A Feedback Quantization Scheme Leveraging Fairness and Throughput for Heterogeneous Multi-User Diversity Systems

Yusuf Soydan and Çağatay Candan

Abstract—We propose a feedback quantization scheme for downlink multiuser diversity systems. The scheme is designed for heterogeneous users, i.e., users operating under different levels of mean signal-to-noise ratios (SNRs). The proposed scheme equally distributes the channel access time to user clusters and, at the same time, maximizes the sum-rate capacity under equal-channel-access-time constraint. The method is compared to a normalized SNR scheduler, full-feedback scheduler, and finite-rate feedback scheduler without fairness consideration to illustrate the performance gaps between different modes of operation.

Index Terms—Fairness, feedback quantization, multiuser diversity, sum capacity.

I. Introduction

Multiuser diversity can yield a significant system-wide throughput for systems containing many users operating under severe channel fluctuations [1]–[3]. Channel fluctuations can be due to fading, as in Rayleigh-fading channels, or can artificially be generated by introducing some randomness at the transmitter side, as in opportunistic communication [3]. Irrespective of how channel fluctuations arise, multiuser diversity utilizes the diversity of signal-to-noise ratios (SNRs) across the system by scheduling the best user in the system at all times. This approach has been shown to maximize the sum rate of the system [1], [2].

In this paper, we focus on the *feedback reduction* and *fairness* issues of multiuser diversity systems. The goal of multiuser diversity is to provide service to the best user (i.e., the user with the highest SNR) among all users. To achieve this goal, users need to estimate the channel and feed back their SNR values to the base station [3]. In practice, the instantaneous SNR of the users has to be quantized to finitely many levels.

Different schemes for SNR quantization have been studied in the literature. In [4]–[6], the effect of the number of quantization bits on throughput and delay has been studied. In [7], the probability of incorrect identification of the best user and its effect on throughput has been studied. In [8], the quantization levels to maximize system throughput have been given for homogeneous users, i.e., when all users have the same mean SNR value. In [9], different schemes for homogeneous users have been compared. Different from existing studies in the literature, we present a quantized feedback system for users with varying levels of mean SNR, i.e., heterogeneous users.

Fairness is another issue of concern for multiuser diversity systems. When users admitted in the system have significantly different levels of mean SNR, multiuser diversity schemes naturally favor users with

high mean SNR values over those with low mean SNR values. It should be noted that the fairness and sum-rate maximization goals are antithetical. Therefore, joint optimization of both criteria is not possible. However, in practice, it is desirable to present a service guarantee to the users admitted in the system. As an example, consider two groups of users, each having ten users; the first group has a mean SNR of 5 dB, whereas the second group has 10-dB mean SNR. When the sum-rate maximization goal is exercised, the group with the lower mean SNR value receives very little service. On the opposite extreme, when all users are scheduled on a round-robin basis, the channel access time is equally distributed among the users, but the system operates at a poor cumulative throughput. A reasonable leverage between throughput and fairness goals is desirable.

In this paper, we propose a clustering-based scheduling method. The system is designed by imposing a constraint, which introduces fairness in the channel utilization time of clusters. The fairness criteria studied in this paper is the egalitarian distribution of channel access time among the clusters. This goal allows concurrently setting data services with different qualities of service requirements, which are called differentiated services. By adjusting the parameters of the system (under the fair channel utilization constraint), the sum capacity is maximized. The system operates as follows: The users are first grouped into clusters. The users of each cluster are given a set of thresholds for quantization of their instantaneous SNR levels. Users quantize and feed the instantaneous SNR levels back to the base station. After reception of feedback from all users, the base station schedules service in two stages. In the first stage, the base station randomly selects a cluster among *the best clusters*. Then, the base station selects a user among *the best users* in the selected cluster. The details of the process are explained in the next section. The thresholds of each cluster are adjusted, so that the probability of having one or more users exceeding the kth threshold of each cluster is the same. This constraint establishes the fairness leg of the scheme. The sum-rate maximization goal is achieved by optimizing over the probability value for exceeding the kth threshold. With this scheme, the channel access times or channel utilization of each cluster is guaranteed to be the same. The system is shown to operate at a significantly higher throughput over the round-robin system.

This paper is organized as follows: Section II describes the system model and the proposed method. Section III describes the method for the threshold selection and explains how the tradeoff between fairness and throughput is resolved according to the proposed methodology. Section IV presents the results of numerical comparisons, and this paper concludes with discussions and further remarks.

II. System Model

We assume that users have Rayleigh-fading channels with possibly different mean SNR values. The SNR of each user is denoted by Γ and taken as exponentially distributed. The cumulative distribution function for SNR is then $F_\Gamma(\gamma) = 1 - e^{-\gamma/\overline{\gamma}}$, where $\overline{\gamma}$ is the mean value. We assume that the system contains T clusters. Cluster-(t) contains $M^{(t)}$ users, and all users in a cluster are assumed to be independent and identically distributed, i.e., the users have the same mean SNR, which is denoted by $\overline{\gamma}^{(t)}$. As in [4] and [8], the aggregate capacity is calculated with the assumption that the users are to transmit with the rate of $\log_2(1 + \gamma)$.

The base station assigns $K - 1$ thresholds to each cluster, as shown in Fig. 1. Users in a cluster quantize their instantaneous SNR values according to the thresholds (q_k's) and feed the result back to the base station. The quantization thresholds required by the users are
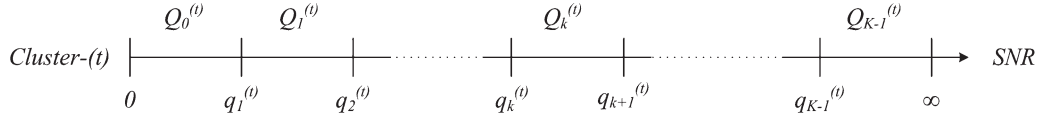
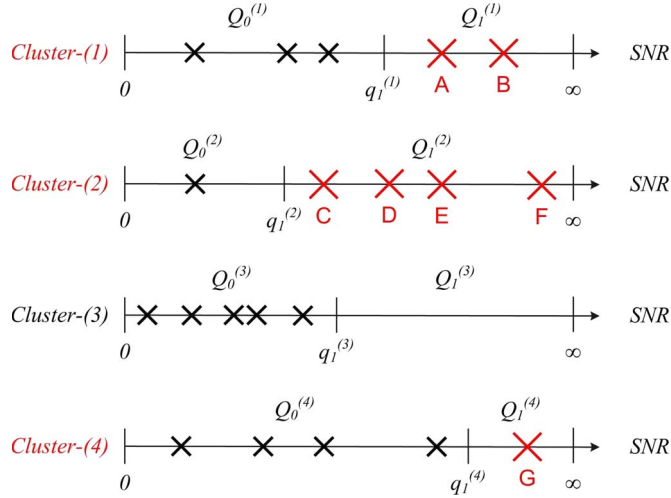Fig. 1.   Quantization levels of instanteneous SNR.

Fig. 2.   Proposed scheme with four clusters, each having a single threshold.

calculated by the base station using the proposed technique and announced to the users.

Upon reception of the quantized instantaneous SNR values, the base station operates as given here.

1) The base station examines the received feedback from all users and determines the users with the largest quantized SNR value. We call the users declaring the largest quantized SNR level in a time slot as the best users of that slot.
2) The base station constructs the list of clusters that the best users belong to.
3) The base station randomly selects (with equal probability) a cluster from the cluster list constructed in step 2.
4) The base station randomly selects (with equal probability) one of the best users in the cluster selected in step 3 and schedules service to this user.

We present the following example to illustrate the scheme. A system with four clusters each using a single threshold is shown in Fig. 2. The users with the labels $A-G$ are candidates for the service, i.e., these users form the set of best users. The base station randomly selects one cluster among the first, second, and fourth clusters, and then randomly selects one of the best users from the selected cluster. It should be noted that, if the fourth cluster is selected, then user G surely receives the service since there is no other user competing for the service in this cluster. Note that the users of the proposed system are not aware of the clusters that they belong to, but they only need the assigned threshold values to implement the scheme.

We present a method for the selection of thresholds under this regime. Different from existing methods in the literature, the proposed method is explicitly designed for heterogeneous users. Our goal is to design a method that benefits from multiuser diversity and, at the same time, has a service guarantee that can be given to each user. With the proposed system, the thresholds are selected such that the clusters have equal channel access times and each user in each cluster takes an equal share of cluster channel access time.

III. FAIR SCHEDULING WITH THROUGHPUT MAXIMIZATION

We examine the fairness issue first. Following the discussion of fairness, the sum-rate maximization goal is described.

Fairness: We focus on cluster-(t), whose quantization levels are illustrated as in Fig. 1. First, we define the no-feedback event. Cluster-(t) is in the no-feedback state when all of its users have instantaneous SNR that is smaller than the first threshold. The probability of this event is denoted by P_{01}, i.e.,

$$P_{01} = \left[F_{\Gamma^{(i)}} \left(q_1^{(t)} \right) \right]^{M^{(t)}}$$

$$= \left[1 - e^{-q_1^{(t)}/\overline{\gamma}^{(t)}} \right]^{M^{(t)}}. \tag{1}$$

It can be noted that the no-feedback event is equivalent to the event of the maximum SNR value of $M^{(t)}$ users in cluster-(t) being less than $q_1^{(t)}$. It can also be noted that a given P_{01} value uniquely determines $q_1^{(t)}$ through the relation

$$q_1^{(t)} = -\overline{\gamma}^{(t)} \ln \left(1 - (P_{01})^{1/M^{(t)}} \right). \tag{2}$$

The single-threshold scheme with two quantization levels $Q_0^{(t)}$ and $Q_1^{(t)}$, as shown in Fig. 1, using the proposed methodology, can be designed as follows: We impose a fixed no-feedback probability value to each cluster. By setting the no-feedback probability to a fixed value (i.e., $P_{01} = 0.3$), the maximum SNR value of each cluster has the same probability to fall in the first interval (Q_0) shown in Fig. 1. According to the proposed method, the base station randomly selects a cluster from the list of best clusters (step 3 of the method); therefore, the channel utilization of each cluster is the same for all clusters. This constraint provides the fairness leg of the system in terms of cluster access probabilities.

Note that it is possible to define other tradeoffs to resolve the conflict between fairness and sum-capacity goals. With the proposed system, the channel access times of each cluster are guaranteed to be the same in the long run. Provided that each cluster contains the same number of users, the channel utilization of each user in all clusters is also the same. It should be noted that any value for the no-feedback probability (instead of $P_{01} = 0.3$) can be used without any change in the fairness argument.

The freedom in the choice of P_{01} can be used to optimize the sum capacity. This results in a single-variable optimization problem, which can be written as follows:

$$\widehat{P}_{01} = \arg \max_{P_{01}} C_{\text{sum}}(P_{01}) \qquad \text{subject to } 0 \leq P_{01} < 1. \tag{3}$$

Here, C_{sum} is the sum capacity of the system whose derivation is given after the conclusion of the fairness discussion.

When a second threshold $q_2^{(t)}$ is added to the system, the number of quantization levels increases to three, i.e., $Q_0^{(t)}$, $Q_1^{(t)}$, and $Q_2^{(t)}$.

We denote the probability of the maximum instantaneous SNR of cluster-(t) falling in the interval $[q_1^{(t)}, q_2^{(t)})$ as P_{12}, i.e.,

$$P_{12} = \left[F_{\Gamma^{(i)}} \left(q_2^{(t)} \right) \right]^{M^{(t)}} - \left[F_{\Gamma^{(i)}} \left(q_1^{(t)} \right) \right]^{M^{(t)}}$$

$$= \left[1 - e^{-q_2^{(t)}/\overline{\gamma}^{(t)}} \right]^{M^{(t)}} - \left[1 - e^{-q_1^{(t)}/\overline{\gamma}^{(t)}} \right]^{M^{(t)}}. \quad (4)$$

As before, we propose to fix P_{12} to a constant value across the clusters. The second threshold can be calculated from (4) as follows:

$$q_2^{(t)} = -\overline{\gamma}^{(t)} \ln \left(1 - (P_{01} + P_{12})^{1/M^{(t)}} \right). \quad (5)$$

It can be noted from (2) and (5) that threshold calculations are coupled. In other words, P_{01} affects both the first and second thresholds. The coupling of coefficients significantly complicates the optimization procedure. To simplify the optimization process, we use a suboptimal term-by-term optimization procedure. We first find the optimal threshold for the single-threshold case by solving (3) and then substitute the optimal value found $(\widehat{P}_{01})$ in the optimization problem having two thresholds

$$\widehat{P}_{12} = \arg \max_{P_{12}} C_{\text{sum}} \left(\widehat{P}_{01}, P_{12} \right)$$

$$\text{subject to} \quad 0 \leq P_{12} < 1 - \widehat{P}_{01}. \quad (6)$$

It should also be noted that the optimization problem described by (3) and (6) is a simple one-variable optimization problem over a finite interval. Simple line search techniques, such as the bisection rule in a finite interval, are sufficient to find the optimal value. Once the optimal probability values are found, the thresholds are calculated using (2) and (5). This simple approach is in contrast with the significantly more difficult numerical optimization problem for thresholds over a semiinfinite interval of $[0, \infty)$.

Sum-Capacity Relation: Let cluster-(r) be the reference cluster, $r \in \{1, 2, 3, \ldots, T\}$, $R^{(r)}$ be the event in which a user (called reference user) in the reference cluster is receiving the service when there are no other clusters in the system, and $\Gamma^{(r)}$ be the random variable for the instantaneous SNR of the reference user. The probability of the reference user in the reference cluster with the instantaneous SNR in the kth quantization level to receive the service in the absence of other clusters is then

$$\mathbf{Pr} \left(R^{(r)} \text{ and } \Gamma^{(r)} \in Q_k^{(r)} \right)$$

$$= \sum_{m^{(r)}=0}^{M^{(r)}-1} \frac{1}{m^{(r)}+1} \binom{M^{(r)}-1}{m^{(r)}} \times \left[\mathbf{Pr} \left(\Gamma^{(r)} \in Q_k^{(r)} \right) \right]^{m^{(r)}}$$

$$\times \left[\mathbf{Pr} \left(\Gamma^{(r)} \in \bigcup_{l<k} Q_l^{(r)} \right) \right]^{M^{(r)}-m^{(r)}-1} \quad (7)$$

where

$$\mathbf{Pr} \left(\Gamma^{(r)} \in Q_k^{(r)} \right) = F_{\Gamma^{(r)}} \left(q_{k+1}^{(r)} \right) - F_{\Gamma^{(r)}} \left(q_k^{(r)} \right) \quad (8)$$

$$\mathbf{Pr} \left(\Gamma^{(r)} \in \bigcup_{l<k} Q_l^{(r)} \right) = F_{\Gamma^{(r)}} \left(q_k^{(r)} \right). \quad (9)$$

The variable $m^{(r)}$ in (7) is the number of cluster-(r) users, in addition to the reference user, that are in the same interval as the reference user. $M^{(r)}$ is the total number of users in cluster-(r).

When there are other clusters competing with cluster-(r), the probability of service for the reference user changes as follows: Assume that there is a competing cluster, which is denoted by cluster-(i), and let $I_k^{(i)}$ be an indicator function taking values 0 or 1. $I_k^{(i)}$ takes the value of 0, when all users in cluster-(i) have SNRs that are less than the kth quantization level of cluster-(i). In this case, none of the cluster-(i) users presents a competition to the reference user. The probability of this event is given as follows:

$$\mathbf{Pr} \left(I_k^{(i)} = 0 \right) = \left[\mathbf{Pr} \left(\Gamma^{(i)} \in \bigcup_{l<k} Q_l^{(i)} \right) \right]^{M^{(i)}}. \quad (10)$$

The indicator function $I_k^{(i)}$ takes the value of 1 when at least one user of cluster-(i) has SNR value falling in the interval $Q_k^{(i)}$. In this case, there is at least one competing party from cluster-(i) for the service. The probability of this event is

$$\mathbf{Pr} \left(I_k^{(i)} = 1 \right) = \sum_{m^{(i)}=1}^{M^{(i)}} \binom{M^{(i)}}{m^{(i)}} \mathbf{Pr} \left[\Gamma^{(i)} \in Q_k^{(i)} \right]^{m^{(i)}}$$

$$\times \left[\mathbf{Pr} \left(\Gamma^{(i)} \in \bigcup_{l<k} Q_l^{(i)} \right) \right]^{M^{(i)}-m^{(i)}}. \quad (11)$$

If event A is defined as scheduling the service to the reference user, the probability of A when the user has instantaneous SNR falling in the kth quantization level can then be calculated as follows:

$$\mathbf{Pr} \left(A | \Gamma^{(r)} \in Q_k^{(r)} \right)$$

$$= \mathbf{Pr} \left(R^{(r)} \text{ and } \Gamma^{(r)} \in Q_k^{(r)} \right) \sum_{i^{(1)}=0}^{1}$$

$$\times \left(\mathbf{Pr} \left(I_k^{(1)} = i^{(1)} \right) \times \ldots \times \sum_{i^{(r-1)}=0}^{1} \right.$$

$$\times \left(\mathbf{Pr} \left(I_k^{(r-1)} = i^{(r-1)} \right) \times \sum_{i^{(r+1)}=0}^{1} \right.$$

$$\times \left(\mathbf{r} \left(I_k^{(r+1)} = i^{(r+1)} \right) \times \ldots \times \sum_{i^{(T)}=0}^{1} \right.$$

$$\left. \left. \left. \times \left(\mathbf{Pr} \left(I_k^{(T)} = i^{(T)} \right) \frac{1}{1 + \sum_{\substack{z=1 \\ z \neq r}}^{T} i^{(z)}} \right) \right) \right) \right).$$

$$(12)$$

Note that (12) exhaustively counts all possible ways of the reference user receiving service in the presence of $T-1$ competing clusters.

Equipped with the probability of receiving service for the reference user, we can immediately write the following capacity relation for the
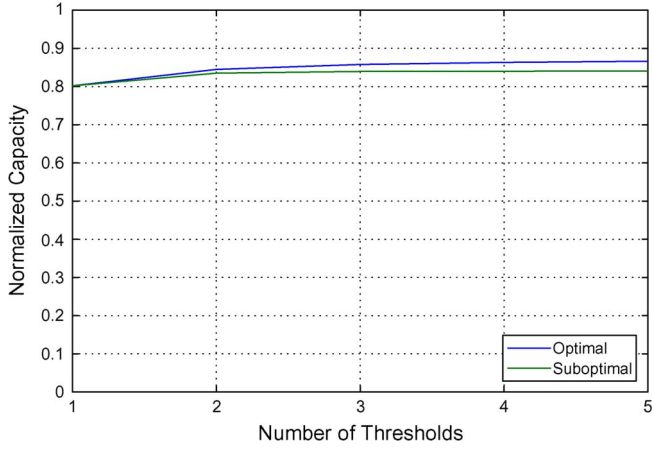
Fig. 3. Sum capacity of the proposed scheme with optimal and suboptimal threshold selection. The capacity values are normalized to those of the full-feedback (unquantized) scheme.

service delivered to the reference user:

$$C_{\text{per-user}}^{(r)} = \sum_{k=0}^{K-1} \mathbf{Pr}\left(A|\Gamma^{(r)} \in Q_k^{(r)}\right)$$

$$\int_{Q_k^{(r)}} \frac{1}{2} \log_2\left(1+\gamma\right) f_{\Gamma^{(r)}}\left(\gamma\right) d\gamma. \quad (13)$$

It should be noted that the unit of (13) is bits per channel use per user. This relation is closely connected to the relations presented for homogeneous users by Floren *et al.* [8].

The aggregate capacity of the system, which is the addition of per-user capacities, is equal to

$$C_{\text{sum}} = \sum_{t=1}^{T} M^{(t)} C_{\text{per-user}}^{(t)}. \quad (14)$$

Channel Utilization: The channel access probability for the reference user can be calculated as follows:

$$P_{\text{access-user}}^{(r)} = \sum_{k=0}^{K-1} \mathbf{Pr}\left(A|\Gamma^{(r)} \in Q_k^{(r)}\right) \int_{Q_k^{(r)}} f_{\Gamma^{(r)}}(\gamma)d\gamma. \quad (15)$$

The access probability of cluster-(r) is then simply

$$P_{\text{access-cluster}}^{(r)} = M^{(r)} \times P_{\text{access-user}}^{(r)}. \quad (16)$$

This completes the discussion of sum-rate and access probability calculations for the proposed system. In the next section, we examine the performance of the proposed method.

IV. NUMERICAL COMPARISONS

On the threshold optimization: The proposed method uses a suboptimal but efficient method for threshold optimization. As described earlier, the suboptimal method implements $K-1$ (which is the number of thresholds) 1-D search operations. However, the optimal way to find the thresholds is through the joint optimization of $K-1$ thresholds. In Fig. 3, we compare the aggregate capacity of the systems whose thresholds are found with optimal and suboptimal methods. In this comparison, the first cluster has five users with 5-dB mean SNR, and the second cluster has ten users with 10-dB mean SNR. The aggregate capacity values for different numbers of thresholds are normalized with the full-feedback aggregate capacity, which is the capacity of
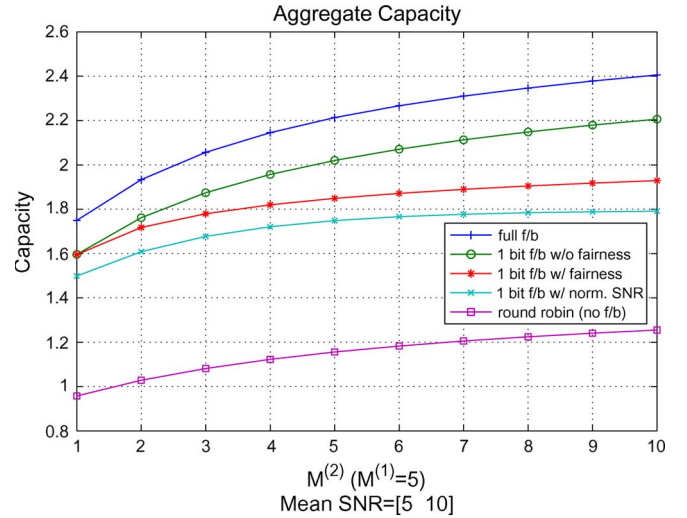

Fig. 4. Aggregate capacity comparison for different systems.

the unquantized system. From Fig. 3, we observe that both methods yield the same performance for the case of a single-threshold system as expected, and when there is more than one threshold, the joint optimization yields marginally better performance than the suboptimal method at an increased cost of optimization.

Comparisons with other methods: Our goal is to examine the performance gaps between different modes of operation and the proposed operational regime. In this comparison, the first cluster operates at 5-dB mean SNR and contains five users. The second cluster operates at 10-dB mean SNR, and the number of users in this cluster varies from 1 to 10.

Fig. 4 shows the sum-rate capacity of five different methods. The first method is the full-feedback scheme. This system assumes that all users present their SNR feedback without quantization. The aggregate capacity of this method is an upper bound for all quantization-based systems with or without fairness consideration. The other methods use 1-bit quantization. The second method does not consider any fairness issue, and its sole optimization goal is to maximize the aggregate capacity. The thresholds required for this method can be set with any single-objective nonlinear optimization routine. We present this method to serve as an upper bound for all 1-bit quantization schemes with or without fairness consideration. The third scheme is the proposed technique. The fourth method is called the normalized SNR method and uses the mean SNR value of each user as the corresponding threshold [8]. With this scheme, when the instantaneous SNR of a user is above the mean SNR, the user requests service by sending a 1-bit feedback. It should be noted that, with this method, users with different mean SNR values are identically treated. Hence, this method can suffer from sum-capacity loss if the variation of the mean SNR values is significant. The fifth method is the round-robin scheme, which schedules users one after another without considering the channel conditions.

It can be noted from Fig. 4 that the sum rate reduces by 0.2 bit/channel use when feedback is reduced from full feedback to 1 bit. As stated before, the curve for 1-bit feedback without any fairness consideration should be interpreted as the capacity bound for 1-bit schemes with fairness considerations. There is another 0.2-bit/channel-use sacrifice when described notion of fairness is imposed on the 1-bit quantization system. The normalized SNR scheme has 0.1-bit/channel-use less capacity than the proposed scheme. As expected, the capacity loss increases when the disparity between the clusters increases.

In Fig. 5, the cluster channel access probabilities of each system are given. The cluster channel access probabilities serve as the service guarantee that can be provided to each cluster. It can be noted that,
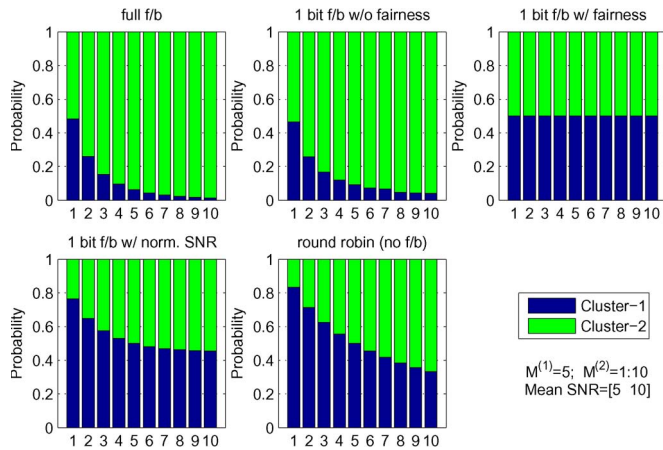
Fig. 5. Cluster channel access probabilities for the systems compared in Fig. 4.
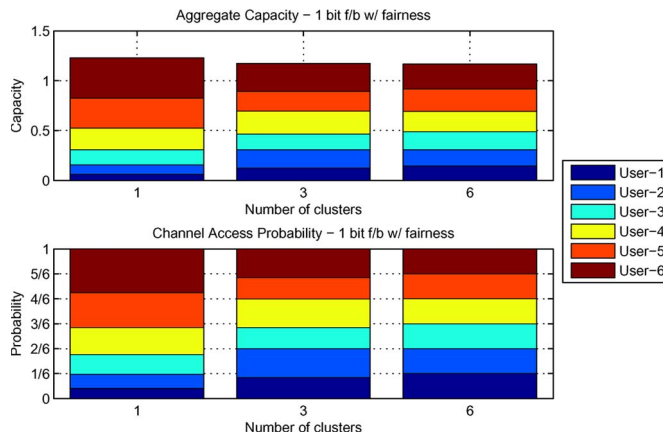

Fig. 6. Aggregate capacity and channel access probability comparison for multiple clusters.

for the systems without fairness consideration (full-feedback method and 1-bit feedback without fairness consideration), the stronger cluster monopolizes the channel utilization, and the users of a weak cluster can receive very little service. The proposed system assigns an equal time share to both clusters by design and presents a significantly improved overall throughput. This figure clearly illustrates the service time fairness of each cluster when the proposed method is utilized and loss of fairness in the sense of cluster service time for various other scenarios. Further comparisons can be found in [10].

On the clustering process: The proposed method aims to set the feedback thresholds, so that the channel access times are equally distributed among the clusters. As an illustrative example, let us assume that there are six users with different mean SNR values in the system. When six users are assigned to six different clusters (with one user per cluster), then the proposed method equally distributes the channel access time to each user. In contrast to this assignment, when all six users are assigned to a single cluster, then the proposed system sets the thresholds of each user to maximize the sum capacity, and the user with the highest mean SNR value gets the most channel access time. A reasonable approach is to group users with similar mean SNR values in the same cluster and create as many clusters as required by the issued quality of service guarantee. For example, users with high quality of service requirements can individually be placed in a cluster or can be combined with a few other users with similar requirements. This would enable frequent channel access to a select group of users.

In Fig. 6, we study the mentioned example to further illustrate the details of user clustering. Here, there are six users, and user-i has

i-dB mean SNR value. As can be seen from Fig. 6, both the capacity and access time of individual users vary, depending on the clustering process. User-6, which is the user with the strongest channel, gets a significant share of channel access time when there is a single cluster, but the same user gets an equal share for the channel access when six clusters are utilized.

Fig. 6 can also be interpreted as an illustration of the conflicting goals of maximizing the aggregate capacity and providing fairness in terms of channel access time. From the view point of aggregate capacity maximization, the number of clusters should be as few as possible; from the view point of channel access time, the number of clusters should be as high as possible. The conflict is resolved by trading off one goal with another.

V. CONCLUSION

In this paper, we have presented a practical feedback quantization scheme to realize fair scheduling with throughput maximization. Different from various methods in the literature, the case of heterogeneous users having different levels of mean SNR has been examined. With the proposed method, the quantization thresholds have been selected such that the user clusters have equal channel access time and each user in each cluster takes an equal share of the cluster channel access time. Under this condition, the parameters of the system are optimized to provide the maximum sum rate. It has been shown that the system provides a significant gain over round-robin system and provides a gain comparable to the systems that do not consider fairness in their designs.

The proposed system can also be used to provide different levels of quality of service to a group users by adjusting the population of each cluster. When there are five users with the same or close-enough mean SNR values, placing each user in a cluster results in an equal distribution of channel access time for each user. Placing a single user in cluster-1 and the rest in cluster-2 results in 50 channel access time for user-1 and 12.5 channel access time for the other users. Such an operation allows concurrently setting services with different quality-of-service requirements.

REFERENCES

[1] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE ICC*, Jun. 1995, vol. 1, pp. 331–335.

[2] D. N. C. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory*, 1997, p. 27.

[3] 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.

[4] D. Gesbert and M. Alouini, "How much feedback is multiuser diversity really worth?" in *Proc. Int. Conf. Commun.*, Jun. 2004, pp. 234–238.

[5] M. Johansson, "Benefits of multiuser diversity with limited feedback," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2003, pp. 155–159.

[6] Y. Al-Harti, A. Tewfik, and M. S. Alouini, "Multiuser diversity with quantized feedback," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 330–337, Jan. 2007.

[7] A. P. T. Lau and F. R. Kchinschang, "Feedback quantization strategies for multiuser diversity systems," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1386–1400, Apr. 2007.

[8] F. Floren, O. Edfors, and B. A. Molin, "The effect of feedback quantization on the throughput of a multiuser diversity scheme," in *Proc. IEEE GLOBECOM*, Dec. 2003, vol. 1, pp. 497–501.

[9] F. Floren, O. Edfors, and B. A. Molin, "Throughput analysis of three multiuser diversity scheme," in *Proc. IEEE Veh. Technol. Conf.*, 2003, pp. 2686–2690.

[10] Y. Soydan, "Cluster based user scheduling schemes to exploit multiuser diversity in wireless broadcast channels," M.S. thesis, Dept. Elect. Electron. Eng., Middle East Tech. Univ., Ankara, Turkey, 2008.