

# Opportunistic Beamforming and Scheduling for OFDMA Systems

Patrick Svedman, *Student Member, IEEE*, Sarah Kate Wilson, *Senior Member, IEEE*, Leonard J. Cimini, Jr., *Fellow, IEEE*, and Björn Ottersten, *Fellow, IEEE*

**Abstract**—Orthogonal frequency-division multiple access (OFDMA) is an attractive technique for exploiting multiuser diversity in the downlink of a cellular system. This paper addresses three problems in multiuser diversity for OFDMA systems. First, we propose a way to significantly reduce the amount of channel state information (CSI) feedback without sacrificing performance too much, by selective and adaptive feedback. Second, we propose a way to increase the cell throughput and fairness by applying an opportunistic beamforming scheme to orthogonal frequency-division multiplexing. This beamforming scheme increases the frequency fading rate, which increases the multiuser diversity effect. Thirdly, we deal with the issue of fairness and quality-of-service (QoS) in opportunistic systems by proposing a modified proportional fair (PF) scheduler for OFDMA. Key features in the scheduler are that it incorporates QoS classes into the PF scheduler and that it has a tunable fairness level. Extensive simulation results are presented to evaluate the performance of the proposed schemes. The opportunistic beamforming scheme performed well in comparison with several other schemes. The modified PF scheduler was able to give users different QoS, based on their requirements, while still exploiting multiuser diversity.

**Index Terms**—Multiple antennas, multiuser diversity, OFDM, scheduling, wireless system design.

## I. INTRODUCTION

OPPORTUNISTIC systems use adaptive modulation instead of power control to achieve the target error rates. By scheduling users with good instantaneous channel conditions, exploiting multiuser diversity, high system throughput can be achieved [1]. In the downlink of an opportunistic system with frequency-selective channels, orthogonal frequency-division multiple access (OFDMA) is suitable because users can be scheduled on orthogonal frequency bands. This enables the exploitation of multiuser diversity in the frequency domain, i.e., users can be scheduled also on their *frequency* fading peaks [2], [3]. In this paper, we deal with some of the problems with opportunistic OFDMA. We propose an adaptive reduced-feedback scheme to cope with the significant amount of channel state

information (CSI) feedback required in a frequency-division duplexing (FDD) opportunistic OFDMA system. Furthermore, we propose an opportunistic beamforming scheme for OFDMA in order to increase the cell throughput and increase fairness. Fairness and QoS guarantees are usually weak points in opportunistic systems. We propose a modified proportional fair (PF) scheduler for orthogonal frequency-division multiplexing (OFDM) that addresses these weaknesses. The scheduler exploits multiuser diversity, but also tries to meet individual user requirements on bit rates and delays. The fairness of the scheduler is tunable; furthermore, a way to couple the scheduler and the beamformer to help the weakest users is proposed.

Opportunistic beamforming uses multiple antennas at the transmitter to increase the temporal fading rate of the individual users [1]. This can help slowly fading users to be scheduled more often. In addition, the fading rate of the intercell interference (ICI) is increased, which is called opportunistic nulling. The basic idea of opportunistic beamforming is that the base station forms a random beam that is changed for each transmission block. Users are then scheduled based on the reported supportable rates. In [1], the concept of using opportunistic beamforming for frequency-selective fading channels using OFDMA is outlined. We extend the idea of [1] by showing how opportunistic beamforming can be applied to OFDMA in practice. Also in [4], the extension of opportunistic beamforming to parallel channels is considered, but without introducing the same randomness in the frequency domain.

One of the main problems with FDD opportunistic OFDMA systems is the large amount of feedback required from the users. Because users can be scheduled on different frequency subbands, users must feed back measurement information about each subband. We propose to reduce the feedback by grouping adjacent subcarriers into clusters [5] and only feeding back information about the strongest clusters. Additionally, we observe that the suitable feedback rate per user depends on the number of users, and we design the adaptive feedback scheme accordingly. Alternatively, the feedback load can be reduced by feeding back information only from users with channel quality above a certain predefined threshold [6]. Clustered OFDMA and multiuser diversity were also studied in [2] and [3], where the authors showed an increase in spectral efficiency as the number of users grew. We propose the use of identical beamforming weights on all subcarriers within each cluster and independent weights between the clusters. This keeps the correlation high between the subcarriers within the clusters so that feedback of only one value is sufficient, e.g., the supportable rate of the weakest subcarrier within the cluster. Furthermore, by having different beam-

Paper approved by T. F. Wong, the Editor for Wideband and Multiple Access Wireless Systems of the IEEE Communications Society. Manuscript received October 18, 2004; revised July 30, 2006. This paper was presented in part at the Vehicular Technology Conference, Milano, Italy, May 2004, and in part at the Vehicular Technology Conference, Los Angeles, CA, September 2004.

P. Svedman and B. Ottersten are with the KTH School of Electrical Engineering, SE-100 44 Stockholm, Sweden (e-mail: patricks@ee.kth.se).

S. K. Wilson is with the Department of Electrical Engineering, Santa Clara University, Santa Clara, CA 95053 USA.

L. J. Cimini Jr. is with the Electrical and Computer Engineering Department, University of Delaware, Newark, DE 19716 USA.

Digital Object Identifier 10.1109/TCOMM.2007.896082

forming weights in the different clusters, the frequency fading can be increased, which is desirable.

In this paper, we also consider passing CSI from the scheduler to the opportunistic beamformer. The scheduler knows which are the strongest clusters of the users, conditioned on the current beamforming weights. If we assume positive temporal correlation between transmission blocks, the beamformer can increase the fairness by retaining the beamforming weights on some clusters.

To exploit multiuser diversity, a proper scheduler has to be used. To maximize the cell throughput, the user with the highest supportable rate should always be selected. An alternative to this potentially very unfair scheduler is the PF scheduler [7]. It offers a compromise between fairness and cell throughput. Examples of modifications of the single-carrier PF scheduler are given in [8] and [9]. In [10], several OFDMA extensions to the single-carrier PF scheduler are considered. We will use the recommended scheme in [10] as a comparison. In this paper, we propose a modified PF (M-PF) scheduler for OFDMA systems. It differs from previous PF schedulers in that it can accommodate different quality-of-service (QoS) classes and that it has a tunable fairness level. Many results on subcarrier, bit, and power allocation schemes for multiuser OFDM are available, for instance [11]–[15]. Optimization problems of different forms are solved to satisfy the minimum rate requirements of the different users. However, these schemes only consider resource allocation for one transmission block. Hence, resources may be wasted on a demanding user that is in a deep fade on most subcarriers. The philosophy of the PF scheduler is different in that it can let a user wait several transmission blocks if it is fading unfavorably. The M-PF scheduler proposed in this paper tries to satisfy the different rate requirements of the different users, while at the same time exploiting multiuser diversity. This can lead to many users not reaching their target rates. On the other hand, the system can accommodate more users that are near their target rates, increasing the multiuser diversity effect.

Recently, resource-allocation schemes that aim at user satisfaction instead of QoS have emerged, e.g., [16]. Also, the business model of the operator influences the desirable behavior of the scheduler [17]. A good scheduler in a system with few high-paying users is not necessarily good in a system with many moderately-paying users. The PF scheduler is more suitable in the latter scenario. Recent studies on the satisfaction of internet application users indicate that the satisfaction of an already well-served user increases only marginally by increasing the service level even further [18], [19]. However, if the service level is decreased below some level, the satisfaction level drops significantly. This is well captured by the logarithmic *utility function* which connects user satisfaction with the QoS. In fact, PF maximizes the sum utility when the utility function of each user is the logarithm of the average rate. We show that the M-PF scheduler proposed in this paper corresponds to a more general log-like utility function, to which the PF utility function is a special case.

The M-PF scheduler allows users to have different utility functions, defined through their target bit rates and target delays. Simulation results show that most users within a QoS class are equally distributed in terms of bit rates and delay. If the QoS

class is well defined for a particular application and most users are near their target QoS, the average user's satisfaction level will be high.

The three topics of this paper, feedback design, opportunistic beamforming, and scheduling, are closely coupled. The performance of a PF scheduled system heavily depends on the kind of CSI that is provided through the feedback as well as the fading rates of the users. The assumptions on feedback influence the beamforming design. This motivates a joint study of these three items.

The outline of the paper is as follows. In Section II, we propose an adaptive low-rate feedback scheme for opportunistic OFDMA, and, in Section III, we propose an opportunistic beamforming scheme. In Section IV an M-PF scheduler for OFDMA is presented. We evaluate the proposed feedback, beamforming, and scheduling methods using simulations in Section V. The paper is concluded in Section VI.

## II. ADAPTIVE LOW-RATE FEEDBACK FOR OPPORTUNISTIC OFDMA

In an FDD opportunistic OFDMA system, the users have to feed back subcarrier measurement information to the scheduler. The amount of feedback can be very large if there are many users and subcarriers in the system. In Section II-A, we propose a way to reduce the amount of feedback. We use the correlation between adjacent OFDM subcarriers and the fact that the information about the strongest subcarriers is most valuable to the scheduler. Furthermore, the amount of feedback needed from each user depends on the number of users in the system. Therefore, the feedback scheme is made adaptive in Section II-B. This section assumes that there is *one* channel quality measure per subcarrier and user. In Section III, we deal with vector-valued channels in connection with opportunistic beamforming. The feedback schemes of this section can, however, still be applied since the users experience scalar channels.

### A. Reduced Feedback by Clustering and Selection

Consider a system with  $K$  active users using OFDMA with  $N$  subcarriers. Let  $M_{\text{mod}}$  be the number of modulation and coding modes (i.e., rates) that the system supports. To feed back the supportable rate of each subcarrier and user would result in  $KN \log_2(M_{\text{mod}})$  bits of feedback. The feedback of all this information every transmission block can for many scenarios create an overwhelming overhead. To reduce the amount of overhead without sacrificing too much in performance, several measures can be taken. Since the correlation between adjacent subcarriers generally is high, the  $N$  subcarriers can be divided into  $Q$  clusters of  $R$  adjacent subcarriers, which can be used as feedback units. The feedback information is a measure of the channel quality in the cluster, for instance, the *minimum* subcarrier supportable rate within the cluster. In a well-designed system, the cluster size  $R$  is chosen so that the subcarriers within one cluster are highly correlated. Having small clusters offers better feedback accuracy for the subcarriers in the cluster but does not reduce feedback much. Having large clusters, on the other hand, reduces feedback more, but increases the risk of users feeding back supportable rates lower than necessary for some subcarriers. Hence, finding a suitable cluster size

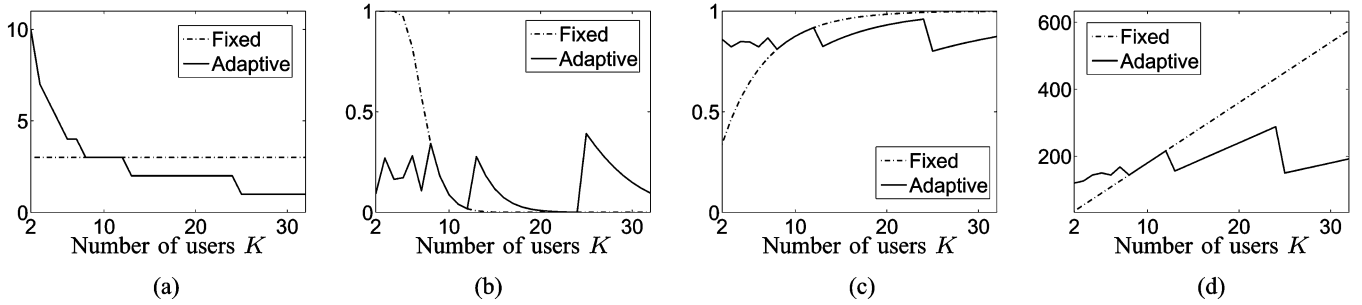


Fig. 1. In (a)–(d), the adaptive feedback rate scheme is compared with a fixed feedback scheme as a function of the number of users,  $K$ . (a) The feedback per user (clusters). (b) The probability that the spectral usage is below 75%. (c) The expected spectral usage. (d) Sum feedback of all users (bits per block).

in practice involves many parameters, for instance, channel statistics, the adaptive modulation properties and uplink signaling constraints, requiring detailed system simulations. Still, analysis of the channel statistics can give some hints on suitable cluster sizes, for instance, by computing the variance of the subcarrier gains within a cluster or by finding the probability that the minimum subcarrier gain within a cluster is less than, say, 3 dB, below the average of the cluster [20].

Furthermore, an opportunistic scheduler usually does not schedule users on their weakest clusters. Hence, the amount of feedback information can be further reduced by letting each user feed back information only about its  $S$  strongest clusters. This, however, introduces the need to also feed back the indices of  $S$  clusters, which gives a total feedback rate of  $KS \log_2(QM_{\text{mod}})$  bits per transmission block.

### B. Adaptive Feedback

If there are few active users and they feed back information about only a few clusters ( $K$  and  $S$  small), there is a high probability that the scheduler receives no information about some clusters. In opportunistic OFDMA systems, it is often not feasible to transfer transmit power from such clusters to ones that users have fed back. In practical systems, transmit power is rarely limited by a total power constraint, but rather a spectrum mask. Furthermore, increasing the rate of a user by allocating more power to a cluster can be problematic. The feedback supportable rates are based on the uniform power levels during training, both for the signal and for the interference. In addition, letting the scheduler assign other rates than the feedback supportable rates introduces additional feedforward overhead.

Before the adaptive feedback scheme is presented, spectral usage needs to be defined.

**Definition 1 (Spectral Usage):** Each user  $k$  independently feeds back information about a subset  $\mathcal{S}_k$  of the  $Q$  clusters, with  $|\mathcal{S}_k| = S$ . The spectral usage,  $\bar{U}$ , is a random variable defined as  $\bar{U} = (U_K/Q)$ , where  $U_K = |\bigcup_{k=1}^K \mathcal{S}_k|$ . The spectral usage is the fraction of all clusters that at least one user has fed back.

The expected spectral usage  $E[\bar{U}]$  is a function of  $K$ ,  $S$ , and  $Q$ , and is derived in Appendix I. Based on the expected spectral usage, an adaptive feedback scheme can be formulated. By fixing  $Q$ ,  $E[\bar{U}]$  can be made fairly constant by adapting the feedback rate  $S$  to the number of users  $K$ . If there are few active users, they each feed back more information. If there are many

active users, they each feed back less information. The scheme is described in the following.

- Offline, determine the desired function  $S = f(K)$  to fulfill a target spectral usage. The target can, for instance, be in either of the following equivalent forms (see Appendix I):
  - 1)  $E[\bar{U}|S, K, Q] = \bar{U}_{\text{target}}$  or
  - 2)  $\Pr(\bar{U} \leq \bar{U}_{\text{threshold}}|S, K, Q) = P_{\text{target}}$ .

The design parameters  $\bar{U}_{\text{target}}$ ,  $\bar{U}_{\text{threshold}}$  and  $P_{\text{target}}$  are used to trade off spectral usage with feedback overhead. Since  $S$  is an integer, the target can only be met approximately.

- Each user must feed back at least one cluster, no matter how many users that are active.
- Depending on the number of active users, the basestation broadcasts the number of clusters  $S$  each user is to feed back. Since the number of users typically is a slowly changing parameter, this induces very little feedforward overhead. To reduce the complexity of the feedback protocol, the adaptive feedback function  $S = f(K)$  can be additionally quantized.

This scheme has several positive effects. The aggregate feedback rate as a function of the number of users can be kept fairly constant. For a very high number of active users, they each feed back only their best cluster, which makes sense. For few users, they each feed back many clusters, so that a large part of the bandwidth can be used.

Note that this adaptive scheme can also be applied on a subcarrier level, if clustered feedback is not used. Fig. 1 illustrates the adaptive feedback scheme. In this example, the number of clusters  $Q = 8$  and the number of modulation levels  $M_{\text{mod}} = 8$ . For the fixed feedback scheme, each user feeds back the modulation level and index of the three strongest clusters ( $S = 3$ ). For the adaptive feedback rate,  $S$  is chosen so that the probability that less than 80% of the clusters are fed back is  $\approx 20\%$ . In Fig. 1(a), the resulting  $S$  for the adaptive scheme is plotted together with the fixed  $S$  for the fixed scheme. For few users ( $K < 7$ ), each user feeds back more information than in the fixed scheme, but less when there are many users ( $K > 12$ ). In Fig. 1(b), the target probability is shown. The abrupt changes in the probability for the adaptive scheme for large  $K$  is due to the integer granularity of  $S$ . From Fig. 1(c), it is clear that the adaptive feedback scheme results in a relatively constant expected spectral usage. The total feedback rate for all users in bits per scheduling block is depicted in Fig. 1(d). Since the feedback

rate per user is fixed for the fixed scheme, the total feedback rate increases linearly with the number of users. For the adaptive scheme, the total feedback rate remains fairly constant for the whole range of  $K$ .

### III. OPPORTUNISTIC BEAMFORMING FOR CLUSTERED OFDMA

In an opportunistic OFDMA system, it is favorable if the users fade, both in time and frequency. It is difficult to combine multi-user diversity gain and fairness in a system with only stationary and flat-fading users. The opportunistic beamforming scheme proposed in this section is a way to increase the fading rates of the users in both time and frequency. The scheme is based on the subcarrier clustering described in Section II-A.

We assume a basestation with  $M$  transmit antennas and  $K$  users with one receive antenna each. The basestation can obtain information about the downlink channels only through feedback. To reduce the amount of feedback from the users to the basestation scheduler, the feedback scheme proposed in the previous section is considered. Assuming synchronous basestations and omitting the time index, the  $k$ th user that is associated with basestation 0 receives, on the  $n$ th subcarrier

$$y_{n,k} = \mathbf{h}_{n,k}^i \mathbf{b}_n^0 \sqrt{P} x_n^0 + \underbrace{\sum_{i=1}^I \mathbf{h}_{n,k}^i \mathbf{b}_n^i \sqrt{P} x_n^i}_{\text{ICI}} + z_{n,k} \quad (1)$$

where  $x_n^i$  is the unit-energy transmitted symbol from basestation  $i$ ,  $z_{n,k}$  is complex-valued additive white Gaussian noise with variance  $\sigma_z^2$ ,  $\mathbf{h}_{n,k}^i \in \mathbb{C}^{M \times 1}$  is the baseband frequency response vector from the  $M$  antennas of the  $i$ th basestation to the  $k$ th user on the  $n$ th subcarrier,  $\mathbf{b}_n^i \in \mathbb{C}^{M \times 1}$  is the transmitter beamforming vector for the  $n$ th subcarrier and  $i$ th basestation, and  $P$  is the transmit power per subcarrier. The elements of  $\mathbf{b}_n$  are i.i.d. complex variables with uniformly distributed phase over  $[0, 2\pi)$  and magnitude over  $[0, 1)$  and normalized so that  $\|\mathbf{b}_n\|_2 = 1$ . Let  $H_{n,k}^i = \mathbf{h}_{n,k}^i \mathbf{b}_n^i$  denote the complex-valued effective baseband frequency response from basestation  $i$  to user  $k$  on subcarrier  $n$ , so that

$$y_{n,k} = H_{n,k}^0 \sqrt{P} x_n^0 + \sum_{i=1}^I H_{n,k}^i \sqrt{P} x_n^i + z_{n,k}. \quad (2)$$

Hence, the users effectively experience scalar channels on all subcarriers. Consequently, the opportunistic beamforming is transparent to the users, which do not need to know the number of transmit antennas. To increase the frequency fading rate, we propose a clustered beamforming (CL-BF) design for  $\mathbf{b}_n$ . For CL-BF, the beamforming vectors  $\mathbf{b}_n$  are identical within one cluster, but independent between the clusters. Hence, the subcarrier correlation within the clusters is maintained in the effective channel, whereas the correlation between the clusters in the effective channel is reduced. By having different beamforming weights in different clusters, we can increase the frequency fading rate. This is a direct extension of the idea in [1], where different beamforming weights were used in different transmission blocks to increase the temporal fading rate. Note that in an OFDMA system without subcarrier clustering,

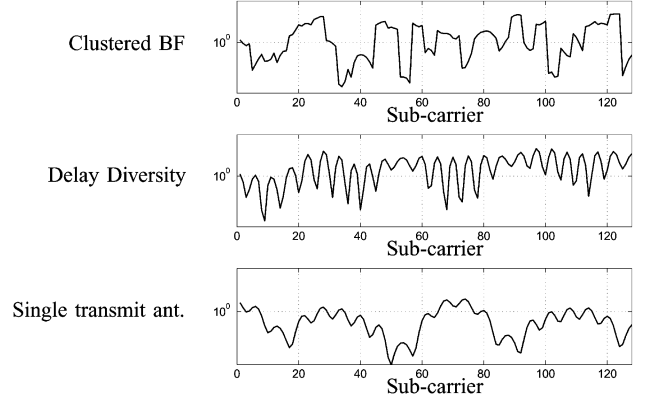


Fig. 2. Example of effective gains of a user for different transmitter schemes. The lowest plot shows the gains for a SISO channel. By adding one transmit antenna and DD in the middle plot, more frequency variability is induced. In the uppermost plot, the effective gains for CL-BF are shown. The cluster size  $R$  is four subcarriers.

the independent beamformers should be applied on a subcarrier basis.

For coherent data reception, the users need to estimate only the effective channels  $H_{n,k}^0$ , treating interference as noise. For the supportable rate estimation for the feedback, the users need to estimate the signal-to-interference-plus-noise ratio (SINR) of each subcarrier. Fortunately, the interference channels do not need to be estimated, since only the combined interference plus noise power is of interest. Hence, the number of parameters to be estimated at the receivers is the same as for opportunistic single-input single-output (SISO) OFDMA. As a final note, it is greatly beneficial in an opportunistic beamforming system if the training periods with new sets of beams are synchronized between the cells. This is necessary for the feedback from the users to be accurate and for the opportunistic nulling to take place [1].

As a comparison to the proposed CL-BF, the equal beamforming (EQ-BF) design of  $\mathbf{b}_n$  will also be used in the simulations. For EQ-BF,  $\mathbf{b}_n = \mathbf{b} \forall n$ . Hence, only one inverse fast Fourier transform (IFFT) has to be performed at the transmitter, compared with  $M$  for CL-BF, but no extra frequency fading is induced. The beamforming weights are constant during one transmission block, which consists of  $N_{tb}$  OFDM symbols, but change from block to block, just as in single-carrier opportunistic beamforming.

A second alternative to CL-BF is delay diversity (DD) [21]. In a DD system, the same signal is transmitted on all antennas, but time-delayed on a subset of the antennas. This results in an artificially longer channel impulse response and more diversity also in the frequency domain. The advantage of such an approach is its low complexity. The first drawback of a DD solution is that the length of the cyclic prefix has to be increased to incorporate the longest length of the artificial impulse response of any user. The second drawback is that it increases the frequency variability not only between the clusters, but also within the clusters. This leads to less accuracy in the feedback rates.

To illustrate the effect of clustered opportunistic beamforming and DD, the effective gains across the subcarriers of a user are shown in Fig. 2. For the clustered opportunistic beamforming and DD, two transmit antennas are used.

#### IV. QoS-AWARE PF SCHEDULING FOR OFDMA

In opportunistic systems, there is a tradeoff between high cell throughput and fairness. By always scheduling the user with the highest SINR, the highest cell throughput is achieved. Some users with low SINR may, however, never be scheduled. The PF scheduler partially solves this problem [1]. In future communication systems, users will likely have different requirements, for instance, real-time data, voice, or background downloading of data. This is not taken into account by the PF algorithm, which treats all users equally. In this section, we modify the PF algorithm to take user requirements in terms of bit rates and delays into account. The M-PF scheduling algorithm is designed for a clustered OFDMA downlink, with partial channel knowledge at the transmitter according to the feedback scheme proposed in Section II. Henceforth, we call this algorithm the M-PF scheduler. Note that the M-PF scheduler could be applied in a single-carrier system as well with small modifications. Note that the transmit buffers of the active users are assumed to contain data.

##### A. Preliminaries

Assume that each user  $k$  belongs to a QoS class with parameters  $R_k$  and  $T_k$ , the target rate in bits/s and averaging window time in seconds, respectively. The averaging window time  $T_k$  reflects the target delay of the QoS class. Long delay windows allow longer periods of not being scheduled, whereas a short delay window will help schedule the user more often.

The CSI at the transmitter at block  $m$  is  $S$  instantaneous rate values for each of the  $K$  users corresponding to the minimum supportable rate in a cluster. This information is gathered in the set of supportable rates  $\{C_{q,k}(m)\}$ , where for each  $k$  only  $S$  rates are nonzero (the ones that were fed back).  $S$  is the number of fed-back clusters and  $Q$  is the number of clusters.

In the following,  $\bar{X}$  denotes the average measure over the historical time window,  $\tilde{X}$  denotes the measure over only the current block that is to be scheduled, i.e., what the scheduler can influence, and  $\hat{X}$  denotes the *new* average measure, including both the historical time window and the current block.

The average rate  $\bar{R}_k(m)$  reflects the average bit rate of user  $k$  during the historical averaging window time  $T_k$ . The scheduled rate of user  $k$ ,  $\tilde{R}_k(m)$  is to be decided by the scheduler

$$\tilde{R}_k(m) = \sum_{q=1}^Q \tilde{R}_{q,k}(m) = R \sum_{q=1}^Q \gamma_{q,k}(m) C_{q,k}(m)$$

where  $R$  is the number of subcarriers per cluster,  $\tilde{R}_{q,k}(m)$  is the instantaneous scheduled rate of user  $k$  on cluster  $q$ , and  $\gamma_{q,k}(m)$  is the indicator function with the property that  $\gamma_{q,k}(m) = 1$  if user  $k$  is scheduled on cluster  $q$  during block  $m$ , and otherwise, zero. Since at most one user can be scheduled per subcarrier,  $\sum_k \gamma_{q,k}(m) \in \{0, 1\}$ .

The average rate *after* scheduling is assumed to be computed as

$$\hat{R}_k(m) = \alpha_k \tilde{R}_k(m) + \beta_k \bar{R}_k(m) \quad (3)$$

where  $\alpha_k$  and  $\beta_k$  are weighting factors with the property that  $\alpha_k + \beta_k = 1$ . The weighting factors are user-specific, since

they may depend on the length of the averaging time window, which is user-specific.

For each user, the scheduler keeps track of the average *relative* rate of each user

$$\bar{B}_k(m) = \frac{\bar{R}_k(m)}{R_k}$$

$\bar{B}_k(m)$  is a measure of how well the user has met its rate requirements in the historical averaging window time,  $T_k$ :

- $\bar{B}_k(m) > 1$ : User  $k$  has exceeded its target rate;
- $\bar{B}_k(m) = 1$ : User  $k$  has just met its target rate;
- $\bar{B}_k(m) < 1$ : User  $k$  has undershot its target rate.

As for the absolute rate, the instantaneous and new average relative rates are defined, respectively, as

$$\begin{aligned} \tilde{B}_k(m) &= \frac{\tilde{R}_k(m)}{R_k} \\ \hat{B}_k(m) &= \frac{\hat{R}_k(m)}{R_k} = \alpha_k \tilde{B}_k(m) + \beta_k \bar{B}_k(m). \end{aligned}$$

##### B. Utility Function

Utility functions are used to describe the connection between user utility, or satisfaction, and QoS parameters, e.g., average rate. The classical PF scheduler maximizes the sum of the users' utilities if the utility function  $U(\hat{R}_k) = \ln(\hat{R}_k)$  [7]. The logarithmic utility function indicates that users with low average rates benefit more in utility from being scheduled than users with high average rates. In this section, we propose a more general utility function, that also is user-specific, to capture the heterogeneous user requirements.

Consider the class of concave and differentiable utility functions

$$U_k^{\kappa}(\hat{B}_k) = \frac{R_k \beta_k^{\kappa}}{\alpha_k} \frac{1}{1 - \kappa} (\hat{B}_k^{1-\kappa} - 1) \quad (4)$$

where  $\kappa \in [0, 1) \cup (1, \infty)$  is a *fairness* parameter, and the time index  $m$  is left out for brevity. In previous work, utility functions are usually functions of the average user bit rate. Here, the utility function depends on the average relative rate  $\hat{B}_k$ , which is natural if the users use different applications with different target bit rates corresponding to an acceptable user satisfaction level. Furthermore, the user-specific weighting constant in the utility function,  $R_k \beta_k^{\kappa} / \alpha_k$ , scales the utility function according to the QoS parameters of the user.

In Appendix II, we show that (4) simplifies to several known utility functions when the QoS parameters  $R_k$ ,  $\alpha_k$  and  $\beta_k$  are equal for all users. When  $\kappa \rightarrow 1$ , (4) simplifies to the utility function of a PF scheduler. When  $\kappa = 0$ , (4) simplifies to the utility function of the maximum sum-rate scheduler. When  $\kappa \rightarrow \infty$ , (4) corresponds to the utility function of the max-min scheduler [22].

Note that the appropriate utility function of a particular user depends on the type of application used. We believe that the utility functions of a large range of wireless applications can be captured by the general form in (4). Also note that the fairness parameter  $\kappa$  is common for all users and can only be tuned on a system level.

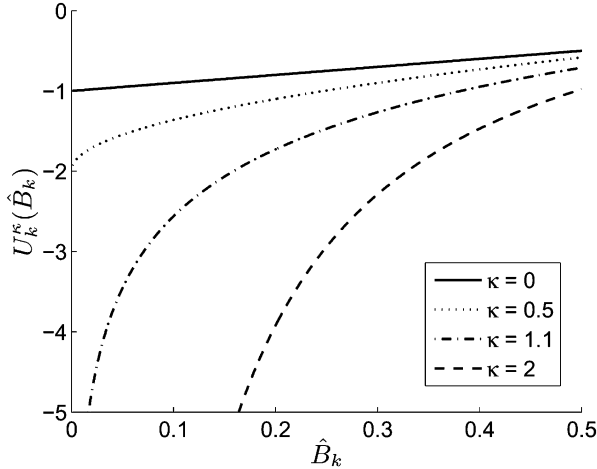


Fig. 3. Effect of the fairness parameter  $\kappa$  on the utility function.

In Fig. 3, the effect of the fairness parameter on the utility function is visualized. The QoS parameters are set to  $(R_k/\alpha_k) = 1$  and  $\beta_k = 0.99$ . The utility function for  $\kappa = 0$  is linear, corresponding to the linear increase of utility with the average rate for all rates. Increasing  $\kappa$  gives a more steep function for small average rates. This corresponds to a scheduler that prefers to allocate users with low average rates, since a great increase in utility can be achieved. The choice of  $\kappa$  for a practical system depends on the desired resource allocation behavior, which in turn depends on the type of traffic that is expected. For more elastic traffic [23], i.e., traffic that tolerates packet delays gracefully, for instance, file downloads, the fairness parameter can be set to a low value. For traffic that relies on a steady data flow, a higher  $\kappa$  should be selected. Note that the delay behavior is also controlled by the user-specific QoS parameters in the utility function.

### C. M-PF Scheduler

Assume that  $\beta_k \bar{R}_k(m) \gg \alpha_k \tilde{R}_k(m)$ . Then, the scheduler that maximizes the sum-utility

$$\sum_{k=1}^K U_k^\kappa(\hat{B}_k(m)) \quad (5)$$

where  $U_k^\kappa(\hat{B}_k(m))$  is defined in (4), is the following. For each cluster  $q$ , the scheduler chooses a user according to the criterion

$$k_q^*(m) = \arg \max_k \left\{ \frac{C_{q,k}(m)}{\bar{B}_k(m)^\kappa} \right\}. \quad (6)$$

This claim is proven Appendix II. In practice, a small positive regularization term can be added to the denominator to stabilize the system for the cases when  $\bar{B}_k(m) = 0$ . The assumption can be motivated since  $\beta_k \gg \alpha_k$  when the historical averaging window time is much larger than one transmission block. It is reasonable to assume that the instantaneous scheduled rate  $\tilde{R}_k(m)$  is not much greater than the historical average rate  $\bar{R}_k(m)$ , since an OFDMA system offers many more scheduling opportunities than the “all-or-nothing” allocation of single-carrier systems, resulting in a steadier data flow for the users.

### D. Exploiting Temporal Channel Correlation in the Scheduler and Beamformer

For slowly moving users, the temporal channel correlation between transmission blocks is high. If CL-BF is used, this can be exploited by the scheduler to help users that are well below their target rates. By keeping the beamforming weights for the user’s strongest cluster during the next transmission block, the probability that the cluster is strong also during the next transmission block is increased.

- 1) A user  $k$  is considered to be well below its target rate if  $B_k < \nu$ , where  $\nu < 1$  is a threshold value.
- 2) For each user  $k$  that is well below its target rate, keep the beamforming weights on the strongest cluster of the user during the next transmission block.

Note that this method does not require knowledge of the users’ speeds.

## V. SIMULATION RESULTS

In order to evaluate the proposed schemes, simulations of a multicell system have been performed. First, the simulation environment and assumptions are described and then the results are presented and discussed.

### A. Simulation Environment and Assumptions

To evaluate the performance of the proposed schemes, the downlink of an FDD system with seven cells has been simulated. Results are collected from one cell which is surrounded by six interfering cells. All basestations are assumed to be synchronous and use the same frequency band for the downlink, making the training period simultaneous for all cells. All basestations use the same beamforming and scheduling schemes. The users are assumed to perfectly estimate the SINR on all subcarriers, and the feedback delay is assumed to be zero. The adaptive feedback rate is such that the probability that less than 80% of the subcarriers can be assigned is 0.2 (see Section II).

If we assume that the interference is Gaussian, the downlink channel consists of parallel Gaussian subchannels. In the simulations, the supportable number of bits per OFDM symbol for the  $n$ th subcarrier and the  $k$ th user in cell 0 is estimated using the gap approximation

$$\frac{1}{2} \log_2 \left( 1 + \frac{\text{SINR}}{\Gamma} \right) = \frac{1}{2} \log_2 \left( 1 + \frac{|H_{n,k}^0|^2}{\Gamma \left( \frac{\sum_{i=1}^I |H_{n,k}^i|^2 + \sigma_z^2}{P} \right)} \right) \quad (7)$$

where the notation is as in (2) and  $\Gamma$  is the gap corresponding to a bit-error rate (BER) of  $10^{-4}$  for quadrature amplitude modulation (QAM) [24]. This means that it is assumed that the adaptive modulation can also handle a noninteger number of bits per symbol. It is assumed that the scheduled users achieve the rates they estimate from (7).

The system parameters are described in Table I. The 3GPP spatial channel model for multiple-input multiple-output (MIMO) simulations for urban environments is used [25]. It is modified for omnidirectional antennas by changing the antenna gain pattern to be uniform. The users are uniformly distributed over the cells and their speeds are uniformly distributed between 0 and 100 km/h. Temporal correlation over the simulated OFDM symbols is added using the Jakes model [26].

TABLE I  
SIMULATION PARAMETERS (IN SOME FIGURES  $\kappa$  IS VARIED)

Sampling frequency	4 MHz
Number of sub-carriers $N$	128
Cyclic prefix length	12 $\mu$ s
Total OFDM symbol period $T_s$	32+12=44 $\mu$ s
Sub-carrier spacing	31.25 kHz
Total bandwidth	4 MHz
SNR at cell boundary	3 dB
Total transmit power $P_{tot}$	500 W
Carrier frequency	1900 MHz
Number of transmit antennas	2
Antenna separation	4 $\lambda$
Maximum user speed	100 km/h
Cell radius	1 km
Distance between base-stations	2 km
Number of interfering base-stations	6
Cluster-size $R$	2 sub-carriers
Fairness parameter $\kappa$	0.5
Weight-keeping threshold $\nu$	0.8
Transmission block $N_{tb}$	16 OFDM symbols
Simulation time	1536 $T_s \approx 68$ ms

Three user classes are considered: one class with high bit-rate requirements and a short delay window, corresponding to a user using a real-time application; one class with medium bit-rate and delay requirements; and one class with a low bit rate and a long delay window, corresponding to a user downloading data in the background. Half of the users belong to Class 3 and one quarter of the users belong to Classes 1 and 2 each. Note that, in general, each user can have individual rate and delay requirements.

- Class 1:  $R_k = 256$  kb/s and  $T_k = 20N_{tb}T_s = 0.9$  ms.
- Class 2:  $R_k = 128$  kb/s and  $T_k = 40N_{tb}T_s = 1.8$  ms.
- Class 3:  $R_k = 64$  kb/s and  $T_k = 60N_{tb}T_s = 2.6$  ms.

In this paper, the average user bit rate  $\bar{R}_k(m)$  is computed for the historical averaging window time  $T_k = N_k N_{tb} T_s$ , where  $N_k \in \{20, 40, 60\}$  is the length of the historical averaging window in blocks. Hence, the user-specific update parameters can be found as

$$\alpha_k = \frac{1}{N_k + 1}, \quad \beta_k = \frac{N_k}{N_k + 1}.$$

In many other papers, for instance, [1], the average user bit rate is updated using an exponentially weighted lowpass filter with slower decay to zero, which could also be used here. However, using the true average bit rate emphasizes scheduling of users that have not been scheduled during the time window.

The opportunistic beamforming schemes with PF schedulers are compared with a smart antenna scheme with round-robin scheduling (Conv. BF RR). In the smart antenna scheme, conventional beamforming is used on all OFDM subcarriers to transmit data to one user at a time. This requires detailed CSI at the transmitter, which is also assumed. The basestations have no channel knowledge of the users in the adjacent cells, so interference nulling is not possible. The transmitter computes the beamforming weights for *each subcarrier* so that the received signals combine coherently at the receive antenna,  $\mathbf{b}_n = \mathbf{h}_n^* / \|\mathbf{h}_n\|_2$ . For the conventional beamforming, waterfilling of the transmit power across the subcarriers is done, giving the stronger subcarriers more power. This may result in some subcarriers not being used. For the smart antenna scheme, round-robin scheduling is used, with fixed allocated

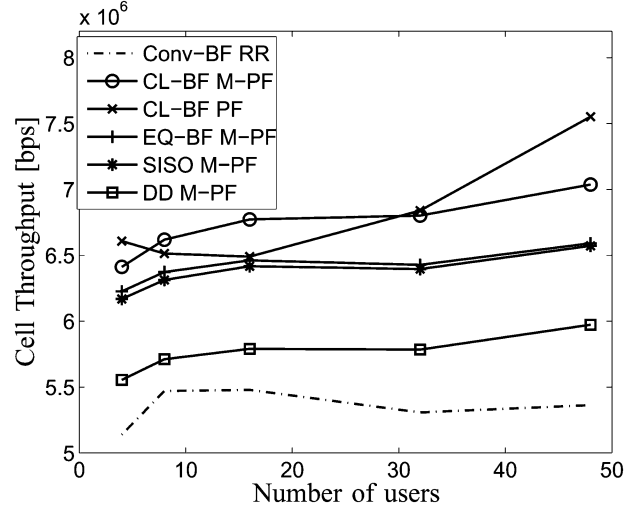


Fig. 4. Cell throughput as a function of the number of users in each cell for two transmit antennas. Conventional beamforming with round-robin scheduling, CL-BF with M-PF and PF scheduling, EQ-BF on all clusters, single transmit antenna, and DD with M-PF scheduling are compared.

time slots for each user. Users with high-rate requirements will be assigned correspondingly more time slots. One time slot here is equal to one transmission block.

The M-PF scheduler proposed in this paper is also compared with the OFDMA PF extension recommended in [10], which will be called standard PF. Note that this scheduler assumes homogeneous users, in terms of bit rate and delay requirements. To evaluate the performance of CL-BF, it is compared with EQ-BF weights across the subcarriers, DD (see Section III), and single-antenna transmission (SISO).

Delay is an important QoS measure. In the simulations, we have defined the maximum scheduling delay a user experiences as the maximum time between the reception of two consecutive packets. We have defined a packet to be 512 bits. A user which is rarely scheduled will have a high maximum delay. Delay is also a function of the rate of the user. Users with very low bit rates will experience longer delays, even if they are scheduled each OFDM symbol.

## B. Simulation Results

The cell throughput is the sum of the bit rates of all users in one cell. In Fig. 4, the cell throughput for different beamforming and scheduling schemes is compared as a function of the number of users in each cell. The schemes based on CL-BF perform best. The low cell throughput of the smart antenna scheme is due to the inability of the round-robin scheduler to avoid ICI. For the opportunistic schemes, the subcarriers with high ICI are probably not even fed back. In an environment free of ICI, the smart antenna scheme outperforms the opportunistic systems for less than 48 users [27]. The cell throughput of the DD scheme decreases due to the increased cyclic prefix overhead. There is only a small difference in total throughput between single-antenna transmission and equal opportunistic beamforming on all clusters (EQ). The reason is that for Rayleigh fading channels (which the channel model closely resembles) with relatively fast fading, opportunistic beamforming does not change the fading statistics [1]. The small gain comes from the slowly fading users

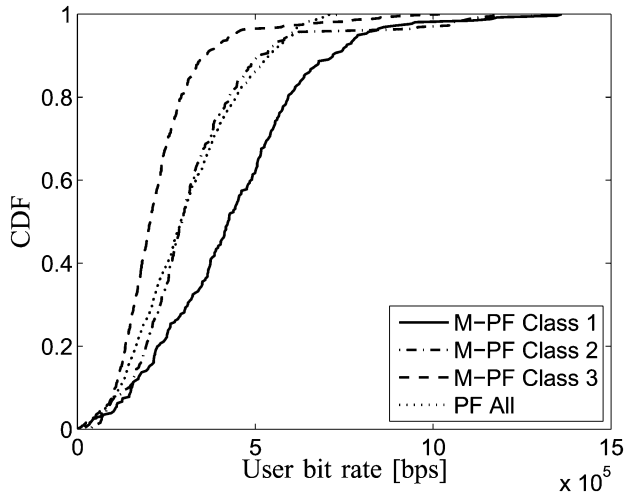


Fig. 5. CDF of the bit rates of the 32 users for CL-BF with M-PF and PF scheduling. Note that the CDF for PF is for all users. The cell throughputs of the two schemes are nearly the same for 32 users.

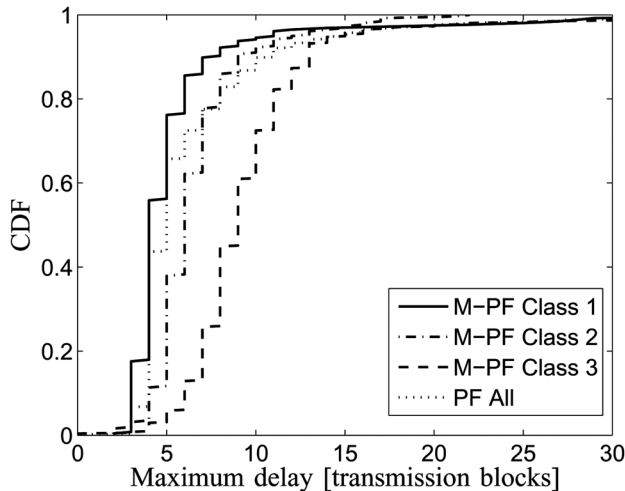


Fig. 6. CDFs of the maximum delays of the 32 users in the cell for CL-BF with M-PF and PF scheduling.

that can be scheduled more often for EQ. For CL-BF, however, the fading rate in both time and frequency is increased, resulting in more efficient scheduling.

The M-PF scheduler is designed to give users the rates they require, but not much more. This is not possible in the standard PF scheduler. To evaluate how well this discrimination works, user rate cumulative distribution functions (CDFs) are plotted in Fig. 5 when there are 32 users per cell. The cell throughputs for CL-BF, M-PF, and CL-BF PF are nearly the same for 32 users (see Fig. 4). However, the M-PF scheduler manages to differentiate the rates of the different QoS classes, as can be seen in Fig. 5. The rates of the users within the QoS classes are fairly concentrated. This indicates that the system is relatively fair. If the rate requirements of the classes are properly set and the system is not overloaded, most users will also be satisfied. Since users are not treated differently by the PF algorithm, the CDF is shown for all users.

In Fig. 6, the CDFs of maximum delay for CL-BF, M-PF, and CL-BF PF are plotted. For all three classes and also for PF, most users experience a maximum delay below 10 transmission

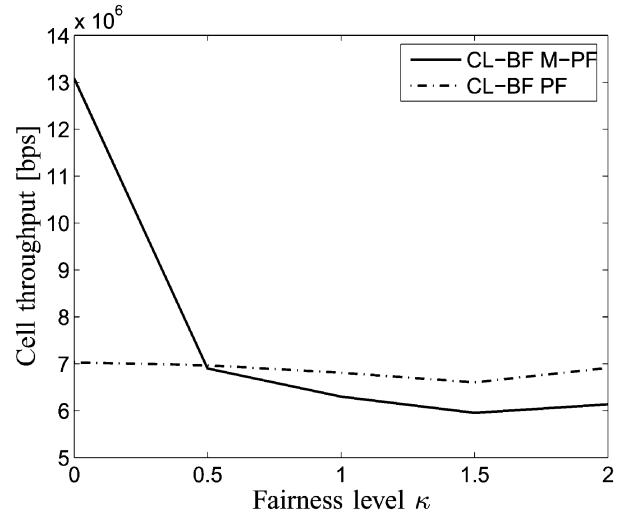


Fig. 7. Cell throughput for 32 users as a function of the fairness parameter  $\kappa$  is plotted for CL-BF with M-PF scheduling. As a comparison, the throughput of CL-BF with standard PF is plotted. The standard PF has no fairness parameter.

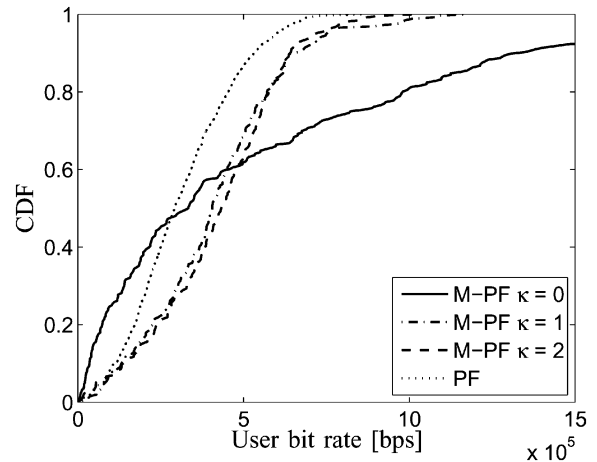


Fig. 8. User-rate CDFs for QoS class 1 for the CL-BF with M-PF scheme are plotted for various fairness levels  $\kappa$ . There are 32 users in the cell.

blocks, which equals 7 ms. The low delay can be explained by the many time-frequency slots that are available to the scheduler (at most  $Q$  users can be scheduled in the same OFDM symbol). The maximum delay of users in QoS Class 3 (background download) is larger, which is acceptable.

To evaluate the effect of the fairness parameter  $\kappa$  in the M-PF scheduler, the cell throughput is plotted as a function of  $\kappa$  in Fig. 7. The cell throughput of the PF scheduler (which has  $\kappa = 1$ ) is plotted as a comparison. As expected, the highest cell throughput is achieved with  $\kappa = 0$ , when the strongest user is always scheduled, without taking fairness into account. User rate and delay CDFs are plotted in Figs. 8 and 9. A higher fairness level gives more equally distributed users within the QoS classes.

To further help the weakest users, the M-PF scheduler interacts with the beamformer, as described in Section IV-D. The beamforming weights of the strongest cluster of the weakest users are kept to the next transmission block. The effect of this method can be seen in Fig. 10, which is a zoom-in on the CDF of the weakest users of QoS Class 1 with a target rate of 256 kb/s. The gain of this method is not larger because the ICI changes



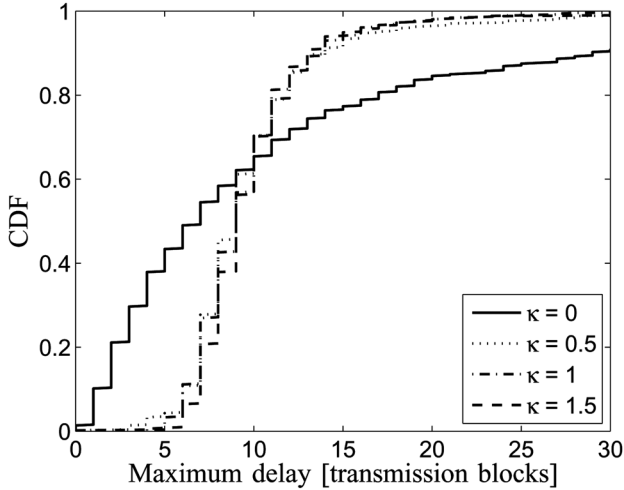


Fig. 9. Maximum delay CDFs for QoS class 3 for the CL-BF with M-PF scheme are plotted for various fairness levels  $\kappa$ . A delay of 10 transmission blocks is about 7 ms. There are 32 users in the cell.

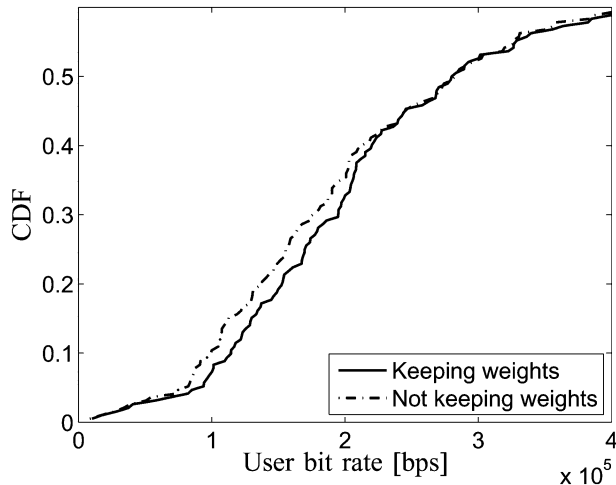


Fig. 10. CDFs for the 60% of the users in class 1 with the lowest rates. CL-BF with M-PF with beamforming weight keeping is compared with CL-BF with M-PF without beamforming weight keeping (see Section IV-D). Beamforming weight keeping means keeping of the beamforming weights for the reported strongest cluster of the weakest users during consecutive transmission blocks. The threshold for considering a user weak is here  $0.8 \times$  target rate, which is 205 kb/s.

between transmission blocks. Even if the basestation keeps the beamforming weights on the strongest cluster, the ICI might change radically.

## VI. CONCLUSIONS

Opportunistic OFDMA is a promising candidate for FDD downlinks in future wireless broadband systems. Users can be

scheduled on the orthogonal subcarriers where they experience the most favorable fading. Such systems pose some problems which we address in this paper. We propose a feedback scheme in which we use some inherent properties of OFDM to significantly reduce the amount of feedback, as well as adapt to the number of users. Furthermore, we address the problem of slowly fading and flat-fading users by proposing an opportunistic beamforming scheme which randomizes the channel in both time and frequency. To meet requirements on fairness and users with different requirements, we extend the standard PF algorithm to include classes of users with different QoS requirements in terms of bit rate and delay. We also introduce a fairness parameter in the PF scheduler. Additionally, we propose a method to couple the opportunistic beamformer with the scheduler. Based on feedback measurements, the scheduler can order the beamformer to keep the beams in some parts of the frequency band. This can help users that have problems meeting their QoS requirements. These methods have all been evaluated by means of simulation, using the 3GPP spatial channel model. The proposed opportunistic beamforming scheme performed well in comparison with several other schemes, for instance, a smart antenna scheme and a DD scheme. The M-PF scheduler was able to give users different QoS, based on their requirements, while still exploiting multiuser diversity.

## APPENDIX I

In this appendix, the expected spectral usage and the probability that less than  $u \in \{S, \dots, Q\}$  different clusters are fed back from  $K$  users are derived ( $K \geq 1$ ). These result can be used to find a proper adaptive feedback rate as a function of the number of users  $K$ .

It is assumed that the users independently feed back CSI about a subset  $\mathcal{S}_k$  of the  $Q$  clusters, with  $|\mathcal{S}_k| = S$ . Let the random variable  $U_k$  denote the number of different clusters, out of the  $Q$  clusters, that have been fed back when  $k$  users have made their pick, i.e.,  $U_k = |\bigcup_{i=1}^k \mathcal{S}_i|$ . The probability that less than  $S$  clusters are picked is zero, as is the probability that more than  $Q$  are picked. The nonzero probabilities can be stacked in a vector and computed recursively from

$$\mathbf{P}_k = \begin{pmatrix} \Pr(U_k = S) \\ \Pr(U_k = S + 1) \\ \vdots \\ \Pr(U_k = Q) \end{pmatrix} = \mathbf{A} \begin{pmatrix} \Pr(U_{k-1} = S) \\ \Pr(U_{k-1} = S + 1) \\ \vdots \\ \Pr(U_{k-1} = Q) \end{pmatrix} = \mathbf{A} \mathbf{P}_{k-1}$$

where the lower triangular matrix  $\mathbf{A}$  denotes the probability transitions when one more user picks clusters, as shown in the equation at the bottom of the page. The elements of  $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} \Pr(U_k = S | U_{k-1} = S) & & & \\ \Pr(U_k = S + 1 | U_{k-1} = S) & \ddots & & 0 \\ & \vdots & & \\ \Pr(U_k = Q | U_{k-1} = S) & \dots & \Pr(U_k = Q | U_{k-1} = Q) & \end{pmatrix}$$

are computed as shown in the equation at the bottom of the page. Since the first user ( $k = 1$ ) picks  $S$  different clusters

$$\mathbf{P}_1 = \begin{pmatrix} P(U_1 = S) \\ P(U_1 = S + 1) \\ \vdots \\ P(U_1 = Q) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\mathbf{P}_K = \mathbf{A}^{K-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The probability that  $u$  or fewer different clusters are fed back by  $K$  users can be computed as

$$\Pr(U_K \leq u) = \sum_{i=0}^{u-S} \mathbf{P}_K(i)$$

where  $\mathbf{P}_K(i)$  denotes the  $i$ th element of  $\mathbf{P}_K$ . Finally, the expected spectral usage can be computed as

$$E\left[\frac{U_K}{Q}\right] = \frac{1}{Q} \sum_{u=S}^Q u \Pr(U_K = u) = \frac{1}{Q} \sum_{u=S}^Q u \mathbf{P}_K(u - S).$$

Given  $Q$  and  $K$ , to find the feedback rate  $S$  so that  $\Pr(U_K \leq u)$  or  $E[U_K/Q]$  remains fairly constant for all  $K$ , a full search of all integer  $S \in \{1, \dots, Q - 1\}$  can be done. Doing this for all feasible values of  $K$  yields the function  $S = f(K)$ .

## APPENDIX II

In this appendix, we elaborate on the utility function (4) and derive the scheduler in Section VI.

*Utility Function Simplifications:* The proposed general utility function in (4) is restated here

$$U_k^\kappa(\hat{B}_k) = \frac{R_k \beta_k^\kappa}{\alpha_k} \frac{1}{1 - \kappa} (\hat{B}_k^{1-\kappa} - 1). \quad (8)$$

When the QoS parameters of all users are equal and the fairness parameter  $\kappa$  equals zero, the utility function becomes linear and equal for all users

$$U_k^\kappa(\hat{B}_k) = U^0(\hat{B}_k) = C_0(\hat{R}_k - 1)$$

which corresponds to the classical maximum rate scheduler, with  $C_0$  being a constant.

When the QoS parameters of all users are equal and the fairness parameter approaches 1

$$\begin{aligned} \lim_{\kappa \rightarrow 1} U_k^\kappa(\hat{B}_k) &= \lim_{\kappa \rightarrow 1} U^\kappa(\hat{B}_k) \\ &= \lim_{\kappa \rightarrow 1} C_1 \frac{1}{1 - \kappa} (\hat{B}_k^{1-\kappa} - 1) \\ &= \lim_{\kappa \rightarrow 1} C_1 \frac{1}{1 - \kappa} \\ &\times \left( (1 - \kappa) \ln \hat{B}_k + \frac{1}{2!} ((1 - \kappa) \ln \hat{B}_k)^2 + \dots \right) \\ &= C_1 \ln \hat{B}_k = C_1 \ln \hat{R}_k - D_1 \end{aligned}$$

the utility function corresponds to that of the classical PF scheduler, with  $C_1$  and  $D_1$  being constants. In the third step, the power series expansion of  $\hat{B}_k^{1-\kappa}$  was used. When  $\kappa \rightarrow \infty$  and the QoS parameters of the users are equal, the utility function corresponds to that of the max-min scheduler, which will be clarified later in Appendix II-C.

*Scheduling Metric Derivation:* In the following, the claim in Section IV-C is established. The problem of maximizing (5) is complex, since it involves a joint user allocation on all clusters. Here, it is relaxed by assuming that  $\beta_k \bar{R}_k(m) \gg \alpha_k \tilde{R}_k(m)$ , which is motivated in Section IV-C.

Denote the optimal instantaneous relative rates  $\{\tilde{B}_{q,k}^*(m)\}$ . The optimal allocation has the following property: given the allocation of all other clusters, the  $q$ th cluster is allocated to the user that achieves the highest utility increase. In other words

$$\begin{aligned} k_q^*(m) &= \arg \max_{k'} \sum_k U_k^\kappa \\ &\times \left( \alpha_k \sum_{i \neq q} \tilde{B}_{i,k}^*(m) + \delta_{k-k'} \frac{\alpha_k R}{R_k} C_{q,k}(m) + \beta_k \bar{B}_k(m) \right) \\ &= \arg \max_{k'} \sum_k U_k^\kappa \\ &\times \left( \frac{\alpha_k}{R_k} \sum_{i \neq q} \tilde{R}_{i,k}^*(m) + \delta_{k-k'} \frac{\alpha_k R}{R_k} C_{q,k}(m) + \frac{\beta_k}{R_k} \bar{R}_k(m) \right) \end{aligned}$$

where  $\delta_x$  is the Kronecker delta. Denote the potential relative rate in cluster  $q$  for user  $k$ ,  $\tilde{B}_{q,k}(m) = (R/R_k) C_{q,k}(m)$ . The

$$\Pr(U_k = S + t | U_{k-1} = S + r) = \begin{cases} \frac{\binom{S+r}{S-(t-r)} \binom{Q-S-r}{t-r}}{\binom{Q}{S}}, & \text{if } r \leq t \text{ and } t - r \leq S \\ 0, & \text{otherwise} \end{cases}$$

relaxing assumption that  $(\beta_k/R_k)\bar{R}_k(m) \gg \alpha_k\check{R}_k(m) \geq (\alpha_k/R_k)\sum_{i \neq q}\check{R}_{i,k}(m)$  means that the user allocations during the current block have little impact on the new average rates. Hence

$$k_q^*(m) = \arg \max_{k'} \sum_k U_k^\kappa \left( \delta_{k-k'} \alpha_k \check{B}_{q,k}(m) + \beta_k \bar{B}_k(m) \right). \quad (9)$$

This implies that the scheduling problem is separable and can be solved for each cluster separately. Note that  $\delta_{k-k_q} \check{B}_{q,k}(m) = \check{B}_{q,k}^*(m)$ .

Now, consider the problem of scheduling a user for the  $q$ th cluster, based on (9). For all but the scheduled user,  $\check{B}_{q,k}(m) = 0$ . Therefore, the user allocation that maximizes the sum utility in (9) is to choose the user with highest marginal utility increase if scheduled

$$\begin{aligned} k_q^*(m) &= \arg \max_k U_k^\kappa \left( \alpha_k \check{B}_{q,k}(m) + \beta_k \bar{B}_k(m) \right) \\ &\quad - U_k^\kappa \left( \beta_k \bar{B}_k(m) \right) \\ &\approx \alpha_k \check{B}_{q,k}(m) \frac{d}{d\hat{B}_k} U_k^\kappa \left( \beta_k \bar{B}_k(m) \right) \end{aligned} \quad (10)$$

where the first-order linear approximation is used with the assumption again that the rates of the current block perturbs the average rates marginally. From (8), the derivative is easily found as

$$\frac{d}{d\hat{B}_k} U_k^\kappa(\hat{B}_k) = \frac{R_k \beta_k^\kappa}{\alpha_k} \hat{B}_k^{-\kappa}. \quad (11)$$

Combining (10) and (11) gives (6), which concludes the proof.

*Scheduling Metric Simplifications:* The scheduling metric

$$k_q^*(m) = \arg \max_k \left\{ \frac{C_{q,k}(m)}{\bar{B}_k(m)^\kappa} \right\} \quad (12)$$

can be simplified by setting the target bit rate  $R_k$  equal for all users, i.e.,  $R_k = R_0$ . Then,  $\bar{B}_k(m) = \bar{R}_k(m)$ , and (12) simplifies to

$$k_q^*(m) = \arg \max_k \left\{ \frac{C_{q,k}(m)}{\bar{R}_k(m)^\kappa} \right\}.$$

Setting  $\kappa = 0$  gives the maximum throughput scheduler, which confirms the result from Appendix II-A

$$k_q^*(m) = \arg \max_k \{C_{q,k}(m)\}.$$

Setting  $\kappa = 1$  gives the PF scheduling metric

$$k_q^*(m) = \arg \max_k \left\{ \frac{C_{q,k}(m)}{\bar{R}_k(m)} \right\}.$$

Letting  $\kappa \rightarrow \infty$ , the user  $k$  with lowest  $\bar{R}_k(m)$  is scheduled

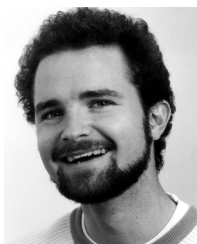
$$k_q^*(m) = \lim_{\kappa \rightarrow \infty} \arg \max_k \left\{ \frac{C_{q,k}(m)}{\bar{R}_k(m)^\kappa} \right\} = \arg \min_k \{ \bar{R}_k(m) \}$$

since  $C_{q,k}(m)$  is bounded. This corresponds to the max-min scheduler.

## REFERENCES

- [1] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.
- [2] M. Sternad, T. Otosson, A. Ahlén, and A. Svensson, "Attaining both coverage and high spectral efficiency with adaptive OFDM," in *Proc. IEEE Veh. Technol. Conf.*, Oct. 2003, pp. 2486–2490.
- [3] W. Wang, T. Otosson, M. Sternad, A. Ahlén, and A. Svensson, "Impact of multiuser diversity and channel variability on adaptive OFDM," in *Proc. IEEE Veh. Technol. Conf.*, Oct. 2003, pp. 547–551.
- [4] R. Laroia, J. Li, S. Rangan, and M. Srinivasan, "Enhanced opportunistic beamforming," in *Proc. IEEE Veh. Technol. Conf.*, Oct. 2003, pp. 1762–1766.
- [5] L. J. Cimini, B. Daneshrad, and N. R. Sollenberger, "Clustered OFDM with transmitter diversity and coding," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 1996, pp. 703–707.
- [6] D. Gesbert and M.-S. Alouini, "How much feedback is multi-user diversity really worth?," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2004, vol. 1, pp. 234–238.
- [7] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, Jan. 1997.
- [8] G. Barriac and J. Holtzman, "Introducing delay sensitivity into the proportional fair algorithm for CDMA downlink scheduling," in *Proc. Int. Symp. Spread-Spectrum Tech. Appl.*, 2002, pp. 652–656.
- [9] K. Kim, H. Kim, and Y. Han, "A proportional fair scheduling algorithm with QoS and priority in 1XEV-DO," in *Proc. IEEE Int. Conf. Pers., Indoor, Mobile Radio Commun.*, Sep. 2002, pp. 2239–2243.
- [10] W. Anchun, X. Liang, Z. Shidong, X. Xibin, and Y. Tan, "Dynamic resource management in the fourth generation wireless systems," in *Proc. IEEE Int. Conf. Commun. Technol.*, Apr. 2003, pp. 1095–1098.
- [11] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.
- [12] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. IEEE Veh. Technol. Conf.*, May 2000, pp. 1085–1089.
- [13] Y. Zhang and K. B. Letaief, "Multiuser subcarrier and bit allocation with adaptive cell selection for OFDM transmission," in *Proc. IEEE Int. Conf. Commun.*, May 2002, pp. 861–865.
- [14] Y. Zhang and K. B. Letaief, "Optimizing power and resource management for multiuser MIMO/OFDM systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2003, pp. 179–183.
- [15] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems," *IEEE Trans. Broadcast.*, vol. 49, no. 4, pp. 362–370, Dec. 2003.
- [16] G. Song and Y. Li, "Adaptive subcarrier and power allocation in OFDM based on maximizing utility," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2003, pp. 905–909.
- [17] L. Badia, M. Lindstrom, J. Zander, and M. Zorzi, "Demand and pricing effects on the radio resource allocation of multimedia communication systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2003, pp. 4116–4121.
- [18] Z. Jiang, H. Mason, B. J. Kim, N. K. Shankaranarayanan, and P. Henry, "A subjective survey of user experience for data applications for future cellular wireless networks," in *Proc. IEEE Symp. Appl. Internet*, Jan. 2001, pp. 167–175.
- [19] N. Enderle and X. Lagrange, "User satisfaction models and scheduling algorithms for packet-switched services in umts," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2003, pp. 1704–1709.
- [20] P. Svedman, S. K. Wilson, L. J. Cimini, and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDM," in *Proc. IEEE Veh. Technol. Conf.*, May 2004, vol. 4, pp. 1878–1882.
- [21] Y. Li, J. C. Chang, and N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high-rate data wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 1233–1243, Jul. 1999.

- [22] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [23] S. Shenker, "Fundamental design issues for the future internet," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1176–1188, Sep. 1995.
- [24] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding—Part II: Coding results," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2595–2604, Oct. 1995.
- [25] Spatial channel model for multiple input multiple output simulations (V6.1.0), TR 25.996 3GPP, 2003 [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/25996.htm>
- [26] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol. 29, pp. 1162–1163, Jun. 1993.
- [27] P. Svedman, S. K. Wilson, and B. Ottersten, "A QOS-aware proportional fair scheduler for opportunistic OFDM," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2004, vol. 1, pp. 558–562.



**Patrick Svedman** (S'03) received the M.Sc. degree in electrical engineering in 1999 from the Royal Institute of Technology (KTH), Stockholm, Sweden, where he is currently working toward the Ph.D. degree.

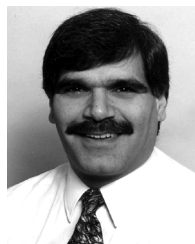
In 1999–2001, he worked at Nokia Networks with digital ASIC design for 3G. His research interests include wireless multiuser communications with an emphasis on OFDMA and multiple-antenna systems.



**Sarah Kate Wilson** (M'87–SM'99) received the AB degree in mathematics from Bryn Mawr College, Bryn Mawr, PA, in 1979, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1987 and 1994, respectively.

She has worked in industry as a programmer/analyst, research engineer, and project manager. She was an Assistant Professor at Purdue University and the Lulea University of Technology (LTU) and was a guest professor/researcher at the Royal Institute of Technology in Sweden. Currently, she is an Assistant Professor and a David Packard Fellow in the Electrical Engineering Department, Santa Clara University, Santa Clara, CA.

Dr. Wilson has served as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and is currently an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE COMMUNICATIONS LETTERS.



**Leonard J. Cimini, Jr.** (S'77–M'82–SM'89–F'00) received the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1982.

He was with AT&T, first in Bell Labs and then AT&T Labs, for 20 years. His research has concentrated on lightwave and wireless communications, and his main emphasis has been on devising techniques for overcoming the bit-rate limitations imposed by the radio environment. In this context, he pioneered the application of OFDM to the emerging field of wireless communications. He has been a Professor in the Electrical and Computer Engineering Department, University of Delaware, Newark, since 2002.

Dr. Cimini has been very active within the IEEE, and he was the founding Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS: WIRELESS COMMUNICATIONS SERIES. Currently, among other activities, he is the Chair of the Emerging Technologies Committee of ComSoc and a Member At-Large of the Board of Governors. He was elected a Fellow of the IEEE in 2000 for contributions to the theory and practice of high-speed wireless communications.



**Björn Ottersten** (S'87–M'89–SM'99–F'04) was born in Stockholm, Sweden, in 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986. In 1989, he received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

He has held research positions at the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, and the Katholieke Universiteit Leuven, Leuven, Belgium. During 1996–1997, he was Director of Research at ArrayComm Inc, San Jose, CA. In 1991 he was appointed Professor of Signal Processing at the Royal Institute of Technology (KTH), Stockholm, Sweden, and he is currently Dean of the School of Electrical Engineering at KTH. From 1992 to 2004, he was Head of the Department for Signals, Sensors, and Systems at KTH. He is also a visiting professor at the University of Luxembourg. His research interests include wireless communications, stochastic signal processing, sensor array processing, and time series analysis.

Dr. Ottersten has served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and a member of the editorial board of the *EURASIP Journal of Applied Signal Processing*. He is currently editor-in-chief of the *EURASIP Signal Processing Journal* and a member of the editorial board of *IEEE Signal Processing Magazine*. He received the Signal Processing Society Paper Award in 1993 and 2001.