Research Article

Multimode Transmission in Network MIMO Downlink with Incomplete CSI

Nima Seifi,[1] Mats Viberg (EURASIP Member),[1] Robert W. Heath Jr.,[2] Jun Zhang,[3] and Mikael Coldrey[4]

[1] *Department of Signals and Systems, Chalmers University of Technology, 412 96 Gothenburg, Sweden*

[2] *Wireless Networking and Communications Group (WNCG), Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-0240, USA*

[3] *Deptartment of Electronic and Computer Engineering, Hong Kong University of Science and Technology, (HKUST), Clear Water Bay, Kowloon, Hong Kong*

[4] *Ericsson Research, Ericsson AB, 417 56 Gothenburg, Sweden*

Correspondence should be addressed to Nima Seifi, nima.seifi@chalmers.se

We consider a cooperative multicell MIMO (a.k.a *network MIMO*) downlink system with multiantenna base stations (BSs), which are connected to a central unit and communicate with multiantenna users. In such a network, obtaining perfect channel state information (CSI) of all users at the central unit to exploit opportunistic scheduling requires a substantial amount of feedback and backhaul signaling. We propose a scheduling algorithm based only on the knowledge of the average SNR at each user from all the cooperating BSs, denoted as *incomplete CSI*. Multimode transmission is applied that is able to adaptively adjust the number of data streams transmitted to each user. Utilizing the results of random matrix theory, an analytical framework is proposed to approximate the ergodic rate of each user with different number of data streams. Using these ergodic rates, a joint user and mode selection algorithm is proposed, where only the scheduled users need to feed back instantaneous CSI. Simulation results demonstrate that the developed analytical framework provides a good approximation for a practical number of antennas. While substantially reducing the feedback overhead, it is shown that the proposed scheduling algorithm performs closely to the opportunistic scheduling algorithm that requires instantaneous CSI feedback from all users.

1. Introduction

Recently, cooperative multicell transmission (also called *network MIMO*) has been proposed as an efficient way to suppress the intercell interference and increase the downlink capacity of cellular systems [1–3]. In one way of realizing a network MIMO system, multiple base stations (BSs) are connected to a central unit via backhaul links. The central unit coordinates BSs and performs joint scheduling and signal processing operations. Assuming no limitations with regards to the capacity, error, and delay in the backhaul, and upon the availability of perfect channel state information (CSI) of all users at the central unit, the network MIMO system in the downlink is equivalent to a MIMO broadcast channel with per-BS power constraint (PBPC) [4–7].

Many of the previous studies on network MIMO assume full coordination over the whole system, which is not practical (if not impossible). First, the backhaul links connecting the BSs and the central unit are subject to transmission error [8] and delay [9]. They also have limited capacity which confines the amount of data and CSI sharing [10–12]. Second, connecting a large number of BSs for joint processing is of high complexity, which has motivated the development of coordination strategies at a local scale [13, 14]. Third, obtaining perfect CSI from all the users at the central unit, which is indispensable to achieve the full diversity or multiplexing gains, results in a substantial training and feedback overhead [15–17].

In this paper, we consider local BS cooperation and focus on the third limitation, that is, the substantial overhead to

obtain the CSI for each active user at the central unit. To this end, we assume the backhaul links to be perfect and leave the study of the effect of imperfect backhaul to future work. We propose a framework that enables scheduling based only on the knowledge of the average received SNR at each user from all the cooperating BSs, denoted as *incomplete CSI*. This reduces the overhead both on the feedback channel and the backhaul, since only the selected users (usually a small number) have to feedback their instantaneous CSI for precoder design.

1.1. Related Work. The scheduling problem in the single-cell multiuser MIMO downlink has been widely investigated under various precoding and beamforming strategies [18–22]. The total number of transmitted data streams in such systems is upper bounded by the number of BS antennas under a linear precoding framework. Therefore, if the total number of receive antennas in the system is greater than the total number of transmit antennas, the scheduling will consist of selecting both users and the number of data streams or modes (note that the term "mode" used in this paper denotes the number of data streams for a given user rather than the number of active users, as in [22] or different MIMO transmission techniques, such as spatial multiplexing/diversity mode [23]) to each user. Such multimode transmission improves performance by allowing a dynamic allocation of the transmission resources among the users [24]. User and mode selection in a network MIMO system is more challenging than in its single-cell counterpart. The increased number of users and BSs in the network MIMO makes the CSI requirement daunting.

Acquiring CSI at the central unit is one of the limiting factors for a practical network MIMO system. The availability of perfect CSI, however, has a cardinal role in exploiting the spatial degrees of freedom in such systems [13]. In practice, CSI for the downlink is obtained through some form of training and feedback. In time-division duplexing (TDD) systems, the CSI is obtained at each BS using the channel reciprocity (see, e.g., [25]). In frequency-division duplexing (FDD) systems, since uplink and downlink take place in widely separated frequency bands, the downlink CSI is fed back via some explicit feedback channels (see e.g., [26]). This places a significant burden on the uplink feedback channel. The feedback overhead increases with the number of BSs, users, antennas, and subcarriers and can easily occupy the whole uplink resources. Furthermore, in both TDD and FDD systems, the CSI should be forwarded from the BSs to the central unit which limits the backhaul resources for data transmission.

The tradeoff between the resources dedicated to CSI overhead and data transmission in the backhaul has been recently studied in [15–17], where several multicell system architectures were compared. It was further shown that the downlink performance of network MIMO systems is mainly limited by the inevitable acquisition of CSI rather than by limited backhaul capacity.

As a solution to this limitation, some authors [27, 28] have proposed strategies based on local CSI at the BSs and statistical CSI at the central unit, whereas others [29, 30] consider to serve only certain subsets of users with multiple BSs. In [31], a decentralized cooperation framework has been proposed in which all the necessary processing is performed in a truly distributed manner among the BSs without the need of any CSI exchange with the central unit. Several BS cooperation strategies have been studied which consider the combination of limited-capacity backhaul and imperfect CSI [32, 33].

1.2. Contributions. In this paper, we develop a scheduling algorithm for a network MIMO system with multiantenna users. To reduce the feedback overhead, we adopt a two-step scheduling process: the first step is joint user and mode selection, and the second step is the feedback and precoder design which only involves the users selected in the first step. The two-step multimode transmission strategy was also proposed in the single-cell MIMO downlink in [22], with single-antenna users and imperfect CSI at the BS. The main contributions are as follows.

Ergodic Rate Analysis. We propose an analytical framework to compute an accurate approximation for the ergodic rate of each user with different number of data streams, based only on the knowledge of incomplete CSI for any given location. Essentially, the aggregate channel from multiple distributed cooperating BSs can be well approximated as coming from a single super BS. This enables an efficient method to evaluate the performance of network MIMO systems without the need for extensive and computationally intensive Monte-Carlo simulations.

Joint User and Mode Selection Algorithm. We use the derived ergodic user rates as a metric to perform user and mode selection, which is suitable for the data application without stringent delay constraint. Since the ergodic rate for each user is obtained only based on incomplete CSI, the small-scale fading is not exploited in the proposed strategy. Therefore, it does not provide small-scale multiuser diversity gain, but instead, it omits the need for feedback of instantaneous CSI from a large number of users for scheduling by more than 93%. It is also shown that the performance of the proposed user and mode-selection strategy is very close to the opportunistic scheduling based on instantaneous CSI feedback from all the users.

1.3. Organization. The rest of the paper is organized as follows: the system model and transmission strategy are described in Section 2. In Section 3, some mathematical preliminaries, which are useful throughout the paper, are presented. An analytical framework to derive an approximation for the ergodic rate of each user at different modes is proposed in Section 4. A greedy joint user and mode selection algorithm based on the derived ergodic rates of each user is described in Section 5. The performance of the proposed user and mode-selection algorithm is evaluated in Section 6. Finally, Section 7 concludes the paper and discusses the future work.

1.4. Notation. Scalars are denoted by lowercase letters, vectors denoted by boldface lowercase letters, and matrices denoted by boldface uppercase letters. $(\cdot)^{\top}$, $(\cdot)^{*}$, $\det(\cdot)$, $\log(\cdot)$, and $\mathrm{tr}(\cdot)$ denote transpose, complex conjugate transpose, determinant, 2-base logarithm, and trace of a matrix, respectively; $\mathbb{U}(m, n)$ is the collection of $m \times n$ unitary matrices with unit-norm orthogonal columns. $\mathbb{E}[\cdot]$ is the statistical expectation. $[\mathbf{\Phi}]_{:(m:n)}$ denotes the matrix obtained by choosing $n - m + 1$ columns from $\mathbf{\Phi}$ starting from the mth column. $[\mathbf{\Phi}]_{(m,n)}$ denotes the $m \times n$ upper-left corner of a square matrix $\mathbf{\Phi}$. $\lambda_i(\mathbf{\Phi})$ and $\lambda_{\min}(\mathbf{\Phi})$ denote the ith ordered and the smallest eigenvalue of $\mathbf{\Phi}\mathbf{\Phi}^{*}$, respectively. $\|\mathbf{x}\|$ denotes the Euclidean norm of a complex vector $\mathbf{x}$, and $|\mathcal{S}|$ is the cardinality of a set $\mathcal{S}$; $\dim(\cdot)$ is the dimensionality operator. Further, $\lfloor \cdot \rfloor$ denotes the floor operation, $\mathbf{\Phi} \otimes \mathbf{\Psi}$ denotes the Kronecker product of the two matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$, and C_m^n denotes the combination of n choosing m.

2. System Model

2.1. Network MIMO Structure. The network MIMO system considered in this paper comprises B cells, each of which has a BS with N_t antennas and K_b users, each equipped with N_r antennas, for $K_b \geq 1$ and $b = 1, 2, \ldots, B$. The total number of *active* users in the system is denoted as $K = \sum_{b=1}^{B} K_b$. Users in different locations of the cellular coverage are subject to distance-dependent pathloss and shadowing. A narrowband frequency-flat fading channel is considered. We consider the downlink transmission. The following are the key assumptions made in this paper.

Assumption 1. All the B cooperating BSs are interconnected via a central unit with the use of backhaul links with infinite capacity such that they can fully share CSI and user data.

With this assumption, all the cooperating BSs form a distributed antenna array that can perform joint scheduling and transmission.

Assumption 2. The number of antennas at each BS is greater than that of each user, that is, $N_t \geq N_r$.

Due to space constraints, user terminals can only have a small number of antennas, which makes $N_t \geq N_r$ a reasonable assumption. Therefore, each user can have at most N_r data streams.

Assumption 3. For the scheduling phase, the central unit relies on the knowledge of incomplete CSI of all users (scheduling CSI). For the transmission phase, however, the central unit has perfect knowledge of the singular values and the corresponding right singular vectors of the selected eigenmodes of each selected user (transmission CSI) for precoding design.

This assumption of scheduling CSI significantly reduces the feedback and backhaul signaling overhead for scheduling. The transmission CSI assumption is due to the transmission strategy employed in this paper (see Section 2.3),

which reduces the transmission CSI required for precoding design with respect to the strategies which requires the complete channel matrix of each selected user. The transmission CSI can be reduced even more using limited feedback techniques [34–36], which we will not explore in this paper.

2.2. Received Signal Model. The aggregate channel matrix of user k from all the B cooperating BSs can be written as

$$\mathbf{H}_k = \left[\sqrt{\rho_{k,1}} \mathbf{H}_{k,1} \quad \cdots \quad \sqrt{\rho_{k,B}} \mathbf{H}_{k,B} \right], \qquad (1)$$

where $\mathbf{H}_{k,b} \in \mathbb{C}^{N_r \times N_t}$ represents the small-scale fading channel matrix and $\rho_{k,b}$ is the large-scale fading channel coefficient that captures the distance-dependent pathloss including shadowing for user k from the bth BS.

We denote the $N_t \times 1$ transmit signal vector from the bth BS as $\mathbf{x}_b$. Therefore, the $BN_t \times 1$ aggregate transmit signal vector from all the B cooperating BSs can be written as

$$\mathbf{x} = \left[\mathbf{x}_1^{\top} \quad \cdots \quad \mathbf{x}_B^{\top} \right]^{\top}. \qquad (2)$$

The discrete-time complex baseband signal received by the kth user is given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k, \qquad (3)$$

where $\mathbf{n}_k$ is the noise vector at the kth user, with entries that are independent and identically distributed complex Gaussian with zero mean and unit variance, denoted as i.i.d $\mathcal{CN}(0, 1)$.

2.3. Transmission Strategy. To simultaneously transmit multiple spatially multiplexed streams to multiple users, we adopt a linear precoding strategy called multiuser eigenmode transmission (MET) (the framework developed in this paper, however, can be used with any other linear precoding in which the precoding matrix for each user is dependent only on the other users' channels) [37]. The MET approach enables the number of data streams for each user to be adaptively selected and at the same time avoids the complexity of joint iterative precoder/equalizer design [38]. Denote $\mathcal{K}$ as the set of served users at a given time interval and assign indices $k = 1, \ldots, |\mathcal{K}|$. Denote $\mathcal{L}_k$ as the set of eigenmodes selected for transmission to user k, which are indexed from 1 to ℓ_k, where $\ell_k = |\mathcal{L}_k|$. Under a linear precoding framework, the total number of data streams in the downlink, denoted as the *system transmission mode* (STM), is upper bounded by the number of transmit antennas and can be written as

$$L = \sum_{j=1}^{|\mathcal{K}|} \ell_j, \qquad (4)$$

where $1 \leq L \leq BN_t$. The aggregate transmitted signal is given by

$$\mathbf{x} = \sum_{k=1}^{|\mathcal{K}|} \mathbf{T}_k \mathbf{d}_k, \qquad (5)$$

where $\mathbf{T}_k \in \mathbb{C}^{BN_t \times \ell_k}$ is the precoding matrix and $\mathbf{d}_k$ denotes the ℓ_k dimensional signal vector for user k. It is assumed that each user k at a given time slot is able to perfectly estimate its channel matrix $\mathbf{H}_k$ without any error. Furthermore, each user k performs a singular value decomposition (SVD) on its channel as $\mathbf{H}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k$. We denote the ith singular value and the corresponding left and right singular vectors of $\mathbf{H}_k$ as $\sigma_{k,i}$, $\mathbf{u}_{k,i}$, and $\mathbf{v}_{k,i}$, respectively. We also assume that the singular values in $\mathbf{\Sigma}_k$ are arranged in the descending order; that is, $\sigma_{k,1} \geq \sigma_{k,2} \geq \cdots \geq \sigma_{k,N_r}$. The kth user's receiver is a linear equalizer given by $[\mathbf{U}_k]^*_{:(1:\ell_k)}$. Using (3) and (5), the postprocessed signal $\mathbf{r}_k$ after applying the linear equalizer is given by

$$\mathbf{r}_k = \mathbf{F}_k \mathbf{T}_k \mathbf{d}_k + \mathbf{F}_k \sum_{j=1, j \neq k}^{|\mathcal{K}|} \mathbf{T}_j \mathbf{d}_j + \mathbf{w}_k, \tag{6}$$

where $\mathbf{F}_k = [\mathbf{U}_k]^*_{:(1:\ell_k)} \mathbf{H}_k$ and $\mathbf{w}_k$ is the processed noise, which is still white since the equalizer is a unitary matrix. In the case of perfect knowledge of $\mathbf{F}_1, \ldots, \mathbf{F}_{|\mathcal{K}|}$, denote the aggregate interference matrix as $\tilde{\mathbf{H}}_k = [\mathbf{F}_1^* \cdots \mathbf{F}_{k-1}^* \mathbf{F}_{k+1}^* \cdots \mathbf{F}_{|\mathcal{K}|}^*]^*$. To suppress the interuser interference, the constraint $\mathbf{F}_j \mathbf{T}_k = \mathbf{0}$ must be satisfied for $k \neq j$ [39]. This requires that $\mathbf{T}_k$ lies in the null space of $\tilde{\mathbf{H}}_k$. With this constraint satisfied, the second term on the right hand side of the equality in (6) becomes zero. Denote the total number of interfering data streams for user k from the other $(|\mathcal{K}| - 1)$ selected users as $\tilde{\ell}_k = \sum_{j=1, j \neq k}^{|\mathcal{K}|} \ell_j$. As a result, there are only $BN_t - \tilde{\ell}_k$ spatial degrees of freedom available at the transmitter side to support spatial multiplexing for user k, and therefore, $\ell_k \leq \min\{N_r, BN_t - \tilde{\ell}_k\}$. In [40], it was shown that the precoding matrix $\mathbf{T}_k$ can be written as a cascade of two precoding matrices $\mathbf{B}_k$ and $\mathbf{D}_k$, that is, $\mathbf{T}_k = \mathbf{B}_k \mathbf{D}_k$, where the $BN_t \times (BN_t - \tilde{\ell}_k)$ matrix $\mathbf{B}_k$ removes the interuser interference. Denote the SVD of $\tilde{\mathbf{H}}_k$ as

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k \tilde{\mathbf{\Sigma}}_k \begin{bmatrix} \tilde{\mathbf{V}}_k^{(1)} & \tilde{\mathbf{V}}_k^{(0)} \end{bmatrix}^*, \tag{7}$$

where $\tilde{\mathbf{V}}_k^{(0)}$ corresponds to the right singular vectors of $\tilde{\mathbf{H}}_k$ associated with the null modes. One natural choice is $\mathbf{B}_k = \tilde{\mathbf{V}}_k^{(0)}$. As a matter of fact, $\mathbf{F}_k \mathbf{B}_k$ is the effective interuser interference-free channel for user k. The $(BN_t - \tilde{\ell}_k) \times \ell_k$ matrix $\mathbf{D}_k$ is used for parallelization. Denote the SVD of the effective channel for user k as

$$\mathbf{F}_k \mathbf{B}_k = \overline{\mathbf{U}}_k \overline{\mathbf{\Sigma}}_k \begin{bmatrix} \overline{\mathbf{V}}_k^{(1)} & \overline{\mathbf{V}}_k^{(0)} \end{bmatrix}^*, \tag{8}$$

where $\overline{\mathbf{V}}_k^{(1)}$ denotes the right singular vectors of $\mathbf{F}_k \mathbf{B}_k$ corresponding to the first ℓ_k nonzero singular values. The optimum choice of $\mathbf{D}_k$ is then $\mathbf{D}_k = \overline{\mathbf{V}}_k^{(1)}$ [41].

3. Mathematical Preliminaries

In this section, we present some mathematical preliminaries from matrix variate distributions and random matrix theory which prove useful in the analysis to follow. For more detailed discussions, the readers are referred to [42–44].

Definition 1. Let $\mathbf{Z}$ denote a $q \times p$ complex matrix with $q \leq p$ and a common covariance matrix $\mathbf{C} = \mathbb{E}\{\mathbf{z}_j \mathbf{z}_j^*\}$ for all j, where $\mathbf{z}_j$ is the jth column vector of $\mathbf{Z}$. The elements of two columns $\mathbf{z}_i$ and $\mathbf{z}_j$ are assumed to be mutually independent. If the elements of $\mathbf{Z}$ are identically distributed as $\mathcal{CN}(0, 1)$ such that $\mathbb{E}\{\mathbf{Z}\} = \mathbf{0}$, then the Hermitian matrix $\mathbf{ZZ}^*$ is a central Wishart matrix with p degrees of freedom and covariance matrix $\mathbf{C}$, denoted as $\mathbf{ZZ}^* \sim \mathcal{CW}_q(p, \mathbf{C})$.

3.1. Approximation of a Linear Combination of Wishart Matrices. Let $\mathbf{Y}_s \sim \mathcal{CW}_q(p_s, \mathbf{C}_s)$ for $s = 1, 2, \ldots, S$ be mutually independent central Wishart matrices. Consider a linear combination

$$\mathbf{Y} = \sum_{s=1}^{S} \alpha_s \mathbf{Y}_s, \quad \alpha_s > 0. \tag{9}$$

The distribution of $\mathbf{Y}$ can be approximated by the distribution of another Wishart matrix as $\hat{\mathbf{Y}} \sim \mathcal{CW}_q(\hat{p}, \hat{\mathbf{C}})$ [42, page 124], where $\hat{p}$ is the equivalent degrees of freedom given by

$$\hat{p} = \left[\frac{\det\left(\sum_{s=1}^{S} \alpha_s p_s \mathbf{C}_s\right)^{(2q)}}{\det\left(\sum_{s=1}^{S} \alpha_s^2 p_s (\mathbf{C}_s \otimes \mathbf{C}_s)\right)} \right]^{1/q^2}, \tag{10}$$

and $\hat{\mathbf{C}}$ denotes the equivalent covariance matrix written as

$$\hat{\mathbf{C}} = \frac{1}{\hat{p}} \sum_{s=1}^{S} \alpha_s p_s \mathbf{C}_s. \tag{11}$$

Now, if $p_1 = \cdots = p_S = p$ and $\mathbf{C}_1 = \cdots = \mathbf{C}_S = \mathbf{C}$, then (10) can be rewritten as

$$\hat{p} = \left[\frac{\left(p \sum_{s=1}^{S} \alpha_s\right)^{2q^2} \det(\mathbf{C})^{(2q)}}{\left(p \sum_{s=1}^{S} \alpha_s^2\right)^{q^2} \det(\mathbf{C} \otimes \mathbf{C})} \right]^{1/q^2}. \tag{12}$$

Using the determinant property of the Kronecker product [42, Chapter 3], that is, $\det(\mathbf{C} \otimes \mathbf{C}) = \det(\mathbf{C})^{2q}$, in (12), we can obtain

$$\hat{p} = p \left[\frac{\left(\sum_{s=1}^{S} \alpha_s\right)^2}{\sum_{s=1}^{S} \alpha_s^2} \right]. \tag{13}$$

By substituting (13) in (11), it then holds that

$$\hat{\mathbf{C}} = \left(\frac{\sum_{s=1}^{S} \alpha_s^2}{\sum_{s=1}^{S} \alpha_s} \right) \mathbf{C}. \tag{14}$$

Finally, recall from Definition 1 that the condition $q \leq \hat{p}$ should hold for the Wishart distribution $\mathcal{CW}_q(\hat{p}, \hat{\mathbf{C}})$ to be meaningful. In the following theorem, the upper and lower bound for $\hat{p}$ are obtained.

Theorem 1. Assume that $\alpha_s \geq 0$ for $s = 1, \ldots, S$ and also that at least one of the α_s's is nonzero. If $\hat{p}$ is defined in (13), then $p \leq \hat{p} \leq Sp$. Furthermore, the upper bound equality happens when $\alpha_1 = \alpha_2 = \cdots = \alpha_S$, while the lower bound equality holds when $\exists! s : \alpha_s > 0$ ($\exists!$ means there exists one and only one).

Proof. See Appendix A.   □

3.2. Truncation of Random Unitary Matrices and Jacobi Ensemble

Definition 2. If $\mathbf{X} \sim \mathcal{CW}_m(n_1, \mathbf{C})$ and $\mathbf{Y} \sim \mathcal{CW}_m(n_2, \mathbf{C})$ are independent complex Wishart matrices, then $\mathbf{J} = \mathbf{X}(\mathbf{X} + \mathbf{Y})^{-1}$ is called a complex Jacobi matrix.

It is shown in [45, Proposition 4.1] that $\mathbf{J}$ has the same distribution as that of $[\mathbf{U}]_{(q,p)}[\mathbf{U}]_{(q,p)}^*$, where $\mathbf{U} \in \mathbb{U}(n,n)$ with $q = m$, $p = n_1$, and $n = n_1 + n_2$. Therefore, the eigenvalues of $\mathbf{J}$ are the same as those of $[\mathbf{U}]_{(q,p)}[\mathbf{U}]_{(q,p)}^*$. The distribution of the extreme eigenvalues of the complex Jacobi ensemble is derived in [44].

4. Ergodic Rate Analysis

In this section, we derive an approximation for the ergodic rate of each user k at different modes ℓ_k. To assist the analysis, we assume that the elements of $\mathbf{H}_{k,b}$ are distributed such that $\mathbf{H}_{k,b}\mathbf{H}_{k,b}^* \sim \mathcal{CW}_{N_r}(N_t, \mathbf{C})$ for $b = 1, \ldots, B$. Let the precoding matrix for user k be written as

$$\mathbf{T}_k = \begin{bmatrix} \mathbf{T}_{k,1}^\top & \cdots & \mathbf{T}_{k,B}^\top \end{bmatrix}^\top, \quad (15)$$

where $\mathbf{T}_{k,b}$ denote the precoding applied at the bth BS for user k, such that the transmitted signal from the bth BS can be written as $\mathbf{x}_b = \sum_{k=1}^{|\mathcal{K}|} \mathbf{T}_{k,b} d_k$. Assuming MET and the practical per-BS power constraint (PBPC) with STM equal to L, the ergodic rate of a user k with ℓ_k data streams using (6) can be expressed as

$$\overline{R}_k(\ell_k, L) = \mathbb{E}\left[\max_{\mathbf{Q}_k} \log \det\left(\mathbf{I}_{\ell_k} + \mathbf{F}_k \mathbf{T}_k \mathbf{Q}_k \mathbf{T}_k^* \mathbf{F}_k^* \right) \right], \quad (16)$$

subject to

$$\text{tr}\left(\mathbb{E}\left[\mathbf{T}_{k,b} \mathbf{Q}_k \mathbf{T}_{k,b}^* \right] \right) \le P, \quad \text{for } b = 1, \ldots, B, \quad (17)$$

where $\mathbf{Q}_k = \mathbb{E}[\mathbf{d}_k \mathbf{d}_k^*]$ is the power allocation matrix for user k and P is the power constraint at each BS. Since the total power constraint (TPC) over all the BSs is less restrictive, the performance under TPC is equal or better than that under PBPC. It has also been shown that there is only a marginal rate loss of PBPC to TPC [13]. Therefore, for simplicity and analytical tractability, we assume TPC and equal power allocation among all the L data streams in the downlink, that is, $\mathbf{Q}_k = (BP/L)\mathbf{I}_{\ell_k}$. The ergodic rate in (16), can be written as [37]

$$\hat{R}_k(\ell_k, L) = \mathbb{E}\left[\log \det\left(\mathbf{I}_{\ell_k} + \frac{BP}{L} \overline{\mathbf{U}}_k [\overline{\boldsymbol{\Sigma}}_k]_{:(1:\ell_k)} [\overline{\boldsymbol{\Sigma}}_k]_{:(1:\ell_k)}^* \overline{\mathbf{U}}_k^* \right) \right]$$

$$\overset{(a)}{=} \mathbb{E}\left[\sum_{i=1}^{\ell_k} \log\left(1 + \frac{BP}{L} \lambda_i(\mathbf{F}_k \mathbf{B}_k) \right) \right], \quad (18)$$

where (a) follows using the matrix identity $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$. Therefore, the ergodic rate of a user k at mode ℓ_k depends on the distributions of $\lambda_i(\mathbf{F}_k \mathbf{B}_k)$ for all i. In order to compute the distribution of $\lambda_i(\mathbf{F}_k \mathbf{B}_k)$ in the network MIMO case, we provide the following result.

Lemma 1. *Assume $\widetilde{\ell}_k = L - \ell_k$, where L is the STM and ℓ_k is the mode of user k. Furthermore, assume $\mathbf{B}_k \in \mathbb{U}(BN_t, (BN_t - \widetilde{\ell}_k))$ is the matrix that projects the channel of user k onto the null space of other users and is independent of $\mathbf{F}_k$. Assume $\mathbf{F}_k$ and $\mathbf{B}_k$ have SVDs given by $\mathbf{F}_k = \mathbf{U}_{\mathbf{F}_k}\boldsymbol{\Sigma}_{\mathbf{F}_k}\mathbf{V}_{\mathbf{F}_k}^*$ and $\mathbf{B}_k = \mathbf{U}_{\mathbf{B}_k}\boldsymbol{\Sigma}_{\mathbf{B}_k}\mathbf{V}_{\mathbf{B}_k}^*$, respectively. It then holds that*

$$\lambda_i(\mathbf{F}_k \mathbf{B}_k) \ge \lambda_i(\mathbf{F}_k) \lambda_{\min}\left(\left[\mathbf{V}_{\mathbf{F}_k}^* \mathbf{U}_{\mathbf{B}_k} \right]_{(i, (BN_t - \widetilde{\ell}_k))} \right), \quad \forall i. \quad (19)$$

Proof. See Appendix B. □

We denote $\lambda_{\min}([\mathbf{V}_{\mathbf{F}_k}^* \mathbf{U}_{\mathbf{B}_k}]_{(i, (BN_t - \widetilde{\ell}_k))})$ with $\lambda_{\min}$ hereafter in the paper for the ease of notation. Denote the joint probability density function (pdf) of $\lambda_i(\mathbf{F}_k)$ and $\lambda_{\min}$ as $f_{(\lambda_i(\mathbf{F}_k), \lambda_{\min})}(\lambda, \lambda')$, and let $f_{(\lambda_i(\mathbf{H}_k))}(\lambda)$ and $f_{(\lambda_{\min})}(\lambda')$ denote the marginal pdf of $\lambda_i(\mathbf{F}_k)$ and $\lambda_{\min}$, respectively. Using the result of Lemma 1 in (18) and the approximation $\log(1 + x) \approx \log(x)$, we can get an approximation for the ergodic rate for user k as

$$\hat{R}_k(\ell_k, L) \approx \mathbb{E}\left[\sum_{i=1}^{\ell_k} \log\left(\frac{BP}{L} \lambda_i(\mathbf{F}_k)\lambda_{\min} \right) \right]$$

$$= \sum_{i=1}^{\ell_k} \int_0^1 \int_0^\infty \log\left(\frac{BP}{L} \lambda \lambda' \right) f_{(\lambda_i(\mathbf{F}_k), \lambda_{\min})}(\lambda, \lambda') d\lambda \, d\lambda'$$

$$= \sum_{i=1}^{\ell_k} \int_0^1 \int_0^\infty \log\left(\frac{BP}{L} \right) f_{(\lambda_i(\mathbf{F}_k), \lambda_{\min})}(\lambda, \lambda') d\lambda \, d\lambda'$$

$$+ \sum_{i=1}^{\ell_k} \int_0^1 \int_0^\infty \log(\lambda) f_{(\lambda_i(\mathbf{F}_k), \lambda_{\min})}(\lambda, \lambda') d\lambda \, d\lambda'$$

$$+ \sum_{i=1}^{\ell_k} \int_0^1 \int_0^\infty \log(\lambda') f_{(\lambda_i(\mathbf{F}_k), \lambda_{\min})}(\lambda, \lambda') d\lambda \, d\lambda'$$

$$\overset{(a)}{=} \ell_k \log\left(\frac{BP}{L} \right) + \sum_{i=1}^{\ell_k} \int_0^\infty \log(\lambda) f_{(\lambda_i(\mathbf{H}_k))}(\lambda) d\lambda$$

$$+ \sum_{i=1}^{\ell_k} \int_0^1 \log(\lambda') f_{(\lambda_{\min})}(\lambda') d\lambda', \quad (20)$$

where (a) follows from the fact that $\lambda_i(\mathbf{F}_k) = \lambda_i(\mathbf{H}_k)$ for all i, which results from (6).

The elements of the aggregate channel matrix $\mathbf{H}_k$ in (1) for any realization of the kth user location do not follow an i.i.d complex Gaussian distribution in general, due to the different large-scale channel coefficients to different BSs. Assuming $\mathbf{H}_{k,m}$ and $\mathbf{H}_{k,n}$ are mutually independent for all $m \ne n$, $\mathbf{H}_k \mathbf{H}_k^* = \sum_{b=1}^B \rho_{k,b} \mathbf{H}_{k,b} \mathbf{H}_{k,b}^*$ is a linear combination of central Wishart matrices, and according to the results in Section 3, its distribution can be approximated with that of another Wishart matrix $\hat{\mathbf{H}}_k \hat{\mathbf{H}}_k^* \sim \mathcal{W}_{N_r}(\hat{N}_{t,k}, \hat{\rho}_k \mathbf{C})$, where $\hat{N}_{t,k}$ and $\hat{\rho}_k$ are obtained using (13) and (14) as (since

the dimensions of $\hat{\mathbf{H}}_k$ must be integers, we use $\lfloor x + 0.5 \rfloor$ to round x to the nearest integer)

$$\hat{N}_{t,k} = p \left\lfloor \frac{\left(\sum_{b=1}^{B} \rho_{k,b} \right)^2}{\sum_{b=1}^{B} \rho_{k,b}^2} + 0.5 \right\rfloor, \qquad (21)$$

and

$$\hat{\rho}_k = \left(\frac{\sum_{b=1}^{B} \rho_{k,b}^2}{\sum_{b=1}^{B} \rho_{k,b}} \right). \qquad (22)$$

Remark 1. We note that $\hat{N}_{t,k}$ is a function of $\rho_{k,b}$ for $b = 1, \ldots, B$, which depends on the position of the user k. Therefore, for user k at any given position in the cell, the $N_r \times \hat{N}_{t,k}$ i.i.d channel matrix $\hat{\mathbf{H}}_k \sim \mathcal{CN}(0, \hat{\rho}_k \mathbf{C})$ can be interpreted as if the user is communicating with one super BS with $\hat{N}_{t,k}$ transmit antennas and the equivalent large-scale channel coefficient $\hat{\rho}_k$. Furthermore, according to Theorem 1, the maximum of $\hat{N}_{t,k}$ is BN_t, which corresponds to positions where $\rho_{k,1} = \cdots = \rho_{k,B}$. At other positions, however, where user k experiences larger $\rho_{k,b}$ values from some of the BSs and smaller from the others, $\hat{N}_{t,k}$ will be smaller than BN_t. It can be concluded that $\hat{N}_{t,k}$ is determined mainly by those BSs to which the user has largest $\rho_{k,b}$ values, and those are the ones that help the cooperation and are actually *seen* by the user.

Since the distribution of $\mathbf{H}_k \mathbf{H}_k^*$ is approximated with the distribution of another Wishart matrix $\hat{\mathbf{H}}_k \hat{\mathbf{H}}_k^*$, we have $f_{\lambda_i(\mathbf{H}_k)}(\lambda) \approx f_{\lambda_i(\hat{\mathbf{H}}_k)}(\lambda)$. The distribution $f_{\lambda_i(\hat{\mathbf{H}}_k)}(\lambda)$ for $i = 1, \ldots, N_r$ for the uncorrelated central case is given in [46] as (the general framework developed in this paper, however, is applicable to arbitrarily correlated channels. We only express the result for the uncorrelated case for simplicity)

$$f_{\lambda_i(\hat{\mathbf{H}}_k)}^{\text{uc}}(\lambda) = G_{\text{uc}} \sum_{n=1}^{N_r} \sum_{m=1}^{N_r} (-1)^{n+m} \lambda^{n+m-2+\hat{N}_{t,k}-N_r} e^{-\lambda} |\mathbf{\Omega}^{\text{uc}}|, \qquad (23)$$

$$\text{for } i = 1, \ldots, N_r,$$

where G_{uc} is given by

$$G_{\text{uc}} = \left[\prod_{i=1}^{N_r} \left(\hat{N}_{t,k} - i \right)! \prod_{j=1}^{N_r} (N_r - j)! \right]^{-1}. \qquad (24)$$

The (i, j)th element of $\mathbf{\Omega}$ is written as

$$\omega_{i,j} = \begin{cases} \gamma\left(\alpha_{i,j}^{(n)(m)} + \hat{N}_{t,k} - N_r + 1, \lambda \right) & \text{for } i = 1, \\ \Gamma\left(\alpha_{i,j}^{(n)(m)} + \hat{N}_{t,k} - N_r + 1, \lambda \right) & \text{for } i = N_r, \\ \left(\alpha_{i,j}^{(n)(m)} + \hat{N}_{t,k} - N_r \right)! \zeta_{N_r,1} & \text{otherwise}, \end{cases} \qquad (25)$$

where $\Gamma(a, b) \triangleq \int_b^\infty x^{a-1} e^{-x} dx$ and $\gamma(a, b) \triangleq \int_0^b x^{a-1} e^{-x} dx$ denote the upper and lower incomplete Gamma functions, respectively, $\alpha_{i,j}^{(n)(m)}$ is given by

$$\alpha_{i,j}^{(n)(m)} = \begin{cases} i + j - 2 & \text{if } i < n \text{ and } j < m, \\ i + j & \text{if } i \geq n \text{ and } j \geq m, \\ i + j - 1 & \text{otherwise}, \end{cases} \qquad (26)$$

and

$$\zeta_{a,b} = \prod_{j=0}^{b-1} (a - j)^{-1/a-b}. \qquad (27)$$

In order to find $f_{\lambda_{\min}}(\lambda')$, we note that the multiplication of two unitary matrices is another unitary matrix, that is, $\mathbf{V}_{\mathbf{F}_k}^* \mathbf{U}_{\mathbf{B}_k} \in \mathbb{U}(BN_t, BN_t)$ [47]. Therefore, $[\mathbf{V}_{\mathbf{F}_k}^* \mathbf{U}_{\mathbf{B}_k}]_{(i,(BN_t - \tilde{\ell}_k))}$ is a truncated unitary matrix. As mentioned in Section 3.2, for any Wishart distributed unitary matrix with Haar measure $\mathbf{A} \in \mathbb{U}(n, n)$, the multiplication $[\mathbf{A}]_{(q,p)} [\mathbf{A}]_{(q,p)}^*$ for $q \leq p \leq n$, has the same distribution as a complex Jacobi ensemble [45, Proposition 4.1]. The distribution of the minimum eigenvalues of the complex Jacobi ensemble is obtained in [44, Equation 3.2]

$$f_{\lambda_{\min}}(\lambda') = \frac{\Gamma_i(BN_t) \Gamma_i(i)}{\Gamma_i\left(i + \tilde{\ell}_k \right) \Gamma_i\left(BN_t - \tilde{\ell}_k \right)}$$

$$\times i\tilde{\ell}_k \lambda'^{(BN_t - \tilde{\ell}_k - i - 2)} (1 - \lambda')^{(i\tilde{\ell}_k - 1)}$$

$$\times {}_2F_1^{(1)}\left(\tilde{\ell}_k - 1, i - BN_t + \tilde{\ell}_k + 2; \tilde{\ell}_k + i; (1 - \lambda') \mathbf{I}_{i-1} \right), \qquad (28)$$

where $\Gamma_m(c) = \pi^{m(m-1)/2} \prod_{j=1}^{m} \Gamma(c - j + 1)$ denotes the multivariate Gamma function, $\Gamma(a) \triangleq \int_0^\infty x^{a-1} e^{-x} dx$ is the Gamma function, and ${}_2F_1^{(1)}(\tilde{\ell}_k, i - BN_t + \tilde{\ell}_k; \tilde{\ell}_k + i; (1 - \lambda') \mathbf{I}_i)$ is a hypergeometric function of a matrix argument [44, 48]. Based on (23) and (28), we can evaluate (20) numerically.

To verify the accuracy of the approximation in (20), we consider a hexagonal cellular layout with cell sectoring. By using 120-degree sectoring in each cell, every 3 neighboring cells can coordinate with each other to serve users in the shadow area shown in Figure 1. The number of transmit antennas is chosen to be $N_t = 4$, which is the value currently implemented in wireless standards such as 3GPP LTE, and $N_r = 2$. We randomly place two users in each cell sector. The pathloss model is based on scenario C2 of the WINNER II specifications [49]. The large-scale fading is modeled as lognormal with standard deviation of 8 dB. The edge SNR is defined to be the received SNR at the edge of the cell, assuming that one BS transmits at full power while all other BSs are off, accounting for pathloss but ignoring shadowing and small-scale fading.

Figure 2 depicts the ergodic rate of a sample user k for $\ell_k = 1, 2$ and different values of L (for the clarity of the plot we consider only some values of L. The performance for the other values can be concluded easily). The results correspond
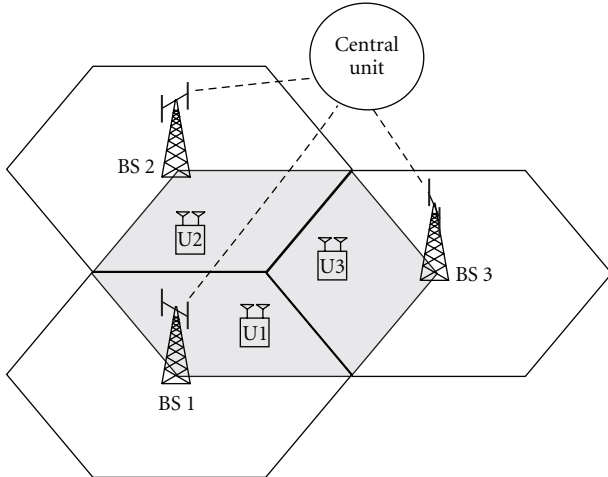
FIGURE 1: A network MIMO system with 3 BSs, connected to a central unit via backhaul links.
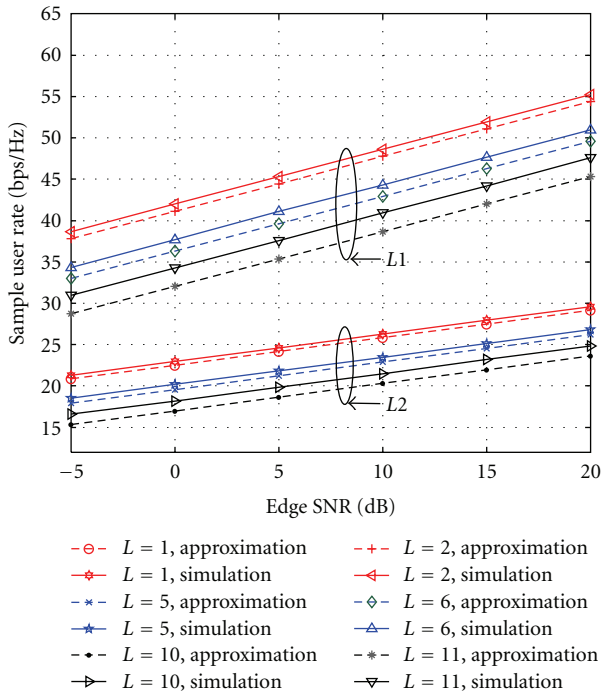

FIGURE 2: Ergodic rate of a sample user k versus edge SNR for a given realization of user locations, different values of L, and $\ell_k = 1, 2$. The dashed lines indicate the rate obtained from the lower bound approximation, that is, $\hat{R}_k(\ell_k, L)$, while the solid lines represent the rate obtained from the simulation, that is, $\overline{R}_k(\ell_k, L)$.

to one random snapshot of user locations when for any given L and ℓ_k other users are assigned with 1 or 2 data streams. It is also assumed that no user is within a normalized distance of 0.2 from its closest BS for the pathloss model to be valid. It is observed that the lower bound approximation in (20) is very close to the simulation results obtained by Monte-Carlo simulations using (16) over the full range of edge SNR. The difference between the approximation and the achieved rate
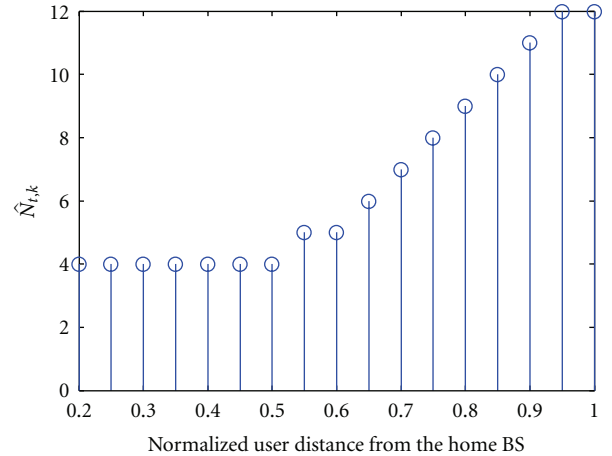

FIGURE 3: Equivalent number of transmit antenna $\hat{N}_{t,k}$ versus the normalized distance from the home BS for a sample user k moving along the line that connects the BS 3 to the center of the shaded hexagon in Figure 1.

is small enough to consider the approximation good enough for scheduling as explained in the next section.

To justify the argument in Remark 1, in Figure 3, we plot $\hat{N}_{t,k}$ versus the normalized distance from the home BS for a sample user k moving along the line that connects the BS 3 to the center of the shaded hexagon in Figure 1. It is observed that within a normalized distance of 0.5 from the BS 3, $\hat{N}_{t,k} = 4$ which results from the fact that $\rho_{k,3}$ is much larger than $\rho_{k,1}$ and $\rho_{k,2}$, and therefore, only the BS 3 is seen by this user. As the user moves toward the center of the shaded hexagon, $\rho_{k,1}$ and $\rho_{k,2}$ increase but $\rho_{k,3}$ decreases, resulting an increase in $\hat{N}_{t,k}$. Indeed at the center of the shaded hexagon $\hat{N}_{t,k} = 12$, which means all the 3 BSs are seen by the user and actually can be helpful in the cooperation. Therefore, BS cooperation is not very helpful for the cell interior user, and only edge users get most of the benefit. This can be used to design BS cooperation.

5. Downlink Scheduling: Joint User and Mode Selection

In this section, downlink scheduling for multimode transmission is discussed. The total number of streams in the system under study is upper bounded by BN_t, and normally $KN_r \gg BN_t$, so at each scheduling phase a subset of users, and the preferred mode of each user must be selected for transmission. In multimode transmission, the number of data streams for each user is adaptively selected, which allows to efficiently exploit the available degrees of freedom in the channel using *multimode diversity* (multimode diversity is a form of selection diversity among users with multiple antennas, which enables the scheduler to perform selection not only among the users (multiuser diversity), but also among the different eigenmodes of each user) [24].

In a system with heterogenous users, the goal of the downlink scheduling is to make the system operate at a rate

point of its ergodic achievable rate region such that a suitable concave and increasing *network utility function* $g(\cdot)$ of the user individual ergodic rates is maximized [50].

Let $\mathcal{M} = \{1, 2, \ldots, KN_r\}$ denote the set consisting of all possible modes for all users, and let $\mathcal{S}_i$ be a subset of $\mathcal{M}$ with $|\mathcal{S}_i| \leq BN_t$. Let $\mathcal{K}_i$ denote the set of users with at least one selected mode in $\mathcal{S}_i$. The downlink scheduling problem we wish to solve is defined as

$$\mathcal{S} = \max_{\mathcal{S}_i} g\left(\left\{\overline{R}_k(\ell_k, |\mathcal{S}_i|)\right\}_{k \in \mathcal{K}_i}\right),$$

$$\text{subject to} \quad \sum_{k=1}^{|\mathcal{K}_i|} \ell_k = |\mathcal{S}_i|, \quad 1 \leq \ell_k \leq N_r. \tag{29}$$

To solve (29) through brute-force exhaustive search over $\mathcal{S}_i$, $\sum_{m=1}^{BN_t} C_m^{KN_r}$ combinations must be checked. Furthermore, for each combination $\mathcal{S}_i$, the knowledge of $\{\overline{R}_k(\ell_k, |\mathcal{S}_i|)\}_{k \in \mathcal{K}_i}$ is required, which is not easy to compute in general. This is a computationally complex problem if $KN_r \gg BN_t$ and very difficult to implement.

5.1. Low-Complexity User and Mode Selection. To reduce the computational complexity and at the same time exploit the benefits of multimode transmission, we propose a low-complexity joint user and mode selection algorithm. To simplify the computation of $\{\overline{R}_k(\ell_k, |\mathcal{S}_i|)\}_{k \in \mathcal{K}_i}$ for any given $\mathcal{S}_i$, we propose to use the approximations obtained in (20) for $\overline{R}_k(\ell_k, |\mathcal{S}_i|)$ instead of the exact fomula in (16). This enables the analytical computation of $g(\{\widehat{R}_k(\ell_k, |\mathcal{S}_i|)\}_{k \in \mathcal{K}_i})$ for any given $\mathcal{S}_i$, with only the knowledge of the average SNR of each users to all the BSs, and avoids the complexity of precoding matrix computations. Therefore, it not only reduces the computational complexity at the central unit but also removes the need for any instantaneous CSI feedback at the expense of sacrificing the small-scale fading multiuser diversity.

One way to reduce the complexity associated with exhaustive search is to treat (29) as a relaxed optimization problem, that is, to greedily select data streams which maximize the network utility function. Toward this goal, we gradually increase L from 1 to BN_t. For any given L, the approximate ergodic rate for the next unselected eigenmode of all users is computed using (20). The algorithm continues until either $L = BN_t$ or the network utility function starts to decrease. Once the user and mode selection is done, only selected users need to feedback the singular values and the corresponding singular vectors of their selected eigenmodes for precoder design. Therefore, the proposed scheduling algorithm is of low complexity and is suitable for application when delay is not a stringent constraint or when the feedback resources is limited. The resulting algorithm is summarized in Algorithm 1.

5.2. Network Utility Function. We focus on two special cases of network utility function, namely, the *ergodic sum rate* and the *sum log ergodic rate*. To perform maximum sum rate

scheduling (MSRS) for a given $\mathcal{S}_i$, the per-cell ergodic sum rate utility function is defined as

$$g_{\text{MSRS}}\left(\left\{\widehat{R}_k(\ell_k, |\mathcal{S}_i|)\right\}_{k \in \mathcal{K}_i}\right) = \frac{1}{B} \sum_{k=1}^{|\mathcal{K}_i|} \widehat{R}_k(\ell_k, |\mathcal{S}_i|). \tag{30}$$

To introduce fairness by performing proportional fairness scheduling (PFS) for a given $\mathcal{S}_i$, the per-cell sum log ergodic rate utility function is defined as [50]

$$g_{\text{PFS}}\left(\left\{\widehat{R}_k(\ell_k, |\mathcal{S}_i|)\right\}_{k \in \mathcal{K}_i}\right) = \frac{1}{B} \sum_{k=1}^{|\mathcal{K}_i|} \log\left(\widehat{R}_k(\ell_k, |\mathcal{S}_i|)\right). \tag{31}$$

6. Simulation Results and Key Observations

In this section, the performance of the proposed user and mode-selection strategy is evaluated via Monte-Carlo simulations. The assumptions for the cellular layout, pathloss, shadowing, and the number of antennas are given in Section 4. We drop $K_1 = K_2 = K_3 = 10$ users randomly according to a uniform distribution in each cell.

Inspired from [49], we follow a drop-based simulation. In this approach, a drop corresponds to one realization for user locations, during which the large-scale fading parameters as well as velocity and direction of travel for users, are practically constant. Therefore, each user can only undergo small-scale fading at each location. Furthermore, large-scale fading parameters are realized independently from drop to drop. This method does not take into account the time evolution of the channel. The main advantage of it is the simplicity of the simulation. We run 1000 drops for user locations. At the beginning of each drop, all the users feedback their average SNR from all the cooperating BSs (in real systems, such update is not frequent and only occurs when users move around) and the set $\mathcal{S}$ is obtained using the Algorithm 1. For each obtained $\mathcal{S}$ at each drop, 1000 realizations are simulated with independent small-scale channel states.

6.1. Sum Rates for Different Systems. Figure 4, compares the ergodic sum rate of the proposed strategy with both MSRS and PFS to that of single-user transmission (SUT) and opportunistic scheduling based on instantaneous CSI (OSICSI). In SUT, only one user with the best ergodic rate is selected and served at each scheduling interval. For the detailed information about the OSICSI algorithm, see [37]. It is shown in Figure 4 that the approximate sum rate is quite close to the achieved one with MSRS. It can also be observed that the achieved sum rate for PFS is very close to that of MSRS. It is further shown that the proposed algorithm for both MSRS and PFS performs much better than SUT and achieves a large fraction of the sum rate of OSICSI over a practical range of edge SNR values. For example, at an edge SNR of 10 dB, it achieves 80% and 68% of the sum rate of opportunistic scheduling with MSRS and PFS, respectively. In Figure 5, the sum log ergodic rate versus edge SNR is plotted. It is observed that the approximate and simulated curves are in good agreement.
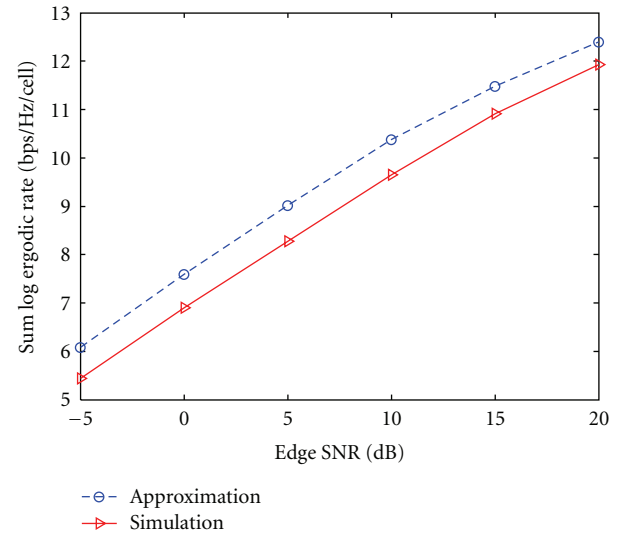
```
(1) Initialization: L = 1, 𝒮 = ∅, g(∅) = 0
(2) while L ≤ min(KN_r, BN_t) do
(3)     r(L) = 0, v = 0, 𝒮_L = ∅, 𝒦_L = ∅, R̂_k(i, L) = 0, ℓ_k = 0, u_k = 0 for k = 1, …, K and i = 1, …, N_r
(4)     while v ≤ L do
(5)         for k = 1 to K do
(6)             Compute s_k ← g({R̂_j(ℓ_j, L)}_{j∈𝒦_{L-1}, j≠k} ∪ {R̂_k(ℓ_k + 1, L)}) from (20)
(7)         end for
(8)         k max ← arg max_k s_k, ℓ_{k max} ← ℓ_{k max} + 1
(9)         r(L) ← s(k max), u_{k max} ← 1, v ← v + 1
(10)    end while
(11)    for k = 1 to K do
(12)        if u_k ≠ 0 then
(13)            𝒮_L = 𝒮_L ∪ {ℓ_k} and 𝒦_L = 𝒦_L ∪ {k}
(14)        end if
(15)    end for
(16)    if r(L) < r(L − 1) then
(17)        𝒮 = 𝒮_{L-1}, break
(18)    end if
(19)    L ← L + 1
(20) end while
```

ALGORITHM 1: Pseudocode for the proposed algorithm.
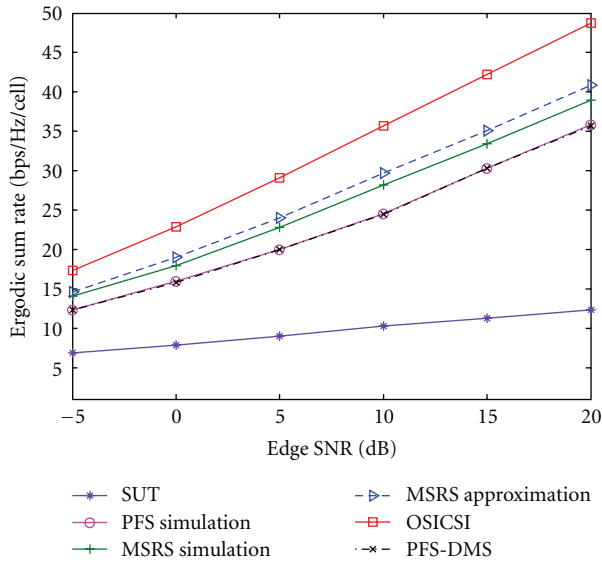
FIGURE 4: Comparison of average sum rate versus edge SNR for $B = 3$, $K = 30$, $N_t = 4$, and $N_r = 2$.

FIGURE 5: Comparison of sum log ergodic rate versus edge SNR for $B = 3$, $K = 30$, $N_t = 4$, and $N_r = 2$.

To compare the performance of MSRS and PFS, we have divided the distance between users and their home BS into bins of width 0.1 of the cell radius, and for each bin, the fraction of times over all the 1000 drops that a user in that bin has been scheduled is plotted in Figure 6. It is observed that with PFS, the activity fraction of users at farther distances from the BS is higher as compared to that using MSRS. On the other hand, the simulation results show that the PFS algorithm almost always chooses 1 data streams for each served user, while in MSRS 2 data streams are also selected about 50% of the time that 1 stream is selected.

Therefore, in PFS, choosing the users without considering the problem of the mode selection (dominant-mode transmission) seems to be the relevant strategy, at least in the chosen scenario. The sum rate of PFS with only dominant selection (PFS-DMS) is also plotted in Figure 4. It can be seen that the sum rate of PFS-DMS matches very well with that of PFS with user and mode selection.

6.2. Feedback Analysis. In this section, we compare the amount of feedback required by the proposed algorithm with
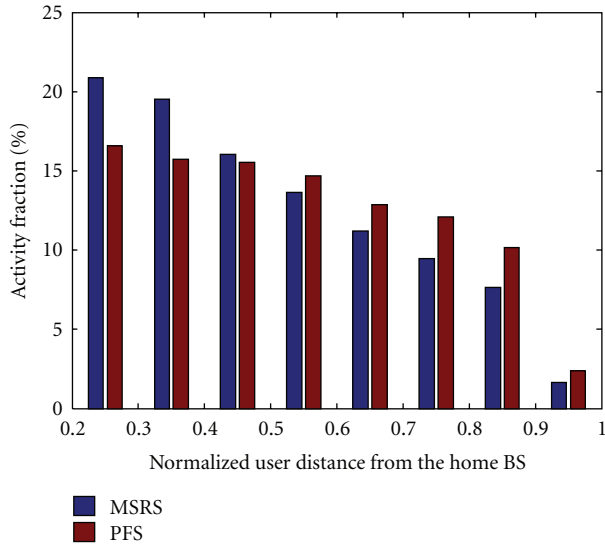
FIGURE 6: Comparison of activity fraction in percentage for both MSRS and PFS for edge SNR of 10 dB.
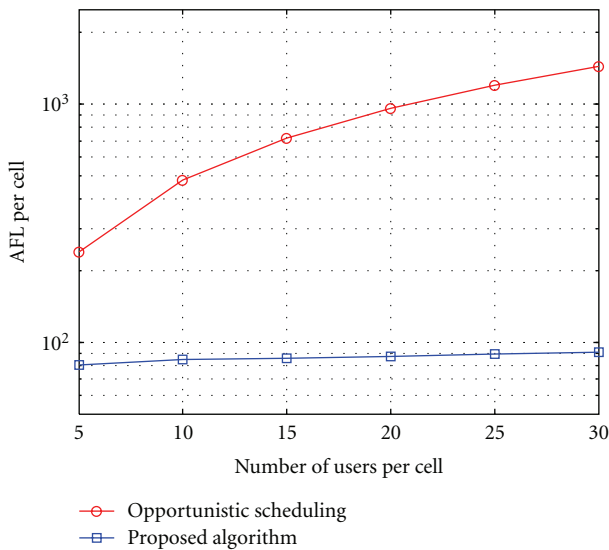
BN_t, corresponding to their selected eigenmodes at each small-scale realization for precoding design. If the number of realizations of small-scale channel states within each drop is assumed to be T (1000 in this paper), the AFL is given by

$$F_{\text{prop}} = \left(\frac{BK}{T}\right) + \mathbb{E}_\mathcal{S}\left[\sum_{k=1}^{|\mathcal{K}|} \ell_k(1 + 2BN_t)\right]. \qquad (32)$$

Since set $\mathcal{S}$ changes from drop to drop, the AFL is obtained by averaging over many drops. For the opportunistic scheduling based on instantaneous CSI, however, all K users should feed back their respective $N_r \times BN_t$ complex channel matrix in all the T realizations in each drop. Therefore, we have

$$F_{\text{opp}} = 2BKN_tN_r. \qquad (33)$$

In Figure 7, the AFL per cell versus the number of users per cell is compared for the proposed scheduling algorithm and the opportunistic scheduling based on instantaneous CSI. It is observed that for the proposed algorithm, the AFL increases very slowly with the number of users, since no matter how big the number of users is in the cell, the number of served users will be limited by the maximum number of transmit antennas. The only overhead of increasing the number of users is the average SNR values that should be fed back at the beginning of each drop, which is negligible with respect to the amount fed back during a drop. For the opportunistic scheduling, however, the AFL increases linearly with the number of users since it has to feed back the CSI for all the users at each realization. It is observed that for $K_1 = K_2 = K_3 = 30$, AFL is decreased by more than 93% which makes the proposed algorithm attractive in such kind of scenarios.

FIGURE 7: Comparison of AFL per cell for the proposed algorithm and the opportunistic scheduling based on instantaneous CSI with edge SNR of 10 dB.

7. Conclusion

In this paper, we propose an analytical framework to approximate ergodic rates of users with different modes in a network MIMO system, based only on the knowledge of received average SNR from all the cooperating BSs at each user, called incomplete CSI in this paper. Based on the derived approximate ergodic rates, the problem of downlink scheduling with both MSRS and PFS is addressed. The proposed scheduling algorithm significantly reduces the feedback amount and performs close to the opportunistic scheduling based on instantaneous CSI. It is of particular interest for applications where there is a total feedback overhead constraint and/or when there is no stringent delay constraint. It is also shown that by introducing fairness, the probability of selecting higher modes for each users decreases significantly, which results in dominant mode transmission (beamforming) to each user.

that by opportunistic scheduling based on instantaneous CSI. We define the average feedback load (AFL) as the average number of real coefficients that are fed back to the central unit during each drop normalized to the number of small-scale realizations within that drop. For the proposed algorithm (the results are almost the same for both MSRS and PFS since the AFL depends on the total number of transmitted data stream, that is, STM, which is the same on average for both schemes. We only plot the result for MSRS here), once the set $\mathcal{S}$ is obtained at the beginning of a drop, each selected user $k \in \mathcal{K}$ sends back ℓ_k real-valued singular values and ℓ_k complex-valued right singular vectors of size

Appendices

A. Proof of Theorem 1

In order to prove the lower bound for $\hat{p}$, we write

$$
\begin{aligned}
\hat{p} &= p\left[\frac{\left(\sum_{s=1}^{S}\alpha_s\right)^2}{\sum_{s=1}^{S}\alpha_s^2}\right] \\
&= p\left[\frac{\sum_{s=1}^{S}\alpha_s^2 + \sum_{i=1}^{S}\sum_{j=1, j\neq i}^{S}\alpha_i\alpha_j}{\sum_{s=1}^{S}\alpha_s^2}\right] \\
&\overset{(a)}{\geq} p\left[\frac{\sum_{s=1}^{S}\alpha_s^2}{\sum_{s=1}^{S}\alpha_s^2}\right] = p,
\end{aligned}
\tag{A.1}
$$

where (a) follows from the fact that $\alpha_s > 0$ for all s. The equality for the lower bound holds only when $\sum_{i=1}^{S}\sum_{j=1, j\neq i}^{S}\alpha_i\alpha_j = 0$, and this happens only when $\exists! s : \alpha_s > 0$. The proof for the upper bound is established using the Cauchy-Schwarz inequality. For some $\beta_s > 0$ for all s, we have

$$
\left(\sum_{s=1}^{S}\alpha_s\beta_s\right)^2 \leq \left(\sum_{s=1}^{S}\alpha_s^2\right)\left(\sum_{s=1}^{S}\beta_s^2\right).
\tag{A.2}
$$

Now, setting $\beta_1 = \cdots = \beta_S = 1$, (A.2) can be rewritten as

$$
\left(\sum_{s=1}^{S}\alpha_s\right)^2 \leq S\left(\sum_{s=1}^{S}\alpha_s^2\right).
\tag{A.3}
$$

The upper bound for $\hat{p}$ using (A.3) is now obtained as

$$
\hat{p} \leq p\left[\frac{S\left(\sum_{s=1}^{S}\alpha_s^2\right)}{\sum_{s=1}^{S}\alpha_s^2}\right] = Sp.
\tag{A.4}
$$

The equality condition for the Cauchy-Schwarz inequality is only satisfied when the vectors $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_S]$ and $\boldsymbol{\beta} = [\beta_1 \cdots \beta_S]$ are linearly dependent; that is, $\boldsymbol{\alpha} = x\boldsymbol{\beta}$ for some scalar x. Since in our case $\boldsymbol{\beta} = [1 \cdots 1]$, it results that the equality holds only when $\alpha_1 = \cdots = \alpha_S$. This completes the proof of the Theorem.

B. Proof of Lemma 1

First, notice that for values of $i > \ell_k$ or $i > BN_t - \tilde{\ell}_k$, $\lambda_i(\mathbf{F}_k)$, or $\lambda_i(\mathbf{B}_k)$ are defined as zero. Hence, the argument of this lemma is obvious for these cases. Now, we prove the argument for $i \leq \ell_k$. For notational convenience, we define $m = BN_t - \tilde{\ell}_k$. According to "max-min" half of the Courant-Fischer Theorem [47], for any subspace $\mathcal{M} \subset \mathbb{C}^{BN_t - \tilde{\ell}_k}$ with dimention i, we have

$$
\lambda_i(\mathbf{F}_k\mathbf{B}_k) = \max_{\mathcal{M}, \dim(\mathcal{M})=i} \min_{\substack{\mathbf{x} \in \mathcal{M} \\ \|\mathbf{x}\| = 1}} \|\mathbf{F}_k\mathbf{B}_k\mathbf{x}\|^2.
\tag{B.5}
$$

Now, let us define $\mathcal{V} = \langle \mathbf{v}_1(\mathbf{B}_k), \mathbf{v}_2(\mathbf{B}_k), \ldots, \mathbf{v}_i(\mathbf{B}_k)\rangle$, where $\mathbf{v}_j(\mathbf{B}_k)$ denotes the jth column of $\mathbf{V}_{\mathbf{B}_k}$ and $\langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\rangle$ is the span of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. We have

$$
\begin{aligned}
\lambda_i(\mathbf{F}_k\mathbf{B}_k) &\geq \min_{\substack{\mathbf{x} \in \mathcal{V} \\ \|\mathbf{x}\| = 1}} \|\mathbf{F}_k\mathbf{B}_k\mathbf{x}\|^2 \\
&\overset{(a)}{=} \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \left\|\mathbf{F}_k\mathbf{U}_{\mathbf{B}_k}\boldsymbol{\Sigma}_{\mathbf{B}_k}\mathbf{x}'\right\|^2 \\
&\overset{(b)}{=} \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \left\|\mathbf{F}_k\mathbf{U}_{\mathbf{B}_k}\mathbf{I}_{BN_t \times m}\mathbf{x}'\right\|^2 \\
&= \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \left\|\mathbf{F}_k[\mathbf{U}_{\mathbf{B}_k}]_{:(1:m)}\mathbf{x}'\right\|^2 \\
&\overset{(c)}{=} \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \left\|\boldsymbol{\Sigma}_{\mathbf{F}_k}\mathbf{V}_{\mathbf{F}_k}^*[\mathbf{U}_{\mathbf{B}_k}]_{:(1:m)}\mathbf{x}'\right\|^2 \\
&\overset{(d)}{\geq} \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \left\|[\boldsymbol{\Sigma}_{\mathbf{F}_k}]_{:(1,i)}[\mathbf{V}_{\mathbf{F}_k}]_{:(1:i)}^*[\mathbf{U}_{\mathbf{B}_k}]_{:(1:m)}\mathbf{x}'\right\|^2 \\
&\overset{(e)}{\geq} \min_{\substack{\mathbf{x}' \in \mathcal{V} \\ \|\mathbf{x}'\| = 1}} \lambda_i(\mathbf{F}_k)\left\|\left[\mathbf{V}_{\mathbf{F}_k}^*\mathbf{U}_{\mathbf{B}_k}\right]_{(i,m)}\mathbf{x}'\right\|^2 \\
&\overset{(f)}{=} \lambda_i(\mathbf{F}_k)\lambda_{\min}\left(\left[\mathbf{V}_{\mathbf{F}_k}^*\mathbf{U}_{\mathbf{B}_k}\right]_{(i,m)}\right).
\end{aligned}
\tag{B.6}
$$

Here, (a) follows from the fact that (i) $\mathbf{x} \in \mathcal{V}$ is equivalent to $\mathbf{x}' \in \mathcal{V}$ for some $\mathbf{x}' = \mathbf{V}_{\mathbf{B}_k}^*\mathbf{x}$; and (ii) $\|\mathbf{x}\|^2 = \|\mathbf{x}'\|^2$. (b) follows from the fact that the singular values of a unitary matrix are all equal to 1. (c) follows from the fact that $\|\mathbf{U}_{\mathbf{F}_k}\mathbf{y}'\| = \|\mathbf{y}'\|$, where $\mathbf{y}' = \boldsymbol{\Sigma}_{\mathbf{F}_k}\mathbf{V}_{\mathbf{F}_k}^*[\mathbf{U}_{\mathbf{B}_k}]_{:(1:m)}\mathbf{x}'$. (d) follows from the fact that defining $\mathbf{z}' = [\mathbf{U}_{\mathbf{B}_k}]_{:(1:m)}\mathbf{x}'$, we have

$$
\begin{aligned}
\left\|\boldsymbol{\Sigma}_{\mathbf{F}_k}\mathbf{V}_{\mathbf{F}_k}^*\mathbf{z}'\right\|^2 &= \left\|[\boldsymbol{\Sigma}_{\mathbf{F}_k}]_{:(1:i)}[\mathbf{V}_{\mathbf{F}_k}]_{:(1:i)}^*\mathbf{z}'\right\|^2 \\
&+ \left\|[\boldsymbol{\Sigma}_{\mathbf{F}_k}]_{:(i+1:BN_t)}^{1/2}[\mathbf{V}_{\mathbf{F}_k}]_{:(i+1:BN_t)}^*\mathbf{z}'\right\|^2.
\end{aligned}
\tag{B.7}
$$

(e) follows from the fact that with $\mathbf{q}' = [\mathbf{V}_{\mathbf{F}_k}^*\mathbf{U}_{\mathbf{B}_k}]_{(i,m)}\mathbf{x}'$, we have $\|[\boldsymbol{\Sigma}_{\mathbf{F}_k}]_{:(1:i)}\mathbf{q}'\|^2 \geq \lambda_i(\mathbf{F}_k)\|\mathbf{q}'\|^2$. Finally, (f) follows from the Rayleigh-Ritz Theorem [47]. Equation (B.6) completes the proof of the Lemma.

Acknowledgments

and in part by SSF within the Strategic Research Center Charmant.

References

[1] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: overcoming intercell interference in indoor wireless systems," in *Proceedings of the 41th Asilomar Conference on Signals, Systems and Computers*, pp. 83–87, Pacific Grove, Calif, USA, November 2007.

[2] S. Shamai, O. Somekh, and B. M. Zaidel, "Multi-cell communications: an information theoretic perspective," in *Proceedings of the Joint Workshop on Communications and Coding*, Florence, Italy, October 2004.

[3] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, Article ID 5089979, pp. 2983–2989, 2009.

[4] F. Boccardi and H. Huang, "Zero-forcing precoding for the MIMO broadcast channel under per-antenna power constraints," in *Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '06)*, pp. 1–5, Cannes, France, July 2006.

[5] K. Karakayali, R. Yates, G. Foschini, and R. Valenzuela, "Optimum zero-forcing beamforming with per-antenna power constraints," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 101–105, Nice, France, June 2007.

[6] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, 2007.

[7] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1435–1445, 2010.

[8] J. Hoydis, M. Kobayashi, and M. Debbah, "On the optimal number of cooperative base stations in network MIMO," *The IEEE Transactions on Signal Processing*, submitted.

[9] T. Tamaki, K. Seong, and J. M. Cioffi, "Downlink MIMO systems using cooperation among base stations in a slow fading channel," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 4728–4733, Glasgow, Scotland, June 2007.

[10] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Downlink multicell processing with limited-backhaul capacity," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 840814, pp. 1–10, 2009.

[11] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, "Uplink macro diversity of limited backhaul cellular network," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3457–3478, 2009.

[12] R. Zakhour and D. Gesbert, "Optimized data sharing in multicell MIMO with finite backhaul capacity," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '10)*, Austin, Tex, USA, June 2010.

[13] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, "Networked MIMO with clustered linear precoding," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, Article ID 4907459, pp. 1910–1921, 2009.

[14] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 190–204, 2009.

[15] S. A. Ramprashad and G. Caire, "Cellular vs. network MIMO: a comparison including the channel state information overhead," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '09)*, pp. 878–884, Tokyo, Japan, September 2009.

[16] S. A. Ramprashad, G. Caire, and H. C. Papadopoulos, "Cellular and network MIMO architectures: MU-MIMO spectral efficiency and costs of channel state information," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, pp. 1811–1818, Pacific Grove, Calif, USA, November 2009.

[17] G. Caire, S. A. Ramprashad, and H. C. Papadopoulos, "Rethinking network MIMO: cost of CSIT, performance analysis, and architecture comparisons," in *Proceedings of the Information Theory and Applications Workshop (ITA '10)*, pp. 386–395, San Diego, Calif, USA, May 2010.

[18] G. Dimić and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3857–3868, 2005.

[19] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[20] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 506–522, 2005.

[21] V. K. N. Lau, "Proportional fair space-time scheduling for wireless communications," *IEEE Transactions on Communications*, vol. 53, no. 8, p. 747, 2005.

[22] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. Heath Jr., "Multi-mode transmission for the MIMO broadcast channel with imperfect channel state information," *IEEE Transactions on Communications*, submitted.

[23] D. J. Love and R. W. Heath Jr., "Multimode precoding for MIMO wireless systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3674–3687, 2005.

[24] R. Chen, Z. Shen, J. G. Andrews, and R. W. Heath Jr., "Multimode transmission for multiuser MIMO systems with block diagonalization," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3294–3302, 2008.

[25] J. Jose, A. E. Ashikhmin, P. Whiting, and S. Vishwanath, "Linear precoding for multi-user multiple antenna TDD systems," *IEEE Transactions on Communications*, submitted.

[26] H. Shirani-Mehr and G. Caire, "Channel state feedback schemes for multiuser MIMO-OFDM downlink," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2713–2723, 2009.

[27] M. Kobayashi, S. Yang, M. Debbah, and J. Belfiore, "Outage efficient strategies for network MIMO with partial CSIT," *The IEEE Transactions on Signal Processing*, submitted.

[28] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4298–4310, 2010.

[29] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. E. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, Article ID 586878, pp. 1–19, 2008.

[30] P. Marsch and G. Fettweis, "On multicell cooperative transmission in backhaul-constrained cellular systems," *Annales*

des Telecommunications/Annals of Telecommunications, vol. 63, no. 5-6, pp. 253–269, 2008.

[31] A. Papadogiannis, E. Hardouin, and D. Gesbert, "Decentralising multicell cooperative processing: a novel robust framework," *EUEASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 890685, pp. 1–10, 2009.

[32] P. Marsch and G. Fettweis, "On downlink network MIMO under a constrained backhaul and imperfect channel knowledge," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1–16, Honolulu, Hawaii, USA, December 2009.

[33] P. Marsch and G. Fettweis, "Uplink CoMP under a constrained backhaul and imperfect channel knowledge," *IEEE Transactions on Wireless Communications*, submitted.

[34] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.

[35] K. Huang, R. W. Heath Jr., and J. G. Andrews, "Space division multiple access with a sum feedback rate constraint," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3879–3891, 2007.

[36] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, Article ID 4641957, pp. 1473–1482, 2008.

[37] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 3, pp. 17–20, Honolulu, Hawaii, USA, 2007.

[38] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1969–1973, 2004.

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

[40] S. Shim, J. S. Kwak, R. W. Heath Jr., and J. G. Andrews, "Block diagonalization for multi-user MIMO with other-cell interference," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, Article ID 4570233, pp. 2671–2681, 2008.

[41] R. Chen, R. W. Heath Jr., and J. G. Andrews, "Transmit selection diversity for unitary precoded multiuser spatial multiplexing systems with linear receivers," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 1159–1171, 2007.

[42] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*, Chapman & Hall/CRC, New York, NY, USA, 2000.

[43] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[44] I. Dumitriu and P. Koev, "Distributions of the extreme eigenvalues of beta-Jacobi random matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 1, pp. 1–6, 2008.

[45] P. J. Forrester, "Quantum conductance problems and the Jacobi ensemble," *Journal of Physics A*, vol. 39, no. 22, pp. 6861–6870, 2006.

[46] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of wishart matrices," *IEEE Transactions on Communications*, vol. 57, no. 4, pp. 1050–1060, 2009.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1990.

[48] P. Koev and A. Edelman, "The efficient evaluation of the hypergeometric function of a matrix argument," *Mathematics of Computation*, vol. 75, no. 254, pp. 833–846, 2006.

[49] IST-WINNER D1.1.2 P. Kyösti et al., "WINNER II channel models," ver. 1.1, September 2007, www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf.

[50] H. Shirani-Mehr, G. Caire, and M. J. Neely, "MIMO downlink scheduling with non-perfect channel state knowledge," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2055–2066, 2010.