

ELEN MACEDO LOBATO

**ANÁLISE ESTATÍSTICA DO ALGORITMO ε LMS NO
DOMÍNIO TRANSFORMADO EM AMBIENTES
ESTACIONÁRIOS E NÃO-ESTACIONÁRIOS**

**FLORIANÓPOLIS
2006**

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA

ANÁLISE ESTATÍSTICA DO
ALGORITMO ϵ LMS NO DOMÍNIO
TRANSFORMADO EM AMBIENTES
ESTACIONÁRIOS E
NÃO-ESTACIONÁRIOS

Monografia submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para a aprovação da tese de Doutorado
em Engenharia Elétrica.

ELEN MACEDO LOBATO

Florianópolis, Maio de 2006.

ANÁLISE ESTATÍSTICA DO ALGORITMO ϵ LMS NO DOMÍNIO TRANSFORMADO EM AMBIENTES ESTACIONÁRIOS E NÃO-ESTACIONÁRIOS

Elen Macedo Lobato

‘Esta Tese foi julgada adequada para obtenção do Título de Doutor em Engenharia Elétrica, Área de Concentração em *Comunicações e Processamento de Sinais*, e aprovada em sua forma final pelo programa de Pós-graduação em Engenharia Elétrica da universidade Federal de Santa Catarina.’

Prof. Rui Seara, Dr.

Orientador

Prof. Alexandre Trofino Neto, Dr.

Coordenador do curso de Pós-graduação em Engenharia Elétrica

Banca Examinadora:

Prof. Rui Seara, Dr.

Presidente

Orlando José Tobias, Dr.

Co-Orientador

Prof. Jacques Szczupak, Ph.D.

Prof. João Marcos Travassos Romano, Dr.

Prof. Bartolomeu Ferreira Uchôa Filho, Ph.D.

Aos meus pais, Jacilda e Sebastião,
ao meu irmão, Lucival, e
ao meu noivo, Cláudio.

AGRADECIMENTOS

Deixo registrados meus sinceros agradecimentos:

A Deus, pelas oportunidades que me foram dadas para a conclusão deste trabalho.

À minha mãe, Jacilda, e à minha irmã, Edilene, por não me deixarem desistir mesmo diante de todas as dificuldades encontradas.

Ao meu noivo, Cláudio, pelo incentivo, carinho, apoio, dedicação e compreensão ao longo dos anos em que o presente trabalho foi desenvolvido.

À família Pires Merlin, pelo carinho com que me acolheram nesse igual período.

Ao meu amigo e colega, Sergio Penedo, pelo incentivo e disposição em discutir comigo vários tópicos de filtragem adaptativa, que não é sua área de conhecimento em Engenharia Elétrica.

Ao meu amigo, Ronaldo Zampolo, pelas longas conversas sobre Espiritismo, Belém (minha terra natal) e Processamento de Sinais.

Ao meu amigo, Gabriel Medina, pela companhia no Restaurante Universitário e nas aulas de forró.

Ao Prof. Rui Seara, pela amizade, motivação, apoio e dedicação, fundamentais para a brilhante orientação deste trabalho de tese.

Ao Prof. Orlando Tobias, por todo o apoio e valiosas sugestões ao longo das inúmeras conversas que tivemos enquanto me co-orientou no trabalho.

Aos Profs. Jacques Szczupak, João Marcos Travassos Romano e Bartolomeu Ferreira Uchôa Filho, pelos comentários e sugestões para o melhoramento do manuscrito de tese.

Ao meu amigo, Elton, cuja participação na realização desse trabalho, editando figuras, revendo artigos e manifestando sua sensata e transparente opinião em diversos momentos do curso, foi de extrema importância para a boa finalização deste documento de tese.

À CAPES pelo apoio financeiro recebido, à UFSC e ao LINSE por toda infra-estrutura concedida para a realização do trabalho.

A todos que de alguma forma contribuíram para a conclusão deste trabalho de tese.

Resumo da Tese apresentada à UFSC como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica.

ANÁLISE ESTATÍSTICA DO ALGORITMO ϵ LMS NO DOMÍNIO TRANSFORMADO EM AMBIENTES ESTACIONÁRIOS E NÃO-ESTACIONÁRIOS

Elen Macedo Lobato

Maio/2006

Orientador: Rui Seara, Dr.

Co-orientador: Orlando José Tobias, Dr.

Área de Concentração: Comunicações e Processamento de Sinais.

Palavras-chave: Algoritmo LMS no domínio transformado, filtragem adaptativa, modelagem estatística.

Número de Páginas: 111.

RESUMO: Este trabalho apresenta uma análise estatística do algoritmo LMS (*least-mean-square*) no domínio transformado (LMS-DT) tanto para ambientes estacionários quanto não-estacionários, resultando em um modelo mais preciso do que aqueles apresentados na literatura atual. A motivação para analisar tal algoritmo vem do fato de que este apresenta uma taxa de convergência mais alta para sinais correlacionados quando comparado com outros algoritmos adaptativos tendo complexidade computacional similar. Tal fato o torna bastante competitivo devido ao grande número de aplicações que consideram sinais de entrada coloridos. O algoritmo LMS-DT apresenta uma etapa de transformação ortogonal, que promove uma separação do sinal de entrada em sinais ocupando bandas de frequência distintas. As amostras intrabandas são correlacionadas, sendo esta correlação mais alta à medida que o número de bandas aumenta. Em nosso conhecimento, não existe um modelo estatístico que forneça uma solução geral e precisa levando em conta tal correlação. Assim, este trabalho propõe um modelo estatístico para o algoritmo LMS-DT considerando as correlações existentes intrabandas. A partir deste modelo, são obtidos parâmetros de projeto, tais como passos de adaptação máximo e ótimo, como também o desajuste do algoritmo. Por meio de simulações numéricas, constata-se uma boa concordância entre os resultados obtidos através do método de Monte Carlo e aqueles fornecidos pelo modelo estatístico proposto, tanto para sinais de entrada Gaussianos brancos quanto coloridos.

Abstract of Thesis presented to UFSC as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering.

STATISTICAL ANALYSIS OF THE TRANSFORM-DOMAIN ϵ LMS ALGORITHM FOR BOTH STATIONARY AND NONSTATIONARY ENVIRONMENT

Elen Macedo Lobato

May/2006

Advisor: Rui Seara, Dr.

Co-advisor: Orlando José Tobias, Dr.

Area of Concentration: Comunicações e Processamento de Sinais.

Palavras-chave: Transform-domain LMS algorithm, adaptive filtering, statistical modeling.

Number of Pages: 111.

ABSTRACT: This work presents a statistical analysis of the transform-domain least-mean-square (TDLMS) algorithm for both stationary and nonstationary environment, resulting in a more accurate model than those discussed in the current open literature. The motivation to analyze such an algorithm comes from the fact that this presents, for correlated signals, a higher convergence speed as compared with other adaptive algorithms that possess a similar computational complexity. Such a fact makes it a highly competitive alternative to applications considering colored input signals. The TDLMS algorithm has an orthogonal transformation stage, providing a separation of the input signal into different frequency bands. The intra-band samples are correlated, being the larger the number of bands, the higher is the correlation. Up to our knowledge, there is no other statistical model of this adaptive algorithm, providing a general and accurate solution, taking into account such correlations. In this way, this work proposes an accurate model allowing for these existing intra-band correlations. Project parameters are obtained from the statistical model, such as upper bound for the step size, optimum step-size value, and algorithm misadjustment. Through numerical simulations, a good agreement between the Monte Carlo method and the predictions from the proposed statistical model is verified for both white and colored Gaussian input signals.

SUMÁRIO

1. Introdução.....	1
1.1 MOTIVAÇÃO DO TRABALHO	5
1.2 CONTRIBUIÇÕES DO TRABALHO	5
1.3 TRABALHOS PUBLICADOS	7
1.4 ESTRUTURA DA MONOGRAFIA	8
2. Algoritmos Adaptativos para Filtragem Linear	9
2.1 CONCEITOS BÁSICOS DE FILTRAGEM ADAPTATIVA.....	9
2.2 ALGORITMO LMS	16
2.2.1 DERIVAÇÃO DA EQUAÇÃO DE ATUALIZAÇÃO DOS COEFICIENTES DO ALGORITMO LMS	16
2.2.2 PROPRIEDADES DO ALGORITMO LMS	17
2.3 ALGORITMO RLS	18
2.3.1 DERIVAÇÃO DA EQUAÇÃO DE ATUALIZAÇÃO DOS COEFICIENTES DO ALGORITMO RLS	18
2.3.2 PROPRIEDADES DO ALGORITMO RLS	19
2.4 ALGORITMOS NO DOMÍNIO TRANSFORMADO	21
2.4.1 ALGORITMO LMS EM BLOCOS NO DOMÍNIO TRANSFORMADO	21
2.4.2 ALGORITMO LMS NO DOMÍNIO TRANSFORMADO.....	21
2.5 COMPARAÇÕES ENTRE OS ALGORITMOS LMS, RLS E LMS-DT.....	22
2.6 CONCLUSÕES	25
3. Algoritmo ϵLMS no Domínio Transformado.....	26
3.1 DESCRIÇÃO DO ALGORITMO LMS NO DOMÍNIO TRANSFORMADO	27
3.2 EFEITOS DA TRANSFORMAÇÃO/NORMALIZAÇÃO NA SUPERFÍCIE DE DESEMPENHO	31
3.3 EFEITOS DA TRANSFORMAÇÃO E NORMALIZAÇÃO NA MATRIZ DE AUTOCORRELAÇÃO	34
3.4 ÍNDICE DE AVALIAÇÃO ρ	36
3.5 PROPRIEDADE DO PARTICIONAMENTO EM FREQUÊNCIA	38

3.6	VETOR DE COEFICIENTES ÓTIMO NO DOMÍNIO TRANSFORMADO E NORMALIZADO	40
3.7	ESTIMATIVA DE POTÊNCIA MÉDIA	42
3.8	CONCLUSÕES.....	44
4.	Análise Estatística do Algoritmo ϵLMS-DT	46
4.1	ANÁLISE ESTATÍSTICA	47
4.1.1	MOMENTO DE PRIMEIRA ORDEM DO VETOR $\mathbf{w}_T(n)$	47
4.1.2	REGIME PERMANENTE DE $\mathbf{w}_T(n)$	51
4.1.3	CURVA DE APRENDIZAGEM E MOMENTO DE SEGUNDA ORDEM DO VETOR $\mathbf{v}_T(n)$	51
4.1.4	LIMITE SUPERIOR PARA O PASSO DE ADAPTAÇÃO.....	54
4.1.5	DESAJUSTE.....	57
4.2	CONCLUSÕES.....	59
5.	Resultados de Simulações: caso estacionário	60
5.1	EXEMPLO 1.....	61
5.2	EXEMPLO 2.....	64
5.3	EXEMPLO 3.....	66
5.4	EXEMPLO 4.....	68
5.5	CONCLUSÕES.....	71
6.	Análise Estatística do LMS-DT Operando em Ambientes Não-Estacionários	73
6.1	FORMULAÇÃO DO PROBLEMA DE IDENTIFICAÇÃO PARA PLANTA VARIANTE NO TEMPO.....	74
6.2	ANÁLISE ESTATÍSTICA	75
6.2.1	CONSIDERAÇÕES DE ANÁLISE	76
6.2.2	MOMENTO DE PRIMEIRA ORDEM DE $\mathbf{v}_T(n)$	76
6.2.3	MOMENTO DE SEGUNDA ORDEM DE $\mathbf{v}_T(n)$	77
6.2.4	CURVA DE APRENDIZAGEM.....	79
6.2.5	ERRO EM EXCESSO	79

6.2.6	DESAJUSTE	81
6.2.7	PASSO DE ADAPTAÇÃO ÓTIMO	81
6.2.8	GRAU DE NÃO-ESTACIONARIDADE	83
6.3	CONCLUSÕES.....	84
7.	Resultados de Simulações: caso não-estacionário.....	85
7.1	EXEMPLO 1.....	87
7.2	EXEMPLO 2.....	89
7.3	EXEMPLO 3.....	92
7.4	EXEMPLO 4.....	93
7.5	EXEMPLO 5.....	94
7.6	CONCLUSÕES.....	96
8.	Comentários e Conclusões Finais	98
	Apêndices	102
	Referências Bibliográficas	108

LISTA DE FIGURAS

Figura 2.1- Diagrama de blocos de um filtro adaptativo transversal	11
Figura 2.2- Diagrama de blocos de um combinador linear adaptativo.....	11
Figura 2.3- Função de desempenho.....	14
Figura 2.4- Exemplo 1. Curvas de aprendizagem dos algoritmos LMS, RLS e LMS-DT..	23
Figura 2.5- Exemplo 2. Curvas de aprendizagem dos algoritmos LMS, RLS e LMS-DT..	24
Figura 3.1- Esquema de filtragem adaptativa no domínio transformado, com a etapa de normalização explícita	28
Figura 3.2- Esquema de filtragem adaptativa no domínio transformado	31
Figura 3.3- Curvas de nível das superfícies de desempenho do filtro adaptativo. (a) Domínio da seqüência. (b) Domínio transformado sem normalização. (c) Domínio transformado com normalização.....	33
Figura 3.4- Superfície da matriz de autocorrelação do sinal de entrada. (a) Domínio da seqüência. (b) Domínio transformado (DCT) sem normalização. (c) Domínio transformado (DCT) com normalização.....	35
Figura 3.5- Curvas de resposta em freqüência dos filtros associados à matriz de transformação DCT. (a) $N = 4$; (b) $N = 8$; (c) $N = 32$ e (d) $N = 64$	39
Figura 3.6- Diagrama em blocos para a determinação da solução de Wiener.....	40
Figura 3.7- Comportamento do filtro passa-baixas para diferentes valores de γ . $\gamma = 0.9$, $\gamma = 0.95$ e $\gamma = 0.99$ estão associados às curvas de $H_1(z)$, $H_2(z)$ e $H_3(z)$, respectivamente	43
Figura 4.1. Limite superior para passo de adaptação considerando-se um sinal de entrada correlacionado para diversos valores de N e M . (*) Experimentalmente obtidos. (Linhas sólidas) predição (4.29). (a) $M = 32$, (b) $M = 64$	57

Figura 5.1. Exemplo 1. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada cinza) simulação MC de [26]. (Linha contínua preta) modelo de [26]. (Linha tracejada preta) simulação MC. (Linha contínua cinza-escuro) modelo proposto. (Linha contínua cinza-claro) modelo de [28]. (b) Curvas de EQM. (Linha irregular cinza) simulação MC de [26]. (Linha contínua preta) modelo de [26]. (Linha irregular preta) simulação MC. (Linha contínua cinza-claro) modelo de [28]. (Linha contínua cinza-escuro) modelo proposto 62

Figura 5.2. Exemplo 1. Desempenho dos modelos para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha sólida preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha sólida cinza) modelo proposto para $0,1\mu_{\max}$ 63

Figura 5.3. Exemplo 2. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 413$. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto 64

Figura 5.4. Exemplo 2. Desempenho do modelo para $0,1\mu_{\max}$ e sinal de entrada correlacionado com $\chi = 413$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto. (b) Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto 65

Figura 5.5. Exemplo 3. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 220$. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto. (a) $0,3\mu_{\max}$ e (b) $0,1\mu_{\max}$ 67

Figura 5.6. Exemplo 3. Desempenho do modelo para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ com sinal de entrada correlacionado com $\chi = 220$. (a) Comportamento médio para os coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\max}$ 68

Figura 5.7. Exemplo 4. Comparações entre modelos para sinal de entrada não-correlacionado. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto. (a) $0,3\mu_{\max}$ e (b) $0,1\mu_{\max}$ 69

Figura 5.8. Exemplo 4. Desempenho do modelo para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ com sinal branco. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\max}$ 70

Figura 7.1- Exemplo 1. Curvas de EQM com uma dispersão dos autovalores igual a 81 e $\alpha = 2$ com SNR de 60 dB. (Linhas irregulares) simulações MC. (Linhas contínuas cinza-escuro) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)]. (a) $0,3\mu_{\text{ótimo}}$. (b) $0,1\mu_{\text{ótimo}}$ 88

Figura 7.2- Exemplo 1. Desempenho do modelo para $0,3\mu_{\text{ótimo}}$ e $0,1\mu_{\text{ótimo}}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\text{ótimo}}$. (b) Curvas de EQM. (Linha irregular cinza)

simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\text{ótimo}}$ 89

Figura 7.3- Exemplo 2. Curvas de EQM com uma dispersão dos autovalores igual a 81 e $\alpha = 1$ para SNR de 60 dB. (Linhas irregulares) simulação MC. (Linhas contínuas cinza-escuro) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)]. (a) $0,3\mu_{\text{ótimo}}$, (b) $0,1\mu_{\text{ótimo}}$ 90

Figura 7.4- Exemplo 2. Desempenho do modelo para $0,3\mu_{\text{ótimo}}$ e $0,1\mu_{\text{ótimo}}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\text{ótimo}}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha sólida preta) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha irregular preta) simulação MC. (Linha sólida cinza) modelo proposto para $0,1\mu_{\text{ótimo}}$ 91

Figura 7.5- Exemplo 3. Resultados dos modelos com uma dispersão dos autovalores igual a 356 e $\alpha = 1$, SNR = 60dB e $0,3\mu_{\text{ótimo}}$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto. (Linha contínua cinza) modelo de [30]. (b) Curvas de EQM. (Linhas irregulares) simulação MC. (Linhas contínuas preta) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)] 93

Figura 7.6- Exemplo 4. Curvas de EQM com uma dispersão dos autovalores igual a 132, $\alpha = 1$ e SNR = 60dB. (Linhas irregulares) simulação MC. (Linhas contínuas) modelo proposto Eq. (6.16). (a) $M = 8$, (b) $M = 16$, (c) $M = 32$ e (d) $M = 64$ 94

Figura 7.7- Exemplo 5. Resultados dos modelos com uma dispersão dos autovalores igual a 356 e $\alpha = 1$ para SNR de 60 dB e $0,3\mu_{\text{ótimo}}$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo de [30]. (Linha contínua cinza) modelo proposto. (b) Curvas de EQM. (Linhas

irregulares) simulação MC. (Linhas contínuas preta) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)]..... 95

Figura A.1. Verificação da aproximação (4.3) para $N = 64$ e $M = \{32, 64, 96, 128\}$. (○)

$\mathbf{E}_1 = E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$. (◇) $\mathbf{E}_2 = E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$ 103

Figura A.2. Verificação da aproximação (4.3) para $N = 128$ e $M = \{32, 64, 96, 128\}$.

(○) $\mathbf{E}_1 = E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$ e (◇) $\mathbf{E}_2 = E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$ 103

LISTA DE TABELAS

Tabela 5.1- Desajuste \mathcal{M} referente às simulações	71
Tabela 7.1- Desajuste \mathcal{M} referente às simulações	96

CAPÍTULO 1

INTRODUÇÃO

Os primeiros trabalhos sobre filtragem adaptativa (FA) foram realizados por volta da década de 1950. Desde então, a FA tem se tornado uma ferramenta muito útil no processamento de sinais em que não se tem o total conhecimento *a priori* de suas características estatísticas.

O problema de filtragem, sob a ótica estatística, teve seu início em 1931 com as primeiras contribuições de Wiener para a teoria de filtragem ótima. Essa contribuição consistiu na formulação do problema de predição linear, em que foi obtido um sistema de equações (chamadas equações de Wiener-Hopf [1]-[3]), as quais resultam, após sua resolução, nos coeficientes do filtro ótimo. O caso mais geral é em termos de duas funções de correlação: a função de autocorrelação do sinal de entrada e a função de correlação cruzada entre o sinal de entrada e o sinal desejado. Todavia, a implementação prática de tal abordagem é inadequada visto que nem sempre se tem o conhecimento *a priori* das características estatísticas dos sinais envolvidos no processo. Por outro lado, através de algoritmos adaptativos, podemos tornar tal solução viável do ponto de vista de implementação prática.

Dentre os principais algoritmos de adaptação conhecidos da literatura, o algoritmo LMS é um dos mais usados, devido à sua simplicidade de implementação e robustez [1]-[3]. Esse algoritmo foi proposto por Widrow e Hoff em 1959 visando sua aplicação em um esquema de reconhecimento de padrões, conhecido como elemento linear adaptativo e

referido na literatura por *adaline*. O algoritmo LMS consiste em dois processos: uma filtragem, seguida pelo cálculo do erro de estimação entre a saída do filtro adaptativo e a saída desejada; e um processo de adaptação, que envolve a obtenção dos coeficientes do filtro em função do erro de estimação obtido.

Apesar de sua simplicidade e robustez, o algoritmo LMS apresenta um importante inconveniente: possui uma convergência lenta para sinais de entrada correlacionados. Para superar tal limitação, vários algoritmos alternativos têm sido propostos na literatura [1]-[6]. Muitos desses, como é o caso do algoritmo RLS (*recursive least-squares*), apresentam um elevado esforço computacional. Dessa forma, isso motiva a busca por outras abordagens, que acelerem a taxa de convergência do algoritmo LMS, sem um aumento significativo de sua complexidade computacional. Uma classe desses algoritmos é o algoritmo LMS no domínio transformado (LMS-DT).

O algoritmo LMS-DT foi primeiramente introduzido por Narayan *et al.* [4], consistindo em um filtro adaptativo cujas entradas são pré-processadas pela transformada discreta de Fourier (*discrete Fourier transform* – DFT) e cujos passos de adaptação são ajustados em função dos níveis de energia do sinal de entrada. O pré-processamento do sinal de entrada através da DFT tem por objetivo descorrelacionar as amostras desse sinal que, quando combinado com um processo de normalização do passo de adaptação, resulta em um aumento de velocidade de convergência do algoritmo. É importante ressaltar que a escolha da DFT como descorrelacionador não é a mais adequada, por essa não ser uma transformação ortogonal ótima¹. Entretanto, como a utilização de uma transformada ótima é inviável na prática, outras transformações ortogonais sub-ótimas, independentes do sinal de entrada, têm sido consideradas como descorrelacionadores de sinais. Alguns exemplos de transformadas sub-ótimas são: transformada discreta de cosseno (*discrete cosine transform* – DCT); transformada discreta de seno (*discrete sine transform* – DST); transformada discreta de Hartley (*discrete Hartley transform* – DHT); transformada de Walsh-Hadamard (*Walsh-Hadamard transform* – WHT) e transformada discreta *wavelet* (*discrete wavelet transform* – DWT) [2], [7]-[9]. Tais transformadas foram analisadas por Boroujeny e Gazor [10], [11] em termos de sua capacidade de ortogonalização. Foi mostrado que quanto melhor for a capacidade de ortogonalização de uma transformada,

¹ Define-se como transformação ótima aquela que descorrelaciona completamente as amostras de um sinal. Um exemplo, é a transformada de Karhunen-Loève (*Karhunen-Loève transform* - KLT), que projeta os dados da entrada sobre os autovetores de sua matriz de autocorrelação. Todavia, como a KLT requer um conhecimento *a priori* de tais autovetores, ela torna-se de pouco interesse em aplicações práticas [2].

maior será a taxa de convergência do algoritmo LMS-DT em relação ao LMS convencional. Entretanto, embora se verifique uma melhora na convergência, o algoritmo LMS no domínio transformado e o algoritmo LMS convencional apresentam o mesmo comportamento em regime permanente [12].

A taxa de convergência do algoritmo LMS-DT considerando diferentes tipos de transformadas foi analisada por Lee e Un [13], [14]. Foi mostrado que para sinais de entrada correlacionados a velocidade de convergência do algoritmo no domínio transformado sempre é superior à obtida com o LMS convencional. Entretanto, o valor mínimo de erro quadrático médio (*mean-square error* - MSE) obtido é indiferente à escolha da transformação utilizada.

Boroujeny e Gazor [15], [16] mostraram que quando são usadas transformações apresentando melhor capacidade de ortogonalização do sinal de entrada, menor será a sensibilidade de desempenho do algoritmo LMS-DT com respeito à aritmética de precisão finita considerada.

Em [17], foi mostrado que o algoritmo LMS-DT se comporta para o caso estacionário, essencialmente, como um algoritmo de mínimos quadrados.

No que diz respeito à aplicabilidade prática, Zhongwei e Tahernezehadi [18] investigaram o desempenho dos algoritmos adaptativos no domínio transformado em cancelamento de eco acústico e observaram que o algoritmo LMS-DT apresenta resultados comparativos com os algoritmos usualmente considerados para tal fim e ainda com a vantagem de apresentar uma mais baixa complexidade computacional. Medley *et al.* [19] relataram uma aplicação do algoritmo LMS-DT para a tarefa de supressão de interferências em sistemas de comunicação de banda larga, obtendo como conclusão que a melhoria na convergência obtida pela normalização de potência valida a eficácia prática do algoritmo LMS-DT. Posteriormente, em [20], foram discutidos os problemas de implementação do algoritmo LMS-DT em relação ao LMS convencional, observando-se que quanto menor o passo de adaptação melhores são os resultados obtidos.

Com respeito à análise estatística do algoritmo LMS-DT para sinais Gaussianos, o principal obstáculo na obtenção das expressões analíticas reside na determinação do valor esperado dos termos que contêm o passo de adaptação variante no tempo. Um outro algoritmo cujo passo de adaptação também varia com o tempo é o algoritmo LMS normalizado (NLMS). No caso do algoritmo NLMS, vários pesquisadores o analisaram e obtiveram modelos para os momentos de primeira e segunda ordens do vetor de

coeficientes do filtro adaptativo, considerando a dependência do passo de adaptação com o sinal de entrada [21]-[23]. No caso do algoritmo LMS-DT, um procedimento comum usado para resolver tal problema é considerar que o passo de adaptação seja invariante no tempo [24], [25]. Essa hipótese simplificativa reduz consideravelmente a complexidade de sua análise estatística. No entanto, o modelo estatístico resultante não prediz adequadamente o processo real de operação do algoritmo em questão. Por outro lado, quando é considerada a natureza variante do passo de adaptação, da mesma forma que no caso do algoritmo NLMS, o algoritmo LMS-DT apresenta uma complicada análise matemática para derivar os momentos de primeira e segunda ordens do vetor de coeficientes do filtro adaptativo. Uma prática comum nesse caso é adotar hipóteses simplificativas visando tornar mais simples o tratamento matemático do cálculo dos valores esperados que contêm o passo de adaptação variante no tempo. Em [26]-[31], uma estratégia considerada é assumir que, além de existir uma independência estatística inter-bandas (o que é esperado), existe também uma independência intrabanda. Entretanto, no caso do algoritmo LMS-DT, essa última conjectura não pode ser considerada válida, visto que o processo de filtragem associado a cada banda inerentemente correlaciona a sequência processada. Tal correlação torna-se maior à medida que o número de bandas aumenta. Dessa forma, uma análise que considere a independência estatística intrabanda não representa com precisão a realidade estatística do processo. Assim, o caminho para uma solução mais precisa seria considerar as correlações existentes internamente a cada uma das bandas. Em [21], uma análise estatística apresentada considera essa abordagem; no entanto, o problema esbarra na solução de uma integral de alta ordem, conhecida na literatura como integral hiperelíptica ou integral Abeliana. Soluções fechadas para essa integral existem para ordens menores ou iguais a quatro e para ordens maiores do que quatro apenas para casos especiais. Sendo assim, desde então, tal caminho de solução não tem sido buscado. Em [33], uma outra alternativa foi considerada para suplantiar as dificuldades de uma análise estatística do algoritmo LMS-DT; entretanto, o resultado da análise apresentada serve apenas para regime permanente. Recentemente em [34], para o caso do algoritmo NLMS, o problema da dificuldade de solução de uma integral hiperelíptica de alta ordem, oriunda da análise de tal algoritmo, foi superado através do uso de um método aproximado da supramencionada integral e que produz excelentes resultados.

Neste trabalho de tese, propõe-se uma análise estatística do algoritmo LMS-DT

para sinais de entrada com distribuição Gaussiana considerando a dependência estatística do sinal intrabanda, a qual recai na solução de integrais hiperelípticas de alta ordem. O caminho aqui seguido para resolver tais integrais é similar ao considerado em [34]. Sendo assim, tem-se uma análise estatística mais precisa do que aquelas até então apresentadas na literatura para o algoritmo em questão.

1.1- MOTIVAÇÃO DO TRABALHO

Considerando a grande aplicabilidade do algoritmo LMS-DT, uma modelagem estatística precisa e confiável de seu comportamento é plenamente justificável. Desde seu surgimento, muito esforço de pesquisa tem sido feito na direção de se obter um modelo que reúna as qualidades acima mencionadas. Para tal, é desejável que o modelo estatístico do algoritmo em questão seja o mais completo possível, devendo incluir ambientes estacionários, não-estacionários, como também considerar a dependência estatística do sinal intrabanda proveniente do processo de transformação ortogonal.

A importância da modelagem estatística está no fato de se poder contar com ferramentas que permitam prever o comportamento do algoritmo para diversas condições de operação. Por exemplo, em algumas aplicações, caso não houvesse modelos estatísticos, poderíamos ser forçados a despendar em simulações considerável tempo que poderiam ser horas, dias ou até meses, dependendo da aplicação. Um adequado modelo estatístico poderia nos fornecer uma gama importante de parâmetros do algoritmo (passo de adaptação ótimo e máximo, desajuste, taxa de convergência, dentre outros) tanto para projetos como para avaliação de operação, exigindo um tempo praticamente insignificante comparado com aquele requerido na ausência de um modelo.

1.2- CONTRIBUIÇÕES DO TRABALHO

Como apresentado anteriormente, os trabalhos publicados deixam uma importante lacuna no que diz respeito à obtenção de um modelo estatístico do algoritmo LMS-DT

mais preciso, considerando tanto ambientes estacionários como não-estacionários. Dessa forma, este trabalho contribui com o estado-da-arte nos seguintes tópicos:

- Análise estatística do algoritmo LMS-DT para ambientes estacionários;

São apresentadas expressões mais precisas para os momentos de primeira e segunda ordens do vetor de coeficientes do filtro LMS-DT, bem como para o erro quadrático médio (EQM). Essa análise estatística foi desenvolvida considerando-se um de sinal de entrada com distribuição Gaussiana, adaptação lenta e uma janela de observação relativamente grande, parâmetro esse usado na etapa de normalização do passo de adaptação.

- Análise de estabilidade e desempenho do algoritmo LMS-DT;

A partir da análise estatística, pode-se obter o limite superior para o passo de adaptação que determine a estabilidade do algoritmo. Também, pode-se obter uma expressão para o desajuste do referido algoritmo. Isso permite avaliar de forma quantitativa o algoritmo em questão.

- Análise estatística do algoritmo LMS-DT para ambientes não-estacionários;

As expressões obtidas para o caso estacionário também são derivadas para o caso não-estacionário. Para realizar essa análise são adotadas as mesmas considerações que no caso estacionário.

- Análise de estabilidade e desempenho do algoritmo LMS-DT operando em ambiente não-estacionário;

Com base nos momentos de primeira e segunda ordens, são derivadas as expressões para o desajuste e passo de adaptação ótimo.

1.3- TRABALHOS PUBLICADOS

A seguir são listadas as publicações resultantes deste trabalho de tese.

1. E. M. Lobato, O. J. Tobias, R. Seara, "Análise Estatística do Algoritmo LMS no Domínio Transformado". XX Simpósio Brasileiro de Telecomunicações - SBT 2003. pp. 1-6, Rio de Janeiro, RJ, Outubro 2003.
2. E. M. Lobato, O. J. Tobias, R. Seara, "Modelo Analítico do Algoritmo ϵ LMS no Domínio Transformado para Sinais de Entrada Gaussianos". XXI Simpósio Brasileiro de Telecomunicações - SBT 2004. pp. 1-6, Belém, PA, Setembro 2004.
3. E. M. Lobato, O. J. Tobias, R. Seara, "A Stochastic Model for the Transform-Domain LMS Algorithm". 12th European Signal Processing Conference - EUSIPCO 2004. pp. 1833-1836, Viena, Áustria, Setembro 2004.
4. E. M. Lobato, O. J. Tobias, R. Seara, "Stochastic Modelling of the Transform Domain ϵ LMS Algorithm for a Time-Varying Environment". 13th European Signal Processing Conference (EUSIPCO 2005). pp. 1-4, Antalya, Turquia, Setembro 2005.
5. E. M. Lobato, O. J. Tobias, R. Seara, "Análise Estatística do Algoritmo ϵ LMS no Domínio Transformado em Ambientes Não-estacionários". XXII Simpósio Brasileiro de Telecomunicações (SBrT 2005). pp. 817-822, Campinas, SP, Setembro 2005.
6. E. M. Lobato, O. J. Tobias, R. Seara, "Stochastic Model for the NLMS Algorithm with Correlated Gaussian Data". IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006). Vol III, pp. 760-763, Toulouse, France, Maio 2006.

1.4- ESTRUTURA DA MONOGRAFIA

O trabalho de pesquisa aqui apresentado é organizado como segue.

O Capítulo 2 apresenta, de forma sintética, uma revisão dos algoritmos adaptativos mais usados em filtragem adaptativa, estabelecendo-se um paralelo com o algoritmo LMS-DT. O objetivo desse capítulo é mostrar a eficiência do LMS-DT para sinais de entrada correlacionados, justificando assim o estudo estatístico proposto.

O Capítulo 3 apresenta os conceitos básicos do algoritmo LMS-DT. Apresenta-se a equação recursiva do vetor de coeficientes do filtro LMS-DT, bem como são mostrados os efeitos da operação de ortogonalização e normalização dos dados de entrada sobre a superfície de desempenho do filtro e sobre a matriz de autocorrelação do sinal de entrada. O desempenho do filtro LMS-DT depende da capacidade de ortogonalização da transformada usada. Assim, é importante considerar uma breve discussão que avalie o desempenho das transformações ortogonais.

O Capítulo 4 traz, de forma detalhada, as considerações realizadas para a modelagem estatística do algoritmo LMS-DT. Com tais considerações, podem-se determinar o valor médio do vetor de coeficientes do filtro adaptativo e o valor do EQM para sinais de entrada Gaussianos (brancos e coloridos), bem como as expressões para o desajuste e para o passo de adaptação máximo.

No Capítulo 5, analisam-se os resultados do modelo obtido a partir das equações apresentadas no Capítulo 4 e aqueles obtidos via simulação numérica.

No Capítulo 6, são apresentadas as expressões dos momentos de primeira e segunda ordens do vetor de coeficientes do filtro LMS-DT, considerando-se um ambiente não-estacionário. É também apresentada a expressão para a determinação do desajuste e o passo de adaptação ótimo para o algoritmo, envolvendo alguns aspectos dos ambientes não-estacionários.

No Capítulo 7, a concordância do modelo apresentado no Capítulo 6 é verificada através de simulações Monte Carlo.

Por fim, no Capítulo 8, são apresentadas as conclusões deste trabalho de tese.

CAPÍTULO 2

ALGORITMOS ADAPTATIVOS PARA FILTRAGEM LINEAR

Neste capítulo são apresentados os conceitos básicos sobre filtros adaptativos e as principais características de dois consagrados algoritmos de filtragem adaptativa, o algoritmo LMS e o RLS. A finalidade de tal discussão está no fato de que o algoritmo LMS-DT para ambientes estacionários comporta-se, essencialmente, como um algoritmo de mínimos quadrados. Além disso, o LMS-DT apresenta algumas importantes características similares às do algoritmo LMS, tais como robustez, simplicidade e baixa complexidade computacional. Também será apresentada neste capítulo para fins de comparação, uma aplicação em identificação de sistemas, utilizando-se os seguintes algoritmos de adaptação: RLS, LMS convencional e LMS-DT.

2.1- CONCEITOS BÁSICOS DE FILTRAGEM ADAPTATIVA

Um sistema adaptativo é um sistema cuja saída é modificada de tal forma que seu comportamento seja alterado de acordo com algum critério desejado.

Os sistemas adaptativos costumam apresentar todas ou algumas das seguintes características [3]:

1. Adaptação automática quando operam em ambientes, ou junto com sistemas, que variam com o tempo (não-estacionários);

2. Podem ser treinados para desempenhar tarefas específicas, tais como filtragem e tomada de decisão. Ou seja, podem sintetizar sistemas responsáveis por essas tarefas através de um processo de treinamento;
3. Devido às características 1 e 2, não requerem procedimentos de síntese elaborados, usualmente necessários para sistemas não-adaptativos. Os sistemas adaptativos tendem a se autoprojetar;
4. Após terem sido treinados com um finito e, freqüentemente, pequeno número de sinais (ou padrões) de treino, podem extrapolar um modelo de comportamento para lidar com novas situações;
5. Dentro de certos limites, podem se auto-restaurar; ou seja, se adaptar “ao redor” de certos tipos de defeitos “internos”;
6. Podem representar sistemas lineares bem como não-lineares. Um exemplo de um sistema adaptativo não-linear é a rede neural;
7. Em geral, são mais complexos e difíceis de se analisar do que os sistemas não-adaptativos, mas oferecem a possibilidade de apresentar um desempenho substancialmente melhor quando as características dos sinais de entrada são desconhecidas ou variantes no tempo.

A propriedade essencial e principal de um algoritmo adaptativo é seu desempenho auto-ajustável e variante no tempo.

A estrutura mais usada na implementação de filtros adaptativos é a estrutura transversal, ilustrada pelo diagrama de blocos da Figura 2.1. Pode-se observar que o filtro adaptativo possui uma entrada única $x(n)$ e uma saída $y(n)$. A seqüência $d(n)$ é denominada sinal desejado. A saída $y(n)$ é gerada como sendo uma combinação linear das amostras da seqüência $x(n)$, conforme a seguinte equação [1]-[3]:

$$y(n) = \sum_{i=0}^{N-1} w_i(n)x(n-i), \quad (2.1)$$

onde $w_i(n)$ são os coeficientes do filtro adaptativo e $N-1$ é a ordem do referido filtro. Os coeficientes (pesos) $w_i(n)$ (variantes no tempo) são controlados pelo algoritmo adaptativo.

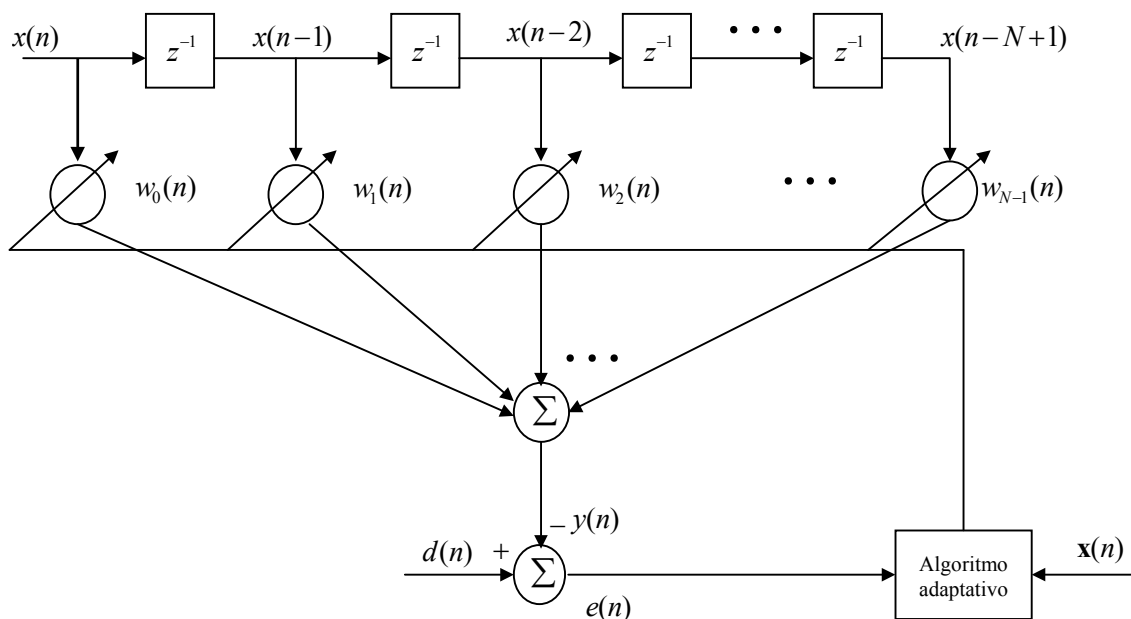


Figura 2.1- Diagrama de blocos de um filtro adaptativo transversal.

Uma outra estrutura típica de filtro adaptativo tem sua forma ilustrada pelo diagrama da Figura 2.2, chamado de combinador linear. Sua saída é uma combinação ponderada de diferentes sinais de entrada, conforme expresso na equação seguinte:

$$y(n) = \sum_{i=0}^{N-1} w_i(n)x_i(n). \quad (2.2)$$

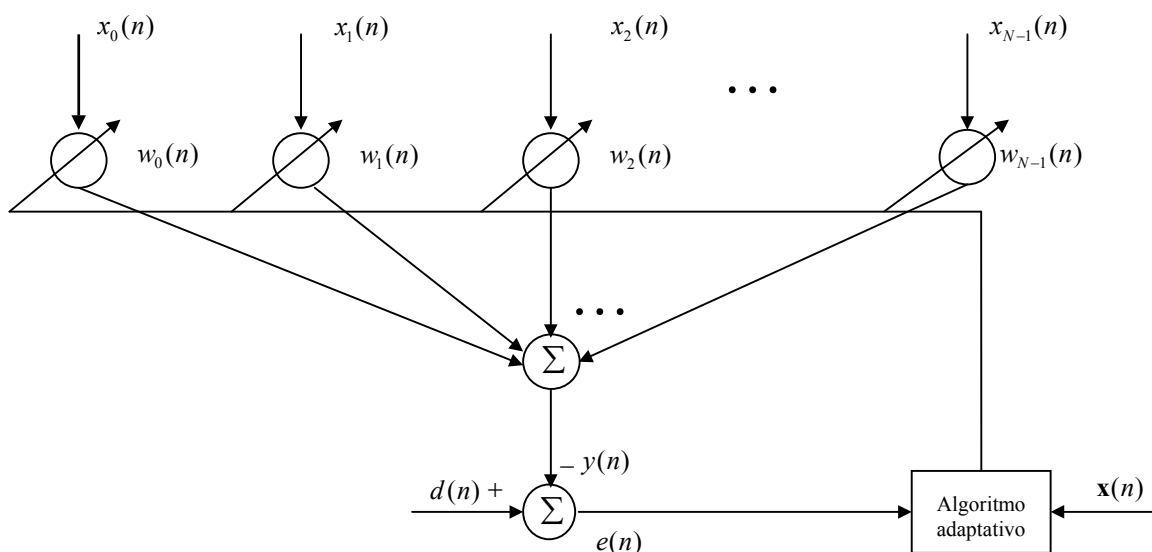


Figura 2.2- Diagrama de blocos de um combinador linear adaptativo.

A estrutura do combinador linear é mais geral do que a transversal. Sendo que uma é obtida da outra escolhendo-se a i -ésima entrada como sendo: $x_i(n) = x(n-i)$.

A estrutura das Figuras 2.1 e 2.2 é não-recursiva, ou seja, a determinação da saída do filtro não envolve qualquer mecanismo de realimentação. A Figura 2.1 é também referida como um filtro de resposta ao impulso finita (FIR), uma vez que sua resposta ao impulso é de duração finita.

Filtros de resposta ao impulso infinita (IIR) também são usados em filtragem adaptativa. Entretanto, devido às muitas dificuldades envolvidas em sua adaptação, seu emprego em aplicações práticas é ainda bastante limitado. O principal problema de tais estruturas é o controle de estabilidade, uma vez que seus pólos podem se posicionar fora do círculo de raio unitário devido ao processo de adaptação. Além disso, a função desempenho (ou seja, o erro quadrático médio como função dos coeficientes do filtro) de um filtro IIR normalmente possui mínimos locais. Isso pode resultar na convergência do filtro para algum desses mínimos e não para o mínimo global desejado da função desempenho. Por outro lado, a função representativa do erro quadrático médio para uma estrutura FIR é uma função quadrática com um único mínimo, podendo ser facilmente encontrado através de vários algoritmos para tal fim.

O objetivo do algoritmo adaptativo é determinar o valor do vetor de coeficientes do filtro, minimizando iterativamente algum critério de erro (função custo). Então, primeiramente, vamos determinar a solução ótima a partir do critério de erro quadrático médio dado por [1]-[3]

$$\begin{aligned}\xi &= E[e^2(n)] \\ &= E\{[d(n) - y(n)]^2\}.\end{aligned}\tag{2.3}$$

O sinal de saída $y(n)$ pode então ser expresso como o produto interno entre os vetores de coeficientes e de entrada. Assim, $y(n) = \mathbf{w}^T \mathbf{x}(n)$, onde o sobrescrito T denota o vetor transposto; $\mathbf{x}(n) \triangleq [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$ representa o vetor do sinal de entrada e $\mathbf{w} \triangleq [w_0 \ w_1 \ \dots \ w_{N-1}]^T$, o vetor constituído pelos coeficientes do filtro, que nesse momento está sendo considerado invariante no tempo.

O erro quadrático médio (EQM) definido em (2.3), considerando um ambiente estacionário, pode ser desenvolvido como:

$$\begin{aligned} \text{EQM} = \xi &= E\{[d(n) - y(n)]^2\} \\ &= E[d^2(n)] + \mathbf{w}^T \mathbf{R} \mathbf{w} - 2\mathbf{p}^T \mathbf{w}, \end{aligned} \quad (2.4)$$

onde \mathbf{R} é a matriz de autocorrelação do sinal de entrada,

$$\mathbf{R} \triangleq E[\mathbf{x}(n)\mathbf{x}^T(n)], \quad (2.5)$$

e \mathbf{p} é o vetor correlação cruzada entre o sinal de entrada e o sinal desejado,

$$\mathbf{p} \triangleq E[\mathbf{x}(n)d(n)]. \quad (2.6)$$

A função de desempenho ξ é uma função quadrática dos coeficientes e assume a forma de um hiperparabolóide, a qual está ilustrada na Figura 2.3 para o caso de dois coeficientes. Os cortes da superfície de desempenho $\xi = \text{constante}$ são hiperelipsóides. A orientação e a forma dessas elipses dependem dos autovalores da matriz de autocorrelação do sinal de entrada \mathbf{R} . No caso bidimensional, se os dois autovalores são muito diferentes, as elipses são estreitas e longas, enquanto que se os autovalores forem iguais as elipses se degeneram em círculos.

O vetor de coeficientes que minimiza a função desempenho ξ corresponde ao ponto inferior da superfície. É obtido matematicamente tomando-se a derivada de ξ em relação ao vetor de coeficientes, igualando-se a zero, e resolvendo-se para \mathbf{w} . A solução $\mathbf{w}_{\text{ótimo}}$ é dada por

$$\mathbf{w}_{\text{ótimo}} = \mathbf{R}^{-1} \mathbf{p}. \quad (2.7)$$

Esta equação é uma expressão matricial da equação de Wiener-Hopf. O valor de erro quadrático médio mínimo é obtido substituindo-se \mathbf{w} por $\mathbf{w}_{\text{ótimo}}$ em (2.4), resultando em

$$\xi_{\text{min}} = E[d^2(n)] + (\mathbf{R}^{-1} \mathbf{p})^T \mathbf{R} (\mathbf{R}^{-1} \mathbf{p}) - 2\mathbf{p}^T (\mathbf{R}^{-1} \mathbf{p}) = E[d^2(n)] - \mathbf{p}^T \mathbf{w}_{\text{ótimo}}. \quad (2.8)$$

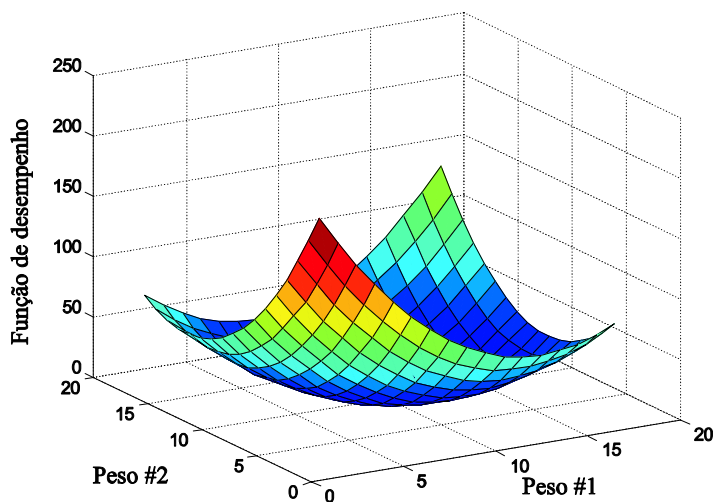


Figura 2.3- Função de desempenho

A função de desempenho ξ pode ser reescrita como

$$\xi = \xi_{\min} + (\mathbf{w} - \mathbf{w}_{\text{ótimo}})^T \mathbf{R} (\mathbf{w} - \mathbf{w}_{\text{ótimo}}). \quad (2.9)$$

Na Figura 2.3, os parâmetros utilizados para gerar a curva conforme (2.9) são:

- $\xi_{\min} = 0.001$;
- $\mathbf{w}_{\text{ótimo}} = [0.9 \quad 0.64]^T$;
- $\mathbf{R} = \begin{bmatrix} 1 & 0.3789 \\ 0.3789 & 1 \end{bmatrix}$.

Geralmente, na prática, as características estatísticas dos sinais envolvidos não são conhecidas *a priori*, o que leva à utilização de algoritmos adaptativos necessária para se obter uma solução alternativa (equivalente) à expressão (2.7). São muitos os algoritmos adaptativos encontrados na literatura. A escolha do algoritmo de adaptação deve considerar os seguintes parâmetros [1]-[3]:

1. Taxa de convergência: definida como o número de iterações necessárias para que o algoritmo, perante sinais estacionários, se aproxime da solução ótima;
2. Desajuste: é um parâmetro adimensional, definido como a diferença entre o erro quadrático médio do algoritmo, obtido em regime permanente, e o erro quadrático médio mínimo, obtido da solução de Wiener, normalizada em relação ao último erro considerado;
3. *Tracking*: capacidade do algoritmo em acompanhar as variações estatísticas do sinal de entrada em um meio não-estacionário.
4. Robustez: capacidade de o algoritmo operar satisfatoriamente com sinais de entrada mal condicionados;
5. Complexidade computacional: refere-se ao número de operações a efetuar para completar uma iteração do algoritmo. Esta é uma das características importantes que determina a viabilidade do algoritmo ser implementado para operação em situações reais.
6. Estabilidade numérica: dependendo do algoritmo de adaptação, poderá surgir instabilidade numérica quando esse for implementado em um processador de sinais. A acumulação dos erros de quantização, devido à precisão finita, poderá desviar o algoritmo da situação de operação obtida para precisão infinita. É possível que o desvio seja de uma natureza tal que os erros de quantização se acumulem sem limite, caracterizando o algoritmo como numericamente instável.

Os algoritmos adaptativos disponíveis na literatura são muitos, podendo-se dividi-los basicamente em duas classes [1]:

- ii) Algoritmos de gradiente estocástico;
- iii) Algoritmos de mínimos quadrados.

Para os algoritmos baseados no gradiente estocástico, a função custo é definida como sendo o valor esperado da diferença ao quadrado entre o sinal desejado e a saída do filtro adaptativo. Tais algoritmos têm como principal vantagem a baixa complexidade. O algoritmo mais popular dessa família é o algoritmo LMS [1].

No caso dos algoritmos de mínimos quadrados, a função custo é definida como sendo a soma dos quadrados dos erros ponderados. Tais algoritmos têm como vantagem a baixa sensibilidade à natureza do sinal de entrada e uma maior velocidade de convergência comparativamente com os algoritmos de gradiente estocástico. Apresentam como principal desvantagem uma complexidade computacional elevada. O algoritmo mais popular dessa família é o algoritmo RLS [1].

2.2- ALGORITMO LMS

O algoritmo LMS (*least-mean-square*), proposto por Widrow e Hoff em 1959, é o mais popular e o mais largamente usado algoritmo de adaptação. Nesta seção, esse algoritmo será brevemente discutido.

2.2.1- DERIVAÇÃO DA EQUAÇÃO DE ATUALIZAÇÃO DOS COEFICIENTES DO ALGORITMO LMS

O algoritmo LMS minimiza a função desempenho $\xi[\mathbf{w}(n)]$ usando o método do gradiente estocástico. Isto é, a cada iteração, os coeficientes são adaptados proporcionalmente a uma estimativa do gradiente da função de desempenho.

Seja $\nabla(n)$ o gradiente da função de desempenho e $\hat{\nabla}(n)$ sua estimativa. O gradiente é então determinado por [2]

$$\nabla(n) = \frac{d\{\xi[\mathbf{w}(n)]\}}{d\mathbf{w}(n)} = \frac{dE\{[d(n) - \mathbf{w}^T(n)\mathbf{x}(n)]^2\}}{d\mathbf{w}(n)}. \quad (2.10)$$

O gradiente estimado $\hat{\nabla}(n)$ é obtido omitindo-se a operação de valor esperado na equação anterior, daí o nome “gradiente estocástico”, resultando em

$$\begin{aligned}\hat{\nabla}(n) &= \frac{d[d(n) - \mathbf{w}^T(n)\mathbf{x}(n)]^2}{d\mathbf{w}(n)} \\ &= -2[d(n) - \mathbf{w}^T(n)\mathbf{x}(n)]\mathbf{x}(n) = -2e(n)\mathbf{x}(n).\end{aligned}\quad (2.11)$$

Ajustando-se os coeficientes proporcionalmente ao gradiente instantâneo, o algoritmo LMS percorre a superfície de desempenho em uma trajetória irregular, cujo curso médio é o caminho da descida mais íngreme. A principal motivação por trás de tal aproximação é diminuir o custo computacional do processo de determinação do vetor de coeficientes ótimo.

A partir da (2.11), tem-se a equação de atualização do vetor de coeficientes do algoritmo LMS, dada pela seguinte expressão [1]-[3]:

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\hat{\nabla}(n) \\ &= \mathbf{w}(n) + 2\mu e(n)\mathbf{x}(n),\end{aligned}\quad (2.12)$$

onde μ é denominado passo de adaptação. Tal parâmetro é uma constante que comanda a velocidade de convergência do algoritmo.

2.2.2- PROPRIEDADES DO ALGORITMO LMS

Abaixo são listadas algumas propriedades do algoritmo LMS [1]-[3]:

1. A superfície de desempenho do algoritmo LMS é um hiperparabolóide em um espaço N -dimensional, com um único ponto de mínimo (solução ótima).
2. A taxa de convergência é afetada pela razão $\lambda_{\max}/\lambda_{\min}$, onde λ_{\max} é o máximo autovalor e λ_{\min} é o mínimo autovalor da matriz autocorrelação do sinal de entrada. Para um sinal de entrada branco, tal razão vale 1, permitindo ao algoritmo LMS convergir para tal sinal mais rapidamente do que para sinais que apresentem uma razão $\lambda_{\max}/\lambda_{\min} > 1$.
3. Há um compromisso entre velocidade de convergência e erro em excesso.

4. Para valores bastante pequenos de tamanho de passo de adaptação, a convergência é muito lenta. Para valores elevados, a convergência estável não é sempre possível. Em resumo, o tamanho do passo representa um importante parâmetro no processo de adaptação LMS.

2.3- ALGORITMO RLS

O algoritmo RLS (*recursive least square*) implementa recursivamente uma solução exata dos mínimos quadrados. Foi mostrado anteriormente que a solução de Wiener para filtros adaptativos de comprimento finito é dada por $\mathbf{w}_{\text{ótimo}} = \mathbf{R}^{-1}\mathbf{p}$. Em cada iteração, o algoritmo RLS estima recursivamente \mathbf{R}^{-1} e \mathbf{p} baseado em todas as informações passadas e computa o vetor de coeficientes como $\mathbf{w}(n) = \mathbf{R}^{-1}(n)\mathbf{p}(n)$.

2.3.1- DERIVAÇÃO DA EQUAÇÃO DE ATUALIZAÇÃO DOS COEFICIENTES DO ALGORITMO RLS

A cada instante de tempo n , uma estimativa de \mathbf{R} e \mathbf{p} é dada por [1]

$$\begin{aligned}\mathbf{R}(n) &= \sum_{i=1}^n \beta^{n-i} \mathbf{x}(i)\mathbf{x}^T(i) \\ &= \beta\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n)\end{aligned}\tag{2.13}$$

e

$$\begin{aligned}\mathbf{p}(n) &= \sum_{i=1}^n \beta^{n-i} \mathbf{x}(i)d(i) \\ &= \beta\mathbf{p}(n-1) + x(n)d(n),\end{aligned}\tag{2.14}$$

onde β é uma constante, denominada fator de esquecimento, tal que $\beta \in [0,1]$. Geralmente é escolhida muito próximo da unidade, porém ligeiramente menor por razões de estabilidade. A estimativa do vetor de coeficientes ótimo $\mathbf{w}_{\text{ótimo}}$ é dada por

$$\mathbf{w}(n) = \mathbf{R}^{-1}(n)\mathbf{p}(n). \quad (2.15)$$

Aplicando-se então o Lema de Inversão de Matrizes ¹em $\mathbf{R}(n)$, tem-se:

$$\mathbf{R}^{-1}(n) = \beta^{-1}\mathbf{R}^{-1}(n-1) - \frac{\beta^{-2}\mathbf{R}^{-1}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)}{1 + \beta^{-1}\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)}. \quad (2.16)$$

Substituindo-se (2.14) e (2.16) em (2.15), após algumas manipulações, chega-se à expressão recursiva do algoritmo RLS [1]-[3]:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{\beta^{-1}\mathbf{R}^{-1}(n-1)\mathbf{x}(n)}{1 + \beta^{-1}\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)}\alpha(n), \quad (2.17)$$

onde $\alpha(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{x}(n)$.

2.3.2- PROPRIEDADES DO ALGORITMO RLS

A expressão recursiva do algoritmo RLS pode ser reescrita como segue:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)\alpha(n), \quad (2.18)$$

onde $\mu(n) = \frac{\beta^{-1}}{1 + \beta^{-1}\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)}$.

¹ $(\mathbf{\Gamma} + \alpha\mathbf{u}\mathbf{u}^T)^{-1} = \mathbf{\Gamma}^{-1} - \frac{\alpha}{1 + \alpha\mathbf{u}^T\mathbf{\Gamma}^{-1}\mathbf{u}}\mathbf{\Gamma}^{-1}\mathbf{u}\mathbf{u}^T\mathbf{\Gamma}^{-1}$

onde $\mathbf{\Gamma}$ é uma matriz não singular, \mathbf{u} é um vetor e α um escalar [2].

A partir de (2.18), pode-se avaliar a operação de decorrelação efetuada pelo algoritmo RLS nos dados de entrada: o gradiente estocástico $\mathbf{x}(n)\alpha(n)$ é pré-multiplicado por uma estimativa da inversa da matriz de autocorrelação $\mathbf{R}^{-1}(n-1)$, que decorrelaciona as amostras do sinal de entrada do filtro adaptativo. Essa decorrelação, juntamente com a expressão do passo de adaptação $\mu(n)$, reduz a sensibilidade do algoritmo aos autovalores da matriz de autocorrelação do vetor de entrada, aumentando sua taxa convergência quando comparada à do algoritmo LMS. Por outro lado, a pré-multiplicação por $\mathbf{R}^{-1}(n-1)$ pode deteriorar a estabilidade do filtro adaptativo caso a matriz $\mathbf{R}(n)$ seja mal condicionada.

Com respeito a capacidade de *tracking* para ambientes não-estacionários, visto que o vetor de coeficientes do algoritmo RLS é baseado em todas as amostras passadas do sinal de entrada, se as estatísticas desse sinal mudam com o tempo, torna-se mais difícil para o algoritmo se ajustar de acordo com tais mudanças. Portanto, o algoritmo RLS apresenta uma menor capacidade de *tracking* comparado ao algoritmo LMS, cujos coeficientes são adaptados unicamente tomando em consideração o gradiente estocástico corrente [1].

Além disso, o algoritmo RLS apresenta uma alta complexidade computacional devido ao manuseio de matrizes no processo de adaptação. A cada iteração, $O(N^2)$ operações são requeridas para cada coeficiente adaptado, enquanto que somente $O(N)$ operações são necessárias para o algoritmo LMS.

Em suma, enquanto o algoritmo RLS possui a vantagem de uma mais rápida taxa de convergência e baixa sensibilidade aos autovalores da matriz de autocorrelação do sinal de entrada, em contraste ele exibe uma alta complexidade computacional e uma baixa eficiência quanto à capacidade de *tracking* em ambientes não-estacionários quando comparado ao algoritmo LMS. Por outro lado, o algoritmo LMS é intrinsecamente lento por não decorrelacionar o sinal de entrada antes do processo de filtragem adaptativo. Uma solução para tal problema é pré-processar o sinal de entrada do filtro LMS com uma transformação que não dependa desse sinal. A decorrelação obtida será somente aproximada; porém, o custo computacional de $O(N)$ e a robustez e capacidade de *tracking* serão preservados. Esses algoritmos são conhecidos na literatura por algoritmo LMS no domínio transformado ou algoritmo LMS no domínio da frequência.

2.4- ALGORITMOS NO DOMÍNIO TRANSFORMADO

O nome “algoritmo LMS no domínio transformado” é algumas vezes ambíguo. Ele tem sido usado na literatura para designar duas diferentes categorias de algoritmos: algoritmo LMS em blocos, implementados no domínio da frequência e algoritmos LMS “não-em blocos”, cujas entradas são transformadas para o domínio da frequência antes do processo de filtragem. Nesta seção, são apresentadas brevemente as principais características das duas famílias de algoritmos [6].

2.4.1- ALGORITMO LMS EM BLOCOS NO DOMÍNIO TRANSFORMADO

Nesta categoria, é assumido que o sinal de saída, o sinal de saída desejado e o sinal de erro são armazenados em vetores. O vetor de saída bem como o gradiente do erro podem ser melhor estimados no domínio da frequência do que no domínio do tempo, uma vez que ambas as formas de representação resultam na convolução de dois vetores. Portanto, esse processo, para cada iteração, envolve a determinação de uma transformada de Fourier, uma transformada inversa de Fourier e operações de manipulação de vetores, tais como concatenação de vetores para calcular uma convolução. Sendo assim, algoritmos LMS em blocos no domínio transformado exibem uma alta complexidade computacional. Em contrapartida, possuem uma taxa de convergência elevada.

2.4.2- ALGORITMO LMS NO DOMÍNIO TRANSFORMADO

Esta família de algoritmos foi introduzida por Narayan *et al.* [4], consistindo simplesmente em um filtro LMS cujas entradas são pré-processadas por uma DFT e cujos passos de adaptação (um para cada coeficiente) são ajustados em função da potência média do sinal de entrada. Como consequências dessas duas etapas de processamento, têm-se: decorrelação do sinal de entrada, realizado pela transformação ortogonal e o aumento da taxa de convergência, provocado pela normalização do passo de adaptação. Uma vez que a DFT não é um decorrelacionador perfeito, esta estrutura não apresenta uma taxa de

convergência tão boa quanto a do algoritmo RLS, que usa a inversa da matriz de autocorrelação do sinal de entrada para descorrelacionar tal sinal. Entretanto, por não se necessitar da estimativa de \mathbf{R}^{-1} , ganha-se em robustez, além de se obter uma baixa complexidade computacional.

A equação de adaptação dos coeficientes é dada por [2]

$$\mathbf{w}_T(n+1) = \mathbf{w}_T(n) + 2\boldsymbol{\mu}(n)e(n)\mathbf{x}_T(n), \quad (2.19)$$

onde $\mathbf{x}_T(n)$ representa o vetor do sinal de entrada transformado, resultante da operação $\mathbf{T}\mathbf{x}(n)$, sendo \mathbf{T} a matriz de transformação ortogonal; $\boldsymbol{\mu} = \mu\mathbf{D}^{-1/2}$, sendo $\mathbf{D} = \text{diag}[\sigma_1^2 \ \sigma_2^2 \ \dots \ \sigma_{N-1}^2]$ e σ_i^2 é a variância de cada sub-banda, podendo ser calculada como segue:

$$\sigma_i^2 = \frac{\mathbf{X}_i^T \mathbf{X}_i}{M} \quad i = 1, 2, \dots, N, \quad (2.20)$$

onde $\mathbf{x}_i = [x_i(n) \ x_i(n-1) \ \dots \ x_i(n-M+1)]$ e M é o tamanho da janela de observação.

No próximo capítulo, o algoritmo LMS-DT será apresentado detalhadamente.

2.5- COMPARAÇÕES ENTRE OS ALGORITMOS LMS, RLS E LMS-DT

Nesta seção, são apresentados os resultados de simulações numéricas, considerando um problema de identificação de sistemas, com o intuito de avaliar o desempenho dos algoritmos LMS, RLS e LMS-DT. O sinal de entrada é correlacionado, obtido a partir de um processo auto-regressivo de segunda ordem AR(2), dado por

$$x(n) = \alpha_1 x(n-1) + \alpha_2 x(n-2) + v(n), \quad (2.20)$$

onde $v(n)$ é um ruído branco com média nula e variância σ_v^2 , α_1 e α_2 são os coeficientes AR(2), $\alpha_1 = 0,1833$ e $\alpha_2 = -0,85$. O ruído de medição $z(n)$ possui uma variância de

$\sigma_z^2 = 10^{-4}$, o que resulta em uma relação sinal-ruído $\text{SNR} = 40\text{dB}$. Os coeficientes da planta são obtidos a partir de $\text{sinc}(i/N)$ com $i = 1, 2, \dots, N$.

Exemplo 1: Dispersão dos autovalores da matriz de autocorrelação igual a 81 e número de coeficientes da planta igual à ordem do filtro adaptativo $N = 8$.

1) Algoritmo LMS

- Passo de adaptação $\mu = 1/60$.

2) Algoritmo RLS

- Fator de esquecimento $\beta = 0.999$.

3) Algoritmo LMS-DT

- Passo de adaptação $\mu = 1/60$;
- Janela para o cálculo da potência média $M = 32$;
- Matriz de transformação DCT.

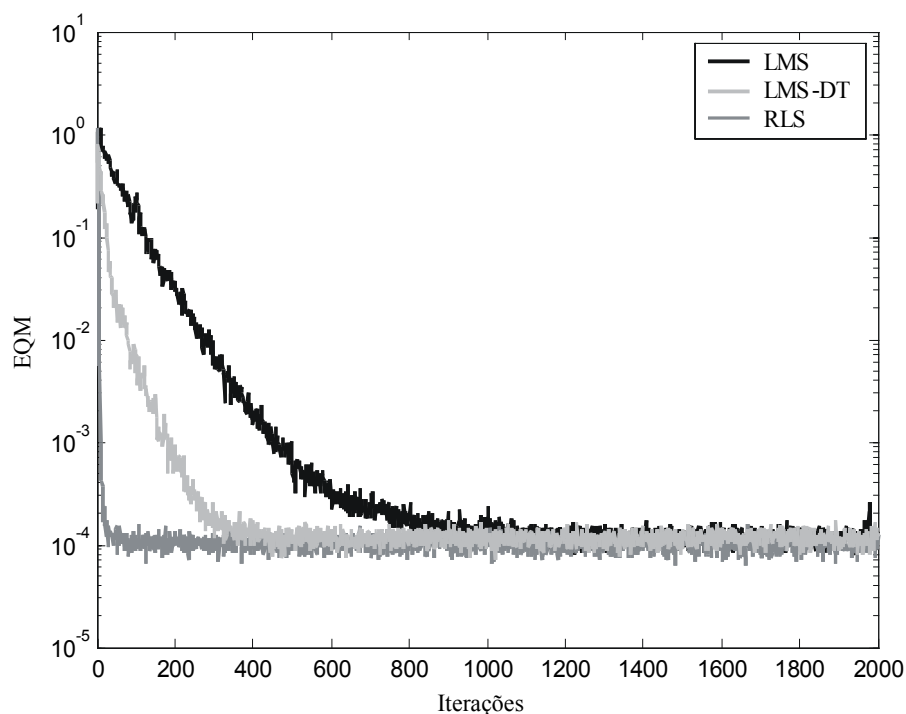


Figura 2.4- Exemplo 1. Curvas de aprendizagem dos algoritmos LMS, RLS e LMS-DT.

Exemplo 2: Dispersão dos autovalores da matriz de autocorrelação igual a 197 e número de coeficientes da planta igual à ordem do filtro adaptativo $N = 8$.

1) Algoritmo LMS:

- $\mu = 1/120$.

2) Algoritmo RLS:

- $\beta = 0.999$.

3) Algoritmo LMS-DT:

- $\mu = 1/120$;
- $M = 32$;
- matriz de transformação DCT.

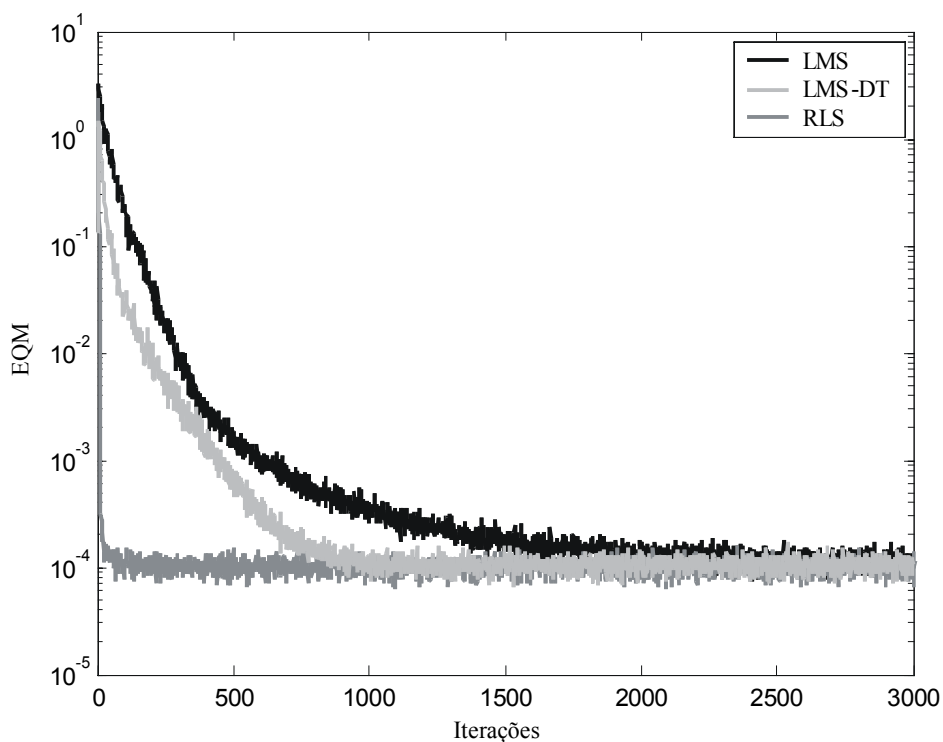


Figura 2.5- Exemplo 2. Curvas de aprendizagem dos algoritmos LMS, RLS e LMS-DT.

As Figuras 2.4 e 2.5 ilustram o comportamento do erro quadrático médio dos algoritmos avaliados. Pode ser observado que o algoritmo RLS apresenta uma mesma taxa de convergência para diferentes valores de dispersão dos autovalores da matriz de autocorrelação (diferentes valores de correlação do sinal de entrada) como já esperado. Nos dois exemplos, o algoritmo RLS apresenta uma taxa de convergência superior aos demais, sendo que o LMS-DT converge mais rapidamente do que o LMS. Vale lembrar que a velocidade de convergência do algoritmo LMS-DT está relacionada com a capacidade de ortogonalização da transformada utilizada para descorrelacionar o sinal de entrada.

2.6- CONCLUSÕES

Neste capítulo, fez-se uma breve apresentação dos conceitos básicos de filtragem adaptativa, enfatizando os dois algoritmos mais usados, o LMS e o RLS. O objetivo dessa abordagem se deve ao fato de que o algoritmo LMS-DT, objeto de estudo estatístico dessa tese, agrega algumas qualidades do RLS e do LMS: isto é, alta velocidade de convergência para sinais de entrada correlacionados e baixa complexidade computacional.

É importante ressaltar que o objetivo dos exemplos apresentados neste capítulo é comparar o comportamento do LMS-DT com os dois algoritmos que apresentam comportamentos opostos de velocidade de convergência para sinais de entrada correlacionados. A partir dos exemplos, em que os passos de adaptação do LMS e do LMS-DT são os mesmos, visando enfatizar apenas o efeito de se considerar o domínio transformado, pôde-se verificar que o algoritmo LMS-DT apresenta uma velocidade de convergência intermediária, quando comparada com os algoritmos RLS e LMS. No próximo capítulo, o algoritmo LMS-DT será estudado em mais detalhes.

CAPÍTULO 3

ALGORITMO ϵ LMS NO DOMÍNIO TRANSFORMADO

Dentre os algoritmos baseados no gradiente, o algoritmo LMS tem sido o mais empregado em aplicações práticas devido à sua simplicidade e robustez. Por outro lado, sua taxa de convergência é altamente dependente da natureza do sinal de entrada [1]-[3], ou seja, quanto mais correlacionado o sinal de entrada for, maior será a degradação da taxa de convergência do algoritmo. Para suplantar tal limitação, uma alternativa é utilizar técnicas de pré-processamento do sinal de entrada, visando geralmente seu branqueamento [3]-[6].

Neste trabalho, considera-se uma abordagem que consiste no pré-processamento do sinal de entrada através de uma transformada ortogonal. Visto que a decorrelação completa é obtida apenas com o uso da transformada de Karhunen-Loève (KLT), a qual requer o conhecimento das características estatísticas do sinal de entrada, o emprego de tal transformação torna-se inviável na prática. Dessa forma, trabalhos de pesquisa na área de filtragem adaptativa têm estudado outras transformações que se aproximam do desempenho da KLT [6], uma delas é a DCT [1], [2], [4]-[6].

Neste capítulo são mostrados alguns aspectos importantes, discutidos na literatura, sobre filtros adaptativos no domínio transformado. A Seção 3.1 apresenta a derivação da equação recursiva de atualização dos coeficientes, considerando que tal algoritmo é uma forma normalizada do algoritmo LMS. Na Seção 3.2, é mostrada a superfície de desempenho após os procedimentos de transformação ortogonal e normalização de potência [2], [5]. Assim, visualizaremos as mudanças ocorridas na geometria de tal

superfície, levando a um aumento de velocidade de convergência do algoritmo LMS-DT. Na Seção 3.3, são apresentados os efeitos da ortogonalização/normalização de potência do sinal de entrada [6], conduzindo a um melhor entendimento do resultado dessas operações na velocidade de convergência do algoritmo LMS-DT. Na Seção 3.4 é mostrada uma figura de mérito, o índice de avaliação ρ [2], capaz de fornecer alguma medida quantitativa de comparação entre as diferentes transformadas ortogonais. A Seção 3.5 apresenta a propriedade do particionamento em frequência das transformadas ortogonais. Na Seção 3.6 é apresentado o efeito da normalização de potência na análise de Wiener no domínio transformado. A Seção 3.7 traz o procedimento adotado para estimar a potência média, usado na etapa de normalização do passo de adaptação. Finalmente, na Seção 3.8, são apresentadas algumas conclusões acerca da importância da filtragem adaptativa no domínio transformado, abordada neste trabalho.

3.1- DESCRIÇÃO DO ALGORITMO LMS NO DOMÍNIO TRANSFORMADO

O algoritmo LMS no domínio transformado é composto por três estágios [6], como ilustrado na Figura 3.1. Primeiramente, o sinal de entrada é processado por uma matriz de transformação ortogonal unitária \mathbf{T} . Assim,

$$\mathbf{x}_T(n) = \mathbf{T}\mathbf{x}(n), \quad (3.1)$$

onde $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$ representa o vetor de amostras do sinal de entrada, e $\mathbf{x}_T(n) = [x_0(n) \ x_1(n) \ \dots \ x_{N-1}(n)]^T$ é o respectivo sinal transformado. Uma vez fixada a ordem do filtro N , a matriz de transformação terá dimensão $N \times N$, onde as linhas (ou colunas) são vetores ortonormais [35]. Portanto,

$$\mathbf{T}\mathbf{T}^T = \mathbf{T}^T\mathbf{T} = \mathbf{I}. \quad (3.2)$$

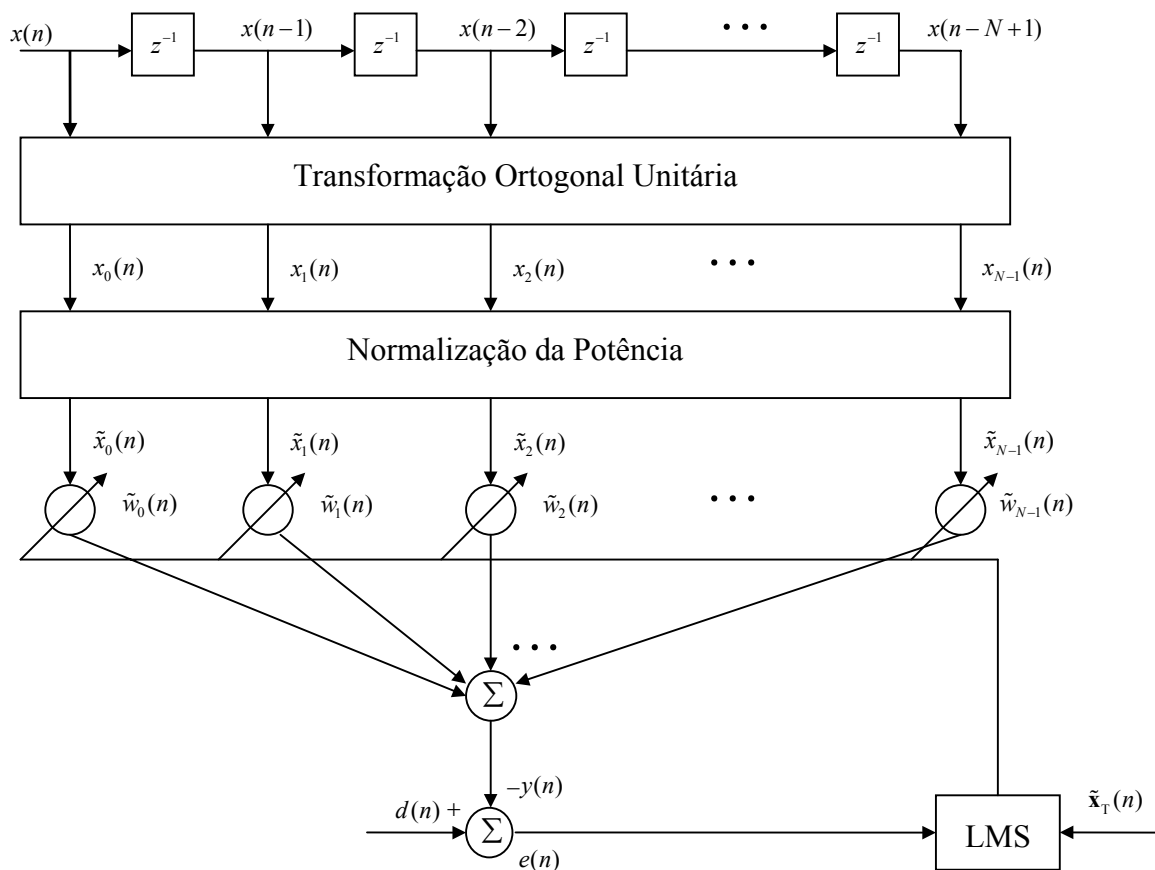


Figura 3.1- Esquema de filtragem adaptativa no domínio transformado, com a etapa de normalização explícita.

No segundo estágio do diagrama ilustrado na Figura 3.1, os sinais transformados são então normalizados pela raiz quadrada de sua potência média [6]. Assim, o sinal resultante desses dois processos (ortogonalização e normalização) é utilizado como entrada de um combinador linear adaptativo e o algoritmo LMS é usado para se obter os coeficientes do filtro. A equação de atualização dos coeficientes é dada pela seguinte expressão [1]-[3]:

$$\tilde{\mathbf{w}}_T(n+1) = \tilde{\mathbf{w}}_T(n) - \mu \nabla J(n), \quad (3.3)$$

onde $\tilde{\mathbf{w}}_T(n) = [\tilde{w}_0(n) \ \tilde{w}_1(n) \ \dots \ \tilde{w}_{N-1}(n)]^T$ é o vetor de coeficientes do filtro adaptativo no domínio transformado e normalizado e μ o passo de adaptação. A função custo $J(n)$ é definida por

$$J(n) = e^2(n), \quad (3.4)$$

e ∇ é o operador gradiente definido como

$$\nabla = \left[\frac{\partial}{\partial \tilde{w}_0(n)} \quad \frac{\partial}{\partial \tilde{w}_1(n)} \quad \dots \quad \frac{\partial}{\partial \tilde{w}_{N-1}(n)} \right]^T. \quad (3.5)$$

Da Figura 3.1, tem-se o sinal de erro determinado pela diferença entre o sinal desejado e o sinal de saída do filtro adaptativo. Assim,

$$e(n) = d(n) - y(n) = d(n) - \tilde{\mathbf{w}}_T^T(n) \tilde{\mathbf{x}}_T(n), \quad (3.6)$$

onde $\tilde{\mathbf{x}}_T(n) = [\tilde{x}_0(n) \quad \tilde{x}_1(n) \quad \dots \quad \tilde{x}_{N-1}(n)]^T$ é o vetor do sinal de entrada transformado e normalizado, determinado através da seguinte expressão [2]:

$$\tilde{\mathbf{x}}_T(n) \triangleq \mathbf{D}^{-1/2} \mathbf{x}_T(n), \quad (3.7)$$

onde $\mathbf{D} = \text{diag}[\sigma_0^2 \quad \sigma_1^2 \quad \dots \quad \sigma_{N-1}^2]$ é uma matriz diagonal constituída pelos valores de potência média relativos a cada sub-banda.

Calculando-se o gradiente da função custo, tem-se

$$\nabla J(n) = \nabla e^2(n) = 2e(n) \nabla e(n). \quad (3.8)$$

Usando o operador gradiente, definido por (3.5), o sinal de erro obtido de (3.6) e com o auxílio de (3.8), obtém-se

$$\nabla e^2(n) = -2e(n) \tilde{\mathbf{x}}_T(n). \quad (3.9)$$

Assim, pode-se determinar a equação de adaptação dos coeficientes do algoritmo LMS-DT. Portanto, através da substituição de (3.9) em (3.3), tem-se

$$\tilde{\mathbf{w}}_T(n+1) = \tilde{\mathbf{w}}_T(n) + 2\mu e(n)\tilde{\mathbf{x}}_T(n). \quad (3.10)$$

Na prática, a etapa de normalização de potência do sinal transformado é efetuada através do passo de adaptação que varia com a potência média em cada sub-banda de frequência [2].

Partindo-se então do princípio da conservação da energia, tem-se que o erro de adaptação deve ser mantido após qualquer operação efetuada no sinal de entrada. Assim,

$$y(n) = \mathbf{x}_T^T(n)\mathbf{w}_T(n) = \tilde{\mathbf{x}}_T^T(n)\tilde{\mathbf{w}}_T(n) = \mathbf{x}_T^T(n)\mathbf{D}^{-1/2}\tilde{\mathbf{w}}_T(n), \quad (3.11)$$

o que implica em $\mathbf{w}_T(n) = \mathbf{D}^{-1/2}\tilde{\mathbf{w}}_T(n)$, onde $\mathbf{w}_T(n)$ é o vetor de coeficientes transformados. Dessa forma, pode-se determinar a equação de adaptação dos coeficientes do algoritmo LMS-DT ilustrado na Figura 3.2. Assim, considerando-se a relação entre $\mathbf{w}_T(n)$ e $\tilde{\mathbf{w}}_T(n)$, e usando (3.10), obtém-se [2]

$$\begin{aligned} \mathbf{w}_T(n+1) &= \mathbf{D}^{-1/2}\tilde{\mathbf{w}}_T(n+1) \\ &= \mathbf{D}^{-1/2}[\tilde{\mathbf{w}}_T(n) + 2\mu e(n)\tilde{\mathbf{x}}_T(n)] \\ &= \mathbf{D}^{-1/2}\tilde{\mathbf{w}}_T(n) + 2\mu e(n)\mathbf{D}^{-1/2}\mathbf{D}^{-1/2}\mathbf{x}_T(n) \\ &= \mathbf{w}_T(n) + 2\mu\mathbf{D}^{-1}e(n)\mathbf{x}_T(n). \end{aligned} \quad (3.12)$$

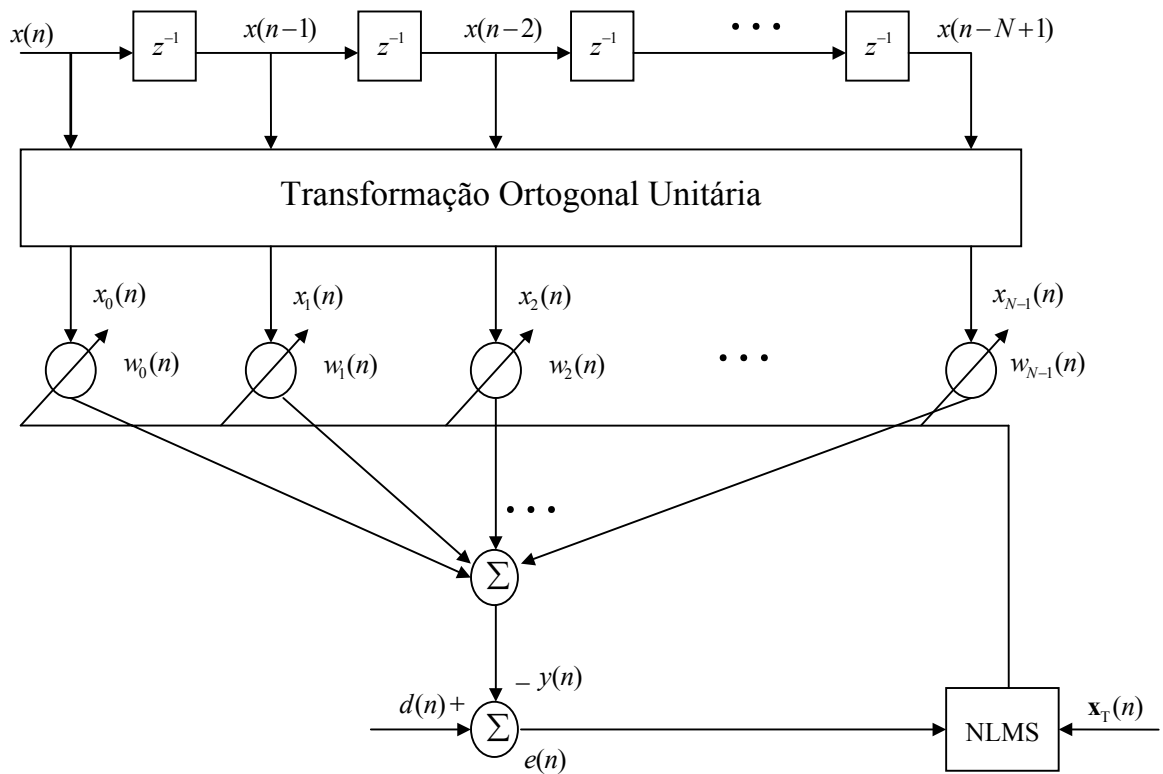


Figura 3.2- Esquema de filtragem adaptativa no domínio transformado.

Note de (3.12) que agora no lugar do algoritmo LMS usa-se o algoritmo LMS normalizado (NLMS), considerando a potência média de cada sub-banda (ver Figura 3.2).

3.2- EFEITOS DA TRANSFORMAÇÃO/NORMALIZAÇÃO NA SUPERFÍCIE DE DESEMPENHO

Nesta seção, será ilustrada a superfície de desempenho após as operações de transformação e normalização [2], [5]. Tal superfície pode ser escrita da seguinte forma [3]:

$$\xi[\mathbf{v}(n)] = \xi_{\min} + \mathbf{v}^T \mathbf{R} \mathbf{v}, \quad (3.13)$$

onde $\mathbf{R} \triangleq E[\mathbf{x}(n)\mathbf{x}^T(n)]$, como definida anteriormente, é a matriz de autocorrelação do sinal de entrada e $\mathbf{v}(n) \triangleq \mathbf{w}(n) - \mathbf{w}_{\text{ótimo}}$ é o vetor de erros nos coeficientes, onde $\mathbf{w}_{\text{ótimo}}$ é o vetor ótimo.

Para uma melhor visualização de tal procedimento, considere-se um filtro com dois coeficientes. Assim, assumindo-se um sinal de entrada correlacionado, a matriz de autocorrelação é dada por

$$\mathbf{R} = \begin{bmatrix} 1 & 0,9 \\ 0,9 & 1 \end{bmatrix}. \quad (3.14)$$

Para tais condições, (3.13) se torna

$$\xi(v_0(n), v_1(n)) = \xi_{\min} + v_0^2(n) + v_1^2(n) + 1,8v_0(n)v_1(n). \quad (3.15)$$

Se não for aplicada a normalização de potência nas amostras transformadas, a superfície de desempenho é igual a [2]

$$\xi_T[\mathbf{v}_T(n)] = \xi_{\min} + \mathbf{v}_T^T(n)\mathbf{R}_T\mathbf{v}_T(n), \quad (3.16)$$

onde $\mathbf{R}_T \triangleq E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)] = \mathbf{TRT}^T$. Utilizando-se, por exemplo, a transformação DCT com

$$\mathbf{T} = \begin{bmatrix} 0.707 & 0.707 \\ 0.707 & -0.707 \end{bmatrix}, \text{ tem-se como resultado}$$

$$\mathbf{R}_{T,\text{DCT}} = \begin{bmatrix} 1,9 & 0 \\ 0 & 0,1 \end{bmatrix}. \quad (3.17)$$

Assim, a superfície de desempenho é expressa por

$$\xi_T(v_{T,0}(n), v_{T,1}(n)) = \xi_{\min} + 1,9v_{T,0}^2(n) + 0,1v_{T,1}^2(n). \quad (3.18)$$

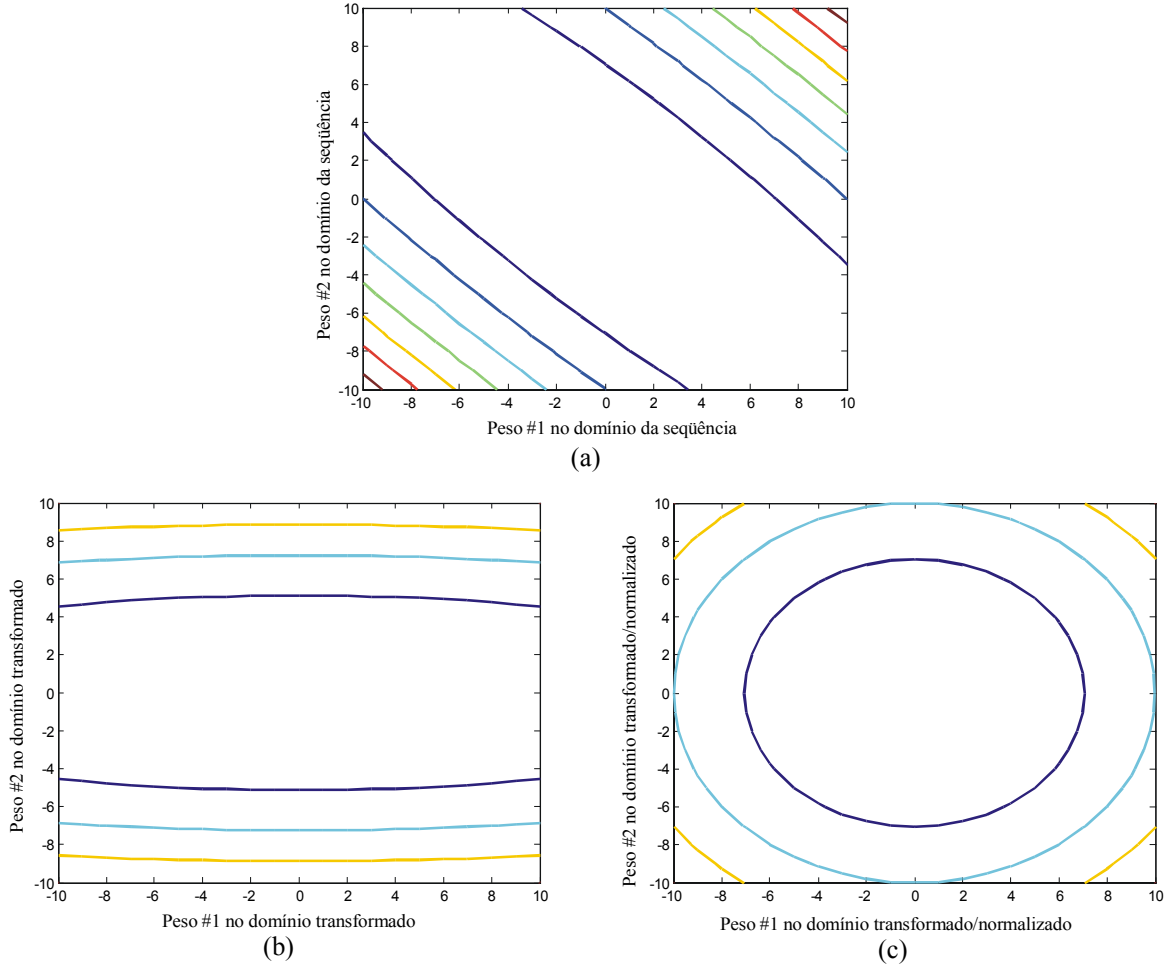


Figura 3.3- Curvas de nível das superfícies de desempenho do filtro adaptativo. (a) Domínio da seqüência. (b) Domínio transformado sem normalização. (c) Domínio transformado com normalização.

As Figuras 3.3(a) e 3.3(b) ilustram as curvas de nível da superfície de desempenho no domínio da seqüência [equação (3.15)] e no domínio transformado [equação (3.18)], respectivamente. A partir da Figura 3.3(b), pode-se observar que o efeito da transformação unitária é rotacionar a superfície, posicionando-a aproximadamente alinhada em relação aos eixos coordenados, visto que $\mathbf{T}\mathbf{T}^T = \mathbf{I}$. Assim,

$$\begin{aligned}
 \xi[\mathbf{v}(n)] &= \xi_{\min} + \mathbf{v}^T(n)\mathbf{R}_x\mathbf{v}(n) \\
 &= \xi_{\min} + \mathbf{v}^T(n)\mathbf{T}^T\mathbf{T}\mathbf{R}_x\mathbf{T}^T\mathbf{T}\mathbf{v}(n) \\
 &= \xi_{\min} + \mathbf{v}_T^T(n)\mathbf{R}_T\mathbf{v}_T(n) = \xi_T(\mathbf{v}_T(n)),
 \end{aligned} \tag{3.19}$$

onde $\mathbf{v}(n)$ e $\mathbf{v}_T(n)$ estão relacionados pela equação $\mathbf{v}_T(n) = \mathbf{T}\mathbf{v}(n)$. A transformação ortogonal não altera a velocidade de convergência do algoritmo LMS, pois a não

modificação da geometria da superfície de desempenho significa que não se promove qualquer alteração nos autovalores da matriz de autocorrelação do sinal de entrada. Quando os elementos de $\mathbf{x}_T(n)$ são normalizados para a potência unitária, a superfície de desempenho é representada como segue [2]:

$$\tilde{\xi}_T[\tilde{\mathbf{v}}_T(n)] = \xi_{\min} + \tilde{\mathbf{v}}_T^T(n) \tilde{\mathbf{R}}_T \tilde{\mathbf{v}}_T(n), \quad (3.20)$$

onde $\tilde{\mathbf{R}}_T \triangleq \mathbf{D}^{-1/2} \mathbf{R}_T \mathbf{D}^{-1/2}$.

Utilizando-se a matriz de autocorrelação no domínio transformado dada por (3.17) e considerando que $\mathbf{D} = \text{diag}[\mathbf{R}_T]$, tem-se

$$\tilde{\xi}_T[\tilde{v}_{T,0}(n), \tilde{v}_{T,1}(n)] = \xi_{\min} + \tilde{v}_{T,0}^2(n) + \tilde{v}_{T,1}^2(n). \quad (3.21)$$

A partir da Figura 3.3(c), pode-se visualizar que a normalização reduz a excentricidade da superfície de desempenho [2], [5]. Essa operação não-unitária modifica os autovalores da matriz de autocorrelação, resultando em uma velocidade de convergência maior do algoritmo LMS-DT.

3.3- EFEITOS DA TRANSFORMAÇÃO E NORMALIZAÇÃO NA MATRIZ DE AUTOCORRELAÇÃO

O objetivo desta seção é analisar o efeito de “branqueamento” do sinal de entrada sobre sua matriz de autocorrelação. O processo de ortogonalização transforma a matriz de autocorrelação \mathbf{R} , que é uma matriz Toeplitz¹, em uma matriz não-Toeplitz \mathbf{R}_T . Se a transformação é ótima, ou seja, representa um descorrelacionador perfeito, \mathbf{R}_T é uma matriz diagonal. Na prática, as transformadas utilizadas não são descorrelacionadores perfeitos, o que implica que a matriz de autocorrelação \mathbf{R}_T não seja uma matriz diagonal [6]. No entanto, com uma escolha adequada de transformação ortogonal, os elementos fora

¹ Uma matriz quadrada é dita Toeplitz se todos os elementos de sua diagonal principal e também todos os outros elementos de quaisquer outras diagonais, paralelas à diagonal principal, são iguais [1].

da diagonal principal têm magnitudes muito pequenas, quando comparadas com aquelas dos elementos da diagonal principal, podendo até ser desconsiderados em casos práticos.

A partir de (3.7), pode-se obter a matriz de autocorrelação, após a transformação e normalização de potência, da seguinte forma:

$$\begin{aligned}\tilde{\mathbf{R}}_T &\triangleq E[\mathbf{D}^{-1/2} \mathbf{x}_T(n) \mathbf{x}_T^T(n) \mathbf{D}^{-1/2}] \\ &= E[\mathbf{D}^{-1/2}] E[\mathbf{x}_T(n) \mathbf{x}_T^T(n)] E[\mathbf{D}^{-1/2}] \\ &= \text{diag}(\mathbf{R}_T)]^{-1/2} \mathbf{R}_T [\text{diag}(\mathbf{R}_T)]^{-1/2}.\end{aligned}\quad (3.22)$$

Se a matriz de autocorrelação do sinal de entrada transformado \mathbf{R}_T é uma matriz diagonal, isso implica que a matriz de autocorrelação do sinal de entrada transformado/normalizado $\tilde{\mathbf{R}}_T$ é forçosamente uma matriz identidade.

Na Figura 3.4, pode-se observar a transformação da matriz de autocorrelação do sinal de entrada em cada etapa do processo de “branqueamento”.

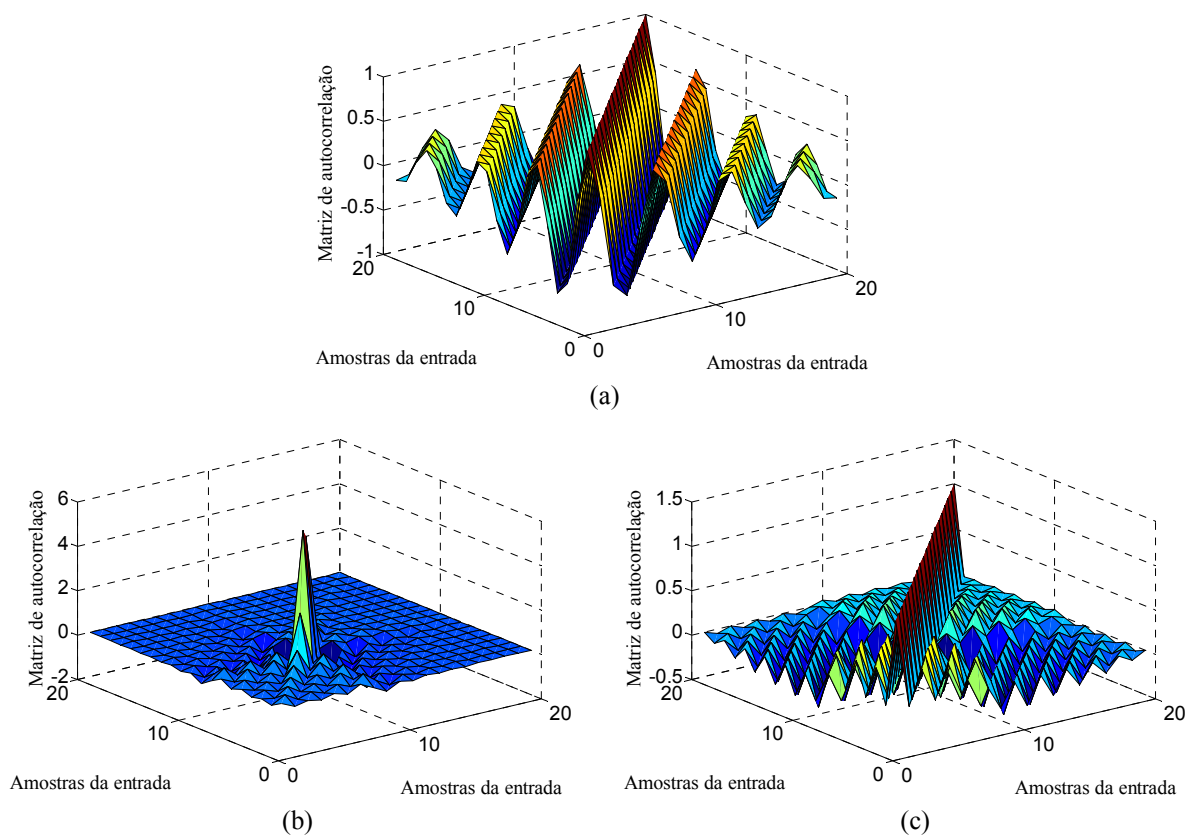


Figura 3.4- Superfície da matriz de autocorrelação do sinal de entrada. (a) Domínio da seqüência. (b) Domínio transformado (DCT) sem normalização. (c) Domínio transformado (DCT) com normalização.

3.4- ÍNDICE DE AVALIAÇÃO ρ

Uma forma de se avaliar o efeito das transformações ortogonais, relacionando-as com as características de convergência do algoritmo LMS, é através do **índice de avaliação** $\rho(\mathbf{R})$, apresentado em [2] e definido como:

$$\rho(\mathbf{R}) = \left(\frac{\underline{\lambda}_a}{\underline{\lambda}_g} \right)^N, \quad (3.23)$$

onde $\underline{\lambda}_a$ e $\underline{\lambda}_g$ representam, respectivamente, a média aritmética e a média geométrica dos autovalores de \mathbf{R} . O valor de $\rho(\mathbf{R})$ é sempre maior ou igual a 1, se aproximando da unidade quando todos os autovalores de \mathbf{R} assumem, aproximadamente, o mesmo valor; caso contrário, afasta-se da unidade. Usando as identidades [2]

$$\sum_i \lambda_i = \text{tr}[\mathbf{R}], \quad (3.24)$$

e

$$\prod_i \lambda_i = \det[\mathbf{R}], \quad (3.25)$$

onde $\text{tr}[\mathbf{R}]$ é o traço de \mathbf{R} e $\det[\mathbf{R}]$ é o determinante de \mathbf{R} , tem-se

$$\rho(\mathbf{R}) = \frac{\{\text{tr}[\mathbf{R}]\}^N}{N^N \det[\mathbf{R}]}. \quad (3.26)$$

O índice $\rho(\cdot)$ não depende somente da dispersão da matriz \mathbf{R} , tendo autovalores máximo e mínimo denotados por λ_{\max} e λ_{\min} , respectivamente, mas também da distribuição dos demais autovalores compreendidos entre λ_{\max} e λ_{\min} . Entretanto, a tendência geral é que para um alto valor de dispersão dos autovalores de \mathbf{R} , resulta em um alto valor do índice $\rho(\cdot)$ [2].

Analisando-se o comportamento de $\rho(\cdot)$ após o uso da transformada ortogonal, obtém-se:

$$\rho(\mathbf{R}_T) = \frac{\{\text{tr}[\mathbf{R}_T]\}^N}{N^N \det[\mathbf{R}_T]} = \frac{\{\text{tr}[\mathbf{TRT}^T]\}^N}{N^N \det[\mathbf{TRT}^T]}. \quad (3.27)$$

Usando as propriedades matriciais:

- Se $\mathbf{A}_{m \times n}$ e $\mathbf{B}_{n \times m}$, $\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]$;
- Se \mathbf{A} e \mathbf{B} são matrizes quadradas, $\det[\mathbf{AB}] = \det[\mathbf{BA}] = \det[\mathbf{A}]\det[\mathbf{B}]$, tem-se

$$\rho(\mathbf{R}_T) = \frac{\{\text{tr}[\mathbf{T}^T \mathbf{TR}]\}^N}{N^N \det[\mathbf{T}^T \mathbf{TR}]} = \rho(\mathbf{R}). \quad (3.28)$$

Através de (3.28), verifica-se que a transformação ortogonal (não considerando normalização de potência) mantém os autovalores de \mathbf{R} e \mathbf{R}_T , resultado esse bem conhecido da literatura [2].

Considerando, agora, a etapa de normalização após a transformação, o índice $\rho(\cdot)$ pode ser escrito como:

$$\rho(\tilde{\mathbf{R}}_T) = \frac{\{\text{tr}[\tilde{\mathbf{R}}_T]\}^N}{N^N \det[\tilde{\mathbf{R}}_T]}. \quad (3.29)$$

Como $\text{tr}[\tilde{\mathbf{R}}_T] = N$, tem-se:

$$\begin{aligned} \rho(\tilde{\mathbf{R}}_T) &= \frac{1}{\det[\mathbf{D}^{-1/2} \mathbf{R}_T \mathbf{D}^{-1/2}]} \\ &= \frac{1}{\det[\hat{\mathbf{D}}^{-1}] \det[\mathbf{R}_T]}. \end{aligned} \quad (3.30)$$

Se $\det(\mathbf{D}) \neq 0$, e utilizando a propriedade $\det(\mathbf{D}^{-1}) = [\det(\mathbf{D})]^{-1}$, pode-se escrever:

$$\begin{aligned}
\rho(\tilde{\mathbf{R}}_T) &= \frac{\det(\mathbf{D})}{\det(\mathbf{R}_T)} = \frac{\det[\text{diag}(\mathbf{R}_T)]}{\det(\mathbf{R}_T)} \\
&= \frac{\det[\text{diag}(\mathbf{R}_T)][\text{tr}(\mathbf{R}_T)/N]^N}{\{\text{tr}[\text{diag}(\mathbf{R}_T)]/N\}^N \det(\mathbf{R}_T)} \\
&= \frac{\rho(\mathbf{R}_T)}{\rho[\text{diag}(\mathbf{R}_T)]}.
\end{aligned} \tag{3.31}$$

A partir de (3.31), pode-se ver que se a matriz \mathbf{R}_T é aproximadamente diagonal, então $\rho(\tilde{\mathbf{R}}_T) = 1$.

3.5- PROPRIEDADE DO PARTICIONAMENTO EM FREQUÊNCIA

A transformada linear genérica $s_{T,k}(n)$ de uma seqüência $s(\cdot)$ de N amostras $[s(n) \ s(n-1) \ \dots \ s(n-N+1)]$ é definida como [36]

$$s_{T,k}(n) = \sum_{l=0}^{N-1} t_{kl} s(n-l), \quad k = 0, \dots, N-1, \tag{3.32}$$

onde k define os diferentes elementos do vetor $s_T(n)$, e t_{kl} representa os elementos da matriz transformação correspondente. Tal transformação pode ser realizada por uma implementação via banco de filtros FIR, onde os coeficientes do filtro são os termos que constituem a matriz transformação. A função de transferência do correspondente filtro é dada por

$$T_k(z) = \sum_{l=0}^{N-1} t_{kl} z^{-l}. \tag{3.33}$$

A Figura 3.5 ilustra a resposta em frequência de (3.33), considerando como transformação ortogonal a DCT², $k = 0$ e quatro diferentes valores de N {4, 8, 32, 64}. Pode ser observado que o processo de ortogonalização conseguirá um melhor particionamento em frequência do sinal a ser processado para valores de N mais elevados (com $N \geq 8$ tem-se já resultados satisfatórios). Isso se deve ao fato de que as respostas em magnitude de cada filtro, associado à DCT, apresentarem os lóbulos laterais mais atenuados em relação ao lóbulo principal à medida que o valor de N aumenta. Além disso, os filtros tornam-se mais seletivos também com o aumento de N .

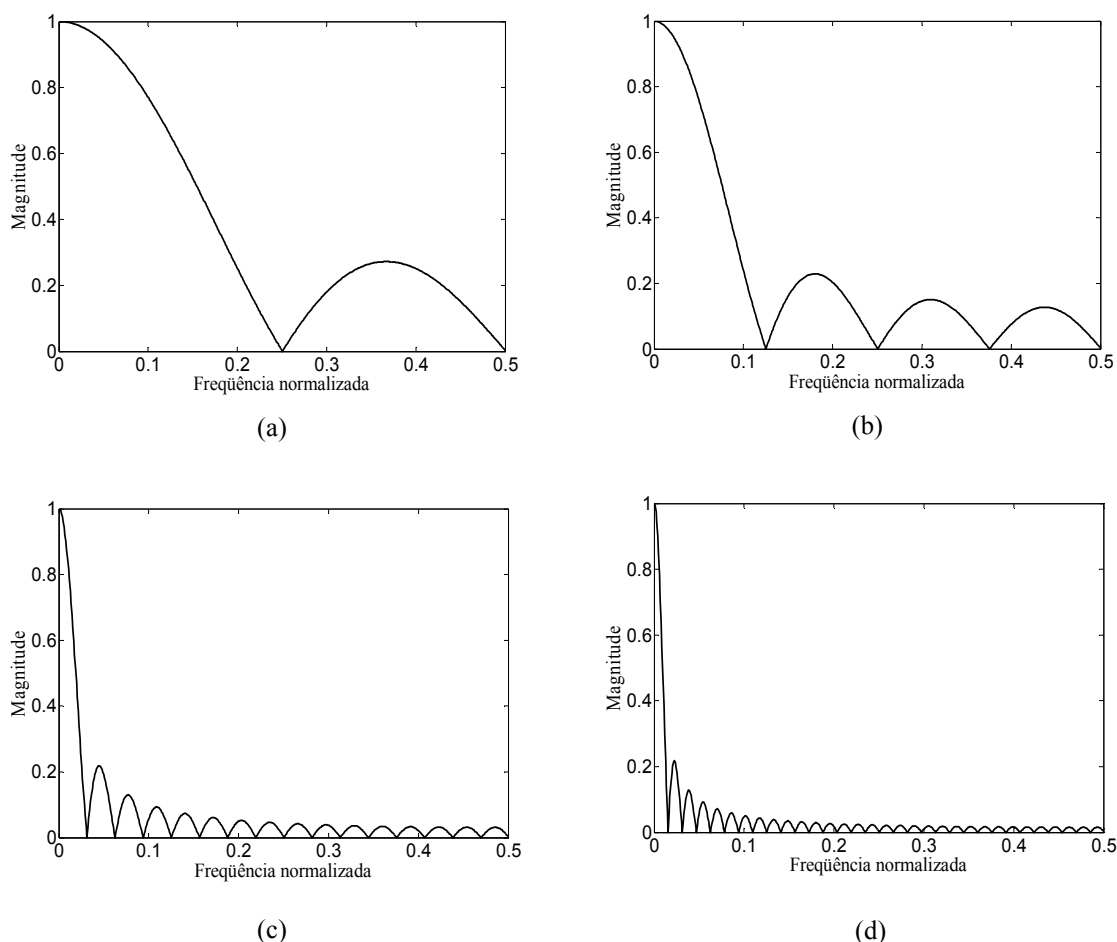


Figura 3.5- Curvas de resposta em frequência dos filtros associados à matriz de transformação DCT. (a) $N = 4$; (b) $N = 8$; (c) $N = 32$ e (d) $N = 64$.

² DCT utilizada [2]:

$$t_{kn} = \begin{cases} \frac{1}{\sqrt{N}}, & k = 0, \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N}, & 1 \leq k \leq N-1. \end{cases} \quad \text{para } 0 \leq n \leq N-1.$$

3.6- VETOR DE COEFICIENTES ÓTIMO NO DOMÍNIO TRANSFORMADO E NORMALIZADO

Nesta seção é obtido o vetor de coeficientes ótimo no domínio transformado, incluindo a etapa de normalização de potência média (conforme Figura 3.6).

De acordo com a Figura 3.1, a expressão do sinal de erro é dada por

$$e(n) = d(n) - \tilde{\mathbf{w}}_T^T \tilde{\mathbf{x}}_T(n) \quad (3.41)$$

Elevando-se ao quadrado (3.41), tem-se o erro quadrático instantâneo, dado por

$$e^2(n) = d^2(n) - 2\tilde{\mathbf{w}}_T^T \tilde{\mathbf{x}}_T(n)d(n) + \tilde{\mathbf{w}}_T^T \tilde{\mathbf{x}}_T(n)\tilde{\mathbf{x}}_T^T(n)\tilde{\mathbf{w}}_T. \quad (3.42)$$

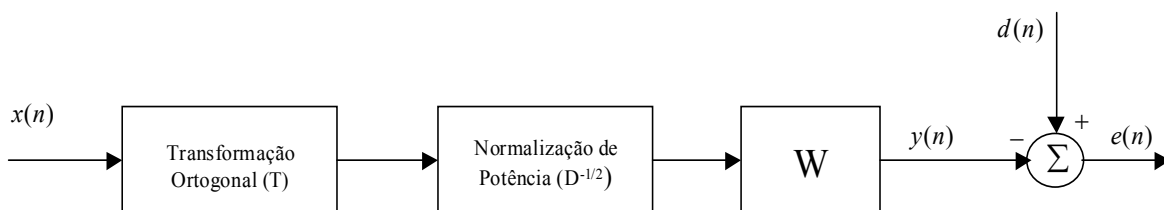


Figura 3.6- Diagrama em blocos para a determinação da solução de Wiener.

Tomando-se o valor esperado em ambos os lados de (3.42), obtém-se o erro quadrático médio, dado por

$$\xi = E[e^2(n)] = E[d^2(n) - 2\tilde{\mathbf{w}}_T^T \tilde{\mathbf{x}}_T(n)d(n) + \tilde{\mathbf{w}}_T^T \tilde{\mathbf{x}}_T(n)\tilde{\mathbf{x}}_T^T(n)\tilde{\mathbf{w}}_T]. \quad (3.43)$$

Considerando $d(n)$ e $\tilde{\mathbf{x}}_T(n)$ estatisticamente estacionários e correlacionados, obtém-se:

$$\xi = E[d^2(n)] - 2\tilde{\mathbf{w}}_T^T \tilde{\mathbf{p}}_T + \tilde{\mathbf{w}}_T^T \tilde{\mathbf{R}}_T \tilde{\mathbf{w}}_T. \quad (3.44)$$

onde, como antes definido, $\tilde{\mathbf{R}}_T = E[\tilde{\mathbf{x}}_T(n)\tilde{\mathbf{x}}_T^T(n)]$ é a matriz de autocorrelação do sinal de entrada transformado e normalizado e $\tilde{\mathbf{p}}_T = E[\tilde{\mathbf{x}}_T(n)d(n)]$ é o vetor de correlação cruzada entre $\tilde{\mathbf{x}}_T(n)$ e $d(n)$.

Para determinar o valor de $\tilde{\mathbf{w}}_T$ que conduz ao mínimo erro quadrático médio, deve-se determinar o $\nabla \xi = 0$, o que resulta em

$$\nabla \xi = 2\tilde{\mathbf{R}}_T \tilde{\mathbf{w}}_{T_{\text{ótimo}}} - 2\tilde{\mathbf{p}}_T = 0, \quad (3.45)$$

portanto

$$\tilde{\mathbf{w}}_{T_{\text{ótimo}}} = \tilde{\mathbf{R}}_T^{-1} \tilde{\mathbf{p}}_T. \quad (3.46)$$

Pode-se, pelo desmembramento da equação anterior, obter uma relação entre o vetor ótimo no domínio transformado $\tilde{\mathbf{w}}_{T_{\text{ótimo}}}$ e o vetor ótimo no domínio não transformado, $\mathbf{w}_{\text{ótimo}}$. Assim,

$$\begin{aligned} \tilde{\mathbf{w}}_{T_{\text{ótimo}}} &= [\mathbf{D}^{-1/2} \mathbf{T} \mathbf{R}_x \mathbf{T}^T \mathbf{D}^{-1/2}]^{-1} E[d(n) \mathbf{D}^{-1/2} \mathbf{T} \mathbf{x}(n)] \\ &= \mathbf{D}^{1/2} \mathbf{T}^{-T} \mathbf{R}_x^{-1} \mathbf{T}^{-1} \mathbf{D}^{1/2} \mathbf{D}^{1/2} \mathbf{T} E[\mathbf{x}(n) d(n)] \\ &= \mathbf{D}^{1/2} \mathbf{T} \mathbf{w}_{\text{ótimo}}. \end{aligned} \quad (3.47)$$

De (3.47), observa-se que o vetor ótimo está intimamente relacionado às características estatísticas do sinal de entrada e do sinal desejado. Deve-se ressaltar que a operação de transformação ortogonal realizada no sinal de entrada é refletida nos coeficientes do filtro adaptativo, de forma que esses englobem a operação de transformação ortogonal inversa.

Agora, para se obter o mínimo erro atingível, substitui-se (3.46) em (3.44), resultando em

$$\begin{aligned}
\tilde{\xi}_{\min_T} &= E[d^2(n)] - \tilde{\mathbf{p}}_T^T \tilde{\mathbf{R}}_T^{-1} \tilde{\mathbf{p}}_T \\
&= E[d^2(n)] - (\mathbf{D}^{-1/2} \mathbf{T} \mathbf{p}_x)^T (\mathbf{D}^{-1/2} \mathbf{T} \mathbf{R}_x \mathbf{T}^T \mathbf{D}^{-1/2})^{-1} (\mathbf{D}^{-1/2} \mathbf{T} \mathbf{p}_x) \\
&= E[d^2(n)] - \mathbf{p}_x^T \mathbf{R}_x^{-1} \mathbf{p}_x \\
&= \xi_{\min}.
\end{aligned} \tag{3.48}$$

De (3.48), pode-se concluir que o mínimo erro permanece inalterado após os processos de transformação ortogonal e de normalização de potência.

3.7- ESTIMATIVA DE POTÊNCIA MÉDIA

Para a implementação da equação (3.12), necessita-se de uma estimativa da potência média do sinal na i -ésima sub-banda, ou seja, de cada termo que constitui a matriz \mathbf{D} . Uma forma de se determinar a potência média do sinal de entrada é através da filtragem passa-baixas do correspondente sinal de potência instantânea [36].

A função de transferência do filtro passa-baixas geralmente utilizada para estimar a potência média é dada por

$$H(z) = \frac{(1-\gamma)z}{(z-\gamma)}, \tag{3.49}$$

onde $0 < \gamma < 1$, sendo escolhidos valores próximos da unidade. Essa constante exerce um papel importante na estimativa, pois quanto mais próxima da unidade ela for, melhor será a estimativa de potência média. A Figura 3.7 ilustra a resposta em frequência (3.49) para três diferentes valores de γ , designados por $H_1(z)$, $H_2(z)$ e $H_3(z)$.

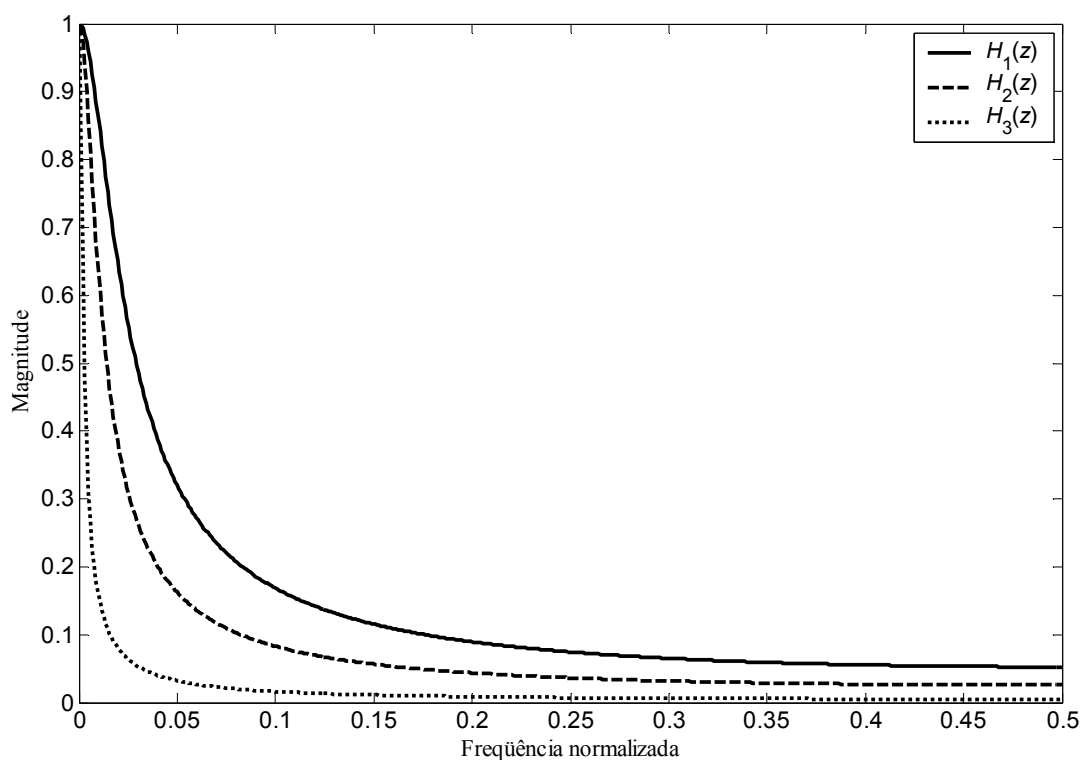


Figura 3.7- Comportamento do filtro passa-baixas para diferentes valores de γ . $\gamma = 0.9$, $\gamma = 0.95$ e $\gamma = 0.99$ estão associados às curvas de $H_1(z)$, $H_2(z)$ e $H_3(z)$, respectivamente.

A equação de diferenças associada à equação (3.49) [2] é

$$\hat{\sigma}_i^2(n) = \gamma \hat{\sigma}_i^2(n-1) + (1-\gamma)x_i^2(n), \quad i = 0, \dots, N-1, \quad (3.50)$$

onde $\hat{\sigma}_i^2(n)$ e $x_i^2(n)$ representam, em cada i -ésima sub-banda, respectivamente, uma estimativa da variância do sinal e a potência instantânea.

Outra forma de obter a potência média é assumir que o processo que gera o sinal de entrada seja ergódico. Assim, pode-se definir uma estimativa da matriz de autocorrelação do sinal de entrada transformado como

$$\hat{\mathbf{R}}_x(n) = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_T(i) \mathbf{x}_T^T(i). \quad (3.51)$$

Equivalentemente, tem-se

$$\hat{\sigma}_i^2(n) = \frac{1}{n} \sum_{k=1}^n x_i^2(k), \quad i = 0, \dots, N-1. \quad (3.52)$$

Para uma janela de observação de M amostras, (3.52) pode ser aproximada por

$$\hat{\sigma}_i^2(n) \cong \frac{1}{M} \sum_{k=0}^{M-1} x_i^2(n-k), \quad i = 0, \dots, N-1. \quad (3.53)$$

Adicionando uma constante $\varepsilon \ll 1$ (parâmetro de regularização) tal que não afete o comportamento do algoritmo e assegure a estabilidade para os instantes em que o valor da potência média seja muito próximo de zero, tem-se a estimativa de potência média, dada por

$$\tilde{\sigma}_i^2(n) = \hat{\sigma}_i^2(n) + \varepsilon. \quad (3.54)$$

Neste trabalho, a opção adotada para estimar a potência média é obtida por (3.53)

3.8- CONCLUSÕES

Neste capítulo, foi analisado o comportamento do filtro adaptativo LMS no domínio transformado salientando-se sua relevância quando o sinal de entrada apresenta amostras correlacionadas. É mostrado o aumento da velocidade de convergência obtido após as etapas de transformação ortogonal e normalização de potência, quando o algoritmo LMS-DT é comparado ao LMS convencional. É importante ressaltar que a transformação objetiva descorrelacionar o sinal de entrada, ao passo que a normalização busca atenuar a dispersão dos autovalores da matriz de autocorrelação desse sinal.

Destaca-se também que, conforme [1], o algoritmo LMS-DT pode ser interpretado com uma forma normalizada do algoritmo LMS convencional. Dessa maneira, os procedimentos de derivação da equação recursiva de atualização dos coeficientes do filtro são análogos para os dois casos.

Com relação à análise de Wiener, conclui-se que a solução ótima no domínio transformado, assim como no domínio não transformado, está intimamente relacionada às estatísticas do sinal de entrada e do sinal desejado. Quanto ao mínimo erro, observa-se que esse não se altera com a aplicação da transformação ortogonal, ou seja, o mapeamento do sinal não traz benefícios nem prejuízos, o que é ratificado pelo princípio de conservação de energia.

Foi observado também que se o sinal de entrada é obtido de um processo estocástico ergódico, isso permite adotar uma estratégia de estimativa da potência média que simplifica a análise estatística do algoritmo, como será visto no próximo capítulo.

CAPÍTULO 4

ANÁLISE ESTATÍSTICA DO ALGORITMO ϵ LMS-DT

Algumas análises estatísticas do algoritmo LMS-DT encontradas na literatura consideram que a matriz \mathbf{D} , definida no Capítulo 3, responsável pela etapa de normalização de potência, é invariante no tempo [24], [25]. Essa consideração simplifica consideravelmente o tratamento matemático envolvido para a obtenção do modelo estocástico. No entanto, as expressões do modelo resultante não levam em conta a natureza variante no tempo do parâmetro do passo de adaptação. Como visto anteriormente, o estágio de normalização de potência fornece um passo de adaptação normalizado similar ao do algoritmo LMS normalizado (NLMS), que é variante no tempo. Portanto, da mesma forma como no caso do algoritmo NLMS [21]-[23], o algoritmo LMS-DT também apresenta um procedimento matemático complicado para se derivar os momentos de primeira e segunda ordens dos coeficientes do filtro adaptativo. Usando-se algumas considerações, obtém-se uma análise razoavelmente simplificada, em particular, para os cálculos dos valores esperados dos termos envolvendo o passo de adaptação variante no tempo. Em [26]-[31], um modelo analítico para o algoritmo LMS-DT é apresentado, para o qual é usado o Princípio da Média [37] como uma hipótese simplificativa. Além disso, o modelo discutido em [26] é preciso apenas para condição de adaptação extremamente lenta. Ainda nessa análise, assim como em [27]-[31], é feito o uso da independência estatística intrabanda em cada banda do processo de transformação ortogonal. Levando-se em conta essa prévia discussão, um

modelo estocástico para o algoritmo LMS-DT pode ser ainda considerado como uma questão aberta na literatura da área.

Neste capítulo, é apresentada uma análise estatística do algoritmo LMS-DT, admitindo a real natureza estatística do sinal que compõe cada sub-banda do sinal processado pela transformação ortogonal. As expressões são derivadas para os momentos de primeira e segunda ordens do vetor de coeficientes do filtro adaptativo para entradas com distribuição Gaussiana, sem a necessidade de expressões auxiliares como em [26]. Considerações simplificativas são usadas permitindo um tratamento matemático menos penoso, porém fornecendo uma boa predição para o comportamento estatístico do referido algoritmo. As expressões teóricas para o passo de adaptação máximo e para o desajuste do algoritmo são também derivadas.

4.1- ANÁLISE ESTATÍSTICA

4.1.1- MOMENTO DE PRIMEIRA ORDEM DO VETOR $\mathbf{w}_T(n)$

Substituindo-se \mathbf{D} por $\mathbf{D}(n)$ em (3.12), a equação de evolução dos coeficientes pode ser reescrita como

$$\mathbf{w}_T(n+1) = \mathbf{w}_T(n) + 2\mu\mathbf{D}^{-1}(n)e(n)\mathbf{x}_T(n), \quad (4.1)$$

onde $e(n) = d(n) - \mathbf{w}_T^T(n)\mathbf{x}_T(n) + z(n)$ é o erro de estimativa e $z(n)$ é o ruído de medição, decorrelacionado de qualquer outro sinal envolvido no processo, com distribuição de probabilidade Gaussiana, média nula e variância σ_z^2 . Tomando-se o valor esperado de (4.1), obtém-se

$$E[\mathbf{w}_T(n+1)] = E[\mathbf{w}_T(n)] + 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)d(n)] - 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)[\mathbf{w}_T(n)]] \quad (4.2)$$

Para se determinar os valores esperados em (4.2), adotaram-se as seguintes considerações:

H1)

$\mathbf{w}_T(n)$ e $\mathbf{x}_T(n)$ são estatisticamente independentes [1];

H2)

$\mathbf{D}^{-1}(n)$ e $\mathbf{x}_T(n)\mathbf{x}_T^T(n)$ são processos conjuntamente estacionários, tal que $\mathbf{D}^{-1}(n)$ varia lentamente em relação a $\mathbf{x}_T(n)\mathbf{x}_T^T(n)$. O Princípio da Média [37] pode assim ser adotado. Tal princípio diz: sejam dois processos $f[a(n)]$ e $f[b(n)]$, conjuntamente estacionários, com $b(n)$ variando lentamente em relação a $a(n)$, então as variáveis aleatória $a(n)$ e $b(n)$ são quase independentes. Portanto, pode-se escrever que $E\{f[a(n)]f[b(n)]\} \approx E\{f[a(n)]\}E\{f[b(n)]\}$.

Os valores esperados em (4.2) são agora obtidos como segue:

Para se determinar $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$, considera-se que o i -ésimo elemento da diagonal da matriz $\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)$ é dado por $[x_i^2(n)][\sigma_i^2(n) + \varepsilon]^{-1}$, onde $\sigma_i^2(n)$ é obtida de (3.53). Fazendo-se a análise do i -ésimo elemento $[x_i^2(n)M][x_i^2(n) + x_i^2(n-1) + \dots + x_i^2(n-M+1) + M\varepsilon]^{-1}$, para um valor elevado de M , a influência do termo $x_i^2(n)$ sobre a parcela do denominador não se faz tão significativa em relação à influência que esse termo exerce sobre a parcela do numerador. Assim, pode-se dizer que os elementos da matriz $\mathbf{D}^{-1}(n)$ variam lentamente em relação aos elementos da matriz $\mathbf{x}_T(n)\mathbf{x}_T^T(n)$, permitindo utilizar H2, que conduz à seguinte expressão (ver Apêndice A). Assim,

$$E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)] \approx E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)] = E[\mathbf{D}^{-1}(n)]\mathbf{R}_T. \quad (4.3)$$

Similarmente,

$$E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)d(n)] \approx E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)d(n)] = E[\mathbf{D}^{-1}(n)]\mathbf{p}_T. \quad (4.4)$$

Utilizando H1 e substituindo-se (4.3) e (4.4) em (4.2), tem-se

$$E[\mathbf{w}_T(n+1)] = \{\mathbf{I} - 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}E[\mathbf{w}_T(n)] + 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{p}_T. \quad (4.5)$$

A derivação do momento de primeira ordem do vetor $\mathbf{w}_T(n)$ é concluída determinando-se o valor esperado $E[\mathbf{D}^{-1}(n)]$. Para isso, é proposto neste trabalho o procedimento apresentado a seguir.

Determinação do valor esperado $E[\mathbf{D}^{-1}(n)]$

Considerando um elemento qualquer da matriz $\mathbf{D}^{-1}(n)$ para $\varepsilon \ll 1$, tem-se a seguinte aproximação

$$\frac{1}{d_i(n)} = \frac{M}{\sum_{k=0}^{M-1} x_i^2(n-k) + \varepsilon M} \approx \frac{M}{\sum_{k=0}^{M-1} x_i^2(n-k)} - \varepsilon \frac{M^2}{\left(\sum_{k=0}^{M-1} x_i^2(n-k)\right)^2}. \quad (4.6)$$

Determinando-se o valor esperado em ambos os lados de (4.6), obtém-se

$$E\left[\frac{1}{d_i(n)}\right] \approx E\left[\frac{M}{\sum_{k=0}^{M-1} x_i^2(n-k)}\right] - \varepsilon E\left[\frac{M^2}{\left(\frac{1}{M} \sum_{k=0}^{M-1} x_i^2(n-k)\right)^2}\right]. \quad (4.7)$$

A solução existente para os valores esperados como este de (4.7) é obtida considerando que o processo $\{x_i(n)\}$ seja independente, Gaussiano e de média nula [22]-[23], [26]-[31] e [38]. Dessa forma, o tratamento matemático requerido é bastante simplificado. No entanto, para sequências correlacionadas $\{x_i(n)\}$ a consideração prévia não produz bons resultados. Neste trabalho de tese, não foi imposta qualquer restrição no que diz respeito às características estatística da sequência $\{x_i(n)\}$ (correlacionada ou não-correlacionada), que é a condição central para se obter um modelo estatístico mais acurado do algoritmo LMS-DT. Assim, o valor esperado de (4.7) é determinado usando uma função densidade de probabilidade multivariada Gaussiana como em [21] e [22], resultando na necessidade

de se computar uma integral hiperelíptica de alta ordem. A seguir, tem-se uma expressão geral que engloba os dois valores esperados requeridos por (4.7).

$$E \left\{ \frac{M^p}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^p} \right\} = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{M \text{ fold}} \frac{M^p}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^p} f(\mathbf{x}_i) d\mathbf{x}_i, \quad (4.8)$$

onde $f(\mathbf{x}_i)$ representa a função densidade de probabilidade multivariada Gaussiana do vetor $\mathbf{x}_i(n)$. Note que os valores esperados em (4.7) são obtidos para $p=1$ e $p=2$ em (4.8). Na literatura, a referida integral possui solução fechada para $M \leq 4$, denominada como integral elíptica [39]; em um caso mais geral para $M > 4$ ela recebe o nome de integral hiperelíptica ou Abeliana [39], apresentando solução apenas para certas condições especiais [21], [23], [26]-[31]. Aqui, fez-se uso de um método para resolver a integral hiperelíptica sem adotar quaisquer restrições sobre os dados considerados. Para obter uma solução para tal integral, observou-se que os pólos da função a ser integrada, à medida que a janela de observação M aumenta, poderiam se agrupados aos pares por apresentarem valores muito próximos. Sendo assim, a integral hiperelíptica se torna uma integral que pode ser facilmente calculada usando o método de frações parciais. Esse procedimento adotado para a obtenção de uma expressão fechada para a referida integral, apesar de envolver uma aproximação, apresenta resultados bastante satisfatórios. Através de simulações, comparando resultados para o valor esperado $E[\mathbf{D}^{-1}(n)]$ obtidos via simulação Monte Carlo e os obtidos a partir do modelo, constatou-se que o erro entre eles é da ordem 10^{-4} e diminui à medida que a janela de observação M aumenta. Detalhes do método proposto são apresentados no Apêndice B. Usando tal método, os valores esperados em (4.7) podem ser aproximados por

$$E \left[\frac{M}{\sum_{k=0}^{M-1} x_i^2(n-k)} \right] \approx \frac{-M}{\sqrt{a_M}} \sum_{q=1}^{M/2} A_q \ln(-\omega'_q), \quad (4.9)$$

e

$$E \left\{ \frac{M^2}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^2} \right\} \approx \frac{M^2}{\sqrt{a_M}} \sum_{q=1}^{M/2} A_q [-\omega'_q \ln(-\omega'_q) + \omega'_q]. \quad (4.10)$$

Em (4.9) e (4.10), os parâmetros ω'_q , a_M , e A_q estão definidos em detalhes no Apêndice B.

4.1.2- REGIME PERMANENTE DE $\mathbf{w}_T(n)$

O valor em regime permanente do vetor de coeficientes do filtro adaptativo é determinado de $\mathbf{w}_\infty = \lim_{n \rightarrow \infty} E[\mathbf{w}_T(n)]$. Assumindo-se que (4.5) convirja, verifica-se que $\lim_{n \rightarrow \infty} E[\mathbf{w}_T(n+1)] = \lim_{n \rightarrow \infty} E[\mathbf{w}_T(n)] = \mathbf{w}_\infty$. De (4.5) e após algumas manipulações matemáticas, obtém-se

$$\mathbf{w}_\infty = \mathbf{R}_T^{-1} \mathbf{p}_T. \quad (4.11)$$

Comparando-se (4.11) com a expressão equivalente apresentada em [2], pode-se ver que a constante ε não provoca qualquer alteração no valor em regime permanente do vetor de coeficientes. Assim, no comportamento médio, os coeficientes do filtro LMS-DT convergem para a conhecida solução de Wiener.

4.1.3- CURVA DE APRENDIZAGEM E MOMENTO DE SEGUNDA ORDEM DO VETOR $\mathbf{v}_T(n)$

A curva de aprendizagem é definida como a evolução ao longo do tempo do erro quadrático médio. Para facilitar a obtenção de tal curva é necessário definir o vetor de erro dos coeficientes como $\mathbf{v}_T(n) = \mathbf{w}_T(n) - \mathbf{w}_{T_{\text{ótimo}}}$. Assim, tem-se o sinal de erro expresso como

$$\begin{aligned}
e(n) &= d(n) - \mathbf{x}_T^T(n) \mathbf{v}_T(n) - \mathbf{x}_T^T(n) \mathbf{w}_{T_{\text{ótimo}}} + z(n) \\
&= e_o(n) - \mathbf{x}_T^T(n) \mathbf{v}_T(n),
\end{aligned} \tag{4.12}$$

onde $e_o(n)$, o erro de estimativa, é dado por

$$e_o(n) = d(n) - \mathbf{x}_T^T(n) \mathbf{w}_{T_{\text{ótimo}}} + z(n). \tag{4.13}$$

Elevando-se ao quadrado ambos os lados de (4.12) e calculando-se o valor esperado da expressão resultante, obtém-se

$$E[e^2(n)] = E[e_o^2(n)] - 2E[e_o(n) \mathbf{x}_T^T(n) \mathbf{v}_T(n)] + E[\mathbf{v}_T^T(n) \mathbf{x}_T(n) \mathbf{x}_T^T(n) \mathbf{v}_T(n)]. \tag{4.14}$$

Pelo Princípio da Ortogonalidade $E[e_o(n) \mathbf{x}_T^T(n)] = 0$ [1]; usando-se então H1 e relembrando que $z(n)$ é um sinal descorrelacionado de qualquer outro sinal envolvido no processo (Seção 4.1.1), a curva de aprendizagem é dada por

$$\begin{aligned}
E[e^2(n)] &= e_{\min} + E[\mathbf{v}_T^T(n) \mathbf{x}_T(n) \mathbf{x}_T^T(n) \mathbf{v}_T(n)] \\
&= e_{\min} + \text{tr}\{\mathbf{R}_T E[\mathbf{v}_T(n) \mathbf{v}_T^T(n)]\},
\end{aligned} \tag{4.15}$$

onde $e_{\min} = E[e_o^2(n)]$ caracteriza o mínimo erro em regime permanente.

Note que (4.15) é completamente determinada se o momento de segunda ordem $E[\mathbf{v}_T(n) \mathbf{v}_T^T(n)]$ for conhecido. Tal momento é também denominado matriz de covariância do vetor de erro $\mathbf{K}(n)$, ou seja, $\mathbf{K}(n) = E[\mathbf{v}_T(n) \mathbf{v}_T^T(n)]$. Para obter $\mathbf{v}_T(n)$, subtrai-se $\mathbf{w}_{T_{\text{ótimo}}}$ de ambos os lados de (4.1) e então se determina o produto externo $\mathbf{v}_T(n) \mathbf{v}_T^T(n)$. Agora, tomando-se o valor esperado de ambos os lados da expressão resultante, obtém-se

$$\begin{aligned}
\mathbf{K}(n+1) = & \overbrace{E\{[\mathbf{I} - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n)\mathbf{v}_T^T(n)[\mathbf{I} - 2\mu\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]\}}^{\text{A}} \\
& + \overbrace{E\{[\mathbf{I} - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n)2\mu\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)[e_o(n) + z(n)]\}}^{\text{B}} \\
& + \overbrace{E\{2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)[z(n) + e_o(n)]\mathbf{v}_T^T(n)[\mathbf{I} - 2\mu\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]\}}^{\text{C}} \\
& + \overbrace{2\mu E\{2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)[z(n) + e_o(n)]\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)[z(n) + e_o(n)]\}}^{\text{D}}.
\end{aligned} \tag{4.16}$$

Em (4.16), a determinação do valor esperado dos termos que contêm $\mathbf{D}^{-1}(n)$ merecem especial atenção. Usando-se então as considerações H1 e H2, os termos A, B, C, e D são obtidos como segue:

$$\begin{aligned}
\text{Termo A: } & E\{[\mathbf{I} - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n)\mathbf{v}_T^T(n)[\mathbf{I} - 2\mu\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]\} \\
& = E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)] - 2\mu E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\
& \quad - 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)] \\
& \quad + 4\mu^2 \underbrace{E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]}_{\text{A1}} \\
& = \mathbf{K}(n) - 2\mu \mathbf{K}(n)\mathbf{R}_T E[\mathbf{D}^{-1}(n)] - 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T \mathbf{K}(n) \\
& \quad + 4\mu^2 \underbrace{E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]}_{\text{A1}}
\end{aligned} \tag{4.17}$$

Adotando agora o teorema da fatoração de momentos de quarta ordem de variáveis aleatórias Gaussianas [1], tem-se o valor esperado A1 de (4.17) dado por

$$\begin{aligned}
& E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\
& \cong E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{D}^{-1}(n)] \\
& = E[\mathbf{D}^{-1}(n)]\{2E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)] \\
& \quad + E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\text{tr}\{E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]\}\}E[\mathbf{D}^{-1}(n)] \\
& = E[\mathbf{D}^{-1}(n)]\{2\mathbf{R}_T \mathbf{K}(n)\mathbf{R}_T + \mathbf{R}_T \text{tr}[\mathbf{R}_T \mathbf{K}(n)]\}E[\mathbf{D}^{-1}(n)],
\end{aligned} \tag{4.18}$$

$$\begin{aligned}
\text{Termo B: } & E\{[\mathbf{I} - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]2\mu\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)[z(n) + e_o(n)]\} \\
& = 2\mu E[\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)z(n) + \mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)e_o(n)] \\
& \quad - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)z(n) \\
& \quad - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)e_o(n) \\
& = 0.
\end{aligned} \tag{4.19}$$

$$\begin{aligned}
\text{Termo C: } & E\{2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)[z(n) + e_o(n)][\mathbf{v}_T^T(n) - 2\mu\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]\} \\
& = 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n) - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n) \\
& \quad + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)e_o(n)\mathbf{v}_T^T(n) - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)e_o(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\
& = 0.
\end{aligned} \tag{4.20}$$

$$\begin{aligned}
\text{Termo D: } & E\{2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)2\mu\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)[z^2(n) + 2z(n)e_o(n) + e_o^2(n)]\} \\
& = E[4\mu^2\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)z^2(n) + 4\mu^2\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)2z(n)e_o(n) \\
& \quad + 4\mu^2\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)e_o^2(n)] \\
& = 4\mu^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_T E[\mathbf{D}^{-1}(n)]\sigma_z^2 + 4\mu^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_T E[\mathbf{D}^{-1}(n)]E[e_o^2(n)].
\end{aligned} \tag{4.21}$$

Substituindo-se (4.17) até (4.21) em (4.16), a expressão recursiva para a matriz de covariância do vetor de erro dos coeficientes é

$$\begin{aligned}
\mathbf{K}(n+1) & = \mathbf{K}(n) - 2\mu\mathbf{K}(n)\mathbf{R}_T E[\mathbf{D}^{-1}(n)] - 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T \mathbf{K}(n) \\
& \quad + 4\mu^2 E[\mathbf{D}^{-1}(n)]\{2\mathbf{R}_T \mathbf{K}(n)\mathbf{R}_T + \mathbf{R}_T \text{tr}[\mathbf{R}_T \mathbf{K}(n)]\} E[\mathbf{D}^{-1}(n)] \\
& \quad + 4\mu^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_T E[\mathbf{D}^{-1}(n)]e_{\min}.
\end{aligned} \tag{4.22}$$

onde a solução da expressão $E[\mathbf{D}^{-1}(n)]$ já foi apresentada na Seção 4.1.1 (ver detalhes no Apêndice B).

4.1.4- LIMITE SUPERIOR PARA O PASSO DE ADAPTAÇÃO

O limite superior para o passo de adaptação representa um parâmetro importante de análise do desempenho do algoritmo no que se refere à sua estabilidade, por se tratar de um valor de passo para o qual o algoritmo obtém máxima convergência sem apresentar um comportamento instável. A seguir, mostra-se o procedimento usado para determinar o limite superior de passo de adaptação.

Sejam as seguintes matrizes $\mathbf{C}(n) = \mathbf{Q}^T \mathbf{K}(n) \mathbf{Q}$ e $\mathbf{\Lambda}_T = \mathbf{Q}^T \mathbf{R}_T \mathbf{Q}$, onde $\mathbf{\Lambda}_T$ é uma matriz diagonal que contém os autovalores λ_i de \mathbf{R}_T e \mathbf{Q} é a matriz ortogonal que contém os autovetores de \mathbf{R}_T . Assim, pode-se escrever (4.22) como segue:

$$\begin{aligned} \mathbf{C}(n+1) = & \mathbf{C}(n) - 2\mu \{ \mathbf{C}(n) \mathbf{\Lambda}_T \mathbf{B} + \mathbf{B} \mathbf{\Lambda}_T \mathbf{C}(n) \} \\ & + 4\mu^2 \mathbf{B} \{ 2\mathbf{\Lambda}_T \mathbf{C}(n) \mathbf{\Lambda}_T + \mathbf{\Lambda}_T \text{tr}[\mathbf{\Lambda}_T \mathbf{C}(n)] \} \mathbf{B} \\ & + 4\mu^2 \mathbf{B} \mathbf{\Lambda}_T \mathbf{B} e_{\min}, \end{aligned} \quad (4.23)$$

onde $\mathbf{B} = \mathbf{Q}^T E[\mathbf{D}^{-1}(n)] \mathbf{Q}$ (vamos considerar na análise a seguir que \mathbf{B} é uma matriz diagonal. Essa consideração pode ser feita, uma vez que os elementos fora da diagonal são muito pequenos). De [40], tem-se que $c_{ij}^2(n) \leq c_{ii}(n) c_{jj}(n)$, com $c_{ij}(n)$ denotando os elementos da matriz $\mathbf{C}(n)$. A convergência de (4.23) é determinada do comportamento dos elementos da diagonal de $\mathbf{C}(n)$. Definindo-se os vetores $\mathbf{c}(n) = [c_{11}(n) \ c_{22}(n) \ c_{33}(n) \ \dots \ c_{NN}(n)]^T$ e $\boldsymbol{\lambda} = [\lambda_0 \ \lambda_1 \ \dots \ \lambda_{N-1}]^T$, onde $\mathbf{c}(n)$ é um vetor constituído pelos elementos da diagonal da matriz $\mathbf{C}(n)$, pode-se escrever a seguinte relação:

$$\text{diag} \{ \mathbf{B} \mathbf{\Lambda}_T \text{tr}[\mathbf{\Lambda}_T \mathbf{C}(n)] \mathbf{B} \} = \mathbf{B}^2 \boldsymbol{\lambda} \boldsymbol{\lambda}^T \mathbf{c}(n). \quad (4.24)$$

Expressando-se (4.23) em termos de $\mathbf{c}(n)$, obtém-se

$$\mathbf{c}(n+1) = \mathbf{F} \mathbf{c}(n) + 4\mu^2 e_{\min} \mathbf{B}^2 \boldsymbol{\lambda}, \quad (4.25)$$

onde

$$\mathbf{F} = \mathbf{I} - 4\mu \mathbf{B} \mathbf{\Lambda}_T + 8\mu^2 \mathbf{B}^2 \mathbf{\Lambda}_T^2 + 4\mu^2 \mathbf{B}^2 \boldsymbol{\lambda} \boldsymbol{\lambda}^T. \quad (4.26)$$

A condição necessária e suficiente para convergência de (4.25) depende das magnitudes dos autovalores da matriz \mathbf{F} , que devem ser menores do que a unidade. Pode-se forçar tal condição através de uma adequada seleção do parâmetro μ . Usando-se

$\lambda'_j = b_j \lambda_j$ e seguindo o mesmo procedimento usado em [40], as magnitudes dos autovalores de (4.26) tornam-se menores do que a unidade se o passo de adaptação μ satisfizer as seguintes condições necessárias e suficientes:

$$0 < \mu < \frac{1}{2b_j \lambda_j}, \quad j = 1, 2, \dots, N, \quad (4.27)$$

e

$$\sum_{j=1}^N \frac{\mu b_j \lambda_j}{1 - 2\mu b_j \lambda_j} < 1. \quad (4.28)$$

A partir de (4.27) e (4.28), obtém-se o passo máximo. O passo máximo deverá satisfazer a seguinte condição:

$$0 < \mu_{\max} \leq \frac{k}{2(b_j \lambda_j)_{\max}}, \quad (4.29)$$

onde k é determinado substituindo em (4.28) μ por μ_{\max} , obtido de (4.29), como segue:

$$\sum_{j=1}^N \left(\frac{\frac{k}{(b_j \lambda_j)_{\max}} b_j \lambda_j}{1 - 2 \frac{k}{(b_j \lambda_j)_{\max}} b_j \lambda_j} \right) = 1. \quad (4.30)$$

As Figs. 4.1 e 4.2 mostram o limite superior do passo de adaptação obtido de (4.29) para diversos valores de $N = \{2, 3, \dots, 32\}$ e $M = \{32, 64\}$, usando-se uma transformação DCT [2], $\varepsilon = 0,001$ e $\sigma_x^2 = 1$. Para se validar (4.29), são plotados alguns pontos [marcados como (*)] obtidos de simulações. O passo máximo experimental é escolhido como sendo aquele que ainda assegure a convergência do algoritmo para apenas uma realização (rodada). A Fig. 4.1 ilustra um caso de sinal de entrada colorido. O sinal correlacionado é obtido de um

processo AR(2) com coeficientes $a_1 = 0,1833$ e $a_2 = -0,85$, e dispersão dos autovalores da matriz de autocorrelação de 5,22 para $N = 2$ e 530 para $N = 32$, $\varepsilon = 0,001$, e $\sigma_x^2 = 1$.

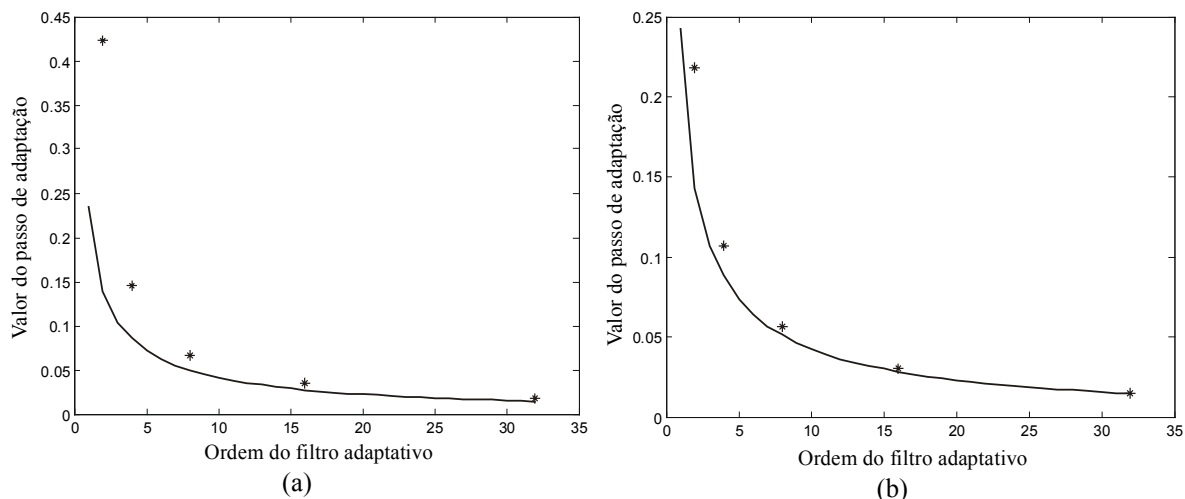


Figura 4.1- Limite superior para passo de adaptação considerando-se um sinal de entrada correlacionado para diversos valores de N e M . (*) Experimentalmente obtidos. (Linhas sólidas) predição (4.29). (a) $M = 32$, (b) $M = 64$.

A partir dos resultados apresentados nas Figuras 4.1(a) e (b), quanto maior o valor da janela de observação M melhor é a concordância entre os resultados obtidos pelo modelo e por simulação. Isso se deve ao fato de que, nessas condições, o Princípio da Média se torna cada vez mais preciso. Além disso, para valores de N grandes ($N > 32$), os resultados também são satisfatórios. Isso é devido ao fato de que para esses casos o processo de decorrelação torna-se mais eficiente, ou seja, o sinal de entrada tende para um sinal “branco”.

4.1.5- DESAJUSTE

O desajuste é um parâmetro adimensional que mede o grau de diferença entre o erro obtido pelo algoritmo adaptativo e o erro obtido através da solução de Wiener devido ao uso da estimativa do gradiente. A seguir é apresentado o procedimento para o cálculo do desajuste.

A curva de aprendizagem do algoritmo LMS-DT é dada por

$$E[e^2(n)] = e_{\min} + \text{tr}\{\mathbf{R}_T E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]\}. \quad (4.31)$$

Uma vez que $\text{tr}\{\mathbf{R}_T E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]\}$ envolve somente termos da diagonal, pode-se rescrever a (4.31) como

$$E[e^2(n)] = e_{\min} + \boldsymbol{\lambda}^T \mathbf{c}(n). \quad (4.32)$$

Assumindo-se a convergência do algoritmo adaptativo, o erro quadrático médio (EQM) em regime permanente é dado por

$$E[e^2(\infty)] = e_{\min} + \boldsymbol{\lambda}^T \mathbf{c}(\infty), \quad (4.33)$$

Determinando-se o valor de (4.25) quando $n \rightarrow \infty$, obtém-se

$$\mathbf{c}(\infty) = 4\mu^2 e_{\min} (\mathbf{I} - \mathbf{F})^{-1} \mathbf{B}^2 \boldsymbol{\lambda}. \quad (4.34)$$

Em (4.33), o termo e_{\min} corresponde ao mínimo EQM; assim, $\boldsymbol{\lambda}^T \mathbf{c}(\infty)$ representa o EQM em excesso. O desajuste é então definido como

$$\mathcal{M} = \frac{E[e^2(\infty)] - e_{\min}}{e_{\min}}. \quad (4.35)$$

Substituindo-se (4.33) e (4.34) em (4.35), obtém-se

$$\mathcal{M} = 4\mu^2 \boldsymbol{\lambda}^T (\mathbf{I} - \mathbf{F})^{-1} \mathbf{B}^2 \boldsymbol{\lambda}. \quad (4.36)$$

Novamente, seguindo [40] com $\lambda'_j = b_j \lambda_j$, tem-se

$$\mathcal{M} = \frac{\sum_{i=1}^N \frac{\mu b_i \lambda_i}{(1 - 2\mu b_i \lambda_i)}}{1 - \sum_{i=1}^N \frac{\mu b_i \lambda_i}{(1 - 2\mu b_i \lambda_i)}}. \quad (4.37)$$

A precisão de (4.37) é verificada nas tabelas apresentadas no Capítulo de Resultados de Simulações, para o caso estacionário.

4.2- CONCLUSÕES

Neste capítulo foi apresentada uma análise estatística para o algoritmo LMS-DT. Para se obter tal análise, algumas hipóteses simplificativas foram adotadas acerca do sinal de entrada ou a respeito do passo de adaptação. Em relação ao sinal de entrada, uma consideração foi quanto a sua distribuição de probabilidade. A distribuição adotada foi a Gaussiana, em virtude de simplificar bastante muitas expressões de valores esperados. Com respeito à adaptação, para obter o modelo adotou-se a adaptação lenta. É importante ressaltar que tal análise é independente do tipo de transformação ortogonal utilizada.

CAPÍTULO 5

RESULTADOS DE SIMULAÇÕES: CASO ESTACIONÁRIO

Para avaliar a qualidade do modelo estatístico obtido para o algoritmo LMS-DT, alguns exemplos são apresentados, considerando-se um problema de identificação de sistemas. Para tal, utilizou-se um sinal de entrada correlacionado, obtido a partir de um processo auto-regressivo de segunda ordem AR(2), dado por

$$x(n) = v(n) - \alpha_1 x(n-1) - \alpha_2 x(n-2), \quad (5.1)$$

onde $v(n)$ é um ruído branco com variância σ_v^2 . Os coeficientes AR são $a_1 = 0,1833$ e $a_2 = -0,85$. O ruído de medição $z(n)$ possui uma variância $\sigma_z^2 = 10^{-4}$ (SNR = 40 dB). A constante ε , considerada em (3.54), é igual a 0,001. As simulações Monte Carlo (MC) são obtidas para uma média de 500 realizações independentes. Para todos os exemplos, os valores de passo de adaptação usados são $0,1\mu_{\max}$ e $0,3\mu_{\max}$, onde μ_{\max} (valor máximo de passo de adaptação) é obtido de (4.29).

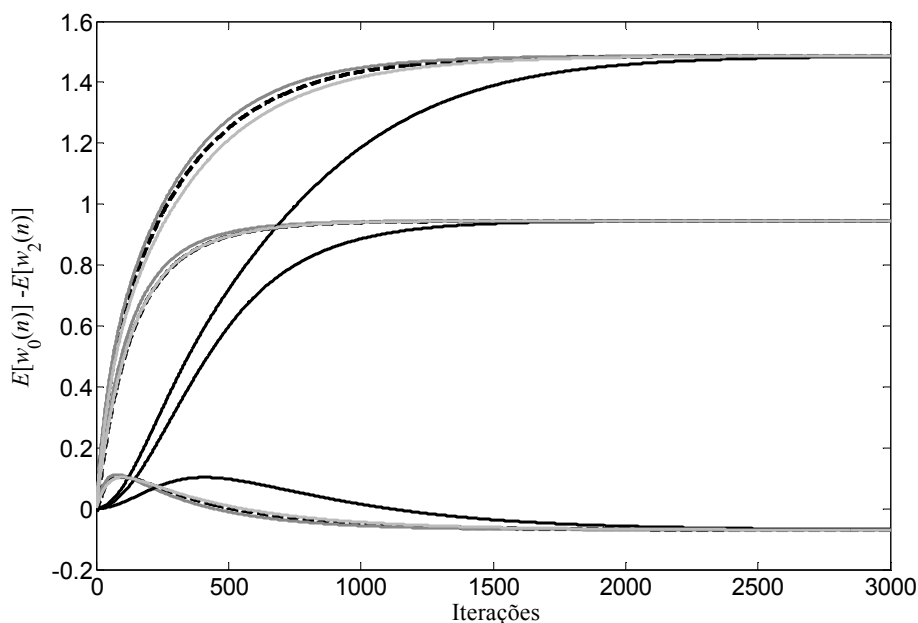
É importante aqui ressaltar que o modelo proposto deveria ser confrontado, para efeito de comparação de desempenho, com os modelos discutidos em [26]-[29]. Entretanto, os modelos apresentados em [28] e [29] são bastante similares em filosofia e comportamento, e o modelo fornecido em [27] é apenas uma versão simplificada do modelo discutido em [28], pois o primeiro não considera o parâmetro de regularização ε . Portanto,

apenas os modelos dados em [26] e [28] serão considerados para verificação entre os modelos estatísticos do algoritmo LMS-DT avaliado neste trabalho de tese.

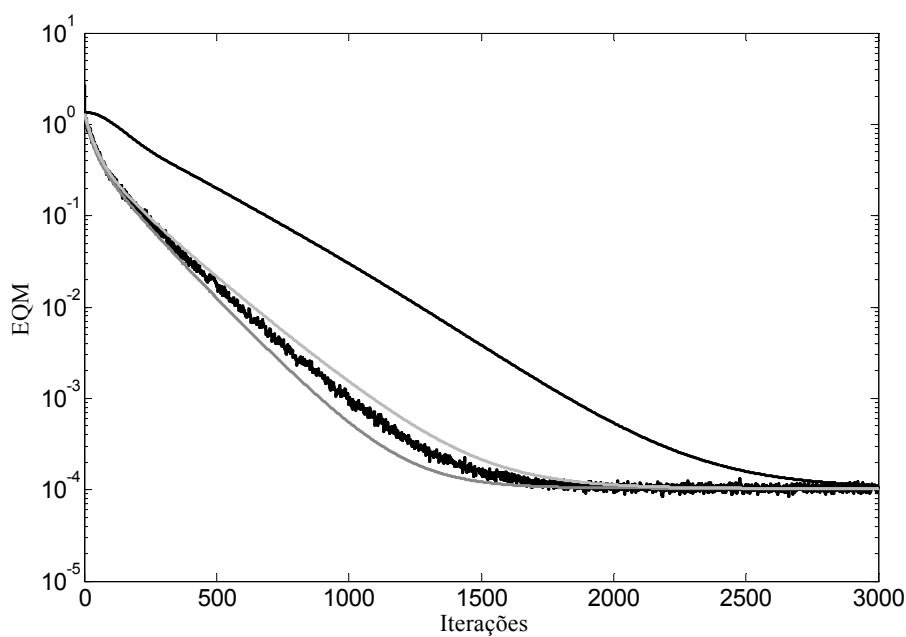
5.1- *EXEMPLO 1*

As condições para este exemplo são: sinal de entrada correlacionado com dispersão dos autovalores da matriz de autocorrelação do sinal de entrada igual a $\chi = 81$; $N = 8$ e $M = 32$. A transformação ortogonal usada é a DCT e os coeficientes da planta são obtidos através da função $\text{sinc}(i/8)$ para $i = 1, 2, \dots, 8$. O passo de adaptação máximo obtido de (4.29) é $\mu_{\max} = 0,054$. A Fig. 5.1 mostra os resultados de simulação MC, as predições teóricas obtidas a partir de [26] e [28] e as obtidas através do modelo proposto [equações (4.5), (4.15) e (4.22)]. Para os resultados dessa figura, considera-se um valor de passo de adaptação igual a $\mu_{\max}/15$. Da Fig. 5.1, pode-se observar que o modelo proposto em [26] não prediz adequadamente o comportamento médio dos coeficientes e nem a curva de EQM do algoritmo adaptativo em estudo para o regime transiente. Por outro lado, o modelo proposto em [28] exibe um comportamento muito melhor do que aquele dado pelo modelo proposto em [26]. Sendo assim, o modelo de [26] não será mais incluído na comparação para os exemplos subseqüentes considerados nesta tese. Com respeito ao modelo aqui proposto, pode ser observado que ele é mais preciso do que o modelo obtido em [28].

A Fig. 5.2 compara o desempenho do modelo proposto para dois valores de μ ($0,1\mu_{\max}$ e $0,3\mu_{\max}$). Dessa figura, pode-se observar uma boa concordância entre os resultados obtidos via simulação MC e aqueles obtidos a partir do modelo proposto.

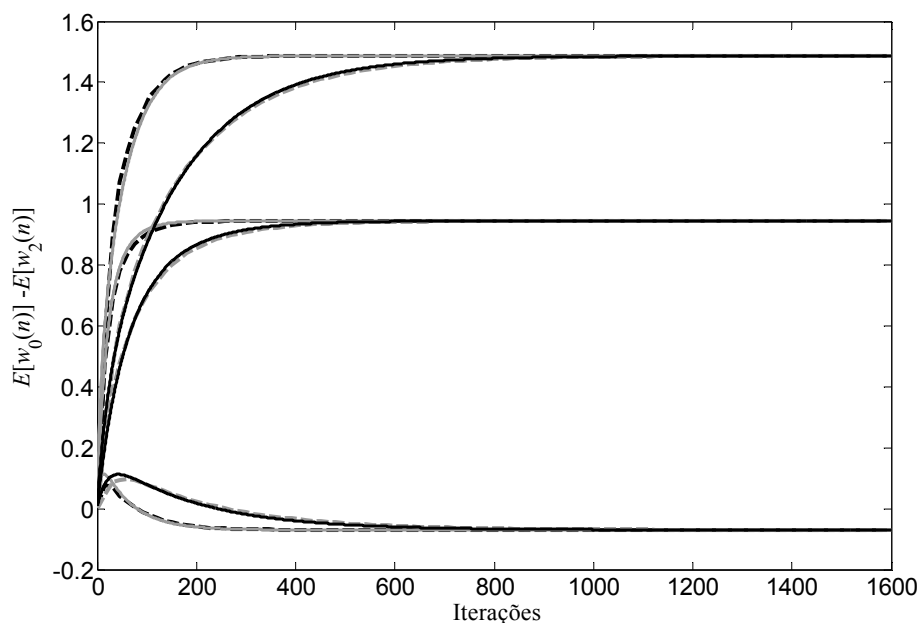


(a)

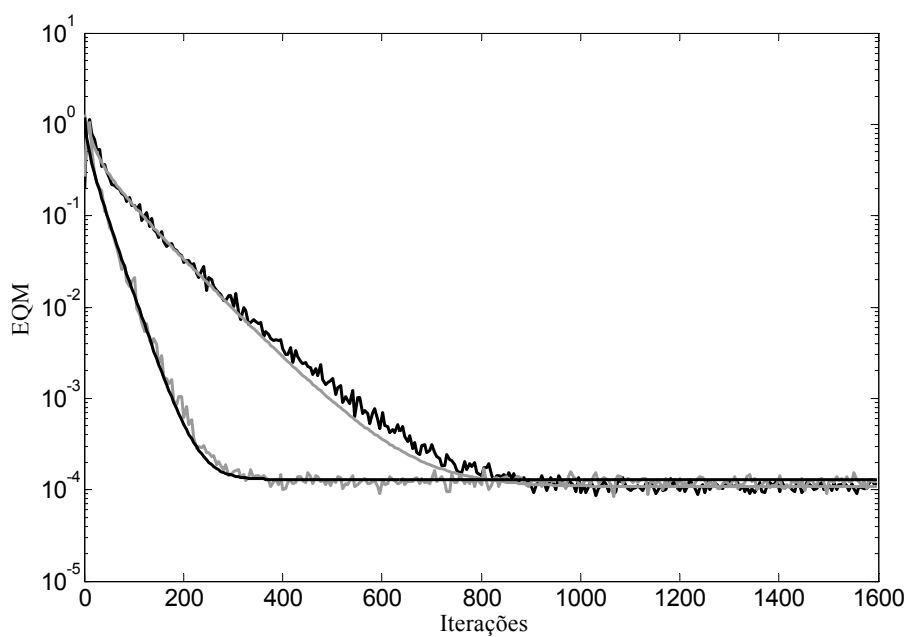


(b)

Figura 5.1- Exemplo 1. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada cinza) simulação MC de [26]. (Linha contínua preta) modelo de [26]. (Linha tracejada preta) simulação MC. (Linha contínua cinza-escuro) modelo proposto. (Linha contínua cinza-claro) modelo de [28]. (b) Curvas de EQM. (Linha irregular cinza) simulação MC de [26]. (Linha contínua preta) modelo de [26]. (Linha irregular preta) simulação MC. (Linha contínua cinza-claro) modelo de [28]. (Linha contínua cinza-escuro) modelo proposto.



(a)



(b)

Figura 5.2- Exemplo 1. Desempenho do modelo proposto para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\max}$.

5.2- EXEMPLO 2

Este exemplo considera um sinal de entrada correlacionado com $\chi = 413$, $N = 8$ e $M = 64$. Os coeficientes da planta são obtidos da função $\text{sinc}(i/8)$ com $i = 1, 2, \dots, 8$. Para esse exemplo, o valor máximo de passo de adaptação é $\mu_{\max} = 0,0794$. A Fig. 5.3 ilustra uma comparação entre os resultados obtidos via simulação MC, os obtidos pelo modelo apresentado em [28] e os obtidos com o modelo proposto, para $0,1\mu_{\max}$. Por sua vez, a Fig. 5.4 ilustra apenas os resultados de simulações MC e os obtidos com o modelo proposto [dados por (4.5), (4.15) e (4.22)] para $0,1\mu_{\max}$. Para esse exemplo, pode ser novamente observada uma boa concordância entre os resultados de simulação e obtidos com o modelo proposto.

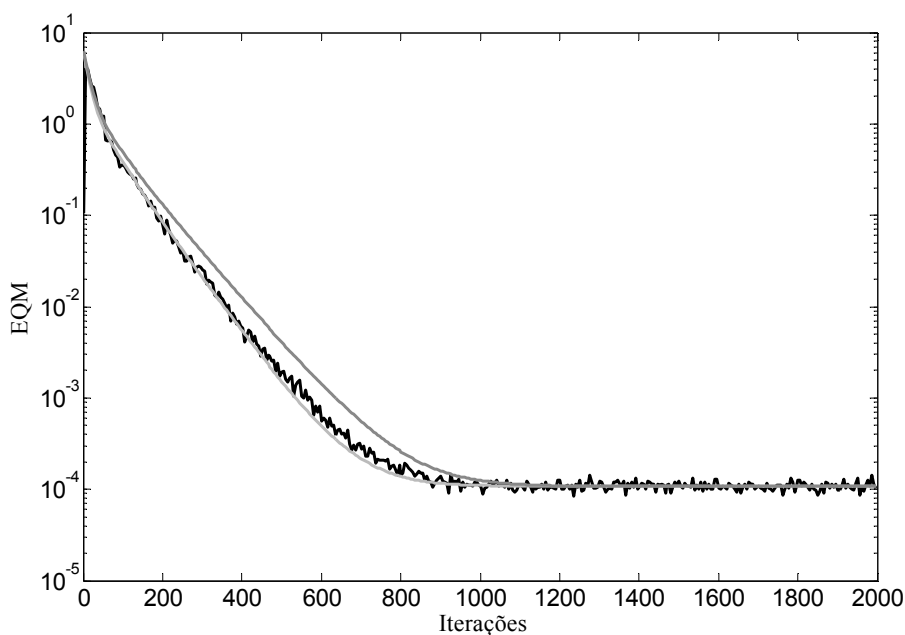
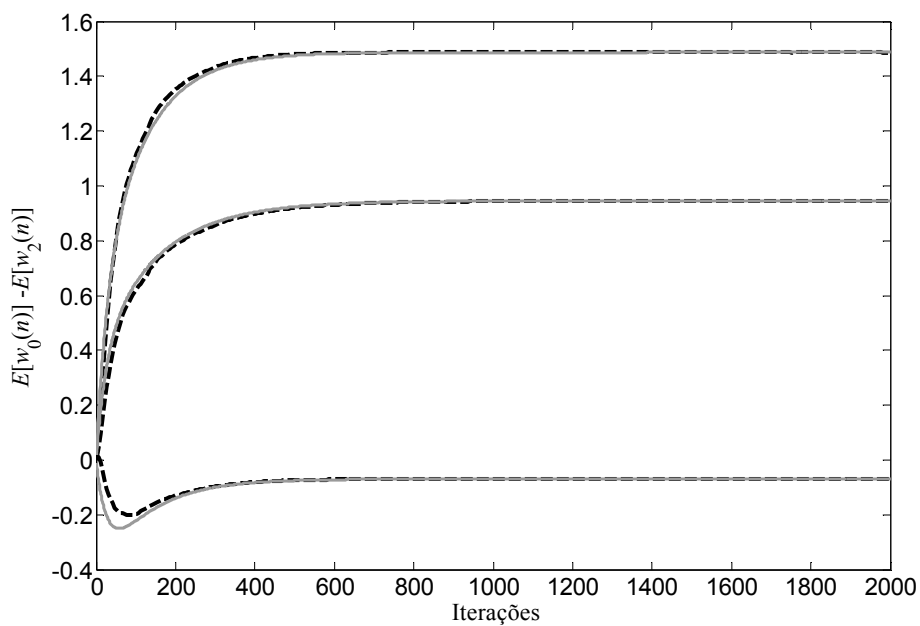
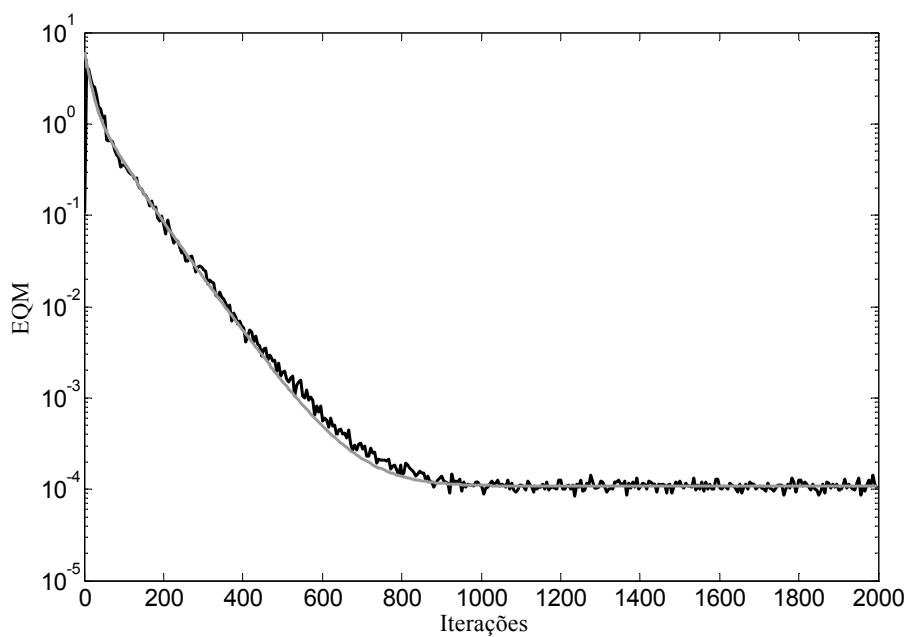


Figura 5.3- Exemplo 2. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 413$. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto.



(a)

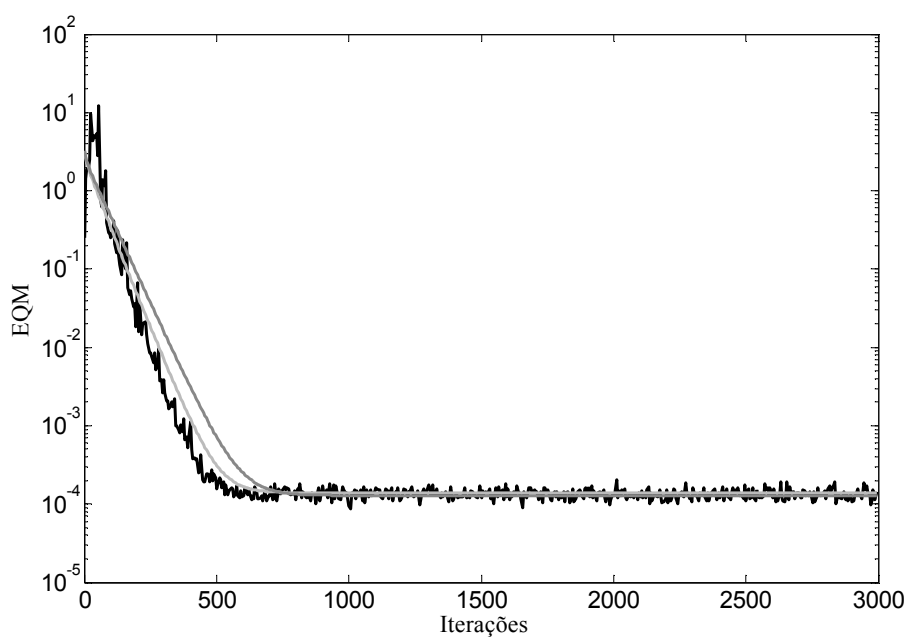


(b)

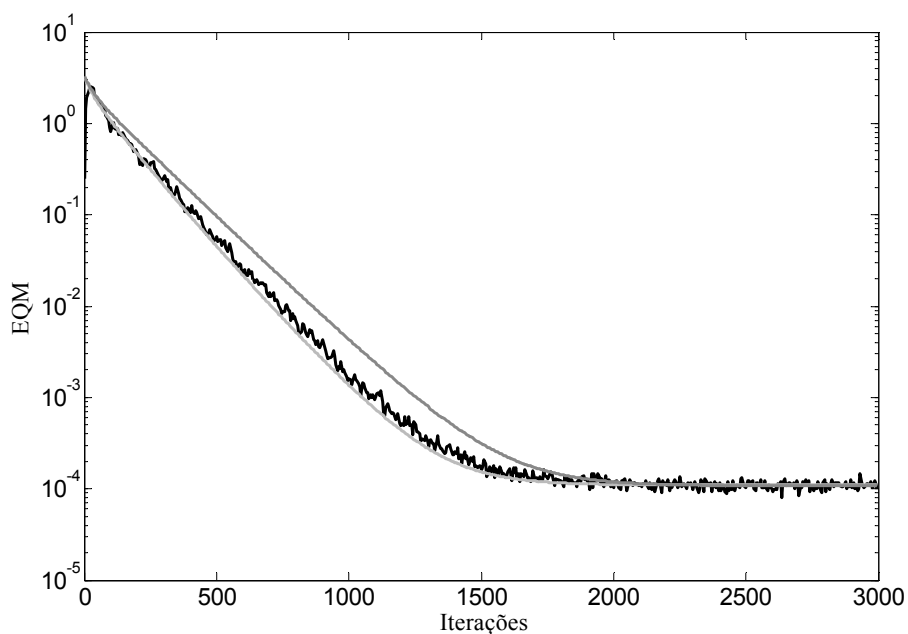
Figura 5.4- Exemplo 2. Desempenho do modelo proposto para $0,1\mu_{\max}$ e sinal de entrada correlacionado com $\chi = 413$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto. (b) Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto.

5.3- EXEMPLO 3

Os coeficientes da planta são obtidos da função $\text{sinc}(i/32)$ com $i = 1, 2, \dots, 32$. O sinal de entrada é correlacionado com dispersão dos autovalores $\chi = 220$. O tamanho do filtro adaptativo é $N = 32$ e o tamanho da janela de observação $M = 64$. A transformação ortogonal usada é a DCT, e $\mu_{\max} = 0,0221$. A Fig. 5.5 mostra uma comparação entre os resultados do EQM obtidos através de simulação MC, os obtidos com o modelo apresentado em [28] e os resultados do modelo proposto para $0,1\mu_{\max}$ e $0,3\mu_{\max}$. As Fig. 5.6(a) e 5.6(b) mostram os resultados obtidos a partir da simulação MC e do modelo proposto, considerando-se $0,1\mu_{\max}$ e $0,3\mu_{\max}$, para o comportamento médio dos coeficientes do filtro adaptativo e a curva de aprendizagem, respectivamente. Novamente, uma boa concordância entre simulação MC e modelo proposto pode ser verificada, principalmente para condição de adaptação lenta.

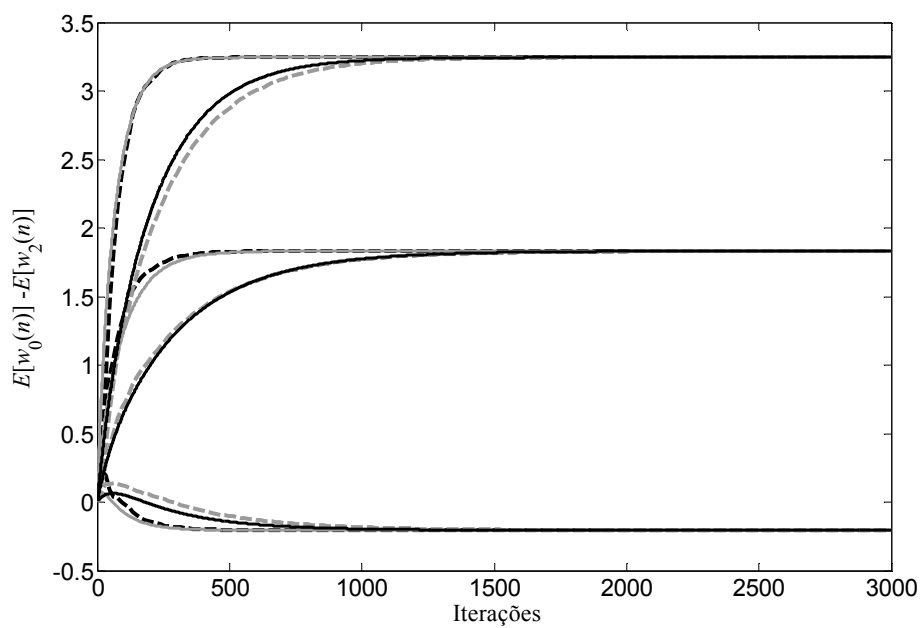


(a)



(b)

Figura 5.5- Exemplo 3. Comparações entre modelos para sinal de entrada correlacionado com $\chi = 220$. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto. (a) $0,3\mu_{\max}$ e (b) $0,1\mu_{\max}$.



(a)

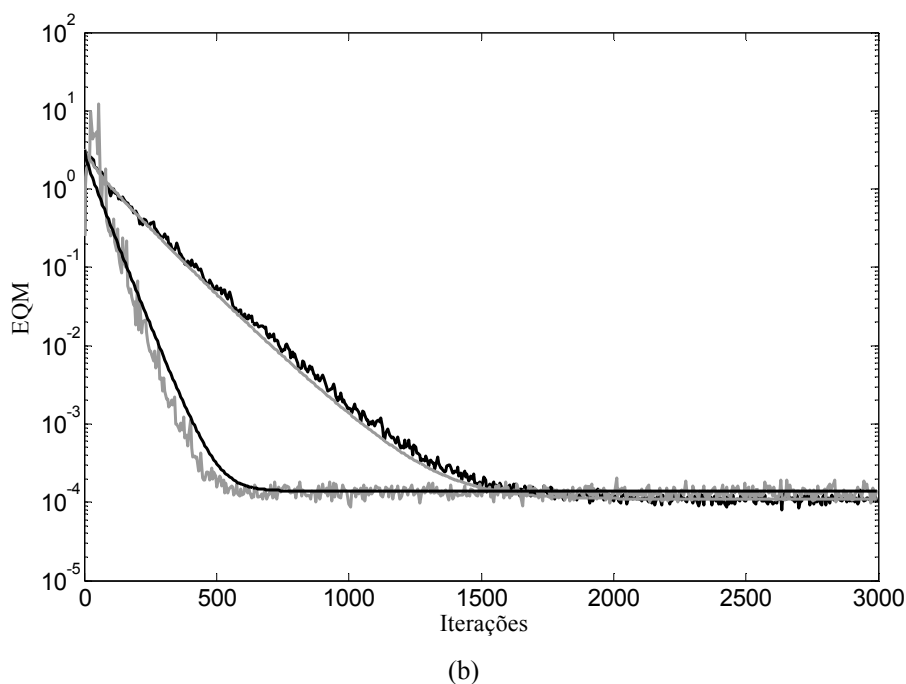
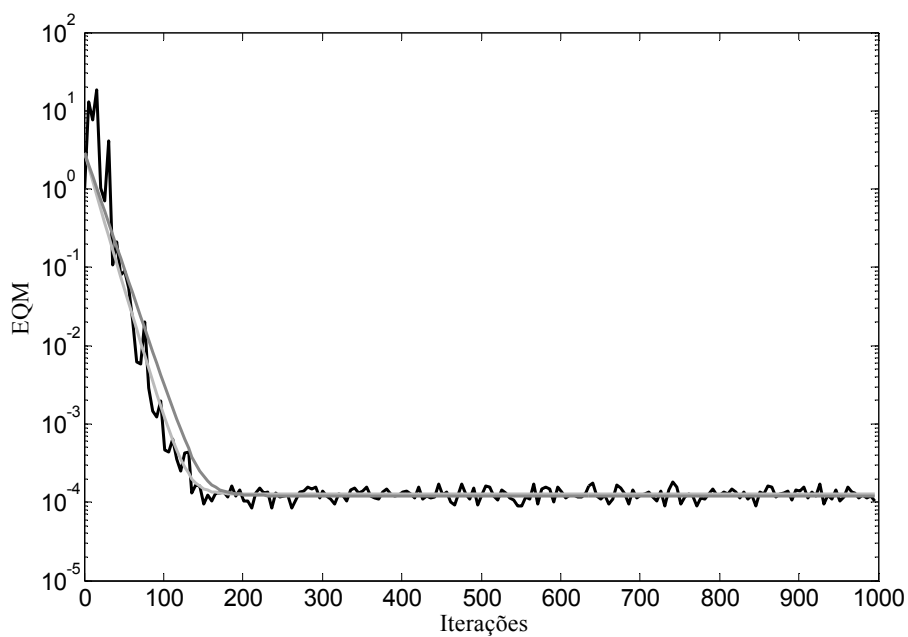


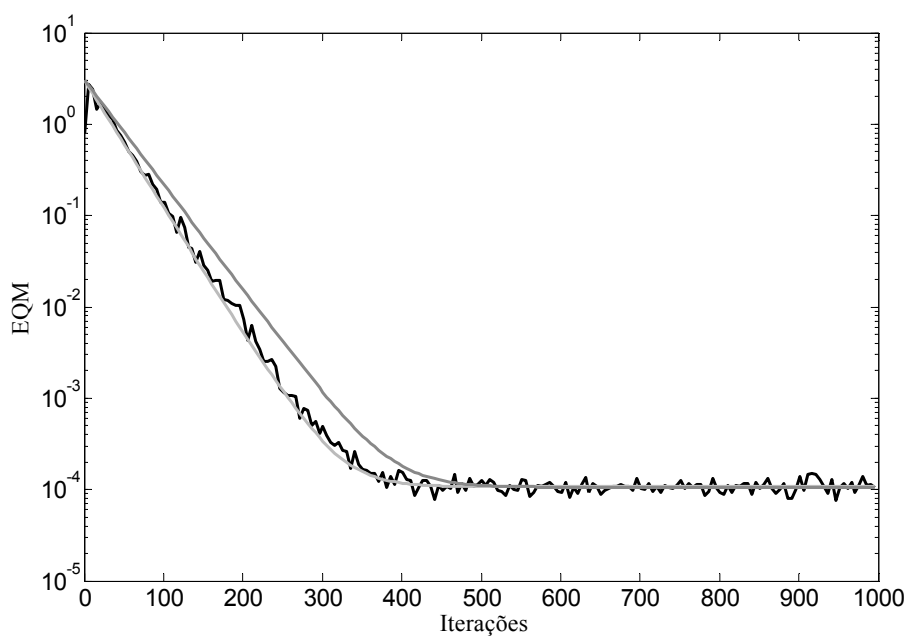
Figura 5.6- Exemplo 3. Desempenho do modelo proposto para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ com sinal de entrada correlacionado com $\chi = 220$. (a) Comportamento médio para os coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\max}$.

5.4- EXEMPLO 4

Neste exemplo, são utilizados os mesmos dados do Exemplo 1; entretanto, o sinal de entrada é um ruído branco. A transformação ortogonal é a DCT. Na Fig. 5.7 são mostrados os resultados do EQM obtidos através de simulação MC, os obtidos com o modelo apresentado em [28] e os resultados obtidos com o modelo proposto para $0,1\mu_{\max}$ e $0,3\mu_{\max}$. As Fig. 5.8(a) e 5.8(b) ilustram a simulação MC e o modelo proposto, considerando-se $0,1\mu_{\max}$ e $0,3\mu_{\max}$, para o comportamento médio dos coeficientes do filtro adaptativo e a curva de aprendizagem, respectivamente. Novamente, o modelo proposto apresenta uma boa precisão.

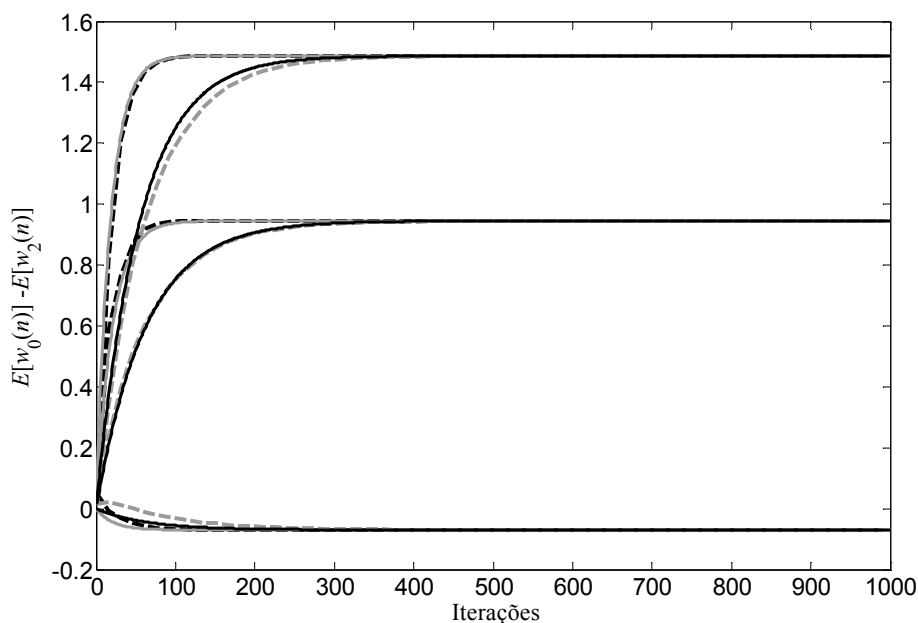


(a)

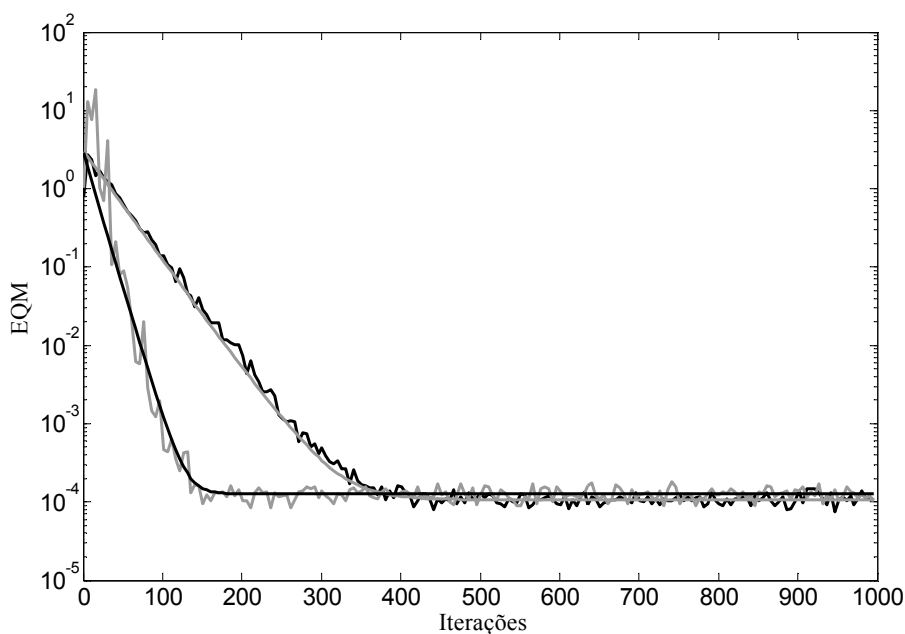


(b)

Figura 5.7- Exemplo 4. Comparações entre modelos para sinal de entrada não-correlacionado. Curvas de EQM. (Linha irregular preta) simulação MC. (Linha contínua cinza-escuro) modelo de [28]. (Linha contínua cinza-claro) modelo proposto. (a) $0,3\mu_{\max}$ e (b) $0,1\mu_{\max}$.



(a)



(b)

Figura 5.8- Exemplo 4. Desempenho do modelo proposto para $0,3\mu_{\max}$ e $0,1\mu_{\max}$ com sinal de entrada branco. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\max}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\max}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\max}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\max}$.

Tabela 5.1- Desajuste \mathcal{M} referente às simulações

		Desajuste \mathcal{M}	
		Experimental	Teórico
Exemplo 1	$0,3\mu_{\max}$	0.23	0.20
	$0,1\mu_{\max}$	0.06	0.06
Exemplo 2	$0,1\mu_{\max}$	0.08	0.08
Exemplo 3	$0,3\mu_{\max}$	0.35	0.37
	$0,1\mu_{\max}$	0.09	0.10
Exemplo 4	$0,3\mu_{\max}$	0.20	0.19
	$0,1\mu_{\max}$	0.06	0.06

Pode ser observado da Tabela 5.1 (obtida dos exemplos anteriormente apresentados) que os valores de desajustes preditos por (4.37) e aqueles obtidos experimentalmente, para diferentes valores de μ , apresentam uma concordância bastante satisfatória entre si, principalmente para adaptação lenta.

Após simulações exaustivas, observamos que o modelo proposto apresenta também resultados consistentes para diferentes ordens de filtros adaptativos como também comprimentos de janela de observação M .

5.5- CONCLUSÕES

Neste capítulo, alguns exemplos foram apresentados com o intuito de se avaliar o modelo estatístico do algoritmo LMS-DT aqui proposto. Primeiramente, fez-se uma comparação entre os resultados obtidos via simulação MC, os obtidos com os modelos apresentados em [26] e [28] e os obtidos a partir do modelo proposto. Os resultados obtidos para o Exemplo 1 mostram que, em regime permanente, todos os modelos apresentam bons resultados. Entretanto, em regime transiente, o modelo proposto prediz com maior precisão o comportamento real do algoritmo LMS-DT.

No caso do Exemplo 2, a principal questão analisada foi o desempenho do modelo para um sinal de entrada altamente correlacionado. Assim como para o Exemplo 1, dois passos de adaptação foram usados. A partir dos resultados, pôde-se verificar o excelente casamento entre os resultados obtidos via método de Monte Carlo e resultados obtidos com o modelo proposto. Além disso, foi novamente verificado que o modelo proposto apresenta resultados mais precisos, quando comparado com aqueles do modelo apresentado em [28].

O objetivo do Exemplo 3 foi analisar o comportamento do modelo para uma ordem diferente de filtro adaptativo. Para tal, usou-se $N = 32$. A partir dos resultados, pôde-se verificar o bom casamento entre os resultados de simulação Monte Carlo e os obtidos com o modelo proposto, tanto para uma condição adaptação lenta quanto para um passo de adaptação um pouco maior $0,3\mu_{\max}$. Um acentuado descasamento, em regime transitório, pode ser observado para os resultados obtidos com o modelo apresentado em [28].

Os exemplos acima discutidos consideram um sinal de entrada correlacionado. Assim, para uma análise mais completa, o Exemplo 4 utiliza um sinal de entrada ruído branco. Dois passos de adaptação, $0,1\mu_{\max}$ e $0,3\mu_{\max}$, são considerados. Observou-se, novamente, um excelente casamento entre os resultados obtidos a partir da simulação Monte Carlo e os obtidos com o modelo proposto.

Os exemplos apresentados foram escolhidos para ilustrar o comportamento do modelo proposto; entretanto, várias outras simulações foram realizadas e, como conclusão geral, pode-se dizer que o modelo proposto apresenta resultados excelentes, quando comparado com modelos apresentados na literatura, para diferentes ordens de filtros adaptativos, adaptação lenta, diferentes graus de correlação entre as amostras do sinal de entrada, como também diferentes janelas de observação M .

CAPÍTULO 6

ANÁLISE ESTATÍSTICA DO ALGORITMO LMS-DT OPERANDO EM AMBIENTE NÃO-ESTACIONÁRIO

A análise estatística do algoritmo LMS-DT para ambientes estacionários tem sido bastante estudada. Uma prova disso são os diversos trabalhos apresentados na literatura [24]-[31]. Os Capítulos 4 e 5 aprofundaram vários estudos para o LMS-DT, para os quais a condição de fundo é considerar que o sistema em questão seja estacionário, tanto em relação aos sinais envolvidos quanto em relação à planta. Em muitas aplicações, os sinais de análise não são estacionários; dessa forma, uma nova abordagem de análise tem que ser considerada. Em tal situação, o algoritmo adaptativo deve seguir as mudanças do ambiente, esse processo é conhecido como *tracking*.

A análise estatística do LMS-DT operando em ambientes não-estacionários, em nosso conhecimento, possui um único breve estudo apresentado por Kim e De Wilde [26]. Assim, este capítulo tem o propósito de desenvolver uma análise do algoritmo LMS-DT considerando que a planta seja variante no tempo. Dessa forma, buscamos dar mais uma contribuição a esse relevante problema ainda pouco explorado na literatura.

O conteúdo desse capítulo é organizado como segue. A Seção 6.1 apresenta o modelo usado para a planta variante no tempo. Na Seção 6.2, são derivadas as expressões analíticas para os momentos de primeira e segunda ordens do vetor de coeficientes do filtro adaptativo do algoritmo LMS-DT operando em ambientes não-estacionários. Finalmente, na Seção 6.3 algumas conclusões deste capítulo são apresentadas.

6.1- FORMULAÇÃO DO PROBLEMA DE IDENTIFICAÇÃO DE SISTEMA PARA PLANTA VARIANTE NO TEMPO

O sinal desejado, considerando um problema de identificação de sistema no qual a planta é variante no tempo, é dado por [1], [2]

$$d(n) = \mathbf{x}^T(n) \mathbf{w}^o(n) + z(n), \quad (6.1)$$

onde $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]^T$ é o vetor do sinal de entrada, $\{x(n)\}$ é um processo estacionário com média nula e distribuição Gaussiana; $\mathbf{w}^o(n) = [w_1^o(n) \ w_2^o(n) \ \dots \ w_{N-1}^o(n)]^T$ é o vetor da planta variante no tempo. O ruído de medição $z(n)$ possui média nula e variância σ_z^2 , sendo decorrelacionado de qualquer outro sinal no sistema. O vetor da planta variante no tempo no domínio transformado é dado por $\mathbf{w}_T^o(n) = \mathbf{T} \mathbf{w}^o(n) = [w_{T,1}^o(n) \ w_{T,2}^o(n) \ \dots \ w_{T,N-1}^o(n)]^T$, onde \mathbf{T} é matriz de transformação ortogonal. Similarmente, o vetor do sinal de entrada no domínio transformado é $\mathbf{x}_T(n) = \mathbf{T} \mathbf{x}(n) = [x_{T,0}(n) \ x_{T,1}(n) \ \dots \ x_{T,N-1}(n)]^T$. Usando-se os vetores transformados, pode-se reescrever (6.1) como segue:

$$d(n) = \mathbf{x}_T^T(n) \mathbf{w}_T^o(n) + z(n). \quad (6.2)$$

O propósito do algoritmo adaptativo é seguir as variações de $\mathbf{w}^o(n)$, que é governado pela expressão

$$\mathbf{w}_T^o(n+1) = \mathbf{w}_T^o(n) + \mathbf{g}(n), \quad (6.3)$$

onde o vetor $\mathbf{g}(n)$ é o processo de perturbação da planta, possuindo média zero e variância σ_g^2 .

6.2- ANÁLISE ESTATÍSTICA

Nesta seção, as expressões analíticas dos momentos de primeira e segunda ordens do vetor de coeficientes do filtro adaptativo em um ambiente não-estacionário são derivadas. A equação de adaptação dos coeficientes no domínio transformado é dada por [2]

$$\mathbf{w}_T(n+1) = \mathbf{w}_T(n) + 2\mu\mathbf{D}^{-1}(n)e(n)\mathbf{x}_T(n), \quad (6.4)$$

onde $\mathbf{w}_T(n) = [w_{T,0}(n) \ w_{T,1}(n) \ \cdots \ w_{T,N-1}(n)]^T$ representa o vetor de coeficientes do filtro adaptativo e $\mathbf{D}^{-1}(n) = \text{diag}[\sigma_0^2(n) \ \sigma_1^2(n) \ \cdots \ \sigma_{N-1}^2(n)]$ é a matriz responsável pela normalização de potência. O sinal de erro é determinado como

$$e(n) = d(n) - \mathbf{x}_T^T(n)\mathbf{w}_T(n). \quad (6.5)$$

Substituindo-se (6.5) em (6.4) e usando (6.2), obtém-se

$$\mathbf{w}_T(n+1) = \mathbf{w}_T(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)[\mathbf{x}_T^T(n)\mathbf{w}_T(n) + z(n) - \mathbf{x}_T^T(n)\mathbf{w}_T(n)] \quad (6.6)$$

Definindo-se o vetor erro de coeficiente no domínio transformado como $\mathbf{v}_T(n+1) = \mathbf{w}_T(n+1) - \mathbf{w}_T^o(n+1)$, (6.6) pode ser reescrita como

$$\begin{aligned} \mathbf{v}_T(n+1) = & \mathbf{v}_T(n) + \mathbf{w}_T^o(n) - \mathbf{w}_T^o(n+1) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{w}_T^o(n) \\ & + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n) - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)[\mathbf{v}_T(n) + \mathbf{w}_T^o(n)] \end{aligned} \quad (6.7)$$

Agora, substituindo-se (6.3) em (6.7), obtém-se a expressão recursiva em termos de $\mathbf{v}_T(n)$ como segue:

$$\begin{aligned} \mathbf{v}_T(n+1) &= \mathbf{v}_T(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n) - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n) - \mathbf{g}(n) \\ &= [\mathbf{I} - 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\mathbf{v}_T(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n) - \mathbf{g}(n) \end{aligned} \quad (6.8)$$

O passo seguinte é determinar os momentos de primeira e segunda ordens de (6.8).

6.2.1- CONSIDERAÇÕES DE ANÁLISE

Para proceder com a análise estatística, as seguintes considerações são adotadas:

- i) $\mathbf{g}(n)$ é Gaussiano, média nula, estacionário com matriz de autocorrelação dada por $\mathbf{G} = E[\mathbf{g}(n)\mathbf{g}^T(n)] = \sigma_q^2\mathbf{I}$ e $\mathbf{g}(n)$ é estatisticamente independente de qualquer outro sinal no sistema.
- ii) $\mathbf{x}_T(n)$ é Gaussiano, média nula, estacionário com matriz de autocorrelação \mathbf{R}_T .
- iii) $\mathbf{v}_T(n)$ e $\mathbf{x}_T(n)$ são estatisticamente independentes.
- iv) $\mathbf{D}^{-1}(n)$ e $\mathbf{x}_T(n)\mathbf{x}_T^T(n)$ são processos conjuntamente estacionários, tal que $\mathbf{D}^{-1}(n)$ varia lentamente com relação a $\mathbf{x}_T(n)\mathbf{x}_T^T(n)$. Esta consideração permite utilizar o Princípio da Média [37] (condição H2 apresentada no Capítulo 4).

Agora, pode-se então continuar com a derivação do modelo.

6.2.2- MOMENTO DE PRIMEIRA ORDEM DE $\mathbf{v}_T(n)$

Tomando-se o valor esperado de ambos os lados de (6.8), obtém-se

$$\begin{aligned} E[\mathbf{v}_T(n+1)] &= E[\mathbf{v}_T(n)] - 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)] \\ &\quad + 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)] - E[\mathbf{g}(n)] \end{aligned} \quad (6.9)$$

Usando-se as considerações (i)-(iv) da Seção 6.2.1, obtém-se

$$E[\mathbf{v}_T(n+1)] \cong \{\mathbf{I} - 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}E[\mathbf{v}_T(n)]. \quad (6.10)$$

onde a solução da expressão $E[\mathbf{D}^{-1}(n)]$ já foi apresentada na Seção 4.1.1 (ver detalhes no Apêndice B).

6.2.3- MOMENTO DE SEGUNDA ORDEM DE $\mathbf{v}_T(n)$

A expressão do momento de segunda ordem para o vetor de erro dos coeficientes, no domínio transformado, é o valor esperado $\mathbf{K}(n) = E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]$. Determinando-se o produto externo $\mathbf{v}_T(n)\mathbf{v}_T^T(n)$ e tomando-se o valor esperado de ambos os lados da expressão resultante, obtém-se

$$\begin{aligned} \mathbf{K}(n+1) &= E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)] + 2\mu E[z(n)\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\ &\quad \overbrace{-2\mu E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] - E[\mathbf{v}_T(n)\mathbf{g}^T(n)]}^A \\ &\quad + 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n)] \\ &\quad \overbrace{+4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z^2(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]}^B \\ &\quad - 4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\ &\quad - 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{g}^T(n)] \\ &\quad \overbrace{-2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)]}^C \\ &\quad - 4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)z(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\ &\quad \overbrace{+4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)]}^D \\ &\quad + 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{g}^T(n)] - E[\mathbf{g}(n)\mathbf{v}_T^T(n)] \\ &\quad - 2\mu E[z(n)\mathbf{g}(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] \\ &\quad + 2\mu E[\mathbf{g}(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] + E[\mathbf{g}(n)\mathbf{g}^T(n)] \quad (6.11) \end{aligned}$$

Consideraram-se primeiramente os termos que contêm $z(n)$, exceto aquele com $z^2(n)$, e utilizando a definição de $z(n)$, tem-se

- $E[z(n)\mathbf{v}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] = 0$.
- $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n)] = 0$.
- $E[z(n)\mathbf{g}(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] = 0$.
- $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{g}^T(n)] = 0$.
- $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] = 0$.
- $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)z(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] = 0$.

Usando-se as considerações de análise (i)-(iv), os valores esperados A, B, C, e D são:

$$\begin{aligned} \text{A. } -2\mu E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] &= -2\mu E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{D}^{-1}(n)] \\ &= -2\mu\mathbf{K}(n)\mathbf{R}_T E[\mathbf{D}^{-1}(n)] \end{aligned}$$

$$\begin{aligned} \text{B. } 4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)z^2(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] &= 4\mu^2 E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{D}^{-1}(n)]E[z^2(n)] \\ &= 4\mu^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_T E[\mathbf{D}^{-1}(n)]\sigma_z^2 \end{aligned}$$

$$\begin{aligned} \text{C. } -2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)] &= -2\mu E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)] \\ &= -2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\mathbf{K}(n) \end{aligned}$$

$$\begin{aligned} \text{D. } 4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{D}^{-1}(n)] &= \\ 4\mu^2 E[\mathbf{D}^{-1}(n)] \{2E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)] & \\ + E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\text{tr}[E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]]\} E[\mathbf{D}^{-1}(n)] & \\ = 4\mu^2 E[\mathbf{D}^{-1}(n)] \{2\mathbf{R}_T\mathbf{K}(n)\mathbf{R}_T + \mathbf{R}_T\text{tr}[\mathbf{K}(n)\mathbf{R}_T]\} E[\mathbf{D}^{-1}(n)] & \end{aligned}$$

Assim, com estes resultados, pode-se reescrever (6.11) como segue:

$$\begin{aligned} \mathbf{K}(n+1) &= \mathbf{K}(n) - 2\mu\mathbf{K}(n)\mathbf{R}_T E[\mathbf{D}^{-1}(n)] - 2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\mathbf{K}(n) \\ &\quad + 4\mu^2 E[\mathbf{D}^{-1}(n)] \{2\mathbf{R}_T\mathbf{K}(n)\mathbf{R}_T + \mathbf{R}_T\text{tr}[\mathbf{R}_T\mathbf{K}(n)]\} E[\mathbf{D}^{-1}(n)] \quad (6.12) \\ &\quad + 4\mu^2\sigma_z^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_T E[\mathbf{D}^{-1}(n)] + \mathbf{G} \end{aligned}$$

onde a solução da expressão $E[\mathbf{D}^{-1}(n)]$ está apresentada no Apêndice B.

6.2.4- CURVA DE APRENDIZAGEM

De (6.2) e (6.5), tem-se que o sinal de erro é dado por

$$\begin{aligned} e(n) &= \mathbf{x}_T^T(n)\mathbf{w}_T^o(n) + z(n) - \mathbf{x}_T^T(n)[\mathbf{v}_T(n) + \mathbf{w}_T^o(n)] \\ &= z(n) - \mathbf{x}_T^T(n)\mathbf{v}_T(n). \end{aligned} \quad (6.13)$$

Elevando-se ao quadrado (6.13), tomando-se o valor esperado de ambos os lados da expressão resultante, e da definição de $z(n)$, obtém-se:

$$E[e^2(n)] = E[z^2(n)] + E[\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)]. \quad (6.14)$$

O segundo termo em (6.14) é manipulado como segue:

$$\begin{aligned} E[\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)\mathbf{v}_T(n)] &= \text{tr}\{E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]\} \\ &= \text{tr}\{\mathbf{R}_T E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]\}. \end{aligned} \quad (6.15)$$

Relembrando-se que $\mathbf{K}(n) = E[\mathbf{v}_T(n)\mathbf{v}_T^T(n)]$ e substituindo-se (6.15) em (6.14), a curva de aprendizagem é descrita por

$$E[e^2(n)] = \sigma_z^2 + \text{tr}\{\mathbf{R}_T\mathbf{K}(n)\}. \quad (6.16)$$

6.2.5- ERRO EM EXCESSO

A expressão do erro em excesso é definida por [1], [2]

$$\xi_{\text{exc}} = \text{tr}[\mathbf{R}_T\mathbf{K}(\infty)]. \quad (6.17)$$

Assumindo-se que o regime permanente tenha sido atingido, ou seja $\lim_{n \rightarrow \infty} \mathbf{K}(n+1) = \lim_{n \rightarrow \infty} \mathbf{K}(n) = \mathbf{K}(\infty)$, e após algumas manipulações algébricas, pode-se expressar (6.12) como segue:

$$\begin{aligned} \mu \{E[\mathbf{D}^{-1}(n)]\}^{-1} \mathbf{K}(\infty) \mathbf{R}_T E[\mathbf{D}^{-1}(n)] + \mu \mathbf{R}_T \mathbf{K}(\infty) &= 4\mu^2 \mathbf{R}_T \mathbf{K}(\infty) \mathbf{R}_T E[\mathbf{D}^{-1}(n)] + 0.5 \{E[\mathbf{D}^{-1}(n)]\}^{-1} \mathbf{G} \\ &+ 2\mu^2 \mathbf{R}_T \text{tr}[\mathbf{R}_T \mathbf{K}(\infty)] E[\mathbf{D}^{-1}(n)] + 2\mu^2 \sigma_z^2 \mathbf{R}_T E[\mathbf{D}^{-1}(n)] \end{aligned} \quad (6.18)$$

Computando-se o traço de ambos os lados de (6.18) e usando-se a identidade $\text{tr}[\mathbf{AB}] = \text{tr}[\mathbf{BA}]$ [35], as seguintes relações são obtidas:

- $\text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{K}(\infty) \mathbf{R}_T E[\mathbf{D}^{-1}(n)]\} = \text{tr}[\mathbf{R}_T E[\mathbf{D}^{-1}(n)] \{E[\mathbf{D}^{-1}(n)]\}^{-1} \mathbf{K}(\infty)] = \text{tr}[\mathbf{R}_T \mathbf{K}(\infty)]$
- $\text{tr}\{\mathbf{R}_T \mathbf{K}(\infty) \mathbf{R}_T E[\mathbf{D}^{-1}(n)]\} = \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T \mathbf{K}(\infty) \mathbf{R}_T\}$
- $\text{tr}\{\mathbf{R}_T E[\mathbf{D}^{-1}(n)]\} = \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T\}$

Então,

$$\begin{aligned} 2\xi_{\text{exc}} &= 4\mu \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T \mathbf{K}(\infty) \mathbf{R}_T\} + \frac{1}{2\mu} \text{tr}[\{E[\mathbf{D}^{-1}(n)]\}^{-1} \mathbf{G}] \\ &+ 2\mu \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T\} \xi_{\text{exc}} + 2\mu \sigma_z^2 \text{tr}[\mathbf{D}^{-1}(n) \mathbf{R}_T] \end{aligned} \quad (6.19)$$

Através de simulações, pode-se verificar que $2\mu \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T \mathbf{K}(\infty) \mathbf{R}_T\} \approx 0$. Então, (6.19) torna-se

$$\xi_{\text{exc}} = \frac{1}{1 - \mu \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T\}} \left\{ \frac{1}{4\mu} \text{tr}[\{E[\mathbf{D}^{-1}(n)]\}^{-1} \mathbf{G}] + \mu \sigma_z^2 \text{tr}\{E[\mathbf{D}^{-1}(n)] \mathbf{R}_T\} \right\}. \quad (6.20)$$

A partir de (6.20), pode-se verificar que o erro em excesso, quando o algoritmo opera em ambientes não-estacionários, é acrescido de uma parcela referente ao ruído $\mathbf{g}(n)$. Além disso, essa parcela torna-se mais evidente quanto mais lenta for a adaptação. Dessa forma, para ambientes não-estacionários, não se pode dizer que quanto menor o passo, menor será o erro em excesso, como obtido para ambientes estacionários [ver Eq. (4.37)].

6.2.6- DESAJUSTE

O desajuste \mathcal{M} é obtido de (6.20) e dado por

$$\mathcal{M} = \frac{\xi_{\text{exc}}}{\xi_{\text{min}}} = \frac{1}{1 - \mu \text{tr}\{E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}} \left\{ \frac{1}{4\mu} \sigma_z^{-2} \text{tr}\{[E[\mathbf{D}^{-1}(n)]]^{-1}\mathbf{G}\} + \mu \text{tr}\{E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\} \right\} \quad (6.21)$$

Pode-se considerar o desajuste do algoritmo composto pela soma de dois termos. Assim,

$$\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 \quad (6.22)$$

onde

$$\mathcal{M}_1 = \frac{\mu \text{tr}\{E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}}{1 - \mu \text{tr}\{E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}} \quad (6.23)$$

e

$$\mathcal{M}_2 = \frac{\frac{1}{4\mu} \sigma_z^{-2} \text{tr}\{[E[\mathbf{D}^{-1}(n)]]^{-1}\mathbf{G}\}}{1 - \mu \text{tr}\{E[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}}. \quad (6.24)$$

Pode ser observado que quando o algoritmo LMS-DT opera em ambiente não-estacionário, o desajuste é acrescido de uma parcela [(6.24)] relacionada ao ruído $\mathbf{g}(n)$ do processo [uma vez que (6.23) é equivalente a (4.37)].

6.2.7- PASSO DE ADAPTAÇÃO ÓTIMO

O passo de adaptação ótimo, que minimize o EQM em excesso e, conseqüentemente, o desajuste do algoritmo LMS-DT, é obtido derivando-se (6.20) em

relação a μ e igualando-se o resultado a zero. Assim, fazendo-se $a = \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}$ e $b = \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\}^{-1}\mathbf{G}$, tem-se

$$\frac{\partial \xi_{\text{exc}}}{\partial \mu} = \frac{\partial}{\partial \mu} \left[\frac{1}{1-\mu a} \left(\frac{b}{4\mu} + \mu \sigma_z^2 a \right) \right] = 0 \quad (6.25)$$

Logo,

$$\begin{aligned} \frac{\partial \xi_{\text{exc}}}{\partial \mu} &= \frac{-b(4-8\mu a)}{(4\mu-4\mu^2 a)^2} + \frac{\sigma_z^2 a(1-\mu a) - \mu a \sigma_z^2 (-a)}{(1-\mu a)^2} \\ &= \frac{-b(4-8\mu a)}{(4\mu-4\mu^2 a)^2} + \frac{\sigma_z^2 a(1-\mu a) - \mu a \sigma_z^2 (-a)}{(1-\mu a)^2} \\ &= \frac{8\mu ab - 4b + 16\mu^2 \sigma_z^2 a}{(1-\mu a)^2} = 0 \end{aligned} \quad (6.26)$$

Assim, obtém-se

$$8\mu ab - 4b + 16\mu^2 \sigma_z^2 a = 0 \quad (6.27)$$

Resolvendo-se a equação de segundo grau, chega-se ao valor de passo que minimiza o EQM. Portanto,

$$\begin{aligned} \mu_{\text{ótimo}} &= \frac{-ab + \sqrt{(ab)^2 + 4\sigma_z^2 ab}}{4\sigma_z^2 a} = \frac{-b}{4\sigma_z^2} + \frac{\sqrt{(ab)^2 + 4\sigma_z^2 ab}}{4\sigma_z^2 a} \\ &= \frac{-\text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\}^{-1}\mathbf{G}}{4\sigma_z^2} \\ &\quad + \frac{\sqrt{(\text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\mathbf{R}_T\} \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\}^{-1}\mathbf{G})^2 + 4\sigma_z^2 \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\mathbf{R}_T\} \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\}^{-1}\mathbf{G}}}{4\sigma_z^2 \text{tr}\{\mathbf{E}[\mathbf{D}^{-1}(n)]\mathbf{R}_T\}} \end{aligned} \quad (6.28)$$

A partir de (6.28), pode-se verificar que o valor do passo ótimo, quando o algoritmo opera em ambiente não-estacionário, apresenta influência do ruído $\mathbf{g}(n)$.

6.2.8- GRAU DE NÃO-ESTACIONARIDADE

O grau de não-estacionaridade, definido por α , mede a capacidade do algoritmo em acompanhar o processo [1]. Assim,

$$\alpha = \left(\frac{E\left[|\mathbf{g}^T(n)\mathbf{x}_T(n)|^2\right]}{E[|z(n)|^2]} \right)^{1/2}. \quad (6.29)$$

A partir de (6.29), pode-se verificar que tanto a potência média do sinal de entrada quanto a potência média do ruído de medição influem diretamente no grau de não-estacionaridade.

Considerando-se a independência entre $\mathbf{g}(n)$ e $\mathbf{x}_T(n)$, pode-se então escrever [1]

$$\begin{aligned} E\left[|\mathbf{g}^H(n)\mathbf{x}_T(n)|^2\right] &= E[\mathbf{g}^H(n)\mathbf{x}_T(n)\mathbf{x}_T^H(n)\mathbf{g}(n)] \\ &= \text{tr}\{E[\mathbf{g}^H(n)\mathbf{x}_T(n)\mathbf{x}_T^H(n)\mathbf{g}(n)]\} \\ &= E\{\text{tr}[\mathbf{g}^H(n)\mathbf{x}_T(n)\mathbf{x}_T^H(n)\mathbf{g}(n)]\} \\ &= E\{\text{tr}[\mathbf{g}(n)\mathbf{g}^H(n)\mathbf{x}_T(n)\mathbf{x}_T^H(n)]\} \\ &= \text{tr}\{E[\mathbf{g}(n)\mathbf{g}^H(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^H(n)]\} \\ &= \text{tr}[\mathbf{G}\mathbf{R}_T], \end{aligned} \quad (6.30)$$

onde \mathbf{G} é a matriz de autocorrelação do vetor ruído do processo $\mathbf{g}(n)$. Substituindo-se (6.30) em (6.29) e lembrando-se que σ_z^2 é a variância de $z(n)$, obtém-se

$$\alpha = \frac{1}{\sigma_z} [\text{tr}(\mathbf{G}\mathbf{R}_T)]^{1/2} = \frac{1}{\sigma_z} [(\sigma_0^2 + \sigma_1^2 + \dots + \sigma_{N-1}^2)\sigma_g^2]^{1/2} = \frac{\sigma_g}{\sigma_z} [\text{tr}(\mathbf{R}_T)]^{1/2}. \quad (6.31)$$

Relacionando-se então o desajuste do algoritmo com o grau de estacionaridade, tem-se

$$\mathcal{M} = \frac{\xi_{\text{exc}}}{\xi_{\text{min}}} \geq \frac{\text{tr}[\mathbf{G}\mathbf{R}_T]}{\sigma_z} = \alpha^2. \quad (6.32)$$

Assim, o desajuste impõe um limite superior para o grau de não-estacionaridade. A partir de (6.32), podem-se tecer algumas conclusões:

- 1) Para variações estatísticas lentas, α é pequeno. Isso, por sua vez, quer dizer que é possível construir um filtro adaptativo capaz de acompanhar um sistema variante no tempo com pequeno desajuste.
- 2) Quando as variações estatísticas são rápidas, α pode ser maior do que 1. Nesse caso, o desajuste produzido pelo filtro adaptativo excede 100 %, o que quer dizer que não há vantagens em se utilizar um filtro adaptativo para resolver o problema de *tracking*.

6.3- CONCLUSÕES

Este capítulo apresentou uma análise estatística do algoritmo LMS-DT operando em ambientes não-estacionários. Assim como para o caso de ambientes estacionários, foram necessárias algumas considerações simplificativas para o desenvolvimento de tal análise, como por exemplo, sinal de entrada com distribuição Gaussiana e passo de adaptação lenta.

CAPÍTULO 7

RESULTADOS DE SIMULAÇÕES: CASO NÃO-ESTACIONÁRIO

O modelo estatístico do algoritmo LMS-DT para o caso não-estacionário aqui proposto é aplicado a um problema de identificação de sistemas, no qual a precisão do modelo é avaliada para sinais de entrada correlacionados com distribuição Gaussiana, obtidos de um processo AR(2), dado por

$$x(n) = a_1x(n-1) + a_2x(n-2) + v(n), \quad (7.1)$$

onde $v(n)$ é um ruído branco com variância σ_v^2 de modo que a variância de $x(n)$ seja igual a 1, a_1 e a_2 são os coeficientes auto-regressivos, com $a_1 = 0,1833$ e $a_2 = -0,85$. A variância do ruído de medição $z(n)$ usada é 10^{-6} (SNR = 60dB). Note que dado os parâmetros σ_z^2 e α , bem como a transformação ortogonal, através de (6.31) pode ser determinada a variância do processo de perturbação. Todas as simulações Monte Carlo são obtidas de uma média de 500 realizações independentes. Para todos os exemplos considerados, os valores de passo de adaptação são: $0,3\mu_{\text{ótimo}}$ e $0,1\mu_{\text{ótimo}}$, onde $\mu_{\text{ótimo}}$ é obtido de (6.28). Os coeficientes variantes no tempo da planta são determinados pela seguinte expressão:

$$\mathbf{w}_T^o(n+1) = \mathbf{w}_T^o(n) + \mathbf{g}(n), \quad (7.2)$$

onde,

$$\mathbf{w}_T^o(0) = \mathbf{T}[\text{sinc}(0) \text{ sinc}(1/N) \dots \text{sinc}(N-1/N)]^T. \quad (7.3)$$

Os elementos de $\mathbf{g}(n)$ são amostras de um processo com características de ruído branco com uma variância obtida a partir de (6.31), dada por

$$\sigma_g^2 = \left(\frac{\alpha \sigma_z}{[\text{tr}(\mathbf{R}_T)]^{1/2}} \right)^2. \quad (7.3)$$

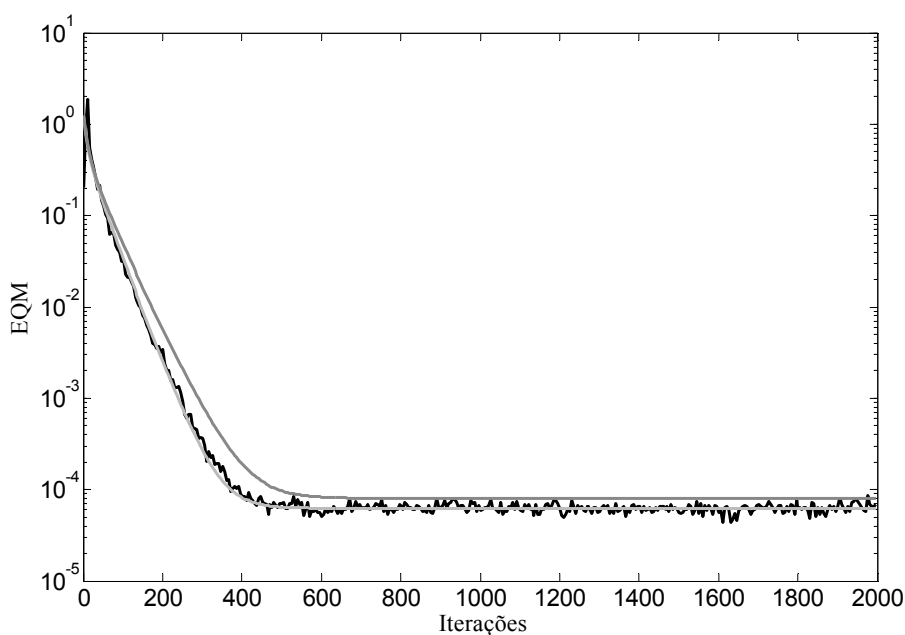
Aqui, a DCT é também usada como transformação ortogonal [2]. O passo de adaptação μ considerado nos exemplos estão relacionados com o valor de passo ótimo $\mu_{\text{ótimo}}$ para o qual o algoritmo adaptativo apresenta o mínimo erro em excesso. A constante $\varepsilon = 0,001$ é também considerada para todos os exemplos.

É importante ressaltar que as variações produzidas nos coeficientes do filtro adaptativo, devido às variações dos coeficientes da planta, não são nitidamente notadas nas figuras dos exemplos mostrados a seguir, pois essas descrevem o comportamento médio dos coeficientes do filtro adaptativo. Devemos nos lembrar aqui que a parcela de perturbação $\mathbf{g}(n)$ de (7.2) corresponde a um processo de média nula. Além do mais, como o objetivo desse capítulo é avaliar o modelo estatístico proposto para o algoritmo LMS-DT, as simulações MC são tomadas considerando um número elevado de realizações independentes, o que também atenua as variações produzidas nos coeficientes do filtro adaptativo. Outro ponto que deve ser ressaltado diz respeito ao modelo utilizado para efeito de comparação com o modelo proposto. Os dois modelos apresentados em [30]-[31] deveriam ser usados para comparar com o modelo proposto. Entretanto, os modelos discutidos em [30] e [31] são idênticos filosoficamente e em desempenho. Portanto, apenas o modelo apresentado em [30] é considerado nos exemplos discutidos a seguir.

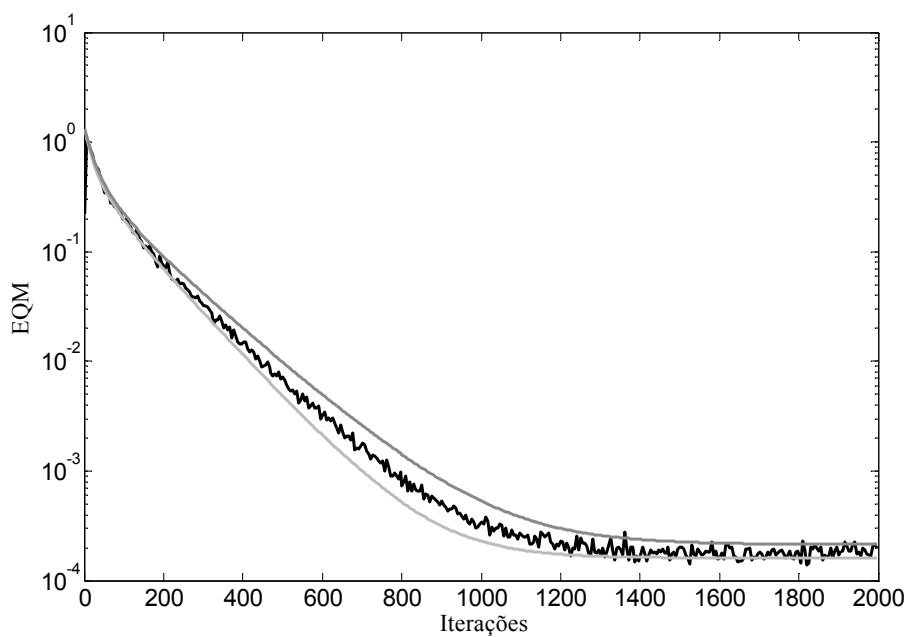
7.1- EXEMPLO 1

Para este exemplo, os seguintes parâmetros são usados: $N=8$ em (7.3), $M=32$ e $\mu_{\text{ótimo}}=0,0467$ [de (6.28)]. A dispersão dos autovalores da matriz de autocorrelação do sinal de entrada igual a 81 e o grau de não-estacionaridade em (6.31) é $\alpha=2$. A Fig. 7.1 ilustra uma comparação entre os resultados de EQM obtidos via simulação MC, obtidos a partir do modelo apresentado em [30] e obtidos pelo modelo proposto, para $0,1\mu_{\text{ótimo}}$ e $0,3\mu_{\text{ótimo}}$. Para esse exemplo, verifica-se que os resultados obtidos a partir do modelo proposto apresentam uma melhor concordância com os resultados de simulação MC do que aqueles obtidos através do modelo de [30].

A Fig. 7.2 compara o desempenho do modelo proposto para dois valores de μ ($0,1\mu_{\text{ótimo}}$ e $0,3\mu_{\text{ótimo}}$). Dessa figura, pode-se constatar também concordância muito boa entre os resultados obtidos de simulação e os obtidos a partir do modelo proposto.

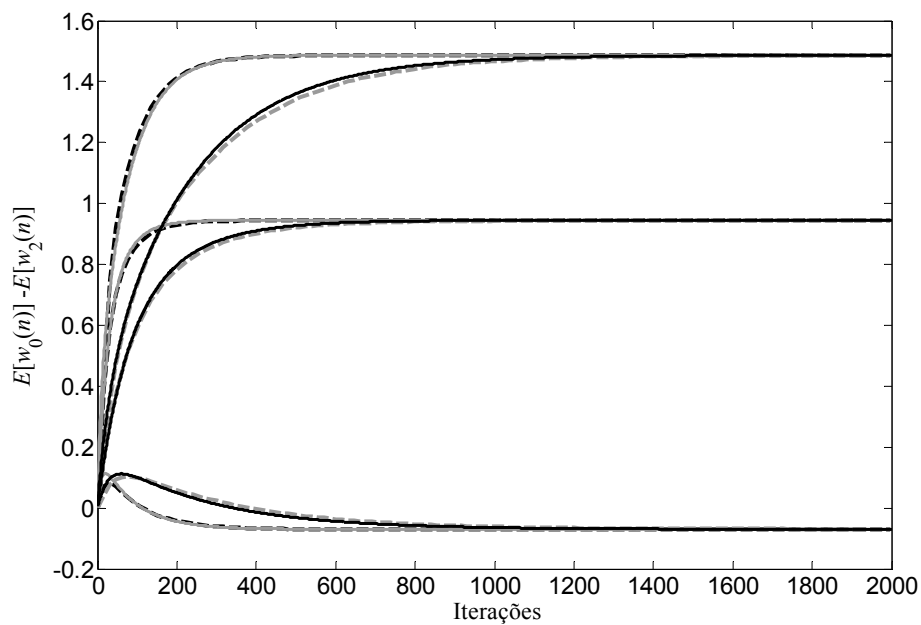


(a)

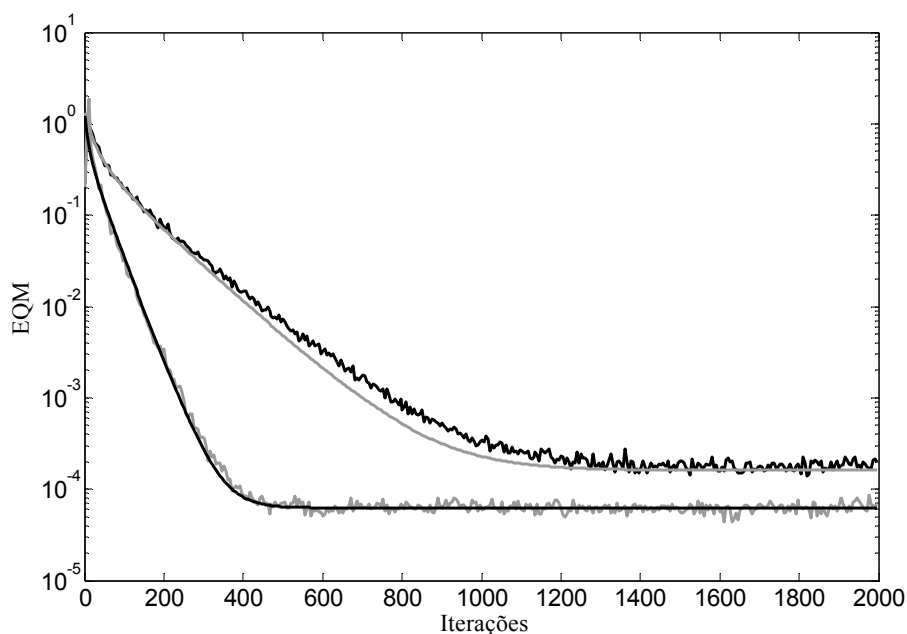


(b)

Figura 7.1- Exemplo 1. Curvas de EQM para uma dispersão dos autovalores igual a e $\alpha = 2$ com SNR de 60 dB. (Linhas irregulares) simulação MC. (Linhas contínuas cinza-escuro) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)]. (a) $0,3\mu_{\text{ótimo}}$ · (b) $0,1\mu_{\text{ótimo}}$ ·



(a)

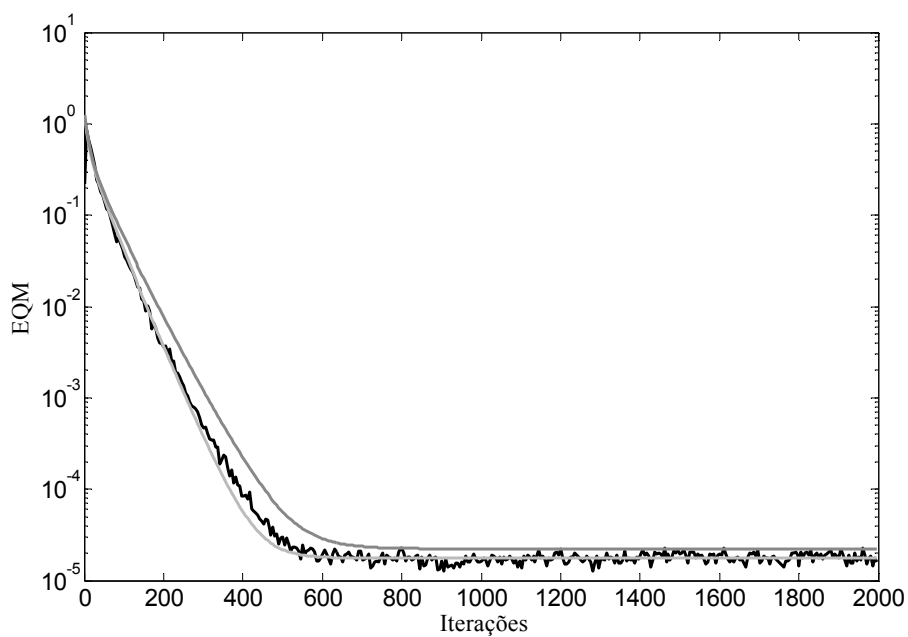


(b)

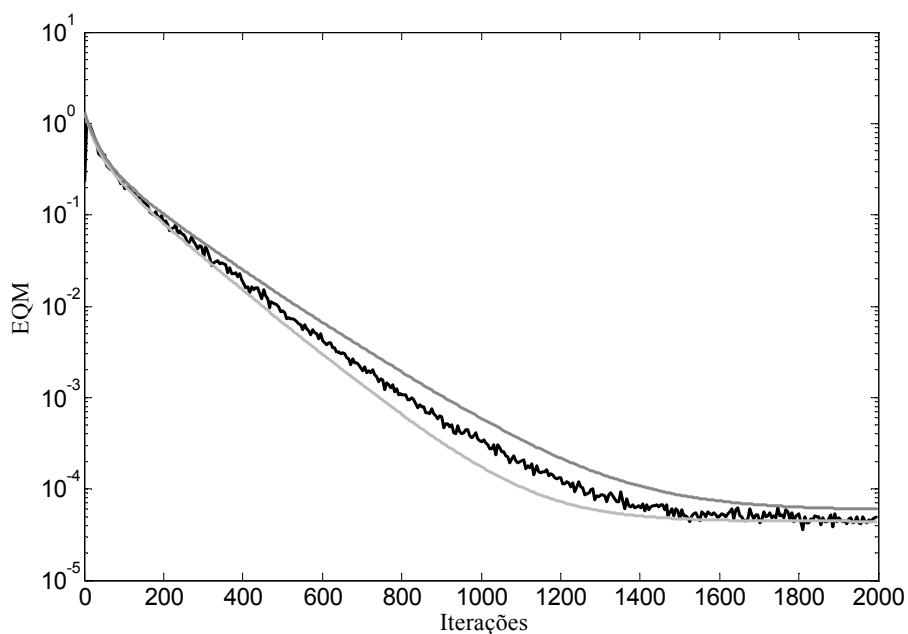
Figura 7.2- Exemplo 1. Desempenho do modelo proposto para $0,3\mu_{\text{ótimo}}$ e $0,1\mu_{\text{ótimo}}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\text{ótimo}}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha irregular preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,1\mu_{\text{ótimo}}$.

7.2- EXEMPLO 2

Para este exemplo, são usados os mesmos parâmetros do Exemplo 1, mas com um grau de não-estacionaridade diferente $\alpha = 1$ e $\mu_{\text{ótimo}} = 0,043$. Os resultados obtidos são mostrados nas Figs. 7.3 e 7.4, seguindo o mesmo padrão de apresentação das Figs. 7.1 e 7.2. Novamente, pode-se observar uma predição muito boa dos resultados obtidos do modelo proposto, quando comparada com os resultados do modelo discutido em [30].

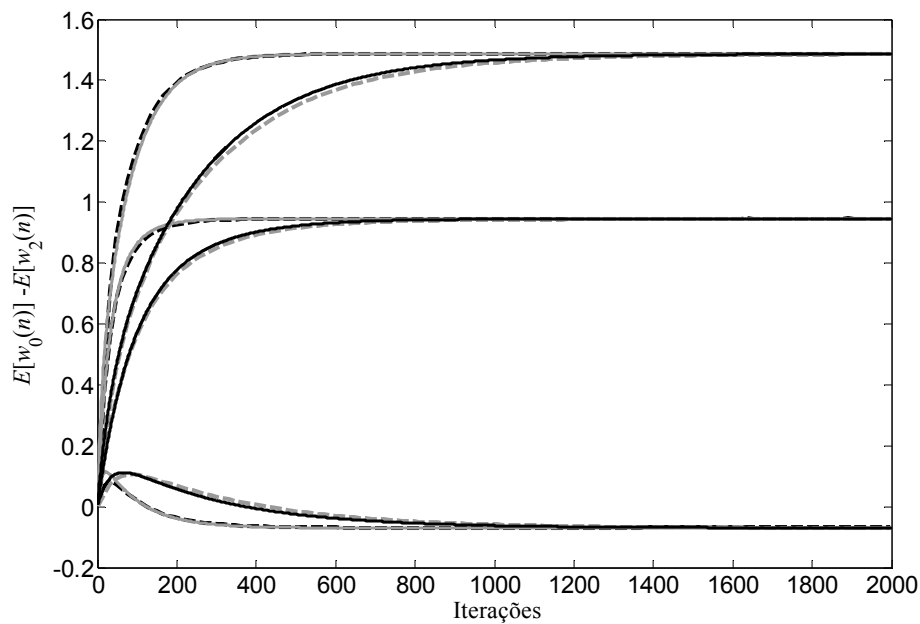


(a)

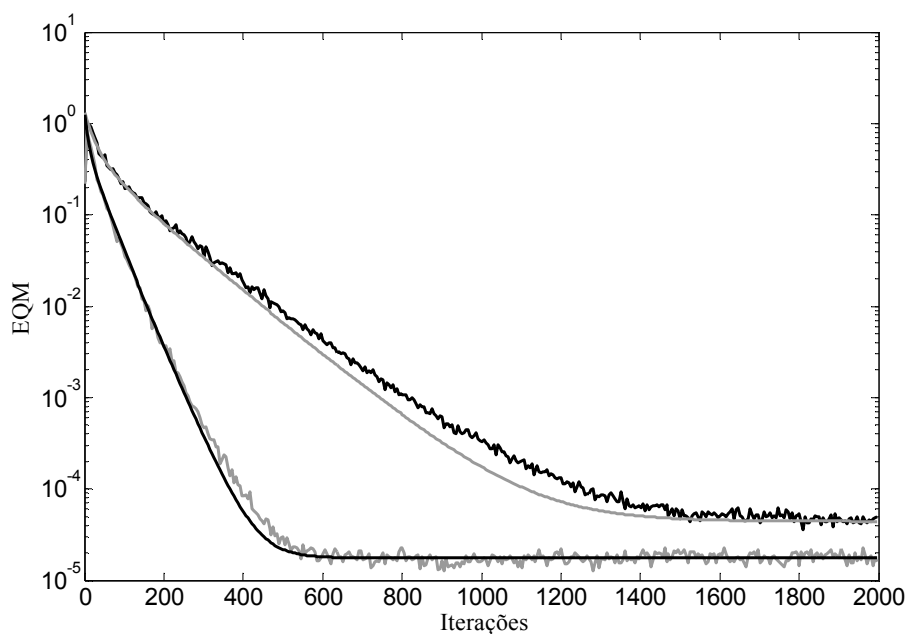


(b)

Figura 7.3- Exemplo 2. Curvas de EQM com uma dispersão dos autovalores igual a 81 e $\alpha = 1$ para SNR de 60 dB. (Linhas irregulares) simulação MC. (Linhas contínuas cinza-escuro) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)]. (a) $0,3\mu_{\text{ótimo}}$, (b) $0,1\mu_{\text{ótimo}}$.



(a)

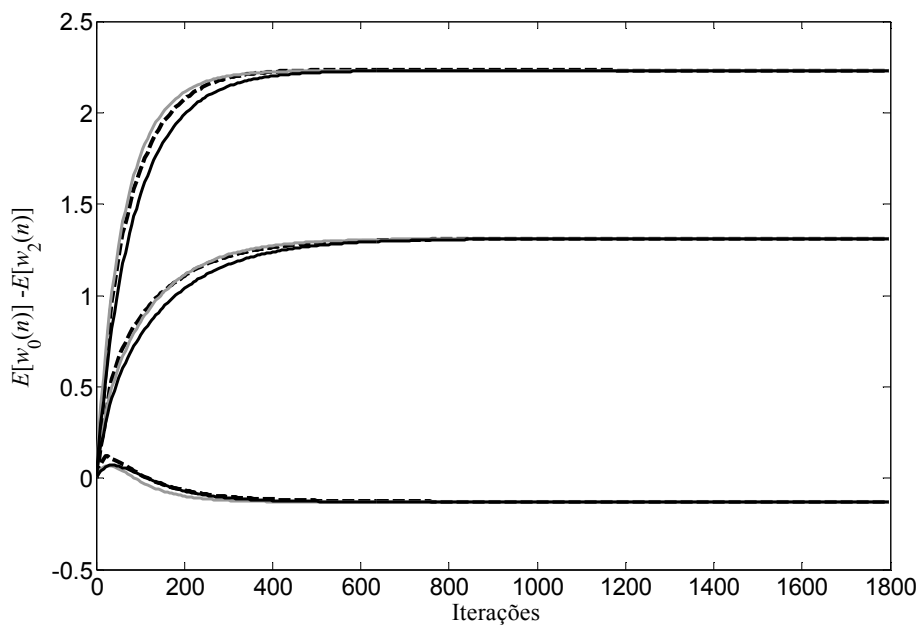


(b)

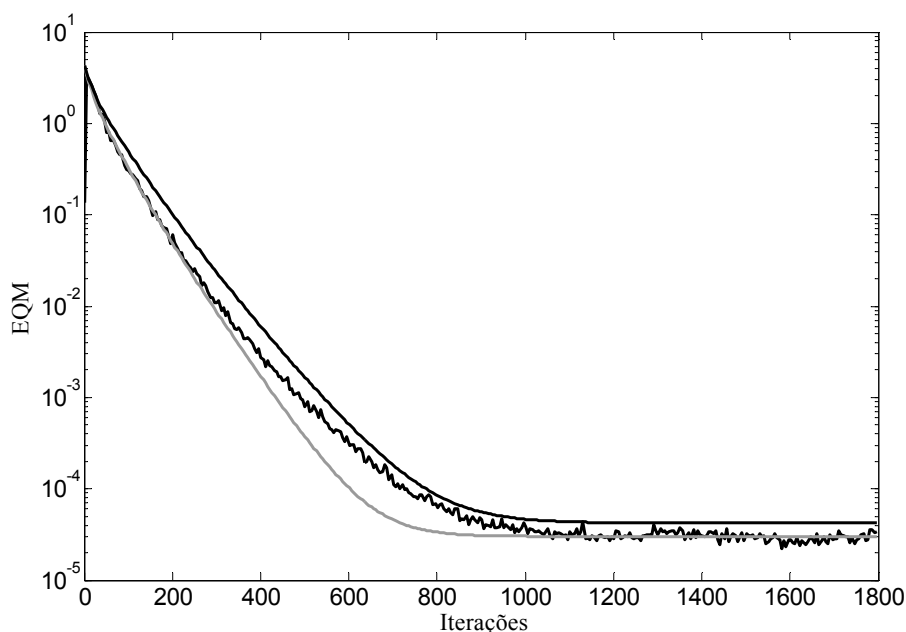
Figura 7.4- Exemplo 2. Desempenho do modelo proposto para $0,3\mu_{\text{ótimo}}$ e $0,1\mu_{\text{ótimo}}$ para sinal de entrada correlacionado com $\chi = 81$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua cinza) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha tracejada cinza) simulação MC. (Linha contínua preta) modelo proposto para $0,1\mu_{\text{ótimo}}$. (b) Curvas de EQM. (Linha irregular cinza) simulação MC. (Linha sólida preta) modelo proposto para $0,3\mu_{\text{ótimo}}$. (Linha irregular preta) simulação MC. (Linha sólida cinza) modelo proposto para $0,1\mu_{\text{ótimo}}$.

7.3- EXEMPLO 3

Para este exemplo foram usados os seguintes parâmetros: $N = 16$ em (7.3), $M = 32$, e $\mu_{\text{ótimo}} = 0.0217$. O passo de adaptação é $0.3\mu_{\text{ótimo}}$. A dispersão dos autovalores da matriz de autocorrelação do sinal de entrada é 356 e o grau de não-estacionaridade em (6.31) é $\alpha = 1$. A Fig. 7.5 compara os resultados obtidos através de simulação MC, modelo proposto e modelo apresentado em [30], tanto para comportamento dos pesos quanto para o EQM. Pode ser também verificada aqui a boa concordância existente entre os resultados de simulação MC e do modelo proposto.



(a)



(b)

Figura 7.5- Exemplo 3. Resultados dos modelos com uma dispersão dos autovalores igual a 356 e $\alpha = 1$, $SNR = 60$ dB e $0,3\mu_{\text{ótimo}}$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0, 1, 2$. (Linha tracejada preta) simulação MC. (Linha contínua preta) modelo de [30]. (Linha contínua cinza-claro) modelo proposto. (b) Curvas de EQM. (Linhas irregulares) simulação MC. (Linhas contínuas preta) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)].

7.4- EXEMPLO 4

Neste exemplo, mostra-se o desempenho do modelo para diferentes tamanhos de janelas de observação $M = \{8, 16, 32, 64\}$. Para esse exemplo a planta apresenta 8 coeficientes ($N = 8$), obtidos de (7.3) e $SNR = 60$ dB. Os passos de adaptação são $\mu_{\text{ótimo}} = 0.0263$ ($M = 8$), $\mu_{\text{ótimo}} = 0.0353$ ($M = 16$), $\mu_{\text{ótimo}} = 0.0432$ ($M = 32$) and $\mu_{\text{ótimo}} = 0.0489$ ($M = 64$). O passo de adaptação considerado em para cada caso é $0.3\mu_{\text{ótimo}}$. A dispersão dos autovalores da matriz de autocorrelação do sinal de entrada é 132 e $\alpha = 1$. A Fig. 7.6 mostra os resultados para o EQM obtidos a partir de simulação MC e os obtidos do modelo proposto. Pode ser observado que a precisão do modelo melhora com o aumento da janela de observação.

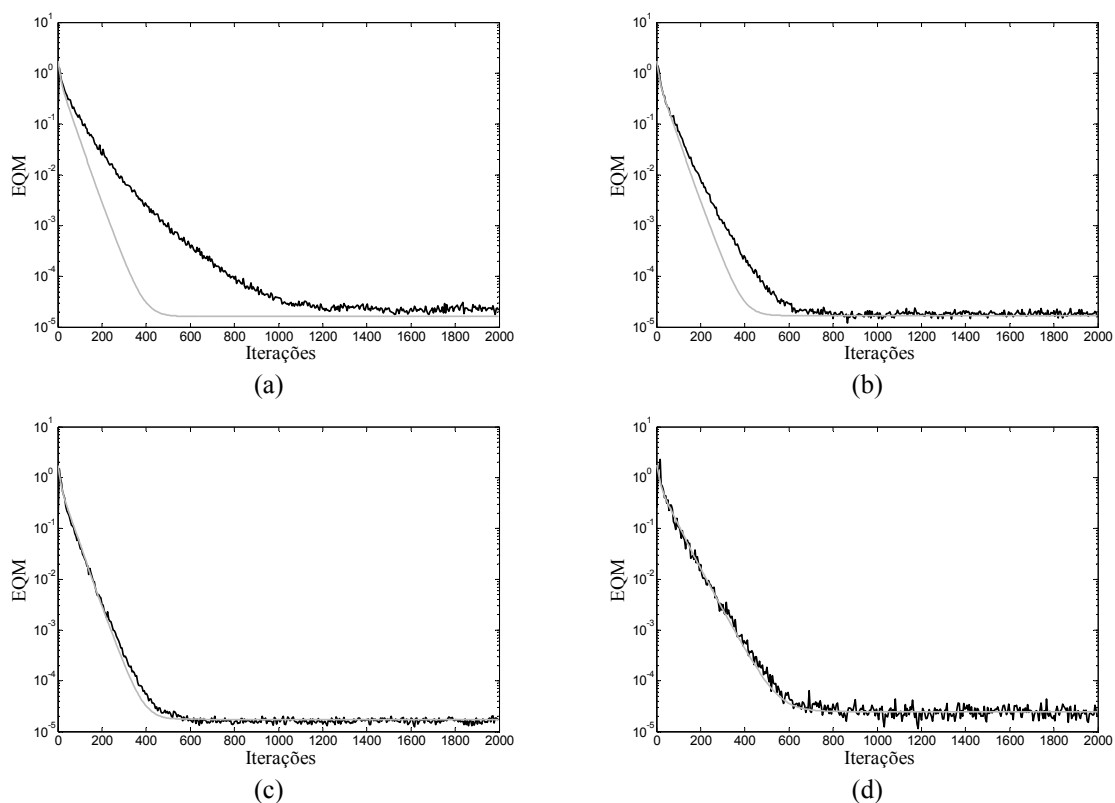
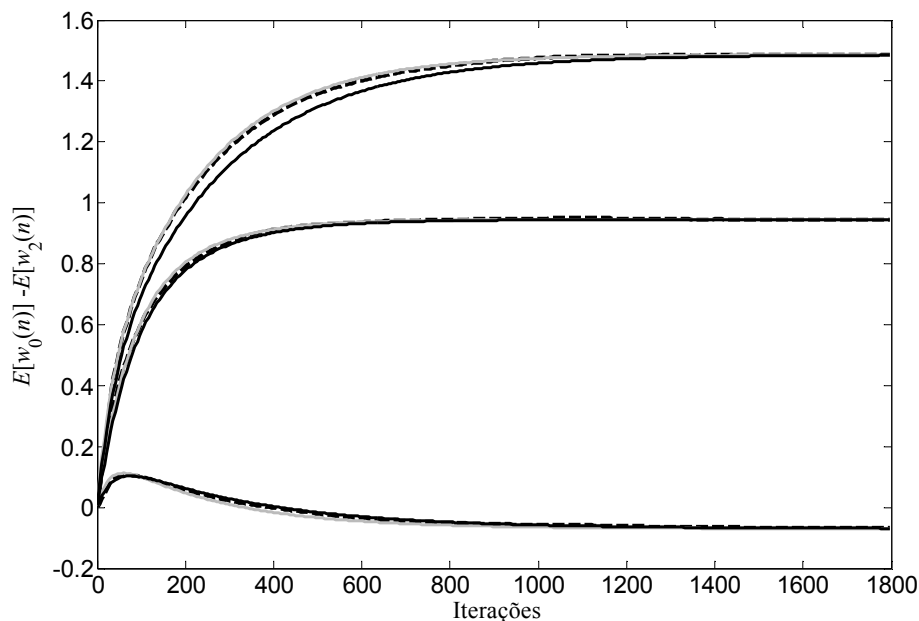


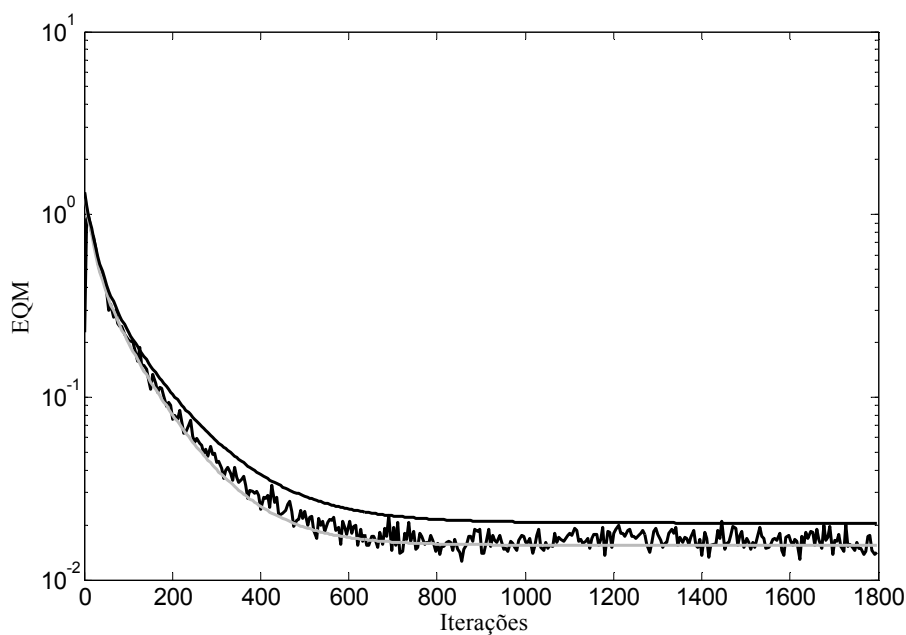
Figura 7.6- Exemplo 4. Curvas de EQM com uma dispersão dos autovalores igual a 132, $\alpha=1$ e $SNR=60\text{dB}$. (Linhas irregulares) simulações MC. (Linhas contínuas) modelo proposto Eq. (6.16). (a) $M=8$, (b) $M=16$, (c) $M=32$ e (d) $M=64$.

7.5- EXEMPLO 5

Para este exemplo foram usados os seguintes parâmetros: $N=8$ em (7.3), $M=32$, e $\mu_{\text{ótimo}}=0.0483$. O passo de adaptação usado é $0,1\mu_{\text{ótimo}}$. A dispersão dos autovalores da matriz de autocorrelação do sinal de entrada é 81 e o grau de não-estacionaridade em (6.31) é $\alpha=20$. A Fig. 7.7 compara os resultados obtidos para o momento de primeira ordem e para o EQM através de simulação MC, modelo apresentado em [30] e modelo proposto. Pode ser observado que, mesmo considerando-se um elevado grau de não-estacionaridade, o modelo proposto ainda apresenta excelentes resultados.



(a)



(b)

Figura 7.7- Exemplo 5. Resultados dos modelos com uma dispersão dos autovalores igual a 81 e $\alpha = 20$ para SNR de 60 dB e $0,1\mu_{\text{ótimo}}$. (a) Comportamento médio dos coeficientes $E[w_i(n)]$, $i = 0,1,2$. (Linha tracejada preta) simulação MC. (Linha contínua preta) modelo de [30]. (Linha contínua cinza-claro) modelo proposto. (b) Curvas de EQM. (Linhas irregulares) simulação MC. (Linhas contínuas preta) modelo de [30]. (Linhas contínuas cinza-claro) modelo proposto [Eq. (6.16)].

Tabela 7.1- Desajuste \mathcal{M} referente às simulações

		Desajuste \mathcal{M}	
		Experimental	Teórico
Exemplo 1	$0,3\mu_{\text{ótimo}}$	62,5	58,47
	$0,1\mu_{\text{ótimo}}$	178,5	157,18
Exemplo 2	$0,3\mu_{\text{ótimo}}$	16,5	15,80
	$0,1\mu_{\text{ótimo}}$	45,8	42,63
Exemplo 3	$0,3\mu_{\text{ótimo}}$	28,9	28,42
Exemplo 5	$0,1\mu_{\text{ótimo}}$	15689	15220

7.6- CONCLUSÕES

Este capítulo apresentou alguns exemplos considerando um problema de identificação de sistema através dos quais os modelos estatísticos do LMS-DT para ambientes não-estacionários foram avaliados.

O objetivo do Exemplo 1 foi verificar o comportamento do modelo para um grau de não-estacionaridade um pouco elevado e considerando as amostras do sinal de entrada fracamente correlacionadas. A partir das curvas apresentadas, considerando passos de adaptação de $0,1\mu_{\text{ótimo}}$ e $0,3\mu_{\text{ótimo}}$, pôde-se verificar uma boa concordância entre os resultados obtidos via simulação Monte Carlo e aqueles obtidos através do modelo proposto.

O Exemplo 2 adota os mesmos parâmetros do Exemplo 1, exceto o grau de não-estacionaridade. Pôde-se também verificar que os resultados obtidos a partir do modelo proposto concordam satisfatoriamente com aqueles de simulação Monte Carlo. No Exemplo 5, esse bom comportamento é novamente constatado. Com isso, pôde verificar que o modelo consegue prever adequadamente o comportamento médio dos coeficientes do filtro LMS-DT, bem como seu EQM, para diferentes graus de não-estacionaridade.

Com o Exemplo 3, buscou-se verificar o comportamento do modelo perante um sinal altamente correlacionado e uma maior ordem de filtro adaptativo. A partir dos

resultados para $0,1\mu_{\text{ótimo}}$ e $0,3\mu_{\text{ótimo}}$, pôde-se verificar um bom casamento entre os resultados obtidos via simulação e aqueles considerando o modelo proposto. Assim, pôde-se constatar que o modelo é satisfatório para diferentes graus de correlação das amostras do sinal de entrada.

O propósito do Exemplo 4 foi verificar o comportamento do modelo para diferentes valores de janela de observação M , usando-se um sinal de entrada fortemente correlacionado e grau de não-estacionaridade igual à unidade. Como foi visto, através das figuras desse exemplo, quanto maior o valor de M melhor a precisão dos resultados do modelo proposto. Contudo, observou-se que para $M > 32$ os resultados já tornam-se satisfatórios.

Com respeito à Tabela 7.1, referente ao desajuste, pôde-se observar uma precisão satisfatória entre os resultados teóricos e os obtidos via simulação, principalmente, para a condição de adaptação lenta e para menores graus de não-estacionaridade.

CAPÍTULO 8

COMENTÁRIOS E CONCLUSÕES FINAIS

Neste trabalho, foi desenvolvido um modelo estatístico mais preciso do algoritmo LMS-DT considerando ambientes estacionários como também não-estacionários. Como verificado na literatura da área, tal algoritmo apresenta para sinais de entrada correlacionados uma taxa de convergência superior à do algoritmo LMS convencional. Esse incremento de taxa de convergência é decorrente do pré-processamento do sinal de entrada obtido por uma transformação ortogonal seguida de uma normalização de potência. Outro ponto evidenciado na literatura é a baixa complexidade computacional do algoritmo LMS-DT quando comparada à do algoritmo RLS. Assim, o algoritmo LMS-DT se torna uma interessante alternativa para aplicações em que o sinal de entrada é correlacionado.

A análise estatística do algoritmo LMS-DT apresenta um tratamento matemático complexo quando é considerada a natureza variante do passo de adaptação decorrente da normalização de potência requerida no processo. A alternativa encontrada para simplificar a matemática envolvida nos cálculos dos valores esperados (requeridos pelo modelo), quando é considerado o passo de adaptação variante, é a utilização do Princípio da Média, permitindo assim isolar-se o termo de valor esperado $E[\mathbf{D}^{-1}(n)]$. No entanto, para que isso seja válido o tamanho da janela de observação, considerada no cálculo da potência média, deve ser grande ($M > 32$). Outra dificuldade encontrada, na análise estatística do referido algoritmo, reside na obtenção do valor esperado $E[\mathbf{D}^{-1}(n)]$ que leve em conta a

dependência estatística intrabanda proveniente do processo de transformação ortogonal. Um grande esforço foi despendido na obtenção de tal solução. Isso se deve ao fato de que o cálculo do requerido valor esperado envolve a solução de uma integral até então sem uma solução geral fechada. Tal integral é conhecida na literatura como integral hiperelíptica. A estratégia adotada para suplantiar tal problema foi a utilização de um método aproximado, apresentado em detalhes no Apêndice B. Através de extensivas simulações, verifica-se que, mesmo sendo uma aproximação, a abordagem escolhida para resolver tal integral conduz a excelentes resultados.

Uma vez superado o ponto crucial da análise estatística do referido algoritmo, pôde-se obter um modelo que representa o comportamento do momento de primeira e segunda ordens dos coeficientes do filtro LMS-DT. Para que essa análise fosse realizada, algumas hipóteses simplificativas foram adotadas, como por exemplo, sinal de entrada com distribuição de probabilidade Gaussiana e adaptação lenta.

Alguns exemplos foram apresentados com o intuito de se avaliar o modelo proposto para ambientes estacionários. A partir dos resultados, algumas conclusões podem ser tiradas:

- Para adaptação lenta, ambos os momentos de primeira e segunda ordens dos coeficientes do filtro conseguem prever satisfatoriamente o seu comportamento médio, tanto em regime permanente quanto em regime transitório;
- Para uma adaptação mais rápida, em regime transitório, há uma ligeira discordância entre os resultados obtidos via simulação e aqueles obtidos através do modelo proposto, resultado esse esperado, pois o modelo torna-se mais preciso à medida que o passo de adaptação diminui;
- Com respeito ao grau de correlação das amostras do sinal de entrada, pôde ser observado que o modelo estatístico apresenta resultados satisfatórios para diferentes graus de correlação;
- Outro ponto que deve ser ressaltado é o tamanho da janela de observação para o cálculo da potência média. Através dos exemplos apresentados, pôde ser verificado que o tamanho de janela usada determina a precisão do modelo

estatístico. Quanto maior for o tamanho da janela, melhor será a precisão do modelo. Esse resultado também era esperado, pois nessa condição o Princípio da Média se faz mais válido;

- Os valores de desajuste obtidos a partir do modelo correspondem apropriadamente aos obtidos através de simulação;
- Para todos os exemplos apresentados, o modelo proposto sempre levou a melhores resultados do que os modelos que não consideram haver dependência estatística intrabanda proveniente do processo de transformação ortogonal.

Assim como para os ambientes estacionários, algumas considerações foram adotadas também para ambientes não-estacionários, como por exemplo, sinal de entrada com distribuição Gaussiana e adaptação lenta.

Para validar o modelo proposto, alguns exemplos foram apresentados. Tomando-se por base os resultados obtidos a partir desses exemplos, chega-se às seguintes conclusões:

- Os momentos de primeira e segunda ordens dos coeficientes do filtro LMS-DT apresentaram resultados satisfatórios quando comparados com os aqueles obtidos por simulação MC;
- Pôde ser verificado que o comportamento do modelo estatístico é muito bom para diferentes graus de não-estacionaridade;
- Quanto ao grau de correlação das amostras do sinal de entrada, o comportamento do modelo estatístico proposto também é satisfatório;
- Com respeito ao tamanho da janela de observação para o cálculo da potência média, pôde ser observado, através das simulações, que quanto maior for o tamanho da janela maior será a precisão do modelo. Isso se deve ao fato de o Princípio da Média, nessas condições, ter maior validade;
- Os valores de desajuste obtidos a partir do modelo concordam satisfatoriamente com os obtidos via simulação numérica;

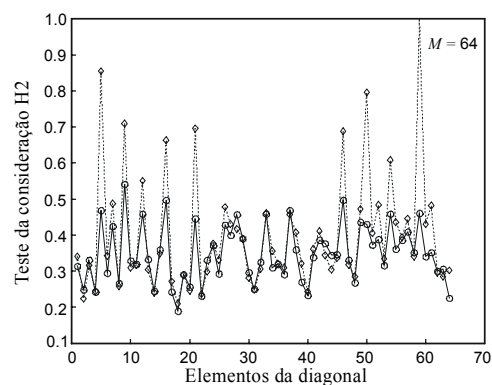
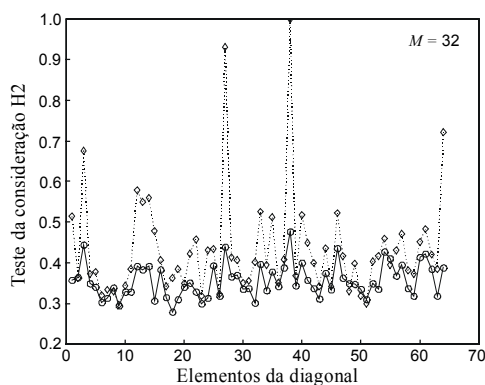
- Para todos os exemplos considerados, foi observado que o modelo aqui proposto apresentou melhores resultados em relação aos modelos que consideram a independência estatística intrabanda.

Uma proposta para a continuação deste trabalho tese seria calcular os valores esperados que envolvem o passo de adaptação variante no tempo sem envolver o uso do Princípio da Média, como também aprimorar os modelos aqui obtidos para uma condição de adaptação mais rápida.

APÊNDICE A

VERIFICAÇÃO DE (4.3)

A consideração H2 permite usar a aproximação $E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)] \approx E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$. Para verificar a precisão de tal aproximação, algumas simulações são apresentadas para diferentes valores de N e M . As simulações ilustradas nas Figs. A.1 e A.2 consideram $N = \{64, 128\}$ e $M = \{32, 64, 96, 128\}$ e sinal de entrada colorido com dispersão dos autovalores igual a 931. Nas figuras, os pontos marcados como (\circ) e (\diamond) representam os elementos da diagonal principal das matrizes $\mathbf{E}_1 = E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$ e $\mathbf{E}_2 = E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)]$, respectivamente. As linhas conectando os pontos servem apenas para melhorar a visualização das figuras. De tais curvas, um certo desvio pode ser verificado para todos os valores de M . Como esperado, tal desvio diminui à medida que M cresce. Não obstante, os efeitos de tal desvio nas predições do primeiro e segundo momentos é atenuado, uma vez que os termos em que (4.3) é utilizada são ponderados por μ ou μ^2 . Em adição, pode ser observado que para comprimentos de filtro adaptativo $N = 128$ e $M > 64$, é obtida uma aproximação mais aceitável e uma melhor estimativa de potência.



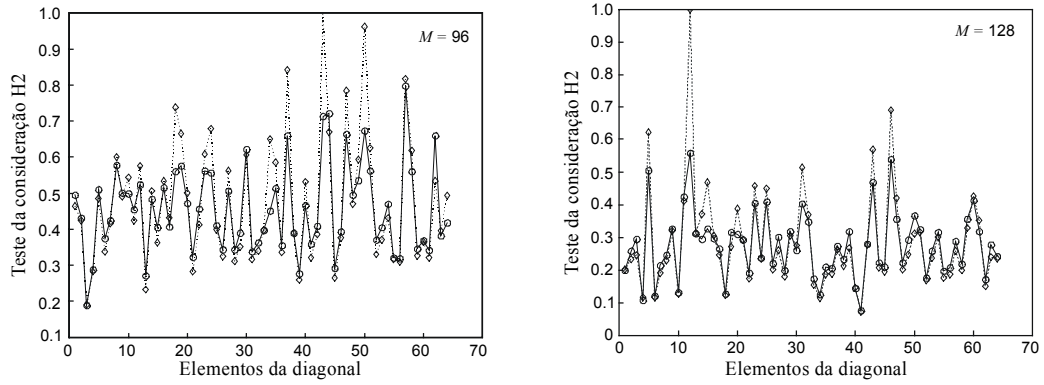


Figura A.1. Verificação da aproximação (4.3) para $N = 64$ e $M = \{32, 64, 96, 128\}$. (O)

$$\mathbf{E}_1 = E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)]. \quad (\diamond) \quad \mathbf{E}_2 = E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)].$$

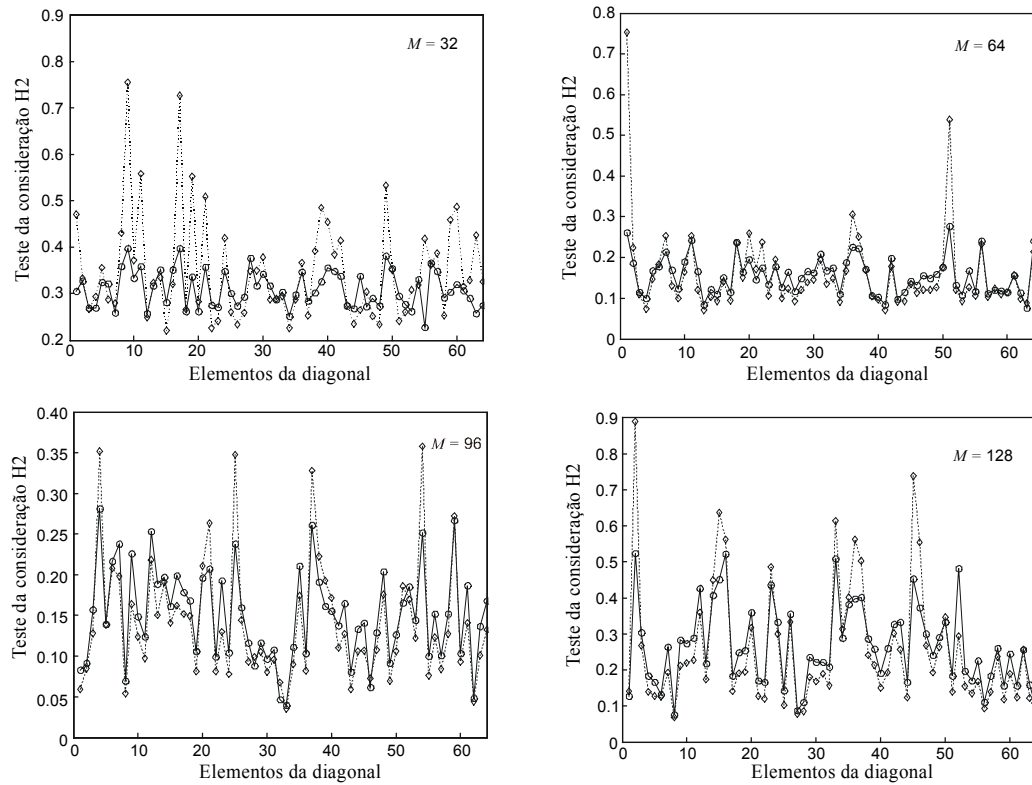


Figura A.2. Verificação da aproximação (4.3) para $N = 128$ e $M = \{32, 64, 96, 128\}$. (O)

$$\mathbf{E}_1 = E[\mathbf{D}^{-1}(n)\mathbf{x}_T(n)\mathbf{x}_T^T(n)] \text{ e } (\diamond) \quad \mathbf{E}_2 = E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_T(n)\mathbf{x}_T^T(n)].$$

APÊNDICE B

DETERMINAÇÃO DE $E \left[\frac{1}{d_i(n)} \right]$

O valor esperado desejado é obtido resolvendo-se a seguinte integral múltipla:

$$E \left\{ \frac{M^p}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^p} \right\} = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{M \text{ fold}} \frac{M^p}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^p} f(\mathbf{x}_i) d\mathbf{x}_i, \quad (\text{B.1})$$

onde $f(\mathbf{x}_i)$ representa a função densidade multivariada Gaussiana do vetor de entrada $\mathbf{x}_i(n)$, dada por

$$f(\mathbf{x}_i) = \frac{1}{(2\pi)^{M/2} [\det(\mathbf{R}_{\mathbf{x}_i})]^{1/2}} e^{-\frac{\mathbf{x}_i^T \mathbf{R}_{\mathbf{x}_i}^{-1} \mathbf{x}_i}{2}}, \quad (\text{B.2})$$

$\det(\cdot)$ denota o determinante da matriz e $\mathbf{R}_{\mathbf{x}_i}$ é a matriz de autocorrelação do sinal transformado correspondente à i -ésima sub-banda. Substituindo-se (B.2) em (B.1) e definindo-se a função $F(\omega)$ como

$$F(\omega) = \frac{M^p}{(2\pi)^{M/2} [\det(\mathbf{R}_{\mathbf{x}_i})]^{1/2}} \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{M \text{ fold}} \frac{1}{(\mathbf{x}_i^T \mathbf{x}_i)^p} e^{-\omega(\mathbf{x}_i^T \mathbf{x}_i)} e^{-\frac{\mathbf{x}_i^T \mathbf{R}_{\mathbf{x}_i}^{-1} \mathbf{x}_i}{2}} d\mathbf{x}_i, \quad (\text{B.3})$$

tem-se que o valor esperado (B.1) é obtido de (B.3) fazendo $\omega = 0$. Assim,

$$E \left\{ \frac{M^p}{[\mathbf{x}_i^T(n)\mathbf{x}_i(n)]^p} \right\} = F(\omega)|_{\omega=0} = F(0). \quad (\text{B.4})$$

Agora, diferenciando-se p vezes (B.3) e através de algumas manipulações algébricas,

usando a propriedade da função densidade que diz que $\underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{M \text{ fold}} f(\mathbf{x}_i) d\mathbf{x}_i = 1$, obtém-se

$$\frac{d^p F(\omega)}{d\omega^p} = \frac{(-1)^p M^p}{\sqrt{\det(\mathbf{B}^{-1}\mathbf{R}_{\mathbf{x}_i})}} \quad (\text{B.5})$$

para

$$\mathbf{B}^{-1} = 2\omega\mathbf{I} + \mathbf{R}_{\mathbf{x}_i}^{-1}, \quad (\text{B.6})$$

onde \mathbf{I} representa a matriz identidade. Então, substituindo-se (B.6) em (B.5) e integrando p vezes, (B.3) torna-se

$$F(\omega) = (-1)^p M^p \underbrace{\int \dots \int}_{p \text{ fold}} \frac{1}{[\det(\mathbf{I} + 2\omega\mathbf{R}_{\mathbf{x}_i})]^{1/2}} \underbrace{d\omega \dots d\omega}_{p \text{ fold}} + C. \quad (\text{B.7})$$

O valor da constante C é obtido da condição $\lim_{\omega \rightarrow \infty} F(\omega) = 0$. O ponto central agora é resolver a integral hiperelíptica (B.7). Para tal, faz-se uso de uma forma particular da matriz de autocorrelação $\mathbf{R}_{\mathbf{x}_i}$ e considera-se os seguintes passos:

- i) Decompor $\mathbf{R}_{\mathbf{x}_i}$ como $\mathbf{R}_{\mathbf{x}_i} = \mathbf{Q}_{\mathbf{x}_i} \mathbf{\Lambda}_{\mathbf{x}_i} \mathbf{Q}_{\mathbf{x}_i}^T$, onde $\mathbf{Q}_{\mathbf{x}_i}$ é a matriz dos autovetores de $\mathbf{R}_{\mathbf{x}_i}$ e $\mathbf{\Lambda}_{\mathbf{x}_i}$ é uma matriz diagonal que contém os autovalores λ_i de $\mathbf{R}_{\mathbf{x}_i}$. O denominador de (B.7) é então escrito como um polinômio de grau M . Assim,

$$\det(\mathbf{I} + 2\omega\mathbf{R}_{x_i}) = \prod_{k=1}^M (1 + 2\omega\lambda_k), \quad (\text{B.8})$$

onde o coeficiente a_M do polinômio correspondente ao termo de maior grau ω^M é dado por $a_M = 2^M \prod_{k=1}^M \lambda_k$ e as raízes desse polinômio são obtidas a partir de $\omega_k = -1/(2\lambda_k)$ para $k = 1, 2, \dots, M$.

ii) Substituir cada par de raízes adjacentes de (B.8) pelo negativo correspondente a sua média geométrica. Assim,

$$\omega'_q = -\sqrt{\omega_{2q-1}\omega_{2q}}, \quad q = 1, 2, \dots, M/2 \text{ e } M \text{ par}. \quad (\text{B.9})$$

Deste modo, agora (B.8) possui raízes com multiplicidade 2 e a raiz quadrada em (B.7) pode então ser eliminada. O passo (ii) é o ponto crucial para permitir a computação de uma integral Abeliana de alta ordem. Assim, (B.7), para $p = 1$, pode ser agora aproximada por

$$\int \frac{-M d\omega}{[\det(\mathbf{I} + \beta\mathbf{R}_{x_i})]^{1/2}} \approx \frac{-M}{\sqrt{a_M}} \int \frac{d\omega}{(\omega - \omega'_1) \cdots (\omega - \omega'_{M/2})} + C. \quad (\text{B.10})$$

iii) Agora, a integral de (B.10) pode ser facilmente resolvida através de uma expansão em frações parciais, resultando em

$$\int \frac{-M d\omega}{[\det(\mathbf{I} + \omega\mathbf{R}_{x_i})]^{1/2}} \approx \frac{-M}{\sqrt{a_M}} \sum_{q=1}^{M/2} A_q \ln(\omega - \omega'_q) + C, \quad (\text{B.11})$$

onde

$$A_q = \frac{1}{\prod_{\substack{j=1 \\ j \neq q}}^{M/2} (\omega'_q - \omega'_j)}, \quad q = 1, 2, \dots, M/2. \quad (\text{B.12})$$

Usando-se o método proposto, os valores esperados requeridos para a análise estatística do algoritmo LMS-DT são agora dados como segue:

para $p = 1$ em (B.7) tem-se

$$E \left[\frac{M}{\sum_{k=0}^{M-1} x_i^2(n-k)} \right] = F(\omega)|_{\omega=0} \approx \frac{-M}{\sqrt{a_M}} \sum_{q=1}^{M/2} A_q \ln(-\omega'_q), \quad (\text{B.13})$$

e, similarmente, para $p = 2$

$$E \left\{ \frac{M^2}{\left[\sum_{k=0}^{M-1} x_i^2(n-k) \right]^2} \right\} \approx \frac{M^2}{\sqrt{a_M}} \sum_{q=1}^{M/2} A_q [-\omega'_q \ln(-\omega'_q) + \omega'_q]. \quad (\text{B.14})$$

As expressões (B.13) e (B.14) são novos resultados para determinação do valor esperado de (4.7).

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] S. Haykin, Adaptive Filter Theory, 3^a ed. Englewood Cliffs, NJ. Ed. Prentice-Hall, 1996.
- [2] B. Farhang-Boroujeny, Adaptive Filters: Theory and Applications, Matlab Exercises, National University of Singapore, John Wiley & Sons, 1998.
- [3] B. Widrow and S.D. Stearns, Adaptive Signal Processing, Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [4] S.S. Narayan, A.M. Peterson and M.J. Narasimha, "Transform Domain LMS Algorithm," IEEE Transactions on Acoustics, Speech and Signal Processing, 31(3), 1983, pp. 609-615.
- [5] D.F. Marshall, W.K. Jenkins and J.J. Murphy, "The use of orthogonal transforms for improving performance of adaptive filters," IEEE Trans. on Circuits and Systems, 36(4), 1989, pp. 474-484.
- [6] F. Beaufays, "Transform-Domain Adaptive Filters: an Analytical Approach," IEEE Trans. On Signal Processing, 43(2), 1995, pp. 422-431.
- [7] B. Farhang-Boroujeny, "Order of N Complexity Transform domain adaptive filters," IEEE International Symposium on Circuits and Systems, vol.2, 1994, pp. 37-40.
- [8] W.K. Jenkins and J.K. Yun, "Analysis of Fixed Point Roundoff Effects in Transform Domain LMS Adaptive Filters," European Conference on Circuit Theory and Design, 1989, pp. 228-232.
- [9] S. Rahardja and S. Falkowski, "Comparative Study of Discrete Orthogonal Transforms in Adaptive Signal Processing," IEICE Transactions Fundamentals, E82-A (8), 1999, pp. 1386-1390.
- [10] B. Farhang-Boroujeny and S. Gazor, "Selection of Orthonormal Transforms for Improving the Performance of the Transform Domain Normalised LMS Algorithm," IEE Proceedings-F, vol.139, no. 5, Oct. 1992, pp. 327-335..
- [11] B. Farhang-Boroujeny and S. Gazor, "Performance of LMS-Based Adaptive Filters in Tracking a Time-Varying Plant," IEEE Transactions on Signal Processing, vol.44, Nov. 1996, pp. 2868-2871.
- [12] B. Farhang-Boroujeny and S. Gazor, "Performance Analysis of Transform Domain

- Normalized LMS Algorithm,” International Conference on Acoustics, Speech and Signal Processing, vol.3, Apr. 1991, pp. 2133-2136.
- [13] J. Lee and C. Un, “Performance of Transform-Domain LMS Adaptive Digital Filters,” IEEE Transactions on Acoustics, Speech and Signal Processing, vol.34, Jun. 1986, pp. 499-510.
- [14] J. Lee and C. Un, “On the Convergence Behaviour of Frequency-Domain LMS Adaptive Digital Filters,” IEEE International Conference on Acoustics, Speech and Signal Processing, vol.9, Mar. 1984, pp. 467-470.
- [15] B. Farhang-Boroujeny, “Quantization Effects in Transform Domain Adaptive Filters,” International Conference on Circuits and Systems, ‘Communications on the Move’, vol.1, Nov. 1992, pp. 25-29.
- [16] B. Farhang-Boroujeny and S. Gazor, “Quantization Effects in Transform-Domain LMS Algorithm,” IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol.39, Jan. 1992, pp. 1-7.
- [17] B. Widrow and E. Walach, “On the Statistical Efficiency of the LMS Algorithm with Nonstationary Inputs,” IEEE Transactions on Information Theory, vol. 30 (2), Mar. 1984, pp. 211-221.
- [18] Z. Zhongwei and M.Tahemezhadi. “Transform Domain LMS Algorithm in Fullband and Subband Based Acoustic Echo Cancellation,” IEEE 39th Midwest Symposium on Circuits and Systems, Vol. 2, Aug. 1996, pp. 795-798.
- [19] M. J. Medley, G. J. Saulnier and P. K. Das, “Interference Suppression Using Transform Domain LMS Adaptive Filtering,” IEEE Digital Signal Processing Workshop Proceedings, Sep. 1996, pp. 117 –120.
- [20] M. J. Medley and G. J. Saulnier, “Transform Domain LMS Convergence in Tone Interference,” MILCOM 97 Proceedings, Vol. 2, Nov 1997, pp. 872-876.
- [21] N. J. Bershad, “Analysis of the Normalized LMS Algorithm with Gaussian Inputs,” IEEE Transactions on Acoustic, Speech, and Processing, Vol. 34 (4), Agosto 1986, pp.793-806.
- [22] M. Tarrab and A. Feuer, “Convergence and Performance Analysis of the Normalized LMS algorithm with Uncorrelated Gaussian Data”, IEEE Transactions on Information Theory, Vol. 34 (4), Julho 1988, pp. 680-691.
- [23] M. H. Costa and J.C.M. Bermudez, “An Improved Model for the Normalized LMS Algorithm with Gaussian Inputs and Large Number of Coefficients,” IEEE

- International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, 2002, pp. 1385-1388.
- [24] S.J. Chern, J.C. Horng and K.M. Wong, "The Performance of the Hybrid LMS Adaptive Algorithm," Elsevier Signal Processing 44, 1995, pp. 67-88.
- [25] S. Hosur and A.H. Tewfik, "Wavelet Transform Domain Adaptive FIR Filtering," IEEE Transactions on Signal Processing, vol. 45 (3), 1997, pp. 617-630.
- [26] D. I. Kim and P. De Wilde, "Performance Analysis of the DCT-LMS Adaptive Filtering Algorithm," Elsevier Signal Processing, vol. 80, 2000, pp. 1629-1654.
- [27] E. M. Lobato, O. J. Tobias, R. Seara, "Análise Estatística do Algoritmo LMS no Domínio Transformado". XX Simpósio Brasileiro de Telecomunicações - SBT 2003. pp. 1-6, Rio de Janeiro, RJ, Outubro 2003.
- [28] E. M. Lobato, O. J. Tobias, R. Seara, "Modelo Analítico do Algoritmo ϵ LMS no Domínio Transformado para Sinais de Entrada Gaussianos". XXI Simpósio Brasileiro de Telecomunicações - SBT 2004. pp. 1-6, Belém, PA, Setembro 2004.
- [29] E. M. Lobato, O. J. Tobias, R. Seara, "A Stochastic Model for the Transform-Domain LMS Algorithm". 12th European Signal Processing Conference – EUSIPCO 2004. pp. 1833-1836, Viena, Áustria, Setembro 2004.
- [30] E. M. Lobato, O. J. Tobias, R. Seara, "Stochastic Modelling of the Transform Domain ϵ LMS Algorithm for a Time-Varying Environment". 13th European Signal Processing Conference (EUSIPCO 2005). pp. 1-4, Antalya, Turquia, Setembro 2005.
- [31] E. M. Lobato, O. J. Tobias, R. Seara, "Análise Estatística do Algoritmo ϵ LMS no Domínio Transformado em Ambientes Não-estacionários". XXII Simpósio Brasileiro de Telecomunicações (SBrT 2005). pp. 817-822, Campinas, SP, Setembro 2005.
- [32] A. Feuer and E. Weinstein, "Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data", IEEE Trans. on Acoust., Speech and Signal Processing, 33 (1), 1985, pp. 222-230.
- [33] K. Mayyas, "Mean-Square Analysis of a Variable Step Size Transform Domain LMS Adaptive Algorithm", Processings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, 2004.
- [34] E. M. Lobato, O. J. Tobias, R. Seara, " Stochastic Model for the NLMS Algorithm

with Correlated Gaussian Data”. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006). Vol III, pp. 760-763, Toulouse, France, Maio 2006.

- [35] J. L. Boldrini et al., *Álgebra Linear*. 3^a ed., Editora Harbra, 2001.
- [36] S. Haykin and B.V. Veen, *Sinais e Sistemas*, 1^a ed., Bookman Press, 2001.
- [37] C.G. Samson and U. Reddy, “Fixed Point Error Analysis of the Normalized Ladder Algorithm,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(5), 1983, pp. 1177-1191.
- [38] A. Papoullis, *Probability, Random Variables and Stochastic Processes*, 3^a ed., McGraw-Hill, 1991.
- [39] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Sixth ed., Academic Press, 2000
- [40] A. Feuer and E. Weinstein, “Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data”, *IEEE Trans. on Acoust., Speech and Signal Processing*, 33 (1), 1985, pp. 222-230.