Wireless Network Virtualization in Uplink Coordinated Multi-Cell MIMO Systems

Ahmed F. Almehdhar*, Min Dong‡, Ben Liang*, Gary Boudreau§, Yahia Ahmed§
*Department of Electrical and Computer Engineering, University of Toronto, Canada
‡Department of Electrical, Computer and Software Engineering, Ontario Tech University, Canada
§Ericsson Canada Inc., Canada

***Abstract*—We consider wireless network virtualization (WNV) in the uplink of a coordinated multi-cell system, where multiple service providers (SPs) operate in virtually isolated networks managed by an infrastructure provider (InP). The InP provides service isolation among the SPs by exploiting the spatial structure in MIMO communication. We jointly optimize the uplink receive beamforming at the base stations (BSs) and the transmit power of the SPs' subscribing users, by alternating between two subproblems that both admit efficient closed-form solutions. We then propose a distributed implementation that solves each subproblem among the BSs without the need for a central controller. We show that the distributed approach requires significantly less communication overhead compared with the centralized one, especially when the system is not overloaded. Our simulation results under typical wireless networking environments demonstrate that the proposed solution enables effective network virtualization, to support the independent operation of multiple SPs over multiple cells, without losing communication efficiency compared with non-virtualized network operation. Furthermore, it substantially outperforms traditional WNV based on strict resource separation, especially for systems with a large number of antennas or a large number of SPs.**

I. Introduction

Wireless communication service providers (SPs) face substantial hurdles to market entry owing to high initial capital expenses and deployment costs. Traditionally, a new SP needs to build new and full infrastructure that consists of a large number of base stations (BSs) to provide wide coverage, capacity, and performance that are similar to those provided by existing SPs. Building such infrastructure requires large capital. To address this issue, wireless network virtualization (WNV) has been put forward as a framework for different SPs to share the network's physical resources.

A WNV system consists of the SPs and an infrastructure provider (InP) that owns and manages the network's physical resources and splits them into virtual slices. These virtual slices are leased to the SPs, which in turn utilize them to provide services to their subscribing users. An SP demands services from the InP without needing knowledge of the existence of the other SPs. Although multiple SPs share the same infrastructure, none of them is expected to consider inter-SP interference in their design for the demands. Thus, it is the job of the InP to provide *service isolation*, i.e., to satisfy the demands of each SP without affecting the other SPs.

This work was supported in part by Ericsson, the Natural Sciences and Engineering Research Council of Canada, and Mitacs.

Although virtualization has been well studied for wired networks [1]–[5], WNV is more complicated, with the need to share both the hardware and the radio spectrum, and with new challenges arising in guaranteeing service isolation under wireless interference [6]. To achieve service isolation among the SPs in a wireless network, most existing works propose strict separation of the physical resources, an approach rooted in the traditional solution for wired network virtualization. This strict separation could be in the form of dividing the time, frequency spectrum, resource blocks, or antennas among different SPs [7]–[12]. However, strict separation lacks flexibility and can lead to inefficient resource utilization and severe loss of system throughput. Other works proposed non-strict resource separation between the SPs, so that they can cause interference to one another. This interference was ignored in [13] and [14], while non-orthogonal multiple access (NOMA) was used in [15]–[17] to handle the interference between users. None of the techniques in [13]–[17] allow the SPs to tailor their service demands for individual users, so they do not achieve full virtualization.

It was first proposed in [18] to provide service isolation among the SPs using multiple-input multiple-output (MIMO) signal processing techniques while they share all physical resources of a base station. Beamforming was used also in [19] and [20] to provide service isolation among the users of different SPs, by minimizing the expected deviation between the InP's supply and the SPs' demands, in offline and online setups, respectively. All these papers considered the virtualization of the wireless *downlink*. However, the problem of *uplink* WNV is equally important.

The beamforming solutions developed in [18]–[20] cannot be applied to the uplink. Furthermore, to achieve service isolation in the uplink we need to additionally manage the transmit powers of the users, to effectively reduce their interference with each other. The authors of [21] studied uplink WNV with beamforming in a single cell, and they further showed that fully-cooperative multi-cell WNV is a direct extension of their single-cell solution. However, a fully-cooperative multi-cell system requires that all mobile users communicate with all BSs, which can incur prohibitive communication and control overhead [22], [23].

Therefore, in this work, we consider a more practical multi-cell scheme where the BSs use *coordinated* multi-cell communication [22]–[25]. In such a scheme, inter-cell interference can be a determining factor on the performance of the communication system. Hence, we propose new designs for the coordinated multi-cell WVN framework where inter-cell interference, both among users of the same SP and among users of different SPs, is implicitly accounted for and suppressed. We jointly design the uplink receive beamforming vectors of each BS and the transmit power of user devices, for the InP to supply the signals demanded by the SPs while suppressing the inter-SP interference and inter-cell interference. The contributions of this paper are summarized below:

- We formulate the above uplink coordinated multi-cell MIMO WNV as a joint beamforming and power control optimization problem, to minimize the system-wide deviation between the SPs' demanded received signals and the actual received signals supplied by the InP. Our formulation allows all SPs to simultaneously enjoy the full physical resources available at the InP while providing them with the required service isolation. (Section II)
- Observing that the formulated joint optimization problem is biconvex, we solve it by first decomposing it into two subproblems, one for the receive beamforming and the other for the transmit power of users. We derive closed-form solutions to both subproblems. Then, a solution to the joint optimization problem is obtained via alternating optimization, which guarantees convergence to a partial optimum. (Section III)
- We further propose a distributed approach to implement the proposed solution without any central controller. This is achieved by decomposing the original problem into cell-separable subproblems, in two different forms, for beamforming and for power control. The proposed distributed implementation provides a solution equivalent to the centralized approach, yet with much lower communication overhead. (Section IV)
- Our simulation results indicate that the proposed virtualization solution provides per-user data rate similar to that of a multi-cell system without virtualization, while providing strong service isolation among the SPs. It also substantially outperforms the traditional virtualization scheme of strictly separating the physical resources among the SPs. (Section V)

II. System Model and Problem Formulation

A. Uplink Coordinated Multi-Cell WNV

We study multi-cell uplink communication in WNV with M SPs. The BSs are governed by an InP that performs virtualization of the system, i.e., slicing the system into M virtual networks, each for an SP. The main objective of virtualization is service isolation as defined in Section I.

Without loss of generality, we focus on any one multiple-access channel that is shared among all SPs. Suppose that SP m serves K_m single-antenna users in this shared channel. Let $K_{\text{tot}} = \sum_{m=1}^{M} K_m$ be the total number of users. We consider C cells where this channel is in use. Each cell has one BS, and the InP has control over all BSs. We further assume that BS c has N_c antennas. Let K_c be the number of users in cell c, and $K_{c,m}$ of them are subscribing to SP m. Then $K_{\text{tot}} = \sum_{c=1}^{C} K_c = \sum_{c=1}^{C}\sum_{m=1}^{M} K_{c,m}$.

Each SP assembles certain demands to be fulfilled by the InP, so that its subscribing users achieve some desired performance, e.g., maximum sum-rate or fairness. The SPs design their demands in ignorance of each other, and thus, it is the responsibility of the InP to supply the requested demands of all SPs while managing the wireless interference among them, i.e., providing service isolation. More precisely, each SP m requests that the InP uses a set of beamforming vectors $\mathbf{w}_{c,m,i} \in \mathbb{C}^{N_c \times 1}$, $\forall i \in \{1, \cdots, K_{c,m}\}$ to decode the messages of its users in cell c, and that its users transmit their signals with powers $\mathbf{p}_{c,m} = \left[p_{c,m,1}, p_{c,m,2}, \cdots, p_{c,m,K_{c,m}}\right]^T \in \mathbb{R}^{K_{c,m}}$. We further assume that the SPs design their demands in a distributed fashion, i.e., each SP designs its demands in cell c without considering the interference coming from its own users in other cells. The inter-cell interference is left for the InP to handle.

Under the control of the InP, the BSs cooperate to better mitigate the interference between cells. We assume that the cooperation between BSs is limited to sharing only their channel state information (CSI), beamforming vectors, and their users' transmit powers. Thus, unlike fully-cooperative multi-cell systems that require extremely high levels of control and communication overhead [22], [23], here the BSs do not cooperate in decoding each other's users' signals. Moreover, the BSs treat out-of-cell signals as interference that cannot be exploited and needs to be mitigated. This corresponds to the *coordinated* multi-cell communication scheme in standard wireless networks [22]–[25].

In the following, we define the desired and actual received signals of all users, comprising the demands of the SPs and the supply of the InP.

1) SPs' Demands: Let $\mathbf{x}_{c,m} = [x_{c,m,1},\, x_{c,m,2},\, \cdots,\, x_{c,m,K_m}]^T \in \mathbb{C}^{K_{c,m}}$ be the transmitted symbol vector of the users of SP m in cell c. Without loss of generality, we set $\mathbb{E}\Big\{\mathbf{x}_{c,m}\Big\} = \mathbf{0}$ and $\mathbb{E}\Big\{\mathbf{x}_{c,m}\mathbf{x}_{c,m}^H\Big\} = \mathbf{I}_{K_{c,m}}$. From SP m's perspective, it desires that the received signal vector at the BS of cell c is

$$\begin{aligned}\mathbf{y}_{c,m}^{\text{desired}} &= \sum_{j=1}^{K_m} \mathbf{h}_{cc,m,j}\sqrt{p_{c,m,j}}\,x_{c,m,j} + \mathbf{n}_c \\ &= \mathbf{H}_{cc,m}\text{diag}(\bar{\mathbf{q}}_{c,m})\mathbf{x}_{c,m} + \mathbf{n}_c, \qquad (1)\end{aligned}$$

where $\bar{\mathbf{q}}_{c,m} = \left[\sqrt{p_{c,m,1}}, \cdots, \sqrt{p_{c,m,K_m}}\right]^T \in \mathbb{R}^{K_{c,m}}$ is the signal amplitude vector set by SP m for its users in cell c, $\mathbf{H}_{cc,m} = \left[\mathbf{h}_{cc,m,1}, \mathbf{h}_{cc,m,2}, \cdots, \mathbf{h}_{cc,m,K_{c,m}}\right] \in \mathbb{C}^{N_c \times K_{c,m}}$ is the channel matrix from the users of SP m that are located in cell c to the different antenna elements of the BS in cell c, and $\mathbf{n}_c \sim \mathcal{CN}(0, \sigma_n^2\mathbf{I}_{N_c})$ is the additive noise at the antennas of BS c. Note that the received signals at BS c in (1) only accounts for the signals of those users in cell c, meaning that the SP does not need to know what occurs outside of cell

c. Then, by the design of SP m, the decoded received signal vector of all users in cell c subscribed to SP m is given by

$$\begin{aligned}\hat{\mathbf{x}}_{c,m}^{\text{desired}} &= \mathbf{W}_{c,m}\mathbf{y}_{c,m}^{\text{desired}} \\ &= \mathbf{W}_{c,m}\mathbf{H}_{cc,m}\text{diag}(\bar{\mathbf{q}}_{c,m})\mathbf{x}_{c,m} + \mathbf{W}_{c,m}\mathbf{n}_c, \end{aligned} \tag{2}$$

where $\mathbf{W}_{c,m} = \left[\mathbf{w}_{c,m,1}, \mathbf{w}_{c,m,2}, \cdots, \mathbf{w}_{c,m,K_{c,m}}\right]^T \in \mathbb{C}^{K_{c,m}\times N_c}$ is the desired beamforming matrix designed by SP m to be used at BS c in order to decode the messages of its users in that cell.

Considering all M SPs, the desired decoded received signal vector from all SPs in cell c is given by

$$\hat{\mathbf{x}}_{c}^{\text{desired}} = \mathbf{D}_c\text{diag}(\bar{\mathbf{q}}_c)\mathbf{x}_c + \mathbf{W}_c\mathbf{n}_c, \tag{3}$$

where $\mathbf{x}_c = [\mathbf{x}_{c,1}^T, \cdots, \mathbf{x}_{c,M}^T]^T \in \mathbb{C}^{K_c}$ is the transmitted symbol vector of users in cell c, $\bar{\mathbf{q}}_c = [\bar{\mathbf{q}}_{c,1}^T, \cdots, \bar{\mathbf{q}}_{c,M}^T]^T \in \mathbb{R}^{K_c}$ denotes their transmitted signal amplitudes, $\mathbf{D}_c$ is a block diagonal matrix representing the virtualization demands in cell c and is given by $\mathbf{D}_c = \text{blkdiag}\left\{\mathbf{W}_{c,1}\mathbf{H}_{cc,1}, \cdots, \mathbf{W}_{c,M}\mathbf{H}_{cc,M}\right\}$, and $\mathbf{W}_c = [\mathbf{W}_{c,1}^T, \cdots, \mathbf{W}_{c,M}^T]^T$ is a stacking of the beamforming matrices designed by all SPs in cell c. By stacking (3) over all C cells, we can finally write the desired decoded received signal from all users in the system as

$$\hat{\mathbf{x}}^{\text{desired}} = \mathbf{D}\text{diag}(\bar{\mathbf{q}})\mathbf{x} + \mathbf{W}\mathbf{n}, \tag{4}$$

where $\mathbf{D} = \text{blkdiag}\left\{\mathbf{D}_1, \cdots, \mathbf{D}_C\right\}$, $\bar{\mathbf{q}} = [\bar{\mathbf{q}}_1^T, \bar{\mathbf{q}}_2^T, \cdots, \bar{\mathbf{q}}_C^T]^T$, $\mathbf{x} = [\mathbf{x}_1^T, \cdots, \mathbf{x}_C^T]^T$, $\mathbf{W} = \text{blkdiag}\left\{\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_C\right\}$, and $\mathbf{n} = [\mathbf{n}_1^T, \cdots, \mathbf{n}_C^T]^T$.

The desired signal in (4), designed by the SPs, does not account for the interference between SPs, i.e., intra-cell inter-SP interference, or the interference across cells. These types of interference occur due to the fact that all SPs use the same time-frequency resources, while they are being oblivious to one another. Therefore, the desired signal is unrealistic and cannot be directly achieved. In WNV, it is the job of the InP to utilize its physical resources to approximate the desired signal as best as it can.

2) InP's Supply: The InP makes the actual beamforming design and chooses the user transmit powers to achieve its goal of satisfying those demands.

In coordinated multi-cell communication, each BS receives signals from all users. Thus, the actual received signal vector at the BS of cell c is given by

$$\mathbf{y}_c^{\text{actual}} = \mathbf{H}_c\text{diag}(\mathbf{q})\mathbf{x} + \mathbf{n}_c, \tag{5}$$

where $\mathbf{H}_c = [\mathbf{H}_{1c}, \mathbf{H}_{2c}, \cdots, \mathbf{H}_{Cc}] \in \mathbb{C}^{N_c\times K}$, and $\mathbf{H}_{lc} = [\mathbf{H}_{lc,1}, \mathbf{H}_{lc,2}, \cdots, \mathbf{H}_{lc,M}] \in \mathbb{C}^{N_c\times K_l}$ is the channel matrix between the users in cell l and the BS in cell c, and $\mathbf{q}$ is the signal amplitude vector set by the InP for all users. BS c applies a set of beamforming vectors to decode the messages of its users in cell c only. Thus, the actual decoded received signal vector at BS c is given by

$$\hat{\mathbf{x}}_c^{\text{actual}} = \mathbf{V}_c\mathbf{y}_c^{\text{actual}} = \mathbf{V}_c\mathbf{H}_c\text{diag}(\mathbf{q})\mathbf{x} + \mathbf{V}_c\mathbf{n}_c, \tag{6}$$

where $\mathbf{V}_c \in \mathbb{C}^{K_c\times N_c}$ is the actual beamforming matrix designed by the InP for the BS of cell c to decode the messages of the users in the cell. If we let $\hat{\mathbf{x}}^{\text{actual}} = [\hat{\mathbf{x}}_1^{\text{actual}^T}, \cdots, \hat{\mathbf{x}}_C^{\text{actual}^T}]^T$ be the actual decoded received signal vector of all cells, we can write it as

$$\begin{aligned}\hat{\mathbf{x}}^{\text{actual}} &= \text{blkdiag}\left\{\mathbf{V}_1, \mathbf{V}_2, \cdots, \mathbf{V}_C\right\}(\mathbf{H}\text{diag}(\mathbf{q})\mathbf{x} + \mathbf{n}) &(7)\\ &= \mathbf{V}\left(\mathbf{H}\text{diag}(\mathbf{q})\mathbf{x} + \mathbf{n}\right), &(8)\end{aligned}$$

where $\mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, \cdots, \mathbf{H}_C^T]^T \in \mathbb{C}^{\sum_c N_c\times K}$ is the overall channel matrix from all users to all BSs, and $\mathbf{V} = \text{blkdiag}\left\{\mathbf{V}_1, \mathbf{V}_2, \cdots, \mathbf{V}_C\right\}$ is a block-diagonal beamforming matrix. It is worth to point out that the form of the supply above does not depend on M, the number of SPs the InP serves.

Although the InP provides services directly to the users, these services should be based on the demands of the SPs of those users. That is, the InP designs the beamforming matrices $\{\mathbf{V}_c\}_{c=1}^C$ and the user transmit power vector $\mathbf{q}$ based on what the different SPs demand for their users. This is expressed in the problem formulation below.

B. WNV Problem Formulation

As an inherent characteristic of WNV, the InP aims to supply the demands requested by the different SPs, which may be based on some prior agreements between the InP and the SPs. The demands, as described in (4), are fully characterized by the receive beamforming matrices and user transmit powers, i.e., $\mathbf{W}_m$ and $\bar{\mathbf{q}}_m$ $\forall m$. Noting that the form in (4) represents perfect isolation between SPs since no interference is present, it is a logical choice for the InP to aim at making $\hat{\mathbf{x}}^{\text{actual}}$ as close to $\hat{\mathbf{x}}^{\text{desired}}$ as possible. We consider the expected l_2-norm deviation as a metric to measure how far the two vectors are from each other, which is given by

$$f(\mathbf{V},\mathbf{q}) = \mathbb{E}\left\{\left\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\right\}, \tag{9}$$

where the expectation is taken over $\mathbf{x}$ and $\mathbf{n}$.

Thus, the InP aims to solve the following optimization problem:

$$\begin{aligned}\min_{\mathbf{V},\mathbf{q}} \quad & f(\mathbf{V}, \mathbf{q}) &(10a)\\ \text{s.t.} \quad & \mathbf{V} \text{ is block diagonal}, &(10b)\\ & 0 \preccurlyeq \mathbf{q} \preccurlyeq \mathbf{q}_{\max}. &(10c)\end{aligned}$$

As seen above, the InP jointly optimizes the beamforming matrix and the transmit powers to minimize the expected deviation. The constraint in (10b) imposes the block diagonal structure on $\mathbf{V}$, and the power constraint in (10c) gives the InP the permission to use any power value below the users' maximum available power. We remark that another practically meaningful variation of this constraint is to prevent the InP from assigning powers that are greater than the requested powers, which can be reflected by replacing constraint (10c) with $0 \preccurlyeq \mathbf{q} \preccurlyeq \bar{\mathbf{q}}$. The solutions in Sections III and IV can be easily modified to facilitate this case.

Next, we first consider a simpler centralized solution, i.e., all BSs are connected to a central controller (CC) and all decisions and design parameters are computed at the CC before they are sent back to the BSs for implementation. Later in Section IV, we propose distributed implementation that removes the need for the CC and achieves the same solution with lower communication overhead.

III. Proposed Solution for Uplink Coordinated Multi-Cell WNV

The first step into tackling problem (10) is to simplify the deviation expression in the objective. We have

$$\begin{aligned} f(\mathbf{V},\mathbf{q}) &= \mathbb{E}\Big\{\big\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\big\|_2^2\Big\} \\ &= \mathbb{E}\Big\{\Big\|\left(\mathbf{V}\mathbf{H}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\right)\mathbf{x} + \left(\mathbf{V}-\mathbf{W}\right)\mathbf{n}\Big\|_2^2\Big\} \\ &= \|\mathbf{V}\mathbf{H}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2\|\mathbf{V}-\mathbf{W}\|_F^2, \quad (11) \end{aligned}$$

where the last line is obtained using the properties $\|\mathbf{a}\|_2^2 = \mathbf{a}^H\mathbf{a} = \text{tr}\Big(\mathbf{a}^H\mathbf{a}\Big)$, $\|\mathbf{A}\|_F^2 = \text{tr}\Big(\mathbf{A}\mathbf{A}^H\Big)$, and $\mathbb{E}\{\text{tr}(\cdot)\} = \text{tr}(\mathbb{E}\{\cdot\})$. Due to the multiplication operation in the first term in (11), problem (10) is non-convex in the decision variables $\mathbf{V}$ and $\mathbf{q}$, so it cannot be solved via regular convex optimization techniques. However, this problem is biconvex [26], i.e., it is convex in $\mathbf{V}$ and $\mathbf{q}$ but not jointly in both.

Therefore, our approach to solving problem (10) is by decomposing it into a beamforming subproblem to optimize $\mathbf{V}$ and a power control subproblem to optimize $\mathbf{q}$. In the following subsections, we show that there is a closed-form solution to each of them. Then, we use an alternating optimization approach to find a partial optimum of the original joint optimization problem.

A. Beamforming Subproblem

Here, we treat $\mathbf{q}$ in problem (10) as a constant and find the optimal beamforming matrix $\mathbf{V}^\star$ by solving the following beamforming subproblem:

$$\begin{aligned} \min_{\mathbf{V}} \ & \|\mathbf{V}\mathbf{H}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2\|\mathbf{V}-\mathbf{W}\|_F^2 \quad (12a) \\ \text{s.t.} \ & \mathbf{V} \text{ is block diagonal.} \quad (12b) \end{aligned}$$

The block-diagonal structure of $\mathbf{V}$ makes it difficult to deal with (12) as it is. Hence, we rewrite this problem by incorporating the constraint (12b) into the objective

$$\begin{aligned} f(\mathbf{V},\mathbf{q}) = \sum_{c=1}^{C}\Bigg(& \|\mathbf{V}_c\mathbf{H}_{cc}\text{diag}(\mathbf{q}_c) - \mathbf{D}_c\text{diag}(\bar{\mathbf{q}}_c)\|_F^2 \\ & + \sum_{l\neq c}^{C}\|\mathbf{V}_c\mathbf{H}_{lc}\text{diag}(\mathbf{q}_l)\|_F^2 + \sigma_n^2\|\mathbf{V}_c-\mathbf{W}_c\|_F^2\Bigg). \quad (13) \end{aligned}$$

The first term in (13) represents the noise-free deviation between the supply and demand in cell c, which, combined with the third term, gives the exact local deviation in cell c assuming the cell is isolated. This is identical to the single-cell case [21], except that here we are summing over all the cells. The second term quantifies the interference received by the BS in cell c from out-of-cell users. This quantity is sometimes called the signal leakage, since it represents unwanted signals that are leaked into cell c from other cells.

We can interpret the signal leakage in (13) as a supply to no demand, and hence (13) can be further simplified as

$$\begin{aligned} f(\mathbf{V},\mathbf{q}) = \sum_{c=1}^{C} & \big\|\mathbf{V}_c\mathbf{H}_c\text{diag}(\mathbf{q}) - \overline{\mathbf{D}}_c\text{diag}(\bar{\mathbf{q}})\big\|_F^2 \\ & + \sigma_n^2\|\mathbf{V}_c-\mathbf{W}_c\|_F^2, \quad (14) \end{aligned}$$

where $\overline{\mathbf{D}}_c = [\mathbf{0}_{K_c\times\sum_{j=1}^{c-1}K_j}, \mathbf{D}_c, \mathbf{0}_{K_c\times\sum_{j=c+1}^{C}K_j}]$, and $\mathbf{0}_{A\times B}$ is a zero-valued matrix of size $A\times B$. The form in (14) indicates that the expected deviation in the system is the sum of the individual deviations of its constituent cells. Thus, we have converted problem (12) to an unconstrained convex problem with a differentiable objective function in (14). A unique minimum can be found by finding the values of $\{\mathbf{V}_c\}_{c=1}^{C}$ that make the gradient of the objective equal to zero.

Let $\mathbf{Q} = \text{diag}(\mathbf{q})$, and $\bar{\mathbf{Q}} = \text{diag}(\bar{\mathbf{q}})$. We start by writing the Frobenius norms in $f(\mathbf{V},\mathbf{q})$ as traces to make the objective easier to differentiate:

$$\begin{aligned} f(\mathbf{V},\mathbf{q}) = \sum_{c=1}^{C}\Big\{ & \text{tr}\Big(\mathbf{V}_c\mathbf{H}_c\mathbf{Q}\mathbf{Q}^H\mathbf{H}_c^H\mathbf{V}_c^H\Big) \\ & - \text{tr}\Big(\overline{\mathbf{D}}_c\bar{\mathbf{Q}}\mathbf{Q}^H\mathbf{H}_c^H\mathbf{V}_c^H\Big) - \text{tr}\Big(\mathbf{V}_c\mathbf{H}_c\mathbf{Q}\bar{\mathbf{Q}}^H\overline{\mathbf{D}}_c^H\Big) \\ & + \text{tr}\Big(\overline{\mathbf{D}}_c\bar{\mathbf{Q}}\bar{\mathbf{Q}}^H\overline{\mathbf{D}}_c^H\Big) + \sigma_n^2\Big(\text{tr}\Big(\mathbf{V}_c\mathbf{V}_c^H\Big) \\ & - \text{tr}\Big(\mathbf{W}_c\mathbf{V}_c^H\Big) - \text{tr}\Big(\mathbf{V}_c\mathbf{W}_c^H\Big) + \text{tr}\Big(\mathbf{W}_c\mathbf{W}_c^H\Big)\Big)\Big\}. \end{aligned}$$

Differentiating $f(\mathbf{V},\mathbf{q})$ with respect to $\mathbf{V}_c^\dagger$, the complex conjugate of $\mathbf{V}_c$, gives

$$\begin{aligned} \frac{\partial f(\mathbf{V},\mathbf{q})}{\partial \mathbf{V}_c^\dagger} = \mathbf{V}_c\mathbf{H}_c\mathbf{Q}\mathbf{Q}^H\mathbf{H}_c^H - \overline{\mathbf{D}}_c\bar{\mathbf{Q}}\mathbf{Q}^H\mathbf{H}_c^H & \\ + \sigma_n^2\mathbf{V}_c - \sigma_n^2\mathbf{W}_c, & \end{aligned}$$

where we have used trace differentiation rules [27]. Since the matrix $\left(\mathbf{H}_c\mathbf{Q}\mathbf{Q}^H\mathbf{H}_c^H + \sigma_n^2\mathbf{I}\right)$ is positive definite, we solve $\frac{\partial f(\mathbf{V},\mathbf{q})}{\partial \mathbf{V}_c^\dagger} = \mathbf{0}$ and obtain a closed-form expression for $\mathbf{V}_c^\star$:

$$\begin{aligned} \mathbf{V}_c^\star &= \left(\overline{\mathbf{D}}_c\bar{\mathbf{Q}}\mathbf{Q}^H\mathbf{H}_c^H + \sigma_n^2\mathbf{W}_c\right)\left(\mathbf{H}_c\mathbf{Q}\mathbf{Q}^H\mathbf{H}_c^H + \sigma_n^2\mathbf{I}\right)^{-1} \\ &= \left(\mathbf{D}_c\bar{\mathbf{Q}}_c\mathbf{Q}_c^H\mathbf{H}_{cc}^H + \sigma_n^2\mathbf{W}_c\right)\left(\mathbf{H}_c\mathbf{Q}\mathbf{Q}^H\mathbf{H}_c^H + \sigma_n^2\mathbf{I}\right)^{-1}. \end{aligned} \quad (15)$$

B. Power Control Subproblem

In this part, we treat $\mathbf{V}$ in problem (10) as a constant and find the optimal user transmit powers $\mathbf{q}^\star$ by solving the following power control subproblem:

$$\begin{aligned} \min_{\mathbf{q}} \ & \|\mathbf{V}\mathbf{H}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2\|\mathbf{V}-\mathbf{W}\|_F^2 \quad (16a) \\ \text{s.t.} \ & 0 \preccurlyeq \mathbf{q} \preccurlyeq \mathbf{q}_{\max}. \quad (16b) \end{aligned}$$

This is a constrained convex optimization problem with a differentiable objective. Strong duality holds since Slater's condition is trivially satisfied.

We solve this problem by studying the KKT conditions [28]. To do so, we start by letting $\mathbf{A} = \mathbf{V}\mathbf{H}$. Now, the objective can be written as $f(\mathbf{V}, \mathbf{q}) = \|\mathbf{A}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2 \|\mathbf{V} - \mathbf{W}\|_F^2$. The Lagrangian of this problem is

$$\begin{aligned}\mathcal{L}(\mathbf{q}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= \|\mathbf{A}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2 \|\mathbf{V} - \mathbf{W}\|_F^2 \\ &\quad + \boldsymbol{\lambda}^T(\mathbf{q} - \mathbf{q}_{\max}) - \boldsymbol{\mu}^T\mathbf{q} \\ &= \sum_{j=1}^{K} q_j^2 \mathbf{a}_j^H \mathbf{a}_j - \sum_{j=1}^{K} \bar{q}_j q_j \mathbf{d}_j^H \mathbf{a}_j - \sum_{j=1}^{K} \bar{q}_j q_j \mathbf{a}_j^H \mathbf{d}_j \\ &\quad + C + \boldsymbol{\lambda}^T(\mathbf{q} - \mathbf{q}_{\max}) - \boldsymbol{\mu}^T\mathbf{q}, \end{aligned} \tag{17}$$

where $\mathbf{a}_j$ and $\mathbf{d}_j$ are respectively the j^{th} columns of $\mathbf{A}$ and $\mathbf{D}$, $C = \sum_{j=1}^{K} \bar{q}_j^2 \mathbf{d}_j^H \mathbf{d}_j + \sigma_n^2 \|\mathbf{V} - \mathbf{W}\|_F^2$ is the sum of the terms that are independent of $\mathbf{q}$, and $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are the Lagrange multipliers associated with the constraints in (16b). Differentiating $\mathcal{L}(\mathbf{q}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ with respect to every q_j and setting it to zero gives the following stationarity condition:

$$2q_j^\star \|\mathbf{a}_j\|_2^2 = 2\bar{q}_j \Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) - (\lambda_j^\star - \mu_j^\star) \quad \forall j. \tag{18}$$

The other KKT conditions for $(\mathbf{q}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$ to form a globally optimal solution of problem (16) are

$$\mathbf{q}^\star \preccurlyeq \mathbf{q}_{\max}, \tag{19}$$

$$\mathbf{q}^\star \succcurlyeq 0, \tag{20}$$

$$\boldsymbol{\lambda}^\star \succcurlyeq 0, \tag{21}$$

$$\boldsymbol{\mu}^\star \succcurlyeq 0, \tag{22}$$

$$\lambda_j^\star (q_j^\star - q_{j_{\max}}) = 0, \quad \forall j, \tag{23}$$

$$\mu_j^\star q_j^\star = 0, \quad \forall j, \tag{24}$$

where equations (19)-(20) and (21)-(22) are, respectively, the primal and dual feasibility conditions, and (23)-(24) are the complementary slackness conditions.

To solve for $\mathbf{q}$, we look into the different possibilities of $q_j^\star$, $\lambda_j^\star$, and $\mu_j^\star$ that satisfy the KKT conditions.

1) $\lambda_j^\star > 0$: From (23), $q_j^\star = q_{j\max}$, and from (24), $\mu_j^\star = 0$. By (18) this implies that $\bar{q}_j \Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) > q_{j\max} \|\mathbf{a}_j\|_2^2$.
2) $\lambda_j^\star = 0$: We have two cases for $\mu_j^\star$:
   i) $\mu_j^\star = 0$: By (18), $q_j^\star$ is given by
   $$q_j^\star = \frac{\Re\big(\mathbf{d}_j^H \mathbf{a}_j\big)}{\|\mathbf{a}_j\|_2^2} \bar{q}_j$$
   and from (19) and (20) we have $0 \le \bar{q}_j \Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) \le q_{j\max} \|\mathbf{a}_j\|_2^2$.
   ii) $\mu_j^\star > 0$: From (24), $q_j^\star = 0$. Then (18) implies that $\Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) < 0$.

By combining these different cases, we have the following closed-form expression for the optimal $q_j^\star$:

$$q_j^\star = \begin{cases} \min\left\{ \frac{\Re(\mathbf{d}_j^H \mathbf{a}_j)}{\|\mathbf{a}_j\|_2^2} \bar{q}_j, q_{j\max} \right\}, & \Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) \ge 0, \\ 0, & \Re\big(\mathbf{d}_j^H \mathbf{a}_j\big) < 0. \end{cases} \tag{25}$$

Algorithm 1 Proposed Solution to Problem (10)

Input: $\mathbf{D}$, $\bar{\mathbf{q}}$, $\mathbf{W}$, ϵ
Output: $\mathbf{V}^\star$, $\mathbf{q}^\star$
Initialize: $\mathbf{q}^{(0)}$, $i \leftarrow 0$
Compute $\mathbf{V}^{\star(0)}$ from $\mathbf{q}^{(0)}$ using (15)
Compute $f^{(0)} = f(\mathbf{V}^{\star(0)}, \mathbf{q}^{(0)})$ using (11)
repeat
    $i \leftarrow i + 1$
    Compute $\mathbf{q}^{\star(i)}$ from $\mathbf{V}^{\star(i-1)}$ using (25)
    Compute $\mathbf{V}^{\star(i)}$ from $\mathbf{q}^{\star(i)}$ using (15)
    Compute $f^{(i)} = f(\mathbf{V}^{\star(i)}, \mathbf{q}^{\star(i)})$ using (11)
until $\frac{f^{(i-1)} - f^{(i)}}{f^{(i)}} \le \epsilon$ ▷ (Convergence)
Set $\mathbf{V}^\star \leftarrow \mathbf{V}^{\star(i)}$, $\mathbf{q}^\star \leftarrow \mathbf{q}^{\star(i)}$

C. Solution to the Joint Optimization Problem

Now that we have optimally solved problem (12) and problem (16), we employ an alternating optimization approach to find a solution to problem (10). The detailed steps are provided in Algorithm 1.

Theorem 1. *Algorithm 1 guarantees convergence to a partial optimum of problem (10).*

Proof. Since problem (10) is biconvex, and we have found optimal solutions to the subproblems, each iteration of the algorithm results in a lower value of the objective function, i.e., $f^{(i+1)} \le f^{(i)}$. Therefore, convergence is guaranteed by the monotone convergence theorem. Furthermore, at convergence, $f(\mathbf{V}^\star, \mathbf{q}^\star) \le f(\mathbf{V}, \mathbf{q}^\star)$ and $f(\mathbf{V}^\star, \mathbf{q}^\star) \le f(\mathbf{V}^\star, \mathbf{q})$ for all feasible $\mathbf{V}$ and $\mathbf{q}$, so $(\mathbf{V}^\star, \mathbf{q}^\star)$ is a partial optimum [26]. ■

IV. Distributed Implementation

A. Motivation

Direct implementation of Algorithm 1 requires centralized operation, with a CC that collects all CSI and the SPs' demands in each cell. The amount of data to be sent to the CC for processing can be large. In particular, the CSI from the BS of cell c to all users in the system is contained in $\mathbf{H}_c$, which is a complex matrix of size $N_c \times K_{\text{tot}}$. In addition to the CSI, each BS sends the SPs' demands to the CC, represented by the complex matrix $\mathbf{W}_c$ of size $K_c \times N_c$ and the real vector $\bar{\mathbf{q}}_c$ of length K_c. Thus, each BS needs to send $2K_{\text{tot}}N_c + 2K_cN_c + K_c$ real coefficients to the CC, totaling $\sum_c 2K_{\text{tot}}N_c + 2K_cN_c + K_c$ from all BSs. This quantity can be prohibitive especially in future large-scale MIMO systems.

B. Proposed Distributed Implementation

Instead, we propose a distributed coordinated approach to solve problem (10) by computing the decision variables locally at the BSs. To this end, we only require the BSs to communicate with each other, either via direct links or through some backhaul connection, to share partial information as needed. The proposed distributed coordinated approach gives an equivalent solution to the centralized approach in the previous section.

Aiming at solving the equivalent problem

$$\min_{\{\mathbf{V}_c\},\{\mathbf{q}_c\}} \quad \mathbb{E}\Big\{\left\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\Big\} \tag{26a}$$

$$\text{s.t.} \;\; 0 \preccurlyeq \mathbf{q} \preccurlyeq \mathbf{q}_{\max}, \tag{26b}$$

in a distributed manner, we consider the expression of the system deviation presented in (13). Noting that the problem is biconvex in terms of $\{\mathbf{V}_c\}_{c=1}^C$ and $\{\mathbf{q}_c\}_{c=1}^C$, we tackle each of the following subproblems separately.

1) Distributed Beamforming Solution to Problem (12): Note that the deviation form in (13) is separable in $\mathbf{V}_c$, i.e., $\mathbf{V}_c$ appears only once in the outer summation. This implies that the problem of minimizing $\mathbb{E}\Big\{\left\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\Big\}$ over $\{\mathbf{V}_c\}_{c=1}^C$ is equivalent to minimizing each entry of the outer summation. In other words, to solve the beamforming subproblem in (12), each BS needs to solve

$$\min_{\mathbf{V}_c} \|\mathbf{V}_c\mathbf{H}_{cc}\text{diag}(\mathbf{q}_c) - \mathbf{D}_c\text{diag}(\bar{\mathbf{q}}_c)\|_F^2 + \sum_{l\neq c}^{C} \|\mathbf{V}_c\mathbf{H}_{lc}\text{diag}(\mathbf{q}_l)\|_F^2 + \sigma_n^2 \|\mathbf{V}_c - \mathbf{W}_c\|_F^2. \tag{27}$$

This convex problem has the same closed-form solution presented in (15).

2) Distributed Power Control Solution to Problem (16): For the power control subproblem, we note that the deviation in (13) is not separable in $\mathbf{q}_c$. Nonetheless, it can be made separable as follows. Note that the issue of non-separation arises in the signal leakage term only. By manipulating the indices, we can rewrite the total signal leakage as

$$\sum_{c=1}^{C}\sum_{l\neq c}^{C} \|\mathbf{V}_c\mathbf{H}_{lc}\text{diag}(\mathbf{q}_l)\|_F^2 = \sum_{c=1}^{C}\sum_{l\neq c}^{C} \|\mathbf{V}_l\mathbf{H}_{cl}\text{diag}(\mathbf{q}_c)\|_F^2.$$

This has a nice interpretation in terms of inter-cell interference, corresponding to the fact that the total leakage going to all cells equals to the total leakage leaving all cells.

With this expression for the signal leakage, we can rewrite (13) to reflect this change as

$$\mathbb{E}\Big\{\left\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\Big\} = \sum_{c=1}^{C}\Bigg(\|\mathbf{V}_c\mathbf{H}_{cc}\text{diag}(\mathbf{q}_c) - \mathbf{D}_c\text{diag}(\bar{\mathbf{q}}_c)\|_F^2 + \sum_{l\neq c}^{C} \|\mathbf{V}_l\mathbf{H}_{cl}\text{diag}(\mathbf{q}_c)\|_F^2 + \sigma_n^2 \|\mathbf{V}_c - \mathbf{W}_c\|_F^2\Bigg). \tag{28}$$

This form is separable in $\mathbf{q}_c$, and thus solving

$$\min_{\mathbf{q}_c} \;\; \|\mathbf{V}_c\mathbf{H}_{cc}\text{diag}(\mathbf{q}_c) - \mathbf{D}_c\text{diag}(\bar{\mathbf{q}}_c)\|_F^2 + \sum_{l\neq c}^{C} \|\mathbf{V}_l\mathbf{H}_{cl}\text{diag}(\mathbf{q}_c)\|_F^2 + \sigma_n^2 \|\mathbf{V}_c - \mathbf{W}_c\|_F^2 \tag{29a}$$

$$\text{s.t.} \;\; 0 \preccurlyeq \mathbf{q}_c \preccurlyeq \mathbf{q}_{\max}, \tag{29b}$$

locally in each BS is equivalent to solving problem (16). The stationarity condition in (18) is replaced with

$$2q_{c,j}^\star\Bigg(\|\mathbf{a}_{cc,j}\|_2^2 + \sum_{l\neq c}\|\mathbf{a}_{cl,j}\|_2^2\Bigg) = 2\bar{q}_{c,j}\Re\big(\mathbf{a}_{cc,j}^H\mathbf{d}_{c,j}\big) - (\lambda_j^\star - \mu_j^\star) \;\; \forall j, \tag{30}$$

where $\mathbf{a}_{cl,j}$, and $\mathbf{d}_{c,j}$ represent the j^{th} column of the matrices $\mathbf{A}_{cl} = \mathbf{V}_l\mathbf{H}_{cl}$, and $\mathbf{D}_c$, respectively. Similar to how we have solved (16), here the KKT conditions (19)-(24), and (30) results in the following closed-form solution to problem (29):

$$q_{c,j}^\star = \begin{cases} \min\left\{\dfrac{\Re\big(\mathbf{a}_{cc,j}^H\mathbf{d}_{c,j}\big)}{\Big(\|\mathbf{a}_{cc,j}\|_2^2 + \sum\limits_{l\neq c}\|\mathbf{a}_{cl,j}\|_2^2\Big)}\bar{q}_{c,j}, q_{j\max}\right\}, & \Re\big(\mathbf{a}_{cc,j}^H\mathbf{d}_{c,j}\big) \geq 0, \\ 0, & \Re\big(\mathbf{a}_{cc,j}^H\mathbf{d}_{c,j}\big) < 0. \end{cases} \tag{31}$$

3) Distributed Alternating Optimization: The above distributed solutions to the subproblems are used in alternating optimization to solve problem (10). The algorithm steps are given in Algorithm 2. Note that the convergence of Algorithm 2 is guaranteed by Theorem 1. However, since the alternating optimization steps in Algorithm 2 are carried out over multiple communication rounds among the BSs, we need to carefully analyze the resultant communication overhead.

Algorithm 2 Distributed Algorithm for Problem (10)

Input: $\{\mathbf{D}_c\}$, $\{\bar{\mathbf{q}}_c\}$, $\{\mathbf{W}_c\}$, ϵ
Output: $\{\mathbf{V}_c^\star\}$, $\{\mathbf{q}_c^\star\}$
Initialize: $\{\mathbf{q}_c^{(0)}\}$, $i \leftarrow 0$
1: BS c computes $\mathbf{V}_c^{\star(0)}$ using (15), $\forall c$
2: Compute $f^{(0)} = f(\mathbf{V}^{\star(0)}, \mathbf{q}^{(0)})$ using (11)
3: **repeat**
4: $\quad i \leftarrow i + 1$
5: $\quad$ BS c shares $\|\mathbf{a}_{lc,j}\|_2$ with BS $l, \forall c, l \neq c, j \in [1 \cdots K_l]$
6: $\quad$ BS c computes $\mathbf{q}_c^{\star(i)}$ using (31), $\forall c$
7: $\quad$ BS c shares $\mathbf{q}_c^{\star(i)}$ with other BSs
8: $\quad$ BS c computes $\mathbf{V}_c^{\star(i)}$ using (15), $\forall c$
9: $\quad$ Compute $f^{(i)} = f(\mathbf{V}^{\star(i)}, \mathbf{q}^{\star(i)})$ using (11)
10: **until** $\frac{f^{(i-1)} - f^{(i)}}{f^{(i)}} \leq \epsilon$ $\qquad \triangleright$ (Convergence)
11: Set $\{\mathbf{V}_c^\star\} \leftarrow \{\mathbf{V}_c^{\star(i)}\}$, $\{\mathbf{q}_c^\star\} \leftarrow \{\mathbf{q}_c^{\star(i)}\}$

C. Communication Overhead

The BSs need to acquire all required information to compute (15) and (31) locally. Other than $\mathbf{Q}$, the expression in (15) contains information that are available at BS c. Thus, to compute $\mathbf{V}_c$, BS c needs to collect the power values of the users in other cells. That is, it needs to receive $K_{\text{tot}} - K_c$ real values. The total over all cells is $\sum_c K_{\text{tot}} - K_c = (C-1)K_{\text{tot}}$. For the power control solution in (31), the quantity $\mathbf{A}_{cl}$ is not present at BS c. This information needs to be transmitted from BS l to BS c, for all $l \neq c$. However, for the power computation, BS c does not need the entire matrix $\mathbf{A}_{cl}$, but only the norms of the columns of $\mathbf{A}_{cl}$. Thus, only K_c real values are sent from BS l to BS c, or a total of $(C-1)K_c$ real values received by BS c. The total over all cells is $\sum_c (C-1)K_c = (C-1)K_{\text{tot}}$. Hence, a total of $2(C-1)K_{\text{tot}}$ coefficients are shared between BSs for one iteration in Algorithm 2. If the alternating optimization approach requires T iteration to converge, the total number of shared coefficients is $2(C-1)K_{\text{tot}}T$. Note that the demands are not shared since they are computed locally at the BSs.

Compared with the centralized approach, Algorithm 2 can significantly reduce the communication overhead. If we assume a uniform distribution of users among cells, i.e., $K_c = K = \frac{K_{\text{tot}}}{C}$, and all BSs have the same number of antenna elements $N_c = N$, the ratio of the required overhead between the distributed and centralized approaches is:

$$\begin{aligned}\delta &= \frac{2(C-1)K_{\text{tot}}T}{\sum_c 2K_{\text{tot}}N_c + 2K_cN_c + K_c} \\ &= \frac{2(C-1)K_{\text{tot}}T}{2CK_{\text{tot}}N + 2K_{\text{tot}}N + K_{\text{tot}}} \\ &= \frac{2(C-1)T}{2(C+1)N+1} \leq \frac{T}{N}.\end{aligned}$$

This indicates that when $T \leq N$, using the distributed solution requires less overhead. As an example, consider the scenario of $N = 128$ antennas, $K_{\text{tot}} = 140$ users, and $C = 7$ cells. In this case, the centralized solution requires $286,860$ real parameters to be sent through the backhaul to the CC. In contrast, as will be shown in the simulation results in Section V, the distributed approach only requires on average 5161 real parameters to be transmitted between BSs.

V. Simulation Results

We conduct simulation in Matlab to study the performance of the proposed coordinated multi-cell WNV method. We also investigate the reduction in communication overhead by distributed implementation.

A. Simulation Setup

We consider a network of hexagonal cells each of radius 500 m with a BS at the center. We adopt the wrap-around implementation of cells to emulate uniform interference for all cells [29], with $C = 7$ neighboring cells. Unless otherwise specified, we set the number of SPs to $M = 4$ as default. Each SP m has $K_m = \frac{K_{\text{tot}}}{M}$ users, and they are evenly split among all cells, i.e., $K_{c,m} = \frac{K_{\text{tot}}}{7M}$. The users in each cell are uniformly distributed in space.

We model the channel from user k to each BS as a Rayleigh fading channel given by $\mathbf{h}_k = \beta_k^{1/2}\mathbf{g}_k$. Here, β_k is the large-scale fading coefficient that captures both pathloss and shadowing and is given as $10\log_{10}\beta_k = -31.54 - 33\log_{10}(d_k) + Z_k$, where d_k is the Euclidean distance from user k to the BS, and $Z_k \sim \mathcal{CN}(0, \sigma_z^2)$ is the shadowing at that user with $\sigma_z = 8$ dB; and $\mathbf{g}_k \sim \mathcal{CN}(0, \mathbf{I})$ denotes the small-scale fading. The users share a bandwidth of $B = 1$ MHz for transmission with a power budget of $p_{\text{max}} = q_{\text{max}}^2 = 27$ dBm for each user. We set the noise power spectral density to $N_0 = -174$ dBm/Hz, and the noise figure to $N_F = 2$ dB.

As an example of the virtualization requirements, we assume that the SPs set their demands with zero-forcing (ZF) beamforming and full power transmission. Hence, SP m's demand in cell c is given by the beamforming matrix $\mathbf{W}_{c,m} = (\mathbf{H}_{cc,m}^H\mathbf{H}_{cc,m})^{-1}\mathbf{H}_{cc,m}^H$ and the signal amplitude vector $\bar{\mathbf{q}}_{c,m} = \mathbf{q}_{\text{max}}$. This choice of power is known to maximize the sum-rate when ZF beamforming is used [30].

We study two important metrics. First, the normalized expected deviation, which indicates the quality of network virtualization, given by

$$\begin{aligned}&\frac{\mathbb{E}\{\left\|\hat{\mathbf{x}}^{\text{actual}} - \hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\}}{\mathbb{E}\{\left\|\hat{\mathbf{x}}^{\text{desired}}\right\|_2^2\}} \\ &\quad = \frac{\|\mathbf{VH}\text{diag}(\mathbf{q}) - \mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2\|\mathbf{V}-\mathbf{W}\|_F^2}{\|\mathbf{D}\text{diag}(\bar{\mathbf{q}})\|_F^2 + \sigma_n^2\|\mathbf{W}\|_F^2}.\end{aligned}$$

The second metric is the average per-user rate normalized by the system bandwidth, which is given by $R = \frac{1}{B}\frac{1}{K_{\text{tot}}}\sum_{k=1}^{K_{\text{tot}}} B_k \log_2(1 + \text{SINR}_k)$, where B_k is the bandwidth for user k, and SINR_k is the signal-to-interference plus noise ratio (SINR) of user k given by

$$\text{SINR}_k = \frac{\left|\mathbf{v}_k^T\mathbf{h}_k\right|^2 q_k^2}{\sum\limits_{j\in\mathcal{B}_k}\left|\mathbf{v}_k^T\mathbf{h}_j\right|^2 q_j^2 + \sigma_n^2\|\mathbf{v}_k\|_2^2},$$

where $\mathbf{v}_k$ is the receive beamforming vector for user k, and $\mathcal{B}_k$ is the set of users that share bandwidth with user k other than k itself.

B. Comparison Benchmarks

We compare the performance of our coordinated multi-cell WNV approach with three other methods. 1) A fully centralized non-virtualized approach, referred to as "Non-virtualized", where the InP uses full communication bandwidth to simultaneously serve all users with ZF beamforming and full power. 2) A common WNV method based on strict resource separation, referred to as "FD-WNV", where service isolation is performed by allocating different frequency bands to different SPs and dividing the bandwidth B equally among them. In FD-WNV, each SP uses ZF beamforming and maximum user transmit power. In both Non-virtualized and FD-WNV systems, full pilot ZF is used where each BS tries to suppress the interference coming from all other cells [31]. 3) A distributed non-coordinated multi-cell WNV, referred to as "Non-coordinated", where each BS uses the solutions developed in [21] for the single cell.

In the following results, our coordinated multi-cell WNV method is referred to as "Proposed". In addition, when comparing the proposed centralized and distributed approaches we refer to them as "Centralized" and "Distributed". For the proposed coordinated WNV, we initialize the power values in Algorithms 1 and 2 with full power, i.e., $\mathbf{q}^{(0)} = \mathbf{q}_{\text{max}}$, and we set $\epsilon = 0.01$. Note that both algorithms have the same performance but differ only in their communication overhead.

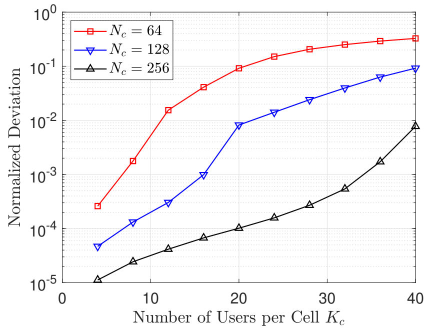


Fig. 1: Normalized deviation vs. K_c and N_c.

C. Deviation between SPs Demands and InP's Supply

Fig. 1 presents the normalized deviation between the InP's supply and the SPs' demands in the proposed coordinated WNV method. This figure indicates how well the proposed approach fulfills its main goal, i.e., service isolation. We observe that with a practical number of antennas, the proposed method can keep the deviation small. Recall that the SPs' demands correspond to an idealized setting where there is no inter-SP interference, as if each SP owned a separate copy of the entire network infrastructure. This suggests that, through proper beamforming design and user power control strategy, there is an opportunity to substantially increase system efficiency by limiting the deviation from the SPs' demands. This observation is further confirmed in terms of the average per-user rate in the results below.

D. Average Data Rate

Fig. 2 shows the average per-user rate with coordinated BSs versus the number of users in each cell K_c for various numbers of BS antennas N_c. As expected, we see a monotonically decreasing per-user rate in all systems. This figure shows a clear gap between the performance of FD-WNV and our proposed approach within the typical region of operation, i.e., $\sum_c K_c \leq N_c$. Even with overloaded systems, i.e., $\sum_c K_c > N_c$, the proposed approach still provides higher spectral efficiency compared with FD-WNV as long as $\sum_c K_c \ggg N_c$. Although the bandwidth separation in FD-WNV guarantees no inter-SP interference, the smaller bandwidth allocated to each SP causes a huge drop in the users' rates.

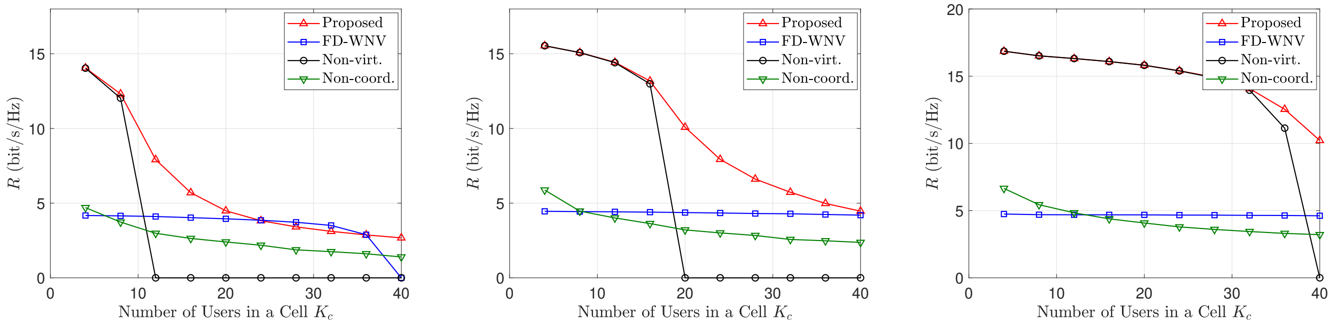
Furthermore, the proposed method outperforms even the non-virtualized method over a wide range of K_c values. This is clear in Fig. 2 when $\sum_c K_c$ is slightly less than N_c. Note that, unlike the non-virtualized method, our method performs virtualization. It achieves average rates that are at least as high as those achieved by the non-virtualized method, even when $\sum_c K_c \ll N_c$. We remark that the non-virtualized method here, which uses full-pilot ZF, is not defined when $\sum_c K_c > N_c$, an undesirable behavior for a beamforming strategy. The proposed coordinated methods, however, do not have such limitation. Fig. 2 also shows that the proposed coordinated WNV approach is far superior to the one without coordination in all settings. The poor performance of the non-coordinated solution is mainly due to the InP disregarding the inter-cell interference.

In Fig. 3, we show the average per-user rate as the number of SPs M varies from 1 to 7, while the total number of users remains at $K_{\text{tot}} = 56$. This figure illustrates that the performance of the proposed approach does not deteriorate as more SPs subscribe to the InP's services. This matches our expectation from Section II-A that the InP supplies an actual signal vector that does not depend on M. In contrast, we observe a drastic drop in the performance of FD-WNV. This is due to its strict separation of frequency bands between different SPs, which is a highly inefficient means to achieve service isolation. Further, note that the non-virtualized system is constant, which is not surprising since it neither performs virtualization nor deals with the SPs. It merely applies ZF beamforming to users regardless of their SPs. Moreover, the performance gap between the proposed and non-coordinated methods is also apparent in this figure.

E. Communication Overhead

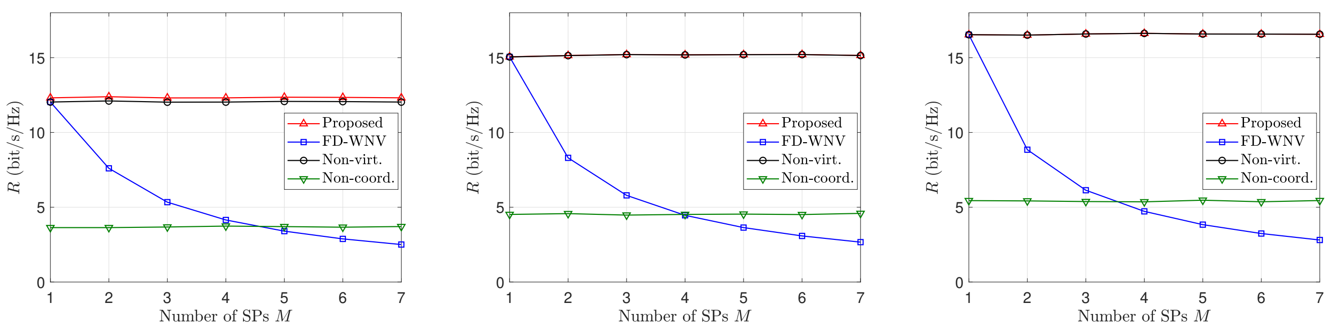
Fig. 4 shows the communication overhead versus K_c in both centralized and distributed approaches. We first note that for the centralized method, the size of the communication overhead increases with N_c, which is expected since the number of parameters is, as we have shown in Section IV-A, a linear function of the number of BS antennas. This has been frequently reported as an impediment in centralized multi-cell systems [32]–[34]. Fig. 4 suggests that as long as $\sum_c K_c \leq N_c$, the distributed solution requires much less communication overhead compared with the centralized solution. However, when $\sum_c K_c > N_c$ and interference dominates, the number of iterations required for the distributed algorithm to converge grows significantly, causing the communication overhead to increase. As K_c increases further, we observe that the number of iterations required for Algorithm 2 to converge starts to decrease, and therefore the communication overhead starts to decrease. It is expected that the number of antennas will increase to the thousands in future MIMO systems, which will allow the systems to avoid operating in the overloaded region. As we have observed here, in that case the distributed algorithm will converge faster and the communication overhead will be kept much smaller than that of the centralized solution, since it is not a function of N_c.

Recall that the results reported in Fig. 4 for the distributed approach is based on the default convergence threshold $\epsilon = 0.01$ in Algorithm 2. However, since the communication overhead depends on T, choosing different values for ϵ would lead to different amounts of communication overhead. Fig. 5 presents the communication overhead versus ϵ for different values of N_c when the number of users is kept at $K_c = 36$ per cell. This figure confirms that relaxing the convergence criterion lowers the communication overhead. Furthermore, in Fig. 6, we study the effect ϵ has on the optimization objective, i.e., the normalized deviation. Choosing a smaller ϵ means that precisely locating the optimum is of greater importance.



(a) $N_c = 64$ (b) $N_c = 128$ (c) $N_c = 256$

Fig. 2: Average per-user rate vs. K_c and N_c



(a) $N_c = 64$ (b) $N_c = 128$ (c) $N_c = 256$

Fig. 3: Average per-user rate vs. M and N_c.

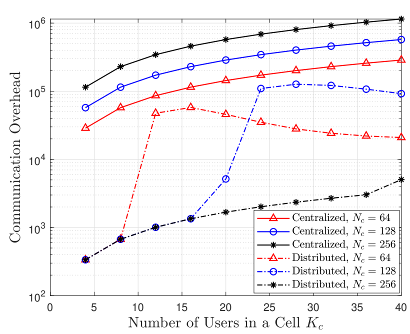


Fig. 4: Comm. overhead vs. K_c and N_c.

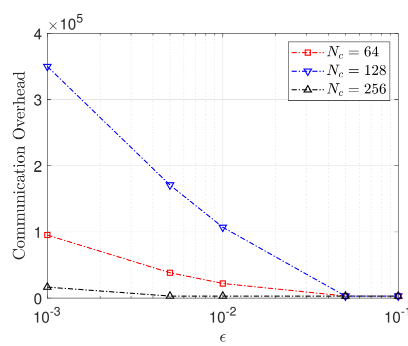


Fig. 5: Comm. overhead vs. ϵ and N_c.

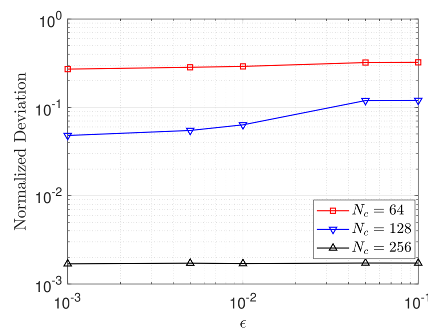


Fig. 6: Normalized deviation vs. ϵ and N_c.

This figure shows that as ϵ increases, the normalized deviation increases. Furthermore, the difference in the normalized deviation when $\epsilon \in [0.001\ \ 0.01]$ is minimal. This is the main reason behind our choice of $\epsilon = 0.01$ in the previous results.

VI. Conclusions

We have considered joint receive beamforming and power control optimization to minimize the expected deviation between the virtual demands from the SPs and the actual supply from the InP, in the uplink of a coordinated multi-cell WNV system. We decompose the problem into two alternating subproblems and derive closed-form solutions to both. We have further proposed an equivalent distributed solution, which can significantly reduce the communication overhead compared with the centralized one. We have observed substantial performance benefits by the proposed method. In our examples with 4 SPs and typical numbers of antennas and users, the proposed virtualization method provides 3 to 4 times the user data rate compared with traditional virtualization approaches that depend on strict resource separation among the SPs, and even higher performance gains for larger numbers of SPs. Furthermore, the benefit of virtualization can be achieved without loss of communication efficiency, as the proposed solution attains data rates similar to or higher than classical systems that do not provide virtualization.

REFERENCES

[1] X. Wang, P. Krishnamurthy, and D. Tipper, "Wireless network virtualization," in *Proceedings of the IEEE International Conference on Computer Communications and Networks (ICNC)*, 2013, pp. 818–822.

[2] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 358–380, 2015.

[3] N. M. Mosharaf Kabir Chowdhury and R. Boutaba, "Network virtualization: state of the art and research challenges," *IEEE Communications Magazine*, vol. 47, no. 7, pp. 20–26, 2009.

[4] N. M. K. Chowdhury and R. Boutaba, "A survey of network virtualization," *Computer Networks*, vol. 54, no. 5, pp. 862–876, 2010.

[5] A. Fischer, J. F. Botero, M. T. Beck, H. de Meer, and X. Hesselbach, "Virtual network embedding: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1888–1906, 2013.

[6] M. Richart, J. Baliosian, J. Serrat, and J.-L. Gorricho, "Resource slicing in virtual wireless networks: A survey," *IEEE Transactions on Network and Service Management*, vol. 13, no. 3, pp. 462–476, 2016.

[7] M. I. Kamel, L. B. Le, and A. Girard, "LTE wireless network virtualization: Dynamic slicing via flexible scheduling," in *Proceedings of the IEEE Vehicular Technology Conference (VTC-Fall)*, 2014.

[8] V. Jumba, S. Parsaeefard, M. Derakhshani, and T. Le-Ngoc, "Resource provisioning in wireless virtualized networks via massive-MIMO," *IEEE Wireless Communications Letters*, vol. 4, no. 3, pp. 237–240, 2015.

[9] S. Parsaeefard, R. Dawadi, M. Derakhshani, and T. Le-Ngoc, "Joint user-association and resource-allocation in virtualized wireless networks," *IEEE Access*, vol. 4, pp. 2738–2750, 2016.

[10] Z. Chang, Z. Han, and T. Ristaniemi, "Energy efficient optimization for wireless virtualized small cell networks with large-scale multiple antenna," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1696–1707, 2017.

[11] K. Zhu and E. Hossain, "Virtualization of 5G cellular networks as a hierarchical combinatorial auction," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2640–2654, 2016.

[12] G. Liu, F. R. Yu, H. Ji, and V. C. M. Leung, "Distributed resource allocation in virtualized full-duplex relaying networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8444–8460, 2016.

[13] S. Parsaeefard, R. Dawadi, M. Derakhshani, T. Le-Ngoc, and M. Baghani, "Dynamic resource allocation for virtualized wireless networks in massive-MIMO-aided and fronthaul-limited C-RAN," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9512–9520, 2017.

[14] Y. Liu, M. Derakhshani, S. Parsaeefard, S. Lambotharan, and K.-K. Wong, "Antenna allocation and pricing in virtualized massive MIMO networks via Stackelberg game," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5220–5234, 2018.

[15] D. Tweed, S. Parsaeefard, M. Derakhshani, and T. Le-Ngoc, "Dynamic resource allocation for MC-NOMA VWNs with imperfect SIC," in *Proceedings of the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017.

[16] D. Tweed and T. Le-Ngoc, "Dynamic resource allocation for uplink MIMO NOMA VWN with imperfect SIC," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2018.

[17] C. Liu, Y. Tao, and S. Xing, "Interference management in NOMA-enabled virtualized wireless networks," *Wireless Networks*, vol. 28, pp. 1457 – 1474, 2022.

[18] M. Soltanizadeh, B. Liang, G. Boudreau, and S. H. Seyedmehdi, "Power minimization in wireless network virtualization with massive MIMO," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018.

[19] J. Wang, M. Dong, B. Liang, G. Boudreau, and H. Abou-zeid, "Distributed coordinated precoding for MIMO cellular network virtualization," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 106–120, 2022.

[20] J. Wang, B. Liang, M. Dong, and G. Boudreau, "Online multicell coordinated MIMO wireless network virtualization with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 455–10 471, 2022.

[21] A. F. Almehdhar, B. Liang, M. Dong, Y. Ahmed, and G. Boudreau, "Beamforming and power control for wireless network virtualization in uplink MIMO systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2024.

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

[23] L. Venturino, N. Prasad, and X. Wang, "Coordinated linear beamforming in downlink multi-cell wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1451–1461, 2010.

[24] 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 Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1408, 2010.

[25] D. W. H. Cai, T. Q. S. Quek, C. W. Tan, and S. H. Low, "Max-min sinr coordinated multipoint downlink transmission—duality and algorithms," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5384–5395, 2012.

[26] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: a survey and extensions," *Mathematical Methods of Operations Research*, vol. 66, pp. 373–407, 2007.

[27] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Technical University of Denmark, 2012.

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

[29] A. K. Dinnis and J. S. Thompson, "The effects of including wraparound when simulating cellular wireless systems with relaying," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, 2007.

[30] H. V. Cheng, E. Björnson, and E. G. Larsson, "Optimal pilot and payload power control in single-cell massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2363–2378, 2017.

[31] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, 2016.

[32] K. Li, R. R. Sharan, Y. Chen, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized baseband processing for massive mu-MIMO systems," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, no. 4, pp. 491–507, 2017.

[33] X. Zhao, M. Li, Y. Liu, T.-H. Chang, and Q. Shi, "Communication-efficient decentralized linear precoding for massive mu-MIMO systems," *IEEE Transactions on Signal Processing*, vol. 71, pp. 4045–4059, 2023.

[34] J. Rodríguez Sánchez, F. Rusek, O. Edfors, M. Sarajlić, and L. Liu, "Decentralized massive MIMO processing exploring daisy-chain architecture and recursive algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 687–700, 2020.