Lecture 3: Multiple Access Channels

Jun Chen

1 Capacity Region of Multiaccess Channels:

Channel model: $p(y|x_1, x_2, \cdots, x_L)$. Let

$$\mathcal{R}(X_{\mathcal{I}_L}) = \left\{ \sum_{i \in \mathcal{A}} R_i \leq I(X_{\mathcal{A}}; Y | X_{\mathcal{A}^c}), \forall \mathcal{A} \subseteq \mathcal{I}_L \right\},$$

where $\mathcal{I}_L = \{1, 2, \cdots, L\}$. The capacity region $\mathcal{C}$ is given by

$$\mathcal{C} = \operatorname{conv} \left(\bigcup_{\prod_{i=1}^L p(x_i)} \mathcal{R}(X_{\mathcal{I}_L}) \right).$$

Coding Scheme: Encoder i generates 2^{nR_i} codewords $\{\mathbf{X}_i(j)\}_{j=1}^{2^{nR_i}}$, $i \in \mathcal{I}_L$. Suppose $\mathbf{X}_1(1)$, $\mathbf{X}_2(1), \cdots, \mathbf{X}_L(1)$ are transmitted, and $\mathbf{Y}$ is the output. Decoder then picks the codewords that are jointly typical with $\mathbf{Y}$. Clearly, with high probability $\mathbf{X}_2(1), \cdots, \mathbf{X}_L(1)$, and $\mathbf{Y}$ are jointly typical. Furthermore,

$$\begin{aligned} & \mathbb{P}((\mathbf{X}_1(i_1), \cdots, \mathbf{X}_j(i_j), \mathbf{X}_{j+1}(1), \cdots, \mathbf{X}_L(1), \mathbf{Y}) \text{ are jointly typical, for } i_1, \cdots, i_j \neq 1) \\ \leq \; & 2^{n \sum_{k=1}^{j} R_i} 2^{-nI(\mathbf{X}_{\mathcal{I}_j}; \mathbf{Y} | \mathbf{X}_{\mathcal{I}_j^c})} \to 0 \end{aligned}$$

provided $\sum_{k=1}^{j} R_i < I(\mathbf{X}_{\mathcal{I}_j}; \mathbf{Y} | \mathbf{X}_{\mathcal{I}_j^c})$. Other rate constraints can be similarly derived.

Remark: We have shown in the first lecture that if $\mathbf{U}$ and $\mathbf{V}$ are generated according to the marginal distribution of $p(u, v)$, then the probability that $\mathbf{U}$ and $\mathbf{V}$ are jointly typical is approximately $2^{-nI(U;V)}$. If we know that $\mathbf{U}$ and $\mathbf{V}$ are respectively jointly typical with $\mathbf{W}$, then the probability that $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$ are jointly typical is approximately $2^{-nI(U;V|W)}$. In this case, we can think that $\mathbf{U}$ and $\mathbf{V}$ are generated according to the marginal conditional distributions $p(u|w)$ and $p(v|w)$ respectively.

It is clear from the proof that the codewords are jointly decoded. This may cause certain practical issue due to the high decoding complexity. We now proceed to study the low-complexity coding schemes by exploiting the combinatorial and geometric structures of the multiaccess capacity region.

2 Combinatorial and Geometric Structures:

Definition 2.1. *Let $f : 2^{\mathcal{I}_L} \to \mathcal{R}_+$ be a set function. The polyhedron*

$$\mathcal{B}(f) \triangleq \left\{ (x_1, \cdots, x_L) : \sum_{i \in \mathcal{S}} x_i \leq f(\mathcal{S}), \forall \mathcal{S} \subseteq \mathcal{I}_L, x_i \geq 0, \forall i \right\} \tag{1}$$

is a polymatroid if the set function f satisfies

1. *$f(\emptyset) = 0$ (normalized).*

2. *$f(\mathcal{S}) \leq f(\mathcal{T})$ if $\mathcal{S} \subset \mathcal{T}$ (nondecreasing).*

3. *$f(\mathcal{S}) + f(\mathcal{T}) \geq f(\mathcal{S} \cup \mathcal{T}) + f(\mathcal{S} \cap \mathcal{T})$ (submodular).*

If f satisfies the three properties, f is called a rank function.

If π is a permutation on the set $\mathcal{I}_L$, define the vector $\mathbf{v}(\pi) \in \mathcal{R}^L$ by $v_{\pi(1)}(\pi) = f(\pi(1))$ and

$$v_{\pi(i)}(\pi) = f(\{\pi(1), \cdots, \pi(i)\}) - f(\{\pi(1), \cdots, \pi(i-1)\})$$

for $i = 2, 3, \cdots, L$.

Theorem 2.1. *Let $\mathcal{B}(f)$ be a polymatroid. Then $v(\pi)$ is a vertex of $\mathcal{B}(f)$ for every permutation π. Also, any vertex of $\mathcal{B}(f)$ strictly inside the positive orthant must be $\mathbf{v}(\pi)$ for some π. Moreover, if λ is a given in $\mathcal{R}_+^L$, then a solution of the optimization problem*

$$\max \lambda \cdot \mathbf{x} \quad \textit{subject to } \mathbf{x} \in \mathcal{B}(f)$$

is attained at a point $\mathbf{v}(\pi^*)$, where the π^* is any permutation such that $\lambda_{\pi^*(1)} \geq \cdots \geq \lambda_{\pi^*(L)}$. Conversely, suppose f is a set function and $\mathcal{B}(f)$ is the polyhedron defined in (1). Then if $\mathbf{v}(\pi) \in \mathcal{B}(f)$ for every permutation π, then $\mathcal{B}(f)$ is a polymatroid.

Remark:

1. To understand the "submodular" condition, let's replace "$\geq$" by "$>$". Suppose there exists a vector $(x_1, \cdots, x_L) \in \mathcal{B}(f)$ with $\sum_{i \in \mathcal{S}} x_i = f(\mathcal{S})$ and $\sum_{i \in \mathcal{T}} x_i = f(\mathcal{T})$ for some $\mathcal{S}$, $\mathcal{T}$ such that $\mathcal{S} \not\subseteq \mathcal{T}$, $\mathcal{T} \not\subseteq \mathcal{S}$. Then we have

$$\begin{aligned} \sum_{i \in \mathcal{S}} x_i + \sum_{i \in \mathcal{T}} x_i &= f(\mathcal{S}) + f(\mathcal{T}) \\ &> f(\mathcal{S} \cup \mathcal{T}) + f(\mathcal{S} \cap \mathcal{T}) \\ &= \sum_{i \in \mathcal{S} \cup \mathcal{T}} x_i + \sum_{i \in \mathcal{S} \cap \mathcal{T}} x_i \\ &= \sum_{i \in \mathcal{S}} x_i + \sum_{i \in \mathcal{T}} x_i, \end{aligned}$$

a contradiction. So when $(x_1, \cdots, x_L)$ meets constraints $f(\mathcal{S})$ and $f(\mathcal{T})$, we must have either $\mathcal{S} \subseteq \mathcal{T}$ or $\mathcal{T} \subseteq \mathcal{S}$. Since a vertex of $\mathcal{B}(f)$ must meet L constraints, the sets associated with these L constraints must be of the form $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots \subset \mathcal{A}_L$. There are totally $L!$ ways.

2. For vertex $\mathbf{v}(\pi)$, we have

$$\begin{aligned} \sum_{i=1}^{k} v_{\pi(i)}(\pi) &= f(\{\pi(1), \cdots, \pi(k)\}) \\ &\geq \sum_{i=1}^{k} x_{\pi(i)} \quad \forall \mathbf{x} \in \mathcal{B}(f). \end{aligned}$$

3. we can write

$$\sum_{i=1}^{L}\lambda_i x_i=\sum_{j=1}^{L}(\lambda_{\pi^*(L-j+1)}-\lambda_{\pi^*(L-j+2)})\sum_{i=1}^{L-j+1}x_{\pi^*(i)},$$

   where $\lambda_{\pi^*(L+1)}=0$. See why the maximum is attained at $\mathbf{v}(\pi^*)$?

Theorem 2.2. *For any independent distribution $\prod_{i=1}^{L}p(x_i)$ on the inputs, the polyhedron $\mathcal{R}(X_{\mathcal{I}_L})$ is a polymatroid.*

Proof. The "normalized" part is obvious. For $\mathcal{S}\subset\mathcal{T}$, we have

$$\begin{aligned}I(X_{\mathcal{S}};Y|X_{\mathcal{S}^c}) &= H(Y|X_{\mathcal{S}^c})-H(Y|X_{\mathcal{I}_L})\\ &\leq H(Y|X_{\mathcal{T}^c})-H(Y|X_{\mathcal{I}_L})\\ &= I(X_{\mathcal{T}};Y|X_{\mathcal{T}^c}),\end{aligned}$$

which gives the "nondecreasing" property.

To verify the "submodular" property, we have

$$\begin{aligned}&I(X_{\mathcal{S}};Y|X_{\mathcal{S}^c})+I(X_{\mathcal{T}};Y|X_{\mathcal{T}^c})\\ =\;& H(X_{\mathcal{S}}|X_{\mathcal{S}^c})-H(X_{\mathcal{S}}|X_{\mathcal{S}^c},Y)+H(X_{\mathcal{T}}|X_{\mathcal{T}^c})-H(X_{\mathcal{T}}|X_{\mathcal{T}^c},Y)\\ =\;& H(X_{\mathcal{S}})-H(X_{\mathcal{S}}|X_{\mathcal{S}^c},Y)+H(X_{\mathcal{T}})-H(X_{\mathcal{T}}|X_{\mathcal{T}^c},Y)\\ =\;& H(X_{\mathcal{S}\cap\mathcal{T}})-H(X_{\mathcal{S}}|X_{\mathcal{S}^c},Y)+H(X_{\mathcal{S}\cup\mathcal{T}})-H(X_{\mathcal{T}}|X_{\mathcal{T}^c},Y)\\ =\;& H(X_{\mathcal{S}\cap\mathcal{T}})-H(X_{\mathcal{S}\setminus\mathcal{T}}|X_{\mathcal{S}^c},Y)-H(X_{\mathcal{S}\cap\mathcal{T}}|X_{(\mathcal{S}\cap\mathcal{T})^c},Y)+H(X_{\mathcal{S}\cup\mathcal{T}})-H(X_{\mathcal{T}}|X_{\mathcal{T}^c},Y)\\ \geq\;& H(X_{\mathcal{S}\cap\mathcal{T}})-H(X_{\mathcal{S}\setminus\mathcal{T}}|X_{(\mathcal{S}\cup\mathcal{T})^c},Y)-H(X_{\mathcal{S}\cap\mathcal{T}}|X_{(\mathcal{S}\cap\mathcal{T})^c},Y)+H(X_{\mathcal{S}\cup\mathcal{T}})-H(X_{\mathcal{T}}|X_{\mathcal{T}^c},Y)\\ =\;& H(X_{\mathcal{S}\cap\mathcal{T}})-H(X_{\mathcal{S}\cap\mathcal{T}}|X_{(\mathcal{S}\cap\mathcal{T})^c},Y)+H(X_{\mathcal{S}\cup\mathcal{T}})-H(X_{\mathcal{S}\cup\mathcal{T}}|X_{(\mathcal{S}\cup\mathcal{T})^c},Y)\\ =\;& I(X_{\mathcal{S}\cap\mathcal{T}};Y|X_{(\mathcal{S}\cap\mathcal{T})^c})+I(X_{\mathcal{S}\cup\mathcal{T}};Y|X_{(\mathcal{S}\cup\mathcal{T})^c}).\end{aligned}$$

□

The vertices of $\mathcal{R}(X_{\mathcal{I}_L})$ are thus given by

$$\begin{aligned}R_{\pi(i)}(\pi) &= I(X_{\pi(1)},\cdots,X_{\pi(i)};Y|X_{\pi(i+1)},\cdots,X_{\pi(L)})-I(X_{\pi(1)},\cdots,X_{\pi(i-1)};Y|X_{\pi(i)},\cdots,X_{\pi(L)})\\ &= I(X_{\pi(i)};Y|X_{\pi(i+1)},\cdots,X_{\pi(L)}),\quad i=1,2,\cdots,L-1\\ R_{\pi(L)}(\pi) &= I(X_{\pi(L)};Y)\end{aligned}$$

for every π. From the above expressions, we can see that each vertex $\mathbf{R}(\pi)$ is associated with a successive decoding scheme: First decode the codeword from user $\pi(L)$, then the codeword from user $\pi(L-1)$, and so on. Note: we have

$$\begin{aligned}I(X_{\mathcal{S}};Y|X_{\mathcal{S}^c}) &= H(X_{\mathcal{S}})-H(X_{\mathcal{S}}|Y,X_{\mathcal{S}^c})\\ &\geq H(X_{\mathcal{S}})-H(X_{\mathcal{S}}|Y)\\ &= I(X_{\mathcal{S}};Y).\end{aligned}$$

We can view $I(X_{\mathcal{S}};Y)$ as the achieve rate from user group $\mathcal{S}$ to the receiver without any side information, and $I(X_{\mathcal{S}};Y|X_{\mathcal{S}^c})$ as the achievable rate with decoded codewords from user group $\mathcal{S}^c$ as the side information. So by exploiting the side information retrieved from the decoded codewords, one can improve the communication rate of the remaining users.

Since we have

$$\sum_{i=1}^{L} R_i(\pi) = I(X_{\mathcal{I}_L}; Y),$$

for any vertex $\mathbf{R}(\pi)$, the sum-rate constraint is always achievable. Define

$$\mathcal{D}(X_{\mathcal{I}_L}) = \left\{ \mathbf{R} \in \mathcal{R}(X_{\mathcal{I}_L}) : \sum_{i=1}^{L} R_i = I(X_{\mathcal{I}_L}; Y) \right\}.$$

We refer to $\mathcal{D}(X_{\mathcal{I}_L})$ as the dominant face of $\mathcal{R}(X_{\mathcal{I}_L})$. Note: $\mathcal{D}(X_{\mathcal{I}_L})$ is always non-empty.

Since any rate tuple strictly inside of $\mathcal{R}(X_{\mathcal{I}_L})$ is inferior to some point on the dominant face in the sense of communication rate, henceforth we shall only focus on $\mathcal{D}(X_{\mathcal{I}_L})$. Clearly, $\mathcal{D}(X_{\mathcal{I}_L})$ is the convex hull of vertices $\mathbf{R}(\pi)$, $\forall \pi$. Therefore, besides using the joint decoding scheme in Section 1, we can also achieve the whole region via timesharing vertices (and thus using the successive decoding scheme).

For non-degenerate cases (which we will assume from now on), all the vertices are distinct (thus totally $L!$ vertices with each corresponding to a permutation π) and the dimension of $\mathcal{D}(X_{\mathcal{I}_L})$ is $L-1$ (Note: $\mathcal{R}(X_{\mathcal{I}_L})$ is of dimension L, $\mathcal{D}(X_{\mathcal{I}_L})$ is the intersection of $\mathcal{R}(X_{\mathcal{I}_L})$ with a hyperplane). Now we shall take a close look at the structure of the dominant face.

For $\emptyset \subset \mathcal{A} \subset \mathcal{I}_L$, define the hyperplane

$$\mathcal{H}(\mathcal{A}) = \left\{ \mathcal{R}_{\mathcal{I}_L} \in \mathcal{R}^L : \sum_{i \in \mathcal{A}} R_i = I(X_{\mathcal{A}}; Y) \right\}$$

and let $\mathcal{F}_{\mathcal{A}} = \mathcal{H}(\mathcal{A}) \cap \mathcal{D}(X_{\mathcal{I}_L})$. If $\emptyset \subset \mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots \subset \mathcal{A}_k \subset \mathcal{I}_L$, is a telescopic sequence of subsets, then $\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}$ is a face of $\mathcal{D}(X_{\mathcal{I}_L})$. Conversely, every face of $\mathcal{D}(X_{\mathcal{I}_L})$ can be written in this form. Let $\mathcal{B}_i = \mathcal{A}_i - \mathcal{A}_{i-1}$, $i = 1, 2, \cdots, k+1$, where we set $\mathcal{A}_0 = \emptyset$ and $\mathcal{A}_{k+1} = \mathcal{I}_L$. Let Ξ be the set of permutation π on $\mathcal{I}_L$ such that

$$\left\{ \pi\left(\sum_{j=0}^{k-i} |\mathcal{B}_{k+1-j}| + 1 \right), \cdots, \pi\left(\sum_{j=0}^{k+1-i} |\mathcal{B}_{k+1-j}| \right) \right\} = \mathcal{B}_i, \quad i = 1, 2, \cdots, k+1.$$

Each permutation $\pi \in \Xi$ is associated with a vertex of $\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}$ and vice versa. Hence $\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}$ has totally $|\Xi| = \prod_{i=1}^{k+1} (|\mathcal{B}_i|!)$ vertices. Moreover, we have $dim(\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}) = L-k-1$. We can also count that the number of faces of dimension $L-k-1$ is $\binom{n}{k}(k+1)!$, $k = 0, 1, \cdots, L-1$.

The decoding order for rate tuples on $\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}$: First decode user group $\mathcal{B}_1$, then $\mathcal{B}_2$, and so on.

3 Rate Splitting:

So far we have two methods to achieve rate tuples on the dominant face other than those $L!$ vertices. The first method to achieve these difficult rate tuples was time sharing between vertices. This approach can require as many as L successive coding schemes[1], each scheme

[1]By Carathéodory's fundamental theorem, any point in the convex closure of a connected commpact set $\mathcal{A}$ in a d dimensional Euclidean space can be represented as a convex combination of $d+1$ or fewer points in the original set $\mathcal{A}$.

requiring L coding steps. The second approach to achieve these rate tuples is joint decoding of all users. This is very difficult to implement in practice since random codes have a decoding complexity of the order of $2^{nI(X_{\mathcal{I}_L};Y)}$, where n is the block length. We now show that any rate tuple in $\mathcal{R}(W_{\mathcal{I}_L})$ can be achieved by a low-complexity successive coding scheme with at most $2L-1$ steps.

Let's see a simple example first. Let (R_1, R_2) be a rate pair on the dominant face of $\mathcal{R}(X_{\mathcal{I}_2})$. We must have

$$R_1 + R_2 = I(X_1, X_2; Y),$$
$$R_1 \in [I(X_1; Y), I(X_1; Y|X_2)],$$
$$R_2 \in [I(X_2; Y), I(X_2; Y|X_1)].$$

Introduce an auxiliary random variable X_2' such that $X_2' \to X_2 \to (X_1, Y)$. It is clear that we can find a class of transition probability $p_\epsilon(x_2'|x_2)$ such that $I(X_2; X_2')$ varies from 0 to $H(X_2)$ as ϵ goes from 0 to 1. Hence for X_2' associated with certain ϵ, we have

$$\begin{aligned} R_1 &= I(X_1; Y|X_2'), \\ R_2 &= I(X_2'; Y) + I(X_2; Y|X_1, X_2'). \end{aligned}$$

An intuitive interpretation of the above expression: User 1 generates a codebook X_1 with $2^{nI(X_1;Y|X_2')}$ codewords. User 2 first generates a codebook X_2' with $2^{nI(X_2';Y)}$ codewords; then for each codeword in codebook X_2', generates a superposition codebook $X_2(X_2')$ with $2^{nI(X_2;Y|X_1,X_2')}$. Receiver first decodes X_2', then X_1, and finally X_2.

We now generalize this scheme to the general case.

Suppose rate tuple $R_{\mathcal{I}_L}$ is on the dominant face of $\mathcal{R}(X_{\mathcal{I}_L})$. Let $X_L' \to X_L \to (X_{\mathcal{I}_{L-1}}, Y)$ such that

$$\sum_{i\in\mathcal{A}} R_i \geq I(X_{\mathcal{A}}; Y|X_L'), \quad \forall \mathcal{A} \subseteq \mathcal{I}_{L-1}$$

and $\exists \mathcal{T} \subseteq \mathcal{I}_{L-1}$ with

$$\sum_{i\in\mathcal{T}} R_i = I(X_{\mathcal{T}}; Y|X_L').$$

Let $R_i' = R_i$, $i = 1, 2, \cdots, L-1$, and $R_L' = R_L - I(X_L', Y)$. It is easy to verify that

$$\begin{aligned} \sum_{i\in\mathcal{T}} R_i' &= I(X_{\mathcal{T}}; Y|X_L'), \\ \sum_{i\in\mathcal{T}} R_i' &= I(X_{\mathcal{I}_L}; Y) - I(X_{\mathcal{T}}; Y|X_L') - I(X_L', Y) \\ &= I(X_{\mathcal{I}_L}, X_L'; Y) - I(X_{\mathcal{T}}; Y|X_L') - I(X_L', Y) \\ &= I(X_{\mathcal{I}_L}; Y|X_L') - I(X_{\mathcal{T}}; Y|X_L') \\ &= I(X_{\mathcal{T}^c}; Y|X_{\mathcal{T}}, X_L'). \end{aligned}$$

So the receiver first decodes X_L', then $X_{\mathcal{T}}$, and finally $X_{\mathcal{T}^c}$. By induction, if $L-1$ users need $2(L-1)-1$ successive coding steps, then L users need $2|\mathcal{T}| - 1 + 2|\mathcal{T}^c| - 1 + 1 = 2L-1$ steps. Note: $\max(|\mathcal{T}|, |\mathcal{T}^c|) \leq L-1$.

Note: $2L-1$ is just an upper bound, for rate tuples on the boundary of $\mathcal{D}(X_{\mathcal{I}_L})$, the coding steps can be further reduced. Specifically, for $R_{\mathcal{I}_L} \in \{\mathcal{F}_{\mathcal{A}_1} \cap \mathcal{F}_{\mathcal{A}_2} \cap \cdots \cap \mathcal{F}_{\mathcal{A}_k}\}$, one can observe that the splitting can be carried out in $\mathcal{B}_i$, $i = 1, 2, \cdots, k+1$ separately. So the number of successive coding steps is $\sum_{i=1}^{k+1}(2|\mathcal{B}_i| - 1) = 2L - k - 1$. Geometrically, for rate tuples on a d-dimensional face of $\mathcal{D}(X_{\mathcal{I}_L})$, $L + d$ coding steps suffice.

It can be further shown that each user gets split at most twice. But the proof is quite lengthy. It will be very nice if someone can come up with a simple proof.

Although the rate splitting method gives us a low-complexity decoding scheme, it is not optimal in the sense of error exponent because the first decoding step becomes the bottle neck. We will study this in detail when we analyze the diversity-multiplexing tradeoff in multiaccess channels.

4 Resource Allocation:

Channel model: $Y(t) = \sum_{i=1}^{L} \sqrt{H_i(t)} X_i(t) + Z(t)$. $X_i(t)$ and $H_i(t)$ are the transmitted waveform and the fading process of the ith user, and $Z(t)$ is white Gaussian noise with variance σ^2. We assume that the fading processes for all users are jointly stationary and ergodic. User i is also subject to an average transmitter power constraint of $\overline{P}_i$.

Fixed fading: $H_i(t) = h_i$ for all t.

$$\mathcal{C}(\mathbf{h}, \overline{\mathbf{P}}) = \left\{ R_{\mathcal{I}_L} : \sum_{i \in \mathcal{A}} R_i \leq \frac{1}{2} \log \left(1 + \frac{\sum_{i \in \mathcal{A}} h_i \overline{P}_i}{\sigma^2} \right), \forall \mathcal{A} \subseteq \mathcal{I}_L \right\},$$

where $\mathbf{h} = (h_1, \cdots, h_L)$ and $\overline{\mathbf{P}} = (\overline{P}_1, \cdots, \overline{P}_L)$.

Fading processes are known at receiver but not transmitter:

$$\mathcal{C}(\overline{\mathbf{P}}) = \left\{ R_{\mathcal{I}_L} : \sum_{i \in \mathcal{A}} R_i \leq \mathbb{E}_{\mathbf{H}} \left[\frac{1}{2} \log \left(1 + \frac{\sum_{i \in \mathcal{A}} H_i \overline{P}_i}{\sigma^2} \right) \right], \forall \mathcal{A} \subseteq \mathcal{I}_L \right\},$$

where $\mathbf{H} = (H_1, \cdots, H_L)$ is a random vector having the stationary distribution of the joint fading process.

The capacity region under dynamic resource allocation:

A power-control policy $\mathcal{P}$ is a mapping from the fading state space to $\mathcal{R}_+^L$. Given a joint fading state $\mathbf{h}$ for the users, $\mathcal{P}_i(\mathbf{h})$ can be interpreted as the transmitter power allocated to user i. Let

$$\mathcal{C}(\mathcal{P}) = \left\{ R_{\mathcal{I}_L} : \sum_{i \in \mathcal{A}} R_i \leq \mathbb{E}_{\mathbf{H}} \left[\frac{1}{2} \log \left(1 + \frac{\sum_{i \in \mathcal{A}} H_i \mathcal{P}_i(\mathbf{H})}{\sigma^2} \right) \right], \forall \mathcal{A} \subseteq \mathcal{I}_L \right\}.$$

We have

$$\mathcal{C}(\overline{\mathbf{P}}) = \bigcup_{\mathcal{P} \in \mathcal{F}} \mathcal{C}(\mathcal{P})$$

where $\mathcal{F}$ is the set of all feasible power control policies satisfying the average power constraint

$$\mathcal{F} = \{\mathcal{P} : E_{\mathbf{H}}[\mathcal{P}_i(\mathbf{H})] \leq \overline{P}_i, \forall i\}.$$

Reminder: The multiplexing method can be replaced by dynamic power control.

A rate allocation policy $\mathcal{R}$ is a mapping from the set of joint fading states to $\mathcal{R}_+^L$; for each fading state $\mathbf{h}$, $\mathcal{R}_i(\mathbf{h})$ can be interpreted as the rate allocated to user i while the users are in state $\mathbf{h}$.

Theorem 4.1. *The boundary surface of $\mathcal{C}(\overline{\mathbf{P}})$ is the closure of all points $R^*_{\mathcal{I}_L}$ such that $R^*_{\mathcal{I}_L}$ is a solution to the optimization problem*

$$\max_{R_{\mathcal{I}_L}} \mu \cdot R_{\mathcal{I}_L} \text{ subject to } R_{\mathcal{I}_L} \in \mathcal{C}(\overline{\mathbf{P}})$$

for some positive $\mu \in \mathcal{R}^L_+$. For a given μ, $R^*_{\mathcal{I}_L}$ is a solution to the above problem if and only if there exists a $\lambda \in \mathcal{R}^L_+$, rate allocation policy $\mathcal{R}(\cdot)$, and power control policy $\mathcal{P}(\cdot)$ such that for every joint fading state $\mathbf{h}$, $(\mathcal{R}(\mathbf{h}), \mathcal{P}(\mathbf{h}))$ is a solution to the optimization problem

$$\max_{\mathbf{r},\mathbf{p}} \mu \cdot \mathbf{r} - \lambda \cdot \mathbf{p} \text{ subject to } \mathbf{r} \in \mathcal{C}(\mathbf{h}, \mathbf{p})$$

and $\mathbb{E}_{\mathbf{H}}[\mathcal{R}_i(\mathbf{H})] = R^*_i$, $\mathbb{E}_{\mathbf{H}}[\mathcal{P}_i(\mathbf{H})] = \overline{P}_i$, $i = 1, 2, \cdots, L$.

Proof. Introducing Lagrangian multipliers $\lambda \in \mathcal{R}^L_+$, we write the optimization problem as

$$\max_{\mathbf{r},\mathbf{p}} \mu \cdot R_{\mathcal{I}_L} - \lambda \cdot \mathbb{E}_{\mathbf{H}}[\mathcal{P}(\mathbf{H})] \text{ subject to } R_{\mathcal{I}_L} \in \mathcal{C}(\mathcal{P}), \mathcal{P} \in \mathcal{F}.$$

To get the optimal rate allocation policy $\mathcal{R}$ for a fixed power allocation policy $\mathcal{P}$, we invoke Theorem 2.1. Let π be the permutation corresponding to a decreasing ordering of the components of the vector μ. By the polymatroid structure of $\mathcal{C}(\mathcal{P})$, for any given power control $\mathcal{P}$, $\mu \cdot R_{\mathcal{I}_L}$ is maximized at

$$\begin{aligned} R_{\pi(1)} &= \mathbb{E}_{\mathbf{H}}\left[\frac{1}{2}\log\left(1 + \frac{H_{\pi(1)}\mathcal{P}_{\pi(1)}(\mathbf{H})}{\sigma^2}\right)\right] \\ R_{\pi(k)} &= \mathbb{E}_{\mathbf{H}}\left[\frac{1}{2}\log\left(1 + \frac{H_{\pi(k)}\mathcal{P}_{\pi(k)}(\mathbf{H})}{\sigma^2 + \sum_{i=1}^{k-1} H_{\pi(i)}\mathcal{P}_{\pi(i)}}\right)\right], \quad k = 2, \cdots, L. \end{aligned}$$

Thus the optimization problem becomes

$$\max_{\mathcal{P}} \mu_1 \mathbb{E}_{\mathbf{H}}\left[\frac{1}{2}\log\left(1 + \frac{H_{\pi(1)}\mathcal{P}_{\pi(1)}(\mathbf{H})}{\sigma^2}\right)\right] + \sum_{k=2}^{L} \mu_k \mathbb{E}_{\mathbf{H}}\left[\frac{1}{2}\log\left(1 + \frac{H_{\pi(k)}\mathcal{P}_{\pi(k)}(\mathbf{H})}{\sigma^2 + \sum_{i=1}^{k-1} H_{\pi(i)}\mathcal{P}_{\pi(i)}}\right)\right] - \lambda \cdot \mathbb{E}_{\mathbf{H}}[\mathcal{P}(\mathbf{H})].$$

Since the permutation π does not depend on $\mathbf{H}$, we can optimize the power control policy $\mathbf{p} = \mathcal{P}(\mathbf{h})$ for every fading state $\mathbf{h}$, i.e.,

$$\max_{\mathbf{p}} \mu_1 \frac{1}{2}\log\left(1 + \frac{h_{\pi(1)}p_{\pi(1)}}{\sigma^2}\right) + \sum_{k=2}^{L} \mu_k \frac{1}{2}\log\left(1 + \frac{h_{\pi(k)}p_{\pi(k)}}{\sigma^2 + \sum_{i=1}^{k-1} h_{\pi(i)}p_{\pi(i)}}\right) - \lambda \cdot \mathbf{p}.$$

This is, in turn, equivalent to

$$\max_{\mathbf{r},\mathbf{p}} \mu \cdot \mathbf{r} - \lambda \cdot \mathbf{p} \text{ subject to } \mathbf{r} \in \mathcal{C}(\mathbf{h}, \mathbf{p}).$$

□

To get the maximum sum-rate, we let $\mu_1 = \cdots = \mu_L = 1$. The optimization is simplified to

$$\max_{\mathbf{p}} \frac{1}{2}\log\left(1 + \frac{\mathbf{h} \cdot \mathbf{p}}{\sigma^2}\right) - \lambda \cdot \mathbf{p}.$$

Take derivative with respect to p_i, we get

$$\frac{h_i}{2(\sigma^2 + \mathbf{h} \cdot \mathbf{p})} - \lambda_i, \quad i = 1, \cdots, L.$$

Since $\frac{\lambda_1}{H_1} \neq \cdots \neq \frac{\lambda_L}{H_L}$ with probability 1, this means that in the optimal strategy, at most one user is allowed to transmit at any given fading state. The optimal power control policy $\mathcal{P}^*$ can readily calculated to be

$$\mathcal{P}^*(\mathbf{h}, \lambda) = \begin{cases} (\frac{1}{2\lambda_i} - \frac{\sigma^2}{h_i})^+, & \text{if } h_i > \frac{\lambda_i}{\lambda_j} h_j \text{ for all } j \\ 0, & \text{otherwise}, \end{cases}$$

where λ is determined by the power constraint. Note that this power control gives rise to a time-division multiple-access strategy.

References

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Information Theory.* Budapest, Hungary: Hungarian Acad. Sci., 1971, pp. 23-52.

[2] H. Liao, "Multiple access channels," Ph.D. dissertation, Dept. Elec. Eng., Univ. Hawaii, Honolulu, 1972.

[3] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: Wiley, 1991.

[4] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, to be published by Cambridge University Press. Online: http://degas.eecs.berkeley.edu/dtse.

[5] J. Edmonds, "Submodular functions, matroids and certain polyhedra, in *Combinatorial structures and their applications* (R. Guy, H. Hanani, N. Sauer, and J. Schonheim, eds.), pp. 69-87, Gordon and Breach, New York, 1970. (Proc. Calgary Int. Conf. 1969).

[6] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," presented at the International Conference on Communications, Seattle, WA, June 1995.

[7] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels-part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Info. Theory*, vol. 44, pp. 2796-2815, Nov. 1998.

[8] B. Rimoldi and R. Urbanke, "On the structure of the dominant face of multiple access channels," *Proc. Information Theory and Communications Workshop*, June 20-25, 1999, Kruger National Park, South Africa, pp. 12-14.

[9] A. B. Carleial, "On the capacity of multiple-terminal communication networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, Aug. 1975.

[10] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple- access channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 364-375, Mar. 1996.

[11] A. J. Grant, B. Rimoldi, R. L. Urbanke, and P. A. Whiting, "Rate-splitting multiple access for discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 873-890, Mar. 2001.

[12] B. Rimoldi, “Generalized time sharing: a low-complexity capacity-achieving multiple- access technique,” *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2432-2442, Sept. 2001.

[13] D. Tse, P. Viswanath, and L. Zheng ”Diversity-multiplexing tradeoff in multiple access channels” , *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1859-1874, Sept. 2004.