

Differentiated Rate Scheduling for Gaussian Broadcast Channels

MASOUD SHARIF, AMIR F. DANA and BABAK HASSIBI

Department of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125

Abstract—In this paper, we consider a fading broadcast channel where users have different rate demands. In particular, we assume users are divided into M groups, each group of which requires the same rate, and where the ratio of the rates of the groups are given. The transmitter would like to maximize the throughput (sum of the rates to all users) while maintaining the rational rate constraints. In general, this problem appears to be computationally intractable since the ergodic capacity region is described as the convex hull of (an infinite) set of rates. In this paper, we therefore, focus on the asymptotic regime of many users (large n) where explicit results can be found. In particular, we propose three scheduling schemes to provide the rational rate constraints namely, weighted opportunistic (WO), time division opportunistic (TO), and superposition coding (SC). The WO scheduling is a generalization of the opportunistic scheduling in which we transmit to only the user that has the maximum weighted signal to noise ratio (SNR). In TO, each group has its own time slot in which the transmitter chooses the user with the best SNR from the corresponding group. Superposition coding is the one that achieves the capacity region. For each scheduling we give explicit scheme to guarantee the rational rate constraints. We also analyze the throughput loss due to rate constraints for different schemes. In particular, we show that the throughput loss compared to the maximum throughput (i.e., the sum rate capacity without any rate constraints) tends to zero for large n , and finally, we analyze the convergence rate of all the schemes.

I. INTRODUCTION

The down-link scheduling in cellular systems is known to be one major bottleneck for future broadband wireless communication. Information theoretic results on broadcast channels provide the limits for the maximum achievable rates¹ for each receiver [6], [2], [3]. For example in a homogeneous network, if the transmitter wants to maximize the throughput (or the sum of the rates to all the receivers²), the best strategy is to transmit to the user with the best channel condition at each channel use. This is the so called “opportunistic” transmission strategy [5].

For a homogenous network, opportunistic scheduling, however, treats all the users equally. In systems which are provisioned to provide differentiated services to different users, the transmitter has to give different services (or rates) to different subsets of receivers, and yet at the same time, maximize the throughput [1].

In this paper, we are interested in analyzing differentiated rate scheduling schemes for broadcast channels. In particular

¹Here we assume the channel is ergodic and rate refers to the average rate over all channel realizations.

²We use users and receivers interchangeably.

we assume receivers are divided into M groups where the ratios of the rates of users in different groups are given. Given these ratios, the overriding question in this paper is to devise schemes that provide the rational rate constraints and yet at the same time maximize the throughput of the system. We are also interested to see how much throughput loss the transmitter would incur by imposing such constraints.

II. PROBLEM FORMULATION

We consider a scalar Gaussian broadcast system with n receivers. We also assume the channels between the transmitter and the receivers are block fading, i.e., the channels remain fixed over the coherence interval of T . Therefore, we can write the relationship between X , the transmitted signal, and Y_i , the received signal at receiver i , as

$$Y_i = h_i X_i + W_i \quad (1)$$

where h_i is the channel coefficient between the transmitter and the i -th receiver and W_i is the additive noise. h_i and $W_i(t)$ are i.i.d. complex circularly symmetric Gaussian random variable with zero mean and unit variance, $CN(0, 1)$. In terms of channel knowledge, we assume that h_i is known perfectly at the transmitter and the receiver. We further assume that the transmitter is subject to a short-term power constraint power, so that the transmitted signal in every state must satisfy power constraint P . We denote the (average) rate of the i 'th user, $i = 1, \dots, n$, over channel realizations by R_i .

In this paper, we are interested in analyzing differentiated rate scheduling schemes for broadcast systems with large number of users, i.e., $n \gg 1$. We consider a partition of the receivers in M groups $\mathcal{G}_1, \dots, \mathcal{G}_M$, where different groups require different rates from the transmitter. We also assume that the size of the groups are all of the same order and hence, the cardinality of \mathcal{G}_k is $\alpha_k n$ where M and α_k 's are fixed numbers such that $\sum_{i=1}^M \alpha_i = 1$.

Assuming that the average rate of a user in the k 'th group denoted by R^k ,³ without loss of generality, we may assume $R^1 < \dots < R^M$. We further impose the constraint that the average rate of a user in the i 'th group is β_i times the average rate of a user in the M 'th group. In other words, we are interested

³Throughout the paper, we use superscript k to refer to any user in \mathcal{G}_k .

in the optimization problem

$$\max \sum_{i=1}^n R_i \quad (2)$$

$$\text{subject to } \frac{R^k}{R^M} = \beta_k, \quad k = 1, \dots, M, \quad (3)$$

where $\beta_1 < \dots < \beta_{M-1} < \beta_M = 1$ and are fixed numbers independent of n . It is clear that the best operating point in the ergodic capacity region is the intersection of the boundary of the capacity region $\mathcal{C}_{BC}(R_1, \dots, R_n)$ with the line defined in (3). While this is easy enough to state, it is not so easy to do since \mathcal{C}_{BC} is implicitly defined as the convex hull of an infinite set of rates. Moreover, it is not so easy to see how any point on the boundary of the capacity region can be mapped back to a schedule.

Therefore, in this paper, we will look at the performance of three scheduling schemes to provide the rational rate constraints in (3), namely, weighted opportunistic, time-division opportunistic, and superposition coding. In order to obtain explicit solutions, we will assume that the number of users n is large. Weighted opportunistic scheduling is the one that transmits to only the user that has the best weighted channel conditions. Of course, the weights should be determined in such a way that (3) is satisfied. In the time-division scheme we allow the transmitter to divide each channel use to non-equal time slots. At the k 'th slot, the transmitter sends to the receiver with the best channel condition from group \mathcal{G}_k . Finally, superposition coding is the one that achieves the ergodic capacity region [6], [2].

III. WEIGHTED OPPORTUNISTIC SCHEDULING

In this section, we consider a scheduling in which we are sending to only one user at each channel use that has the maximum weighted signal to noise ratio, i.e.,

$$\max_{1 \leq i \leq n} \mu_i |h_i|^2. \quad (4)$$

Since the rates of the users within the same group are the same, it is clear that we only need to set M different μ_i 's corresponding to M groups. The first question would be to figure out the mapping between the β_k 's in (3) and μ_k 's. In other words, what values of μ_i 's give the desired ratios β_i 's? Secondly, we are also interested to obtain the loss we incur on the throughput (i.e., sum of the rates) of the system by imposing (3) and using this scheduling.

In order to find the rate of a user in \mathcal{G}_1 , i.e. $R^{1,w}$, using weighted opportunistic scheduling, we may use (4) to write

$$R^{1,w} = \int_0^\infty \log(1 + Px) e^{-x} (1 - e^{-x})^{\alpha_1 n - 1} \times \prod_{j=2}^M (1 - e^{-\frac{\mu_j}{\mu_1} x})^{\alpha_j n} dx. \quad (5)$$

Equation (5) follows by noting that a user in \mathcal{G}_1 is chosen if its own channel is better than $\alpha_1 n - 1$ users in \mathcal{G}_1 and its weighted channel is better than $\alpha_i n$ users in \mathcal{G}_i for $i = 2, \dots, M$.

Analyzing $R^{1,w}$ for any n involves the exponential integral function and therefore explicitly finding the solution for μ_i 's

to guarantee β_i 's becomes cumbersome and numerically intractable. Therefore, we consider the regime of large number of users and find μ_i 's that satisfy (3) in this regime.

Lemma 1. *Suppose M and α_i 's are fixed and n grows. Then if μ_i 's are chosen such that*

$$\mu_i = 1 + \frac{\log \beta_i}{\log n}, \quad i = 1, \dots, M, \quad (6)$$

then $\lim_{n \rightarrow \infty} \frac{R_i^{1,w}}{R^{1,w}} = \beta_i$ for $i = 1, \dots, M - 1$. Furthermore

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n R_i^w}{\log \log n} = 1. \quad (7)$$

Proof: Assuming (6) holds, we first prove that $R^{1,w} = \frac{\alpha_1 \beta_1 \log \log n}{\sum_{k=1}^M \alpha_k \beta_k}$. We can write (5) as,

$$\begin{aligned} R^{1,w} &= \int_0^\infty \log(1 + Px) e^{-x} e^{-\sum_{k=1}^M \alpha_k n \log(1 - e^{-\frac{\mu_k}{\mu_1} x})} dx \\ &= \int_0^\infty \log(1 + Px) e^{-x - \sum_{k=1}^M \alpha_k n e^{-\frac{\mu_k}{\mu_1} x} + O(n e^{-2 \frac{\mu_1}{\mu_M} x})} dx \\ &= \int_0^\infty \log(1 + Px) e^{-x - e^{-x} \sum_{k=1}^M \alpha_k e^{-\frac{x \log(\beta_1/\beta_k)}{\log n} + O(n^{-\gamma})}} dx \end{aligned}$$

where γ is some constant bigger than 1. We now consider three regions for the integral above, namely, between $\log an \pm 4 \log \log n$, greater than $\log an + 4 \log \log n$, and less than $\log an - 4 \log \log n$ where $a = \frac{\sum_{k=1}^M \alpha_k \beta_k}{\beta_1}$. We can bound the second integral by noting that the average rate to the first group is at most of the order $\log \log n$ and therefore an upper would be $\log \log n$ multiplied by the probability that $|h_1|^2$ is greater than $\log an + 4 \log \log n$ condition on the fact that $|h_1|^2 \geq \frac{\mu_i}{\mu_1} |h_i|^2$ for $i = 2, \dots, n$. Hence the second integral can be written as

$$\begin{aligned} &\int_{\log an + 4 \log \log n}^\infty \log(1 + Px) e^{-x - e^{-x} \sum_{k=1}^M \alpha_k e^{-\frac{x \log(\beta_1/\beta_k)}{\log n} + O(n e^{-2 \frac{\mu_1}{\mu_M} x})}} dx \\ &= O\left(\log \log n \int_{\log an + 4 \log \log n}^\infty e^{-x - e^{-x} \sum_{k=1}^M \alpha_k e^{-\log(\beta_1/\beta_k) \frac{x}{\log n}} dx}\right) \\ &= O\left(\frac{\log \log n}{n(\log n)^4}\right) \end{aligned}$$

Similarly the integral for the third region becomes $O\left(\frac{\log \log n}{n} e^{-(\log n)^4}\right)$.

We now focus on evaluating the first integral. In fact we may change the variables to $y = x - \log an$ to get,

$$\begin{aligned} &\int_{\log an - 4 \log \log n}^{\log an + 4 \log \log n} \log(1 + Px) e^{-x - e^{-x} \sum_{k=1}^M \alpha_k e^{-\log(\beta_1/\beta_k) \frac{x}{\log n}}} dx \\ &= \log \log n \int_{\log an - 4 \log \log n}^{\log an + 4 \log \log n} e^{-x - e^{-x} \sum_{k=1}^M \alpha_k e^{-\log(\beta_1/\beta_k) \frac{x}{\log n}}} dx \\ &= \frac{\log \log n}{an} \int_{-4 \log \log n}^{4 \log \log n} e^{-y} e^{-e^{-y}} dy \\ &= \frac{\beta_1 \log \log n}{n \sum_{k=1}^M \alpha_k \beta_k} \left(1 - O\left(\frac{1}{(\log n)^4}\right)\right). \quad (8) \end{aligned}$$

It is quite straightforward to write the rate for users in the other groups in a similar way. Therefore for any group, we obtain

$$R^{k,w} = \frac{\beta_k}{\sum_{j=1}^M \alpha_j \beta_j} \frac{\log \log n}{n} + O\left(\frac{\log \log n}{n \log^4 n}\right). \quad (9)$$

Clearly the ratios of the rates satisfy (3) in the limit of large n and also the first order term in the throughput of this scheme is $\log \log n$ that leads to the second part of the Lemma. \square

Lemma 1 asserts that the average rates of users are quite sensitive to the change of μ_i 's. In order to further understand the impact of a change in μ_i 's on the rates, we consider a two group system. Following the methodology in the proof of Lemma 1, we can prove the following results. If

$$\frac{\mu_1}{\mu_2} = 1 - o\left(\frac{1}{\log n}\right) \implies \lim_{n \rightarrow \infty} \frac{R^{1,w}}{R^{2,w}} = 1. \quad (10)$$

where $R^{i,w}$ is as defined in Lemma 1. Moreover, if

$$\frac{\mu_1}{\mu_2} = c < 1 \implies \lim_{n \rightarrow \infty} \frac{R^{1,w}}{R^{2,w}} = 0 \quad (11)$$

where c is a constant independent of n .

We further look into the throughput loss that we would incur by differentiated rate scheduling. From Lemma 1, it is clear that the first order of the sum rate (i.e., $\log \log n$) remains unchanged. In the next Lemma we show that the difference of the sum rate capacity (maximum throughput) and the throughput of this schemes converges to zero. We also obtain the convergence rate.

Lemma 2. *Suppose M and α_i 's are fixed and n grows. Then,*

$$\int_0^\infty n \log(1+Px) e^{-x} (1-e^{-x})^{n-1} - \sum_{i=1}^n R_i^w = \Theta\left(\frac{(\log \log n)^3}{(\log n)^2}\right) \quad (12)$$

where the first term denotes the sum rate capacity achieved by sending to the user with the best channel condition at each channel use.

Proof: We prove this lemma for the case of $M = 2$ for the sake of brevity, however it is quite straightforward to generalize to $M > 2$. We first divide the integral that represents the difference of the throughputs into three regions as we did in the proof of Lemma 1. The first term would be the integral when $\log(n/2) - 4 \log \log n \leq x \leq \log(n/2) + 4 \log \log n$ which is denoted by region \mathcal{A} . We can then write the first term as,

$$\begin{aligned} & \int_{\mathcal{A}} \frac{n}{2} \log(1+Px) e^{-x} (1-e^{-x})^{n/2-1} \left\{ (1-e^{-x})^{n/2-1} \right. \\ & \left. - (1-e^{-\frac{\mu_1}{\mu_2}x})^{n/2} - (1-e^{-\frac{\mu_1}{\mu_2}x})^{n/2} \right\} dx \\ & = \Theta\left(\log \log n \int_{-4 \log \log n}^{4 \log \log n} e^{-y} \left(1 - \frac{e^{-y}}{n/2}\right)^{n/2-1} \left\{ \left(1 - \frac{e^{-y}}{n/2}\right)^{n/2} \right. \right. \\ & \left. \left. - (1-e^{-\frac{\mu_1}{\mu_2}(y+\log n/2)})^{n/2} - (1-e^{-\frac{\mu_1}{\mu_2}(y+\log n/2)})^{n/2} \right\} dy\right) \end{aligned}$$

Assuming that $\epsilon = \frac{\log \log n}{\log n}$, we can write the above equation as,

$$\begin{aligned} & \Theta\left(\log \log n \left(2e^{-\frac{1}{\log^{\gamma} n}} - \frac{e^{-\frac{1}{\log^{\gamma} n}(1+\beta_1(1+\epsilon))}}{1+\beta_1(1+\epsilon)} \right. \right. \\ & \left. \left. - \frac{e^{-\frac{1}{\log^{\gamma} n}(1+\frac{1}{\beta_1}(1-\epsilon))}}{1+\frac{1}{\beta_1}(1-\epsilon)}\right)\right) \\ & \Theta(\epsilon^2 \log \log n) = \Theta\left(\frac{(\log \log n)^3}{(\log n)^2}\right) \quad (13) \end{aligned}$$

It is quite straightforward to show that the other terms would contribute $O\left(\frac{\log \log n}{(\log n)^{2\gamma}}\right)$. This completes the proof. \square

IV. TIME-DIVISION OPPORTUNISTIC SCHEDULING

Another, in fact simpler, approach to guarantee the rate constraints is to do time-sharing between different groups by dividing each channel use of duration T to M slots of different lengths⁴. The i 'th slot is dedicated to the i 'th group and the transmitter chooses the receiver with the best channel conditions for transmission from \mathcal{G}_i . Intuitively, we should be able to achieve the rational rate constraints if we divide the slots into the same ratios. (Lemma 3 shows this is the case.)

Here we denote the rate of a user in \mathcal{G}_k using the this scheme by $R^{k,t}$, similarly the rate of the i 'th user will be denoted by R_i^t . In the next lemma, we show that if the cardinality of all groups is of the order of n , we can construct the length of slots such that (3) is satisfied and the throughput of the scheme scales like $\log \log n$.

Lemma 3. *Suppose M and α_i 's are fixed. Also, let l_i be the length of the i 'th slot and is equal to*

$$\frac{l_i}{T} = \frac{\alpha_i \beta_i}{\sum_{i=1}^M \beta_i \alpha_i} \quad i = 1, \dots, M. \quad (14)$$

Then $\lim_{n \rightarrow \infty} \frac{R^{i,t}}{R^{M,t}} = \beta_i$ for $i = 1, \dots, M-1$. Moreover,

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n R_i^t}{\log \log n} = 1. \quad (15)$$

Proof: It is quite easy to show that at the i 'th slot in which the transmitter sends information to the user with the best channel condition among users in \mathcal{G}_i , the total rate of information sent to \mathcal{G}_i is equal to

$$\frac{\alpha_i \beta_i}{\sum_{i=1}^M \beta_i \alpha_i} \log \log \alpha_i n \quad (16)$$

Therefore since users in each group are equally likely to be chosen, the rate to a user in the i 'th group is equal to $\frac{1}{\alpha_i n}$ times the rate to \mathcal{G}_i as in (16). This proves the first part of the Lemma. The second part of the lemma follows by noting that $\sum_{k=1}^M = 1$ and α_i 's are fixed and are not vanishing to zero. \square

It is worth mentioning that in the time-division scheme at the i 'th slot, the transmitter will certainly not transmit to any

⁴Instead of one channel use, one might divide every K channel uses to slots of different lengths

user outside group i even if the user has much better channel condition. Therefore although simpler, one may guess that the time-division scheme has a lower throughput than the weighted opportunistic one. The next Lemma proves again the throughput of this scheme converges to the sum rate capacity, however, the convergence rate for the time-division scheme is a polynomially larger than that of the weighted opportunistic scheduling (i.e., $\Theta\left(\frac{\log \log n}{\log n}\right)$ versus $\Theta\left(\frac{(\log \log n)^3}{(\log n)^2}\right)$).

Lemma 4. *Suppose M and α_i 's are fixed. Also, let l_i chosen as in (14). Then*

$$\int_0^\infty n \log(1+Px) e^{-x} (1-e^{-x})^{n-1} - \sum_{i=1}^n R_i^t = \Theta\left(\frac{\log \log n}{\log n}\right) \quad (17)$$

Proof: Here we present the proof for the special case of having two groups of the same size. The proof however can be generalized easily and we omit it for the sake of brevity.

The difference of the throughputs can be written as,

$$\int_0^\infty n \log(1+Px) e^{-x} \left\{ (1-e^{-x})^{n-1} - \frac{1}{2}(1-e^{-x})^{n/2-1} \right\} dx$$

We can now follow the same approach as in the proof of Lemma 1 and expand the integral to three parts, namely, $\log n/2 \pm 4 \log \log n$, larger than $\log n/2 + 4 \log \log n$, and smaller than $\log n/2 - 4 \log \log n$. The last two integrals lead to $O\left(\frac{\log \log n}{\log^4 n}\right)$.

The first region can be then evaluated as,

$$\begin{aligned} & \int_{\log n/2 - 4 \log \log n}^{\log n/2 + 4 \log \log n} n \log(1+Px) e^{-x} \\ & \quad \times \left\{ (1-e^{-x})^{n-1} - \frac{1}{2}(1-e^{-x})^{n/2-1} \right\} dx \\ = & \Theta\left(\log \log n \int_{-4 \log \log n}^{4 \log \log n} e^{-y} \left(1 - \frac{e^{-y}}{n/2}\right)^{n/2-1} \right. \\ & \quad \times \left. \left\{ \left(1 - \frac{e^{-y}}{n/2}\right)^{n/2} - 1 \right\} dy\right) \\ = & \Theta\left(\log \log n \int_{-4 \log \log n}^{4 \log \log n} e^{-y} e^{-e^{-y}} \left(e^{-e^{-y}} - \frac{1}{2}\right) dy\right) \\ = & \Theta\left(\frac{\log \log n}{\log n}\right) \end{aligned} \quad (18)$$

This completes the proof for the two group case. The generalization to the M group case follows by using the same technique. \square

In the next section, we look into a scheme that employs superposition coding and clearly leads to the best throughput as we actually work on the boundary of the capacity region. As the analysis becomes complicated, we just consider two groups and obtain a scheduling that maximizes the throughput while the rational rate constraints of (3) are satisfied. It is worth mentioning that the ergodic capacity region of a broadcast channel with two users has been studied in [2], here we look at a generalization of the result of [2] in which we have n users divided into two groups.

V. SUPERPOSITION CODING

In this section, we analyze the performance of superposition coding for the case when there are only two groups of users $\mathcal{G}_1, \mathcal{G}_2$ with equal sizes that have different rate demands. We assume that the average rate provided to a user in the first group is required to be $\beta > 1$ times the rate provided to a user in the second group.

In order to maximize the rate (sum-rate) while keeping the ratio of different group rates fixed and equal to β , we need to find the point on the boundary of the capacity region of the Gaussian broadcast channel with short-term power constraint P that satisfies the differentiated rate constraint. We know that every boundary point is the solution to the maximization problem

$$\max_{(R_1, \dots, R_n) \in \mathcal{C}_{BC}} \sum_{i=1}^n \mu_i R_i$$

for some positive values μ_1, \dots, μ_n . In our case because of the symmetry among the users in each group, the values of μ_i s will be the same for the users in the same group. Therefore, we only need to characterize the boundary points that are the maximizing solution to the problem

$$\max_{(R_1, \dots, R_n) \in \mathcal{C}_{BC}} \mu_1 \left(\sum_{i \in \mathcal{G}_1} R_i \right) + \mu_2 \left(\sum_{i \in \mathcal{G}_2} R_i \right)$$

for $\mu_1, \mu_2 > 0$. The following lemma characterizes such boundary points. The proof of this lemma uses the duality of the broadcast channel and the multi-access channel for scalar channels [7]. We do not include the proof for the sake of brevity.

Lemma 5. *Consider a scalar Gaussian broadcast system with the model described in Section II. Consider the following optimization problem*

$$\max_{(R_1, \dots, R_n) \in \mathcal{C}_{BC}} \mu_1 \left(\sum_{i \in \mathcal{G}_1} R_i \right) + \mu_2 \left(\sum_{i \in \mathcal{G}_2} R_i \right) \quad (19)$$

where \mathcal{C}_{BC} is the ergodic capacity region of broadcast channel with short-term power constraint P and $\mu_1 \geq \mu_2$ are two positive numbers. Then the solution of the above optimization problem is

$$\begin{aligned} \frac{n}{2} R_i &= \mathbb{E}(\log(1+Px) | \mu_1 x \geq \mu_2 y) + \\ & \mathbb{E}\left(\log\left(\frac{(\mu_1 - \mu_2)y(1+Px)}{\mu_1(y-x)}\right) | (x, y) \in \mathcal{R}\right) \end{aligned}$$

for $i \in \mathcal{G}_1$. Similarly, for $i \in \mathcal{G}_2$, we have

$$\begin{aligned} \frac{n}{2} R_i &= \mathbb{E}(\log(1+Py) | \mu_1 x \leq \mu_2 y) - \\ & \mathbb{E}\left(\log\left(\frac{(\mu_1 - \mu_2)x(1+Py)}{\mu_2(y-x)}\right) | (x, y) \in \mathcal{R}\right) \end{aligned} \quad (20)$$

where $x = \max_{i \in \mathcal{G}_1} |h_i|^2$, $y = \max_{i \in \mathcal{G}_2} |h_i|^2$ and region \mathcal{R} is defined as

$$\mathcal{R} = \{(x, y) \in R^+ \times R^+ | 0 \leq \frac{\mu_2}{(\mu_1 - \mu_2)x} - \frac{\mu_1}{(\mu_1 - \mu_2)y} \leq P\}.$$

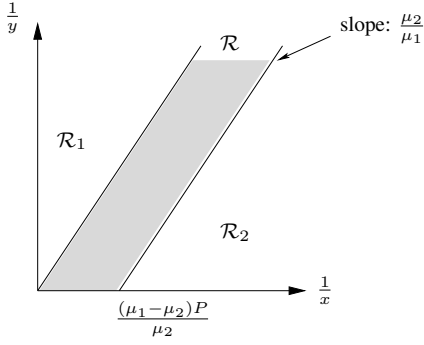


Fig. 1. The decision region for power allocation in the superposition coding in two group case: If $(x, y) \in \mathcal{R}_1$, all the power is allocated to best user of group one. If $(x, y) \in \mathcal{R}_2$, all the power is allocated to best user of group two. If $(x, y) \in \mathcal{R}$ then power is split between the best users of the two groups as in (21).

The power allocation policy corresponding to the superposition coding that maximizes (19) can be interpreted as follows: for each realization of channel coefficients, we only send to the user with the best channel in each group. Letting $x = \max_{i \in \mathcal{G}_1} |h_i|^2$, $y = \max_{i \in \mathcal{G}_2} |h_i|^2$ we have the following possibilities:

- 1) If $\mu_1 x \geq \mu_2 y$, we assign all the power to the best user of the first group.
- 2) If $0 \leq \frac{\mu_2}{(\mu_1 - \mu_2)x} - \frac{\mu_1}{(\mu_1 - \mu_2)y} \leq P$ then we split the power between the two best users in the two groups as

$$P_x = \frac{(\mu_1 - \mu_2)Pxy + \mu_1 x - \mu_2 y}{\mu_1(y - x)}, \quad P_y = P - P_x. \quad (21)$$

- 3) If $\frac{\mu_2}{(\mu_1 - \mu_2)x} - \frac{\mu_1}{(\mu_1 - \mu_2)y} > P$, all the power is assigned to the best user in the second group.

We have plotted the decision region for power allocation in the $(\frac{1}{x}, \frac{1}{y})$ region in Fig. V. In the weighted opportunistic scheduling the power allocation policy would be to send to the best user in the first group if (x, y) is in \mathcal{R}_1 and to send to the best user in the second group if (x, y) is in $\mathcal{R} \cup \mathcal{R}_2$.

The question that is remained to be answered is to figure out how to choose μ_1 and μ_2 such that the rate constraint in (3) satisfied. This is answered in the next lemma whose proof we omit for lack of space.

Lemma 6. Suppose $\beta < 1$ is fixed, $\mu_1 = 1$, and $\mu_2 = 1 - \frac{1}{(\log n)^\alpha}$ where $\alpha = \frac{\beta+3}{\beta+1}$ (i.e., $1 \leq \alpha \leq 2$), then

$$\lim_{n \rightarrow \infty} \frac{R^1}{R^2} = \beta. \quad (22)$$

Finally we look into the throughput loss due to the constraint of (3) using superposition coding. Using Lemma 2, it is clear that the loss should tend to zero for large n and also the convergence rate should be faster than $\frac{(\log \log n)^3}{(\log n)^2}$. In the next lemma we prove that the convergence rate can not be faster than $\frac{(\log \log n)^{1+2\alpha}}{(\log n)^{2\alpha}}$ where $1 \leq \alpha \leq 2$ is a fixed number.

Lemma 7. Suppose $\beta > 1$ is fixed and μ_1, μ_2 are chosen as in

Lemma 6. Then

$$\int_0^\infty n \log(1 + Px) e^{-x} (1 - e^{-x})^{n-1} - \sum_{i=1}^n R_i = \Omega\left(\frac{(\log \log n)^{1+2\alpha}}{(\log n)^{2\alpha}}\right) \quad (23)$$

Proof: Here is the outline of the proof. We can write the throughput under constraints of (3) using (20) as,

$$\begin{aligned} & \alpha_1 n \int_0^\infty \log(1 + Px) e^{-x} (1 - e^{-x})^{\alpha_1 n - 1} (1 - e^{-\frac{\mu_1}{\mu_2} x})^{\alpha_2 n} dx \\ & + \alpha_2 n \int_0^\infty \log(1 + Px) e^{-x} (1 - e^{-x})^{\alpha_2 n - 1} (1 - e^{-\frac{\mu_2}{\mu_1} x})^{\alpha_1 n} dx \\ & + E \log\left(\frac{(1 + Px)y}{(1 + Py)x} \mid (x, y) \in \mathcal{R}\right) \end{aligned}$$

In fact the first two terms are the same as the throughput of the weighted opportunistic scheduling with $\mu_1 = 1$ and μ_2 are chosen as in Lemma 6. Therefore similar to the proof of Lemma 2 we can show that the difference of the sum rate capacity and the first two terms tends to zero like $\frac{(\log \log n)^{1+2\alpha}}{(\log n)^{2\alpha}}$. The third term however can be easily shown to be a positive when $(x, y) \in \mathcal{R}$. Therefore, the difference of the sum rate capacity and the throughput of this scheme can not tend to zero faster than $\frac{(\log \log n)^{1+2\alpha}}{(\log n)^{2\alpha}}$. This complete the proof of the lemma. \square

VI. CONCLUSION

We considered a fading broadcast channel with n users demanding for different rates. In our model we assumed users are divided into M groups each one is demanding different rates and where the ratios of the rates of the groups are given. Users in each groups have the same rate requirement. We considered the problem of scheduling to users to maximize the throughput of the system while maintaining the rational rate constraints. As the problem in general seems to be computationally intractable, we focused on the regime of large number of users. Three different scheduling are proposed, namely, weighted opportunistic, time-division opportunistic and super position coding. We gave explicit scheduling to guarantee the rate constraints. We further showed that the throughput loss due to these constraints tends to zero for all three schemes as $n \rightarrow \infty$ and we obtained their convergence rates.

REFERENCES

- [1] N. Jindal and A. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," *IEEE Trans. Info. Theory*, vol. 49, no. 11, pp. 2895–2909, 2003.
- [2] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. I. ergodic capacity," *IEEE Trans. Info. Theory*, vol. 47, no. 3, pp. 1083–1102, 2001.
- [3] H. Viswanathan, S. Venkatesa, and H. Huang, "Downlink capacity evaluation of cellular networks with known interference cancellation," *IEEE Jour. Selec. Areas in Comm.*, vol. 21, no. 5, pp. 802–811, 2003.
- [4] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Info. Theory*, vol. 51, no. 2, pp. 506–522, 2005.
- [5] R. Knopp and P. Humblet, "Information capacity and power control in single cell multiuser communications," in *Proc. IEEE Inter. Conf. Comm.*, vol. 1, pp. 331–335, 1995.
- [6] T. Cover, "Broadcast Channels," *IEEE Trans. Info. Theory*, vol. 18, no. 1, pp. 2–14, 1972.
- [7] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels" *IEEE Trans. Info. Theory*, vol. 50, no. 5, pp. 768–783, 2004.