

ORLANDO JOSÉ TOBIAS

**ANÁLISE ESTATÍSTICA DO COMPORTAMENTO
DO ALGORITMO LMS FILTRADO**

**FLORIANÓPOLIS
ANO 1999**

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

**ANÁLISE ESTATÍSTICA DO
COMPORTAMENTO DO ALGORITMO
LMS FILTRADO**

Tese submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para a
obtenção do grau de Doutor em Engenharia Elétrica.

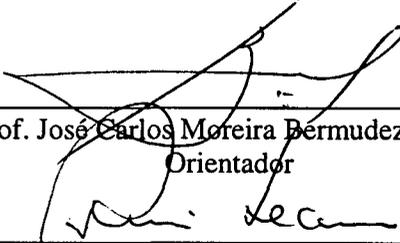
ORLANDO JOSÉ TOBIAS

Florianópolis, Outubro de 1999.

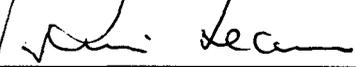
ANÁLISE ESTATÍSTICA DO COMPORTAMENTO DO ALGORITMO LMS FILTRADO

Orlando José Tobias

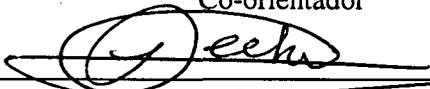
‘Esta Tese foi julgada adequada para obtenção do Título de Doutor em Engenharia Elétrica, Área de Concentração em *Sistemas de Informação*, e aprovada em sua forma final pelo Curso de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.’



Prof. José Carlos Moreira Bermudez, Ph.D.
Orientador

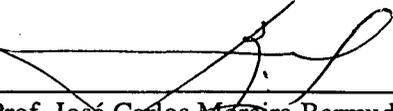


Prof. Rui Seara, Dr.
Co-orientador

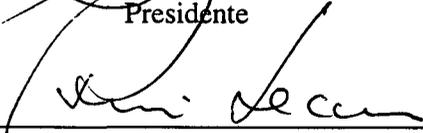


Prof. Idemar Cassana Decker, D.Sc.
Coordenador do Curso de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:



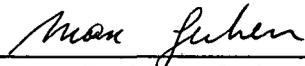
Prof. José Carlos Moreira Bermudez, Ph.D.
Presidente



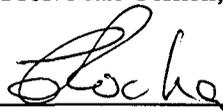
Prof. Rui Seara, Dr.



Prof. José João de Espíndola, Ph.D.



Prof. Max Gerken, Dr.



Prof. Carlos Aurélio Faria da Rocha, Dr.



Prof. Vitor H. Nascimento, Ph.D.

AGRADECIMENTOS

A meus pais Catalina e Orlando que, com enorme amor e sacrifício, fizeram com que eu pudesse chegar onde cheguei.

A Patrícia e Diego pela alegria de viver.

À sociedade brasileira, representada pela CAPES, pela concessão de uma bolsa de estudos, permitindo assim a realização deste trabalho.

Aos amigos do LINSE pela agradável convivência.

Aos professores Bermudez e Rui Seara pela orientação.

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

ANÁLISE ESTATÍSTICA DO COMPORTAMENTO DO ALGORITMO LMS FILTRADO

Orlando José Tobias

Outubro/1999

Orientador: José Carlos Moreira Bermudez, PhD.

Co-orientador: Rui Seara, Dr.

Área de Concentração: Sistemas de Informação.

Palavras-chave: Algoritmos adaptativos, Algoritmo LMS filtrado, Controle ativo de ruído.

Número de Páginas: 135.

RESUMO: Neste trabalho, é apresentada uma análise do comportamento estocástico do algoritmo LMS filtrado (FX-LMS). É obtida uma expressão recursiva para os momentos de primeira e de segunda ordem dos coeficientes do filtro adaptativo para sem a consideração de adaptação lenta. O modelo proposto não depende do tipo do sinal de referência, podendo ser sendo ruído branco ou colorido. Também, não é utilizada a clássica suposição da teoria da independência. A solução de Wiener é determinada explicitamente em função das estatísticas da entrada e as respostas ao impulso dos caminhos primário e secundário. É mostrado, também, que em regime permanente o valor esperado dos coeficientes do filtro convergem para a solução de Wiener para o caso em que o caminho secundário é estimado sem erro. Simulações mostram um excelente casamento com o comportamento predito pelo modelo teórico proposto.

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 BEHAVIOR OF THE FILTERED-LMS ALGORITHM

Orlando José Tobias

October / 1999

Advisor: José Carlos Moreira Bermudez, PhD.

Co-advisor: Rui Seara, Dr.

Area of Concentration: Information systems.

Keywords: Adaptive algorithms, Filtered LMS algorithm, Active noise control.

Number of Pages: 135.

ABSTRACT: In this work is presented an statistical analysis of the filtered-LMS algorithm (FX-LMS). We derive recursive expressions for the first and second moments of the filter weights without the common assumption of slow adaptation. The proposed model does not depend of the nature of the reference signal, white noise or a correlated process. Also, the model is derived without invoking the classical assumption of the independence theory. The optimum weight vector is determined as a function of the input signal statistics and the impulse responses of the primary and secondary paths. Simulation results show a very close agreement with the behavior predicted by the theoretical model.

SUMÁRIO

1. INTRODUÇÃO	1
MOTIVAÇÃO DO PRESENTE TRABALHO	1
1.1. INTRODUÇÃO AO CONTROLE DE RUÍDO ACÚSTICO	2
1.2. ADAPTABILIDADE	5
1.3. ALGORITMO ADAPTATIVO	6
1.4. ALGORITMO LMS FILTRADO	6
1.5. ESTIMAÇÃO DE S	8
1.6. ESTADO ATUAL DA ARTE	10
1.7. ORGANIZAÇÃO DA TESE	13
2. MOMENTO DE PRIMEIRA ORDEM	14
INTRODUÇÃO	14
2.1. CONSIDERAÇÕES INICIAIS - DIAGRAMA EM BLOCOS.	14
2.2. EQUACIONAMENTO DO SISTEMA.	17
2.2.1. Sinal de erro.	17
2.2.2. Atualização dos coeficientes do filtro adaptativo.	17
2.3. SOLUÇÃO ÓTIMA.	18
2.3.1. Princípio da ortogonalidade.	22
2.3.2. Superfície de erro – Mínimo EMQ	23
2.3.3. Forma canônica da superfície de erro.	25
2.4. CÁLCULO DO MOMENTO DE PRIMEIRA ORDEM	27
2.4.1. Redução ao LMS clássico	29

2.5. REGIME PERMANENTE QUANDO $\hat{S} \neq S$	30
2.5.1. Estimação do erro em excesso para o regime permanente.	31
2.6. VETOR DE ERRO NOS COEFICIENTES.	32
2.7. EFEITO DOS ERRO DE ESTIMAÇÃO DO CAMINHO SECUNDÁRIO.	34
2.8. RESULTADOS DE SIMULAÇÕES.	36
2.8.1. Avaliação da suposição (II).	36
2.8.2. Avaliação do modelo: Caso 1.	46
2.8.2. Avaliação do modelo: Caso 2.	49
2.8.2. Avaliação do modelo: Caso 3.	52
2.9. CONCLUSÕES.	55
3. MOMENTO DE SEGUNDA ORDEM	56
INTRODUÇÃO.	56
3.1. ALGORITMO DLMS.	57
3.1.1. Sinal de erro.	57
3.1.2. Equação de atualização dos coeficientes.	58
3.1.3. Vetor de coeficientes ótimo.	59
3.1.4. Erro médio quadrático.	59
3.1.5. Valor em regime permanente do vetor de coeficientes.	60
3.1.6. Vetor erro nos coeficientes em relação a \mathbf{W}_∞	61
3.1.7. Estudo do comportamento médio dos coeficientes quando $\hat{D} \neq D$	61
3.1.8. Erro médio quadrático em função de $\mathbf{V}(n)$	63
3.1.9. Cálculo de $E[\mathbf{V}(n-D)\mathbf{V}^T(n-D)]$	63
3.1.10. Estudo do caso com estimação perfeita do atraso.	70

3.1.11. Estudo da estabilidade do algoritmo DLMS.	75
3.1.12. Análise do caso geral ($\hat{D} \neq D$).	77
3.1.13. Análise do DLMS utilizando a teoria da independência.	78
3.1.14. Conversão do caso sob análise no caso estudado por Long et al.	79
3.2. RESULTADOS DE SIMULAÇÕES PARA O ALGORITMO DLMS.	84
3.2.1. Simulação 1.	84
3.2.2. Simulação 2.	86
3.2.3. Simulação 3.	86
3.2.4. Simulação 4.	88
3.3. CASO GERAL.	90
3.3.1. Sinal de erro.	91
3.3.2. Erro médio quadrático.	91
3.4. RESULTADOS EXPERIMENTAIS.	94
3.4.1. Simulação 1.	94
3.4.2. Simulação 2.	96
3.4.3. Simulação 3.	97
3.4.4. Simulação 4.	98
3.5. CONCLUSÕES.	99
4. APLICAÇÕES	101
Exemplo 4.1.	101
Exemplo 4.2.	105
Exemplo 4.3.	108
Exemplo 4.4.	112
Exemplo 4.5.	116

5. CONCLUSÕES	119
6. REFERÊNCIAS BIBLIOGRÁFICAS	121
APÊNDICE A	123
APÊNDICE B	125
Matriz de correlação.	125
Simplificação da Equação (b.4).	128
Determinação da Equação associada à (b.4).	129
APÊNDICE C	125
Cálculo do momento de quarta ordem de uma variável gaussiana. . .	133

Introdução

1

MOTIVAÇÃO DO PRESENTE TRABALHO

O princípio básico de funcionamento de um filtro adaptativo é a capacidade de variar seus parâmetros, através de um algoritmo de controle, de forma que convirjam a uma solução que minimize uma determinada função custo. A ampla aceitação dos filtros adaptativos é evidenciada pela vasta literatura pertinente, como também pelas mais diversas situações em que estes vem sendo aplicados. Dentre os algoritmos de controle o mais utilizado é, sem dúvidas, o algoritmo LMS (Least Mean-Square), principalmente pela sua simplicidade e fácil implementação.

Nos últimos anos, a área de controle ativo de ruído acústico e vibrações vem recebendo muita atenção por parte de pesquisadores e empresas envolvidas com este problema. O desempenho de um sistema de controle ativo é determinado, fundamentalmente, pelo algoritmo de controle ou algoritmo de adaptação. A efetividade de um algoritmo adaptativo depende de sua adequação ao problema sendo tratado e de um bom projeto. Tanto a adequação ao problema quanto a qualidade do projeto dependem, no entanto, do conhecimento que o projetista tem das propriedades de cada algoritmo considerado. Em geral, este conhecimento é obtido através de modelos analíticos que permitem prever o comportamento do sistema, ao invés de avaliar seu desempenho por meios puramente experimentais. Modelos analíticos permitem a inferência sobre o funcionamento do algoritmo em diversas situações, o estudo sobre a adequação do algoritmo à aplicação, a comparação dos desempenhos de diferentes algoritmos e o estudo da estabilidade. Portanto, tais modelos viabilizam projetos mais rápidos e melhores.

Uma variante do algoritmo LMS, o LMS filtrado ou FXLMS (Filtered X LMS), é o algoritmo mais utilizado em aplicações de controle ativo. Neste caso, o sinal de referência é filtrado para compensar os efeitos de operações de filtragem presentes no caminho de adaptação. Essas filtragens modificam substancialmente o comportamento do algoritmo.

Por este motivo, as clássicas e consagradas aproximações matemáticas que permitem analisar o algoritmo LMS não são extensíveis ao algoritmo FXLMS.

Neste trabalho apresentamos uma análise estatística do comportamento do algoritmo LMS filtrado. Esta análise dá origem a um modelo analítico que permite prever o comportamento do algoritmo FXLMS. A abordagem utilizada para analisar este algoritmo, e que representa uma alternativa importante em relação às abordagens usuais, foi a consideração de que a operação de filtragem no caminho de adaptação enfatiza a importância da correlação entre vetores de entrada consecutivos. Essa abordagem contradiz claramente a suposição clássica, utilizada para a análise do algoritmo LMS, que inclui-se na chamada teoria da independência [6]. Nessa teoria, considera-se que diferentes vetores de entrada são estatisticamente independentes. Tal consideração fornece muito bons resultados para o algoritmo LMS clássico. No entanto, a simples extensão para o caso filtrado leva a resultados que não refletem o comportamento real do novo algoritmo. Simulações numéricas e suas previsões usando o modelo proposto, validam a análise proposta. Os modelos analíticos obtidos para o comportamento dos momentos de primeira e segunda ordem dos coeficientes do filtro adaptativo constituem-se na principal contribuição desta tese.

1.1 INTRODUÇÃO AO CONTROLE ATIVO DE RUÍDO ACÚSTICO

O termo Controle Ativo de Ruído Acústico¹ descreve a supressão de um campo sonoro não desejado através da superposição com outro campo sonoro de igual magnitude porém de fase invertida. Esta idéia não é nova, sendo Paul Lueg [1] o autor do primeiro artigo publicado em 1936 sobre este assunto. Através dos anos, esta área tem recebido muita atenção e diversas implementações da idéia original já foram exploradas. Algumas destas podem ser encontradas em artigos [2] e livros [3-4, 25].

Para entender o porquê da atenção recebida, é importante compreender as vantagens do controle ativo quando confrontado com as técnicas tradicionais de redução de ruído (controle passivo). Nas técnicas passivas faz-se uso de materiais absorventes. Estes materiais são muito efetivos para altas frequências, geralmente acima de 150 Hz. No

¹ Nesta seção estamos nos referindo ao controle ativo de ruído acústico, no entanto o caso de controle ativo de vibrações é completamente equivalente em relação aos princípios básicos. As principais diferenças situam-se nos transdutores utilizados.

entanto, para obter-se uma apreciável redução do ruído, as dimensões dos dispositivos devem ser da ordem do comprimento da onda da freqüência a ser atenuada. Assim, para freqüências inferiores a 150 Hz, deve-se utilizar dispositivos de tamanho e peso excessivos, tornando esta solução muito onerosa e, às vezes, impraticável. Com o advento das novas gerações de processadores digitais, as técnicas de controle ativo de ruído (e vibrações) tornaram-se possíveis. Sua utilização concentra-se na faixa de baixas freqüências, já que o controle ativo de ruído para altas freqüências é, no momento, de implementação difícil, devido principalmente à alta velocidade requerida para processar a grande quantidade de informação. Além disso, nesta faixa de freqüências o controle passivo é uma solução consagrada pela sua eficácia e baixo custo. O controle ativo em baixas freqüências representa uma solução viável e de baixo custo para a maioria dos casos práticos, vindo preencher o espaço para o qual o controle clássico é deficiente.

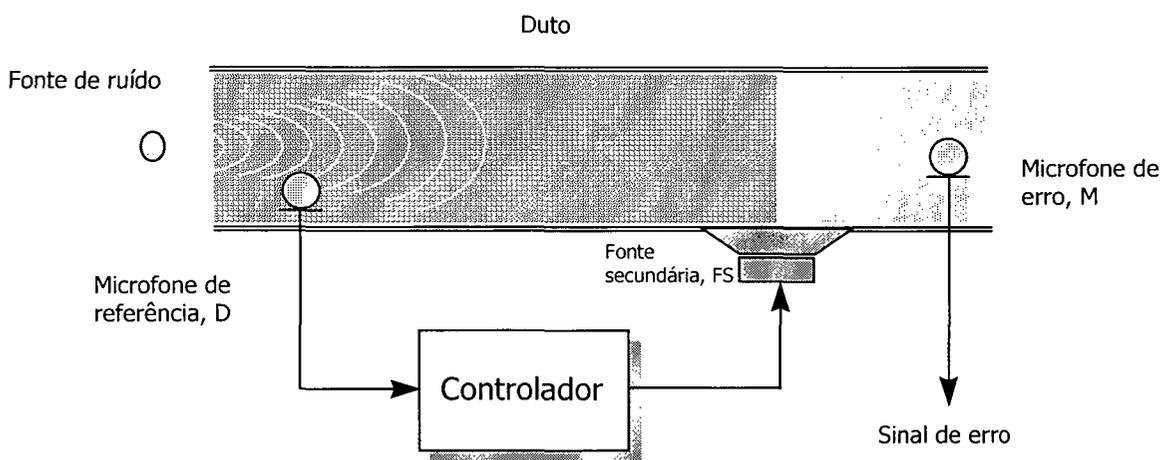


Fig. 1.1. Diagrama esquemático do sistema de controle ativo de ruído em dutos.

A Fig. 1.1 apresenta o diagrama esquemático de um sistema de controle ativo de ruído acústico em um duto. Em geral, um sistema de controle ativo para esta aplicação consiste de um transdutor D (que pode ser um microfone, chamado de *microfone de referência*) para adquirir o sinal do ruído que deseja-se cancelar (proveniente da fonte primária); um sistema de processamento ou controlador para tratar o sinal adquirido pelo transdutor D e produzir o sinal que alimenta uma fonte secundária FS (ou sistema de fontes secundárias, dependendo da aplicação), que é geralmente um alto-falante que irradia o *anti-ruído*. Se o sinal recebido pelo microfone de referência for correlacionado com o ruído originado na

fonte primária, o sinal de referência fornece a informação necessária ao controlador sobre o ruído primário antes que este chegue ao ponto de cancelamento, sendo esta uma condição necessária para o caso de um controlador causal. O campo sonoro residual é monitorado, à direita da fonte secundária, através de um outro transdutor M (ou *microfone de erro*). Nesta aplicação, além da redução de custo, este sistema possui a vantagem de não obstruir o fluxo normal do ar. Outras aplicações em que esta técnica vem sendo utilizada com sucesso são: redução de ruído em cabina de aviões e no interior de carros, capacetes e fones de ouvido.

Em geral, sistemas como o da Fig. 1.1 podem ser representados esquematicamente pelo seguinte diagrama em blocos:

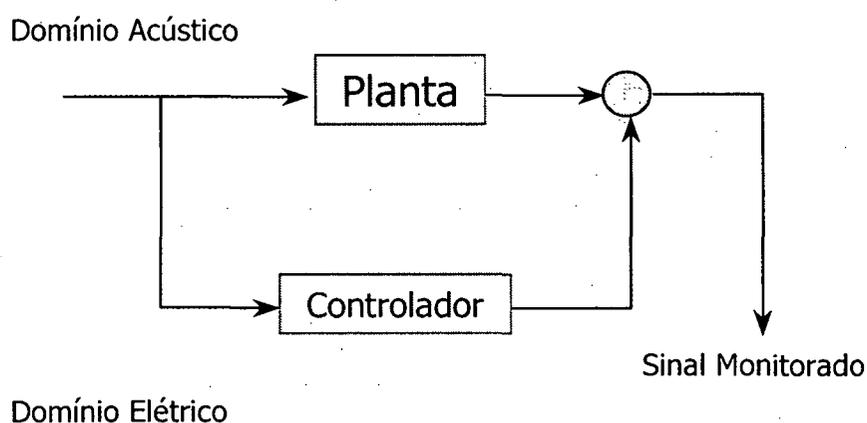


Fig. 1.2. Diagrama em blocos de uma aplicação de controle ativo.

Neste diagrama, a planta modela a resposta (ou o sistema) do trecho de duto que vai desde o detector até a fonte secundária. Assim, o diagrama da Fig. 1.2. pode ser visto como um problema de identificação de sistemas. No entanto, é interessante notar os diferentes domínios em que se encontram as transferências e os sinais envolvidos, assim como o ponto de soma que encontra-se no domínio acústico, como indicado na figura. Este fato representa a mais importante diferença em relação ao esquema tradicional de identificação de sistemas.

1.2. ADAPTABILIDADE

Alterações nas condições físicas do meio (tais como temperatura, umidade, etc.) e no fluxo de ar produzem, inevitavelmente, mudanças na resposta ao impulso da planta em função do tempo. Isto significa que o desempenho do sistema de controle tende a deteriorar-se se não for ajustado adequadamente em função destas mudanças. É nesse ponto, que é interessante introduzir a propriedade de *adaptabilidade* ao controlador. Um sistema de controle ativo adaptativo² de ruído acústico tem a habilidade de identificar mudanças nas características do duto pela adaptação da função de transferência do controlador. Estes tipos de sistemas, com uma certa capacidade de *aprender*, são conhecidos como *sistemas adaptativos*. Um sistema destes, pelo menos em teoria, pode convergir ao ponto ótimo de funcionamento partindo do completo desconhecimento do meio em que está envolvido. Isto faz com que o controle ativo utilize como ferramenta a teoria de *sistemas adaptativos*, permitindo que estes sistemas possam ser utilizados numa ampla faixa de aplicações.

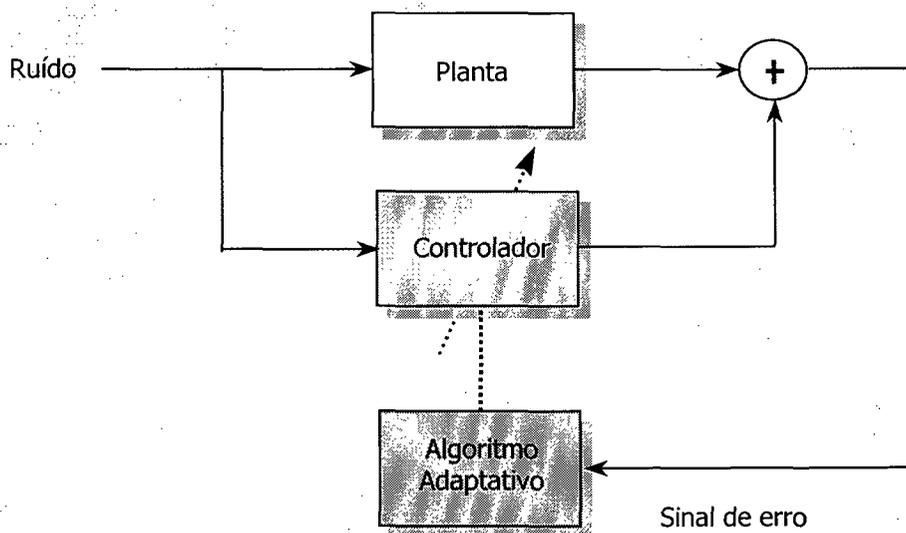


Fig. 1.3. Diagrama em blocos de um sistema adaptativo.

² Na literatura, existe certa confusão em relação à classificação controle ativo e adaptativo. Existem controladores ativos projetados segundo a teoria clássica de controle [3, 25] sem capacidade de adaptação. Quando o controlador modifica seus parâmetros através de um algoritmo adaptativo, utilizamos o termo controle ativo adaptativo.

Para representar um sistema com estas características devemos alterar o diagrama em blocos da Fig. 1.2. Em geral, utiliza-se o sinal monitorado (ou também chamado de sinal de erro) como medida de discrepância entre as atuais respostas do controlador e do sistema desconhecido. O sinal de erro é utilizado como entrada para um algoritmo que modifica de forma contínua os parâmetros do controlador (filtro adaptativo) a fim de atingir o ponto de funcionamento ótimo. O diagrama em blocos que considera esta situação é apresentado na Fig. 1.3.

1.3 ALGORITMO ADAPTATIVO.

Existe uma grande variedade de algoritmos adaptativos descritos na literatura. Na seleção do algoritmo devem ser considerados aspectos como velocidade de convergência, robustez, capacidade para acompanhar as variações do sistema desconhecido e complexidade computacional. Em geral, a escolha do algoritmo baseia-se em um compromisso entre desempenho e facilidade de implementação. Para o caso de controle de ruído acústico ou de vibrações, esta escolha é crucial já que são necessários filtros com grande número de coeficientes. O algoritmo adaptativo mais utilizado pela sua simplicidade é o algoritmo LMS (Least Mean Squares), [5-6] implementado através de filtros FIR. Para sermos mais rigorosos em relação ao diagrama em blocos apresentado na Fig. 1.3, o mesmo deve ser complementado com as transferências decorrentes da passagem dos sinais entre os diferentes domínios (elétrico-acústico). Isto será brevemente discutido na próxima seção.

1.4 ALGORITMO LMS FILTRADO

A utilização do algoritmo LMS supõe que o sinal de erro seja diretamente disponível, sendo originado pela diferença entre o sinal da fonte primária e a saída do filtro adaptativo. O processo de adaptação requer que o gradiente do erro médio quadrático seja aproximado pelo produto do valor instantâneo do sinal de erro e o sinal de entrada do filtro (sinal de referência). No entanto, as características particulares da configuração de um sistema de controle ativo adaptativo de ruído fazem com que o diagrama da Fig. 1.3 deva ser modificado a fim de considerar as seguintes diferenças em relação ao caso clássico: o ponto de soma representa a região do espaço (domínio acústico) onde ocorre

a superposição dos sinais acústicos oriundos das fontes primária (ruído não desejado) e secundária (antiruído). Outra diferença diz respeito ao sinal de erro, que corresponde ao sinal elétrico obtido do microfone de erro, a partir do campo sonoro residual.

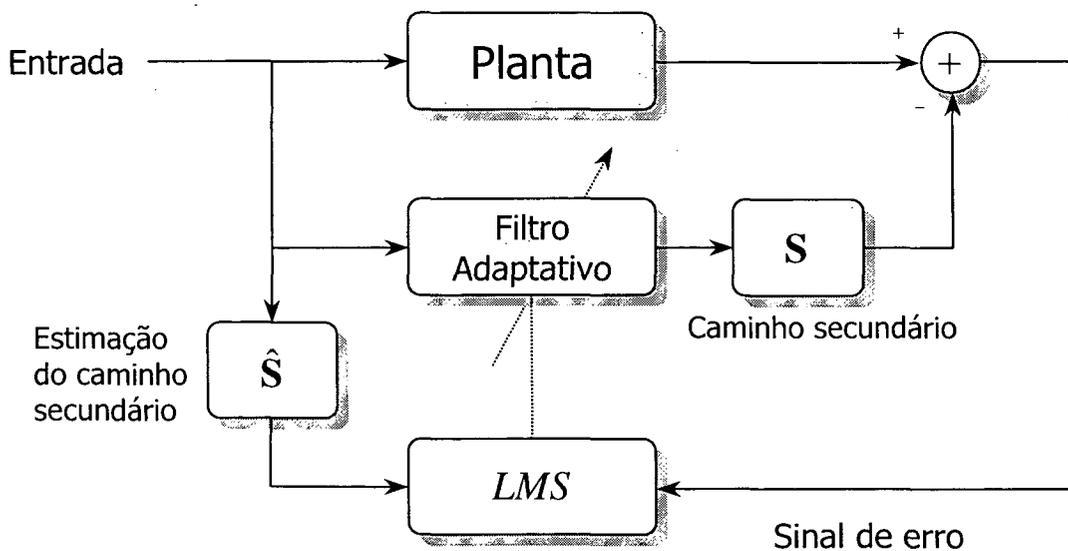


Fig. 1.4. Diagrama em blocos do algoritmo LMS filtrado.

Nos sistemas de controle ativo adaptativo de ruído acústico, o sinal de erro de cancelamento acústico não pode ser acessado diretamente. Ele é acessado como um sinal elétrico apenas à saída de um transdutor (normalmente um microfone). Da mesma forma, o sinal de saída do filtro adaptativo não pode ser empregado diretamente na operação de cancelamento acústico. Este sinal elétrico é primeiramente processado por um sistema que o converte em um sinal acústico. Estas conversões de natureza dos sinais dão origem ao que chamamos de *caminho secundário do sinal*, o qual é composto pelas respostas do seguintes elementos: *caminho secundário*, que é composto pelas respostas dos seguintes elementos: o conversor digital-analógico (D/A) em série com o filtro adaptativo, filtro de reconstrução, amplificador de potência, alto-falante, caminho acústico que vai do alto-falante até o microfone de erro, pré-amplificador, filtro anti-recobrimento e, finalmente, conversor analógico-digital (A/D) que fornece o sinal de erro para o algoritmo. O caminho secundário, denotado por S , afeta o comportamento do algoritmo podendo causar dificuldades no processo de convergência [14-15]. Portanto, seu efeito deve ser compensado, a fim de evitar a instabilidade do algoritmo adaptativo. Existem várias

configurações possíveis que podem ser utilizadas para compensar o efeito de S . Uma delas propõe a inclusão de um filtro inverso S^{-1} em série com S a fim de remover seu efeito. Porém, a solução mais utilizada consiste na utilização de um filtro \hat{S} com resposta ao impulso igual à de S para filtrar o sinal de referência. Na prática, \hat{S} é uma estimativa da resposta ao impulso S . Finalmente, uma descrição mais precisa do diagrama em blocos da Fig. 1.3, representando o esquema do algoritmo conhecido por LMS Filtrado ou FXLMS³, é mostrada na Fig. 1.4.

1.5 ESTIMAÇÃO DE S

O algoritmo LMS filtrado requer, para seu funcionamento, o conhecimento da resposta do filtro S , chamado de caminho secundário. Assumindo que as características de S são invariantes no tempo, é possível modelar sua resposta através de um procedimento *off-line* numa primeira etapa de treinamento. Ao final desta etapa, o modelo \hat{S} obtido é incluído no sistema para operação normal. Na etapa de treinamento, um ruído branco é utilizado como sinal de entrada, uma vez que possui uma densidade espectral constante para todas as frequências. A montagem experimental para a determinação da estimativa de S é mostrada na Fig. 1.5 [4].

A modelagem *off-line* é uma ferramenta muito útil para a determinação da resposta \hat{S} . Entretanto, possui o problema de não acompanhar possíveis mudanças do caminho secundário. Para os casos em que a resposta do caminho secundário possa ser variante no tempo, é desejável realizar tal modelagem de forma contínua (estimação em tempo real (*on-line*)). Outra opção a ser empregada, ao invés da estimação contínua, é a estimação em determinados intervalos de tempo. Supondo que \hat{S} varie lentamente, as adaptações do controlador e do estimador do caminho secundário podem ser consideradas separadamente. Como exemplo, a Fig. 1.6 mostra o diagrama em blocos para o caso de estimação *on-line* proposto por Widrow e Stearns [5].

³ Sigla em inglês, derivada de filtered-X LMS.

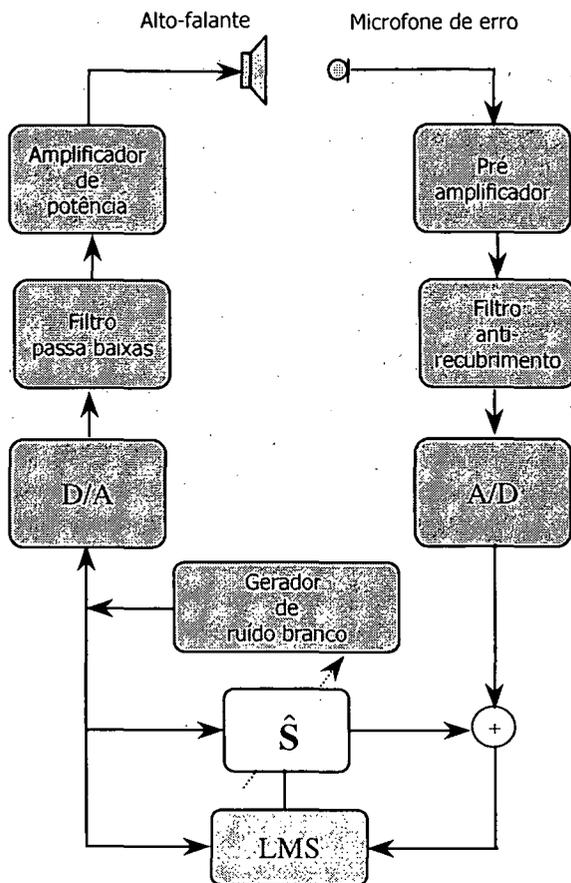


Fig. 1.5. Montagem experimental para a estimação de S .

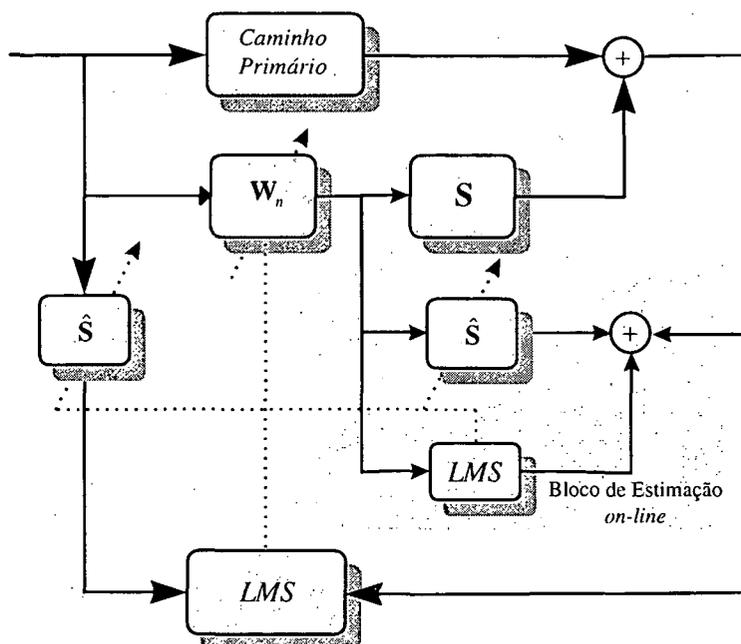


Fig. 1.6. Diagrama em blocos de um sistema com estimação *on-line* de S .

Como vimos, a utilização do algoritmo FXLMS surge de forma natural devido às características particulares do sistema de controle ativo adaptativo. Portanto, existe a necessidade do detalhado conhecimento do comportamento do algoritmo para as diversas condições de funcionamento.

1.6 ESTADO ATUAL DA ARTE

Os primeiros a proporem a utilização do algoritmo LMS filtrado foram Widrow e Stearns [5]. A utilização deste algoritmo foi dentro do contexto da teoria de controle adaptativo, não sendo fornecida qualquer análise do seu comportamento. É citado, também, nessa referência, um exemplo com a utilização do algoritmo LMS filtrado para o controle de ruído acústico. Burgess [15], paralelamente a Widrow e Stearns, desenvolveu o algoritmo LMS filtrado aplicado ao problema de controle ativo de ruído acústico num duto. Em [15], a única análise realizada por foi a obtenção das expressões de atualização dos coeficientes do filtro adaptativo. Uma publicação recente [16], apresentou uma análise estatística do comportamento deste algoritmo sendo calculados os momentos de primeira e de segunda ordens. Entretanto, a análise em [16] aborda o problema do ponto de vista clássico, utilizando a teoria da independência [17-18]. Como demonstram os resultados obtidos a partir da análise proposta nesta tese, as equações apresentadas em [16] fornecem resultados que não descrevem o real comportamento do algoritmo.

Análises do comportamento estatístico do algoritmo LMS freqüentemente empregam um conjunto de aproximações denominado "Teoria da Independência" [5,18]. Deste conjunto, a aproximação talvez mais importante é a de que diferentes vetores de entrada são independentes. Esta pode ser uma situação real nos sistemas adaptativos denominados *combinadores lineares*, nos quais cada coeficiente do filtro adaptativo atua sobre uma seqüência diferente de entrada. No caso de sistemas usando filtros adaptativos transversais com uma única entrada, vetores de entrada consecutivos de dimensão N têm $(N-1)$ componentes em comum. Assim, tais vetores são altamente correlacionados.

Apesar de desprezar a correlação entre os vetores de entrada, a aplicação da Teoria da Independência leva a resultados muito bons quando aplicada à análise do comportamento estatístico do algoritmo LMS usando passos de adaptação reduzidos. Alguns trabalhos mais recentes analisam o comportamento do algoritmo LMS . A

modelação considerando esta correlação acarreta um tratamento matemático bastante complexo [19,21]. Em [19] é descrita uma análise para o LMS clássico que considera a correlação dos dados de entrada. Nessa análise, verifica-se que é possível obter predições mais exatas de seu comportamento estatístico e, também, que os limites de estabilidade obtidos para a taxa de aprendizagem são mais severos do que os resultantes da consideração de independência. O custo desta descrição mais exata é um importante aumento da complexidade matemática do modelo. Como exemplo disto, o trabalho reporta que são necessárias 28000 equações para descrever o comportamento de um filtro de 6 coeficientes. Contudo, o algoritmo LMS tem a possibilidade de ser estudado à luz das duas abordagens, fornecendo muitos bons resultados em ambos os casos. Porém, no caso do LMS filtrado, verifica-se que um bom modelo depende da consideração da correlação entre vetores de entrada, como será mostrado mais adiante.

Uma outra análise do LMS filtrado é apresentada em [4, pag. 65]. Nesse trabalho, é feita a suposição de que os coeficientes do filtro adaptativo variam lentamente e de que os sistemas que compõem o caminho adaptativo são lineares e invariantes no tempo, possibilitando assim a troca da ordem entre os mesmos. Neste caso, o diagrama em blocos é dado por:

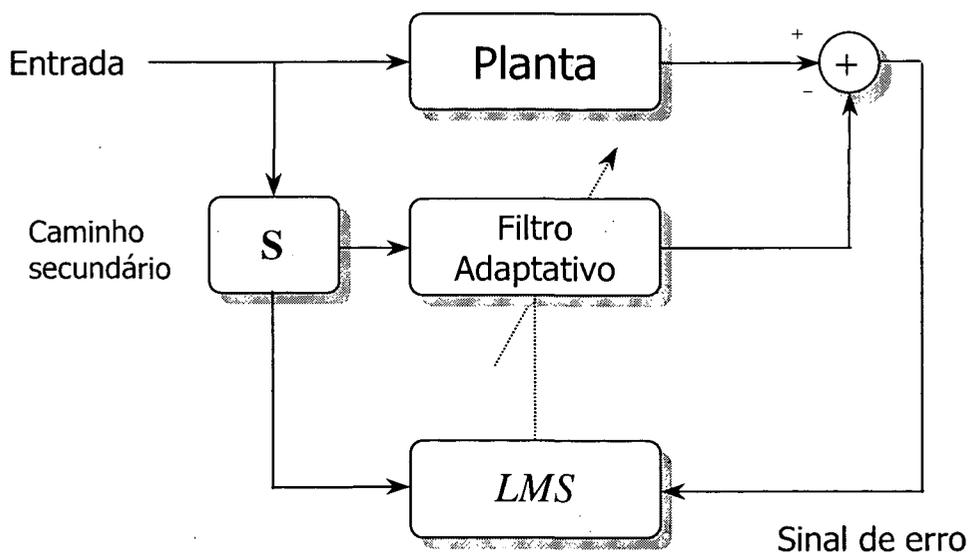


Fig. 1.7. Diagrama em blocos modificado do algoritmo LMS filtrado.

Esta versão fornece bons resultados para o caso em que o caminho secundário e sua estimação são iguais e com uma taxa de aprendizagem (μ) muito pequena. Esse modelo torna-se limitado para o estudo da influência dos erros na estimação do caminho secundário ou para o estudo do comportamento do algoritmo para taxas de aprendizagem maiores. Em [22] é utilizado o modelo da Fig. 1.4, mas, com a inversão das posições do filtro adaptativo e do caminho secundário. Nesse trabalho, o objetivo é determinar a influência dos erros na estimação de S sobre o comportamento do algoritmo. Também é determinada uma expressão para descrever o comportamento dinâmico dos pesos baseada no intercâmbio das posições, a qual não modela o comportamento real. Outra deficiência do modelo da Fig. 1.7 pode ser apontada quando for necessário considerar uma não-linearidade (devido principalmente ao alto-falante). Nesse caso, não é possível efetuar a troca entre caminho secundário e filtro adaptativo. Portanto, para esse último caso a análise resulta impraticável. A abordagem proposta nesta tese não impõe grandes limitações ao valor da taxa de aprendizagem e não utiliza a simplificação decorrente da troca de posição dos blocos (Fig. 1.7). Isto permite uma análise mais realista do desempenho do algoritmo [30,33-34]. Um outro resultado importante é obtido para o caso especial em que o caminho secundário representa um atraso de um número determinado de amostras. Tal situação é conhecida na literatura por algoritmo DLMS (Delayed-LMS) [9, 26, 29, 32]. Neste caso, os trabalhos que conhecemos da literatura também utilizam a Teoria da Independência. Assim, não é possível estudar o efeito de erros na estimação do atraso nem determinar o comportamento dos coeficientes do filtro nesta condição. Com o novo modelo proposto, é possível obter dois resultados importantes: demonstrar porque o algoritmo DLMS não converge com ruído branco na entrada e em condições de erro na estimação do atraso e, em segundo lugar, sob similar condição, o mesmo converge se a entrada for correlacionada.

A proposta deste trabalho é o desenvolvimento de modelos para o comportamento dos momentos de primeira e de segunda ordens do algoritmo LMS filtrado. Os novos modelos permitem derivar um conjunto de importantes parâmetros de projeto (velocidade de convergência, erro mínimo, desajuste, etc.), oferecendo ao projetista a possibilidade de estudar o desempenho do algoritmo nas mais diversas condições de funcionamento.

1.7 ORGANIZAÇÃO DA TESE

No Capítulo 2, abordamos o cálculo do momento de primeira ordem do algoritmo LMS filtrado e apresentamos resultados de simulações mostrando a validade da análise proposta.

O momento de segunda ordem é tratado no Capítulo 3, no qual é apresentado primeiramente o caso mais simples, que é o algoritmo DLMS, seguido da análise do caso geral. Ao final de cada caso analisado são apresentados resultados de simulações.

O Capítulo 4 é dedicado a uma análise exploratória, no domínio da frequência, dos efeitos das operações de filtragem sobre o desempenho do algoritmo FXLMS.

Finalmente, no Capítulo 5 são apresentadas as conclusões deste trabalho, assim como apontadas algumas perspectivas para próximos trabalhos.

Momento de Primeira Ordem

2

INTRODUÇÃO

Neste capítulo é apresentada a análise do momento de primeira ordem (comportamento médio) dos coeficientes do filtro adaptativo usando o algoritmo FXLMS. Nessa análise, não é utilizada a chamada teoria da independência [18], usualmente empregada no estudo do algoritmo LMS clássico e em estudos recentes do algoritmo LMS filtrado [2-4,16,22]. Com esta nova abordagem é obtida, também, a solução ótima e o valor em regime permanente dos coeficientes do filtro adaptativo. Um parâmetro importante para o estudo do comportamento do algoritmo, o vetor de erro nos coeficientes em relação ao valor em regime permanente, é também examinado. Finalmente, resultados de simulações são mostrados a fim de comprovar a qualidade do modelo proposto. Estes resultados são também comparados com os resultados obtidos utilizando a teoria da independência. A comparação mostra claramente que a teoria da independência não fornece resultados aceitáveis.

2.1 CONSIDERAÇÕES INICIAIS – DIAGRAMA EM BLOCOS

A Fig. 2.1 mostra o diagrama esquemático para uma aplicação de controle de ruído acústico. A Fig. 2.2, apresenta o diagrama em blocos para o algoritmo FXLMS [2-5], o qual pode ser usado para modelar o sistema da Fig. 2.1. Na Fig. 2.2, W^o é o sistema desconhecido, $W(n)$ é o filtro adaptativo e S , \hat{S} são filtros lineares invariantes no tempo. Com relação à Fig. 2.1, W^o modela a resposta desconhecida, localizada entre os pontos A e B. S modela a resposta entre C e B e \hat{S} modela o correspondente filtro de compensação. A resposta entre A e E é desconsiderada em relação a outros efeitos de filtragem. Também é desconsiderada a realimentação acústica existente entre D e E [4]. Na literatura, a resposta de S é denominada resposta ao impulso do caminho secundário. O filtro \hat{S} é projetado para ser uma estimativa da resposta do caminho secundário

[2, capítulos 3 e 7]. No entanto, na prática, uma estimativa perfeita é muito difícil de ser obtida. Portanto, o caso mais geral corresponde ao de estimação imperfeita. Todas as respostas ao impulso são modeladas através de filtros lineares com resposta ao impulso finita, filtros FIR. Então, os sinais e respostas ao impulso envolvidos na presente análise, são:

- Sinal de referência $x(n)$: estacionário e com variância σ_x^2 ;
- Ruído de medição $z(n)$: ruído branco estacionário e não correlacionado com nenhum outro sinal do sistema, possuindo variância σ_z^2 ;
- \mathbf{W}^o : vetor representando a resposta ao impulso do sistema a ser identificado. Em controle ativo de ruído acústico, representa a resposta do caminho acústico;
- Os sistemas \mathbf{W}^o , \mathbf{S} e $\hat{\mathbf{S}}$ são modelados através de filtros FIR;
- $\mathbf{W}(n)$: vetor das amostras da resposta ao impulso do filtro adaptativo;
- $e(n)$: sinal de erro;
- $y(n)$: saída do filtro adaptativo;
- $y_f(n)$: saída do bloco em série com $\mathbf{W}(n)$;
- $x_f(n)$: saída do bloco de estimação do caminho secundário;
- $d(n)$: saída desejada;
- $\mathbf{X}_f(n)$: vetor $\mathbf{X}(n)$ filtrado (saída de $\hat{\mathbf{S}}$).



Fig. 2.1. Diagrama esquemático para o caso de controle de ruído acústico.

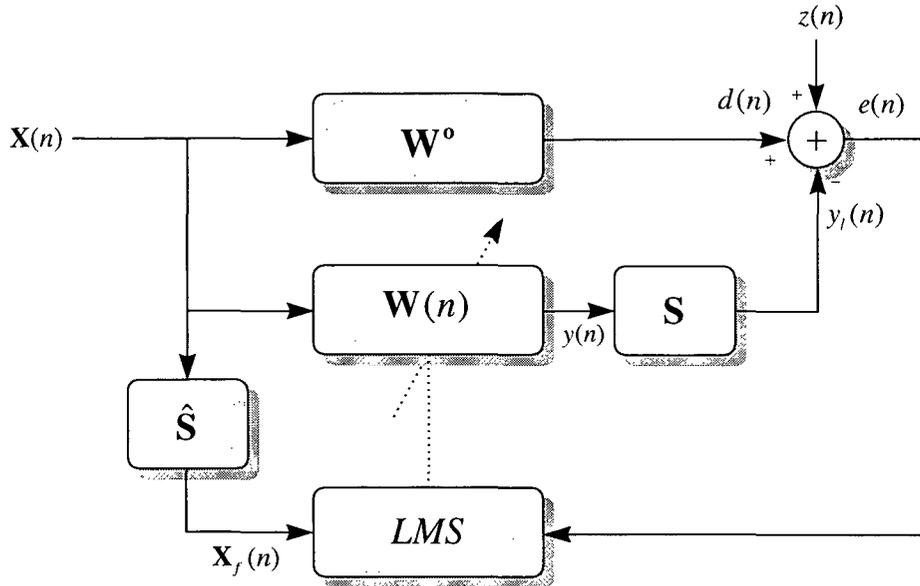


Fig. 2.2. Diagrama em blocos do algoritmo LMS filtrado.

Os vetores utilizados são definidos pela seguintes estruturas:

$$\text{vetor } \mathbf{X}(n): \quad \mathbf{X}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$$

$$\text{vetor } \mathbf{W}(n): \quad \mathbf{W}(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T$$

$$\text{vetor } \mathbf{S}: \quad \mathbf{S} = [s_0, s_1, \dots, s_{M-1}]^T$$

$$\text{vetor } \hat{\mathbf{S}}: \quad \hat{\mathbf{S}} = [\hat{s}_0, \hat{s}_1, \dots, \hat{s}_{\hat{M}-1}]^T$$

$$\text{vetor } \mathbf{W}^o: \quad \mathbf{W}^o = [w_0^o, w_1^o, \dots, w_{N-1}^o]^T$$

$$\text{vetor } \mathbf{X}_f(n): \quad \mathbf{X}_f(n) = [x_f(n), x_f(n-1), \dots, x_f(n-N+1)]^T$$

Note que é assumido igual dimensão para os vetores \mathbf{W}^o e $\mathbf{W}(n)$, com o objetivo de facilitar a notação. No caso que um desses vetores possua menor número de componentes, o mesmo poderá ser preenchido com zeros a fim de igualar as dimensões. Os vetores \mathbf{S} e $\hat{\mathbf{S}}$ possuem dimensões M e \hat{M} , respectivamente.

2.2. EQUACIONAMENTO DO SISTEMA

Em função das definições anteriormente apresentadas, temos que:

$$x_f(n) = \sum_{i=0}^{\hat{M}-1} \hat{s}_i x(n-i) \quad (2.1)$$

e

$$\mathbf{X}_f(n) = \sum_{i=0}^{\hat{M}-1} \hat{s}_i \mathbf{X}(n-i) \quad (2.2)$$

2.2.1 Sinal de erro

Considerando o sistema da Fig. 2.2 temos:

$$e(n) = d(n) - y_1(n) + z(n), \quad (2.3)$$

onde:

$$d(n) = \mathbf{X}^T(n) \mathbf{W}^o, \quad (2.4)$$

$$y(n) = \mathbf{X}^T(n) \mathbf{W}(n) = \mathbf{W}^T(n) \mathbf{X}(n), \quad (2.5)$$

$$y_1(n) = \sum_{i=0}^{M-1} s_i y(n-i) = \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \mathbf{W}(n-i), \quad (2.6)$$

substituindo (2.6) em (2.3), temos:

$$e(n) = d(n) - \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \mathbf{W}(n-i) + z(n). \quad (2.7)$$

A Eq. (2.7), mostra que o erro é uma função do sinal desejado, $d(n)$, e de M vetores consecutivos do sinal de entrada e do coeficientes do filtro adaptativo.

2.2.2. Atualização dos coeficientes do filtro adaptativo

A equação de atualização dos coeficientes do filtro adaptativo usando o algoritmo LMS filtrado [4, 30], é dada por:

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) + \mu e(n) \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{X}(n-j) \right) \\ &= \mathbf{W}(n) + \mu e(n) \mathbf{X}_j(n), \end{aligned} \quad (2.8)$$

onde μ representa a taxa de aprendizagem. Se substituirmos (2.7) em (2.8) temos a expressão recursiva final para atualização dos coeficientes do filtro adaptativo:

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) + \mu \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{X}(n-j) \mathbf{X}^T(n) \mathbf{W}^o \right. \\ &\quad \left. - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{X}(n-j) \mathbf{X}^T(n-i) \mathbf{W}(n-i) + \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{X}(n-j) z(n) \right). \end{aligned} \quad (2.9)$$

2.3. SOLUÇÃO ÓTIMA

Para a estrutura convencional não filtrada, a solução de Wiener é dada por [5,6]:

$$\mathbf{W}_{wiener} = \mathbf{R}_0^{-1} \mathbf{P}. \quad (2.10)$$

Onde $\mathbf{R}_0 = E[\mathbf{X}(n)\mathbf{X}^T(n)]$ e $\mathbf{P} = E[d(n)\mathbf{X}(n)]$. Na presente análise, a presença do filtro \mathbf{S} no caminho secundário modifica a expressão da solução de ótima para o problema de filtragem linear. Portanto, é interessante determinar a expressão do vetor ótimo para este caso. Para a determinação da solução ótima para o caso filtrado, faremos uso do diagrama em blocos da Fig. 2.3.

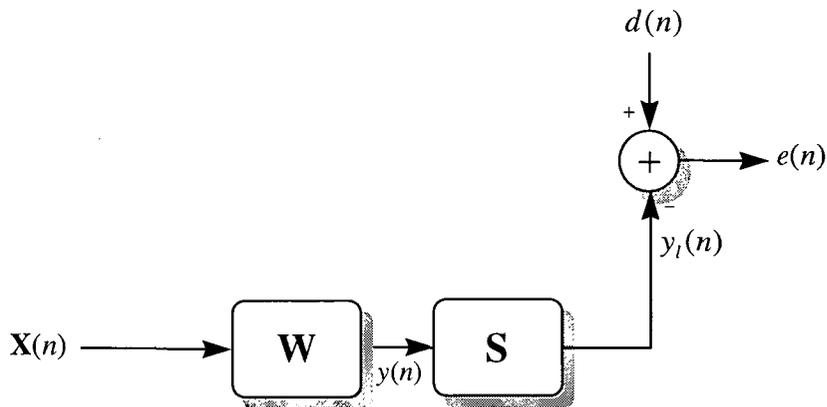


Fig. 2.3. Diagrama em blocos para a determinação da solução de Wiener.

A expressão do erro instantâneo obtido partir da Fig. 2.3 tem a mesma forma que a Eq.(2.7). No entanto, note-se que agora o vetor \mathbf{W} de coeficientes é feito fixo. Com esta condição, a expressão para o erro instantâneo ao quadrado é:

$$e^2(n) = d^2(n) - 2 \left[d(n) \sum_{j=0}^{M-1} s_j \mathbf{X}^T(n-j) \right] \mathbf{W} + \mathbf{W}^T \left[\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{X}(n-j) \mathbf{X}^T(n-i) \right] \mathbf{W} . \quad (2.11)$$

Tomando o valor esperado em (2.11), temos uma expressão para o erro médio quadrático:

$$E[e^2(n)] = E[d^2(n)] - 2 \sum_{j=0}^{M-1} s_j E[d(n) \mathbf{X}^T(n-j)] \mathbf{W} + \mathbf{W}^T \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j E[\mathbf{X}(n-j) \mathbf{X}^T(n-i)] \mathbf{W} . \quad (2.12)$$

Chamando $\mathbf{R}_{i-j} = E[\mathbf{X}(n-j) \mathbf{X}^T(n-i)]$ e $\mathbf{P}_j = E[d(n) \mathbf{X}^T(n-j)]$, a Eq. (2.12) pode ser expressa como:

$$E[e^2(n)] = E[d^2(n)] - 2 \left[\sum_{j=0}^{M-1} s_j \mathbf{P}_j^T \right] \mathbf{W} + \mathbf{W}^T \left[\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{R}_{i-j} \right] \mathbf{W} . \quad (2.13)$$

Para obtermos uma expressão mais compacta de (2.13), definimos as seguintes quantidades:

$$\tilde{\mathbf{P}}_s = \sum_{j=0}^{M-1} s_j \mathbf{P}_j , \quad (2.14a)$$

e

$$\tilde{\mathbf{R}}_{ss} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{R}_{i-j} . \quad (2.14b)$$

Note que $\tilde{\mathbf{R}}_{ss}$ corresponde à matriz de autocorrelação do sinal $\mathbf{X}(n)$ filtrado por \mathbf{S} . Da mesma forma, $\tilde{\mathbf{P}}_s$ corresponde ao vetor correlação cruzada entre $d(n)$ e o sinal $\mathbf{X}(n)$ filtrado por \mathbf{S} . Substituindo (2.14) em (2.13) e usando a notação $J = E[e^2(n)]$, temos:

$$J = E[d^2(n)] - 2\tilde{\mathbf{P}}_s^T \mathbf{W} + \mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W}. \quad (2.15)$$

É interessante observar que a função custo, Eq. (2.15), para a estrutura filtrada possui a mesma forma que a do caso clássico. De forma análoga ao caso clássico, para obtermos o vetor de coeficientes ótimo \mathbf{W}_{ot} que minimiza o erro médio quadrático, derivamos (2.15) em relação à variável \mathbf{W} e igualamos a derivada a zero. Com este procedimento determinamos as equações normais, expressas a seguir:

$$\tilde{\mathbf{R}}_{ss} \mathbf{W}_{ot} = \tilde{\mathbf{P}}_s, \quad (2.16)$$

onde \mathbf{W}_{ot} é o vetor ótimo de coeficientes, o qual minimiza J . Pela definição da matriz $\tilde{\mathbf{R}}_{ss}$ é fácil observar que a mesma é uma matriz simétrica, já que $\mathbf{R}_{i-j} = \mathbf{R}_{j-i}^T$. Assumindo que a inversa de $\tilde{\mathbf{R}}_{ss}$ existe, os coeficientes ótimos são dados por:

$$\mathbf{W}_{ot} = \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{P}}_s. \quad (2.17)$$

Note que, novamente, a expressão (2.17) é do mesmo tipo que a obtida para a estrutura convencional. No entanto, a diferença principal encontra-se na definição das matrizes $\tilde{\mathbf{R}}_{ss}$ e $\tilde{\mathbf{P}}_s$, onde fica em evidência o efeito do filtro em série com o filtro adaptativo. A expressão (2.17) foi apresentada em [22], onde a mesma foi derivada de forma heurística.

É interessante expressar a relação entre a solução de Wiener, Eq. (2.17), e a resposta ao impulso da planta física, \mathbf{W}^o , para o caso da Fig. 2.2. Se considerarmos a definição de $\tilde{\mathbf{P}}_s$, temos que:

$$\tilde{\mathbf{P}}_s = \sum_{i=0}^{M-1} s_i \mathbf{P}_i = \sum_{i=0}^{M-1} s_i E[d(n)\mathbf{X}(n-i)]. \quad (2.18)$$

A partir da Fig. 2.2 temos que $d(n) = \mathbf{X}^T(n)\mathbf{W}^o + z(n)$, substituindo $d(n)$ em (2.18), temos:

$$\begin{aligned} \tilde{\mathbf{P}}_s &= \sum_{i=0}^{M-1} s_i E[\mathbf{X}(n-i)\mathbf{X}^T(n)\mathbf{W}^o + z(n)\mathbf{X}(n-i)] = \sum_{i=0}^{M-1} s_i E[\mathbf{X}(n-i)\mathbf{X}^T(n)]\mathbf{W}^o \\ &= \left[\sum_{i=0}^{M-1} s_i \mathbf{R}_{-i} \right] \mathbf{W}^o, \end{aligned} \quad (2.19a)$$

ou, definindo $\tilde{\mathbf{R}}_s = \sum_{i=0}^{M-1} s_i \mathbf{R}_{-i}$ temos:

$$\tilde{\mathbf{P}}_s = \tilde{\mathbf{R}}_s \mathbf{W}^o. \quad (2.19b)$$

De (2.17) e (2.19), temos a relação que estamos buscando:

$$\tilde{\mathbf{R}}_{ss} \mathbf{W}_{ot} = \tilde{\mathbf{R}}_s \mathbf{W}^o. \quad (2.20)$$

Assumindo que a inversa de $\tilde{\mathbf{R}}_{ss}$ existe, então

$$\mathbf{W}_{ot} = \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{R}}_s \mathbf{W}^o. \quad (2.21)$$

2.3.1. Princípio da ortogonalidade

Da mesma forma que para a estrutura convencional, também pode-se obter a expressão do princípio de ortogonalidade para a estrutura com o filtro S. Para isto, vamos partir da Eq. (2.16). Expandindo os termos, temos:

$$\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]\mathbf{W}_{or} - \sum_{i=0}^{M-1} s_i E[d(n)\mathbf{X}(n-i)] = 0. \quad (2.22)$$

Esta expressão pode ser reescrita como:

$$E\left[d(n)\sum_{i=0}^{M-1} s_i \mathbf{X}(n-i)\right] - E\left[\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{X}(n-i)\mathbf{X}^T(n-j)\right]\mathbf{W}_{or} = 0, \quad (2.23)$$

ou ainda como:

$$\begin{aligned} & E\left[\left(d(n) - \sum_{j=0}^{M-1} s_j \mathbf{X}^T(n-j)\mathbf{W}_{or}\right)\left(\sum_{i=0}^{M-1} s_i \mathbf{X}(n-i)\right)\right] \\ & = E\left[e_0\left(\sum_{i=0}^{M-1} s_i \mathbf{X}(n-i)\right)\right] = 0, \end{aligned} \quad (2.24)$$

onde:

$$e_0 = d(n) - \sum_{j=0}^{M-1} s_j \mathbf{X}^T(n-j)\mathbf{W}_{or}, \quad (2.25)$$

é o sinal de erro quando o filtro opera na condição ótima, ou seja, para $\mathbf{W} = \mathbf{W}_{or}$. Note que (2.24) corresponde à ortogonalidade entre o sinal de erro e um sinal correspondente à filtragem de $x(n)$ pelo filtro S. Este sinal não existe fisicamente na estrutura da Fig. 2.3, já que $x(n)$ é aplicado à entrada do filtro W. O princípio da ortogonalidade provê a condição necessária e suficiente para que a função custo, o erro médio quadrático, atinja o mínimo valor. Tal condição corresponde ao erro e_0 , ortogonal ao vetor $\left(\sum_{i=0}^{M-1} s_i \mathbf{X}(n-i)\right)$, que é a versão filtrada do vetor de entrada. Este resultado também pode

ser obtido usando-se as propriedades de linearidade e invariância no *tempo* dos filtros \mathbf{W} e \mathbf{S} na Fig. 2.3 [22]; podendo-se inverter a ordem entre \mathbf{W} e \mathbf{S} .

2.3.2. Superfície de erro – Mínimo erro médio quadrático

A função custo J , definida pela equação (2.13) é, da mesma forma que na estrutura clássica, uma função quadrática do vetor de coeficientes do filtro, \mathbf{W} . Utilizando a expressão (2.17) em (2.15) é possível determinar o valor mínimo para J , erro médio quadrático mínimo, sendo sua expressão dada por:

$$J_{\min} = \min \{ E[e^2(n)] \} = E[z^2(n)] + E[d^2(n)] - \tilde{\mathbf{P}}_s^T \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{P}}_s. \quad (2.26)$$

Note-se que em (2.26) é adicionado um termo representando um ruído de medição, dado por $z(n)$. Usando (2.17), obtém-se:

$$J_{\min} = E[z^2(n)] + E[d^2(n)] - \tilde{\mathbf{P}}_s^T \mathbf{W}_{ot}. \quad (2.27)$$

Diminuindo (2.27) de (2.15) e usando (2.16), chega-se à seguinte expressão:

$$J = J_{\min} + (\mathbf{W} - \mathbf{W}_{ot})^T \tilde{\mathbf{R}}_{ss} (\mathbf{W} - \mathbf{W}_{ot}). \quad (2.28)$$

Nas Figs. 2.4 e 2.5, são mostradas as superfícies de erro para os casos onde a entrada é ruído branco e ruído colorido, respectivamente. A superfície cinza corresponde à superfície para o caso filtrado e a quadriculada para o caso clássico. A partir dessas figuras pode-se observar o efeito da filtragem, correlacionando o sinal de entrada e, portanto, produzindo uma superfície com gradiente mais suave do que no caso não filtrado. Na medida em que o sinal de entrada for mais correlacionado (Fig. 2.5), a tendência é que a superfície de desempenho resulte mais achatada (as curvas de nível são elipses). Este fato está diretamente relacionado com a velocidade de convergência mais lenta, tanto para o caso clássico quanto para o caso filtrado para o qual o efeito é ainda mais pronunciado. Na Fig. 2.4.b, é mostrado um detalhe da região correspondente ao ponto mínimo da superfície de

erro. O gráfico mostra que a presença do filtro na malha de adaptação pode aumentar o erro mínimo da superfície de desempenho.

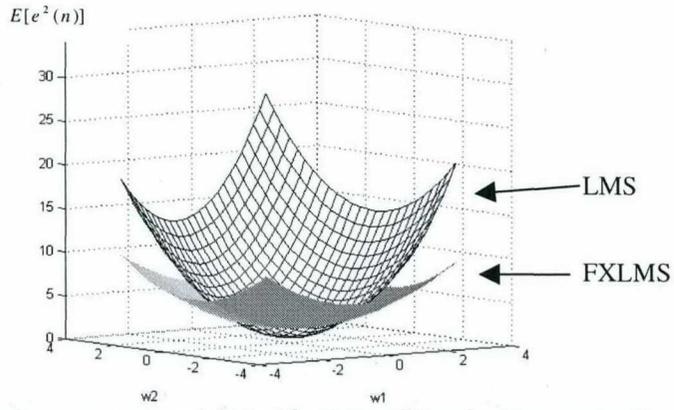


Fig. 2.4.a. Superfície de erro para o LMS clássico e filtrado com entrada sendo ruído branco.

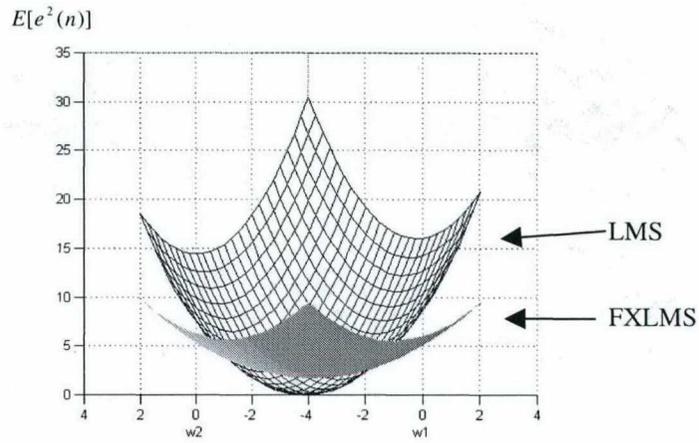


Fig. 2.4.b. Vista frontal da Fig. 2.4.a.

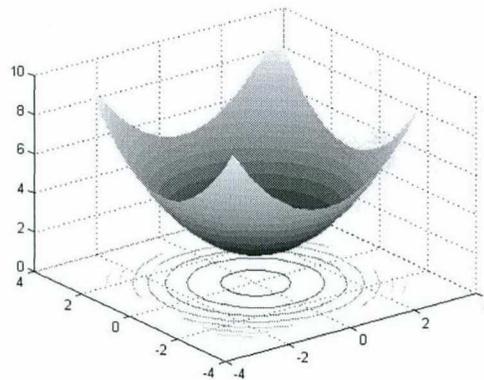


Fig. 2.4.c. Curvas de nível da Fig. 2.4.a para o caso filtrado.

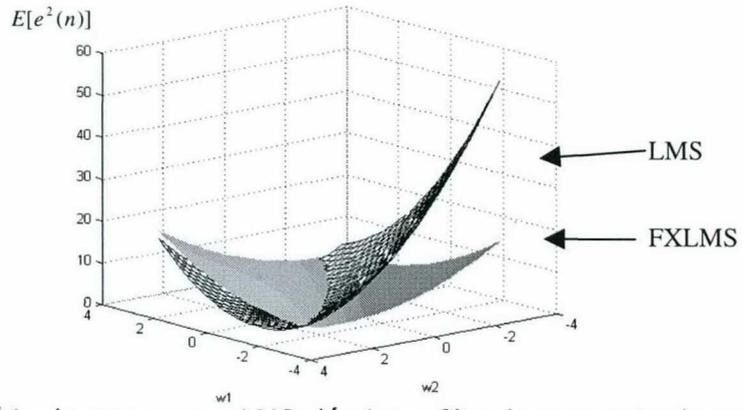


Fig. 2.5.a. Superfície de erro para o LMS clássico e filtrado com entrada sendo ruído colorido.

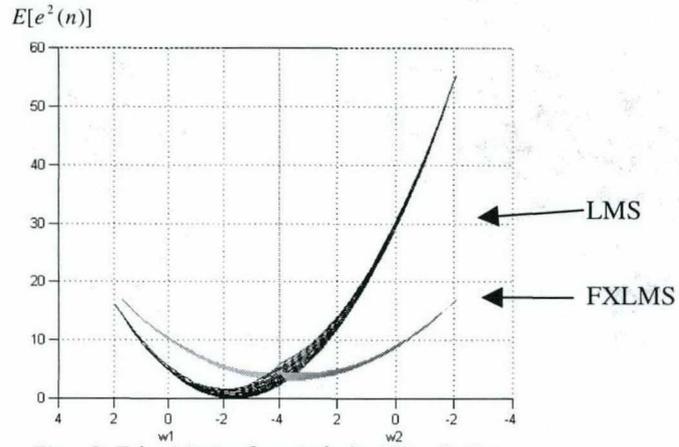


Fig. 2.5.b. Vista frontal da Fig. 2.5.a.

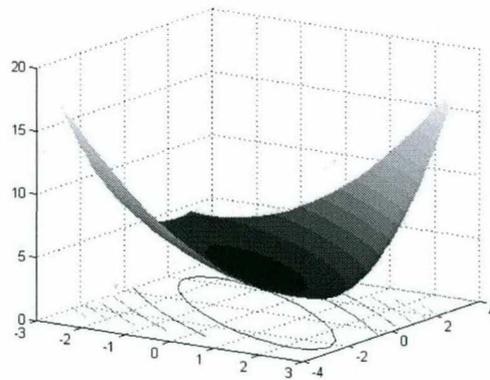


Fig. 2.5.c. Curvas de nível da Fig. 2.5.a para o caso filtrado.

2.3.3. Forma canônica da superfície de erro

A matriz $\tilde{\mathbf{R}}_{ss}$ definida em (2.14b) pode ser expressa em função de seus autovalores e autovetores [5-6]. Tal decomposição é apresentada a seguir:

$$\tilde{\mathbf{R}}_{ss} = \mathbf{Q}\tilde{\Lambda}\mathbf{Q}^T, \quad (2.29)$$

onde \mathbf{Q} é a matriz cujas colunas são os autovetores de $\tilde{\mathbf{R}}_{ss}$ e $\tilde{\Lambda} = \text{diag}(\lambda_0, \dots, \lambda_{N-1})$ é a matriz dos autovalores de $\tilde{\mathbf{R}}_{ss}$. Substituindo (2.29) em (2.28) tem-se:

$$J = J_{\min} + (\mathbf{W} - \mathbf{W}_{ot})^T \mathbf{Q}\tilde{\Lambda}\mathbf{Q}^T (\mathbf{W} - \mathbf{W}_{ot}), \quad (2.30)$$

chamando $\tilde{\mathbf{V}} = \mathbf{Q}^T (\mathbf{W} - \mathbf{W}_{ot})$, a expressão quadrática da superfície de erro pode ser reescrita na sua forma canônica como sendo:

$$J = J_{\min} + \tilde{\mathbf{V}}^T \tilde{\Lambda} \tilde{\mathbf{V}}, \quad (2.31)$$

ou em função dos autovalores λ_k ,

$$J = J_{\min} + \sum_{k=1}^{N-1} \lambda_k |\tilde{v}_k|^2, \quad (2.32)$$

onde \tilde{v}_k corresponde à k -ésima componente do vetor $\tilde{\mathbf{V}}$. Da mesma forma que no caso do LMS clássico, os componentes do vetor $\tilde{\mathbf{V}}$ estão contidos nos eixos principais da superfície de erro. Isto pode ser demonstrado calculando o gradiente da função custo,

$$\nabla J = 2\tilde{\Lambda}\tilde{\mathbf{V}} = 2[\lambda_0 \tilde{v}_0, \lambda_1 \tilde{v}_1, \dots, \lambda_{N-1} \tilde{v}_{N-1}]^T. \quad (2.33)$$

De (2.33) observa-se que se apenas um componente \tilde{v}_k for diferente de zero, o vetor gradiente encontra-se sobre esse eixo. Também pode-se perceber de (2.33) a

importância dos autovalores da matriz $\tilde{\mathbf{R}}_{ss}$; analisando o gradiente de J em relação a qualquer eixo principal \tilde{v}_k , pode se escrever:

$$\frac{\partial J}{\partial \tilde{v}_k} = 2\lambda_k \tilde{v}_k, \quad (2.34)$$

ou ainda, a derivada segunda

$$\frac{\partial^2 J}{\partial \tilde{v}_k^2} = 2\lambda_k, \quad k = 0, 1, \dots, N-1. \quad (2.35)$$

A última expressão mostra que a derivada segunda de J em relação a \tilde{v}_k é igual a duas vezes o autovalor correspondente nessa direção. Então, pode-se enunciar a seguinte propriedade:

Os autovalores da matriz $\tilde{\mathbf{R}}_{ss}$ correspondem à derivada segunda da superfície de erro J em relação aos eixos principais de J .

No caso da estrutura clássica, a mesma propriedade é válida para os autovalores da matriz de autocorrelação da entrada [5]. Como conclusão, podemos dizer que do ponto de vista da superfície de desempenho, as estruturas clássica e filtrada são bastante semelhantes em relação a suas propriedades. Para passarmos de uma para a outra, devemos substituir \mathbf{R} por $\tilde{\mathbf{R}}_{ss}$, o que corresponde a substituir o sinal $\mathbf{X}(n)$ por sua versão filtrada pelo filtro S .

2.4. CÁLCULO DO MOMENTO DE PRIMEIRA ORDEM

Nesta seção, é determinado um modelo para a evolução do valor esperado dos coeficientes do filtro adaptativo, $E[\mathbf{W}(n)]$. Em análises do comportamento estatístico do algoritmo LMS, utiliza-se freqüentemente a teoria da independência [18]. Esta teoria assume a independência de vetores de entrada em diferentes instantes de tempo, como por exemplo $\mathbf{X}(n)$ e $\mathbf{X}(n-1)$ [5-7]. Esta hipótese, embora sabidamente incorreta, muitas vezes leva a uma simplificação significativa do equacionamento matemático, sem prejuízo

da precisão na determinação do comportamento do algoritmo. Este é o caso do algoritmo LMS com aprendizagem suficientemente lenta [5-7]. No caso do algoritmo LMS filtrado, entretanto, a presença dos sistemas S e \hat{S} introduz uma correlação entre vetores de entrada que não pode ser ignorada. Assim, a teoria da independência não pode ser usada na análise do comportamento do algoritmo LMS filtrado.

A correlação entre vetores de entrada pode ser facilmente observada a partir da análise de dois vetores de entrada consecutivos, como é mostrado a seguir:

$$\mathbf{X}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-N+1) \end{bmatrix} \quad \text{e} \quad \mathbf{X}(n-1) = \begin{bmatrix} x(n-1) \\ x(n-2) \\ \vdots \\ x(n-N) \end{bmatrix}$$

Estes vetores compartilham $N - 1$ componentes com apenas um componente diferente. Tomando o valor esperado em (2.9), temos:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] + \mu \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j E[\mathbf{X}(n-j)d(n)] - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)\mathbf{W}(n-i)] + \sum_{j=0}^{\hat{M}-1} \hat{s}_j E[\mathbf{X}(n-j)z(n)] \right) \quad (2.36)$$

Análises prévias [4,16,22] têm utilizado a teoria da independência (TI) [6,18] para determinar os valores esperados da Eq. (2.36). Isto significa que $\mathbf{X}(n-i)$ e $\mathbf{X}(n-j)$ são estatisticamente independentes para $i \neq j$. Esta condição nem sempre é verdadeira para sistemas representados por filtros transversais. No entanto, a TI é de fundamental importância para a maioria das análises estatísticas do LMS, sendo razoavelmente exata para uma taxa de aprendizagem pequena [19]. No caso filtrado, a atualização do vetor de coeficientes depende das correlações entre sucessivos vetores de entrada. Então, a TI não pode ser a base para uma análise estatística do comportamento do algoritmo FXLMS, exceto para casos muito particulares (por exemplo, quando a entrada é uma combinação linear de vetores independentes [26]). Através de exemplos vamos poder observar que o modelo obtido com a TI não consegue descrever corretamente o valor esperado dos

coeficientes do filtro adaptativo, mesmo para caminhos secundários com respostas ao impulso contendo poucos coeficientes.

Sem a utilização da TI a análise do FXLMS torna-se muito complexa, mesmo para o momento de primeira ordem [19]. Como pode-se observar, o valor esperado no terceiro termo da Eq. (2.36) depende das versões presentes e passadas do vetor de entrada, e não apenas daquele que está presente no instante de tempo considerado. Isto acarreta uma grande dificuldade matemática, se o que se pretende é uma análise exata. Neste ponto, deve-se recorrer a simplificações que resultem numa formulação matemática acessível e, ao mesmo tempo, mantenham a qualidade do modelo. De forma intuitiva, pode-se assumir como hipótese simplificadora que as correlações entre diferentes vetores de entrada são mais importantes do que as correlações entre os vetores de entrada e o vetor de coeficientes. Tal intuição é consolidada através de extensivas simulações numéricas em que foram comparados os valores esperados $E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)\mathbf{W}(n-i)]$ e $E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)]E[\mathbf{W}(n-i)]$ para vários valores de i , j , \mathbf{W}^o , \mathbf{S} e $\hat{\mathbf{S}}$. Na seção correspondente aos resultados experimentais são apresentadas simulações que ilustram esta suposição. Então, pode-se assumir que:

$$\text{II.i)} E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)|\mathbf{W}(n)\mathbf{W}(n-1)\dots\mathbf{W}(n-N+1)] \approx E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)], \forall i, j.$$

Assim,

$$\begin{aligned} \text{II.ii)} E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)\mathbf{W}(n)] &= E\{E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)|\mathbf{W}(n), \mathbf{W}(n-1), \dots]\mathbf{W}(n-i)\} \\ &= E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)]E[\mathbf{W}(n-i)], \forall i, j \end{aligned}$$

Note que (II.i) ou (II.ii) é menos restritiva do que a TI [18], que assume que vetores de entrada para instantes de tempo diferentes são independentes. Usando (II.ii) a correlação é preservada. Esta hipótese, porém, é menos restritiva do que a hipótese de independência entre os vetores de entrada e os vetores de coeficientes. Indicando à matriz de correlação do sinal de entrada por $\mathbf{R}_{i-j} = E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)]$, e chamando $\mathbf{P}_j = E[d(n)\mathbf{X}(n-j)]$ o vetor de correlação cruzada entre o sinal de entrada e o sinal desejado, a Eq. (2.36) pode ser rescrita como:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} E[\mathbf{W}(n-i)] + \mu \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j. \quad (2.37)$$

A fim de comparar os modelos, a mesma análise utilizando a TI leva à seguinte expressão:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] - \mu \sum_{i=0}^{\min(M, \hat{M})-1} s_i \hat{s}_i \mathbf{R}_{i-i} E[\mathbf{W}(n-i)] + \mu \hat{s}_0 \mathbf{P}_0; \quad (2.38)$$

A comparação fica mais simples se rescrevermos (2.37) na seguinte forma:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] - \mu \sum_{i=0}^{\min(M, \hat{M})-1} s_i \hat{s}_i \mathbf{R}_{i-i} E[\mathbf{W}(n-i)] + \mu \hat{s}_0 \mathbf{P}_0 - \mu \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ i \neq j}}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} E[\mathbf{W}(n-i)] + \mu \sum_{j=1}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j. \quad (2.39)$$

Desta forma, na segunda linha da Eq. (2.39) são evidenciados os termos não considerados pela TI.

2.4.1 Redução ao LMS clássico

A partir da equação (2.37), pode-se obter a expressão do momento de primeira ordem para o LMS clássico, em que \mathbf{S} e $\hat{\mathbf{S}}$ não representam operações de filtragem. Para isto devemos fazer:

$$s_0 = \hat{s}_0 = 1 \quad \text{e} \quad s_i = \hat{s}_j = 0, \quad 1 \leq i \leq M-1 \quad \text{e} \quad 1 \leq j \leq \hat{M}-1.$$

Com isto (2.37) fica da seguinte maneira:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] - \mu \mathbf{R}_0 E[\mathbf{W}(n)] + \mu \mathbf{P} = (\mathbf{I} - \mu \mathbf{R}_0) E[\mathbf{W}(n)] + \mu \mathbf{P}, \quad (2.40)$$

que é a conhecida expressão do momento de primeira ordem para os coeficientes do filtro adaptativo utilizando o algoritmo LMS clássico [5,6], sendo $\mathbf{R}_0 = E[\mathbf{X}(n)\mathbf{X}^T(n)]$ e $\mathbf{P} = E[d(n)\mathbf{X}(n)]$.

2.5. REGIME PERMANENTE QUANDO $\hat{\mathbf{S}} \neq \mathbf{S}$

A partir da expressão (2.37), que é dada em função dos coeficientes do filtro estimado $\hat{\mathbf{S}}$, determina-se o valor dos coeficientes em regime permanente ($n \rightarrow \infty$). Para n suficientemente grande, e assumindo que o algoritmo converge, observa-se que:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] = E[\mathbf{W}(n-1)] = \dots = E[\mathbf{W}(n-T)], \quad T \in \mathbf{Z}^+.$$

Aplicando esta condição na Eq. (2.37), e definindo o valor médio do vetor dos coeficientes em infinito como sendo $\mathbf{W}_\infty = \lim_{n \rightarrow \infty} E[\mathbf{W}(n)]$, temos que:

$$\sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} \mathbf{W}_\infty = \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j, \quad (2.41a)$$

ou

$$\tilde{\mathbf{R}}_{s\hat{s}} \mathbf{W}_\infty = \tilde{\mathbf{P}}_{\hat{s}}, \quad (2.41b)$$

onde

$$\tilde{\mathbf{R}}_{s\hat{s}} = \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} \quad (2.42a)$$

e

$$\tilde{\mathbf{P}}_{\hat{s}} = \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j. \quad (2.42b)$$

Determinando \mathbf{W}_∞ em (2.41b) para $\mathbf{R}_{s\hat{s}}$ não-singular, substituindo $\tilde{\mathbf{P}}_{\hat{s}} = \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o$ e utilizando (2.21), pode-se expressar a relação entre o valor em regime permanente e o valor ótimo dos coeficientes, quando $\hat{\mathbf{S}} \neq \mathbf{S}$. Tal relação é dada por:

$$\mathbf{W}_\infty = \tilde{\mathbf{R}}_{s\hat{s}}^{-1} \tilde{\mathbf{R}}_{\hat{s}} \tilde{\mathbf{R}}_{s\hat{s}}^{-1} \tilde{\mathbf{R}}_{s\hat{s}} \mathbf{W}_{or}. \quad (2.43)$$

No caso de uma identificação perfeita ($\hat{S} = S$) $\tilde{\mathbf{R}}_{s\hat{s}} = \tilde{\mathbf{R}}_{ss}$ e $\tilde{\mathbf{R}}_{\hat{s}} = \tilde{\mathbf{R}}_s$. Assim, o vetor ótimo é pré-multiplicado pela matriz identidade e \mathbf{W}_∞ corresponderá à solução ótima. Caso contrário, \mathbf{W}_{ot} é pré-multiplicado por uma matriz cheia, o que significa que cada componente de \mathbf{W}_∞ será uma combinação linear dos componentes do vetor ótimo.

2.5.1. Estimação do erro em excesso para o regime permanente

Para μ suficientemente pequeno, é possível obter uma estimação do erro em excesso para o regime permanente a partir do momento de primeira ordem. O erro em excesso [6] é definido como sendo a diferença entre o erro médio quadrático, produzido pelo algoritmo adaptativo, e o J_{\min} , que acontece quando se trabalha na condição ótima ($\mathbf{W} = \mathbf{W}_{ot}$). Assim,

$$J_{exc} \cong E[e^2(n) | \mathbf{W} = \mathbf{W}_\infty]_{\hat{S} \neq S} - J_{\min} \quad (2.44)$$

ou,

$$J_{exc} \cong \mathbf{W}_{ot}^T \left(\tilde{\mathbf{R}}_{s\hat{s}}^{-1} \tilde{\mathbf{R}}_s \tilde{\mathbf{R}}_s^{-1} \tilde{\mathbf{R}}_{ss} - \mathbf{I} \right) \tilde{\mathbf{R}}_{ss} \left(\tilde{\mathbf{R}}_{s\hat{s}}^{-1} \tilde{\mathbf{R}}_s \tilde{\mathbf{R}}_s^{-1} \tilde{\mathbf{R}}_{ss} - \mathbf{I} \right) \mathbf{W}_{ot} \quad (2.45)$$

Da mesma forma que na seção anterior, $\tilde{\mathbf{R}}_{s\hat{s}} = \tilde{\mathbf{R}}_{ss}$ e $\tilde{\mathbf{R}}_{\hat{s}} = \tilde{\mathbf{R}}_s$ se a estimação do caminho secundário for perfeita ($\hat{S} = S$). Neste caso, o erro em excesso é igual a zero. A validade desta aproximação será verificada nos exemplos apresentados ao final do capítulo.

2.6. VETOR DE ERRO NOS COEFICIENTES

Em análises estatísticas do desempenho de algoritmos adaptativos, o comportamento do vetor de coeficientes do filtro adaptativo é freqüentemente estudado através de sua posição em relação ao vetor ótimo no espaço de soluções. No caso do algoritmo LMS, por exemplo, estuda-se o comportamento do vetor erro nos coeficientes em relação à solução de Wiener para cada iteração [5,6]. Esta análise é bastante conveniente porque, no caso do algoritmo LMS clássico, a solução de Wiener corresponde à solução ótima e coincide com o valor médio do vetor dos coeficientes após a convergência. Assim, as expressões

matemáticas que determinam a evolução do vetor erro nos coeficientes tornam-se mais simples do que as expressões que expressam a evolução do vetor de pesos diretamente.

No caso do algoritmo LMS Filtrado, a presença dos filtros \mathbf{S} e $\hat{\mathbf{S}}$ modifica substancialmente o comportamento do algoritmo. Conforme mostrado anteriormente, a solução de Wiener não pode ser atingida usando-se um filtro adaptativo com resposta ao impulso de duração finita. Assim, o filtro ótimo \mathbf{W}_{ot} não coincide com a solução de Wiener. Além disso, a expressão (2.43) mostra que o valor médio do vetor dos coeficientes só atinge \mathbf{W}_{ot} após a convergência quando $\hat{\mathbf{S}} = \mathbf{S}$, uma situação praticamente impossível em implementações reais. Assim, a melhor forma de simplificar as expressões que determinam o comportamento do algoritmo é estudar a evolução do vetor dos coeficientes em relação ao seu valor em regime permanente \mathbf{W}_{∞} .

Definindo

$$\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}_{\infty} \quad (2.46)$$

e subtraindo \mathbf{W}_{∞} em ambos lados de (2.9), temos:

$$\begin{aligned} \mathbf{V}(n+1) = \mathbf{V}(n) + \mu & \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{X}(n-j) d(n) \right. \\ & \left. - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{X}(n-j) \mathbf{X}^T(n-i) (\mathbf{V}(n-i) + \mathbf{W}_{\infty}) + \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{X}(n-j) z(n) \right). \end{aligned} \quad (2.47)$$

Tomando valor esperado em (2.47), e utilizando as suposições (II.i) e (II.ii) para a variável $\mathbf{V}(n)$, obtém-se a seguinte expressão:

$$\begin{aligned} E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] + \mu & \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j E[\mathbf{X}(n-j) d(n)] \right. \\ & - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j E[\mathbf{X}(n-j) \mathbf{X}^T(n-i) \mathbf{V}(n-i)] \\ & \left. - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j E[\mathbf{X}(n-j) \mathbf{X}^T(n-i)] \mathbf{W}_{\infty} \right). \end{aligned} \quad (2.48)$$

O termo contendo o valor esperado de $\mathbf{X}(n-j)$ e $z(n)$ não aparece em (2.48), porque, por hipótese, $z(n)$ é de média zero e não correlacionado com nenhum outro sinal. Escrevendo (2.48) em função das matrizes \mathbf{R}_{i-j} , temos:

$$E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] + \mu \left(\sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} E[\mathbf{V}(n-i)] - \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} \mathbf{W}_\infty \right). \quad (2.49)$$

A partir da igualdade $\sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} \mathbf{W}_\infty = \sum_{j=0}^{\hat{M}-1} \hat{s}_j \mathbf{P}_j$, Eq. (2.41), a expressão (2.49) pode ser colocada em uma forma mais simples,

$$E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{\hat{M}-1} s_i \hat{s}_j \mathbf{R}_{i-j} E[\mathbf{V}(n-i)]. \quad (2.50)$$

A Eq.(2.50) descreve o comportamento do valor médio do vetor erro nos coeficientes. Note que (2.50) é uma equação de diferenças vetorial de ordem M , não tendo uma solução analítica simples, já que em geral as N equações são acopladas. Uma outra característica que diferencia o caso filtrado do caso clássico é que (2.50) não pode ser diagonalizada. Isto acontece porque, o termo com duplo somatório contém termos (matrizes) não diagonalizáveis, mesmo que a entrada seja ruído branco. Em [22] foi considerada tal diagonalização a partir da suposição de aprendizado muito lento. Em outras palavras, a aproximação utilizada em [22] foi $E[\mathbf{V}(n-i)] \approx E[\mathbf{V}(n)]$, reduzindo assim (2.50) a uma equação de diferenças de primeira ordem. Nesta análise não fazemos esta aproximação. Finalmente, neste trabalho também não fazemos nenhuma consideração em relação à natureza dos sinais, ruído branco ou colorido, que geram a matriz \mathbf{R}_{i-j} . Portanto, a Eq. (2.50) é válida para qualquer tipo de sinal de entrada, tal como será comprovado através de resultados experimentais.

2.7 EFEITO DOS ERROS DE ESTIMAÇÃO DO CAMINHO SECUNDÁRIO

Na maioria das aplicações práticas é impossível obter uma estimativa perfeita do caminho secundário \mathbf{S} . Então, como parâmetro de projeto, é importante obter uma avaliação dos

efeitos dos erros cometidos na estimação da resposta S . O efeito de tais erros foi previamente estudado para sinais de referência periódicos [4, 27, 22, 28]. Estes erros também foram investigados em [16] para o algoritmo DLMS (*Delayed-LMS*) com ruído branco de entrada. Em [16] foi utilizado um modelo analítico baseado na teoria da independência. Esta seção apresenta o estudo dos efeitos dos erros de estimação de S no comportamento do valor médio dos coeficientes para o algoritmo FXLMS utilizando ruído branco como sinal de entrada.

Considere o problema de filtragem linear da Fig. 2.3, onde \mathbf{W} e S são assumidos lineares e invariantes no tempo. Neste caso podemos inverter a ordem das respostas \mathbf{W} e S sem alterar o comportamento do sistema. Na nova configuração fica claro que o vetor \mathbf{W}_{or} corresponde à solução de Wiener do problema de filtragem linear com entrada dada por:

$$\mathbf{X}_f(n) = [x_f(n) \ x_f(n-1) \ \cdots \ x_f(n-N+1)]^T = \sum_{i=0}^{M-1} s_i \mathbf{X}(n-i) \quad (2.51)$$

Pelo princípio da ortogonalidade, o sinal de erro é ortogonal à entrada $x_f(n)$, Eq. (2.51), para $\mathbf{W} = \mathbf{W}_{or}$. Considerando a Fig. 2.2, assumindo por simplicidade que $\hat{M} = M$, usando a Eq. (2.7) e definindo o vetor de erro nos coeficientes como sendo $\bar{\mathbf{V}}(n) = \mathbf{W}(n) - \mathbf{W}_{or}$, o sinal de erro é escrito como sendo:

$$e(n) = e_o(n) - \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \bar{\mathbf{V}}(n-i), \quad (2.52)$$

onde:

$$e_o(n) = d(n) + z(n) - \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \mathbf{W}_{or} \quad (2.53)$$

é o erro na solução ótima. Modelando os erros de estimação através do modelo $\hat{s}_i = s_i + \delta_i$, $i = 0, \dots, M-1$, com $\Delta = [\delta_0 \ \delta_1 \ \cdots \ \delta_{M-1}]^T$, sendo estatisticamente independente de qualquer entrada, substituindo (2.52) em (2.8) e expandindo os termos

como funções de $\mathbf{V}(n)$ e $\mathbf{X}_f(n)$, o valor esperado condicionado em Δ é dado pela seguinte expressão¹:

$$\begin{aligned}
 E[\bar{\mathbf{V}}(n+1)|\Delta] = & E[\bar{\mathbf{V}}(n)|\Delta] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{R}_{i-j} E[\bar{\mathbf{V}}(n-i)|\Delta] \\
 & + \mu E[e_o(n) \mathbf{X}_\delta(n)|\Delta] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i \delta_j \mathbf{R}_{i-j} E[\bar{\mathbf{V}}(n-i)|\Delta],
 \end{aligned} \tag{2.54}$$

onde $\mathbf{X}_\delta(n) = \sum_{i=0}^{M-1} \delta_i \mathbf{X}(n-i)$. O regime permanente do valor esperado do vetor erro é dado por:

$$E[\bar{\mathbf{V}}(\infty)|\Delta] = (\tilde{\mathbf{R}}_{ss} + \tilde{\mathbf{R}}_{s\delta})^{-1} (\tilde{\mathbf{P}}_\delta - \tilde{\mathbf{R}}_{s\delta} \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{P}}_s) = (\tilde{\mathbf{R}}_{ss} + \tilde{\mathbf{R}}_{s\delta})^{-1} (\tilde{\mathbf{P}}_\delta - \tilde{\mathbf{R}}_{s\delta} \mathbf{W}_{opt}), \tag{2.55}$$

onde,

$$\tilde{\mathbf{P}}_\delta = \sum_{i=0}^{M-1} \delta_i E[d(n) \mathbf{X}(n-i)|\Delta] \text{ e } \tilde{\mathbf{R}}_{s\delta} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i \delta_j \mathbf{R}_{i-j}.$$

As Eqs. (2.54) e (2.55) mostram a maneira em que os erros de estimação do caminho secundário afetam os regimes transitórios e permanente do algoritmo FXLMS. Note que para estimação *on-line* o valor esperado não é afetado se os δ_i 's possuem média zero [16] (o valor esperado de (2.54) neste caso reduz-se a (2.50)). Por outro lado, os erros alteram a velocidade de convergência e originam um desvio do regime permanente para uma estimação *off-line* (mesmos δ_i 's para todas as realizações - não aleatório). Também, a Eq. (2.54) reduz-se à forma (2.50) para $\Delta = \mathbf{0}$ e $\hat{\mathbf{S}} = \mathbf{S}$.

2.8. RESULTADOS DE SIMULAÇÕES

Nesta seção são apresentados resultados de simulações que avaliam a qualidade do novo modelo analítico proposto. Primeiramente são mostradas simulações que ilustram a

¹ O Apêndice A apresenta os detalhes da obtenção da Eq. (2.54).

suposição (II)-(seção 2.4); e que permite separar o cálculo dos valores esperados a fim de viabilizar o tratamento matemático. Em seguida, são mostrados diferentes exemplos onde são apresentadas as curvas correspondentes à evolução temporal do valor médio dos coeficientes do filtro adaptativo obtidas a partir da simulação e do modelo. Também, é realizada uma comparação entre o novo modelo e o modelo clássico, que utiliza a teoria da independência, a fim de mostrar o efeito de não considerar a correlação entre os vetores de entrada. Finalmente, é mostrada a evolução do erro médio quadrático a fim de verificar a aproximação (2.44) para determinar o erro em excesso.

2.8.1 Avaliação da suposição (II)

Para ilustrar a suposição (II) vamos utilizar uma planta representada pela seguinte resposta ao impulso: $\mathbf{W}^o = [1,5; 1,0; -0,7]$. O caminho secundário é modelado pela resposta ao impulso $\mathbf{S} = [0,2; 0,8]$ e assumimos a condição de estimação perfeita $\hat{\mathbf{S}} = \mathbf{S}$. O máximo valor para a taxa de aprendizagem, determinado de forma experimental, é $\mu_{\max} = 0,5$. Como sinal de entrada é utilizado um ruído branco gaussiano de variância $\sigma_x^2 = 1$. A fim de testar a suposição realizada são utilizados os seguintes conjuntos de valores para a taxa de aprendizagem e o ruído de medição (branco e gaussiano): $\mu = 0,1$ e $\sigma_z^2 = 0,0001$ (SNR 40dB); $\mu = 0,1$ e $\sigma_z^2 = 0,01$ (SNR 20dB); $\mu = 0,25$ e $\sigma_z^2 = 0,0001$ (SNR 40dB); $\mu = 0,25$ e $\sigma_z^2 = 0,01$ (SNR 20dB).

As Figs. 2.6a e 2.6b mostram a evolução temporal do valor médio dos coeficientes do filtro adaptativo para os dois valores de μ (0,1 e 0,25) e com SNR de 40 dB. Nessas figuras pode-se observar o excelente casamento entre a simulação e o modelo proposto, para os dois diferentes valores da taxa de aprendizagem. O modelo que utiliza a teoria da independência, claramente, não reflete o comportamento real do valor médio dos coeficientes.

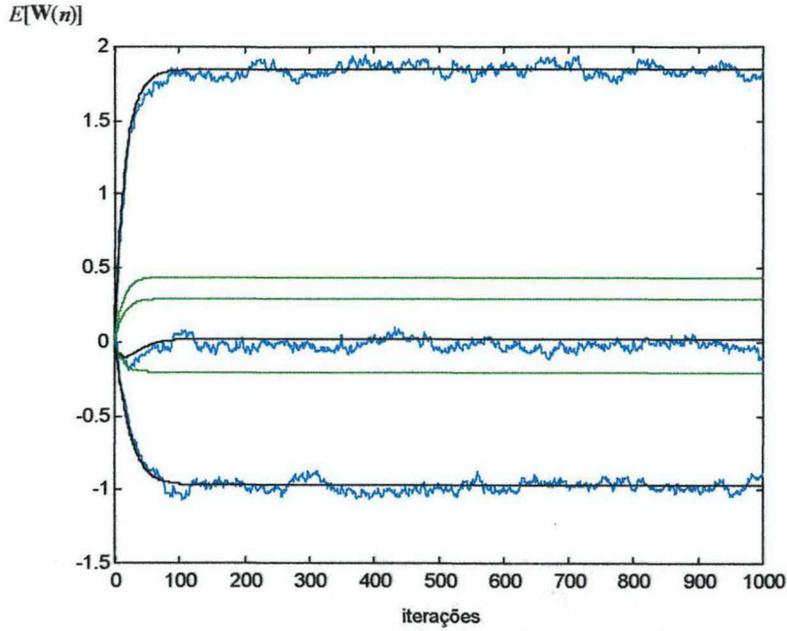


Fig. 2.6a . Comportamento do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,1$ e SNR 40dB. $E[\mathbf{W}(n)]$ obtida usando (2.37) (linha preta), simulação (linha azul) obtida pela média de 100 realizações e modelo baseado na teoria da independência Eq. (2.38) (linha verde).

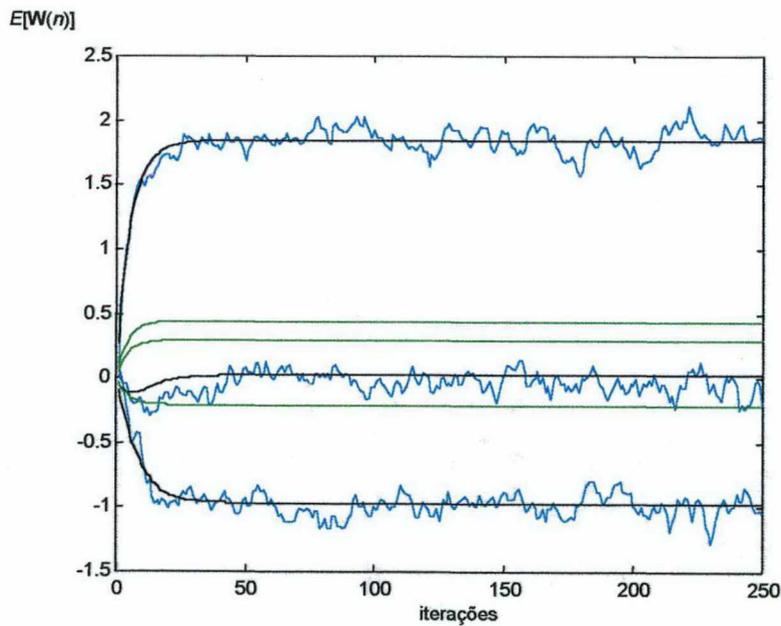


Fig. 2.6b. Comportamento do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,25$ e SNR 40dB. $E[\mathbf{W}(n)]$ obtida usando (2.37) (linha preta), simulação (linha azul) obtida pela média de 100 realizações e modelo baseado na teoria da independência Eq. (2.38) (linha verde).

As Figs. 2.7a, 2.7b e 2.7c ilustram a suposição (II.ii)-(seção 2.4) para diferentes valores de i e j e $\mu = 0,1$. Essas curvas são obtidas através de simulações fazendo-se a média de

100 realizações. Nessas figuras são mostradas duas curvas. Uma delas corresponde ao cálculo do valor esperado dado por $E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{W}(n-j)]$ e a outra realizando o cálculo dos valores esperados obtido de forma separada, ou seja $E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{W}(n-j)]$. A partir dos resultados obtidos, isto é, o bom casamento entre as curvas, podemos assumir que a suposição realizada é aceitável, permitindo o desenvolvimento matemático do modelo proposto.

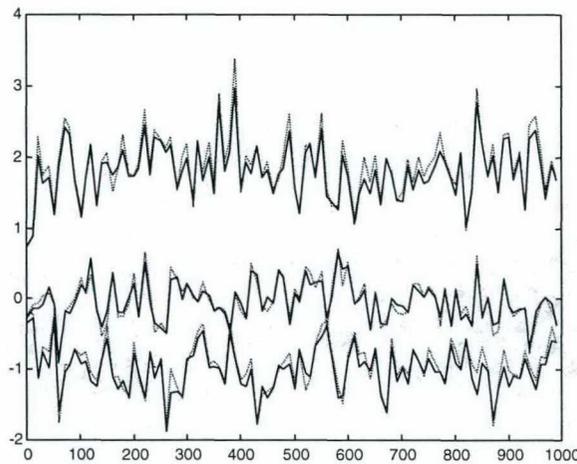


Fig. 2.7a. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}(n)\mathbf{W}(n)]$ (linha cheia). $\mathbf{R}_0 E[\mathbf{W}(n)]$ (linha pontilhada).

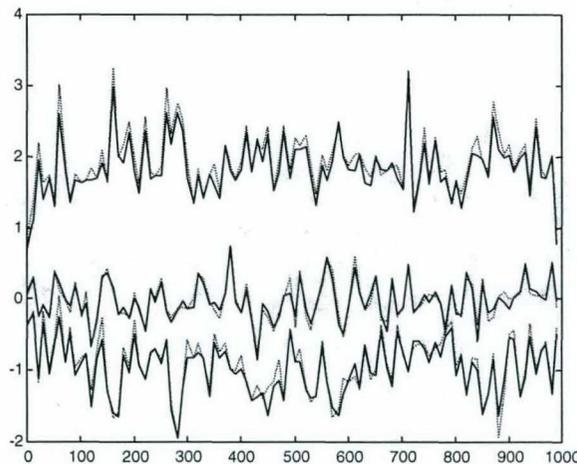


Fig. 2.7b. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}(n-1)\mathbf{W}(n-1)]$ (linha cheia). $\mathbf{R}_1 E[\mathbf{W}(n-1)]$ (linha pontilhada).

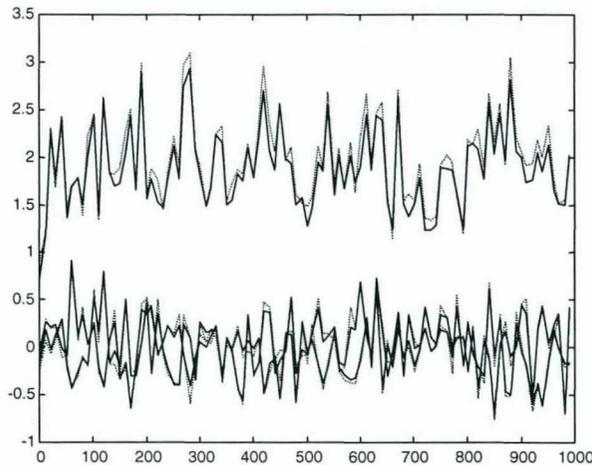


Fig. 2.7c. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-2)\mathbf{W}(n-2)]$ (linha cheia). $\mathbf{R}_2 E[\mathbf{W}(n-2)]$ (linha pontilhada).

As Figs. 2.8a, 2.8b e 2.8c ilustram a suposição (II.ii)-(seção 2.4) para diferentes valores de i e j e $\mu = 0,25$. Essas curvas são obtidas através de simulações fazendo-se a média de 100 realizações. Note que para o valor de μ utilizado, que representa o 50% do valor limite, as duas curvas apresentam diferenças significativas. Isto significa que a suposição começa a perder validade à medida que a taxa de aprendizagem assume valores que estão perto da instabilidade. No entanto, o desvio das suposições não parece afetar o momento de primeira ordem dos coeficientes do filtro adaptativo. Por outro lado, como comprovaremos mais adiante, o modelo para o momento de segunda ordem evidencia este afastamento entre as curvas, resultando em que o modelo teórico proposto já não acompanha a simulação.

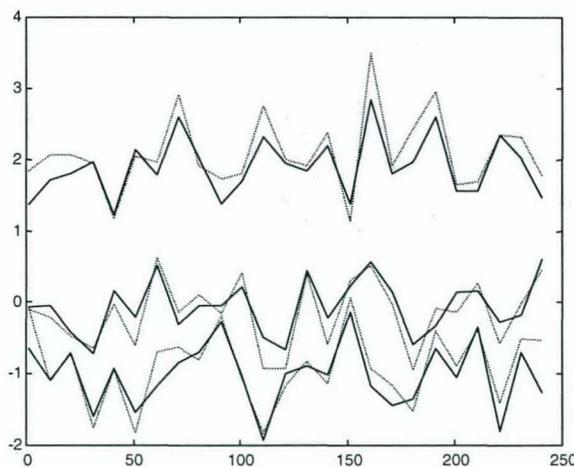


Fig. 2.8a. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}^T(n)\mathbf{W}(n)]$ (linha cheia). $\mathbf{R}_0 E[\mathbf{W}(n)]$ (linha pontilhada).

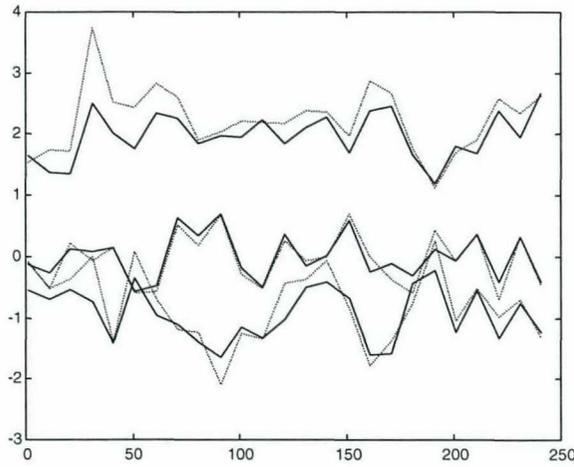


Fig. 2.8b. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-1)\mathbf{W}(n-1)]$ (linha cheia). $\mathbf{R}_1 E[\mathbf{W}(n-1)]$ (linha pontilhada).

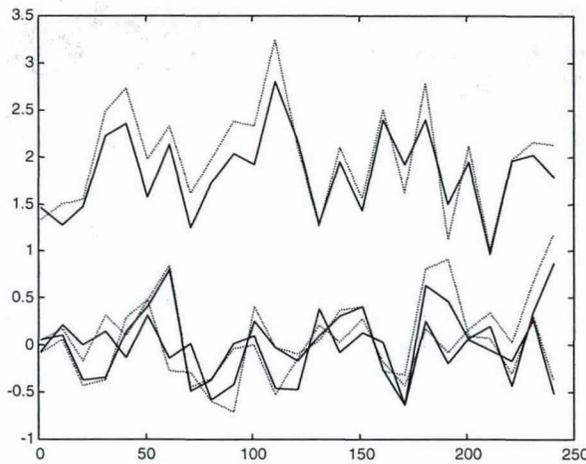


Fig. 2.8c. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 40dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-2)\mathbf{W}(n-2)]$ (linha cheia). $\mathbf{R}_2 E[\mathbf{W}(n-2)]$ (linha pontilhada).

A Fig. 2.9 mostra a evolução temporal do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,1$ e com SNR de 20 dB. Nessa figura pode-se observar que o aumento da SNR não modificou o casamento apresentado entre a simulação e o modelo proposto, para a taxa de aprendizagem utilizada. O modelo que utiliza a teoria da independência, claramente, não reflete o comportamento real do valor médio dos coeficientes.

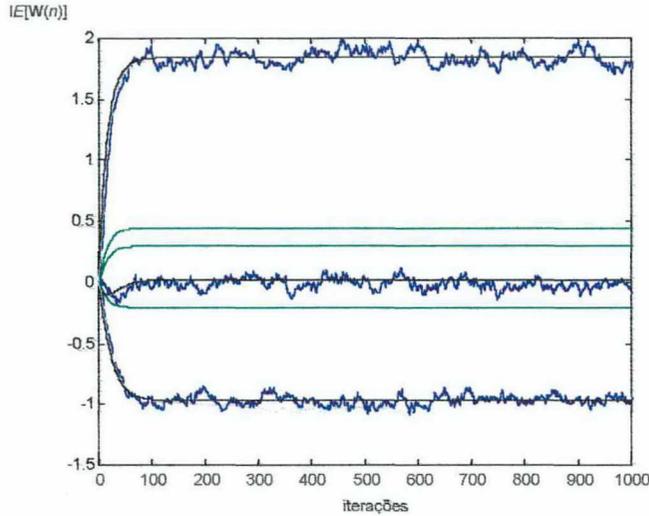


Fig. 2.9. Comportamento do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,1$ e SNR 20dB. $E[\mathbf{W}(n)]$ obtida usando (2.37) (linha preta), simulação (linha azul) obtida pela média de 100 realizações e modelo baseado na teoria da independência Eq. (2.38) (linha verde).

As Figs. 2.10a, 2.10b e 2.10c ilustram a suposição (II.ii) para diferentes valores de i e j e $\mu = 0,1$. A SNR utilizada é de 20 dB. Essas curvas são obtidas através de simulações fazendo-se a média de 100 realizações. Nessas figuras são mostradas duas curvas. Uma delas corresponde ao cálculo do valor esperado dado por $E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{W}(n-j)]$ e a outra realizando o cálculo dos valores esperados realizado em forma separada, ou seja $E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{W}(n-j)]$. Novamente, dos resultados obtidos para uma SNR diferente, vemos que o bom casamento entre as curvas continua sendo aceitável, validando assim a suposição realizada.

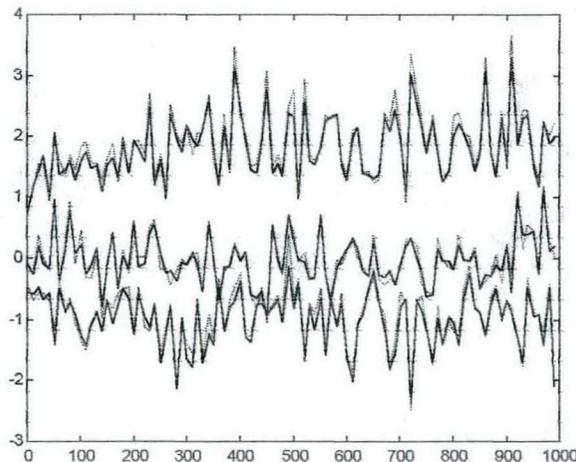


Fig. 2.10a. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n)\mathbf{W}(n)]$ (linha cheia). $\mathbf{R}_0 E[\mathbf{W}(n)]$ (linha pontilhada).

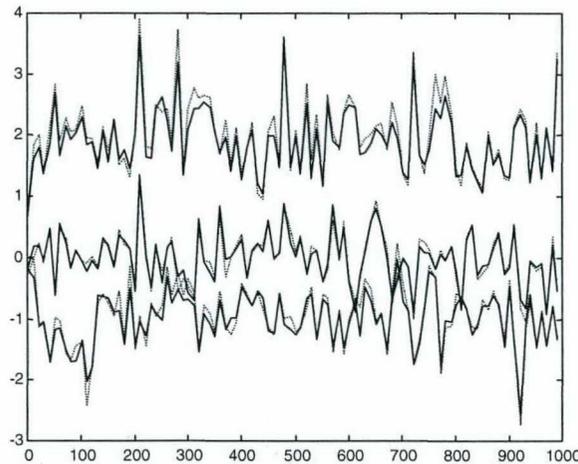


Fig. 2.10b. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-1)\mathbf{W}(n-1)]$ (linha cheia). $\mathbf{R}_1 E[\mathbf{W}(n-1)]$ (linha pontilhada).

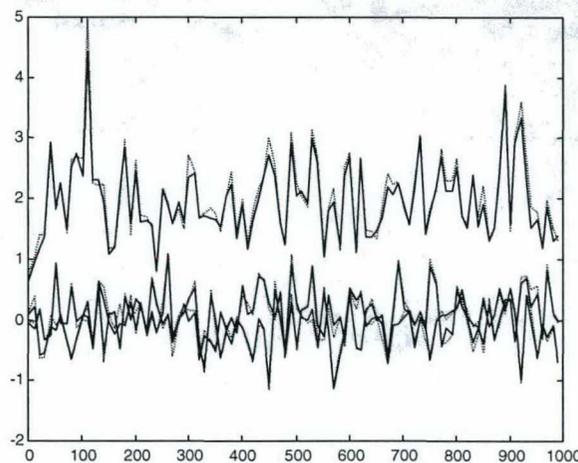


Fig. 2.10c. Teste da suposição (II.ii) para $\mu = 0,1$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-2)\mathbf{W}(n-2)]$ (linha cheia). $\mathbf{R}_2 E[\mathbf{W}(n-2)]$ (linha pontilhada).

A Fig. 2.11 mostra a evolução temporal do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,25$, com SNR de 20 dB. Dessa figura pode-se observar que o aumento da SNR não modificou o casamento apresentado entre a simulação e o modelo proposto, para os valores de taxa de aprendizagem utilizados. O modelo que utiliza a teoria da independência, claramente, não reflete o comportamento real do valor médio dos coeficientes.

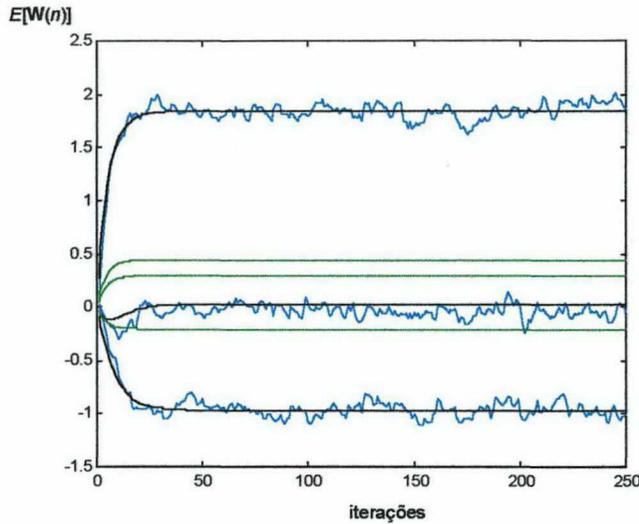


Fig. 2.11. Comportamento do valor médio dos coeficientes do filtro adaptativo para $\mu = 0,25$ e SNR 20dB. $E[\mathbf{W}(n)]$ obtida usando (2.37) (linha preta), simulação (linha azul) obtida pela média de 50 realizações e modelo baseado na teoria da independência Eq. (2.38) (linha verde).

As Figs. 2.12a, 2.12b e 2.12c ilustram a suposição (II.ii) para diferentes valores de i e j e $\mu = 0,25$ e SNR 20dB. Essas curvas são obtidas através de simulações fazendo-se a média de 100 realizações. Novamente, para o valor de μ utilizado, que representa 50% do valor limite, as duas curvas apresentam diferenças significativas. Isto significa que a suposição começa a perder validade na medida em que a taxa de aprendizagem assume valores que estão perto da instabilidade e não devido ao valor da SNR.

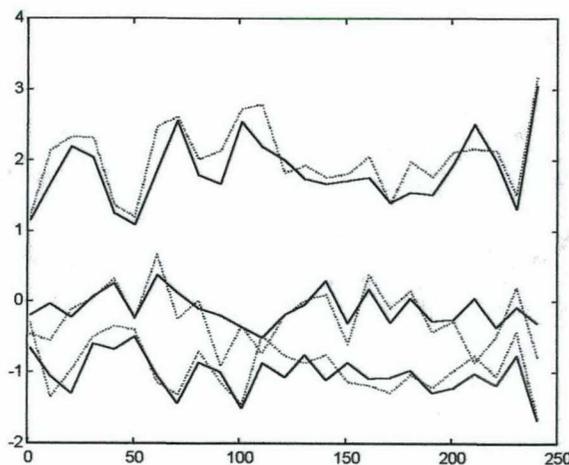


Fig. 2.12a. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n)\mathbf{W}(n)]$ (linha cheia). $\mathbf{R}_0 E[\mathbf{W}(n)]$ (linha pontilhada).

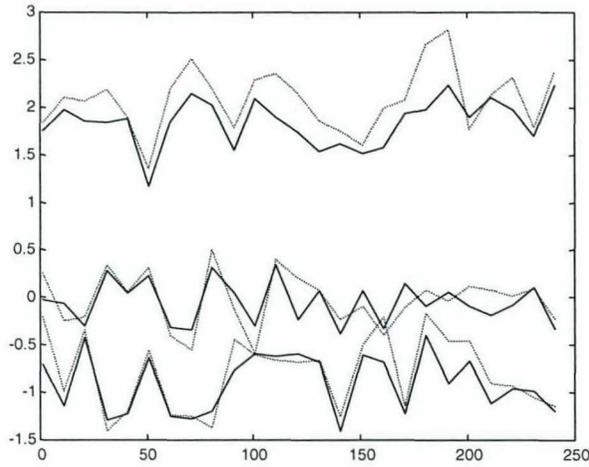


Fig. 2.12b. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-1)\mathbf{W}(n-1)]$ (linha cheia). $\mathbf{R}_1 E[\mathbf{W}(n-1)]$ (linha pontilhada).

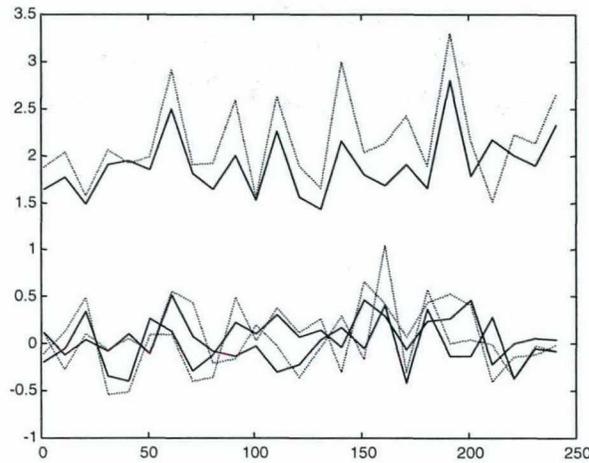


Fig. 2.12c. Teste da suposição (II.ii) para $\mu = 0,25$ e SNR 20dB. $E[\mathbf{X}(n)\mathbf{X}^T(n-2)\mathbf{W}(n-2)]$ (linha cheia). $\mathbf{R}_2 E[\mathbf{W}(n-2)]$ (linha pontilhada).

2.8.2 Avaliação do modelo: Caso 1

Para avaliar o modelo correspondente ao momento de primeira ordem dos coeficientes do filtro adaptativo, como planta utilizaremos a função "coseno levantado" com 11 amostras, sendo sua resposta ao impulso dada por:

$$\mathbf{W}^o = [1,0; 0,8415; 0,4546; 0,0470; -0,1892; -0,1918; -0,0466; 0,0939; 0,1237; 0,0458; -0,0544].$$

O caminho secundário é modelado pela resposta ao impulso $\mathbf{S} = [1,5; 1,0; -0,7]$, a taxa de aprendizagem é $\mu = 0,002$ ($(\mu_{\max})_{LMS} = 0,0606$) e assumimos a condição de estimação perfeita $\hat{\mathbf{S}} = \mathbf{S}$. O sinal de entrada é um ruído branco gaussiano de variância $\sigma_x^2 = 1$. O ruído de medição é um ruído branco gaussiano de variância $\sigma_z^2 = 0,0001$. Na Fig. 2.13a é mostrada a relação entre as bandas passantes correspondentes à planta e ao caminho secundário. Considerando a banda passante o valor de freqüências correspondente a 3 dB de atenuação, constata-se que o caminho secundário tem uma banda passante maior do que a da planta.

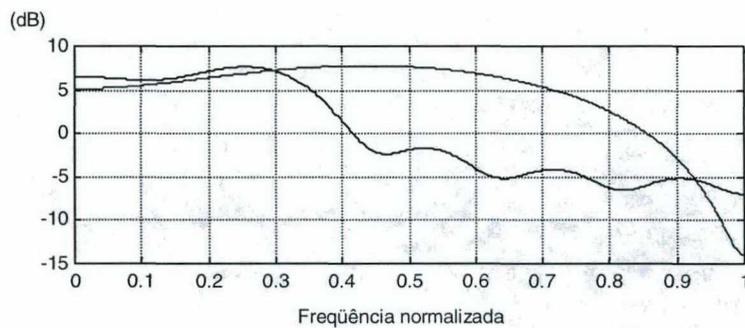


Fig. 2.13a. Relação entre as bandas passantes da planta (linha preta) e o caminho secundário (linha azul).

Nas Figs. 2.13b e 2.13c, é mostrado o comportamento dos coeficientes do filtro adaptativo obtido por simulação, o qual é comparado com o modelo proposto (Eqs. (2.37) e (2.50)). A diferença entre as figuras reside no número de realizações utilizadas na simulação. Como pode-se observar, à medida em que se aumenta o número de realizações, tanto mais a simulação se aproxima do modelo, tal como se devia esperar.

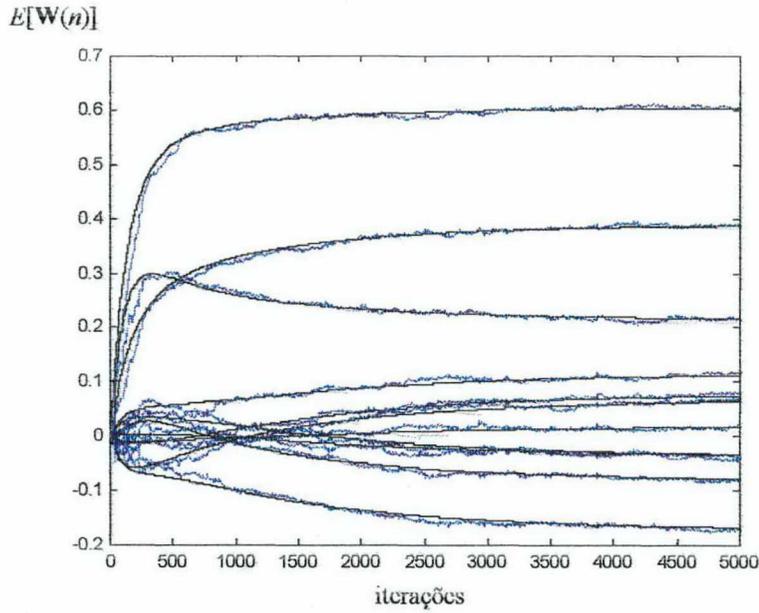


Fig. 2.13b. Uma realização dos coeficientes do filtro adaptativo, simulação (linha azul) e modelo proposto (linha preta) Eqs. (2.37) e (2.50).

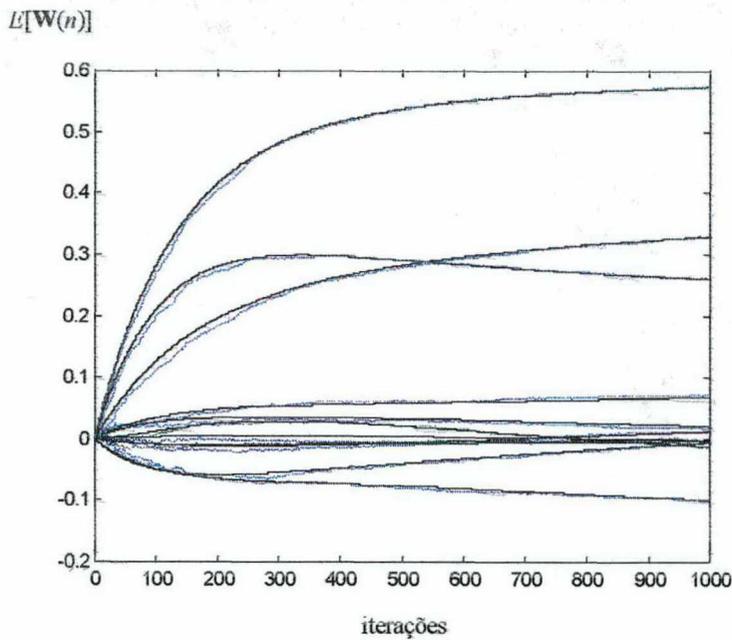


Fig. 2.13c. Valor esperado (10 realizações) dos coeficientes do filtro adaptativo, primeiras 1000 iterações. Simulação (linha azul) e modelo proposto (linha preta) Eqs. (2.37) e (2.50).

A Fig. 2.13d mostra a comparação entre o modelo proposto Eq. (2.37) e o modelo que utiliza a teoria da independência Eq. (2.38). Como pode-se observar, este último não consegue prever de forma adequada o comportamento dos coeficientes devido ao efeito dos termos não considerados indicados pela Eq. (2.39). A predição para o valor em regime permanente da Eq. (2.43) é

$$\mathbf{W}_\infty = [0,6034; 0,2165; 0,3877; -0,0784; 0,0638; -0,1677; 0,0752; -0,0337; 0,1117; -0,0328; 0,0164]$$

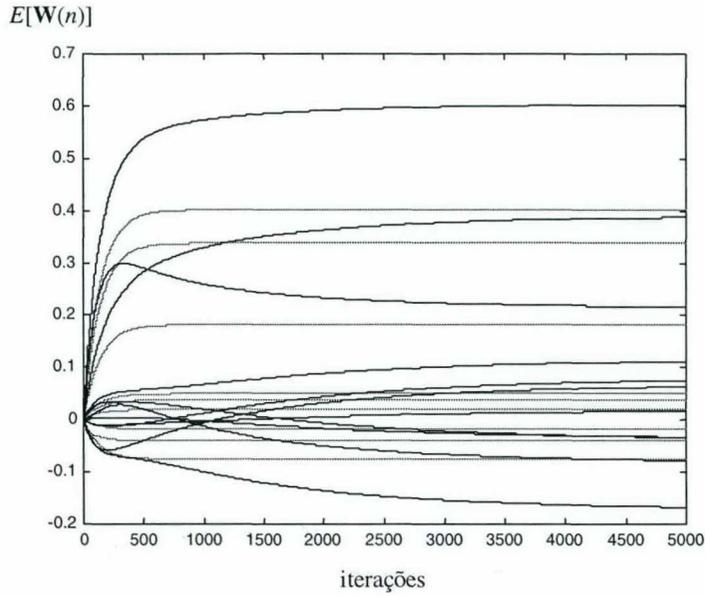


Fig. 2.13d. Comparação entre o modelo proposto Eq. (2.37) (linha preta) e o modelo baseado na teoria da independência Eq. (2.38) (linha verde).

A Fig. 2.13e mostra a evolução do erro médio quadrático obtido pela média de 50 realizações; nessa figura é possível verificar a validade da aproximação apresentada pela Eq. (2.44).

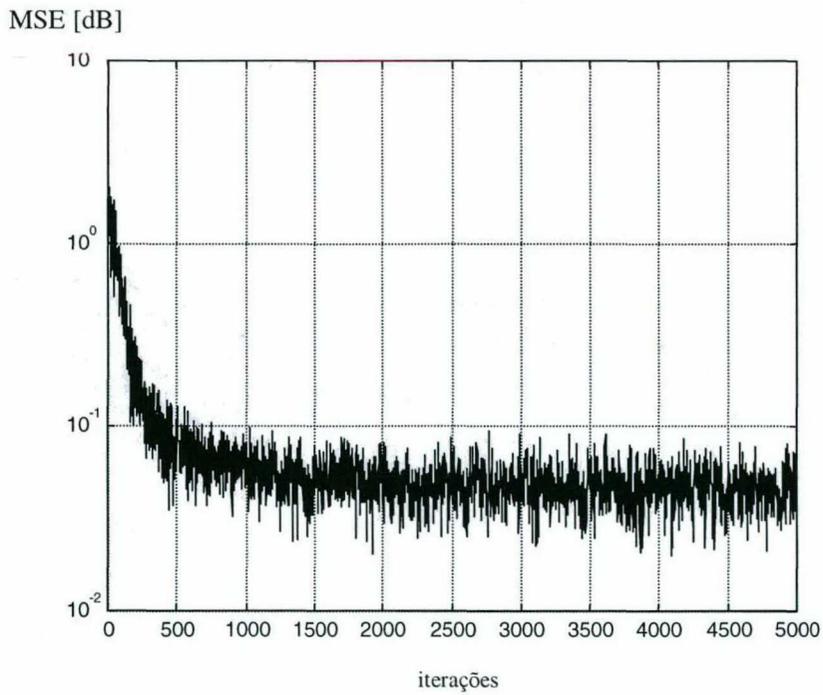


Fig. 2.13e. Estimação do erro em excesso, medido 0.0479; estimado pela Eq. (2.44) 0.0481.

2.8.3 Avaliação do modelo: Caso 2

Neste exemplo é estudado o caso de estimação imperfeita. A planta é representada pela resposta ao impulso $\mathbf{W}^o = [1,5; 1,0; -0,7]$, o caminho secundário é modelado pela resposta ao impulso $\mathbf{S} = [1,0; -0,5]$ e sua estimação por $\hat{\mathbf{S}} = [1,0; 0,5]$, a taxa de aprendizagem é $\mu = 0,003$ ($(\mu_{\max})_{LMS} = 0,222$). O sinal de entrada é um ruído branco gaussiano de variância $\sigma_x^2 = 1$. O ruído de medição é um ruído branco gaussiano de variância $\sigma_z^2 = 0,0001$. Na Fig. 2.14a são mostradas as bandas passantes da planta, do caminho secundário e da estimação do caminho secundário. Note que \mathbf{S} é um sistema passa-altas e $\hat{\mathbf{S}}$ é um sistema passa-baixas.

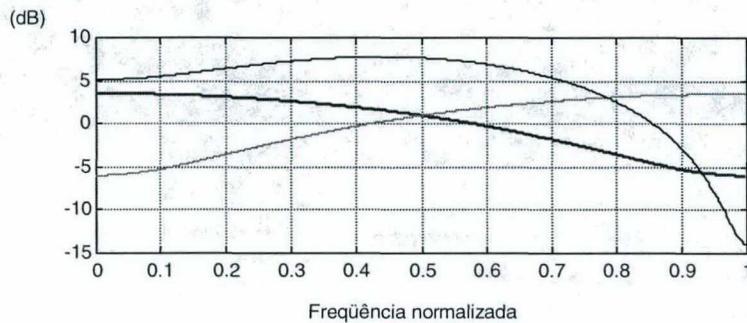


Fig. 2.14a. Relação entre as bandas passantes da planta (linha preta), caminho secundário (linha verde) e estimação (linha azul).

Na Fig. 2.14b são mostrados a evolução do valor médio dos coeficientes resultantes da simulação, obtida através da média de 10 realizações, e os coeficientes determinados pelo modelo proposto nas Eqs. (2.37) e (2.50). A Fig. 2.14c apresenta uma comparação do modelo da Eq. (2.38) e (2.50) com o modelo gerado usando a teoria da independência. Pode-se novamente observar a qualidade do novo modelo para prever o comportamento dos valor esperado dos coeficientes quando comparado com o modelo obtido pela TI. A Fig. 2.14c não inclui as curvas da simulação para uma melhor clareza. A Fig. 2.14b já mostrou a semelhança entre a simulação e os resultados usando o novo modelo. O valor em regime permanente dado pela Eq. (2.43) é $\mathbf{W}_\infty = [1,5137; 1,7294; 0,2196]$, que pode ser verificado na Fig. 2.14b.

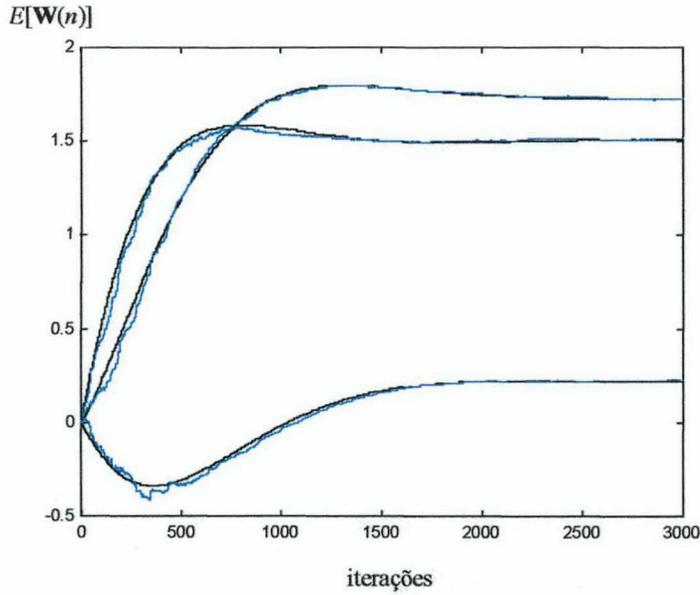


Fig. 2.14b. Valor esperado (10 realizações) dos coeficientes do filtro adaptativo. Simulação (linha azul) e modelo proposto (linha preta) Eqs. (2.37), (2.50).

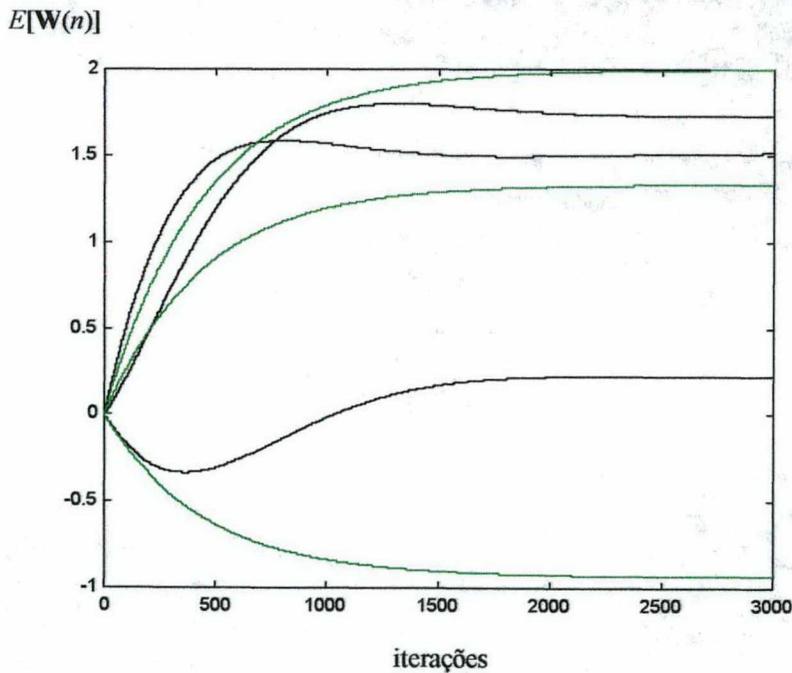


Fig. 2.14c. Comparação entre o modelo proposto Eq. (2.37) (linha preta) e o modelo baseado na teoria da independência Eq. (2.38) (linha verde).

A Fig. 2.14d mostra a evolução do erro médio quadrático obtido pela média de 50 realizações; nessa figura é possível verificar a validade da aproximação obtida pela Eq. (2.44).

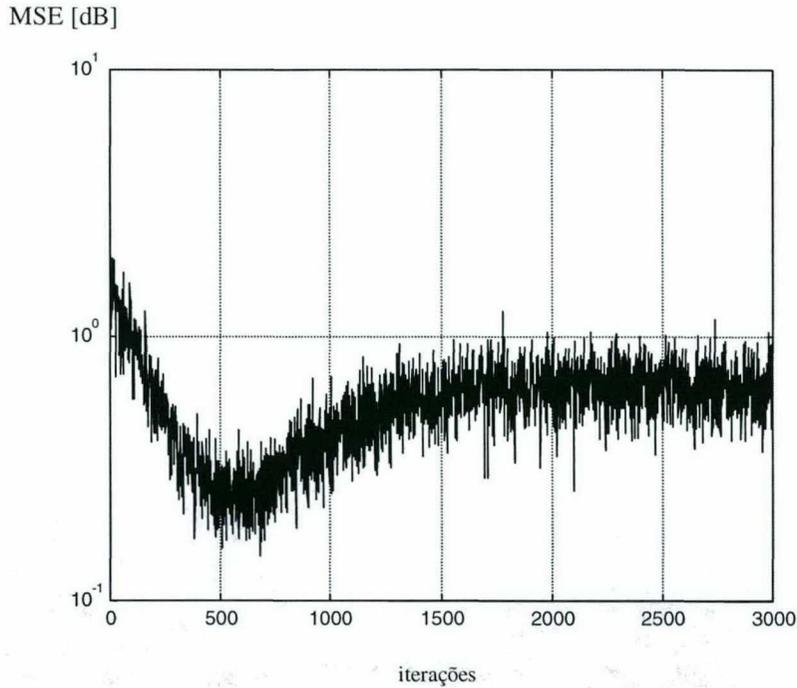


Fig. 2.14d. Estimação do erro em excesso; medido 0.6582; estimado pela Eq. (2.44) 0.6581.

2.8.4 Avaliação do modelo: Caso 3

Neste exemplo é considerada uma entrada correlacionada. O sinal correlacionado é obtido a partir de um modelo AR (autorregressivo) de segunda ordem, dado pela seguinte expressão:

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + u(n)$$

onde $u(n)$ é um ruído branco gaussiano com variância unitária. Os coeficientes do processo usado são $a_1 = -0,1950$ e $a_2 = 0,95$, a dispersão dos autovalores da matriz de correlação de $x(n)$ é $\chi = 1,22$. A planta e o caminho secundário são representados pelas respostas ao impulso dadas por $\mathbf{W}^o = [1,5; 1,0; -0,7]^T$, $\hat{\mathbf{S}} = \mathbf{S} = [1,0; -0,5]^T$, respectivamente; e $\mu = 0,003$ ($(\mu_{\max})_{LMS} = 0,222$). Na Fig. 2.15a é mostrada a relação entre as bandas passantes da planta e do caminho secundário.

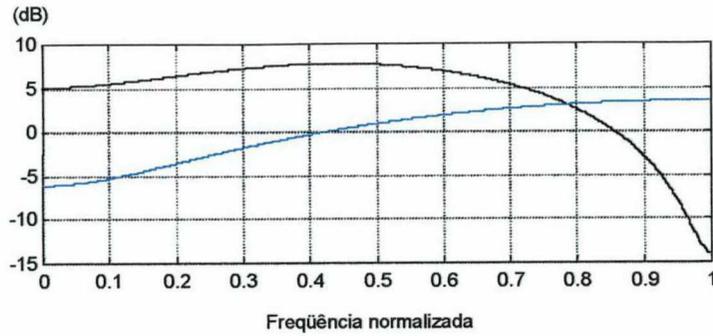


Fig. 2.15a. Relação entre as bandas passantes da planta (linha preta) e do caminho secundário (linha azul).

Dado que não foi feita hipótese alguma em relação às propriedades estatísticas do sinal de referência além da estacionaridade, o modelo proposto é válido para sinais brancos ou correlacionados. Tal fato pode ser observado pela Fig. 2.15b. Para o caso correlacionado, a TI também não proporciona uma descrição correta da evolução dos coeficientes, como fica claro da comparação realizada na Fig. 2.15c.

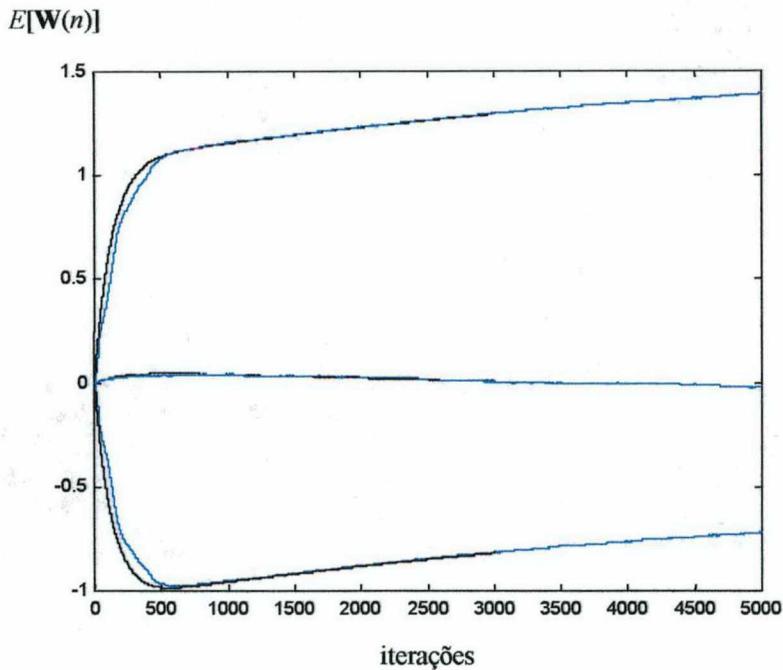


Fig. 2.15b. Valor esperado (10 realizações) dos coeficientes do filtro adaptativo para entrada correlacionada. Simulação (linha azul) e modelo proposto (linha preta) Eq. (2.37).

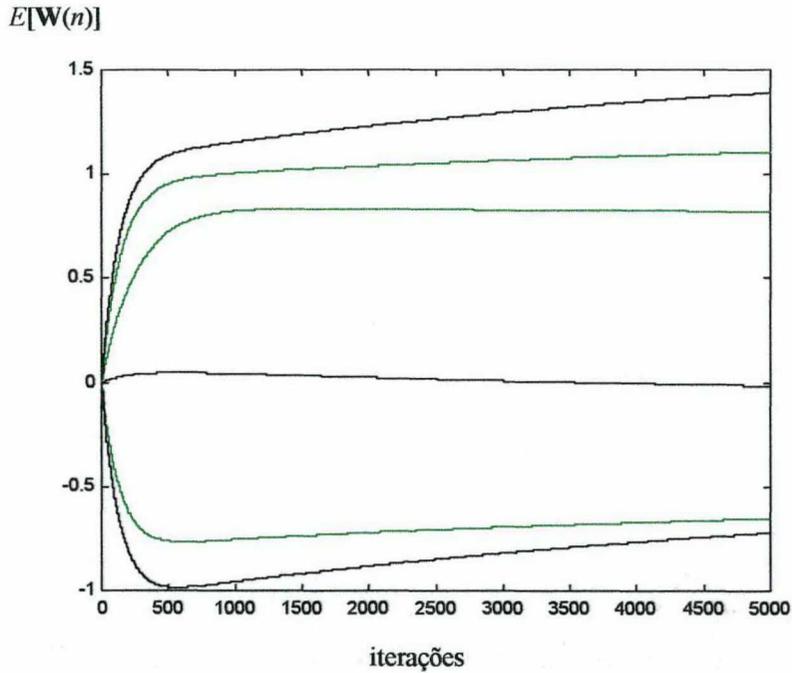


Fig. 2.15c. Comparação entre o modelo proposto Eq. (2.37) (linha cheia), e o modelo baseado na teoria da independência Eq. (2.38) (linha pontilhada).

O valor de regime permanente dado pela Eq. (2.43) é $\mathbf{W}_\infty = [1,6198; -0,0726; -0,4925]$. Esse resultado pode ser verificado na Fig. 2.15d que mostra a evolução dos coeficientes até atingir o regime permanente.

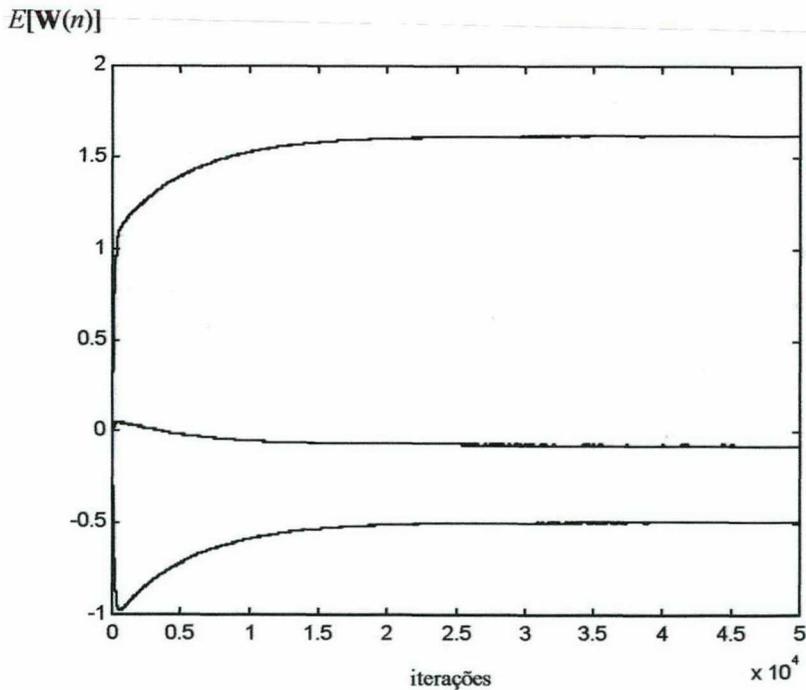


Fig. 2.15d. Evolução temporal do valor esperado (10 realizações) dos coeficientes do filtro adaptativo até atingir o regime permanente.

A Fig. 2.15e mostra a evolução do erro médio quadrático obtido pela média de 50 realizações; nessa figura é possível verificar a validade da aproximação obtida em (2.44).

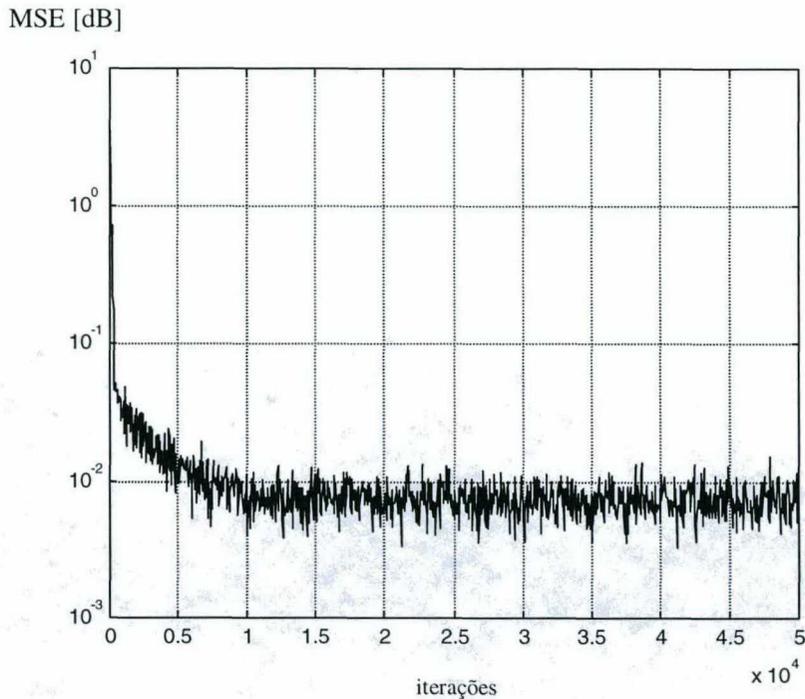


Fig. 2.15e. Estimação do erro em excesso; medido 0.0075; estimado pela Eq. (2.44) 0.0075.

2.9. CONCLUSÕES

Este capítulo apresentou a análise do comportamento médio do vetor de coeficientes para o algoritmo LMS filtrado. O novo modelo não utiliza a teoria da independência, levando em consideração a correlação entre diferentes vetores do sinal de referência. Nenhuma hipótese é feita com relação à velocidade de adaptação ou ao grau de correlação das amostras do sinal de entrada.

O comportamento do valor esperado do vetor de coeficientes é determinado por uma equação de diferenças recursiva como função da estatística dos sinais de entrada. A partir dos resultados de simulações, pode-se observar que o modelo proposto prediz o comportamento médio dos coeficientes do filtro adaptativo, quando o algoritmo LMS é utilizado para atualizar os coeficientes. Também constatou-se que o modelo é válido para sinais brancos ou coloridos. Em relação às análises disponíveis na literatura, mostrou-se que para obter um modelo representativo para o comportamento do valor esperado dos

coeficientes do filtro adaptativo, a teoria da independência não pode ser aplicada, e as correlações existentes entre os vetores de entrada devem ser levadas em consideração. Desta forma, o modelo proposto, que considera as correlações existentes entre do vetor de entrada, consegue prever perfeitamente o comportamento do valor médio dos coeficientes. Também é obtida uma expressão para estimar o erro médio quadrático em excesso assumindo uma taxa de aprendizagem pequena. Com esta ferramenta podemos prever o erro em excesso que o algoritmo pode atingir para uma determinada aplicação.

Para desenvolver o modelo proposto é realizada a suposição (II)-(seção 2.4), com o objetivo de facilitar o tratamento matemático para obter o modelo do momento de primeira ordem. Esta suposição é verificada experimentalmente, sendo apresentados diferentes resultados de simulações que confirmam tal suposição. É verificado que para valores da taxa de aprendizagem perto do valor limite a suposição começa a perder validade. No entanto, raramente numa aplicação prática, o algoritmo LMS filtrado é utilizado com valores tão altos para a taxa de aprendizagem. Mas, como pode ser observado nas figuras correspondentes, a predição do comportamento do valor médio dos coeficientes continua sendo de boa qualidade.

No capítulo seguinte é analisado o momento de segunda ordem (matriz de covariância). Primeiramente, será analisado o caso em que os filtros correspondentes ao caminho secundário e sua estimação possuem apenas um coeficiente, ou seja, representam um determinado número de atrasos em amostras; que é o caso do algoritmo DLMS (*Delayed-LMS*). Através desta análise, que apresenta uma matemática mais simples, são estabelecidos os conceitos necessários para se desenvolver o modelo de cálculo do momento de segunda ordem do caso geral em que os sistemas do caminho secundário e sua estimação realizam operações de filtragem.

Momento de Segunda Ordem

3

INTRODUÇÃO

Neste capítulo é apresentada uma análise correspondente ao momento de segunda ordem (matriz de correlação) dos coeficientes do filtro adaptativo usando o algoritmo LMS filtrado. Novamente, aqui não é empregada a teoria da independência para a obtenção da expressão recursiva da matriz de correlação. As expressões obtidas são válidas para sinais de entrada gaussianos, podendo ser brancos ou coloridos. Como resultado da análise é obtida uma equação recursiva para o comportamento da matriz de correlação do vetor de erro dos coeficientes do filtro adaptativo. Esta equação permite a determinação de uma expressão analítica para o comportamento do erro médio quadrático.

Com o objetivo de facilitar a exposição do desenvolvimento matemático necessário para a obtenção do novo modelo, vamos abordar primeiramente o caso em que as respostas ao impulso do caminho secundário e sua estimação possuem apenas um coeficiente. Este caso particular corresponde ao algoritmo conhecido na literatura como *Delayed-LMS*, abreviado aqui pela sigla *DLMS*. Como as expressões matemáticas são consideravelmente mais simples para esse caso, pode-se realizar um estudo mais completo do que no caso geral. Uma vez obtidas as expressões finais desejadas, as mesmas são estendidas para o caso geral, que é nosso principal objetivo. Na literatura concernente ao algoritmo *DLMS* existe a afirmação, não demonstrada, de que o algoritmo não converge se a estimação do atraso não for perfeita. Neste capítulo poderemos constatar, de forma teórica, quando a entrada for um ruído branco, porque o algoritmo não converge [26] com $\hat{D} \neq D$. No entanto, se a entrada for correlacionada ou ruído colorido, o algoritmo apresenta convergência para um adequado valor da taxa de convergência. Tais afirmações poderão ser verificadas na Seção 3.1.8. É importante ressaltar que estes fatos são possíveis de serem colocados em evidência porque não estamos utilizando a teoria da independência.

3.1. ALGORITMO DLMS

Nesta seção apresentamos a análise do algoritmo DLMS. O diagrama em blocos do sistema adaptativo é apresentado na Fig. 3.1. Note que este sistema corresponde ao algoritmo LMS filtrado no caso em que o caminho secundário é representado por apenas um coeficiente. \mathbf{W}^o é o vetor de amostras da resposta ao impulso do sistema desconhecido. $\mathbf{W}(n)$ é o vetor de coeficientes do filtro adaptativo. D e \hat{D} representam os atrasos, em amostras, que sofre o sinal que entra em cada um desses blocos.

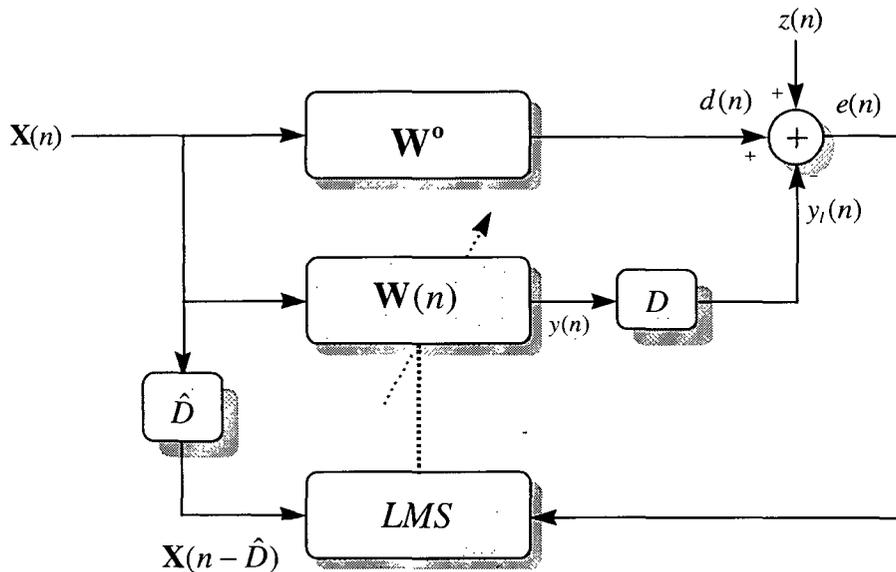


Fig. 3.1. Diagrama em blocos do algoritmo DLMS.

3.1.1. Sinal de erro:

A partir da Fig. 3.1, o sinal de erro é dado por:

$$\begin{aligned}
 e(n) &= d(n) - y_l(n) + z(n) = d(n) - y(n - D) + z(n) \\
 &= d(n) - \mathbf{X}^T(n - D)\mathbf{W}(n - D) + z(n).
 \end{aligned}
 \tag{3.1}$$

Note que agora a saída do caminho secundário corresponde à saída do filtro adaptativo atrasada em D amostras. Esta expressão pode ser derivada de (2.7) fazendo $i = D$, $s_D = 1$ e $s_i = 0$ para $i \neq D$.

3.1.2. Equação de atualização dos pesos:

A equação de atualização dos coeficientes do filtro adaptativo para o caso analisado é obtida a partir de (2.8), sendo dada por:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu e(n) \mathbf{X}(n - \hat{D}). \quad (3.2)$$

Substituindo (3.1) em (3.2), temos:

$$\begin{aligned} \mathbf{W}(n+1) = \mathbf{W}(n) + \mu & \left(d(n) \mathbf{X}(n - \hat{D}) - \mathbf{X}(n - \hat{D}) \mathbf{X}^T(n - D) \mathbf{W}(n - D) \right. \\ & \left. + z(n) \mathbf{X}(n - \hat{D}) \right). \end{aligned} \quad (3.3)$$

Utilizando as mesmas considerações (II.i e II.ii, p. 28) aplicadas para obter (2.37), o valor esperado de (3.3) é dado por:

$$\begin{aligned} E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] + \mu & \left(E[d(n) \mathbf{X}(n - \hat{D})] - E[\mathbf{X}(n - \hat{D}) \mathbf{X}^T(n - D)] E[\mathbf{W}(n - D)] \right. \\ & \left. + E[z(n) \mathbf{X}(n - \hat{D})] \right), \end{aligned} \quad (3.4)$$

ou, definindo $\mathbf{R}_{D-\hat{D}} = E[\mathbf{X}(n - \hat{D}) \mathbf{X}^T(n - D)]$ e $\mathbf{P}_{\hat{D}} = E[d(n) \mathbf{X}(n - \hat{D})]$, podemos rescrever (3.4) como:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] + \mu \left(\mathbf{P}_{\hat{D}} - \mathbf{R}_{D-\hat{D}} E[\mathbf{W}(n - D)] \right) \quad (3.5)$$

3.1.3. Vetor de coeficientes ótimo

O vetor de coeficientes ótimo pode ser obtido de (2.16). Para este caso, onde apenas um coeficiente do filtro S é diferente de zero, $\tilde{\mathbf{R}}_{ss}$ se transforma em \mathbf{R}_0 e $\tilde{\mathbf{P}}_s$, em \mathbf{P}_D . Então o vetor ótimo é dado pela expressão:

$$\mathbf{R}_0 \mathbf{W}_{ot} = \mathbf{P}_D \quad (3.6)$$

com $\mathbf{P}_D = E[d(n)\mathbf{X}(n-D)]$.

3.1.4. Erro médio quadrático $E[e^2(n)]$

Elevando $e(n)$ em (3.1) ao quadrado, temos:

$$e^2(n) = d^2(n) - 2d(n)\mathbf{X}^T(n-D)\mathbf{W}(n-D) + \mathbf{W}^T(n-D)\mathbf{X}(n-D)\mathbf{X}^T(n-D)\mathbf{W}(n-D) + z^2(n).$$

Para determinarmos o valor esperado de $e^2(n)$, fazemos as seguintes considerações simplificadoras:

$$\text{III.i) } E[d(n)\mathbf{X}(n-i)\mathbf{W}(n-i)] \cong E[d(n)\mathbf{X}(n-i)]E[\mathbf{W}(n-i)]$$

$$\text{III.ii) } E[\mathbf{X}^T(n-i)\mathbf{W}(n-i)\mathbf{X}^T(n-j)\mathbf{W}(n-j)] \cong \text{tr}\{E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{W}(n-j)\mathbf{W}^T(n-i)]\}$$

Note que (III.i) e (III.ii) são uma extensão da aproximação (II) empregada na Seção 2.4 para o cálculo do momento de primeira ordem. Basicamente, através dessas suposições, estamos atribuindo maior importância para a correlação entre vetores do sinal de entrada do que para a correlação existente entre o sinal de entrada e o vetor de coeficientes. Usando estas aproximações temos:

$$E[e^2(n)] = E[d^2(n)] - 2E[d(n)\mathbf{X}^T(n-D)]E[\mathbf{W}(n-D)] + E[\mathbf{W}^T(n-D)\mathbf{X}(n-D)\mathbf{X}^T(n-D)\mathbf{W}(n-D)] + E[z^2(n)], \quad (3.7)$$

ou,

$$J(n) = E[e^2(n)] = E[d^2(n)] - 2E[d(n)\mathbf{X}^T(n-D)]E[\mathbf{W}(n-D)] + \text{tr}\{\mathbf{R}_0 E[\mathbf{W}(n-D)\mathbf{W}^T(n-D)]\} + \sigma_z^2, \quad (3.8)$$

onde $\mathbf{R}_0 = E[\mathbf{X}(n-D)\mathbf{X}^T(n-D)]$, de acordo com a definição $\mathbf{R}_{i-j} = E[\mathbf{X}(n-j)\mathbf{X}^T(n-i)]$.

3.1.5. Valor em regime permanente do vetor de coeficientes

Dado que a expressão (3.6) não leva em consideração a possibilidade de um erro na estimação do atraso D , é importante obter uma expressão para o valor em regime permanente que considere essa situação. Calculando $\lim_{n \rightarrow \infty} E[\mathbf{W}(n)] = \mathbf{W}_\infty$ na Eq. (3.5), e resolvendo a equação para \mathbf{W}_∞ , podemos obter uma expressão para o regime permanente dos coeficientes do filtro, que é expressa a seguir:

$$\mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty = \mathbf{P}_{\hat{D}} \quad (3.9)$$

Esta relação vai possibilitar a obtenção de expressões mais simples, tanto para o momento de primeira ordem (tal como é o caso da expressão (2.50)) quanto para o momento de segunda ordem.

Da relação entre \mathbf{W}_{ot} e \mathbf{W}_∞ , podemos fazer os seguintes comentários. O vetor \mathbf{W}_{ot} corresponde ao mínimo da superfície de erro (EMQ) e sobre ele não recai nenhum efeito do algoritmo ou da estimação de D dada por \hat{D} . No caso de $\mathbf{W}_\infty = \lim_{n \rightarrow \infty} E[\mathbf{W}(n)]$, o mesmo acarreta dois efeitos; o efeito da estimação incorreta \hat{D} , e alguma polarização que pode ocorrer devido ao algoritmo. Analisando (3.9) quando $\hat{D} = D$ obtemos a expressão (3.6). Isto significa que $\lim_{n \rightarrow \infty} E[\mathbf{W}(n)]$ é uma estimativa não polarizada de \mathbf{W}_{ot} quando $\hat{D} = D$.

3.1.6. Vetor erro nos coeficientes em relação a \mathbf{W}_∞

Definindo o vetor erro nos coeficientes como sendo $\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}_\infty$ e utilizando a Eq. (3.3), temos:

$$\mathbf{V}(n) = \mathbf{V}(n-1) + \mu \left(d(n-1)\mathbf{X}(n-\hat{D}-1) - \mathbf{X}(n-\hat{D}-1)\mathbf{X}^T(n-D-1)\mathbf{V}(n-D-1) - \mathbf{X}(n-\hat{D}-1)\mathbf{X}^T(n-D-1)\mathbf{W}_\infty + z(n-1)\mathbf{X}(n-\hat{D}-1) \right), \quad (3.10)$$

a expressão (3.6) pode ser considerada como um caso particular de (3.9) quando $\hat{D} = D$, da mesma forma que em (3.4), o valor esperado de (3.10), é dado por:

$$E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] - \mu \mathbf{R}_{D-\hat{D}} E[\mathbf{V}(n-D)] - \mu \mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty + \mu \mathbf{P}_{\hat{D}}. \quad (3.11)$$

Com a ajuda da relação (3.9), (3.11) pode ser simplificada da seguinte forma:

$$E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] - \mu \mathbf{R}_{D-\hat{D}} E[\mathbf{V}(n-D)]. \quad (3.12)$$

A equação de diferenças (3.12) descreve o comportamento do valor médio do vetor de erro nos coeficientes do filtro adaptativo, usando o algoritmo DLMS. Através desta equação podem-se tirar conclusões interessantes em relação à convergência do algoritmo, algumas das quais são mencionadas na literatura, mas não demonstradas. Tal é o assunto da próxima seção.

3.1.7. Estudo do comportamento médio dos coeficientes quando $\hat{D} \neq D$

É bem conhecida da literatura a afirmação que na condição $\hat{D} \neq D$, o algoritmo DLMS não converge [26]. Através da presente análise, e principalmente pela não utilização da teoria da independência, estamos em condições de demonstrar que esta afirmação é válida unicamente para ruído branco gaussiano. Para uma entrada correlacionada gaussiana (ruído colorido gaussiano), tal afirmação não é verdadeira. Para ilustrar estas afirmações, faremos uso de dois exemplos. Suponha que o filtro adaptativo e a planta possuem dois coeficientes, e seja $\hat{D} = 0$ e $D = 1$. Então, para ruído branco de entrada, a

partir da Eq. (3.5) e utilizando $d(n) = \mathbf{X}^T(n)\mathbf{W}^o$, temos que $\mathbf{R}_{D-\hat{D}}$ e $\mathbf{P}_{\hat{D}}$ são dados pelas seguintes expressões:

$$\mathbf{R}_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ e } \mathbf{P}_0 = E[\mathbf{X}(n)\mathbf{X}^T(n)]\mathbf{W}^o = \mathbf{R}_0\mathbf{W}^o = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_0^o \\ w_1^o \end{pmatrix} = \begin{pmatrix} w_0^o \\ w_1^o \end{pmatrix}.$$

Substituindo em (3.5), temos:

$$E\left[\begin{pmatrix} w_0^{n+1} \\ w_1^{n+1} \end{pmatrix}\right] = E\left[\begin{pmatrix} w_0^n \\ w_1^n \end{pmatrix}\right] + \mu \left(\begin{pmatrix} w_0^o \\ w_1^o \end{pmatrix} - \mathbf{R}_1 E\left[\begin{pmatrix} w_0^{n-1} \\ w_1^{n-1} \end{pmatrix}\right] \right) = E\left[\begin{pmatrix} w_0^n \\ w_1^n \end{pmatrix}\right] + \mu \left(\begin{pmatrix} w_0^o \\ w_1^o \end{pmatrix} - E\left[\begin{pmatrix} 0 \\ w_0^{n-1} \end{pmatrix}\right] \right)$$

Dessa última equação vemos que o primeiro componente de $\mathbf{W}(n)$ é permanentemente adicionado com a quantidade μw_0^o , portanto, a magnitude desse componente é monotonicamente crescente. Por outro lado, o segundo componente de $\mathbf{W}(n)$ é modificado pelo primeiro componente, que é indefinidamente crescente, resultando assim na não convergência do filtro adaptativo, qualquer que seja o valor de μ . De (3.12) verifica-se a não convergência sempre que uma linha de $\mathbf{R}_{D-\hat{D}}$ é nula e $\mathbf{V}(0) \neq 0$.

Consideremos agora uma entrada correlacionada, também para o caso de ordem dois. Para este caso as matrizes $\mathbf{R}_{D-\hat{D}}$ e $\mathbf{P}_{\hat{D}}$ adotam a seguinte forma:

$$\mathbf{R}_1 = \begin{pmatrix} r_1 & r_2 \\ r_0 & r_1 \end{pmatrix} \text{ e } \mathbf{P}_0 = E[\mathbf{X}(n)\mathbf{X}^T(n)]\mathbf{W}^o = \mathbf{R}_0\mathbf{W}^o = \begin{pmatrix} r_0 & r_1 \\ r_1 & r_0 \end{pmatrix} \mathbf{W}^o,$$

onde, em geral, $r_i, i = 0, 1, \dots$ são não nulos. Substituindo esses valores em (3.5), é fácil observar que a equação resultante não apresenta os problemas do caso com ruído branco. Neste caso, pela aplicação de (3.9), pode-se obter o valor do vetor de coeficientes em regime permanente. Na Seção 3.2, apresentando resultados de simulações, são mostradas as curvas obtidas para estes exemplos.

3.1.8. Erro médio quadrático em função de $\mathbf{V}(n)$

Utilizando a definição de $\mathbf{V}(n)$ (Seção 3.1.6) e a Eq. (3.7), pode-se escrever uma expressão para o erro médio quadrático como sendo uma função do vetor de erro nos coeficientes, que é dado por:

$$\begin{aligned}
 J(n) = & E[d^2(n)] - 2\mathbf{P}_D^T E[\mathbf{V}(n-D)] - 2\mathbf{P}_D^T \mathbf{W}_\infty \\
 & + \text{tr}\{\mathbf{R}_0 \mathbf{W}_\infty \mathbf{W}_\infty^T\} + \text{tr}\{\mathbf{R}_0 E[\mathbf{V}(n-D)] \mathbf{W}_\infty^T\} \\
 & + \text{tr}\{\mathbf{R}_0 \mathbf{W}_\infty E[\mathbf{V}^T(n-D)]\} \\
 & + \text{tr}\{\mathbf{R}_0 E[\mathbf{V}(n-D) \mathbf{V}^T(n-D)]\} + \sigma_z^2,
 \end{aligned} \tag{3.13}$$

ou, pela utilização das propriedades do traço de uma matriz, temos:

$$\begin{aligned}
 J(n) = & E[d^2(n)] - 2\mathbf{P}_D^T E[\mathbf{V}(n-D)] - 2\mathbf{P}_D^T \mathbf{W}_\infty \\
 & + \mathbf{W}_\infty^T \mathbf{R}_0 \mathbf{W}_\infty + \mathbf{W}_\infty^T \mathbf{R}_0 E[\mathbf{V}(n-D)] \\
 & + E[\mathbf{V}^T(n-D)] \mathbf{R}_0 \mathbf{W}_\infty \\
 & + \text{tr}\{\mathbf{R}_0 E[\mathbf{V}(n-D) \mathbf{V}^T(n-D)]\} + \sigma_z^2.
 \end{aligned} \tag{3.14}$$

Como pode-se observar, o erro médio quadrático é função de $E[\mathbf{V}(n)]$, já determinado em (3.12), e da matriz de correlação $E[\mathbf{V}(n) \mathbf{V}^T(n)]$, objeto de estudo da próxima seção.

3.1.9. Cálculo de $E[\mathbf{V}(n-D) \mathbf{V}^T(n-D)]$

Partindo da expressão do vetor $\mathbf{V}(n)$, uma expressão para $\mathbf{V}(n-D)$ é obtida da seguinte forma. Seja a expressão (3.10) repetida abaixo:

$$\begin{aligned}
 \mathbf{V}(n) = & \mathbf{V}(n-1) + \mu \left(d(n-1) \mathbf{X}(n-\hat{D}-1) - \mathbf{X}(n-\hat{D}-1) \mathbf{X}^T(n-D-1) \mathbf{V}(n-D-1) \right. \\
 & \left. - \mathbf{X}(n-\hat{D}-1) \mathbf{X}^T(n-D-1) \mathbf{W}_\infty + z(n-1) \mathbf{X}(n-\hat{D}-1) \right).
 \end{aligned}$$

Se substituirmos a variável n por $n-D$, a expressão para $\mathbf{V}(n-D)$ é dada por:

$$\begin{aligned}
 \mathbf{V}(n-D) = & \mathbf{V}(n-D-1) + \mu \left(d(n-D-1) \mathbf{X}(n-\hat{D}-D-1) \right. \\
 & - \mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{V}(n-2D-1) \\
 & - \mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{W}_\infty \\
 & \left. + z(n-D-1) \mathbf{X}(n-\hat{D}-D-1) \right). \tag{3.15}
 \end{aligned}$$

efetuando o produto $\mathbf{V}(n-D)\mathbf{V}^T(n-D)$ e tomando o valor esperado utilizando (III.i) e (III.ii), temos:

$$\begin{aligned}
 E[\mathbf{V}(n-D)\mathbf{V}^T(n-D)] = & E[\mathbf{V}(n-D-1)\mathbf{V}^T(n-D-1)] \\
 & + \mu E[\mathbf{V}(n-D-1)] E[d(n-D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & - \mu E[\mathbf{V}(n-D-1)\mathbf{V}^T(n-2D-1)] \mathbf{R}_{\hat{D}-D} \\
 & - \mu E[\mathbf{V}(n-D-1)] \mathbf{W}_\infty^T \mathbf{R}_{\hat{D}-D} \\
 & + \mu E[d(n-D-1)\mathbf{X}(n-\hat{D}-D-1)] E[\mathbf{V}^T(n-D-1)] \\
 & + \mu^2 E[d^2(n-D-1)\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & - \mu^2 E[d(n-D-1)\mathbf{X}(n-\hat{D}-D-1)\mathbf{V}^T(n-2D-1)\mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & - \mu^2 E[d(n-D-1)\mathbf{X}(n-\hat{D}-D-1)\mathbf{W}_\infty^T \mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & - \mu \mathbf{R}_{D-\hat{D}} E[\mathbf{V}(n-2D-1)\mathbf{V}^T(n-D-1)] \\
 & - \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{V}(n-2D-1)d(n-D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & + \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{V}(n-2D-1)\mathbf{V}^T(n-2D-1)\mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & + \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{V}(n-2D-1)\mathbf{W}_\infty^T \mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & - \mu \mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty E[\mathbf{V}^T(n-D-1)] \\
 & - \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{W}_\infty d(n-D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & + \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{W}_\infty \mathbf{V}^T(n-2D-1)\mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & + \mu^2 E[\mathbf{X}(n-\hat{D}-D-1)\mathbf{X}^T(n-2D-1)\mathbf{W}_\infty \mathbf{W}_\infty^T \mathbf{X}(n-2D-1)\mathbf{X}^T(n-\hat{D}-D-1)] \\
 & + \mu^2 \mathbf{R}_0 \sigma_z^2. \tag{3.16}
 \end{aligned}$$

Com o objetivo de simplificar a forma de (3.16), vamos introduzir a seguinte notação: $\mathbf{K}(n-i, n-j) = E[\mathbf{V}(n-i)\mathbf{V}^T(n-j)]$, que também é utilizada em [6]. Com isso, a Eq. (3.16) fica da seguinte forma:

$$\begin{aligned}
 \mathbf{K}(n-D, n-D) &= \mathbf{K}(n-D-1, n-D-1) \\
 &+ \mu E[\mathbf{V}(n-D-1)] \mathbf{P}_{\hat{D}}^T \\
 &- \mu \mathbf{K}(n-D-1, n-2D-1) \mathbf{R}_{\hat{D}-D} \\
 &- \mu E[\mathbf{V}(n-D-1)] \mathbf{W}_x^T \mathbf{R}_{\hat{D}-D} \\
 &+ \mu \mathbf{P}_{\hat{D}} E[\mathbf{V}^T(n-D-1)] \\
 &+ \mu^2 E[d^2(n-D-1) \mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &- \mu^2 E[d(n-D-1) \mathbf{X}(n-\hat{D}-D-1) \mathbf{V}^T(n-2D-1) \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &- \mu^2 E[d(n-D-1) \mathbf{X}(n-\hat{D}-D-1) \mathbf{W}_x^T \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &- \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}(n-2D-1, n-D-1) \\
 &- \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{V}(n-2D-1) d(n-D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &+ \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{V}(n-2D-1) \mathbf{V}^T(n-2D-1) \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &+ \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{V}(n-2D-1) \mathbf{W}_x^T \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &- \mu \mathbf{R}_{D-\hat{D}} \mathbf{W}_x E[\mathbf{V}^T(n-D-1)] \\
 &- \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{W}_x d(n-D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &+ \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{W}_x \mathbf{V}^T(n-2D-1) \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &+ \mu^2 E[\mathbf{X}(n-\hat{D}-D-1) \mathbf{X}^T(n-2D-1) \mathbf{W}_x \mathbf{W}_x^T \mathbf{X}(n-2D-1) \mathbf{X}^T(n-\hat{D}-D-1)] \\
 &+ \mu^2 \mathbf{R}_0 \sigma_z^2.
 \end{aligned} \tag{3.17}$$

Da mesma forma como (3.11) foi simplificada utilizando (3.9), pode ser aplicado à Eq. (3.17) o mesmo procedimento a fim de obter-se uma forma mais compacta através de (3.9). O processo de simplificação é realizado identificando termos que estejam relacionados através de (3.9). Para isto procedemos da seguinte forma, pós-multiplicando ambos lados de (3.9) por $E[\mathbf{V}^T(n-D-1)]$, temos:

$$\mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty E[\mathbf{V}^T(n-D-1)] = \mathbf{P}_{\hat{D}} E[\mathbf{V}^T(n-D-1)],$$

o que significa que o quinto e o décimo terceiro termos de (3.17) podem ser cancelados por serem iguais. O mesmo acontece com os segundo e quarto termos. Para determinar os valores esperados contendo termos de quarta ordem (sexto, sétimo, oitavo, décimo, décimo primeiro, décimo segundo, décimo quarto, décimo quinto e décimo sexto) é utilizada a equação para momentos de quarta ordem de um processo gaussiano de média zero [6-7, 31]. O desenvolvimento é mostrado detalhadamente no Apêndice C. Uma vez calculados estes valores esperados e, utilizando novamente (3.9), é possível cancelar os

seguintes termos: sexto, sétimo, oitavo, décimo, décimo segundo, décimo quarto, décimo quinto e décimo sexto. Eliminados esses termos, a Eq. (3.17) pode ser rescrita como:

$$\begin{aligned}
 \mathbf{K}(n-D, n-D) &= \mathbf{K}(n-D-1, n-D-1) \\
 &- \mu \mathbf{K}(n-D-1, n-2D-1) \mathbf{R}_{\hat{D}-D} \\
 &- \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}(n-2D-1, n-D-1) \\
 &+ \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}(n-2D-1, n-2D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}(n-2D-1, n-2D-1) \} \right) \\
 &+ \mu^2 \mathbf{R}_0 \sigma_z^2
 \end{aligned} \tag{3.18}$$

Com o objetivo de simplificar ainda mais a expressão da matriz de correlação, adotaremos a seguinte notação $\mathbf{K}_{ij}(n) = \mathbf{K}(n-i, n-j)$. Esta notação não foi diretamente aplicada em (3.16) para não prejudicar a clareza do texto. Note que com a nova notação, atrasos são expressos através da modificação dos índices i e j de $\mathbf{K}_{ij}(n)$. Isto possibilita a redução do modelo a um número finito de equações recursivas. Desta forma, (3.18) é rescrita como:

$$\begin{aligned}
 \mathbf{K}_{D,D}(n) &= \mathbf{K}_{D,D}(n-1) - \mu \mathbf{K}_{D,2D}(n-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{2D,D}(n-1) \\
 &+ \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D}(n-D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}_{D,D}(n-D-1) \} \right) \\
 &+ \mu^2 \mathbf{R}_0 \sigma_z^2.
 \end{aligned} \tag{3.19}$$

A Eq. (3.19) ainda pode ser trabalhada a fim de que fique mais compacta. O primeiro passo é utilizar a relação entre atrasos e os índices de tempo das matrizes de correlação $\mathbf{K}_{ij}(n)$. Considere a seguinte igualdade:

$$\mathbf{K}_{D,D}(n) = E[\mathbf{V}(n-D) \mathbf{V}^T(n-D)] = \mathbf{K}_{0,0}(n-D). \tag{3.20}$$

Isto significa que $\mathbf{K}_{D,D}(n)$ e $\mathbf{K}_{0,0}(n-D)$ representam a mesma variável. Assim, podemos reduzir pelo menos um dos índices de todas as variáveis de (3.19) para o valor zero através da igualdade $\mathbf{K}_{i,j}(n) = \mathbf{K}_{i-\Delta, j-\Delta}(n-\Delta)$, o que nos permite escrever:

$$\begin{aligned}
 \mathbf{K}_{0,0}(n-D) &= \mathbf{K}_{0,0}(n-D-1) - \mu \mathbf{K}_{0,D}(n-D-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,0}(n-D-1) \\
 &\quad + \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{0,0}(n-2D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}_{0,0}(n-2D-1) \} \right) \\
 &\quad + \mu^2 \mathbf{R}_0 \sigma_z^2.
 \end{aligned} \quad (3.21)$$

Fazendo $n = n + D$ em (3.21), temos:

$$\begin{aligned}
 \mathbf{K}_{0,0}(n) &= \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,D}(n-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,0}(n-1) \\
 &\quad + \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{0,0}(n-D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}_{0,0}(n-D-1) \} \right) \\
 &\quad + \mu^2 \mathbf{R}_0 \sigma_z^2.
 \end{aligned} \quad (3.22)$$

Note que no caso da variável $\mathbf{K}_{2D,D}(n)$ em (3.19) apenas um dos índices é reduzido a zero. A recursão (3.22) é função de $\mathbf{K}_{0,0}(n)$, $\mathbf{K}_{0,D}(n)$ e $\mathbf{K}_{D,0}(n)$. Logo, precisamos determinar expressões explícitas para $\mathbf{K}_{0,D}(n)$ e $\mathbf{K}_{D,0}(n)$. Pela sua definição sabemos que essas variáveis estão relacionadas entre si por: $\mathbf{K}_{D,0}(n) = \mathbf{K}_{0,D}^T(n)$. Assim, precisamos determinar uma expressão para calcular apenas uma dessas duas variáveis. Escolhendo arbitrariamente $\mathbf{K}_{0,D}(n)$, pós-multiplicamos ambos os lados de (3.10) por $\mathbf{V}^T(n-D)$ e tomamos o valor esperado, obtendo:

$$\begin{aligned}
 E[\mathbf{V}(n) \mathbf{V}^T(n-D)] &= E[\mathbf{V}(n-1) \mathbf{V}^T(n-D)] + \mu (\mathbf{P}_{\hat{D}} E[\mathbf{V}^T(n-D)] \\
 &\quad - \mathbf{R}_{D-\hat{D}} E[\mathbf{V}(n-D-1) \mathbf{V}^T(n-D)] \\
 &\quad - \mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty E[\mathbf{V}^T(n-D)]),
 \end{aligned} \quad (3.23)$$

usando a relação (3.9) e empregando a notação adotada, (3.23) é expressa como:

$$\mathbf{K}_{0,D}(n) = \mathbf{K}_{0,D-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D-1}(n-1). \quad (3.24)$$

Neste ponto, estamos em condições de caracterizar a matriz de correlação do vetor de erro nos coeficientes através do seguinte sistema de equações recursivas:

$$\begin{aligned} \mathbf{K}_{0,0}(n) &= \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,D}(n-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,0}(n-1) \\ &\quad + \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{0,0}(n-D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}_{0,0}(n-D-1) \} \right) \\ &\quad + \mu^2 \mathbf{R}_0 \sigma_z^2, \end{aligned} \quad (3.25a)$$

$$\mathbf{K}_{0,D}(n) = \mathbf{K}_{0,D-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D-1}(n-1), \quad (3.25b)$$

$$\mathbf{K}_{0,D}(n) = \mathbf{K}_{D,0}^T(n). \quad (3.25c)$$

Antes de continuarmos com o manuseio da Eq. (3.25), é útil ilustrar com um exemplo sua utilização. Suponhamos que $D = 2$ e seja \hat{D} qualquer valor arbitrário. Então, temos:

Eq. (3.25a):

$$\begin{aligned} \mathbf{K}_{0,0}(n) &= \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,2}(n-1) \mathbf{R}_{\hat{D}-2} - \mu \mathbf{R}_{2-\hat{D}} \mathbf{K}_{2,0}(n-1) \\ &\quad + \mu^2 \left(2 \mathbf{R}_{2-\hat{D}} \mathbf{K}_{0,0}(n-2-1) \mathbf{R}_{\hat{D}-2} + \mathbf{R}_0 \text{tr} \{ \mathbf{R}_0 \mathbf{K}_{0,0}(n-2-1) \} \right) \\ &\quad + \mu^2 \mathbf{R}_0 \sigma_z^2, \end{aligned} \quad (3.26a)$$

vemos que $\mathbf{K}_{0,0}(n)$ depende de suas versões atrasadas $\mathbf{K}_{0,0}(n-1)$ e $\mathbf{K}_{0,0}(n-3)$, assim como de $\mathbf{K}_{0,2}(n-1)$ e $\mathbf{K}_{2,0}(n-1)$. Para calcular esses últimos termos, recorreremos novamente para à Eq. (3.25b), obtendo:

$$\begin{aligned} \mathbf{K}_{0,2}(n) &= \mathbf{K}_{0,2-1}(n-1) - \mu \mathbf{R}_{2-\hat{D}} \mathbf{K}_{2,2-1}(n-1) \\ &= \mathbf{K}_{0,1}(n-1) - \mu \mathbf{R}_{2-\hat{D}} \mathbf{K}_{2,1}(n-1). \end{aligned}$$

Para determinarmos $\mathbf{K}_{0,2}(n)$ devemos conhecer $\mathbf{K}_{0,1}(n-1)$ e $\mathbf{K}_{2,1}(n-1)$. Como:

$\mathbf{K}_{2,1}(n-1) = \mathbf{K}_{1,0}(n-2)$, a expressão anterior é rescrita como:

$$\mathbf{K}_{0,2}(n) = \mathbf{K}_{0,1}(n-1) - \mu \mathbf{R}_{2-\hat{D}} \mathbf{K}_{1,0}(n-2). \quad (3.26b)$$

Aplicando novamente (3.25) temos que $\mathbf{K}_{0,1}(n)$ é dado por:

$$\mathbf{K}_{0,1}(n) = \mathbf{K}_{0,0}(n-1) - \mu \mathbf{R}_{2-\hat{D}} \mathbf{K}_{1,0}(n-1), \quad (3.26c)$$

finalmente, de (3.25c) temos que: $\mathbf{K}_{0,1}(n) = \mathbf{K}_{1,0}^T(n)$. Desta forma, fica completa a descrição do cálculo em forma recursiva da matriz de correlação. Em virtude da forma da Eq. (3.25b), originaram-se as Eqs. (3.26) para determinação de $\mathbf{K}_{0,2}(n)$. Em geral, para o cálculo de $\mathbf{K}_{0,D}(n)$ com (3.25b), para um valor qualquer de D , é gerado o seguinte conjunto de equações:

$$\begin{aligned} \mathbf{K}_{0,D}(n) &= \mathbf{K}_{0,D-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D-1}(n-1), \\ \mathbf{K}_{0,D-1}(n) &= \mathbf{K}_{0,D-2}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D-1,D-2}(n-1), \\ &\cdot \\ &\cdot \\ \mathbf{K}_{0,1}(n) &= \mathbf{K}_{0,0}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{1,0}(n-1) \end{aligned}$$

Tal conjunto pode ser representado pela seguinte expressão:

$$\mathbf{K}_{0,D-j}(n) = \mathbf{K}_{0,D-j-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D-j,D-j-1}(n-1) \text{ com } j = 0, \dots, D-1. \quad (3.27)$$

Pela utilização de $\mathbf{K}_{D-j,D-j-1}(n) = \mathbf{K}_{1,0}(n-D+j+1)$, (3.27) pode ser escrita como:

$$\mathbf{K}_{0,D-j}(n) = \mathbf{K}_{0,D-j-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{1,0}(n-D+j) \text{ com } j = 0, \dots, D-1. \quad (3.28)$$

Neste ponto, estamos em condições de escrever o sistema de equações que descreve o comportamento da matriz de correlação do vetor de erro nos coeficientes do filtro adaptativo, para o algoritmo DLMS:

$$\begin{aligned} \mathbf{K}_{0,0}(n) = & \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,D}(n-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,0}(n-1) \\ & + \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{0,0}(n-D-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \left\{ \mathbf{R}_0 \mathbf{K}_{0,0}(n-D-1) \right\} \right) \\ & + \mu^2 \mathbf{R}_0 \sigma_z^2, \end{aligned} \quad (3.29a)$$

$$\mathbf{K}_{0,D-j}(n) = \mathbf{K}_{0,D-j-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{1,0}(n-D+j) \quad \text{com } j = 0, \dots, D-1, \quad (3.29b)$$

$$\mathbf{K}_{r,s}(n) = \mathbf{K}_{s,r}^T(n). \quad (3.29c)$$

É interessante notar que (3.29) também pode ser escrita como um conjunto de equações de diferenças de ordem 1 em várias variáveis. Podemos rescrever (3.29) da seguinte forma:

$$\begin{aligned} \mathbf{K}_{0,0}(n) = & \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,D}(n-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,0}(n-1) \\ & + \mu^2 \left(2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D}(n-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr} \left\{ \mathbf{R}_0 \mathbf{K}_{D,D}(n-1) \right\} \right) \\ & + \mu^2 \mathbf{R}_0 \sigma_z^2, \end{aligned} \quad (3.30a)$$

$$\mathbf{K}_{0,D-j}(n) = \mathbf{K}_{0,D-j-1}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D-j,D-j-1}(n-1) \quad \text{com } j = 0, \dots, D-1, \quad (3.30b)$$

$$\mathbf{K}_{r,s}(n) = \mathbf{K}_{s,r}^T(n), \quad (3.30c)$$

$$\mathbf{K}_{r,s}(n) = \mathbf{K}_{r-\Delta,s-\Delta}(n-\Delta). \quad (3.30d)$$

Esta representação pode ser interessante para formulações matriciais do comportamento do algoritmo. Em especial, tal representação será útil no estudo da estabilidade do algoritmo.

3.1.10. Estudo do caso com estimação perfeita do atraso

Nesta seção vamos considerar o caso de estimação perfeita do atraso, $\hat{D} = D$. Esse caso é importante porque permite estabelecer o comportamento do algoritmo em uma situação ideal, a qual pode ser usada como referência para casos mais práticos. Com esta condição simplificada o tratamento matemático do problema permite chegar a algumas conclusões relativas à estabilidade do algoritmo. Para esse caso o sistema (3.30) (essa representação

foi escolhida arbitrariamente, não afetando o resultado final desta seção) adota a seguinte forma:

$$\begin{aligned} \mathbf{K}_{0,0}(n) = & \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,D}(n-1) \mathbf{R}_0 - \mu \mathbf{R}_0 \mathbf{K}_{D,0}(n-1) \\ & + \mu^2 \left(2 \mathbf{R}_0 \mathbf{K}_{D,D}(n-1) \mathbf{R}_0 + \mathbf{R}_0 \text{tr} \left\{ \mathbf{R}_0 \mathbf{K}_{D,D}(n-1) \right\} \right) \\ & + \mu^2 \mathbf{R}_0 \sigma_z^2, \end{aligned} \quad (3.31a)$$

$$\mathbf{K}_{0,D-j}(n) = \mathbf{K}_{0,D-j-1}(n-1) - \mu \mathbf{R}_0 \mathbf{K}_{D-j,D-j-1}(n-1) \quad \text{com } j = 0, \dots, D-1, \quad (3.31b)$$

$$\mathbf{K}_{r,s}(n) = \mathbf{K}_{s,r}^T(n), \quad (3.31c)$$

$$\mathbf{K}_{r,s}(n) = \mathbf{K}_{r-1,s-1}(n-1). \quad (3.31d)$$

Pela utilização de uma transformação ortogonal, podemos expressar a seguinte relação:

$$\mathbf{Q}^T \mathbf{R}_0 \mathbf{Q} = \mathbf{\ddot{E}},$$

onde \mathbf{Q} é a matriz cujas colunas são os autovetores de \mathbf{R}_0 e $\mathbf{\ddot{E}} = \text{diag}(\lambda_0, \dots, \lambda_{N-1})$ é a matriz dos autovalores de \mathbf{R}_0 . Pré e pós-multiplicando o sistema de equações (3.31) por \mathbf{Q}^T e \mathbf{Q} , respectivamente, e utilizando a notação:

$$\mathbf{Q}^T \mathbf{K}_{i,j}(n) \mathbf{Q} = \mathbf{C}_{i,j}(n)$$

podemos escrever:

$$\begin{aligned} \mathbf{C}_{0,0}(n) = & \mathbf{C}_{0,0}(n-1) - \mu \mathbf{C}_{0,D}(n-1) \Lambda - \mu \Lambda \mathbf{C}_{D,0}(n-1) \\ & + \mu^2 \left(2 \Lambda \mathbf{C}_{D,D}(n-1) \Lambda + \Lambda \text{tr} \left\{ \Lambda \mathbf{C}_{D,D}(n-1) \right\} \right) \\ & + \mu^2 \Lambda \sigma_z^2 \end{aligned} \quad (3.32a)$$

$$\mathbf{C}_{0,D-j}(n) = \mathbf{C}_{0,D-j-1}(n-1) - \mu \Lambda \mathbf{C}_{D-j,D-j-1}(n-1) \quad \text{com } j = 0, \dots, D-1 \quad (3.32b)$$

$$\mathbf{C}_{r,s}(n) = \mathbf{C}_{s,r}^T(n) \quad (3.32c)$$

$$\mathbf{C}_{r,s}(n) = \mathbf{C}_{r-1,s-1}(n-1) \quad (3.32d)$$

Se aplicarmos a mesma transformação à Eq. (3.8), obtemos:

$$J(n) = E[e^2(n)] = E[d^2(n)] - 2E[d(n)\mathbf{X}^T(n-D)]E[\mathbf{W}(n-D)] + \text{tr}\{\Lambda\mathbf{C}_{0,0}(n)\} + \sigma_z^2.$$

Agora vemos que $J(n)$ depende do *traço* ($\Lambda\mathbf{C}_{0,0}(n)$). Isso mostra que o comportamento do algoritmo pode ser estudado a partir dos elementos da diagonal da matriz $\mathbf{C}_{0,0}(n)$. Definindo o vetor $c_{i,j}(n)$ contendo os elementos da diagonal da matriz $\mathbf{C}_{i,j}(n)$, e seja o vetor λ que contém a diagonal de Λ . Temos, a partir de (3.32):

$$c_{0,0}(n) = c_{0,0}(n-1) - \mu c_{0,D}(n-1)\Lambda - \mu\Lambda c_{D,0}(n-1) + \mu^2 \left(2\Lambda c_{D,D}(n-1)\Lambda + \Lambda \text{tr}\{\Lambda c_{D,D}(n-1)\} \right) + \mu^2 \Lambda \sigma_z^2, \quad (3.33a)$$

$$c_{0,D-j}(n) = c_{0,D-j-1}(n-1) - \mu\Lambda c_{D-j,D-j-1}(n-1) \text{ com } j = 0, \dots, D-1, \quad (3.33b)$$

$$c_{r,s}(n) = c_{s,r}(n), \quad (3.33c)$$

$$c_{r,s}(n) = c_{r-1,s-1}(n-1). \quad (3.33d)$$

Utilizando as seguintes propriedades, pode-se simplificar as equações (3.33):

- i) Como $\mathbf{C}_{0,D}(n) = \mathbf{C}_{D,0}^T(n)$, $c_{0,D}(n) = c_{D,0}(n)$,
- ii) $\lambda\lambda^T c_{D,D}(n) = \text{diag}\{\Lambda \text{tr}\{\Lambda c_{D,D}(n)\}\}$.

Com isto, podemos rescrever (3.33) como:

$$c_{0,0}(n) = c_{0,0}(n-1) - 2\mu\Lambda c_{0,D}(n-1) + 2\mu^2 \Lambda^2 c_{D,D}(n-1) + \mu^2 \lambda\lambda^T c_{D,D}(n-1) + \mu^2 \lambda \sigma_z^2, \quad (3.34a)$$

$$c_{0,D-j}(n) = c_{0,D-j-1}(n-1) - \mu\Lambda c_{D-j,D-j-1}(n-1) \text{ com } j = 0, \dots, D-1, \quad (3.34b)$$

$$c_{r,s}(n) = c_{r-1,s-1}(n-1). \quad (3.34c)$$

Com o objetivo do estudo da estabilidade, ainda é possível manipular as Eqs. (3.34) para colocar em evidência o atraso D e, desta forma, calcular o limite superior da taxa de

aprendizagem (μ) que assegura a convergência do sistema de equações (3.34). Para isto, vamos estudar a Eq. (3.34b). Seja um atraso D arbitrário, que produz o seguinte conjunto de equações:

$$j = 0, \quad c_{0,D}(n) = c_{0,D-1}(n-1) - \mu \Lambda c_{D,D-1}(n-1), \quad (3.35a)$$

$$j = 1, \quad c_{0,D-1}(n) = c_{0,D-2}(n-1) - \mu \Lambda c_{D-1,D-2}(n-1), \quad (3.35b)$$

$$\vdots$$

$$j = D-2, \quad c_{0,2}(n) = c_{0,1}(n-1) - \mu \Lambda c_{2,1}(n-1), \quad (3.35c)$$

$$j = D-1, \quad c_{0,1}(n) = c_{0,0}(n-1) - \mu \Lambda c_{1,0}(n-1). \quad (3.35d)$$

Substituindo (3.35d) em (3.35c), temos que (3.35c) pode ser escrita como sendo:

$$j = D-2, \quad c_{0,2}(n) = c_{0,0}(n-2) - \mu \Lambda c_{1,0}(n-2) - \mu \Lambda c_{2,1}(n-1) \quad (3.35c)$$

de (3.34c) sabemos que $c_{2,1}(n-1) = c_{1,0}(n-2)$, então (3.35c) adota a seguinte forma:

$$j = D-2, \quad c_{0,2}(n) = c_{0,0}(n-2) - 2\mu \Lambda c_{1,0}(n-2) \quad (3.35c)$$

Continuando com o processo de substituição, é fácil obter a seguinte expressão equivalente para a (3.35a), que é dada por:

$$j = 0, \quad c_{0,D}(n) = c_{0,0}(n-D) - D\mu \Lambda c_{1,0}(n-D). \quad (3.36)$$

Substituindo (3.36) em (3.34a), temos:

$$c_{0,0}(n) = c_{0,0}(n-1) - 2\mu \Lambda (c_{0,0}(n-D-1) - D\mu \Lambda c_{1,0}(n-D-1))$$

$$+ 2\mu^2 \Lambda^2 c_{D,D}(n-1) \quad (3.37a)$$

$$+ \mu^2 \lambda \lambda^T c_{D,D}(n-1) + \mu^2 \lambda \sigma_z^2,$$

ou,

$$\begin{aligned}
 c_{0,0}(n) = & c_{0,0}(n-1) - 2\mu\Lambda c_{0,0}(n-D-1) + 2\mu^2 D\Lambda^2 c_{1,0}(n-D-1) \\
 & + 2\mu^2 \Lambda^2 c_{D,D}(n-1) \\
 & + \mu^2 \lambda \lambda^T c_{D,D}(n-1) + \mu^2 \lambda \sigma_z^2 .
 \end{aligned} \tag{3.37b}$$

De (3.37b) vemos que agora são necessárias apenas duas equações para caracterizar o comportamento da matriz de correlação do vetor erro nos dos coeficientes: (3.37b) e a expressão associada (3.35d), para calcular a variável $c_{1,0}(n)$. Então, as equações que descrevem o presente caso são:

$ \begin{aligned} c_{0,0}(n) = & c_{0,0}(n-1) - 2\mu\Lambda c_{0,0}(n-D-1) + 2\mu^2 D\Lambda^2 c_{1,0}(n-D-1) \\ & + 2\mu^2 \Lambda^2 c_{D,D}(n-1) \\ & + \mu^2 \lambda \lambda^T c_{D,D}(n-1) + \mu^2 \lambda \sigma_z^2 , \end{aligned} $	(3.37b)
$c_{1,0}(n) = c_{0,0}(n-1) - \mu\Lambda c_{1,0}(n-1) .$	(3.35d)

É interessante expressar (3.37b) e (3.35d) em forma matricial. Para isto utilizaremos um exemplo a fim de ilustrar o procedimento. Seja $D=1$, então de (3.37b) e (3.35d) são obtidas as seguintes expressões:

$$\begin{aligned}
 c_{0,0}(n) = & c_{0,0}(n-1) - 2\mu\Lambda c_{0,0}(n-1-1) + 2\mu^2 \Lambda^2 c_{1,0}(n-1-1) \\
 & + 2\mu^2 \Lambda^2 c_{1,1}(n-1) + \mu^2 \lambda \lambda^T c_{1,1}(n-1) + \mu^2 \lambda \sigma_z^2 , \\
 c_{1,0}(n) = & c_{0,0}(n-1) - \mu\Lambda c_{1,0}(n-1)
 \end{aligned}$$

pela Eq. (3.34c) temos que: $c_{1,0}(n-2) = c_{2,1}(n-1)$ ou $c_{2,1}(n) = c_{1,0}(n-1)$ e $c_{1,1}(n) = c_{0,0}(n-1)$. Desta forma, podemos escrever o seguinte conjunto de equações:

$$\begin{aligned}
 c_{0,0}(n) = & c_{0,0}(n-1) + 2\mu^2 \Lambda^2 c_{2,1}(n-1) \\
 & + (-2\mu\Lambda + 2\mu^2 \Lambda^2 + \mu^2 \lambda \lambda^T) c_{1,1}(n-1) \\
 & + \mu^2 \lambda \sigma_z^2 , \\
 c_{1,0}(n) = & c_{0,0}(n-1) - \mu\Lambda c_{1,0}(n-1) \\
 c_{2,1}(n) = & c_{1,0}(n-1) \\
 c_{1,1}(n) = & c_{0,0}(n-1)
 \end{aligned}$$

ou, pode-se escrever na seguinte forma matricial:

$$\begin{pmatrix} c_{0,0}(n) \\ c_{1,0}(n) \\ c_{1,1}(n) \\ c_{2,1}(n) \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 & -2\mu\Lambda + 2\mu^2\Lambda^2 + \mu^2\lambda\lambda^T & 2\mu^2D\Lambda^2 \\ \mathbf{I} & -\mu\Lambda & 0 & 0 \\ \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 \end{pmatrix} \begin{pmatrix} c_{0,0}(n-1) \\ c_{1,0}(n-1) \\ c_{1,1}(n-1) \\ c_{2,1}(n-1) \end{pmatrix} + \begin{pmatrix} \mu\lambda\sigma_z^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Para um D genérico obtemos a seguinte expressão:

$$\begin{pmatrix} c_{0,0}(n) \\ c_{0,1}(n) \\ \vdots \\ \vdots \\ c_{D-1,D-2}(n) \end{pmatrix} = \begin{pmatrix} \mathbf{I} & 0 & \dots & 0 & 0 & 0 & \dots & -2\mu\Lambda + 2\mu^2\Lambda^2 + \mu^2\lambda\lambda^T & 2\mu^2D\Lambda^2 \\ \mathbf{I} & -\mu\Lambda & \dots & 0 & & & & & 0 \\ \mathbf{I} & 0 & \vdots & 0 & & & & & \vdots \\ \mathbf{I} & 0 & & 0 & & & & & 0 \\ \vdots & & \vdots & & & & & & \vdots \\ 0 & & \dots & \mathbf{I} & 0 & \dots & & & 0 \end{pmatrix} \begin{pmatrix} c_{0,0}(n-1) \\ c_{0,1}(n-1) \\ \vdots \\ \vdots \\ c_{D-1,D-2}(n-1) \end{pmatrix} + \begin{pmatrix} \mu\lambda\sigma_z^2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.38)$$

ou, em forma equivalente,

$$\mathbf{c}_n = \mathbf{A}\mathbf{c}_{n-1} + \mathbf{u} \quad (3.39)$$

A evolução do erro médio quadrático (EMQ) fica completamente determinada usando-se as equações (3.39) e (3.12) em (3.13) ou (3.14). As informações relativas à estabilidade do algoritmo podem ser obtidas a partir da matriz \mathbf{A} , especificamente de seus autovalores. Dado que estes últimos dependem de um parâmetro variável, que é a taxa de aprendizagem μ , devemos obter um limite superior para esta taxa. Através deste limite vamos garantir que os autovalores sejam menores do que a unidade, assegurando assim a convergência de (3.39). Tal é o assunto da próxima seção.

3.1.11. Estudo da estabilidade do algoritmo DLMS

Como já foi mencionado, a estabilidade do algoritmo DLMS pode ser estudada a partir dos autovalores da matriz \mathbf{A} em (3.39). Como estes autovalores são função da taxa de aprendizagem μ , devemos impor os limites para μ de forma que os autovalores sejam menores do que a unidade, que é a condição de estabilidade [23,24]. Para esta análise,

devemos observar a primeira linha de \mathbf{A} . Note, que os elementos de \mathbf{A} são matrizes. Se os elementos dessa primeira linha de matrizes satisfazem a condição:

$$1 - 2\mu\lambda_i + \mu^2 \left(2D\lambda_i^2 + 2\lambda_i^2 + \lambda_i \sum_{j=1}^N \lambda_j \right) < 1 \text{ com } i = 1, \dots, N, \quad (3.40)$$

onde N é a dimensão do vetor de entrada. Então de (3.40) podemos obter o seguinte limite superior para μ :

$$\mu < \frac{2}{2(D+1)\lambda_i + \sum_{j=1}^N \lambda_j}, \quad (3.41)$$

O menor valor deste limite é determinado pelo máximo autovalor λ_{\max} . Então:

$$\mu < \frac{2}{2(D+1)\lambda_{\max} + \sum_{j=1}^N \lambda_j} \quad (3.42)$$

Uma expressão similar a (3.40) é obtida em [21, pag. 172] para o algoritmo LMS clássico. Para o caso de ruído branco gaussiano como sinal de entrada, os autovalores são todos iguais a $\lambda_j = \sigma_x^2, \forall j$. Com isso (3.42) assume a seguinte forma:

$$\mu < \frac{2}{[2(D+1) + N]\sigma_x^2} \quad (3.43)$$

Em [9,29,31] são apresentados limites de estabilidade para o caso em que o atraso se encontra no caminho do erro, e não em cascata com o filtro adaptativo como é o nosso caso. Na Fig. 3.2 são mostradas as curvas correspondentes aos limites dessas referências e da Eq.(3.43). Com o objetivo de comparar resultados, foram utilizadas as mesmas condições que em [29], ou seja, entrada ruído branco e $N=100$.

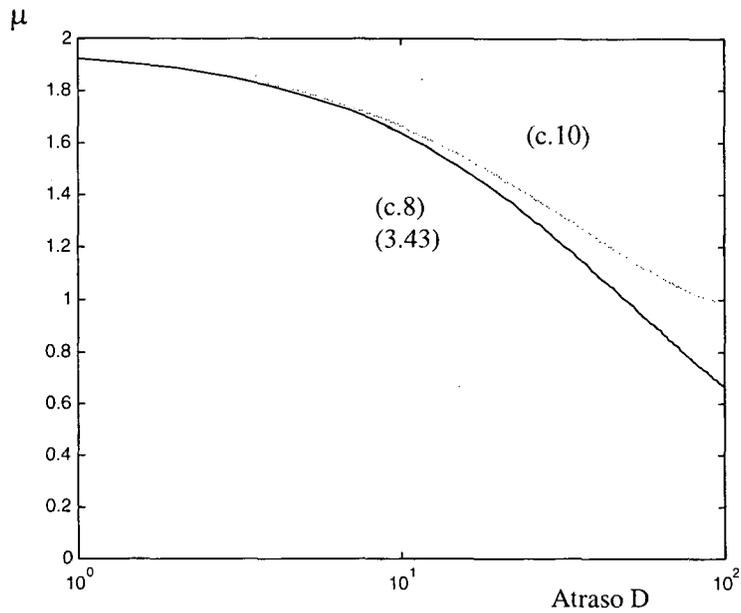


Fig. 3.2. Gráfico das Eqs.(c.8) e (c.10) de [29] e (3.43).

Como pode-se observar da figura, os limites dados pelas Eqs. (c.8) de [29] e (3.43) são iguais. No entanto, as equações (c.8) e (c.10) foram obtidas em [29] através de uma série de aproximações. Em [32] um limite superior é obtido a partir do momento de primeira ordem. Este limite não foi mostrado na Fig. 3.2, uma vez que é muito superior aos aqui apresentados. Este resultado está de acordo com o esperado, uma vez que o limite obtido a partir de momentos de segunda ordem é mais restritivo do que o obtido através do momento de primeira ordem.

3.1.12. Análise do caso geral ($\hat{D} \neq D$)

O caso mais geral $\hat{D} \neq D$, representado pelas Eqs. (3.30), pode ser trabalhado da mesma forma que em (3.35) e (3.37), de modo que o sistema seja reduzido para apenas duas equações. Logo, seguindo o mesmo procedimento, o sistema (3.30) pode ser rescrito da seguinte forma:

$$\begin{aligned}
 \mathbf{K}_{0,0}(n) = & \mathbf{K}_{0,0}(n-1) - \mu \mathbf{K}_{0,0}(n-D-1) \mathbf{R}_{\hat{D}-D} - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{0,0}(n-D-1) \\
 & + \mu^2 D \mathbf{R}_{D-\hat{D}} (\mathbf{K}_{1,0}(n-D-1) + \mathbf{K}_{0,1}(n-D-1)) \mathbf{R}_{\hat{D}-D} \\
 & + \mu^2 (2 \mathbf{R}_{D-\hat{D}} \mathbf{K}_{D,D}(n-1) \mathbf{R}_{\hat{D}-D} + \mathbf{R}_0 \text{tr}\{\mathbf{R}_0 \mathbf{K}_{D,D}(n-1)\}) \\
 & + \mu^2 \mathbf{R}_0 \sigma_z^2,
 \end{aligned} \tag{3.44a}$$

$$\mathbf{K}_{0,1}(n) = \mathbf{K}_{0,0}(n-1) - \mu \mathbf{R}_{D-\hat{D}} \mathbf{K}_{1,0}(n-1), \tag{3.44b}$$

$$\mathbf{K}_{0,1}(n) = \mathbf{K}_{1,0}^T(n). \tag{3.44c}$$

Como pode-se observar, o sistema de Eqs. (3.44) não pode ser manipulado como feito na Seção 3.1.11 com o sistema (3.31). Isto é, aqui não é possível a utilização de uma transformação ortogonal aplicada sobre a matriz de correlação da entrada, já que não temos uma matriz única para diagonalizar. Neste caso, as matrizes para serem diagonalizadas são: $\mathbf{R}_{D-\hat{D}}$, $\mathbf{R}_{\hat{D}-D}$ e \mathbf{R}_0 , o que não é possível realizar através de uma única transformação. Esta dificuldade não permite avançar no tratamento matemático de (3.44), a fim de realizar um estudo analítico da estabilidade do algoritmo, como na Seção 3.1.12.

3.1.13. Análise do DLMS utilizando a teoria da independência

Como as análises do comportamento do algoritmo DLMS, disponíveis na literatura, baseiam-se na teoria da independência, é importante compararmos os resultados aqui obtidos com os propostos na literatura. Assim, nesta seção vamos determinar o equacionamento do caso DLMS estudado à luz da teoria da independência (TI). O primeiro ponto a ser analisado é a Eq. (3.4), rescrita aqui por conveniência:

$$\begin{aligned}
 E[\mathbf{W}(n+1)] = & E[\mathbf{W}(n)] + \mu \left(E[d(n)\mathbf{X}(n-\hat{D})] - E[\mathbf{X}(n-\hat{D})\mathbf{X}^T(n-D)] E[\mathbf{W}(n-D)] \right. \\
 & \left. + E[z(n)\mathbf{X}(n-\hat{D})] \right),
 \end{aligned} \tag{3.4}$$

Segundo a TI os componentes dos vetores de entrada são independentes para instantes de tempo diferentes. Então, o valor esperado $E[\mathbf{X}(n-\hat{D})\mathbf{X}^T(n-D)]$, para $\hat{D} \neq D$, é zero. Dado que o sinal desejado é $d(n) = \mathbf{X}^T(n)\mathbf{W}^o$, o vetor de correlação cruzada

$\mathbf{P}_{\hat{D}} = E[d(n)\mathbf{X}(n - \hat{D})] = E[\mathbf{X}(n - \hat{D})\mathbf{X}^T(n)]\mathbf{W}^o$ também é zero. Com estes resultados, a Eq. (3.4) fica reduzida à seguinte expressão:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)].$$

Isto significa que sob a ótica da TI, o desenvolvimento posterior à Eq. (3.4) não existe. Mesmo para o caso mais simples de estimação perfeita do atraso, $\hat{D} = D$, já que o vetor $\mathbf{P}_{\hat{D}} = \mathbf{P}_D$ ($\mathbf{P}_D = E[\mathbf{X}(n-D)\mathbf{X}^T(n)]\mathbf{W}^o = 0$) é igual a zero, o que leva a um modelo errado para o comportamento do valor médio dos coeficientes. Com o fato de \mathbf{P}_D ser zero, não poderíamos calcular o valor ótimo, como em (3.6), nem as equações seguintes, vetor de erro nos coeficientes, etc.

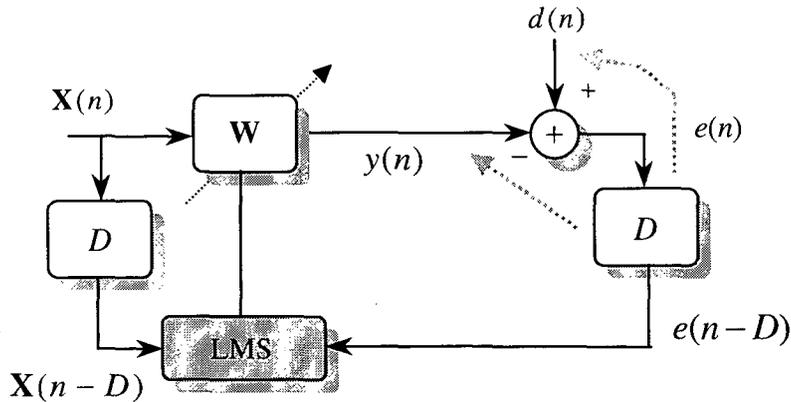
Como resumo, pode-se concluir que a não utilização da TI permitiu a elaboração de um modelo matemático para descrever adequadamente o comportamento do sistema adaptativo, utilizando o algoritmo DLMS.

3.1.14. Conversão do caso sob análise no caso estudado por Long et al.

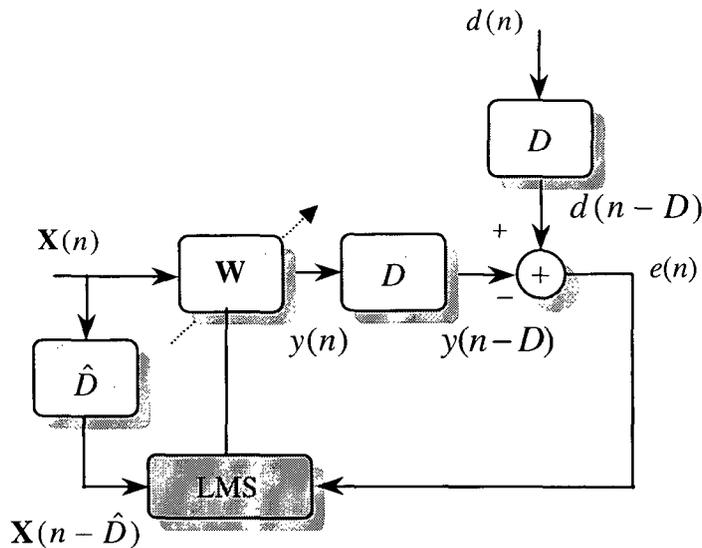
O artigo de Long et al., [9]¹, pode ser considerado com sendo um trabalho pioneiro em relação à análise do algoritmo DLMS, tendo como parâmetro disto o grande número de vezes que o mesmo tem sido referenciado. Nesta seção, comparamos a análise realizada neste trabalho com os resultados de [9,29]. Para tal, vamos modificar a estrutura proposta em [29], obtendo-se assim, uma estrutura que se aproxime do caso que estamos estudando. Tal modificação pode ser melhor compreendida pelas seguintes transformações de blocos.

Diagrama da estrutura proposta em [9,29]:

¹ Posteriormente foi publicada uma correção deste trabalho, Referência [29].



Um diagrama equivalente pode ser obtido deslocando o bloco correspondente ao atraso na direção indicada pelas setas. Também vamos substituir por \hat{D} o bloco correspondente à estimação do atraso, o que confere maior generalidade ao estudo realizado em [9,29] (onde $\hat{D} = D$). O diagrama em blocos resultante é o seguinte:



A partir desse diagrama, vemos que o caso analisado em [9,29] corresponde a substituímos $d(n)$ por $d(n-D)$ em nossas equações. As equações modificadas são apontadas a seguir. De (3.1) obtemos:

$$e(n) = d(n-D) - \mathbf{X}^T(n-D)\mathbf{W}(n-D) + z(n). \quad (3.45)$$

As Eqs. (3.2), (3.3) e (3.4), resultam em:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] + \mu \left(E[d(n-D)\mathbf{X}(n-\hat{D})] - E[\mathbf{X}(n-\hat{D})\mathbf{X}^T(n-D)]E[\mathbf{W}(n-D)] + E[z(n)\mathbf{X}(n-\hat{D})] \right), \quad (3.46)$$

Chamando $\mathbf{P}_{D-\hat{D}} = E[d(n-D)\mathbf{X}(n-\hat{D})]$, podemos rescrever a expressão equivalente para a (3.5), dada por:

$$E[\mathbf{W}(n+1)] = E[\mathbf{W}(n)] + \mu (\mathbf{P}_{D-\hat{D}} - \mathbf{R}_{D-\hat{D}} E[\mathbf{W}(n-D)]) \quad (3.47)$$

A Eq. (3.6), é modificada para

$$\mathbf{R}_0 \mathbf{W}_{ot} = \mathbf{P}_0 \quad (3.49)$$

onde $\mathbf{P}_0 = E[d(n-D)\mathbf{X}(n-D)]$.

A Eq. (3.9) resulta em:

$$\mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty = \mathbf{P}_{D-\hat{D}} \quad (3.50)$$

Note que para estimação perfeita do atraso, (3.50) reduz-se a (3.49). É importante observar, também, que o valor em regime permanente (3.50) é diferente do obtido de (3.9).

As Eqs. (3.10) e (3.11) são também afetadas; no entanto, considerando a relação (3.50), a Eq. (3.12) é mantida inalterada.

As Eqs. (3.13) ou (3.14), resultam em:

$$\begin{aligned}
 J(n) = & E[d^2(n)] - 2\mathbf{P}_0^T E[\mathbf{V}(n-D)] - 2\mathbf{P}_0^T \mathbf{W}_\infty \\
 & + \mathbf{W}_\infty^T \mathbf{R}_0 \mathbf{W}_\infty + \mathbf{W}_\infty^T \mathbf{R}_0 E[\mathbf{V}(n-D)] \\
 & + E[\mathbf{V}^T(n-D)] \mathbf{R}_0 \mathbf{W}_\infty \\
 & + \text{tr} \left\{ \mathbf{R}_0 E[\mathbf{V}(n-D) \mathbf{V}^T(n-D)] \right\} + \sigma_z^2.
 \end{aligned} \tag{3.51}$$

As Eqs. (3.15) a (3.17) são modificadas. No entanto, aplicando-se (3.50), verifica-se que o sistema de equações que descreve o momento de segunda ordem não apresenta alteração. Assim, continuam valendo as Eqs. (3.37b) e (3.35d) para o caso de estimação perfeita e, as (3.44) para o caso mais geral $\hat{D} \neq D$.

As equações que permitem descrever o comportamento do algoritmo DLMS para os casos da Fig. 3.1 e o estudado em [9,29] são resumidas nas tabelas I e II.

Tabela I.

Caso	Valor dos coeficientes em regime permanente	
	$\hat{D} = D$	$\hat{D} \neq D$
DLMS Fig. 3.1	$\mathbf{R}_0 \mathbf{W}_{ot} = \mathbf{P}_D$	$\mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty = \mathbf{P}_{\hat{D}}$
DLMS Long et al. [9,29]	$\mathbf{R}_0 \mathbf{W}_{ot} = \mathbf{P}_0$	Não considerado
DLMS da Fig. 3.1 Substituindo $d(n)$ por $d(n-D)$	$\mathbf{R}_0 \mathbf{W}_{ot} = \mathbf{P}_0$	$\mathbf{R}_{D-\hat{D}} \mathbf{W}_\infty = \mathbf{P}_{D-\hat{D}}$

Tabela II.

Caso	Equações que descrevem o comportamento do algoritmo utilizando o vetor erro nos coeficientes $V(n)$					
	$\hat{D} = D$			$\hat{D} \neq D$		
	M 1	M 2	EMQ	M 1	M 2	EMQ
DLMS Fig. 3.1	(3.12) subs. $\mathbf{R}_{D-\hat{D}}$ por \mathbf{R}_0	(3.37b) (3.35d)	(3.14) subs. \mathbf{W}_∞ por \mathbf{W}_{ot}	(3.12)	(3.44a) (3.44b) (3.44c)	(3.14)
DLMS Long et al. [9,29]	(3.12) subs. $\mathbf{R}_{D-\hat{D}}$ por \mathbf{R}_0	(3.37b) (3.35d)	(3.14) subs. \mathbf{W}_∞ por \mathbf{W}_{ot}	Não considera da	Não considera da	Não considera da
DLMS da Fig. 3.1 Substituindo $d(n)$ por $d(n-D)$	(3.12) subs. $\mathbf{R}_{D-\hat{D}}$ por \mathbf{R}_0	(3.37b) (3.35d)	(3.14) subs. \mathbf{W}_∞ por \mathbf{W}_{ot}	(3.12)	(3.44a) (3.44b) (3.44c)	(3.14)

Como as equações de diferenças que descrevem o momento de segunda ordem para o caso $\hat{D} = D$ não são modificadas, o limite de estabilidade dado pela Eq. (3.42) é o mesmo para qualquer um dos casos da Tabela II.

Assim, a análise apresentada nesta seção inclui as análises do algoritmo DLMS presentes na literatura, permitindo ainda o estudo do caso de erro na estimação do atraso. Uma outra vantagem desta análise é a forma mais compacta das equações que descrevem o momento de segunda ordem. A expressão apresentada em [9,29] tem infinitos termos.

3.2 RESULTADOS DE SIMULAÇÕES PARA O ALGORITMO DLMS

Nesta seção são apresentados resultados de simulações a fim de avaliar a modelagem proposta. Primeiramente, são mostrados dois exemplos relativos à estabilidade do algoritmo DLMS quando $\hat{D} \neq D$ (Seção 3.1.8). Os exemplos restantes mostram as evoluções temporais do vetor erro nos coeficientes e do EMQ correspondentes à simulação e ao modelo proposto.

3.2.1 Simulação 1:

A planta utilizada para este exemplo é $\mathbf{W}^o = [-1,5; 1; -0,7]$. Assumimos estimação imperfeita do atraso no caminho secundário, sendo $D = 1$ e sua estimação $\hat{D} = 0$. A taxa de aprendizagem é $\mu = 0,01$ ($(\mu_{\max})_{LMS} = 0,2$). O sinal de entrada é um ruído colorido gaussiano com variância $\sigma_x^2 = 1$. O ruído de medição é um ruído branco gaussiano com variância $\sigma_z^2 = 0,0001$. Os coeficientes do processo AR que gera o sinal de entrada são $a_1 = -0,195$ e $a_2 = 0,95$. A dispersão dos autovalores é $\chi = 1,22$ e os elementos da matriz de correlação são $r_0 = 1$, $r_1 = 0,1$, $r_2 = -0,9305$ e $r_3 = 0,2764$.

A Fig. 3.2a apresenta o comportamento médio dos coeficientes do filtro adaptativo, obtidos por simulação e usando o modelo dado pela Eq. (3.5). Da Eq. (3.9) temos que o valor em regime permanente é $\mathbf{W}_\infty = [2897; -596,5; 3048,1]$, o qual pode ser verificado nesta figura.

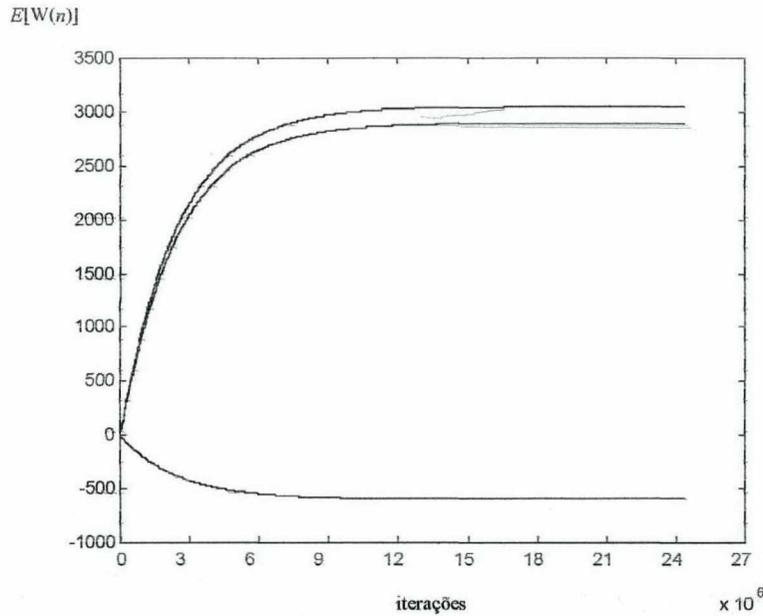


Fig. 3.2a. Evolução dos coeficientes do filtro adaptativo para entrada correlacionada e usando o algoritmo DLMS.

Na Fig. 3.2b é mostrado um detalhe da Fig. 3.2a correspondente às primeiras 5000 iterações. A simulação, em linha pontilhada, é obtida pela média de 20 realizações. Pode-se observar que o modelo proposto prediz satisfatoriamente o comportamento oscilatório dos coeficientes do filtro. Este comportamento não pode ser previsto usando a TI, a qual leva a um comportamento médio exponencial para os coeficientes.

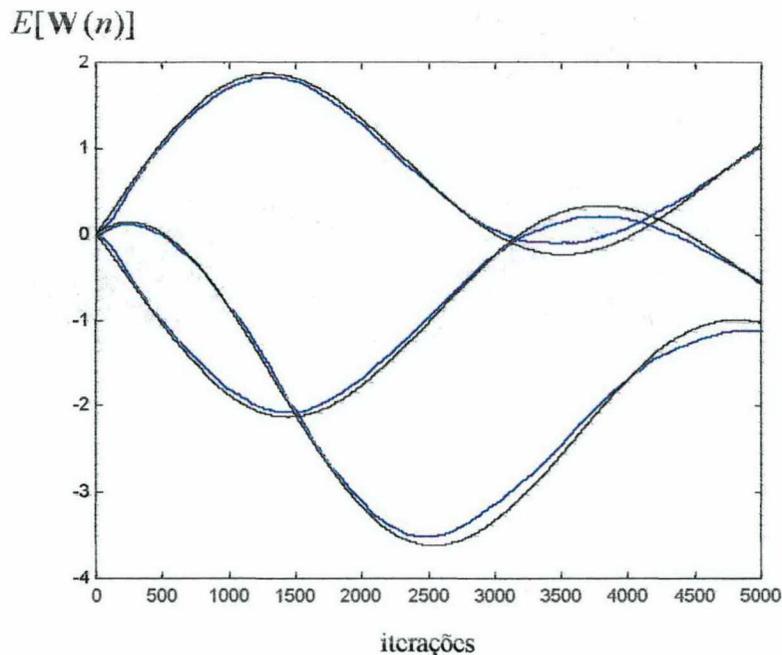


Fig. 3.2b. Detalhe das primeiras 5000 iterações da Fig. 3.2a. Simulação (linha azul) e modelo Eq. (3.5) (linha preta).

3.2.2 Simulação 2:

Neste exemplo, a planta é dada por $\mathbf{W}^o = [1; 0,5]$. O atraso do caminho secundário é $D=1$ e sua estimação $\hat{D}=0$. A taxa de aprendizagem utilizada é $\mu=0,01$. Como sinal de entrada usamos um ruído branco gaussiano com variância $\sigma_x^2=1$. O ruído de medição é, também, branco e gaussiano com variância $\sigma_z^2=0,0001$.

A Fig. 3.3, apresenta o comportamento médio dos coeficientes de acordo com (3.5). Verifica-se que o algoritmo não converge, conforme previsto na seção 3.18. A impossibilidade de convergência para $\hat{D} \neq D$ com entrada ruído branco foi comentada em [26], mais não prevista teoricamente.

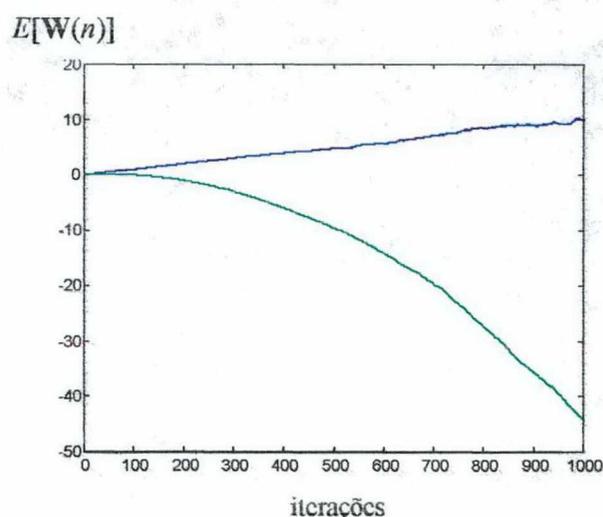


Fig. 3.3. Evolução dos coeficientes do filtro adaptativo, caso DLMS, com entrada não correlacionada. Coeficiente $w_0(n)$ (linha azul) e $w_1(n)$ (linha verde).

3.2.3 Simulação 3:

Para este exemplo utilizamos como planta: $\mathbf{W}^o = [-0,197; 0,1179; 0,8; 0,19; -0,0197]$.

Os atraso do caminho secundário e sua estimação são $D=\hat{D}=1$. A taxa de aprendizagem utilizada é $\mu=0,005$, sendo o máximo valor dado pela Eq. (3.43)

$\mu_{\max}=0,222$. O sinal de entrada é um ruído branco gaussiano com variância $\sigma_x^2=1$. O

ruído de medição adicionado é branco e gaussiano com variância $\sigma_z^2 = 0,0001$. A partir da Eq. (3.9), o valor em regime permanente é $\mathbf{W}_\infty = [0,1179; 0,8; 0,19; -0,0197; 0]$.

A Fig. 3.4.a ilustra a evolução temporal do vetor erro nos coeficientes. Assim, pode-se avaliar o uso da Eq. (3.9). Os componentes do vetor erro nos coeficientes (Eq. (3.12)) tendem a zero com o decorrer das iterações.

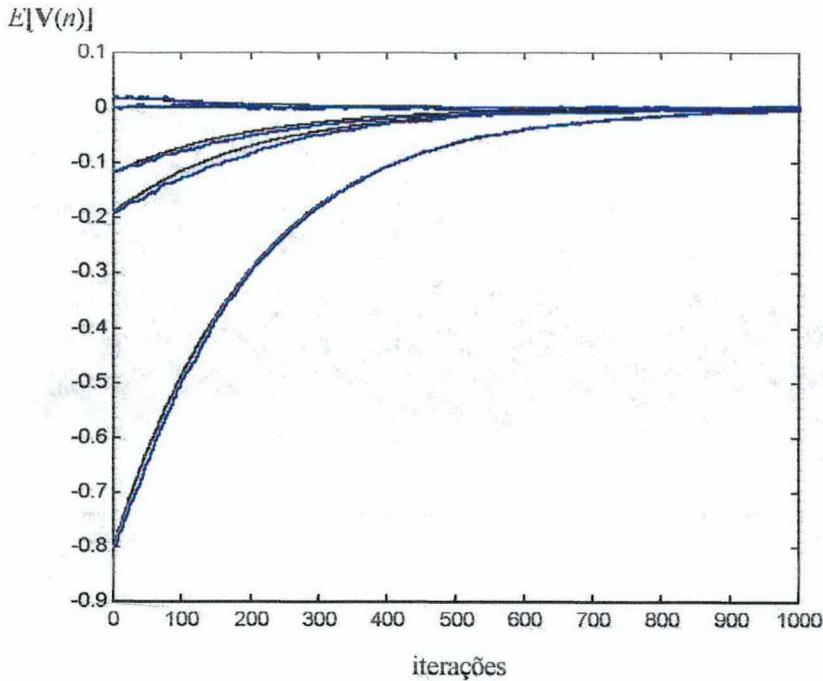


Fig. 3.4a. Evolução dos coeficientes do vetor erro nos coeficientes. Simulação (linha azul), modelo Eq. (3.12) (linha preta).

Na Fig. 3.4b é mostrado o erro médio quadrático obtido a partir das Eqs. (3.12), (3.37b), (3.35d) e (3.14). Desta figura, pode-se avaliar o uso da Eq. (2.44), a qual estima o erro cometido pelo algoritmo.

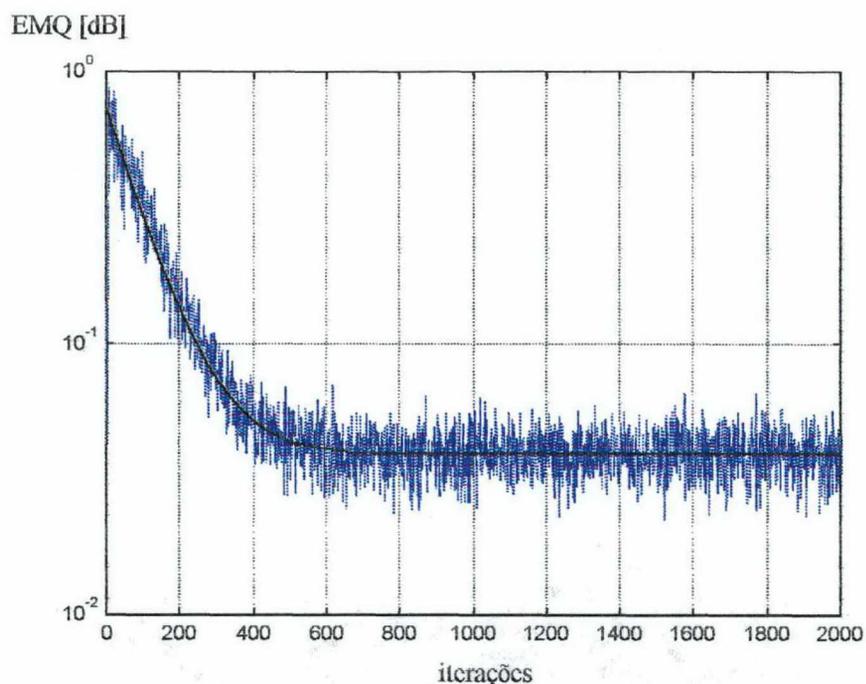


Fig. 3.4b. Erro médio quadrático. Simulação (linha azul), modelo (linha preta). Erro medido 0,0404, erro estimado pela Eq. (2.44) 0,0403.

3.2.4 Simulação 4:

Neste exemplo são utilizadas as mesmas respostas da Simulação 3. No entanto, aumentou-se o valor do atraso para: $D = \hat{D} = 2$. A taxa de aprendizagem é $\mu = 0,005$ sendo o máximo valor dado pela Eq. (3.43) $\mu_{\max} = 0,1818$. O sinal de entrada é um ruído branco gaussiano com variância $\sigma_x^2 = 1$. O ruído de medição utilizado tem variância $\sigma_z^2 = 0,0001$. Da Eq. (3.9), o valor em regime permanente dos coeficientes do filtro adaptativo é $\mathbf{W}_\infty = [0,8; 0,19; -0,0197; 0; 0]$, o qual pode ser verificado a partir da Fig. 3.5a, uma vez que o vetor erro nos coeficientes (Eq. (3.12)) tende a zero com o decorrer das iterrações.

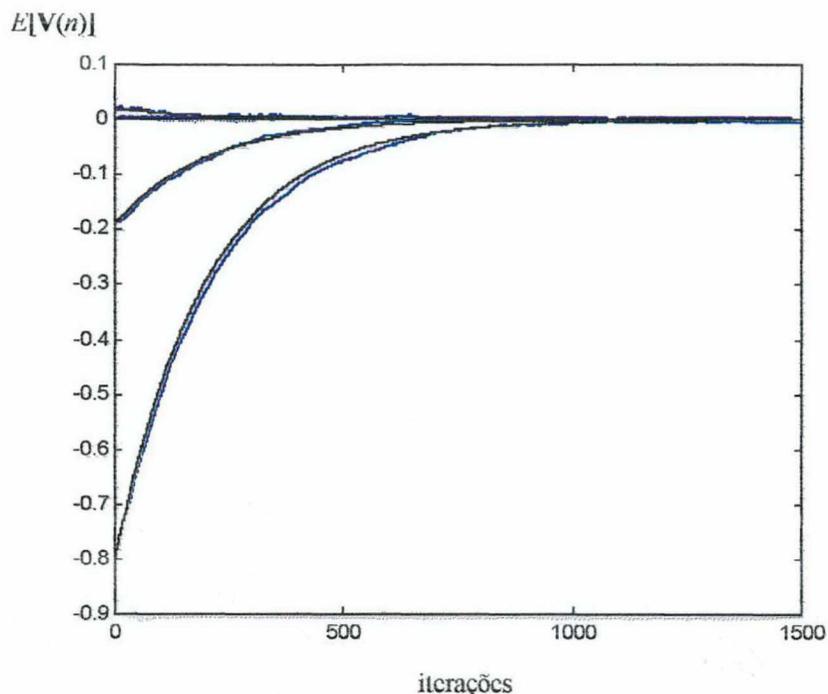


Fig. 3.5a. Evolução dos coeficientes do vetor erro nos coeficientes. Simulação (linha azul), modelo Eq. (3.12) (linha preta).

Na Fig. 3.5b é mostrado o erro médio quadrático obtido a partir das equações (3.12), (3.37b), (3.35d) e (3.14). Desta figura, pode-se avaliar o uso da Eq. (2.44), a qual estima o erro cometido pelo algoritmo.

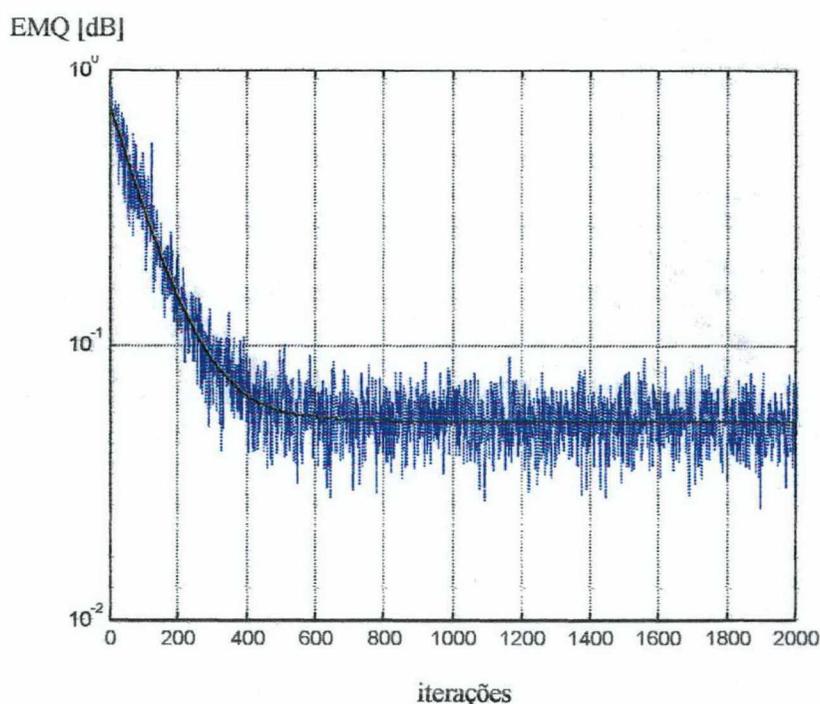


Fig. 3.5b. Erro médio quadrático. Simulação (linha azul), modelo (linha preta). Erro medido 0,0536, erro estimado pela Eq. (2.44) 0,054.

Da comparação entre as Figs. 3.4b e 3.5b, observa-se que, à medida que o atraso aumenta, o EMQ também aumenta, como se devia esperar. Este fato também é previsto pela Eq. (2.44).

3.3 CASO GERAL

Esta seção estende a análise da seção 3.2 para o estudo do comportamento dos momentos de segunda ordem dos coeficientes do filtro adaptativo correspondente ao caso mais geral. Neste caso os blocos correspondentes ao caminho secundário e sua estimação realizam operações de filtragem.

Por conveniência, é repetida a figura correspondente ao diagrama em blocos do algoritmo LMS filtrado da Fig. 3.6. Da mesma forma que na Fig. 3.1, W^o é o sistema desconhecido e $W(n)$ é o filtro adaptativo. Como foi anteriormente mencionado, a diferença em relação ao caso DLMS, reside em que S e \hat{S} são modelados agora por filtros FIR lineares.

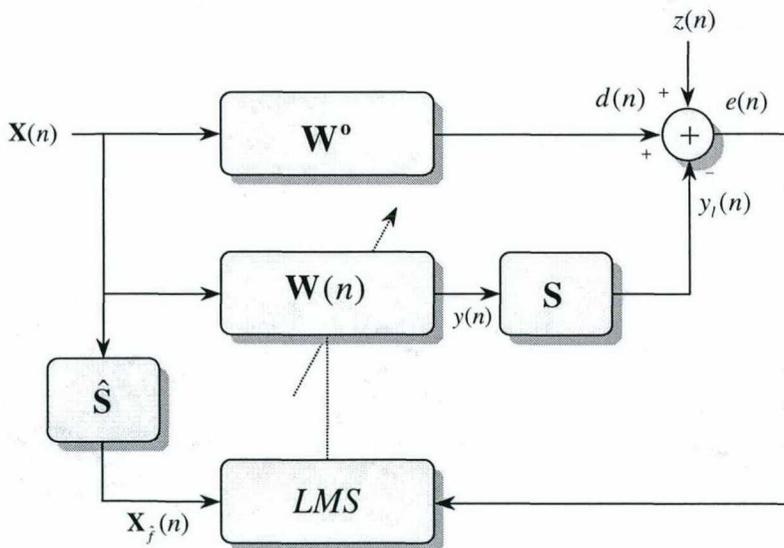


Figura 3.6. Diagrama em blocos do algoritmo LMS filtrado.

Tal como na seção anterior, vamos começar determinando a expressão do erro instantâneo, para depois calcular o valor esperado do erro instantâneo ao quadrado. Uma

vez obtida esta última expressão, podemos conhecer a forma da matriz de correlação que devemos determinar.

3.3.1. Sinal de erro

A partir do diagrama em blocos da Fig 3.6, o sinal de erro é dado por:

$$e(n) = d(n) - y_l(n) + z(n), \quad (3.45)$$

onde,

$$d(n) = \mathbf{X}^T(n) \mathbf{W}^o, \quad (3.46)$$

$$y(n) = \mathbf{X}^T(n) \mathbf{W}(n) = \mathbf{W}^T(n) \mathbf{X}(n), \quad (3.47)$$

$$y_l(n) = \sum_{i=0}^{M-1} s_i y(n-i) = \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \mathbf{W}(n-i). \quad (3.48)$$

Substituindo (3.47) e (3.48) em (3.45), obtém-se a expressão para o erro instantâneo, dada por:

$$e(n) = d(n) - \sum_{i=0}^{M-1} s_i \mathbf{X}^T(n-i) \mathbf{W}(n-i) + z(n). \quad (3.49)$$

3.3.2. Erro Médio Quadrático

Elevando ao quadrado a Eq. (3.49), temos:

$$\begin{aligned} e^2(n) &= d^2(n) - 2 \sum_{i=0}^{M-1} s_i d(n) \mathbf{X}^T(n-i) \mathbf{W}(n-i) \\ &\quad + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{X}^T(n-i) \mathbf{W}(n-i) \mathbf{X}^T(n-j) \mathbf{W}(n-j) \\ &\quad + 2d(n)z(n) - 2 \sum_{i=0}^{M-1} s_i z(n) \mathbf{X}^T(n-i) \mathbf{W}(n-i) + z^2(n). \end{aligned} \quad (3.50)$$

Tomando o valor esperado em ambos lados da Eq. (3.50), temos:

$$\begin{aligned}
 E[e^2(n)] &= E[d^2(n)] - 2 \sum_{i=0}^{M-1} s_i E[d(n) \mathbf{X}^T(n-i) \mathbf{W}(n-i)] \\
 &\quad + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j E[\mathbf{X}^T(n-i) \mathbf{W}(n-i) \mathbf{X}^T(n-j) \mathbf{W}(n-j)] \\
 &\quad - 2 \sum_{i=0}^{M-1} s_i E[z(n) \mathbf{X}^T(n-i) \mathbf{W}(n-i)] \\
 &\quad + 2E[d(n)z(n)] + E[z^2(n)].
 \end{aligned} \tag{3.51}$$

Para obtermos a expressão do erro médio quadrático (EMQ), a partir da Eq. (3.51) vamos utilizar as suposições (III.i) e (III.ii), resultando na seguinte expressão:

$$\begin{aligned}
 E[e^2(n)] &= E[d^2(n)] - 2 \sum_{i=0}^{M-1} s_i \mathbf{P}_i^T E[\mathbf{W}(n-i)] \\
 &\quad + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \text{tr} \{ \mathbf{R}_{j-i} E[\mathbf{W}(n-j) \mathbf{W}^T(n-i)] \} + \sigma_z^2.
 \end{aligned} \tag{3.52}$$

Expressando a Eq. (3.52) em função do vetor erro nos coeficientes $\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}_\infty$, obtém-se a seguinte expressão:

$$\begin{aligned}
 E[e^2(n)] &= E[d^2(n)] - 2 \sum_{i=0}^{M-1} s_i \mathbf{P}_i^T E[\mathbf{V}(n-i)] - 2 \sum_{i=0}^{M-1} s_i \mathbf{P}_i^T \mathbf{W}_\infty \\
 &\quad + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \text{tr} \{ \mathbf{R}_{j-i} E[\mathbf{V}(n-j) \mathbf{W}_\infty^T] \} + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \text{tr} \{ \mathbf{R}_{j-i} \mathbf{W}_\infty E[\mathbf{V}^T(n-i)] \} \\
 &\quad + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \text{tr} \{ \mathbf{R}_{j-i} \mathbf{W}_\infty \mathbf{W}_\infty^T \} + \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \text{tr} \{ \mathbf{R}_{j-i} E[\mathbf{V}(n-j) \mathbf{V}^T(n-i)] \} + \sigma_z^2.
 \end{aligned} \tag{3.53}$$

A partir da Eq. (3.53), da mesma forma que em (3.14), pode-se observar que para descrever o comportamento do erro médio quadrático, deve-se conhecer o valor esperado do vetor erro $E[\mathbf{V}(n)]$, já determinado no Capítulo 2 (Eq. (2.50)), e a matriz de correlação $E[\mathbf{V}(n-j) \mathbf{V}^T(n-i)]$. De maneira análoga ao cálculo de (3.17), o primeiro passo é calcular o produto $\mathbf{V}(n-j) \mathbf{V}^T(n-i)$, como em (3.16), utilizando (2.47). Em

seguida, calcula-se o valor esperado utilizando (III) e, empregando a notação adotada para escrever (3.19), a expressão recursiva para $\mathbf{K}_{j,i}(n)$ pode ser escrita como:

$$\begin{aligned}
 \mathbf{K}_{j,i}(n) &= \mathbf{K}_{j,i}(n-1) \\
 &- \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{K}_{j,k+i}(n-1) \mathbf{R}_{l-k} - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k+j,i}(n-1) \\
 &+ \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{r=0}^{M-1} \sum_{m=0}^{\hat{M}-1} s_k \hat{s}_l s_r \hat{s}_m (\mathbf{R}_{k-l} \mathbf{K}_{k+j,r+i}(n-1) \mathbf{R}_{m-r} \\
 &\quad + \mathbf{R}_{r+i-l-j} \mathbf{K}_{r+i,k+j}(n-1) \mathbf{R}_{m+i-k-j} \\
 &\quad + \mathbf{R}_{m+i-l-j} \text{tr}\{\mathbf{R}_{r+i-k-j} \mathbf{K}_{k+j,r+i}(n-1)\}) \\
 &+ \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k s_l \mathbf{R}_{l+i-k-j} E[z(n-i-1)z(n-j-1)].
 \end{aligned} \tag{3.54}$$

No Apêndice B são detalhados os passos para a obtenção de (3.54). Para descrever completamente a variável $\mathbf{K}_{j,i}(n)$, devem-se conhecer também as variáveis $\mathbf{K}_{j,k+i}(n) = E[\mathbf{V}(n-j)\mathbf{V}^T(n-k-i)]$ e $\mathbf{K}_{k+j,i}(n) = E[\mathbf{V}(n-k-j)\mathbf{V}^T(n-i)]$. O cálculo de $\mathbf{K}_{j,k+i}(n)$ ou $\mathbf{K}_{k+j,i}(n)$ é realizado como em (3.23). Partindo de (2.47) determina-se, por exemplo, $\mathbf{V}(n-j)$. Pós-multiplicando ambos lados por $\mathbf{V}^T(n-k-i)$, o valor esperado $E[\mathbf{V}(n-j)\mathbf{V}^T(n-k-i)]$ resulta a seguinte expressão:

$$\mathbf{K}_{j,k+i}(n) = \mathbf{K}_{j,k+i-1}(n-1) - \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{K}_{m+j,k+i-1}(n-1) \tag{3.55}$$

Após a aplicação do processo de redução dos índices das matrizes $\mathbf{K}_{j,i}(n)$, procedendo como em (3.21), obtém-se o seguinte sistema de equações que descrevem o comportamento da matriz de correlação:

$$\begin{aligned}
 \mathbf{K}_{0,0}(n) &= \mathbf{K}_{0,0}(n-1) \\
 &- \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{K}_{0,k}(n-1) \mathbf{R}_{l-k} - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k,0}(n-1) \\
 &+ \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{r=0}^{M-1} \sum_{m=0}^{\hat{M}-1} s_k \hat{s}_l s_r \hat{s}_m (\mathbf{R}_{k-l} \mathbf{K}_{k,r}(n-1) \mathbf{R}_{m-r} \\
 &\quad + \mathbf{R}_{r-l} \mathbf{K}_{r,k}(n-1) \mathbf{R}_{m-k} \\
 &\quad + \mathbf{R}_{m-l} \text{tr}\{\mathbf{R}_{r-k} \mathbf{K}_{k,r}(n-1)\}) \\
 &+ \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} s_k s_l \mathbf{R}_{l-k} \sigma_z^2,
 \end{aligned} \tag{3.56a}$$

$$\mathbf{K}_{0,\Delta}(n) = \mathbf{K}_{0,\Delta-1}(n-1) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k,\Delta-1}(n-1), \quad \Delta = 1, \dots, M-1 \tag{3.56b}$$

As Eqs. (3.56) são completamente análogas às (3.25). No Apêndice B, é descrito o procedimento para a determinação das Equações (3.56a) e (3.56b).

3.4 RESULTADOS EXPERIMENTAIS

Nesta seção são apresentados resultados de simulações que avaliam a qualidade do novo modelo proposto. São mostradas as curvas correspondentes à evolução temporal do erro médio quadrático obtidas por simulação e, a partir do novo modelo proposto Eqs. (3.53) e (3.56).

3.4.1 Simulação 1:

A planta utilizada neste exemplo é dada pela seguinte resposta ao impulso:

$$\mathbf{W}^o = [1; 0,8415; 0,4546; 0,0470; -0,1892; -0,1918; -0,0466; 0,0939; 0,1237; 0,0458; -0,0544].$$

O caminho secundário e sua estimação são representados pela resposta ao impulso $\mathbf{S} = \hat{\mathbf{S}} = [1; 0,5]$. A taxa de aprendizagem é $\mu = 0,005$ ($(\mu_{\max})_{LMS} = 0,0909$). Como sinal de entrada utilizou-se um ruído branco gaussiano com variância $\sigma_x^2 = 1$ e como ruído de medição foi utilizado um ruído branco gaussiano com variância $\sigma_z^2 = 0,0001$.

A Fig. 3.4.1a ilustra a relação entre as bandas passantes das funções de transferência correspondentes à planta e ao caminho secundário.

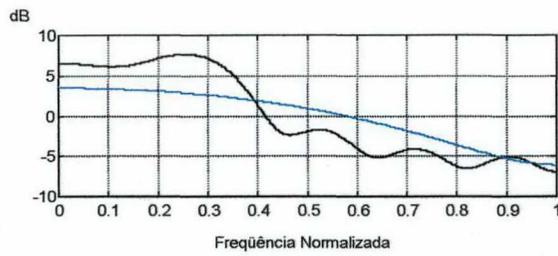


Fig. 3.4.1a. Relação entre as bandas passantes da planta (linha preta) e do caminho secundário (linha azul).

Na Fig. 3.4.1b, é mostrado o erro médio quadrático obtido por simulação e o calculado a partir das Eqs. (3.56), (3.53) e (2.50). Na mesma figura pode ser avaliado o uso da Eq. (2.44) que oferece uma estimação do erro em excesso. A partir da simulação o erro em excesso medido é $7,3410^{-4}$ e através da Eq. (2.44) obtém-se $7,4410^{-4}$, o que representa uma aceitável predição.

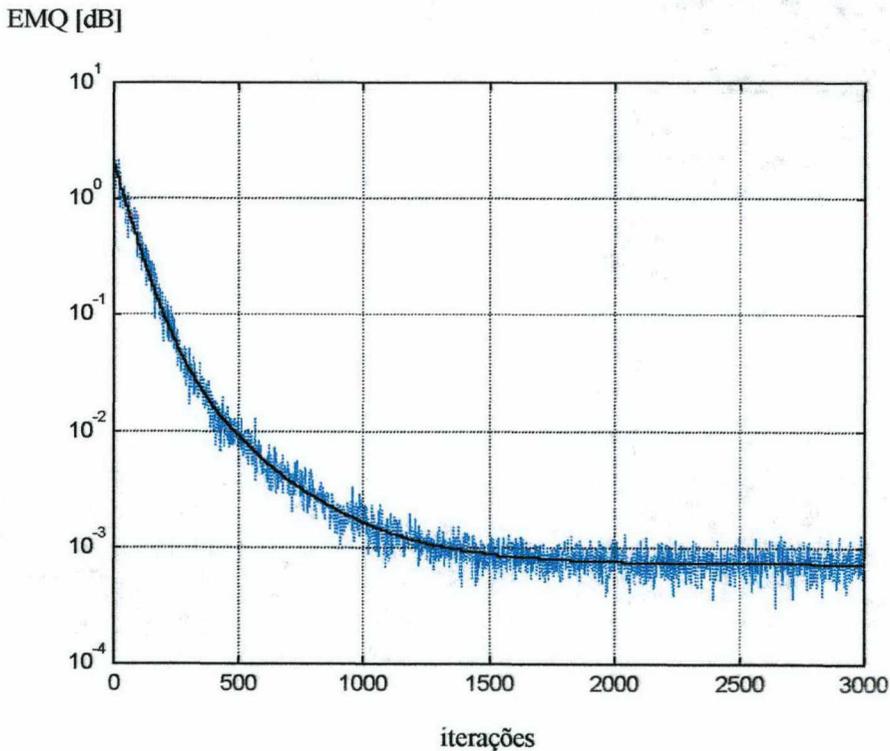


Fig. 3.4.1b. Erro médio quadrático: simulação (linha azul) e modelo Eq. (3.56) (linha preta). Erro medido $7,341 10^{-4}$, erro estimado pela Eq. (2.44) $7,441 10^{-4}$.

3.4.2 Simulação 2:

A planta para este exemplo é dada por $\mathbf{W}^o = [-0,02; 0,1179; 0,8; 0,1179; -0,02]$. O caminho secundário é $\mathbf{S} = [1; 0,5]$ e sua estimativa é $\hat{\mathbf{S}} = [1; -0,5]$. A taxa de aprendizagem é $\mu = 0,005$ ($(\mu_{\max})_{LMS} = 0,2$). Como sinal de entrada foi utilizado um ruído branco gaussiano com variância $\sigma_x^2 = 1$. O ruído de medição é branco e gaussiano com variância $\sigma_z^2 = 0,0001$.

A Fig. 3.4.2a mostra a relação entre as bandas passantes correspondentes à planta, ao caminho secundário e à estimação do caminho secundário. Neste exemplo foi utilizada como estimação do caminho secundário, uma transferência com características opostas a \mathbf{S} com o objetivo de verificar se o modelo (Eq. (3.56)) prediz o comportamento do EMQ nesta situação.

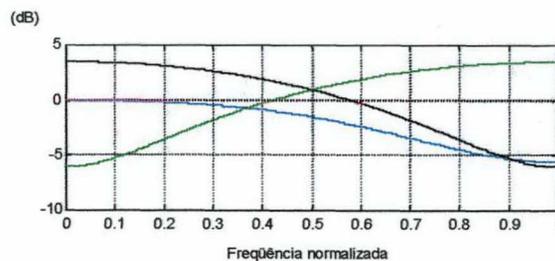


Fig. 3.4.2a. Relação entre as bandas passantes da planta (linha preta), do caminho secundário (linha azul) e de sua estimação (linha verde).

A partir da Fig. 3.4.2b, pode-se observar que a falta de casamento entre \mathbf{S} e sua estimação se traduz num aumento do erro mínimo. Da simulação temos que o erro medido em regime permanente é 0,0060. Usando (2.44) obtemos 0,0058, que corresponde a uma boa indicação do erro que será atingido pelo algoritmo.

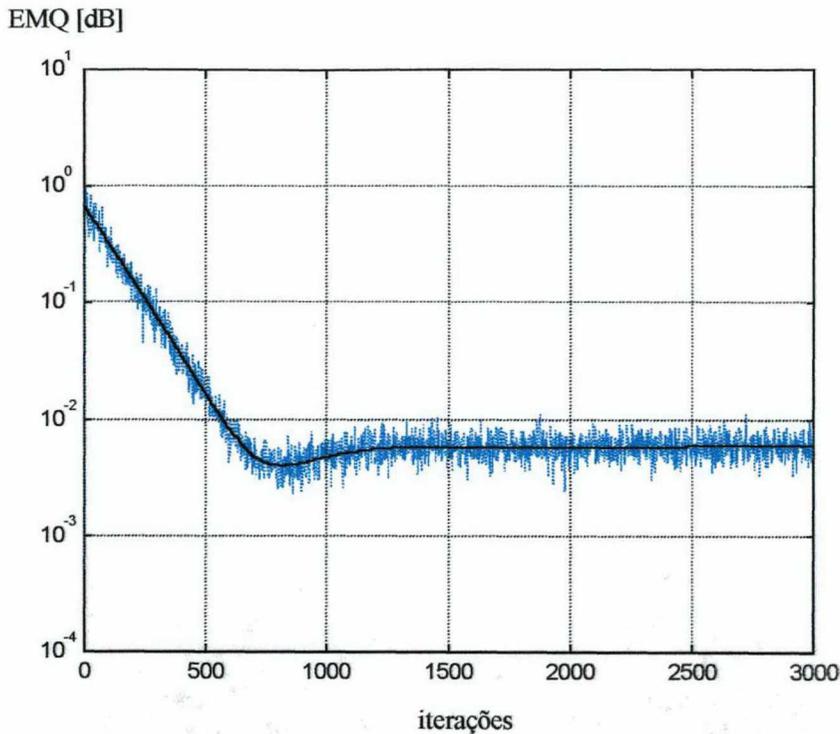


Fig. 3.4.2b. Erro médio quadrático: simulação (linha azul) e modelo Eq. (3.56) (linha preta). Erro medido 0,006, erro estimado pela Eq. (2.44) 0,0058.

3.4.3 Simulação 3:

A planta utilizada é $\mathbf{W}^o = [-1,5; 1; -0,7]$. O caminho secundário e sua estimação são dados por $\mathbf{S} = \hat{\mathbf{S}} = [1; 0,5]$. A taxa de aprendizagem utilizada é $\mu = 0,008$ ($(\mu_{\max})_{LMS} = 0,2$). Como sinal de entrada foi utilizado um ruído colorido gaussiano com variância $\sigma_x^2 = 1$. Como ruído de medição foi utilizado um ruído branco gaussiano com variância $\sigma_z^2 = 0,0001$. O sinal correlacionado é gerado através de um processo AR com coeficientes dados por $a_1 = -0,195$ e $a_2 = 0,95$. A dispersão dos autovalores da matriz de correlação é $\chi = 1,22$. Os elementos da matriz de correlação são dados por $r_0 = 1$, $r_1 = 0,1$, $r_2 = -0,9305$ e $r_3 = 0,2764$.

A Fig. 3.4.3a apresenta as bandas passantes correspondentes à planta, ao caminho secundário e ao filtro que cria o processo correlacionado gaussiano.

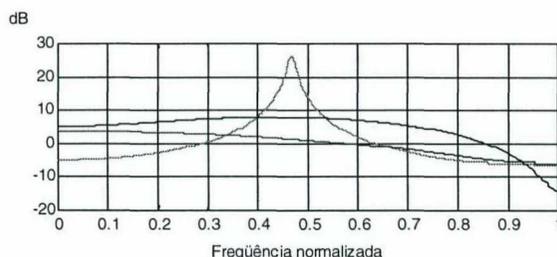


Fig. 3.4.3a. Relação entre as bandas passantes da planta (linha preta), do caminho secundário (linha azul) e do filtro que cria a entrada correlacionada (linha vermelha).

Na Fig. 3.4.3b é mostrado o erro médio quadrático. Desta figura pode-se observar o bom casamento entre a predição, obtida a partir do modelo proposto (Eq (3.56)), e o resultado da simulação.

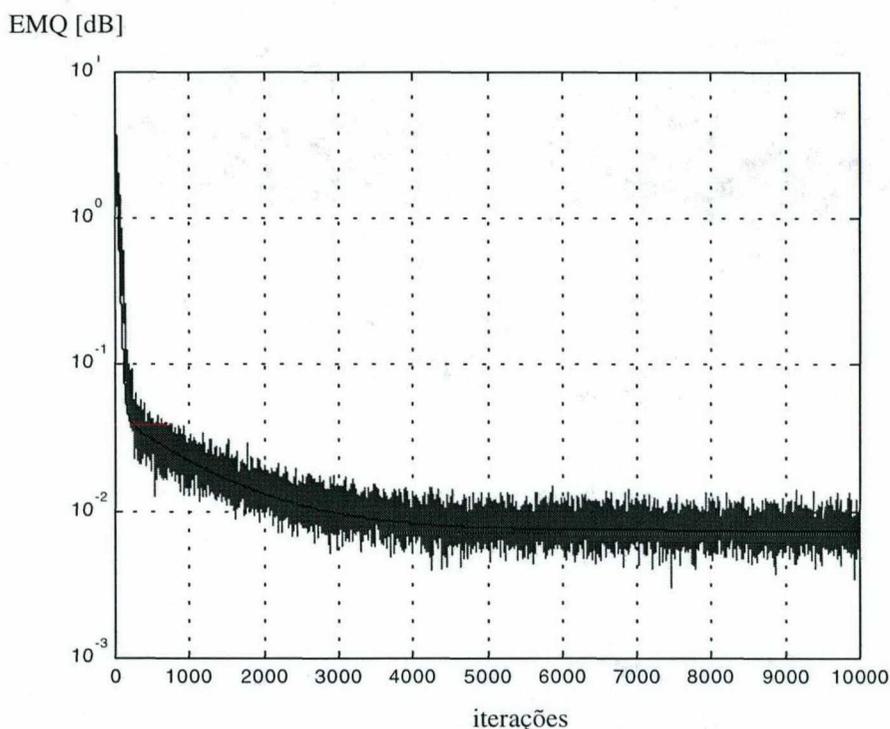


Fig. 3.4.3b. Erro médio quadrático: simulação (linha azul) e modelo Eq. (3.56) (linha preta). Erro medido 0,0079, erro estimado pela Eq. (2.44) 0,00791.

3.4.4 Simulação 4:

Neste exemplo vamos comparar o modelo proposto com o modelo obtido utilizando a teoria da independência. Para isto, vamos utilizar os mesmos dados que na simulação 1. A Fig. 3.4.4 ilustra o erro médio quadrático obtido a partir de (3.56) e o mesmo baseado na teoria da independência. Para obter este último modelo devemos zerar as matrizes

R_{i-j} quando $i \neq j$. Isto ocasiona que muito termos sejam desconsiderados, tanto na expressão do cálculo do valor esperado quanto nas equações do cálculo da matriz de correlação. Desta figura, pode-se observar que a predição do comportamento do erro médio quadrático, obtida a partir do modelo que utiliza a teoria da independência, não representa o comportamento real.

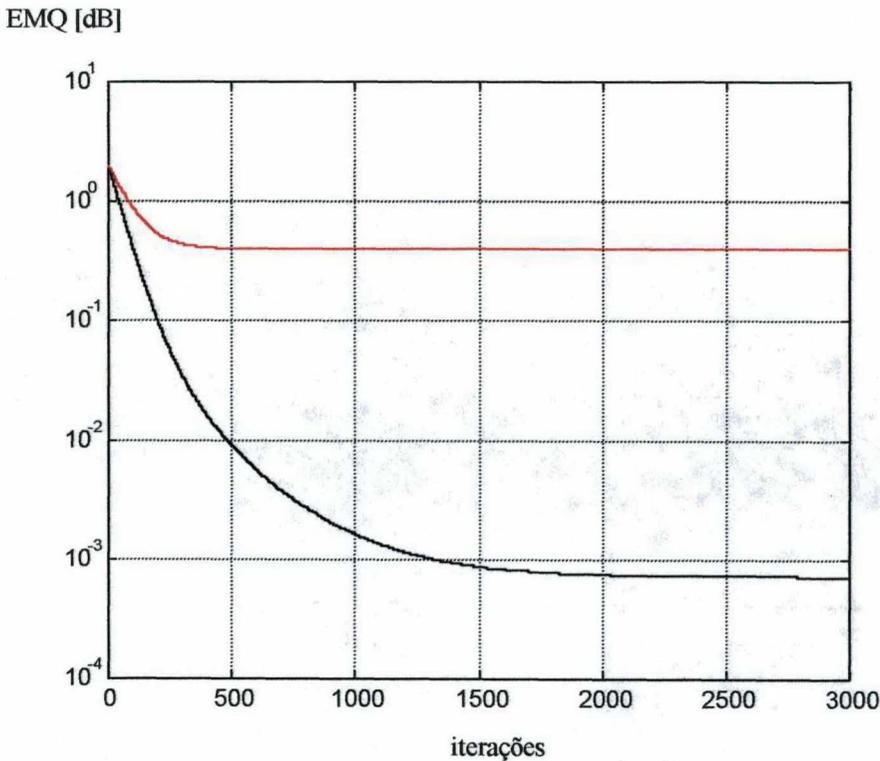


Fig. 3.4.4. Erro médio quadrático: modelo utilizando a teoria da independência (linha vermelha) e modelo proposto Eq. (3.56) (linha preta).

3.5 CONCLUSÕES

Este capítulo apresentou uma análise do comportamento dos momentos de segunda ordem dos coeficientes para o algoritmo LMS filtrado. O novo modelo não utiliza a teoria da independência, levando em consideração a correlação entre os diferentes vetores do sinal de referência. A partir do apresentado neste capítulo, pode-se observar a importância de não considerar a teoria da independência na determinação do modelo analítico. Sem a utilização desta simplificação, conseguimos estudar o que acontece quando existe um erro na estimação do atraso no caso do algoritmo DLMS, para diferentes tipos de entradas: ruído branco gaussiano e ruído colorido gaussiano. As conclusões que resultaram deste estudo não poderiam ser obtidas à luz da teoria da

independência, tal como foi mostrado na Seção 3.1.8 e na Seção 3.1.14. Foram obtidas as equações recursivas matriciais que governam o comportamento da matriz de correlação dos coeficientes do filtro adaptativo em função da estatística do sinal de entrada, para estimação perfeita do atraso e para o caso de erro de estimação, para o algoritmo DLMS. Mostrou-se, também, que a análise proposta deste algoritmo é uma análise geral, e que os estudos presentes na literatura podem ser considerados como casos particulares. Finalmente, foram obtidas as equações recursivas matriciais para os momentos de segunda ordem para o caso geral. Neste caso, a dificuldade matemática não permitiu avançar nos cálculos visando à obtenção de um limite superior para a taxa de aprendizagem. Através de resultados de simulações foi possível avaliar a qualidade do novo modelo proposto. Na próxima seção são apresentados exemplos onde o comportamento do algoritmo (velocidade de convergência e erro em regime permanente) são observados em relação aos tipos de respostas em frequência que apresentam a planta, o caminho secundário e sua estimação.

Aplicações

4

Neste Capítulo são apresentados alguns exemplos de aplicação onde as respostas ao impulso utilizadas representam situações que podem ser encontradas na prática. Em todos os casos é comparado o modelo teórico proposto com a simulação. O objetivo principal deste capítulo é fornecer um conhecimento do desempenho que pode-se esperar do algoritmo, através da apresentação de uma série de exemplos. O conhecimento que pretendemos abordar consiste em relacionar o número necessário de iterações para atingir o regime permanente e o erro médio quadrático, que o algoritmo fornece nessa condição, com os tipos de resposta em frequência da planta, do caminho secundário e de sua estimação.

Exemplo 4.1

Neste exemplo é utilizada uma planta com uma resposta em frequência do tipo passa baixas, modelada pela seguinte resposta ao impulso:

$$W^o = [0,0061; 0,0343; 0,0954; 0,1671; 0,2; 0,1671; 0,0954; 0,0343; 0,0061]$$

Este tipo de resposta em frequência é usualmente apresentado pelo caminho primário em aplicações de controle ativo de ruído acústico. O caminho secundário apresenta uma característica de resposta em frequência do tipo passa faixa de banda larga, apresentando uma atenuação importante na região de baixas frequências. Esta atenuação é uma característica do transdutor utilizado, que é geralmente um alto-falante, possuindo uma resposta deficiente para frequências muito baixas. Para avaliar o que acontece nesta situação, uma resposta ao impulso do caminho secundário com tais características é dada por:

$$S = [-0,0801; 0,2561; -0,058; 0,2523; -0,4996; -0,333; 0,6396; -0,2824; 0,108]$$

A Fig. 4.1a mostra a relação entre as bandas passantes da planta e do caminho secundário. Nessa figura pode-se observar o mencionado em relação às características das transferências utilizadas. Neste exemplo assumimos estimação perfeita, $\hat{S} = S$. Outros parâmetros utilizados são: taxa de aprendizagem $\mu = 0,005$, sendo o valor limite determinado em forma experimental, para o qual o algoritmo não converge, $\mu_{\max} = 0,075$. O sinal de entrada é um ruído branco gaussiano com variância $\sigma_x^2 = 1$ e o ruído de medição é um ruído branco gaussiano com variância $\sigma_z^2 = 0,0001$.

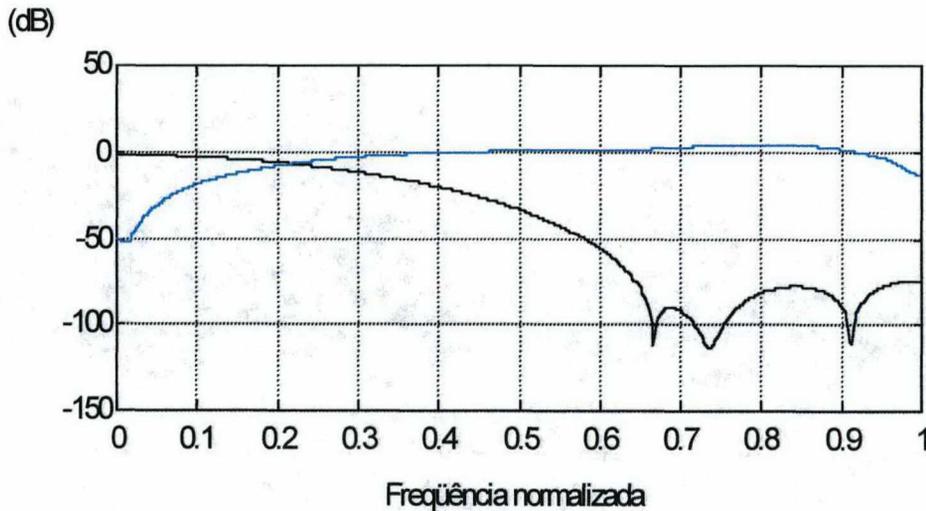


Fig. 4.1a. Resposta em frequência da planta (linha preta) e do caminho secundário (linha azul).

Na Fig. 4.1b é mostrado o comportamento dos coeficientes do filtro adaptativo obtido pela média de 50 realizações, até atingir o regime permanente. Também, nessa figura podemos verificar a expressão (2.21), que prediz o valor em regime permanente dos coeficientes do filtro, para o caso $\hat{S} = S$, sendo dado por:

$$W_{ot} = [-0,0269; 0,0547; 0,185; 0,2922; 0,3306; 0,3085; 0,2491; 0,1687; 0,0856]$$

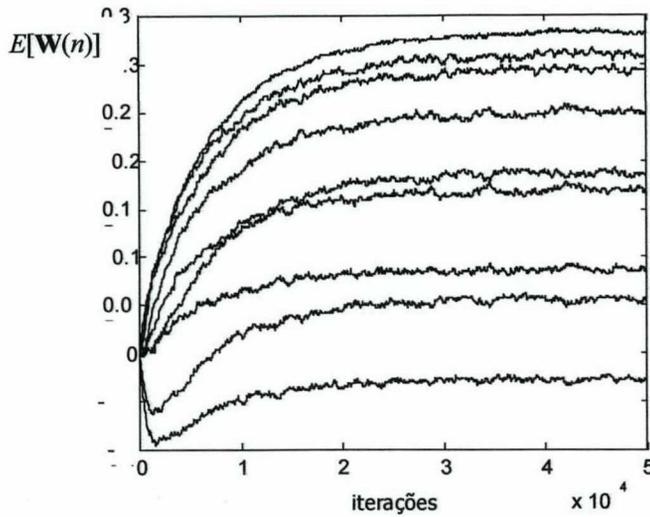


Fig. 4.1b. Valor médio dos coeficientes do filtro adaptativo, obtido através de 50 realizações.

Na Fig. 4.1c é comparado o resultado da simulação com o modelo (Eqs. (2.37)), para as primeiras 12000 iterações. Dessa figura pode-se observar o excelente casamento entre a simulação e o novo modelo proposto. Esta figura também mostra a evolução dos coeficientes do filtro adaptativo utilizando o modelo obtido usando a teoria da independência (Eq. (2.38)). Note que a predição do comportamento dos coeficientes deste modelo está longe do que acontece na prática.

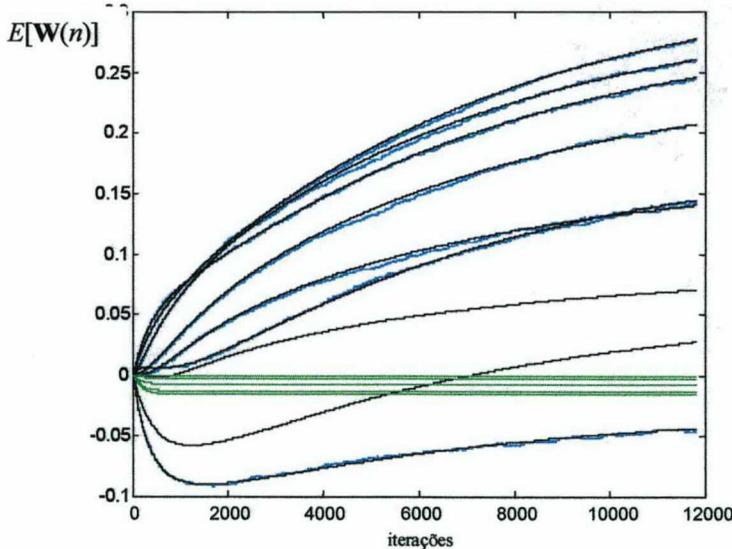


Fig. 4.1c. Detalhe do comportamento transiente dos coeficientes do filtro adaptativo. Simulação (linha azul), modelo Eq. (2.37) (linha preta) e modelo utilizando a teoria da independência Eq. (2.38) (linha verde).

A Fig. 4.1d mostra o erro médio quadrático obtido da simulação e a partir do modelo utilizando as Eqs. (3.51) e (3.56). Também, podemos verificar a Eq. (2.44) que prediz o erro mínimo, sendo erro medido igual a 0,093 e da Eq. (2.44) 0,0935.

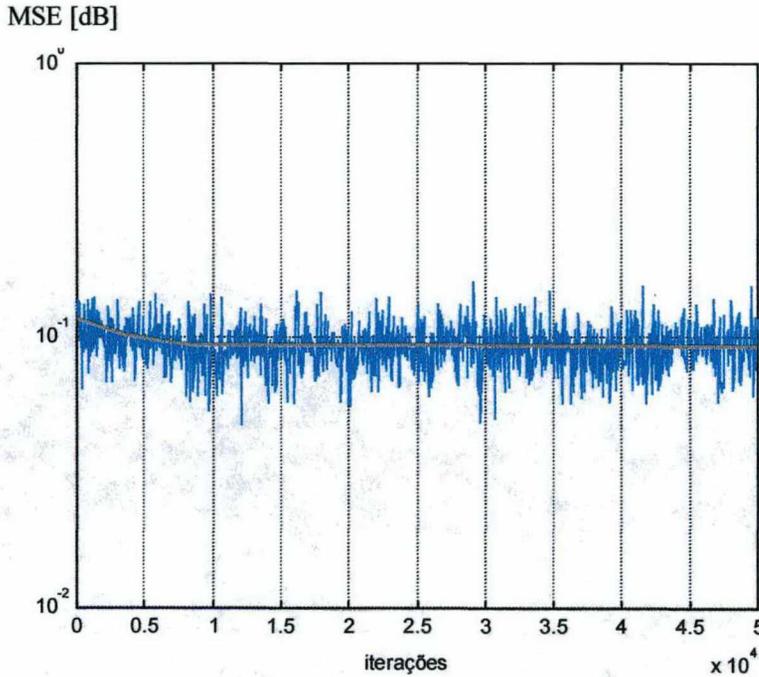


Fig. 4.1d. Erro médio quadrático simulação (linha azul) e modelo Eqs. (3.51) e (3.56) (linha cinza).

A Fig. 4.1e apresenta a resposta em frequência da convolução entre a resposta do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha), junto com as correspondentes à planta (linha preta), ao caminho secundário (linha azul) e do filtro adaptativo em regime permanente (linha verde). Note-se que o caminho de adaptação, dado pela convolução entre o filtro adaptativo e caminho secundário, não consegue modelar a planta na região de frequências baixas, estando limitado pela característica do caminho secundário nessa região. Entretanto, o filtro adaptativo acompanha à planta até o ponto em que a atenuação desta última começa ser importante, a partir de, aproximadamente, o valor de frequência normalizada 0,4.

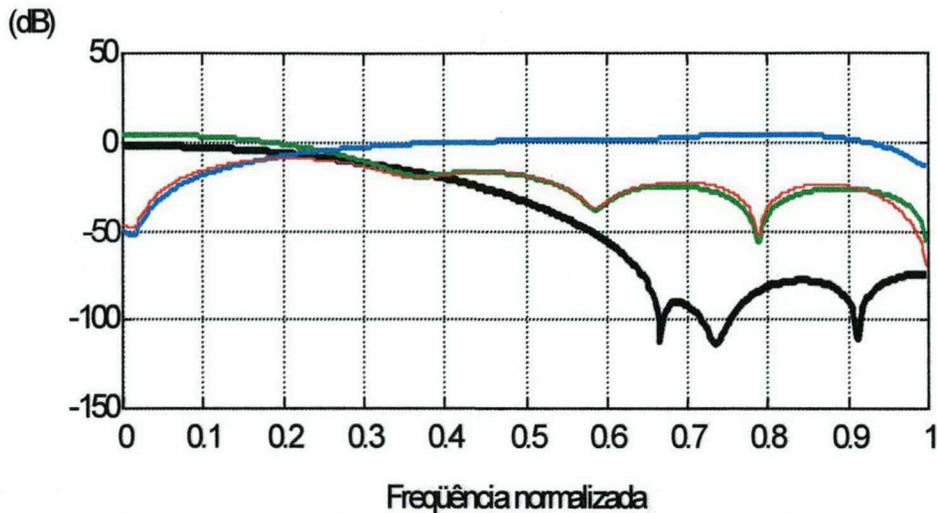


Fig. 4.1e. Respostas em frequência da planta (linha preta), caminho secundário (linha azul), convolução do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha) e o filtro adaptativo (linha verde).

A partir dos resultados obtidos, pode-se inferir que a lenta convergência apresentada e o elevado erro residual estão relacionados com a resposta em frequência dos sistemas envolvidos. A “janela” de frequências onde o algoritmo recebe energia é determinada pela combinação do caminho secundário, na região inferior, e pelo caminho primário na região superior. Esta situação também se reflete no erro médio quadrático em excesso que, como pode-se observar na Fig. 4.1d, possui um valor alto em regime permanente. Nos próximos exemplos vamos ver o que acontece quando se aumenta o valor da taxa de aprendizagem, quando a estimação do caminho secundário não é perfeita e, no exemplo 4.4, vamos ver o comportamento do algoritmo quando o caminho secundário não impõe restrições na faixa de frequências do caminho primário, ou seja, quando aumentamos o tamanho da janela de frequências.

Exemplo 4.2

Neste exemplo vamos investigar a influencia da taxa de aprendizagem no comportamento do algoritmo e, também, avaliar o modelo proposto para outro valor de μ . Para isto, são utilizadas as mesmas condições que no Exemplo 4.1, sendo a única diferença para este caso o valor da taxa de aprendizagem, sendo dada por $\mu = 0,03$, o que representa o 40% do valor máximo. Em relação ao valor anteriormente utilizado, o novo valor é 6 vezes

maior portanto é nosso interesse avaliar a Eq. (2.44) com o objetivo de determinar se a predição do erro continua valendo.

Dado que a planta, o caminho secundário e sua estimação são as mesmas que no Exemplo 4.1, a figura correspondente às respostas em frequência é a Fig. 4.1a. Na Fig. 4.2a é comparado o resultado da simulação com o modelo (Eqs. (2.37)). Dessa figura pode-se observar o excelente casamento entre a simulação e o novo modelo proposto, para os diferentes valores de μ utilizados. A Fig.4.2a, também mostra a evolução dos coeficientes do filtro adaptativo utilizando o modelo obtido da teoria da independência Eq. (2.38). Note que a predição do comportamento dos coeficientes deste modelo está longe do que acontece na prática. Também podemos verificar a Eq. expressão (2.43), que prediz o valor em regime permanente dos coeficientes do filtro, sendo dado por:

$$\mathbf{W}_{ot} = [-0,0269; 0,0547; 0,185; 0,2922; 0,3306; 0,3085; 0,2491; 0,1687; 0,0856]$$

O valor médio dos coeficientes em regime permanente, obviamente, é o mesmo que no exemplo 4.1. No entanto, da mesma forma que no algoritmo LMS clássico, as flutuações do valor do coeficientes ao redor do valor médio são maiores na medida em que a taxa de aprendizagem aumenta. Portanto, neste exemplo foi efetuada a média de 100 realizações a fim de obter as curvas, obtidas da simulação, com a menor flutuação possível. Em função do valor de μ observamos que os coeficientes atingem o regime permanente, aproximadamente, na iteração 6000; sendo que antes deviam transcorrer aproximadamente 50000 iterações para o regime permanente.

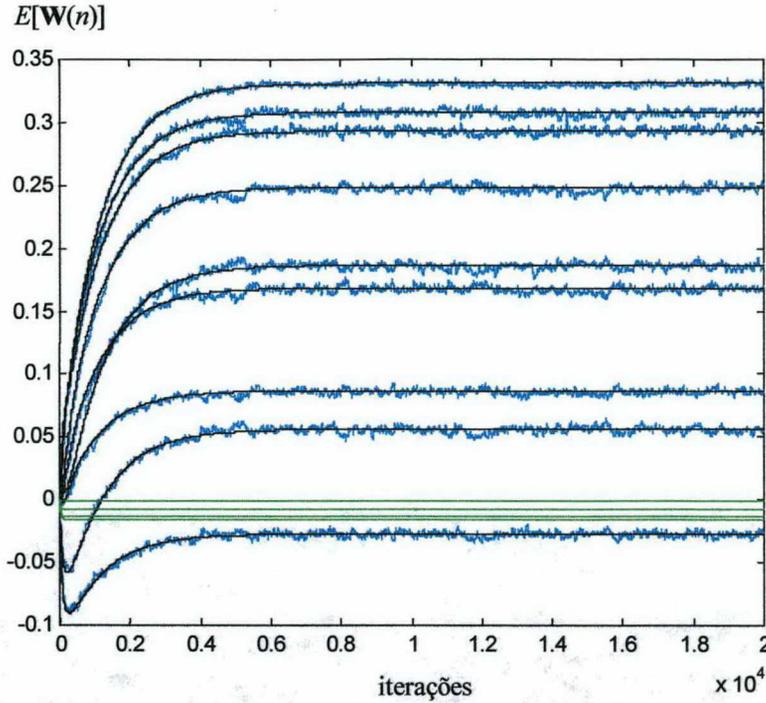


Fig. 4.2a. Evolução dos coeficientes do filtro adaptativo. Simulação (linha azul), modelo Eq. (2.37) (linha preta) e modelo utilizando a teoria da independência Eq. (2.38) (linha verde).

A Fig. 4.2b mostra o erro médio quadrático obtido da simulação e a partir do modelo utilizando as Eqs. (3.51) e (3.56). Também, verificamos que a predição do erro mínimo da Eq. (2.44) começa a se afastar do valor real, na medida em que o valor de μ aumenta. Isso é esperado, já (2.44) foi obtida para valores pequenos da taxa de aprendizagem, sendo erro o medido 0,0941 e o dado pela Eq. (2.44) igual a 0,0935. Note-se que o EMQ medido, em regime permanente, aumentou em relação ao caso do exemplo anterior. Este fato está de acordo com o que também acontece com o LMS clássico, ou seja, um aumento na taxa de aprendizagem produz um aumento do erro médio quadrático residual.

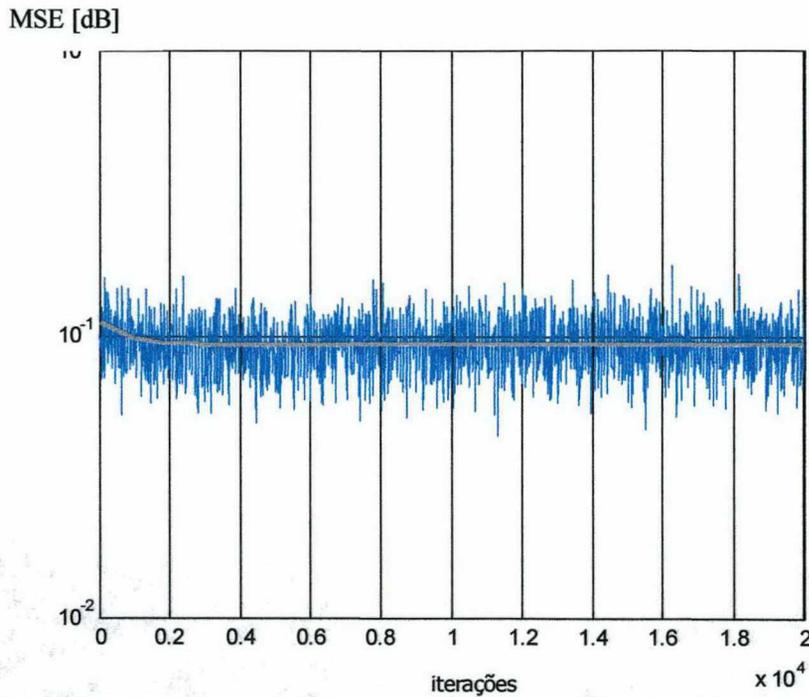


Fig. 4.2b. Erro médio quadrático simulação (linha azul) e modelo (linha cinza) Eqs. (3.51) e (3.56).

Em relação à suposição (II)-(seção 2.4), tal como verificamos no capítulo 2, para o valor da taxa de aprendizagem utilizada que pode ser considerada alta, não registrou-se nenhum efeito apreciável no modelo proposto, no sentido de que o mesmo continua acompanhando a simulação. Tendo visto o que acontece quando se aumenta o valor da taxa de aprendizagem, no próximo exemplo vamos analisar o efeito dos erros na estimação do caminho secundário.

Exemplo 4.3

Neste exemplo vamos investigar a o que acontece quando a estimação do caminho secundário não é perfeita. Para isto, são utilizadas as mesmas condições que no exemplo 4.1 e 4.2. Neste caso, a diferença reside em que vamos admitir erro na estimação do caminho secundário. A fim de comparar as respostas impulsivas do caminho secundário e sua estimação, as mesmas são colocadas a seguir:

$$S = [-0,0801; 0,2561; -0,058; 0,2523; -0,4996; -0,333; 0,6396; -0,2824; 0,108]$$

$$\hat{S} = [-0,1; 0,23; 0,0; 0,27; -0,45; -0,4; 0,6; -0,3; 0,0]$$

Os outros parâmetros são: taxa de aprendizagem $\mu = 0,03$, para este caso o valor máximo, determinado experimentalmente, continua sendo 0,075. Como sinal de entrada utilizamos um ruído branco e gaussiano com variância $\sigma_x^2 = 1$, a variância do ruído branco gaussiano aditivo é $\sigma_z^2 = 0,0001$. A Fig. 4.3a mostra a relação entre as bandas passantes da planta, o caminho secundário e sua estimação. Nessa figura pode-se observar o mencionado em relação às características das transferências utilizadas.

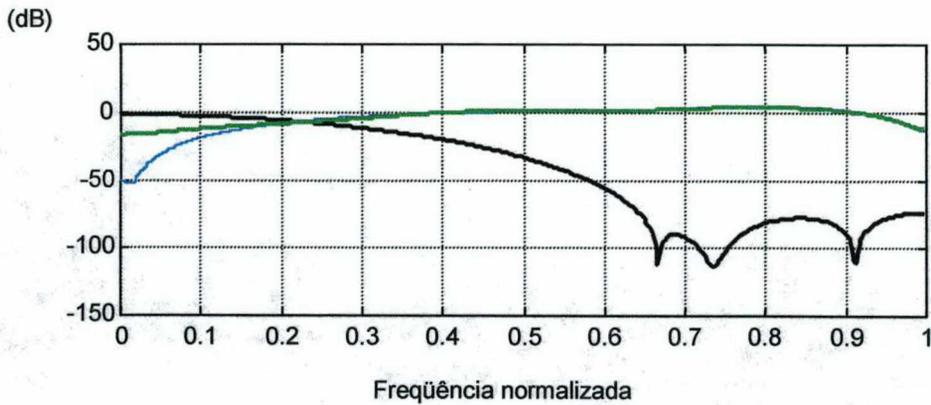


Fig. 4.3a. Resposta em frequência da planta (linha preta), do caminho secundário (linha azul) e estimação (linha verde).

Na Fig. 4.3b é comparado o resultado da simulação, através da média de 100 realizações, com o modelo (Eqs. (2.37)). Dessa figura pode-se observar o excelente casamento entre a simulação e o novo modelo proposto para a situação analisada. A Fig.4.3b, também mostra a evolução do valor médio dos coeficientes do filtro adaptativo utilizando o modelo obtido usando a teoria da independência (Eq. (2.38)). Note que a predição do comportamento dos coeficientes deste modelo está longe de prever o real comportamento dos coeficientes. Também podemos verificar a expressão (2.43), que prediz o valor em regime permanente dos coeficientes do filtro, para o caso $\hat{S} \neq S$, sendo dado por:

$$W_{\infty} = [-0,0468; 0,0193; 0,1527; 0,272; 0,3318; 0,330; 0,285; 0,204; 0,1128]$$

Em função do erro de estimação, note que os valores de regime permanente são diferentes em relação aos obtidos nos exemplos anteriormente apresentados. Também, pode-se perceber o menor número de iterações necessárias para entrar em regime permanente. Este fato, que pode ser considerado como sendo interessante, depende da qualidade da estimação.

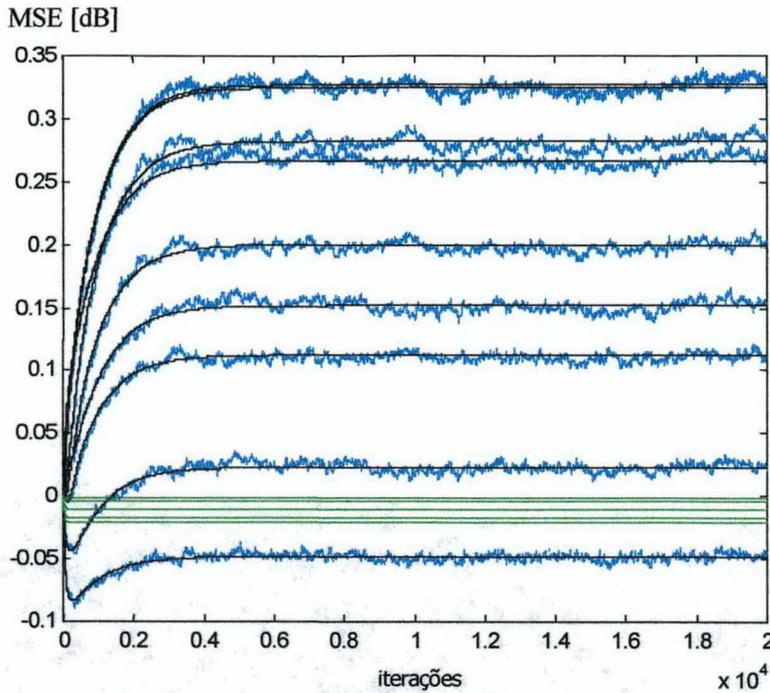


Fig. 4.3b. Detalhe do comportamento transiente dos coeficientes do filtro adaptativo. Simulação (linha azul), modelo Eq. (2.37) (linha preta) e modelo utilizando a teoria da independência Eq. (2.38) (linha verde).

A Fig. 4.3c mostra o erro médio quadrático obtido da simulação e a partir do modelo utilizando as Eqs. (3.51) e (3.56). Também verificamos que a predição do erro mínimo da Eq. (2.44). Para este caso, temos que o erro medido é 0,0964 e, a partir da Eq. (2.44) obtemos 0,0953. Comparando com o exemplo 4.2, vemos que o EMQ em regime permanente aumentou devido a dois fatores: a falta de casamento entre o caminho secundário e sua estimação, e ao μ mais elevado.

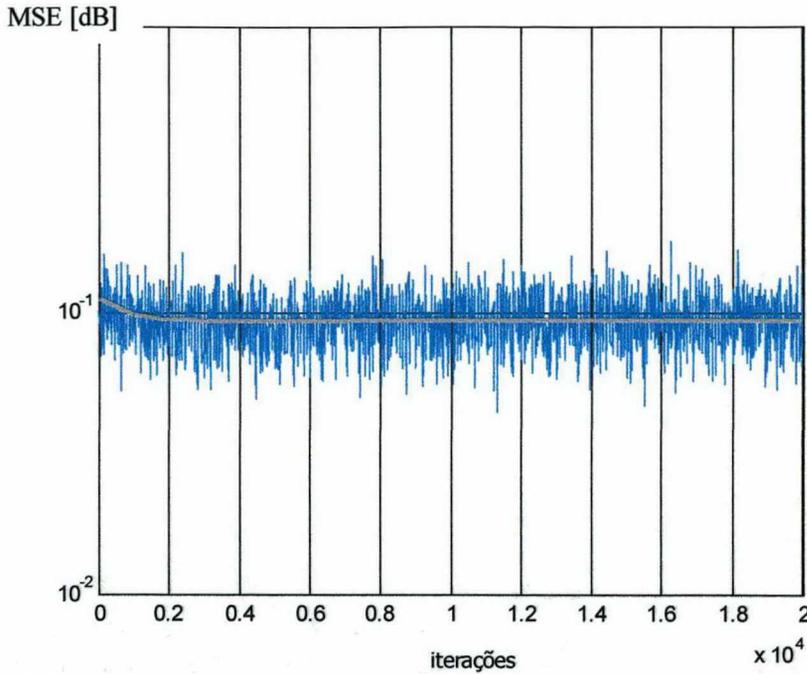


Fig. 4.3c. Erro médio quadrático simulação (linha azul) e modelo (linha cinza) Eqs. (3.51) e (3.56).

A Fig. 4.3d apresenta a resposta em frequência da convolução entre a resposta do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha), junto com as correspondentes à planta (linha grossa), ao caminho secundário (linha azul), à estimação (linha verde) e ao filtro adaptativo (linha verde clara). Note que o caminho de adaptação está sempre limitado pelo caminho secundário. O motivo da lenta convergência e do erro elevado reside na faixa estreita de frequências, ao redor da frequência 0,2, onde caminho de adaptação permite passar a máxima energia.

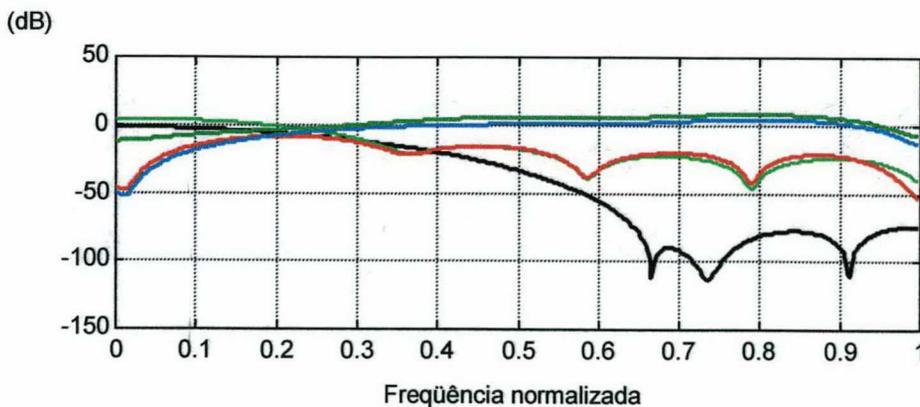


Fig. 4.3d Respostas em frequência da planta (linha preta), caminho secundário (linha azul), estimação (linha verde), convolução do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha) e o filtro adaptativo (linha verde clara).

Verifica-se que o tipo de resposta em frequência do caminho secundário tem forte influência tanto na velocidade de convergência quanto no EMQ mínimo que o algoritmo fornece. Isso será comprovado, no próximo exemplo vamos comprovar este fato.

Exemplo 4.4

Para investigar a influência do caminho secundário, neste caso usamos uma característica de resposta em frequência do tipo passa baixa menos restritiva que a da planta. A resposta ao impulso dada por é $S = [1; 0,5]$. A planta para este exemplo é a mesma que vem sendo utilizada nos exemplos anteriormente apresentados. Neste exemplo assumimos estimação perfeita, $\hat{S} = S$. Os outros parâmetros utilizados são: taxa de aprendizagem $\mu = 0,005$. O valor para o qual o algoritmo não converge é 0,08, determinado de forma experimental. O sinal de entrada é um ruído branco gaussiano com variância $\sigma_x^2 = 1$, a variância do ruído branco gaussiano aditivo $\sigma_z^2 = 0,0001$. Estes parâmetros são iguais aos do exemplo 4.1 a fim de deixar como única variável o caminho secundário. A Fig. 4.1a mostra a relação entre as bandas passantes da planta e a do caminho secundário. Nessa figura pode-se observar que o caminho secundário não impõe restrição alguma para a banda passante do caminho primário como nos exemplos anteriores.

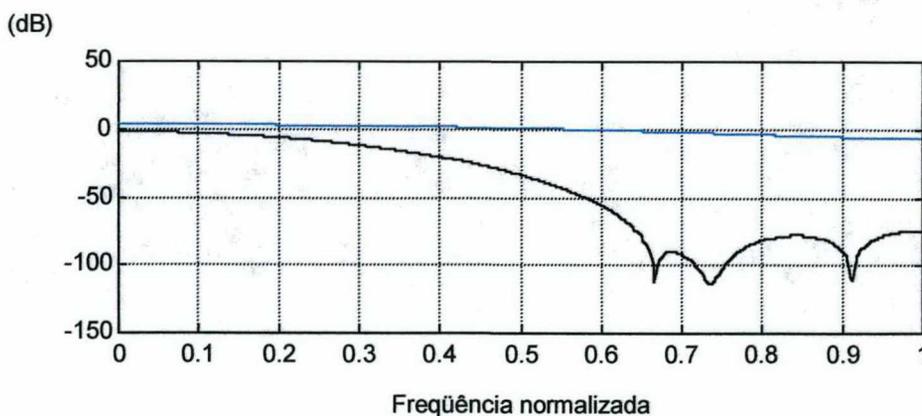


Fig. 4.4a. Resposta em frequência da planta (linha preta) e do caminho secundário (linha azul).

Na Fig. 4.4b é mostrada a evolução do valor médio dos coeficientes do filtro adaptativo obtidos a partir da simulação (linha azul), através da média de 10 realizações, e do

modelo (linha preta) dado pela Eq. (2.37). Dessa figura pode-se observar o excelente casamento entre a simulação e o novo modelo proposto. Também nessa figura podemos verificar a expressão (2.21), que prediz o valor em regime permanente dos coeficientes do filtro para o caso $\hat{S} = S$, sendo dado por:

$$\mathbf{W}_{ot} = [0,0062; 0,0311; 0,0798; 0,1273; 0,1364; 0,099; 0,0459; 0,0114; 0,0002]$$

A Fig.4.4b, também mostra a evolução valor médio dos coeficientes do filtro adaptativo utilizando o modelo obtido da teoria da independência. Note que a predição do comportamento dos coeficientes deste modelo não reflete o comportamento real dos mesmos. Pode-se observar que mesmo para um valor de μ pequeno, como no exemplo 4.1, o regime permanente começa a partir de, aproximadamente, $n=500$, lembrando que no primeiro exemplo a convergência ocorria após 50000 iterações.

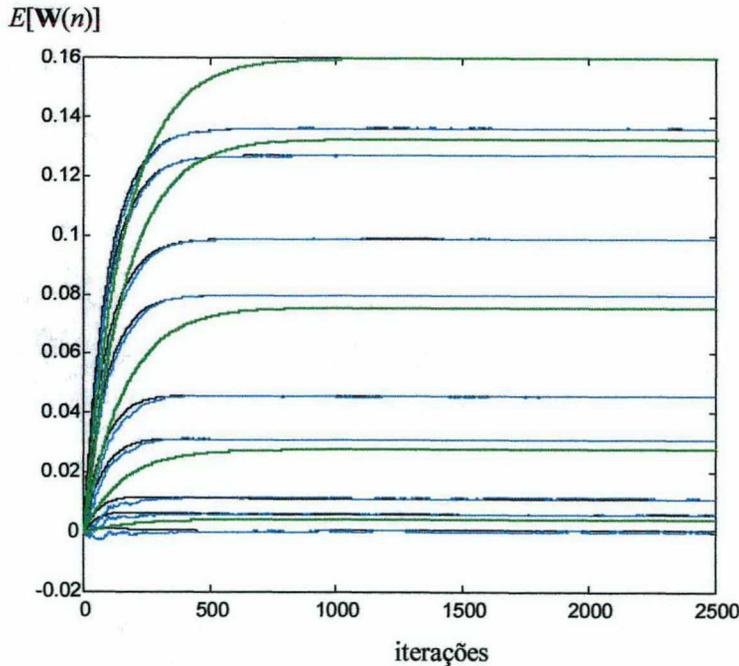


Fig. 4.4b. Comportamento dos coeficientes do filtro adaptativo. Simulação (linha azul), modelo Eq. (2.37) (linha preta) e modelo utilizando a teoria da independência Eq. (2.38) (linha verde).

A Fig. 4.4c mostra o erro médio quadrático obtido da simulação e a partir do modelo utilizando as Eqs. (3.51) e (3.56). Também verificamos a Eq. (2.44) que prediz o erro mínimo, sendo o erro medido $1,0314 \cdot 10^{-4}$ e da Eq. (2.44) temos $1,029 \cdot 10^{-4}$. Comparando com os exemplos antes apresentados, notamos o menor número de iterações necessárias para atingir o regime permanente, assim como o menor erro em regime permanente que o algoritmo fornece. Este fato é devido à característica do caminho secundário, já que o mesmo não atenua o sinal de entrada para nenhuma frequência contida na banda passante do caminho primário.

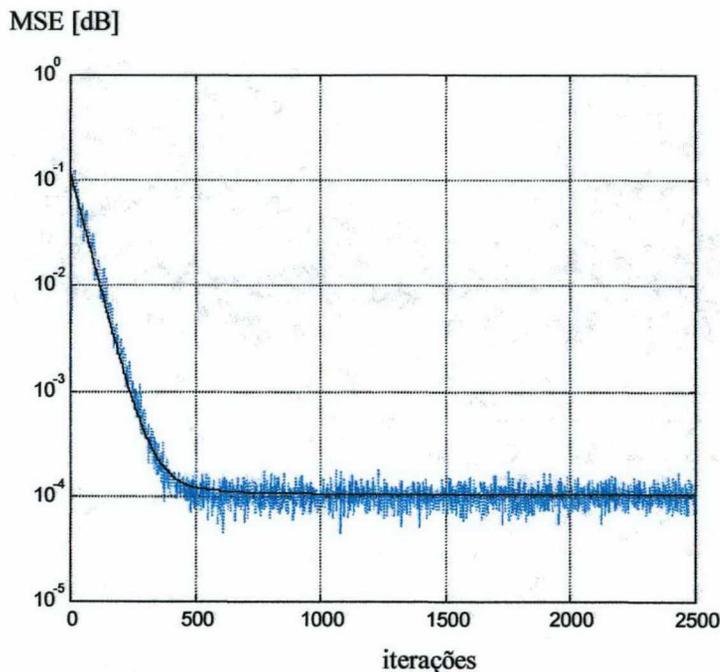


Fig. 4.4c. Erro médio quadrático simulação (linha azul) e modelo (linha preta) Eqs. (3.51) e (3.56).

Na Fig. 4.4d são apresentadas as respostas em frequência dadas pela convolução entre a resposta do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha), da planta (linha preta), o caminho secundário (linha azul) e do filtro adaptativo (linha verde). Pelo fato que o caminho secundário tem uma banda passante que inclui a do caminho primário, o resultado é que o caminho de adaptação consegue modelar muito bem a planta para quase toda a faixa de frequências na banda passante do caminho primário. Nos exemplos anteriores, o caminho secundário atuava como fator limitante.

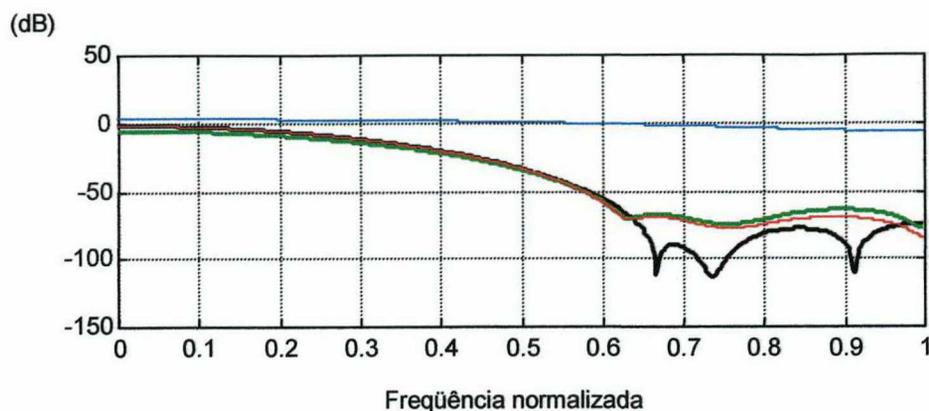


Fig. 4.4d. Respostas em frequência da planta (linha preta), caminho secundário (linha azul), convolução do filtro adaptativo em regime permanente e o caminho secundário (linha vermelha) e o filtro adaptativo (linha verde).

A partir dos exemplos até aqui apresentados, podemos ver que as relações entre as bandas passantes da planta, do caminho secundário e de sua estimação influenciam fortemente o comportamento do algoritmo quanto à velocidade de convergência e ao EMQ mínimo. Para o próximo exemplo vamos investigar o que acontece com o modelo quando são utilizados valores de μ perto do limite superior, em função do que tínhamos observado em relação à suposição (II) no capítulo 2 para estes valores de μ .

Exemplo 4.5

Neste exemplo é investigada a utilização de um valor da taxa de aprendizagem perto do valor limite, para o qual o algoritmo não converge. Para isto foram utilizados os dados do exemplo 4.4. A única diferença é, então, o valor da taxa de aprendizagem, sendo seu valor $\mu = 0,07$. O valor selecionado μ corresponde ao 87% do valor para o qual o algoritmo não converge.

A figura correspondente às respostas em frequência é a Fig. 4.4a, já que as respostas ao impulso utilizadas não foram alteradas. Na Fig. 4.5a é mostrada a evolução do valor médio dos coeficientes do filtro adaptativo obtidos a partir da simulação (linha azul), através da média de 500 realizações, e do modelo (linha preta) dado pela Eq. (2.37). Dessa figura pode-se observar o excelente casamento entre a simulação e o novo modelo proposto mesmo para o elevado valor de μ utilizado. Também, nessa figura podemos verificar a expressão (2.21), que prediz o valor em regime permanente dos coeficientes do filtro para o caso $\hat{S} = S$, sendo dado por:

$$\mathbf{W}_{ot} = [0,0062; 0,0311; 0,0798; 0,1273; 0,1364; 0,099; 0,0459; 0,0114; 0,0002]$$

A Fig.4.4b, também mostra a evolução dos coeficientes do filtro adaptativo utilizando o modelo obtido da teoria da independência. Note que a predição do comportamento dos coeficientes deste modelo não reflete o comportamento real.

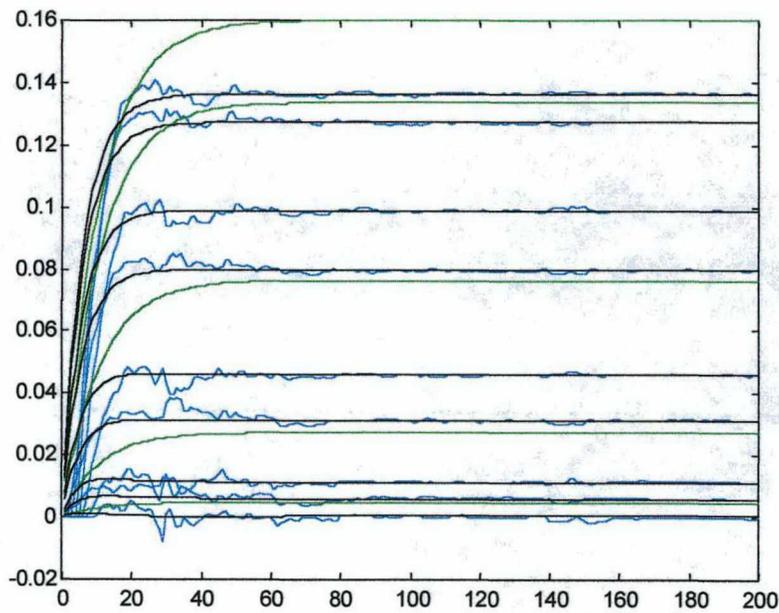


Fig. 4.5a. Comportamento médio dos coeficientes do filtro adaptativo. Simulação (linha azul), modelo Eq. (2.37) (linha preta) e modelo utilizando a teoria da independência Eq. (2.38) (linha verde).

A Fig. 4.5b mostra o erro médio quadrático obtido da simulação e a partir do modelo utilizando as Eqs. (3.51) e (3.56). Também verificamos a Eq. (2.44) que prediz o erro mínimo, sendo o erro medido $5,38 \cdot 10^{-4}$ e da Eq. (2.44) $1,029 \cdot 10^{-4}$. Tal como devíamos esperar, a predição afasta-se do valor real na medida em que o valor da taxa de aprendizagem assume valores significativos, tomando como referência o valor para o qual o algoritmo não apresenta convergência.

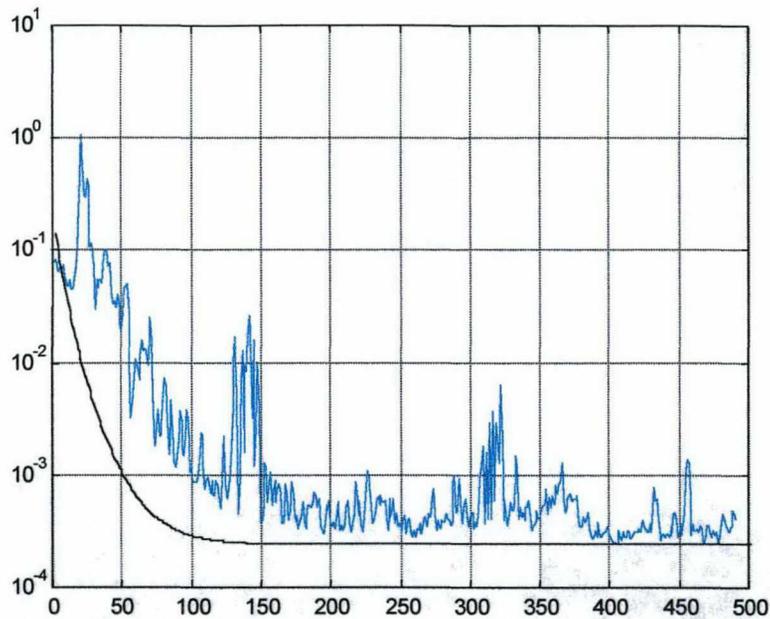


Fig. 4.5b. Erro médio quadrático simulação (linha azul) e modelo (linha preta) Eqs. (3.51) e (3.56).

Neste exemplo pode-se comprovar que as suposições realizadas para desenvolver o novo modelo funcionam muito bem para o momento de primeira ordem dos coeficientes do filtro adaptativo, mesmo para valores elevados de μ . O mesmo não acontece, entretanto, para o momento de segunda ordem, onde a violação da hipótese faz com que o modelo para este momento não acompanhe a simulação. Entretanto, em nenhuma aplicação prática é utilizada uma taxa de aprendizagem de valor próximo ao limite de estabilidade. Portanto, o modelo proposto pode ser utilizado para prever o comportamento do algoritmo LMS filtrado para as aplicações práticas usuais.

Conclusões 5

O objetivo deste trabalho foi o estudo do comportamento estatístico do algoritmo LMS filtrado. Este estudo levou à determinação de modelos analíticos para o comportamento dos momentos de primeira e de segunda ordem dos coeficientes do filtro adaptativo. Estes modelos permitem descrever e prever o comportamento deste algoritmo, tanto no regime transiente quanto para o regime permanente. Também foi estudada a influência do tipo de resposta da planta e do caminho secundário exercem sobre o desempenho do algoritmo. Esta influência não é descrita na literatura técnica existente. Outro problema tratado foi o efeito dos erros de estimação do caminho secundário sobre o comportamento do sistema. Quando o tratamento matemático permitiu, foi obtido um limite superior para a taxa de convergência. Este foi o caso do algoritmo DLMS.

Os modelos matemáticos desenvolvidos nesta tese foram obtidos sem invocar a suposição usual, empregada nas análises existentes na literatura, da teoria da independência. Isto representa uma novidade na análise de um problema que utiliza uma variante do algoritmo LMS. Ao longo deste trabalho, demonstrou-se que a teoria da independência não produz resultados satisfatórios na análise do algoritmo FXLMS. Como o caminho secundário correlaciona as amostras oriundas do sinal de entrada, tais correlações devem ser consideradas. Como além da distribuição gaussiana (no caso dos momentos de segunda ordem) não foi feita nenhuma suposição simplificadora em relação ao tipo de sinal de entrada, os modelos desenvolvidos são válidos para ruído branco ou ruído colorido gaussiano.

Em relação ao algoritmo DLMS, foi possível refutar, através do modelo proposto, o fato que o algoritmo DLMS não converge se a estimação do atraso não for perfeita. Com o novo modelo pode-se verificar que o algoritmo não converge para uma entrada sendo ruído branco gaussiano. Até o presente trabalho, este caso somente podia ser verificado de forma experimental. Para uma entrada colorida e gaussiana demonstrou-se que o algoritmo também converge na condição de estimação imperfeita do atraso. Na literatura

não é relatado este comportamento, já que os modelos existentes não permitem predizê-lo. Também, foi demonstrado que os novos modelos correspondentes ao momento de primeira e segunda ordem dos coeficientes do filtro adaptativo são mais gerais que os que se encontram na literatura. Adicionalmente, as equações teóricas são expressas de maneira mais compacta.

Como proposta para futuros trabalhos pode-se citar, como ponto principal, o estudo da estabilidade para o caso geral e para o caso do algoritmo DLMS com erro na estimação do atraso. Outra possibilidade de trabalho futuro seria uma tentativa do aproveitamento dos modelos aqui propostos no desenvolvimento de novos algoritmos adaptativos para o controle ativo de ruído ou vibrações.

Referências Bibliográficas

6

- [1] P. Lueg, "Process of silencing Sound Oscillations," Patente número: 2.043.416, U.S., 1936.
- [2] S.J. Elliott e P.A. Nelson, "Active control noise," *IEEE Signal Proc. Mag.*, pp. 12-35, out. 1993.
- [3] P.A. Nelson e S.J. Elliott, *Active control of sound*, Academic Press, 1990.
- [4] Sen M. Kuo e Dennis R. Morgan, *Active noise control systems*, John Wiley & Sons, 1996.
- [5] B. Widrow and S. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1985.
- [6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1986.
- [7] L. Horowitz e K. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Acoust. Speech, Signal Proc.*, vol-29, pp. 722-736, jun. 1981.
- [8] A. Feuer e E. Weinstein, "Convergence analysis of LMS filters with uncorrelated gaussian data," *IEEE Acoust. Speech, Signal Proc.*, vol-33, pp. 222-230, fev.1985.
- [9] G. Long, F. Ling e J. Proakis, "The LMS algorithm with delayed coefficient adaptation," *IEEE Acoust. Speech, Signal Proc.*, vol-37, pp. 1397-1405, set. 1989.
- [10] A. Feuer e E. Weinstein, "The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes," *IEEE Trans. on Signal Proc.*, Vol. 41, No. 3, pp. 1149-1160, mar. 1993.
- [11] J. M. Bermudez e N. Bershad, "A nonlinear analytical model for the quantized LMS algorithms-The arbitrary step size case," *IEEE Trans. on Signal Proc.*, Vol. 44, No. 5, pp. 1175-1183, mai. 1996.
- [12] K. Mayyas e T. Aboulnasr, "Leaky LMS algorithms: MSE analysis for gaussian data," *IEEE Trans. on Signal Proc.*, Vol. 45, No. 4, pp. 927-934, abr. 1997.
- [13] W. Wu e P. Chen, "Adaptive AR modelling in white gaussian noise," *IEEE Trans. on Signal Proc.*, Vol. 45, No. 5, pp. 1184-1192, mai. 1997.
- [14] D. R. Morgan, "An analysis of multiple correlation cancellation loop with a filter in the auxiliary path," *IEEE Acoust. Speech, Signal Proc.*, vol-28, pp. 454-467, 1980.
- [15] J. C. Burgess, "Active adaptive sound control in a duct: A computer simulation," *J. Acoust. Soc. Am.*, Vol. 70, pp. 715-726, 1981.
- [16] E. Bjarnanson, "Analysis of the filtered-X LMS algorithm", *IEEE Trans. On Speech and Audio Processing*, Vol. 3, Nro. 6, Nov. 1995.
- [17] A. Gersho, "Adaptive equalization of highly dispersive channels for data

- transmission," *Bell Syst. Tech. J.*, Vol. 48, pp. 55-70, Jan. 1969.
- [18] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, Vol. 58, pp. 963-993, Maio-Junio. 1979.
- [19] S. Douglas e W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. on Signal Proc.*, Vol. 43, No. 12, Dec. 1995.
- [20] E. Kreyszig, *Advanced Engineering Mathematics*, John Wiley & Sons, pp. 1007-1009, Sétima edição, 1993.
- [21] O. Macchi, *Adaptive Processing*, John Wiley & Sons, 1995.
- [22] S. Snyder e C. Hansen, "The effect of transfer function estimation errors on the filtered-X LMS algorithm," *IEEE Trans. on Signal Proc.*, Vol. 42, No. 4, Abr. 1994.
- [23] S. M. Ahn, "Stability of a matrix polynomial in discrete systems," *IEEE Trans. on Automat. Control*, Vol. AC-27, No. 5, Out. 1982.
- [24] Khanh T. Ngo e Kelvin T. Erickson, "Stability of discrete-time matrix polynomials," *IEEE Trans. on Automat. Control*, Vol. 42, No. 4, Abr. 1997.
- [25] P.A. Nelson e S.J. Elliott, *Active control of vibrations*, Academic Press, 1990.
- [26] M. Rupp e R. Frenzel, "Analysis of LMS and NLMS algorithms with delayed coefficient update under the presence of spherically invariant processes," *IEEE Trans. on Signal Proc.*, Vol. 42, No. 3, Mar. 1994.
- [27] S.J. Elliott, I.M. Stothers and P.A. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, pp. 1423-1434, October 1987.
- [28] C. C. Boucher. S. J. Elliott and P. A. Nelson, "Effect of errors in the plant model on the performance of algorithms for adaptive feedforward control," *IEE Proceedings-F: Radar Signal Processing*, 138, pp. 313-319, August 1991.
- [29] G. Long, F. Ling e J. Proakis, Corrections to "The LMS algorithm with delayed coefficient adaptation," *IEEE Trans. on Signal Proc.*, vol-40, no. 1, pp. 230-232, jan. 1992.
- [30] O.J. Tobias, J.C. Bermudez, N. Bershad e R. Seara, "Mean weight behavior of the Filtered-X LMS algorithm," *Proc. ICASSP'98*, pp. 3545-3548, 1998.
- [31] R. Deutsch, *Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1965.
- [32] P. Kabal, "The stability of adaptive minimum mean square error equalizers using delayed adjustment," *IEEE Trans. on Comm.*, vol. 31, no. 3, pp. 430-432, mar. 1983.
- [33] O.J. Tobias, J.C. Bermudez, N. Bershad e R. Seara, "Second moment analysis of the Filtered-X LMS algorithm," *Proc. ICASSP'99*, pp. 1873-1876, 1999.
- [34] O.J. Tobias, J.C. Bermudez, N. Bershad e R. Seara, "New Results and Insights for the Filtered-X LMS algorithm," *Proc. ASILOMAR*, 1998.
- [35] W.A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: a general study, analysis, and critique," *Signal Proc.*, vol. 6, pp. 113-133, 1984.

Apêndice A

OBTENÇÃO DE (2.54)

Substituindo (2.52) em (2.9) e expandindo os termos em função de $\mathbf{V}(n)$ e $\mathbf{X}_{\hat{f}}(n)$, tem-se:

$$E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i \hat{s}_j E[\mathbf{X}(n-j) \mathbf{X}^T(n-i) \mathbf{V}(n-i)] + \mu E[e_o(n) \mathbf{X}_{\hat{f}}(n)]. \quad (\text{a.1})$$

Utilizando (2.2), (2.22) e (2.53), o último termo de (a.1) pode ser escrito como

$$E[e_o(n) \mathbf{X}_{\hat{f}}(n)] = \tilde{\mathbf{P}}_{\hat{s}} - \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{P}}_s. \quad (\text{a.2})$$

A Eq. (a.2) é zero se e somente se $\hat{\mathbf{S}} = \mathbf{S}$. Para $\hat{s}_i = s_i + \delta_i$, $i = 0, \dots, M$, $\mathbf{X}_{\hat{f}}(n)$ pode ser colocado da seguinte forma:

$$\mathbf{X}_{\hat{f}}(n) = \sum_{i=0}^{M-1} s_i \mathbf{X}(n-i) + \sum_{i=0}^{M-1} \delta_i \mathbf{X}(n-i) = \mathbf{X}_f(n) + \mathbf{X}_{\delta}(n). \quad (\text{a.3})$$

Tomando o valor esperado em (a.3) e aplicando o princípio da ortogonalidade, obtém-se:

$$E[e_o(n)\mathbf{X}_f(n)] = E[e_o(n)\mathbf{X}_f(n)] + E[e_o(n)\mathbf{X}_\delta(n)] = E[e_o(n)\mathbf{X}_\delta(n)]. \quad (\text{a.4})$$

Logo, desconsiderando as correlações entre $\mathbf{V}(n-i)$ e $\mathbf{X}(n-j)$, Eq. (a.1) é rescrita da seguinte forma:

$$\begin{aligned} E[\mathbf{V}(n+1)] = E[\mathbf{V}(n)] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i s_j \mathbf{R}_{i-j} E[\mathbf{V}(n-i)] \\ + \mu E[e_o(n)\mathbf{X}_\delta(n)] - \mu \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} s_i \delta_j \mathbf{R}_{i-j} E[\mathbf{V}(n-i)]. \end{aligned} \quad (\text{a.5})$$

Apêndice B

MATRIZ DE CORRELAÇÃO

A Equação (3.54), dada pelo valor esperado $E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)]$, é obtida da seguinte forma. Partindo da expressão recursiva para o cálculo de $\mathbf{V}(n)$, Eq. (2.47), temos:

$$\begin{aligned} \mathbf{V}(n-j) = & \mathbf{V}(n-j-1) + \mu \left(\sum_{k=0}^{\hat{M}-1} \hat{s}_k \mathbf{X}(n-k-j-1) d(n-j-1) \right. \\ & - \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{X}(n-l-j-1) \mathbf{X}^T(n-k-j-1) \mathbf{W}_\infty \\ & - \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{X}(n-l-j-1) \mathbf{X}^T(n-k-j-1) \mathbf{V}(n-k-j-1) \\ & \left. + \sum_{l=0}^{\hat{M}-1} \hat{s}_l \mathbf{X}(n-l-j-1) z(n-j-1) \right), \end{aligned} \quad (\text{b.1})$$

e,

$$\begin{aligned} \mathbf{V}^T(n-i) = & \mathbf{V}^T(n-i-1) + \mu \left(\sum_{k=0}^{\hat{M}-1} \hat{s}_k d(n-i-1) \mathbf{X}^T(n-k-i-1) \right. \\ & - \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{W}_\infty^T \mathbf{X}(n-k-i-1) \mathbf{X}^T(n-l-i-1) \\ & - \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{V}^T(n-k-i-1) \mathbf{X}(n-k-i-1) \mathbf{X}^T(n-l-i-1) \\ & \left. + \sum_{l=0}^{\hat{M}-1} \hat{s}_l \mathbf{X}^T(n-l-i-1) z(n-i-1) \right). \end{aligned} \quad (\text{b.2})$$

Tendo as duas expressões, o próximo passo é calcular o produto $\mathbf{V}(n-j)\mathbf{V}^T(n-i)$ e tomar o valor esperado. Então:

$$\begin{aligned}
 & E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)] = \\
 & (1) E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-i-1)] \\
 & (2) + \mu \sum_{k=0}^{\hat{M}-1} \hat{s}_k E[\mathbf{V}(n-j-1)d(n-i-1)\mathbf{X}^T(n-k-i-1)] \\
 & (3) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l E[\mathbf{V}(n-j-1)\mathbf{W}_x^T \mathbf{X}(n-k-i-1)\mathbf{X}^T(n-l-i-1)] \\
 & (4) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-k-i-1)\mathbf{X}(n-k-i-1)\mathbf{X}^T(n-l-i-1)] \\
 & (5) + \mu \sum_{k=0}^{\hat{M}-1} \hat{s}_k E[\mathbf{X}(n-j-k-1)d(n-j-1)\mathbf{V}^T(n-i-1)] \\
 & (6) + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l E[\mathbf{X}(n-j-k-1)d(n-j-1)d(n-i-1)\mathbf{X}^T(n-l-i-1)] \\
 & (7) - \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{M-1} \sum_{m=0}^{\hat{M}-1} \hat{s}_k s_l \hat{s}_m E[\mathbf{X}(n-j-k-1)d(n-j-1)\mathbf{W}_x^T \mathbf{X}(n-l-i-1)\mathbf{X}^T(n-m-i-1)] \\
 & (8) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{\hat{M}-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-k-1)d(n-j-1)\mathbf{V}^T(n-l-i-1)\mathbf{X}(n-l-i-1)\mathbf{X}^T(n-m-i-1)] \\
 & (9) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{W}_x \mathbf{V}^T(n-i-1)] \\
 & (10) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{\hat{M}-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{W}_x d(n-i-1)\mathbf{X}^T(n-m-i-1)] \\
 & (11) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{W}_x \mathbf{W}_x^T \mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\
 & (12) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\
 & \quad \cdot \mathbf{W}_x \mathbf{V}^T(n-m-i-1)\mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\
 & (13) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1)\mathbf{V}(n-k-j-1)\mathbf{V}^T(n-i-1)] \\
 & (14) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{\hat{M}-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{V}(n-k-j-1)d(n-i-1)\mathbf{X}^T(n-m-i-1)] \\
 & (15) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\
 & \quad \cdot \mathbf{V}(n-j-k-1)\mathbf{W}_x^T \mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \tag{b.3} \\
 & (16) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\
 & \quad \cdot \mathbf{V}(n-k-j-1)\mathbf{V}^T(n-m-i-1)\mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\
 & (17) + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l E[\mathbf{X}(n-k-j-1)\mathbf{X}^T(n-l-i-1)z(n-i-1)z(n-j-1)].
 \end{aligned}$$

Na equação (b.3) não foram colocados os termos contendo $z(n)$, uma vez é independente de qualquer sinal, e que é um processo com valor médio zero. Utilizando as suposições:

$$i) E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{W}(n-i)] \cong E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{V}(n-i)] \cong \mathbf{R}_{j-i}E[\mathbf{V}(n-i)],$$

$$ii) E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{V}(n-j)\mathbf{V}^T(n-i)] \cong E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)]E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)] \\ \cong \mathbf{R}_{j-i}E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)].$$

Utilizando também a definição da matriz de correlação cruzada entre o sinal de entrada e o sinal desejado $\mathbf{P}_{j-i} = E[d(n-i)\mathbf{X}(n-j)]$, a equação (b.3) pode ser escrita como:

$$\begin{aligned} & E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)] = \\ & (1) E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-i-1)] \\ & (2) + \mu \sum_{k=0}^{M-1} \hat{s}_k E[\mathbf{V}(n-j-1)] \mathbf{P}_k^T \\ & (3) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} s_k \hat{s}_l E[\mathbf{V}(n-j-1)] \mathbf{W}_x^T \mathbf{R}_{l-k} \\ & (4) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} s_k \hat{s}_l E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-k-i-1)] \mathbf{R}_{l-k} \\ & (5) + \mu \sum_{k=0}^{M-1} \hat{s}_k \mathbf{P}_k E[\mathbf{V}^T(n-i-1)] \\ & (6) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \hat{s}_k \hat{s}_l E[\mathbf{X}(n-j-k-1)d(n-j-1)d(n-i-1)\mathbf{X}^T(n-l-i-1)] \\ & (7) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \hat{s}_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-k-1)d(n-j-1)\mathbf{W}_x^T \mathbf{X}(n-l-i-1)\mathbf{X}^T(n-m-i-1)] \\ & (8) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-k-1)d(n-j-1)\mathbf{V}^T(n-l-i-1)\mathbf{X}(n-l-i-1)\mathbf{X}^T(n-m-i-1)] \\ & (9) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{W}_x E[\mathbf{V}^T(n-i-1)] \\ & (10) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{W}_x d(n-i-1)\mathbf{X}^T(n-m-i-1)] \\ & (11) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \sum_{r=0}^{M-1} \hat{s}_k \hat{s}_l \hat{s}_m \hat{s}_r E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{W}_x \mathbf{W}_x^T \mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\ & (12) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \sum_{r=0}^{M-1} s_k \hat{s}_l \hat{s}_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\ & \quad \cdot \mathbf{W}_x \mathbf{V}^T(n-m-i-1)\mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\ & (13) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} s_k \hat{s}_l \mathbf{R}_{k-l} E[\mathbf{V}(n-k-j-1)\mathbf{V}^T(n-i-1)] \\ & (14) - \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} s_k \hat{s}_l \hat{s}_m E[\mathbf{X}(n-j-l-1)\mathbf{X}^T(n-k-j-1)\mathbf{V}(n-k-j-1)d(n-i-1)\mathbf{X}^T(n-m-i-1)] \\ & (15) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \sum_{r=0}^{M-1} s_k \hat{s}_l \hat{s}_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\ & \quad \cdot \mathbf{V}(n-j-k-1)\mathbf{W}_x^T \mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\ & (16) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \sum_{r=0}^{M-1} s_k \hat{s}_l \hat{s}_m \hat{s}_r E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-k-j-1) \\ & \quad \cdot \mathbf{V}(n-k-j-1)\mathbf{V}^T(n-m-i-1)\mathbf{X}(n-m-i-1)\mathbf{X}^T(n-r-i-1)] \\ & (17) + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \hat{s}_k \hat{s}_l \mathbf{R}_{l+i-k-l} E[z(n-i)z(n-j)] \end{aligned} \tag{b.4}$$

Simplificação da Equação (b.4)

Para simplificar a (b.4) é utilizada a relação, obtida a partir da condição de regime permanente para os coeficientes do filtro adaptativo Eq. (2.41), e colocada a seguir:

$$\sum_{a=0}^{M-1} \sum_{b=0}^{\hat{M}-1} s_a \hat{s}_b \mathbf{R}_{a-b} \mathbf{W}_\infty = \sum_{c=0}^{\hat{M}-1} \hat{s}_c \mathbf{P}_{-c}. \quad (\text{b.5})$$

Como exemplo vamos considerar o termo (2) dado por $\sum_{k=0}^{\hat{M}-1} \hat{s}_k E[\mathbf{V}(n-j-1)] \mathbf{P}_k^T$. Vemos que este termo tem a mesma forma que o lado direito de (b.5), se o mesmo for transposto e pre-multiplicado por $E[\mathbf{V}(n-j-1)]$. Então, transpondo e pre-multiplicando ambos membros de (b.5) por $E[\mathbf{V}(n-j-1)]$ tem-se:

$$\sum_{a=0}^{M-1} \sum_{b=0}^{\hat{M}-1} s_a \hat{s}_b E[\mathbf{V}(n-j-1)] \mathbf{W}_\infty^T \mathbf{R}_{b-a} = \sum_{c=0}^{\hat{M}-1} \hat{s}_c E[\mathbf{V}(n-j-1)] \mathbf{P}_c^T.$$

Dessa igualdade, pode-se observar que tem a mesma forma que os termos (2) and (3), isto significa que ambos podem ser cancelados. Pelo mesmo procedimento é, também, possível cancelar os termos (5) e (9). Os termos contendo momentos de quarta ordem são calculados pela utilização do teorema para momento de quarta ordem de uma variável gaussiana [7,31], desenvolvido no Apêndice C e, expresso a seguir:

$$E[\mathbf{X}(n-i) \mathbf{X}^T(n-j) \mathbf{A} \mathbf{B}^T \mathbf{X}(n-k) \mathbf{X}^T(n-l)] = \mathbf{R}_{j-i} \mathbf{A} \mathbf{B}^T \mathbf{R}_{l-k} + \mathbf{R}_{k-i} \mathbf{B} \mathbf{A}^T \mathbf{R}_{l-j} + \mathbf{R}_{l-i} \text{tr}\{\mathbf{R}_{k-j} \mathbf{B} \mathbf{A}^T\}$$

Onde \mathbf{A} e \mathbf{B} são vetores e \mathbf{R}_{m-i} são matrizes de correlação da entrada, calculadas pelo valor esperado $E[\mathbf{X}(n-m) \mathbf{X}^T(n-l)]$. Pela aplicação deste teorema e da igualdade (b.5) são eliminados todos os termos que contem as variáveis \mathbf{W}_∞ , $d(n)$ e $E[\mathbf{V}(n-m)]$, $\forall m$. Isto significa que são cancelados os seguintes termos: (6), (7), (8), (10), (11), (12), (14) e (15). Portanto, a expressão recursiva final para o cálculo da matriz de correlação para os coeficientes do filtro adaptativo é dada por:

$$\begin{aligned}
 E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)] &= E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-i-1)] \\
 &\quad - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-k-i-1)] \mathbf{R}_{l-k} \\
 &\quad - \mu \sum_{k=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_r \mathbf{R}_{k-r} E[\mathbf{V}(n-k-j-1)\mathbf{V}^T(n-i-1)] \\
 &\quad + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r (\mathbf{R}_{k-l} E[\mathbf{V}(n-k-j-1)\mathbf{V}^T(n-m-i-1)] \mathbf{R}_{r-m} \\
 &\quad \quad + \mathbf{R}_{m+i-l-j} E[\mathbf{V}(n-m-i-1)\mathbf{V}^T(n-k-j-1)] \mathbf{R}_{r+i-k-j} \\
 &\quad \quad + \mathbf{R}_{r+i-l-j} \text{tr}\{E[\mathbf{V}(n-m-i-1)\mathbf{V}^T(n-k-j-1)] \mathbf{R}_{m+i-k-j}\}) \\
 &\quad + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l \mathbf{R}_{l+i-j-k} E[z(n-i-1)z(n-j-1)].
 \end{aligned} \tag{b.6}$$

Utilizando a notação $\mathbf{K}_{j,i} = E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)]$ (b.6) pode ser rescrita como:

$$\begin{aligned}
 \mathbf{K}_{j,i}(n) &= \mathbf{K}_{j,i}(n-1) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{K}_{j,k+i}(n-1) \mathbf{R}_{l-k} - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k+j,i}(n-1) \\
 &\quad + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r (\mathbf{R}_{k-l} \mathbf{K}_{k+j,m+i}(n-1) \mathbf{R}_{r-m} \\
 &\quad \quad + \mathbf{R}_{m+i-l-j} \mathbf{K}_{m+i,k+j}(n-1) \mathbf{R}_{r+i-k-j} \\
 &\quad \quad + \mathbf{R}_{r+i-l-j} \text{tr}\{\mathbf{K}_{m+i,k+j}(n-1) \mathbf{R}_{m+i-k-j}\}) \\
 &\quad + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l \mathbf{R}_{l+i-j-k} E[z(n-i-1)z(n-j-1)].
 \end{aligned} \tag{b.7}$$

Determinação da Equação associada à (b.4)

Da mesma forma que em (3.22), onde a matriz de correlação com mesmo índice dependia de uma matriz com índices diferentes, em (b.7) acontece a mesma situação. Quando foi tratada a Eq. (3.22), foi necessária a introdução da Eq. (3.23) a fim de determinar todas as variáveis que descrevem a matriz de correlação. Por este motivo, vamos precisar contar com uma expressão associada, equivalente à (3.23), para complementar (b.7). Novamente, para obter esta equação vamos partir da expressão de $\mathbf{V}(n-j)$, dada pela equação (2.47). Então, se pós-multiplicarmos ambos os lados de (2.47) por $\mathbf{V}^T(n-k-i)$ e tomamos o valor esperado, a expressão resultante é:

$$\begin{aligned}
 E[\mathbf{V}(n-j)\mathbf{V}^T(n-k-i)] &= E[\mathbf{V}(n-j-1)\mathbf{V}^T(n-k-i)] \\
 &+ \mu \sum_{m=0}^{\hat{M}-1} \hat{s}_m E[\mathbf{X}(n-m-j-1)d(n-j-1)E[\mathbf{V}^T(n-k-i)]] \\
 &- \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-m-j-1)]E[\mathbf{V}(n-m-j-1)\mathbf{V}^T(n-k-i)] \\
 &- \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l E[\mathbf{X}(n-l-j-1)\mathbf{X}^T(n-m-j-1)]\mathbf{W}_\infty E[\mathbf{V}^T(n-k-i)] \\
 &+ \mu \sum_{m=0}^{\hat{M}-1} \hat{s}_m E[\mathbf{X}(n-m-j-1)E[z(n-j-1)]E[\mathbf{V}^T(n-k-i)]].
 \end{aligned} \tag{b.8}$$

Utilizando a notação corrente, e considerando que o valor esperado da variável $z(n)$ é zero, por ser um processo de média zero, (b.8) pode ser rescrita como:

$$\begin{aligned}
 \mathbf{K}_{j,k+i}(n) &= \mathbf{K}_{j,k+i-1}(n-1) + \mu \sum_{m=0}^{\hat{M}-1} \hat{s}_m \mathbf{P}_{-m} E[\mathbf{V}^T(n-k-i)] \\
 &- \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{K}_{m+j,k+i-1}(n-1) - \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{W}_\infty E[\mathbf{V}^T(n-k-i)].
 \end{aligned} \tag{b.9}$$

A expressão (b.9) pode ser simplificada pelo uso da relação (b.5). Para isto se pós-multiplicarmos ambos os lados de (b.5) por $E[\mathbf{V}^T(n-k-i)]$, pode-se observar que o segundo e o último termo em (b.9) são iguais. Portanto, podemos rescrever a seguinte expressão para (b.9):

$$\mathbf{K}_{j,k+i}(n) = \mathbf{K}_{j,k+i-1}(n-1) - \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{K}_{m+j,k+i-1}(n-1). \tag{b.10}$$

Com as equações (b.7) e (b.10) estamos em condições de modelar o comportamento do momento de segunda ordem do vetor erro nos coeficientes do filtro adaptativo, utilizando o algoritmo LMS filtrado. A fim de acondicionar os índices, tal como foi realizado para obter-se as equações (3.30), devemos manipular os índices das matrizes de correlação. A forma em que estes devem ser trabalhados é obtida a partir da Eq. (3.53). Nessa expressão, aparece o cálculo do traço de $\mathbf{K}_{j,i}(n) = E[\mathbf{V}(n-j)\mathbf{V}^T(n-i)]$, onde os índices

i e j podem assumir duas condições, serem iguais ou diferentes. Em função desta premissa, o procedimento é o seguinte:

Considere que os índices são iguais. Seja $i = j = D$, substituindo esta condição em (b.7), temos:

$$\begin{aligned}
 \mathbf{K}_{D,D}(n) = & \mathbf{K}_{D,D}(n-1) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{K}_{D,k+D}(n-1) \mathbf{R}_{l-k} - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k+D,D}(n-1) \\
 & + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r (\mathbf{R}_{k-l} \mathbf{K}_{k+D,m+D}(n-1) \mathbf{R}_{r-m} \\
 & \quad + \mathbf{R}_{m-l} \mathbf{K}_{m+D,k+D}(n-1) \mathbf{R}_{r-k} \\
 & \quad + \mathbf{R}_{r-l} \text{tr}\{\mathbf{K}_{m+D,k+D}(n-1) \mathbf{R}_{m-k}\}) \\
 & + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l \mathbf{R}_{l-k} E[z(n-D-1)z(n-D-1)]
 \end{aligned} \tag{b.11}$$

Utilizando a relação $\mathbf{K}_{D,D}(n) = \mathbf{K}_{0,0}(n-D)$, chamando $\sigma_z^2 = E[z(n)z(n)]$ e fazendo $n-D = n$, (b.11) é rescrita como:

$$\begin{aligned}
 \mathbf{K}_{0,0}(n) = & \mathbf{K}_{0,0}(n-1) - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{K}_{0,k}(n-1) \mathbf{R}_{l-k} - \mu \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_k \hat{s}_l \mathbf{R}_{k-l} \mathbf{K}_{k,0}(n-1) \\
 & + \mu^2 \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{m=0}^{M-1} \sum_{r=0}^{\hat{M}-1} s_k \hat{s}_l s_m \hat{s}_r (\mathbf{R}_{k-l} \mathbf{K}_{k,m}(n-1) \mathbf{R}_{r-m} \\
 & \quad + \mathbf{R}_{m-l} \mathbf{K}_{m,k}(n-1) \mathbf{R}_{r-k} \\
 & \quad + \mathbf{R}_{r-l} \text{tr}\{\mathbf{K}_{m,k}(n-1) \mathbf{R}_{m-k}\}) \\
 & + \mu^2 \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{\hat{M}-1} \hat{s}_k \hat{s}_l \mathbf{R}_{l-k} E[z(n-1)z(n-1)]
 \end{aligned} \tag{b.12}$$

De (b.12) vemos que $\mathbf{K}_{0,0}(n)$ depende das matrizes $\mathbf{K}_{0,k}(n)$ e $\mathbf{K}_{k,0}(n)$. Entretanto, as matrizes $\mathbf{K}_{k,m}(n)$ sempre poderão ser colocadas na forma $\mathbf{K}_{0,k}(n)$ ou $\mathbf{K}_{k,0}(n)$ através da relação $\mathbf{K}_{k,m}(n) = \mathbf{K}_{k-m,0}(n-m)$ se $k > m$, ou $\mathbf{K}_{k,m}(n) = \mathbf{K}_{0,m-k}(n-k)$ em caso contrario. Tais matrizes são versões atrasadas de $\mathbf{K}_{0,1}(n)$ ou $\mathbf{K}_{1,0}(n)$, as que são

determinadas manipulando os índices de (b.10). Para isto, e sem perda de generalidade, vamos continuar com a suposição $i = j = D$, mas agora aplicada em (b.10), obtendo:

$$\mathbf{K}_{D,k+D}(n) = \mathbf{K}_{D,k+D-1}(n-1) - \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{K}_{m+D,k+D-1}(n-1) , \quad (\text{b.13})$$

ou,

$$\mathbf{K}_{0,k}(n) = \mathbf{K}_{0,k-1}(n-1) - \mu \sum_{m=0}^{M-1} \sum_{l=0}^{\hat{M}-1} s_m \hat{s}_l \mathbf{R}_{m-l} \mathbf{K}_{m,k-1}(n-1) . \quad (\text{b.14})$$

Onde pode-se constatar que (b.12) e (b.14) correspondem a (3.54), como desejávamos demonstrar.

Apêndice C

CÁLCULO DO MOMENTO DE QUARTA ORDEM DE UMA VARIÁVEL GAUSSIANA

No Capítulo 3, foi utilizado o teorema para os momentos de quarta ordem de uma variável gaussiana. Este teorema foi aplicado para determinar os valores esperados contendo termos de quarta ordem dentro da expressão da matriz de correlação. Na literatura de filtragem adaptativa [6,7,16,29,35], este teorema é formulado utilizando a teoria da independência. No presente trabalho, este teorema foi reformulado sem a utilização desta teoria, preservando assim a correlação entre os vetores de entrada para os diferentes instantes de tempo, e cuja expressão é escrita a seguir:

$$E[\mathbf{X}(n-i)\mathbf{X}^T(n-j)\mathbf{A}\mathbf{B}^T\mathbf{X}(n-k)\mathbf{X}^T(n-l)] = \mathbf{R}_{j-i}\mathbf{A}\mathbf{B}^T\mathbf{R}_{l-k} + \mathbf{R}_{k-i}\mathbf{B}\mathbf{A}^T\mathbf{R}_{l-j} + \mathbf{R}_{l-i}\text{tr}\{\mathbf{R}_{k-j}\mathbf{B}\mathbf{A}^T\}, \quad (\text{c.1})$$

onde $\mathbf{X}(n)$, \mathbf{A} e \mathbf{B} são vetores coluna de dimensão $N \times 1$. Para obter-se a equação (c.1) procede-se da seguinte maneira. Sejam as matrizes \mathbf{D} , \mathbf{F} e \mathbf{G} , definidas como sendo:

$$\begin{aligned} \mathbf{D} &= \mathbf{X}(n-i)\mathbf{X}^T(n-j), \\ \mathbf{F} &= \mathbf{A}\mathbf{B}^T, \\ \mathbf{G} &= \mathbf{X}(n-k)\mathbf{X}^T(n-l), \end{aligned}$$

e seja \mathbf{M} , a matriz definida pelo produto dessas matrizes, então:

$$\mathbf{M} = \mathbf{D}\mathbf{F}\mathbf{G},$$

sendo seus elementos definidos pela relação:

$$(\mathbf{M})_{rp} = \sum_{q=1}^N \sum_{m=1}^N (\mathbf{D})_{rm} (\mathbf{F})_{mq} (\mathbf{G})_{qp}.$$

Ou, em forma equivalente

$$(\mathbf{M})_{rp} = \sum_{q=1}^N \sum_{m=1}^N x(n-i-r)x(n-j-m)(\mathbf{F})_{mq} x(n-k-q)x(n-l-p). \quad (\text{c.2})$$

Quando é aplicado o valor esperado em ambos os lados da equação (c.2), temos:

$$E[(\mathbf{M})_{rp}] = \sum_{q=1}^N \sum_{m=1}^N E[x(n-i-r)x(n-j-m)x(n-k-q)x(n-l-p)](\mathbf{F})_{mq}. \quad (\text{c.3})$$

Segundo [7], o valor esperado é dado por:

$$\begin{aligned} E[(\mathbf{M})_{rp}] &= \sum_{q=1}^N \sum_{m=1}^N E[x(n-i-r)x(n-j-m)x(n-k-q)x(n-l-p)](\mathbf{F})_{mq} = \\ &= \sum_{q=1}^N \sum_{m=1}^N \left(E[x(n-i-r)x(n-j-m)]E[x(n-k-q)x(n-l-p)](\mathbf{F})_{mq} \right. \\ &\quad + E[x(n-i-r)x(n-k-q)]E[x(n-j-m)x(n-l-p)](\mathbf{F})_{mq} \\ &\quad \left. + E[x(n-i-r)x(n-l-p)]E[x(n-j-m)x(n-k-q)](\mathbf{F})_{mq} \right). \end{aligned} \quad (\text{c.4})$$

Expressando a equação (c.4) em forma matricial, podem-se identificar as seguintes estruturas:

$$\begin{aligned} e1) \quad & \sum_{q=1}^N \sum_{m=1}^N E[x(n-i-r)x(n-j-m)x(n-k-q)x(n-l-p)](\mathbf{F})_{mq} = \\ & \sum_{q=1}^N \sum_{m=1}^N E[(\mathbf{X}(n-i)\mathbf{X}^T(n-j))_{rm}](\mathbf{F})_{mq} E[(\mathbf{X}(n-k)\mathbf{X}^T(n-l))_{pq}] = \mathbf{R}_{j-i} \mathbf{F} \mathbf{R}_{l-k}, \end{aligned}$$

$$\begin{aligned}
 e2) \quad & \sum_{q=1}^N \sum_{m=1}^N E[x(n-i-r)x(n-k-q)x(n-j-m)x(n-l-p)](\mathbf{F})_{mq} = \\
 & \sum_{q=1}^N \sum_{m=1}^N E[(\mathbf{X}(n-i)\mathbf{X}^T(n-k))_{rq}] (\mathbf{F})_{qm} E[(\mathbf{X}(n-j)\mathbf{X}^T(n-l))_{mq}] = \mathbf{R}_{k-i} \mathbf{F}^T \mathbf{R}_{l-j},
 \end{aligned}$$

$$\begin{aligned}
 e3) \quad & \sum_{q=1}^N \sum_{m=1}^N E[x(n-i-r)x(n-l-p)x(n-j-m)x(n-k-q)](\mathbf{F})_{mq} = \\
 & \sum_{q=1}^N \sum_{m=1}^N E[(\mathbf{X}(n-i)\mathbf{X}^T(n-l))_{rp}] (\mathbf{F})_{mq} E[(\mathbf{X}(n-j)\mathbf{X}^T(n-k))_{mq}] = \\
 & E[(\mathbf{X}(n-i)\mathbf{X}^T(n-l))_{rp}] \sum_{q=1}^N \sum_{m=1}^N (\mathbf{F})_{mq} E[(\mathbf{X}(n-j)\mathbf{X}^T(n-k))_{mq}] = \mathbf{R}_{l-i} \text{tr}\{\mathbf{F}\mathbf{R}_{k-j}\}.
 \end{aligned}$$

Juntando os termos (e1), (e2) e (e3), obtêm-se a equação (c.1), que desejava-se constatar.