Interference Alignment Under Limited Feedback for MIMO Interference Channels

Rajesh T. Krishnamachari and Mahesh K. Varanasi, *Fellow, IEEE*

***Abstract*—This paper analyzes multiple-input, multiple-output interference channels where each receiver knows its channels from all the transmitters and feeds back this information using a limited number of bits to all the other terminals. It is shown that as long as the feedback bit rate scales sufficiently fast with the signal-to-noise ratio, the transmitters can use an interference alignment strategy by treating the quantized channel estimates as being perfect to achieve the sum degrees of freedom of the interference channel attainable with perfect and global channel state information. A tradeoff between the feedback rate and the achievable degrees of freedom is established by showing that a slower scaling of feedback rate for any one user leads to commensurately fewer degrees of freedom for that user alone. It is then shown that under the same fixed transmission strategy but with random quantization, the above mentioned sufficient condition on the feedback scaling rate to attain a given sum degrees of freedom (up to the maximum attainable) is also necessary in this setting.**

***Index Terms*—Composite grassmann manifold, finite-rate feedback, interference alignment, interference channel, MIMO, quantization.**

I. Introduction

The importance of the role played by interference management in wireless networks has prompted many researchers to analyze the interference channel from an information-theoretic perspective. While the capacity region remains unknown, many insightful characterizations have been developed for the two-user interference channel through inner/outer bounds and approximations (cf. [1]–[6] and references therein). For K-user interference channels with $K > 2$, the conventional wisdom was that if the strength of the interference is comparable to the actual signal, orthogonalizing the signaling dimensions between users was best providing $\frac{1}{K}$ degrees of freedom per user in a K-user interference channel. Abandoning this 'cake-cutting' approach, [7] demonstrated the achievability of the sum degrees of freedom for this channel of $\frac{K}{2}$ through what has come to be known as interference alignment (IA). This surprising result has spurred further research in this area (cf. [8]), including the analysis of the variation of spectral efficiency with the number of users under limited multipath in [9], thereby demonstrating the necessity of scaling bandwidth sufficiently fast with K for the IA scheme of [7].

The requirement of perfect and global channel state information at the transmitters and receivers (CSITR) by the IA scheme in [7] is, of course, practically unrealizable for a time-variant or frequency-selective system and this issue has recently begun to receive considerable attention ([10]–[12]). For a frequency-selective single-input single-output (SISO) channel with L taps in the channel between any pair of nodes and P as the total power available with the transmitting sources, the limited feedback scenario has been analyzed in [12] where it was shown that for sufficiently large L, the full spatial multiplexing gain of $\frac{K}{2}$ can still be obtained as long as the feedback rate is scaled as $K(L-1)\log P$ bits per receiver. This scheme employed a Grassmannian line quantization—as in [13]—of the channel vectors by the channel-aware receivers and their feedback using a limited number of bits broadcast to all the nodes in the network.

For the K-user MIMO channel with M_t antennas at each transmitter and M_r antennas at each receiver, the sum degrees of freedom under perfect (and global) CSITR was bounded in [14] within close lower and upper bounds which coincided whenever the ratio $\frac{\max\{M_t,M_r\}}{\min\{M_t,M_r\}}$ was an integer. In particular, this means that the spatial multiplexing gain is precisely known for both the single-input multiple-output (SIMO) and the multiple-input single-output (MISO) cases. In this paper, we analyze the frequency-selective MIMO interference channel under the regime of limited feedback and demonstrate the achievability of the same degrees of freedom as the original perfect-CSITR IA scheme as long as each receiver uses no less than $N_f = \min\{M_t, M_r\}(K\min\{M_t, M_r\}-1)(RL-1)\log P$ bits, where $R = \lfloor \frac{\max\{M_t,M_r\}}{\min\{M_t,M_r\}} \rfloor$. In each case, a codebook over the composite Grassmann manifold is employed to jointly quantize the normalized channel directions. Our result extends the previous work of [12] to multi-antenna systems and also offers an improvement in the feedback scaling rate required when specialized to their single-antenna case. With the previous pre-print version of our results in [15] as a benchmark, the authors in [16] addressed the question of improving the achievable rate at finite P in a more complex setting which requires the additional iterative computation of pre-quantization filters at the receivers. The alternative of analog channel state feedback was also recently studied in [17].

We extend our analysis by proving that if any user i fed back $\alpha_i \cdot N_f (0 < \alpha_i \leq 1)$ bits, then user i achieves degrees of

Manuscript received June 20, 2012; revised October 29, 2012 and March 17, 2013; accepted April 04, 2013. Date of publication May 20, 2013; date of current version July 10, 2013 . The associate editor coordinating the review of this manuscript and approving it for publication was Prof. David Love. This work was supported in part by NSF under Grant CCF-0728955 and by KAU Grant 4-15 1432/HICI.

R. T. Krishnamachari was with the Department of Electrical, Computer and Energy Engineering, University of Colorado, Boulder, CO 80309-0425 USA. He is now employed in the financial services industry in New York, NY, USA (e-mail address: Rajesh.Krishnamachari@Ccolorado.edu).

M. K. Varanasi is with the Department of Electrical, Computer and Energy Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: varanasi@colorado.edu).

Digital Object Identifier 10.1109/TSP.2013.2264055

freedom equal to α_i times what that user would achieve with global and perfect CSITR. Interestingly therefore, a slower scaling of feedback rate for any one user leads to commensurately fewer degrees of freedom for that user alone. We also provide a partial converse to these achievability results by computing an upper bound on the achievable sum degrees of freedom at a given scaling of feedback rate within the ambit of the interference alignment strategy used in the achievability proof, when the channel vectors are individually quantized using a random code and the degrees of freedom of the channel are interpreted as averages over the ensemble of all random codes.

Notation: $\mathbb{R}$ and $\mathbb{C}$ represent the real and complex fields, respectively. If $z \in \mathbb{C}$, then z^c represents its complex conjugate. The superscripts ^{H} and ^{t} represent the hermitian conjugate and the transpose of a matrix, respectively. Extending the conjugate notation to vectors, we denote as $v^c = [v_1^c, v_2^c, \ldots, v_n^c]^t$ when $v = [v_1, v_2, \ldots, v_n]^t$. We use square brackets [.] for time indices and circular brackets (.) for frequency indices. The symbols $\circ$ and $\otimes$ represent the Hadamard and Kronecker products of matrices, respectively. We use the term 'const' to denote any constant independent of the power P, whose value might change from equation to equation. $\mathrm{CN}(0, \sigma^2)$ represents the circularly symmetric complex normal distribution with the mean being equal to zero and variance equal to σ^2. All logarithms in this paper are taken with respect to base two. $A =_d B$ denotes that the random variables A and B have the same distribution. The uniform distribution over a manifold S is denoted by $\mathrm{Unif}(S)$. $1_{[a,b]}(x)$ is the indicator function that is one over the interval $[a, b]$ and zero elsewhere. The gamma and beta functions are given by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ and $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, respectively. The beta (α, β) distribution refers to the probability density function $f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} 1_{[0,1]}(x)$.

II. IA With Limited Feedback for SIMO/MISO Systems

The K-user $1 \times R$ SIMO interference channel with one antenna at each transmitter and $M_r = R$ antennas at each receiver as well as the corresponding $R \times 1$ MISO channel (here $M_t = R$) was studied under the perfect CSITR assumption by [14] who used IA to achieve the sum degrees of freedom of this channel at $d_{sum} = \frac{R}{R+1}.K$, when $K > R$. While the scheme in [14] is agnostic to whether coding is performed across time instances or frequencies, we shall restrict ourselves in this paper to coding over different frequency tones obtained through a DFT (Discrete Fourier Transform)-based transformation of an underlying frequency selective (or multi-tap) channel, since this facilitates the construction of a *causal* feedback procedure.

A. System Model and Theorem Statement

Throughout this section, we shall deal with the $1 \times R$ SIMO channel and transfer the results at the end to the $R \times 1$ MISO case using the idea of reciprocity of alignment [18]. There are K single-antenna sources $S_1, S_2, \ldots, S_K$ with a message each for the respective multi-antenna destinations $D_1, D_2, \ldots, D_K$. The L-tap channel between the source S_k and the destination D_i is given by $\mathbf{h}_{i,k}[l] \in \mathbb{C}^{R\times 1}, l \in \{0, 1, \ldots, L-1\}$. For our analysis, we will assume an idealistic scenario where these coefficients are drawn independently and with identical distribution (i.i.d.) from a continuous distribution and are bounded with probability one. We would also assume that the channel values $\mathbf{h}_{i,k}[l]$ do not change during the transmission of the signal. The additive noise is assumed to be varying as i.i.d. random variables across both time and space with the distribution $\mathrm{CN}(0, N_o)$.

The receivers are assumed to have perfect knowledge of their respective channels, i.e. each destination D_i knows $\mathbf{h}_{i,k}^t[0], \mathbf{h}_{i,k}^t[1], \ldots, \mathbf{h}_{i,k}^t[L-1]\ \forall k \in \{1, 2, \ldots K\}$. As in [12], we assume there are dedicated error-free broadcast links from the destinations to every other node in the network. During an initial channel feedback phase, each receiver broadcasts its CSI using N_f bits of feedback. This is followed by the data transmission phase. The maximal rate of communication between S_i and D_i such that the probability of error is driven to zero as the block-length goes to infinity is denoted as R_i with $R_{sum} \triangleq \sum_{i=1}^K R_i$. The sum degrees of freedom is defined as $d_{sum} \triangleq \lim_{P\to\infty} \frac{R_{sum}}{\log P}$, where P is the total power constraint on the transmitters.

Theorem 1: For a K-user $1 \times R$ SIMO or $R \times 1$ MISO channel with $K > R$, where each pair of nodes in the network has an L-tap frequency selective channel between them, it is possible to achieve the full spatial multiplexing gain of $\frac{KR}{R+1}$ as long as each destination scales its feedback rate faster than $(K-1)(RL-1)\log P$ bits of feedback, where P represents the total power available with the transmitting nodes of the network.

The proof is split into the succeeding three subsections.

B. Quantization on Composite Grassmann

We shall find in the succeeding subsection that our feedback naturally lies on a surface called the composite Grassmann manifold. In this subsection, we bound the maximum distortion under quantization by a sphere-packing code over the composite Grassmann manifold. Recall that the Grassmann manifold $G_{n,k}$ is the collection of all the k-dimensional subspaces within the n-dimensional Euclidean space $\mathbb{C}^n$.

It shall be seen that to realize the complete spatial multiplexing gain of the system, it is important to quantize and feedback $K-1$ unit norm vectors of RL length vectors (and not just their span) for the construction of the input beamforming vectors via the scheme of [14] employed here. One viable choice would be quantize each of the $K-1$ vectors individually over the $G_{RL,1}$ manifold and we describe this further in Section II-E. Another choice would be to quantize them jointly using a source code over the $(K-1)^{\text{st}}$ direct product of the $G_{RL,1}$ space.

In [19] and [20], the composite Grassmann manifold $G_{n,k}^m$ is formed by taking the direct sum of m copies of the Grassmann manifold $G_{n,k}$, i.e.

$$G_{n,k}^m = \bigoplus_{m \text{ copies}} G_{n,k}.$$

For $G_{n,1}$, the chordal distance between two points is defined as $d_c^2(\mathbf{x}, \mathbf{y}) \triangleq 1 - |\mathbf{x}^H \mathbf{y}|^2$. One can extend this distance to $G_{n,1}^m$ as follows: If $\mathbf{X}, \mathbf{Y} \in G_{n,1}^m$, then $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots, \mathbf{x}_m]$, and

$\mathbf{Y} \triangleq [\mathbf{y}_1, \ldots, \mathbf{y}_m]$, where $\mathbf{x}_i, \mathbf{y}_i \in G_{n,1}\ \forall i \in \{1, 2, \ldots, m\}$. The squared distance between $\mathbf{X}$ and $\mathbf{Y}$ is then given by

$$d^2(\mathbf{X}, \mathbf{Y}) \triangleq \sum_{i=1}^{m} d_c^2(\mathbf{x}_i, \mathbf{y}_i).$$

In our analysis, we would need only $G_{RL,1}^{K-1}$; so we limit the succeeding discussion in this subsection to it.

Define a ball of radius δ around a point $\mathbf{X} \in G_{RL,1}^{K-1}$ as

$$B_{\mathbf{X}}(\delta) = \left\{ \mathbf{Y} \in G_{RL,1}^{K-1} | d(\mathbf{X}, \mathbf{Y}) \leq \delta \right\}.$$

Since $G_{RL,1}^{K-1}$ is a homogenous space, the volume of a ball is independent of the choice of $\mathbf{X}$. Hence, we can denote a ball of radius δ in the composite Grassmann manifold as $B(\delta)$ without any reference to its central point. Based on our distance metric, consider a maximal packing of spheres on the composite Grassmann manifold $G_{RL,1}^{K-1}$ such that the minimum distance between the centers of any two spheres is more than δ. This is the precise analogue of Grassmannian sphere-packing in [21]. We denote this maximal packing code as our quantization codebook $\mathcal{C}_{sph}$.

A realization $\mathbf{x} \in G_{RL,1}^{K-1}$ shall be encoded using N_f bits corresponding to the index of the codeword in $\mathcal{C}_{sph}$ closest to it, i.e. the quantized version of $\mathbf{x}$ shall be

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{c}_i \in \mathcal{C}_{sph}} \left[d(\mathbf{x}, \mathbf{c}_i) \right].$$

The maximum distortion is then defined as

$$\triangle_{\max} \triangleq \max_{\mathbf{x} \in G_{RL,1}^{K-1}} \left[d(\mathbf{x}, \hat{\mathbf{x}}) \right].$$

Lemma 2: For the code $\mathcal{C}_{sph}$, the maximum distortion experienced by a realization is bounded by the minimum distance of the code as follows:

$$\triangle_{\max} \leq \delta \leq \frac{\text{const}}{2^{\frac{N_f}{2(K-1)(RL-1)}}}. \tag{1}$$

Proof: The proof readily follows by combining a volume result of [20] with the quantization analysis of [12]. The number of codewords, denoted by $|\mathcal{C}_{sph}|$, is constrained by the well-known Hamming bound as follows:

$$|\mathcal{C}_{sph}| \leq \frac{1}{\mu\left(B(\delta)\right)},$$

where $\mu(B(\delta))$ represents the normalized volume of a ball in the composite Grassmann manifold, defined as $\mu(B(\delta)) = \frac{\mathrm{Vol}(B(\delta))}{\mathrm{Vol}(G_{RL,1}^{K-1})}$. For $G_{RL,1}^{K-1}$, Theorem 7 in [20] provides the normalized ball volume, up to a multiplicative factor of $(1 + O(\delta^2))$, as

$$\mu\left(B(\delta)\right) = \frac{\Gamma^{K-1}(RL)}{\Gamma\left((K-1)(RL-1)+1\right)} \delta^{2(K-1)(RL-1)}.$$

The exponent of the δ term, i.e., $2(K-1)(RL-1)$ is the real dimension of the manifold. Substituting $2^{N_f} \triangleq |\mathcal{C}_{sph}|$ and ignoring the multiplicative $(1 + O(\delta^2))$ factor under the regime of high feedback rates, we write

$$2^{N_f} \leq \frac{\text{const}}{\delta^{2(K-1)(RL-1)}} \Rightarrow \delta \leq \frac{\text{const}}{2^{\frac{N_f}{2(K-1)(RL-1)}}}.$$

■

We shall observe later that setting $\frac{1}{2^{\frac{N_f}{2(K-1)(RL-1)}}}$ to be equal to $\frac{1}{\sqrt{P}}$, which results in the feedback scaling rate being given by $N_f = (K-1)(RL-1)\log P$ is sufficient to achieve the maximum sum degrees of freedom of the SIMO channel.

C. Feedback and Reconstruction Procedure

The use of an OFDM-type cyclic signal model transforms our L-tap frequency selective channel into N parallel frequency flat channels, which can be succinctly described in matrix format as

$$\overline{\mathbf{y}}_i = \mathbf{H}_{i,i}\overline{\mathbf{x}}_i + \sum_{k \neq i} \mathbf{H}_{i,k}\overline{\mathbf{x}}_k + \overline{\mathbf{z}}_i. \tag{2}$$

Here, $\overline{\mathbf{y}}_i \in \mathbb{C}^{NR \times 1}$ can be decomposed as $[\mathbf{y}_i(0)^H, \mathbf{y}_i(1)^H, \ldots, \mathbf{y}_i(N-1)^H]^H$, where $\mathbf{y}_i(r) \in \mathbb{C}^{R \times 1}$ is the channel output on the r-th tone at the i-th receiver. The channel input $\overline{\mathbf{x}}_i \in \mathbb{C}^{N \times 1}$ and the i.i.d. noise $\overline{\mathbf{z}}_i \in \mathbb{C}^{NR \times 1}$ can be decomposed analogously. The impulse response matrix is given by $\mathbf{H}_{i,k} \triangleq \mathrm{diag}\{\mathbf{h}_{i,k}^c(0), \cdots, \mathbf{h}_{i,k}^c(N-1)\}$, where $\mathbf{h}_{i,k}(r) \in \mathbb{C}^{R \times 1}$ is the channel vector from S_k to D_i on the r-th tone.

The destination node D_i constructs an RL-length vector as

$$\mathbf{t}_{i,k} \triangleq \frac{\mathrm{vec}(\mathbf{T}_{i,k})}{||\mathrm{vec}(\mathbf{T}_{i,k})||},$$

where

$$\mathbf{T}_{i,k} \triangleq \begin{pmatrix} \mathbf{h}_{i,k}^t[0] \\ \mathbf{h}_{i,k}^t[1] \\ \vdots \\ \mathbf{h}_{i,k}^t[L-1] \end{pmatrix} \triangleq \left(\mathbf{c}_{i,k}[1], \mathbf{c}_{i,k}[2], \ldots, \mathbf{c}_{i,k}[R]\right).$$

The i-th receiver quantizes $\mathbf{Q}_i \triangleq [\mathbf{t}_{i,1}, \mathbf{t}_{i,2}, \ldots, \mathbf{t}_{i,i-1}, \mathbf{t}_{i,i+1}, \ldots \mathbf{t}_{i,K}]$ over the 2^{N_f}-level codebook $\mathcal{C}_{sph}$ on the composite Grassmann $G_{RL,1}^{K-1}$ to obtain $\hat{\mathbf{Q}}_i = [\hat{\mathbf{t}}_{i,1}, \hat{\mathbf{t}}_{i,2}, \ldots, \hat{\mathbf{t}}_{i,i-1}, \hat{\mathbf{t}}_{i,i+1}, \ldots, \hat{\mathbf{t}}_{i,K}]$, which is fed back to all the other terminals.

We can now form an $L \times R$ matrix $\hat{\mathbf{T}}_{i,k}$ as $\mathrm{vec}^{-1}(\hat{\mathbf{t}}_{i,k})$ to get

$$\hat{\mathbf{T}}_{i,k} \triangleq \begin{pmatrix} \hat{\mathbf{h}}_{i,k}^t[0] \\ \hat{\mathbf{h}}_{i,k}^t[1] \\ \vdots \\ \hat{\mathbf{h}}_{i,k}^t[L-1] \end{pmatrix} \triangleq \left(\hat{\mathbf{c}}_{i,k}[1], \hat{\mathbf{c}}_{i,k}[2], \ldots, \hat{\mathbf{c}}_{i,k}[R]\right).$$

Let $\hat{\mathbf{c}}_{i,k}(m)$ represent the N-point DFT of $\hat{\mathbf{c}}_{i,k}[m], m \in \{1, 2, \ldots, R\}$. These are arranged as shown to form an $N \times R$ matrix,

$$\hat{\mathbf{F}}_{i,k} \triangleq \begin{pmatrix} \hat{\mathbf{h}}^t_{i,k}(0) \\ \hat{\mathbf{h}}^t_{i,k}(1) \\ \vdots \\ \hat{\mathbf{h}}^t_{i,k}(N-1) \end{pmatrix} \triangleq (\hat{\mathbf{c}}_{i,k}(1), \hat{\mathbf{c}}_{i,k}(2), \ldots, \hat{\mathbf{c}}_{i,k}(R)).$$

The equivalent matrix in the unquantized domain is $\mathbf{F}_{i,k}$ with rows $\mathbf{c}_{i,k}(.)$ being the N-point DFT of the corresponding row $\mathbf{c}_{i,k}[.]$ in $\mathbf{T}_{i,k}$.

$$\mathbf{F}_{i,k} \triangleq \begin{pmatrix} \mathbf{h}^t_{i,k}(0) \\ \mathbf{h}^t_{i,k}(1) \\ \vdots \\ \mathbf{h}^t_{i,k}(N-1) \end{pmatrix} \triangleq (\mathbf{c}_{i,k}(1), \mathbf{c}_{i,k}(2), \ldots, \mathbf{c}_{i,k}(R)).$$

A naive extension of the quantization strategy in the SISO analysis in [12] would be to quantize the R vectors individually using $G_{L,1}$. Such an idea, if applicable, would lead to a smaller feedback scaling rate than suggested by Theorem 1. However, this idea fails since such a Grassmannian quantization disregards the magnitudes of the different vectors and amounts to post-multiplying the $\mathbf{T}_{i,k}$ matrix by $\operatorname{diag}\{\frac{1}{\|\mathbf{c}_{i,k}[1]\|}, \ldots, \frac{1}{\|\mathbf{c}_{i,k}[R]\|}\}$ prior to its representation by a Grassmannian codeword. The lack of knowledge of this diagonal matrix however precludes the design of input beamforming vectors of the IA scheme of [14]. On the other hand, quantizing $\mathbf{t}_{i,k}$ on $G_{RL,1}$ causes merely the loss of a single scalar, namely $\|\mathrm{vec}(\mathbf{T}_{i,k})\|$, the unavailability of which has no impact on computing the IA beamforming vectors.

Following [14], let $n \in \mathbb{N}$ and

$$\tau \triangleq KR(K - R - 1). \tag{3}$$

The number of tones over which we shall code, or the length of each super-symbol in the terminology of [7], is given by

$$N \triangleq (R+1)(n+1)^\tau, \tag{4}$$

coding over which the user i shall be shown to achieve the d_i degrees of freedom where

$$d_i \triangleq \begin{cases} R(n+1)^\tau & i \in \{1, 2, \ldots, R+1\}; \\ R(n)^\tau & i \in \{R+2, \ldots, K\} \end{cases}. \tag{5}$$

The choice of N is explained in [14], and is hence not repeated here.

The source S_k formulates d_k independent symbols—$x^1_k, \ldots, x^{d_k}_k \in \mathbb{C}$, which are sent along the directions $\mathbf{v}^1_k, \ldots, \mathbf{v}^{d_k}_k \in \mathbb{C}^{N \times 1}$ as follows:

$$\overline{\mathbf{x}}_k \triangleq \sum_{m=1}^{d_k} x^m_k \mathbf{v}^m_k, \tag{6}$$

where $\|\mathbf{v}^m_k\| = 1$ and $E(|x^m_k|^2) = \frac{P}{K.d_k}$.

For our IA scheme, we impose—in line with the analysis in [12]—three constraints on the input beamformers $\mathbf{v}^m_k$ and their corresponding output combiners $\mathbf{u}^m_k$.

$$\begin{aligned} &(\mathbf{u}^m_i)^H \, \hat{\mathbf{H}}_{i,k} \, (\mathbf{v}^p_k) = 0, \quad \forall i \neq k \in \{1, 2, \ldots, K\} \\ &\text{and } \forall m \in \{1, 2, \ldots, d_i\}, p \in \{1, 2, \ldots, d_k\}, \\ &\text{where } \hat{\mathbf{H}}_{i,k} \triangleq \operatorname{diag}\{\hat{\mathbf{h}}^c_{i,k}(0), \cdots, \hat{\mathbf{h}}^c_{i,k}(N-1)\}. \end{aligned} \tag{7}$$

$$\begin{aligned} &(\mathbf{u}^m_i)^H \, \mathbf{H}_{i,i} \, (\mathbf{v}^p_i) = 0, \\ &\forall i \in \{1, 2, \ldots, K\} \text{ and } \forall m \neq p \in \{1, 2, \ldots, d_i\}. \end{aligned} \tag{8}$$

$$\begin{aligned} &\left|(\mathbf{u}^m_i)^H \, \mathbf{H}_{i,i} \, (\mathbf{v}^m_i)\right| \geq c > 0, \\ &\forall i \in \{1, 2, \ldots, K\}, m \in \{1, 2, \ldots, d_i\}. \end{aligned} \tag{9}$$

One way to construct the input beamformers and output combiners would be to mimic the scheme of [14]. The only modification required in the construction of $\mathbf{v}^p_k$ is to use the frequency domain equivalents of the quantized values of the channel parameters rather than the actual parameters themselves, i.e., to substitute $\hat{\mathbf{H}}_{i,k}, k \neq i$ in place of $\mathbf{H}_{i,k}$. Once $\mathbf{v}^p_k, \ \forall k, p$ are obtained, receiver i chooses its output combiners $\mathbf{u}^p_k$ to be in a subspace orthogonal to the interference caused by other users under the assumption of the intervening channel matrix being $\hat{\mathbf{H}}_{i,k}$. This yields (7). A potential source of interference at receiver i arises from the need to zero-force the $d_i - 1$ messages from transmitter i other than the one being currently decoded. However, invoking our assumption that receiver i has perfect knowledge of both $\mathbf{H}_{i,i}$ and $\mathbf{v}^p_i$, we can construct the output beamformers $\mathbf{u}^m_i$ to satisfy (8). However, since $\hat{\mathbf{H}}_{i,k}$ is used in the above construction in place of the actual channel matrix $\mathbf{H}_{i,k}$, we have $(\mathbf{u}^m_i)^H \mathbf{H}_{i,k} (\mathbf{v}^p_k) \neq 0$. The implication of this is that the interference caused by the other users cannot be perfectly nulled at receiver i.

Note that the construction procedure for the input beamformers does not require the knowledge of $\mathbf{H}_{i,i}$ for any i. Further, in the construction of the output combiners as well, we do not require the knowledge of $\mathbf{H}_{j,j}, j \neq i$ at receiver i. This validates our approach of not quantizing $\mathbf{H}_{i,i}$ at each receiver.

Since we would need N to approach infinity to obtain the maximum sum degrees of freedom, we would require the number of taps L also to scale proportionately fast with the auxiliary parameter n. If such a condition is not imposed, the channel coefficients would not be independent across different tones and hence would violate a necessary condition needed in the proof of the CSITR case considered in [14]. Alternatively, motivated by the approximate equivalence of the number of tones N and the number of taps L in practice, we could model the channel vectors $\mathbf{h}_{i,k}(m) \ \forall m$ as directly being explicitly independent of each other. This corresponds to the bandwidth scaling requirement of [9] and the tap scaling requirement of [12].

D. Achievability Proof

Having established the feedback procedure, bounds on quantization error and some properties of the beamforming vectors, we can now analyze the rate expression. Our aim is to show that scaling the feedback rate N_f as $(K-1)(RL-1)\log P$ is sufficient to bound the interference terms in the rate expression. The destination D_i projects the received signal $\overline{\mathbf{y}}_i$ onto the d_i directions given by $\mathbf{u}_i^m \in \mathbb{C}^{RN\times 1}, m \in \{1,2,\ldots,d_i\}$,

$$\begin{aligned}(\mathbf{u}_i^m)^H \overline{\mathbf{y}}_i &= (\mathbf{u}_i^m)^H \mathbf{H}_{i,i}\mathbf{v}_i^m \mathbf{x}_i^m + \sum_{p\neq m} (\mathbf{u}_i^m)^H \mathbf{H}_{i,i}\mathbf{v}_i^p \mathbf{x}_i^p \\ &\quad + \sum_{k\neq i}\sum_{p=1}^{d_k} (\mathbf{u}_i^m)^H \mathbf{H}_{i,k}\mathbf{v}_k^p \mathbf{x}_k^p + (\mathbf{u}_i^m)^H \overline{\mathbf{z}}_i. \quad (10)\end{aligned}$$

Let us define $\mathbf{1}_R \triangleq [1,1,\ldots,1]^t \in \mathbb{R}^{R\times 1}$,

$$\mathbf{w}_{i,k} \triangleq \left[\mathbf{h}_{i,k}^H(0), \mathbf{h}_{i,k}^H(1), \ldots, \mathbf{h}_{i,k}^H(N-1)\right]^H \in \mathbb{C}^{RN\times 1},$$

and $\hat{\mathbf{w}}_{i,k} \triangleq [\hat{\mathbf{h}}_{i,k}^H(0), \hat{\mathbf{h}}_{i,k}^H(1), \ldots, \hat{\mathbf{h}}_{i,k}^H(N-1)]^H \in \mathbb{C}^{RN\times 1}$.

Define $(\mathbf{u}_i^m)^c \circ (\mathbf{v}_k^p \otimes \mathbf{1}_R)$ as $\mathbf{b}_{i,k}^{m,p} \in \mathbb{C}^{RN\times 1}$. Then, we can denote $(\mathbf{u}_i^m)^H \mathbf{H}_{i,k}\mathbf{v}_k^p$ as $\mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}$. Using this notation,

$$\begin{aligned}(\mathbf{u}_i^m)^H \overline{\mathbf{y}}_i &= \mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,m}\mathbf{x}_i^m + \sum_{p\neq m} \mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,p}\mathbf{x}_i^p \\ &\quad + \sum_{k\neq i}\sum_{p=1}^{d_k} \mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}\mathbf{x}_k^p + (\mathbf{u}_i^m)^H \overline{\mathbf{z}}_i. \quad (11)\end{aligned}$$

We choose the input symbols $\mathbf{x}_i^m$ to be i.i.d. $\mathrm{CN}(0, \frac{P}{Kd_i})$, and since the receiver knows $\mathbf{b}_{i,i}^{m,m}$ and $\mathbf{w}_{i,i}$, it can treat the other interference as noise to obtain a rate of

$$R_i = \frac{1}{N}\sum_{m=1}^{d_i} \log\left(1 + \frac{\frac{P}{Kd_i}\left|\mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,m}\right|^2}{I_{i,1} + I_{i,2} + N_o}\right), \quad (12)$$

where

$$I_{i,1} \triangleq \sum_{p\neq m} \frac{P}{Kd_i}\left|\mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,p}\right|^2, \quad (13)$$

and

$$I_{i,2} \triangleq \sum_{k\neq i}\sum_{p=1}^{d_k} \frac{P}{Kd_k}\left|\mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}\right|^2. \quad (14)$$

$I_{i,1}$ is the self-interference (treated as noise) caused by transmitter i involving messages other than the one being currently decoded by receiver i. $I_{i,2}$ is the inter-user interference due to transmitters other than transmitter i.

Our three conditions on the vectors $\mathbf{u}_i^m$ and $\mathbf{v}_i^m$, expressed before in (7)–(9), can be now be re-formulated now as

$$\hat{\mathbf{w}}_{i,k}^H \mathbf{b}_{i,k}^{m,p} = 0 \; \forall k \neq i, \quad \forall m,p, \quad (15)$$

$$\mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,p} = 0 \; \forall i, m \neq p, \quad (16)$$

$$\left|\mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,m}\right| \geq c > 0 \; \forall i, m. \quad (17)$$

In particular, (16) implies that the self-interference terms vanish for all values of P, i.e.

$$I_{i,1} = 0.$$

Note that the corresponding $I_{i,1}$ is not zero in the SISO analysis of [12]. Thus, not quantizing $\mathbf{H}_{i,i}$ not only requires a lower feedback rate but it also yields a higher information rate at all finite values of P.

The aim of the analysis below is to bound the inter-user interference term $I_{i,2}$ by a positive constant independent of the power P. This would ensure that the interference remains bounded in the relevant signal dimensions allowing for the attainment of the complete spatial multiplexing gain.

We observe that $\mathbf{w}_{i,i}$, $\hat{\mathbf{w}}_{i,i}$ and $\mathbf{b}_{i,i}^{m,p}$ are all RN-length vectors and that

$$\begin{aligned}\|\hat{\mathbf{w}}_{i,k}\|^2 &= \sum_{n=0}^{N-1}\left\|\hat{\mathbf{h}}_{i,k}(n)\right\|^2 = \sum_{m=1}^{R}\|\hat{\mathbf{c}}_{i,k}(m)\|^2 \\ &= \sum_{m=1}^{R}\|\hat{\mathbf{c}}_{i,k}[m]\|^2 = \|\hat{\mathbf{t}}_{i,k}\|^2 = 1.\end{aligned}$$

The third equality above follows from Parseval's theorem. Note that $\mathbf{w}_{i,k}$, $\hat{\mathbf{w}}_{i,k}$ and $\mathbf{b}_{i,k}^{m,p}$ all lie in the space $\mathbb{C}^{RN}$; and that $\hat{\mathbf{w}}_{i,k}$ and $\mathbf{b}_{i,k}^{m,p}$ are orthogonal to each other. This implies that one can start with $\hat{\mathbf{w}}_{i,k}$ and $\frac{\mathbf{b}_{i,k}^{m,p}}{\|\mathbf{b}_{i,k}^{m,p}\|}$ as two basis vectors for the space $\mathbb{C}^{RN}$, and complete the basis by finding other vectors using the usual Gram-Schmidt procedure. By invoking the conservation of energy principle on the representation of $\mathbf{w}_{i,k}$ as a sum of its projections along this particular basis, we can see that

$$\|\mathbf{w}_{i,k}\|^2 \geq \left|(\mathbf{w}_{i,k})^H \hat{\mathbf{w}}_{i,k}\right|^2 + \left|(\mathbf{w}_{i,k})^H \frac{\mathbf{b}_{i,k}^{m,p}}{\left\|\mathbf{b}_{i,k}^{m,p}\right\|}\right|^2.$$

Using this inequality, we can write

$$\begin{aligned}&\frac{P}{Kd_k}\cdot\left|\mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}\right|^2 \\ &\leq \frac{P}{Kd_k}\cdot\left\|\mathbf{b}_{i,k}^{m,p}\right\|^2\left(\|\mathbf{w}_{i,k}\|^2 - \left|(\mathbf{w}_{i,k})^H \hat{\mathbf{w}}_{i,k}\right|^2\right) \\ &= \frac{P}{Kd_k}\cdot\left\|\mathbf{b}_{i,k}^{m,p}\right\|^2\|\mathbf{w}_{i,k}\|^2\left(1 - \left|\frac{(\mathbf{w}_{i,k})^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{w}}_{i,k}\right|^2\right).\end{aligned}$$

By using Parseval's result, we can see that

$$\begin{aligned}\|\mathbf{w}_{i,k}\|^2 &= \sum_{n=0}^{N-1}\|\mathbf{h}_{i,k}(n)\|^2 = \sum_{m=1}^{R}\|\mathbf{c}_{i,k}(m)\|^2 \\ &= \sum_{m=1}^{R}\|\mathbf{c}_{i,k}[m]\|^2 = \|\mathrm{vec}(\mathbf{T}_{i,k})\|^2.\end{aligned}$$

Using this observation on $\|\mathbf{w}_{i,k}\|$, we can re-write

$$\begin{aligned}\frac{(\mathbf{w}_{i,k})^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{w}}_{i,k} &= \sum_{j=1}^{R}\frac{(\mathbf{c}_{i,k}(j))^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{c}}_{i,k}(j) \\ &= \sum_{m=1}^{R}\frac{(\mathbf{c}_{i,k}[m])^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{c}}_{i,k}[m] = \mathbf{t}_{i,k}^H\hat{\mathbf{t}}_{i,k}. \quad (18)\end{aligned}$$

In the above calculations, the first step comes from viewing the inner product of $\mathbf{w}_{i,k}$ and $\hat{\mathbf{w}}_{i,k}$ as the sum of inner products between the columns of matrices $\mathbf{F}_{i,k}$ and $\hat{\mathbf{F}}_{i,k}$. The second step follows from Parseval's theorem and reduces the term to corresponding columns of $\mathbf{T}_{i,k}$ and $\hat{\mathbf{T}}_{i,k}$. The last step follows from the definition of $\hat{\mathbf{t}}_{i,k}$ and the observation above that $\|\mathbf{w}_{i,k}\| = \|\mathrm{vec}(\mathbf{T}_{i,k})\|$.

This leads to a series of simple manipulations below where we invoke the definition of our distance metric over the Composite Grassmann to upper bound (up to a multiplicative constant) the $|\mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}|^2$ term by the maximum quantization error over the manifold.

$$\begin{aligned}
&\frac{P}{Kd_k} \cdot \left|\mathbf{w}_{i,k}^H \mathbf{b}_{i,k}^{m,p}\right|^2 \\
&\leq \frac{P}{Kd_k} \cdot \left\|\mathbf{b}_{i,k}^{m,p}\right\|^2 \|\mathbf{w}_{i,k}\|^2 \left(1 - \left|\frac{(\mathbf{w}_{i,k})^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{w}}_{i,k}\right|^2\right) \\
&= \frac{P}{Kd_k} \cdot \left\|\mathbf{b}_{i,k}^{m,p}\right\|^2 \|\mathbf{w}_{i,k}\|^2 \cdot \left(1 - \left|\mathbf{t}_{i,k}^H \hat{\mathbf{t}}_{i,k}\right|^2\right) \\
&= \frac{P}{Kd_k} \cdot \left\|\mathbf{b}_{i,k}^{m,p}\right\|^2 \|\mathbf{w}_{i,k}\|^2 \cdot d_c^2(\mathbf{t}_{i,k}, \hat{\mathbf{t}}_{i,k}) \\
&\leq \frac{P}{Kd_k} \cdot \left\|\mathbf{b}_{i,k}^{m,p}\right\|^2 \|\mathbf{w}_{i,k}\|^2 \cdot \sum_{j=1, j\neq i}^{K} d_c^2(\mathbf{t}_{i,j}, \hat{\mathbf{t}}_{i,j}) \\
&= \frac{P}{Kd_k} \cdot \left\|\mathbf{b}_{i,k}^{m,p}\right\|^2 \|\mathbf{w}_{i,k}\|^2 \cdot d^2(\mathbf{Q}_i, \hat{\mathbf{Q}}_i) \\
&\leq P.\ \text{const.}\ \triangle_{\max}^2 = \text{const}
\end{aligned}$$

Hence, the inter-user interference $I_{i,2}$ is bounded independent of the power P. Now, (12) yields

$$\begin{aligned}
d_{sum} &= \lim_{P\to\infty} \frac{R_{sum}}{\log P} \\
&\geq \sum_{i=1}^{K}\sum_{m=1}^{d_i} \lim_{P\to\infty} \frac{\log\left(1 + \frac{\frac{P}{Kd_i}\left|\mathbf{w}_{i,i}^H \mathbf{b}_{i,i}^{m,m}\right|^2}{I_{i,1}+I_{i,2}+N_o}\right)}{N \log P} \\
&= \frac{\sum_{i=1}^{K} d_i}{N} \\
&= \frac{(R+1)R(n+1)^\tau + (K-R-1)Rn^\tau}{(R+1)(n+1)^\tau}.
\end{aligned}$$

The last step above follows from (3), (4) and (5). Taking the supremum over all values of n, we obtain the desired result that $d_{sum} = \frac{RK}{R+1}$.

The reciprocity of alignment idea developed in [18] and employed in [14] allows for the analogous construction of the input and output IA beamformers satisfying (7)–(9) for the $R \times 1$ MISO channel as well. From the quantization perspective, both the number of parameters to be quantized and the variable $Q_i \in G_{RL,1}^{K-1}$ remain unaltered and the same feedback scaling rate is applied for the $R \times 1$ MISO scenario as well. The only minor change lies in the row vector-column vector exchange in the system model and the consequent modifications required in the manipulation of the rate expression which are straight-forward. This concludes the proof of Theorem 1.

E. Grassmann Quantization

While quantization over the composite Grassmann manifold is preferred over multiple quantizations over individual Grassmann manifolds for the conceptual clarity of collating all necessary feedback information into a single variable, it is not indispensable for the proof of Theorem 1. In case a suitable code over the composite Grassmann is not found or if complexity of decoding is an issue, then multiple quantizations over the individual Grassmann manifold could be used at the expense of a somewhat lower rate than in the joint quantization scenario as explained below. Note that, ignoring the multiplicative $(1 + O(\delta^2))$ factor, the normalized volume of a geodesic ball of radius δ in both $G_{RL,1}$ and $G_{RL,1}^{K-1}$ can be written in the format $\mu(B(\delta)) = c\delta^D$, where for $G_{RL,1}$, $c = 1$ and $D = 2(RL-1)$ and for $G_{RL,1}^{K-1}$, $c = \frac{\Gamma^{K-1}(RL)}{\Gamma((K-1)(RL-1)+1)}$ and $D = 2(K-1)(RL-1)$. Explicitly computing the constant involved in (1), we get

$$\triangle_{max}^2 \leq \frac{4}{(c\, 2^{N_f})^{\frac{2}{D}}}.$$

Based on this result, it is seen that joint quantization over $G_{RL,1}^{K-1}$ yields a tighter bound than individual quantization over $G_{RL,1}$. Let us define $b_{max}^2 = \max_{i,k,m,p} \|\mathbf{b}_{i,k}^{m,p}\|^2$, and $h_{max}^2 = \max_{i,k} \|\mathbf{H}_{i,k}\|^2$. Note that both these quantities are independent of the power P. We can now bound the inter-user interference as

$$I_{i,2} \leq \frac{4\, b_{max}^2 h_{max}^2}{K} c_3,$$

where $c_3 = (K-1)$ for quantization over $G_{RL,1}$ as compared to

$$c_3 = \left[\frac{\Gamma^{K-1}(RL)}{\Gamma\left((K-1)(RL-1)+1\right)}\right]^{-\frac{1}{(K-1)(RL-1)}}$$

for quantization over $G_{RL,1}^{K-1}$ which is smaller than $K-1$. As the value of K increases, the bound for joint quantization under the composite Grassmann manifold gets progressively tighter in the sense that the difference between the c_3's increases with K.

III. Extensions to MIMO Systems and Lower Feedback Scaling Rates

A. IA for the MIMO Case With Limited Feedback

Under the assumption of global CSITR, [14] demonstrated the achievability of $d_{sum} = \frac{R}{R+1} \cdot \min\{M_t, M_r\}K$ for a $M_t \times M_r$ K-user interference channel with $K > R$, where $R = \lfloor \frac{\max\{M_t, M_r\}}{\min\{M_t, M_r\}} \rfloor$.

Theorem 3: In a K-user interference channel where each transmitter has M_t antennas, each receiver has M_r antennas, $K > R = \lfloor \frac{\max\{M_t, M_r\}}{\min\{M_t, M_r\}} \rfloor$, and each transmit/receive pair of nodes has an L-tap frequency selective channel between them, an IA scheme under limited feedback achieves the same degrees of freedom as the IA scheme with CSITR if the receiver employs at least $\min\{M_t, M_r\}(K \min\{M_t, M_r\} - 1)(RL - 1)\log P$ bits of feedback.

Proof: Due to the reciprocity of alignment idea [18], we can concentrate our attention on the case of $M_t \leq M_r$ without loss of generality. From each receiver node, we discard $M_r - RM_t$ antennas. Now, by restricting the cooperation allowed amongst the transmitters and the receivers, we can treat this K-user $M_t \times RM_t$ channel as a SIMO KM_t-user $1 \times R$ channel. As long as each user-utilizing composite (or individual) Grassmann quantization-sends $(KM_t - 1).(RL - 1).\log P$ bits of feedback, we know from the preceding section that a spatial muliplexing gain of $\frac{KM_tR}{R+1}$ can be achieved. This value of $\frac{KM_tR}{R+1}$ matches with the desired inner bound achieved with perfect channel knowledge in [14]. Combining back the receivers to get the K-user channel, we conclude that each node needs to transmit at least $M_t(KM_t - 1)(RL - 1)\log P$ bits of feedback. ■

There is thus an interesting connection between the feedback scaling rates obtained above and the dimension of the quantization manifold encapsulating all the feedback information. In the SISO case [12], the feedback information can be seen to be represented by a point on the composite Grassmannian $G^K_{L,1}$, and the pre-log factor in the feedback scaling rate was found to be half the real dimension of this manifold. In the SIMO $1 \times M_r$ case, we have represented the feedback information by a point on the manifold $G^{K-1}_{M_rL,1}$, whereas in the MIMO $M_t \times M_r$ case, we need $\min\{M_t, M_r\}$ points on the manifold $G^{K\min\{M_t,M_r\}-1}_{RL,1}$. Correspondingly, we have found the pre-log factors in the SIMO and MIMO cases to be $\frac{1}{2}\dim(G^{K-1}_{M_rL,1})$ and $\min\{M_t, M_r\} \times \frac{1}{2}\dim(G^{K\min\{M_t,M_r\}-1}_{RL,1})$, respectively. A general result connecting the feedback scaling rate and the dimension of the quantization manifold is not currently known. However, a conjecture is provided and discussed in the Appendix, thereby also providing a heuristic explanation for the above results. If this connection did indeed exist, then for a given achievable scheme, the identification of the manifold of smallest dimension in which the feedback information (that is sufficient to achieve the full degrees of freedom) lies would be the key to solving the question of the minimum feedback scaling rate needed for that scheme to achieve its full degrees of freedom. It is also an interesting question as to whether there are alternative constructions that are degrees of freedom optimal but require less feedback information in the sense described above. The next section tackles the question of what guarantees can be given with smaller feedback scaling rates for the IA scheme of [14].

B. Impact of Smaller Feedback Scaling Rate

\It would be desirable to have a tradeoff between a flexible feedback rate for an individual user and the degrees of freedom obtained thereby, with the system-level strategy being optimized for achieving the Pareto-optimal point maximizing the sum degrees of freedom of the network. This tradeoff provides flexibility to the system designer in terms of allocation of resources for feedback and obtainable rates is given by the theorem below.

We denote the degrees of freedom achieved by user i as $\tilde{d}_i \triangleq \frac{d_i}{N}$, where $d_i \triangleq \lim_{P\to\infty} \frac{R_i}{\log P}$ and R_i is the rate achieved by user i over N blocks.

Theorem 4: If the feedback rate employed by receiver i in the MIMO interference channel (including the SIMO/MISO models) scales as $\alpha_i.N_f$, for some $0 < \alpha_i \leq 1$, then user i can achieve $\alpha_i.\tilde{d}_i$ degrees of freedom by using interference alignment.

Proof: We analyze below for a K-user SIMO $1 \times R$ channel; the corresponding MIMO results follow similarly. Intuitively, the quantization error is proportional to $P^{1-\alpha_i}$ and this acts as a principal component of the interference faced by user i, leading to a reduction of $(1 - \alpha_i)\tilde{d}_i$ degrees of freedom from user i. With b^2_{max} and h^2_{max} as defined at the end of Section II, we have

$$\begin{aligned} I_{i,2} &\leq b^2_{max}.\ h^2_{max}.\ \text{const.}\ P.\sum_{k=1;k\neq i}^{K} d_c^2(\mathbf{t}_{i,k}, \hat{\mathbf{t}}_{i,k}) \\ &\leq P.\ \text{const.}\ \triangle^2_{\max,i}, \end{aligned}$$

where $\triangle_{\max,i}$ is the maximum quantization error possible while using the codebook of receiver i with $2^{\alpha_i N_f}$ representations on $G^K_{RL,1}$. As in (1), we have

$$\begin{aligned} \triangle_{\max,i} &\leq \frac{\text{const}}{2^{\frac{\alpha_i N_f}{2(K-1)(RL-1)}}} \\ \implies \triangle^2_{\max,i} &\leq \frac{\text{const}}{2^{\frac{\alpha_i (K-1)(RL-1)\log P}{(K-1)(RL-1)}}} = \frac{\text{const}}{P^{\alpha_i}} \end{aligned}$$

This leads to $I_{i,2} \leq \text{const}.P^{1-\alpha_i}$ which, using the lower bound for R_i from (12) (with $I_{i,1} = 0$), implies

$$MG(R_i) \geq \frac{1}{N}\left(d_i - (1 - \alpha_i)d_i\right) = \frac{\alpha_i d_i}{N} = \alpha_i \tilde{d}_i,$$

where $MG(x)$ denotes the multiplexing gain of the quantity x defined as $MG(x) = \lim_{P\to\infty} \frac{x}{\log P}$. ■

In particular, if $\alpha_1 = \alpha_2 = \ldots \alpha_K = \alpha$ (say), then one can realize a sum degrees of freedom of $d_{sum} = \alpha.\frac{RK}{K+1}$ in a K-user SIMO $1 \times R$ channel. A similar result would also hold for the MIMO $M_t \times M_r$ channel. Also, note that while the rate achieved by user i would depend upon $\{\alpha_j\}_{j=1}^K$, the degrees of freedom depends only on α_i as long as the other α_j's are positive.

Corollary 5: The full spatial multiplexing gain of the IA scheme of [14] is attained if the feedback scaling rate is changed from the N_f of Theorem 3 to $N_f - o(\log P)$.

The proof of this corollary is left to the reader. It follows in a straightforward manner from the proofs of Theorems 1 and 3.

IV. Random Quantization

While a general upper bound for the sum degrees of freedom attainable over the interference channel at a given feedback scaling rate is unknown, we consider the limited feedback IA scheme that is similar to that of the previous sections but that uses a random quantization codebook. Under such random vector quantization we show that the tradeoff between the attainable sum degrees of freedom (obtained after averaging over the ensemble of random codebooks) and feedback scaling rate of Theorems 3 and 4 cannot be improved, thus lending a measure of tightness to the sufficient condition on the feedback scaling rate of those theorems to attain any given sum degrees of freedom (up to the maximum attainable). Since the extension of results on the SIMO channel to the MISO/MIMO cases follows a straightforward computation, we concentrate on the SIMO setup of Section II below. The constraints on the feedback and transmission strategy are as follows:

1) As before, we restrict ourselves to treating the quantized channel estimates as being perfect within the ambit of the linear IA strategy of [14].
2) The i-th transmitter sends i.i.d. messages x_i^m distributed as $\mathrm{CN}(0, \frac{P}{Kd_i})$.
3) The i-th receiver individually quantizes each vector $\mathbf{v}_{i,k}$ $\forall k \in \{1,\ldots,i-1,i+1,\ldots,K\}$ using a random code $\mathcal{C}_{rand}$ uniformly distributed over $G_{RL,1}$. The rate and the sum degrees of freedom metrics are evaluated as averages over the ensemble of all random codes.
4) Denoting the feedback rate by N_f, we assume that each vector $\mathbf{v}_{i,k}$ is quantized using $B = \frac{N_f}{K-1}$ bits.

For notational clarity, let us denote the full sum degrees of freedom for the interference channel as d_{sum}^{CSITR} and the feedback scaling rate found earlier (in Theorems 1 and 3) to be sufficient for attaining it to be N_f^{ideal}.

In the lemma below, $\mathbf{w}_{i,k}$ and $\hat{\mathbf{w}}_{i,k}$ are as defined in Section II-D for the SIMO scenario. While $\mathbf{w}_{i,k}$ remains a constant vector, $\hat{\mathbf{w}}_{i,k}$ is now a random unit norm vector with the randomness arising out of the use of a code $\mathcal{C}_{rand} = \{\mathbf{q}_1,\ldots,\mathbf{q}_{2^B}\}$ with $\mathbf{q}_i \sim \mathrm{Unif}(G_{RL,1})$ $\forall i \in \{1,\ldots,2^B\}$. We also note that the expectation below is over the ensemble of all random codes and not over the channel distribution.

Lemma 6: Let $Z \triangleq d_c^2(\frac{\mathbf{w}_{i,k}}{\|\mathbf{w}_{i,k}\|}, \hat{\mathbf{w}}_{i,k})$. Then,

$$E_{\mathcal{C}rand}(\log(Z)) = -\frac{\log e}{RL-1}\sum_{n=1}^{2^B}\frac{1}{n}.$$

Proof: Recalling the calculation in (18), we can write

$$Z = 1 - \left|\frac{\mathbf{w}_{i,k}^H}{\|\mathbf{w}_{i,k}\|}\hat{\mathbf{w}}_{i,k}\right|^2 = 1 - \left|\mathbf{t}_{i,k}^H\hat{\mathbf{t}}_{i,k}\right|^2.$$

Let $\mathbf{q}_1$ be a unit norm RL-length vector uniformly distributed over the $G_{RL,1}$ manifold that is independent of the vector $\mathbf{v}_{i,k}$. For any fixed unitary matrix $\mathbf{U}$, we have $\mathbf{q}_1 =_d \mathbf{U}\mathbf{q}_1$, which implies that

$$\left|\mathbf{t}_{i,k}^H\mathbf{q}_1\right|^2 =_d \left|\mathbf{t}_{i,k}^H(\mathbf{U}\mathbf{q}_1)\right|^2 = \left|(\mathbf{U}^H\mathbf{t}_{i,k})^H\mathbf{q}_1\right|^2.$$

Since the above equality holds for all unitary matrices, it continues to hold if $\mathbf{U}$ is independent of $\mathbf{q}_1$ and Haar distributed over the unitary group $\mathcal{U}(RL)$. Noting that $\mathbf{U}^H\mathbf{t}_{i,k}$ is uniformly distributed over $G_{RL,1}$, we conclude that

$$\left|\mathbf{t}_{i,k}^H\mathbf{q}_1\right|^2 =_d \left|\mathbf{q}_2^H\mathbf{q}_1\right|^2,$$

where $\mathbf{q}_2$ is another unit norm RL-length vector uniformly distributed over the $G(RL,1)$ manifold that is independent of $\mathbf{q}_1$. Based on the discussion in [22] (see page 5048), we can conclude that $1-|\mathbf{q}_2^H\mathbf{q}_1|^2$, and consequently $1-|\mathbf{t}_{i,k}^H\mathbf{q}_1|^2$, has the $\mathrm{beta}(RL-1,1)$ distribution. Since Z is defined as

$$Z = \min_{\mathbf{q}_i \in \mathcal{C}_{rand}}\left|\mathbf{t}_{i,k}^H\mathbf{q}_i\right|^2,$$

it is the minimum of 2^B i.i.d. $\mathrm{beta}(RL-1,1)$ random variables, and from Lemma 3 of [22], we get

$$E_{\mathcal{C}_{rand}}(\log(Z)) = -\frac{\log_2 e}{RL-1}\sum_{n=1}^{2^B}\frac{1}{n}.$$

■

In the discussion below, we slightly redefine the notion of sum degrees of freedom. We view it as an average sum degrees of freedom obtained through an averaging over the ensemble of all random codes.

Theorem 7: Under the constraints described above, a receiver scaling its feedback rate at $\alpha.N_f^{ideal}$ for some $0 < \alpha \leq 1$ cannot attain more than $\alpha.d_{sum}^{CSITR}$ sum degrees of freedom.

Proof: We prove the theorem only for the SIMO scenario, since its extension to the MISO/MIMO cases is straightforward. The proof follows an analogous result for the broadcast channel in [22]. Note that, the expression for the sum degrees of freedom can be split into two parts as

$$d_{sum} = d_{sum}^1 - d_{sum}^2,$$

where

$$d_{sum}^1 = \lim_{P,N\to\infty} E_{\mathcal{C}_{rand}}\frac{1}{N\log P} \times \sum_{i=1}^{K}\sum_{m=1}^{d_i}\log\left(N_o + I_{i,2} + \frac{P}{Kd_i}\left|\mathbf{w}_{i,i}^H\mathbf{b}_{i,i}^{m,m}\right|^2\right),$$

and

$$d_{sum}^2 = \lim_{P,N\to\infty} E_{\mathcal{C}_{rand}}\frac{1}{N\log P}\sum_{i=1}^{K}\sum_{m=1}^{d_i}\log(N_o + I_{i,2}).$$

By noting that for all values of i,k,m and p, $|\mathbf{w}_{i,k}^H\mathbf{b}_{i,k}^{m,p}|^2 \leq h_{max}^2 b_{max}^2 \leq \mathrm{const}$, we get

$$d_{sum}^1 \leq \lim_{P,N\to\infty} E_{\mathcal{C}_{rand}}\sum_{i=1}^{K}\sum_{m=1}^{d_i}\frac{\log(\mathrm{const} + P\mathrm{const})}{N\log P} = d_{sum}^{CSITR}.$$

This motivates us to interpret d_{sum}^1 as the sum degrees of freedom under perfect global knowledge of the channel states at the transmitters, and d_{sum}^2 as the loss in degrees of freedom attributable to imprecise channel knowledge arising from limited feedback. To obtain a lower bound on d_{sum}^2, we can ignore the positive N_o term in the numerator, and by further isolating the P factor within $I_{i,2}$, we get

$$d_{sum}^2 \geq d_{sum}^{CSITR} - \lim_{P,N\to\infty} E_{\mathcal{C}_{rand}} \sum_{i=1}^{K}\sum_{m=1}^{d_i}\frac{\log\left(\sum_{k\neq i}\sum_{p=1}^{d_k}\frac{1}{Kd_k}\left|\mathbf{w}_{i,k}^H\mathbf{b}_{i,k}^{m,p}\right|^2\right)}{N\log P}.$$

Combining the expressions for d_{sum}^1 and d_{sum}^2, and by replacing $\mathbf{w}_{i,k}$ by $\frac{\mathbf{w}_{i,k}}{\|\mathbf{w}_{i,k}\|}$, we obtain a bound on d_{sum} as

$$d_{sum} \le -\lim_{P,N\to\infty} \sum_{i=1}^{K}\sum_{m=1}^{d_i} E_{\mathcal{C}_{rand}} \frac{\log\left(\sum_{k\neq i}\sum_{p=1}^{d_k}\frac{1}{Kd_k}\left|\frac{\mathbf{H}_{i,k}^H}{\|\mathbf{H}_{i,k}\|}\mathbf{b}_{i,k}^{m,p}\right|^2\right)}{N\log P}.$$

We note that one can always find a unit-norm vector $\mathbf{s}\in\mathbb{C}^{RN\times 1}$ satisfying $\mathbf{s}^H\hat{\mathbf{w}}_{i,k}=0$, such that

$$\frac{\mathbf{w}_{i,k}}{\|\mathbf{w}_{i,k}\|}=\sqrt{1-Z}\hat{\mathbf{w}}_{i,k}+\sqrt{Z}\mathbf{s}.$$

This implies that

$$\begin{aligned}\left|\frac{\mathbf{w}_{i,k}^H}{\|\mathbf{w}_{i,k}\|}\mathbf{b}_{i,k}^{m,p}\right|^2 &= (1-Z)\left|\hat{\mathbf{w}}_{i,k}^H\mathbf{b}_{i,k}^{m,p}\right|^2+Z\left|\mathbf{s}^H\mathbf{b}_{i,k}^{m,p}\right|^2\\ &= Z|\mathbf{s}^H\mathbf{b}_{i,k}^{m,p}|^2.\end{aligned}$$

$$d_{sum} \le -\lim_{P,N\to\infty} E_{\mathcal{C}_{rand}} \sum_{i=1}^{K}\sum_{m=1}^{d_i}\frac{\log(Z)+\log\left(\sum_{k\neq i}\sum_{p=1}^{d_k}\left|\mathbf{s}^H\mathbf{b}_{i,k}^{m,p}\right|^2\right)}{N\log P}.$$

Note that while the magnitude of the quantization error decreases with the power P, the error direction $\mathbf{s}$ remains independent of P. Hence, the expression simplifies in the regime of high P to

$$\begin{aligned}d_{sum} &\le -\lim_{P,N\to\infty} E_{\mathcal{C}_{rand}}\sum_{i=1}^{K}\sum_{m=1}^{d_i}\frac{\log(Z)}{N\log P}\\ &= -\,d_{sum}^{CSITR}\lim_{P\to\infty}\frac{E_{\mathcal{C}_{rand}}\log(Z)}{\log P}\\ &= -\,d_{sum}^{CSITR}\lim_{P\to\infty}\frac{-\log e}{RL-1}\sum_{k=1}^{2^B}\frac{1}{k}\frac{1}{\log P}\\ &\le d_{sum}^{CSITR}\lim_{P\to\infty}\frac{1}{RL-1}\frac{\mathrm{const}+\log(2^B)}{\log P}\\ &= d_{sum}^{CSITR}\lim_{P\to\infty}\frac{\frac{B}{RL-1}}{\log P}.\end{aligned}$$

In the calculations above, we have used the fact that $\sum_{k=1}^{2^B}\frac{1}{k}=1+\sum_{k=2}^{2^B}\frac{1}{k}=1+\int_1^{2^B}\sum_{k=2}^{2^B}1_{[\frac{1}{k-1},\frac{1}{k})}(x)dx\le 1+\int_1^{2^B}\frac{1}{x}dx=1+\log_e(x)$.

Thus, if $N_f=\alpha(K-1)(RL-1)\log P$ or equivalently, $B=\alpha(RL-1)\log P$, we can upper bound the sum degrees of freedom as

$$d_{sum}\le\alpha\, d_{sum}^{CSITR}.$$

■

V. Conclusion

Multi-user interference channels with multiple antennas at each node are analyzed in a frequency-selective setup, wherein the receivers with knowledge of their respective channels quantize the channel directions using a code-book on the composite Grassmann or multiple individual manifolds and broadcast them to all other nodes at a rate that scales sufficiently fast with the power constraint on the nodes. It is shown that an interference alignment scheme based on treating these channel estimates as being perfect is sufficient to attain the same sum degrees of freedom as the interference alignment implementation utilizing perfect channel state information at all the nodes. We also demonstrate a continuous tradeoff whereby an individual user can opt for a slower scaling of feedback bits and obtain proportionally lower degrees of freedom. Under random vector quantization of individual channel vectors, the minimum feedback scaling rate needed to guarantee the attainment of the full (or given) sum degrees of freedom within the ambit of the interference alignment strategy is found. A information theoretic converse establishing an upper bound on the achievable sum degrees of freedom without constraining the transmission or feedback strategies is of interest but not known at this time and is a subject for further investigation.

Appendix

In this appendix, we provide a heuristic explanation for why feedback scaling rate N_f required for emulating CSITR performance may be of the form $\frac{N}{m}\log P$ for general quantization manifolds, with N being the dimension of the quantization manifold and m an integer.

Let us denote the channel parameter whose knowledge enables the system to realize CSITR-like performance as q, and the surface on which it takes values as the manifold M. If the real dimension of M is given by $\dim M=N$, then we know that around each point on the manifold, we can establish a local coordinate system with N variables. Consider such a coordinate chart around the optimal value of q, denoted as q_{opt}. With knowledge of q_{opt}, suppose a receiver feeds back this information using N_f bits. By following the argument in the paper, one can bound the maximum distortion under quantization using a 2^{N_f}-level sphere-packing codebook on the manifold as $\triangle_{\max}\le\frac{\mathrm{const}}{2^{\frac{N_f}{N}}}$. The receiver quantizes q_{opt} as some $\hat{q}$. It is the difference between q_{opt} and $\hat{q}$ that can potentially prevent us from achieving the ideal CSITR degrees of freedom. So the aim would be to find an appropriate feedback scaling rate that collapses the difference $q_{opt}-\hat{q}$ sufficiently fast to prevent any loss of degrees of freedom. In the regime of high N_f, it is reasonable to assume that $\hat{q}$ would be close to q_{opt} and hence we can take $\hat{q}$ to lie within the same local coordinate system that covers q_{opt}. Isolating the key part of the performance measure affected by the inaccuracy in q as $f(q)$, we get a first-order Taylor approximation as

$$f(\hat{q})\approx f(q_{opt})-\left|(\nabla f)^t(\triangle q)\right|,$$

where ∇f represents the gradient of the function f and $\triangle q$ represents the displacement of $\hat{q}$ from q_{opt}. Note that this is a coarse approximation chosen primarily for illustration; to make it rigorous, one would have to choose a normal coordinate system based at q_{opt} and calculate distances along radial geodesics emanating from it. For illustration purposes, one can visualize $f(q)$ to be the achievable rate of the system, which increases with the transmit power P. Then $f(q_{opt})$ represents the CSITR value of the achievable rate and $f(\hat{q})$ represents the same under finite rate feedback. It is reasonable to expect $f(q_{opt})$, $f(\hat{q})$ and ∇f to be functions of P.

We want to bound the difference between $f(\hat{q})$ and $f(q_{opt})$ by a constant independent of P; as we expect to see both $f(\hat{q})$ and $f(q_{opt})$ increase monotonically with P. The variation of the (∇f) with P determines the behavior required of the $\triangle q$ term. If (∇f) varies as $P^{\frac{1}{m}}$, then we need $\triangle q$ to vary as $\frac{1}{P^{\frac{1}{m}}}$. This is ensured by

$$\|\triangle q\|^m \leq \frac{1}{P} \Rightarrow \triangle^m_{\max} \leq \frac{1}{P}$$
$$\times \Rightarrow \frac{\text{const}}{2^{\frac{mN_f}{N}}} \leq \frac{1}{P} \Rightarrow N_f \approx \frac{N}{m} \log P \text{ as } P \text{ increases.}$$

Thus, if N_f varies as $\frac{N}{m}\log P$ with N being the dimension of the quantization manifold, then the product $(\nabla f)^t(\triangle q)$ becomes independent of the power P. Then the term $\frac{f(q_{opt})-f(\hat{q})}{\log P}$ vanishes as P increases enabling us to attain the CSITR value for the degrees of freedom measure. This informal argument provides an intuitive reason for our results on the feedback scaling rate in this paper. Making the argument precise and checking its validity in detail is a subject of future enquiry.

References

[1] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[2] T. S. Han and K. Kobayashi, "A new achievable region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[3] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 2, pp. 581–586, Feb. 2004.

[4] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.

[5] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[6] S. Karmakar and M. K. Varanasi, "The capacity region of the MIMO interference channel and its reciprocity to within a constant gap," *IEEE Trans. Inf. Theory*, 2013, 10.1109/TIT.2013.2256231, preprint.

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

[8] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. n Theory*, vol. 7, no. 1, pp. 1–136, 2011.

[9] L. Grokop, D. N. C. Tse, and R. D. Yates, "Interference alignment for line-of-sight channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5820–5839, Sep. 2011.

[10] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," presented at the IEEE Int. Conf. Commun., Dresden, Germany, 2009.

[11] A. S. Motahari, S. Oveis-Gharan, M. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single-antenna systems," *IEEE Trans. Inf. Theory* Nov. 2009 [Online]. Available: http://arxiv.org/abs/0908.2282

[12] J. Thukral and H. Bolcskei, "Interference alignment with limited feedback," in *IEEE Intnl. Symp. Inf. Theory*, Seoul, South Korea, 2009.

[13] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finte-rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[14] T. Gou and S. A. Jafar, "Degrees of freedom for the K user M × N MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.

[15] R. T. Krishnamachari and M. K. Varanasi, Interference Alignment Under Limited Feedback for the MIMO Interference Channel Nov. 2009 [Online]. Available: http://arxiv.org/pdf/0911.5509

[16] J. S. Kim, S. H. Moon, S. R. Lee, and I. Lee, "A new channel quantization strategy for MIMO interference alignment with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 358–366, Jan. 2012.

[17] O. E. Ayach and R. H. Heath, "Interference alignment with analog channel state feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 626–636, Feb. 2012.

[18] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[19] W. Dai, "Communications With Finite Rate Feedback and Quantization on Grassmann Manifolds," Ph.D. dissertation, Univ. Colorado, Boulder, CO, USA, 2007.

[20] W. Dai, B. Rider, and Y. Liu, "Joint beamforming for multiaccess MIMO systems with finite rate feedback," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2618–2628, May 2009.

[21] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes etc: Packings in Grassmannian spaces," *J. Exp. Math.*, vol. 5, pp. 139–159, 1996.

[22] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

Rajesh T. Krishnamachari received his B. Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 2003. He worked as a R&D Engineer in the telecommunications industry between 2003-2005. In January 2006, he began his doctoral studies at the Electrical, Computer and Energy Engineering Department at the University of Colorado, Boulder, CO, from where he received his Ph.D. degree in December, 2011. Since 2012, he has been employed in the financial services industry in New York, NY.

Mahesh K. Varanasi (S'87–M'89–SM'95–F'10) received the Ph.D.degree in electrical and computer engineering from Rice University, Houston, TX, in 1989.

He joined what is now the Electrical, Computer and Energy Engineering Department of the University of Colorado in Boulder, CO, USA, as an Assistant Professor in 1989. In the same department, he was an Associate Professor between 1996–2001 and has been a full Professor since 2001. He is currently an affiliated faculty member in the Department of Mathematics and the Department of Applied Mathematics at the University of Colorado, and in the Department of Computer Science at the King Abdulaziz University, Jeddah, KSA. His research and teaching interests are in the areas of network communication and information theory, wireless communications and coding, detection and estimation theory, and signal processing. He has published on a variety of topics in these fields and is a Highly Cited Researcher in the "Computer Science" category according to the ISI Web of Science. He served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS during 2007–2009.