ISIT'2013 1569719391

On Fading MAC Channels With Asymmetric CSI

Kamal Singh, Sibi Raj B Pillai, Bikash K Dey
Department of Electrical Engineering
Indian Institute of Technology Bombay
{kamalsingh,bsraj,bikash}@ee.iitb.ac.in

***Abstract*—We consider a distributed MAC setting with blockwise flat fading links and full receiver CSI (channel state information). Of the L transmitters, a subset is assumed to have knowledge of the global CSI vector in each block, whereas the remaining users have access only to their respective link qualities, i.e. each one in the latter set is unaware of the quality of other links. Outage is not allowed in any communication block. We propose efficient power-allocation and rate-adaptation strategies which are sum-rate optimal when users in each subset observe identical fading distributions chosen from a class, which includes the popular Rayleigh, Ricean etc.**

I. INTRODUCTION

The capacity region of fading multiple access channels depends on the availability of channel state information (CSI). In order to characterize the impact of varying levels of transmitter CSI in a slow fading MAC channel, this paper considers a particular model of transmitter CSI availability, which we term as the *asymmetric CSI model*. More specifically, we consider blockwise flat fading links and partition the transmitters into two groups. A set S_f of users have access to the global CSI vector in every block[1]. The remaining users know just their respective fading values in each block, we denote this set by S_d. The latter CSIT model is also known as MAC with distributed CSI in literature [1]. The distributed setup was motivated by its practical relevance, and is analyzed under the title 'channel gain available at respective encoders and the decoder' in [1] (page 590), see also [2].

The transmitters can effectively utilize the CSI by adapting their transmit power and transmission-rates based on the available information. In a slow fading MAC, a reasonable objective is to design a distributed access scheme which will guarantee outage-free operation in each block, thus coding and decoding have to take place within the coherence period. Furthermore, the dynamic power-allocation has to respect an allowed long-term average expenditure. The feasible average (over blocks) rate-pairs in this setting is called the *power-controlled adaptive capacity region*, a term coined in [1]. When the channel gains are only available at the respective encoders and the decoder (distributed CSI), a single letter characterization of the power controlled capacity region was obtained in [2], which expresses the capacity region as an optimization problem in terms of power-allocation functions. Optimal sum-capacity achieving power allocations for the distributed CSI MAC was explicitly determined for a large class of symmetric fading models in [3]. The interesting case of non-identical channel statistics appear in [4]. While these solutions are useful, the current problem is inherently different due to the global CSI availability at some terminals, and needs significantly different approaches. In fact, the asymmetric CSI problem relates more to the degraded CSI MAC setup in [5]. When there is only one transmitter in each user-set, our model indeed becomes a two-user degraded CSI MAC.

In a two-user degraded CSI MAC, the CSI available to one transmitter is a physically degraded version of that available to the other. [5] obtained a single letter characterization of the capacity region for this model, which for the Gaussian case can again be cast as an optimization problem in terms of power-allocation functions. Our problem/solution differs from this model in three aspects:

- when there are two or more users in the set S_d, we no longer have the degraded CSI setup.
- our interest is the adaptive capacity, where the communication in each block has to be outage-free (as in [2]), whereas [5] analyzes the ergodic rates.
- even in the two user degraded CSI MAC model, we seek explicit solutions to the optimal power control.

Recall that the ergodic capacity is achieved by coding over a large number of blocks, which is simply not possible in the slow-fading, opportunistic settings that we consider. In spite of the above differences, the resulting adaptive sum-capacity for our asymmetric CSI model can be expressed in terms of the ergodic sum-capacity of a two-user degraded CSI MAC. Thus, to solve the problem at hand, we determine the optimal power allocation schemes for the degraded CSI MAC under several fading models (for example Rayleigh, Rice), a result of independent interest.

The power-allocation and rate-adaptation schemes that are proposed in this paper are sum-rate optimal (achieves *power-controlled adaptive sum-capacity*) when the links in each user-set have identical statistics, chosen from a class which includes several popular fading models. While this assumption can be relaxed in several directions, we limit the discussion to this special case, which is not only of practical relevance, but also gives intuitive results. Thus, the users in the set S_d have identical channel statistics with cdf $\Psi_d(h)$, whereas each user in S_f encounters fading with cdf $\Psi_f(h)$. The fading gains of different users are assumed to be independent.

The asymmetric CSI model is a forward step in comprehensively understanding the more general CSI settings in slow fading models. It also has some practical relevance. The model can effectively capture cognitive access in networks, where the primary user is unaware of the channel coefficients of the secondary. On the other hand, the secondary user is not

[1] index f stands for full-CSI, and d for *distributed*.

allowed to disrupt the primary, and may have to glean the full CSI vector. A secondary-aware protocol, without explicit global CSI at the primary, can support both users at the same time, creating optimal opportunistic access and overall throughput. A similar situation occurs when a set of users does not care too much about expending resources to garner the CSI of other links, due to resource constraints, or viewing other links to be of secondary importance.

The organization of the paper is as follows. Section II will introduce the system model, along with some definitions and notations. In Section III, we will find the sum-capacity of a two user MAC with asymmetric CSI. In here one user has full CSI, while the other knows only its own CSI. We demonstrate the optimal water-filling power allocation and the corresponding rate allocations. This is then extended to multiple users, with L_f users in S_f and $L_d = L - L_f$ users in S_d, in Section IV. Conclusions and future work are presented in Section V.

II. System Model

Consider a $L-$user fading MAC given by,

$$Y = \sum_{i=1}^{L} H_i X_i + Z\,,$$

where users transmit real-valued signals X_i, $1 \leq i \leq L$, and encounter real-valued multiplicative fades H_i. The additive noise Z is an i.i.d normalized Gaussian process independent of X and H. The fading space $\mathcal{H}_i$ of the i-th user is the set of values taken by H_i, and the joint fading space $\bar{\mathcal{H}}$ is the set of values taken by the joint fading state $\bar{H} = (H_1, H_2, \cdots, H_L)$. We assume that the fading processes H_i are independent, and their distributions are known to all the transmitters and the receiver. The slow-fading nature of the medium is captured by a block-fading model, in which the fading vector remains constant within a block and varies independently across blocks. The full CSI is available at the receiver.

As mentioned earlier, we partition the L users into two sets. Each user in the set S_f has access to the full global CSI. The remaining users, denoted as S_d has access only to distributed CSI, i.e. they know their respective fading gains and have no more idea than the statistics of the other links. The transmitters have the freedom to adapt their rate (and power) according to the available knowledge of the channel conditions. However, we demand that the rate choices should ensure no outage in every transmission block for all users. Denoting $\bar{h} = (h_1, \cdots, h_L)$, the power allocation function at user i can be written as,

$$P_i(\bar{h}) = \begin{cases} P_i(h_1, h_2, \cdots, h_L),\ i \in S_f \\ P_i(h_i),\ i \in S_d. \end{cases} \tag{1}$$

The rate choices in these cases are denoted as

$$R_i(\bar{h}) = \begin{cases} R_i(h_1, h_2, \cdots, h_L),\ i \in S_f \\ R_i(h_i),\ i \in S_d. \end{cases} \tag{2}$$

The next few definitions closely follow the ones in [6].

Definition 1. *A power-rate strategy is a collection of mappings $(P_i, R_i) : \bar{\mathcal{H}} \longmapsto \mathbb{R}^+ \times \mathbb{R}^+$; $i = 1, 2, \cdots, L$. Thus, in the global fading-state $\bar{H}$, the i^{th} user expends power $P_i(\bar{H})$ and employs a codebook of rate $R_i(\bar{H})$.*

Let $C_{MAC}(\bar{h}, \bar{P}(\bar{h}))$ denote the capacity region of a Gaussian multiple-access channel with fixed channel gains of $\bar{h}$ and respective power allocations $\bar{P}(\bar{h}) = (P_1(\bar{h}), \cdots, P_L(\bar{h}))$. We know that,

$$C_{MAC}(\bar{h}, \bar{P}(\bar{h})) = \Bigg\{ \bar{R} \in \{\mathbb{R}^+\}^L : \forall S \subseteq \{1, 2, \cdots, L\} \\ \sum_{i \in S} R_i \leq \frac{1}{2} \log\left(1 + \sum_{i \in S} |h_i|^2 P_i(\bar{h})\right) \Bigg\}. \tag{3}$$

Definition 2. *A power-rate strategy is* feasible *if it satisfies the average power constraints for each user i.e.*
$\forall i \in \{1, 2, \cdots, L\}, \quad \mathbb{E}_H P_i(\bar{H}) \leq P_i^{avg}$.

Definition 3. *A power-rate strategy is termed as* outage-free *if it never results in outage i.e.*

$$\forall \bar{h} \in \bar{\mathcal{H}}, (R_1(\bar{h}), \cdots, R_L(\bar{h})) \in C_{MAC}(\bar{h}, \bar{P}(\bar{h})).$$

Let Θ_{MAC} be the collection of all feasible outage-free power-rate strategies. The relevant averages are evaluated over the joint fading pdf $\Psi(\bar{h}) = \prod_{i \in S_f} \Psi_f(h_i) \prod_{j \in S_d} \Psi_d(h_j)$. The results in this paper are illustrated for the special case where the underlying fading distributions are Rayleigh. This not only has wide applications, but also allows a clear exposition of the nature of the results. Extensions to many other continuous valued distributions are immediate, with the structural results remaining intact. Discrete fading states will need slightly different treatment, particularly in the high SNR regime, we do not consider it here.

Remark 4. *For the ease of exposition, we also assume that all users in each of the sets to have the same average transmit power constraint. This restriction can be relaxed in a straightforward manner.*

Definition 5. *The **adaptive sum-capacity** $C_{\bar{\Psi}}$ is the maximum average sum-rate achievable, i.e.*

$$C_{\bar{\Psi}} = \max_{\theta \in \Theta_{MAC}} E_{\bar{H}}(\sum_{i=1}^{L} R_i(\bar{H})).$$

Remark 6. *It should be noted that the above definition of adaptive sum-capacity is not exactly a capacity in the Shannon sense[1], as we consider coding within a coherence period.*

Since the receiver has perfect CSI and transmitters know their respective CSI, the capacity expressions depend only on the magnitude of the fading coefficients. Thus, from now on wards, we will consider only the fading magnitudes. With slight abuse of notation, we will denote the magnitude of fading as $h_i, 1 \leq i \leq L$.

In order to find the adaptive sum-capacity $C_{\bar{\Psi}}$, we first consider a two-user asymmetric CSI model with one user in each subset. In this case, the system becomes a MAC with degraded CSI [5]. We will explicitly solve the power control problem formulation posed in [5].

III. Two-User MAC with Asymmetric CSI

With only two users in the asymmetric setup, though we are interested in evaluating the *adaptive sum-capacity*, this notion coincides with the ergodic sum-capacity in [5]. It also turns out that the more general solution can be stated in terms of the two-user result (see Theorem 9). W.l.o.g assume that user 1 has full CSI, while user 2 knows only its own CSI. Any given power-rate allocation strategy θ will achieve an average sum-rate of

$$T_\theta = \int d\Psi(h_1,h_2)\left(R_1(h_1,h_2)+R_2(h_2)\right).$$

Our objective is to find $\max_\theta T_\theta$ which we denote by $C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, P_d^{avg})$. This notation will come handy later in Theorem 9. As mentioned earlier, we demonstrate the results for the case where $\Psi_f(\cdot)$ and $\Psi_d(\cdot)$ are Rayleigh distributed, their parameters can vary.

The optimal strategy in the asymmetric CSI model is quite different from the conventional water-filling strategies in the presence of full CSI [7]. We will show that for every fading gain h_2 of the second user, there exists a threshold $\gamma(h_2)$ such that the first user allocates power only when $h_1 \geq \gamma(h_2)$.

Theorem 7. *There exists a positive constant λ_1 and a non-negative function $\gamma(\cdot)$ such that*

$$\max_\theta T_\theta = \frac{1}{2}\mathbb{E}\log\left(1+h_1^2P_1^*(h_1,h_2)+h_2^2P_2^*(h_2)\right) \tag{4}$$

where

$$P_2^*(h_2) = \left(\frac{\gamma^2(h_2)}{\lambda_1 h_2^2} - \frac{1}{h_2^2}\right)^+ \mathbb{1}_{\{h_2\neq 0\}}. \tag{5}$$

$$P_1^*(h_1,h_2) = \left(\frac{1}{\lambda_1} - \frac{\gamma^2(h_2)}{\lambda_1 h_1^2}\right)^+, \tag{6}$$

and

$$\int d\Psi(\bar{h})P_i^*(\bar{h}) = P_i^{avg},\ i=1,2. \tag{7}$$

Proof: We will first construct an upperbound to the adaptive sum-capacity, which is shown to be tight later. Notice that for every block, the MAC sum-rate bound in (3) implies

$$R_1(h_1,h_2)+R_2(h_2) \leq \frac{1}{2}\log(1+h_1^2P_1(h_1,h_2)+h_2^2P_2(h_2)). \tag{8}$$

The expected value of the RHS in (8) can be maximized under the average power constraints of P_f^{avg} and P_d^{avg} at users 1 and 2 respectively. We implicitly use the individual rate constraints of the participating users and maximize only the sum-rate bound. The maximum value indeed matches the optimal rate given in the theorem. The details are shown in Appendix A.

On the other hand, the sum-rate given in Theorem 7 can be achieved. To this end, let us always decode user 1 first and allow a clean channel to user 2. Thus, user 2 can choose the maximal allowed rate for a transmit power of $P_2^*(h_2)$. User 1, by knowing both channels, selects a power of $P_1^*(h_1,h_2)$ and operates at the corner point of the dominant face of the MAC capacity region such that its message can be decoded by treating transmissions from user 2 as noise. Clearly the chosen sum-rate matches that in (4), thus achieving the promised two-user sum-capacity $C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, P_d^{avg})$. ■

One drawback of the current result is that the power-allocation is parametrized in terms of λ_1 and the threshold functions $\gamma(h_2)$. We now present an iterative procedure to evaluate the adaptive sum-capacity in a small number of steps.

A. Computational Algorithm

The procedure relies on the derivation of Theorem 7 given in Appendix A. Let $P_1(h_1,h_2)$ and $P_2(h_2)$ denote the power allocations given in (5) – (6). The algorithm, parametrized in terms of a positive variable α, is as follows.

1) Choose an initial value of $\alpha=1$, a small step-size δ and approximation error tolerance ϵ.
2) For every h_2, find a threshold $\gamma(h_2)$ such that
$$\alpha = h_2^2\int \min\{\frac{1}{\gamma^2(h_2)}, \frac{1}{h^2}\}d\Psi_f(h).$$
3) Find λ_1 such that $\int P_2(h)d\Psi_d(h) = P_d^{avg}$.
4) Find $\hat{P}_1 = \int P_1(h_1,h_2)d\Psi_f(h_1)d\Psi_d(h_2)$.
5) If
$$\begin{cases}\left(\hat{P}_1 - P_1\right) > \epsilon \text{ then } \alpha=\alpha+\delta \text{ ; go back to 2}\\ \left(\hat{P}_1 - P_1\right) < -\epsilon \text{ then } \alpha=\alpha-\delta \text{ ; go back to 2}\end{cases}$$

It can be shown that for some $\epsilon > 0$ and a properly chosen step-size δ , the proposed algorithm converges to the correct solution, the proof is omitted due to space constraints, see [8]. Figure 1 illustrates the results obtained by running the algorithm for the case where Ψ_f and Ψ_d are normalized Rayleigh distributed. Clearly, the sum-capacity for the asymmetric CSI system is very close to the full CSIT model, showing its utility.

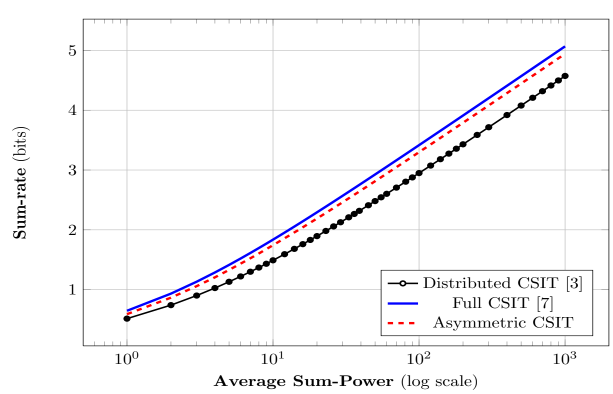

Fig. 1. Sum-capacity for pdfs $\Psi_f'(h) = \Psi_d'(h) = 2he^{-h^2}$, $P_f^{avg} = P_d^{avg}$

IV. Asymmetric MAC with many users

While extending the 2 user example to more users in the set S_d, guaranteeing outage-free operation is a little more involved. Nevertheless, we demonstrate the sum-capacity achieving schemes in this section.

Consider L_f users in the set S_f and the remaining users in S_d. Without loss of generality, index the users in S_f by $\{1,\cdots,L_f\}$. Recall also that we assumed identical distributions and average powers for all users in each partition. The adaptive sum-capacity $C_{\bar{\Psi}}$ is

$$\begin{aligned} C_{\bar{\Psi}} &= \max\left(\int d\Psi(\bar{h})\sum_{i=1}^{L_f} R_i(\bar{h}) + \sum_{j>L_f}\int d\Psi_d(h_j)R_j(h_j)\right) \\ &= \max\left(\int d\Psi(\bar{h})\sum_{i=1}^{L_f} R_i(\bar{h}) + \int d\Psi_d(h)\sum_{j>L_f} R_j(h)\right). \end{aligned} \tag{9}$$

We prove next that we can limit ourselves to those strategies where in every block, only one user from the set S_f transmits and the rest remain silent.

Lemma 8. *In order to achieve $C_{\bar{\Psi}}$, it is sufficient to consider time-sharing of users in S_f.*

Proof: For a given fading vector and respective transmit powers, the $L-$user MAC capacity region is a polymatroid. For any rate-tuple in the dominant face of this polymatroid, let the sum-rate be

$$R_{sum}^{dom} \triangleq \frac{1}{2}\log\left(1+\sum_{i\in S_f} h_i^2 P_i(\bar{h}) + \sum_{j\in S_d} h_j^2 P_j(\bar{h})\right).$$

Consider the total-rate for the users in S_d. If

$$\sum_{i\in S_d} R_i(h_i) < \frac{1}{2}\log\left(1+\frac{\sum_{i=L_f+1}^{L} h_i^2 P_i(h_i)}{1+\sum_{j=1}^{L_f} h_j^2 P_j(\bar{h})}\right), \tag{10}$$

then clearly $\sum_{i=1}^{L} R_i < R_{sum}^{dom}$. This has a useful implication on decoding, that the users in the set S_d can now be jointly decoded by treating all transmissions from S_f as noise, see the denominator of (10). On subtracting the signals from S_d, the users in S_f will encounter a clean channel devoid of any interference from S_d, and it is well known that the 'best-user-transmits' policy is optimal among S_f [7].

On the other hand, suppose that (10) is not true, and the chosen global rate-vector is on the dominant face, then reallocating all the available power of S_f to the best user in this set may only improve the received power and hence the sum-rate. This is indeed a form of TDM, among S_f, and the lemma is proved. ■

Notice that the argument can be modified to accommodate different channel statistics and power-constraints among S_f. The assumptions of identical distribution and average powers for S_f considerably simplify the exposition, that we can replace all the users in S_f by a single one having a channel state of $h = \max\{h_1,\cdots,h_{L_f}\}$, equipped with the total average sum-power of the set S_f. Thus, we will now consider only a single transmitter in S_f with cdf Ψ_f, which is the equivalent cdf of the highest channel coefficient. We now state the adaptive sum-capacity for this model. Recall our notation $C_{sum}(\cdot,\cdot,\cdot,\cdot)$ from Section III.

Theorem 9. *For $L_f = 1$, the adaptive sum-capacity of a L-user MAC with asymmetric CSI is $C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, (L-1)P_d^{avg})$.*

Proof: We first show that the sum-capacity given in the theorem is an upperbound to achievable throughput $C_{\bar{\Psi}}$ in (9). A communication scheme meeting the upperbound will then be presented, thus proving the theorem. The upper and lower bounds are presented in lemmas 10 and 11 respectively. ■

Lemma 10.

$$C_{\bar{\Psi}} \leq C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, (L-1)P_d^{avg}). \tag{11}$$

Proof: Similar to (9), the average sum-rate T_θ of any scheme θ can be bounded as

$$\begin{aligned} T_\theta &= \int R_1(\bar{h})d\Psi(\bar{h}) + \int d\Psi_d(h)\sum_{i=2}^{L} R_i(h) && (12) \\ &\leq \int R_1(h_1,h_2,0,\cdots,0)d\Psi(h_1,h_2,h_3,\cdots,h_L) \\ &\qquad + \int d\Psi_d(h_2)\sum_{i=2}^{L} R_i(h_2) && (13) \\ &= \int \left(R_1(h_1,h_2) + \sum_{i=2}^{L} R_i(h_2)\right) d\Psi(h_1,h_2) \\ &\leq C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, (L-1)P_d^{avg}). && (14) \end{aligned}$$

Here, the first inequality follows from the fact that the rate of user 1 can only possibly increase if the some of the other links are made inactive, i.e. they have zero gain. The final expression is true since a single user observing channel state h_2 and having average power $(L-1)P_d^{avg}$ can do whatever $L-1$ MAC users observing the same channel can do. Thus the system has sum-capacity equivalent to a 2 user model, i.e. $C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, (L-1)P_d^{avg})$ as given in Section III. ■

An achievable scheme with sum-rate $C_{\bar{\Psi}}$ is proposed now.

Lemma 11.

$$\max_\theta T_\theta \geq C_{sum}(\Psi_f, P_f^{avg}, \Psi_d, (L-1)P_d^{avg}) - \epsilon, \forall \epsilon > 0.$$

Proof: Our assumption of equal average powers for users in S_d makes it straightforward to specify a communication scheme. Imagine a Time Division (TDM) scheme for the $L-1$ users in S_d. Under TDM in S_d, the system is equivalent to a two user MAC, albeit the users can scale their power up to make-up for the inactivity period. Thus each user in S_d appears as if its average active power is $(L-1)P_d^{avg}$. Notice that the RHS is nothing but the adaptive sum-capacity of a two-user system, hence the lemma. ■

Remark 12. *Though Theorem 9 is stated for equal average powers for the users in the set S_d, it easily extends to unequal powers by replacing $(L-1)P_d^{avg}$ by the available average sum-power of the set S_d.*

V. Conclusion

We have shown throughput optimal power-rate allocation strategies in a MAC with asymmetric CSI. While Theorem 7 is shown for a specific distribution like Rayleigh, it applies to all distributions with $E\frac{1}{h^2} = \infty$, which includes many popular fading models. In addition, the result can be extended to general continuous and discrete distributions by considering the possibility of a zero threshold, which however makes the exposition cumbersome. The considered model is somewhat special, where the CSI is either full or distributed. Our future work will consider more general CSI models.

References

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[2] C.-S. Hwang, M. Malkin, A. El Gamal, and J. M. Cioffi, "Multiple-access channels with distributed channel state information," in *International Symposium on Information Theory ISIT*, june 2007, pp. 1561 –1565.

[3] Y. Deshpande, S. R. B. Pillai, and B. K. Dey, "On the sum capacity of multiaccess block-fading channels with individual side informatioin," in *IEEE Information Theory Workshop, Paraty*, 2011.

[4] S. Sreekumar, B. K. Dey, and S. R. B. Pillai, "Adaptive sum-capacity in presence of distributed csi for non-identical links," in *Proc. ISIT*, 2013.

[5] Y. Cemal and Y. Steinberg, "The multiple-access channel with partial state information at the encoders," *Information Theory, IEEE Transactions on*, vol. 51, no. 11, pp. 3992 – 4003, nov. 2005.

[6] K. Iyer, S. R. B. Pillai, and B. K. Dey, "Power controlled adaptive sum capacity in the presence of distributed csi," in *IEEE ISITA*, 2012.

[7] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *ICC '95 Seattle*, vol. 1, Jun. 1995, pp. 331 –335.

[8] K. Singh, S. R. B. Pillai, and B. K. Dey, "On fading mac channels with asymmetric csi," *In preparation*, 2013.

Appendix A
Two user sum-capacity

Let us employ Lagrange multipliers λ_1 and λ_2 for each of the power constraints to obtain the cost function

$$\int \log(1+h_1^2P_1(h_1,h_2)+h_2^2P_2(h_2))d\Psi(h_1,h_2) \\ -\lambda_1\int P_1(h_1,h_2)d\Psi(h_1,h_2) - \lambda_2\int P_2(h_2)d\Psi_d(h).$$

The derivatives with respect to the power allocation functions has to be zero for optimality, whenever non-zero power is allocated. Thus whenever $P_1 > 0$,

$$\frac{h_1^2}{1+h_1^2P_1(h_1,h_2)+h_2^2P_2(h_2)} = \lambda_1, \tag{15}$$

and whenever $P_2 > 0$,

$$\int_{h_1:P_1(h_1,h_2)>0}\frac{h_2^2}{1+h_1^2P_1(h_1,h_2)+h_2^2P_2(h_2)}d\Psi_f(h_1) \\ +\int_{h_1:P_1(h_1,h_2)=0}\frac{h_2^2}{1+h_2^2P_2(h_2)}d\Psi_f(h_1) = \lambda_2. \tag{16}$$

For a given value of h_2 let us denote by $\tilde{H}_1(h_2)$, the set of h_1 for which $P_1(h_1,h_2) > 0$ in a given strategy. Thus, (16) can be re-written as,

$$\frac{h_2^2}{1+h_2^2P_2(h_2)}\int\limits_{\mathcal{H}_1\setminus\tilde{H}_1(h_2)} d\Psi_f(h_1) + \int\limits_{\tilde{H}_1(h_2)}\frac{\lambda_1 h_2^2}{h_1^2}d\Psi_f(h_1) = \lambda_2, \tag{17}$$

where we also used (15), which is valid for the set $\tilde{H}_1(h_2)$. From (15),

$$P_1(h_1,h_2) = \frac{1}{\lambda_1} - \frac{1+h_2^2P_2(h_2)}{h_1^2}. \tag{18}$$

The last two equations can be solved for $P_1(h_1,h_2)$ and $P_2(h_2)$. Alternately, since the problem is convex, we can find λ_1 and λ_2 such that the above equations are satisfied. To this end, let us choose a non-negative threshold value $\gamma(h_2)$ such that $P_1(h_1,h_2) = 0$ whenever $h_1 \leq \gamma(h_2)$. For notational convenience, we will drop the argument and denote $\gamma(h_2)$ as γ. With this, (17) becomes

$$\frac{\Psi_f(\gamma)}{1+h_2^2P_2(h_2)} + \lambda_1 c(\gamma) = \frac{\lambda_2}{h_2^2}, \tag{19}$$

where

$$c(\gamma) = \int_\gamma^\infty \frac{d\Psi_f(h)}{h^2}$$

The power allocation functions are

$$P_2(h_2) = \left(\frac{\Psi_f(\gamma)}{\lambda_2 - \lambda_1 h_2^2 c(\gamma)} - \frac{1}{h_2^2}\right)^+ \tag{20}$$

$$P_1(h_1,h_2) = \left(\frac{1}{\lambda_1} - \frac{\Psi_f(\gamma)h_2^2/h_1^2}{\lambda_2 - \lambda_1 h_2^2 c(\gamma)}\right)^+ \tag{21}$$

In order to simplify the expressions, we now use our assumption of Rayleigh fading, which guarantees the existence of some γ for every h_2 such that $P_1(\gamma,h_2) = 0$. This follows from the fact that $E\frac{1}{h^2} = \infty$ for distributions like Rayleigh. Equating $P_1(\gamma,h_2) = 0$, we get

$$\frac{\lambda_2}{\lambda_1} = \frac{\Psi_f(\gamma)h_2^2}{\gamma^2} + h_2^2 c(\gamma). \tag{22}$$

Rewriting the integrals in RHS,

$$\frac{\lambda_2}{\lambda_1} = h_2^2\int_h \min\{\frac{1}{\gamma^2},\frac{1}{h^2}\}d\Psi_f(h). \tag{23}$$

This equation will enable the computation of γ for every h_2, for a given value of $\frac{\lambda_2}{\lambda_1}$. Eliminating λ_2 from the power allocation equations using (22), we get

$$P_2(h_2) = \left(\frac{\gamma^2}{\lambda_1 h_2^2} - \frac{1}{h_2^2}\right)^+ \text{and } P_1(h_1,h_2) = \left(\frac{1}{\lambda_1} - \frac{\gamma^2}{\lambda_1 h_1^2}\right)^+ \tag{24}$$

Acknowledgments

This work was supported by the Bharti Center for Communication at IIT Bombay.