A Rapid Convergent Max-SINR Algorithm for Interference Alignment Based on Principle Direction Search

Zhilu Wu[1], Lihui Jiang[1], Guanghui Ren[1], Gangyi Wang[2], Nan Zhao[3], and Yaqin Zhao[1]
[1] School of Electronics and Information Engineering, Harbin Institute of Technology
Harbin, Heilongjiang 150001- P.R.China
[e-mail: wuzhilu@hit.edu.cn, jianglihui@hit.edu.cn, rgh@hit.edu.cn, yaqinzhao@hit.edu.cn]
[2] School of Instrumentation Science and Opto-electronics Engineering, Beihang University
Beijing 100191 - P.R.China
[e-mail: WangGangyi@buaa.edu.cn]
[3] School of Information and Communication Engineering, Dalian University of Technology
Dalian, Liaoning 116024 - P.R.China
[e-mail: zhaonan@dlut.edu.cn]
*Corresponding author: Zhilu Wu

Abstract

The maximal signal-to-interference-plus-noise ratio (Max-SINR) algorithm for interference alignment (IA) has received considerable attention for its high sum rate achievement in the multiple-input multiple-output (MIMO) interference channel. However, its complexity may increase dramatically when the number of users approaches the IA feasibility bound, and the number of iterations and computational time may become unacceptable. In this paper, we study the properties of the Max-SINR algorithm thoroughly by presenting theoretical insight into the algorithm and by providing the potential of reducing the overall computational cost. Furthermore, a novel IA algorithm based on the principle direction search is proposed, which can converge more rapidly than the conventional Max-SINR method. In the proposed algorithm, it searches along the principle direction, which is found to approximately point to the convergence values, and can approach the convergence solutions rapidly. In addition, the closed-form solution of the optimal step size can be formulated in the sense of minimal interference leakage. Simulation results demonstrate that the proposed algorithm outperforms the conventional minimal interference leakage and Max-SINR algorithms in terms of the convergence rate while guaranteeing the high throughput of IA networks.

1. Introduction

In current multiuser wireless networks where some resources such as frequency, time, and space are usually shared among different users, interference appears and becomes the bottleneck of the throughput of the system. Many scholars have focused on interference management techniques to eliminate the interference [1] and to maximize the sum rate [2], and interference alignment (IA) is a novel scheme that can mitigate the interference and thus increase the total throughput significantly [3]. Cadambe and Jafar showed that the degrees of freedom (DoFs) of the multiple-input multiple-output (MIMO) communication system scale proportionally to the number of users by employing IA in the high signal-to-noise ratio (SNR) regime in [4]. The essential idea of IA is to separate the desired signal and interference into different signal subspaces by using the proper precoding matrices and to reconstruct the former as well as to suppress the latter by using the proper decoding matrices [5]. Due to its promising performance in eliminating the interference and increasing the sum rate, IA has become an intensive research topic in recent years [6-8].

The feasibility condition for the *K*-user IA system of *M* transmitting antennas, *N* receiving antennas, and *d* DoFs was provided by Yetis *et al*. in [9] as $K \leq (M+N)/d$-1. However, the closed-form solutions to IA can be obtained only in some very special situations [10], and finding the optimal IA solutions still remains an open problem. Schmidt *et al*. analyzed these challenges in [11] and noted that the number of IA solutions increases dramatically with the number of users and antennas. In addition, the sum rate that different IA solutions can achieve varies considerably, and a brutal search is obviously infeasible due to the large computational cost. Furthermore, this problem was proved to be non-convex and NP-hard in [12]. Therefore, many scholars have developed different strategies to determine the sub-optimal IA solutions. Gomadam *et al*. developed the minimizing interference leakage (MinIL) algorithm in [13] to seek perfect IA points, and its convergence was proved. However, the MinIL algorithm does not consider the direct channel and might suffer from the loss of the desired signal power. Therefore, a maximal signal-to-interference-plus-noise ratio (Max-SINR) algorithm was proposed in [14], which aims at maximizing the SINR of each data stream by designing proper precoding and decoding matrices. The minimum mean square error (MMSE) criterion was introduced in [15] and [16] to optimize the sum rate in the broadcast and interference channels, respectively. Shrestha *et al*. developed the maximal signal-to-leakage-and-noise ratio (Max-SLNR) algorithm in [17], which has the advantages of best sum capacity in low and moderate SNR regimes, stable performance with respective to antenna order, and not requiring channel reciprocity. The gradient descent method was leveraged in [18] to maximize the sum rate. An excellent survey on the current IA techniques was conducted by Schmidt *et al*. in [11], where the objectives, convergence performance, and sum rate of different beamforming strategies are compared in great detail.

Among the current IA algorithms, the Max-SINR method can achieve a high sum rate and has received considerable attention [11, 19]. As noted in [11], the Max-SINR algorithm can achieve the maximum DoFs and can converge faster than many other algorithms such as MMSE. In addition, the throughput that the Max-SINR algorithm achieves is considerably higher than that of the MinIL algorithm in the low and medium SNR regimes. However, the iterations and computational time required by the Max-SINR algorithm might increase dramatically when the number of users *K* is close to the IA feasibility bound $(M+N)/d$-1, especially in the situations of larger *M*, *N*, and *K*. In practical systems, the precoding and decoding matrices should be obtained before the channel state information (CSI) changes, and

longer calculating times might cause a misalignment and loss of sum rate when the outdated precoding and decoding matrices are applied.

In this paper, we focus on developing the low-complexity Max-SINR algorithm with rapid convergence. The main contributions of this paper are summarized as follows: 1. The behavior of the Max-SINR algorithm in the neighborhood of the convergence solution is analyzed thoroughly, and it is found that the relationship of the consecutive iteration deviations can be approximately expressed by linear transformations. 2. The principle direction of the transformations is introduced, which can approximately point to the convergence solution. 3. A novel principle direction search (PDS) algorithm is proposed, which can increase the convergence rate and reduce the overall computational cost significantly.

The remaining parts of the paper are organized as follows. In Section 2, the system model is presented. The properties of the traditional Max-SINR algorithm are investigated in Section 3. The proposed PDS algorithm is developed in Section 4. In Section 5, simulation results are provided, followed by the conclusions in Section 6.

Notation: $\text{Re}\{\mathbf{A}\}$, $\text{Im}\{\mathbf{A}\}$, $\mathbf{A}^{\text{T}}$, $\mathbf{A}^{*}$, $\mathbf{A}^{\text{H}}$, $\text{Tr}[\mathbf{A}]$, $\|\mathbf{A}\|$, and $\mathbf{A}_{(l)}$ represent the real part, imaginary part, transpose, conjugate, conjugate transpose, trace, Frobenius norm, and the lth column of matrix $\mathbf{A}$, respectively. $\mathbf{A}_{k}^{[j]}$ denotes the value of matrix $\mathbf{A}_k$ at the jth iteration. Particularly, $\mathbf{A}_{k(l)}^{[j]}$ is the lth column of matrix $\mathbf{A}_k$ obtained at the jth iteration, and $\mathbf{A}_{k(l)}^{\text{H}}$ is the conjugate transpose of $\mathbf{A}_{k(l)}$. We use $\mathbb{R}$ and $\mathbb{C}$ to represent the sets of real and complex numbers, respectively. $\mathbf{I}$, $\mathbb{C}^{M \times N}$, and $\mathbb{CN}(\boldsymbol{\mu}, \boldsymbol{\Delta})$ denote the identity matrix, $M \times N$ complex matrix, and complex Gaussian vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Delta}$, respectively. The notation $(M \times N, d)^K$ is employed to represent an MIMO interference channel with K users, M transmitting antennas, N receiving antennas, and d data streams.

2. System Model

In this paper, we mainly concentrate on the $(M \times N, d)^K$ channels. For the kth user, the transmitted symbol vector $\mathbf{s}_k \in \mathbb{C}^{d \times 1}$ is first precoded by the transmitting beamforming matrix $\mathbf{P}_k \in \mathbb{C}^{M \times d}$, and the transmitted signal vector $\mathbf{x}_k \in \mathbb{C}^{M \times 1}$ can be obtained as:

$$\mathbf{x}_k = \mathbf{P}_k \mathbf{s}_k, k = 1, 2, ..., K. \tag{1}$$

The transmitted signal vectors of all users will be sent to the receivers, and the received signal vector of the kth user $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$ can be expressed as:

$$\mathbf{y}_k = \sum_{j=1}^{K} \mathbf{H}_{kj} \mathbf{x}_j + \mathbf{z}_k, k = 1, 2, ..., K \tag{2}$$

where $\mathbf{z}_k \in \mathbb{C}^{N \times 1}$ represents the received noise vector with $\mathbf{z}_k \sim \mathbb{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ and $\mathbf{H}_{kj} \in \mathbb{C}^{N \times M}$ denotes the channel coefficient matrix from transmitter j to receiver k. In this paper, perfect CSI is assumed to be known by each user, and the reader can refer to [20] and [21] for more information on CSI. The receiver k employs the decoding matrix $\mathbf{D}_k \in \mathbb{C}^{N \times d}$ to extract the desired symbol vector and to suppress the interference from other users. The final reconstructed symbol vector of the kth user can be attained as:

$$\hat{\mathbf{s}}_k = \mathbf{D}_k^{\text{H}} \mathbf{y}_k = \mathbf{D}_k^{\text{H}} \mathbf{H}_{kk} \mathbf{P}_k \mathbf{s}_k + \mathbf{D}_k^{\text{H}} \sum_{j=1, j \neq k}^{K} \mathbf{H}_{kj} \mathbf{P}_j \mathbf{s}_j + \mathbf{D}_k^{\text{H}} \mathbf{z}_k \tag{3}$$

where $\mathbf{D}_k^{\mathrm{H}}\mathbf{H}_{kk}\mathbf{P}_k\mathbf{s}_k$, $\mathbf{D}_k^{\mathrm{H}}\sum_{j=1,j\neq k}^{K}\mathbf{H}_{kj}\mathbf{P}_j\mathbf{s}_j$, and $\mathbf{D}_k^{\mathrm{H}}\mathbf{z}_k$ are the desired symbol vector, residual interference, and noise, respectively.

The interference in the network can be eliminated completely if the precoding and decoding matrices satisfy the following conditions:

$$\mathbf{D}_k^{\mathrm{H}}\mathbf{H}_{kj}\mathbf{P}_j=\mathbf{0};\ k,j\in\{1,2,...,K\};k\neq j \tag{4}$$

$$\mathrm{rank}(\mathbf{D}_k^{\mathrm{H}}\mathbf{H}_{kk}\mathbf{P}_k)=d \tag{5}$$

$$\mathbf{P}_k^{\mathrm{H}}\mathbf{P}_k=\mathbf{I},\mathbf{D}_k^{\mathrm{H}}\mathbf{D}_k=\mathbf{I}. \tag{6}$$

However, when perfect IA cannot be achieved, interference leakage appears, and it can be given by:

$$L=\sum_{k=1}^{K}\mathrm{Tr}[\mathbf{D}_k^{\mathrm{H}}(\sum_{j=1,j\neq k}^{K}\mathbf{H}_{kj}\mathbf{P}_j\mathbf{P}_j^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}})\mathbf{D}_k]. \tag{7}$$

In addition, the desired signal and the interference-plus-noise covariance matrices of the lth data stream of receiver k are defined, respectively, as [14]:

$$\mathbf{A}_{kl}=\mathbf{H}_{kk}\mathbf{P}_{k(l)}\mathbf{P}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}},l=1,2,...,d \tag{8}$$

$$\mathbf{B}_{kl}=\sum_{j=1,j\neq k}^{K}\mathbf{H}_{kj}\mathbf{P}_j\mathbf{P}_j^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}+\sum_{r=1,r\neq l}^{d}\mathbf{H}_{kk}\mathbf{P}_{k(r)}\mathbf{P}_{k(r)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}+\sigma^2\mathbf{I}. \tag{9}$$

Thus the SINR of the lth data stream of receiver k can be calculated as:

$$SINR_{kl}=(\mathbf{D}_{k(l)}^{\mathrm{H}}\mathbf{A}_{kl}\mathbf{D}_{k(l)})/(\mathbf{D}_{k(l)}^{\mathrm{H}}\mathbf{B}_{kl}\mathbf{D}_{k(l)}). \tag{10}$$

According to the Max-SINR algorithm [14], the decoding vector that maximizes (10) can be given by:

$$\mathbf{D}_{k(l)}=(\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)})/\|\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}\|. \tag{11}$$

Using the channel reciprocity, the channel coefficients matrices, precoding matrices, and decoding matrices in the reciprocal channel are defined, respectively, as:

$$\overleftarrow{\mathbf{H}}_{ij}=\mathbf{H}_{ji}^{\mathrm{H}},\ \overleftarrow{\mathbf{P}}_i=\mathbf{D}_i,\ \text{and}\ \overleftarrow{\mathbf{D}}_i=\mathbf{P}_i;\ i,j\in\{1,2,...,K\} \tag{12}$$

Similarly, the desired signal and the interference-plus-noise covariance matrices in the reciprocal channel, $i.e.$, $\overleftarrow{\mathbf{A}}_{kl}$ and $\overleftarrow{\mathbf{B}}_{kl}$, can be attained accordingly.

In this paper, we focus on increasing the convergence rate and reducing the computational complexity of the conventional Max-SINR algorithm in the high SNR regime. The properties of the Max-SINR algorithm are studied, and a rapid convergent IA algorithm is proposed.

3. Properties of the Traditional Max-SINR Algorithm

The Max-SINR algorithm is a promising method to approach the channel capacity, and the sum rate it achieves outperforms many other algorithms. In the low to medium SNR regimes where the sum rate is dominated by noise instead of interference, the Max-SINR algorithm can achieve a significantly higher sum rate than many other algorithms. However, in the high SNR regime, the throughput will be determined by interference rather than noise, and the Max-SINR algorithm has to align the interference properly with considerably more iterations and computational time. The large computational cost and long iteration time might constitute

a major impediment to its application in practical systems. Therefore, the properties of the Max-SINR algorithm in the high SNR regime will be studied in this section to provide the theoretical foundation of the later proposed rapid convergent algorithm.

3.1 The mechanism of the Max-SINR Algorithm

The Max-SINR algorithm, as shown in Algorithm 1, leverages the channel reciprocity and maximizes the SINR of each data stream alternatively. It is mainly composed of two procedures, *i.e.*, the forward and backward iterations. The most important point of the algorithm lies in the design of the precoding and decoding matrices. Gomadam *et al*. [14] did not provide the matrices design mechanism, and we present Theorem 1 to explain that the maximal SINR problem is essentially the optimization of the generalized Rayleigh quotient of matrices $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$.

Algorithm 1: Max-SINR Algorithm

---

1. Set the maximal iterations as $I_{\max}$ and the initial precoding matrices as $\mathbf{P}_k^{[0]}$ with $\|\mathbf{P}_{k(l)}^{[0]}\|=1$ ($k=1, 2, …, K; l=1, 2, …, d$). The variable i is used to represent the iterations and is initialized as $i=1$.

2. (Forward Iteration) Calculate the interference-plus-noise covariance matrices $\mathbf{B}_{kl}^{[i]}$ in the original channel according to (9), *i.e.,*

$$\mathbf{B}_{kl}^{[i]} = \sum_{j=1, j\neq k}^{K} \mathbf{H}_{kj}\mathbf{P}_j^{[i-1]}(\mathbf{P}_j^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}} + \sum_{r=1, r\neq l}^{d} \mathbf{H}_{kk}\mathbf{P}_{k(r)}^{[i-1]}(\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}} + \sigma^2\mathbf{I}. \quad (13)$$

Formulate the decoding matrix $\mathbf{D}_k^{[i]}$, the *l*th column of which can be attained according to (11), *i.e.*,

$$\mathbf{D}_{k(l)}^{[i]} = (\mathbf{B}_{kl}^{[i]})^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}^{[i-1]} / \|(\mathbf{B}_{kl}^{[i]})^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}^{[i-1]}\|. \quad (14)$$

3. (Backward Iteration) Calculate the interference-plus-noise covariance matrices $\tilde{\mathbf{B}}_{kl}^{[i]}$ in the reciprocal channel, *i.e.*,

$$\tilde{\mathbf{B}}_{kl}^{[i]} = \sum_{j=1, j\neq k}^{K} \mathbf{H}_{jk}^{\mathrm{H}}\mathbf{D}_j^{[i]}(\mathbf{D}_j^{[i]})^{\mathrm{H}}\mathbf{H}_{jk} + \sum_{r=1, r\neq l}^{d} \mathbf{H}_{kk}^{\mathrm{H}}\mathbf{D}_{k(r)}^{[i]}(\mathbf{D}_{k(r)}^{[i]})^{\mathrm{H}}\mathbf{H}_{kk} + \sigma^2\mathbf{I}. \quad (15)$$

Formulate the precoding matrix $\mathbf{P}_k^{[i]}$, the *l*th column of which can be expressed as:

$$\mathbf{P}_{k(l)}^{[i]} = (\tilde{\mathbf{B}}_{kl}^{[i]})^{-1}\mathbf{H}_{kk}^{\mathrm{H}}\mathbf{D}_{k(l)}^{[i]} / \|(\tilde{\mathbf{B}}_{kl}^{[i]})^{-1}\mathbf{H}_{kk}^{\mathrm{H}}\mathbf{D}_{k(l)}^{[i]}\|. \quad (16)$$

4. Update $i\leftarrow i+1$. If $i > I_{\max}$ or the algorithm converges, stop and output the results; otherwise, go to step 2).

Theorem 1: The unit-norm vector $\mathbf{D}_{k(l)}$ in (11) is the generalized eigenvector associated with the maximal generalized eigenvalue of matrices $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$. In particular, $SINR_{kl}$ in (10) is the maximal generalized eigenvalue. Similarly, this theorem holds for the reciprocal channel.
Proof: From (8) and (11), we have

$$\mathbf{A}_{kl}\mathbf{D}_{k(l)} = \mathbf{H}_{kk}\mathbf{P}_{k(l)}\mathbf{P}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}} \Box \frac{\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}}{\|\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}\|} = \frac{\mathbf{H}_{kk}\mathbf{P}_{k(l)}}{\|\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}\|}\mathbf{P}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\mathbf{B}_{kl}^{-1}\mathbf{H}_{kk}\mathbf{P}_{k(l)}. \quad (17)$$

Define

$$\theta_{kl} = \mathbf{P}_{k(l)}^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}} \mathbf{B}_{kl}^{-1} \mathbf{H}_{kk} \mathbf{P}_{k(l)} \in \mathbb{C} \tag{18}$$

and (17) can be rewritten as:

$$\mathbf{A}_{kl} \mathbf{D}_{k(l)} = \theta_{kl} \frac{\mathbf{H}_{kk} \mathbf{P}_{k(l)}}{\| \mathbf{B}_{kl}^{-1} \mathbf{H}_{kk} \mathbf{P}_{k(l)} \|} = \theta_{kl} \frac{\mathbf{B}_{kl} \mathbf{B}_{kl}^{-1} \mathbf{H}_{kk} \mathbf{P}_{k(l)}}{\| \mathbf{B}_{kl}^{-1} \mathbf{H}_{kk} \mathbf{P}_{k(l)} \|}. \tag{19}$$

Substitute (11) into (19), and the following equation can be satisfied:

$$\mathbf{A}_{kl} \mathbf{D}_{k(l)} = \theta_{kl} \mathbf{B}_{kl} \mathbf{D}_{k(l)}. \tag{20}$$

Therefore, θ_{kl} is the generalized eigenvalue of matrices $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$ with the corresponding generalized eigenvector of $\mathbf{D}_{k(l)}$. Particularly, θ_{kl} is the generalized Rayleigh quotient and equates to $SINR_{kl}$.

Left-multiply both sides of (20) by $\mathbf{B}_{kl}^{-1}$, and the generalized eigenvalue problem can be equivalent to the ordinary eigenvalue problem, *i.e.*,

$$\mathbf{B}_{kl}^{-1} \mathbf{A}_{kl} \mathbf{D}_{k(l)} = \theta_{kl} \mathbf{D}_{k(l)}. \tag{21}$$

Thus, θ_{kl} and $\mathbf{D}_{k(l)}$ are the eigenvalue and the associated eigenvector of matrix $\mathbf{B}_{kl}^{-1} \mathbf{A}_{kl}$, respectively. As $\mathbf{H}_{kk} \mathbf{P}_{k(l)} \in \mathbb{C}^{N \times 1}$, we have rank($\mathbf{H}_{kk} \mathbf{P}_{k(l)}$)=rank($\mathbf{H}_{kk} \mathbf{P}_{k(l)}$)$^{\mathrm{H}}$=1. Then,

$$\mathrm{rank}(\mathbf{A}_{kl}) \leq \min\{\mathrm{rank}(\mathbf{H}_{kk} \mathbf{P}_{k(l)}), \mathrm{rank}(\mathbf{H}_{kk} \mathbf{P}_{k(l)})^{\mathrm{H}}\} = 1. \tag{22}$$

Therefore, $\mathrm{rank}(\mathbf{A}_{kl}) = 1$. Because $\mathrm{rank}(\mathbf{B}_{kl}^{-1} \mathbf{A}_{kl}) = \mathrm{rank}(\mathbf{A}_{kl}) = 1$, $\mathbf{B}_{kl}^{-1} \mathbf{A}_{kl}$ has only one nonzero eigenvalue. Equivalently, there is only one nonzero generalized eigenvalue of matrices $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$. As θ_{kl} equates to $SINR_{kl}$, it can be shown that $\theta_{kl} > 0$. Therefore, θ_{kl} is the only nonzero generalized eigenvalue of $\mathbf{A}_{kl}$ and $\mathbf{B}_{kl}$. Obviously, θ_{kl} is the maximal Rayleigh quotient and the maximal SINR. Similarly, the proof above can be extended to the reciprocal channel to show that $\mathbf{P}_{k(l)}$ serves as the generalized eigenvector associated with the maximal generalized eigenvalue of matrices $\overline{\mathbf{A}}_{kl}$ and $\overline{\mathbf{B}}_{kl}$.

Theorem 1 provides a way for us to change the SINR optimization into a generalized eigenvalue problem, and some interesting properties can also be explored based on this theorem.

3.2 The Principle Direction of the Max-SINR Algorithm

In this section, we will develop the principle direction of the Max-SINR method, which will serve as the foundation of the latter proposed rapid convergent algorithm. Based on Theorem 1, we continue to develop Theorem 2 to show that the relationship for the consecutive iteration deviations of the Max-SINR algorithm can be approximately expressed by iteration-independent linear transformations. Properties of the transformations are later studied in Theorem 3 and Property 4. From these theorems, the principle direction of the transformations is then introduced in Theorem 5, and its mechanism for increasing the convergence rate is verified. Finally, because the principle direction cannot be attained directly, a convenient method to approximate the principle direction is developed in Theorem 6.

Before introducing Theorem 2, the following notations have to be defined beforehand. Define the convergence values of $\mathbf{P}_k^{[i]}$ and $\mathbf{D}_k^{[i]}$ as $\tilde{\mathbf{P}}_k$ and $\tilde{\mathbf{D}}_k$, respectively. The deviations from the ith iteration to the convergence value, *i.e.*, $\Delta \mathbf{P}_k^{[i]}$ and $\Delta \mathbf{D}_k^{[i]}$, satisfy

$$\mathbf{P}_k^{[i]} = \tilde{\mathbf{P}}_k + \Delta\mathbf{P}_k^{[i]}, k = 1, 2, ..., K \tag{23}$$

$$\mathbf{D}_k^{[i]} = \tilde{\mathbf{D}}_k + \Delta\mathbf{D}_k^{[i]}. \tag{24}$$

Next, several column vectors are constructed to represent the overall precoding and decoding vectors as follows:

$$\hat{\mathbf{p}}^{[i]} = \begin{bmatrix} \mathbf{P}_{1(1)}^{[i]} \\ \mathbf{P}_{1(2)}^{[i]} \\ ... \\ \mathbf{P}_{1(d)}^{[i]} \\ ... \\ \mathbf{P}_{K(1)}^{[i]} \\ \mathbf{P}_{K(2)}^{[i]} \\ ... \\ \mathbf{P}_{K(d)}^{[i]} \end{bmatrix}, \hat{\tilde{\mathbf{p}}} = \begin{bmatrix} \tilde{\mathbf{P}}_{1(1)} \\ \tilde{\mathbf{P}}_{1(2)} \\ ... \\ \tilde{\mathbf{P}}_{1(d)} \\ ... \\ \tilde{\mathbf{P}}_{K(1)} \\ \tilde{\mathbf{P}}_{K(2)} \\ ... \\ \tilde{\mathbf{P}}_{K(d)} \end{bmatrix}, \Delta\hat{\mathbf{p}}^{[i]} = \begin{bmatrix} \Delta\mathbf{P}_{1(1)}^{[i]} \\ \Delta\mathbf{P}_{1(2)}^{[i]} \\ ... \\ \Delta\mathbf{P}_{1(d)}^{[i]} \\ ... \\ \Delta\mathbf{P}_{K(1)}^{[i]} \\ \Delta\mathbf{P}_{K(2)}^{[i]} \\ ... \\ \Delta\mathbf{P}_{K(d)}^{[i]} \end{bmatrix}, \hat{\mathbf{d}}^{[i]} = \begin{bmatrix} \mathbf{D}_{1(1)}^{[i]} \\ \mathbf{D}_{1(2)}^{[i]} \\ ... \\ \mathbf{D}_{1(d)}^{[i]} \\ ... \\ \mathbf{D}_{K(1)}^{[i]} \\ \mathbf{D}_{K(2)}^{[i]} \\ ... \\ \mathbf{D}_{K(d)}^{[i]} \end{bmatrix}, \hat{\tilde{\mathbf{d}}} = \begin{bmatrix} \tilde{\mathbf{D}}_{1(1)} \\ \tilde{\mathbf{D}}_{1(2)} \\ ... \\ \tilde{\mathbf{D}}_{1(d)} \\ ... \\ \tilde{\mathbf{D}}_{K(1)} \\ \tilde{\mathbf{D}}_{K(2)} \\ ... \\ \tilde{\mathbf{D}}_{K(d)} \end{bmatrix}, \Delta\hat{\mathbf{d}}^{[i]} = \begin{bmatrix} \Delta\mathbf{D}_{1(1)}^{[i]} \\ \Delta\mathbf{D}_{1(2)}^{[i]} \\ ... \\ \Delta\mathbf{D}_{1(d)}^{[i]} \\ ... \\ \Delta\mathbf{D}_{K(1)}^{[i]} \\ \Delta\mathbf{D}_{K(2)}^{[i]} \\ ... \\ \Delta\mathbf{D}_{K(d)}^{[i]} \end{bmatrix}. \tag{25}$$

In addition, we can separate the real and imaginary parts of the vectors above to construct the associated real vectors, *i.e.*,

$$\bar{\mathbf{p}}^{[i]} = \begin{bmatrix} \text{Re}\{\hat{\mathbf{p}}^{[i]}\} \\ \text{Im}\{\hat{\mathbf{p}}^{[i]}\} \end{bmatrix}, \bar{\tilde{\mathbf{p}}} = \begin{bmatrix} \text{Re}\{\hat{\tilde{\mathbf{p}}}\} \\ \text{Im}\{\hat{\tilde{\mathbf{p}}}\} \end{bmatrix}, \Delta\bar{\mathbf{p}}^{[i]} = \begin{bmatrix} \text{Re}\{\Delta\hat{\mathbf{p}}^{[i]}\} \\ \text{Im}\{\Delta\hat{\mathbf{p}}^{[i]}\} \end{bmatrix} \in \mathbb{R}^{2KMd\times1} \tag{26}$$

$$\bar{\mathbf{d}}^{[i]} = \begin{bmatrix} \text{Re}\{\hat{\mathbf{d}}^{[i]}\} \\ \text{Im}\{\hat{\mathbf{d}}^{[i]}\} \end{bmatrix}, \bar{\tilde{\mathbf{d}}} = \begin{bmatrix} \text{Re}\{\hat{\tilde{\mathbf{d}}}\} \\ \text{Im}\{\hat{\tilde{\mathbf{d}}}\} \end{bmatrix}, \Delta\bar{\mathbf{d}}^{[i]} = \begin{bmatrix} \text{Re}\{\Delta\hat{\mathbf{d}}^{[i]}\} \\ \text{Im}\{\Delta\hat{\mathbf{d}}^{[i]}\} \end{bmatrix} \in \mathbb{R}^{2KNd\times1}. \tag{27}$$

It is obvious that the following relations can be satisfied:

$$\bar{\mathbf{p}}^{[i]} = \bar{\tilde{\mathbf{p}}} + \Delta\bar{\mathbf{p}}^{[i]} \tag{28}$$

$$\bar{\mathbf{d}}^{[i]} = \bar{\tilde{\mathbf{d}}} + \Delta\bar{\mathbf{d}}^{[i]}. \tag{29}$$

The change in the overall deviations $\Delta\bar{\mathbf{p}}^{[i]}$ and $\Delta\bar{\mathbf{d}}^{[i]}$ is studied, and the following theorem can be given:

Theorem 2: In the high SNR regime, if the Max-SINR algorithm can converge to a near IA point, there exists iteration-independent transformations $\mathbf{T}_P \in \mathbb{R}^{2MKd\times 2MKd}$ and $\mathbf{T}_D \in \mathbb{R}^{2NKd\times 2NKd}$, respectively, that exert on the deviation vectors $\Delta\bar{\mathbf{p}}^{[i]}$ and $\Delta\bar{\mathbf{d}}^{[i]}$. When $\bar{\mathbf{p}}^{[i]}$ and $\bar{\mathbf{d}}^{[i]}$ are located within the small neighborhoods of $\bar{\tilde{\mathbf{p}}}$ and $\bar{\tilde{\mathbf{d}}}$, respectively, the following approximation can be obtained:

$$\Delta\bar{\mathbf{p}}^{[i+1]} \approx \mathbf{T}_P \Delta\bar{\mathbf{p}}^{[i]} \tag{30}$$

$$\Delta\bar{\mathbf{d}}^{[i+1]} \approx \mathbf{T}_D \Delta\bar{\mathbf{d}}^{[i]}. \tag{31}$$

Proof: See Appendix A.

It is obvious that $\mathbf{T}_D$ and $\mathbf{T}_P$ determine the behavior of the Max-SINR algorithm in the neighborhoods around the convergence solutions. Therefore, a thorough analysis on the transformations will be executed in Theorem 3 and Property 4 to provide theoretical insight

into the algorithm and to explore the potential of increasing the convergence rate of the Max-SINR algorithm.

Theorem 3: The transformations $\mathbf{T}_D$ and $\mathbf{T}_P$ of the Max-SINR algorithm have the same eigenvalues.
Proof: For matrices $\mathbf{E} \in \mathbb{C}^{m \times n}$ and $\mathbf{F} \in \mathbb{C}^{n \times m}$, $\mathbf{EF}$ has the same eigenvalues as $\mathbf{FE}$ [22]. Therefore, $\mathbf{T}_D = \mathbf{T}_{PD} \mathbf{T}_{DP}$ and $\mathbf{T}_P = \mathbf{T}_{DP} \mathbf{T}_{PD}$ have the same eigenvalues.

In the $(M \times M, d)^K$ case, define the eigenvalues of $\mathbf{T}_D$ and $\mathbf{T}_P$ as $\{\lambda_1, \lambda_2, \ldots, \lambda_{2MKd}\}$ and the associated eigenvectors as $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{2MKd}\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{2MKd}\}$, respectively. For the high complexity of $\mathbf{T}_D$ and $\mathbf{T}_P$, their properties are mainly studied by simulations, and one important feature can be observed as follows:

Property 4: In the $(M \times M, d)^K$ channel, the eigenvalues and eigenvectors of $\mathbf{T}_D$ and $\mathbf{T}_P$ are real and satisfy:

$$\lambda_1 > \lambda_2 \geq \ldots \geq \lambda_{2MKd} \tag{32}$$

$$\text{rank}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{2MKd}) = 2MKd \tag{33}$$

$$\text{rank}(\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{2MKd}) = 2MKd \tag{34}$$

In particular, the last several eigenvalues tend to be zeros.

Here, we refer to the eigenvectors associated with the maximal eigenvalue, *i.e.*, $\mathbf{e}_1$ and $\mathbf{f}_1$, as the principle directions. When $\mathbf{T}_D$ and $\mathbf{T}_P$ are successively exerted on $\Delta\bar{\mathbf{d}}^{[i]}$ and $\Delta\bar{\mathbf{p}}^{[i]}$, respectively, the following theorem can be obtained.

Theorem 5: For the $(M \times M, d)^K$ channel in the high SNR regimes, if the Max-SINR algorithm can converge to a near IA point, when $\bar{\mathbf{d}}^{[i]}$ and $\bar{\mathbf{p}}^{[i]}$ are located within a small region of $\bar{\bar{\mathbf{d}}}$ and $\bar{\bar{\mathbf{p}}}$, they will converge approximately along the principle directions of $\mathbf{T}_D$ and $\mathbf{T}_P$, respectively.
Proof: From Property 4, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{2MKd}\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{2MKd}\}$ can serve as the bases of the $2MKd$ real space. Therefore, $\Delta\bar{\mathbf{d}}^{[i-1]}$ and $\Delta\bar{\mathbf{p}}^{[i-1]}$ can be expanded as:

$$\Delta\bar{\mathbf{d}}^{[i-1]} = \sum_{l=1}^{2MKd} g_l \mathbf{e}_l \tag{35}$$

$$\Delta\bar{\mathbf{p}}^{[i-1]} = \sum_{l=1}^{2MKd} h_l \mathbf{f}_l. \tag{36}$$

At the $(i+j)$th iteration, from (30), (31), (35), and (36), $\Delta\bar{\mathbf{d}}^{[i+j]}$ and $\Delta\bar{\mathbf{p}}^{[i+j]}$ can be given by:

$$\Delta\bar{\mathbf{d}}^{[i+j]} \approx \mathbf{T}_D^{j+1} \Delta\bar{\mathbf{d}}^{[i-1]} = \sum_{l=1}^{2MKd} g_l \mathbf{T}_D^{j+1} \mathbf{e}_l = \sum_{l=1}^{2MKd} g_l \lambda_l^{j+1} \mathbf{e}_l \tag{37}$$

$$\Delta\bar{\mathbf{p}}^{[i+j]} \approx \mathbf{T}_P^{j+1} \Delta\bar{\mathbf{p}}^{[i-1]} = \sum_{l=1}^{2MKd} h_l \mathbf{T}_P^{j+1} \mathbf{f}_l = \sum_{l=1}^{2MKd} h_l \lambda_l^{j+1} \mathbf{f}_l. \tag{38}$$

From (32), when j is sufficiently large, λ_1 will take the dominant position, *i.e.*,

$$\lambda_1^{j+1} >> \lambda_2^{j+1} >> \ldots >> \lambda_{2MKd}^{j+1}. \tag{39}$$

Then, $\Delta\bar{\mathbf{d}}^{[i+j]}$ and $\Delta\bar{\mathbf{p}}^{[i+j]}$ can be approximated by the principle directions as follows:

$$\Delta\overline{\mathbf{d}}^{[i+j]} \approx g_1 \lambda_1^{j+1} \mathbf{e}_1 \tag{40}$$

$$\Delta\overline{\mathbf{p}}^{[i+j]} \approx h_1 \lambda_1^{j+1} \mathbf{f}_1. \tag{41}$$

From (28), (29), (40), and (41), $\overline{\mathbf{d}}^{[i+j]}$ and $\overline{\mathbf{p}}^{[i+j]}$ can be written as:

$$\overline{\mathbf{d}}^{[i+j]} = \overline{\overline{\mathbf{d}}} + \Delta\overline{\mathbf{d}}^{[i+j]} \approx \overline{\overline{\mathbf{d}}} + g_1 \lambda_1^{j+1} \mathbf{e}_1 \tag{42}$$

$$\overline{\mathbf{p}}^{[i+j]} = \overline{\overline{\mathbf{p}}} + \Delta\overline{\mathbf{p}}^{[i+j]} \approx \overline{\overline{\mathbf{p}}} + h_1 \lambda_1^{j+1} \mathbf{f}_1. \tag{43}$$

Thus, $\overline{\mathbf{d}}^{[i+j]}$ and $\overline{\mathbf{p}}^{[i+j]}$ will converge to $\overline{\overline{\mathbf{d}}}$ and $\overline{\overline{\mathbf{p}}}$ approximately along the principle directions of $\mathbf{e}_1$ and $\mathbf{f}_1$, respectively.

If the principle directions can be obtained, we can search along these directions and go to the destination directly. However, the principle directions require knowledge of the convergence values, which is impossible to attain beforehand. Nevertheless, the following theorem provides a way for us to approximate the principle directions.

Theorem 6: In the high SNR regime, if the Max-SINR algorithm can converge to a near IA point, when $\overline{\mathbf{d}}^{[i]}$ and $\overline{\mathbf{p}}^{[i]}$ are located within a small region of $\overline{\overline{\mathbf{d}}}$ and $\overline{\overline{\mathbf{p}}}$, respectively, the principle directions can be approximated by the difference in the consecutive iteration results of the Max-SINR algorithm. In particular, the convergence solutions can be approximated as:

$$\tilde{\mathbf{D}}_k \approx \mathbf{D}_k^{[i]} + s(\mathbf{D}_k^{[i]} - \mathbf{D}_k^{[i-1]}), s \in \mathbb{R}, k=1,2,\ldots,K \tag{44}$$

$$\tilde{\mathbf{P}}_k \approx \mathbf{P}_k^{[i]} + s(\mathbf{P}_k^{[i]} - \mathbf{P}_k^{[i-1]}). \tag{45}$$

Proof: At the $(i+j+1)$th iteration, from (42) and (43), the decoding and precoding vectors can be expressed as:

$$\overline{\mathbf{d}}^{[i+j+1]} \approx \overline{\overline{\mathbf{d}}} + g_1 \lambda_1^{j+2} \mathbf{e}_1 \tag{46}$$

$$\overline{\mathbf{p}}^{[i+j+1]} \approx \overline{\overline{\mathbf{p}}} + h_1 \lambda_1^{j+2} \mathbf{f}_1. \tag{47}$$

Subtract (46) by (42) and (47) by (43), and we have

$$\overline{\mathbf{d}}^{[i+j+1]} - \overline{\mathbf{d}}^{[i+j]} \approx g_1 \lambda_1^{j+1} (\lambda_1 - 1) \mathbf{e}_1 \tag{48}$$

$$\overline{\mathbf{p}}^{[i+j+1]} - \overline{\mathbf{p}}^{[i+j]} \approx h_1 \lambda_1^{j+1} (\lambda_1 - 1) \mathbf{f}_1. \tag{49}$$

Therefore, the principle directions can be approximated by the difference in the consecutive iterations results. Substitute (48) and (49) into (46) and (47), respectively, and the convergence solutions can be approximated as:

$$\overline{\overline{\mathbf{d}}} \approx \overline{\mathbf{d}}^{[i+j+1]} - g_1 \lambda_1^{j+2} \mathbf{e}_1 \approx \overline{\mathbf{d}}^{[i+j+1]} - [\lambda_1 / (\lambda_1 - 1)](\overline{\mathbf{d}}^{[i+j+1]} - \overline{\mathbf{d}}^{[i+j]}) \tag{50}$$

$$\overline{\overline{\mathbf{p}}} \approx \overline{\mathbf{p}}^{[i+j+1]} - h_1 \lambda_1^{j+2} \mathbf{e}_1 \approx \overline{\mathbf{p}}^{[i+j+1]} - [\lambda_1 / (\lambda_1 - 1)](\overline{\mathbf{p}}^{[i+j+1]} - \overline{\mathbf{p}}^{[i+j]}). \tag{51}$$

Replace $i+j+1$ with i and $-\lambda_1/(\lambda_1-1)$ with s. Then, (50) and (51) can be expressed as:

$$\overline{\overline{\mathbf{d}}} \approx \overline{\mathbf{d}}^{[i]} + s(\overline{\mathbf{d}}^{[i]} - \overline{\mathbf{d}}^{[i-1]}) \tag{52}$$

$$\overline{\overline{\mathbf{p}}} \approx \overline{\mathbf{p}}^{[i]} + s(\overline{\mathbf{p}}^{[i]} - \overline{\mathbf{p}}^{[i-1]}). \tag{53}$$

Substitute (25) through (27) into (52) and (53), then (44) and (45) can be formulated.

Theorem 6 provides a direct way to approach the convergence solutions. It is expected that the iterations and computational time can be reduced by searching along the principle directions instead of the circuitous route of the Max-SINR algorithm. Based on the analysis in this section, the rapid convergent maximal SINR algorithm is straightforward.

4. The Principle Direction Search Algorithm

The original Max-SINR algorithm has satisfactory performance of a high sum rate. However, in the high SNR regime where the throughput is mainly limited by interference instead of noise, the Max-SINR algorithm will have to align the interference properly at the expense of a large amount of iterations and computational time. Particularly, the complexity will become significantly larger when the number of users is close to the IA feasibility bound ($M+N$)/d-1. Moreover, the complexity of the Max-SINR algorithm will increase with the number of users and antennas. These disadvantages might serve as a bottleneck for the application of the Max-SINR algorithm. In this section, a highly efficient maximal SINR algorithm, namely, a principle direction search algorithm, is proposed based on the theoretical analysis in Section 3 with the advantage of considerably less complexity than the traditional Max-SINR algorithm.

4.1 The Procedure of the PDS Algorithm

The procedure of the PDS algorithm is stated in Algorithm 2. As shown in Theorem 5, the principle direction search can be employed to increase the convergence rate of the Max-SINR algorithm when the current iteration results are located within a small neighborhood of the convergence solutions. In the proposed algorithm, PDS will be started after the intermediate sum rate is higher than the product of the parameter η and the initial sum rate R_1. The principle directions are attained by the difference in the intermediate results as shown in (44) and (45). In addition, PDS will be executed every ω iterations because some iterations have to be performed so that (39) can be satisfied and the principle direction will point to the convergence solution more accurately.

Algorithm 2: PDS Algorithm

1. Set the maximal iterations as $I_{\max}$ and the initial precoding matrices as $\mathbf{P}_k^{[0]}$ with $\| \mathbf{P}_{k(l)}^{[0]} \|=1$ (k=1, 2, …, K; l=1, 2, …, d). Initialize the parameters η and ω. The variable i is used to represent the iterations with an initial value of 1. The other auxiliary variables are initialized as *flag*=0 and i_s=0.

2. (Forward Iteration) Calculate $\mathbf{B}_{kl}^{[i]}$ in the original channel according to (13), and formulate the decoding matrices $\mathbf{D}_k^{[i]}$ according to (14).

3. (Backward Iteration) Calculate $\bar{\mathbf{B}}_{kl}^{[i]}$ in the reciprocal channel according to (15), and formulate the precoding matrices $\mathbf{P}_k^{[i]}$ according to (16).

4. If *flag*=1 and (i-i_s) mod ω=0, go to step 5) for PDS; otherwise, go to step 6).

5. (PDS) Update the precoding and decoding matrices as follows:
   a) Calculate the principle directions as:
$$\delta\mathbf{P}_k = \mathbf{P}_k^{[i]} - \mathbf{P}_k^{[i-1]}, k = 1, 2, ..., K \tag{54}$$
$$\delta\mathbf{D}_k = \mathbf{D}_k^{[i]} - \mathbf{D}_k^{[i-1]}. \tag{55}$$
   b) Calculate the step size s, and the corresponding details will be provided in "The Step Size Calculation" section.
   c) Calculate the new precoding and decoding matrices as:
$$\breve{\mathbf{P}}_k = \mathbf{P}_k^{[i]} + s\delta\mathbf{P}_k \tag{56}$$
$$\breve{\mathbf{D}}_k = \mathbf{D}_k^{[i]} + s\delta\mathbf{D}_k. \tag{57}$$

d) Normalize $\check{\mathbf{P}}_k$ and $\check{\mathbf{D}}_k$ so that $\| \check{\mathbf{P}}_{k(l)} \|=1$ and $\| \check{\mathbf{D}}_{k(l)} \|=1$, $l=1,\ 2,\ ...,\ d$.
Update $\mathbf{P}_k^{[i]} \leftarrow \check{\mathbf{P}}_k$, $\mathbf{D}_k^{[i]} \leftarrow \check{\mathbf{D}}_k$, $i \leftarrow i+1$, and go to step 2).

6. If *flag*=1, go to step 7); otherwise, compute the sum rate and record it as R_i. Set *flag*=1 and $i_s=i$ when $R_i>\eta R_1$. (Remark: *flag* and i_s are updated only once.)

7. Update $i \leftarrow i+1$. If $i > I_{\max}$ or the algorithm converges, stop and output the results; otherwise, go to step 2).

4.2 The Step Size Calculation

The step size calculation is one of the crucial procedures in the PDS algorithm and plays an important role in increasing the convergence rate. It is desirable that the optimal step size in the sense of maximal SINR or sum rate could be provided analytically. However, the closed-form solution of the optimal step size under this criterion remains unknown. Nevertheless, in the high SNR regime, the sum rate is mainly limited by the interference, and minimal interference leakage can be employed as the optimization objective. In addition, the optimal step size along the principle direction under the minimal interference leakage criterion can be formulated analytically. Therefore, instead of maximizing the sum rate, we minimize the interference leakage by choosing the proper step size.

As shown in (7), (54), (55), (56), and (57), when $\check{\mathbf{P}}_k$ and $\check{\mathbf{D}}_k$ are applied, the interference leakage can be calculated as:

$$
\begin{aligned}
L &= \sum_{k=1}^{K} \mathrm{Tr}[\check{\mathbf{D}}_k^{\mathrm{H}} (\sum_{l=1,l\neq k}^{K} \mathbf{H}_{kl}\check{\mathbf{P}}_l\check{\mathbf{P}}_l^{\mathrm{H}}\mathbf{H}_{kl}^{\mathrm{H}})\check{\mathbf{D}}_k] \\
&= \sum_{k=1}^{K} \mathrm{Tr}\{(\mathbf{D}_k^{[i]} + s\delta\mathbf{D}_k)^{\mathrm{H}}[\sum_{l=1,l\neq k}^{K} \mathbf{H}_{kl}(\mathbf{P}_l^{[i]} + s\delta\mathbf{P}_l)(\mathbf{P}_l^{[i]} + s\delta\mathbf{P}_l)^{\mathrm{H}}\mathbf{H}_{kl}^{\mathrm{H}}](\mathbf{D}_k^{[i]} + s\delta\mathbf{D}_k)\} \\
&= \beta_4 s^4 + \beta_3 s^3 + \beta_2 s^2 + \beta_1 s + \beta_0
\end{aligned} \tag{58}
$$

where

$$
\beta_4 = \sum_{k=1}^{K} \sum_{l=1,l\neq k}^{K} \mathrm{Tr}[\mathbf{X}_{kl}\mathbf{X}_{kl}^{\mathrm{H}}] \tag{59}
$$

$$
\beta_3 = \sum_{k=1}^{K} \sum_{l=1,l\neq k}^{K} 2\mathrm{Tr}[\mathrm{Re}(\mathbf{V}_{kl}\mathbf{X}_{kl}^{\mathrm{H}} + \mathbf{W}_{kl}\mathbf{X}_{kl}^{\mathrm{H}})] \tag{60}
$$

$$
\beta_2 = \sum_{k=1}^{K} \sum_{l=1,l\neq k}^{K} \mathrm{Tr}[\mathbf{V}_{kl}\mathbf{V}_{kl}^{\mathrm{H}} + \mathbf{W}_{kl}\mathbf{W}_{kl}^{\mathrm{H}} + 2\mathrm{Re}(\mathbf{V}_{kl}\mathbf{W}_{kl}^{\mathrm{H}} + \mathbf{X}_{kl}\mathbf{U}_{kl}^{\mathrm{H}})] \tag{61}
$$

$$
\beta_1 = \sum_{k=1}^{K} \sum_{l=1,l\neq k}^{K} \mathrm{Tr}[2\mathrm{Re}(\mathbf{U}_{kl}\mathbf{V}_{kl}^{\mathrm{H}} + \mathbf{U}_{kl}\mathbf{W}_{kl}^{\mathrm{H}})] \tag{62}
$$

$$
\beta_0 = \sum_{k=1}^{K} \sum_{l=1,l\neq k}^{K} \mathrm{Tr}[\mathbf{U}_{kl}\mathbf{U}_{kl}^{\mathrm{H}}] \tag{63}
$$

$$
\mathbf{U}_{kl} = [(\mathbf{D}_k^{[i]})^{\mathrm{H}} \mathbf{H}_{kl}]\mathbf{P}_l^{[i]} \tag{64}
$$

$$
\mathbf{V}_{kl} = [(\mathbf{D}_k^{[i]})^{\mathrm{H}}\mathbf{H}_{kl} - (\mathbf{D}_k^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kl}]\mathbf{P}_l^{[i]} \tag{65}
$$

$$
\mathbf{W}_{kl} = [(\mathbf{D}_k^{i})^{\mathrm{H}}\mathbf{H}_{kl}](\mathbf{P}_l^{[i]} - \mathbf{P}_l^{[i-1]}) \tag{66}
$$

$$\mathbf{X}_{kl} = [(\mathbf{D}_k^{[i]})^{\mathrm{H}} \mathbf{H}_{kl} - (\mathbf{D}_k^{[i-1]})^{\mathrm{H}} \mathbf{H}_{kl}](\mathbf{P}_l^{[i]} - \mathbf{P}_l^{[i-1]}). \tag{67}$$

It is obvious that $\beta_4 > 0$, and the quartic function (58) must have the global minimal value. Let the derivatives of (58) be zero, *i.e.*,

$$dL / ds = 4\beta_4 s^3 + 3\beta_3 s^2 + 2\beta_2 s + \beta_1 = 0. \tag{68}$$

The cubic equation (68) has three real solutions at most, and the optimal step size is chosen as the real solution that has the smallest function value.

5. Simulation Results

In this section, we evaluate the performances of the proposed PDS algorithm. We assume narrowband frequency non-selective fading channels. Except for the performance evaluation in **Fig. 2**, all of the other experiments are averaged over 250 channel realizations. Each element of the channel coefficients is independent and complies with the $\mathbb{CN}(0,1)$ distribution. The parameters of the PDS algorithm are first analyzed and selected through simulations in section 5.1. Then, a comparison of the PDS, MinIL, and Max-SINR algorithms is executed in section 5.2.

5.1 Parameter Analysis

As shown in Algorithm 2, the parameters η and ω can affect the convergence rate of the PDS algorithm. The closed-form solutions of the optimal parameters remain unknown, and simulations are employed to determine them in **Fig. 1**.

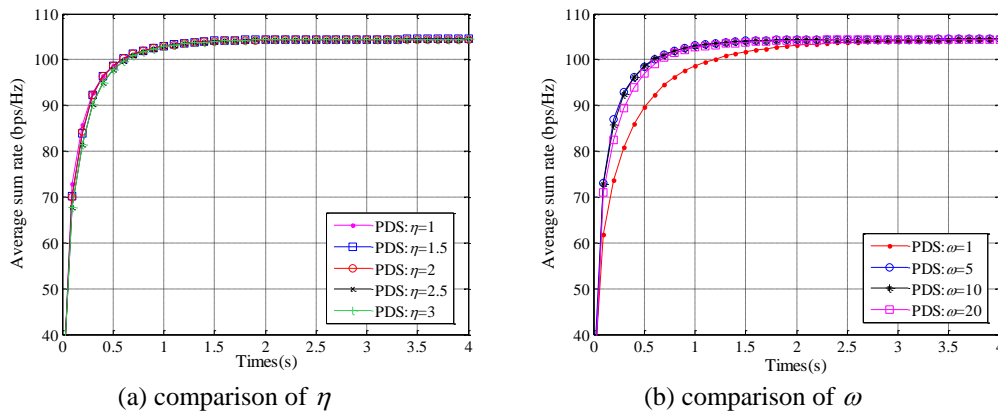(a) comparison of η (b) comparison of ω

Fig. 1. Convergence of the average sum rate of the PDS algorithm in $(5 \times 5, 1)^9$ under different parameters with SNR=30 dB

The parameter η determines when the PDS procedure starts. As shown in Theorem 5, the principle direction will approximately point to the convergence solution only when the intermediate iteration results are within a small region of the convergence point. Because the convergence solution is not available beforehand, it is impossible to measure the distance between the intermediate and the convergence points directly. Nevertheless, the sum rate can be leveraged as one type of measure of the distance because it increases as the algorithm approaches the destination. Therefore, we start PDS when the current sum rate is larger than the product of η and R_1. In this experiment, η is selected as 1, 1.5, 2, 2.5, and 3, and ω is preset to be 10. The convergence of the average sum rate with SNR=30 dB in $(5 \times 5, 1)^9$ is depicted in

Fig. 1 (a). As shown in the figure, the convergence curves with different η nearly overlap, indicating that the algorithm is insensitive to this parameter. In the simulations afterward, the parameter is chosen as η=1.5.

The parameter ω controls the interval for executing PDS. As shown in (39), (40), and (41), the maximal eigenvalue and the principle direction will take the dominating position only after several iterations are executed. Therefore, PDS is performed every ω iterations. In this experiment, ω is selected as 1, 5, 10, and 20, and η is preset to be 1.5. The convergence of the average sum rate with SNR=30 dB in $(5\times5, 1)^9$ is depicted in **Fig. 1** (b). As shown in the figure, ω=1 has the worst performance because the non-principle eigenvectors in (37) and (38) cannot be ignored and the principle direction cannot point to the convergence solution properly when only one iteration is performed. On the other hand, if ω is set to be too large, as in the ω=20 case, the PDS algorithm will fail to increase the convergence rate in time, and the performance of ω=20 is inferior to the performances of ω=5 and ω=10. Because ω=5 has the best convergence rate, it will be chosen as the selected interval afterward.

5.2 Comparison of Different Algorithms

In this section, the PDS, MinIL, and Max-SINR algorithms are compared. To gain an overview on the efficiency of the PDS algorithm, the convergence curves of the interference leakage and the sum rate during the first several iterations are provided in **Fig. 2** for a randomly generated channel realization. In contrast to the smooth and slow variance on the interference leakage and the sum rate of the MinIL and Max-SINR algorithms, there is an instant decrease in interference leakage and an instant increase in sum rate after PDS is applied. Therefore, the PDS algorithm can suppress the interference leakage and increase the sum rate more effectively compared with the MinIL and Max-SINR algorithms.

(a) interference leakage evaluation   (b) sum rate evaluation
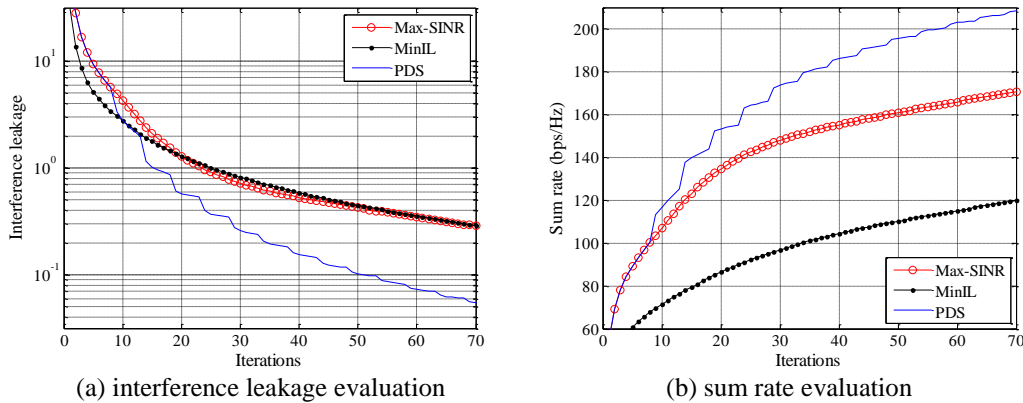
Fig. 2. Convergence of the interference leakage and sum rate of the PDS, MinIL, and Max-SINR algorithms for a randomly generated channel realization in $(10\times10, 1)^{19}$ with SNR=30 dB

To further compare the convergence rate of the algorithms, the average interference leakage and the sum rate with respect to the computational time for $(5\times5, 1)^9$ are depicted in **Fig. 3**. As shown in the figures, the average interference leakage of the MinIL and Max-SINR algorithms decreases in the same pattern during the first several seconds and reaches the level of 3×10^{-3} at 2 s. In contrast, the PDS algorithm has a higher convergence rate and can suppress the interference leakage to the level of 2×10^{-4} at 2 s. Both the PDS and Max-SINR algorithms converge at the same level of interference leakage, while the former algorithm takes only 3 s compared with the 20 s required by the latter algorithm. Although the MinIL algorithm can

achieve a lower level of interference leakage after 8.5 s, its sum rate is considerably inferior to those of the PDS and Max-SINR algorithms. The average sum rate of the PDS algorithm converges at approximately 2 s with a convergence value of 104 bps/Hz, while the Max-SINR algorithm takes as long as 9 s to reach the same sum rate.

The proposed algorithm can be applied not only to the single data stream channel but also to the multi-data stream scenarios. The convergence of the average sum rate of the algorithms is compared in the $(10\times10, 2)^9$ and $(10\times10, 4)^4$ channels in **Fig. 4**. As shown in **Fig. 4** (a), the Max-SINR algorithm requires as much as 19.5 s to reach the sum rate of 190 bps/Hz, while the PDS algorithm only takes 6.5 s to reach the same sum rate. For the scenario of even more data streams as shown in **Fig. 4** (b), the PDS algorithm can achieve the highest sum rate and exhibit considerably faster convergence rate than the Max-SINR algorithm.

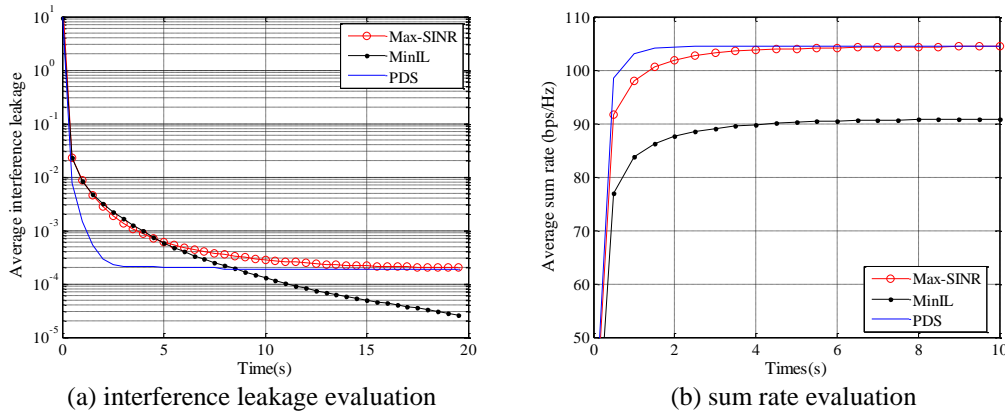(a) interference leakage evaluation   (b) sum rate evaluation

Fig. 3. Convergence of the average interference leakage and sum rate of the PDS, MinIL, and Max-SINR algorithms in $(5\times5, 1)^9$ with SNR=30 dB
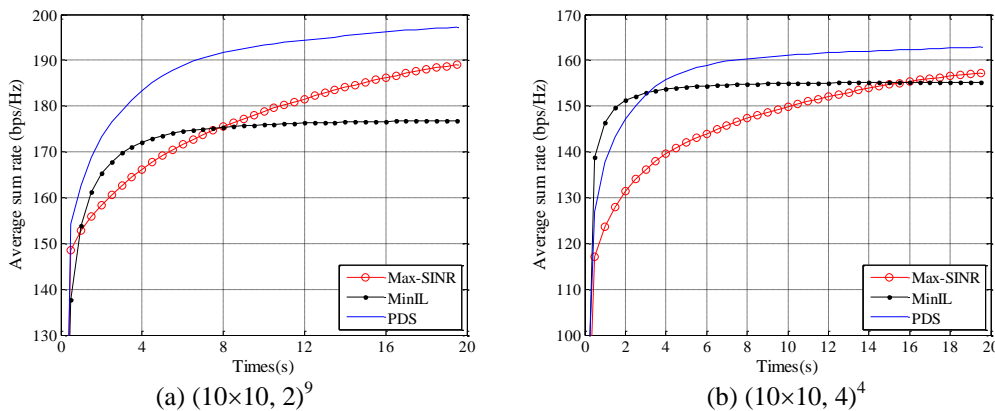

(a) $(10\times10, 2)^9$ (b) $(10\times10, 4)^4$

Fig. 4. Convergence of the average sum rate of the PDS, MinIL, and Max-SINR algorithms in $(10\times10, 2)^9$ and $(10\times10, 4)^4$ with SNR=30 dB

Although we are mainly concerned with the high SNR regime, it is desirable to evaluate the sum rate in terms of different SNRs. As shown in **Fig. 5** (a), the average sum rates of the PDS, MinIL, and Max-SINR algorithms with respect to SNR are compared under sufficient iterations of 10,000. It can be seen that both the PDS and Max-SINR algorithms can achieve the same sum rate, which is higher than those of the MinIL algorithm from the low to high

SNR regimes. However, these sufficient iterations are nearly impossible for practical applications, and limited iterations are often required. As revealed in **Fig. 5** (b), under 100 iterations, both the PDS and Max-SINR algorithms can achieve a higher sum rate than the MinIL algorithm for all SNRs. In the low SNR regime, the PDS and Max-SINR algorithms share the same sum rate. However, the former algorithm can achieve a higher sum rate in the high SNR regime. In addition, the difference in the sum rate between the PDS and Max-SINR algorithms increases with an increase in SNR, from a difference of 2 bps/Hz at SNR=20 dB to 23 bps/Hz at SNR=40 dB. When the number of iterations allowed is even smaller as shown in **Fig. 5** (c), the advantage of the PDS algorithm over the MinIL and Max-SINR algorithms is even more significant. Therefore, the proposed PDS algorithm has practical significance in that it can achieve a higher sum rate than the Max-SINR algorithm under limited iterations.
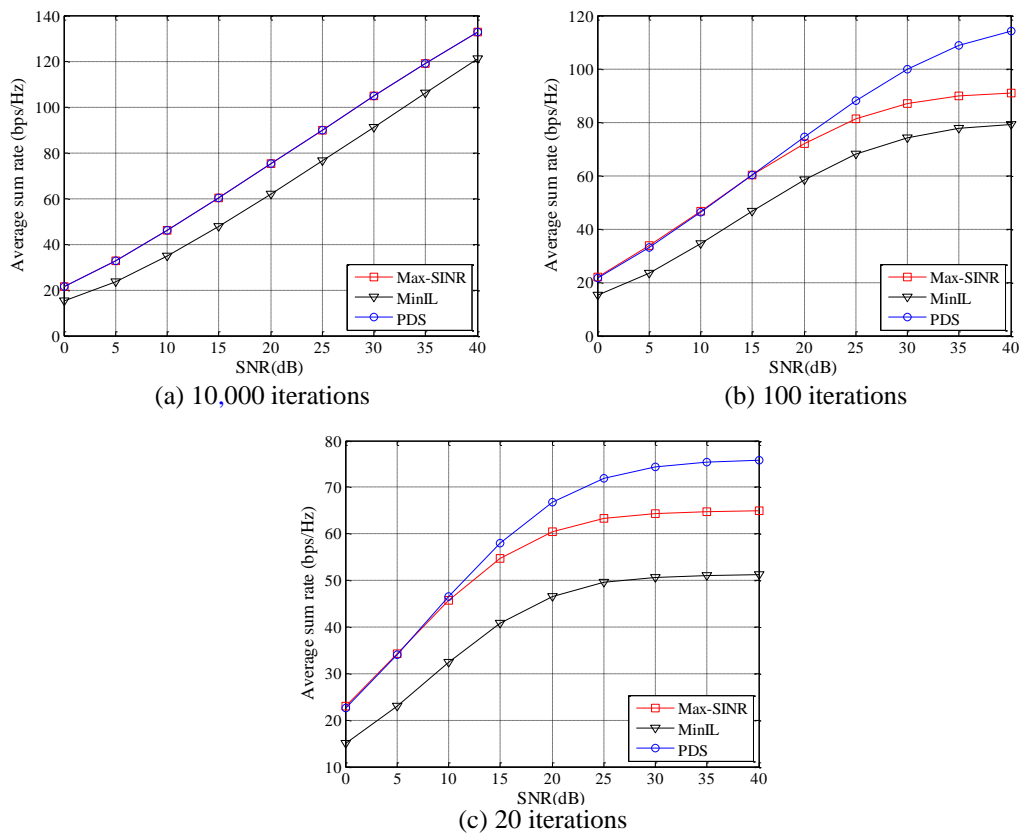

(a) 10,000 iterations   (b) 100 iterations

(c) 20 iterations

Fig. 5. Average sum rate versus SNR under iterations of 10,000, 100, and 20 for $(5{\times}5, 1)^9$

To further evaluate the performances of the algorithms, the average iterations and computational time under different numbers of users, data streams, and antennas are listed in **Table 1**. We consider the $(M{\times}M, d)^K$ channel with $K=2M/d$-1, the maximal user number in the IA networks [9]. All three algorithms are run with sufficient iterations of 10,000, and the final sum rate R can be obtained. Then, we calculate the iterations that are required to reach the sum rate of $0.9R$ and $0.99R$ as $I_{0.9}$ and $I_{0.99}$, respectively. By averaging $I_{0.9}$, $I_{0.99}$, and R over 250 realizations of channel coefficients, we attain $I_{AV0.9}$, $I_{AV0.99}$, and R_{AV} for each algorithm. Similarly, we can attain $T_{AV0.9}$ and $T_{AV0.99}$ for the computational time. In addition, the ratios of the average iteration and computational time between PDS and Max-SINR algorithms are provided in parentheses under the corresponding values of the PDS algorithm in **Table 1**. For

the single data stream situation, the PDS and Max-SINR algorithms can achieve the same sum rate, which is comparably higher than those of the MinIL algorithm. From the perspective of average iterations, the PDS algorithm can reach the sum rate of $0.9R_{AV}$ and $0.99R_{AV}$ with only 29~30% and 25~33% of the iterations required by the Max-SINR algorithm. Furthermore, in terms of computational time, the PDS algorithm can reach the sum rate of $0.9R_{AV}$ and $0.99R_{AV}$ with only 51~55% and 44~61% of the time required by the Max-SINR algorithm. For the multi-data stream situation, the PDS algorithm can achieve a higher sum rate than the Max-SINR algorithm with lower computational complexity and more rapid convergence.

Table 1 Statistics on the average iterations, computational time, and sum rate of different algorithms for $(M \times M, d)^K$ with SNR=30 dB

(M, d, K)	Algorithms	$I_{AV0.9}$	$I_{AV0.99}$	$T_{AV0.9}$ (s)	$T_{AV0.99}$ (s)	R_{AV} (bps/Hz)
(5, 1,9)	Max-SINR	2.04×10^2	1.18×10^3	6.81×10^{-1}	3.95×10^0	105
	MinIL	2.49×10^2	1.03×10^3	1.01×10^0	4.16×10^0	91
	PDS	6.00×10^1 (0.29)	2.95×10^2 (0.25)	3.48×10^{-1} (0.51)	1.74×10^0 (0.44)	105
(6, 1, 11)	Max-SINR	2.41×10^2	1.75×10^3	1.14×10^0	8.28×10^0	131
	MinIL	2.91×10^2	1.54×10^3	1.68×10^0	8.85×10^0	111
	PDS	6.93×10^1 (0.29)	5.18×10^2 (0.30)	5.82×10^{-1} (0.51)	4.44×10^0 (0.54)	131
(9, 1, 17)	Max-SINR	2.85×10^2	2.33×10^3	3.29×10^0	2.67×10^1	212
	MinIL	3.95×10^2	2.43×10^3	5.75×10^0	3.53×10^1	172
	PDS	8.61×10^1 (0.30)	7.75×10^2 (0.33)	1.78×10^0 (0.54)	1.63×10^1 (0.61)	212
(10, 1, 19)	Max-SINR	3.06×10^2	2.20×10^3	4.27×10^0	3.06×10^1	239
	MinIL	4.02×10^2	2.43×10^3	6.55×10^0	3.96×10^1	191
	PDS	9.24×10^1 (0.30)	6.05×10^2 (0.27)	2.33×10^0 (0.55)	1.54×10^1 (0.50)	239
(12, 2, 11)	Max-SINR	2.00×10^3	6.99×10^3	4.02×10^1	1.40×10^2	250
	MinIL	2.93×10^2	2.08×10^3	4.82×10^0	3.38×10^1	216
	PDS	6.97×10^2 (0.35)	4.82×10^3 (0.69)	1.92×10^1 (0.48)	1.31×10^2 (0.93)	255
(12, 4, 5)	Max-SINR	2.04×10^3	6.92×10^3	2.08×10^1	6.99×10^1	210
	MinIL	2.34×10^2	2.04×10^3	1.17×10^0	1.01×10^1	195
	PDS	7.07×10^2 (0.35)	5.01×10^3 (0.72)	8.75×10^0 (0.42)	6.16×10^1 (0.88)	216

6. Conclusions

In this paper, we have focused on reducing the execution time and computational complexity of the traditional Max-SINR algorithm for IA. The properties of the Max-SINR method have been investigated, and the principle direction has been introduced, which is found to approximately point to the convergence solutions. A rapid convergent low-complexity PDS algorithm has been proposed that approaches the convergence point directly along the principle direction. Furthermore, the analytical form of the optimal step size along this direction has been formulated under the minimal interference leakage criterion. Numerical results have verified that the proposed PDS algorithm can achieve a significantly higher convergence rate and lower overall computational complexity than the conventional MinIL and Max-SINR algorithms while maintaining the high sum rate. This paper has developed a

rapid convergent algorithm and provided theoretical insight into the Max-SINR algorithm.

Acknowledgement

Appendix A: Proof of Theorem 2

Proof: As shown by Schmidt *et al*. [11], the Max-SINR algorithm can suppress the interference leakage to approximately zero and approach the IA solutions in the high SNR regime, *i.e.*,

$$\tilde{\mathbf{D}}_j^{\mathrm{H}} \mathbf{H}_{ji} \tilde{\mathbf{P}}_i \approx \mathbf{0}, i \neq j. \tag{A.1}$$

Then, the desired signal and the interference-plus-noise covariance matrices of the *l*th data stream of the *k*th user for the forward and backward channels at the convergence point can be expressed as:

$$\tilde{\mathbf{A}}_{kl} = \mathbf{H}_{kk} \tilde{\mathbf{P}}_{k(l)} \tilde{\mathbf{P}}_{k(l)}^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}} \tag{A.2}$$

$$\overline{\tilde{\mathbf{A}}}_{kl} = \mathbf{H}_{kk}^{\mathrm{H}} \tilde{\mathbf{D}}_{k(l)} \tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}} \mathbf{H}_{kk} \tag{A.3}$$

$$\tilde{\mathbf{B}}_{kl} = \sum_{j=1, j \neq k}^{K} \mathbf{H}_{kj} \tilde{\mathbf{P}}_j \tilde{\mathbf{P}}_j^{\mathrm{H}} \mathbf{H}_{kj}^{\mathrm{H}} + \sum_{r=1, r \neq l}^{d} \mathbf{H}_{kk} \tilde{\mathbf{P}}_{k(r)} (\tilde{\mathbf{P}}_{k(r)})^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}} + \sigma^2 \mathbf{I} \tag{A.4}$$

$$\overline{\tilde{\mathbf{B}}}_{kl} = \sum_{j=1, j \neq k}^{K} \mathbf{H}_{jk}^{\mathrm{H}} \tilde{\mathbf{D}}_j \tilde{\mathbf{D}}_j^{\mathrm{H}} \mathbf{H}_{jk} + \sum_{r=1, r \neq l}^{d} \mathbf{H}_{kk}^{\mathrm{H}} \tilde{\mathbf{D}}_{k(r)} (\tilde{\mathbf{D}}_{k(r)})^{\mathrm{H}} \mathbf{H}_{kk} + \sigma^2 \mathbf{I}. \tag{A.5}$$

Define the generalized eigenvalues and the associated generalized eigenvectors of $\tilde{\mathbf{A}}_{kl}$ and $\tilde{\mathbf{B}}_{kl}$ as $\{\chi_{kl\_1}, \chi_{kl\_2}, ..., \chi_{kl\_N}\}$ and $\{\mathbf{v}_{kl\_1}, \mathbf{v}_{kl\_2}, ..., \mathbf{v}_{kl\_N}\}$, respectively. Similarly, define the generalized eigenvalues and the corresponding generalized eigenvectors of $\overline{\tilde{\mathbf{A}}}_{kl}$ and $\overline{\tilde{\mathbf{B}}}_{kl}$ as $\{\bar{\chi}_{kl\_1}, \bar{\chi}_{kl\_2}, ..., \bar{\chi}_{kl\_M}\}$ and $\{\bar{\mathbf{v}}_{kl\_1}, \bar{\mathbf{v}}_{kl\_2}, ..., \bar{\mathbf{v}}_{kl\_M}\}$, respectively. From Theorem 1, there is only one nonzero generalized eigenvalue, *i.e.*,

$$\chi_{kl\_1} > 0, \chi_{kl\_2} = \chi_{kl\_3} = ... = \chi_{kl\_N} = 0 \tag{A.6}$$

$$\bar{\chi}_{kl\_1} > 0, \bar{\chi}_{kl\_2} = \bar{\chi}_{kl\_3} = ... = \bar{\chi}_{kl\_M} = 0. \tag{A.7}$$

From Theorem 1, $\tilde{\mathbf{D}}_{k(l)}$ should be the generalized eigenvector associated with the maximal generalized eigenvalue of $\tilde{\mathbf{A}}_{kl}$ and $\tilde{\mathbf{B}}_{kl}$; $\tilde{\mathbf{P}}_{k(l)}$ should be the generalized eigenvector associated with the maximal generalized eigenvalue of $\overline{\tilde{\mathbf{A}}}_{kl}$ and $\overline{\tilde{\mathbf{B}}}_{kl}$, *i.e.*,

$$\tilde{\mathbf{P}}_{k(l)} = \bar{\mathbf{v}}_{kl\_1} \tag{A.8}$$

$$\tilde{\mathbf{D}}_{k(l)} = \mathbf{v}_{kl\_1}. \tag{A.9}$$

Substitute (23) into (13), and $\mathbf{B}_{kl}^{[i]}$ can be expressed as

$$\mathbf{B}_{kl}^{[i]} = \sum_{j=1, j\neq k}^{K} \mathbf{H}_{kj}(\tilde{\mathbf{P}}_j + \Delta\mathbf{P}_j^{[i-1]})(\tilde{\mathbf{P}}_j + \Delta\mathbf{P}_j^{[i-1]})^{\mathrm{H}} \mathbf{H}_{kj}^{\mathrm{H}}$$

$$+ \sum_{r=1, r\neq l}^{d} \mathbf{H}_{kk}(\tilde{\mathbf{P}}_{k(r)} + \Delta\mathbf{P}_{k(r)}^{[i-1]})(\tilde{\mathbf{P}}_{k(r)} + \Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}} + \sigma^2\mathbf{I} \qquad (A.10)$$

$$= \tilde{\mathbf{B}}_{kl} + \Delta\mathbf{B}_{kl}^{[i]}$$

where $\Delta\mathbf{B}_{kl}^{[i]}$ is given by

$$\Delta\mathbf{B}_{kl}^{[i]} = \sum_{j=1, j\neq k}^{K} \mathbf{H}_{kj}[\tilde{\mathbf{P}}_j(\Delta\mathbf{P}_j^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_j^{[i-1]}\tilde{\mathbf{P}}_j^{\mathrm{H}} + \Delta\mathbf{P}_j^{[i-1]}(\Delta\mathbf{P}_j^{[i-1]})^{\mathrm{H}}]\mathbf{H}_{kj}^{\mathrm{H}}$$

$$+ \sum_{r=1, r\neq l}^{d} \mathbf{H}_{kk}[\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_{k(r)}^{[i-1]}\tilde{\mathbf{P}}_{k(r)}^{\mathrm{H}} + \Delta\mathbf{P}_{k(r)}^{[i-1]}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}]\mathbf{H}_{kk}^{\mathrm{H}} \qquad (A.11)$$

From (8) and (23), the desired signal covariance matrix can be written as

$$\mathbf{A}_{kl}^{[i]} = \mathbf{H}_{kk}\mathbf{P}_{k(l)}^{[i-1]}(\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}}$$

$$= \mathbf{H}_{kk}(\tilde{\mathbf{P}}_{k(l)} + \Delta\mathbf{P}_{k(l)}^{[i-1]})(\tilde{\mathbf{P}}_{k(l)} + \Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}} \mathbf{H}_{kk}^{\mathrm{H}} \qquad (A.12)$$

$$= \tilde{\mathbf{A}}_{kl} + \Delta\mathbf{A}_{kl}^{[i]}$$

where $\Delta\mathbf{A}_{kl}^{[i]}$ is given by

$$\Delta\mathbf{A}_{kl}^{[i]} = \mathbf{H}_{kk}[\tilde{\mathbf{P}}_{k(l)}(\Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_{k(l)}^{[i-1]}\tilde{\mathbf{P}}_{k(l)}^{\mathrm{H}} + \Delta\mathbf{P}_{k(l)}^{[i-1]}(\Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}}]\mathbf{H}_{kk}^{\mathrm{H}}. \qquad (A.13)$$

When $\overline{\mathbf{p}}^{[i]}$ is located within a small neighborhood of $\overline{\overline{\mathbf{p}}}$, $\Delta\mathbf{P}_{k(l)}^{[i-1]}$ is a small item, and the high order terms in (A.11) and (A.13) can be omitted. Thus, $\Delta\mathbf{A}_{kl}^{[i]}$ and $\Delta\mathbf{B}_{kl}^{[i]}$ can be approximated as

$$\Delta\mathbf{A}_{kl}^{[i]} \approx \mathbf{H}_{kk}[\tilde{\mathbf{P}}_{k(l)}(\Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_{k(l)}^{[i-1]}\tilde{\mathbf{P}}_{k(l)}^{\mathrm{H}}]\mathbf{H}_{kk}^{\mathrm{H}} \qquad (A.14)$$

$$\Delta\mathbf{B}_{kl}^{[i]} \approx \sum_{j=1, j\neq k}^{K} \mathbf{H}_{kj}[\tilde{\mathbf{P}}_j(\Delta\mathbf{P}_j^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_j^{[i-1]}\tilde{\mathbf{P}}_j^{\mathrm{H}}]\mathbf{H}_{kj}^{\mathrm{H}}$$

$$+ \sum_{r=1, r\neq l}^{d} \mathbf{H}_{kk}[\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}} + \Delta\mathbf{P}_{k(r)}^{[i-1]}\tilde{\mathbf{P}}_{k(r)}^{\mathrm{H}}]\mathbf{H}_{kk}^{\mathrm{H}} \qquad (A.15)$$

From Theorem 1, the decoding vector $\mathbf{D}_{k(l)}^{[i]}$ in the forward iteration is the generalized eigenvector associated with the maximal generalized eigenvalue of $\mathbf{A}_{kl}^{[i]}$ and $\mathbf{B}_{kl}^{[i]}$. From the eigenvalue perturbation theory, when $\Delta\mathbf{A}_{kl}^{[i]}$ and $\Delta\mathbf{B}_{kl}^{[i]}$ are very small, $\mathbf{D}_{k(l)}^{[i]}$ can be approximated as

$$\mathbf{D}_{k(l)}^{[i]} \approx (1 - \frac{\mathbf{v}_{kl\_1}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\mathbf{v}_{kl\_1}}{2})\mathbf{v}_{kl\_1} + \sum_{n=2}^{N} \frac{\mathbf{v}_{kl\_n}^{\mathrm{H}}(\Delta\mathbf{A}_{kl}^{[i]} - \chi_{kl\_1}\Delta\mathbf{B}_{kl}^{[i]})\mathbf{v}_{kl\_1}}{\chi_{kl\_1} - \chi_{kl\_n}}\mathbf{v}_{kl\_n}. \qquad (A.16)$$

Substitute (A.6) and (A.9) into (A.16), and $\mathbf{D}_{k(l)}^{[i]}$ can be rewritten as

$$\mathbf{D}_{k(l)}^{[i]} \approx (1 - \frac{\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}}{2})\tilde{\mathbf{D}}_{k(l)} + \sum_{n=2}^{N} \frac{\mathbf{v}_{kl\_n}^{\mathrm{H}}(\Delta\mathbf{A}_{kl}^{[i]} - \chi_{kl\_1}\Delta\mathbf{B}_{kl}^{[i]})\tilde{\mathbf{D}}_{k(l)}}{\chi_{kl\_1}}\mathbf{v}_{kl\_n} = \tilde{\mathbf{D}}_{k(l)} + \Delta\mathbf{D}_{k(l)}^{[i]}$$

$$(A.17)$$

where $\Delta\mathbf{D}_{k(l)}^{[i]}$ is given by:

$$\Delta\mathbf{D}_{k(l)}^{[i]} = -\frac{\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}}{2}\tilde{\mathbf{D}}_{k(l)} + \sum_{n=2}^{N}\frac{\mathbf{v}_{kl\_n}^{\mathrm{H}}(\Delta\mathbf{A}_{kl}^{[i]} - \chi_{kl\_1}\Delta\mathbf{B}_{kl}^{[i]})\tilde{\mathbf{D}}_{k(l)}}{\chi_{kl\_1}}\mathbf{v}_{kl\_n}. \qquad (A.18)$$

Substitute (A.14) and (A.15) into (A.18), and the terms $\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}$, $\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{A}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}$, and $\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}$ in (A.18) can be expanded as:

$$\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)} \approx \sum_{j=1, j\neq k}^{K}[\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kj}\tilde{\mathbf{P}}_{j}(\Delta\mathbf{P}_{j}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kj}\Delta\mathbf{P}_{j}^{[i-1]}\tilde{\mathbf{P}}_{j}^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]$$

$$+ \sum_{r=1, r\neq l}^{d}[\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}\Delta\mathbf{P}_{k(r)}^{[i-1]}\tilde{\mathbf{P}}_{k(r)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]$$

$$(A.19)$$

$$\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{A}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)} \approx \mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(l)}(\Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\Delta\mathbf{P}_{k(l)}^{[i-1]}\tilde{\mathbf{P}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} \quad (A.20)$$

$$\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)} \approx \sum_{j=1, j\neq k}^{K}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kj}\tilde{\mathbf{P}}_{j}(\Delta\mathbf{P}_{j}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kj}\Delta\mathbf{P}_{j}^{[i-1]}\tilde{\mathbf{P}}_{j}^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]$$

$$+ \sum_{r=1, r\neq l}^{d}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\Delta\mathbf{P}_{k(r)}^{[i-1]}\tilde{\mathbf{P}}_{k(r)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}].$$

$$(A.21)$$

When the interference leakage at the convergence point is assumed to be approximately zero as shown in (A.1), the terms $\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kj}\tilde{\mathbf{P}}_{j}$, $\tilde{\mathbf{P}}_{j}^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}$, $\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(r)}$, and $\tilde{\mathbf{P}}_{k(r)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}$ in (A.19) and (A.21) can be ignored. Then, $\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}$ and $\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)}$ can be rewritten as

$$\tilde{\mathbf{D}}_{k(l)}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)} \approx 0 \qquad (A.22)$$

$$\mathbf{v}_{kl\_n}^{\mathrm{H}}\Delta\mathbf{B}_{kl}^{[i]}\tilde{\mathbf{D}}_{k(l)} \approx \sum_{j=1, j\neq k}^{K}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kj}\tilde{\mathbf{P}}_{j}(\Delta\mathbf{P}_{j}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]$$

$$+ \sum_{r=1, r\neq l}^{d}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}].$$

$$(A.23)$$

Substitute (A.20), (A.22), and (A.23) into (A.18), and we have

$$\Delta\mathbf{D}_{k(l)}^{[i]} \approx \frac{1}{\chi_{kl\_1}}\sum_{n=2}^{N}\{[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(l)}(\Delta\mathbf{P}_{k(l)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)} + \mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\Delta\mathbf{P}_{k(l)}^{[i-1]}\tilde{\mathbf{P}}_{k(l)}^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]\mathbf{v}_{kl\_n}\}$$

$$-\sum_{n=2}^{N}\{\sum_{j=1, j\neq k}^{K}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kj}\tilde{\mathbf{P}}_{j}(\Delta\mathbf{P}_{j}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kj}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]\mathbf{v}_{kl\_n}\}$$

$$-\sum_{n=2}^{N}\{\sum_{r=1, r\neq l}^{d}[\mathbf{v}_{kl\_n}^{\mathrm{H}}\mathbf{H}_{kk}\tilde{\mathbf{P}}_{k(r)}(\Delta\mathbf{P}_{k(r)}^{[i-1]})^{\mathrm{H}}\mathbf{H}_{kk}^{\mathrm{H}}\tilde{\mathbf{D}}_{k(l)}]\mathbf{v}_{kl\_n}\}.$$

$$(A.24)$$

The items that depend on the iterations i come from $\Delta\mathbf{P}_{j(l)}^{[i-1]}$ and $(\Delta\mathbf{P}_{j(l)}^{[i-1]})^{\mathrm{H}}$ with j=1, 2, …, K and l=1, 2, …, d. Consequently, $\Delta\mathbf{D}_{k(l)}^{[i]}$ can be represented by the linear combination of $\mathrm{Re}\{\Delta\mathbf{P}_{j(l)}^{[i-1]}\}$ and $\mathrm{Im}\{\Delta\mathbf{P}_{j(l)}^{[i-1]}\}$. Therefore, there exists an iteration-independent transformation $\mathbf{T}_{PD}\in\mathbb{R}^{2NKd\times2MKd}$, so that

$$\Delta\overline{\mathbf{d}}^{[i]} \approx \mathbf{T}_{PD}\Delta\overline{\mathbf{p}}^{[i-1]}. \tag{A.25}$$

Similarly in the backward iteration, there exists an iteration-independent transformation $\mathbf{T}_{DP}\in\mathbb{R}^{2MKd\times2NKd}$, and $\Delta\overline{\mathbf{p}}_{k}^{[i]}$ can be approximated as

$$\Delta\overline{\mathbf{p}}^{[i]} \approx \mathbf{T}_{DP}\Delta\overline{\mathbf{d}}^{[i]}. \tag{A.26}$$

From (A.25) and (A.26), the relationship of the precoding and the decoding vectors between two successive iterations can be depicted as

$$\Delta\overline{\mathbf{d}}^{[i]} \approx \mathbf{T}_{PD}\mathbf{T}_{DP}\Delta\overline{\mathbf{d}}^{[i-1]} = \mathbf{T}_{D}\Delta\overline{\mathbf{d}}^{[i-1]} \tag{A.27}$$

$$\Delta\overline{\mathbf{p}}^{[i]} \approx \mathbf{T}_{DP}\mathbf{T}_{PD}\Delta\overline{\mathbf{p}}^{[i-1]} = \mathbf{T}_{P}\Delta\overline{\mathbf{p}}^{[i-1]} \tag{A.28}$$

where

$$\mathbf{T}_{D} = \mathbf{T}_{PD}\mathbf{T}_{DP} \in\mathbb{R}^{2NKd\times2NKd} \tag{A.29}$$

$$\mathbf{T}_{P} = \mathbf{T}_{DP}\mathbf{T}_{PD} \in\mathbb{R}^{2MKd\times2MKd}. \tag{A.30}$$

References

[1] H. Zhang, X. Chu, W. Ma, W. Zheng and X. Wen, "Resource allocation with interference mitigation in OFDMA femtocells for co-channel deploymenta," *EURASIP J. Wireless Commun. Netw*, vol. 2012, no. 1, pp. 289, 2012. Article (CrossRef Link)

[2] R. Shrestha and J. M. Kim, "Proportional resource allocation in OFDMA systems based on weighted normalized user's CSI," *IEICE Transactions on Communications*, vol. E95B, no. 6, pp. 2137-2140, 2012. Article (CrossRef Link)

[3] N. Zhao, F. R. Yu, H. Sun, A. Nallanathan and H. Yin, "A novel interference alignment scheme based on sequential antenna switching in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5008-5021, 2013. Article (CrossRef Link)

[4] 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, 2008. Article (CrossRef Link)

[5] H. Sung, S.-H. Park, K.-J. Lee and I. Lee, "Linear precoder designs for *K*-user interference channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 291-301, 2010. Article (CrossRef Link)

[6] H. Chae, K. Kim, R. Ran and D. K. Kim, "A single feedback based interference alignment for three-user MIMO interference channels with limited feedback," *KSII Trans. on internet and information systems*, vol. 7, no. 4, pp. 692-710, 2013. Article (CrossRef Link)

[7] N. Zhao, F. R. Yu and V. C. M. Leung, "Opportunistic Communications in Interference Alignment Networks with Wireless Power Transfer," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 88-95, 2015. Article (CrossRef Link)

[8] N. Zhao, F. R. Yu and H. Sun, "Adaptive Energy-Efficient Power Allocation in Green Interference Alignment Wireless Networks," *IEEE Trans. Veh. Technol.*, to appear. Article (CrossRef Link)

[9] C. M. Yetis, T. Gou, S. A. Jafar and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771-4782, 2010. Article (CrossRef Link)

[10] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE*

Trans. Veh. Technol., vol. 60, no. 1, pp. 206-218, 2011. Article (CrossRef Link)

[11] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig and W. Utschick, "Comparison of distributed beamforming algorithms for MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3476-3489, 2013. Article (CrossRef Link)

[12] M. Razaviyayn, M. Sanjabi and Z.-Q. Luo, "Linear transceiver design for interference alignment: complexity and computation," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2896-2910, 2012. Article (CrossRef Link)

[13] K. Gomadam, V. R. Cadambe and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1-6, 2008. Article (CrossRef Link)

[14] 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, 2011. Article (CrossRef Link)

[15] S. S. Christensen, R. Agarwal, E. Carvalho and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792-4799, 2008. Article (CrossRef Link)

[16] H. Shen, B. Li, M. Tao and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480-3489, 2010. Article (CrossRef Link)

[17] R. Shrestha, I. Bae and J. M. Kim, "A leakage-based solution for interference alignment in MIMO interference channel networks," *KSII Trans. on internet and information systems*, vol. 8, no. 2, pp. 424-442, 2014. Article (CrossRef Link)

[18] I. Santamaria, O. Gonzalez, R. W. Heath and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. of IEEE GLOBECOM*, pp. 1-6, 2010. Article (CrossRef Link)

[19] T. Y. Xu and X. G. Xia, "A diversity analysis for distributed interference alignment using the Max-SINR algorithm," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1857-1868, 2014. Article (CrossRef Link)

[20] X. Han, A. de Almeida and Z. Yang, "Channel estimation for MIMO multi-relay systems using a tensor approach," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 163, 2014. Article (CrossRef Link)

[21] X. Chen and C. Yuen, "Performance Analysis and Optimization for Interference Alignment Over MIMO Interference Channels With Limited Feedback," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1785-1795, 2014. Article (CrossRef Link)

[22] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, 1st Edition, Oxford Univ. Press, New York, 1988.

Zhilu Wu is a professor in the School of Electronics and Information Engineering and the dean of the Information Engineering Department at Harbin Institute of Technology (HIT), Harbin, China. He received the B.S. degree in electronic instrument and measurement technology in 1983, M.E. degree in signal and information processing in 1989, and Ph.D. in information and communication engineer in 2008, respectively, from HIT. His research interests include wireless communication, interference alignment, and software radio.

Lihui Jiang received the B.S. degree in science and technology of optical information in 2009 from South China University of Technology, Guangzhou, China. He received the M.E. degree in signal and information processing in 2011 from HIT. He is currently working toward the Ph.D. degree in information and communication at HIT. His research interests include wireless communication, interference alignment, continuous phase modulation, and machine learning.

Guanghui Ren is a professor in the School of Electronics and Information Engineering at HIT. He received the B.S. degree in electronic instrument and measurement technology in 1984 and M.E. degree in signal and information processing in 2003 from HIT. His research interests include wireless communication, interference alignment, image processing, automatic test systems, and parallel computing.

Gangyi Wang is currently a lecturer in the School of Instrumentation Science and Opto-electronics Engineering at Beihang University, Beijing, China. He received the B.S. degree in electronics and information engineering in 2006, the M.E. degree, and Ph.D. degree in information and communication engineering in 2008 and 2013, respectively, from HIT. His research interests include computer vision, interference alignment, and embedded systems.

Nan Zhao is currently an associate professor in the School of Information and Communication Engineering at Dalian University of Technology, Liaoning, China. He received the B.S. degree in electronics and information engineering in 2005, the M.E. degree in signal and information processing in 2007, and the Ph.D. degree in information and communication engineering in 2011, respectively, from HIT. He serves as an Area Editor for the journal "AEÜ -International Journal of Electronics and Communications". He served (or is serving) as a technical program committee (TPC) member for a number of international conferences, e.g., Globecom 2013 & 2014, GreenCom 2013, VTC 2012, etc. His recent research interests include interference alignment, green communications, cognitive radio, and optical communications.

Yaqin Zhao is a professor in the School of Electronics and Information Engineering at HIT. She received the B.S. degree in electronics and information engineering in 1998, M.E. degree in signal and information processing in 2000, and Ph.D. in information and communication engineering in 2008, respectively, from HIT. Her research interests include wireless communication, interference alignment, software radio, and CDMA.