Symmetry and Asymmetry of MIMO Fading Channels

Emmanuel Abbe, Emre Telatar, *Member, IEEE*, and Lizhong Zheng, *Member, IEEE*

***Abstract*— We consider ergodic coherent MIMO channels, and characterize the optimal input distribution under general distributions of the fading matrix. We describe how symmetries in the fading matrix distribution and the constraint set, described naturally with group structures, imply the symmetric properties of the optimal input distribution. We use this approach to give a general condition under which the white Gaussian input is optimal, and to compute the optimal input for MISO Ricean channel. In contrast, we also investigate the Kronecker model, in which case we show how an asymmetric structure in the problem is also preserved in the optimal input, leading to a new water-filling solution.**

I. Introduction

Vector fading channel of the form (1) has been studied thoroughly in the recent years. The optimal input distribution of x is, however, only characterized in some special cases. In this paper, we try to strengthen the known results on the capacity achieving input distributions by exploiting the symmetric properties of the channel.

$$y = Hx + n \tag{1}$$

We assume in (1) that $x \in \mathcal{X} = \mathbb{C}^t$, and $y \in \mathcal{Y} = \mathbb{C}^r$, where t and r are the number of transmit and receive antennas, respectively. We assume that at each time i, the $r \times t$ matrix H_i is drawn from an ergodic process having marginal probability measure μ_H. In this paper, we will focus on the case where H_i is perfectly known at the receiver, i.e., the coherent communication model. We assume the additive noise n is drawn i.i.d from a complex circularly-symmetric gaussian (C.C.S.G.) random variable of covariance matrix K, independently from the H_i's.

Moreover, the inputs $\{x_i\}$ are constrained in the following way. If we want to use a code $\mathcal{C} = \{c(1), \dots, c(M)\} \subset \mathcal{X}^n$, $n \geq 1$, then the code has to satisfy that for every $1 \leq m \leq M$: $\frac{1}{n}\sum_{i=1}^n c(m)_i c(m)_i^* \in \mathrm{D}_t$, where $\mathrm{D}_t \subset H_+(t)$ is a given compact set — we use $H_+(n)$ to denote the set of hermitian positive definite matrices of size n.

A particular example of such a constraint is when one consider D_t to be $\{A \in H_+(t) | \mathrm{tr} A \leq P\}$, for a given $P \in \mathbb{R}$. This is equivalent to ask for $\mathbb{E}X^*X \leq P$ and is called the *total power constraint*. An *individual power constraint* can also be considered, i.e. when $\mathbb{E}|Xi|^2 \leq P_i$, for given $P_i \in \mathbb{R}$, $1 \leq i \leq t$, then the set D_t would be $\{A \in H_+(t) | A_{ii} \leq P_i, 1 \leq i \leq t\}$.

E. Abbe and L. Zheng are with the Massachusetts Institute of Technology, Cambridge MA 02139, USA, and E. Telatar is with the EPFL, Lausanne 1015, CH. (email: eabbe@mit.edu,lizhong@mit.edu,emre.telatar@epfl.ch).

Let C be the capacity of this channel under this general constraint. Then, denoting by X a random vector (r.ve.) in $\mathbb{C}^t$, we know from standard information theoretic arguments that

$$C(\mu_H, \mathrm{D}_t) = \max_{X:\mathbb{E}XX^* \in \mathrm{D}_t} I(X; Y, H)$$

Under the coherent assumption, the optimal input of x is alway Gaussian. The optimal covariance matrix of x, is however not characterized for the general cases. In the special case that H has i.i.d. C.C.S.G. entries, with a total power constraint of P, it is shown in [1] that the optimal input covariance matrix is white, $\mathbb{E}XX^* = \frac{P}{t} I_t$, and the capacity increase linearly with $\min t, r$. The proof of this statement uses the concavity of the mutual information function, as well as the symmetry of the channel. In a nutshell, since any rotation of the input distribution would not change the resulting mutual information, the optimal input has to be isotropically distributed.

The main questions addressed in the current paper is how to address the symmetry of a vector channel, and how does the symmetry in the channel affect the optimal input distribution. We argue that group structures are the natural tool for this task. Using the notion of groups, we can charaterize a number of different symmetries offered by the vector channel, most of which are weaker than the special case in [1]. We explore how does such symmetric structures translate into the symmetry in the optimal input. In particular, we prove a more general condition, than [1], for the optimal input distribution to be white Gaussian.

Moreover, we also present some partial results when such symmetric property is absent from the channel.

II. General expression of the capacity

Definition 1: We define the optimal inputs by

$$X_{\mathrm{opt}}(\mu_H, \mathrm{D}_t) = \arg\max_{X \in \mathrm{C}_t} I(X; HX + N, H),$$

where $\arg\max_{X\in \mathrm{C}_t} f(X)$, for a real function f, denotes the set of the elements x satisfying $f(x) \geq f(y)$, $\forall y \in \mathrm{C}_t$. We now use the assumptions we made on the channel to give a more specific expression for the capacity and the optimal inputs. The fact that the gaussian distribution maximizes the entropy under a covariance constraint lead to the following result.

Proposition 1: Let

$$\psi : Q \in \mathrm{D}_t \mapsto \mathbb{E}^{\mu_H} \log\det(I + K^{-1}HQH^*) \in \mathbb{R}, \qquad (2)$$

which we call the mutual information function. Then, according to previous definitions and assumptions, we have

$$X_{\mathrm{opt}}(\mu_H, \mathrm{D}_t) \sim \mathcal{N}_{\mathbb{C}^t}(Q_{\mathrm{opt}}),$$

where

$$Q_{\mathrm{opt}}(\mu_H, \mathrm{D}_t) = \arg\max_{Q \in \mathrm{D}_t} \psi(Q)$$

and

$$C(\mu_H, \mathrm{D}_t) = \max_{Q \in \mathrm{D}_t} \psi(Q).$$

n D_t. We can also argue that we can restrict the domain of maximization to $\partial\mathrm{D}_t$, since the derivative of $\log\det(I + K^{-1}HQH^*)$ with respect to any entries of Q is monotone (by concavity of $Q \rightsquigarrow \log\det(I + K^{-1}HQH^*)$).

III. Symmetries

A. Quantifying symmetries

Assume that the channel has the same output distribution when sending any input X or a permuted version of it, say, PX, where P is a permutation matrix.

$$Y = HX + N \Longleftrightarrow Y = H(PX) + N.$$

Then, we talk about a *symmetry of the channel* with respect to that transformation P. But from previous equivalence, this is to say that

$$HP \overset{(d)}{\simeq} H.$$

Remarks:

1) This type of invariance has a natural group structure: assume you have the invariance $HP_i \overset{(d)}{\simeq} H$ for a set of matrices P_i, then clearly $HP_iP_j \overset{(d)}{\simeq} H$ and if P_i is invertible $HP_i^{-1} \overset{(d)}{\simeq} H$. Thus this invariance still holds for the group generated by this set.
2) In order to compare X and PX, we need to ensure that PX is satisfying the considered contraint too, i.e. its covariance matrix $P(\mathbb{E}XX^*)P^*$ has to belong to D_t as well.
3) Other group than the permutations might be of interest, for example if we want to consider situation where the symmetry is expressed by keeping the channel equivalent whether we send an input X or a modified version of it where some component's signs have been flipped, then the group of diagonal matrices with 1 and -1 is the appropriate group.

These remarks motivate the following definitions.

Definition 2: Let G be a group in $M_n(\mathbb{C})$.

1) A random matrix is G-invariant (on the right) if $Hg \overset{(d)}{\simeq} H$, $\forall g \in G$.
2) A set of matrices $D \subset M_n(\mathbb{C})$ is invariant in G-conjugation if $gQg^{-1} \in D$, $\forall Q \in D, g \in G$.
3) A function $\Psi : D \to \mathbb{R}$ is G-invariant if D is invariant in G-conjugation and if $\Psi(gQg^{-1}) = \Psi(Q)$, $\forall Q \in D, g \in G$.

Note that only subgroups of unitary matrices are of our interest regarding our MIMO channel setting, because the mutual information function evaluated at Q depends on the distribution of HQH^*. Examples of functions which are invariant in G-conjugation for unitary subgroups are all functions of the form $x \rightsquigarrow \mathbb{E}f(MxM^*)$ where f is any measurable function and M is a random matrix that is G-invariant on the right. The reason for which a "conjugation" invariance for unitary subgroups is relevant in our MIMO settings is a consequence of the fact that we are working with second order moment constraint, which implies that the mutual information has precisely the above described form (cf. (2)). Finally, examples of groups in $M_n(\mathbb{C})$ are $U(n)$, which is the largest group we will consider, and its subgroups $\Sigma(n)$ and $\Pi(n)$ (with the usual matrix multiplication), defined as:

1) $U(n)$: the unitary group of size n,
2) $\Pi(n)$: the group of permutation matrices of size n,
3) $\Sigma(n)$: the diagonal matrices group with 1 and -1 of size n.

We now gave a definition to quantify symmetries in the problem, through **the group of invariance** of H and D, or equivalently of ψ, the question is then: how do we use this invariance in order to get knowledge on the optimal input? In the next section we will see that this is done through the commutant.

B. Exploiting symmetries

Definition 3: The commutant of G is defined by the algebra $\mathrm{Comm}(G) = \{A \in M_n(\mathbb{C}) | Ag = gA, \forall g \in G\} = \{A \in M_n(\mathbb{C}) | A = gAg^{-1}, \forall g \in G\}$.

We start with a trivial observation linking the commutant and G-invariant functions.

Lemma 1: Let $G \subset M_n(\mathbb{C})$ be a group and $D \subset M_n(\mathbb{C})$. Let $\Psi : D \to \mathbb{R}$ a G-invariant function having a unique maximizer Q_{opt}. Then $Q_{\mathrm{opt}} \in \mathrm{Comm}(G) \cap D$.

Proof: We have $\psi(Q_{\mathrm{opt}}) = \psi(gQ_{\mathrm{opt}}g^{-1})$, $\forall g \in G$. We conclude by the uniqueness of the maximizer. ■

Note that the bigger is the group the smaller is the commutant, which is what we expect in order to exploit symmetries.

C. Symmetries in MIMO

We rewrite one last time the previous observations for our MIMO channel context.

Proposition 2: Let a MIMO channel as defined in the introduction and let G be a subgroup of $U(t)$. If

- the constraint set D_t is invariant in G-conjugation,
- the fading matrix distribution μ_H is G-invariant,

then

$$Q_{\mathrm{opt}} \in \mathrm{Comm(G)} \cap \mathrm{D_t}.$$

Also note that if G_1, G_2 are two groups in $U(t)$ and if D_t is invariant in G_1-conjugation whereas μ_H is G_2-invariant, then

$$Q_{\mathrm{opt}} \in \mathrm{Comm}(G_1 \cap G_2) \cap \mathrm{D}_t.$$

We will now see some specific applications of previous proposition. The cases that we will consider are dealing with the following commutants:

$$\mathrm{Comm}(\Sigma(n)) \text{ is the set of diagonal matrices in } M_n(\mathbb{C}),$$
$$\mathrm{Comm}(\Pi(n)) = \{\alpha I_n + \beta J_n | \alpha, \beta \in \mathbb{C}\}, \text{ where } J_n = 1^{n\times n},$$
$$\mathrm{Comm}(U(n)) = \{\alpha I_n | \alpha \in \mathbb{C}\}.$$

Corollary 1: Total power constraint
For a given $P \in \mathbb{R}_+$, we consider $Q \in \mathcal{D}_t = \{Q \in H_+(t) | \mathrm{tr}(Q) \leq P\}$. If μ_H is invariant in G-conjugation for a subgroup G of $U(t)$, then $Q_c \in \mathrm{Comm}(G) \cap \mathcal{D}_t$.

Simply observe that $\mathcal{D}_t$ is invariant in $U(t)$-conjugation. Two interesting cases of subgroups of $U(t)$ are $\Pi(t)$ and $\Sigma(t)$. From what we saw in the examples of the commutant, if we consider a distribution μ_H invariant under $\Sigma(t)$, then Q_c is diagonal and if it is invariant under $\Pi(t)$, then Q_c will have the same value for all components inside the diagonal ($\frac{P}{t}$ if one works in $\mathcal{D}_t$) and also the same value for all elements outside the diagonal, as long as it stays a positive definite matrix. Examples of $\Sigma(t)$-invariant random matrices are matrices with independent symmetric entries (symmetric means that $H_{ij} \overset{(d)}{\simeq} -H_{ij}$) and examples of $\Pi(t)$-invariant ones are matrices with i.i.d. entries or jointly gaussian entries having a covariance matrix of the form $\alpha I_{rt} + \beta J_{rt}$.

Corollary 2: Still considering $Q \in \mathcal{D}_t$, if H is $\Pi(t)\Sigma(t)$-invariant, which is for example the case when H_{ij} are i.i.d. and $H_{ij} \overset{(d)}{\simeq} -H_{ij}$, $\forall 1 \leq i \leq r, 1 \leq j \leq t$, then $Q_c = \frac{P}{t} I_t$.

Note that we did not assume that the entries of H are gaussian (which would be a particular case of this) in order to get $\frac{P}{t} I_t$ as a maximizer. Also note that the group $\Pi(t)$ could be replaced by $C(t)$, the group of cycling permutations, and we would get the same conclusion. Generally, this will be true as long as we have a group of invariance G such that $\mathrm{Comm}(G)$ is reduced to the multiple's of the identity.

Corollary 3: Local power constraint
If X is constrained by $\mathbb{E}|X_i|^2 \leq P_i$ for given $P_i \in \mathbb{R}_+$, $\forall 1 \leq i \leq t$, and if H is $\Sigma(t)$-invariant, then $Q_c = \mathrm{diag}(P_1, \ldots, P_t)$.

The constraint $\mathbb{E}|X_i|^2 \leq P_i$ implies that $Q \in \widetilde{\mathcal{D}}_t = \{Q \in H_+(t) | Q_{ii} \leq P_i, \forall 1 \leq i \leq t\}$, now we no longer have that $\widetilde{\mathcal{D}}_t$ is invariant in $U(t)$-conjugation, but we still have, for example, invariance in $\Sigma(t)$-conjugation.

Conclusion: As it has been illustrated in previous example, the problem of symmetries should be generally approached in the following way: first identify the invariance property of the domain D_t in which we are working (we saw examples of total and local power constrain (see corollaries 2 and 3), several intermediary cases are possible), then identify the invariance property of the fading matrix distribution μ_H, once we have these two groups of invariance, we know that we can restrict our search of Q_c to matrices commuting with these groups and staying in D_t. Which means that **the commutant is summarizing the information given by the symmetries in the problem**.

IV. Asymmetries in MIMO

Let us consider now the following specific channel.

Definition 4: The Kronecker channel
We consider the constraint set $\mathcal{D}_t = \{Q \in H_+(t) | \mathrm{tr}(Q) \leq P\}$, and $H = AWB$, where

- $A \in M_r(\mathbb{C})$ non-zero,
- $B \in M_t(\mathbb{C})$ non-zero,
- W is a $r \times t$ random matrix being $U(t)$-invariant on the right.

Question: do we still have some symmetries for such a channel? Any matrix multiplication on the right of H will be first acting on the fixed matrix B, therefore our invariance properties seem to be compromised. But before facing the asymmetric structure of this channel, let us bring back some symmetries in the problem.

A. Bringing back the symmetries

As we just saw, in some situations, a symmetric structure is not clearly existing. But with appropriate transformation, one can bring some symmetries back into the problem. We now give an example of how to carry out such a procedure for the Kronecker channel. One can also tackle the Ricean channel defined by $H = A + W$, where A is a deterministic matrix, in a similar way. For both cases, the following simple observation is used.

Lemma 2: Let $\Psi : D \to \mathbb{R}$ with a unique maximizer Q_{opt} and such that D is invariant in G-conjugation. Then, for any $M \in G$, we have

$$M^* Q_{\mathrm{opt}} M = \arg\max_{Q \in D} \Psi(M Q M^*).$$

Because we are now dealing with the Kronecker channel, the mutual information function ψ is given by

$$Q \in \mathcal{D} \mapsto \psi(Q) = \mathbb{E} \log\det(I + K^{-1} AWBQB^*W^*A^*).$$

We now denote the SVD of B by $B = U_B \,\mathrm{diag}(b) V_B^*$, where $U_B, V_B \in U(t)$ and $b \in \mathbb{R}_+^t$. Using our previous lemma, we can choose $M = V_B$, in order to get that

$$V_B^* Q_{\mathrm{opt}} V_B = \arg\max_{Q \in D} \psi(V_B Q V_B^*) \qquad (3)$$

The advantage of getting to the last maximization problem is that

$$Q \rightsquigarrow \psi(V_B Q V_B^*)$$

is a $\Sigma(t)$-invariant function and thus we can restrict our maximization to matrices being diagonal (with trace smaller than P). In other words, we showed the following observations:

Remark 1: the eigenvectors of Q_{opt} are the right-eigenvectors of B and its eigenvalues $q_{\mathrm{opt}} = (q_1^{\mathrm{opt}}, \ldots, q_t^{\mathrm{opt}})$ are given by

$$q_{\mathrm{opt}} = \arg \max_{q \in \mathbb{R}_+ \text{ s.t. } \sum_{i=1}^t q_i \leq P} \mathbb{E} \log\det(I + K^{-1} AW \,\mathrm{diag}(q_1 b_1^2, \ldots, q_t b_t^2)) W^* A^*).$$

B. Asymmetries for the Kronecker channel

Suppose that the matrix B in the Kronecker channel is diagonal, i.e. $B = \mathrm{diag}(b)$, with $b \in \mathbb{R}_+^t$. Then we know that the optimal covariance matrix is diagonal but we do not know what are the value of the diagonal elements. Now assume that $b_1 \leq \ldots \leq b_n$, can we then expect that the optimal covariance matrix should preserve this ordering in some sense?

We will now analyze these kinds of questions. We will present two propositions that will help describe the optimal input for the Kronecker channel. We know that if the random matrix H were replaced by the deterministic matrix B, the optimal input covariance has eigenvalues q_i^{opt} given via **"water-filing"** on the singular values of B (cf. [1]). Two particular properties of the "water-filing" solution are the following.

1) **Monotonicity:** if $b_i \geq b_j$ then $q_i^{\mathrm{opt}} \geq q_j^{\mathrm{opt}}$ (with equality if $b_i = b_j$)
2) **On/Off threshold:** if b_{i+1} is sufficiently bigger than b_i, then we might end up by sharing the whole power P on the $t-i$ biggest components of b.

We will see in the next two propositions that these two properties are preserved in the Kronecker model.

We start with the monotonicity result.

Proposition 3: We have

$$Q_{\mathrm{opt}} = V_B \,\mathrm{diag}(q_{\mathrm{opt}}) V_B^*$$

where q_{opt} satisfies

$$q_i^{\mathrm{opt}} \geq 0, \;\; \textstyle\sum_{i=1}^t q_i^{\mathrm{opt}} = P,$$
$$q_i^{\mathrm{opt}} \geq q_j^{\mathrm{opt}} \text{ if } b_i > b_j, \text{ and } q_i^{\mathrm{opt}} = q_j^{\mathrm{opt}} \text{ if } b_i = b_j.$$

Note: If $B = I_t$ and μ_W is G-invariant on the right with $G \leq U(t)$, then $Q_{\mathrm{opt}} \in \mathrm{Comm}(G) \cap \mathcal{D}_t$.

Remark: This proposition says that the eigenvectors of Q_{opt} are the right-eigenvectors of B (which has been shown in previous section) and that its eigenvalues are monotonically distributed with respect to the singular values of B.

In order to prove this result, we need a preliminary lemma. Let $\lambda_1(M) \leq \ldots \leq \lambda_n(M)$ denote the ordered eigenvalues of any matrix $M \in H(n)$ — we use $H(n)$ to denote the set of hermitian matrices of size n.

Lemma 3: Let $n \geq 1$, $P \in H_+(n)$ and $H \in H(n)$. We then have,

$$\lambda_k(H+P) > \lambda_k(H), \;\; \forall k = 1, \ldots, n.$$

Proof of proposition 3. The initial expression of the mutual information function for this channel is

$$\psi(Q) = \mathbb{E} \log\det(I + K^{-1} AWBQB^*W^*A^*).$$

First note that A affects the function ψ in the same way as K^{-1}, in other words, we could consider one of this two matrix to be the identity, for example, assume i.i.d. components for the noise and set $\tilde{A} = K^{-\frac{1}{2}} A$. If $B = I_r$ any invariance properties on the right for μ_W will be preserved for AW, thus the note after the proposition is a direct consequence of proposition 2.

The first part of the proposition is proved in the previous section, let us now look at the eigenvalues. We have

$$\begin{aligned} q_{\mathrm{opt}} &= \arg \max_{q \in \mathbb{R}_+ \text{ s.t. } \sum_{i=1}^t q_i \leq P} \mathbb{E} \log\det(I \\ &\quad + AW \,\mathrm{diag}(q_1 b_1^2, \ldots, q_t b_t^2)) W^* A^*). \end{aligned}$$

Thus we will consider from now on

$$\psi : q \rightsquigarrow \mathbb{E} \log\det(I + AW \,\mathrm{diag}(q_1 b_1^2, \ldots, q_t b_t^2)) W^* A^*).$$

Now observe that if $b_i = b_j$ then ψ is $\Pi(t)_{ij}$-invariant, where $\Pi(t)_{ij}$ is the subgroup of permutations keeping the diagonal elements different than i and j invariant (transposition), thus we get from proposition 2 that $q_i^{\mathrm{opt}} = q_j^{\mathrm{opt}}$.

Now, let $P' = P - \sum_{i=3}^t q_i^{\mathrm{opt}}$, such that $q_1^{\mathrm{opt}} + q_2^{\mathrm{opt}} = P'$. We will show that if $b_1 > b_2$, then for any $0 \leq P' \leq P$,

$$\partial_{q_1} \psi(q) \big|_{(\frac{P'}{2}, \frac{P'}{2}, q_3^{\mathrm{opt}}, \ldots, q_t^{\mathrm{opt}})} > \partial_{q_2} \psi(q) \big|_{(\frac{P'}{2}, \frac{P'}{2}, q_3^{\mathrm{opt}}, \ldots, q_t^{\mathrm{opt}})},$$

which, by the concavity of ψ, implies that

$$q_1^{\mathrm{opt}} > q_2^{\mathrm{opt}}.$$

By symmetry of the problem, this clearly implies the result for any components i and j (other than 1 and 2).

We have

$$\psi(q) = \mathbb{E} \log\det(I + \sum_{i=1}^t q_i b_i^2 A w_i (A w_i)^*)$$

where w_i is the i-th column of W. For an invertible matrix M, we have the formula $\partial_{m_{ij}} \log\det(M) = (M^{-1})_{ji}$, therefore we have

$$\partial_{q_j} \psi(q) = \\ b_j^2 \mathbb{E} \,\mathrm{tr} \Big(I + \textstyle\sum_{i=1}^t q_i b_i^2 A w_i (A w_i)^* \Big)^{-1} A w_j (A w_j)^*.$$

Let us denote $X_i = A w_i (A w_i)^*$, which are hermitian positive semidefinite matrices, as well as $\Big(I + \sum_{i=1}^t q_i b_i^2 X_i \Big)$ which is in addition positive definite and invertible. We define $Z = \sum_{i=3}^t q_i b_i^2 X_i$ and $Z_{\mathrm{opt}} = \sum_{i=3}^t q_i^{\mathrm{opt}} b_i^2 X_i$, we then rewrite

$$\begin{aligned} \partial_{q_1} \psi(q) &= b_1^2 \mathbb{E} \,\mathrm{tr} \big(I + q_1 b_1^2 X_1 + q_2 b_2^2 X_2 + Z \big)^{-1} X_1 \quad (4) \\ \partial_{q_2} \psi(q) &= b_2^2 \mathbb{E} \,\mathrm{tr} \big(I + q_1 b_1^2 X_1 + q_2 b_2^2 X_2 + Z \big)^{-1} X_2 \\ &= b_2^2 \mathbb{E} \,\mathrm{tr} \big(I + q_1 b_1^2 X_2 + q_2 b_2^2 X_1 + Z \big)^{-1} X_1 \quad (5) \end{aligned}$$

where in the last line we interchanged the random matrices X_1 and X_2, as W is $\Pi(t)$-invariant. To conclude the proof, we have to show that if $b_1 > b_2$

$$\begin{aligned} & b_1^2 \mathbb{E} \,\mathrm{tr} \Big(I + \tfrac{P'}{2} b_1^2 X_1 + \tfrac{P'}{2} b_2^2 X_2 + Z_{\mathrm{opt}} \Big)^{-1} X_1 \\ & > b_2^2 \mathbb{E} \,\mathrm{tr} \Big(I + \tfrac{P'}{2} b_1^2 X_2 + \tfrac{P'}{2} b_2^2 X_1 + Z_{\mathrm{opt}} \Big)^{-1} X_1, \end{aligned}$$

for any $0 \leq P' \leq 1$. This is clearly verified in the scalar case ($r = 1$). In the matrix case, more steps (using the previous lemma) are required to show that the result hold.

We now present an On/Off threshold result.

Proposition 4: Let $b_1 \leq b_2 \leq \ldots \leq b_t$. We assume that $r = 1$, $w_{1j} \overset{i.i.d.}{\sim} \mathcal{N}_{\mathbb{C}}(1), \forall 1 \leq j \leq t$. Then, for all $j = 1, \ldots, t$, there exists $\bar{b}(b_j) \geq 0$ such that

$$\text{if } b_{j+1} > \bar{b}(b_j) \text{ then } q_i^{\text{opt}} = 0, \ \ \forall i = 1, \ldots, j.$$

Comments: We will see that one can take $\bar{b}(b_j) = \sqrt{\frac{\bar{a}(Pb_j^2)}{P}}$, where $\bar{a}$ is given by the reciprocal of the function $\frac{1}{F} - 1$, with $F(a) = \mathbb{E}\frac{1}{1+aX}$, which is also known as the Ei or exponential function. The previous result says the following, if there is a value b_{j+1} such that Pb_{j+1}^2 is bigger than $\bar{a}(Pb_j^2)$, we then know that the optimal q_i^{opt} are zero for $i = 1, \ldots, j$. In other words, if some of these "gains" (b_i's) are two small compared to some others, we switch off the corresponding antennas.

Proof: In this setting we have

$$\psi(q) = \mathbb{E} \log(1 + P \sum_{i=1}^{t} q_i d_i X_i),$$

with $b_i^2 = d_i$, $X_i \overset{i.i.d.}{\sim} \mathcal{E}(1), \forall 1 \leq i \leq t$ and $q \in \Theta(t) = \{x \in \mathbb{R}_+^t | \sum_{i=1}^t x_i = 1\}$. Let $Z_j = 1 + P \sum_{i \neq j, j+1} q_i X_i$. We then have

$$\partial_{q_j} \psi(q) = \mathbb{E} \frac{P d_j X_j}{Z_j + P d_j X_j + P d_{j+1} X_{j+1}}.$$

Let $0 < T \leq 1$ and $p^{(j)}$ be a vector with $p_{j+1}^{(j)} = T$, $p_j^{(j)} = 0$ and thus $\sum_{i \neq j, j+1} p_i^{(j)} = 1 - T$. From the concavity of ψ, if

$$\partial_{q_j} \psi(q)|_{q=p^{(j)}} < \partial_{q_{j+1}} \psi(q)|_{q=p^{(j)}}, \ \ \forall 0 < T \leq 1, \qquad (6)$$

then $q_i^{\text{opt}} = 0$, $\forall i = 1, \ldots, j$. Now, (6) becomes

$$\mathbb{E} \frac{Z + TPd_j X_j}{Z + TPd_{j+1} X_{j+1}} < 1, \ \ \forall 0 < T \leq 1,$$

so if for all $z \geq 1$ and $0 < T \leq 1$ we have

$$\mathbb{E} \frac{z + TPd_j X_j}{z + TPd_{j+1} X_{j+1}} = \mathbb{E} \frac{z/T + Pd_j X_j}{z/T + Pd_{j+1} X_{j+1}} < 1,$$

we are done. Last inequality is equivalent to

$$\mathbb{E} \frac{1}{z + Pd_{j+1} X_{j+1}} < \frac{1}{z + Pd_j}, \ \ \forall z \geq 1.$$

Let $F(a) = \mathbb{E}\frac{1}{1+aX}$, $a_{j+1} = Pd_{j+1}$ and $a_j = Pd_j$, we now wonder when

$$F(a_{j+1}/z) < \frac{1}{1 + a_j/z}, \ \ \forall z \geq 1.$$

For a given $\beta \in \mathbb{R}_+$, let $\alpha(\beta)$ be the smallest number satisfying $F(\alpha(\beta)) \leq \frac{1}{1+\beta}$. Then if for any possible a_j, $\bar{a}(a_j) = \sup_{z \geq 1} z\alpha(a_j/z) < +\infty$, we deduce that for $a_j > \bar{a}(a_j)$, we satisfy $F(a_{j+1}/z) < \frac{1}{1+a_j/z}$, $\forall z \geq 1$. One can check that α is convex and that it is a continuous increasing function with $\alpha(0) = 0$. Therefore

$$\alpha(a_j/z) = \alpha(a_j/z + 0(1 - 1/z)) \leq \alpha(a_j)/z + 0$$

and thus

$$z\alpha(a_j/z) \leq \alpha(a_j), \ \ \forall z \geq 1$$

which implies that $\bar{a}(a_j) = \sup_{z \geq 1} z\alpha(a_j/z) = \alpha(a_j)$. And we conclude by setting $\bar{b}(b_j) = \sqrt{\frac{\bar{a}(Pb_j^2)}{P}}$.

■

The function $\bar{b}^2(\cdot)$ is continuous convex and increasing with $\bar{b}(0) = 0$, a derivative of 1 at 0 and of 0 at infinity.

In the following graphic, the inverse of the function $\bar{b}^2$ is plotted. Let us now look at some numerical example, we

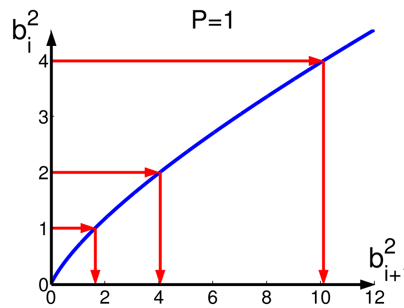

Fig. 1. Inverse of $\bar{b}^2$, $P = 1$.

assume that we have $r = 6$ receiving antennas and $P = 1$. If for example B is such that its singular values have squared given by $(b_1^2, \ldots, b_6^2) = (1, 1, 2, 3, 4, 11)$, it can be seen from figure 1 that b_6 exceeds the On/Off threshold of b_5 and thus the optimal power allocation is $(q_1^{\max}, \ldots, q_6^{\max}) = (0, \ldots, 0, 1)$, i.e. in this case we solved the problem. If we had $(b_1^2, \ldots, b_6^2) = (1, 1, 2, 5, 6, 8)$, then the previous situation does not hold anymore, but b_4 exceeds the On/Off threshold of b_3 and we are reduced to a half-dimensional optimization problem for the values of b_4, b_5 and b_6.

Conclusion: The optimal power allocation is not achieved in the same way than for the case of a deterministic fading matrix B, but still we preserve the same properties, indeed the monotonicity and the on/off threshold (with more specific hypotheses).

References

[1] E. Telatar, *Capacity of Multi-antenna Gaussian Channels*, European Trans. on Telecommunications, Vol. 10, No 6, 1999, pp. 585-595.

[2] V. Marchenko and L. Pastur, *Distribution of eigenvalues in certain sets of random matrices*, Math. USSR-Sbornik 1 (1967), no. 4, pp. 457-483.