

UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

ANTONIO ALISSON PESSOA GUIMARÃES

SOBRE LIMITANTES DE CAPACIDADE DE UM CANAL MIMO COM
DESVANECIMENTO DE RICE:
UMA ABORDAGEM VIA TEORIA DE MAJORIZAÇÃO

FORTALEZA – CEARÁ
JULHO – 2010

ANTONIO ALISSON PESSOA GUIMARÃES

SOBRE LIMITANTES DE CAPACIDADE DE UM CANAL MIMO COM
DESVANECIMENTO DE RICE:
UMA ABORDAGEM VIA TEORIA DE MAJORIZAÇÃO

Dissertação submetida à coordenação do Programa de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Teleinformática.

Área de concentração: Sinais e Sistemas.

Orientador: Prof. Dr. Charles Casimiro Cavalcante

FORTALEZA – CEARÁ

JULHO – 2010

Ficha catalográfica elaborada pela Biblioteca do curso de Matemática da UFC.

G963s Guimarães, Antonio Alisson Pessoa
Sobre limitantes de capacidade de um canal MIMO com
desvanecimento de Rice: uma abordagem via teoria de majorização/
Antonio Alisson Pessoa Guimarães – Fortaleza – 2010.
111 f.; il.; enc.

Dissertação (Mestrado) - Universidade Federal do Ceará,
Departamento de Engenharia de Teleinformática, Fortaleza, 2010.
Orientador: Prof. Dr. Charles Casimiro Cavalcante.
Área de concentração: Sinais e Sistemas.

1. Teleinformática. 2. Sinais e Sistemas. 3. Sistemas MIMO.
4. Processamento de sinais. I. Cavalcante, Charles Casimiro
(orient.). II. Universidade Federal do Ceará – Programa de Pós-
Graduação em Engenharia de Teleinformática. III. Título.

CDD 621.38

ANTÔNIO ALISSON PESSOA GUIMARÃES

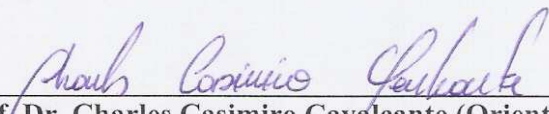
SOBRE LIMITANTES DE CAPACIDADE DE UM CANAL MIMO COM DESVANECIMENTO DE RICE: UMA ABORDAGEM VIA TEORIA DE MAJORIZAÇÃO

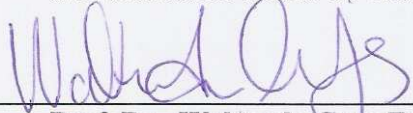
Dissertação submetida à Coordenação do Programa de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Teleinformática.

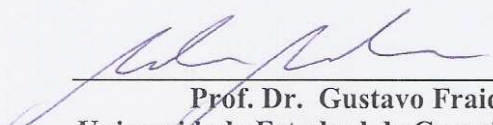
Área de concentração SINAIS e SISTEMAS.

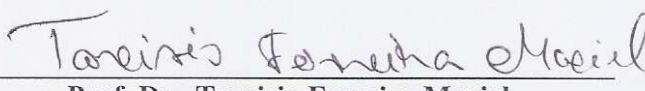
Aprovada em 09/07/2010.

BANCA EXAMINADORA


Prof. Dr. Charles Casimiro Cavalcante (Orientador)
Universidade Federal do Ceará -UFC


Prof. Dr. Walter da Cruz Freitas Jr.
Universidade Federal do Ceará -UFC


Prof. Dr. Gustavo Fraidenraich
Universidade Estadual de Campinas - UNICAMP


Prof. Dr. Tarcisio Ferreira Maciel
Universidade Federal do Ceará -UFC

Dedico este trabalho com muito carinho, aos meus pais, Eduardo e Fátima, que são meus maiores incentivadores, e às minhas irmãs Artemis e Artemízia, as quais sou imensamente grato por serem minhas fiéis companheiras em todos os momentos.

AGRADECIMENTOS

A Deus pelo dom da vida, pelas muitas graças concedidas, dentre elas a concretização deste trabalho e a benção especial de ver meu pai totalmente recuperado de uma cirurgia feita no coração, a qual foi realizada durante este curso de mestrado.

Aos meus pais **Eduardo** e **Fátima**, e às minhas irmãs, **Artemis** e **Artemízia**, por todo amor e incentivos incansáveis, pois sempre me deram um ombro amigo para dividir os bons e difíceis momentos da vida. Esta vitória também é de vocês!

Agradeço aos meus queridos e já falecidos avós **João** e **Odília** por todo o carinho, conselhos e exemplos de vida.

À minha noiva **Dayane** pelo amor, paciência e compreensão, principalmente nos momentos em que eu estava distante (física e mentalmente). Pela expectativa que dividimos em cada etapa e o entusiasmo com que comemoramos cada conquista. Esta caminhada não teria sido a mesma sem você!

Meus sinceros agradecimentos ao meu orientador e amigo prof. **Charles Casimiro Cavalcante** que, com paciência, humildade e dedicação foi um incentivador nas buscas, coerente nas palavras, incansável nas inúmeras releituras e reescritas que esse trabalho exigiu, até ser finalizado. Não posso deixar de agradecer a receptividade com que tem comigo desde nosso primeiro encontro, a confiança no meu trabalho e incentivos para superar os desafios da vida e da ciência. Além disso, a minha eterna gratidão por ter sido solidário e, sobretudo, compreensivo nos difíceis momentos ocasionados por problemas de

saúde na minha família, deixando-me à disposição para dar o apoio necessário ao meus familiares.

Também gostaria de agradecer ao **GTEL** (Grupo de Pesquisa em Telecomunicações Sem Fio), o qual me cedeu toda estrutura e apoio para o desenvolvimento desta obra, bem como, todos que fazem parte deste grupo, como funcionários, pesquisadores e professores, em especial, ao prof. **Tarcisio Maciel** por ter me propiciado importantes reflexões e sugestões para o enriquecimento deste trabalho.

Não poderia deixar de esquecer da minha amiga do GTEL, **Lígia Sousa**, diante de sua simpatia marcante, sempre esteve disponível a tirar minhas dúvidas nos aspectos teóricos, como também o fundamental suporte computacional. Além de tudo, teve uma contribuição significativa no desenvolvimento desta obra, principalmente, com indicações de artigos nesta mesma linha de pesquisa.

Aos **professores** do curso de mestrado de Engenharia de Teleinformática os quais contribuíram para o meu crescimento profissional, em especial, para o prof. **José Tarcísio**, um dos responsáveis pela minha transição de aluno do curso de Matemática para aluno de Engenharia, e que também solidarizou-se com meu momento de fragilidade.

Faço um especial agradecimento à **banca examinadora** pelas importantes contribuições para os encaminhamentos finais do trabalho aqui proposto.

À **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior** (CAPES) e à **Fundação Cearense de Amparo à Pesquisa** (FUNCAP) pelas bolsas de estudos concedidas.

Enfim, a todos aqueles que, diretamente ou indiretamente, contribuíram para que este sonho se concretizasse.

Muito obrigado!!!

Alisson Guimarães.

*“Ao sonharmos com o impossível,
começamos a torná-lo possível – o
impossível de hoje é a ciência de
amanhã.”*

Marcelo Gleiser

RESUMO

Uma das formas de se medir o desempenho de sistema comunicação de um canal MIMO (*Multiple-Input Multiple-Output*, do inglês) é através da capacidade ergódica ou capacidade de Shannon, a qual representa a máxima eficiência espectral. No entanto, dependendo da distribuição estatística das entradas do canal, pode ser inviável o estudo computacional, devido principalmente a complexidade de obtenção de tal resultado. Portanto, em geral, para contornar esta situação trabalha-se com aproximações à capacidade através de limitantes, inferior e superior, como também em casos de alto e baixo regimes da relação sinal-ruído (SNR, do inglês). Por sua vez, a obtenção de uma expressão fechada para tais limitantes ainda é um grande desafio. Diante desta problemática, esta dissertação tem por objetivo a obtenção de limitantes, bem como aproximações à capacidade em alto e baixo regimes de SNRs, de um canal MIMO com desvanecimento de Rice cujos elementos são descorrelacionados e a matriz da componente linha de visada tem posto arbitrário, o qual não excede a quantidade mínima entre o número de antenas transmissoras e receptoras. Além da hipótese de que o transmissor não tem qualquer conhecimento dos estados do canal (CSI, do inglês). Para tal, é proposto o uso de uma ferramenta matemática ainda pouco explorada na literatura, com relação a este assunto, chamada de teoria de majorização. Dependendo do tipo de problema, o método consiste em definir uma função Schur-côncava específica com relação à capacidade ergódica, a qual aplicada em vetores previamente majorizados permite obter desigualdades numéricas e, consequentemente, limitantes e aproximações à capacidade. Simulações computacionais são realizadas, e no caso do limitante superior, para evidenciar a potencialidade do método, são feitas algumas comparações com resultados já conhecidos da literatura.

Palavras-chave: Capacidade ergódica, canal com desvanecimento de Rice, sistemas MIMO, teoria de majorização.

ABSTRACT

One way of measuring the performance of communication system of a MIMO channel (Multiple-Input Multiple-Output) is through the capacity or ergodic Shannon capacity, which represents the highest spectral efficiency. However, depending on the statistical distribution of the inputs, its computational study may not be feasible mainly due to the complexity of obtaining such an outcome. Therefore, in general, to cope with drawback we use approaches to capacity by limiting, above and below, in cases of high and low regimes of signal to noise ratio (SNR). In turn, obtaining a closed expression for these bounds is still a great challenge. In face this problem, this thesis aims at finding bounds and approximations to the capacity at high and low SNRs regimes in a MIMO fading channel with Ricean fading, whose elements are uncorrelated and the matrix component of the line-of-sight has an arbitrary rank, which does not exceed the minimum between the number of transmitting and receiving antennas. In addition, the transmitter has no knowledge of the channel states. To this end, we propose the use of a mathematical tool not much explored in the literature with regard to this matter, called majorization theory. Depending on the type of problem, the method consist in defining a specific Schur-concave function with respect to the ergodic capacity, which previously applied majorized vectors allowing to obtain numerical inequalities and thereby limits and approximations to capacity. Computer simulations are carried out and, in the case of upper capacity bounds some comparisons with known results from literature are made to highlight the potential of the method.

Keywords: Ergodic capacity, Ricean channel, MIMO systems, majorization theory.

LISTA DE FIGURAS

1	Representação de um canal MIMO $n_R \times n_T$	26
2	Representação da decomposição de um canal MIMO com CSI no transmissor.	30
3	Princípio do algoritmo <i>water-filling</i>	32
4	Propagação multipercurso.	38
5	Interpretação geométrica de majorização de vetores.	46
6	Interpretação geométrica de um conjunto contido em \mathcal{D}_2	47
7	Comparação entre limitantes superiores para $\Omega = 1$ e $0 \leq \text{SNR} \leq 25$ dB, com incremento de 5 dB.	76
8	Comparação entre limitantes superiores para os cenários 1×1 , 2×2 , 4×4 e 8×8	76
9	Comparação entre limitante superior C_{up} e capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$	77
10	Comparação entre limitante superior C_{up} e capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$	78
11	Comparação entre limitante superior C_{up} e as capacidades ergódicas $\bar{C}_{3 \times 3}$ e $\bar{C}_{3 \times 5}$ para os cenários 3×3 e 3×5 , respectivamente. O fator de Rice utilizado é $K = 3$	78
12	Comparação entre erros relativos de aproximação do limitante C_{up} à capacidade ergódica \bar{C} num canal de ordem 3×5	79

13	Comparação entre os limitantes superiores C_{up} e $C_{\text{up}}^{\text{int}}$ e a capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$	80
14	Comparação entre os limitantes superiores C_{up} e $C_{\text{up}}^{\text{int}}$ e a capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$	81
15	Comparação entre os limitantes superiores C_{up} e $C_{\text{up}}^{\text{int}}$ e a capacidade ergódica \bar{C} para o cenário 3×5 , com fator de Rice $K = 3$	81
16	Aproximações à capacidade ergódica nos casos 2×2 , 4×4 e 8×8 , com fator de Rice $K = 3$, para alto regime SNR.	82
17	Comparação entre o limitante inferior C_{lo} e a capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$	85
18	Comparação entre o limitante inferior C_{lo} e a capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$	85
19	Comparação entre os limitantes inferiores C_{lo} e $C_{\text{lo}}^{\text{int}}$ e a capacidade ergódica \bar{C} para o cenário 8×8 , com fator de Rice $K = 3$	86
20	Aproximações à capacidade ergódica nos casos 2×2 , 4×4 e 8×8 , com fator de Rice $K = 3$, para baixo regime SNR.	88
21	Comparação com limitantes superiores de Jin <i>et. al</i> (2007) para os cenários 1×1 e 2×2 , com $K = 0$	92
22	Comparação com limitantes superiores de Jin <i>et. al</i> (2007) para os cenários 4×4 e 8×8 , com $K = 0$	93
23	Comparação com limitantes superiores de Jin <i>et. al</i> (2007) para os cenários 1×1 e 2×2 , com $K = 3$	93
24	Comparação com limitantes superiores de Jin <i>et. al</i> (2007) para os cenários 4×4 e 8×8 , com $K = 3$	94

LISTA DE ACRÔNIMOS

BER	Taxa de erro de bit
CSI	Estado de informação do canal
pdf	Função densidade de probabilidade
i.i.d.	Independentes e identicamente distribuídas
MIMO	<i>Multiple-Input Multiple-Output</i>
MSE	Erro quadrático médio
SISO	<i>Single-Input Single-Output</i>
SNR	Relação sinal-ruído
SVD	Decomposição em valores singulares
v.a.	Variável aleatória

SÍMBOLOS E NOTAÇÕES

Nesta seção são apresentados os símbolos e notações utilizadas neste trabalho. De uma forma geral, os escalares serão representados por letras minúsculas em itálico, os vetores por letras romanas minúsculas em negrito e as matrizes por letras romanas maiúsculas, também em negrito. Outras convenções estão listadas abaixo:

Conjuntos e Espaços Vetoriais

\mathbb{N}	Conjunto dos números naturais
\mathbb{R}	Conjunto dos números reais
\mathbb{R}_+	Conjunto dos números reais não-negativos
$\mathbb{R}^{m \times n}$	Conjunto das matrizes de ordem $m \times n$ cujos elementos são números reais
\mathbb{R}^n	Espaço vetorial cujos elementos (x_1, x_2, \dots, x_n) são listas ordenadas de números reais. Em termos matriciais, este espaço é identificado por $\mathbb{R}^{n \times 1}$
\mathbb{R}_+^n	$\mathbb{R}_+^n = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n ; x_i \geq 0, i = 1, 2, \dots, n\}$
\mathcal{D}	$\mathcal{D} = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n ; x_1 \geq x_2 \geq \dots \geq x_n\}$
\mathcal{D}_+	$\mathcal{D}_+ = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}_+^n ; x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$
\mathbb{C}	Conjunto dos números complexos
$\mathbb{C}^{m \times n}$	Conjunto das matrizes de ordem $m \times n$ cujos elementos são números complexos
\mathbb{C}^n	Espaço vetorial cujos elementos (z_1, z_2, \dots, z_n) são listas ordenadas de números complexos. Em termos matriciais, este espaço é identificado por $\mathbb{C}^{n \times 1}$
\mathcal{N}	$\mathcal{N} = \{1, 2, \dots, n\} \subset \mathbb{N}$

\mathcal{P}	Conjunto das partes de \mathcal{N}
$P_i(\mathcal{N})$	Subconjunto de \mathcal{P} , cujos elementos são todas as combinações de i termos tomados n a n

Escalares Especiais

z^*	Conjugado complexo de z
$\pi(i)$	i -ésimo elemento da permutação do conjunto $\{1, 2, \dots, k\} \subset \mathbb{N}$
K	Fator de Rice
n_T	Número de antenas transmissoras
n_R	Número de antenas receptoras
P_T	Potência máxima de transmissão
Ω	Potência média dos elementos da matriz do canal
ρ	Relação sinal-ruído
C	Capacidade de um canal com múltiplas antenas
\bar{C}	Capacidade ergódica de canal
$\bar{C}_{n_R \times n_T}$	Capacidade ergódica de canal de ordem $n_R \times n_T$
C_{up}	Limitante superior de um canal MIMO com desvanecimento de Rice
$C_{\text{up}}^{\text{int}}$	Limitante superior de um canal MIMO com desvanecimento de Rice cujo valor está compreendido entre \bar{C} e C_{up} , isto é, $\bar{C} \leq C_{\text{up}}^{\text{int}} \leq C_{\text{up}}$
C_{lo} e $C_{\text{lo}(2)}$	Limitantes inferiores de um canal MIMO com desvanecimento de Rice
$n!$	Fatorial de n
$\binom{m}{n}$	Número binomial, $\binom{m}{n} = \frac{m!}{n!(m-n)!}$

Vetores e Matrizes

\mathbf{x}	Vetor \mathbf{x}
$[\mathbf{x}] = (x_{[1]}, x_{[2]}, \dots, x_{[n]})$	Vetor obtido de $\mathbf{x} = (x_1, x_2, \dots, x_n)$ por reordenação decrescente das coordenadas, isto é, $x_{[1]} \geq \dots \geq x_{[n]}$
$\mathbf{1}$	Vetor cujas coordenadas são todas iguais a 1, isto é, $\mathbf{1} = (1, 1, \dots, 1)$
$\pi(\mathbf{x})$	Vetor permutação das coordenadas de \mathbf{x}
$\ \mathbf{x}\ $	Norma euclidiana do vetor \mathbf{x}
\mathbf{X}	Matriz \mathbf{X}

$\mathbf{X} = [x_{ij}]$	Representação de uma matriz \mathbf{X} com relação as suas entradas
\mathbf{I}_n	Matriz identidade de ordem $n \times n$
\mathbf{X}^T	Matriz/vetor transposta de \mathbf{X}
\mathbf{X}^H	Matriz/vetor transposta conjugada de \mathbf{X}
\mathbf{X}^{-1}	Matriz inversa de \mathbf{X}
$\mathbf{X}^{1/2}$	Matriz raiz quadrada de \mathbf{X}
$\mathbf{X} \otimes \mathbf{Y}$	Produto de Kronecker
$\text{vec}(\mathbf{X})$	Vetor obtido por empilhamento das colunas da matriz \mathbf{X}
\mathbf{H}	Matriz de canal MIMO de ordem $n_T \times n_R$
\mathbf{H}_w	Matriz de canal do tipo Rayleigh
$\text{diag}(\mathbf{x})$	Matriz diagonal cujos elementos da diagonal principal são as coordenadas do vetor \mathbf{x}
$\mathbf{R}_{\mathbf{x}}$	Matriz de covariância de um vetor aleatório \mathbf{x}
$\mathbf{R}_{\mathbf{H}}$	Matriz de covariância do canal MIMO
\mathbf{R}_{Tx}	Matriz de covariância no transmissor
\mathbf{R}_{Rx}	Matriz de covariância no receptor
$\text{d}(\mathbf{X})$	Vetor cujas coordenadas são os elementos da diagonal principal da matriz quadrada \mathbf{X}
$\lambda(\mathbf{X})$	Vetor cujas coordenadas são os autovalores da matriz hermitiana \mathbf{X}
$\det(\mathbf{X})$	Determinante da matriz \mathbf{X}
$\text{posto}(X)$	Posto da matriz \mathbf{X}
$\text{tr}(\mathbf{X})$	Traço da matriz \mathbf{X}
$\mathbf{x} \leq \mathbf{y}$	$x_i \leq y_i$, para $i = 1, 2, \dots, n$. Sendo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ e $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$
$\mathbf{x} \prec \mathbf{y}$	Operação de majorização: o vetor \mathbf{x} é majorizado por \mathbf{y}

Estatística e Distribuições

X	Variável aleatória X
$\mathbb{E}\{X\}$ ou μ_X	Esperança da variável aleatória X
$\mathbb{E}_X\{f(X)\}$	Esperança da variável aleatória $Y = f(X)$ com respeito a X
$\text{var}(X)$	Variância da variável aleatória X
$\text{cov}(X, Y)$	Covariância entre as variáveis X e Y
$\mathbb{E}_X\{f(X)\}$	Esperança da variável aleatória $Y = f(X)$ com respeito a X

$p_R(\cdot)$	Função densidade de probabilidade da variável aleatória R
$X \sim \mathcal{N}(\mu, \sigma^2)$	Variável gaussiana de média μ e variância σ^2
$\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Vetor aleatório $\mathbf{x} \in \mathbb{C}^{n \times 1}$ gaussiano complexo circularmente simétrico de vetor-média $\boldsymbol{\mu} \in \mathbb{C}^{n \times 1}$ e matriz de covariância $\boldsymbol{\Sigma}$ de ordem $n \times n$
$\mathbf{X} \sim \mathcal{CN}(\boldsymbol{\Lambda}, \boldsymbol{\Phi} \otimes \boldsymbol{\Omega})$	Matriz aleatória $\mathbf{X} \in \mathbb{C}^{m \times n}$ gaussiana complexa circularmente simétrica de matriz-média $\boldsymbol{\Lambda} \in \mathbb{C}^{n \times 1}$ e matriz de covariância $\boldsymbol{\Phi} \otimes \boldsymbol{\Omega}$ de ordem $mn \times mn$

Funções e outros operadores

$\exp(\cdot)$	Função exponencial
$\log_b(\cdot)$	Função logarítmica de base b
$\max\{\cdot\}$	Função máximo
$(x)^+$	$\max\{x, 0\}$
$\min\{\cdot\}$	Função mínimo
$ \cdot $	Valor absoluto de um escalar
$H(\cdot)$	Entropia
$H(\cdot \cdot)$	Entropia condicional
$I(\cdot; \cdot)$	Informação mútua
$I_0(\cdot)$	Função de Bessel modificada de primeira espécie de ordem zero
$M_X(\cdot)$	Função geradora de momentos da variável aleatória X

Outros símbolos

\approx	Aproximadamente igual a ...
\triangleq	Definido por ...
\ll	Muito menor que ...
\square	Final de demonstração de teorema
\triangleleft	Final de exemplo

SUMÁRIO

1	Introdução	20
1.1	Organização da dissertação e contribuições	22
2	Introdução aos Sistemas MIMO	24
2.1	O sistema MIMO	24
2.1.1	Modelo de canal MIMO	25
2.2	Capacidade de um sistema MIMO	26
2.2.1	Transmissor sem informação do canal MIMO	28
2.2.2	Transmissor com informação do canal MIMO	30
2.2.3	Transmissor com informação do canal MIMO: caso da potência ótima	31
2.2.4	Capacidade Ergódica	32
2.3	Modelo analíticos de canais MIMO	33
2.3.1	Modelo i.i.d.	34
2.3.2	Modelo de Kronecker	35
2.4	Canal com desvanecimento de Rice e de Rayleigh	37
2.4.1	Características gerais de propagação	37
2.4.2	Canal com desvanecimento de Rayleigh	40

2.4.3	Canal com desvanecimento de Rice	41
2.5	Conclusões e síntese	43
3	Teoria de Majorização	44
3.1	Definições e resultados básicos	46
3.1.1	Propriedades básicas	48
3.1.2	Matriz Duplamente Estocástica	50
3.1.3	Schur-convexidade e Schur-concavidade	51
3.2	Majorização e Álgebra Linear	58
3.3	Majorização e Otimização	61
3.4	Majorização Estocástica	64
3.5	Conclusões e Síntese	66
4	Aproximações e Limitantes à Capacidade Ergódica	67
4.1	Considerações iniciais	68
4.1.1	Descrição do modelo	68
4.1.2	Formulação do problema	69
4.1.3	Transformação da matriz do canal	69
4.2	Limitante superior à capacidade ergódica	72
4.2.1	Aproximação à capacidade ergódica em alto regime de SNR	82
4.3	Limitante inferior à capacidade ergódica	83
4.3.1	Aproximação à capacidade ergódica em baixo regime SNR	86
4.4	Estudo de um caso assintótico	89
4.5	Comparações com dados da literatura: caso do limitante superior	90
4.6	Conclusões e Síntese	94
5	Conclusões e Perspectivas	95

Apêndice A Teoria de majorização: outras considerações	98
A.1 Um estudo sobre o conjunto \mathcal{D}_n	98
A.2 Demonstração da propriedade (p8) das relações básicas de majorização . .	100
A.3 Composição de funções: um caso específico de preservação da propriedade de Schur-convexidade	101
Apêndice B Complemento da Demonstração do Teorema 4.1	103
Apêndice C Complemento da Demonstração do Teorema 4.4	105
C.1 Parte 1	105
C.2 Parte 2	106
Referências Bibliográficas	108

CAPÍTULO 1

INTRODUÇÃO

Desde a primeira transmissão via rádio, as comunicações sem fio tiveram uma evolução notável. De um modo geral, as tecnologias sem fio convergem basicamente para os mesmos objetivos: transmitir o máximo de informação, num curto período de tempo e com melhor qualidade possível. Nesse sentido, tivemos grandes avanços na modulação, codificação e processamento de sinais para maximizar a eficiência espectral e mantendo a probabilidade de erro tão pequena quanto se queira. No entanto, esta eficiência espectral é limitada pela capacidade de Shannon (eficiência espectral máxima) (SHANNON, 1948, 1949).

Portanto, uma das formas para se medir o rendimento de um canal é através da capacidade, a qual representa o máximo de eficiência espectral. Desta forma, a utilização de múltiplas antenas no transmissor e no receptor, chamada comumente de MIMO (*multiple-input multiple-output*), ganhou rapidamente popularidade, devido principalmente aos trabalhos de Winters (1987), Telatar (1999) e Foschini e Gans (1998), os quais propuseram melhorias de desempenho na capacidade do canal de Rayleigh sem aumentar a potência de transmissão ou a largura de banda utilizada, situações estas, não proporcionadas pelos sistemas SISO (*Single-Input Single-Output*).

Foi Winters (1987), portanto, na década de 1990, quem propôs expressões matemáticas para a capacidade de sistema MIMO equivalentes às expressões desenvolvidas por Shannon (1948, 1949) para os canais SISO. Em seu trabalho, ele destacou a importância dos autovalores da matriz de potência do canal no cálculo da capacidade e mostrou como

estes se relacionam diretamente com as características físicas do canal em questão.

No entanto, mesmo com o resultado da fórmula da capacidade conhecido e dependendo do tipo de canal ou do cenário, em termos computacionais, por exemplo, fica por muitas vezes inviável a implementação de tal resultado, devido em geral a complexidade da referida fórmula. Desta forma, buscam-se limitantes inferior e/ou superior, os quais proporcionam aproximações à capacidade.

Por sua vez, não é tarefa simples a obtenção de expressões fechadas para tais limitantes. Além disso, se as expressões obtidas forem tão complexas quanto à capacidade em questão, em termos práticos, pode ser que não se apresente nenhum ganho ou vantagem em considerar uma ou outra situação.

Na literatura, no caso do canal com desvanecimento de Rice, que é o foco desta dissertação, existem importantes trabalhos em se tratando da determinação de limitantes para a capacidade. Por exemplo, McKay e Collings (2005) e Jin, Gao e You (2007) obtêm limitantes à capacidade no caso de um canal MIMO espacialmente correlacionado considerando as seguintes hipóteses: o posto da matriz da componente de linha de visada (LOS, do inglês) é igual a 1; o receptor tem perfeito conhecimento dos estados do canal (CSI, do inglês) e o transmissor não dispõe de CSI. Por sua vez, Jayaweera e Poor (2005), seguindo basicamente as mesmas condições, abordam a situação em que os elementos da matriz são descorrelacionados. Enquanto, Lebrun *et al.* (2006) tratam de casos assintóticos.

No entanto, grande parte destes trabalhos, dentre outros que tratam de expressões fechadas à limitantes de capacidade, envolvem geralmente em seus desenvolvimentos a matriz de potência do canal, a partir da distribuição não-central de Wishart (JAYAWEERA; POOR, 2005), a qual tem um tratamento matemático complexo, principalmente, no que se refere a obtenção da função densidade de probabilidade conjunta dos autovalores de tal matriz (FRAIDENRAICH.; LEVEQUE; CIOFFI, 2008).

Em Zhong, Wong e Jin (2009), apresentam um método alternativo, via teoria de majorização, o qual permite a obtenção de limitantes sem que seja necessário o uso direto da distribuição não-central de Wishart.

Portanto, seguindo a proposta de Zhong, Wong e Jin (2009), esta dissertação tem como objetivo a obtenção de limitantes inferior e superior de um canal de desvanecimento de Rice, admitindo-se que a matriz da componente de linha de visada tem posto não excedente a quantidade mínima entre o número de antenas transmissoras e receptoras.

Além disso, considera-se que apenas o receptor tem perfeito conhecimento dos estados do canal.

1.1 Organização da dissertação e contribuições

O presente trabalho está estruturado em capítulos, os quais encontram-se organizados da seguinte forma:

Capítulo 2: Introdução aos Sistemas MIMO

Neste capítulo são revisados alguns fundamentos relativos a sistemas MIMO dando ênfase ao estudo da capacidade do canal. Além disso, são feitos breves comentários a respeito das características gerais de propagação levando-se em consideração os canais com desvanecimento de Rayleigh e de Rice.

Capítulo 3: Teoria de Majorização

Apresentação de conceitos e resultados básicos relativos à teoria de majorização, tais como: relação de majorização entre vetores, matriz duplamente estocástica, a definição de função Schur-côncava e Schur-convexa, a condição de Schur a qual permite caracterizar se uma determinada função é Schur-côncava ou Schur-convexa, dentre outras condições. Além disso, é apresentado um conjunto de aplicações para ilustrar a potencialidade da ferramenta, sendo estas comumente encontradas em problemas de comunicação sem fio.

As principais contribuições¹ neste capítulo se referem a formulação de teoremas relacionando teoria de majorização com problemas de otimização. Este estudo faz parte de um trabalho aceito no *International Telecommunications Symposium 2010* (ITS 2010) com o título: *A (Very) Brief Survey on Optimization Methods for Wireless Communication Systems*. Uma outra contribuição se refere a obtenção de limitantes à capacidade ergódica de um determinado canal MIMO aplicando-se o estudo de majorização estocástica.

Capítulo 4: Aproximações e Limitantes à Capacidade Ergódica

Neste capítulo, via teoria de majorização, são apresentados limitantes, inferior e superior, à capacidade ergódica de um determinado canal de desvanecimento de Rice, bem

¹Os resultados do presente estudo, apresentados como teoremas, encontram-se destacados ao longo do texto.

como, aproximações à capacidade em alto e baixo regime da relação sinal-ruído. A estrutura deste capítulo é dada em termos de teoremas, no entanto, com o objetivo de comprovar a análise feita matematicamente, são apresentados alguns resultados de simulação computacional e comparações com resultados já consagrados da literatura.

Capítulo 5: Conclusões e Perspectivas

Além das conclusões, este capítulo é dedicado a perspectivas e trabalhos futuros relacionados a esta dissertação.

Apêndice A

Neste apêndice, são detalhados alguns resultados de teoria de majorização discutidos no Capítulo 3.

Apêndices B e C

Com a finalidade de deixar mais claro o entendimento do capítulo 4, o qual apresenta teoremas não demonstrados em sua totalidade, tais apêndices complementam tais provas.

CAPÍTULO 2

INTRODUÇÃO AOS SISTEMAS MIMO

Este capítulo procura discorrer sobre os principais conceitos utilizados no modelo proposto nesta dissertação. Inicialmente, considera-se um breve estudo sobre o canal MIMO de uma forma geral. Na sequência, realiza-se uma análise determinística e estocástica da capacidade do canal, tomando-se como referência a teoria da informação (COVER; THOMAS, 1991). Em seguida, é estudada a capacidade nas situações em que o transmissor não possui informação acerca do canal de propagação e quando apenas o transmissor dispõe desta informação. Posteriormente, abordam-se os modelos analíticos usualmente estudados na literatura: o modelo independente e identicamente distribuído (i.i.d.) e o de Kronecker. Por fim, analisa-se algumas características dos canais com desvanecimento de Rayleigh e de Rice.

2.1 O sistema MIMO

Em um sistema MIMO com n_T antenas de transmissão e n_R antenas de recepção, $n_T n_R$ subcanais são gerados entre o transmissor e o receptor. A matriz de resposta ao impulso $\mathbf{H}(\tau, t)$ de um canal MIMO é expressa da seguinte forma (PAULRAJ; NABAR; GORE, 2003):

$$\mathbf{H}(\tau, t) = \begin{bmatrix} h_{1,1}(\tau, t) & h_{1,2}(\tau, t) & \cdots & h_{1,n_T}(\tau, t) \\ h_{2,1}(\tau, t) & h_{2,2}(\tau, t) & \cdots & h_{2,n_T}(\tau, t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{n_R,1}(\tau, t) & h_{n_R,2}(\tau, t) & \cdots & h_{n_R,n_T}(\tau, t) \end{bmatrix}. \quad (2.1)$$

sendo que, cada um dos números complexos $h_{ij}(\tau, t)$ representa a resposta impulsiva variante no tempo entre a j -ésima antena transmissora e a i -ésima antena receptora. A relação entrada-saída do sistema tem o seguinte modelo matemático:

$$\mathbf{y}(t) = \int_{\tau} \mathbf{H}(\tau, t) \mathbf{x}(\tau - t) d\tau + \mathbf{n}(t) = \mathbf{H}(\tau, t) * \mathbf{x}(t) + \mathbf{n}(t) \quad (2.2)$$

em que, $\mathbf{y}(t) = [y_1(t) \ y_2(t) \ \cdots \ y_{n_R}(t)]^T \in \mathbb{C}^{n_R \times 1}$, é o vetor de símbolos recebidos durante um dado uso do canal, $\mathbf{x}(t) = [x_1(t) \ x_2(t) \ \cdots \ x_{n_T}(t)]^T \in \mathbb{C}^{n_T \times 1}$ é o vetor de símbolos transmitidos por cada antena, $\mathbf{n}(t) = [n_1(t) \ n_2(t) \ \cdots \ n_{n_R}(t)]^T \in \mathbb{C}^{n_R \times 1}$ é o vetor ruído gaussiano branco e, por fim, o símbolo $*$ representa o operador convolução. Se o canal é invariante no tempo, a matriz do canal depende apenas do atraso, assim, $\mathbf{H}(\tau, t) = \mathbf{H}(\tau)$ e a Equação (2.2) fica simplificada por:

$$\mathbf{y}(t) = \int_{\tau} \mathbf{H}(\tau) \mathbf{x}(\tau - t) d\tau + \mathbf{n}(t) = \mathbf{H}(\tau) * \mathbf{x}(t) + \mathbf{n}(t). \quad (2.3)$$

2.1.1 Modelo de canal MIMO

Considerando que o canal apresenta um comportamento constante na frequência, tem-se $\mathbf{H}(\tau) = \mathbf{H}$ e, portanto, a Equação (2.3) fica apenas com a seguinte representação matricial (OESTGES; CLERCKX, 2007)

$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t). \quad (2.4)$$

Desta forma, a saída no sistema no instante t não depende de tempos passados e os índices temporais podem ser eliminados da Equação (2.4), ficando apenas

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (2.5)$$

A Figura 1 ilustra um modelo MIMO arbitrário $n_R \times n_T$.

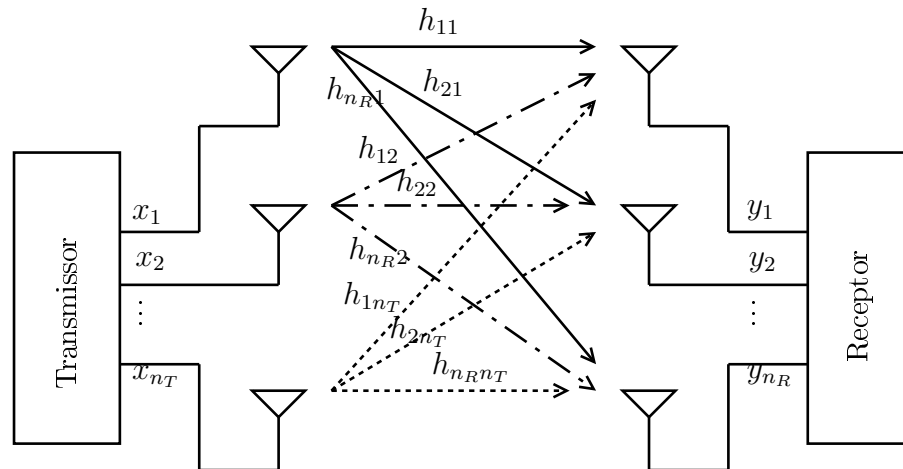


Figura 1: Representação de um canal MIMO $n_R \times n_T$.

2.2 Capacidade de um sistema MIMO

Para o estudo de capacidade que se segue, um canal MIMO sem memória será dado por

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2.6)$$

sendo $\mathbf{R}_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ a matriz de covariância do vetor de símbolos transmitidos, em que $\mathbb{E}\{\cdot\}$ é o operador esperança. A potência total é limitada pela constante P_T independentemente do número de antenas transmissoras. Além disso, serão acrescentadas as seguintes condições: as entradas da matriz do canal \mathbf{H} são variáveis aleatórias gaussianas circularmente simétricas, o vetor ruído tem distribuição $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathbf{n}})$, isto é, também segue uma distribuição gaussiana complexa circularmente simétrica cujo vetor-média é nulo e matriz de covariância é dada por $\mathbf{R}_{\mathbf{n}}$, por fim, os vetores de símbolos transmitidos e ruído são considerados independentes.

Do ponto de vista da teoria da informação, a definição da capacidade do canal apresentado no modelo em (2.6) é dada por (COVER; THOMAS, 1991):

$$C = \max_{p(\mathbf{x})} I(\mathbf{x}; \mathbf{y}) \quad (2.7)$$

em que \mathbf{x} e \mathbf{y} são variáveis aleatórias representando, respectivamente, a entrada e a saída do canal e $I(\mathbf{x}; \mathbf{y})$ representa a informação mútua entre \mathbf{x} e \mathbf{y} . A Equação (2.7) diz que a informação mútua é maximizada sobre todas as possíveis distribuições estatísticas $p(\mathbf{x})$. Sendo $H(\cdot)$ a entropia, que representa a medida da quantidade média de informação que

uma variável aleatória contém, a informação mútua $I(\mathbf{x}; \mathbf{y})$ pode ser representada em função da entropia como segue

$$I(\mathbf{x}; \mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}), \quad (2.8)$$

em que $H(\mathbf{y}|\mathbf{x})$ indica a entropia condicional entre as variáveis \mathbf{x} e \mathbf{y} . Desta forma, a informação mútua entre os vetores de entrada e saída do sistema depende das propriedades do canal que relaciona \mathbf{x} e \mathbf{y} e das propriedades de \mathbf{x} através da distribuição $p(\mathbf{x})$.

Dado que os vetores \mathbf{x} e \mathbf{n} são assumidos independentes, tem-se $H(\mathbf{n}|\mathbf{x}) = H(\mathbf{n})$, além disso

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}) &= H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}) \\ &= H(\mathbf{y}) - H(\mathbf{H}\mathbf{x} + \mathbf{n}|\mathbf{x}) \\ &= H(\mathbf{y}) - H(\mathbf{n}|\mathbf{x}) \\ &= H(\mathbf{y}) - H(\mathbf{n}). \end{aligned} \quad (2.9)$$

Portanto, da Equação (2.9) conclui-se que a maximização da informação mútua $I(\mathbf{x}, \mathbf{y})$ se restringe à maximização da entropia $H(\mathbf{y})$ (PAULRAJ; NABAR; GORE, 2003).

A entropia $H(\mathbf{y})$, no entanto, é maximizada, para uma variável fixa, quando \mathbf{y} for uma distribuição gaussiana complexa circularmente simétrica de média zero. Por outro lado, admitindo-se $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_y)$, esta consideração implica que o vetor de símbolos transmitidos \mathbf{x} também segue uma distribuição gaussiana de média nula, isto é, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_x)$ (PAULRAJ; NABAR; GORE, 2003). Diante destas hipóteses, a matriz de covariância de \mathbf{y} é dada por:

$$\begin{aligned} \mathbf{R}_y &= \mathbb{E}\left\{(\mathbf{H}\mathbf{x} + \mathbf{n})(\mathbf{H}\mathbf{x} + \mathbf{n})^H\right\} \\ &= \mathbb{E}\{\mathbf{H}\mathbf{x}\mathbf{x}^H\mathbf{H}^H\} + \mathbb{E}\{\mathbf{n}\mathbf{n}^H\} \\ &= \mathbf{H}\mathbf{R}_x\mathbf{H}^H + \mathbf{R}_n. \end{aligned} \quad (2.10)$$

Além disso, as entropias dos vetores \mathbf{y} e \mathbf{n} são expressas por (PAULRAJ; NABAR; GORE, 2003)

$$H(\mathbf{y}) = \log_2 [\det (\pi e \mathbf{R}_y)] \quad (2.11)$$

$$H(\mathbf{n}) = \log_2 [\det (\pi e \mathbf{R}_n)]. \quad (2.12)$$

Portanto,

$$\begin{aligned}
I(\mathbf{x}; \mathbf{y}) &= H(\mathbf{y}) - H(\mathbf{n}) \\
&= \log_2 [\det (\pi e \mathbf{R}_y)] - \log_2 [\det (\pi e \mathbf{R}_n)] \\
&= \log_2 \left[\frac{\det (\pi e \mathbf{R}_y)}{\det (\pi e \mathbf{R}_n)} \right] \\
&= \log_2 [\det (\mathbf{R}_y \mathbf{R}_n^{-1})] \\
&= \log_2 [\det ((\mathbf{H} \mathbf{R}_x \mathbf{H}^H + \mathbf{R}_n) \mathbf{R}_n^{-1})] \\
&= \log_2 [\det (\mathbf{H} \mathbf{R}_x \mathbf{H}^H \mathbf{R}_n^{-1} + \mathbf{I}_{n_R})]
\end{aligned} \tag{2.13}$$

Diante do resultado anterior, a capacidade do canal MIMO de (2.7) é dada por

$$C = \max_{\text{tr}(\mathbf{R}_x) \leq P_T} \left\{ \log_2 [\det (\mathbf{I}_{n_R} + \mathbf{H} \mathbf{R}_x \mathbf{H}^H \mathbf{R}_n^{-1})] \right\} \text{ bps/Hz.} \tag{2.14}$$

2.2.1 Transmissor sem informação do canal MIMO

Quando o transmissor não tem informação dos estados de canal (CSI, do inglês), geralmente considera-se que a distribuição de potência é uniforme a partir das n_T antenas transmissoras. A matriz de covariância do sinal transmitido \mathbf{R}_x pode ser escolhida estatisticamente não-preferencial à qualquer subcanal, isto é, $\mathbf{R}_x = \frac{P_T}{n_T} \mathbf{I}_{n_T}$ (CAVALCANTE; CAVALCANTI; FREITAS JR., 2008). Considerando-se que o ruído é descorrelacionado em cada ramo de recepção, então a respectiva matriz de covariância é dada por $\mathbf{R}_n = \sigma^2 \mathbf{I}_{n_R}$, sendo σ^2 a variância do ruído \mathbf{n} . Diante destas considerações, a capacidade do canal MIMO dada pela Equação (2.14), sem CSI no transmissor, fica determinada da seguinte forma

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T} \mathbf{H} \mathbf{H}^H \right) \right] \text{ bps/Hz} \tag{2.15}$$

ou simplesmente,

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{\rho}{n_T} \mathbf{H} \mathbf{H}^H \right) \right] \text{ bps/Hz.} \tag{2.16}$$

sendo $\rho = \frac{P_T}{\sigma^2}$ a relação sinal-ruído (SNR) em cada ramo da recepção.

A partir dos resultados (2.15) ou (2.16), observa-se que a capacidade do canal depende da matriz semi-definida positiva e hermitiana $\mathbf{H} \mathbf{H}^H \in \mathbb{C}^{n_R \times n_R}$. Pelo teorema espectral, esta matriz é diagonalizável (HORN; JOHNSON, 1985), isto é, existe uma matriz unitária \mathbf{U} de ordem $n_R \times n_R$ tal que $\mathbf{H} \mathbf{H}^H = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$, sendo $\mathbf{\Lambda}$ uma matriz diagonal formada

pelos autovalores de $\mathbf{H}\mathbf{H}^H$ que são todos não-negativos e, sem perda de generalidade considera-se que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_R}$. Em termos simbólicos, esta matriz é representada da seguinte maneira: $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_{n_R}\}$.

A capacidade (2.15) fica, portanto, expressa por

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T} \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \right) \right]. \quad (2.17)$$

Ao assumir que o posto da matriz do canal é igual a r , ou seja, $\text{posto}(\mathbf{H}) = r$, seguem-se os seguintes resultados (HORN; JOHNSON, 1985)

- (i) $r \leq \min\{n_T, n_R\}$
- (ii) $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$
- (iii) $\lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_{n_R} = 0$.

Pelo fato da matriz \mathbf{U} ser unitária e $\det(\mathbf{I}_m + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A})$ para quaisquer que sejam as matrizes \mathbf{A} e \mathbf{B} de ordens $m \times n$ e $n \times m$, respectivamente, a igualdade em (2.17) torna-se

$$C = \log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T} \mathbf{\Lambda} \right) \right]. \quad (2.18)$$

Por sua vez,

$$\det \left(\mathbf{I}_{n_R} + \frac{P_T}{\sigma^2 n_T} \mathbf{\Lambda} \right) = \prod_{i=1}^{n_R} \left(1 + \frac{P_T}{\sigma^2 n_T} \lambda_i \right) = \prod_{i=1}^r \left(1 + \frac{P_T}{\sigma^2 n_T} \lambda_i \right) \quad (2.19)$$

assim, a capacidade do canal MIMO, sem CSI no transmissor, é expressa por:

$$C = \sum_{i=1}^r \log_2 \left[1 + \frac{P_T}{\sigma^2 n_T} \lambda_i \right]. \quad (2.20)$$

Portanto, a capacidade do canal MIMO é o somatório das capacidades de r canais SISO independentes, em que os ganhos de canal SISO são os autovalores da matriz $\mathbf{H}\mathbf{H}^H$ e cuja potência total é distribuída uniformemente por cada antena transmissora (PAULRAJ; NABAR; GORE, 2003). Além disso, como o ganho de capacidade depende dos autovalores da matriz $\mathbf{H}\mathbf{H}^H$, então se estes apresentarem valores pequenos ou nulos, pode ser impossível a transmissão num canal SISO associado.

2.2.2 Transmissor com informação do canal MIMO

No caso em que o transmissor tem informação sobre o canal de propagação, o estudo da capacidade fica dependendo da transformação do canal MIMO em subcanais SISO paralelos como será descrito a seguir.

No que diz Biglieri *et al.* (2007), primeiramente, faz-se a decomposição em valores singulares (SVD, do inglês) da matriz \mathbf{H}

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \quad (2.21)$$

sendo que, $\mathbf{U} \in \mathbb{C}^{n_R \times n_R}$ e $\mathbf{V} \in \mathbb{C}^{n_T \times n_T}$ são matrizes unitárias, enquanto $\mathbf{\Sigma} \in \mathbb{R}_+^{n_R \times n_T}$ é uma matriz diagonal representada da seguinte forma (PAULRAJ; NABAR; GORE, 2003)

$$\mathbf{\Sigma} = \left[\begin{array}{c|c} \mathbf{D} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \quad (2.22)$$

com $\mathbf{D} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_r\} \in \mathbb{R}_+^{r \times r}$ e $\sigma_i = \sqrt{\lambda_i}$. Em seguida, pré-multiplica-se o sinal entrada $\tilde{\mathbf{x}}$ pela matriz \mathbf{V} , depois pós-multiplica-se o sinal de saída \mathbf{y} do canal por \mathbf{U}^H , conforme ilustra o esquema abaixo.

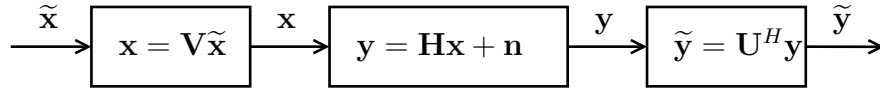


Figura 2: Representação da decomposição de um canal MIMO com CSI no transmissor.

Diante destas hipóteses, o vetor $\tilde{\mathbf{y}} \in \mathbb{C}^{n_R \times 1}$ fica determinado por

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}^H \mathbf{y} \\ &= \mathbf{U}^H (\mathbf{H}\mathbf{x} + \mathbf{n}) \\ &= \mathbf{U}^H (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H (\mathbf{V}\tilde{\mathbf{x}}) + \mathbf{n}) \\ &= \mathbf{\Sigma}\tilde{\mathbf{x}} + \mathbf{U}^H \mathbf{n} \\ &= \mathbf{\Sigma}\tilde{\mathbf{x}} + \tilde{\mathbf{n}} \end{aligned} \quad (2.23)$$

em que $\tilde{\mathbf{n}} = \mathbf{U}^H \mathbf{n}$. Como $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$, nota-se que:

- (i) $\mathbb{E}\{\|\mathbf{x}\|^2\} = \mathbb{E}\{\|\tilde{\mathbf{x}}\|^2\}$
- (ii) $\text{posto}(\mathbf{R}_{\mathbf{x}}) = \text{posto}(\mathbf{R}_{\tilde{\mathbf{x}}})$, pois $\mathbf{R}_{\mathbf{x}} = \mathbf{V}\mathbf{R}_{\tilde{\mathbf{x}}}\mathbf{V}^H$.

Além disso, desde que \mathbf{U} é unitária e $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_R})$, o vetor $\tilde{\mathbf{n}}$ segue a mesma

distribuição de \mathbf{n} , isto é, $\tilde{\mathbf{n}} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_R})$ (BIGLIERI *et al.*, 2007). Agora usando-se o fato de que $\mathbf{\Sigma}$ é uma matriz diagonal, a Equação (2.23) pode ser decomposta em r subcanais SISO paralelos da forma

$$\tilde{y}_i = \sqrt{\lambda_i} \tilde{x}_i + \tilde{n}_i, \quad i = 1, 2, \dots, r. \quad (2.24)$$

A capacidade do canal MIMO é a soma das capacidades individuais dos r subcanais SISO paralelos e é expressa por (PAULRAJ; NABAR; GORE, 2003)

$$C = \sum_{i=1}^r \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T} \lambda_i \right] \quad (2.25)$$

sendo os coeficientes de ponderação $p_i = \mathbb{E}\{|\tilde{x}_i|^2\}$ ($i = 1, 2, \dots, r$) e $\sum_{i=1}^r p_i = n_T$.

2.2.3 Transmissor com informação do canal MIMO: caso da potência ótima

Conforme visto anteriormente, a equação (2.25) corresponde à capacidade do canal MIMO com CSI no transmissor, porém, não necessariamente, esta capacidade é máxima. Contudo, para o estudo da otimalidade da capacidade, considera-se C como uma função real cujo domínio é o conjunto das r -uplas de coeficientes de ponderação obtidos de cada matriz de covariância dos vetores de símbolos transmitidos. Assim, para cada vetor $\tilde{\mathbf{x}}$ transmitido, tem-se uma matriz $\mathbf{R}_{\tilde{\mathbf{x}}}$ de posto r que define uma r -upla correspondente de coeficientes de ponderação. Com estas considerações, tem-se o seguinte problema de otimização não-linear (PAULRAJ; NABAR; GORE, 2003)

$$\begin{aligned} & \max_{\{p_k\}_{k=1}^r} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T} \lambda_i \right] \right\}, \\ & \text{sujeito a } \sum_{i=1}^r p_i = n_T. \end{aligned} \quad (2.26)$$

No entanto, este é um problema de otimização convexa, pois a função objetivo e a restrição são funções côncavas. Desta forma, garante-se a existência de um ponto ótimo $\mathbf{p}^{\text{opt}} = (p_1^{\text{opt}}, p_2^{\text{opt}}, \dots, p_r^{\text{opt}})$ que maximiza a função objetivo (BOYD; VANDENBERGHE, 2004). Ao resolver o problema (2.26) pelo método dos multiplicadores de Lagrange encontra-se

$$p_i^{\text{opt}} = \left(\mu - \frac{n_T \sigma^2}{P_T \lambda_i} \right)^+, \quad i = 1, 2, \dots, r. \quad (2.27)$$

em que μ é uma constante chamada de nível d'água e $(x)^+ = \max\{x, 0\}$ (BIGLIERI *et al.*, 2007). Noutras palavras, considerando-se os r sub-canais paralelos, e que cada um apresenta um nível de potência comum μ , o cálculo dos coeficientes de ponderação p_i^{opt} , operação denominada de *water-filling*, é traduzida pela seguinte relação (PAULRAJ; NABAR; GORE, 2003; BIGLIERI *et al.*, 2007):

$$\frac{\sigma^2 n_T}{P_T \lambda_1} + p_1^{\text{opt}} = \frac{\sigma^2 n_T}{P_T \lambda_2} + p_2^{\text{opt}} = \dots = \frac{\sigma^2 n_T}{P_T \lambda_r} + p_r^{\text{opt}} = \mu. \quad (2.28)$$

No entanto, se $\frac{\sigma^2 n_T}{P_T \lambda_i} \geq \mu$, então o coeficiente de ponderação associado é nulo. Portanto, da Equação (2.28), conclui-se que canais com maior ganho λ_i recebem mais potência, enquanto que canais com pior ganho recebem menos potência. A Figura 3 ilustra o princípio do algoritmo *water-filling*.

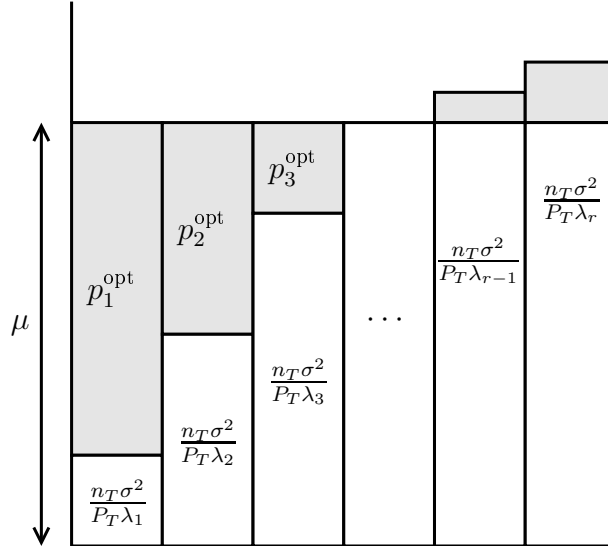


Figura 3: Princípio do algoritmo *water-filling*.

2.2.4 Capacidade Ergódica

Quando o canal MIMO apresenta-se como matriz aleatória, também chamada de canal estocástico, implica que a capacidade do canal também é uma variável aleatória. Neste caso, quando cada subcanal corresponde a uma realização independente de \mathbf{H} , a capacidade ergódica ou capacidade de Shannon, representa a média da taxa de informação sobre as distribuições dos elementos do canal. Noutros termos, a capacidade ergódica é a máxima informação mútua sobre todos os estados do canal e é dada por (PAULRAJ;

NABAR; GORE, 2003; CAVALCANTE; CAVALCANTI; FREITAS JR., 2008):

$$\bar{C} = \mathbb{E} \left\{ \max_{p(\mathbf{x}) : \text{tr}(\mathbf{R}_{\mathbf{x}}) \leq P_T} I(\mathbf{x}; \mathbf{y}) \right\}. \quad (2.29)$$

Os resultados das respectivas capacidades são semelhantes em se tratando de matriz do canal determinística ou estocástica, assim, no caso do canal ser desconhecido no transmissor a capacidade ergódica é dada por

$$\bar{C} = \mathbb{E} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{P_T}{\sigma^2 n_T} \lambda_i \right] \right\}, \quad (2.30)$$

e quando transmissor conhece perfeitamente toda informação sobre o canal, tem-se

$$\bar{C} = \mathbb{E} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T} \lambda_i \right] \right\}. \quad (2.31)$$

No entanto, no caso da capacidade máxima com CSI no transmissor, a capacidade ergódica é

$$\bar{C}_{\text{opt}} = \mathbb{E} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{p_i^{\text{opt}} P_T}{\sigma^2 n_T} \lambda_i \right] \right\}, \quad (2.32)$$

sendo p_i^{opt} obtido pelo problema de otimização convexa (2.26).

2.3 Modelo analíticos de canais MIMO

Os modelos analíticos (não-físicos) são baseados em parâmetros estatísticos do canal de propagação. A modelagem estatística do canal procura estabelecer apenas as relações matemáticas entre os parâmetros e as estatísticas destes parâmetros e, em geral, são modelos propícios à simulação e proporcionam uma caracterização precisa do canal nas situações para as quais os parâmetros foram determinados (ALMERS *et al.*, 2007). Por outro lado, estes modelos apresentam uma visão limitada das características de propagação dos canais MIMO e dependem do equipamento de medição, da largura de banda, da configuração dos arranjos das antenas e da altura destas (CAVALCANTE; CAVALCANTI; FREITAS JR., 2008). Dentre os modelos analíticos existentes, para modelagem do canal MIMO, o presente trabalho dará destaque aos modelos (ALMERS *et al.*, 2007):

- ▶ i.i.d. (independente e identicamente distribuído);
- ▶ Kronecker.

Estes modelos inserem em seus desenvolvimentos a matriz

$$\mathbf{R}_{\mathbf{H}} = \mathbb{E}\left\{\text{vec}(\mathbf{H}) \text{vec}(\mathbf{H})^H\right\} \in \mathbb{C}^{n_R n_T \times n_R n_T} \quad (2.33)$$

chamada matriz de correlação do canal a qual é hermitiana e semi-definida positiva (YU *et al.*, 2004). De uma maneira geral, dada uma matriz $\mathbf{A} \in \mathbb{C}^{m \times n}$ tal que $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n]$, o vetor $\text{vec}(\mathbf{A}) \in \mathbb{C}^{mn \times 1}$ é obtido pelo empilhamento das colunas de \mathbf{A} , isto é, $\text{vec}(\mathbf{A}) = [\mathbf{a}_1^T \ \mathbf{a}_2^T \ \cdots \ \mathbf{a}_n^T]^T$ (HORN; JOHNSON, 1985).

2.3.1 Modelo i.i.d.

No modelo independente e identicamente distribuído (i.i.d.), assume-se que não existe nenhuma correlação entre os sinais nas antenas transmissoras e nas antenas receptoras, desta forma a matriz de correlação do canal MIMO é dada por $\mathbf{R}_{\mathbf{H}} = \sigma^2 \mathbf{I}_{n_R n_T}$. Noutros termos, os elementos da matriz \mathbf{H} são descorrelacionados e estatisticamente independentes e, além disso, apresentam a mesma variância σ^2 .

O modelo i.i.d. consiste apenas num parâmetro σ^2 (potência do canal) e é frequentemente utilizado para considerações teóricas na análise de sistemas MIMO. Este modelo é bastante simples mas só é válido para ambientes com elevado espalhamento multipercurso (ALMERS *et al.*, 2007).

O modelo clássico de canal MIMO i.i.d., também chamado de espacialmente branco, é definido quando os elementos da matriz canal \mathbf{H} possuem média zero, distribuição gaussiana complexa circularmente simétrica e variância unitária, ou seja, $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{n_R} \otimes \mathbf{I}_{n_T})$, sendo $\mathbb{E}\{\mathbf{H}\} = \mathbf{0}$ e $\mathbf{R}_{\mathbf{H}} = \mathbf{I}_{n_R} \otimes \mathbf{I}_{n_T}$, em que o símbolo \otimes representa o produto de Kronecker. Tradicionalmente, quando o canal é espacialmente branco, a matriz do canal é representada por \mathbf{H}_w e esta é chamada de matriz espacialmente branca ou simplesmente matriz branca (PAULRAJ; NABAR; GORE, 2003; ZARBOUTI; TSOULOS; KAKLAMANI, 2006).

Considerando-se que os elementos de \mathbf{H}_w são representados por \tilde{h}_i , seguem-se abaixo algumas propriedades estatísticas da matriz branca (HAYKIN, 2001):

$$\mathbb{E}\{\tilde{h}_i\} = 0, \quad (2.34)$$

$$\mathbb{E}\{|\tilde{h}_i|^2\} = 1, \quad (2.35)$$

$$\mathbb{E}\{\tilde{h}_i \tilde{h}_j^*\} = 0, \text{ se } i \neq j, \quad (2.36)$$

$$\mathbb{E}\{\tilde{h}_{s_1}^* \tilde{h}_{s_2}^* \cdots \tilde{h}_{s_k}^* \tilde{h}_{t_1} \tilde{h}_{t_2} \cdots \tilde{h}_{t_p}\} = 0, \text{ se } k \neq p, \quad (2.37)$$

$$\mathbb{E}\{\tilde{h}_{s_1}^* \tilde{h}_{s_2}^* \cdots \tilde{h}_{s_k}^* \tilde{h}_{t_1} \tilde{h}_{t_2} \cdots \tilde{h}_{t_p}\} = \prod_{i=1}^k \mathbb{E}\{\tilde{h}_{s_{\pi(i)}}^* \tilde{h}_{t_i}\}, \text{ se } k = p. \quad (2.38)$$

sendo π uma permutação do conjunto $\{1, 2, \dots, k\} \subset \mathbb{N}$ e $\pi(i)$ o i -ésimo elemento da permutação.

2.3.2 Modelo de Kronecker

No caso dos elementos da matriz \mathbf{H} serem correlacionados, o modelo de Kronecker modela o canal pela equação (PAULRAJ; NABAR; GORE, 2003)

$$\text{vec}(\mathbf{H}) = \mathbf{R}_{\mathbf{H}}^{1/2} \text{vec}(\mathbf{H}_w), \quad (2.39)$$

sendo $\mathbf{H}_w \in \mathbb{C}^{n_R \times n_T}$ a matriz espacialmente branca e $\mathbf{R}_{\mathbf{H}} \in \mathbb{C}^{n_R n_T \times n_R n_T}$ a matriz de covariância do canal. De forma equivalente, este modelo propõe a obtenção da matriz do canal, \mathbf{H} , a partir de duas matrizes de correlação, sendo uma para o transmissor, $\mathbf{R}_{\mathbf{T}_x} = \mathbb{E}\{\mathbf{H}^H \mathbf{H}\} \in \mathbb{C}^{n_T \times n_T}$, e outra para o receptor, $\mathbf{R}_{\mathbf{R}_x} = \mathbb{E}\{\mathbf{H} \mathbf{H}^H\} \in \mathbb{C}^{n_R \times n_R}$, se as estatísticas de desvanecimento no transmissor e no receptor forem assumidas independentes, conforme mostra a equação abaixo (YU *et al.*, 2004)

$$\mathbf{H} = \mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T. \quad (2.40)$$

No entanto, as matrizes de correlação $\mathbf{R}_{\mathbf{T}_x}$ e $\mathbf{R}_{\mathbf{R}_x}$ são semi-definidas positivas e hermitianas.

De acordo com a Equação (2.40), é possível analisar o posto da matriz do canal em função dos postos das matrizes de correlação no transmissor e receptor. De fato, pelo teorema espectral, as matrizes $\mathbf{R}_{\mathbf{T}_x}$ e $\mathbf{R}_{\mathbf{R}_x}$ são diagonalizáveis, assim existem matrizes unitárias \mathbf{U}_T e \mathbf{U}_R tais que $\mathbf{R}_{\mathbf{T}_x} = \mathbf{U}_T \mathbf{D}_T \mathbf{U}_T^H$ e $\mathbf{R}_{\mathbf{R}_x} = \mathbf{U}_R \mathbf{D}_R \mathbf{U}_R^H$, em que \mathbf{D}_T e \mathbf{D}_R são matrizes diagonais formadas pelos autovetores das matrizes $\mathbf{R}_{\mathbf{T}_x}$ e $\mathbf{R}_{\mathbf{R}_x}$, respectivamente. Como os elementos da matriz espacialmente branca são totalmente decorrelacionados, segue-se que esta matriz tem posto completo, isto é, $\text{posto}(\mathbf{H}_w) = \min\{n_R, n_T\}$. Desta

forma, dois casos são considerados:

- $\text{posto}(\mathbf{H}_w) = n_T$;
- $\text{posto}(\mathbf{H}_w) = n_R$.

Para o primeiro caso, tem-se:

$$\begin{aligned}
\text{posto}(\mathbf{H}) &= \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \\
&\leq \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \right), \text{posto} \left(\mathbf{H}_w \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \right\} \\
&= \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \right), \text{posto} \left(\left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \right\} \\
&= \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \right), \text{posto} \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right) \right\} \\
&= \min \{ \text{posto}(\mathbf{R}_{\mathbf{R}_x}), \text{posto}(\mathbf{R}_{\mathbf{T}_x}) \}
\end{aligned} \tag{2.41}$$

Para o segundo caso, basta notar que

$$\text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \leq \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \right), \text{posto} \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right) \right\} \tag{2.42}$$

e

$$\text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \right) = \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \right) = \text{posto}(\mathbf{R}_{\mathbf{R}_x}). \tag{2.43}$$

Assim, o resultado segue análogo ao caso anterior. De fato,

$$\begin{aligned}
\text{posto}(\mathbf{H}) &= \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \\
&\leq \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \mathbf{H}_w \right), \text{posto} \left(\left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right)^T \right) \right\} \\
&= \min \left\{ \text{posto} \left(\mathbf{R}_{\mathbf{R}_x}^{1/2} \right), \text{posto} \left(\mathbf{R}_{\mathbf{T}_x}^{1/2} \right) \right\} \\
&= \min \{ \text{posto}(\mathbf{R}_{\mathbf{R}_x}), \text{posto}(\mathbf{R}_{\mathbf{T}_x}) \}
\end{aligned} \tag{2.44}$$

Portanto, para qualquer situação tem-se,

$$\text{posto}(\mathbf{H}) \leq \min \{ \text{posto}(\mathbf{R}_{\mathbf{R}_x}), \text{posto}(\mathbf{R}_{\mathbf{T}_x}) \}, \tag{2.45}$$

isto é, o posto da matriz do canal fica restrito ao mínimo entre os postos das matrizes de correlação do transmissor e receptor.

Com as Equações (2.39) e (2.40), a matriz de correlação do canal pode ser decomposta nas duas matrizes de correlação, $\mathbf{R}_{\mathbf{T}_x}$ e $\mathbf{R}_{\mathbf{R}_x}$, da seguinte forma, (ZHANG; PALOMAR;

OTTERSTEN, 2008; YU *et al.*, 2004)

$$\mathbf{R}_H = \mathbf{R}_{T_x} \otimes \mathbf{R}_{R_x}. \quad (2.46)$$

Para tal, faz-se uso do resultado $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$, dentre outros relacionados ao produto de Kronecker (HARVILLE, 2008), ou seja,

$$\begin{aligned} \mathbf{R}_H &= \mathbb{E} \left\{ \text{vec}(\mathbf{H}) \text{vec}(\mathbf{H})^H \right\} \\ &= \mathbb{E} \left\{ \text{vec} \left(\mathbf{R}_{R_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{T_x}^{1/2} \right)^T \right) \text{vec} \left(\mathbf{R}_{R_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{T_x}^{1/2} \right)^T \right)^H \right\} \\ &= \mathbb{E} \left\{ \left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right) \text{vec}(\mathbf{H}_w) \left[\left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right) \text{vec}(\mathbf{H}_w) \right]^H \right\} \\ &= \left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right) \mathbb{E} \left\{ \text{vec}(\mathbf{H}_w) \text{vec}(\mathbf{H}_w)^H \right\} \left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right)^H \\ &= \left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right) \left(\mathbf{R}_{T_x}^{1/2} \otimes \mathbf{R}_{R_x}^{1/2} \right) \\ &= \left(\mathbf{R}_{T_x}^{1/2} \mathbf{R}_{T_x}^{1/2} \right) \otimes \left(\mathbf{R}_{R_x}^{1/2} \mathbf{R}_{R_x}^{1/2} \right) \\ &= \mathbf{R}_{T_x} \otimes \mathbf{R}_{R_x} \end{aligned} \quad (2.47)$$

o que mostra o resultado em (2.46).

2.4 Canal com desvanecimento de Rice e de Rayleigh

2.4.1 Características gerais de propagação

Na propagação no meio sem fio, o sinal transmitido pode chegar ao receptor através de muitos percursos (multipercursos ou multicaminhos) e a origem destes multipercursos deve-se a reflexões, refrações e difrações que originam múltiplas componentes do sinal, tal como ilustra a Figura 4. Assim, o sinal recebido é constituído por um número aleatório de componentes multipercurso em que cada um tem amplitude, fase, desvio Doppler e atraso aleatórios. Além disso, se o transmissor ou receptor estiverem em movimento, esses fenômenos de propagação são variáveis no tempo e provocarão desvanecimento (*fading* em inglês) que são variações na amplitude do sinal (HAYKIN; MOHER, 2008).

Normalmente, o sinal está sujeito a dois tipos de variações:

- ▶ desvanecimento lento;
- ▶ desvanecimento rápido.

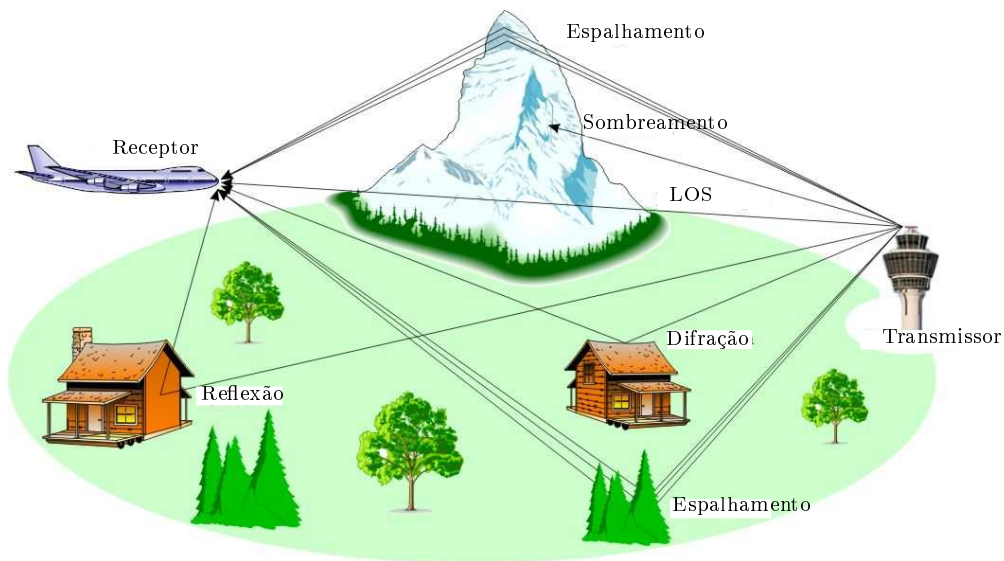


Figura 4: Propagação multipercursos.

O desvanecimento lento, também conhecido como sombreamento (*shadowing* em inglês), é causado por diferentes obstáculos de grandes dimensões encontrados entre transmissor e receptor. No entanto, como geralmente o ambiente não se altera bruscamente, num curto período de tempo, estas variações são consideradas lentas. O sombreamento é modelado por uma variável aleatória (v.a.) com distribuição log-normal (GOLDSMITH, 2005).

O desvanecimento rápido está associado ao multipercursos. Neste caso, com pequenas alterações no ambiente ou no posicionamento do receptor há uma variação rápida na amplitude do sinal num curto período de tempo. No entanto, quando o ambiente sem fio é caracterizado por uma série de obstáculos que geram uma grande quantidade de componentes multipercursos, ou seja, quando não há linha de visada (LOS, do inglês) entre o transmissor e o receptor, a envoltória do sinal recebido pode ser descrita estatisticamente usando a distribuição de Rayleigh. Por outro lado, quando se tem linha de visada entre o transmissor e o receptor, o desvanecimento rápido é descrito estatisticamente por uma distribuição de Rice (ZARBOUTI; TSOULOS; KAKLAMANI, 2006).

Na prática, para cada componente multipercursos, o ângulo de chegada é uniformemente distribuído no intervalo $[0, 2\pi]$ e as componentes em fase, $I(t)$, e em quadratura, $Q(t)$, são consideradas independentes. Além disso, assume-se que estas componentes possuem mesma variância σ^2 .

A envoltória do sinal recebido é definida por (PROAKIS, 1983)

$$r(t) = \sqrt{I(t)^2 + Q(t)^2}, \quad (2.48)$$

e quando entre o receptor e o transmissor não existe LOS, esta segue a distribuição de Rayleigh cuja função densidade de probabilidade (pdf) é dada por

$$p_{\text{Rayleigh}}(z) = \begin{cases} \frac{z}{\sigma^2} \exp\left(-\frac{z^2}{2\sigma^2}\right), & \text{se } z \geq 0, \\ 0, & \text{se } z < 0. \end{cases} \quad (2.49)$$

Por sua vez, na presença de linha de visada, a envoltória segue a distribuição de Rice cuja pdf é

$$p_{\text{Rice}}(z) = \begin{cases} \frac{z}{\sigma^2} \exp\left(-\frac{z^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{zs}{\sigma^2}\right), & \text{se } z \geq 0, \\ 0, & \text{se } z < 0, \end{cases} \quad (2.50)$$

em que, o parâmetro $s \in \mathbb{R}_+$ corresponde a potência da componente da linha de visada e $I_0(x) \triangleq \frac{1}{\pi} \int_0^\pi \exp(x \cos \phi) d\phi$ é a função de Bessel modificada de primeira espécie de ordem zero. A potência média do sinal recebido é dada por $\mathbb{E}\{r^2\} = s^2 + 2\sigma^2$, a qual será simplesmente denotada pela constante por Ω .

De acordo com a Equação (2.50), quando s tende a zero, a distribuição de Rice tende para a distribuição de Rayleigh, ou seja,

$$\lim_{s \rightarrow 0} p_{\text{Rice}} = p_{\text{Rayleigh}}. \quad (2.51)$$

Usualmente, a distribuição de Rice também é descrita em termos de uma constante K definida por $K = \frac{s^2}{2\sigma^2}$, chamada de fator de Rice, conforme mostra a seguinte pdf (GOLDSMITH, 2005)

$$p_{\text{Rice}}(z) = \begin{cases} \frac{2(K+1)}{\Omega} z \exp\left[-\frac{(K+1)z^2}{\Omega} - K\right] I_0\left(2z\sqrt{\frac{K(K+1)}{\Omega}}\right), & \text{se } z \geq 0 \\ 0, & \text{se } z < 0. \end{cases} \quad (2.52)$$

Neste caso, se K tende a zero, a distribuição Rice transforma-se na distribuição Rayleigh e, se $K \rightarrow \infty$, então o canal não possui desvanecimento.

2.4.2 Canal com desvanecimento de Rayleigh

Segundo Paulraj, Nabar e Gore (2003), a matriz do canal MIMO $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{n_R \times n_T}$ é chamada de canal de Rayleigh quando:

- (H1) o modelo analítico considerado neste canal é i.i.d. (ver Seção 2.3.1);
- (H2) cada elemento h_{ij} segue uma distribuição gaussiana complexa circularmente simétrica
- (H3) h_{ij} tem média nula e variância unitária;
- (H4) a envoltória de cada elemento h_{ij} segue uma distribuição de Rayleigh.

Do exposto acima, segue que cada um dos elementos h_{ij} não tem correlação com os outros, além disso, a matriz do canal torna-se a matriz espacialmente branca, ou seja, $\mathbf{H} = \mathbf{H}_w$.

Quando o número de antenas transmissoras é igual ao número de antenas receptoras, isto é, $n_T = n_R = n$, e considerando-se o caso assintótico $n \rightarrow \infty$, a capacidade ergódica sobre o canal de Rayleigh torna-se determinística e cresce linearmente com o número de antenas para uma SNR fixa (PAULRAJ; NABAR; GORE, 2003). De fato, a capacidade ergódica neste caso é dada pela Equação (2.29), ou seja,

$$\bar{C} = \mathbb{E}\left\{\log_2\left[\det\left(\mathbf{I}_n + \frac{\rho}{n}\mathbf{H}_w\mathbf{H}_w^H\right)\right]\right\}. \quad (2.53)$$

Pela lei forte dos grandes números (LEON-GARCIA, 1994)

$$\text{para } n \rightarrow \infty, \text{ isto implica que } \frac{1}{n}\mathbf{H}_w\mathbf{H}_w^H \rightarrow \mathbf{I}_n, \quad (2.54)$$

assim, quando $n \rightarrow \infty$ segue que

$$\bar{C} \rightarrow \log_2 [\mathbf{I}_n + \rho \mathbf{I}_n] = \log_2 (1 + \rho)^n = n \log_2 (1 + \rho), \quad (2.55)$$

o que prova o resultado.

2.4.3 Canal com desvanecimento de Rice

No caso do canal de Rice, o sinal recebido é decomposto em duas componentes: uma com relação a matriz de linha de visada e a outra a partir das reflexões dos multipercursos (KHALIGHI *et al.*, 2001). Desta forma, se os elementos da matriz \mathbf{H}_{Rice} são descorrelacionados, o canal de Rice é modelado por:

$$\mathbf{H}_{\text{Rice}} = \sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}} + \sqrt{\frac{\Omega}{K+1}} \mathbf{H}_w, \quad (2.56)$$

sendo $\bar{\mathbf{H}} \in \mathbb{C}^{n_R \times n_T}$ uma matriz determinística, cujos elementos possuem potência normalizada, K é o fator de Rice, a constante Ω é a potência média de cada elemento da matriz do canal e, por fim, \mathbf{H}_w é a matriz espacialmente branca. Segundo Zarbouti *et al.* (ZARBOUTI; TSOULOS; KAKLAMANI, 2006), a matriz $\bar{\mathbf{H}}$ é definida por

$$\bar{\mathbf{H}} = e^{j\phi_0} \begin{bmatrix} 1 & e^{j\theta} & \dots & e^{j(n_T-1)\theta} \\ e^{-j\theta} & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & e^{j\theta} \\ e^{-j(n_R-1)\theta} & e^{-j(n_R-2)\theta} & \dots & 1 \end{bmatrix} \quad (2.57)$$

em que ϕ_0 corresponde mudança de fase do sinal na propagação da antena de transmissão para a antena de recepção correspondente e θ é a mudança de fase entre os elementos vizinhos da matriz.

Diante da Equação (2.56), observar-se que a esperança do canal de Rice é dada por $\mathbb{E}\{\mathbf{H}_{\text{Rice}}\} = \sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}}$, além disso, com relação à matriz de covariância do canal tem-se $\mathbf{R}_{\mathbf{H}_{\text{Rice}}} = \frac{\Omega}{K+1} \mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}$. Com efeito, definindo $\text{cov}(X, Y)$ a covariância entre as v.a.'s complexas arbitrárias X e Y , sabe-se que (PAPOULIS, 1991)

$$\text{cov}(X, Y) = \mathbb{E}\{XY^*\} - \mathbb{E}\{X\}\mathbb{E}\{Y^*\}. \quad (2.58)$$

Aplicando este resultado aos elementos da matriz \mathbf{H} , observemos que

$$\text{cov}(h_{ij}, h_{kl}) = \mathbb{E}\{h_{ij}h_{kl}^*\} - \mathbb{E}\{h_{ij}\}\mathbb{E}\{h_{kl}^*\}, \quad (2.59)$$

sendo

$$\begin{aligned} \mathbb{E}\{h_{ij}h_{kl}^*\} &= \mathbb{E}\left\{\left(\sqrt{\frac{\Omega K}{K+1}}\bar{h}_{ij} + \sqrt{\frac{\Omega}{K+1}}\tilde{h}_{ij}\right) \cdot \left(\sqrt{\frac{\Omega K}{K+1}}\bar{h}_{kl}^* + \sqrt{\frac{\Omega}{K+1}}\tilde{h}_{kl}^*\right)\right\} \\ &= \frac{\Omega K}{K+1}\bar{h}_{ij}\bar{h}_{kl}^* + \frac{\Omega}{K+1}\mathbb{E}\{\tilde{h}_{ij}\tilde{h}_{kl}^*\} \end{aligned} \quad (2.60)$$

e

$$\begin{aligned} \mathbb{E}\{h_{ij}\}\mathbb{E}\{h_{kl}^*\} &= \mathbb{E}\left\{\sqrt{\frac{\Omega K}{K+1}}\bar{h}_{ij} + \sqrt{\frac{\Omega}{K+1}}\tilde{h}_{ij}\right\} \cdot \mathbb{E}\left\{\sqrt{\frac{\Omega K}{K+1}}\bar{h}_{kl}^* + \sqrt{\frac{\Omega}{K+1}}\tilde{h}_{kl}^*\right\} \\ &= \frac{\Omega K}{K+1}\bar{h}_{ij}\bar{h}_{kl}^*. \end{aligned} \quad (2.61)$$

Assim,

$$\text{cov}(h_{ij}, h_{kl}) = \frac{\Omega}{K+1}\mathbb{E}\{\tilde{h}_{ij}\tilde{h}_{kl}^*\}. \quad (2.62)$$

Desta forma, se $i = k$ e $j = l$, tem-se $\mathbb{E}\{\tilde{h}_{ij}\tilde{h}_{kl}^*\} = \mathbb{E}\{|\tilde{h}_{ij}|^2\} = 1$ e, conseqüentemente, $\text{cov}(h_{ij}, h_{ij}) = \text{var}(h_{ij}) = \frac{\Omega K}{K+1}$. Agora, no caso de $i \neq k$ ou $j \neq l$, tem-se $\text{cov}(h_{ij}, h_{kl}) = 0$. Diante destas considerações, a matriz de covariância do canal é dada por

$$\mathbf{R}_{\mathbf{H}_{\text{Rice}}} = \frac{\Omega}{K+1}\mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}. \quad (2.63)$$

Por sua vez, a matriz do canal segue uma distribuição gaussiana¹ (JAYAWEERA; POOR, 2005; JIN; GAO; YOU, 2007), ou seja,

$$\mathbf{H}_{\text{Rice}} \sim \mathcal{CN}\left(\sqrt{\frac{\Omega K}{K+1}}\bar{\mathbf{H}}, \frac{\Omega}{K+1}\mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}\right). \quad (2.64)$$

Situação semelhante ocorre quando os elementos da matriz do canal de Rice são correlacionados, nesta análise tem-se (ZHANG; PALOMAR; OTTERSTEN, 2008)

¹Tal análise pode ser verificada a partir da função geradora de momentos em cada elemento da matriz do canal (MAGALHÃES, 2004).

$$\mathbf{H}_{\text{Rice}} = \sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}} + \sqrt{\frac{\Omega}{K+1}} \mathbf{R}_{\text{R}_x}^{1/2} \mathbf{H}_w \left(\mathbf{R}_{\text{T}_x}^{1/2} \right)^T, \quad (2.65)$$

sendo \mathbf{R}_{R_x} e \mathbf{R}_{T_x} as matrizes de correlação no receptor e no transmissor, respectivamente. Analogamente ao caso anterior, segue que (JIN; GAO; YOU, 2007)

$$\mathbf{H}_{\text{Rice}} \sim \mathcal{CN} \left(\sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}}, \frac{\Omega}{K+1} \mathbf{R}_{\text{T}_x} \otimes \mathbf{R}_{\text{R}_x} \right). \quad (2.66)$$

2.5 Conclusões e síntese

Como o objetivo desta dissertação é estudar a capacidade ergódica de um canal MIMO com desvanecimento de Rice, este capítulo deteve-se em explorar os fundamentos relativos a sistemas MIMO, com definições e modelos, dando ênfase ao estudo da capacidade do canal, bem como as principais características do canal de desvanecimento de Rice. Ademais, as notações utilizadas nas formulações de tais modelos, dentre outras, serão utilizadas no restante deste trabalho.

Particularmente, analisando a capacidade ergódica da Equação (2.31), a qual é obtida quando o transmissor conhece perfeitamente toda informação sobre o canal, nota-se que tal expressão está em função dos autovalores de uma matriz hermitiana específica. Ao considerar, portanto, o referido resultado como uma função real de várias variáveis e aplicando-se a teoria de majorização, que será apresentada no próximo capítulo, esta hipótese terá grande impacto no capítulo 4.

CAPÍTULO 3

TEORIA DE MAJORIZAÇÃO

A teoria de majorização foi desenvolvida a partir da expansão da teoria de inequações matemáticas. A princípio, esta teoria consiste em estabelecer uma comparação entre dois vetores de \mathbb{R}^n a partir da re-ordenação decrescente de suas coordenadas. Alguns resultados já eram conhecidos, tais como: a *curva de Lorentz* (1905), que consiste em determinar como as distribuições de renda ou riqueza podem ser comparadas a partir de uma determinada população; o *princípio de transferência* (1920) que também atuava no contexto de distribuição de renda; a *desigualdade de Hadamard*, apresentada por Schur em 1923, dentre outros resultados (MARSHALL; OLKIN, 1979). Portanto, com o intuito de formalizar tais idéias, dentre outras, o livro *Inequalities* de Hardy, Littlewood e Pólya da década de 1930, foi o primeiro a unificar os assuntos existentes, apresentando, então, definições, notações e o desenvolvimento de resultados dentro deste novo formalismo matemático (HARDY; LITTLEWOOD; PÓLYA, 1934).

A partir da publicação do trabalho de Hardy, Littlewood e Pólya (1934), o interesse sobre o assunto foi se popularizando na comunidade científica e, naturalmente, outros importantes resultados foram surgindo. No ano de 1970, Marshall e Olkin documentaram alguns destes, num trabalho monográfico, de aproximadamente 100 páginas gerado a partir de discussões ocorridas em seminários apresentados na Universidade de Cambridge nos anos de 1967 e 1968. Este documento foi expandido e originou em 1979 o livro *Inequalities: Theory of Majorization and its Applications*, que é considerado a principal referência no assunto (MARSHALL; OLKIN, 1979).

A teoria de majorização apresenta interessantes resultados aplicados à álgebra linear, principalmente, envolvendo autovalores de matrizes reais e complexas com relação à desigualdades. Desta forma, esta se inseriu como uma aliada à resolução de problemas de várias áreas do conhecimento, inclusive em sistemas de comunicação sem fio, que é o foco deste trabalho. Zhang, Palomar e Ottersten (2008) enumeram algumas aplicações da teoria de majorização, em sistemas MIMO, com relação à problemas de otimização, tais como:

- ▶ minimização da soma ponderada de erros quadráticos médios (MSEs, do inglês);
- ▶ minimização do produto de MSEs;
- ▶ maximização da soma ponderada de SNRs;
- ▶ maximização do produto de SNRs;
- ▶ maximização da média harmônica de SNRs;
- ▶ minimização da taxa de erro de bit (BER) médio no caso de constelações iguais, dentre outros.

A principal contribuição da teoria de majorização, em grande parte destes problemas, está na reformulação do problema inicial, ou seja, a partir de certas manipulações, transforma-se um problema geralmente não-convexo em convexo (PALOMAR; JIANG, 2006; PALOMAR; LAGUNAS; CIOFFI, 2004). Ademais, existem outras aplicações em sistemas MIMO via teoria majorização. Por exemplo, Zhong, Wong e Jin (2009) obtêm limitantes, inferior e superior, para a capacidade ergódica no caso de um canal de distribuição Nakagami- m , Zhou *et al.* (2008) estudam o impacto da correlação espacial sobre a capacidade ergódica de um canal MIMO. Outras aplicações, portanto, podem ser encontradas em (PALOMAR; JIANG, 2006; JORSWIECK; BOCHE, 2007).

O objetivo deste capítulo é, portanto, apresentar uma revisão da literatura sobre os principais aspectos da teoria de majorização, dando ênfase ao estudo matemático envolvido que, por sua vez, se aplica aos sistemas de comunicação sem fio. Uma vez que a literatura sobre o assunto, em geral, não está direcionada para os problemas de telecomunicações, acredita-se que esta revisão possa contribuir com os trabalhos futuros sobre o assunto.

Os fundamentos sobre a teoria de majorização, apresentados neste trabalho, estão baseados em referências que podem ser consideradas clássicas sobre o assunto: *Inequalities*:

Theory of Majorization and its Applications de Marshall e Olkin (1979) e *Matrix Analysis* de Bhatia (1997).

3.1 Definições e resultados básicos

A relação de majorização \prec permite comparar dois vetores $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ de \mathbb{R}^n a partir de suas coordenadas, para tal deve-se considerar os vetores $[\mathbf{x}] = (x_{[1]}, x_{[2]}, \dots, x_{[n]})$ e $[\mathbf{y}] = (y_{[1]}, y_{[2]}, \dots, y_{[n]})$ também de \mathbb{R}^n , obtidos de \mathbf{x} e \mathbf{y} , respectivamente, por reordenação de suas coordenadas numa sequência decrescente, isto é, $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[n]}$ e $y_{[1]} \geq y_{[2]} \geq \dots \geq y_{[n]}$. Diz que o \mathbf{x} é majorizado por \mathbf{y} ou que \mathbf{y} majoriza \mathbf{x} , em que se escreve $\mathbf{x} \prec \mathbf{y}$, se as seguintes condições forem satisfeitas:

$$\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, \quad 1 \leq k \leq n-1, \quad (3.1)$$

$$\sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}. \quad (3.2)$$

Noutras palavras, o vetor \mathbf{y} majoriza \mathbf{x} se as coordenadas de \mathbf{y} são mais dispersas ou “espalhadas” que as de \mathbf{x} (MARSHALL; OLKIN, 1979). Além disso, estas coordenadas estão sujeitas à restrição de apresentarem mesmo somatório. Para ilustrar esta definição, considera-se, por exemplo, os vetores $\mathbf{x} = (3; 3; 3; 3; 3)$, $\mathbf{y} = (5; 4; 3; 2; 1)$ e $\mathbf{z} = (7; 5; 2; 0,8; 0,2)$ de \mathbb{R}^5 , nota-se que as condições (3.1) e (3.2) são satisfeitas e $\mathbf{x} \prec \mathbf{y} \prec \mathbf{z}$. A Figura 5 mostra, portanto, que as coordenadas de \mathbf{z} apresentam comportamento mais “desordenado” que as de \mathbf{x} e \mathbf{y} respectivamente, e, por sua vez, as coordenadas de \mathbf{y} são mais desordenadas que as do vetor \mathbf{x} .

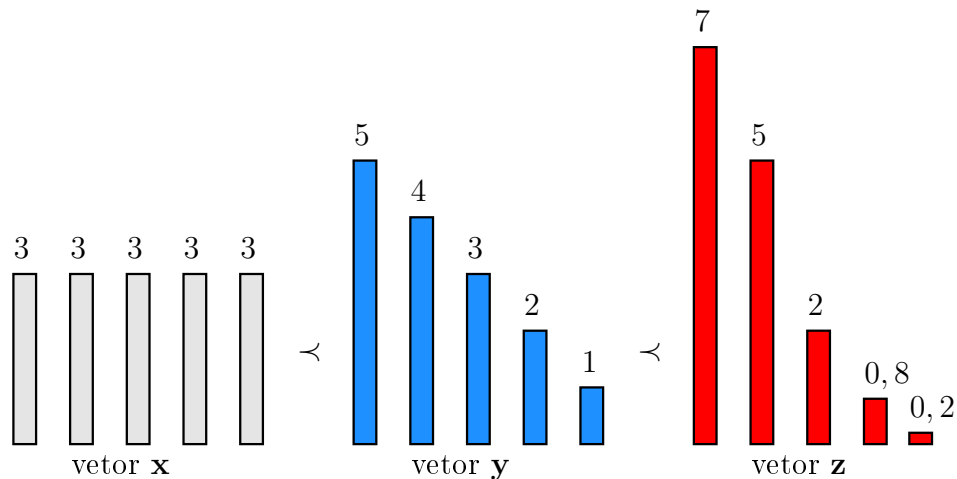


Figura 5: Interpretação geométrica de majorização de vetores.

No entanto, nem todos os vetores de \mathbb{R}^n podem ser comparados via relação de majorização, mesmo satisfazendo a condição de igualdade da soma de suas coordenadas. Por exemplo, os vetores $\mathbf{x} = (1; 2; 3)$ e $\mathbf{y} = (3; 5; 1, 2; 1, 3)$, neste $\sum_{i=1}^3 x_{[i]} = \sum_{i=1}^3 y_{[i]} = 6$, porém \mathbf{x} não majoriza \mathbf{y} , pois $x_{[1]} < y_{[1]}$, e \mathbf{y} não majoriza \mathbf{x} , já que $x_{[1]} + x_{[2]} > y_{[1]} + y_{[2]}$.

Um outro fato interessante trata-se quando $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$. Neste porém, não necessariamente, implica que $\mathbf{x} = \mathbf{y}$, como pode ser verificado pelos vetores distintos $\mathbf{x} = (1; 2; 3)$ e $\mathbf{y} = (3; 2; 1)$. Todavia, esta igualdade é satisfeita quando os vetores pertencem ao conjunto convexo $\mathcal{D}_n = \{\mathbf{x} \in \mathbb{R}^n; x_1 \geq x_2 \geq \dots \geq x_n\}$ ¹. A Figura 6 ilustra um conjunto \mathcal{A} contido em \mathcal{D}_2 .

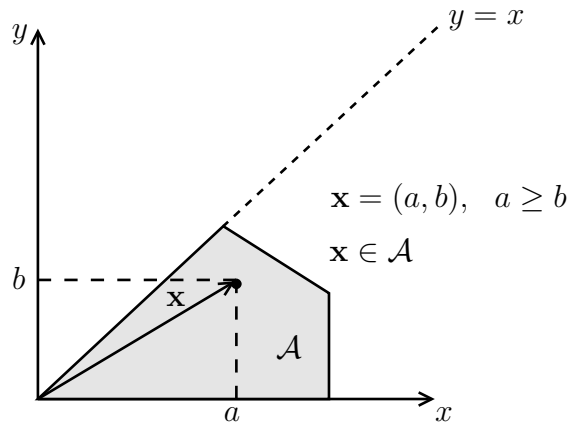


Figura 6: Interpretação geométrica de um conjunto contido em \mathcal{D}_2 .

Notemos que, se $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$ em \mathcal{A} , então $\mathbf{x} = \mathbf{y}$. De fato, como $\mathbf{x} = (x_1, x_2)$ e $\mathbf{y} = (y_1, y_2)$ pertencem ao conjunto \mathcal{A} , tem-se $x_1 \geq x_2$ e $y_1 \geq y_2$. Além disso, por $\mathbf{x} \prec \mathbf{y}$, segue que

$$x_1 \leq y_1, \quad (3.3)$$

$$x_1 + x_2 = y_1 + y_2. \quad (3.4)$$

Por sua vez, por $\mathbf{y} \prec \mathbf{x}$, tem-se

$$y_1 \leq x_1, \quad (3.5)$$

$$y_1 + y_2 = x_1 + x_2. \quad (3.6)$$

Analisando as inequações apresentadas em (3.3) e (3.5), obtém-se $x_1 = y_1$ e, conseqüentemente, da igualdade (3.4) segue que $x_2 = y_2$. Portanto, os vetores \mathbf{x} e \mathbf{y} são iguais.

¹Os detalhes a respeito destes resultados, encontram-se no Apêndice A.

3.1.1 Propriedades básicas

A seguir são listadas algumas propriedades da relação de majorização, as quais são basicamente consequências diretas da definição. Sejam, portanto, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$, $\mathbf{z} = (z_1, z_2, \dots, z_n)$ e $\mathbf{1} = (1, 1, \dots, 1)$ vetores de \mathbb{R}^n , segue que:

(p1): $\mathbf{x} \prec \mathbf{x}$ (reflexividade);

(p2): se $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{z}$, então $\mathbf{x} \prec \mathbf{z}$ (transitividade);

(p3): se $\mathbf{x} \prec \mathbf{y}$, então $k\mathbf{x} \prec k\mathbf{y}$, para todo $k \geq 0$;

(p4): $\Pi_1 \mathbf{x} \prec \mathbf{x} \prec \Pi_2 \mathbf{x}$, para quaisquer que sejam as matrizes de permutação Π_1 e Π_2 ;

(p5): se $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$, então $[\mathbf{x}] = [\mathbf{y}]$;

(p6): $\frac{1}{n} \mathbf{1} \prec (1, 0, \dots, 0)$;

(p7): $\mathbf{x} \prec (X, 0, \dots, 0)$, se $x_i \geq 0$ e $\sum_{i=1}^n x_i = X$;

(p8): $\frac{X}{n} \mathbf{1} \prec \mathbf{x}$, se $x_i \geq 0$ e $\sum_{i=1}^n x_i = X$;

(p9): se $\mathbf{x} \prec \mathbf{y}$, então $\min \{x_i\} \geq \min \{y_i\}$;

(p10) se $\mathbf{x} \prec \mathbf{1}$, então $\mathbf{x} = \mathbf{1}$.

Em (p1) e (p4), por exemplo, tem-se casos triviais de majorização. No caso da propriedade (p2), por hipótese, tem-se $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{z}$, desta forma:

$$\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, \quad 1 \leq k \leq n-1 \quad (3.7)$$

$$\sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}, \quad (3.8)$$

$$\sum_{i=1}^k y_{[i]} \leq \sum_{i=1}^k z_{[i]}, \quad 1 \leq k \leq n-1, \quad (3.9)$$

$$\sum_{i=1}^n y_{[i]} = \sum_{i=1}^n z_{[i]}, \quad (3.10)$$

nota-se, portanto, das desigualdades (3.7) e (3.9) que $\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k z_{[i]}$, para $1 \leq k \leq n-1$, e das igualdades (3.8) e (3.10) que $\sum_{i=1}^n x_{[i]} = \sum_{i=1}^n z_{[i]}$, desta forma, $\mathbf{x} \prec \mathbf{z}$. No caso de (p3), basta multiplicar as expressões (3.7) e (3.8) por $k \geq 0$. Em (p5),

como $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$, tem-se $x_{[i]} = y_{[i]}$, para todo $i = 1, 2, \dots, n$, o que prova o resultado. Para a propriedade (**p6**), usa-se o seguinte artifício (MARSHALL; OLKIN, 1979):

$$\begin{aligned} \frac{1}{n} \mathbf{1} &= \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n} \right) \prec \left(\frac{1}{n-1}, \frac{1}{n-1}, \dots, \frac{1}{n-1}, 0 \right) \\ &\prec \left(\frac{1}{n-2}, \frac{1}{n-2}, \dots, 0, 0 \right) \\ &\vdots \\ &\prec \left(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0, 0 \right) \\ &\prec (1, 0, \dots, 0). \end{aligned} \tag{3.11}$$

Em (**p7**), ao tomar o vetor $[\mathbf{x}] = (x_{[1]}, x_{[2]}, \dots, x_{[n]})$, verifica-se que $\sum_{i=1}^n x_i = \sum_{i=1}^n x_{[i]} = X$. Além disso, $\sum_{i=1}^k x_{[i]} \leq X$, para todo $1 \leq k \leq n-1$. Assim, diante destas considerações conclui-se que $\mathbf{x} \prec (X, 0, \dots, 0)$. A verificação da propriedade (**p8**) encontra-se no Apêndice A. No item (**p9**), por hipótese, sabe-se que $\mathbf{x} \prec \mathbf{y}$. Desta forma, para $k = n-1$, tem-se

$$\sum_{i=1}^{n-1} x_{[i]} \leq \sum_{i=1}^{n-1} y_{[i]}. \tag{3.12}$$

Agora, ao acrescentar o termo $x_{[n]} + y_{[n]}$ a ambos os membros da desigualdade acima, obtém-se

$$y_{[n]} + \sum_{i=1}^n x_{[i]} \leq x_{[n]} + \sum_{i=1}^n y_{[i]}, \tag{3.13}$$

ou seja, $y_{[n]} \leq x_{[n]}$. Noutros termos, $\min \{x_i\} \geq \min \{y_i\}$. Por fim, vem (**p10**). Por hipótese, o vetor \mathbf{x} é majorizado pelo vetor $\mathbf{1}$, então pela propriedade (**p9**) tem-se $x_{[n]} \geq 1$. Por outro lado, por definição, sabe-se que $x_{[1]} \leq 1$. Diante destas considerações, temos que

$$1 \geq x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[k]} \geq \dots \geq x_{[n]} \geq 1, \tag{3.14}$$

ou seja, $x_{[k]} = 1$, para qualquer que seja $k = 1, 2, \dots, n$. Assim, $\mathbf{x} = \mathbf{1}$, o que completa as verificações das propriedades.

Vale ressaltar ainda que, a partir das propriedades básicas já apresentadas, outras relações de majorização são obtidas. Por exemplo, aplicando-se a propriedade de transitividade estudada em (**p2**), nas relações de majorização ilustradas em (**p8**) e (**p7**), nesta ordem, obtém-se um importante resultado, o qual será aplicado na sequência desta

dissertação, a saber:

$$\frac{X}{n} \mathbf{1} \prec \mathbf{x} \prec (X, 0, \dots, 0), \quad \text{se } x_i \geq 0 \text{ e } \sum_{i=1}^n x_i = X. \quad (3.15)$$

O trabalho de Jorswieck e Boche (2007), apresenta a relação de majorização como uma medida de correlação espacial de um canal MIMO com n_T antenas transmissoras. Especificamente, ao serem consideradas duas matrizes arbitrárias de correlação no transmissor \mathbf{R}_T^1 e \mathbf{R}_T^2 , cujos respectivos autovalores são $\lambda_1^{T,1} \geq \lambda_2^{T,1} \geq \dots \geq \lambda_{n_T}^{T,1} \geq 0$ e $\lambda_1^{T,2} \geq \lambda_2^{T,2} \geq \dots \geq \lambda_{n_T}^{T,2} \geq 0$, a matriz de correlação \mathbf{R}_T^1 é considerada mais correlacionada que \mathbf{R}_T^2 , quando o vetor de autovalores de \mathbf{R}_T^1 majoriza o vetor de autovalores de \mathbf{R}_T^2 , isto é,

$$\lambda(\mathbf{R}_T^2) \prec \lambda(\mathbf{R}_T^1), \quad (3.16)$$

sendo $\lambda(\mathbf{R}_T^1) = (\lambda_1^{T,1}, \lambda_2^{T,1}, \dots, \lambda_{n_T}^{T,1})$ e $\lambda(\mathbf{R}_T^2) = (\lambda_1^{T,2}, \lambda_2^{T,2}, \dots, \lambda_{n_T}^{T,2})$.

3.1.2 Matriz Duplamente Estocástica

Segundo Marshall e Olkin (1979), uma matriz real quadrada \mathbf{A} é chamada de matriz *estocástica* quando todos os elementos são não-negativos e a soma dos termos de cada linha é igual a 1. Ademais, se a soma de cada coluna for também igual a 1, esta chama-se *duplamente estocástica*². Portanto, se uma matriz quadrada $\mathbf{A} = [a_{ij}]$ de ordem $n \times n$ é duplamente estocástica tem-se:

$$a_{ij} \geq 0, \quad \forall i, j, \quad (3.17)$$

$$\sum_{i=1}^n a_{ij} = 1, \quad \forall 1 \leq j \leq n, \quad (3.18)$$

$$\sum_{j=1}^n a_{ij} = 1, \quad \forall 1 \leq i \leq n. \quad (3.19)$$

Os teoremas seguintes, os quais são apresentados sem demonstração, ilustram a relação entre majorização de vetores e matriz duplamente estocástica.

Teorema 3.1 (MARSHALL; OLKIN, 1979, Teorema 2.A.4) Uma matriz \mathbf{A} , de ordem $n \times n$, é duplamente estocástica se, e somente se, $\mathbf{A}\mathbf{x} \prec \mathbf{x}$, para todo vetor \mathbf{x} em \mathbb{R}^n .

²A origem da denominação *matriz duplamente estocástica* é comentada em (MARSHALL; OLKIN, 1979, p. 19)

Teorema 3.2 (Hardy, Littlewood e Pólya, 1929) (MARSHALL; OLKIN, 1979, Teorema 2.B.2)

Sejam \mathbf{x} e \mathbf{y} vetores de \mathbb{R}^n . Uma condição necessária e suficiente para $\mathbf{x} \prec \mathbf{y}$ é que existe uma matriz duplamente estocástica \mathbf{A} de ordem $n \times n$, tal que $\mathbf{x} = \mathbf{A}\mathbf{y}$.

O objetivo da inserção da matriz duplamente estocástica à teoria, é simplesmente facilitar o desenvolvimento de outros resultados que virão na sequência. Especificamente, do Teorema 3.1, além de propiciar um critério, o qual nos permite determinar se uma matriz é duplamente estocástica, este também, a partir de uma matriz duplamente conhecida, nos garante a existência de importantes relações de majorização³. Por sua vez, o Teorema 3.2 nos possibilita a substituição de uma relação de majorização por uma igualdade vetorial, sendo que cada coordenada de um vetor é combinação convexa do outro (BOYD; VANDENBERGHE, 2004).

3.1.3 Schur-convexidade e Schur-concavidade

Neste tópico serão apresentadas duas funções reais definidas num conjunto $\mathcal{A} \subset \mathbb{R}^n$, a saber: função Schur-convexa e Schur-côncava, bem como algumas propriedades e exemplos.

Uma função real $f : \mathbb{R} \rightarrow \mathbb{R}$ é dita não-decrescente quando para dois valores quaisquer x_1 e x_2 pertencentes ao domínio da função, tais que $x_1 \leq x_2$, tem-se $f(x_1) \leq f(x_2)$, ou seja, a ordem do sinal de desigualdade permanece. No caso da função ser não-crescente, a ordem do sinal de desigualdade é mudada em relação àquela considerada aos elementos do domínio. De maneira análoga, Schur (1923) generalizou este conceito de preservação de ordem para o caso de uma função ter como domínio um conjunto em que os elementos possam ser comparados via majorização, ou seja, uma função $\varphi : \mathcal{A} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é chamada de Schur-convexa se

$$\mathbf{x} \prec \mathbf{y} \text{ em } \mathcal{A}, \text{ então } \varphi(\mathbf{x}) \leq \varphi(\mathbf{y}). \quad (3.20)$$

Por sua vez, a função $\varphi(\cdot)$ é dita ser Schur-côncava se

$$\mathbf{x} \prec \mathbf{y} \text{ em } \mathcal{A}, \text{ então } \varphi(\mathbf{x}) \geq \varphi(\mathbf{y}). \quad (3.21)$$

Conforme comenta Marshall e Olkin (1979), uma definição mais plausível para esta “preservação” de ordem seria função Schur-crescente e Schur-decrescente, respectivamente.

³Uma aplicação desta resultado encontra-se no Apêndice A.

No entanto, por questões históricas esta definição é preservada. A seguir são apresentados dois exemplos que reforçam o entendimento das definições dadas:

Exemplo 3.1 A função máximo de várias variáveis $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(\mathbf{x}) = \max \{x_1, x_2, \dots, x_n\}$ é Schur-convexa, enquanto que, a função mínimo $g : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $g(\mathbf{x}) = \min \{x_1, x_2, \dots, x_n\}$ é Schur-côncava. De fato, sejam \mathbf{x} e \mathbf{y} dois vetores de \mathbb{R}^n tais que $\mathbf{x} \prec \mathbf{y}$. Por definição tem-se $x_{[1]} \leq y_{[1]}$ e pela propriedade (p9) segue que $y_{[n]} \leq x_{[n]}$. Por outro lado, como $f(\mathbf{x}) = x_{[1]}$, $f(\mathbf{y}) = y_{[1]}$, $g(\mathbf{x}) = x_{[n]}$ e $g(\mathbf{y}) = y_{[n]}$, segue que $f(\mathbf{x}) \leq f(\mathbf{y})$ e $g(\mathbf{y}) \leq g(\mathbf{x})$, ou seja, a função $f(\cdot)$ é Schur-convexa enquanto $g(\cdot)$ é Schur-côncava. \triangleleft

Exemplo 3.2 A função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(\mathbf{x}) = \sum_{i=1}^n x_i$ é tanto Schur-convexa quanto Schur-côncava. De fato, se $\mathbf{x} \prec \mathbf{y}$, tem-se

$$f(\mathbf{x}) = \sum_{i=1}^n x_i = \sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]} = \sum_{i=1}^n y_i = f(\mathbf{y}),$$

portanto, $f(\cdot)$ é Schur-convexa e Schur-côncava. \triangleleft

Agora serão apresentados alguns critérios os quais permitem identificar se uma função é Schur-convexa ou Schur-côncava sem a necessidade da aplicação direta da definição, que por muitas vezes, devido a complexidade dos cálculos, torna inviável tal análise. Portanto, a fim de verificar a aplicabilidade da matriz duplamente estocástica, apresentaremos em detalhes a demonstração do próximo teorema. No entanto, com o propósito de deixar o referido capítulo menos extenso e enfadonho, omitiremos as demonstrações de alguns teoremas⁴.

Teorema 3.3 (Schur, 1923) (MARSHALL; OLKIN, 1979, Proposição 3.C.1) Se $I \subset \mathbb{R}$ é um intervalo e $g : I \rightarrow \mathbb{R}$ uma função convexa, então $\varphi : I^n \rightarrow \mathbb{R}$, definida por

$$\varphi(\mathbf{x}) = \sum_{i=1}^n g(x_i) \tag{3.22}$$

é Schur-convexa. Analogamente, se $g(\cdot)$ é côncava, então $\varphi(\cdot)$ é Schur-côncava.

⁴As demonstrações as quais foram omitidas podem ser encontradas nas referências citadas no início deste capítulo.

Demonstração: Se $\mathbf{x} \prec \mathbf{y}$ em I^n , então pelo Teorema 3.2, existe uma matriz duplamente estocástica \mathbf{A} tal que $\mathbf{x} = \mathbf{A}\mathbf{y}$. Desta forma, tem-se

$$x_i = \sum_{j=1}^n a_{ij}y_j, \text{ para } i = 1, 2, \dots, n, \quad (3.23)$$

sendo, $a_{ij} \geq 0$, $\sum_{i=1}^n a_{ij} = \sum_{j=1}^n a_{ij} = 1$. Como $g(\cdot)$ é uma função convexa e cada coordenada do vetor \mathbf{x} é combinação convexa das coordenadas de \mathbf{y} , então, pela desigualdade de Jensen (BOYD; VANDENBERGHE, 2004), tem-se a seguinte desigualdade

$$g(x_i) = g\left(\sum_{j=1}^n a_{ij}y_j\right) \leq \sum_{j=1}^n a_{ij}g(y_j). \quad (3.24)$$

Por fim, observa-se que $\varphi(\mathbf{x}) \leq \varphi(\mathbf{y})$. De fato,

$$\begin{aligned} \varphi(\mathbf{x}) &= \sum_{i=1}^n g(x_i) \leq \sum_{i=1}^n \sum_{j=1}^n a_{ij}g(y_j) = \sum_{j=1}^n \sum_{i=1}^n a_{ij}g(y_j) \\ &= \sum_{j=1}^n g(y_j) \sum_{i=1}^n a_{ij} = \sum_{j=1}^n g(y_j) \\ &= \varphi(\mathbf{y}). \end{aligned} \quad (3.25)$$

Portanto, a função $\varphi(\cdot)$ é Schur-convexa. \square

Exemplo 3.3 (BHATIA, 1997, Exemplo II.3.11) Uma função bastante estudada na teoria da informação é a função entropia $H : \mathbb{R}_+^n \rightarrow \mathbb{R}$, definida por $H(\mathbf{p}) = -\sum_{i=1}^n p_i \ln p_i$, sendo $\sum_{i=1}^n p_i = 1$. Ao definir a função $g(x) = -x \ln x$, tem-se que esta é côncava e, aplicando-se o teorema anterior, conclui-se que a função entropia é Schur-côncava. Noutros termos, se $\mathbf{x} \prec \mathbf{y}$, então $H(\mathbf{y}) \leq H(\mathbf{x})$. Em particular, pela propriedade apresentada na Equação (3.15), garante-se que $\frac{1}{n} \mathbf{1} \prec \mathbf{p} \prec \mathbf{e}_1$, sendo $\mathbf{e}_1 = (1, 0, \dots, 0)$. Então aplicando-se a função H , obtém-se os seguintes limitantes à entropia:

$$H(\mathbf{e}_1) \leq H(\mathbf{p}) \leq H\left(\frac{1}{n} \mathbf{1}\right).$$

\triangleleft

A seguir, apresentamos as definições de *função simétrica* e *função de classe \mathcal{C}^1* , as quais nos permitirão um melhor entendimento nas propostas sugeridas pelos teoremas que se seguem. Especificamente, sendo $I \subset \mathbb{R}$ um intervalo da reta, então de acordo com Bhatia (1997), uma função $f : I^n \rightarrow \mathbb{R}$ ($n > 1$) é chamada de *simétrica* se para toda matriz permutação $\mathbf{\Pi}$ de ordem $n \times n$ e para todo $\mathbf{x} \in I^n$ tem-se $f(\mathbf{\Pi}\mathbf{x}) = f(\mathbf{x})$. Por

sua vez, a função $f(\cdot)$ é dita de *classe* \mathcal{C}^1 se todas as derivadas parciais existem e são contínuas em I^n (LIMA, 2000).

Teorema 3.4 (MARSHALL; OLKIN, 1979, Proposição 3.C.2) Seja $I \subset \mathbb{R}$ um intervalo. Se $\varphi : I^n \rightarrow \mathbb{R}$ é uma função simétrica e convexa, então $\varphi(\cdot)$ é Schur-convexa. Analogamente, se $\varphi(\cdot)$ é simétrica e côncava, então $\varphi(\cdot)$ é Schur-côncava.

Exemplo 3.4 Para ilustrar o teorema acima, considere a função $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ definida por $f(\mathbf{x}) = \max \{x_1, x_2, \dots, x_n\}$ a qual é convexa (BOYD; VANDENBERGHE, 2004). Então dada uma matriz permutação $\mathbf{\Pi}$, tem-se

$$f(\mathbf{\Pi x}) = \max \{x_{\pi_1}, x_{\pi_2}, \dots, x_{\pi_n}\} = \max \{x_1, x_2, \dots, x_n\} = f(\mathbf{x}).$$

Portanto, $f(\cdot)$ é Schur-convexa. ◁

O teorema seguinte é conhecido como *condição de Schur*, a qual aplica-se no caso da função ser simétrica e possuir todas as derivadas parciais de primeira ordem no ponto dado. Este é considerado como o método mais simples na verificação se uma função é Schur-convexa ou Schur-côncava.

Teorema 3.5 (Condição de Schur) (BHATIA, 1997, Teorema II.3.14) Seja $I \subset \mathbb{R}$ um intervalo e $\varphi : I^n \rightarrow \mathbb{R}$ uma função simétrica de classe \mathcal{C}^1 . Então $\varphi(\cdot)$ é Schur-convexa se, e somente se,

$$(x_i - x_j) \left(\frac{\partial \varphi}{\partial x_i} - \frac{\partial \varphi}{\partial x_j} \right) \geq 0, \text{ para todo, } i, j \text{ com } 1 \leq i, j \leq n. \quad (3.26)$$

Por sua vez, $\varphi(\cdot)$ é Schur-côncava se, e somente se,

$$(x_i - x_j) \left(\frac{\partial \varphi}{\partial x_i} - \frac{\partial \varphi}{\partial x_j} \right) \leq 0, \text{ para todo, } i, j \text{ com } 1 \leq i, j \leq n. \quad (3.27)$$

Em particular, os resultados acima são válidos para $i = 2$ e $j = 1$. Desta forma, para que a função $\varphi(\cdot)$ seja Schur-convexa, basta verificar se $(x_2 - x_1) \left(\frac{\partial \varphi}{\partial x_2} - \frac{\partial \varphi}{\partial x_1} \right) \geq 0$. No caso das coordenadas dos vetores estarem em ordem decrescente, esta verificação resume-se a $\left(\frac{\partial \varphi}{\partial x_2} - \frac{\partial \varphi}{\partial x_1} \right) \geq 0$.

Exemplo 3.5 No contexto de comunicação sem fio, algumas funções são bastante estudadas em problemas de alocação de recursos para sistemas de múltiplos usuários ou sistemas de múltiplas antenas. Pode-se citar, como exemplo, a função $C : \mathcal{D}_n \rightarrow \mathbb{R}$ definida por $C(\mathbf{p}) = \ln(1 + \rho \sum_{i=1}^n p_i \alpha_i)$, com $\alpha \in \mathcal{D}_n$ e $\rho > 0$ (JORSWIECK; BOCHE, 2007). Di-

ante das hipóteses levantadas, tal função é Schur-convexa. De fato, tomando-se $\mathbf{\Pi}$, uma matriz de permutação qualquer, tem-se

$$C(\mathbf{\Pi p}) = \ln \left(1 + \sum_{i=1}^n p_{\pi_i} \right) = \ln \left(1 + \sum_{i=1}^n p_i \right) = C(\mathbf{p})$$

ou seja, $C(\cdot)$ é uma função simétrica. Além disso, como \mathbf{p} e $\boldsymbol{\alpha}$ pertencem ao conjunto \mathcal{D}_n , segue que $p_1 \geq p_2 \geq \dots \geq p_n \geq 0$ e $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq 0$ e fazendo-se os referidos cálculos chega-se em

$$\frac{\partial C}{\partial p_2} - \frac{\partial C}{\partial p_1} = \frac{\rho \alpha_2}{1 + \rho \sum_{i=1}^n p_i \alpha_i} - \frac{\rho \alpha_1}{1 + \rho \sum_{i=1}^n p_i \alpha_i} \leq 0,$$

o que mostra que a função $C(\cdot)$ é Schur-côncava. \triangleleft

Um outro critério, para caracterização de funções Schur-côncava e Schur-convexa, envolvendo diferenciabilidade de funções é descrito pelo seguinte teorema.

Teorema 3.6 (MARSHALL; OLKIN, 1979, Proposição 3.H.2) Seja $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ uma função definida por $\varphi(\mathbf{x}) = \sum_{i=1}^n g_i(x_i)$, sendo cada função $g_i : \mathbb{R} \rightarrow \mathbb{R}$ diferenciável. Então, $\varphi(\cdot)$ é Schur-convexa sobre \mathcal{D} se, e somente se,

$$g'_i(a) \geq g'_{i+1}(b), \quad \text{sempre que, } a \geq b, \quad \text{com } i = 1, 2, \dots, n-1. \quad (3.28)$$

Analogamente, a função $\varphi(\cdot)$ é Schur-côncava sobre o conjunto \mathcal{D} se, e somente se, $g'_i(a) \leq g'_{i+1}(b)$, sempre que, $a \geq b$.

Exemplo 3.6 Uma aplicação direta do Teorema 3.6 é encontrada em (PALOMAR; CIOFFI; LAGUNAS, 2003). O referido trabalho investiga um problema de alocação de potência robusta e, além disso, apresenta a seguinte função real $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ definida por $\varphi(\mathbf{x}) = -\sum_{i=1}^n \ln(1 + x_i \alpha_i)$, com $0 \leq \alpha_i \leq \alpha_{i+1}$, a qual é Schur-convexa. De fato, sejam a e b números reais, tais que $a \geq b$. Agora considere a função real $g_i(x) = -\ln(1 + \alpha_i x)$, para $i = 1, 2, \dots, n-1$. Então, derivando-se a função $g_i(\cdot)$ nos pontos a e b , tem-se

$$g'_i(a) = -\frac{\alpha_i}{1 + \alpha_i a} \quad \text{e} \quad g'_{i+1}(b) = -\frac{\alpha_{i+1}}{1 + \alpha_{i+1} b}. \quad (3.29)$$

Por sua vez, como $a \geq b$ e $\alpha_i \geq \alpha_{i+1}$, segue que

$$\frac{\alpha_{i+1}}{1 + \alpha_{i+1} b} \geq \frac{\alpha_i}{1 + \alpha_i a}, \quad (3.30)$$

isto é, $g'_i(a) \geq g'_{i+1}(b)$, o que comprova o resultado. \triangleleft

Outros resultados interessantes com relação a Schur-convexidade e a Schur-concavidade envolvem a composição de funções. Na literatura, alguns problemas de otimização fazem uso deste recurso, principalmente, aplicados à função custo (PALOMAR; JIANG, 2006; JIANG; PALOMAR; VARANASI, 2007; BERGMAN *et al.*, 2008; PALOMAR, 2005). Vejamos os casos que aparecem com maior frequência nos problemas de comunicação sem fio⁵. Seja, portanto, o diagrama abaixo:

$$\mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{h} \mathbb{R}$$

$$\psi = h \circ f$$

sendo, $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$ e $f_i : \mathbb{R}^m \rightarrow \mathbb{R}$, $i = 1, 2, \dots, n$. Assim, ao definir $\psi(\mathbf{x}) = h(f(\mathbf{x})) = h(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$, em que cada função $f_i(\cdot)$ é Schur-convexa, e $h : \mathbb{R}^n \rightarrow \mathbb{R}$ é crescente⁶, tem-se que a função $\psi(\cdot)$ preserva a Schur-convexidade. De fato, sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, tais que $\mathbf{x} \prec \mathbf{y}$. Como cada função $f_i(\cdot)$ é Schur-convexa, segue que $f_i(\mathbf{x}) \leq f_i(\mathbf{y})$. Por sua vez, $h(\cdot)$ é crescente, então, diante desta hipótese, conclui-se que $h(f(\mathbf{x})) \leq h(f(\mathbf{y}))$, isto é, $\psi(\mathbf{x}) \leq \psi(\mathbf{y})$. Noutros termos, $\psi(\cdot)$ é Schur-convexa.

Além deste resultado, uma outra composição importante considerada é quando cada função $f_i(\mathbf{x}) = g(x_i)$, sendo $g : \mathbb{R} \rightarrow \mathbb{R}$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$, conforme ilustra o diagrama a seguir.

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n \xrightarrow{h} \mathbb{R}$$

$$\psi = h \circ f$$

Contudo, se $h(\cdot)$ for Schur-convexa e crescente e $g(\cdot)$ for convexa, então a função composta $\psi(\cdot)$ também é Schur-convexa⁷.

Exemplo 3.7 Jiang, Palomar e Varanasi (2007) apresentam a seguinte situação: sejam as funções $h : \mathbb{R}^n \rightarrow \mathbb{R}$ Schur-convexa e crescente e $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$, definida por $\psi(\mathbf{x}) = h(e^{x_1}, e^{x_2}, \dots, e^{x_n})$. Diante destas hipóteses e do fato que a função $g(x) = e^x$ é convexa, garante-se que a função $\psi(\cdot)$ também é Schur-convexa. \triangleleft

Neste tópico, foram apresentados alguns critérios que permitem caracterizar se uma função é Schur-convexa ou Schur-côncava. No entanto, estes critérios não determinam se a referida função é convexa ou côncava, pois convexidade e Schur-convexidade, por exemplo,

⁵A referência (MARSHALL; OLKIN, 1979) apresenta uma lista detalhada, com várias situações, envolvendo composição de funções.

⁶Uma função $h : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita ser *crescente*, se $\mathbf{a} \leq \mathbf{b}$ em \mathbb{R}^n (ou $a_i \leq b_i$ para cada $i = 1, 2, \dots, n$), implicar em $h(\mathbf{a}) \leq h(\mathbf{b})$ (MARSHALL; OLKIN, 1979, 16.A).

⁷A verificação desta afirmação encontra-se no Apêndice A.

não possuem definições equivalentes. A seguir é apresentada uma função Schur-convexa que não é convexa.

Exemplo 3.8 (BHATIA, 1997, Exemplo II.3.15) Seja $I = (0, 1)$ um intervalo aberto e sejam as funções $f : I \rightarrow \mathbb{R}$, dada por $f(x) = \ln\left(\frac{1}{x} - 1\right)$, e $\Phi : \mathcal{A} \rightarrow \mathbb{R}$, definida por

$$\Phi(x_1, x_2) = \ln\left(\frac{1}{x_1} - 1\right) + \ln\left(\frac{1}{x_2} - 1\right),$$

sendo $\mathcal{A} = \{(x_1, x_2) \in I^2 ; x_1 + x_2 \leq 1\}$. A função $f(\cdot)$ é convexa no intervalo $I_1 = (0, \frac{1}{2}]$ e não-convexa em $I_2 = [\frac{1}{2}, 1)$. De fato,

$$f'(x) = (x^2 - x)^{-1} \quad \text{e} \quad f''(x) = -\left(\frac{1}{(x^2 - x)^2} \cdot \frac{1}{2x - 1}\right).$$

Em I_1 tem-se que a função $f(\cdot)$ é convexa, pois $f''(x) \geq 0$ e no intervalo I_2 , a função $f(\cdot)$ é côncava (não-convexa), já que neste caso, $f''(x) \leq 0$. Diante destas considerações, a função $f(\cdot)$ não é convexa no intervalo I . Consequentemente, $\Phi(\cdot)$ também não é convexa em \mathcal{A} . De fato, dados $a, b \in I_2$, tais que $a + b \leq 1$. Note que o ponto (a, b) pertence ao conjunto \mathcal{A} . No entanto, $\Phi(\cdot)$ é soma de duas funções côncavas (de mesmo domínio). Desta forma, garante-se que $\Phi(\cdot)$ é uma função côncava (BOYD; VANDENBERGHE, 2004), isto é, a referida função é não-convexa.

Por outro lado, $\Phi(x_1, x_2) = \Phi(x_2, x_1)$, ou seja, $\Phi(\cdot)$ é simétrica. De fato, aplicando-se o critério de Schur (3.5), tem-se

$$(x_2 - x_1) \left(\frac{\partial \Phi}{\partial x_2} - \frac{\partial \Phi}{\partial x_1} \right) = \frac{(x_2 - x_1)^2 [1 - (x_1 + x_2)]}{x_1 x_2 (1 - x_1)(1 - x_2)} \geq 0,$$

ou seja, $\Phi(\cdot)$ é Schur-convexa. Portanto, $\Phi(\cdot)$ é uma função Schur-convexa e não-convexa.

◁

Este capítulo apresentou definições e resultados relativos à teoria de majorização diante de um aspecto geral. Contudo, as próximas seções serão destinadas à aplicação destes resultados à alguns problemas de álgebra linear, otimização e variáveis aleatórias, com a finalidade de apresentar outros resultados que serão fundamentais no desenvolvimento desta obra.

3.2 Majorização e Álgebra Linear

O trabalho de Issai Schur sobre majorização foi motivado pela descoberta de que os valores singulares de uma matriz hermitiana e semi-definida positiva majorizam os elementos da diagonal principal (MARSHALL; OLKIN, 1979). Esta descoberta, portanto, forneceu uma nova abordagem a *desigualdade de Hadamard*. Além desta, outras importantes desigualdades, envolvendo valores singulares e teoria das matrizes de uma forma geral, surgiram no desenvolvimento da teoria, principalmente, com o auxílio da álgebra linear.

Esta seção dará ênfase a resultados que se aplicam aos sistemas de comunicação sem fio, principalmente, naqueles que frequentemente aparecem no estudo de capacidade de canal (JORSWIECK; BOCHE, 2007; ZHONG; WONG; JIN, 2009). Portanto, por questão de conveniência, ao tomar uma matriz hermitiana $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{n \times n}$, serão considerados dois vetores, a saber: $\lambda(\mathbf{H}) = (\lambda_1, \lambda_2, \dots, \lambda_n)$ e $d(\mathbf{H}) = (h_{11}, h_{22}, \dots, h_{nn})$. As coordenadas do primeiro vetor, representam os autovalores de \mathbf{H} em ordem decrescente, isto é, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, enquanto as coordenadas do segundo são as entradas da diagonal principal de \mathbf{H} .

Teorema 3.7 (Inequação de Schur) (MARSHALL; OLKIN, 1979, Teorema 16.B.1) Se $\mathbf{H} \in \mathbb{C}^{n \times n}$ uma matriz hermitiana, então o vetor $d(\mathbf{H})$ é majorizado por $\lambda(\mathbf{H})$, isto é,

$$d(\mathbf{H}) \prec \lambda(\mathbf{H}). \quad (3.31)$$

Demonstração: Como \mathbf{H} é uma matriz hermitiana, existe uma matriz unitária $\mathbf{U} = [u_{ij}]$ de ordem n tal que $\mathbf{H} = \mathbf{U}\mathbf{D}_\lambda\mathbf{U}^H$, sendo \mathbf{D}_λ uma matriz diagonal cujos elementos da diagonal principal são os autovalores de \mathbf{H} , em ordem decrescente, isto é, $\mathbf{D}_\lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ (HORN; JOHNSON, 1985). Os elementos da diagonal principal da matriz hermitiana são dados por

$$h_{ii} = \sum_{j=1}^n |u_{ij}|^2 \lambda_j, \quad i = 1, 2, \dots, n. \quad (3.32)$$

Considerando a matriz $\mathbf{P} = [|u_{ij}|^2] \in \mathbb{R}_+^{n \times n}$, tem-se que esta matriz é duplamente estocástica. Além disso, os vetores $d(\mathbf{H})$ e $\lambda(\mathbf{H})$ estão relacionados pelo seguinte produto matricial:

$$d(\mathbf{H}) = \mathbf{P} \lambda(\mathbf{H}). \quad (3.33)$$

Então, pelo Teorema (3.2)

$$d(\mathbf{H}) \prec \lambda(\mathbf{H}). \quad (3.34)$$

□

No desenvolvimento da proposição acima, foi visto que $\mathbf{H} = \mathbf{U}\mathbf{D}_\lambda\mathbf{U}^H$, desta forma, os traços das matrizes \mathbf{H} e \mathbf{D}_λ são iguais, ou seja, $\text{tr}(\mathbf{H}) = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n h_{ii}$. Diante deste resultado, e da hipótese de que $h_{ii} \geq 0$, algumas consequências diretas são retiradas do teorema *Inequação de Schur* e de propriedades estudadas anteriormente. Vejamos, portanto, algumas situações (PALOMAR; JIANG, 2006):

$$(c1) \quad \frac{\text{tr}(\mathbf{H})}{n} \mathbf{1} \prec d(\mathbf{H}) \prec \lambda(\mathbf{H});$$

$$(c2) \quad \frac{\text{tr}(\mathbf{H})}{n} \mathbf{1} \prec d(\mathbf{U}^H\mathbf{H}\mathbf{U}) = \lambda(\mathbf{H}).$$

Das relações apresentadas em (c1), a relação de majorização $\frac{\text{tr}(\mathbf{H})}{n} \mathbf{1} \prec d(\mathbf{H})$ é devido a propriedade (p8) vista na Seção 3.1, enquanto a outra, é a desigualdade de Schur. Por sua vez, para (c2) apliquemos o resultado estudado em (c1). Para tal, observemos que $\mathbf{H} = \mathbf{U}\mathbf{D}_\lambda\mathbf{U}^H$ (ver demonstração do Teorema 3.7). Assim, por (c1) obtém-se

$$\frac{\text{tr}(\mathbf{U}^H\mathbf{H}\mathbf{U})}{n} \mathbf{1} \prec d(\mathbf{U}^H\mathbf{H}\mathbf{U}). \quad (3.35)$$

No entanto, $\text{tr}(\mathbf{U}^H\mathbf{H}\mathbf{U}) = \text{tr}(\mathbf{H})$ e $d(\mathbf{U}^H\mathbf{H}\mathbf{U}) = d(\mathbf{D}_\lambda) = \lambda(\mathbf{H})$. Assim,

$$\frac{\text{tr}(\mathbf{H})}{n} \mathbf{1} \prec d(\mathbf{U}^H\mathbf{H}\mathbf{U}) = \lambda(\mathbf{H}). \quad (3.36)$$

A próxima proposição trata-se do teorema chamado *Inequação de Hadamard* a qual é uma importante ferramenta no estudo de limitantes para capacidade de canais MIMO:

Teorema 3.8 (Inequação de Hadamard) (BHATIA, 1997, Teorema II.3.17)

Se $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{n \times n}$ é uma matriz hermitiana semi-definida positiva, então

$$\det(\mathbf{H}) \leq \prod_{i=1}^n h_{ii}. \quad (3.37)$$

Demonstração: Considerando a função convexa $g(x) = -\ln x$ e $\Phi(\mathbf{x}) = \sum_{i=1}^n g(x_i)$, tem-se, pelo Teorema de Schur 3.3, que a função $\Phi(\cdot)$ é Schur-convexa. Além disso, pela inequação de Schur (Teorema 3.7) sabe-se que $d(\mathbf{H}) \prec \lambda(\mathbf{H})$. Então, diante destas considerações, obtém-se

$$\Phi(d(\mathbf{H})) \leq \Phi(\lambda(\mathbf{H})). \quad (3.38)$$

No entanto,

$$\Phi(d(\mathbf{H})) = -\sum_{i=1}^n \ln(h_{ii}) = -\ln\left(\prod_{i=1}^n h_{ii}\right), \quad (3.39)$$

$$\Phi(\lambda(\mathbf{H})) = -\sum_{i=1}^n \ln(\lambda_i) = -\ln\left(\prod_{i=1}^n \lambda_i\right). \quad (3.40)$$

Desta forma, por substituição de $\Phi(d(\mathbf{H}))$ e $\Phi(\lambda(\mathbf{H}))$ na desigualdade em (3.38), segue que

$$\ln\left(\prod_{i=1}^n \lambda_i\right) \leq \ln\left(\prod_{i=1}^n h_{ii}\right). \quad (3.41)$$

Portanto,

$$\det(\mathbf{H}) = \prod_{i=1}^n \lambda_i \leq \prod_{i=1}^n h_{ii}. \quad (3.42)$$

o que prova o teorema. \square

Corolário 3.1 (BHATIA, 1997, Exercício II.3.18) Se $\mathbf{A} \in \mathbb{C}^{m \times n}$ é uma matriz complexa, então

$$\det(\mathbf{A}\mathbf{A}^H) \leq \prod_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \quad (3.43)$$

Demonstração: Fazendo-se $\mathbf{H} = \mathbf{A}\mathbf{A}^H \in \mathbb{C}^{m \times m}$, tem-se que esta matriz é hermitiana e semi-definida positiva (HORN; JOHNSON, 1985). Então, pelo teorema anterior, tem-se $\det(\mathbf{H}) \leq \prod_{i=1}^m h_{ii}$. Além disso,

$$h_{ii} = \sum_{j=1}^n |a_{ij}|^2, \text{ para } i = 1, 2, \dots, m. \quad (3.44)$$

Por outro lado, $\det(\mathbf{H}) = \det(\mathbf{A}\mathbf{A}^H) \leq \prod_{i=1}^m h_{ii}$. Fazendo-se, portanto, a devida substituição chega-se a

$$\det(\mathbf{A}\mathbf{A}^H) \leq \prod_{i=1}^m \sum_{j=1}^n |a_{ij}|^2. \quad (3.45)$$

\square

A *inequação de Hadamard* é bastante encontrada na literatura em se tratando de obtenção de limitantes à capacidade de canal. Por exemplo, o trabalho de Zarbouti, Tsoulos e Kaklamani (2006) obtém limitantes de capacidade de um canal MIMO com desvanecimento de Rayleigh diante de algumas restrições.

Contudo, a *inequação de Schur* é o resultado central desta seção, pois por esta relação de majorização e por funções Schur-côncavas convenientemente definidas será possível a

obtenção de limitantes e aproximações à capacidade de um canal MIMO com desvanecimento de Rice que é a proposta deste trabalho.

A próxima seção trata de alguns problemas envolvendo teoria de majorização e otimização. No entanto, mesmo não sendo utilizada ao longo deste trabalho, a referida seção tem o propósito de apresentar a relação entre as duas teorias, como também uma contribuição à literatura no assunto em questão, uma vez que novos teoremas são propostos. Ressalta-se ainda que este tema é de relevante importância em problemas de comunicações móveis (PALOMAR; JIANG, 2006; PALOMAR; CIOFFI; LAGUNAS, 2003; PALOMAR; LAGUNAS; CIOFFI, 2004; PALOMAR, 2005).

3.3 Majorização e Otimização

Foi mencionado, na introdução deste capítulo, que a teoria de majorização é uma importante ferramenta na resolução de problemas de otimização aplicadas a sistemas de comunicação sem fio. Nesta seção, serão apresentados alguns resultados relativos à otimização via função Schur-côncava e Schur-convexa (JORSWIECK; BOCHE, 2007; PALOMAR; CIOFFI; LAGUNAS, 2003; PALOMAR, 2005).

Teorema 3.9 (JORSWIECK; BOCHE, 2007, Teorema 2.21) Seja a função Schur-côncava $f : \mathcal{D}_+ \rightarrow \mathbb{R}$ e considere o seguinte problema de otimização

$$\begin{aligned} & \max_{\mathbf{x} \in \mathcal{D}_+} f(\mathbf{x}) \\ & \text{sujeito à } \sum_{i=1}^n x_i = P. \end{aligned} \tag{3.46}$$

Então, o ponto $\bar{\mathbf{x}} = \frac{P}{n} \mathbf{1}$ é solução para o problema. Por sua vez, no caso do problema de minimização

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{D}_+} f(\mathbf{x}) \\ & \text{sujeito à } \sum_{i=1}^n x_i = P \end{aligned} \tag{3.47}$$

o ponto ótimo é dado por $\underline{\mathbf{x}} = (P, 0, \dots, 0)$.

Demonstração: No problema de maximização, o ponto $\bar{\mathbf{x}}$ satisfaz a restrição dada, pois $\sum_{i=1}^n \bar{x}_i = P$. Além disso, pela propriedade (p8) da Seção 3.1, tem-se $\bar{\mathbf{x}} \prec \mathbf{x}$, para todo $\mathbf{x} \in \mathcal{D}_+$, com $\sum_{i=1}^n x_i = P$. Portanto, como a função $f(\cdot)$ é Schur-côncava tem-se $f(\bar{\mathbf{x}}) \geq f(\mathbf{x})$, ou seja, $\bar{\mathbf{x}}$ é ponto ótimo.

No caso do problema de minimização, devido a propriedade (p7), tem-se $\mathbf{x} \prec \underline{\mathbf{x}}$, para todo $\mathbf{x} \in \mathcal{D}_+$, e novamente, como $f(\cdot)$ é Schur-côncava, segue que $\bar{\mathbf{x}}$ é ponto de mínimo, pois $f(\mathbf{x}) \geq f(\underline{\mathbf{x}})$. Em resumo, para todo $\mathbf{x} \in \mathcal{D}_+$ satisfazendo a restrição dada, tem-se

$$f(\underline{\mathbf{x}}) \leq f(\mathbf{x}) \leq f(\bar{\mathbf{x}}), \quad (3.48)$$

ou seja, $\underline{\mathbf{x}}$ é ponto ótimo para o problema de minimização, enquanto $\bar{\mathbf{x}}$ é ponto ótimo para o problema de maximização. \square

Se a restrição do problema de maximização de (3.46) for relaxada para a condição $\sum_{i=1}^n x_i \leq P$, o problema de otimização terá mesma solução se a função objetivo for crescente⁸. O seguinte teorema formaliza esta afirmação.

Teorema 3.10 Seja a função Schur-côncava e crescente $f : \mathcal{D}_+ \rightarrow \mathbb{R}$ e considere o seguinte problema de otimização

$$\begin{aligned} & \max_{\mathbf{x} \in \mathcal{D}_+} f(\mathbf{x}) \\ & \text{sujeito à } \sum_{i=1}^n x_i \leq P. \end{aligned} \quad (3.49)$$

Portanto, o ponto $\bar{\mathbf{x}} = \frac{P}{n}\mathbf{1}$ é ótimo.

Demonstração: Pelo teorema anterior, para todo $\mathbf{x} \in \mathcal{D}_+$ tal que $\sum_{i=1}^n x_i = P$, tem-se $f(\mathbf{x}) \leq f(\bar{\mathbf{x}})$. Agora, seja $\mathbf{a} \in \mathcal{D}_+$ tal que $\sum_{i=1}^n a_i < P$, assim existe $k > 0$, tal que $\sum_{i=1}^n a_i + k = P$. Considerando o vetor $\mathbf{b} = (a_1 + k, a_2, a_3, \dots, a_n)$, tem-se que $\mathbf{a} \leq \mathbf{b}$ e como a função objetivo é crescente, segue que $f(\mathbf{a}) \leq f(\mathbf{b})$. Por fim, como $f(\mathbf{b}) \leq f(\bar{\mathbf{x}})$, conclui-se que $f(\mathbf{a}) \leq f(\bar{\mathbf{x}})$, ou seja, $\bar{\mathbf{x}}$ é ponto ótimo. \square

Analogamente, para o problema de minimização, tem-se o seguinte resultado

⁸Ver definição de função crescente em \mathbb{R}^n na referência (MARSHALL; OLKIN, 1979, 16.A)

Teorema 3.11 Seja a função Schur-côncava e decrescente $f : \mathcal{D}_+ \rightarrow \mathbb{R}$ e considere o seguinte problema de otimização

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{D}_+} f(\mathbf{x}) \\ \text{sujeito à } \sum_{i=1}^n x_i \leq P. \end{aligned} \quad (3.50)$$

O ponto ótimo, portanto, é dado por $\underline{\mathbf{x}} = (P, 0, \dots, 0)$.

Demonstração: Novamente pelo Teorema 3.10, sabe-se que para todo $\mathbf{x} \in \mathcal{D}_+$, tal que $\sum_{i=1}^n x_i = P$, tem-se $f(\underline{\mathbf{x}}) \leq f(\mathbf{x})$. Tomando-se um ponto arbitrário $\mathbf{a} \in \mathcal{D}_+$, tal que $\sum_{i=1}^n a_i < P$, então existe $k > 0$, o qual nos assegura $\sum_{i=1}^n a_i + k = P$. Considerando agora, o vetor $\mathbf{b} = (a_1 + k, a_2, a_3, \dots, a_n)$, segue então que $\mathbf{a} \leq \mathbf{b}$, e como a função objetivo é decrescente, tem-se $f(\mathbf{b}) \leq f(\mathbf{a})$. Por sua vez, $f(\underline{\mathbf{x}}) \leq f(\mathbf{b})$, isto devido ao fato de que \mathbf{b} é ponto viável ao problema. Diante destas considerações, conclui-se $f(\underline{\mathbf{x}}) \leq f(\mathbf{a})$. Logo, para todo $\mathbf{x} \in \mathcal{D}_+$, tem-se $f(\underline{\mathbf{x}}) \leq f(\mathbf{x})$, isto é, $\underline{\mathbf{x}}$ é ponto ótimo do referido problema. \square

Exemplo 3.9 Num problema de alocação de potência de um canal MIMO, Palomar, Cioffi e Lagunas (2003) apresentam o seguinte problema de otimização:

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{D}_+} f(\mathbf{x}) = \sum_{i=1}^n \ln(1 + x_i \alpha_i), \quad \text{com } 0 \leq \alpha_i \leq \alpha_{i+1} \\ \text{sujeito à } \sum_{i=1}^n x_i \leq P. \end{aligned} \quad (3.51)$$

Este é o típico problema proposto pelo Teorema 3.10. A verificação que a função objetivo é crescente é uma tarefa relativamente simples. Além disso, ao considerar $g_i(x) = \ln(1 + k_i x)$, com $k_i > 0$, tal função é côncava para $i = 1, 2, \dots, n$ (ver Exemplo 3.6). Então, pelo Teorema 3.6, a função $f(\cdot)$ é Schur-côncava, logo o ponto $\bar{\mathbf{x}} = \frac{P}{n} \mathbf{1}$ é solução (uniforme) do problema e o valor máximo da função é $f(\frac{P}{n} \mathbf{1}) = \sum_{i=1}^n \ln(1 + \frac{1}{n} \alpha_i)$. \triangleleft

Na próxima seção, serão definidas funções Schur-côncavas e Schur-convexas com relação à vetores de variáveis aleatórias, as quais serão fundamentais no andamento deste trabalho. Além disso, como forma motivação e entendimento do método, serão obtidos limitantes à capacidade ergódica de um determinado canal MIMO.

3.4 Majorização Estocástica

Quando se estuda capacidade de um canal, na verdade, se está interessado na quantidade máxima de informação que pode ser transmitida por um canal. No entanto, este canal pode ser determinístico ou estocástico conforme visto no capítulo anterior. No caso do canal ser de natureza estocástica, os resultados sobre capacidade ergódica apresentam o operador esperança $\mathbb{E}\{\cdot\}$ na sua composição. Portanto, é neste sentido que se insere a *majorização estocástica*, a qual é uma extensão da teoria de majorização, porém aplicada a problemas envolvendo entes estatísticos como variável aleatória e média. No entanto, este trabalho dará ênfase a apenas um caso específico desta teoria, que consiste na análise de Schur-concavidade de funções do tipo $G : \mathcal{D}_+ \subset \mathbb{R}^k \rightarrow \mathbb{R}$ definidas por

$$G(\mathbf{w}) = \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_i w_i] \right\} \quad (3.52)$$

sendo $\alpha \geq 0$, $\beta_i \geq 0$ constantes reais e w_1, w_2, \dots, w_k variáveis aleatórias de mesma distribuição de probabilidade. Maiores detalhes a respeito de majorização estocástica encontram-se em (MARSHALL; OLKIN, 1979; JORSWIECK; BOCHE, 2007).

O interesse com relação à função da Equação (3.52) reside no fato de que esta se apresenta em cenários práticos relevantes nos sistemas de comunicação sem fio. Por exemplo: o caso do transmissor não ter informações sobre o estado do canal (CSI, do inglês) e no caso do transmissor conhecer perfeitamente o estado do canal. Ademais, conforme estudado no capítulo anterior, estes processos resultam em equações distintas para a capacidade ergódica. Admitindo-se, portanto, o modelo de canal MIMO da Equação (2.4), as capacidades ergódicas são dadas por:

$$\bar{C}_{\text{sCSI}} = \mathbb{E} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{P_T}{\sigma^2 n_T} \lambda_i \right] \right\} \quad (3.53)$$

$$\bar{C}_{\text{pCSI}} = \mathbb{E} \left\{ \sum_{i=1}^r \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T} \lambda_i \right] \right\} \quad (3.54)$$

em que, \bar{C}_{sCSI} e \bar{C}_{pCSI} são as respectivas capacidades do canal sem CSI e com CSI.

O teorema seguinte, formaliza as considerações feitas para Equação (3.52).

Teorema 3.12 Seja a função $G : \mathcal{D}_+ \subset \mathbb{R}^k \rightarrow \mathbb{R}$ definida por

$$G(\mathbf{w}) = \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_i w_i] \right\}. \quad (3.55)$$

Se w_1, w_2, \dots, w_k são variáveis aleatórias identicamente distribuídas, com $\alpha \geq 0$ e $\beta_1 \geq \beta_2 \geq \dots \geq \beta_k \geq 0$, então esta função é Schur-côncava.

Demonstração: A verificação deste resultado consiste numa aplicação direta da condição Schur (Teorema 3.5). Sejam, portanto, $\mathbf{\Pi}$ uma matriz permutação e os vetores $\mathbf{w} = (w_1, w_2, \dots, w_k)$ e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)$, assim

$$\mathbf{\Pi w} = (w_{\pi_1}, w_{\pi_2}, \dots, w_{\pi_k}), \quad \mathbf{\Pi \beta} = (\beta_{\pi_1}, \beta_{\pi_2}, \dots, \beta_{\pi_k}) \quad \text{e} \quad (3.56)$$

$$G(\mathbf{\Pi w}) = \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_i w_{\pi_i}] \right\}. \quad (3.57)$$

Como as variáveis são identicamente distribuídas, tem-se

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_i w_{\pi_i}] \right\} &= \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_{\pi_i} w_{\pi_i}] \right\} \\ &= \mathbb{E} \left\{ \sum_{i=1}^k \log_2 [1 + \alpha \beta_i w_i] \right\} \\ &= G(\mathbf{w}), \end{aligned} \quad (3.58)$$

ou seja, a função $G(\cdot)$ é simétrica. Por outro lado,

$$\frac{\partial G}{\partial w_2} - \frac{\partial G}{\partial w_1} = \frac{1}{\ln 2} \cdot (\beta_2 - \beta_1) \cdot \mathbb{E} \left\{ \frac{1}{1 + \alpha \beta_2 w_2} - \frac{1}{1 + \alpha \beta_1 w_1} \right\}, \quad (3.59)$$

como $\beta_2 - \beta_1 \leq 0$ e $\mathbb{E} \left\{ \frac{1}{1 + \alpha \beta_2 w_2} - \frac{1}{1 + \alpha \beta_1 w_1} \right\} \geq 0$, segue que $\frac{\partial G}{\partial w_2} - \frac{\partial G}{\partial w_1} \leq 0$. Portanto, pelo critério de Schur, a função $G(\cdot)$ é Schur-côncava. \square

Em particular, no caso do canal sem CSI, considerando-se a capacidade ergódica dada

na Equação (3.53) como uma função real sobre \mathcal{D}_+ , isto é, $\bar{C} : \mathcal{D}_+ \rightarrow \mathbb{R}$ tem-se que $\bar{C}(\cdot)$ é uma função Schur-côncava. Noutros termos, se $\mathbf{x} \prec \mathbf{y}$, então $\bar{C}(\mathbf{y}) \leq \bar{C}(\mathbf{x})$.

O próximo exemplo esclarece a aplicação do Teorema 3.12 com a obtenção de limitantes à capacidade ergódica, no caso do transmissor não ter informações sobre o estado do canal, mediante algumas restrições.

Exemplo 3.10 Seja $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ a matriz do canal MIMO dado na Equação (2.4). Além disso, considerando-se $r = n_R = n_T$ e que $\lambda_1, \lambda_2, \dots, \lambda_r$ são os autovalores da matriz $\mathbf{H}\mathbf{H}^H$, tais que $\sum_{i=1}^r \lambda_i = r$, então diante da função $\bar{C} : \mathcal{D}_+ \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e das relações de majorizações apresentadas em (3.15), que garante $\mathbf{1} \prec (\lambda_1, \lambda_2, \dots, \lambda_r) \prec (r, 0, \dots, 0)$, tem-se

$$\bar{C}(r, 0, \dots, 0) \leq \bar{C}(\lambda_1, \lambda_2, \dots, \lambda_r) \leq \bar{C}(1, 1, \dots, 1)$$

ou seja,

$$\log_2 \left(1 + \frac{P_T}{\sigma^2} \right) \leq \bar{C}_{\text{sCSI}} \leq r \log_2 \left(1 + \frac{P_T}{\sigma^2 r} \right).$$

◁

3.5 Conclusões e Síntese

Este capítulo introduziu conceitos e alguns resultados fundamentais da teoria de majorização aplicados a problemas de comunicação sem fio. No entanto, na perspectiva deste trabalho contribuir com a literatura no assunto, o capítulo em questão foi escrito com certo rigor matemático, abordando situações que são comumente encontradas em se tratando de problemas de capacidade ergódica de canais MIMO.

Após a apresentação dos conceitos e propriedades básicas da relação de majorização, uma atenção especial foi dada ao estudo de funções Schur-côncava e Schur-convexa as quais estão presentes em problemas de sistemas de comunicação, conforme descritos na introdução deste capítulo.

Por fim, apresentamos a majorização estocástica a qual nos possibilita a obtenção de limitantes (inferior e superior) da capacidade ergódica considerando um determinado canal MIMO, motivando portanto, o estudo do próximo capítulo, o qual se propõem a obter limitantes no caso de canal com desvanecimento de Rice.

CAPÍTULO 4

APROXIMAÇÕES E LIMITANTES À CAPACIDADE ERGÓDICA

Este capítulo trata da proposta central desta dissertação, que consiste em determinar limitantes, superior e inferior, à capacidade ergódica de um canal MIMO com desvanecimento de Rice via teoria de majorização. Ademais, serão apresentadas aproximações à capacidade ergódica nos casos de baixos e altos valores de SNRs.

No que diz respeito aos limitantes em canais com desvanecimento de Rice, em McKay e Collings (2005) e Jin, Gao e You (2007) são apresentados limites inferiores e superiores à capacidade ergódica de canais correlacionados, enquanto, Jayaweera e Poor (2005) tratam de canais descorrelacionados. No entanto, em qualquer caso, o tratamento no assunto é ainda um grande desafio devido à dificuldade de se obter expressões fechadas para a capacidade ergódica. Esta complexidade se deve, principalmente, ao desenvolvimento da função de distribuição conjunta dos autovalores da matriz $\mathbf{H}\mathbf{H}^H$ (JAYAWEERA; POOR, 2005; JIN; GAO; YOU, 2007).

Recentemente, Zhong, Wong e Jin (2009) têm mostrado que é possível obter tais limitantes, sem lidar diretamente com a distribuição conjunta de autovalores. O êxito desta proposta foi proporcionado, principalmente, devido a utilização da teoria de majorização, apresentada no capítulo anterior, a qual os possibilitou especificar uma comparação dos elementos da diagonal principal da matriz $\mathbf{H}\mathbf{H}^H$ com seus respectivos valores singulares.

O desenvolvimento deste capítulo, portanto, se fundamenta basicamente na proposta

sugerida por Zhong, Wong e Jin (2009), proporcionando desta forma um método alternativo na obtenção de limitantes à capacidade ergódica no caso de um canal de desvanecimento de Rice.

4.1 Considerações iniciais

Para o desenvolvimento deste trabalho, nesta seção serão especificadas todas as hipóteses e restrições sobre o modelo de sistema em questão MIMO, bem como a formulação do problema.

4.1.1 Descrição do modelo

O modelo utilizado para o canal com múltiplas antenas consiste de um transmissor com n_T antenas e um receptor com n_R antenas, conforme mencionado na Equação (2.4), isto é,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (4.1)$$

No entanto, o estado de informação sobre o canal é perfeitamente conhecido pelo receptor e desconhecido no transmissor. Além disso, considera-se que o número de antenas receptoras não excede o número de antenas transmissoras, portanto, $n_R \leq n_T$.

O vetor de símbolos recebidos é dado por $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$. Por sua vez, $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$ representa o vetor de símbolos transmitidos por cada antena o qual satisfaz a restrição total de potência P_T , ou seja, $\mathbb{E}\{\|\mathbf{x}\|^2\} \leq P_T$. O vetor ruído $\mathbf{n} \in \mathbb{C}^{n_R \times 1}$ segue uma distribuição gaussiana circularmente simétrica cujo vetor média é nulo e matriz de covariância é dada por $\mathbf{R}_n = N_0 \mathbf{I}_{n_R}$, ou noutros termos, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N_0 \mathbf{I}_{n_R})$.

Os elementos da matriz do canal $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{n_R \times n_T}$ são considerados decorrelacionados e de potência média constante¹, isto é, $\mathbb{E}\{|h_{ij}|^2\} = \Omega$. Por sua vez, a matriz do canal segue o modelo de desvanecimento de Rice, conforme apresentado na Equação (2.56), ou seja:

$$\mathbf{H} = \sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}} + \sqrt{\frac{\Omega}{K+1}} \mathbf{H}_w, \quad (4.2)$$

em que K é o fator de Rice, $\bar{\mathbf{H}} = [\bar{h}_{ij}]$ é uma matriz determinística, de posto igual a L , cujos elementos têm potência normalizada, ou seja, $\mathbb{E}\{|\bar{h}_{ij}|^2\} = 1$. Por fim, $\mathbf{H}_w = [\tilde{h}_{ij}]$

¹Em termos computacionais, este trabalho considera tal potência como unitária, i.e, $\Omega = 1$.

representa a matriz espacialmente branca. Desta forma,

$$\mathbf{H}_w \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}). \quad (4.3)$$

Diante destas hipóteses sobre a matriz do canal, esta segue uma distribuição gaussiana complexa de média $\mathbb{E}\{\mathbf{H}\} = \sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}}$ e matriz covariância $\mathbf{R}_{\mathbf{H}} = \frac{\Omega}{K+1} \mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}$. Noutros termos (JAYAWEERA; POOR, 2005; JIN; GAO; YOU, 2007),

$$\mathbf{H} \sim \mathcal{CN}\left(\sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}}, \frac{\Omega}{K+1} \mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R}\right). \quad (4.4)$$

Por fim, em se tratando da capacidade ergódica, como o transmissor não tem informação dos estado do canal, então pelo que foi estudado nas Seções 2.2.1 e 2.2.4 tem-se

$$\bar{C} = \mathbb{E}\left\{\log_2 \left[\det \left(\mathbf{I}_{n_R} + \frac{\rho}{n_T} \mathbf{H}\mathbf{H}^H\right)\right]\right\}, \quad (4.5)$$

ou simplesmente,

$$\bar{C} = \mathbb{E}\left\{\sum_{i=1}^r \log_2 \left[1 + \frac{\rho}{n_T} \lambda_i\right]\right\}, \quad (4.6)$$

sendo $\rho = \frac{P_T}{N_0}$ a relação sinal-ruído (SNR), $r = \text{posto}(\mathbf{H}\mathbf{H}^H)$ e λ_i é autovalor da matriz $\mathbf{H}\mathbf{H}^H$, para todo $i = 1, 2, \dots, r$. Observar-se que, no caso matriz do canal ter posto completo, isto é, $\text{posto}(\mathbf{H}) = n_R$, tem-se (HORN; JOHNSON, 1985):

$$r = \text{posto}(\mathbf{H}\mathbf{H}^H) = \text{posto}(\mathbf{H}^H) = \text{posto}(\mathbf{H}) = n_R. \quad (4.7)$$

4.1.2 Formulação do problema

Conforme descrito no início deste capítulo, o objetivo central deste trabalho é a determinação de limitantes, inferior e superior, à capacidade ergódica do canal mostrada na Equação (4.6) do canal de desvanecimento de Rice conforme a Equação (4.2), via teoria de majorização, mediante as hipóteses levantadas na Seção 4.1.1. Além disso, serão apresentadas aproximações à capacidade para baixos e altos valores de SNRs.

4.1.3 Transformação da matriz do canal

Para o desenvolvimento da nossa proposta, faremos uma transformação na matriz de canal com a finalidade de simplificar as demonstrações dos teoremas que se seguem. Seja,

portanto, a matriz $\mathbf{W} = [w_{ij}] \in \mathbb{C}^{n_R \times n_T}$ definida por:

$$\mathbf{W} = \sqrt{\frac{1+K}{\Omega}} \mathbf{H}. \quad (4.8)$$

Diante desta transformação, tem-se os seguintes resultados:

(t1): $\text{posto}(\mathbf{W}\mathbf{W}^H) = \text{posto}(\mathbf{H}\mathbf{H}^H)$.

(t2): $\mathbf{W} = \sqrt{K} \bar{\mathbf{H}} + \mathbf{H}_w$.

Demonstração: Como a matriz \mathbf{H} é dada conforme a Equação (4.2) segue que:

$$\mathbf{W} = \sqrt{\frac{1+K}{\Omega}} \mathbf{H} = \sqrt{\frac{1+K}{\Omega}} \left(\sqrt{\frac{\Omega K}{K+1}} \bar{\mathbf{H}} + \sqrt{\frac{\Omega}{K+1}} \mathbf{H}_w \right) \quad (4.9)$$

$$= \sqrt{K} \bar{\mathbf{H}} + \mathbf{H}_w. \quad (4.10)$$

□

(t3): $\mathbf{W} \sim \mathcal{CN}(\sqrt{K} \bar{\mathbf{H}}, \mathbf{I}_{n_T} \otimes \mathbf{I}_{n_R})$

Demonstração: Pela propriedade (t2), tem-se $w_{ij} = \sqrt{K} \bar{h}_{ij} + \tilde{h}_{ij}$, sendo $\tilde{h}_{ij} \sim \mathcal{N}(0, 1)$. Para verificar que w_{ij} é uma variável gaussiana de média $\sqrt{K} \bar{h}_{ij}$ e variância igual a 1, analisemos a respectiva função geradora de momentos (MAGALHÃES, 2004), isto é,

$$\begin{aligned} M_{w_{ij}}(t) &\triangleq \mathbb{E}\{e^{tw_{ij}}\} = \mathbb{E}\{e^{t(\sqrt{K} \bar{h}_{ij} + \tilde{h}_{ij})}\} \\ &= e^{t\sqrt{K} \bar{h}_{ij}} \mathbb{E}\{e^{t\tilde{h}_{ij}}\} \\ &= e^{t\sqrt{K} \bar{h}_{ij}} M_{\tilde{h}_{ij}}(t) \\ &= e^{t\sqrt{K} \bar{h}_{ij}} e^{t^2/2}. \end{aligned} \quad (4.11)$$

Portanto, $w_{ij} \sim \mathcal{N}(\sqrt{K} \bar{h}_{ij}, 1)$ e conseqüentemente \mathbf{W} tem uma distribuição gaussiana de média $\mathbb{E}\{\mathbf{W}\} = \sqrt{K} \bar{\mathbf{H}}$. Por fim, analisemos a matriz de covariância de \mathbf{W} , ou seja,

$$\begin{aligned}
\mathbf{R}_{\mathbf{W}} &\triangleq \mathbb{E}\left\{\text{vec}(\mathbf{W})\text{vec}(\mathbf{W})^H\right\} \\
&= \mathbb{E}\left\{\text{vec}\left(\sqrt{\frac{1+K}{\Omega}}\mathbf{H}\right)\text{vec}\left(\sqrt{\frac{1+K}{\Omega}}\mathbf{H}\right)^H\right\} \\
&= \frac{1+K}{\Omega}\mathbb{E}\left\{\text{vec}(\mathbf{H})\text{vec}(\mathbf{H})^H\right\} \\
&= \frac{1+K}{\Omega}\mathbf{R}_{\mathbf{H}} \\
&= \frac{1+K}{\Omega}\frac{\Omega}{1+K}\mathbf{I}_{n_T}\otimes\mathbf{I}_{n_R} \\
&= \mathbf{I}_{n_T}\otimes\mathbf{I}_{n_R}.
\end{aligned} \tag{4.12}$$

Conclui-se, portanto, $\mathbf{W} \sim \mathcal{CN}\left(\sqrt{K}\bar{\mathbf{H}}, \mathbf{I}_{n_T}\otimes\mathbf{I}_{n_R}\right)$. \square

(t4): $\mathbb{E}\{|w_{ij}|^2\} = K + 1$.

Demonstração:

$$\begin{aligned}
\mathbb{E}\{|w_{ij}|^2\} &= \mathbb{E}\{w_{ij}w_{ij}^*\} = \mathbb{E}\left\{\sqrt{\frac{K+1}{\Omega}}h_{ij}\sqrt{\frac{K+1}{\Omega}}h_{ij}^*\right\} \\
&= \frac{K+1}{\Omega}\mathbb{E}\{|h_{ij}|^2\}.
\end{aligned} \tag{4.13}$$

Como $\mathbb{E}\{|h_{ij}|^2\} = \Omega$, segue portanto o resultado. \square

(t5): A capacidade ergódica dada na Equação (4.5) pode ser representada por

$$\bar{C} = \mathbb{E}\left\{\log_2\det\left(\mathbf{I}_{n_R} + \frac{\rho\Omega}{n_T(K+1)}\mathbf{W}\mathbf{W}^H\right)\right\}, \tag{4.14}$$

ou simplesmente,

$$\bar{C} = \mathbb{E}\left\{\sum_{i=1}^r\log_2\left[1 + \frac{\rho\Omega}{n_T(K+1)}\lambda_i^w\right]\right\}. \tag{4.15}$$

sendo $\lambda_1^w, \lambda_2^w, \dots, \lambda_r^w$ autovalores da matriz $\mathbf{W}\mathbf{W}^H$.

Demonstração: A primeira expressão é decorrente da igualdade $\mathbf{H} = \sqrt{\frac{\Omega}{K+1}}\mathbf{W}$, enquanto a segunda é uma consequência direta de resultados estudados nas Seções 2.2.1 e 2.2.4.

\square

4.2 Limitante superior à capacidade ergódica

Fazendo uso da teoria de majorização, obteremos nesta seção um limitante superior à capacidade ergódica do canal MIMO $\mathbf{W} \in \mathbb{C}^{n_R \times n_T}$, com desvanecimento de Rice. Conforme já foi comentado anteriormente, esta teoria prevê um tratamento matemático diferente sobre o assunto, evitando-se o uso direto da distribuição estatística não-central de Wishart. Especificamente, o método consiste em definirmos uma função de Schur-côncava em termos dos autovalores e elementos da diagonal da matriz hermitiana $\mathbf{W}\mathbf{W}^H$, conforme mostra o teorema a seguir cujas hipóteses já foram descritas na seção anterior.

Teorema 4.1 A capacidade ergódica de um canal MIMO com desvanecimento de Rice é limitada superiormente por

$$\bar{C} \leq C_{\text{up}} = n_R \log_2(1 + \rho\Omega). \quad (4.16)$$

Demonstração: Seja $\lambda(\mathbf{W}\mathbf{W}^H) = (\lambda_1^w, \lambda_2^w, \dots, \lambda_{n_R}^w)$ o vetor de autovalores e seja $d(\mathbf{W}\mathbf{W}^H) = (d_1, d_2, \dots, d_{n_R})$ o vetor dos elementos da diagonal principal da matriz $\mathbf{W}\mathbf{W}^H$. Por questão de conveniência, considera-se que as coordenadas destes vetores estejam dispostas em ordem decrescente, isto é, $\lambda_1^w \geq \lambda_2^w \geq \dots \geq \lambda_{n_R}^w$ e $d_1 \geq d_2 \geq \dots \geq d_{n_R}$. No entanto, como posto $(\mathbf{W}\mathbf{W}^H) = r$, tem-se

$$\lambda_1^w \geq \lambda_2^w \geq \dots \geq \lambda_r^w > 0 \quad \text{e} \quad \lambda_{r+1}^w = \lambda_{r+2}^w = \dots = \lambda_{n_R}^w = 0. \quad (4.17)$$

Por outro lado, considera-se a função real $\phi : \mathcal{D}_+ \subset \mathbb{R}^{n_R} \rightarrow \mathbb{R}$ definida por

$$\phi(\mathbf{x}) = \mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} x_i \right] \right\}. \quad (4.18)$$

Pelo Teorema 3.7 o vetor $d(\mathbf{W}\mathbf{W}^H)$ é majorizado por $\lambda(\mathbf{W}\mathbf{W}^H)$, ou seja, $d(\mathbf{W}\mathbf{W}^H) \prec \lambda(\mathbf{W}\mathbf{W}^H)$. Além disso, como a função $\phi(\cdot)$ é Schur-côncava (ver Teorema 3.12), tem-se

$$\phi(\lambda(\mathbf{W}\mathbf{W}^H)) \leq \phi(d(\mathbf{W}\mathbf{W}^H)), \quad (4.19)$$

ou equivalentemente,

$$\mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} \lambda_i^w \right] \right\} \leq \mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right] \right\}. \quad (4.20)$$

Encontramos desta forma, um limitante superior à capacidade ergódica, ou seja,

$$\bar{C} \leq \mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right] \right\}, \quad (4.21)$$

pois, $\bar{C} = \mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} \lambda_i^w \right] \right\}$. Igualdade esta já apresentada na propriedade **(t5)** da seção anterior.

Por sua vez, tomando-se o segundo membro da desigualdade na Equação (4.21) tem-se:

$$\mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right] \right\} = \mathbb{E} \left\{ \log_2 \left[\prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right] \right\}. \quad (4.22)$$

Aplicando a desigualdade de Jensen na Equação (4.22) segue que (MARSHALL; OLKIN, 1979; BOYD; VANDENBERGHE, 2004):

$$\begin{aligned} \bar{C} &\leq \mathbb{E} \left\{ \log_2 \left[\prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right] \right\} \\ &\leq \log_2 \left[\mathbb{E} \left\{ \prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right\} \right]. \end{aligned} \quad (4.23)$$

Por outro lado, sabe-se que $d_i = \sum_{j=1}^{n_T} |w_{ij}|^2$, para $i = 1, 2, \dots, n_R$ e que, pela propriedade **(t4)**, $\mathbb{E} \{d_i\} = n_T(K+1)$. Desta forma, a esperança do produto das variáveis aleatórias d_1, d_2, \dots, d_{n_R} é dada por:

$$\begin{aligned} \mathbb{E} \left\{ \prod_{i=1}^{n_R} d_i \right\} &= \sum_{j_1=1}^{n_T} \sum_{j_2=1}^{n_T} \cdots \sum_{j_{n_R}=1}^{n_T} \mathbb{E} \left\{ \prod_{k=1}^{n_R} |w_{kj_k}|^2 \right\} \\ &= (n_T \cdot (K+1))^{n_R}. \end{aligned}$$

Além disso²,

$$\mathbb{E} \left\{ \prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right\} = (1 + \rho\Omega)^{n_R}. \quad (4.24)$$

Substituindo, portanto, o operador esperança por $(1 + \rho\Omega)^{n_R}$ em (4.23) tem-se:

$$\begin{aligned} \bar{C} &\leq \log_2 (1 + \rho\Omega)^{n_R} \\ &= n_R \log_2 (1 + \rho\Omega), \end{aligned} \quad (4.25)$$

o que prova o resultado. \square

O limitante superior proposto pelo Teorema 4.1 também pode ser obtido a partir da *inequação de Hadamard* (3.37). Para aplicação de tal resultado, basta verificar que a matriz hermitiana $\mathbf{P} = \mathbf{I}_{n_R} + \alpha \mathbf{W} \mathbf{W}^H$ é semi-definida positiva, sendo $\alpha = \frac{\rho\Omega}{n_T(K+1)}$. Com efeito, para todo vetor $\mathbf{x} \in \mathbb{C}^{n_R \times 1}$ não-nulo, tem-se

$$\begin{aligned} \mathbf{x}^H \mathbf{P} \mathbf{x} &= \mathbf{x}^H (\mathbf{I}_{n_R} + \alpha \mathbf{W} \mathbf{W}^H) \mathbf{x} \\ &= \mathbf{x}^H \mathbf{x} + \alpha \mathbf{x}^H (\mathbf{W} \mathbf{W}^H) \mathbf{x} \\ &= \mathbf{x}^H \mathbf{x} + \alpha (\mathbf{x}^H \mathbf{W}) (\mathbf{W}^H \mathbf{x}) \\ &= \mathbf{x}^H \mathbf{x} + \alpha (\mathbf{W}^H \mathbf{x})^H (\mathbf{W}^H \mathbf{x}) \\ &= \|\mathbf{x}\|^2 + \alpha \|\mathbf{W}^H \mathbf{x}\|^2 \\ &\geq 0. \end{aligned} \quad (4.26)$$

Verifica-se, desta forma, que a matriz hermitiana $\mathbf{P} = [p_{ij}]$ é semi-definida positiva. Então, pela *desigualdade de Hadamard*, garante-se que $\det(\mathbf{P}) \leq \prod_{i=1}^{n_R} p_{ii}$. Equivalentemente,

$$\det(\mathbf{I}_{n_R} + \alpha \mathbf{W} \mathbf{W}^H) \leq \prod_{i=1}^{n_R} (1 + \alpha d_i), \quad (4.27)$$

²Os detalhes a respeito destas últimas igualdades encontram-se no Apêndice B.

sendo que $d_i = \sum_{j=1}^{n_T} |w_{ij}|^2$, para $i = 1, 2, \dots, n_R$. A partir deste último resultado, aplica-se a função $\log_2(\cdot)$ e operador esperança $\mathbb{E}\{\cdot\}$, nesta ordem, obtendo-se portanto

$$\begin{aligned} \bar{C} &= \mathbb{E} \left\{ \log_2 \det \left(\mathbf{I}_{n_R} + \frac{\rho\Omega}{n_T(K+1)} \mathbf{W}\mathbf{W}^H \right) \right\} \\ &\leq \mathbb{E} \left\{ \log_2 \left[\prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right] \right\}. \end{aligned} \quad (4.28)$$

Notemos que esta desigualdade aparece em (4.23). Então, repetindo os resultados obtidos a partir de tal desigualdade, chega-se em

$$\bar{C} \leq C_{\text{up}} = n_R \log_2(1 + \rho\Omega), \quad (4.29)$$

o que comprova a validade de tal argumento.

O resultado obtido no Teorema 4.1 mostra que o limitante superior C_{up} , independe do número de antenas transmissoras e do fator de Rice. Além disso, para valores fixos de Ω e SNR, a função C_{up} cresce linearmente com o número de antenas receptoras. Para ilustrar este fato, a Figura 7 abaixo considera um sistema MIMO com desvanecimento de Rice conforme descrito na Equação (4.2), com $\Omega = 1$ e $0 \leq \text{SNR} \leq 25$ dB, com incremento de 5 dB .

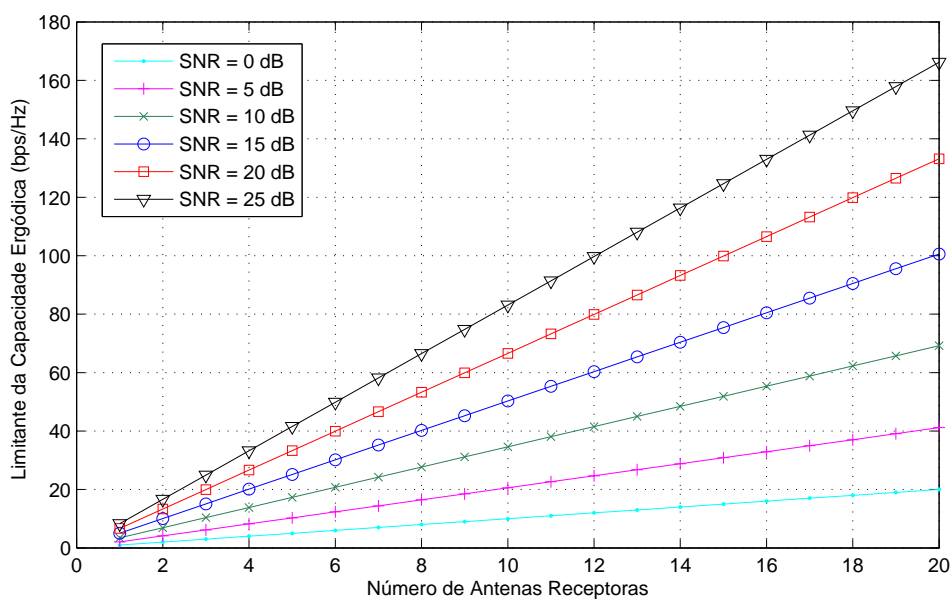


Figura 7: Comparação entre limitantes superiores para $\Omega = 1$ e $0 \leq \text{SNR} \leq 25$ dB, com incremento de 5 dB.

Analisando, desta vez, o limitante C_{up} como função dos valores das SNRs, a Figura 8 faz um comparativo entre os valores dos limitantes superiores à capacidade ergódica, de um sistema MIMO com desvanecimento de Rice, em que o número antenas transmissoras é igual ao número de receptoras, admitindo-se os seguintes cenários: 1×1 , 2×2 , 4×4 e 8×8 .

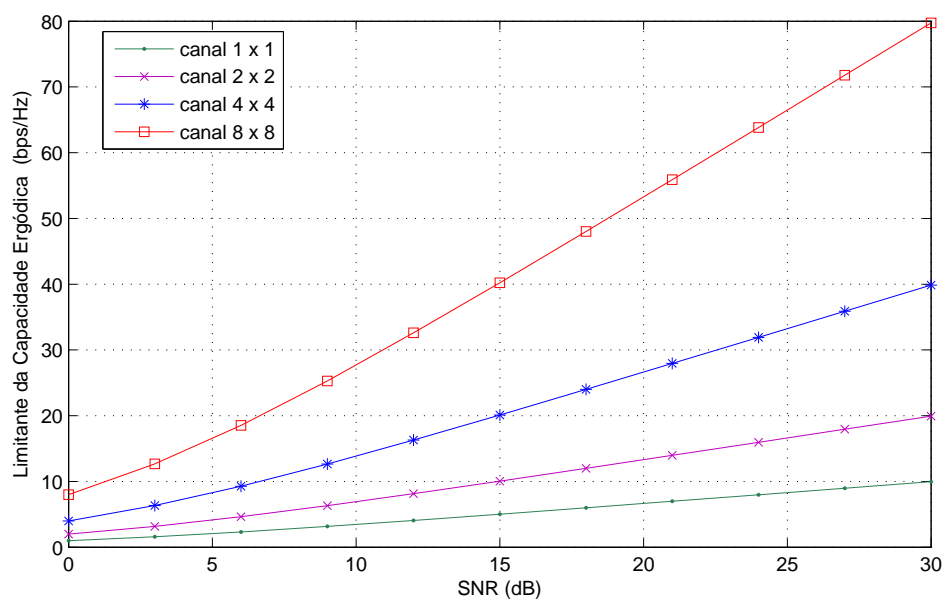


Figura 8: Comparação entre limitantes superiores para os cenários 1×1 , 2×2 , 4×4 e 8×8 .

Nas Figuras 9, 10 e 11 seguintes, são feitas comparações do limitante superior obtido neste trabalho com a expressão da capacidade ergódica \bar{C} apresentada na Propriedade (t5). Para a primeira figura, a matriz do canal considerada tem ordem 2×2 , para a segunda, a ordem da matriz é 4×4 , enquanto para a Figura 11, o cenário considerado é 3×5 . Além disso, em todos os casos o fator de Rice considerado é $K = 3$. O objetivo deste comparativo é ilustrar, para um determinado valor da relação sinal-ruído fixado, a proximidade dos valores do limitante em relação ao valor da capacidade ergódica.

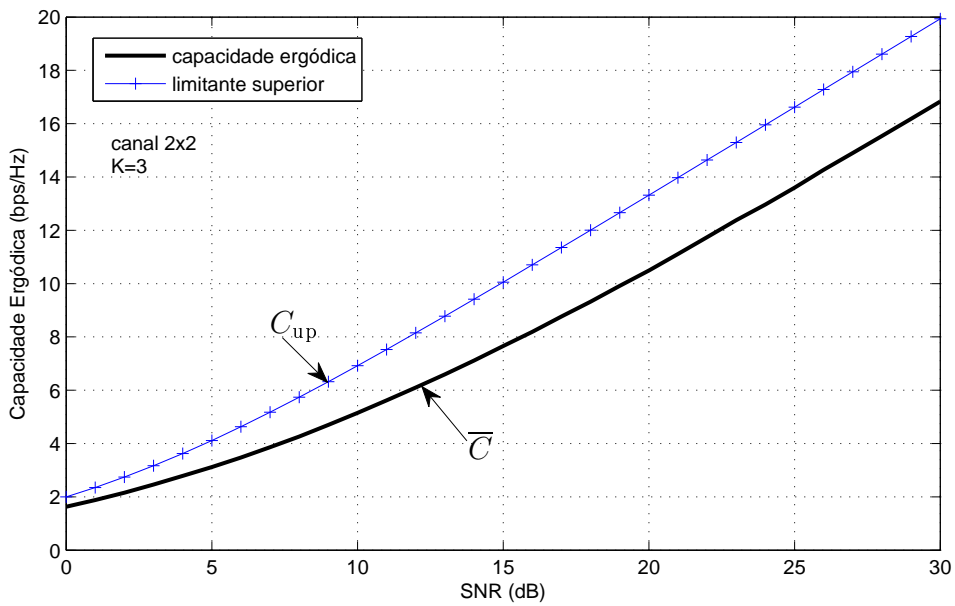


Figura 9: Comparação entre limitante superior C_{up} e capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$.

Nota-se, nesta figura como nas outras que se seguem, que os valores da capacidade e do limitante são bem próximos, para pequenos valores de SNRs, porém estes tendem a se distanciar com o crescimento do valor da relação sinal-ruído e com o aumento do número de antenas receptoras.

Convém ressaltar que o resultado obtido para C_{up} independe do número de antenas transmissoras. Portanto, os canais de ordem $n_R \times n_T$, com $n_R \neq n_T$, e $n_R \times n_R$ ambos possuem o mesmo limitante superior. Por exemplo, os canais de ordem 3×5 e 3×3 possuem mesmo limitante superior, cuja curva apresenta-se na Figura 11.

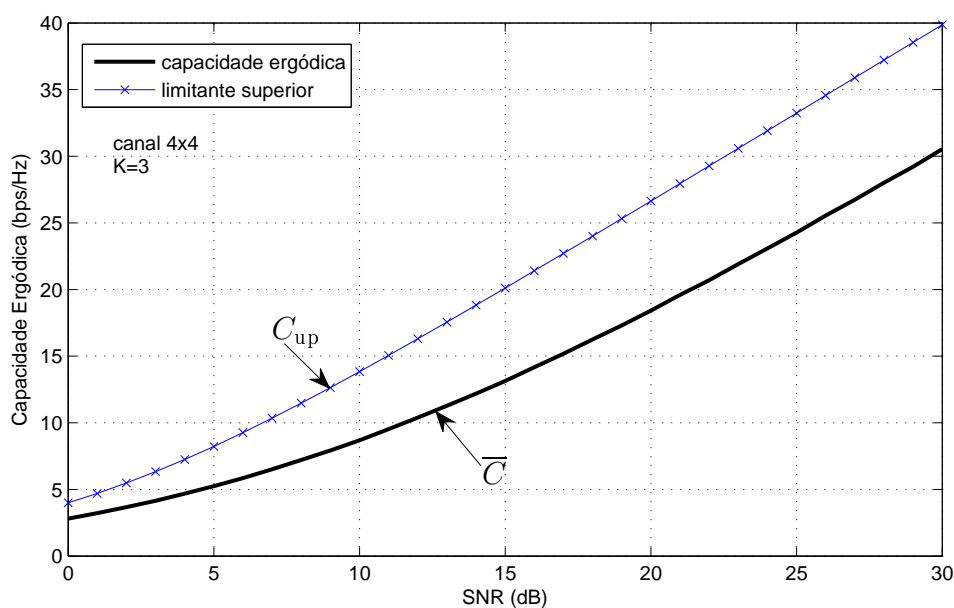


Figura 10: Comparação entre limitante superior C_{up} e capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$.

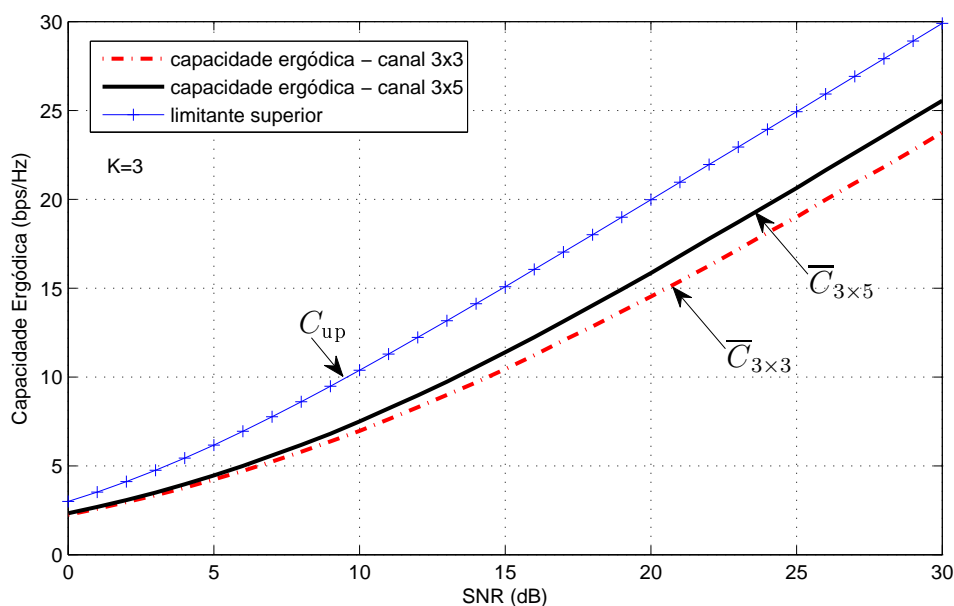


Figura 11: Comparação entre limitante superior C_{up} e as capacidades ergódicas $\bar{C}_{3 \times 3}$ e $\bar{C}_{3 \times 5}$ para os cenários 3×3 e 3×5 , respectivamente. O fator de Rice utilizado é $K = 3$.

Com o propósito de quantificar o quão próximo está o limitante superior proposto pelo Teorema 4.1 da capacidade ergódica, a próxima figura apresenta o erro relativo desta aproximação, o qual permite-nos fazer a referida análise. Em termos de simulação, as curvas foram geradas a partir do acréscimo da constante de Rice e fixando-se alguns valores de SNR.

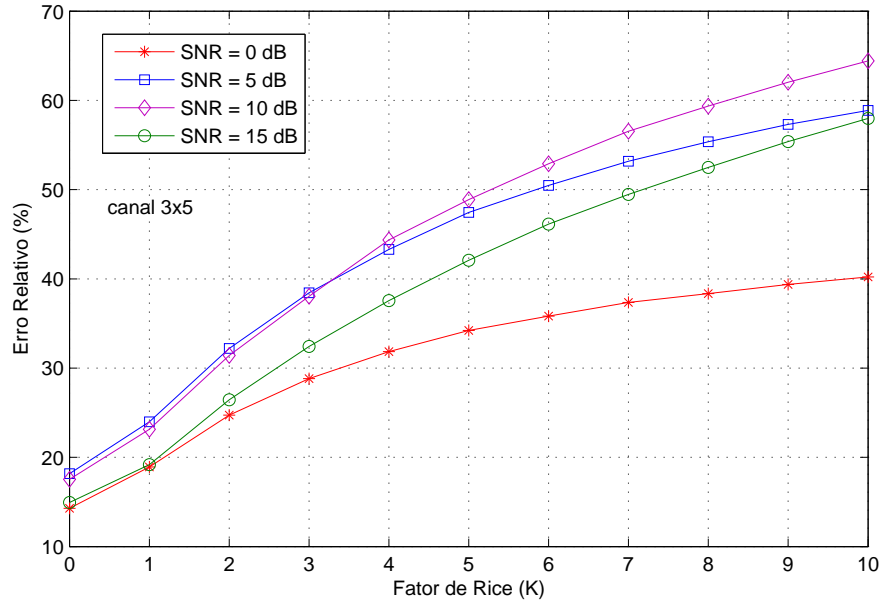


Figura 12: Comparação entre erros relativos de aproximação do limitante C_{up} à capacidade ergódica \bar{C} num canal de ordem 3×5 .

No exemplo dado, observar-se que a taxa de erro cresce juntamente com o acréscimo do fator de Rice, para o caso de um canal de ordem 3×5 . Desta forma, no canal com desvanecimento de Rayleigh ($K = 0$), tem-se que o erro relativo de aproximação fica em torno de 15% na relação sinal-ruído de 0 dB e 15 dB e, em torno de 18%, para SNR de 5 dB e 10 dB.

Analisando desta vez o maior valor de SNR em questão ($K = 10$), observa-se que o erro de aproximação, no caso de 0 dB, fica em torno de 40%, enquanto, para os outros valores da relação sinal-ruído, tem-se uma elevada taxa de erro que. Conforme mostra a partir, tais taxas apresentam-se a partir 57%, aproximadamente.

Ainda com o desenvolvimento do Teorema 4.1, obtivemos um outro limitante superior, o qual representaremos por $C_{\text{up}}^{\text{int}}$ (ver Equação (4.19)). Especificamente,

$$C_{\text{up}}^{\text{int}} = \mathbb{E} \left\{ \sum_{i=1}^{n_R} \log_2 \left[1 + \frac{\rho \Omega}{n_T (K + 1)} d_i \right] \right\}, \quad (4.30)$$

sendo d_1, d_2, \dots, d_{n_R} os elementos da diagonal principal da matriz $\mathbf{W}\mathbf{W}^H$. Além disso,

$$d_i = \sum_{j=1}^{n_T} |w_{ij}|^2, \quad \text{para } i = 1, 2, \dots, n_R. \quad (4.31)$$

Diante dos cálculos apresentados no teorema, verifica-se que tal limitante apresenta valores intermediários entre a capacidade ergódica \bar{C} e o limitante C_{up} . Desta maneira, a curva do referido limitante superior tem uma melhor aproximação à capacidade ergódica do que C_{up} . Para verificar esta afirmação, apresentamos três situações em que os canais possuem ordem 2×2 , 4×4 e 3×5 , com fator de Rice $K = 3$, conforme ilustram as figuras a seguir.

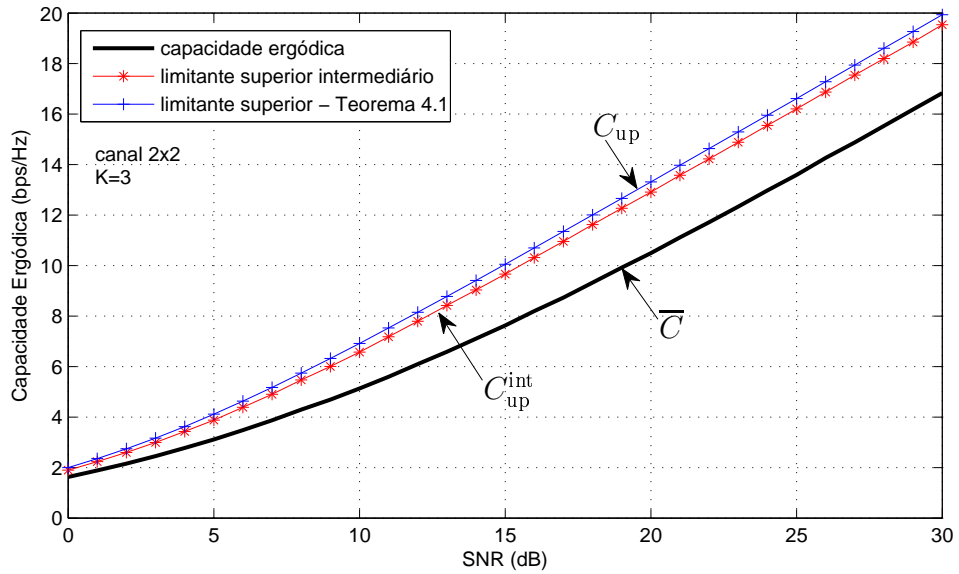


Figura 13: Comparação entre os limitantes superiores C_{up} e $C_{\text{up}}^{\text{int}}$ e a capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$.

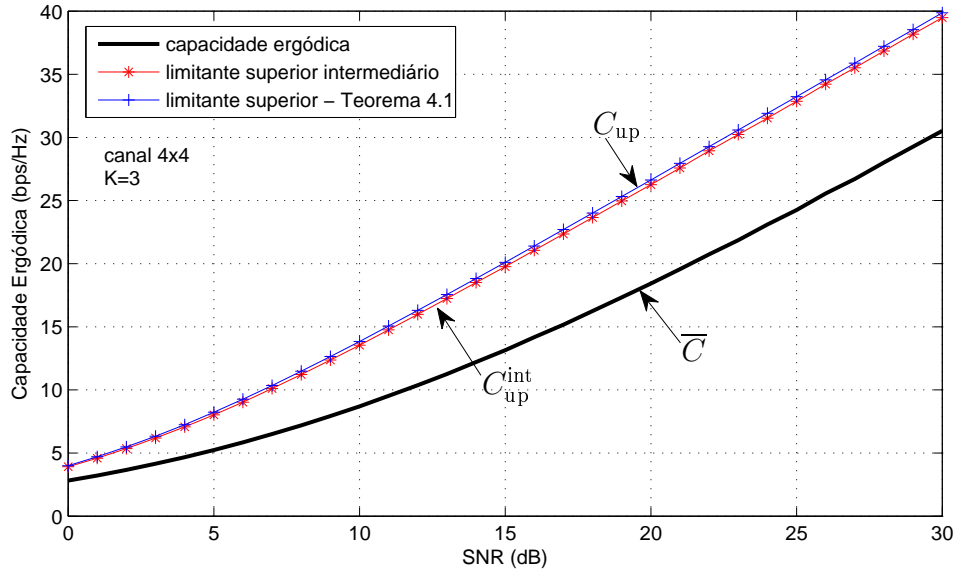


Figura 14: Comparação entre os limitantes superiores C_{up} e C_{up}^{int} e a capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$.

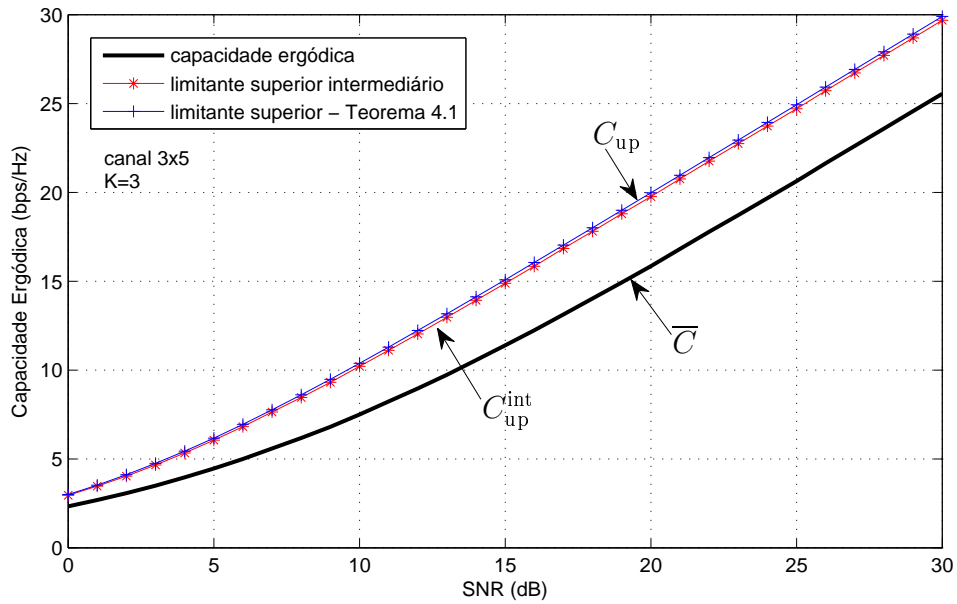


Figura 15: Comparação entre os limitantes superiores C_{up} e C_{up}^{int} e a capacidade ergódica \bar{C} para o cenário 3×5 , com fator de Rice $K = 3$.

Notemos, a partir dos resultados simulados, que as curvas, as quais são representadas pelo limitantes superiores C_{up} e C_{up}^{int} , apresentam uma excelente aproximação entre si. No entanto, o limitante C_{up} apresenta-se como um resultado determinístico, dependendo de uma pequena quantidade de parâmetros. Conseqüentemente, o limitante superior à capacidade ergódica C_{up} requer menos recurso de hardware do que C_{up}^{int} , por este último depender de simulação Monte Carlo para ser gerado.

4.2.1 Aproximação à capacidade ergódica em alto regime de SNR

Para altos valores de SNRs, a função $n_R \log_2(1 + \rho\Omega)$ pode ser aproximada por $n_R \log_2(\rho\Omega)$ (PAULRAJ; NABAR; GORE, 2003). Consequentemente, a capacidade ergódica \bar{C} pode ser aproximada por

$$\bar{C} \approx n_R (\log_2 \rho + \log_2 \Omega). \quad (4.32)$$

Na Figura 16 são ilustradas as capacidades ergódicas e as suas respectivas aproximações para os casos de canais de ordem 2×2 , 4×4 e 8×8 , com fator de Rice $K = 3$ e $\Omega = 1$. Verifica-se, que as melhores aproximações são detectadas quando há uma pequena quantidade de antenas receptoras envolvidas, nas situações em que há igualdade entre os números de antenas transmissoras e receptoras. Portanto, o aumento dos valores de SNR e do número de antenas receptoras, implica num distanciamento entre a respectiva curva da capacidade ergódica e a sua aproximação quase linear.

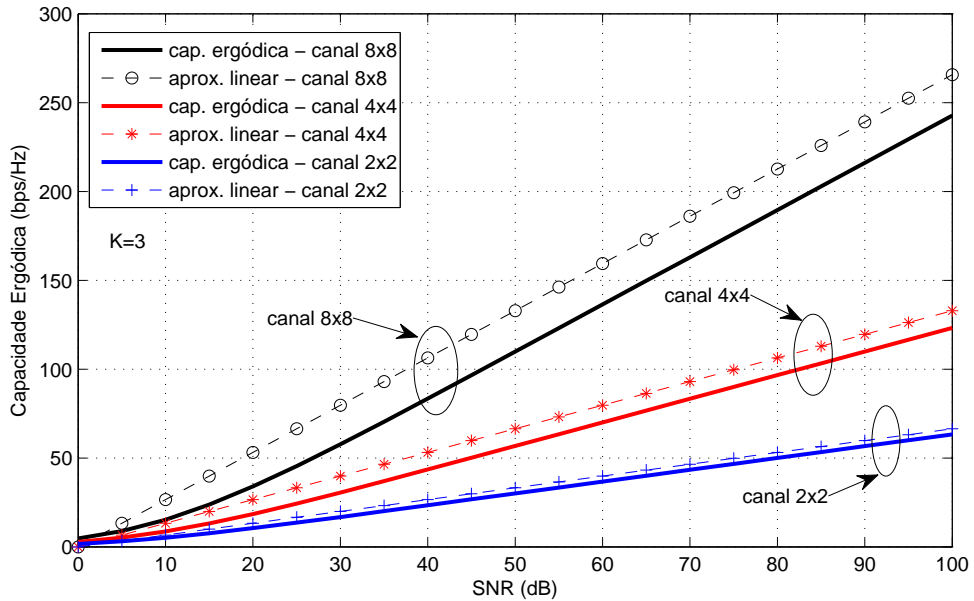


Figura 16: Aproximações à capacidade ergódica nos casos 2×2 , 4×4 e 8×8 , com fator de Rice $K = 3$, para alto regime SNR.

Na próxima seção, são obtidos dois limitantes inferiores à capacidade: C_{10} e C_{10}^{int} . No entanto, o ponto de partida para a obtenção do primeiro limitante também toma-se como referência a função Schur-côncava $\phi(\cdot)$ apresentada em (4.18).

4.3 Limitante inferior à capacidade ergódica

De maneira semelhante à seção anterior, apresentaremos um limitante inferior à capacidade ergódica do canal via teoria de majorização, bem como uma aproximação à capacidade para baixos valores de SNRs.

Teorema 4.2 A capacidade ergódica de um canal MIMO com desvanecimento de Rice é limitada inferiormente por

$$C_{\text{lo}} = \mathbb{E} \left\{ \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} \lambda_{\text{max}}^w \right] \right\} \leq \bar{C}, \quad (4.33)$$

sendo λ_{max}^w o maior autovalor da matriz $\mathbf{W}\mathbf{W}^H$.

Demonstração: Por questão de simplicidade fazemos $\alpha = \frac{\rho\Omega}{n_T(K+1)}$. Pela Propriedade (p7) da Seção 3.1.1, sabe-se que

$$d(\mathbf{W}\mathbf{W}^H) = (\lambda_1^w, \lambda_2^w, \dots, \lambda_{n_R}^w) \prec \left(\sum_{i=1}^{n_R} \lambda_i^w, 0, \dots, 0 \right) = \mathbf{S}. \quad (4.34)$$

Como a função $\phi(\cdot)$ na Equação (4.18) é Schur-côncava, tem-se

$$\phi(\mathbf{S}) \leq \phi(d(\mathbf{W}\mathbf{W}^H)) = \bar{C}, \quad (4.35)$$

ou seja,

$$\mathbb{E} \left\{ \log_2 \left[1 + \alpha \sum_{i=1}^{n_R} \lambda_i^w \right] \right\} \leq \bar{C}. \quad (4.36)$$

Por sua vez, para todo $k = 1, 2, \dots, n_R$, tem-se

$$1 + \alpha \lambda_k^w \leq 1 + \alpha \sum_{i=1}^{n_R} \lambda_i^w. \quad (4.37)$$

Em particular, o resultado é válido para o maior autovalor da matriz $\mathbf{W}\mathbf{W}^H$, isto é,

$$1 + \alpha \lambda_{\max}^w \leq 1 + \alpha \sum_{i=1}^{n_R} \lambda_i^w, \quad (4.38)$$

e como a função $\log_2(\cdot)$ é crescente, segue que

$$\log_2 [1 + \alpha \lambda_{\max}^w] \leq \log_2 \left[1 + \alpha \sum_{i=1}^{n_R} \lambda_i^w \right]. \quad (4.39)$$

Portanto, aplicando-se o operador esperança na desigualdade obtida em (4.39) tem-se que (MAGALHÃES, 2004),

$$\begin{aligned} \mathbb{E} \{ \log_2 (1 + \alpha \lambda_{\max}^w) \} &\leq \mathbb{E} \left\{ \log_2 \left[1 + \alpha \sum_{i=1}^{n_R} \lambda_i^w \right] \right\} \\ &\leq \bar{C}, \end{aligned} \quad (4.40)$$

ou seja,

$$C_{\text{lo}} = \mathbb{E} \left\{ \log_2 \left[1 + \frac{\rho \Omega}{n_T (K + 1)} \lambda_{\max}^w \right] \right\} \leq \bar{C}, \quad (4.41)$$

o que completa a prova. □

Analisando o comportamento das curvas que representam a capacidade ergódica \bar{C} e seu respectivo limitante inferior C_{lo} , há um ascendente distanciamento entre as referidas curvas, com relação ao crescimento dos valores de SNR. Além disso, para altos valores SNR, o limitante inferior C_{lo} tem um comportamento quase linear. As Figuras 17 e 18 ilustram estas características para o caso de canais, de ordens 2×2 e 4×4 , com desvanecimento de Rice de fator $K = 3$.

No desenvolvimento do Teorema 4.2, também obtivemos um outro limitante inferior à capacidade ergódica (ver Equação (4.36)), o qual será de chamado de $C_{\text{lo}}^{\text{int}}$. Tal resultado está compreendido entre C_{lo} e \bar{C} para um valor de SNR fixado. Especificamente, tal limitante é dado por

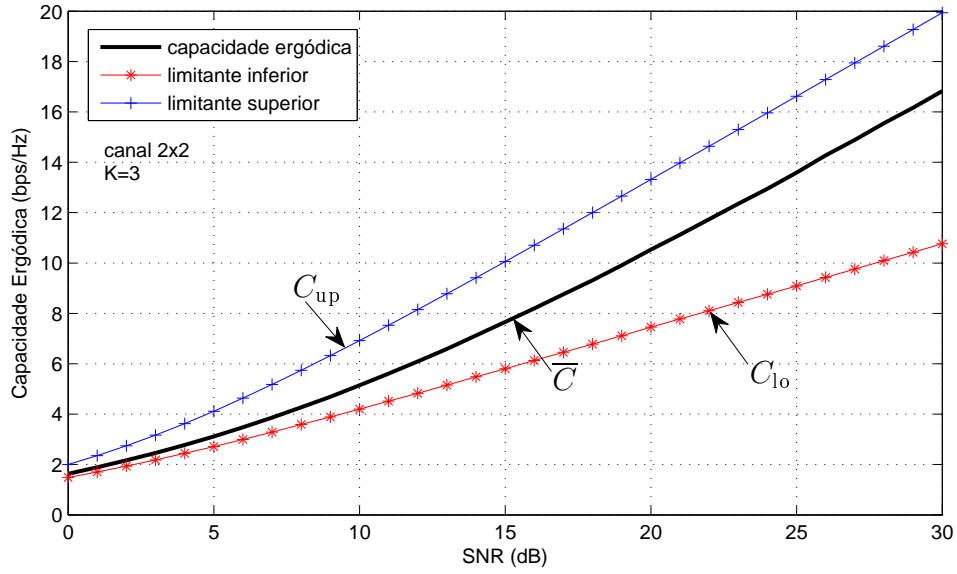


Figura 17: Comparação entre o limitante inferior C_{lo} e a capacidade ergódica \bar{C} para o cenário 2×2 , com fator de Rice $K = 3$.

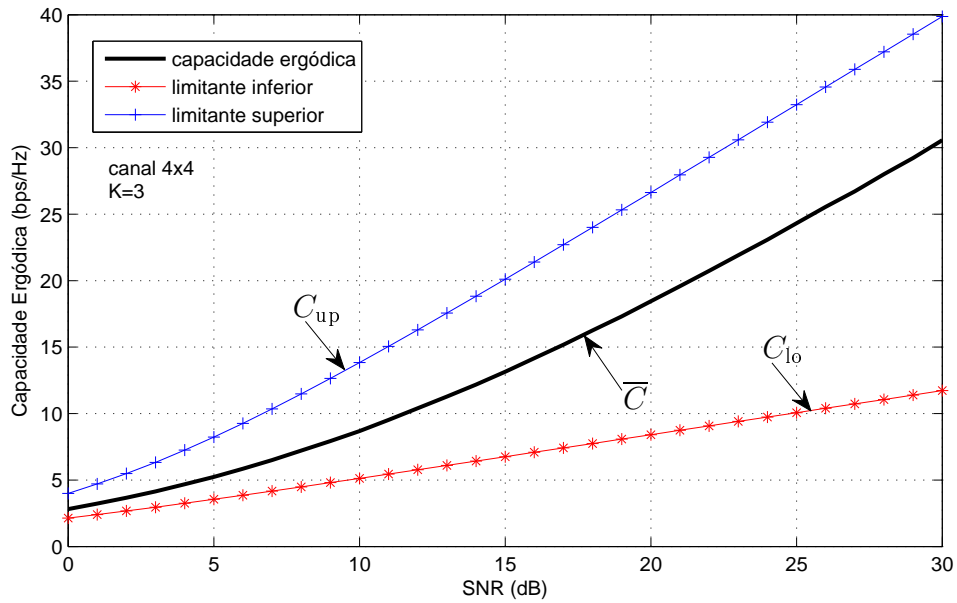


Figura 18: Comparação entre o limitante inferior C_{lo} e a capacidade ergódica \bar{C} para o cenário 4×4 , com fator de Rice $K = 3$.

$$C_{lo}^{int} = \mathbb{E} \left\{ \log_2 \left[1 + \frac{\rho\Omega}{n_T(K+1)} \sum_{i=1}^{n_R} \lambda_i^w \right] \right\}. \quad (4.42)$$

Para apresentar o resultado graficamente, consideramos um canal, de ordem 8×8 , com desvanecimento de Rice de fator $K = 3$, conforme ilustra a Figura 19. Nota-se

que as curvas determinadas por tais limitantes são bem aproximadas. Porém, devido a esta aproximação, com o aumento dos valores de SNR, as curvas relativas a capacidade ergódica e o limitante tendem a um acentuado distanciamento. Por outro lado, para baixos valores de SNR, tem-se uma excelente aproximação à capacidade ergódica.

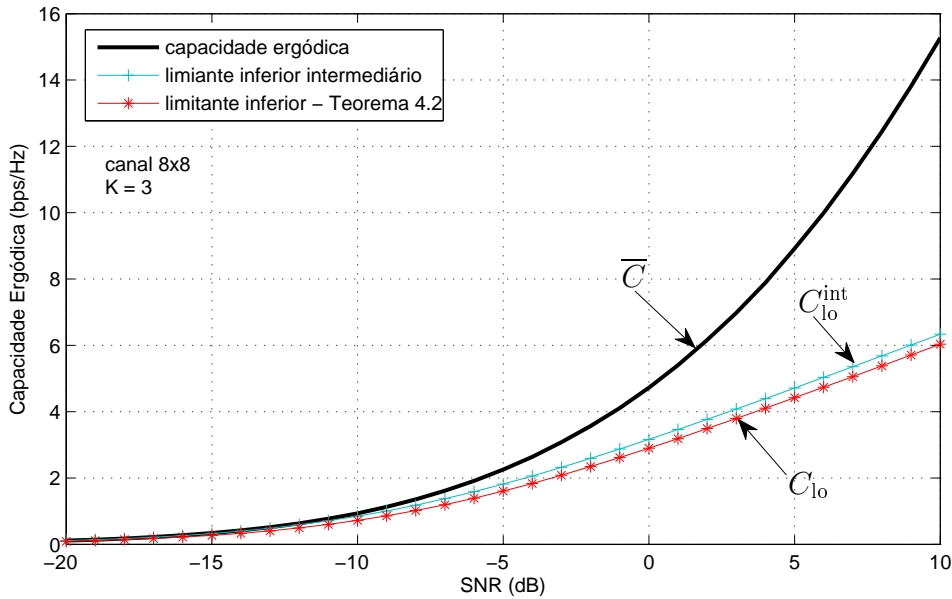


Figura 19: Comparação entre os limitantes inferiores C_{lo} e C_{lo}^{int} e a capacidade ergódica \bar{C} para o cenário 8×8 , com fator de Rice $K = 3$.

4.3.1 Aproximação à capacidade ergódica em baixo regime SNR

O próximo teorema, considerando pequenos valores para SNRs dará uma aproximação à capacidade ergódica a qual dependerá apenas de termos determinísticos como a constante Ω , o número de antenas receptoras e os valores da relação sinal-ruído.

Teorema 4.3 A capacidade ergódica de um canal com desvanecimento de Rice, num baixo regime SNR, pode ser aproximada por

$$\bar{C} \approx n_R \rho \Omega \log_2(e). \quad (4.43)$$

Demonstração: Novamente, por questão de simplicidade chamemos $\beta = \frac{\Omega}{n_T(K+1)}$. Assim, seguindo a mesma idéia de se definir funções específicas as quais são aplicadas na *inequação de Schur*, consideremos portanto, a função real $\varphi : \mathcal{D}_+ \subset \mathbb{R}^{n_R} \rightarrow \mathbb{R}$ definida por

$$\varphi(\mathbf{x}) = \mathbb{E} \left\{ \sum_{i=1}^{n_R} f(x_i) \right\}. \quad (4.44)$$

Desde que $f(\cdot)$ é linear, tem-se que $\varphi(\cdot)$ é uma função Schur-côncava e Schur-convexa. Adicionando-se ao fato que $d(\mathbf{W}\mathbf{W}^H) \prec \lambda(\mathbf{W}\mathbf{W}^H)$, tem-se

$$\varphi(\lambda(\mathbf{W}\mathbf{W}^H)) = \varphi(d(\mathbf{W}\mathbf{W}^H)). \quad (4.45)$$

Noutros termos,

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=1}^{n_R} \lambda_i^w \right\} &= \mathbb{E} \left\{ \sum_{i=1}^{n_R} d_i \right\} \\ &= \sum_{i=1}^{n_R} \mathbb{E}\{d_i\} \\ &= \sum_{i=1}^{n_R} n_T(K+1) \\ &= n_R n_T(K+1). \end{aligned} \quad (4.46)$$

Por outro lado, para baixo regime de SNR, os valores de ρ são tomados suficientemente pequenos. Assim, de acordo com Paulraj, Nabar e Gore (2003), para $\rho \ll 1$ e para $k > 0$, a função $g(\rho) = \log_2(1 + k\rho)$ é aproximada pela função linear $f(\rho) = (k \log_2 e)\rho$. Desta forma, a função $g(\rho) = \log_2[1 + \rho\beta \sum_{i=1}^{n_R} \lambda_i^w]$ é aproximada pela função linear

$$f(\rho) = \left(\beta \log_2 e \sum_{i=1}^{n_R} \lambda_i \right) \rho. \quad (4.47)$$

Assim, a esperança $\mathbb{E}\{\log_2[1 + \rho\beta \sum_{i=1}^{n_R} \lambda_i^w]\}$ é aproximada por

$$\mathbb{E} \left\{ \left(\beta \log_2(e) \sum_{i=1}^{n_R} \lambda_i^w \right) \rho \right\}. \quad (4.48)$$

Contudo,

$$\begin{aligned}
\mathbb{E} \left\{ \left(\beta \log_2(e) \sum_{i=1}^{n_R} \lambda_i^w \right) \rho \right\} &= \beta \rho \log_2(e) \mathbb{E} \left\{ \sum_{i=1}^{n_R} \lambda_i^w \right\} \\
&= \beta \rho \log_2(e) n_R n_T (K + 1) \\
&= \frac{\Omega}{n_T (K + 1)} \rho \log_2(e) n_R n_T (K + 1) \\
&= n_R \rho \Omega \log_2(e).
\end{aligned} \tag{4.49}$$

Portanto, a capacidade ergódica \bar{C} pode ser aproximada por $n_R \rho \Omega \log_2(e)$, ou seja,

$$\bar{C} \approx n_R \rho \Omega \log_2(e), \tag{4.50}$$

o que completa a prova. \square

Para a figura que se segue, foram considerados alguns canais de ordem quadrada, cujas curvas foram simuladas sobre o intervalo de SNRs de -20 dB a -5 dB, com incremento de 1 dB. Com decréscimo dos valores de SNR, as aproximações tendem sobrepor-se a curva da respectiva capacidade ergódica.

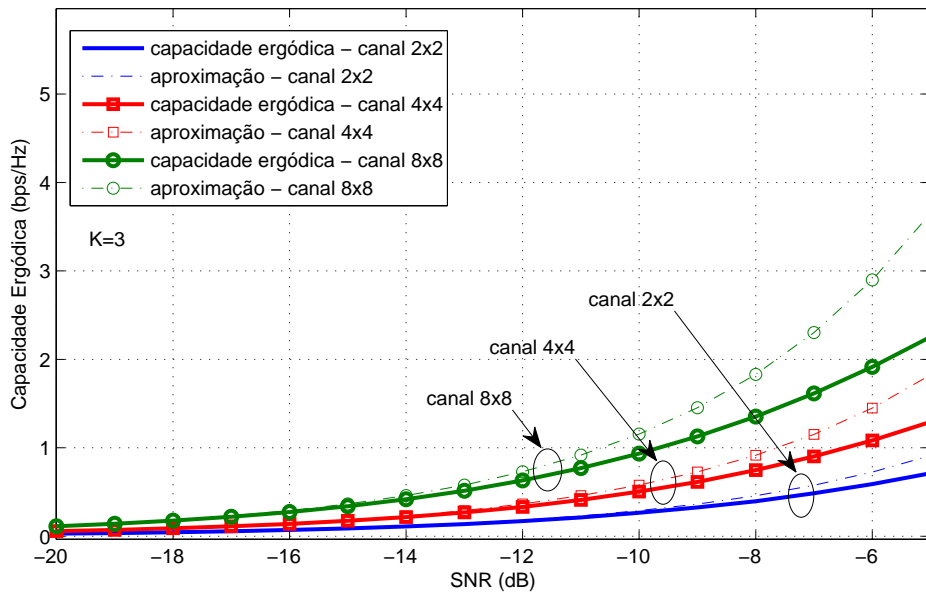


Figura 20: Aproximações à capacidade ergódica nos casos 2×2 , 4×4 e 8×8 , com fator de Rice $K = 3$, para baixo regime SNR.

Uma consequência direta da demonstração deste teorema consiste na determinação da esperança da variável aleatória λ_i^w , para qualquer que seja $i = 1, 2, \dots, n_R$, no caso dos autovalores $\lambda_1^w, \lambda_2^w, \dots, \lambda_{n_R}^w$ serem variáveis identicamente distribuídas.

Como as variáveis em questão são i.i.d., façamos $\mathbb{E}\{\lambda_i^w\} = \mathbb{E}\{\lambda^w\}$. Assim,

$$\mathbb{E} \left\{ \sum_{i=1}^{n_R} \lambda_i^w \right\} = \mathbb{E} \left\{ \sum_{i=1}^{n_R} d_i \right\} = n_R n_T (K + 1). \quad (4.51)$$

Por outro lado,

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=1}^{n_R} \lambda_i^w \right\} &= \sum_{i=1}^{n_R} \mathbb{E}\{\lambda_i^w\} \\ &= n_R \mathbb{E}\{\lambda^w\}. \end{aligned} \quad (4.52)$$

Assim, das igualdades em (4.51) e (4.52) conclui-se que $\mathbb{E}\{\lambda^w\} = n_T(K + 1)$.

4.4 Estudo de um caso assintótico

Nesta seção, faremos uma comparação entre o limitante superior proposto pelo Teorema 4.2 com o resultado assintótico de limitante de capacidade ergódica (o qual será apresentado na sequência), no caso de canais cujos elementos são independentes. Desta forma, seguiremos basicamente as mesmas considerações estudadas na Seção 2.4.2, porém, neste caso para canais que seguem o desvanecimento de Rice.

Portanto, considerando que as entradas do canal \mathbf{H} são independentes, tem-se que as entradas de \mathbf{W} são também independentes. Então, pela lei forte dos grandes números (MAGALHÃES, 2004), para $n_T \rightarrow \infty$, obtém-se pela Equação (4.8)

$$\frac{1}{n_T} \mathbf{W} \mathbf{W}^H = \frac{1}{n_T} \frac{K+1}{\Omega} \mathbf{H} \mathbf{H}^H \rightarrow \frac{K+1}{\Omega} \mathbf{I}_{n_R}. \quad (4.53)$$

Desta maneira, a capacidade ergódica apresentada em (t5), resume-se a

$$\bar{C} = \mathbb{E} \{ \log_2 \det (\mathbf{I}_{n_R} + \rho \mathbf{I}_{n_R}) \}, \quad (4.54)$$

ou simplesmente,

$$\bar{C} = n_R \log_2(1 + \rho), \quad (4.55)$$

expressão esta equivalente à apresentada em (2.55).

Observemos, portanto, que a partir do caso assintótico $n_T \rightarrow \infty$, temos que a capacidade ergódica é rigorosamente igual ao limitante C_{up} proposto pelo Teorema 4.2, quando $\Omega = 1$. Noutros termos, quando as potências médias das entradas do canal \mathbf{H} são unitárias.

4.5 Comparações com dados da literatura: caso do limitante superior

Nesta seção, com relação ao limitante superior à capacidade ergódica, faremos comparações do resultado obtido neste trabalho, nos casos de canais com desvanecimento de Rice e de Rayleigh, com os resultados propostos em Jin, Gao e You (2007), seguindo basicamente as mesmas hipóteses.

Jin, Gao e You (2007, Teorema 4) investigou o limitante superior do canal MIMO com desvanecimento de Rice seguindo, basicamente, as condições levantadas na Seção 4.1.1. No entanto, a metodologia proposta foi a partir do desenvolvimento da matriz complexa não-central de Wishart. Seguindo, portanto, as notações e considerações desta dissertação, a expressão proposta por Jin, Gao e You (2007), para o posto da matriz $\bar{\mathbf{H}}$ igual a 1, é dada pela seguinte expressão:

$$C_{U_1} = \log_2 \left[\sum_{i=0}^{n_R} \binom{n_R}{i} \left(\frac{\rho}{n_T} \right)^i \frac{n_T!}{(n_T - i)!} \frac{1 + iK}{(K + 1)^i} \right]. \quad (4.56)$$

Consequentemente, no caso do canal com desvanecimento de Rayleigh tem-se

$$C_{U_1}^{K=0} = \log_2 \left[\sum_{i=0}^{n_R} \binom{n_R}{i} \left(\frac{\rho}{n_T} \right)^i \frac{n_T!}{(n_T - i)!} \right]. \quad (4.57)$$

O próximo teorema mostra que o limitante superior C_{U_1} obtido por Jin, Gao e You (2007) não excede o limitante proposto por este trabalho, diante da hipótese de que $\Omega = 1$ e posto $(\bar{\mathbf{H}}) = 1$.

Teorema 4.4 Se $\Omega = 1$ e posto $(\bar{\mathbf{H}}) = 1$, então $C_{U_1} \leq C_{\text{up}}$.

Demonstração: Primeiramente, observar-se que

$$\sum_{i=0}^{n_R} \binom{n_R}{i} \left(\frac{\rho}{n_T}\right)^i \frac{n_T!}{(n_T-i)!} \frac{1+iK}{(K+1)^i} = \sum_{i=0}^{n_R} \binom{n_R}{i} \rho^i \frac{n_T!}{(n_T)^i (n_T-i)!} \frac{1+iK}{(K+1)^i}. \quad (4.58)$$

Como $\frac{n_T!}{(n_T)^i (n_T-i)!} \frac{1+iK}{(K+1)^i} \leq 1$, para todo i inteiro não negativo (ver prova no Apêndice C), tem-se

$$\binom{n_R}{i} \rho^i \frac{n_T!}{(n_T)^i (n_T-i)!} \frac{1+iK}{(K+1)^i} \leq \binom{n_R}{i} \rho^i \quad (4.59)$$

e, conseqüentemente,

$$\begin{aligned} \sum_{i=0}^{n_R} \binom{n_R}{i} \rho^i \frac{n_T!}{(n_T)^i (n_T-i)!} \frac{1+iK}{(K+1)^i} &\leq \sum_{i=0}^{n_R} \binom{n_R}{i} \rho^i \\ &= (1+\rho)^{n_R}. \end{aligned} \quad (4.60)$$

Sabendo-se que $\log_2(\cdot)$ é uma função crescente segue que

$$\begin{aligned} C_{U_1} = \log_2 \left[\sum_{i=0}^{n_R} \binom{n_R}{i} \left(\frac{\rho}{n_T}\right)^i \frac{n_T!}{(n_T-i)!} \frac{1+iK}{(K+1)^i} \right] &\leq \log_2 (1+\rho)^{n_R} \\ &= n_R \log_2(1+\rho) \end{aligned} \quad (4.61)$$

Portanto, para $\Omega = 1$, tem-se

$$C_{U_1} \leq C_{\text{up}}, \quad (4.62)$$

qualquer que seja $K \geq 0$. □

As figuras que se seguem, ilustram o resultado do Teorema 4.4 admitindo-se canais MIMO de desvanecimento de Rayleigh e de Rice, com fator $K = 0$ e $K = 3$, respectivamente, cujas dimensões são 1×1 , 2×2 , 4×4 e 8×8 .

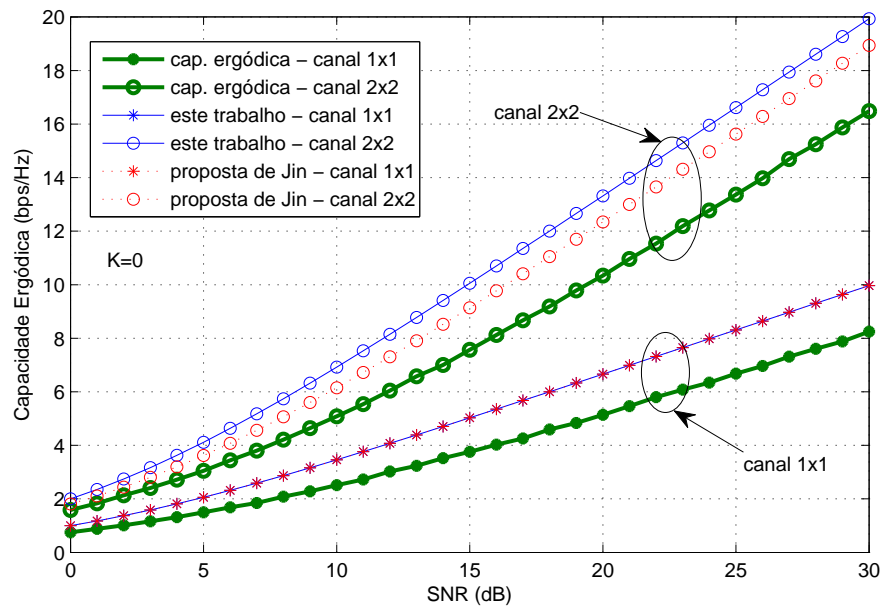


Figura 21: Comparação com limitantes superiores de Jin *et. al* (2007) para os cenários 1×1 e 2×2 , com $K = 0$.

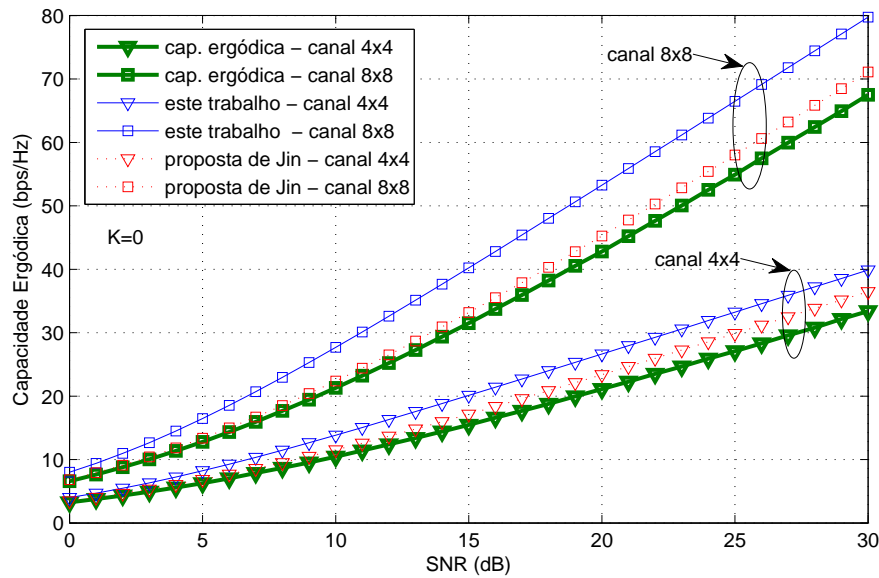


Figura 22: Comparação com limitantes superiores de Jin *et. al* (2007) para os cenários 4×4 e 8×8 , com $K = 0$.

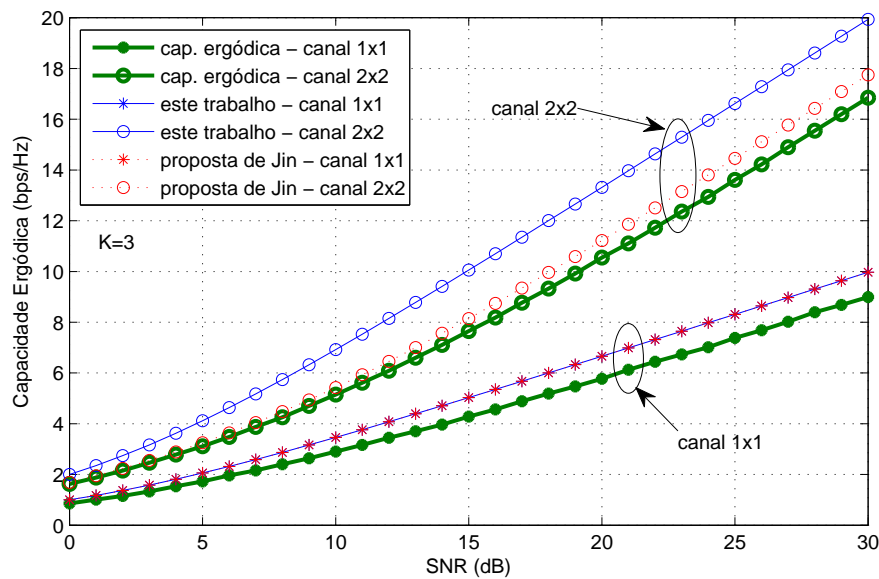


Figura 23: Comparação com limitantes superiores de Jin *et. al* (2007) para os cenários 1×1 e 2×2 , com $K = 3$.

As Figuras 21, 22, 23 e 24 enfatizam que, para um dado valor fixado da relação sinal-ruído, o limitante de capacidade proposta por Jin, Gao e You (2007), não excede o limitante superior obtido por este trabalho.

Nota-se também, para um valor de SNR fixado e diante um mesmo cenário, que os valores de capacidade são mais próximos quando o canal considerado tem desvanecimento Rayleigh. Além disso, pela proposta tomada como referência, o crescimento do fator

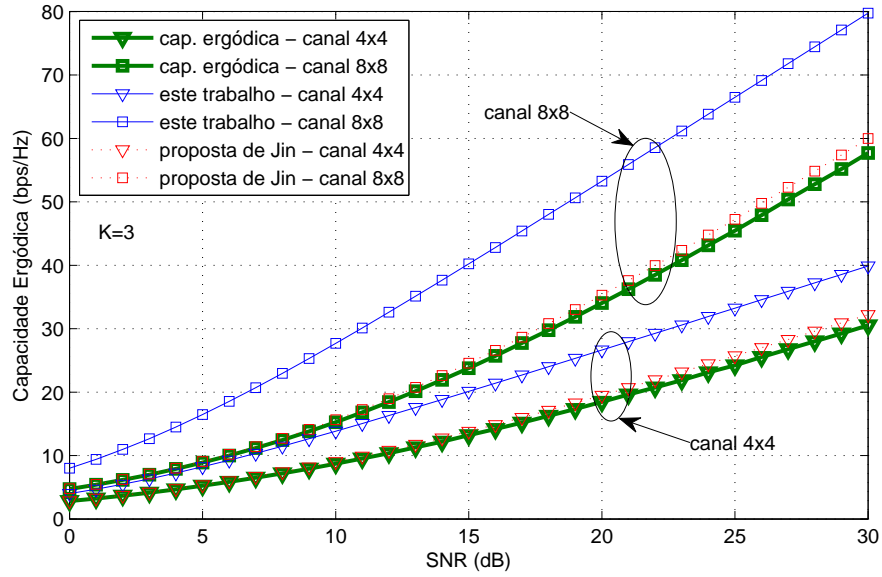


Figura 24: Comparação com limitantes superiores de Jin *et. al* (2007) para os cenários 4×4 e 8×8 , com $K = 3$.

de Rice implica no decrescimento do limitante de capacidade, situação esta não proporcionada por este trabalho, o qual preserva os limitantes independentemente do valor da constante K .

4.6 Conclusões e Síntese

Este capítulo mostrou a potencialidade da teoria de majorização na obtenção de limitantes, inferior e superior, à capacidade ergódica de um canal MIMO com desvanecimento de Rice, evitando desta forma, a aplicação direta da distribuição estatística não-central de Wishart, a qual é comumente encontrada na literatura referente a este estudo.

Especificamente, para o desenvolvimento deste capítulo, com o propósito de minimizar os cálculos envolvidos, foi feita uma transformação da matriz do canal de \mathbf{H} para \mathbf{W} . Desta forma, mediante algumas outras manipulações matemáticas, foram obtidos limitantes à capacidade com a aplicação da teoria de majorização a qual nos possibilitou obter relações entre determinados vetores via função Schur-côncava. Além disso, em alto e baixo regimes de SNRs foram apresentados aproximações à capacidade ergódica.

Por fim, no caso do limitante superior, foram feitas algumas comparações com resultados conhecidos da literatura, no entanto, foi mostrado que estes não excedem o limitante proposto por este trabalho o qual independe do fator de Rice do número de antenas transmissoras envolvidas.

CAPÍTULO 5

CONCLUSÕES E PERSPECTIVAS

Esta dissertação foi dedicada a obtenção de limitantes, inferior e superior, e aproximações à capacidade de um canal MIMO com desvanecimento de Rice, diante de uma ferramenta matemática ainda pouco explorada, chamada de teoria de majorização. No entanto, a motivação principal para a utilização desta técnica reside no fato da dificuldade de se obter expressões fechadas à capacidade ergódica, bem como, expressões para limitantes. Esta dificuldade, como ilustra a literatura sobre o assunto, se deve, principalmente, ao desenvolvimento da função de distribuição conjunta dos autovalores da matriz de Wishart não-central.

Ao longo do nosso estudo foi considerado um canal MIMO, com desvanecimento de Rice, espacialmente descorrelacionado com matriz da componente de linha de visada de posto igual a um valor arbitrário não excedente a quantidade mínima do número de antenas transmissoras e receptoras. Além disso, o receptor foi considerado com perfeito conhecimento dos estados do canal, enquanto o transmissor sem qualquer tipo de informação. Ademais, determinamos uma outra decomposição da matriz do canal, em função da matriz de linha de visada e da matriz espacialmente branca, com o propósito de minimizar os cálculos envolvidos.

Diante das considerações tomadas e com a aplicação da teoria de majorização, a qual nos possibilitou definir funções Schur-côncava específicas aplicadas em vetores previamente majorizados, foi possível a obtenção de tais limitantes e aproximações à capacidade ergódica, sendo neste último caso, a investigação foi dada a partir de baixo e altos regimes

de SNRs. Das expressões obtidas, no caso do limitante superior, foi visto que tal resultado independe do número de antenas receptoras consideradas no sistema e da constante de Rice, enquanto, no caso do limitante inferior, a expressão independe do número de antenas receptoras. Vale ressaltar que em ambas as situações, a simulação numérica é facilmente implementada e no caso do limitante superior tem-se uma expressão determinística, ou seja, esta não depende de entes estatísticos.

Como consequência dos resultados analíticos, foram feitas algumas simulações numéricas em termos dos limitantes. Nos estudos apresentados, notou-se que as curvas da capacidade ergódica se distanciam das curvas dos limitantes, tanto no caso do limitante superior com no inferior, com acréscimo da relação sinal-ruído. No entanto, obtivemos excelentes aproximações à capacidade nas situações de altos e baixos regimes SNR para os cenários apresentados.

Ainda em termos computacionais, fizemos uma comparação do limitante superior deste trabalho com um resultado já consagrado na literatura, o qual obteve um limitante superior utilizando uma outra ferramenta matemática. Para tal comparação, consideramos os mesmos cenários, além disso, restringimos o posto da matriz componente de linha de visada ser igual 1, com a finalidade de se enquadrar com as hipóteses levantadas pela nossa referência. Contudo, observou-se que nossa proposta também limita superiormente a proposta sugerida por nossa referência. No entanto, o limitante obtido por nosso trabalho tem uma expressão mais simples e depende de menos parâmetros do que o apresentado como referência.

Destacamos, portanto, como principais contribuições desta dissertação: o entendimento da teoria de majorização; a definição de funções Schur-côncavas e/ou Schur-convexas específicas que possam ser aplicadas em determinados vetores majorizados, com a finalidade de transformar a relação de majorização numa relação de desigualdade; a aplicação desta em problemas de comunicação móvel; a obtenção de limitantes e aproximações à capacidade ergódica em canal MIMO.

No que se refere às contribuições deste trabalho, pode-se considerar também os teoremas propostos sobre problemas de otimização os quais se enquadram em problemas envolvendo capacidade de canal.

Como perspectivas de estudos futuros visualizamos, nesta mesma linha de pesquisa, a obtenção de limitantes da capacidade ergódica de um canal MIMO com desvanecimento de Rice, cujos elementos são correlacionados, como também considerando outras distribuições estatísticas para as entradas do canal: distribuição Gama, Beta, Log-Normal, Weibull,

dentre outras.

Neste contexto, outros assuntos também apresentam potenciais perspectivas de investigação, tais como: o estudo de capacidade em canais do tipo *broadcast* e *relay* e a determinação de funções de custo para o tratamento de problemas de otimização.

APÊNDICE A

TEORIA DE MAJORIZAÇÃO: OUTRAS CONSIDERAÇÕES

Neste apêndice serão detalhadas algumas considerações levantadas no Capítulo 3, a respeito da teoria de majorização. Em A.1, verificaremos que o conjunto $\mathcal{D}_n = \{\mathbf{x} \in \mathbb{R}^n; x_1 \geq x_2 \geq \dots \geq x_n\}$ é convexo e que a igualdade entre vetores é satisfeita, isto é, $\mathbf{x} = \mathbf{y}$, quando ocorre as relações de majorização $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$ em \mathcal{D}_n . Na seção A.2, provaremos a propriedade (p8) das relações básicas de majorização. Por fim, em A.3 mostraremos que uma determinada classe de composição de funções preserva a propriedade de Schur-convexidade.

A.1 Um estudo sobre o conjunto \mathcal{D}_n

Primeiramente, verifiquemos que o conjunto \mathcal{D}_n é convexo. Boyd e Vandenberghe (2004), um conjunto \mathcal{C} é chamado de *conjunto convexo* se para quaisquer $\mathbf{u}, \mathbf{v} \in \mathcal{C}$ e para qualquer $t \in [0, 1]$, tem-se $t\mathbf{u} + (1-t)\mathbf{v} \in \mathcal{C}$. Desta forma, dados arbitrariamente $\mathbf{x} = (x_1, x_2, \dots, x_n)$ e $\mathbf{y} = (y_1, y_2, \dots, y_n)$ em \mathcal{D}_n e $t \in [0, 1]$, mostremos que $t\mathbf{x} + (1-t)\mathbf{y} \in \mathcal{D}_n$.

De fato, como $x_i \geq x_{i+1}$, para $i = 1, 2, \dots, n-1$, então

$$tx_i \geq tx_{i+1}. \tag{A.1}$$

Analogamente, segue que

$$(1-t)y_i \geq (1-t)y_{i+1}. \quad (\text{A.2})$$

Diante das inequações apresentadas em (A.1) e (A.2), tem-se que

$$tx_i + (1-t)y_i \geq tx_{i+1} + (1-t)y_{i+1}, \quad (\text{A.3})$$

para $i = 1, 2, \dots, n-1$. Noutros termos, $t\mathbf{x} + (1-t)\mathbf{y} \in \mathcal{D}_n$ o que comprova o resultado.

Agora, mostremos que se $\mathbf{x} \prec \mathbf{y}$ e $\mathbf{y} \prec \mathbf{x}$ em \mathcal{D}_n , então $\mathbf{x} = \mathbf{y}$. De fato, como $\mathbf{x} \prec \mathbf{y}$, segue que

$$\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i, \quad 1 \leq k \leq n-1, \quad (\text{A.4})$$

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i. \quad (\text{A.5})$$

Por sua vez, como $\mathbf{y} \prec \mathbf{x}$, tem-se

$$\sum_{i=1}^k y_i \leq \sum_{i=1}^k x_i, \quad 1 \leq k \leq n-1, \quad (\text{A.6})$$

$$\sum_{i=1}^n y_i = \sum_{i=1}^n x_i. \quad (\text{A.7})$$

Assim, das inequações apresentadas em (A.4) e (A.6) obtém-se

$$\sum_{i=1}^k y_i = \sum_{i=1}^k x_i, \quad 1 \leq k \leq n-1, \quad (\text{A.8})$$

ou seja, $x_j = y_j$, para $j = 1, 2, \dots, n-1$. Basta, portanto, mostrar que $x_n = y_n$. Da igualdade em (A.5), observar-se que

$$\sum_{i=1}^{n-1} x_i + x_n = \sum_{i=1}^{n-1} y_i + y_n. \quad (\text{A.9})$$

Como $x_j = y_j$, para $j = 1, 2, \dots, n-1$, segue que $\sum_{i=1}^{n-1} x_i = \sum_{i=1}^{n-1} y_i$ e, conseqüentemente, na igualdade apresentada em (A.9) obtém-se $x_n = y_n$. Portanto, $x_j = y_j$, para $j = 1, 2, \dots, n$, ou seja, $\mathbf{x} = \mathbf{y}$.

A.2 Demonstração da propriedade (p8) das relações básicas de majorização

A referida propriedade propõem que

$$\frac{X}{n} \mathbf{1} \prec \mathbf{x}, \text{ se } x_i \geq 0 \text{ e } \sum_{i=1}^n x_i = X. \quad (\text{A.10})$$

Demonstração: Pela definição de relação de majorização, devemos mostrar que

$$k \frac{X}{n} \leq \sum_{i=1}^k x_{[i]}, \quad 1 \leq k \leq n-1, \quad (\text{A.11})$$

$$X = \sum_{i=1}^n x_{[i]}. \quad (\text{A.12})$$

Suponhamos por absurdo que exista $k \in \{1, 2, \dots, n-1\}$, tal que

$$k \frac{X}{n} > \sum_{i=1}^k x_{[i]}. \quad (\text{A.13})$$

Diante desta inequação segue que

$$\begin{aligned} kX &> n \sum_{i=1}^k x_{[i]} \\ &\geq n \sum_{i=1}^n x_{[i]} \\ &= nX. \end{aligned} \quad (\text{A.14})$$

Ou seja, $k > n$, um absurdo. Logo, a condição dada em (A.11) é verdadeira. Por outro lado,

$$X = \sum_{i=1}^n x_i = \sum_{i=1}^n x_{[i]}, \quad (\text{A.15})$$

o que comprova a igualdade dada na Equação (A.12). Logo, a relação de majorização proposta em (p8) é satisfeita.

A.3 Composição de funções: um caso específico de preservação da propriedade de Schur-convexidade

Seja a seguinte composição de funções,

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n \xrightarrow{h} \mathbb{R}$$

$\psi = h \circ f$

em que, $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$ e $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \dots, n$. Então, pela composição, temos que

$$\psi(\mathbf{x}) = h(f(\mathbf{x})) = h(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x})). \quad (\text{A.16})$$

Agora, suponha que exista uma função real $g : \mathbb{R} \rightarrow \mathbb{R}$, tal que

$$f_i(\mathbf{x}) = f_i(x_1, x_2, \dots, x_n) = g(x_i), \quad \text{para } i = 1, 2, \dots, n. \quad (\text{A.17})$$

Contudo, se $h(\cdot)$ for Schur-convexa e crescente e $g(\cdot)$ for convexa, então a função composta $\psi(\cdot)$ também é Schur-convexa. De fato, dados os vetores \mathbf{x} e \mathbf{y} em \mathbb{R}^n , tais que $\mathbf{x} \prec \mathbf{y}$. Pelo Teorema 3.2, existe uma matriz $\mathbf{A} = [a_{ij}]$, de ordem $n \times n$, duplamente estocástica, tal que $\mathbf{x} = \mathbf{A}\mathbf{y}$. Noutros termos, cada coordenada do vetor \mathbf{x} é combinação convexa das coordenadas do vetor \mathbf{y} , ou seja,

$$x_i = \sum_{j=1}^n a_{ij} y_j, \quad \text{para cada } i = 1, 2, \dots, n. \quad (\text{A.18})$$

Como $g(\cdot)$ é uma função convexa, então pela desigualdade de Jensen, tem-se

$$g(x_i) \leq \sum_{j=1}^n a_{ij} g(y_j), \quad \text{para cada } i = 1, 2, \dots, n. \quad (\text{A.19})$$

Assim,

$$(g(x_1), g(x_2), \dots, g(x_n)) \leq \left(\sum_{j=1}^n a_{1j} g(y_j), \sum_{j=1}^n a_{2j} g(y_j), \dots, \sum_{j=1}^n a_{nj} g(y_j) \right), \quad (\text{A.20})$$

e como $h(\cdot)$ é uma função crescente, segue que

$$h(g(x_1), g(x_2), \dots, g(x_n)) \leq h\left(\sum_{j=1}^n a_{1j}g(y_j), \sum_{j=1}^n a_{2j}g(y_j), \dots, \sum_{j=1}^n a_{nj}g(y_j)\right). \quad (\text{A.21})$$

Por outro lado, como $(g(y_1), g(y_2), \dots, g(y_n)) \in \mathbb{R}^n$ e $\mathbf{A} = [a_{ij}]$ é uma matriz duplamente estocástica, então pelo Teorema 3.1, garante-se que

$$\mathbf{A} \cdot \begin{bmatrix} g(y_1) \\ g(y_2) \\ \vdots \\ g(y_n) \end{bmatrix} \prec \begin{bmatrix} g(y_1) \\ g(y_2) \\ \vdots \\ g(y_n) \end{bmatrix}, \quad (\text{A.22})$$

ou seja,

$$\left(\sum_{j=1}^n a_{1j}g(y_j), \sum_{j=1}^n a_{2j}g(y_j), \dots, \sum_{j=1}^n a_{nj}g(y_j)\right) \prec (g(y_1), g(y_2), \dots, g(y_n)). \quad (\text{A.23})$$

Por sua vez, como a função é Schur-convexa, tem-se

$$h\left(\sum_{j=1}^n a_{1j}g(y_j), \sum_{j=1}^n a_{2j}g(y_j), \dots, \sum_{j=1}^n a_{nj}g(y_j)\right) \leq h(g(y_1), g(y_2), \dots, g(y_n)). \quad (\text{A.24})$$

Portanto, das inequações apresentadas em (A.21) e (A.24) obtém-se

$$\psi(\mathbf{x}) = h(g(x_1), g(x_2), \dots, g(x_n)) \leq h(g(y_1), g(y_2), \dots, g(y_n)) = \psi(\mathbf{y}), \quad (\text{A.25})$$

isto é, $\psi(\mathbf{x}) \leq \psi(\mathbf{y})$, o que conclui que a função $\psi(\cdot)$ é Schur-convexa.

APÊNDICE B

COMPLEMENTO DA DEMONSTRAÇÃO DO TEOREMA 4.1

Este apêndice tem por finalidade mostrar a seguinte igualdade:

$$\mathbb{E} \left\{ \prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right\} = (1 + \rho\Omega)^{n_R}. \quad (\text{B.1})$$

Para tal, primeiramente, definamos os conjuntos $\mathcal{N} = \{1, 2, \dots, n\} \subset \mathbb{N}$, \mathcal{P} o conjunto das partes de \mathcal{N} e $P_i(\mathcal{N}) \subset \mathcal{P}$ cujos elementos são todas as combinações de i termos tomados n a n . Por exemplo, se $n = 2$, tem-se $P_1(\mathcal{N}) = \{\{1\}, \{2\}\}$ e $P_2(\mathcal{N}) = \{\{1, 2\}\}$. No caso de $n = 3$, segue que $P_1(\mathcal{N}) = \{\{1\}, \{2\}, \{3\}\}$, $P_2(\mathcal{N}) = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ e $P_3(\mathcal{N}) = \{\{1, 2, 3\}\}$ e assim sucessivamente.

Voltando, portanto, ao nosso problema, tem-se $\mathcal{N} = \{1, 2, \dots, n_R\}$ e para simplificar os cálculos que se seguem, façamos $\alpha = \frac{\rho\Omega}{n_T(K+1)}$ e $\Delta = \mathbb{E} \left\{ \prod_{i=1}^{n_R} \left(1 + \frac{\rho\Omega}{n_T(K+1)} d_i \right) \right\}$. Então, ao desenvolver a esperança na Equação (B.1) tem-se

$$\Delta = 1 + \alpha \sum_{\{i\} \in P_1(\mathcal{N})} \mathbb{E}\{d_i\} + \alpha^2 \sum_{\{i,j\} \in P_2(\mathcal{N})} \mathbb{E}\{d_i d_j\} + \dots + \alpha^{n_R} \sum_{\{i_1, i_2, \dots, i_{n_R}\} \in P_{n_R}(\mathcal{N})} \mathbb{E}\{d_{i_1} d_{i_2} \dots d_{i_{n_R}}\} \quad (\text{B.2})$$

Sabe-se da Seção 4.1 que

- i. $\mathbb{E}\{|w_{ij}|^2\} = K + 1$.
- ii. $d_i = \sum_{j=1}^{n_T} |w_{ij}|^2$ para $i = 1, 2, \dots, n_R$.

iii. $\mathbb{E}\{d_i\} = n_T(K + 1)$

Então, diante destes resultados obtêm-se as seguintes igualdades

$$\begin{aligned}
\mathbb{E}\{d_{i_1}d_{i_2}\cdots d_{i_p}\} &= \mathbb{E}\left\{\sum_{j_1=1}^{n_T}|w_{i_1j_1}|^2\sum_{j_2=1}^{n_T}|w_{i_2j_2}|^2\cdots\sum_{j_p=1}^{n_T}|w_{i_pj_p}|^2\right\} \\
&= \mathbb{E}\left\{\sum_{j_1=1}^{n_T}\sum_{j_1=2}^{n_T}\cdots\sum_{j_p=1}^{n_T}|w_{i_1j_1}|^2|w_{i_2j_2}|^2\cdots|w_{i_pj_p}|^2\right\} \\
&= \sum_{j_1=1}^{n_T}\sum_{j_1=2}^{n_T}\cdots\sum_{j_p=1}^{n_T}\mathbb{E}\{|w_{i_1j_1}|^2|w_{i_2j_2}|^2\cdots|w_{i_pj_p}|^2\} \\
&= \sum_{j_1=1}^{n_T}\sum_{j_1=2}^{n_T}\cdots\sum_{j_p=1}^{n_T}(K + 1)^p \\
&= (n_T(K + 1))^p.
\end{aligned} \tag{B.3}$$

Desta forma,

$$\begin{aligned}
\Delta &= 1 + \binom{n_R}{1}\alpha n_T(K + 1) + \binom{n_R}{2}\alpha^2(n_T(K + 1))^2 + \cdots + \\
&\quad + \binom{n_R}{n_R}\alpha^{n_R}(n_T(K + 1))^{n_R}
\end{aligned} \tag{B.4}$$

Como $\alpha = \frac{\rho\Omega}{n_T(K+1)}$, então

$$\begin{aligned}
\alpha^k(n_T(K + 1))^k &= \left(\frac{\rho\Omega}{n_T(K + 1)}\right)^k (n_T(K + 1))^k \\
&= (\rho\Omega)^k.
\end{aligned} \tag{B.5}$$

Portanto, o somatório na Equação (B.4) fica simplesmente dado por

$$\begin{aligned}
\Delta &= 1 + \binom{n_R}{1}\rho\Omega + \binom{n_R}{2}(\rho\Omega)^2 + \cdots + \binom{n_R}{n_R}(\rho\Omega)^{n_R} \\
&= \sum_{i=0}^{n_R} \binom{n_R}{i}(\rho\Omega)^i \\
&= (1 + \rho\Omega)^{n_R}.
\end{aligned} \tag{B.6}$$

Conclui-se finalmente que $\Delta = (1 + \rho\Omega)^{n_R}$. \square

APÊNDICE C

COMPLEMENTO DA DEMONSTRAÇÃO DO TEOREMA 4.4

Este apêndice, o qual consiste em duas seções, complementa a prova do Teorema 4.4. A primeira seção consiste em mostrar que $\frac{n_T!}{(n_T)^i(n_T-i)!} \leq 1$, enquanto na segunda seção, prova-se que $\frac{1+iK}{(K+1)^i} \leq 1$, para qualquer i inteiro não-negativo e para qualquer $K \geq 0$.

Diante, portanto, das duas demonstrações conclui-se que $\frac{n_T!}{(n_T)^i(n_T-i)!} \frac{1+iK}{(K+1)^i} \leq 1$.

C.1 Parte 1

Seja $n \in \mathbb{Z}_+$ e $i = 0, 1, 2, \dots, n$, então

$$\frac{n!}{n^i(n-i)!} \leq 1. \quad (\text{C.1})$$

A prova será pelo princípio de indução finita aplicado em i . Para $i = 0$ a expressão dada é exatamente igual a 1. Agora, suponhamos que a referida desigualdade seja válida (hipótese de indução). Provemos, portanto, que para $i + 1$ tem-se:

$$\frac{n!}{n^{i+1}(n-(i+1))!} \leq 1. \quad (\text{C.2})$$

De fato, notemos que

$$\begin{aligned} \frac{n!}{n^{i+1}(n-(i+1))!} &= \frac{n!}{n^{i+1}((n-i)-1)!} \\ &= \frac{n!(n-i)}{n^{i+1}(n-i)((n-i)-1)!} \\ &= \frac{n-i}{n} \frac{n!}{n^i(n-i)!}. \end{aligned} \tag{C.3}$$

Pela hipótese de indução tem-se $\frac{n!}{n^i(n-i)!} \leq 1$, desta forma,

$$\begin{aligned} \frac{n-i}{n} \frac{n!}{n^i(n-i)!} &\leq \frac{n-i}{n} \\ &\leq 1. \end{aligned} \tag{C.4}$$

Portanto,

$$\frac{n!}{n^{i+1}(n-(i+1))!} \leq 1 \tag{C.5}$$

o que conclui a prova. \square

C.2 Parte 2

Seja $K \in \mathbb{R}_+$ e $i = 0, 1, 2, \dots, n$, então

$$\frac{1+iK}{(K+1)^i} \leq 1. \tag{C.6}$$

Novamente fazendo-se uso do princípio de indução finita, tem-se, para $i = 0$ a expressão dada é igual a 1. Agora, pela hipótese de indução, supõem-se que a desigualdade dada seja válida para i . Basta provar, portanto, que para $i + 1$, tem-se

$$\frac{1+(i+1)K}{(K+1)^{i+1}} \leq 1, \tag{C.7}$$

ou equivalentemente,

$$1+(i+1)K \leq (K+1)^{i+1}. \tag{C.8}$$

Obervemos que

$$\begin{aligned} 1+(i+1)K &= (1+iK) + K \\ &\leq (K+1)^i + K, \end{aligned} \tag{C.9}$$

sendo esta última desigualdade devido a hipótese de indução. No entanto,

$$(K + 1)^i + K \leq (K + 1)^{i+1} \quad (\text{C.10})$$

para todo $i = 0, 1, 2, \dots, n$. De fato, se existisse algum i tal que

$$(K + 1)^i + K > (K + 1)^{i+1} \quad (\text{C.11})$$

teríamos

$$\frac{K}{(K + 1)^i} + 1 > K + 1, \quad (\text{C.12})$$

ou seja,

$$\frac{K}{(K + 1)^i} > K. \quad (\text{C.13})$$

Se $K = 0$, tem-se que $0 > 0$, um absurdo. Logo $K > 0$, desta forma,

$$(K + 1)^i < 1, \quad (\text{C.14})$$

que também é um absurdo. Tem-se portanto que

$$\frac{K}{(K + 1)^i} + 1 > K + 1 \quad (\text{C.15})$$

o que prova o resultado. \square

REFERÊNCIAS BIBLIOGRÁFICAS

ALMERS, P.; BONEK, E.; BURR, A.; CZINK, N.; DEBBAH, M.; DEGLI-ESPOSTI, V.; HOFSTETTER, H.; KYÖSTI, P.; LAURENSEN, D.; MATZ, G.; MOLISCH, A. F.; OESTGES, C.; ÖZCELİK, H. Survey of Channel and Radio Propagation Models for Wireless MIMO Systems. *EURASIP Journal on Wireless Communications and Networking*, 2007.

BERGMAN, S.; JÄRMYR, S.; OTTERSTEN, B.; JORSWIECK, E. Optimization with skewed majorization constraints: Application to MIMO systems. In: *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*. Cannes: [s.n.], 2008. p. 1–6.

BHATIA, R. *Matrix Analysis*. New York: Springer, 1997.

BIGLIERI, E.; CALDERBANK, R.; CONSTANTINIDES, A.; GOLDSMITH, A.; PAULRAJ, A.; POOR, H. V. *MIMO Wireless Communications*. New York: Cambridge University Press, 2007.

BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. New York: Cambridge University Press, 2004.

CAVALCANTE, C. C.; CAVALCANTI, F. R. P.; FREITAS JR., W. da C. Sistemas de Comunicações Móveis Multiportadoras e Multiantenas: Processamento de Sinais e Alocação de Recursos. In: . Rio de Janeiro: Mini-curso apresentado no XXVI Simpósio Brasileiro de Telecomunicações (SBrT2008), 2008.

COVER, T. M.; THOMAS, J. A. *Elements of Information theory*. New Jersey: John Wiley & Sons, Inc., 1991.

- FOSCHINI, G. J.; GANS, M. J. On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas. *Wireless Pers. Commun.*, v. 6, p. 311–335, 1998.
- FRAIDENRAICH., G.; LEVEQUE, O.; CIOFFI, J. M. On the MIMO Channel Capacity for the Nakagami- m Channel. *IEEE Transactions on Information Theory*, v. 54, n. 8, p. 3752–3757, Aug. 2008.
- GOLDSMITH, A. *Wireless Communications*. New York: Cambridge University Press, 2005.
- HARDY, G. H.; LITTLEWOOD, L. E.; PÓLYA, G. *Inequalities*. London: Cambridge University Press, 1934.
- HARVILLE, D. A. *Matrix Algebra From a Statistician's Perspective*. New York: Springer, 2008.
- HAYKIN, S. *Adaptive Filter Theory*. 4. ed. New Jersey: Prentice Hall, 2001.
- HAYKIN, S.; MOHER, M. *Sistemas modernos de comunicações wireless*. Porto Alegre: Bookman, 2008.
- HORN, R. A.; JOHNSON, C. R. *Matrix Analysis*. New York: Cambridge University Press, 1985.
- JAYAWEERA, S. K.; POOR, H. V. On the Capacity of Multiple Antenna Systems in Rician Fading. *IEEE Trans. Commun.*, v. 4, n. 3, p. 1102–1111, 2005.
- JIANG, Y.; PALOMAR, D. P.; VARANASI, M. K. Precoder Optimization for Nonlinear MIMO Transceiver Based on Arbitrary Cost Function. In: *Information Sciences and Systems, 2007. CISS '07. 41st Annual Conference on*. Baltimore, MD: [s.n.], 2007. p. 119–124.
- JIN, S.; GAO, X.; YOU, X. On the Ergodic Capacity of Rank-1 Ricean-Fading MIMO Channels. *IEEE Transactions on Information Theory*, v. 53, n. 2, p. 502–517, Feb. 2007.
- JORSWIECK, E.; BOCHE, H. Majorization and Matrix-Monotone Functions in Wireless Communications. *Found. Trends Commun. Inf. Theory*, Found. Trends Commun. Inf. Theory, v. 3, n. 6, p. 553–701, 2007.

- KHALIGHI, M.; BROSSIER, J.-M.; JOURDAIN, G.; RAOOF, K. On Capacity of Rician MIMO channels. In: *Personal, Indoor and Mobile Radio Communications, 2001 12th IEEE International Symposium on*. San Diego, CA: [s.n.], 2001. v. 1, p. A-150-A-154.
- LEBRUN, G.; FAULKNER, M.; SHAFI, M.; SMITH, P. J. MIMO Ricean Channel Capacity: An Asymptotic Analysis. *IEEE Trans. Wireless Commun.*, v. 5, n. 6, p. 1343-1350, Jun. 2006.
- LEON-GARCIA, A. *Probability and Random Process for Electrical Engineering*. 2. ed. New York: Addison Wesley, 1994.
- LIMA, E. L. *Curso de Análise*. 6. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada (IMPA), 2000.
- MAGALHÃES, M. N. *Probabilidade e Variáveis Aleatórias*. 2. ed. São Paulo: Edusp, 2004.
- MARSHALL, A. W.; OLKIN, I. *Theory of Majorization and Its Applications*. New York: Academic Press, 1979.
- MCKAY, M. R.; COLLINGS, I. B. Capacity bounds for correlated Rician MIMO channels. In: *IEEE International Conference on Communications (ICC)*. Seoul, Korea: [s.n.], 2005. v. 2, p. 772-776.
- OESTGES, C.; CLERCKX, B. *MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design*. London: Academic Press, 2007.
- PALOMAR, D. P. Convex Primal Decomposition for Multicarrier Linear MIMO Transceivers. *IEEE Transactions on Signal Processing*, v. 53, n. 12, p. 4661-4674, Dec. 2005.
- PALOMAR, D. P.; CIOFFI, J. M.; LAGUNAS, M. A. Uniform Power Allocation in MIMO Channels: A Game-Theoretic Approach. *IEEE Transactions on Information Theory*, v. 49, n. 7, p. 1707-1727, 2003.
- PALOMAR, D. P.; JIANG, Y. MIMO Transceiver Design via Majorization Theory. *Found. Trends Commun. Inf. Theory*, Found. Trends Commun. Inf. Theory, v. 3, n. 4-5, p. 331-551, 2006.
- PALOMAR, D. P.; LAGUNAS, M. A.; CIOFFI, J. M. Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints. *IEEE Transactions on Signal Processing*, v. 52, n. 5, p. 1179-1197, May. 2004.

- PAPOULIS, A. *Probability, Random Variables and Stochastic Processes*. 3. ed. New York: McGraw-Hill, Inc., 1991.
- PAULRAJ, A.; NABAR, R.; GORE, D. *Introduction to Space-Time Wireless Communications*. New York: Cambridge University Press, 2003.
- PROAKIS, J. *Digital Communications*. 4. ed. New York: McGraw-Hill Education, 1983.
- SHANNON, C. E. A Mathematical Theory of Communication. *The Bell System Technical Journal*, v. 27, p. 379–423 and 623–656, 1948.
- _____. Communication in the presence of noise. *Proceeding of the IRE and waves and electrons*, p. 10–21, 1949.
- TELATAR, I. E. Capacity of Multi-Antenna Gaussian Channels. *Europ. Trans. Telecommun.*, v. 10, n. 6, p. 585–595, Nov. 1999.
- WINTERS, J. On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment. *IEEE Journal on Selected Areas on Communications*, v. 5, p. 871–878, 1987.
- YU, K.; BENGTSSON, M.; OTTERSTEN, B.; MCNAMARA, D.; KARLSSON, P.; BEACH, M. Modeling of Wide-Band MIMO Radio Channels Based on NLoS Indoor Measurements. *IEEE Trans. on Vehicular Technology*, v. 53, n. 3, p. 655–665, May. 2004.
- ZARBOUTI, D.; TSOULOS, G.; KAKLAMANI, D. Theory and Practice of MIMO Wireless Communications Systems. In: TSOULOS, G. (Ed.). *MIMO System Technology for Wireless Communications*. [S.l.]: Taylor & Francis Group, 2006.
- ZHANG, X.; PALOMAR, D. P.; OTTERSTEN, B. Statistically Robust Design of Linear MIMO Transceivers. *IEEE Transactions on Signal Processing*, v. 56, n. 8, p. 3678–3689, Aug. 2008.
- ZHONG, C.; WONG, K.-K.; JIN, S. Capacity Bounds for MIMO Nakagami- m Fading Channels. *IEEE Transactions on Signal Processing*, v. 57, n. 9, p. 3613–3623, Sep. 2009.
- ZHOU, X.; LAMAHEWA, T. A.; SADEGHI, P.; DURRANI, S. Capacity of MIMO systems: Impact of spatial correlation with channel estimation errors. In: *11th IEEE Singapore International Conference on Communication Systems, ICCS 2008*. Guangzhou: [s.n.], 2008. p. 817–822.