

DISSERTAÇÃO DE MESTRADO

**DESENVOLVIMENTO DE ESTRATÉGIA DE PRÉ-CODIFICAÇÃO
E DE PREDIÇÃO DE CANAL PARA O ENLACE DIRETO
DE SISTEMAS MIMO MULTIUSUÁRIO**

João Paulo Leite

Brasília, Agosto de 2009

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**DESENVOLVIMENTO DE ESTRATÉGIA DE
PRÉ-CODIFICAÇÃO E DE PREDIÇÃO DE CANAL
PARA O ENLACE DIRETO DE SISTEMAS MIMO
MULTIUSUÁRIO**

JOÃO PAULO LEITE

ORIENTADOR: PAULO HENRIQUE PORTELA DE CARVALHO

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

PUBLICAÇÃO: PPGENE.DM - 394A/09

BRASÍLIA/DF: AGOSTO - 2009

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

DESENVOLVIMENTO DE ESTRATÉGIA DE PRECODIFICAÇÃO E
DE PREDIÇÃO DE CANAL PARA O ENLACE DIRETO DE
SISTEMAS MIMO MULTIUSUÁRIO

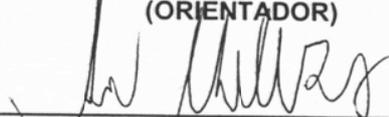
JOÃO PAULO LEITE

DISSERTAÇÃO DE MESTRADO ACADÊMICO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE
BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE.

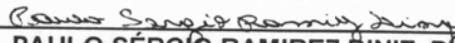
APROVADA POR:



ROBSON DOMINGOS VIEIRA, Dr., ENE/UNB
(ORIENTADOR)



ANDRÉ NOLL BARRETO, Dr., ENE/UNB
(EXAMINADOR INTERNO)



PAULO SÉRGIO RAMIREZ DINIZ, Dr., COPPE/UFRJ
(EXAMINADOR EXTERNO)

BRASÍLIA, 03 DE AGOSTO DE 2009.

FICHA CATALOGRÁFICA

LEITE, JOÃO PAULO

Desenvolvimento de Estratégia de Pré-codificação e de Predição de Canal para o Enlace Direto de Sistemas MIMO Multiusuário [Distrito Federal] 2009.

xx, 111p., 210 x 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2009).

Dissertação de Mestrado - Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Elétrica.

- | | |
|------------------------------|-------------------------|
| 1. Sistemas MIMO | 2. Pré-codificação |
| 3. Predição de Canal | 4. Filtragem Adaptativa |
| 5. Comunicações Multiusuário | 6. Enlace Direto |
| I. ENE/FT/UnB | II. Título (série) |

REFERÊNCIA BIBLIOGRÁFICA

LEITE, J. P. (2009). Desenvolvimento de Estratégia de Pré-codificação e de Predição de Canal para o Enlace Direto de Sistemas MIMO Multiusuário. Dissertação de Mestrado em Engenharia Elétrica, Publicação PPGENE.DM - 394A/09, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 111p.

CESSÃO DE DIREITOS

AUTOR: João Paulo Leite.

TÍTULO: Desenvolvimento de Estratégia de Pré-codificação e de Predição de Canal para o Enlace Direto de Sistemas MIMO Multiusuário.

GRAU: Mestre

ANO: 2009

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

João Paulo Leite
SMBS chácara 02 casa 01
71080-000, Guará II - DF - Brasil.

Dedicatória

Aos meus pais, João e Rosângela, que forneceram todas as condições para que eu pudesse chegar aonde cheguei, e nunca mediram esforços para me oferecer uma formação de qualidade.

Aos meus avós, Jorge e Rosa, pelo exemplo de vida.

Ao meu irmão, Pedro, por seu companheirismo.

João Paulo Leite

Agradecimentos

Agradeço ...

Ao professor Dr. Paulo Henrique Portela de Carvalho, pela oportunidade de trabalho e pela paciência durante os vários anos de convivência. Sua dedicação e cordialidade foram fundamentais para meu crescimento acadêmico.

Ao professor Dr. Robson Domingos Vieira, por me apresentar o fascinante universo das comunicações por múltiplas antenas. Este trabalho não seria possível sem seu conhecimento, disposição e amizade.

Aos funcionários do Departamento de Engenharia Elétrica (ENE), pelo convívio agradável e presteza no atendimento.

Aos colegas do Laboratório de Estruturas de Micro-ondas e Ondas Milimétricas (LEMOM), em especial ao Alex Helder e ao Charles Costa, pela amizade e pelos momentos de distração.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo suporte financeiro.

A Deus, por ter me permitido grandes oportunidades, e por ter colocado em meu caminho pessoas verdadeiramente iluminadas e inspiradas.

João Paulo Leite

RESUMO

Técnicas de comunicação que utilizam múltiplas antenas (MIMO, *multiple-input, multiple-output*) são uma importante área de pesquisa, que têm atraído considerável atenção. Em especial, os sistemas MIMO multiusuário são capazes de combinar o aumento de capacidade proporcionado por sistemas MIMO e, quando informação sobre o estado de canal (CSI, *channel state information*) está disponível no transmissor, as propriedades espaciais do canal podem ser exploradas por meio da pré-codificação, de forma que os diversos usuários podem ser multiplexados espacialmente. Entretanto, a natureza variante no tempo do canal rádio móvel não é levada em consideração ao analisar o desempenho dos esquemas de pré-codificação.

Este trabalho apresenta como primeira contribuição uma proposta de pré-codificador para o enlace direto de sistemas MIMO multiusuário que utiliza o conceito de perturbação contínua. Os resultados de simulação mostram que a técnica proposta apresenta desempenho superior ao das técnicas lineares atualmente relatadas na literatura.

A segunda contribuição diz respeito à influência da informação desatualizada sobre o estado do canal no desempenho das técnicas de pré-codificação. É proposto um preditor adaptativo de canal que utiliza o algoritmo *set-membership affine projection* (SM-AP) como forma de compensar a desatualização da CSI causada pela variação temporal do canal de comunicação. Os resultados mostram que o preditor SM-AP, além de exibir baixo custo computacional, apresenta desempenho superior aos algoritmos clássicos NLMS (*normalized least mean square*) e RLS (*recursive least squares*).

ABSTRACT

Multiple-input, multiple-output communication techniques are an important area of research, and have attracted considerable attention. In particular, multiuser MIMO systems are capable of offering the higher link capacity of MIMO systems and, when channel state information (CSI) is available at the transmitter, the spatial properties of the channel can be exploited by precoding, so that multiple users can be spatially multiplexed. However, the time-varying nature of the mobile-radio channel is not taken into account when analyzing the performance of the precoding schemes.

The first contribution of this work is the proposal of a precoding technique based on the concept of continuous perturbation. Simulation results show that the proposed technique outperforms the linear precoding techniques currently reported in the literature.

The second contribution concerns the influence of the outdated CSI in the performance of the precoding techniques. An adaptive channel predictor based on the set-membership affine projection (SM-AP) filtering is proposed as a way to compensate the outdated CSI caused by the temporal variations of the communication channel. The results show that the SM-AP predictor has low computational cost and it has superior performance when compared to the classical algorithms NLMS (normalized least mean square) and RLS (recursive least squares).

SUMÁRIO

1	INTRODUÇÃO	1
1.1	CONTEXTUALIZAÇÃO	1
1.2	DEFINIÇÃO DO PROBLEMA.....	2
1.3	OBJETIVOS	3
1.4	ORGANIZAÇÃO DO MANUSCRITO.....	3
2	ANÁLISE DE CAPACIDADE DE SISTEMAS MIMO MULTIUSUÁRIO	5
2.1	INTRODUÇÃO	5
2.2	MODELO DO SISTEMA	6
2.3	CAPACIDADE	10
2.3.1	CAPACIDADE DO CANAL MIMO MAC	10
2.3.2	CODIFICAÇÃO DIRTY PAPER.....	11
2.3.3	CAPACIDADE DO CANAL MIMO BC.....	12
2.3.4	DUALIDADE MAC-BC	13
2.3.5	SOMAS DAS CAPACIDADES	14
2.4	CONCLUSÃO.....	16
3	TÉCNICAS DE PRÉ-CODIFICAÇÃO LINEARES E NÃO LINEARES PARA O CANAL MIMO BC	17
3.1	INTRODUÇÃO	17
3.2	MODELO DE SISTEMA	18
3.3	ZERO-FORCING	20
3.4	REGULARIZAÇÃO DE INVERSA.....	23
3.5	BLOCK DIAGONALIZATION	27
3.6	PRÉ-CODIFICAÇÃO TOMLINSON-HARASHIMA	31
3.6.1	MODELO DO SISTEMA	31
3.6.2	PRÉ-CODIFICAÇÃO ZERO-FORCING TOMLINSON-HARASHIMA	35
3.6.3	PRÉ-CODIFICAÇÃO MINIMUM MEAN SQUARE ERROR TOMLINSON-HARASHIMA ..	36
3.6.4	A ORDENAÇÃO DOS USUÁRIOS	40
3.7	VECTOR PERTURBATION	43
3.8	MMSE VECTOR PERTURBATION.....	49
3.9	PROPOSTA DE PRÉ-CODIFICADOR	53
3.10	INFLUÊNCIA DA INFORMAÇÃO IMPERFEITA DE ESTADO DE CANAL SOBRE A PRÉ-CODIFICAÇÃO.....	55
3.11	CONCLUSÃO.....	56
4	ALGORITMOS DE FILTRAGEM ADAPTATIVA	58
4.1	INTRODUÇÃO	58
4.2	SOLUÇÃO DE WIENER.....	59

4.3	FILTROS ADAPTATIVOS	61
4.3.1	MÉTODO DO GRADIENTE DESCENDENTE E MÉTODO DE NEWTON	62
4.4	ALGORITMO LMS	63
4.5	ALGORITMO NLMS	64
4.5.1	ALGORITMO AFFINE PROJECTION	64
4.6	ALGORITMO RLS	67
4.7	ALGORITMO SET-MEMBERSHIP AFFINE PROJECTION	68
4.7.1	FILTRAGEM SET-MEMBERSHIP	68
4.7.2	SET-MEMBERSHIP AFFINE PROJECTION	69
4.7.3	ESCOLHA DO VETOR DE PARÂMETROS	70
4.8	PROPOSTA DE PREDITOR DE CANAL	72
4.8.1	MODELO DO SISTEMA	72
4.8.2	METODOLOGIA DE PREDIÇÃO	74
4.9	CONCLUSÃO.....	76
5	RESULTADOS DE SIMULAÇÕES.....	77
5.1	INTRODUÇÃO	77
5.2	CENÁRIOS DE SIMULAÇÃO.....	77
5.3	DESEMPENHOS DOS PREDITORES DE CANAL	79
5.4	DESEMPENHO DOS PRÉ-CODIFICADORES	81
5.4.1	RESULTADOS DE REFERÊNCIA	82
5.4.2	CENÁRIOS COM PREDITORES DE CANAL	86
5.5	CONCLUSÃO.....	94
6	CONCLUSÕES.....	95
6.1	TRABALHOS FUTUROS	96
	REFERÊNCIAS BIBLIOGRÁFICAS.....	97
	ANEXOS.....	106
I	LONG TERM EVOLUTION.....	107
II	MODELO DE CANAL SCM.....	110

LISTA DE FIGURAS

2.1	Canais de <i>downlink</i> e <i>uplink</i> em um sistema celular.	5
2.2	Matrizes de canal para o cenário MIMO multiusuário.	6
2.3	Diagrama de blocos para os canais (a) MIMO BC e (b) MIMO MAC.	7
2.4	Região de Capacidade do canal MIMO MAC com $n_{Rk} = 1, k = 1, 2$	11
2.5	Codificação <i>dirty paper</i> para canal escalar monousuário.	11
2.6	Codificação de K usuários utilizando codificação <i>dirty paper</i>	12
2.7	Região de capacidade do canal MIMO BC para $M_T = 2$ e $K = 2$. Os valores numéricos foram obtidos considerando $\mathbf{H}_1 = [1 \ 0,4]$ e $\mathbf{H}_2 = [0,4 \ 1]$ para $P_T = 10$. Modificado de [1]. ...	15
2.8	Soma das capacidades do canal MIMO BC.	15
3.1	Ilustração da região de capacidade de um sistema multiusuário considerando o problema <i>near-far</i> . A soma das capacidades pode penalizar determinados usuários, dependendo da forma assumida pela região de capacidade.	19
3.2	Hiato entre a soma das capacidades da pré-codificação <i>zero-forcing</i> (ZF) e o limite dado pelo DPC para $K = 10$ usuários.	21
3.3	Comparação da soma de capacidades em função de K da pré-codificação <i>zero-forcing</i> e o limite teórico oferecido pelo DPC, para uma razão sinal-ruído de $\rho = 10$ dB.	22
3.4	Comportamento dos quatro maiores autovalores de $(\mathbf{H}\mathbf{H}^H)^{-1}$ em função de K	23
3.5	Comparação da soma das capacidades em função de K das técnicas de pré-codificação <i>zero-forcing</i> e regularização de inversa para uma relação sinal-ruído de $\rho = 10$ dB.	25
3.6	Comparação da soma das capacidades em função de ρ das técnicas de pré-codificação <i>zero-forcing</i> e regularização de inversa para $K = 10$ usuários.	25
3.7	Comparação das curvas de BER dos esquemas de pré-codificação por inversão de canal e regularização de inversa para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.	26
3.8	Comparação das curvas de BER dos esquemas de pré-codificação por inversão de canal e regularização de inversa para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.	27
3.9	Comparação da soma das capacidades em função de ρ da técnica de pré-codificação <i>block diagonalization</i> para $M_T = 6$ antenas transmissoras e diferentes configurações de número de usuários e de antenas nos receptores.	30
3.10	Comparação da soma das capacidades em função de ρ da técnica de pré-codificação <i>block diagonalization</i> para $M_T = 10$ antenas transmissoras e diferentes configurações de número de usuários e de antenas nos receptores.	30
3.11	Estrutura do pré-codificador Tomlinson-Harashima para o enlace direto de sistemas MIMO multiusuário.	31
3.12	Implementação do operador módulo por meio de uma função dente de serra.	32
3.13	Descrição linearizada do operador módulo.	32
3.14	Constelação QPSK estendida. É mostrada a constelação original, \mathcal{A} e algumas de suas extensões periódicas utilizando $\tau = 4$. Os símbolos módulo-congruentes são representados pelos mesmos marcadores.	33

3.15	Função do operador módulo.	34
3.16	Comparação das curvas de BER dos esquemas de pré-codificação ZF THP e MMSE THP para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.	38
3.17	Comparação das curvas de BER dos esquemas de pré-codificação ZF THP e MMSE THP para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.	38
3.18	Comparação das técnicas de pré-codificação ZF THP e MMSE THP com os algoritmos lineares em um cenário com $K = 10$ usuários e modulação QPSK.	39
3.19	Comparação das técnicas de pré-codificação ZF THP e MMSE THP com os algoritmos lineares em um cenário com $K = 10$ usuários e modulação 16QAM.	39
3.20	Desempenho dos pré-codificadores ZF THP e MMSE THP com ordenação de usuários dada pelo algoritmo <i>best-first</i>	42
3.21	Interpretação geométrica da técnica de pré-codificação <i>vector perturbation</i>	45
3.22	Comparação das curvas de BER dos esquemas de pré-codificação para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.	46
3.23	Comparação das curvas de BER dos esquemas de pré-codificação para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.	46
3.24	Interpretação geométrica da redução de reticulado utilizando duas bases diferentes. A primeira base, mais ortogonal, é mostrada em (a), e outro exemplo de base é mostrado em (b).	49
3.25	Comparação das curvas de BER dos esquemas de pré-codificação com $K = 4$ usuários utilizando modulação QPSK.	51
3.26	Comparação das curvas de BER dos esquemas de pré-codificação com $K = 10$ usuários utilizando modulação QPSK.	51
3.27	Comparação das curvas de BER dos esquemas de pré-codificação com $K = 4$ usuários utilizando modulação 16QAM.	52
3.28	Comparação das curvas de BER dos esquemas de pré-codificação com $K = 10$ usuários utilizando modulação 16QAM.	52
3.29	Desempenho do pré-codificador proposto comparado aos algoritmos lineares de pré-codificação para a modulação QPSK.	54
3.30	Desempenho do pré-codificador proposto comparado aos algoritmos lineares de pré-codificação para a modulação 16QAM.	55
3.31	Comparação das curvas de BER dos esquemas de pré-codificação <i>vector perturbation</i> e regularização de inversa sob CSI imperfeita para $K = 4$ usuários utilizando modulação 16QAM.	56
4.1	Forma direta de implementação de um filtro FIR.	58
4.2	Diagrama de blocos do filtro de Wiener. Os sinais $s(n)$ e $d(n)$ são estatisticamente relacionados, e o objetivo do filtro é produzir a estimativa $r(n)$ de $d(n)$ que possui o menor erro quadrático médio.	59
4.3	Estrutura geral de um filtro adaptativo.	62
4.4	Hiperplanos associados à atualização dos coeficientes do filtro adaptativo utilizando o algoritmo <i>affine projection</i> com $N_P = 2$	67

4.5	Interpretação geométrica do algoritmo <i>set-membership affine projection</i> que resulta em erro <i>a posteriori</i> constante.....	71
4.6	Estrutura adaptativa aplicada à predição de sinal.	72
5.1	Ambiente utilizado nas simulações, destacando (a) o MIMO BC e a presença do canal de <i>feedback</i> e (b) a estrutura do sistema receptor.	78
5.2	Convergência dos preditores adaptativos para um horizonte de predição de 1 símbolo OFDM.	80
5.3	Convergência dos preditores adaptativos para um horizonte de predição de 28 símbolos OFDM.	80
5.4	MSE de Predição versus horizonte de predição.	81
5.5	Capacidade de rastreamento dos preditores SM-AP, NLMS e RLS em um ambiente não estacionário. As características do canal mudam a cada 100 símbolos OFDM.	82
5.6	Desempenho do pré-codificador <i>zero-forcing</i> sem a utilização de um preditor de canal.	83
5.7	Desempenho do pré-codificador <i>vector perturbation</i> sem a utilização de um preditor de canal.	84
5.8	Desempenho do pré-codificador proposto sem a utilização de um preditor de canal.	85
5.9	Desempenho do algoritmo de pré-codificação <i>zero-forcing</i> utilizando a modulação QPSK. ..	86
5.10	Desempenho do algoritmo de pré-codificação <i>zero-forcing</i> utilizando a modulação 16QAM.	87
5.11	Desempenho do algoritmo de pré-codificação <i>vector perturbation</i> utilizando a modulação QPSK.	88
5.12	Desempenho do algoritmo de pré-codificação <i>vector perturbation</i> utilizando a modulação 16QAM.	89
5.13	Desempenho do algoritmo de pré-codificação proposto utilizando a modulação QPSK.	90
5.14	Desempenho do algoritmo de pré-codificação proposto utilizando a modulação 16QAM.	91
5.15	Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando a técnica de pré-codificação <i>zero-forcing</i> e as modulações QPSK e 16QAM.	92
5.16	Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando o algoritmo de pré-codificação proposto e as modulações QPSK e 16QAM.	93
5.17	Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando a técnica de pré-codificação <i>vector perturbation</i> e as modulações QPSK e 16QAM.	93
I.1	Estrutura de quadro no enlace direto do padrão LTE para uma largura de banda de 10 MHz.	107
I.2	Alocação de subportadoras que carregam símbolos pilotos no enlace direto do padrão LTE para quatro antenas de transmissão.....	109
II.1	Modelo espacial do 3GPP para simulações MIMO.	110
II.2	Exemplo de realização do canal para diferentes velocidades do móvel.	111

LISTA DE TABELAS

5.1	Especificação dos parâmetros do enlace direto do padrão 3GPP-LTE.....	77
5.2	Especificação dos valores de parâmetros para o modelo de canal WINNER's SCM.	79
5.3	Cenários para a simulação do desempenho dos pré-codificadores.....	82

LISTA DE SÍMBOLOS

Siglas

3GPP	<i>Third Generation Partnership Project</i>
AP	<i>Affine Projection</i>
AWGN	<i>Additive White Gaussian Noise</i>
BC	<i>Broadcast Channel</i>
BD	<i>Block Diagonalization</i>
BER	<i>Bit Error Rate</i>
CDMA	<i>Code-division Multiple Access</i>
CSI	<i>Channel State Information</i>
DFE	<i>Decision Feedback Equalizer</i>
DFT	<i>Discrete Fourier Transform</i>
DPC	<i>Dirty Paper Coding</i>
FDD	<i>Frequency Division Duplexing</i>
ESPRIT	<i>Estimation of Signal Parameters via Rotational Invariance Techniques</i>
FIR	<i>Finite Impulse Response</i>
GSM	<i>Global System for Mobile Communications</i>
IMT-A	<i>International Mobile Telecommunications - Advanced</i>
LLL	<i>Lenstra-Lenstra-Lovász</i>
LMMSE	<i>Linear Minimum Mean Square Error</i>
LMS	<i>Least Mean Square</i>
LTE	<i>Long Term Evolution</i>
MAC	<i>Multiple Access Channel</i>
MIMO	<i>Multiple-input, Multiple-output</i>
MISO	<i>Multiple-input, Single-output</i>
MMSE	<i>Minimum Mean Square Error</i>
MPAM	<i>M-ary Pulse Amplitude Modulation</i>
MUI	<i>Multiuser Interference</i>
NLMS	<i>Normalized Least Mean Square</i>
OFDM	<i>Orthogonal Frequency-Division Multiplexing</i>
OFDMA	<i>Orthogonal Frequency-Division Multiple Access</i>
QAM	<i>Quadrature Amplitude Modulation</i>
QPSK	<i>Quaternary Phase Shift Keying</i>
SDMA	<i>Space-Division Multiple Access</i>
SIC	<i>Successive Interference Cancellation</i>
SINR	<i>Signal-to-interference-plus-noise Ratio</i>
SISO	<i>Single-input, Single-output</i>
SNR	<i>Signal-to-noise Ratio</i>
RI	<i>Regularized Inverse</i>
RLS	<i>Recursive Least Squares</i>
SCM	<i>Spatial Channel Model</i>
SM-AP	<i>Set-Membership Affine Projection</i>
SM-APVDR	<i>Set-Membership Affine Projection with Variable Data-reuse</i>
SM-PUAP	<i>Set-Membership Partial Update Affine Projection</i>

TDD	<i>Time Division Duplexing</i>
TDMA	<i>Time-division Multiple Access</i>
THP	<i>Tomlinson-Harashima Precoding</i>
TTI	<i>Transmission Time Interval</i>
UMTS	<i>Universal Mobile Telecommunications System</i>
V-BLAST	<i>Vertical-Bell Laboratories Layered Space-Time</i>
VP	<i>Vector Perturbation</i>
WiMAX	<i>Worldwide Interoperability for Microwave Access</i>
WLAN	<i>Wireless Local Area Network</i>
WSSUS	<i>Wide Sense Stationary Uncorrelated Scattering</i>
ZF	<i>Zero-Forcing</i>
ZMSW	<i>Zero-Mean Spatially White</i>

Símbolos Matemáticos

\mathbf{A}^T	Transposto da matriz \mathbf{A}
\mathbf{A}^*	Conjugado da matriz \mathbf{A}
\mathbf{A}^H	Hermitiano (conjugado transposto) da matriz \mathbf{A}
$\text{Tr}(\mathbf{A})$	Traço da matriz \mathbf{A}
\mathbf{A}^{-1}	Inversa da matriz \mathbf{A}
\mathbf{A}^\dagger	Pseudo-inversa da matriz \mathbf{A}
$\lfloor a \rfloor$	Truncamento, maior inteiro que não seja maior que a
\mathbb{E}	Operador média
$\ \mathbf{A}\ ^2$	Norma de Frobenius da matriz \mathbf{A}
$ \mathbf{A} $	Determinante da matriz \mathbf{A}
$\text{Re}(a)$	Parte real do número complexo a
$\text{Co}()$	Região convexa
\log	Logaritmo na base 10
$\text{diag}(a_1, \dots, a_n)$	Matriz diagonal de dimensão $n \times n$ cujos elementos são $a_1 \dots a_n$
j	Unidade imaginária, $\sqrt{-1}$
\mathbb{C}	Conjunto dos números complexos
\mathbf{I}	Matriz identidade
$\mathbf{0}$	Matriz identicamente nula
e	Número de Euler, $e = 2,718281828 \dots$

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

As comunicações sem fio apresentaram, nas últimas décadas, um desenvolvimento muito grande. Um dos motivos que impulsionou tal desenvolvimento foi o crescimento das comunicações móveis pessoais. As novas aplicações e serviços são cada vez mais elaborados e específicos, e demandam grande velocidade de transmissão e melhor qualidade de serviço. A citar como exemplo, os sistemas de telefonia móvel de quarta geração, denominados IMT-A (*International Mobile Telecommunications - Advanced*), encontram-se em fase de especificação pela União Internacional de Telecomunicações e prevêem taxas de transmissão de até 1 Gbps para usuários de baixa mobilidade, e de até 100 Mbps para usuários de alta mobilidade, e possuem como premissa a otimização da utilização do espectro eletromagnético [2].

Os requisitos impostos pelas novas gerações dos sistemas de comunicações móveis não podem ser satisfeitos simplesmente aumentando-se a largura de banda destinada à transmissão (pois o espectro eletromagnético é um recurso caro e escasso) ou aumentando-se a potência de transmissão (devido a limitações dos equipamentos de rádio e ao aumento proibitivo dos níveis de interferência). Neste contexto, as comunicações que utilizam MIMO (do inglês, *multiple-input multiple-output*) são uma importante área de pesquisa para os sistemas de próxima geração devido ao seu potencial para promover o aumento da capacidade do enlace de transmissão (por meio da multiplexação espacial dos fluxos (*streams*) transmitidos) e aumentar a qualidade do sinal recebido (por exemplo, explorando diversidade por meio de códigos espaço-temporais ou espaço-frequenciais) [3].

As técnicas que visam a utilização de múltiplas antenas no transmissor e no receptor foram inicialmente investigadas e exploradas em sistemas monousuário [4]. Entretanto, para aplicações do tipo WLAN (do inglês, *wireless local area network*) ou telefonia móvel celular, nas quais uma estação rádio-base deve se comunicar com vários usuários simultaneamente, as técnicas MIMO podem ser utilizadas em sistemas multiusuário. Como consequência, o estudo de sistemas MIMO multiusuário emergiu recentemente como um importante tópico de investigação [5].

A premissa de sistemas MIMO multiusuário consiste em utilizar o domínio espacial (ou seja, os graus de liberdade introduzidos pelas múltiplas antenas) para compartilhar o canal de comunicação entre os usuários. Apesar desta técnica de múltiplo acesso envolver custos adicionais de *hardware*, como o aumento no número de antenas e filtros utilizados, ela não envolve qualquer custo adicional em largura de banda, diferentemente do que ocorre com as técnicas de múltiplo acesso por divisão do tempo, TDMA (do inglês, *time division multiple access*), ou múltiplo acesso por divisão de código, CDMA (do inglês, *code division multiple access*) [6, 7]. Apesar desse fato, os protocolos clássicos de múltiplo acesso podem ser utilizados conjuntamente com sistemas MIMO multiusuário.

Ainda que a área de comunicações multiusuário seja bem fundamentada e bastante estudada [8], a adição de múltiplas antenas nos terminais de transmissão e/ou recepção dá origem a diversas questões e problemas específicos. De acordo com o histórico levantado em [5], o desenvolvimento de técnicas

de transmissão e recepção para o enlace reverso de sistemas MIMO multiusuário surge como uma generalização dos algoritmos utilizados nos sistemas MIMO monousuário considerando a presença de múltiplos usuários no sistema. Na verdade, técnicas de detecção multiusuário que utilizam o cancelamento sucessivo de interferência [9] (SIC, do inglês, *successive interference cancellation*) são capazes de atingir a capacidade do enlace reverso de sistemas MIMO multiusuário, e a complexidade requerida é essencialmente a mesma dos algoritmos utilizados nos sistemas MIMO monousuário [10].

A situação é consideravelmente diferente para o enlace direto de sistemas MIMO multiusuário [11]. Neste, conforme mostrado em [12, 13], a estratégia ótima de transmissão (no sentido de alcançar a capacidade do canal) envolve o conceito de codificação *dirty paper* [14] para o cancelamento de interferência entre os usuários. Apesar de ótimo sob o ponto de vista de teórico, a implementação de técnicas de codificação *dirty paper* é extremamente complexa, o que estimulou o desenvolvimento de estratégias subótimas de pré-codificação lineares e não lineares. Exemplos da primeira são os algoritmos *zero-forcing* e regularização de inversa [15], e exemplos da segunda são as técnicas de pré-codificação Tomlinson-Harashima e *vector perturbation* [16, 17].

1.2 DEFINIÇÃO DO PROBLEMA

O cenário MIMO multiusuário começou a ser intensamente investigado devido a algumas vantagens que ele oferece quando comparado a sistemas MIMO monousuário. Tais sistemas possuem a capacidade de combinar o aumento de capacidade promovido por sistemas MIMO com os benefícios de supressão de interferência presentes em sistemas múltiplo acesso por divisão espacial, ou SDMA (do inglês, *space-division multiple access*). Como exemplo, pode-se afirmar que os sistemas MIMO multiusuário [18]:

- permitem ganho de múltiplo acesso proporcional ao número de antenas na estação rádio base graças aos esquemas de multiplexação multiusuário;
- apresentam maior imunidade aos problemas de propagação que afligem os sistemas MIMO monousuário, como ambientes pobres em espalhadores ou transmissão em linha de visada;
- permitem que ganhos de multiplexação espacial na estação rádio-base sejam obtidos sem a necessidade de múltiplas antenas nos receptores, tornando possível o desenvolvimento de terminais menores e mais baratos. Os custos maiores de implementação encontram-se no lado da infraestrutura do sistema.

Para que seja possível tirar proveito das vantagens supracitadas, informação precisa sobre o estado do canal, ou CSI (do inglês, *channel state information*) deve estar presente no transmissor, especialmente no enlace direto de sistemas MIMO multiusuário. Esta é uma diferença fundamental entre sistemas monousuário e multiusuário: apesar da CSI no transmissor não ser essencial em ambientes monousuário (exceto possivelmente para valores muito baixos de relação sinal-ruído) [4], a CSI é de importância crítica para qualquer aplicação em sistemas MIMO multiusuário, quer seja para aplicação de técnicas de pré-codificação e pré-cancelamento de interferência, quer seja para a utilização de algoritmos de seleção e escalonamento de usuários [19].

Grande parte dos trabalhos publicados sobre o enlace direto de sistemas MIMO multiusuário supõem que tanto transmissor quanto receptor possuem CSI perfeita. Esta pode ser obtida diretamente quando o canal de comunicação varia lentamente com o tempo, como, por exemplo, em ambientes de propagação *indoor*, mas é muito mais difícil de ser obtida em situações nas quais os usuários apresentam alta mobilidade. Apesar de várias técnicas de MIMO monousuário propostas assumirem modelos simples e quase-estáticos para as características de rádio propagação, a natureza variante no tempo do canal deve ser considerada ao tratar de sistemas MIMO multiusuário.

1.3 OBJETIVOS

A suposição mais comum ao tratar de sistemas MIMO multiusuário é quanto à presença de CSI perfeita tanto no transmissor quanto no receptor. Uma análise da penalidade imposta pela informação imperfeita ou desatualizada sobre o comportamento do canal é de grande valia para projetistas de sistemas.

Como primeiro objetivo, este trabalho propõe um esquema de pré-codificação que tem como base o conceito de perturbação contínua. Sob determinadas condições, o pré-codificador proposto apresenta desempenho melhor que as técnicas lineares apresentadas, mesmo sob a influência de CSI desatualizada.

O segundo objetivo é analisar, por meio de simulações computacionais, o comportamento do desempenho de alguns dos algoritmos de pré-codificação utilizados no enlace direto de sistemas MIMO multiusuário sob a influência de CSI imperfeita e desatualizada. Para tal, é considerado um modelo realista de canal, padronizado pelo 3GPP (*Third Generation Partnership Project*) para investigação de desempenho de sistemas MIMO, e um modelo realista de sistema, que tem como base as especificações do enlace direto do padrão 3GPP-LTE (*Long Term Evolution*).

Este trabalho também apresenta a proposta de um preditor de canal adaptativo que utiliza o algoritmo *set-membership affine projection* como forma de compensar a desatualização de CSI causada pela natureza variante no tempo do canal radio móvel. O desempenho do preditor proposto é comparado com os algoritmos clássicos utilizados em filtragem adaptativa, a citar, os algoritmos *normalized least mean square* e *recursive least squares*.

1.4 ORGANIZAÇÃO DO MANUSCRITO

A dissertação está dividida em seis capítulos. O primeiro é constituído pela presente introdução.

No capítulo 2, são apresentados os resultados fundamentais de Teoria de Informação referentes à capacidade do enlace direto e do enlace reverso de sistemas MIMO multiusuário.

O capítulo 3 trata das técnicas lineares e não lineares de pré-codificação utilizadas para cancelamento de interferência no enlace direto de sistemas MIMO multiusuário. Nesse capítulo, é também apresentada a técnica de pré-codificação proposta, e seu desempenho é comparado com o das técnicas de pré-codificação lineares apresentadas.

Em seguida, o capítulo 4 fornece uma introdução aos filtros adaptativos, apresentando os algoritmos

clássicos que são utilizados para aplicações que envolvem filtragem adaptativa. Ao final, é proposta uma estrutura de preditor de canal para sistemas MIMO multiusuário que utiliza o algoritmo *set-membership affine projection*.

Os resultados de simulação que mostram o desempenho dos pré-codificadores e preditores de canal são deixados para o capítulo 5.

O trabalho é finalizado com as conclusões, no capítulo 6, bem como com as propostas para trabalhos futuros.

2 ANÁLISE DE CAPACIDADE DE SISTEMAS MIMO MULTIUSUÁRIO

2.1 INTRODUÇÃO

Um canal multiusuário é qualquer canal que deve ser compartilhado entre diversos usuários. Na literatura [8], o canal multiusuário é classificado quanto a direção de transmissão, a citar o canal de *downlink*, também denominado canal de *broadcast* ou canal direto [20], e o canal de *uplink*, também chamado de canal de múltiplo acesso ou canal reverso [21]. A maioria dos sistemas de comunicação é bidirecional, consistindo de canal direto e canal reverso. De forma a evitar interferência entre os dois canais, estes são separados de forma ortogonal, tipicamente utilizando os domínios do tempo ou da frequência. Esta separação é denominada duplexação. Em particular, na duplexação por divisão no tempo, TDD (do inglês, *time-division duplexing*), atribuem-se diferentes *slots* no tempo para transmissão do enlace direto e do enlace reverso, e na duplexação por divisão na frequência, FDD (do inglês, *frequency-division duplexing*), atribuem-se bandas de frequência distintas para transmissão dos sinais dos enlaces direto e reverso.

No contexto de utilização de múltiplas antenas, existem dois modelos de canal MIMO multiusuário: o canal MIMO MAC (do inglês, *multiple access channel*), referente ao enlace reverso, e o canal MIMO BC (do inglês, *broadcast channel*), referente ao enlace direto.

O MIMO MAC consiste em vários transmissores com múltiplas antenas transmitindo para um receptor com múltiplas antenas. O MIMO BC consiste de um transmissor com múltiplas antenas transmitindo para vários receptores com múltiplas antenas. Em arquiteturas de redes celulares, o MIMO MAC modela o canal dos usuários móveis para a estação rádio-base, e o MIMO BC modela o canal da estação rádio-base para os dispositivos móveis. Este exemplo é mostrado na Fig. 2.1.

Contrariamente aos métodos com os quais a maioria dos sistemas são projetados, os benefícios em termos de capacidade que podem ser conseguidos transmitindo-se de forma simultânea para múltiplos usuários (no canal direto) ou múltiplos usuários transmitindo simultaneamente (no canal reverso) sem qualquer separação nos domínios do tempo, frequência ou código podem ser muito grandes, pois é possível

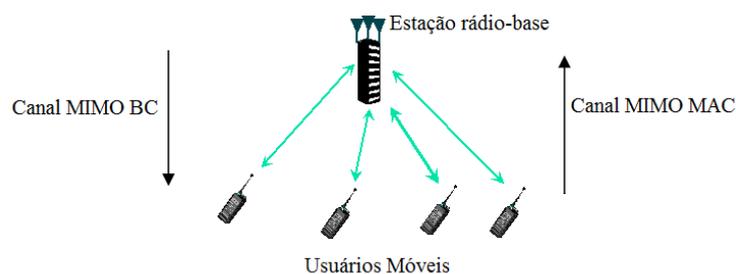


Figura 2.1: Canais de *downlink* e *uplink* em um sistema celular.

tirar proveito da dimensão espacial introduzida pela utilização de múltiplas antenas [22].

Como as tecnologias que fazem uso de múltiplas antenas devem se tornar cada dia mais comuns, a exemplo dos padrões WiMAX [23, 24] e 3GPP-LTE [25, 26], é fundamental a compreensão dos limites impostos por estes canais utilizando as ferramentas desenvolvidas em Teoria da Informação.

Este capítulo apresenta o modelo matemático que será utilizado para tratar formalmente os canais MIMO MAC e MIMO BC. Serão ainda apresentados os resultados de capacidade dos canais MIMO multiusuário.

2.2 MODELO DO SISTEMA

Para descrever os canais MIMO MAC e MIMO BC, será considerado um cenário unicelular, com uma estação rádio-base que transmite simultaneamente para K usuários ativos na célula. A estação rádio-base possui M_T antenas, o k -ésimo usuário possui n_{Rk} antenas, e o número total de antenas receptoras é dado por $N_R = \sum_{k=1}^K n_{Rk}$. Denotaremos por $\mathbf{H}_k(n) \in \mathbb{C}^{n_{Rk} \times M_T}$ a matriz na forma

$$\mathbf{H}_k(n) = \begin{bmatrix} H_{1,1}^k(n) & \cdots & H_{1,M_T}^k(n) \\ \vdots & \ddots & \vdots \\ H_{n_{Rk},1}^k(n) & \cdots & H_{n_{Rk},M_T}^k(n) \end{bmatrix}, \quad (2.1)$$

que corresponde aos ganhos do canal entre as antenas da estação rádio-base e as antenas do k -ésimo usuário em determinado símbolo n , conforme mostrado na Fig. 2.2. O elemento $h_{i,j}^k$ dessa matriz representa o ganho entre a j -ésima antena da estação rádio-base e a i -ésima antena de recepção do usuário k .

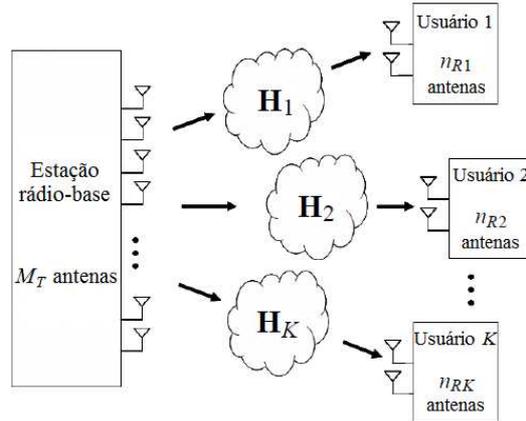


Figura 2.2: Matrizes de canal para o cenário MIMO multiusuário.

É comum assumir que os coeficientes da matriz \mathbf{H}_k são modelados por meio de variáveis aleatórias gaussianas complexas independentes e identicamente distribuídas (no espaço), de média nula e variância unitária. Esta hipótese é referida como ZMSW (do inglês, *zero-mean spatially white*), e dá origem ao desvanecimento Rayleigh [27].

Nesta dissertação, a indexação dos elementos em n será omitida quando esta for óbvia pelo contexto, especialmente nas próximas seções.

Tendo como base o modelo apresentado na Eq. (2.1), supõe-se que o canal de comunicação não é seletivo em frequência, apresentando desvanecimento plano em toda a banda de transmissão. Nos sistemas de próxima geração, este pressuposto, a princípio, não vigora. Canais de banda larga são seletivos na frequência, apresentando resposta em amplitude que pode variar muito dentro da banda ocupada pelo sistema para transmissão, gerando interferência intersimbólica. Entretanto, os sistemas de próxima geração utilizam OFDM (do inglês, *orthogonal frequency-division multiplexing*) e OFDMA (do inglês, *orthogonal frequency-division multiple access*), que dividem um canal seletivo em frequência em diversos subcanais que apresentam desvanecimento plano, nos quais o modelo apresentado é válido. Portanto, é possível implementar algoritmos MIMO de processamento separadamente para cada subportadora [22].

Por motivos de modelagem, será considerado que o canal no enlace reverso é o mesmo utilizado no enlace direto (isto é, o sistema é do tipo TDD), de forma que a matriz de ganhos do canal entre o k -ésimo usuário e a estação rádio-base é dada por \mathbf{H}_k^H . A verdadeira resposta do canal para o caso TDD no enlace reverso está relacionada a resposta no enlace direto por meio da transposição, sendo dada por \mathbf{H}_k^T . Entretanto, para simplicidade no tratamento matemático, a matriz conjugada transposta é utilizada. Esta restrição não afeta os resultados obtidos para a capacidade do canal ¹.

Além disso, supõe-se que a estação rádio-base possui M_T antenas tanto no enlace direto (onde são consideradas como antenas de transmissão) quanto no enlace reverso (onde são consideradas como antenas de recepção). O mesmo é verdade para n_{Rk} , o número de antenas presentes no k -ésimo usuário móvel, tanto no enlace reverso (onde são consideradas como antenas de transmissão) como no enlace direto (onde são consideradas como antenas de recepção).

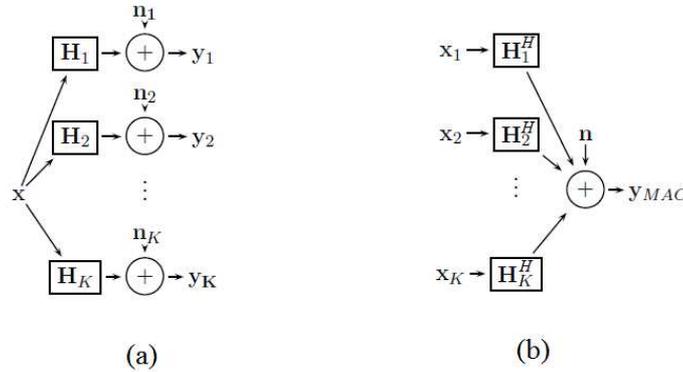


Figura 2.3: Diagrama de blocos para os canais (a) MIMO BC e (b) MIMO MAC.

A Fig. 2.3 mostra o diagrama de blocos que será utilizado para a representação dos sinais envolvidos em um sistema MIMO multiusuário. O canal MIMO MAC será representado da seguinte forma: em um dado símbolo n , o k -ésimo transmissor enviará o vetor de símbolos $\mathbf{x}_k(n) \in \mathbb{C}^{n_{Rk} \times 1}$ para o receptor. $\mathbf{y}_{MAC}(n) \in \mathbb{C}^{M_T \times 1}$ é o sinal recebido e $\mathbf{n}(n) \in \mathbb{C}^{M_T \times 1}$ é o ruído aditivo gaussiano presente no receptor. O sinal recebido é matematicamente representado por

¹É importante observar ainda que, além do caso TDD possuir tratamento matemático mais simples, é um modelo útil para introduzir o conceito de dualidade entre os canais de múltiplo acesso e de *broadcast*, a ser apresentado posteriormente. Entretanto, o sistema não precisa operar em TDD para que os resultados apresentados sejam válidos.

$$\begin{aligned}
\mathbf{y}_{MAC}(n) &= \sum_{k=1}^K \mathbf{H}_k^H(n) \mathbf{x}_k(n) + \mathbf{n}(n) \\
&= \begin{bmatrix} \mathbf{H}_1^H(n) & \dots & \mathbf{H}_K^H(n) \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(n) \\ \vdots \\ \mathbf{x}_K(n) \end{bmatrix} + \mathbf{n}(n).
\end{aligned} \tag{2.2}$$

O ruído apresenta matriz de covariância dada por $\mathcal{R}_{\mathbf{n}} \triangleq \mathbb{E} \{ \mathbf{n}(n) \mathbf{n}^H(n) \} = \sigma_n^2 \mathbf{I}$, e é, portanto, espacialmente branco.

No canal MIMO MAC, cada transmissor está sujeito a uma restrição individual quanto à sua máxima potência de transmissão P_k , isto é, definindo-se a matriz de covariância do sinal transmitido pelo usuário k como $\mathcal{R}_{\mathbf{x}_k} \triangleq \mathbb{E} [\mathbf{x}_k(n) \mathbf{x}_k^H(n)]$, tem-se $\text{Tr}(\mathcal{R}_{\mathbf{x}_k}) \leq P_k$, para $k = 1, \dots, K$.

O modelo matemático utilizado no cenário MIMO BC é o seguinte: a estação rádio-base pode desejar transmitir dados com diferentes taxas para cada um dos usuários. Isto pode ser conseguido escolhendo-se de forma apropriada a modulação a ser utilizada ou mudando o número de fluxos (*streams*) de dados que serão transmitidos para cada usuário. Neste caso, m_k denota o número de fluxos a serem transmitidos simultaneamente para o usuário k . Os valores de m_1, \dots, m_K não dependerão apenas da taxa de transmissão desejada, mas também da potência de transmissão disponível, do número de antenas de transmissão e do número de antenas de recepção em cada móvel. Tipicamente, $m_k \leq n_{Rk}$ se não for considerada nenhuma forma adicional de codificação para multiplexação espacial. Vale também a restrição $\sum_k m_k \leq M_T$. O vetor de fluxos para o usuário k será representado por $\mathbf{u}_k(n) \in \mathbb{C}^{m_k \times 1}$.

Em um instante de símbolo n , o sinal que é transmitido para o usuário k é um vetor $\mathbf{x}_k(n) \in \mathbb{C}^{M_T \times 1}$. Em muitos casos, o sinal transmitido é uma função linear dos símbolos do fluxo de dados, isto é, $\mathbf{x}_k(n) = \mathbf{W}_k \mathbf{u}_k(n)$, onde as colunas de \mathbf{W}_k , $\mathbf{W}_k = [\mathbf{w}_{1k}, \dots, \mathbf{w}_{m_k k}]$ correspondem aos pesos (*beamformers*) de cada símbolo. Também podem ser consideradas situações que empregam um mapeamento não linear \mathbf{f}_k dos símbolos aos dados efetivamente transmitidos, de forma que $\mathbf{x}_k(n) = \mathbf{f}_k(\mathbf{u}_k(n))$. O papel dos pesos e funções de mapeamento não lineares não é apenas de organizar os fluxos dos diferentes usuários nas antenas de transmissão, mas sua utilização é fundamental para o cancelamento de interferência e técnicas de pré-codificação que visam atingir a capacidade do canal, como será visto posteriormente, ao serem abordadas as estratégias de pré-codificação no canal MIMO BC.

O ruído aditivo gaussiano em cada antena receptora é representado pelo vetor $\mathbf{n}_k(n) \in \mathbb{C}^{n_{Rk} \times 1}$, cuja matriz de covariância é $\mathcal{R}_{\mathbf{n}_k} \triangleq \mathbb{E} \{ \mathbf{n}_k(n) \mathbf{n}_k^H(n) \} = \sigma_n^2 \mathbf{I}$, em que supõe-se que as estatísticas do ruído (sua média e variância) são as mesmas para todos os usuários. O sinal recebido pelo k -ésimo usuário, $\mathbf{y}_k(n) \in \mathbb{C}^{n_{Rk} \times 1}$, é matematicamente representado por

$$\begin{aligned}
\mathbf{y}_k(n) &= \sum_{i=1}^K \mathbf{H}_k(n) \mathbf{x}_i(n) + \mathbf{n}_k(n) \\
&= \mathbf{H}_k(n) \mathbf{x}_k(n) + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{H}_k(n) \mathbf{x}_i(n) + \mathbf{n}_k(n).
\end{aligned} \tag{2.3}$$

Ao observar a Eq. (2.3), fica claro que o usuário k recebe não apenas o sinal desejado $\mathbf{x}_k(n)$, mas também uma componente de interferência devido à presença dos demais usuários, denominada interferência multiusuário (MUI, do inglês, *multiuser interference*) ou interferência interusuário.

A Eq. (2.3) pode ser reescrita de forma compacta, fazendo:

$$\begin{aligned} \mathbf{y}(n) &= \begin{bmatrix} \mathbf{y}_1(n) \\ \vdots \\ \mathbf{y}_K(n) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1(n) \\ \vdots \\ \mathbf{H}_K(n) \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} \mathbf{n}_1(n) \\ \vdots \\ \mathbf{n}_K(n) \end{bmatrix} \\ &= \mathbf{H}(n)\mathbf{x}(n) + \mathbf{n}(n), \end{aligned} \quad (2.4)$$

em que

$$\mathbf{H}(n) = [\mathbf{H}_1^H(n) \dots \mathbf{H}_K^H(n)]^H \quad (2.5)$$

é a matriz combinada de canal, e representa a concatenação de todos os ganhos de canal de todos os receptores considerados, e

$$\mathbf{x}(n) = \sum_{k=1}^K \mathbf{f}_k(\mathbf{u}_k(n)) \quad (2.6)$$

é o sinal efetivamente transmitido pela estação rádio-base. Se uma combinação linear dos símbolos é transmitida, a Eq. (2.6) pode ser escrita como:

$$\begin{aligned} \mathbf{x}(n) &= \begin{bmatrix} \mathbf{W}_1 & \dots & \mathbf{W}_K \end{bmatrix} \begin{bmatrix} \mathbf{u}_1(n) \\ \vdots \\ \mathbf{u}_K(n) \end{bmatrix} \\ &= \sum_{k=1}^K \mathbf{W}_k \mathbf{u}_k(n). \end{aligned} \quad (2.7)$$

A matriz de covariância do sinal \mathbf{x} é dada por $\mathcal{R}_{\mathbf{x}} \triangleq \mathbb{E}[\mathbf{x}(n)\mathbf{x}^H(n)]$. O transmissor está sujeito a uma determinada restrição na potência máxima de transmissão, denominada P_T , o que implica em $\text{Tr}(\mathcal{R}_{\mathbf{x}}) \leq P_T$.

Supondo que a perda de percurso (*path loss*) pode ser desprezada, de forma que a variância dos coeficientes da resposta do canal seja unitária, define-se a relação sinal-ruído como a razão:

$$\rho = \frac{P_T}{\sigma_n^2} \quad (2.8)$$

2.3 CAPACIDADE

O conceito de capacidade foi introduzido por Claude Shannon em [28, 29]. Para o caso monousuário, a capacidade C do canal é a máxima taxa com a qual a informação pode ser transmitida de forma confiável. Demonstrou-se que, para qualquer taxa de transmissão $R < C$, existe um código de taxa R que pode ser utilizado para transmitir a informação com probabilidade de erro arbitrariamente pequena.

A definição de capacidade de um canal MIMO é semelhante: consiste em um número real que representa o limite para comunicação livre de erros: qualquer taxa estritamente menor que a capacidade é alcançável [30].

Em canais multiusuário, a capacidade do canal não é mais representada por um número, mas sim por uma região K -dimensional, já que diferentes taxas estão associadas aos múltiplos usuários (de acordo com a política de alocação de recursos, por exemplo). A região de capacidade é definida como o conjunto de todos os valores de taxas (R_1, R_2, \dots, R_K) que podem ser simultaneamente alcançadas, com probabilidade de erro arbitrariamente pequena, por todos os K usuários.

A seguir, serão consideradas as regiões de capacidade dos canais MIMO MAC e MIMO BC.

2.3.1 Capacidade do Canal MIMO MAC

Para um dado conjunto de restrições quanto à potência de transmissão $\mathbf{P} = (P_1, \dots, P_K)$, a capacidade do canal MIMO MAC é dada por [31, 32, 33]:

$$\mathcal{C}_{MAC}(\mathbf{P}; \sigma_n^2; \mathbf{H}^H) \triangleq \bigcup_{\substack{\mathcal{R}_{\mathbf{x}_k} \geq 0, \\ \text{Tr}(\mathcal{R}_{\mathbf{x}_k}) \leq P_k \forall k}} \left\{ (R_1, \dots, R_K) : \sum_{i \in S} R_i \leq \log_2 \left| \mathbf{I} + \sum_{i \in S} \frac{1}{\sigma_n^2} \mathbf{H}_i^H \mathcal{R}_{\mathbf{x}_k} \mathbf{H}_i \right| \forall S \subseteq \{1, \dots, K\} \right\}, \quad (2.9)$$

em que a variável S refere-se a um subconjunto de $\{1, \dots, K\}$. A região é alcançada da seguinte forma: o k -ésimo usuário transmite um vetor gaussiano aleatório \mathbf{x}_k de média nula e matriz de covariância espacial $\mathcal{R}_{\mathbf{x}_k}$. Cada conjunto de matrizes de covariância $(\mathcal{R}_{\mathbf{x}_1}, \dots, \mathcal{R}_{\mathbf{x}_K})$ corresponde a um poliedro K -dimensional, denotado por:

$$\left\{ (R_1, \dots, R_K) : \sum_{i \in S} R_i \leq \log \left| \mathbf{I} + \sum_{i \in S} \frac{1}{\sigma_n^2} \mathbf{H}_i^H \mathcal{R}_{\mathbf{x}_k} \mathbf{H}_i \right| \forall S \subseteq \{1, \dots, K\} \right\}, \quad (2.10)$$

e a região de capacidade é dada pela união (tomada sobre todas as matrizes de covariância que satisfazem à restrição de potência) de todos esses poliedros.

A Fig. 2.4 mostra a região de capacidade considerando $K = 2$ usuários, cada um com uma antena. Neste caso, as matrizes de covariância se reduzem a simples escalares iguais à restrição de potência. Para este caso, a região de capacidade é representada por um pentágono.

Determinar o conjunto de matrizes de covariância $(\mathcal{R}_{\mathbf{x}_1}, \dots, \mathcal{R}_{\mathbf{x}_K})$ que satisfaz a Eq. (2.9) é um

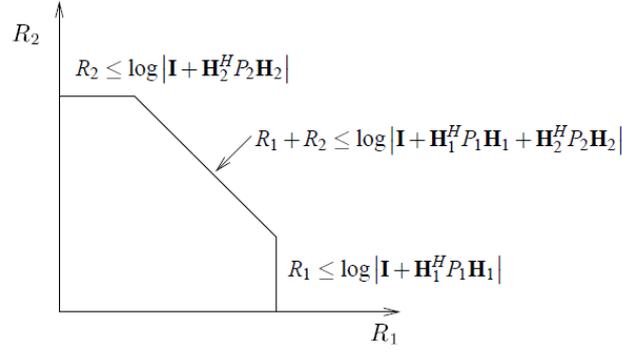


Figura 2.4: Região de Capacidade do canal MIMO MAC com $n_{Rk} = 1, k = 1, 2$.

problema de otimização convexo, e há, portanto, técnicas numéricas eficientes para resolvê-lo [34].

2.3.2 Codificação Dirty Paper

Antes de apresentar os resultados de capacidade para o canal MIMO BC, é necessário introduzir o conceito de codificação *dirty paper* (DPC, do inglês, *dirty paper coding*) [14], originado em 1983. Consideremos inicialmente o caso de um canal escalar com ruído gaussiano e interferência, representado por

$$y = x + \mathcal{I} + \bar{n}, \quad (2.11)$$

em que x é uma palavra-código utilizada para transmitir a informação u para o usuário, \mathcal{I} é a componente de interferência (de comportamento gaussiano e independente do sinal transmitido), \bar{n} representa o ruído gaussiano, de média nula e variância σ_n^2 , e y é o sinal recebido.

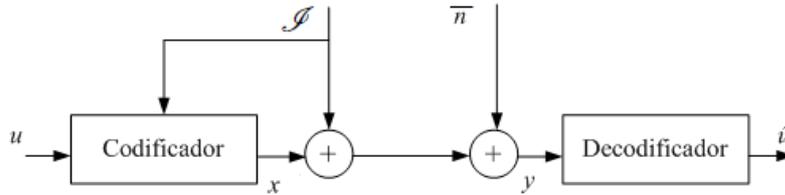


Figura 2.5: Codificação *dirty paper* para canal escalar monousuário.

Supondo que o transmissor, mas não o receptor, conhece a componente de interferência de forma não causal (isto é, antes do início da transmissão) e determinística, conforme mostrado na Fig. 2.5, e a potência média de transmissão é P_T , então a capacidade desse canal será dada por

$$C = \log_2 \left(1 + \frac{P_T}{\sigma_n^2} \right), \quad (2.12)$$

que é a mesma de um sistema no qual a componente de interferência não está presente (não importante qual seja a potência da interferência). Em outros termos, o transmissor é capaz de determinar uma palavra-código x de tal forma que o receptor não "veja" a interferência.

Em [35, 36, 37], é mostrado que o resultado obtido pode ser utilizado para um comportamento qualquer da interferência, não necessariamente gaussiano.

Apesar do resultado supor um canal do tipo escalar, a extensão para o caso MIMO é imediata, uma vez que este pode ser decomposto em vários canais SISO paralelos por meio de uma decomposição em valores singulares [38].

2.3.3 Capacidade do Canal MIMO BC

Quando o transmissor (neste caso, a estação rádio-base) possui mais de uma antena, $M_T > 1$, o canal MIMO BC é dito não degradado, isto é, os receptores não podem ser ordenados de acordo com a qualidade de seus canais, já que existirão diferentes ganhos de canal associados a cada uma das antenas de transmissão (e possivelmente de recepção, caso $n_{Rk} > 1$). Ao contrário do canal MAC, uma expressão geral para a região de capacidade do canal de *broadcast* não degradado não é conhecida. Na verdade, esta ainda é uma das perguntas fundamentais de Teoria da Informação multiusuário que ainda não foi respondida. Entretanto, em [11], é proposta uma região de taxas alcançáveis para este tipo de canal utilizando DPC, e, em [12], foi mostrado que a região proposta corresponde de fato à região de capacidade do canal MIMO BC.

A região de capacidade do canal MIMO BC é baseada no conceito de codificação *dirty paper*, cuja premissa básica é que, se o transmissor possui conhecimento perfeito e não causal da interferência gerada por determinado usuário, então é possível que a interferência seja "pré-subtraída" no transmissor de tal forma que a potência de transmissão não seja alterada [35].

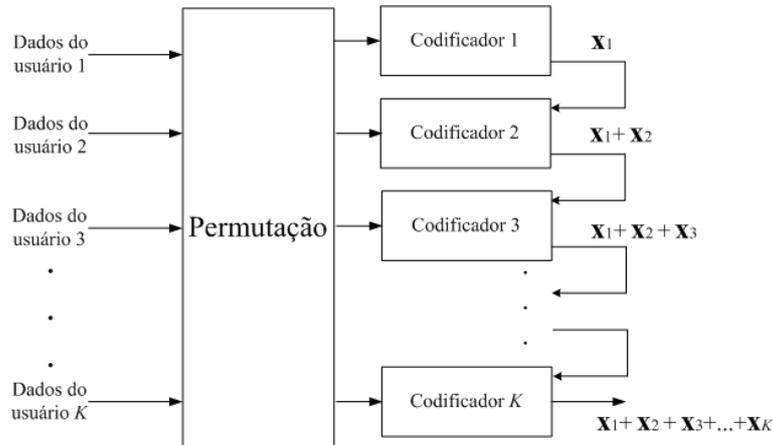


Figura 2.6: Codificação de K usuários utilizando codificação *dirty paper*.

No canal MIMO BC, o DPC pode ser aplicado no transmissor escolhendo-se apropriadamente as palavras-código dos diferentes usuários, conforme ilustrado na Fig. 2.6. Os usuários são codificados de modo sucessivo. O transmissor escolhe uma palavra-código \mathbf{x}_1 para o usuário 1. A palavra-código para o segundo usuário, \mathbf{x}_2 , será escolhida em função de \mathbf{x}_1 , de forma que a interferência gerada por \mathbf{x}_1 não seja vista pelo usuário 2. De forma similar, é escolhida uma palavra-código \mathbf{x}_3 para o usuário 3, de tal forma que ele não veja os sinais destinados aos usuários 1 e 2 como interferência. O procedimento é repetido para todos os K usuários. Ao codificar o k -ésimo usuário, o transmissor já conhece a interferência que será causada pelos usuários previamente codificados $(1, \dots, k - 1)$. Com efeito, a interferência gerada pelos

usuários pode ser subtraída do sinal que será enviado.

Como pode ser percebido, a ordem na qual os usuários são codificados exercerá influência no resultado final, e também deve ser otimizada. Seja então $\pi(\cdot)$ uma permutação dos K usuários e $\mathcal{R} = [\mathcal{R}_{\mathbf{x}_1}, \dots, \mathcal{R}_{\mathbf{x}_K}]$, o conjunto de matrizes positivas semidefinidas com $\text{Tr}(\mathcal{R}_{\mathbf{x}_1} + \dots + \mathcal{R}_{\mathbf{x}_K}) \leq P_T$, que representa as matrizes de covariância. Sob a utilização do DPC, se o usuário $\pi(1)$ é codificado primeiro, seguido do usuário $\pi(2)$, e assim por diante, então o seguinte vetor de capacidades é alcançável:

$$\mathbf{R}(\pi, \mathcal{R}) = (R_{\pi(1)}, \dots, R_{\pi(K)}) : R_{\pi(k)} = \frac{\log \left| \mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{H}_{\pi(k)} \left(\sum_{j=1}^k \mathcal{R}_{\mathbf{x}_{\pi(j)}} \right) \mathbf{H}_{\pi(k)}^H \right|}{\log \left| \mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{H}_{\pi(k)} \left(\sum_{j=1}^{k-1} \mathcal{R}_{\mathbf{x}_{\pi(j)}} \right) \mathbf{H}_{\pi(k)}^H \right|}. \quad (2.13)$$

A região de capacidade é a fronteira convexa da região definida por todos os vetores de capacidade sobre todas as permutações possíveis e sobre todas as matrizes de covariância que satisfazem a restrição de potência:

$$\mathcal{C}_{BC}(P_T; \mathbf{H}; \sigma_n^2) \triangleq \text{Co} \left(\bigcup_{\pi, \mathcal{R}} \mathbf{R}(\pi, \mathcal{R}) \right), \quad (2.14)$$

em que $\mathbf{R}(\pi, \mathcal{R})$ é dado pela Eq. (2.13) e $\text{Co}(\cdot)$ denota a fronteira convexa da região. O sinal transmitido é dado por $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_K$, e as matrizes de covariância são na forma $\mathcal{R}_{\mathbf{x}_k} \triangleq \mathbb{E}[\mathbf{x}_k \mathbf{x}_k^H]$. Uma consequência do DPC é que $\mathbf{x}_1, \dots, \mathbf{x}_K$ não são correlacionados, e, conseqüentemente, é válida a igualdade $\text{Tr}(\mathcal{R}_{\mathbf{x}}) = \text{Tr}[\mathcal{R}_{\mathbf{x}_1} + \dots + \mathcal{R}_{\mathbf{x}_K}] \leq P_T$.

Um detalhe importante a ser percebido é que a expressão definida pela Eq. (2.14) não é uma função nem côncava, nem convexa das matrizes de covariância. Isso torna a tarefa de encontrar a região de capacidade do canal MIMO BC extremamente difícil, porque uma busca sobre todo o espaço de matrizes de covariância deve ser realizada. Entretanto, explorando a dualidade existente entre os canais MIMO BC e MIMO MAC, essa tarefa se torna menos complicada, como será considerado na próxima seção.

Cabe ainda mais uma observação. Apesar do DPC ser de grande interesse teórico devido ao fato de ser um esquema que é capaz de atingir a capacidade do canal, sua implementação é extremamente difícil devido à sua natureza complexa, e requer estratégias sofisticadas de codificação e decodificação [39, 40]. A principal dificuldade está na natureza não linear do processo de codificação multiestágio para a pré-subtração de interferência. Complexidade adicional é originada no processo de determinar as matrizes de autocovariância que maximizem determinado valor de taxa. Além disso, muitos sistemas podem não possuir poder computacional suficiente para os cálculos iterativos com matrizes.

2.3.4 Dualidade MAC-BC

O canais MIMO BC e MIMO MAC aparentemente são muito similares. Entretanto, há três diferenças fundamentais entre os dois modelos. Em primeiro lugar, no enlace direto existe um termo vetorial de ruído aditivo associado a cada receptor, enquanto que no enlace reverso existe apenas uma componente

vetorial de ruído (já que existe apenas uma entidade receptora). Uma segunda diferença fundamental é que, no enlace direto, apenas uma restrição de potência é aplicável ao transmissor. No enlace reverso, há diferentes restrições associadas a cada usuário. Em terceiro lugar, tanto sinal como interferência associados ao usuário k sofrem a influência do mesmo canal no enlace direto (ver Eq. (2.3)), enquanto que no enlace reverso, sinal e interferência sofrem a ação de diferentes canais.

Dois canais MIMO MAC e MIMO BC de K -usuários são ditos duais se [41, 13, 42]:

1. As respostas impulsivas \mathbf{H}_k , $k = 1, \dots, K$ do canal no enlace direto são as mesmas do enlace reverso para todo k ;
2. Cada receptor no enlace direto possui as mesmas estatísticas de ruído, sendo estas as mesmas do receptor no enlace reverso;
3. A restrição de potência de transmissão P_T no enlace direto é igual à soma das restrições de potência individuais P_k no enlace reverso, isto é, $P_T = \sum_{k=1}^K P_k$.

A partir do conceito de dualidade, em [43, 44], é mostrado que a região de capacidade do canal MIMO BC com restrição de potência P_T e matriz combinada de canal \mathbf{H} é igual à união das regiões de capacidade do canal MAC dual com matriz de canal \mathbf{H}^H , em que esta é tomada sobre o conjunto de todas as restrições individuais de potência P_k que somam P_T . Matematicamente,

$$\mathcal{C}_{BC}(P_T; \mathbf{H}; \sigma_n^2) = \bigcup_{(P_1, \dots, P_K): \sum_{k=1}^K P_k = P_T} \mathcal{C}_{MAC}((P_1, \dots, P_K); \sigma_n^2; \mathbf{H}^H), \quad (2.15)$$

em que $\mathcal{C}_{BC}(P_T; \mathbf{H}; \sigma_n^2)$ é dado pela Eq. (2.14) e $\mathcal{C}_{MAC}((P_1, \dots, P_K); \sigma_n^2; \mathbf{H}^H)$ é dado pela Eq. (2.9).

A Fig. 2.7 mostra a região de capacidade do canal MIMO BC com dois usuários, cada um com uma antena, e uma estação rádio-base com duas antenas. A região foi obtida a partir da relação de dualidade com o canal MIMO MAC, em que a região de capacidade do canal MIMO BC corresponde à fronteira da união das regiões, e cada linha no interior da região corresponde à região de capacidade de um canal MIMO MAC cuja soma das potências equivale a P_T .

É importante lembrar que a região de capacidade do MIMO BC é extremamente difícil de ser calculada, haja vista que não é uma função nem côncava, nem convexa, das matrizes de covariância. Já a região de capacidade do canal MIMO MAC pode ser obtida de forma eficiente utilizando-se técnicas convencionais de otimização convexa. Na verdade, basta determinar as matrizes de covariância para o caso MIMO MAC e então transformá-las nas matrizes de covariância correspondentes ao canal MIMO BC, conforme definido pelas relações apresentadas em [45].

2.3.5 Somas das Capacidades

Outra medida comumente utilizada para avaliar a capacidade de canais multiusuário, e em especial o canal MIMO BC, é a soma das capacidades (do inglês, *sum capacity*), que pode ser derivada da região de capacidade. A soma das capacidades é o máximo valor possível da soma de todas as taxas tomadas sobre os vetores que pertencem à região de capacidade, e expressa por meio de

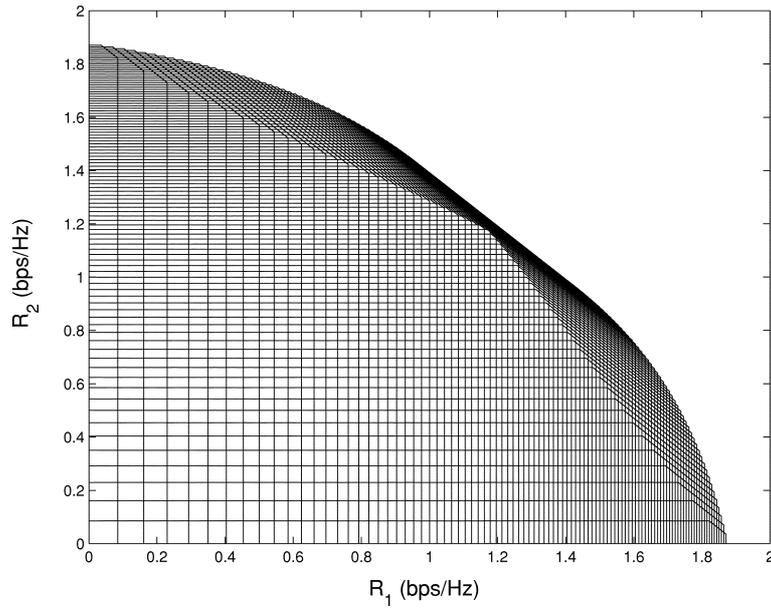


Figura 2.7: Região de capacidade do canal MIMO BC para $M_T = 2$ e $K = 2$. Os valores numéricos foram obtidos considerando $\mathbf{H}_1 = [1 \ 0,4]$ e $\mathbf{H}_2 = [0,4 \ 1]$ para $P_T = 10$. Modificado de [1].

$$C_{BC}^{sum} = C_{DPC}^{sum} = \max \sum_{k=1}^K R_k. \quad (2.16)$$

A Eq. (2.16) representa o *throughput* máximo do sistema, sendo mais fácil de ser caracterizado do que uma região K -dimensional. O ponto que representa a soma das capacidades é mostrado na Fig. 2.8 com relação à região de capacidade para o canal MIMO BC com $M_T = 2$ antenas e dois usuários, cada um com uma antena.

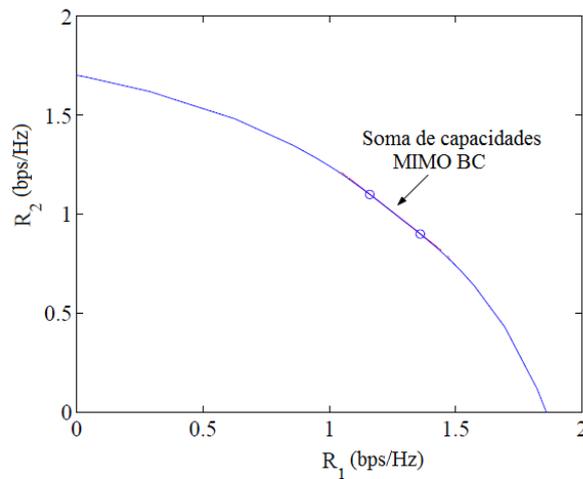


Figura 2.8: Soma das capacidades do canal MIMO BC.

Se todos os receptores possuem apenas uma antena, a somas das capacidades se reduz a [13]

$$C_{DPC}^{sum} = \max_{\text{Tr}(\mathbf{D}) \leq P_T} \log \left| \mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{H} \mathbf{D} \mathbf{H}^H \right|, \quad (2.17)$$

em que \mathbf{D} é uma matriz diagonal positiva semidefinida de dimensão $K \times K$. A Eq. (2.17) é bastante semelhante à fórmula de capacidade de um sistema MIMO ponto-a-ponto (monousuário) com M_T antenas transmissoras e K antenas receptoras, no qual CSI está disponível apenas no receptor [30]. Esta comparação torna fácil perceber que a soma das capacidades aumenta linearmente com $\min(M_T, K)$.

O fato mais impressionante deste resultado é que a falta de cooperação entre os receptores não reduz a capacidade do sistema. Na verdade, o canal MIMO BC experimenta o mesmo ganho de multiplexação espacial que o caso MIMO ponto-a-ponto. Entretanto, neste, múltiplas antenas são necessárias tanto no transmissor quanto no receptor para que sejam observados os ganhos lineares de capacidade, enquanto que, naquele, a presença de múltiplas antenas só é necessária no lado transmissor (resultado similar pode ser mostrado para o caso MIMO MAC [33]).

É importante observar que, apesar de a soma das capacidades não apresentar todas as informações necessárias sobre o canal multiusuário, ela ainda assim é útil para analisar o desempenho dos algoritmos de transmissão e pré-codificação, a serem apresentados no capítulo 3.

2.4 CONCLUSÃO

Neste capítulo, foi apresentado um modelo matemático para o tratamento dos canais multiusuário MIMO BC e MIMO MAC. Foram apresentados resultados fundamentais em teoria da informação sobre a capacidade dos canais MIMO multiusuário, sendo ainda considerada a dualidade MAC-BC para a obtenção da região de capacidade do caso MIMO BC.

Foi ainda visto que o DPC, enquanto ótimo sob o ponto de vista teórico para se atingir a região de capacidade do canal de *broadcast*, é um conceito de Teoria da Informação que se mostrou extremamente complexo de ser implementado na prática. A estrutura (*framework*) proposta pelo DPC mostra que, por meio da codificação sucessiva dos usuários, é possível elaborar um esquema de pré-codificação que permite alcançar a região de capacidade do canal MIMO BC. Entretanto, não é dito como obter as palavras-código que serão utilizadas.

Devido a este fato, no próximo capítulo, serão abordadas técnicas subótimas de pré-codificação no transmissor, que utilizam técnicas de processamento lineares e não lineares e que, sob determinadas condições, podem alcançar assintoticamente a capacidade do canal MIMO BC.

3 TÉCNICAS DE PRÉ-CODIFICAÇÃO LINEARES E NÃO LINEARES PARA O CANAL MIMO BC

3.1 INTRODUÇÃO

A pré-codificação é uma estratégia interessante quando a transmissão é realizada sobre um canal que, além do ruído aditivo, sofre também com um termo adicional de interferência, especialmente se este é conhecido pelo transmissor.

Os algoritmos de pré-codificação para o enlace direto de um sistema MIMO multiusuário podem ser classificados segundo diversos critérios, de acordo com os objetivos que são almejados: maximizar a soma das capacidades, eliminar a interferência interusuário, alcançar algum objetivo específico de qualidade de serviço ou restrição de relação sinal-ruído, e devem levar em conta a presença de múltiplas antenas nos receptores e a existência de múltiplos fluxos de dados a serem transmitidos para cada usuário, para citar alguns exemplos.

É interessante observar que uma técnica de pré-codificação que maximize a capacidade de apenas um usuário pode resultar em níveis de interferência inaceitáveis para os demais usuários, tornado seus enlaces de comunicação inutilizáveis, conforme mostram os resultados de [46]. Maximizar a capacidade de um dos usuários implica em alocar 100% da potência disponível para ele, o que resultaria em uma taxa de transmissão nulas para os demais usuários.

Os algoritmos considerados neste capítulo visam atingir o máximo *throughput* do sistema, ao mesmo tempo que tentam eliminar completamente a interferência interusuário, e será dada mais ênfase ao caso que recebe mais atenção na literatura: usuários com uma antena receptora.

As técnicas apresentadas podem ser divididas em dois grandes grupos: lineares e não lineares.

A pré-codificação linear é uma generalização dos algoritmos clássicos de SDMA, no qual são determinadas diferentes matrizes de pré-codificação para cada usuário. Os pré-codificadores são projetados tendo como base a informação sobre o estado de canal de todos os usuários. A abordagem mais direta é referida como inversão de canal, ou pré-codificação *zero-forcing*, que consiste em utilizar um conjunto de pesos (*beamformers*) no transmissor de forma a inverter, no transmissor, o efeito do canal sobre os símbolos de informação e eliminar completamente a interferência interusuário em todos os receptores. Em seguida, será abordada a técnica de regularização de inversa, uma das formas encontradas para se contornar o problema inerente da pré-codificação *zero-forcing*, que é o aumento no nível da potência do sinal transmitido. Será considerada de forma sucinta a técnica de *block diagonalization*, que é uma generalização da técnica de inversão de canal, utilizada sobretudo quando os usuários possuem múltiplas antenas.

Apesar de serem capazes de eliminar completamente a interferência entre os múltiplos usuários, e serem computacionalmente simples, as técnicas lineares não são ótimas no sentido de maximizar a soma das capacidades. Como já mencionado, técnicas baseadas em DPC são capazes de atingir valores de

capacidade mais próximos dos limites do canal MIMO BC. Técnicas baseadas em DPC são diferentes dos casos considerados anteriormente pois a informação que é transmitida é uma função não linear tanto dos símbolos de informação como da interferência presente no sistema.

Em seguida serão estudados os algoritmos não lineares, que tratam de implementações subótimas da codificação *dirty paper* que podem se aproximar da soma das capacidades (e em alguns casos, efetivamente alcançá-la) [47, 48]. Serão descritas duas técnicas: a pré-codificação Tomlinson-Harashima [49], e a técnica de *vector perturbation* [50]. Ambas utilizam como função de mapeamento não linear o operador módulo, que é possivelmente a mais simples implementação subótima do DPC. Por meio da pré-codificação sobre um reticulado previamente definido, limitam-se os valores de excursão do sinal a ser transmitido após a pré-codificação como forma de controlar a potência de transmissão.

Tendo como base o conceito de pré-codificação pela introdução de um vetor de perturbação nos dados a serem transmitidos como forma de melhorar o desempenho de sistemas MIMO multiusuário, será feita uma proposta de pré-codificador para o canal MIMO BC, e seu desempenho será comparado com as técnicas lineares clássicas anteriormente apresentadas. Por meio de simulações, será mostrado que o pré-codificador proposto apresenta desempenho melhor do que as técnicas lineares tratadas, mas não exibe o mesmo ganho proporcionado pela técnica de *vector perturbation*.

Finalmente, será iniciado o estudo dos efeitos da informação imperfeita sobre o estado do canal nas técnicas de pré-codificação apresentadas.

3.2 MODELO DE SISTEMA

Será revisitado de forma breve o modelo apresentado para o canal MIMO BC. Utilizando as Eq. (2.3), (2.4) e (2.7), o sinal recebido pelo usuário k pode ser expresso por meio de

$$\begin{aligned} \mathbf{y}_k &= \sum_{i=1}^K \mathbf{H}_k \mathbf{W}_i \mathbf{u}_i + \mathbf{n}_k \\ &= \mathbf{H}_k \mathbf{W}_k \mathbf{u}_k + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{H}_k \mathbf{W}_i \mathbf{u}_i + \mathbf{n}_k. \end{aligned} \quad (3.1)$$

Supondo que cada usuário possui apenas uma antena de recepção, e que o transmissor enviará apenas um símbolo de informação u_k para o usuário k , isto é, $m_k = n_{Rk} = 1$, $k = 1, \dots, K$, o modelo apresentado pela Eq. (3.1) se reduz a

$$y_k = \mathbf{H}_k \mathbf{w}_k u_k + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{H}_k \mathbf{w}_i u_i + n_k, \quad (3.2)$$

em que \mathbf{H}_k é o vetor coluna $1 \times M_T$ que representa os ganhos de canal entre a antena do usuário k e as M_T antenas da estação rádio-base, e o vetor de símbolos transmitidos \mathbf{x} está relacionado ao vetor de símbolos de informação \mathbf{u} por meio de $\mathbf{x} = \mathbf{W}\mathbf{u}$, em que $\mathbf{x} = [x_1, \dots, x_K]^T$, $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$, $\mathbf{u} = [u_1, \dots, u_K]^T$.

A matriz combinada de canal será expressa por meio de $\mathbf{H} = [\mathbf{H}_1^T, \dots, \mathbf{H}_K^T]^T$. Por simplicidade, será suposto ainda que todos os símbolos do vetor \mathbf{u} são originados do mesmo alfabeto de símbolos (por exemplo, QPSK ou 16QAM).

Conforme abordado no capítulo 2, a relação sinal-ruído será definida por meio da expressão

$$\rho = \frac{P_T}{\sigma_n^2}, \quad (3.3)$$

em que P_T representa a potência de transmissão e σ_n^2 é a variância da componente de ruído branco gaussiano.

Antes de prosseguir com a apresentação dos algoritmos, é importante expor algumas observações. As técnicas de pré-codificação que serão apresentadas procuram maximizar o *throughput* do sistema. Entretanto, operar no ponto da soma das capacidades nem sempre leva a uma solução desejável ou justa, e pode não ser, necessariamente, o objetivo do projetista do sistema. Para ilustrar e explicar de forma intuitiva esta situação, será utilizada a Fig. 3.1, que mostra duas regiões de capacidade para um cenário MIMO BC com dois usuários. A maior delas refere-se ao caso em que os dois usuários possuem aproximadamente a mesma capacidade máxima, e a menor delas, marcada em cinza, refere-se ao caso em que a capacidade máxima difere para os dois usuários. Mais especificamente, o usuário 2 possui menor capacidade, pois seu canal de comunicação encontra mais atenuação quando comparado ao canal do usuário 1. Esta situação é referida na literatura como problema *near-far* [27].

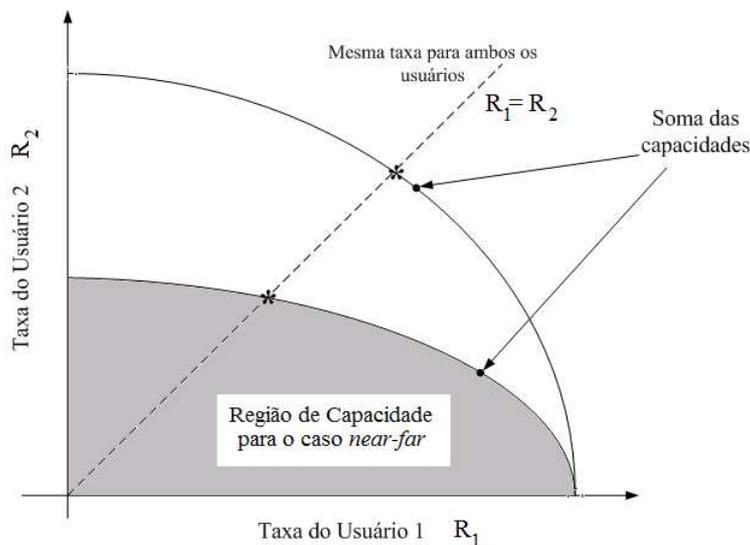


Figura 3.1: Ilustração da região de capacidade de um sistema multiusuário considerando o problema *near-far*. A soma das capacidades pode penalizar determinados usuários, dependendo da forma assumida pela região de capacidade.

Em qualquer uma das situações apresentadas, o ponto da soma das capacidades (representado pelo marcador "•") não implica em alocação igualitária de taxas de transmissão para ambos os usuários (situação esta representada pelo ponto "*"). A situação é ainda mais grave quando se tem o problema *near-far*, uma vez que a maximização da capacidade do sistema vem às custas do usuário 2, além de o sistema apresentar reduzida soma das capacidades quando comparado ao caso em que os usuários possuem aproximadamente o mesmo ganho de canal.

Uma alternativa que pode ser utilizada nos casos que apresentam o problema *near-far* é trabalhar com algoritmos de seleção de usuários [51, 52], ou procurar satisfazer alguma meta de qualidade de serviço para cada usuário. O problema de reunir as restrições de qualidade de serviço e mínima potência de transmissão é denominado controle de potência (ou ainda, balanceamento de interferência) no enlace direto. Métodos iterativos para solução do problema são consideradas em [53, 54]. Entretanto, essas soluções não serão descritas nesta dissertação. A abordagem utilizada é sob a ótica de algoritmos de processamento, e considera-se que não existe o problema *near-far*.

3.3 ZERO-FORCING

A inversão de canal, quando realizada pelo transmissor, é denominada pré-codificação *zero-forcing* (ZF) [55, 56]. Os *beamformers* dos usuários são escolhidos de forma a satisfazer $\mathbf{H}_k \mathbf{w}_i = 0 \forall k \neq i$, de forma que toda a interferência entre os usuários é eliminada. A matriz de pré-codificação $\mathbf{W} = \mathbf{W}_{ZF}$ é dada pela pseudo-inversa de Moore-Penrose, de forma que $\mathbf{W}_{ZF} \mathbf{H} = \mathbf{I}$. Sua definição é:

$$\begin{aligned} \mathbf{W}_{ZF} &= \mathbf{H}^\dagger \\ &= \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1}. \end{aligned} \quad (3.4)$$

Neste caso, o vetor efetivamente transmitido será expresso por

$$\begin{aligned} \mathbf{x} &= \frac{1}{\sqrt{\gamma}} \mathbf{W}_{ZF} \mathbf{u} \\ &= \frac{1}{\sqrt{\gamma}} \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1} \mathbf{u}. \end{aligned} \quad (3.5)$$

A constante de normalização γ é definida de tal forma que a potência de transmissão esteja restrita ao valor P_T , isto é, $\|\mathbf{x}\|^2 = P_T$. Logo, $\gamma = \frac{1}{P_T} \|\mathbf{W} \mathbf{u}\|^2$. Sem perda de generalidade, utilizar-se-á que $P_T = 1$. Por meio da Eq. (3.5), pode-se expressar γ como

$$\gamma = \mathbf{u}^H (\mathbf{H} \mathbf{H}^H)^{-1} \mathbf{u}. \quad (3.6)$$

O caso $K < M_T$ não é interessante em termos de maximização de capacidade, pois nem todas as dimensões espaciais oferecidas pelas múltiplas antenas serão aproveitadas. Basta lembrar que, conforme mostrado no capítulo 2, o aumento de capacidade no canal MIMO BC se dá com $\min(M_T, K)$ quando cada receptor possui uma antena. Será dada mais atenção, portanto, apenas aos casos em que $M_T = K$. Nesta situação, a Eq. (3.4) reduz-se a

$$\mathbf{W}_{ZF} = \mathbf{H}^{-1}, \quad (3.7)$$

e o vetor transmitido será expresso por

$$\mathbf{x} = \frac{1}{\sqrt{\gamma}} \mathbf{H}^{-1} \mathbf{u}. \quad (3.8)$$

Com essa abordagem, toda a interferência multiusuário é eliminada, e o problema é reduzido ao caso de K canais SISO escalares, nos quais o k -ésimo usuário recebe apenas a componente u_k do vetor de símbolos \mathbf{u} , além da componente de ruído gaussiano. A partir das Eq. (3.2) e (3.4), podemos então escrever

$$y_k = \frac{1}{\sqrt{\gamma}} u_k + n_k. \quad (3.9)$$

Apesar de fornecer um modo bastante simples de realizar a pré-codificação, o *zero-forcing* apresenta algumas desvantagens e limitações. A primeira é que o valor da constante de normalização depende do vetor de dados, que varia de transmissão para transmissão. Esta limitação pode ser contornada escolhendo γ de tal forma que a potência média seja dada por P_T , isto é,

$$\gamma = \frac{1}{P_T} \text{Tr} \left[(\mathbf{H}\mathbf{H}^H)^{-1} \right], \quad (3.10)$$

supondo que a transmissão dos usuários é independente e os símbolos escolhidos possuem potência média unitária. Entretanto, a média dada pela Eq. (3.10) pode não existir, conforme será exposto.

Além disso, conforme mostrado em [57], existirá um hiato (*gap*) entre a soma de capacidades da pré-codificação *zero-forcing* e o limite oferecido pelo DPC. Este hiato pode ser analiticamente determinado por meio da expressão:

$$C_{DPC}^{sum} - C_{ZF}^{sum} = (\log e) \sum_{i=1}^{K-1} \frac{i}{M_T - i}. \quad (3.11)$$

A Fig. 3.2 ilustra o resultado da Eq. (3.11) considerando um ambiente com $K = 10$ usuários.

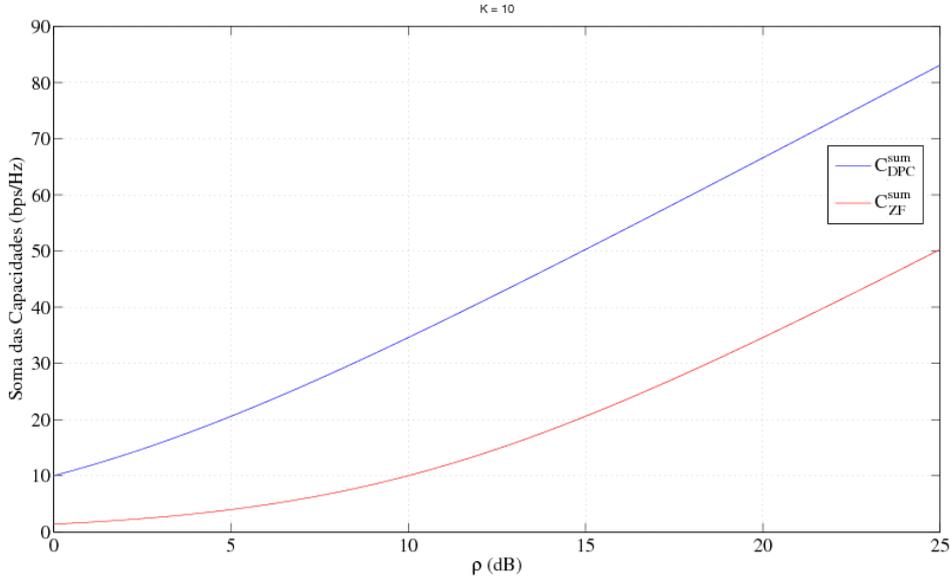


Figura 3.2: Hiato entre a soma das capacidades da pré-codificação *zero-forcing* (ZF) e o limite dado pelo DPC para $K = 10$ usuários.

Um problema ainda mais sério surge quando a matriz combinada de canal é mal condicionada. Nesta

situação, pelo menos um dos autovalores de $(\mathbf{H}\mathbf{H}^H)^{-1}$ será muito grande, e a relação sinal-ruído nos receptores será muito baixa devido ao fator γ , pois este assumirá um valor muito grande. Esta situação é análoga ao fenômeno de *noise enhancement* que ocorre nos equalizadores de canal do tipo *zero-forcing* [58] aplicados aos sistemas MIMO ponto-a-ponto. Entretanto, diferentemente deste caso, o desempenho no cenário MIMO multiusuário é prejudicado até mesmo quando os elementos de \mathbf{H} são independentes e identicamente distribuídos [59], o que corresponde a um ambiente rico em espalhadores.

Em [59], é mostrado que γ é uma variável aleatória cuja função densidade de probabilidade é dada por

$$p_\gamma(\gamma) = K \frac{\gamma^{K-1}}{(1+\gamma)^{K+1}} \quad (3.12)$$

A distribuição da Eq. (3.12) possui média infinita. Como consequência deste fato, a soma das capacidades, quando é utilizada pré-codificação do tipo *zero-forcing*, não aumenta linearmente com K , diferentemente do que ocorre com o limite da capacidade, dado pela Eq. (2.17). Pode-se mostrar que para $K = M_T$, é válido o seguinte resultado assintótico, para $K \rightarrow \infty$ [59]:

$$\mathcal{C}_{ZF}^{sum} \underset{K \rightarrow \infty}{=} \rho \log e. \quad (3.13)$$

A Fig. 3.3 mostra o limite de capacidade supracitados. Como é possível observar, a medida que K tende a infinito, a soma de capacidades da pré-codificação *zero-forcing* se aproxima do limite dado pela Eq. (3.13).

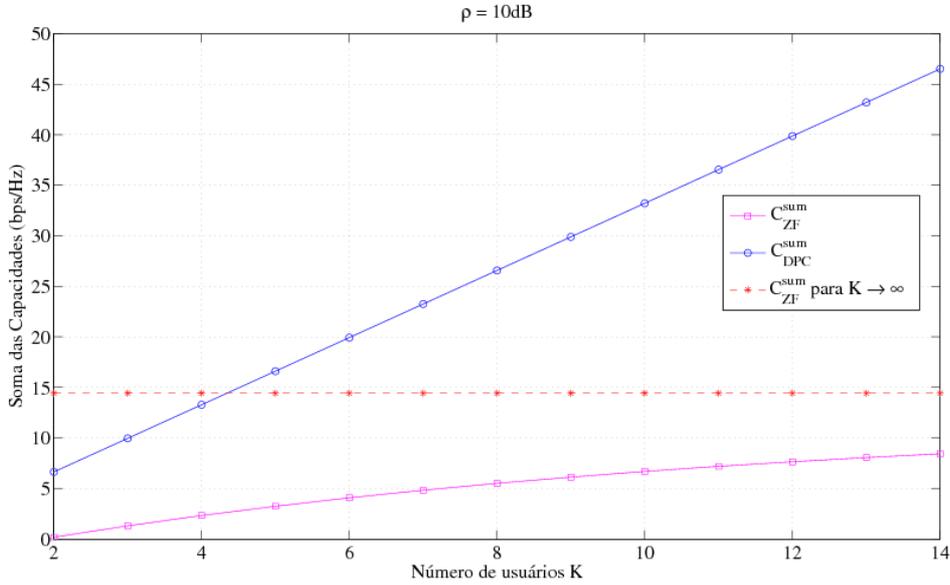


Figura 3.3: Comparação da soma de capacidades em função de K da pré-codificação *zero-forcing* e o limite teórico oferecido pelo DPC, para uma razão sinal-ruído de $\rho = 10$ dB.

A explicação para o fato pode ser obtida ao observar os autovalores de $(\mathbf{H}\mathbf{H}^H)^{-1}$. Conforme é mostrado em [59], o menor autovalor de $\mathbf{H}\mathbf{H}^H$, λ_1 , possui função densidade de probabilidade exponencial $p_{\lambda_1}(\lambda_1) = K e^{-K\lambda_1}$. O maior autovalor de $(\mathbf{H}\mathbf{H}^H)^{-1}$, λ_1^{-1} , apresentará distribuição gama-inversa, cuja média é infinita. Como consequência, os demais $K - 1$ autovalores são significativamente bem

comportados, conforme pode ser visto na simulação mostrada na Fig. 3.4. Nesta figura, é exibido o comportamento dos quatro maiores autovalores da matriz $(\mathbf{H}\mathbf{H}^H)^{-1}$ em função de K . Foram realizadas 5.000 simulações, e os autovalores estão normalizados por um fator K . Além disso, utilizou-se a suposição ZMSW para gerar os coeficientes presentes na matriz combinada de canal. O maior autovalor possui comportamento errático para todo K , consequência do fato de sua média ser infinita (pois ele possui distribuição gamma-inversa), e é claramente várias ordens de magnitude maior que os demais autovalores. Logo, qualquer técnica que procure melhorar o desempenho da inversão de canal deve procurar reduzir os efeitos do maior autovalor de $(\mathbf{H}\mathbf{H}^H)^{-1}$ sobre o vetor pré-codificado \mathbf{x} . Este pressuposto será utilizado na pré-codificação por regularização de inversa e também na técnica não linear denominada *vector perturbation*.

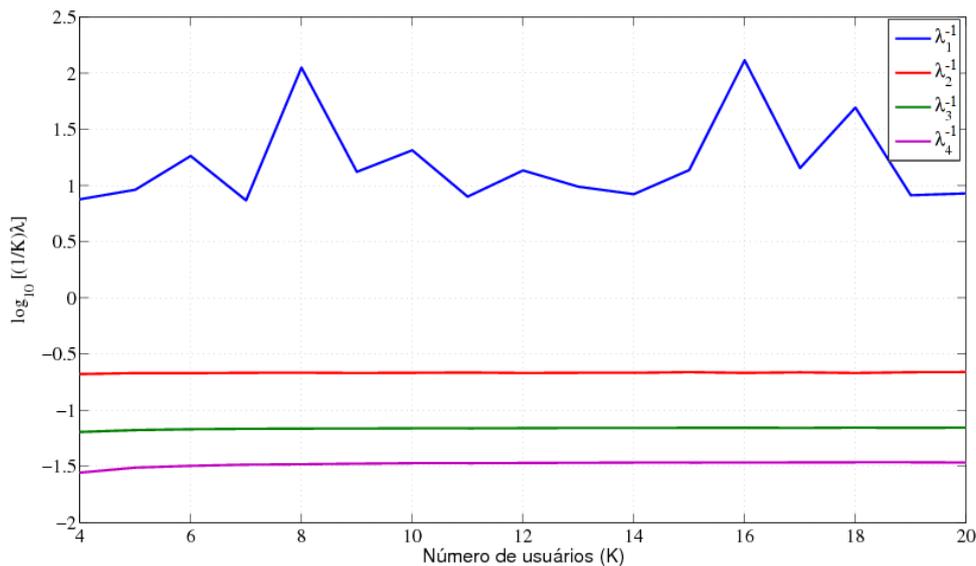


Figura 3.4: Comportamento dos quatro maiores autovalores de $(\mathbf{H}\mathbf{H}^H)^{-1}$ em função de K .

3.4 REGULARIZAÇÃO DE INVERSA

Uma abordagem comumente utilizada para reduzir os efeitos da inversão de uma matriz mal condicionada é regularizar a matriz inversa adicionando a esta um múltiplo da matriz identidade. Se o ruído é espacialmente branco e um valor para a constante de regularização é escolhido apropriadamente, esta abordagem é equivalente a uma conformação de feixes (*beamforming*) utilizando o critério de minimizar o erro quadrático médio. Esta abordagem é denominada regularização de inversa (RI, do inglês, *regularized inverse*).

Aplicando este princípio no lado transmissor, tem-se o vetor transmitido

$$\begin{aligned}
 \mathbf{x} &= \frac{1}{\sqrt{\gamma}} \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \varsigma \mathbf{I})^{-1} \mathbf{u} \\
 &= \frac{1}{\sqrt{\gamma}} \mathbf{W}_{IR} \mathbf{u},
 \end{aligned} \tag{3.14}$$

em que ς é denominado parâmetro de regularização. A presença de um valor não nulo para ς indica que o transmissor não será capaz de eliminar completamente a interferência interusuário. O sinal recebido pelo usuário k não é mais simplesmente uma versão de u_k afetada por um fator multiplicativo, conforme expresso na Eq. (3.9), mas também incluirá um termo de interferência residual multiusuário.

Para avaliar a quantidade de interferência que é introduzida, pode-se utilizar a autodecomposição $\mathbf{H}\mathbf{H}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$, em que \mathbf{Q} representa a matriz dos autovetores de $\mathbf{H}\mathbf{H}^H$ e $\mathbf{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_K]$ é uma matriz diagonal cujas entradas são os autovalores de $\mathbf{H}\mathbf{H}^H$, organizados de forma decrescente, com $\lambda_1 < \lambda_2 < \dots < \lambda_K$. Dessa forma, o vetor recebido, após sofrer a influência do canal, será expresso por

$$\mathbf{H}\mathbf{x} = \frac{1}{\sqrt{\gamma}} \mathbf{Q} \frac{\mathbf{\Lambda}}{\mathbf{\Lambda} + \varsigma \mathbf{I}} \mathbf{Q}^H \mathbf{u}. \quad (3.15)$$

O sinal recebido pelo usuário k pode ser escrito como

$$[\mathbf{H}\mathbf{x}]_k = \begin{bmatrix} q_{k,1} \frac{\lambda_1}{\lambda_1 + \varsigma} & \dots & q_{k,K} \frac{\lambda_K}{\lambda_K + \varsigma} \end{bmatrix} \begin{bmatrix} q_{1,1}^* & \dots & q_{K,1}^* \\ \vdots & \ddots & \vdots \\ q_{1,K}^* & \dots & q_{K,K}^* \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix}, \quad (3.16)$$

em que $q_{m,n}$ representa o elemento da matriz \mathbf{Q} na m -ésima linha e n -ésima coluna. Efetuando os produtos matriciais, tem-se

$$y_k = \left(\frac{1}{\sqrt{\gamma}} \right) \left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \varsigma} |q_{k,i}|^2 \right) u_k + w'_k. \quad (3.17)$$

A componente w'_k corresponde ao termo de ruído aditivo gaussiano somado à interferência residual, dada pelos termos da Eq. (3.16) que envolvem u_l , com $l \neq k$.

Naturalmente, a quantidade de interferência gerada depende de ς . Quando $\varsigma = 0$, volta-se para o caso da pré-codificação *zero-forcing* (Eq. (3.7)). Não importa o quanto a inversa da Eq. (3.14) seja mal condicionada, ela pode se comportar tão bem quanto desejado, bastando para isso escolher ς grande o suficiente. Entretanto, a quantidade de interferência gerada será tão maior quanto maior for o valor da constante de regularização. Para esta situação, está caracterizada uma solução de compromisso entre o bom condicionamento da inversa da matriz combinada de canal (que tem impacto direto sobre o fator de normalização γ) e o nível de interferência residual que será gerado.

Uma métrica apropriada para a escolha de ς é maximizar a relação sinal-ruído mais interferência, *SINR* (do inglês, *signal to interference-plus-noise ratio*). Em [59], é mostrado que, para valores grandes de K e supondo que $\mathbb{E}[\mathbf{u}\mathbf{u}^H] = \mathbf{I}$ afim de simplificar os cálculos, a *SINR* pode ser aproximada por

$$SINR \approx \frac{\left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \varsigma} \right)^2}{K^2 \sigma_n^2 \left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \varsigma} \right)^2 + K \left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \varsigma} \right)^2 - \left(\sum_{i=1}^K \frac{\lambda_i}{\lambda_i + \varsigma} \right)^2}. \quad (3.18)$$

Tomando a derivada da Eq.(3.18) em relação ς , a *SINR* é maximizada para $\varsigma_{opt} = K\sigma_n^2$, independente-

mente dos autovalores $\lambda_1, \dots, \lambda_K$. Apesar das vantagens oferecidas sobre a pré-codificação *zero-forcing* (como será mostrado nos próximos parágrafos), a desvantagem da regularização de inversa é que ela exige a estimação da variância do ruído gaussiano.

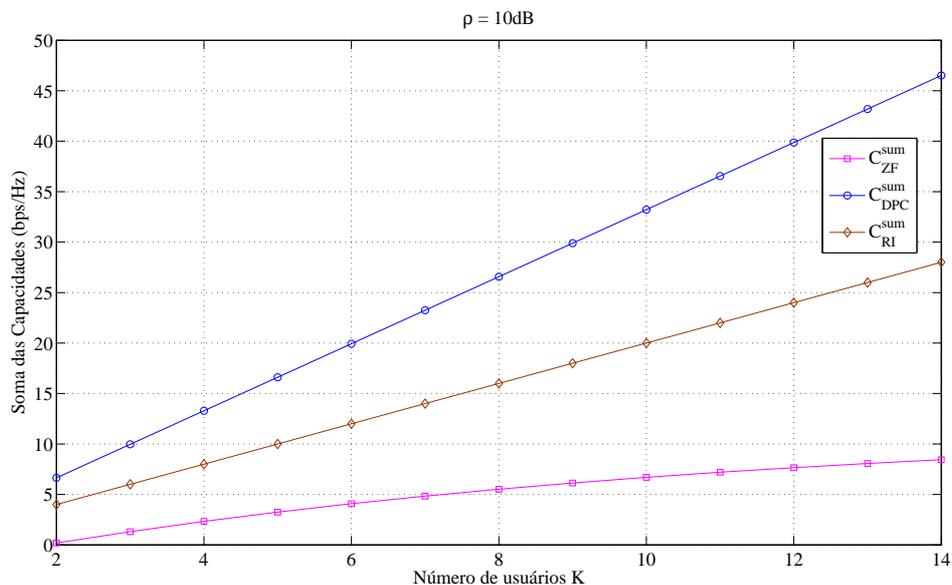


Figura 3.5: Comparação da soma das capacidades em função de K das técnicas de pré-codificação *zero-forcing* e regularização de inversa para uma relação sinal-ruído de $\rho = 10$ dB.

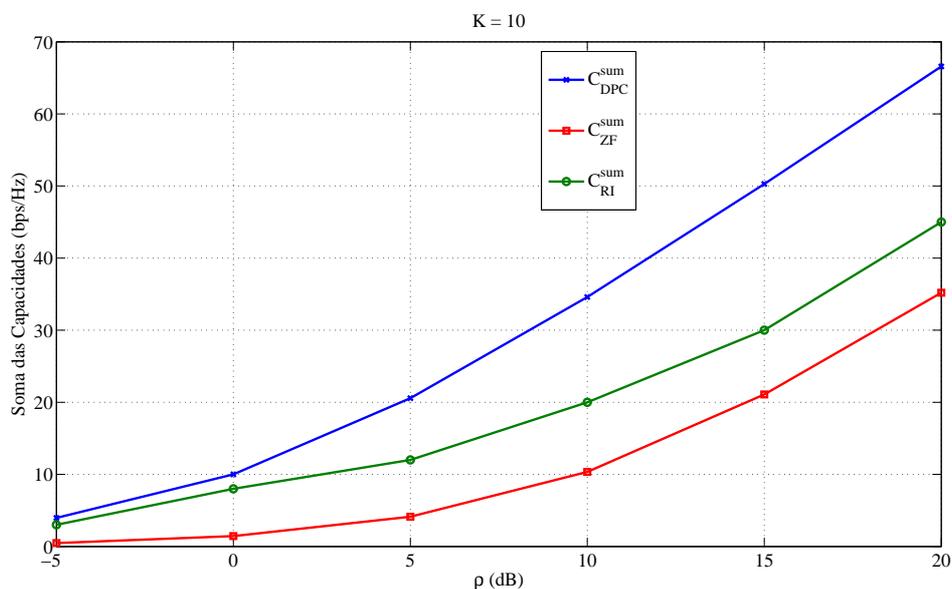


Figura 3.6: Comparação da soma das capacidades em função de ρ das técnicas de pré-codificação *zero-forcing* e regularização de inversa para $K = 10$ usuários.

Diferentemente da pré-codificação por inversão do canal, a soma das capacidades da pré-codificação por regularização de inversa aumenta linearmente com K , apesar da inclinação da curva ser diferente, conforme está mostrado na Fig. 3.5 mantendo $\rho = 10$ dB e variando o número de usuários K .

A regularização de inversa representa uma grande melhora sobre a simples inversão de canal, entretanto

permanece o hiato na capacidade entre este esquema e o limite teórico oferecido pelo DPC, especialmente no regime de relação sinal-ruído elevada. A Fig. 3.6 mostra a soma das capacidades da pré-codificação *zero-forcing* e da pré-codificação por regularização de inversa variando-se a relação sinal-ruído e mantendo o número de usuários fixo em $K = 10$. Para valores baixos de ρ , a soma das capacidades da pré-codificação por regularização de inversa se aproxima de C_{DPC}^{sum} , enquanto que para valores mais altos de ρ , a tendência é que a soma das capacidades se aproxime de C_{ZF}^{sum} , o que é consequência da própria definição de ς . Para $\rho \rightarrow \infty$, $\varsigma \rightarrow 0$, e retornamos ao caso da pré-codificação *zero-forcing*.

Como última análise, também será considerada a curva de desempenho da BER (do inglês, *bit error rate*, ou taxa de erro de *bit*) média dos sistemas que utilizam *zero-forcing* e regularização de inversa. As curvas foram obtidas por meio de simulação, utilizando a técnica de Monte Carlo, e o canal apresenta desvanecimento do tipo Rayleigh sem os efeitos da perda de percurso. Para fornecer uma comparação entre os sistemas, utilizou-se como parâmetro de desempenho a figura de mérito dos sistemas digitais, definida como a razão entre a energia média de bit e a densidade espectral de potência do ruído [60], que relaciona-se à ρ por meio de

$$\frac{E_b}{N_o} = \frac{\rho}{\log_2 M}. \quad (3.19)$$

Os resultados estão mostrados na Fig. 3.7 para a modulação 16QAM e na Fig. 3.8 para a modulação QPSK, variando-se o número K de usuários. Enquanto que o desempenho da pré-codificação *zero-forcing* piora com aumento do número de usuários, o desempenho da pré-codificação por regularização de inversa melhora ligeiramente para valores maiores de K .

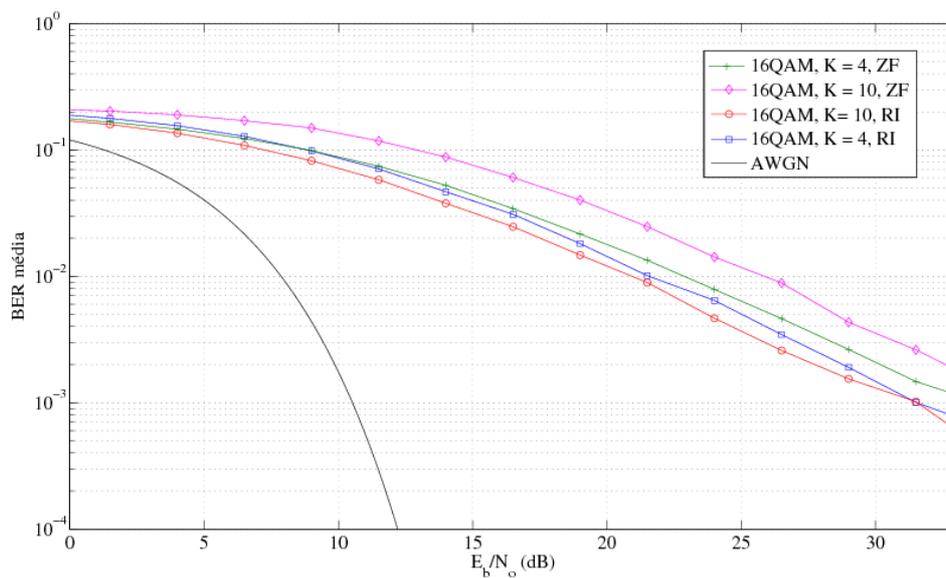


Figura 3.7: Comparação das curvas de BER dos esquemas de pré-codificação por inversão de canal e regularização de inversa para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.

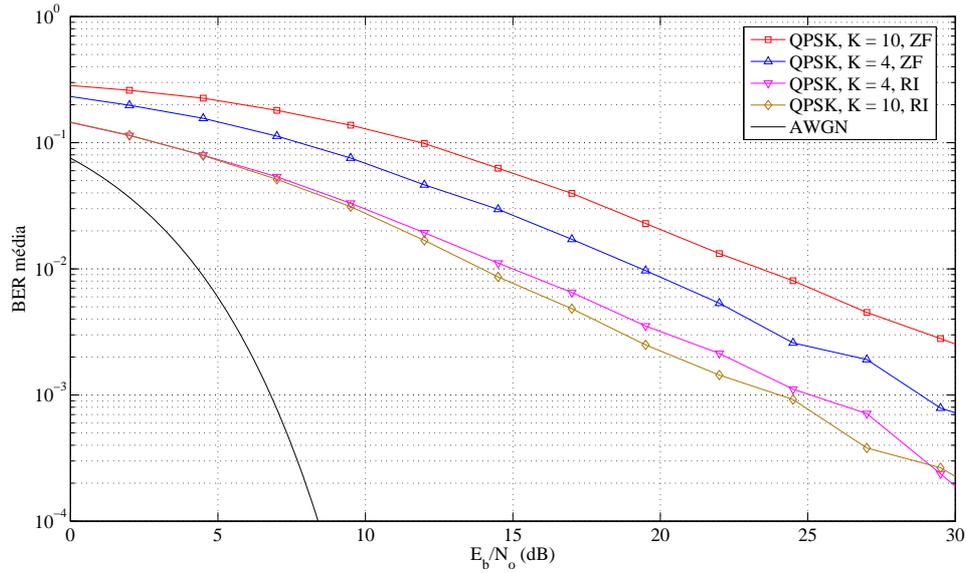


Figura 3.8: Comparação das curvas de BER dos esquemas de pré-codificação por inversão de canal e regularização de inversa para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.

3.5 BLOCK DIAGONALIZATION

Até este ponto, supôs-se que cada usuário possui apenas uma antena de recepção. Os algoritmos considerados anteriormente podem ser estendidos de maneira trivial para receptores com múltiplas antenas, considerando cada antena como um usuário separado, desde que $N_R \leq M_T$, o número de total de antenas de recepção (considerando todos os K usuários) seja menor ou igual ao número de antenas de transmissão.

Apesar de permitir arquiteturas extremamente simples de recepção, esta abordagem ignora a capacidade dos receptores de discriminar espacialmente seus fluxos de dados (*streams*), e é viável apenas quando o número de usuários é pequeno. Neste caso específico, cada antena seria considerada um "usuário virtual", e cada fluxo transmitido seria decodificada de forma independente em cada antena de recepção, como se a transmissão fosse realizada por N_R canais SISO paralelos. Apesar de essa abordagem resultar em receptores bastante simples, em geral ela leva a um desempenho subótimo pois, se múltiplas antenas estão presentes nos equipamentos dos usuários, estes são capazes de discriminar espacialmente seus próprios fluxos [61]. Como consequência, haverá um aumento significativo do hiato entre a capacidade alcançável por esses sistemas e o limite teórico do DPC. Logo, em vez de tentar diagonalizar completamente o canal (que corresponde à situação apresentada anteriormente), pode-se encontrar uma solução bloco diagonal: esta abordagem remove a interferência interusuário, mas preserva a interferência gerada entre os próprios fluxos do usuário.

A discussão anterior será formalizada. Em vez de forçar $\mathbf{H}\mathbf{W}$ na Eq. (3.4) a ser igual a matriz identidade, uma alternativa é fazê-la bloco diagonal. Dessa forma, a interferência interusuário é eliminada, mas é requerido que cada usuário realize algum tipo de processamento (demultiplexação espacial) de forma a separar e decodificar cada uma dos fluxos destinados a ele. Para ser preciso, o objetivo é encontrar a matriz de pré-codificação $\mathbf{W} = \mathbf{W}_{BD}$ de forma que

$$\mathbf{H}\mathbf{W}_{DB} = \begin{bmatrix} \mathbf{M}_1 & & \\ & \ddots & \\ & & \mathbf{M}_K \end{bmatrix}. \quad (3.20)$$

A matriz \mathbf{M}_k é de dimensão $n_{Rk} \times n_{Rk}$, supondo que são transmitidos $m_k = n_{Rk}$ fluxos para o usuário k . Entretanto, algumas das colunas de \mathbf{M}_k podem ser nulas se $m_k < n_{Rk}$. Entre os critérios que podem ser utilizados para determinar \mathbf{M}_k , destaca-se o algoritmo apresentado em [15, 62], denominado *block diagonalization* (BD), que é capaz de alcançar a soma das capacidades sob a restrição bloco diagonal imposta pela Eq. (3.20).

Em primeiro lugar, define-se a matriz $\tilde{\mathbf{H}}_k$ de dimensão $(N_R - n_{Rk}) \times M_T$:

$$\tilde{\mathbf{H}}_k = \left[\mathbf{H}_1^T \quad \dots \quad \mathbf{H}_{k-1}^T \quad \mathbf{H}_{k+1}^T \quad \dots \quad \mathbf{H}_K^T \right]^T, \quad (3.21)$$

que corresponde a matriz combinada de canal de todos os usuários, com exceção do k -ésimo.

Se o posto de $\tilde{\mathbf{H}}_k$ é \tilde{L}_k , então seu espaço nulo possui dimensão $M_T - \tilde{L}_k \geq n_{Rk}$. Por meio de uma decomposição em valores singulares, pode-se reescrever a Eq. (3.21) como

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k \tilde{\Sigma}_k \left[\tilde{\mathbf{V}}_k^{(1)} \quad \tilde{\mathbf{V}}_k^{(0)} \right]^H. \quad (3.22)$$

Na Eq. (3.22), $\tilde{\mathbf{V}}_k^{(0)}$ contém os últimos $M_T - \tilde{L}_k$ vetores singulares direitos de $\tilde{\mathbf{H}}_k$, formando uma base ortogonal para seu espaço nulo. Consequentemente, suas colunas são candidatas à matriz de pré-codificação do k -ésimo usuário, \mathbf{W}_{BDk} .

Seja então \bar{L}_k o posto do produto $\tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k^{(0)}$. Para que a transmissão para o usuário k ocorra livre de interferência, $\bar{L}_k \geq 1$ é uma condição necessária. Conforme mostrado em [63], uma condição suficiente é garantir que pelo menos uma linha de \mathbf{H}_k seja linearmente independente das linhas de $\tilde{\mathbf{H}}_k$. Uma forma de satisfazer tal condição é evitar multiplexar espacialmente usuários cujas matrizes de canal possuam elevado grau de correlação.

Supondo que a condição de independência é satisfeita para todos os usuários, $\tilde{\mathbf{V}}_k^{(0)}$ pode possuir mais *beamformers* do que o número de fluxos que o k -ésimo usuário pode suportar. Uma combinação ótima desses vetores de *beamformers* deve ser encontrada para formar \mathbf{W}_{BDk} , sendo este um problema já bastante explorado em sistemas MIMO ponto-a-ponto [64]. Para resolvê-lo, pode-se utilizar novamente a decomposição em valores singulares de $\tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k^{(0)}$:

$$\tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k^{(0)} = \begin{bmatrix} \mathbf{U}_k^{(1)} & \mathbf{U}_k^{(0)} \end{bmatrix} \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_k^{(1)} & \mathbf{V}_k^{(0)} \end{bmatrix}^H, \quad (3.23)$$

em que Σ_k é uma matriz diagonal quadrada $L_k \times L_k$ que contém os autovalores de $\tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k^{(0)}$, e $\mathbf{V}_k^{(1)}$ contém os L_k autovetores associados aos autovalores não-nulos de $\tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k^{(0)}$. As $L_k \leq n_{Rk}$ colunas do produto $\tilde{\mathbf{V}}_k^{(0)} \mathbf{V}_k^{(1)}$ representam os pesos que maximizam a capacidade do usuário k , sujeito à restrição de não gerar nenhuma interferência nos demais usuários.

A matriz de pré-codificação \mathbf{W}_{BD} será da seguinte forma:

$$\mathbf{W}_{BD} = \left[\tilde{\mathbf{V}}_1^{(0)} \mathbf{V}_1^{(1)} \quad \dots \quad \tilde{\mathbf{V}}_K^{(0)} \mathbf{V}_K^{(1)} \right] \mathbf{P}^{\frac{1}{2}}, \quad (3.24)$$

em que $\mathbf{W}_{BDk} = \mathbf{V}_k^{(0)} \mathbf{V}_k^{(1)}$. A matriz \mathbf{P} é do tipo diagonal e seus elementos representam a potência alocada para cada um dos usuários.

Se a matriz de pré-codificação \mathbf{W}_{BD} é dada pela Eq. (3.24), a soma das capacidade do método da pré-codificação por *block diagonalization* será dada por

$$C_{BD}^{sum} = \max_{\mathbf{P}, \text{Tr}(\mathbf{P})=P_T} \log \left| \mathbf{I} + \frac{1}{\sigma_n^2} \mathbf{\Sigma}^2 \mathbf{P} \right|, \quad (3.25)$$

com

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & & \\ & \ddots & \\ & & \mathbf{\Sigma}_K \end{bmatrix}. \quad (3.26)$$

A escolha ótima para as entradas da matriz \mathbf{P} é encontrada utilizando o algoritmo de *water-filling* nos elementos diagonais de $\mathbf{\Sigma}$, sujeito a restrição de potência P_T [65].

Conforme mostrado em [57], haverá um hiato entre a soma de capacidades do BD e do limite dado pelo DPC. Admitindo que todos os usuários possuem o mesmo número de antenas, dado por N_R/K , este hiato entre as somas de capacidades é calculado por meio da expressão

$$C_{DPC}^{sum} - C_{BD}^{sum} = \log_2 e \left(\sum_{k=0}^{K-1} \sum_{n=0}^{\frac{N_R}{K}-1} \sum_{i=k}^{\frac{N_R}{K}} \frac{1}{M_T - n - i} \right). \quad (3.27)$$

A Eq. (3.27) é simplificada para a Eq. (3.11) quando $N_R/K = 1$, isto é, a técnica de pré-codificação *zero-forcing* é um caso particular da pré-codificação por *block diagonalization*.

Apesar de ambas as técnicas, *zero-forcing* e *block diagonalization*, fornecerem o mesmo ganho de multiplexação espacial, existirá diferença no *throughput* final do sistema. Para fornecer alguma intuição sobre o significado deste fato, será tomado o caso em que M_T é mantido constante, N_R/K aumenta mas N_R é mantido constante, fazendo com que K diminua. Em outras palavras, o número de antenas por receptor aumenta, mas o número total de antenas de recepção N_R se mantém contante, diminuindo a quantidade K de usuários. Neste caso, o hiato entre a soma das capacidades diminui, conforme mostrado nos resultados de simulações. Na Fig. 3.9, tem-se $M_T = 6$ antenas e, na Fig. 3.10, tem-se o caso para o qual $M_T = 10$ antenas de transmissão. A notação $\{3, 3\}$, indica que há $K = 2$ usuários, cada um com $n_{R1} = n_{R2} = 3$ antenas de recepção. Pelo comportamento das curvas de capacidade, pode-se inferir que a soma das capacidades do DPC é a mesma de um canal MIMO $N_R \times M_T$. Entretanto, a soma de capacidades do *block diagonalization* é K vezes a capacidade de um canal MIMO $N_R/K \times (M_T - N_R + 1)$, enquanto que a capacidade do *zero-forcing* será de N_R canais MISO $(M_T - N_R + 1) \times 1$.

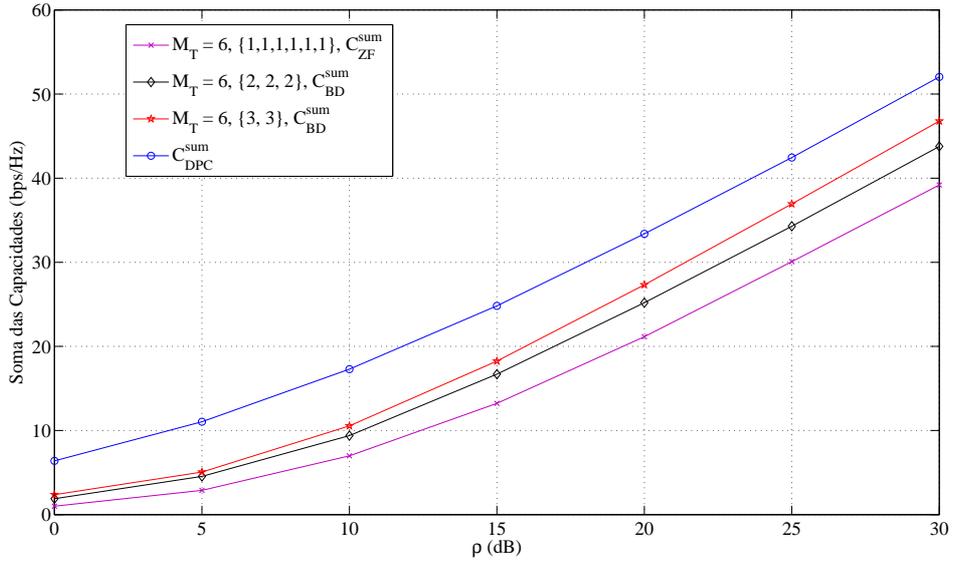


Figura 3.9: Comparação da soma das capacidades em função de ρ da técnica de pré-codificação *block diagonalization* para $M_T = 6$ antenas transmissoras e diferentes configurações de número de usuários e de antenas nos receptores.

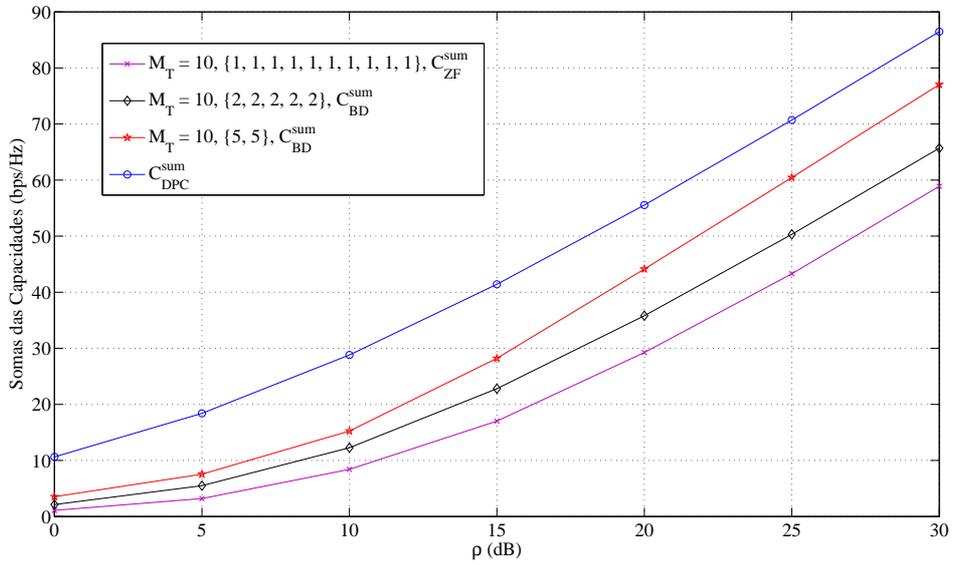


Figura 3.10: Comparação da soma das capacidades em função de ρ da técnica de pré-codificação *block diagonalization* para $M_T = 10$ antenas transmissoras e diferentes configurações de número de usuários e de antenas nos receptores.

3.6 PRÉ-CODIFICAÇÃO TOMLINSON-HARASHIMA

3.6.1 Modelo do Sistema

A pré-codificação Tomlinson-Harashima foi inicialmente proposta de maneira independente por Tomlinson [66] e Harashima [67] para eliminar a interferência intersimbólica em canais seletivos em frequência em constelações do tipo MPAM (do inglês, *M-ary pulse amplitude modulation*) utilizando a aritmética modular. A estratégia proposta leva ainda a estrutura de equalização DFE (do inglês, *decision feedback equalizer*) para o transmissor, de forma a eliminar a propagação de erros causada por detecções errôneas dos símbolos recebidos [58], pois o transmissor conhece, *a priori*, os símbolos a serem transmitidos. Sua extensão para pré-equalização em canais MIMO foi considerada em [68, 69, 70, 71, 72], e o modelo do sistema está apresentado na Fig. 3.11.

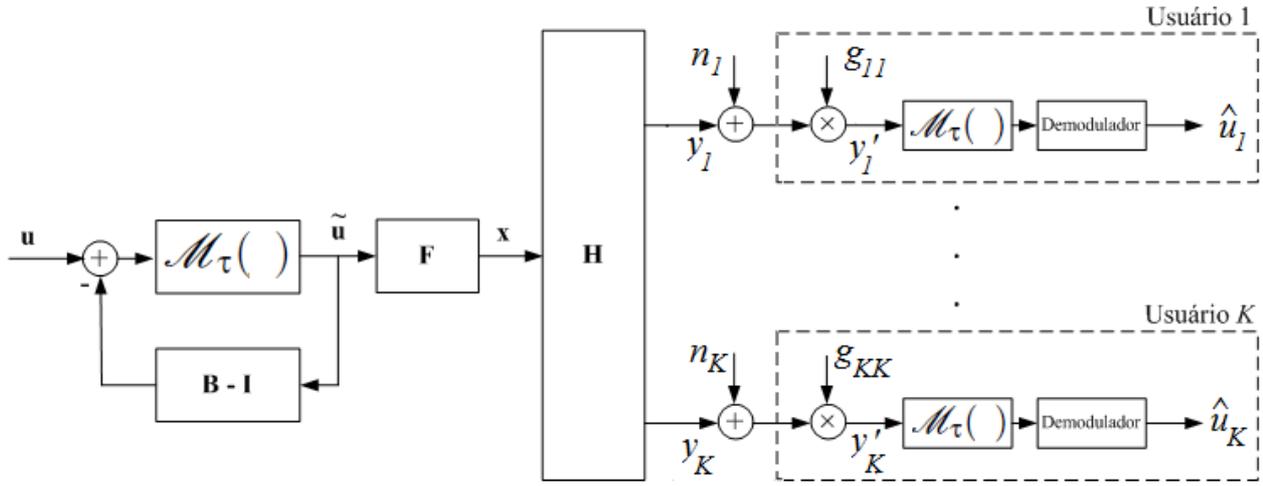


Figura 3.11: Estrutura do pré-codificador Tomlinson-Harashima para o enlace direto de sistemas MIMO multiusuário.

A matriz \mathbf{F} , de dimensão $M_T \times K$, tem como função tornar a interferência espacialmente causal, o que significa que \tilde{u}_k sofre a interferência de \tilde{u}_l , $l < k$, mas não de \tilde{u}_i , $i > k$. Esta matriz é equivalente ao filtro de *feedforward* de equalizadores clássicos DFE quando utilizados no receptor [58].

A interferência causal (isto é, a componente de interferência de \tilde{u}_i em \tilde{u}_k , com $i < k$), é eliminada pelo filtro de *feedback* $\mathbf{B} - \mathbf{I}$, com a matriz \mathbf{B} triangular inferior, de dimensão $K \times K$ e com elementos unitários em sua diagonal principal [73]. A operação do filtro de *feedback* é realizada de forma sucessiva, do símbolo u_1 para o símbolo u_K , e pode ser descrita como

$$\tilde{u}_k = \mathcal{M}_\tau \left(u_k - \sum_{i=1}^{k-1} b_{k,i} u_i \right), \quad (3.28)$$

em que $b_{k,i}$ é o elemento da linha k e coluna i da matriz \mathbf{B} , e \mathcal{M}_τ denota o operador módulo τ , definido por

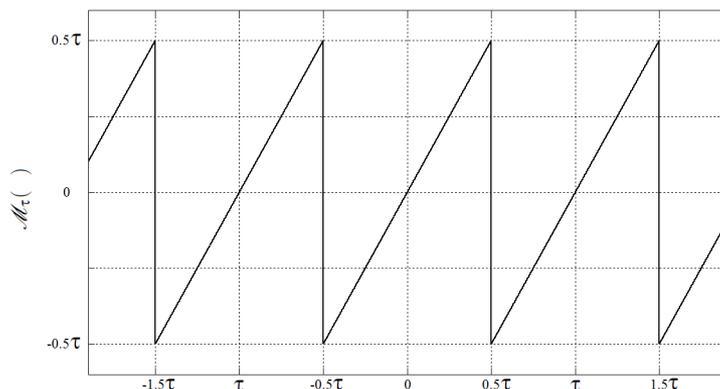


Figura 3.12: Implementação do operador módulo por meio de uma função dente de serra.

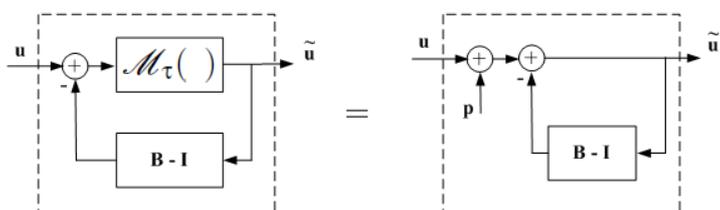


Figura 3.13: Descrição linearizada do operador módulo.

$$\mathcal{M}_\tau(u_k) = u_k - \left\lfloor \frac{u_k + \tau/2}{\tau} \right\rfloor \tau, \quad (3.29)$$

em que τ é um valor real positivo denominado constante do módulo. O gráfico do operador módulo é mostrado na Fig. 3.12. Seu funcionamento é da seguinte forma: se em sua entrada é apresentado um valor maior que $\tau/2$, τ é subtraído repetidas vezes do valor de entrada até que o resultado produzido em sua saída seja menor que $\tau/2$. Se na entrada do operador é apresentado um valor menor que $-\tau/2$, τ é adicionado de forma sucessiva até que o resultado observado em sua saída seja maior ou igual a $-\tau/2$. Em outras palavras, adicionam-se múltiplos inteiros de τ à entrada para limitá-la ao intervalo $[-\tau/2, \tau/2]$.

O pré-codificador gera um vetor de dados efetivo $\mathbf{u} + \mathbf{p}$ como entrada para \mathbf{B}^{-1} (representada pelo laço de realimentação na Fig. 3.13), em que \mathbf{p} representa o vetor de pré-codificação. Ele pode ser escolhido de forma que as entradas de $\tilde{\mathbf{u}} = \mathbf{B}^{-1}(\mathbf{u} + \mathbf{p})$ pertençam a determinada faixa estabelecida de valores [70]. Isto dá origem à descrição linearizada do operador módulo, mostrada na Fig. (3.13).

Por exemplo, se uma constelação quadrada MQAM é utilizada, os elementos u_k são oriundos do conjunto

$$\mathcal{A} = \left\{ u^I + ju^Q : u^I, u^Q \in \left\{ \pm 1, \pm 3, \dots, \pm\sqrt{M} - 1 \right\} \right\}, \quad (3.30)$$

em que u^I e u^Q referem-se, respectivamente, às componentes em fase e em quadratura do símbolo u e j é a unidade imaginária, $j = \sqrt{-1}$. A componente p_k de \mathbf{p} deve satisfazer à condição $p_k = \tau(a_k + b_k j)$, $a_k, b_k \in \mathbb{Z}$, e podem ser escolhidas de tal forma que $\tilde{u}_k^I, \tilde{u}_k^Q \in [-\sqrt{M}, \sqrt{M}]$, em que \tilde{u}_k^I e \tilde{u}_k^Q

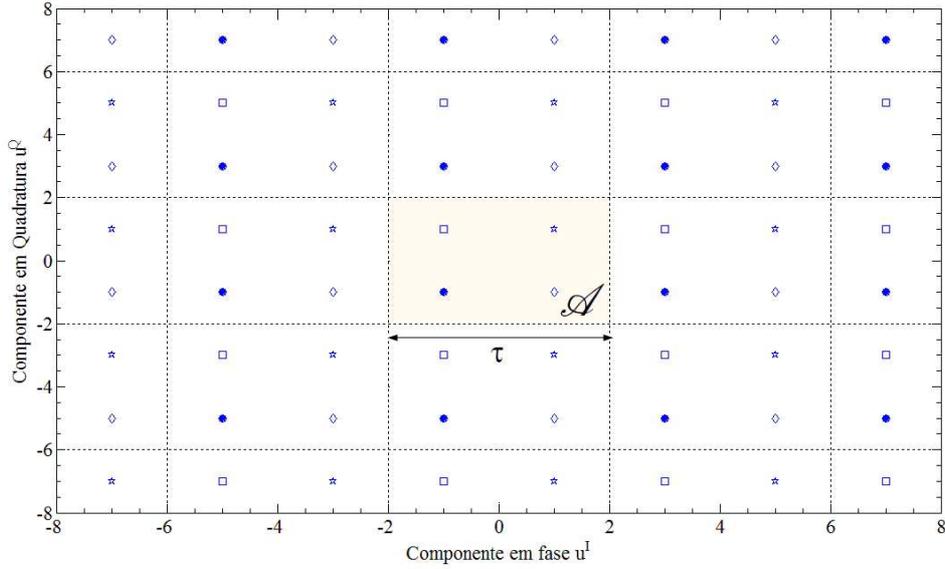


Figura 3.14: Constelação QPSK estendida. É mostrada a constelação original, \mathcal{A} e algumas de suas extensões periódicas utilizando $\tau = 4$. Os símbolos módulo-congruentes são representados pelos mesmos marcadores.

representam, respectivamente, as componentes em fase e em quadratura de \tilde{u}_k , o k -ésimo elemento do vetor $\tilde{\mathbf{u}}$. Neste caso, o operador módulo é aplicado separadamente nas partes real e imaginária de u_k . Por exemplo, suponha a utilização de um alfabeto 4QAM, no qual as partes real e imaginária dos símbolos podem assumir os valores ± 1 . Escolhendo $\tau = 4$, a saída do operador módulo sempre estará entre -2 e $+2$. Uma entrada $2,2 - 3,1j$ seria consequentemente mapeada em $-1,8 + 0,9j$. Neste caso, o módulo adicionou o valor $-4 + 4j$ ao símbolo de sua entrada, que corresponde a $4(-1 + 1j) = \tau(-1 + 1j)$.

É interessante notar ainda que, conforme descrito em [74, 75, 76], a pré-codificação por aritmética modular é baseada no conceito de espaço de sinais equivalentes. É gerada uma constelação estendida, ou um reticulado (do inglês, *lattice*), \mathcal{A}' por meio da chamada extensão periódica da constelação original:

$$\tilde{\mathcal{A}}' = \mathcal{A} + \tau(a + jb), \quad a, b \in \mathbb{Z} \quad (3.31)$$

O conceito de constelação estendida é mostrado na Fig. 3.14 para a modulação QPSK utilizando $\tau = 4$, em que os símbolos módulo-congruentes são representados pelos mesmos marcadores. Dois símbolos u e v são ditos módulo-congruentes se $u = \mathcal{M}_\tau(v)$.

Uma interpretação alternativa para a pré-codificação Tomlinson-Harashima é a seguinte: a partir de \mathcal{A}' , são selecionados os elementos $u_k + p_k$ do vetor de dados efetivos que são módulo-congruentes a u_k e que minimizam o módulo de \tilde{u}_k . Logo, a pré-codificação Tomlinson-Harashima é uma extensão dos algoritmos de pré-codificação lineares apresentados, porém utilizando os sinais do reticulado \mathcal{A}' , em vez do conjunto \mathcal{A} .

Naturalmente, a escolha para o valor de τ afeta o desempenho do sistema. Valores muito grandes são interessantes para mitigar os efeitos do ruído no receptor, já que as extensões da constelação original estarão posicionadas mais distantes umas das outras (ver Eq. (3.31)). Pequenos valores de τ podem

provocar a sobreposição das constelações, de forma que a decodificação biunívoca do símbolo recebido não será possível, o que contribuiria para aumentar a taxa de erro do sistema. De forma geral, pequenos valores de τ são vantajosos pois permitem um reticulado mais denso e mais flexível, mas ainda assim deve ser escolhido de forma a permitir decodificação sem ambiguidades. Em [16], é sugerido o valor de

$$\tau = 2 \left(\mathcal{A}_{\max} + \frac{\Delta}{2} \right), \quad (3.32)$$

em que definiu-se $\mathcal{A}_{\max} = \max(u^I, u^Q)$, e Δ é o menor espaçamento entre dois símbolos da constelação. Por exemplo, para a constelação da modulação QPSK, $\mathcal{A}_{\max} = 1$, $\Delta = 2$ e $\tau = 4$. Já para constelação da modulação 16QAM, $\mathcal{A}_{\max} = 3$, $\Delta = 2$ e $\tau = 8$, conforme já considerado anteriormente.

No receptor, o operador módulo também deve estar presente para que \mathbf{u} seja recuperado a partir de $\mathbf{u} + \mathbf{p}$, conforme mostra a Fig. 3.15, ou seja, desfazer os efeitos do operador módulo presente no transmissor. Os fatores de escala $g_{1,1}, g_{2,2}, \dots, g_{K,K}$ são determinados no transmissor, e devem também ser enviados aos receptores.

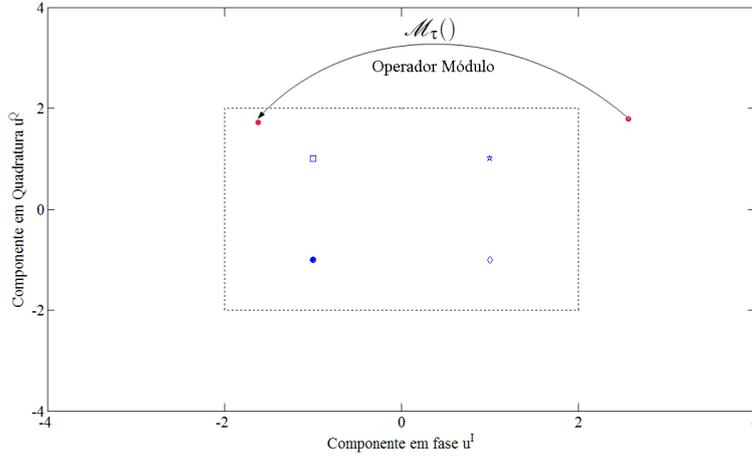


Figura 3.15: Função do operador módulo.

Apesar de os valores de saída $\tilde{\mathbf{u}}$ estarem limitados em amplitude, a pré-codificação Tomlinson-Harashima provoca aumento na potência média de transmissão (ou, de forma equivalente, perda na relação sinal-ruído, devido à presença da restrição de potência P_T no transmissor). Esta perda, denominada *shapping loss*, pode ser mensurada por meio da razão

$$\Gamma_{loss} = \frac{\mathbb{E} \left[\|\tilde{\mathbf{u}}\|^2 \right]}{\mathbb{E} \left[\|\mathbf{u}\|^2 \right]}. \quad (3.33)$$

Em [70], mostrou-se que para constelações quadradas, a perda de pré-codificação pode ser aproximada pela razão

$$\Gamma_{loss} = \frac{M}{M-1}. \quad (3.34)$$

Para pequenos valores da ordem M da modulação, a perda Γ_{loss} introduzida não é desprezível, mas

decai rapidamente conforme o tamanho do alfabeto utilizado pela modulação aumenta. Por exemplo, foi encontrado o valor de $\Gamma_{loss} = 1,25$ dB para as modulações BPSK e QPSK, e de $\Gamma_{loss} = 0,28$ dB para 16QAM [77].

Para o sistema representado na Fig. 3.11, será introduzida a matriz diagonal \mathbf{G} , que contém os fatores de escala utilizados pelos receptores para a detecção dos símbolos recebidos. Em outras palavras, define-se a matriz $\mathbf{G} = \text{diag}[g_{1,1}, g_{2,2}, \dots, g_{K,K}]$, de forma que, na entrada do operador módulo, o sinal recebido por cada usuário, dado pelos os elementos y'_1, y'_2, \dots, y'_K , pode ser expresso por

$$\begin{aligned} \mathbf{y}' &= \mathbf{G}\mathbf{y} \\ &= \mathbf{G}\mathbf{H}\mathbf{F}\mathbf{B}^{-1}\mathbf{u} + \mathbf{G}\mathbf{n} \\ &= \mathbf{G}\mathbf{H}\mathbf{F}\mathbf{B}^{-1}\mathbf{u} + \tilde{\mathbf{n}}, \end{aligned} \quad (3.35)$$

em que $\mathbf{y}' = [y'_1, \dots, y'_K]^T$ e $\tilde{\mathbf{n}} = [\tilde{n}_1, \dots, \tilde{n}_K]^T = \mathbf{G}\mathbf{n}$.

Nas próximas seções, serão apresentados os critérios utilizados para a determinação das matrizes de pré-codificação \mathbf{B} , \mathbf{F} e \mathbf{G} .

3.6.2 Pré-codificação Zero-Forcing Tomlinson-Harashima

Se é desejado que toda a interferência interusuário seja eliminada, então

$$\mathbf{G}\mathbf{H}\mathbf{F}\mathbf{B}^{-1} = \mathbf{I} \quad (3.36)$$

é uma condição necessária. As matrizes \mathbf{B} , \mathbf{F} e \mathbf{G} podem ser encontradas por meio da decomposição QR da matriz \mathbf{H}^H , expressa por meio de

$$\mathbf{H}^H = \begin{bmatrix} \mathbf{Q}' & \tilde{\mathbf{Q}}' \end{bmatrix} \begin{bmatrix} \mathbf{R}' \\ \mathbf{0}_{(M_T-K) \times K} \end{bmatrix} \quad (3.37)$$

em que \mathbf{Q}' é uma matriz $M_T \times K$ cujas colunas são ortonormais, $\tilde{\mathbf{Q}}'$ é $M_T \times (M_T - K)$, \mathbf{R}' é uma matriz triangular superior $K \times K$ cujo elemento da i -ésima linha e j -ésima coluna é denotado por $r'_{i,j}$, e $\mathbf{0}_{(M_T-K) \times K}$ é a matriz $(M_T - K) \times K$ cujos elementos são todos nulos. Logo, conforme demonstrado em [49], as matrizes procuradas serão dadas por:

$$\mathbf{G} = \text{diag} \left[r'_{1,1}, r'_{2,2}, \dots, r'_{K,K} \right], \quad (3.38)$$

$$\mathbf{B} = \mathbf{G}\mathbf{R}'^H \quad (3.39)$$

e

$$\mathbf{F} = \mathbf{Q}'. \quad (3.40)$$

A partir da Eq. (3.35), da Eq. (3.36) e da descrição linearizada do operador módulo, mostrada na Fig. 3.13

$$\mathbf{y}' = \mathbf{u} + \mathbf{p} + \tilde{\mathbf{n}}, \quad (3.41)$$

de forma que toda a interferência é eliminada, e o operador módulo no receptor eliminará a componente \mathbf{p} .

A matriz de autocovariância do ruído será dada por

$$\begin{aligned} \mathbb{E} [\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H] &= \sigma_n^2 \mathbf{G}^2 \\ &= \text{diag} \left[\frac{\sigma_n^2}{r'_{1,1}}, \dots, \frac{\sigma_n^2}{r'_{K,K}} \right]. \end{aligned} \quad (3.42)$$

As Eq. (3.38)-(3.40), juntamente com a restrição dada pela Eq. (3.36), definem o pré-codificador *zero-forcing* Tomlinson-Harashima (ZF THP).

3.6.3 Pré-codificação Minimum Mean Square Error Tomlinson-Harashima

Uma outra abordagem para a obtenção das matrizes de pré-codificação é utilizar o critério de minimizar o erro quadrático médio. O pré-codificador, neste caso, é denominado MMSE THP.

Pode-se definir o vetor erro \mathbf{e} como a diferença entre o vetor \mathbf{y}' na entrada do operador módulo e o vetor de dados efetivos $\mathbf{u} + \mathbf{p}$:

$$\begin{aligned} \mathbf{e} &= \mathbf{y}' - (\mathbf{u} + \mathbf{p}) \\ &= (\mathbf{GHF} - \mathbf{B}) \tilde{\mathbf{u}} + \mathbf{Gn}. \end{aligned} \quad (3.43)$$

A matriz de covariância do erro será então

$$\begin{aligned} \mathcal{R}_{\mathbf{e}} &= \mathbb{E} [\mathbf{e}\mathbf{e}^H] \\ &= (\mathbf{GHF} - \mathbf{B}) \mathcal{R}_{\tilde{\mathbf{u}}} (\mathbf{GHF} - \mathbf{B})^H + \sigma_n^2 \mathbf{G}\mathbf{G}^H, \end{aligned} \quad (3.44)$$

em que $\mathcal{R}_{\tilde{\mathbf{u}}}$ é a matriz de covariância do vetor $\tilde{\mathbf{u}}$. O critério MMSE busca minimizar $\text{Tr} [\mathcal{R}_{\mathbf{e}}]$ com respeito as matrizes \mathbf{B} , \mathbf{F} e \mathbf{G} . Devido às restrições já apresentadas, \mathbf{B} deve ser uma matriz triangular unitária e \mathbf{G} deve ser diagonal, já que não existe o processamento cooperativo entre os receptores no canal MIMO BC. Ainda é desejado que a matriz \mathbf{F} possua colunas ortogonais e não afete a potência de transmissão. As condições são satisfeitas fazendo $\mathbf{F}\mathbf{F}^H = \mathbf{I}$, da mesma forma como ocorre com o ZF THP.

Pelo princípio da ortogonalidade [78], a solução MMSE deve satisfazer

$$\mathbb{E} [\mathbf{e}\mathbf{y}^H] = \mathbf{0}. \quad (3.45)$$

Logo,

$$\mathbb{E} [\mathbf{G}\mathbf{y}\mathbf{y}^H - \mathbf{B}\tilde{\mathbf{u}}\mathbf{y}^H] = \mathbf{0}, \quad (3.46)$$

$$\mathbf{G} (\mathbf{H}\mathbf{F}\mathcal{R}_{\tilde{\mathbf{u}}}\mathbf{F}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}) = \mathbf{B}\mathcal{R}_{\tilde{\mathbf{u}}}\mathbf{F}^H\mathbf{H}^H. \quad (3.47)$$

Supondo que a matriz de autocovariância de $\tilde{\mathbf{u}}$ seja dada por $\mathcal{R}_{\tilde{\mathbf{u}}} = \sigma_{\tilde{\mathbf{u}}}^2\mathbf{I}$, e introduzindo o fator $\xi = \sigma_n^2/\sigma_{\tilde{\mathbf{u}}}^2$, a Eq. (3.47) se torna

$$\mathbf{G} (\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H + \xi\mathbf{I}) = \mathbf{B}\mathbf{F}^H\mathbf{H}^H. \quad (3.48)$$

Por meio da fatoração de Cholesky de $(\mathbf{H} + \xi\mathbf{H}^{\dagger H})(\mathbf{H} + \xi\mathbf{H}^{\dagger H})^H$, é possível obter uma matriz \mathbf{R} triangular inferior, de dimensão $K \times K$, que satisfaz

$$\mathbf{R}\mathbf{R}^H = (\mathbf{H} + \xi\mathbf{H}^{\dagger H})(\mathbf{H} + \xi\mathbf{H}^{\dagger H})^H. \quad (3.49)$$

Conforme mostrado em [49], a solução para o MMSE THP é formada pelas matrizes

$$\mathbf{G} = \text{diag} [r'_{1,1}, r'_{2,2}, \dots, r'_{K,K}], \quad (3.50)$$

$$\mathbf{B} = \mathbf{G}\mathbf{R}, \quad (3.51)$$

e

$$\mathbf{F} = (\mathbf{H} + \xi\mathbf{H}^{\dagger}) (\mathbf{R}^{-1})^H. \quad (3.52)$$

É ainda possível calcular a matriz de autocovariância do erro quadrático médio. Substituindo as soluções encontradas, dadas pelas Eq. (3.50)-(3.52), na Eq. (3.44), obtém-se o resultado

$$\mathcal{R}_e = \sigma_n^2\mathbf{G}^2 + \sigma_n^2\xi\mathbf{G}(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{G}. \quad (3.53)$$

A Fig. 3.16 mostra o desempenho das técnicas de pré-codificação ZF THP e MMSE THP para a modulação QPSK, e a Fig. 3.17 mostra o desempenho das técnicas de pré-codificação ZF THP e MMSE THP para a modulação 16QAM, variando-se o número de usuários. Ao contrário da técnica *zero-forcing*, é possível perceber que o aumento do número de usuários contribui para o melhor desempenho da técnica, apesar do ganho ainda ser muito próximo do conseguido pela técnica de pré-codificação por regularização de inversa.

Na Fig. 3.18 é mostrado o desempenho das técnicas baseadas em pré-codificação Tomlinson-Harashima quando comparadas aos algoritmos lineares para um cenário com 10 usuários utilizando a modulação QPSK. Como é possível observar, o ganho com relação às técnicas de *zero-forcing* e regularização de inversa é desprezível. Este fato é justificado pela perda de relação sinal-ruído introduzida

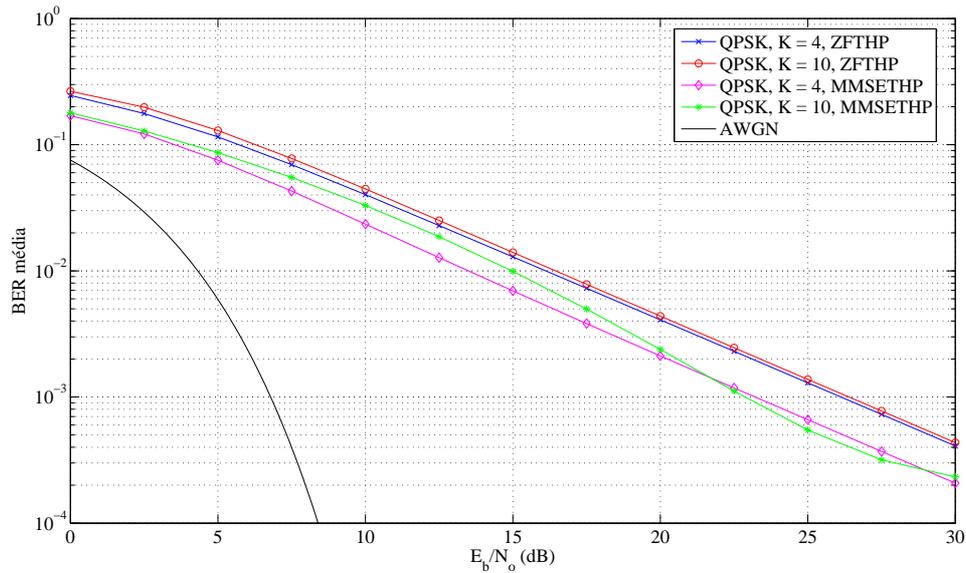


Figura 3.16: Comparação das curvas de BER dos esquemas de pré-codificação ZF THP e MMSE THP para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.

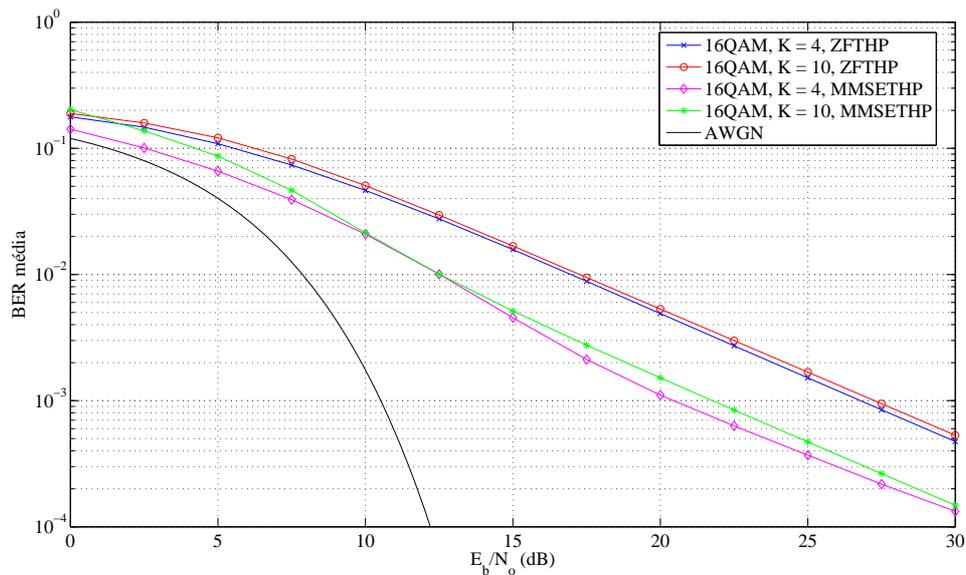


Figura 3.17: Comparação das curvas de BER dos esquemas de pré-codificação ZF THP e MMSE THP para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.

pele *shaping loss*, que será mais acentuado ao se utilizar a modulação QPSK. A Fig. 3.19 ilustra um cenário similar, porém utilizando a modulação 16QAM. Neste caso, é possível inferir que o ganho é significativo quando comparado aos algoritmos lineares. Este resultado sugere que a técnica de pré-codificação Tomlinson-Harashima depende não apenas da informação sobre o estado do canal, mas seu desempenho está intimamente ligado ao tipo de modulação (e ao alfabeto de símbolos da modulação) que é utilizado, já que ela é baseada no conceito de pré-codificação sobre um reticulado.

Além disso, por utilizar pré-codificação sucessiva dos usuários, a técnica Tomlinson-Harashima é

sensível à ordem com a qual os usuários são pré-codificados, como será abordado na próxima seção.

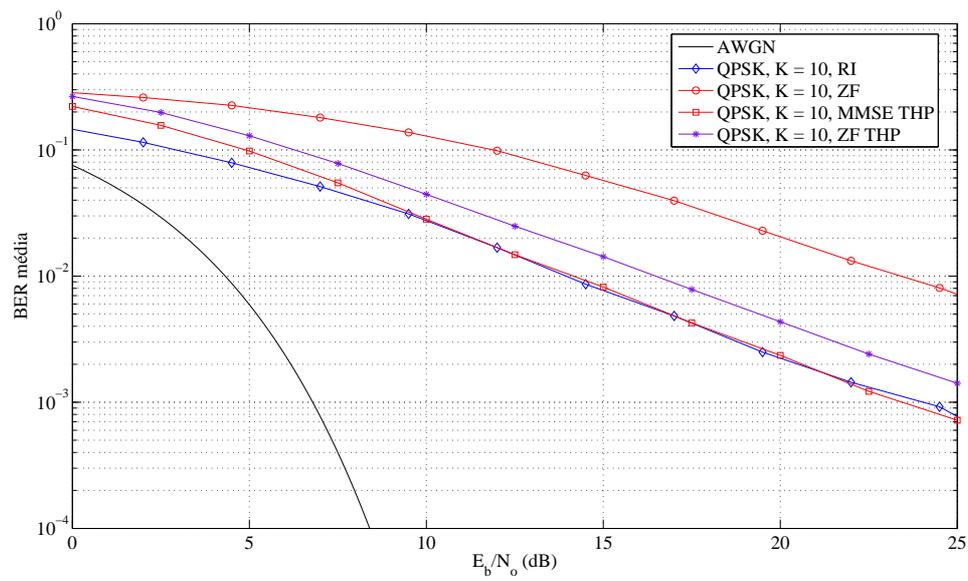


Figura 3.18: Comparação das técnicas de pré-codificação ZF THP e MMSE THP com os algoritmos lineares em um cenário com $K = 10$ usuários e modulação QPSK.

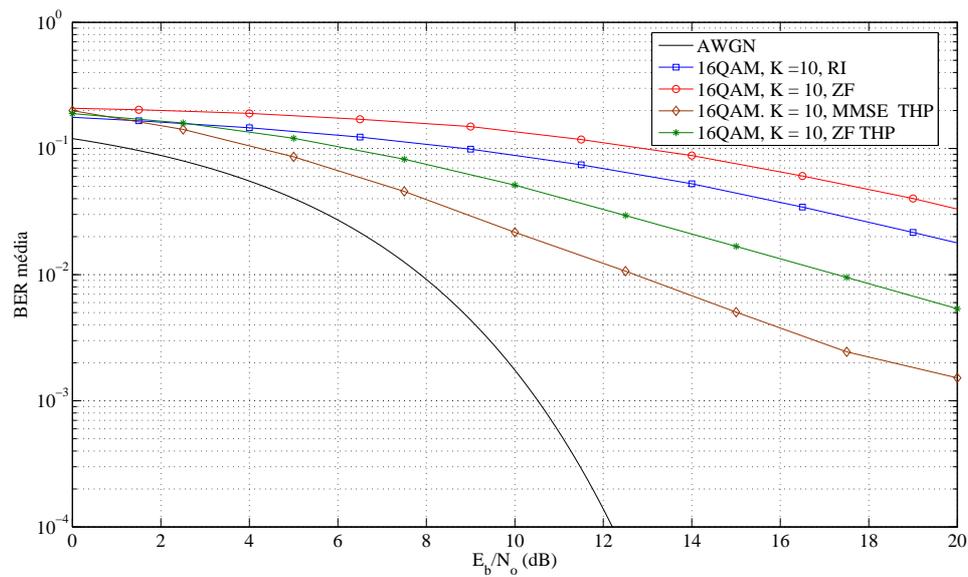


Figura 3.19: Comparação das técnicas de pré-codificação ZF THP e MMSE THP com os algoritmos lineares em um cenário com $K = 10$ usuários e modulação 16QAM.

3.6.4 A Ordenação dos Usuários

Para as técnicas de pré-codificação ZF THP e MMSE THP, apresentadas nas seções anteriores e dadas pelas Eq. (3.38)-(3.40) e Eq. (3.50)-(3.52), respectivamente, pode-se mostrar que a ordem das colunas da matriz de canal \mathbf{H} não exercerá influência nem nas matrizes de processamento \mathbf{B} e \mathbf{G} , nem na matriz de autocovariância do ruído (para o caso ZF THP), dada pela Eq. (3.42), ou na matriz de autocovariância do erro quadrático médio (para o caso MMSE THP), dada pela Eq. (3.53) [79]. Logo, a ordenação das colunas da matriz combinada de canal não afeta o desempenho do sistema.

O mesmo já não é verdade para a ordenação das linhas da matriz \mathbf{H} . Uma mudança na ordem das linhas da matriz combinada de canal modificará as matrizes \mathbf{B} e \mathbf{G} , bem como as matrizes de autocovariância do ruído ou do erro quadrático médio. Consequentemente, o desempenho das técnicas de pré-codificação ZF THP e MMSE THP pode ser melhorado por meio de uma ordenação conveniente das linhas de \mathbf{H} pois, conforme exposto em [49], o desempenho da BER do sistema estará limitado pelo usuário que possui ruído de maior variância (para o pré-codificador ZF THP) ou pelo usuário que possui o maior erro quadrático médio (para o pré-codificador MMSE THP).

Para perceber de forma intuitiva como a ordem das linhas da matriz \mathbf{H} afeta o desempenho do sistema, basta lembrar que supôs-se que o símbolo u_k é destinado ao usuário k . Para a estrutura considerada na Fig. 3.11, se nenhum ordenamento é aplicado, $\mathbf{u} = [u_1, u_2, \dots, u_K]$ é o vetor de dados na entrada do pré-codificador e, uma vez que a matriz \mathbf{B} é triangular inferior, a subtração sucessiva da interferência interusuário é realizada de u_1 para u_K , conforme expresso pela Eq. (3.28). Como a pré-codificação Tomlinson-Harashima é baseada no conceito de DPC, a ordem com a qual os usuários são codificados alterará o desempenho do sistema, conforme comentado no capítulo 2.

Antes de prosseguir, definir-se-á como métrica de desempenho \mathbb{M}_i o valor do i -ésimo elemento da diagonal da matriz de autocovariância do ruído, dada pela Eq. (3.42), ou o valor do i -ésimo elemento da diagonal da matriz de autocovariância do erro quadrático médio, dada pela Eq. (3.53). A primeira será considerada ao utilizar-se a pré-codificação ZF THP, e a segunda será utilizada ao tratar da pré-codificação MMSE THP.

A ordenação ótima é aquela que gera os menores valores possíveis de métricas de desempenho, \mathbb{M}_i , e é representada por $\pi = [\pi(1), \pi(2), \dots, \pi(K)]$, de tal forma que os usuários são pré-codificados de $\pi(1)$ para $\pi(K)$, o que dá origem ao vetor $\mathbf{u} = [u_{\pi(1)}, u_{\pi(2)}, \dots, u_{\pi(K)}]$ na entrada do pré-codificador.

Utilizando os conceitos desenvolvidos em [80, 81, 82] para detecção em sistemas MIMO monousuário, é possível mostrar que a ordenação ótima dos usuários é alcançada por meio do algoritmo *best-first* [49], detalhado em Algoritmo 1.

É interessante observar que, de acordo com o passo 4 do Algoritmo 1, para encontrar o i -ésimo elemento da ordenação π , é necessário realizar uma busca sobre todas as permutações possíveis entre os usuários que ainda não foram ordenados. Logo, a complexidade do algoritmo *best-first* aumenta rapidamente com o número de usuários e, por esse motivo, algoritmos subótimos de ordenação, baseados no conceito de otimizações sucessivas, podem ser encontrados em [83, 84].

Naturalmente, antes de realizar a pré-codificação, uma vez determinada a ordenação π , deve-se definir a matriz combinada de canal ordenada como

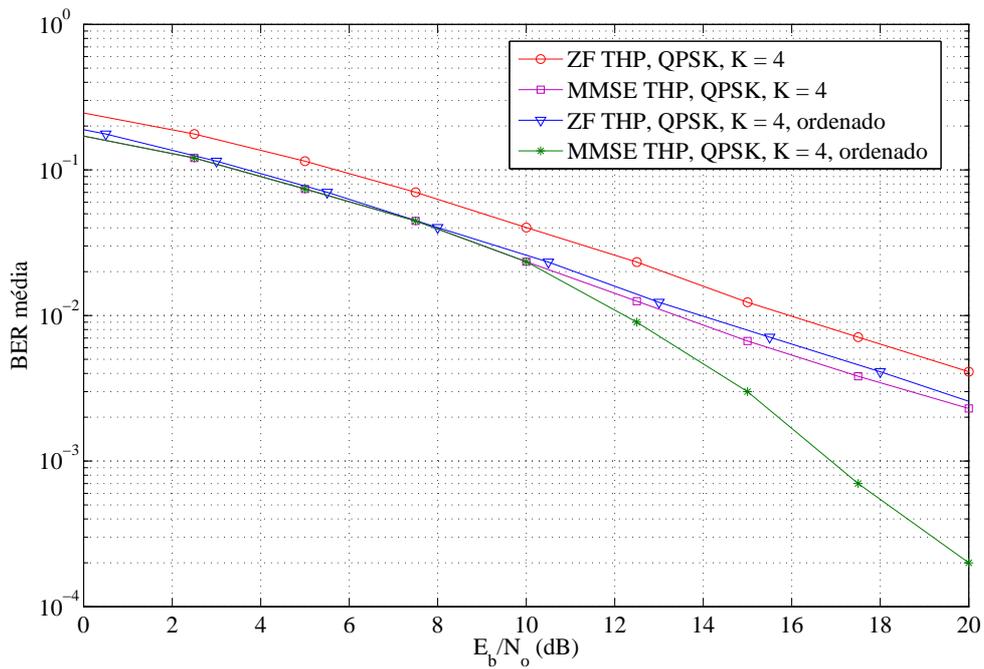
Algoritmo 1 Algoritmo *Best-First*

1. Inicialize o contador $i = 1$
 2. Crie o conjunto de K usuários $U_\pi = \{1, 2, \dots, K\}$
 3. Para $i < K$:
 4. Para todos os usuários que pertencem a U_π , encontre o usuário $\beta \in U_\pi$ que possui a menor métrica \mathbb{M}_i ao ocupar a posição i da ordenação π
 5. Coloque o usuário β , determinado no passo 4, na posição i da ordenação π , isto é, $\pi(i) = \beta$
 6. Exclua β do conjunto U_π
 7. Incremente i
 8. Fim
-

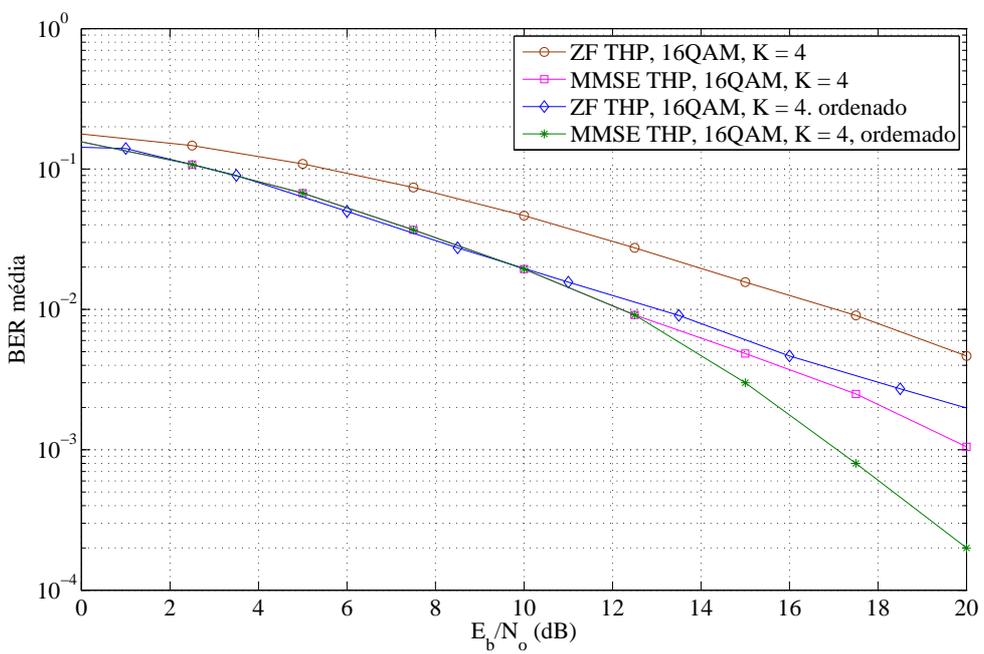
$$\mathbf{H}_\pi = \left[\mathbf{H}_{\pi(1)}^T, \mathbf{H}_{\pi(2)}^T, \dots, \mathbf{H}_{\pi(K)}^T \right]^T, \quad (3.54)$$

e as matrizes de processamento \mathbf{B} , \mathbf{F} e \mathbf{G} devem ser geradas utilizando \mathbf{H}_π ao invés de \mathbf{H} , e serão denotadas por \mathbf{B}_π , \mathbf{F}_π e \mathbf{G}_π .

Os resultados de simulação dos pré-codificadores baseados em Tomlinson-Harashima utilizando a ordenação de usuários por meio do algoritmo *best-first* são mostrados na Fig. 3.20 (a) para a modulação QPSK e na Fig. 3.20 (b) para a modulação 16QAM. Para ambas as modulações, a ordenação de usuários resulta em ganho de relação sinal-ruído para o algoritmo ZF THP. Este ganho é de aproximadamente 2 dB para a modulação QPSK, e de 4 dB para a modulação 16QAM. Para o pré-codificador MMSE THP, é caracterizado ganho de diversidade, pois há mudança na inclinação da curva de BER, especialmente no regime de SNR elevada. Nesta situação, tem-se um ganho superior a 4 dB para uma BER da ordem de 10^{-3} para ambas as modulações quando comparadas aos pré-codificadores Tomlinson-Harashima quando não utilizam a ordenação de usuários.



(a) Modulação QPSK, $K = 4$ usuários



(b) Modulação 16QAM, $K = 4$ usuários

Figura 3.20: Desempenho dos pré-codificadores ZF THP e MMSE THP com ordenação de usuários dada pelo algoritmo *best-first*.

3.7 VECTOR PERTURBATION

Conforme exposto na seção 3.3, a perda de capacidade inerente à pré-codificação *zero-forcing* está no aumento da potência do vetor transmitido \mathbf{x} (ou, de forma equivalente, na constante de normalização γ , que assumirá valores muito grandes devido ao comportamento do maior autovalor da matriz \mathbf{H}^{-1} quando $K = M_T$). Na seção 3.4, foi mostrada uma forma de regularização da matriz inversa que tenta controlar os efeitos da excursão dos valores deste maior autovalor. Entretanto, ainda permanece o hiato entre a soma das capacidades alcançadas por essas técnicas de pré-codificação e o limite de capacidade do canal MIMO BC. Ainda foi exposto que qualquer técnica que vise melhorar o desempenho da inversão de canal deve procurar reduzir os efeitos do maior autovalor de \mathbf{H}^{-1} sobre o vetor pré-codificado.

A idéia da técnica *vector perturbation* (VP) é "perturbar" o vetor de símbolos \mathbf{u} por um outro vetor de perturbação \mathbf{p} de tal forma que $\mathbf{u} + \mathbf{p}$ esteja na mesma direção do autovetor associado ao menor autovalor de $(\mathbf{H}\mathbf{H}^H)^{-1}$ [16, 85]. É imposta a restrição de que o receptor deve ser capaz de decodificar $\mathbf{u} + \mathbf{p}$ sem o conhecimento da componente de perturbação \mathbf{p} que é introduzida. Para isso, será utilizado novamente o operador módulo nos receptores, e a perturbação \mathbf{p} será escolhida de tal forma que ela satisfaça:

$$\begin{aligned}\mathbf{p} &= \tau (\mathbf{a} + j\mathbf{b}) \\ &= \tau \mathcal{L},\end{aligned}\tag{3.55}$$

em que $\mathcal{L} = \mathbf{a} + j\mathbf{b}$, \mathbf{a} , \mathbf{b} são vetores de inteiros (isto é, seus elementos pertencem ao conjunto dos números inteiros) e τ é uma constante real, já definida pela Eq. (3.32), quando a pré-codificação Tomlinson-Harashima foi considerada.

Utilizando a matriz inversa do canal como geradora de *beamformers*, a determinação de \mathbf{p} é baseada na solução do seguinte problema de otimização [16]:

$$\begin{aligned}\mathcal{L} &= \arg \min_{\mathcal{L}'} \|\mathbf{W}_{ZF} (\mathbf{u} + \tau \mathcal{L}')\|^2 \\ &= \arg \min_{\mathcal{L}'} (\mathbf{u} + \tau \mathcal{L}')^H (\mathbf{H}\mathbf{H}^H)^{-1} (\mathbf{u} + \tau \mathcal{L}'),\end{aligned}\tag{3.56}$$

que consiste em minimizar a potência do vetor pré-codificado (não normalizado), dado pela expressão $\mathbf{W}_{ZF}(\mathbf{u} + \mathbf{p})$, sob a restrição de eliminar completamente a interferência interusuário, motivo pelo qual é utilizada a matriz de pré-codificação \mathbf{W}_{ZF} .

A Eq. (3.56) é um problema de mínimos quadrados em um reticulado inteiro (do inglês, *integer-lattice least-squares*) [86, 87, 88], e existem diversos algoritmos propostos para resolvê-lo, como, por exemplo, o método de Fincke e Pohst [89], já amplamente utilizado nos receptores do tipo *sphere decoder* para detecção espaço-temporal em sistemas MIMO monousuário [90]. Não serão apresentados os detalhes de operação e implementação dos algoritmos, pois fogem aos objetivos da dissertação. Basta citar que eles evitam a busca exaustiva por todo o reticulado definido por $\tau \mathbf{H}^{-1}$.

Uma análise detalhada do pré-codificador *vector perturbation* é complexa e trabalhosa, especialmente pelo fato de ser difícil obter a distribuição de probabilidade de γ quando é utilizada a perturbação definida pela Eq. (3.55). Apesar disso, em [16], é fornecida uma análise menos rigorosa e mais intuitiva do seu funcionamento, e será apresentada nos próximos parágrafos.

O processo de pré-codificação procura alinhar $\mathbf{x} = \mathbf{u} + \mathbf{p}$ aos vetores singulares da matriz inversa do canal. Tomando $(\mathbf{H}\mathbf{H}^H)^{-1} = \mathbf{Q}\mathbf{\Lambda}^{-1}\mathbf{Q}^H$, pode-se expressar a função objetivo da Eq. (3.56) como

$$\gamma = \|\mathbf{W}_{ZF}(\mathbf{u} + \mathbf{p})\|^2 = \sum_{k=1}^K \frac{1}{\lambda_k} \delta_k^2, \quad (3.57)$$

com $\delta_k = |\mathbf{q}_k^H(\mathbf{u} + \mathbf{p})|$, \mathbf{q}_k é a k -ésima coluna da matriz \mathbf{Q} . Supõe-se ainda que $\lambda_1 < \lambda_2 < \dots < \lambda_K$. A técnica de *vector perturbation* procura minimizar a Eq. (3.57) buscando o vetor de inteiros \mathcal{L} , conforme expresso pela Eq. (3.55), que resulte no menor módulo possível de \mathbf{x} . De fato, a escolha $\mathbf{p} = 0$ conduz ao menor valor para o módulo de $\mathbf{u} + \mathbf{p}$. Ao escolher \mathcal{L} não nulo, certamente aumenta-se a norma do vetor perturbado, mas a norma de $\mathbf{H}^{-1}(\mathbf{u} + \mathbf{p})$, o vetor pré-codificado, é reduzida no processo.

A Eq. (3.57) é minimizada quando $\delta_1 = \dots = \delta_{k-1} = 0$ e $\delta_k = \|\mathbf{u} + \mathbf{p}\|^2$. Basta observar a forma com que os autovalores de $(\mathbf{H}\mathbf{H}^H)^{-1}$ foram arranjados. Entretanto, o fato de forçar o vetor \mathcal{L} a possuir elementos inteiros nem sempre permitirá que $\mathbf{u} + \mathbf{p}$ seja perfeitamente ortogonal às bases $\mathbf{q}_1, \dots, \mathbf{q}_{K-1}$ e paralelo a \mathbf{q}_K , mas apenas grosseiramente orientado nessas direções.

É mostrado em [85] que $\mathbb{E}[\gamma]$ é limitada inferiormente por

$$\mathbb{E}[\gamma] \geq \frac{ev}{K} \mathbb{E}[\|\mathbf{u} + \mathbf{p}\|^2], \quad (3.58)$$

com

$$v = \mathbb{E} \left[\left(\prod_{k=1}^K \frac{\delta_k^2}{c^2} \right)^{1/K} \right], \quad (3.59)$$

e

$$c = (1/K) \mathbb{E}[\|\mathbf{u} + \mathbf{p}\|^2]. \quad (3.60)$$

A igualdade na Eq. (3.58) é alcançada quando

$$\mathbb{E} \left[\frac{\delta_1^2}{\lambda_1} \right] = \dots = \mathbb{E} \left[\frac{\delta_K^2}{\lambda_K} \right] = \frac{evc^2}{K}. \quad (3.61)$$

Logo, para minimizar $\mathbb{E}[\gamma]$, é necessário orientar $\mathbf{u} + \mathbf{p}$ na direção de cada autovetor de forma inversamente proporcional ao valor dos autovalor associado ao respectivo autovetor.

A técnica de *vector perturbation* também possui interpretação geométrica interessante [91], análoga à pré-codificação Tomlinson-Harashima. Quando o transmissor possui conhecimento sobre a resposta do canal, ele pode eliminar completamente a interferência multiusuário, transmitindo o vetor $\mathbf{H}^{-1}\mathbf{u}$ em vez de \mathbf{u} , de forma que o receptor poderá detectar seu símbolo sem a influência de nenhuma interferência. Entretanto, a pré-multiplicação por \mathbf{H}^{-1} aumenta a potência do vetor pré-codificado, conforme visto para o caso *zero-forcing*. Como consequência, a constelação utilizada para transmitir os símbolos é "alongada", conforme mostra a Fig. 3.21 (a).

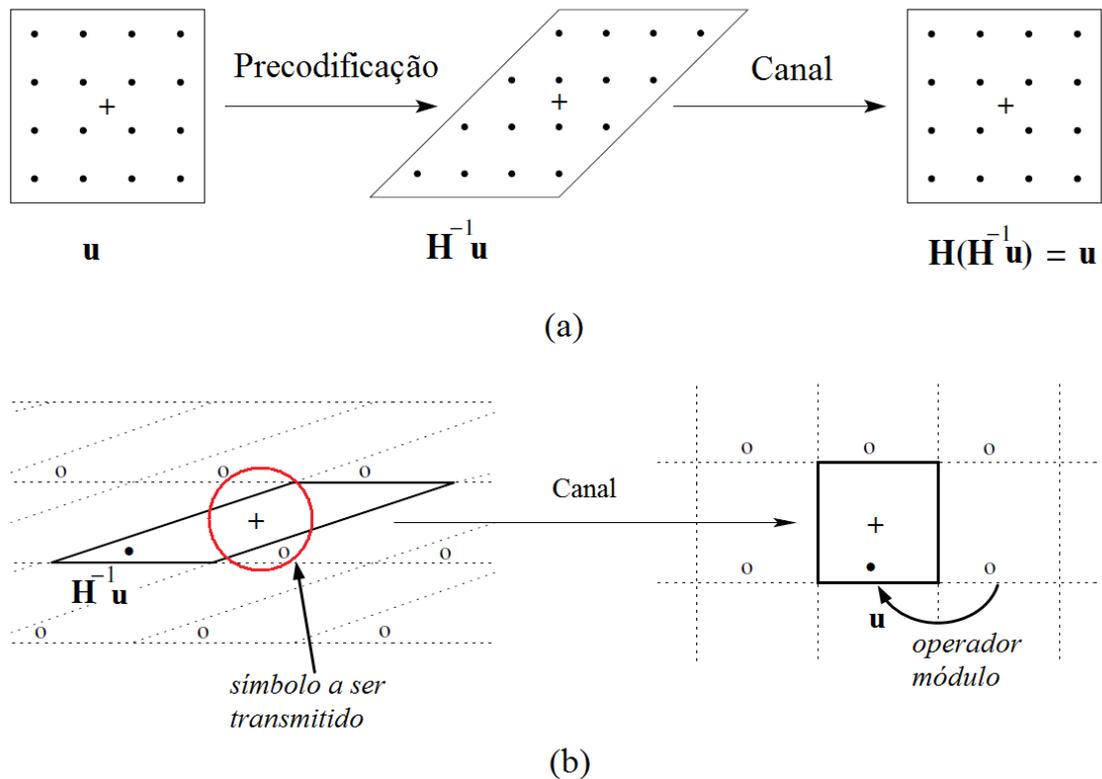


Figura 3.21: Interpretação geométrica da técnica de pré-codificação *vector perturbation*.

Pode-se então utilizar o conceito de extensão periódica da constelação, introduzido na pré-codificação Tomlinson-Harashima, para determinar o conjunto de símbolos módulo-congruentes aos vetores do reticulado formado por $\tau \mathbf{H}^{-1}$, conforme mostra a Fig. 3.21 (b), na qual o sinal mensagem original é representado por "•". Qualquer um dos pontos módulo-congruente à mensagem original (marcados por "o") pode ser utilizado para transmiti-la, pois eles formam um reticulado junto à constelação original. Dentre todos os (infinitos) pontos que podem ser escolhidos, será tomado aquele que se encontra mais próximo à origem, pois esta escolha acarreta na menor potência possível de transmissão. Graças à presença do operador de módulo, o receptor, ao receber qualquer um dos "o", é capaz de mapeá-lo em "•", o ponto dentro da constelação original. Cabe ao algoritmo de busca sobre o reticulado inteiro encontrar em qual constelação deslocada (ou seja, determinar o vetor \mathcal{L}) está o símbolo módulo-congruente mais próximo da origem.

O desempenho do pré-codificador baseado em *vector perturbation* é mostrado na Fig. 3.22 para a modulação QPSK, e na Fig. 3.23 para a modulação 16QAM, considerando os cenários com $K = 4$ usuários e $K = 10$ usuários. A melhora no desempenho é evidente com o aumento do número de usuários no sistema, e sua superioridade com relação às técnicas anteriormente expostas é clara.

Apesar de ser extremamente poderosa, a pré-codificação por *vector perturbation* é também computacionalmente complexa e muito onerosa [92]. Sua complexidade cresce exponencialmente com a dimensão K do problema. Existem várias formas de reduzir o custo computacional da busca em um reticulado inteiro, como, por exemplo, técnicas que possuem como base a decomposição QR ou V-BLAST (do inglês, *Vertical-Bell Laboratories Layered Space-Time*) [17], ou valendo-se de algoritmos de redução de reticulados (*lattice-reduction algorithms*) [93]. Por serem temas consideravelmente amplos, que fogem ao

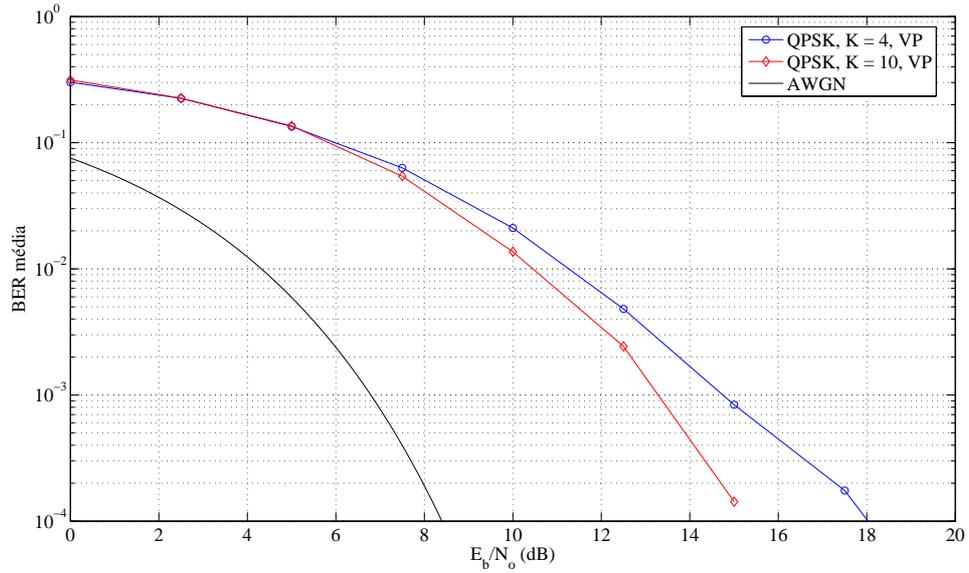


Figura 3.22: Comparação das curvas de BER dos esquemas de pré-codificação para $K = 4$ usuários e $K = 10$ usuários utilizando modulação QPSK.

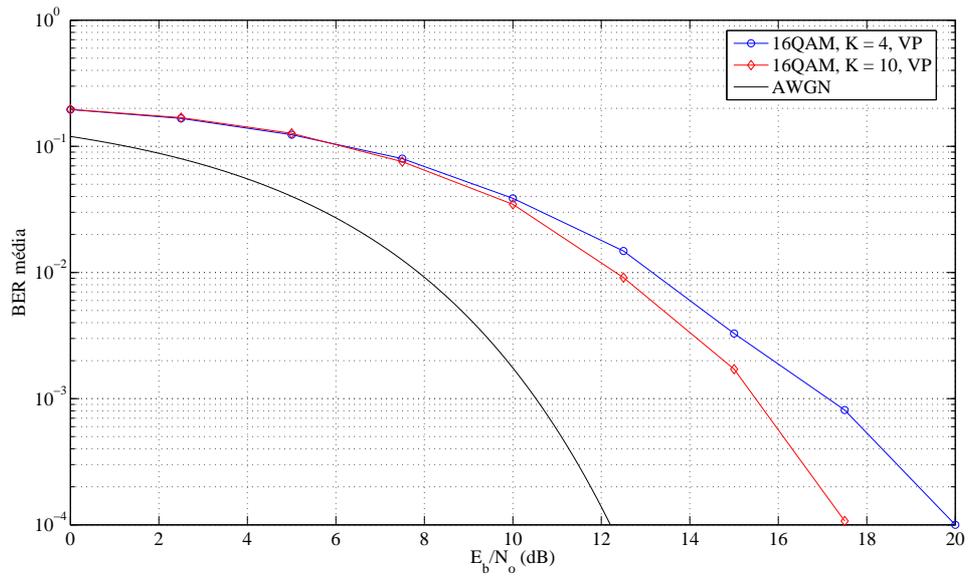


Figura 3.23: Comparação das curvas de BER dos esquemas de pré-codificação para $K = 4$ usuários e $K = 10$ usuários utilizando modulação 16QAM.

objetivo desta dissertação, eles serão discutidos brevemente.

Uma das formas de determinar o vetor de perturbação de forma sub-ótima é aplicar sucessivas vezes o operador módulo em um esquema similar ao oferecido pela pré-codificação de Tomlinson-Harashima. O método utiliza a decomposição QR da matriz de canal \mathbf{H} . Seja $\mathbf{H} = \mathbf{Q}\mathbf{R}$, em que \mathbf{R} é uma matriz do tipo triangular inferior, \mathbf{Q} é uma matriz unitária e $\alpha = 1/(1 + \sigma_n^2)$ [16] é um termo proporcional à relação sinal-ruído. O vetor $\tilde{\mathbf{x}} = [\tilde{x}_1, \dots, \tilde{x}_K]$ é gerado de forma sucessiva, conforme apresentado em [16]:

$$\begin{aligned}
\tilde{x}_1 &= u_1 \\
\tilde{x}_2 &= \mathcal{M}_\tau \left(u_2 - \alpha \frac{r_{21}}{r_{22}} \tilde{x}_1 \right) \\
&\vdots \\
\tilde{x}_K &= \mathcal{M}_\tau \left(u_K - \alpha \sum_{i=1}^{K-1} \frac{r_{Ki}}{r_{KK}} \tilde{x}_i \right),
\end{aligned} \tag{3.62}$$

que pode ser reescrita em termos do vetor de inteiros \mathcal{L} que é adicionado ao sinal:

$$\tilde{\mathbf{x}} = ((1 - \alpha)\mathbf{I} + \alpha\mathbf{R})^{-1} (\mathbf{u} + \tau\mathcal{L}). \tag{3.63}$$

A partir da Eq. (3.62), é formado o vetor $\mathbf{x} = \mathbf{Q}^H \tilde{\mathbf{x}}$, que é normalizado e transmitido. O parâmetro α aumenta a SINR para cada usuário, de forma similar ao parâmetro ς utilizado na pré-codificação por regularização de inversa [16].

Algoritmos baseados na decomposição QR não são capazes de explorar a máxima ordem de diversidade (a inclinação da curva de BER será diferente da técnica original), apesar de possuírem como vantagem o reduzido custo computacional, além de serem amplamente conhecidos e exaustivamente analisados na literatura [85].

Ao contrário da técnica de otimização sucessiva apresentada, as técnicas de redução de reticulado são capazes de atingir a mesma diversidade da técnica *vector perturbation*, mas apresentam perda de relação sinal-ruído. As técnicas de redução de reticulado mais conhecidas são o algoritmo de Lenstra-Lenstra-Lovász (LLL) [94] e o algoritmo de Brun's [95]. Estes algoritmos são mais eficientes pois, em vez de utilizarem como base do reticulado os vetores \mathbf{H}^{-1} , eles buscam um conjunto de vetores de base $(\mathbf{H}^{-1})'$ que possuam melhores características de ortogonalidade (este conceito será exposto nos próximos parágrafos), obtendo uma base que permite uma busca aproximada e mais eficiente sobre o reticulado [96].

A decomposição LLL consiste em, a partir da base original, $\mathbf{W}_{ZF} = \mathbf{H}^{-1}$, obter uma base reduzida \mathbf{W}' de forma tal que $\mathbf{W}' = \mathbf{W}_{ZF}\mathbf{T}$, em que \mathbf{T} é uma matriz unimodular, isto é, seus elementos são números complexos que possuem elementos inteiros em suas partes real e imaginária, e seu determinante é unitário.

O defeito de ortogonalidade (*orthogonality defect*) da base \mathbf{W}' é definido como [95]

$$\delta(\mathbf{W}') = \frac{1}{|\mathbf{W}'|} \prod_{k=1}^K \|\mathbf{w}'_k\|^2, \tag{3.64}$$

em que \mathbf{w}'_k é a k -ésima coluna de \mathbf{W}' . A Eq. (3.64) provê uma maneira de mensurar a ortogonalidade dos vetores que formam a base de um reticulado por meio da razão entre o produto das normas dos vetores da base do reticulado e o volume da célula primitiva do reticulado que eles definem. É possível mostrar que $\delta(\mathbf{W}') \geq 1$, sendo a igualdade alcançada quando os vetores da base são ortogonais [97]. O objetivo da decomposição LLL é encontrar uma base \mathbf{W}' de tal forma que $\delta(\mathbf{W}') < \delta(\mathbf{W})$. Utilizando um abuso de linguagem, pode-se dizer que os vetores da nova base são "mais ortogonais" entre si que os elementos da base original, pois apresentam menor defeito de ortogonalidade.

Dessa forma, a função de custo dada pela Eq. (3.56) pode ser reescrita como

$$\begin{aligned}\left\| \mathbf{W}_{ZF} (\mathbf{u} + \tau \mathcal{L}') \right\|^2 &= \left\| \mathbf{W}' \mathbf{T} \mathbf{T}^{-1} (\mathbf{u} + \tau \mathcal{L}') \right\|^2 \\ &= \left\| \tilde{\mathbf{W}} (\mathbf{T}^{-1} \mathbf{u} + \tau \mathbf{T}^{-1} \mathcal{L}') \right\|^2 \\ &= \left\| \tilde{\mathbf{W}} (\tilde{\mathbf{u}} + \tau \tilde{\mathcal{L}}') \right\|^2,\end{aligned}\tag{3.65}$$

em que foi definido

$$\tilde{\mathbf{u}} = \mathbf{T}^{-1} \mathbf{u}\tag{3.66}$$

e

$$\tilde{\mathcal{L}}' = \mathbf{T}^{-1} \mathcal{L}'.\tag{3.67}$$

Como a matriz \mathbf{T} é unimodular, a solução do problema dado pela Eq. (3.56) é equivalente a resolver primeiro o problema

$$\mathcal{L} = \arg \min_{\mathcal{L}'} \left\| \tilde{\mathbf{W}} (\tilde{\mathbf{u}} + \tau \tilde{\mathcal{L}}') \right\|^2,\tag{3.68}$$

e depois calcular

$$\mathbf{p} = \tau \mathbf{T} \mathcal{L}.\tag{3.69}$$

Por exemplo, a chamada estimativa de Babai [96] é obtida ao fazer $\mathbf{W}' = \mathbf{I}$, a matriz identidade, e a solução fornecida pelas Eq. (3.66) e Eq. (3.67) será dada pela expressão

$$\mathbf{p} = \tau \mathbf{T} \left\lfloor \frac{\mathbf{T}^{-1} \mathbf{u}}{\tau} \right\rfloor.\tag{3.70}$$

Uma vantagem do algoritmo LLL é que ele apresenta tempo de execução polinomial em relação à dimensão K do problema.

Da mesma forma que a técnica *vector perturbation*, a redução de reticulado também possui interpretação geométrica. Um exemplo simples de reticulado bidimensional está mostrado na Fig. 3.24. O mesmo reticulado é representado por duas bases diferentes, na Fig. 3.24 (a) e na Fig. 3.24 (b), e essas bases são mostradas por meio de setas vermelhas localizadas no canto superior esquerdo de cada figura. As regiões ótimas de decisão (denominadas células de Voronoi, e mostradas na cor laranja) são hexágonos, e correspondem à solução do problema de busca em um reticulado inteiro utilizando os algoritmos que fornecem a solução exata para o problema.

As regiões de decisão fornecidas pela Eq. (3.69), para cada uma das bases reduzidas mostradas na figura, são representadas por paralelogramos, e estão mostradas na cor verde no canto superior direito de cada figura. A comparação entre ambas as regiões (da solução exata e da solução aproximada) é fornecida

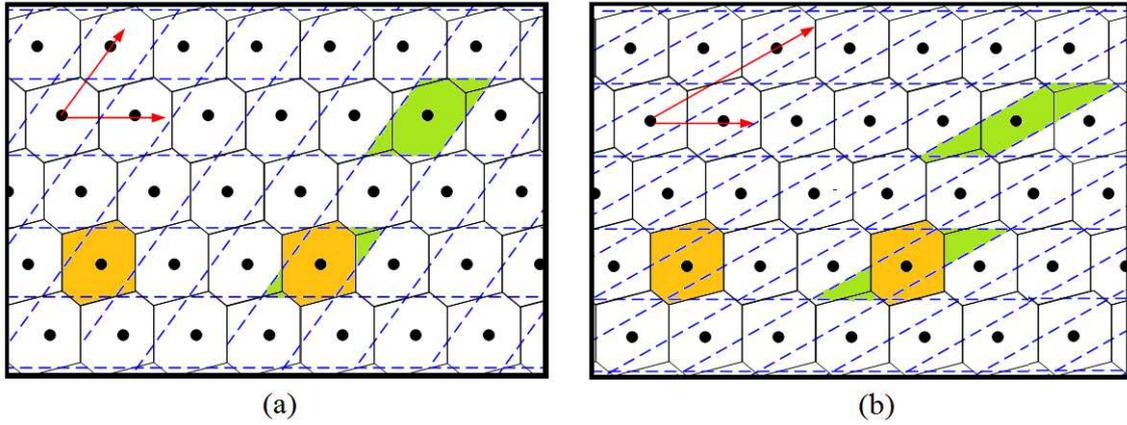


Figura 3.24: Interpretação geométrica da redução de reticulado utilizando duas bases diferentes. A primeira base, mais ortogonal, é mostrada em (a), e outro exemplo de base é mostrado em (b).

no canto inferior direito inferior de cada diagrama.

A partir do exemplo, duas observações importantes podem ser feitas: em primeiro lugar, a solução aproximada geralmente não alcançará a solução ótima, independentemente da escolha da base. Em segundo lugar, uma boa escolha de base pode melhorar a precisão da aproximação fornecida pela redução de reticulado. A Fig. 3.24 (a) consiste em uma boa escolha para a base reduzida, pois as regiões de decisão ótima e aproximada apresentam larga sobreposição. Por outro lado, a escolha da base representada na Fig. 3.24 (b) é uma escolha ruim, pois as regiões de decisão diferem de forma significativa. Nesta situação, a base representada na Fig. 3.24 (a) apresenta menor defeito de ortogonalidade do que a base mostrada na Fig. 3.24 (b).

3.8 MMSE VECTOR PERTURBATION

É sabido que a solução MMSE representa uma relação de compromisso entre o nível de interferência residual que será gerado pelo pré-codificador e a atenuação do vetor transmitido [58]. Sob esta premissa, foi apresentada a técnica de pré-codificação por regularização de inversa.

Como já era esperado, é possível utilizar, em vez da matriz inversa do canal, a matriz inversa regularizada para realizar a pré-codificação *vector perturbation*. Este esquema foi considerado em [98] e é denominado MMSE *vector perturbation* (MMSE VP). Neste caso, a matriz de pré-codificação é dada por

$$\mathbf{W}_{MMSE} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \frac{\sigma_n^2}{\sigma_u^2} \mathbf{I} \right)^{-1}, \quad (3.71)$$

em que σ_u^2 representa a potência média dos símbolos da constelação utilizada, e pode ser determinado por meio de sua matriz de autocovariância, por hipótese dada por $\mathcal{R}_{\mathbf{u}} = \sigma_u^2 \mathbf{I}$. O fator σ_n^2/σ_u^2 pode ser interpretado como possuindo a mesma função do parâmetro de regularização ζ utilizado na técnica de pré-codificação por regularização de inversa, porém considerando que a constelação da qual os símbolos de informação são originados não precisa apresentar, necessariamente, potência média unitária.

O problema de otimização dado pela Eq.(3.56) será agora escrito como:

$$\begin{aligned}\mathcal{L} &= \arg \min_{\mathcal{L}'} \left\| \mathbf{W}_{MMSE} \left(\mathbf{u} + \tau \mathcal{L}' \right) \right\|^2 \\ &= \arg \min_{\mathcal{L}'} \left(\mathbf{u} + \tau \mathcal{L}' \right)^H \left(\mathbf{H}\mathbf{H}^H + \frac{\sigma_n^2}{\sigma_u^2} \mathbf{I} \right)^{-1} \left(\mathbf{u} + \tau \mathcal{L}' \right).\end{aligned}\quad (3.72)$$

Utilizando a fatoração de Cholesky, pode-se determinar uma matriz \mathbf{L} que satisfaça

$$\left(\mathbf{H}\mathbf{H}^H + \frac{\sigma_n^2}{\sigma_u^2} \mathbf{I} \right) = \mathbf{L}^H \mathbf{L}, \quad (3.73)$$

e a Eq. (3.72) pode ser reescrita como:

$$\begin{aligned}\mathcal{L} &= \arg \min_{\mathcal{L}'} \left(\mathbf{u} + \tau \mathcal{L}' \right)^H \mathbf{L}^H \mathbf{L} \left(\mathbf{u} + \tau \mathcal{L}' \right) \\ &= \arg \min_{\mathcal{L}'} \left\| \mathbf{L} \left(\mathbf{u} + \tau \mathcal{L}' \right) \right\|^2\end{aligned}\quad (3.74)$$

Ao observar a forma da Eq. (3.74), o problema se reduz a uma busca em um reticulado inteiro. Nesta situação específica, o reticulado é gerado por $\tau \mathbf{L}$, e procura-se pelo vetor de inteiros que corresponde ao ponto do reticulado mais próximo de $-\mathbf{L}\mathbf{u}$. Em geral, este não será o ponto que resulta no menor fator de normalização γ (devido à presença do parâmetro de regularização), mas ainda assim é observada uma melhora no desempenho com relação a técnica *vector perturbation*.

As Figuras 3.25 a 3.28 mostram, de forma comparativa, o desempenho de todos os pré-codificadores considerados, inclusive o MMSE VP. Em primeiro lugar, observa-se que ambos os pré-codificadores baseados em *vector perturbation* apresentam melhor desempenho que os pré-codificadores lineares para regimes de relação sinal-ruído mais elevada, e o desempenho do MMSE VP é ligeiramente superior ao desempenho do VP. Entretanto, o desempenho do *vector perturbation*, para valores de relação sinal-ruído mais baixos, é inferior ao das técnicas lineares, especialmente para modulações de ordem mais alta (16QAM) e número de usuários maior. A mesma observação já não é válida para os pré-codificadores baseados na técnica Tomlinson-Harashima, cujo desempenho é superior para valores mais baixos de E_b/N_o . Apesar disso, à medida em que a relação sinal-ruído aumenta, tanto VP com MMSE VP apresentam ordem de diversidade maior do que a apresentada pelos pré-codificadores ZF THP e MMSE THP.

Como comentário final, a pré-codificação utilizando o operador módulo é a forma mais simples de implementação do DPC, envolvendo um simples reticulado cúbico [70]. Melhoras significativas no desempenho podem ser obtidas ao serem considerados reticulados mais complexos, de dimensão maior [35, 99]. Particularmente, estas técnicas baseadas em reticulados maiores podem melhorar o desempenho das técnicas não-lineares na região de relação sinal-ruído mais baixa, onde as técnicas lineares apresentam desempenho ligeiramente superior do que as técnicas não lineares (em especial, VP e MMSE VP).

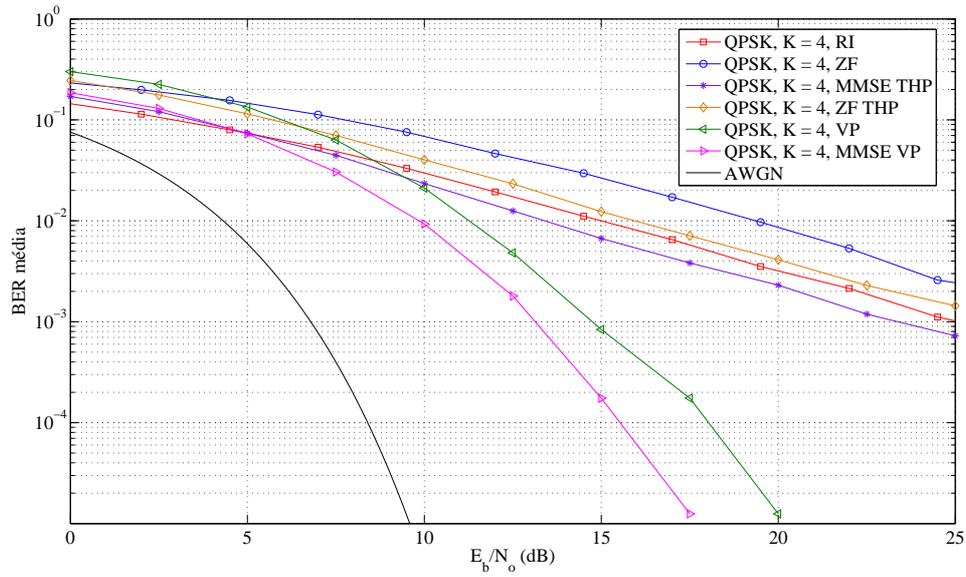


Figura 3.25: Comparação das curvas de BER dos esquemas de pré-codificação com $K = 4$ usuários utilizando modulação QPSK.

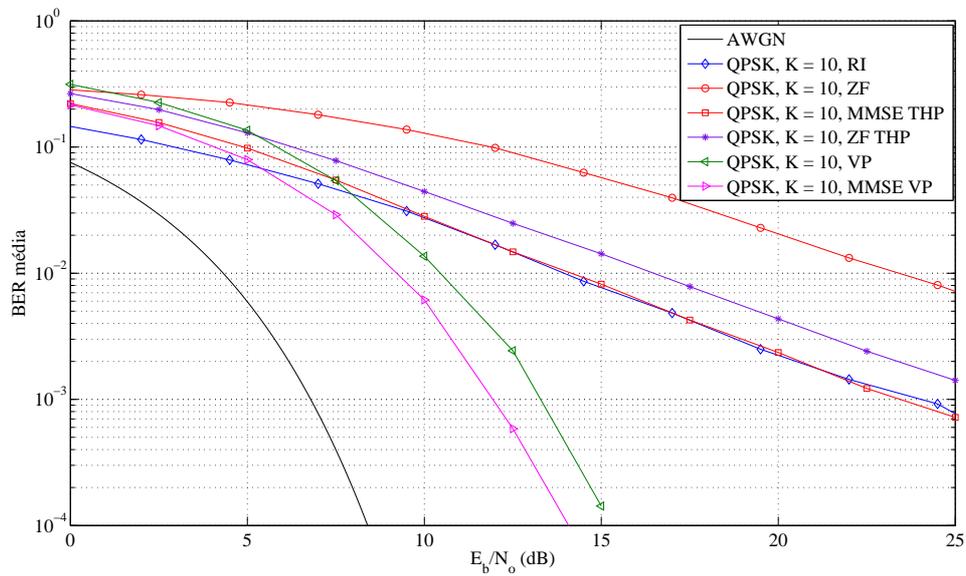


Figura 3.26: Comparação das curvas de BER dos esquemas de pré-codificação com $K = 10$ usuários utilizando modulação QPSK.

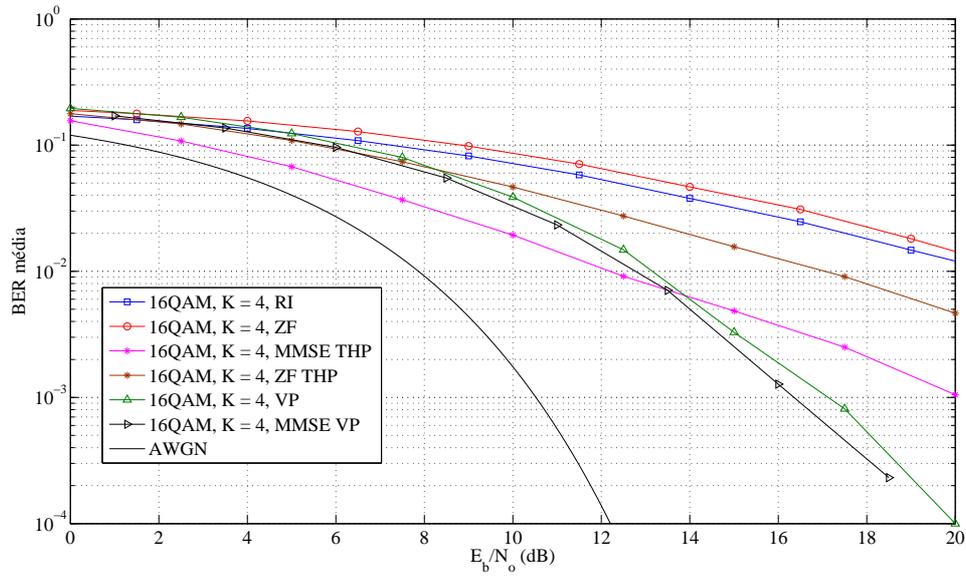


Figura 3.27: Comparação das curvas de BER dos esquemas de pré-codificação com $K = 4$ usuários utilizando modulação 16QAM.

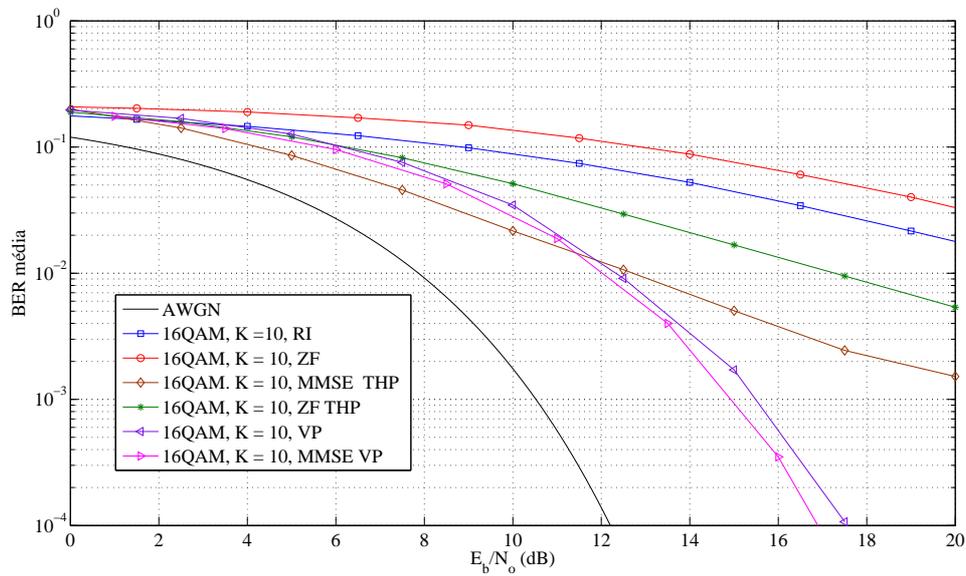


Figura 3.28: Comparação das curvas de BER dos esquemas de pré-codificação com $K = 10$ usuários utilizando modulação 16QAM.

3.9 PROPOSTA DE PRÉ-CODIFICADOR

Perturbar o vetor de dados (no caso da pré-codificação *vector perturbation*) ou a inversa da matriz combinada de canal (para o caso de regularização de inversa) é uma forma de controlar o valor que a constante de normalização γ poderá assumir, de forma a reduzir a perda de relação sinal-ruído no receptor.

Nesta seção, é proposto um algoritmo de pré-codificação baseado no conceito de perturbação contínua, apresentado em [100]. Diferentemente da técnica *vector perturbation*, na qual o vetor de perturbação é escolhido a partir de um reticulado inteiro e pode ser removido por meio do operador módulo no receptor, a perturbação contínua não pode ser completamente eliminada pelos receptores, e deve ser tratada como interferência. Naturalmente, uma métrica utilizada para se determinar o vetor de perturbação será a relação sinal-ruído mais interferência (SINR).

A idéia é formalizada do seguinte modo. Em primeiro lugar, o vetor de informação é perturbado por um vetor \mathbf{p}' , denominado vetor de perturbação contínua. Os elementos desse vetor podem assumir qualquer valor real, ao passo que, na pré-codificação por *vector perturbation*, os elementos do vetor de perturbação \mathbf{p} devem ser múltiplos de τ .

Diferentemente do esquema proposto em [101], no qual a matriz de pré-codificação é a pseudo-inversa da matriz combinada de canal, a matriz de pré-codificação do método proposto é \mathbf{W}_{MMSE} , a inversa regularizada, definida pela Eq. (3.71).

Dessa forma, o vetor de dados recebidos \mathbf{y} relaciona-se ao vetor de símbolos de informação \mathbf{u} por meio de

$$\mathbf{y} = \frac{1}{\sqrt{\gamma}} \mathbf{H} \mathbf{W}_{MMSE} (\mathbf{u} + \mathbf{p}') + \mathbf{n}. \quad (3.75)$$

Utilizando a autodecomposição da matriz $\mathbf{H} \mathbf{H}^H$, conforme apresentado na seção 3.4, a relação sinal-ruído mais interferência pode ser reescrita como

$$SINR = \frac{\left(\sum_{k=1}^K \frac{\lambda_k}{\lambda_k + \zeta} \right)^2}{\| \mathbf{H} \mathbf{W}_{MMSE} \mathbf{p}' \|^2 + \gamma K \sigma_n^2 + \sum_{k=1}^K w'_k}, \quad (3.76)$$

em que w'_k se refere à interferência residual causada pela regularização de inversa. É possível mostrar que o máximo da Eq. (3.76) ocorre quando o termo

$$\| \mathbf{H} \mathbf{W}_{MMSE} \mathbf{p}' \|^2 + \gamma K \sigma_n^2 \quad (3.77)$$

é mínimo. Entretanto, como \mathbf{p}' é contínuo (pode assumir qualquer valor real), não é necessária a utilização de algoritmos de complexidade elevada, como o *sphere decoder*, para determinar o valor ótimo de \mathbf{p}' (no sentido de maximizar a Eq. (3.76)), ao contrário do que ocorre com a técnica *vector perturbation*, na qual a perturbação \mathbf{p} assume valores discretos. A otimização da Eq. (3.77) pode ser feita de forma analítica. Tomando sua derivada em relação a \mathbf{p}' , e após algumas manipulações, o vetor de perturbação contínua será

dado por:

$$\mathbf{p}' = -K\sigma_n^2 (\mathbf{W}_{MMSE}^H \mathbf{H}^H \mathbf{W}_{MMSE} \mathbf{H} + K\sigma_n^2 \mathbf{W}_{MMSE}^H \mathbf{W}_{MMSE})^{-1} \mathbf{W}_{MMSE}^H \mathbf{W}_{MMSE} \mathbf{u} \quad (3.78)$$

A Fig. 3.29 mostra o desempenho do pré-codificador proposto quando comparado às técnicas lineares de *zero-forcing* e regularização de inversa para a modulação QPSK, e a Fig. 3.30 mostra o desempenho do pré-codificador para a modulação 16QAM. Apesar de ser incapaz de ter desempenho comparável aos esquemas de pré-codificação não lineares, pois não é baseado no conceito de codificação *dirty paper*, ele apresenta ganho significativo quando comparado aos outros algoritmos lineares. É possível observar um ganho de 1,5 dB quando comparado à pré-codificação por regularização de inversa sob modulação 16QAM, e de aproximadamente 1 dB sobre a regularização de inversa sob modulação QPSK, sendo estes ganhos aparentemente independentes do número K de usuários.

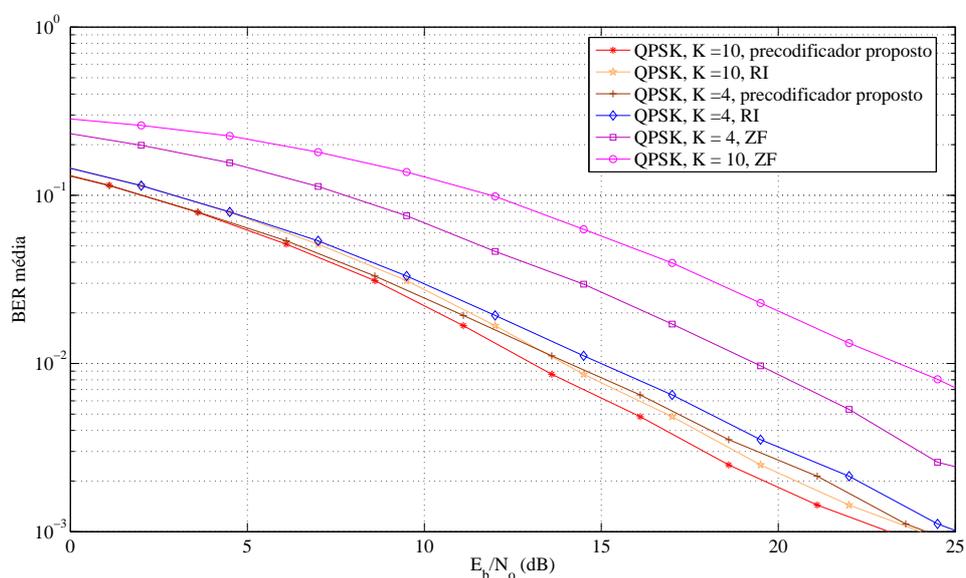


Figura 3.29: Desempenho do pré-codificador proposto comparado aos algoritmos lineares de pré-codificação para a modulação QPSK.

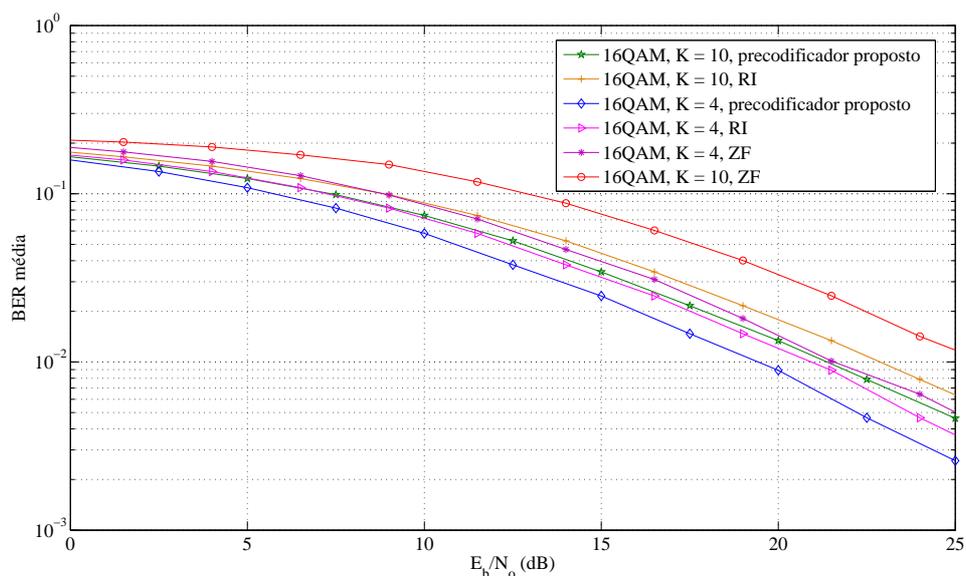


Figura 3.30: Desempenho do pré-codificador proposto comparado aos algoritmos lineares de pré-codificação para a modulação 16QAM.

3.10 INFLUÊNCIA DA INFORMAÇÃO IMPERFEITA DE ESTADO DE CANAL SOBRE A PRÉ-CODIFICAÇÃO

Até este momento, foi suposto que o transmissor possui conhecimento perfeito das informações sobre o estado de canal, ou CSI (do inglês, *channel state information*) de cada usuário, e dessa forma ele pode determinar as matrizes de pré-codificação. Entretanto, em termos práticos, esta hipótese dificilmente será verdadeira.

O transmissor pode obter CSI de duas formas diferentes. Em sistemas de comunicações baseados em FDD, o receptor utiliza um canal de retorno (*feedback*) para enviar ao transmissor as informações de estado do canal. Em sistemas TDD, o transmissor pode, a princípio, obter essas informações diretamente, por meio de sequências de treinamento do enlace reverso, desde que o canal seja recíproco¹. Cada esquema possui vantagens e desvantagens em termos de latência, custos computacionais e perda de capacidade [102].

Independente da forma de duplexação utilizada, devido ao atraso introduzido pelo canal de retorno, ou à natureza variante do canal no tempo, ou até mesmo por atrasos que são gerados pelos próprios algoritmos de processamento (como estimação e decodificação de canal e pré-codificação), as informações de estado de canal podem estar desatualizadas, e esta desfasagem pode degradar severamente o desempenho dos sistemas considerados.

Como exemplo, seja o resultado de simulação mostrado na Fig. 3.31. É mostrado o desempenho

¹Neste ponto, uma observação é pertinente: mesmo que os *slots* de tempo alocados para a transmissão nos enlaces reverso e direto possuam duração comparável ao tempo de coerência do canal, os canais de ambos os enlaces pode ainda assim não ser recíproco, pois o comportamento da interferência nos dois enlaces é diferente, bem como as cadeias de radio frequência (RF) dos transmissores e receptores, o comportamento do ruído, etc. Estes fatos contribuem para que, mesmo operando em TDD e desvanecimento lento, os canais dos enlaces direto e reverso não sejam recíprocos.

da modulação 16QAM utilizando os pré-codificadores *vector perturbation* e regularização de inversa sob condições em que o transmissor não dispõe de informação perfeita sobre \mathbf{H} , mas tem apenas acesso à variável $\mathbf{H} + \delta\mathbf{H}$, em que $\delta\mathbf{H}$ é uma variável aleatória gaussiana que modela os efeitos da CSI imperfeita. O modelo considerado é bastante simples e será substituído por considerações mais realistas nos próximos capítulos. O objetivo principal agora é apenas mostrar a influência da CSI imperfeita no processo de pré-codificação.

Como pode ser percebido, para que os ganhos oferecidos pelas técnicas de pré-codificação no ambiente MIMO BC possam ser explorados plenamente, a resposta do canal tem de ser estimada com antecedência (no caso de atrasos introduzidos pelo sistema ou ao comportamento variante no tempo da resposta do canal), e filtros preditores podem ser utilizados para tal propósito, conforme será abordado no próximo capítulo.

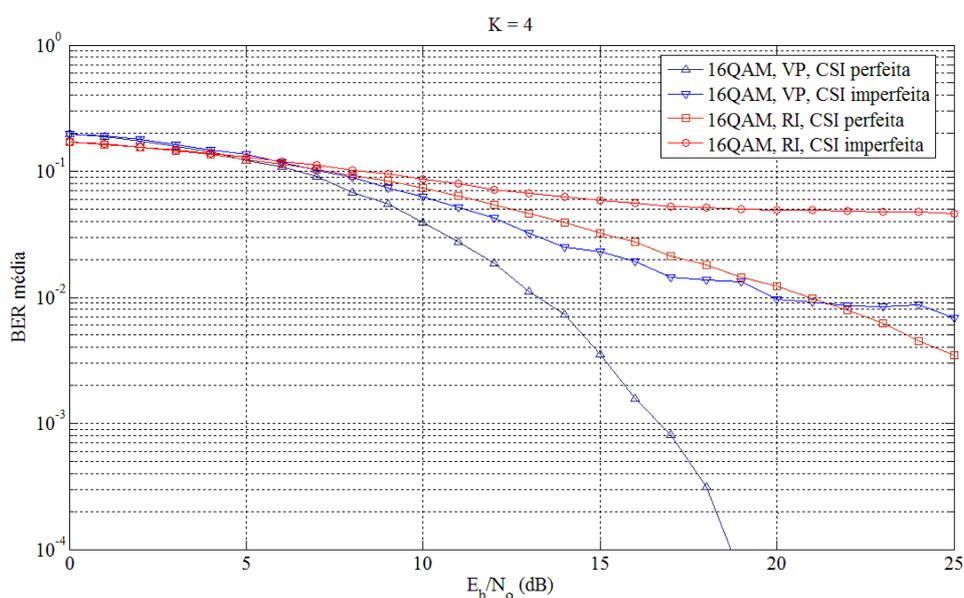


Figura 3.31: Comparação das curvas de BER dos esquemas de pré-codificação *vector perturbation* e regularização de inversa sob CSI imperfeita para $K = 4$ usuários utilizando modulação 16QAM.

3.11 CONCLUSÃO

Este capítulo abordou, sob a ótica de processamento de sinais, os diversos algoritmos de pré-codificação propostos para o canal de *broadcast* de um sistema MIMO multiusuário. As técnicas lineares consideradas foram três: *zero-forcing*, regularização de inversa e *block diagonalization*, esta uma generalização do conceito de inversão de canal. As técnicas não lineares apresentadas foram baseadas no conceito de pré-codificação sobre um reticulado, e para tal utilizavam o operador módulo. Foram analisadas as técnicas de pré-codificação *zero-forcing* Tomlinson-Harashima, MMSE Tomlinson-Harashima, *vector perturbation* e MMSE *vector perturbation*. Para todos os algoritmos, foram apresentadas suas vantagens, desvantagens e uma análise comparativa.

Apresentou-se uma proposta de pré-codificador tendo como base o conceito de perturbação contínua

e pré-codificação utilizando a solução MMSE. Apesar de o seu desempenho não exibir a mesma ordem de diversidade das técnicas não lineares apresentadas, ele garante ganho significativo sobre as técnicas lineares consideradas, que não consideram a introdução de qualquer perturbação.

Finalmente, foi feita uma breve apresentação dos efeitos da informação imperfeita sobre o estado do canal sobre o desempenho de alguns pré-codificadores, e foi sugerido que algoritmos adaptativos podem ser utilizados para prover CSI atualizada para o transmissor, sendo possível reduzir os impactos da informação desatualizada sobre o estado de canal nas técnicas de pré-codificação do canal MIMO BC.

No próximo capítulo será realizado um estudo dos algoritmos clássicos de filtragem adaptativa que podem ser utilizados para realizar a predição da resposta do canal.

4 ALGORITMOS DE FILTRAGEM ADAPTATIVA

4.1 INTRODUÇÃO

Conforme tratado na seção 3.10 do capítulo 3, a informação desatualizada sobre o estado do canal pode degradar severamente o desempenho das técnicas de pré-codificação apresentadas e, para compensar os atrasos introduzidos pelo sistema, faz-se necessário o uso de estratégias de predição de canal.

Apesar de várias abordagens para a predição de canal serem propostas, como, por exemplo, modelos autoregressivos, soma de senóides e técnicas baseadas em subespaços vetoriais [103], optou-se pela utilização de filtros digitais adaptativos pois, conforme exposto em [103], esta abordagem tem se mostrado mais adequada para modelagens mais realistas do canal de comunicação rádio móvel.

Os filtros digitais são os equivalentes dos analógicos no contexto de processamento digital de sinais. A saída $r(n)$ de um filtro digital não recursivo no instante n está relacionada com sua entrada $s(n)$ pela equação de diferenças finitas [104]

$$r(n) = \sum_{i=0}^{N_c-1} c_i s(n-i), \quad (4.1)$$

em que $c_0, c_1, \dots, c_{N_c-1}$ são os N_c coeficientes do filtro, que representam sua resposta impulsional. Como esta resposta possui comprimento finito, o filtro é dito do tipo FIR (do inglês, *finite impulse response*). Uma das vantagens de trabalhar com filtros de resposta finita ao impulso é que eles são naturalmente estáveis [105].

A Eq. (4.1) pode ser implementada por meio da estrutura de forma direta representada na Fig. 4.1, conhecida como *tapped-delay line*.

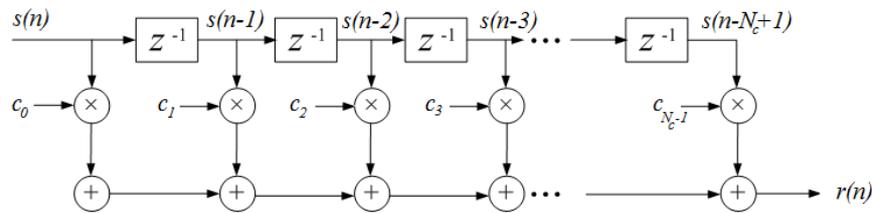


Figura 4.1: Forma direta de implementação de um filtro FIR.

Por questões de brevidade de notação, definiremos os vetores $\mathbf{s}(n) = [s(n), s(n-1), \dots, s(n-N_c+1)]^T$ e $\mathbf{c} = [c_0^*, c_1^*, \dots, c_{N_c-1}^*]^T$, de forma que a convolução discreta definida pela Eq. (4.1) pode ser expressa por meio do produto matricial

$$r(n) = \mathbf{c}^H \mathbf{s}(n). \quad (4.2)$$

É importante perceber que, nos filtros digitais convencionais, os coeficientes do filtro são escolhidos durante a fase de projeto e permanecem constantes durante toda operação do filtro. Entretanto, há determinadas aplicações que não são solucionadas adequadamente com a utilização de filtros de coeficientes fixos. Nestes casos, os denominados filtros adaptativos podem ser utilizados. Este tipo de filtro possui a habilidade de modificar sua resposta de forma automática, por meio de um algoritmo pré-definido, melhorando e ajustando seu desempenho durante a fase de operação.

Neste capítulo, serão apresentados os algoritmos clássicos utilizados em filtragem adaptativa. Serão mostrados os algoritmos clássicos *least mean square*, *normalized least mean square*, *recursive least squares* e *affine projection*. Suas características não serão discutidas de forma exaustiva. Tal tratamento pode ser encontrado em [106, 107]. Em seguida, será apresentado o algoritmo *set-membership affine projection*, e serão discutidas suas vantagens sobre os algoritmos considerados anteriormente. Finalmente será proposta uma estrutura de predição adaptativa de canal aplicada a sistemas MIMO-OFDM multiusuário utilizando o algoritmo *set-membership affine projection*.

4.2 SOLUÇÃO DE WIENER

A estimação de um sinal a partir de um outro é um dos problemas mais importantes em processamento de sinais, e engloba uma gama muito variada de aplicações. Em 1940, Norbert Wiener, em seu trabalho pioneiro [108], formulou o problema de projetar um filtro que produziria a estimativa ótima $r(n)$ de um sinal $s(n)$ a partir de uma observação ou medição ruidosa $d(n)$ [109], conforme mostrado na Fig. 4.2.

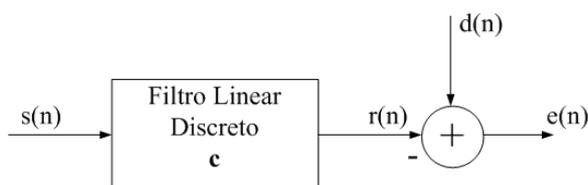


Figura 4.2: Diagrama de blocos do filtro de Wiener. Os sinais $s(n)$ e $d(n)$ são estatisticamente relacionados, e o objetivo do filtro é produzir a estimativa $r(n)$ de $d(n)$ que possui o menor erro quadrático médio.

Dependendo da forma com a qual os sinais $s(n)$ e $d(n)$ estão relacionados, tem-se a definição de diferentes tipos de problemas e aplicações de filtragem. Por exemplo, em aplicações de predição (previsão das amostras futuras de uma sequência através do emprego de informações passadas), $d(n) = s(n + 1)$, e \mathbf{c} representa a resposta de um filtro causal, então o filtro de Wiener torna-se um preditor linear. Neste caso, o filtro produzirá uma predição (estimativa) de $s(n + 1)$ em termos de uma combinação linear dos valores passados de $s(n)$. Em aplicações de equalização (recuperação do conteúdo espectral de uma informação processada por um outro sistema), se $s(n) = d(n) \otimes g(n)$, com $g(n)$ a resposta impulsional de um sistema linear invariante ao deslocamento, e \otimes indica a operação de convolução, o filtro de Wiener se torna um equalizador. Outras aplicações incluem a suavização e a identificação de sistemas [109].

A saída do filtro é dada pela Eq. (4.2). O sinal de erro, $e(n)$, é definido como a diferença entre o sinal desejado $d(n)$ e a saída do filtro:

$$e(n) = d(n) - r(n) = d(n) - \mathbf{c}^H \mathbf{s}(n). \quad (4.3)$$

O problema de projeto consiste em, a partir da entrada $s(n)$ e do sinal desejado $d(n)$, encontrar o conjunto de coeficientes \mathbf{c} que minimiza alguma função objetivo J . Para o caso considerado por Wiener, esta função objetivo é o erro quadrático médio, $J(e(n)) = \mathbb{E} [|e(n)|^2]$.

O quadrado do módulo do erro é dado por

$$|e(n)|^2 = |d(n)|^2 - 2\text{Re} [\mathbf{c}^H(n) \mathbf{s}(n) \mathbf{d}^*(n)] + \mathbf{c}^H(n) \mathbf{s}(n) \mathbf{s}^H(n) \mathbf{c}(n), \quad (4.4)$$

de forma que a função objetivo pode ser expressa como

$$\begin{aligned} J(e(n)) &= \mathbb{E} [|e(n)|^2] \\ &= \mathbb{E} [|d(n)|^2] - 2\text{Re} [\mathbf{c}^H \mathcal{R}_{sd}] + \mathbf{c}^H \mathcal{R}_s \mathbf{c}, \end{aligned} \quad (4.5)$$

para a qual se definiu

$$\mathcal{R}_{sd} = \mathbb{E} [\mathbf{s}(n) \mathbf{d}^*(n)] \quad (4.6)$$

como o vetor de correlação cruzada entre o sinal desejado e o sinal de entrada, e

$$\mathcal{R}_s = \mathbb{E} [\mathbf{s}(n) \mathbf{s}^H(n)] \quad (4.7)$$

é a matriz de autocovariância do sinal de entrada.

A função objetivo definida pela Eq. (4.5) é uma função quadrática nos coeficientes do filtro, e consiste na superfície de um hiperparabolóide de N_c dimensões. Dessa forma, ela possui um mínimo global e nenhum mínimo local. Para determinar o valor mínimo da função de custo, obtém-se primeiramente seu gradiente, $\nabla_{\mathbf{c}^*} J$. Após algumas manipulações, apresentadas em [106], é possível mostrar que

$$\begin{aligned} \nabla_{\mathbf{c}^*} \mathbb{E} [|e(n)|^2] &= \frac{\partial \mathbb{E} [|e(n)|^2]}{\partial \mathbf{c}^*} \\ &= \mathbb{E} [-e^*(n) \mathbf{s}(n)] \\ &= -\mathcal{R}_{sd} + \mathcal{R}_s \mathbf{c}. \end{aligned} \quad (4.8)$$

O vetor de coeficientes ótimos \mathbf{c}_o é aquele para o qual o gradiente da Eq. (4.8) se anula:

$$\mathbf{c}_o = \mathcal{R}_s^{-1} \mathcal{R}_{sd}, \quad (4.9)$$

conhecida como equação de Wiener-Hopf.

Em termos práticos, determinar os valores de coeficientes dados pela Eq. (4.9) pode ser uma tarefa difícil. Primeiramente, porque é necessário o conhecimento das estatísticas (representadas pelas matrizes de autocovariância e correlação cruzada) dos sinais envolvidos. Além disso, o cálculo dos

coeficientes envolve uma inversão matricial, que é computacionalmente dispendioso, especialmente para filtros que possuem um número elevado de coeficientes. Finalmente, a solução de Wiener supõe que os sinais são estacionários no sentido amplo, uma hipótese que nem sempre é válida ou verdadeira. Caso esta condição não seja satisfeita, o filtro de Wiener pode não ser realizável ou não possuir desempenho ótimo.

4.3 FILTROS ADAPTATIVOS

Se os sinais $s(n)$ e $d(n)$ não são estacionários, os coeficientes \mathbf{c} do filtro de Wiener dependerão de n , e a resposta do filtro será variante no tempo. Neste caso, a Eq. (4.2) será reescrita como:

$$\begin{aligned} r(n) &= \sum_{i=0}^{N_c-1} c_i(n)s(n-i) \\ &= \mathbf{c}^H(n)\mathbf{s}(n), \end{aligned} \quad (4.10)$$

em que $c_i(n)$ é o valor do i -ésimo coeficiente do filtro no instante n , e $\mathbf{c}(n) = [c_0^*(n), c_1^*(n), \dots, c_{N_c-1}^*(n)]^T$. Em muitos aspectos, o projeto de filtros variantes no tempo é mais difícil do que o projeto do filtro de Wiener invariante no tempo, já que para cada instante n , é necessário o conhecimento das estatísticas dos sinais envolvidos para a determininação do conjunto ótimo de coeficientes.

Entretanto, o problema pode ser consideravelmente simplificado se, ao invés de minimizar o erro quadrático médio para cada instante n , considerarmos, para a atualização dos coeficientes, uma solução na forma [109]

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mu\Delta\mathbf{c}(n), \quad (4.11)$$

em que $\Delta\mathbf{c}(n)$ é um fator de correção aplicado ao valor atual dos coeficientes para formar um novo conjunto de estimações para os coeficientes, e μ é uma constante que garante a convergência do algoritmo para \mathbf{c}_o . Esta equação é a chave para o projeto de filtros adaptativos que serão desenvolvidos neste capítulo.

A Eq. (4.11) define um filtro adaptativo, cuja estrutura está mostrada na Fig. 4.3. A saída do filtro é comparada com o sinal desejado (referência). Através da observação do erro de estimação $e(n)$, um algoritmo ajusta os coeficientes do filtro de forma a minimizar alguma função objetivo.

O principal componente de um filtro adaptativo é o algoritmo que define como o fator de correção $\Delta\mathbf{c}(n)$ será determinado. Qualquer que seja o algoritmo utilizado, ele deve satisfazer a algumas propriedades [109]:

1. em um ambiente estacionário, o filtro adaptativo deve produzir um conjunto de coeficientes que convirja para a solução de Wiener;
2. não é necessário o conhecimento das estatísticas dos sinais envolvidos;
3. o filtro deve ser capaz de adaptar-se à mudanças nas estatísticas dos sinais e rastrear a solução conforme evolui o tempo.

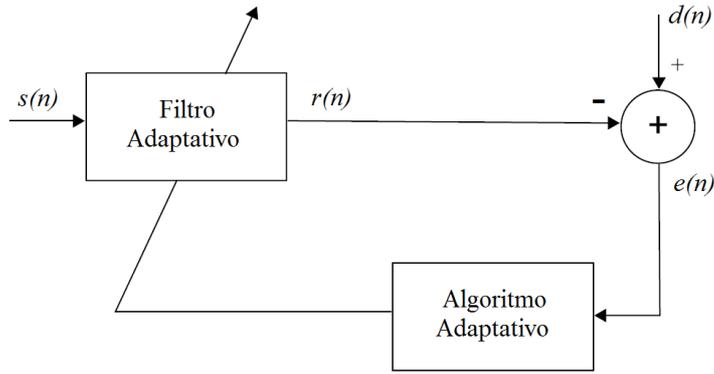


Figura 4.3: Estrutura geral de um filtro adaptativo.

O desempenho de um filtro adaptativo será determinado através de sua velocidade de convergência, que é a taxa com a qual o algoritmo atinge a vizinhança da solução ótima, seu desajuste, que é a diferença relativa entre o erro final do algoritmo e o erro mínimo proporcionado pela solução de Wiener, sua capacidade de acompanhar as variações do sinal de entrada em um ambiente de estatísticas não estacionárias (rastreamento, ou *tracking*), sua complexidade computacional, que é o número de operações realizadas em cada iteração, e suas características numéricas de estabilidade.

Nas próximas seções, serão mostradas algumas formas de se obter a componente de correção $\Delta \mathbf{c}(n)$, que diferenciam os algoritmos adaptativos.

4.3.1 Método do Gradiente Descendente e Método de Newton

De forma geral, o fator de correção $\Delta \mathbf{c}(n)$ deve ser escolhido de tal forma que $|e(n)| < |e(n-1)|$, garantindo que a função objetivo J será monotonicamente decrescente. É possível mostrar que, para que esta condição seja satisfeita, o fator de correção deve possuir a forma [110]

$$\Delta \mathbf{c}(n) = -\mathbf{M} \nabla_{\mathbf{c}^*} J(e(n)), \quad (4.12)$$

em que \mathbf{M} é uma matriz positiva semidefinida qualquer.

Escolhendo $\mathbf{M} = \mathbf{I}$ (a matriz identidade), tem-se o chamado método do gradiente descendente, cuja interpretação geométrica é bastante intuitiva. Como é sabido, o gradiente de uma função determina sua direção de crescimento. Os algoritmos baseados na técnica do gradiente descendente buscam a solução na direção oposta à indicada pelo vetor gradiente, que corresponde à direção de decaimento da função. A equação de atualização dos coeficientes pode ser expressa matematicamente por meio de

$$\mathbf{c}(n+1) = \mathbf{c}(n) - \mu \nabla_{\mathbf{c}^*} J(e(n)). \quad (4.13)$$

De acordo com a Eq. (4.8), o gradiente pode ser escrito como

$$-\nabla_{\mathbf{c}^*} J(e(n)) = \mathcal{R}_{sd} - \mathcal{R}_s \mathbf{c}(n), \quad (4.14)$$

o que conduz à seguinte expressão para a atualização dos coeficientes do filtro:

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mu [\mathcal{R}_{sd} - \mathcal{R}_s \mathbf{c}(n)]. \quad (4.15)$$

Uma outra alternativa consiste em escolher \mathbf{M} como o inverso da matriz Hessiana, $\mathbf{M} = [\nabla_{\mathbf{c}}^2 J(e(n))]^{-1}$, como forma de acelerar a convergência dos algoritmos. Dessa forma, obtem-se o chamado método de Newton, cuja fórmula de recursão é, após alguma manipulação [110], expresso por meio de

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mu [\varepsilon \mathbf{I} + \mathcal{R}_s]^{-1} [\mathcal{R}_{sd} - \mathcal{R}_s \mathbf{c}(n)], \quad (4.16)$$

em que ε é um fator de regularização utilizado para garantir a existência da matriz inversa $[\mathcal{R}_s]^{-1}$.

Os algoritmos definidos pelas Eq. (4.15) e (4.16) ainda dependem do conhecimento das matrizes de \mathcal{R}_{sd} e \mathcal{R}_s . Nas próximas seções, serão apresentados algoritmos que utilizam aproximações estocásticas para as estatísticas dos sinais de entrada e desejado, e por essa razão são denominados métodos do gradiente estocástico (do inglês, *stochastic-gradient*).

4.4 ALGORITMO LMS

O algoritmo denominado LMS (*least mean square*), proposto em 1959 por Widrow e Hoff, é possivelmente o algoritmo de filtragem adaptativa mais utilizado na prática devido à sua simplicidade. Ele procura aproximar as matrizes \mathcal{R}_s e \mathcal{R}_{sd} por seus valores instantâneos

$$\hat{\mathcal{R}}_s = \mathbf{s}(n)\mathbf{s}^H(n) \quad (4.17)$$

e

$$\hat{\mathcal{R}}_{sd} = \mathbf{s}(n)d^*(n). \quad (4.18)$$

Substituindo as Eq. (4.17)-(4.18) na Eq. (4.15), a fórmula de recursão para o gradiente descendente se torna

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mu e^*(n)\mathbf{s}(n), \quad (4.19)$$

O passo de adaptação μ deve ser escolhido de forma a garantir a convergência da recursão definida pela Eq. (4.19). Pode-se mostrar que condição necessária e suficiente para a convergência do método do gradiente descendente (e, em especial, o LMS) é dada por [107]

$$0 < \mu < \frac{2}{\text{Tr}(\mathcal{R}_s)}, \quad (4.20)$$

que depende do conhecimento dos autovalores da matriz de autocovariância de $\mathbf{s}(n)$.

Pode-se mostrar, ainda, que a escolha do passo de adaptação é uma solução de compromisso entre a velocidade de convergência e o desajuste alcançado pelo sistema em regime permanente (após a convergência dos coeficientes) [107]. Valores pequenos de μ geram desajuste menor, mas requerem um número de iterações maior para a convergência do algoritmo.

Apesar desses inconvenientes, o algoritmo LMS apresenta complexidade computacional extremamente reduzida, $O(N_c)$, o que o tornou bastante popular quando comparado aos demais algoritmos de filtragem adaptativa que serão apresentados ao longo deste capítulo.

4.5 ALGORITMO NLMS

O algoritmo NLMS (do inglês, *normalized least mean square*) pode ser obtido a partir do método de Newton, dado pela Eq. (4.16), utilizando as aproximações dadas pelas Eq. (4.17) e (4.18) para as matrizes de autocovariância e de correlação cruzada, da mesma forma que foi feito para o algoritmo LMS.

Utilizando o lema da inversão matricial, é possível obter o seguinte resultado para a equação de atualização dos coeficientes:

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \frac{\mu}{\varepsilon + \mathbf{s}^H(n)\mathbf{s}(n)} e^*(n)\mathbf{s}(n). \quad (4.21)$$

A partir da Eq. (4.21), pode-se inferir que o algoritmo NLMS é uma variação do LMS, porém utilizando um passo de adaptação variável expresso por

$$\mu(n) = \frac{\mu}{\varepsilon + \mathbf{s}^H(n)\mathbf{s}(n)} = \frac{\mu}{\varepsilon + \|\mathbf{s}(n)\|^2}, \quad (4.22)$$

que leva em consideração as variações do nível do sinal de entrada do filtro, isto é, o passo de adaptação é normalizado relativamente à potência instantânea do sinal de entrada em cada iteração. É uma alternativa que procura tornar o algoritmo mais eficiente no sentido de reduzir o tempo de convergência dos coeficientes do filtro, porém aumentando seu desajuste.

É possível ainda mostrar que, por ser baseado no método de Newton, o passo de adaptação fixo μ deve apresentar valores no intervalo $0 < \mu < 2$ para garantir a convergência do algoritmo. Diferentemente do LMS, a escolha de μ independente dos autovalores de \mathcal{R}_s , o que aumenta a robustez do método quanto a escolha do passo de convergência μ .

Apesar de exigir um número maior de operações, pois envolve o cálculo de $\|\mathbf{s}(n)\|^2$, a complexidade do método pode ainda ser aproximada por $O(N_c)$.

4.5.1 Algoritmo Affine Projection

Os algoritmos LMS e NLMS foram obtidos utilizando-se aproximações instantâneas simples para as matrizes de autocovariância e covariância cruzada \mathcal{R}_s e \mathcal{R}_{sd} . Algoritmos de melhor desempenho (em termos de desajuste e velocidade de convergência), mas custo computacional mais elevado, podem ser obtidos se

forem utilizadas aproximações mais sofisticadas para as quantidades supracitadas.

O algoritmo *affine projection* (AP) é uma generalização do algoritmo NLMS. A cada iteração, em vez de utilizar apenas o regressor $\mathbf{s}(n)$ e a observação $d(n)$ para a atualização dos coeficientes, utilizam-se os N_P regressores e observações mais recentes, dados respectivamente por $\mathbf{s}(n-i+1)$ e $d(n-i+1)$, $i = 1, \dots, N_P$ (usualmente, mas não necessariamente, $N_P \leq N_c$). Os valores \mathcal{R}_s e \mathcal{R}_{sd} serão aproximados pelas expressões

$$\begin{aligned}\hat{\mathcal{R}}_s &= \frac{1}{N_P} \left(\sum_{i=n-N_P+1}^n \mathbf{s}(i)\mathbf{s}^H(i) \right) \\ &= \frac{1}{N_P} \mathbf{S}(n)\mathbf{S}^H(n)\end{aligned}\quad (4.23)$$

e

$$\begin{aligned}\hat{\mathcal{R}}_{sd} &= \frac{1}{N_P} \left(\sum_{i=n-N_P+1}^n d^*(i)\mathbf{s}(i) \right) \\ &= \frac{1}{N_P} \mathbf{S}(n)\mathbf{d}^*(n),\end{aligned}\quad (4.24)$$

para as quais convencionou-se

$$\mathbf{S}(n) \triangleq \begin{bmatrix} \mathbf{s}(n) & \mathbf{s}(n-1) & \dots & \mathbf{s}(n-N_P+1) \end{bmatrix}\quad (4.25)$$

e

$$\mathbf{d}(n) = \begin{bmatrix} d(n) \\ d(n-1) \\ \vdots \\ d(n-N_P+1) \end{bmatrix}.\quad (4.26)$$

A atualização dos coeficientes, utilizando o método de Newton e o lema da inversão matricial, será dada por

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mu \mathbf{S}(n) [\varepsilon \mathbf{I} + \mathbf{S}^H(n)\mathbf{S}(n)]^{-1} \mathbf{e}^*(n),\quad (4.27)$$

em que $\mathbf{e}^*(n)$ corresponde ao vetor de erros de estimação, dado por

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{S}^T(n)\mathbf{c}^*(n).\quad (4.28)$$

É possível mostrar que escolhendo $\varepsilon = 0$ e $\mu = 1$, o algoritmo AP pode ser obtido ao resolver o problema de otimização [110]:

$$\begin{aligned} \min \|\mathbf{c}(n+1) - \mathbf{c}(n)\|^2 \\ \text{sujeito a } \mathbf{d}(n) - \mathbf{S}^T(n)\mathbf{c}^*(n+1) = 0, \end{aligned} \quad (4.29)$$

cuja solução será dada pela Eq. (4.27):

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mathbf{S}(n) [\mathbf{S}^H(n)\mathbf{S}(n)]^{-1} \mathbf{e}^*(n). \quad (4.30)$$

Em outras palavras, o algoritmo atualiza os coeficientes de modo que os erros *a posteriori*, dados por $\mathbf{d}(n) - \mathbf{S}^T(n)\mathbf{c}^*(n+1)$, são simultaneamente forçados para zero. A interpretação geométrica das Eq. (4.29) e (4.30) é a seguinte: pode-se dizer que existem infinitos vetores \mathbf{c} que satisfazem a condição $d(n) = \mathbf{c}^H \mathbf{s}(n)$. Todos os valores de \mathbf{c} que satisfazem a condição anterior formam o hiperplano $\mathcal{M}(n)$. Para $N_P > 1$, a recursão definida pela Eq. (4.30) força as seguintes N_P igualdades:

$$\begin{aligned} d(n) &= \mathbf{c}^H(n+1)\mathbf{s}(n) \\ d(n-1) &= \mathbf{c}^H(n+1)\mathbf{s}(n-1) \\ &\vdots \\ d(n-N_P+1) &= \mathbf{c}^H(n+1)\mathbf{s}(n-N_P+1) \end{aligned} \quad (4.31)$$

Para cada par $d(n-i), \mathbf{s}(n-i)$, existem infinitos vetores \mathbf{c} que satisfazem $d(n-i) = \mathbf{c}^H \mathbf{s}(n-i)$, cada um definindo um hiperplano $\mathcal{M}(n-i)$. O vetor $\mathbf{c}(n+1)$ que é calculado pelo algoritmo *affine projection* se encontra na interseção dos últimos N_P hiperplanos correspondentes, isto é,

$$\mathbf{c}(n+1) \in \left(\bigcap_{i=n-N_P+1}^n \mathcal{M}(i) \right). \quad (4.32)$$

A Fig. 4.4 mostra esta construção para o caso $N_P = 2$. A estimativa $\mathbf{c}(n)$ se encontra no hiperplano $\mathcal{M}(n-1)$, enquanto que a estimativa atualizada $\mathbf{c}(n+1)$ se encontra na interseção $\mathcal{M}(n-1) \cap \mathcal{M}(n)$.

Quando $N_P = 1$, tem-se o caso especial NLMS, para o qual a atualização do vetor de coeficientes é vista como uma projeção afim unidimensional. No algoritmo *affine projection*, as projeções são realizadas em múltiplas dimensões e, conforme o número de dimensões aumenta, a velocidade de convergência aumenta também. Entretanto, a complexidade no algoritmo *affine projection* será $O(N_P^2 N_c)$, que depende do número de projeções N_P a serem realizadas, sendo consideravelmente mais complexo que o NLMS.

Pode-se afirmar, a partir das considerações anteriores, que o algoritmo *affine projection* permite um compromisso entre velocidade de convergência e baixa complexidade computacional. Com o ajuste adequado da quantidade N_P de dados a serem reutilizados, é possível obter desempenho melhor que o dos algoritmos RLS (que será apresentado na próxima seção) e NLMS em termos de velocidade de convergência, porém o nível de desajuste do algoritmo pode aumentar.

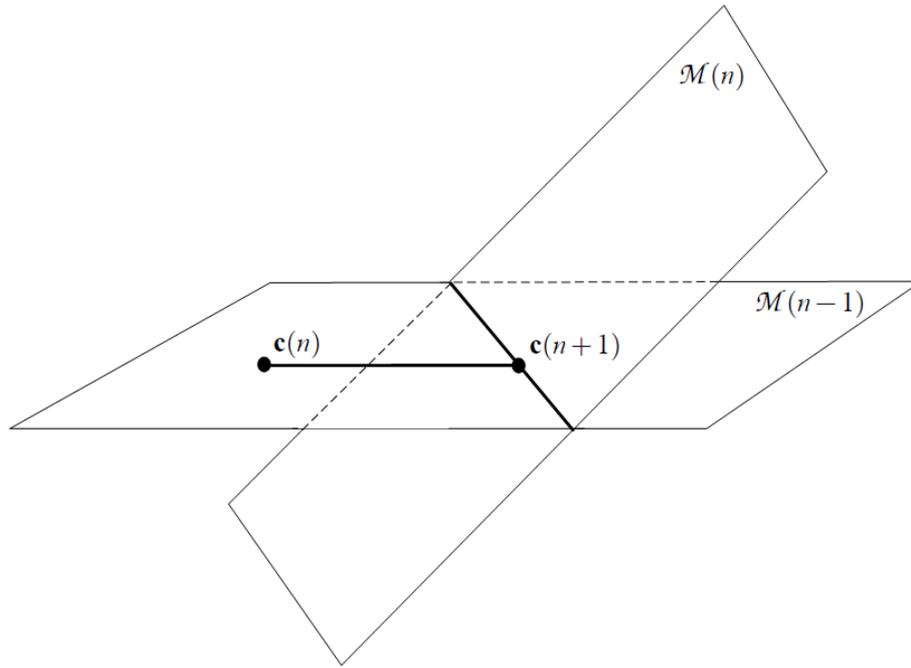


Figura 4.4: Hiperplanos associados à atualização dos coeficientes do filtro adaptativo utilizando o algoritmo *affine projection* com $N_P = 2$.

4.6 ALGORITMO RLS

Um outro exemplo de algoritmo que emprega uma aproximação um pouco mais sofisticada para as matrizes de autocovariância e correlação cruzada é o RLS (do inglês, *recursive least squares*). Apesar de o algoritmo poder ser derivado como solução exata do problema clássico de estimação por mínimos quadrados, aqui ele será apresentado como uma outra versão de aproximação estocástica para o método de Newton [78, 110].

A partir da equação de recursão do método de Newton, dada pela Eq. (4.16), a componente expressa por $\mathcal{R}_{sd} - \mathcal{R}_s \mathbf{c}(n)$ será estimada por meio da aproximação instantânea

$$\mathcal{R}_{sd} - \mathcal{R}_s \mathbf{c}(n) \approx \mathbf{s}(n) [d^*(n) - \mathbf{s}^H(n) \mathbf{c}(n)], \quad (4.33)$$

e \mathcal{R}_s será substituída pela média ponderada

$$\hat{\mathcal{R}}_s = \frac{1}{n+1} \sum_{i=0}^n \lambda^{n-i} \mathbf{s}(i) \mathbf{s}^H(i) \quad (4.34)$$

para algum escalar λ , denominado fator de esquecimento, e ajustado no intervalo $0 < \lambda \leq 1$.

A Eq. (4.34) utiliza uma média de todos os vetores de entrada até o instante n para o cômputo de uma estimativa para a matriz de autocovariância do sinal de entrada. Para $\lambda = 1$, todos os vetores regressores $\mathbf{s}(n)$ ponderam da mesma maneira o cálculo de $\hat{\mathcal{R}}_s$. Escolhendo um valor de $\lambda \neq 1$, introduz-se memória finita no algoritmo, pois os regressores mais recentes possuirão maior peso no cálculo dessa estimativa, e amostras passadas possuirão peso menor, sendo possível "esquecer" amostras que estejam em um passado remoto e dar maior relevância às amostras mais recentes.

Escolhendo $\mu = \mu(n) = 1/(n + 1)$, $\varepsilon = \varepsilon(n) = \lambda^{n+1}\varepsilon/(n + 1)$, e utilizando o lema da inversão matricial, são obtidas as seguintes equações para a atualização dos coeficientes do filtro adaptativo:

$$\Psi(n + 1) = \lambda^{-1} \left[\Psi(n) - \lambda^{-1} \frac{\Psi(n)\mathbf{s}(n)\mathbf{s}^H(n)\Psi(n)}{1 + \lambda^{-1}\mathbf{s}^H(n)\Psi(n)\mathbf{s}(n)} \right], \quad (4.35)$$

$$\mathbf{c}(n + 1) = \mathbf{c}(n) + \Psi(n + 1) [d(n) - \mathbf{c}^H(n)\mathbf{s}(n)]^*. \quad (4.36)$$

O algoritmo RLS apresenta elevada velocidade de convergência, que independente da diferença entre os autovalores da matriz de autocovariância do sinal de entrada (diferentemente do que ocorre com o algoritmo LMS). Em regime permanente, o desajuste do algoritmo pode ser muito próximo de zero. Por outro lado, ele apresenta complexidade computacional mais elevada, $O(N_c^2)$, e pode também apresentar problemas de estabilidade e rastreamento, de acordo com o valor escolhido para o fator de esquecimento [109].

4.7 ALGORITMO SET-MEMBERSHIP AFFINE PROJECTION

4.7.1 Filtragem Set-Membership

Os algoritmos convencionais de filtragem considerados nas seções anteriores estimam os coeficientes $\mathbf{c}(n)$ do filtro adaptativo de forma que o erro de estimação $e(n)$ seja feito pequeno suficiente, de forma a minimizar determinada função objetivo.

A filtragem *set-membership*, introduzida em [111], tem como objetivo alcançar determinado valor limite ν na estimativa do módulo do erro, $|e(n)|$. Qualquer coeficiente estimado que resulte em um erro menor do que o especificado é uma solução aceitável. Logo, a solução para o problema da filtragem *set-membership* é um conjunto no espaço dos coeficientes, em vez de apenas um ponto dentro desse espaço.

Seja então $\mathcal{H}(n)$, denominado *constraint set*, o conjunto que contém todos os vetores \mathbf{c} para os quais o erro de saída no instante n tem limite superior (em magnitude) dado por ν , ou seja,

$$\mathcal{H}(n) = \{ \mathbf{c} \in \mathbb{C}^{N_c \times 1} : |d(n) - \mathbf{c}^H \mathbf{s}(n)| \leq \nu \}. \quad (4.37)$$

O *set membership* $\psi(n)$ é definido como a interseção dos *constraint sets* até o instante n :

$$\psi(n) = \bigcap_{i=1}^n \mathcal{H}(i). \quad (4.38)$$

A escolha do limiar ν é um parâmetro de projeto e está relacionado à taxa de atualização dos coeficientes do filtro após a convergência, o erro de estado estacionário, o desajuste e o nível de ruído presente no sistema. Como pode ser observado na Eq. (4.37), se o limiar escolhido for muito pequeno, o conjunto $\psi(n)$ pode ser vazio. Se o limiar for muito grande, estimativas inconsistentes podem ser obtidas [106]. Para uma escolha apropriada do valor de ν , algumas regras práticas são sugeridas na literatura de *set-membership* [112].

Uma importante consequência da definição dada pela Eq. (4.37) é que, se a estimativa dos coeficientes $\mathbf{c}(n)$ produz um erro inferior ao limiar especificado, então não é necessário realizar uma atualização dos coeficientes, pois

$$\mathbf{c}(n+1) = \mathbf{c}(n), \text{ se } |d(n) - \mathbf{c}^H(n)\mathbf{s}(n)| \leq \nu. \quad (4.39)$$

Com esta estratégia, é possível reduzir, na média, a complexidade computacional dos algoritmos que utilizam o conceito de *set membership*, pois os coeficientes não serão, necessariamente, atualizados em cada iteração do processo definido pela Eq. (4.11).

4.7.2 Set-Membership Affine Projection

O algoritmo *set-membership affine projection* (SM-AP), apresentado em [113], é uma extensão do *framework* definido pela filtragem *set-membership* para o algoritmo *affine projection*. Conforme já exposto anteriormente, o algoritmo *affine projection* permite um compromisso entre velocidade de convergência e baixa complexidade computacional. Entretanto, o nível de desajuste do algoritmo AP pode aumentar com a convergência acelerada. Uma das formas de contornar este problema é a utilização do conceito de filtragem *set-membership*. Diferentemente dos algoritmos considerados até o momento, não há conflito entre a velocidade de convergência, o desajuste e a complexidade computacional no algoritmo SM-AP [114].

A Eq. (4.38) sugere que vários *constraint sets* podem ser utilizados para determinar a atualização dos coeficientes do filtro. Será considerado o caso em que os coeficientes atualizados pertencem ao conjunto formado pelos últimos N_P *constraint sets*. Dessa forma, é utilizado o conceito de reaproveitamento de regressores passados, que foi introduzido pelo algoritmo adaptativo *affine projection*.

Em primeiro lugar, o *set membership* definido pela Eq. (4.38) pode ser expresso como

$$\begin{aligned} \psi(n) &= \bigcap_{i=1}^{n-N_P} \mathcal{H}(i) \quad \bigcap_{l=n-N_P+1}^n \mathcal{H}(l) \\ &= \psi^{n-N_P}(n) \cap \psi^{N_P}(n), \end{aligned} \quad (4.40)$$

em que $\psi^{n-N_P}(n)$ é a interseção dos primeiros $n - N_P$ *constraint sets* e $\psi^{N_P}(n)$ é a interseção dos últimos N_P *constraint sets*. Para o algoritmo *set membership affine projection*, os coeficientes atualizados devem estar contidos nos últimos N_P *constraint sets*, o que significa que $\mathbf{c}(n+1) \in \psi^{N_P}(n)$.

Seja $\mathcal{M}(n-i+1)$ o hiperplano que contém todos os vetores \mathbf{c} que satisfazem a igualdade definida por $d(n-i+1) - \mathbf{c}^H \mathbf{s}(n-i+1) = g_i(n)$, $i = 1, \dots, N_P$, $g_i(n)$ é um ponto do hiperplano $\mathcal{M}(n-i+1)$. Se $|g_i(n)| \leq \nu$, então $\mathcal{M}(n-i+1) \in \mathcal{H}(n-i+1)$.

O seguinte critério de otimização é utilizado para determinar a atualização dos coeficientes sempre que $\mathbf{c}(n) \notin \psi^{N_P}(n)$:

$$\begin{aligned} \min \|\mathbf{c}(n+1) - \mathbf{c}(n)\|^2 \\ \text{sujeito a } \mathbf{d}(n) - \mathbf{S}^T(n)\mathbf{c}^*(n+1) = \mathbf{g}(n), \end{aligned} \quad (4.41)$$

para o qual definiu-se o vetor de parâmetros

$$\mathbf{g}(n) = \begin{bmatrix} g_1(n) \\ g_2(n) \\ \vdots \\ g_{n-N_P+1}(n) \end{bmatrix}. \quad (4.42)$$

O problema dado pela Eq. (4.41) pode ser resolvido por meio dos multiplicadores de Lagrange. Sua solução resulta na seguinte equação geral de atualização:

$$\mathbf{c}(n+1) = \begin{cases} \mathbf{c}(n) + \mathbf{S}(n) [\mathbf{S}^H(n)\mathbf{S}(n)]^{-1} [\mathbf{e}^*(n) - \mathbf{g}^*(n)], & \text{se } |e(n)| > \nu \\ \mathbf{c}(n), & \text{caso contrário} \end{cases} \quad (4.43)$$

com

$$\mathbf{e}(n) = \begin{bmatrix} e(n) \\ \epsilon(n-1) \\ \vdots \\ \epsilon(n-N_P+1) \end{bmatrix}, \quad (4.44)$$

em que $\epsilon(n-i) = d(n-i) - \mathbf{c}^H(n)\mathbf{s}(n-i)$ é denominado erro *a posteriori* na iteração $n-i$.

4.7.3 Escolha do vetor de parâmetros

A definição dos valores de $g_i(n)$ é fundamental para o desempenho do algoritmo, e existe um número infinito de escolhas para o vetor de parâmetros $\mathbf{g}(n)$.

Por exemplo, para $\mathbf{g}(n) = \mathbf{0}$, tem-se o algoritmo *affine projection* com $\mu = 1$ e $\varepsilon = 0$ quando os coeficientes $\mathbf{c}(n) \notin \psi^{N_P}(n)$. Comparado ao algoritmo *affine projection* convencional, é observada redução da complexidade computacional dada a natureza seletiva do algoritmo para atualização dos coeficientes. Se $\mathbf{g}(n) = \mathbf{0}$ e $\nu = 0$, tem-se o algoritmo *affine projection* para a atualização dos coeficientes, conforme expresso pela Eq. (4.27).

Uma versão particularmente simples pode ser obtida escolhendo-se $g_i(n) = \epsilon(n-i+1)$, para os valores de $i = 2, \dots, N_P$, e $g_1(n) = \nu \frac{e(n)}{|e(n)|}$, de forma que a solução esteja na fronteira mais próxima de $\mathcal{H}(n)$. A partir destas escolhas, ter-se-á [113]

$$\mathbf{e}_{ap}(n) - \mathbf{g}(n) = \begin{bmatrix} e(n) \\ \epsilon(n-1) \\ \epsilon(n-2) \\ \vdots \\ \epsilon(n-N_P+1) \end{bmatrix} - \begin{bmatrix} e(n) \frac{\nu}{|e(n)|} \\ \epsilon(n-1) \\ \epsilon(n-2) \\ \vdots \\ \epsilon(n-N_P+1) \end{bmatrix} = \begin{bmatrix} e(n) \left[1 - \frac{\nu}{|e(n)|}\right] \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (4.45)$$

e a Eq. (4.43) de atualização dos coeficientes será dada por

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \mathbf{S}(n) [\mathbf{S}^H(n)\mathbf{S}(n)]^{-1} e(n)\kappa(n)\mathbf{v}_1, \quad (4.46)$$

em que

$$\kappa(n) = \begin{cases} 1 - \frac{\nu}{|e(n)|}, & \text{se } |e(n)| > \nu \\ 0, & \text{caso contrário} \end{cases} \quad (4.47)$$

e

$$\mathbf{v}_1 = [1 \ 0 \ \dots \ 0]^T. \quad (4.48)$$

A interpretação geométrica da situação apresentada pela Eq. (4.46) é mostrada na Fig. 4.5. Procura-se minimizar a distância euclidiana $\|\mathbf{c}(n+1) - \mathbf{c}(n)\|^2$ sujeito à restrição de que $\mathbf{c}(n+1) \in \psi^{N_P}(n)$, de forma que o erro *a posteriori* na iteração $n-i$ é mantido constante para $i = 1, \dots, N_P$.

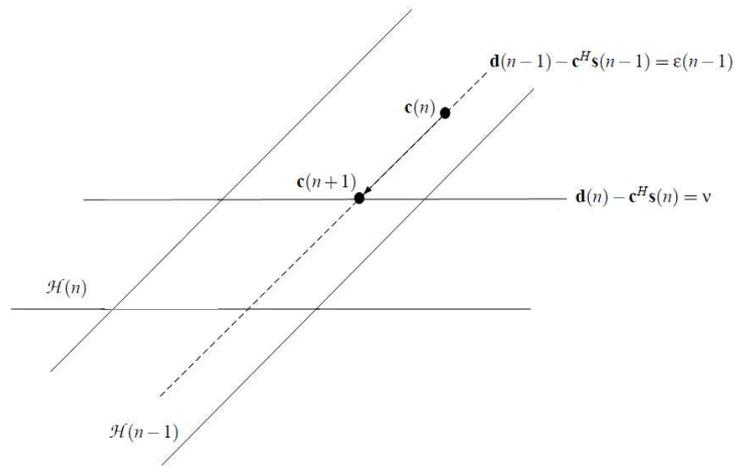


Figura 4.5: Interpretação geométrica do algoritmo *set-membership affine projection* que resulta em erro *a posteriori* constante.

4.8 PROPOSTA DE PREDITOR DE CANAL

Conforme exposto no capítulo 3, para que os ganhos oferecidos pelas técnicas de pré-codificação possam ser plenamente explorados, a resposta do canal tem de ser estimada com antecedência, uma vez que informação desatualizada sobre o estado do canal pode comprometer seriamente o desempenho do sistema. Nesta seção, é proposto um preditor adaptativo baseado no algoritmo *set-membership affine projection* para sistemas MIMO-OFDM multiusuário.

Na aplicação de predição, o filtro adaptativo fornece uma estimativa futura para o valor do sinal de entrada. O valor $s(n)$ presente na entrada do sistema é utilizado como resposta desejada $d(n)$, e os valores passados $s(n - \ell)$ correspondem à entrada do filtro adaptativo, conforme mostrado na Fig. 4.6. O parâmetro ℓ é chamado horizonte de predição, e denota o número de amostras para as quais a resposta do canal é predita à frente do instante n .

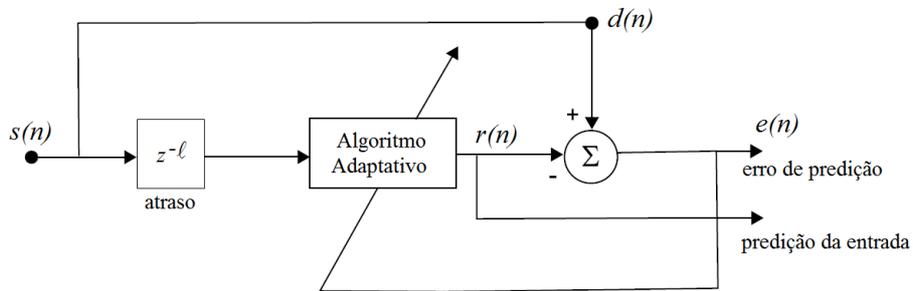


Figura 4.6: Estrutura adaptativa aplicada à predição de sinal.

4.8.1 Modelo do sistema

Conforme exposto, será considerado um ambiente com K usuários, cada um com uma antena de recepção, e uma estação rádio-base com M_T antenas de transmissão. A escolha de usuários com apenas uma antena de recepção é justificada pelo fato de ser o caso mais considerado na literatura, conforme exposto no capítulo 3. Entretanto, o preditor de canal apresentado pode ser facilmente estendido para o caso em que os usuários apresentem mais de uma antena em seus terminais.

É utilizada a técnica OFDM para a transmissão com N_{DFT} subportadoras (uma discussão detalhada do OFDM pode ser encontrada em [115, 1]). Dessa forma, $H_{1,I}^k(n, \mathcal{P})$ é a resposta do canal entre a antena de recepção do k -ésimo usuário e a I -ésima antena de transmissão, $I = 1, \dots, M_T$, na \mathcal{P} -ésima subportadora do n -ésimo símbolo OFDM, com $\mathcal{P} = 0, \dots, N_{DFT} - 1$.

Como cada subportadora define um canal não seletivo em frequência [1], é possível implementar algoritmos MIMO de processamento separadamente para cada subportadora. O modelo de sinal apresentado no capítulo 2 passa agora a ser indexado por subportadora.

No modelo a ser apresentado, supõe-se que a resposta do canal pode variar consideravelmente entre símbolos OFDM diferentes, mas não dentro do mesmo símbolo. Conside-se ainda que uma extensão cíclica (também denominada prefixo cíclico) é inserida em cada símbolo OFDM de forma a eliminar a interferência intersimbólica causada pelo canal de comunicação. Neste caso, $y_k(n, \mathcal{P})$ representa o sinal

recebido pelo k -ésimo usuário na \mathcal{P} -ésima subportadora do n -ésimo símbolo OFDM recebido, expresso por

$$\begin{aligned}
y_k(n, \mathcal{P}) &= \mathbf{H}_k(n, \mathcal{P})\mathbf{W}_k(n, \mathcal{P})u_k(n, \mathcal{P}) + \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{H}_k(n, \mathcal{P})\mathbf{W}_i(n, \mathcal{P})u_i(n, \mathcal{P}) + n_k(n, \mathcal{P}) \\
&= \mathbf{H}(n, \mathcal{P})\mathbf{W}(n, \mathcal{P})\mathbf{u}(n, \mathcal{P}) + n_k(n, \mathcal{P}) \\
&= \mathbf{H}(n, \mathcal{P})\mathbf{x}(n, \mathcal{P}) + n_k(n, \mathcal{P}),
\end{aligned} \tag{4.49}$$

em que $n_k(n, \mathcal{P})$ é o termo de ruído aditivo gaussiano na antena de recepção do k -ésimo usuário na subportadora \mathcal{P} do n -ésimo símbolo OFDM,

$$\mathbf{H}_k(n, \mathcal{P}) = \begin{bmatrix} H_{1,1}^k(n, \mathcal{P}) & \dots & H_{1,M_T}^k(n, \mathcal{P}) \end{bmatrix} \tag{4.50}$$

é a matriz de canal do k -ésimo usuário,

$$\mathbf{H}(n, \mathcal{P}) = \begin{bmatrix} \mathbf{H}_1(n, \mathcal{P}) & \dots & \mathbf{H}_K(n, \mathcal{P}) \end{bmatrix}^T \tag{4.51}$$

é a matriz combinada de canal,

$$\mathbf{W}(n, \mathcal{P}) = \begin{bmatrix} \mathbf{W}_1(n, \mathcal{P}) & \dots & \mathbf{W}_K(n, \mathcal{P}) \end{bmatrix} \tag{4.52}$$

é a matriz de pesos para a pré-codificação,

$$\mathbf{u}(n, \mathcal{P}) = \begin{bmatrix} u_1(n, \mathcal{P}) & \dots & u_K(n, \mathcal{P}) \end{bmatrix}^T \tag{4.53}$$

representa o vetor de símbolos de informação dos usuários, e

$$\begin{aligned}
\mathbf{x}(n, \mathcal{P}) &= \mathbf{W}(n, \mathcal{P})\mathbf{u}(n, \mathcal{P}) \\
&= \begin{bmatrix} \mathbf{W}_1(n, \mathcal{P}) & \dots & \mathbf{W}_K(n, \mathcal{P}) \end{bmatrix} \begin{bmatrix} u_1(n, \mathcal{P}) \\ \vdots \\ u_K(n, \mathcal{P}) \end{bmatrix}
\end{aligned} \tag{4.54}$$

é o vetor de símbolos pré-codificados, e expressa o que é efetivamente transmitido pela estação rádio base. Ao derivar a Eq. (4.49), utilizou-se um algoritmo de pré-codificação linear, expresso pela matriz de pesos $\mathbf{W}(n, \mathcal{P})$. Entretanto, o modelo é facilmente estendido para algoritmos não lineares escrevendo o sinal efetivamente transmitido pela estação rádio-base como $\mathbf{x}(n, \mathcal{P}) = \mathbf{f}(\mathbf{u}(n, \mathcal{P}))$, em que \mathbf{f} é uma função qualquer do vetor de símbolos \mathbf{u} . Para os pré-codificadores lineares, tem-se o caso particular em que a função \mathbf{f} é dada por $\mathbf{f}(\mathbf{u}(n, \mathcal{P})) = \mathbf{W}(n, \mathcal{P})\mathbf{u}(n, \mathcal{P})$.

4.8.2 Metodologia de Predição

Uma das abordagens utilizadas para a predição de canais MIMO é realizá-la separadamente para cada coeficiente de canal [116, 117]. Logo, a predição dos coeficientes da matriz $\mathbf{H}_k(n, \mathcal{P})$ pode ser decomposta em M_T preditores, um para cada resposta de canal referente, a cada antena de transmissão. Apesar de ser um esquema sub-ótimo [118], pois deixa de explorar a correlação que pode existir entre as informações presentes nas múltiplas antenas, ele reduz consideravelmente a complexidade computacional exigida quando comparado a algoritmos que utilizam modelos de predição auto-regressivos e técnicas como o ESPRIT (*estimation of signal parameters via rotational invariance techniques*), conforme apresentado em [119].

Entretanto, mesmo realizando a predição separadamente para cada uma das M_T antenas de transmissão, ainda é necessário predizer a resposta do canal para cada subportadora. Utilizar preditores individuais para cada subportadora é inviável sob o ponto de vista prático para um número muito grande de subportadoras [120, 121]. Entretanto, este problema pode ser contornado realizando a predição de canal do domínio do tempo [122].

Pode-se escrever que a resposta em frequência $H_{1,I}^k(n, \mathcal{P})$ do canal está relacionada com sua resposta impulsional $h_{1,I}^k(n, l)$ por meio de uma DFT (do inglês, *discrete Fourier transform*, ou transformada discreta de Fourier), expressa como

$$H_{1,I}^k(n, \mathcal{P}) = \sum_{l=0}^L h_{1,I}^k(n, l) e^{-j2\pi Pl/N_{DFT}}, \quad (4.55)$$

em que l foi utilizado para indexar determinado multipercorso da resposta impulsional do canal, com $0 \leq l \leq L$. Como consequência, L representa o máximo espalhamento de retardo do canal com $L + 1$ multipercursos.

Na prática, não é conhecida a resposta verdadeira do canal $h_{1,I}^k(n, l)$, mas apenas uma estimativa $\hat{h}_{1,I}^k(n, l)$,

$$\hat{h}_{1,I}^k(n, l) = h_{1,I}^k(n, l) + \tilde{e}, \quad (4.56)$$

em que \tilde{e} representa o erro de estimação na resposta do canal, que é função da razão sinal-ruído na entrada do sistema receptor e do método de estimação escolhido [123].

Como demonstrado em [122], sob a hipótese de que a resposta impulsional do canal pode ser caracterizada por um processo estacionário no sentido amplo com espalhadores não correlatados (WSSUS - *wide sense stationary uncorrelated scattering*), a predição de canal pode ser realizada decompondo sua resposta impulsional em $L + 1$ preditores paralelos (a demonstração deste fato pode ser encontrada em [124, 125]), expressos por

$$\hat{h}_{1,I}^k(n + \ell, l) = \mathbf{c}_{I,k}^H(n, l) \hat{\mathbf{h}}_{1,I}^k(n, l), \quad l = 0, \dots, L, \quad (4.57)$$

em que

$$\hat{\mathbf{h}}_{1,I}^k(n, l) = \left[\hat{h}_{1,I}^k(n, l) \quad \hat{h}_{1,I}^k(n-1, l) \quad \dots \quad \hat{h}_{1,I}^k(n-N_c+1, l) \right]^T \quad (4.58)$$

representa o vetor de regressores e

$$\mathbf{c}_{I,k}(n, l) = \begin{bmatrix} c_{0,I,k}(n, l)^* \\ c_{1,I,k}(n, l)^* \\ \vdots \\ c_{N_c-1,I,k}(n, l)^* \end{bmatrix} \quad (4.59)$$

contém os N_c coeficientes do preditor, conforme dado pela Eq. (4.2).

A determinação dos coeficientes $\mathbf{c}_{I,k}(n, l)$ é realizada utilizando-se o algoritmo *set-membership affine projection*, expresso pelas Eq. (4.46) - (4.48). Sua escolha é justificada pelo fato de o algoritmo apresentar boa velocidade de convergência e reduzida taxa de atualização de coeficientes devido à sua natureza seletiva. Neste caso, tem-se as seguintes equações de atualização para os coeficientes do filtro preditor:

$$\mathbf{c}_{I,k}(n+1, l) = \mathbf{c}_{I,k}(n, l) + \hat{\mathbf{H}}_{1,I}^k(n, l) \left[\left(\hat{\mathbf{H}}_{1,I}^k(n, l) \right)^H \hat{\mathbf{H}}_{1,I}^k(n, l) \right]^{-1} e_{I,k}(n, l) \kappa_{I,k}(n, l) \mathbf{v}_1, \quad (4.60)$$

em que

$$\hat{\mathbf{H}}_{1,I}^k(n, l) = \left[\hat{\mathbf{h}}_{1,I}^k(n, l) \quad \hat{\mathbf{h}}_{1,I}^k(n-1, l) \quad \dots \quad \hat{\mathbf{h}}_{1,I}^k(n-N_p+1, l) \right], \quad (4.61)$$

$$e_{I,k}(n, l) = \hat{\mathbf{h}}_{1,I}^k(n, l) - \mathbf{c}_{I,k}^H(n, l) \hat{\mathbf{h}}_{1,I}^k(n-l, l), \quad (4.62)$$

$$\kappa_{I,k}(n, l) = \begin{cases} 1 - \frac{\nu}{|e_{I,k}(n, l)|}, & \text{se } |e_{I,k}(n, l)| > \nu_{I,k} \\ 0, & \text{caso contrário} \end{cases} \quad (4.63)$$

O valor $\nu_{I,k}$ é utilizado para definir o *constraint set* do k -ésimo usuário associado à I -ésima antena do transmissor (estação rádio-base).

A Eq. (4.62) representa uma aproximação para o erro de predição, visto que ela é baseada na resposta estimada do canal, e não na resposta verdadeira, conforme mostrou a Eq. (4.56). Entretanto, para valores altos de relação sinal-ruído, o erro introduzido por esta aproximação pode ser considerado desprezível, e não comprometerá de modo muito severo o desempenho do preditor.

Além disso, os coeficientes do preditor são inicializados segundo

$$\mathbf{c}_{I,k}(n, l) = [1 \ 0 \ \dots \ 0]^T, \quad n = 0, \dots, \ell-1, \quad (4.64)$$

de forma que $\hat{h}_{1,I}^k(n+\ell, l) = \hat{h}_{1,I}^k(n, l)$ para $n = 0, \dots, \ell-1$.

A resposta predita em cada subportadora, $\hat{H}_{1,I}^k(n+\ell, \mathcal{P})$, pode ser obtida por meio de uma DFT

$$\hat{H}_{1,I}^k(n + \ell, \mathcal{P}) = \sum_{l=0}^L \hat{h}_{1,I}^k(n + \ell, l) e^{-j2\pi \mathcal{P}l/N_{DFIT}}. \quad (4.65)$$

4.9 CONCLUSÃO

Neste capítulo, descreveram-se os algoritmos clássicos de filtragem adaptativa LMS, NLMS, AP e RLS, procurando enfatizar as características computacionais dos algoritmos, como complexidade, velocidade de convergência e desajuste. Realizou-se uma breve descrição da filtragem do tipo *set membership*, que permite reduzir a complexidade computacional dos algoritmos que a utilizam por meio da diminuição do número de atualizações dos coeficientes do filtro. Como aplicação do conceito de *set membership*, abordou-se o algoritmo *set-membership affine projection*. Em último lugar, foi proposto um preditor adaptativo de canal para sistemas MIMO-OFDM no cenário multiusuário que utiliza o algoritmo SM-AP.

Conforme discutido no capítulo 3, filtros preditores podem ser utilizados como forma de compensar a presença de CSI desatualizada no transmissor. Dessa forma, os ganhos oferecidos pelas técnicas de pré-codificação para o canal MIMO BC podem ser plenamente explorados. No próximo capítulo, serão apresentados resultados de simulações que mostram o desempenho de alguns dos algoritmos adaptativos considerados neste capítulo ao serem utilizados na predição da resposta do canal rádio móvel em sistemas MIMO multiusuário. As simulações do próximo capítulo também incluem os efeitos da informação imperfeita de estado de canal no desempenho de alguns dos pré-codificadores apresentados no capítulo 3, bem como seu desempenho ao utilizar alguns dos algoritmos adaptativos para realizar a predição da resposta do canal de comunicação como forma de compensar a CSI desatualizada.

5 RESULTADOS DE SIMULAÇÕES

5.1 INTRODUÇÃO

Após considerar a descrição das técnicas de pré-codificação utilizadas no canal MIMO BC, realizada no capítulo 3, e dos algoritmos clássicos de filtragem adaptativa, no capítulo 4, serão apresentados os resultados de simulação para o desempenho dos preditores e dos pré-codificadores em um cenário de simulação um pouco mais realista. Para tal, será utilizado o enlace direto do padrão 3GPP-LTE [126]. Para modelar a evolução temporal do canal de rádio e o atraso introduzido pelo canal de retorno, utilizar-se-á o modelo espacial de canal padronizado pelo 3GPP para simulação de sistemas MIMO [127].

Inicialmente, serão apresentados os cenários de simulação e a modelagem utilizada. Em seguida, serão mostrados os resultados de simulação do desempenho dos algoritmos de filtragem adaptativa utilizados na aplicação de predição da resposta do canal. Finalmente, o desempenho dos pré-codificadores será apresentado, considerando a utilização dos algoritmos de predição como forma de compensar o atraso introduzido pelo canal de retorno.

5.2 CENÁRIOS DE SIMULAÇÃO

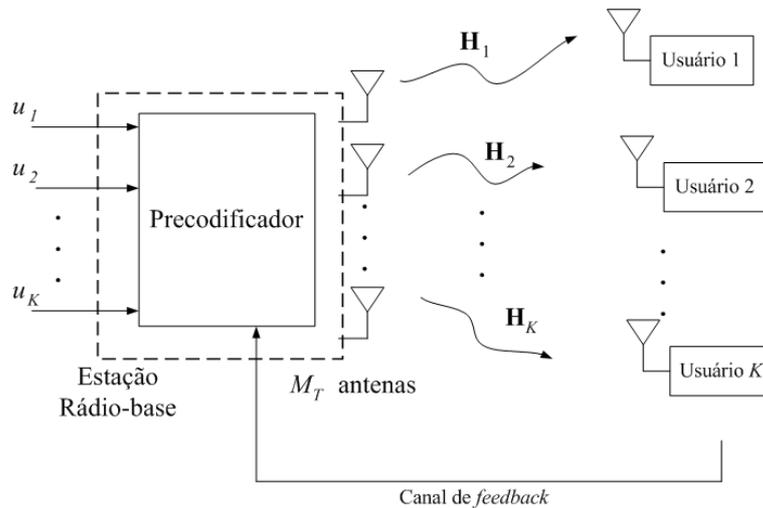
O sistema simulado está ilustrado na Fig. 5.1 (a). Tem-se o canal MIMO BC com $K = 4$ usuários, cada um com uma antena de recepção, e uma estação rádio base com $M_T = 4$ antenas de transmissão. A escolha desse número de antenas é justificada pelo fato de que o padrão LTE, como está atualmente especificado, define a utilização de até quatro antenas na estação rádio-base.

A formatação da informação dentro do quadro de rádio é feita de acordo com as especificações do padrão 3GPP-LTE sem a utilização de codificação de canal, e estão mostradas na Tabela 5.1. Detalhes adicionais quanto às características do quadro de rádio do enlace direto do padrão LTE podem ser encontrados no Anexo I. Conforme exposto no capítulo 3, a pré-codificação é realizada separadamente em cada uma das subportadoras dos símbolos OFDM.

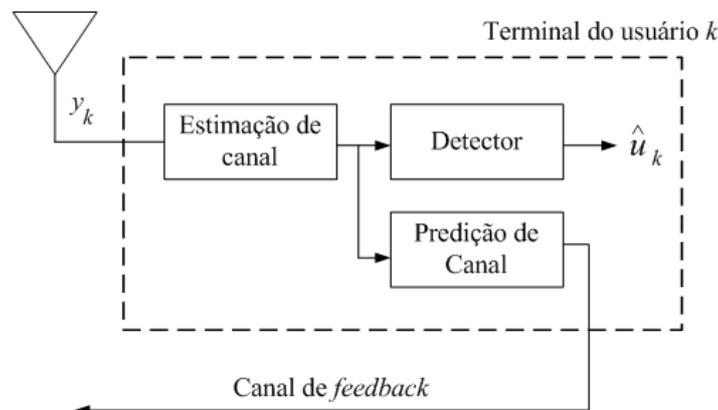
Tabela 5.1: Especificação dos parâmetros do enlace direto do padrão 3GPP-LTE.

Parâmetros	Valores
Largura de banda	10 MHz
Número de pontos da FFT	1.024
Número de subportadoras utilizadas para transmissão de informação	600
Espaçamento entre subportadoras	15 kHz
Comprimento do prefixo cíclico	72 amostras (curto)
Símbolos OFDM por quadro de rádio	140
Modulação	QPSK, 16QAM

O diagrama de blocos dos sistemas receptores está mostrado de forma simplificada na Fig. 5.1 (b). A partir do sinal recebido e da posição dos símbolos pilotos presentes no quadro LTE, é realizada a estimação da resposta do canal por meio do algoritmo LMMSE (do inglês, *linear minimum mean square error*), apresentado em [128, 129, 130]. A partir da resposta estimada, é realizado o processo de predição da resposta de canal. O resultado do algoritmo de predição é então enviado ao transmissor por meio de um canal de retorno (*feedback*) analógico, [131], isto é, a informação enviada por ele não é quantizada. O transmissor utilizará então esta informação para realizar a pré-codificação para o novo conjunto de símbolos.



(a)



(b)

Figura 5.1: Ambiente utilizado nas simulações, destacando (a) o MIMO BC e a presença do canal de *feedback* e (b) a estrutura do sistema receptor.

Para a simulação do canal rádio móvel, foi utilizado o modelo padronizado pelo 3GPP, denominado *Spatial Channel Model* (SCM) [127]. O modelo foi implementado utilizando um conjunto de funções em MATLAB denominado WINNER's SCM [132]. Apesar de vários modelos de canal serem sugeridos na literatura, como, por exemplo, o modelo de Jakes [133], optou-se por um modelo um pouco mais sofisticado e realista, com o objetivo de obter o comportamento variante no tempo do canal o mais próximo possível de um canal real. Os parâmetros do modelo SCM que foram utilizados nas simulações estão mostrados na Tabela 5.2, e uma breve descrição das características do modelo SCM é realizada no Anexo II. Não foi utilizado nenhum modelo para a perda de percurso, de forma que todos os usuários apresentam a mesma potência média de recepção.

Tabela 5.2: Especificação dos valores de parâmetros para o modelo de canal WINNER's SCM.

Parâmetros	Valores
Frequência central da portadora	2 GHz
Velocidade do móvel	5 m/s, 30 m/s
Número de antenas na estação rádio-base	4
Número de antenas na estação móvel	1
Ambiente	Microurbano (<i>Urban Micro</i>)
Número de multipercursos	19

5.3 DESEMPENHOS DOS PREDITORES DE CANAL

Nesta seção, será considerado o desempenho dos preditores NLMS, RLS e SM-AP. As curvas de aprendizado, mostradas nas Fig. 5.2 a 5.5, foram obtidas pela média de 1.000 experimentos, cada um consistindo de 5 quadros de rádio LTE, totalizando 700 símbolos OFDM por experimento.

Os valores utilizados para os diferentes parâmetros do canal estão especificados na Tabela 5.2 e foram utilizados em todas as simulações, a menos que seja indicado o contrário. A velocidade dos usuários foi de 30 m/s, e a razão sinal-ruído foi fixada em 20 dB (para fins de estimação de canal).

O número de coeficientes para os preditores foi $N_c = 10$ em todos os algoritmos, e o número de *constraint sets* reutilizados pelo SM-AP foi $N_P = 3$. O algoritmo RLS utilizou fator de esquecimento $\lambda = 0,99$ e o passo de adaptação do algoritmo NLMS foi $\mu = 0,5$. Estes valores foram escolhidos pois encontram-se relatados em [122]. A escolha do limiar ν para definição dos *constraint sets* do algoritmo SM-AP, conforme definido no capítulo 4, foi $\nu = 2\sqrt{\sigma_n^2}$ [112].

A Fig. 5.2 mostra as curvas de aprendizagem dos preditores para um horizonte de predição ℓ de um símbolo OFDM. Para permitir uma comparação justa entre os algoritmos, também é mostrado o desempenho do algoritmo RLS para $N_c = 3$ coeficientes. Essa escolha resulta em uma complexidade computacional da mesma ordem de grandeza da do SM-AP com $N_P = 3$. O NLMS apresentou a menor velocidade de convergência (50 símbolos OFDM) e o maior erro quadrático médio de predição, de aproximadamente -28 dB. A convergência do RLS com $N_c = 10$ e do SM-AP foram similares, apesar de o RLS exibir MSE ligeiramente inferior ao SM-AP. Comparando o RLS com $N_c = 3$ e do SM-AP, este exibiu menor tempo de convergência que aquele.

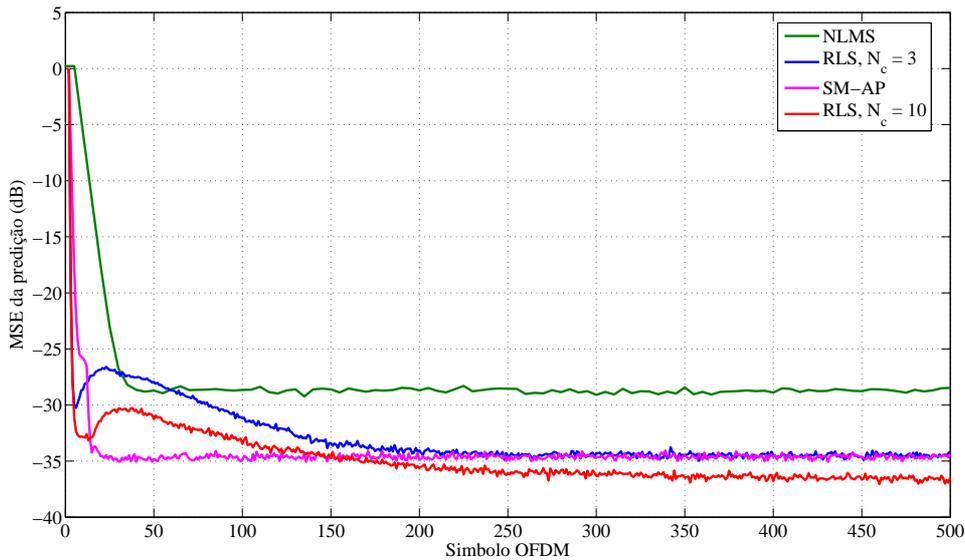


Figura 5.2: Convergência dos preditores adaptativos para um horizonte de predição de 1 símbolo OFDM.

A Fig. 5.3 mostra as curvas de aprendizagem dos preditores para um horizonte de predição ℓ de 28 símbolos OFDM. Duas opções foram consideradas para o valor do fator de esquecimento: $\lambda = 0,99$ e $\lambda = 0,6$. Como é possível perceber, esta escolha pode ser crítica para o desempenho do algoritmo RLS, tanto em termos do MSE de predição como da velocidade de convergência. Valores maiores para o fator de esquecimento podem degradar severamente o desempenho do RLS, especialmente para valores maiores de horizonte de predição.

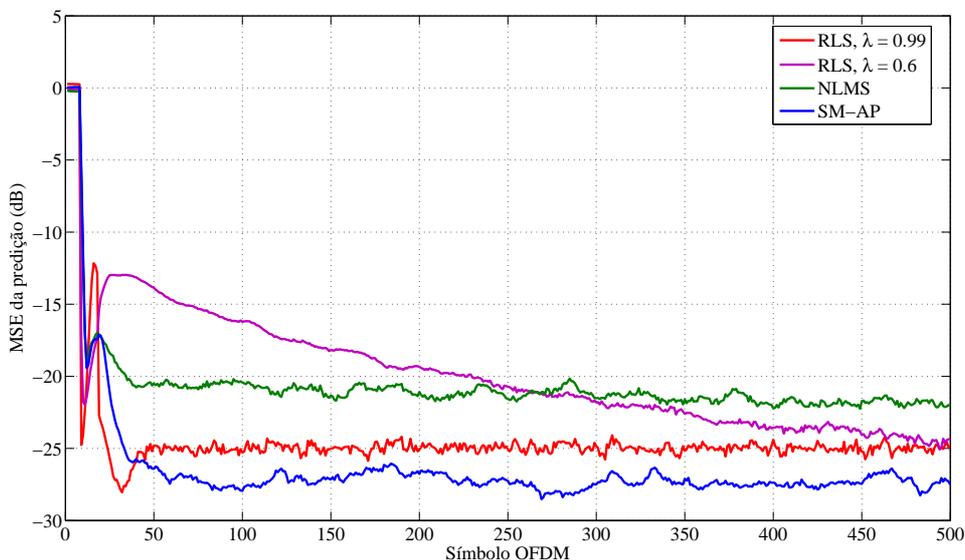


Figura 5.3: Convergência dos preditores adaptativos para um horizonte de predição de 28 símbolos OFDM.

A Fig. 5.4 mostra o comportamento do MSE, após a convergência dos coeficientes dos preditores, em função do horizonte de predição. O SM-AP apresenta desempenho superior a ambos os algoritmos de predição, e seu MSE é aproximadamente 1dB inferior ao apresentado pelo RLS utilizando fator de

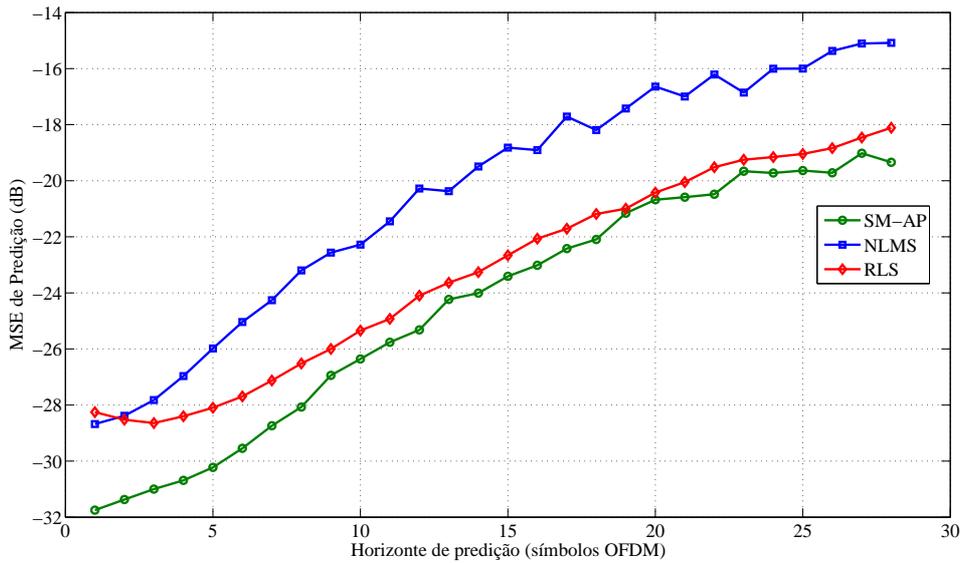


Figura 5.4: MSE de Predição versus horizonte de predição.

esquecimento $\lambda = 0,6$.

Na Fig. 5.5, a capacidade de rastreamento dos preditores adaptativos em um canal de comportamento não estacionário é investigada. Foi considerado um ambiente não estacionário onde as características do canal mudam a cada 100 símbolos OFDM. Um desvio Doppler de 50 Hz é considerado nos primeiros 100 símbolos OFDM. Em seguida, este valor é alterado para 100 Hz e mantido por mais 100 símbolos. Os 100 símbolos seguintes apresentam desvio Doppler de 150 Hz e os últimos 100 símbolos foram simulados utilizando um desvio Doppler de 200 Hz. A convergência mais lenta do algoritmo RLS é justificada pela influência do fator de esquecimento. O preditor baseado no algoritmo NLMS exibiu erro quadrático médio aproximadamente 5 dB maior que o apresentado pelo SM-AP.

De forma geral, os preditores RLS e SM-AP apresentaram desempenho semelhante para o horizonte de predição de um símbolo OFDM, sendo que o preditor SM-AP apresenta menor custo computacional. Para horizontes de predição maiores que um símbolo OFDM, o preditor SM-AP apresenta melhor desempenho que o RLS. Isto se deve ao fato de que o algoritmo RLS não é robusto com respeito a escolha do fator de esquecimento. Como consequência, seu desempenho degrada-se mais rapidamente, sobretudo em canais variantes no tempo.

5.4 DESEMPENHO DOS PRÉ-CODIFICADORES

Nesta seção, será investigado o comportamento de três algoritmos de pré-codificação: *zero-forcing*, *vector perturbation*, e o pré-codificador proposto (baseado no conceito de perturbação contínua). As especificações do sistema estão listadas na Tabela 5.1.

Serão mostrados resultados para quatro cenários diferentes, apresentados na Tabela 5.3. Os cenários 1 e 2 representam um ambiente de usuários de baixa mobilidade, e os cenários 3 e 4 representam um ambiente

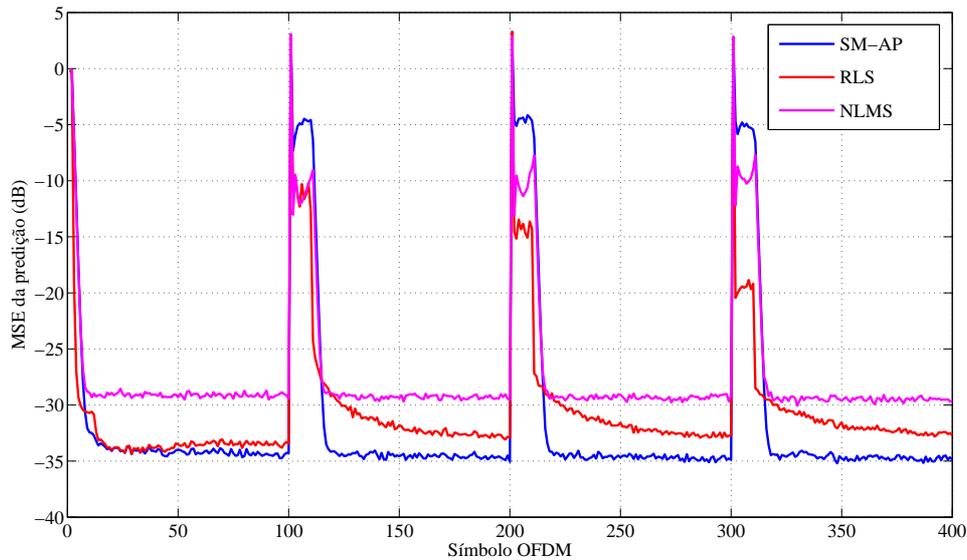


Figura 5.5: Capacidade de rastreamento dos preditores SM-AP, NLMS e RLS em um ambiente não estacionário. As características do canal mudam a cada 100 símbolos OFDM.

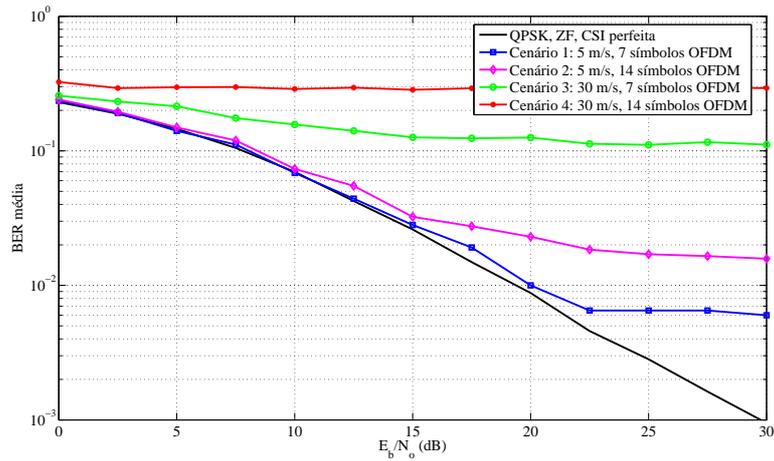
cujos usuários apresentam alta mobilidade. São considerados dois valores para o atraso introduzido pelo canal de retorno: 7 símbolos OFDM (que correspondem a um *slot* no padrão 3GPP-LTE) e 14 símbolos OFDM (que correspondem ao intervalo de um subquadro no padrão 3GPP-LTE). O atraso de 1 ms (14 símbolos OFDM) corresponde a um valor mais realista de atraso no canal de retorno, visto que corresponde ao intervalo de tempo de transmissão, ou TTI (do inglês, *transmission time interval*) utilizado pelo padrão LTE. O TTI corresponde ao período em que são tomadas as decisões quanto ao escalonamento dos usuários e os blocos de transporte são transferidos pela camada física da interface de rádio.

Tabela 5.3: Cenários para a simulação do desempenho dos pré-codificadores.

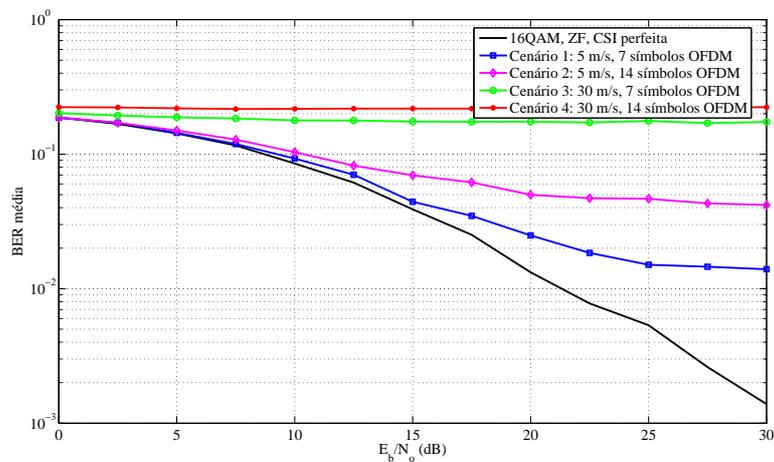
Cenário	Velocidade do usuário	Atraso do canal de retorno
1	5 m/s	7 símbolos OFDM
2	5 m/s	14 símbolos OFDM
3	30 m/s	7 símbolos OFDM
4	30 m/s	14 símbolos OFDM

5.4.1 Resultados de Referência

As Fig. 5.6, 5.7 e 5.8 mostram o desempenho dos algoritmos de pré-codificação *zero-forcing*, *vector perturbation* e o pré-codificador proposto, em cada um dos quatro cenários definidos na Tabela 5.3, para as modulações QPSK e 16QAM, sem a utilização de um preditor de canal para compensar os atrasos introduzidos pelo canal de retorno. Nestes casos, a pré-codificação é realizada no transmissor utilizando informação desatualizada sobre o estado do canal. Mais especificamente, a pré-codificação é realizada utilizando a resposta do canal atrasada por um período igual ao atraso introduzido pelo canal de retorno (7 símbolos ou 14 símbolos, de acordo com o cenário).



(a) Modulação QPSK



(b) Modulação 16QAM

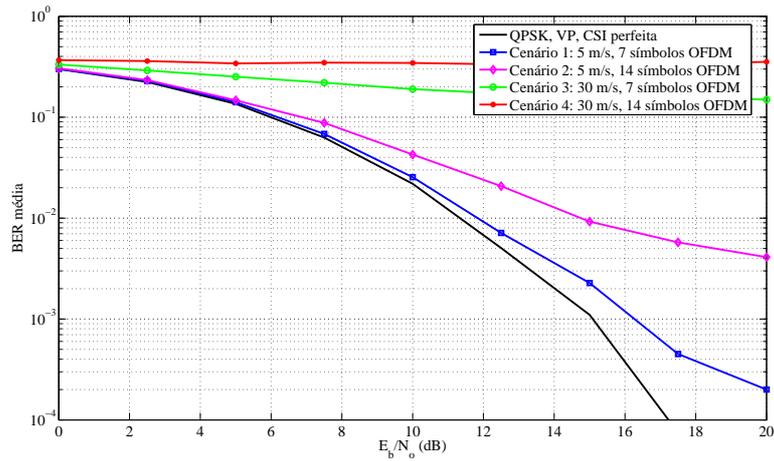
Figura 5.6: Desempenho do pré-codificador *zero-forcing* sem a utilização de um preditor de canal.

Serão analisados inicialmente os cenários 1 e 2, de baixa mobilidade. A modulação QPSK se mostra mais robusta que a 16QAM, para todos os algoritmos de pré-codificação considerados. O pré-codificador proposto apresenta ganho de 3 dB quando comparado ao algoritmo *zero-forcing* para a modulação QPSK e 1dB na modulação 16QAM, considerando a presença de CSI perfeita no transmissor.

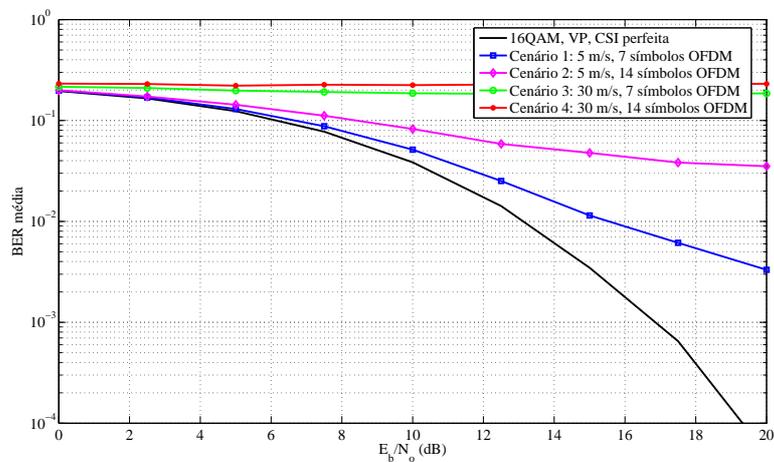
É interessante perceber que, nos cenários de baixa mobilidade, o desempenho dos pré-codificadores não sofre degradação muito grande (de 1 dB a 3 dB) para valores de $\frac{E_b}{N_o}$ até 15 dB. Nestas situações, o uso de um preditor de canal poderia ser dispensado.

Este fato é justificado ao observar o tempo de coerência do canal em cada um dos cenários. De acordo com os resultados apresentados em [134], para a velocidade de 5 m/s, o tempo de coerência do canal é de aproximadamente 2,2 ms. Isto significa que durante o período de 28 símbolos OFDM, a resposta do canal apresenta valor de correlação temporal alto (superior a 0,9), e a informação sobre o estado do canal leva um número maior de símbolos para se tornar desatualizada.

A partir do valor de $\frac{E_b}{N_o} = 15$ dB, tanto o pré-codificador proposto como o pré-codificador *zero-forcing*



(a) Modulação QPSK



(b) Modulação 16QAM

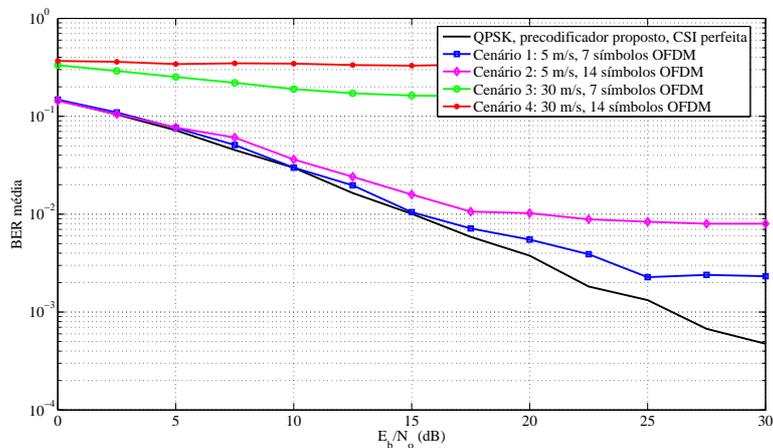
Figura 5.7: Desempenho do pré-codificador *vector perturbation* sem a utilização de um preditor de canal.

passam a apresentar saturação no valor da BER. Entretanto, o *error floor* exibido pelo pré-codificador proposto é menor que o apresentado pelo pré-codificador do tipo *zero-forcing* para a modulação QPSK.

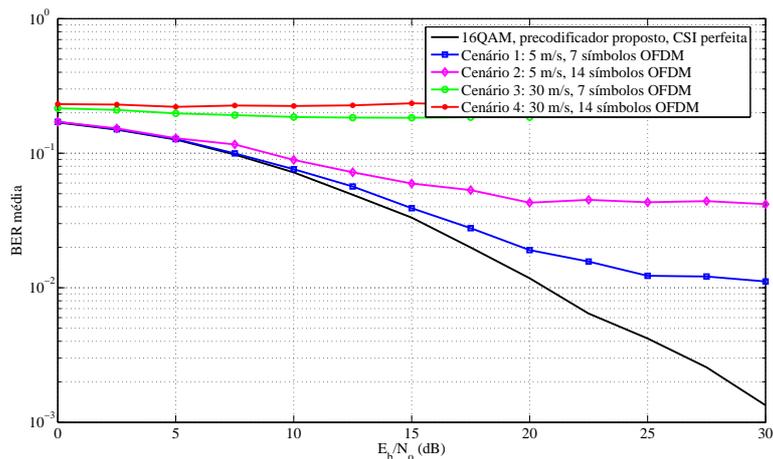
A técnica de pré-codificação *vector perturbation* apresenta a menor degradação no desempenho da BER quando comparada as demais técnicas consideradas, conforme mostra a Fig. 5.7. Entretanto, para a modulação 16QAM e um atraso de 14 símbolos OFDM, o pré-codificador *vector perturbation* também começa a exibir este comportamento de saturação de BER em um ambiente de baixa mobilidade, mas ele ocorre para valores mais elevados de relação sinal-ruído. A Fig. 5.7 (b) mostra que o *error floor* de $3 \cdot 10^{-2}$ surge para valores de $\frac{E_b}{N_o}$ superiores a 18 dB.

Nos cenários 1 e 2, cogitou-se que a utilização de um preditor de canal poderia ser dispensada para alguns valores de relação sinal-ruído, já que a degradação que é introduzida não é muito significativa. O mesmo já não pode ser dito ao tratar de cenários que possuem usuários de alta mobilidade (cenários 3 e 4). Todos os algoritmos de pré-codificação têm seu desempenho severamente comprometido, resultando em limiares altos de saturação de BER, da ordem de 10^{-1} , mesmo para baixos valores de relação sinal-ruído, conforme mostram as Fig. 5.6, 5.7 e 5.8.

Este resultado também pode ser explicado por meio das ferramentas apresentadas em [134]. Considerando a velocidade de 30 m/s, o tempo de coerência do canal é aproximadamente 0,323 ms, o que equivale a 4 símbolos OFDM no enlace direto do padrão LTE. Logo, mesmo para um atraso do canal de retorno de 7 símbolos OFDM, a informação sobre o estado do canal já estará consideravelmente desatualizada, de forma a interferir severamente no desempenho das técnicas de pré-codificação.



(a) Modulação QPSK



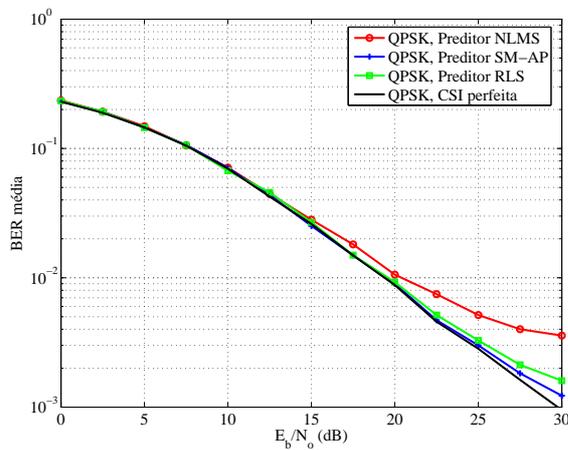
(b) Modulação 16QAM

Figura 5.8: Desempenho do pré-codificador proposto sem a utilização de um preditor de canal.

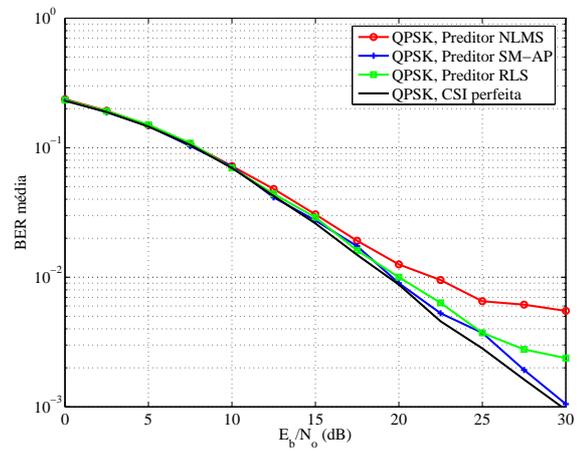
5.4.2 Cenários com preditores de canal

Nesta seção, será mostrado o desempenho dos algoritmos de pré-codificação em operação conjunta com os algoritmos de previsão de canal, sendo estes utilizados para compensar o atraso introduzido pelo canal de retorno, isto é, mitigar os efeitos da informação desatualizada sobre o estado do canal, conforme foi mostrado na Fig. 5.1 (b). Nos cenários 1 e 3, foi adotado um horizonte de previsão de $\ell = 7$ símbolos OFDM, e nos cenários 2 e 4, escolheu-se $\ell = 14$ símbolos OFDM, que correspondem aos valores de atraso introduzidos pelo canal de retorno em cada um dos cenários.

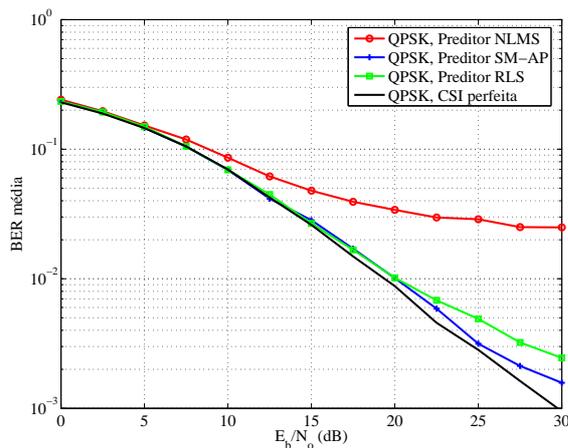
Os resultados ao utilizar a modulação QPSK estão mostrados nas Fig. 5.9, 5.11 e 5.13 para os algoritmos de pré-codificação *zero-forcing*, *vector perturbation* e o algoritmo proposto, respectivamente, e nas Fig. 5.10, 5.12 e 5.14 para a modulação 16QAM. De forma geral, pode-se afirmar que os algoritmos de previsão são capazes de melhorar o desempenho dos pré-codificadores quando estes operam em um ambiente cuja informação sobre o estado do canal não é perfeita, principalmente nos cenários 3 e 4, que se referem a usuários de alta mobilidade. Basta comparar os resultados obtidos com seus respectivos cenários de referência, mostrados na seção 5.4.1.



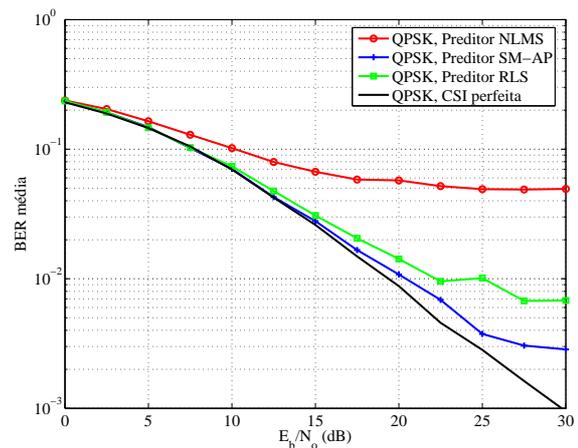
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso



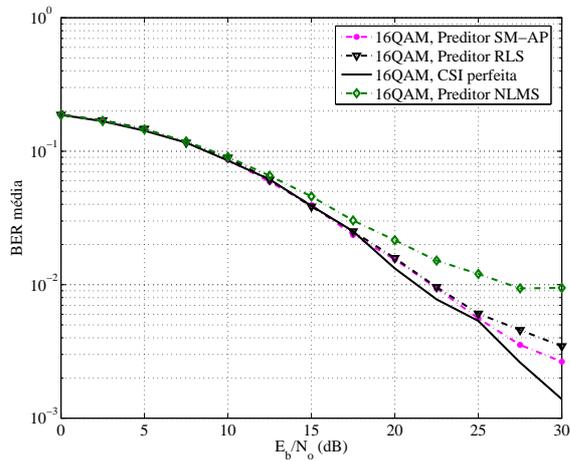
(c) Cenário 3: 30 m/s - 7 símbolos de atraso



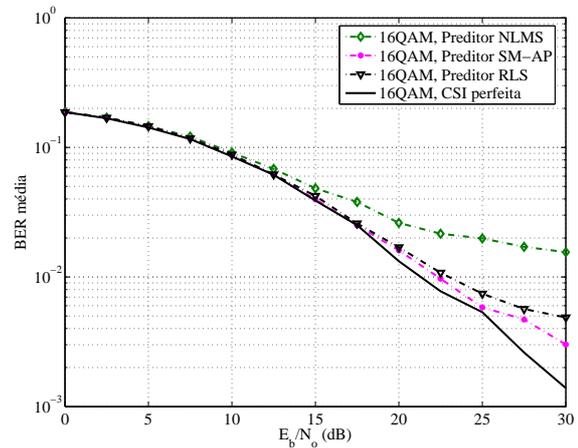
(d) Cenário 4: 30 m/s - 14 símbolos de atraso

Figura 5.9: Desempenho do algoritmo de pré-codificação *zero-forcing* utilizando a modulação QPSK.

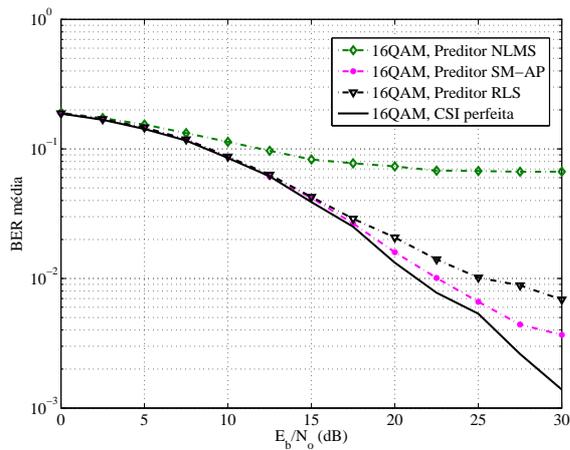
O algoritmo NLMS exibe o pior desempenho em todos os cenários e técnicas de modulação e pré-codificação analisados. Para a pré-codificação *zero-forcing* operando com modulação QPSK, o sistema exibe comportamento de saturação no valor da BER para horizontes de predição a partir de 7 símbolos OFDM mesmo no cenário de baixa mobilidade, conforme mostrado na Fig. 5.9. Ao considerar os cenários 3 e 4, de alta mobilidade, o *error floor* tende a aumentar. Comportamento semelhante é observado para a modulação 16QAM, exibida na Fig. 5.10, para a qual o valor de saturação da BER é superior aos casos considerados anteriormente, pois esta modulação é menos robusta à CSI imperfeita.



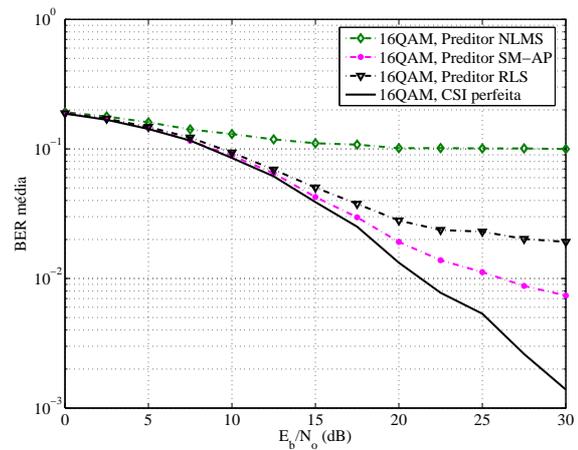
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso

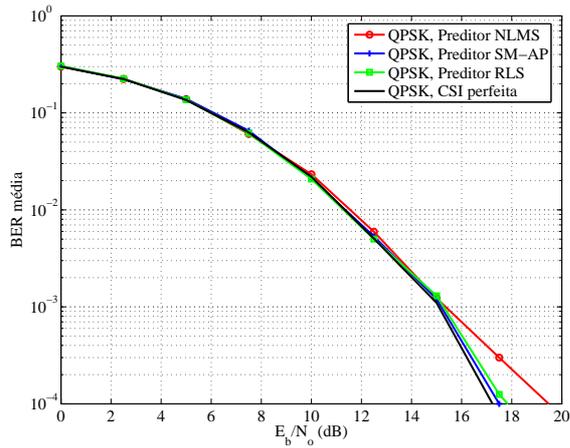


(c) Cenário 3: 30 m/s - 7 símbolos de atraso

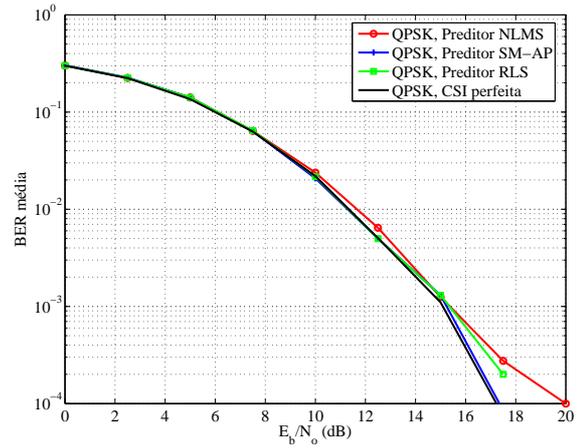


(d) Cenário 4: 30 m/s - 14 símbolos de atraso

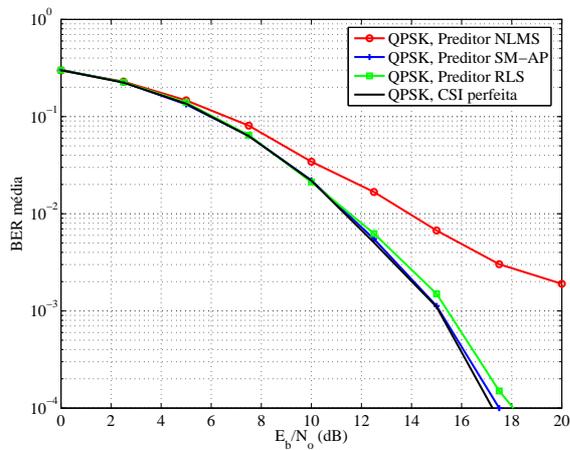
Figura 5.10: Desempenho do algoritmo de pré-codificação *zero-forcing* utilizando a modulação 16QAM.



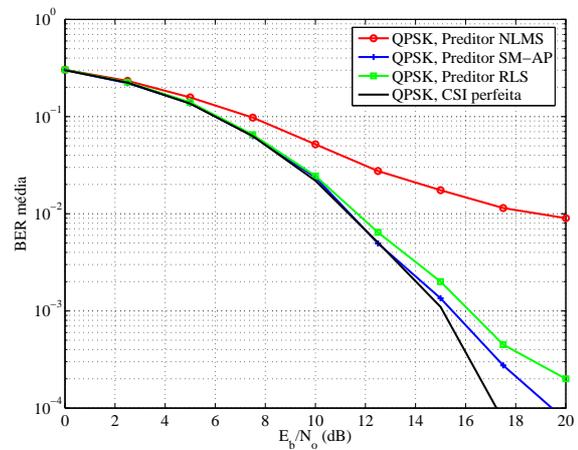
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso



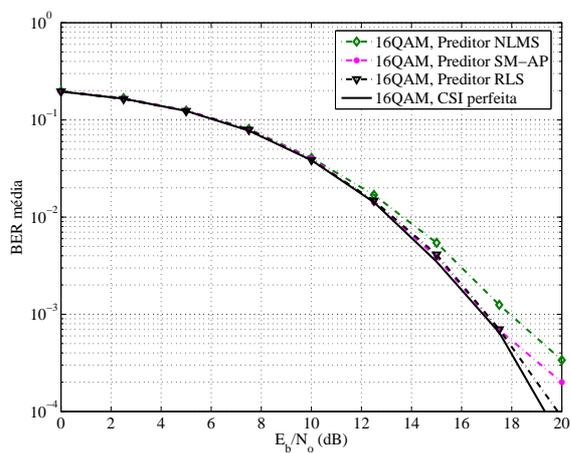
(c) Cenário 3: 30 m/s - 7 símbolos de atraso



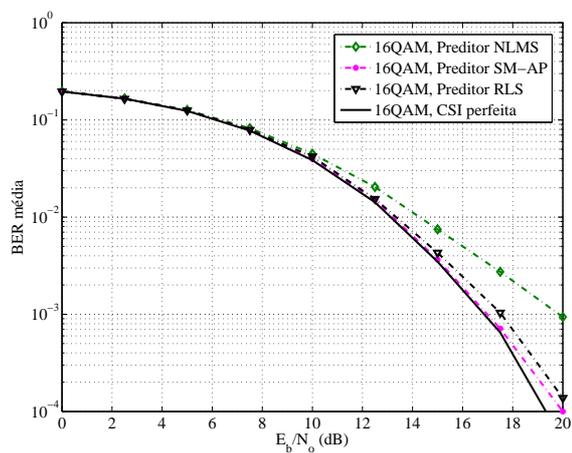
(d) Cenário 4: 30 m/s - 14 símbolos de atraso

Figura 5.11: Desempenho do algoritmo de pré-codificação *vector perturbation* utilizando a modulação QPSK.

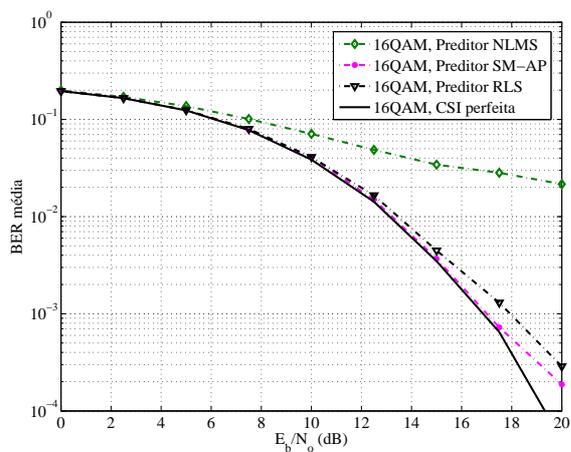
Conforme mostra a Fig. 5.11, a técnica de pré-codificação *vector perturbation* é menos sensível aos erros de predição ao utilizar a modulação QPSK. Para os cenários 1 e 2, todos os algoritmos adaptativos exibem comportamento semelhante. Na região de relação sinal-ruído mais elevada, o desempenho do NLMS começa a exibir degradação, e, nos cenários 3 e 4, o comportamento da saturação da BER fica evidente, o que não ocorre com o RLS e o SM-AP. Entretanto, ao utilizar a modulação 16QAM, mostrado na Fig. 5.12, o algoritmo NLMS exibe *error floor* muito elevado, e sua utilização não é indicada nesta situação. Os algoritmos RLS e SM-AP exibem comportamento semelhante, entretanto a superioridade do SM-AP é revelada nos cenários de maior mobilidade [135, 136].



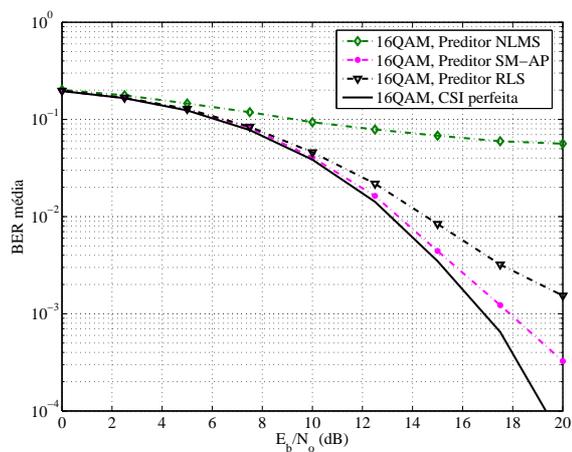
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso



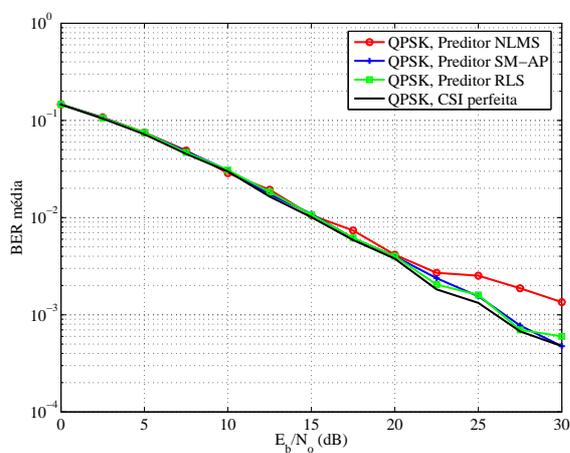
(c) Cenário 3: 30 m/s - 7 símbolos de atraso



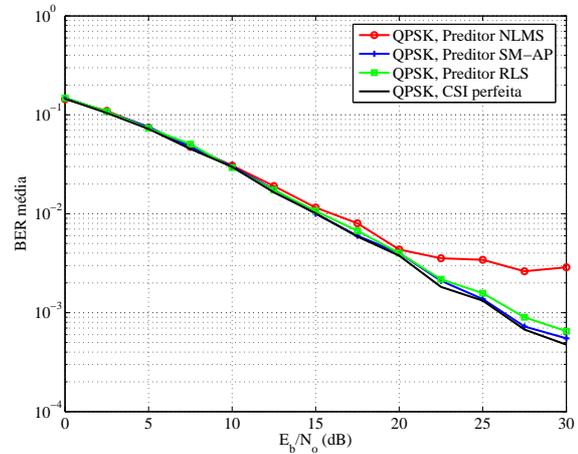
(d) Cenário 4: 30 m/s - 14 símbolos de atraso

Figura 5.12: Desempenho do algoritmo de pré-codificação *vector perturbation* utilizando a modulação 16QAM.

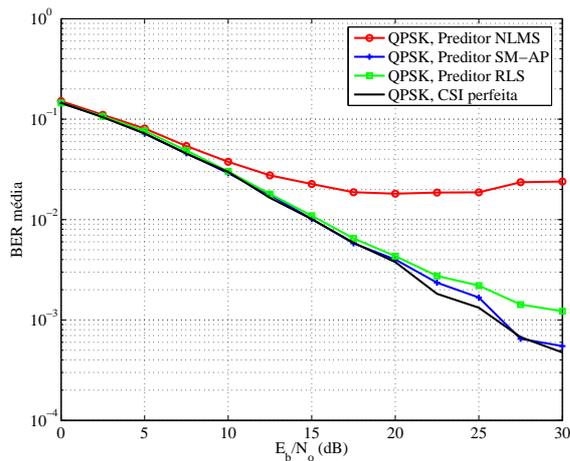
O algoritmo de pré-codificação proposto é superior ao *zero-forcing* em todos os cenários utilizando a modulação QPSK, de acordo com os resultados exibidos pela Fig. 5.13. Para a modulação 16QAM, cujos resultados são mostrados na Fig. 5.14, o desempenho da técnica de pré-codificação proposta é ligeiramente superior ao da técnica *zero-forcing*. Entretanto, ele não é tão robusto à CSI desatualizada quanto o pré-codificador *vector perturbation*.



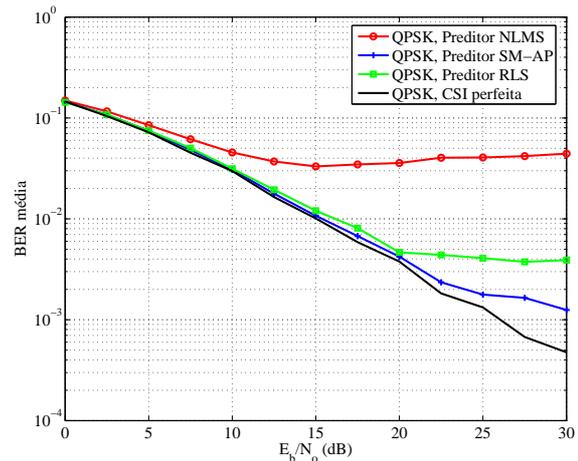
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso

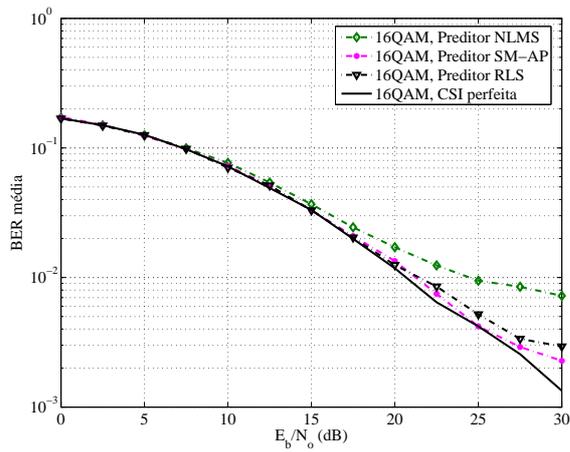


(c) Cenário 3: 30 m/s - 7 símbolos de atraso

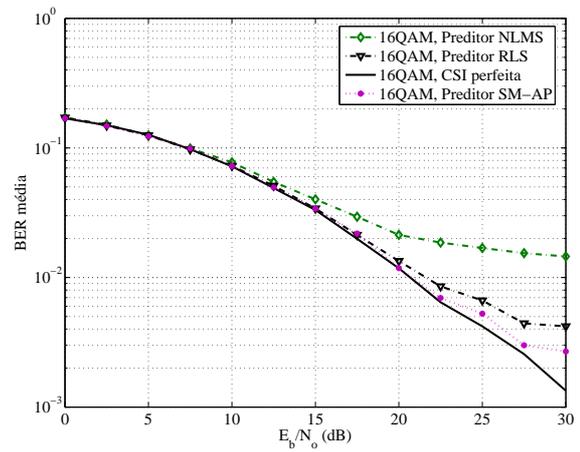


(d) Cenário 4: 30 m/s - 14 símbolos de atraso

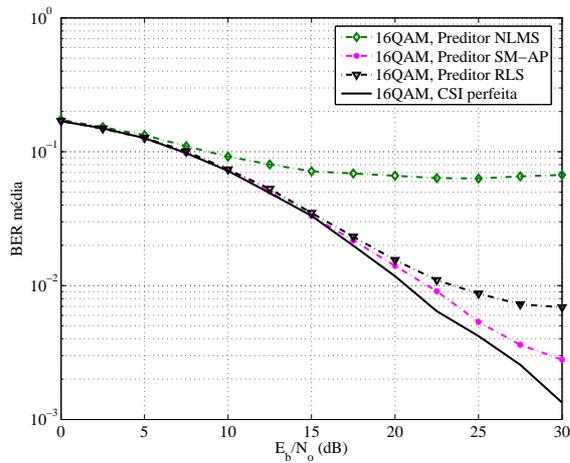
Figura 5.13: Desempenho do algoritmo de pré-codificação proposto utilizando a modulação QPSK.



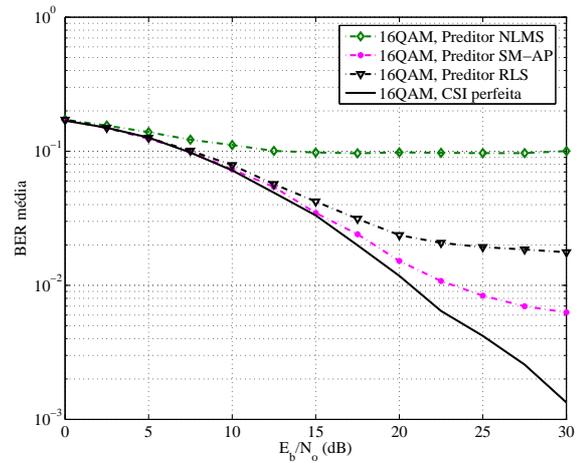
(a) Cenário 1: 5 m/s - 7 símbolos de atraso



(b) Cenário 2: 5 m/s - 14 símbolos de atraso



(c) Cenário 3: 30 m/s - 7 símbolos de atraso



(d) Cenário 4: 30 m/s - 14 símbolos de atraso

Figura 5.14: Desempenho do algoritmo de pré-codificação proposto utilizando a modulação 16QAM.

Como última análise, as Fig. 5.15 a 5.17 mostram o desempenho do preditor baseado no algoritmo RLS para a escolha $N_c = 3$ coeficientes em um ambiente de alta mobilidade (cenário 4). Esta escolha foi feita pois, neste caso, a complexidade do algoritmo RLS apresenta a mesma ordem da grandeza da apresentada pelo SM-AP com $N_P = 3$ *constraint sets* reutilizados. As figuras mostram que o SM-AP possui desempenho superior ao algoritmo RLS quando ambos apresentam aproximadamente o mesmo custo computacional. Por conveniência, os resultados obtidos anteriormente para os filtros preditores também são mostrados.

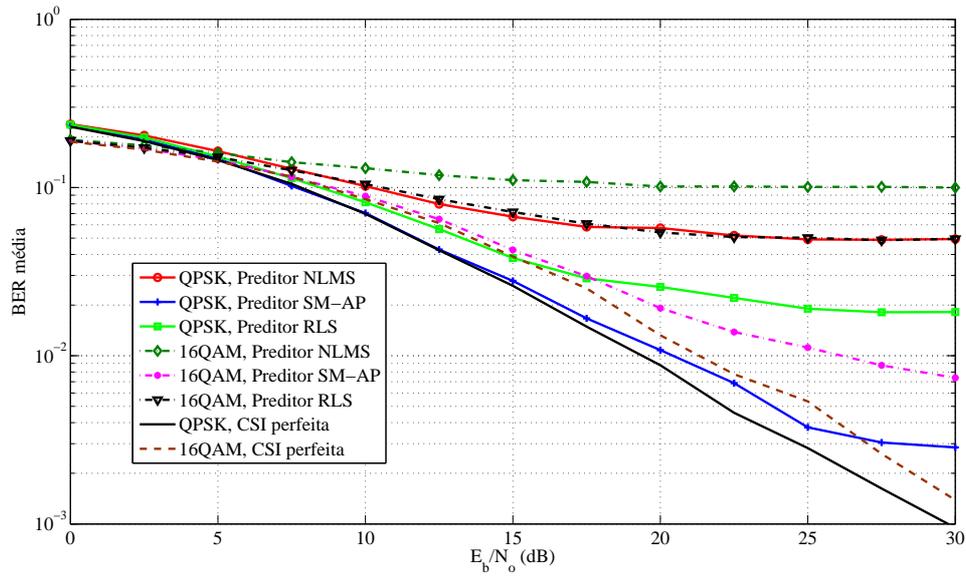


Figura 5.15: Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando a técnica de pré-codificação *zero-forcing* e as modulações QPSK e 16QAM.

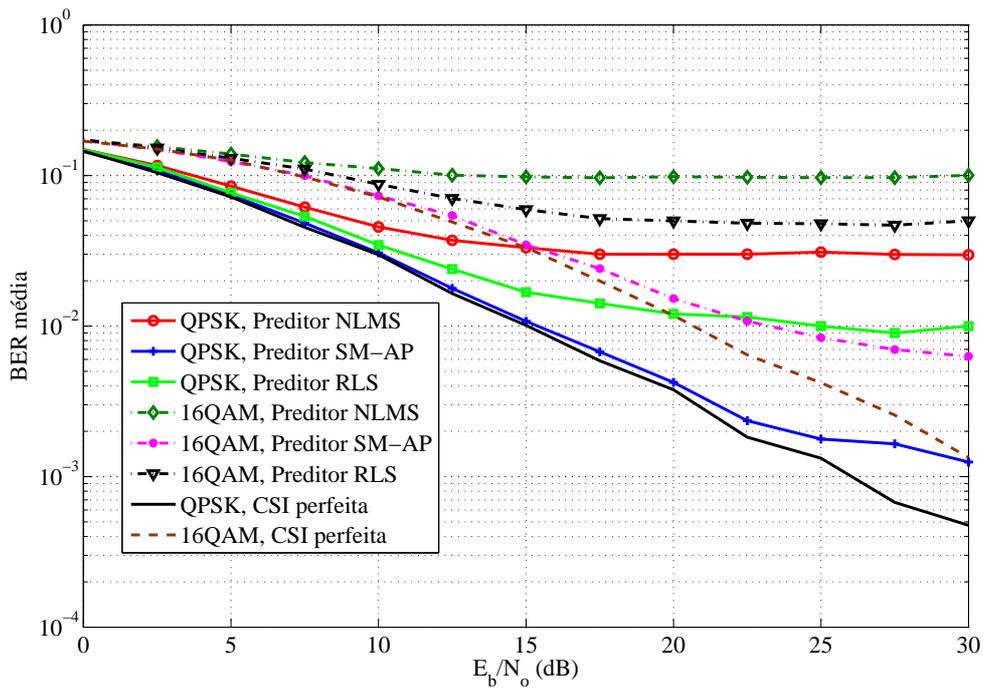


Figura 5.16: Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando o algoritmo de pré-codificação proposto e as modulações QPSK e 16QAM.

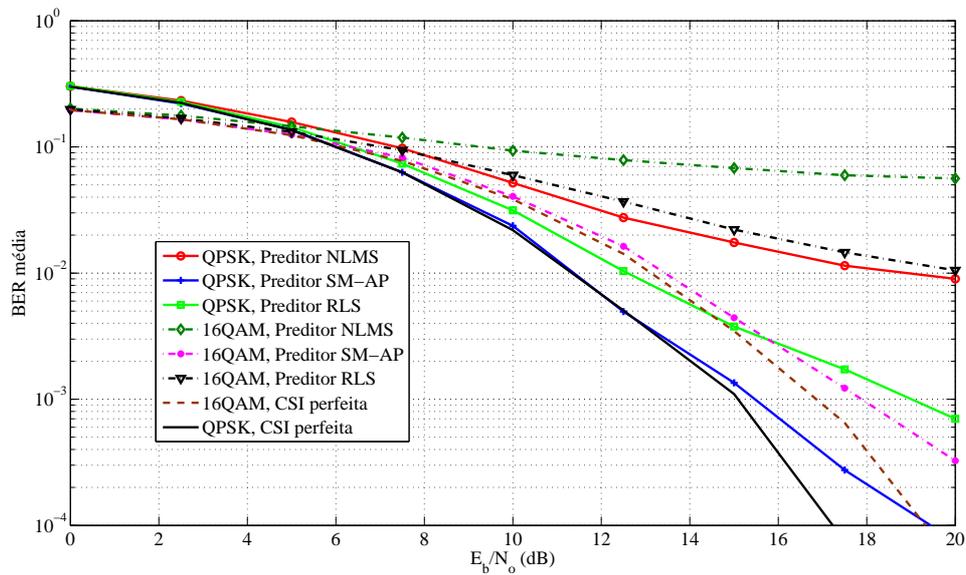


Figura 5.17: Desempenho do preditor RLS com $N_c = 3$ coeficientes, no cenário 4 (velocidade de 30 m/s e atraso de 14 símbolos OFDM), utilizando a técnica de pré-codificação *vector perturbation* e as modulações QPSK e 16QAM.

5.5 CONCLUSÃO

Neste capítulo, o desempenho das técnicas de pré-codificação e filtragem adaptativa, apresentadas nos capítulos 3 e 4, foram avaliadas por meio de simulações em nível de enlace utilizando as especificações do enlace direto do padrão 3GPP-LTE e um modelo realista de canal rádio móvel, padronizado pelo 3GPP para simulações de sistemas MIMO.

As simulações realizadas permitem inferir que, para usuários de alta mobilidade, a utilização de esquemas de pré-codificação é inviável sem a operação conjunta com algoritmos de predição de canal. A informação atualizada sobre o estado do canal é de fundamental importância para que as técnicas de pré-codificação possam exibir bom desempenho.

Os resultados ainda mostraram que o preditor de canal baseado no algoritmo NLMS exibiu o pior desempenho em todos os cenários de simulação considerados. Para horizontes de predição pequenos, o algoritmo SM-AP apresentou desempenho semelhante ao RLS, entretanto este não se mostrou robusto quanto à escolha do fator esquecimento. A superioridade do algoritmo SM-AP foi revelada em cenários que possuíam usuários de alta mobilidade e eram exigidos valores maiores de horizonte de predição. O preditor SM-AP, além de melhor desempenho, apresentou custo computacional menor que o do RLS.

6 CONCLUSÕES

Este trabalho abordou o desempenho de algoritmos de pré-codificação utilizados no enlace direto de sistemas MIMO multiusuário sob a influência de informação imperfeita sobre o estado de canal.

O capítulo 1 apresentou as motivações do trabalho. Para que os sistemas MIMO multiusuário possam tirar proveito das vantagens oferecidas pela presença das múltiplas antenas, é necessária a presença de informação sobre o estado do canal no transmissor. Entretanto, devido à natureza variante no tempo do canal rádio móvel e aos atrasos presentes nos sistemas de comunicação, a CSI pode estar desatualizada ao realizar a pré-codificação. Nestes casos, algoritmos de previsão de canal podem ser utilizados para compensar o efeito negativo desses atrasos no desempenho dos sistemas.

O capítulo 2 apresentou resultados referentes à capacidade do enlace direto e do enlace reverso de sistemas MIMO multiusuário. Foi mostrado que, no enlace direto, a capacidade é alcançada ao se utilizar pré-codificação do tipo *dirty paper*. Por se tratar de um esquema muito difícil de ser implementado, o capítulo 3 tratou das técnicas lineares e não lineares de pré-codificação utilizadas para cancelamento de interferência.

Foi proposto um método pré-codificação linear que utiliza um vetor de perturbação contínua como forma de melhorar o desempenho das técnicas de pré-codificação lineares apresentadas. Os resultados de simulação mostraram que o pré-codificador proposto apresenta desempenho superior ao das técnicas lineares atualmente descritas na literatura. Por não ser baseado no conceito de DPC, o pré-codificador proposto apresenta pior desempenho que as técnicas não lineares Tomlinson-Harashima e *vector perturbation*. Entretanto, diferentemente da técnica *vector perturbation*, que utiliza um vetor de perturbação discreto, e para a qual é necessária a utilização de algoritmos de custo computacional bastante elevado para a determinação do vetor de perturbação, a perturbação contínua pode ser determinada de forma analítica, e com custo computacional consideravelmente menor.

O capítulo 4 introduziu os conceitos que são frequentemente utilizados em filtragem adaptativa, e foram descritos os algoritmos clássicos LMS, NLMS, RLS e AP, além do SM-AP. Foi ainda proposta uma estrutura de preditor adaptativo de canal para sistemas MIMO-OFDM multiusuário que utiliza o algoritmo *set-membership affine projection*.

No capítulo 5, foram apresentados resultados de simulações do desempenho dos pré-codificadores e preditores de canal em um cenário mais realista, que utilizava um modelo de canal padronizado pelo 3GPP para simulações MIMO, operava sob as especificações do enlace direto do padrão 3GPP-LTE e incluía os efeitos da CSI desatualizada no transmissor. Foi possível inferir que, para usuários de alta mobilidade, o desempenho dos esquemas de pré-codificação é muito prejudicado pela informação desatualizada sobre o estado de canal, mesmo para pequenos horizontes de previsão, e a utilização de técnicas de pré-codificação é inviável sem a operação conjunta com preditores de canal.

O preditor de canal proposto, que utiliza o algoritmo adaptativo SM-AP, apresentou desempenho semelhante ao do RLS para usuários que apresentam baixa mobilidade, entretanto este não apresentou desempenho satisfatório em ambientes cujos usuários apresentam alta mobilidade. Nestes cenários, a superioridade do algoritmo SM-AP foi revelada, especialmente para horizontes de previsão mais amplos.

Além disso, ao escolher os parâmetros dos filtros preditores de tal forma que os algoritmos SM-AP e RLS possuíssem complexidade computacional da mesma ordem de grandeza, este apresentou desempenho inferior àquele.

6.1 TRABALHOS FUTUROS

Diversos aspectos podem ser explorados e investigados na área de comunicação MIMO multiusuário. Algumas sugestões para a extensão deste trabalho são:

- estudo e implementação de técnicas de pré-codificação para canais seletivos no tempo ou na frequência, como as apresentadas em [137], assim como a investigação do efeito da CSI desatualizada sobre o desempenho dessas técnicas;
- a investigação e a proposta de outros filtros e estimadores estocásticos, como, por exemplo, o filtro de Kalman, para realizar a predição da resposta do canal em sistemas MIMO multiusuário, e sua influência na operação dos pré-codificadores apresentados;
- a utilização de outros algoritmos do tipo SM-AP para realizar a predição de canal como, por exemplo, os algoritmos *set-membership affine projection with variable data reuse* (SM-APVDR) [138] e *set-membership partial-update affine projection* (SM-PUAP) [139], que buscam reduzir de forma mais acentuada a complexidade computacional do *set-membership affine projection* por meio da utilização de um número variável de projeções (fator de reúso dos dados) ou por meio da atualização de apenas uma parte dos coeficientes do filtro preditor;
- a investigação de como as imperfeições de rádio frequência (como, por exemplo ruído de fase, amplificadores não lineares e acoplamento indutivo entre as múltiplas antenas) [140] afetam o desempenho das técnicas de pré-codificação apresentadas;
- verificar o desempenho das técnicas de pré-codificação ao operar sobre um canal de realimentação digital, no qual a informação sobre o estado do canal é quantizada antes de ser enviada ao transmissor e sofre a influência de um canal de retorno não ideal;
- considerar a influência de CSI imperfeita nas técnicas de redução de *feedback* do canal de realimentação e nos algoritmos de seleção de usuários, como, por exemplo, o *framework* proposto e apresentado em [141];
- estudo e simulação, em nível sistêmico, de técnicas que explorem a diversidade multiusuário no enlace direto de sistemas MIMO multiusuário [142], como os algoritmos propostos em [143, 144], bem como a influência da CSI desatualizada nesses sistemas.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] GOLDSMITH, A. *Wireless Communications*. New York, NY: Cambridge University Press, 2005.
- [2] ITU. *Recommendation ITU-R M.1645: Vision, framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000*. [S.l.].
- [3] GESBERT, D. et al. From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems. *IEEE Journal on Selected Areas in Communications*, v. 21, n. 3, p. 281–302, April 2003.
- [4] GOLDSMITH, A. J. et al. Capacity Limits of MIMO Channels. *IEEE Journal on Selected Areas in Communications*, v. 21, p. 684–702, June 2003.
- [5] GESBERT, D. et al. From Single User to Multiuser Communications: Shifting the MIMO Paradigm. *IEEE Signal Processing Magazine*, v. 24, n. 5, p. 36–46, September 2007.
- [6] JINDAL, N.; GOLDSMITH, A. Dirty Paper Coding vs. TDMA for MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, v. 51, n. 5, p. 1783–1794, May 2004.
- [7] SHARIF, M.; HASSIBI, B. A comparison of Time-sharing, DPC, and Beamforming for MIMO Broadcast Channels with Many Users. *IEEE Transactions on Communications*, v. 55, n. 1, p. 11–15, January 2007.
- [8] COVER, T. M.; THOMAS, J. A. *Elements of Information Theory*. Hoboken, New Jersey: Wiley-Interscience, 2006.
- [9] VERDU, S. *Multiuser Detection*. New York, NY: Cambridge University Press, 1998.
- [10] GUESS, T.; VARANASI, M. K. Multi-user Decision-feedback Receivers for the General Gaussian Multiple-access Channel. *Proceedings of the 34th Allerton Conf. on Commun., Control and Computing*, p. 190–199, October 1996.
- [11] CAIRE, G.; SHAMAI, S. On the Achievable Throughput of a Multi-antenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, v. 43, p. 1691–1706, July 2003.
- [12] WEINGARTEN, H.; STEINBERG, Y.; SHAMAI, S. The Capacity Region of the Gaussian MIMO Broadcast Channel. In: *Proceedings Conference on Information Sciences and Systems (CISS)*. [S.l.: s.n.], 2004.
- [13] YU, W.; CIOFFI, J. Sum capacity of a Gaussian Vector Broadcast Channel. *IEEE Transactions on Information Theory*, v. 50, n. 9, p. 1875–1892, September 2004.
- [14] COSTA, M. H. M. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, v. 29, p. 439–441, May 1983.
- [15] SPENCER, Q. H.; SWINDLEHURST, A. L.; HAARDT, M. Zero-forcing Methods for Downlink Spatial Multiplexing in Multi-user MIMO Channels. *IEEE Transactions on Signal Processing*, v. 52, n. 2, p. 461–471, February 2004.

- [16] PEEL, C. B.; HOCHWALD, B. M.; SWINDLEHURST, A. L. A Vector-Perturbation Technique for Near-capacity Multi-antenna Multi-user Communication - Part II: Perturbation. *IEEE Transactions on Communications*, v. 53, n. 3, p. 537–544, March 2005.
- [17] TU, Z.; BLUM, R. Multiuser Diversity for a Dirty Paper Approach. *IEEE Communications Letters*, v. 7, n. 8, p. 370–372, August 2003.
- [18] BOLCSKEI, H. et al. *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. New York, NY: Cambridge University Press, 2008.
- [19] SPENCERA, Q. H. et al. An Introduction to the Multi-user MIMO Downlink. *IEEE Communications Magazine*, v. 42, n. 10, p. 60–67, October 2004.
- [20] COVER, T. Broadcast Channels. *IEEE Transactions on Information Theory*, v. 18, n. 1, p. 2–14, January 1972.
- [21] SHANNON, C. Two-way Communication Channels. In: *Proceedings of 4th Berkeley Symp. Math. Statist. and Prob.* [S.l.: s.n.], 1961.
- [22] TSE, D.; VISWANATH, P. *Fundamentals of Wireless Communication*. New York, NY: Cambridge University Press, 2005.
- [23] ANDREWS, J. G.; GHOSH, A.; MUHAMED, R. *Fundamentals of WiMAX*. Upper Saddle River, NJ: Prentice Hall PTR, 2007.
- [24] NUAYMI, L. *WiMAX: Technology for Broadband Wireless Access*. West Sussex, England: Wiley, 2007.
- [25] SESIA, S.; TOUFIK, I.; BAKER, M. *LTE The UMTS Long Term Evolution*. West Sussex, England: Wiley, 2009.
- [26] DAHLMAN, E. et al. *3G Evolution*. Jordan Hill, Oxford: Academic Press, 2008.
- [27] GESBERT, D. et al. Outdoor MIMO Wireless Channels: Models and Performance Prediction. *IEEE Transactions on Communications*, v. 50, n. 12, p. 1926–1934, December 2002.
- [28] SHANNON, C. A Mathematical Theory of Communications. In: *Bell Sys. Tech. J.* [S.l.: s.n.], 1948. p. 379–423.
- [29] SHANNON, C. Communications in the Presence of Noise. In: *Proceedings of IRE*. [S.l.: s.n.], 1949. p. 10–21.
- [30] FOSCHINI, G. J.; GANS, M. J. On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas. In: *Wireless Personal Commun. : Kluwer Academic Press*. [S.l.: s.n.], 1998. p. 311–335.
- [31] VERDU, S. Multiple-access Channels with Memory with and without Frame Synchronism. *IEEE Transactions on Information Theory*, v. 5, p. 605–619, May 1989.
- [32] TELATAR, E. Capacity of Multi-antenna Gaussian Channels. *European Transactions on Information Theory*, v. 10, n. 6, p. 585–596, November 1999.

- [33] YU, W. et al. Iterative Water-filling for Vector Multiple-access Channels. *IEEE Transactions on Information Theory*, v. 1, p. 145–152, January 2004.
- [34] BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. [S.l.]: Cambridge University Press, 2003.
- [35] EREZ, U.; SHAMAI, S.; ZAMIR, R. Capacity and Lattice Strategies for Cancelling Known Interference. In: *International Symposium on Information Theory and its Applications*. [S.l.: s.n.], 2000. p. 681–684.
- [36] ZAMIR, R.; FEDER, M. On Lattice Quantization Noise. *IEEE Transactions on Information Theory*, v. 42, n. 7, p. 1152–1159, July 1996.
- [37] URBANKE, R.; RIMOLDI, B. Lattice Codes Can Achieve Capacity on AWGN Channel. *IEEE Transactions on Information Theory*, v. 44, n. 1, p. 273, January 1998.
- [38] KIM, J.; CIOFFI, J. Spatial Multiuser Access with Antenna Diversity Using Singular Value Decomposition. In: *Proceedings of International Conference on Communications*. [S.l.: s.n.], 2000.
- [39] EREZ, U.; BRINK, S. Approaching the Dirty Paper Limit for Canceling Known Interference. In: *Proc. Allerton Conf. Communications, Control and Computing*. [S.l.: s.n.], 2003.
- [40] YU, W.; VARODAYAN, D.; CIOFFI, J. Trellis and Convolutional Precoding for Transmitter-based Interference Pre-subtraction, volume = 53, year = 2005. *IEEE Transactions on Communications*, p. 1220–1230, July.
- [41] JINDAL, N.; VISHWANATAH, S.; GOLDSMITH, A. J. On the duality of Gaussian Multiple Access and Broadcast Channels. *IEEE Transactions on Information Theory*, v. 5, p. 768–783, May 2004.
- [42] VISHWANATAH, P.; TSE, D. Sum capacity of the Vector Gaussian Broadcast Channel and Uplink-downlink Duality. *IEEE Transactions on Information Theory*, v. 8, p. 1912–1921, August 2003.
- [43] VISHWANATH, S.; JINDAL, N.; GOLDSMITH, A. Duality, Achievable Rates and Sum Capacity of Gaussian MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, v. 49, n. 10, p. 2658–2668, August 2003.
- [44] EREZ, U.; SHAMAI, S.; ZAMIR, R. Capacity and Lattice Strategies for Cancelling Known Interference. In: *Proceedings of Int. Symp. Inform. Theory and its Applications*. [S.l.: s.n.], 2000. p. 681–684.
- [45] JINDAL, N. et al. Sum Power Iterative Water-filling for Multi-antenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, v. 4, p. 1570–1579, April 2005.
- [46] ZHANG, J. et al. Linear Transmitter Precoding Design for Downlink of Multiuser MIMO Systems. *IEE Electronic Letters*, v. 41, n. 14, p. 811–813, July 2005.
- [47] YU, W.; CIOFFI, J. Trellis Precoding for the Broadcast Channel. In: *2001 IEEE GlobeCom*. [S.l.: s.n.], 2001. p. 1344–1348.
- [48] CAIRE, G.; SHAMAI, S. On the Achievable Throughput of a Multi-antenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, v. 49, n. 7, p. 1691–1706, July 2003.

- [49] LIU, J.; KRZYMIEN, W. A. Improved Tomlinson-Harashima Precoding for the Downlink of Multiple Antenna Multi-user Systems. *IEEE Wireless Communications and Networking Conference, 2005*, v. 1, n. 1, p. 466–472, March 2005.
- [50] SHI, S.; SCHUBERT, M. Precoding and Power Loading for Multi-antenna Broadcast Channels. In: *Conference on Information Sciences and Systems*. [S.l.: s.n.], 2004.
- [51] BAYESTEH, A.; KHANDANI, A. On the User Selection for MIMO Broadcast Channels. *IEEE ISIT*, p. 2325–2329, September 2005.
- [52] WANG, J.; LOVE, D. J.; ZOLTOWSKI, M. D. User Selection for the MIMO Broadcast Channel with a Fairness Constraint. *IEEE ICASSP*, v. 3, p. 9–12, April 2007.
- [53] RASHID-FARROKHI, F.; LIU, K. R.; TASSIULAS, L. Transmit Beamforming and Power Control for Cellular Wireless Systems. *IEEE Journal on Selected Areas in Communications*, v. 16, n. 8, p. 1437–1450, October 1998.
- [54] VISOTSKY, E.; MADHOW, U. Optimum Beamforming Using Transmit Antenna Arrays. *IEEE Vehicular Technology Conference*, v. 1, p. 851–856, May 1999.
- [55] GERLACH, D.; PAULRAJ, A. Adaptive Transmitting Antenna Arrays with Feedback. *IEEE Signal Processing Letters*, v. 1, n. 10, p. 150–152, October 1994.
- [56] HAUSTEIN, T. et al. Performance of MIMO Systems with Channel Inversion. In: *IEEE Vehicular Technology Conference*. [S.l.: s.n.], 2002.
- [57] LEE, J.; JUNDAL, N. Dirty Paper Coding vs. Linear Precoding for MIMO Broadcast Channels. In: *Fortieth Asilomar Conference on Signals, Systems and Computers, 2006*. [S.l.: s.n.], 2006. p. 779–783.
- [58] PROAKIS, J. G. *Digital Communications*. Avenue of Americas, New York: McGraw-Hill Science Engineering Math, 2000.
- [59] PEEL, C. B.; HOCHWALD, B. M.; SWINDLEHURST, A. L. A Vector-Perturbation Technique for Near Capacity Multi-antenna Multi-user Communication - Part I: Channel Inversion and Regularization. *IEEE Transactions on Communications*, v. 53, n. 1, p. 195–202, January 2005.
- [60] SKLAR, B. *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ: Prentice Hall PTR, 2001.
- [61] FOSCHINI, G. J. Layered Space-time Architecture for Wireless Communication in a Fading Environment when Using Multi-element Antennas. *Bell Labs Tech. Journal*, v. 1, n. 2, p. 41–59, Autumn 1996.
- [62] STANKOVIC, V.; HAARDT, M.; FUCHS, M. Combination of Block Diagonalization and THP Transmit Filtering for Downlink Beamforming in Multi-user MIMO Systems. In: *European Conference on Wireless Technology (ECWT 2004)*. [S.l.: s.n.], 2004.
- [63] HORN, R. A.; JOHNSON, C. R. *Matrix Analysis*. New York, NY: Cambridge University Press, 1985.
- [64] RALEIGH, G. G.; CIOFFI, J. M. Spatio-temporal Coding for Wireless Communication. *IEEE Transactions on Communications*, v. 46, p. 357–366, March 1998.

- [65] CHUAH, C.-N.; KAHN, J. M.; TSE, D. Capacity of Multi-antenna Array Systems in Indoor Wireless Environment. *IEEE Global Telecommunications Conference, 1998 GLOBECOM 98*, v. 4, n. 4, p. 1894–1899, 1998.
- [66] TOMLINSON, M. New Automatic Equalizer Employing Modulo Arithmetic. *Electronic Letters*, v. 7, n. 2, March 1971.
- [67] MIYAKAWA, H.; HARASHIMA, H. Information Transmission Rate in Matched Transmission Systems with Peak Transmitting Power Limitation. *Electronic Letters*, v. 7, n. 2, p. 138–139, August 1972.
- [68] FISCHER, R. et al. Space-Time Transmission using Tomlinson-Harashima Precoding. In: *4th International ITG Conference on Source and Channel Coding*. [S.l.: s.n.], 2002.
- [69] SHENOUDA, M. B.; DAVIDSON, T. N. A Framework for Designing MIMO Systems with Decision Feedback Equalization or Tomlinson-Harashima Precoding. *IEEE Journal on Selected Areas in Communications*, v. 26, n. 2, p. 401–411, February 2008.
- [70] FISCHER, R. F. H. *Precoding and Signal Shaping for Digital Transmission*. New York: John Wiley & Sons, 2002.
- [71] WINDPASSINGER, C.; VENCEL, T.; FISCHER, R. F. H. Precoding and Loading for Blast-like Systems. In: *IEEE International Conference on Communications*. [S.l.: s.n.], 2003. p. 3061–3065.
- [72] JOHAM, M.; BREHMER, J.; UTSCHICK, W. MMSE Approaches to Multiuser Spatio-temporal Tomlinson-Harashima Precoding. In: *5th International ITG Conference on Source and Channel Coding (ITG SCC04)*. [S.l.: s.n.], 2004. p. 387–394.
- [73] GOLUB, G.; LOAN, C. V. *Matrix Computations*. Baltimore, MD.: Johns Hopkins University Press, 1996.
- [74] CONWAY, J. H.; SLOANE, N. J. A. *Sphere Packings, Lattices and groups*. New York, NY: Springer Verlag, 1988.
- [75] FORNEY, G. D. Coset Codes - Part I: Introduction and Geometrical Classification. *IEEE Transactions on Information Theory*, p. 1123–1151, September 1988.
- [76] FORNEY, G. D. Coset Codes - Part II: Binary Lattices and Related Codes. *IEEE Transactions on Information Theory*, p. 1152–1187, September 1988.
- [77] WINDPASSINGER, C. et al. Precoding in Multi-antenna and Multiuser Communications. *IEEE Transactions on Wireless Communications*, v. 3, n. 4, p. 1305–1316, July 2004.
- [78] KAY, S. M. *Fundamentals of Statistical Processing, Volume I: Estimation Theory*. Upper Saddle River, New Jersey: Prentice Hall PTR, 1993.
- [79] STRANG, G. *Linear Algebra and Its Applications*. Belmont, CA: Thomson, Brooks and Cole, 2005.
- [80] WOLNIANSKY, P. et al. V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. In: *Proc. ISSSE 98*. [S.l.: s.n.], 1998. p. 295–300.

- [81] HASSIBI, B. An efficient square-root algorithm for BLAST. In: *Proc. ICASSP 00*. [S.l.: s.n.], 2000. p. 737–740.
- [82] FOSCHINI, G. et al. Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multielement Arrays. *IEEE J. Select. Areas Commun.*, v. 17, p. 1841–1852, November 1999.
- [83] KUSUME, K.; JOHAM, M.; BAUCH, W. U. ang G. Efficient Tomlinson-Harashima Precoding for Spatial Multiplexing on Flat MIMO Channel. In: *ICC - IEEE International Conference on Communications 2005*. [S.l.: s.n.], 2005. p. 2021– 2025.
- [84] FUNG, C.-H. F.; YU, W.; LIM, T. J. Precoding for the Multiantenna Downlink: Multiuser SNR Gap and Optimal User Ordering. *IEEE Transactions on Communications*, v. 55, n. 1, p. 188–197, January 2007.
- [85] HOCHWALD, B. M.; PEEL, C. B. Vector Precoding for the Multi-antenna, Multi-user Channel. In: *Allerton Conference on Communication, Control, and Computing*. [S.l.: s.n.], 2003.
- [86] DAMEN, M. O.; GAMAL, H. E.; CAIRE, G. On Maximum-likelihood Detection and the Search for the Closest Lattice Point. *IEEE transactions on Information Theory*, v. 49, p. 2389–2402, October 2003.
- [87] AGRELL, E. et al. Closest Point Searches in Lattices. *IEEE transactions on Information Theory*, v. 28, p. 2201–2214, August 2002.
- [88] MOW, W. H. Universal Lattice Decoding: Principle and Recent Advances. *Wireless Communications and Mobile Computing*, v. 3, n. 5, p. 553–569, 2003.
- [89] FINCKE, U.; POHST, M. Improved Methods for Calculating Vectors of Short Lengths in a Lattice, Including a Complexity Analysis. In: *Proc. ACM Symp. Theory Comput.* [S.l.: s.n.], 1985. p. 463–471.
- [90] DAMEN, M. O.; CHKEIF, A.; BELFIORE, J. Lattice Code and Decoder for Space-time Codes. *IEEE Communication Letters*, v. 4, p. 161–163, May 200.
- [91] AIRY, M. et al. Transmit Precoding for the Multiple Antenna Broadcast Channel. In: *IEEE 63rd Vehicular Technology Conference*. [S.l.: s.n.], 2006. p. 1396–1400.
- [92] HASSIBI, B.; VIKALO, H. On the Expected Complexity of Integer Least Squares Problems. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.* [S.l.: s.n.], 2002. p. 1497–1500.
- [93] WINDPASSINGER, C.; FISCHER, R. F. H.; HUBER, J. B. Lattice-Reduction-Aided Broadcast Precoding. *IEEE Transactions on Communications*, v. 52, n. 12, p. 2057–2060, December 2004.
- [94] LENSTRA, H. W. L. A. K.; LOVASZ, L. Factoring Polynomials with Rational Coefficients. *Mathematische Annalen*, v. 261, p. 515–534, 1982.
- [95] SEETHALER, D.; MATZ, G. Efficient vector perturbation in multi-antenna multi-user systems based on Approximate Integer Relations. In: *Proc. EUSIPCO 2006*. [S.l.: s.n.], 2006.
- [96] LIU, F.; JIANG, L.; HE, C. Low Complexity MMSE Vector Precoding Using Lattice Reduction for MIMO Systems. *IEEE International Conference on Communications ICC 2007*, p. 2598–2603, June 2007.

- [97] YAP, C.-K. *Fundamental Problems of Algorithmic Algebra*. Oxford, New York: Oxford University Press, 2000.
- [98] SCHMIDT, D. A.; JOHAM, M.; UTSCHICK, W. Minimum Mean Square Error Vector Precoding. *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005*, v. 1, p. 107–111, September 2005.
- [99] HOCHWALD, B.; BRINK, S. ten. Achieving Near-capacity on a Multiantenna Channel. *IEEE Transactions on Communications.*, v. 51, p. 389–399, March 2003.
- [100] KIM, E. Y.; CHUN, J. Optimum Vector Perturbation Minimizing Total MSE in Multiuser MIMO Downlink. *IEEE International Conference on Communications ICC 2006*, v. 9, p. 4242–4247, June 2006.
- [101] CHUA, W. S.; YUEN, C.; CHIN, F. A Continuous Vector-Perturbation for Multi-antenna Multi-user Communication. *IEEE 65th Vehicular Technology Conference, VTC2007 Spring*, p. 1806–1810, April 2007.
- [102] MARZETTA, T. L.; HOCHWALD, B. M. Fast Transfer of Channel State Information in Wireless Systems. *IEEE Transactions on Signal Processing*, v. 54, n. 4, p. 1268–1278, April 2006.
- [103] DUEL-HALLEN, A. Fading Channel Prediction for Mobile Radio Adaptive Transmission Systems. In: *Proceedings of the IEEE*. [S.l.: s.n.], 2007. p. 2299–2313.
- [104] PROAKIS, J. G.; MANOLAKIS, D. K. *Digital Signal Processing*. Upper Saddle River, New Jersey: Prentice Hall, 2006.
- [105] OPPENHEIM, A. V.; SCHAFER, R. W.; BUCK, J. R. *Discrete-Time Signal Processing*. Upper Saddle River, New Jersey: Prentice Hall, 1999.
- [106] DINIZ, P. S. R. *Adaptive Filtering: Algorithms and Practical Implementation*. [S.l.]: Kluwer Academic Publishers, 2002.
- [107] HAYKIN, S. *Adaptive Filter Theory*. Upper Saddle River, New Jersey: Prentice Hall, 2001.
- [108] WIENER, N. *Extrapolation, Interpolation and Smoothing of Stationary Time-Series with Engineering Applications*. Cambridge, NA: MIT Press, 1949.
- [109] HAYES, M. H. *Statistical Digital Signal Processing and Modeling*. [S.l.]: Wiley, 1996.
- [110] SAYED, A. H. *Fundamentals of Adaptive Filtering*. Hoboken, New Jersey: Wiley-IEEE Press, 2003.
- [111] GOLLAMUDI, S. et al. Set-Membership Filtering and a Set-Membership Normalized LMS Algorithm with an Adaptive Step Size. *IEEE Signal Processing Letters*, v. 5, n. 5, p. 111–114, May 1998.
- [112] GALDINO, J.; JR, J. A.; CAMPOS, M. de. A Set-Membership NLMS Algorithm with Time-Varying Error Bound. *Proc. IEEE Intern. Symposium on Circuits and Systems*, p. 277–280, May 2006.
- [113] WERNER, S.; DINIZ, P. Set-Membership Affine Projection Algorithm. *IEEE Signal Processing Letters*, v. 8, n. 8, p. 231–235, August 2001.

- [114] WERNER, S. *Reduced Complexity Adaptive Filtering with Applications to Communications Systems*. Tese (Doutorado) — Helsinki University of Technology, November 2002.
- [115] XIONG, F. *Digital Modulation Techniques*. Norwood, MA, USA: Artech House Publishers, 2006.
- [116] EKMAN, T. *Prediction of Mobile Radio Channels: Modeling and Design*. Tese (Doutorado) — Uppsala University, October 2002.
- [117] RAMYA, T. R.; BHASHYAM, S. On Using Channel Prediction in Adaptive Beamforming Systems. *2nd International Conference on Communication Systems Software and Middleware, 2007. COMSWARE 2007*, p. 1–6, January 2007.
- [118] SVANTESSON, T.; SWINDLEHURST, A. L. A Performance Bound for Prediction of MIMO Channels. *IEEE Transactions on Signal Processing*, v. 54, n. 2, p. 520–529, February 2006.
- [119] VANDERPYPEN, J.; SCHUMACHER, L. MIMO Channel Prediction Using ESPRIT Based Techniques. *The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC 07*, p. 1–5, September 2007.
- [120] DUEL-HALLEN, A.; HU, S.; HALLEN, H. Long Range Prediction of Fading Channels. *IEEE Signal Processing Magazine*, v. 17, n. 3, p. 62–75, May 2000.
- [121] FORENZA, A.; JR, R. W. H. Link Adaptation and Channel Prediction in Wireless OFDM Systems. *45th Midwest Symposium on Circuits and Systems*, p. 211–214, August 2002.
- [122] SCHAFHUBER, D.; MATZ, G. MMSE and Adaptive Prediction of Time-Varying Channels for OFDM Systems. *IEEE Transactions on Wireless Communications*, v. 4, n. 3, p. 593–602, March 2005.
- [123] BOS, A. van den. *Parameter Estimation for Scientists and Engineers*. Hoboken, New Jersey: Wiley-Interscience, 2007.
- [124] MATZ, G. *A time-frequency calculus for time-varying systems and nonstationary processes with applications*. Tese (Doutorado) — Vienna University of Technology, November 2000.
- [125] SCHAFHUBER, D. *Wireless OFDM Systems: Channel Prediction and System Capacity*. Tese (Doutorado) — Vienna Technical University, March 2003.
- [126] 3GPP. *3GPP TR 36.211 V8.5.0 - Physical Channels and Modulation*. [S.l.], 2008.
- [127] 3GPP. *3GPP TR 25.996 V8.5.0 Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations*. [S.l.], 2003.
- [128] LARSSON, D. *Analysis of Channel Estimation Methods for OFDMA*. Dissertação (Mestrado) — KTH School of Electrical Engineering, December 2006.
- [129] SOMASEGARAN, L. *Channel Estimation and Prediction in UMTS LTE*. Dissertação (Mestrado) — Aalborg University, June 2007.
- [130] ZEMEN, T. *OFDM Multi-User Communication Over Time-Variant Channels*. Tese (Doutorado) — Technische Universität Wien, July 2004.

- [131] CAIRE, G.; SHIRANI-MEHR, H. Feedback Schemes for multiuser MIMO-OFDM Downlink. *Information Theory and Applications Workshop, 2008*, p. 33–40, February 2008.
- [132] SALO, J. et al. *MATLAB implementation of the 3GPP Spatial Channel Model*. [S.l.], 2005. Online <http://www.tkk.fi/Units/Radio/scm/>.
- [133] DENT, P.; BOTTOMLEY, G. E.; CROFT, T. Jakes Fading Model Revisited. *IEEE Electronic Letters*, v. 29, p. 1162–1163, June 1993.
- [134] FLEURY, B. H. An Uncertainty Relation for WSS Processes and Its Application to WSSUS Systems. *IEEE Transactions on Wireless Communications*, v. 44, n. 12, p. 1632–1634, December 1996.
- [135] LEITE, J. P.; VIEIRA, R. D.; CARVALHO, P. H. P. de. OFDM Channel Prediction Using Set-Membership Affine Projection Algorithm in Time-Varying Wireless Channel. In: *IEEE 10th Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2009) (July 2009)*. [S.l.: s.n.], 2009.
- [136] LEITE, J. P.; VIEIRA, R. D.; CARVALHO, P. H. P. de. Predição de Canal em Sistemas OFDM Utilizando o Algoritmo Set-Membership Affine Projection. In: *XXVII Simpósio Brasileiro de Telecomunicações - SBrT 2009*. [S.l.: s.n.], 2009.
- [137] BREHMER, J. et al. Reduced-complexity Linear and Nonlinear Precoding for Frequency-Selective MIMO Channels. *60th Vehicular Technology Conference: VTC 2004 - Fall*, v. 5, p. 3684–3688, September 2004.
- [138] WERNER, S.; MOREIRA, J. E. W.; DINIZ, P. S. R. Set-Membership Affine Projection Algorithm with Variable Data-Reuse Factor. In: *IEEE International Symposium in Circuits and Systems - ISCAS 2006*. [S.l.: s.n.], 2006. p. 261–264.
- [139] DINIZ, P. S. R.; PINTO, G. O.; HJORUNGNES, A. Data Selective Partial-Update Affine Projection Algorithm. In: *IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2008*. [S.l.: s.n.], 2008. p. 3833–3836.
- [140] FETTWEIS, G. et al. Dirty RF: A New Paradigm. *International Journal of Wireless Information Networks*, v. 14, n. 2, p. 133–148, June 2007.
- [141] YOO, T.; JINDAL, N.; GOLDSMITH, A. Multi-antenna Downlink Channels with Limited Feedback and User Selection. *IEEE Journal on Selected Areas in Communications*, v. 25, n. 7, p. 1478–1491, September 2007.
- [142] JALALI, A.; PADOVANI, R.; PANKAJ, R. Data Throughput of CDMA-HDR a High Efficiency High Data Rate Personal Communication Wireless System. *IEEE 51st Vehicular Technology Conference Proceedings. VTC 2000*, v. 3, p. 1854–1858, 2000.
- [143] VISWANATH, P.; TSE, D. N. C.; LAROIA, R. Opportunistic Beamforming Using Dumb Antennas. *IEEE Transactions on Information Theory*, v. 48, n. 6, p. 1277–1294, June 2002.
- [144] SPENCER, Q. H.; SWINDLEHURST, A. L. Channel Allocation in Multi-user MIMO Wireless Communications Systems. *IEEE International Conference on Communications 2004*, v. 5, p. 3035–3039, June 2004.

I. LONG TERM EVOLUTION

O LTE é uma proposta de evolução para o UMTS (*Universal Mobile Telecommunications System*), sistema de telefonia celular de terceira geração. Elaborado e discutido por uma parceria denominada 3GPP (*Third Generation Partnership Project*), o LTE busca manter a competitividade da linha evolutiva que advém do sistema GSM (*Global System for Mobile Communications*).

Neste anexo, será considerada uma breve descrição da camada física do enlace direto do padrão LTE operando em FDD para uma largura de banda de 10 MHz e prefixo cíclico curto, que corresponde ao cenário utilizado para realizar as simulações apresentadas no capítulo 5.

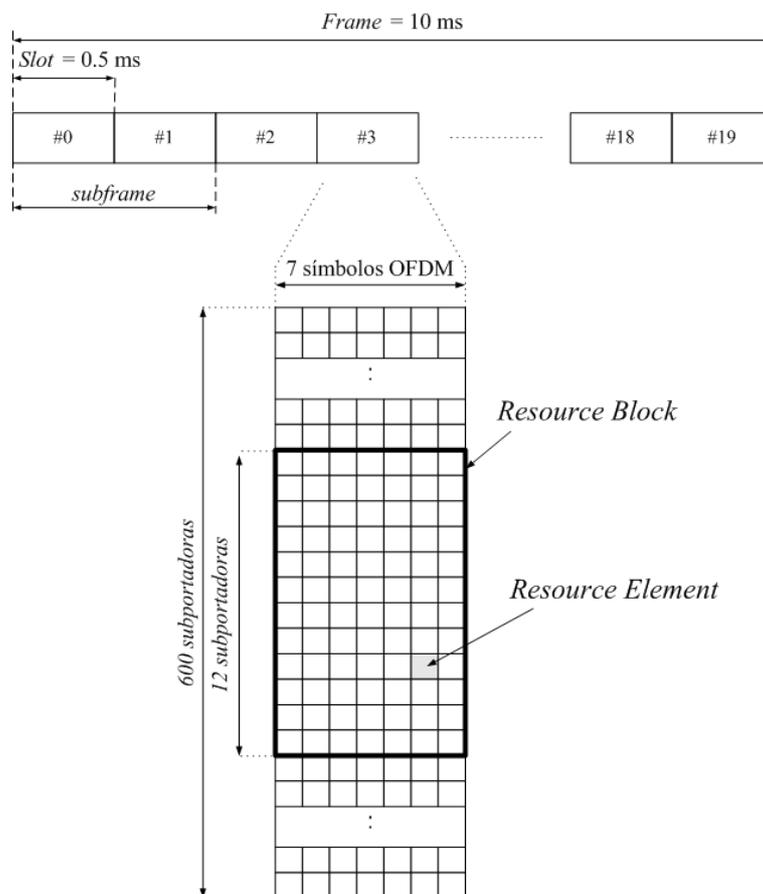


Figura I.1: Estrutura de quadro no enlace direto do padrão LTE para uma largura de banda de 10 MHz.

A técnica de múltiplo acesso utilizada no enlace direto do sistema LTE é o OFDMA. Como consequência, o quadro de rádio (*frame*) possui dimensões tanto no domínio do tempo como no domínio da frequência, conforme mostrado na Fig. I.1. O quadro de rádio tem duração de 10 ms e é composto por 20 subquadros, cada um com 1 ms de duração. Cada subquadro, por sua vez, é constituído de 2 *slots*, cada um com duração de 0,5 ms.

Logo, pela descrição apresentada no parágrafo anterior, o quadro de rádio é formado por 20 *slots*, numerados de 0 a 19, conforme mostrado na Fig. I.1. Cada *slot* contém de 7 símbolos OFDM. Para a banda de transmissão de 10 MHz, são utilizadas 600 subportadoras, com espaçamento de 15 kHz entre

elas.

A menor unidade de recurso recebe a denominação de elemento de recurso (*resource element*), e corresponde a uma subportadora dentro de um símbolo OFDM. O bloco de recursos (*resource block*) é definido como o conjunto de 7 símbolos OFDM contíguos no domínio do tempo e de 12 subportadoras adjacentes no domínio da frequência, e consiste na menor unidade de recursos que pode ser alocada para um usuário.

A distribuição das subportadoras que carregam símbolos pilotos, utilizados para a estimação de canal, é mostrada na Fig. I.2 para o caso em que a estação rádio-base opera utilizando quatro antenas de transmissão. Pode-se perceber que há símbolos de referência no primeiro e no quarto símbolo OFDM de um *slot*. Existe um espaçamento de seis subportadoras entre os símbolos de referência, sendo que os símbolos de referência localizados no quarto símbolo OFDM de um *slot* são alocados com um *offset* de quatro subportadoras com relação ao primeiro símbolo do *slot*.

Mais especificamente, no padrão LTE, a estação rádio-base pode possuir uma, duas ou quatro antenas, e quando duas ou mais antenas são utilizadas para transmissão, os símbolos de referência são posicionados de tal forma que sejam ortogonais, de forma a não interferir uns nos outros. Essa ortogonalidade é obtida garantindo que nenhum sinal seja transmitido nos elementos de recurso que são utilizados por uma determinada antena para a transmissão de símbolos de referência. Estes são marcados na Fig. I.2 como as portadoras não utilizadas.

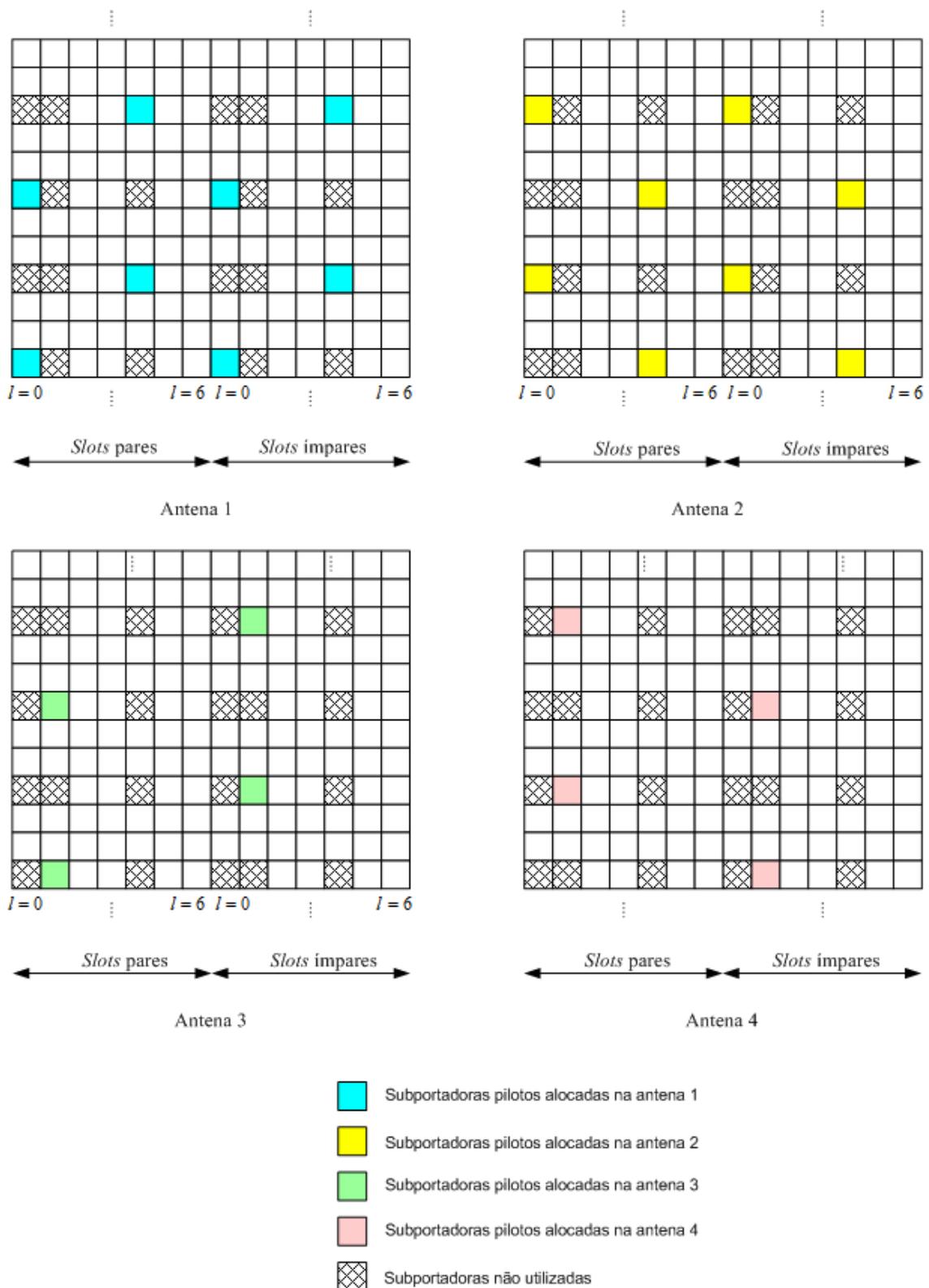
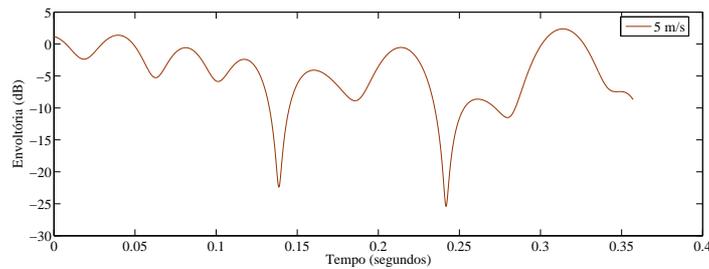


Figura I.2: Alocação de subportadoras que carregam símbolos pilotos no enlace direto do padrão LTE para quatro antenas de transmissão.

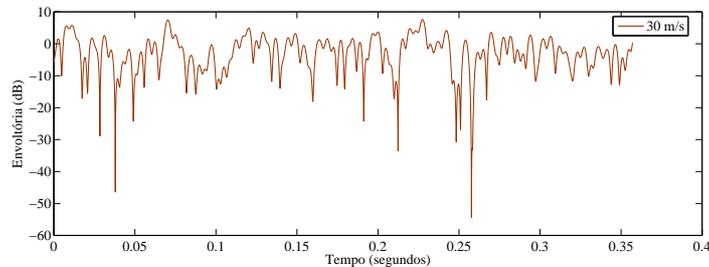
- P_n é a potência do n -ésimo multipercorso;
- σ_s é a desvio padrão da componente de sombreamento (*shadowing*);
- $\theta_{n,m,AoD}$ é o ângulo de partida do m -ésimo *subpath* do n -ésimo multipercorso;
- $\theta_{n,m,AoA}$ é o ângulo de chegada do m -ésimo *subpath* do n -ésimo multipercorso;
- $G_{BS}(\theta_{n,m,AoD})$ é o ganho de cada elemento do conjunto de antenas na estação rádio-base;
- $G_{MS}(\theta_{n,m,AoA})$ é o ganho de cada elemento do conjunto de antenas no terminal móvel;
- k é o número de onda, dado por $\frac{2\pi}{\lambda}$, em que λ é o comprimento de onda;
- d_s é a distância entre a antena s do conjunto de antenas da estação rádio-base e a antena de referência (para a qual $s = 1$);
- d_u é a distância entre a antena u do conjunto de antenas do móvel e a antena de referência (para a qual $u = 1$);
- $\Phi_{n,m}$ é a fase do m -ésimo *subpath* do n -ésimo multipercorso;
- $\|\mathbf{v}\|$ é o módulo do vetor velocidade;
- θ_v é o ângulo do vetor velocidade.

O modelo também apresenta funcionalidades adicionais, como, por exemplo, a modelagem dos efeitos de propagação considerando a polarização utilizada pelas antenas e a propagação em linha de visada.

Um exemplo de realização de canal para $u = s = n = 1$ é mostrado na Fig. II.2 para diferentes valores de velocidade do terminal móvel. A figura exibe o módulo da resposta $h_{u,s,n}(t)$.



(a) 5 m/s



(b) 30 m/s

Figura II.2: Exemplo de realização do canal para diferentes velocidades do móvel.

