solution is easily obtained by setting $d_{\text{sum}} = 1$. The interesting problem is when $K > 2$. In the general case we divide the transmission time (consisting of n time snapshots) among different transmission blocks and in each transmission block we activate a subset of transmitters $\mathcal{K}_r$.

Definition 7: We define the set $\mathcal{G}_u$ as follows:

$$\mathcal{G}_u = \{\mathcal{K}_{u_1}, \mathcal{K}_{u_2}, \ldots, \mathcal{K}_{u_{l_u}}\}\ \mathcal{K}_{u_t} \subseteq \mathcal{K} = \{1, \ldots, K\}, 1 \leq t \leq l_u. \tag{32}$$

In our proposed method, depending on the set $\mathcal{G}_u$, the transmission time is divided into $|\mathcal{G}_u| = l_u$ distinct blocks and each block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}, 1 \leq t \leq l_u$ contains $|\mathcal{B}_t^{\mathcal{K}_{u_t}}| \leq n$ time snapshots. In each transmission block we activate a subset of transmitters $\mathcal{T}^{(u_t)} = \{\mathrm{TX}_{k_1^{u_t}}, \ldots, \mathrm{TX}_{k_r^{u_t}}\},\ 1 \leq r \leq K$ to transmit their messages to their corresponding receivers. The set of active transceivers $\{k_1^{u_t}, \ldots, k_r^{u_t}\}$ in the transmission block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}, 1 \leq t \leq l_u$ is equivalent to the set of $\mathcal{K}_{u_t}$ in $\mathcal{G}_u$. In other words, $\mathcal{K}_{u_t} = \{k_1^{u_t}, \ldots, k_r^{u_t}\}$ shows the indexes of the active transceivers at the transmission block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}, 1 \leq t \leq l_u$. The transmission block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$ for every set of $\mathcal{G}_u$ is defined as follows:

$$\mathcal{B}_1^{\mathcal{K}_{u_1}} = \left\{i | i \in [1:m], m \notin \mathcal{C}'_{\mathcal{K}_{u_1}}\right\},\ \mathcal{B}_t^{\mathcal{K}_{u_t}} = \left\{i | i \in \left[\max\left(\mathcal{B}_{t-1}^{\mathcal{K}_{u_{t-1}}}\right) + 1 : m\right], m \notin \mathcal{C}'_{\mathcal{K}_{u_t}}\right\}, \tag{33}$$

where all the members of the set $\mathcal{B}_t^{\mathcal{K}_{u_t}}$ are consecutive. So, $\max \mathcal{B}_t^{\mathcal{K}_{u_t}} - \min \mathcal{B}_t^{\mathcal{K}_{u_t}} = |\mathcal{B}_t^{\mathcal{K}_{u_t}}| - 1$.

As an example consider we have $K = 4$ and $\mathcal{G}_1 = \{\{1, 3, 4\}, \{1, 2, 3, 4\}\}$ where $u = 1$. In this case we have two transmission blocks of $\mathcal{B}_1^{\mathcal{K}_{1_1}}$ and $\mathcal{B}_2^{\mathcal{K}_{1_2}}$ which indicate that at the first and the second transmission block the active transmitters have been selected from $\mathcal{T}^{1_1} = \{\mathrm{TX}_1, \mathrm{TX}_3, \mathrm{TX}_4\}$ and $\mathcal{T}^{1_2} = \{\mathrm{TX}_1, \mathrm{TX}_2, \mathrm{TX}_3, \mathrm{TX}_4\}$, respectively. Similarly, if $\mathcal{G}_2 = \{\{3, 4\}, \{1, 2, 4\}, \{1\}\}$, since $|\mathcal{G}_2| = 3$, we have 3 transmission block. At the first transmission block we have 2 active transmitters of TX_3 and TX_4, at the second transmission block we have 3 active transmitters of $\mathrm{TX}_i, i = 1, 2, 4$ and at the third transmission block we only have one active transmitter of TX_1.

Generally for the transmission time consisted of n time snapshots, the set $\mathcal{G}_u$ with $|\mathcal{G}_u| \leq n$ can be chosen from n^Q different sets where $Q \overset{(a)}{=} 2^K - 1$. The equality of (a) comes from this fact that in each transmission snapshot we can activate a subset of transmitters from the set of $\mathcal{K}$ except $\varnothing$, consequently, we should have, $u \leq n^Q$.

Remark 2: From definition 4, the length of each transmission block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$ should be selected such that all the cross-links among transceivers with the indexes in the set of $\mathcal{K}_{u_t}$ have not any altering point. In other words, there is not any changing point in t^{th} transmission block of cross channels among active users.

The maximum value of d_{sum}, mathematically can be calculated from the following optimization problem:

$$\max_{\mathcal{G}_u, d_i(t)} \sum_{t=1}^{|\mathcal{G}_u|} \sum_i d_i(t),\ 1 \leq u \leq n^Q, i \in \mathcal{K}_{u_t}, \tag{34}$$

subject to the following constraints at each block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$:

$$d_i(t) + \max_{j \in \mathcal{K}_{u_t}, j \neq i} d_j(t) \leq |\mathcal{B}_t^{\mathcal{K}_{u_t}}| \text{ and } d_i(t) \leq |\mathcal{B}_t^{\mathcal{K}_{u_t}}| - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[ii]}\right), \tag{35}$$

Lemma 2: For the $K-$user IC in a random block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$, there exists $\bar{\mathbf{V}}^{[k_1^{u_t}]}, \ldots, \bar{\mathbf{V}}^{[k_r^{u_t}]}$ such that all the conditions of (35) *are satisfied.*

Proof: We provide the proof by finding proper precoder matrices at transmitters. Let at the block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$ we have:

$$\bar{\mathbf{V}}^{[k_1^{u_t}]}(t) = \left[\mathbf{v}_1 \ \ldots \ \mathbf{v}_{d_{k_1^{u_t}}}\right], \ldots, \bar{\mathbf{V}}^{[k_r^{u_t}]}(t) = \left[\mathbf{v}_1 \ \ldots \ \mathbf{v}_{d_{k_r^{u_t}}}\right], \tag{36}$$

where $\mathbf{v}_j$ is a random vector with nonzero elements and the size of $|\mathcal{B}_t^{\mathcal{K}_{u_t}}| \times 1$. Now, assume $\bar{\mathbf{V}}^{[k_m^{u_t}]}(t)$ is a precoder matrix with the maximum rank of $d_{k_m^{u_t}}$ among the active transmitters with indices in $\mathcal{K}_{u_t}$. In other words, $\max_{j \in \mathcal{K}_{u_t}} \text{rank}\left(\bar{\mathbf{V}}^{[j]}(t)\right) = d_{k_m^{u_t}}$.

Since at each transmission block all the cross-links have the constant elements, all the shared basis vectors at different receivers of $\text{RX}_j, j \in \mathcal{K}_{u_t}, j \neq k_m^{u_t}$, are in the space spanned by matrix $\bar{\mathbf{V}}^{[k_m^{u_t}]}(t)$. Similarly, at the $\text{RX}_{k_m^{u_t}}$, since all the cross channels among active transceivers have constant values, all the interference signals are aligned with each other. Now, we should show that all the desired signal are linearly independent of the interference signals. Considering the first lemma, from (35), we can conclude that the desired signal space at each active receiver is linearly independent of the interference signals. So, the proof has been completed. ■

The optimum vector for the $K-$tuple $(d_1, \ldots, d_K)$ to maximized d_{sum} can be obtained from a simple integer programming and searching algorithm (see Algorithm 1). Although, this searching algorithm shows a new scheme to use the channel changing pattern to partially or perfectly align interference signals, this algorithm cannot give us enough intuition to find out how it can be efficient. In the next section we find the average achievable sum DoF to analyze the performance of our technique.

V. Average Achievable Sum DoF and Numerical Results

A. Average achievable sum DoF:

In this section, we analyze the average achievable sum DoF $\bar{d}_{\text{sum}}$. We can model all signaling channels with the state diagram which has been shown in Fig. 1. In this figure in each time snapshot the state of the channel remains in its previous state with the probability of p_{ij} and changes with the probability of $1 - p_{ij}$. It is clear that if $p_{i_1 i_2} = p_{i_1 i_1}$, all the direct and the cross channels have the same statistical characteristic. From Definition 4, at the transmission block of

Algorithm 1 A search algorithm for finding optimum value of $K-$tuple $(d_1, \ldots, d_K)$

Input: $K, n,\ \mathcal{C}^{[ij]},\ i, j \in \mathcal{K}$

Set : $Q = 2^K - 1,\ d^* = 0$

for $u = 1$ to 2^Q **do**

$\mathcal{B}_1^{\mathcal{K}_{u_1}} = \{i | i \in [1:m]\,, m \notin \mathcal{C}'_{\mathcal{K}_{u_1}}\}$

for $t = 2$ to $|\mathcal{G}_u|$ **do**

$\mathcal{B}_t^{\mathcal{K}_{u_t}} = \{i | i \in \left[\max\left(\mathcal{B}_{t-1}^{\mathcal{K}_{u_{t-1}}}\right) + 1 : m\right], m \notin \mathcal{C}'_{\mathcal{K}_{u_t}}\}$

$M_t = 0$

for $k = 1$ to $k = K$ **do**

$E_k(t) = |\mathcal{B}_t^{\mathcal{K}_{u_t}}| - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[kk]}\right)$

end for

$\mathbf{E}(t) = (E_1(t), \ldots, E_k(t))$

for $(e_1(t), \ldots, e_K(t)) = (0, \ldots, 0)$ to $(n, \ldots, n)$ **do**

if $\sum_{j=1}^{K} e_j > M_t$, $(e_1(t), \ldots, e_K(t)) \leq \mathbf{E}(t)$ and $e_j + \max\limits_{k \in \mathcal{K}_{u_t} - j} e_k \leq |\mathcal{B}_t^{\mathcal{K}_{u_t}}|$ **then**

$(d_1(t), \ldots, d_K(t)) = (e_1(t), \ldots, e_K(t))$

$M_t = \sum_{j \in K_{u_t}} d_j(t)$

end if

end for

end for

if $\sum_{t=1}^{|\mathcal{G}_u|} M_t \geq d^*$ **then**

$d^* = \sum_{t=1}^{|\mathcal{G}_u|} M_t$

$\mathbf{d}(t) = (d_1(t), \ldots, d_K(t))$

end if

end for

Output: $\mathbf{d}(t)$

$\mathcal{B}_t^{\mathcal{K}_{u_t}}$ all the cross channels have a constant value. The probability that we have a transmission block of $\mathcal{B}_t^{\mathcal{K}_{u_t}}$ with the n^* time snapshots can be calculated from the following relation:

$$P\left(|\mathcal{B}_t^{\mathcal{K}_{u_t}}| = n^*\right) = \prod_{k_i \neq k_j} p_{k_i k_j}^{(n^*-1)}, k_i, k_j \in \mathcal{K}_{u_t}, \tag{37}$$

where $(n^* - 1)$ comes from the fact that during n^* time snapshots we have $n^* - 1$ transition points. During these transition points, all the cross channels should have a constant value. In (37), k_i and k_j indicate the indexes of transmitters and receivers in the set of $\mathcal{K}_{u_t}$. Let's define the vector $\mathbf{E}_{\mathcal{K}_{u_t}}$ as follows:

$$\mathbf{E}_{\mathcal{K}_{u_t}} = (e_{k_1}, e_{k_2}, \ldots, e_{k_r})\,, \tag{38}$$

where $e_{k_i} = n^* - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right),\ k_i \in \mathcal{K}_{u_t}$. In fact, e_{k_i} indicates the second constraint of (35) on the achievable DoF of d_{k_i}, in other words we should have $d_{k_i} \leq e_{k_i} = n^* - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right),\ k_i \in \mathcal{K}_{u_t}$ or equivalently, $\mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right) \leq n^* - d_{k_i},\ k_i \in \mathcal{K}_{u_t}$.

Now, we want to calculate the probability that $\mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right)$ has lower or equivalent value to the term of $n^* - d_{k_i},\ k_i \in \mathcal{K}_{u_t}$. In this case the states of the diagonal channel matrix $\bar{\mathbf{H}}_t^{[k_i k_i]}$

can be modeled with the vector $\mathbf{S}^{[k_i]} = \left(s_1^{[k_i]}, \ldots, s_{L^{[k_i]}}^{[k_i]}\right)$ where $1 \le s_i^{[k_i]} \le n^*$, $0 \le L^{[k_i]} \le n^*$. The value of $s_i^{[k_i]}$ indicates the number of time snapshots which is occupied by the i^{th} state of the channel between TX_{k_i} and RX_{k_i}. The value of $L^{[k_i]}$ in the vector $\mathbf{S}^{[k_i]} = \left(s_1^{[k_i]}, \ldots, s_{L^{[k_i]}}^{[k_i]}\right)$ indicates the number of blocks in which the diagonal matrix of $\bar{\mathbf{H}}_t^{[k_i k_i]}$ during n^* time snapshots has constant value. It is clear that the vector $\mathbf{S}^{[k_i]}$ has the following relation with the channel changing pattern of $\mathcal{C}^{[k_i k_i]}$:

$$s_1^{[k_i]} = c_1^{[k_i k_i]} - 1,\ \ s_l^{[k_i]} = c_l^{[k_i k_i]} - c_{l-1}^{[k_i k_i]}, \tag{39}$$

and $s_{L^{[k_i]}}^{[k_i]} = n^* - c_{L^{[k_i]}-1}^{[k_i k_i]} + 1$. As an example, for the channel matrix of $\bar{\mathbf{H}}^{[11]}$ in the table I, we conclude that $\mathbf{S}^{[11]} = (1,1,2,2,2,2)\,,\ L^{[1]} = 6$, $\mathbf{S}^{[22]} = (1,3,2,1,3)\,,\ L^{[2]} = 5$ and $\mathbf{S}^{[33]} = (1,2,2,3,2)\,,\ L^{[3]} = 5$. Now we wish to calculate the probability that during a transmission block consisted of n^* time snapshots the direct channel of $\bar{\mathbf{H}}_t^{[k_i k_i]}$ has $L^{[k_i]}$ states can be calculated as follows:

$$p\left(|\mathbf{S}^{[k_i k_i]}| = L^{[k_i]}\right) = \binom{n^*-1}{L^{[k_i]}-1} p_{k_i k_i}^{\sum_{l=1}^{L^{[k_i]}} s_l^{[k_i k_i]} - 1} \left(1 - p_{k_i k_i}\right)^{L^{[k_i]}-1} \overset{(a)}{=} \binom{n^*-1}{L^{[k_i]}-1} p_{k_i k_i}^{n^* - L^{[k_i]}} \left(1 - p_{k_i k_i}\right)^{L^{[k_i]}-1}, \tag{40}$$

where $|\mathbf{S}^{[k_i k_i]}| = L^{[k_i]}$ indicates the number of blocks in the direct channel of $\bar{\mathbf{H}}_t^{[k_i k_i]}$ and (a) comes from the fact that at each transmission block, the value of $\sum_{l=1}^{L^{[k_i]}} s_l^{[k_i k_i]}$ equals to the number of time snapshots of n^*. If d_{k_i} is the sum DoF of the k_i-th user, from second constraint of (35) we should have $d_{k_i} \le n^* - \mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right)$. It is clear if $L^{[k_i]}$ is the number of channel states in $\mathbf{S}^{[k_i]} = \left(s_1^{[k_i]}, \ldots, s_{L^{[k_i]}}^{[k_i]}\right)$, the value of $\mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right)$ can be calculated from the vector $\mathbf{S}^{[k_i]}$ with the following relation:

$$\mathcal{F}\left(\bar{\mathbf{H}}_t^{[k_i k_i]}\right) = \max_{1 \le l \le L^{[k_i]}} s_l^{[k_i k_i]}, \tag{41}$$

where $\max_{1 \le l \le L^{[k_i]}} s_l^{[k_i k_i]}$ indicates the maximum block length in which the channel matrix $\bar{\mathbf{H}}_t^{[k_i k_i]}$ has the constant value. From (40), we can conclude that the probability of occurance all the state vectors with $L^{[k_i]}$ states are similar and equals to $\binom{n^*-1}{L^{[k_i]}-1} p_{k_i k_i}^{n^* - L^{[k_i]}} \left(1 - p_{k_i k_i}\right)^{L^{[k_i]}-1}$.

Using second constraint of (35) and (40), we want to calculate the probability that all the values of the $s_l^{[k_i k_i]}, 1 \le l \le L^{[k_i]}$, have lower value than $n^* - d_{k_i}$. We define the set of events $\mathcal{X}^{[k_i k_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)$ as follows:

$$\mathcal{X}^{[k_i k_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right) = \left\{ s_l^{[k_i k_i]} \mid \max_l s_l^{[k_i k_i]} \le n^* - d_{k_i}, \sum_{l=1}^{L^{[k_i]}} s_l^{[k_i k_i]} = n^*, s_l \ge 1 \right\}. \tag{42}$$

From [25], we can easily show that the cardinality of the set $\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)$ has the following value:

$$|\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)| = \sum_{j=0}^{L^{[k_i]}} (-1)^j \binom{L^{[k_i]}}{j} \binom{n^* - j(n^* - d_{k_i}) - 1}{L^{[k_i]} - 1}. \tag{43}$$

Let $\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right)$ be defined as follows:

$$\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right) = \left\{ s_l^{[k_ik_i]} \mid \sum_{l=1}^{L^{[k_i]}} s_l^{[k_ik_i]} = n^*, s_l \geq 1 \right\}, \tag{44}$$

similar to (43) from [25], the cardinality of the set $\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right)$ has the value of $\binom{n^*-1}{L^{[k_i]}-1}$. From (40), all the events in the set of $\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right)$ have the same probability of occurrence. So, the probability of occurring each member of the set $\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)$ can be calculated by dividing the cardinality of the set $\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)$ to the cardinality of the set $\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right)$ multiplied by the probability that we have $L^{[k_i]}$ states (40). So, we have:

$$p\left(\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)\right) = p\left(|\mathbf{S}^{[k_ik_i]}| = L^{[k_i]}\right) \frac{|\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)|}{|\mathcal{Z}^{[k_ik_i]}\left(n^*, L^{[k_i]}\right)|} \tag{45}$$

$$= p_{k_ik_i}^{n^*-L^{[k_i]}} \left(1 - p_{k_ik_i}\right)^{L^{[k_i]}-1} |\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)|, \tag{46}$$

So, from (46) and (43) we have:

$$p\left(\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)\right) = p_{k_ik_i}^{n^*-L^{[k_i]}} \left(1 - p_{k_ik_i}\right)^{L^{[k_i]}-1} \sum_{j=0}^{L^{[k_i]}} (-1)^j \binom{L^{[k_i]}}{j} \binom{n^* - j(n^* - d_{k_i}) - 1}{L^{[k_i]} - 1}. \tag{47}$$

The probability that in a transmission block consisted of n^* time snapshots, the $|\mathcal{K}_{u_t}|$-tuple of $(d_{k_1}, \ldots, d_{k_r})$ from the active transceivers with the indexes in the set of $\mathcal{K}_{u_t}$ lies in the achievable DoF region of our proposed method can be derived as follows:

$$p\left(d_{k_1}, \ldots, d_{k_r} | \mathcal{K}_{u_t}, |\mathcal{B}_t^{\mathcal{K}_{u_t}}| = n^*\right) = \prod_{k_i \in \mathcal{K}_{u_t}} \sum_{L^{[k_i]}=1}^{n^*} p\left(\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)\right) \; i, j \in \mathcal{K}_{u_t}. \tag{48}$$

Therefore, from (37) we have:

$$p\left(d_{k_1}, \ldots, d_{k_r} | \mathcal{K}_{u_t}\right) = \prod_{i \neq j} p_{ij}^{(n^*-1)} \prod_{k_i \in \mathcal{K}_{u_t}} \sum_{L^{[k_i]}=1}^{n^*} p\left(\mathcal{X}^{[k_ik_i]}\left(n^*, L^{[k_i]}, d_{k_i}\right)\right) \; i, j \in \mathcal{K}_{u_t}, \tag{49}$$

where the term $\prod_{i \neq j} p_{ij}^{(n^*-1)}$ indicates the probability of $P\left(|\mathcal{B}_t^{\mathcal{K}_{u_t}}|\right)$ as it has been calculated in (37). Therefore, the average transmission sum DoF can be calculated as follows:

$$\bar{d}_{\text{sum}} = \sum_{1 \leq n^* \leq \infty, \mathcal{K}_{u_t} \subseteq \mathcal{K}} \left(p\left(d_{k_1}, \ldots, d_{k_r} | \mathcal{K}_{u_t}\right) \frac{1}{n^*} \sum_{k_i \in \mathcal{K}_{u_t}} d_{k_i} \right), 1 \leq n^* \leq \infty, \tag{50}$$

where $d_{k_i} + \max_{k_j \in \mathcal{K}_{u_t}, k_j \neq k_i} d_{k_j} \leq n^*$. Fig. 2 shows the numerical results for the case where the probability of channel changing pattern for the cross channels are 0.8 and the probability of channel changing pattern for the direct channels ranges from 0.1 to 0.9. As it has been indicated

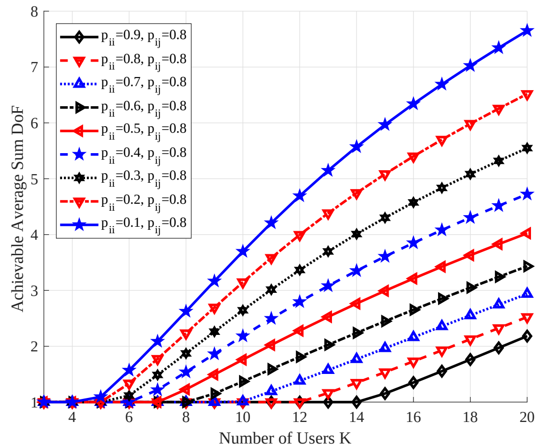


Fig. 2: This figure shows the average achievable sum DoF as a function of number of user. As the number of users increases the chance of finding proper channel conditions increase and we achieve to higher average achievable sum DoF.

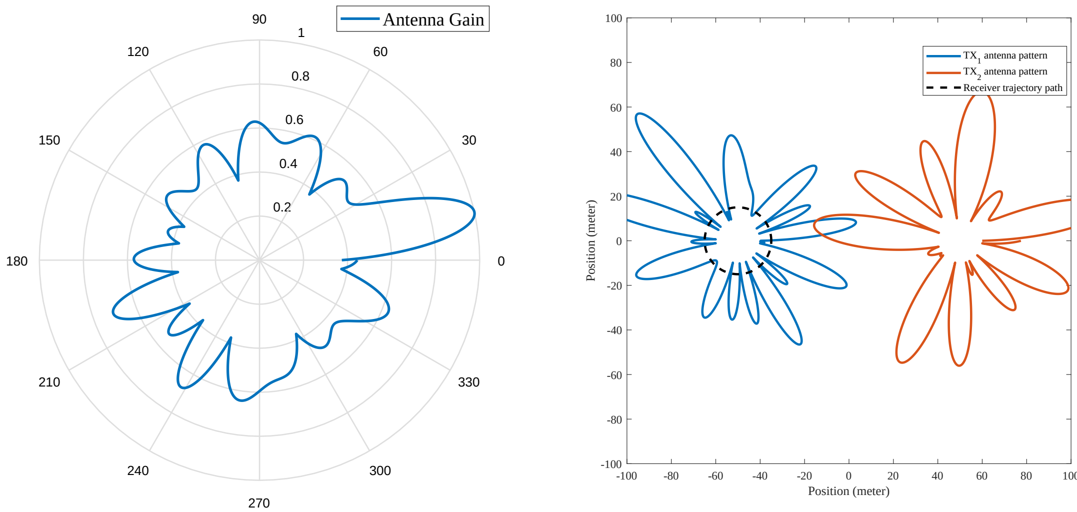
when the direct channel has higher fluctuations and variability, we can achieve higher average achievable sum DoF. Therefore, the performance of our scheme can be increased by generating such conditions that increase the chance of channel variations for the direct links while the cross links stay constant. In the next section, we propose some means to achieve this goal.

VI. On the practicality of channel variation assumption

In this work, we assumed variations in antenna radiation patterns result in different coherence times for different links. Here, we motivate this assumption for mobile users, and present an antenna design that will create such patterns for stationary users.

A. Mobile users

Consider a K-user IC in which receivers have a random and time-varying physical location. We assume each transmitter is equipped with an antenna whose pattern exhibits high fluctuations similar to Fig. 3(a). As an example, Fig. 3(b) shows a mobile receiver with a circular trajectory around its desired transmitter with the antenna structure of Fig. 3(a). This receiver is also affected by an interfering transmitter. As it indicated in this figure, the receiver along its circular path, experiences many channel variations from the desired transmitter but the interference link generates lower variations. For cellular applications, we can assume that all desired links have



(a) (b)

Fig. 3: (a) Antenna pattern structure at the transmitter with high radiation pattern fluctuations; (b) Two transmitters and a receiver that moves along a circular trajectory path.

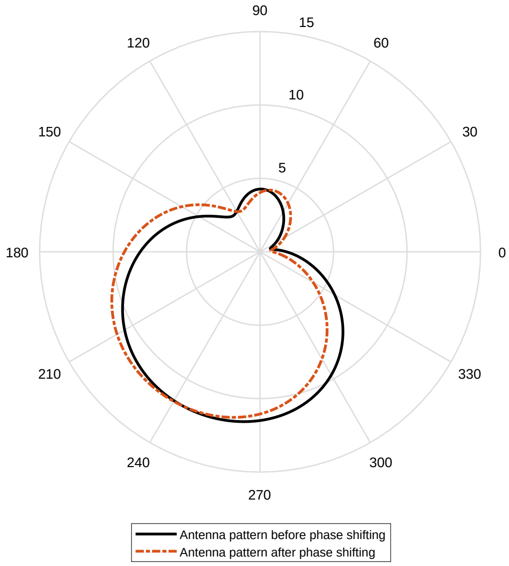
lower distances than interfering links. Therefore, receivers experience the desired links with higher fluctuation rates.

B. Stationary users

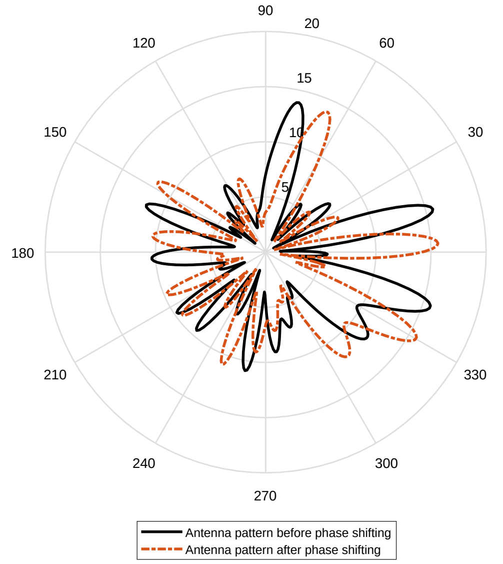
In this case, we consider the 2-dimensional (2D) antenna structure of Fig. 5(a), which is fed by a power divider and a number of phase shifter elements. These phase shifter elements are controlled by a simple algorithm at the transmitter to control and rotate the transmission beam of the antenna with the small rotation angle of $\Delta\phi$. Fig. 4(a) and Fig. 4(b) depict the transmission beam of our antenna structure. For propagation angles close to the transmitter antenna plane, the transmission beam has a smooth shape. On the other hand, for angles perpendicular to the transmitter antenna plane, the transmission beam has a lot of fluctuations. The transmission gain of the antenna as a function of space angles of θ and ϕ can be calculated as follows:

$$|A(\theta, \phi)| = |\sum_{m=0}^{M} \sum_{n=0}^{N} \alpha(m,n) e^{\sqrt{-1}(\pi d \sin(\theta)((m-0.5)\cos(\phi)+(n-0.5)\sin(\phi))+\psi(m,n))}|, \tag{51}$$

where $\alpha(m,n) \in \{0,1\}$ and $\psi(m,n)$ is a random phase generated by each element of the transmitter antenna. From (51), when the value of θ is close to zero, the antenna gain has the lowest sensitivity to the value of ϕ, while when θ is close to $\frac{\pi}{2}$, it has highest sensitivity to ϕ.



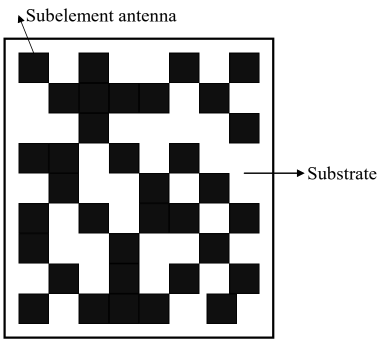
(a)



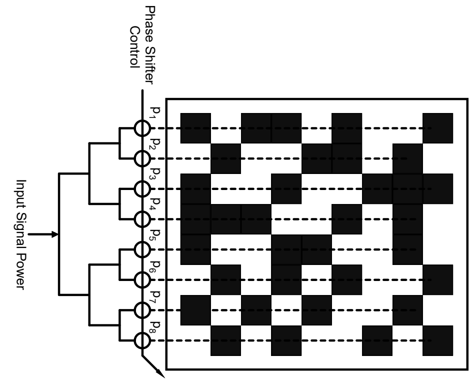
(b)

Fig. 4: (a) Transmission pattern of the antenna array with the propagation angle $\theta = 10$ (close to the plane of antenna structure) and small variation of the rotation angle (for $\Delta\phi = 10$): in this case there is a small change between these propagation patterns; (b) Similar to (a), but for propagation angle $\theta = 90$ (for $\Delta\phi = 10$), we can see a lot of differences between the original and the rotated patterns.

The transmission beam of the antenna can be rotated using phase shifter elements $\{p_0, \ldots, p_n\}$ depicted in Fig. 5(b). The reception area can be covered by properly installing such antenna arrays. One can install these antenna structures such that the receivers close to the transmitter antenna, receive the signal with angles close to $\theta = 90$, and the receivers further away, receive the signal with an angle close to $\theta = 0$. Therefore, as it has been indicated in Fig. 4(a) and Fig. 4(b), closer receivers will have a higher sensitivity to the pattern rotation of the antenna and see their corresponding channel with higher fluctuation rates. In cellular networks, this a realistic assumption as receivers in a cell have a closer distance to their corresponding transmitters in comparison to the interfering transmitters in the adjacent cells operating in the same frequency bandwidth. Therefore, there is no need for the receivers to have mobility, and proper conditions for the receivers can be achieved by rotating the transmitter beam using electronic phase shifter elements with a small $\Delta\phi$. Motivated by our discussion in this section, in the following section, we propose a blind interference alignment scheme that does not need to access channel changing



(a)



(b)

Fig. 5: (a) A simple structure of transmitter antenna with many randomly-positioned elements. These elements can be considered as parasitic element or exited separately by proper feeding network with random amplitude and phase; (b) A feeding structure for the antenna. The total power is divided by a proper power divider. This structure can generate a random beam which can rotated by some phase shifter elements. These phase shifters can be controlled by a simple algorithm to rotate the transmission angle.

pattern to maximize average transmission rate.

VII. Blind Opportunistic Interference Alignment

So far, we assumed that all transmitters have access to non-causal knowledge of all channel changing patterns. Although non-causality makes the scheme impractical, it provides us with valuable insights that enable us to develop a practical solution under certain assumptions. In this section, we consider the case in which transmitters have no access to the channel state information or channel changing patterns, and using superposition coding try to maximize their average transmission rate. At all of the transmitters $\mathrm{TX}_k, 1 \leq k \leq K$, we use a multi-layer encoding scheme consists of M distinct layers of $l_i^{[k]}, 1 \leq i \leq M$. In this coding procedure, every layers of $l_i^{[k]}, 1 \leq i \leq M$ are represented by a Gaussian random variable $X_i^{[k]},\ 1 \leq i \leq M$. Without accessing channel state information at transmitters, we are going to maximize the average transmission rate. These transmission layers are designed such that the decoder based on the state of channel changing pattern structure can recover some parts of the transmitted data.

A. Encoding method

The encoding strategy has the following steps:

1) Let $X_i^{[k]}, 1 \leq i \leq M$ be M Gaussian distributed continuous random variables with limited power constraint.

2) At TX_k, we partition our message $W^{[k]}$ into M sub-messages of $\left(W_1^{[k]}, \cdots, W_M^{[k]}\right)$, and we send each sub-message via a single transmission layer. For each layer of $l_i^{[k]}, 1 \leq i \leq M$ consisted of n time snapshots, we consider $\left(2^{nR_i^{[k]}}, \frac{n}{2}\right)$ code for the Gaussian channel with a limited power constraint. We generate codewords of $x_{i1}^{[k]\frac{n}{2}}(1), \ldots, x_{i1}^{[k]\frac{n}{2}}\left(2^{\frac{n}{2}R_i^{[k]}}\right)$ as follows:

$$x_{i1}^{[k]\frac{n}{2}}\left(2^{nR_i^{[k]}}\right) = \left(x_{i1}^{[k]}\left(2^{nR_i^{[k]}}\right), \ldots, x_{i\frac{n}{2}}^{[k]}\left(2^{nR_i^{[k]}}\right)\right)^{\mathrm{T}}, \tag{52}$$

3) At j-th transmission time, TX_k computes $X_j^{[k]} = \sum_{i=1}^{M} x_{ij}^{[k]}(W_i^{[k]})$ of M independent layers and concatenates them to generate the matrix $\bar{\mathbf{x}}^{[k]} = \left[X_1^{[k]}, \ldots, X_{\frac{n}{2}}^{[k]}\right]^T$.

4) Transmitters use a random full rank matrix of $\bar{\mathbf{V}}$ with the size of $n \times \frac{n}{2}$ as their precoder matrices. We can design $\bar{\mathbf{V}}$ such that $\lim_{n\to\infty} \frac{1}{n}\mathrm{tr}\left(\bar{\mathbf{V}}\bar{\mathbf{x}}^{[k]}(\bar{\mathbf{V}}\bar{\mathbf{x}}^{[k]})^H\right) = P_t$, where $\mathrm{tr}\left(\mathbf{A}\right)$ computes the trace of $\mathbf{A}$.

5) The output of the encoder at TX_k, $\bar{\mathbf{X}}^{[k]}$, is an $n \times 1$ column vector with the following relation:

$$\bar{\mathbf{X}}^{[k]} = \bar{\mathbf{V}}\bar{\mathbf{x}}^{[k]}, \tag{53}$$

and for i^{th} layer we have the following constraint:

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{tr}\left(\bar{\mathbf{V}}x_{i1}^{[k]\frac{n}{2}}\left(W_i^{[k]}\right)\left(\bar{\mathbf{V}}x_{i1}^{[k]\frac{n}{2}}\left(W_i^{[k]}\right)\right)^H\right) = P_i^{[k]}, \tag{54}$$

where $P_i^{[k]}$ [1] indicates the total transmission power for i^{th} layer at TX_k and $\sum_{i=1}^{M} P_i^{[k]} = P_t$, which shows that the summation of transmission power for all of the layers is equal to P_t. In other words, for the combination of all layers we have the following constrain:

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{tr}\left(\bar{\mathbf{X}}^{[k]}\left(\bar{\mathbf{X}}^{[k]}\right)^H\right) = \sum_{i=1}^{M} P_i^{[k]} = P_t, \tag{55}$$

B. Model of the received signal

Fig. 6, shows a simple model for reception space at a receiver for 3-user interference channel. In this case we assume that we have deployed the antenna structure of previous section and $p_{ij} >> p_{ii}, i \neq j$, it means that the probability of changing channel value for the cross channels is much lower than direct channels. As it is indicated in this figure all the interference signals are aligned at the receiver, while depend on the value of $\mathcal{F}\left(\bar{\mathbf{H}}\right)$ the direct channel signal space has some free interference dimensions which are linearly independent from the interference signals. The decoder decodes the transmitted signal using following zero forcing decoding matrix of $\bar{\mathbf{D}}$

[1] For simplifying the notation all the superscript in [.] indicates the indexes of the users.

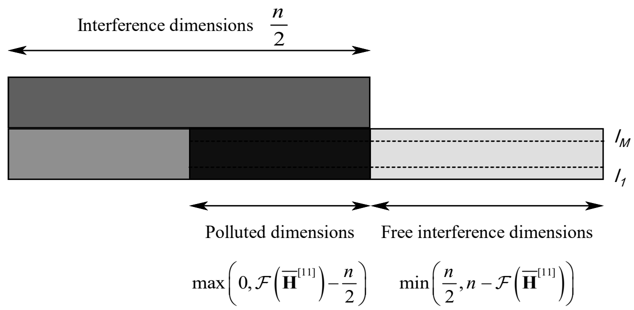


Fig. 6: The reception space of a receiver includes three different subspaces: the desired space which is free of interference and dimension of $\min\left(\frac{n}{2}, n - \mathcal{F}(\bar{\mathbf{H}})\right)$; the desired space polluted by interference signals with dimension $\max\left(0, \mathcal{F}(\bar{\mathbf{H}}) - \frac{n}{2}\right)$; the interference space.

$= \left(\bar{\mathbf{I}} - \bar{\mathbf{V}}(\bar{\mathbf{V}}^H\bar{\mathbf{V}})^{-1}\bar{\mathbf{V}}^H\right)$. If we assume during n transmission time snapshots all the cross links have a constant value, we have:

$$\bar{\mathbf{D}}\bar{\mathbf{H}}^{[ij]}\mathbf{V}\bar{\mathbf{x}}^{[j]} = \bar{\mathbf{H}}^{[ij]}\bar{\mathbf{D}}\mathbf{V}\bar{\mathbf{x}}^{[j]} = \bar{\mathbf{H}}^{[ij]}\left(\bar{\mathbf{I}} - \bar{\mathbf{V}}(\bar{\mathbf{V}}^H\bar{\mathbf{V}})^{-1}\bar{\mathbf{V}}^H\right)\mathbf{V}\bar{\mathbf{x}}^{[j]} \tag{56}$$

$$= \bar{\mathbf{H}}^{[ij]}\left(\mathbf{V} - \bar{\mathbf{V}}(\bar{\mathbf{V}}^H\bar{\mathbf{V}})^{-1}\bar{\mathbf{V}}^H\mathbf{V}\right)\bar{\mathbf{x}}^{[j]} = \bar{\mathbf{0}} \tag{57}$$

Since all the transmitters use the same precoder of $\bar{\mathbf{V}}$, for the encoding scheme of the previous subsection where all the transmitted signals have the same number of dimension of $\frac{n}{2}$, the number of free interference subspace dimensions can be calculated from the first lemma as follows:

$$\operatorname{rank}([\mathbf{V}\ \mathbf{H}\mathbf{V}]) - \operatorname{rank}([\mathbf{V}]) = \min\left(d_v + \min\left(d_v, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right) - d_v, \tag{58}$$

Substituting $d_v = \frac{n}{2}$ in the above equation we have:

$$\operatorname{rank}([\mathbf{V}\ \mathbf{H}\mathbf{V}]) - \operatorname{rank}([\mathbf{V}]) = \min\left(\frac{n}{2} + \min\left(\frac{n}{2}, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), n\right) - \frac{n}{2} \tag{59}$$

$$= \min\left(\min\left(\frac{n}{2}, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right), \frac{n}{2}\right) = \min\left(\frac{n}{2}, n - \mathcal{F}\left(\bar{\mathbf{H}}\right)\right) \tag{60}$$

C. Decoding method

To prevent the repetition, we consider the decoding operation at the first receiver and other receivers have the same analysis. We assume all the direct channel matrices are known at their corresponding receivers and the channel matrices give no gain to the received signals. In other words, we assume that for all the channel links $\det\left(\bar{\mathbf{H}}^{[ij]}\left(\bar{\mathbf{H}}^{[ij]}\right)^H\right) = 1$.

After applying zero forcing matrix at the first receiver, from (60) we have $\min\left(\frac{n}{2}, n - \mathcal{F}\left(\bar{\mathbf{H}}^{[11]}\right)\right)$ free interference dimensions. Let $\bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}$ be the received sequence at the first receiver, then based on the following successive decoding strategy steps we can decode a part of transmitted data:

1) *First layer decoding:* Based on $\bar{\mathbf{H}}^{[ij]}$ and the decoder matrix of $\bar{\mathbf{D}}$, the decoder finds that $\widehat{w}_1$ is

sent if there exists a unique message such that 2-tuple $n-$sequence of $\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \in \mathcal{T}_\epsilon^n$, otherwise it declares error in the first decoding step.

2) *Second layer decoding:* After decoding the first layer and if such $\widehat{w}_1$ is found, the decoder tries to find $\widehat{w}_2$ such that the 3-tuple $n-$sequence of $\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1), \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{21}^{[1]\frac{n}{2}}(\widehat{w}_2), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \in \mathcal{T}_\epsilon^n$, otherwise it declares error in the second decoding stage.

3) *i-th layer decoding:* After finding unique messages $\widehat{w}_j, \cdots, \widehat{w}_{i-1}$ in the previous steps, the decoder finds $\widehat{w}_i$ such that $(i+1)-$tuple of $n-$sequences $\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(\widehat{w}_i), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right)$ $\in \mathcal{T}_\epsilon^n$, otherwise it declares error in the i-th decoding step.

D. Error probability analysis

In this subsection, we derive a set of upper-bounds on the transmission rate of the i-th layer to guarantee the decodability. More precisely, we focus on receiver one, and we show that if

$$R_i^{[1]} \leq \frac{i}{2}\log\left(1 + \frac{P_i^{[1]}}{N + \sum_{j=i+1}^{M} P_j^{[1]}}\right),\ M = \frac{n}{2} \tag{61}$$

then, the message in the i-th transmission layer can be decoded with vanishing error probability. We note that in (61), N is the noise power after zero-forcing at the receiver of interest. Let $\mathcal{E}_i$ be an event that decoding the i-th layer is erroneous. Thanks to the symmetry of the random codebook generation,

$$P\left(\mathcal{E}_i\right) = P\left(\mathcal{E}_i | W_1^{[1]} = 1, \ldots, W_M^{[1]} = 1\right) = P\left(\widehat{W}_i^{[1]} \neq 1 | W_1^{[1]} = 1, \ldots, W_M^{[1]} = 1\right), \tag{62}$$

where $\widehat{W}_i^{[1]}$ indicates the i-th decoded message at the receiver. The decoder in the first step of decoding makes an error if one or both of the following events occur:

$$\mathcal{E}_{11} = \left\{\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \notin \mathcal{T}_\epsilon^n\right\},\ \mathcal{E}_{12} = \left\{\exists \widehat{w}_1 \neq 1 :\ \left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \in \mathcal{T}_\epsilon^n\right\}. \tag{63}$$

By the law of large numbers (LLN), the probability of the first term, $P(\mathcal{E}_{11})$, tends to zero as $n \to \infty$. From the packing lemma, the probability of the second event, $P(\mathcal{E}_{12})$, also tends to zero as n goes to infinity as long as $R_1 \leq I\left(X_1; Y\right)$. So, if $R_1^{[1]} \leq I\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1); \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right)$ and $n \to \infty$, the value of $P(\mathcal{E}_1) = P(\mathcal{E}_{11} \cup \mathcal{E}_{12})$ goes to zero. After decoding the first transmission layer, the receiver subtracts $\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(\widehat{w}_1)$ from the received signal.

Now, assume the decoder has recovered messages $W_1^{[1]}, \ldots, W_{i-1}^{[1]}$ correctly. In the i-th step, the decoder makes an error if one or both of the following events occur:

$$\mathcal{E}_{i1} = \left\{\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(1), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \notin \mathcal{T}_\epsilon^n\right\}, \tag{64}$$

$$\mathcal{E}_{i2} = \left\{\exists \widehat{w}_i \neq 1 :\ \left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(\widehat{w}_i), \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) \in \mathcal{T}_\epsilon^n\right\}. \tag{65}$$

Again by LLN, $P(\mathcal{E}_{i1}) \to 0$ as $n \to \infty$, and by using packing lemma if we have:

$$R_i^{[1]} \le I\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(\widehat{w}_1); \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n} | \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right), \tag{66}$$

$P(\mathcal{E}_{i2})$ tends to zero as n goes to infinity. Thus, as $n \to \infty$, the value of $P(\mathcal{E}_i) = P(\mathcal{E}_{i1} \cup \mathcal{E}_{i2})$ goes to zero. Also, we can assume that for decoding each layer of i we have i free dimensions at the receiver. In other words, at RX_1 and during the i^{th} decoding stage, we assume that the direct channel between TX_1 and RX_1 has such condition that after multiplying the received signal by zero-forcing matrix of $\bar{\mathbf{D}}$, we have i free interference dimensions. If we cannot find at least i interference dimensions at receiver, then, the decoder failes to decode transmitted data of i^{th} layer. In this encoding-decoding strategy, we can assume that the number of layers is $\frac{n}{2}$ which is the half of the number of time snapshots in each transmission block ($M = \frac{n}{2}$). The transmission rate of i^{th} layer, satisfying the above constraints, can be upper-bounded as follows:

$$\begin{aligned}
R_i^{[1]} &\le I\left(\bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(\widehat{w}_1); \bar{\mathbf{D}}\mathbf{Y}_1^{[1]n} | \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right) && (67)\\
&= h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_1^{[1]}, \ldots, \bar{\mathbf{D}}\bar{\mathbf{Y}}_n^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right)\\
&\quad - h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_1^{[1]}, \ldots, \bar{\mathbf{D}}\bar{\mathbf{Y}}_n^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(1)\right) && (68)\\
&= \sum_{s_j \in \mathcal{S}} h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_j}^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_1}^{[1]}, \ldots, \bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_{j-1}}^{[1]}, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right)\\
&\quad + \underbrace{\sum_{s \in \mathcal{S}^c} h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_{s}^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{Y}}_{\mathcal{S}}^{[1]}, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right)}_{\stackrel{(a)}{=} o_1(\log(P_t))}\\
&\quad - \sum_{s_j \in \mathcal{S}} h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_j}^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_1}^{[1]}, \ldots, \bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_{j-1}}^{[1]}, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{(i-1)1}^{[1]\frac{n}{2}}(1)\right)\\
&\quad \underbrace{- \sum_{s_j \in \mathcal{S}^c} h\left(\bar{\mathbf{D}}\bar{\mathbf{Y}}_{s_j}^{[1]} | \bar{\mathbf{D}}\bar{\mathbf{Y}}_{1\mathcal{S}}^{[1]}, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{11}^{[1]\frac{n}{2}}(1), \ldots, \bar{\mathbf{D}}\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}x_{i1}^{[1]\frac{n}{2}}(1)\right)}_{\stackrel{(a)}{=} o_2(\log(P_t))} && (69)\\
&\stackrel{(b)}{=} |\mathcal{S}|/2 \log_2 2\pi e\left(\sum_{j=i}^{M} P_j^{[1]} + N\right) - |\mathcal{S}|/2 \log_2 2\pi e\left(\sum_{j=i+1}^{M} P_j^{[1]} + N\right) + o'(\log(P_t)) && (70)\\
&= |\mathcal{S}|/2 \log_2\left(1 + \frac{P_i^{[1]}}{\sum_{j=i+1}^{M} P_j^{[1]} + N}\right) + o'(\log(P_t)) && (71)
\end{aligned}$$

where $\bar{\mathbf{D}}\bar{\mathbf{Y}}_{1s}^{[1]}$ is the s^{th} row of $\bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}$; the members of set $\mathcal{S}$ in the above equations are the rows of matrix $\bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}$, for $\mathrm{rank}\left(\bar{\mathbf{D}}\mathbf{Y}_1^{[1]n}\right) = |\mathcal{S}|$, and where all other rows in $\mathcal{S}^c = \{1, \ldots, n\} - \mathcal{S}$ can be reconstructed (up to the noise term) from the linear combination of these rows; (a) comes from the fact that any row with an index in $\mathcal{S}^c$ can be calculated from the rows in $\mathcal{S}$ within noise power; (b) comes from the distribution of input data which is Gaussian with limited power and the random nature of precoder matrix which holds the i.i.d distribution of input sequence, and that we obtain no power gain from the channel gains and the decoding operation (*i.e.*, we assumed $\det\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)=1$ and $\det\left(\bar{\mathbf{D}}\bar{\mathbf{D}}^H\right)=1$). As we discussed before, the decoder can decode the transmitted data from i^{th} layer if the number of interference-free dimensions at i^{th} stage is equal to i. Therefore, the decoder can decode the transmitted data from i^{th} layer iff $|\mathcal{S}|$ equals to the value of i. Finally, based on the argument presented above, we conclude that:

$$R_i^{[1]} \leq \frac{i}{2}\log\left(1+\frac{P_i^{[1]}}{N+\sum_{j=i+1}^{M}P_j^{[1]}}\right)+o'\left(\log\left(P_t\right)\right),\ M=\frac{n}{2}, \tag{72}$$

where N is the noise power after zero-forcing at the receiver. Intuitively, at the first receiver, we can treat interference from all other transmission layers as noise, and the value of i indicates the number of interference-free dimensions at the i^{th} decoding stage. If P_t is large, then, we can neglect the effect of $o'\left(\log\left(P_t\right)\right)$ in our equations. Therefore, depending on different channel states, receivers will be able to decode a part of transmitted data. If the number of interference-free dimensions at the receiver is l_s, we can except that the receiver can decode the transmission layers from layer one to the l_s^{th} layer. In the next subsection, based on (59), we divide the total transmission power among $M=\frac{n}{2}$ layers to maximize the average transmission power.

E. Maximize average transmission rate:

Referring previous subsection, the average transmission rate per transmission time can be calculated from the following relation:

$$\bar{R}_{\text{avg}}=\frac{1}{n}\sum_{i=1}^{\frac{n}{2}}\frac{i}{2}\log\left(1+\frac{P_i^{[1]}}{N+\sum_{j=i+1}^{M}P_j^{[1]}}\right)F_I\left(\frac{n}{2}-i\right) \tag{73}$$

where $F_I\left(\frac{n}{2}-i\right)=p\left(\min\left(\frac{n}{2},n-\mathcal{F}\left(\bar{\mathbf{H}}\right)\right)\geq i\right)$. When the value of n has enough large value, the value of $\frac{i}{n}, 1\leq i\leq\frac{n}{2}$ and the function $P_i^{[1]}$ can be represented by the new parameters of $0\leq z\leq\frac{1}{2}$ and $\rho(z)dz$, respectively. In this case we consider that $P(1/2-z)=\int_z^{\frac{1}{2}}\rho(x)dx$. So, $P\left(0\right)=\int_{\frac{1}{2}}^{\frac{1}{2}}\rho(x)dx=0$ and $P\left(z=1/2\right)=\int_0^{\frac{1}{2}}\rho(x)dx=P_t^{[1]}$. Therefore, when the value of n has enough large value, the equation 73 can be represented as follows:

$$\bar{R}_{\text{avg}}=\frac{1}{2\ln 2}\int_0^{\frac{1}{2}}\frac{z\rho(1/2-z)F_Z\left(1/2-z\right)}{N+P(1/2-z)}dz \tag{74}$$

where $F_Z\left(\frac{1}{2}-z\right)=F_I\left(\frac{n}{2}-\lfloor nz\rfloor\right),\ 0\leq z\leq\frac{1}{2}$ and $\rho(z)=\frac{dP(z)}{dz}$. The following theorem help us to maximize the above average transmission rate. For simplifying our notation we subside $\frac{1}{2}-z$ with u and therefore we have:

$$\bar{R}_{\text{avg}}=\frac{1}{2\ln 2}\int_{\frac{1}{2}}^{0}\frac{-(1/2-u)\rho(u)F_U\left(u\right)}{N+P(u)}du=\frac{1}{2\ln 2}\int_0^{\frac{1}{2}}\frac{(1/2-u)\rho(u)F_U\left(u\right)}{N+P(u)}du \tag{75}$$

Theorem 1: A necessary condition for the function $y(u)$ to be an extremum of:

$$\int_{z_1}^{z_2} D\left(y, y^{'}, u\right) du, \tag{76}$$

is that y satisfies the following Euler differential equation of $D_y - \frac{dD_{y'}}{du} = 0,\ z_1 \leq u \leq z_2$, where the subscripts denote the partial derivatives with respect to corresponding arguments [26], [27].
In our problem we have:

$$D(y, y', u) = \frac{1}{2\ln 2} \frac{(1/2 - u)\rho(u) F_U\left(u\right)}{N + P(u)} \tag{77}$$

where $y(u) = P(u)$ and $y'(u) = \rho(u)$. In this case we have:

$$D_y = \frac{1}{2\ln 2} \frac{-(1/2 - u)\rho(u) F_U(u)}{\left(N + P(u)\right)^2},\ D_{y'} = \frac{1}{2\ln 2} \frac{(1/2 - u) F_U(u)}{N + P(u)}. \tag{78}$$

So, we have $\frac{dD_{y'}}{du} = \frac{1}{2\ln 2} \frac{(-F_U(u) + (1/2 - u) f_U(u))(N + P(u))}{(N + P(u))^2}$, where $f_U(u) = \frac{dF_U(u)}{du}$, and from the first theorem we can conclude that:

$$(1/2 - u)\rho(u) F_U(u) + (-F_U(u) + (1/2 - u) f_U(u))(N + P(u)) = 0. \tag{79}$$

So, we have:

$$\frac{dP(u)}{N + P(u)} = -\frac{(-F_U(u) + (1/2 - u) f_U(u)) du}{(1/2 - u) F_U(u)}, \tag{80}$$

and finally:

$$\ln\left(N + P(u)\right) = -\ln\left((1/2 - u) F_U(u)\right) + C. \tag{81}$$

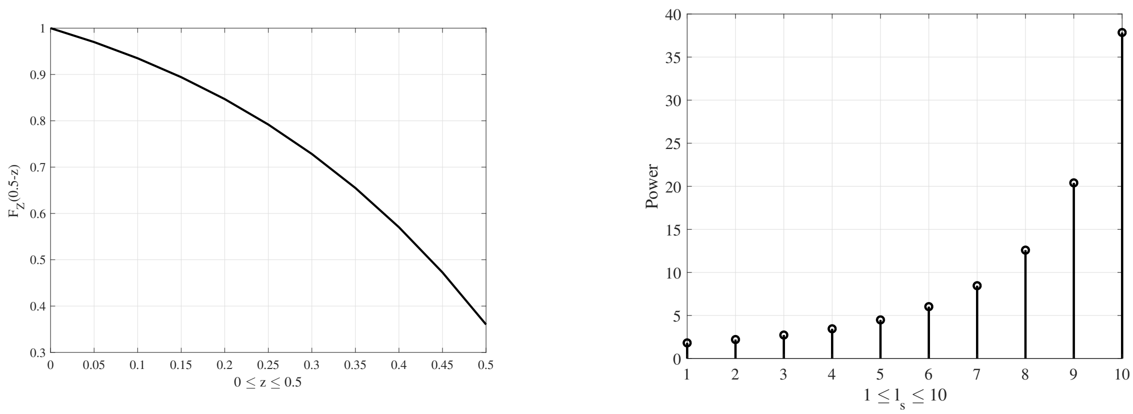
Therefore:

$$P(1/2 - z) = \frac{C}{z F_Z(1/2 - z)} - N. \tag{82}$$

where $P(1/2 - z = 0) = 0$. So, the value of C can be calculated by setting $z = \frac{1}{2}$ from the above equation we can set $C = \frac{N}{2} F_Z\left(0\right)$, and the function $P(z)$ can be calculated as follows:

$$P(z) = \frac{N}{2(1/2 - z) F_Z(z)} F_Z\left(0\right) - N. \tag{83}$$

When $P(z_0) \geq P_t^{[1]}$, we have $P(z > z_0) = P_t^{[1]}$. From (43), the value of $F_Z(z)$ for the precoder length of $n = 20$ and the probability of changing direct channel of $p_{11} = 0.9$, has been depicted in Fig. 7a, and the function $P(z)$ can be calculated based on the calculated function of $F_Z(z)$. By finding $P(z)$, the transmission power for each transmission layer can be calculated from $P_j^{[1]} = P_Z\left(\frac{j-1}{n}\right) - P_Z\left(\frac{j}{n}\right)$. For a case where $P_t^{[1]} = 100$, $n = 20$ and $N = 1$ the total transmission power for each transmission layer has been depicted in Fig. 7b.



(a)

(b)

Fig. 7: (a) Finding the cumulative distribution $F_Z(z)$ from $F_I(i)$. The value of $F_I(i)$ can be calculated from $F_I(i) = \sum_{L=1}^{n} p\left(\mathcal{X}^{[11]}\left(n, L, n-i\right)\right)$; (b) The allocated power for each transmission layer of the proposed scheme. In this figure, we assume that $n = 20$ and $p_{11} = 0.9$.

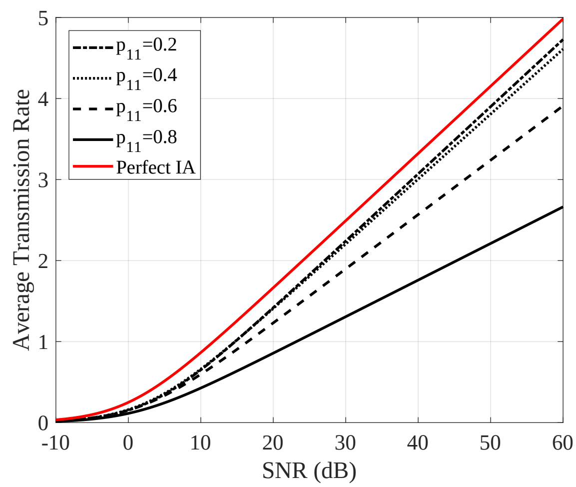


Fig. 8: Average achievable rate per user in a $K-$user IC for $p_{ii} = p_{11} \in \{0.2, 0.4, 0.6, 0.8\}$ and $n = 20$. We assume that all cross channels due to our implemented antenna structure have constant values. This figure shows that when the direct channels have higher fluctuations, we can achieve a higher average transmission rate.

F. Numerical Results:

In this subsection, we consider an interference channel and demonstrate the gain of our methodology for a relatively small transmission block length. We note that many prior IA schemes, require long block-lengths to achieve the desired rates. With the radiation pattern of Fig. 3(a), we can divide the transmission time into sufficient segments in which cross links have constant values. In this example, we assume a transmission block-length of $n = 20$. Fig. 8 shows the average achievable transmission rate for different values of p_{ii} using multi-layer encoding strategy of Section VII. This figure indicates that when p_{11} decreases, meaning that the fluctuation rate of direct links increases, then, the chance of finding proper channel states to align interfering signals and to receive the desired signal linearly independent of interference signals increase.

VIII. Conclusion

In this paper, we investigate the problem of interference alignment for the $K-$user IC. The proposed scheme allows transceivers to achieve higher average sum rate without knowledge of channel state information at transmitters. In fact, in the proposed scheme of this paper, all the transmitters in a hunting mode strategy wait for proper channel conditions to transmit their data with the highest available sum DoF. We showed that with an antenna structure deployed at transmitters one can generate conditions to increase the chance of alignment conditions. For a more practical scenario in which the transmitters have not accessed to the channel variation time snapshots, we proposed a multi-layer encoding strategy. In this method based on cumulative distribution of $F_I\left(\frac{n}{2}-i\right) = p\left(\min\left(\frac{n}{2}, n-\mathcal{F}\left(\bar{\mathbf{H}}\right)\right) \geq i\right)$, we divide the total transmission power among many layers to achieve highest achievable sum rate. At the receivers based on the channel conditions, the decoder can decode a part of transmitted information from the desired transmitter.

References

[1] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings*, vol. 3, pp. 1745–1749 vol.3, May 2001.

[2] A. Tolli, H. Pennanen, and P. Komulainen, "On the value of coherent and coordinated multi-cell transmission," in *2009 IEEE International Conference on Communications Workshops*, pp. 1–5, June 2009.

[3] S. A. Jafar, "Topological interference management through index coding," *IEEE Transactions on Information Theory*, vol. 60, pp. 529–568, January 2014.

[4] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the k user interference channel," *IEEE Transactions on information Theory*, vol. 54, pp. 3425–3441, August 2008.

[5] S. W. Choi, S. A. Jafar, and S.-Y. Chung, "On the beamforming design for efficient interference alignment," *IEEE Communications Letters*, vol. 13, no. 11, pp. 847–849, 2009.

[6] M. Johnny and M. R. Aref, "An efficient precoder size for interference alignment of the K-user interference channel," *IEEE Communications Letters*, vol. 21, no. 9, pp. 1941–1944, 2017.

[7] M. A. Maddah-Ali, S. A. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Transactions on Information Theory*, vol. 54, pp. 3457–3470, August 2008.

[8] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom for the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, August 2008.

[9] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2783–2794, 2011.

[10] O. El Ayach and R. W. Heath, "Interference alignment with analog channel state feedback," *IEEE Transactions on Wireless Communications*, vol. 11, no. 2, pp. 626–636, 2011.

[11] M. A. Maddah-Ali and D. N. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Transactions on Information Theory*, vol. 58, pp. 4418–4431, July 2012.

[12] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6.3, pp. 228–240, June 2012.

[13] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity region of the miso broadcast channels with delayed CSIT," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2913–2924, 2016.

[14] D. Castanheira, A. Silva, and A. Gameiro, "Retrospective interference alignment: Degrees of freedom scaling with distributed transmitters," *IEEE Transactions on Information Theory*, vol. 63, pp. 1721–1730, March 2017.

[15] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6093–6130, 2014.

[16] S. A. Jafar, "Blind interference alignment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 216–227, 2012.

[17] M. Johnny and M. R. Aref, "BIA for the k-user interference channel using reconfigurable antenna at receivers," *IEEE Transactions on Information Theory*, pp. 1–1, 2019.

[18] O. El Ayach, S. W. Peters, and R. W. Heath, "The practical challenges of interference alignment," *IEEE Wireless Communications*, vol. 20, no. 1, pp. 35–42, 2013.

[19] G. Caire and S. Shamai, "On achievable rates in a multi-antenna gaussian broadcast channel," in *Proceedings. 2001 IEEE International Symposium on Information Theory (IEEE Cat. No.01CH37252)*, pp. 147–, June 2001.

[20] S. A. Jafar, "Too much mobility limits the capacity of wireless ad hoc networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 3954–3965, Nov 2005.

[21] T. Gou, C. Wang, and S. A. Jafar, "Aiming perfectly in the dark-blind interference alignment through staggered antenna switching," *IEEE Transactions on Signal Processing*, vol. 59, pp. 2734–2744, June 2011.

[22] M. Johnny and M. R. Aref, "Sum degrees of freedom for the k-user interference channel using antenna switching," in *WSA 2017; 21th International ITG Workshop on Smart Antennas*, pp. 1–6, March 2017.

[23] C. Wang, "Degrees of freedom characterization: The 3-user siso interference channel with blind interference alignment," *IEEE Communications Letters*, vol. 18, pp. 757–760, May 2014.

[24] A. Vahid and M. Johnny, "Low-complexity practical implementation of blind interference alignment using reconfigurable antennas," *U.S. Patent no. 62,888,697*, Aug 2019.

[25] J. H. Van Lint, R. M. Wilson, and R. M. Wilson, *A course in combinatorics*. Cambridge university press, 2001.

[26] I. M. Gelfand, R. A. Silverman, *et al.*, *Calculus of variations*. Courier Corporation, 2000.

[27] M. Johnny and M. R. Aref, "A multi-layer encoding and decoding strategy for binary erasure channel," *IEEE Transactions on Information Theory*, vol. 65, pp. 4143–4151, July 2019.