

Joint queue control and user scheduling in MIMO broadcast channel under zero-forcing multiplexing

Feng She¹, Wen Chen^{2,*},[†], Hanwen Luo² and Damao Yang²

¹*College of Electronic & Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China*

²*Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, People's Republic of China*

SUMMARY

This paper studies the problem of queue control and user scheduling in multi-antenna broadcast (downlink) systems under zero forcing beamforming transmit strategy. In the system, we assume that the data packet arrives randomly to the buffered transmitter. By taking the broadcast channel as a controlled queueing system, we deduce the property of queue control function that maximizes the weighted system throughput while guarantees the delay fairness among users. We also present a low-complexity user selection algorithm with the consideration of queue state and channel state together. Simulation results show that the joint queue control and user selection policy can achieve considerable fairness and stability among users. Copyright © 2009 John Wiley & Sons, Ltd.

Received 19 September 2007; Revised 3 April 2009; Accepted 4 June 2009

KEY WORDS: MIMO; broadcast channel; downlink scheduling; spatial multiplex

1. INTRODUCTION

Multiple-input multiple-output (MIMO) system is well motivated for wireless communications through fading channels because of the potential improvement in transmit rate or diversity gain [1]. It is well known that multiple antennas can be easily deployed at base station in cellular systems. However, mobile terminals usually have a small number of antennas due to the size and cost constraint. Thus, it may appear that we cannot obtain significant capacity benefit from the multiple transmit antennas. This is true with the transmit strategy of time division multiple access [2].

*Correspondence to: Wen Chen, Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, People's Republic of China.

[†]E-mail: wenchen@sjtu.edu.cn

Contract/grant sponsor: NSF China; contract/grant numbers: #60572157, #60672067

Contract/grant sponsor: NSF Shanghai; contract/grant number: 062R1401

Contract/grant sponsor: Shanghai-Canada NRC; contract/grant number: #065No7112

Contract/grant sponsor: Cultivation Fund of the Key Scientific and Technical Innovation Project; contract/grant number: #706022

To solve the problem, multiuser must be served simultaneously. One way to accomplish this is called dirty paper coding (DPC), which is a multiuser encoding strategy based on interference presubtraction [3]. But DPC is of high complexity, and hence a much simple transmit strategy, zero forcing beamforming (ZFBF) technique has been proposed for space division multiple access to remove the cochannel interference in MIMO downlink systems [4–7].

In ZFBF multiplexing, an issue that has not been treated thoroughly in the literature is how to control data queue in transmitter. This is the problem of resource allocation (such as power and bandwidth) and user scheduling in order to perform desirably with respect to the criteria such as fairness or throughput [8]. This problem has attracted great interest in the recent years [9–11].

Under random packet arrival, the notation of ‘fairness’ is replaced by the notation of ‘stability’ [8, 12, 13]. The system goal is to stabilize the K transmission queues whenever the arrival rate is inside the stability region in the system.

User scheduling algorithms proposed in the literatures most commonly ignore queuing and randomness in packet arrivals and hence cannot offer stability guarantee. This is true in some scheduling algorithms that aim to satisfy a fairness criteria, such as proportional fair scheduling [2].

A guiding work that incorporates randomness and stability issues has been presented in [13], where the network capacity region is defined as the region of the input data rates making system stable. It is shown that this region is achievable by a maximum weight matching (weights related to queue length). Based on those definitions, Neely *et al.* [12] consider a broadcast scenario under time division and demonstrate a schedule algorithm that achieves the system capacity region. Similarly, Vishwanathan and Kumaran [14] show that a throughput optimal policy is a maximum weight matching in the form of $\max \sum_i \alpha_i q_i r_i$, where q_i and r_i are the queue length and service rate of the i th user, respectively, α_i is a nonnegative constant. In addition, in broadcast system, Airy *et al.* [15] compare several user scheduling policies.

In this paper we formulate the general broadcast scheduling problem, where packet arrivals into the users queues are random processes. It is established that a throughput optimal schedule in the downlink channel is the one that chooses a group of users and corresponding rates from the DPC region to maximize $g(q) \cdot r$, where $g(q)$ is a function of queue length q , and r is transmit rate. However, this solution is of high complexity, not only due to the complexity of obtaining the DPC region but also due to searching for an ordered subset of users out of total users. Therefore, ZFBF is an attractive technique in multiplexing scenario as it does not incur the user ordering problem and has lower coding complexity. Thus, in this paper we consider ZFBF strategy that demonstrates good performance with low complexity. The contributions of this paper are as follows

1. We give the property of the queue control function which can guarantee queue equal among the active users.
2. We have compared the fairness performance by using different queue control functions.
3. We give the upper bound of the weighted throughput with queue control function.
4. We propose a low-complexity user scheduling algorithm under ZFBF multiplexing.

This paper is organized as follows. We outline the system model and the transmit strategy of ZFBF in Section 2. In Section 3, the queue and power control policy are presented. We analyze the scheduling algorithm under ZFBF multiplexing in Section 4. The simulation results are presented in Section 5. Finally, we conclude this paper in Section 6.

Notations used in this paper are as follows: $(\cdot)^T$ denotes matrix transpose, $(\cdot)^H$ denotes matrix conjugate-transpose, $E[\cdot]$ denotes statistical expectation, $(\cdot)^\dagger$ is pseudoinverse matrix and $\text{tr}(\cdot)$ is matrix trace.

2. SYSTEM MODEL AND MULTIUSER TRANSMIT STRATEGIES

2.1. Multi-user broadcast channel model

We consider a single cell MIMO BC system with a single base station supporting data traffic to K users. As shown in Figure 1, the base station is with N_t transmit antennas and each of the user terminal has a single receive antenna. We assume $K \geq N_t$. For simplicity, we assume that all the users experience independent fading. We use a simple channel model where the channel gain from a transmit antenna to a user is described by a zero-mean circularly symmetric complex Gaussian random variable [2]. Here, we assume that $\mathbf{h}_k \in \mathbb{C}^{1 \times N_t}$ is the multiple-input single-output (MISO) channel gain matrix to the k th user. Thus, the signal received by the user k is given by

$$\mathbf{y}_k = \mathbf{h}_k \mathbf{x} + \mathbf{n}_k, \quad k = 1, \dots, K \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the transmit signal vector with a power constraint $\text{tr}(E[\mathbf{x}\mathbf{x}^H]) \leq P$, and \mathbf{n}_k is the complex Gaussian noise with unit variance per vector component (i.e. $E[\mathbf{n}\mathbf{n}^H] = \mathbf{I}$).

At transmitter, we employ the ZFBF transmit strategy. In ZFBF, the scheduler first selects an active user set $S \subset 1, 2, \dots, K$, where $|S| \leq N_t$. Then, the transmitter assigns different beamforming direction to each data stream in such a way that the interference at each receiver is completely suppressed.

Denote $h_i, i = 1, \dots, |S|$ as the channel to the i th active user and define $H(S) = [h_1^T, \dots, h_{|S|}^T]$ as the channel matrix of active user set. The transmit signal is expressed as

$$\mathbf{x} = \sum_{i=1}^{|S|} \sqrt{P_i} w_i s_i \quad (2)$$

where s_i , w_i , and P_i are data symbol, beamforming vector's entry, and transmit power for the i th active user, respectively. Then the received signal at the i th active user is given by

$$y_i = \sqrt{P_i} h_i w_i s_i + \sum_{j=1, j \neq i}^{|S|} \sqrt{P_j} h_j w_j s_j + n_j \quad (3)$$

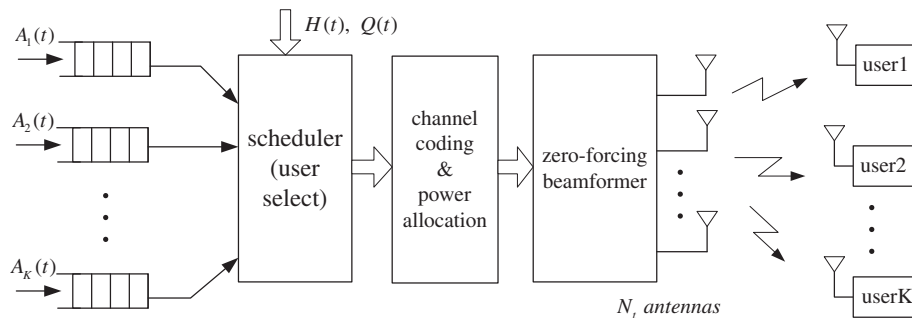


Figure 1. Block diagram of multi-antenna broadcast system with queue control.

and the sum rate achieved by ZFBF is [4]

$$R_{\text{BF}} = \max_{w_k, P_k} \log \left(\frac{1 + \sum_{j=1}^{|S|} P_j |h_k w_j|^2}{1 + \sum_{j=1, j \neq k}^{|S|} P_j |h_k w_j|^2} \right) \quad (4)$$

subject to

$$\sum_{k=1}^K \|w_k\|^2 P_k \leq P \quad (5)$$

In this paper, we assume that the base station has perfect channel state information (CSI) of all the downlink channels, while each of the users only has the CSI of its own downlink channel and does not know the CSI of downlink channel of other users.

Figure 1 also depicts a cellular network with K flows transmitter buffers in the base station. In Figure 1, $H(t)$, $Q(t)$, and $A_i(t)$ are CSI, queue state information and data arrival rate of user i , respectively. We assume that time is slotted and let $q_k(t)$ denote the size of the k th queue at the beginning of the time slot t , $a_k(t)$ denotes the number of arrivals to queue i in time slot t , and $r_k(t)$ denote the amount of service rate offered to queue k in slot t . We assume that each of these parameters can only be nonnegative. Then the evolution of the size of the k th queue is given by

$$q_k(t+1) = (q_k(t) + a_k(t) - r_k(t))^+ \quad (6)$$

where $(y)^+ = \max\{y, 0\}$, and $k = 1, \dots, K$.

2.2. Multi-antenna ZFBF

In this subsection, ZFBF in multi-antenna broadcast channel (BC) is introduced. It is well known that multiple transmit antennas can potentially yield an N_t -fold increase in the sum capacity, where N_t is the number of transmit antennas. Yoo and Goldsmith [2] show that employing ZFBF to a set of N_t nearly orthogonal users with large channel norms is asymptotically optimal as the number of users grows large.

In ZFBF, the beamforming matrix takes a pseudoinverse of the channel matrix. By this way, a respective data stream of each transmit antenna can be decoupled and thus sent to its desired user without interfering with other users. In multiuser MISO case, we first select a user subset S to be served together, where subset size $|S| \leq N_t$, and then build up the corresponding channel matrix $H(S)$ of the user set. The beamforming matrix $\mathcal{W}(S)$ is written as

$$\mathcal{W}(S) = H(S)^H (H(S)H(S)^H)^{-1} \quad (7)$$

As a result, the achievable throughput of ZFBF for a given user set S is given by

$$R_{\text{ZFBF}}(S) = \max_{\sum_{i \in S} b_i^{-1} P_i \leq P} \sum_{i \in S} \log_2(1 + P_i) \quad (8)$$

where

$$b_i = \frac{1}{[(H(S)H(S)^H)^{-1}]_{i,i}} \quad (9)$$

is the effective channel gain of the i th user and the power P_i is obtained by waterfilling.

Note that the b_i in (9) has an important geometric interpretation as noted in [4], i.e. the distance squared of the i th user channel from the span of every other user channels in the activation set.

3. RESOURCE ALLOCATION AND STABILITY

3.1. Resource allocation with random arrival

All the previously investigated transmission schemes can incorporate the adaptive max-stability policy of [8, 12]. Then the policy is to maximize the weighted sum of instantaneous rates $\max \sum_{k \in S} g(q_k(t))r_k(t)$, where $g(q_k(t))$ is the function of queue length and S is the active user set in time slot t .

Therefore, we shall investigate the algorithms for the weighted-sum-rate maximization under ZF strategy, which is the most attractive scheme in practice. Packets for the i th source are stored in the i th buffer until they are served. The transmission power $P_i(t)$ and rate $r_i(t)$ used by transmitter i at time t are to be dynamically allocated so as to optimize throughput and delay.

We now explicitly pose the dynamic resource allocation problem. For each i , $i = 1, \dots, K$, let $q_i(t)$ be the number of untransmitted bits (unfinished work) in queue i at time t . Consider a stationary controller which at any time channel state $H(t)$ and queue state $q(t)$ outputs power allocation $P(t)$ and rate allocation $r(t)$. The controller first chooses a power control policy

$$P = \mathcal{P}(H, q) \quad (10)$$

Here, $P_i = \mathcal{P}_i(H, q)$ is the power allocated to transmitter i in response to fading state H and queue state q . Next, the controller chooses a rate allocation policy \mathcal{R} which satisfies the power constraint. For a given \mathcal{P} , the controller is allowed to allocate any rate in the BC region induced by the power policy \mathcal{P} . Then the rate control policy is

$$r = \mathcal{R}(H, q) \in C_{BC}(H, P) \quad (11)$$

where $C_{BC}(h, p)$ is given by (24).

3.2. Stability and throughput optimization

ZFBF inverts the channel matrix at the transmitter so that orthogonal channels between transmitter and receivers are created. It is then possible to encode users individually, as opposed to the more complex long-block-vector coding generally needed to implement DPC.

In practical wireless communication systems, large queue lengths will either result in large delays or lead to large amount of packet losses in the system when the buffer size is limited. Both will negatively affect the quality of service experienced by users. Our goal is to achieve fairness operation that stabilizes the system. In other words, our work is to keep equal queue length among active users even though each user has different arrive rates.

The queue and power control problem is given by

$$\begin{aligned} \max \quad & \sum_{i \in S} g(q_i) \cdot r_i \\ \text{s.t.} \quad & \sum_{i \in S} P_i \leq P, \quad q_1 \leq q_2 \leq \dots \leq q_n \end{aligned} \quad (12)$$

where n is user set size.

Proposition 1

Let $g(q)$ be the function of queue length q . Then to reach equal queue length among active users, the second-order deviation of queue function should be

$$\frac{d^2}{dq_i^2}g(q_i)=0$$

where q_i is the queue length variable of the i th user.

Proof

From (12), the Lagrangian is

$$L(q, r, \mu, v) = \sum_{i \in S} g(q_i) \cdot r_i + \mu \left(\sum_{i \in S} P_i - P \right) + \sum_{i \in S} v_i (q_i - q_{i+1}) \quad (13)$$

Then its subgradient is given by

$$\frac{\partial}{\partial q_i} L(q, r, \mu, v) = \frac{d}{dq_i} g(q_i) \cdot r_i + (v_i - v_{i-1}) = 0 \quad (14)$$

We define $q_{n+1}=0$ and $v_0=0$, then

$$\begin{aligned} \sum_{i \in S} v_i (q_i - q_{i+1}) &= \sum_{i \in S} (v_i q_i - v_i q_{i+1}) \\ &= v_1 q_1 - v_1 q_2 + v_2 q_2 - v_2 q_3 + \cdots + v_i q_i - v_i q_{i+1} + \cdots + v_n q_n - v_n q_{n+1} \\ &= q_1 (v_1 - v_0) + q_2 (v_2 - v_1) + \cdots + q_i (v_i - v_{i-1}) + \cdots + q_n (v_n - v_{n-1}) \\ &= \sum_{i \in S} (v_i - v_{i-1}) q_i \end{aligned} \quad (15)$$

Equation (14) implies that

$$(v_i - v_{i-1}) = - \left(\frac{d}{dq_i} g(q_i) \right) \cdot r_i \quad (16)$$

Equations (13), (15), and (16) imply that

$$L(q, r, \mu, v) = \sum_{i \in S} g(q_i) \cdot r_i + \mu \left(\sum_{i \in S} P_i - P \right) - \sum_{i \in S} \frac{d}{dq_i} g(q_i) \cdot r_i \cdot q_i \quad (17)$$

Then, its subgradient is given by

$$\begin{aligned} \frac{\partial}{\partial q_i} L(q, r, \mu, v) &= \left(\frac{d}{dq_i} g(q_i) \right) r_i - \frac{d}{dq_i} \left(\frac{d}{dq_i} g(q_i) r_i q_i \right) \\ &= \frac{dg(q_i)}{dq_i} r_i - \frac{d^2 g(q_i)}{dq_i^2} r_i q_i - \frac{dg(q_i)}{dq_i} r_i \\ &= - \frac{d^2 g(q_i)}{dq_i^2} r_i q_i = 0 \end{aligned} \quad (18)$$

because $r_i q_i \neq 0$, thus

$$\frac{d^2}{dq_i^2} g(q_i) = 0 \quad (19)$$

□

Thus, (19) gives us sense to define a queue function $g(q)$ to guarantee the equal size for the active user subset.

While in real wireless network, there will be $n \ll K$, where n is the number of active users in one time slot and K is the number of total users in the system. Because the above optimal queue control function can only make the queue length of active user set to reach equal in one time slot, thus, Proposition 1 cannot guarantee the fairness among all users. Thus, to achieve better fairness among users, it is necessary to do some adjustment on the queue function $g(q)$ presented in Proposition 1. Here, we define $g(q) = q^m$, $m \geq 0$ as a queue function. Then the system optimization problem become

$$\max q^m(t) \cdot r(t) \quad (20)$$

Note that when $m = 0$, (20) is the throughput maximization problem without fairness consideration and a classic stability problem when $m = 1$. Because $m = 0$ and 1 are two special cases, we take a more general case q^m in user scheduling, and select proper m for desired performance.

4. JOINT SCHEDULING POLICY OF ZFBF MULTIPLEXING

In this section, we first analyze the upper bound of weighted throughput with queue function, and then present a low-complexity user scheduling algorithm.

4.1. Upper bound of weighted throughput

Define $f(q, r) = \sum_{i \in S} g(q_i) r_i$, our goal is to

$$\max f(q, r) = \max \sum_{i \in S} g(q_i) r_i \quad (21)$$

where S is the active user set. Similar to the analysis in [16], we can get the following theorem.

Theorem 1

With a total power constraint, we have

$$\max f(q, r) = \max \sum_{i \in S} g(q_i) r_i \leq \max \left(\sum_{i \in S} g(q_i) \right) \left(\log_2 \left(1 + \frac{P}{\text{tr}(W_S^{-1})} \right) + D(g(q_{Si}) \| b_{Si}) \right) \quad (22)$$

where $W_S = H(S)H(S)^H$,

$$b_i = \frac{1}{(W_S^{-1})_{i,i}}, \quad \sum_{i \in S} P_i / b_i \leq P, \quad g(q_{Si}) = \frac{g(q_i)}{\sum_{i \in S} g(q_i)}, \quad b_{Si} = \frac{b_i}{\sum_{i \in S} b_i}$$

and $D(\cdot \| \cdot)$ denotes the Kullback Leibler distance.

Proof

Under ZFBF, our problem becomes,

$$f(q, r) = \max_{\sum_i P_i / b_i \leq P} \sum_{i \in S} g(q_i) \log_2(1 + P_i) \quad (23)$$

Therefore, the Lagrangian is:

$$L(q, p, \lambda) = \sum_{i \in S} g(q_i) \log_2(1 + P_i) - \frac{1}{\lambda} \left(\sum_{i \in S} P_i / b_i - P \right) \quad (24)$$

Then, we get

$$\frac{\partial L}{\partial P_i} = g(q_i) \frac{1}{1 + P_i} - \frac{1}{\lambda b_i} = 0 \quad (25)$$

Then we have

$$P_i = \lambda g(q_i) b_i - 1 \quad \text{where} \quad \sum_{i \in S} \left(\lambda g(q_i) - \frac{1}{b_i} \right)^+ = P \quad (26)$$

Noting that

$$\lambda = \frac{P + \sum_{i \in S} \frac{1}{b_i}}{\sum_{i \in S} g(q_i)}$$

we get

$$f(q, S) = \sum_{i \in S} g(q_i) \log_2(1 + P_i) = \sum_{i \in S} g(q_i) \log_2(b_i g(q_i)) \quad (27)$$

By the cofactor expansion for the inverse, we have,

$$b_i = \frac{\det(H(S))H(S)^\dagger}{\det(M_{i,i}(H(S)H(S)^\dagger))} \triangleq \frac{A(S)}{M_{i,i}(S)} \quad (28)$$

Returning to (27), we get

$$\begin{aligned} f(q, S) &= \sum_{i \in S} g(q_i) \log_2(1 + P_i) \\ &= \left(\sum_{i \in S} g(q_i) \right) \log_2(\lambda) + \sum_{i \in S} g(q_i) \log_2(g(q_i) b_i) \\ &\leq \left(\sum_{i \in S} g(q_i) \right) \left[\log_2 \left(\frac{P + \frac{1}{A(S)} \sum_{i \in S} M_{i,i}(S)}{\sum_{i \in S} g(q_i)} \right) + \sum_{i \in S} g(q_i) \log_2(g(q_i) \frac{A(S)}{M_{i,i}(S)}) \right] \\ &= \left(\sum_{i \in S} g(q_i) \right) \left[\log_2 \left(\frac{P A(S)}{\sum_{i \in S} M_{i,i}(S)} + 1 \right) + \sum_{i \in S} g(q_i) \log_2 \left(\frac{g(q_i)}{b_i} \right) \right] \\ &= \left(\sum_{i \in S} g(q_i) \right) \left[\log_2 \left(\frac{P}{\sum_{i \in S} \frac{1}{b_i}} + 1 \right) + D(g(q_i) \| b_i) \right] \end{aligned} \quad (29)$$

□

This theorem provides significant insight into user selection for ZFBF multiplexing for maximize the weighted throughput. Particularly, (29) may be broken up into three main parts, i.e. the queuing gain $\sum_{i \in S} g(q_i)$, the geometry gain,

$$\log_2 \left(\frac{P}{\sum_{i \in S} \frac{1}{b_i}} + 1 \right)$$

and the pairing gain $D(g(q_{Si}) \| b_{Si})$ [16]. Here, we consider the queue control and access control in multi-user scheduling, which is more practical in communication systems. The work that considers queue control in multi-user scheduling extend the work in [16], and present new insight into multiuser communication. Since the queuing gain captures the significance of the magnitude of the queue state, the geometry gain captures the effects of the channels, the two parts are main properties in (29).

4.2. Selecting semi-orthogonal user set with waterfilling

For a given set of channel gains and corresponding queue length, the weighted sum rate is maximized by an orthogonal user set. As mentioned above, such a set of user has the largest conditional ‘geometry gain’. In addition, since the channel matrix is orthogonal, zero-forcing incurs no power penalty for inverting the channel and thus can approach the DPC rate region [2].

While in practice, for finite user number K , the probability of existence of an orthogonal set is zero. Thus, we consider ‘nearly’ orthogonal user sets. To be precise, define two vectors v_1 and v_2 to be α -orthogonal if

$$\frac{|v_1 v_2^H|}{\|v_1\| \cdot \|v_2\|} \leq \alpha \quad (30)$$

Here, we construct a low-complexity user group by combing the semi-orthogonal user selection and waterfilling (SUSWF) algorithm, as outlined next.

(1) Initialization:

$$S_1 = \{1, \dots, K\}, \quad i = 1, \quad S_0 = \phi(\text{empty set}) \quad (31)$$

(2) For each user $k \in S_i$, calculate g_k , the component of H_k orthogonal to the space spanned by $\{g_1, \dots, g_{i-1}\}$

$$g_k = H_k - \sum_{j=1}^{i-1} \frac{H_k g_j^H}{\|g_j\|^2} g_j \quad (32)$$

When $i = 1$, let $g_k = H_k$.

(3) Select the i th user as follows:

$$s_i = \arg \max_k g(q_k) \log_2 \left(1 + \frac{P}{N_t} \|g_k\|^2 \right) \quad (33)$$

$$S_0 = S_0 \cup \{s_i\}, \quad h_i = h_{s_i}, \quad g_i = g_{s_i} \quad (34)$$

(4) Do waterfilling among selected users and reject the users that are under the waterlevel in this step [17].

- (5) If $|S_0| < N_t$, then calculate S_{i+1} , which is the set of unselected users which is α -semiorthogonal to g_i

$$S_{n+1} = \left\{ k \in S_n, k \neq s_i \mid \frac{|h_k g_i^H|}{\|h_k\| \|g_i\|} < \alpha \right\} \quad (35)$$

- (5) Increase i by 1.

where α is a small positive constant that is defined as (30). If S_{n+1} is nonempty and $|S_0| < N_t$, then go to step 2. Otherwise, the algorithm is finished.

User scheduling algorithm based on semi-orthogonality has been proposed in [18, 19]. A similar algorithm to the SUS has been proposed in [2]. While in [2], stability and queue control are not considered. Furthermore, we added waterfilling process in user selection. Dimić and Sidiropoulos [17] point out that in the algorithm of user selection, the water level will decrease in the next iteration. Thus, the users rejected in this iteration will be rejected in the next iteration. Therefore, the advantage of step 4 is that it can reduce the complexity (running time) by eliminating those weaker users in this subsequent iteration.

5. SIMULATION RESULT AND DISCUSSION

In this section, numerical results for the queue control of weighted SUSWF are presented. The considered multiuser MIMO system in BC with the number of transmit antenna is $N_t = 4$ and each user has a single receive antenna. We assume that the power in the simulation is $P = 15$ dB. With spatial multiplexing, the number of data streams does not exceed the number of transmit antennas. We also assume that the channels between different transmit and receive antennas are independent.

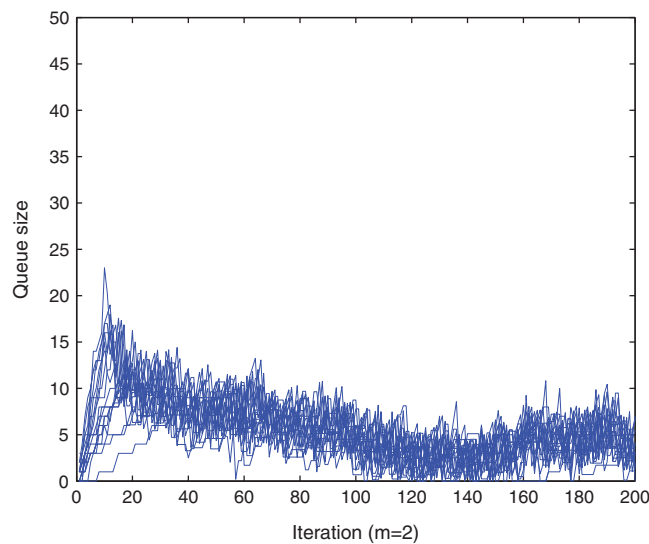


Figure 2. The queue length in multi-antenna broadcast channel with $m = 2$.

To be simple, we define a special queue function $g(q) = q^m$ and consider 20 users in simulation. We assume that at least one user queues are served in one time slot. In this scenario, the scheduling algorithm is equivalent to serving the users that solve $f(t) = \arg \max_{i \in S} \sum q^m(t) r(t)$. We consider each queue with mutually independent arrival processes such that $A_k(t)$ is with rate λ_k (bits/channel use), and let the rate vectors, $(\lambda_1(t), \dots, \lambda_K(t))$ be randomly created between $[0, 1]$. We also assume that the number of arrivals in each slot to each queue is Poissian distributed.

A. Experiment 1: The first experiment tests the average queue length with $m = 2$ in queue control function q^m . In this experiment, the queue control carry out from the 15th iteration. Figure 2 shows that by queue controlling function, the system can become stable after few times of iteration.

B. Experiment 2: This experiment is about the fairness among multi-users. Here, we use the variance value of the multi-user queue length to denote the fairness. In other words, large variance value implies poor fairness while small variance value denotes good fairness performance among users. In this experiment, we calculate the normalized variance value with different m . Figure 3 shows that the variance decreases with the increment of m . Thus, to get the fairness among users, the selected m should not be very small.

C. Experiment 3: The third experiment is about the average sum throughput of different m . Owing to the variance value being very large when $m < 1$, which is shown in the above experiment, we only consider the scenario of $m \geq 1$ in this experiment. By 10 000 times of iteration, Figure 4 shows that the average throughput is almost the same, which implies that $m \geq 1$ is a good option. The unit of unit of throughput in Figure 4 is bits/Hz/s.

D. Experiment 4: This experiment tests the average queue length with different m . We only consider $m \geq 1$ as the above experiment. We calculate the average queue size by 10 000 times of iteration. Figure 5 shows that the sum queue length increases slightly with the increment of m .

E. Experiment 5: The last experiment is about the stability of the system with queue control. In this experiment, we test the input rate and the output rate, and we also test the queue size of each user. We carry queue control from the 20th iteration. From Figure 6, we can see that the system with queue control is stable, and the input rate and output rate also show the equilibrium.

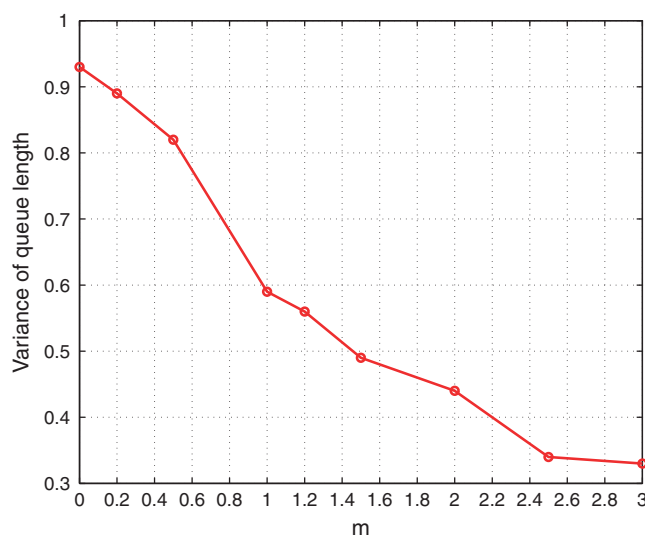


Figure 3. Variance of queue length with different m .

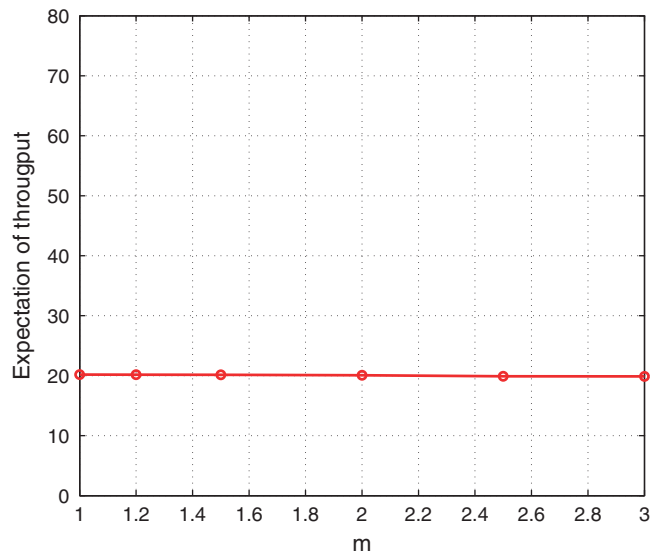


Figure 4. The expectation of throughput with different m .

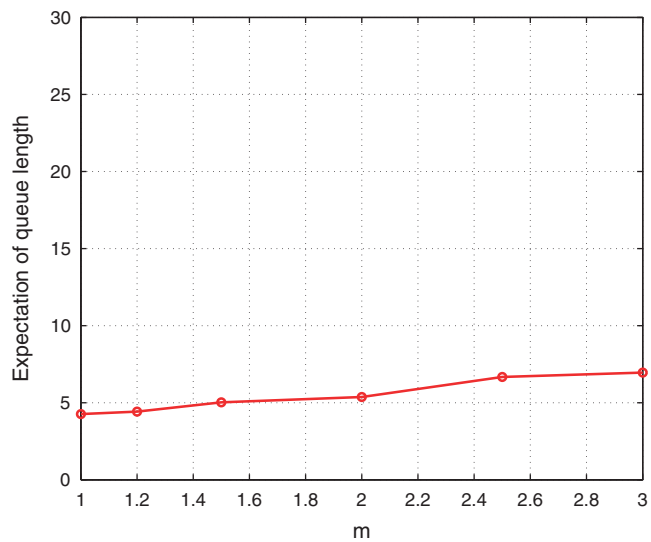


Figure 5. The expectation of queue length with different m .

6. CONCLUSION

In this paper, we have presented a framework of joint queue control and user scheduling. The function of queue length should be correctly selected. Based on the queue length function, we

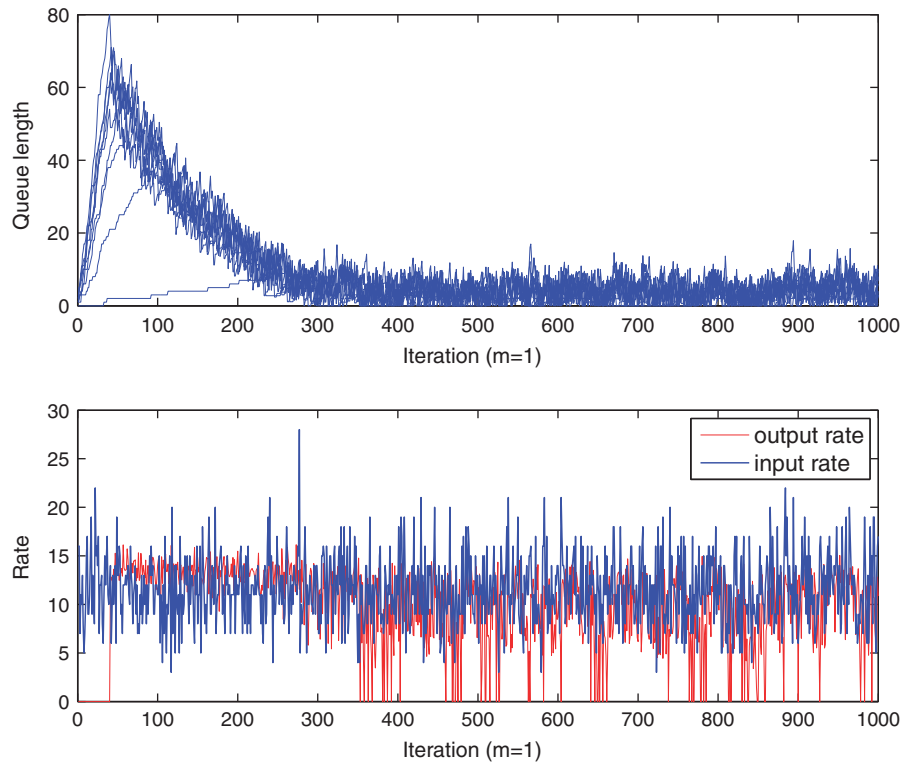


Figure 6. The input–output throughput and queue length with $m = 1$.

derive the upper bound of the weighted throughput. By jointly considering channel gains and queue states, we present a low-complexity user selection algorithm which is named as SUSWF.

In simulation experiments, we select a special queue control function as $g(q) = q^m$ in simulation, and the simulation results show that queue control policy is critical in fairness user scheduling. In this paper, we take a special case of queue control function, while the design of the queue function is still very challenging and interesting for future works.

ACKNOWLEDGEMENTS

The authors would like to thank to Da Mao Yang and Hong Xing Li for their helpful discussion. This work is supported by NSF China #60572157, #60672067, NSF Shanghai 062R1401, Shanghai-Canada NRC #065No7112, Cultivation Fund of the Key Scientific and Technical Innovation Project #706022.

REFERENCES

1. Alamouti SM. A simple diversity techniques for wireless communications. *IEEE Journal on Selected Areas in Communications* 1998; **16**:1451–1458.
2. Yoo T, Goldsmith A. On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming. *IEEE Journal on Selected Areas in Communications* 2006; **24**(3):528–541.

3. Costa MHM. Writing on dirty paper. *IEEE Transactions on Information Theory* 1983; **29**(3):439–441.
4. Caire G, Shamai S. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Transactions on Information Theory* 2003; **49**(7):1691–1706.
5. Jindal N, Goldsmith A. Dirty-paper coding versus TDMA for MIMO broadcast channels. *IEEE Transactions on Information Theory* 2005; **51**(5):1783–1794.
6. Spencer QH, Swindlehurst AL, Haardt M. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Transactions on Signal Processing* 2004; **52**(2):461–471.
7. Spencer QH, Peel CB, Swindlehurst AL, Haardt M. An introduction to the multi-user MIMO downlink. *IEEE Communications Magazine* 2004; **42**(10):60–67.
8. Yeh EM, Cohen AS. Information theory, queueing, and resource allocation in multi-user fading communications. *Proceedings of the 2004 CISS*. Princeton: NJ, 2004.
9. Andrews M, Kumaran K, Ramannan K, Stolyar A, Whiting P. Providing quality of service over a shared wireless link. *IEEE Communications Magazine* 2001; **39**(2):150–154.
10. Viswanath P, Tse D, Laroia R. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory* 2002; **48**(6):1277–1294.
11. Weingarten H, Steinberg Y, Shamai S. The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Transactions on Information Theory* 2006; **52**(9):3936–3964.
12. Neely M, Modiano E, Rohrs C. Power allocation and routing in multibeam satellites with time-varying channels. *IEEE/ACM Transactions on Networking* 2003; **11**(1):138–152.
13. Tassiulas L, Ephremides A. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control* 1992; **37**(12):1936–1948.
14. Vishwanathan H, Kumaran K. Rate scheduling in multiple antenna downlink wireless systems. *IEEE Transactions on Communications* 2005; **53**(4):645–655.
15. Airy M, Shakkottai S, Heath JRW. Limiting queueing models for scheduling in multi-user mimo wireless systems. *Proceedings of the 2nd IASTED Conference on Communications, Internet and Info Technology*, Scottsdale, AZ, 2003.
16. Swannack C, Uysal-Biyikoglu E, Wornell GW. Low complexity scheduling for maximizing throughput in the MIMO broadcast channel. *Proceedings of the 42nd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2004.
17. Dimić G, Sidiropoulos ND. On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm. *IEEE Transactions on Signal Processing* 2005; **53**(10):3857–3868.
18. Tu Z, Blum R. Multiuser diversity for a dirty paper approach. *IEEE Communications Letters* 2003; **7**(8):370–372.
19. Zhang R, Liang YC, Cioffi J. Throughput comparison of wireless downlink transmission schemes with multiple antennas. *Proceedings of the IEEE International Conference on Communications*, Seoul, Korea, vol. 4, 16–20 May 2005; 2700–2704.

AUTHORS' BIOGRAPHIES



Feng She received the BS and MS from Central South University and Tongji University respectively, and received the PhD from Shanghai Jiaotong University, Shanghai, China in 2008. He was with Huawei Technologies as a researcher from 2008 to 2009. Then he joined Shanghai University of Engineering Sciences, China, as an assistant professor in the Department of Electronic Engineering. Dr She's research interests include wireless communication and networking for multiple antenna system, with current focus on MIMO communication systems, multiuser communication theory, cross-layer design and optimization, and interference management.



Wen Chen received the BS and MS from Wuhan University, and received the PhD from the University of Electro-Communications, Tokyo, Japan in 1999. He was a researcher of Japan Society for the Promotion of Sciences (JSPS) from 1999 through 2001. Then he joined the University of Alberta, Canada, starting as a post-doctoral fellow in Information Research Lab and continuing as a research associate in the Department of Electrical and Computer Engineering. Since 2006, he has been a full professor in the Department of Electronic Engineering, Shanghai Jiaotong University, China, where he is also the director of Institute for Signal Processing and Systems. Dr Chen was awarded the Ariyama Memorial Research Prize in 1997, the PIMS Post-Doctoral Fellowship in 2001. He received the honors of 'New Century Excellent Young Researcher in China' in 2006 and 'the Pujiang Excellent Investigator in Shanghai' in 2007. He was elected the vice general secretary of Shanghai Institute of Electronics in 2008. He is on the

editorial board of the International Journal of Wireless Communications and Networking, and serves Journal of Communications, Journal of Computers and EURASIP Journal on Wireless Communications and Networking as the guest editor. He is the Technical Program Committee chair for IEEE-ICCSC2008 and the General Conference Chair for IEEE-ICIS2009. He has published more than 60 papers in IEEE and IEICE journals and conferences. His interests include cooperative communications and network coding.



Hanwen Luo received the BS from Shanghai Jiaotong University. He is a full professor in the Department of Electronic Engineering, Shanghai Jiaotong University, China. He has published more than 20 papers in IEEE and IEICE journals and conferences. His interests include wireless communications and networks.



Damao Yang received the BS from Donghua University, Shanghai, China in 2006. Now he is a post-graduate student in the Department of Electronic Engineering, Shanghai Jiaotong University, China. His interests include wireless communications and networks.