Virtual Cell Beamforming in Cooperative Networks

-Technical Report-

Jihwan Kim, *Student Member, IEEE,* Hyang-Won Lee, *Member, IEEE,* Song Chong, *Member, IEEE*

Abstract

We study the coordinated transmission problem in cooperative cellular networks where a cluster of base stations forms a *virtual cell* to serve a mobile station (MS). The performance of such a MS-centric virtual cell network is dictated by the beamformer that enables to suppress interference, however designing a beamformer is highly challenging due to the coupled nature of interference and desired signals under arbitrarily formed virtual cells. We develop a new formulation of the beamforming problem for sum-rate maximization in virtual cell networks, and analyze the structure of its optimal solutions. Based on this analysis, we develop a beamforming algorithm that can balance between desired signal maximization and interference minimization, so as to maximize the sum-rate. We show through extensive simulations that our *balanced beamforming* algorithm mitigates edge user effect and outperforms existing algorithms in various scenarios where virtual cells are allowed to overlap.

Index Terms

Wireless Networks, MIMO System, Base Station Cooperation, Virtual Cell, Overlapped Cluster, Multi-cell Beamforming, Sum-rate Maximization

Jihwan Kim and Song Chong are with the Department of Electrical Engineering, KAIST, Daejeon, Republic of Korea. (e-mail: kimji@netsys.kaist.ac.kr, songchong@kaist.edu)

Hyang-Won Lee is with the Department of Internet & Multimedia Engineering, Konkuk University, Seoul, Republic of Korea. (e-mail: leehw@konkuk.ac.kr.)

I. Introduction

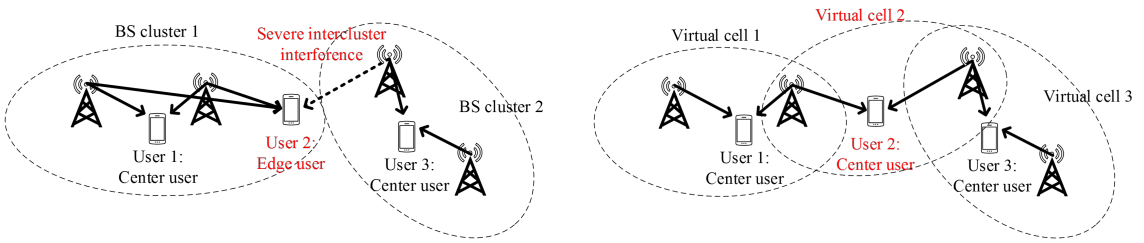
The demand for wireless capacity has been rapidly increasing due to the explosive growth of mobile devices, so the next generation wireless systems such as 5G will have to support a thousand times more traffic than it has to today [1]. A number of solutions were proposed and are being used in practice to address this issue. Recently, there has been a considerable amount of attention drawn to the ultra-dense network (UDN) where a huge number of small cells are deployed close to mobile stations, so as to boost the network capacity [2]–[5]. However, such a UDN approach could only offer a marginal performance improvement for cell boundary users due to severe intercell interference, as base stations (BSs) come close to each other. It is thus indispensable for base stations to cooperate in order to manage intercell interference [6].

BS cooperation techniques for interference mitigation are generally called the Coordinated Multi-Point or CoMP in the LTE standard [7], [8]. Above all, the joint transmission, in which multiple BSs form a cluster[1] and cooperatively transmit to a mobile station (MS), is considered the most promising solution for managing intercell interference, since a potential interferer can work as a serving base station transmitting useful signal. Hence, in such a coordinated multi-cell transmission, a mobile station can receive data from multiple BSs as if there is a "virtual cell" serving the MS. We will thus refer to the joint transmission in CoMP as the *virtual-cell approach*. Although the concept of virtual-cell approach has been accepted generally, it still remains to fully tap its potentials, and thus advanced cooperation techniques are required for the future ultra-dense network where a large number of small cells coordinate their transmissions.

One of the key components of virtual-cell approach is the clustering policy that determines a set or cluster[2] of BSs that will coordinate their transmissions for MSs. Conceivably, how to form a cluster can significantly affect the performance of boundary users which is one of the main performance metrics that

[1]Ideally, all the BSs in the network can cooperate, forming one super-BS. This approach achieves high capacity by perfectly managing intercell interference, but it is impractical since it requires heavy data sharing and signaling overhead in the backhaul network [6], [9]. Thus, a practical approach is to limit the size of a cluster.

[2]We will use cluster and virtual cell interchangeably throughout the paper.



(a) Overlapped clusters are not allowed.

(b) MS-centric clustering where overlapped clusters are allowed.

Fig. 1. Various clustering policies: dashed circles are clusters.

the virtual-cell approach aims at improving. When overlapped clusters are not allowed as in Fig. 1(a), each cluster manages inter-cell interference within the cluster, but cluster edge usres can still experience severe inter-cluster interference. On the other hand, in MS-centric clustering where overlapped clusters are allowed (Fig. 1(b)), the notion of cell or cluster boundaries is essentially eliminated. This can substantially enhance the performance of MSs that would otherwise suffer from boundary effect. Hence, it is important to allow clusters of BSs to be formed arbitrarily based on the needs of MSs.

Another key component of virtual-cell approach is the *coordinated transmission policy* that determines (for a given clustering) the transmit signals of clusters so as to optimize a certain performance metric. Since a virtual cell network can be viewed as a networked multi-input and multi-output (MIMO) system where a mobile station can receive from multiple base stations and a base station can transmit to multiple mobile stations concurrently, the studies of coordinated transmissions in virtual cell networks are largely based on the transmission techniques used in MIMO systems. In a typical MIMO system, multiple data streams are concurrently transmitted from the same base station, and as a consequence, mobile stations can experience severe interference. To address this problem, the orientation of signals is controlled so that interference can be suppressed at the receiver side. This is called the *beamforming technique*, which is comprised of the *precoding* for beam direction decision and and *power scaling* for beam amplitude decision. It is well-known that how to design a beamformer has a profound effect on the performance of MIMO systems.

There has been a large body of work that deals with beamforming in the context of virtual cell networks.

In particular, a number of beamforming algorithms have been proposed with various objectives such as maximizing desired signal [10], nullifying interference [11]–[14], minimizing power subject to SINR constraints [15], [16], balancing the SINR[3] among all mobile stations [17], maximizing virtual SINR [18], [19] and minimizing weighted mean-squared error [20], [21]. Most of these algorithms are specialized to a limited scenario (More details are discussed in Section II), thereby making it difficult to use in the virtual cell network that can have arbitrary clusters based on the needs of MSs with allowing overlap and so experience various SNR regimes and interference environments.

In this paper, we focus on the beamforming (i.e., precoding and power scaling) algorithm that works well with various MS-centric clustering policies. In the coordinated transmission, the beam for an MS is formed by signals, each coming from a different BS, and thus designing a beamformer requires a complicated joint optimization of beam direction and amplitude across multiple BS power budgets and antennas in different BSs. We develop a new formulation of the sum-rate maximization problem that can be decomposed into the precoding problem, which determines only the direction of the beam for each MS, and the power scaling problem, which determines only the amplitude of each beam subject to power budget constraints. This decomposition alleviates the difficulty caused by arbitrarily formed clusters and provides an insight into the structure of optimal precoding. In particular, we show that the optimal precoding is an eigenvector of a certain matrix (called the balancing matrix; see Section IV-B). Based on this analysis, we develop a precoding algorithm for maximizing the sum-rate. For power scaling, we develop a cluster-level water-filling algorithm by applying the idea of iterative water-filling [22]–[24]. We verify through extensive simulations that our algorithms outperform the existing algorithms [10], [14], [18], [22], [25], i) over various SNR and interference environments, ii) under overlapped/non-overlapped clustering policies, and iii) without global channel state information. There still remain several issues to be address in order for coordinated multi-cell transmission techniques to be used in practice, e.g., due to heavy signal overhead and synchronization. Nevertheless, our results in this paper demonstrate the

[3]The term "balancing" in [17] is used to indicate that it achieves fair data rates among users. On the other hand, the term "balancing" in our work means that our algorithm seeks a balance between desired signal maximization and interference minimization.

potentials of coordinated virtual-cell transmission for boosting the network capacity and mitigating cell boundary effect in future cellular systems such as cloud RAN [26].

The rest of the paper is organized as follows. In Section II, we discuss related work. In Section III, we introduce the system model and formulate the beamforming problem. In Section IV, we decompose the beamforming problem into precoding problem and power scaling problem, and propose our balanced beamforming algorithm. In Section V, we evaluate the performance of our balanced beamforming algorithm through extensive simulations.

II. Related Work

As conventional beamforming schemes, there are maximal ratio combining (MRC) [10], which maximizes desired signal, and zero forcing (ZF) [11]–[14], which nullify interference toward other streams. While these schemes can be implemented easily, they are applied to limited scenarios. For example, nullifying interference is an effective way to increase the sum-rate when interference is dominant. However, if the noise is dominant as in the low SNR regime, suppressing (negligible) interference has marginal impact on the performance. Likewise, MRC works well only in the low SNR regime and shows poor performance in the high SNR regime [27].

There are other beamforming schemes aiming at sum-rate maximization in the single-cell transmission [16], [20] and the multi-cell transmission [13], [14], [18], [19], [21], [25]. However, since the problem is non-convex, most of the works in this context focus on developing a suboptimal solution. In [20], the weighted sum-rate maximization problem is transformed to an equivalent weighted sum mean-squared error (MSE) minimization problem, which is more tractable than the original formulation. Based on the new formulation, a precoding algorithm is derived that converges to a local optimal solution. Park et al. [19] propose a beamforming scheme based on virtual SINR (VSINR), defined as a signal to leakage power plus noise ratio, in coordinated beamforming and two-cell joint transmission systems. The VSINR approximates the actual SINR only when interference is symmetric, and thus, using the VSINR may cause serious performance degradation in the case of asymmetric interference. The beamforming

scheme proposed in [19] resolves this problem by adaptively weighting leakage power. In [16], the authors propose a joint scheduling, precoding and power allocation to maximize the network utility by applying the precoding scheme in [15] that minimizes power consumption subject to SINR constraints.

Unlike the single-cell transmission setting where each beamformer is designed within a single power budget and antennas in the same BSs, in the coordinated multi-cell transmission, the beamformer for a data stream should be designed across multiple power budgets and antennas in different BSs. As a consequence, beamforming schemes for single-cell transmission cannot be directly applied to the coordinated multi-cell transmission setting, especially under overlapped clusters. In [13], the authors propose a block diagonalization (BD) based beamforming in the coordinated multi-cell transmission with non-overlapped clusters. Their key idea is to precancel other-cluster interference if a cluster edge user is served in the neighbor cluster. In addition, three power allocation methods are developed, however they can only be used with BD or when interference is completely canceled. Kaviani et al. [14] modify the BD precoding scheme to support the coordinated transmission with overlapped clustering. Nevertheless, to fully mitigate interference [13], [14], the precoding can potentially reduce desired signal power. In [18], the authors also propose a beamforming scheme for the overlapped cluster environment, which maximizes the clustered virtual SINR (CVSINR) with a power allocation proportional to the channel gain. However, similar to the VSINR method, the CVSINR scheme may perform poorly if interference is asymmetric. Bogale et al. [28] propose centralized/distributed algorithms based on MMSE method to maximize the weighted sum-rate. Their algorithms are developed with the assumption that every BS in the network cooperates as a single super-cell. By adapting the WMMSE framework in [20], Hong et al. [21] propose a joint clustering and beamforming algorithm in heterogeneous networks. In [25], by adapting Wiener filter, Hui proposes a power negotiation algorithm, called iterative Sum-Min algorithm, to resolve potential conflicts in power demand among overlapped clusters. The Sum-Min algorithm is simple, distributed, and well-fitted to any precoding schemes even in the overlapped cluster environment.

III. System Model

We consider a downlink virtual cell network consisting of a set $\mathcal{M}$ of BSs and a set $\mathcal{K}$ of MSs. To ignore the scheduling effect, we assume that every MS in $\mathcal{K}$ is a scheduled user and $\mathcal{K}$ is fixed throughout the paper. Each BS $m \in \mathcal{M}$ has N_T antennas and each MS $k \in \mathcal{K}$ has a single antenna[4]. We assume the universal frequency reuse, so there can be severe interference from one data stream to another. The BSs are all connected (through a backhaul network) to a central controller (as in cloud RAN [26] where multiple BSs are connected to and controlled by a central entity), which calculates the beamformer for all BSs. Using beamforming, interference between concurrently transmitted data streams can be suppressed.

We assume that there is a set of BSs for each MS to form a virtual cell, i.e., each MS has its virtual cell. We denote the BS cluster for MS k by $\mathcal{M}_k \subseteq \mathcal{M}$ and the set of MSs whose virtual cell is formed by BS m by $\mathcal{K}_m \subseteq \mathcal{K}$. The BSs within a cluster fully share user data, and thus they can perform a coordinated transmission by beamforming across all the BS antennas in the cluster. In other words, each BS m transmits to a set $\mathcal{K}_m$ of MSs and each MS k receives from a set $\mathcal{M}_k$ of BSs. In this paper, we assume that $\mathcal{M}_k$ (and thus $\mathcal{K}_m$) is predetermined by the underlying clustering policy. Since we do not assume any restriction on $\mathcal{M}_k$, it is possible for them to overlap, i.e., there can be $k, k' \in \mathcal{K}$ such that $\mathcal{M}_k \cap \mathcal{M}_{k'} \neq \emptyset$ for $\mathcal{M}_k \neq \mathcal{M}_{k'}$. It is the most general setting in the context of BS cooperation.

Let $\mathbf{h}_{km} \in \mathbb{C}^{1\times N_T}$ be the channel vector from BS m to MS k, where $\mathbb{C}^{a\times b}$ is the set of $a \times b$ complex matrices. The channel vector is given by $\mathbf{h}_{km} = [\alpha_{km,1}, \alpha_{km,2}, ..., \alpha_{km,N_T}]\sqrt{\rho_{km}}$ where $\alpha_{km,n}, n = 1, ..., N_T$ is the multi-antenna channel coefficient and ρ_{km} is a large-scale fading, such as path loss and shadowing, from BS m to MS k. We assume that every MS can perfectly measure their channel vectors and feed them back to the central controller. This is a typical assumption in the study of coordinated multi-cell transmission, and we discuss the case of imperfect (or local) channel state information in Sections IV-E and V-F.

The beamformer is comprised of transmit power scaling factor and precoding. Denote by p_{mk} and

[4]Although an MS may have multiple antennas, if they fix the receive beamformer, the MS can be regarded as a single-antenna user [16], [20].

$\mathbf{v}_{mk} \in \mathbb{C}^{N_T \times 1}$ the transmit power scaling factor and precoding vector for the stream of data signal $s_k \in \mathbb{C}$ from BS m to MS k, respectively. Each BS m has a power budget $\bar{P}_m$ and the following peak power constraint,

$$\sum_{k \in \mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_{mk} \leq \bar{P}_m, \forall m \in \mathcal{M}, \tag{1}$$

where $\|\cdot\|$ denotes the Euclidean norm.

The transmitted signal of BS m (denoted by $\mathbf{x}_m \in \mathbb{C}^{N_T \times 1}$) is given by

$$\mathbf{x}_m = \sum_{k \in \mathcal{K}_m} \mathbf{v}_{mk} \sqrt{p_{mk}} s_k, \tag{2}$$

and the received signal of MS k (denoted by $y_k \in \mathbb{C}$) is

$$y_k = \underbrace{\sum_{m \in \mathcal{M}_k} \mathbf{h}_{km} \mathbf{v}_{mk} \sqrt{p_{mk}} s_k +}_{\text{desired signal}}$$
$$\underbrace{\sum_{j \neq k, j \in \mathcal{K}} \sum_{m \in \mathcal{M}_j} \mathbf{h}_{km} \mathbf{v}_{mj} \sqrt{p_{mj}} s_j}_{\text{interference}} + z_k, \tag{3}$$

where $z_k \in \mathbb{C}$ is the additive complex white Gaussian noise with distribution $\mathcal{CN}(0, \sigma_k^2)$.

The signal to interference plus noise ratio (SINR) for MS k is thus expressed as

$$\gamma_k(\mathbf{p}, \mathbf{v}) = \frac{|\sum_{m \in \mathcal{M}_k} \mathbf{h}_{km} \mathbf{v}_{mk} \sqrt{p_{mk}}|^2}{\sum_{j \neq k, j \in \mathcal{K}} |\sum_{m \in \mathcal{M}_j} \mathbf{h}_{km} \mathbf{v}_{mj} \sqrt{p_{mj}}|^2 + \sigma_k^2}. \tag{4}$$

For given precoding $\mathbf{v}$ and power scaling $\mathbf{p}$, the achievable data rate for MS k is assumed to be $\log(1 + \gamma_k(\mathbf{p}, \mathbf{v}))$.

We study the following beamforming problem with arbitrary MS set $\mathcal{K}$ and BS clustering $\mathcal{M}_k, \forall k \in \mathcal{K}$:

[P-SRM] Sum-Rate Maximization Problem

$$\max_{\mathbf{p} \geq 0, \mathbf{v}} \sum_{k \in \mathcal{K}} w_k \log_2(1 + \gamma_k(\mathbf{p}, \mathbf{v})) \tag{5}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_{mk} \leq \bar{P}_m, \forall m \in \mathcal{M} \tag{6}$$

where w_k is the weight of MS k. The objective function is the general weighted sum-rate of mobile stations. By adjusting the weights w_k, the sum-rate maximization can be used to accomplish various

objectives. For example, if one desires to maximize the system capacity, equal weights can be used. Other objectives with respect to long-term throughput can also be achieved using the sum-rate maximization. For instance, if $w_k = [1/R_k(t)]^\alpha$ where $R_k(t)$ is the average throughput of MS k at time t, the long-term throughput achieves alpha-fairness among MSs [29]. In particular, as alpha increases, max-min fairness can be achieved.

Note again that no assumptions are imposed on BS clustering to form virtual cells in **[P-SRM]**. Thus, the optimal beamforming of **[P-SRM]** maximizes the sum-rate whether the virtual cells are disjoint or non-disjoint and whether the MSs are in the low or high SNR regime. However, solving **[P-SRM]** directly is not easy due to its non-convexity and the coupled nature of $\mathbf{p}$ and $\mathbf{v}$. Our goal in this paper is to develop a beamforming algorithm for maximizing the sum-rate, by analyzing the optimal solution of **[P-SRM]**.

IV. Beamforming for Sum-Rate Maximization

To solve **[P-SRM]**, we reformulate and decompose the problem into precoding and power scaling subproblems. We analyze the properties of the subproblems, and develop a beamforming algorithm for sum-rate maximization. Furthermore, we discuss the case of imperfect CSI, in which MSs measure the channel gains from a limited set of BSs.

*A. Reformulation and Decomposition of **[P-SRM]***

As seen in the formulation **[P-SRM]**, designing a beamformer requires a complicated joint optimization of beam direction and amplitude across multiple BS power budgets and antennas in different BSs. To circumvent the difficulty, we reformulate the problem **[P-SRM]** so that the beam direction and amplitude design problems can be decoupled. First, the precodings for MS k are merged into a single vector $\bar{\mathbf{v}}_k = \left[\mathbf{v}_{mk}^H, \forall m \in \mathcal{M}_k\right]^H$, where $(\cdot)^H$ is the conjugate transpose. In addition, the power scaling factor for MS k is equally controlled by the BSs in $\mathcal{M}_k$ as $p_{mk} = p_k, \forall m \in \mathcal{M}_k$. Such a mobile-centric approach for precoder design and power scaling eases the difficulty caused by arbitrarily formed clusters.

Consider the following beamforming problem:

[P-RSRM] Reformulated SRM

$$\max_{\mathbf{p}\succeq 0,\mathbf{v}} \sum_{k\in\mathcal{K}} w_k \log_2(1+\gamma_k(\mathbf{p},\mathbf{v})) \tag{7}$$

$$\text{s.t. } \|\bar{\mathbf{v}}_k\| = 1, \forall k \in \mathcal{K} \tag{8}$$

$$\sum_{k\in\mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k \leq \bar{P}_m, \forall m \in \mathcal{M} \tag{9}$$

Although the constraints in **[P-RSRM]** are different from those in **[P-SRM]**, the two problems are equivalent.

Proposition 1: **[P-RSRM]** is equivalent to **[P-SRM]**.

Proof: Since the objective functions in both problems are the same, and p_{mk} and $\mathbf{v}_{mk}$ always appear together in the form of $\mathbf{v}_{mk}\sqrt{p_{mk}}$ in $\gamma_k(\mathbf{p},\mathbf{v})$ as seen in (4), we just need to show that every possible $\mathbf{v}_{mk}\sqrt{p_{mk}}$ satisfying (6) can be represented by $\bar{\mathbf{v}}_k$ and p_k subject to (8) and (9), and vice versa.

For any $(\hat{\mathbf{p}},\hat{\mathbf{v}})$ satisfying (6), we can pick $(\tilde{\mathbf{p}},\tilde{\mathbf{v}})$ as follows:

$$\tilde{p}_{mk} = \tilde{p}_k = \sum_{m\in\mathcal{M}_k} \|\hat{\mathbf{v}}_{mk}\|^2 \hat{p}_{mk}, \forall \mathcal{M}_k, k \in \mathcal{K}, \tag{10}$$

$$\tilde{\mathbf{v}}_{mk} = \sqrt{\frac{\hat{p}_{mk}}{\sum_{m\in\mathcal{M}_k} \|\hat{\mathbf{v}}_{mk}\|^2 \hat{p}_{mk}}}\hat{\mathbf{v}}_{mk}, \forall \mathcal{M}_k, k \in \mathcal{K}. \tag{11}$$

Then, $(\tilde{\mathbf{p}},\tilde{\mathbf{v}})$ satisfies $\hat{\mathbf{v}}_{mk}\sqrt{\hat{p}_{mk}} = \tilde{\mathbf{v}}_{mk}\sqrt{\tilde{p}_{mk}}$ subject to (8) and (9). The converse is obvious. ■

By the generalized Benders Decomposition [30], **[P-RSRM]** can be rewritten as

$$\max_{\|\bar{\mathbf{v}}_k\|^2=1,\forall k\in\mathcal{K}} \left[\max_{\mathbf{p}\succeq 0 \text{ s.t. } (9)} \sum_{k\in\mathcal{K}} w_k \log(1+\gamma_k(\mathbf{p},\mathbf{v})) \right]. \tag{12}$$

The outer problem finds a precoder $\bar{\mathbf{v}}_k$ for each MS k, and this corresponds to determining the direction of the beam for MS k. The inner problem finds a transmit power scaling factor p_k for each MS k, and this corresponds to determining the amplitude of the beam for MS k. We will solve the two subproblems alternately to find a beamformer. Note that even if the value of p_k changes, the pointing direction of precoder $\bar{\mathbf{v}}_k$ remains unchanged, and hence, its relationship with channel conditions (such as orthogonality) is unaffected. This decoupling property is critical to our understanding of the structure of optimal precoding and power scaling. Note that this cannot be achieved by the original formulation **[P-SRM]**.

B. Precoding Design for Fixed Power Allocation

We start with the precoding subproblem given a power scaling factor $\mathbf{p}$. First, define a vector $\bar{\mathbf{h}}_{kj} = \left[\mathbf{h}_{km}\sqrt{p_{mj}}, \forall m \in \mathcal{M}_j\right] \in \mathbb{C}^{|\mathcal{M}_j|N_T \times 1}$, i.e., $\bar{\mathbf{h}}_{kj}$ is the aggregated channel states from the BSs in $\mathcal{M}_j$ to MS k. The precoding subproblem in (12) can be rewritten as

[P-Pre] Optimal Precoding Problem

$$\max_{\bar{\mathbf{v}}} \sum_{k\in\mathcal{K}} w_k \log_2(1+\gamma_k(\bar{\mathbf{v}})) \tag{13}$$

$$\text{s.t. } \gamma_k(\bar{\mathbf{v}}) = \frac{|\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2}{\sum_{j\neq k, j\in\mathcal{K}} |\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2 + \sigma_k^2}, \forall k \in \mathcal{K} \tag{14}$$

$$\|\bar{\mathbf{v}}_k\|^2 = 1, \forall k \in \mathcal{K} \tag{15}$$

where $\bar{\mathbf{v}} = [\bar{\mathbf{v}}_k, k \in \mathcal{K}]$. This problem is still non-convex and hard to solve in general. To circumvent the difficulty, we characterize the conditions for optimal precoding, and use the optimality conditions to develop a precoding algorithm for maximizing the sum-rate.

First, the Lagrangian function of **[P-Pre]** is defined as

$$L_{\text{Pre}}(\bar{\mathbf{v}}, \lambda) = \sum_{k\in\mathcal{K}} w_k \log_2 \left(\frac{\sum_{j\in\mathcal{K}} |\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2 + \sigma_k^2}{\sum_{j\neq k, j\in\mathcal{K}} |\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2 + \sigma_k^2} \right) + \sum_{k\in\mathcal{K}} \lambda_k \left(1 - \|\bar{\mathbf{v}}_k\|^2\right) \tag{16}$$

where λ_k is the Lagrangian multiplier for constraint (15). After differentiating with respect to $\bar{\mathbf{v}}_k$, we obtain

$$\frac{\partial L_{\text{Pre}}(\bar{\mathbf{v}}, \lambda)}{\partial \bar{\mathbf{v}}_k} = \frac{w_k/\ln 2}{\sum_{i\in\mathcal{K}} |\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2 + \sigma_k^2} \left(\bar{\mathbf{h}}_{kk}^H \bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k\right) - \sum_{j\in\mathcal{K}, j\neq k} \frac{\gamma_j(\mathbf{v}) w_j/\ln 2}{\sum_{i\in\mathcal{K}} |\bar{\mathbf{h}}_{ji}\bar{\mathbf{v}}_i|^2 + \sigma_j^2} \left(\bar{\mathbf{h}}_{jk}^H \bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k\right) - \lambda_k \bar{\mathbf{v}}_k. \tag{17}$$

By the Karush-Kuhn-Tucker (KKT) condition, the optimal precoding must satisfy

1) $\bar{\mathbf{v}}_k^H \bar{\mathbf{v}}_k = 1, \forall k \in \mathcal{K}$

2) $\left(\mathbf{A}_k(\bar{\mathbf{v}}) - \mathbf{B}_k(\bar{\mathbf{v}})\right) \bar{\mathbf{v}}_k = \lambda_k \bar{\mathbf{v}}_k, \forall k \in \mathcal{K}$

where $\mathbf{A}_k(\bar{\mathbf{v}}) = w_{k,A}\left(\bar{\mathbf{h}}_{kk}^H \bar{\mathbf{h}}_{kk}\right)$ and $\mathbf{B}_k(\bar{\mathbf{v}}) = \sum_{j\neq k} w_{j,B}\left(\bar{\mathbf{h}}_{jk}^H \bar{\mathbf{h}}_{jk}\right)$ with $w_{k,A} = \frac{w_k/\ln 2}{\sum_{i\in\mathcal{K}} |\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2+\sigma_k^2}$ and $w_{j,B} = \frac{\gamma_j(\bar{\mathbf{v}})w_j/\ln 2}{\sum_{i\in\mathcal{K}} |\bar{\mathbf{h}}_{ji}\bar{\mathbf{v}}_i|^2+\sigma_j^2}$. Note that both $\mathbf{A}_k(\cdot)$ and $\mathbf{B}_k(\cdot)$ are Hermitian matrices, i.e., $\mathbf{A}_k^H = \mathbf{A}_k$ and $\mathbf{B}_k^H = \mathbf{B}_k$, so they have only real eigenvalues.

The conditions 1) and 2) show that each $\bar{\mathbf{v}}_k^*$ in the optimal precoding $\bar{\mathbf{v}}^* = [\bar{\mathbf{v}}_k^*, \forall k \in \mathcal{K}]$ must be an eigenvector of Hermitian matrix $\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*)$. We call $\mathbf{A}_k - \mathbf{B}_k$ the *balancing matrix* and will discuss its role later. Motivated by this result, we develop a precoding algorithm that finds an eigenvector of the balancing matrix. Although there are a number of algorithms for finding the eigenvectors of Hermitian matrices, in our case, the balancing matrix $\mathbf{A}_k - \mathbf{B}_k$ is a function of $\bar{\mathbf{v}}$, making it hard to find such a vector $\bar{\mathbf{v}}^*$. Our idea is to solve a fixed point equation iteratively whose limit point is an eigenvector.

Algorithm 1 Eigenprecoding Algorithm

1: **procedure** EIGENPRECODING($\mathbf{p}, \bar{\mathbf{v}}$)
2: Initialize $\bar{\mathbf{h}}_{kj}, \forall k, j \in \mathcal{K}$ for given $\mathbf{p}$
3: **repeat**
4: Update $\mathbf{A}_k$ and $\mathbf{B}_k$, $\forall k \in \mathcal{K}$ for given $\mathbf{v}$
5: $\bar{\mathbf{v}}_k \leftarrow \arg\max_{\|\bar{\mathbf{u}}_k\|^2=1} \bar{\mathbf{u}}_k^H(\mathbf{A}_k - \mathbf{B}_k)\bar{\mathbf{u}}_k$, $\forall k \in \mathcal{K}$
6: **until** $\bar{\mathbf{v}}$ converges or max # of iterations
7: **return** $\bar{\mathbf{v}}$
8: **end procedure**

In particular, our precoding algorithm (shown in Algorithm 1) updates the solution by the eigenvector corresponding to the largest eigenvalue of the current balancing matrix and constructs a new balancing matrix using the updated eigenvector. This is repeated until it converges. The largest eigenvalue and its eigenvector can be computed efficiently using the well-known algorithm such as QR method [31]. Then, each iteration of our precoding algorithm requires $O((|\mathcal{K}|N)^2 + |\mathcal{K}|N^3)$ computations where N is the maximum size of precoding vectors. For a constant cluster size (i.e., constant N), the complexity becomes $O(|\mathcal{K}|^2)$.

While the KKT conditions only require the optimal precoding to be an eigenvector, our algorithm seeks to find the largest eigenvalue and its eigenvector. The reason is that, under some conditions, the algorithm finding the largest eigenvalue can find an optimal precoding.

Proposition 2: Every limit point of the sequence $\{\bar{\mathbf{v}}\}$ generated by the iteration in Algorithm 1 satisfies the KKT conditions.

Proof: Let $\bar{\mathbf{v}}^*$ be a limit point. Then,

$$\bar{\mathbf{v}}_k^* = \arg \max_{\|\bar{\mathbf{u}}_k\|^2=1} \bar{\mathbf{u}}_k^H (\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*))\bar{\mathbf{u}}_k, \ \forall k \in \mathcal{K}. \tag{18}$$

This implies that $\bar{\mathbf{v}}_k^*$ is an eigenvector corresponding to the largest eigenvalue of $\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*)$. Thus, $\bar{\mathbf{v}}^*$ satisfies the KKT conditions. ■

Therefore, whenever Algorithm 1 returns a precoding vector $\bar{\mathbf{v}}$ after it has converged, $\bar{\mathbf{v}}$ is a KKT point. Furthermore, the limit point is an optimal solution of **[P-Pre]** in some cases. Let $\bar{\bar{\mathbf{v}}}_k = [\bar{\mathbf{v}}_k^{\Re T} \bar{\mathbf{v}}_k^{\Im T}]^T$ where $\bar{\mathbf{v}}_k^{\Re}$ and $\bar{\mathbf{v}}_k^{\Im}$ are real part and imaginary part of $\bar{\mathbf{v}}_k$, respectively, i.e., $\bar{\mathbf{v}}_k = \bar{\mathbf{v}}_k^{\Re} + \bar{\mathbf{v}}_k^{\Im} i$. Note that $\bar{\bar{\mathbf{v}}}$ is a double-sized real vector and a one-to-one mapping of $\bar{\mathbf{v}}$. The objective function in **[P-Pre]** also can be defined for the real vector $\bar{\bar{\mathbf{v}}}$ such that $g(\bar{\bar{\mathbf{v}}}) = \sum_{k\in\mathcal{K}} w_k \log_2(1+\gamma_k(\bar{\bar{\mathbf{v}}}))$ where $\gamma_k(\bar{\bar{\mathbf{v}}}) = \frac{(\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Im})^2 + (\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Re})^2}{\sum_{j\neq k}(\bar{\mathbf{h}}_{kj}^{\Re}\bar{\mathbf{v}}_j^{\Re} - \bar{\mathbf{h}}_{kj}^{\Im}\bar{\mathbf{v}}_j^{\Im})^2 + (\bar{\mathbf{h}}_{kj}^{\Re}\bar{\mathbf{v}}_j^{\Im} + \bar{\mathbf{h}}_{kj}^{\Im}\bar{\mathbf{v}}_j^{\Re})^2 + \sigma_k^2}$ instead of (14).

Lemma 1: $\nabla_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{v}}})^T \bar{\bar{\mathbf{u}}} = \sum_{k\in\mathcal{K}} 2\Re\left(\bar{\mathbf{v}}_k^H(\mathbf{A}_k(\bar{\mathbf{v}}) - \mathbf{B}_k(\bar{\mathbf{v}}))\bar{\mathbf{u}}\right)$, where $\bar{\bar{\mathbf{v}}}$ and $\bar{\bar{\mathbf{u}}}$ are corresponding double-sized real vectors of $\bar{\mathbf{v}}$ and $\bar{\mathbf{u}}$, respectively, and $\Re(x)$ is the real part of x.

Proof: This is a straightforward calculation, and the details can be found in [32]. ■

Proposition 3: Suppose that $\bar{\mathbf{v}}^*$ is a limit point of the sequence $\{\bar{\mathbf{v}}\}$ generated by the iteration in Algorithm 1 and C is the maximum eigenvalue of Hessian $\nabla^2_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{u}}})$ for all $\bar{\bar{\mathbf{u}}}$ such that $\|\bar{\bar{\mathbf{u}}}\| \leq 2|\mathcal{K}|$. Then, $\bar{\mathbf{v}}^*$ is a global optimal solution of **[P-Pre]**, if $\lambda_k \geq \frac{C}{2}$, $\forall k \in \mathcal{K}$ where $\lambda_k = \bar{\mathbf{v}}_k^{*H}(\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*))\bar{\mathbf{v}}_k^*$.

Proof: Let $\bar{\mathbf{v}}$ be a feasible solution of **[P-Pre]**, i.e., $\|\bar{\mathbf{v}}_k\| = 1, \forall k \in \mathcal{K}$. Let $\bar{\bar{\mathbf{v}}}$ and $\bar{\bar{\mathbf{v}}}^*$ be corresponding real vectors of $\bar{\mathbf{v}}$ and $\bar{\mathbf{v}}^*$, respectively. By Taylor's theorem,

$$\begin{aligned} g(\bar{\bar{\mathbf{v}}}) = g(\bar{\bar{\mathbf{v}}}^*) &+ \nabla_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{v}}}^*)^T(\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*) \\ &+ \frac{1}{2}(\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*)^T \nabla^2_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{u}}})(\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*) \end{aligned} \tag{19}$$

where $\bar{\bar{\mathbf{u}}}$ is a convex combination of $\bar{\bar{\mathbf{v}}}$ and $\bar{\bar{\mathbf{v}}}^*$. By Lemma 1, we obtain

$$g(\bar{\bar{\mathbf{v}}}) = g(\bar{\bar{\mathbf{v}}}^*) + \sum_{k\in\mathcal{K}} 2\Re\left(\bar{\mathbf{v}}_k^{*H}(\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*))(\bar{\mathbf{v}}_k - \bar{\mathbf{v}}_k^*)\right) + \frac{1}{2}(\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*)^T \nabla^2_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{u}}})(\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*). \tag{20}$$

Since $\bar{\mathbf{v}}_k^*$ is an eigenvector corresponding to the maximum eigenvalue of $\mathbf{A}_k(\bar{\mathbf{v}}^*) - \mathbf{B}_k(\bar{\mathbf{v}}^*)$ and C is an upperbound on the eigenvalue of $\nabla^2_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{u}}})$, we obtain

$$\begin{aligned} g(\bar{\bar{\mathbf{v}}}) &\leq g(\bar{\bar{\mathbf{v}}}^*) - \sum_{k\in\mathcal{K}} 2\lambda_k(1 - \Re(\bar{\mathbf{v}}_k^{*H}\bar{\mathbf{v}}_k)) + \frac{C}{2}\|\bar{\bar{\mathbf{v}}} - \bar{\bar{\mathbf{v}}}^*\|^2 \\ &\leq g(\bar{\bar{\mathbf{v}}}^*) - \sum_{k\in\mathcal{K}} 2\lambda_k(1 - \Re(\bar{\mathbf{v}}_k^{*H}\bar{\mathbf{v}}_k)) + \frac{C}{2}\sum_{k\in\mathcal{K}} 2 - 2\Re(\bar{\mathbf{v}}_k^{*H}\bar{\mathbf{v}}) \\ &= g(\bar{\bar{\mathbf{v}}}^*) - \sum_{k\in\mathcal{K}} (2\lambda_k - C)(1 - \Re(\bar{\mathbf{v}}_k^{*H}\bar{\mathbf{v}}_k)) \leq g(\bar{\bar{\mathbf{v}}}^*). \end{aligned} \tag{21}$$

The last inequality follows from the fact that $\lambda_k \geq \frac{C}{2}, \forall k \in \mathcal{K}$ and $\Re(\bar{\mathbf{v}}_k^{*H}\bar{\mathbf{v}}_k) \leq 1, \forall k \in \mathcal{K}$. Consequently, $g(\bar{\bar{\mathbf{v}}}^*) \geq g(\bar{\bar{\mathbf{v}}}), \forall \bar{\bar{\mathbf{v}}}$ such that corresponding $\bar{\mathbf{v}}$ satisfies $\|\bar{\mathbf{v}}_k\|^2 = 1$, $\forall k \in \mathcal{K}$. This proves that $\bar{\mathbf{v}}^*$ is a global optimal solution of **[P-Pre]**. ■

Although it is hard to characterize when the assumption of Proposition 3 holds, it is clear that finding the maximum eigenvalue increases the chance of obtaining an optimal precoding.

***Interpretation of Eigenprecoding Algorithm*:** Our algorithm operates based on the two matrices $\mathbf{A}_k$ and $\mathbf{B}_k$. Specifically, it seeks to maximize the quadratic potential function $\bar{\mathbf{v}}_k^H(\mathbf{A}_k - \mathbf{B}_k)\bar{\mathbf{v}}_k$, which can be written as $\max \bar{\mathbf{v}}_k^H\mathbf{A}_k\bar{\mathbf{v}}_k - \bar{\mathbf{v}}_k^H\mathbf{B}_k\bar{\mathbf{v}}_k$. Hence, it maximizes $\bar{\mathbf{v}}_k^H\mathbf{A}_k\bar{\mathbf{v}}_k$ while minimizing $\bar{\mathbf{v}}_k^H\mathbf{B}_k\bar{\mathbf{v}}_k$. Recall that $\mathbf{A}_k$ is a weighted covariance matrix of $\bar{\mathbf{h}}_{kk}$, given by $\mathbf{A}_k = w_{k,A}\bar{\mathbf{h}}_{kk}^H\bar{\mathbf{h}}_{kk}$. It is thus clear that maximizing $\bar{\mathbf{v}}_k^H\mathbf{A}_k\bar{\mathbf{v}}_k = w_{k,A}|\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2$ has the effect of increasing the desired signal strength of MS k. Notice that this is what the maximal ratio combining does.

On the other hand, $\mathbf{B}_k$ is the sum of weighted covariance matrices of $\bar{\mathbf{h}}_{jk}$, given by $\mathbf{B}_k = \sum_{j\neq k} w_{j,B}\bar{\mathbf{h}}_{jk}^H\bar{\mathbf{h}}_{jk}$. Each term in $\mathbf{B}_k$ is closely related to the interference from transmitted signal for MS k to other MSs. Consequently, minimizing $\bar{\mathbf{v}}_k^H\mathbf{B}_k\bar{\mathbf{v}}_k$ has the effect of reducing interference. To be more precise, $w_{j,B}$ plays an important role in determining the amount of reduction in interference to each MS. Recall that

$w_{j,B} = \frac{\gamma_j(\bar{\mathbf{v}})/\ln 2}{\sum_{i\in\mathcal{K}}|\bar{\mathbf{h}}_{ji}\bar{\mathbf{v}}_i|^2+\sigma_j^2}$. Suppose that two MSs a and b have the same SINR, e.g., $\gamma_a = \gamma_b = 1$, but the signal strength of a is smaller than b, e.g., $|\bar{\mathbf{h}}_{aa}\bar{\mathbf{v}}_a|^2 = 1$ and $|\bar{\mathbf{h}}_{bb}\bar{\mathbf{v}}_b|^2 = 10$. Then, $w_{a,B} = \frac{1/1}{2}$, which is larger than $w_{b,B} = \frac{10/10}{20}$. Now suppose that interference to both MSs increases by 1. Then, the SINR of MS a reduces by half, i.e., $\gamma_a = \frac{1}{2}$ whereas the SINR of MS b remains almost the same, i.e., $\gamma_b = \frac{10}{11}$. In other words, the SINR change is more sensitive to interference variation for high values of $w_{j,B}$. Following this interpretation, $w_{j,B}$ is called the *interference sensitivity*. Note again that $\mathbf{B}_k$ is the sum of channel covariance matrices from MS k to other MSs weighted by individual interference sensitivities. Consequently, when minimizing $\bar{\mathbf{v}}_k^H\mathbf{B}_k\bar{\mathbf{v}}_k$, interference to MSs with higher interference sensitivity tends to be more reduced than that to MSs with lower sensitivity; so that too much decrease in the SINR of sensitive MSs can be avoided. This part is similar to the zero forcing precoding (and its variants) in the sense that the primary focus is on the interference minimization.

It is clear that in the low SNR regime where the noise power σ_k^2 is dominant, the MRC (maximizing desired signal strength) works well since canceling interference has marginal effect. In contrast, in the high SNR regime where the noise power is negligible, the ZF (canceling interference) performs well since interference is dominant in the denominator of SINR expression. By using the balancing matrix, our precoding algorithm adaptively (to SNR regime) selects a compromise between the two extreme solutions, and as a consequence performs uniformly well over various SNR regimes. We verify this through simulations in Section V.

***Convergence of Eigenprecoding Algorithm*:** It is hard to prove the convergence of the proposed algorithm in general case. Instead, we derive the condition for convergence in the case of two users. This condition provides an insight on the convergence in general case.

Assume that there are two MSs k and j whose channels satisfy the conditions, $\|\bar{\mathbf{h}}_{kk}\| \geq \|\bar{\mathbf{h}}_{jk}\|$ and $\|\bar{\mathbf{h}}_{jj}\| \geq \|\bar{\mathbf{h}}_{kj}\|$. This assumption states that intended channels are better than interference channels (to others). Let $\mathbf{z}_1 = \frac{\bar{\mathbf{h}}_{kk}^H}{\|\bar{\mathbf{h}}_{kk}\|}$, $\mathbf{z}_1' = \frac{\bar{\mathbf{h}}_{jj}^H}{\|\bar{\mathbf{h}}_{jj}\|}$, $\mathbf{z}_2 = \frac{\bar{\mathbf{h}}_{jk}^H-(\mathbf{z}_1^H\bar{\mathbf{h}}_{jk}^H)\mathbf{z}_1}{\|\bar{\mathbf{h}}_{jk}^H-(\mathbf{z}_1^H\bar{\mathbf{h}}_{jk}^H)\mathbf{z}_1\|}$ and $\mathbf{z}_2' = \frac{\bar{\mathbf{h}}_{kj}^H-(\mathbf{z}_1'^H\bar{\mathbf{h}}_{kj}^H)\mathbf{z}_1'}{\|\bar{\mathbf{h}}_{kj}^H-(\mathbf{z}_1'^H\bar{\mathbf{h}}_{kj}^H)\mathbf{z}_1'\|}$. Here, $\mathbf{z}_1$ (or $\mathbf{z}_1'$) is a normalized vector of $\bar{\mathbf{h}}_{kk}^H$ (or $\bar{\mathbf{h}}_{jj}^H$), and $\mathbf{z}_2$ (or $\mathbf{z}_2'$) is a normal vector perpendicular to $\bar{\mathbf{h}}_{kk}^H$ (or $\bar{\mathbf{h}}_{jj}^H$).

Additionally, we assume that $|\bar{\mathbf{h}}_{jk}\mathbf{z}_1| \leq |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|$ and $|\bar{\mathbf{h}}_{kj}\mathbf{z}_1'| \leq |\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|$. This condition is satisfied if intended channels and interference channels are closer to orthogonal than parallel. Then, we can prove the following result.

Proposition 4: The sequence $\bar{\mathbf{v}}$ generated by the iteration in Algorithm 1 converges if channels satisfy the following condition,

$$|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2|\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2\frac{C_1C_2}{\sigma^2} + |\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2|\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2\frac{C_1'C_2'}{\sigma^2} < \frac{1}{2},$$

where

$$\begin{aligned}
C_1 &= \left(2/\sigma^2 + 1/\|\bar{\mathbf{h}}_{kk}\|^2\right)^3 \left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right)^3, \\
C_1' &= \left(2/\sigma^2 + 1/\|\bar{\mathbf{h}}_{jj}\|^2\right)^3 \left(\|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + 1\right)^3, \\
C_2 &= \|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + \|\bar{\mathbf{h}}_{jj}\|^2 + 1, \\
C_2' &= \|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + \|\bar{\mathbf{h}}_{kk}\|^2 + 1.
\end{aligned}$$

Proof: If the channels satisfy the above condition, the update of $\bar{\mathbf{v}}$ given by $\bar{\mathbf{v}}_k \leftarrow \arg\max_{\|\bar{\mathbf{u}}_k\|^2=1} \bar{\mathbf{u}}_k^H(\mathbf{A}_k - \mathbf{B}_k)\bar{\mathbf{u}}_k, \forall k \in \mathcal{K}$, becomes a contraction mapping. For details, refer to Appendix in our technical report [32]. ■

The above convergence condition is satisfied when $|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|$ and $|\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|$ are sufficiently small, i.e., the absolute value of the inner product of interference channels and intended channels are sufficiently small. This happens when the amplitude of interference channels is sufficiently small or when interference channels and intended channels are nearly orthogonal. In this case, interference is negligible or maximizing intended signal strength leads to suppressing interference. As a consequence, our algorithm, which seeks a balance between signal maximization and interference minimization, may quickly converge by merely finding a precoding vector $\bar{\mathbf{v}}$ close to $[\bar{\mathbf{h}}_{kk}\bar{\mathbf{h}}_{jj}]^T$. In fact, it is easy to check that our algorithm converges in one step if interference channels are zero or perfectly orthogonal to intended channels. Note also that the convergence condition is satisfied if the noise power σ^2 is sufficiently large. In this case, interference is negligible and our algorithm would focus on signal maximization so that it eventually converges.

In summary, the above convergence condition is consistent with our intuition, and we expect that our algorithm will converge in general case under similar conditions. Although our proof is limited to the case of two users, simulation results show that our algorithm converges for most of the scenarios where there are more than two MSs. Note further that the condition is only sufficient and we found that our algorithm converges even when the condition is not satisfied.

C. Power Allocation for Fixed Precoding Vector

Next, we solve the power scaling problem, given a precoding vector $\bar{\mathbf{v}}$. Denote by $g_{kj} = |\sum_{m\in\mathcal{M}_j} \mathbf{h}_{km}\mathbf{v}_{mj}|^2$ the aggregated channel gain from the BSs in $\mathcal{M}_j$ to MS k. As mentioned in Section IV-A, we assume the cluster-level power scaling factor control as $p_{mk} = p_k, \forall m \in \mathcal{M}_k$. The power scaling subproblem in (12) is written as

[P-Pow] Optimal Power Allocation Problem

$$\max_{\mathbf{p}\succeq 0} \sum_{k\in\mathcal{K}} w_k \log_2(1+\gamma_k(\mathbf{p})) \tag{22}$$

$$\gamma_k(\mathbf{p}) = \frac{g_{kk}p_k}{\sum_{j\neq k, j\in\mathcal{K}} g_{kj}p_j + \sigma_k^2}, \forall k \in \mathcal{K} \tag{23}$$

$$\sum_{k\in\mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k \leq \bar{P}_m, \forall m \in \mathcal{M}. \tag{24}$$

This is a non-convex optimization problem extensively studied in the literature. Our power scaling algorithm is based on the well-known iterative water-filling (WF) algorithm [22]–[24]. The key idea is to regard interference from MS k to others and interference from others to MS k as constants and compute the transmit power scaling factor for MS k using the KKT condition. The new power scaling factor changes interference levels, and the power scaling factor is recomputed with the new interference levels. This iteration continues until it converges.

To apply this idea to our problem, define the Lagrangian function as

$$\begin{aligned} L_{\text{Pow}}(\mathbf{p},\mu) = &\sum_{k\in\mathcal{K}} w_k \log_2\left(\frac{\sum_{j\in\mathcal{K}} g_{kj}p_j + \sigma_k^2}{\sum_{j\neq k, j\in\mathcal{K}} g_{kj}p_j + \sigma_k^2}\right) \\ &+ \sum_{m\in\mathcal{M}} \mu_m \left(\bar{P}_m - \sum_{k\in\mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k\right), \end{aligned} \tag{25}$$

where μ_m is the Lagrangian multiplier for the power budget constraint of BS m (24). After differentiating with respect to p_k, we obtain

$$\frac{\partial L_{\text{Pow}}(p,\mu)}{\partial p_k} = \frac{w_k/\ln 2}{g_{kk}p_k + \sum_{i\in\mathcal{K},i\neq k} g_{ki}p_i + \sigma_k^2} g_{kk} - \sum_{j\in\mathcal{K},j\neq k} \frac{\gamma_j w_j/\ln 2}{\sum_{i\in\mathcal{K}} g_{ji}p_i + \sigma_j^2} g_{jk} - \sum_{m\in\mathcal{M}_k} \mu_m \|\mathbf{v}_{mk}\|^2. \quad (26)$$

By using the KKT condition, it can be shown that the optimal power scaling factor must satisfy the following conditions:

1) $\mu_m \left(\bar{P}_m - \sum_{k\in\mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k\right) = 0, \forall m \in \mathcal{M}$

2) $\mu_m \geq 0, \sum_{k\in\mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k \leq \bar{P}_m, \forall m \in \mathcal{M}$

3) $t_k(\mathbf{p}) + \sum_{m\in\mathcal{M}_k} \|\mathbf{v}_{mk}\|^2 \mu_m = \frac{w_k/\ln 2}{p_k + r_k(\mathbf{p})}, \forall k \in \mathcal{K}$

where $t_k(\mathbf{p}) = \sum_{j\neq k, j\in\mathcal{K}} \tilde{w}_j g_{kj}$ with $\tilde{w}_j = \frac{\gamma_j w_j/\ln 2}{\sum_{i\in\mathcal{K}} g_{ji}p_i + \sigma_j^2}$ and $r_k(\mathbf{p}) = \frac{\sum_{i\neq k, i\in\mathcal{K}} g_{ki}p_i + \sigma^2}{g_{kk}}$. Solving 3) for p_k and applying the non-negativity constraint yield

$$p_k = \left[\frac{w_k/\ln 2}{t_k(\mathbf{p}) + \sum_{m\in\mathcal{M}_k} \|\mathbf{v}_{mk}\|^2 \mu_m} - r_k(\mathbf{p})\right]^+, \forall k \quad (27)$$

where $[\cdot]^+$ denotes the projection onto the set of non-negative real numbers. Since $t_k(\mathbf{p})$ and $r_k(\mathbf{p})$ are also defined by $\mathbf{p}$, it is hard to directly find $\mathbf{p}$ satisfying (27). Following the idea of iterative WF, our cluster-level WF algorithm in Algorithm 2 finds $\mathbf{p}$ using (27) for given t_k and r_k, updates t_k and r_k by the new $\mathbf{p}$, and repeats this until it converges. Note that solving (27) for given t_k and r_k requires to determine the values of μ_m subject to the constraints (24). Since p_k is a monotonic function of μ_m, the multi-dimensional bisection search can be used to solve (27) for given t_k and r_k (see steps 6-12). Then, each iteration of our power scaling algorithm requires $O(|\mathcal{K}|^2 + |\mathcal{M}||\mathcal{K}| \log_2(\bar{\mu}/\delta))$ computations where $\bar{\mu}$ is an upper bound on dual variables and δ is a small positive constant for the termination condition.

Note that r_k is large when MS k experiences high interference, and t_k is large when MS k generates severe interference to other MSs. Therefore, p_k solving (27) is inversely proportional to t_k and r_k, and thus becomes small when MS k experiences and generates high interference, because in this case allocating high power to MS k would result in the decrease or marginal increase in the sum-rate.

Algorithm 2 Cluster-level WF Algorithm

procedure CLUSTERWF($\mathbf{p}, \bar{\mathbf{v}}$)

  Initialize $g_{jk}, \forall j, k \in \mathcal{K}$ for given $\bar{\mathbf{v}}$

  repeat

    Update t_k and r_k, $\forall k \in \mathcal{K}$ for given $\mathbf{p}$

    Set $\mu_m = \frac{\bar{\mu}}{2}, [a_m, b_m] = [0, \bar{\mu}], \quad \forall m \in \mathcal{M}$

    repeat

      $p_k \leftarrow \left[\frac{w_k}{t_k + \sum_{m \in \mathcal{M}_k} \|\mathbf{v}_{mk}\|^2 \mu_m} - r_k \right]^+, \forall k \in \mathcal{K}$

      $\mu_m \leftarrow \frac{a_m + b_m}{2}, \quad \forall m \in \mathcal{M}$

      $P_m \leftarrow \sum_{k \in \mathcal{K}_m} \|\mathbf{v}_{mk}\|^2 p_k, \forall m \in \mathcal{M}$

      if $P_m > \bar{P}_m$ then $a_m \leftarrow \mu_m, \forall m \in \mathcal{M}$

      if $P_m < \bar{P}_m$ then $b_m \leftarrow \mu_m, \forall m \in \mathcal{M}$

    until $\bar{P}_m - P_m < \delta$ or $\sum_m \mu_m < \delta$

  until $\mathbf{p}$ converge or max # of iterations

  return $\mathbf{p}$

end procedure

The convergence of iterative water-filling type algorithms such as Cluster-level WF algorithm is difficult to prove in general case [24]. However, such a form of iterative water-filling algorithm is known to converge when interference is weak (see [24] and references therein). We expect that our Cluster-level WF algorithm converges under a similar weak interference condition. Note that this is somewhat consistent with the convergence condition of Eigenprecoding algorithm. In simulations, we could observe that the algorithm converges most of the time, even when interference is fairly strong.

D. Description of Iterative Balanced Beamforming

We now present our joint power scaling and precoding algorithm, called the iterative balanced beamforming (IBB), in Algorithm 3. The MRC precoding scheme and equal power allocation are used for initialization. In the first iteration, $\bar{\mathbf{v}}$ is updated based on the current precoding vector and power scaling

factor by using eigenprecoding algorithm (in Algorithm 1), and next, under the updated precoding vector and given power scaling factor, $\mathbf{p}$ is updated using cluster-level WF algorithm (in Algorithm 2). In the next iteration, $\bar{\mathbf{v}}$ and $\mathbf{p}$ are updated in the same manner, with the newly updated precoding vector and power scaling factor. This procedure is repeated until the precoding vector $\bar{\mathbf{v}}$ and power scaling factor $\mathbf{p}$ converge or the maximum number of iterations is reached. If Eigenprecoding algorithm finds an optimal solution of **[P-Pre]** for given power scaling factor and Cluster-level WF algorithm finds an optimal solution of **[P-Pow]** for given precoding vector, then IBB algorithm generates a bounded monotone increasing sequence of the weighted sum-rate, and thus it converges. Although it is hard to prove when two algorithms find optimal solutions of each problem, simulation results show that for most of the time it converges in about 10-20 iterations, even when the maximum number of iterations in eigenprecoding and cluster-level WF algorithms is set to 1.

Algorithm 3 Iterative Balanced Beamforming

1: Initialize Precoder $\bar{\mathbf{v}}$ & Power $\mathbf{p}$
2: **repeat**
3: $\quad\bar{\mathbf{v}} \leftarrow$EIGENPRECODING$(\mathbf{p}, \bar{\mathbf{v}})$ ▷ Algorithm 1
4: $\quad\mathbf{p} \leftarrow$CLUSTERWF$(\mathbf{p}, \bar{\mathbf{v}})$ ▷ Algorithm 2
5: **until** $\bar{\mathbf{v}}$ & $\mathbf{p}$ converge or max # of iterations

E. IBB with Local Information

The IBB algorithm requires global channel state information (CSI) from every BS to every MS in order to mitigate interference to other MSs all over the network. However, since local interference from neighbor cells or clusters is dominant [23], it may be sufficient to take into account interference toward neighbor cells or clusters instead of the whole network.

To reduce overhead for channel measurement and feedback, we introduce C-cluster $\hat{\mathcal{M}}_k \supseteq \mathcal{M}_k$ for each MS k, defined as the set of BSs that MS k meaures CSI from. Note that at least CSI from BSs in $\mathcal{M}_k$ is required for the coordinated multi-cell transmission. Define the actually measured channel vector

$\hat{\mathbf{h}}_{km}$ as

$$\hat{\mathbf{h}}_{km} = \begin{cases} \mathbf{h}_{km}, & \text{if } m \in \hat{\mathcal{M}}_k, \\ \overrightarrow{0}_{1\times N_T}, & \text{if } m \notin \hat{\mathcal{M}}_k. \end{cases} \tag{28}$$

Clearly, the channel vector $\hat{\mathbf{h}}_{km}$ provides an accurate approximation of the actual channel vector if the BSs in $\mathcal{M}\backslash\hat{\mathcal{M}}_k$ are far from MS k. Hence, in this case, our algorithm will perform as well as in the case of perfect CSI.

With local information, we redefine the aggregated channel from the BSs in $\mathcal{M}_j$ to MS k for precoding and power scaling as

$$\bar{\mathbf{h}}_{kj}^{\text{loc}} = \left[\hat{\mathbf{h}}_{km}\sqrt{p_{mj}}, \forall m \in \mathcal{M}_j\right],\ g_{kj}^{\text{loc}} = \left| \sum_{m\in\mathcal{M}_j} \hat{\mathbf{h}}_{km}\mathbf{v}_{mj} \right|^2.$$

Accordingly, the matrix $\mathbf{B}_k$ and the value t_k should be redefined as

$$\mathbf{B}_k^{\text{loc}} = \sum_{j\neq k, j\in\mathcal{K}} w_{j,B}(\bar{\mathbf{h}}_{jk}^{\text{loc}H}\bar{\mathbf{h}}_{jk}^{\text{loc}}),\ t_k^{\text{loc}} = \sum_{j\neq k, j\in\mathcal{K}} \tilde{w}_j g_{kj}^{\text{loc}}.$$

Note that $\mathbf{A}_k$ and r_k are unchanged. With these modified parameters, our algorithm works exactly the same way as in the case of perfect CSI. In Section V-F, we study the impact of imperfect (or local) CSI on the performance of our algorithm, and show that with a moderate amount of CSI, our algorithm performs comparably to the case of perfect CSI.

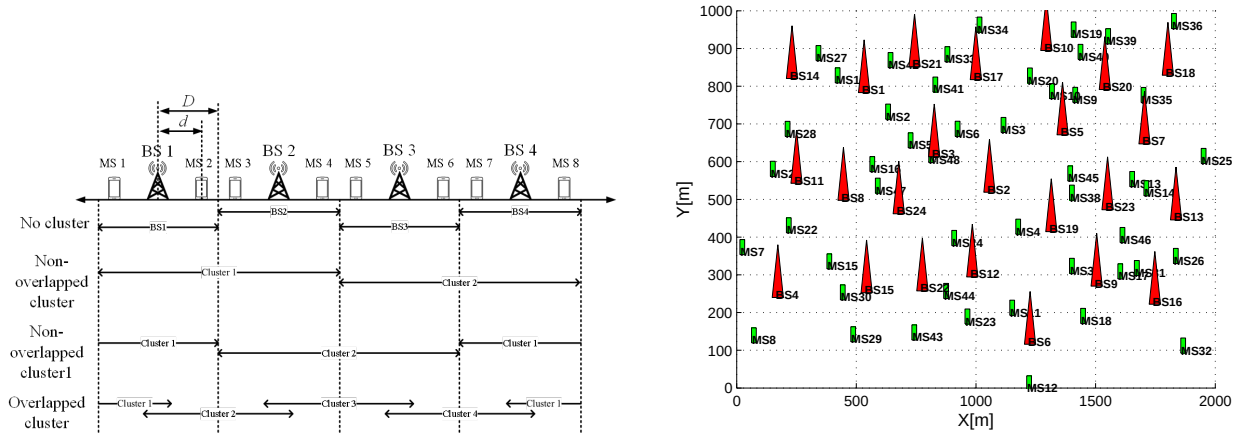
V. Performance Evaluation

In this section, we evaluate the performance of the proposed IBB algorithm in various scenarios. In particular, we demonstrate that our algorithm outperforms existing algorithms under overlapped virtual cells as well as non-overlap virtual cells, and in various SNR regimes. In addition, we confirm the potential of overlapped virtual cells for enhancing edge user performance.

A. Simulation Setup

We use the two topologies shown in Fig. 2. First, in the linear topology in Fig. 2(a), 4 BSs and 8 MSs are located linearly. The distance between MS and its nearest BS is d and the cell radius is D. In

the large topology in Fig. 2(b), 24 BSs and 48 MSs are randomly located in the 2 km$\times$1 km plane. In both topologies, each BS has 23dBm power budget and is equipped with 2 antennas, and each MS is equipped with a single antenna. The system bandwidth is 5 MHz. The standard deviation of shadowing is 8 dB, and the path loss is given by $128.1 + 37.6 \log_{10}(d)$ where d is the distance from BS to MS in km. We use the wrap-around technique to eliminate the boundary effect. Each entry of multi-antenna channel coefficients, $[\alpha_{km,1}, \alpha_{km,2}, ...\alpha_{km,N_T}]$, is modeled as a zero mean complex Gaussian random variable. We assume that the background noise is -169 dBm/Hz and the noise figure is 7dB. In **[P-SRM]**, the weight of each MS is equal to 1. As a performance metric, we use the throughput during 500 time slots instead of the instantaneous data rate. Note that the instantaneous data rate is not suitable for the performance comparison among individual MSs since some users may yield too high or too low data rate under a given channel state due to the randomness. Channel states are newly generated for every time slot.



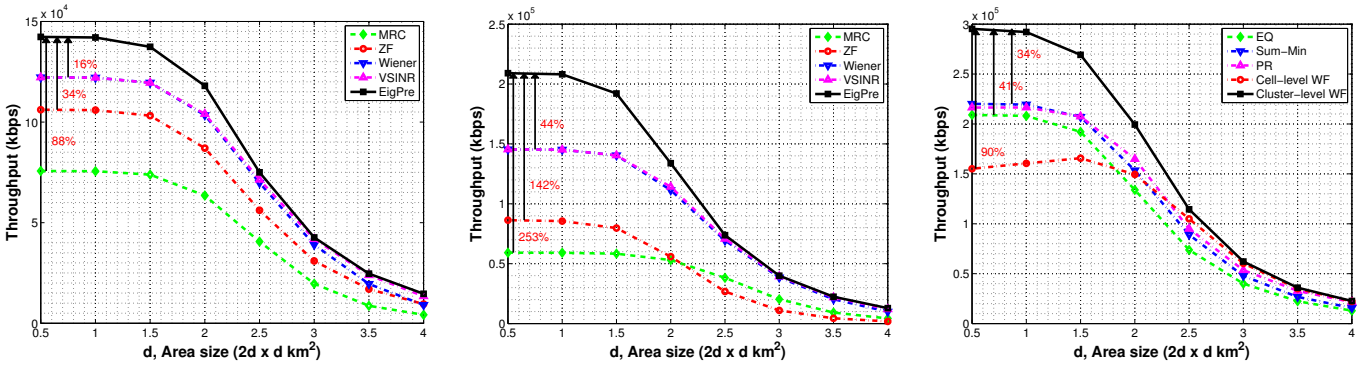
(a) Linear topology

(b) Large topology

Fig. 2. Simulation topologies

B. Performance Comparison with Existing Algorithms

First, we compare our algorithms with existing algorithms using the large topology in Fig. 2(b). The topology is scaled up and down in order to simulate various SINR environments. For instance, as the scaling factor d increases, the distance (and the network size as well) between every pair of MSs/BSs increases by a factor of d, which would reduce signal and interference due to path loss. If d decreases,



(a) Sum-rate comparison of precoding schemes with equal power allocation: no clustering

(b) Sum-rate comparison of precoding schemes with equal power allocation: clustering with size 2

(c) Sum-rate comparison of power allocation schemes with eigenprecoding algorithm: clustering with size 2

Fig. 3. Sum-rate comparison of beamforming schemes in the large topology

everything gets closer, thereby increasing signal and interference. The following precoding algorithms are compared with our eigenprecoding algorithm (EigPre) assuming equal power allocation:

- MRC: Maximal ratio combining that maximizes desired signal [10]
- ZF: Zero-forcing that nullifies interference [14]
- Wiener: MMSE based Wiener filter [25]
- VSINR: Virtual SINR maximization [18]

In addition, our power scaling algorithm (cluster-level WF) is compared to the following power allocation algorithms:

- EQ: Equal power allocation, i.e., $p_{mk} = \frac{1}{|\mathcal{K}_m|}\bar{P}_m$
- Sum-Min: Power allocation by iterative sum-min algorithm [25]
- PR: Power allocation proportional to effective channel gain [18], i.e., $p_{mk} = \frac{\|\mathbf{h}_{km}\|^2}{\sum_{j\in\mathcal{K}_m}\|\mathbf{h}_{jm}\|^2}\bar{P}_m$
- Cell-level WF: Conventional water-filling algorithm[5] for the single-cell transmission [22], [23]

We assume a heuristic clustering method such that each MS selects x nearest BSs, in which case x is the cluster size. Selected BSs form a virtual cell for each MS, and it can be overlapped with other virtual cells.

[5]To apply the conventional water-filling algorithm in the coordinated multi-cell transmission, we approximate the norm of sum by the sum of norm, as $|\sum_{m\in\mathcal{M}_k}\mathbf{h}_{km}\mathbf{v}_{mk}\sqrt{p_{mk}}|^2 \approx \sum_{m\in\mathcal{M}_k}|\mathbf{h}_{km}\mathbf{v}_{mk}|^2 p_{mk}$.

Figs. 3(a) and 3(b) show the throughput of each precoding algorithm under equal power allocation. Our precoding algorithm achieves much higher throughput than the other algorithms without and with coordinated multi-cell transmission. Note that the throughput of MRC and ZF decrease slightly with clustering. This result shows that merely adopting coordinated multi-cell transmission (or clustering) without appropriate precoding and power allocation may worsen the network performance. Although the throughput of Wiener filter and VSINR is enhanced with clustering, it is outperformed by our algorithm.

Our cluster-level WF also significantly improves the throughput compared to the other power allocation algorithms as shown in Fig. 3(c). Although the cell-level WF can achieve near-optimal performance in the single-cell transmission [22], [23], it performs even worse than the equal power allocation in the dense scenarios (i.e., small d). This is because the individual control of p_{mk} can change the direction of the beam that is supposedly good computed by the precoding algorithm. Note that our cluster-level WF does not change the direction of given beam. Fig. 3 also shows that all the algorithms achieve about the same throughput in the large scenario (i.e., large d), since the noise term which is not controllable becomes relatively dominant. Overall, our eigenprecoding algorithm and cluster-level WF algorithm outperform the existing algorithms in various SINR environments.

C. Evaluation of Eigenprecoding over the Whole SNR Regime

Next, we demonstrate that our precoding algorithm can balance signal and interference adaptively to SNR values so as to maximize the throughput over the whole SNR regime. We use the linear topology in Fig. 2(a). To change SNR, we vary the noise power. To focus on the precoding performance, we assume the equal power allocation and the single-cell transmission. Again, our algorithm is compared to the following four precoding schemes, MRC [10], ZF [14], Wiener filter [25] and VSINR [18].

Fig. 4 shows three notable results. First, MRC achieves larger throughput than ZF in the low SNR regime. Second, ZF achieves larger total throughput than MRC in the high SNR regime. Last, Wiener filter, VSINR and our algorithm are superior to both MRC and ZF over the whole SNR regime. The reason for observed results is quite intuitive. In the low SNR regime, the denominator of SINR is dominated

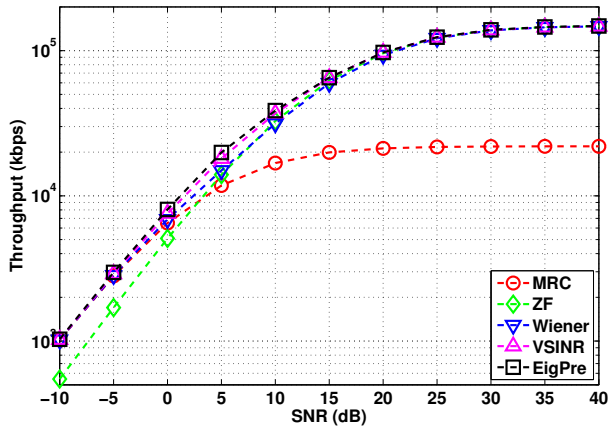


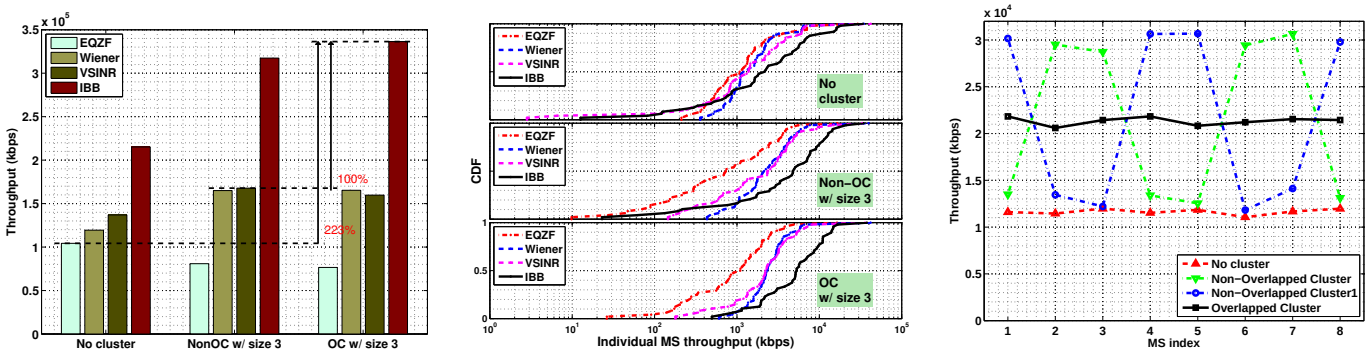
Fig. 4. Sum-rate comparison of precoding schemes for the whole SNR regime in the linear topology with $d = 40$ and $D = 100$

by the noise, and hence, most of the benefits are achieved by maximizing the desired signal rather than minimizing (negligible) interference. On the contrary, in the high SNR regime, minimizing interference becomes critical in sum-rate maximization as it is a dominant factor. As discussed in Section IV-B, our algorithm can choose a balance between maximizing desired signal and minimizing interference, and as a result, achieves high throughput regardless of SNR values.

It is worth mentioning that although VSINR finds a solution as good as ours over the whole SNR regime, this is a restricted scenario where the virtual SINR well approximates the actual SINR. If MSs are located asymmetrically, then intercell interference will also be asymmetric, and as a consequence, the throughput decreases due to inaccurate approximations. In fact, this was shown in Fig. 3 where our algorithm performs better than VSINR.

D. Evaluation of IBB under Various Clustering Policies

We demonstrate the performance of our algorithm under the three clustering policies to form virtual cells including no cluster, non-overlapped clusters (Non-OC) with size 3 and overlapped clusters (OC) with average size 3. For non-overlapped clustering, we divide the 2 km$\times$1 km plane in Fig. 2(b) into eight 0.5 km$\times$0.5 km squares, in each of which BSs form one cluster. For overlapped clustering, we use the same heuristic policy in Section V-B. The zero forcing with equal power allocation (EQZF) [14], Wiener filter with sum-min algorithm based power allocation (Wiener) [25] and virtual SINR maximizing with proportional power allocation (VSINR) [18] are used for comparison with our iterative balanced



(a) System throughput achieved by three beamforming algorithms under three clustering policies in the large topology

(b) Individual MS throughput CDF of three beamforming algorithms under three clustering policies in the large topology

(c) Individual MS throughput under four clustering policies in the linear topology with $d = 80$ and $D = 100$: IBB

Fig. 5. Throughput analysis with various clustering policies

beamforming (IBB) algorithm. In addition, we study the effect of overlapped virtual cells on the edge users.

Fig. 5(a) shows that our algorithm significantly improves the system throughput compared to the other algorithms all over the three clustering policies. In particular, the result with overlapped clustering shows the highest gain. Fig. 5(b) depicts the cumulative distribution function (CDF) of individual MS throughput. Our algorithm with overlapped clustering offers far better edge user performance[6] than any other combinations of beamforming and clustering except Wiener (bottom of Fig. 5(b)). However, Wiener is outperformed by IBB in total system throughput. This is an impressive result since it is difficult to simultaneously enhance both efficiency and fairness in the wireless network.

To highlight the effect of allowing overlapped virtual cells, we run the same simulation in the linear topology with four clustering policies, denoted by No cluster, Non-overlapped cluster, Non-overlapped cluster1 and Overlapped cluster (see Fig 2(a) for how BSs are clustered). Fig. 5(c) shows that all the three clustering policies except no clustering achieve almost equal system throughput. However, under the two non-overlapped clustering policies, the throughput of edge users (MSs 1, 4, 5 and 8 in non-overlapped

[6]As shown in Fig. 5(b), edge user performance of our algorithm with no clustering or non-overlapped clustering may be worse than the other algorithms. This is because our objective is the sum-rate maximization without fairness constraints. The issue of fairness is studied in the next section.

cluster, and MSs 2, 3, 6 and 7 in non-overlapped cluster1) exhibits marginal increase compared with the case of no clustering. On the other hand, the edge user performance is significantly enhanced with overlapped clustering. In some scenarios such as the one in Fig. 2(a), the notion of edge users is blurred under overlapped clustering, since each MS may form a cluster of any BSs that will perform coordinated transmissions for the MS. This is why every MS achieves nearly equal throughput without much loss of system throughput. In summary, the results in Fig. 5 show the potential of overlapped virtual cells for boosting capacity with fairness guarantee, and the necessity of beamforming algorithms that can work well with overlapped virtual cells, such as ours.

E. Evaluation of IBB under different fairness criteria

In the previous section, we showed that overlapped clustering can help increase fairness among MSs, even when the objective is to maximize capacity without fairness consideration. In this section, we further study the fairness property of our algorithm. We adopt the α-proportional fair policy [29] by setting the weight of each MS k to $w_k = (1/R_k)^\alpha$ where R_k is the running average throughput of MS k and α is a fairness parameter. The simulation is run for four fairness parameters (0, 0.5, 1 and 5) under three different clustering policies in the large topology.

Fig. 6 and TABLE I show that, as α increases, the throughput allocation becomes fairer while the total throughput decreases. Note from TABLE I that fairness index and 5 percentile user throughput are improved via clustering. It is clear that for a given clustering policy, the value of alpha controls the trade-off between fairness and total throughput. To achieve better fairness by using large alpha, the total throughput is reduced, and vice versa. It is remarkable that for a given value of alpha, both fairness and total throughput are enhanced via clustering, compared with the case of no clustering. This again demonstrates that coordinated multi-cell transmissions can potentially be a promising solution for boosting system throughput as well as fairness among MSs. Note that for most of the cases, overlapped clustering outperforms non-overlapped clustering in both throughput and fairness.

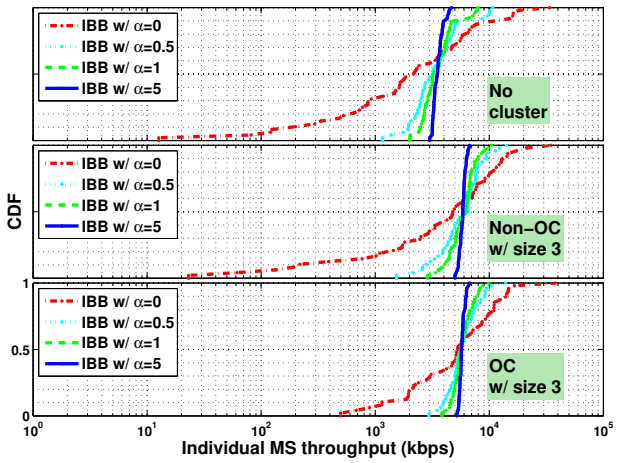


Fig. 6. Individual MS throughput CDF of IBB with different fairness criteria and three clustering policies in the large topology

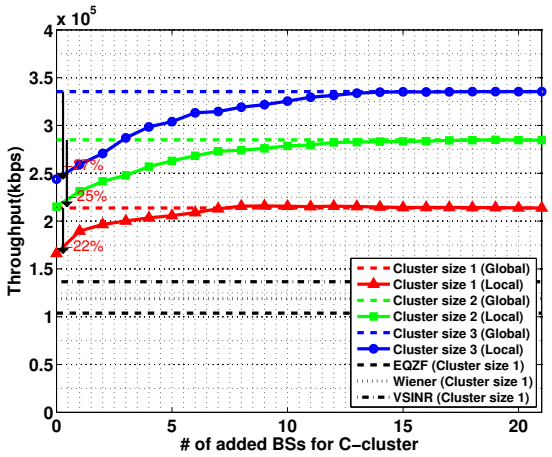


Fig. 7. Impact of local channel state information under three different cluster sizes in the large topology

F. Impact of Local Channel State Information

So far, we have assumed perfect CSI, which may be infeasible in practice. In this section, we study the impact of the amount of CSI on the performance of our algorithm. Recall that the C-cluster of MS k, $\hat{\mathcal{M}}_k$, is the set of BSs which MS k measures CSI from. To gradually increase the amount of CSI, we add a BS one by one to the C-cluster of each MS, starting from the smallest C-cluster, $\hat{\mathcal{M}}_k = \mathcal{M}_k$. The added BS is the nearest BS among the BSs not yet included in $\hat{\mathcal{M}}_k$. The simulation is run for cluster sizes from 1 to 3.

As expected, the throughput increases as the amount of CSI increases for all of the three cluster sizes

[7]Jain's fairness index J is defined as $J = \frac{(\sum_{k\in\mathcal{K}} R_k)^2}{|\mathcal{K}| \sum_{k\in\mathcal{K}} (R_k)^2}$ where R_k is the throughput of MS k. J is a positive number no greater than 1, and the larger the value of J, the fairer the throughput allocation.

TABLE I

TOTAL THROUGHPUT, FAIRNESS INDEX AND 5 PERCENTILE USER AVERAGE THROUGHPUT UNDER DIFFERENT FAIRNESS PARAMETERS α AND CLUSTERING POLICIES IN THE LARGE TOPOLOGY

Fairness parameter (α)	0	0.5	1	5
	Total throughput (Mbps)			
No cluster	215	186	180	171
Non-OC with size 3	317	295	290	281
OC with size 3	336	295	287	278
	Jain's fairness index[7] [33]			
No cluster	0.345	0.741	0.875	0.989
Non-OC with size 3	0.514	0.849	0.935	0.996
OC with size 3	0.562	0.901	0.959	0.997
	5 percentile user average throughput (Mbps)			
No cluster	0.047	1.268	2.021	3.074
Non-OC with size 3	0.042	1.838	3.132	5.064
OC with size 3	0.589	3.217	4.005	5.263

(see Fig. 7), since it is possible to beamform toward a direction that suppresses more interference. When beams are formed using only the minimum amount of CSI, there is a performance degradation of 22%-27%. However, even in this case, our algorithm performs better than zero-forcing with the equal power allocation (EQZF), Wiener filter with sum-min algorithm based power allocation (Wiener) and virtual SINR maximizing with proportional power allocation (VSINR). Note that with CSI from 6 additional BSs, our algorithm achieves the sum-rate comparable to that with global CSI. This result shows that with only a moderate amount of information our algorithm is able to find a beamforming that effectively suppresses interference and increases the throughput.

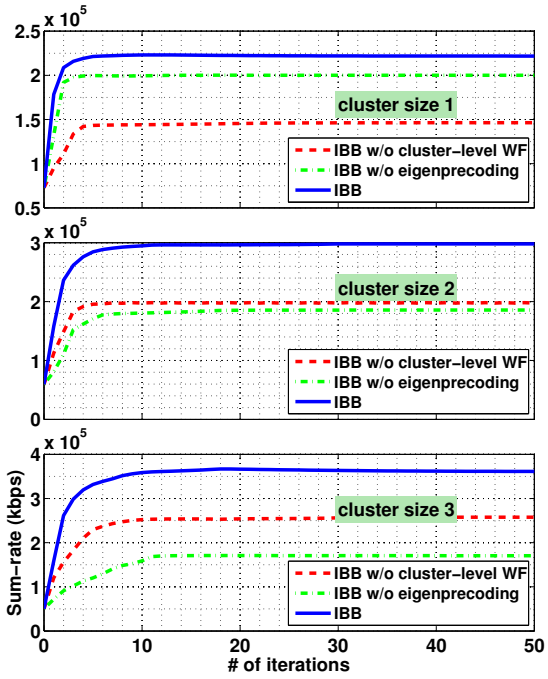


Fig. 8. Convergence of IBB (combined precoding and power scaling iterations), IBB without cluster-level WF (only precoding iteratons) and IBB without eigenprecoding (only power scaling iterations) in the large topology

G. Convergence of IBB

Finally, we test the convergence of our IBB algorithm and its variants: IBB without cluster-level WF, which has only iterations for precoding, and IBB without eigenprecoding, which has only iterations for power scaling. For a randomly generated channel state, the simulation is run over cluster sizes from 1 to 3 in the large topology.

Fig. 8 shows that our IBB algorithm and its variants require a few iterations for convergence. Furthermore, significant throughput improvement is achieved in about 5 iterations, and after that, there is a marginal improvement. Thus, if the running time is of concern, we may run the algorithm for a small number of iterations (before converging), without significant loss of sum-rate compared with the fully convergent case. Fig. 8 also shows that the precoding algorithm becomes more important than the power scaling algorithm as the cluster size increases.

VI. Conclusion

In this paper, we studied the beamforming problem for sum-rate maximization in the virtual cell network. We reformulated the sum-rate maximization problem into an equivalent joint precoder design

and power scaling problem. For precoding, we introduced the so-called balancing matrix that captures the interplay between signal maximization and interference minimization toward sum-rate maximization. Our precoding algorithm finds an eigenvector of the so-called balancing matrix. Moreover, we proposed a cluster-level water-filling algorithm that works well with overlapped clustering. Simulation results show that the proposed beamforming algorithm outperforms the existing algorithms in various BS cooperation environments. Using our algorithm, we also demonstrated the potential of overlapped virtual cells for boosting capacity while enhancing edge user performance, which is an attractive result for future cellular systems.

REFERENCES

[1] Cisco System, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017*. A Cisco White Paper, Feb. 2013.

[2] Nokia, *Nokia Solutions and Networks Looking ahead to 5G*. NSN White paper, Dec. 2013.

[3] Ericsson, *5G Radio Access, Research and Vision*. Ericsson White paper, June 2013.

[4] *White Paper on 5G radio network architecture*. Radio Access and Spectrum, Future Networks Cluster White paper, http://www.ict-ijoit.eu, March 2014.

[5] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan, "Trends in small cell enhancements in lte advanced," *Commun. Magazine, IEEE*, vol. 51, no. 2, pp. 98–105, 2013.

[6] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell mimo cooperative networks: A new look at interference," *IEEE JSAC*, vol. 28, no. 9, pp. 1380–1408, 2010.

[7] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in lte-advanced: deployment scenarios and operational challenges," *Commun. Magazine, IEEE*, vol. 50, no. 2, pp. 148–155, 2012.

[8] T. . 3GPP, "Coordinated multi-point operation for lte physical layer aspects," Dec. 2011.

[9] T. Biermann, L. Scalia, C. Choi, H. Karl, and W. Kellerer, "Comp clustering and backhaul limitations in cooperative cellular mobile access netowrks," *Pervasive and Mobile Computing*, vol. 8, no. 5, pp. 662–681, Oct. 2012.

[10] T. K. Y. Lo, "Maximum ratio transmission," in *Communications, 1999. ICC '99. 1999 IEEE International Conference on*, vol. 2, 1999, pp. 1310–1314.

[11] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *Signal Processing, IEEE Transactions on*, vol. 52, no. 2, pp. 461–471, 2004.

[12] Z. Shen, R. Chen, J. Andrews, R. Heath, and B. Evans, "Sum capacity of multiuser mimo broadcast channels with block diagonalization," *Wireless Commun., IEEE Transactions on*, vol. 6, no. 6, pp. 2040–2045, 2007.

[13] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. Heath, "Networked mimo with clustered linear precoding," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 4, pp. 1910–1921, 2009.

[14] S. Kaviani and W. Krzymien, "Multicell scheduling in network mimo," in *proc. IEEE GLOBECOM*, 2010, pp. 1–5.

[15] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1748–1759, 2010.

[16] W. Yu, T. Kwon, and C. Shin, "Multicell coordination via joint scheduling, beamforming and power spectrum adaptation," in *INFOCOM*, 2011, pp. 2570–2578.

[17] A. Tolli, H. Pennanen, and P. Komulainen, "Sinr balancing with coordinated multi-cell transmission," in *IEEE Wireless Communications and Networking Conference*, 2009, pp. 1–6.

[18] J. Gong, S. Zhou, L. Geng, M. Zheng, and Z. Niu, "A novel precoding scheme for dynamic base station cooperation with overlapped clusters," *IEICE Trans. Commu.*, vol. 96, no. 2, pp. 656–659, Feb. 2013.

[19] S.-H. Park, H. Park, H. Kong, and I. Lee, "New beamforming techniques based on virtual sinr maximization for coordinated multi-cell transmission," *IEEE Trans. Wirel. Comm.*, vol. 11, no. 3, pp. 1034–1044, 2012.

[20] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted mmse approach to distributed sum-utility maximization for a mimo interfering broadcast channel," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4331–4340, 2011.

[21] M. Hong, R. Sun, H. Baligh, and Z.-Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE JSAC*, vol. 31, no. 2, pp. 226–240, 2013.

[22] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *Signal Processing, IEEE Transactions on*, vol. 53, no. 2, pp. 686–695, 2005.

[23] K. Son, S. Lee, Y. Yi, and S. Chong, "Refim: A practical interference management in heterogeneous wireless access networks," *Selected Areas in Commun., IEEE Journal on*, vol. 29, no. 6, pp. 1260–1272, 2011.

[24] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Information Theory and Applications Workshop*, 2007, pp. 414–420.

[25] D. Hui, "Distributed precoding with local power negotiation for coordinated multi-point transmission," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1–5.

[26] C. Mobile, "C-ran: the road towards green ran," *White Paper*, 2011.

[27] E. Larsson, D. Danev, and E. Jorswieck, "Asymptotically optimal transmit strategies for the multiple antenna interference channel," in *Annual Allerton Conference*, 2008, pp. 708–714.

[28] T. Bogale and L. Vandendorpe, "Weighted sum rate optimization for downlink multiuser mimo coordinated base station systems: Centralized and distributed algorithms," *Signal Processing, IEEE Transactions on*, vol. 60, no. 4, pp. 1876–1889, April 2012.

[29] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000. [Online]. Available: http://dx.doi.org/10.1109/90.879343

[30] A. Geoffrion, "Generalized benders decomposition," *Journal of Optimization Theory and Applications*, vol. 10, no. 4, 1972.

[31] C. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics, 2000.

[32] J. Kim, H.-W. Lee, and S. Chong, "Virtual cell beamforming in cooperative networks," *Technical Report*, 2014. [Online]. Available: http://netsys.kaist.ac.kr/publication/papers/Resources/R7.pdf

[33] R. Jain, *The art of computer systems performance analysis*. John Wiley & Sons Chichester, 1991, vol. 182.

[34] H. K. Khalil, *Nonlinear Systems, Third Edition*. New Jersey: Prentice Hall, 2002.

APPENDIX

A. Proof of Lemma 1

First, note that

$$|\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2 = (\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Re} - \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Im})^2 + (\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Im} + \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Re})^2.$$

Thus, we have

$$\frac{\partial|\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2}{\partial\bar{\mathbf{v}}_i^{\Re}} = 2\bar{\mathbf{h}}_{ki}^{\Re T}(\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Re} - \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Im}) + 2\bar{\mathbf{h}}_{ki}^{\Im T}(\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Im} + \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Re})$$
$$\frac{\partial|\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2}{\partial\bar{\mathbf{v}}_i^{\Im}} = -2\bar{\mathbf{h}}_{ki}^{\Im T}(\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Re} - \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Im}) + 2\bar{\mathbf{h}}_{ki}^{\Re T}(\bar{\mathbf{h}}_{ki}^{\Re}\bar{\mathbf{v}}_i^{\Im} + \bar{\mathbf{h}}_{ki}^{\Im}\bar{\mathbf{v}}_i^{\Re}).$$

Let $T_k = \sum_{i\in\mathcal{K}}|\bar{\mathbf{h}}_{ki}\bar{\mathbf{v}}_i|^2 + \sigma_k^2$. Then,

$$\frac{\partial g(\bar{\bar{\mathbf{v}}})}{\partial\bar{\mathbf{v}}_k^{\Re}} = 2\frac{w_k/\log 2}{T_k}\left(\bar{\mathbf{h}}_{kk}^{\Re T}(\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Im}) + \bar{\mathbf{h}}_{kk}^{\Im T}(\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Re})\right)$$
$$- 2\sum_{j\neq k}\frac{\gamma_j w_j/\log 2}{T_j}\left(\bar{\mathbf{h}}_{jk}^{\Re T}(\bar{\mathbf{h}}_{jk}^{\Re}\bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{jk}^{\Im}\bar{\mathbf{v}}_k^{\Im}) + \bar{\mathbf{h}}_{jk}^{\Im T}(\bar{\mathbf{h}}_{jk}^{\Re}\bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{jk}^{\Im}\bar{\mathbf{v}}_k^{\Re})\right),$$

$$\frac{\partial g(\bar{\bar{\mathbf{v}}})}{\partial\bar{\mathbf{v}}_k^{\Im}} = 2\frac{w_k/\log 2}{T_k}\left(-\bar{\mathbf{h}}_{kk}^{\Im T}(\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Im}) + \bar{\mathbf{h}}_{kk}^{\Re T}(\bar{\mathbf{h}}_{kk}^{\Re}\bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{kk}^{\Im}\bar{\mathbf{v}}_k^{\Re})\right)$$
$$- 2\sum_{j\neq k}\frac{\gamma_j w_j/\log 2}{T_j}\left(-\bar{\mathbf{h}}_{jk}^{\Im T}(\bar{\mathbf{h}}_{jk}^{\Re}\bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{jk}^{\Im}\bar{\mathbf{v}}_k^{\Im}) + \bar{\mathbf{h}}_{jk}^{\Re T}(\bar{\mathbf{h}}_{jk}^{\Re}\bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{jk}^{\Im}\bar{\mathbf{v}}_k^{\Re})\right).$$

The directional derivative is then written as

$$\begin{aligned}
\nabla_{\bar{\bar{\mathbf{v}}}} g(\bar{\bar{\mathbf{v}}})^T \bar{\bar{\mathbf{u}}} &= \sum_k \left(\frac{\partial g(\bar{\bar{\mathbf{v}}}^*)}{\partial \bar{\mathbf{v}}_k^{\Re}} \bar{\mathbf{u}}_k^{\Re} + \frac{\partial g(\bar{\bar{\mathbf{v}}}^*)}{\partial \bar{\mathbf{v}}_k^{\Im}} \bar{\mathbf{u}}_k^{\Im} \right) \\
&= 2 \sum_k \frac{w_k / \log 2}{T_k} \left\{ (\bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{v}}_k^{\Im}) \bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{u}}_k^{\Re} + (\bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{v}}_k^{\Re}) \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{u}}_k^{\Re} \right. \\
&\quad \left. - (\bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{v}}_k^{\Im}) \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{u}}_k^{\Im} + (\bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{kk}^{\Im} \bar{\mathbf{v}}_k^{\Re}) \bar{\mathbf{h}}_{kk}^{\Re} \bar{\mathbf{u}}_k^{\Im} \right\} \\
&\quad - 2 \sum_k \sum_{j \neq k} \frac{\gamma_j w_j / \log 2}{T_j} \left\{ (\bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{v}}_k^{\Im}) \bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{u}}_k^{\Re} + (\bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{v}}_k^{\Re}) \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{u}}_k^{\Re} \right. \\
&\quad \left. - (\bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{v}}_k^{\Re} - \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{v}}_k^{\Im}) \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{u}}_k^{\Im} + (\bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{v}}_k^{\Im} + \bar{\mathbf{h}}_{jk}^{\Im} \bar{\mathbf{v}}_k^{\Re}) \bar{\mathbf{h}}_{jk}^{\Re} \bar{\mathbf{u}}_k^{\Im} \right\} \\
&= 2 \sum_k \Re \left(\bar{\mathbf{v}}_k^H (\mathbf{A}_k(\bar{\mathbf{v}}) - \mathbf{B}_k(\bar{\mathbf{v}}) \bar{\mathbf{u}}_k \right).
\end{aligned}$$

■

B. Proof of Proposition 4

The key idea to prove this proposition is to express the solution in each iteration as a function of the previous solution, and to derive the conditions that the iteration is a contraction mapping.

Recall that there are two MSs k and j. Then, the balancing matrix $\mathbf{A}_k(\bar{\mathbf{v}}) - \mathbf{B}_k(\bar{\mathbf{v}})$ for MS k is given by

$$\mathbf{A}_k(\bar{\mathbf{v}}) - \mathbf{B}_k(\bar{\mathbf{v}}) = w_{k,A} \bar{\mathbf{h}}_{kk}^H \bar{\mathbf{h}}_{kk} - w_{j,B} \bar{\mathbf{h}}_{jk}^H \bar{\mathbf{h}}_{jk},$$

where

$$\begin{aligned}
w_{k,A} &= \frac{w_k / \log 2}{|\bar{\mathbf{h}}_{kk} \bar{\mathbf{v}}_k|^2 + |\bar{\mathbf{h}}_{kj} \bar{\mathbf{v}}_j|^2 + \sigma_k^2} \\
w_{j,B} &= \frac{\gamma_j(\bar{\mathbf{v}}) w_j / \log 2}{|\bar{\mathbf{h}}_{jk} \bar{\mathbf{v}}_k|^2 + |\bar{\mathbf{h}}_{jj} \bar{\mathbf{v}}_j|^2 + \sigma_j^2}.
\end{aligned}$$

The balancing matrix for MS j can be similarly written as

$$\mathbf{A}_j(\bar{\mathbf{v}}) - \mathbf{B}_j(\bar{\mathbf{v}}) = w_{j,A} \bar{\mathbf{h}}_{jj}^H \bar{\mathbf{h}}_{jj} - w_{k,B} \bar{\mathbf{h}}_{kj}^H \bar{\mathbf{h}}_{kj}.$$

The iteration t of our precoding algorithm can be written as

$$\bar{\mathbf{v}}_l(t+1) = \arg \max_{||\bar{\mathbf{u}}_l|| = 1} \bar{\mathbf{u}}_l^H (\mathbf{A}_k(\bar{\mathbf{v}}(t)) - \mathbf{B}_k(\bar{\mathbf{v}}(t))) \bar{\mathbf{u}}_l, l = k, j.$$

That is, our precoding algorithm finds an eigenvector corresponding to the maximum eigenvalue of $\mathbf{A}_l(\bar{\mathbf{v}}(t)) - \mathbf{B}_l(\bar{\mathbf{v}}(t))$ for each MS l. This dominant eigenvector can be easily calculated owing to the special structure of the balancing matrix. For brevity, we will omit time indices and dependence of $w_{k,A}, w_{j,B}, \mathbf{A}, \mathbf{B}$ on the precoding vector whenever it is clear from the context.

First, note that the balancing matrix $\mathbf{A}_k - \mathbf{B}_k$ is the sum of two rank-1 matrices, and hence, its rank is at most 2. Since it is Hermitian as well, the matrix can be diagonalized using an orthogonal matrix. To do this, consider an orthonormal basis $\mathbf{z}_1$, $\mathbf{z}_2$, ..., $\mathbf{z}_{N_T}$ (the size of the balancing matrix is $N_T \times N_T$ where N_T is the number of transmit antennas) such that

$$\mathbf{z}_1 = \frac{\bar{\mathbf{h}}_{kk}^H}{||\bar{\mathbf{h}}_{kk}^H||}$$
$$\mathbf{z}_2 = \frac{\bar{\mathbf{h}}_{jk}^H - (\mathbf{z}_1^H \bar{\mathbf{h}}_{jk}^H)\mathbf{z}_1}{||\bar{\mathbf{h}}_{jk}^H - (\mathbf{z}_1^H \bar{\mathbf{h}}_{jk}^H)\mathbf{z}_1||}.$$

Assume that the rest of the vectors have been chosen appropriately. Then, it is easy to check that the matrix $\mathbf{A}_k - \mathbf{B}_k$ can be represented as

$$\mathbf{A}_k - \mathbf{B}_k = a^2 \mathbf{z}_1 \mathbf{z}_1^H - |b|^2 \mathbf{z}_1 \mathbf{z}_1^H - c^2 \mathbf{z}_2 \mathbf{z}_2^H - bc \mathbf{z}_2 \mathbf{z}_1^H - b^H c \mathbf{z}_1 \mathbf{z}_2^H,$$

where $a = \sqrt{w_{k,A}}||\bar{\mathbf{h}}_{kk}^H||$, $b = \sqrt{w_{j,B}}\mathbf{z}_1^H \bar{\mathbf{h}}_{jk}^H$ and $c = \sqrt{w_{j,B}}\mathbf{z}_2^H \bar{\mathbf{h}}_{jk}^H$. Note that only b is a complex number, and a and c are real.

Consequently, the balancing matrix can be represented in the above orthonormal basis as

$$(\mathbf{A}_k - \mathbf{B}_k) \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \cdots & \mathbf{z}_{N_T} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \cdots & \mathbf{z}_{N_T} \end{bmatrix}}_{\triangleq \mathbf{Z}} \underbrace{\begin{bmatrix} a^2 - |b|^2 & -bc & 0 & \cdots & 0 \\ -b^H c & -c^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}}_{\triangleq \mathbf{W}}.$$

Thus, the matrices $\mathbf{A}_k - \mathbf{B}_k$ and $\mathbf{W}$ are similar and share eigenvalues so that if $(\lambda, \mathbf{q})$ is an eigenpair of $\mathbf{W}$, then $(\lambda, \mathbf{Zq})$ is an eigenpair of $\mathbf{A}_k - \mathbf{B}_k$. Therefore, finding the dominant eigenvector of $\mathbf{A}_k - \mathbf{B}_k$ boils down to finding the dominant eigenvector of $\mathbf{W}$.

The matrix $\mathbf{W}$ has only two nonzero eigenvalues, and the larger of the two is

$$\lambda = \frac{a^2 - |b|^2 - c^2 + \sqrt{(a^2 - |b|^2 - c^2)^2 + 4a^2c^2}}{2},$$

the corresponding eigenvector is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{-a^2+|b|^2-c^2-\sqrt{(a^2-|b|^2-c^2)^2+4a^2c^2}}{2} \\ b^H c \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The dominant eigenvector of $\mathbf{A}_k - \mathbf{B}_k$ is then given as

$$\mathbf{Z}\mathbf{x} = x_1\mathbf{z}_1 + x_2\mathbf{z}_2.$$

Thus, the precoding vector for MS k in iteration t is updated as

$$\bar{\mathbf{v}}_k(t+1) = \frac{1}{\sqrt{|x_1(\bar{\mathbf{v}}(t))|^2 + |x_2(\bar{\mathbf{v}}(t))|^2}}(x_1(\bar{\mathbf{v}}(t))\mathbf{z}_1 + x_2(\bar{\mathbf{v}}(t))\mathbf{z}_2), \tag{29}$$

since we require $||\bar{\mathbf{v}}_k|| = 1$. Note that x_1 and x_2 are the functions of the previous precoding vector $\bar{\mathbf{v}}(t)$, and thus, we use the notation $x_1(\bar{\mathbf{v}}(t))$ and $x_2(\bar{\mathbf{v}}(t))$. On the other hand, the vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ remain constant throughout the iterations.

The precoding vector for MS j can be similarly calculated as

$$\bar{\mathbf{v}}_j(t+1) = \frac{1}{\sqrt{|x_1'(\bar{\mathbf{v}}(t))|^2 + |x_2'(\bar{\mathbf{v}}(t))|^2}}(x_1'(\bar{\mathbf{v}}(t))\mathbf{z}_1' + x_2'(\bar{\mathbf{v}}(t))\mathbf{z}_2'), \tag{30}$$

where

$$\mathbf{z}'_1 = \frac{\bar{\mathbf{h}}_{jj}^H}{||\bar{\mathbf{h}}_{jj}^H||}$$
$$\mathbf{z}'_2 = \frac{\bar{\mathbf{h}}_{kj}^H - (\mathbf{z}_1'^H \bar{\mathbf{h}}_{kj}^H)\mathbf{z}'_1}{||\bar{\mathbf{h}}_{kj}^H - (\mathbf{z}_1'^H \bar{\mathbf{h}}_{kj}^H)\mathbf{z}'_1||}$$
$$x'_1 = \frac{-a'^2 + |b'|^2 - c'^2 - \sqrt{(a'^2 - |b'|^2 - c'^2)^2 + 4a'^2c'^2}}{2}$$
$$x'_2 = b'^H c'$$
$$a' = \sqrt{w_{j,A}}||\bar{\mathbf{h}}_{jj}^H||$$
$$b' = \sqrt{w_{k,B}}\mathbf{z}_1^H \bar{\mathbf{h}}_{kj}^H$$
$$c' = \sqrt{w_{k,B}}\mathbf{z}_2^H \bar{\mathbf{h}}_{kj}^H.$$

So far, we have found a closed form solution of iteration t in our precoding algorithm. The next step is to apply the following lemma to find the conditions for contraction mapping. However, the following lemma applies to real-valued functions of real variables, and thus we transform the above update equations into the form to which we can apply the following lemma.

Lemma 2 (Lemma 3.1 of [34]): Let $f : D \to \mathbb{R}^m$ be continuous for some domain $D \subset \mathbb{R}^n$. Suppose that the Jacobian $[\partial f/\partial x]$ exists and is continuous on D. If there exists a constant $L \geq 0$ such that

$$\left|\left|\frac{\partial f(x)}{\partial x}\right|\right| \leq L, \forall x \in D,$$

then

$$||f(x) - f(y)|| \leq L||x - y||, \forall x, y \in D.$$

For a complex vector $\mathbf{z}$, let $\mathbf{z} = \mathbf{z}^{\Re} + \mathbf{z}^{\Im} i$, i.e., $\mathbf{z}^{\Re}$ and $\mathbf{z}^{\Im}$ are the vectors of real part and imaginary part of $\mathbf{z}$. Define $\bar{\mathbf{z}} = \begin{bmatrix} \mathbf{z}^{\Re} \\ \mathbf{z}^{\Im} \end{bmatrix}$. We represent the update in (29) and (30) using this augmented notation of complex vector. First, recall that $x_1 = \frac{-a^2+|b|^2-c^2-\sqrt{(a^2-|b|^2-c^2)^2+4a^2c^2}}{2}$ and $x_2 = b^H c$. Let $b = w_{j,B}(\alpha + \beta i)$. Thus, we have $\alpha = \Re(\bar{\mathbf{h}}_{jk}\mathbf{z}_1\mathbf{z}_2^H\bar{\mathbf{h}}_{jk}^H)$ and $\beta = \Im(\bar{\mathbf{h}}_{jk}\mathbf{z}_1\mathbf{z}_2^H\bar{\mathbf{h}}_{jk}^H)$. Then, the update equation (29) can be rewritten

as

$$\bar{\bar{\mathbf{v}}}_k(t+1) = \frac{1}{\sqrt{|x_1(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1(\bar{\bar{\mathbf{v}}}(t)) \begin{bmatrix} \mathbf{z}_1^{\Re} \\ \mathbf{z}_1^{\Im} \end{bmatrix} + w_{j,B}(\bar{\bar{\mathbf{v}}}(t)) \begin{bmatrix} \alpha \mathbf{z}_2^{\Re} - \beta \mathbf{z}_2^{\Im} \\ \beta \mathbf{z}_2^{\Re} + \alpha \mathbf{z}_2^{\Im} \end{bmatrix} \right),$$

where we abuse the notation so that $x_l(\bar{\mathbf{v}}(t)) = x_l(\bar{\bar{\mathbf{v}}}(t)), l = 1, 2$. Similarly, we have

$$\bar{\bar{\mathbf{v}}}_j(t+1) = \frac{1}{\sqrt{|x_1'(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2'(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1'(\bar{\bar{\mathbf{v}}}(t)) \begin{bmatrix} \mathbf{z}_1'^{\Re} \\ \mathbf{z}_1'^{\Im} \end{bmatrix} + w_{k,B}(\bar{\bar{\mathbf{v}}}(t)) \begin{bmatrix} \alpha' \mathbf{z}_2'^{\Re} - \beta' \mathbf{z}_2'^{\Im} \\ \beta' \mathbf{z}_2'^{\Re} + \alpha' \mathbf{z}_2'^{\Im} \end{bmatrix} \right),$$

where $\alpha' = \Re(\bar{\mathbf{h}}_{kj} \mathbf{z}_1' \mathbf{z}_2'^H \bar{\mathbf{h}}_{kj}^H)$ and $\beta' = \Im(\bar{\mathbf{h}}_{kj} \mathbf{z}_1' \mathbf{z}_2'^H \bar{\mathbf{h}}_{kj}^H)$. Let $\tilde{\mathbf{z}}_2^{\Re} = \alpha \mathbf{z}_2^{\Re} - \beta \mathbf{z}_2^{\Im}$, $\tilde{\mathbf{z}}_2^{\Im} = \beta \mathbf{z}_2^{\Re} - \alpha \mathbf{z}_2^{\Im}$, $\tilde{\mathbf{z}}_2'^{\Re} = \alpha' \mathbf{z}_2'^{\Re} - \beta' \mathbf{z}_2'^{\Im}$ and $\tilde{\mathbf{z}}_2'^{\Re} = \beta' \mathbf{z}_2'^{\Re} - \alpha' \mathbf{z}_2'^{\Im}$. Combining the above two equations yields

$$\bar{\bar{\mathbf{v}}}(t+1) = \begin{bmatrix} \bar{\bar{\mathbf{v}}}_k(t+1) \\ \bar{\bar{\mathbf{v}}}_j(t+1) \end{bmatrix} = f(\bar{\bar{\mathbf{v}}}(t)) = \begin{bmatrix} f_1(\bar{\bar{\mathbf{v}}}(t)) \\ f_2(\bar{\bar{\mathbf{v}}}(t)) \\ f_3(\bar{\bar{\mathbf{v}}}(t)) \\ f_4(\bar{\bar{\mathbf{v}}}(t)) \end{bmatrix},$$

where

$$\begin{aligned}
f_1(\bar{\bar{\mathbf{v}}}(t)) &= \frac{1}{\sqrt{|x_1(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1(\bar{\bar{\mathbf{v}}}(t)) \mathbf{z}_1^{\Re} + w_{j,B}(\bar{\bar{\mathbf{v}}}(t)) \tilde{\mathbf{z}}_2^{\Re} \right) \\
f_2(\bar{\bar{\mathbf{v}}}(t)) &= \frac{1}{\sqrt{|x_1(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1(\bar{\bar{\mathbf{v}}}(t)) \mathbf{z}_1^{\Im} + w_{j,B}(\bar{\bar{\mathbf{v}}}(t)) \tilde{\mathbf{z}}_2^{\Im} \right) \\
f_3(\bar{\bar{\mathbf{v}}}(t)) &= \frac{1}{\sqrt{|x_1'(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2'(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1'(\bar{\bar{\mathbf{v}}}(t)) \mathbf{z}_1'^{\Re} + w_{k,B}(\bar{\bar{\mathbf{v}}}(t)) \tilde{\mathbf{z}}_2'^{\Re} \right) \\
f_4(\bar{\bar{\mathbf{v}}}(t)) &= \frac{1}{\sqrt{|x_1'(\bar{\bar{\mathbf{v}}}(t))|^2 + |x_2'(\bar{\bar{\mathbf{v}}}(t))|^2}} \left(x_1'(\bar{\bar{\mathbf{v}}}(t)) \mathbf{z}_1'^{\Im} + w_{k,B}(\bar{\bar{\mathbf{v}}}(t)) \tilde{\mathbf{z}}_2'^{\Im} \right).
\end{aligned}$$

Let L be an upper bound on the norm of the Jacobian of f, i.e.,

$$\left|\left| \frac{\partial f(\bar{\bar{\mathbf{v}}})}{\partial \bar{\bar{\mathbf{v}}}} \right|\right| \leq L \text{ for every feasible precoding vector } \bar{\bar{\mathbf{v}}}.$$

By Lemma 2, we have

$$||\bar{\bar{\mathbf{v}}}(t+1) - \bar{\bar{\mathbf{v}}}(t)|| = ||f(\bar{\bar{\mathbf{v}}}(t)) - f(\bar{\bar{\mathbf{v}}}(t-1))|| \leq L||\bar{\bar{\mathbf{v}}}(t) - \bar{\bar{\mathbf{v}}}(t-1)||.$$

Therefore, the update equation $\bar{\bar{\mathbf{v}}}(t+1) = f(\bar{\bar{\mathbf{v}}}(t))$ is a contraction mapping as long as $L < 1$. It now remains to find the conditions for the upper bound L to be smaller than 1.

We divide the Jacobian of f into four components as follows

$$\frac{\partial f(\bar{\bar{\mathbf{v}}})}{\partial \bar{\bar{\mathbf{v}}}} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}, \tag{31}$$

where

$$\mathbf{F}_{11} = \begin{bmatrix} \frac{\partial f_1(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial f_1(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Im}} \\ \frac{\partial f_2(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial f_2(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}, \mathbf{F}_{12} = \begin{bmatrix} \frac{\partial f_1(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial f_1(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Im}} \\ \frac{\partial f_2(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial f_2(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix}, \mathbf{F}_{21} = \begin{bmatrix} \frac{\partial f_3(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial f_3(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Im}} \\ \frac{\partial f_4(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial f_4(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}, \mathbf{F}_{22} = \begin{bmatrix} \frac{\partial f_3(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial f_3(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Im}} \\ \frac{\partial f_4(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial f_4(\bar{\mathbf{v}})}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix}.$$

Each component is $2N_T \times 2N_T$ matrix, and through some calculations, they are expressed as follows

---

$$\begin{aligned}
\mathbf{F}_{11} =& \begin{bmatrix} \mathbf{z}_1^{\Re} \\ \mathbf{z}_1^{\Im} \end{bmatrix} \left[\frac{\partial}{\partial \bar{\mathbf{v}}_k^{\Re}} \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}} \right)^T \frac{\partial}{\partial \bar{\mathbf{v}}_k^{\Im}} \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}} \right)^T \right] + \begin{bmatrix} \tilde{\mathbf{z}}_2^{\Re} \\ \tilde{\mathbf{z}}_2^{\Im} \end{bmatrix} \left[\frac{\partial}{\partial \bar{\mathbf{v}}_k^{\Re}} \left(\frac{w_{j,B}}{\sqrt{x_1^2+x_2^2}} \right)^T \frac{\partial}{\partial \bar{\mathbf{v}}_k^{\Im}} \left(\frac{w_{j,B}}{\sqrt{x_1^2+x_2^2}} \right)^T \right], \\
=& \mathbf{Y}_0 \begin{bmatrix} y_1 & y_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \left(\frac{\partial w_{k,A}}{\partial \bar{\mathbf{v}}_k^{\Re}} \right)^T & \left(\frac{\partial w_{k,A}}{\partial \bar{\mathbf{v}}_k^{\Im}} \right)^T \\ \left(\frac{\partial w_{j,B}}{\partial \bar{\mathbf{v}}_k^{\Re}} \right)^T & \left(\frac{\partial w_{j,B}}{\partial \bar{\mathbf{v}}_k^{\Im}} \right)^T \end{bmatrix} \\
=& \mathbf{Y}_0 \begin{bmatrix} y_1 & y_2 \\ 0 & 1 \end{bmatrix} \left\{ \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{kk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{kk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix} \begin{bmatrix} \frac{-1}{(S_k+I_k)^2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{jk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{jk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{-S_j \left(\frac{1}{I_j} + \frac{1}{S_j+I_j} \right)}{I_j(S_j+I_j)} \end{bmatrix} \right\}, \qquad (32)
\end{aligned}$$

$$\mathbf{F}_{12} = \mathbf{Y}_0 \begin{bmatrix} y_1 & y_2 \\ 0 & 1 \end{bmatrix} \left\{ \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{kj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{kj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix} \begin{bmatrix} \frac{-1}{(S_k+I_k)^2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{jj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{jj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{(S_j+I_j)^2} \end{bmatrix} \right\}, \qquad (33)$$

$$\mathbf{F}_{21} = \mathbf{Y}_0' \begin{bmatrix} y_1' & y_2' \\ 0 & 1 \end{bmatrix} \left\{ \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{jk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{jk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix} \begin{bmatrix} \frac{-1}{(S_j+I_j)^2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{kk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{kk} \bar{\mathbf{v}}_k|^2}{\partial \bar{\mathbf{v}}_k^{\Im}} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{(S_k+I_k)^2} \end{bmatrix} \right\}, \qquad (34)$$

$$\mathbf{F}_{22} = \mathbf{Y}_0' \begin{bmatrix} y_1' & y_2' \\ 0 & 1 \end{bmatrix} \left\{ \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{jj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{jj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix} \begin{bmatrix} \frac{-1}{(S_j+I_j)^2} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{\partial |\bar{\mathbf{h}}_{kj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Re}} & \frac{\partial |\bar{\mathbf{h}}_{kj} \bar{\mathbf{v}}_j|^2}{\partial \bar{\mathbf{v}}_j^{\Im}} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{-S_k \left(\frac{1}{I_k} + \frac{1}{S_k+I_k} \right)}{I_k(S_k+I_k)} \end{bmatrix} \right\}, \qquad (35)$$

where

$$\mathbf{Y}_0 = (x_1^2 + x_2^2)^{-\frac{3}{2}} \begin{bmatrix} x_2^2\mathbf{z}_1^{\Re} - x_1 w_{j,B}\tilde{\mathbf{z}}_2^{\Re} & x_1^2\tilde{\mathbf{z}}_2^{\Re} - x_1|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2|\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 w_{j,B}\mathbf{z}_1^{\Re} \\ x_2^2\mathbf{z}_1^{\Im} - x_1 w_{j,B}\tilde{\mathbf{z}}_2^{\Im} & x_1^2\tilde{\mathbf{z}}_2^{\Im} - x_1|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2|\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 w_{j,B}\mathbf{z}_1^{\Im} \end{bmatrix},$$

$$\mathbf{Y}_0' = (x_1'^2 + x_2'^2)^{-\frac{3}{2}} \begin{bmatrix} x_2'^2\mathbf{z}_1'^{\Re} - x_1' w_{k,B}\tilde{\mathbf{z}}_2'^{\Re} & x_1'^2\tilde{\mathbf{z}}_2'^{\Re} - x_1'|\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2|\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2 w_{k,B}\mathbf{z}_1'^{\Re} \\ x_2'^2\mathbf{z}_1'^{\Im} - x_1' w_{k,B}\tilde{\mathbf{z}}_2'^{\Im} & x_1'^2\tilde{\mathbf{z}}_2'^{\Im} - x_1'|\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2|\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2 w_{k,B}\mathbf{z}_1'^{\Im} \end{bmatrix},$$

$$y_1 = \frac{1}{2}\left[-\bar{a} + \frac{-\bar{a}^2 w_{k,A} + \bar{a}\bar{b} w_{j,B}}{\sqrt{\bar{a}^2 w_{k,A}^2 + \bar{c}^2 w_{j,B}^2 - 2\bar{a}\bar{b} w_{k,A} w_{j,B}}}\right], y_2 = \frac{1}{2}\left[\bar{b} + \frac{-\bar{c}^2 w_{j,B} + \bar{a}\bar{b} w_{k,A}}{\sqrt{\bar{a}^2 w_{k,A}^2 + \bar{c}^2 w_{j,B}^2 - 2\bar{a}\bar{b} w_{k,A} w_{j,B}}}\right],$$

where $\bar{a} = \|\bar{\mathbf{h}}_{kk}\|^2$, $\bar{b} = |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 - |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2$, and $\bar{c} = \|\bar{\mathbf{h}}_{jk}\|^2$,

$$y_1' = \frac{1}{2}\left[-\bar{a}' + \frac{-\bar{a}'^2 w_{j,A} + \bar{a}'\bar{b}' w_{k,B}}{\sqrt{\bar{a}'^2 w_{j,A}^2 + \bar{c}'^2 w_{k,B}^2 - 2\bar{a}'\bar{b}' w_{j,A} w_{k,B}}}\right], y_2' = \frac{1}{2}\left[\bar{b}' + \frac{-\bar{c}'^2 w_{k,B} + \bar{a}'\bar{b}' w_{j,A}}{\sqrt{\bar{a}'^2 w_{j,A}^2 + \bar{c}'^2 w_{k,B}^2 - 2\bar{a}'\bar{b}' w_{j,A} w_{k,B}}}\right],$$

where $\bar{a}' = \|\bar{\mathbf{h}}_{jj}\|^2$, $\bar{b}' = |\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2 - |\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2$, and $\bar{c}' = \|\bar{\mathbf{h}}_{kj}'\|^2$,

$$S_k = |\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2, I_k = |\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2 + \sigma^2, S_j = |\bar{\mathbf{h}}_{jj}\bar{\mathbf{v}}_j|^2, I_j = |\bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k|^2 + \sigma^2$$

The upper bound of Jacobian (31) can be given by

$$\left\| \frac{\partial f(\bar{\bar{\mathbf{v}}})}{\partial \bar{\bar{\mathbf{v}}}} \right\| \leq \|\mathbf{F}_{11}\| + \|\mathbf{F}_{12}\| + \|\mathbf{F}_{21}\| + \|\mathbf{F}_{22}\| \tag{36}$$

and now we find the upper bound of each component matrix. Since $\mathbf{F}_{22}$ and $\mathbf{F}_{21}$ are similar to $\mathbf{F}_{11}$ and $\mathbf{F}_{12}$, we only describe deriving the upper bound for two componen matrices. From (32) and (33), we obtain

$$\begin{aligned} \|\mathbf{F}_{11}\| &\leq \|\mathbf{Y}_0\||y_1|\frac{1}{(S_k+I_k)^2}\left\|\begin{bmatrix} \frac{\partial|\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Re}} & \frac{\partial|\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}\right\| \\ &\quad + \|\mathbf{Y}_0\|(|y_2|+1)\frac{S_j}{I_j(S_j+I_j)}\left(\frac{1}{I_j}+\frac{1}{S_j+I_j}\right)\left\|\begin{bmatrix} \frac{\partial|\bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Re}} & \frac{\partial|\bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}\right\|, \end{aligned} \tag{37}$$

$$\begin{aligned} \|\mathbf{F}_{12}\| &\leq \|\mathbf{Y}_0\||y_1|\frac{1}{(S_k+I_k)^2}\left\|\begin{bmatrix} \frac{\partial|\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2}{\partial\bar{\mathbf{v}}_k^{\Re}} & \frac{\partial|\bar{\mathbf{h}}_{kj}\bar{\mathbf{v}}_j|^2}{\partial\bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}\right\| \\ &\quad + \|\mathbf{Y}_0\|(|y_2|+1)\frac{1}{(S_j+I_j)^2}\left\|\begin{bmatrix} \frac{\partial|\bar{\mathbf{h}}_{jj}\bar{\mathbf{v}}_j|^2}{\partial\bar{\mathbf{v}}_k^{\Re}} & \frac{\partial|\bar{\mathbf{h}}_{jj}\bar{\mathbf{v}}_j|^2}{\partial\bar{\mathbf{v}}_k^{\Im}} \end{bmatrix}\right\|. \end{aligned} \tag{38}$$

In (37) and (38), the upper bound of terms by S_k and I_k (or S_j and I_j) for all $\bar{\mathbf{v}}$ is given by,

$$\frac{1}{I_k} \leq \frac{1}{\sigma^2}, \frac{1}{S_k+I_k} \leq \frac{1}{\sigma^2}, \frac{S_k}{S_k+I_k} \leq 1,$$

and $\|[\frac{\partial|\bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Re}}\frac{\partial|\bar{\mathbf{h}}_{jk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k^{\Im}}]\|$ is equivalent to $\|\frac{\partial|\bar{\mathbf{h}}_{kk}\bar{\mathbf{v}}_k|^2}{\partial\bar{\mathbf{v}}_k}\|$, which is upper-bounded by $2\|\bar{\mathbf{h}}_{kk}\|^2$. Then, the upper bound in (37) and (38) yields

$$\|\mathbf{F}_{11}\| \leq \frac{1}{\sigma^4}\|\mathbf{Y}_0\|\left(2|y_1|\|\bar{\mathbf{h}}_{kk}\|^2 + 4(|y_2|+1)\|\bar{\mathbf{h}}_{jk}\|^2\right), \tag{39}$$

$$\|\mathbf{F}_{12}\| \leq \frac{1}{\sigma^4}\|\mathbf{Y}_0\|\left(2|y_1|\|\bar{\mathbf{h}}_{kj}\|^2 + 2(|y_2|+1)\|\bar{\mathbf{h}}_{jj}\|^2\right). \tag{40}$$

Next, we derive the upper bound of $|y_1|$, $|y_2|$ and $\|\mathbf{Y}_0\|$. Since $\bar{b} = |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 - |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 \leq \bar{c} = \|\bar{\mathbf{h}}_{jk}\|^2$, the denomiator in y_1 satisfies the condition $\bar{a}^2w_{k,A}^2 + \bar{c}^2w_{j,B}^2 - 2\bar{a}\bar{b}w_{k,A}w_{j,B} \geq (\bar{a}w_{k,A} - \bar{b}w_{j,B})^2$. Then, we obtain

$$|y_1| \leq \frac{1}{2}\bar{a} + \frac{1}{2}\left|\frac{-\bar{a}^2w_{k,A} + \bar{a}\bar{b}w_{j,B}}{\bar{a}w_{k,A} - \bar{b}w_{j,B}}\right| = \bar{a} = \|\bar{\mathbf{h}}_{kk}\|^2. \tag{41}$$

In case of y_2, we derive the upper bound by dividing positive and negative. If $y_2 > 0$, similarly to y_1, we obtain

$$\begin{aligned}|y_2| = y_2 &\leq \frac{1}{2}\bar{b} - \frac{\bar{b}^2w_{j,B} - \bar{a}\bar{b}w_{k,A}}{2\sqrt{\bar{a}^2w_{k,A}^2 + \bar{c}^2w_{j,B}^2 - 2\bar{a}\bar{b}w_{k,A}w_{j,B}}} \\ &\leq \frac{1}{2}\bar{b} + \frac{1}{2}\left|\frac{\bar{b}^2w_{j,B} - \bar{a}\bar{b}w_{k,A}}{\bar{a}w_{k,A} - \bar{b}w_{j,B}}\right| = \bar{b} = \|\bar{\mathbf{h}}_{jk}\|^2.\end{aligned} \tag{42}$$

The first inequality follows from the fact that $\bar{b} \leq \bar{c}$. If $y_2 < 0$, $\bar{c}^2w_{j,B} - 2\bar{a}\bar{b}w_{k,A} > 0$ and $\bar{a}^2w_{j,B} - 2\bar{a}\bar{b}w_{k,A} > 0$. Then, in the denominator in y_2, $\sqrt{\bar{a}^2w_{k,A}^2 + \bar{c}^2w_{j,B}^2 - 2\bar{a}\bar{b}w_{k,A}w_{j,B}}$ is low-bounded by $\bar{c}w_{j,B}$ and $\bar{a}w_{k,A}$. Then, we obtain the upper bound of y_2 as follows

$$\begin{aligned}|y_2| &\leq \frac{1}{2}\bar{b} + \frac{1}{2}\left|\frac{\bar{b}^2w_{j,B}}{\bar{c}w_{j,B}}\right| + \left|\frac{\bar{a}\bar{b}w_{k,A}}{\bar{a}w_{k,A}}\right| \leq \bar{b} + \frac{1}{2}\bar{c} = \frac{3}{2}|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 - \frac{1}{2}|\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 \\ &\leq \frac{3}{2}|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 - \frac{1}{2}|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 = |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2.\end{aligned} \tag{43}$$

The last ineqaulity follows from the assumpation that $|\bar{\mathbf{h}}_{jk}\mathbf{z}_1| \leq |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|$. Since the upper bound of positive y_2 in (42) is greater than (43), we use $\|\bar{\mathbf{h}}_{jk}\|^2$ as an upper bound of y_2.

Now we derive the upper bound of $\mathbf{Y}_0$.

$$
\begin{aligned}
\|\mathbf{Y}_0\| &= \left\| (x_1^2+x_2^2)^{-\frac{3}{2}} \left\{ |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 w_{j,B} \begin{bmatrix} \mathbf{z}_1^{\Re} \\ \mathbf{z}_1^{\Im} \end{bmatrix} \begin{bmatrix} w_{j,B} & -x_1 \end{bmatrix} + x_1 \begin{bmatrix} \tilde{\mathbf{z}}_2^{\Re} \\ \tilde{\mathbf{z}}_2^{\Im} \end{bmatrix} \begin{bmatrix} -w_{j,B} & x_1 \end{bmatrix} \right\} \right\| \\
&\leq (x_1^2+x_2^2)^{-\frac{3}{2}} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 w_{j,B} \left\| \begin{bmatrix} \mathbf{z}_1^{\Re} \\ \mathbf{z}_1^{\Im} \end{bmatrix} \right\| \left\| \begin{bmatrix} w_{j,B} \\ -x_1 \end{bmatrix} \right\| + (x_1^2+x_2^2)^{-\frac{3}{2}} |x_1| \left\| \begin{bmatrix} \tilde{\mathbf{z}}_2^{\Re} \\ \tilde{\mathbf{z}}_2^{\Im} \end{bmatrix} \right\| \left\| \begin{bmatrix} w_{j,B} \\ -x_1 \end{bmatrix} \right\| \\
&= (x_1^2+x_2^2)^{-\frac{3}{2}} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 w_{j,B} \left\| \begin{bmatrix} w_{j,B} \\ -x_1 \end{bmatrix} \right\| + (x_1^2+x_2^2)^{-\frac{3}{2}} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 |x_1| \left\| \begin{bmatrix} w_{j,B} \\ -x_1 \end{bmatrix} \right\| \\
&\leq (x_1^2+x_2^2)^{-\frac{3}{2}} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 (w_{j,B}+|x_1|)^2 \qquad (44)
\end{aligned}
$$

From the assumption that $b^2 = |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 \leq c^2 = |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2$, we derive the lower bound of $x_1^2+x_2^2$ as follows

$$
\begin{aligned}
x_1^2+x_2^2 =& \frac{1}{2}(a^4+b^4+c^4-2a^2b^2+2a^2c^2+2b^2c^2) \\
&+ \frac{1}{2}(a^2-b^2+c^2)\sqrt{a^4+b^4+c^4-2a^2b^2+2a^2c^2+2b^2c^2} \\
\geq& \frac{1}{2}a^4 + \frac{1}{2}(a^2-b^2+c^2)\sqrt{a^4} \geq a^4 = w_{k,A}^2 \|\bar{\mathbf{h}}_{kk}\|^4, \qquad (45)
\end{aligned}
$$

and the upper bound of x_1 as follows

$$
\begin{aligned}
|x_1| \leq& \frac{1}{2}\left(\bar{a}w_{k,A} + \bar{b}w_{j,B} + \sqrt{\bar{a}^2 w_{k,A}^2 + \bar{c}^2 w_{j,B}^2 - 2\bar{a}\bar{b}w_{k,A}w_{j,B}} \right) \\
\leq& \frac{1}{2}\left(\bar{a}w_{k,A} + \bar{b}w_{j,B} + \sqrt{(\bar{a}w_{k,A} + \bar{c}w_{j,B})^2} \right) \leq \bar{a}w_{k,A} + \bar{c}w_{j,B}, \\
=& w_{k,A}\|\bar{\mathbf{h}}_{kk}\|^2 + w_{j,B}\|\bar{\mathbf{h}}_{jk}\|^2. \qquad (46)
\end{aligned}
$$

Plugging (45) and (46) into (44) yields

$$
\begin{aligned}
\|\mathbf{Y}_0\| &\leq \frac{1}{w_{k,A}^3 \|\bar{\mathbf{h}}_{kk}\|^6} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 \left(w_{k,A}\|\bar{\mathbf{h}}_{kk}\|^2 + w_{j,B}\|\bar{\mathbf{h}}_{jk}\|^2 + w_{j,B} \right)^2. \\
&\leq \left(\frac{2\|\bar{\mathbf{h}}_{kk}\|^2 + \sigma^2}{\|\bar{\mathbf{h}}_{kk}\|^2} \right)^3 \frac{\left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1 \right)^2}{\sigma^4} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2. \qquad (47)
\end{aligned}
$$

Here, for all $\bar{\mathbf{v}}$, $w_{k,A}$ and $w_{j,B}$ is upper-bounded by $\frac{1}{\sigma^2}$, and $w_{k,A}$ is lower-bounded by $\frac{1}{2\|\bar{\mathbf{h}}_{kk}\|^2+\sigma^2}$.

From (41), (42) and (47), we can obtain the upper bound of Jacobian components in (39) and (40) as follows

$$\begin{aligned}
\|\mathbf{F}_{11}\| &\leq \frac{1}{\sigma^4}\|\mathbf{Y}_0\| \left(2\|\bar{\mathbf{h}}_{kk}\|^4 + 4(\|\bar{\mathbf{h}}_{jk}\|^2+1)\|\bar{\mathbf{h}}_{jk}\|^2\right) \\
&\leq \frac{2}{\sigma^4}\|\mathbf{Y}_0\| \left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right)^2 \\
&= 2\left(\frac{2}{\sigma^2} + \frac{1}{\|\bar{\mathbf{h}}_{kk}\|^2}\right)^3 \frac{\left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right)^4}{\sigma^2} |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2, \\
\|\mathbf{F}_{12}\| &\leq \frac{1}{\sigma^4}\|\mathbf{Y}_0\| \left(2\|\bar{\mathbf{h}}_{kk}\|^2|\bar{\mathbf{h}}_{kj}\|^2 + 2(\|\bar{\mathbf{h}}_{jk}\|^2+1)\|\bar{\mathbf{h}}_{jj}\|^2\right) \\
&\leq \frac{2}{\sigma^4}\|\mathbf{Y}_0\| \left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right) \|\bar{\mathbf{h}}_{jj}\|^2 \\
&= 2\left(\frac{2}{\sigma^2} + \frac{1}{\|\bar{\mathbf{h}}_{kk}\|^2}\right)^3 \frac{\left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right)^3}{\sigma^2} \|\bar{\mathbf{h}}_{jj}\|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2.
\end{aligned}$$

Similarly, we can derive the upper bound of $\|\mathbf{F}_{21}\|$ and $\|\mathbf{F}_{22}\|$ as follows

$$\begin{aligned}
\|\mathbf{F}_{21}\| &\leq 2\left(\frac{2}{\sigma^2} + \frac{1}{\|\bar{\mathbf{h}}_{jj}\|^2}\right)^3 \frac{\left(\|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + 1\right)^4}{\sigma^2} |\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2 |\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2, \\
\|\mathbf{F}_{22}\| &\leq 2\left(\frac{2}{\sigma^2} + \frac{1}{\|\bar{\mathbf{h}}_{jj}\|^2}\right)^3 \frac{\left(\|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + 1\right)^3}{\sigma^2} \|\bar{\mathbf{h}}_{kk}\|^2 |\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2 |\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2.
\end{aligned}$$

Consequently, Jacobian of f in (31) is upper-bounded by (36) as follows

$$\left\| \frac{\partial f(\bar{\bar{\mathbf{v}}})}{\partial \bar{\bar{\mathbf{v}}}} \right\| \leq 2|\bar{\mathbf{h}}_{jk}\mathbf{z}_1|^2 |\bar{\mathbf{h}}_{jk}\mathbf{z}_2|^2 \frac{C_1 C_2}{\sigma^2} + 2|\bar{\mathbf{h}}_{kj}\mathbf{z}_1'|^2 |\bar{\mathbf{h}}_{kj}\mathbf{z}_2'|^2 \frac{C_1' C_2'}{\sigma^2}, \tag{48}$$

where

$$\begin{aligned}
C_1 &= \left(2/\sigma^2 + 1/\|\bar{\mathbf{h}}_{kk}\|^2\right)^3 \left(\|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + 1\right)^3, \\
C_1' &= \left(2/\sigma^2 + 1/\|\bar{\mathbf{h}}_{jj}\|^2\right)^3 \left(\|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + 1\right)^3, \\
C_2 &= \|\bar{\mathbf{h}}_{kk}\|^2 + \|\bar{\mathbf{h}}_{jk}\|^2 + \|\bar{\mathbf{h}}_{jj}\|^2 + 1, \\
C_2' &= \|\bar{\mathbf{h}}_{jj}\|^2 + \|\bar{\mathbf{h}}_{kj}\|^2 + \|\bar{\mathbf{h}}_{kk}\|^2 + 1.
\end{aligned}$$

Thus, if the upper bound in (48) is smaller than 1, the update $\bar{\bar{\mathbf{v}}}(t)$ is a contraction mapping, and thus it converges. ■

Jihwan Kim received the B.S. and M.S. degree in electrical engineering and computer science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2009 and 2011, respectively. He is currently a Ph.D candidate at KAIST. His research interests are in the areas of WLANs and cellular network optimization.

Hyang-Won Lee (S'02-M'08) received the BS, MS, and PhD degrees in all electrical engineering and computer science from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2001, 2003, and 2007, respectively. He was a postdoctoral research associate at the Massachusetts Institute of Technology during 2007-2011, and a research assistant professor at KAIST between 2011 and 2012. He is currently an assistant professor in the Department of Internet and Multimedia Engineering, Konkuk University, Seoul, Republic of Korea. His research interests are in the areas of congestion control, stochastic network control, robust optimization and reliable network design.

Song Chong is a Professor in the Department of Electrical Engineering at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, and was the Head of Communication and Computing Division of the department in 2009-2010. He is the founding Director of KAIST-LGE 5G Mobile Communications & Networking Research Center funded by LG Electronics. Prior to joining KAIST in March 2000, he was with the Performance Analysis Department, AT&T Bell Laboratories, New Jersey, as a Member of Technical Staff. His current research interests include wireless networks, mobile networks and systems, network data analytics, distributed algorithms, and cross-layer control and optimization. He has published over 120 papers in top journals and conferences. He is currently an Associate Editor of IEEE/ACM Transactions on Networking and IEEE Transactions on Mobile Computing, and has served on the Technical Program Committee of a number of leading international conferences including IEEE INFOCOM, ACM MobiCom, ACM CoNEXT, ACM MobiHoc and IEEE ICNP. He serves on the Steering Committee of WiOpt and was the General Chair of WiOpt '09. He received the IEEE William R. Bennett Prize in 2013, given to the best original paper published in IEEE/ACM Transactions on Networking in 2011-2013, and the IEEE SECON Best Paper Award in 2013. He received the B.S. and M.S. degrees from Seoul National University and the Ph.D. degree from the University of Texas at Austin, all in electrical engineering.