Interference Alignment and Degrees of Freedom of the K User Interference Channel

Viveck R. Cadambe, Syed A. Jafar
Electrical Engineering and Computer Science
University of California Irvine,
Irvine, California, 92697, USA
Email: vcadambe@uci.edu, syed@uci.edu

Abstract

The main result of this paper is that in the fully connected K user wireless interference channel, each user can simultaneously achieve half of the capacity that he could have achieved in the absence of all interference. If we can control the channel coefficients then this result holds at any signal-to-noise-ratio (SNR). If the channel coefficients are randomly generated by nature and vary across channel uses then the result holds at high SNR, so that the K user interference channel has $K/2$ degrees of freedom. The result reveals the fallacy of the cake-cutting interpretation of orthogonal medium access because, essentially, it shows that everyone can get one half of the cake. The key to the result is the idea of interference alignment that takes advantage of the fact that every receiver sees a different alignment of signal vectors. Thus, signals are designed to cast overlapping shadows at the receivers where they constitute interference while they remain distinguishable at the receivers where they are desired.

Index Terms

Capacity, Degrees of Freedom, Interference Alignment, Interference Channel, MIMO, Multiplexing

V. Cadambe and S. Jafar are with the Center of Pervasive Communications and Computing (CPCC), Department of Electrical Engineering and Computer Science, University of California Irvine, Irvine, CA, 92697 USA e-mail: vcadambe@uci.edu, syed@uci.edu. This work was supported in part by the National Science Foundation under CAREER Grant 0546860 and by DARPA under ITMANET Grant UTA06-793. This work was presented in part at the 45th Annual Allerton Conference on Communications, Control and Computing, 2007.

I. INTRODUCTION

The capacity of wireless networks is the much sought after "Holy Grail" of network information theory [1]. Since exact capacity characterizations for most network communication scenarios appear intractable, the key to understanding the capacity limits is to pursue capacity approximations. The number of degrees of freedom (also known as the capacity pre-log or the multiplexing gain [2]) of a network provides a capacity approximation $C(\rho) = d\log(\rho) + o(\log(\rho))$, where d is the total number of degrees of freedom, $C(\rho)$ is the sum rate capacity and ρ is the signal-to-noise-power-ratio (SNR, or equivalently the total transmit power of all nodes while the local noise power at each node is normalized to unity). The accuracy of this approximation approaches 100% as ρ increases. By de-emphasizing local noise[1] relative to the signal (and interference) power, the degrees of freedom perspective is able to directly address the principal bottleneck on the capacity of wireless networks - the interference between concurrent transmissions. It is therefore a natural tool to explore the capacity of wireless networks. Degrees of freedom are known for centralized networks (multiple access [4] and broadcast channels [5]–[7]) and for the two user interference channel [8], [9]. However, prior to this work, the degrees of freedom for interference networks with more than 2 users have not been found. The best known outerbound states that the K user interference network cannot have more than $K/2$ degrees of freedom. It has been conjectured that the outerbound is loose and the interference network has only 1 degree of freedom [8] for any number of users. The unresolved gap between the inner and outer bounds highlights our lack of understanding of the capacity of wireless networks. It is this open problem that we pursue in this paper.

The Fallacy of the Cake-Cutting Interpretation of Orthogonal Medium Access

To gain some intuition into the degrees of freedom bounds for interference networks, let us first consider a simpler problem - designing an interference free system. Suppose we wish to design and operate a fully connected interference network in such a way that interference is completely avoided. We say a network is "fully connected" if all transmitters are heard by all receivers, i.e., all channel coefficients are non-zero. Since interference is not allowed, the users' channel access must be orthogonal. The traditional view of interference-free medium access is that the total channel resource should be divided among users so that each user gets a fraction of it and the sum of all these fractions is equal to one. Intuitively, the innerbound on the degrees of freedom reflects this cake-cutting view of orthogonal access as it states that the sum of the degrees of freedom achieved by all the users' rates cannot be more than one. The outerbound, on the other hand, points to a scenario where every user is able to access one half of the channel resource free from interference, thus producing a total of $K/2$ degrees of freedom. In terms of the cake-cutting analogy the outerbound implies a scenario where everyone gets half the cake. While it is easy to attribute this impossible scenario to a loose outerbound, we find that the problem lies not with the tightness of the outerbound but rather with the cake-cutting interpretation of orthogonal medium access allocation.

The main insight is conveyed through the example illustrated in Figure 1. Consider a wireless interference network that consists of K transmitters and K receivers, and where each transmitter has one independent message for its corresponding receiver. While the classical Gaussian interference channel model ignores signal propagation delays, for the purpose of this toy example let us depart from the classical model and explicitly consider propagation delays τ_{ij} between receiver j and transmitter i, $\forall i, j \in \{1, 2, \cdots, K\}$. In particular, let us assume that all the nodes are

[1]De-emphasizing the local noise is also the philosophy behind the deterministic channel model proposed in [3] which has been very successful in providing capacity approximations within a few bits for multiple access, broadcast, relay and two user interference channels.

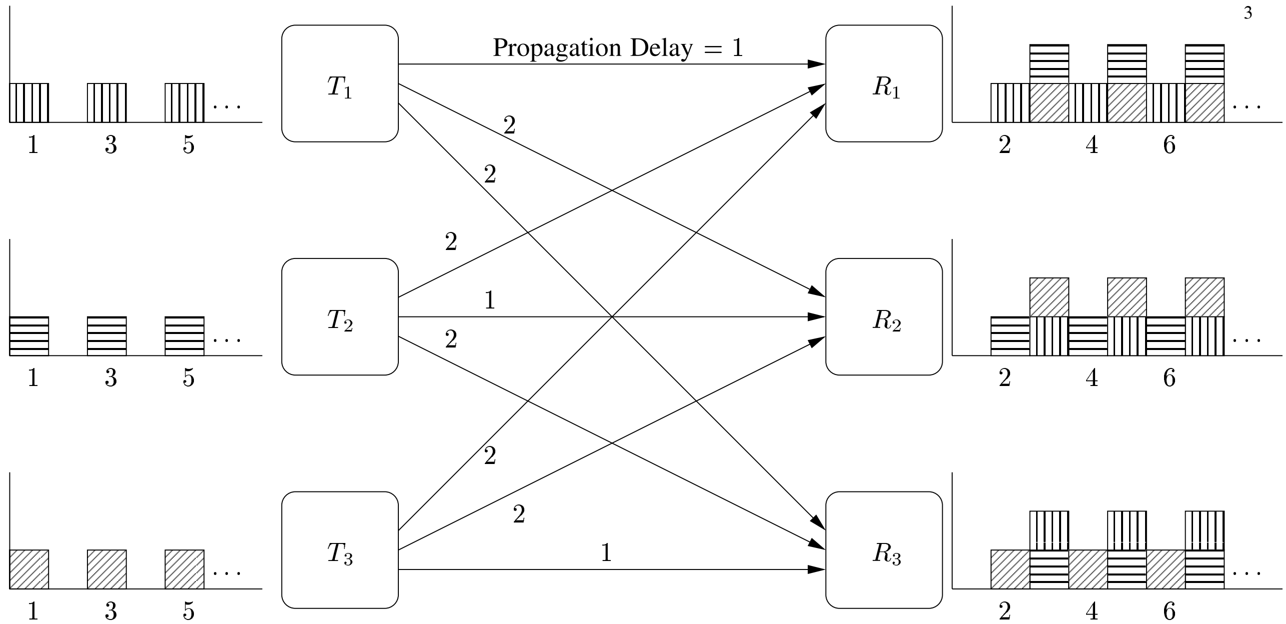

Fig. 1. Orthogonal Channel Access Example (Everyone gets half the cake)

located in such a way that the propagation delay is equal to one symbol duration for all desired signal paths and two symbol durations for all paths that carry interference signals. In other words,

$$\tau_{ij} = \begin{cases} 1, & i = j, \\ 2, & i \neq j, \end{cases}$$

where one unit of time corresponds to one symbol duration. On this channel, suppose each transmitter transmits only during odd time slots and is silent during the even time slots. Let us consider what happens at receiver 1. The symbols sent from its desired transmitter (transmitter 1) are received *free from interference* during the even time slots and all the undesired (interference) transmissions are received simultaneously during the odd time slots. Thus, surprisingly, each user is able to access the channel one-half of the time with no interference from other users. The fallacy of the cake-cutting interpretation of orthogonal channel allocation is obvious in this example as we are able to divide the "cake" in such a way that each user gets one-half of it, even when the number of users $K > 2$. The key to this counter-intuitive example is *interference alignment*.

Interference Alignment

The cake cutting analogy fails in the preceding example because of the *relativity of alignment* - i.e., the alignment of signal vector spaces is relative to the observer (the receiver). This is similar in spirit to the "relativity of simultaneity" in physics which states that the order of observed events depends on the observers' frame of reference[2]. Two transmitters may appear to be accessing the channel simultaneously to one receiver while they appear to be orthogonal to another receiver. Since each receiver has a different view, there exist scenarios where each receiver,

[2] We note, however, that in physics, the term "relativity of simultaneity" is now associated exclusively with the theory of relativity which concerns itself with frames of reference that are in motion with respect to each other and do not allow an absolute definition of time.

from its own perspective, appears to be privileged relative to others. The goal of interference alignment is to create such scenarios in a wireless network. Specifically, interference alignment refers to a construction of signals in such a manner that the they cast overlapping shadows at the receivers where they constitute interference while they remain distinguishable at the receivers where they are desired.

The idea of interference alignment evolved out of the degrees of freedom investigations on the 2 user MIMO X channel [10]–[12] and the compound broadcast channel [13]. The 2 user X channel is a communication system with 2 transmitters, 2 receivers and 4 independent messages, one from each transmitter to each receiver. Taking advantage of the multiple access channel (MAC) and the broadcast channel (BC) components contained within the X channel, Maddah-Ali, Motahari and Khandani proposed an elegant coding scheme (the MMK scheme) in [10] for the 2 user MIMO X channel. The MMK scheme naturally combines successive deoding and dirty paper coding, the optimal schemes for the constituent MAC and BC. Interestingly, the MMK scheme achieves $\lfloor \frac{4}{3}M \rfloor$ degrees of freedom on the 2 user X channel when all nodes are equipped with M antennas. The key to this result is the implicit interference alignment that is facilitated by the iterative optimization of transmit precoding and receive combining vectors. The first explicit interference alignment scheme is presented in [11] where it is shown that dirty paper coding and successive decoding are not required to achieve the maximum degrees of freedom on the 2 user MIMO X channel. The achievability of $\frac{4}{3}M$ degrees of freedom and the converse are established in [12]. Interference alignment is used in [12], [14] to obtain innerbounds on the degrees of freedom region of the MIMO X channel. Interference alignment is also a key ingredient of the degrees of freedom characterization of the compound broadcast channel in [13].

Degrees of Freedom of the K user Interference Channel

The main result of this paper is that the K user interference channel with single antennas at all nodes has $K/2$ degrees of freedom, when the channel coefficients vary across symbols and are drawn randomly from a continuous distribution. An interesting implication of this result is that:

"*At high SNR, every user in a wireless interference network is (simultaneously) able to achieve one half of the capacity that he could achieve in the absence of all interference.*"

The result shows that the capacity of wireless networks has been previously underestimated. For example, at high SNR the capacity is higher by 50%, 900%, and 4900% than prior belief for networks with 3, 20, and 100 interfering users, respectively. Interference is one of the principal challenges faced by wireless networks. However, we find that with perfect channel knowledge the time/frequency varying interference channel is not interference limited. In fact, after the first two users, additional users do not compete for degrees of freedom and each additional user is able to achieve $1/2$ degree of freedom without reducing the degrees of freedom of previously existing users. What makes this result even more remarkable is that linear scaling of degrees of freedom with users is achieved *without cooperation* in the form of message sharing that may allow MIMO behavior. If nodes have multiple antennas, and even if there is a different number of antennas at each node, we show that every user in the interference network will still achieve at least half of the capacity that he could achieve in the absence of all interference. Thus, the interference penalty is not more than half the degrees of freedom.

The result has the same flavor as the toy example with propagation delays that we presented earlier in this section. In both cases the conclusion is that *everyone gets half the cake.* However, there is a significant difference between the toy example and the degrees of freedom result. Note that in the toy example, we constructed an artificial channel where the delays were carefully selected to facilitate interference alignment. Even if there are no delays involved, one can construct a similar example by choosing the values of the channel coefficients. For example, if all the desired channels have real channel coefficients and all the interfering channel coefficients are purely imaginary, then all the transmitters can send real signals and the receivers will be able to discard all the interference simply by discarding the imaginary part of the received signal. The degrees of freedom result on the other hand is for channels whose coefficients are random, i.e. selected by nature so that we have no control over the channel coefficients. While interference alignment over random time varying channel coefficients is perhaps more interesting than the toy example, note that there is a penalty involved with the inability to control the channel coefficients. In the toy example, every user achieves half his interference-free capacity at *any* SNR. With random channel coefficients the users' rates suffer a penalty, but the penalty is $o(\log(\rho))$, i.e. it becomes a negligible fraction of the users' rates at high SNR. Indeed, we expect that the rate penalty will increase with the number of users, so that it will take higher and higher SNR to approach half of each user's capacity as the number of users increases. The degrees of freedom perspective is too coarse to capture this penalty and therefore does not reveal this competition among users. In this sense the picture presented by the degrees of freedom result is optimistic.

The achievability proofs in this paper are based on explicit constructions of interference alignment schemes. Similar to the case of the X channel in [12], interference alignment is achieved through joint beamforming in both space and time over multiple symbol extensions of the time varying channel. However, there are some unique aspects to the interference alignment schemes used in this paper. On the 2 user X channel, finite symbol extensions are sufficient to achieve the outerbound on the degrees of freedom. The interference alignment schemes constructed in this paper for the K user interference channel (with single antenna nodes) do not exactly achieve the outerbound on the degrees of freedom. Instead, by using longer symbol extensions we are able to approach arbitrarily close to the outerbound. Intuitively, this can be understood as follows. In order to achieve exactly $K/2$ degrees of freedom over a finite symbol extension, every receiver must be able to partition its observed signal space into two subspaces of equal size, one of which is meant for the desired signals and the other is the "waste basket" for all the interference terms. Moreover, the vector spaces corresponding to the interference contributed by all undesired transmitters must exactly align at every receiver within the waste basket which has the same size as each of the interference signals. It turns out this problem is overconstrained and does not admit a solution. We circumvent this problem by allowing some overflow space (a few extra symbols) for interference terms that do not align perfectly. Fortunately, we find that the size of the overflow space becomes a negligible fraction of the total number of dimensions as we increase the size of the signal space. Thus, for any $\epsilon > 0$ it is possible to align interference to the extent that the achieved degrees of freedom are within ϵ of the outerbound. The tradeoff is that the smaller the value of ϵ, the larger the number of symbols (time slots) needed to recover a fraction $1 - \epsilon$ of the outerbound value per symbol. As an example, consider the $K = 3$ user interference channel. We are able to achieve $3n + 1$ degrees of freedom over a $2n + 1$ symbol extension of the channel so that the degrees of freedom per symbol equal $\frac{3n+1}{2n+1}$, for any positive integer n. By choosing n large enough we can approach arbitrarily close to the outerbound of $3/2$ degrees of freedom. The case of $n = 1$ is shown in Figure 2. The figure illustrates how $3n + 1 = 4$ degrees of freedom are achieved over a $2n + 1 = 3$ symbol extension of the channel with $K = 3$ single antenna users, so that a total of $4/3$ degrees of freedom are achieved per channel use. User 1 achieves 2 degrees of freedom by transmitting two independently coded streams along the beamforming vectors $\mathbf{v}_1^{[1]}, \mathbf{v}_2^{[1]}$ while users 2 and 3 achieve one degree of freedom by sending their independently encoded data streams along the beamforming vectors $\mathbf{v}^{[2]}, \mathbf{v}^{[3]}$, respectively.

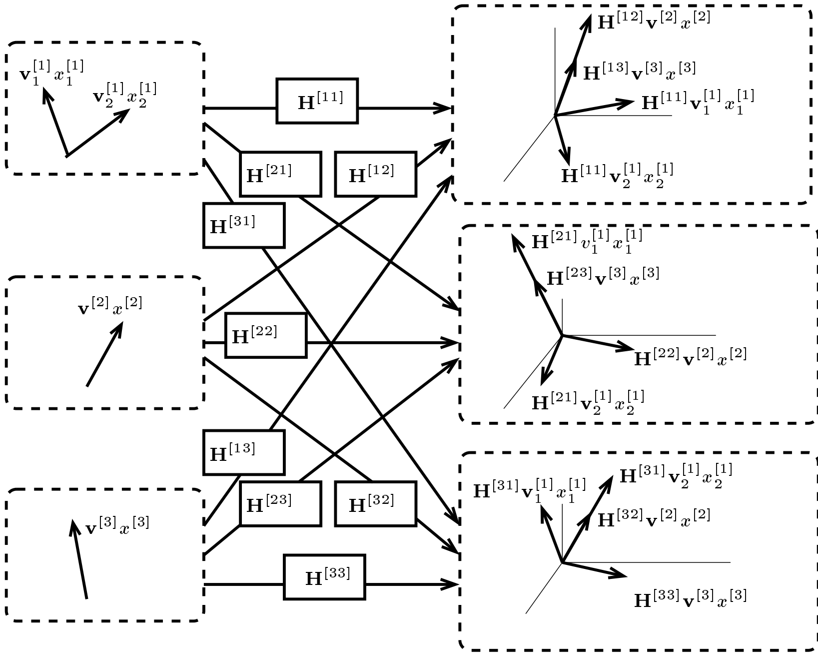

Fig. 2. Interference alignment on the 3 user interference channel to achieve $4/3$ degrees of freedom

Let us pick $\mathbf{v}^{[2]}$ be the 3×1 vector of all ones.

$$\mathbf{v}^{[2]} = \mathbf{1}_{3\times 1}.$$

The remaining beamforming vectors are chosen as follows.

- At receiver 1, the interference from transmitters 2 and 3 are perfectly aligned.

$$\mathbf{H}^{[12]}\mathbf{v}^{[2]} = \mathbf{H}^{[13]}\mathbf{v}^{[3]} \Rightarrow \mathbf{v}^{[3]} = \left(\mathbf{H}^{[13]}\right)^{-1}\mathbf{H}^{[12]}\mathbf{1}_{3\times 1}.$$

- At receiver 2, the interference from transmitter 3 aligns itself along one of the dimensions of the two-dimensional interference signal from transmitter 1.

$$\mathbf{H}^{[23]}\mathbf{v}^{[3]} = \mathbf{H}^{[21]}\mathbf{v}_1^{[1]} \Rightarrow \mathbf{v}_1^{[1]} = \left(\mathbf{H}^{[21]}\right)^{-1}\mathbf{H}^{[23]}\left(\mathbf{H}^{[13]}\right)^{-1}\mathbf{H}^{[12]}\mathbf{1}_{3\times 1}.$$

- Similarly, at receiver 3, the interference from transmitter 2 aligns itself along one of the dimensions of interference from transmitter 1.

$$\mathbf{H}^{[32]}\mathbf{v}^{[2]} = \mathbf{H}^{[31]}\mathbf{v}_2^{[1]} \Rightarrow \mathbf{v}_2^{[1]} = \left(\mathbf{H}^{[31]}\right)^{-1}\mathbf{H}^{[32]}\mathbf{1}_{3\times 1}.$$

Note that in order to deliver a capacity that grows as $\log(\rho)$, i.e., in order to carry one degree of freedom, it is not necessary for a beamforming vector to be orthogonal to the interference. It suffices if the beamforming vector is linearly independent of the basis vectors of the interference signal space. Also, note that the construction of beamforming vectors for interference alignment is not unique. For example, $\mathbf{v}^{[2]}$ could be any random vector instead of the all ones vector. Moreover, at receiver 2, the interference from transmitter 3, $\mathbf{H}^{[23]}\mathbf{v}^{[3]}$ does not necessarily have to align with one of the beams received from transmitter 1. It only needs to lie within the 2 dimensional space spanned by the two beams received from transmitter 1.

$$\mathbf{H}^{[23]}\mathbf{v}^{[3]} \in \text{span}\left[\mathbf{H}^{[21]}\mathbf{v}_1^{[1]} \qquad \mathbf{H}^{[21]}\mathbf{v}_2^{[1]}\right].$$

Similarly, at receiver 3 we only need

$$\mathbf{H}^{[32]}\mathbf{v}^{[2]} \in \text{span}\left[\mathbf{H}^{[31]}\mathbf{v}_1^{[1]} \qquad \mathbf{H}^{[31]}\mathbf{v}_2^{[1]}\right].$$

Since in this work our interest is only in the degrees of freedom we do not consider the optimization of beamforming vectors over these possibilities.

We make no claim that the interference alignment schemes used in this work achieve the maximum degrees of freedom for finite channel extensions or even that they are close to optimal except in the asymptotic sense of the degrees of freedom per orthogonal time and frequency dimension. The only purpose they serve is to show that as we allow the size of the signal vector space to grow larger (more frequency slots or more time slots), each user will be able to access nearly half of the total number of dimensions, free from interference, with a rate penalty that becomes negligible fraction of the users' rates at high SNR. A systematic construction of the most efficient interference alignment schemes over a fixed number of dimensions is an open problem that we believe to be one of the key steps to making interference alignment practical in a wireless network. Multiple antennas appear to be quite important in this regard. We show through an example that if multiple antennas are present at each node, then it may be possible to obtain more efficient interference alignment schemes. Specifically, we find that with $K = 3$ users, if all nodes have $M > 1$ antennas then the outerbound of $3M/2$ degrees of freedom can be exactly achieved with at most a 2 symbol extension. An investigation of interference alignment schemes with multiple antenna nodes for $K > 3$ users remains an interesting direction for future work.

II. System Model

Consider the K user interference channel, comprised of K transmitters and K receivers. We assume coding may occur over multiple orthogonal frequency and time dimensions and the rates as well as the degrees of freedom are normalized by the number of orthogonal time and frequency dimensions. Each node is equipped with only one antenna (multiple antenna nodes are considered later in this paper). The channel output at the k^{th} receiver over the f^{th} frequency slot and the t^{th} time slot is described as follows:

$$Y^{[k]}(f,t) \quad = \quad H^{[k1]}(f)X^{[1]}(f,t) + H^{[k2]}(f)X^{[2]}(f,t) + \cdots + H^{[kK]}(f)X^{[K]}(f,t) + Z^{[k]}(f,t),$$

where, $k \in \{1, 2, \cdots, K\}$ is the user index, $f \in \mathbb{N}$ is the frequency slot index, $t \in \mathbb{N}$ is the time slot index, $Y^{[k]}(f,t)$ is the output signal of the k^{th} receiver, $X^{[k]}(f,t)$ is the input signal of the k^{th} transmitter, $H^{[kj]}(f)$ is the channel fade coefficient from transmitter j to receiver k over the f^{th} frequency slot and $Z^{[k]}(f,t)$ is the additive white Gaussian noise (AWGN) term at the k^{th} receiver. The channel coefficients vary across frequency slots but are assumed constant in time. We assume all noise terms are i.i.d. (independent identically distributed) zero mean complex Gaussian with unit variance. We assume all channel coefficients $H^{[kj]}(f)$ are known to all transmitters and receivers. To avoid degenerate channel conditions (e.g. all channel coefficients are equal or channel coefficients are equal to either zero or infinity) we assume that the channel coefficient values are drawn i.i.d. from a continuous distribution and the absolute value of all the channel coefficients is bounded between a non-zero minimum value and a finite maximum value. Since the channel values are assumed constant in time, the time index t is sometimes suppressed for compact notation.

We assume that transmitters $1, 2, \cdots, K$ have independent messages $W_1, W_2, \cdots, W_K$ intended for receivers $1, 2, \cdots, K$, respectively. The total power across all transmitters is assumed to be equal to ρ per orthogonal time and frequency dimension. We indicate the size of the message set by $|W_i(\rho)|$. For codewords spanning $f_0 \times t_0$ channel uses (i.e. using f_0 frequency slots and t_0 time slots), the rates $R_i(\rho) = \frac{\log |W_i(\rho)|}{f_0 t_0}$ are achievable if the probability of error for all messages can be simultaneously made arbitrarily small by choosing an appropriately large $f_0 t_0$.

The capacity region $\mathcal{C}(\rho)$ of the three user interference channel is the set of all *achievable* rate tuples $\mathbf{R}(\rho) = (R_1(\rho), R_2(\rho), \cdots, R_K(\rho))$.

A. Degrees of Freedom

Similar to the degrees of freedom region definition for the MIMO X channel in [12] we define the degrees of freedom region $\mathcal{D}$ for the K user interference channel as follows:

$$\mathcal{D} = \Bigg\{ (d_1, d_2, \cdots, d_K) \in \mathbb{R}_+^K : \forall (w_1, w_2, \cdots, w_K) \in \mathbb{R}_+^K$$
$$w_1 d_1 + w_2 d_2 + \cdots + w_K d_K \leq \limsup_{\rho \to \infty} \left[\sup_{\mathbf{R}(\rho) \in \mathcal{C}(\rho)} [w_1 R_1(\rho) + w_2 R_2(\rho) + \cdots + w_K R_K(\rho)] \frac{1}{\log(\rho)} \right] \Bigg\}. \tag{1}$$

III. Degrees of Freedom for the K User Interference Channel

The following theorem presents the main result of this section.

Theorem 1: The number of degrees of freedom per orthogonal time and frequency dimension for the K user interference channel with single antennas at all nodes is $K/2$.

$$\max_{\mathcal{D}} d_1 + d_2 + \cdots + d_K \quad = \quad K/2. \tag{2}$$

A. Converse for Theorem 1

The converse argument for the theorem is a simple extension of the outerbounds presented in [8], [9] which are themselves based on Carleial's outerbound [15]. However, because we assume that the channel coefficients vary across frequency (or time) our model is different from these works which focus on constant channel coefficients. For the sake of completeness we derive the converse in this section.

The converse follows from the following lemma which provides an outerbound on the degrees of freedom region of the K user interference channel.

Lemma 1:

$$\max_{\mathcal{D}} d_i + d_j \leq \limsup_{\rho \to \infty} \sup_{\mathbf{R}(\rho) \in \mathcal{C}(\rho)} \frac{R_i(\rho) + R_j(\rho)}{\log(\rho)} \leq 1, \qquad \forall i, j \in \{1, 2, \cdots, K\}, i \neq j. \tag{3}$$

To obtain the converse result of Theorem 1, simply add all the inequalities from Lemma 1. This gives us:

$$\max_{\mathcal{D}} \sum_{i,j \in \{1,2,\cdots,K\}, i \neq j} (d_i + d_j) \quad \leq \quad \sum_{i,j \in \{1,2,\cdots,K\}, i \neq j} 1 \tag{4}$$
$$\Rightarrow \max_{\mathcal{D}} d_1 + d_2 + \cdots + d_K \quad \leq \quad K/2. \tag{5}$$

The proof of Lemma 1 (for the general case where all nodes have M antennas) is provided in Appendix I. A sketch of the proof is provided here. Without loss of generality, let us focus on case $i = 1, j = 2$. In order to obtain the corresponding outerbound, consider any reliable coding scheme for the K user interference channel. Now, suppose we eliminate messages $W_3, W_4, \cdots, W_K$, i.e., we use a pre-determined sequence of bits known to all the transmitters and receivers in place of these messages so that $R_3 = R_4 = \cdots = R_K = 0$. Then all receivers can subtract out the signals received from transmitters $3, 4, \cdots, K$. This is equivalent to a two user interference channel, where receiver 1 and 2 receive signals only from transmitters 1, 2 and decode messages W_1 and W_2 respectively. Next we argue that the two user interference channel with frequency selective channel coefficients (or even time varying channel coefficients) and a single antenna at each node can only have one degree of freedom when the coefficients are restricted to take values over a continuous set bounded away from zero and infinity. This argument proceeds as follows.

Let us provide message W_1 to receiver 2. Because receiver 2 has complete knowledge of all channel coefficients and the message W_2, we can eliminate the channel between transmitter 1 and receiver 2. Because the coding scheme is a reliable coding scheme by assumption, receiver 1 is also capable (with high probability) of decoding W_1, its desired message. In that case, we can also eliminate the channel from transmitter 1 to receiver 1. Then we end up with each receiver seeing only transmitter 2's signal with noise. For each channel use, we make sure that receiver 1 has the better channel by reducing noise variance if necessary. Thus, the signal at receiver 2 is a degraded version of the signal at receiver 1 and both receivers have the same side information (the message W_1 and knowledge of all the channel coefficients). We argue that if receiver 2 can decode its message W_2, receiver 1 must also be able to decode W_2 with a high probability. Finally, the closing argument is that since receiver 1 (with possibly reduced noise) is able to decode both messages W_1, W_2 for any reliable coding scheme, the rates $R_1(\rho), R_2(\rho)$ must lie in the capacity region of the multiple access channel from transmitters $1, 2$ to receiver 1 with reduced noise at the receiver. But since this receiver has only 1 antenna and reducing the noise variance (by a finite amount that depends only on the channel coefficients and not on the SNR ρ) does not affect the degrees of freedom, the total degrees of freedom cannot be more than 1 per orthogonal time and frequency dimension. This gives us the desired outerbound of (3) for the case $i = 1, j = 2$.

■

B. *Achievability Proof for Theorem 1 with $K = 3$*

The achievability proof is presented next. Since the proof is rather involved, we present first the constructive proof for $K = 3$. The proof for general $K \geq 3$ is then provided in Appendix II.

We show that $(d_1, d_2, d_3) = (\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$ lies in the degrees of freedom region $\forall n \in \mathbb{N}$. Since the degrees of freedom region is closed, this automatically implies that

$$\max_{(d_1,d_2,d_3)\in\mathcal{D}} d_1 + d_2 + d_3 \geq \sup_n \frac{3n+1}{2n+1} = \frac{3}{2}.$$

This result, in conjunction with the converse argument proves the theorem.

To show that $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$ lies in $\mathcal{D}$, we construct an interference alignment scheme using only $2n + 1$ frequency slots. We collectively denote the $2n + 1$ symbols transmitted over the first $2n + 1$ frequency slots at each time instant as a supersymbol. We call this the $(2n + 1)$ symbol extension of the channel. With the extended channel, the signal vector at the k^{th} user's receiver can be expressed as

$$\bar{\mathbf{Y}}^{[k]} = \bar{\mathbf{H}}^{[k1]}\bar{\mathbf{X}}^{[1]} + \bar{\mathbf{H}}^{[k2]}\bar{\mathbf{X}}^{[2]} + \bar{\mathbf{H}}^{[k3]}\bar{\mathbf{X}}^{[3]} + \bar{\mathbf{Z}}^{[k]}, \quad k \in \{1,2,3\}.$$

where $\bar{\mathbf{X}}^{[k]}$ is a $(2n+1)\times 1$ column vector representing the $2n+1$ symbol extension of the transmitted symbol $X^{[k]}$, i.e

$$\bar{\mathbf{X}}^{[k]}(t) \triangleq \begin{bmatrix} X^{[k]}(1,t) \\ X^{[k]}(2,t) \\ \vdots \\ X^{[k]}(2n+1,t) \end{bmatrix}.$$

Similarly $\bar{\mathbf{Y}}^{[k]}$ and $\bar{\mathbf{Z}}^{[k]}$ represent $2n+1$ symbol extensions of the $Y^{[k]}$ and $Z^{[k]}$ respectively. $\bar{\mathbf{H}}^{[kj]}$ is a diagonal $(2n+1)\times(2n+1)$ matrix representing the $2n+1$ symbol extension of the channel i.e

$$\bar{\mathbf{H}}^{[kj]} \triangleq \begin{bmatrix} H^{[kj]}(1) & 0 & \dots & 0 \\ 0 & H^{[kj]}(2) & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \cdots & H^{[kj]}(2n+1) \end{bmatrix}.$$

Recall that we assume that the channel coefficient values for each frequency slot are chosen independently from a continuous distribution. Thus, all the diagonal channel matrices $\bar{\mathbf{H}}^{[kj]}$ are comprised of all distinct diagonal elements with probability 1.

We show that $(d_1,d_2,d_3) = (n+1,n,n)$ is achievable on this extended channel implying that $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$ lies in the degrees of freedom region of the original channel.

In the extended channel, message W_1 is encoded at transmitter 1 into $n+1$ independent streams $x_m^{[1]}(t), m = 1,2,\dots,(n+1)$ sent along vectors $\mathbf{v}_m^{[1]}$ so that $\bar{\mathbf{X}}^{[1]}(t)$ is

$$\bar{\mathbf{X}}^{[1]}(t) = \sum_{m=1}^{n+1} x_m^{[1]}(t)\mathbf{v}_m^{[1]} = \bar{\mathbf{V}}^{[1]}\mathbf{X}^{[1]}(t),$$

where $\mathbf{X}^{[1]}(t)$ is a $(n+1)\times 1$ column vector and $\bar{\mathbf{V}}^{[1]}$ is a $(2n+1)\times(n+1)$ dimensional matrix. Similarly W_2 and W_3 are each encoded into n independent streams by transmitters 2 and 3 as $\mathbf{X}^{[2]}(t)$ and $\mathbf{X}^{[3]}(t)$ respectively.

$$\bar{\mathbf{X}}^{[2]}(t) = \sum_{m=1}^{n} x_m^{[2]}(t)\mathbf{v}_m^{[2]} = \bar{\mathbf{V}}^{[2]}\mathbf{X}^{[2]}(t),$$

$$\bar{\mathbf{X}}^{[3]}(t) = \sum_{m=1}^{n} x_m^{[3]}(t)\mathbf{v}_m^{[3]} = \bar{\mathbf{V}}^{[3]}\mathbf{X}^{[3]}(t).$$

The received signal at the i^{th} receiver can then be written as

$$\bar{\mathbf{Y}}^{[i]}(t) = \bar{\mathbf{H}}^{[i1]}\bar{\mathbf{V}}^{[1]}\mathbf{X}^{[1]}(t) + \bar{\mathbf{H}}^{[i2]}\bar{\mathbf{V}}^{[2]}\mathbf{X}^{[2]}(t) + \bar{\mathbf{H}}^{[i3]}\bar{\mathbf{V}}^{[3]}\mathbf{X}^{[3]}(t) + \bar{\mathbf{Z}}^{[i]}(t).$$

In this achievable scheme, receiver i eliminates interference by zero-forcing all $\bar{\mathbf{V}}^{[j]}, j \neq i$ to decode W_i. At receiver 1, $n+1$ desired streams are decoded after zero-forcing the interference to achieve $n+1$ degrees of freedom. To obtain $n+1$ interference free dimensions from a $2n+1$ dimensional received signal vector $\bar{\mathbf{Y}}^{[1]}(t)$, the dimension of the interference should be not more than n. This can be ensured by perfectly aligning the interference from transmitters 2 and 3 as follows.

$$\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]} = \bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}. \tag{6}$$

At the same time, receiver 2 zero-forces the interference from $\bar{\mathbf{X}}^{[1]}$ and $\bar{\mathbf{X}}^{[3]}$. To extract n interference-free dimensions from a $2n+1$ dimensional vector, the dimension of the interference has to be not more than $n+1$. i.e.

$$\text{rank}\left(\begin{bmatrix}\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]} & \bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}\end{bmatrix}\right) \leq n+1.$$

This can be achieved by choosing $\bar{\mathbf{V}}^{[3]}$ and $\bar{\mathbf{V}}^{[1]}$ so that

$$\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]} \prec \bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}, \tag{7}$$

where $\mathbf{P} \prec \mathbf{Q}$, means that the set of column vectors of matrix $\mathbf{P}$ is a subset of the set of column vectors of matrix $\mathbf{Q}$. Similarly, to decode W_3 at receiver 3, we wish to choose $\bar{\mathbf{V}}^{[2]}$ and $\bar{\mathbf{V}}^{[1]}$ so that

$$\bar{\mathbf{H}}^{[32]}\bar{\mathbf{V}}^{[2]} \prec \bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[1]} \tag{8}$$

Thus, we wish to pick vectors $\bar{\mathbf{V}}^{[1]}$, $\bar{\mathbf{V}}^{[2]}$ and $\bar{\mathbf{V}}^{[3]}$ so that equations (6), (7), (8) are satisfied. Note that the channel matrices $\bar{\mathbf{H}}^{[ij]}$ have a full rank of $2n+1$ almost surely. Since multiplying by a full rank matrix (or its inverse) does not affect the conditions represented by equations (6), (7) and (8), they can be equivalently expressed as

$$\begin{aligned} \mathbf{B} &= \mathbf{TC}, & (9)\\ \mathbf{B} &\prec \mathbf{A}, & (10)\\ \mathbf{C} &\prec \mathbf{A}, & (11) \end{aligned}$$

where

$$\begin{aligned} \mathbf{A} &= \bar{\mathbf{V}}^{[1]}, & (12)\\ \mathbf{B} &= (\bar{\mathbf{H}}^{[21]})^{-1}\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}, & (13)\\ \mathbf{C} &= (\bar{\mathbf{H}}^{[31]})^{-1}\bar{\mathbf{H}}^{[32]}\bar{\mathbf{V}}^{[2]}, & (14)\\ \mathbf{T} &= \bar{\mathbf{H}}^{[12]}(\bar{\mathbf{H}}^{[21]})^{-1}\bar{\mathbf{H}}^{[23]}(\bar{\mathbf{H}}^{[32]})^{-1}\bar{\mathbf{H}}^{[31]}(\bar{\mathbf{H}}^{[13]})^{-1}. & (15) \end{aligned}$$

Note that $\mathbf{A}$ is a $(2n+1)\times(n+1)$ matrix. $\mathbf{B}$ and $\mathbf{C}$ are $(2n+1)\times n$ matrices. Since all channel matrices are invertible, we can choose $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ so that they satisfy equations (9)-(11) and then use equations (12)-(15) to find $\bar{\mathbf{V}}^{[1]}$,$\bar{\mathbf{V}}^{[2]}$ and $\bar{\mathbf{V}}^{[3]}$. $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are picked as follows. Let $\mathbf{w}$ be the $(2n+1)\times 1$ column vector

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

We now choose $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ as:

$$\begin{aligned} \mathbf{A} &= [\mathbf{w} \;\; \mathbf{Tw} \;\; \mathbf{T}^2\mathbf{w} \;\; \dots \;\; \mathbf{T}^n\mathbf{w}],\\ \mathbf{B} &= [\mathbf{Tw} \;\; \mathbf{T}^2\mathbf{w} \;\; \dots \;\; \mathbf{T}^n\mathbf{w}],\\ \mathbf{C} &= [\mathbf{w} \;\; \mathbf{Tw} \;\; \dots \;\; \mathbf{T}^{n-1}\mathbf{w}]. \end{aligned}$$

It can be easily verified that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ satisfy the three equations (9)-(11). Therefore, $\bar{\mathbf{V}}^{[1]}$, $\bar{\mathbf{V}}^{[2]}$ and $\bar{\mathbf{V}}^{[3]}$ satisfy the interference alignment equations in (6), (7) and (8).

Now, consider the received signal vectors at Receiver 1. The desired signal arrives along the $n+1$ vectors $\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]}$ while the interference arrives along the n vectors $\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}$ and the n vectors $\bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}$. As enforced by equation (6) the interference vectors are perfectly aligned. Therefore, in order to prove that there are $n+1$ interference free dimensions it suffices to show that the columns of the square, $(2n+1)\times(2n+1)$ dimensional matrix

$$\left[\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]} \quad \bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}\right]. \tag{16}$$

are linearly independent almost surely. Multiplying by the full rank matrix $(\bar{\mathbf{H}}^{[11]})^{-1}$ and substituting the values of $\bar{\mathbf{V}}^{[1]}, \bar{\mathbf{V}}^{[2]}$, equivalently we need to show that almost surely

$$\mathbf{S} \stackrel{\triangle}{=} \left[\mathbf{w} \;\; \mathbf{T}\mathbf{w} \;\; \mathbf{T}^2\mathbf{w} \;\; \ldots \;\; \mathbf{T}^n\mathbf{w} \;\; \mathbf{D}\mathbf{w} \;\; \mathbf{D}\mathbf{T}\mathbf{w} \;\; \mathbf{D}\mathbf{T}^2\mathbf{w} \;\; \ldots \;\; \mathbf{D}\mathbf{T}^{n-1}\mathbf{w}\right] \tag{17}$$

has linearly independent column vectors where $\mathbf{D} = (\bar{\mathbf{H}}^{[11]})^{-1}\bar{\mathbf{H}}^{[12]}$ is a diagonal matrix. In other words, we need to show $\det(\mathbf{S}) \neq 0$ with probability 1. The proof is obtained by contradiction. If possible, let $\mathbf{S}$ be singular with non-zero probability. i.e, $\Pr(|\mathbf{S}| = 0) > 0$. Further, let the diagonal entries of $\mathbf{T}$ be $\lambda_1, \lambda_2, \ldots \lambda_{2n+1}$ and the diagonal entries of $\mathbf{D}$ be $\kappa_1, \kappa_2 \ldots \kappa_{2n+1}$. Then the following equation is true with non-zero probability.

$$|\mathbf{S}| = \begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^n & \kappa_1 & \kappa_1\lambda_1 & \ldots & \kappa_1\lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^n & \kappa_2 & \kappa_2\lambda_2 & \ldots & \kappa_2\lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{2n+1} & \lambda_{2n+1}^2 & \ldots & \lambda_{2n+1}^n & \kappa_{2n+1} & \kappa_{2n+1}\lambda_{2n+1} & \ldots & \kappa_{2n+1}\lambda_{2n+1}^{n-1} \end{vmatrix} = 0.$$

Let C_{ij} indicate the cofactor of the ith row and jth column of $|\mathbf{S}|$. Expanding the determinant along the first row, we get

$$|\mathbf{S}| = 0 \Rightarrow C_{11} + \lambda_1 C_{12} + \ldots \lambda_1^n C_{1(n+1)} + \kappa_1\left[C_{1(n+2)} + \lambda_1 C_{1(n+3)} + \ldots + \lambda_1^{n-1} C_{1(2n+1)}\right] = 0.$$

None of 'co-factor' terms C_{1j} in the above expansion depend λ_1 and κ_1. If all values other than κ_1 are given, then the above is a linear equation in κ_1. Now, $|\mathbf{S}| = 0$ implies one of the following two events

1) κ_1 is a root of the linear equation.
2) All the coefficients forming the linear equation in κ_1 are equal to 0, so that the singularity condition is trivially satisfied for all values of κ_1.

Since κ_1 is a random variable drawn from a continuous distribution, the probability of κ_1 taking a value which is equal to the root of this linear equation is zero. Therefore, the second event happens with probability greater than 0 and we can write,

$$\Pr\left(|\mathbf{S}| = 0\right) > 0 \quad \Rightarrow \quad \Pr(C_{1(n+2)} + \lambda_1 C_{1(n+3)} + \ldots + \lambda_1^{n-1} C_{1(2n+1)} = 0) > 0.$$

Consider the equation

$$C_{1(n+2)} + \lambda_1 C_{1(n+3)} + \ldots + \lambda_1^{n-1} C_{1(2n+1)} = 0.$$

Since the terms C_{1j} do not depend on λ_1, the above equation is a polynomial of degree n in λ_1. Again, as before, there are two possibilities. The first possibility is that λ_1 takes a value equal to one of the n roots of the above equation. Since λ_1 is drawn from a continuous distribution, the probability of this event happening is zero. The second possibility is that all the coefficients of the above polynomial are zero with non-zero probability and we can write

$$\Pr(C_{1(n+2)} + \ldots + \kappa_1 \lambda_1^n C_{1(2n+1)} = 0) > 0 \Rightarrow \Pr(C_{1(2n+1)} = 0) > 0.$$

We have now shown that if the determinant of the $(2n+1) \times (2n+1)$ matrix $\mathbf{S}$ is equal to 0 with non-zero probability, then the determinant of following $2n \times 2n$ matrix (obtained by stripping off the first row and last column of $\mathbf{S}$) is equal to 0 with non-zero probability.

$$\det \begin{bmatrix} 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^n & \kappa_2 & \kappa_2\lambda_2 & \ldots & \kappa_2\lambda_2^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{2n+1} & \lambda_{2n+1}^2 & \ldots & \lambda_{2n+1}^n & \kappa_{2n+1} & \kappa_{2n+1}\lambda_{2n+1} & \ldots & \kappa_{2n+1}\lambda_{2n+1}^{n-2} \end{bmatrix} = 0$$

with probability greater than 0. Repeating the above argument and eliminating the first row and last column at each stage we get

$$\det \begin{bmatrix} 1 & \lambda_{n+1} & \lambda_{n+1}^2 & \ldots & \lambda_{n+1}^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{2n+1} & \lambda_{2n+1}^2 & \ldots & \lambda_{2n+1}^n \end{bmatrix} = 0$$

with probability greater than 0. But this is a Vandermonde matrix and its determinant

$$\prod_{n+1 \leq i < j \leq 2n+1} (\lambda_i - \lambda_j)$$

is equal to 0 only if $\lambda_i = \lambda_j$ for some $i \neq j$. Since λ_i are drawn independently from a continuous distribution, they are all distinct almost surely. This implies that $\Pr(|\mathbf{S}| = 0) = 0$.

Thus, the $n+1$ vectors carrying the desired signal at receiver 1 are linearly independent of the n interference vectors which allows the receiver to zero force interference and obtain $n+1$ interference free dimensions, and therefore $n+1$ degrees of freedom for its message.

At receiver 2 the desired signal arrives along the n vectors $\bar{\mathbf{H}}^{[22]}\bar{\mathbf{V}}^{[2]}$ while the interference arrives along the $n+1$ vectors $\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}$ and the n vectors $\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}$. As enforced by equation (7) the interference vectors $\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}$ are perfectly aligned within the interference vectors $\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}$. Therefore, in order to prove that there are n interference free dimensions at receiver 2 it suffices to show that the columns of the square, $(2n+1) \times (2n+1)$ dimensional matrix

$$\begin{bmatrix} \bar{\mathbf{H}}^{[22]}\bar{\mathbf{V}}^{[2]} & \bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]} \end{bmatrix} \tag{18}$$

are linearly independent almost surely. This proof is quite similar to the proof presented above for receiver 1 and is therefore omitted to avoid repetition. Using the same arguments we can show that both receivers 2 and 3 are able to zero force the $n+1$ interference vectors and obtain n interference free dimensions for their respective desired signals so that they each achieve n degrees of freedom.

Thus we established the achievability of $d_1 + d_2 + d_3 = \frac{3n+1}{2n+1}$ for any n. This scheme, along with the converse automatically imply that

$$\sup_{(d_1,d_2,d_3) \in \mathcal{D}} d_1 + d_2 + d_3 = \frac{3}{2}.$$

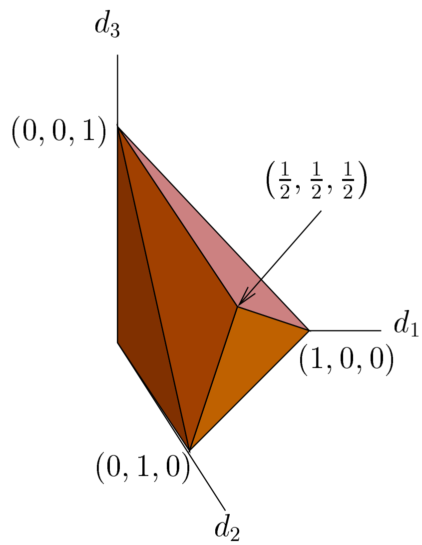

Fig. 3. Degrees of Freedom Region for the 3 user interference channel

C. The Degrees of Freedom Region for the 3 *User Interference Channel*

Theorem 2: The degrees of freedom region of the 3 user interference channel is characterized as follows:

$$\begin{aligned}\mathcal{D} = \{(d_1, d_2, d_3) : &\\ d_1 + d_2 &\leq 1\\ d_2 + d_3 &\leq 1\\ d_1 + d_3 &\leq 1\}.\end{aligned} \tag{19}$$

Proof: The converse argument is identical to the converse argument for Theorem 1 and is therefore omitted. We show achievability as follows. Let $\mathcal{D}'$ be the degrees of freedom region of the 3 user interference channel. We need to prove that $\mathcal{D}' = \mathcal{D}$. We show that $\mathcal{D} \subset \mathcal{D}'$ which along with the converse proves the stated result.

The points $K = (0,0,1)$, $L = (0,1,0)$, $J = (1,0,0)$ can be verified to lie in $\mathcal{D}'$ through trivial achievable schemes. Also, Theorem 1 implies that $N = (\frac{1}{2},\frac{1}{2},\frac{1}{2})$ lies in $\mathcal{D}'$ (Note that this is the only point which achieves a total of $\frac{3}{2}$ degrees of freedom and satisfies the inequalities in (19). Consider any point $(d_1,d_2,d_3) \in \mathcal{D}$ as defined by the statement of the theorem. The point (d_1,d_2,d_3) can then be shown to lie in a convex region whose corner points are $(0,0,0)$, J, K, L and N. i.e (d_1,d_2,d_3) can be expressed as a convex combination of the end points (see Fig. 3).

$$(d_1,d_2,d_3) = \alpha_1(1,0,0) + \alpha_2(0,1,0) + \alpha_3(0,0,1) + \alpha_4(\frac{1}{2},\frac{1}{2},\frac{1}{2}) + \alpha_5(0,0,0),$$

where the constants α_i are defined as follows.

	α_1	α_2	α_3	α_4	α_5
$d_1 + d_2 + d_3 \leq 1$	d_1	d_2	d_3	0	$1 - d_1 - d_2 - d_3$
$d_1 + d_2 + d_3 > 1$	$\frac{d_1-d_2-d_3+1}{2}$	$\frac{d_2-d_1-d_3+1}{2}$	$\frac{d_3-d_1-d_2+1}{2}$	$d_1 + d_2 + d_3 - 1$	0

It is easily verified that the values of α_i are non-negative for all $(d_1,d_2,d_3) \in \mathcal{D}$ and that they add up to one. Thus, all points in $\mathcal{D}$ are convex combinations of achievable points J, K, L, N and $(0,0,0)$. Since convex combinations are achievable by time sharing between the end points, this implies that $\mathcal{D} \subset \mathcal{D}^{'}$. Together with the converse, we have $\mathcal{D} = \mathcal{D}^{'}$ and the proof is complete. ∎

Note that the proof presented above uses coding over multiple frequency slots where the channel coefficients take distinct values. We now examine the possible ramifications of this assumption both from a theoretical as well as a practical perspective.

The assumption of frequency selective channels is intriguing because it is not clear if $K/2$ degrees of freedom will be achieved with constant channels that do not vary across time or frequency slots. Therefore the validity of the Host-Madsen-Nosratinia conjecture [8] remains partly undetermined. On the one hand the number of degrees of freedom is a discontinuous measure as evident from the point to point channel where it represents the rank of the channel matrix. Therefore the constant coefficient case may be of limited significance. On the other hand, the constant channel case may shed light on novel interference alignment schemes such as interference alignment in the signal "level" dimension as demonstrated for certain special cases in [16] and interference alignment through lattice codes as demonstrated for the one-sided interference channel in [17].

Next we discuss the relationship between degrees of freedom and an $\mathcal{O}(1)$ capacity characterization.

D. The $\mathcal{O}(1)$ Capacity Approximation

The degrees of freedom d provide a capacity approximation that is accurate within $o(\log(\rho))$, i.e.,

$$C(\rho) = d\log(\rho) + o(\log(\rho)). \tag{20}$$

In general, a capacity approximation within $\mathcal{O}(1)$ is more accurate than an approximation within $o(\log(\rho))$. However, it turns out that in many cases the two are directly related. For example, it is well known that for the full rank MIMO channel with M input antennas and N output antennas, transmit power ρ and i.i.d. zero mean unit variance additive white Gaussian noise (AWGN) at each receiver, the capacity $C(\rho)$ may be expressed as:

$$C(\rho) = \min(M, N)\log(1+\rho) + \mathcal{O}(1) = d\log(1+\rho) + \mathcal{O}(1). \tag{21}$$

A similar relationship between the degrees of freedom and the $\mathcal{O}(1)$ capacity characterization also holds for the MIMO multiple access channel, the MIMO broadcast channel, and the two user MIMO interference channel. For the MIMO MAC and BC, the outerbound on sum capacity obtained from full cooperation among the distributed nodes is $d\log(1+\rho)+\mathcal{O}(1)$. The innerbound obtained from zero forcing is also $d\log(1+\rho)+\mathcal{O}(1)$ so that we can write $C(\rho) = d\log(1+\rho)+\mathcal{O}(1)$. For the two user MIMO interference channel and the 2 user MIMO X channel the outerbound is obtained following an extension of Carleial's outerbound [15] which results in a MIMO MAC channel. The innerbound is obtained from zero forcing. Since both of these bounds are within $\mathcal{O}(1)$ of $d\log(1+\rho)$ we can similarly write $C(\rho) = d\log(1+\rho)+\mathcal{O}(1)$. However, consider the K user interference channel with single antennas at each node. In this case we have only shown:

$$(K/2-\epsilon)\log(1+\rho) + \mathcal{O}(1) \leq C(\rho) \leq (K/2)\log(1+\rho) + \mathcal{O}(1),\ \forall \epsilon > 0. \tag{22}$$

Consider a hypothetical capacity function $C(\rho) = K/2\log(1+\rho) - c\sqrt{\log(1+\rho^2)}$. Such a capacity function would also satisfy the inner and outerbounds provided above for the K user interference channel and has $D = K/2$ degrees of freedom. However, this hypothetical capacity function does not have a $\mathcal{O}(1)$ capacity characterization equal to $\overline{C}(\rho) = K/2\log(1+\rho)$ as the difference between $C(\rho)$ and $\overline{C}(\rho)$ is unbounded. To claim that the $\mathcal{O}(1)$ capacity of the K user interference channel is $(K/2)\log(1+\rho)$ we need to show an innerbound of $(K/2)\log(1+\rho)+\mathcal{O}(1)$.

Since our achievable schemes are based on interference alignment and zero forcing, the natural question to ask is whether an interference alignment and zero forcing based scheme can achieve exactly $K/2$ degrees of freedom. The following explanation uses the $K = 3$ case to suggest that the answer is negative.

Consider an achievable scheme that uses a M symbol extension of the channel. Now, consider a point $(\alpha_1, \alpha_2, \alpha_3)$ that can be achieved over this extended channel using interference alignment and zero-forcing alone. If possible, let the total degrees of freedom over this extended channel be $3M/2$. i.e. $\alpha_1 + \alpha_2 + \alpha_3 = 3M/2$. It can be argued along the same lines as the converse part of Theorem 1 that (α_i, α_j) is achievable in the 2 user interference channel for $\forall (i, j) \in \{(1, 2), (2, 3), (3, 1)\}$. Therefore

$$\alpha_1 + \alpha_2 \leq M,$$

$$\alpha_2 + \alpha_3 \leq M,$$

$$\alpha_1 + \alpha_3 \leq M.$$

It can be easily seen that the only point $(\alpha_1, \alpha_2, \alpha_3)$ that satisfies the above inequalities and achieves a total of $3M/2$ degrees of freedom is $(\frac{M}{2}, \frac{M}{2}, \frac{M}{2})$. Therefore, any scheme that achieves a total of $3M/2$ degrees of freedom over the extended channel achieves the point $(\frac{M}{2}, \frac{M}{2}, \frac{M}{2})$.

We assume that the messages W_i are encoded along $M/2$ independent streams similar to the coding scheme in the proof of Theorem 1 i.e.

$$\bar{\mathbf{X}}^{[i]} = \sum_{m=1}^{M/2} x_m^{[i]} \mathbf{v}_m^{[i]} = \bar{\mathbf{V}}^{[i]} \mathbf{X}^{[i]}.$$

Now, at receiver 1, to decode an $M/2$ dimensional signal using zero-forcing, the dimension of the interference has to be at most $M/2$. i.e.,

$$\text{rank}[\bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]} \quad \bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}] = M/2. \tag{23}$$

Note that since $\bar{\mathbf{V}}^{[2]}$ has $M/2$ linearly independent column vectors and $\bar{\mathbf{H}}^{[12]}$ is full rank with probability 1, $\text{rank}(\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}) = M/2$. Similarly the dimension of the interference from transmitter 3 is also equal to $M/2$. Therefore, the two vector spaces on the left hand side of equation (23) must have full intersection, i.e.,

$$\text{span}(\bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}) = \text{span}(\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}), \tag{24}$$

$$\text{span}(\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}) = \text{span}(\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}) \text{ (At receiver 2) }, \tag{25}$$

$$\text{span}(\bar{\mathbf{H}}^{[32]}\bar{\mathbf{V}}^{[2]}) = \text{span}(\bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[1]}) \text{ (At receiver 3) }, \tag{26}$$

where $\text{span}(\mathbf{A})$ represents the space spanned by the column vectors of matrix $\mathbf{A}$ The above equations imply that

$$\text{span}(\bar{\mathbf{H}}^{[13]}(\bar{\mathbf{H}}^{[23]})^{-1}\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}) = \text{span}(\bar{\mathbf{H}}^{[12]}(\bar{\mathbf{H}}^{[32]})^{-1}\bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[1]}),$$

$$\Rightarrow \text{span}(\bar{\mathbf{V}}^{[1]}) = \text{span}(\mathbf{T}\bar{\mathbf{V}}^{[1]}),$$

where $\mathbf{T} = (\bar{\mathbf{H}}^{[13]})^{-1}\bar{\mathbf{H}}^{[23]}(\bar{\mathbf{H}}^{[21]})^{-1}\bar{\mathbf{H}}^{[12]}(\bar{\mathbf{H}}^{[32]})^{-1}\bar{\mathbf{H}}^{[31]}$. The above equation implies that there exists at least one eigenvector $\mathbf{e}$ of $\mathbf{T}$ in $\text{span}(\bar{\mathbf{V}}^{[1]})$. Note that since all channel matrices are diagonal, the set of eigenvectors of all channel matrices, their inverses and their products are all identical to the set of column vectors of the identity matrix. i.e vectors of the form $[0\ 0\ \ldots\ 1\ \ldots\ 0]^T$. Therefore $\mathbf{e}$ is an eigenvector for all channel matrices. Since $\mathbf{e}$ lies in span($\bar{\mathbf{V}}^{[1]}$), equations (24)-(26) imply that,

$$\begin{aligned} \mathbf{e} &\in \text{span}(\bar{\mathbf{H}}^{[ij]}\bar{\mathbf{V}}^{[i]}), \forall i,j \in \{1,2,3\}, \\ \Rightarrow \mathbf{e} &\in \text{span}(\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]}) \cap \text{span}(\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}). \end{aligned}$$

Therefore, at receiver 1, the desired signal $\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]}$ is *not* linearly independent with the interference $\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[2]}$. Therefore, receiver 1 cannot decode W_1 completely by merely zero-forcing the interference signal. Evidently, interference alignment in the manner described above cannot achieve exactly $3/2$ degrees of freedom on the 3 user interference channel with a single antenna at all nodes.

We explore this interesting aspect of the 3 user interference channel further in the context of multiple antenna nodes. Our goal is to find out if exactly $3M/2$ degrees of freedom may be achieved with M antennas at each node. As shown by the following theorem, indeed we can achieve exactly $3M/2$ degrees of freedom so that the $\mathcal{O}(1)$ capacity characterization for $M > 1$ is indeed related to the degrees of freedom as $\overline{C}(\rho) = (3M/2)\log(1+\rho)$.

IV. Degrees of Freedom for the Interference Channel with Multiple Antenna Nodes

A. *The 3 user interference channel with $M > 1$ antennas at each node and constant channel coefficients*

The 3 user MIMO interference channel is interesting because in this case we show that we can achieve exactly $3M/2$ degrees of freedom with *constant* channel matrices, i.e., multiple frequency slots are not required. This gives us a lowerbound on sum capacity of $3M/2\log(1+\rho)+\mathcal{O}(1)$. Since the outerbound on sum capacity is also $3M/2\log(1+\rho)+\mathcal{O}(1)$ we have an $\mathcal{O}(1)$ approximation to the capacity of the 3 user MIMO interference channel with $M > 1$ antennas at all nodes.

Theorem 3: In a 3 user interference channel with $M > 1$ antennas at each transmitter and each receiver and constant coefficients, the sum capacity $C(\rho)$ may be characterized as:

$$C(\rho) = (3M/2)\log(1+\rho) + \mathcal{O}(1). \tag{27}$$

The outerbound follows directly from [9] which shows that the 2 user interference channel with M antennas at each node and constant channel coefficients (only one frequency slot) has only M degrees of freedom. In the 3 user case, we eliminate one message at a time to obtain inequalities $d_1 + d_2 \leq M, d_2 + d_3 \leq M, d_1 + d_3 \leq M$. Adding up all three inequalities we obtain the converse.

The proof is presented in Appendices III and IV.

B. *The K user Interference Channel with Multiple Antenna Nodes*

Theorem 3 in the preceding section shows that $3M/2$ degrees of freedom are achievable on the 3 user interference channel with $M > 1$ antennas at each node using only *one* frequency slot. It is not known if the result can be extended to $K > 3$ users. However, by coding over multiple frequency and time slots it is easy to find the degrees of freedom for the K user interference channel with M antennas at each node. The result follows directly from Theorem 1 and is presented in the following Corollary.

Corollary 1: The K user interference channel with M antennas at each node has $KM/2$ spatial degrees of freedom per orthogonal time and frequency dimension.

Proof: The converse for Corollary 1 is already derived in Appendix 1 in equation (29).

Achievability of Corollary 1 is also straightforward. Suppose we view each of the M co-located antennas at a node as a separate node. In other words we do not allow joint processing of signals obtained from the co-located antennas. Then, instead of a K user interference channel with M antenna nodes we obtain a KM user interference channel with single antenna nodes. But the result of Theorem 1 establishes that $KM/2$ degrees of freedom are achievable on this interference channel. Thus, we can also achieve $KM/2$ degrees of freedom per orthogonal time and frequency dimension on the K user interference channel with M antenna nodes. ■

Lastly, let us consider the most general K user interference channel where each node is equipped with possibly different number of antennas. In this case also a lower bound on the degrees of freedom is directly established from the result of Theorem 1. The following Corollary states this result.

Corollary 2: The total degrees of freedom per orthogonal time and frequency dimension for the K user interference channel where transmitter i has M_i^T antennas and receiver i has M_i^R antennas $\forall i \in \{1, 2, \cdots, K\}$ is bounded below as:

$$d_1 + d_2 + \cdots + d_K \geq \frac{1}{2} \sum_{i=1}^{K} \min(M_i^T, M_i^R). \tag{28}$$

Thus, no more than half the degrees of freedom are lost on the K user interference channel with multiple antenna nodes.

Proof: The achievability proof is straightforward as, once again, the i^{th} transmitter receiver pair can be replaced with $\min(M_i^T, M_i^R)$ single antenna transmitter and receiver nodes by only allowing distributed processing of signals at each antenna and discarding the remaining antennas. Thus, we obtain an interference channel with $\sum_{i=1}^{K} \min(M_i^T, M_i^R)$ transmitters and receivers, each equipped with only a single antenna. The achievability of $\frac{1}{2} \sum_{i=1}^{K} \min(M_i^T, M_i^R)$ degrees of freedom on this interference channel then follows from the result of Theorem 1. ■

Note that Corollary 2 only establishes an innerbound and is not always tight. For example, consider the 2 user interference channel where each transmitter has 2 antennas while each receiver has only 1 antenna. While Corollary 2 only shows achievability of 1 degree of freedom for this channel, it is known that this interference channel has 2 degrees of freedom [9]. However, Corollary 2 is interesting because it shows that interference cannot reduce the degrees of freedom of the interference channel by more than half compared to when each transmitter and receiver is able to operate without interference from other users.

V. Conclusion

We have shown that with perfect channel knowledge the K user interference channel has (almost surely) $K/2$ degrees of freedom when the fading coefficients are generated from a continuous distribution. Conventional wisdom has so far been consistent with the Host-Madsen-Nosratinia conjecture that distributed interfering systems cannot have more than 1 degree of freedom and therefore the best known outerbound of $K/2$ has not been considered significant. This pessimistic outlook has for long invited researchers to try to prove that more than 1 degree of freedom is not possible while ignoring the $K/2$ outerbound. The present result shifts the focus onto the outerbound by proving that it is tight for fading channels if perfect and global channel knowledge is available. Thus, the present result could guide future research along an optimistic path in the same manner that MIMO technology has shaped our view of the capacity of a wireless channel.

While this work shows the potential benefits of interference alignment, several challenges must be overcome before these benefits translate into practice. One key issue is the assumption of global channel knowledge. While a node may acquire channel state information for its own channels, it is much harder to learn the channels between other pairs of nodes with which this node is not directly associated. On the other hand, global channel knowledge may not be necessary if there is a feedback channel through which the receivers can guide the transmitters into aligned configurations in real time by applying incremental corrections.

Another issue is the high SNR nature of the degrees of freedom result. As mentioned earlier, the high SNR restriction is the penalty paid for the inability to control the channel coefficients. In this context it is important to investigate if there is a tradeoff between the degrees of freedom and the power offset. In other words, by sacrificing some of the degrees of freedom it may be possible to gain better control of the effective channels. For example, extra antennas can be used to create controlled fluctuations in the effective channels in the manner of opportunistic beamforming [18]. While the pre-log term is the principal determinant of capacity as SNR approaches infinity, $O(1)$ terms are quite significant at SNR values of practical interest. Therefore, sacrificing some degrees of freedom (the pre-log term) in exchange for $O(1)$ improvements (less power offset) may be necessary in practice.

The need for long symbol extensions is another potential problem that needs to be overcome for interference alignment schemes to become practical. For example, if we think in terms of frequency slots (e.g. OFDM carriers), the number of frequency slots that we need to use in order for each user to recover close to half of his interference free capacity (per frequency slot) can be too high. There are several possibilities to overcome this challenge. First, one may construct supersymbols over time instead of frequency so instead of a large number of carriers we only need a large number of time slots, which is already the case for capacity achieving codes. Second, by using multiple antennas we may be able to significantly reduce the required number of time/frequency slots per supersymbol, as shown in this paper for the case of 3 users. The third solution is to utilize interference alignment in other dimensions. A particularly promising approach in this direction is interference alignment over each symbol rather than over supersymbols. An interference alignment scheme is proposed in [16] that utilizes signal ”levels” to create the vector space within which interference may be aligned.

From a theoretical perspective, the present work offers significant insights into the capacity of wireless networks and the nature of coding schemes that are needed to achieve that capacity. The lack of understanding of interference networks prior to this work is evident from the large gap between the previously best known inner and outerbounds on the degrees of freedom. As we close this gap, we also reveal the fallacy of the cake cutting interpretation of medium access. Equally significant is the realization that the interference channel with more than 2 users presents many new challenges and opportunities (interference alignment) that are not as evident in the 2 user case. Thus, the 2 user interference channel model is signficantly lacking in its ability to represent and to provide insights into larger wireless interference networks.

The key insight of this paper is the role of interference alignment in a wireless network. From a capacity perspective the idea of interference alignment reaffirms the need for structured codes in wireless networks, also pointed out by [19]. For the single user point to point Gaussian channel it is well known that the capacity can be achieved through random (Gaussian) codebooks as well as through structured (lattice) codes. There is a growing realization that structured codes, optional for the single user case, may be necessary for approaching the capacity of networks. In an interference network when we design one user’s codebook we are also designing the interference/noise that will be seen by other users. Having structure in the interference may therefore be necessary. It is the structure imposed on the transmitted signals that facilitates interference alignment in this work. The intuition from this work is that since random codes will not automatically align themselves, structured codes will be necessary for wireless networks. Indeed interference alignment at the codeword level has been shown to be optimal in the capacity sense in [17] and in the degrees of freedom sense in [16] for some interesting cases. A combination of Han-Kobayashi [20] type achievable schemes and structured codes is one of the most promising avenues in the quest for the capacity of wireless networks.

ACKNOWLEDGMENT

We would like to acknowledge many helpful discussions with Shlomo Shamai that have guided this work.

APPENDIX I
CONVERSE FOR LEMMA 1

We present the proof for the case that all nodes are equipped with M antennas. In this case the statement of the lemma becomes:

$$\max_{\mathcal{D}} d_i + d_j \leq \limsup_{\rho\to\infty} \sup_{\mathbf{R}(\rho)\in\mathcal{C}(\rho)} \frac{R_i(\rho)+R_j(\rho)}{\log(\rho)} \leq M, \qquad \forall i,j \in \{1,2,\cdots,K\}, i \neq j. \tag{29}$$

We conside the case $i = 1, j = 2$ and eliminate messages $W_3, W_4, \cdots, W_K$, leaving us with a 2 user MIMO interference channel with frequency selective channel coefficients. Since the converse argument holds with or without channel variations, we do not distinguish between time and frequency dimensions. Instead, to simplify the notation for this proof, we map the time and frequency slot index (f,t) onto a channel use index $(f,t) \to n$. The channel input-output equations are written equivalently as:

$$\mathbf{Y}^{[1]}(n) = \mathbf{H}^{[11]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \mathbf{Z}^{[1]}(n), \tag{30}$$

$$\mathbf{Y}^{[2]}(n) = \mathbf{H}^{[21]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[22]}(n)\mathbf{X}^{[2]}(n) + \mathbf{Z}^{[2]}(n). \tag{31}$$

With probability one the channel matrices are invertible. So we can equivalently write

$$\mathbf{Y}^{[1]}(n) = \mathbf{H}^{[11]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \mathbf{Z}^{[1]}(n), \tag{32}$$

$$\mathbf{Y}^{[2]'}(n) = \mathbf{H}^{[12]}(n)\left(\mathbf{H}^{[22]}(n)\right)^{-1}\mathbf{H}^{[21]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \mathbf{Z}^{[2]'}(n). \tag{33}$$

where

$$\mathbf{Z}^{[1]}(n) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_M\right), \tag{34}$$

$$\mathbf{Z}^{[2]'}(n) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}^{[12]}(n)\left(\mathbf{H}^{[22]}(n)\right)^{-1}\left(\mathbf{H}^{[22]}(n)\right)^{-\dagger}\left(\mathbf{H}^{[12]}(n)\right)^{\dagger}\right). \tag{35}$$

Since the capacity of the interference channel depends only on the noise marginals, we assume without loss of generality that:

$$\mathbf{Z}^{[1]}(n) = \overline{\mathbf{Z}}(n) + \overline{\mathbf{Z}}^{[1]}(n), \tag{36}$$

$$\mathbf{Z}^{[2]'}(n) = \overline{\mathbf{Z}}(n) + \overline{\mathbf{Z}}^{[2]}(n), \tag{37}$$

where

$$\overline{\mathbf{Z}}(n) \sim \mathcal{CN}\left(\mathbf{0}, \alpha(n)\mathbf{I}_M\right), \tag{38}$$

$$\overline{\mathbf{Z}}^{[1]}(n) \sim \mathcal{CN}\left(\mathbf{0}, (1-\alpha(n))\mathbf{I}_M\right), \tag{39}$$

$$\overline{\mathbf{Z}}^{[2]}(n) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{H}^{[12]}(n)\left(\mathbf{H}^{[22]}(n)\right)^{-1}\left(\mathbf{H}^{[22]}(n)\right)^{-\dagger}\left(\mathbf{H}^{[12]}(n)\right)^{\dagger} - \alpha(n)\mathbf{I}_M\right), \tag{40}$$

and

$$\alpha(n) = \min\left(1, \lambda_{\min}\left(\mathbf{H}^{[12]}(n)\left(\mathbf{H}^{[22]}(n)\right)^{-1}\left(\mathbf{H}^{[22]}(n)\right)^{-\dagger}\left(\mathbf{H}^{[12]}(n)\right)^{\dagger}\right)\right)$$

is strictly positive with probability one. Here $\lambda_{\min}(A)$ refers to the smallest eigenvalue of matrix A. $\overline{\mathbf{Z}}(n)$, $\overline{\mathbf{Z}}^{[1]}(n)$, $\overline{\mathbf{Z}}^{[2]}(n)$ are mutually independent and jointly Gaussian.

Consider any reliable coding scheme for this interference channel, spanning N channel uses. We use the notation $[A(n)]_1^N$ to indicate the vector of values taken by variable $A(n)$ for $n = 1, 2, \cdots, N$. Starting from Fano's inequality, we have

$$\begin{aligned}
R_1(\rho) + R_2(\rho) &\leq \frac{1}{N} I(W_1; \left[\mathbf{Y}^{[1]}(n)\right]_1^N) + \frac{1}{N} I(W_2; W_1, \left[\mathbf{Y}^{[2]'}(n)\right]_1^N) + \epsilon, && (41)\\
&= \frac{1}{N} I(W_1; \left[\mathbf{Y}^{[1]}(n)\right]_1^N) + \frac{1}{N} I(W_2; \left[\mathbf{Y}^{[2]'}(n)\right]_1^N \Big| W_1, \left[\mathbf{X}^{[1]}(n)\right]_1^N) + \epsilon,\\
&\leq \frac{1}{N} I(W_1; \left[\mathbf{Y}^{[1]}(n)\right]_1^N) + \frac{1}{N} I(W_2; \left[\mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \overline{\mathbf{Z}}(n)\right]_1^N | W_1, \left[\mathbf{X}^{[1]}(n)\right]_1^N) + \epsilon && (42)\\
&\leq \frac{1}{N} I(W_1; \left[\mathbf{H}^{[11]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \overline{\mathbf{Z}}(n)\right]_1^N)\\
&\quad + \frac{1}{N} I(W_2; \left[\mathbf{H}^{[11]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \overline{\mathbf{Z}}(n)\right]_1^N \Big| W_1, \left[\mathbf{X}^{[1]}(n)\right]_1^N) + \epsilon, && (43)\\
&\leq \frac{1}{N} I(W_1, W_2; \left[\mathbf{H}^{[11]}(n)\mathbf{X}^{[1]}(n) + \mathbf{H}^{[12]}(n)\mathbf{X}^{[2]}(n) + \overline{\mathbf{Z}}(n)\right]_1^N) + \epsilon, && (44)\\
&\leq M\log(\rho) + o(\log(\rho)), && (45)
\end{aligned}$$

where the last step follows from the known result that the sum capacity of a multiple access channel with an M antenna receiver can only contribute at most M degree of freedom. Thus, we have $d_1 + d_2 \leq M$. Similarly, for any $i, j \in \{1, 2, \cdots, K\}, i \neq j$ we obtain $d_i + d_j \leq M$. Finally, adding up all the outerbounds in (3), we obtain the converse statement for the degrees of freedom of the K user interference channel with M antennas at each node and frequency selective channel coefficients.

$$\max_{\mathcal{D}} d_1 + d_2 + \cdots + d_K \quad \leq \quad MK/2. \tag{46}$$

APPENDIX II
ACHIEVABILITY FOR THEOREM 1 FOR ARBITRARY K

Let $N = (K-1)(K-2) - 1$. We show that $(d_1(n), d_2(n), \ldots d_K(n))$ lies in the degrees of freedom region of the K user interference channel for any $n \in \mathbb{N}$ where

$$\begin{aligned}
d_1(n) &= \frac{(n+1)^N}{(n+1)^N + n^N},\\
d_i(n) &= \frac{n^N}{(n+1)^N + n^N}, \qquad i = 2, 3 \ldots K.
\end{aligned}$$

This implies that

$$\max_{(d_1, d_2, \ldots d_K) \in \mathcal{D}} d_1 + d_2 + \cdots d_K \geq \sup_n \frac{(n+1)^N + (K-1)n^N}{(n+1)^N + n^N} = K/2.$$

We provide an achievable scheme to show that $((n+1)^N, n^N, n^N \ldots n^N)$ lies in the degrees of freedom region of an $M_n = (n+1)^N + n^N$ symbol extension of the original channel which automatically implies the desired result. In the extended channel, the signal vector at the k^{th} user's receiver can be expressed as

$$\bar{\mathbf{Y}}^{[k]}(t)=\sum_{j=1}^{K}\bar{\mathbf{H}}^{[kj]}\bar{\mathbf{X}}^{[j]}(t)+\bar{\mathbf{Z}}^{[k]}(t),$$

where $\bar{\mathbf{X}}^{[j]}$ is an $M_n \times 1$ column vector representing the M_n symbol extension of the transmitted symbol $X^{[k]}$, i.e.,

$$\bar{\mathbf{X}}^{[j]}(t) \triangleq \begin{bmatrix} X^{[j]}(1,t) \\ X^{[j]}(2,t) \\ \vdots \\ X^{[j]}(M_n,t)) \end{bmatrix}.$$

Similarly $\bar{\mathbf{Y}}^{[k]}$ and $\bar{\mathbf{Z}}^{[k]}$ represent M_n symbol extensions of the $Y^{[k]}$ and $Z^{[k]}$ respectively. $\bar{\mathbf{H}}^{[kj]}$ is a diagonal $M_n \times M_n$ matrix representing the M_n symbol extension of the channel i.e.,

$$\bar{\mathbf{H}}^{[kj]} \triangleq \begin{bmatrix} H^{[kj]}(1) & 0 & \dots & 0 \\ 0 & H^{[kj]}(2) & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \cdots & H^{[kj]}(M_n) \end{bmatrix}.$$

Recall that the diagonal elements of $\bar{\mathbf{H}}^{[kj]}$ are drawn independently from a continuous distribution and are therefore distinct with probability 1.

In a manner similar to the $K=3$ case, message W_1 is encoded at transmitter 1 into $(n+1)^N$ independent streams $x_m^{[1]}(t), m=1,2,\ldots,(n+1)^N$ along vectors $\mathbf{v}_m^{[1]}$ so that $\bar{\mathbf{X}}^{[1]}(t)$ is

$$\bar{\mathbf{X}}^{[1]}(t)=\sum_{m=1}^{(n+1)^N} x_m^{[1]}(t)\mathbf{v}_m^{[1]}=\bar{\mathbf{V}}^{[1]}\mathbf{X}^{[1]}(t),$$

where $\mathbf{X}^{[1]}(t)$ is a $(n+1)^N \times 1$ column vector and $\bar{\mathbf{V}}^{[1]}$ is a $M_n \times (n+1)^N$ dimensional matrix. Similarly $W_i, i \neq 1$ is encoded into n^K independent streams by transmitter i as

$$\bar{\mathbf{X}}^{[i]}(t)=\sum_{m=1}^{n^N} x_m^{[i]}(t)\mathbf{v}_m^{[i]}=\bar{\mathbf{V}}^{[i]}\mathbf{X}^{[i]}(t).$$

The received signal at the i^{th} receiver can then be written as

$$\bar{\mathbf{Y}}^{[i]}(t)=\sum_{j=1}^{K}\bar{\mathbf{H}}^{[ij]}\bar{\mathbf{V}}^{[j]}\mathbf{X}^{[j]}(t)+\bar{\mathbf{Z}}^{[i]}(t).$$

All receivers decode the desired signal by zero-forcing the interference vectors. At receiver 1, to obtain $(n+1)^N$ interference free dimensions corresponding to the desired signal from an $M_n=(n+1)^N+n^N$ dimensional received signal vector $\bar{\mathbf{Y}}^{[1]}$, the dimension of the interference should be not more than n^N. This can be ensured by perfectly aligning the interference from transmitters $2,3\ldots K$ as follows

$$\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}=\bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[3]}=\bar{\mathbf{H}}^{[14]}\bar{\mathbf{V}}^{[4]}=\ldots=\bar{\mathbf{H}}^{[1K]}\bar{\mathbf{V}}^{[K]}. \tag{47}$$

At the same time, receiver 2 zero-forces the interference from $\bar{\mathbf{X}}^{[i]}, i \neq 2$. To extract n^N interference-free dimensions from a $M_n=(n+1)^N+n^N$ dimensional vector, the dimension of the interference has to be not more than $(n+1)^N$.

This can be achieved by choosing $\bar{\mathbf{V}}^{[i]}, i \neq 2$ so that

$$\begin{aligned} \bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]} &\prec \bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}, \\ \bar{\mathbf{H}}^{[24]}\bar{\mathbf{V}}^{[4]} &\prec \bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}, \\ &\vdots \\ \bar{\mathbf{H}}^{[2K]}\bar{\mathbf{V}}^{[K]} &\prec \bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[1]}. \end{aligned} \tag{48}$$

Notice that the above relations align the interference from $K-2$ transmitters within the interference from transmitter 1 at receiver 2. Similarly, to decode W_i at receiver i when $i \neq 1$ we wish to choose $\bar{\mathbf{V}}^{[i]}$ so that the following $K-2$ relations are satisfied.

$$\bar{\mathbf{H}}^{[ij]}\bar{\mathbf{V}}^{[j]} \prec \bar{\mathbf{H}}^{[i1]}\bar{\mathbf{V}}^{[1]}, j \notin \{1, i\}. \tag{49}$$

We now wish to pick vectors $\bar{\mathbf{V}}^{[i]}, i = 1, 2 \ldots K$ so that equations (47), (48) and (49) are satisfied. Since channel matrices $\bar{\mathbf{H}}^{[ij]}$ have a full rank of M_n almost surely, equations (47), (48) and (49) can be equivalently expressed as

$$\bar{\mathbf{V}}^{[j]} = \mathbf{S}^{[j]}\mathbf{B} \qquad j = 2, 3, 4 \ldots K \qquad \text{At receiver 1} \tag{50}$$

$$\left.\begin{aligned} \mathbf{T}_3^{[2]}\mathbf{B} = \mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ \mathbf{T}_4^{[2]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ &\vdots \\ \mathbf{T}_K^{[2]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \end{aligned}\right\} \quad \text{At receiver 2} \tag{51}$$

$$\left.\begin{aligned} \mathbf{T}_2^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ \mathbf{T}_3^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ &\vdots \\ \mathbf{T}_{i-1}^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ \mathbf{T}_{i+1}^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \\ &\vdots \\ \mathbf{T}_K^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[1]} \end{aligned}\right\} \quad \text{At receiver i where } i = 3 \ldots K \tag{52}$$

where

$$\mathbf{B} = (\bar{\mathbf{H}}^{[21]})^{-1}\bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[3]}, \tag{53}$$

$$\mathbf{S}^{[j]} = (\bar{\mathbf{H}}^{[1j]})^{-1}\bar{\mathbf{H}}^{[13]}(\bar{\mathbf{H}}^{[23]})^{-1}\bar{\mathbf{H}}^{[21]}, \qquad j = 2, 3, \ldots K, \tag{54}$$

$$\mathbf{T}_j^{[i]} = (\bar{\mathbf{H}}^{[i1]})^{-1}\bar{\mathbf{H}}^{[ij]}\mathbf{S}^{[j]} \quad i, j = 2, 3 \ldots K, j \neq i. \tag{55}$$

Note that $\mathbf{T}_3^{[2]} = \mathbf{I}$, the $M_n \times M_n$ identity matrix. We now choose $\bar{\mathbf{V}}^{[1]}$ and $\mathbf{B}$ so that they satisfy the $(K-2)(K-1) = N+1$ relations in (51)-(52) and then use equations in (50) to determine $\bar{\mathbf{V}}^{[2]}, \bar{\mathbf{V}}^{[3]} \ldots \bar{\mathbf{V}}^{[K]}$. Thus, our goal is to find matrices $\bar{\mathbf{V}}^{[1]}$ and $\mathbf{B}$ so that

$$\mathbf{T}_j^{[i]}\mathbf{B} \prec \bar{\mathbf{V}}^{[1]}$$

for all $i, j = \{2, 3 \ldots K\}, i \neq j$.

Let $\mathbf{w}$ be the $M_n \times 1$ column vector

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

We need to choose $n^{(K-1)(K-2)-1} = n^N$ column vectors for $\mathbf{B}$. The sets of column vectors of $\mathbf{B}$ and $\bar{\mathbf{V}}^{[1]}$ are chosen to be equal to the sets B and $\bar{V}^{[1]}$ where,

$$B = \left\{ \left(\prod_{m,k\in\{2,3,\ldots K\}, m\neq k, (m,k)\neq(2,3)} \left(\mathbf{T}_k^{[m]}\right)^{\alpha_{mk}} \right) \mathbf{w} : \forall \alpha_{mk} \in \{0,1,2\ldots n-1\} \right\},$$

$$\bar{V}^{[1]} = \left\{ \left(\prod_{m,k\in\{2,3,\ldots K\}, m\neq k, (m,k)\neq(2,3)} \left(\mathbf{T}_k^{[m]}\right)^{\alpha_{mk}} \right) \mathbf{w} : \forall \alpha_{mk} \in \{0,1,2\ldots n\} \right\}.$$

For example, if $K = 3$ we get $N = 1$. $\mathbf{B}$ and $\bar{\mathbf{V}}^{[1]}$ are chosen as

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} \mathbf{w} & \mathbf{T}_2^{[3]}\mathbf{w} & \ldots & (\mathbf{T}_2^{[3]})^{n-1}\mathbf{w} \end{bmatrix} \\ \bar{\mathbf{V}}^{[1]} &= \begin{bmatrix} \mathbf{w} & \mathbf{T}_2^{[3]}\mathbf{w} & \ldots & (\mathbf{T}_2^{[3]})^{n}\mathbf{w} \end{bmatrix} \end{aligned}$$

To clarify the notation further, consider the case where $K = 4$. Assuming $n = 1$, B consists of exactly one element i.e $B = \{\mathbf{w}\}$. The set $\bar{V}^{[1]}$ consists of all $2^N = 2^5 = 32$ column vectors of the form $(\mathbf{T}_4^{[2]})^{\alpha_{24}}(\mathbf{T}_2^{[3]})^{\alpha_{32}}(\mathbf{T}_4^{[3]})^{\alpha_{24}}(\mathbf{T}_3^{[4]})^{\alpha_{43}}(\mathbf{T}_2^{[4]})^{\alpha_{42}}\mathbf{w}$
where all $\alpha_{24}, \alpha_{32}, \alpha_{34}, \alpha_{42}, \alpha_{43}$ take values $0, 1$. B and $\bar{V}^{[1]}$ can be verified to have n^N and $(n+1)^N$ elements respectively.

$\bar{\mathbf{V}}^{[i]}, i = 2, 3 \ldots K$ are chosen using equations (50). Clearly, for $(i,j) = (2,3)$,

$$\mathbf{T}_j^{[i]}\mathbf{B} = \mathbf{B} \prec \bar{\mathbf{V}}^{[1]}$$

Now, for $i \neq j, i, j = 2 \ldots K, (i,j) \neq (2,3)$,

$$\begin{aligned} \mathbf{T}_j^{[i]}B &= \Bigg\{ \left(\prod_{m,k\in\{2,3,\ldots N\}, m\neq k, (m,k)\neq(2,3)} \left(\mathbf{T}_k^{[m]}\right)^{\alpha_{mk}} \right) \mathbf{w} : \\ &\qquad \forall (m,k) \neq (i,j), \alpha_{mk} \in \{0,1,2\ldots n-1\}, \alpha_{ij} \in \{1,2,\ldots n\} \Bigg\} \\ \Rightarrow \mathbf{T}_j^{[i]}B &\in \bar{V}^{[1]} \\ \Rightarrow \mathbf{T}_j^{[i]}\mathbf{B} &\prec \bar{\mathbf{V}}^{[\mathbf{1}]}. \end{aligned}$$

Thus, the interference alignment equations (50)-(52) are satisfied.

Through interference alignment, we have now ensured that the dimension of the interference is small enough. We now need to verify that the components of the desired signal are linearly independent of the components of the interference so that the signal stream can be completely decoded by zero-forcing the interference. Consider the received signal vectors at Receiver 1. The desired signal arrives along the $(n+1)^N$ vectors $\bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]}$. As enforced by equations (50), the interference vectors from transmitters $3, 4 \ldots K$ are perfectly aligned with the interference from transmitter 2 and therefore, all interference arrives along the n^N vectors $\bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]}$. In order to prove that there are $(n+1)^N$ interference free dimensions it suffices to show that the columns of the square, $M_n \times M_n$ dimensional matrix

$$\begin{bmatrix} \bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]} & \bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]} \end{bmatrix} \tag{56}$$

are linearly independent almost surely. Multiplying the above $M_n \times M_n$ matrix with $(\bar{\mathbf{H}}^{[11]})^{-1}$ and substituting for $\bar{\mathbf{V}}^{[1]}$ and $\bar{\mathbf{V}}^{[2]}$, we get a matrix whose lth row has entries of the forms

$$\prod_{(m,k)\in\{2,3\ldots K\}, m\neq k, (m,k)\neq(2,3)} (\lambda_l^{mk})^{\alpha_{mk}}$$

and

$$d_l \prod_{(m,k)\in\{2,3\ldots K\}, m\neq k, (m,k)\neq(2,3)} (\lambda_l^{mk})^{\beta_{mk}},$$

where $\alpha_{mk} \in \{0, 1, \ldots n-1\}$ and $\beta_{mk} \in \{0, 1, \ldots n\}$ and λ_l^{mk}, d_l are drawn independently from a continuous distribution. The same iterative argument as in section III-B can be used. i.e. expanding the corresponding determinant along the first row, the linear independence condition boils down to one of the following occurring with non-zero probability

1) d_1 being equal to one of the roots of a linear equation
2) The coefficients of the above mentioned linear equation being equal to zero

Thus the iterative argument can be extended here, stripping the last row and last column at each iteration and the linear independence condition can be shown to be equivalent to the linear independence of a $n^N \times n^N$ matrix whose rows are of the form $\prod_{(m,k)\in\{2,3\ldots K\}, m\neq k} (\lambda_l^{mk})^{\alpha_{mk}}$ where $\alpha_{pq} \in \{0, 1, \ldots n-1\}$. Note that this matrix is a more general version of the Vandermonde matrix obtained in section III-B. So the argument for the $K = 3$ case does not extend here. However, the iterative procedure which eliminated the last row and the last column at each iteration, can be continued. For example, expanding the determinant along the first row, the singularity condition simplifies to one of

1) λ_l^{mk} being equal to one of the roots of a finite degree polynomial
2) The coefficients of the above mentioned polynomial being equal to zero

Since the probability of condition 1 occurring is 0, condition 2 must occur with non-zero probability. Condition 2 leads to a polynomial in another random variable λ_l^{pq} and thus the iterative procedure can be continued until the linear independence condition is shown to be equivalent almost surely to a 1×1 matrix being equal to 0. Assuming, without loss of generality, that we placed the $\mathbf{w}$ in the first row (this corresponds to the term $\alpha_{mk} = 0, \forall(m, k)$), the linear independence condition boils down to the condition that $1 = 0$ with non-zero probability - an obvious contradiction. Thus the matrix

$$\begin{bmatrix} \bar{\mathbf{H}}^{[11]}\bar{\mathbf{V}}^{[1]} & \bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[2]} \end{bmatrix}$$

can be shown to be non-singular with probability 1.

Similarly, the desired signal can be chosen to be linearly independent of the interference at all other receivers almost surely. Thus $(\frac{(n+1)^N}{(n+1)^N+n^N}, \frac{n^N}{(n+1)^N+n^N}, \cdots, \frac{n^N}{(n+1)^N+n^N})$ lies in the degrees of freedom region of the K user interference channel and therefore, the K user interference channel has $K/2$ degrees of freedom.

APPENDIX III
PROOF OF THEOREM 3 FOR M EVEN

Proof: To prove achievability we first consider the case when M is even. Through an achievable scheme, we show that there are $M/2$ non-interfering paths between transmitter i and receiver i for each $i = 1, 2, 3$ resulting in a total of $3M/2$ paths in the network.

Transmitter i transmits message W_i for receiver i using $M/2$ independently encoded streams over vectors $\mathbf{v}^{[i]}$ i.e

$$\mathbf{X}^{[i]}(t) = \sum_{m=1}^{M/2} x_m^{[i]}(t)\mathbf{v}_m^{[i]} = \mathbf{V}^{[i]}\mathbf{X}^i(t), i = 1, 2, 3.$$

The signal received at receiver i can be written as

$$\mathbf{Y}^{[i]}(t) = \mathbf{H}^{[i1]}\mathbf{V}^{[1]}\mathbf{X}^1(t) + \mathbf{H}^{[i2]}\mathbf{V}^{[2]}\mathbf{X}^2(t) + \mathbf{H}^{[i3]}\mathbf{V}^{[3]}\mathbf{X}^3(t) + \mathbf{Z}_i(t).$$

All receivers cancel the interference by zero-forcing and then decode the desired message. To decode the $M/2$ streams along the column vectors of $\mathbf{V}^{[i]}$ from the M components of the received vector, the dimension of the interference has to be less than or equal to $M/2$. The following three interference alignment equations ensure that the dimension of the interference is equal to $M/2$ at all the receivers.

$$\begin{aligned} \text{span}(\mathbf{H}^{[12]}\mathbf{V}^{[2]}) &= \text{span}(\mathbf{H}^{[13]}\mathbf{V}^{[3]}), & (57)\\ \mathbf{H}^{[21]}\mathbf{V}^{[1]} &= \mathbf{H}^{[23]}\mathbf{V}^{[3]}, & (58)\\ \mathbf{H}^{[31]}\mathbf{V}^{[1]} &= \mathbf{H}^{[32]}\mathbf{V}^{[2]}, & (59)\end{aligned}$$

where span$(\mathbf{A})$ represents the vector space spanned by the column vectors of matrix $\mathbf{A}$ We now wish to choose $\mathbf{V}^{[i]}, i = 1, 2, 3$ so that the above equations are satisfied. Since $\mathbf{H}^{[ij]}, i, j \in \{1, 2, 3\}$ have a full rank of M almost surely, the above equations can be equivalently represented as

$$\begin{aligned} \text{span}(\mathbf{V}^{[1]}) &= \text{span}(\mathbf{E}\mathbf{V}^{[1]}), & (60)\\ \mathbf{V}^{[2]} &= \mathbf{F}\mathbf{V}^{[1]}, & (61)\\ \mathbf{V}^{[3]} &= \mathbf{G}\mathbf{V}^{[1]}, & (62)\end{aligned}$$

where

$$\begin{aligned} \mathbf{E} &= (\mathbf{H}^{[31]})^{-1}\mathbf{H}^{[32]}(\mathbf{H}^{[12]})^{-1}\mathbf{H}^{[13]}(\mathbf{H}^{[23]})^{-1}\mathbf{H}^{[21]},\\ \mathbf{F} &= (\mathbf{H}^{[32]})^{-1}\mathbf{H}^{[31]},\\ \mathbf{G} &= (\mathbf{H}^{[23]})^{-1}\mathbf{H}^{[21]}.\end{aligned}$$

Let $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_M$ be the M eigenvectors of $\mathbf{E}$. Then we set $\mathbf{V}_1$ to be

$$\mathbf{V}^{[1]} = \begin{bmatrix} \mathbf{e}_1 & \ldots & \mathbf{e}_{(M/2)} \end{bmatrix}.$$

Then $\mathbf{V}^{[2]}$ and $\mathbf{V}^{[3]}$ are found using equations (60)-(62). Clearly, $\mathbf{V}^{[i]}, i = 1, 2, 3$ satisfy the desired interference alignment equations (57)-(59). Now, to decode the message using zero-forcing, we need the desired signal to be linearly independent of the interference at the receivers. For example, at receiver 1, we need the columns of $\mathbf{H}^{[11]}\mathbf{V}^{[1]}$ to be linearly independendent with the columns of $\mathbf{H}^{[21]}\mathbf{V}^{[2]}$ almost surely. i.e we need the matrix below to be of full rank almost surely

$$\begin{bmatrix} \mathbf{H}^{[11]}\mathbf{V}^{[1]} & \mathbf{H}^{[12]}\mathbf{V}^{[2]} \end{bmatrix}.$$

Substituting values for $\mathbf{V}^{[1]}$ and $\mathbf{V}^{[2]}$ in the above matrix, and multiplying by full rank matrix $(\mathbf{H}^{[11]})^{-1}$, the linear independence condition is equivalent to the condition that the column vectors of

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots \mathbf{e}_{(M/2)} & \mathbf{K}\mathbf{e}_1 & \ldots \mathbf{K}\mathbf{e}_{(M/2)} \end{bmatrix}$$

are linearly independent almost surely, where $\mathbf{K} = (\mathbf{H}^{[11]})^{-1}\mathbf{H}^{[12]}\mathbf{F}$.

This is easily seen to be true because $\mathbf{K}$ is a random (full rank) linear transformation. To get an intuitive understanding of the linear independence condition, consider the case of $M = 2$. Let $\mathcal{L}$ represent the line along which lies the first eigenvector of the random 2×2 matrix $\mathbf{E}$. The probability of a random rotation (and scaling) $\mathbf{K}$ of $\mathcal{L}$ being collinear with $\mathcal{L}$ is zero.

Using a similar argument, we can show that matrices

$$\left[\begin{array}{cc} \mathbf{H}^{[22]}\mathbf{V}^{[2]} & \mathbf{H}^{[21]}\mathbf{V}^{[1]} \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc} \mathbf{H}^{[33]}\mathbf{V}^{[3]} & \mathbf{H}^{[31]}\mathbf{V}^{[1]} \end{array}\right]$$

have a full rank of M almost surely and therefore receivers 2 and 3 can decode the $M/2$ streams of $\mathbf{V}^{[2]}$ and $\mathbf{V}^{[3]}$ using zero-forcing. Thus, a total $3M/2$ interference free transmissions per channel-use are achievable with probability 1 and the proof is complete. ∎

APPENDIX IV
PROOF OF THEOREM 3 FOR M ODD

Proof: Consider a two time-slot symbol extension of the channel, with the same chanel coefficients over the two symbols. It can be expressed as

$$\bar{\mathbf{Y}}^{[k]} = \bar{\mathbf{H}}^{[k1]}\bar{\mathbf{X}}^{[1]} + \bar{\mathbf{H}}^{[k2]}\bar{\mathbf{X}}^{[2]} + \bar{\mathbf{H}}^{[k3]}\mathbf{X}^{[3]} + \bar{\mathbf{Z}}^{[k]}$$

where $\bar{\mathbf{X}}^{[i]}$ is a $2M \times 1$ vector that represents the two symbol extension of the transmitted $M \times 1$ symbol symbol $\mathbf{X}^{[k]}$, i.e

$$\bar{\mathbf{X}}^{[k]}(t) \triangleq \left[\begin{array}{c} \mathbf{X}^{[k]}(1, 2t+1) \\ \mathbf{X}^{[k]}(1, 2t+2) \end{array}\right]$$

where $\mathbf{X}^{[k]}(t)$ is an $M \times 1$ vector representing the vector transmitted at time slot t by transmitter k. Similarly $\bar{\mathbf{Y}}^{[k]}$ and $\bar{\mathbf{Z}}^{[k]}$ represent the two symbol extensions of the the received symbol $\mathbf{Y}^{[k]}$ and the noise vector $\mathbf{Z}^{[k]}$ respectively at receiver i. $\bar{\mathbf{H}}^{[ij]}$ is a $2M \times 2M$ block diagonal matrix representing the extension of the channel.

$$\bar{\mathbf{H}}^{[ij]} \triangleq \left[\begin{array}{cc} \mathbf{H}^{[ij]}(1) & 0 \\ 0 & \mathbf{H}^{[ij]}(1) \end{array}\right].$$

We will now show (M, M, M) lies in the degrees of freedom region of this extended channel channel with an achievable scheme, implying that that a total of $3M/2$ degrees of freedom are achievable over the original channel. Transmitter k transmits message W_i for receiver i using M independently encoded streams over vectors $\mathbf{v}^{[k]}$ i.e

$$\bar{\mathbf{X}}^{[k]} = \sum_{m=1}^{M} x_m^{[k]}\mathbf{v}_m^{[k]} = \bar{\mathbf{V}}^{[k]}\mathbf{X}^{[k]},$$

where $\bar{\mathbf{V}}^{[k]}$ is a $2M \times M$ matrix and $\bar{\mathbf{X}}[k]$ is a $M \times 1$ vector representing M independent streams. The following three interference alignment equations ensure that the dimension of the interference is equal to M at receivers 1,2 and 3.

$$\begin{aligned} \text{rank}[\bar{\mathbf{H}}^{[21]}\bar{\mathbf{V}}^{[2]}] &= \text{rank}[\bar{\mathbf{H}}^{[31]}\bar{\mathbf{V}}^{[3]}], && (63) \\ \bar{\mathbf{H}}^{[12]}\bar{\mathbf{V}}^{[1]} &= \bar{\mathbf{H}}^{[32]}\bar{\mathbf{V}}^{[3]}, && (64) \\ \bar{\mathbf{H}}^{[13]}\bar{\mathbf{V}}^{[1]} &= \bar{\mathbf{H}}^{[23]}\bar{\mathbf{V}}^{[2]}. && (65) \end{aligned}$$

The above equations imply that

$$\begin{aligned}\text{span}(\bar{\mathbf{V}}^{[1]}) &= \text{span}(\bar{\mathbf{E}}\bar{\mathbf{V}}^{[1]}), \quad (66)\\ \bar{\mathbf{V}}^{[2]} &= \bar{\mathbf{F}}\bar{\mathbf{V}}^{[1]}, \quad (67)\\ \bar{\mathbf{V}}^{[3]} &= \bar{\mathbf{G}}\bar{\mathbf{V}}^{[1]}, \quad (68)\end{aligned}$$

where

$$\begin{aligned}\mathbf{E} &= (\mathbf{H}^{[13]})^{-1}\mathbf{H}^{[23]}(\mathbf{H}^{[21]})^{-1}\mathbf{H}^{[31]}(\mathbf{H}^{[32]})^{-1}\mathbf{H}^{[12]},\\ \mathbf{F} &= (\mathbf{H}^{[13]})^{-1}\mathbf{H}^{[23]},\\ \mathbf{G} &= (\mathbf{H}^{[12]})^{-1}\mathbf{H}^{[32]},\end{aligned}$$

and $\bar{\mathbf{E}}$, $\bar{\mathbf{F}}$ and $\bar{\mathbf{G}}$ are $2M \times 2M$ block-diagonal matrices representing the $2M$ symbol extension of $\mathbf{E}$, $\mathbf{F}$ and $\mathbf{G}$ respectively. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_M$, be the eigen vectors of $\mathbf{E}$. Then, we pick $\bar{\mathbf{V}}^{[1]}$ to be

$$\bar{\mathbf{V}}^{[1]} = \begin{bmatrix} \mathbf{e}_1 & 0 & \mathbf{e}_3 & \ldots & 0 & \mathbf{e}_M \\ 0 & \mathbf{e}_2 & 0 & \ldots & \mathbf{e}_{M-1} & \mathbf{e}_M \end{bmatrix}. \quad (69)$$

As in the even M case, $\bar{\mathbf{V}}^{[2]}$ and $\bar{\mathbf{V}}^{[3]}$ are then determined by using equations (66)-(68).

Now, we need the desired signal to be linearly independent of the interference at all the receivers. At receiver 1, the desired linear independence condition boils down to

$$\text{span}(\bar{\mathbf{V}}^{[1]}) \cap \text{span}(\bar{\mathbf{K}}\bar{\mathbf{V}}^{[1]}) = \{0\}$$

where $\mathbf{K} = (\mathbf{H}^{11})^{-1}\mathbf{H}^{[21]}(\mathbf{F})^{-1}$ and $\bar{\mathbf{K}}$ is the two-symbol diagonal extension of $\mathbf{K}$. Notice that $\mathbf{K}$ is an $M \times M$ matrix. The linear independence condition is equivalent to saying that all the columns of the following $2M \times 2M$ matrix are independent.

$$\begin{bmatrix} \mathbf{e}_1 & 0 & \mathbf{e}_3 & \ldots & 0 & \mathbf{e}_M & \mathbf{K}\mathbf{e}_1 & 0 & \mathbf{K}\mathbf{e}_3 & \ldots & 0 & \mathbf{K}\mathbf{e}_M \\ 0 & \mathbf{e}_2 & 0 & \ldots & \mathbf{e}_{M-1} & \mathbf{e}_M & 0 & \mathbf{K}\mathbf{e}_2 & 0 & \ldots & \mathbf{K}\mathbf{e}_{M-1} & \mathbf{K}\mathbf{e}_M. \end{bmatrix} \quad (70)$$

We now argue that the probability of the columns of the above matrix being linearly dependent is zero. Let $\mathbf{c}_i, i = 1, 2 \ldots 2M$ denote the columns of the above matrix. Suppose the columns $\mathbf{c}_i$ are linearly dependent, then

$$\exists \alpha_i \quad \text{s.t} \quad \sum_{i=1}^{2M} \alpha_i \mathbf{c}_i = 0$$

Let

$$\begin{aligned}\mathbf{P} &= \{\mathbf{e}_1, \mathbf{e}_3 \ldots \mathbf{e}_{M-2}, \mathbf{K}\mathbf{e}_1, \ldots \mathbf{K}\mathbf{e}_{M-2}\},\\ \mathbf{Q} &= \{\mathbf{e}_2, \mathbf{e}_4 \ldots \mathbf{e}_{M-1}, \mathbf{K}\mathbf{e}_2, \ldots \mathbf{K}\mathbf{e}_{M-1}\}.\end{aligned}$$

Now, there are two possibilities

1) $\alpha_M = \alpha_{2M} = 0$. This implies that either one of the following sets of vectors is linearly dependent. Note that both sets are can be expressed as the union of
   a) A set of $\lfloor (M/2) \rfloor$ eigen vectors of $\mathbf{E}$
   b) A random transformation $\mathbf{K}$ of this set.

   An argument along the same lines as the even M case leads to the conclusion that the probability of the union of the two sets listed above being linearly dependent in a M dimensional space is zero.

2) $\alpha_{2M} \neq 0$ or $\alpha_M \neq 0$. This implies that

$$\alpha_M \mathbf{e}_M + \alpha_{2M}\mathbf{K}\mathbf{e}_M \in \text{span}(\mathbf{P}) \cap \text{span}(\mathbf{Q})$$

$$\Rightarrow \text{span}(\{\mathbf{K}\mathbf{e}_M, \mathbf{e}_M\}) \cap \text{span}(\mathbf{P}) \cap \text{span}(\mathbf{Q}) \neq \{0\}.$$

Also,

$$\text{rank}(\text{span}(\mathbf{P}) \cup \text{span}(\mathbf{Q})) = \text{rank}(\mathbf{P}) + \text{rank}(\mathbf{Q}) - \text{rank}(\mathbf{P} \cap \mathbf{Q}),$$

$$\Rightarrow \text{rank}(\mathbf{P} \cap \mathbf{Q}) = 2M - 2 - \text{rank}(\text{span}(\mathbf{P}) \cup \text{span}(\mathbf{Q})).$$

Note that $\mathbf{P}$ and $\mathbf{Q}$ are $M-1$ dimensional spaces. (The case where their dimensions are less than $M-1$ is handled in the first part). Also, $\mathbf{P}$ and $\mathbf{Q}$ are drawn from completely different set of vectors. Therefore, the union of $\mathbf{P}, \mathbf{Q}$ has a rank of M almost surely. Equivalently span$(\mathbf{P}) \cap$ span$(\mathbf{Q})$ has a dimension of $M-2$ almost surely. Since the set $\{\mathbf{e}_M, \mathbf{K}\mathbf{e}_M\}$ is drawn from an eigen vector $\mathbf{e}_M$ that does not exist in either $\mathbf{P}$ or $\mathbf{Q}$, the probability of the 2 dimensional space span$(\{\mathbf{e}, \mathbf{K}\mathbf{e}_M\})$ intersecting with the $M-2$ dimensional space $\mathbf{P} \cap \mathbf{Q}$ is zero. For example, if $M = 3$, let L indicate the line formed by the intersection of the the two planes span$(\{\mathbf{e}_1, \mathbf{K}\mathbf{e}_1\})$ and span$(\{\mathbf{e}_2, \mathbf{K}\mathbf{e}_2\})$. The probability that line L lies in the plane formed by span$(\{\mathbf{e}_3, \mathbf{K}\mathbf{e}_3\})$. Thus, the probability that the desired signal lies in the span of the interference is zero at receiver 1. Similarly, it can be argued that the desired signal is independent of the interference at receivers 2 and 3 almost surely. Therefore (M, M, M) is achievable over the two-symbol extended channel. Thus $3M/2$ degrees of freedom are achievable over the 3 user interference channel with M antenna at each transmitting and receiving node.

■

REFERENCES

[1] S. Toumpis, "Wireless ad hoc networks," in *IEEE Sarnoff Symposium Princeton NJ*, April 2004.

[2] L. Zheng and D. N. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the non-coherent multi-antenna channel,," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb 2002.

[3] A. S. Avestimehr, S. Diggavi, and D. Tse, "A deterministic approach to wireless relay networks," Oct 2007, arXiv:cs.IT/0710.3777.

[4] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1859–1874, Sept. 2004.

[5] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. on Information Theory*, vol. 50, pp. 1875–1892, Sept. 2004.

[6] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, pp. 1912–1921, Aug 2003.

[7] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inform. Theory*, pp. 2895–2909, Oct. 2003.

[8] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. of ISIT*, 2005.

[9] S. Jafar and M. Fakhereddin, "Degrees of freedom for the mimo interference channel," *IEEE Transactions on Information Theory*, vol. 53, pp. 2637–2642, July 2007.

[10] M. Maddah-Ali, A. Motahari, and A. Khandani, "Signaling over MIMO multi-base systems - combination of multi-access and broadcast schemes," in *Proc. of ISIT*, pp. 2104–2108, 2006.

[11] S. Jafar, "Degrees of freedom on the MIMO X channel- optimality of the MMK scheme," Tech. Report, Sep. 2006, arXiv:cs.IT/0607099v2.

[12] S. Jafar and S. Shamai, "Degrees of freedom region for the MIMO X channel," in *arXiv:cs.IT/0607099v3*, May 2007.

[13] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel," in *Proceedings of Annual Information Theory and Applications Workshop UCSD*, Jan 2007.

[14] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over X channel: Signaling and performance analysis," in *Tech. Report. UW-ECE-2006-27, University of Waterloo*, December 2006.

[15] A. B. Carleial, “Interference channels,” *IEEE Trans. Inform. Theory*, vol. 24, no. 1, pp. 60–70, 1978.

[16] S. Jafar, S. Shamai, and V. Cadambe, “Interference alignment on the deterministic channel and application to gaussian networks,” Submitted to 2008 ISIT.

[17] G. Bresler, A. Parekh, and D. Tse, “Approximate capacity of the many-to-one interference channel,” Sep. 2007. Allerton Conference.

[18] P. Viswanath, D. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1277–1294, June 2002.

[19] B. Nazer and M. Gastpar, “The case for structured random codes in network communication theorems,” in *IEEE Information Theory Workshop*, Sep 2007.

[20] T. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Inform. Theory*, vol. 27, pp. 49–60, Jan 1981.

Syed Ali Jafar (S’ 99- M’04) received the B. Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT), Delhi, India in 1997, the M.S. degree in Electrical Engineering from California Institute of Technology (Caltech), Pasadena USA in 1999, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA USA in 2003. He was a summer intern in the Wireless Communications Group of Lucent Bell Laboratories, Holmdel, NJ, in 2001. He was an engineer in the Satellite Networks Division of Hughes Software Systems from 1997-1998 and a senior engineer at Qualcomm Inc., San Diego, CA in 2003. He is currently an Assistant Professor in the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory and wireless communications.

Dr. Jafar received the NSF CAREER award in 2006. He is the recipient of the 2006 UC Irvine Engineering Faculty of the Year award for excellence in teaching. Dr. Jafar serves as the Editor for Wireless Communication Theory and CDMA for the IEEE Transactions on Communications.

Viveck R. Cadambe (S’ 05) received the B. Tech. and M.S. degrees in Electrical Engineering from Indian Institute of Technology, Chennai, in 2006. He is currently working toward the Ph.D. degree at the University of California Irvine. Mr. Cadambe is a recipient of the UC Irvine CPCC graduate fellowship for the year 2007-08. His research interests include multiuser information theory and wireless networks.