Fair Licensed Spectrum Sharing Between Two MNOs Using Resource Optimization in Multi-Cell Multi-User MIMO Networks

Tianchen Wang®, Associate Member, IEEE, and Raviraj Adve®, Fellow, IEEE

Abstract—Licensed spectrum sharing has been a promised approach to provide mobile network operators (MNOs) with required spectrum at times of increased traffic, or to improve the mobile user data rate with limited spectral resources. In this paper, we investigate the use of multi-user multiple-input, multiple-output (MIMO) techniques to enable licensed spectrum sharing. Specifically, we present a fair spectrum sharing system between two MNOs in multi-cell multi-user MIMO networks. We impose fairness by ensuring that each MNO receives spectrum in proportion to the amount it contributes. We formulate a constrained optimization problem to determine resource allocation and user scheduling across two MNOs. Since the problem is non-convex, we develop an algorithm to provide an effective solution through fractional programming and block coordinate descent. Our numerical results illustrate that the proposed spectrum sharing scheme can achieve up to 60% improvement in terms of the average user rate among the two operators while ensuring that neither MNO is exploited for participating in the sharing mechanism. This improvement is in relation to the baseline of each MNO using multi-user MIMO communications on its own. In addition, most users, especially cell-center users close to the BSs, take advantage of our proposed spectrum sharing framework.

Index Terms—Spectrum sharing, resource allocation, user scheduling, coordinated beamforming, fairness, MIMO.

I. INTRODUCTION

A LONG with the development of modern wireless communications, the number of mobile devices has increased explosively over the years. However, the spectrum which plays a fundamental role in wireless communications remains a scarce resource [2]. Nonetheless, the spectrum is under-utilized, partly because of licensing spectrum on an exclusive basis by mobile network operators (MNOs). In this regard, allowing MNOs to share their licensed spectrum is an innovative way to approach the above problem. In addition, sharing spectrum could allow MNOs access to extra spectrum in times of increased traffic. We begin here with a comprehensive review of the relevant literature.

Licensed spectrum sharing can be categorized into heterogeneous spectrum sharing and homogeneous spectrum sharing [2]. In heterogeneous sharing, the spectrum is shared between different wireless communication systems (i.e., mobile and non-mobile systems). On the other hand, homogeneous sharing denotes the spectrum sharing within the same communication system (i.e., sharing among multiple MNOs). Licensed-Assisted Access (LAA) and Licensed Shared Access (LSA) are two typical heterogeneous spectrum sharing schemes, where the spectrum is shared between MNOs and an incumbent spectrum user. LAA is an LTE-unlicensed spectrum sharing scheme which allows MNOs to incorporate the 5 GHz unlicensed band together with their licensed spectrum to deliver wireless service to the end users.

In [3], the authors propose an auction scheme through Lyapunov optimization for the LTE-unlicensed users to share the same unlicensed spectrum with Wi-Fi in an opportunistic manner. It maximizes the overall social welfare and guarantees bounded drop rate and delay for the participating BSs. In [4], the authors apply game theory to study spectrum sharing within unlicensed bands among multiple strategic operators. The dynamics of the spectrum sharing market are investigated through an entry game. A dynamic sharing scheme among participating operators is proposed to work with the time-varying traffic.

Licensed Shared Access (LSA) is a spectrum sharing concept which enables MNOs (i.e., licensee operators (LOs)) to access additional spectrum resource from other spectrum owners (i.e., incumbent operators (IOs)) on a licensed shared base [5]. The authors in [6] propose a framework to discourage the LOs’ misbehaviour (i.e., non-granted access to IO’s spectrum) in LSA networks. The framework consists of a Dedicated Sensing Network (DSN) to sense the misbehaviour, a scoring system to penalize the misbehaviour and a resource allocation scheme to determine the priority in using the shared resource based on the score of each LO. Although the proposed scheme could successfully discourage the misbehaviour among the LOs, it could decrease the usage of the shared resource when severe penalty scheme is applied.

Homogeneous spectrum sharing among operators can be classified into mutual renting and spectrum pooling, when spectrum is the only shared resource. For mutual renting, each
MNO can rent their unutilized spectrum to other MNOs for additional revenue, and thus increases the spectrum utilization efficiency. For instance, in [7], the authors investigate how secondary network operators (i.e., buyers) should acquire spectrum access from primary network operators (i.e., sellers) based on the characteristics (i.e., cost and quality) of the available bands. Two optimization problems are formulated to study the questions on how SNOs could minimize cost while ensuring targeted quality-of-service (QoS), and how SNOs could achieve maximal utility gain with limited budget, respectively. However, mutual renting is an opportunistic spectrum sharing mechanism which is beneficial to relieve temporal spectrum shortage, yet cannot improve the whole system throughput in the long run.

In the case of spectrum pooling, in addition to the spectrum exclusively allocated to specific MNOs, there is assumed to exist an unallocated band spectrum operated by the national regulatory authority which is to be shared by MNOs. Alternatively, multiple operators agree to use their aggregate spectrum in an shared manner, where there is a resource manager, or a meta-operator to manage the spectrum allocation among the operators. Our proposed work falls into this category of spectrum pooling where two MNOs will agree to share spectrum, but exploit concepts from multiuser multiple input, multiple output (MIMO) networks to simultaneously use the same spectrum resources.

The spectrum pooling approach can be either cooperative or non-cooperative. The difference between the two schemes is whether the coordination of the spectrum allocation among MNOs is real-time or not. With non-cooperative spectrum pooling, the operators access the spectrum pool following predefined regulations or protocols. In [8], the authors propose three resource sharing models between two operators by overflow management. The proposed models reduce the blocking probability for both operators with low traffic intensity. If only one operator suffers from heavy traffic, it can attain a reduction in blocking probability as a tradeoff with a marginal increase in the blocking probability of the other operator. The work in [8] is revisited in [9], where the authors propose two different rules in deciding the amount of spectrum to be shared based on the same model. The degradation of blocking probability for the operator with lower traffic intensity (the lender) during the sharing has been minimized. However, the scenario when the traffic intensities for both operator are high is not investigated in either work.

Game theory is a useful tool to investigate the behaviour of MNOs when the spectrum sharing is non-cooperative. In [10], the authors investigate spectrum sharing amongst operators for device-to-device communications. They formulate a non-cooperative game to model the bargain of spectrum contributions of each participating operator to the shared spectrum pool. A unique equilibrium can be achieved by the proposed Jacobi-play strategy update and all operators can achieve performance gains compared with the non-sharing case. However, the gain and cost is not guaranteed to be proportional (i.e., although an operator could improve its performance, it may need to contribute more spectrum to the pool than others). We attempt to address this potential disparity.

The authors in [11] apply matching game theory to study non-orthogonal spectrum assignment among operators to maximize the sum rate of all operators. In the matching game, the distribution of transmit power at each BS is investigated using Q-learning. The study reveals that spectrum allocation contributes more to the gain in the sum rate than power allocation. However, the authors do not study the benefit brought to each operator (i.e., operators may not to be incentivized to join in spectrum sharing).

In [12], the authors used non-cooperative repeated games to model the interactions between two operators in a small (femto) cell. At each term, each operator proposes the number of component carriers (CCs) to borrow and lease among the K CCs in the spectrum pool. The CCs can be used exclusively by one operator once the number of CCs it borrows is no more than the number of CCs the other operator is willing to grant. The proposed model outperforms the case where there is no coordination between the operators. The same repeated game model as in [12] is used to study spectrum sharing in multi-operator heterogeneous networks in [13]. The spectrum is shared between the micro-cells and femto-cells of different operators. However, in [12] and [13], it is assumed that the operators have similar loads and are rational during the games, which is not realistic.

Our work, though, falls in the category of cooperative spectrum pooling wherein, spectrum is allocated to different operators in a centralized manner through a central controller or a resource manager. This cooperation requires an agreement amongst the MNOs involved to pool their resources for mutual benefit, such as when they face high traffic demands. In [14], the authors propose an algorithm for the spectrum manager to allocate the spectrum between two networks within a single cell, which improves the effective spectrum utilisation while maintaining the quality of service (QoS) of each network under a fluctuated traffic environment. The work shows that sharing spectrum within a single cell by multiple operators could improve the spectrum utilisation efficiency compared with fixed spectrum allocation schemes. However, this work does not guarantee fairness between the two networks in spectrum usage.

In [15], the authors investigate the spectrum sharing among MNOs at very short timescales with an instantaneous, dynamic sharing framework. A spectrum sharing manager regulates the resource allocation among the BSs of different MNOs. The proposed sharing scheme achieves increased spectral efficiency and reduced over-the-air latency compared to existing technologies (i.e., Wi-Fi). However, only a general coordination protocol is proposed without considering the wireless transmission specifications (e.g., path loss). In [16], the authors form a resource allocation problem that investigates the allocation of spectrum and BS density in an ultra-dense network (UDN) to minimize the cost of all operators while maintaining the average user rate. The tradeoff between spectrum usage and the density of operating BSs during the peak hours is illustrated. However, the case when both spectrum and BS density are limited is not studied in the work.

The authors in [17] reveal the superiority of non-orthogonal spectrum sharing over orthogonal spectrum sharing in terms
of user throughput. Orthogonal spectrum sharing refers to the scenario when a certain frequency band cannot be simultaneously accessed by multiple operators. On the other hand, with non-orthogonal spectrum sharing, operators can transmit signals to their users through the same frequency band concurrently, which brings out the issue of inter-operator interference. In [18], the authors investigate the potential gain of non-orthogonal spectrum sharing with a multiple input, single-output model. The authors conclude that only a limited improvement can be achieved through non-orthogonal spectrum sharing. However, the authors assume that BSs from both operators are co-located in each cell, which is unlikely in a real-world BS deployment.

Fairness among operators in accessing the spectrum is another concern in spectrum sharing and of importance in this paper. In [19], the author proposes a scheme for MNOs to share their spectrum within indoor small cells. Spectrum from all MNOs assemble a spectrum pool to be shared fairly by all users in a small cell. The amount each MNO pays for the spectrum is proportional to the spectrum usage of its own users. The performance of the proposed scheme is limited by the co-channel interference; under maximum interference, the spectrum sharing scheme does not improve the performance from the no-sharing case.

In summary, although importance of spectrum sharing is well documented, a cooperative spectrum sharing scheme that considers fairness between MNOs with real-world interference patterns has not been investigated. Furthermore, few works have investigated how the advent of physical layer technologies, such as cooperative MIMO systems, can help encourage spectrum sharing and enable maximizing the efficiency of spectrum usage. In this regard, the contributions of this paper are:

- We investigate non-orthogonal spectrum sharing under a deployment where BSs from different MNOs are not co-located for a multi-cell multi-user MIMO network, within the realm of cooperative spectrum pooling.
- We propose a spectrum sharing optimization framework with inter-operator coordinated beamforming between two MNOs which allows for simultaneous use by both while making fairness a key factor in determining the usage of the shared spectrum between the operators. In this regard, we exploit the beamforming capabilities of multi-antenna BSs to optimize spectrum allocation.
- We develop an algorithm based on fractional programming to manage the resource allocation and user scheduling among the two MNOs. These well-established techniques provide smooth convergence to stationary points of the optimization problem.
- To the best of the authors’ knowledge, this is the first work to investigate fair licensed spectrum sharing in multi-cell multi-user MIMO networks, from the perspective of centralized optimization. The proposed spectrum sharing mechanism can be incorporated into the modern LTE and 5G networks, especially in the sub-6GHz range. It is worth emphasizing that our work explores what is theoretically possible and would require significant additional developments for practical implementation. We hope to spark continued investigations in this field. Specifically, in this work, we assume perfect knowledge of both the inter-operator and intra-operator channel state information (CSI) among all of the users. The acquisition and sharing of CSI in cooperative networks is an active research area. Furthermore, in this paper, we focus on maximizing the spectral efficiency of the operators; specifically, we do not include economic incentives to participate.

The paper is organized as follows: in Section II, we present the system model at hand and devise the spectrum sharing scheme. Based on the sharing scheme, we formulate an optimization problem for the resource allocation and user scheduling. In Section III, we investigate and develop an algorithm to solve the proposed optimization problem. The performance of the spectrum sharing scheme is evaluated in Section IV. Section V ends the paper with some conclusions.

Notation: within this paper, we use the lower case letters to denote scalars (e.g., α, β), bold lower-case letter to denote vectors (e.g., v). We use the bold upper-case letter to denote matrices (e.g., D). For a vector x, ||x|| refers to its Euclidean norm. For a matrix A, A^† denotes the inverse of the matrix, and A^† refers to its conjugate transpose. Regarding to the sets of numbers, \( \mathbb{R} \) denotes the set of real numbers, \( \mathbb{C} \) denotes the set of complex numbers, and \( \mathbb{Z} \) denotes the set of integers. We use the symbol \( Re \{ \cdot \} \) to indicate the real part of a complex term. Finally, I is an identity matrix.

II. System Model and Problem Formulation

We consider the downlink of a multi-cell multi-user MIMO network comprising two operators sharing the same coverage area, denoted as OP1 and OP2 respectively. While our formulation can be extended to as many operators as desired, here we focus on two operators for the ease of exposition. The spectrum of OP1 and OP2 can be divided into \( N_1 \) and \( N_2 \) frequency slots, each of equal bandwidth individually. The operating area of each operator is divided into non-overlapping cells. Each operator has a base station (BS) located at the center of each cell to transmit signals to its own users within the cell using the frequency slots. We assume the cells of the two operators are diagonally overlapped, where the BSs of one operator stand at the corners of cells of the other operator. We use wraparound with a cluster of \( B \) cells to avoid network-edge issues and achieve more realistic results [20]. The distribution of users in each cluster is independent of each other. In addition, all wireless channels between BSs and users in a cluster are assumed independent. Finally, we assume there are \( K_1 \) users among all cells of OP1 and \( K_2 \) users among all cells of OP2.

For either operator, each BS is equipped with \( M \) transmit antennas, and each user is equipped with one receive antenna. We use \( H_{1,i,j} = [h_{1,i,j,1}, h_{1,i,j,2}, \ldots, h_{1,i,j,M}] \in \mathbb{C}^{M \times K_1} \) to denote the downlink channel between the BS at cell \( i \) of OP1 and users at cell \( j \) of OP1 in the \( j \)th frequency slot (\( 1 \leq j \leq N_1 + N_2 \)). Here, \( h_{1,i,j,k} \triangleq \sqrt{\beta_{1,i,j}^k} \cdot g_{1,i,j}^k \in \mathbb{C}^{M \times 1} \), where \( \beta_{1,i,j}^k = \left( \frac{d_{1,i,j}}{d_0} \right)^{-\alpha} \) accounts for the path-loss between the BS and the \( k \)th user belonging to OP1 and \( g_{1,i,j}^k \in \mathbb{C}^{M \times 1} \) accounts for the associated small-scale
fading. Within the expression of path-loss, \( d_{i,1}^k \) denotes the corresponding distances between the BS and the \( k^{th} \) user at the \( i^{th} \) cell, \( d_0 \) is the reference distance and \( \alpha \) is the pass-loss exponent [21]. \( g_{1,1,j}^m \), which denotes the small-scale fading between the \( m^{th} \) antenna and the \( k^{th} \) user, is a zero-mean, unit-variance complex Gaussian random variable, i.e., \( g_{1,1,j}^m \sim \mathcal{CN}(0,1) \). \( h_{2,1,1,j}^m \), \( h_{2,1,1,j}^m \) and \( h_{2,1,1,j}^m \) are similar, where the first subscript indicates which operator the BS belongs to while the second indicates the operator associated with the user.

With \( M \) transmit antennas per BS, within each cell, using linear processing, at most \( M \) users can be scheduled simultaneously in a single frequency slot [22]. Thus, there are up to \( M \) beamformers to be designed at each BS to facilitate the signal transmission to the scheduled users within each frequency slot. We use \( \mathbf{v}_{1,1,j}^m \in \mathbb{C}^{M \times 1} \) to denote the \( m^{th} \) beamformer at cell \( i \) with frequency slot \( j \). Moreover, we use \( \mathbf{u}_{1,1,j}^m \in \mathbb{Z}^{K_1 \times 1} \) to illustrate the user scheduling variables on the \( m^{th} \) beam through the \( j^{th} \) frequency slot in the \( i^{th} \) cell. Specifically, \( \mathbf{u}_{1,1,j}^m \) is a length-\( K_1 \) vector of comprising \( K_1-1 \) zeros and at most a single one indicating the scheduled user (i.e., if the \( k^{th} \) user is scheduled, then only the \( k^{th} \) entry of the vector \( \mathbf{u}_{1,1,j}^m \) is 1).

The implementation of spectrum sharing between the two operators is illustrated in Fig. 1. We apply the method of spectrum pooling with a spectrum pool comprising the \( N_1+N_2 \) frequency slots from both OP1 and OP2. This aggregate spectrum can be accessed by both operators in all cells. However, the spectrum is allocated to the two operators dynamically through a central controller, which manages the resource allocation and user scheduling among all cells from both operators. Specifically, before data transmission, both operators share their information (e.g., CSI, user distribution, etc.) to the central controller, as illustrated in the red arrows. Based on the information provided from the two operators, the central controller then determines the optimal beamforming vectors, user scheduling, and spectrum allocations for each operator, in order to maximize the overall performance of the system. The determined beamforming vectors, user scheduling, and spectrum allocations are sent back to the operators, as shown in the blue arrows, and thereafter used in the data transmissions to the users. Distributed implementations and reducing the overhead to implement our approach is outside the scope of this paper.

To impose fairness, the total number of frequency slots that can be used by each operator within each cell simultaneously is restricted, i.e., the operator contributing less to the spectrum pool is restricted to use fewer frequency slots than the other. We let OP1 to be the operator that provides more frequency slots (\( N_1 \geq N_2 \)) within the system, as such it has the right to access the whole spectrum pool (i.e., altogether \( N_1 + N_2 \) frequency slots). We impose proportional spectrum access fairness, i.e., OP2 can only use \( \frac{N_2}{N_1+N_2} \) of the frequency slots. In this way, the ratio of the amount of spectrum accessed by the two operators in spectrum sharing, remains consistent with the ratio of the amount of spectrum they contribute. Further assigning more spectrum to OP2 will disincentivize OP1 from participating, as the relative difference between the quantity of spectrum they can access narrows compared with the case of no spectrum sharing. On the other hand, OP2 may be reluctant to take part in spectrum sharing if fewer frequency slots are assigned to it, since the relative gap between the amount of spectrum it can access and the amount of spectrum OP1 can access enlarges. It is worth mentioning that, the proportional fairness here is simply one of many possible fairness metrics; any fairness condition, which incentivizes both operators to participate in spectrum sharing, can be applied. We use \( s_{1,i,j} \), \( s_{2,i,j} \in \{0,1\} \) to indicate that whether the \( j^{th} \) frequency slot is assigned to OP1 and OP2 respectively in the \( i^{th} \) cell. Note that, importantly, the same frequency slot may be used by both operators. Furthermore, it is not necessary for OP2 to use the same frequency slots in all its \( B \) cells. For each user, the intra-operator interference arises from the beams from all BSs within the \( B \) cells other than the beam on which the user is scheduled. Importantly, since both operators can access all slots, inter-operator interference arises from all \( B \) cells of the other operator.

---

1As mentioned, to focus on the performance upper bound, we assume knowledge of all channels at a central controller. Estimating the channels and disseminating this information is outside the scope of this paper. Standard channel MIMO estimation techniques may be used, however, this assumes coordination of training across both operators.
Accordingly, the intra-operator interference to the $m^{th}$ scheduled user through the $j^{th}$ frequency slot in the $i^{th}$ cell for OP1 and OP2 could be formulated respectively as:

\[
I_{1,1,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{1,i',j} \times |(H_{1,i',1,j} u_{1,i,j}^{m})|^{2} \]  \hspace{1cm} (1a)
\[
I_{2,2,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{2,i',j} \times |(H_{2,i',2,j} u_{2,i,j}^{m})|^{2} \]  \hspace{1cm} (1b)
\[
I_{1,2,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{1,i',j} \times |(H_{1,i',2,j} u_{1,i,j}^{m})|^{2} \]  \hspace{1cm} (2a)
\[
I_{1,2,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{1,i',j} \times |(H_{1,i',2,j} u_{1,i,j}^{m})|^{2} \]  \hspace{1cm} (2b)

Here, $(H_{1,i',1,j} u_{1,i,j}^{m})$ is a $M \times 1$ vector denoting the channel between the BS in the $(i')^{th}$ cell of the same operator and the scheduled user. For each user, only the inter-operator interference from the $B$ closest BSs of the other operator is considered. Thus, the inter-operator interference shares a similar form as the intra-operator interference in (1):

\[
I_{1,1,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{1,i',j} \times |(H_{1,i',1,j} u_{1,i,j}^{m})|^{2} \]  \hspace{1cm} (3a)
\[
I_{2,2,i,j} = \sum_{i' = 1}^{B} \sum_{m' = 1}^{M} s_{2,i',j} \times |(H_{2,i',2,j} u_{2,i,j}^{m})|^{2} \]  \hspace{1cm} (3b)

Using these interference expressions, the signal-to-interference-plus-noise ratio (SINR) for the user scheduled on the $m^{th}$ beam on the $j^{th}$ frequency slot in the $i^{th}$ cell of OP1 is given by

\[
\Gamma_{i,j}^{m} = \frac{s_{1,i,j} \times |(H_{1,i,1,j} u_{1,i,j}^{m})|^{2}}{\sigma^{2} + I_{1,1,i,j} + I_{2,1,i,j}} \]  \hspace{1cm} (4)

with an associated data rate of $R_{i,j}^{m} = BW \log(1 + \Gamma_{i,j}^{m})$. A similar expression can be written for users of OP2 as well. Consequently, the sum-rate of the whole system is given by $\sum_{i,j,m} (R_{1,i,j}^{m} + R_{2,i,j}^{m})$. Individual weights for the rate can be easily incorporated [23].

The central controller that manages the beamforming for each frequency slot in each cell and schedules users for both operators to maximize the sum-rate of all users within the system. Then, the objective function at hand is:

\[
f_{0}(s_{1}, s_{2}, u_{1}, u_{2}, v_{1}, v_{2}) = \max_{s_{1}, s_{2}, u_{1}, u_{2}, v_{1}, v_{2}} \sum_{i=1}^{N_{1}+N_{2}} \sum_{j=1}^{K} \left( \log(1 + \frac{s_{1,i,j} \times |(H_{1,i,1,j} u_{1,i,j}^{m})|^{2}}{\sigma^{2} + I_{1,1,i,j} + I_{2,1,i,j}}) + \log(1 + \frac{s_{2,i,j} \times |(H_{2,i,2,j} u_{2,i,j}^{m})|^{2}}{\sigma^{2} + I_{2,1,i,j} + I_{1,2,i,j}}) \right) \]  \hspace{1cm} (5)

Optimal resource allocation and user scheduling, therefore, requires the solution to the following constrained optimization problem:

\[
\max_{s_{1}, s_{2}, u_{1}, u_{2}, v_{1}, v_{2}} f_{0}(s_{1}, s_{2}, u_{1}, u_{2}, v_{1}, v_{2}) \]  \hspace{1cm} (5a)

subject to

\[
\sum_{m=1}^{M} \|v_{1,i,j,m}^{m}\|^{2} \leq P_{1}, \]  \hspace{1cm} (5b)
\[
\sum_{m=1}^{M} \|v_{2,i,j,m}^{m}\|^{2} \leq P_{2}, \quad \forall i, j; \]  \hspace{1cm} (5c)
\[
\sum_{k=1}^{K} (u_{1,i,j,m}^{m}) k = 1, \quad \sum_{k=1}^{K} (u_{2,i,j,m}^{m}) k = 1, \quad \forall i, j, m; \]  \hspace{1cm} (5d)
\[
\sum_{j=1}^{N_{1}+N_{2}} s_{1,i,j} \leq N_{1} + N_{2}, \quad \forall i; \]  \hspace{1cm} (5e)
\[
\sum_{j=1}^{N_{1}+N_{2}} s_{2,i,j} \leq \frac{N_{2}}{N_{1}} \cdot (N_{1} + N_{2}), \quad \forall i \]  \hspace{1cm} (5f)

Among the constraints in (5), (5b) imposes a power constraint on each frequency slot from both operators. The constraint (5c) enforces that each beam is allocated to at most one user. Moreover, the constraint (5d) ensures that one user can only be assigned to at most one beam within each frequency slot. (5e) and (5f) restrict the spectrum usage by each operator to guarantee the fairness between the two participating operators (while (5e) is redundant, we include it for completeness.)

### III. PROPOSED SOLUTION APPROACH

Solving the above constrained optimization problem is challenging since the objective function (5a) is non-convex on the beamformer $v$ and discontinuous on $u$ and $s$. Fortunately, we can formulate an iterative optimization scheme to tackle the problem based on the well-established technique of fractional programming [23], [24]. By means of fractional programming, the objective function (5a) can be made concave. Fractional programming uses the fact that the following sum-of-logarithms optimization problem:

\[
\max_{x} f(x) = \sum_{i} \log \left( 1 + \frac{A_{i}(x)}{B_{i}(x)} \right) \]  \hspace{1cm} (6)

has an equivalent sum-of-ratios form:

\[
\max_{x, \gamma} f_{x}(x, \gamma) = \sum_{i} \left( \log(1 + \gamma_{i}) - \gamma_{i} + \frac{(1 + \gamma_{i}) A_{i}(x)}{A_{i}(x) + B_{i}(x)} \right) \]  \hspace{1cm} (7)

where $\gamma_{i} = \frac{A_{i}(x)}{B_{i}(x)}$ is an auxiliary variable [24].

**Proof:** By introducing an auxiliary variable $\gamma_{i} = \frac{A_{i}(x)}{B_{i}(x)}$, the original sum-of-logarithms optimization problem (6) can
be rewritten as:
\[
\begin{align*}
\max_{x, \gamma} f(\gamma) &= \sum_i \log(1 + \gamma_i) \\
\text{subject to } \gamma_i &= \frac{A_i(x)}{B_i(x)} \tag{7}
\end{align*}
\]

The Lagrange dual function of the optimization problem (7) is:
\[
\begin{align*}
g(\nu) &= \max_{x, \gamma} \mathcal{L}(\gamma, \nu) \\
&= \max_{x, \gamma} \sum_i \left( \log(1 + \gamma_i) - \nu_i \left( \gamma_i - \frac{A_i(x)}{B_i(x)} \right) \right) \tag{8}
\end{align*}
\]

Since \(\mathcal{L}(\gamma, \nu)\) in (8) is concave differentiable over \(\gamma_i\) when \(x\) is fixed, the optimal \(\gamma_i^*\) can be achieved by setting \(\frac{\partial \mathcal{L}(\gamma, \nu)}{\partial \gamma_i} = 0\), which yields \(\gamma_i^* = \frac{1}{\nu_i} - 1\). Replacing the optimal \(\gamma_i^*\) back into (8), yields:
\[
\begin{align*}
g(\nu) &= \sum_i \left( \log \left( \frac{1}{\nu_i} \right) - \nu_i \left( \frac{1}{\nu_i} - 1 - \frac{A_i(x)}{B_i(x)} \right) \right) \tag{9}
\end{align*}
\]

The dual optimization problem of (7) is then:
\[
\begin{align*}
\min_{\nu} g(\nu) &= \min_{\nu} \sum_i \left( \log \left( \frac{1}{\nu_i} \right) - \nu_i \left( \frac{1}{\nu_i} - 1 - \frac{A_i(x)}{B_i(x)} \right) \right) \tag{10}
\end{align*}
\]

Since \(g(\nu)\) in (10) is convex differentiable over \(\nu_i\), the optimal \(\nu_i^*\) can be achieved by setting \(\frac{\partial g(\nu)}{\partial \nu_i} = 0\), which yields \(\nu_i^* = \frac{B_i(x)}{A_i(x)}\).

Since there is only one equality constraint in (7), Slater's condition is satisfied and thus strong duality holds, which implies:
\[
\begin{align*}
\max_{x, \gamma} f(\gamma) &= \max_{x, \gamma} \mathcal{L}(\gamma, \nu^*) \\
&= \max_{x, \gamma} \sum_i \left( \log(1 + \gamma_i) - \gamma_i + \frac{(1 + \gamma_i)A_i(x)}{A_i(x) + B_i(x)} \right) \tag{11}
\end{align*}
\]

Thus, the objective function (4) can be converted into the form of sum-of-ratios (12), shown at the bottom of the page, where \(I_{2,1,i,j}\) and \(I_{1,2,i,j}\) are specified in (2a) and (2b) respectively, and,
\[
\begin{align*}
\gamma_{m,i,j}^* &= \frac{s_{1,i,j}}{A_{1,i,j}} |(H_{1,i,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2 \tag{13}
\end{align*}
\]

The value of \(\gamma_{2,i,j}^*\) can be obtained similarly as \(\gamma_{m,i,j}^*\) in (13). Note that the new objective function \(f_\nu(\cdot)\) is not convex at this point, as the variables (i.e., \(x, u, v\)) appear in both the numerators and denominators among the fractions. Thus, \(f_\nu(\cdot)\) needs to be further transformed.

\[\text{A. Quadratic Transform for Multiple-Ratio Fractional Programming}\]

the following sum-of-ratio optimization problem:
\[
\begin{align*}
\max_x \sum_i A_i(x) \tag{14a} \\
\text{subject to } x \in X \tag{14b}
\end{align*}
\]

is equivalent to the convex optimization problem:
\[
\begin{align*}
\max_{x,y} \sum_i (2y_i \sqrt{A_i(x)} - y_i^2 B_i(x)) \tag{15a} \\
\text{subject to } x \in X \tag{15b}
\end{align*}
\]

where \(y\) is an auxiliary variable. Here, \(A_i(x)\) is a non-negative function and \(B_i(x)\) is a positive function. With the optimal auxiliary variable \(y^*\), both (14) and (15a) share the same optimal solution \(x^*\) as well as the same optimum of the objective functions.

\[
\begin{align*}
\frac{f_\nu(s_1, s_2, u_1, u_2, v_1, v_2)}{\sigma^2 + \sum_{i'}^B \sum_{i''}^M |H_{1,i',1,j}|^2 |H_{2,i',2,j}|^2 |(H_{1,i',1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} + \frac{f_\nu(s_2, s_1, u_1, u_2, v_1, v_2)}{\sigma^2 + \sum_{i'}^B \sum_{i''}^M |H_{1,i''1,j}|^2 |H_{2,i''2,j}|^2 |(H_{1,i''1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} = \left(1 + \gamma_{m,i,j}^*\right) |s_{1,i,j}| |(H_{1,i,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2 + \left(1 + \gamma_{m,i,j}^*\right) |s_{2,i,j}| |(H_{2,i,j}u_{2,i,j}^{m})|^2 v_{2,i,j}^m |^2 + I_{1,2,i,j}
\end{align*}
\]

\[
\begin{align*}
\frac{f_\nu(s_1, s_2, u_1, u_2, v_1, v_2, y_1, y_2)}{\sum_{i'}^B \sum_{i''}^M |H_{1,i',1,j}|^2 |H_{2,i',2,j}|^2 |(H_{1,i',1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} + \frac{f_\nu(s_2, s_1, u_1, u_2, v_1, v_2, y_1, y_2)}{\sum_{i'}^B \sum_{i''}^M |H_{1,i''1,j}|^2 |H_{2,i''2,j}|^2 |(H_{1,i''1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} = \left(1 + \gamma_{m,i,j}^*\right) |s_{1,i,j}| |(H_{1,i,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2 + \left(1 + \gamma_{m,i,j}^*\right) |s_{2,i,j}| |(H_{2,i,j}u_{2,i,j}^{m})|^2 v_{2,i,j}^m |^2 + I_{1,2,i,j}
\end{align*}
\]

\[
\begin{align*}
\frac{f_\nu(s_1, s_2, u_1, u_2, v_1, v_2, y_1, y_2)}{\sum_{i'}^B \sum_{i''}^M |H_{1,i',1,j}|^2 |H_{2,i',2,j}|^2 |(H_{1,i',1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} + \frac{f_\nu(s_2, s_1, u_1, u_2, v_1, v_2, y_1, y_2)}{\sum_{i'}^B \sum_{i''}^M |H_{1,i''1,j}|^2 |H_{2,i''2,j}|^2 |(H_{1,i''1,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2} = \left(1 + \gamma_{m,i,j}^*\right) |s_{1,i,j}| |(H_{1,i,j}u_{1,i,j}^{m})|^2 v_{1,i,j}^m |^2 + \left(1 + \gamma_{m,i,j}^*\right) |s_{2,i,j}| |(H_{2,i,j}u_{2,i,j}^{m})|^2 v_{2,i,j}^m |^2 + I_{1,2,i,j}
\end{align*}
\]
The detailed proof of the preceding quadratic transform is presented in [24]. Thus, by incorporating the quadratic transform to the sum of ratios terms, the objective function in (12) can be reformulated correspondingly (16), shown at the bottom of the previous page.

Here, \( f_q \) is a concave function of \( y \), such that the optimal solution for \( y_{1,i,j}^m \) can be achieved by setting \( \frac{\partial f_q}{\partial y_{1,i,j}} = 0 \) with all other variables fixed, i.e.,

\[
y_{1,i,j}^m = \frac{s_{1,i,j}^1(v_{1,i,j}^m) \gamma_{1,i,j}}{\sigma^2 + \sum_{i'v' = 1}^{B} \sum_{m' = 1}^{M} s_{1,i',j}^1 \gamma_{1,i',j} \left| v_{1,i',j}^m \right|^2 + \sum_{i'v' = 1}^{B} \sum_{m'' = 1}^{M} s_{2,i',j} \left| v_{2,i',j}^m \right|^2} \left( H_{1,i,j}^m u_{1,i,j}^m \right) \left( \frac{1 + \gamma_{1,i,j}^m}{\gamma_{1,i,j}} \right)^{1/2} \left( H_{2,i,j}^m u_{2,i,j}^m \right) \left| v_{2,i,j}^m \right|^2 \right)
\]

(17)

A corresponding expression for \( y_{2,i,j}^m \) can be attained. The objective function \( f_q \) is concave in \( v \) as well. Thus, with fixed \( s, u, \gamma \) and \( y \), the optimal solution of \( v \) can be achieved by solving the following constrained convex optimization problem:

\[
\max_{v_1, v_2} f_q(v_1, v_2)
\]

subject to (5b)

(18a)

The method of Lagrange multipliers allows us to solve the convex optimization problem (18), where the Lagrangian over \( v \) is

\[
L_q(v_1, v_2) = f_q(v_1, v_2, \ldots) - \sum_{i,j} \left( \lambda_{1,i,j} \left( \sum_{m=1}^{M} (v_{1,i,j}^m)^{v_{1,i,j}^m} - P_1 \right) \right) + \left( \lambda_{2,i,j} \left( \sum_{m=1}^{M} (v_{2,i,j}^m)^{v_{2,i,j}^m} - P_2 \right) \right)
\]

(19)

By solving \( \frac{\partial L_q}{\partial v_{1,i,j}^m} = 0 \), we can achieve a closed form solution for the vector \( v_{1,i,j}^m \):

\[
v_{1,i,j}^m = s_{1,i,j} D_{1,i,j}^{-1} \sqrt{1 + \gamma_{1,i,j}^m} (y_{1,i,j}^m)^{(H_{1,i,j}^m u_{1,i,j}^m)^{(H_{2,i,j}^m u_{2,i,j}^m)}}
\]

(20)

where

\[
D_{1,i,j} = \lambda_{1,i,j} I + s_{1,i,j} \sum_{i'v' = 1}^{B} \sum_{m' = 1}^{M} \left| y_{1,i',j}^m \right|^2 \times (H_{1,i,j}^m u_{1,i,j}^m)^{(H_{1,i',j}^m u_{1,i',j}^m)}
\]

\[
+ s_{1,i,j} \sum_{i'v' = 1}^{B} \sum_{m' = 1}^{M} \left| y_{2,i',j}^m \right|^2 \times (H_{1,i,j}^m u_{1,i,j}^m)^{(H_{2,i',j}^m u_{2,i',j}^m)}
\]

(21)

The Lagrange multiplier \( \lambda_{1,i,j} \) can be obtained through bi-section search on the power. The optimal \( v_{2,i,j}^m \) can be obtained similarly.

Since the interference to each scheduled user depends only on the beamformers at each BS regardless of the users scheduled in the beams, the optimization of user scheduling can be decoupled from the optimization of beamformers. Importantly, the user scheduling variable \( u \) can be optimized individually keeping the other variables fixed. Specifically, optimal user scheduling \( u^* \) can be achieved through the following integer optimization problem with respect to \( u \):

\[
\max_{u_1, u_2} f_q(u_1, u_2)
\]

subject to (5c), (5d)

(22a)

where all the other variables are kept fixed. To help simplify the objective function (22a), we introduce a new function \( R(\cdot) \) to include all terms and variables within \( f_q \) with regard to the \( i^{th} \) cell and the \( j^{th} \) frequency slot of a single operator,

\[
R(u_{1,i,j}) = \sum_{m=1}^{M} \left\{ (1 + \gamma_{1,i,j}) - \gamma_{1,i,j} - |y_{1,i,j}^m|^2 \sigma^2
\right.
\]

\[
+ 2Re \{ (y_{1,i,j}^m)^{(H_{1,i,j}^m u_{1,i,j}^m)} \}
\]

\[
\left. + \frac{B}{2} \sum_{i'v'=1}^{B} \sum_{m''=1}^{M} s_{2,i',j} \left| (H_{2,i,j}^m u_{2,i,j}^m)^{(v_{2,i,j}^m)} \right|^2 \right\}
\]

(23)

such that \( f_q \) could be re-written as:

\[
f_q(u_1, u_2) = \sum_{i=1}^{B} \sum_{j=1}^{N_1+N_2} \left( R(u_{1,i,j}) + R(u_{2,i,j}) \right)
\]

(24)

Since \( R(u_{1,i,j}) \) and \( R(u_{2,i,j}) \) are independent of the user scheduling in the other cells and frequency slots of both operators, the optimal solution for (22a) can be obtained by solving \( \max_{u_{1,i,j}} R(u_{1,i,j}) \) and \( \max_{u_{2,i,j}} R(u_{2,i,j}) \), \( v_{i,j} \) separately with constraints (5c) (5d). In each of the above constrained optimization problems on \( R(\cdot) \) can be described as a maximum weighted bipartite matching problem where we need to obtain the optimal one-to-one matches between the beams and the users, as shown in Fig. 2.

In the complete bipartite graph, the nodes in the right denote the \( M \) beams and the nodes in the left denote the users to be scheduled. The weight of the edge between the \( k^{th} \) user node and the \( m^{th} \) beam node is \( W_{k,m} \). Considering the optimization for the \( i^{th} \) cell and \( j^{th} \) frequency slot of OP1 we have,

\[
W_{k,m} = \log(1 + \gamma_{1,i,j}^m) - \gamma_{1,i,j}^m - |y_{1,i,j}^m|^2 \sigma^2
\]

\[
+ 2Re \{ (y_{1,i,j}^m)^{(1 + \gamma_{1,i,j}^m)} \}
\]

\[
\left. + \frac{B}{2} \sum_{i'v'=1}^{B} \sum_{m''=1}^{M} s_{2,i',j} \left| (H_{2,i,j}^m u_{2,i,j}^m)^{(v_{2,i,j}^m)} \right|^2 \right\}
\]

(25)

Unlike in (23), \( \tilde{u}^k \) is used instead of \( u_{1,i,j}^m \) in (25). Here, \( \tilde{u}^k \in \mathbb{R}^{K1 \times 1} \) is a unit vector with the \( k^{th} \) element equaling to one. Using this vector, \( W_{k,m} \) represents the contribution to
Based on the solution, the optimal and one user can only be scheduled with at most one beam. Constraints (5c) (5d), i.e., each beam must have a user assigned the user nodes and the beam nodes, the solution satisfies the constraints (5c) (5d), i.e., each beam must have a user assigned and one user can only be scheduled with at most one beam. Based on the solution, the optimal \((u_{i,i,j}^m)^*\) can be achieved by setting one to the \(k^{th}\) term from a zero vector of length \(K_1\), if the \(k^{th}\) user is scheduled within the \(m^{th}\) beam.

With regard to the optimal selection of frequency slots \(s^*\) that maximizes \(f_g\), we observe that \(f_g\) can be rearranged into a linear form with respect to \(s\) when the other variables are fixed:

\[
f_g(s_1, s_2) = B \sum_{i=1}^{N} \sum_{j=1}^{Q} 2Re\{(y_{1,i,j}^m)^\dagger \sqrt{1 + \gamma_{1,i,j}^m} \langle v_{1,i,j}^m, H_{1,i,j} u_{1,i,j}^m \rangle \}
- \sum_{i'=1}^{B} \sum_{m'=1}^{M} |y_{i',i,j}^m|^2 \sum_{m=1}^{M} |\langle H_{1,i',j} u_{1,i,j}^m, v_{1,i,j}^m \rangle|^2
- \sum_{i'=1}^{B} \sum_{m'=1}^{M} |y_{i,j,i'}^m|^2 \sum_{m=1}^{M} |\langle H_{1,i,j} u_{1,i,j}^m, v_{1,i,j}^m \rangle|^2
+ \text{const}
\]

where

\[
Q_{1,i,j} = \sum_{m=1}^{M} 2Re\{\langle y_{1,i,j}^m, v_{1,i,j}^m \rangle \langle H_{1,i,j} u_{1,i,j}^m \rangle \}
- \sum_{i'=1}^{B} \sum_{m'=1}^{M} |y_{i',i,j}^m|^2 \sum_{m=1}^{M} |\langle H_{1,i',j} u_{1,i,j}^m, v_{1,i,j}^m \rangle|^2
- \sum_{i'=1}^{B} \sum_{m'=1}^{M} |y_{i,j,i'}^m|^2 \sum_{m=1}^{M} |\langle H_{1,i,j} u_{1,i,j}^m, v_{1,i,j}^m \rangle|^2
\]

(27)

Importantly, this implies that the optimizations on \(s_1\) and \(s_2\) can be independent, since the objective function (26) is linear in both \(s_1\) and \(s_2\). Meanwhile, the constraints (5e) and (5f) regulate the usage of the spectrum for the two operators individually. Then the optimal \(s_1^*\) and \(s_2^*\) can be achieved by solving the following optimization problems individually:

\[
\max_{s_1} \sum_{i=1}^{B} \sum_{j=1}^{N} Q_{1,i,j} s_{1,i,j} \quad (28a)
\]

subject to (5e), \(s_{1,i,j} \in \{0,1\}\) (28b)

and

\[
\max_{s_2} \sum_{i=1}^{B} \sum_{j=1}^{N} Q_{2,i,j} s_{2,i,j} \quad (29a)
\]

subject to (5f), \(s_{2,i,j} \in \{0,1\}\) (29b)

(28) and (29) can be solved effectively through integer programming. However, the optimal \(s_1\) and \(s_2\) can also be achieved in a greedy manner per cell. We first sort the frequency slots in descending order with respect to the values of \(Q\), and then set the corresponding \(s = 1\) if \(Q\) is positive for the first \(N_c\) sorted frequency slots. Here, \(N_c\) denotes the maximum number of frequency slots assigned to an operator, i.e.,

\[
N_c = \begin{cases} 
N_1 & \text{OP1} \\
\frac{N_2}{N_1} \cdot (N_1 + N_2) & \text{OP2}
\end{cases}
\]

Claim 1: For the problems in (28) and (29), i.e., keeping all other variables fixed, the frequency slot selection variable \(s^*\) obtained through the proposed greedy algorithm above is the globally optimal solution.

**Proof:** Assume there exists one frequency slot selection \(s'\) such that it reaches a superior outcome than \(s^*\). This would imply

\[
\sum_{j=1}^{N} Q_j s'_j > \sum_{j=1}^{N} Q_j s^*_j
\]

(31)

The relationship between the left term and the right term in (31) can be understood through the sign of the difference between them, i.e., \(\sum_{j=1}^{N} Q_j s'_j - \sum_{j=1}^{N} Q_j s^*_j\), which could be re-formulated as the following,

\[
\sum_{j=1}^{N} Q_j (s'_j - s^*_j) + \sum_{j=1}^{N} Q_j (s^*_j - s^*_j)
\]

(32)

In (32), the terms with positive and negative values of \(Q\) are separated. The first scenario to be considered is when the number of positive \(Q\) is no more than \(N_c\), i.e., \(\sum_{j=1}^{N} Q_j (Q_j > 0) \leq N_c\). In this case, \(s^*_j = 1\) only if \(Q_j > 0\) according to the proposed greedy approach. Then, \(\sum_{j=1}^{N} Q_j (s'_j - s^*_j) = \sum_{j=1}^{N} Q_j (s'_j - 1) \leq 0\), since \(s'_j \leq 1\) and \(Q_j > 0\). Identically, \(\sum_{j=1}^{N} Q_j (s^*_j - s^*_j) = \sum_{j=1}^{N} Q_j (s^*_j - 0) \leq 0\), since \(s^*_j \geq 0\) and \(Q_j < 0\) here. Thus, the value of (32) is not positive when \(\sum_{j=1}^{N} Q_j (Q_j > 0) \leq N_c\).
The other scenario is when there are more than $N_c$ positive $Q$s, i.e., $\sum_{j=1}^{N} Q_j (Q_j > 0) > N_c$. Similar as the previous case, $s_j^* = 1$ only if $Q_j > 0$, such that $\sum_{j=1}^{N} Q_j (s_j^* - s_j^0) = 0$. Moreover, the $Q_j$ is among the top $N_c$ largest, $\sum_{j=1}^{N} Q_j (s_j^* - s_j^0) = \sum_{Q_j > 0} Q_j s_j^* - \sum_{Q_j > 0} Q_j s_j^0 < 0$. Hence, the value of (32) is not positive as well when $\sum_{j=1}^{N} Q_j (Q_j > 0) > N_c$. In conclusion, $\sum_{j=1}^{N} Q_j s_j^* - \sum_{Q_j > 0} Q_j s_j^0 \leq 0$, which contradicts (31). Therefore, there is no solution better than $s^*$.

Based on the optimization of each individual variable as indicated above, we can gather all the optimization steps into an iterative optimization algorithm, as presented in Algorithm 1. The algorithm implements block coordinate descent to determine the resource allocation and user scheduling within the system. In each step, one type of variable is optimized keeping the others fixed. Since the objective function is upper bounded and each step can only improve on the value of the objective function, $f_q$, the algorithm converges in a non-decreasing manner.

In Algorithm 1, for a single iteration, the time complexity of Step 4 - 6 is $O(BNM)$, as the corresponding parameters (i.e., $\gamma$, $y$, and $v$) are updated for every user in all cells with all frequency slots. The Hungarian algorithm is applied in user scheduling in each cell and with each frequency slot, resulting in the total time complexity of $O(BNM^2K)$ for Step 7. In Step 8, for each cell, the value of $Q$ is calculated per frequency slot. The values of $Q$s are then sorted in the greedy algorithm used to determine the optimal values of $s$. Thus in total, the time complexity for Step 8 is $O(B(N + N \log(N)))$. Consequently, the computational complexity of the proposed algorithm is given by $O(BNM + BNM^2K + B(N + N \log(N)))$ for every iteration.

To summarize our contribution here, we have formulated an optimization problem that encourages spectrum sharing while imposing fairness constraints. The problem is solved using block coordinate descent as listed in Algorithm 1. The output of the algorithm is, in each cell, the frequency slots used by each operator, the users (of each operator) scheduled and the associated beamforming weights (which, implicitly includes power allocation).

### IV. Simulation Results and Performance Evaluation

In this section, we illustrate the gains made possible by our resource allocation scheme. The numerical values of the parameters used in the simulations are given in Table I. We fix the number of frequency slots provided by OP1 at 16 (i.e., $N_1 = 16$), and vary the number of frequency slots provided by OP2 from 4 up to 16; this allows us to investigate the impact of unequal contributions to the spectrum pool. Note that, as per the constraint in (5f), if it contributes fewer than 4 frequency slots, OP2 would not receive any additional frequency slots, i.e., operators must contribute some minimum amount for a fair spectrum sharing to be possible. The number of users from OP1 and OP2 within each cell is set to be $M$ times the number of frequency slots provided by OP1 and OP2 respectively (i.e., $K_1 = M \times N_1$ and $K_2 = M \times N_2$). In our case, we assume each BS is deployed with $M = 8$ antennas.

Fig. 3a illustrates the simple layout we consider in simulation, but a more sophisticated cell structure does not change our framework. As seen in the figure, with wraparound, a cluster of $B$ cells is surrounded by another $B - 1$ cloned clusters, and so on (we use $B = 9$ in our simulation). Thus, the cellular layout of each operator consists of infinitely duplicated clusters with identical distributions of users, as well as channels between each BS and each user. Importantly, as seen in the four colored areas in Fig. 3b, different areas within any cell of either operator will suffer completely different sources of their main inter-operator interference.

To evaluate the performance of the proposed spectrum sharing mechanism, we compare the spectrum efficiency (defined as $\text{(sum rate)}/(\text{number of frequency slots})$, as well as the average user rate of each operator achieved through the following cases:

- **No spectrum sharing**: Each operator uses its own spectrum exclusively as in the conventional wireless communications framework. The resource allocation and user scheduling for each operator are achieved individually using the algorithm presented in [23].
- **Spectrum sharing with same number of frequency slots as provided**: The spectrum pool is divided into two disjoint sets comprising $N_1$ and $N_2$ frequency slots respectively, which are then allocated to the operators in accordance with the number of frequency slots they

| VARIABLE | VALUE |
|----------|-------|
| Cell size per side | 2000 m |
| BS transmit power | $P_1 = P_2 = 43$ dBm |
| Frequency slot bandwidth | 1.25 MHz |
| Noise power spectral density | -174 dBm/Hz |
| Path-loss exponent | $\alpha = 3.7$ |
| Reference distance | $d_0 = 0.4$ m |
provide. Note that, in this case, the same frequency slots are allocated across all $B$ cells, such that there is no inter-operator interference (as in the no spectrum sharing case). Effectively, in this scheme, the operators exchange their frequency slots in order to achieve maximum rate across all cells.

*Fair spectrum sharing:* The spectrum sharing framework proposed in Section II, where OP1 can transmit signals over all frequency slots, while OP2 can merely access to at most $\lfloor \frac{N_1}{N_2} \cdot (N_1 + N_2) \rfloor$ frequency slots simultaneously, as in constraint (5e).

*Full spectrum sharing:* Both operators can transmit signals over all frequency slots together in this case, ignoring the contributions to the spectrum pool. In this case, the constraints (5f) is turned into $\sum_{j=1}^{N_1+N_2} s_{2,j} \leq N_1 + N_2$, $\forall i$, as the constraint (5e).

Due to the discrete variables (i.e., $u$, $s$) within the optimization problem, the optimal solution is sensitive to the initialization of the variables. Based on a suggestion in [23], we use a greedy initialization for $u$ based on the channels between the BSs and the users, i.e., the $M$ users with the best $M$ channels are scheduled within each frequency slot. The beamforming vector $v$ for each scheduled user is matched to the channel, i.e., proportional to the corresponding channel and normalized to satisfy the total power constraint per frequency slot.

The objective function has an upper bound and each step in the algorithm improves its value, hence, the algorithm converges. In our experience, the proposed algorithm converges within 20 iterations of updating the variables among the $B$ cells of both operators; Fig. 4 represents a typical convergence curve for one realization of channels and users, i.e., one run of the algorithm.

Fig. 5 plots the spectrum efficiency, defined as the aggregate data rate (including both OP1 and OP2) per frequency slot. Specifically, the figure plots the spectrum efficiency as a function of the number of frequency slots contributed by OP2 ($N_2$). We observe that the proposed spectrum sharing mechanism can increase the spectrum efficiency by up to 60%, compared to the case where there is no spectrum sharing, when both operators make equal contributions to the spectrum.
Fig. 7. Rate gain compared with no spectrum sharing.

The average user rate among users from OP1 and OP2 are presented in Fig. 6a and Fig. 6b, respectively. Moreover, the gain of fair spectrum sharing and full spectrum sharing over no spectrum sharing in terms of the user average rate of each operator is given in Fig. 7. Interestingly, under the proposed fair spectrum sharing scheme, it is OP1, not OP2, that experiences a higher gain compared with the full sharing case. This is because, under fair spectrum sharing, OP1 will still have nearly exclusive access to several frequency slots in each cell due to the restrictions on the usage of OP2. In addition, OP1 always achieves a better gain than OP2. This suggests OP1 has an incentive to participate even if it contributes more than OP2. When there is a wide difference between the number of frequency slots provided by OP2 \( N_2 \) and OP1 \( N_1 \), the average user rate of OP2 decreases compared to the “No Sharing” case. This is because under this circumstance, there is more benefit towards the objective function \( f_o \) through minimizing the inter-operator interference from OP2 to OP1, compared to maximizing the rate contributed by OP2. Nonetheless, as the contribution of the spectrum pool from OP2 expands, its gain in average user
rate raises accordingly. Thus, the proposed spectrum sharing scheme encourages the powerless operator (i.e., OP2 in our system) to provide as much spectrum as possible. On the other hand, the proposed spectrum sharing framework prevents the powerful operator (i.e., OP1 in our system) to get exploited by the powerless operator, especially compared with the full spectrum sharing mechanism.

When one operator is allocated with the whole spectrum pool (i.e., OP1 in fair spectrum sharing and both operators in full spectrum sharing), the plot of the gain in average user rate at the other operator is consistent with the plot of the gain in frequency slots, as illustrated in Fig. 7. This indicates that the decrease in the average user rate per frequency slot due to the inter-operator interference is finite, such that the powerless operator could estimate the gain from the proposed spectrum sharing mechanism through the drop in average user rate per frequency slot and the raise in number of frequency slots; this observation helps it determine whether to participate in spectrum sharing especially when it can merely provide much fewer frequency slots than the powerful operator.

Fig. 8a and Fig. 8b show the average number of scheduled users per cell for OP1 and OP2 respectively. There is a general improvement for fair spectrum sharing and full spectrum sharing over no spectrum sharing, which illustrates the increased connectivity within the network. However, the growth in average scheduled users is less than the enhancement in average user rate. This implies that, in the no-spectrum sharing case, the proposed spectrum sharing framework mostly increases the rate of the users which are already scheduled rather than scheduling much more users within each cell. This is due to our focus on the sum-rate.

The detailed information of user scheduling with the proposed spectrum sharing framework when the number of frequency slots provided by OP2 is minimal (i.e., \(N_2 = 4\)) and maximal (i.e., \(N_2 = 16\)) are illustrated in Fig. 9 and 10, respectively, compared to the case of no spectrum sharing which is presented in Fig. 9. Within these plots, each dot and the filled circle represents a user within each cell that is not scheduled or scheduled, respectively. The colors of the filled circles reflect the number of frequency slots used to transmit signals to the users. The corresponding amount of the spectrum increases as the color changes from blue to red. In general, the system will allocate more frequency slots to users that are closer to the BSs. The users that are closer to the corners of the cells are less likely be scheduled especially when both operators contribute equally to the spectrum pool, when inter-operator interference exists within every frequency slot among the cells.

The rate changes with respect to each user for fair spectrum sharing, full spectrum sharing and spectrum sharing with same number of frequency slots compared with the case of no spectrum sharing within the simulation are illustrated in Fig. 10 and Fig. 13. In Fig. 10, OP2 contributes few frequency slots than OP1 (i.e., \(N_2 = 4\)). While in Fig. 13, OP2 provides the same number of frequency slots as OP1 (i.e., \(N_2 = 16\)). Each dot within the figures represents a user within the system. The black dots depict the users without rate changes, and the filled circles with warm colors (i.e., from red to yellow) and cold colors (i.e., blue) represent the users with rate improvement and reduction, respectively. The lighter the color and larger the circle, the greater the rate changes. For the case of spectrum sharing with same number of frequency slots, the number of users with increased rate and the number of users with decreased rate are comparable, which leads to the minor change of the overall average user rate.

In regards to fair spectrum sharing, when \(N_2 = 4\), the users of OP1 who are close to the BSs achieve most of the rate improvement. While most scheduled users of OP2 suffer from rate loss, as stated above, due to the reduce of inter-operator
interference to OP1. It is worth noting that some of the users of OP1 near the corner of its corresponding cell (i.e., closer to the BS of OP2) still get scheduled or achieve rate gain in this case. On the other hand, when both operators provide
Fig. 12. Rate changes among users ($N_2 = 4$).
Fig. 13. Rate changes among users ($N_2 = 16$).
the same amount of spectrum, the users near the corners of the cell are not scheduled, as shown in Fig. 13. Towards full spectrum sharing, most of the scheduled users of OP2 gain extra rate when $\lambda_2 = 4$. However, for OP1, only the users close to the BSs achieve a slight rate improvement in most cases.

V. CONCLUSION

The paper develops a fair spectrum sharing system between two MNOs with inter-operator coordinated beamforming in multi-cell multi-user MIMO networks. Our goal is to understand how an optimization framework and MIMO technologies help in enhancing the benefits of spectrum sharing. In this regard, we formulate an optimization problem to determine the resource allocation and user scheduling within the system. We develop an algorithm to tackle the problem through fractional programming and block coordinate descent. The simulation results illustrate that both MNOs could achieve up to 60% improvement in terms of the average user rate with the proposed fair spectrum sharing framework. Interference suppression allows for both operators to use the same frequency slots simultaneously; indeed, the superiority of spectrum sharing mainly comes from the increase in the number of frequency slots used by each operator. Importantly, our results show that both operators have an incentive to take part in the spectrum sharing mechanism proposed, when the gap between the amount of spectrum they provide is not extremely large.

In this paper, to focus on the performance upper bound, we assume perfect channel knowledge between the operators within the system, which is impractical in real-world communication networks. The next steps would be to incorporate channel estimation and sharing of channel state information to obtain a more realistic network. However, we believe our results provide an incentive to investigate spectrum sharing that exploits recent advances in optimization and resource allocation for MIMO wireless networks.

REFERENCES

[1] T. Wang and R. Adve, “Fair licensed spectrum sharing between two MNOs using resource optimization,” in Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops), Jun. 2021, pp. 1–6.
[2] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moesner, “Licensed spectrum sharing schemes for mobile operators: A survey and outlook,” IEEE Commun. Surveys Tuts., vol. 18, no. 4, pp. 2591–2623, 4th Quart., 2016.
[3] Z. Jiang and S. Mao, “Interoperator opportunistic spectrum sharing in LTE-unlicensed,” IEEE Trans. Veh. Technol., vol. 66, no. 6, pp. 5217–5228, Jun. 2017.
[4] F. Teng, D. Guo, and M. L. Honig, “Sharing of unlicensed spectrum by strategic operators,” IEEE J. Sel. Areas Commun., vol. 35, no. 3, pp. 668–679, Mar. 2017.
[5] M. D. Mueck, V. Frascolla, and B. Badic, “Licensed shared access—State-of-the-art and current challenges,” in Proc. 1st Int. Workshop Cogn. Cellular Syst. (CCS), Sep. 2014, pp. 1–5.
[6] C. Galioti, G. K. Papageorgiou, K. Voulgaris, M. M. Butt, N. Marchetti, and C. B. Papadias, “Unlocking the deployment of spectrum sharing with a policy enforcement framework,” IEEE Access, vol. 6, pp. 11793–11803, 2018.
[7] M. Asaduzzaman, R. Abozariba, and M. Patwary, “Dynamic spectrum sharing optimization and post-optimization analysis with multiple operators in cellular networks,” IEEE Trans. Wireless Commun., vol. 17, no. 3, pp. 1589–1603, Mar. 2018.
[8] R. Abozariba, M. Asaduzzaman, and M. Patwary, “Radio resource sharing framework for cooperative multipoint operator networks with dynamic overflow modeling,” IEEE Trans. Veh. Technol., vol. 66, no. 3, pp. 2433–2447, Mar. 2017.
[9] S. Dhavane and M. Z. A. Khan, “Impact of spectrum sharing rules in multipoint communication system,” in Proc. 10th Int. Conf. Commun. Syst. Netw. (COMSNETS), Jan. 2018, pp. 447–449.
[10] B. Cho, K. Koufos, R. Jiattti, and S.-L. Kim, “Co-primary spectrum sharing for inter-operator device-to-device communication,” IEEE J. Sel. Areas Commun., vol. 35, no. 1, pp. 91–105, Jan. 2017.
[11] T. Sanguanpuak, S. Gurucharrya, N. Rajatheva, M. Bennis, and M. Latva-Aho, “Multi-operator spectrum sharing for small cell networks: A matching game perspective,” IEEE Trans. Wireless Commun., vol. 16, no. 6, pp. 3761–3774, Jun. 2017.
[12] B. Singh, K. Koufos, O. Tirkkonen, and R. Berry, “Co-primary inter-operator spectrum sharing over a limited spectrum pool using repeated games,” in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2015, pp. 1494–1499.
[13] B. Singh, K. Koufos, O. Tirkkonen, and R. Jiattti, “Repeated spectrum sharing games in multi-operator heterogeneous networks,” in Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. ( DySPAN), Sep. 2015, pp. 221–228.
[14] Y. Chung and Z. Tsai, “Performance evaluation of dynamic spectrum sharing for two wireless communication networks,” IET Commun., vol. 4, no. 4, pp. 452–462, 2010.
[15] J. Leon, R. D. Ford, V. V. Ratanam, J. Cho, and J. Zhang, “Coordinated dynamic spectrum sharing for 5G and beyond cellular networks,” IEEE Access, vol. 7, pp. 111592–111604, 2019.
[16] G. P. Koukoudakis and P. Soldati, “Joint network density and spectrum sharing in multi-operator colocated ultra-dense networks,” in Proc. 7th Int. Conf. Modern Circuits Syst. Technol. (MOCAST), May 2018, pp. 1–4.
[17] R. Lijins, H. Zhang, I. Noppen, L. Yu, E. Karipidis, and K. Borner, “System-level assessment of non-orthogonal spectrum sharing via transmit beamforming,” in Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring), Jun. 2013, pp. 1–6.
[18] J. Lindblom and E. G. Larsson, “Does non-orthogonal spectrum sharing in the same cell improve the sum-rate of wireless operators?” in Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Oct. 2012, pp. 6–10.
[19] R. K. Saha, “Nationwide spectrum sharing of mobile network operators with indoor small cells,” in Proc. DySPAN, 2019, pp. 1–2.
[20] A. K. Diniss and J. S. Thompson, “The effects of including wraparound when simulating cellular wireless systems with relaying,” in Proc. IEEE 65th Veh. Technol. Conf., Oct. 2007, pp. 914–918.
[21] A. Goldsmith, Wireless Communications. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[22] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, Jr., and B. L. Evans, “Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization,” IEEE Trans. Signal Process., vol. 54, no. 9, pp. 3658–3663, Sep. 2006.
[23] A. A. Khan, R. Adve, and W. Yu, “Optimizing downlink resource allocation in multiuser MIMO networks via fractional programming and the Hungarian algorithm,” IEEE Trans. Wireless Commun., vol. 19, no. 8, pp. 5162–5175, Aug. 2020.
[24] K. Shen and W. Yu, “Fractional programming for communication systems—Part II: Uplink scheduling via matching,” IEEE Trans. Signal Process., vol. 66, no. 10, pp. 2631–2644, May 2018.
[25] R. J. Wilson, Introduction to Graph Theory. Hoboken, NJ, USA: Wiley, 1986.
[26] H. W. Kuhn, “The Hungarian method for the assignment problem,” Naval Res. Logistics, vol. 52, pp. 7–21, Feb. 2005. [Online]. Available: https://ideas.repec.org/a/wly/navres/v52y2005i1p7-21.html.
[27] D. P. Bertsekas, “The auction algorithm: A distributed relaxation method for the assignment problem,” Ann. Oper. Res., vol. 14, no. 1, pp. 105–123, Dec. 1988.
Tianchen Wang (Associate Member, IEEE) was born in Yangzhou, China. He received the B.A.Sc. and M.A.Sc. degrees in electrical and computer engineering from the University of Toronto, ON, Canada, in 2019 and 2021, respectively.

Raviraj Adve (Fellow, IEEE) was born in Bombay, India. He received the B.Tech. degree in electrical engineering from IIT, Bombay, in 1990, and the Ph.D. degree from Syracuse University in 1996. From 1997 to August 2000, he worked at Research Associates for Defense Conversion Inc., on contract with the Air Force Research Laboratory, Rome, NY, USA. In August 2000, he joined the Faculty of the University of Toronto, where he is currently a Professor. His research interests include analysis and design techniques for cooperative and heterogeneous networks, energy harvesting networks, and signal processing techniques for radar and sonar systems. His thesis received the Syracuse University Outstanding Dissertation Award. He has received the 2009 Fred Nathanson Young Radar Engineer of the Year Award.