



FILTRAGEM ADAPTATIVA COM SOLUÇÃO ESPARSA PARA ARRANJOS DE SENSORES

José Francisco de Andrade Junior

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Elétrica.

Orientadores: Marcello Luiz Rodrigues de
Campos

José Antonio Apolinário Junior

Rio de Janeiro

Março de 2015

FILTRAGEM ADAPTATIVA COM SOLUÇÃO ESPARSA PARA ARRANJOS
DE SENSORES

José Francisco de Andrade Junior

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Marcello Luiz Rodrigues de Campos, Ph.D.

Prof. José Antonio Apolinário Junior, D.Sc.

Prof. Paulo Sergio Ramirez Diniz, Ph.D.

Prof. Vitor Heloiz Nascimento, Ph.D.

Prof. Marco Antonio Grivet Mattoso Maia, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2015

Andrade Junior, José Francisco de

Filtragem Adaptativa com Solução Esparsa para Arranjos de Sensores / José Francisco de Andrade Junior. – Rio de Janeiro: UFRJ/COPPE, 2015.

XIV, 82 p.: il.; 29,7cm.

Orientadores: Marcello Luiz Rodrigues de Campos

José Antonio Apolinário Junior

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2015.

Referências Bibliográficas: p. 73 – 76.

1. Filtragem Adaptativa Esparsa. 2. Arranjo de Sensores. 3. Norma L_1 . 4. Esparsidade. I. Campos, Marcello Luiz Rodrigues de *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

À minha família

Agradecimentos

O sucesso é ir de fracasso em fracasso sem perder o entusiasmo.

Winston Churchill

Em primeiro lugar, a Deus por tudo que conquistei na vida.

À minha esposa Ozaneide, pelo inestimável apoio, dedicação e compreensão, sem os quais eu não seria possível a conclusão deste trabalho; e ao meu filho Gabriel pelas horas revitalizantes de convívio.

Aos meus pais, Andrade e Clotilde Fortunato (*in memoriam*), pelo zelo e dedicação a mim conferidos. Aos meus avós Benedito (*in memoriam*) e Nazaré de Andrade, pelo amor, pela minha educação, que, na falta da minha mãe, para mim foram tudo.

Aos meus Amigos e Irmãos de Infância, pelo constante incentivo e compreensão, sem os quais eu não teria chegado até aqui.

De forma muito especial, aos meus Orientadores e amigos, Professores Marcello Campos e José Antonio Apolinário Jr., pelo incentivo, compreensão, confiança, amizade, ensinamentos, paciência e forma tranquila com que conduziram a árdua tarefa de orientar.

Aos membros da Banca examinadora, por terem aceitado o convite para fazer parte deste processo de avaliação.

Ao amigo Marcelo Ventura, pelas proveitosas discussões sobre os tópicos de Processamento de Sinais.

À colega Catia Valdman, pelo companheirismo e proveitosas discussões sobre os tópicos ministrados nas disciplinas de Processamento de Sinais.

Aos professores do Programa de Engenharia Elétrica Amit Bhaya e Sergio Lima Netto, pelos excelentes cursos ministrados.

À Marinha do Brasil, por ter permitido e apoiado a realização deste curso em regime parcial.

Ao Projeto COPPE_{TE}X por contribuir para a edição em nível profissional desta Tese.

Finalmente, agradeço a todos, incluindo professores e funcionários do PEE, que, de alguma forma, colaboraram para a consecução deste trabalho.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

FILTRAGEM ADAPTATIVA COM SOLUÇÃO ESPARSA PARA ARRANJOS DE SENSORES

José Francisco de Andrade Junior

Março/2015

Orientadores: Marcello Luiz Rodrigues de Campos

José Antonio Apolinário Junior

Programa: Engenharia Elétrica

Neste trabalho é proposta uma família contendo três algoritmos adaptativos com restrições baseada no algoritmo ℓ_1 -norm *Linearly-Constrained Least-Mean-Square* (ℓ_1 -CLMS), cuja característica diferencial é a flexibilidade decorrente da coexistência de restrições de norma ℓ_1 e lineares em um mesmo algoritmo. A família de algoritmo proposta se destina à solução de problemas envolvendo síntese de arranjos de sensores com um grande número de elementos e é baseada na minimização do erro médio quadrático sujeita às restrições lineares e de norma unitária (ℓ_1). Dentre os algoritmos da família, existe uma versão normalizada e outra ponderada-normalizada.

O desempenho desses algoritmos é investigado na síntese de arranjos adaptativos de sensores com elevado número de sensores sob a presença de ruído gaussiano e na identificação de sistemas com diferentes níveis de esparsidade.

Foi verificado que, quando aplicada à síntese de arranjos adaptativos de sensores, a restrição de norma ℓ_1 permite atingir configurações esparsas para os arranjos finais. Para problemas envolvendo identificação de sistemas esparsos, a versão ponderada do algoritmo ℓ_1 -CNLMS apresenta um desempenho superior em relação aos demais algoritmos propostos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

SPARSE FILTERING SOLUTIONS FOR SENSOR ARRAYS

José Francisco de Andrade Junior

March/2015

Advisors: Marcello Luiz Rodrigues de Campos

José Antonio Apolinário Junior

Department: Electrical Engineering

This Thesis proposes a family with three linearly constrained with ℓ_1 -norm adaptive algorithms. The algorithm that gives rise to this family is the ℓ_1 -norm *Linearly-Constrained Least-Mean-Square*. The algorithms proposed by this work have as distinguishing feature the flexibility resulting from the coexistence between ℓ_1 -norm and linear constraints. As practical application for the proposed algorithms, we have synthesis of sensors arrays with a large number of sensors based on minimization of mean squared error subjected to constraints. Among the family of algorithms, there is a normalized version and a weighted-normalized version.

The performance of these algorithms is investigated on the synthesis of adaptive arrays of sensors with a large number of sensors in the presence of Gaussian noise and on the system identification problem for varying degrees of sparsity.

When applied to the adaptive sensor array synthesis problem, the ℓ_1 -norm constraint allows settings to achieve the final sparse arrays. For applications involving identification of sparse systems, the weighted version ℓ_1 -WCNLMS algorithm has superior performance when compared to the other proposed algorithms.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiv
1 Introdução	1
1.1 Filtragem adaptativa no contexto dos arranjos de sensores	1
1.2 Uso de restrições de norma ℓ_1 como um caminho viável para soluções esparsas	4
1.3 Propósito deste Trabalho	8
1.4 Estrutura da tese	9
1.5 Notação	11
2 Fundamentos de Arranjos de Sensores	12
2.1 Modelo do Sinal de Entrada	13
2.2 Diagrama de Radiação do Arranjo de Sensores (<i>Beam pattern</i>) . . .	15
2.3 Conformação Adaptativa de Feixe	16
2.4 Conformador de Feixe de Variância Mínima Com Restrições Lineares	17
3 Algoritmos da Família ℓ_1-CLMS	20
3.1 Revisão: Algoritmos CLMS e CNLMS	20
3.1.1 Algoritmo CLMS	20
3.1.2 Algoritmo CNLMS	21
3.2 Algoritmo ℓ_1 -CLMS	22
3.3 Algoritmo ℓ_1 -CNLMS	25
3.4 Algoritmo ℓ_1 -WCNLMS	26
3.5 Complexidade Computacional	30
4 Análises de Estabilidade e Convergência	31
4.1 Análise de Primeira Ordem	32
4.2 Análise de Segunda Ordem	35
4.3 Excesso de MSE e Desajuste	39

5	Síntese de Arranjos de Sensores	45
5.1	Arranjo Hexagonal Padrão (<i>Standard Hexagonal Array</i> - SHA) para Satélites Geoestacionários Militares na Banda- <i>X</i>	46
5.2	Arranjo Hexagonal (HAA) para Comunicações por Satélite na Banda- <i>Ku</i>	49
5.3	Arranjo Uniforme Linear	55
5.3.1	ULA: Simulação I	56
5.3.2	ULA: Simulação II	57
5.3.3	ULA: Simulação III	58
5.4	Conclusão do Capítulo	60
6	Simulações Adicionais	62
6.1	Sistema Esperso Complexo	62
6.2	Sistema Compressível Complexo	64
6.3	Excesso de MSE e Desajuste	65
6.4	Conclusão do Capítulo	69
7	Conclusões	71
	Referências Bibliográficas	73
A	Minimum-disturbance Approach	77

Lista de Figuras

1.1	Modelagem genérica para um Arranjo Adaptativo Linear	2
1.2	Comparação entre as soluções obtidas com os algoritmos LASSO (azul) e Ridge (vermelho) em relação a solução <i>Least Square</i> ordinária. As restrições referentes aos algoritmos LASSO e Ridge são representadas pelas Bolas ℓ_1 (losango) e ℓ_2 (círculo), respectivamente.	6
1.3	Projeção ortogonal de solução <i>Least Square</i> sobre um das faces de um hiper cubo no \mathbb{R}^3 . Comparando-se as solução $\mathbf{w}_{\text{LASSO1}}$ e $\mathbf{w}_{\text{LASSO2}}$, pode-se observar, que a partir de um ajuste adequado na restrição de norma ℓ_1 , a solução final torna-se esparsa.	7
2.1	Geometria Genérica de um Arranjo Planar.	13
2.2	Geometria de um Arranjo Linear Uniforme (ULA).	14
2.3	Exemplo de Diagrama de Radiação para arranjo tipo ULA com 20 sensores e direção interesse igual a 90°	16
2.4	Estrutura de <i>Beamforming</i> para sinais de banda estreita.	17
3.1	Resultado das aproximações da penalidade de norm ℓ_0 , em função de β , por meio das funções de ponderação $\phi_1(\mathbf{w})$ e $\phi_2(\mathbf{w})$. Quando comparado com o resultado obtido pela penalidade de norma ℓ_1 , i.e., $\phi(\mathbf{w}) = \ \mathbf{w}\ _1$, observa-se um notável ganho em favor do emprego das funções de ponderação para valores de w_i próximos à origem.	29
5.1	Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -CNLMS e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -CNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	47
5.2	Síntese do arranjo SHA com 91 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -CNLMS e CNLMS.	48

5.3	Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -WCNLMS e respectivos <i>beam patterns</i> computados utilizando os algoritmos ℓ_1 -WCNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	48
5.4	Síntese do arranjo SHA com 91 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtidos por meio dos algoritmos ℓ_1 -WCNLMS e CNLMS.	49
5.5	Ilustração de feixes de cobertura em uma porção do footprint de um satélite geostacionário. Considera-se, neste caso, uma taxa de reuso de frequências de 1 : 4, cuja função é garantir que nenhum feixe adjacente utilize uma mesma frequência ou um mesmo grupo de frequências.	50
5.6	Sínteses para uma arranjo tipo HAA via algoritmos L_1 -CNLMS, CNLMS e solução LCMV, observado de um zênite $\phi = 47.5^\circ$	51
5.7	Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -CNLMS e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -CNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	52
5.8	Síntese do arranjo HAA com 397 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -CNLMS e CNLMS.	53
5.9	Sínteses para uma arranjo tipo HAA via algoritmos L_1 -WCNLMS, CNLMS e solução LCMV, observado de um zênite $\phi = 47.5^\circ$	53
5.10	Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -WCNLMS e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -WCNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	54
5.11	Síntese do arranjo HAA com 397 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -WCNLMS e CNLMS.	55

5.12	Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.62t_o$ e $\beta = 1$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	56
5.13	Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.58t_o$ e $\beta = 2$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	57
5.14	Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos <i>beam patterns</i> computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.56t_o$ e $\beta = 2$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.	58
5.15	Síntese do arranjo ULA com 20 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS e CNLMS na Simulação III.	59
6.1	MSE obtido para os algoritmos L_1 -CNLMS, L_1 -WCNLMS e CNLMS, quando da identificação de planta complexa esparsa com 50 coeficientes e sinal de referência contaminado com ruído de medição de variância $\sigma_N^2 = 10^{-2}$	63
6.2	MSE obtido para os algoritmos L_1 -CNLMS, L_1 -WCNLMS e CNLMS, quando da identificação de planta complexa compressível com 50 coeficientes e sinal de referência contaminado com ruído de medição de variância $\sigma_N^2 = 10^{-2}$	64
6.3	Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 0,1$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.	65

6.4	Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 0,5$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.	66
6.5	Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 1,0$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.	67
6.6	Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 1,5$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.	68
6.7	$J(\infty)$ e Desajuste calculados por meio das formulações aproximadas proposta pelas Eqs. (4.33) e (4.36), respectivamente, para uma planta complexa com 20 coeficientes, passo de convergência fixo $\mu_0 = 1$, e ruído de medição com variância $\sigma_N^2 = 10^{-4}$. Nesta simulação foram executadas 400 realizações independentes.	69

Lista de Tabelas

1.1	Notação	11
3.1	Número de Operações Complexas por Iteração ($L = 1$)	30
5.1	Parâmetros empregados nas simulações das Seções 5.1 e 5.2	47
5.2	Parâmetro empregados nas simulações da Seção 5.3	60

Capítulo 1

Introdução

Nos sistemas de telecomunicações modernos, assim como numa série de outras aplicações, como radares, sonares, radio-telescópios, sistemas de teleconferência, equipamentos para aquisição de vídeo e imagem (câmeras), dentre outros, o uso de arranjo de sensores vem se tornando cada vez mais comum. Existe, portanto, uma demanda pela otimização e soluções que atendam o crescente nível de qualidade de serviço exigida por esses sistemas. Sinais de interferência podem impor graves inconvenientes e limitações para sistemas que não possuem dispositivos capazes atenuar os seus efeitos, quer devido à falta de capacidade computacional, quer devido às limitações de fonte de energia (aqueles alimentados por baterias, por exemplo). Desta forma, existe uma crescente demanda por sistemas que consigam mitigar os efeitos de sinais de interferência e, que ao mesmo tempo, otimizam o número de sensores ativos em um arranjo.

Em decorrência do cenário onde coexistem sinais de interesse e sinais interferentes ser dinâmico, sobretudo para sistemas de comunicações móveis, os algoritmos para filtragem adaptativa esparsa assumem o papel de candidatos ideais para solucionar o problema apresentado.

1.1 Filtragem adaptativa no contexto dos arranjos de sensores

Ao longo das últimas décadas, a filtragem adaptativa está cada mais vez mais presente nos sistemas de comunicações, em especial com o adventos das comunicações celulares, dos sistemas multiusuário, das comunicações móveis via satélite e, no campo militar, dos sistemas de Guerra Eletrônica. Todos esses sistemas dependem de antenas eficientes e, desta forma, têm os arranjos de sensores como uma atraente opção em termos de sistemas irradiantes com flexibilidade capaz de atender essa demanda. Decorrente da características dinâmicas dos sistemas supracitados, tais

arranjos necessitam ser reconfigurados dinamicamente e, como já mencionado, uma solução já consagrada é o emprego de filtragem adaptativa espacial [1].

Por filtragem adaptativa, entende-se o emprego de filtros, cujos coeficientes podem ser ajustados de forma dinâmica em função de parâmetros variantes no tempo impostos pelo meio externo ao filtro. Esses parâmetros permitem expressar quantitativamente sinais de erro, que são usados como elementos de ajuste da função de transferência de malha aberta do filtro. Casos típicos de filtragem adaptativa utilizam a minimização do valor médio quadrático do sinal de erro, que é obtido com base na dinâmica do sinal de entrada, tomando-se como base um sinal de referência.

Restringindo-se ao contexto dos Arranjos de Sensores, ilustra-se, por meio da Fig. 1.1, um modelo genérico de um arranjo de sensores adaptativo linear. O sinal de saída y_k , portanto, é comparado a um sinal de referência d_k , gerando, então um sinal de erro e_k , cujo valor é minimizado pelo algoritmo adaptativo até um valor aceitável à solução do problema. Para alguns algoritmos, escolhe-se como figura de mérito a ser minimizada, o valor médio quadrático do erro (*Mean-Square Error - MSE*).

Portanto, neste modelo, o sinal de erro é responsável pelo controle dinâmico da planta em malha aberta do sistema. Por sua vez, os coeficientes computados pelo filtro, representados pelos pesos w_i , são responsáveis pelo ajuste do arranjo, condição que o transforma em arranjo de sensores adaptativo.

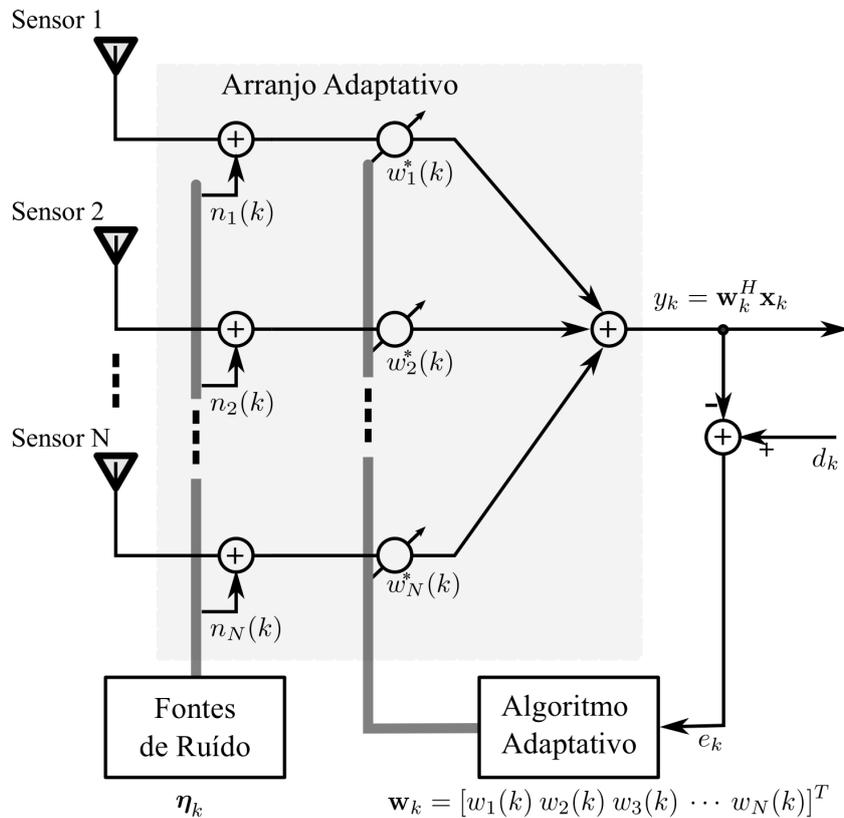


Figura 1.1: Modelagem genérica para um Arranjo Adaptativo Linear

Nas aplicações voltadas aos arranjos de sensores, os algoritmos empregados dividem-se em duas grandes classes: autodidatas (ou cegos) e treinados. Os algoritmos treinados fazem uso de um sinal de referência previamente conhecido ou estimado localmente. Fazem parte desta classe os filtros de Wiener-Hopf (Wiener) [2], e, dentre outros, os seguintes algoritmos clássicos:

- *Least Mean Square* (LMS) [3];
- *Constrained Least Mean Square* (CLMS) [4] ;
- *Constrained Normalized Least Mean Square* (CNLMS) [5] ;
- *Recursive Least Square* (RLS) [3]; e
- *Constrained Affine Projection* (CAP) [6].

Na prática, quando o problema envolve arranjos de sensores adaptativos, é comum se fazer uso da energia do sinal de saída do arranjo, y_k , como sinal de erro, assumindo-se o sinal de referência d_k como nulo. Desta maneira, a solução se dá por meio da minimização da energia do sinal na saída do arranjo, dispensado, portanto, sinal de referência.

A classe de filtros autodidatas não utiliza qualquer sinal de referência para executar o processamento do sinal. Ao invés de uma sequência de treinamento ou sinal de referência, esses tipo de algoritmo calcula seus coeficientes com base em alguma característica conhecida do sinal. Como exemplos de algoritmos autodidatas clássicos podem ser citados:

- *Constant Modulus Algorithm* (CMA) [2];
- *Normalized constant-modulus algorithm* (NCMA) [7] ;
- *Blind constrained constant modulus (CCM) adaptive algorithms* [8];
- *Least Square - Constant Modulus Algorithm* (LS-CMA) [9]; e

Quando se conhece informações *a priori* sobre as soluções que se deseja obter, uma outra classe de filtragem pode ser utilizada. Esta classe de filtros, denominada filtragem com restrições, impõe penalidades à função objetivo, de forma que as soluções obtidas incorporem as informações previamente conhecidas. Em geral, as soluções com restrições são obtidas a partir de projeções ortogonais do vetor de coeficientes (sem restrição) sobre o hiperplano da restrição [2] ou, quando existirem mais de uma restrição, a partir de projeções ortogonais sobre a intersecção dos diversos hiperplanos de restrições. Uma abordagem de filtragem adaptativa com restrições bastante conhecida é a denominada *Generalized Sidelobe Canceller* (GSC),

cuja estrutura decompõe e o problema em dois, um com restrições lineares, outro sem restrições, dividindo o vetor de coeficientes a ser calculado em duas componentes [10].

Como exemplo de aplicação prática para filtragem adaptativa com restrições, pode-se citar a solução de um problema de arranjo de sensores [10], cujo conjunto de restrições lineares pode ser empregado para se obter coeficientes que assegurem ganhos nulos em determinadas direções do diagrama de radiação (*beam pattern*) e um ganho unitário na direção de um sinal de interesse específico. A imposição de ganhos nulos em algumas direções se faz necessária, por exemplo, para minimizar sinais interferentes (*jammers*), cujas direções sejam conhecidas. Esta situação acontece na prática quando se deseja suprimir sinais interferentes nos sistemas de comunicações militares táticas [11], que se desenvolvem em um teatro de operações onde coexistem as emissões das Forças aliadas e inimigas. As Forças inimigas podem fazer uso de sistemas interferidores (*jammers* ou *electronic attack*), cuja função é bloquear as emissões de uma determinada rede de comunicações aliada. A forma clássica de mitigar essas fontes de interferência, garantindo o uso do espectro eletromagnético à tropa aliada, se dá por meio de sistemas de comunicações com salto de frequência (*Frequency Hopping - FH*) [11], também denominados sistemas com Medidas de Proteção Eletrônica (MPE) ou *Electronic Protective Measures* (EPM) [11]. As contramedidas clássicas, no entanto, realizadas por meio de MPE, possuem um grande inconveniente que é o aumento da indiscrição eletromagnética, que se dá em decorrência do aumento da taxa de utilização do espectro eletromagnético provocado pelo uso de FH. Esse inconveniente permite precisar a posição da estação emissora, quase sempre, resultando na sua neutralização por meios bélicos. Portanto, o uso de supressão de interferências por meio da síntese dinâmica do diagrama de radiação do arranjo é uma solução adequada e de maior aceitabilidade para este problema.

1.2 Uso de restrições de norma ℓ_1 como um caminho viável para soluções esparsas

A busca de soluções esparsas se dá por diversas razões, mas pode-se elencar que as principais são a robustez estatística, a economia de recursos computacionais e a possibilidade de implementação de circuitos eletrônicos mais econômicos em termos de consumo médio de energia.

Em geral, a obtenção de soluções esparsas se dá por meio de algoritmos que impõem restrições com o propósito de encolher a norma da solução ou pela projeção sobre a Bola ℓ_1 [12]. Intuitivamente, quando se busca esparsidade, pensa-se logo na escolha seletiva de variáveis de interesse. Este problema pode ser resolvido matema-

ticamente por meio de regressão com restrição de norma ℓ_0 ¹ que, na prática, se dá pela redução da cardinalidade da solução, que, em termos de filtragem adaptativa, significa aumentar o número de coeficientes nulos. A solução deste problema, além de não convexa, requer uma busca combinatória em todos os subconjuntos dos k elementos passíveis de pertencer ao conjunto soluções, isto é; $C_k^n = \frac{n!}{k!(n-k)!}$ possibilidades de busca, e, portanto, apresenta complexidade computacional de ordem $\mathcal{O}(e^N)$ [13], o que inviabiliza a solução do problema.

Como alternativa ao emprego de restrições de norma ℓ_0 , pode-se impor uma restrição de norma ℓ_1 ², cuja solução natural resulta em um conjunto esparso de solução. Este problema, conhecido como *Least Absolute Shrinkage and Selection Operator* (LASSO), foi introduzindo inicialmente em [14].

Com o propósito permitir um entendimento intuitivo e o porquê de as restrições de norma ℓ_1 conduzirem a soluções esparsas, será apresentado um exemplo simples em \mathbb{R}^2 , abordando inicialmente uma restrição de norma quadrática, que é o problema de regressão, conhecido como *ridge regression* [15].

$$\min_{\mathbf{w}} \|y - \mathbf{w}^T \mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{w}\|_2^2 \leq \delta, \quad (1.1)$$

onde $\mathbf{w} = [w_1 \ w_2]^T$ e δ é o raio do disco centrado na origem, cujo lugar geométrico contém as soluções possíveis para \mathbf{w} .

A solução para o problema formulado pela Eq. (1.1) possui solução pertencente ao domínio $\|\mathbf{w}\|_2^2 \leq \delta$, onde δ , para $\mathbf{w} \in \mathbb{R}^2$, é o raio do disco centrado na origem dos eixos cartesianos ou raio da Bola³ para espaços vetoriais de dimensões superiores.

Na Fig. 1.2, é ilustrada a representação geométrica para o problema formulado por meio da Eq. (1.1) e, adicionalmente, representa-se também a Bola correspondente à norma $\|\mathbf{w}\|_1$ (Bola ℓ_1), na forma de um quadrado rotacionado de 45° em relação ao eixo das abscissas (losango).

Em decorrência da geometria da Bola ℓ_2 , obter uma solução que apresente um coeficiente exatamente igual a zero é possível, mas tem baixa probabilidade de ocorrência. Se, no entanto, se buscar incessantemente a esparsidade por meio desta regressão, com a diminuição do valor de restrição δ imposta, i.e., por meio do encolhimento da norma ℓ_2 do vetor de coeficientes, somente quando $\delta = 0$, ter-se-á garantido que os valores dos coeficientes serão exatamente iguais a zero. Por outro

¹Rigorosamente, a norma ℓ_0 não é uma norma no sentido estrito, pois não goza de todas as propriedades matemáticas de uma norma. Na verdade, é a cardinalidade de um conjunto ou vetor e se traduz pelo número de elementos não nulos desse conjunto ou vetor.

²A norma unitária (norma ℓ_1) de um vetor é definida como: $\|\mathbf{w}\|_1 = \sum_{i=1}^N |w_i|$, onde o w_i é i -ésima componente desse vetor.

³Num espaço métrico (\mathbf{X}, d) a Bola aberta de raio δ , centrada num ponto \mathbf{p}_0 , é o conjunto de pontos cuja distância a \mathbf{p}_0 , é inferior a δ , isto é, $\mathbf{B}(\mathbf{p}_0, \delta) = \{\mathbf{v} \in \mathbf{X} : d(\mathbf{v}, \mathbf{p}_0) = \|\mathbf{v} - \mathbf{p}_0\|_2 \leq \delta\}$.

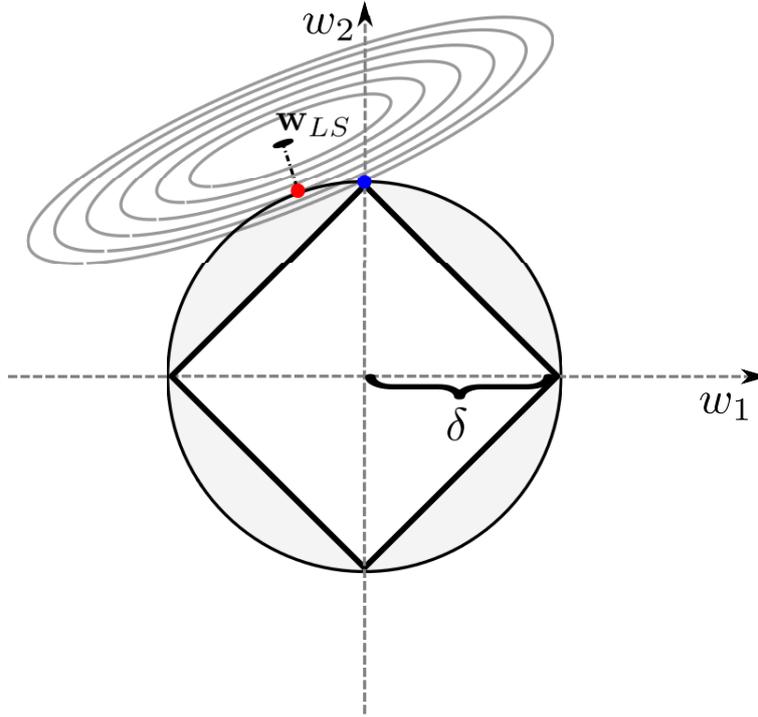


Figura 1.2: Comparação entre as soluções obtidas com os algoritmos LASSO (azul) e Ridge (vermelho) em relação a solução *Least Square* ordinária. As restrições referentes aos algoritmos LASSO e Ridge são representadas pelas Bolas ℓ_1 (losango) e ℓ_2 (círculo), respectivamente.

lado, pode-se reformular o problema empregando a restrição de norma unitária:

$$\min_{\mathbf{w}} \|y - \mathbf{w}^T \mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{w}\|_1 \leq \delta, \quad (1.2)$$

onde $\mathbf{w} = [w_1 \ w_2]^T$ e δ é igual a metade da diagonal do losango, centrado na origem, cujo lugar geométrico contém as soluções possíveis para \mathbf{w} .

Quando o problema é formulado de acordo com a Eq. (1.2), a geometria da Bola ℓ_1 favorece o aparecimento de vetores soluções com alguns coeficientes exatamente iguais a zero. Isto se dá pelo fato de que quando se projeta a solução sem restrições sobre a Bola ℓ_1 , a solução estará localizada em um dos três lugares geométricos do hipercubo, isto é; no vértice, na aresta ou na face. Nas duas primeiras situações, a esparsidade acontece de imediato. Para o caso da projeção sobre a face do hipercubo, mesmo para a situação mais crítica, que ocorre quando o vetor dos coeficientes *Least Square* é ortogonal à face, pode-se notar que uma variação adequada no valor da restrição δ fará com que a nova solução venha pertencer a um vértice ou aresta do hipercubo, conforme ilustrado na Fig. 1.3.

Motivados pelo trabalho apresentado na referência [14], vários autores propuseram diversos algoritmos de filtragem adaptativa esparsa. Dentre os algoritmos propostos nos últimos anos, alguns se baseiam no algoritmo *Least-Mean-Square* (LMS),

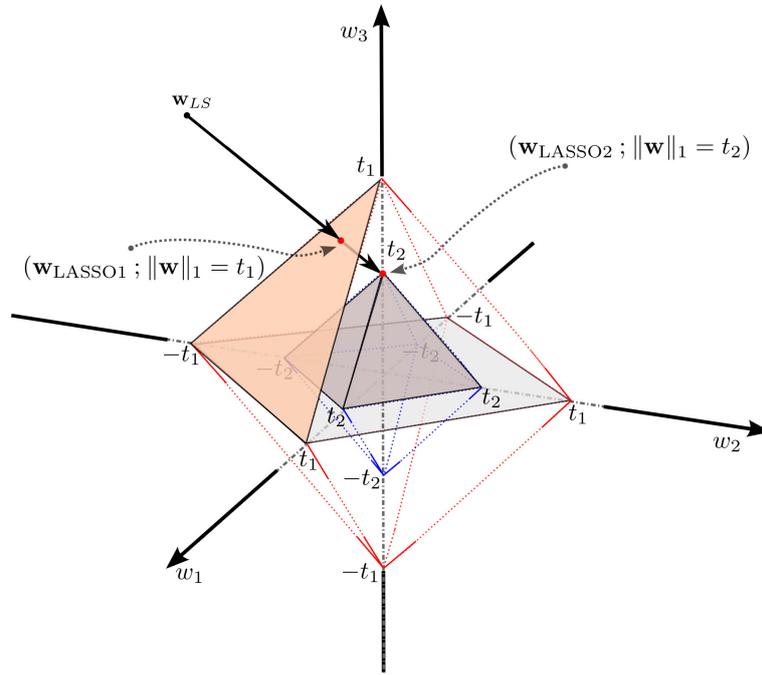


Figura 1.3: Projeção ortogonal de solução *Least Square* sobre um das faces de um hiper-cubo no \mathbb{R}^3 . Comparando-se as solução $\mathbf{w}_{\text{LASSO1}}$ e $\mathbf{w}_{\text{LASSO2}}$, pode-se observar, que a partir de um ajuste adequado na restrição de norma ℓ_1 , a solução final torna-se esparsa.

podendo ser citados os seguintes:

- *The zero-attracting LMS Algorithm* (ZA-LMS) [16];
- *The reweighted zero-attracting LMS* (RZA-LMS) [16];
- *Proportionate Normalized LMS Algorithm* (PNLMS) [17]; e
- *Improved Proportionate Normalized LMS algorithm* (IPNLMS) [18].

O algoritmo ZA-LMS é deduzido a partir da formulação utilizada na dedução do LMS como o acréscimo de uma restrição de norma ℓ_1 à função objetivo quadrática do LMS. O algoritmo RZA-LMS é implementado a partir do ZA-LMS por meio da heurística que substitui a norma ℓ_1 por uma barreira logarítmica e, desta forma, consegue-se uma melhora na performance do algoritmo inicial, quando se considera aplicações envolvendo identificação de sistemas esparsos.

Os algoritmos PNLMS e IPNLMS apresentam rendimento superior ao apresentado pelo algoritmo *Normalized Least Mean Square* (NLMS), quando a aplicação envolve identificação de sistemas esparsos. O emprego desses algoritmos em aplicações como *beamforming* possui o inconveniente de não ser possível fazer uso de restrições lineares e, portanto, não garantir ganho máximo na direção do sinal de interesse.

Uma abordagem importante, que não é baseada no algoritmo LMS e nem faz uso de regularização por meio de função objetivo, foi proposta em [12], denominada *Online Sparse System Identification and Signal Reconstruction Using Projections Onto Weighted ℓ_1 Balls*. Este algoritmo executa, a cada iteração, uma projeção da solução sem restrição sobre a Bola ℓ_1 (ℓ_1 -balls) e, desta forma, garante o atendimento à restrição de norma ℓ_1 .

Algoritmos de Otimização Convexa, como o *Second Order Cone Programming* (SOCP) [19], também podem ser empregados na solução do problema de síntese de arranjos de sensores com restrições lineares e de norma ℓ_1 . Contudo, a despeito de apresentarem maior flexibilidade em relação às restrições, não permitem resolver problemas de arranjos de sensores de forma adaptativa, além de apresentarem grande complexidade computacional em relação aos algoritmos da família LMS, tornando-os inadequados para aplicações em sistemas embarcados, cujas características exijam soluções adaptativas com baixo custo computacional aliadas ao baixo consumo de energia.

Na referência [20], são propostos dois algoritmo adaptativos para o corpo dos números complexos, cujas regularizações de norma ℓ_1 são substituídas por termos re-ponderados, resultando em melhoria de desempenho em termos de taxa de convergência e desajuste. Estes algoritmos são adequados para soluções de problemas como supressão de interferência e arranjos esparsos na presença de sensores com falha.

A aplicação de filtragem adaptativa esparsa à solução de problemas envolvendo arranjos de sensores, por meio da adição de restrições de norma ℓ_1 , ou por outra metodologia qualquer, pode ser vantajosa em razão de dois fatores importantes: o primeiro, que decorre da esparsidade, é o menor consumo de energia, que é viável sempre que a complexidade do cenário envolvendo os sinais de interesse e interferências, permitir uma síntese do *beam pattern* por meio de uma filtragem espacial esparsa; e o segundo se dá pela flexibilidade de autoconfiguração que o sistema irradiante adquire como resultado da adaptabilidade. Além do fato de que, em geral, o emprego de filtragem esparsa apresenta uma taxa de convergência mais rápida quando a planta admite uma solução esparsa.

Especificamente, para o caso onde o número de sensores é bastante elevado, a redução da complexidade computacional e do consumo de energia são parâmetros de projeto com alto grau de importância.

1.3 Propósito deste Trabalho

Este trabalho tem como propósito propor três algoritmos, de uma mesma família, cuja característica principal e comum é a filtragem com restrições lineares e de norma

ℓ_1 . Estes algoritmos se destinam a aplicações envolvendo arranjos adaptativos com um grande número de sensores. Estes algoritmos são baseados na minimização do erro médio quadrático sujeita às restrições lineares e de norma unitária (ℓ_1), conforme ilustrado seguir:

$$\min_{\mathbf{w}} E[|e_k|^2] \quad \text{s.t.} \quad \begin{cases} \mathbf{C}^H \mathbf{w} = \mathbf{z} \\ \|\mathbf{w}\|_1 = t \end{cases} \quad (1.3)$$

onde $e_k = d_k - \mathbf{w}^H \mathbf{x}_k$ é o estimador para o erro, sendo \mathbf{x}_k e d_k o vetor sinal de entrada e sinal desejado, respectivamente. A matriz \mathbf{C} , com dimensão $N \times L$, é denominada matriz de restrições e \mathbf{z} é o vetor de restrições correspondente, cujo número de componentes, L , corresponde ao número de restrições lineares. Para o caso de arranjo de sensores, o sinal desejado d_k , ou de referência, é considerado nulo, i. e., $d_k = 0$, o que resulta na minimização da energia do sinal de saída do arranjo.

Serão apresentados três algoritmos, sendo o primeiro e mais básico deles denominado ℓ_1 -norm Linearly-Constrained Least-Mean-Square (ℓ_1 -CLMS) [21]. Os outros são aos algoritmos correspondentes normalizado e ponderado, denominados ℓ_1 -norm Linearly-Constrained Normalized Least-Mean-Square (ℓ_1 -CNLMS) [22] e ℓ_1 -norm Reweighted Linearly-Constrained Least-Mean-Square (ℓ_1 -WCNLMS), respectivamente.

Um importante diferencial entre os algoritmos propostos por este trabalho e os algoritmos citados na Seção anterior é que, além da esparsidade, a família de algoritmos aqui proposta é também capaz de impor restrições lineares às soluções. Este fato permite uma maior flexibilidade quando da solução de problemas práticos de arranjos de sensores sem no entanto onerar a complexidade computacional, que se mantém da mesma ordem da complexidade apresentada pelo algoritmo LMS.

1.4 Estrutura da tese

Esta tese foi organizada em sete Capítulos e um apêndice, conforme detalhamento a seguir:

- No Capítulo 1, são apresentados os principais assuntos tratados ao longo de todo o trabalho, com introdução dos conceitos basilares sem, no entanto, fazer desenvolvimentos matemáticos ou abordagens teóricas mais elaboradas;
- O Capítulo 2 contém os fundamentos matemáticos e conceitos necessários à modelagem de um arranjo de sensores. Desde o modelo do sinal de entrada, passando pelos conceitos de topologia planar, *beam pattern*, *beamforming*, até a obtenção da solução para o Conformador de Feixe de Variância Mínima com Restrições Lineares (*Linearly Constrained Minimum Variance* -LCMV);

- No Capítulo 3, encontram-se os principais desenvolvimentos e deduções matemáticas das expressões finais dos algoritmos ℓ_1 -CLMS, ℓ_1 -CNLMS e ℓ_1 -WCNLMS. Algoritmos CLMS e CNLMS, considerados importantes para o entendimento da dedução dos algoritmos propostos por este trabalho, também são abordados nesse Capítulo.
- No Capítulo 4, são desenvolvidas as análises de estabilidade de primeira e segunda ordem para o algoritmo ℓ_1 -CNLMS. Decorrente da semelhança existente entre as expressões do algoritmos ℓ_1 -CLMS, ℓ_1 -CNLMS e ℓ_1 -WCNLMS, as análises de estabilidade se restringem ao algoritmo ℓ_1 -CNLMS.
- No Capítulo 5, são apresentados resultados das aplicações dos algoritmos à síntese e a solução de arranjos de sensores, sendo evidenciada a esparsidade obtida para os arranjos de sensores. São também apresentadas comparações, em termos de *beam pattern*, entre os arranjos esparsos obtidos e arranjos uniformes equivalentes.
- No Capítulo 6, são apresentados os resultados de simulações voltadas à comparação de taxa de convergência dos algoritmos ℓ_1 -WCNLMS e ℓ_1 -CNLMS envolvendo a identificação de sistemas esparsos e compressíveis. Neste Capítulo também são apresentadas simulações para validação numérica das expressões aproximadas de $MSE(\infty)$, Excesso de MSE e Desajuste apresentadas na Seção 6.3.
- O Capítulo 7 apresenta as conclusões da tese, sendo também listadas as contribuições do trabalho, bem como possíveis desdobramentos futuros.
- Por fim, no Apêndice, é realizada uma dedução alternativa para o algoritmo ℓ_1 -CNLMS, usando-se a abordagem de *Minimum-disturbance* [23] [24].

1.5 Notação

Nesta Seção serão citadas somente notações e definições de uso extensivo e presentes em todo o corpo deste trabalho. As notações e definições de ocorrência esporádica serão citadas e/ou definidas no momento que aparecerem no texto.

Tabela 1.1: Notação

Notação	Definição
Norma L_p de \mathbf{w}	$\ \mathbf{w}\ _p = \left(\sum_{i=1}^n w_i ^p \right)^{1/p}, 1 \leq p < \infty$
m	escalar denotado por letras minúsculas
\mathbf{v}	vetor representado por letras minúsculas em negrito
\mathbf{A}	matriz representada por letras maiúsculas em negrito
$(\cdot)^H$	operador matricial de transposição conjugada (Hermitiano)
$\text{tr}\{\cdot\}$	operador matricial para função Traço
$E\{\cdot\}$	operador para cálculo do valor esperado
\mathbf{I}_N	matriz identidade de ordem $N \times N$
$\Re(\cdot)$	Parte real de um escalar, vetor ou matriz complexo
$\Im(\cdot)$	Parte imaginária de um escalar, vetor ou matriz complexo

Capítulo 2

Fundamentos de Arranjos de Sensores

As técnicas de processamentos de sinais aplicadas aos arranjos de sensores visam extrair a máxima informação transportada por uma onda por ele captada e idealizada como plana e uniforme. Na prática, os arranjos de sensores atuam como transdutores que transformam ondas (eletromagnéticas ou mecânicas) em sinais descritos em termos de aberturas espaciais e temporais. Em conformidade com a amostragem temporal, que gera o domínio de tempo discreto, a amostragem espacial realizada pelos arranjos de sensores resulta em um domínio denominado espaço discreto e, desta forma, o processamento de sinais aplicado aos arranjos de sensores se caracteriza pelo aspecto de dimensão espaço-tempo.

O processamento responsável pela filtragem no domínio espaço-tempo é denominado conformador de feixe (*beamformer*) e esta técnica apresenta resultados superiores aos de sistemas irradiantes que empregam antenas singelas. Este ganho em desempenho advém da seletividade espacial e do conseqüente ajuste dinâmico do diagrama de radiação (*beam pattern*) do arranjo de sensores.

Por meio de filtragem espaço-temporal com restrições lineares, pode-se ainda selecionar fontes de sinais de interesse e suprimir as demais consideradas como fontes interferentes (*jammers*).

Neste capítulo serão apresentados conceitos matemáticos básicos, contemplando arranjos planares (2D), idealmente formados por sensores isotrópicos e idênticos. A partir das expressões apresentadas, as formulações para arranjos lineares (1D) e espaciais (3D) podem ser obtidas, de forma simples. A abordagem apresentada tem o propósito de construir o arcabouço teórico necessário ao emprego dos algoritmos objeto deste trabalho, não tendo, portanto, a pretensão de apresentar uma revisão completa sobre o assunto arranjo de sensores.

2.1 Modelo do Sinal de Entrada

Considere um arranjo planar, conforme ilustrado pela Fig. 2.1, composto por N antenas (sensores) e sejam L sinais de banda estreita, cujas direções de chegada são caracterizadas pelos seus azimutes (θ_i) e zênites (ϕ_i) [10], matematicamente definidos por $(\theta_1, \phi_1), \dots, (\theta_L, \phi_L)$.

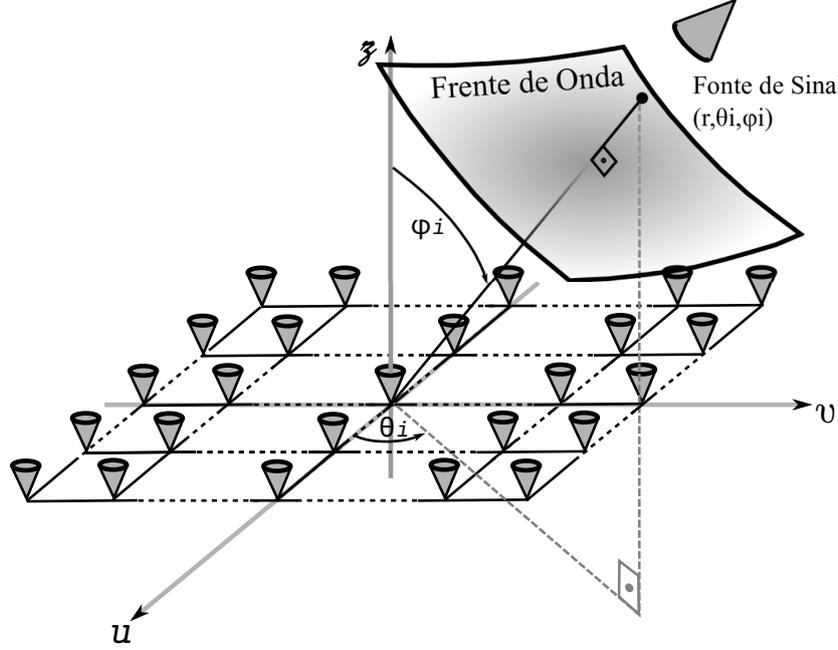


Figura 2.1: Geometria Genérica de um Arranjo Planar.

Então, os sinais observados pelos N sensores, durante durante M amostras (*snapshots*), quando considerados analíticos¹ e discretos no domínio do tempo, podem ser denotados por \mathbf{x}_k , k variando de 1 a M . Portanto, o vetor de sinal \mathbf{x}_k , contendo N elementos para cada valor de k , pode ser escrito como:

$$\mathbf{x}_k = \sum_{i=1}^L \mathbf{a}(\theta_i, \phi_i) \mathfrak{S}_i(k) + \boldsymbol{\eta}(k), \quad k = 1, 2, \dots, M. \quad (2.1)$$

onde $\mathbf{a}(\theta_i, \phi_i)$ representa as colunas da matriz de direcionamento (*steering matrix*) e $\mathfrak{S}_i(k)$ a i -ésima fonte de sinal proveniente da direção (θ_i, ϕ_i) .

Empregando-se a notação matricial, pode-se representar a expressão acima de forma mais compacta, $\mathbf{A}\boldsymbol{\mathfrak{S}}(k) + \boldsymbol{\eta}(k)$, permitindo-se definir uma matriz de dados de entrada \mathbf{X} , de dimensão $N \times M$, como:

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_M] = \mathbf{A}\boldsymbol{\mathfrak{S}} + \boldsymbol{\mathcal{N}}. \quad (2.2)$$

¹Um sinal analítico é uma sequência de valores complexos cuja Transformada de Fourier de Tempo Discreto (DTFT) não possui componentes de frequência negativas. As partes real e imaginária de um sinal analítico são valores reais relacionados entre si pela Transformada de Hilbert [25].

Na Eq. (2.2), a matriz complexa \mathcal{S} , de dimensão $L \times M$, representa o envelope do sinal recebido no centro do arranjo, \mathcal{N} , de dimensão $N \times M$, representa o ruído de observação e $\mathbf{A} = [\mathbf{a}(\phi_1), \dots, \mathbf{a}(\phi_d)]$, de dimensão $N \times L$, é a matriz de direcionamento (*steering matrix*). A função da matriz \mathbf{A} é distribuir para os N sensores os sinais provenientes das L diferentes fontes de sinais, cujas direções (θ_i, ϕ_i) são contabilizadas por meio das diferenças de fase presentes nas colunas da matriz \mathbf{A} :

$$\mathbf{a}(\theta_i, \phi_i) = \exp \left\{ -j \frac{2\pi}{\lambda} \left[\cos(\theta_i) \sin(\phi_i) \mathbf{d}_u + \sin(\theta_i) \sin(\phi_i) \mathbf{d}_v \right] \right\}, \quad i = 1, \dots, L, \quad (2.3)$$

onde \mathbf{d}_u e \mathbf{d}_v são vetores de dimensão $N \times 1$ e representam a posição de cada sensor no arranjo. As componentes de \mathbf{d}_u e \mathbf{d}_v correspondem às coordenadas cartesianas u (abscissa) e v (ordinada), respectivamente, e λ é o comprimento de onda do sinal. Embora a formulação apresentada seja baseada em posicionamento dos sensores sobre uma grade retangular, uma configuração planar qualquer poderá ser implementada por meio da escolha adequada das coordenadas cartesianas.

Para o Arranjo Linear Uniforme (ULA) mostrado na Fig. (2.2), as colunas da matriz de direcionamento \mathbf{A} são calculadas fazendo $\mathbf{d}_v = \mathbf{0}$, $\phi_i = 90^\circ$, $\mathbf{d}_u = [0, d, 2d, \dots, (N-1)d]^T$ e $d = \frac{\lambda}{2}$, portanto:

$$\mathbf{a}(\theta_i) = [1, e^{-j\pi \cos(\theta_i)}, \dots, e^{-j\pi(N-1) \cos(\theta_i)}]^T. \quad (2.4)$$

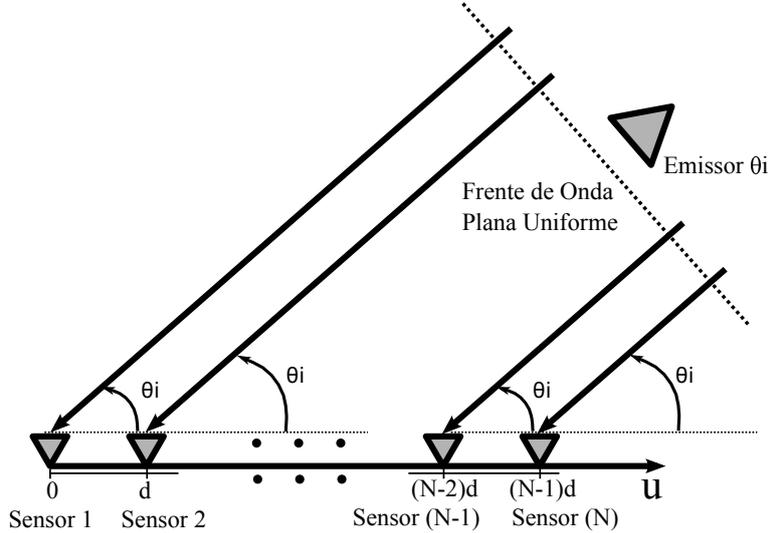


Figura 2.2: Geometria de um Arranjo Linear Uniforme (ULA).

2.2 Diagrama de Radiação do Arranjo de Sensores (*Beam pattern*)

Considerando uma onda plana com comprimento de onda λ , incidindo da direção (θ_s, ϕ_s) e propagando-se através de um arranjo planar, cujos N sensores isotrópicos estão dispostos segundo os pares ordenados $\mathbf{D}_{uv} = [\mathbf{d}_u, \mathbf{d}_v]_{N \times 2}$, pode-se calcular o *beam pattern*, para uma dada direção (θ, ϕ) e um dado vetor de coeficientes \mathbf{w} , como:

$$B(\theta, \phi) = \mathbf{w}^H \exp \left\{ -j \frac{2\pi}{\lambda} \left[\cos(\theta) \sin(\phi) \mathbf{d}_u + \sin(\theta) \sin(\phi) \mathbf{d}_v \right] \right\}. \quad (2.5)$$

Para o arranjo linear uniforme (ULA) apresentado pela Fig. 2.2, pode-se expressar as posições dos sensores como $u_i = (i - 1) \frac{\lambda}{2}$, portanto, a expressão dada pela Eq. (2.5) se reduz a:

$$\begin{aligned} B(\theta) &= \sum_{i=1}^N w_i e^{-j \frac{2\pi}{\lambda} u_i \cos(\theta)} \\ &= \sum_{i=1}^N w_i e^{-j \pi (i-1) \cos(\theta)}, \end{aligned} \quad (2.6)$$

onde w_i é a i -ésima componente do vetor \mathbf{w} .

Na Fig. 2.3 é exemplificado um Diagrama de Radiação para arranjo tipo ULA com 20 sensores e direção de interesse igual a 90° .

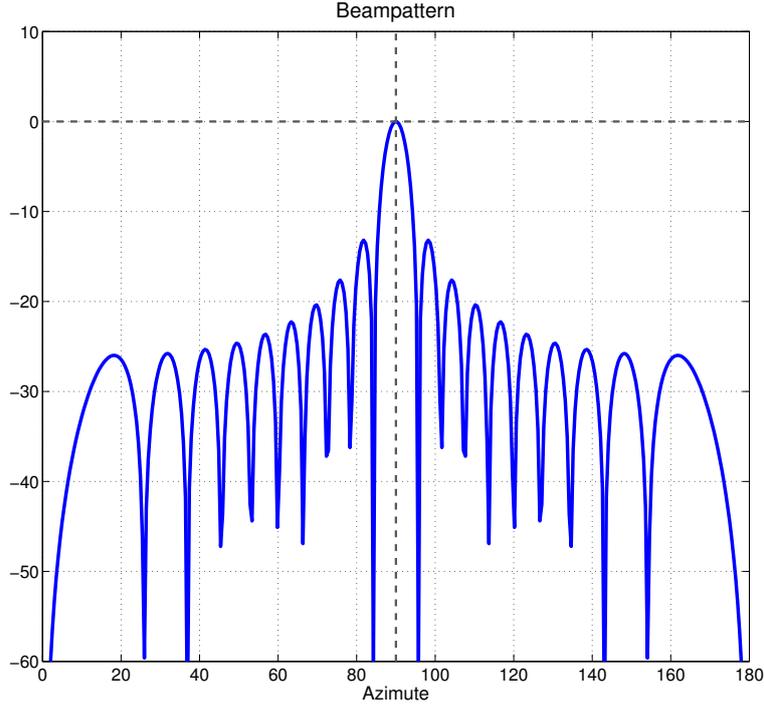


Figura 2.3: Exemplo de Diagrama de Radiação para arranjo tipo ULA com 20 sensores e direção interesse igual a 90° .

2.3 Conformação Adaptativa de Feixe

Conformação Adaptativa de Feixe (*adaptive beamforming*) pode ser definida como o emprego de algoritmos adaptativos [2] para síntese dinâmica de arranjo de sensores, quase sempre buscando maximizar o ganho em uma dada direção de interesse e atenuá-lo nas direções dos sinais interferentes. Para tanto, a técnica empregada é a filtragem espacial, pois embora, os sinais provenientes de diferentes emissores se situem na mesma faixa do espectro, as emissões emanadas por eles, chegam de direções distintas. *A priori*, considera-se que apenas a direção do sinal de interesse é conhecida e portanto, esta informação é usada como uma restrição linear.

A Fig. (2.4) apresenta esquematicamente um *beamformer*. Para o *snapshot* k , a saída do arranjo \mathbf{y}_k é dada pela soma ponderada dos sinais recebidos $\mathfrak{S}_i(k)$ no arranjo e de ruídos $\boldsymbol{\eta}(k)$, conforme Eq. (2.1). Os pesos $\mathbf{w}_k = [w_1(k) w_2(k) w_3(k) \cdots w_N(k)]^T$ são computados de forma a minimizar o erro e_k e a atender as restrições impostas. Quase sempre, o sinal de referência $d(k)$ não está disponível, usa-se então $d(k) = 0$. Desta forma, o algoritmo adaptativo busca minimizar a energia do sinal de saída. No presente modelo, é importante ficar claro que os sinais $\mathfrak{S}_i(k)$ contém tanto o sinal desejado, quanto os sinais interferentes.

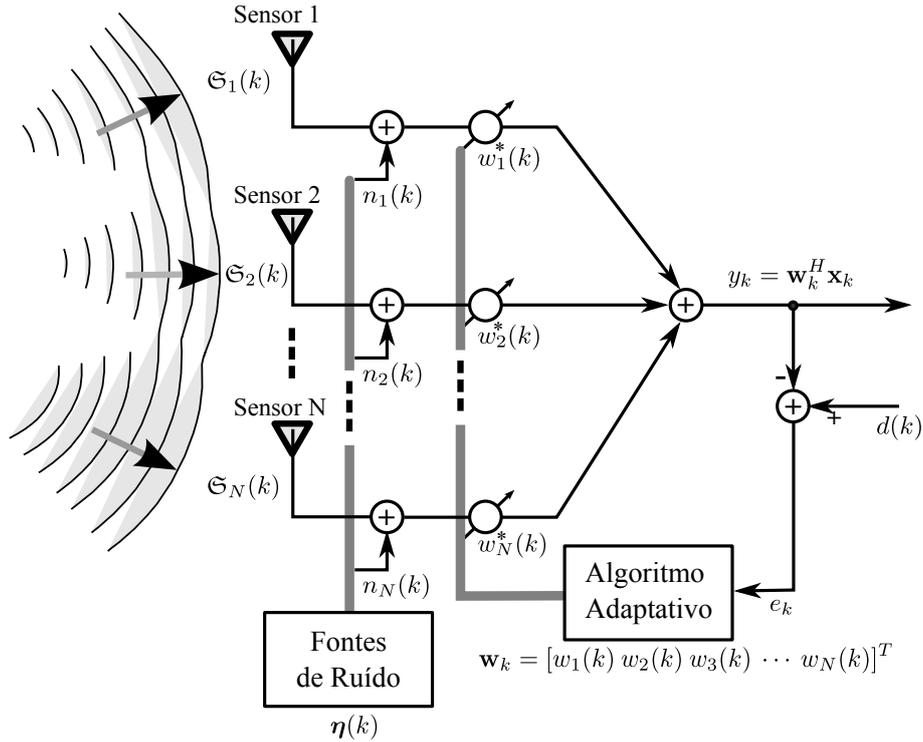


Figura 2.4: Estrutura de *Beamforming* para sinais de banda estreita.

O emprego dos algoritmos adaptativos à solução dos problemas de conformação de feixes é bastante eficiente na atenuação dos sinais interferentes; contudo, de alguma forma, há necessidade de identificar o sinal de interesse, do qual nem sempre se possui informações *a priori*. No caso do sinal de referência não estar disponível, minimiza-se a energia do sinal de saída y_k , ao mesmo tempo que se garante o atendimento às restrições lineares. Alternativamente, podem ser empregados sinais de referência que mantenham algum grau de correlação com o sinal de interesse.

Na Seção 2.4, é apresentada a solução denominada *Linearly Constrained Minimum Variance* (LCMV), cujo resultado é um vetor de coeficientes ótimos que minimiza a variância da saída do filtro sujeito às restrições lineares. A solução LCMV, no entanto, não é adequada para a maioria das aplicações adaptativas, conforme explicado seguir.

2.4 Conformador de Feixe de Variância Mínima Com Restrições Lineares

Da formulação apresentada nas seções anteriores, pode-se observar que um *beamforming* de mínima variância, ou *Linearly Constrained Minimum Variance* (LCMV), é obtido minimizando-se o valor médio da energia do sinal na saída do *beamforming* ao mesmo tempo que as restrições lineares são satisfeitas. Matematicamente, o vetor

ótimo dos coeficientes \mathbf{w}_{LCMV} [10] é obtido resolvendo-se:

$$\min_{\mathbf{w}} E[\mathbf{w}^H \mathbf{x} \mathbf{x}^H \mathbf{w}] = \mathbf{w}^H \mathbf{R} \mathbf{w}, \text{ s.t. } \mathbf{C}^H \mathbf{w} = \mathbf{z}, \quad (2.7)$$

sendo \mathbf{R} a matriz de covariância do sinal de entrada, e \mathbf{C} a matriz de restrições lineares de dimensão $N \times L$ e \mathbf{z} o vetor de restrições correspondente, que possui L componentes, onde L corresponde ao número de restrições.

Fazendo-se uso da técnica de multiplicadores de Lagrange, tem-se:

$$\xi(\mathbf{w}) = \mathbf{w}^H \mathbf{R} \mathbf{w} + \boldsymbol{\lambda}^H (\mathbf{C}^H \mathbf{w} - \mathbf{z}) \quad (2.8)$$

Após avaliar o gradiente complexo da Eq. (2.8) em relação a \mathbf{w} e igualar o seu resultado a zero, pode-se fazer uso das das restrições lineares para obter o sistema:

$$\begin{cases} \frac{\partial \xi(\mathbf{w})}{\partial \mathbf{w}} = \mathbf{R}^T \mathbf{w}^* + \mathbf{C}^* \boldsymbol{\lambda}^* = \mathbf{0} \\ \mathbf{C}^H \mathbf{w} = \mathbf{z}, \end{cases} \quad (2.9)$$

e, então, calcular o valor de $\boldsymbol{\lambda}$ como:

$$\boldsymbol{\lambda} = -\left(\mathbf{C}^H \mathbf{R}^{-1} \mathbf{C}\right)^{-1} \mathbf{z}. \quad (2.10)$$

Substituindo o resultado dado pela Eq. (2.10) na primeira equação do sistema (2.9), chega-se à solução LCMV:

$$\mathbf{w}_{LCMV} = \mathbf{R}^{-1} \mathbf{C} \left(\mathbf{C}^H \mathbf{R}^{-1} \mathbf{C}\right)^{-1} \mathbf{z}, \quad (2.11)$$

tal que $\mathbf{C}^H \mathbf{w}_{LCMV} = \mathbf{z}$, sendo as colunas da matriz \mathbf{C} , para o caso de um ULA, expressas como [2]:

$$\mathbf{c}_i = [1, e^{-j(2\pi/\lambda)d \cos(\theta_i)}, \dots, e^{-j(2\pi/\lambda)d(N-1) \cos(\theta_i)}]^T, \quad (2.12)$$

$$i = 1, \dots, L.$$

A solução apresentada pela Eq. (2.11) é ótima em relação a minimização da energia do sinal de saída. Na Eq. (2.12), o ângulo θ_i indica a direção do sinal de interesse e dos sinais de interferência. Estes sinais são levados em consideração como restrições que, matematicamente, são representadas pelas colunas da matriz de restrição \mathbf{C} e pelas linhas do vetor de restrições \mathbf{z} . Os elementos do vetor \mathbf{z} valem 1 para as direções de interesse e 0 para as direções correspondentes aos sinais interferentes.

A matriz de covariância \mathbf{R} , definida como $E[\mathbf{x}_k \mathbf{x}_k^H]$, pode ser estimada como:

$$\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^H = \mathbf{A}\mathcal{S}\mathcal{S}^H\mathbf{A}^H + \mathcal{N}\mathcal{N}^H; \quad (2.13)$$

porém, é usual estimar a matriz de covariância como uma média temporal [2],

$$\hat{\mathbf{R}} = \frac{1}{M} \sum_{k=1}^M \mathbf{x}_k \mathbf{x}_k^H, \quad (2.14)$$

a qual, à menos de uma constante multiplicativa, corresponde à Eq. (2.13).

Para o cálculo do vetor de coeficientes por meio da Eq. (2.11), é necessário estimar a matriz $\hat{\mathbf{R}}^{-1}$, que na prática, para valores grandes de N , é mal condicionada, além de necessitar de um número de multiplicações proporcional ao cubo do número de coeficientes, i.e., $\mathcal{O}(N^3)$. Como alternativa para superar essa complexidade computacional, são empregados algoritmos adaptativos cuja complexidade computacional, em alguns casos, chega a $\mathcal{O}(N)$.

Capítulo 3

Algoritmos da Família ℓ_1 -CLMS

Neste Capítulo são apresentadas as deduções das expressões finais para algoritmos da família denominada ℓ_1 -norm *Linearly-Constrained Least-Mean-Square* (ℓ_1 -CLMS). Esta família é composta pelos algoritmos ℓ_1 -CLMS, ℓ_1 -CNLMS e ℓ_1 -WCLMS, conforme já mencionado na Seção 1.3.

Os algoritmos propostos por este trabalho visam às restrições lineares e à esparsidade, o que resulta em uma maior flexibilidade quando da aplicação a arranjos de sensores e em um melhor rendimento quando se trata de identificação de sistemas esparsos.

3.1 Revisão: Algoritmos CLMS e CNLMS

Como revisão e servindo como ponto de partida para as deduções dos algoritmos ℓ_1 -CLMS, ℓ_1 -CNLMS, são apresentadas as deduções dos algoritmos CLMS [4] e CNLMS [5].

3.1.1 Algoritmo CLMS

O algoritmo CLMS foi proposto inicialmente em [4] para solucionar problemas relacionados a arranjos de antenas (sensores) em tempo de execução, cuja solução deveria contemplar a direção do sinal de interesse e ao mesmo tempo, atenuar sinais decorrentes de interferências provenientes de outras direções. Esse algoritmo pode ser deduzido por meio do uso da técnica de multiplicadores de Lagrange [2].

Seja $e_k = d_k - \mathbf{w}^H \mathbf{x}_k$ o erro estimado, onde \mathbf{x}_k e d_k são, respectivamente, o sinal de entrada e o sinal desejado; então, o algoritmo CLMS pode ser formulado como

$$\min_{\mathbf{w}} E[|e_k|^2] \quad \text{s.t.} \quad \mathbf{C}^H \mathbf{w} = \mathbf{z}, \quad (3.1)$$

sendo \mathbf{C} a matriz de restrições de dimensão $N \times L$ e \mathbf{z} o vetor de restrições corres-

pondente, que possui L componentes, onde L corresponde ao número de restrições.

Fazendo-se uso da técnica de multiplicadores de Lagrange, tem-se:

$$\xi(\mathbf{w}) = E[|e_k|^2] + \boldsymbol{\lambda}_1^H (\mathbf{C}^H \mathbf{w} - \mathbf{z}). \quad (3.2)$$

Desta forma, a função custo $\xi(\mathbf{w})$ pode ter seu valor minimizado por meio da aplicação do método *Steepest Descent* [3], permitindo que a cada iteração os valores dos coeficientes (solução) possam ser atualizados de acordo com:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \frac{\mu}{2} \nabla_{\mathbf{w}} \xi(\mathbf{w}). \quad (3.3)$$

Infelizmente, este algoritmo não é viável sem que se conheça o valor esperado $E[e_k^* \mathbf{x}_k]$. Para muitos sistemas, a média estatística pode ser aproximada pelo estimador temporal

$$\hat{E}[e_k^* \mathbf{x}_k] = \frac{1}{N} \sum_{i=0}^{N-1} e^*(k-i) \mathbf{x}(k-i), \quad (3.4)$$

onde N representa o número de amostras disponíveis para a estimativa. Uma simplificação seria assumir que apenas uma amostra está disponível ($N = 1$), i. e., $\hat{E}[e_k^* \mathbf{x}_k] \approx e_k^* \mathbf{x}_k$, implicando $\nabla_{\mathbf{w}} \xi(\mathbf{w}) \approx \hat{\nabla}_{\mathbf{w}} \xi(\mathbf{w})$ e, portanto

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \frac{\mu}{2} \hat{\nabla}_{\mathbf{w}} \xi(\mathbf{w}). \quad (3.5)$$

Empregando as devidas manipulações algébricas e encontrado-se o valor de $\boldsymbol{\lambda}_1$ para $\hat{\nabla}_{\mathbf{w}} \xi(\mathbf{w}) = -2e_k^* \mathbf{x}_k + \mathbf{C} \boldsymbol{\lambda}_1 = 0$, obtém-se:

$$\mathbf{w}_{k+1} = \mathbf{P} [\mathbf{w}_k + \mu e_k^* \mathbf{x}_k] + \mathbf{f} \quad (3.6)$$

onde $\mathbf{P} = \mathbf{I}_N - \mathbf{C} (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H$ é uma matriz auxiliar de projeção e $\mathbf{f} = \mathbf{C} (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{z}$ é o vetor com $N \times 1$ componentes responsável pela projeção da solução sem restrições, obtida como $\mathbf{P} [\mathbf{w}_k + \mu e_k^* \mathbf{x}_k]$, sobre o hiperplano das restrições.

3.1.2 Algoritmo CNLMS

O algoritmo CNLMS foi proposto inicialmente em [5] como um versão normalizada do algoritmo CLMS, cujo propósito primário seria o cancelamento de interferências em sistemas de comunicações móveis que fazem uso de Acesso por Multiplexação por Divisão de Código com Sequência Direta (DS-SS).

A dedução deste do algoritmo se dá a partir da expressão obtida para o algoritmo CLMS, Eq. (3.6). Para tanto, por se tratar de um algoritmo normalizado, ter-se-á que calcular o passo de convergência instantâneo μ_k , garantindo que o erro a

a posteriori instantâneo permaneça nulo para quaisquer valores de k [2].

O erro *a posteriori* instantâneo é calculado da Eq. (3.6) como:

$$\begin{aligned}
e_{ap}(k) &= d_k - \mathbf{w}_{k+1}^H \mathbf{x}_k \\
&= [d_k - (\mathbf{P}\mathbf{w}_k + \mathbf{f})^H \mathbf{x}_k] - \mu_k e_k (\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k) \\
&= (d_k - y_k) - \mu_k e_k (\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k) \\
&= e_k (1 - \mu_k \mathbf{x}_k^H \mathbf{P} \mathbf{x}_k)
\end{aligned} \tag{3.7}$$

Portanto, deve-se calcular o valor do passo de convergência que minimize o erro quadrático *a posteriori* instantâneo, conforme

$$\frac{\partial [e_{ap}(k)e_{ap}^*(k)]}{\partial \mu_k^*} = 0, \tag{3.8}$$

$$\frac{\partial [e_{ap}(k)e_{ap}^*(k)]}{\partial \mu_k^*} = \frac{e_{ap}^*(k)}{2} \left\{ \left[\frac{\partial [e_{ap}(k)]}{\partial [\Re(\mu_k^*)]} + j \frac{\partial [e_{ap}(k)]}{\partial [\Im(\mu_k^*)]} \right] \right\}, \tag{3.9}$$

onde $\Re(\mu_k^*)$ and $\Im(\mu_k^*)$ são as partes real e imaginária de μ_k^* , respectivamente.

Resolvendo a Eq. (3.8) para μ_k , chega-se a

$$\mu_k = \frac{\mu_0}{\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k + \gamma}, \tag{3.10}$$

onde μ_0 é um fator de convergência fixo introduzido na expressão para propiciar o controle do desajuste e γ é uma parcela que visa evitar *overflow* quando o valor da energia $\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k$ se tornar muito pequeno.

Uma forma mais simples para encontrar o valor de μ_k , seria resolver diretamente a equação $e_{ap}(k) = 0$, uma vez que a matriz \mathbf{P} é simétrica e positiva semidefinida implicando $\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k \geq 0 \forall k$. Contudo, preferiu-se apresentar a solução mais formal dada pela Eq. (3.8).

A expressão final para o algoritmo CNLMS [5] é portanto:

$$\boxed{\mathbf{w}_{k+1} = \mathbf{P} \left[\mathbf{w}_k + \frac{\mu_0 e_k^*}{\mathbf{x}_k^H \mathbf{P} \mathbf{x}_k + \gamma} \mathbf{x}_k \right] + \mathbf{f}}. \tag{3.11}$$

3.2 Algoritmo ℓ_1 -CLMS

O algoritmo ℓ_1 -norm *Linearly-Constrained Least-Mean-Square* (ℓ_1 -CLMS) [21] é deduzido a partir do algoritmo LMS por meio da imposição de restrições lineares e de norma ℓ_1 . Para para tanto, deve-se definir primeiramente a função custo a ser minimizada com as respectivas restrições.

$$\min_{\mathbf{w}} E[|e_k|^2] \quad \text{s.t.} \quad \begin{cases} \mathbf{C}^H \mathbf{w} = \mathbf{z} \\ \|\mathbf{w}\|_1 = t \end{cases} \quad (3.12)$$

$$\xi(\mathbf{w}) = E[|e_k|^2] + \boldsymbol{\lambda}_1^H (\mathbf{C}^H \mathbf{w} - \mathbf{z}) + \lambda_2 (\|\mathbf{w}\|_1 - t) \quad (3.13)$$

Então, precisa-se encontrar o gradiente de $\xi(\mathbf{w})$, igualá-lo a zero e isolar os valores de $\boldsymbol{\lambda}_1$ e λ_2 :

$$\begin{aligned} \nabla_{\mathbf{w}} \xi(\mathbf{w}) &= \nabla_{\mathbf{w}} \left\{ E[|e_k|^2] + \boldsymbol{\lambda}_1^H (\mathbf{C}^H \mathbf{w} - \mathbf{z}) + \lambda_2 (\|\mathbf{w}\|_1 - t) \right\} \\ &= -2E[e_k \nabla_{\mathbf{w}} e_k] + \mathbf{C} \boldsymbol{\lambda}_1 + \lambda_2 \text{sign}(\mathbf{w}), \end{aligned} \quad (3.14)$$

onde $\text{sign}[w] \triangleq w/|w|$ para $w \in \mathbb{C}$.

Considerando o estimador temporal instantâneo para o gradiente, conforme *modus operandi* empregado na Seção 3.1.1, e expressando o erro a cada instante k como sendo $e_k = d_k - \mathbf{w}^H \mathbf{x}_k$, onde \mathbf{x}_k representa o sinal de entrada e d_k o sinal desejado, tem-se:

$$\hat{\nabla}_{\mathbf{w}} \xi(\mathbf{w}) = -2e_k^* \mathbf{x}_k + \mathbf{C} \boldsymbol{\lambda}_1 + \lambda_2 \text{sign}(\mathbf{w}). \quad (3.15)$$

Substituindo a Eq. (3.15) na Eq. (3.3), tem-se:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \frac{\mu}{2} \left\{ -2e_k^* \mathbf{x}_k + \mathbf{C} \boldsymbol{\lambda}_1 + \lambda_2 \text{sign}[\mathbf{w}(k)] \right\}. \quad (3.16)$$

De maneira a simplificar a notação, doravante, será empregada a definição $\mathbf{s}_k = \text{sign}[\mathbf{w}]$, tal que $\|\mathbf{w}\|_1 = \mathbf{s}_k^H \mathbf{w}$.

Da restrição linear imposta na Eq. (3.12), pode-se expressar $\mathbf{C}^H \mathbf{w}_{k+1} = \mathbf{z}$ e, desta forma, multiplicando-se a Eq. (3.16) por \mathbf{C}^H , obtém-se:

$$\boldsymbol{\lambda}_1 = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \left(2e_k^* \mathbf{x}_k - \lambda_2 \mathbf{s}_k \right). \quad (3.17)$$

Considerando que próximo da convergência os vetores \mathbf{w}_k e \mathbf{w}_{k+1} pertencerão ao mesmo hiperquadrante, propõe-se como válida a aproximação $\mathbf{s}_k^H \mathbf{w}_{k+1} \approx t$. Usando esta aproximação e definindo $t_k = \mathbf{s}_k^H \mathbf{w}_k$, pode-se eliminar os vetores \mathbf{w}_{k+1} e \mathbf{w}_k na Eq. (3.16) multiplicando-a por \mathbf{s}_k^H :

$$\begin{aligned} t &= t_k - \frac{\mu}{2} \left\{ -e_k^* \mathbf{s}_k^H \mathbf{x}_k + \mathbf{s}_k^H \mathbf{C} \boldsymbol{\lambda}_1 \right. \\ &\quad \left. + \lambda_2 \mathbf{s}_k^H \mathbf{s}_k \right\}, \end{aligned} \quad (3.18)$$

usando o fato que $\mathbf{s}_k^H \mathbf{s}_k = N$, após isolar λ_2 na Eq. (3.18), chega-se a

$$\lambda_2 = \left(\frac{-2}{N\mu} \right) e_{L_1}(k) + \frac{2}{N} e_k^* \mathbf{s}_k^H \mathbf{x}_k - \frac{1}{N} \mathbf{s}_k^H \mathbf{C} \boldsymbol{\lambda}_1, \quad (3.19)$$

onde $e_{L_1}(k) = t - t_k$.

Para se chegar a expressão final para o algoritmo ℓ_1 -CLMS, deve-se resolver o sistema formado pelas Eqs. (3.17) and (3.19), obter os valores dos multiplicadores de Lagrange $\boldsymbol{\lambda}_1$ and λ_2 e substitui-los na Eq. (3.16). Após realizar essa resolução, obtêm-se

$$\boxed{\mathbf{w}_{k+1} = \mathbf{P} [\mathbf{w}_k + \mu e_k^* \bar{\mathbf{P}}_k \mathbf{x}_k] + \mathbf{f} + \mathbf{f}_{L_1}(k)}, \quad (3.20)$$

com

$$\begin{cases} \bar{\mathbf{P}}_k = \left[\mathbf{I} - \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right) \mathbf{s}_k^H \right] \mathbf{P}, \\ e_{L_1}(k) = t - \mathbf{s}_k^H \mathbf{w}_k, \\ \mathbf{f}_{L_1}(k) = e_{L_1}(k) \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right). \end{cases}$$

O algoritmo ℓ_1 -CLMS consta da estrutura resumida em Algoritmo 1.

Algorithm 1 : Algoritmo ℓ_1 -CLMS

Inicialização:

$k \leftarrow 1$;

$t \leftarrow t_{opt}$;

$\mu \in [0, 2]$

$\gamma = 1e - 10$;

$\mathbf{P} \leftarrow \mathbf{I}_N - \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H$;

$\mathbf{f} \leftarrow \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{z}$;

$\mathbf{w}(1) \leftarrow \mathbf{f}$;

while ($k < K_{max}$) **do**

$e_k \leftarrow d_k - \mathbf{w}_k^H \mathbf{x}_k$;

$e_{L_1}(k) \leftarrow t - \|\mathbf{w}_k\|_1$;

$\mathbf{s}_k \leftarrow \text{sign}(\mathbf{w}_k)$;

$\bar{\mathbf{P}}_k \leftarrow \left[\mathbf{I}_N - \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right) \mathbf{s}_k^H \right] \mathbf{P}$;

$\mathbf{f}_{L_1} \leftarrow e_{L_1}(k) \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right)$;

$\mathbf{w}_{k+1} \leftarrow \mathbf{P} [\mathbf{w}_k + \mu e_k^* \bar{\mathbf{P}}_k \mathbf{x}_k] + \mathbf{f} + \mathbf{f}_{L_1}(k)$;

$k \leftarrow k + 1$;

end while

3.3 Algoritmo ℓ_1 -CNLMS

O algoritmo ℓ_1 -CNLMS [22] é a versão normalizada do ℓ_1 -CLMS [21]. A versão normalizada, por sua vez, apresenta um desempenho ligeiramente superior, sendo esta a motivação para o seu desenvolvimento.

A exemplo do desenvolvimento apresentado na Seção 3.1.2, a dedução deste algoritmo se dá a partir da expressão obtida para o algoritmo ℓ_1 -CLMS, da pela Eq. (3.20). Para tanto, por se tratar de um algoritmo normalizado, deve-se calcular o passo de convergência instantâneo μ_k , garantindo que o erro *a posteriori* instantâneo permaneça nulo para quaisquer valores de k .

O erro *a posteriori* instantâneo é calculado da Eq. (3.20) como:

$$\begin{aligned}
 e_{ap}(k) &= d_k - \mathbf{w}_{k+1}^H \mathbf{x}_k \\
 &= [d_k - (\mathbf{P}\mathbf{w}_k + \mathbf{f})^H \mathbf{x}_k] - \mu_k^* e_k (\mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k) \\
 &= (d_k - y_k) - \mu_k e_k \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k - \mathbf{f}_{L1}^H(k) \mathbf{x}_k \\
 &= e_k (1 - \mu_k \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k) - \mathbf{f}_{L1}^H(k) \mathbf{x}_k
 \end{aligned} \tag{3.21}$$

Portanto, deve-se calcular o valor do passo de convergência que minimize o erro quadrático *a posteriori* instantâneo:

$$e_{ap}(k) = 0 \implies \mu_k^* = \frac{\mu_0 [e_k - \mathbf{f}_{L1}^H(k) \mathbf{x}_k]}{e_k \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k + \gamma}, \tag{3.22}$$

onde μ_0 é um fator de convergência fixo introduzido na expressão para propiciar o controle do desajuste e γ é uma parcela que visa evitar *overflow* quando o valor da parcela $e_k \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k$ se tornar muito pequeno.

Substituindo o valor instantâneo do passo de convergência calculado pela Eq. (3.22) na Eq. (3.20), obtêm-se a expressão final para o algoritmo ℓ_1 -CNLMS:

$$\boxed{\mathbf{w}_{k+1} = \mathbf{P} \left\{ \mathbf{w}_k + \left[\frac{\mu_0 e_k^* (e_k^* - \mathbf{x}_k^H \mathbf{f}_{L1}(k))}{e_k^* \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k + \gamma} \right] \bar{\mathbf{P}}_k \mathbf{x}_k \right\} + \mathbf{f} + \mathbf{f}_{L1}(k)}. \tag{3.23}$$

Uma solução alternativa, que evitaria o erro no denominador da expressão de μ_k , seria não zerar o erro *a posteriori*, aproximando a Eq. (3.22) por:

$$e_{ap}(k) = -\mathbf{f}_{L1}^H(k) \mathbf{x}_k \implies \mu_k \approx \frac{\mu_0}{\mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k + \gamma}. \tag{3.24}$$

O algoritmo ℓ_1 -CNLMS também pode ser deduzido por meio da abordagem do Distúrbio Mínimo (*Minimum-disturbance approach*) [23] [24].

A despeito de escrita com outras variáveis, a expressão final para o algoritmo nor-

malizado obtida por meio da abordagem do Distúrbio Mínimo é idêntica à Eq. (3.23).

$$\boxed{\mathbf{w}_{k+1} = \mathbf{P} \left[\mathbf{w}_k + \frac{\mu_0}{\tau_k} e_k^* \bar{\mathbf{G}}_k \mathbf{x}_k \right] + \mathbf{f} + \mathbf{f}_{L_1}^{MD}(k)}, \quad (3.25)$$

onde

$$\begin{cases} \tau_k = \mathbf{x}_k^H \mathbf{P} \mathbf{x}_k \\ \mathbf{G}_k = \mathbf{P} [\mathbf{x}_k \mathbf{x}_k^H - \tau_k \mathbf{I}] \mathbf{P} \\ D_k = \mathbf{s}_k^H \mathbf{G}_k \mathbf{s}_k \\ \bar{\mathbf{G}}_k = [\mathbf{I} - (\frac{\mathbf{G}_k \mathbf{s}_k}{D_k}) \mathbf{s}_k^H] \mathbf{P}, \\ \mathbf{f}_{L_1}^{MD}(k) = e_{L_1}(k) (\frac{\mathbf{G}_k \mathbf{s}_k}{D_k}). \end{cases}$$

No Apêndice A, é demonstrada, em detalhes, a identidade entre as Eqs. (3.23) e (3.25).

3.4 Algoritmo ℓ_1 -WCNLMS

O algoritmo ℓ_1 -WCNLMS é a versão ponderada do algoritmo ℓ_1 -CNLMS [22]. Esta versão incorpora o conceito de ponderação [26] que é uma forma eficiente de melhorar o desempenho de algoritmos quando o problema envolvido é identificação de sistemas esparsos. Para entender conceitualmente como a ponderação melhora o desempenho de algoritmos esparsos, devemos ter em mente que, em geral, quando um algoritmo com restrições de norma ℓ_1 , como o ℓ_1 -CNLMS ou ZA-LMS [16], a cada iteração, tentar promover esparsidade por meio do encolhimento dos valores absolutos dos coeficientes do filtro, ele o faz de maneira uniforme para todos os coeficientes, não distinguindo os coeficientes esparsos dos não esparsos. Portanto, quando o sistema for esperso, o rendimento do algoritmo, em termos de MSE, é degradado.

Como base na fundamentação teórica apresentada pelos trabalhos [16] [20] [26], pode-se formular um algoritmo baseado no ℓ_1 -CNLMS, com as mesmas características em termos de restrição e melhor desempenho quando da identificação de sistemas esparsos. Considere o problema proposto como:

$$\min_{\mathbf{w}} E[|e_k|^2] \quad \text{s.t.} \quad \begin{cases} \mathbf{C}^H \mathbf{w} = \mathbf{z}; \\ \|\mathcal{P} \mathbf{w}\|_1 = t, \end{cases} \quad (3.26)$$

onde \mathcal{P} é uma matriz diagonal, cujos elementos p_i são responsáveis por ponderar os coeficientes do vetor \mathbf{w} . De imediato, surge a pergunta: para quais pesos p_i o rendimento do algoritmo será melhorado? Uma possibilidade é fazer uso de pesos cujos valores sejam inversamente proporcionais aos valores absolutos dos coeficientes

do filtro [26]:

$$p_i = \begin{cases} \frac{1}{|w_i|}, & w_i \neq 0 \\ \infty, & w_i = 0. \end{cases} \quad (3.27)$$

Se a solução \mathbf{w} possuir cardinalidade \mathcal{C} , que implica $\|\mathbf{w}\|_0 \leq \mathcal{C}$, então o problema formulado pela Eq. (3.26) terá solução, para encolhimento dos pesos segundo a Eq. (3.27), se existir um número de elementos não esparsos \mathcal{N} , tal que $\mathcal{N} \geq \mathcal{C}$. Por sua vez, considerando o produto de \mathcal{P} e \mathbf{w} , observa-se que maiores valores de p_i forçam a solução \mathbf{w} se concentrar nos índices cujos valores absolutos são pequenos e, por construção, correspondem precisamente aos índices cujos coeficientes não são nulos. Resumidamente, no exemplo acima, pesos de maiores valores favorecem a esparsidade, enquanto que pesos de menor magnitude contribuem para a diminuir o grau de esparsidade da solução. Como efeito da promoção seletiva de esparsidade, tem-se o aumento do rendimento do algoritmo, em termos de velocidade de convergência, e a diminuição da polarização.

Na prática, existem diversas alternativas à escolha de pesos expressa na Eq. (3.27). Neste trabalho, a função de pesos adotada é a função $\phi_1(\mathbf{w}) = \sum_{i=1}^N \arctan(\beta|w_i|)$:

$$\min_{\mathbf{w}} E[|e_k|^2] \quad \text{s.t.} \quad \begin{cases} \mathbf{C}^H \mathbf{w} = \mathbf{z}; \\ \frac{2}{\pi} \sum_{i=1}^N \arctan(\beta|w_i|) = t. \end{cases} \quad (3.28)$$

onde β é o fator que determina a taxa crescimento da função $\phi_1(\mathbf{w})$, servindo, desta forma, como parâmetro para ajuste de taxa de convergência do algoritmo em relação à esparsidade do sistema sob identificação.

Com a escolha de $\phi_1(\mathbf{w}) = \sum_{i=1}^N \arctan(\beta|w_i|)$, o que se busca, intrinsecamente, é o aumento da taxa de crescimento do gradiente nas proximidades da origem. Neste caso, a taxa de crescimento do gradiente $\nabla_{\mathbf{w}} \phi_1(\mathbf{w})$ é inversamente proporcional ao quadrado dos valores absolutos dos coeficientes do filtro, portanto, o algoritmo apresentará uma melhoria de rendimento para as soluções esparsas.

Para a dedução do algoritmo ℓ_1 -WCNLMS, pode-se fazer uso dos resultados obtidos na Seção 3.3 para o algoritmo ℓ_1 -CNLMS, bastando para tanto, substituir o valor de \mathbf{s}_k na expressão final do algoritmo pelo valor do gradiente da função $\phi_1(\mathbf{w})$.

$$\begin{aligned} \nabla_{\mathbf{w}} \phi_1(\mathbf{w}) &= \nabla_{\mathbf{w}} \left\{ \frac{2}{\pi} \sum_{j=1}^N \arctan(\beta|w_j|) \right\} \\ &= \frac{2\beta}{\pi} \left[\frac{\text{sign}[w_1]}{\beta^2|w_1|^2 + 1} \quad \frac{\text{sign}[w_2]}{\beta^2|w_2|^2 + 1} \quad \cdots \quad \frac{\text{sign}[w_N]}{\beta^2|w_N|^2 + 1} \right]^T \end{aligned} \quad (3.29)$$

Portanto a expressão final para o algoritmo é dada por

$$\boxed{\mathbf{w}_{k+1} = \mathbf{P} \left[\mathbf{w}_k + \mu e_k^* \bar{\mathbf{P}}_k \mathbf{x}_k \right] + \mathbf{f} + \mathbf{f}_{L_1}(k)}, \quad (3.30)$$

com:

$$\begin{cases} \mathbf{s}_k = \frac{2\beta}{\pi} \left[\frac{\text{sign}[w_1]}{\beta^2|w_1|^2 + 1} \frac{\text{sign}[w_2]}{\beta^2|w_2|^2 + 1} \cdots \frac{\text{sign}[w_N]}{\beta^2|w_N|^2 + 1} \right]^T, \\ e_{L_1}(k) = t - \mathbf{s}_k^H \mathbf{w}_k, \\ \bar{\mathbf{P}}_k = \left[\mathbf{I} - \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right) \mathbf{s}_k^H \right] \mathbf{P}, \\ \mathbf{f}_{L_1}(k) = e_{L_1}(k) \left(\frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^2} \right). \end{cases}$$

Além de acelerar a convergência do algoritmo, a ponderação também diminui a polarização dos algoritmos com restrição de norma ℓ_1 . Este aumento de rendimento é justificado pelo fato de que, quando se adota uma função de ponderação $\phi(\mathbf{w})$, procura-se fazê-lo de maneira que o resultado seja equivalente a se aproximar um filtro com restrição com norma ℓ_0 por meio de um algoritmo com restrição de norma unitária ponderada, sendo esta a principal razão da escolha de $\phi_1(\mathbf{w}) = \sum_{i=1}^N \arctan(\beta|w_i|)$, que é assintoticamente limitada a 1. Na literatura [16] [20] [26], outras funções usadas para ponderação também são propostas. Dentre elas, merece destaque a função $\phi_2(\mathbf{w}) = \sum_{i=1}^N \ln(\beta|w_i| + 1)$, cujo valor do gradiente é dado por:

$$\begin{aligned} \nabla_{\mathbf{w}} \phi_2(\mathbf{w}) &= \nabla_{\mathbf{w}} \left\{ \frac{2}{\pi} \sum_{j=1}^N \ln(\beta|w_j| + 1) \right\} \\ &= \beta \left[\frac{\text{sign}[w_1]}{\beta|w_1| + 1} \frac{\text{sign}[w_2]}{\beta|w_2| + 1} \cdots \frac{\text{sign}[w_N]}{\beta|w_N| + 1} \right]^T. \end{aligned} \quad (3.31)$$

A função $\phi_1(\mathbf{w})$ apresenta uma maior inclinação, para valores de $\mathbf{w} \approx \mathbf{0}$, quando comparada com $\phi_2(\mathbf{w})$, i.e., para valores de w_i próximos a origem, uma penalidade de norma ℓ_0 será melhor aproximada por $\phi_1(\mathbf{w})$ do que por $\phi_2(\mathbf{w})$ ou $\|\mathbf{w}\|_1$. A Fig. 3.1 apresenta resultados para aproximação da restrição de norma ℓ_0 por meio das funções de ponderação $\phi_1(\mathbf{w})$, $\phi_2(\mathbf{w})$ e o resultado clássico obtido por meio $\phi(\mathbf{w}) = \|\mathbf{w}\|_1$.

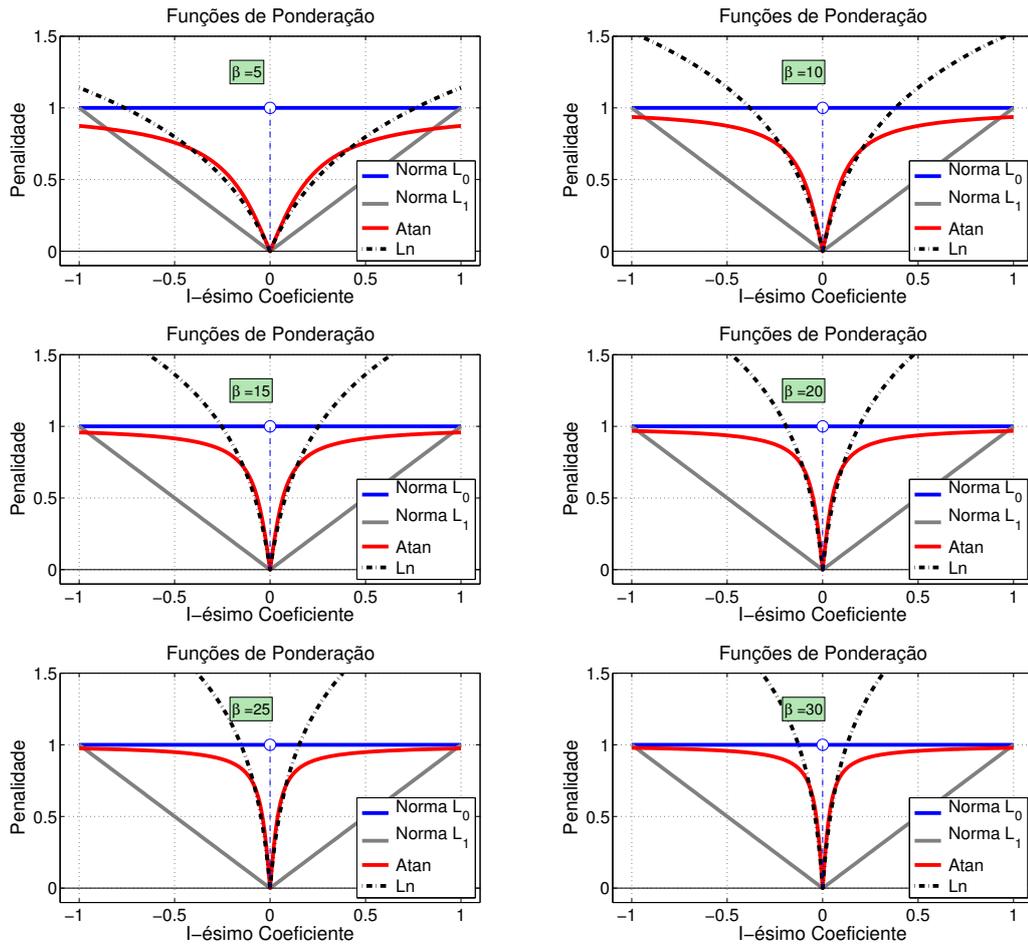


Figura 3.1: Resultado das aproximações da penalidade de norma l_0 , em função de β , por meio das funções de ponderação $\phi_1(\mathbf{w})$ e $\phi_2(\mathbf{w})$. Quando comparado com o resultado obtido pela penalidade de norma l_1 , i.e., $\phi(\mathbf{w}) = \|\mathbf{w}\|_1$, observa-se um notável ganho em favor do emprego das funções de ponderação para valores de w_i próximos à origem.

Na Fig 3.1, observa-se que a escolha do valor de β influencia sobremaneira na curva da função de ponderação na região do domínio próxima a zero (origem). Grandes valores de β produzem nas funções de ponderação acentuadas inclinações na região da origem, ao passo que, para regiões distantes da origem a curva começa a apresentar uma declinação menos acentuada tendendo a se tornar constante. O efeito, em termos de esparsidade, é que valores grandes de β , da ordem de grandeza da ordem do filtro, promovem esparsidade com maior intensidade, mas restringem o domínio que contempla esparsidade a um pequeno intervalo de valores absolutos, i. e., apenas coeficientes com valores absolutos dentro desse intervalo serão candidatos à esparsidade. A diminuição do valor de β , por sua vez, provoca um aumento no domínio na faixa de valores absolutos dos coeficientes que podem ser contemplados

com esparsidade, mas o faz com forma mais lenta.

Baseado na explicação do parágrafo anterior, conclui-se que:

- Para solução de problemas cujos sinais admitem alguma compressão, deve-se fazer uso de valores de β pequenos ou moderados, pois desta forma, o domínio é ampliado e pequenas componentes do sinal poderão ser contempladas; e
- Para identificação de sistemas esparsos, deve-se fazer uso de valores grandes de β , acima de 10, por exemplo, pois nesses casos, se está interessado em zerar os coeficientes cujos valores absolutos são próximos a zero.

3.5 Complexidade Computacional

Nesta Seção são apresentados os números de operações complexas para cada algoritmo proposto neste trabalho.

Na Tabela 3.1, são apresentados os números de operações complexas por iteração, para cada um dos algoritmos propostos neste trabalho. A depender da forma de implementação, o número de operações pode divergir ligeiramente do apresentado, porém a complexidade computacional do algoritmo permanece da ordem $\mathcal{O}(N)$. Isto indica que os algoritmos propostos apresentam a mesma complexidade do algoritmo CNLMS .

Tabela 3.1: Número de Operações Complexas por Iteração ($L = 1$)

	Adições	Multiplicações	Divisões
ℓ_1 -CLMS	$14N + 2$	$11N + 3$	$N + 1$
ℓ_1 -CNLMS	$14N + 3$	$12N + 5$	$N + 2$
ℓ_1 -WCNLMS	$16N + 5$	$13N + 6$	$N + 3$
CNLMS	$3N - 3$	$3N + 1$	1

Capítulo 4

Análises de Estabilidade e Convergência

Buscando provar que o algoritmo ℓ_1 -CNLMS é estável e convergente, neste capítulo são apresentadas as análises de estabilidade de primeira e segunda ordens [2] e as expressões para o Excesso de MSE e Desajuste [2]. Em decorrência da expressão final do algoritmo ℓ_1 -WCNLMS, Eq. (3.30), ser obtida da expressão do algoritmo ℓ_1 -CNLMS, por meio da substituição de \mathbf{s}_k pelo gradiente temporal da função de ponderação, que é limitada em norma e que não introduz novos autovalores, a análise apresentada neste Capítulo é aplicável à toda família de algoritmos apresentada no Capítulo 3.

Inicialmente, algumas considerações serão feitas no tocante ao sinal de referência, ruído e vetores de coeficientes ótimos. O sinal de referência d_k é escrito em função de dois coeficientes ótimos, i.e, neste trabalho, devido à coexistência de restrições lineares e restrição de norma ℓ_1 , dois vetores de coeficientes ótimos são considerados: \mathbf{w}_o , que é o vetor contendo os coeficientes ótimos, cujo cálculo leva em consideração uma dada restrição de norma ℓ_1 ; e o vetor \mathbf{w}_U , que é o vetor ótimo para a situação em que a restrição de norma ℓ_1 é igual ao valor da norma unitária de \mathbf{w}_o , onde matematicamente, esta situação é expressa como $t = \|\mathbf{w}_o\|_1$.

Portanto, a modelagem utilizada para o ruído leva em consideração duas parcelas de ruído: a primeira, descorrelacionada com o sinal de entrada, aqui representada por n_k , é associada ao sinal de referência e ao vetor \mathbf{w}_U ; a segunda parcela, η_k , correlacionada com o sinal de entrada, é a parcela responsável pela polarização do algoritmo e rege o excesso de MSE quando $t \neq \|\mathbf{w}_o\|_1$. Assim, a parcela do ruído correlacionada como o sinal de entrada pode ser escrita em função da parcela descorrelacionada, dos vetores de coeficientes ótimos e do sinal de entrada, conforme

a observamos seguir:

$$\begin{aligned} e_k^* &= [d_k - \mathbf{w}_k^H \mathbf{x}_k]^* = (\mathbf{x}_k^H \mathbf{w}_U + n_k^*) - \mathbf{x}_k^H \mathbf{w}_k \\ &= \mathbf{x}_k^H (\mathbf{w}_o - \mathbf{w}_k) + \eta_k, \end{aligned} \quad (4.1)$$

onde $\eta_k = n_k^* + \mathbf{x}_k^H (\mathbf{w}_U - \mathbf{w}_o)$.

4.1 Análise de Primeira Ordem

Para proceder a análise de primeira ordem, deve-se, primeiramente, substituir o passo de convergência fixo da Eq. (3.20), pela versão do passo de convergência variável calculado por meio da Eq. (3.23), obtendo:

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{P} \mathbf{w}_k + \left(\frac{\mu_0}{\zeta_k} \right) (e_k^* - \mathbf{x}_k^H \mathbf{f}_{L1}(k)) \mathbf{P} \bar{\mathbf{P}}_k \mathbf{x}_k + \mathbf{f} + \mathbf{f}_{L1}(k) \\ &= \mathbf{P} \mathbf{w}_k + \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k e_k^* - \left(\frac{\mu_0 \mathbf{x}_k^H \mathbf{f}_{L1}(k)}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \\ &\quad + \mathbf{f} + \mathbf{f}_{L1}(k), \end{aligned} \quad (4.2)$$

onde $\zeta_k = \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k + \gamma$, onde a parcela γ é um número pequeno e positivo, reintroduzido na expressão que contém a energia do sinal de entrada, como o propósito de garantir que ζ_k não assuma valores nulos na hipótese do vetor \mathbf{x}_k pertencer ao Espaço Nulo da matriz de projeção $\bar{\mathbf{P}}_k$.

Usando o resultado apresentado na Eq. (4.2) e o fato que o vetor de restrições lineares \mathbf{f} pode ser escrito em função de \mathbf{w}_o , i. e., $\mathbf{f} = [\mathbf{I} - \mathbf{P}] \mathbf{w}_o$, pode-se reescrever a Eq. (4.2), como se segue:

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{P} [\mathbf{w}_k - \mathbf{w}_o] + \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k [\mathbf{x}_k^H (\mathbf{w}_o - \mathbf{w}_k) + \eta_k] \\ &\quad - \left(\frac{\mu_0 \mathbf{x}_k^H \mathbf{f}_{L1}(k)}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k + \mathbf{f}_{L1}(k) + \mathbf{w}_o. \end{aligned} \quad (4.3)$$

Definindo o erro do filtro adaptativo referente ao vetor de coeficientes ótimos

com restrições como $\Delta \mathbf{w}(k) = \mathbf{w}_k - \mathbf{w}_o$, chega-se a:

$$\begin{aligned}
\mathbf{w}_{k+1} &= \mathbf{P} \Delta \mathbf{w}(k) + \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k [\eta_k - \mathbf{x}_k^H \Delta \mathbf{w}(k)] \\
&\quad - \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \mathbf{x}_k^H \mathbf{f}_{L1}(k) + \mathbf{f}_{L1}(k) + \mathbf{w}_o \\
&= \mathbf{P} \Delta \mathbf{w}(k) + \mathbf{w}_o - \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \mathbf{x}_k^H \Delta \mathbf{w}(k) \\
&\quad + \left\{ \mathbf{I} - \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \mathbf{x}_k^H \right\} \mathbf{f}_{L1}(k) + \mathbf{f}(\eta_k),
\end{aligned} \tag{4.4}$$

com $\mathbf{f}(\eta_k) = \eta_k \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k$.

A fim de permitir uma representação mais compacta para o vetor $\mathbf{f}_{L1}(k)$, pode-se definir variáveis auxiliares e, então, reescrevê-la como:

$$\begin{aligned}
\mathbf{f}_{L1}(k) &= e_{L1}(k) \mathbf{v}_k, \\
&= (t - \mathbf{s}_k^H \mathbf{w}_k) \mathbf{v}_k, \\
&= (t - \mathbf{s}_k^H \Delta \mathbf{w}(k) - \mathbf{s}_k^H \mathbf{w}_o) \mathbf{v}_k, \\
&= t \mathbf{v}_k - \mathbf{P} \mathbf{M}_0(k) (\Delta \mathbf{w}(k) + \mathbf{w}_o),
\end{aligned} \tag{4.5}$$

com

$$\begin{cases} \mathbf{v}_k = \frac{\mathbf{P} \mathbf{s}_k}{\|\mathbf{P} \mathbf{s}_k\|^{2+\gamma}}, \\ \mathbf{M}_0(k) = \frac{\mathbf{P} \mathbf{s}_k \mathbf{s}_k^H}{\|\mathbf{P} \mathbf{s}_k\|^{2+\gamma}}. \end{cases}$$

A parcela γ é um número pequeno e positivo, cujo propósito é garantir que \mathbf{v}_k se mantenha limitado em norma, mesmo na hipótese do vetor \mathbf{s}_k pertencer ao Espaço Nulo da matriz de projeção \mathbf{P} .

Após substituir o resultado expresso pela Eq. (4.5) na Eq. (4.4), obtém-se:

$$\begin{aligned}
\mathbf{w}_{k+1} &= \mathbf{P} \Delta \mathbf{w}(k) - \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \mathbf{x}_k^H \Delta \mathbf{w}(k) \\
&\quad + t \mathbf{v}_k - \mathbf{P} \mathbf{M}_0(k) \Delta \mathbf{w}(k) - \mathbf{P} \mathbf{M}_0(k) \mathbf{w}_o \\
&\quad - \left\{ \left(\frac{\mu_0}{\zeta_k} \right) \bar{\mathbf{P}}_k \mathbf{x}_k \mathbf{x}_k^H \right\} \left\{ t \mathbf{v}_k - \mathbf{P} \mathbf{M}_0(k) \Delta \mathbf{w}(k) - \right. \\
&\quad \quad \left. \mathbf{P} \mathbf{M}_0(k) \mathbf{w}_o \right\} + \mathbf{f}(\eta_k) + \mathbf{w}_o.
\end{aligned} \tag{4.6}$$

Subtraindo-se \mathbf{w}_o de ambos membros da Eq. (4.6) e sabendo-se que $\mathbf{P} \bar{\mathbf{P}}_k = \bar{\mathbf{P}}_k$ e que $\mathbf{P} \mathbf{M}_0 = \mathbf{M}_0$, chega-se a:

$$\begin{aligned}\Delta \mathbf{w}(k+1) &= \mathbf{P} \left[\mathbf{I} - \mu_0 \bar{\mathbf{P}}_k \left(\frac{\mathbf{x}_k \mathbf{x}_k^H}{\zeta_k} \right) \right] \left[\mathbf{I} - \mathbf{M}_0(k) \right] \Delta \mathbf{w}(k) \\ &+ \left[\mathbf{I} - \mu_0 \bar{\mathbf{P}}_k \left(\frac{\mathbf{x}_k \mathbf{x}_k^H}{\zeta_k} \right) \right] \mathbf{v}_k (t - \mathbf{s}_k^H \mathbf{w}_0) + \mathbf{f}(\eta_k),\end{aligned}\tag{4.7}$$

cujo valor esperado é expresso como:

$$\begin{aligned}E[\Delta \mathbf{w}(k+1)] &= E \left\{ \mathbf{P} \left[\mathbf{I} - \mu_0 \mathbf{M}_1(k) \right] \left[\mathbf{I} - \mathbf{M}_0(k) \right] \Delta \mathbf{w}(k) \right. \\ &\quad \left. + \left[\mathbf{I} - \mu_0 \mathbf{M}_1(k) \right] \mathbf{v}_k \delta_k + \mathbf{f}(\eta_k) \right\},\end{aligned}\tag{4.8}$$

com

$$\begin{cases} \delta_k = t - \mathbf{s}_k^H \mathbf{w}_0, \\ \mathbf{M}_1(k) = \bar{\mathbf{P}}_k \left(\frac{\mathbf{x}_k \mathbf{x}_k^H}{\mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k + \gamma} \right). \end{cases}$$

A despeito de não ser trivial se chegar a uma expressão fechada para o cálculo do valor esperado dado pela Eq. (4.8), fazendo uso do princípio da independência estatística (*Independence Assumption*) [27] e investigando os autovalores das matrizes \mathbf{M}_0 e \mathbf{M}_1 , é possível determinar se o algoritmo aqui analisado converge em média.

Inicialmente, para a condição normalizada, aquela cujo erro *a posteriori* é nulo ($\mu_0 = 1$), parte-se da premissa que, próximo à convergência, o sinal de entrada \mathbf{x}_k e o vetor \mathbf{s}_k são estatisticamente independentes, podendo, então, garantir que o algoritmo converge em média se os autovalores das matrizes \mathbf{M}_0 e \mathbf{M}_1 permanecerem dentro do círculo unitário, uma vez que o vetor \mathbf{v}_k e o escalar δ_k são limitados em norma para quaisquer valores de k . Neste caso, a validade da independência estatística se dá pelo fato de que, próximo da convergência, é razoável considerar que o vetor de coeficientes do filtro permanece sobre o mesmo hiperquadrante, não mudando se conjunto de sinais dado por \mathbf{s}_k , ou, pelo menos, mudando \mathbf{s}_k a uma taxa bem mais lenta que o sinal de entrada \mathbf{x}_k .

Desprezando a influência de γ para efeito de autovalores, pode-se considerar que as matrizes \mathbf{M}_0 e \mathbf{M}_1 são do tipo $\mathcal{M} = \frac{\mathbf{u}\mathbf{v}^T}{\mathbf{v}^T\mathbf{u}}$ e, portanto, possuem posto unitário tendo apenas um autovalor não nulo e igual a 1. Para se garantir a convergência do algoritmo, uma análise mais cuidadosa deve ser realizada nos termos da Eq. (4.9). Deve-se, para tanto, encontrar a faixa de valores de μ_0 para os quais os autovalores das \mathbf{AB} permanecem dentro do círculo unitário.

$$\mathbf{AB} = \underbrace{\left[\mathbf{I} - \mu_0 \mathbf{M}_1(k) \right]}_{\mathbf{A}} \underbrace{\left[\mathbf{I} - \mathbf{M}_0(k) \right]}_{\mathbf{B}}.\tag{4.9}$$

$$\begin{aligned}
\text{tr}[\mathbf{AB}] &= \text{tr} \left[\mathbf{I} - \mathbf{M}_0(k) - \mu_0 \mathbf{M}_1(k) + \mu_0 \mathbf{M}_1(k) \mathbf{M}_0(k) \right] \\
&= N - 1 - \mu_0 + \mu_0 \text{tr} \left[\mathbf{M}_1(k) \mathbf{M}_0(k) \right] \\
&= N - 1 - \mu_0 + \mu_0 \text{tr} \left[\mathbf{M}_0(k) \mathbf{M}_1(k) \right] \\
&= N - \mu_0 - 1,
\end{aligned} \tag{4.10}$$

com $\mathbf{M}_0(k) \mathbf{M}_1(k) = \mathbf{0}$.

Seja a matriz \mathbf{A} inversível, com posto completo e igual a N , e a matriz \mathbf{B} idempotente, implicando posto dessa matriz igual ao seu traço, i.e., $\text{tr}[\mathbf{B}] = \text{rank}(\mathbf{B}) = N - 1$. Como a matriz $\mathbf{M}_1(k)$ tem posto unitário, a matriz \mathbf{A} possui $N - 1$ autovalores iguais a 1 e um autovalor, ainda desconhecido, igual λ_0 . Por outro lado, a matriz \mathbf{B} possui posto igual a $N - 1$ e goza de todas as propriedades de uma matriz de projeção. Portanto, o posto do produto \mathbf{AB} é igual ao posto de \mathbf{A} reduzido de um, implicando $N - 1$ autovalores não nulos, dentre os quais $N - 2$ autovalores são iguais a 1. Como o traço de uma matriz é igual ao somatório de seus autovalores:

$$\begin{aligned}
\text{tr}[\mathbf{AB}] &= \sum_{i=1}^N \lambda_i = \underbrace{(1 + 1 + \cdots + 1)}_{(N-2) \text{ termos}} + \lambda_0 + 0 \\
&= N - 2 + \lambda_0.
\end{aligned} \tag{4.11}$$

Todavia, da Eq. (4.10), tem-se que $\text{tr}[\mathbf{AB}] = N - \mu_0 - 1$, que implica $\lambda_0 = 1 - \mu_0$. A partir do valor de λ_0 , chega-se aos autovalores do produto \mathbf{AB} como sendo: 0, 1 and $\lambda_0 = 1 - \mu_0$. Assim, o algoritmo converge em média se $|\lambda_0| < 1$, resultando em:

$$0 < \mu_0 < 2. \tag{4.12}$$

Como os algoritmos L_1 -WCNLMS e L_1 -CNLMS produzem pares de matrizes \mathbf{A} e \mathbf{B} , cujos conjuntos de autovalores são idênticos, o resultados obtido pela Eq. (4.12) é igualmente aplicável aos dois algoritmos.

4.2 Análise de Segunda Ordem

Nesta Seção, é apresentada a análise de segunda ordem para o algoritmo ℓ_1 -CNLMS, que é realizada por meio da investigação dos autovalores das estatísticas de segunda ordem obtidas para o vetor de erros. Para tanto, as seguintes premissas são assumidas:

- a) Para valores grandes de k , o vetor $\Delta \mathbf{w}(k + 1)$ converge para zero na média;

- b) Os vetores de sinal dos coeficientes \mathbf{s}_k , e de coeficientes \mathbf{w}_k , próximo à convergência, são estatisticamente independentes; e
- c) O ruído η_k e o sinal de entrada \mathbf{x}_k possuem distribuição gaussiana com médias zero e variâncias σ_N^2 e σ_x^2 , respectivamente.

A partir da Eq. (4.7), pode-se expressar a matriz de covariância para o vetor $\Delta \mathbf{w}(k+1)$ como:

$$\begin{aligned}
\text{cov}[\Delta \mathbf{w}(k+1)] &= E[\Delta \mathbf{w}(k+1)\Delta \mathbf{w}^H(k+1)] = \\
&E\left\{\mathbf{f}(\eta_k)\right\}E\left\{(\mathbf{PAB}\Delta \mathbf{w}(k))^H + (\mathbf{A}\mathbf{v}_k\delta_k)^H\right\} \\
&+ E\left\{(\mathbf{PAB}\Delta \mathbf{w}(k)) + (\mathbf{A}\mathbf{v}_k\delta_k)\right\}E\left\{\mathbf{f}^H(\eta_k)\right\} \\
&+ E\left\{(\mathbf{PAB})\Delta \mathbf{w}(k)\Delta \mathbf{w}^H(k)(\mathbf{PAB})^H\right\} \\
&+ E\left\{[(\mathbf{PAB})\Delta \mathbf{w}(k)(\mathbf{A}\mathbf{v}_k\delta_k)^H]^H\right\} \\
&+ E\left\{(\mathbf{PAB})\Delta \mathbf{w}(k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\right\} \\
&+ E\left\{(\mathbf{A}\mathbf{v}_k\delta_k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\right\} \\
&+ E\left\{\mathbf{f}(\eta_k)\mathbf{f}^H(\eta_k)\right\},
\end{aligned} \tag{4.13}$$

$$\text{com } E[\mathbf{f}(\eta_k)] = E\left[\eta_k \begin{pmatrix} \mu_0 \\ \zeta_k \end{pmatrix} \bar{\mathbf{P}}_k \mathbf{x}_k\right] \approx 0.$$

Agrupando os termos semelhantes, após algumas manipulações algébricas, chega-se a:

$$\begin{aligned}
\text{cov}[\Delta \mathbf{w}(k+1)] &\approx E\left\{(\mathbf{A}\mathbf{v}_k\delta_k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\right\} \\
&+ E\left\{(\mathbf{PAB})\Delta \mathbf{w}(k)\Delta \mathbf{w}^H(k)(\mathbf{PAB})^H\right\} \\
&+ E\left\{(\mathbf{PAB})\Delta \mathbf{w}(k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\right\} \\
&+ [(\mathbf{PAB})\Delta \mathbf{w}(k)(\mathbf{A}\mathbf{v}_k\delta_k)^H]^H \\
&+ E\left\{\mathbf{f}(\eta_k)\mathbf{f}^H(\eta_k)\right\}.
\end{aligned} \tag{4.14}$$

Inicialmente, cada termo da Eq. (4.14) é analisado em separado para facilitar os cálculos. Como não é trivial calcular os valores esperados dos termos dessa equação, algumas premissas e simplificações são adotadas para possibilitar investigar os

autovalores de cada termo interno ao operador $E\{\cdot\}$. Decorrente dessa não trivialidade, é adotada a estratégia de avaliar os autovalores dos argumentos dos operadores $E\{\cdot\}$, sem no entanto, avaliar os valores médios propriamente ditos dos termos da equação. Apesar desta análise não contemplar corretamente o cálculo da matriz de covariância, ela permite, ainda assim, determinar a estabilidade do algoritmo e o cálculo da faixa de valores do passo de convergência para a qual o algoritmo é estável.

- Primeiro Termo:

Aplicando-se o princípio da independência estatística [2] [28] ao primeiro termo constante do lado direito da Eq. (4.14), chega-se a:

$$E\{(\mathbf{A}\mathbf{v}_k\delta_k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\} = |\delta_k|^2 E\{\mathbf{A}\mathbf{v}_k\mathbf{v}_k^H\mathbf{A}^H\}. \quad (4.15)$$

Na Eq. (4.15), calculando-se o traço para o termo interno ao operador $E\{\cdot\}$ e aplicando o teorema que permite expressar o traço de um produto de matrizes fatores como uma desigualdade do produto dos traços das matrizes fatores [29]¹, obtém-se:

$$\begin{aligned} \text{tr}\{(\mathbf{A}\mathbf{v}_k\delta_k)(\mathbf{A}\mathbf{v}_k\delta_k)^H\} &= |\delta_k|^2 \text{tr}\{\mathbf{A}^H\mathbf{A}\mathbf{v}_k\mathbf{v}_k^H\} \\ &\leq |\delta_k|^2 \text{tr}\{\mathbf{A}^H\mathbf{A}\} \text{tr}\{\mathbf{v}_k\mathbf{v}_k^H\} \\ &= \frac{|\delta_k|^2}{\|\mathbf{P}\mathbf{s}_k\|^2 + \gamma} (\mu_0^2 - 2\mu_0 + N), \end{aligned} \quad (4.16)$$

com $\text{tr}\{\mathbf{v}_k\mathbf{v}_k^H\} = \frac{1}{\|\mathbf{P}\mathbf{s}_k\|^2 + \gamma}$ e $\text{tr}\{\mathbf{A}^H\mathbf{A}\} = \mu_0^2 - 2\mu_0 + N$.

- Segundo Termo:

Para o cálculo do traço das quantidades dentro do operador $E\{\cdot\}$ do Segundo Termo presente do lado direito da Eq. (4.14), aplica-se o princípio da independência estatística [2] [28] e o teorema que permite expressar o traço de um produto de matrizes fatores como uma desigualdade do produto dos traços das matrizes fatores [29]:

$$\begin{aligned} \text{tr}\left\{(\mathbf{PAB}) \text{cov}[\Delta\mathbf{w}(k)] (\mathbf{PAB})^H\right\} &\leq |\text{tr}\{(\mathbf{PAB})\}|^2 \text{tr}\left\{\text{cov}[\Delta\mathbf{w}(k)]\right\} \\ &= |N - \mu_0 - L|^2 \text{tr}\left\{\text{cov}[\Delta\mathbf{w}(k)]\right\}, \end{aligned} \quad (4.17)$$

¹ $\text{tr}\{\mathbf{Y}_1\mathbf{Y}_2\cdots\mathbf{Y}_N\} \leq \text{tr}\{\mathbf{Y}_1\}\text{tr}\{\mathbf{Y}_2\}\cdots\text{tr}\{\mathbf{Y}_N\}$.

com

$$\begin{cases} \text{tr}\{\mathbf{AB}\} = N - \mu_0 - 1, \\ \text{tr}\{(\mathbf{PAB})\} = N - \mu_0 - L, \\ L = \text{número de restrições lineares.} \end{cases}$$

- Terceiro Termo:

Fazendo as mesmas considerações realizadas para o Primeiro e Segundo Termos, pode-se, então, calcular o traço das quantidades dentro do operador $E\{\cdot\}$ do Terceiro Termo presente do lado direito da Eq. (4.14):

$$\begin{aligned} & \text{tr}\left\{ (\mathbf{PAB}) \Delta \mathbf{w}(k) (\mathbf{A} \mathbf{v}_k \delta_k)^{\text{H}} \right. \\ & \quad \left. + [(\mathbf{PAB}) \Delta \mathbf{w}(k) (\mathbf{A} \mathbf{v}_k \delta_k)^{\text{H}}]^{\text{H}} \right\} \\ & \leq 2\Re\left\{ \text{tr}\{(\mathbf{PAB})\} \right\} \Re\left\{ \text{tr}\{\Delta \mathbf{w}(k) (\mathbf{A} \mathbf{v}_k \delta_k)^{\text{H}}\} \right\} \\ & = 2\{N - \mu_0 - L\} \Re\left\{ \text{tr}\{\Delta \mathbf{w}(k) (\mathbf{A} \mathbf{v}_k \delta_k)^{\text{H}}\} \right\}. \end{aligned} \tag{4.18}$$

- Quarto Termo:

Para o Quarto Termo, fazendo as mesmas considerações realizadas para os termos anteriores, pode-se, então, calcular o traço das quantidades dentro do operador $E\{\cdot\}$ como:

$$\begin{aligned} \text{tr}\{\mathbf{f}(\eta_k) \mathbf{f}^{\text{H}}(\eta_k)\} &= \text{tr}\left\{ \left[\eta_k \begin{pmatrix} \mu_0 \\ \zeta_k \end{pmatrix} \bar{\mathbf{P}}_k \mathbf{x}_k \right] \left[\eta_k \begin{pmatrix} \mu_0 \\ \zeta_k \end{pmatrix} \bar{\mathbf{P}}_k \mathbf{x}_k \right]^{\text{H}} \right\} \\ &= \frac{\mu_0^2 |\eta_k|^2}{(\|\bar{\mathbf{P}}_k \mathbf{x}_k\|^2 + \gamma)^2} \text{tr}\left\{ (\bar{\mathbf{P}}_k \mathbf{x}_k) (\mathbf{x}_k^{\text{H}} \bar{\mathbf{P}}_k^{\text{H}}) \right\} \\ &= \frac{\mu_0^2 |\eta_k|^2}{(\|\bar{\mathbf{P}}_k \mathbf{x}_k\|^2 + \gamma)^2} \|\bar{\mathbf{P}}_k \mathbf{x}_k\|^2 \\ &\approx \frac{\mu_0^2 |\eta_k|^2}{\|\bar{\mathbf{P}}_k \mathbf{x}_k\|^2 + \gamma}. \end{aligned} \tag{4.19}$$

- Cálculo de μ_0 :

Com base nas Eqs. (4.15), (4.18) e (4.19), levando-se em consideração os termos limitados em norma, para se garantir a convergência do algoritmo, é necessário calcular a faixa de μ_0 , para a qual os autovalores da matriz \mathbf{PAB} , se mantêm dentro do círculo unitário, i.e., $\lambda_i \in [0,1] \forall i$.

Seguindo o mesmo raciocínio empregado na Análise de Primeira Ordem, pode-se afirmar que a matriz \mathbf{PAB} terá posto $N - L$, o que significa que dentre os seus autovalores há $N - L$ autovalores não nulos, dentre os quais $N - L - 1$ são iguais a 1 e um autovalor é igual a λ_0 . Desta forma, como o traço de uma matriz é igual ao somatório dos seus autovalores:

$$\begin{aligned} \text{tr}[\mathbf{PAB}] &= \sum_{i=1}^N \lambda_i = \underbrace{(1 + 1 + \cdots + 1)}_{(N-L-1) \text{ termos}} + \lambda_0 + \underbrace{(0 + 0 + \cdots + 0)}_{(L) \text{ termos}} \\ &= N - L + \lambda_0 - 1. \end{aligned} \quad (4.20)$$

O $\text{tr}[\mathbf{PAB}] = N - L - \mu_0$, que implica $\lambda_0 = 1 - \mu_0$. A partir do valor de λ_0 , chega-se aos autovalores do produto \mathbf{PAB} como sendo: 0, 1 and $\lambda_0 = 1 - \mu_0$. Assim, o algoritmo converge em média se $|\lambda_0| < 1$, resultando em:

$$0 < \mu_0 < 2. \quad (4.21)$$

4.3 Excesso de MSE e Desajuste

Nesta seção são obtidas as expressões aproximadas para o Excesso de MSE e o Desajuste [2]. Inicialmente, são definidos três conceitos relacionados ao Erro Médio Quadrático apresentado por um filtro e que são utilizados nesta seção:

- a) **MSE mínimo**: está relacionado à parte do sistema que não foi possível identificar, e decorre da existência de ruído não correlacionado ao sinal de entrada \mathbf{x}_k . Pode também ocorrer quando o filtro possui número de coeficientes inferior ao necessário para modelar o sistema desconhecido;
- b) **Excesso de MSE**: na média, o resultado obtido para algoritmos baseados em LMS converge para uma solução ótima \mathbf{w}_o , porém, os valores instantâneos obtidos para o vetor de coeficientes \mathbf{w}_k sofrem desvios instantâneos do ponto ótimo dado por $\Delta\mathbf{w}_k = (\mathbf{w}_k - \mathbf{w}_o)$. Estes desvios decorrem da característica tipicamente ruidosa do estimador do gradiente. O efeito desse desvio instantâneo se dá por meio do aumento ou Excesso de MSE. Especificamente, para algoritmos com restrição de norma ℓ_1 , uma contribuição adicional para composição do Excesso de MSE, proporcional a $\|\mathbf{w}_U - \mathbf{w}_0\|^2$, deve ser considerada; e
- c) **Desajuste (M)**: este parâmetro representa a relação entre o Excesso de MSE e o MSE mínimo. Desta forma, serve como figura de mérito para comparar vários algoritmos empregados em filtragem adaptativa.

Tomando como base o algoritmo ℓ_1 -CNLMS, da Eq. (4.2), pode-se expressar o quadrado do valor do erro instantâneo, no instante k , como:

$$\begin{aligned}
|e_k|^2 &= [d_k - \mathbf{w}_k^H \mathbf{x}_k]^* [d_k - \mathbf{w}_k^H \mathbf{x}_k] \\
&= \{ [\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k)]^H \mathbf{x}_k + n_k^* \}^* \{ [\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k)]^H \mathbf{x}_k + n_k^* \} \\
&= \mathbf{x}_k^H [\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k)] [\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k)]^H \mathbf{x}_k + 2\Re \left\{ n_k [\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k)]^H \mathbf{x}_k \right\} + |n_k|^2,
\end{aligned} \tag{4.22}$$

onde $\mathbf{\Delta}_0 = (\mathbf{w}_U - \mathbf{w}_o)$.

Expandindo os termos da Eq. (4.22) e tomando o valor esperado do resultado, obtêm-se:

$$\begin{aligned}
E\{|e_k|^2\} &= E\left\{ \mathbf{x}_k^H \mathbf{\Delta}_0 \mathbf{\Delta}_0^H \mathbf{x}_k \right\} + E\left\{ \mathbf{x}_k^H \mathbf{\Delta} \mathbf{w}(k) \mathbf{\Delta} \mathbf{w}^H(k) \mathbf{x}_k \right\} \\
&\quad + 2\Re \left\{ E\left[e_0 \mathbf{\Delta} \mathbf{w}^H(k) \mathbf{x}_k \right] \right\} + 2\Re \left\{ E\left[n_k (\mathbf{\Delta}_0 + \mathbf{\Delta} \mathbf{w}(k))^H \mathbf{x}_k \right] \right\} + \sigma_N^2,
\end{aligned} \tag{4.23}$$

com $e_0^* = \mathbf{\Delta}_0^H \mathbf{x}_k$.

Em decorrência da independência estatística e do princípio da ortogonalidade, é importante observar que o terceiro e o quarto termos da Eq. (4.23) são nulos. Assim, pode-se reescrever esta equação como:

$$J(k) = E\{|e_k|^2\} = E\left\{ \mathbf{x}_k^H \text{cov}[\mathbf{\Delta} \mathbf{w}(k)] \mathbf{x}_k \right\} + E\left\{ \mathbf{x}_k^H \mathbf{K}_0 \mathbf{x}_k \right\} + \sigma_N^2, \tag{4.24}$$

onde $\mathbf{K}_0 = \mathbf{\Delta}_0 \mathbf{\Delta}_0^H$.

Tomando o traço da Eq. (4.24) e levando em consideração que o traço do valor esperado é o valor esperado do traço, chega-se a:

$$J(k) = E\left\{ \text{tr} \left[\mathbf{x}_k^H \text{cov}[\mathbf{\Delta} \mathbf{w}(k)] \mathbf{x}_k \right] \right\} + E\left\{ \text{tr} \left[\mathbf{x}_k^H \mathbf{K}_0 \mathbf{x}_k \right] \right\} + \sigma_N^2 \tag{4.25}$$

Como a função traço de um produto é invariante às permutações cíclicas dos fatores, i.e., $\text{tr}[\mathbf{VZ}] = \text{tr}[\mathbf{ZV}]$, tem-se:

$$\begin{aligned}
J(k) &= \text{tr} \left\{ \mathbf{R} \text{cov}[\mathbf{\Delta} \mathbf{w}(k)] \right\} + \text{tr} \left\{ \mathbf{R} \mathbf{K}_0 \right\} + \sigma_N^2 \\
&= J_U(k) + J_C + J_{\min}.
\end{aligned} \tag{4.26}$$

onde:

$$\begin{cases} \mathbf{R} &= E\{\mathbf{x}_k \mathbf{x}_k^H\}, \\ J_U(k) &= \text{tr}\left\{\mathbf{R}\text{cov}[\Delta \mathbf{w}(k)]\right\}, \\ J_C &= \text{tr}\left\{\mathbf{R}\mathbf{K}_0\right\}, \\ J_{\min} &= \sigma_N^2. \end{cases}$$

Na Eq. (4.26), o termo J_C é responsável pela parcela de MSE provocada pela restrição de norma ℓ_1 ; neste caso, representado pela variância do sinal de entrada ponderada pelo valor quadrático do desvio entre os vetores ótimos com e sem restrição de norma ℓ_1 . O termo J_{\min} corresponde ao MSE da solução do filtro de Wiener [3] e o termo $J_U(k)$ representa a parcela de MSE correspondente ao caso sem restrições (*unconstrained*).

Para se chegar a uma expressão para o Excesso de MSE, o caminho natural seria obter uma expressão simplificada para $J_U(k) = \text{tr}\{\mathbf{R}\text{cov}[\Delta \mathbf{w}(k)]\}$ por meio da Eq. (4.14); mas, devido à complexidade dessa expressão, é proposta neste trabalho uma abordagem alternativa por meio da heurística simplificada apresentada em [31].

De acordo com a heurística simplificada supracitada, o MSE pode ser expresso em função do erro médio quadrático mínimo e da média do Excesso de MSE:

$$J = J_{\min} + \left\{ \text{Média do Excesso de MSE} \right\}. \quad (4.27)$$

O Excesso de MSE é originado dos desvios instantâneos que o vetor \mathbf{w}_k sofre em torno do ponto ótimo, i.e., por $\Delta \mathbf{w}_k = (\mathbf{w}_k - \mathbf{w}_o)$. Estes desvios decorrem da característica tipicamente ruidosa do estimador do gradiente e, especificamente para algoritmos com restrição de norma ℓ_1 , uma contribuição adicional para composição do Excesso de MSE, proporcional a $\|\mathbf{w}_U - \mathbf{w}_0\|^2$, deve ser considerada. Esta contribuição se encontra representada na Eq. (4.26) pelo termo $J_C = \text{tr}\left\{\mathbf{R}\mathbf{K}_0\right\}$.

Neste trabalho, a média do Excesso de MSE será aproximada pela soma de parcelas ponderadas referentes à potência do ruído inerente ao estimador do gradiente e ao MSE introduzido pela restrição de norma ℓ_1 :

$$J(\infty) \approx \alpha_1 J_U(\infty) + \alpha_2 J_C + J_{\min}, \quad (4.28)$$

onde os fatores α_1 e α_2 são introduzidos com o propósito de compensar o erro decorrente da heurística proposta. Estes fatores foram obtidos empiricamente por meio de interpolação de valores simulados, fazendo-se variar o passo de convergência

fixo μ_0 e o número de coeficientes N .

Como os termos $J_U(\infty)$ e J_C são independentes, pode-se aplicar o princípio da superposição, fazendo $J_C = 0$, para estimar $J_U(\infty)$ com base nas características de decaimento exponencial da curva de aprendizagem do algoritmo LMS [3]:

$$\alpha_1 J_U(\infty) = \alpha_1 \mu_0 \sigma_N^2 \sum_{j=1}^N \frac{\lambda_j}{2 - \mu_0 \lambda_j}. \quad (4.29)$$

Na Eq. (4.29), os autovalores λ_j são os mesmo que determinam a estabilidade do algoritmo ℓ_1 -CNLMS e, desta forma, são os autovalores da matriz **PAB**. Substituindo os autovalores da matriz **PAB** na Eq. (4.29) e este resultado na Eq. (4.28), obtém-se:

$$J(\infty) \approx \alpha_1 \mu_0 \sigma_N^2 \frac{(N - L - 1)(\mu_0^2 - \mu_0 + 2) + (2 - \mu_0)(1 - \mu_0)}{(2 - \mu_0)(\mu_0^2 - \mu_0 + 2)} + \alpha_2 \text{tr}\{\mathbf{RK}_0\} + \sigma_N^2. \quad (4.30)$$

Os fatores α_1 e α_2 são empregados para compensar as aproximações decorrentes do emprego da heurística proposta (motivada por [31]) no cálculo do $J(\infty)$ para o algoritmo ℓ_1 -CNLMS. Seus valores foram ajustados numericamente, aproximando $\text{tr}\{\mathbf{RK}_0\}$ por $N\sigma_x^2 \text{tr}\{\mathbf{K}_0\}$, até se conseguir um resultado satisfatório para uma vasta gama de N , μ_0 e t :

$$\begin{cases} \alpha_1 = \frac{1}{N\sigma_x^2}; \text{ e} \\ \alpha_2 = \frac{1 + \mu_0^2}{N\sigma_x^2}, \text{ para } 0 < \mu_0 < 1,90. \end{cases}$$

Na Eq. (4.30), resta calcular a parcela $\text{tr}\{\mathbf{K}_0\}$, cujo valor, *a priori*, não se conhece, porque é definida em termos dos vetores de coeficientes ótimos \mathbf{w}_U e \mathbf{w}_0 . Pode-se, no entanto, buscar uma aproximação assumindo-se as seguintes premissas:

- a) O Excesso de MSE será calculado considerando valores de $k \rightarrow \infty$; e
- b) Para $k \rightarrow \infty$, os vetores ótimos \mathbf{w}_U e \mathbf{w}_0 pertencerão a um mesmo hiperquadrante, o que implica em ambos vetores compartilharem o mesmo conjunto de sinais.

Pré-multiplicando-se \mathbf{K}_0 por \mathbf{s}_k^H e pós-multiplicando-se por \mathbf{s}_k , obtém-se:

$$\begin{aligned} \mathbf{s}_k^H \mathbf{K}_0 \mathbf{s}_k &= \text{tr} \left\{ \mathbf{K}_0 \mathbf{s}_k \mathbf{s}_k^H \right\} \\ &\leq N \text{tr} \{ \mathbf{K}_0 \}. \end{aligned} \quad (4.31)$$

A matriz \mathbf{K}_0 pode ser expressa em função dos vetores \mathbf{w}_U e \mathbf{w}_0 , conforme a seguir:

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathbf{s}_k^H \mathbf{K}_0 \mathbf{s}_k &= \lim_{k \rightarrow \infty} (\mathbf{s}_k^H \mathbf{\Delta}_0) (\mathbf{s}_k^H \mathbf{\Delta}_0)^H \\ &= \|\mathbf{s}_\infty^H (\mathbf{w}_U - \mathbf{w}_0)\|^2 \\ &\approx \|\mathbf{s}_\infty^H \mathbf{w}_U - t\|^2 \\ &\approx (\|\mathbf{w}_U\|_1 - t)^2 \\ &= (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2, \end{aligned} \quad (4.32)$$

onde $t = \alpha \|\mathbf{w}_U\|_1$ e α é o fator de “encolhimento”, tal que $\alpha \in (0,1)$.

Das Eqs. (4.31) e (4.31), observa-se que $N \text{tr} \{ \mathbf{K}_0 \} \approx (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2$. Substituindo este resultado na Eq. (4.30), chega-se em:

$$\begin{aligned} J(\infty) &\approx \alpha_1 \mu_0 \sigma_N^2 \frac{(N - L - 1)(\mu_0^2 - \mu_0 + 2) + (2 - \mu_0)(1 - \mu_0)}{(2 - \mu_0)(\mu_0^2 - \mu_0 + 2)} \\ &\quad + \alpha_2 \sigma_x^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2 + \sigma_N^2. \end{aligned} \quad (4.33)$$

O valor da norma $\|\mathbf{w}_U\|_1$ pode ser estimado por meio das características de energia da planta ou sistema sob filtragem. A expressão aproximada para o excesso de MSE para o algoritmo ℓ_1 -CNLMS, considerando-se valores de $k \rightarrow \infty$, é dado por:

$$\begin{aligned} J_{\text{exc}} &= J(\infty) - J_{\text{min}} \\ &\approx \alpha_1 \mu_0 \sigma_N^2 \frac{(N - L - 1)(\mu_0^2 - \mu_0 + 2) + (2 - \mu_0)(1 - \mu_0)}{(2 - \mu_0)(\mu_0^2 - \mu_0 + 2)} + \alpha_2 \sigma_x^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2. \end{aligned} \quad (4.34)$$

Para valores de $\mu_0 \ll 1$, J_{exc} é calculado como:

$$\begin{aligned} J_{\text{exc}} &\approx \alpha_1 \mu_0 \sigma_N^2 \frac{(N - L - 1)}{2} + \alpha_2 \sigma_x^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2 \\ &\approx \frac{\alpha_1 \mu_0 \sigma_N^2}{2} \{ \text{tr}(\mathbf{PAB}) - 1 \} + \alpha_2 \sigma_x^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2. \end{aligned} \quad (4.35)$$

O Desajuste é definido como a relação entre o valor do Excesso de MSE no estado permanente $J_{\text{ex}}(\infty)$ e o valor do MSE mínimo J_{min} , assim:

$$\begin{aligned} \mathcal{M} &= \frac{J_{\text{exc}}(\infty)}{J_{\text{min}}} \\ &\approx \alpha_1 \mu_0 \frac{(N - L - 1)(\mu_0^2 - \mu_0 + 2) + (2 - \mu_0)(1 - \mu_0)}{(2 - \mu_0)(\mu_0^2 - \mu_0 + 2)} + \alpha_2 \left(\frac{\sigma_x}{\sigma_N} \right)^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2. \end{aligned} \quad (4.36)$$

Para valores de $\mu_0 \ll 1$, o Desajuste \mathcal{M} pode ser aproximado como:

$$\mathcal{M} \approx \frac{\alpha_1 \mu_0}{2} \{\text{tr}(\mathbf{PAB}) - 1\} + \alpha_2 \left(\frac{\sigma_x}{\sigma_N} \right)^2 (1 - \alpha)^2 \|\mathbf{w}_U\|_1^2. \quad (4.37)$$

Para o caso particular, quando não se considera a restrição de norma ℓ_1 , i.e., para $\alpha = 1$, o termo J_C se torna nulo e os resultados obtidos pelas Eqs. (4.35) e (4.37), para $\mu_0 \ll 1$, podem ser reescritos como:

$$J_{\text{exc}} \approx \frac{\mu_0 \sigma_N^2}{2N \sigma_x^2} \{\text{tr}(\mathbf{PAB}) - 1\}. \quad (4.38)$$

$$\mathcal{M} \approx \frac{\mu_0}{2N \sigma_x^2} \{\text{tr}(\mathbf{PAB}) - 1\}. \quad (4.39)$$

Na Seção 6.3, são apresentadas simulações que permitem validar as aproximações adotadas para o cálculo das expressões do Excesso de MSE e Desajuste.

Capítulo 5

Síntese de Arranjos de Sensores

Neste capítulo são apresentados resultados para síntese de arranjos esparsos. Para tanto, procurou-se realizar simulações de arranjos de sensores empregados na prática, escolhendo os sistemas de comunicações militares via satélite como aplicação. Como simulação final, é apresentada a síntese de Arranjo Uniforme Linear (*Uniform Linear Array* (ULA) sob severo cenário de interferência.

Os sistemas via satélite, em geral, apresentam limitações em termos de sistemas de alimentação. Uma outra característica específica de sistemas de comunicações militares via satélite é a grande flutuação no número de usuários, ora alto, ora baixo. Além do citado, nas comunicações militares via satélite há a real possibilidade de interferências intencionais de fontes emissoras, cuja localização pode estar fora do teatro de operações, muitas em outros territórios.

Uma forma de mitigar os problemas de fonte de energia limitada seria fazer uso de arranjo de sensores adaptativos, cuja solução esparsa, obtida por meio de restrições de norma ℓ_1 , permita otimizar o número de sensores ativos de acordo com a demanda de usuários e sinais interferentes,

Já o problema das interferências intencionais (*jammers*) pode ser resolvido por meio do emprego de arranjos adaptativos com restrições lineares, onde se garantiria um ganho de 0 dB na direção das estações dos usuários, sendo esses fixos ou móveis, e os sinais interferentes seriam normalmente atenuados pela conservação da energia do filtro.

Como os algoritmos propostos neste trabalho atendem simultaneamente às necessidades supracitadas, esses algoritmos são adequados à solução do problema. Nas seções seguintes, são apresentados resultados de síntese de arranjo de sensores fazendo uso dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS¹. Os resultados das simulações apresentados nesta seção foram obtidos considerando 200 realizações independentes.

¹Não serão contempladas simulações que empreguem o algoritmo ℓ_1 -CLMS, cujo propósito de sua apresentação foi tão somente servir de alicerce teórico para a dedução das versões normalizada e ponderada.

Neste ponto, é importante discorrer sobre o limite inferior para restrição de norma ℓ_1 empregada para os algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS. Nas aplicações de síntese de arranjo de sensores, que, usualmente, estabelecem como restrição linear ganho unitário na direção do sinal de interesse, o limite inferior para essa restrição vale 1. Esta afirmação baseia-se no fato de que os valores absolutos dos elementos das colunas da matriz de restrições \mathbf{C} valem 1, portanto, para que a restrição $\mathbf{C}^H \mathbf{w} = 1$ seja satisfeita, o limite inferior de $\|\mathbf{w}\|_1$ deve ser $t=1$. Este limite corresponde à norma ℓ_1 da solução trivial para um arranjo do tipo ULA quando não há presença de interferências e a direção do sinal de interesse é 90° .

Um outro ponto a ser destacado é o procedimento de escolha dos passos fixos de convergência μ_0 para as simulações apresentadas neste trabalho. Como os resultados são apresentados para diferentes algoritmos, o critério de escolha de μ_0 foi o MSE, i.e., os valores de μ_0 são ajustados de maneira que o MSE obtido por cada algoritmo seja o mesmo.

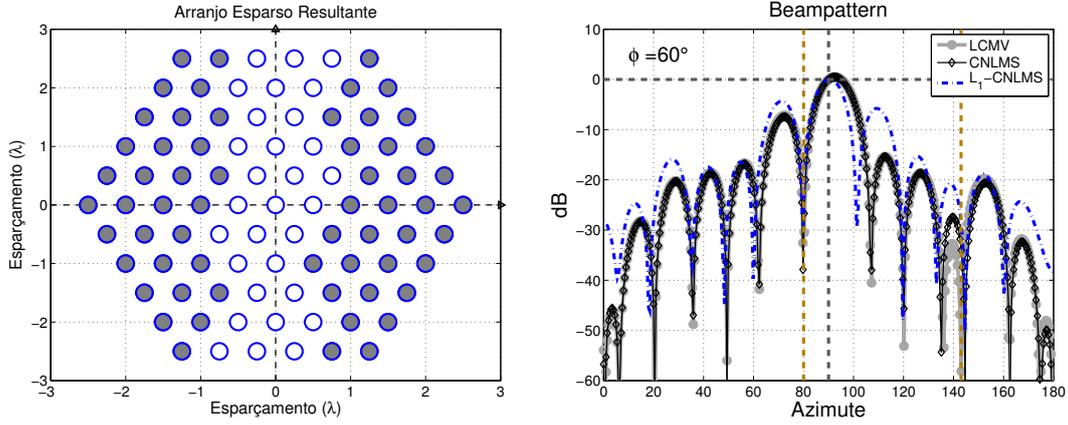
5.1 Arranjo Hexagonal Padrão (*Standard Hexagonal Array - SHA*) para Satélites Geoestacionários Militares na Banda- X

A primeira síntese de arranjos contempla um arranjo de antenas (sensores) do tipo SHA empregado em sistemas de comunicações militares via satélites geoestacionários que operam na banda- X [10] [32].

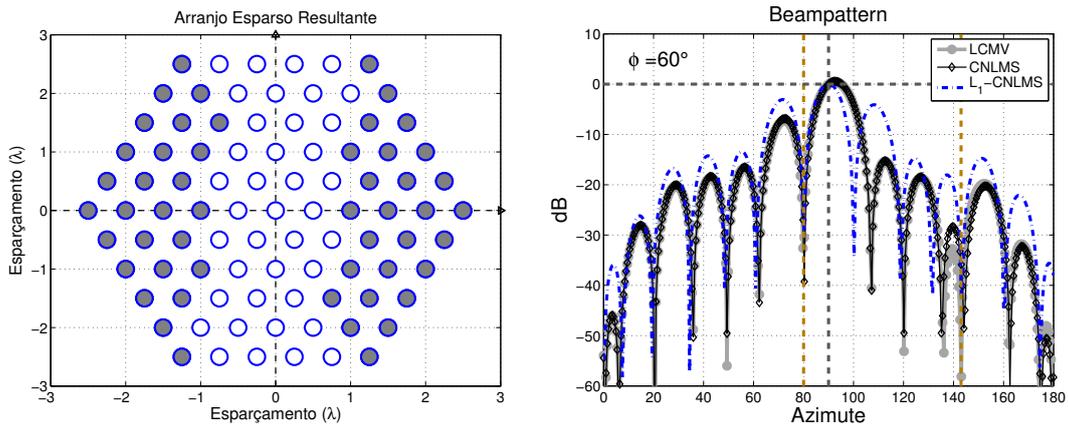
Este arranjo do tipo SHA é formado por 6 elementos em cada aresta, resultando num total de 91 elementos, conforme pode ser observado na Fig. 5.1. Como sinais de entrada, ou sinais recebidos pelo arranjo, são considerados três sinais analíticos de banda estreita com modulação do tipo QPSK [33]. Os demais parâmetros empregados na simulação e síntese são listados na Tabela 5.1

Tabela 5.1: Parâmetros empregados nas simulações das Seções 5.1 e 5.2

Parâmetro	Simulação - SHA	Simulação - HAA
μ_0 (CNLMS)	2×10^{-3}	2×10^{-2}
μ_0 (ℓ_1 -CNLMS)	1×10^{-3}	1×10^{-2}
μ_0 (ℓ_1 -CWNLMS)	1×10^{-3}	1×10^{-2}
β (ℓ_1 -WCNLMS)	6	15
Separação dos Elementos	$\frac{\lambda}{2}$	λ
$\ \mathbf{w}_{LCMV}\ _1$	1,10	1,50
Restrição de Norma ℓ_1 (t)	1,00	1,05
Frequências dos Sinais	8 GHz	12 GHz
Sinal de Interesse (θ_S, ϕ_S)	(90°;60°)	(90°;45°)
Signal SNR	20 dB	20 dB
Interferência 1 (θ_1, ϕ_1)	(80°;60°)	(87,50°; 45°)
Interferência 2 (θ_2, ϕ_2)	(143°; 60°)	(88,75°; 47,50°)
Interferência 3 (θ_3, ϕ_3)	-	(91,25°; 47,50°)
Interferência 4 (θ_4, ϕ_4)	-	(92,50°; 45°)
Interferências 1, 2, 3, 4 SNR	40 dB	40 dB



(a) Na iteração $k = 15 \times 10^3$



(b) Na iteração $k = 30 \times 10^3$

Figura 5.1: Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -CNLMS e respectivos *beampatterns* computados utilizando-se os algoritmos ℓ_1 -CNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

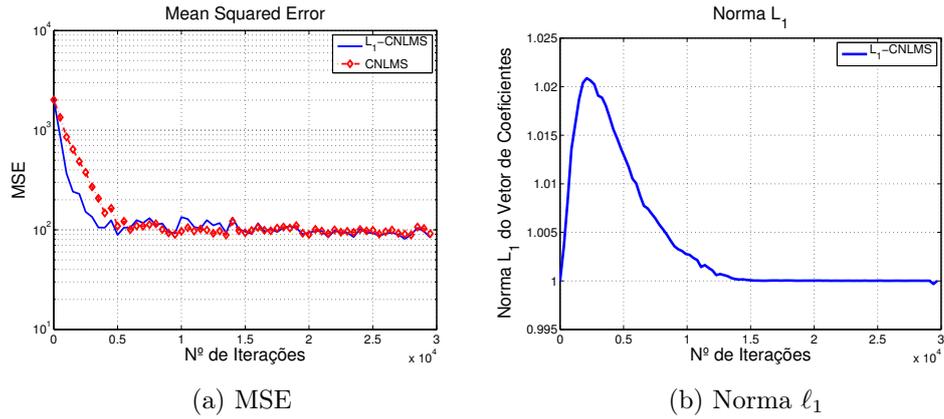


Figura 5.2: Síntese do arranjo SHA com 91 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -CNLMS e CNLMS.

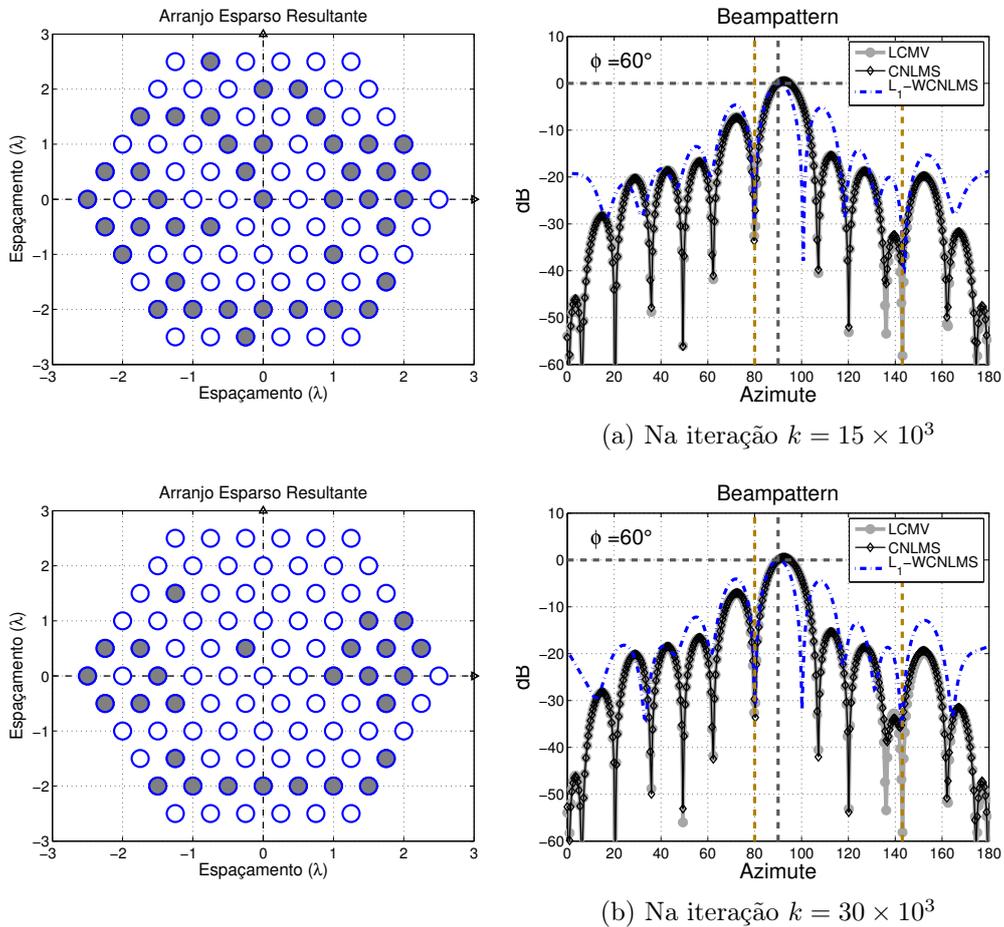


Figura 5.3: Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -WCNLMS e respectivos *beam patterns* computados utilizando os algoritmos ℓ_1 -WCNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

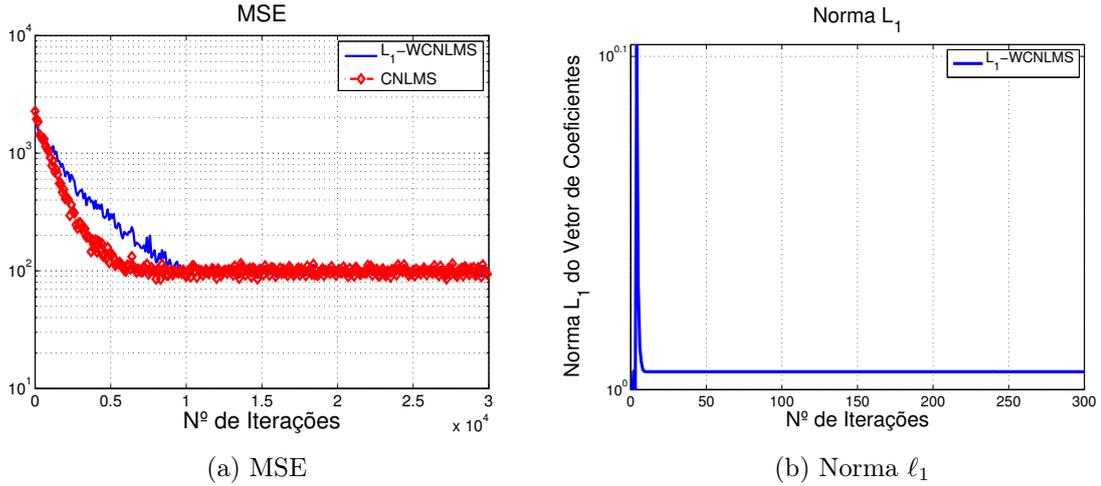


Figura 5.4: Síntese do arranjo SHA com 91 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparso obtidos por meio dos algoritmos ℓ_1 -WCNLMS e CNLMS.

Os resultados apresentados pelas Figs 5.1 e 5.3 representam arranjos adaptativos sintetizados pelos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, respectivamente. Esses resultados são comparados, em termos de MSE e diagramas de radiação, como o resultado obtido pelo algoritmo CNLMS. Foram atingidos níveis de esparsidade de 33% e 70%, para os algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, respectivamente. Observou-se também, que os resultados, em termos de *beam pattern* e MSE, se mantiveram aceitáveis, mesmo para o caso mais esparso.

5.2 Arranjo Hexagonal (HAA) para Comunicações por Satélite na Banda- Ku

Esta simulação tem como propósito sintetizar um arranjo esparso a partir de um arranjo do tipo HAA, normalmente, usado em satélites de comunicações que operam na Banda- Ku . Cada aresta do arranjo hexagonal possui 12 elementos, perfazendo um total de 397 elementos para o arranjo.

O esquema proposto, baseado em um arranjo de antenas de um satélite de comunicações real [34], deve ser capaz de iluminar seu *footprint*² com 64 feixes, cujo ângulo médio (centro-a-centro) entre os feixes que compartilham um mesmo grupo de frequências vale 2.5° . De maneira a evitar que os feixes adjacentes partilhem do mesmo grupo de frequências, emprega reuso de frequências a uma taxa de 1 : 4. Apesar da taxa de reuso de frequências, decorrente da estreita separação entre os

²Footprint de um satélite de comunicações é a área, na terra, iluminada pelo sistema irradiante do satélite. De maneira prática, corresponde à porção da terra onde os transponders do satélite oferecem cobertura.

feixes que utilizam o mesmo grupo de frequências, sob certas condições de potência, ganho das antenas e ângulo de visada, pode haver interferências mútuas provocadas pelos usuários do sistema, conforme ilustrado na Fig. 5.5.

Os satélites de comunicações geoestacionários empregam arranjo de sensores, cujos elementos (antenas) apresentam ganhos altos. Desta forma, alguns arranjos utilizam antenas de abertura cujas dimensões físicas não permitem uma separação de meio comprimento de onda. Nesta simulação, o espaçamento utilizado é de um comprimento de onda. Como sinais de entrada e interferências, são considerados cinco sinais analíticos de banda estreita como modulação QPSK [33], conforme disposição apresentada na Fig. 5.5.

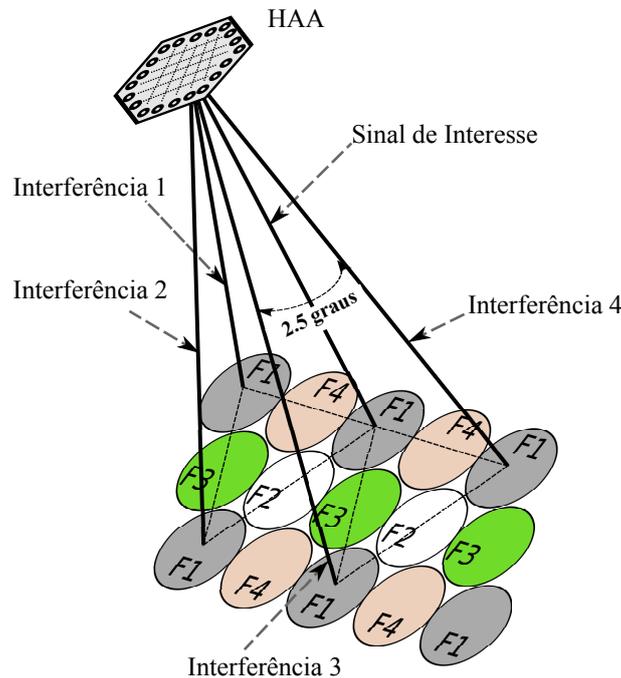


Figura 5.5: Ilustração de feixes de cobertura em uma porção do footprint de um satélite geoestacionário. Considera-se, neste caso, uma taxa de reuso de frequências de 1 : 4, cuja função é garantir que nenhum feixe adjacente utilize uma mesma frequência ou um mesmo grupo de frequências.

Com esta simulação pretende-se também mostrar o efeito da interferência co-canal de sinais de banda estreita, bem como a atenuação que o algoritmo impõe aos sinais interferentes.

As sínteses dos arranjos adaptativos esparsos obtidos com os algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS constam do conjunto de figuras 5.7 - 5.11. Esses resultados são comparados, em termos de MSE e diagramas de radiação, como o resultado obtido pelo algoritmo CNLMS. Foram atingidos níveis de esparsidade de 61% e 80%, para os algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, respectivamente. Observou-se também que os resultados, em termos de *beam pattern* e MSE, se mantiveram aceitáveis, mesmo para o caso mais esparsos, uma vez que mantiveram o ganho unitário na

direção de interesse, ao mesmo tempo que atenuaram os sinais interferentes.

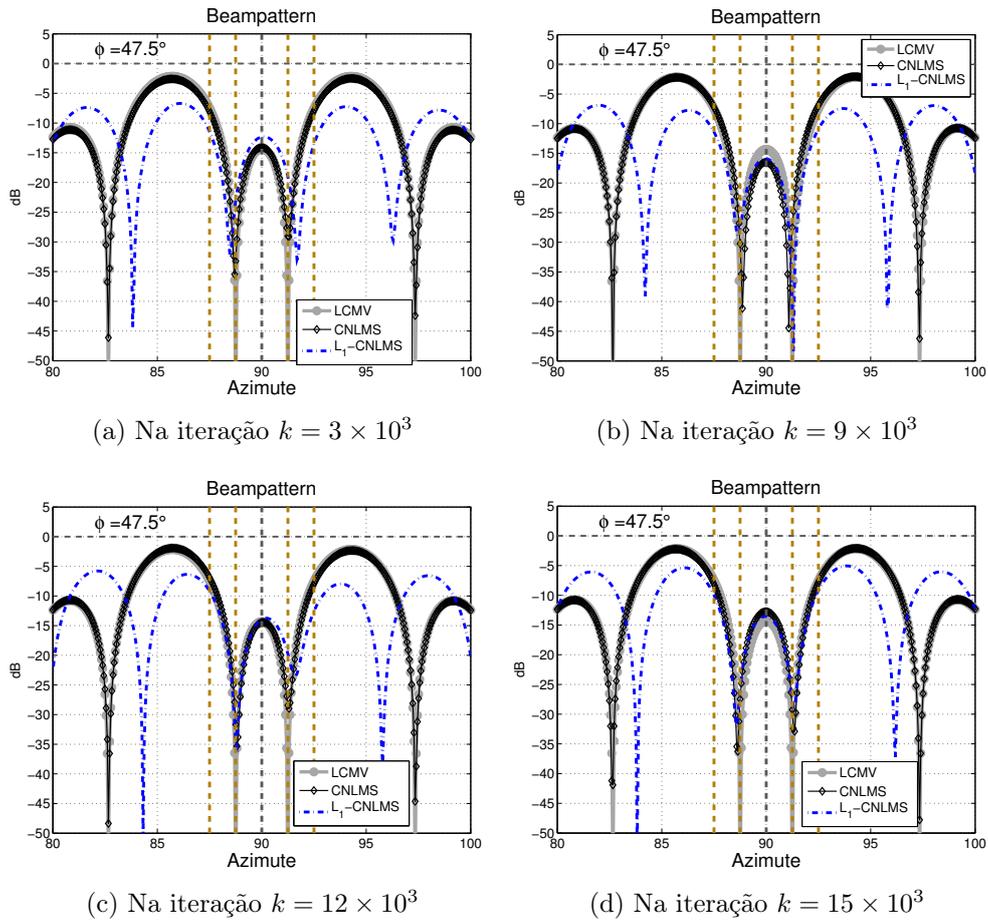


Figura 5.6: Sínteses para uma arranjo tipo HAA via algoritmos L_1 -CNLMS, CNLMS e solução LCMV, observado de um zênite $\phi = 47.5^\circ$.

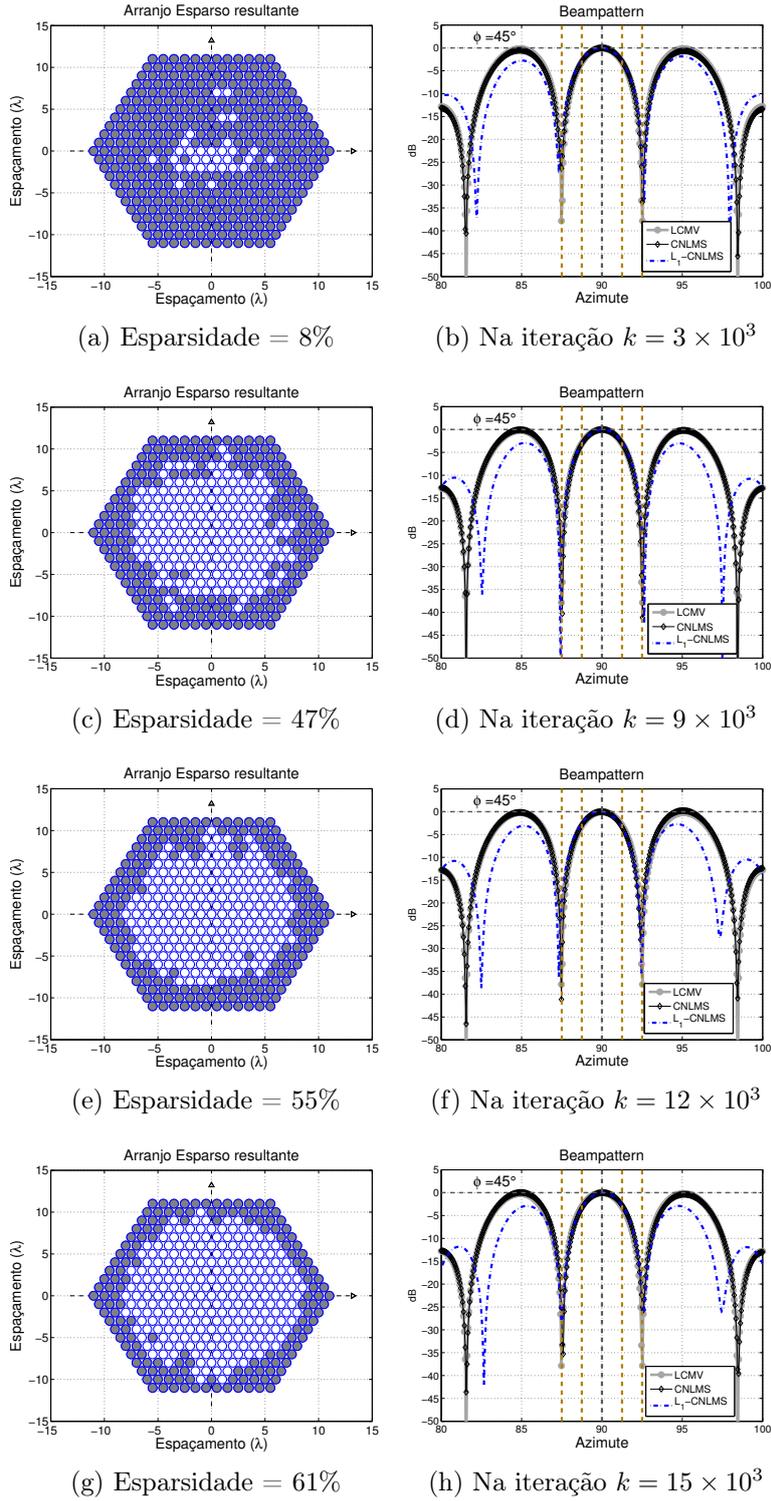


Figura 5.7: Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -CNLMS e respectivos *beam patterns* computados utilizando-se os algoritmos ℓ_1 -CNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

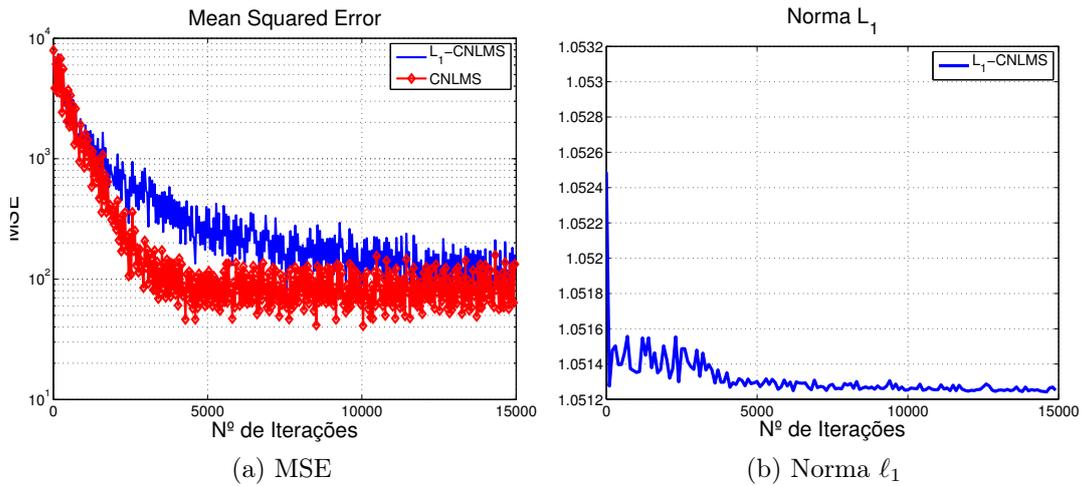


Figura 5.8: Síntese do arranjo HAA com 397 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -CNLMs e CNLMs.

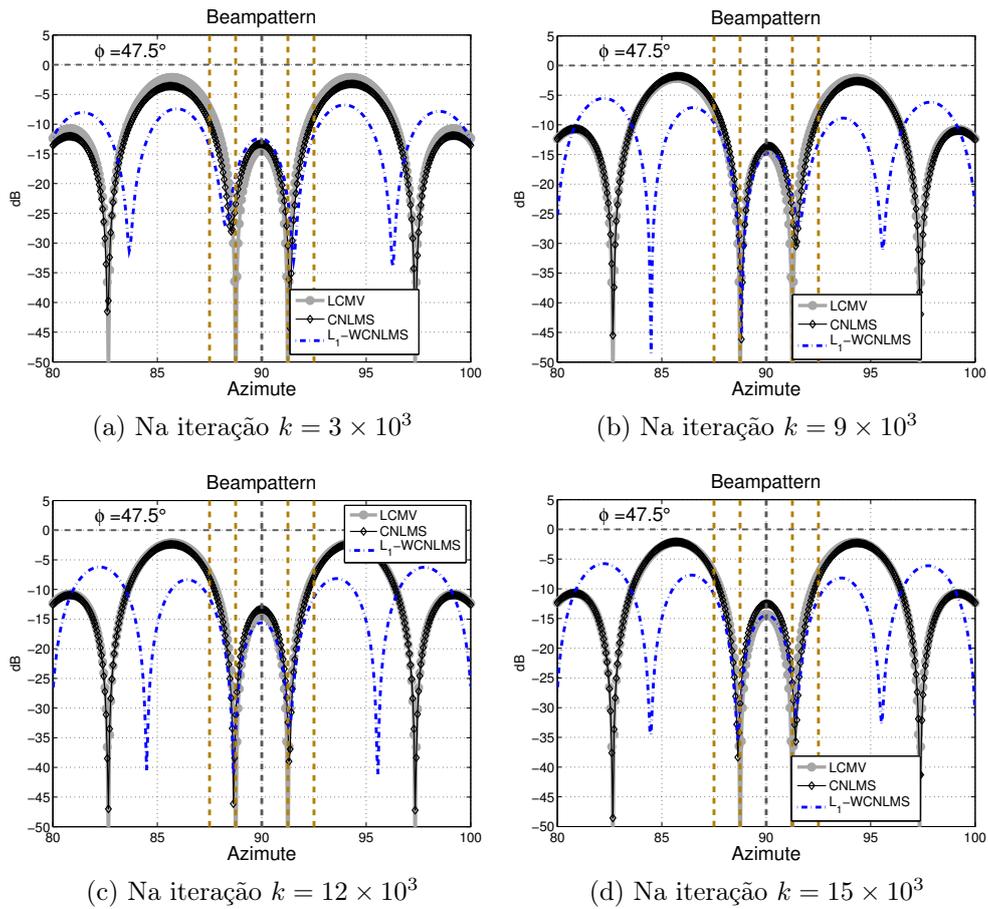


Figura 5.9: Sínteses para uma arranjo tipo HAA via algoritmos L_1 -WCNLMs, CNLMs e solução LCMV, observado de um zênite $\phi = 47.5^\circ$.

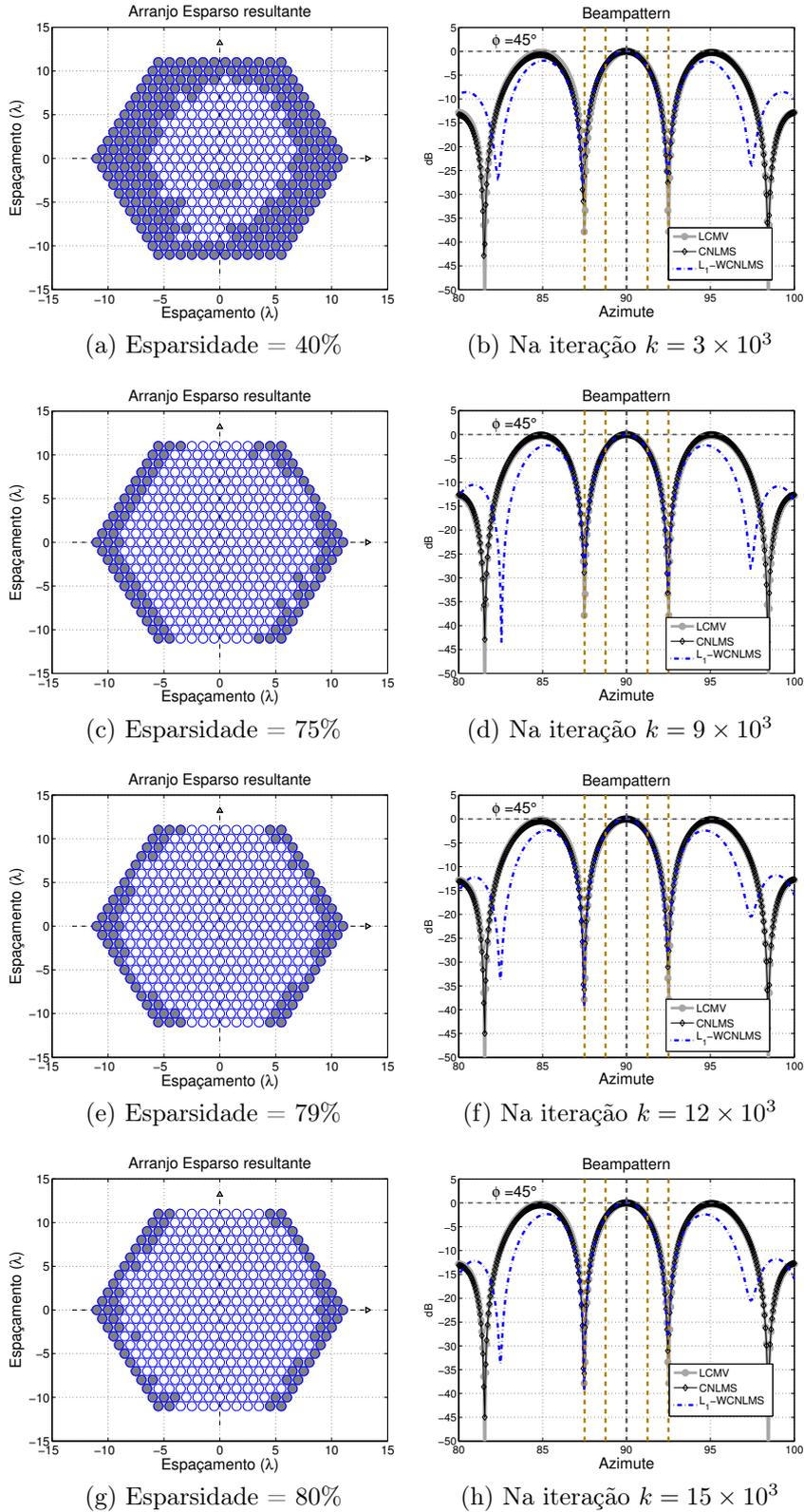


Figura 5.10: Arranjos esparsos sintetizados por meio do algoritmo ℓ_1 -WCNLMS e respectivos *beampatterns* computados utilizando-se os algoritmos ℓ_1 -WCNLMS, CNLMS e solução LCMV. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

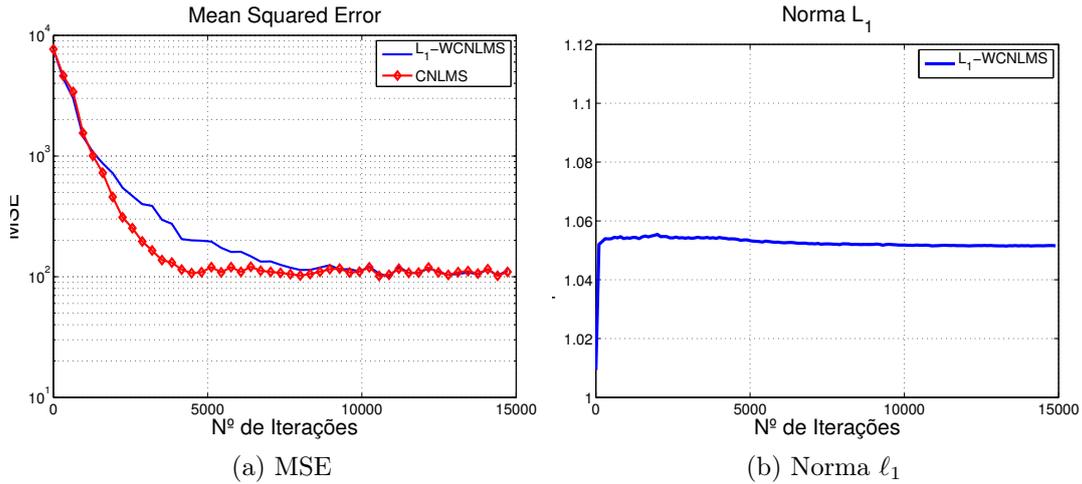


Figura 5.11: Síntese do arranjo HAA com 397 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos ℓ_1 -WCNLMS e CNLMS.

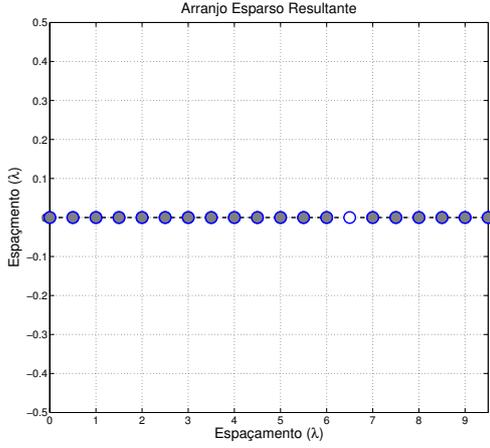
5.3 Arranjo Uniforme Linear

Esta seção apresenta a síntese de um Arranjo Uniforme Linear (*Uniform Linear Array* - ULA) com 20 elementos, cujo sinal de interesse é um envelope complexo na presença de 6 *jammers* (interferidores). Com o intuito de simular um cenário severo de interferência, dentre os seis sinais interferentes, dois sinais possuem ângulos de incidência quase ortogonais ($\pm 3^\circ$). Três valores diferentes de restrição de norma ℓ_1 são utilizados, sendo que o menor deles está bem próximo ao limite $t = 1$.

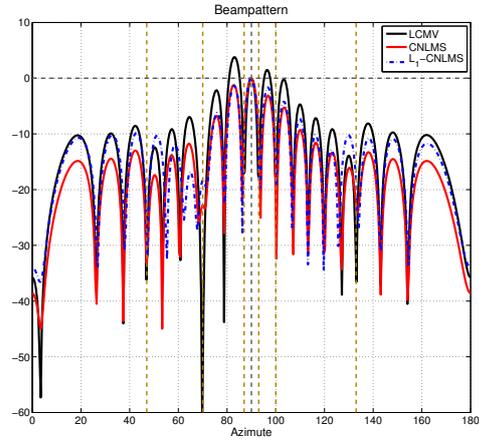
As simulações I, II e III apresentam níveis de esparsidade distintos em decorrência de diferentes valores de β e de t . Os resultados das simulações apresentados nesta seção foram obtidos considerando 200 realizações independentes.

Na Tabela 5.2, estão listados os demais parâmetros das simulações desta seção.

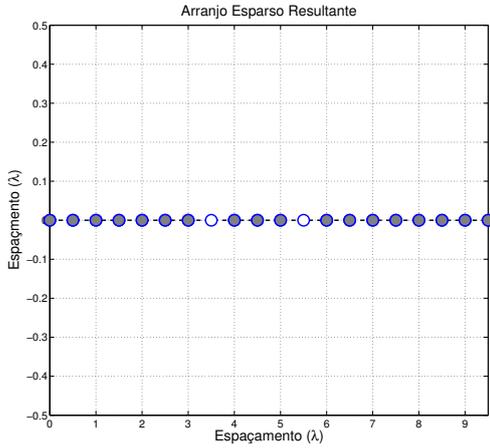
5.3.1 ULA: Simulação I



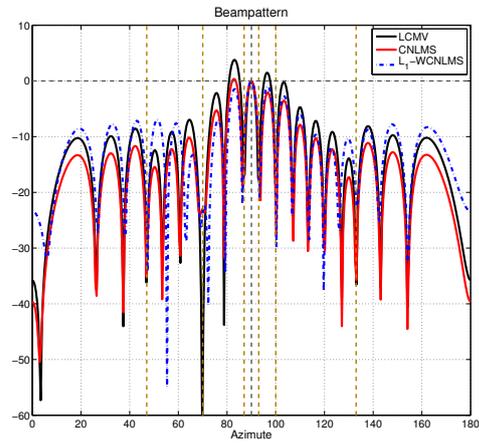
(a) Esparsidade = 5%



(b) ℓ_1 -CNLMS, iteração $k = 30 \times 10^3$



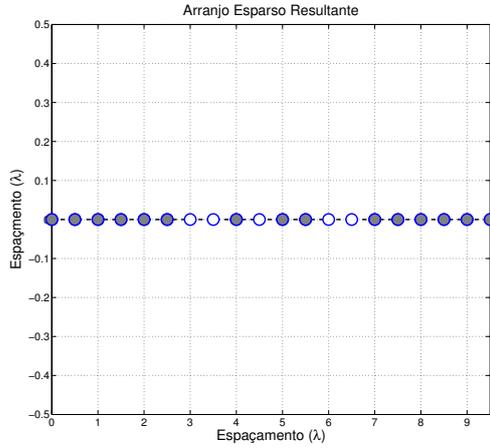
(c) Esparsidade = 10%



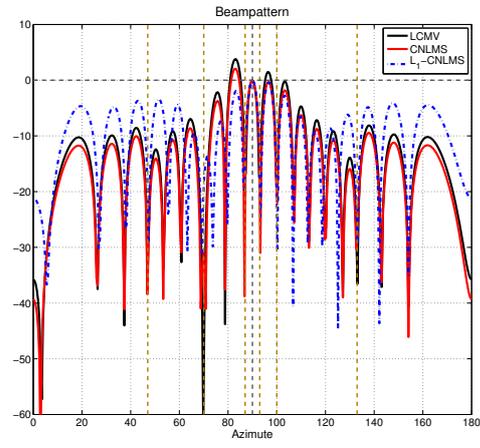
(d) ℓ_1 -WCNLMS, iteração $k = 30 \times 10^3$

Figura 5.12: Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos *beampatterns* computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.62t_o$ e $\beta = 1$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

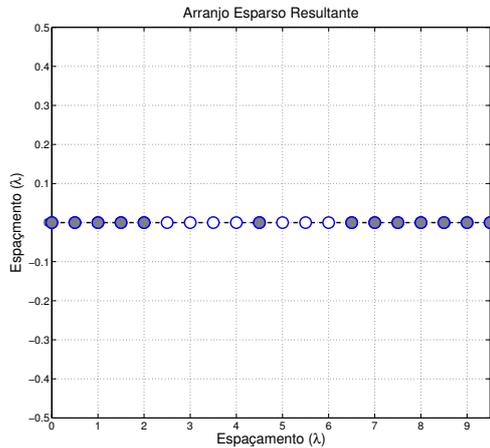
5.3.2 ULA: Simulação II



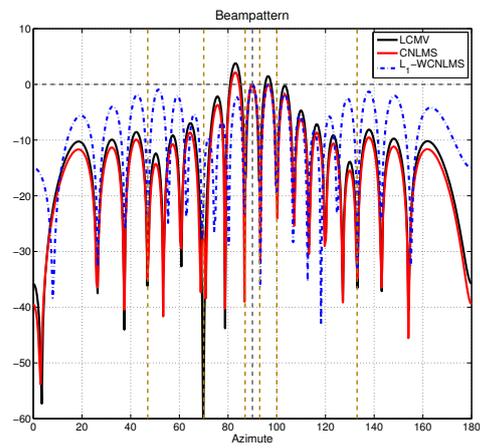
(a) Esparsidade = 25%



(b) ℓ_1 -CNLMS, iteração $k = 30 \times 10^3$



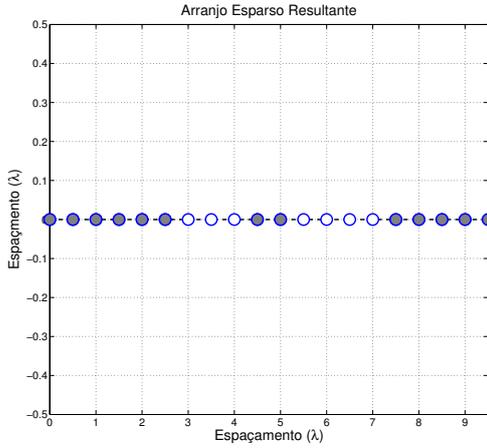
(c) Esparsidade = 35%



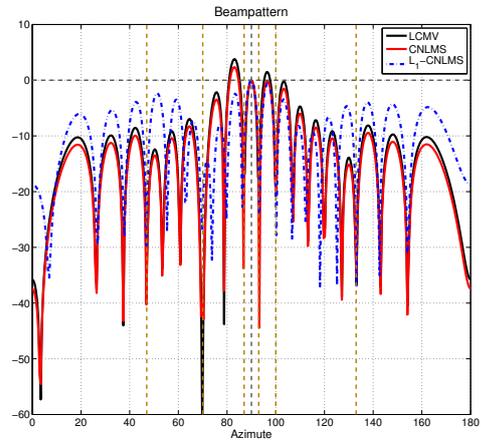
(d) ℓ_1 -WCNLMS, iteração $k = 30 \times 10^3$

Figura 5.13: Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos *beampatterns* computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.58t_o$ e $\beta = 2$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.

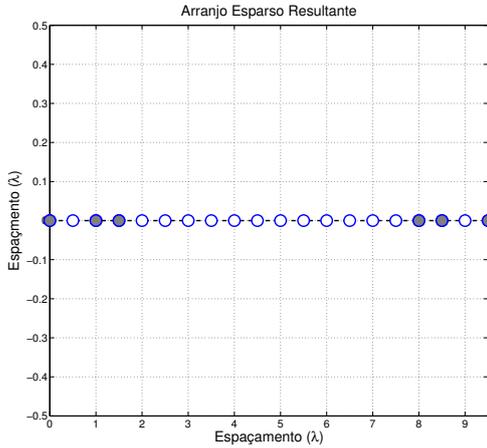
5.3.3 ULA: Simulação III



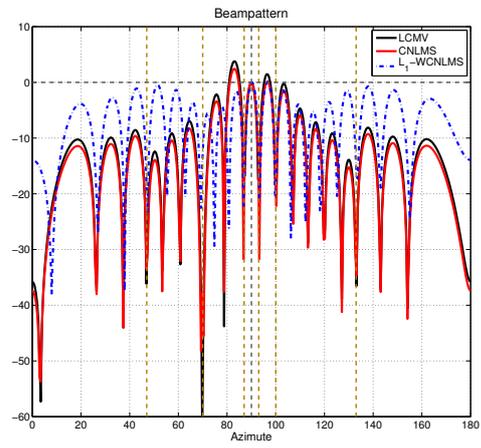
(a) Esparsidade = 35%



(b) ℓ_1 -CNLMS, iteração $k = 30 \times 10^3$

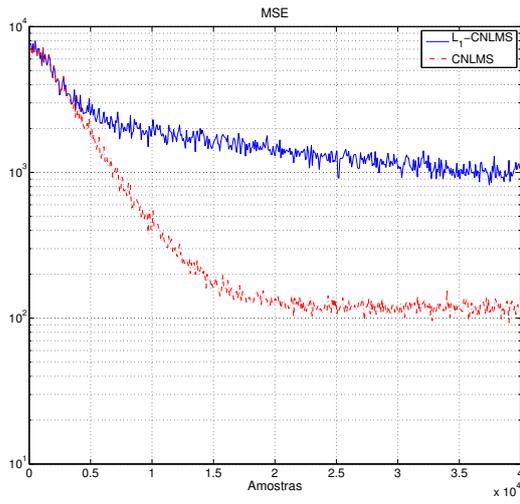


(c) Esparsidade = 70%

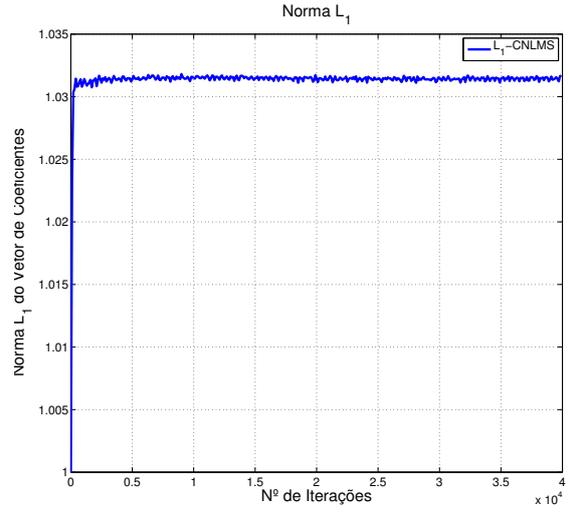


(d) ℓ_1 -WCNLMS, iteração $k = 30 \times 10^3$

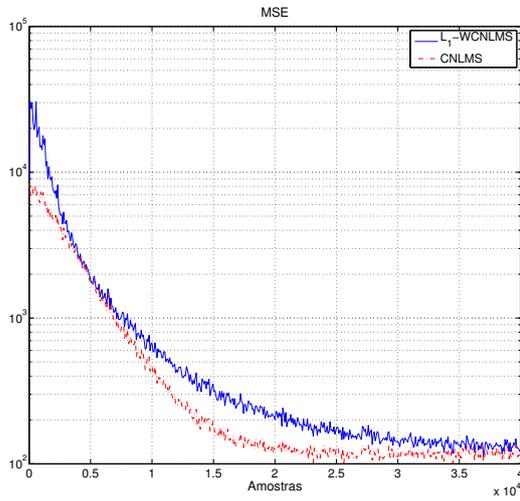
Figura 5.14: Arranjos esparsos sintetizados por meio dos algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS, e respectivos *beampatterns* computados utilizando-se os algoritmos ℓ_1 -CNLMS, ℓ_1 -WCNLMS, CNLMS e solução LCMV, com restrição de norma ℓ_1 igual a $t = 0.56t_o$ e $\beta = 2$. Os elementos denotados por círculos brancos, representam aqueles elementos que foram desligados, isto é, elementos cujo coeficiente é nulo. Na disposição dos sensores sobre o plano, toma-se como referência para o azimute o eixo das abscissas e o sentido anti-horário.



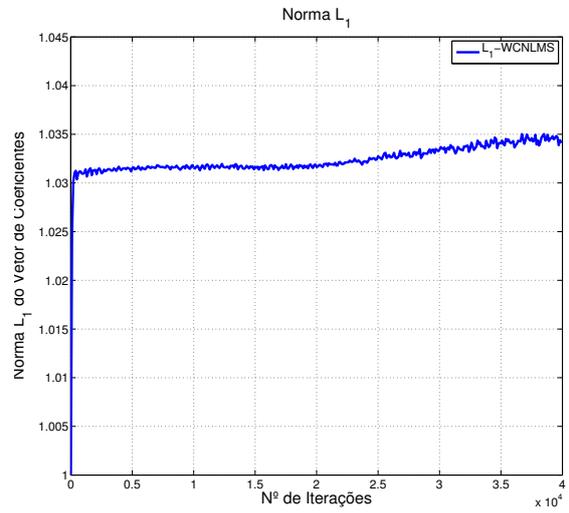
(a) MSE



(b) Norma ℓ_1



(c) MSE



(d) Norma ℓ_1

Figura 5.15: Síntese do arranjo ULA com 20 elementos: Erro Médio Quadrático (MSE) e Norma ℓ_1 para o vetor de coeficientes esparsos obtido por meio dos algoritmos l_1 -CNLMS, l_1 -WCNLMS e CNLMS na Simulação III.

Tabela 5.2: Parâmetro empregados nas simulações da Seção 5.3

Parâmetros	I	II	III
μ_0 (CNLMS)	7×10^{-3}	1×10^{-3}	7×10^{-3}
μ_0 (ℓ_1 -CNLMS)	7×10^{-3}	7×10^{-3}	7×10^{-3}
μ_0 (ℓ_1 -WCNLMS)	7×10^{-3}	7×10^{-3}	7×10^{-3}
β (ℓ_1 -WCNLMS)	1	2	2
Separação dos Elementos	$\frac{\lambda}{2}$	$\frac{\lambda}{2}$	$\frac{\lambda}{2}$
Restrições de Norma ℓ_1	$0.62t_o$	$0.58t_o$	$0.56t_o$
$\ \mathbf{w}_{LCMV}\ _1$	1,82	1,82	1,82
Frequência dos Sinais (kHz)	1	1	1
Pot. Sinal SNR (dB)	20	20	20
Pot. interferências SNR (dB)	40	40	40
Sinal de Interesse (θ_S)	90°	90°	90°
Interferência 1 (θ_1)	87°	87°	87°
Interferência 2 (θ_2)	93°	93°	93°
Interferência 3 (θ_3)	70°	70°	70°
Interferência 4 (θ_4)	110°	110°	110°
Interferência 5 (θ_5)	15°	15°	15°
Interferência 6 (θ_6)	155°	155°	155°

5.4 Conclusão do Capítulo

Conforme pode ser observado por meio dos resultados apresentados neste Capítulo, os algoritmos ℓ_1 -CNLMS e ℓ_1 -WCNLMS permitem sintetizar arranjos de sensores planares e lineares por meio de soluções esparsas. Pôde-se observar também, que mesmo para elevados níveis de esparsidade, a restrição linear de ganho unitário na direção do sinal de interesse, bem como a atenuação dos sinais interferentes, foram satisfeitas.

Com base nesses resultados, também foi notado um desempenho bem superior do algoritmo ℓ_1 -WCNLMS, em termos de esparsidade, em relação à versão não ponderada ℓ_1 -CNLMS. Contudo, não se observa um desempenho diferenciado entre esses algoritmos quanto à velocidade de convergência. Este fato decorre dos arranjos, em princípio, não serem sistemas esparsos, se comportando, nesses casos, como sistemas compressíveis. Para problemas envolvendo arranjo de sensores, pode-se medir a compressibilidade por meio da norma ℓ_1 da solução LCMV, que apresentando valores maiores que 1, indica que a solução pode ser encolhida, e, portanto, o arranjo pode ser considerado compressível uma vez que o limite inferior para $\|\mathbf{w}\|_1$ é igual a

1. Estimar a compressibilidade de um arranjo de sensores, baseando-se na norma ℓ_1 da solução LCMV, se justifica pelo fato da solução \mathbf{w}_{LCMV} ser uma solução ótima em termos de mínima potência do sinal. Portanto, a compressibilidade nos arranjos de sensores implica a possibilidade de encolhimento da solução, permitindo impor-lhe esparsidade por meio do estabelecimento de um limiar de valor absoluto para ativar ou desativar o sensor correspondente.

Em decorrência do algoritmo ℓ_1 -WCNLMS não atuar de maneira uniforme em todos os coeficientes do vetor de pesos, tem-se uma configuração do arranjo esparsa resultante menos uniforme e, por vezes, aleatória, mantendo-se, no entanto, o grau de esparsidade de acordo com os parâmetros de inicialização do algoritmo. Por outro lado, o algoritmo ℓ_1 -CNLMS produz um arranjo esparsa mais uniforme, atendendo a um padrão que depende, além dos parâmetros de inicialização do algoritmo, das direções do sinal de interesse e do conjunto de interferidores.

Na Seção 5.3, pode ser observado que, tanto o valor imposto para restrição t , como o valor de β , influenciam sobremaneira no “encolhimento” da solução e, conseqüentemente, no nível de esparsidade do arranjo final.

Ainda sobre a Seção 5.3, outra importante observação é que os diagramas de radiação resultantes apresentaram um aumento no nível dos lóbulos secundários. A explicação para este fato é a necessidade de redistribuição de energia nas direções mais afastadas do centro do arranjo. Isto é necessário para que se possa atenuar os dois sinais interferentes, cujos ângulos de incidência são muito próximo a 90° , sem no entanto, comprometer o ganho unitário na direção do sinal de interesse 90° .

Capítulo 6

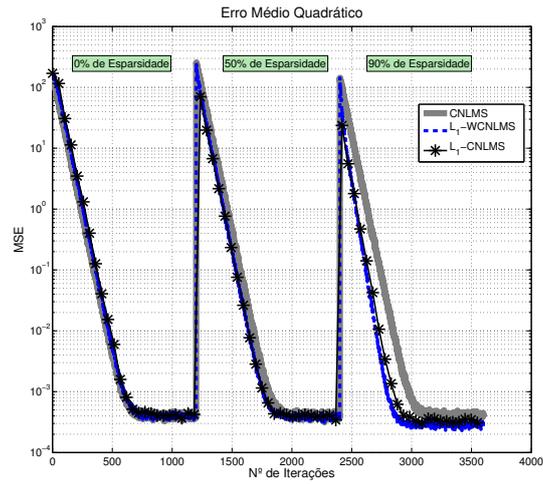
Simulações Adicionais

Neste capítulo são evidenciadas as principais diferenças, em termos de convergência, existentes entre aos algoritmos L_1 -CNLMS e L_1 -WCNLMS, quando da identificação de sistemas esparsos e compressíveis [35]. Também são apresentados os valores de $MSE(\infty)$ e Desajuste, calculados por meio das expressões aproximadas dadas pelas Eqs. (4.33) e (4.36). Esses resultados são comparados com os resultados produzidos pelo algoritmo ℓ_1 -CNLMS para filtros de várias ordens e passos fixos de convergência diferentes.

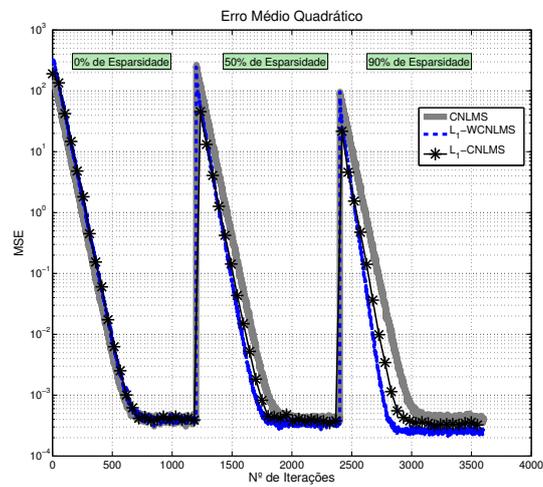
6.1 Sistema Esparso Complexo

Nesta seção, considera-se o problema de identificação de um sistema que, ao longo do tempo, tem o nível de esparsidade do filtro (coeficientes) assumindo três valores distintos: 0%, 50% e 90%. Na solução deste problema, são empregados os algoritmos L_1 -CNLMS e L_1 -WCNLMS, tendo como resultados um quadro comparativo da velocidade de convergência, em termos de MSE, para valores de β iguais a: 0,1; 1; e 7.

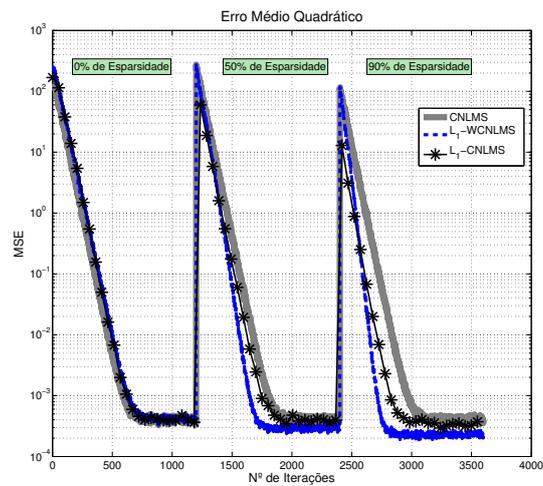
Cabe destacar ainda que o filtro utilizado possui 50 coeficientes complexos, gerados aleatoriamente cada realização e, para cada terço do número total de iterações, um número de coeficientes correspondente à esparsidade desejada é anulado. Assume-se também, que o sinal de referência d_k foi contaminado com um ruído aditivo Gaussiano de medição, cuja variância vale $\sigma_N^2 = 10^{-4}$. Os passos de convergência μ_0 utilizados nesta simulação são iguais a 1. Na Fig. (6.1), são apresentados os resultados de MSE obtidos pelos algoritmos L_1 -CNLMS e L_1 -WCNLMS.



(a) $\beta = 0,1$



(b) $\beta = 1$

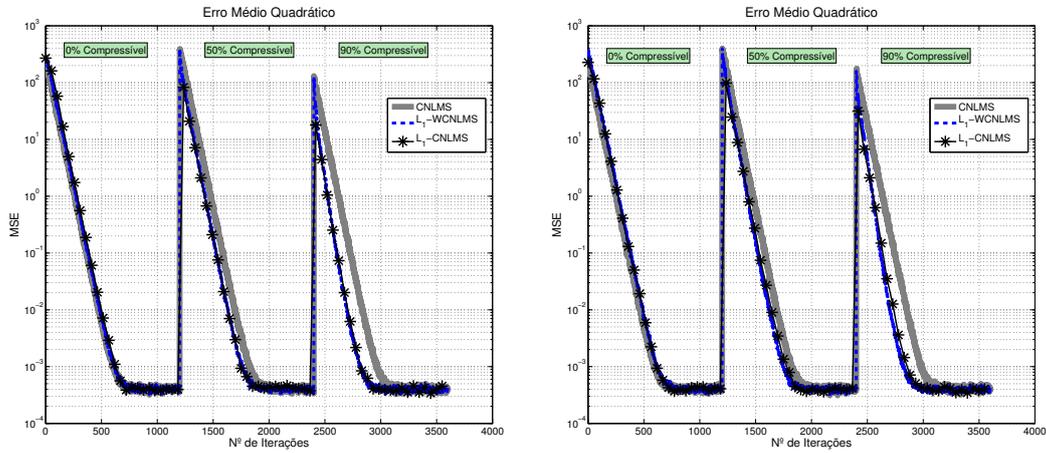


(c) $\beta = 7$

Figura 6.1: MSE obtido para os algoritmos L_1 -CNLMS, L_1 -WCNLMS e CNLMS, quando da identificação de planta complexa esparsa com 50 coeficientes e sinal de referência contaminado com ruído de medição de variância $\sigma_N^2 = 10^{-2}$.

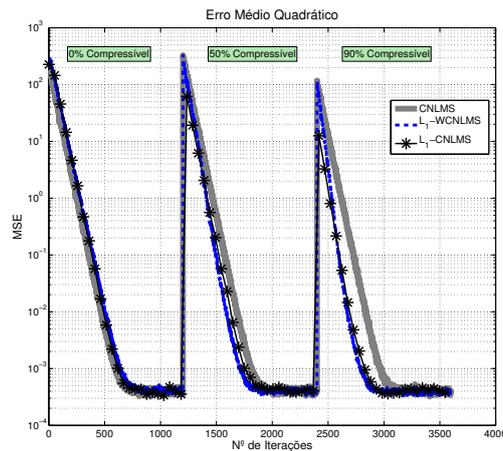
6.2 Sistema Compressível Complexo

Nesta seção, os algoritmos L_1 -CNLMS e L_1 -WCNLMS são avaliados quanto à velocidade de convergência, para valores de β iguais a: 0,1; 1; e 7. A planta de sistema a ser identificada nesta simulação, tem o seu nível de compressibilidade alterado a cada terço do número total de iterações. Este sistema é idêntico ao apresentado na Seção 6.1, exceto pelo fato de que os coeficientes nulos são substituídos por ruído Gaussiano de variância $\sigma_N^2 = 10^{-4}$ [35]. Cabe destacar que o filtro utilizado possui 50 coeficientes complexos, gerados aleatoriamente a cada realização. Assume-se também, que o sinal de referência d_k foi contaminado com um ruído aditivo Gaussiano de medição, cuja variância é $\sigma_N^2 = 10^{-4}$. Na Fig. (6.2), são apresentados os resultados de MSE obtidos pelos algoritmos L_1 -CNLMS e L_1 -WCNLMS.



(a) $\beta = 0,1$

(b) $\beta = 1$

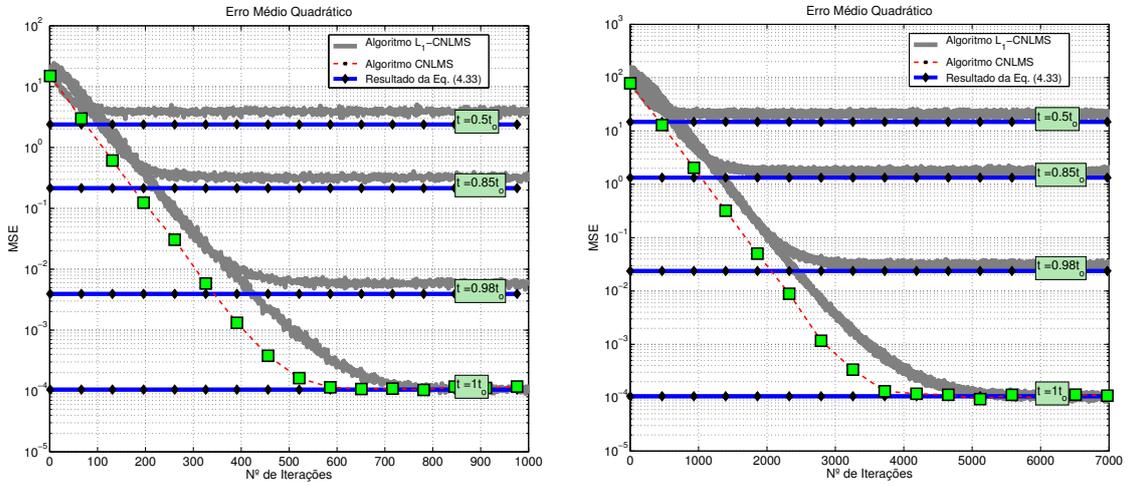


(c) $\beta = 7$

Figura 6.2: MSE obtido para os algoritmos L_1 -CNLMS, L_1 -WCNLMS e CNLMS, quando da identificação de planta complexa compressível com 50 coeficientes e sinal de referência contaminado com ruído de medição de variância $\sigma_N^2 = 10^{-2}$.

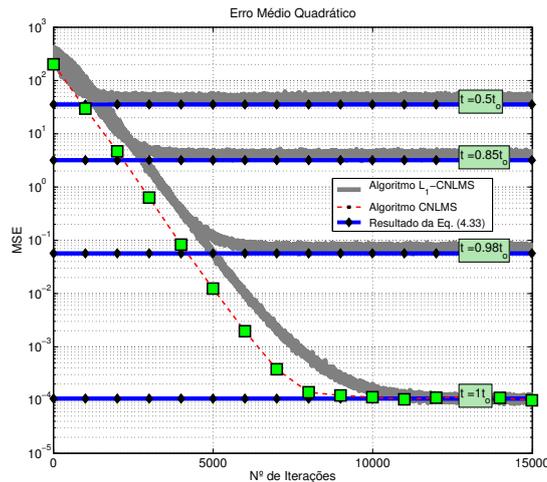
6.3 Excesso de MSE e Desajuste

Nesta seção são apresentadas simulações para $J(\infty) = MSE(\infty)$ e Desajuste, calculados por meio das Eqs. (4.33) e (4.36). Os resultados para $J(\infty)$ são comparados com aqueles obtidos pelo algoritmo L_1 -CNLMS, conforme consta das Figuras 6.3-6.6. Essas simulações empregaram plantas complexas com 10, 50 e 100 coeficientes e um total de 100 realizações foi computado. Foi assumida também a presença de um ruído de medição com variância $\sigma_N^2 = 10^{-4}$. Para o caso de algoritmos com restrições de norma ℓ_1 , $J(\infty)$ é função do quadrado do “encolhimento”, que é representado na Eq. (4.33) pela parcela $(1 - \alpha)^2 \|\mathbf{w}_U\|_1^2$. Portanto, as simulações desta Seção são computadas para valores distintos de restrição de norma ℓ_1 .



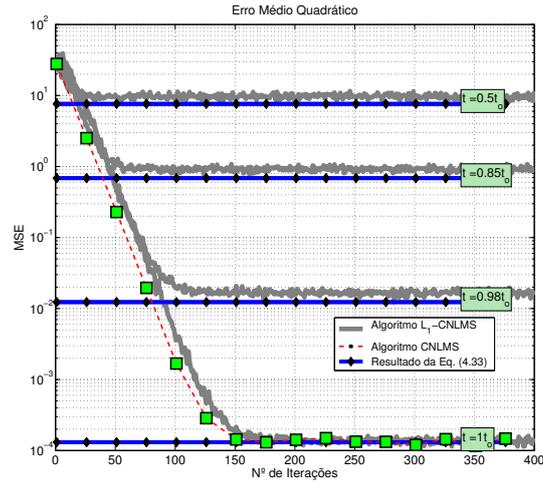
(a) $\mu_0 = 0,1$; $N = 10$.

(b) $\mu_0 = 0,1$; $N = 50$.

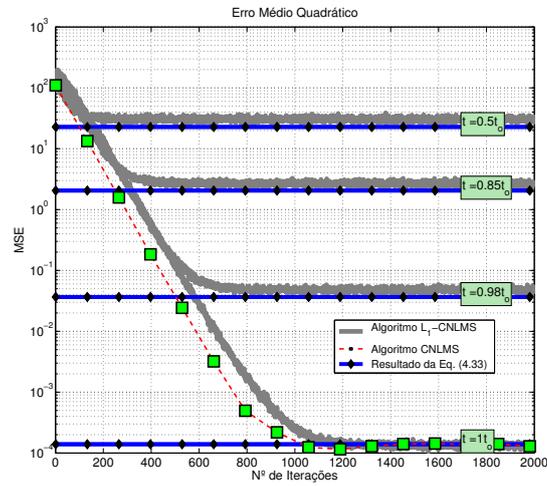


(c) $\mu_0 = 0,1$; $N = 100$.

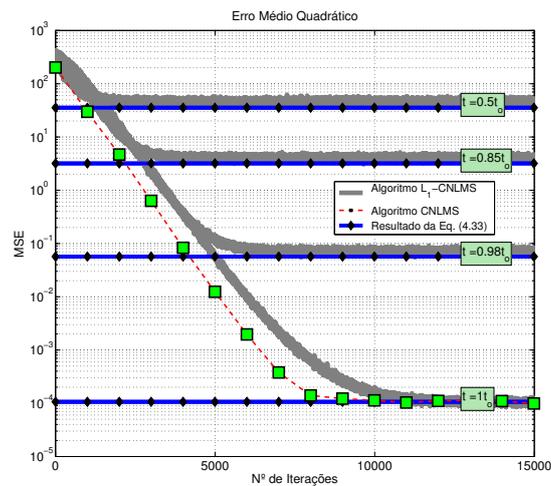
Figura 6.3: Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 0,1$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.



(a) $\mu_0 = 0,5$; $N = 10$.

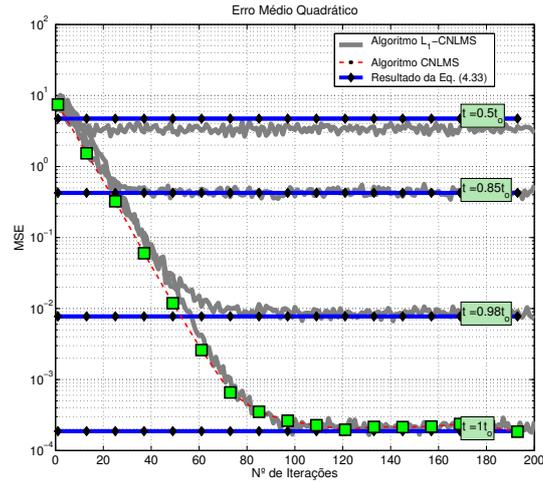


(b) $\mu_0 = 0,5$; $N = 50$.

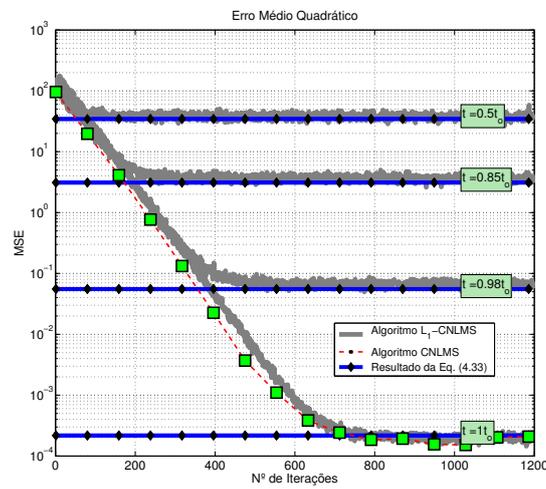


(c) $\mu_0 = 0,5$; $N = 100$.

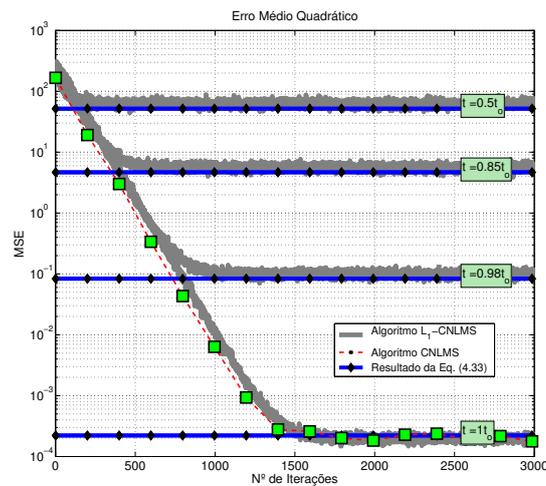
Figura 6.4: Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 0,5$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.



(a) $\mu_0 = 1,0$; $N = 10$.

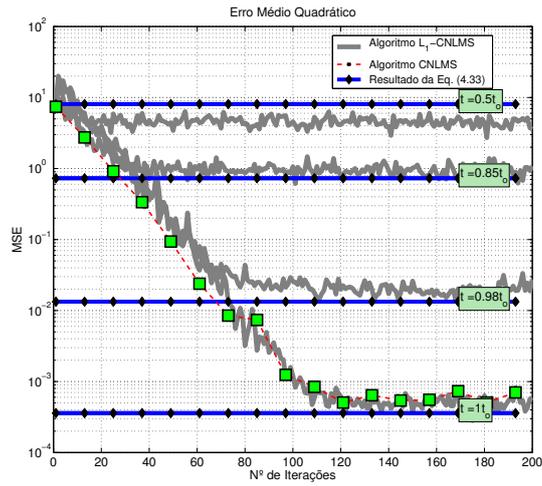


(b) $\mu_0 = 1,0$; $N = 50$.

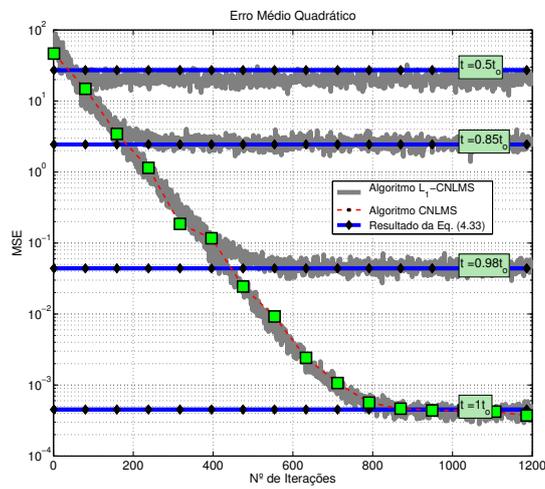


(c) $\mu_0 = 1,0$; $N = 100$.

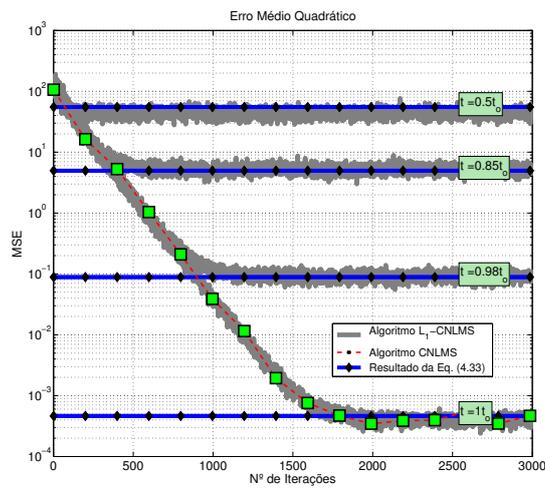
Figura 6.5: Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 1,0$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.



(a) $\mu_0 = 1,5$; $N = 10$.



(b) $\mu_0 = 1,5$; $N = 50$.



(c) $\mu_0 = 1,5$; $N = 100$.

Figura 6.6: Resultados comparativos entre o valor teórico de $J(\infty)$, calculado para $\mu_0 = 1,5$ por meio da Eq. (4.33), e os valores simulados para as curvas de aprendizagem obtidos pelo algoritmo ℓ_1 -CNLMS.

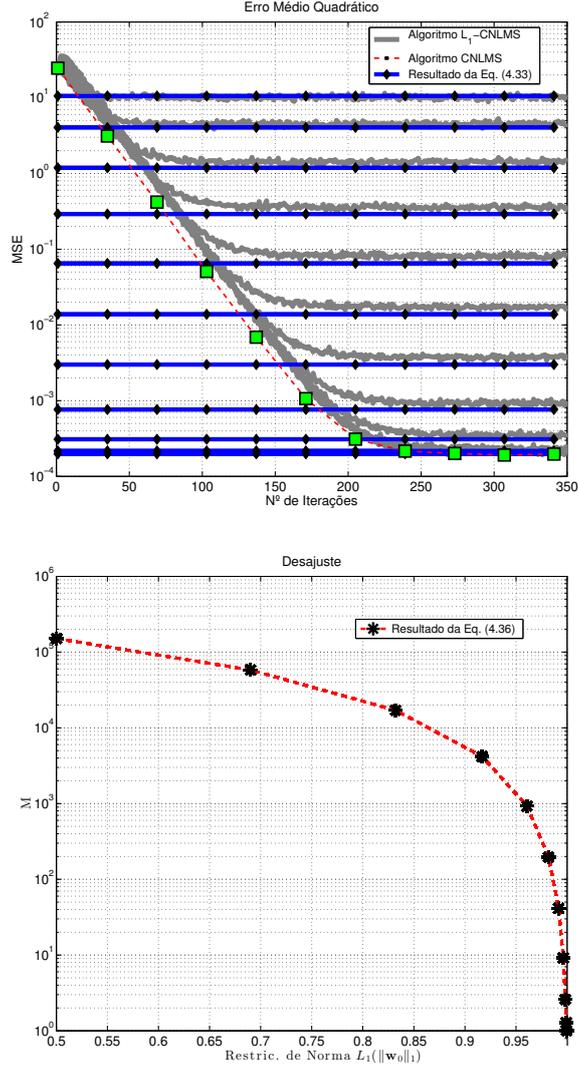


Figura 6.7: $J(\infty)$ e Desajuste calculados por meio das formulações aproximadas proposta pelas Eqs. (4.33) e (4.36), respectivamente, para uma planta complexa com 20 coeficientes, passo de convergência fixo $\mu_0 = 1$, e ruído de medição com variância $\sigma_N^2 = 10^{-4}$. Nesta simulação foram executadas 400 realizações independentes.

Na Fig. 6.7, observa-se uma boa correspondência entre as aproximação propostas, $J(\infty)$ e Desajuste, e a curva de aprendizagem obtida diretamente do algoritmo ℓ_1 -CNLMS.

As aproximações empregadas nas simulações apresentadas nesta Seção, se tornam mais precisas à medida que a ordem do filtro aumenta, conforme os resultados apresentados pelas Figuras 6.3-6.5.

6.4 Conclusão do Capítulo

Conforme pode ser observado por meio dos resultados apresentados neste Capítulo, o algoritmo ℓ_1 -WCNLMS demonstrou um rendimento superior ao do algoritmo ℓ_1 -

CNLMS à medida que aumenta a esparsidade ou compressibilidade dos filtros sob identificação. Este fato já havia sido observado de forma indireta nos resultados obtidos no Capítulo 5, lá traduzidos na forma de nível de esparsidade do arranjo final.

Quando se considera sistemas compressíveis, o rendimento do algoritmo ℓ_1 -WCNLMS cai um pouco em relação aos resultados obtidos para sistemas esparsos. No entanto, essa queda é mais acentuada somente para sistemas, cujo percentual de coeficientes compressíveis está dentro de uma faixa próxima ou inferior a 50%. Para essa faixa, o rendimento dos algoritmos ℓ_1 -WCNLMS e ℓ_1 -CNLMS se equivalem. O mesmo acontece para valores pequenos de β (ver Fig. 6.2a).

A última seção apresenta os resultados aproximados obtidos por meio das Eqs. (4.33) e (4.36), que evidenciam a influencia do “encolhimento” de norma ℓ_1 sobre o valores das figuras de mérito $J(\infty)$, $J(\infty)_{\text{exc}}$ e Desajuste. Cabe destacar ainda, que essas aproximações tornam-se mais precisas à medida que a ordem do filtro aumenta.

Capítulo 7

Conclusões

Este trabalho, cuja meta inicialmente pretendida era a síntese de arranjos de sensores esparsos por meio de filtragem adaptativa com restrições, apresentou como principal contribuição o desenvolvimento de algoritmos que combinam a simplicidade dos algoritmos derivados do LMS como a possibilidade de se ter, ao mesmo tempo, filtragem adaptativa com restrições lineares e regularização de norma ℓ_1 . Três versões de algoritmos foram apresentadas: o ℓ_1 -CLMS, que deu origem aos demais e por isso foi pouco explorado, servindo apenas de alicerce teórico para o desenvolvimento dos seus derivados, denominados ℓ_1 -CNLMS e ℓ_1 -WCNLMS.

A versão ℓ_1 -CNLMS é normalizada e portanto difere do algoritmo ℓ_1 -CLMS no que se refere ao valor do passo de convergência, que deixa de ser fixo para ter um valor instantâneo que é função dos erro instantâneo, do erro de norma ℓ_1 e da energia do sinal de entrada. Esta versão apresenta bons resultados quando o problema abordado possui algum grau de compressibilidade, como foi o caso de algumas configurações de sensores, cujas normas das soluções LCMV é maior que a unidade.

A versão ℓ_1 -WCNLMS, que foi desenvolvida a partir do algoritmo ℓ_1 -CNLMS, apresenta portanto, características de ponderação e normalização. Os resultados produzidos por esse algoritmo são superiores quando se trata da identificação de sistemas esparsos e da capacidade de gerar esparsidade. O seu rendimento, em termos de velocidade de convergência, cai à medida que a esparsidade da planta diminui. Essa degradação de rendimento, no entanto, não significa diminuição da sua capacidade de induzir esparsidade em sistemas compressíveis, como foi possível constatar por meio das sínteses de alguns arranjos esparsos apresentadas no Capítulo 5.

Baseando-se nos resultados obtidos nos Capítulos 5 e 6, observa-se que os algoritmos propostos permitem atingir um nível de esparsidade atrativo, o que os torna adequados à solução de problemas práticos onde as fontes de energia são críticas, como no caso dos sistemas de comunicações móveis satelitais, sistemas de táticos de guerra-eletrônica, dentre outros.

Especificamente, se tratando de Filtragem Adaptativa com Solução Esparsa para Arranjos de Sensores, quando os sensores são inúmeras antenas alimentadas por um número igual de *Low Noise Blocks* (LNB), chegar-se à esparsidade significa obter um menor número de LNB ativos e, portanto; economia de energia, maior autonomia etc.

Destarte, a solução de, em um mesmo algoritmo se ter restrições lineares e de norma ℓ_1 , proporciona uma atraente opção para dotar os arranjos adaptativos de sensores, além das vantagens já citadas, de maior flexibilidade.

Como sugestão para continuidade dos assuntos desenvolvidos por este trabalho, pode-se buscar o desenvolvimento de esquemas de adaptação para os valores de norma ℓ_1 e β , pois desta forma, em ambientes não estacionários, onde os níveis de esparsidade e compressibilidade podem variar com maior frequência, um esquema adaptativo para atualização desses parâmetros poderá ser trazer ganhos ao rendimentos dos algoritmos apresentados nesta tese.

Referências Bibliográficas

- [1] VAN VEEN, B., BUCKLEY, K. “Beamforming: a versatile approach to spatial filtering”, *ASSP Magazine, IEEE*, v. 5, n. 2, pp. 4–24, April 1988. ISSN: 0740-7467. doi: 10.1109/53.665.
- [2] DINIZ, P. S. R. *Adaptive Filtering: Algorithms and Practical Implementation*. 3rd ed. New York, USA, Springer, October 2010. ISBN: 1441940626.
- [3] HAYKIN, S. *Adaptive Filter Theory (4th Edition)*. New Jersey, USA, Prentice Hall, September 2001. ISBN: 0130901261.
- [4] FROST III, O. L. “An algorithm for linearly constrained adaptive array processing”, *Proceedings of the IEEE*, v. 60, n. 8, pp. 926–935, August 1972. ISSN: 0018-9219. doi: 10.1109/PROC.1972.8817.
- [5] APOLINÁRIO JR., J. A., WERNER, S., DINIZ, P. S. R., et al. “Constrained normalized adaptive filtering for CDMA mobile communications”. In: *Proceedings of the 9th European Signal Processing Conference (EUSIPCO 1998)*, Rhodes, Greece, September 1998.
- [6] DE CAMPOS, M. L. R., APOLINÁRIO, JR., J. A. “The Constrained Affine Projection Algorithm — Development and Convergence Issues”. In: *First Balkan Conference on Signal Processing, Communications, Circuits and Systems*, pp. 1–4, 2000.
- [7] JONES, D. “A normalized constant-modulus algorithm”. In: *Signals, Systems and Computers, 1995. 1995 Conference Record of the Twenty-Ninth Asilomar Conference on*, v. 1, pp. 694–697 vol.1, Oct 1995. doi: 10.1109/ACSSC.1995.540639.
- [8] DE LAMARE, R., DINIZ, P. “Blind Adaptive Interference Suppression Based on Set-Membership Constrained Constant-Modulus Algorithms With Dynamic Bounds”, *Signal Processing, IEEE Transactions on*, v. 61, n. 5, pp. 1288–1301, March 2013. ISSN: 1053-587X. doi: 10.1109/TSP.2012.2229995.

- [9] AGEE, B. “The least-squares CMA: A new technique for rapid correction of constant modulus signals”. In: *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86.*, v. 11, pp. 953–956, Apr 1986. doi: 10.1109/ICASSP.1986.1168852.
- [10] VAN TREES, H. L. *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. 1st ed. New York, USA, Wiley-Interscience, March 2002. ISBN: 0471093904.
- [11] FRATER, M., RYAN, M. *Electronic Warfare for the Digitized Battlefield*. Norwood, MA, USA, Artech House, Inc., 2001. ISBN: 1580532713.
- [12] KOPSINIS, Y., SLAVAKIS, K., THEODORIDIS, S. “Online Sparse System Identification and Signal Reconstruction Using Projections Onto Weighted ℓ_1 Balls”, *IEEE Transactions on Signal Processing*, v. 59, n. 3, pp. 936–952, March 2011. ISSN: 1053-587X. doi: 10.1109/TSP.2010.2090874.
- [13] ELAD, M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. 1st ed. New York, USA, Springer Publishing Company, Incorporated, 2010. ISBN: 144197010X, 9781441970107.
- [14] TIBSHIRANI, R. “Regression Shrinkage and Selection via the LASSO”, *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 58, n. 1, pp. 267–288, 1996. ISSN: 00359246. doi: 10.2307/2346178.
- [15] HOERL, A. E., KENNARD, R. W. “Ridge Regression: Biased Estimation for Nonorthogonal Problems”, *Technometrics*, v. 12, pp. 55–67, 1970.
- [16] CHEN, Y., GU, Y., HERO III, A. O. “Sparse LMS for system identification”. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, Ann Arbor, USA, April 2009.
- [17] DUTTWEILER, D. “Proportionate normalized least-mean-squares adaptation in echo cancelers”, *IEEE Transactions on Speech and Audio Processing*, v. 8, n. 5, pp. 508–518, September 2000. ISSN: 1063-6676. doi: 10.1109/89.861368.
- [18] PALEOLOGU, C., BENESTY, J., CIOCHINA, S. “An improved proportionate NLMS algorithm based on the l_0 norm”. In: *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 309–312, March 2010. doi: 10.1109/ICASSP.2010.5495903.

- [19] LEBRET, H., BOYD, S. “Antenna array pattern synthesis via convex optimization”, *IEEE Transactions on Signal Processing*, v. 45, n. 3, pp. 526–532, March 1997.
- [20] NETO, F., NASCIMENTO, V., ZAKHAROV, Y., et al. “Adaptive re-weighting homotopy for sparse beamforming”. In: *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, pp. 1287–1291, September 2014.
- [21] DE ANDRADE JR., J. F., DE CAMPOS, M. L. R., APOLINÁRIO JR., J. A. “An L_1 -Norm Linearly Constrained LMS Algorithm Applied to Adaptive Beamforming”. In: *Proceedings of the 7th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2012)*, Hoboken, USA, June 2012.
- [22] DE ANDRADE JR., J. F., DE CAMPOS, M. L. R., APOLINÁRIO JR., J. A. “An L_1 -constrained normalized LMS algorithm and its application to thinned adaptive antenna arrays.” In: *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013)*, Vancouver, Canada, May 2013.
- [23] CASTOLDI, F. T., DE CAMPOS, M. L. R. “Minimum-disturbance description for the development of adaptation algorithms and a new leakage least squares algorithm”. In: *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009)*, Washington, USA, April 2009.
- [24] CASTOLDI, F. T., CAMPOS, M. L. R. “Application of a Minimum-Disturbance Description to Constrained Adaptive Filters”, *IEEE Signal Processing Letters*, v. 20, n. 12, pp. 1215–1218, 2013.
- [25] SMITH, J. O. *Mathematics of the Discrete Fourier Transform (DFT)*. <http://www.w3k.org/books/>, W3K Publishing, 2007.
- [26] CANDÈS, E. J., WAKIN, M. B., BOYD, S. P. “Enhancing Sparsity by Reweighted L_1 Minimization”, *J. Fourier Anal. Appl.*, v. 14, pp. 877–905, 2008.
- [27] RUPP, M., BUTTERWECK, H.-J. “Overcoming the independence assumption in LMS filtering”. In: *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, v. 1, pp. 607–611 Vol.1, November 2003. doi: 10.1109/ACSSC.2003.1291983.
- [28] PAPOULIS, A. *Probability, random variables, and stochastic processes*. 3rd ed. New York, McGraw-Hill, 1991.

- [29] YANG, X. “A Matrix Trace Inequality”, *Journal of Mathematical Analysis and Applications*, v. 250, n. 1, pp. 372–374, October 2000.
- [30] REED, I. “On a moment theorem for complex Gaussian processes”, *Information Theory, IRE Transactions on*, v. 8, n. 3, pp. 194–195, April 1962. ISSN: 0096-1000. doi: 10.1109/TIT.1962.1057719.
- [31] WIDROW, B., WALACH, E. *Adaptive Inverse Control - A Signal Processing Approach*. Reissue edition ed. Hoboken, NJ, Wiley, 2008.
- [32] LIANG, G., GONG, W. B., LIU, H. J., et al. “Development of 61-channel digital beamforming (DBF) transmitter array for mobile satellite communication”, *Progress In Electromagnetics Research (PIER)*, v. 97, pp. 177–195, 2009.
- [33] PROAKIS, J. *Digital Communications*. McGraw-Hill series in electrical and computer engineering : communications and signal processing. New York, USA, McGraw-Hill, 2001. ISBN: 9780071181839.
- [34] VIGANO, M., TOSO, G., SELLERI, S., et al. “GA optimized thinned hexagonal arrays for satellite applications”. In: *Proceedings of 2007 IEEE Antennas and Propagation Society International Symposium*, Honolulu, USA, June 2007.
- [35] LIMA, M. V. S. *Energy-Efficient Adaptive Filters*. Doutorado, UFRJ/COPPE/Programa de Engenharia Elétrica, 2013.

Apêndice A

Minimum-disturbance Approach

Na referência [23], é proposta uma abordagem diferente para dedução de expressões de algoritmos adaptativos denominada *Minimum-disturbance Approach* (MDA). O ponto chave desta abordagem é expressar as funções custo determinísticas por meio de funções quadráticas do distúrbio mínimo dos coeficientes sujeitas (ou não) à restrições. Com o propósito de deduzir o algoritmo L_1 -CNLMS usando MDA, três restrições serão consideradas:

$$\min_{\mathbf{w}} \|\mathbf{w} - \mathbf{w}_k\|_2^2 \quad \text{s.t.} \quad \begin{cases} d_k = \mathbf{x}_k^H \mathbf{w}, \\ \mathbf{C}^H \mathbf{w} = \mathbf{z}, \\ \|\mathbf{w}\|_1 = t. \end{cases} \quad (\text{A.1})$$

Fazendo uso da técnica de multiplicadores de Lagrange para funções custo complexas, pode-se escrever a função a ser minimizada como:

$$\begin{aligned} \xi(\mathbf{w}) = & \|\mathbf{w} - \mathbf{w}_k\|_2^2 + \lambda_1^* [\mathbf{x}_k^H \mathbf{w} - d_k] \\ & + \boldsymbol{\lambda}_2^T [\mathbf{C}^H \mathbf{w} - \mathbf{z}] + \lambda_3 [\|\mathbf{w}\|_1 - t]. \end{aligned} \quad (\text{A.2})$$

Resolver o problema de minimização estabelecido pela Eq. (A.1), é equivalente a resolver $\frac{\partial \xi(\mathbf{w})}{\partial \mathbf{w}} = 0$:

$$\begin{aligned} \frac{\partial \xi(\mathbf{w})}{\partial \mathbf{w}} = & 2(\mathbf{w} - \mathbf{w}_k) + \lambda_1 \mathbf{x}_k + \mathbf{C}^* \boldsymbol{\lambda}_2 \\ & + \lambda_3 \text{sign}[\mathbf{w}] = 0, \end{aligned} \quad (\text{A.3})$$

$$\mathbf{w} = \mathbf{w}_k - \frac{1}{2} \left\{ \lambda_1 \mathbf{x}_k + \mathbf{C}^* \boldsymbol{\lambda}_2 + \lambda_3 \text{sign}[\mathbf{w}] \right\}. \quad (\text{A.4})$$

Para eliminar os multiplicadores de Lagrange λ_1 , $\boldsymbol{\lambda}_2$ and λ_3 , se faz necessária a aplicação das restrições à Eq. (A.4).

Multiplicado esta equação pela matriz \mathbf{C}^H , chega-se

$$-\frac{1}{2}\mathbf{C}^*\boldsymbol{\lambda}_2 = \frac{\lambda_1}{2}[\mathbf{I} - \mathbf{P}]\mathbf{x}_k + \frac{\lambda_3}{2}[\mathbf{I} - \mathbf{P}]\mathbf{s}, \quad (\text{A.5})$$

onde $\mathbf{s} = \text{sign}[\mathbf{w}]$.

Substituindo este resultado na própria Eq. (A.4), é possível eliminar o multiplicado desconhecido $\boldsymbol{\lambda}_2$:

$$\mathbf{w} = \mathbf{w}_k - \frac{\mathbf{P}}{2} \left\{ \lambda_1 \mathbf{x}_k + \lambda_3 \mathbf{s} \right\}. \quad (\text{A.6})$$

No intuito de isolar o multiplicador λ_1 , pode-se multiplicar a Eq. (A.6) pelo sinal de entrada \mathbf{x}_k^H , obtendo:

$$-\frac{1}{2}\lambda_1 = \frac{1}{\tau_k} \left\{ e_k^* + \frac{\lambda_3}{2} \mathbf{x}_k^H \mathbf{P} \mathbf{s} \right\}, \quad (\text{A.7})$$

com $e_k = d_k - \mathbf{w}^H \mathbf{x}_k$ e $\tau_k = \mathbf{x}_k^H \mathbf{P} \mathbf{x}_k$.

Substituindo-se o valor de λ_1 na Eq. (A.4), chega-se a:

$$\mathbf{w} = \mathbf{w}_k + \frac{1}{\tau_k} \left\{ \mathbf{P} \mathbf{x}_k e_k^* + \frac{\lambda_3}{2} \mathbf{G}_k \mathbf{s} \right\}, \quad (\text{A.8})$$

onde $\mathbf{G}_k = \mathbf{P}[\mathbf{x}_k \mathbf{x}_k^H - \tau_k \mathbf{I}] \mathbf{P}$.

Da mesma forma que foi feito na seção 3.2, pode-se considerar que próximo da convergência os vetores \mathbf{w}_k e \mathbf{w} pertencerão ao mesmo hiperquadrante. Propõe-se, então, como válida a aproximação $\mathbf{s}_k^H \mathbf{w} \approx t$, assim $\mathbf{s}_k \approx \mathbf{s} = \text{sign}[\mathbf{w}]$. Usando-se esta aproximação e definindo $t_k = \mathbf{s}_k^H \mathbf{w}_k$, pode-se eliminar os vetores \mathbf{w} e \mathbf{w}_k na Eq. (A.8) multiplicando-a por \mathbf{s}_k^H .

Para isolar λ_3 da Eq. (A.8), pode-se multiplicar esta equação por \mathbf{s}_k^H e empregar as definições e aproximações supracitadas:

$$t = t_k + \frac{1}{\tau_k} \left\{ \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k e_k^* + \frac{1}{2} D_k \lambda_3 \right\}, \quad (\text{A.9})$$

$$\lambda_3 = \frac{2\tau_k}{D_k} \left\{ e_{L_1}(k) - \frac{1}{\tau_k} \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k e_k^* \right\}, \quad (\text{A.10})$$

com $e_{L_1}(k) = t - t_k$ e $D_k = \mathbf{s}_k^H \mathbf{G}_k \mathbf{s}_k$.

Substituindo na Eq. (A.8) o valor de λ_3 calculado por meio da Eq. (A.10) e fazendo uso da escolha conveniente $\mathbf{w} = \mathbf{w}_{k+1}$, chega-se a expressão para o algoritmo

L_1 -CNLMS:

$$\mathbf{w}_{k+1} = \mathbf{P} \left[\mathbf{w}_k + \frac{\mu_0}{\tau_k} \bar{\mathbf{G}}_k \mathbf{x}_k e_k^* \right] + \mathbf{f} + \mathbf{f}_{L_1}^{MD}(k), \quad (\text{A.11})$$

com

$$\begin{cases} \bar{\mathbf{G}}_k = \left[\mathbf{I} - \left(\frac{\mathbf{G}_k \mathbf{s}_k}{D_k} \right) \mathbf{s}_k^H \right] \mathbf{P}, \\ \mathbf{f}_{L_1}^{MD}(k) = \left(\frac{\mathbf{G}_k \mathbf{s}_k}{D_k} \right) e_{L_1}(k). \end{cases}$$

O fator μ_0 é um fator de convergência fixo introduzido na expressão para propiciar o controle do desajuste. Se o valor deste fator é igual a unidade, o erro quadrático instantâneo *a posteriori* é nulo de acordo com a primeira restrição imposta na Eq. (A.1).

Agora, se faz necessário provar a identidade entre as Eqs. (3.23) e (A.11) e, desta forma, provar que ambas correspondem ao mesmo algoritmo.

Após simples manipulações algébricas, a Eq. (3.23) pode ser reescrita como:

$$\begin{aligned} \mathbf{w}_{k+1} = & \mathbf{P} \left\{ \mathbf{w}_k + \frac{\mu_0}{\zeta_k} \bar{\mathbf{P}}_k \mathbf{x}_k e_k^* \right\} + \mathbf{f} \\ & + \left\{ \mathbf{P} \mathbf{s}_k - \frac{1}{\zeta_k} \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \bar{\mathbf{P}}_k \mathbf{x}_k \right\} \frac{e_{L_1}(k)}{Q_k}, \end{aligned} \quad (\text{A.12})$$

com

$$\begin{cases} Q_k = \mathbf{s}_k^H \mathbf{P} \mathbf{s}_k = \|\mathbf{P} \mathbf{s}_k\|^2, \\ \zeta_k = \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k. \end{cases}$$

Por sua vez, a expressão obtida via MDA, Eq. (A.11), pode ser reescrita como:

$$\mathbf{w}_{k+1} = \mathbf{P} \left[\mathbf{w}_k + \frac{\mu_0}{\tau_k} \bar{\mathbf{G}}_k \mathbf{x}_k e_k^* \right] + \mathbf{f} + \left(\frac{\mathbf{G}_k}{D_k} \right) \mathbf{s}_k e_{L_1}(k), \quad (\text{A.13})$$

com

$$\begin{cases} \mathbf{G}_k = \mathbf{P} [\mathbf{x}_k \mathbf{x}_k^H - \tau_k \mathbf{I}] \mathbf{P}, \\ D_k = \mathbf{s}_k^H \mathbf{G}_k \mathbf{s}_k = \|\mathbf{G}_k \mathbf{s}_k\|^2, \\ \bar{\mathbf{G}}_k = \left[\mathbf{I} - \frac{\mathbf{G}_k \mathbf{s}_k \mathbf{s}_k^H}{D_k} \right] \mathbf{P}. \end{cases}$$

Portanto, para que as Eqs. (3.23) e (A.11) sejam iguais, devemos provar que:

$$\begin{cases} \frac{1}{\zeta_k} \bar{\mathbf{P}}_k \mathbf{x}_k = \frac{1}{\tau_k} \bar{\mathbf{G}}_k \mathbf{x}_k, \\ \frac{1}{Q_k} \left\{ \mathbf{P} \mathbf{s}_k - \frac{1}{\zeta_k} \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \bar{\mathbf{P}}_k \mathbf{x}_k \right\} = \left(\frac{\mathbf{G}_k}{D_k} \right) \mathbf{s}_k. \end{cases} \quad (\text{A.14})$$

Primeiro Passo:

Para provar as igualdades estabelecidas pelo sistema de equações (A.14), neste primeiro momento, deve-se estabelecer uma relação entre os escalares D_k , Q_k and ζ_k :

$$\begin{aligned} D_k &= \mathbf{s}_k^H (\mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P}) \mathbf{s}_k - \tau_k Q_k, \\ \zeta_k &= \mathbf{x}_k^H \bar{\mathbf{P}}_k \mathbf{x}_k \\ &= \mathbf{x}_k^H \left[\mathbf{I} - \frac{\mathbf{P} \mathbf{s}_k \mathbf{s}_k^H}{Q_k} \right] \mathbf{P} \mathbf{x}_k \\ &= \frac{1}{Q_k} \left[\tau_k Q_k - \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \right]. \end{aligned}$$

Como $\mathbf{s}_k^H \mathbf{P} \mathbf{x}_k$ e $\mathbf{s}_k^H \mathbf{P} \mathbf{x}_k$ pertencem ao corpo dos números reais, o produto $\mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k$ pode ser reescrito como:

$$\mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k = \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k.$$

Usando o resultado acima, é possível escrever ζ_k em função de D_k e Q_k :

$$\begin{aligned} \zeta_k &= \frac{1}{Q_k} \left[\tau_k Q_k - \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \right] \\ &= \frac{1}{Q_k} \left[\tau_k Q_k - \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \right] = -\frac{D_k}{Q_k} \\ \zeta_k &= -\left[\frac{D_k}{Q_k} \right]. \end{aligned} \tag{A.15}$$

Segundo Passo:

Como segundo passo, será investigada a primeira identidade do sistema de equações (A.14):

$$\begin{aligned} \frac{1}{\zeta_k} \bar{\mathbf{P}}_k \mathbf{x}_k &= \frac{Q_k \left[\left(\frac{Q_k \mathbf{I} - \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H}{Q_k} \right) \mathbf{P} \mathbf{x}_k \right]}{\left[\tau_k Q_k - \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \right]} \\ &= \frac{1}{D_k} \left[\mathbf{P} \mathbf{s}_k \mathbf{s}_k^H - Q_k \mathbf{I} \right] \mathbf{P} \mathbf{x}_k. \\ \frac{1}{\zeta_k} \bar{\mathbf{P}}_k \mathbf{x}_k &= \frac{1}{D_k} \left[\mathbf{P} \mathbf{s}_k \mathbf{s}_k^H - Q_k \mathbf{I} \right] \mathbf{P} \mathbf{x}_k. \end{aligned} \tag{A.16}$$

Do do segundo membro da primeira identidade do sistema de equações (A.14),

tem-se:

$$\begin{cases} \frac{1}{\tau_k} \bar{\mathbf{G}}_k \mathbf{x}_k = \frac{1}{\tau_k D_k} \left\{ D_k \mathbf{I} - (\mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} - \tau_k \mathbf{P}) \mathbf{s}_k \mathbf{s}_k^H \right\} \mathbf{P} \mathbf{x}_k \\ = \frac{1}{\tau_k D_k} \left\{ [\mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k] \mathbf{P} \mathbf{x}_k - [\mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k] \mathbf{P} \mathbf{x}_k \right. \\ \left. + \tau_k \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k - \tau_k Q_k \mathbf{P} \mathbf{x}_k \right\} \\ = \frac{1}{D_k} \left[\mathbf{P} \mathbf{s}_k \mathbf{s}_k^H - Q_k \mathbf{I} \right] \mathbf{P} \mathbf{x}_k. \end{cases}$$

O resultado obtido acima é idêntico ao segundo membro da Eq. (A.16), portanto, a primeira identidade do sistema de equações (A.14) foi verificada:

$$\boxed{\frac{1}{\tau_k} \bar{\mathbf{G}}_k \mathbf{x}_k = \frac{1}{\zeta_k} \bar{\mathbf{P}}_k \mathbf{x}_k}. \quad (\text{A.17})$$

Terceiro Passo:

Como terceiro passo, será investigada a segunda identidade do sistema de equações (A.14).

Após algumas manipulações algébricas, o segundo membro da segunda identidade pode ser escrito como:

$$\left(\frac{\mathbf{G}_k}{D_k} \right) \mathbf{s}_k = \left(\frac{1}{D_k} \right) \left\{ \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} - \tau_k \mathbf{P} \right\} \mathbf{s}_k \quad (\text{A.18})$$

Do primeiro membro da segunda identidade do sistema de equações (A.14), tem-se:

$$\begin{cases} \frac{1}{Q_k} \left\{ \mathbf{P} \mathbf{s}_k - \frac{1}{\zeta_k} \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \bar{\mathbf{P}} \mathbf{x}_k \right\} = \frac{1}{Q_k} \left\{ \mathbf{P} \mathbf{s}_k - \frac{1}{\zeta_k} \bar{\mathbf{P}} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \right\} \\ = \frac{1}{Q_k \zeta_k} \left\{ \zeta_k \mathbf{P} \mathbf{s}_k - \left[\frac{Q_k \mathbf{I} - \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H}{Q_k} \right] \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \right\} \\ = \frac{1}{D_k} \left\{ \frac{D_k}{Q_k} \mathbf{P} \mathbf{s}_k + \left[\frac{Q_k \mathbf{I} - \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H}{Q_k} \right] \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \right\} \\ = \frac{1}{D_k} \left\{ \frac{D_k}{Q_k} \mathbf{P} \mathbf{s}_k + \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k - \frac{1}{Q_k} \mathbf{P} \mathbf{s}_k \mathbf{s}_k^H \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k \right\} \\ = \frac{1}{D_k} \left\{ \frac{1}{Q_k} \left[\mathbf{s}_k^H (\mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P}) \mathbf{s}_k - \tau_k Q_k \right] \mathbf{P} \mathbf{s}_k + \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} \mathbf{s}_k - \frac{1}{Q_k} \left[\mathbf{s}_k^H (\mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P}) \mathbf{s}_k \right] \mathbf{P} \mathbf{s}_k \right\} \\ = \frac{1}{D_k} \left\{ \mathbf{P} \mathbf{x}_k \mathbf{x}_k^H \mathbf{P} - \tau_k \mathbf{P} \right\} \mathbf{s}_k. \end{cases}$$

O resultado anterior obtido é idêntico ao segundo membro da Eq. (A.18), por-

tanto, a segunda identidade do sistema de equações (A.14) foi verificada:

$$\boxed{\left(\frac{\mathbf{G}_k}{D_k}\right)\mathbf{s}_k = \frac{1}{Q_k} \left\{ \mathbf{P}\mathbf{s}_k - \frac{1}{\zeta_k} \mathbf{x}_k^H \mathbf{P}\mathbf{s}_k \bar{\mathbf{P}}\mathbf{x}_k \right\}}. \quad (\text{A.19})$$

Neste apêndice foi demonstrada a identidade entre Eqs. (3.23) e (A.11), concluindo-se, portanto, que a abordagem aqui apresentada produziu o mesmo resultado obtido na seção 3.3.