Scheduling for Massive MIMO with Few Excess Antennas

Stefan Dierks and Niklas Jünger
Institute for Communications Engineering,
Technische Universität München,
Munich, Germany
Email: {stefan.dierks, niklas.juenger}@tum.de

***Abstract*—Scheduling all user equipment jointly and using a linear precoder is optimal for massive MIMO with a very large number of base station antennas. Scheduling becomes more important as the number of base station antennas is reduced. We determine the regime where scheduling provides gains for massive MIMO scenarios with linear precoding methods. We apply semi-orthogonal user selection to massive MIMO and propose a scheduling algorithm with smaller complexity.**

I. Introduction

The capacity of multiple-input multiple-output (MIMO) communication systems is achieved by dirty paper coding (DPC) [1]. However linear precoding methods are often preferred because they are simpler and achieve good rates. The performance of DPC is approached even more closely by combining linear precoding and scheduling [2].

We refer to conventional MIMO as communication systems where the number of base station (BS) antennas is less or equal to the number of served user equipments (UEs). A conventional MIMO BS serves a subset of all UEs at any time. The scheduling algorithm selects the subset according to an objective (e.g. maximize sum rate or fairness). Many scheduling algorithms have been proposed [3]. A widely considered scheduling approach is semi-orthogonal user selection (SUS) [2] that was developed for conventional MIMO.

Massive MIMO is a key idea to increase the spectral efficiency of new mobile communication standards (e.g. 5G). It refers to systems where the number of BS antennas exceeds the number of served UEs. For massive MIMO with sufficient randomness in the channel and sufficiently many BS antennas the channels hardens [4]. This means that scheduling all UEs and using a linear transmission scheme like zero forcing beamforming (ZFBF) is optimal. In this regime scheduling does not provide gains. However scheduling improves performance when the channels to the UEs are correlated [2] or when the excess of BS antennas is small [5].

We analyze for a fixed number of UEs the gain of scheduling with increasing numbers of BS antennas using the suboptimal transmission scheme ZFBF. We obtain results for uncorrelated and correlated fading channels. As expected ZFBF and all UEs being scheduled at all time instants approaches capacity (achieved by DPC) as the number of BS antennas increases. In the regime of few excess BS antennas compared to served UEs the gap to capacity is large. However this regime is favorable for practical implementations as more antennas mean higher cost and higher space consumption. Scheduling helps to bridge the gap to capacity.

We apply SUS to massive MIMO and achieve a performance close to capacity. We propose a scheduling algorithm with smaller complexity. The simulation results show that its performance is similar.

In [6] scheduling serves a different purpose. The idea, which is called Joint Spatial Division and Multiplexing (JSDM), is to partition UEs into groups based on their channel's covariance matrices. This allows dividing precoding into two stages. In the pre-beamforming stage the groups are separated using the dominant eigenvectors of each group's channel covariance matrix. The dimensionality of the effective channels after the pre-beamforming stage is reduced compared to the original channel. The precoding of the second stage combats the inter-group interference based on the effective channels after the first stage. The required channel state information (CSI) is reduced as the first stage requires longterm statistical CSI only.

The works [7] and [8] on scheduling for massive MIMO communication systems are based on JSDM. There the focus is somewhat different to our work as the number of UEs is larger than the number of BS antennas. The task of the scheduler is to select UEs which form well separated groups for JSDM. In [8] it is also shown that random beamforming [9] performs poorly for a finite number of UEs.

In [10] a fixed number of heterogeneous UEs are scheduled based on the norm of their instantaneous CSI. The work targets scenarios where the excess of BS antennas to served UEs is large in contrast to our work. The channel is then asymptotically orthogonal and transmitting to the UEs with the largest channel norm is optimal.

II. System Model

Consider a MIMO broadcast channel (BC) with N single antenna UEs. The number of BS antennas is M. Let $M \geq K$ to obtain a massive MIMO scenario.

The received signal of the k-th UE is

$$y_k = \mathbf{h}_k^{\mathsf{H}} \mathbf{x} + z_k \qquad k = 1, \ldots, K \tag{1}$$

where $\mathbf{h}_k$ is the vector of channel coefficients from the BS to the k-th UE, $\mathbf{x}$ is the transmitted signal at the BS and z_k

is independent proper complex thermal Gaussian noise with variance σ_N^2. $[\cdot]^\mathsf{H}$ denotes the complex conjugate transpose.

The antenna correlation at the BS is

$$\mathbf{h}_k = \mathbf{h}_{\text{i.i.d.}} \mathbf{R}_{\text{BS}}^{\frac{1}{2}} \quad (2)$$

where $\mathbf{h}_{\text{i.i.d.}}$ has i.i.d. zero-mean, unit variance proper complex Gaussian entries and $\mathbf{R}_{\text{BS}}$ is the antenna correlation matrix.

We use two channel correlation models. In the uncorrelated fading channel model the antenna correlation matrix is an identity matrix $\mathbf{R}_{\text{BS}} = \mathbf{I}$. We use the correlated fading channel model from [11], where we assume that the UEs are located at the "broadside" of the BS antenna array. The scatterers are located around each UE with an angular spread Δ and the antenna spacing is $\lambda/2$.

The set of UEs scheduled is

$$\mathcal{S} \subset \{1, \ldots, K\}. \quad (3)$$

The received signals of the UEs scheduled in $\mathcal{S}$ are

$$\mathbf{y}(\mathcal{S}) = \mathbf{H}(\mathcal{S})\,\mathbf{x} + \mathbf{z} \quad (4)$$

where $\mathbf{H} = [\mathbf{h}_{(1)}, \ldots, \mathbf{h}_{(|\mathcal{S}|)}]^\mathsf{H}$ is the collection of channel vectors of the $|\mathcal{S}|$ scheduled UEs and $\mathbf{z} = [z_1, \ldots, z_{|\mathcal{S}|}]^\mathsf{T}$. For linear precoding the transmitted signals $\mathbf{x}$ are

$$\mathbf{x} = \mathbf{W}(\mathcal{S})\,\mathbf{s}(\mathcal{S}) \quad (5)$$

where $\mathbf{W}(\mathcal{S}) = [\mathbf{w}_{(1)}, \ldots, \mathbf{w}_{(|\mathcal{S}|)}]$ is the matrix of the precoding vectors and $\mathbf{s}(\mathcal{S})$ is the vector of unit variance transmit symbols.

The BS has an average sum power constraint

$$\sum_{i \in \mathcal{S}} P_i \leq P_{\text{sum}} \quad (6)$$

where P_i is the power allocated for the i-th UE and P is the maximal sum power. Perfect CSI at all nodes is assumed.

A. Zero-Forcing Beamforming

For ZFBF the linear precoders are determined according to an interference zero forcing objective. The optimal solution given a sum power constraint is the pseudo-inverse combined with a power allocation [12]

$$\mathbf{W}(\mathcal{S}) = \mathbf{H}(\mathcal{S})^\mathsf{H} \left(\mathbf{H}(\mathcal{S})\,\mathbf{H}(\mathcal{S})^\mathsf{H}\right)^{-1} \text{diag}(\mathbf{p}(\mathcal{S})) \quad (7)$$

where $\mathbf{p} = [P_{(1)}, \ldots, P_{(|\mathcal{S}|)}]$ is the power allocation vector of non-negative reals. With this choice of precoding matrix the received signals are

$$\mathbf{y}(\mathcal{S}) = \mathbf{H}(\mathcal{S})\,\mathbf{W}(\mathcal{S})\,\mathbf{s}(\mathcal{S}) + \mathbf{z} \quad (8)$$
$$= \text{diag}(\mathbf{p}(\mathcal{S}))\,\mathbf{s}(\mathcal{S}) + \mathbf{z}. \quad (9)$$

The rate achieved with ZFBF and schedule $\mathcal{S}$ is

$$R(\mathcal{S}) = \max_{\mathbf{p}} \sum_{i \in \mathcal{S}} \log(1 + P_i) \quad (10)$$
$$\text{s.t.} \sum_{i \in \mathcal{S}} \gamma_i^{-1} P_i \leq P_{\text{sum}} \quad (11)$$

where

$$\gamma_i = \frac{1}{\|\mathbf{w}_i\|^2}. \quad (12)$$

Note that the values γ_i depend on $\mathcal{S}$. We determine the optimal choice of $\mathbf{p}(\mathcal{S})$ by waterfilling.

III. Scheduling

The optimal schedule is found by exhaustive search. We calculate the achieved sum rates of all combinations of UEs and find the combination with the maximal sum rate

$$\mathcal{S}_{\text{opt}} = \arg \max_{\mathcal{S} \subset \{1,\ldots,K\}} R(\mathcal{S}). \quad (13)$$

This approach is limited to few UEs as the number of combinations increases exponentially with the number of UEs. In the following we describe the original SUS algorithm and our proposed algorithm.

A. Semi-orthogonal User Selection

The SUS algorithm [2] was designed for conventional MIMO where the number of BS antennas is less or equal to the number of served UEs. It finds a suboptimal user group with the objective of maximizing the sum rate. It achieves the same asymptotic performance as DPC [2].

The idea for selecting the user group is as follows: First the UE with the largest channel norm is scheduled. In each following iteration the SUS algorithm schedules the UE with the largest orthogonal component to the subspace spanned by the already scheduled UEs. The key novelty of the SUS algorithm is that after each iteration the semi-orthogonality of each unscheduled UE to the current scheduled UE is determined. When the degree of semi-orthogonality is too small the UE is removed from the set of unscheduled UEs. The result of the algorithm is a set of scheduled UEs. The channel vectors of the UEs are as orthogonal as possible to each other and their norms are as large as possible.

The steps to find a suboptimal user group are [2]:

Step 1: Initialization:

$$\mathcal{T}_1 = \{1, \ldots, K\} \quad (14)$$
$$i = 1 \quad (15)$$
$$\mathcal{S} = \emptyset. \quad (16)$$

Step 2: For each UE $k \in \mathcal{T}_i$, calculate $\mathbf{g}_k$, the component of $\mathbf{h}_k$ orthogonal to the subspace spanned by $\{\mathbf{g}_{(1)}, \ldots, \mathbf{g}_{(i-1)}\}$:

$$\mathbf{g}_k = \mathbf{h}_k - \sum_{j=1}^{i-1} \frac{\mathbf{h}_k \mathbf{g}_{(j)}^*}{\|\mathbf{g}_{(j)}\|^2} \mathbf{g}_{(j)}. \quad (17)$$

When $i = 1$, this implies $\mathbf{g}_{(k)} = \mathbf{h}_{(k)}$.

Step 3: Select the i-th UE as follows:

$$\pi(i) = \arg \max_{k \in \mathcal{T}_i} \|\mathbf{g}_k\| \quad (18)$$
$$\mathcal{S} \leftarrow \mathcal{S} \cup \{\pi(i)\} \quad (19)$$
$$\mathbf{h}_{(i)} = \mathbf{h}_{\pi(i)} \quad (20)$$
$$\mathbf{g}_{(i)} = \mathbf{g}_{\pi(i)}. \quad (21)$$

Step 4: If $|\mathcal{S}| < \min\{M, K\}$, then calculate $\mathcal{T}_{i+1}$, the set of UE semi-orthogonal to $\mathbf{g}_{(i)}$:

$$\mathcal{T}_{i+1} = \left\{ k \in \mathcal{T}_i,\ k \neq \pi(i) \mid \frac{|\mathbf{h}_k \mathbf{g}^*_{(i)}|}{\|\mathbf{h}_k\| \|\mathbf{g}_{(i)}\|} < \alpha \right\} \tag{22}$$

$$i \leftarrow i + 1 \tag{23}$$

where α is a small positive constant. If the set $\mathcal{T}_{i+1}$ is nonempty and the cardinality of $\mathcal{S}$ fullfills $|\mathcal{S}| < \min\{M, K\}$, then go to Step 2. Otherwise the algorithm terminates.

The variable α characterizes the degree of required semi-orthogonality between two channel vectors. For smaller α more UEs are removed. The optimal α is determined with numerical simulations.

B. Massive MIMO Pair-wise SUS Algorithm

The SUS algorithm presented in the previous section starts with an empty set of scheduled UEs and adds a UE in every step. A natural question is whether one can improve the performance and/or the complexity by initializing the set of scheduled UEs as all UEs for a massive MIMO scenario.

We propose an approach that we call massive MIMO pair-wise semi-orthogonal user selection (pair-wise SUS). The first schedule is all UEs. At each iteration the scheduling algorithm finds the UE pair with the smallest degree of orthogonality. From this pair the UE with the smaller channel vector norm is removed. This continues until the degrees of orthogonality between the remaining UEs are large enough.

The steps of the pair-wise SUS algorithm are:
Step 1: Initialization:

$$\mathcal{S}_0 = \{1, \ldots, K\} \tag{24}$$

$$i = 1. \tag{25}$$

Step 2: Determine and store the degrees of orthogonality $\beta_{k,j}$ between all UE channel pairs $j \neq k$:

$$\beta_{k,j} = \frac{|\mathbf{h}_k \mathbf{h}^*_j|}{\|\mathbf{h}_k\| \|\mathbf{h}_j\|}. \tag{26}$$

Note that $\beta_{k,j} = \beta_{j,k}$.
Step 3: Find pair $\mathcal{P}_i$ of the UEs scheduled at the i-iteration $\mathcal{S}_i$ with the smallest degree of orthogonality:

$$\mathcal{P}_i = \{k, j\} = \arg \max_{k,j \in \mathcal{S}_i} \beta_{k,j}. \tag{27}$$

If $\beta_{\mathcal{P}_i} < \beta_{\text{min}}$ the algorithm terminates.
Step 4: Select the i-th UE to be eliminated as follows:

$$\pi(i) = \arg \min_{r \in \mathcal{P}_i} \|\mathbf{h}_r\| \tag{28}$$

$$\mathcal{S}_{i+1} = \{k \in \mathcal{S}_i \mid k \neq \pi(i)\} \tag{29}$$

$$i \leftarrow i + 1. \tag{30}$$

If $|\mathcal{S}_i| > 1$ go to Step 3.

The value β_{min} is a small positive constant. It characterizes the allowed degree of semi-orthogonality between two channel vectors. For smaller β_{min} more UEs are removed. The optimal β_{min} is determined with numerical simulations.

Note that removing a UE based on the orthogonality to individual other UEs is usually suboptimal when maximizing the sum rate.

C. Complexity Analysis

The complexity of the two algorithms differs only in the selection of the scheduled UEs, while the complexity of the zero-forcing beamforming is one $|S| \times |S|$ matrix inversion and one water filling over $|S|$ UEs. This is rather small compared to the user scheduling [2]. We compare the required number of multiplications for the two scheduling algorithms.

In [2] the complexity of the SUS algorithm is found to be upperbounded by

$$C_{\text{SUS}} \leq (C_{\text{mm}} + C_{\text{vn}} + C_{\text{ip}}) \sum_{i=1}^{K} |\mathcal{T}_i| \tag{31}$$

where C_{mm} is the complexity of one $M \times (M \times M)$ vector-matrix multiplication in step 2 of the SUS algorithm, C_{vn} is the complexity of one vector 2-norm in step 3 and C_{ip} is the complexity of one normalized inner-product in step 4. We bound

$$K \leq \sum_{i=1}^{K} |\mathcal{T}_i| \leq K^2. \tag{32}$$

The complexity of the pair-wise SUS algorithm is upper bounded by

$$C_{\text{SUS}} \leq C_{\text{ip}} \frac{K(K-1)}{2} + C_{\text{vn}} K. \tag{33}$$

In step 2 of the pair-wise SUS algorithm a normalized inner-product calculation is required for each UE pair. Recall that step 2 is executed once and that the calculation is invariant to permutation of the UE pair. Hence the number of required normalized inner-product calculations is $K(K-1)/2$. The number of vector 2-norm calculations in step 3 is upper bounded by K.

The bound (32) is usually closer to the lower bound [2]. Hence the number of inner products is larger for the pair-wise SUS algorithm. On the other no vector-matrix multiplications are required for the pair-wise SUS algorithm. As the number of BS antennas M is larger than the number of UEs K we conclude that the complexity of the pair-wise SUS algorithm is smaller than the SUS algorithm.

IV. Simulation Results

Consider a communication system with 10 UEs and one BS. The number of BS antennas is varied between 10 and 100. The SNR is 10 dB. In this SNR regime maximum ratio transmission (MRT) is outperfomed by ZFBF. We average over 500 channel realizations. The capacity is obtained as in [1].

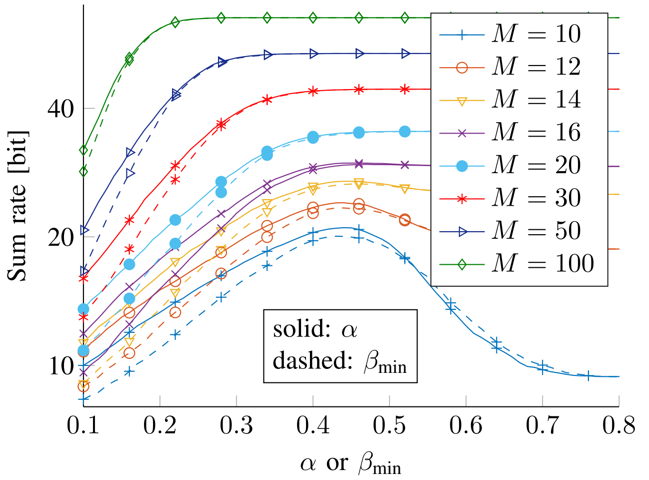

Fig. 1. Sum rate as a function of α for M BS antennas, i.i.d. channel coefficients, 10 UEs and an SNR of 10 dB. Note that the y-axis is logarithmic.

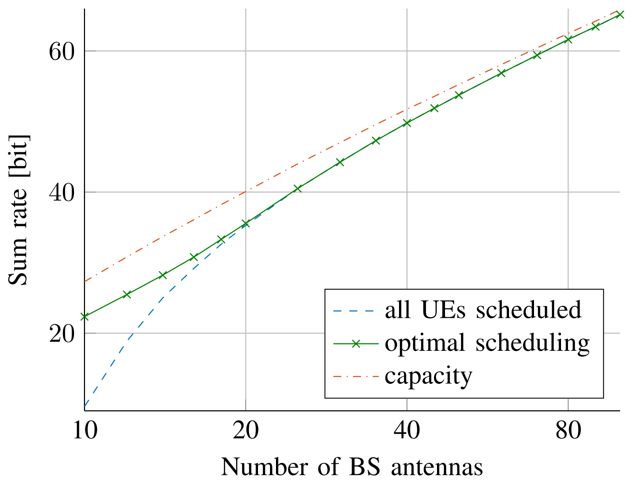

Fig. 2. Capacity and suboptimal linear precoding rates for i.i.d. channel coefficients, 10 UEs and an SNR of 10 dB. Note that the x-axis is logarithmic.

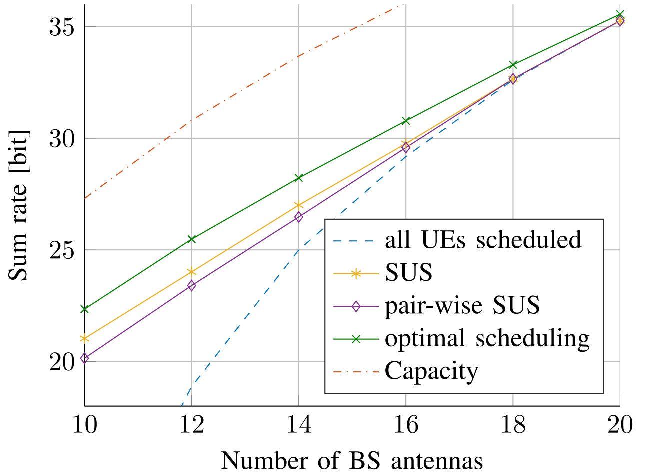

Fig. 3. Comparison of scheduling algorithms for i.i.d. channel coefficients, 10 UEs and an SNR of 10 dB. Note that the x-axis is linear.

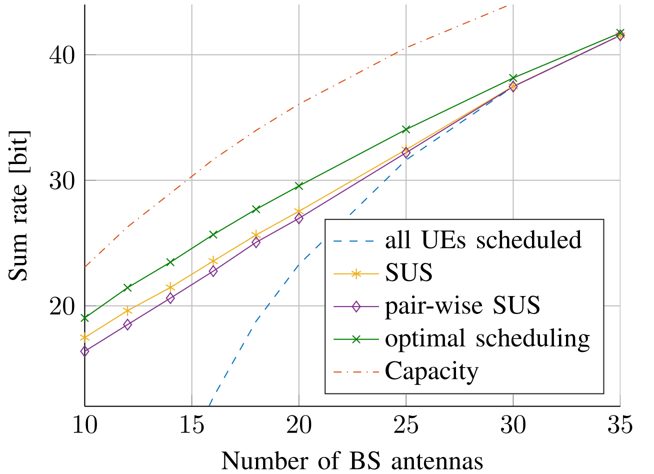

Fig. 4. Comparison of scheduling algorithms for correlated fading with an angular spread of $\Delta = 30°$, 10 UEs and an SNR of 10 dB. Note that the x-axis is linear.

A. Uncorrelated Channel Fading

For uncorrelated fading the channel coefficients are i.i.d. zero-mean circularly symmetric complex Gaussian random variables. The optimal degree of required semi-orthogonality α and $\beta_{\min}$ used in the two algorithms are determined numerically as in Figure 1. The optimal degrees of semi-orthogonality are almost the same for both algorithms. Note that we have to determine the optimal values for each parameter set.

Figure 2 shows the sum-rate versus the number of BS antennas. The gap between optimal scheduling and capacity for a small number of BS antennas is due to suboptimal ZFBF precoding. As the number of BS antennas increases the gap vanishes. For more than twice as many BS antennas as served UEs all scheduling algorithms perform the same. Here scheduling all UEs is optimal. Hence scheduling helps save costs by operating efficiently in the regime of less than twice as many BS antennas as served UEs.

Figure 3 shows the sum rates for the different scheduling algorithms for the regime between 10 and 20 BS antennas. All presented scheduling algorithms achieve a performance close to optimal scheduling. The gap between the proposed algorithm and SUS is small.

B. Correlated Channel Fading

We assume a correlated fading as in [11] with an angular spread $\Delta = 30$. The optimal α and $\beta_{\min}$ are again determined numerically for each parameter set. Here again the suboptimal ZFBF precoding approaches capacity with increasing number of BS antennas. In Figure 4 the sum rate is shown for the for the regime between 10 and 35 BS antennas. Note that with increasing M the suboptimal ZFBF precoding approaches capacity. For this correlated scenario 3.5-times as many BS antennas as served UEs are required for scheduling all UEs to be optimal. Again scheduling helps to bridge the gap. Both scheduling algorithms perform similarly.

V. Conclusions

We compared the performance of the SUS algorithm, the proposed algorithm, optimal scheduling, and capacity in a massive MIMO scenario. We show that for a smaller excess of BS antennas to served UEs, scheduling provides gains over scheduling all UEs. The proposed scheduling algorithm performs similar as SUS, while its complexity is smaller.

Acknowledgment

S. Dierks was supported by the German Ministry of Education and Research in the framework of an Alexander von Humboldt Professorship.

References

[1] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.

[2] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[3] W. Ajib and D. Haccoun, "An overview of scheduling algorithms in MIMO-based fourth-generation wireless systems," *IEEE Netw.*, vol. 19, no. 5, pp. 43–48, Sep. 2005.

[4] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] S. Dierks, W. Zirwas, M. Jäger, B. Panzner, and G. Kramer, "MIMO and Massive MIMO - Analysis for a Local Area Scenario," in *Proc. of the 23nd European Signal Processing Conference (EUSIPCO)*, Sep. 2015, pp. 2496–2500.

[6] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing - The Large-Scale Array Regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[7] Y. Xu, G. Yue, N. Prasad, S. Rangarajan, and S. Mao, "User grouping and scheduling for large scale mimo systems with two-stage precoding," in *IEEE International Conference on Communications (ICC)*, Jun. 2014, pp. 5197–5202.

[8] G. Lee and Y. Sung, "Asymptotically optimal simple user scheduling for massive MIMO downlink with two-stage beamforming," in *IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2014, pp. 60–64.

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

[10] S. L. H. Nguyen, T. Le-Ngoc, and A. Ghrayeb, "Robust Channel Estimation and Scheduling for Heterogeneous Multiuser Massive MIMO Systems," in *Proc. of 20th European Wireless Conference*, May 2014, pp. 1–6.

[11] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[12] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-Forcing Precoding and Generalized Inverses," *IEEE Trans. Signal Proc.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.