Efficient User Selection Algorithms for Multiuser MIMO Systems with Zero-Forcing Dirty Paper Coding

Youxiang Wang, Soojung Hur, Yongwan Park, and Jeong-Hee Choi

Abstract: **This paper investigates the user selection problem of successive zero-forcing precoded multiuser multiple-input multiple-output (MU-MIMO) downlink systems, in which the base station and mobile receivers are equipped with multiple antennas. Assuming full knowledge of the channel state information at the transmitter, dirty paper coding (DPC) is an optimal precoding strategy, but practical implementation is difficult because of its excessive complexity. As a suboptimal DPC solution, successive zero-forcing DPC (SZF-DPC) was recently proposed; it employs partial interference cancellation at the transmitter with dirty paper encoding. Because of a dimensionality constraint, the base station may select a subset of users to serve in order to maximize the total throughput. The exhaustive search algorithm is optimal; however, its computational complexity is prohibitive. In this paper, we develop two low-complexity user scheduling algorithms to maximize the sum rate capacity of MU-MIMO systems with SZF-DPC. Both algorithms add one user at a time. The first algorithm selects the user with the maximum product of the maximum column norm and maximum eigenvalue. The second algorithm selects the user with the maximum product of the minimum column norm and minimum eigenvalue. Simulation results demonstrate that the second algorithm achieves a performance similar to that of a previously proposed capacity-based selection algorithm at a high signal-to-noise (SNR), and the first algorithm achieves performance very similar to that of a capacity-based algorithm at a low SNR, but both do so with much lower complexity.**

Index Terms: **Multiuser multi-input and multi-output (MU-MIMO), precoding, user selection, successive zero-forcing (SZF) dirty paper coding (DPC).**

I. INTRODUCTION

Multiuser multiple-input multiple-output (MU-MIMO) systems have attracted much attention because of their large spectral efficiencies, especially in downlink broadcast scenarios. Dirty paper coding (DPC) [1] has been shown to achieve the sum-rate capacity of a MIMO broadcast channel [2], [3]. However, DPC is difficult to implement in practical systems because of its excessive complexity. Hence, suboptimal precoding schemes with lower complexity have been developed, such as zero-forcing DPC (ZF-DPC) [2], [4], and zero-forcing beamforming (ZF-BF) [5] for single antenna users; block diagonalization (BD) [6] and successive zero-forcing (SZF) precoding [6], [7] have been proposed for multiple antenna users. In ZF-BF, the transmitter transmits signals multiplied by the beamforming weight vectors to cancel interference on the receive side. However, this technique suffers from enhanced power noise, and it cannot be easily generalized to users with multiple receiving antennas. BD transmission scheme was proposed for the multiuser downlink channel assuming users have multiple receive antennas [6]. It is an efficient linear precoding technique for eliminating inter-user interference with relatively low complexity, but the achievable throughput is significantly reduced compared to a system that employs DPC. The SZF precoding scheme is an efficient method of balancing encoding complexity and system throughput.

To avoid interference when the number of users is large in a downlink MU-MIMO scenario, the maximum number of users that could be supported by the base station (BS) is limited by the number of transmit antennas and receive antennas each user has. The transmitter can schedule its transmission to those users with favorable channel fading conditions to improve the system throughput [4]. Several algorithms related to user selection in downlink MU-MIMO systems have been considered. An exhaustive search over all possible user subsets is the optimal algorithm, but it requires highly complex calculation. Several suboptimal schemes have been proposed to reduce this computational complexity. Two suboptimal user selection schemes were proposed based on BD were proposed [8], but their computational complexity is still too high. Chen and Wang [9] developed max-max and max-min scheduling algorithms for a MIMO downlink system that uses the simple spatial multiplexing at the transmitter and ZF processing at the receiver. Both algorithms require that ZF be performed at the receiver in order to calculate the output signal-to-noise (SNR) of each subchannel for all users. The receiver sends the maximal subchannel output SNR to the transmitter for the max-max algorithm and sends the minimal subchannel output SNR to the transmitter for the max-min algorithm. A suboptimal scheduling algorithm called semi-orthogonal user selection (SUS) algorithm is based on ZF-BF [5]. However, this algorithm does not always select the best user group. When the orthogonality of the channel vectors of the selected part is not high, the projection of a user channel may cause interference in another user's channel.

In this paper, we consider successive ZF-DPC (SZF-DPC) for the case of users with multiple antennas. We assume that full channel knowledge is available at the transmitter and at each receiver. This scheme can be viewed as an extension of ZF-DPC for users with multiple receive antennas. It is similar to ZF-DPC when each user has one receive antenna, which can achieve an asymptotical sum rate capacity as that of the opti-

Manuscript received October 1, 2009 ; approved for publication by Gerhard P. Fettweis, Guest Editor, May 4, 2010.

This research was supported by the Yeungnam University research grants in 2007.

Y. Wang, S. Hur, and Y. Park are with the Department of Information and Communication Engineering, Yeungnam University, 214-1 Daedong, Gyeongsan-si, 712-749, Korea, email: didibird726@gmail.com, {sjheo, ywpark}@ynu.ac.kr.

J.-H. Choi is with the School of Computer and Communication Engineering, Daegu University, Jillyang, Gyeongsan-si, 712-714, Korea, email: choijh@biho.taegu.ac.kr.

1229-2370/11/$10.00 © 2011 KICS

mal DPC scheme with any ordered set of users at a high SNR and with optimal user ordering at a low SNR [7]. As mentioned above, user scheduling is necessary to avoid interference when the system contains many active users. In this context, two suboptimal algorithms with reduced complexity are proposed for SZF precoding, assuming perfect channels state information regarding all users at the transmitter. The first algorithm, max-max subchannel selection, selects the user with the maximum product of the maximum channel norm and maximum eigenvalue. The second, max-min subchannel selection, selects the user with the maximum product of the minimum channel norm and minimum eigenvalues. Both algorithms aim to maximize the system capacity. Since the user selection criterion is simple and easy to run recursively, the computational complexity is lower than that of an exhaustive search or the scheduling algorithms in [8].

This paper is organized as follows. Section II introduces the MU-MIMO downlink system model, reviews the SZF-DPC algorithm, and derives the sum rate capacity of a system that uses SZF-DPC. The proposed scheduling algorithms are described in Section III. Section IV presents simulation results that demonstrate the improved performance offered by the proposed schemes and provides a complexity analysis of the proposed algorithms and other existing algorithms. Conclusions are given in Section V.

We use the following notations in this paper. Uppercase bold letters represent matrices and lowercase bold letters represent vectors. We use $(\cdot)^T$ for a matrix transpose and $(\cdot)^H$ for a conjugate transpose. The notation $\mathrm{diag}(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_N)$ represents a block-diagonal matrix with diagonal matrix elements $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_N$.

II. SYSTEM MODEL AND SUCCESSIVE ZERO-FORCING DIRTY PAPER CODING

We consider the downlink of a MU-MIMO system with K users having $N_1, N_2, \cdots, N_K$ receive antennas and a BS with M transmit antennas. We assume a spatially uncorrelated flat Rayleigh fading channel between each user and the BS. On the transmit side, the transmit data symbol vector of each user, i.e., $\mathbf{s}_k \in \mathcal{C}^{L_k}$, $k = 1, \cdots, K$ is passed through a certain transmit precoding matrix $\mathbf{W}_k \in \mathcal{C}^{M \times L_k}$ before it is launched into the downlink channel. The received signal vector for user k can then be described as

$$\mathbf{y}_k = \mathbf{H}_k \sum_{i=1}^{K} \mathbf{W}_i \mathbf{s}_i + \mathbf{n}_k, \qquad k = 1, \cdots, K \tag{1}$$

where $\mathbf{y}_k \in \mathcal{C}^{N_k \times 1}$ is the received vector; $\mathbf{H}_k \in \mathcal{C}^{N_k \times M}$ denotes the downlink channel matrix of user k, the elements of which are zero-mean unit-variance independent and identically distributed (i.i.d) complex Gaussian random variables; $\mathbf{n}_k \sim \mathcal{CN}(0, \mathbf{I}_{N_k})$ is an additive Gaussian noise vector; and $\mathbf{I}_{N_k}$ is an $N_k \times N_k$ identity matrix.

To avoid interference, the BD algorithm designs the precoding matrix $\mathbf{W}_k$ into the null space of all the other users' channels $\{\mathbf{H}_j\}_{j \neq k}$, such that

$$\mathbf{H}_j \mathbf{W}_k = 0, \ \forall j \neq k \text{ and } 1 \leq k, \ j \leq K. \tag{2}$$

This decomposes the multiuser channel into equivalent single-user channels, and the received signal vector $\mathbf{y}_k$ is reduced to

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{W}_k \mathbf{s}_k + \mathbf{n}_k. \tag{3}$$

To maximize the system's throughput under the BD method, the set of precoding matrices should satisfy the total power constraint of the transmitted signals.

A. Successive Zero-Forcing Dirty Paper Coding

Because of the complexity of DPC, Caire and Shamai [3] proposed ZF-DPC for single-antenna users. ZF-DPC uses QR decomposition combined with DPC at the transmitter of optimizing the sum capacity. SZF-DPC can be used to pre-eliminate inter-user interference when each user has multiple receive antennas; its operation is similar to that of ZF-DPC. At a high SNR, for a given user order π, SZF-DPC has been shown to achieve the sum capacity of DPC for that user order [7]. However, optimal user ordering is required for SZF-DPC to achieve sum capacity of DPC at a low SNR [10].

The user precoding order is considered in SZF-DPC. Given an ordered set of users with an order π, for each user $j \in \{2, 3, \cdots, K\}$ the precoding matrix $\mathbf{W}_{\pi(j)}$ is designed such that it lies in the null space of the aggregate channel $\bar{\mathbf{H}}^{j-1}$ of $j-1$ previously precoded users' channels, where

$$\bar{\mathbf{H}}^{j-1} = [\mathbf{H}_{\pi(1)}^T \ \ \mathbf{H}_{\pi(2)}^T \ \ \cdots \ \ \mathbf{H}_{\pi(j-1)}^T]^T. \tag{4}$$

Then, the received signal of the jth user is

$$\begin{aligned} \mathbf{y}_{\pi(j)} = & \mathbf{H}_{\pi(j)} [\mathbf{W}_{\pi(j)} \mathbf{s}_{\pi(j)} + \sum_{i<j} \mathbf{W}_{\pi(i)} \mathbf{s}_{\pi(i)} \\ & + \sum_{i>j} \mathbf{W}_{\pi(i)} \mathbf{s}_{\pi(i)}] + \mathbf{n}_{\pi(j)}. \end{aligned} \tag{5}$$

For each j, the term $\sum_{i>j} \mathbf{W}_{\pi(i)} \mathbf{s}_{\pi(i)}$ is canceled by the above subspace constraint of the precoding matrices. The encoder considers the interference signal $\sum_{i<j} \mathbf{W}_{\pi(i)} \mathbf{s}_{\pi(i)}$ caused by users $i < j$ to be known noncausally, and it can be canceled by DPC. If we apply singular value decomposition (SVD) to matrix $\bar{\mathbf{H}}^{j-1}$, then

$$\bar{\mathbf{H}}^{j-1} = \bar{\mathbf{U}}^{j-1} \bar{\mathbf{\Sigma}}^{j-1} [\bar{\mathbf{V}}^{(j-1,1)} \ \ \bar{\mathbf{V}}^{(j-1,0)}]^H \tag{6}$$

where matrix $\bar{\mathbf{V}}^{(j-1,1)}$ denotes the right singular mateix consisting of the singular vectors corresponding to nonzero singular values of $\bar{\mathbf{H}}^{j-1}$ and the columns in $\bar{\mathbf{V}}^{(j-1,0)}$ form the basis set for the null space of $\bar{\mathbf{H}}^{j-1}$. Then, the precoding matrix of the jth user $\mathbf{W}_{\pi(j)}$ is is constrained to lie on the subspace defined by $\bar{\mathbf{V}}^{(j-1,0)}$.

B. Sum Rate Capacity of SZF-DPC

Any rate in the capacity region of the broadcast channel (achieved with DPC) was shown to be within the capacity region of the dual media access controller (achieved with successive decoding) under the sum power constraint [2]. Therefore, the sum-rate capacities of the system with SZF-DPC are given by

$$C_K^{\text{SZF-DPC}} = \max\Big\{\text{logdet}\Big[\mathbf{I} + \sum_{j=1}^{K}(\mathbf{H}_{\pi(j)}\bar{\mathbf{V}}^{(j-1,0)})^H \mathbf{Q}_{\pi(j)}(\mathbf{H}_{\pi(j)}\bar{\mathbf{V}}^{(j-1,0)})\Big]\Big\}, \quad (7)$$

$$\text{for } \mathbf{Q}_{\pi(j)} : \mathbf{Q}_{\pi(j)} \geq 0 \text{ and } \sum_{j\in\{1,2,\cdots,K\}} \text{Tr}(\mathbf{Q}_{\pi(j)}) \leq P$$

where $\mathbf{Q}_{\pi(j)}$ is the matrix obtained by waterfilling power allocation over the singular values of the block-diagonal matrix

$$\hat{\mathbf{H}}_K = \text{diag}\Big\{\mathbf{H}_{\pi(1)}\bar{\mathbf{V}}^{(0,0)}, \mathbf{H}_{\pi(2)}\bar{\mathbf{V}}^{(1,0)}, \cdots, \mathbf{H}_{\pi(K)}\bar{\mathbf{V}}^{(K-1,0)}\Big\} \quad (8)$$

under power constraint P. $\bar{\mathbf{V}}^{(0,0)} = \mathbf{I}$ is the null space matrix of the first user. To simplify the notation, we let $\tilde{\mathbf{H}}_j = \mathbf{H}_{\pi(j)}\bar{\mathbf{V}}^{(j-1,0)}$. Let $\tilde{\mathbf{h}}_j^i$ denote the ith column of $\tilde{\mathbf{H}}_j$ and $\ell_i = \sum_{j=1}^{K}(\tilde{\mathbf{h}}_j^i)^H\mathbf{Q}_{\pi(j)}\tilde{\mathbf{h}}_j^i$ for $i = 1, 2, \cdots, M$. Letting $\mathbf{\Theta} = \sum_{j=1}^{K}\tilde{\mathbf{H}}_j^H\mathbf{Q}_{\pi(j)}\tilde{\mathbf{H}}_j + \mathbf{I}$, we have

$$\mathbf{\Theta} = \begin{pmatrix} \ell_1+1 & \cdots & \cdots & \cdots \\ \cdots & \ell_2+1 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \cdots \\ \cdots & \cdots & \cdots & \ell_M+1 \end{pmatrix}. \quad (9)$$

Then, we can obtain the sum rate capacity as

$$C_K^{\text{SZF-DPC}} \leq \log\Big[\prod_{i=1}^{M}(\ell_i+1)\Big]. \quad (10)$$

Using Hadamard's inequality reveals that $\mathbf{\Theta}$ is a positive definite [11]. Let λ_j^l denotes the lth eigenvalue of $\mathbf{Q}_{\pi(j)}$ and $\lambda_j^{\max} = \max_l \lambda_j^l$, $\lambda_j^{\min} = \min_l \lambda_j^l$. Also, let $h_i^{\max} = \max_j \|\tilde{\mathbf{h}}_j^i\|^2$ and $h_i^{\min} = \min_j \|\tilde{\mathbf{h}}_j^i\|^2$. Then,

$$\sum_{j=1}^{K}\|\tilde{\mathbf{h}}_j^i\|^2\lambda_j^{\min} \leq \ell_i \leq \sum_{j=1}^{K}\|\tilde{\mathbf{h}}_j^i\|^2\lambda_j^{\max}$$

$$\Rightarrow h_i^{\min}\sum_{j=1}^{K}\lambda_j^{\min} \leq \ell_i \leq h_i^{\max}\sum_{j=1}^{K}\lambda_j^{\max}$$

$$\text{for } i = 1, 2, \cdots, M \quad (11)$$

where $\|\cdot\|$ denotes the Euclidean norm and the first inequality above is in view of the Rayleigh-Ritz theorem [10]. $\|\tilde{\mathbf{h}}_j^i\|^2$ is actually the effective channel power gain if maximum ratio combining is employed at the receiver for the link between user j and base antenna i [10]. Here ℓ_i has the upper bound $h_i^{\max}\sum_{j=1}^{K}\lambda_j^{\max}$ and lower bound $h_i^{\min}\sum_{j=1}^{K}\lambda_j^{\min}$; a consideration of (10)reveals that these bounds will limit the achievable capacity $C_K^{\text{SZF-DPC}}$. In the next section, we derive the user selection algorithms using these bounds.

III. USER SELECTION ALGORITHMS

In a MU-MIMO downlink systems with a large number of users contending for services, the BS needs to select a subset of users from among all the users in order to obtain multiuser diversity and maximize the sum capacity. An exhaustive search for selecting the best set of users is very complex. The SUS algorithm was proposed to select users having one receive antenna on the basis of both the channel norm and a spatial property (orthogonality) [5]. This algorithm is similar to the greedy algorithm proposed in [3], but is different in that it uses an orthogonality threshold for selecting the users in each step.

In this paper, we propose two user scheduling algorithms to maximize the sum rate capacity. The proposed scheduling schemes, like the methods proposed in [8], aim to maximize the total throughput while keeping the complexity low. To simplify the problem, we assume that each user employs an equal number of receive antennas, i.e., $N_k = N$, $k = 1, 2, \cdots, K$. With this assumption, the BS can simultaneously transmit to at most $K_0 = \lceil M/N \rceil$ users, where $\lceil x \rceil$ denotes the smallest integer that is larger than or equal to x.

A. Max-Max Subchannel (MMS) Scheduling

Let Ω denote the set of users who have not been allocated and $\pi(i)$ denote the user index allocated in the ith iteration. The first user $\pi(1)$ is selected using the max-max subchannel (MMS) criterion. Considering (10) and (11), we use the max-max subchannel of a matrix because it gives the maximum value of the capacity upper bound from the eigenmodes of the equivalent single-user channel. SVD is applied to the selected user's channel to obtain the null space matrix $\bar{\mathbf{V}}^{(0,0)}$, which is needed in subsequent steps of the algorithm. Then, the effective channel of the remaining unselected users can be constructed as $\hat{\mathbf{H}}_k = \mathbf{H}_k\bar{\mathbf{V}}^{(0,0)}$, and from those users, it can be found that the user $\pi(2)$ which has the maximum product of maximum channel norm and maximum eigenvalue. Next, the BS generates the second precoding matrix $\bar{\mathbf{V}}^{(1,0)}$, which is the null space of $\bar{\mathbf{H}}^1 = [\mathbf{H}_{\pi(1)}^T\ \mathbf{H}_{\pi(2)}^T]$, and it constructs the effective channel matrix of the remaining unselected users. The algorithm terminates when all K_0 users are selected. This is described as follows.

1) Initialization.

Set $\Omega = \{1, 2, \cdots, K\}$, $\mathcal{Q} = \phi$
$h_k^{\max} = \max_{k\in\Omega}\|\mathbf{h}_k^i\|^2$, $i = 1, \cdots, M$,
$\lambda_k^{\max} = \max_{k\in\Omega}\lambda_k^l$, $\quad l = 1, \cdots, N$.
Select a user $\pi(1) = \text{argmax}_{k\in\Omega}\, h_k^{\max}\lambda_k^{\max}$.
Set $\mathcal{Q} = \mathcal{Q} + \{\pi(1)\}$, $\Omega = \Omega\backslash\{\pi(1)\}$.
Perform SVD on $\mathbf{H}_{\pi(1)}$,
$\mathbf{H}_{\pi(1)} = \bar{\mathbf{U}}^0\bar{\Sigma}^0[\bar{\mathbf{V}}^{(0,1)}\ \bar{\mathbf{V}}^{(0,0)}]^H$.

2) For $j = 2 : K_0$, select the jth user as follows.

$$\bar{\mathbf{H}}^{j-1} = [\mathbf{H}_{\pi(1)}^T\mathbf{H}_{\pi(2)}^T\mathbf{H}_{\pi(j-1)}^T]^T = \bar{\mathbf{U}}^{(j-1)}\bar{\Sigma}^{(j-1)}[\bar{\mathbf{V}}^{(j-1,1)}\ \bar{\mathbf{V}}^{(j-1,0)}]^H.$$

Find precoding matrix $\bar{\mathbf{V}}^{(j-1,0)}$.
Obtain the effective channel $\hat{\mathbf{H}}_k = \mathbf{H}_k\bar{\mathbf{V}}^{(j-1,0)}$.
$h_k^{\max} = \max_{k\in\Omega}\|\hat{\mathbf{h}}_k^i\|^2$, $i = 1, \cdots, M$,

$\lambda_k^{\max} = \max_{k\in\Omega} \lambda_k^l, \quad l = 1,\cdots,N.$

Select a user $\pi(j) = \operatorname{argmax}_{k\in\Omega} h_k^{\max}\lambda_k^{\max}$.

Set $\mathcal{Q} = \mathcal{Q} + \{\pi(j)\}; \Omega = \Omega\backslash\{\pi(j)\}$.

End

3) Output: The selected user set $\mathcal{Q}$.

B. Max-min Subchannel(MNS) Scheduling

The algorithm is summarized as follows.

1) Initialization.

Set $\Omega = \{1, 2, \cdots, K\}$, $\mathcal{Q} = \phi$.

$h_k^{\min} = \min_{k\in\Omega} \|\mathbf{h}_k^i\|^2, \ i = 1, \cdots, M,$

$\lambda_k^{\min} = \min_{k\in\Omega} \lambda_k^l, \quad l = 1, \cdots, N.$

Select a user $\pi(1) = \operatorname{argmax}_{k\in\Omega} h_k^{\min}\lambda_k^{\min}$.

Set $\mathcal{Q} = \mathcal{Q} + \{\pi(1)\}; \Omega = \Omega\backslash\{\pi(1)\}$.

Perform SVD on $\mathbf{H}_{\pi(1)}$,

$\mathbf{H}_{\pi(1)} = \bar{\mathbf{U}}^0\bar{\mathbf{\Sigma}}^0[\bar{\mathbf{V}}^{(0,1)}\bar{\mathbf{V}}^{(0,0)}]^H$.

2) For $j = 2 : K_0$, select the jth user as follows.

$$\bar{\mathbf{H}}^{j-1} = [\mathbf{H}_{\pi(1)}^T \quad \mathbf{H}_{\pi(2)}^T \quad \cdots \quad \mathbf{H}_{\pi(j-1)}^T]^T = \bar{\mathbf{U}}^{(j-1)}\bar{\Sigma}^{(j-1)}[\bar{\mathbf{V}}^{(j-1,1)} \ \bar{\mathbf{V}}^{(j-1,0)}]^H.$$

Find precoding matrix $\bar{\mathbf{V}}^{(j-1,0)}$.

Obtian effective channel $\hat{\mathbf{H}}_k = \mathbf{H}_k\bar{\mathbf{V}}^{(j-1,0)}$.

$h_k^{\min} = \min_{k\in\Omega} \|\hat{\mathbf{h}}_k^i\|^2, \ i = 1, \cdots, M,$

$\lambda_k^{\min} = \min_{k\in\Omega} \lambda_k^l, \quad l = 1, \cdots, N.$

Select a user $\pi(j) = \operatorname{argmax}_{k\in\Omega} h_k^{\min}\lambda_k^{\min}$.

Set $\mathcal{Q} = \mathcal{Q} + \{\pi(j)\}; \ \Omega = \Omega\backslash\{\pi(j)\}$

End

3) Output: The selected user set $\mathcal{Q}$.

In MMS selection, the upper bound of ℓ_i in (11) is considered. In this MNS scheme, we consider the lower bound of ℓ_i, which bounds the minimum maximal capacity in (10). The user having the maximal minimal subchannel is selected as the first user, $\pi(1)$. SVD is applied to the selected user's channel to obtain the null space matrix $\bar{\mathbf{V}}^{(0,0)}$, then the effective channel of the remaining unselected users can be constructed as $\hat{\mathbf{H}}_k = \mathbf{H}_k\bar{\mathbf{V}}^{(0,0)}$. Next, we find user $\pi(2)$ who has the maximal minimal subchannel, and generates the second precoding matrix $\bar{\mathbf{V}}^{(1,0)}$, which is the null space of $\bar{\mathbf{H}}_1 = [\mathbf{H}_{\pi(1)}^T \quad \mathbf{H}_{\pi(2)}^T]^T$, in order to construct the effective channel matrix of the remaining unselected users. The above steps are repeated until all values of K_0 are selected.

C. Proportional Fair Scheduling Algorithm

The algorithms discussed above focus on maximizing the sum rate capacity, which may starve some users owing to bad channel conditions. Simple round-robin scheduling is a fair scheduling algorithm that provides equal opportunities to all K users. However, since round-robin supports only one user at a time in a time division multiple access (TDMA) fashion, it does not achieve spatial multiplexing gains. To ensure fairness in some systems, the proportional fair scheduling (PFS) algorithm in [12] can be used to achieve fairness among users while exploiting multiuser diversity gains.

In the round-robin algorithm, we recursively apply user selection as described in subsections III-A and III-B. Specifically, we construct the first user group by running the user selection algorithms, and we then select the second group by repeating the same algorithm for the remaining users. This procedure is repeated until no users are left. Each user obtains the same opportunity to occupy the resource in this algorithm, but there will be some loss of sum capacity.

The PFS algorithm can be used to balance the sum capacity and user fairness. We assume that $\bar{R}_k(t)$ is the average throughput achieved by user k up to time t, which is updated using an exponentially weighted low-pass filter,

$$\bar{R}_k(t+1) = \begin{cases} (1-\frac{1}{t_c})\bar{R}_k(t) + \frac{1}{t_c}R_k(t), & k \in \mathcal{Q}, \\ (1-\frac{1}{t_c})\bar{R}_k(t), & k \notin \mathcal{Q} \end{cases} \tag{12}$$

where $R_k(t)$ is the rate of user k during the tth transmission interval, and t_c is the number of time slots over which the throughput of user K is averaged. In this algorithm, we select one user with the largest ratio of $R_k(t)$ to the average throughput $\bar{R}_k(t)$.

IV. SIMULATION RESULTS

In this section, we compare the performance and computational complexity of the following algorithms.

- The capacity-based user selection algorithm [8],
- the max-max and max-min user selection algorithms [9],
- the channel norm-based user selection algorithm,
- the proposed MMS and MNS user selection algorithms.

A. Performance Comparisons of Different Selection Algorithms

Figs. 1–3 compare the sum capacity of the proposed MMS and MNS algorithms, capacity-based algorithm, channel norm-based algorithm, and max-max and max-min algorithms versus the number of users for a MU-MIMO downlink system with $M = 6$ antennas at the BS and $N_k = N = 2$ for each user. Fig. 4 provides the sum capacity versus SNR of the system with $M = 6$ and $N_k = N = 2$. The simulation results were obtained by averaging over 3000 independent channel realizations. Since the channels of the users are assumed to be spatially uncorrelated, the number of users that supported K_0 is three. As shown in Figs. 1, 2, and 4, our proposed MNS algorithm performs slightly better than that of the max-min algorithm and the results of both algorithms are very close to those of capacity-based algorithm at a high SNR regime. Fig. 3 shows that the proposed MMS and max-max algorithms yield results very close to those of the capacity-based algorithm at a low SNR. The performance exchange can be explained as follows. The sum capacity of the SZF-DPC is expressed as $C_{K_0}^{\mathrm{SZF-DPC}} = \sum_{k=1}^{K_0}\sum_{i=1}^{N}\log(1 + p_{k,i}\lambda_{k,i})$, where $p_{k,i}$ is obtained by using waterfilling over block-diagonal channel matrix $\hat{\mathbf{H}}_{K_0}$ with a total power constraint. For one selected user, the capacity is $C_{K_0}^{\mathrm{SZF-DPC}} = \sum_{i=1}^{N}\log(1 + p_{k,i}\lambda_{k,i}) \approx \sum_{i=1}^{N} p_{k,i}\lambda_{k,i}$ at a low SNR. Therefore, the larger the sum of $p_{k,i}\lambda_{k,i}$ is, the higher the sum capacity is at a low SNR. The highest-order statistic of the subchannel output SNR has been shown

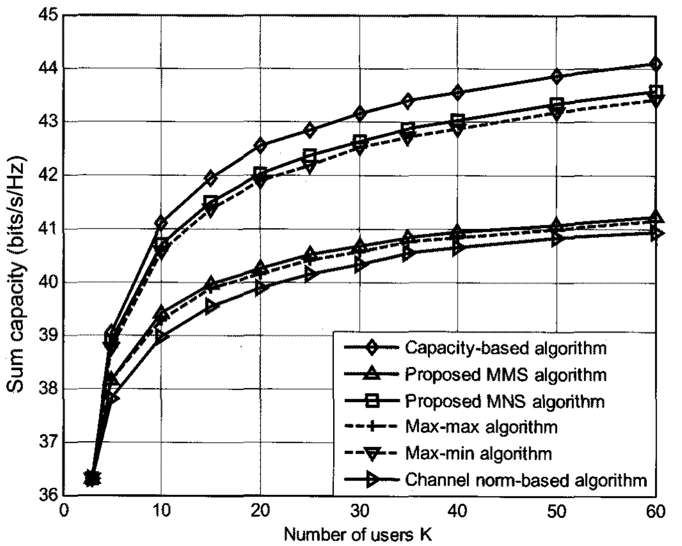


Fig. 1. Sum capacity versus the number of users with $M = 6$, $K_0 = 3$, $N_k = N = 2$, and SNR = 20 dB.

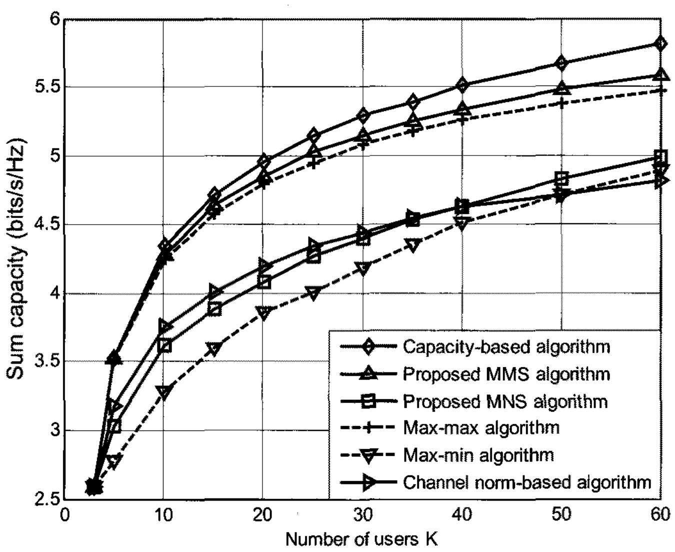


Fig. 3. Sum capacity versus the number of users with $M = 6$, $K_0 = 3$, $N_k = N = 2$, and SNR = 0 dB.

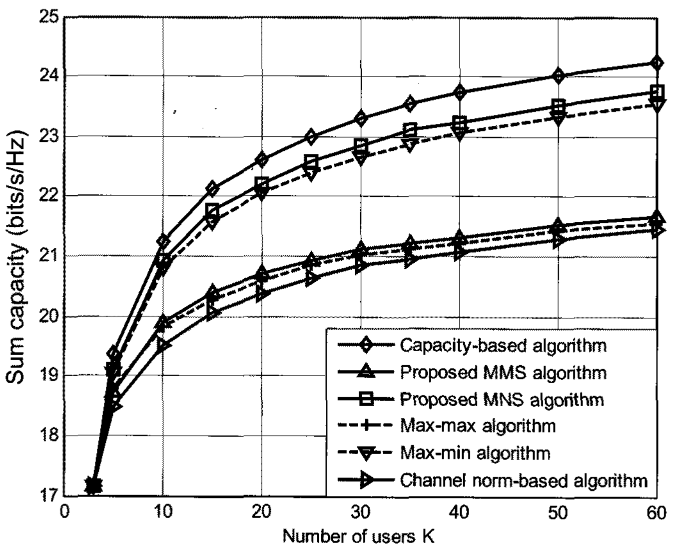


Fig. 2. Sum capacity versus the number of users with $M = 6$, $K_0 = 3$, $N_k = N = 2$, and SNR = 10 dB.

to gain a significant boost under max-min scheduling, whereas the lower-order statistic of the subchannel output SNR improves only slightly [9]. In comparison, max-min scheduling enhances the magnitude of all the subchannels' output SNRs quite uniformly. As the output SNR is correlated with the eigenvalue of the channel matrix, the eigenvalue distribution of the proposed systems has similar properties to those of the proposed MMS and MNS algorithms. Therefore, at a low SNR, the capacity of the proposed MMS algorithm is improved by the significant boost in the highest-order statistic $p_{k,N}\lambda_{k,N}$. At a high SNR, $C_k^{\mathrm{SZF-DPC}} = \sum_{i=1}^{N} \log(1 + p_{k,i}\lambda_{k,i})$ is a concave function, so the contributed throughput from the enhancement of the highest order statistic $p_{k,N}\lambda_{k,N}$ will be suppressed more than for the lower-order statistic $p_{k,i}\lambda_{k,i}$ [9]. Thus, the MMS algorithm yields a lower capacity at a high SNR. From the above analysis, we can conclude that a selection based on the maximum singular value is not appropriate for a ZF receiver because the ZF method is more greatly affected by the minimum singular values, and a high maximum eigenvalue may indicate a much lower minimum eigenvalue, which degrades the performance of ZF processing. At a high SNR, the channel norm-based algorithm performs the worst because it does not consider the spatial correlation between users. At a low SNR, the channel norm-based algorithm outperforms the max-min and MNS algorithms for a small number of users. When the number of users increases to 50, the proposed MNS and max-min algorithms outperform the channel norm-based algorithm. In Fig. 4, we see a negligible difference in performance between the proposed MNS and max-min algorithms because both schemes achieve a performance close that of the capacity-based algorithm. The difference between the proposed MMS and max-max algorithms is also negligible, and both schemes outperform the channel norm-based algorithm. The round-robin scheduling algorithm performed the worst, because it achieves the same fairness for each user without considering their channel conditions. In Fig. 5, we plot the scheduled times each individual user attained under each scheduling strategy in a MU-MIMO downlink system with $M = 6$ and $N_k = N = 2$. We use $K = 30$ users with average received SNRs ranging from 0 dB to 15 dB in a log-linear scale. Each user is identified with a unique user index k $(k = 1, 2, \cdots, K)$ in an increasing SNR order. Thus, user 1 has an SNR of 0 dB, and user 30 has an SNR of 15 dB. The simulation results were obtained by using 3000 independent time slots. Three users are scheduled at each time slot. The sum capacity of each scheduling strategy is presented in Table 1. Fig. 5 and Table 1 show that the proposed user selection algorithms without PFS have the highest sum capacity but the lowest fairness. The proposed user selection algorithm with PFS guarantees fairness at the expense of sum capacity loss. The proposed round-robin algorithm provides deterministic fairness but with the lowest sum capacity. As observed in Fig. 5, the proposed algorithms with PFS and the round-robin algorithm improve the delay characteristics at the expense of sum capacity

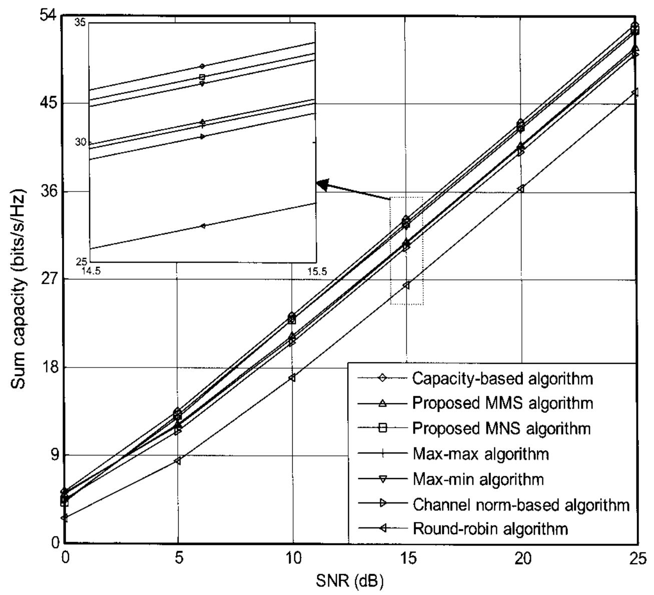


Fig. 4. Sum capacity versus SNR with $M = 6$, $K_0 = 3$, and $N_k = N = 2$.

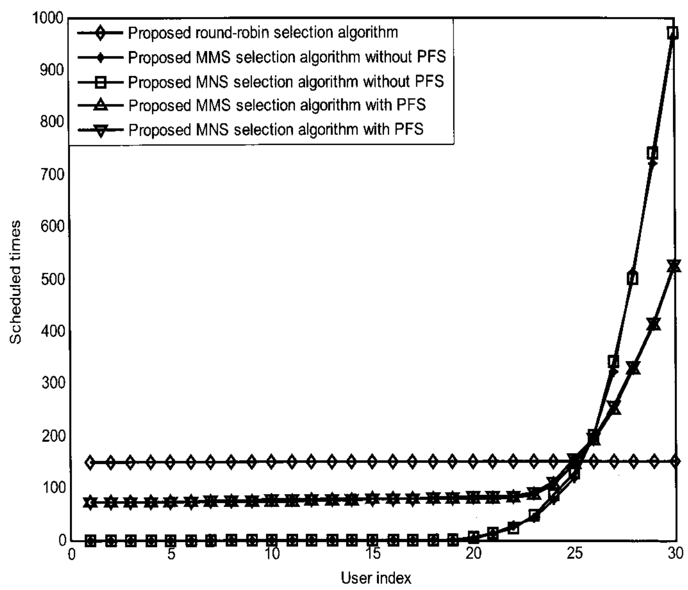


Fig. 5. Fairness comparison of various algorithms with $M = 6$, $K_0 = 3$ $K = 30$, and $N_k = N = 2$.

Table 1. Sum capacity of each scheduling strategy.

Scheduling strategy	Capacity(bps/Hz)
Proposed MNS algorithm without PFS	19.8
Proposed MMS algorithm without PFS	18.21
Proposed MNS algorithm with PFS	15.42
Proposed MMS algorithm with PFS	14.13
Proposed round-robin algorithm	11.36

losses, where more scheduled times indicate a lower delay time.

B. Complexity Analysis

In this section, we compare the computational complexity of our proposed algorithms with capacity-based algorithm and other low-complexity scheduling algorithms in terms of the number of flops required. A flop is defined to be a real floating point operation [8]. A real addition, multiplication, or division is consider to be one flop. Complex addition and multiplication require two and six flops, respectively. For this analysis, we assume that the total number of users is much larger that the number of scheduled users, $K \gg K_0 \gg 1$. Note that $N_k = N$, $\forall k$, and $K_0 = \lceil M/N \rceil$, and the algorithms each schedule the maximum of K_0 users.

For an $m \times n$ complex-valued matrix $\mathbf{A} \in \mathcal{C}^{m\times n}$, we list the complexity of various matrix operations required for our proposed scheduling algorithms.

- The Frobenius norm $\|\mathbf{A}\|_F^2$ requires $2mn$ real multiplications and $2mn$ real additions, therefore, $4mn$ flops are required [14].
- Water-filling over i eigenmodes requires $(i^2 + 3i)/2$ real multiplications, $i^2 + 3i$ real additions, and $(i^2 + 3i)/2$ real divisions. The flop count of water-filling is $2i^2 + 6i$ [8].
- Multiplying an $m\times n$ matrix by an $n\times l$ matrix requires mnl complex additions and mnl complex multiplications, for a total of $8mnl$ flops [8].
- The flop count for SVD of a complex-valued matrix is approximated as $24mn^2 + 48m^2n + 54m^3$ by treating every operation as a complex multiplication [8].

The complexity is denoted as φ. Next, we analyze the computational complexity of the scheduling algorithms presented in this paper.

1) Proposed max-min (max-max) subchannel scheduling algorithm.

- $j = 1$: SVD of $\mathbf{H}_k$ requires $48N^2M + 24NM^2 + 54N^3$ flops, the Frobenius norm of M columns is calculated for the kth user, requiring $4MN$ flops. The total computational complexity of this step is thus about $K(48N_2M + 24NM_2 + 54N_3 + 4MN + 1)$.
- $j \geq 2$: To get $\mathbf{W}_{\pi(j)}$ by SVD requires $48(j-1)^2N^2M + 24(j-1)NM^2 + 54(j-1)^3N^3$ flops. $8NM[M-(j-1)N]$ are required to compute $\bar{\mathbf{H}}_k = \mathbf{H}_k\mathbf{W}$. SVD of $\bar{\mathbf{H}}_k$ requires $48N^2[M-(j-1)N] + 24N[M-(j-1)N]^2 + 54N^3$ flops. The Frobenius norm of each user introduces $4M[M-(j-1)N]$ flops.

Therefore, the computational complexity of the proposed algorithms is approximately

$$\begin{aligned}\varphi_P \leq \sum_{j=2}^{K_0}\{&48(j-1)^2N^2M + 24(j-1)NM^2 + 54(j-1)^3N^3\\&+ [8NM(M-(j-1)N) + 48N^2(M-(j-1)N)\\&+ 24N[M-(j-1)N]^2 + 4M(M-(j-1)N)\\&+ 54N^3 + 1] \times (K-j+1)\}\\&+ K(48N^2M + 24NM^2 + 54N^3 + 4MN + 1)\\&\approx o(M^2NK_0K).\end{aligned}$$

2) Capacity-based scheduling algorithm.

$$\begin{aligned}\varphi_C \leq & \sum_{j=2}^{K_0}\{[48(j-1)^2+48j]N^2M \\ & +[24(j-1)+32j]NM^2+[54(j-1)^3+54j]N^3 \\ & +2j^2N^2+8jN\}\times(K-j+1) \\ & +K(48N^2M+24NM^2+54N^3+2N^2+8N) \\ & \approx o(M^3K_0K). \end{aligned} \tag{13}$$

3) Max-min(max-max) eigenvalue scheduling algorithm.

$$\begin{aligned}\varphi_M \leq & \sum_{j=2}^{K_0}\{48(j-1)^2N^2M+24(j-1)NM^2 \\ & +54(j-1)^3N^3+[8NM(M-(j-1)N) \\ & +48N^2(M-(j-1)N)+24N[M-(j-1)N]^2 \\ & +54N^3]\times(K-j+1)\} \\ & +K(48N^2M+24NM^2+54N^3) \\ & \approx o(M^2NK_0K).\end{aligned}$$

4) Norm-based scheduling algorithm.

$$\begin{aligned}\varphi_N \leq & \sum_{j=2}^{K_0}\{48(j-1)^2N^2M+24(j-1)NM^2 \\ & +54(j-1)^3N^3+[8NM(M-(j-1)N) \\ & +4N(M-(j-1)N]\times(K-j+1)\}+4MNK \\ & \approx o(M^2NK_0K).\end{aligned}$$

Fig. 6 compares the complexity of various scheduling algorithms with respect to the total number of users in a system with six transmit antennas at BS and two receive antennas at each user. The maximum number of simultaneously supportable users is three. It can be observed that, as in the complexity analysis, all the scheduling algorithms have linear complexity with respect to the number of users in the system. Fig. 6 shows that different scheme has different slopes. The channel norm-based algorithm has the lowest complexity among all the schemes. The proposed MMS and MNS algorithms incur a complexity burden similar to that of max-max and max-min algorithms. The capacity-based algorithm incurs largest computational load for all the schemes.

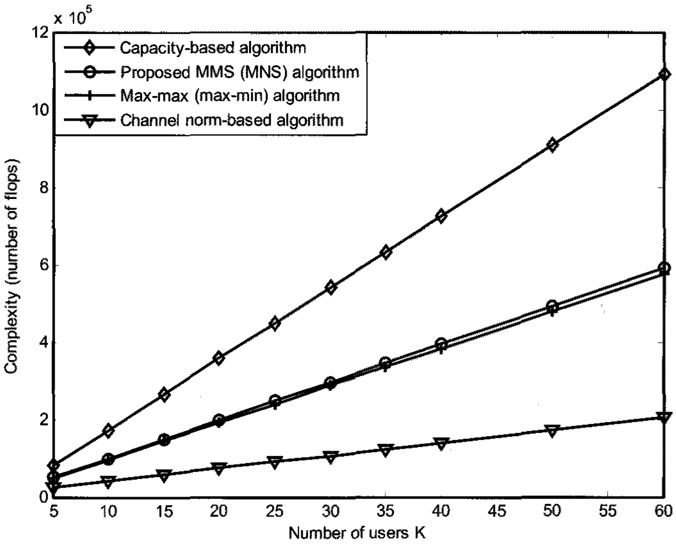


Fig. 6. Comparison of the complexities of various algorithms with different numbers of users, $M = 6$, $K_0 = 3$, and $N_k = N = 2$.

V. CONCLUSIONS

In this paper, we present two suboptimal user selection algorithms for a MU-MIMO downlink system using SZF-DPC at the transmitter. The objective is to obtain the optimal user set to maximize the system's sum rate capacity with low computational complexity. Simulation results show that the performance of the proposed MNS algorithm is superior to that of the previously proposed max-min algorithm and close to that of the capacity-based algorithm at high SNR but with significantly reduced complexity. The proposed MMS algorithm's performance is slightly better than that of the previously proposed max-max algorithm and very close to that of the capacity-based algorithm at low SNR with low complexity.

REFERENCES

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, pp. 439–441, May. 1983.

[2] P. Viswanath and D. Tse, " Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[3] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1691–1706, July 2003.

[4] Z. Tu and R. S. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Commun. lett.*, vol. 7, pp. 307–372, Aug. 2003.

[5] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 528–541, Mar. 2006.

[6] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, pp. 461–471, Feb. 2004.

[7] A. D. Dabbagh and D. J. Love, "Precoding for multiple antenna Gaussian broadcast channels with successive zero-forcing," *IEEE Trans. Signal Process.*, vol. 55, pp. 3837–3850, July 2007.

[8] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans,"Low complexity user selection algorithm for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, pp. 3658–3663, Sept. 2006.

[9] C.-J. Chen and L.-C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1435–1445, Sept. 2007.

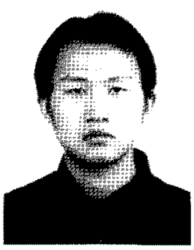
[10] Y. Xie and C. Georghiades, "Some results on the sum-rate capacity of MIMO fading broadcast channels," *IEEE Trans. Wireless Commun.*,vol. 5, pp. 377–383, Feb. 2006.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge University Press, 1985.

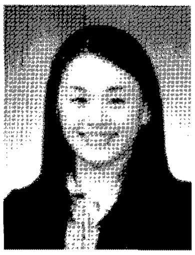
[12] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1277–1294, June 2002.

[13] S. Sigdel and W. A. Krzymien, "Efficient user selection and ordering algorithms for successive zero-forcing precoding for multiuser MIMO downlink," in *Proc. IEEE VTC*, Barcelona, Apr. 2009

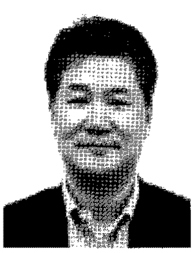
[14] S. Sigdel and W. A. Krzymien, "Simplified user scheduling and antenna selection algorithms for multiuser MIMO orthogonal space-division multiplexing downlink," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1329–1344, Mar. 2009.



Youxiang Wang received the B.S. degree in Microwave Engineering from University of Electronic Science and Technology of China, Chengdu, China, in 2002, and the M.S. and Ph.D. degrees in Information and Communication Engineering from Yeungnam University, Korea, in 2006 and 2010, respectively. He is currently with Research Institute of China Unicom. His research interests include multiple antenna wireless systems, as well as the standardization and implementation of 3GPP LTE and LTE-A.



Soojung Hur received the B.S. degree from Daegu University, Gyeongbuk, Korea, in 2001. She finished M.S. course in Electrical Engineering from San Diego State University of San Diego in 2004. And she received M.S. degree in Information and Communication Engineering from Yeungnam University, Korea in 2007. She is working toward her Ph.D. degree at the Mobile Communication lab, Yeungnam University, Korea. Her current research interests include performance of mobile communication and large area wireless communication techniques.



Yongwan Park received the B.S. and M.S. degrees in Electrical Engineering from Kyungpook University, Daegu, Korea, in 1982 and 1984, respectively, and his M.S. and Ph.D. degrees in Electrical Engineering from State University of New York (Buffalo), U.S.A., in 1989 and 1992, respectively. He joined the California Institute of Technology as a Research Fellow from 1992 to 1993. From 1994 to 1996, he has served as a Chief Researcher for developing IMT-2000 system at SK Telecom, Korea. Since September 1996, he has been a Professor of Information and Communication Engineering at Yeungnam University, Korea. From Jan. 2000 to Feb. 2000, he was a Invited Professor at NTT Mobile Communications Network Inc. (NTT DoCoMo) Wireless Lab, Japan. He was also a Visiting Professor at UC Irvine, U.S.A., in 2003. His current research areas of interest include Beyond 3G/4G system, OFDM system, PAR reduction, interference cancellation, and resource management in wireless communication etc.

Jeong-Hee Choi received the B.S. degree from Kyungbuk National University, Korea, in 1986. And she received M.S. and Ph.D. degrees in Electrical and Computer Engineering from State University of New York at Buffalo in 1989 and 1992, respectively. From 1994 to 1998, she was a Senior Researcher at SK Telecom central research center, Korea. She is currently a Professor of Division of Computer and Communication Engineering at Daegu University, Korea. Her research interests are in the area of radar image processing, inverse scattering, and terrestrial/satellite mobile communications.