Linear Precoder Designs for K-user Interference Channels

Hakjea Sung, *Student Member, IEEE*, Seok-Hwan Park, *Student Member, IEEE*, Kyoung-Jae Lee, *Student Member, IEEE*, and Inkyu Lee, *Senior Member, IEEE*

***Abstract*—This paper studies linear precoding and decoding schemes for K-user interference channel systems. It was shown by Cadambe and Jafar that the interference alignment (IA) algorithm achieves a theoretical bound on degrees of freedom (DOF) for interference channel systems. Based on this, we first introduce a non-iterative solution for the precoding and decoding scheme. To this end, we determine the orthonormal basis vectors of each user's precoding matrix to achieve the maximum DOF, then we optimize precoding matrices in the IA method according to two different decoding schemes with respect to individual rate. Second, an iterative processing algorithm is proposed which maximizes the weighted sum rate. Deriving the gradient of the weighted sum rate and applying the gradient descent method, the proposed scheme identifies a local-optimal solution iteratively. Simulation results show that the proposed iterative algorithm outperforms other existing methods in terms of sum rate. Also, we exhibit that the proposed non-iterative method approaches a local optimal solution at high signal-to-noise ratio with reduced complexity.**

***Index Terms*—Interference channel, interference alignment (IA), linear precoding, minimum mean-square error (MMSE) filtering, gradient decent.**

I. INTRODUCTION

In the past years, researches on information theory have been applied for Gaussian interference channels, and several results have been introduced for special cases [1] [2]. However, the capacity region of interference channels still remains unknown in general [3]. Recently, the investigation of degrees of freedom (DOF) has been of concern in the interference channels [4] [5]. The DOF is defined as

$$\text{DOF} = \lim_{\rho\to\infty} \frac{C_\Sigma(\rho)}{\log\rho}$$

where $C_\Sigma(\rho)$ is the ergodic sum capacity at signal-to-noise ratio (SNR) ρ. By definition, the DOF is equivalent to the multiplexing gain. Hence, this can be used as a leverage on characterizing the system performance.

Manuscript received February 12, 2009; revised July 9, 2009; accepted September 29, 2009. The associate editor coordinating the review of this paper and approving it for publication was Y. Sanada.

This research was supported in part by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2009-2009-C1090-0902-0013), and in part by Seoul R&BD Program (R0911771). This paper was presented in part at the IEEE Global Communications Conference (GLOBECOM), Honolulu, USA, November 2009.

The authors are with the School of Electrical Engineering, Korea University, Seoul, Korea (e-mail: {jaysung, psk1501, kyoungjae, inkyu}@korea.ac.kr).

Digital Object Identifier 10.1109/TWC.2010.01.090221

In [4], an upperbound on the DOF has been determined, and an achievable precoding method named as interference alignment (IA) has been introduced for interference channel systems. For the case where the global channel knowledge is available at each node, the IA methods designs the precoding matrix at each transmitter restricting all interference at every receivers to approximately half of the received signal space, leaving the other half interference free for the desired signal. This promising method successfully achieves a theoretical bound on the DOF for interference channels. However, since this method has approached the precoder design problem only from the DOF point of view, there are still chances to improve the sum rate performance.

In this paper, we propose two linear precoding and decoding methods for the K-user interference channel systems. The focus is on the case where linear operation is considered for the proposed schemes. Then, the interference cancellation techniques are not allowed at both transmitter and receiver, and other users' interference is treated as additive noise. Within this context, the first proposed method employs a simple non-iterative algorithm for identifying the precoding and decoding matrices. In order to achieve the maximum DOF, we determine the basis vectors of precoding matrices based on the modified IA method in which the chordal distance criterion is additionally applied to make the desired signal space and the interference signal space roughly orthogonal to each other. Then, we employ the block interference suppression concept used in single cell multi-user downlink systems [6] [7] for receiver filters and optimize the precoding matrices obtained from the IA algorithm according to the decoding methods such that the resulting individual rate is maximized.

The second proposed method is an iterative solution which shows better sum rate performance than non-iterative solutions with the increased computational complexity. Unlike the non-iterative algorithm which identifies the precoding matrices for satisfying the interference aligning constraints, the proposed algorithm iteratively computes the precoding matrices which maximize the weighted sum rate by applying a gradient descent algorithm [8]. Although the gradient descent algorithm may not guarantee the global optimal solution, a locally maximized sum rate is found. From simulation results, we show that the proposed schemes outperform the traditional method such as orthogonal resource sharing while achieving the theoretical bound on the DOF, and illustrate that the proposed non-iterative method approaches a local optimal solution obtained by an iterative technique at high SNR with lower complexity.

A similar idea with the proposed non-iterative method was introduced in [9] as an improved IA scheme. However, an explicit solution has not been given in [9] as will be shown in this paper. Also, in [10] and [11], iterative processing schemes have been introduced for the interference channel systems. However, the algorithm in [10] optimizes only the transmit covariance matrices. Although the method in [11] determines both precoding and decoding matrices which maximize the signal-to-interference-plus-noise ratio (SINR) using the joint transmitter-receiver optimization, this scheme becomes sub-optimal as compared to the proposed iterative method. This is due to a fact that the scheme in [11] computes the precoding and decoding matrices in a distributed iterative manner, whereas the proposed scheme identifies the precoder with a centralized optimization approach.

This paper is organized as follows: In Section II, we describe a general system model for the K-user interference channel. Section III reviews the conventional IA methods, and in Section IV, we propose a non-iterative algorithm based on the IA method. Section V introduces an iterative weighted sum rate maximization scheme. In Section VI, the simulation results are presented. Finally, the paper is terminated with conclusions in Section VII.

The following notations are used throughout the paper. We employ uppercase boldface letters for matrices and lowercase boldface for vectors. For any general matrix $\mathbf{X}$, $\mathbf{X}^T$, $\mathbf{X}^*$ and $\mathbf{X}^H$ denote the transpose, the conjugate and the conjugate transpose, respectively. $\mathrm{Tr}(\mathbf{X})$ indicates the trace and the Frobenius norm of a matrix $\mathbf{X}$ is $\|\mathbf{X}\|_F^2 = \mathrm{Tr}(\mathbf{X}\mathbf{X}^H)$. $|\mathbf{X}|$ and vec$(\mathbf{X})$ represent the determinant and the stacked columns of a matrix $\mathbf{X}$, respectively.

II. System Model

We consider K-user interference channel systems where K transmitters are transmitting independent data streams to K receivers simultaneously and generating co-channel interference at all receivers as shown in Fig 1. Also, there is no coordination among transmitters and also among receivers. In this system, the jth transmitter is equipped with N_{t_j} antennas and receiver i has N_{r_i} antennas. In the discrete-time complex baseband case, the channel from the jth transmitter to the ith receiver is modeled by the $N_{r_i} \times N_{t_j}$ channel matrix $\mathbf{H}_{ij}$. We assume that the channel information is globally available, i.e., each node perfectly knows all channel coefficients.

First, we define the $d_i \times 1$ data symbol vector for user i as $\mathbf{x}_i$ where d_i is the number of data streams for user i and $d_i \leq N_{t_i}$. Here, we assume that all symbols of $\mathbf{x}_i$ are independently generated with unit variance, i.e., $\mathbb{E}[\mathbf{x}_i\mathbf{x}_i^H] = \mathbf{I}_{d_i}$ for $i = 1, \cdots, K$, and these are only known to one specific transmitter such that no joint transmission such as dirty paper coding [12] is possible. Also, defining the $N_{t_i} \times d_i$ precoding matrix for user i as $\mathbf{P}_i$, the $N_{r_i} \times 1$ received signal vector $\mathbf{y}_i$ is given by

$$\mathbf{y}_i = \mathbf{H}_{ii}\mathbf{P}_i\mathbf{x}_i + \sum_{j=1, j\neq i}^{K} \mathbf{H}_{ij}\mathbf{P}_j\mathbf{x}_j + \mathbf{n}_i \qquad (1)$$

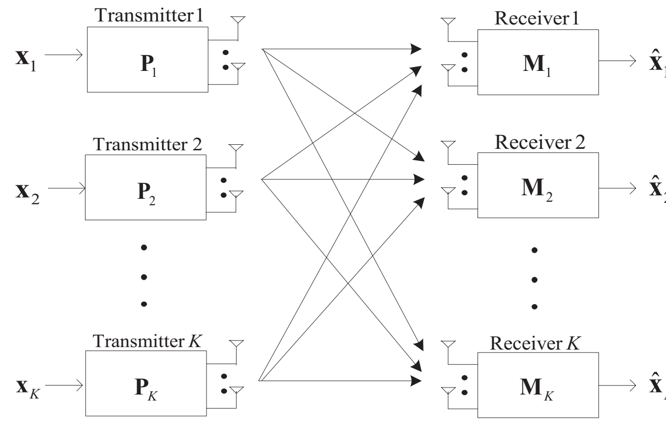

Fig. 1. Structure of K user interference channel systems.

where $\mathbf{n}_i$ denotes the independent and identically distributed (i.i.d.) complex Gaussian noise vector at receiver i with zero mean and $\mathbb{E}[\mathbf{n}_i\mathbf{n}_i^H] = \sigma_n^2\mathbf{I}_{N_{r_i}}$ for $i = 1, \cdots, K$.

In (1), the first term $\mathbf{H}_{ii}\mathbf{P}_i\mathbf{x}_i$ is the desired signal vector sent by the ith transmitter and $\sum_{j\neq i}^{K} \mathbf{H}_{ij}\mathbf{P}_j\mathbf{x}_j$ represents the interference from other transmitters. Also, each transmitter i has to satisfy the transmit power constraint $\mathbb{E}[\|\mathbf{P}_i\mathbf{x}_i\|^2] \leq P_i$ where P_i denotes the maximum average transmitted power of the ith transmitter. Since we assume $\mathbb{E}[\mathbf{x}_i\mathbf{x}_i^H] = \mathbf{I}_{d_i}$, the total transmit power constraint for transmitter i can be expressed as $\mathrm{Tr}(\mathbf{P}_i^H\mathbf{P}_i) \leq P_i$ for $i = 1, \cdots, K$.

In this paper, we focus on linear processing scheme where no interference cancellation is performed and the interference generated by other transmitters is treated as additive noise from each receiver. Then, denoting the $d_i \times N_{r_i}$ decoding matrix of the ith receiver as $\mathbf{M}_i$ which decouples the N_{r_i} interfering received signals into d_i substreams for single symbol detection, the $d_i \times 1$ receive filter output vector of user i can be written as

$$\hat{\mathbf{x}}_i = \mathbf{M}_i\mathbf{y}_i = \mathbf{M}_i\mathbf{H}_{ii}\mathbf{P}_i\mathbf{x}_i + \mathbf{M}_i\sum_{j=1, j\neq i}^{K} \mathbf{H}_{ij}\mathbf{P}_j\mathbf{x}_j + \mathbf{M}_i\mathbf{n}_i. \qquad (2)$$

III. Review of Interference Alignment Methods

In this section, we review the IA algorithm presented in [4]. As a non-iterative linear precoding method, this algorithm enables to achieve the theoretical bound on the DOF with a simple zero-forcing (ZF) filter at each receiver. For the brief review, we illustrate this method for the case of $K = 3$. A more general case can be found in [4] and [13]. This representation will serve as a basis of the proposed non-iterative scheme in Section IV.

A. Interference Alignment for 3-user MIMO case

For multi-input multi-output (MIMO) interference channels, we consider the case where all nodes have M antennas. i.e., $N_{t_i} = N_{r_i} = M$, and M is even. In this case, the achievable DOF are $3M/2$ and the received signal vector of the ith receiver in (1) can be written as

$$\mathbf{y}_i = \mathbf{H}_{i1}\mathbf{P}_1\mathbf{x}_1 + \mathbf{H}_{i2}\mathbf{P}_2\mathbf{x}_2 + \mathbf{H}_{i3}\mathbf{P}_3\mathbf{x}_3 + \mathbf{n}_i$$

where $\mathbf{H}_{ij}$ is an $M \times M$ full rank channel matrix. In order to transmit $3M/2$ total independent data streams, $\mathbf{x}_i$ and $\mathbf{P}_i$ can be an $\frac{M}{2} \times 1$ vector and an $M \times \frac{M}{2}$ matrix, respectively, for $i = 1, 2, 3$.

To decode $M/2$ data streams from the $M \times 1$ received signal vector $\mathbf{y}_i$, the interference signal space should have at most $M/2$ dimension and be linearly independent with the desired signal space. Thus, each precoder has to be designed to satisfy the three interference aligning constraints described as

$$\begin{aligned} \mathrm{span}(\mathbf{H}_{12}\mathbf{P}_2) &= \mathrm{span}(\mathbf{H}_{13}\mathbf{P}_3), \\ \mathrm{span}(\mathbf{H}_{21}\mathbf{P}_1) &= \mathrm{span}(\mathbf{H}_{23}\mathbf{P}_3), \\ \mathrm{span}(\mathbf{H}_{31}\mathbf{P}_1) &= \mathrm{span}(\mathbf{H}_{32}\mathbf{P}_2) \end{aligned} \tag{3}$$

where $\mathrm{span}(\mathbf{X})$ indicates the vector space spanned by the column vectors of $\mathbf{X}$.

Then, to compute the proper precoding matrices, the IA method restricts the above constraints as

$$\begin{aligned} &\mathrm{span}(\mathbf{H}_{12}\mathbf{P}_2) = \mathrm{span}(\mathbf{H}_{13}\mathbf{P}_3), \\ &\mathbf{H}_{21}\mathbf{P}_1 = \mathbf{H}_{23}\mathbf{P}_3, \quad \mathbf{H}_{31}\mathbf{P}_1 = \mathbf{H}_{32}\mathbf{P}_2. \end{aligned} \tag{4}$$

These equations can be equivalently expressed as

$$\begin{aligned} &\mathrm{span}(\mathbf{P}_1) = \mathrm{span}(\mathbf{E}\mathbf{P}_1), \\ &\mathbf{P}_2 = \mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{P}_1, \quad \mathbf{P}_3 = \mathbf{H}_{23}^{-1}\mathbf{H}_{21}\mathbf{P}_1 \end{aligned} \tag{5}$$

where $\mathbf{E} = \mathbf{H}_{31}^{-1}\mathbf{H}_{32}\mathbf{H}_{12}^{-1}\mathbf{H}_{13}\mathbf{H}_{23}^{-1}\mathbf{H}_{21}$.

Finally, we can set $\mathbf{P}_1$ to be

$$\mathbf{P}_1 = [\,\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_{M/2}\,] \tag{6}$$

where $\mathbf{e}_1, \cdots, \mathbf{e}_M$ are the eigenvectors of $\mathbf{E}$. Similarly, we obtain $\mathbf{P}_2$ and $\mathbf{P}_3$ from (5) and (6) for the 3-user MIMO interference channels.

B. Interference Alignment for 3-user SISO case

For single-input single-output (SISO) interference channels, we need a symbol extension of the channel in order to apply preprocessing which aligns the interference signals. The symbol extension can be made in time-slots or frequency-slots. Due to the causality requirement, we consider that the channel between each transmitter and receiver is comprised of $2n + 1$ orthogonal frequency dimensions where n is an arbitrary positive integer in the frequency selective fading case. Then, the received signal vector of the ith receiver can be written as

$$\mathbf{y}_i = \mathbf{H}_{i1}\mathbf{P}_1\mathbf{x}_1 + \mathbf{H}_{i2}\mathbf{P}_2\mathbf{x}_2 + \mathbf{H}_{i3}\mathbf{P}_3\mathbf{x}_3 + \mathbf{n}_i.$$

Here, unlike the MIMO case, $\mathbf{y}_i$ and $\mathbf{n}_i$ are the $(2n + 1) \times 1$ vectors extended along the frequency slots, and $\mathbf{H}_{ij}$ is a $(2n + 1) \times (2n + 1)$ diagonal matrix whose diagonal elements represent the channel coefficient corresponding to each subcarrier. Also, in this case, the DOFs of user 1, user 2 and user 3 are $n + 1$, n and n, respectively. Thus, $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ can be the $(2n + 1) \times (n + 1)$, $(2n + 1) \times n$ and $(2n + 1) \times n$ matrices, respectively.

Similar to the MIMO case, to align the interference, the following constraints are imposed.

$$\mathbf{H}_{12}\mathbf{P}_2 = \mathbf{H}_{13}\mathbf{P}_3, \;\; \mathbf{H}_{23}\mathbf{P}_3 \prec \mathbf{H}_{21}\mathbf{P}_1, \;\; \mathbf{H}_{32}\mathbf{P}_2 \prec \mathbf{H}_{31}\mathbf{P}_1$$

where $\mathbf{X} \prec \mathbf{Y}$ represents that the set of column vectors of $\mathbf{X}$ is a subset of the set of column vectors of $\mathbf{Y}$. The above equations can be rewritten as

$$\mathbf{B} = \mathbf{T}\mathbf{C}, \quad \mathbf{B} \prec \mathbf{A}, \quad \mathbf{C} \prec \mathbf{A} \tag{7}$$

where

$$\begin{aligned} \mathbf{T} &= \mathbf{H}_{12}\mathbf{H}_{21}^{-1}\mathbf{H}_{23}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{H}_{13}^{-1}, \quad \mathbf{A} = \mathbf{P}_1, \\ \mathbf{B} &= \mathbf{H}_{21}^{-1}\mathbf{H}_{23}\mathbf{P}_3, \quad \mathbf{C} = \mathbf{H}_{31}^{-1}\mathbf{H}_{32}\mathbf{P}_2. \end{aligned} \tag{8}$$

To satisfy the above constraints in (12), $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ can be chosen as

$$\mathbf{A} = [\,\mathbf{w} \quad \mathbf{T}\mathbf{w} \quad \mathbf{T}^2\mathbf{w} \quad \cdots \quad \mathbf{T}^n\mathbf{w}\,], \tag{9}$$

$$\mathbf{B} = [\,\mathbf{T}\mathbf{w} \quad \mathbf{T}^2\mathbf{w} \quad \cdots \quad \mathbf{T}^n\mathbf{w}\,], \tag{10}$$

$$\mathbf{C} = [\,\mathbf{w} \quad \mathbf{T}\mathbf{w} \quad \cdots \quad \mathbf{T}^{(n-1)}\mathbf{w}\,] \tag{11}$$

where the $(2n + 1) \times 1$ column vector $\mathbf{w}$ is defined in [4] as

$$\mathbf{w} = [\,1 \quad 1 \quad \cdots \quad 1\,]^T. \tag{12}$$

Finally, we can obtain $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$ using (8) - (12) for the SISO interference channels.

IV. Noniterative Algorithm Based on Interference Alignment

Based on the conventional IA algorithm reviewed in the previous section, we propose non-iterative linear precoder and decoder design methods for interference channel systems. Before addressing this, we first introduce a combination matrix $\mathbb{C}(\mathbf{X})$. Any $m \times l$ matrix $\mathbf{X}$ can be decomposed as $\mathbb{O}(\mathbf{X})$ and $\mathbb{C}(\mathbf{X})$, i.e., $\mathbf{X} = \mathbb{O}(\mathbf{X})\mathbb{C}(\mathbf{X})$ where $\mathbb{O}(\mathbf{X})$ is defined as a matrix which consists of the orthonormal basis vectors that span the column space of $\mathbf{X}$ and $\mathbb{C}(\mathbf{X})$ denotes the combination matrix of $\mathbf{X}$. For instance, denoting the QR decomposition of $\mathbf{X}$ as $\mathbf{X} = \mathbf{Q}\mathbf{R}$, the $m \times l$ matrix $\mathbb{O}(\mathbf{X})$ can be obtained as $\mathbf{Q}\mathbf{Y}$ where $\mathbf{Y}$ is an arbitrary $l \times l$ unitary matrix, and the $l \times l$ matrix $\mathbb{C}(\mathbf{X})$ is given as $\mathbf{Y}^H\mathbf{R}$. Using this property, we can decompose each precoding matrix $\mathbf{P}_i$ as $\mathbf{P}_i = \mathbb{O}(\mathbf{P}_i)\mathbb{C}(\mathbf{P}_i)$.

In some sense, the conventional IA is a method which characterizes only $\mathbb{O}(\mathbf{P}_i)$ since the interference alignment is determined only with $\mathbb{O}(\mathbf{P}_i)$, and there is no consideration on $\mathbb{C}(\mathbf{P}_i)$. The combination matrix $\mathbb{C}(\mathbf{P}_i)$ randomly chosen from the IA method does not affect the DOF. However, this leads to a degradation of the sum rate performance. Thus, we identify the proposed non-iterative scheme with the following procedure. First, the basis vectors of the precoding matrix are determined from the IA method to achieve the maximum DOF, then the receiver filters are chosen to suppress interference among users. Finally, the combination matrix of each precoder is computed according to different receiver filters such that the resulting individual rate is maximized. In what follows, we explain the proposed non-iterative algorithms for the 3-user case, and these schemes can be adopted for more general cases as in [4] and [13].

$$\mathbf{P}_{\rm cd,1} = \underset{\mathbf{P}_1 \prec \text{eig}(\mathbf{E})}{\arg\max} \left\{ d_{\rm cd}(\mathbf{H}_{11}\mathbf{P}_1, \mathbf{H}_{12}\mathbf{P}_2) + d_{\rm cd}(\mathbf{H}_{22}\mathbf{P}_2, \mathbf{H}_{21}\mathbf{P}_1) + d_{\rm cd}(\mathbf{H}_{33}\mathbf{P}_3, \mathbf{H}_{31}\mathbf{P}_1) \right\} \tag{13}$$

$$= \underset{\mathbf{P}_1 \prec \text{eig}(\mathbf{E})}{\arg\max} \left\{ \sqrt{\frac{M}{2} - \|\mathbb{O}(\mathbf{H}_{11}\mathbf{P}_1)^H \mathbb{O}(\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{P}_1)\|_F^2} + \sqrt{\frac{M}{2} - \|\mathbb{O}(\mathbf{H}_{22}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{P}_1)^H \mathbb{O}(\mathbf{H}_{21}\mathbf{P}_1)\|_F^2} + \sqrt{\frac{M}{2} - \|\mathbb{O}\mathbf{H}_{33}(\mathbf{H}_{23}^{-1}\mathbf{H}_{21}\mathbf{P}_1)^H \mathbb{O}(\mathbf{H}_{31}\mathbf{P}_1)\|_F^2} \right\} \tag{14}$$

A. Precoder basis determination using the modified Interference Alignment

As shown in Section III, the conventional IA is a method which determines the signal space, span$(\mathbf{H}_{ij}\mathbf{P}_j)$ for $i, j = 1, 2, 3$, and this has been done under the interference aligning constraints in (3), which means that the conventional IA obtains $\mathbf{P}_j$ only with a consideration of the interference signal space. However, it is desirable to take the desired signal space into account for finding $\mathbf{P}_j$ in order to make the desired signal space, span$(\mathbf{H}_{ii}\mathbf{P}_i)$, and the interference signal space, span$(\mathbf{H}_{ij}\mathbf{P}_j)$ for $j \neq i$, roughly orthogonal to each other. By doing so, we can minimize the interference signals that spill over the desired signal space. To this end, we apply a maximum chordal distance criterion additionally to the conventional IA, and identify the precoder basis vectors.

1) MIMO case: For the MIMO case, the conventional IA computes the precoding matrices by simply constructing $\mathbf{P}_1$ with the first $M/2$ eigenvectors of $\mathbf{E}$ as in (6). Instead, we make $\mathbf{P}_1$ with $M/2$ column vectors selected from M eigenvectors of $\mathbf{E}$ with a maximum chordal distance criterion. Let us define the chordal distance between an $m \times n_1$ matrix $\mathbf{X}_1$ and $m \times n_2$ matrix $\mathbf{X}_2$ for $m \geq n_1, n_2$ as [14]

$$\begin{aligned} d_{\rm cd}(\mathbf{X}_1, \mathbf{X}_2) &= \frac{1}{\sqrt{2}} \|\mathbb{O}(\mathbf{X}_1)\mathbb{O}(\mathbf{X}_1)^H - \mathbb{O}(\mathbf{X}_2)\mathbb{O}(\mathbf{X}_2)^H\|_F \\ &= \sqrt{\frac{n_1 + n_2}{2} - \|\mathbb{O}(\mathbf{X}_1)^H \mathbb{O}(\mathbf{X}_2)\|_F^2}. \end{aligned} \tag{15}$$

Then, we can formulate the selection problem from (5) and (15) as (14). In (14), eig$(\mathbf{E})$ represents the matrix made of M eigenvectors of $\mathbf{E}$. Now we choose the subset which maximizes the chordal distance in (14) out of $\binom{M}{M/2}$ eigenvector subsets of $\mathbf{E}$ as $\mathbf{P}_{\rm cd,1}$. After determining $\mathbf{P}_{\rm cd,1}$, $\mathbf{P}_{\rm cd,2}$ and $\mathbf{P}_{\rm cd,3}$ are also found using (5). Note that, since the interference signal spaces from other users are all aligned, we only consider the distance between the desired signal term $\mathbf{H}_{ii}\mathbf{P}_i$ and one of interference signal terms $\mathbf{H}_{ij}\mathbf{P}_j$ for $i, j = 1, 2, 3$ and $i \neq j$ to formulate the problem in (13).

2) SISO case: In the SISO case, the conventional IA obtains the precoding matrices by setting all elements of $\mathbf{w}$ to 1 as in (12). In contrast, adjusting each element of $\mathbf{w} = [w_1 \ \cdots \ w_{2n+1}]^T$, we can maximize the chordal distance in (13) since all precoding matrices are a function of $\mathbf{w}$ and the interference aligning constraints in (7) are also satisfied with any nonzero complex w_m for $m = 1, \cdots, 2n+1$. Then the problem which identifies the chordal distance maximizing vector $\mathbf{w}_{\rm cd}$ can be written from (15) as

$$\begin{aligned} \mathbf{w}_{\rm cd} = \arg\max_{\mathbf{w}} \Big\{ &\sqrt{N - \|\mathbb{O}(\mathbf{H}_{11}\mathbf{P}_1)^H \mathbb{O}(\mathbf{H}_{12}\mathbf{P}_2)\|_F^2} \\ &+ \sqrt{N - \|\mathbb{O}(\mathbf{H}_{22}\mathbf{P}_2)^H \mathbb{O}(\mathbf{H}_{21}\mathbf{P}_1)\|_F^2} \\ &+ \sqrt{N - \|\mathbb{O}(\mathbf{H}_{33}\mathbf{P}_3)^H \mathbb{O}(\mathbf{H}_{31}\mathbf{P}_1)\|_F^2} \Big\} \end{aligned} \tag{16}$$

where $N = (2n+1)/2$.

However, unlike the selection problem (14) of the MIMO case where the eigenvectors $\mathbf{E}$ of are known, it is not possible to obtain the optimal $\mathbf{w}_{\rm cd}$ directly from the above equation due to the operation in $\mathbb{O}(\mathbf{X})$. Thus, we approximate the above problem as

$$\begin{aligned} \mathbf{w}_{\rm cd} = \arg\min_{\mathbf{w}} \big\{ &\|(\mathbf{H}_{11}\mathbf{P}_1)^T \mathbf{H}_{12}\mathbf{P}_2\|_F^2 + \|(\mathbf{H}_{22}\mathbf{P}_2)^T \mathbf{H}_{21}\mathbf{P}_1\|_F^2 \\ &+ \|(\mathbf{H}_{33}\mathbf{P}_3)^T \mathbf{H}_{31}\mathbf{P}_1\|_F^2 \big\}. \end{aligned} \tag{17}$$

Here, since $\mathbf{T}$ and $\mathbf{H}_{ij}$ have a diagonal structure for $i, j = 1, 2, 3$ from (8) and (11), the first term of the argument $\|(\mathbf{H}_{11}\mathbf{P}_1)^T \mathbf{H}_{12}\mathbf{P}_2\|_F^2$ can be rewritten as

$$\begin{aligned} &\|(\mathbf{H}_{11}\mathbf{P}_1)^T \mathbf{H}_{12}\mathbf{P}_2\|_F^2 \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n+1} \left| \mathbf{w}^T \mathbf{T}^{(l-1)} \mathbf{H}_{11}\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{T}^{(k-1)} \mathbf{w} \right|^2 \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n+1} \left| \mathbb{D}\big(\mathbf{T}^{(l-1)} \mathbf{H}_{11}\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{T}^{(k-1)}\big) \mathbf{w}_{\rm sq} \right|^2 \end{aligned} \tag{18}$$

where $\mathbb{D}(\mathbf{X})$ is defined as a row vector consisting of the diagonal elements of $\mathbf{X}$ and the column vector $\mathbf{w}_{\rm sq}$ is denoted by $\mathbf{w}_{\rm sq} = [w_{\rm sq,1} \ w_{\rm sq,2} \ \cdots \ w_{\rm sq,2n+1}]^T = [w_1^2 \ w_2^2 \ \cdots \ w_{2n+1}^2]^T$. A detailed derivation of (18) is presented in Appendix A. Similarly, the other two terms are given by

$$\begin{aligned} &\|(\mathbf{H}_{22}\mathbf{P}_2)^T \mathbf{H}_{21}\mathbf{P}_1\|_F^2 \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n+1} \left| \mathbb{D}\big(\mathbf{T}^{(k-1)} \mathbf{H}_{31}\mathbf{H}_{32}^{-1}\mathbf{H}_{22}\mathbf{H}_{21}\mathbf{T}^{(l-1)}\big) \mathbf{w}_{\rm sq} \right|^2, \end{aligned} \tag{19}$$

$$\begin{aligned} &\|(\mathbf{H}_{33}\mathbf{P}_3)^T \mathbf{H}_{31}\mathbf{P}_1\|_F^2 \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n+1} \left| \mathbb{D}\big(\mathbf{T}^{k} \mathbf{H}_{21}\mathbf{H}_{23}^{-1}\mathbf{H}_{33}\mathbf{H}_{31}\mathbf{T}^{(l-1)}\big) \mathbf{w}_{\rm sq} \right|^2. \end{aligned} \tag{20}$$

Then, applying (18)-(20) to (17), the optimization problem becomes

$$\hat{\mathbf{w}}_{\rm sq} = \arg\min_{\mathbf{w}_{\rm sq}} \|\boldsymbol{\Omega}\,\mathbf{w}_{\rm sq}\|^2 \tag{21}$$

$$\text{subject to } \|\mathbf{w}_{\rm sq}\| = 1 \tag{22}$$

where $\mathbf{\Omega} = [\mathbf{\Omega}_{[1,1]}^T \cdots \mathbf{\Omega}_{[1,n+1]}^T \cdots \mathbf{\Omega}_{[n,1]}^T \cdots \mathbf{\Omega}_{[n,n+1]}^T]^T$ and the $3 \times (2n+1)$ matrix $\mathbf{\Omega}_{[k,l]}$ is defined as

$$\mathbf{\Omega}_{[k,l]} = \begin{bmatrix} \mathbb{D}\left(\mathbf{T}^{(l-1)}\mathbf{H}_{11}\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{T}^{(k-1)}\right) \\ \mathbb{D}\left(\mathbf{T}^{(k-1)}\mathbf{H}_{31}\mathbf{H}_{32}^{-1}\mathbf{H}_{22}\mathbf{H}_{21}\mathbf{T}^{(l-1)}\right) \\ \mathbb{D}\left(\mathbf{T}^{k}\mathbf{H}_{21}\mathbf{H}_{23}^{-1}\mathbf{H}_{33}\mathbf{H}_{31}\mathbf{T}^{(l-1)}\right) \end{bmatrix}.$$

In order to obtain a non-trivial solution in (21), we impose the constraint that $\mathbf{w}_{\text{sq}}$ is a unit norm vector since the amplitude of $\mathbf{w}$ does not affect the chordal distance in (16). Then the solution of (21) $\hat{\mathbf{w}}_{\text{sq}}$ equals the right singular vector corresponding to the minimum singular value of $\mathbf{\Omega}$, and the optimized vector $\mathbf{w}_{\text{cd}}$ results in $\mathbf{w}_{\text{cd}} = [\sqrt{\hat{w}_{\text{sq},1}}\ \sqrt{\hat{w}_{\text{sq},2}} \cdots \sqrt{\hat{w}_{\text{sq},2n+1}}]^T$. With this $\mathbf{w}_{\text{cd}}$, $\mathbf{P}_{\text{cd},1}$, $\mathbf{P}_{\text{cd},2}$ and $\mathbf{P}_{\text{cd},3}$ for the SISO case are computed from (8) - (12).

Note that, for identifying the optimized vector $\mathbf{w}_{\text{cd}}$, we approximate the exact equation (16) as (17) with the constraint in (22). In spite of this approximation, the modified $\mathbf{w}_{\text{cd}}$ increases the chordal distance in (16) compared to the original $\mathbf{w}$ in (12) since the value of the objective in (17) decreases as the desired signal space becomes orthogonal to the interference, statically. Also this leads to the improved sum rate performance, which will be shown in the simulation section.

B. Precoder optimization methods with two different decoding schemes

Once $\mathbf{P}_i$ is obtained from the signal space constraints, we now optimize its combination matrix with two different decoding schemes, i.e, ZF based and minimum mean-squared error (MMSE) based methods. We will describe the procedure of the methods for the MIMO case. This can be similarly applied to the SISO case.

1) ZF based decoder: First, consider the QR decomposition of $\mathbf{P}_i$ obtained from the IA method as

$$\mathbf{P}_i = \mathbf{Q}_i\mathbf{R}_i \quad \text{for} \quad i = 1, 2, 3$$

where $\mathbf{Q}_i$ is an $M \times \frac{M}{2}$ matrix whose columns form an orthonormal basis for $\mathbf{P}_i$ and $\mathbf{R}_i$ denotes an $\frac{M}{2} \times \frac{M}{2}$ upper triangular matrix. Then, the modified precoder can be formulated as

$$\mathbf{P}_{\text{zf},i} = \mathbf{Q}_i\mathbf{C}_{\text{zf},i} \tag{23}$$

where $\mathbf{C}_{\text{zf},i}$ represents an $\frac{M}{2} \times \frac{M}{2}$ square matrix which satisfies the transmit power constraint $\text{Tr}\,(\mathbf{C}_{\text{zf},i}^H\mathbf{C}_{\text{zf},i}) \le P_i$. Note that the modified precoding matrices $\mathbf{P}_{\text{zf},i}$ for $i = 1, 2, 3$ will not meet the restricting constraints in (4). Instead, these matrices will satisfy the equations in (3) for any matrix $\mathbf{C}_{\text{zf},i}$, and also this combination matrix does not affect the chordal distance in (13) since we have

$$\text{span}(\mathbf{H}_{ij}\mathbf{Q}_j) = \text{span}(\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{C}_{\text{zf},j}) \ \text{ for } \ i, j = 1, 2, 3. \tag{24}$$

In this way, as we relax the constraints in (4) and optimize the precoding matrix in two independent steps, we can identify the non-iterative scheme for interference channel systems.

Now, applying the precoder in (23) to (1), the $M \times 1$ received signal vector $\mathbf{y}_i$ of the ith receiver is written as

$$\mathbf{y}_i = \mathbf{H}_{\text{ef},i1}\mathbf{C}_{\text{zf},1}\mathbf{x}_1 + \mathbf{H}_{\text{ef},i2}\mathbf{C}_{\text{zf},2}\mathbf{x}_2 + \mathbf{H}_{\text{ef},i3}\mathbf{C}_{\text{zf},3}\mathbf{x}_3 + \mathbf{n}_i \tag{25}$$

where the $M \times \frac{M}{2}$ effective channel $\mathbf{H}_{\text{ef},ij}$ is denoted by $\mathbf{H}_{\text{ef},ij} = \mathbf{H}_{ij}\mathbf{Q}_j$. Here, all matrices $\mathbf{H}_{\text{ef},ij}$ for $j = 1, 2, 3$ and $j \neq i$ span the same space due to the interference aligning processing. Thus, we choose one of these matrices arbitrarily and denote the singular value decomposition (SVD) of this as

$$\mathbf{H}_{\text{ef},ij} = [\mathbf{U}_{ij}^{(1)}\ \ \mathbf{U}_{ij}^{(0)}][\mathbf{\Lambda}_{ij}\ \ \mathbf{O}]^T\mathbf{V}_{ij}^H \tag{26}$$

where the matrix $\mathbf{U}_{ij}^{(1)}$ is composed of the first $\frac{M}{2}$ left singular vectors, and the matrix $\mathbf{U}_{ij}^{(0)}$ holds the last $(M - \frac{M}{2})$ left singular vectors. Then, from (26), the $\frac{M}{2} \times M$ interference nulling matrix at the ith receiver which completely eliminates the interference becomes $\bar{\mathbf{M}}_{\text{zf},i} = \mathbf{U}_{ij}^{(0)H}$ [6]. Multiplying this to (25), the $\frac{M}{2} \times 1$ non-interfering received signal vector $\bar{\mathbf{x}}_i$ of the ith receiver can be written as

$$\bar{\mathbf{x}}_i = \bar{\mathbf{M}}_{\text{zf},i}\mathbf{y}_i = \mathbf{H}_{\text{zf},i}\mathbf{C}_{\text{zf},i}\mathbf{x}_i + \bar{\mathbf{M}}_{\text{zf},i}\mathbf{n}_i \tag{27}$$

where the $\frac{M}{2} \times \frac{M}{2}$ block channel matrix $\mathbf{H}_{\text{zf},i}$ is expressed as $\mathbf{H}_{\text{zf},i} = \bar{\mathbf{M}}_{\text{zf},i}\mathbf{H}_{\text{ef},ii}$.

Next, let us denote the SVD of $\mathbf{H}_{\text{zf},i}$ as $\mathbf{H}_{\text{zf},i} = \mathbf{U}_{\text{zf},i}\mathbf{\Lambda}_{\text{zf},i}\mathbf{V}_{\text{zf},i}^H$. Then, the information rate of the ith receiver can be computed as

$$\begin{aligned} R_i^{\text{zf}} &= \log_2\left|\mathbf{I} + \mathbf{H}_{\text{zf},i}\mathbf{C}_{\text{zf},i}\mathbf{C}_{\text{zf},i}^H\mathbf{H}_{\text{zf},i}^H\left(\sigma_n^2\bar{\mathbf{M}}_{\text{zf},i}\bar{\mathbf{M}}_{\text{zf},i}^H\right)^{-1}\right| \\ &= \log_2\left|\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{\Lambda}_{\text{zf},i}\mathbf{V}_{\text{zf},i}^H\mathbf{C}_{\text{zf},i}\mathbf{C}_{\text{zf},i}^H\mathbf{V}_{\text{zf},i}\mathbf{\Lambda}_{\text{zf},i}^H\right|. \end{aligned}$$

Thus, the optimal combination matrix $\mathbf{C}_{\text{zf},i}$ corresponding to the ZF based decoder is given by $\mathbf{C}_{\text{zf},i} = \mathbf{V}_{\text{zf},i}\mathbf{\Sigma}_i^{\frac{1}{2}}$ where the diagonal matrix $\mathbf{\Sigma}_i$ is calculated by using the water-filling solution [15] as

$$\max_{\mathbf{\Sigma}_i}\log_2\left|\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{\Lambda}_{\text{zf},i}^2\mathbf{\Sigma}_i\right| \quad \text{subject to } \text{Tr}(\mathbf{\Sigma}_i) \le P_i.$$

Also applying $\mathbf{U}_{\text{zf},i}^H$ to (27), the received signal streams can be decoupled into $M/2$ parallel streams for single-symbol detection.

Finally, the decoding and precoding matrices for this case are obtained as

$$\begin{aligned} \mathbf{M}_{\text{zf},i} &= \mathbf{U}_{\text{zf},i}^H\bar{\mathbf{M}}_{\text{zf},i}, \\ \mathbf{P}_{\text{zf},i} &= \mathbf{Q}_i\mathbf{V}_{\text{zf},i}\mathbf{\Sigma}_i^{\frac{1}{2}}. \end{aligned}$$

Also, the decoder output symbol vector $\hat{\mathbf{x}}_i$ of the ith receiver can be written as

$$\hat{\mathbf{x}}_i = \mathbf{M}_{\text{zf},i}\mathbf{y}_i = \mathbf{\Lambda}_{\text{zf},i}\mathbf{\Sigma}_i^{\frac{1}{2}}\mathbf{x}_i + \mathbf{M}_{\text{zf},i}\mathbf{n}_i.$$

2) MMSE based decoder: Now we consider the case of the optimal decoding method. For this case, denoting the $M \times \frac{M}{2}$ precoding matrix of the ith transmitter as $\mathbf{P}_{\text{ms},i} = \mathbf{Q}_i\mathbf{C}_{\text{ms},i}$ and the $\frac{M}{2} \times M$ decoding matrix of the ith receiver as $\mathbf{M}_{\text{ms},i}$, and applying these to (2), the corresponding decoder output signal vector $\hat{\mathbf{x}}_i$ of the ith receiver can be written as

$$\begin{aligned} \hat{\mathbf{x}}_i = \mathbf{M}_{\text{ms},i}\mathbf{y}_i &= \mathbf{M}_{\text{ms},i}\mathbf{H}_{\text{ef},ii}\mathbf{C}_{\text{ms},i}\mathbf{x}_i \\ &\quad + \mathbf{M}_{\text{ms},i}\sum_{j\neq i}^{3}\mathbf{H}_{\text{ef},ij}\mathbf{C}_{\text{ms},j}\mathbf{x}_j + \mathbf{M}_{\text{ms},i}\mathbf{n}_i. \end{aligned}$$

From the above equation, the SINR of each stream is expressed as equation (28) where $\mathbf{m}_{\text{ms},i,r}^T$ and $\mathbf{c}_{\text{ms},i,r}$ are the rth

$$\mathrm{SINR}_{i,r} = \frac{|\mathbf{m}_{\mathrm{ms},i,r}\mathbf{H}_{\mathrm{ef},ii}\,\mathbf{c}_{\mathrm{ms},i,r}|^2}{\|\mathbf{m}_{\mathrm{ms},i,r}\|^2\sigma_n^2 + \sum_{l\neq r}^{\frac{M}{2}}|\mathbf{m}_{\mathrm{ms},i,r}\mathbf{H}_{\mathrm{ef},ii}\,\mathbf{c}_{\mathrm{ms},i,l}|^2 + \sum_{j\neq i}^{3}\|\mathbf{m}_{\mathrm{ms},i,r}\mathbf{H}_{\mathrm{ef},ij}\,\mathbf{C}_{\mathrm{ms},j}\|^2} \tag{28}$$

column vectors of $\mathbf{M}_{\mathrm{ms},i}^T$ and $\mathbf{C}_{\mathrm{ms},i}$, respectively [16]. Then, unlike the ZF case which eliminates other users' interference completely, we identify the decoding matrix of each receiver from the following formulation

$$\max_{\mathbf{M}_{\mathrm{ms},i}} \sum_{r=1}^{M/2} \log_2\left(1+\mathrm{SINR}_{i,r}\right). \tag{29}$$

To define the solution of (29), we denote again $\mathbf{M}_{\mathrm{ms},i}$ as $\mathbf{M}_{\mathrm{ms},i} = \tilde{\mathbf{M}}_{\mathrm{ms},i}\bar{\mathbf{M}}_{\mathrm{ms},i}$. Then the $\frac{M}{2}\times M$ matrix $\bar{\mathbf{M}}_{\mathrm{ms},i}$ is given by

$$\bar{\mathbf{M}}_{\mathrm{ms},i} = (\mathbf{H}_{\mathrm{ef},ii}\mathbf{C}_{\mathrm{ms},i})^H \times\Big(\sum_{j\neq i}^{3}\mathbf{H}_{\mathrm{ef},ij}\mathbf{C}_{\mathrm{ms},j}\mathbf{C}_{\mathrm{ms},j}^H\mathbf{H}_{\mathrm{ef},ij}^H + \sigma_n^2\mathbf{I}\Big)^{-1}, \tag{30}$$

and the $\frac{M}{2}\times\frac{M}{2}$ square matrix $\tilde{\mathbf{M}}_{\mathrm{ms},i}$ is determined from the following conditions:

$$\tilde{\mathbf{M}}_{\mathrm{ms},i}\Big[\bar{\mathbf{M}}_{\mathrm{ms},i}\Big(\sum_{j\neq i}^{3}\mathbf{H}_{\mathrm{ef},ij}\mathbf{C}_{\mathrm{ms},j}\mathbf{C}_{\mathrm{ms},j}^H\mathbf{H}_{\mathrm{ef},ij}^H + \sigma_n^2\mathbf{I}\Big)\bar{\mathbf{M}}_{\mathrm{ms},i}^H\Big]\tilde{\mathbf{M}}_{\mathrm{ms},i}^H = \mathbf{I}_{\frac{M}{2}} \tag{31}$$

and

$$\tilde{\mathbf{M}}_{\mathrm{ms},i}\big(\bar{\mathbf{M}}_{\mathrm{ms},i}\mathbf{H}_{\mathrm{ef},ii}\mathbf{C}_{\mathrm{ms},i}\mathbf{C}_{\mathrm{ms},i}^H\mathbf{H}_{\mathrm{ef},ii}^H\bar{\mathbf{M}}_{\mathrm{ms},i}^H\big)\tilde{\mathbf{M}}_{\mathrm{ms},i}^H = \mathbf{D}_i \tag{32}$$

where $\mathbf{D}_i$ is some diagonal matrix.[1] Then, $\tilde{\mathbf{M}}_{\mathrm{ms},i}$ can be calculated using the Cholesky factorization and SVD operation.

For the case of the ZF based decoder, the optimal combination matrix $\mathbf{C}_{\mathrm{zf},i}$ is found using the SVD of the block channel $\mathbf{H}_{\mathrm{zf},i}$ in (27) and the power loading solution, since the interference nulling matrix $\bar{\mathbf{M}}_{\mathrm{zf},i}$ and $\mathbf{H}_{\mathrm{zf},i}$ are not a function of $\mathbf{C}_{\mathrm{zf},i}$. In contrast, in the MMSE case, the optimal $\mathbf{C}_{\mathrm{ms},i}$ cannot be computed in a closed-from solution since $\mathbf{M}_{\mathrm{ms},i}$ depends on the choice of the combination matrices, and then the joint optimization based on iterative water-filling [17] should be needed. To reduce the computational complexity, we determine $\mathbf{C}_{\mathrm{ms},i}$ with an equal power distribution, i.e., $\mathbf{C}_{\mathrm{ms},i} = \sqrt{p_i}\mathbf{X}$ where $\mathbf{X}$ is a unitary matrix and p_i represents the average power of each stream as $p_i = 2P_i/M$, as a sub-optimal strategy. However, this solution approaches an optimal one for high SNR.

In this case, $\mathbf{C}_{\mathrm{ms},j}\mathbf{C}_{\mathrm{ms},j}^H = p_j\,\mathbf{I}$ and, from (30), the interference suppressing matrix $\bar{\mathbf{M}}_{\mathrm{ms},i}$ is given by

$$\bar{\mathbf{M}}_{\mathrm{ms},i} = (\mathbf{H}_{\mathrm{ef},ii})^H\Big(\sum_{j\neq i}^{3}\frac{p_j}{p_i}\mathbf{H}_{\mathrm{ef},ij}\mathbf{H}_{\mathrm{ef},ij}^H + \frac{\sigma_n^2}{p_i}\mathbf{I}\Big)^{-1}. \tag{33}$$

[1]It can be verified that the MMSE based decoder in (30), (31) and (32) is a solution of (29) by checking that the individual sum rate of all streams after the decoder $\mathbf{M}_{\mathrm{ms},i}$ is equal to $\log_2\big|\mathbf{I} + \mathbf{H}_{\mathrm{ef},ii}\mathbf{C}_{\mathrm{ms},i}\mathbf{C}_{\mathrm{ms},i}^H\mathbf{H}_{\mathrm{ef},ii}^H\big(\sum_{j\neq i}^{3}\mathbf{H}_{\mathrm{ef},ij}\mathbf{C}_{\mathrm{ms},j}\mathbf{C}_{\mathrm{ms},j}^H\mathbf{H}_{\mathrm{ef},ij}^H + \sigma_n^2\mathbf{I}\big)^{-1}\big|$.

Also, from (31) and (32), $\tilde{\mathbf{M}}_{\mathrm{ms},i}$ is obtained from $\tilde{\mathbf{M}}_{\mathrm{ms},i} = \mathbf{U}_{\mathrm{ms},i}^H\mathbf{L}_{\mathrm{ms},i}^{-1}$ where $\mathbf{L}_{\mathrm{ms},i}$ is computed using the Cholesky factorization as

$$\bar{\mathbf{M}}_{\mathrm{ms},i}\Big(\sum_{j\neq i}^{3}p_j\mathbf{H}_{\mathrm{ef},ij}\mathbf{H}_{\mathrm{ef},ij}^H + \sigma_n^2\mathbf{I}\Big)\bar{\mathbf{M}}_{\mathrm{ms},i}^H = \mathbf{L}_{\mathrm{ms},i}\mathbf{L}_{\mathrm{ms},i}^H, \tag{34}$$

and $\mathbf{U}_{\mathrm{ms},i}$ is determined from the SVD of $\mathbf{H}_{\mathrm{ms},i}$ as $\mathbf{H}_{\mathrm{ms},i} = \mathbf{U}_{\mathrm{ms},i}\mathbf{\Lambda}_{\mathrm{ms},i}\mathbf{V}_{\mathrm{ms},i}^H$. Here the $\frac{M}{2}\times\frac{M}{2}$ block channel matrix $\mathbf{H}_{\mathrm{ms},i}$ is denoted by $\mathbf{H}_{\mathrm{ms},i} = \mathbf{L}_{\mathrm{ms},i}^{-1}\bar{\mathbf{M}}_{\mathrm{ms},i}\mathbf{H}_{\mathrm{ef},ii}$.

Then the final decoding and precoding matrices are given by

$$\begin{aligned}\mathbf{M}_{\mathrm{ms},i} &= \mathbf{U}_{\mathrm{ms},i}^H\mathbf{L}_{\mathrm{ms},i}^{-1}\bar{\mathbf{M}}_{\mathrm{ms},i},\\ \mathbf{P}_{\mathrm{ms},i} &= \sqrt{p_i}\mathbf{Q}_i\mathbf{V}_{\mathrm{ms},i}.\end{aligned}$$

The decoder output signal vector $\hat{\mathbf{x}}_i$ of the ith receiver can be written as

$$\hat{\mathbf{x}}_i = \sqrt{p_i}\mathbf{\Lambda}_{\mathrm{ms},i}\,\mathbf{x}_i + \mathbf{M}_{\mathrm{ms},i}\sum_{j\neq i}^{3}\mathbf{H}_{ij}\mathbf{P}_{\mathrm{ms},j}\,\mathbf{x}_j + \mathbf{M}_{\mathrm{ms},i}\,\mathbf{n}_i.$$

As shown in this section, the optimal combining matrix can be found only for the case of a sub-optimal decoding method. When an optimal receive filter is applied, even if we aim to optimize the combination matrix instead of the whole precoding matrix, a non-iterative approach would not be feasible. To address this concern, we will introduce an iterative approach in the following section.

V. Iterative Weighted Sum Rate Maximization Algorithm

In this section, we propose an iterative method for K-user interference channels. Unlike the non-iterative algorithm which obtains the precoder basis for satisfying the interference aligning constraints and only the combination matrix according to individual rate maximization criterion, the proposed algorithm identifies the whole precoding matrices iteratively using the weighted sum rate maximization criterion, which results in a local optimal solution. Note that, with a symbol extension as in Section III-B, the following solution can be applied to the SISO case as well as the MIMO case.

First, to identify the precoding matrices $\mathbf{P}_i$ for $i = 1,\cdots,K$, we formulate the weighted sum rate maximization problem on $\mathbf{P}_i$ as

$$\max_{\{\mathbf{P}_1\cdots\mathbf{P}_K\}}\sum_{i=1}^{K}\alpha_i\log_2\Big|\mathbf{I} + \mathbf{H}_{ii}\mathbf{P}_i\mathbf{P}_i^H\mathbf{H}_{ii}^H \times\Big(\sum_{j\neq i}^{K}\mathbf{H}_{ij}\mathbf{P}_j\mathbf{P}_j^H\mathbf{H}_{ij}^H + \sigma_n^2\mathbf{I}\Big)^{-1}\Big|$$

$$\text{subject to}\quad \mathrm{Tr}(\mathbf{P}_i^H\mathbf{P}_i)\le P_i\quad\text{for}\quad i = 1,\cdots,K \tag{35}$$

where α_i represents a weight factor of user i.[2] Since the above maximization problem is not a convex or concave problem with respect to $\mathbf{P}_i$ in general, it cannot be solved analytically. Hence we identify the optimal precoding matrix maximizing the weighted sum rate by deriving the gradient of the weighted sum rate and applying a gradient descent algorithm.

In order to exploit the gradient method, we first convert the problem in (35) into an unconstrained maximization problem. Let us denote $\mathbf{P}_i = \sqrt{\beta_i}\tilde{\mathbf{P}}_i$ where β_i is defined by

$$\beta_i = \frac{P_i}{\mathrm{Tr}(\tilde{\mathbf{P}}_i^H \tilde{\mathbf{P}}_i)} \quad \text{for} \quad i = 1, \cdots, K.$$

Then we substitute $\sqrt{\beta_i}\tilde{\mathbf{P}}_i$ for $\mathbf{P}_i$ in the cost function of (35). Now, the problem which computes the optimum precoding matrices for maximizing the weighted sum rate can be written as

$$\{\bar{\mathbf{P}}_{\mathrm{op},1}, \cdots, \bar{\mathbf{P}}_{\mathrm{op},K}\} = \underset{\{\mathbf{P}_1 \cdots \mathbf{P}_K\}}{\arg\max} \tilde{R}_{\text{wsum}} \tag{36}$$

where $\tilde{R}_{\text{wsum}}$ is defined as

$$\tilde{R}_{\text{wsum}} = \sum_{i=1}^{K} \alpha_i \log_2 \left| \mathbf{I} + \beta_i \mathbf{H}_{ii} \tilde{\mathbf{P}}_i \tilde{\mathbf{P}}_i^H \mathbf{H}_{ii}^H \times \Big(\sum_{j \neq i}^{K} \beta_j \mathbf{H}_{ij} \tilde{\mathbf{P}}_j \tilde{\mathbf{P}}_j^H \mathbf{H}_{ij}^H + \sigma_n^2 \mathbf{I} \Big)^{-1} \right|. \tag{37}$$

Since the weighted sum rate $\tilde{R}_{\text{wsum}}$ is a real valued function, we have $\nabla_{\tilde{\mathbf{P}}_k} \tilde{R}_{\text{wsum}} = 2\partial \tilde{R}_{\text{wsum}} / \partial \tilde{\mathbf{P}}_k^*$ derived as

$$\begin{aligned}
\nabla_{\tilde{\mathbf{P}}_k} \tilde{R}_{\text{wsum}} = {} & \frac{2\beta_k}{\ln 2} \sum_{i=1}^{K} \alpha_i \mathbf{H}_{ik}^H \tilde{\mathbf{\Phi}}_i^{-1} \mathbf{H}_{ik} \tilde{\mathbf{P}}_k \\
& - \frac{2\beta_k^2}{P_k \cdot \ln 2} \sum_{i=1}^{K} \alpha_i \mathrm{Tr}(\tilde{\mathbf{P}}_k^H \mathbf{H}_{ik}^H \tilde{\mathbf{\Phi}}_i^{-1} \mathbf{H}_{ik} \tilde{\mathbf{P}}_k) \tilde{\mathbf{P}}_k \\
& - \frac{2\beta_k}{\ln 2} \sum_{i \neq k}^{K} \alpha_i \mathbf{H}_{ik}^H \tilde{\mathbf{\Pi}}_i^{-1} \mathbf{H}_{ik} \tilde{\mathbf{P}}_k \\
& + \frac{2\beta_k^2}{P_k \cdot \ln 2} \sum_{i \neq k}^{K} \alpha_i \mathrm{Tr}(\tilde{\mathbf{P}}_k^H \mathbf{H}_{ik}^H \tilde{\mathbf{\Pi}}_i^{-1} \mathbf{H}_{ik} \tilde{\mathbf{P}}_k) \tilde{\mathbf{P}}_k
\end{aligned} \tag{38}$$

where the $N_{r_i} \times N_{r_i}$ matrices $\tilde{\mathbf{\Phi}}_i$ and $\tilde{\mathbf{\Pi}}_i$ are defined as

$$\tilde{\mathbf{\Phi}}_i = \sum_{j=1}^{K} \beta_j \mathbf{H}_{ij} \tilde{\mathbf{P}}_j \tilde{\mathbf{P}}_j^H \mathbf{H}_{ij}^H + \sigma_n^2 \mathbf{I}, \tag{39}$$

$$\tilde{\mathbf{\Pi}}_i = \sum_{j \neq i}^{K} \beta_j \mathbf{H}_{ij} \tilde{\mathbf{P}}_j \tilde{\mathbf{P}}_j^H \mathbf{H}_{ij}^H + \sigma_n^2 \mathbf{I}, \tag{40}$$

respectively. A more detailed derivation of (38) is presented in Appendix B.

With the derived gradient expression, we propose an iterative algorithm for solving (36) as follows:

-Initialization:

1) Initialize $\tilde{\mathbf{P}}_i(0)$ as an arbitrary $N_{t_i} \times N_{t_i}$ matrix for $i = 1, \cdots, K$
2) Calculate $\tilde{R}_{\text{wsum},0}$ with the initial matrices $\{\tilde{\mathbf{P}}_1(0), \cdots, \tilde{\mathbf{P}}_K(0)\}$ and set $l = 1$

-Iteration Loop:

3) for k = 1 to K
4) Calculate the gradient $\tilde{\mathbf{G}}_k = \nabla_{\tilde{\mathbf{P}}_k} \tilde{R}_{\text{wsum}}(\tilde{\mathbf{P}}_1(l), \cdots, \tilde{\mathbf{P}}_{k-1}(l), \tilde{\mathbf{P}}_k(l-1), \cdots, \tilde{\mathbf{P}}_K(l-1))$
5) Update $\tilde{\mathbf{P}}_k(l) \leftarrow \tilde{\mathbf{P}}_k(l-1) + \delta \cdot \tilde{\mathbf{G}}_k$
6) end
7) Calculate $\tilde{R}_{\text{wsum},l}$ with the updated precoding matrices $\{\tilde{\mathbf{P}}_1(l), \cdots, \tilde{\mathbf{P}}_K(l)\}$
8) If $|\tilde{R}_{\text{wsum},l} - \tilde{R}_{\text{wsum},l-1}| > \epsilon$, set $l \leftarrow l + 1$ and go back to step 3), otherwise stop the iteration

-Computation of $\bar{\mathbf{P}}_{\mathrm{op},i}$ for $i = 1, \cdots, K$:

9) Set $\tilde{\mathbf{P}}_{\mathrm{op},i} \leftarrow \sqrt{\beta_i}\tilde{\mathbf{P}}_i(l)$ where $\beta_i = P_i / \mathrm{Tr}(\tilde{\mathbf{P}}_i(l)^H \tilde{\mathbf{P}}_i(l))$
10) Set $\bar{\mathbf{P}}_{\mathrm{op},i} \leftarrow \tilde{\mathbf{U}}_i \tilde{\mathbf{\Lambda}}_i$ where $\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{\Lambda}}_i$ are obtained from the SVD of $\tilde{\mathbf{P}}_{\mathrm{op},i}$ as $\tilde{\mathbf{P}}_{\mathrm{op},i} = \tilde{\mathbf{U}}_i [\, \tilde{\mathbf{\Lambda}}_i \; \mathbf{0}] \tilde{\mathbf{V}}_i^H$

In this algorithm, ϵ is the tolerance factor for terminating the iteration and δ denotes the step size. Several line search methods are introduced in [19] to efficiently determine the step size δ. We employ a line search method called Armijo's Rule which provides provable convergence [19]. Also, the $N_{t_i} \times N_{t_i}$ matrix $\tilde{\mathbf{P}}_{\mathrm{op},i}$ calculated from the main loop could be a rank deficient matrix. Thus, we convert this matrix into the $N_{t_i} \times d_i$ matrix $\bar{\mathbf{P}}_{\mathrm{op},i}$ using the SVD of $\tilde{\mathbf{P}}_{\mathrm{op},i}$ where $d_i = \mathrm{rank}(\tilde{\mathbf{P}}_{\mathrm{op},i})$ represents the number of data streams allowed for user i.

The above proposed algorithm exploits a fact that the weighted sum rate increases fastest when the precoding matrices move in the direction of the gradient of the weighted sum rate. Although this algorithm cannot guarantee the global optimal solution due to non-convexity of the maximization problem in (35), a locally maximized weighted sum rate can be found.

After computing the precoding matrices $\bar{\mathbf{P}}_{\mathrm{op},i}$ for $i = 1, \cdots, K$, we now determine the decoding matrices with a similar procedure as in Section IV-B2. Then, the $d_i \times N_{r_i}$ interference suppressing matrix for receiver i is given by

$$\bar{\mathbf{M}}_{\mathrm{op},i} = (\mathbf{H}_{ii} \bar{\mathbf{P}}_{\mathrm{op},i})^H \Big(\sum_{j \neq i}^{K} \mathbf{H}_{ij} \bar{\mathbf{P}}_{\mathrm{op},j} (\mathbf{H}_{ij} \bar{\mathbf{P}}_{\mathrm{op},j})^H + \sigma_n^2 \mathbf{I} \Big)^{-1}.$$

Also, applying the Cholesky factorization as

$$\bar{\mathbf{M}}_{\mathrm{op},i} \Big(\sum_{j \neq i}^{K} \mathbf{H}_{ij} \bar{\mathbf{P}}_{\mathrm{op},j} (\mathbf{H}_{ij} \bar{\mathbf{P}}_{\mathrm{op},j})^H + \sigma_n^2 \mathbf{I} \Big) \bar{\mathbf{M}}_{\mathrm{op},i}^H = \mathbf{L}_{\mathrm{op},i} \mathbf{L}_{\mathrm{op},i}^H$$

and the SVD of $\mathbf{H}_{\mathrm{op},i}$ as $\mathbf{H}_{\mathrm{op},i} = \mathbf{U}_{\mathrm{op},i} \mathbf{\Lambda}_{\mathrm{op},i} \mathbf{V}_{\mathrm{op},i}^H$ where the $d_i \times d_i$ block channel matrix $\mathbf{H}_{\mathrm{op},i}$ is defined as $\mathbf{H}_{\mathrm{op},i} = \mathbf{L}_{\mathrm{op},i}^{-1} \bar{\mathbf{M}}_{\mathrm{op},i} \mathbf{H}_{ii} \bar{\mathbf{P}}_{\mathrm{op},i}$, the weighted sum rate maximizing precoding and decoding matrices are obtained as

$$\begin{aligned}
\mathbf{P}_{\mathrm{op},i} &= \bar{\mathbf{P}}_{\mathrm{op},i} \mathbf{V}_{\mathrm{op},i}, \\
\mathbf{M}_{\mathrm{op},i} &= \mathbf{U}_{\mathrm{op},i}^H \mathbf{L}_{\mathrm{op},i}^{-1} \bar{\mathbf{M}}_{\mathrm{op},i}.
\end{aligned} \tag{41}$$

Finally, applying the above solutions to (2), the decoder output

[2]The weight factor α_i can be chosen according to the state of the packet queues for a max-stability service or different Quality of Service demands of each user [18]. For instance, using equal weight, the problem in (35) can be interpreted as a conventional sum rate maximization. Since the study of a determining method of the weight factor is outside the scope of this paper, we assume that α_i is given in this paper.

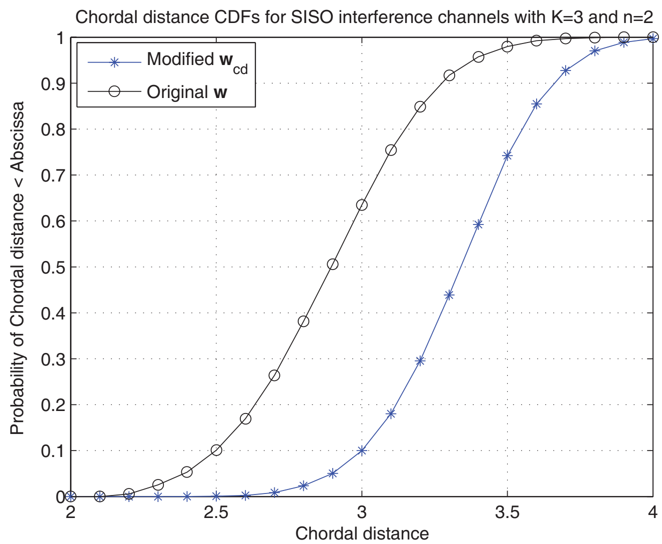

Fig. 2. CDFs of the chordal distance for 3-user SISO interference channel systems.

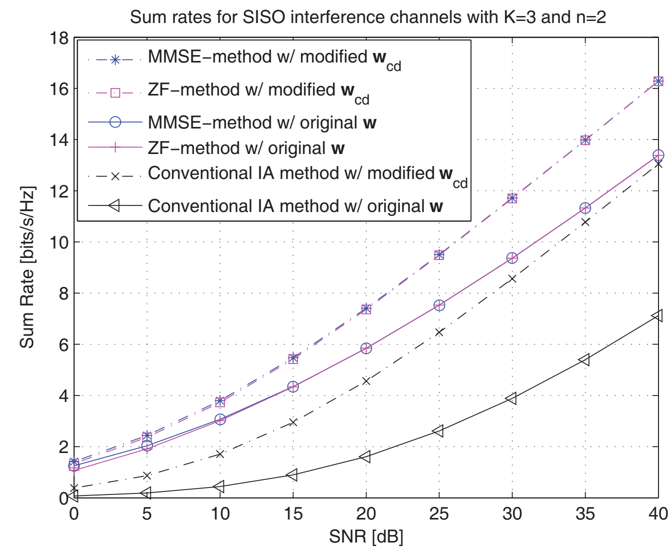

Fig. 3. Sum rates of non-iterative methods for 3-user SISO interference channel systems.

signal vector $\hat{\mathbf{x}}_i$ of the ith receiver can be written as

$$\hat{\mathbf{x}}_i = \mathbf{\Lambda}_{\text{op},i}\,\mathbf{x}_i + \mathbf{M}_{\text{op},i}\sum_{j\neq i}^{K}\mathbf{H}_{ij}\mathbf{P}_{\text{op},j}\,\mathbf{x}_j + \mathbf{M}_{\text{op},i}\,\mathbf{n}_i.$$

Also, since $\mathbf{M}_{\text{op},i}(\sum_{j\neq i}^{K}\mathbf{H}_{ij}\mathbf{P}_{\text{op},j}(\mathbf{H}_{ij}\mathbf{P}_{\text{op},j})^H + \sigma_n^2\mathbf{I})\mathbf{M}_{\text{op},i}^H = \mathbf{I}$ from (41), the weighted sum rate of the proposed iterative scheme is given by

$$R_{\text{wsum}}^{\text{op}} = \sum_{i=1}^{K}\sum_{r=1}^{d_i}\log_2\left(1+\lambda_{\text{op},i,r}^2\right)$$

where $\lambda_{\text{op},i,r}$ is the rth diagonal element of $\mathbf{\Lambda}_{\text{op},i}$.

Note that, when we compute $\bar{\mathbf{P}}_{\text{op},i}$ using the proposed iterative method, we utilize $\tilde{R}_{\text{wsum}}$ in (37) while the joint decoding approach is given on the receiver side. As shown in Section IV-B2, the decoding method in (41) does not change the maximized weighted sum rate obtained from (36) allowing a single-symbol detectable receiver. Hence, by exploiting the iterative gradient descent algorithm for precoder and applying the MMSE based decoder, we can identify the linear processing scheme which maximizes the weighted sum rate for interference channel systems.

VI. Numerical Results

This section evaluates the sum rate performance of various data transmission strategies over interference channel systems. In all simulations, we consider the case of $K = 3$, and assume that each transmitter has the same transmission power constraint P, i.e., $P_i = P$ for $i = 1, \cdots, K$. Then, the SNR is defined as $P/((2n+1)\sigma_n^2)$ and P/σ_n^2 for the SISO and MIMO case, respectively. Also, we assume that the elements of the channel matrix $\mathbf{H}_{ij}$ have an i.i.d. complex Gaussian distribution with zero mean and unit variance, and set n in Section III-B to be 2 for symbol extensions of the SISO case.

First, Fig. 2 presents the comparison of the cumulative distribution functions (CDFs) of the chordal distance (16) for the precoding methods with the modified $\mathbf{w}_{\text{cd}}$ in (21) and

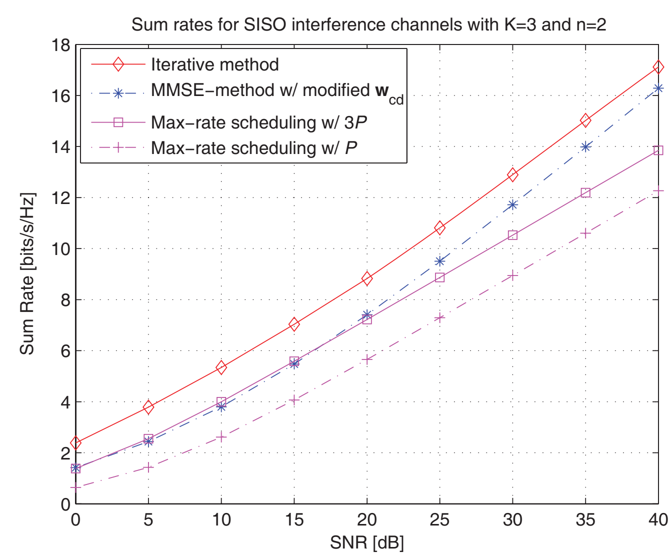

Fig. 4. Comparison of the sum rate as a function of SNR for 3-user SISO interference channel systems.

the original $\mathbf{w}$ in (12) for the SISO interference channels. From this plot, we confirm that the modified $\mathbf{w}_{\text{cd}}$ increases the distance between the desired signal space and the interference space compared to the original $\mathbf{w}$, even though there is an approximation for identifying the optimal $\mathbf{w}_{\text{cd}}$. This increase in the chordal distance leads to a sum rate improvement, which will be shown in the following.

In Figure 3, we illustrate the sum rates of various non-iterative methods as a function of SNR for the SISO interference channels with $n = 2$. This figure shows that the proposed methods achieve much better sum rate performance than the conventional IA algorithm. Also, from the observation on the slopes of the proposed schemes, it is clear that these methods achieve the maximum DOF, which is expected since the proposed non-iterative method fulfills the interference alignment constraints, too. In general, the

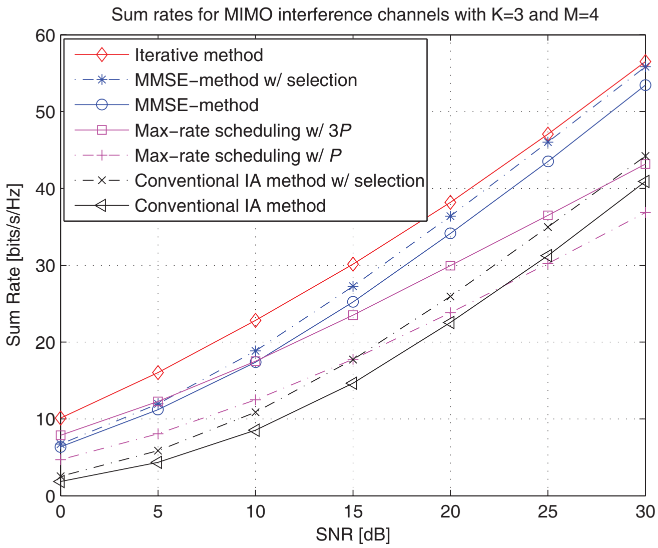

Fig. 5. Comparison of the sum rate as a function of SNR for 3-user MIMO interference channel systems with $M = 4$.

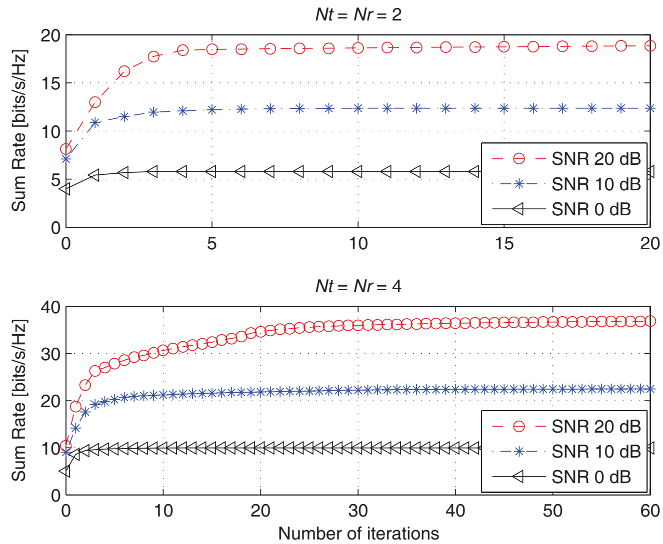

Fig. 6. Convergence of the iterative scheme for 3-user MIMO interference channels.

MMSE based method outperforms the ZF based scheme in low SNR region. However, for interference channel systems, these methods provide almost identical sum rates, since the ZF-method employs an optimal power loading solution while the sub-optimal equal power allocation is applied to the MMSE-method. Also, as predicted in Fig. 2, the schemes with the modified $\mathbf{w}_{\rm cd}$ outperform the methods with the original $\mathbf{w}$.

Fig. 4 is presented to address the sum rate benefits of the proposed algorithms over the traditional scheme such as orthogonalization in the SISO interference channels. To plot the sum rate of the orthogonalization approach, we utilize the max-rate scheduling method which selects only the best user among the K transmitter-receiver pairs in terms of the sum rate at the given channel condition. For a fair comparison, we plot sum rates of the scheduling method with the total power and user power constraints, i.e., the transmit powers of each user are $K \cdot P$ and P, respectively. Also, the sum rate of the iterative method in Section V is calculated using equal weights ($\alpha_i = 1$ for $i = 1, \cdots, K$) to maximize the sum rate corresponding to the best effort service. In this plot, we observe that the iterative algorithm outperforms other methods. As expected from the DOF, the performance gains of the proposed methods over the traditional scheme grow as SNR increases.

In Fig. 5, we consider MIMO interference channels where all transmitters and receivers have 4 antennas ($N_{t_i} = N_{r_i} = M = 4$ for $i = 1, 2, 3$). In the MIMO case, the sum rates of various algorithms exhibit a similar aspect to those of the SISO case. In this plot, it should be emphasized that the performance of the proposed non-iterative scheme with a selection based on maximum chordal distance criterion approaches the maximized sum rate result obtained by the iterative technique at high SNR. This demonstrates the efficiency of the proposed non-iterative algorithm which identifies the precoder basis vectors and its combination matrix separately.

Finally, in order to show the provable convergence of the proposed iterative method, we plot the sum rate of the iterative scheme in terms of the number of iterations for the MIMO interference channel systems in Fig. 6. As shown in this plot, the sum rate of the iterative method grows as the number of iteration increases, and is saturated after some iterations. Also, it can be seen that more iterations are required to converge as the number of antennas and SNR increase since the gap between the rate with random initial matrices and a local optimal point also grows. Note that, the convergence of this iterative algorithm could be improved significantly if it is initialized with the precoding matrices based on the non-iterative method.

VII. Conclusion

In this paper, we have studied linear precoder and decoder design methods for K-user interference channel systems. For a non-iterative approach, a two-stage optimization procedure has been utilized. Then, using the modified IA algorithm, the precoder basis vectors are designed to determine the signal spaces such that the maximum DOF and chordal distance are achieved, and the combination matrix is optimized based on the block interference suppressing receive filters to maximize the individual rate. To identify the weighted sum rate maximization method, an iterative technique has also been proposed by exploiting an iterative gradient descent algorithm and employing the MMSE based decoder. Through numerical simulations, we have demonstrated a local optimal sum rate performance of interference channels using the proposed iterative algorithm. Also, we have illustrated that the performance of the proposed non-iterative scheme approaches the maximized sum rate at high SNR with a low computational complexity.

Appendix A
Derivation of Equation (18)

From (8), (9) and (10), the $(2n+1) \times (n+1)$ matrix $\mathbf{P}_1$ is given by

$$\mathbf{P}_1 = \begin{bmatrix} \mathbf{w} & \mathbf{T}\mathbf{w} & \cdots & \mathbf{T}^n\mathbf{w} \end{bmatrix},$$

and the $(2n+1)\times n$ matrix $\mathbf{P}_2$ is written as

$$\mathbf{P}_2 \;=\; \mathbf{H}_{32}^{-1}\mathbf{H}_{31}\big[\mathbf{w} \quad \mathbf{T}\mathbf{w} \quad \cdots \quad \mathbf{T}^{(n-1)}\mathbf{w}\big].$$

Substituting these to the left hand side of equation (18), the $(n+1)\times n$ matrix $\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2$ can be expressed as

$$\begin{aligned}\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2 = \big[\mathbf{w} \;\; \mathbf{T}\mathbf{w} \;\; \cdots \;\; \mathbf{T}^n\mathbf{w}\big]^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\\ \times\big[\mathbf{w} \;\; \mathbf{T}\mathbf{w} \;\; \cdots \;\; \mathbf{T}^{(n-1)}\mathbf{w}\big].\end{aligned}$$

Here, we denote the (l,k) element of the matrix $\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2$ as $[\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2]_{(l,k)}$. Then, since $\mathbf{T}$ and $\mathbf{H}_{ij}$ are diagonal matrices, $[\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2]_{(l,k)}$ is obtained as

$$\begin{aligned}[\mathbf{P}_1^T\mathbf{H}_{11}^T\mathbf{H}_{12}\mathbf{P}_2]_{(l,k)} &= \mathbf{w}^T\mathbf{T}^{(l-1)}\mathbf{H}_{11}\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{T}^{(k-1)}\mathbf{w}\\ &= \mathbb{D}\big(\mathbf{T}^{(l-1)}\mathbf{H}_{11}\mathbf{H}_{12}\mathbf{H}_{32}^{-1}\mathbf{H}_{31}\mathbf{T}^{(k-1)}\big)\mathbf{w}_{\text{sq}}.\end{aligned}$$

Finally, by definition of the Frobenius norm, $\|(\mathbf{H}_{11}\mathbf{P}_1)^T\mathbf{H}_{12}\mathbf{P}_2\|_F^2$ can be rewritten as equation (18).

Appendix B
Derivation of the Weighted Sum Rate Gradient

From (39) and (40), the weighted sum rate (37) can be expressed as

$$\tilde{R}_{\text{wsum}} = \sum_{i=1}^{K}\alpha_i\log_2|\tilde{\mathbf{\Phi}}_i| - \alpha_i\log_2|\tilde{\mathbf{\Pi}}_i|. \tag{42}$$

Then, from $d(\ln|\mathbf{X}|) = \text{Tr}\{\mathbf{X}^{-1}d(\mathbf{X})\}$, the differential of the weighted sum rate for $\tilde{\mathbf{P}}_k^*$ is given by

$$\begin{aligned}d\tilde{R}_{\text{wsum}} = \frac{1}{\ln 2}\sum_{i=1}^{K}\alpha_i\text{Tr}\big(\tilde{\mathbf{\Phi}}_i^{-1}(d\beta_k\mathbf{H}_{ik}\tilde{\mathbf{P}}_k\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H\\ +\beta_k\mathbf{H}_{ik}\tilde{\mathbf{P}}_k d\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H)\big)\\ -\frac{1}{\ln 2}\sum_{i\neq k}^{K}\alpha_i\text{Tr}\big(\tilde{\mathbf{\Pi}}_i^{-1}(d\beta_k\mathbf{H}_{ik}\tilde{\mathbf{P}}_k\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H\\ +\beta_k\mathbf{H}_{ik}\tilde{\mathbf{P}}_k d\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H)\big).\end{aligned}$$

Also, using $d\{\text{Tr}(\mathbf{X})\} = \text{Tr}\{d(\mathbf{X})\}$, $\text{vec}(d\mathbf{X}) = d\text{vec}(\mathbf{X})$, $\text{Tr}(\mathbf{X}^T\mathbf{Y}) = \text{vec}(\mathbf{X})^T\text{vec}(\mathbf{Y})$ and $\text{Tr}(\mathbf{X}d\mathbf{Y}^H) = \text{Tr}(\mathbf{X}^Td\mathbf{Y}^*)$ [20], the above equation can be rewritten as

$$\begin{aligned}d\tilde{R}_{\text{wsum}} = \Bigg[\frac{\beta_k}{\ln 2}\sum_{i=1}^{K}\alpha_i\,\text{vec}(\mathbf{H}_{ik}^H\tilde{\mathbf{\Phi}}_i^{-1}\mathbf{H}_{ik}\tilde{\mathbf{P}}_k)^T\\ -\frac{\beta_k^2}{P_k\cdot\ln 2}\sum_{i=1}^{K}\alpha_i\text{Tr}(\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H\tilde{\mathbf{\Phi}}_i^{-1}\mathbf{H}_{ik}\tilde{\mathbf{P}}_k)\text{vec}(\tilde{\mathbf{P}}_k)^T\\ -\frac{\beta_k}{\ln 2}\sum_{i\neq k}^{K}\alpha_i\,\text{vec}(\mathbf{H}_{ik}^H\tilde{\mathbf{\Pi}}_i^{-1}\mathbf{H}_{ik}\tilde{\mathbf{P}}_k)^T\\ +\frac{\beta_k^2}{P_k\cdot\ln 2}\sum_{i\neq k}^{K}\alpha_i\text{Tr}(\tilde{\mathbf{P}}_k^H\mathbf{H}_{ik}^H\tilde{\mathbf{\Pi}}_i^{-1}\mathbf{H}_{ik}\tilde{\mathbf{P}}_k)\text{vec}(\tilde{\mathbf{P}}_k)^T\Bigg]d\text{vec}(\tilde{\mathbf{P}}_k^*).\end{aligned}$$

Here we have used $d\beta_k = -\beta_k^2\,\text{Tr}(\tilde{\mathbf{P}}_k d\tilde{\mathbf{P}}_k^H)/P_k$.

After all, the coefficients of $d\text{vec}(\tilde{\mathbf{P}}_k^*)$ in the above equation directly lead to the derivative $\partial\tilde{R}_{\text{wsum}}/\partial\tilde{\mathbf{P}}_k^*$, and the gradient of the weighted sum rate $\nabla_{\tilde{\mathbf{P}}_k}\tilde{R}_{\text{wsum}}$ is then derived as (38).

References

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

[2] I. Sason, “On the achievable rate regions for the Gaussian interference channel,” *IEEE Trans. Inf. Theory*, vol. 50, pp. 1345–1356, Jan. 2004.

[3] R. H. Etkin, D. N. C. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534–5562, Dec. 2008.

[4] V. Cadambe and S. Jafar, “Interference alignment and degrees of freedom of the K-user interference channel,” *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[5] S. Jafar and M. J. Fakherenddin, “Degrees of freedom for MIMO interference channel,” *IEEE Trans. Inf. Theory*, vol. 53, pp. 2637–2642, July 2007.

[6] Q. H. Spencer, A. L. Swindelhurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *IEEE Trans. Signal Process.*, vol. 52, pp. 461–471, Feb. 2004.

[7] H. Sung, S.-R. Lee, and I. Lee, “Generalized channel inversion methods for multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 57, Nov. 2009.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[9] M. Shen, A. Hϕst-Madson, and J. Vidal, “An improved interference alignment scheme for frequency selective channels,” in *Proc. IEEE ISIT*, pp. 559–563, July 2008.

[10] S. Ye and R. S. Blum, “Optimized signaling for MIMO interference systems with feedback,” *IEEE Trans. Signal Process.*, vol. 51, pp. 2839–2848, Nov. 2003.

[11] K. Gomadam, V. Cadambe, and S. Jafar, “Approaching the capacity of wireless networks through distributed interference alignment,” in *Proc. IEEE GLOBECOM*, Dec. 2008.

[12] M. H. M. Costa, “Writing on dirty paper,” *IEEE Trans. Inf. Theory*, vol. 29, pp. 439–441, May 1983.

[13] T. Gou and S. Jafar, “Degrees of freedom of the K user M x N MIMO interference channel,” http://arxiv.org/abs/0809.0099v1.

[14] D. J. Love and R. W. Heath, “Limited feedback unitary precoding for spatial multiplexing systems,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 2967–2976, Aug. 2005.

[15] J. M. Cioffi, “EE379A class note: signal processing and detection,” Stanford Univ.

[16] P. Viswanath and D. N. C. Tse, “Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[17] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, “Iterative water-filling for Gaussian vector multiple-access channels,” *IEEE Trans. Inf. Theory*, vol. 50, pp. 145–152, Jan. 2004.

[18] K. Seong, R. Narasimhan, and J. M. Cioffi, “Queue proportional scheduling via geometric programming in fading broadcast channels,” *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1593–1602, Aug. 2006.

[19] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley & Sons, 3rd ed., 2006.

[20] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, revised ed., 2002.

Hakjea Sung (S’06) received the B.S. and M.S. degrees in electrical engineering from Hongik University, Seoul, Korea, in 1998 and 2000, respectively. Currently, he is working toward the Ph.D. degree at Korea University, Seoul, Korea. Prior to joining Korea University, he has been with the Samsung Electronics, Suwon, Korea, as a research engineer in the mobile communication R&D group, since 2000. His research interests include signal processing techniques for MIMO-OFDM systems and multi-user MIMO wireless networks. He received the Best Paper Award at the IEEE Vehicular Technology Conference in 2009.

Seokhwan Park (S'05) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, Korea, in 2005 and 2007, where he is currently working toward the Ph.D. degree in the School of Electrical Engineering. His research interests include signal processing techniques for MIMO-OFDM systems. He has received the Best Paper Award at the 12th Asia-Pacific conference on Communications, and the IEEE Seoul Section Student Paper Contest.

Kyoung-Jae Lee (S'06) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, Korea, in 2005 and 2007, where he is currently working toward the Ph.D. degree in the School of Electrical Engineering. During the winter of 2006, he worked as an intern at Beceem Communications, Santa Clara, CA, USA. His research interests are in communication theory and signal processing for wireless communications, including MIMO-OFDM systems and wireless relay networks. He received the Gold Paper Award at the IEEE Seoul Section Student Paper Contest in 2007, and the Best Paper Award at the IEEE Vehicular Technology Conference in 2009.

Inkyu Lee (S'92-M'95-SM'01) received the B.S. degree (Hon.) in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1992 and 1995, respectively. From 1991 to 1995, he was a Research Assistant at the Information Systems Laboratory, Stanford University. From 1995 to 2001, he was a Member of Technical Staff at Bell Laboratories, Lucent Technologies, where he studied the high-speed wireless system design. He later worked for Agere Systems (formerly Microelectronics Group of Lucent Technologies), Murray Hill, NJ, as a Distinguished Member of Technical Staff from 2001 to 2002. In September 2002, he joined the faculty of Korea University, Seoul, Korea, where he is currently a Professor in the School of Electrical Engineering. He is now visiting University of Southern California, LA, USA, as a visiting Professor during 2009. He has published over 50 journal papers in IEEE, and has 30 U.S. patents granted or pending. His research interests include digital communications, signal processing, and coding techniques applied to wireless systems with an emphasis on MIMO-OFDM. Dr. Lee currently serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. Also, he has been a Chief Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on 4G Wireless Systems). He received the IT Young Engineer Award as the IEEE/IEEK joint award in 2006. Also he received the Best Paper Award for APCC in 2006 and VTC in 2009.