

Codebook-Based Opportunistic Interference Alignment

Hyun Jong Yang, *Member, IEEE*, Bang Chul Jung, *Senior Member, IEEE*, Won-Yong Shin, *Member, IEEE*, and Arogyaswami Paulraj, *Fellow, IEEE*

Abstract—For the multiple-input multiple-output interfering multiple-access channels (IMACs), opportunistic interference alignment (OIA) using the singular value decomposition (SVD)-based beamforming at each user fundamentally reduces the user scaling condition required to achieve any target DoF, compared to that for the single-input multiple-output IMAC. In this paper, we tackle two practical challenges of the existing SVD-based OIA: 1) the need of full feedforward of the selected users' beamforming weight vectors and 2) a low rate achieved based on the existing zero-forcing receiver. We first propose a codebook-based OIA, in which the weight vectors are chosen from a predefined codebook with a finite size so that information of the weight vectors can be sent to the belonging BS with limited feedforward. We derive the codebook size required to achieve the same user scaling condition as the SVD-based OIA case for both Grassmannian and random codebooks. Surprisingly, it is shown that the derived codebook size is the same for the two codebook methods. Second, we introduce an enhanced receiver at the base stations (BSs) in pursuit of further improving the achievable rate. Assuming no collaboration between the BSs, the interfering links between a BS and the selected users in neighboring cells are difficult to be acquired at the belonging BS. We propose the use of a simple minimum Euclidean distance receiver operating with no information of the interfering links. With the help of the OIA, we show that this new receiver asymptotically achieves the channel capacity as the number of users increases.

Index Terms—Codebook, degrees-of-freedom (DoF), interfering multiple-access channel (IMAC), limited feedforward, opportunistic interference alignment (OIA).

Manuscript received August 08, 2013; revised January 04, 2014 and April 08, 2014; accepted April 08, 2014. Date of publication April 22, 2014; date of current version May 09, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yongming Huang. This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (MSIP) (2012R1A1A1044151) and the Ministry of Education (2013R1A1A2A10004905), and by the 2014 Research Fund(1.130065.01) of UNIST. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Cambridge, MA, USA, July 2012, and is to be presented in part at the IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, Italy, May 2014.

H. J. Yang is with the School of Electrical and Computer Engineering, UNIST, Ulsan 689-798, Republic of Korea (e-mail: hjyang@unist.ac.kr).

B. C. Jung is with the Department of Information and Communication Engineering, Gyeongsang National University, Tongyoung 650-160, Republic of Korea (e-mail: bcjung@gnu.ac.kr).

W.-Y. Shin is with the Department of Computer Science and Engineering, Dankook University, Yongin 448-701, Republic of Korea (e-mail: wyshin@dankook.ac.kr).

A. Paulraj is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305-9510 USA (e-mail: apaulraj@stanford.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2014.2319774

I. INTRODUCTION

INTERFERENCE alignment (IA) [1], [2] is the key ingredient to achieve the optimal degrees-of-freedom (DoF) for a variety of interference channel models. The conventional IA framework, however, has several well-known practical challenges: global channel state information (CSI) and arbitrarily large frequency/time-domain dimension extension. Recently, the concept of opportunistic interference alignment (OIA) was introduced in [3], [4], for the K -cell single-input multiple-output (SIMO) interfering multiple-access channel (IMAC), where there are one M -antenna base station and N users in each cell. In the OIA scheme for the SIMO IMAC, S ($S \leq M$) users amongst the N users are opportunistically selected in each cell in the sense that inter-cell interference is aligned at a pre-defined interference space. Even if several studies have independently addressed one or a few of the practical problems (see [5], [6]), the OIA scheme simultaneously resolves the aforementioned issues. Specifically, the OIA scheme operates with i) local CSI acquired via pilot signaling, ii) no dimension extension in the time/frequency domain, iii) no iterative optimization of precoders, and iv) no coordination between the users or the BSs. It has been shown that there exists a trade-off between the achievable DoF and the number of users, which can be characterized by a *user scaling condition* [4], [7], [8]. Similarly, the analysis of the scaling condition of some system parameters required to achieve a target performance have been widely studied to provide a remarkable insight into the convergence rate to the target performance, e.g., the user scaling condition to achieve target DoF for the IMAC [3], [4], [7], [8], the scaling condition of the number of feedback bits to achieve the optimal DoF for multiple-input multiple-output (MIMO) interference channels [9], [10], and the codebook size scaling condition to achieve the target achievable rate for limited feedback MIMO systems [11]–[13]. For the SIMO IMAC, the OIA scheme asymptotically achieves KS DoF, for $0 < S \leq M$, if the number of per-cell users, N , scales faster than $\text{SNR}^{(K-1)S}$ [4], where SNR denotes the received signal-to-noise ratio (SNR). Note that the optimal DoF is achieved when $S = M$.

For the MIMO IMAC, where each user has L antennas, the user scaling condition to achieve KS DoF can be greatly reduced to $\text{SNR}^{(K-1)S-L+1}$ with the use of singular value decomposition (SVD)-based beamforming at each user, by further minimizing the generating interference level [7]. However, to implement the SVD-based OIA with local CSI and no coordination between the users or the BSs, each beamforming weight vector is computed at each user, and then information of the selected users' weight vectors should be sent to the corresponding

BS for the coherent detection. In addition, although the OIA based on the zero-forcing (ZF) receiver at the BSs is sufficient to achieve the optimal DoF, its achievable rate is in general far below the channel capacity, and the gap increases as the dimension of channel matrices grows. In this paper, we would like to answer the aforementioned two practical issues of the SVD-based OIA.

In recent cellular systems such as the 3GPP Long Term Evolution [14], each selected user should transmit an uplink pilot (known as Sounding Reference Signal in 3GPP systems) so that the corresponding BS estimates the uplink channel matrix, which is widely used for channel quality estimation, downlink signal design assuming the channel reciprocity in time division duplexing (TDD) systems, etc. The effective channel matrices rotated by the weight vectors should also be known so that the BSs perform coherent detection—the matrices can be acquired by the BSs through either of the following two methods: i) additional dedicated time/frequency pilot (known as Demodulation Reference Signal in 3GPP systems [14]), where the pilots are rotated by weight vectors [15], [16] and ii) limited feedforward of the indices of the weight vectors (as included in Downlink Control Information Format 4 [17]). For the first method, however, the system capacity can be degraded as the number of selected users increases due to the increased pilot overhead [18], [19]. For a reliable transmission, the length of pilot signaling also needs to be sufficiently long [20], [21]. Furthermore, in cellular networks, long training sequences or disjoint pilot resources for all users in each cell are required to avoid the pilot contamination coming from the inter-cell interference [22]. For these reasons, practical communication systems such as the 3GPP standard allow highly limited resources for uplink pilot. On the other hand, the second method using the limited feedforward is preferable especially for the MIMO IMAC in the sense that feedforward information can be flexibly multiplexed with uplink data requiring no additional pilot resource. Several studies [18], [19], [23] have addressed the same issues on the feedforward of the weight vectors for multiuser MIMO systems, and have proposed the design of codebook-based precoding matrices.

In the first part of this paper, we introduce a codebook-based OIA scheme, where weight vectors are chosen from a pre-defined codebook with a finite size such that information of the weight vectors of selected users is sent to the corresponding BS via limited feedforward signaling. Two widely-used codebooks, the Grassmannian and random codebooks, are used. Surprisingly, although the granularity of the Grassmannian codebook is higher than that of the random codebook for a given codebook size, our result indicates that for both codebook approaches, the codebook size, required to achieve the same user scaling condition as the SVD-based OIA case, coincides. It is also shown that the required codebook size in bits increases linearly with the number of transmit antennas and logarithmically with the received SNR, i.e., the required codebook size scales as $L \log_2 \text{SNR}$.

In the second part, we propose a receiver design at the BSs in pursuit of improving the achievable rate based on the ZF receiver. The design is challenging in the sense that local CSI and no coordination between any BSs are assumed, thus resulting in no available information of the interfering links at each BS. Thus, the maximum likelihood (ML) decoding is not possible at each BS since the covariance matrix of the effective noise

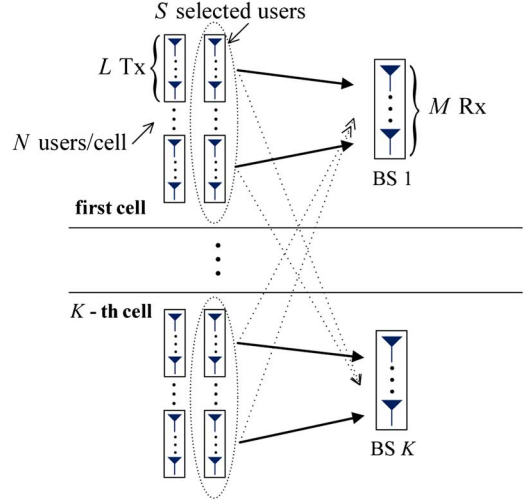


Fig. 1. The MIMO IMAC model.

cannot be estimated due to no information of the interfering links. We propose the use of a simple minimum Euclidean distance receiver, where the ML cost-function is used by assuming the identity noise covariance matrix, which does not require information of the interfering links. We show that this receiver asymptotically achieves the channel capacity as the number of users increases.

Simulation results are provided to justify the derived user and codebook size scaling conditions and to evaluate the performance of the minimum Euclidean distance receiver. A practical scenario, e.g., low SNR, small codebook size, and a small number of users, is taken into account to show the robustness of our scheme.

The remainder of this paper is organized as follows. Section II describes the system and channel models and Section III presents the proposed codebook-based OIA scheme. Section IV derives the user and codebook size scaling conditions of the proposed OIA scheme along with two different codebooks and Section V derives the asymptotic performance of the minimum Euclidean distance receiver. Section VI performs the numerical evaluation. Section VII summarizes the paper with some concluding remarks.

Notations: $\mathbb{C}$ indicates the field of complex numbers. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and the conjugate transpose, respectively. $\mathbf{I}_n$ is the $(n \times n)$ -dimensional identity matrix.

II. SYSTEM AND CHANNEL MODELS

Consider the TDD MIMO IMAC with K cells, each of which consists of a BS with M antennas and N users, each having L antennas, as depicted in Fig. 1. It is assumed that each selected user transmits a single spatial stream. In each cell, S ($S \leq M$) users are selected for uplink transmission. We assume that $L < (K - 1)S + 1$ and $M < KS$, because all inter-cell interference can be completely canceled at the transmitters or receivers otherwise [7]. Let $\mathbf{H}_k^{[i,j]} \in \mathbb{C}^{M \times L}$ denote the channel matrix from user j in the i th cell to BS k . A frequency-flat fading and the reciprocity between uplink and downlink channels are assumed. Each element of $\mathbf{H}_k^{[i,j]}$ is assumed to be an identical and independent complex Gaussian random variable with zero mean and variance $1/L$. In addition,

quasi-static frequency-flat fading is assumed, i.e., channel coefficients are constant during one transmission block and change to new independent instances for every transmission block. User j in the i th cell estimates the uplink channel of its own link, $\mathbf{H}_k^{[i,j]}$ ($k = 1, \dots, K$), via downlink pilot signaling transmitted from the BSs; that is, local CSI is utilized as in [6]. Without loss of generality, the indices of the selected users in each cell are assumed to be $(1, \dots, S)$ for notational simplicity. Then, the received signal at BS i is expressed as:

$$\mathbf{y}_i = \sum_{j=1}^S \mathbf{H}_i^{[i,j]} \mathbf{w}^{[i,j]} x^{[i,j]} + \underbrace{\sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} x^{[k,m]}}_{\text{inter-cell interference}} + \mathbf{z}_i, \quad (1)$$

where $\mathbf{w}^{[i,j]} \in \mathbb{C}^{L \times 1}$ and $x^{[i,j]}$ are the weight vector and transmit symbol with unit average power at user j in the i th cell, respectively, and $\mathbf{z}_i \in \mathbb{C}^{M \times 1}$ denotes the additive white Gaussian noise at BS i , with zero mean and the covariance $N_0 \mathbf{I}_M$.

III. PROPOSED CODEBOOK-BASED OIA: OVERALL PROCEDURE

The proposed scheme essentially follows the same procedure as that of the SVD-based MIMO OIA [7], [8] except for the weight vector design step. For the completeness of our achievability results, we briefly describe the overall procedure for all the steps.

A. Offline Procedure—Reference Basis Broadcasting

The orthogonal reference basis matrix at BS k , to which the received interference vectors are aligned, is denoted by $\mathbf{Q}_k = [\mathbf{q}_{k,1}, \dots, \mathbf{q}_{k,M-S}] \in \mathbb{C}^{M \times (M-S)}$. Here, BS k in the k th cell ($k \in \{1, \dots, K\}$) independently and randomly generates $\mathbf{q}_{k,m} \in \mathbb{C}^{M \times 1}$ ($m = 1, \dots, M-S$) from the M -dimensional sphere. BS k also finds the null space of $\mathbf{Q}_k$, defined by $\mathbf{U}_k = [\mathbf{u}_{k,1}, \dots, \mathbf{u}_{k,S}] \triangleq \text{null}(\mathbf{Q}_k)$, where $\mathbf{u}_{k,i} \in \mathbb{C}^{M \times 1}$ is orthonormal, and then broadcasts it to all users. Note that this process is required only once prior to data transmission and does not need to change with respect to channel instances.

B. Step 1—Weight Vector Design

Let us denote the codebook set consisting of N_f elements as $\mathcal{C}_f \triangleq \{\mathbf{c}_1, \dots, \mathbf{c}_{N_f}\}$, where $\mathbf{c}_1, \dots, \mathbf{c}_{N_f} \in \mathbb{C}^{L \times 1}$ are chosen from the L -dimensional unit sphere. Then, the number of bits to represent $\mathcal{C}_f$ is denoted by $n_f = \lceil \log_2 N_f \rceil$. Let $\mathbf{w}^{[i,j]}$ denote the weight vector at user j in the i th cell. Each user attempts to minimize the leakage of interference (LIF) $\eta^{[i,j]}$ defined by [4], [7]

$$\eta^{[i,j]} = \sum_{k=1, k \neq i}^K \left\| \mathbf{U}_k^H \mathbf{H}_k^{[i,j]} \mathbf{w}^{[i,j]} \right\|^2 = \left\| \mathbf{G}^{[i,j]} \mathbf{w}^{[i,j]} \right\|^2, \quad (2)$$

where $\mathbf{G}^{[i,j]}$ is the stacked interference channel matrix given by

$$\mathbf{G}^{[i,j]} \triangleq \left[\left(\mathbf{U}_1^H \mathbf{H}_1^{[i,j]} \right)^T, \dots, \left(\mathbf{U}_{i-1}^H \mathbf{H}_{i-1}^{[i,j]} \right)^T, \right]$$

$$\left(\mathbf{U}_{i+1}^H \mathbf{H}_{i+1}^{[i,j]} \right)^T, \dots, \left(\mathbf{U}_K^H \mathbf{H}_K^{[i,j]} \right)^T \right]^T. \quad (3)$$

Let us denote the SVD of $\mathbf{G}^{[i,j]}$ by

$$\mathbf{G}^{[i,j]} = \mathbf{E}^{[i,j]} \mathbf{\Sigma}^{[i,j]} \mathbf{V}^{[i,j]H}, \quad (4)$$

where $\mathbf{E}^{[i,j]} \in \mathbb{C}^{(K-1)S \times L}$ and $\mathbf{V}^{[i,j]} \in \mathbb{C}^{L \times L}$ are left- and right-singular vectors of $\mathbf{G}^{[i,j]}$, respectively, consisting of L orthonormal columns, and $\mathbf{\Sigma}^{[i,j]} = \text{diag}(\sigma_1^{[i,j]}, \dots, \sigma_L^{[i,j]})$. Here, $\sigma_m^{[i,j]}$ denotes the m th singular value of $\mathbf{G}^{[i,j]}$, where $\sigma_1^{[i,j]} \geq \dots \geq \sigma_L^{[i,j]}$. Then, $\mathbf{w}^{[i,j]}$ is chosen from $\mathcal{C}_f$ by

$$\mathbf{w}^{[i,j]} = \arg \max_{\mathbf{c} \in \mathcal{C}_f} \left| \left(\mathbf{v}_L^{[i,j]} \right)^H \mathbf{c} \right|^2. \quad (5)$$

Clearly, the weight vector minimizing $\eta^{[i,j]}$ is the L th column of $\mathbf{V}^{[i,j]}$, denoted by $\mathbf{v}_L^{[i,j]}$, and this precoding is subject to the SVD-based OIA.

C. Step 2—User Selection

Each user reports its LIF metric in (2) to the corresponding BS, and then each BS selects S users, having the LIF metrics up to the S th smallest one, amongst N users in the cell. Subsequently, each selected user forwards the index of $\mathbf{w}^{[i,j]}$ in the codebook to the belonging BS.

D. Step 3—Uplink Transmission and Detection

If all the selected users transmit the uplink signals simultaneously, then the received signal at BS i is given by (1). As in the SVD-based OIA [7], the linear ZF detection is sufficient to achieve the maximum DoF. The decision statistics $\mathbf{r}_i$ at BS i is obtained from

$$\mathbf{r}_i = [r_{i,1}, \dots, r_{i,S}]^T \triangleq \mathbf{F}_i^H \mathbf{U}_i^H \mathbf{y}_i, \quad (6)$$

where $\mathbf{F}_i \in \mathbb{C}^{S \times S}$ is the ZF equalizer defined by

$$\mathbf{F}_i = [\mathbf{f}_{i,1}, \dots, \mathbf{f}_{i,S}] \\ \triangleq \left(\left[\mathbf{U}_i^H \mathbf{H}_i^{[i,1]} \mathbf{w}^{[i,1]}, \dots, \mathbf{U}_i^H \mathbf{H}_i^{[i,S]} \mathbf{w}^{[i,S]} \right]^{-1} \right)^H.$$

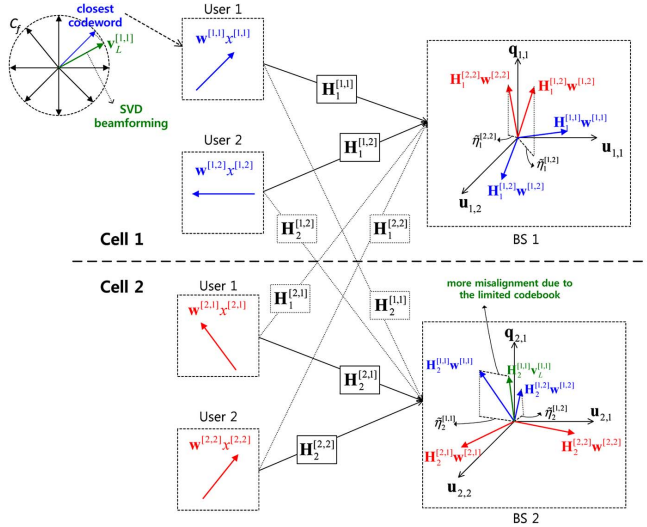
Note that multiplying $\mathbf{U}_i^H$ to $\mathbf{y}_i$ cancels interference aligned at $\mathbf{Q}_i$. The achievable rate $R^{[i,j]}$ is then given by

$$R^{[i,j]} = \log_2 \left(1 + \text{SINR}^{[i,j]} \right) = \log_2 \left(1 + \frac{\text{SNR}}{\|\mathbf{f}_{i,j}\|^2 + \tilde{I}^{[i,j]}} \right), \quad (7)$$

where $\text{SNR} = 1/N_0$ and

$$\tilde{I}^{[i,j]} \triangleq \sum_{k=1, k \neq i}^K \sum_{m=1}^S \left| \mathbf{f}_{i,j}^H \mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \right|^2 \text{SNR}.$$

Fig. 2 illustrates the principle of the proposed signaling for $K = 2$, $M = 3$, and $S = 2$. If interference $\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}$ in (1) is perfectly aligned to the interference basis $\mathbf{Q}_i$, i.e., $\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \in \text{span}(\mathbf{Q}_i)$, then interference in $\mathbf{r}_i$ of (6) vanishes because $\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} = 0$. As illustrated in Fig. 2, the value $\|\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}\|^2$ represents the amount of the signal transmitted from user m in the k th cell to BS i


 Fig. 2. The codebook-based OIA where $K = 2$, $M = 3$, $S = 2$.

that is not aligned to the interference reference basis $\mathbf{Q}_i$. This misalignment becomes higher than the SVD-based OIA case, due to the finite codebook size.

IV. ACHIEVABILITY RESULTS

It was shown in [7] that using the SVD-based OIA scheme leads to a comparatively less number of users required to achieve the maximum DoF in the MIMO IMAC model. In this section, we derive the number of feedforward bits required to achieve the same achievability as the SVD-based OIA case in terms of the DoF and user scaling condition when two different types of codebook-based OIA schemes, i.e., the Grassmannian and random codebook-based OIAs, are used.

In our analysis, we use the total DoF defined as [1]

$$\text{DoF} = \lim_{\text{SNR} \rightarrow \infty} \frac{\sum_{i=1}^K \sum_{j=1}^S R^{[i,j]}}{\log_2 \text{SNR}},$$

where $R^{[i,j]}$ is the achievable rate for user j in the i th cell and $\text{SNR} = 1/N_0$. The maximum achievable DoF under the assumption of selecting S spatial streams per cell is KM [7], which can be obtainable if $S = M$.

In addition, interference decaying rate of the codebook-based OIA is characterized and then compared with the results on DoF achievability. DoF achievability for $L \geq (K-1)S + 1$ is also analyzed.

A. DoF Achievability: Grassmannian Codebook

We start with the following three lemmas which shall be used to establish our main theorem.

Lemma 1: For any given codebook, the LIF metric $\eta^{[i,j]}$ in (2) is upper-bounded by

$$\eta^{[i,j]} \leq \sigma_L^{[i,j]^2} + d^{[i,j]^2} \sigma_1^{[i,j]^2}, \quad (8)$$

where $d^{[i,j]}$ is the residual distance defined by

$$d^{[i,j]} = \sqrt{1 - |\mathbf{w}^{[i,j]} \mathbf{H}^{[i,j]} \mathbf{v}_L^{[i,j]}|^2} \quad (9)$$

and $\sigma_m^{[i,j]}$ is the m th singular value of the stacked interference channel matrix in (3).

Proof: Since $\mathbf{v}_L^{[i,j]}$ is isotropically distributed over the L -dimensional sphere with identically and isotropically distributed (i.i.d.) complex Gaussian channel matrices [24], the weight vector $\mathbf{w}^{[i,j]}$ chosen from a codebook can be written by $\mathbf{w}^{[i,j]} = \sqrt{1 - d^{[i,j]^2}} \mathbf{v}_L^{[i,j]} + d^{[i,j]} \mathbf{t}^{[i,j]}$ [13], [25], where $0 \leq d^{[i,j]} \leq 1$ accounts for the quantization error and $\mathbf{t}^{[i,j]}$ is a unit-norm vector i.i.d. over $\text{null}(\mathbf{v}_L^{[i,j]})$. Then, $\eta^{[i,j]}$ in (2) is bounded by

$$\eta^{[i,j]} = \left\| \sqrt{1 - d^{[i,j]^2}} \mathbf{G}^{[i,j]} \mathbf{v}_L^{[i,j]} + d^{[i,j]} \mathbf{G}^{[i,j]} \mathbf{t}^{[i,j]} \right\|^2 \quad (10)$$

$$\leq (1 - d^{[i,j]^2}) \sigma_L^{[i,j]^2} + d^{[i,j]^2} \left\| \mathbf{G}^{[i,j]} \mathbf{t}^{[i,j]} \right\|^2 \quad (11)$$

$$\leq \sigma_L^{[i,j]^2} + d^{[i,j]^2} \sigma_1^{[i,j]^2}, \quad (12)$$

where (12) follows from $\left\| \mathbf{G}^{[i,j]} \mathbf{t}^{[i,j]} \right\|^2 \leq \sigma_1^{[i,j]^2}$ for any unit-norm vector $\mathbf{t}^{[i,j]}$, which proves the lemma. ■

Now, we further bound the LIF metric for the Grassmannian codebook as follows.

Lemma 2: By using the Grassmannian codebook, $\eta^{[i,j]}$ is further bounded by $\eta^{[i,j]} \leq \sigma_L^{[i,j]^2} + \nu_f \sigma_1^{[i,j]^2}$, where ν_f denotes the number of feedforward bits given by

$$\nu_f = \left(\frac{1}{N_f} \right)^{\frac{1}{(L-1)}} \quad (13)$$

and N_f is the number of elements in the codebook set.

Proof: The Grassmannian codebook $\mathcal{C}_f$ is the set of codewords chosen by the optimal sphere packing for the L -dimensional sphere; namely, the chordal distance of any two codewords is all the same, i.e., $\sqrt{1 - |\mathbf{c}_i^H \mathbf{c}_j|} = d$ for any $i \neq j$ and $d \geq 0$. Based on this property, the Rankin, Gilbert-Varshamov, and Hamming bounds on the distance of the codebook give us [25], [26]

$$d^{[i,j]^2} \leq \min \left\{ \frac{1}{2}, \frac{(L-1)N_f}{2L(N_f-1)}, \left(\frac{1}{N_f} \right)^{\frac{1}{(L-1)}} \right\}. \quad (14)$$

For large N_f , the third term of (14) becomes dominant, thus providing an arbitrarily tight bound. Inserting (14) to (8) proves the lemma. ■

From Lemma 2, we also have the following lemma.

Lemma 3: For the Grassmannian codebook, it follows that $\eta^{[i,j]} \leq \eta_{\text{GC}}^{[i,j]}$ where

$$\eta_{\text{GC}}^{[i,j]} = \begin{cases} C_1 \nu_f \sigma_1^{[i,j]^2} & \text{if } \sigma_L^{[i,j]^2} \leq (1 + \delta) \nu_f \sigma_1^{[i,j]^2} \\ C_2 \sigma_L^{[i,j]^2} & \text{otherwise,} \end{cases} \quad (15)$$

for any constant $\delta \geq 0$ independent of SNR. Here, $C_1 = (2 + \delta)$ and $C_2 = (1 + 1/(1 + \delta))$; thus, $1 \leq C_2 \leq 2$.

Now we are ready to show our first main theorem, which derives the number of feedforward bits, required to achieve the same user scaling condition as the SVD-based OIA case, for the proposed OIA with Grassmannian codebook.

Theorem 1: The codebook-based OIA with the optimal Grassmannian codebook $\mathcal{C}_f$ [24] achieves the user scaling condition of the SVD-based OIA if¹

$$n_f = \Omega(L \log_2 \text{SNR}) \text{ bits.} \quad (16)$$

Moreover, under the condition (16), KS DoF are achievable with high probability if $N = \omega(\text{SNR}^{(K-1)S-L+1})$.

Proof: Let us start from showing the following simple bound on $\text{SINR}^{[i,j]}$ in (7):

$$\text{SINR}^{[i,j]} \geq \frac{\text{SNR}/\|\mathbf{f}_{i,j}\|^2}{1 + I^{[i,j]}}, \quad (17)$$

where $I^{[i,j]} \triangleq \sum_{k=1, k \neq i}^K \sum_{m=1}^S \|\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}\|^2 \text{SNR}$. Suppose that $I^{[i,j]} \leq \epsilon$ for some constant $\epsilon > 0$ independent of the received SNR so that each user achieves 1 DoF. By this principle, we obtain a lower bound on the achievable DoF for the codebook-based OIA as

$$\text{DoF} \geq KS \cdot \mathcal{P}, \quad (18)$$

where

$$\mathcal{P} \triangleq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ I^{[i,j]} \leq \epsilon, \forall \text{ user } j \text{ and BS } i, \right. \\ \left. j \in \mathcal{S} \triangleq \{1, \dots, S\}, i \in \mathcal{K} \triangleq \{1, \dots, K\} \right\}. \quad (19)$$

¹We use the following notation: i) $f(x) = O(g(x))$ means that there exist constants M and m such that $f(x) \leq Mg(x)$ for all $x > m$. ii) $f(x) = o(g(x))$ means that $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$. iii) $f(x) = \Omega(g(x))$ if $g(x) = O(f(x))$. iv) $f(x) = \omega(g(x))$ if $g(x) = o(f(x))$. v) $f(x) = \Theta(g(x))$ if $f(x) = O(g(x))$ and $g(x) = O(f(x))$.

From the fact that the sum of received interference at all the BSs is equivalent to the sum of the LIF metrics [4], [7], i.e.,

$$\sum_{i=1}^K \sum_{k=1, k \neq i}^K \sum_{m=1}^S \left\| \mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[i,j]} \right\|^2 = \sum_{i=1}^K \sum_{j=1}^S \eta^{[i,j]}, \quad (20)$$

we can bound $\mathcal{P}$ as

$$\mathcal{P} \geq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \sum_{i=1}^K \sum_{j=1}^S I^{[i,j]} \leq \epsilon \right\} \\ \geq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \eta^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{KS^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{S} \right\}. \quad (21)$$

At this point, let us choose N_f such that $N_f^{1/(L-1)} \geq \text{SNR}^{(1+\gamma)}$ for $\gamma > 0$, i.e.,

$$n_f = \log_2 N_f \geq (1+\gamma)(L-1) \log_2 \text{SNR}, \quad (22)$$

resulting in (16). From (13) and (22), ν_f is bounded by

$$\nu_f = N_f^{-1/(L-1)} \leq \text{SNR}^{-(1+\gamma)}. \quad (23)$$

Now we consider the LIF-overestimating modification by using the upper bound $\eta_{\text{GC}}^{[i,j]}$ in Lemma 3. From $\eta^{[i,j]} \leq \eta_{\text{GC}}^{[i,j]}$ and (21), we have

$$\mathcal{P} \geq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \eta_{\text{GC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{KS^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{S} \right\}. \quad (24) \\ \geq \mathcal{P}_{\text{GC}} \triangleq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \left[\eta_{\text{GC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{KS^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{S} \right] \right. \\ \left. \& \left[\sigma_L^{[i,j]^2} \geq (1+\delta)\nu_f \sigma_1^{[i,j]^2}, \forall i \in \mathcal{K}, \right. \right. \\ \left. \left. \forall j \in \mathcal{N} \triangleq \{1, \dots, N\} \right] \right\}. \quad (25)$$

$$\mathcal{P}_{\text{GC}} = \lim_{\text{SNR} \rightarrow \infty} \underbrace{\Pr \left\{ \sigma_L^{[i,j]^2} \geq (1+\delta)\nu_f \sigma_1^{[i,j]^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{N} \right\}}_{\triangleq p_c} \\ - \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \left[\text{there exist less than } S \text{ users per cell such that } \eta_{\text{GC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{KS^2} \right] \right. \\ \left. \& \left[\sigma_L^{[i,j]^2} \geq (1+\delta)\nu_f \sigma_1^{[i,j]^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{N} \right] \right\} \quad (26)$$

$$= \lim_{\text{SNR} \rightarrow \infty} p_c - \lim_{\text{SNR} \rightarrow \infty} \sum_{m=0}^{S-1} \binom{N}{m} \left[\Pr \left\{ \eta_{\text{GC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{KS^2} \& (1+\delta)\nu_f \sigma_1^{[i,j]^2} \leq \sigma_L^{[i,j]^2} \right\} \right]^m \\ \times \left[\underbrace{\Pr \left\{ (1+\delta)\nu_f \sigma_1^{[i,j]^2} \leq \sigma_L^{[i,j]^2} \& \frac{\epsilon \text{SNR}^{-1}}{C_2 KS^2} \leq \sigma_L^{[i,j]^2} \right\}}_{\triangleq P_o} \right]^{N-m} \quad (27)$$

$$\geq \lim_{\text{SNR} \rightarrow \infty} p_c - \lim_{\text{SNR} \rightarrow \infty} \sum_{m=0}^{S-1} N^m P_o^{N-m}, \quad (28)$$

From the principle $\Pr(\mathcal{A} \cap \mathcal{B}) = \Pr(\mathcal{B}) - \Pr(\mathcal{A}^c \cap \mathcal{B})$ for sets $\mathcal{A}$ and $\mathcal{B}$, (25) can be rewritten as (26), shown at the bottom of the previous page, where (27) follows from the fact that the statistics of each user is independent of each other, and (28) follows from $\Pr\left\{\eta_{\text{GC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{K S^2} \ \& \ (1+\delta)\nu_f \sigma_1^{[i,j]^2} \leq \sigma_L^{[i,j]^2}\right\} \leq 1$ and $\binom{N}{i} = \frac{N!}{i!(N-i)!} \leq N^i$.

For the rest of the proof, we show that (28) tends to one under certain conditions. In Appendix A, we first show that for given $\gamma > 0$ and $\delta > 0$, it follows that

$$\lim_{\text{SNR} \rightarrow \infty} p_c = 1, \text{ if } N = O\left(\text{SNR}^{(1+\beta)((K-1)S-L+1)}\right) \quad (29)$$

Now we show that the second term of (28) tends to zero as the SNR increases. From the fact that $\Pr\{[A \leq B] \& [C \leq B]\} = \Pr\{[A \leq B]|A \geq C\} \Pr\{A \geq C\} + \Pr\{[C \leq B]|A < C\} \Pr\{A < C\}$ for random variables A , B , and C , the probability P_o can be written as

$$\begin{aligned} P_o &= \Pr\left\{\left((1+\delta)\nu_f \sigma_1^{[i,j]^2} \leq \sigma_L^{[i,j]^2}\right) \mid (1+\delta)\nu_f \sigma_1^{[i,j]^2} \geq \frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2}\right\} \cdot p'_o \\ &\quad + \Pr\left\{\frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2} \leq \sigma_L^{[i,j]^2} \mid (1+\delta)\nu_f \sigma_1^{[i,j]^2} < \frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2}\right\} t(1-p'_o) \end{aligned}$$

where $p'_o = \Pr\left\{(1+\delta)\nu_f \sigma_1^{[i,j]^2} \geq \frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2}\right\}$. From (23), for any given channel instance, we have

$$\begin{aligned} \lim_{\text{SNR} \rightarrow \infty} p'_o &\leq \lim_{\text{SNR} \rightarrow \infty} \Pr\left\{(1+\delta)\text{SNR}^{-(1+\gamma)} \sigma_1^{[i,j]^2} \geq \frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2}\right\} \\ &= \lim_{\text{SNR} \rightarrow \infty} \Pr\left\{\sigma_1^{[i,j]^2} \geq \frac{\epsilon \text{SNR}^\gamma}{(1+\delta)C_2 K S^2}\right\} = 0, \end{aligned}$$

which results in

$$\lim_{\text{SNR} \rightarrow \infty} P_o = \lim_{\text{SNR} \rightarrow \infty} \Pr\left\{\frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2} \leq \sigma_L^{[i,j]^2}\right\}.$$

From [27, Theorem 4], we have

$$\begin{aligned} \Pr\left\{\frac{\epsilon \text{SNR}^{-1}}{C_2 K S^2} \leq \sigma_L^{[i,j]^2}\right\} &= 1 - \alpha \left(\frac{\epsilon}{C_2 K S^2}\right)^\psi \text{SNR}^{-\psi} + o(\text{SNR}^{-\psi}), \quad (30) \end{aligned}$$

where $\psi \triangleq (K-1)S-L+1$ and $\alpha > 0$ is a constant determined by K , S , and L . Applying (30) to (28) yields

$$\begin{aligned} \mathcal{P} \geq \mathcal{P}_{\text{GC}} &\geq \lim_{\text{SNR} \rightarrow \infty} p_c \\ &\quad - \lim_{\text{SNR} \rightarrow \infty} \sum_{i=0}^{S-1} N^m \left[1 - \alpha \left(\frac{\epsilon}{C_2 K S^2}\right)^\psi \text{SNR}^{-\psi} + o(\text{SNR}^{-\psi})\right]^{N-m}. \quad (31) \end{aligned}$$

For given γ and $0 < \beta < \gamma$, let us choose $N = O(\text{SNR}^{(1+\beta)((K-1)S-L+1)})$. Then, from (29), it follows that $\lim_{\text{SNR} \rightarrow \infty} p_c = 1$. On the other hand, the second term of (31) tends to zero because N^m increases polynomially with SNR for given m while

$$\left[1 - \alpha \left(\frac{\epsilon}{C_2 K S^2}\right)^\psi \text{SNR}^{-\psi} + o(\text{SNR}^{-\psi})\right]^{N-m}$$

decreases exponentially with SNR. Thus, $\mathcal{P}$ tends to one, which means that KS DoF are achievable.

As assumed earlier, note that our analysis holds for $\beta < \gamma$. However, it is obvious that assuming either the condition $N = O(\text{SNR}^{(1+\beta)((K-1)S-L+1)})$ for any $\beta \geq \gamma$ or $N = \omega(\text{SNR}^{(K-1)S-L+1})$ leads to the same or higher DoF compared to the case for $0 \leq \beta < \gamma$, due to the fact that increasing N for given n_f and SNR values yields a reduced LIF and thus an increased achievable rate for all the selected users. Since the maximum achievable DoF are upper-bounded by KS for given S , the last argument indicates that KS DoF are achievable if $n_f = \Omega(L \log_2 \text{SNR})$ and $N = \omega(\text{SNR}^{(K-1)S-L+1})$, which completes the proof. ■

Theorem 1 indicates that n_f should scale with SNR so as to achieve the target DoF under the same user scaling condition as the SVD-based OIA case, and that from (22), no more feedforward bits than $(1+\gamma)(L-1) \log_2 \text{SNR}$ are indeed required. The derived n_f scaling condition is proportional to L and $\log_2 \text{SNR}$, which is consistent with the previous results on the number of feedback bits required to avoid performance loss due to the finite codebook size in a variety of limited feedback systems [9], [13], [24].

B. DoF Achievability: Random Codebook

For a random codebook scenario, each element $\mathbf{c}_n$ of $\mathcal{C}_f$ ($n \in \{1, \dots, N_f\}$) is chosen independently and isotropically from the L -dimensional sphere. The following second main theorem shows that the same user scaling condition as the Grassmannian codebook-based OIA case is obtained even with the random codebook-based OIA.

Theorem 2: The codebook-based OIA with a random codebook achieves KS DoF with high probability if $N = \omega(\text{SNR}^{(K-1)S-L+1})$ and $n_f = \Omega(L \log_2 \text{SNR})$ bits.

Proof: Since (17)–(21) also hold for the random codebook approach, we only show that $\mathcal{P}$ in (19) tends to one under two conditions $N = \omega(\text{SNR}^{(K-1)S-L+1})$ and $n_f = \Omega(L \log_2 \text{SNR})$. Unlike the Grassmannian codebook, the residual distance $d^{[i,j]}$ in (9) is now a random variable and thus is unbounded. Note that the cumulative density function (CDF) of the squared chordal distance between any two independent unit random vectors chosen isotropically from the L -dimensional sphere is given by $\beta(L-1, 1)$, where $\beta(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$ is the beta function [13]. Since $d^{[i,j]^2}$ for the random codebook is the minimum of N_f independent random variables with distribution $\beta(L-1, 1)$, the CDF of $d^{[i,j]^2}$ is given by

$$\Pr\{d^{[i,j]^2} \leq z\} = 1 - (1 - z^{L-1})^{N_f}. \quad (32)$$

Now, let us again consider the following modification for given channel instance:

- i) if $d^{[i,j]^2} \leq \text{SNR}^{-(1+\gamma)}$ for all i and j , then the same LIF-overestimating modification as the Grassmannian codebook case is used, where the LIF values are replaced with their upper bounds. Specifically, from Lemma 1, we shall use the following upper bound on $\eta^{[i,j]}$:

$$\eta_{\text{RC}}^{[i,j]} = \begin{cases} C_3 d^{[i,j]^2} \sigma_1^{[i,j]^2} & \text{if } \sigma_L^{[i,j]^2} \leq (1+\delta') d^{[i,j]^2} \sigma_1^{[i,j]^2} \\ C_4 \sigma_L^{[i,j]^2} & \text{otherwise,} \end{cases} \quad (33)$$

for any constant $\delta' \geq 0$ independent of SNR, where $C_3 = (2 + \delta')$ and $C_4 = (1 + 1/(1 + \delta'))$; thus, $1 \leq C_4 \leq 2$,

- ii) otherwise, i.e., if $d^{[i,j]^2} > \text{SNR}^{-(1+\gamma)}$ for any i or j , then we drop the case by assuming 0 DoF for this case.

Let us define the event $\mathcal{D}$ as

$$\mathcal{D} = \left\{ d^{[i,j]^2} \leq \text{SNR}^{-(1+\gamma)}, \right. \\ \left. \forall i \in \mathcal{K} = \{1, \dots, K\}, \forall j \in \mathcal{N} = \{1, \dots, N\} \right\}.$$

From $\eta^{[i,j]} \leq \eta_{\text{RC}}^{[i,j]}$, we have

$$\begin{aligned} \mathcal{P} &\geq \lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \eta^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{K S^2}, \right. \\ &\quad \left. \forall i \in \mathcal{K}, \forall j \in \mathcal{S} = \{1, \dots, S\} \right\} \\ &\geq \mathcal{P}_{\text{RC}} \triangleq \lim_{\text{SNR} \rightarrow \infty} \Pr \{ \mathcal{D} \} \\ &\quad \cdot \Pr \left\{ \eta_{\text{RC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{K S^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{S} | \mathcal{D} \right\}. \end{aligned}$$

From (32) and the inequality $(1-x)^y > 1-xy$ for any $0 < x < 1 < y$, we have

$$\begin{aligned} \Pr \{ \mathcal{D} \} &= \left(1 - \left(1 - \left(\text{SNR}^{-(1+\gamma)} \right)^{L-1} \right)^{N_f} \right)^{KN} \\ &> 1 - KN \left(1 - \text{SNR}^{-(1+\gamma)(L-1)} \right)^{N_f}. \end{aligned} \quad (34)$$

Let us choose N such that N scales polynomially with SNR. If N_f scales faster than $\text{SNR}^{(1+\gamma)(L-1)}$, then the second term of (34) vanishes as the SNR increases, because $(1 - \text{SNR}^{-(1+\gamma)(L-1)})^{N_f}$ decreases exponentially with SNR while N increases polynomially with SNR.

Now recall that for the Grassmannian codebook approach, $d^{[i,j]^2}$ is bounded by $d^{[i,j]^2} \leq \nu_f \leq \text{SNR}^{-(1+\gamma)}$ along with the choice of $n_f \geq (1 + \gamma)(L - 1) \log_2 \text{SNR}$, and that our achievability proof is based on the upper bound on the LIF metric in (15). If $\mathcal{D}$ holds, then the upper bound in (33) is identical to (15), and thus it is not difficult to show that if $N = O(\text{SNR}^{(1+\beta)((K-1)S-L+1)})$ for any $0 < \beta < \gamma$, then

$$\lim_{\text{SNR} \rightarrow \infty} \Pr \left\{ \eta_{\text{RC}}^{[i,j]} \leq \frac{\epsilon \text{SNR}^{-1}}{K S^2}, \forall i \in \mathcal{K}, \forall j \in \mathcal{S} | \mathcal{D} \right\} = 1, \quad (35)$$

as shown in (24)(31). From (34) and (35), choosing the two conditions $n_f = \log_2 N_f \geq (1 + \gamma)(L - 1) \log_2(\text{SNR})$, i.e., $n_f = \Omega(L \log_2(\text{SNR}))$, and $N = O(\text{SNR}^{(1+\beta)((K-1)S-L+1)})$ for any $0 < \beta < \gamma$, the probability $\mathcal{P}_{\text{RC}}$ tends to one for increasing SNR. Note that taking the limit of N polynomially increasing with SNR comes merely from the strict condition of $\mathcal{D}$. Since increasing N for given n_f lowers the LIF and thereby

increases the achievable rate for each selected user, $\mathcal{P}$ tends to one for any $N = \omega(\text{SNR}^{(K-1)S-L+1})$, which completes the proof. ■

Interestingly, Theorem 2 indicates that the required n_f for the random codebook is the same as that for the Grassmannian codebook. This is an encouraging result since analytical construction methods of the Grassmannian codebook for large n_f have been unknown, and even its numerical construction requires excessive computational complexity. We complete the achievability discussion by providing the following remarks.

Remark 1 (Random vs. Grassmannian Codebook): In the previous work on limited feedback systems, the performance analysis has focused on the average SNR or the average rate loss [28]. It has been known that the Grassmannian codebook outperforms the random codebook in the average sense. However, in our OIA framework, the focus is on the asymptotic performance for increasing SNR, and it turns out that the asymptotic behavior is the same for the two codebook approaches. In fact, our result is consistent with the previous work on limited feedback systems (see [29]), where the performance gap between two codebooks was shown to be negligible as the number of feedback bits increases.

Remark 2 (Comparison to the MIMO Broadcasting Channel): For the MIMO broadcasting channel with limited feedback, where the transmitter has L antennas employing the random codebook, it was shown in [13] that the achievable rate loss for each user, denoted by ΔR , coming from the finite size of the codebook is given by $\Delta < \log_2(1 + \text{SNR} \cdot 2^{-n_f/(L-1)})$. Thus, to achieve the maximum DoF for each user, or to make the rate loss negligible as the SNR increases, the term $\text{SNR} \cdot 2^{-n_f/(L-1)}$ should remain constant for increasing SNR. That is, n_f should scale faster than $(L - 1) \log_2 \text{SNR}$. Although the system model and signaling methodology under consideration are different from our setting, Theorems 1 and 2 are consistent with this previous result.

C. Interference Decaying Rate

In this subsection, we examine the interference decaying rate of the proposed codebook-based OIA and characterize required conditions on n_f , and compare them with the results on the DoF achievability discussed in Theorems 1 and 2. In particular, since a DoF is an asymptotic measure in the high SNR regime, the aim is to evaluate the proposed scheme with respect to other parameters such as the number of feedforward bits, n_f , or the number of users per cell, N , for finite SNR. We start with the following lemma.

Lemma 4: The averaged inverse of the scheduling metric for each selected user in the SVD-based OIA with full feedforward, where the weight vector is $\mathbf{v}_L^{[i,j]}$ and the scheduling metric is given by $\eta_{\text{SVD}}^{[i,j]} = \sigma_L^{[i,j]^2}$ [7], is bounded by

$$E \left[\frac{1}{\eta_{\text{SVD}}^{[i,j]}} \right] = E \left[\frac{1}{\sigma_L^{[i,j]^2}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right), \quad (36)$$

where $\tau = (K - 1)S + L - 1$.

Proof: Consider a worse performance case where N users are randomly divided into S subgroups with N/S users each and where one user with the minimum scheduling metric $\eta_{\text{SVD}}^{[i,j]}$ is selected in each subgroup. Compared to the original scheme where S users are selected from N users, this modified scheme yields a

performance lower bound. For this modified scheme, the lemma can be proven by following the footsteps of the proof of [30, Theorem 3]. The detailed proof is provided in Appendix B. ■

The following lemma characterizes the interference decaying rate of the codebook-based OIA with respect to N .

Lemma 5: The averaged inverse of the scheduling metric for each selected user in the codebook-based OIA is bounded by

$$E \left[\frac{1}{\eta^{[i,j]}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right) \cdot \left(1 - \rho(1 + \delta)^\tau \cdot 2^{-\frac{n_f \tau}{(L-1)}} \right)^{KN}, \quad (37)$$

for a Grassmannian codebook for $\delta \geq 0$, where $\tau = (K-1)S - L + 1$ and ρ is a constant determined by K , S , and L (See (76) in Appendix A), and by

$$E \left[\frac{1}{\eta^{[i,j]}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right) \cdot \left(1 - \rho(1 + \delta)^\tau \cdot D^{*\tau} \right)^{KN} \times \left(1 - \left(1 - D^{*L-1} \right)^{2^{n_f}} \right)^{KN} \quad (38)$$

for a random codebook with $D^* > 0$.

Proof: i) Grassmannian Codebook

From Lemma 2, we have

$$E \left[\frac{1}{\eta^{[i,j]}} \right] \geq E \left[\frac{1}{\sigma_L^{[i,j]^2} + \nu_f \sigma_1^{[i,j]^2}} \right] \quad (39)$$

$$\geq E \left[\frac{1}{C_2 \sigma_L^{[i,j]^2}} \right] \cdot p_c, \quad (40)$$

where $C_2 = 1 + 1/(1 + \delta)$ and p_c is defined in (26). Then, p_c is bounded by

$$p_c = \left(\Pr \left\{ \frac{\sigma_1^{[i,j]^2}}{\sigma_L^{[i,j]^2}} \leq (1 + \delta)^{-1} \nu_f^{-1} \right\} \right)^{KN} \quad (41)$$

$$\geq \left(1 - \rho(1 + \delta)^\tau \nu_f^\tau \right)^{KN} \quad (42)$$

$$= \left(1 - \rho(1 + \delta)^\tau \cdot 2^{-\frac{n_f \tau}{(L-1)}} \right)^{KN}, \quad (43)$$

where (42) follows from Lemma 7 in Appendix A, and (43) follows from (13). In addition, ρ is a constant determined by K , S , and L (See Lemma 7).

Now, from Lemma 4, we have $E \left[\frac{1}{C_2 \sigma_L^{[i,j]^2}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right)$. Inserting this result and (43) into (40) gives us (37).

ii) Random Codebook

For $D^* > 0$, define the event $\mathcal{F}$ such that

$$\mathcal{F} = \left\{ d^{[i,j]^2} \leq D^*, \forall i \in \mathcal{K}, \forall j \in \mathcal{N} \right\}. \quad (44)$$

From Lemma 1, we have

$$E \left[\frac{1}{\eta^{[i,j]}} \right] \geq E \left[\frac{1}{\sigma_L^{[i,j]^2} + d^{[i,j]^2} \sigma_1^{[i,j]^2}} \right] \quad (45)$$

$$\geq E \left[\frac{1}{\sigma_L^{[i,j]^2} + D^* \sigma_1^{[i,j]^2}} \right] \cdot \Pr\{\mathcal{F}\} \quad (46)$$

From (32), $\Pr\{\mathcal{F}\}$ is given by

$$\Pr\{\mathcal{F}\} = \left(1 - \left(1 - D^{*L-1} \right)^{N_f} \right)^{KN}. \quad (47)$$

In addition, following the same derivations from (39) to (43) with the replacement $\nu_f \rightarrow D^*$, we can readily show that

$$E \left[\frac{1}{\sigma_L^{[i,j]^2} + D^* \sigma_1^{[i,j]^2}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right) \cdot \left(1 - \rho(1 + \delta)^\tau \cdot D^{*\tau} \right)^{KN}. \quad (48)$$

Inserting (47) and (48) into (46) yields (38). ■

From Lemmas 4 and 5, the following theorem establishes the achievability of the interference decaying rate of the SVD-based OIA.

Theorem 3: The codebook-based OIA achieves the interference decaying rate of the SVD-based OIA with full feedforward, i.e.,

$$E \left[\frac{1}{\eta^{[i,j]}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right), \quad (49)$$

if

$$n_f = \omega(L \log_2 N). \quad (50)$$

Proof: i) Grassmannian Codebook

In (37), the term $\left(1 - \rho(1 + \delta)^\tau \cdot 2^{-\frac{n_f \tau}{(L-1)}} \right)^{KN}$ converges to 1 with respect to N if and only if

$$\lim_{N \rightarrow \infty} \frac{KN}{2^{\frac{n_f \tau}{(L-1)}}} = 0, \quad (51)$$

and vanishes to 0 otherwise. Therefore, choosing $n_f = \omega(L \log_2 N)$, we have $E \left[\frac{1}{\eta^{[i,j]}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right)$.

ii) Random Codebook

In (38), the term $\left(1 - \rho(1 + \delta)^\tau \cdot D^{*\tau} \right)^{KN}$ converges to 1 with respect to N if and only if

$$\lim_{N \rightarrow \infty} \frac{KN}{D^{*- \tau}} = 0, \quad (52)$$

which yields the condition $D^{*-1} = \omega(N^{1/\tau})$. In addition, the term $\left(1 - \left(1 - D^{*L-1} \right)^{N_f} \right)^{KN}$ in (38) can be bounded by

$$\left(1 - \left(1 - D^{*L-1} \right)^{N_f} \right)^{KN} \geq 1 - KN \left(1 - D^{*L-1} \right)^{N_f} \quad (53)$$

for $0 < \left(1 - D^{*L-1} \right)^{N_f} < 1 < KN$ and $KN \left(1 - D^{*L-1} \right)^{N_f} < 1$. The right-hand side of (53) converges to 1 with respect to N_f only with the choice

$$N_f = \omega \left(D^{*-(L-1)} \right). \quad (54)$$

Now, inserting $D^{*-1} = \omega(N^{1/\tau})$ into (54) yields $N_f = \omega(N^{(L-1)/\tau})$, and thus if $n_f = \omega(L \log_2 N)$, then $E \left[\frac{1}{\eta^{[i,j]}} \right] \geq \Theta \left(N^{\frac{1}{\tau}} \right)$. ■

Theorem 3 shows that n_f should increase with respect to $\log_2 N$ for given L to attain the SVD-based OIA in terms of the interference decaying rate with respect to N . On the other hand, from Theorems 1 and 2, n_f needs to increase with respect

to $\log_2$ SNR for given L to achieve the achievable rate performance of the SVD-based OIA with respect to SNR; thus, these two observations are consistent with each other. Furthermore, the exponent τ of SNR in the user scaling condition given by $N = \omega(\text{SNR}^\tau)$ is related to the exponent τ of N in the interference decaying rate given by $\Theta(N^{1/\tau})$. That is, the larger τ , the lower user scaling condition to achieve KS DoF and the faster interference decaying rate.

The following remarks provide further intuition of the results in Theorem 3.

Remark 3: As N increases, the CDF of $\eta^{[i,j]}$ converges to a unit step function, i.e., a degenerate distribution having the value around zero. For a degenerate distribution, $E[\eta^{[i,j]}] = 1/E\left[\frac{1}{\eta^{[i,j]}}\right]$ by Jensen's inequality. In the large N regime, with the choice of (50), we have

$$E\left[\eta^{[i,j]}\right] \approx 1/E\left[\frac{1}{\eta^{[i,j]}}\right] \leq \Theta\left(N^{-\frac{1}{\tau}}\right), \quad (55)$$

which follows from Theorem 3.

Remark 4: From the proofs of Lemma 4 and Theorems 1 and 2, both the user scaling condition to achieve KS DoF and interference decaying rate are determined by the tail distribution of the CDF of a scheduling metric. Specifically, it is not difficult to show that if the scheduling metric of any OIA scheme (or its upper bound) has a tail CDF given by $\Theta(x^{\tilde{\tau}})$ for $\tilde{\tau} > 0$, then the user scaling condition and interference decaying rate are given by $N = \omega(\text{SNR}^{\tilde{\tau}})$ and $E[1/\eta^{[i,j]}] = \Theta(N^{1/\tilde{\tau}})$, respectively; that is, there exists the one-to-one mapping relationship between a user scaling condition and interference decaying rate.

Now the following theorem discusses the convergence of the codebook-based OIA to the SVD-based OIA in terms of the interference decaying rate for increasing n_f for fixed N .

Theorem 4: For fixed N and SNR, the distortion in the scheduling metric of each selected user in the codebook-based OIA, compared to the SVD-based OIA, is bounded by

$$E\left[\eta^{[i,j]} - \eta_{\text{SVD}}^{[i,j]}\right] \leq \Theta\left(2^{-\frac{n_f}{L-1}}\right), \quad (56)$$

where $\eta_{\text{SVD}}^{[i,j]} = \sigma_L^{[i,j]^2}$ denotes the scheduling metric of the SVD-based OIA scheme [7] (see also Lemma 4).

Proof: The scheduling metric is bounded by $\eta^{[i,j]} \leq \sigma_L^{[i,j]^2} + d^{[i,j]^2} \sigma_1^{[i,j]^2}$ from Lemma 1. For a Grassmannian codebook, the proof is immediate from Lemma 2. For a random codebook, the scheduling metric is bounded by

$$E\left[\eta^{[i,j]}\right] = E\left[\sigma_L^{[i,j]^2}\right] + E\left[d^{[i,j]^2}\right] E\left[\sigma_1^{[i,j]^2}\right] \quad (57)$$

$$\leq E\left[\sigma_L^{[i,j]^2}\right] + \Theta\left(2^{-\frac{n_f}{L-1}}\right), \quad (58)$$

where (57) follows from the fact that $d^{[i,j]^2}$ is independent of $\sigma_1^{[i,j]^2}$, and (58) follows from $E[d^{[i,j]^2}] < 2^{-\frac{n_f}{L-1}}$ [13, Lemma 1].

Therefore, $E[\eta^{[i,j]} - \eta_{\text{SVD}}^{[i,j]}]$ is bounded by $\Theta\left(2^{-\frac{n_f}{L-1}}\right)$ for both the codebooks. ■

We conclude the discussion on the interference decaying rate with respect to n_f by providing the following remark.

Authorized licensed use limited to: AJOU UNIVERSITY. Downloaded on March 07, 2025 at 01:38:14 UTC from IEEE Xplore. Restrictions apply.

Remark 5: Recall that the sum of LIF scheduling metrics is the same as the sum of received interference (see (20)). Hence, $E[\eta_{\text{SVD}}^{[i,j]} - \eta^{[i,j]}]$ measures the additional interference level caused by the codebook-based OIA, compared with the SVD-based OIA, due to limited feedforward. Theorem 4 shows that this additional interference vanishes as n_f increases at the rate of $2^{-\frac{n_f}{L-1}}$ for both the codebooks. Therefore, the decaying rate of this additional interference with respect to n_f is dependent only on the dimension of the codebook, L .

V. ASYMPTOTICALLY OPTIMAL RECEIVER DESIGN AT THE BSS

While using the ZF receiver is sufficient to achieve the maximum DoF, we study the design of an enhanced receiver at the BSs in pursuit of improving the achievable rate. Recall that each BS is not assumed to have CSI of the cross-links from the users in the other cells, because no coordination between BSs is assumed. In this section, the main challenge is thus to decode the desired symbols with no CSI of the cross-links at the receivers. For convenience, let us rewrite the received signal at BS i in (1) as

$$\mathbf{y}_i = \tilde{\mathbf{H}}_i^{(c)} \mathbf{x}_i + \sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} x^{[k,m]} + \mathbf{z}_i, \quad (59)$$

where $\tilde{\mathbf{H}}_i^{(c)} \triangleq [\mathbf{H}_i^{[i,1]} \mathbf{w}^{[i,1]}, \dots, \mathbf{H}_i^{[i,S]} \mathbf{w}^{[i,S]}] \in \mathbb{C}^{M \times S}$ and $\mathbf{x}_i \triangleq [x^{[i,1]}, \dots, x^{[i,S]}]^T \in \mathbb{C}^{S \times 1}$. The channel capacity I_C is given by [31]

$$I_C = \log_2 \det \left(\mathbf{R}_c^{-1/2} \tilde{\mathbf{H}}_i^{(c)} \left(\tilde{\mathbf{H}}_i^{(c)} \right)^H \mathbf{R}_c^{-1/2} + \mathbf{I}_M \right), \quad (60)$$

where

$$\mathbf{R}_c = \sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \left(\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \right)^H + N_0 \mathbf{I}_M, \quad (61)$$

which is not available at BS i due to the assumption of unknown inter-cell interfering links. The channel capacity I_C is achievable with the optimal ML decoder

$$\hat{\mathbf{x}}_i^{\text{ML}} = \arg \min_{\mathbf{x}} \left(\mathbf{y}_i - \tilde{\mathbf{H}}_i^{(c)} \mathbf{x} \right)^H \mathbf{R}_c^{-1} \left(\mathbf{y}_i - \tilde{\mathbf{H}}_i^{(c)} \mathbf{x} \right), \quad (62)$$

which is infeasible to implement due to unknown $\mathbf{R}_c$. After nulling interference by multiplying $\mathbf{U}_i$, the received signal is given by

$$\begin{aligned} \tilde{\mathbf{y}}_i &= \mathbf{U}_i^H \mathbf{y}_i \\ &= \tilde{\mathbf{H}}_i \mathbf{x}_i + \underbrace{\sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} x^{[k,m]} + \mathbf{U}_i^H \mathbf{z}_i}_{\tilde{\mathbf{z}}_i}, \end{aligned} \quad (63)$$

where $\tilde{\mathbf{H}}_i \triangleq [\mathbf{U}_i^H \mathbf{H}_i^{[i,1]} \mathbf{w}^{[i,1]}, \dots, \mathbf{U}_i^H \mathbf{H}_i^{[i,S]} \mathbf{w}^{[i,S]}] \in \mathbb{C}^{S \times S}$ and $\tilde{\mathbf{z}}_i \in \mathbb{C}^{S \times 1}$ represents the effective noise. Let us denote

the covariance matrix of the effective noise after interference nulling by

$$\begin{aligned} \mathbf{R} &\triangleq E \{ \tilde{\mathbf{z}}_i \tilde{\mathbf{z}}_i^H \} \\ &= \sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \left(\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \right)^H \\ &\quad + N_0 \mathbf{I}_S. \end{aligned} \quad (64)$$

Then, the ML decoder for the modified channel (63) becomes $\arg \min_{\mathbf{x}} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x})^H \mathbf{R}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x})$, which is also infeasible to implement since the term $\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}$ ($k \in \{1, \dots, i-1, i+1, \dots, K\}, m \in \{1, \dots, S\}$) is not available at BS i .

As an alternative approach, we now introduce the following minimum Euclidean distance receiver after interference nulling at BS i :

$$\hat{\mathbf{x}}_i = \arg_{\mathbf{x}} \min \|\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}\|. \quad (65)$$

It is worth noting that the receiver in (65) is not universally optimal since $\mathbf{R}$ is not an identity matrix for given channel instance. Now, we show the achievable rate based on the use of the receiver in (65). The maximum achievable rate of any sub-optimal receiver, referred to as *mismatch capacity* [32], [33], is lower-bounded by the generalized mutual information, defined as [32], [33]

$$I_{\text{GMI}} = \sup_{\theta \geq 0} I(\theta), \quad (66)$$

where

$$I(\theta) \triangleq E \left[\log_2 \frac{Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)^\theta}{E \left[Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)^\theta | \tilde{\mathbf{y}}_i, \tilde{\mathbf{H}}_i \right]} \right] \tilde{\mathbf{H}}_i \quad (67)$$

and $Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)$ is the decoding metric expressed in probability. The following lemma characterizes the GMI of the decoder with mismatched noise covariance matrix.

Lemma 6: Consider the modified channel (63) and the decoding metric with mismatched noise covariance matrix $\hat{\mathbf{R}}$, given by

$$Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i) = \exp \left(-(\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i)^H \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i) \right). \quad (68)$$

Then, the GMI I_{GMI} based on the metric (68) is given by (66), where $I(\theta)$ for given $\hat{\mathbf{R}}$ is expressed as

$$\begin{aligned} I(\theta) &= -\frac{\theta}{\log 2} \text{tr}(\hat{\mathbf{R}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}^{-1/2}) \\ &\quad + \frac{\theta}{\log 2} \left[\text{tr} \left(\mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \left(\tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{R} \right) \right) \right] + \log_2 \det(\mathbf{\Omega}), \end{aligned} \quad (69)$$

$\mathbf{R}$ is given by (64), and $\mathbf{\Omega} \triangleq \theta \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{I}_S$. ■

Proof: See Appendix C. ■

We remark that if $\mathbf{R} = \hat{\mathbf{R}}$, then it is obvious to show $I_{\text{GMI}} = I(\theta = 1)$. In this case, using (69), $I(\theta = 1)$ can be simplified to $I(\theta = 1) = \log_2 \det(\mathbf{R}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{I}_S)$, which is equal to the channel capacity. The following theorem characterizes the

achievable rate of the proposed minimum Euclidean distance decoder.

Theorem 5 (Asymptotic Capacity): The GMI I_{GMI} of the codebook-based OIA using the minimum Euclidean distance receiver in (65) is given by

$$\begin{aligned} I_{\text{GMI}} &= \sup_{\theta \geq 0} -\frac{\theta}{\log 2} \text{tr} \left(N_0^{-1} \mathbf{R} \right) \\ &\quad + \frac{\theta}{\log 2} \left[\text{tr} \left(N_0^{-1} \mathbf{\Omega}'^{-1} \left(\tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{R} \right) \right) \right] + \log_2 \det(\mathbf{\Omega}'), \end{aligned} \quad (70)$$

which asymptotically achieves the channel capacity I_C if $N = \omega(\text{SNR}^{(K-1)S-L+1})$ and $n_f = \Omega(L \log_2(\text{SNR}))$, where $\mathbf{\Omega}' = \theta N_0^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{I}_S$.

Proof: The decoder in (65) is equivalent to the one that utilizes the decoding metric $Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)$ in (68) with $\hat{\mathbf{R}} = N_0 \mathbf{I}_S$. From Lemma 6, the GMI based the minimum Euclidean distance decoder is thus given by (70).

Now, we show that as N increases, I_{GMI} approaches I_C with increasing SNR. Recall that the user selection and weight vector design are performed such that interference is aligned to the reference basis matrix $\mathbf{Q}_i \in \mathbb{C}^{M \times (M-S)}$ as much as possible. From Theorem 1 and 2, if $N = \omega(\text{SNR}^{(K-1)S-L+1})$ and $n_f = \Omega(L \log_2(\text{SNR}))$, then the sum of received interference aligned to $\mathbf{U}_i$ can be made arbitrarily small with high probability, thereby resulting in

$$\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} \in \text{span}(\mathbf{Q}_i) \quad (71)$$

from the fact that the interference term of (59) is given by $\sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} x^{[k,m]}$. Since $\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} = 0$, the interference term of (63) is canceled out, and thus it follows that $\mathbf{R} \rightarrow \hat{\mathbf{R}} = N_0 \mathbf{I}_S$. In this case, we have

$$I_{\text{GMI}} \rightarrow I_{\text{GMI}}^* = \log_2 \det \left(N_0^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{I}_S \right). \quad (72)$$

Now we prove that I_{GMI}^* in (72) asymptotically achieves I_C . Since $[\mathbf{U}_i, \mathbf{Q}_i] \in \mathbb{C}^{M \times M}$ is an orthogonal matrix, I_{GMI}^* can be rewritten as

$$\begin{aligned} I_{\text{GMI}}^* &= \log_2 \det \left(N_0^{-1} \mathbf{U}_i^H \tilde{\mathbf{H}}_i^{(c)} \left(\tilde{\mathbf{H}}_i^{(c)} \right)^H \mathbf{U}_i + \mathbf{I}_S \right), \\ &= \log_2 \det \left(\begin{bmatrix} N_0^{-1} \mathbf{U}_i^H \tilde{\mathbf{H}}_i^{(c)} \left(\tilde{\mathbf{H}}_i^{(c)} \right)^H \mathbf{U}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \mathbf{I}_M \right) \\ &= \log_2 \det \left([\mathbf{U}_i, \mathbf{Q}_i] \cdot \begin{bmatrix} N_0^{-1} \mathbf{U}_i^H \tilde{\mathbf{H}}_i^{(c)} \left(\tilde{\mathbf{H}}_i^{(c)} \right)^H \mathbf{U}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right. \\ &\quad \left. \cdot [\mathbf{U}_i, \mathbf{Q}_i]^H + \mathbf{I}_M \right) \\ &= \log_2 \det \left(\mathbf{D} \tilde{\mathbf{H}}_i^{(c)} \left(\tilde{\mathbf{H}}_i^{(c)} \right)^H \mathbf{D} + \mathbf{I}_M \right), \end{aligned} \quad (73)$$

where

$$\mathbf{D} = [\mathbf{U}_i, \mathbf{Q}_i] \begin{bmatrix} N_0^{-1/2} \mathbf{I}_S & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{U}_i, \mathbf{Q}_i]^H = N_0^{-1/2} \mathbf{U}_i \mathbf{U}_i^H.$$

From the fact that from (71), $\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}$ can be represented as a linear combination of $\mathbf{q}_{i,1}, \dots, \mathbf{q}_{i,M-S}$, $\mathbf{R}_c$ in (61) is written as

$$\begin{aligned} \mathbf{R}_c &= \sum_{k=1, k \neq i}^K \sum_{m=1}^S \sum_{l=1}^{M-S} \sum_{t=1}^{M-S} \tilde{\xi}_{l,t}^{[k,m]} \mathbf{q}_{i,l} \mathbf{q}_{i,t}^H + N_0 \mathbf{I}_M \\ &= \sum_{l=1}^{M-S} \sum_{t=1}^{M-S} \xi_{l,t} \mathbf{q}_{i,l} \mathbf{q}_{i,t}^H + N_0 \mathbf{I}_M \\ &= \sum_{l=1}^{M-S} \sum_{t=1}^{M-S} \xi_{l,t} \mathbf{q}_{i,l} \mathbf{q}_{i,t}^H + N_0 [\mathbf{Q}_i, \mathbf{U}_i] \cdot [\mathbf{Q}_i, \mathbf{U}_i]^H \\ &= [\mathbf{Q}_i, \mathbf{U}_i] \cdot \begin{bmatrix} \Xi & \mathbf{0} \\ \mathbf{0} & N_0 \mathbf{I}_S \end{bmatrix} \cdot [\mathbf{Q}_i, \mathbf{U}_i]^H, \end{aligned}$$

where $\tilde{\xi}_{l,t}^{[k,m]}$ denotes the component coefficient of $\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}$ on $\mathbf{q}_{i,l} \mathbf{q}_{i,t}^H$, $\xi_{l,t} = \sum_{k=1, k \neq i}^K \sum_{m=1}^S \tilde{\xi}_{l,t}^{[k,m]}$, and

$$\Xi = \begin{bmatrix} \xi_{1,1} + N_0 & \xi_{1,2} & \cdots & \xi_{1,M-S} \\ \xi_{2,1} & \xi_{2,2} + N_0 & \cdots & \xi_{2,M-S} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{M-S,1} & \xi_{M-S,2} & \cdots & \xi_{M-S,M-S} + N_0 \end{bmatrix}.$$

Since each coefficient $\xi_{l,t}$ is chosen from a continuous distribution, Ξ has full rank almost surely, and thus is invertible. Therefore, we get

$$\mathbf{R}_c^{-1/2} = [\mathbf{Q}_i, \mathbf{U}_i] \begin{bmatrix} \Xi^{-1/2} & \mathbf{0} \\ \mathbf{0} & N_0^{-1/2} \mathbf{I}_S \end{bmatrix} [\mathbf{Q}_i, \mathbf{U}_i]^H.$$

For given channel instance, as N_0 decreases, $\mathbf{R}_c^{-1/2}$ becomes

$$\mathbf{R}_c^{-1/2} \rightarrow N_0^{-1/2} \mathbf{U}_i \mathbf{U}_i^H. \quad (74)$$

Applying the asymptotic $\mathbf{R}_c^{-1/2}$ of (74) to (60) finally yields I_C , which is equal to I_{GMI}^* of (73). This completes the proof. ■

As shown in Theorem 5, the minimum Euclidean distance receiver asymptotically achieves the channel capacity even without any coordination between the BSs or users. However, it is worth noting that if the interference alignment level is too low due to small N or n_f to satisfy the conditions in Theorems 1 and 2, then the achievable rate in (70) may be lower than that based on the ZF receiver. Thus, in small N and n_f regimes, there may exist crossovers, where the achievable rate of the two schemes is switched, which will be shown in Section VI via numerical evaluation. We conclude our discussion on the receiver design with the following remark.

Remark 6 (DoF Achievability of the Optimal Receiver): Even with the use of the ML receiver in (62) based on full knowledge of $\mathbf{R}_c$, the user and n_f scaling conditions to achieve KS DoF are the same as those based of the ZF receiver case, which make the amount of interference bounded even for increasing SNR.

VI. NUMERICAL RESULTS

In this section, we run computer simulations to verify the performance of the proposed two types of codebook-based OIA schemes, i.e., the Grassmannian and random codebook-based OIAs, for finite system parameters SNR, N , and n_f . For comparison, the max-SNR scheme was used, in which each user employs eigen-beamforming in terms of maximizing its

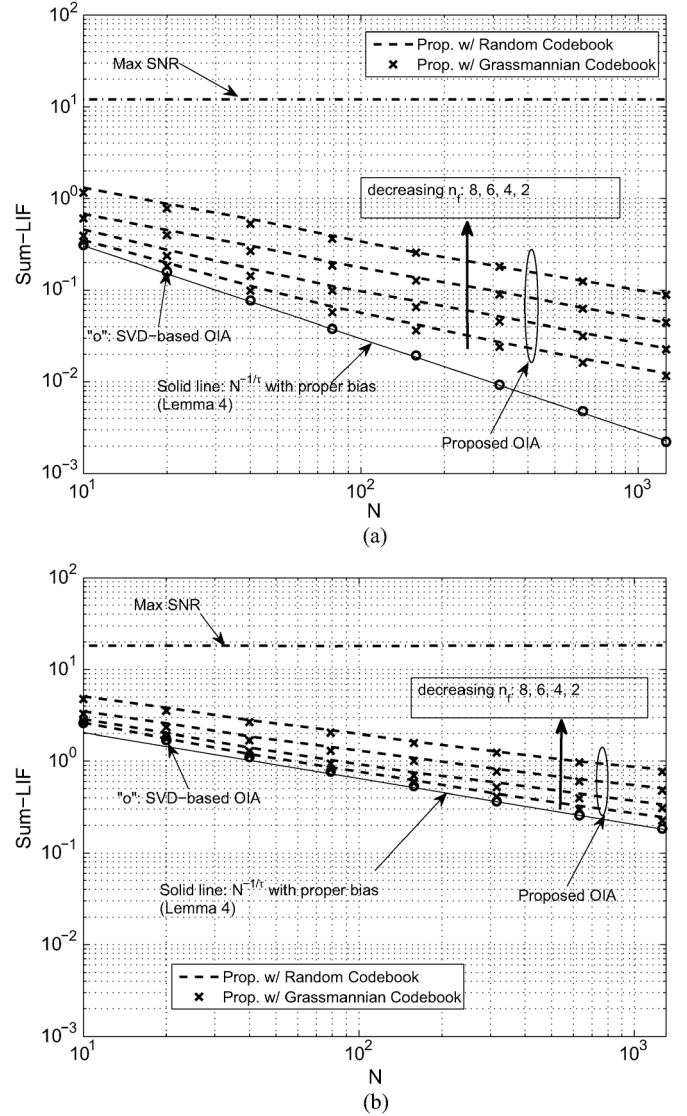


Fig. 3. The sum-LIF versus N when $K = 2$, $M = 3$, and $L = 2$. (a) $S = 2$. (b) $S = 3$.

received SNR and the belonging BS selects the S users who have the SNR values up to the S th largest one. The SVD-based OIA scheme in [7] is also compared as a baseline method. Throughout the simulations, each channel coefficient was chosen independently from a complex Gaussian distribution with zero mean and variance $1/L$. In addition, simulation results were averaged over sufficient channel realizations.

Figs. 3 show a log-log plot of the sum-LIF (or equivalently, the sum of generating interference) versus N for the MIMO IMAC with $K = 2$, $M = 3$, $L = 2$, and (a) $S = 2$ or (b) $S = 3$. From Theorems 1 and 2, the system parameter S governs the trade-off between the achievable DoF and the number of users required to guarantee such DoF [4], [7].² It is seen that the sum-LIF increases as S grows for any given scheme. However, as addressed in [4], [7], note that a smaller LIF does not necessarily leads to a higher achievable rate, especially in the

²While the sum-LIF with $S = 1$ is lowest compared to the cases with $S = 2$ and $S = 3$, the case with $S = 1$ provides the smallest achievable DoF. For more discussions about optimizing S , we refer to [7].

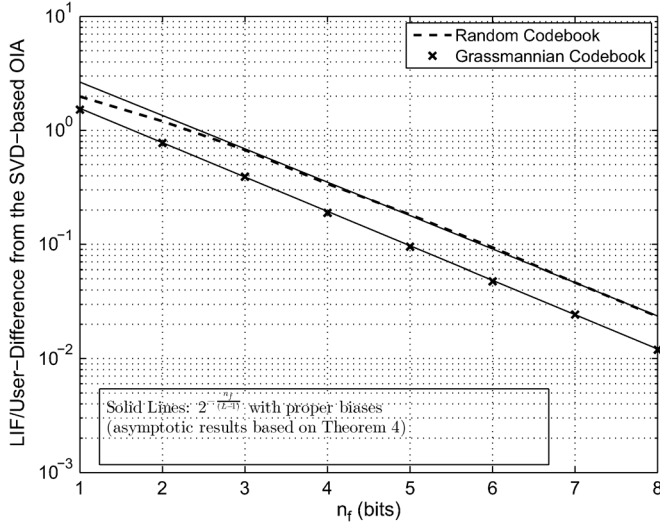


Fig. 4. (LIF/user of the codebook-based OIA—LIF/user of the SVD-based OIA) versus n_f , when $K = 2$, $M = 3$, $L = 2$, $S = 2$, and $N = 100$.

high SNR regime. In addition, the trade-off between n_f and the sum-LIF is clearly seen from Fig. 3. In addition, the figures show that the interference decaying rate (i.e., the slope of the sum-LIF curve) of the SVD-based OIA with full feedforward, $E[\eta_{\text{SVD}}^{[i,j]}]$, approaches to $\Theta(N^{-1/\tau})$ (solid lines), thus yielding the bound in Lemma 4 is tight. Although the sum-LIF level of the codebook-based OIA scheme decreases as n_f increases, it is seen that its decaying rate of the sum-LIF with respect to N cannot attain that of the SVD-based OIA unless n_f increases according to increasing N as shown in Theorem 3.

Fig. 4 shows $E[\eta_{\text{SVD}}^{[i,j]} - \eta^{[i,j]}]$ versus n_f for the two codebooks with the same configuration, where $\eta_{\text{SVD}}^{[i,j]}$ is the scheduling metric of the SVD-based OIA. Recall that $E[\eta_{\text{SVD}}^{[i,j]} - \eta^{[i,j]}]$ measures the additional interference of the codebook-based OIA due to limited feedforward, compared with the SVD-based OIA (see Remark 5). It is seen that the interference decaying rates for both the codebooks are the same as $2^{-\frac{n_f}{L-1}}$ for fixed SNR and N , which is consistent with the result in Theorem 4.

From the observations in Figs. 3 and 4 and Remark 4, we can conjecture that sufficient user scaling conditions found in Section IV also yield necessary conditions for given achievable schemes.

Figs. 5 depicts the achievable rate versus SNR when $K = 2$, $M = 3$, $L = 2$, $S = 2$, and (a) $N = 20$ or (b) $N = 100$. We consider the following four receiver structures for the proposed codebook-based OIA:

- Scheme 1: ZF receiver with the Grassmannian codebook (dashed line)
- Scheme 2: ZF receiver with the random codebook (x)
- Scheme 3: minimum Euclidean distance receiver with the Grassmannian codebook (solid line)
- Scheme 4: minimum Euclidean distance receiver with the random codebook (o)

A relationship between the sum-rate for given N and the number of feedforward bits, n_f , is observed. It is first seen that as $n_f = 8$, the proposed codebook-based OIA schemes closely obtain the achievable rate of the SVD-based OIA. It is also

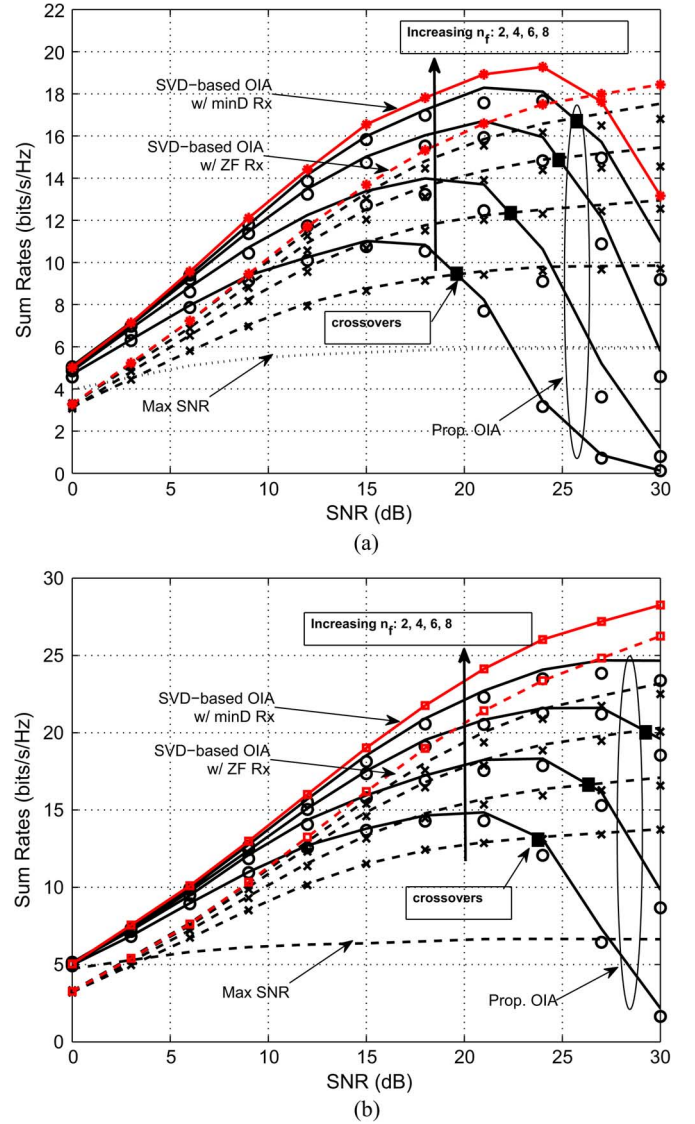


Fig. 5. The achievable rate versus SNR when $K = 2$, $M = 3$, $L = 2$, and $S = 2$. (a) $N = 20$. (b) $N = 100$.

seen that the gain coming from the Grassmannian codebook over the random codebook is marginal. From Theorem 5, we remark that the achievable rate based on the minimum Euclidean distance receiver asymptotically achieves the channel capacity if interference $\mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]}$ is perfectly aligned to $\mathbf{Q}_i$ at BS i ; that is, the covariance matrix of interference in (64), $\sum_{k=1, k \neq i}^K \sum_{m=1}^S \mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]} (\mathbf{U}_i^H \mathbf{H}_i^{[k,m]} \mathbf{w}^{[k,m]})^H$, becomes negligible compared to $N_0 \mathbf{I}_S$ due to the fact that interference is sufficiently aligned for large N . In addition, it is observed that in the low to mid SNR regimes, using the minimum Euclidean distance receiver leads to a higher sum rate than the ZF receiver case even for practical N . However, as the SNR increases beyond a certain point, i.e., in the high SNR regime, the covariance matrix of interference becomes dominant, thus yielding a performance degradation of the minimum Euclidean distance receiver. On the other hand, since the ZF receiver has no such limitation, its achievable rate increases with SNR. In consequence, for given N , there exist crossovers, where the achievable rate of the two schemes is

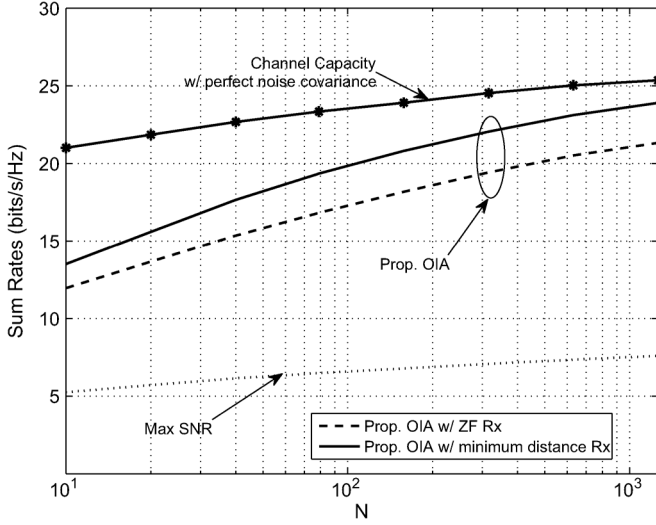


Fig. 6. The achievable rate versus N when $K = 2$, $M = 3$, $L = 2$, $S = 2$, $\text{SNR} = 20$ dB, and $n_f = 6$. The random codebook is used.

switched. It is furthermore seen that when N increases, these crossovers appear at higher SNRs, because our system is less affected by the covariance matrix of interference owing to a better interference alignment.

Fig. 6 shows a log-linear plot of the achievable rate versus N when $K = 2$, $M = 3$, $L = 2$, $S = 2$, $\text{SNR} = 20$ dB, and $n_f = 6$ when the random codebook is used. As shown in Theorem 5, it is seen that the GMI of the codebook-based OIA using the minimum Euclidean distance receiver asymptotically achieves the channel capacity as N increases. On the other hand, the achievable rate of the codebook-based OIA using the ZF receiver exhibits a constant gap even in large N regime, compared to that of the minimum Euclidean distance receiver. This observation is consistent with previous results on the single-user MIMO channel, showing that there exists a constant SNR gap between the channel capacity and the achievable rate based on the ZF receiver in the high SNR regime.

VII. CONCLUSION

For the MIMO IMAC, we have proposed two different types of codebook-based OIA methods and analyzed the codebook size required to achieve the same user scaling condition and DoF as the SVD-based OIA case. We have shown that the required codebook size scaling is the same for both of the random and Grassmannian codebooks. In addition, we have shown that the simple minimum Euclidean distance receiver operating even with no CSI of inter-cell interfering links achieves the channel capacity as N increases. Numerical examples have shown that it suffices for finite n_f to almost obtain the achievable rate of the SVD-based OIA, e.g., the case where $n_f = 8$ and $L = 2$, and that the minimum Euclidean distance receiver enhances the achievable rate based on the ZF receiver especially in the low to mid SNR regimes.

To consider an OIA framework for the MIMO interference broadcast channel (IBC), a careful precoding matrix and receiver design are required. In fact, a few studies have shown that the user scaling condition to achieve KS DoF is SNR^{KM-L} for the K -cell MIMO IBC with full feedforward [30], [34]–[36],

which is higher than $\text{SNR}^{(K-1)S-L+1}$ for the IMAC case. In [37], we shall show that the user scaling condition can indeed be lowered for the MIMO IBC to $\text{SNR}^{(K-1)S-L+1}$ with $n_f = \omega(L \log_2 \text{SNR})$, thus proving the uplink-downlink duality.

APPENDIX A PROOF OF (29)

By using $\nu_f \leq \text{SNR}^{-(1+\gamma)}$ in (23), p_c defined in (26) can be lower-bounded by

$$p_c \geq \Pr \left\{ \sigma_L^{[i,j]^2} \geq (1+\delta) \sigma_1^{[i,j]^2} \text{SNR}^{-(1+\gamma)}, \forall i \in \mathcal{K}, j \in \mathcal{N} \right\} \\ = \left(\Pr \left\{ \frac{\sigma_1^{[i,j]^2}}{\sigma_L^{[i,j]^2}} \leq (1+\delta)^{-1} \text{SNR}^{1+\gamma} \right\} \right)^{KN}, \quad (75)$$

where (75) comes from the fact that $\mathbf{G}^{[i,j]}$ is independent for different i or j , and thus their singular values are independent for different users. Note that $\frac{\sigma_1^{[i,j]^2}}{\sigma_L^{[i,j]^2}}$ is the condition number of $\mathbf{G}^{[i,j]^H} \mathbf{G}^{[i,j]}$. At this point, we introduce the following lemma on the CDF of the condition number.

Lemma 7: The CDF of $\frac{\sigma_1^{[i,j]^2}}{\sigma_L^{[i,j]^2}}$, denoted by $F_c(x)$, is lower-bounded by

$$F_c(x) \geq 1 - \rho x^{-((K-1)S-L+1)}, \quad (76)$$

where ρ is a constant determined by K , S , and L .

Proof: Since each channel coefficient is assumed to be chosen from a continuous distribution, $\mathbf{G}^{[i,j]}$ has full rank almost surely [38]. Moreover, assuming that $(K-1)S > L$, $\mathbf{G}^{[i,j]^H} \mathbf{G}^{[i,j]}$ is the full-rank central Wishart matrix, i.e., $\mathbf{G}^{[i,j]^H} \mathbf{G}^{[i,j]} \sim \mathcal{CW}_L((K-1)S, \mathbf{I}_L)$. Using the high-tail distribution of the complementary CDF in [39, Theorem 4], the CDF is bounded by (76), where $\rho = (\sum_{l=1}^L \kappa_l \mu_l)$. Here, $\kappa_l = \frac{(-1)^{l(l-1)/2}}{((K-1)S-L+1)!}$ and $\mu_l = \int_0^\infty \lambda_1^{((K-1)S-L+1)} g_{L-1, (K-1)S}(\lambda_1) d\lambda_1$, where $\lambda_1 \triangleq \sigma_1^{[i,j]^2}$ and $g_{L-1, (K-1)S}(\cdot)$ denotes the distribution of the largest eigenvalue of a reduced Wishart matrix $\tilde{\mathbf{G}} \sim \mathcal{CW}_{L-1}((K-1)S, \mathbf{I}_{L-1})$. Thus, ρ is determined only by K , S , and L . ■

Now, from Lemma 7, (75) is further bounded by

$$p_c \geq \left(1 - \rho (1+\delta)^{(K-1)S-L+1} \text{SNR}^{-(1+\gamma)((K-1)S-L+1)} \right)^{KN} \\ \geq 1 - \rho KN (1+\delta)^{(K-1)S-L+1} \text{SNR}^{-(1+\gamma)((K-1)S-L+1)}, \quad (77)$$

where (77) follows from $(1-x)^y > 1 - xy$ for any $0 < x < 1 < y$. Therefore, if N scales slower than $\text{SNR}^{(1+\gamma)((K-1)S-L+1)}$ for given γ , i.e., $N = O(\text{SNR}^{(1+\beta)((K-1)S-L+1)})$ where $\beta < \gamma$, then $\lim_{\text{SNR} \rightarrow \infty} p_c \rightarrow 1$ for any given δ , which proves the argument (29).

APPENDIX B PROOF OF LEMMA 4

Let us define β such that

$$\Pr \left\{ \sigma_L^{[i,j]^2} \leq \frac{1}{\beta} \right\} = \frac{S}{N} \iff \Pr \left\{ \frac{1}{\sigma_L^{[i,j]^2}} \geq \beta \right\} = \frac{S}{N}. \quad (78)$$

From [7, Lemma 1], $\Pr \left\{ \sigma_L^{[i,j]^2} \leq \frac{1}{\beta} \right\} = a_0 \beta^{-\tau} + o(\beta^{-\tau})$, where a_0 is a nonzero coefficient, and hence $\beta = \Theta(N^{1/\tau})$. In addition, since $1/\sigma_L^{[i,j]^2}$ of the selected user in each group is the maximum value out of N/S independent random variables, we have

$$\Pr \left\{ \frac{1}{\sigma_L^{[i,j]^2} \Big|_{\text{selected}}} \leq \beta \right\} = \left(1 - \frac{S}{N} \right)^{\frac{N}{S}}. \quad (79)$$

Then, the Markov inequality yields

$$E \left[\frac{1}{\sigma_L^{[i,j]^2} \Big|_{\text{selected}}} \right] \geq \beta \cdot \Pr \left\{ \frac{1}{\sigma_L^{[i,j]^2} \Big|_{\text{selected}}} \geq \beta \right\} \quad (80)$$

$$= \beta \cdot \left(1 - \left(1 - \frac{S}{N} \right)^{\frac{N}{S}} \right) \quad (81)$$

$$= \Theta \left(N^{\frac{1}{\tau}} \right). \quad (82)$$

where (81) follows from (79), and (82) follows from the fact that $\beta = \Theta(N^{1/\tau})$ and that $(1 - \frac{S}{N})^{N/S}$ converges to a constant $1/e$ for increasing N .

APPENDIX C

PROOF OF LEMMA 6

Let us first consider the numerator of the logarithmic term in (67) as follows.

$$\begin{aligned} E \left[\log_2 Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)^\theta | \tilde{\mathbf{H}}_i \right] &= -\frac{\theta}{\log 2} E \left[(\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i)^H \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i) | \tilde{\mathbf{H}}_i \right] \\ &= -\frac{\theta}{\log 2} E \left[\tilde{\mathbf{z}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{z}}_i | \tilde{\mathbf{H}}_i \right] \\ &= -\frac{\theta}{\log 2} \text{tr}(\hat{\mathbf{R}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}^{-1/2}), \end{aligned} \quad (83)$$

where (83) follows from $E[\mathbf{z}^H \mathbf{A} \mathbf{z}] = \text{tr}(\mathbf{A}^{1/2} \mathbf{R} \mathbf{A}^{1/2})$ for any random vector $\mathbf{z}$ with $E\{\mathbf{z} \mathbf{z}^H\} = \mathbf{R}$ and $E\{\mathbf{z}\} = \mathbf{0}$ and for any conjugate symmetric matrix $\mathbf{A}$.

Let us turn to the denominator of the logarithmic term in (67). Here, given that $\tilde{\mathbf{H}}_i$ and $\tilde{\mathbf{y}}_i$ are deterministic, the expectation is taken over $\mathbf{x}_i$. Since $\mathbf{x}_i$ is an S -dimensional complex Gaussian random vector, i.e., the probability density function of $\mathbf{x}_i$ is given by $1/\pi^S \exp(-\|\mathbf{x}_i\|^2)$, we have

$$\begin{aligned} E \left[Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)^\theta | \tilde{\mathbf{y}}_i, \tilde{\mathbf{H}}_i \right] &= \int \exp \left(-\theta (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i)^H \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i) \right) \\ &\quad \cdot \frac{1}{\pi^S} \exp(-\|\mathbf{x}_i\|^2) d\mathbf{x}_i \\ &= \frac{1}{\pi^S} \int \exp \left(-\theta (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i)^H \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i) - \|\mathbf{x}_i\|^2 \right) d\mathbf{x}_i \\ &= \frac{1}{\pi^S} \int \exp(-A) d\mathbf{x}_i. \end{aligned} \quad (84)$$

Here, A can be further expressed as

$$\begin{aligned} A &\triangleq \theta (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i)^H \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{y}}_i - \tilde{\mathbf{H}}_i \mathbf{x}_i) + \|\mathbf{x}_i\|^2 \\ &= \mathbf{x}_i^H \left(\theta \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i + \mathbf{I}_S \right) \mathbf{x}_i - \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \mathbf{x}_i \\ &\quad - \theta \mathbf{x}_i^H \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i + \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i. \end{aligned} \quad (85)$$

Letting $\tilde{\mathbf{\Omega}} = \theta \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i + \mathbf{I}_S$, it follows that

$$A = \left\| \tilde{\mathbf{\Omega}}^{\frac{1}{2}} \mathbf{x}_i - \mathbf{C} \tilde{\mathbf{y}}_i \right\|^2 + \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i - \tilde{\mathbf{y}}_i^H \mathbf{C}^H \mathbf{C} \tilde{\mathbf{y}}_i, \quad (86)$$

where $\mathbf{C}$ is given by $\mathbf{C} = \theta \tilde{\mathbf{\Omega}}^{-1/2} \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1}$, which comes from the equivalence of (85) and (86). Now let us further simplify the last two terms of (86) as

$$\begin{aligned} &\theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i - \tilde{\mathbf{y}}_i^H \mathbf{C}^H \mathbf{C} \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \left(\hat{\mathbf{R}}^{-1} - \theta \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1} \right) \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1/2} \left(\mathbf{I}_S - \theta \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{H}}_i \tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1/2} \right) \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{y}}_i. \end{aligned}$$

Without loss of generality, it follows that $\hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{H}}_i = \mathbf{\Phi} \mathbf{\Lambda} \mathbf{T}^H$, where $\mathbf{\Phi} \in \mathbb{C}^{S \times S}$ and $\mathbf{T} \in \mathbb{C}^{S \times S}$ are orthogonal matrices and $\mathbf{\Lambda}$ is an $(S \times S)$ -dimensional diagonal matrix. Then, we get

$$\tilde{\mathbf{\Omega}} = \theta \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i + \mathbf{I}_S = \mathbf{T}(\theta \mathbf{\Lambda}^2 + \mathbf{I}_S) \mathbf{T}^H. \quad (87)$$

Inserting (87) to $\hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{H}}_i \tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1/2}$ gives us

$$\hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{H}}_i \tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1/2} = \mathbf{\Phi} \mathbf{\Lambda}^2 (\theta \mathbf{\Lambda}^2 + \mathbf{I}_S)^{-1} \mathbf{\Phi}^H,$$

which yields

$$\begin{aligned} &\theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i - \tilde{\mathbf{y}}_i^H \mathbf{C}^H \mathbf{C} \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1/2} \mathbf{\Phi} \left(\mathbf{I}_S - \theta \mathbf{\Lambda}^2 (\theta \mathbf{\Lambda}^2 + \mathbf{I}_S)^{-1} \right) \mathbf{\Phi}^H \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1/2} \mathbf{\Phi} (\theta \mathbf{\Lambda}^2 + \mathbf{I}_S)^{-1} \mathbf{\Phi}^H \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{y}}_i \end{aligned} \quad (88)$$

$$\begin{aligned} &= \theta \tilde{\mathbf{y}}_i^H \hat{\mathbf{R}}^{-1/2} \left(\theta \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H \hat{\mathbf{R}}^{-1/2} + \mathbf{I}_S \right)^{-1} \hat{\mathbf{R}}^{-1/2} \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \left(\theta \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \hat{\mathbf{R}} \right)^{-1} \tilde{\mathbf{y}}_i \\ &= \theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i, \end{aligned} \quad (89)$$

where (88) follows immediately from evaluating the diagonal terms, and (89) follows from $\mathbf{\Omega} \triangleq \theta \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H + \mathbf{I}_S$. Inserting (89) and (86) to (84) gives us

$$\begin{aligned} E \left[Q(\tilde{\mathbf{y}}_i | \mathbf{x}_i)^\theta | \tilde{\mathbf{y}}_i, \tilde{\mathbf{H}}_i \right] &= \frac{1}{\pi^S} \int \exp \left(-\left\| \tilde{\mathbf{\Omega}}^{\frac{1}{2}} \mathbf{x}_i - \mathbf{C} \tilde{\mathbf{y}}_i \right\|^2 - \theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i \right) d\mathbf{x}_i \\ &= \frac{1}{\pi^S} \exp \left(-\theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i \right) \\ &\quad \cdot \int \exp \left(-\left\| \tilde{\mathbf{\Omega}}^{\frac{1}{2}} \mathbf{x}_i - \mathbf{C} \tilde{\mathbf{y}}_i \right\|^2 \right) d\mathbf{x}_i \end{aligned}$$

$$= \frac{1}{\pi^S} \exp \left(-\theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i \right) \pi^S \det(\tilde{\mathbf{\Omega}}^{-1}) \quad (90)$$

$$= \exp \left(-\theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i \right) \det(\mathbf{\Omega}^{-1}), \quad (91)$$

where (90) follows from the fact that for $\mathbf{x}, \mathbf{m} \in \mathbb{C}^{S \times 1}$ and conjugate symmetric $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{S \times S}$,

$$\begin{aligned} & \int \exp(-(\mathbf{A}\mathbf{x} - \mathbf{m})^H \mathbf{B}^{-1} (\mathbf{A}\mathbf{x} - \mathbf{m})) d\mathbf{x} \\ &= \int \exp\left(-(\mathbf{x} - \mathbf{A}^{-1}\mathbf{m})^H \mathbf{A}^H \mathbf{B}^{-1} \mathbf{A} (\mathbf{x} - \mathbf{A}^{-1}\mathbf{m})\right) d\mathbf{x} \\ &= \pi^S \det\left(\mathbf{A}^{-1} \mathbf{B} (\mathbf{A}^H)^{-1}\right), \end{aligned}$$

and (91) follows from $\det(\tilde{\mathbf{\Omega}}) = \det(\mathbf{\Omega})$. Inserting (83) and (91) to (67) gives us

$$\begin{aligned} I(\theta) &= -\frac{\theta}{\log 2} \text{tr}(\hat{\mathbf{R}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}^{-1/2}) \\ &\quad - E\left[\log_2\left(\exp\left(-\theta \tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i\right) \det(\mathbf{\Omega}^{-1})\right) |\tilde{\mathbf{H}}_i\right] \\ &= -\frac{\theta}{\log 2} \text{tr}(\hat{\mathbf{R}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}^{-1/2}) \\ &\quad + \frac{\theta}{\log 2} E\left[\tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i |\tilde{\mathbf{H}}_i\right] + \log_2 \det(\mathbf{\Omega}). \end{aligned}$$

From

$$\begin{aligned} E\left[\tilde{\mathbf{y}}_i^H \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{y}}_i |\tilde{\mathbf{H}}_i\right] &= E\left[\left(\mathbf{x}_i^H \tilde{\mathbf{H}}_i^H + \mathbf{z}_i^H\right) \mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} (\tilde{\mathbf{H}}_i \mathbf{x}_i + \mathbf{z}_i)\right] \\ &= \text{tr}\left(\mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H\right) + \text{tr}\left(\mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \mathbf{R}\right), \end{aligned}$$

we finally have

$$\begin{aligned} I(\theta) &= -\frac{\theta}{\log 2} \text{tr}(\hat{\mathbf{R}}^{-1/2} \mathbf{R} \hat{\mathbf{R}}^{-1/2}) \\ &\quad + \frac{\theta}{\log 2} \left[\text{tr}\left(\mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \tilde{\mathbf{H}}_i \tilde{\mathbf{H}}_i^H\right) + \text{tr}\left(\mathbf{\Omega}^{-1} \hat{\mathbf{R}}^{-1} \mathbf{R}\right)\right] + \log_2 \det(\mathbf{\Omega}), \end{aligned}$$

which completes the proof of the lemma.

REFERENCES

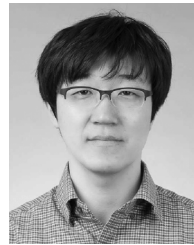
- [1] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [2] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
- [3] B. C. Jung and W.-Y. Shin, "Opportunistic interference alignment for interference-limited cellular TDD uplink," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 148–150, Feb. 2011.
- [4] B. C. Jung, D. Park, and W.-Y. Shin, "Opportunistic interference mitigation achieves optimal degrees-of-freedom in wireless multi-cell uplink networks," *IEEE Trans. Commun.*, vol. 60, no. 7, pp. 1935–1944, Jul. 2012.
- [5] C. Suh and D. Tse, "Interference alignment for cellular networks," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Urbana-Champaign, IL, USA, Sep. 2008, pp. 1037–1044.
- [6] 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.
- [7] H. J. Yang, W.-Y. Shin, B. C. Jung, and A. Paulraj, "Opportunistic interference alignment of MIMO interfering multiple-access channels," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2180–2192, May 2013.
- [8] H. J. Yang, W.-Y. Shin, B. C. Jung, and A. Paulraj, "Opportunistic interference alignment of MIMO IMAC: Effect of user scaling over degrees-of-freedom," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 2646–2650.
- [9] J. Thukral and H. Bölcskei, "Interference alignment with limited feedback," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jul. 2009, pp. 1759–1763.
- [10] R. T. Krishnamachari and M. K. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, USA, Jun. 2010, pp. 619–623.
- [11] B. Mondal and R. W. Heath, Jr., "Performance analysis of quantized beamforming MIMO systems," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4753–4766, Dec. 2006.
- [12] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.
- [13] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [14] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures*, TS 36.213, 3GPP Standard, v.11.2.0.
- [15] L. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.
- [16] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.
- [17] *Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding*, TS 36.212, 3GPP Standard, v.11.2.0.
- [18] C.-B. Chae, D. Mazzarese, T. Inoue, and R. W. Heath, Jr., "Coordinated beamforming for the multiuser MIMO broadcast channel with limited feedforward," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6044–6056, Dec. 2008.
- [19] "Apparatus and method for beamforming with limited feedforward channel in multiple input multiple output wireless communication system," U.S. Patent 7 786 934, Aug. 2010.
- [20] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Channel estimation and linear precoding in multi-cell TDD systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2102–2116, Jun. 2011.
- [21] D. Samardzija, L. Xiao, and N. Mandayam, "Impact of pilot assisted channel state estimation on multiple antenna multiuser TDD systems with spatial filtering," in *Proc. 40th Annu. Conf. Inf. Sci. Syst.*, Lucent Technol. Bell Labs, Holmdel, NJ, Mar. 2006, pp. 381–385.
- [22] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [23] L. Soriano-Equigua, J. Sánchez-García, J. Flores-Troncoso, and R. W. Heath, Jr., "Noniterative coordinated beamforming for multiuser MIMO systems with limited feedforward," *IEEE Signal Process. Lett.*, vol. 18, no. 12, pp. 701–704, Dec. 2011.
- [24] D. J. Love and R. W. Heath, Jr., "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
- [25] W. Dai, Y. E. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, Mar. 2008.
- [26] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.
- [27] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices," *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 424–434, Mar. 2008.
- [28] C. K. Au-Yeung and D. J. Love, "Optimization and tradeoff analysis of two-way limited feedback beamforming systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2570–2579, May 2009.
- [29] B. Khoshnevis, "Multiple-antenna communications with limited channel state information," Ph.D. dissertation, Univ. of Toronto, Toronto, ON, USA, 2011.
- [30] J. Jose, S. Subramanian, X. Wu, and J. Li, "Opportunistic interference alignment in cellular downlink," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput.*, Urbana-Champaign, IL, USA, Oct. 2012, pp. 1529–1545.
- [31] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.

- [32] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.
- [33] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai (Shitz), "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.
- [34] J. H. Lee and W. Choi, "On the achievable dof and user scaling law of opportunistic interference alignment in 3-transmitter MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2743–2753, Jun. 2013.
- [35] H. D. Nguyen, R. Zhang, and H. T. Hui, "Multi-cell random beamforming: Achievable rate and degrees-of-freedom region," *IEEE Trans. Signal Process.*, vol. 61, no. 14, pp. 3532–3544, July 2013.
- [36] H. D. Nguyen, R. Zhang, and H. T. Hui, "Effect of receive spatial diversity on the degrees-of-freedom region in multi-cell random beamforming," [Online]. Available: <http://arxiv.org/abs/1303.5947>, preprint
- [37] H. J. Yang, W.-Y. Shin, B. C. Jung, C. Suh, and A. Paulraj, "Opportunistic downlink interference alignment," [Online]. Available: <http://arxiv.org/abs/1312.7198>, preprint
- [38] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Inst. of Technol., Cambridge, MA, USA, 1989.
- [39] M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex Wishart matrices," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1705–1717, Jun. 2010.



Hyun Jong Yang (M'12) received the B.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2004, and the M.S. and Ph.D. degrees in electrical engineering from KAIST, in 2006 and 2010, respectively. From Aug. 2010 to Aug. 2011, he was a research fellow at Korea Institute Ocean Science Technology (KIOST), Daejeon, Republic of Korea. From Oct. 2011 to Oct. 2012, he worked as a post-doctoral researcher in the Electrical Engineering Department, Stanford

University, Stanford, CA. From Oct. 2012 to Aug. 2013, he was a Staff II Systems Design Engineer, Broadcom Corporation, Sunnyvale, CA, where he developed physical-layer algorithms for LTE-A MIMO receivers. In addition, he was a delegate of Broadcom in 3GPP standard meetings for RAN1 Rel-12 technologies. Since Aug. 2013, he has been an Assistant Professor in the School of Electrical and Computer Engineering, UNIST, Ulsan, Korea. His fields of interests are algorithms and theory for wireless communication and their applications and implementation.



Bang Chul Jung (S'02–M'08–SM'14) received the B.S. degree in Electronics Engineering from Ajou University, Suwon, Korea, in 2002 and the M.S. and Ph.D. degrees in Electrical & Computer Engineering from KAIST, Daejeon, Korea, in 2004 and 2008, respectively. He was a senior researcher/research professor with KAIST Institute for Information Technology Convergence, Daejeon, Korea, from January 2009 to February 2010. He is now an associate professor of Department of Information and Communication Engineering, Gyeongsang National University, Korea. Dr. Jung was the recipient of the Fifth IEEE Communication Society Asia-Pacific Outstanding Young Researcher Award in 2011. His research interests include wireless communications, compressed sensing, interference management, MIMO, multiple access techniques, and radio resource management.



Won-Yong Shin (S'02–M'08) received the B.S. degree in electrical engineering from Yonsei University, Seoul, Korea, in 2002. He received the M.S. and the Ph.D. degrees in electrical engineering and computer science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2004 and 2008, respectively.

From February 2008 to April 2008, he was a Visiting Scholar in the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA. From September 2008 to April 2009, he was with the Brain Korea Institute and CHiPS at KAIST as a Postdoctoral Fellow. From August 2008 to April 2009, he was with the Lumicomm, Inc., Daejeon, Korea, as a Visiting Researcher. In May 2009, he joined Harvard University as a Postdoctoral Fellow and was promoted to a Research Associate in October 2011. Since March 2012, he has been with the Division of Mobile Systems Engineering, College of International Studies, Dankook University, Yongin, Korea, where he is currently an Assistant Professor. His research interests are in the areas of information theory, communications, signal processing, and their applications to multiuser networking issues.



Arogyaswami Paulraj (SM'85–F'91) received the B.E. degree from Naval College of Engineering, Lonavala, India, in 1966 and the Ph.D. degree from the Indian Institute of Technology, Delhi, India, in 1973. He is a pioneer of MIMO wireless communications. He joined Stanford University, Stanford, CA, in 1972, after a 20-year industry/military career in India. He founded two successful semiconductor companies—Iospan Wireless acquired by Intel Corp. in 2003 and Beceem Communications acquired by Broadcom Corp. in 2010. He has authored over 400

research publications and two text books, and is a co-inventor in 52 US patents. Dr. Paulraj is a member of the U.S. National Academy of Engineering, the Royal Swedish Academy of Engineering Sciences, the Academy of Sciences for the Developing World (TWAS), and the Indian National Academy of Engineering. He was the recipient of a number of national awards in India for sonar development and of several IEEE awards, including the Technical Achievement Award from the Signal Processing Society, fellow grade, and several best paper awards. He was awarded the 2011 IEEE Alexander Graham Bell Medal for his work in MIMO.