

MÁRCIO HOLSBACH COSTA

**COMPORTAMENTO DO ALGORITMO FXLMS
EM SISTEMAS COM
CAMINHO SECUNDÁRIO NÃO-LINEAR:
APLICAÇÃO AO CONTROLE ATIVO DE RUÍDO**

FLORIANÓPOLIS

2001

UNIVERSIDADE FEDERAL DE SANTA CATARINA

**PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA ELÉTRICA**

**COMPORTAMENTO DO ALGORITMO FXLMS
EM SISTEMAS COM
CAMINHO SECUNDÁRIO NÃO-LINEAR:
APLICAÇÃO AO CONTROLE ATIVO DE RUÍDO**

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

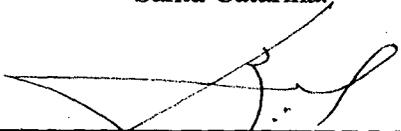
MÁRCIO HOLSBACK COSTA

Florianópolis, maio de 2001.

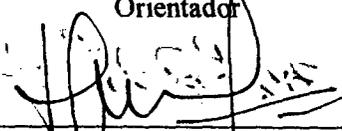
COMPORTAMENTO DO ALGORITMO FXLMS EM SISTEMAS COM CAMINHO SECUNDÁRIO NÃO-LINEAR: APLICAÇÃO AO CONTROLE ATIVO DE RUÍDO

Márcio Holsbach Costa

‘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 Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.’



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

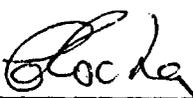


Prof. Edson Roberto de Pieri, Dr.
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:



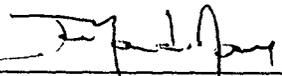
Prof. José Carlos Moreira Bermudez, Ph.D.
Presidente



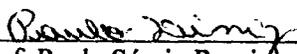
Prof. Carlos Aurélio Faria da Rocha, D.Sc.



Prof. Jefferson Luiz Brum Marques, Ph.D.



Prof. João Marcos Travassos Romano, Dr.



Prof. Paulo Sérgio Ramirez Diniz, Ph.D.

AGRADECIMENTOS

Aos funcionários da UFSC: Elton Fontão, João Marco Ferreira, Marcos Assis e Walter Gontijo pelo auxílio técnico-administrativo.

Ao professor Victorino Piccinini pelo grande auxílio com a língua portuguesa.

Aos professores do programa de pós-graduação em engenharia elétrica pelo ambiente de aprendizado.

À Universidade Católica de Pelotas e à Capes pela oportunidade, incentivo e financiamento.

Aos colegas Alexandre Pino, Cássio Lopes, Maurício Tavares, Orlando Tobias, Pedro Inácio Hübsher e Sérgio Almeida pelas instigantes discussões e contribuições.

Ao professor Neil Bershad pela valiosa ajuda e disposição ao longo de todo o trabalho.

Ao meu orientador José Carlos Bermudez, pela grande atenção, empenho e amizade ao longo destes anos.

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

COMPORTAMENTO DO ALGORITMO FXLMS EM SISTEMAS COM CAMINHO SECUNDÁRIO NÃO-LINEAR: APLICAÇÃO AO CONTROLE ATIVO DE RUÍDO

Márcio Holsbach Costa

Maio/2001

Orientador: José Carlos Moreira Bermudez, Ph.D.

Área de Concentração: Sistemas de Informação.

Palavras-chave: Algoritmo LMS, Controle Ativo de Ruído, Sistemas Não-Lineares.

Número de Páginas: 203.

RESUMO

Este trabalho apresenta um estudo sobre o comportamento do algoritmo LMS Filtrado (FXLMS) quando o caminho secundário (saída do filtro adaptativo) é modelado por um sistema Wiener-Hammerstein. Essa configuração é de especial interesse em aplicações de controle ativo de ruído acústico e vibrações, onde a característica não-linear do tipo saturação representa as distorções introduzidas pelos transdutores e amplificadores de potência. São obtidas equações determinísticas para o comportamento médio dos coeficientes, erro médio quadrático e momentos de segunda ordem para sinais gaussianos. A partir dessas, soluções para o regime permanente são obtidas. Simulações Monte Carlo demonstram excelente concordância com o comportamento previsto pelos modelos teóricos. Os resultados analíticos e as simulações apresentados neste trabalho demonstram o impacto significativo que características não-lineares podem ter sobre o comportamento do sistema adaptativo.

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

BEHAVIOR OF THE FXLMS ALGORITHM IN SYSTEMS WITH A NONLINEAR SECONDARY PATH: APPLICATION TO ACTIVE NOISE CONTROL

Márcio Holsbach Costa

May/2001

Advisor: José Carlos Moreira Bermudez, Ph.D.

Area of Concentration: Information Systems.

Keywords: Adaptive Systems, Acoustic Noise, Nonlinear Systems.

Number of Pages: 203.

ABSTRACT

This work presents a statistical analysis of the Filtered LMS algorithm behavior when the secondary path (output of the adaptive filter) is modeled by a Wiener-Hammerstein system. This system is of special interest for active acoustic noise and vibration control, where the saturation nonlinearity models the nonlinear distortion introduced by the power amplifiers and transducers. Deterministic nonlinear recursions are derived for the mean and second order moments of the adaptive weights and for the mean square error for Gaussian inputs. Steady-state conditions are derived from these results. Monte Carlo simulations show excellent agreement with the behavior predicted by the theoretical models. The analytical and simulation results presented show the significant impact that even little nonlinearities can have on the adaptive filter behavior.

SUMÁRIO

1. INTRODUÇÃO	1
1.1 FILTRAGEM ADAPTATIVA	1
1.1.1 O ALGORITMO LMS	4
1.2 CONTROLE ATIVO DE RUÍDO ACÚSTICO E VIBRAÇÕES	8
1.2.1 INTRODUÇÃO	8
1.2.2 TIPOS DE SISTEMAS ANC	11
1.2.3 ALGORITMOS ADAPTATIVOS PARA SISTEMAS ANC.....	14
1.3 NÃO-LINEARIDADES EM SISTEMAS ANC	17
1.4 FILTROS ADAPTATIVOS SUJEITOS A NÃO-LINEARIDADES.....	19
1.5 OBJETIVOS DESTE TRABALHO.....	20
1.6 JUSTIFICATIVA E RELEVÂNCIA	20
1.7 ESTRUTURA DO TRABALHO	21
2. DEFINIÇÃO DO PROBLEMA.....	22
2.1 INTRODUÇÃO.....	22
2.2 ALGORITMO FXLMS SUJEITO A UM SISTEMA WH.....	26
2.3 CARACTERÍSTICAS DOS SINAIS E DINÂMICA DO SISTEMA	27
2.4 INFLUÊNCIA LINEAR DO CAMINHO SECUNDÁRIO.....	30
2.5 MODELO ANALÍTICO DA NÃO-LINEARIDADE	32
2.6 GRAU DE NÃO-LINEARIDADE DO SISTEMA ANC	37
2.7 CASOS PARTICULARES DE INTERESSE.....	38
2.7.1 SISTEMA WIENER.....	39
2.7.2 SISTEMA SEM MEMÓRIA.....	39
3. ANÁLISE DA SUPERFÍCIE DE DESEMPENHO	41
3.1 DETERMINAÇÃO ANALÍTICA DA SUPERFÍCIE DE DESEMPENHO.....	41
3.1.1 SISTEMA WIENER.....	45
3.1.2 SISTEMA SEM MEMÓRIA.....	46
3.2 PONTOS ESTACIONÁRIOS.....	47

3.2.1 SOLUÇÃO APROXIMADA PARA O CASO GERAL.....	50
3.2.2 SISTEMA WIENER.....	53
3.2.3 SISTEMA SEM MEMÓRIA.....	55
3.3 HESSIANO DA SUPERFÍCIE DE DESEMPENHO	55
3.4 MÍNIMO ERRO MÉDIO QUADRÁTICO	59
3.4.1 SOLUÇÃO APROXIMADA PARA O CASO GERAL.....	60
3.4.2 SISTEMA WIENER.....	60
3.4.3 SISTEMA SEM MEMÓRIA.....	60
4. COMPORTAMENTO MÉDIO DOS COEFICIENTES	62
4.1 ANÁLISE DO COMPORTAMENTO MÉDIO DOS COEFICIENTES.....	63
4.1.1 SISTEMA WIENER.....	66
4.1.2 SISTEMA SEM MEMÓRIA.....	67
4.2 COEFICIENTES EM REGIME PERMANENTE.....	68
4.2.1 SISTEMA WIENER.....	69
4.2.2 SISTEMA SEM MEMÓRIA.....	70
4.3 MODELO SIMPLIFICADO	70
4.4 ANÁLISE DA ESTABILIDADE	72
4.5 INTERPRETAÇÃO DO PARÂMETRO η^2	73
5. ERRO MÉDIO QUADRÁTICO	76
5.1 ANÁLISE DO ERRO MÉDIO QUADRÁTICO.....	76
5.1.1 SISTEMA WIENER.....	81
5.1.2 SISTEMA SEM MEMÓRIA.....	81
5.2 ERRO MÉDIO QUADRÁTICO EM REGIME PERMANENTE.....	82
5.2.1 SISTEMA WIENER.....	83
5.2.2 SISTEMA SEM MEMÓRIA.....	83
5.3 MODELO SIMPLIFICADO	84

6. MOMENTOS DE SEGUNDA ORDEM	86
6.1 ANÁLISE DOS MOMENTOS DE SEGUNDA ORDEM	86
6.1.1 SISTEMA WIENER.....	102
6.1.2 SISTEMA SEM MEMÓRIA.....	103
7. RESULTADOS.....	105
7.1 ANÁLISE DA SUPERFÍCIE DE DESEMPENHO	105
7.1.1 DISTORÇÕES DA SUPERFÍCIE DE DESEMPENHO	105
7.1.2 PONTOS ESTACIONÁRIOS	108
7.1.3 ERRO EM EXCESSO DEVIDO À NÃO-LINEARIDADE.....	114
7.2 COMPORTAMENTO EM REGIME TRANSITÓRIO.....	115
7.2.1 COEFICIENTES E ERRO MÉDIO QUADRÁTICO.....	115
7.2.2 EFEITOS DOS MOMENTOS DE SEGUNDA ORDEM.....	121
7.2.3 EFEITOS DO PASSO DE CONVERGÊNCIA	123
7.2.4 EFEITOS DO GRAU DE CORRELAÇÃO E DA DURAÇÃO DA RESPOSTA DA PLANTA.....	126
7.3 COMPORTAMENTO EM REGIME PERMANENTE	128
7.3.1 POLARIZAÇÃO DOS COEFICIENTES	129
7.3.2 ERRO MÉDIO QUADRÁTICO EM EXCESSO	130
7.3.3 DESAJUSTE EM REGIME PERMANENTE	131
7.3.4 ESTABILIDADE.....	132
8. DISCUSSÃO E CONCLUSÕES.....	134
8.1 PROPOSTAS PARA CONTINUAÇÃO DO TRABALHO.....	136
9. ANEXO 1 – DETERMINAÇÃO DE $E\{g(y_1)Y_2\}$	138
10. ANEXO 2 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)\}$.....	142
11. ANEXO 3 – DETERMINAÇÃO DE $E\{Y_1Y_2^T W_1 W_2^T Y_3 Y_4^T\}$	148

12. ANEXO 4 – DETERMINAÇÃO DE $E\{g(y_1)y_2Y_3Y_4^T\}$	150
13. ANEXO 5 – DETERMINAÇÃO DE $E\{g(y)y^3\}$	154
14. ANEXO 6 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)Y_3Y_4^T\}$	157
15. ANEXO 7 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)y_1^2\}$	162
16. ANEXO 8 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)y_1y_2\}$	172
17. ANEXO 9 – DETERMINAÇÃO DE η^2	181
18. ANEXO 10 - PUBLICAÇÕES	184
19. REFERÊNCIAS BIBLIOGRÁFICAS	186

1. INTRODUÇÃO

O objetivo desta seção é introduzir os principais conceitos necessários à compreensão do trabalho. Basicamente quatro tópicos são abordados: filtragem adaptativa; controle ativo de ruído acústico e vibrações (ANC); não-linearidades em sistemas ANC; e algoritmos adaptativos sujeitos a não-linearidades.

A partir dos tópicos abordados na revisão da bibliografia descrita acima, são traçados os objetivos do trabalho, suas justificativas e relevância. Finalizando, é apresentada a estrutura sob a qual analisamos os problemas levantados.

1.1 Filtragem Adaptativa

O processamento adaptativo de sinais teve um grande impulso a partir da década de 60 através do desenvolvimento do algoritmo LMS (WIDROW e HOFF, 1960). Desde então essa técnica tem sido amplamente explorada em um grande número de aplicações em diferentes áreas, como: sistemas de transmissão de dados (MODEMS, telefonia móvel, comunicações por satélite); processamento de sinais biológicos; radar e sonar; exploração geofísica; compressão de voz; cancelamento de eco; equalização; filtragem; controle; estimação espectral; e, mais recentemente, na área de controle ativo de ruído acústico e de vibrações (WIDROW *et al.*, 1975, HAYKIN, 1991, KUO e MORGAN, 1996).

A filtragem adaptativa tem cativado a atenção de inúmeros pesquisadores por sua capacidade intrínseca de otimização de desempenho, usando pouca informação *a priori* sobre as características do sistema. Essa propriedade é especialmente vantajosa quando o sistema em questão não é completamente descrito ou conhecido em termos de suas características físicas ou dinâmicas (MITRA e KAISER, 1993).

Processamentos *off-line* permitem a utilização de algoritmos matematicamente mais sofisticados, possibilitando uma busca mais apurada da solução ótima. Entretanto, em algumas aplicações em tempo real esse tipo de sofisticação torna-se praticamente inviável.

Basicamente um autômato (ou filtro) adaptativo é uma estrutura ajustável de tal modo que seu desempenho possa ser otimizado por um algoritmo numérico. De uma forma geral, pode-se entender esse autômato como um filtro (geralmente digital) cujos coeficientes são modificados de forma a minimizar uma função objetivo (geralmente determinada em função do sinal de saída desse filtro).

Segundo WIDROW e STEARNS (1985), os sistemas adaptativos possuem todas ou algumas das seguintes características:

- adaptação automática à medida que ocorre a modificação do ambiente e/ou alterações do sistema (**auto-otimização**);
- podem ser treinados para desenvolver uma tarefa específica de filtragem ou decisão, ou seja, podem ser programados através de um processo de treinamento (**autoprogramáveis**);
- em decorrência do item anterior, não necessitam procedimentos elaborados de síntese, são basicamente **autoprojetáveis**;
- podem extrapolar o espaço de conhecimento e lidar com novas situações após o treinamento com um pequeno conjunto de padrões de entrada (**auto-aprendizado**);
- até certo ponto podem reparar a si mesmos, ou seja, podem adaptar-se em regiões próximas da ótima mesmo quando sujeitos a certos tipos de defeitos ou limitações (**auto-regeneráveis**);

- geralmente são mais complexos e difíceis de analisar que sistemas não-adaptativos, mas oferecem a possibilidade de um desempenho substancialmente melhor quando as características do ambiente são desconhecidas ou variantes no tempo.

A partir do exposto pode-se concluir que os autômatos adaptativos são essencialmente sistemas não-lineares e variantes no tempo. Entretanto, os parâmetros de uma certa classe desses filtros tornam-se aproximadamente constantes após algum tempo de funcionamento. Quando isso ocorre, sua ação sobre o sistema torna-se linear. Tais sistemas, chamados de filtros adaptativos lineares, são bastante úteis, tendem a ser matematicamente tratáveis e geralmente são mais fáceis de analisar que a maioria dos autômatos adaptativos.

Em termos de identificação de sistemas lineares pode-se dividi-los em duas classes quanto à forma de sua resposta ao impulso: resposta finita ao impulso (FIR - *finite impulse response*) e resposta infinita ao impulso (IIR - *infinite impulse response*). Existem diferentes estruturas de implementação, conforme a classe a que o sistema analisado pertence (OPPENHEIM e SCHAFER, 1989).

Tratando-se da classe de sistemas adaptativos lineares, podem ser encontrados inúmeros trabalhos de análise, projeto e aplicações. Apenas para citar alguns dos algoritmos mais utilizados, tem-se (GLENZIS *et al.*, 1999):

- LMS;
- LMS Filtrado (FXLMS);
- *Leaky* LMS;
- LMS Normalizado;
- RLS e suas variações.

Atualmente os esforços encontram-se direcionados para a procura de estruturas alternativas e de novos algoritmos. Entretanto, diversas questões permanecem não resolvidas com relação à dinâmica e desempenho desses sistemas, entre elas:

- efeitos de precisão finita (quantização);
- características do sinal de excitação (as análises são geralmente baseadas em considerações de ruído branco gaussiano);
- influências de componentes associados (características não-lineares de transdutores associados, atrasos associados ao sistema analisado, distorções de sinais devido ao *hardware*, etc).

1.1.1 O Algoritmo LMS

Embora escape ao propósito deste trabalho uma descrição e análise detalhada dos diferentes algoritmos adaptativos, faz-se necessária uma introdução rápida ao algoritmo que será a base de nosso estudo: o LMS. Havendo a necessidade de um maior rigor em sua descrição e em seus derivados (FXLMS, NLMS, *Leaky* LMS, ...), diversas referências podem ser utilizadas: MANOLAKIS *et al.* (2000), HAYKIN (1991) e WIDROW e STEARNS (1985).

O algoritmo LMS foi desenvolvido por WIDROW e HOFF em 1959 na Universidade de Stanford e se popularizou por sua simplicidade computacional. Este algoritmo adaptativo utiliza uma estimação instantânea da superfície de desempenho para produzir sucessivas atualizações dos coeficientes através do método do gradiente descendente. Devido a essa característica o LMS é dito pertencente à classe de algoritmos baseados no método do gradiente estocástico e é considerado como padrão na comparação de desempenho entre algoritmos adaptativos.

O LMS pode ser considerado como um sistema de controle realimentado e basicamente consiste na combinação de dois processos básicos:

- processo adaptativo: consiste no ajuste automático de um conjunto de coeficientes;
- processo de filtragem: envolve (a) o cálculo do produto interno entre o vetor de sinais de entrada e o vetor de coeficientes, formando assim um sinal de saída e (b) a geração de um sinal de erro através da comparação entre o sinal produzido pelo algoritmo e um sinal dito desejado. Esse sinal de erro é utilizado em conjunto com o sinal de entrada para atualizar o processo adaptativo.

A figura 1.1 apresenta o diagrama em blocos de um problema de identificação/controlado onde o objetivo é minimizar a potência do sinal de saída $e(n)$ através da especificação do controlador W .

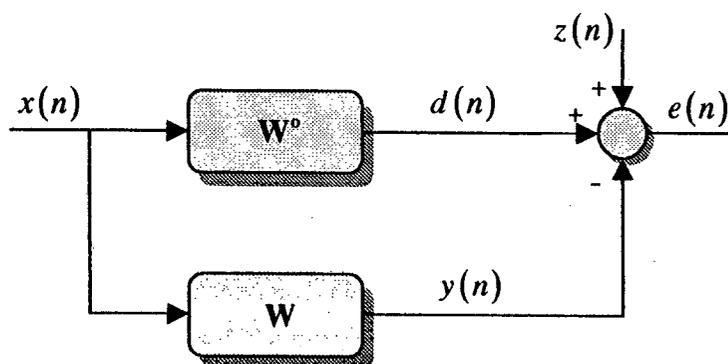


Figura 1.1 – Diagrama em blocos do problema analisado

A figura 1.1 apresenta um sistema W^o , considerado linear, excitado por um sinal $x(n)$ (onde “ n ” representa o instante de tempo), produzindo um sinal de saída $d(n)$. W é o controlador que, sujeito ao mesmo sinal de excitação $x(n)$ ou a um sinal de alguma forma correlacionado a este, produz uma saída denominada de $y(n)$. O sinal $z(n)$ corresponde ao ruído de medição. O objetivo do problema é determinar o controlador *ótimo* W , em um sentido médio quadrático, de forma a minimizar a saída $e(n)$.

O sistema W^o é considerado possuindo resposta finita ao impulso (FIR) e portanto pode ser representado por um filtro transversal, ou, como comumente é chamado, por um combinador linear. Portanto, a saída $d(n)$ pode ser avaliada através da seguinte equação:

$$d(n) = \sum_{j=0}^{N-1} w_j^o x(n-j) \quad (1.1)$$

ou, na forma vetorial, por:

$$d(n) = W^{oT} X(n) \quad (1.2)$$

onde

$$\begin{cases} W^o = [w_0^o & w_1^o & \dots & w_{N-1}^o]^T \\ X(n) = [x(n) & x(n-1) & \dots & x(n-N+1)]^T \end{cases} \quad (1.3)$$

Ou seja, o vetor W^o é constituído por uma série ordenada de coeficientes que corresponde à resposta ao impulso do sistema. O vetor $X(n)$ é o vetor de amostras atrasadas do sinal de excitação (ou, como comumente é chamado, sinal de referência). No próximo instante de tempo “ $n+1$ ” todas as amostras de $X(n)$ são deslocadas para a direita, a amostra mais antiga é descartada e a nova amostra do sinal de excitação ocupa a primeira posição do vetor.

Da mesma forma, tem-se que

$$y(n) = \sum_{j=0}^{N-1} w_j x(n-j) = W^T(n) X(n) \quad (1.4)$$

e, conseqüentemente, o sinal de erro instantâneo é dado por

$$e(n) = d(n) - y(n) = W^{oT} X(n) - W^T(n) X(n) + z(n) \quad (1.5)$$

O erro médio quadrático pode ser avaliado elevando-se (1.5) ao quadrado e tomando-se seu valor esperado. É fácil demonstrar que esse resultado é uma função quadrática do vetor de coeficientes do controlador \mathbf{W} . Portanto, $E\{e^2(n)\}$ é um hiperparabolóide no espaço dos coeficientes de \mathbf{W} . Essa superfície denomina-se **superfície de desempenho** do problema analisado.

Para determinar o conjunto de coeficientes que minimiza (produz o mínimo valor) a superfície de desempenho, pode-se utilizar o método do gradiente descendente:

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \mu_o \nabla E\{e^2(n)\} \quad (1.6)$$

onde $\mu_o = \mu/2$ é uma constante de adaptação denominada de passo de convergência.

Ou seja, partindo-se de um valor inicial arbitrário para $\mathbf{W}(0)$, a cada iteração do algoritmo pequenos passos são dados em direção ao conjunto *ótimo* de coeficientes.

O LMS utiliza uma estimativa do gradiente da superfície de desempenho para realizar a adaptação. Ao invés do erro médio quadrático ele utiliza o erro instantâneo, ou seja:

$$\nabla E\{e^2(n)\} \cong \nabla e^2(n) = 2e(n) \frac{\partial e(n)}{\partial \mathbf{W}(n)} = -2e(n) \mathbf{X}(n) \quad (1.7)$$

Substituindo (1.7) em (1.6) tem-se como resultado a equação de atualização dos coeficientes do LMS:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu e(n) \mathbf{X}(n) \quad (1.8)$$

Ao longo deste trabalho será utilizada a seguinte notação:

- grandezas escalares: letras minúsculas;
- grandezas vetoriais: letras maiúsculas em negrito.

1.2 Controle Ativo de Ruído Acústico e Vibrações

1.2.1 Introdução

Dentro do contexto de abrangência de filtros adaptativos encontra-se a técnica de controle ativo de ruído (LEITCH e TOKHI, 1987, STEVENS e AHUJA, 1991, MORGAN, 1991, ELLIOT e NELSON, 1993, FÜLLER e FLOTOW, 1995, NELSON e ELLIOT, 1995, KUO e MORGAN, 1996, HANSEN, 1997, KUO e MORGAN, 1999).

Basicamente, pode-se descrever o processo de controle de ruído acústico (ANC) como sendo o cancelamento das ondas de pressão (som) em um determinado ponto ou região no espaço, através da criação de um campo acústico artificial de mesma intensidade, porém com fase inversa. A sobreposição de ambos os efeitos ocasiona a diminuição ou o cancelamento dos níveis sonoros, dependendo do desempenho do sistema utilizado (NELSON e ELLIOT, 1995, KUO e MORGAN, 1996).

Os fundamentos desta técnica foram idealizados na década de 30 por Paul Lueg (GUICKING, 1990), porém apenas nos anos 50 uma maior atenção foi dada ao assunto. Atualmente, com o advento dos sistemas dedicados ao processamento de sinais, o controle de ruído acústico tem sido cada vez mais explorado (KUO e MORGAN, 1996, KUO *et al.*, 1996).

Até o surgimento do ANC, o controle de ruído era realizado apenas por meios passivos, como: forrações, espumas, amortecedores de absorção e reflexão, isoladores e silenciadores, entre outros. Porém, as técnicas passivas funcionam bem apenas para médias e altas frequências (acima de 500 Hz). Além disso, quando utilizadas para baixas frequências, soluções passivas podem ser volumosas e possuir peso considerável. Em contraposição, sistemas ativos são eficientes apenas abaixo de 1000 Hz (KUO *et al.*, 1996, ERIKSSON e ZUROSKI, 1997).

Os mecanismos físicos pelos quais a técnica de cancelamento de ruído pode ser

entendida são a interferência destrutiva e a variação da impedância de acoplamento. O cancelamento das ondas de pressão não é trivial e depende tanto da posição espacial como do instante de tempo.

De forma prática, o controle ativo do som é realizado através de sistemas eletroacústicos. Em sua forma mais simples, um sistema de controle (geralmente microprocessado) aciona um alto-falante de forma a produzir um campo sonoro que é uma imagem espelhada (antifase) exata do ruído do ambiente. Essa imagem espelhada é criada a partir da amostragem do ruído através de microfones, ou a partir de um sinal de referência obtido com o auxílio de outros transdutores. Na realidade essa tarefa é mais complicada e exige um grande esforço computacional, na maioria das vezes sem gerar o cancelamento completo do ruído.

Existem dois tipos básicos de enfoque para o controle ativo de ruído: cancelamento ativo de ruído acústico e controle ativo acústico-estrutural. No cancelamento de ruído acústico, os atuadores são fontes acústicas (alto-falantes) que produzem um sinal de som em antifase ao do ambiente, como já explicado. No caso do controle estrutural os atuadores são geralmente transdutores piezoelétricos que modificam o modo de vibração de uma estrutura, alterando, dessa maneira, o ruído irradiado (NELSON e ELLIOT, 1995, KUO e MORGAN, 1996).

O controle ativo de ruído tem um melhor desempenho para sons que possuem um comportamento espacial simples. Um exemplo clássico seria o de ondas de som de baixa frequência propagando-se através de um duto, um problema essencialmente unidimensional. O controle completo de sons com características espaciais complicadas ainda não é possível nos dias de hoje (GUO e PAN, 1995). Porém, as técnicas atuais são bastante eficientes para o controle de ruído em espaços fechados (NELSON *et al.*, 1987, BULLMORE *et al.*, 1987, ELLIOT *et al.*, 1987b), tais como em cabines de aeronaves e de veículos, e em dutos por onde fluem gases (BURGESS, 1981, MASSARANI, 1990, OSÓRIO e NÓBREGA, 1995, NELSON e ELLIOT, 1995, BERNHARD, 1995, KUO e MORGAN, 1996, KUO *et al.*, 1996, KIM *et al.*, 1997). Além do comportamento espacial

do ruído, outro fator de importância é o posicionamento adequado dos transdutores (BAEK e ELLIOT, 1995).

Um dos grandes problemas da redução de ruído em regiões de localização específica é a amplificação do ruído nas imediações do local em questão. Os sistemas de controle tendem a reduzir o ruído localmente ao invés de globalmente. Geralmente, a redução global é obtida apenas para campos sonoros simples, onde o mecanismo primário é o acoplamento de impedâncias. Entretanto, o cancelamento direcional é possível mesmo a elevadas frequências se os atuadores e o sistema de controle estiverem aptos a isso.

Na maioria das aplicações, a descrição completa do sistema a ser controlado não é possível. Além disso, tais sistemas são frequentemente não-estacionários devido a mudanças de temperatura em tubos de exaustão ou a modificações do ambiente causadas pela movimentação de pessoas ou máquinas. Assim, uma grande parte dos sistemas implementados faz uso de filtros adaptativos ao invés de técnicas de controle convencional. Dentre os algoritmos adaptativos mais utilizados pode-se citar o algoritmo LMS e sua variação, o FXLMS (KUO e MORGAN, 1996).

Diversas aplicações desta técnica têm sido sugeridas na literatura. Dentre essas pode-se citar:

- dutos de ventilação (BURGESS, 1981, MASSARANI, 1990);
- sistemas de ar-condicionado;
- eletrodomésticos (aspiradores, máquinas de lavar, refrigeradores,...);
- sistemas automotivos (ELLIOT *et al.*, 1988, BERNHARD, 1995, LY, 1997);
- estruturas espaciais (WOODARD e NAGCHAUDHURI, 1998);
- ambientes industriais;
- equipamentos de ressonância magnética (PLA *et al.*, 1995, CHEN *et al.*, 1999).

1.2.2 Tipos de Sistemas ANC

Pode-se classificar os sistemas ANC em quatro tipos principais, dependendo da forma de construção e dos objetivos que se propõem (NELSON e ELLIOT, 1995, KUO e MORGAN, 1996, KUO *et al.*, 1996):

- sistemas não-realimentados de banda-larga;
- sistemas não-realimentados de banda-estreita;
- sistemas realimentados;
- sistemas de múltiplos canais.

Apresenta-se a seguir um pequeno resumo sobre cada um dos tipos de sistema ANC, assim como seus diagramas em blocos.

- **Sistemas não-realimentados de banda-larga**

A figura 1.2 apresenta o diagrama simplificado de um cancelador não-realimentado de banda-larga. A fonte de ruído é amostrada por um microfone, produzindo um sinal elétrico de referência denominado $x(n)$. O controlador produz um sinal elétrico que aciona um alto-falante para produzir o cancelamento do sinal acústico indesejável. A resposta do controlador é ajustada a partir de $x(n)$ e do sinal de erro $e(n)$ (captado pelo microfone de erro) de modo a aumentar a eficiência do cancelamento.

- **Sistemas não-realimentados de banda-estreita**

Em aplicações onde a fonte de ruído é periódica (ou quase periódica), como é o caso de ruídos produzidos por máquinas rotativas, o microfone de entrada pode ser substituído por sensores não-acústicos, como tacômetros, acelerômetros ou sensores ópticos. A figura 1.3 mostra esse tipo de implementação. O sensor não-acústico fornece um sinal sincronizado com o ruído, permitindo a determinação de sua frequência fundamental. O sinal de erro permite o refinamento (adaptação) do sistema.

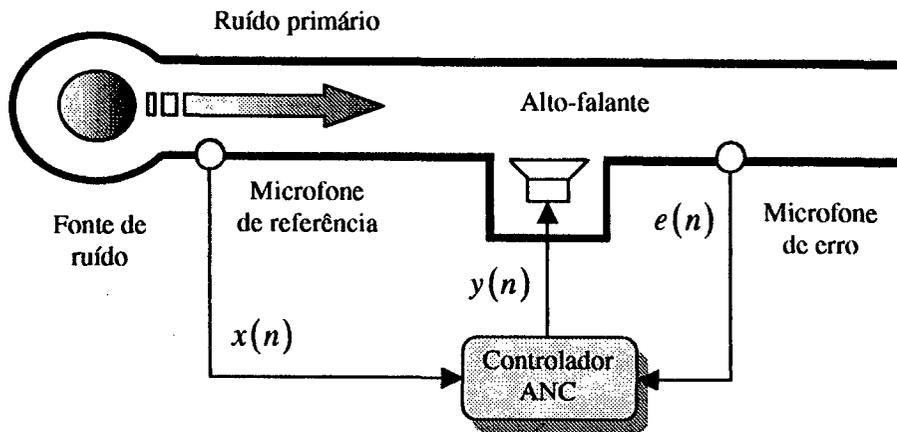


Figura 1.2 - Sistema de controle de ruído não-realimentado de banda-larga em um duto

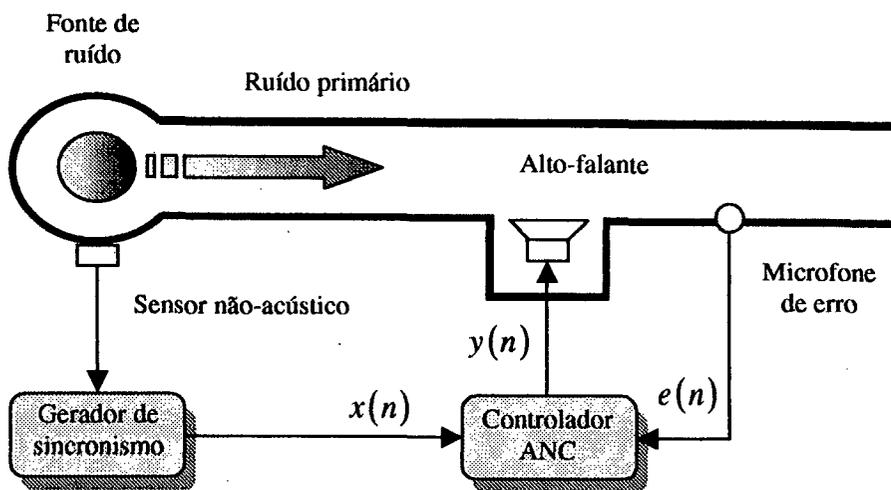


Figura 1.3 - Sistema ANC não-realimentado de banda-estreita

- **Sistemas realimentados**

Este tipo de configuração faz uso de apenas um transdutor acústico, o qual fornece o sinal de erro (figura 1.4). Em alguns casos, a redução do ruído indesejável requer um elevado ganho na malha de realimentação. Assim, os atrasos de fase podem tornar o sistema instável. Essa configuração é mais facilmente controlável no caso de controle de ruído de banda-estreita (KURO e MORGAN, 1996).

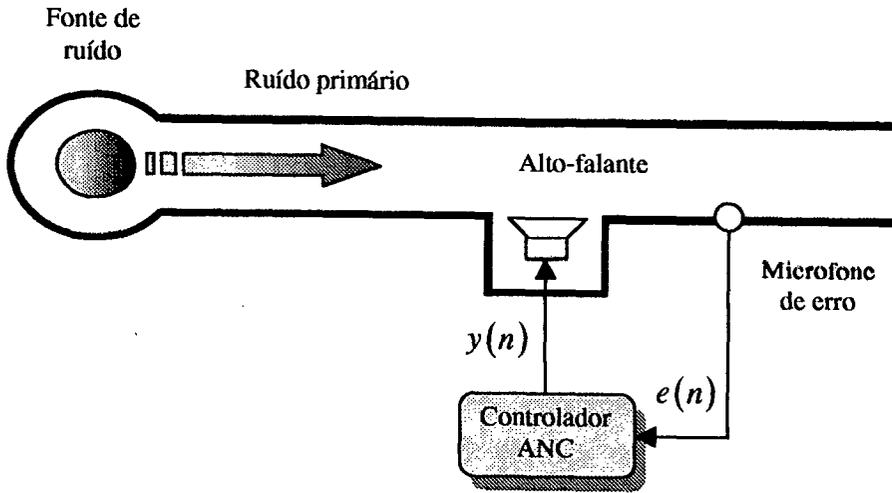


Figura 1.4 - Sistema realimentado de cancelamento de ruído

- **Sistemas de múltiplos canais**

Quando a geometria do problema é complexa, a utilização de apenas uma fonte secundária para o cancelamento do ruído não é suficiente. Nesse caso devem ser utilizados conjuntos de fontes associadas. Um sistema de múltiplos canais é apresentado na figura 1.5. Nesse caso existem diversos sinais de erro que devem ser combinados em uma função objetivo a ser minimizada (ELLIOTT *et al.*, 1992).

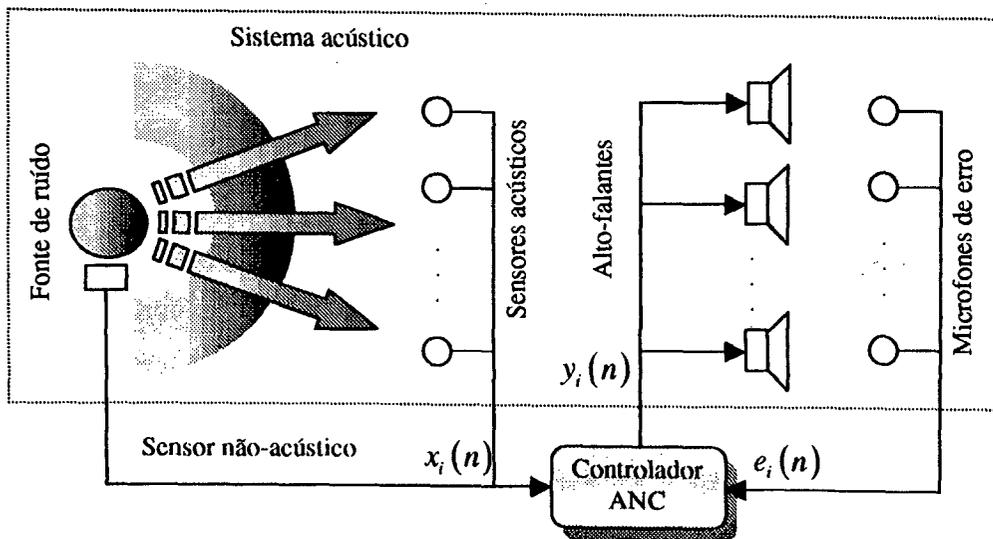


Figura 1.5 - Sistema de cancelamento de ruído de múltiplos canais

1.2.3 Algoritmos Adaptativos para Sistemas ANC

Diversas técnicas têm sido exploradas para a implementação de sistemas ANC. Entre elas pode-se citar:

- técnicas de controle convencional (controladores PID);
- técnicas de controle robusto (TOKHI e LEITCH, 1991);
- filtros adaptativos lineares (VIPPERMAN e BURDISSO, 1995);
- algoritmos genéticos (HANSEN *et al.*, 1996);
- redes neurais (HAYKIN, 1999).

Embora resultados satisfatórios possam ser obtidos em certas aplicações utilizando-se técnicas convencionais de controle, as formas adaptativas têm-se mostrado de grande valia, visto que a maioria dos sistemas ANC não pode ser completamente descrita em termos analíticos ou são não-estacionários.

A complexidade estrutural das redes neurais e dos algoritmos genéticos necessários à solução de problemas práticos dificulta sobremaneira a análise da dinâmica de seu funcionamento. Assim, apesar de fornecerem resultados relevantes, esses podem ser de difícil otimização.

Grande parte dos sistemas ANC é baseada em filtros adaptativos lineares. Essa popularidade deve-se à baixa complexidade estrutural e computacional desses filtros, as quais facilitam a análise de sua dinâmica, permitindo a compreensão, pelo projetista, das diversas variáveis envolvidas na obtenção da solução ótima.

Filtros adaptativos lineares IIR tendem a modelar sistemas físicos complexos com um número menor de coeficientes que os equivalentes FIR, resultando em menor custo computacional por iteração do algoritmo adaptativo. Entretanto, como grande revés, possuem características indesejáveis, como: convergência lenta, existência de mínimos locais e regiões de instabilidade. Em especial, em aplicações de controle de ruído acústico,

dependendo da aplicação, tanto estruturas FIR como IIR podem possuir um grande número de coeficientes em decorrência das características do meio acústico de propagação. Devido a esses fatos, a maioria das aplicações faz uso de filtros FIR (ERIKSSON *et al.*, 1987, LIAVAS e REGALIA, 1998).

Embora existam diferentes estruturas de implementação de filtros FIR (cascata, fase linear, *lattice*, entre outras), com diferentes propriedades, a forma transversal é a mais utilizada em decorrência de sua menor complexidade (custo computacional).

Diversos algoritmos podem ser encontrados na literatura para a implementação de sistemas ANC baseados em filtros adaptativos lineares com estrutura transversal. Dentre eles pode-se citar: o LMS, FXLMS e Leaky FXLMS. O presente trabalho restringe-se à análise do algoritmo FXLMS. O LMS é tratado como um caso particular do FXLMS.

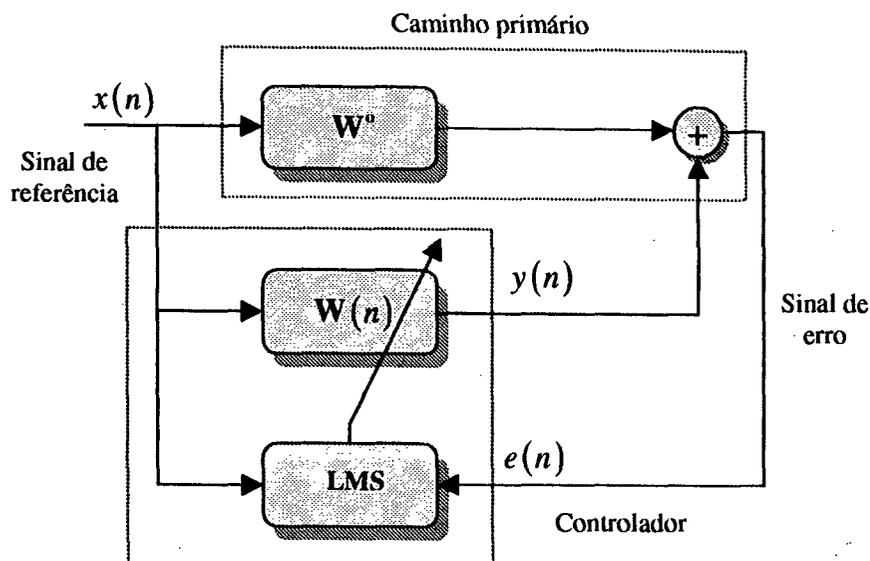


Figura 1.6 - Sistema básico de controle ativo de ruído acústico baseado no algoritmo LMS

A figura 1.6 apresenta o diagrama básico de um sistema de controle baseado no algoritmo LMS. O sistema utiliza um filtro adaptativo $W(n)$ para minimizar o erro médio quadrático na saída do módulo somador. Isso ocorre através da estimação da resposta do caminho primário W^o existente entre os sinais de referência e de erro.

De uma forma simplificada, a saída do filtro adaptativo deve convergir para um sinal próximo ao sinal na saída de W^o . Quando ambos os sinais são combinados, o erro residual tende a anular-se.

Visto que o cancelamento se dá por superposição das ondas de pressão, o sinal da segunda parcela da equação de atualização dos coeficientes (1.8) deve ser negativo.

Um grande problema desse modelo no caso de aplicações ANC é a existência do chamado **caminho acústico secundário** (CHEN *et al.*, 1996). O sinal gerado pelo filtro adaptativo é sujeito a uma série de modificações antes de ser utilizado para cancelar o ruído sonoro. Essas modificações são decorrentes da atuação do conjunto: conversor digital-analógico, pré-amplificadores, filtro de reconstrução, amplificadores de potência, alto-falantes, meio de propagação, microfones, filtros, amplificadores e conversor analógico-digital (KUO e MORGAN, 1996). Uma descrição mais representativa do sistema mencionado pode ser apreciada na figura 1.7.

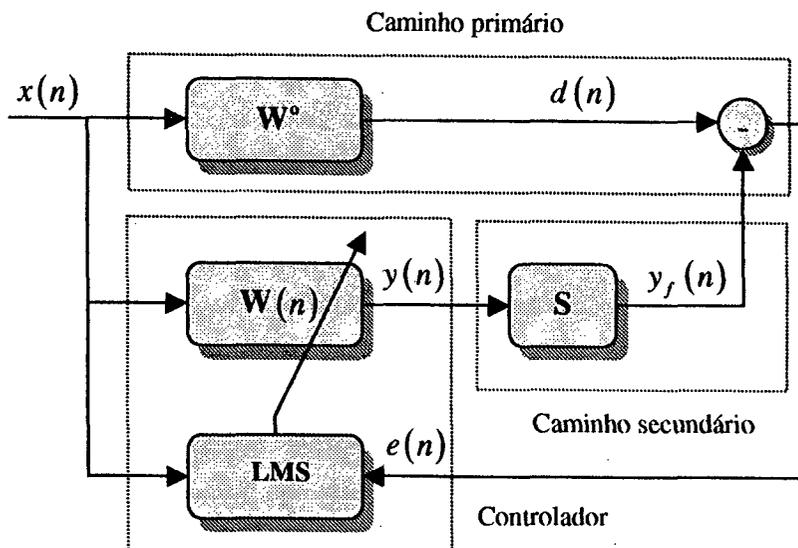


Figura 1.7 - Diagrama em blocos de um sistema básico de controle de ruído levando-se em consideração a influência do caminho acústico secundário.

Com o intuito de compensar a influência de S , o algoritmo FXLMS utiliza uma função \hat{S} , a qual representa uma estimativa do caminho secundário S (KUO e MORGAN,

1996, ZHANG, *et al.*, 2000). Basicamente, sua função é filtrar o sinal de excitação $x(n)$ antes de este ser utilizado na atualização dos coeficientes. Tanto melhor será o desempenho desse algoritmo quanto mais próxima for \hat{S} de S . Esse artifício é utilizado pelo fato de a função S não ser sempre fase mínima, ou seja, não existe necessariamente a função inversa S^{-1} . Dependendo da aplicação, S pode ser estimada antes da etapa de controle, ou pode ser estimada em tempo real durante a atuação do sistema de controle.

O diagrama em blocos da implementação prática de um sistema de controle ativo de ruído acústico em um duto, utilizando o algoritmo FXLMS, pode ser visto na figura 1.8.

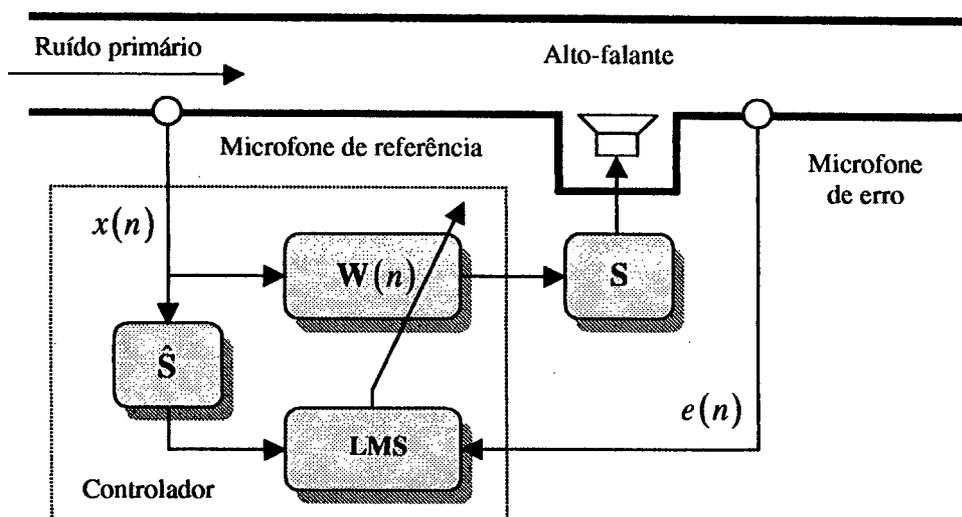


Figura 1.8 - Sistema de controle de ruído em um duto acústico utilizando o FXLMS

Supondo-se que S é linear, então quando $\hat{S} = S$ as características do sistema em regime permanente seriam equivalentes às de um sistema sem memória no caminho secundário, cujo caminho primário é definido pela convolução entre os sistemas W^o e S^{-1} , e sinal de excitação filtrado por S .

1.3 Não-Linearidades em Sistemas ANC

O comportamento de sistemas de controle de ruído baseados nos algoritmos LMS e

FXLMS pode ser estudado através da teoria de filtragem adaptativa. Diversos trabalhos estabelecem critérios de desempenho e estabilidade a partir de modelagens analíticas do comportamento desses sistemas. A maioria desses trabalhos assume que a propagação do som no ar ou de vibrações em estruturas pode ser descrita, em uma primeira instância, como um problema puramente linear. Entretanto, esses sistemas podem incluir componentes com características substancialmente não-lineares. Funções do tipo zona-morta, quantização ou saturação, entre outras, podem ser identificadas no meio físico, em transdutores ou no *hardware* associado (SUTTON e ELLIOT, 1995, TAO e KOKOTOVIC, 1996).

Para exemplificar, muitos dos transdutores (acústicos e de vibração) utilizados em sistemas ANC caracterizam-se por possuírem comportamento não-linear (KLIPPEL, 1990, KLIPPEL, 1992, SCOTT *et al.*, 1996, HEINLE *et al.*, 1998). Esse problema geralmente é minimizado, limitando-se as excursões de sinal relativamente à capacidade dos atuadores (ELLIOT e NELSON, 1993, SUYKENS *et al.*, 1995, HANSEN, 1997, KUO e MORGAN, 1999, DOUGLAS, 1999). Entretanto, por razões econômicas é de grande interesse que as especificações do sistema não sejam superdimensionadas. Um outro elemento não-linear de interesse é a etapa de amplificação de potência do sinal a ser entregue aos transdutores. Essa etapa pode acarretar distorção ou até mesmo saturação do sinal de controle.

Da mesma forma que sistemas ANC são afetados por essas não-linearidades, também diversas outras aplicações de controle/identificação adaptativa que utilizam os mesmos transdutores o são; por exemplo: cancelamento de eco (NOLLET e JONES, 1997, STENGER e RABENSTEIN, 1998, STENGER e KELLERMAN, 2000) e aplicações biomédicas (MELO *et al.*, 1998, MELO *et al.*, 2000).

Considerando a existência dessas influências não-lineares, torna-se importante o aprimoramento dos modelos analíticos existentes. Os novos modelos devem permitir a determinação dos efeitos de não-linearidades representativas do comportamento do sistema físico.

1.4 Filtros Adaptativos Sujeitos a Não-Linearidades

O estudo dos efeitos de não-linearidades no comportamento de algoritmos adaptativos ainda é incipiente. Entretanto, a influência de elementos não-lineares com características como saturação e quantização é de grande interesse por parte do projetista.

Freqüentemente, efeitos não-lineares são modelados por aproximações linearizadas (GITLIN *et al.*, 1973, CARAISCOS e LIU, 1984, ALEXANDER, 1987). Alguns trabalhos recentes desenvolvem modelos não-lineares que consideram a natureza exata da não-linearidade envolvida (BERMUDEZ e BERSHAD, 1996a e 1996b, BERSHAD e BERMUDEZ, 1996a e 1996b). Uma técnica eficiente de análise utiliza não-linearidades do tipo saturação como aproximação analítica de não-linearidades reais (DUTTWEILER, 1982, BERSHAD, 1986, BERSHAD, 1988, SHYNK e BERSHAD, 1991, BERSHAD *et al.* 1993a e 1993b, BERSHAD *et al.*, 1999). Outros trabalhos que devem ser mencionados são: ROWE (1982), DOUGLAS e MENG (1994a) e (1994b).

A utilização de funções não-lineares do tipo saturação possibilita uma adequada modelagem das distorções por limitação de excursão de sinal, tanto em nível de *software* (*overflow*) quanto de *hardware*. Funções desse tipo são úteis também na análise de redes neurais. Muitas dessas redes podem ser entendidas como filtros adaptativos com não-linearidades em suas saídas. A facilidade de manipulação analítica de funções de saturação pode viabilizar a análise de estruturas bastante complexas.

A utilização dessa técnica produziu interessantes conclusões na análise do comportamento do neurônio artificial (SHYNK e BERSHAD, 1991, BERSHAD *et al.*, 1993a e 1993b, BERSHAD *et al.*, 1996, BERSHAD *et al.* 1997), na identificação de sistemas não-lineares com memória (BERSHAD *et al.*, 1999, BERSHAD *et al.*, 2000) e no LMS sujeito à saturações do sinal de erro (BERSHAD, 1988).

Apesar de todos esses trabalhos, até o momento desconhecem-se esforços no sentido de desenvolver modelos analíticos capazes de prever a degradação de desempenho

em sistemas ANC baseados no algoritmo LMS decorrente da influência de não-linearidades.

1.5 Objetivos deste Trabalho

Estudar o comportamento de algoritmos adaptativos baseados no algoritmo LMS sujeitos à influência de não-linearidades através de uma metodologia de análise capaz de fornecer modelos analíticos que possam prever, com adequada precisão, o comportamento de sistemas reais tanto em regime transitório como permanente. Em um segundo momento determinar figuras de mérito capazes de qualificar/quantificar o desempenho desses sistemas em condições determinadas, auxiliando o projetista de sistemas adaptativos.

1.6 Justificativa e Relevância

Diversos sistemas de interesse do ponto de vista da filtragem adaptativa possuem características não-lineares; dentre esses pode-se citar: sistemas biológicos, controle ativo de ruído acústico, redes neurais e sistemas de identificação e controle de um modo geral.

Principalmente nas áreas de controle ativo de ruído e identificação de sistemas, a presença de componentes não-lineares (transdutores e amplificadores) é uma constante. A modelagem analítica usual do comportamento de algoritmos adaptativos desconsidera a influência dessas não-linearidades, o que produz resultados nem sempre condizentes com a situação real.

A literatura técnica especializada apresenta diversos estudos sobre o comportamento de sistemas não-lineares. No entanto, poucos desses estudos podem ser utilizados (ainda que aproximadamente) para modelar a influência de não-linearidades em sistemas de controle ativo de ruído e vibrações que utilizam algoritmos baseados no LMS.

A metodologia utilizada pode ser de grande valia também para o estudo de redes

neurais artificiais.

1.7 Estrutura do Trabalho

Inicialmente, foi apresentada a fundamentação teórica necessária à análise do problema. No capítulo 2 será feita a definição formal do problema genérico abordado neste trabalho. Dessa formulação geral podem ser extraídos, como casos particulares, uma série de problemas de filtragem adaptativa extensivamente estudados na literatura.

O capítulo 3 apresenta a análise da superfície de desempenho do sistema de controle sob o ponto de vista de um problema de estimação média quadrática não-linear. Os capítulos 4, 5 e 6 apresentam a análise estatística propriamente dita, resultando na obtenção de modelos analíticos para o comportamento médio dos coeficientes, evolução do erro médio quadrático e momentos de segunda ordem do vetor de coeficientes. Mostra-se que para o caso limite em que a influência da não-linearidade é desprezível, as expressões obtidas neste trabalho reduzem-se às expressões já descritas na literatura para os casos particulares analisados, supondo-se sistemas puramente lineares. Nesse sentido, o presente trabalho generaliza a modelagem do comportamento estocástico de uma grande classe de sistemas adaptativos lineares e não-lineares.

No capítulo 7 são avaliados os modelos analíticos e apresentadas confrontações entre os resultados teóricos e os resultados de simulações computacionais. No capítulo 8 são feitas observações gerais sobre o trabalho e apresentadas conclusões e propostas para sua continuação. Nos anexos encontram-se vários desenvolvimentos matemáticos descritos ao longo dos capítulos.

2. DEFINIÇÃO DO PROBLEMA

Neste capítulo é formalizada a descrição do problema a ser analisado. São apresentados os modelos analíticos e as principais considerações estatísticas e funcionais que serão utilizados ao longo do trabalho.

2.1 Introdução

Como já descrito no capítulo 1, o algoritmo LMS é o algoritmo mais popular em aplicações de tempo real e tem sido utilizado com sucesso em inúmeras áreas. Entretanto, certos sistemas caracterizam-se por introduzir operações de filtragem no caminho de realimentação do algoritmo (saída do filtro adaptativo). Em aplicações de controle ativo de ruído acústico isso ocorre devido à presença do sistema de atuação acústica (caminho acústico secundário).

Os elementos do caminho acústico secundário criam atrasos que podem causar a degradação do desempenho ou até mesmo a instabilidade do processo de adaptação. Para resolver esse problema foi desenvolvida uma versão modificada do LMS denominada de algoritmo LMS Filtrado (*Filtered-X Least Mean Square*), também chamado de FXLMS (BURGESS, 1981, WIDROW *et al.*, 1981, BJARNASON, 1995, KUO e MORGAN, 1996).

Nesse algoritmo, o sinal de referência é filtrado de modo a compensar os filtros que se encontram dentro do laço de realimentação eletroacústico. Essa filtragem adicional no sistema adaptativo gera correlações entre os sinais e, por conseguinte, complica de forma significativa o comportamento e análise do algoritmo adaptativo. Essas correlações invalidam o uso da teoria da independência (MAZO, 1979), muito utilizada para simplificar a análise estatística do comportamento do algoritmo LMS. Portanto, a análise desse algoritmo torna-se bastante complexa.

Diversos trabalhos tratam da análise do algoritmo FXLMS. SNYDER e HANSEN (1994) avaliaram o comportamento médio dos coeficientes e, posteriormente, BJARNASON (1995) desenvolveu expressões para os momentos de primeira e segunda ordem para o caso de estimação correta do caminho secundário, especializando seus resultados para o caso do algoritmo *Delayed LMS* (caso particular do FXLMS). Entretanto, essas análises basearam-se na teoria da independência. Recentemente, TOBIAS (1999) desenvolveu um modelo analítico mais acurado para o comportamento médio e para os momentos de segunda ordem dos coeficientes do algoritmo FXLMS sem recorrer à teoria da independência.

Outra importante característica de sistemas ANC é a existência de efeitos não-lineares no laço de realimentação do algoritmo em decorrência do comportamento dos amplificadores e transdutores utilizados. Em termos de análise, essa característica é freqüentemente desconsiderada, supondo-se que o sistema seja puramente linear (BERNHARD *et al.*, 1997). Para isso, uma série de precauções deveria ser tomada, como: limitação das excursões dos sinais dentro da região linear dos amplificadores de potência e escolha de atuadores estritamente lineares. Entretanto, isso nem sempre é possível, devido tanto a aspectos econômicos como a operacionais.

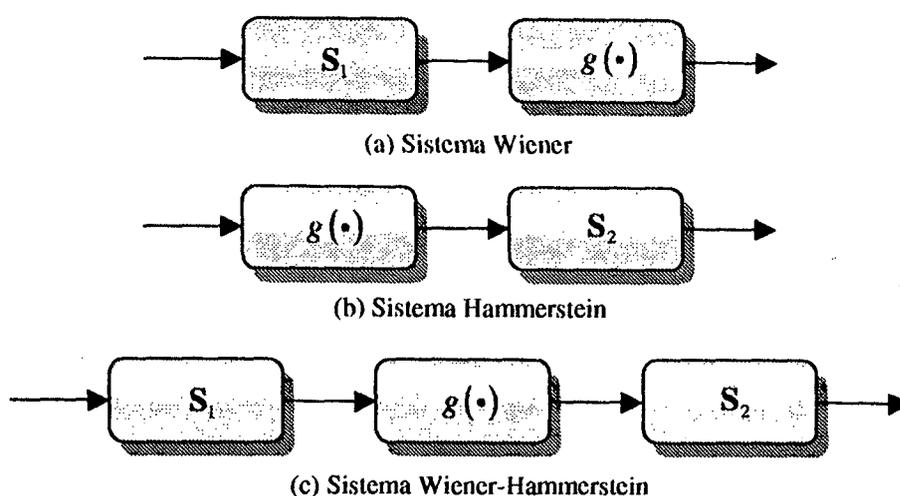


Figura 2.1 – (a) Sistema Wiener; (b) sistema Hammerstein e (c) sistema Wiener-Hammerstein. $g(\cdot)$: função não-linear; S_i : sistema linear

Em decorrência desses fatores, uma modelagem mais adequada de um sistema ANC seria a utilização de um **Sistema Wiener-Hammerstein** (BERSHAD *et al.*, 2000) para o caminho acústico secundário.

O sistema Wiener-Hammerstein (WH) constitui-se da composição de dois modelos dinâmicos lineares interpostos por uma função estática não-linear (AGUIRRE, 2000). Um sistema desse tipo recebe tal nome devido ao fato de que no modelo de Hammerstein a não-linearidade precede a parte linear, enquanto que no de Wiener ocorre o contrário (figura 2.1).

A figura 2.2 apresenta o diagrama em blocos de um sistema ANC utilizando o algoritmo FXLMS e um sistema WH no caminho acústico secundário.

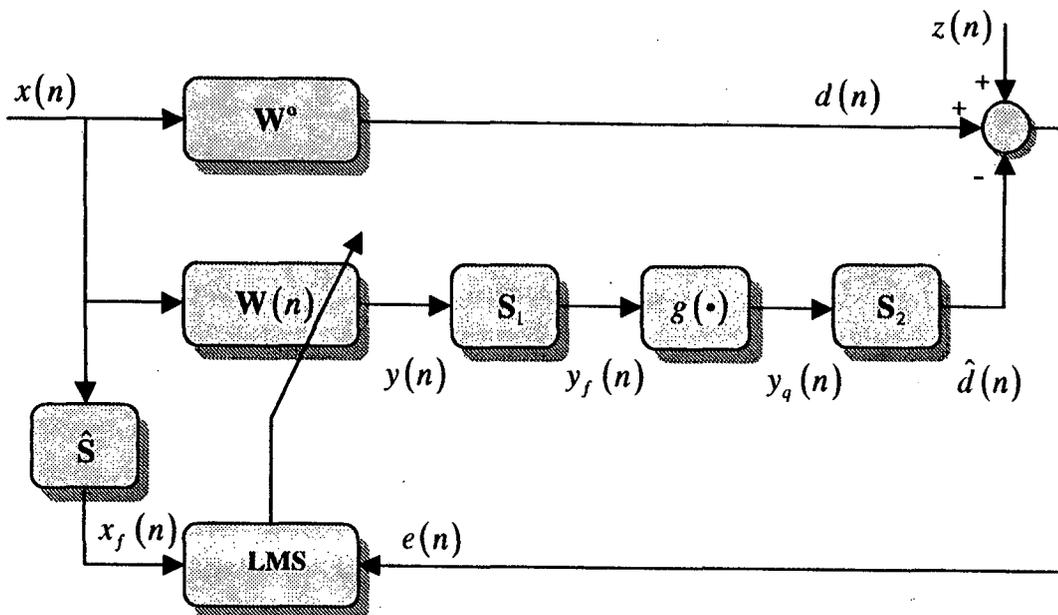


Figura 2.2 - Topologia analisada: algoritmo FXLMS Wiener-Hammerstein (FXLMS-WH).

Os sinais e sistemas envolvidos na descrição do algoritmo FXLMS-WH (figura 2.2) são apresentados a seguir:

- n : instante de tempo.
- $x(n)$: sinal de referência ou de excitação.
- $x_f(n)$: sinal de referência filtrado.
- $e(n)$: sinal de erro.
- $d(n)$: sinal desejado (sinal primário ou saída da planta).
- $z(n)$: ruído de medição (aditivo).
- $y(n)$: saída do filtro adaptativo (sinal de controle).
- $y_f(n)$: saída do filtro adaptativo filtrada (sinal de controle filtrado).
- $y_q(n)$: saída da função não-linear.
- $\hat{d}(n)$: estimativa do sinal desejado (sinal de cancelamento).
- $\mathbf{W}^o = [w_0^o \quad w_1^o \quad \dots \quad w_{N-1}^o]^T$: resposta ao impulso da planta.
- $\mathbf{S}_1 = [s_{1,0} \quad s_{1,1} \quad \dots \quad s_{1,M_1-1}]^T$: primeiro filtro do caminho secundário.
- $\mathbf{S}_2 = [s_{2,0} \quad s_{2,1} \quad \dots \quad s_{2,M_2-1}]^T$: segundo filtro do caminho secundário.
- $\hat{\mathbf{S}} = [\hat{s}_0 \quad \hat{s}_1 \quad \dots \quad \hat{s}_{\hat{M}-1}]^T$: estimativa da parte linear do caminho secundário.
- $\mathbf{W}(n) = [w_0(n) \quad w_1(n) \quad \dots \quad w_{N-1}(n)]^T$: vetor de coeficientes adaptativos.
- $\mathbf{X}(n) = [x(n) \quad x(n-1) \quad \dots \quad x(n-N+1)]^T$: vetor de dados observados.
- $\mathbf{X}_f(n) = [x_f(n) \quad x_f(n-1) \quad \dots \quad x_f(n-\hat{M}+1)]^T$: vetor de dados observados filtrados.
- $g(\bullet)$: função não-linear.

Os vetores $\mathbf{W}(n)$ e \mathbf{W}^o são considerados de mesma dimensão N . Quando de diferentes comprimentos, podem ser acomodados com a inserção de zeros.

Na figura 2.2, o sistema S_1 representa a influência da “parte elétrica” do sistema (filtros, amplificação, conversão D/A), enquanto que S_2 representa os efeitos do caminho de propagação do sinal no meio físico, assim como da “parte elétrica” de aquisição (amplificação, filtros, conversão A/D). O sistema $g(\cdot)$ representa efeitos não-lineares do tipo saturação.

Modificações impostas ao sinal de erro, desde que lineares, podem ser representadas de forma artificial através de sua inclusão como parte dos sistemas W^o e S_2 . A expansão do modelo apresentado para o estudo de sistemas multicanal também é direta (ELLIOT *et al.*, 1987a, SNYDER e HANSEN, 1992, WANG e REN, 1999).

A vantagem do modelo proposto é que, com a popularização de sistemas ANC, implementações de baixo custo poderiam ser mais bem projetadas, tanto do ponto de vista operacional como econômico, visto que atualmente a maioria dos projetistas que utilizam o algoritmo LMS e seus derivados despreza efeitos não-lineares na análise de seu comportamento.

A partir do modelo descrito pretende-se estudar o comportamento do algoritmo FXLMS na existência de efeitos não-lineares no caminho acústico secundário. A influência da não-linearidade é modelada como uma função não-linear de saturação atuando sobre o sinal de controle filtrado.

2.2 Algoritmo FXLMS Sujeito a um Sistema WH

A equação de atualização dos coeficientes do algoritmo FXLMS é dada por (HAYKIN, 1991, KUO e MORGAN, 1996):

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu e(n) \mathbf{X}_f(n) \quad (2.1)$$

onde μ é denominado de passo de convergência ou constante de adaptação (mantido constante durante todo o processo de adaptação) e o vetor de dados filtrado é definido como

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

Portanto, o sinal de erro é dado por

$$\begin{aligned} e(n) &= d(n) + z(n) - \hat{d}(n) \\ &= \mathbf{W}^{oT} \mathbf{X}(n) + z(n) - \sum_{j=0}^{M_2-1} s_{2j} g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \end{aligned} \quad (2.3)$$

Substituindo-se a equação (2.3) em (2.1), tem-se:

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) + \mu \left\{ \mathbf{W}^{oT} \mathbf{X}(n) + z(n) \right. \\ &\quad \left. - \sum_{j=0}^{M_2-1} s_{2j} g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \right\} \mathbf{X}_f(n) \end{aligned} \quad (2.4)$$

A equação (2.4) é a equação de atualização dos coeficientes do algoritmo FXLMS para o caso de um caminho secundário modelado por um sistema Wiener-Hammerstein.

2.3 Características dos Sinais e Dinâmica do Sistema

Considerações simplificadoras são frequentemente necessárias para assegurar a tratabilidade matemática do problema de análise estatística do comportamento de algoritmos adaptativos. Entretanto, as experiências de vários autores nos mostram que os resultados obtidos usando certas considerações podem ser utilizados para prever o comportamento desses algoritmos em condições reais de operação. Logo, os resultados assim obtidos servem como confiáveis guias de projeto (DUTTWEILER, 1982, SHYNK e BERSHAD, 1991, HAYKIN, 1991, BERMUDEZ e BERSHAD, 1996a e 1996b).

Segundo DUTTWEILER (1982), uma análise sob condições mais restritivas permite desenvolver o **sentimento do projetista** quanto ao desempenho de algoritmos adaptativos sob condições reais, e portanto de maior complexidade de análise.

Ao longo deste trabalho, as seguintes hipóteses simplificadoras serão impostas progressivamente:

- o sinal de excitação $x(n)$ é estacionário, gaussiano com média zero e variância σ_x^2 ;
- o ruído aditivo $z(n)$ é estacionário, branco, gaussiano com média zero e variância σ_z^2 e independente de qualquer outro sinal;
- o passo de convergência é suficientemente pequeno de modo que o processo de adaptação ocorra de forma lenta;
- as correlações entre $\mathbf{X}(n-k)$ e $\mathbf{X}(n-l)$ para quaisquer k e l são muito mais relevantes que as correlações entre $\mathbf{X}(n-k)$ e $\mathbf{W}(n-p)$ para quaisquer k e p .

Embora a análise realizada neste trabalho utilize a consideração de que os sinais possam ser correlacionados, o caso particular em que se supõe que os sinais sejam brancos também é de especial interesse.

A suposição de sinais brancos é comum na análise do comportamento de algoritmos adaptativos. Como justificativa para o uso do ruído branco, BERSHAD (1988) cita:

- a característica de ruído branco é representativa de sistemas analógicos de banda-larga amostrados a taxas próximas da de Nyquist e adaptados de forma discreta, como ocorre em sistemas de radar e sonar adaptativos;

- os resultados analíticos dessa suposição são representativos do desempenho relativo de algoritmos para cancelamento de eco em sinais de voz, apesar de os mesmos não serem brancos;
- em transmissão digital de dados, a seqüência de dígitos binários pode ser considerada independente e, conseqüentemente, o modelo de ruído branco é representativo.

A suposição de distribuição gaussiana justifica-se freqüentemente através do teorema do limite central (PAPOULIS, 1991).

Vários autores (WIDROW *et al.*, 1975, BERSHAD, 1988, HAYKIN, 1991) fazem menção à **Teoria da Independência** (MAZO, 1979). Essa teoria supõe que o vetor de amostras $\mathbf{X}(n)$ seja estatisticamente independente ao longo do tempo (embora, em implementações que utilizam um filtro transversal operando como linha de retardo, a cada iteração apenas uma amostra seja atualizada e as outras deslocadas de posição ao longo do vetor). Em decorrência dessa suposição os vetores $\mathbf{X}(n)$ e $\mathbf{W}(n)$ seriam estatisticamente independentes no algoritmo LMS, acarretando uma grande simplificação do processo de análise.

Essa hipótese simplificadora produz bons resultados na previsão do comportamento do algoritmo LMS com passo de adaptação pequeno. No entanto, seu emprego na análise de outros algoritmos deve ser avaliado cuidadosamente. Em alguns casos, o emprego da teoria da independência pode levar a modelos analíticos bastante imprecisos. Este é o caso, por exemplo, do algoritmo FXLMS (TOBIAS, 1999 e TOBIAS *et al.*, 2000).

Neste trabalho não se utiliza a teoria da independência, visto que se considera a correlação entre $\mathbf{X}(n-k)$ e $\mathbf{X}(n-l)$.

2.4 Influência Linear do Caminho Secundário

A existência de operações de filtragem no caminho secundário acarreta um aumento na correlação do sinal de cancelamento ($\hat{d}(n)$), modificando o comportamento do algoritmo LMS. Análises realizadas à luz da teoria da independência desprezam essas correlações. Embora os resultados obtidos ainda mantenham alguma informação acerca da dinâmica do processo de adaptação, TOBIAS (1999) demonstrou que resultados muito mais acurados são obtidos quando essas correlações são levadas em consideração.

No caso de uma implementação utilizando um filtro transversal, a correlação entre dois vetores de dados de entrada quaisquer pode ser entendida intuitivamente através da comparação entre dois vetores consecutivos:

$$\mathbf{X}(n-1) = \begin{bmatrix} x(n-1) \\ x(n-2) \\ \dots \\ x(n-N) \end{bmatrix} \quad \mathbf{X}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \dots \\ x(n-N+1) \end{bmatrix} \quad (2.5)$$

Através de (2.5) verifica-se que ambos os vetores compartilham $N-1$ componentes e portanto possuem um alto grau de similaridade. À medida que o intervalo de tempo entre os dois vetores aumenta, diminuirão progressivamente os elementos comuns a ambos os vetores. De uma forma geral tem-se que

$$\mathbf{X}(n-l) = \begin{bmatrix} x(n-l) \\ x(n-l-1) \\ \dots \\ x(n-N+1-l) \end{bmatrix} \quad \mathbf{X}(n-k) = \begin{bmatrix} x(n-k) \\ x(n-k-1) \\ \dots \\ x(n-N+1-k) \end{bmatrix} \quad (2.6)$$

A correlação entre os vetores em (2.6) é especificada pela matriz de correlação:

$$\mathbf{R}_{k-l} = E\{\mathbf{X}(n-l)\mathbf{X}^T(n-k)\} \quad (2.7)$$

ou, na forma expandida:

$$\mathbf{R}_{k-l} = \begin{bmatrix} E\{x(n-l)x(n-k)\} & E\{x(n-l)x(n-1-k)\} & \dots & E\{x(n-l)x(n-N+1-k)\} \\ E\{x(n-1-l)x(n-k)\} & E\{x(n-1-l)x(n-1-k)\} & \dots & E\{x(n-1-l)x(n-N+1-k)\} \\ \dots & \dots & \dots & \dots \\ E\{x(n-N+1-l)x(n-k)\} & E\{x(n-N+1-l)x(n-1-k)\} & \dots & E\{x(n-N+1-l)x(n-N+1-k)\} \end{bmatrix} \quad (2.8)$$

Considerando a estacionaridade de $x(n)$ em (2.8), as seguintes características da matriz de correlação \mathbf{R}_c podem ser verificadas para $c = k - l$, a partir da matriz de autocorrelação \mathbf{R}_0 :

- $\mathbf{R}_c = \mathbf{R}_{-c}^T$;
- se $c > 0$ então \mathbf{R}_c pode ser obtida através de c deslocamentos das colunas de \mathbf{R}_0 para a esquerda, preenchendo-se as novas colunas com os próximos elementos da função de correlação em cada linha;
- se $c < 0$ então \mathbf{R}_c pode ser obtida através de c deslocamentos das colunas de \mathbf{R}_0 para a direita, preenchendo-se as novas colunas com os próximos elementos da função de correlação em cada linha.

Segundo a teoria da independência, matrizes de correlação \mathbf{R}_c seriam nulas, exceto para $c = 0$. Ou seja, a suposição de que os vetores de dados de entrada são independentes para instantes de tempo diferentes faz com que informação acerca do comportamento do sistema seja desprezada, mesmo quando o sinal de entrada é branco.

Como exemplo, para uma ordem igual a 4, apresentam-se as seguintes matrizes de correlação:

$$\begin{aligned}
 \mathbf{R}_0 &= \begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} & ; & \mathbf{R}_1 = \begin{bmatrix} r_1 & r_2 & r_3 & r_4 \\ r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \end{bmatrix} \\
 \mathbf{R}_2 &= \begin{bmatrix} r_2 & r_3 & r_4 & r_5 \\ r_1 & r_2 & r_3 & r_4 \\ r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \end{bmatrix} & ; & \mathbf{R}_{-2} = \begin{bmatrix} r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \\ r_4 & r_3 & r_2 & r_1 \\ r_5 & r_4 & r_3 & r_2 \end{bmatrix}
 \end{aligned} \tag{2.9}$$

onde $r_{k-l} = E\{x(n-l)x(n-k)\}$.

TOBIAS (1999) demonstrou que a informação pertinente à correlação entre vetores de dados de entrada é importante para a descrição da dinâmica do algoritmo FXLMS. De forma inversa, devido à complexidade do tratamento matemático, as correlações entre vetores de dados e vetores de coeficientes foram desprezadas. Diversas simulações foram realizadas e demonstraram a validade dessa proposição para passos suficientemente pequenos.

2.5 Modelo Analítico da Não-Linearidade

Diversas formas de modelagem para sistemas não-lineares têm sido propostas na literatura (SCHETZEN, 1989, TAO e KOKOTOVIC, 1996, HAYKIN, 1999). Do ponto de vista de modelagem matemática, é desejável que a função (modelo) analítica(o) que representa a não-linearidade possua as seguintes características fundamentais:

- a função $g(\bullet)$ deve estar relacionada ao fenômeno físico que pretende modelar.

Neste trabalho deve representar analiticamente as características não-lineares impostas por um transdutor (por exemplo um alto-falante) e/ou devidas à saturação e distorção decorrentes da operação de amplificadores de potência em

regiões não-lineares, ou ainda à função de ativação de um neurônio artificial (SHYNK e BERSHAD, 1991);

- a função $g(\bullet)$ deve ser construída de tal forma que possibilite a tratabilidade matemática necessária dentro do conjunto de técnicas utilizadas.

Uma função não-linear bastante difundida e amplamente utilizada para representar o tipo de característica física requerida é uma variação da **função erro** (BERSHAD, 1988, BERSHAD, 1990), determinada pela equação:

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (2.10)$$

onde σ^2 é denominado de coeficiente de não-linearidade.

Essa função atende aos dois requisitos necessários (representação do fenômeno físico e tratabilidade matemática). Isso pode ser constatado através da figura 2.3, onde se pode verificar a variação de seu comportamento desde linear até altamente não-linear em função do parâmetro σ^2 e pelas seguintes características:

$$\left\{ \begin{array}{l} \lim_{\sigma^2 \rightarrow \infty} g(y) = y \quad (2.11a) \\ \lim_{\sigma^2 \rightarrow 0} g(y) = \sigma \frac{\sqrt{\pi}}{\sqrt{2}} \text{sinal}(y) \quad (2.11b) \\ \frac{\partial g(y)}{\partial y} = e^{-\frac{y^2}{2\sigma^2}} \quad (2.11c) \end{array} \right.$$

onde $\text{sinal}(\bullet)$ é a função sinal (1 se $\rho > 0$ ou -1 se $\rho < 0$). Outras características convenientes de $g(\bullet)$ são:

- não alterar o valor médio de variáveis com média zero;
- não possuir memória;
- facilitar a determinação de esperanças estatísticas envolvendo funções densidade de probabilidade gaussianas.

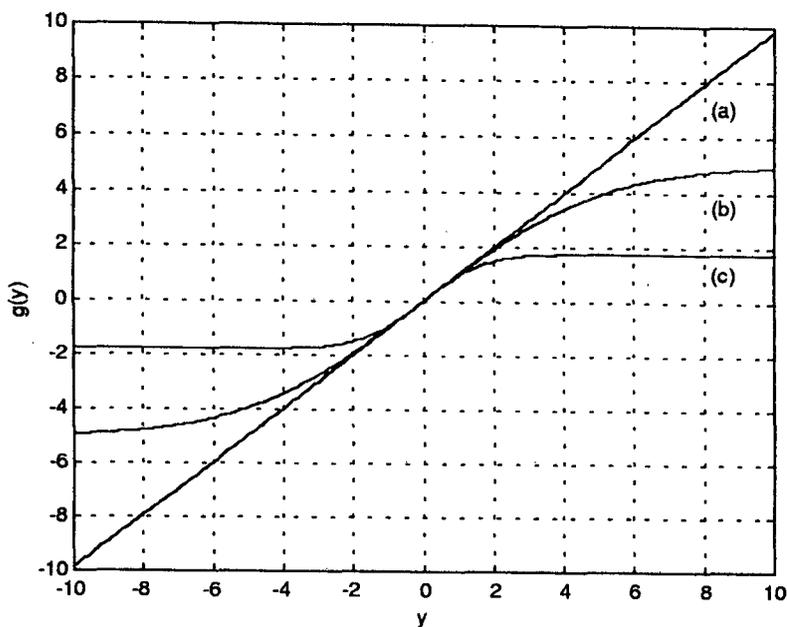


Figura 2.3 - Gráfico da função erro para: (a) $\sigma^2 = 1000$; (b) $\sigma^2 = 16$ e (c) $\sigma^2 = 2$.

Note que para o caso particular (2.11a), onde $\sigma^2 \rightarrow \infty$, a equação (2.4) torna-se idêntica à do algoritmo FXLMS convencional (HAYKIN, 1991).

A multiplicação de $g(\cdot)$ por uma constante arbitrária (escalamento) permite que essa função possa ser ajustada da melhor forma possível ao comportamento real da não-linearidade do sistema, tornando-a mais representativa do fenômeno estudado. Um caso interessante é quando essa constante multiplicativa é função de σ^2 . Denominada neste trabalho de função erro modificada $f(\cdot)$, sua equação é dada por

$$f(y) = \frac{1}{\sigma} \sqrt{\frac{\pi}{2}} \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (2.12)$$

Como resultado desse escalamento, tem-se que

$$\lim_{\sigma^2 \rightarrow \infty} f(y) = \sqrt{\frac{2}{\pi}} \frac{y}{\sigma} \quad (2.13a)$$

$$\lim_{\sigma^2 \rightarrow 0} f(y) = \text{sinal}(y) \quad (2.13b)$$

$$\frac{\partial f(y)}{\partial y} = \frac{1}{\sigma} \sqrt{\frac{\pi}{2}} e^{-\frac{y^2}{2\sigma^2}} \quad (2.13c)$$

Como pode ser verificado através do conjunto das equações (2.13) e da figura 2.4, a escolha do parâmetro σ^2 permite que o comportamento da função erro modificada varie do caso linear até o de uma função sinal (± 1).

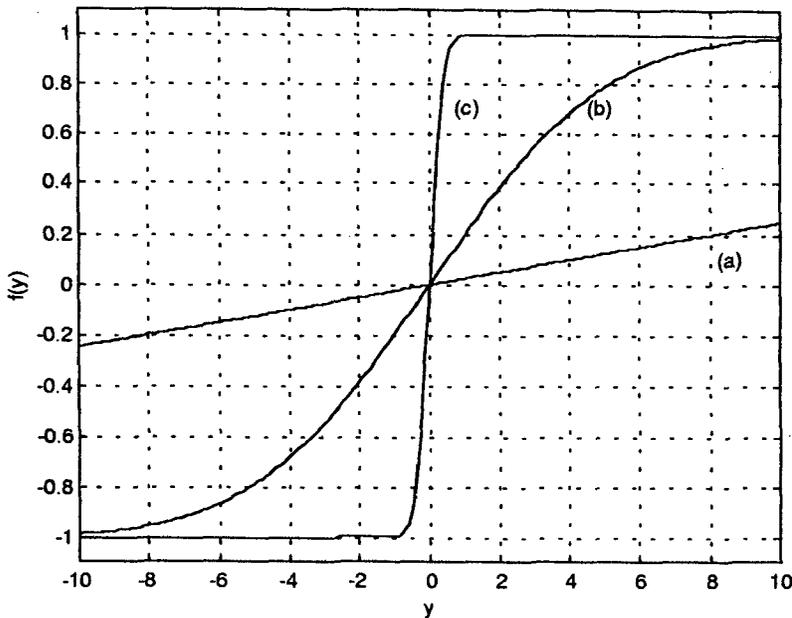


Figura 2.4 - Gráfico da função erro modificada para:

(a) $\sigma^2 = 1000$; (b) $\sigma^2 = 16$ e (c) $\sigma^2 = 0,1$.

No trabalho de SHYNK e BERSHAD (1991) é apontada como importante característica da função erro modificada também o fato de ela, para uma escolha adequada de parâmetros, produzir resultados bastante similares aos da função tangente hiperbólica, comumente utilizada como função de ativação na área de redes neurais (figura 2.5).

A função erro modificada aparece em diversos trabalhos que tratam da análise do comportamento de redes neurais do tipo Perceptron de uma camada (SHYNK e BERSHAD, 1991, BERSHAD *et al.*, 1993a, BERSHAD *et al.*, 1993b).

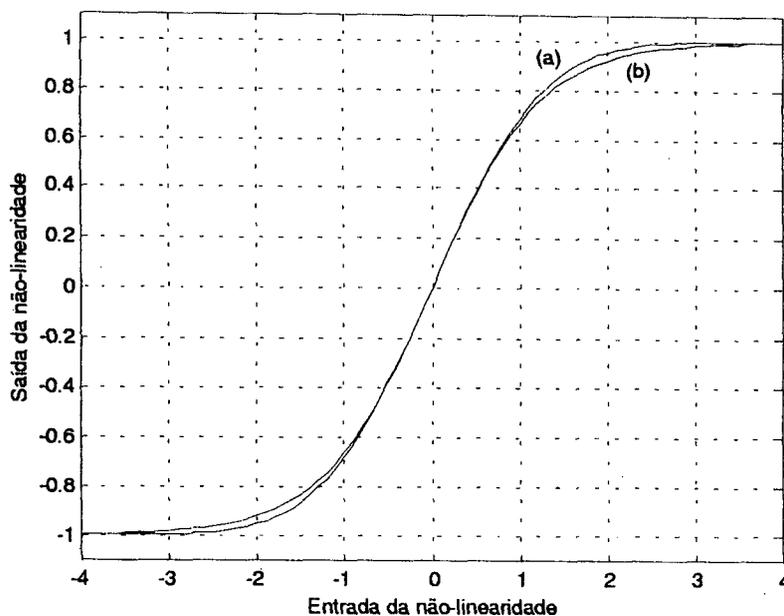


Figura 2.5 - Comparação entre (a) função erro modificada e (b) tangente hiperbólica:

$$y = \tanh\left(\frac{1}{\sigma} \sqrt{\frac{2}{\pi}} x\right), \text{ para } \sigma^2 = 1.$$

Neste trabalho, a característica não-linear do sistema ANC será modelada pela função erro (equação (2.10)). Essa função é a mais adequada para modelar condições de não-linearidade relativamente reduzidas, as quais constituem a situação mais importante em aplicações práticas. Sistemas com alto grau de não-linearidade ($\sigma^2 \rightarrow 0$) podem ser estudados, multiplicando-se a função $g(\cdot)$ por uma constante do tipo A/σ , onde $A \in \mathbb{R}^+$.

Tal artifício evita o aparecimento do fator de atenuação σ na equação (2.11b). Nesse caso, as equações desenvolvidas ao longo deste trabalho são facilmente modificadas através da propagação dessa constante ao longo dos equacionamentos.

2.6 Grau de Não-Linearidade do Sistema ANC

Esta seção apresenta a definição do chamado **grau de não-linearidade do sistema adaptativo**. Esta figura de mérito, utilizada posteriormente, avalia quantitativamente o quanto o processo de adaptação é influenciado pela característica não-linear intrínseca ao sistema.

O grau de não-linearidade do sistema adaptativo é assim definido:

Supondo-se que (1) a função não-linear seja suprimida do sistema adaptativo, de forma que este se torne completamente linear, (2) que exista convergência dos coeficientes e (3) que o sistema adaptativo esteja em regime permanente; o grau de não-linearidade é definido como a razão entre a potência do sinal no ponto em que foi suprimida a função não-linear e a máxima potência de saída possível da função analítica que modela a não-linearidade.

Ou seja:

$$\eta^2 = \frac{\pi}{2} \frac{E\{y_f^2(n)\} \Big|_{\substack{n \rightarrow \infty \\ \sigma^2 \rightarrow \infty}}}{\max\{y_q^2(n)\}} \quad (2.14)$$

Os sinais envolvidos na equação (2.14) podem ser identificados na figura 2.2.

Essa definição permite que o numerador de η^2 esteja associado à parte linear do sistema, e o denominador, à parte não-linear.

Portanto, a equação (2.14) estabelece uma relação entre a máxima potência possível na saída da função não-linear e a potência do sinal que seria necessária para a *minimização* do sinal de erro em um problema de estimação média quadrática linear. Ou seja, η^2 permite inferências sobre a limitação de potência sofrida pelo sinal de cancelamento devido ao efeito de saturação.

No caso em que a função erro $g(\bullet)$ é utilizada, então

$$\max \{y_q^2(n)\} = \max \{g^2(y_f(n))\} = \lim_{y \rightarrow \infty} \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz = \frac{\pi}{2} \sigma^2 \quad (2.15)$$

E portanto:

$$\eta^2 = \frac{1}{\sigma^2} E \{y_f^2(n)\} \Big|_{\sigma^2 \rightarrow \infty} \quad (2.16)$$

O grau de não-linearidade do sistema ANC será utilizado, ao longo deste trabalho, para estudar o comportamento e estabilidade do algoritmo em diferentes condições de operação.

No ANEXO 9 é apresentada uma metodologia para a determinação de η^2 em sistemas reais.

2.7 Casos Particulares de Interesse

A modelagem do caminho secundário por um sistema WH é bastante complexa e representativa de uma grande variedade de aplicações. Entretanto, diversos sistemas adaptativos de interesse prático têm sido modelados na literatura através de estruturas mais simples. Dentre as várias configurações, duas delas se destacam:

2.7.1 Sistema Wiener

O primeiro caso de interesse ocorre quando a influência do sistema S_2 pode ser desconsiderada. Nessa situação, o caminho secundário seria modelado por um **sistema Wiener**, conforme a figura 2.1a.

O modelamento do caminho secundário de sistemas ANC por um sistema Wiener é de grande interesse em aplicações em que as modificações no sinal de cancelamento impostas pelo caminho de propagação são desprezíveis frente à introduzida pelos dispositivos eletrônicos de controle. Um exemplo desse tipo de aplicação é o projeto de protetores auditivos ativos utilizados em aviação militar ou de resgate e salvamento.

A análise desse caso é possível pela simplificação do sistema apresentado na figura 2.2, através da seguinte substituição: $S_2 = [1 \ 0 \ \dots \ 0]^T$ e da troca de variáveis $S_1 \rightarrow S$.

Devido à existência de operações de filtragem decorrentes de S , é utilizado o algoritmo FXLMS para a atualização do vetor de coeficientes.

Recorrendo-se à equação (2.11a), verifica-se que para $\sigma \rightarrow \infty$ a função $g(\bullet)$ comporta-se de maneira puramente linear. Portanto, a utilização dessa condição no sistema Wiener resulta no caso particular do algoritmo FXLMS convencional (TOBIAS, 1999; TOBIAS *et al.*, 2000).

2.7.2 Sistema Sem Memória

Um segundo caso de interesse seria quando as influências de ambos os sistemas lineares do caminho secundário (S_1 e S_2) são desprezíveis. Isso acontece quando o período de amostragem do sistema de controle é grande frente às constantes de tempo de propagação do sinal no meio e do *hardware* associado. Nessa situação o sistema será denominado de **sem memória** (característica referente ao caminho secundário).

Dada a inexistência de operações de filtragem no caminho secundário, é utilizado o algoritmo LMS para a atualização dos coeficientes.

A análise desse sistema é possível através da utilização dos seguintes parâmetros no sistema descrito na figura 2.2: $\hat{\mathbf{S}} = \mathbf{S}_1 = \mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$. Essa consideração equivale à supressão dos sistemas \mathbf{S}_1 , \mathbf{S}_2 e $\hat{\mathbf{S}}$ no diagrama em blocos analisado.

Da mesma forma que para o sistema Wiener, quando $\sigma \rightarrow \infty$ os efeitos não-lineares da função $g(\cdot)$ tornam-se nulos e o equacionamento obtido deve representar o comportamento do algoritmo LMS convencional (HAYKIN, 1991).

No desenvolvimento deste trabalho, expressões gerais são desenvolvidas para o sistema FXLMS-WH apresentado na figura 2.2. Em seguida, expressões específicas são determinadas para os dois casos particulares descritos acima, por simplificação das expressões gerais. É demonstrado também que no caso de $\sigma \rightarrow \infty$ as equações obtidas concordam com os modelos convencionais descritos na literatura para sistemas puramente lineares, caracterizando a generalidade dos modelos desenvolvidos.

3. ANÁLISE DA SUPERFÍCIE DE DESEMPENHO

Neste capítulo será analisada a superfície de desempenho do sistema ANC sob a ótica de um problema de estimação média quadrática não-linear (PAPOULIS, 1991).

A análise da superfície de desempenho possibilita um maior entendimento sobre os fatores envolvidos que afetam o comportamento do sistema, possibilitando a *otimização* das variáveis de controle. O problema do projeto do controlador *ótimo* consiste na determinação do vetor de coeficientes de controle \mathbf{W} que minimize a função objetivo, neste caso o erro médio quadrático (EMQ) na saída do sistema. A figura 3.1 apresenta o diagrama em blocos do sistema de controle estudado.

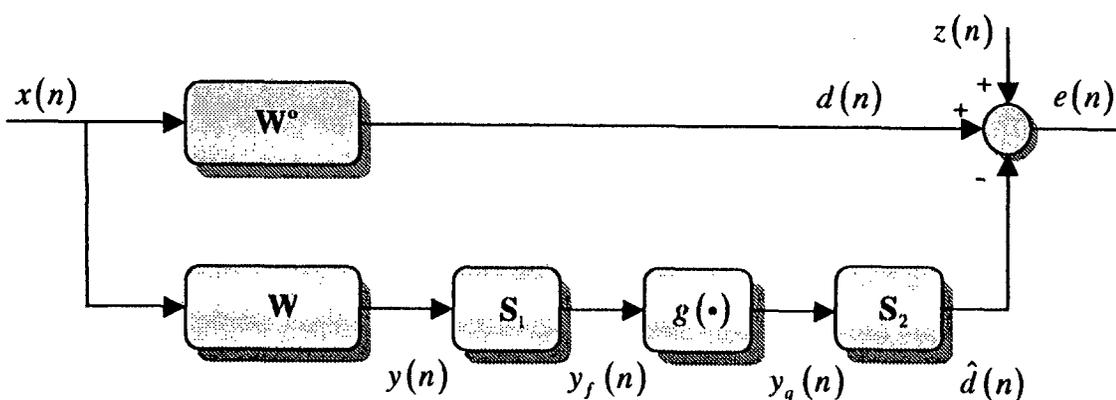


Figura 3.1 – Diagrama em blocos do sistema estudado

A nomenclatura utilizada na figura 3.1 é descrita na seção 2.1.

Posteriormente, os resultados aqui obtidos serão utilizados para avaliar o desempenho do algoritmo FXLMS aplicado ao controlador \mathbf{W} .

3.1 Determinação Analítica da Superfície de Desempenho

Dado um controlador fixo \mathbf{W} na figura 3.1, o sinal de erro é dado por

$$\begin{aligned}
e(n) &= d(n) + z(n) - \hat{d}(n) \\
&= \mathbf{W}^{oT} \mathbf{X}(n) + z(n) - \sum_{j=0}^{M_2-1} s_{2j} g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j) \right]
\end{aligned} \tag{3.1}$$

Elevando-se a equação (3.1) ao quadrado e tomando-se o valor esperado obtém-se:

$$\begin{aligned}
E\{e^2(n)\} &= \mathbf{W}^{oT} E\{\mathbf{X}(n)\mathbf{X}^T(n)\} \mathbf{W}^o + 2\mathbf{W}^{oT} E\{z(n)\mathbf{X}(n)\} + E\{z^2(n)\} \\
&\quad - 2\mathbf{W}^{oT} \sum_{j=0}^{M_2-1} s_{2j} E\left\{g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j) \right] \mathbf{X}(n)\right\} \\
&\quad - 2 \sum_{j=0}^{M_2-1} s_{2j} E\left\{z(n) g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j) \right]\right\} \\
&\quad + \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} E\left\{g \left[\sum_{p=0}^{M_1-1} s_{1p} \mathbf{W}^T \mathbf{X}(n-p-k) \right] g \left[\sum_{q=0}^{M_1-1} s_{1q} \mathbf{W}^T \mathbf{X}(n-q-l) \right]\right\}
\end{aligned} \tag{3.2}$$

Alguns dos valores esperados de (3.2) são de fácil determinação quando consideradas as propriedades do ruído $z(n)$, de forma que:

$$\begin{cases}
E\{\mathbf{X}(n)\mathbf{X}^T(n)\} = \mathbf{R}_0 & (3.3a) \\
E\{z(n)\mathbf{X}(n)\} = E\{z(n)\} \cdot E\{\mathbf{X}(n)\} = 0 & (3.3b) \\
E\{z^2(n)\} = \sigma_z^2 & (3.3c) \\
E\left\{z(n) g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j) \right]\right\} = E\{z(n)\} \cdot E\{g(\cdot)\} = 0 & (3.3d)
\end{cases}$$

Os valores esperados (3.3b) e (3.3d) decorrem do fato de o ruído de medição $z(n)$ possuir média zero e de não ser correlacionado com o sinal de excitação $x(n)$. As equações (3.3a) e (3.3c) são respectivamente a matriz de autocorrelação do sinal de referência e a variância do ruído de medição.

A determinação dos dois valores esperados restantes em (3.2) é mais complicada, já que ambos envolvem operações não-lineares sobre o sinal de referência. A solução desse

problema pode ser encontrada através do Teorema Modificado de Price (BUSSGANG, 1952, PRICE, 1958, PAWULA, 1967).

O valor esperado no quarto termo da equação (3.2) pode ser escrito na forma $E\{g(y_1)Y_2\}$, onde

$$\begin{cases} y_1 = \sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j) \\ \mathbf{Y}_2 = \mathbf{X}(n) \end{cases} \quad (3.4)$$

são variáveis aleatórias gaussianas de média zero, devido às propriedades de $x(n)$.

No ANEXO 1 é demonstrado que, para a função $g(\cdot)$ definida em (2.10),

$$E\{g(y_1)Y_2\} = \frac{E\{y_1 Y_2\}}{\sqrt{\frac{1}{\sigma^2} E\{y_1^2\} + 1}} \quad (3.5)$$

A partir de (3.4), mostra-se facilmente que

$$\begin{cases} E\{y_1^2\} = \mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{q-p} \mathbf{W} \\ E\{y_1 Y_2\} = \sum_{i=0}^{M_1-1} s_{1i} \mathbf{R}_{j+i} \mathbf{W} \end{cases} \quad (3.6)$$

Note-se que a equivalência entre variáveis, em (3.4), é possível somente no caso de o sinal de referência ser considerado gaussiano e média zero e os vetores \mathbf{W} e \mathbf{S}_1 serem fixos. Substituindo (3.4) e (3.6) em (3.5), obtemos o seguinte resultado:

$$E\left\{g\left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T \mathbf{X}(n-i-j)\right] \mathbf{X}(n)\right\} = \frac{\sum_{i=0}^{M_1-1} s_{1i} \mathbf{R}_{j+i} \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{S_1} \mathbf{W} + 1}} \quad (3.7)$$

onde

$$\tilde{\mathbf{R}}_{S_1 S_1} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} \mathbf{R}_{i-k} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} E\{\mathbf{X}(n-k) \mathbf{X}^T(n-i)\} \quad (3.8)$$

A matriz $\tilde{\mathbf{R}}_{S_1 S_1}$ pode ser entendida fisicamente como a matriz de autocorrelação do sinal de referência filtrado pelo sistema S_1 .

Resta ainda um último valor esperado em (3.2) a ser avaliado. Esse termo pode ser escrito como $E\{g(y_1)g(y_2)\}$, onde

$$\begin{cases} y_1 = \sum_{p=0}^{M_1-1} s_{1p} \mathbf{W}^T \mathbf{X}(n-p-k) \\ y_2 = \sum_{q=0}^{M_1-1} s_{1q} \mathbf{W}^T \mathbf{X}(n-q-l) \end{cases} \quad (3.9)$$

são variáveis aleatórias escalares gaussianas de média zero.

No ANEXO 2 demonstra-se, usando (2.10), que

$$E\{g(y_1)g(y_2)\} = \sigma^2 \arcsen\left(\frac{E\{y_1 y_2\}}{\sqrt{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2)}}\right) \quad (3.10)$$

Os valores esperados em (3.10) são iguais a

$$\begin{cases} E\{y_1^2\} = E\{y_2^2\} = \mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} \\ E\{y_1 y_2\} = \mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \mathbf{W} \end{cases} \quad (3.11)$$

Portanto, substituindo-se (3.9) e (3.11) em (3.10), obtém-se:

$$E \left\{ g \left[\sum_{p=0}^{M_1-1} s_{1p} \mathbf{W}^T \mathbf{X}(n-p-k) \right] g \left[\sum_{q=0}^{M_1-1} s_{1q} \mathbf{W}^T \mathbf{X}(n-q-l) \right] \right\} =$$

$$= \sigma^2 \arcsen \left(\frac{\mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + \sigma^2} \right) \quad (3.12)$$

Substituindo-se os resultados (3.3), (3.7) e (3.12) em (3.2), obtém-se finalmente:

$$\xi = E\{e^2(n)\} = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2 s_1}^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + 1}}$$

$$+ \sigma^2 \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \arcsen \left(\frac{\mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + \sigma^2} \right) \quad (3.13)$$

onde $\tilde{\mathbf{R}}_{s_2 s_1}$ também é chamada de matriz de correlação filtrada, porém esta pode ser interpretada como sendo a matriz de correlação cruzada entre o sinal de referência filtrado por \mathbf{S}_1 e o sinal de referência filtrado pelo sistema \mathbf{S}_2 . Ou seja:

$$\tilde{\mathbf{R}}_{s_2 s_1} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} s_{2j} s_{1i} \mathbf{R}_{-i-j} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} s_{2j} s_{1i} E\{\mathbf{X}(n-i-j) \mathbf{X}^T(n)\} \quad (3.14)$$

A equação (3.13) descreve a superfície de desempenho do sistema ANC sujeito a uma não-linearidade (figura 3.1).

3.1.1 Sistema Wiener

De acordo com a seção 2.7.1, a solução para o sistema Wiener pode ser encontrada substituindo-se $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$ e $\mathbf{S}_1 = \mathbf{S}$ em (3.13). Esse procedimento resulta em:

$$\xi_w = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1}} + \sigma^2 \arcsen \left(\frac{\mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + \sigma^2} \right) \quad (3.15)$$

onde

$$\begin{cases} \tilde{\mathbf{R}}_S = \sum_{i=0}^{M-1} s_i \mathbf{R}_{-i} \\ \tilde{\mathbf{R}}_{SS} = \sum_{k=0}^{M-1} \sum_{i=0}^{M-1} s_k s_i \mathbf{R}_{k-i} \end{cases} \quad (3.16)$$

são as matrizes de correlação filtradas desse caso particular.

O caso particular em que $\sigma \rightarrow \infty$ corresponde à superfície de desempenho de um problema de estimação média quadrática linear. Nesse caso,

$$\xi_{wlin} = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{W} + \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} \quad (3.17)$$

A equação (3.17) concorda com o resultado obtido por TOBIAS (1999) para a superfície de desempenho do algoritmo FXLMS convencional.

3.1.2 Sistema Sem Memória

De acordo com a seção 2.7.2, a solução para o sistema sem memória é dada por

$$\xi_{SM} = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_0 \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_0 \mathbf{W} + 1}} + \sigma^2 \arcsen \left(\frac{\mathbf{W}^T \tilde{\mathbf{R}}_0 \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_0 \mathbf{W} + \sigma^2} \right) \quad (3.18)$$

Caso $\sigma \rightarrow \infty$, então

$$\xi_{SMlin} = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o - 2 \mathbf{W}^{oT} \tilde{\mathbf{R}}_0 \mathbf{W} + \mathbf{W}^T \tilde{\mathbf{R}}_0 \mathbf{W} + \sigma_z^2 \quad (3.19)$$

A equação (3.19) concorda com o resultado obtido por HAYKIN (1991).

A partir desses resultados, verifica-se que o equacionamento derivado em (3.13) é uma generalização dos modelos convencionais existentes na literatura.

3.2 Pontos Estacionários

A partir da equação (3.13) que descreve a superfície de desempenho do problema estudado, é de grande interesse determinar os pontos de mínimo da mesma, ou seja, o conjunto de vetores $\tilde{\mathbf{W}}$ que minimizam o erro médio quadrático $E\{e^2(n)\}$.

Para a determinação dos pontos estacionários deriva-se a equação (3.13) em relação ao vetor de coeficientes de controle, igualando-se o resultado a zero. Através dos vetores $\tilde{\mathbf{W}}$ que satisfazem essa condição são encontrados os pontos de gradiente nulo da superfície. Esses pontos podem corresponder a situações de máximo, mínimo ou pontos de sela. A avaliação posterior do Hessiano da superfície de desempenho indicará em qual dessas situações os pontos $\tilde{\mathbf{W}}$ serão classificados.

Em TOBIAS (1999), a superfície de desempenho do algoritmo FXLMS aplicado a um sistema linear é avaliada, e, da mesma forma que para o algoritmo LMS (WIDROW e STEARNS, 1985), essa também possui o formato de um hiperparabolóide, existindo apenas um ponto de mínimo. Para o caso não-linear estudado neste trabalho são esperadas alterações com relação aos resultados obtidos para os sistemas puramente lineares. Entretanto, como a função não-linear utilizada varia suavemente desde a situação linear até um alto grau de não-linearidade, é esperado que também as alterações da superfície de desempenho ocorram de forma suave.

Derivando-se a equação (3.13) e igualando-se o resultado a zero, tem-se:

$$2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + 1}} \right] \Bigg|_{\tilde{\mathbf{W}}} = \sigma^2 \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \frac{\partial}{\partial \mathbf{W}} \arcsen \left(\frac{\mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + \sigma^2} \right) \Bigg|_{\tilde{\mathbf{W}}} \quad (3.20)$$

Inicialmente procede-se à análise da parcela à esquerda da igualdade de (3.20).

Segundo ZWILLINGER (1996):

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \mathbf{X}} \mathbf{A} \mathbf{X} = \mathbf{A}^T \\ \frac{\partial}{\partial \mathbf{X}} \mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{A} \mathbf{X} + \mathbf{A}^T \mathbf{X} \\ \frac{\partial}{\partial x} \frac{1}{u^n} = -\frac{n}{u^{n+1}} \frac{\partial u}{\partial x} \\ \frac{\partial}{\partial \mathbf{X}} \mathbf{U} \mathbf{V} = \frac{\partial \mathbf{U}}{\partial \mathbf{X}} \mathbf{V} + \mathbf{U} \frac{\partial \mathbf{V}}{\partial \mathbf{X}} \end{array} \right. \quad (3.21)$$

Portanto

$$\mathbf{L} \mathbf{E} = 2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + 1}} \right] \Bigg|_{\mathbf{W}=\tilde{\mathbf{W}}} = 2 \frac{(\tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{S_1 S_1} \tilde{\mathbf{W}} + \sigma^2) \tilde{\mathbf{R}}_{S_2 S_1} \mathbf{W}^o - (\mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \tilde{\mathbf{W}}) \tilde{\mathbf{R}}_{S_1 S_1} \tilde{\mathbf{W}}}{\sigma^2 \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{S_1 S_1} \tilde{\mathbf{W}} + 1 \right)^{3/2}} \quad (3.22)$$

onde foi utilizada a propriedade de que $\tilde{\mathbf{R}}_{S_1 S_1}$ é simétrica.

Analisando a parcela à direita da igualdade de (3.20):

$$\mathbf{L} \mathbf{D} = \sigma^2 \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \frac{\partial}{\partial \mathbf{W}} \arcsen \left(\frac{\mathbf{W}^T \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + \sigma^2} \right) \Bigg|_{\mathbf{W}=\tilde{\mathbf{W}}} \quad (3.23)$$

e utilizando-se as seguintes relações (ZWILLINGER, 1996):

$$\begin{cases} \frac{\partial}{\partial u} \arcsen(u) = \frac{1}{\sqrt{1-u^2}} du \\ \frac{\partial}{\partial x} f(u) = \frac{\partial}{\partial u} f(u) \frac{\partial u}{\partial x} \end{cases} \quad (3.24)$$

chega-se a

$$LD = 2 \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \frac{\left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + 1 \right) \mathbf{M}_{k,l} \mathbf{W} - \left(\frac{1}{\sigma^2} \mathbf{W}^T \mathbf{M}_{k,l} \mathbf{W} \right) \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W}}{\left(\left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + 1 \right)^2 - \left(\frac{1}{\sigma^2} \mathbf{W}^T \mathbf{M}_{k,l} \mathbf{W} \right)^2 \right)^{1/2} \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1 s_1} \mathbf{W} + 1 \right)} \quad (3.25)$$

onde

$$\mathbf{M}_{k,l} = \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \quad (3.26)$$

Substituindo-se (3.22) e (3.25) em (3.20), chega-se a

$$\left[\sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \frac{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right) \mathbf{M}_{k,l} - \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \mathbf{M}_{k,l} \tilde{\mathbf{W}} \right) \tilde{\mathbf{R}}_{s_1 s_1}}{\left(\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)^2 - \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \mathbf{M}_{k,l} \tilde{\mathbf{W}} \right)^2 \right)^{1/2}} + \frac{\left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2 s_1}^T \tilde{\mathbf{W}} \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2}} \tilde{\mathbf{R}}_{s_1 s_1} \right] \tilde{\mathbf{W}} = \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2} \tilde{\mathbf{R}}_{s_2 s_1} \mathbf{W}^o \quad (3.27)$$

A equação (3.27) determina as condições para a existência do(s) mínimo(s) da superfície de desempenho. Devido à sua complexidade não foi possível obter uma solução analítica para o caso geral. Entretanto, algoritmos de busca numérica podem ser utilizados para a determinação do(s) ponto(s) e valor(es) do(s) mínimo(s). Inúmeros testes utilizando a equação (3.13) indicam a existência de apenas um mínimo para a superfície de desempenho para uma ampla gama de parâmetros.

Para o caso particular em que o sistema é puramente linear: $\sigma^2 \rightarrow \infty$, $\mathbf{S} = \mathbf{S}_1$ e $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$, a equação (3.27) torna-se:

$$\tilde{\mathbf{W}}_{lin} = \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_S \mathbf{W}^o \quad (3.28)$$

A equação (3.28) expressa o ponto de mínimo da função objetivo quando o sistema é considerado puramente linear. Esse resultado concorda com a equação (2.21) apresentada por TOBIAS (1999).

3.2.1 Solução Aproximada para o Caso Geral

Embora uma solução geral para a equação (3.27) não tenha sido possível, pode-se achar uma solução aproximada para sistemas com grau de não-linearidade reduzido, ou seja, para σ^2 muito elevado.

Esse caso é de grande interesse, pois a maioria dos sistemas ANC são projetados de modo a ter a maior linearidade possível, embora nem sempre sejam completamente lineares. Através de uma solução aproximada pode-se realizar inferências acerca das alterações ocasionadas por pequenas influências não-lineares, permitindo ao projetista o ajuste de seus parâmetros de projeto.

Supondo-se que σ^2 seja suficientemente elevado para que o argumento da função arco-seno na equação (3.13) seja suficientemente pequeno de tal modo que este se encontre na região linear dessa função, pode-se aproximar (3.13) por

$$\xi \cong \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + 1}} + \frac{\mathbf{W}^T \tilde{\mathbf{R}}_{S_2 S_2 S_1} \mathbf{W}}{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W} + 1} \quad (3.29)$$

onde:

$$\tilde{\mathbf{R}}_{S_2 S_2 S_1 S_1} = \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{2k} s_{2l} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \quad (3.30)$$

Para obter os pontos estacionários da equação (3.29), devemos igualar seu gradiente a zero, resultando em

$$-2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2, s_1}^T \mathbf{W}}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1, s_1} \mathbf{W} + 1}} \right] + \frac{\partial}{\partial \mathbf{W}} \left[\frac{\mathbf{W}^T \tilde{\mathbf{R}}_{s_2, s_2, s_1, s_1} \mathbf{W}}{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1, s_1} \mathbf{W} + 1} \right] = \mathbf{0} \quad (3.31)$$

O primeiro termo de (3.31) foi resolvido em (3.22), enquanto que o segundo termo (chamado de \mathbf{T}_2) pode ser calculado utilizando-se (3.21), resultando em

$$\mathbf{T}_2 = \frac{2}{\left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1, s_1} \mathbf{W} + 1 \right)} \tilde{\mathbf{R}}_{s_2, s_2, s_1, s_1} \mathbf{W} - \frac{2 \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_2, s_2, s_1, s_1} \mathbf{W} \right)}{\left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{s_1, s_1} \mathbf{W} + 1 \right)^2} \tilde{\mathbf{R}}_{s_1, s_1} \mathbf{W} \quad (3.32)$$

Substituindo-se os resultados em (3.31) chega-se a

$$\tilde{\mathbf{W}} \cong \left[k_1 \tilde{\mathbf{R}}_{s_2, s_2, s_1, s_1} + k_2 \tilde{\mathbf{R}}_{s_1, s_1} \right]^{-1} \tilde{\mathbf{R}}_{s_2, s_1} \mathbf{W}^o \quad (3.33)$$

onde

$$k_1 = \frac{1}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1, s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2}} \quad (3.34a)$$

$$k_2 = \frac{\left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2, s_1}^T \tilde{\mathbf{W}} \right) \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1, s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2} - \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_2, s_2, s_1, s_1} \tilde{\mathbf{W}} \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1, s_1} \tilde{\mathbf{W}} + 1 \right)^{3/2}} \quad (3.34b)$$

A equação (3.33) é uma primeira aproximação para o(s) ponto(s) estacionário(s) da superfície de desempenho. Entretanto, mesmo supondo-se que exista a matriz inversa, pouco pode ser concluído a partir de (3.33). Portanto, é necessário ainda algum tipo de simplificação para que o resultado possua alguma utilidade para fins de projeto.

Substituindo-se (3.34a) em (3.34b) chega-se à seguinte relação:

$$k_2 = k_1 \cdot \frac{\left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2 s_1}^T \tilde{\mathbf{W}} \right) \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2} - \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1} \tilde{\mathbf{W}} \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)} \quad (3.35)$$

Visto que utilizamos a premissa de que σ^2 é muito elevado, pode-se tirar duas conclusões: (a) $k_1 \cong 1$ e (b) o numerador da fração que multiplica o parâmetro k_1 em (3.35) é muito pequeno, enquanto que o denominador é próximo da unidade. Como conclusão, tem-se que, para σ^2 muito elevado, $k_1 \gg k_2$.

Portanto, se os elementos da matriz de correlação filtrada $\tilde{\mathbf{R}}_{s_1 s_1}$ não forem muito maiores que os elementos de $\tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}$, pode-se fazer a seguinte simplificação:

$$\tilde{\mathbf{W}} \cong \frac{1}{k_1} \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}^{-1} \tilde{\mathbf{R}}_{s_2 s_1} \mathbf{W}^o = \left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{W}} + 1 \right)^{1/2} \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}^{-1} \tilde{\mathbf{R}}_{s_2 s_1} \mathbf{W}^o \quad (3.36)$$

Portanto o ponto estacionário possui a seguinte forma:

$$\tilde{\mathbf{W}} \cong c \cdot \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}^{-1} \tilde{\mathbf{R}}_{s_2 s_1} \mathbf{W}^o \quad (3.37)$$

onde $c = 1/k_1$ é um escalar. Substituindo-se (3.37) em (3.36), chega-se à seguinte relação:

$$c = (c^2 \beta^2 + 1)^{1/2} \quad (3.38)$$

onde

$$\beta^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2 s_1}^T \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}^{-1} \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{R}}_{s_2 s_2 s_1 s_1}^{-1} \tilde{\mathbf{R}}_{s_2 s_1} \mathbf{W}^o \quad (3.39)$$

Por (3.38), sabe-se que $c \in \mathbb{R}^+$. Elevando-se ambos os lados ao quadrado, a única resposta possível é

$$c = \frac{1}{\sqrt{1-\beta^2}} \quad (3.40)$$

Substituindo-se (3.40) em (3.37), chega-se finalmente a

$$\tilde{\mathbf{W}} \equiv \frac{1}{\sqrt{1-\beta^2}} \tilde{\mathbf{R}}_{S_2 S_2 S_1 S_1}^{-1} \tilde{\mathbf{R}}_{S_2 S_1} \mathbf{W}^o \quad (3.41)$$

A equação (3.41) é uma aproximação para a localização do ponto de mínimo da superfície de desempenho quando o sistema é quase linear, ou seja, σ^2 é muito elevado frente às correlações envolvidas nesse desenvolvimento.

Para demonstrar que esse ponto estacionário é realmente de mínimo seria necessário provar que o Hessiano da superfície de desempenho nesse ponto é definido positivo. Em decorrência da complexidade matemática e de o ponto estacionário determinado ser apenas uma aproximação, essa tarefa não foi realizada. Entretanto, inúmeras simulações demonstraram ser essa uma estimativa razoável para o aparente único ponto de mínimo da superfície de desempenho. Além disso, para $\beta^2 = 0$, $\mathbf{S} = \mathbf{S}_1$ e $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$, a equação (3.41) coincide com a solução desenvolvida para o ponto de mínimo da superfície de desempenho do FXLMS linear (TOBIAS, 1999).

3.2.2 Sistema Wiener

Quando as alterações provocadas por \mathbf{S}_2 são desprezíveis com relação ao restante dos componentes do sistema ANC, é possível calcular a solução exata para o único ponto estacionário da superfície de desempenho. Essa solução é de especial interesse na análise de protetores auditivos ativos.

Para $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$ e $\mathbf{S}_1 = \mathbf{S}$, a equação (3.27) torna-se:

$$\tilde{\mathbf{W}}_w = \frac{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}_w^T \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}}_w + 1 \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}_w^T \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}}_w + 1 \right)^{1/2} + \left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_s^T \tilde{\mathbf{W}}_w \right)} \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{R}}_s \mathbf{W}^o \quad (3.42)$$

onde $\tilde{\mathbf{W}}_w$ é o ponto de mínimo quando o caminho secundário é modelado por um sistema Wiener, e $\tilde{\mathbf{R}}_s$ e $\tilde{\mathbf{R}}_{ss}$ são definidas em (3.16).

Portanto, pode-se concluir que a solução possui a seguinte forma (onde c_w é um escalar real):

$$\tilde{\mathbf{W}}_w = c_w \cdot \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{R}}_s \mathbf{W}^o \quad (3.43)$$

Substituindo-se (3.43) em (3.42) e definindo

$$\beta_w^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_s^T \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{R}}_s \mathbf{W}^o \quad (3.44)$$

obtem-se, após algumas simplificações,

$$c_w \sqrt{\frac{c_w^2 \beta_w^2 + 1}{2c_w^2 \beta_w^2 + 1}} = 1 \quad (3.45)$$

Elevando-se ambos os lados de (3.45) ao quadrado, chega-se a uma equação bi-quadrática, e utilizando-se a fórmula de báscara podem ser determinadas as quatro soluções numéricas pertinentes. Entretanto, pode ser mostrado que duas delas serão complexo-conjugadas, e uma terceira será real negativa. Como $c_w \in \mathbb{R}^+$ (segundo (3.45)), existe apenas uma solução possível. Substituindo-a em (3.43), chega-se a

$$\tilde{\mathbf{W}}_w = \sqrt{1 - \frac{1}{2\beta_w^2} + \sqrt{\frac{1}{4\beta_w^4} + 1}} \cdot \tilde{\mathbf{R}}_{ss}^{-1} \tilde{\mathbf{R}}_s \mathbf{W}^o \quad (3.46)$$

A equação (3.46) é a solução para o único ponto estacionário da superfície de desempenho quando o caminho secundário é modelado por um sistema Wiener. A prova matemática de que esse é um ponto de mínimo será desenvolvida na seção 3.3 deste capítulo.

3.2.3 Sistema Sem Memória

Muitos sistemas ANC e outros sistemas não-lineares podem ser modelados supondo-se que o sinal de controle ($y(n)$) não é afetado pelo meio de propagação e *hardware* associado ($\mathbf{S}_2 = \mathbf{S}_1 = [1 \ 0 \ \dots \ 0]^T$). Esse caso corresponde ao algoritmo LMS com uma não-linearidade na saída do filtro adaptativo. Para tais casos, a equação (3.46) pode ser simplificada de tal forma que

$$\tilde{\mathbf{W}}_{SM} = \sqrt{1 - \frac{1}{2\beta_{SM}^2} + \sqrt{\frac{1}{4\beta_{SM}^4} + 1}} \cdot \mathbf{W}^o \quad (3.47)$$

onde

$$\beta_{SM}^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \mathbf{R}_{xx} \mathbf{W}^o \quad (3.48)$$

Para $\sigma^2 \rightarrow \infty$ o sistema comporta-se de maneira linear, e (3.47) tende para o vetor de coeficientes da planta, de acordo com HAYKIN (1991).

3.3 Hessiano da Superfície de Desempenho

Para provar que a equação (3.46) é um ponto de mínimo, é necessário o cálculo do Hessiano da superfície de desempenho e demonstrar que este é definido positivo no ponto $\tilde{\mathbf{W}}$. Como já mencionado, a determinação do Hessiano para o caso geral WH é muito complexa, se possível. Portanto, a demonstração apresentada refere-se ao caso particular do caminho secundário modelado por um sistema Wiener.

A equação da superfície de desempenho para o sistema Wiener pode ser obtida supondo-se $\mathbf{S} = \mathbf{S}_1$ e $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$ na equação (3.13), resultando em

$$\xi_w = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - \frac{2}{\sqrt{\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W} + 1}} \mathbf{W}^{oT} \tilde{\mathbf{R}}_s^T \mathbf{W} + \sigma^2 \arcsen \left(\frac{\mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W}}{\mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W} + \sigma^2} \right) \quad (3.49)$$

Para obter-se o Hessiano é necessário calcular a derivada segunda de (3.49). Aproveitando os resultados obtidos em (3.22) e (3.25) para $\mathbf{S} = \mathbf{S}_1$ e $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$, chega-se a

$$\begin{aligned} \nabla^2 \xi_w(\mathbf{W}) = & -2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}} + 1 \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}} + 1 \right)^{3/2}} \tilde{\mathbf{R}}_s \mathbf{W}^o \right] + 2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{\left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_s^T \tilde{\mathbf{W}} \right)}{\left(\frac{1}{\sigma^2} \tilde{\mathbf{W}}^T \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}} + 1 \right)^{3/2}} \tilde{\mathbf{R}}_{ss} \tilde{\mathbf{W}} \right] \\ & + 2 \frac{\partial}{\partial \mathbf{W}} \left[\frac{1}{\left(\frac{2}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W} + 1 \right)^{1/2} \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{ss} \mathbf{W} + 1 \right)} \tilde{\mathbf{R}}_{ss} \mathbf{W} \right] \end{aligned} \quad (3.50)$$

Utilizando-se as relações descritas em (3.21) e sabendo-se ainda que

$$\frac{\partial}{\partial \mathbf{X}} (a\mathbf{Y}) = \left(\frac{\partial a}{\partial \mathbf{X}} \right) \mathbf{Y}^T + a \frac{\partial \mathbf{Y}}{\partial \mathbf{X}} \quad (3.51)$$

onde a é um escalar e \mathbf{X} e \mathbf{Y} são vetores, após um pouco de esforço matemático chega-se a

$$\begin{aligned}
\nabla^2 \xi_w(\mathbf{W}) = & \frac{2}{\sigma^2 \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^{3/2}} \left[\tilde{\mathbf{R}}_{SS} \mathbf{W} \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T + \tilde{\mathbf{R}}_S \mathbf{W}^o \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \right] \\
& + \left[\frac{\left(\frac{2}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{W} \right)}{\left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^{3/2}} + \frac{2}{\left(\frac{2}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^{1/2} \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)} \right] \tilde{\mathbf{R}}_{SS} \\
& - \left[\frac{\left(\frac{6}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{W} \right)}{\sigma^2 \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^{5/2}} + \frac{\left(\frac{12}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 8 \right)}{\sigma^2 \left(\frac{2}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^{3/2} \left(\frac{1}{\sigma^2} \mathbf{W}^T \tilde{\mathbf{R}}_{SS} \mathbf{W} + 1 \right)^2} \right] \tilde{\mathbf{R}}_{SS} \mathbf{W} \mathbf{W}^T \tilde{\mathbf{R}}_{SS}
\end{aligned} \tag{3.52}$$

Para verificar a concavidade da superfície de erro no suposto ponto de mínimo temos de avaliar o Hessiano na posição dada por (3.43). Substituindo esse resultado em (3.52), resulta:

$$\nabla^2 \xi_w(\tilde{\mathbf{W}}_w) = a \tilde{\mathbf{R}}_{SS} + \frac{b}{\sigma^2} \tilde{\mathbf{R}}_S \mathbf{W}^o \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \tag{3.53}$$

onde

$$\left\{ \begin{aligned}
a &= \frac{2c_w \beta_w^2 (2c_w^2 \beta_w^2 + 1)^{1/2} + 2(c_w^2 \beta_w^2 + 1)^{1/2}}{(2c_w^2 \beta_w^2 + 1)^{1/2} (c_w^2 \beta_w^2 + 1)^{3/2}} \\
b &= \frac{(-2c_w^3 \beta_w^2 + 4c_w)(2c_w^2 \beta_w^2 + 1)^{3/2} - (12c_w^4 \beta_w^2 + 8c_w^2)(c_w^2 \beta_w^2 + 1)^{1/2}}{(2c_w^2 \beta_w^2 + 1)^{3/2} (c_w^2 \beta_w^2 + 1)^{5/2}}
\end{aligned} \right. \tag{3.54}$$

Através de (3.54) é fácil verificar que o parâmetro “a” é positivo para todo β_w^2 . Entretanto, a avaliação numérica de “b” demonstra que o mesmo é sempre negativo para todo β_w^2 . Em decorrência disso, a verificação da positividade do Hessiano não é direta e temos de recorrer a um procedimento alternativo.

HORN e JOHNSON (1999) apresentam sob o número 4.3.4 o seguinte teorema:

“Seja $\mathbf{A} \in \mathbf{M}_n$ uma matriz hermitiana, $z \in \mathbf{C}^N$ um dado vetor e se os autovalores de A e

$\mathbf{A} \pm zz^*$ forem arranjados em ordem crescente, tem-se que $\lambda_k(\mathbf{A}) \leq \lambda_{k+1}(\mathbf{A} \pm zz^*) \leq \lambda_{k+2}(\mathbf{A})$, $k = 1, 2, \dots, N-2$ ” (onde N é o número de autovalores)”.

Mapeando-se essas variáveis, para o caso analisado, obtêm-se as seguintes relações:

$$\begin{cases} \mathbf{A} = a\tilde{\mathbf{R}}_{SS} \\ z = \frac{\sqrt{|b|}}{\sigma} \tilde{\mathbf{R}}_S \mathbf{W}^o \end{cases} \quad (3.55)$$

Supondo-se que $\tilde{\mathbf{R}}_{SS}$ seja definida positiva, o que é razoável, visto que $\tilde{\mathbf{R}}_{SS}$ é a matriz de autocorrelação do sinal de referência filtrado por \mathbf{S} , e sendo o parâmetro “ a ” positivo para todo β_w^2 , então a matriz \mathbf{A} possui apenas autovalores positivos. Portanto temos de nos preocupar apenas com a prova de positividade de um único autovalor da matriz $\nabla^2 \xi_w(\tilde{\mathbf{W}}_w)$.

Sabendo-se que o determinante de uma matriz é igual ao produto de seus autovalores (HORN e JOHNSON, 1999) e que existem $N-1$ autovalores positivos, então o sinal do determinante da equação (3.53) será dado pelo sinal do autovalor restante. Procedendo-se à avaliação do determinante, após certo desenvolvimento matemático chega-se a

$$\det\{\nabla^2 \xi_w(\tilde{\mathbf{W}}_w)\} = \det\left\{a\mathbf{I} + \frac{b}{\sigma^2} \Lambda^{-1/2} \mathbf{Q}^T \tilde{\mathbf{R}}_S \mathbf{W}^o \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{Q} \Lambda^{-1/2}\right\} \cdot \det\{\tilde{\mathbf{R}}_{SS}\} \quad (3.56)$$

onde $\tilde{\mathbf{R}}_{SS} = \mathbf{Q} \Lambda \mathbf{Q}^T$ em decorrência de sua positividade, e $\det\{\mathbf{A} \cdot \mathbf{B}\} = \det\{\mathbf{A}\} \cdot \det\{\mathbf{B}\}$.

O segundo determinante do lado direito de (3.56) é positivo, visto que $\tilde{\mathbf{R}}_{SS}$ é definida positiva; portanto resta apenas avaliar o primeiro determinante.

Sabendo-se que os autovalores da matriz no interior do primeiro determinante de (3.56) são

$$\begin{cases} a & \text{multiplicidade } N-1 \\ \gamma = a + b\beta_w^2 = a + \frac{b}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_S \mathbf{W}^o & \text{multiplicidade } 1 \end{cases} \quad (3.57)$$

e que o determinante dessa matriz é igual ao produto de seus autovalores, tem-se como resultado:

$$\det \left\{ a\mathbf{I} + \frac{b}{\sigma^2} \Lambda^{-1/2} \mathbf{Q}^T \tilde{\mathbf{R}}_S \mathbf{W}^o \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \mathbf{Q} \Lambda^{-1/2} \right\} = a^{N-1} (a + b\beta_w^2) \quad (3.58)$$

Substituindo-se (3.45) e (3.54) em (3.57), obtém-se:

$$\begin{cases} a = \frac{2}{c_w (c_w^2 \beta_w^2 + 1)^{1/2}} \\ \gamma = \frac{4c_w^4 \beta_w^4 + 4c_w^2 \beta_w^2 + 2}{(c_w^2 \beta_w^2 + 1)^2 (2c_w^2 \beta_w^2 + 1)^{3/2}} \end{cases} \quad (3.59)$$

Visto que $a, \gamma > 0 \quad \forall \beta_w^2$, podemos concluir que o Hessiano será definido positivo no ponto $\tilde{\mathbf{W}}_w$, caracterizando um mínimo da superfície de desempenho.

A relação obtida é válida para qualquer \mathbf{S} , mesmo para o caso em que não existem alterações do sinal de controle ($\mathbf{S} = [1 \ 0 \ \dots \ 0]^T$), ou seja, para o caminho secundário modelado por um sistema não-linear sem memória.

3.4 Mínimo Erro Médio Quadrático

O objetivo de um sistema ANC é a diminuição dos níveis sonoros ou vibracionais de um determinado sistema. Portanto, o conhecimento da máxima atenuação possível, segundo uma determinada função objetivo (no caso o erro médio quadrático), para um dado sistema, é de grande importância.

Nesta seção são utilizados os resultados referentes aos pontos de mínimo da superfície de desempenho para avaliar o máximo cancelamento possível para a topologia analisada (figura 3.1) e seus casos particulares.

3.4.1 Solução Aproximada para o Caso Geral

Substituindo-se a equação (3.41) em (3.13) tem-se como resultado:

$$\begin{aligned} \xi_{MIN} \equiv & \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \tilde{\mathbf{R}}_{S_2 S_2 S_1 S_1}^{-1} \tilde{\mathbf{R}}_{S_2 S_1} \mathbf{W}^o \\ & + \sigma^2 \sum_{k=0}^{M_2-1} \sum_{l=0}^{M_2-1} s_{2k} s_{2l} \arccos \left(\frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{S_2 S_1}^T \tilde{\mathbf{R}}_{S_2 S_2 S_1 S_1}^{-T} \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{R}_{p+k-q-l} \tilde{\mathbf{R}}_{S_2 S_2 S_1 S_1}^{-1} \tilde{\mathbf{R}}_{S_2 S_1} \mathbf{W}^o \right) \end{aligned} \quad (3.60)$$

A equação (3.60) avalia, de forma aproximada, o máximo cancelamento possível (mínimo EMQ) obtido pelo controlador.

3.4.2 Sistema Wiener

No caso em que o caminho secundário é modelado por um sistema Wiener, substituindo-se (3.46) em (3.13) para $\mathbf{S} = \mathbf{S}_1$ e $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$, resulta:

$$\xi_{W MIN} = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 + \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_S \mathbf{W}^o \left[\frac{1}{\beta_w^2} \arccos \left(\frac{c_w^2 \beta_w^2}{c_w^2 \beta_w^2 + 1} \right) - \frac{2c_w}{\sqrt{c_w^2 \beta_w^2 + 1}} \right] \quad (3.61)$$

No caso de ausência de efeitos não-lineares ($\beta_w^2 \rightarrow 0$), a equação (3.61) concorda com a equação (2.26) de TOBIAS (1999).

3.4.3 Sistema Sem Memória

O máximo cancelamento possível, no caso do caminho secundário ser modelado

por um sistema não-linear sem memória, é obtido substituindo-se $\mathbf{S} = [1 \ 0 \ \dots \ 0]^T$ em (3.61):

$$\xi_{SM \ MIN} = \mathbf{W}^{oT} \mathbf{R}_{xx} \mathbf{W}^o \left[\frac{1}{\beta_{SM}^2} \arcsen \left(\frac{c_{SM}^2 \beta_{SM}^2}{c_{SM}^2 \beta_{SM}^2 + 1} \right) - \frac{2c_{SM}}{\sqrt{c_{SM}^2 \beta_{SM}^2 + 1}} + 1 \right] + \sigma_z^2 \quad (3.62)$$

No caso de não existirem efeitos não-lineares ($\beta_{SM}^2 \rightarrow 0$), a equação (3.62) concorda com os resultados de HAYKIN (1991).

4. COMPORTAMENTO MÉDIO DOS COEFICIENTES

Neste capítulo é analisado o comportamento médio dos coeficientes do sistema adaptativo apresentado na figura 4.1, tanto durante o período de adaptação como em regime permanente. O sistema S_1 permite levar em consideração a influência da *parte elétrica de atuação* (filtros, amplificação, conversão digital-analógica), enquanto que S_2 representa os atrasos introduzidos pela propagação do sinal de controle no meio e pela *parte elétrica de aquisição* (amplificadores, filtros, conversão analógica-digital). A função $g(\cdot)$ representa os efeitos não-lineares do tipo saturação causados pelos transdutores e amplificadores utilizados. O vetor de coeficientes de adaptação é atualizado a cada iteração pelo algoritmo LMS Filtrado.

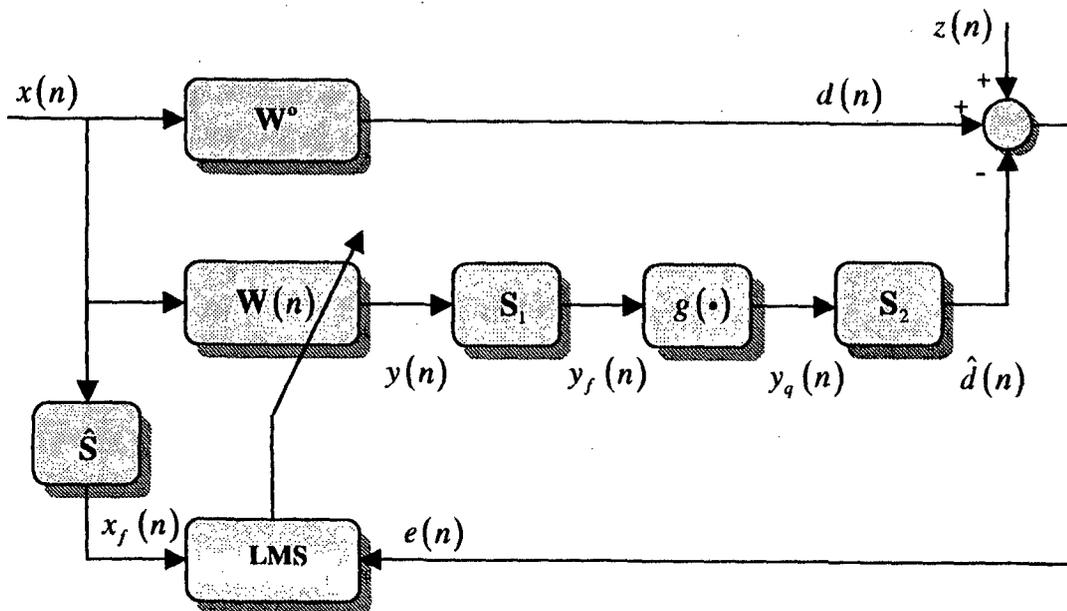


Figura 4.1 - Topologia analisada: algoritmo FXLMS com o caminho secundário modelado por um sistema Wiener-Hammerstein.

As definições das variáveis da figura 4.1 são encontradas na seção 2.1.

No decorrer desse texto são apresentadas equações recursivas para o comportamento médio dos coeficientes em termos de momentos de até segunda ordem.

Através desses resultados são obtidas as condições em regime permanente.

4.1 Análise do Comportamento Médio dos Coeficientes

A equação de atualização dos coeficientes do sistema analisado é dada pela equação (2.4), transcrita a seguir:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu \left\{ \mathbf{W}^{oT} \mathbf{X}(n) + z(n) - \sum_{j=0}^{M_2-1} s_{2j} g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \right\} \mathbf{X}_f(n) \quad (4.1)$$

Em decorrência do comportamento não-linear de $g(\cdot)$, a avaliação direta das esperanças necessárias para a determinação do comportamento médio dos coeficientes não é de fácil realização. Portanto, inicialmente são avaliadas as esperanças condicionadas no conjunto de vetores dado por

$$\mathcal{W} = \{ \mathbf{W}(n), \mathbf{W}(n-1), \dots, \mathbf{W}(n-M+1) \} \quad (4.2)$$

onde $M = M_1 + M_2 - 1$. Tomando-se o valor esperado condicionado em \mathcal{W} da equação (4.1), obtém-se:

$$E\{ \mathbf{W}(n+1) | \mathcal{W} \} = \mathbf{W}(n) + \mu E\{ \mathbf{X}_f(n) \mathbf{X}^T(n) | \mathcal{W} \} \mathbf{W}^o + \mu E\{ z(n) \mathbf{X}_f(n) | \mathcal{W} \} - \mu \sum_{j=0}^{M_2-1} s_{2j} E\left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}_f(n) | \mathcal{W} \right\} \quad (4.3)$$

Na avaliação dos valores esperados em (4.3), considera-se, conforme discutido na seção 2.4, que as correlações entre diferentes vetores de entrada são muito mais importantes do que as correlações entre vetores de entrada e vetores de coeficientes para a determinação do comportamento do algoritmo. Com essa consideração, os dois primeiros valores esperados em (4.3) são de fácil solução, segundo a metodologia clássica de análise

de filtros adaptativos (HAYKIN, 1991). Portanto

$$\left\{ E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{s}} \right. \quad (4.4a)$$

$$\left\{ E\{z(n)\mathbf{X}_f(n)|\mathcal{W}\} = \mathbf{0} \right. \quad (4.4b)$$

onde

$$\tilde{\mathbf{R}}_{\hat{s}} = \sum_{j=0}^{M-1} \hat{s}_j \mathbf{R}_j = \sum_{j=0}^{M-1} \hat{s}_j E\{\mathbf{X}(n-j)\mathbf{X}^T(n)\} \quad (4.5)$$

O primeiro valor esperado de (4.4) representa a matriz de correlação cruzada entre o sinal de referência filtrado pela estimativa da parte linear do caminho secundário e o próprio sinal de referência. O segundo valor esperado de (4.4) resulta do fato de o ruído de medição ser independente do sinal de referência e possuir média zero. Note-se que são desprezadas as correlações entre o vetor de coeficientes e o vetor de dados observados.

O último valor esperado de (4.3) pode ser escrito na forma $E\{g(y_1)\mathbf{Y}_2|\mathcal{W}\}$, onde

$$\left\{ \begin{aligned} y_1 &= \sum_{i=0}^{M-1} s_{1i} \mathbf{W}^T(n-i-j)\mathbf{X}(n-i-j) \\ \mathbf{Y}_2 &= \sum_{k=0}^{M-1} \hat{s}_k \mathbf{X}(n-k) \end{aligned} \right. \quad (4.6a)$$

$$\left. \right\} \quad (4.6b)$$

são variáveis aleatórias gaussianas de média zero quando condicionadas em \mathcal{W} (BENDAT e PIERSOL, 1986).

No ANEXO 1 é demonstrado que

$$E\{g(y_1)\mathbf{Y}_2\} = \frac{E\{y_1\mathbf{Y}_2\}}{\sqrt{\frac{1}{\sigma^2} E\{y_1^2\} + 1}} \quad (4.7)$$

Desprezando-se as correlações entre os vetores de dados observados e vetores de coeficientes em comparação com as correlações entre vetores de dados observados, pode-se determinar as seguintes correlações condicionadas em \mathcal{W} :

$$\left\{ E\{y_1^2|\mathcal{W}\} = \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i}s_{1p} \mathbf{W}^T(n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) \right. \quad (4.8a)$$

$$\left. E\{y_1 Y_2|\mathcal{W}\} = \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i}\hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \right. \quad (4.8b)$$

Substituindo-se (4.6) e (4.8) em (4.7), chega-se a

$$E\left\{g\left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j)\right] \mathbf{X}_f(n) \middle| \mathcal{W}\right\} = \frac{\sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i}\hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j)}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i}s_{1p} \mathbf{W}^T(n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) + 1}} \quad (4.9)$$

Utilizando-se os resultados (4.4) e (4.9) em (4.3), obtém-se como resultado o comportamento médio do vetor de coeficientes condicionado em \mathcal{W} :

$$E\{\mathbf{W}(n+1)|\mathcal{W}\} = \mathbf{W}(n) + \mu \tilde{\mathbf{R}}_s \mathbf{W}^0 - \mu \sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i}\hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j)}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i}s_{1p} \mathbf{W}^T(n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) + 1}} \quad (4.10)$$

Para retirar-se o condicionamento em \mathcal{W} em (4.10) deve-se calcular o valor esperado da equação levando-se em consideração a estatística dos $\mathbf{W}(n-k)$ (para $k=0, 1, \dots, M-1$). Entretanto, a função densidade de probabilidade dos vetores de coeficientes não é conhecida.

Utilizando-se a suposição de que μ seja suficientemente pequeno, pode-se assumir que os coeficientes variam lentamente. Assim, as flutuações de $\mathbf{W}(n-k)$ em torno de sua média $E\{\mathbf{W}(n-k)\}$ possuem um efeito desprezível no comportamento médio dos coeficientes ao longo do tempo. Portanto, faz-se a seguinte aproximação:

$$E\{\mathbf{W}(n+1)\} = E\{\mathbf{W}(n)\} + \mu \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o - \mu \sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{i=0}^{M_1-1} \sum_{k=0}^{M-1} s_{1i} \hat{s}_k \mathbf{R}_{j+i-k} E\{\mathbf{W}(n-i-j)\}}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \text{tr}\{\mathbf{R}_{p-i} \mathbf{K}_{p,i}(n-j)\} + 1}} \quad (4.11)$$

onde

$$\mathbf{K}_{p,i}(n-j) = E\{\mathbf{W}(n-p-j) \mathbf{W}^T(n-i-j)\} \quad (4.12)$$

são as matrizes de correlação dos vetores de coeficientes em diversas iterações. Esses momentos serão desenvolvidos no capítulo 6. O operador $\text{tr}\{\cdot\}$ corresponde ao traço da matriz argumento.

A equação (4.11) é o modelo analítico para o comportamento médio dos coeficientes do sistema apresentado na figura 4.1, supondo-se sinais gaussianos de média nula.

4.1.1 Sistema Wiener

No caso em que as alterações sofridas pela propagação do sinal de controle no meio físico sejam desprezíveis ($\mathbf{S} = \mathbf{S}_1$, $\mathbf{S}_2 = [1 \ 0 \ \dots \ 0]^T$), a equação (4.11) reduz-se a

$$E\{\mathbf{W}_w(n+1)\} = E\{\mathbf{W}_w(n)\} + \mu \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o - \mu \frac{\sum_{i=0}^{M-1} \sum_{k=0}^{M-1} s_i \hat{s}_k \mathbf{R}_{i-k} E\{\mathbf{W}_w(n-i)\}}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M-1} \sum_{p=0}^{M-1} s_i s_p \text{tr}\{\mathbf{R}_{p-i} \mathbf{K}_{w,p,i}(n)\} + 1}} \quad (4.13)$$

onde

$$\mathbf{K}_{w_{p,i}}(n) = E\{\mathbf{W}_w(n-p)\mathbf{W}_w^T(n-i)\} \quad (4.14)$$

são as matrizes de correlação do vetor de coeficientes quando o caminho secundário é modelado por um sistema Wiener.

Caso $\sigma^2 \rightarrow \infty$, o sistema pode ser considerado puramente linear e portanto (4.13) reduz-se a

$$E\{\mathbf{W}_w(n+1)\} = E\{\mathbf{W}_w(n)\} - \mu \sum_{i=0}^{M-1} \sum_{k=0}^{\hat{M}-1} s_i \hat{s}_k \hat{\mathbf{R}}_{i-k} E\{\mathbf{W}_w(n-i)\} + \mu \tilde{\mathbf{R}}_s \mathbf{W}^o \quad (4.15)$$

a qual concorda com a equação (2.37) de TOBIAS (1999).

4.1.2 Sistema Sem Memória

Nas situações em que o sinal de controle está sujeito apenas às alterações causadas pelo sistema não-linear, o equacionamento do comportamento médio dos coeficientes pode ser simplificado substituindo-se $\mathbf{S} = [1 \ 0 \ \dots \ 0]^T$ em (4.13):

$$E\{\mathbf{W}_{SM}(n+1)\} = \left[\mathbf{I} - \mu \frac{1}{\sqrt{\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1}} \mathbf{R}_0 \right] E\{\mathbf{W}_{SM}(n)\} + \mu \mathbf{R}_0 \mathbf{W}^o \quad (4.16)$$

onde

$$\mathbf{K}_{SM}(n) = E\{\mathbf{W}_{SM}(n)\mathbf{W}_{SM}^T(n)\} \quad (4.17)$$

é a equação dos momentos de segunda ordem quando o caminho secundário não possui memória.

Caso $\sigma^2 \rightarrow \infty$, o sistema pode ser considerado puramente linear e portanto (4.16) tende para

$$E\{\mathbf{W}_{SM}(n+1)\} = [\mathbf{I} - \mu \mathbf{R}_0] E\{\mathbf{W}_{SM}(n)\} + \mu \mathbf{R}_0 \mathbf{W}^0 \quad (4.18)$$

que concorda com os resultados obtidos para o algoritmo LMS convencional em HAYKIN (1991).

4.2 Coeficientes em Regime Permanente

Calculando-se o limite quando $n \rightarrow \infty$ da equação (4.11) e supondo-se a existência de convergência de tal modo que $\mathbf{W}_{RP} = \lim_{n \rightarrow \infty} E\{\mathbf{W}(n)\}$, então, após certa manipulação algébrica, obtém-se

$$\mathbf{W}_{RP} = \left(\frac{1}{\sigma^2} \mathbf{W}_{RP}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W}_{RP} + 1 \right)^{1/2} \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^0 \quad (4.19)$$

onde

$$\begin{cases} \tilde{\mathbf{R}}_{\hat{S}} = \sum_{j=0}^{\hat{M}-1} \hat{S}_k \mathbf{R}_{-k} = \sum_{j=0}^{\hat{M}-1} \hat{S}_k E\{\mathbf{X}(n-k) \mathbf{X}^T(n)\} \\ \tilde{\mathbf{R}}_{S_1 S_1} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} \mathbf{R}_{i-k} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} E\{\mathbf{X}(n-k) \mathbf{X}^T(n-i)\} \\ \tilde{\mathbf{R}}_{S_2 \hat{S}_1} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{2j} s_{1i} \hat{S}_k \mathbf{R}_{i+j-k} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{2j} s_{1i} \hat{S}_k E\{\mathbf{X}(n-k) \mathbf{X}^T(n-i-j)\} \end{cases} \quad (4.20)$$

Entretanto, (4.19) pode ser representada da seguinte forma:

$$\mathbf{W}_{RP} = k \cdot \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^0 \quad (4.21)$$

onde k é um escalar real positivo. Substituindo-se (4.21) em (4.19) e igualando-se as parcelas escalares, chega-se a

$$k = (k^2 \eta^2 + 1)^{1/2} \quad (4.22)$$

onde:

$$\eta^2 = \frac{1}{\sigma^2} \mathbf{W}^{0T} \tilde{\mathbf{R}}_{\hat{S}}^T \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-T} \tilde{\mathbf{R}}_{S_1 S_1} \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^0 \quad (4.23)$$

O parâmetro η^2 é analisado na seção 4.5. Resolvendo-se (4.22) para k e sabendo-se que $k \in \mathbb{R}^+$, então existe apenas uma solução possível: $k = 1/\sqrt{1-\eta^2}$. Substituindo-se esse resultado em (4.21), chega-se a

$$\lim_{n \rightarrow \infty} E\{\mathbf{W}(n)\} = \frac{1}{\sqrt{1-\eta^2}} \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (4.24)$$

A equação (4.24) representa o valor médio dos coeficientes em regime permanente, supondo-se que exista convergência, para o sistema ANC descrito na figura 4.1.

4.2.1 Sistema Wiener

No caso em que a influência do sistema S_2 é desprezível, então (4.24) torna-se

$$\lim_{n \rightarrow \infty} E\{\mathbf{W}_W(n)\} = \frac{1}{\sqrt{1-\eta_W^2}} \tilde{\mathbf{R}}_{\hat{S}\hat{S}}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (4.25)$$

onde

$$\eta_W^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{S}}^T \tilde{\mathbf{R}}_{\hat{S}\hat{S}}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (4.26)$$

e

$$\left\{ \begin{aligned} \tilde{\mathbf{R}}_{\hat{S}} &= \sum_{j=0}^{\hat{M}-1} \hat{s}_k \mathbf{R}_{-k} = \sum_{j=0}^{\hat{M}-1} \hat{s}_k E\{\mathbf{X}(n-k)\mathbf{X}^T(n)\} \\ \tilde{\mathbf{R}}_{\hat{S}\hat{S}} &= \sum_{i=0}^{\hat{M}-1} \sum_{k=0}^{\hat{M}-1} s_i \hat{s}_k \mathbf{R}_{i+j-k} = \sum_{i=0}^{\hat{M}-1} \sum_{k=0}^{\hat{M}-1} s_i \hat{s}_k E\{\mathbf{X}(n-k)\mathbf{X}^T(n-i-j)\} \end{aligned} \right. \quad (4.27)$$

Se $\eta_W^2 \rightarrow 0$ (caso linear), então (4.25) reduz-se à expressão (15) de TOBIAS *et al.* (1998), que trata do regime permanente do algoritmo FXLMS convencional. Se, em conjunto, $\hat{\mathbf{S}} = \mathbf{S}$ (perfeita estimação do caminho secundário), então $\tilde{\mathbf{R}}_{\hat{S}\hat{S}} = \tilde{\mathbf{R}}_{SS}$, $\tilde{\mathbf{R}}_{\hat{S}} = \tilde{\mathbf{R}}_S$ e (4.25) reduz-se à expressão dos coeficientes de Wiener com restrição, descrita na equação (18) de TOBIAS *et al.* (1998). Portanto, para $\hat{\mathbf{S}} = \mathbf{S}$ e $\eta_W^2 \neq 0$ o vetor dos coeficientes médios tende para uma versão escalada dos coeficientes de Wiener restringidos.

4.2.2 Sistema Sem Memória

Quando o caminho secundário não possui memória, a equação (4.25) é simplificada, resultando:

$$\lim_{n \rightarrow \infty} E\{\mathbf{W}_{SM}(n)\} = \frac{1}{\sqrt{1-\eta_{SM}^2}} \mathbf{W}^o \quad (4.28)$$

onde:

$$\eta_{SM}^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o \quad (4.29)$$

A equação (4.28) nos mostra que para um sistema sem memória o valor médio do vetor de coeficientes em regime permanente é uma versão escalada da resposta da planta a ser controlada. Para o caso em que o sistema é linear ($\eta_{SM}^2 \rightarrow 0$), então a equação (4.28) concorda com o resultado obtido para o regime permanente do algoritmo LMS convencional (HAYKIN, 1991).

4.3 Modelo Simplificado

As equações que descrevem o comportamento médio dos coeficientes ao longo do processo de adaptação são dadas por (4.11), (4.13) e (4.16). Pode-se verificar que, embora significativamente distintas, essas equações sofrem basicamente a mesma influência decorrente da existência da não-linearidade presente no caminho secundário. De maneira ainda mais explícita, o mesmo ocorre com o valor médio dos coeficientes em regime permanente (equações (4.24), (4.25) e (4.28)), onde a influência da não-linearidade resume-se a um escalamento das respostas que seriam obtidas se os sistemas fossem puramente lineares.

Em aplicações práticas de sistemas ANC, são utilizados um grande número de coeficientes (chegando em algumas aplicações até 5000), fazendo com que os modelos analíticos desenvolvidos aqui possuam uma grande carga computacional em sua implementação (vide capítulo 6). Portanto, um modelo analítico simplificado de baixa

complexidade computacional mas que ainda assim mantenha informações relevantes sobre os efeitos não-lineares é de grande interesse para o projetista de sistemas ANC, de modo que diversos testes e avaliações poderiam ser realizados em um curto espaço de tempo.

Nesta seção são feitas considerações mais restritivas sobre as características dos sinais e sistemas envolvidos no sistema ANC de modo a obter-se um modelo de baixa carga computacional mas ainda assim representativo dos efeitos não-lineares.

Conforme já descrito no capítulo 2, a consideração de sinais gaussianos brancos (não-correlacionados) é comum em várias aplicações e, embora muitas vezes não corresponda à realidade, ainda assim os modelos obtidos mantêm importantes propriedades do sistema analisado.

Supondo-se, então, que (a) o caminho secundário do sistema ANC possa ser considerado sem memória, (b) que o sinal de referência seja considerado branco e gaussiano e (c) que o passo de convergência seja suficientemente pequeno de tal modo que as flutuações dos coeficientes sejam desprezíveis com relação aos valores médios, a equação (4.16) pode ser aproximada por

$$E\{\mathbf{W}_{BR}(n+1)\} = \left[1 - \mu \frac{\sigma_x^2}{\sqrt{\frac{\sigma_x^2}{\sigma^2} E\{\mathbf{W}_{BR}^T(n)\} E\{\mathbf{W}_{BR}(n)\} + 1}} \right] E\{\mathbf{W}_{BR}(n)\} + \mu \sigma_x^2 \mathbf{W}^o \quad (4.30)$$

Considerando-se ainda que a inicialização do vetor de coeficientes seja realizada na origem $\mathbf{W}(0) = [0 \dots 0]^T$ (consideração usualmente utilizada), então pode ser verificado de (4.30) que já na primeira iteração o vetor $E\{\mathbf{W}_{BR}(n)\}$ é colinear com a planta (\mathbf{W}^o) e se mantém nessa direção durante todo o processo de adaptação: $E\{\mathbf{W}_{BR}(1)\} = \mu \sigma_x^2 \mathbf{W}^o$. Visto que o termo que multiplica $E\{\mathbf{W}_{BR}(n)\}$ é um escalar, então (4.30) pode ser escrita da seguinte forma:

$$E\{\mathbf{W}_{BR}(n)\} = k(n)\mathbf{W}^o \quad (4.31)$$

onde $k(n)$ é um escalar. Substituindo-se (4.31) em (4.30), obtém-se uma equação recursiva escalar dada por:

$$k(n+1) = \left[1 - \mu \frac{\sigma_x^2}{\sqrt{\eta_{BR}^2 k^2(n) + 1}} \right] k(n) + \mu \sigma_x^2 \quad (4.32)$$

onde

$$\eta_{BR}^2 = \frac{\sigma_x^2}{\sigma^2} \mathbf{W}^{oT} \mathbf{W}^o \quad (4.33)$$

é o grau de não-linearidade, correspondente à equação (4.29) para $\mathbf{R}_0 = \sigma_x^2 \mathbf{I}$.

Assumindo-se que exista convergência em (4.32), definindo-se $k_{RP} = \lim_{n \rightarrow \infty} k(n)$ e substituindo-se o resultado em (4.31), chega-se a

$$\lim_{n \rightarrow \infty} E\{\mathbf{W}_{BR}(n)\} = \frac{1}{\sqrt{1 - \eta_{BR}^2}} \mathbf{W}^o \quad (4.34)$$

Ou seja, o modelo simplificado, descrito por (4.31), (4.32) e (4.33), ainda mantém informação acerca dos efeitos não-lineares do sistema.

A equação recursiva escalar (4.32) permite que o comportamento de um grande número de coeficientes seja avaliado com um mínimo de gasto computacional. Entretanto, a informação sobre os efeitos das correlações do sinal de referência e de atrasos intrínsecos ao sistema não é mantida. Nesse caso, é necessária a compreensão de que a diminuição da velocidade de convergência do algoritmo está relacionada ao aumento de correlação.

4.4 Análise da Estabilidade

Da mesma forma que para os algoritmos FXLMS e LMS aplicados a sistemas

lineares, as equações recursivas do comportamento médio dos coeficientes para o caso não-linear estudado neste trabalho também têm sua estabilidade governada pelo passo de convergência (HAYKIN, 1991), como pode ser observado através de (4.11) (4.13), (4.16) e (4.32). Pode-se interpretar μ como o fator de realimentação do sistema.

Entretanto, para o caso não-linear um novo fator pode ser identificado como corresponsável pela estabilidade do algoritmo. Observando-se as equações (4.24), (4.25), (4.28) e (4.34), pode-se verificar que à medida que $\eta^2 \rightarrow 1$ o valor médio dos coeficientes cresce sem limites, e para $\eta^2 > 1$ a equação do comportamento médio dos pesos não possui pontos estacionários, independentemente da escolha do passo de convergência.

Esse fator torna-se mais importante quando se considera que na maioria das aplicações práticas faz-se uso de um passo de convergência pequeno, de forma a minimizar o desajuste em regime permanente e sua variabilidade, já que o algoritmo estará sendo aplicado a apenas uma realização do processo que define o sinal de entrada. Além disso, fatores econômicos evitam o superdimensionamento dos componentes do sistema de controle. Portanto, alto-falantes e amplificadores podem ser utilizados ao máximo de suas especificações, o que pode acarretar efeitos altamente não-lineares. Dessa forma, a influência de η^2 sobre a estabilidade do algoritmo pode ser preponderante sobre μ em aplicações práticas.

4.5 Interpretação do Parâmetro η^2

Nesta seção é demonstrado que o parâmetro η^2 , responsável pela estabilidade do algoritmo, é o denominado grau de não-linearidade do sistema adaptativo.

Segundo a equação (2.16) do capítulo 2, o grau de não-linearidade do sistema adaptativo é determinado por

$$\eta^2 = \frac{1}{\sigma^2} E \{ y_f^2(n) \} \Big|_{\substack{n \rightarrow \infty \\ \sigma^2 \rightarrow \infty}} \quad (4.35)$$

Supondo-se que o sistema esteja em regime permanente, equacionando-se o sinal de controle filtrado $y_f(n)$ (na entrada da não-linearidade), elevando-se ao quadrado e tomando-se seu valor esperado, chega-se a

$$E \{ y_f^2(n) \} = \mathbf{W}_{RP}^T \tilde{\mathbf{R}}_{S_1 S_1} \mathbf{W}_{RP} \quad (4.36)$$

onde $\tilde{\mathbf{R}}_{S_1 S_1}$ é a matriz de autocorrelação do sinal $x(n)$ filtrado por S_1 .

A equação (4.36) pode ser interpretada como a potência do sinal de referência $x(n)$ filtrado pelo sistema S_1 e em seguida por \mathbf{W}_{RP} . Como ambos os sistemas são lineares, a ordem de filtragem não altera o resultado.

Supondo-se ainda que o sistema adaptativo seja linear, então o vetor de coeficientes em regime permanente é dado pela equação (4.24), para $\eta^2 = 0$. Substituindo-se esse resultado e (4.36) em (4.35) chega-se a

$$\eta^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{S}}^T \tilde{\mathbf{R}}_{S_2 \hat{S} S_1}^{-T} \tilde{\mathbf{R}}_{S_1 S_1} \tilde{\mathbf{R}}_{S_2 \hat{S} S_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (4.37)$$

A equação (4.37) é idêntica à equação (4.23), responsável pela estabilidade do algoritmo. As equações (4.26), (4.29) e (4.33) descrevem casos particulares de (4.23).

A partir desse desenvolvimento, conclui-se que a figura de mérito denominada de grau de não-linearidade do sistema adaptativo (descrita no capítulo 2) é responsável não apenas pela determinação dos efeitos da não-linearidade sobre o desempenho do algoritmo adaptativo como também representa um papel chave nas condições de estabilidade do mesmo.

No caso em que houver a correta estimação da parte linear do caminho secundário ($\hat{S} = S_1 * S_2$), então a equação (4.37) torna-se idêntica à equação (3.39). Portanto, β^2 em (3.39) corresponde ao grau de não-linearidade no caso de estimação perfeita da parte linear do caminho secundário.

5. ERRO MÉDIO QUADRÁTICO

Neste capítulo é analisado o comportamento do erro médio quadrático do sistema adaptativo apresentado na figura 4.1, tanto durante o período de adaptação como em regime permanente.

São desenvolvidas equações recursivas para o erro médio quadrático em termos de momentos de até segunda ordem para sinais gaussianos. Através desses resultados são obtidas expressões que modelam o comportamento do algoritmo em regime permanente.

5.1 Análise do Erro Médio Quadrático

A equação do sinal de erro (2.3) do sistema apresentado na figura 4.1 é reproduzida a seguir:

$$\begin{aligned} e(n) &= d(n) + z(n) - \hat{d}(n) \\ &= \mathbf{W}^{oT} \mathbf{X}(n) + z(n) - \sum_{j=0}^{M_2-1} s_{2j} \mathcal{G} \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \end{aligned} \quad (5.1)$$

Elevando-se (5.1) ao quadrado e tomando seu valor esperado, condicionado no conjunto de vetores de coeficientes atrasados \mathcal{W} (equação (4.2)), obtém-se:

$$\begin{aligned} E\{e^2(n) | \mathcal{W}\} &= \\ &= \mathbf{W}^{oT} E\{\mathbf{X}(n) \mathbf{X}^T(n) | \mathcal{W}\} \mathbf{W}^o + 2\mathbf{W}^{oT} E\{z(n) \mathbf{X}(n) | \mathcal{W}\} + E\{z^2(n) | \mathcal{W}\} \\ &\quad - 2 \sum_{j=0}^{M_2-1} s_{2j} E\left\{ z(n) \mathcal{G} \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] | \mathcal{W} \right\} \\ &\quad - 2\mathbf{W}^{oT} \sum_{j=0}^{M_2-1} s_{2j} E\left\{ \mathcal{G} \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}(n) | \mathcal{W} \right\} \\ &\quad + E\left\{ \left(\sum_{j=0}^{M_2-1} s_{2j} \mathcal{G} \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \right)^2 | \mathcal{W} \right\} \end{aligned} \quad (5.2)$$

Os quatro primeiros valores esperados de (5.2) são facilmente avaliados em função das características dos sinais envolvidos:

$$E\{\mathbf{X}(n)\mathbf{X}^T(n)|\mathcal{W}\} = \mathbf{R}_0 \quad (5.3a)$$

$$E\{z(n)\mathbf{X}(n)|\mathcal{W}\} = E\{z(n)\} \cdot E\{\mathbf{X}(n)\} = \mathbf{0} \quad (5.3b)$$

$$E\{z^2(n)|\mathcal{W}\} = E\{z^2(n)\} = \sigma_z^2 \quad (5.3c)$$

$$E\left\{z(n)g\left[\sum_{i=0}^{M_1-1} s_{1i}\mathbf{W}^T(n-i-j)\mathbf{X}(n-i-j)\right]\right\} = E\{z(n)|\mathcal{W}\} \cdot E\{g(\cdot)|\mathcal{W}\} = 0 \quad (5.3d)$$

A primeira equação em (5.3) corresponde à matriz de autocorrelação do sinal de referência. O segundo e o quarto valores esperados são nulos em decorrência da independência entre o ruído de medição e o sinal de referência. O terceiro valor esperado corresponde à variância do ruído de medição.

Da mesma forma que para a análise do comportamento médio dos coeficientes, as correlações entre vetores de coeficientes e vetores de dados observados são desprezadas frente às correlações entre vetores de dados.

O quinto termo de (5.2) pode ser resolvido através do Teorema Modificado de Price (BUSSGANG, 1952, PRICE, 1958, PAWULA, 1967). Esse valor esperado é semelhante ao utilizado na avaliação do comportamento médio dos coeficientes (equação (4.7)), com a diferença de que o vetor aleatório \mathbf{Y}_2 e sua correlação com a variável y_1 são dados por

$$\left\{ \begin{array}{l} \mathbf{Y}_2 = \mathbf{X}(n) \end{array} \right. \quad (5.4a)$$

$$\left\{ \begin{array}{l} E\{y_1\mathbf{Y}_2|\mathbf{W}(n)\} = \sum_{i=0}^{M_1-1} s_{1i}\mathbf{R}_{j+i}\mathbf{W}(n-i-j) \end{array} \right. \quad (5.4b)$$

Substituindo-se (5.4a), (5.4b), (4.6a) e (4.8a) em (4.7), chega-se a

$$\begin{aligned}
 E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{li} \mathbf{W}^T (n-i-j) \mathbf{X} (n-i-j) \right] \mathbf{X} (n) \middle| \mathcal{W} \right\} = \\
 \frac{\sum_{i=0}^{M_1-1} s_{li} \mathbf{R}_{j+i} \mathbf{W} (n-i-j)}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{li} s_{lp} \mathbf{W}^T (n-i-j) \mathbf{R}_{p-i} \mathbf{W} (n-p-j) + 1}}
 \end{aligned} \tag{5.5}$$

O último termo de (5.2) pode ser descrito por $E \{ g (y_1) g (y_2) | \mathcal{W} \}$, onde

$$\begin{cases}
 y_1 = \sum_{i=0}^{M_1-1} s_{li} \mathbf{W}^T (n-i-j) \mathbf{X} (n-i-j) \\
 y_2 = \sum_{l=0}^{M_1-1} s_{ll} \mathbf{W}^T (n-l-q) \mathbf{X} (n-l-q)
 \end{cases} \tag{5.6}$$

são variáveis aleatórias gaussianas de média zero quando condicionadas em \mathcal{W} .

No ANEXO 2 é demonstrado que

$$E \{ g (y_1) g (y_2) \} = \sigma^2 \arcsen \left(\frac{E \{ y_1 y_2 \}}{\sqrt{(E \{ y_1^2 \} + \sigma^2) (E \{ y_2^2 \} + \sigma^2)}} \right) \tag{5.7}$$

Negligenciando-se as correlações entre vetores de dados e vetores de coeficientes em comparação às correlações entre vetores de dados, as seguintes correlações podem ser avaliadas:

$$E \{ y_1^2 | \mathcal{W} \} = \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{li} s_{lp} \mathbf{W}^T (n-i-j) \mathbf{R}_{p-i} \mathbf{W} (n-p-j) \tag{5.8a}$$

$$E \{ y_2^2 | \mathcal{W} \} = \sum_{l=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{ll} s_{lr} \mathbf{W}^T (n-l-q) \mathbf{R}_{r-l} \mathbf{W} (n-r-q) \tag{5.8b}$$

$$E \{ y_1 y_2 | \mathcal{W} \} = \sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{li} s_{ll} \mathbf{W}^T (n-i-j) \mathbf{R}_{l+q-i-j} \mathbf{W} (n-l-q) \tag{5.8c}$$

Substituindo-se (5.6), (5.8a), (5.8b) e (5.8c) em (5.7), chega-se a

$$\begin{aligned}
 E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T (n-i-j) \mathbf{X}(n-i-j) \right] g \left[\sum_{l=0}^{M_1-1} s_{1l} \mathbf{W}^T (n-l-q) \mathbf{X}(n-l-q) \right] \middle| \mathcal{W} \right\} = \\
 = \sigma^2 \arcsen \left(\frac{\sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \mathbf{W}^T (n-i-j) \mathbf{R}_{l+q-i-j} \mathbf{W}(n-l-q)}{\left(\sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \mathbf{W}^T (n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) + \sigma^2 \right)^{1/2}} \right. \\
 \left. \frac{1}{\left(\sum_{l=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{1l} s_{1r} \mathbf{W}^T (n-l-q) \mathbf{R}_{r-l} \mathbf{W}(n-r-q) + \sigma^2 \right)^{1/2}} \right) \quad (5.9)
 \end{aligned}$$

Substituindo-se os resultados (5.3), (5.5) e (5.9) em (5.2), obtém-se:

$$\begin{aligned}
 E \{ e^2(n) | \mathcal{W} \} = \mathbf{W}^o \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 \\
 - 2 \sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^o \mathbf{R}_{j+i} \mathbf{W}(n-i-j)}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \mathbf{W}^T (n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) + 1}} \\
 + \sigma^2 \sum_{j=0}^{M_2-1} \sum_{q=0}^{M_2-1} s_{2j} s_{2q} \arcsen \left(\frac{\sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \mathbf{W}^T (n-i-j) \mathbf{R}_{l+q-i-j} \mathbf{W}(n-l-q)}{\left(\sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \mathbf{W}^T (n-i-j) \mathbf{R}_{p-i} \mathbf{W}(n-p-j) + \sigma^2 \right)^{1/2}} \right. \\
 \left. \frac{1}{\left(\sum_{l=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{1l} s_{1r} \mathbf{W}^T (n-l-q) \mathbf{R}_{r-l} \mathbf{W}(n-r-q) + \sigma^2 \right)^{1/2}} \right) \quad (5.10)
 \end{aligned}$$

A equação (5.10) modela o comportamento do erro médio quadrático condicionado em \mathcal{W} . Visto que a estatística dos vetores de coeficientes é desconhecida, alguma

aproximação deve ser feita para retirar-se o condicionamento de (5.10) em \mathcal{W} e obter-se um modelo para o comportamento do erro médio quadrático.

Utilizando o mesmo raciocínio usado na modelagem do comportamento médio dos coeficientes, ou seja, supondo-se que o passo de convergência seja suficientemente pequeno, uma possível aproximação para o valor esperado de (5.10) é a seguinte:

$$\begin{aligned} \xi(n) = & \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^{oT} \mathbf{R}_{j+i} E\{\mathbf{W}(n-i-j)\}}{\sqrt{\frac{1}{\sigma^2} \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \text{tr}\{\mathbf{R}_{p-i} \mathbf{K}_{p,i}(n-j)\} + 1}} \\ & + \sigma^2 \sum_{j=0}^{M_2-1} \sum_{q=0}^{M_2-1} s_{2j} s_{2q} \arccos \left(\frac{\sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \text{tr}\{\mathbf{R}_{l+q-i-j} \mathbf{K}_{l+q,i+j}(n)\}}{\left(\sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \text{tr}\{\mathbf{R}_{p-i} \mathbf{K}_{p,i}(n-j)\} + \sigma^2 \right)^{1/2}} \right. \\ & \left. \frac{1}{\left(\sum_{i=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{1i} s_{1r} \text{tr}\{\mathbf{R}_{r-i} \mathbf{K}_{r,i}(n-q)\} + \sigma^2 \right)^{1/2}} \right) \end{aligned} \quad (5.11)$$

onde

$$\begin{cases} \mathbf{K}_{l+q,i+j}(n) = E\{\mathbf{W}(n-l-q) \mathbf{W}^T(n-i-j)\} \\ \mathbf{K}_{p,i}(n-j) = E\{\mathbf{W}(n-p-j) \mathbf{W}^T(n-i-j)\} \end{cases} \quad (5.12)$$

são as matrizes de correlação dos vetores de coeficientes. Esses momentos de segunda ordem serão determinados no capítulo 6. Os vetores $E\{\mathbf{W}(n-i-j)\}$ são determinados através da equação do comportamento médio dos coeficientes (4.11).

A equação (5.11) descreve o erro médio quadrático ao longo do processo de adaptação quando é utilizado o algoritmo FXLMS, e o caminho secundário é modelado por um sistema Wiener-Hammerstein, supondo-se sinais gaussianos.

5.1.1 Sistema Wiener

Quando o caminho secundário pode ser modelado por um sistema Wiener, a equação (5.11) reduz-se a

$$\begin{aligned} \xi_w(n) = & \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - \frac{2 \sum_{i=0}^{M-1} s_i \mathbf{W}^{oT} \mathbf{R}_{-i} E\{\mathbf{W}_w(n-i)\}}{\sqrt{\frac{1}{\sigma^2} \sum_{j=0}^{M-1} \sum_{i=0}^{M-1} s_j s_i \text{tr}\{\mathbf{R}_{j-i} \mathbf{K}_{w_{i,j}}(n)\} + 1}} \\ & + \sigma^2 \arcsen \left(\frac{\sum_{j=0}^{M-1} \sum_{i=0}^{M-1} s_j s_i \text{tr}\{\mathbf{R}_{j-i} \mathbf{K}_{w_{i,j}}(n)\}}{\sum_{j=0}^{M-1} \sum_{i=0}^{M-1} s_j s_i \text{tr}\{\mathbf{R}_{j-i} \mathbf{K}_{w_{i,j}}(n)\} + \sigma^2} \right) \end{aligned} \quad (5.13)$$

onde $\mathbf{K}_{w_{i,j}}(n) = E\{\mathbf{W}(n-i)\mathbf{W}^T(n-j)\}$ são os momentos de segunda ordem.

No caso de um caminho secundário linear ($\sigma^2 \rightarrow \infty$), é fácil verificar que a equação (5.13) reduz-se à equação (3.52) de TOBIAS (1999), que determina o erro médio quadrático para o algoritmo FXLMS convencional.

5.1.2 Sistema Sem Memória

No caso em que o caminho secundário não possua memória, a equação (5.13) pode ser resumida a

$$\xi_{SM}(n) = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - \frac{2 \mathbf{W}^{oT} \mathbf{R}_{xx} E\{\mathbf{W}_{SM}(n)\}}{\sqrt{\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_{xx} \mathbf{K}_{SM}(n)\} + 1}} + \sigma^2 \arcsen \left(\frac{\text{tr}\{\mathbf{R}_{xx} \mathbf{K}_{SM}(n)\}}{\text{tr}\{\mathbf{R}_{xx} \mathbf{K}_{SM}(n)\} + \sigma^2} \right) \quad (5.14)$$

onde $\mathbf{K}_{SM}(n)$ é a matriz de momentos de segunda ordem.

Para $\sigma^2 \rightarrow \infty$, o sistema sem memória reduz-se à equação (5.14) que é a expressão do erro médio quadrático do algoritmo LMS convencional (HAYKIN, 1991).

5.2 Erro Médio Quadrático em Regime Permanente

Nesta seção será avaliado o erro médio quadrático em regime permanente para as diferentes modelagens do caminho secundário.

Na seção 4.2 foi determinado o valor médio dos coeficientes em regime permanente (equação (4.24)). Substituindo esse valor em (5.11) e assumindo-se que as oscilações (momentos de segunda ordem) sejam desprezíveis frente ao valor médio dos coeficientes,

$$\begin{cases} \lim_{n \rightarrow \infty} E\{\mathbf{W}(n)\} = \mathbf{W}_{RP} \\ \lim_{n \rightarrow \infty} E\{\mathbf{W}(n)\mathbf{W}^T(n)\} \equiv \mathbf{W}_{RP}\mathbf{W}_{RP}^T \end{cases} \quad (5.15)$$

chega-se a:

$$\begin{aligned} \lim_{n \rightarrow \infty} \xi(n) = & \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2\mathbf{W}^{oT} \tilde{\mathbf{R}}_{s_2 s_1}^T \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-1} \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o + \frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-T} \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-1} \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o}{\eta^2} \\ & \cdot \sum_{j=0}^{M_2-1} \sum_{q=0}^{M_2-1} s_{2j} s_{2q} \arcsen \left(\frac{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-T} \sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \mathbf{R}_{i+q-i-j} \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-1} \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o}{\mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-T} \tilde{\mathbf{R}}_{s_1 s_1} \tilde{\mathbf{R}}_{s_2 \hat{s}_1}^{-1} \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o} \cdot \eta^2 \right) \end{aligned} \quad (5.16)$$

onde

$$\begin{cases} \tilde{\mathbf{R}}_{\hat{s}} = \sum_{j=0}^{\hat{M}-1} \hat{s}_k \mathbf{R}_{-k} = \sum_{j=0}^{\hat{M}-1} \hat{s}_k E\{\mathbf{X}(n-k)\mathbf{X}^T(n)\} \\ \tilde{\mathbf{R}}_{s_1 s_1} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} \mathbf{R}_{i-k} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1k} s_{1i} E\{\mathbf{X}(n-k)\mathbf{X}^T(n-i)\} \\ \tilde{\mathbf{R}}_{s_2 s_1} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} s_{2j} s_{1i} \mathbf{R}_{-i-j} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} s_{2j} s_{1i} E\{\mathbf{X}(n-i-j)\mathbf{X}^T(n)\} \\ \tilde{\mathbf{R}}_{s_2 \hat{s}_1} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{2j} s_{1i} \hat{s}_k \mathbf{R}_{i+j-k} = \sum_{j=0}^{M_2-1} \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{2j} s_{1i} \hat{s}_k E\{\mathbf{X}(n-k)\mathbf{X}^T(n-i-j)\} \end{cases} \quad (5.17)$$

A equação (5.16) modela o erro médio quadrático em regime permanente, para o sistema ANC descrito na figura 4.1.

Através de (5.16) verifica-se que o EMQ em regime permanente, ou seja, o nível de cancelamento do sistema ANC, depende diretamente do grau de não-linearidade do sistema.

5.2.1 Sistema Wiener

Quando o caminho secundário puder ser representado por um sistema Wiener, o erro médio quadrático em regime permanente (5.16) torna-se

$$\lim_{n \rightarrow \infty} \xi_{\mathcal{W}}(n) = \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 - 2 \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o + \mathbf{W}^{oT} \tilde{\mathbf{R}}_S^T \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_{SS} \tilde{\mathbf{R}}_{SS}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \left[\frac{1}{\eta_w^2} \arcsen(\eta_w^2) \right] \quad (5.18)$$

onde:

$$\begin{cases} \tilde{\mathbf{R}}_{\hat{S}} = \sum_{j=0}^{\hat{M}-1} \hat{s}_k \mathbf{R}_{-k} = \sum_{j=0}^{\hat{M}-1} \hat{s}_k E \{ \mathbf{X}(n-k) \mathbf{X}^T(n) \} \\ \tilde{\mathbf{R}}_{\hat{S}\hat{S}} = \sum_{i=0}^{M-1} \sum_{k=0}^{\hat{M}-1} s_i \hat{s}_k \mathbf{R}_{i+j-k} = \sum_{i=0}^{M-1} \sum_{k=0}^{\hat{M}-1} s_i \hat{s}_k E \{ \mathbf{X}(n-k) \mathbf{X}^T(n-i-j) \} \end{cases} \quad (5.19)$$

Caso o sistema possua um caminho secundário puramente linear ($\eta_w^2 \rightarrow 0$), situação que corresponde ao algoritmo FXLMS convencional, então (5.18) reduz-se à soma das expressões (2.26) e (2.45) de TOBIAS (1999), que trata do regime permanente do algoritmo FXLMS convencional.

5.2.2 Sistema Sem Memória

Se o caminho secundário pode ser considerado como um sistema sem memória, a equação (5.18) é simplificada, resultando em

$$\lim_{n \rightarrow \infty} \xi_{SM}(n) = \left[\frac{1}{\eta_{SM}^2} \arccos(\eta_{SM}^2) - 1 \right] \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o + \sigma_z^2 \quad (5.20)$$

A análise da equação (5.20) permite concluir que o cancelamento obtido para o sistema sem memória é diretamente dependente do grau de não-linearidade do sistema adaptativo. Para o caso em que o sistema é linear ($\eta_w^2 \rightarrow 0$), então a equação (5.20) concorda com o resultado obtido para o regime permanente do algoritmo LMS convencional (HAYKIN, 1991).

5.3 Modelo Simplificado

No caso do modelo simplificado, discutido na seção 4.3, onde se supõe que o sinal de referência seja branco e a inicialização seja feita na origem ($\mathbf{W}(0) = \mathbf{0}$), a equação do erro médio quadrático é obtida substituindo-se

$$\begin{cases} \mathbf{R}_0 = \sigma_x^2 \mathbf{I} \\ E\{\mathbf{W}_{BR}(n)\} = k(n) \mathbf{W}^o \\ \mathbf{K}_{BR}(n) = k^2(n) \mathbf{W}^{oT} \mathbf{W}^o \end{cases} \quad (5.21)$$

em (5.14), de tal forma que

$$\xi_{BR}(n) = \sigma_x^2 \mathbf{W}^{oT} \mathbf{W}^o \left[1 - \frac{2k(n)}{\sqrt{\eta_{BR}^2 k^2(n) + 1}} + \frac{1}{\eta_{BR}^2} \arccos\left(\frac{\eta_{BR}^2 k^2(n)}{\eta_{BR}^2 k^2(n) + 1}\right) \right] + \sigma_z^2 \quad (5.22)$$

onde

$$k(n+1) = \left[1 - \mu \frac{\sigma_x^2}{\sqrt{\eta_{BR}^2 k^2(n) + 1}} \right] k(n) + \mu \sigma_x^2 \quad (5.23)$$

O erro médio quadrático em regime permanente é obtido substituindo-se

$$\lim_{n \rightarrow \infty} k(n) = \frac{1}{\sqrt{1 - \eta_{BR}^2}} \quad (5.24)$$

em (5.22), resultando em:

$$\lim_{n \rightarrow \infty} \xi_{BR}(n) = \sigma_x^2 \mathbf{W}^{oT} \mathbf{W}^o \left[\frac{1}{\eta_{BR}^2} \arcsen(\eta_{BR}^2) - 1 \right] + \sigma_z^2 \quad (5.25)$$

$$\mathbf{K}_{a,b}(n-c) = \begin{cases} \mathbf{K}_{0,b-a}(n-a-c) & \text{se } (b > a) \\ \mathbf{K}_{0,0}(n-a-c) & \text{se } (b = a) \\ \mathbf{K}_{0,a-b}^T(n-b-c) & \text{se } (b < a) \end{cases} \quad (6.3)$$

Note-se que a matriz $\mathbf{K}_{a,b}(n-c)$ pode ser obtida, para quaisquer índices a , b e c , a partir das matrizes $\mathbf{K}_{0,0}(n)$ e $\mathbf{K}_{0,k}(n)$ calculadas para diferentes atrasos de tempo. Em decorrência desse fato, serão determinadas equações recursivas para essas duas matrizes.

- **Determinação do Comportamento de $\mathbf{K}_{0,k}(n)$ para $k > 0$**

A equação de atualização dos coeficientes (2.4) é descrita por

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu \left\{ \mathbf{W}^{oT} \mathbf{X}(n) + z(n) - \sum_{j=0}^{M_2-1} s_{2j} g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \right\} \mathbf{X}_f(n) \quad (6.4)$$

Pós-multiplicando-se a equação (6.4) por $\mathbf{W}^T(n-k+1)$ e aplicando-se a esperança condicionada em \mathcal{W} (equação (4.2)), tem-se como resultado

$$\begin{aligned} E \left\{ \mathbf{W}(n+1) \mathbf{W}^T(n-k+1) \middle| \mathcal{W} \right\} &= \mathbf{W}(n) \mathbf{W}^T(n-k+1) \\ &+ \mu E \left\{ \mathbf{X}_f(n) \mathbf{X}^T(n) \middle| \mathcal{W} \right\} \mathbf{W}^o \mathbf{W}^T(n-k+1) \\ &- \mu \sum_{j=0}^{M_2-1} s_{2j} E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}_f(n) \middle| \mathcal{W} \right\} \mathbf{W}^T(n-k+1) \end{aligned} \quad (6.5)$$

Para a obtenção da equação (6.5) utiliza-se a suposição de que o sinal de excitação é independente do ruído de medição, e ambos possuem média zero.

As esperanças encontradas em (6.5) já foram avaliadas em (4.4a) e (4.9); substituindo esses resultados, obtém-se

$$E\{\mathbf{W}(n+1)\mathbf{W}^T(n-k+1)|\mathfrak{D}\} = \mathbf{W}(n)\mathbf{W}^T(n-k+1) + \mu\tilde{\mathbf{R}}_s\mathbf{W}^o\mathbf{W}^T(n-k+1) - \mu\sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} \hat{s}_q \mathbf{R}_{j+p-q} \mathbf{W}(n-p-j)\mathbf{W}^T(n-k+1)}{\sqrt{\frac{1}{\sigma^2} \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \mathbf{W}^T(n-p-j)\mathbf{R}_{q-p}\mathbf{W}(n-q-j)+1}} \quad (6.6)$$

Supondo-se μ suficientemente pequeno e substituindo-se a notação na forma de (6.1), aproxima-se o valor esperado de (6.6) por

$$\mathbf{K}_{0,k}(n+1) = \mathbf{K}_{0,k-1}(n) + \mu\tilde{\mathbf{R}}_s\mathbf{W}^o E\{\mathbf{W}^T(n-k+1)\} - \mu\sum_{j=0}^{M_2-1} s_{2j} \frac{\sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} \hat{s}_q \mathbf{R}_{j+p-q} \mathbf{K}_{j+p,k-1}(n)}{\sqrt{\frac{1}{\sigma^2} \sum_{p=0}^{M_1-1} \sum_{q=0}^{M_1-1} s_{1p} s_{1q} \text{tr}\{\mathbf{R}_{q-p}\mathbf{K}_{q,p}(n-j)\}+1}} \quad (6.7)$$

onde $\mathbf{K}_{j+p,k-1}(n)$ e $\mathbf{K}_{q,p}(n-j)$ são obtidas a partir de (6.3).

Examinando-se a equação (6.7), verifica-se que a partir de $\mathbf{K}_{0,0}(n)$ é possível obter-se as matrizes $\mathbf{K}_{0,k}(n+1)$ para $k > 0$.

A equação (6.7) é a solução final para a primeira equação recursiva necessária para a avaliação dos momentos de segunda ordem.

Supondo-se que $\sigma \rightarrow \infty$, a equação (6.7) reduz-se à equação de $\mathbf{K}_{0,k}(n)$, que descreve o comportamento dos momentos de segunda ordem do algoritmo FXLMS convencional (TOBIAS, 1999).

- **Determinação do Comportamento de $\mathbf{K}_{0,0}(n)$**

O primeiro passo para obter-se uma equação recursiva para $\mathbf{K}_{0,0}(n)$ é pós-

multiplicar a equação dos coeficientes (2.4) por sua transposta e tomar-se a esperança condicionada em \mathcal{W} . Procedendo-se dessa maneira, chega-se a

$$E\{\mathbf{W}(n+1)\mathbf{W}^T(n+1)|\mathcal{W}\} = \mathbf{W}(n)\mathbf{W}^T(n) + \mu \left[E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\} \mathbf{W}^o \mathbf{W}^T(n) \right] \quad (\text{T1})$$

$$+ \mu \left[E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\} \mathbf{W}^o \mathbf{W}^T(n) \right]^T \quad (\text{T2})$$

$$+ \mu \left[E\{z(n)\mathbf{X}_f(n)|\mathcal{W}\} \mathbf{W}^T(n) \right] \quad (\text{T3})$$

$$+ \mu \left[E\{z(n)\mathbf{X}_f(n)|\mathcal{W}\} \mathbf{W}^T(n) \right]^T \quad (\text{T4})$$

$$+ 2\mu^2 E\{z(n)\mathbf{W}^{oT} \mathbf{X}(n)\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\} \quad (\text{T5})$$

$$- \mu \left[\sum_{j=0}^{M_2-1} s_{2j} E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}_f(n) \middle| \mathcal{W} \right\} \mathbf{W}^T(n) \right] \quad (\text{T6})$$

$$- \mu \left[\sum_{j=0}^{M_2-1} s_{2j} E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}_f(n) \middle| \mathcal{W} \right\} \mathbf{W}^T(n) \right]^T \quad (\text{T7})$$

$$+ \mu^2 E\{\mathbf{X}_f(n)\mathbf{X}^T(n)\mathbf{W}^o \mathbf{W}^{oT} \mathbf{X}(n)\mathbf{X}_f^T(n)|\mathcal{W}\} \quad (\text{T8})$$

$$+ \mu^2 E\{z^2(n)\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\} \quad (\text{T9})$$

$$- 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{W}^{oT} \mathbf{X}(n)\mathbf{X}_f(n)\mathbf{X}_f^T(n) \middle| \mathcal{W} \right\} \quad (\text{T10})$$

$$- 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} E \left\{ z(n) g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \mathbf{X}_f(n)\mathbf{X}_f^T(n) \middle| \mathcal{W} \right\} \quad (\text{T11})$$

$$+ \mu^2 \sum_{j=0}^{M_2-1} \sum_{q=0}^{M_2-1} s_{2j} s_{2q} E \left\{ g \left[\sum_{i=0}^{M_1-1} s_{1i} \mathbf{W}^T(n-i-j) \mathbf{X}(n-i-j) \right] \cdot g \left[\sum_{l=0}^{M_1-1} s_{1l} \mathbf{W}^T(n-l-q) \mathbf{X}(n-l-q) \right] \mathbf{X}_f(n)\mathbf{X}_f^T(n) \middle| \mathcal{W} \right\} \quad (\text{T12})$$

(6.8)

O valor esperado em (T1) corresponde à matriz $\tilde{\mathbf{R}}_s$ definida na equação (4.5), e por conseguinte (T2) é sua transposta. Os valores esperados (T3), (T4), (T5), (T9) e (T11) são nulos, visto que o ruído de medição e o sinal de excitação são não-correlacionados e de

média zero, de forma que

$$\left\{ \begin{array}{l} E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{s}} \\ E\{z(n)\mathbf{X}_f(n)|\mathcal{W}\} = \mathbf{0} \\ E\{z(n)\mathbf{W}^{oT}\mathbf{X}(n)\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\} = \mathbf{0} \\ E\{z^2(n)|\mathcal{W}\} = \sigma_z^2 \\ E\{\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \\ E\left\{z(n)g\left[\sum_{i=0}^{M_1-1} s_{1i}\mathbf{W}^T(n-i-j)\mathbf{X}(n-i-j)\right]\mathbf{X}_f(n)\mathbf{X}_f^T(n)\right|\mathcal{W}\} = \mathbf{0} \end{array} \right. \quad (6.9)$$

onde

$$\tilde{\mathbf{R}}_{\hat{s}\hat{s}} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} \hat{s}_k \hat{s}_i \mathbf{R}_{i-k} = \sum_{k=0}^{M_1-1} \sum_{i=0}^{M_1-1} \hat{s}_k \hat{s}_i E\{\mathbf{X}(n-k)\mathbf{X}^T(n-i)\} \quad (6.10)$$

A avaliação de $\mathbf{K}_{0,0}(n)$ depende de muitos termos, e, portanto, de forma a reduzir o tamanho das expressões, será utilizada a seguinte notação até o final do capítulo:

$$\left\{ \begin{array}{l} y_1 = \sum_{i=0}^{M_1-1} s_{1i}\mathbf{W}^T(n-i-j)\mathbf{X}(n-i-j) \\ \mathbf{Y}_2 = \mathbf{X}_f(n) \\ y_3 = \mathbf{W}^{oT}\mathbf{X}(n) \\ y_4 = \sum_{l=0}^{M_1-1} s_{1l}\mathbf{W}^T(n-l-q)\mathbf{X}(n-l-q) \end{array} \right. \quad (6.11)$$

Utilizando-se o Teorema Modificado de Price (ANEXO 1), obtêm-se os valores esperados em (T6) e (T7). As expressões de (T6) e (T7) podem ser escritas em termos da convenção expressa em (6.11), na seguinte forma:

$$E\{g(y_1)\mathbf{Y}_2|\mathcal{W}\} = \frac{E\{y_1\mathbf{Y}_2|\mathcal{W}\}}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}} \quad (6.12)$$

O termo (T8) pode ser obtido através do ANEXO 3:

$$\begin{aligned}
 E\{\mathbf{X}_f(n)\mathbf{X}^T(n)\mathbf{W}^o\mathbf{W}^{oT}\mathbf{X}(n)\mathbf{X}_f^T(n)|\mathcal{W}\} = \\
 = 2E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\}\mathbf{W}^o\mathbf{W}^{oT}E\{\mathbf{X}(n)\mathbf{X}_f^T(n)|\mathcal{W}\} \\
 + \mathbf{W}^{oT}E\{\mathbf{X}(n)\mathbf{X}^T(n)|\mathcal{W}\}\mathbf{W}^o \cdot E\{\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\}
 \end{aligned} \tag{6.13}$$

Determinando-se as correlações presentes em (6.13),

$$\begin{cases}
 E\{\mathbf{X}_f(n)\mathbf{X}^T(n)|\mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{s}} \\
 E\{\mathbf{X}_f(n)\mathbf{X}_f^T(n)|\mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \\
 E\{\mathbf{X}(n)\mathbf{X}^T(n)|\mathcal{W}\} = \mathbf{R}_0
 \end{cases} \tag{6.14}$$

pode-se obter o seguinte resultado para (T8):

$$E\{\mathbf{X}_f(n)\mathbf{X}^T(n)\mathbf{W}^o\mathbf{W}^{oT}\mathbf{X}(n)\mathbf{X}_f^T(n)|\mathcal{W}\} = 2\tilde{\mathbf{R}}_{\hat{s}}\mathbf{W}^o\mathbf{W}^{oT}\tilde{\mathbf{R}}_{\hat{s}}^T + \mathbf{W}^{oT}\mathbf{R}_0\mathbf{W}^o\tilde{\mathbf{R}}_{\hat{s}\hat{s}} \tag{6.15}$$

O valor esperado de (T10) pode ser determinado utilizando-se a seguinte expressão desenvolvida no ANEXO 4:

$$\begin{aligned}
 E\{g(y_1)y_3\mathbf{Y}_2\mathbf{Y}_2^T|\mathcal{W}\} = & \left[E\{y_1y_3|\mathcal{W}\}E\{\mathbf{Y}_2\mathbf{Y}_2^T|\mathcal{W}\} + E\{y_1\mathbf{Y}_2|\mathcal{W}\}E\{y_3\mathbf{Y}_2^T|\mathcal{W}\} \right. \\
 & \left. + E\{y_3\mathbf{Y}_2|\mathcal{W}\}E\{y_1\mathbf{Y}_2^T|\mathcal{W}\} \right] \frac{E\{g(y_1)y_1|\mathcal{W}\}}{E\{y_1^2|\mathcal{W}\}} \\
 & + E\{y_1y_3|\mathcal{W}\}E\{y_1\mathbf{Y}_2|\mathcal{W}\}E\{y_1\mathbf{Y}_2^T|\mathcal{W}\} \left[\frac{E\{g(y_1)y_1^3|\mathcal{W}\}}{E\{y_1^2|\mathcal{W}\}^3} - 3 \frac{E\{g(y_1)y_1|\mathcal{W}\}}{E\{y_1^2|\mathcal{W}\}^2} \right]
 \end{aligned} \tag{6.16}$$

Os valores esperados de funções não-lineares encontradas em (6.16) são determinados segundo o ANEXO 5:

$$E\{g(y_1)y_1^3|\mathcal{W}\} = \frac{E\{y_1^2|\mathcal{W}\}^2}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{3/2}} + 2\frac{E\{y_1^2|\mathcal{W}\}^2}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}} \quad (6.17)$$

e do ANEXO 1:

$$E\{g(y_1)y_1|\mathcal{W}\} = \frac{E\{y_1^2|\mathcal{W}\}}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}} \quad (6.18)$$

Substituindo-se (6.17) e (6.18) em (6.16), obtém-se a solução para (T10):

$$E\{g(y_1)y_3Y_2Y_2^T|\mathcal{W}\} = -\frac{E\{y_1y_3|\mathcal{W}\}E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}}{\sigma^2\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{3/2}} \quad (6.19)$$

$$+ \frac{\left[E\{y_1y_3|\mathcal{W}\}E\{Y_2Y_2^T|\mathcal{W}\} + E\{y_1Y_2|\mathcal{W}\}E\{y_3Y_2^T|\mathcal{W}\} + E\{y_3Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\right]}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}}$$

Substituindo-se os resultados (6.9), (6.12), (6.15) e (6.19) em (6.8), chega-se a

$$E\{\mathbf{W}(n+1)\mathbf{W}^T(n+1)|\mathcal{W}\} = \mathbf{W}(n)\mathbf{W}^T(n) + \mu\tilde{\mathbf{R}}_{\hat{s}}\mathbf{W}^o\mathbf{W}^T(n) + \mu\mathbf{W}(n)\mathbf{W}^{oT}\tilde{\mathbf{R}}_{\hat{s}}^T$$

$$- \mu\sum_{j=0}^{M_2-1} s_{2j} \frac{\left[E\{y_1Y_2|\mathcal{W}\}\mathbf{W}^T(n) + \mathbf{W}(n)E\{y_1Y_2^T|\mathcal{W}\}\right]}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}} + 2\mu^2\tilde{\mathbf{R}}_{\hat{s}}\mathbf{W}^o\mathbf{W}^{oT}\tilde{\mathbf{R}}_{\hat{s}}^T$$

$$+ \mu^2\mathbf{W}^{oT}\mathbf{R}_{xx}\mathbf{W}^o\tilde{\mathbf{R}}_{\hat{s}\hat{s}} + \mu^2\sigma_z^2\tilde{\mathbf{R}}_{\hat{s}\hat{s}} + 2\mu^2\sum_{j=0}^{M_2-1} s_{2j} \frac{E\{y_1y_3|\mathcal{W}\}E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}}{\sigma^2\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{3/2}}$$

$$- 2\mu^2\sum_{j=0}^{M_2-1} s_{2j} \frac{\left[E\{y_1y_3|\mathcal{W}\}E\{Y_2Y_2^T|\mathcal{W}\} + E\{y_1Y_2|\mathcal{W}\}E\{y_3Y_2^T|\mathcal{W}\} + E\{y_3Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\right]}{\left(\frac{1}{\sigma^2}E\{y_1^2|\mathcal{W}\}+1\right)^{1/2}}$$

$$+ \mu^2\sum_{j=0}^{M_2-1}\sum_{q=0}^{M_2-1} s_{2j}s_{2q} E\{g(y_1)g(y_4)Y_2Y_2^T\} \quad (6.20)$$

Falta ainda avaliar o último valor esperado de (6.20). Para tanto, utilizamos o ANEXO 6, que estabelece três modelagens diferentes, dependendo das propriedades dos sinais envolvidos:

- caso 1: se $E\{y_1^2\} = 0$ ou $E\{y_2^2\} = 0$;
- caso 2: se $E\{y_1^2\}E\{y_4^2\} - E\{y_1 y_4\}^2 = 0$;
- caso 3: se $E\{y_1^2\}E\{y_4^2\} - E\{y_1 y_4\}^2 \neq 0$.

A necessidade de diferentes soluções para diferentes características das variáveis de entrada decorre da complexidade matemática para a avaliação dos valores esperados envolvidos. Portanto, a análise do problema por partes simplifica o desenvolvimento das soluções.

- Caso 1: A primeira condição acima apenas informa a inexistência de um dos sinais escalares envolvidos (potência nula). Nesse caso, o resultado da função será nulo.
- Caso 2: Se ambos os sinais y_1 e y_2 possuírem potências não nulas e a seguinte condição for verificada, $E\{y_1^2\}E\{y_4^2\} - E\{y_1 y_4\}^2 = 0$, mostra-se no ANEXO 6 que

$$E\{g(y_1)g(y_4)Y_2 Y_2^T | \mathcal{W}\} = \frac{E\{y_1 Y_2 | \mathcal{W}\}E\{y_1 Y_2^T | \mathcal{W}\}}{E\{y_1^2 | \mathcal{W}\}^2} E\{g(y_1)g(y_4)y_1^2 | \mathcal{W}\} + \left[E\{Y_2 Y_2^T | \mathcal{W}\} - \frac{E\{y_1 Y_2 | \mathcal{W}\}E\{y_1 Y_2^T | \mathcal{W}\}}{E\{y_1^2 | \mathcal{W}\}} \right] E\{g(y_1)g(y_4) | \mathcal{W}\} \quad (6.21)$$

- Caso 3: Se $E\{y_1^2\}E\{y_4^2\} - E\{y_1 y_4\}^2 \neq 0$, então a solução para o valor esperado desejado é dada por (ANEXO 6):

$$\begin{aligned}
E\{g(y_1)g(y_4)Y_2Y_2^T|\mathcal{W}\} &= \left[E\{Y_2Y_2^T|\mathcal{W}\} \right. \\
&\quad \frac{E\{y_1^2|\mathcal{W}\}E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4^2|\mathcal{W}\}E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}}{E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2} \\
&\quad \left. + \frac{E\{y_1y_4|\mathcal{W}\} \left[E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2} \right] E\{g(y_1)g(y_4)|\mathcal{W}\} \\
&\quad + \left[\frac{E\{y_4^2|\mathcal{W}\}^2 E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right. \\
&\quad \frac{E\{y_4^2|\mathcal{W}\}E\{y_1y_4|\mathcal{W}\} \left[E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \\
&\quad \left. + \frac{E\{y_1y_4|\mathcal{W}\}^2 E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right] E\{g(y_1)g(y_4)y_1^2|\mathcal{W}\} \\
&\quad + \left[\frac{E\{y_1^2|\mathcal{W}\}^2 E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right. \\
&\quad \frac{E\{y_1^2|\mathcal{W}\}E\{y_1y_4|\mathcal{W}\} \left[E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \\
&\quad \left. + \frac{E\{y_1y_4|\mathcal{W}\}^2 E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right] E\{g(y_1)g(y_4)y_4^2|\mathcal{W}\} \\
&\quad + \left[\frac{E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} \left[E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right. \\
&\quad \left. - 2 \frac{E\{y_1y_4|\mathcal{W}\} \left[E\{y_1^2|\mathcal{W}\}E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4^2|\mathcal{W}\}E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right. \\
&\quad \left. + \frac{E\{y_1y_4|\mathcal{W}\}^2 \left[E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\} + E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\} \right]}{\left(E\{y_1^2|\mathcal{W}\}E\{y_4^2|\mathcal{W}\} - E\{y_1y_4|\mathcal{W}\}^2 \right)^2} \right] E\{g(y_1)g(y_4)y_1y_4|\mathcal{W}\}
\end{aligned}
\tag{6.22}$$

As equações (6.21) e (6.22) expressam $E\{g(y_1)g(y_4)Y_2Y_2^T|\mathcal{W}\}$ em termos de valores esperados de no máximo duas variáveis aleatórias diferentes. Os termos que envolvem expressões não-lineares podem ser encontrados nos ANEXOS 2, 7 e 8 e são descritos a seguir:

$$E\{g(y_1)g(y_4)|\mathcal{W}\} = \sigma^2 \arcsen \left(\frac{E\{y_1y_4|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right)^{1/2} \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right)^{1/2}} \right) \quad (6.23)$$

$$\begin{aligned} E\{g(y_1)g(y_4)y_1^2|\mathcal{W}\} &= \sigma^2 E\{y_1^2|\mathcal{W}\} \arcsen \left(\frac{E\{y_1y_4|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right)^{1/2} \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right)^{1/2}} \right) \\ &+ \frac{\sigma^2 E\{y_1^2|\mathcal{W}\} E\{y_1y_4|\mathcal{W}\} \left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right)}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) \left[\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) - E\{y_1y_4|\mathcal{W}\}^2\right]^{1/2}} \\ &+ \frac{\sigma^4 E\{y_1^2|\mathcal{W}\} E\{y_1y_4|\mathcal{W}\} \left[E\{y_4^2|\mathcal{W}\} + \sigma^2\right] - \left[\sigma^4 + \sigma^2 E\{y_1^2|\mathcal{W}\}\right] E\{y_1y_4|\mathcal{W}\}^3}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) \left[\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) - E\{y_1y_4|\mathcal{W}\}^2\right]^{1/2}} \end{aligned} \quad (6.24)$$

$$\begin{aligned} E\{g(y_1)g(y_4)y_1y_4|\mathcal{W}\} &= \sigma^2 E\{y_1y_4|\mathcal{W}\} \arcsen \left(\frac{E\{y_1y_4|\mathcal{W}\}}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right)^{1/2} \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right)^{1/2}} \right) \\ &+ \frac{\sigma^2 E\{y_1^2|\mathcal{W}\} E\{y_4^2|\mathcal{W}\} \left[\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) - E\{y_1y_4|\mathcal{W}\}^2\right] + \sigma^6 E\{y_1y_4|\mathcal{W}\}^2}{\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) \left[\left(E\{y_1^2|\mathcal{W}\} + \sigma^2\right) \left(E\{y_4^2|\mathcal{W}\} + \sigma^2\right) - E\{y_1y_4|\mathcal{W}\}^2\right]^{1/2}} \end{aligned} \quad (6.25)$$

Para obter-se os momentos de segunda ordem deve-se calcular o valor esperado da equação (6.20), levando-se em consideração a estatística dos $W(n-k)$. Entretanto, a função densidade de probabilidade dos vetores de coeficientes não é conhecida. Utilizando-

se a suposição de que o passo de convergência é suficientemente pequeno, pode-se obter a seguinte aproximação para a equação (6.20):

$$\begin{aligned}
\mathbf{K}_{0,0}(n+1) &= \mathbf{K}_{0,0}(n) + \mu \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o E\{\mathbf{W}^T(n)\} + \mu E\{\mathbf{W}(n)\} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \\
&- \mu \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{E\{y_1 \mathbf{Y}_2 | \mathcal{W}\} \mathbf{W}^T(n)\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} - \mu \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{\mathbf{W}(n) E\{y_1 \mathbf{Y}_2^T | \mathcal{W}\}\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} + 2\mu^2 \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \\
&+ \mu^2 \mathbf{W}^{oT} \mathbf{R}_{xx} \mathbf{W}^o \tilde{\mathbf{R}}_{\hat{s}\hat{s}} + \mu^2 \sigma_z^2 \tilde{\mathbf{R}}_{\hat{s}\hat{s}} + 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{y_1 y_3\} \cdot E\{E\{y_1 \mathbf{Y}_2 | \mathcal{W}\} E\{y_1 \mathbf{Y}_2^T | \mathcal{W}\}\}}{\sigma^2 \left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{3/2}} \\
&- 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{y_1 y_3\} \cdot E\{\mathbf{Y}_2 \mathbf{Y}_2^T\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} - 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{E\{y_1 \mathbf{Y}_2 | \mathcal{W}\} E\{y_3 \mathbf{Y}_2^T | \mathcal{W}\}\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} \\
&- 2\mu^2 \sum_{j=0}^{M_2-1} s_{2j} \frac{E\{E\{y_3 \mathbf{Y}_2 | \mathcal{W}\} E\{y_1 \mathbf{Y}_2^T | \mathcal{W}\}\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} + \mu^2 \sum_{j=0}^{M_2-1} \sum_{q=0}^{M_2-1} s_{2j} s_{2q} E\{g(y_1) g(y_4) \mathbf{Y}_2 \mathbf{Y}_2^T\}
\end{aligned} \tag{6.26}$$

onde

$$E\{g(y_1) g(y_4) \mathbf{Y}_2 \mathbf{Y}_2^T\} = \begin{cases} 0 & \text{se } E\{y_1^2\} = 0 \text{ ou } E\{y_2^2\} = 0 \\ \text{Eq. (6.28) se } E\{y_1^2\} E\{y_2^2\} - E\{y_1 y_2\}^2 = 0 & \\ \text{Eq. (6.29) se } E\{y_1^2\} E\{y_2^2\} - E\{y_1 y_2\}^2 \neq 0 & \end{cases} \tag{6.27}$$

Portanto

$$\begin{aligned}
E\{E\{g(y_1) g(y_4) \mathbf{Y}_2 \mathbf{Y}_2^T | \mathcal{W}\}\} &= \frac{E\{E\{y_1 \mathbf{Y}_2 | \mathcal{W}\} E\{y_1 \mathbf{Y}_2^T | \mathcal{W}\}\}}{E\{y_1^2\}^2} E\{E\{g(y_1) g(y_4) y_1^2 | \mathcal{W}\}\} \\
&+ \left[E\{\mathbf{Y}_2 \mathbf{Y}_2^T\} - \frac{E\{E\{y_1 \mathbf{Y}_2 | \mathcal{W}\} E\{y_1 \mathbf{Y}_2^T | \mathcal{W}\}\}}{E\{y_1^2\}} \right] E\{E\{g(y_1) g(y_4) | \mathcal{W}\}\}
\end{aligned} \tag{6.28}$$

e:

$$\begin{aligned}
E\{E\{g(y_1)g(y_4)Y_2Y_2^T|\mathcal{W}\}\} &= E\{E\{g(y_1)g(y_4)|\mathcal{W}\}\} \cdot [E\{Y_2Y_2^T\} \\
&\quad \frac{E\{y_1^2\}E\{E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{y_4^2\}E\{E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}}{E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2} \\
&\quad + \frac{E\{y_1y_4\}[E\{E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2}] \\
&+ \left[\frac{E\{y_4^2\}^2 E\{E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. \frac{E\{y_4^2\}E\{y_1y_4\}[E\{E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. + \frac{E\{y_1y_4\}^2 E\{E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\}}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right] E\{E\{g(y_1)g(y_4)y_1^2|\mathcal{W}\}\} \\
&+ \left[\frac{E\{y_1^2\}^2 E\{E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\}}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. \frac{E\{y_1^2\}E\{y_1y_4\}[E\{E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. + \frac{E\{y_1y_4\}^2 E\{E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right] E\{E\{g(y_1)g(y_4)y_4^2|\mathcal{W}\}\} \\
&+ \left[\frac{E\{y_1^2\}E\{y_4^2\} [E\{E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. - 2 \frac{E\{y_1y_4\}[E\{y_1^2\}E\{E\{y_4Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{y_4^2\}E\{E\{y_1Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right. \\
&\quad \left. + \frac{E\{y_1y_4\}^2 [E\{E\{y_1Y_2|\mathcal{W}\}E\{y_4Y_2^T|\mathcal{W}\}\} + E\{E\{y_4Y_2|\mathcal{W}\}E\{y_1Y_2^T|\mathcal{W}\}\}]}{(E\{y_1^2\}E\{y_4^2\} - E\{y_1y_4\}^2)^2} \right] \\
&E\{E\{g(y_1)g(y_4)y_1y_4|\mathcal{W}\}\}
\end{aligned}$$

Os valores esperados não-lineares de (6.28) e (6.29) são aproximados da seguinte forma:

$$E\{E\{g(y_1)g(y_4)|\mathcal{W}\}\} = \sigma^2 \arcsen\left(\frac{E\{y_1 y_4\}}{(E\{y_1^2\} + \sigma^2)^{1/2} (E\{y_4^2\} + \sigma^2)^{1/2}}\right) \quad (6.30)$$

$$\begin{aligned} E\{E\{g(y_1)g(y_4)y_1^2|\mathcal{W}\}\} &= \sigma^2 E\{y_1^2\} \arcsen\left(\frac{E\{y_1 y_4\}}{(E\{y_1^2\} + \sigma^2)^{1/2} (E\{y_4^2\} + \sigma^2)^{1/2}}\right) \\ &+ \frac{\sigma^2 E\{y_1^2\} E\{y_1 y_4\} (E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2)}{(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2]^{1/2}} \\ &+ \frac{E\{y_1 y_4\} [\sigma^2 E\{y_1^2\} [-E\{y_1 y_4\}^2 + \sigma^2 [E\{y_4^2\} + \sigma^2]] - \sigma^4 E\{y_1 y_4\}^2]}{(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2]^{1/2}} \end{aligned} \quad (6.31)$$

$$\begin{aligned} E\{E\{g(y_1)g(y_4)y_4^2|\mathcal{W}\}\} &= \sigma^2 E\{y_4^2\} \arcsen\left(\frac{E\{y_1 y_4\}}{(E\{y_1^2\} + \sigma^2)^{1/2} (E\{y_4^2\} + \sigma^2)^{1/2}}\right) \\ &+ \frac{\sigma^2 E\{y_4^2\} E\{y_1 y_4\} (E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2)}{(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2]^{1/2}} \\ &+ \frac{E\{y_1 y_4\} [\sigma^2 E\{y_4^2\} [-E\{y_1 y_4\}^2 + \sigma^2 [E\{y_1^2\} + \sigma^2]] - \sigma^4 E\{y_1 y_4\}^2]}{(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2]^{1/2}} \end{aligned} \quad (6.32)$$

c

$$\begin{aligned} E\{E\{g(y_1)g(y_4)y_1 y_4|\mathcal{W}\}\} &= \sigma^2 E\{y_1 y_4\} \arcsen\left(\frac{E\{y_1 y_4\}}{(E\{y_1^2\} + \sigma^2)^{1/2} (E\{y_4^2\} + \sigma^2)^{1/2}}\right) \\ &+ \frac{\sigma^2 E\{y_1^2\} E\{y_4^2\} [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2] + \sigma^6 E\{y_1 y_4\}^2}{(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) [(E\{y_1^2\} + \sigma^2) (E\{y_4^2\} + \sigma^2) - E\{y_1 y_4\}^2]^{1/2}} \end{aligned} \quad (6.33)$$

A partir desses resultados, podemos verificar que os momentos de segunda ordem do vetor de coeficientes são definidos em termos de correlações simples entre variáveis aleatórias gaussianas. Essas correlações são facilmente avaliadas através das relações de (6.11):

$$\left\{ \begin{array}{l}
 E\{y_1^2 | \mathcal{W}\} = \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \text{tr} \{ \mathbf{R}_{p-i} \mathbf{W}(n-p-j) \mathbf{W}^T(n-i-j) \} \\
 E\{y_1 Y_2 | \mathcal{W}\} = \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i} \hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \\
 E\{y_1 y_3 | \mathcal{W}\} = \mathbf{W}^{oT} \sum_{i=0}^{M_1-1} s_{1i} \mathbf{R}_{j+i} \mathbf{W}(n-i-j) \\
 E\{y_1 y_4 | \mathcal{W}\} = \sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \text{tr} \{ \mathbf{R}_{l+q-i-j} \mathbf{W}(n-l-q) \mathbf{W}^T(n-i-j) \} \\
 E\{\mathbf{Y}_2 \mathbf{Y}_2^T | \mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{S}\hat{S}} \\
 E\{y_3 \mathbf{Y}_2 | \mathcal{W}\} = \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \\
 E\{y_4^2 | \mathcal{W}\} = \sum_{l=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{1l} s_{1r} \text{tr} \{ \mathbf{R}_{r-l} \mathbf{W}(n-r-q) \mathbf{W}^T(n-l-q) \} \\
 E\{y_4 \mathbf{Y}_2 | \mathcal{W}\} = \sum_{l=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1l} \hat{s}_k \mathbf{R}_{q+l-k} \mathbf{W}(n-l-q)
 \end{array} \right. \quad (6.34)$$

Entretanto, preferimos organizar as correlações envolvidas da seguinte forma (para facilitar as aproximações que serão necessárias):

$$\left\{ \begin{aligned}
E\{y_1^2 | \mathcal{W}\} &= \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{li} s_{lp} \text{tr} \{ \mathbf{R}_{p-i} \mathbf{W}(n-p-j) \mathbf{W}^T(n-i-j) \} \\
E\{y_1 Y_2 | \mathcal{W}\} \mathbf{W}^T(n) &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{li} \hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \mathbf{W}^T(n) \\
E\{y_1 y_3 | \mathcal{W}\} &= \mathbf{W}^{oT} \sum_{i=0}^{M_1-1} s_{li} \mathbf{R}_{j+i} \mathbf{W}(n-i-j) \\
E\{y_1 y_4 | \mathbf{W}(n)\} &= \sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{li} s_{ll} \text{tr} \{ \mathbf{R}_{l+q-i-j} \mathbf{W}(n-l-q) \mathbf{W}^T(n-i-j) \} \\
E\{Y_2 Y_2^T | \mathcal{W}\} &= \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \\
E\{y_4^2 | \mathcal{W}\} &= \sum_{i=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{li} s_{lr} \text{tr} \{ \mathbf{R}_{r-i} \mathbf{W}(n-r-q) \mathbf{W}^T(n-l-q) \} \\
E\{y_1 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{li} \hat{s}_k s_{li} \hat{s}_p \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \mathbf{W}^T(n-l-j) \mathbf{R}_{j+l-p}^T \\
E\{y_1 Y_2 | \mathcal{W}\} E\{y_3 Y_2^T | \mathcal{W}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{li} \hat{s}_k \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \mathbf{W}^{oT} \mathbf{R}_{\hat{s}}^T \\
E\{y_1 Y_2 | \mathcal{W}\} E\{y_4 Y_2^T | \mathcal{W}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{li} \hat{s}_k s_{li} \hat{s}_p \mathbf{R}_{j+i-k} \mathbf{W}(n-i-j) \mathbf{W}^T(n-l-q) \mathbf{R}_{q+l-p}^T \\
E\{y_3 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\} &= \mathbf{R}_{\hat{s}} \mathbf{W}^o \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{li} \hat{s}_k \mathbf{W}^T(n-i-j) \mathbf{R}_{j+i-k}^T \\
E\{y_4 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\} &= \sum_{l=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{i=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{li} \hat{s}_k s_{li} \hat{s}_p \mathbf{R}_{q+l-k} \mathbf{W}(n-l-q) \mathbf{W}^T(n-i-j) \mathbf{R}_{j+i-p}^T \\
E\{y_4 Y_2 | \mathcal{W}\} E\{y_4 Y_2^T | \mathcal{W}\} &= \sum_{l=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{i=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{li} \hat{s}_k s_{li} \hat{s}_p \mathbf{R}_{q+l-k} \mathbf{W}(n-l-q) \mathbf{W}^T(n-i-q) \mathbf{R}_{q+i-p}^T
\end{aligned} \right. \tag{6.35}$$

Tomando-se o valor esperado de (6.35) e supondo-se que o processo de adaptação ocorre de forma lenta, obtêm-se as seguintes aproximações para os valores esperados necessários:

$$\left\{ \begin{aligned}
E\{y_1^2\} &= \sum_{i=0}^{M_1-1} \sum_{p=0}^{M_1-1} s_{1i} s_{1p} \text{tr}\{\mathbf{R}_{p-i} \mathbf{K}_{p,i}(n-j)\} \\
E\{E\{y_1 Y_2 | \mathcal{W}\} \mathbf{W}^T(n)\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i} \hat{s}_k \mathbf{R}_{j+i-k} \mathbf{K}_{i+j,0}(n) \\
E\{y_1 y_3\} &= \mathbf{W}^{oT} \sum_{i=0}^{M_1-1} s_{1i} \mathbf{R}_{j+i} E\{\mathbf{W}(n-i-j)\} \\
E\{y_1 y_4\} &= \sum_{i=0}^{M_1-1} \sum_{l=0}^{M_1-1} s_{1i} s_{1l} \text{tr}\{\mathbf{R}_{q+l-j-i} \mathbf{K}_{q+l,j+i}(n)\} \\
E\{\mathbf{Y}_2 \mathbf{Y}_2^T\} &= \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \\
E\{y_4^2\} &= \sum_{l=0}^{M_1-1} \sum_{r=0}^{M_1-1} s_{1l} s_{1r} \text{tr}\{\mathbf{R}_{r-l} \mathbf{K}_{r,l}(n-q)\} \\
E\{E\{y_1 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{1i} \hat{s}_k s_{1l} \hat{s}_p \mathbf{R}_{j+i-k} \mathbf{K}_{i,l}(n-j) \mathbf{R}_{j+l-p}^T \\
E\{E\{y_1 Y_2 | \mathcal{W}\} E\{y_3 Y_2^T | \mathcal{W}\}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i} \hat{s}_k \mathbf{R}_{j+i-k} E\{\mathbf{W}(n-i-j)\} \mathbf{W}^{oT} \mathbf{R}_{\hat{s}}^T \\
E\{E\{y_1 Y_2 | \mathcal{W}\} E\{y_4 Y_2^T | \mathcal{W}\}\} &= \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{l=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{1i} \hat{s}_k s_{1l} \hat{s}_p \mathbf{R}_{j+i-k} \mathbf{K}_{j+i,q+l}(n) \mathbf{R}_{q+l-p}^T \\
E\{E\{y_3 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\}\} &= \mathbf{R}_{\hat{s}} \mathbf{W}^o \sum_{i=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} s_{1i} \hat{s}_k E\{\mathbf{W}^T(n-i-j)\} \mathbf{R}_{j+i-k}^T \\
E\{E\{y_4 Y_2 | \mathcal{W}\} E\{y_1 Y_2^T | \mathcal{W}\}\} &= \sum_{l=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{i=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{1l} \hat{s}_k s_{1i} \hat{s}_p \mathbf{R}_{q+l-k} \mathbf{K}_{q+l,j+i}(n) \mathbf{R}_{j+i-p}^T \\
E\{E\{y_4 Y_2 | \mathcal{W}\} E\{y_4 Y_2^T | \mathcal{W}\}\} &= \sum_{l=0}^{M_1-1} \sum_{k=0}^{\hat{M}-1} \sum_{i=0}^{M_1-1} \sum_{p=0}^{\hat{M}-1} s_{1l} \hat{s}_k s_{1i} \hat{s}_p \mathbf{R}_{q+l-k} \mathbf{K}_{l,i}(n-q) \mathbf{R}_{q+i-p}^T
\end{aligned} \right. \tag{6.36}$$

Como resultado final, $\mathbf{K}_{0,0}(n)$ pode ser obtido através da substituição de (6.27), (6.28), (6.29), (6.30), (6.31), (6.32), (6.33) e (6.36) em (6.26).

Note-se que $\mathbf{K}_{0,k}(n)$ e $\mathbf{K}_{0,0}(n)$ são acopladas através da relação em (6.3). Essas três equações permitem a avaliação dos momentos de segunda ordem do algoritmo FXLMS quando o caminho secundário é modelado por um sistema Wiener-Hammerstein.

6.1.1 Sistema Wiener

No caso em que o caminho secundário possa ser modelado por um sistema Wiener, os momentos de segunda ordem podem ser bastante simplificados, diminuindo em muito a carga computacional. Evitando-se os passos intermediários, os resultados finais para os momentos de segunda ordem são dados por

$$\begin{aligned}
\mathbf{K}_{w_{0,0}}(n+1) &= \mathbf{K}_{w_{0,0}}(n) + \mu E\{\mathbf{W}_w(n)\} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T + \mu \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o E\{\mathbf{W}_w^T(n)\} \\
&+ \mu^2 \sigma_z^2 \tilde{\mathbf{R}}_{\hat{s}\hat{s}} + 2\mu^2 \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T \\
&+ \mu^2 \mathbf{W}^{oT} \mathbf{R}_{xx} \mathbf{W}^o \tilde{\mathbf{R}}_{\hat{s}\hat{s}} - \frac{\mu \left[\sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p \hat{s}_q \mathbf{R}_{q-p} \mathbf{K}_{w_{0,p}}^T(n) + \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p \hat{s}_q \mathbf{K}_{w_{0,p}}(n) \mathbf{R}_{q-p}^T \right]}{\sqrt{\frac{1}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + 1}} \\
&+ \frac{2\mu^2 \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} \sum_{k=0}^{M-1} \sum_{l=0}^{\hat{M}-1} \sum_{r=0}^{M-1} s_r s_k \hat{s}_l \hat{s}_q E\{\mathbf{W}_w^T(n-r)\} \mathbf{R}_r \mathbf{W}^o \mathbf{R}_{l-k} \mathbf{K}_{w_{p,k}}^T(n) \mathbf{R}_{p-q}}{\sigma^2 \left(\frac{1}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + 1 \right)^{3/2}} \\
&- \frac{2\mu^2}{\sqrt{\frac{1}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + 1}} \cdot \left\{ \sum_{r=0}^{M-1} s_r E\{\mathbf{W}_w^T(n-r)\} \mathbf{R}_r \mathbf{W}^o \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \right. \\
&+ \left. \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p \hat{s}_q \mathbf{R}_{q-p} E\{\mathbf{W}_w(n-p)\} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{s}}^T + \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p \hat{s}_q \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o E\{\mathbf{W}_w^T(n-p)\} \mathbf{R}_{p-q} \right\} \\
&+ \mu^2 \sigma^2 \arcsen \left(\frac{\sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\}}{\sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + \sigma^2} \right) \tilde{\mathbf{R}}_{\hat{s}\hat{s}} \\
&+ \frac{2\mu^2 \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} \sum_{r=0}^{M-1} \sum_{k=0}^{\hat{M}-1} s_p \hat{s}_q s_r \hat{s}_k \mathbf{R}_{q-p} \mathbf{K}_{w_{p,r}}(n) \mathbf{R}_{r-k}}{\left(\frac{1}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + 1 \right) \sqrt{\frac{2}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w_{q,p}}(n)\} + 1}}
\end{aligned} \tag{6.37}$$

e:

$$\begin{aligned}
\mathbf{K}_{w_0,k}(n+1) &= \mathbf{K}_{w_0,k-1}(n) + \mu \tilde{\mathbf{R}}_{\hat{s}} \mathbf{W}^o E\{\mathbf{W}_w^T(n-k+1)\} \\
&\quad - \mu \frac{\sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p \hat{s}_q \mathbf{R}_{q-p} \mathbf{K}_{w,p,k-1}(n)}{\sqrt{\frac{1}{\sigma^2} \sum_{p=0}^{M-1} \sum_{q=0}^{\hat{M}-1} s_p s_q \text{tr}\{\mathbf{R}_{q-p} \mathbf{K}_{w,p,q}(n)\} + 1}}
\end{aligned} \tag{6.38}$$

As aproximações necessárias para a avaliação de (6.37) e (6.38), através da suposição de um processo de adaptação lento (sem o conhecimento da estatística do vetor de coeficientes adaptativos), são as seguintes:

$$\begin{cases}
\mathbf{W}_w(n-r) \cong E\{\mathbf{W}_w(n-r)\} \\
\mathbf{W}_w(n-a) \mathbf{W}_w^T(n-b) \cong \mathbf{K}_{w,a,b}(n) \\
\mathbf{W}_w^T(n-a) \mathbf{R}_{p-q} \mathbf{W}_w(n-b) \cong \text{tr}\{\mathbf{R}_{p-q} \mathbf{K}_{w,a,b}(n)\} \\
\mathbf{W}_w^T(n-r) \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \mathbf{W}_w(n-a) \mathbf{W}_w^T(n-b) \mathbf{R}_0 \cong E\{\mathbf{W}_w^T(n-r)\} \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \mathbf{K}_{w,a,b}(n) \mathbf{R}_0
\end{cases} \tag{6.39}$$

Para $\sigma^2 \rightarrow \infty$ (caso linear), e expressando-se (6.37) e (6.38) em termos do vetor erro dos coeficientes, essas equações reduzem-se às equações (3.56a) e (3.56b) de TOBIAS (1999) desenvolvidas para FXLMS convencional.

6.1.2 Sistema Sem Memória

No caso de um caminho secundário sem memória, os momentos de segunda ordem podem ser expressos por apenas uma equação. Simplificando-se (6.37) e utilizando-se as seguintes aproximações (suposição de passo de adaptação pequeno),

$$\begin{cases}
\mathbf{W}(n) \cong E\{\mathbf{W}(n)\} \\
\mathbf{W}(n) \mathbf{W}^T(n) \cong \mathbf{K}(n) \\
\mathbf{W}^T(n) \mathbf{R}_0 \mathbf{W}(n) \cong \text{tr}\{\mathbf{R}_0 \mathbf{K}(n)\} \\
\mathbf{W}^T(n) \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \mathbf{W}(n) \mathbf{W}^T(n) \mathbf{R}_0 \cong E\{\mathbf{W}^T(n)\} \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \mathbf{K}(n) \mathbf{R}_0
\end{cases} \tag{6.40}$$

chega-se a

$$\begin{aligned}
\mathbf{K}_{SM}(n+1) = & \mathbf{K}_{SM}(n) + \mu E\{\mathbf{W}_{SM}(n)\} \mathbf{W}^{oT} \mathbf{R}_0 + \mu \mathbf{R}_0 \mathbf{W}^o E\{\mathbf{W}_{SM}^T(n)\} \\
& - \frac{\mu}{\sqrt{\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1}} [\mathbf{R}_0 \mathbf{K}_{SM}(n) + \mathbf{K}_{SM}(n) \mathbf{R}_0] \\
& + \mu^2 \sigma_z^2 \mathbf{R}_0 + 2\mu^2 \mathbf{R}_0 \mathbf{W}^o \mathbf{W}^{oT} \mathbf{R}_0 + \mu^2 \mathbf{W}^{oT} \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \\
& + \frac{2\mu^2 E\{\mathbf{W}_{SM}^T(n)\} \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 \mathbf{K}_{SM}(n) \mathbf{R}_0}{\sigma^2 \left(\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1 \right)^{3/2}} \\
& \frac{2\mu^2 (E\{\mathbf{W}_{SM}^T(n)\} \mathbf{R}_0 \mathbf{W}^o \mathbf{R}_0 + \mathbf{R}_0 E\{\mathbf{W}_{SM}(n)\} \mathbf{W}^{oT} \mathbf{R}_0 + \mathbf{R}_0 \mathbf{W}^o E\{\mathbf{W}_{SM}^T(n)\} \mathbf{R}_0)}{\left(\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1 \right)^{1/2}} \\
& + \mu^2 \sigma^2 \arcsen \left(\frac{\text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\}}{\text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + \sigma^2} \right) \mathbf{R}_0 \\
& + \frac{2\mu^2 \mathbf{R}_0 \mathbf{K}_{SM}(n) \mathbf{R}_0}{\left(\frac{1}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1 \right) \sqrt{\frac{2}{\sigma^2} \text{tr}\{\mathbf{R}_0 \mathbf{K}_{SM}(n)\} + 1}}
\end{aligned} \tag{6.41}$$

No caso em que o sistema puder ser considerado como linear ($\sigma^2 \rightarrow \infty$), então a equação (6.41) reduz-se ao resultado obtido para os momentos de segunda ordem do LMS convencional (HAYKIN, 1991).

7. RESULTADOS

Neste capítulo são apresentados resultados de simulações e procedimentos numéricos que demonstram a validade dos modelos analíticos desenvolvidos nos capítulos anteriores. Os parâmetros utilizados nos exemplos foram escolhidos com o intuito de apresentar, de forma clara e concisa, a influência da não-linearidade no algoritmo.

Várias outras confrontações entre resultados analíticos e de simulação podem ser encontradas nos artigos descritos no ANEXO 10.

7.1 Análise da Superfície de Desempenho

Nesta seção são realizadas simulações e procedimentos numéricos com o intuito de demonstrar a validade e extrair informação dos modelos teóricos obtidos durante a análise da superfície de desempenho (capítulo 3).

7.1.1 Distorções da Superfície de Desempenho

A equação (3.13) representa o modelo analítico para a superfície de desempenho do sistema estudado. A partir de (3.13) podem ser realizadas inferências sobre os efeitos da não-linearidade no sistema adaptativo, ou seja, sobre as distorções que podem influenciar a dinâmica do processo de adaptação e os níveis de cancelamento em regime permanente.

Uma forma de verificar essas distorções é avaliar (3.13) em função dos coeficientes. Visto que esse procedimento resulta em uma figura multidimensional, é comum restringir-se o número de coeficientes a dois, de modo a possibilitar sua visualização, permitindo assim uma estimativa das características de um sistema com maior número de dimensões. No caso de vários coeficientes podem ser utilizadas projeções bidimensionais da superfície de desempenho.

Seja o seguinte sistema (figura 3.1): $\sigma_z^2 = 10^{-4}$; $S_2 = [0,6 \ 0,4 \ -0,2 \ 0,1]^T$; $S_1 = [0,8 \ 0,2 \ 0,1]^T$; $W^o = [1 \ 1]^T$; sinal de excitação correlacionado de potência unitária e dispersão de autovalores da matriz de correlação igual a 3,8.

A figura 7.1 apresenta quatro superfícies de desempenho (mesmos parâmetros mas diferentes σ^2) traçadas em um mesmo gráfico. Note que o efeito da não-linearidade aumenta com a diminuição de σ^2 .

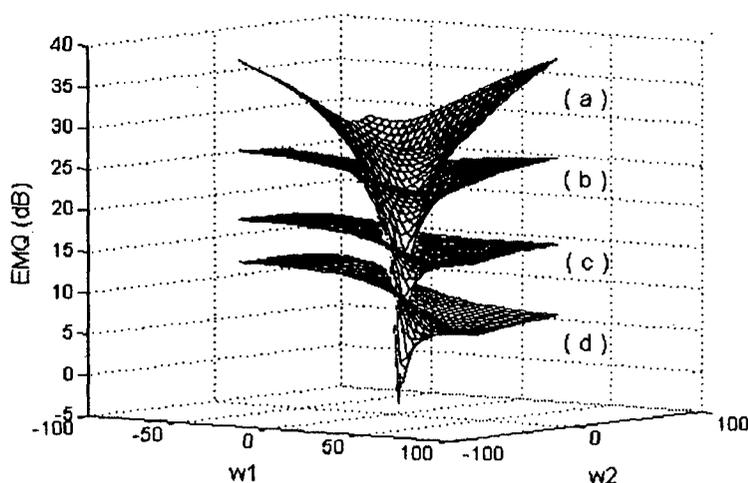


Figura 7.1 – Superfícies de desempenho (equação (3.13)) para: (a) $\sigma^2 = 5 \cdot 10^{10}$; (b) $\sigma^2 = 500$; (c) $\sigma^2 = 50$ e (d) $\sigma^2 = 10$.

Através da figura 7.1 pode ser notado que, à medida que σ^2 diminui, ocorre uma deformação de tal forma que as regiões mais distantes da posição do mínimo sofrem uma maior diminuição de seu gradiente. Essa deformação não ocorre de forma simétrica ao redor do ponto de mínimo (figura 7.1d) mesmo quando se supõe a inexistência de memória no caminho secundário e que o sinal de excitação é não-correlacionado.

De uma forma geral, para coeficientes de controle muito distantes do ponto de mínimo da superfície, à medida que σ^2 diminui, o erro médio quadrático diminui, como pode ser visto nas regiões de paralisia (gradiente reduzido) da figura 7.1. Isso deve-se ao

fato de que para coeficientes muito distantes dos *ótimos*, a potência do sinal $y_f(n)$ (figura 3.1) é muito elevada. Quanto mais não-linear for o sistema (mais baixo o limiar de saturação) mais essa potência será limitada.

Em aplicações de acionamento de equipamentos elétricos de grande potência é comum a utilização de sistemas não-lineares do tipo saturação na saída da malha de atuação. O objetivo é limitar o sinal de controle em casos de instabilidade.

Procedimentos de *otimização* aplicados ao sistema analisado podem ter seu desempenho prejudicado em função de sua inicialização, caso ocorra em uma região de baixo gradiente.

A figura 7.2 apresenta três superfícies de desempenho avaliadas em uma pequena região no entorno de seus pontos de mínimo. Os parâmetros do sistema são os mesmos descritos para a figura 7.1. Os valores de σ^2 foram escolhidos de forma a facilitar a visualização.

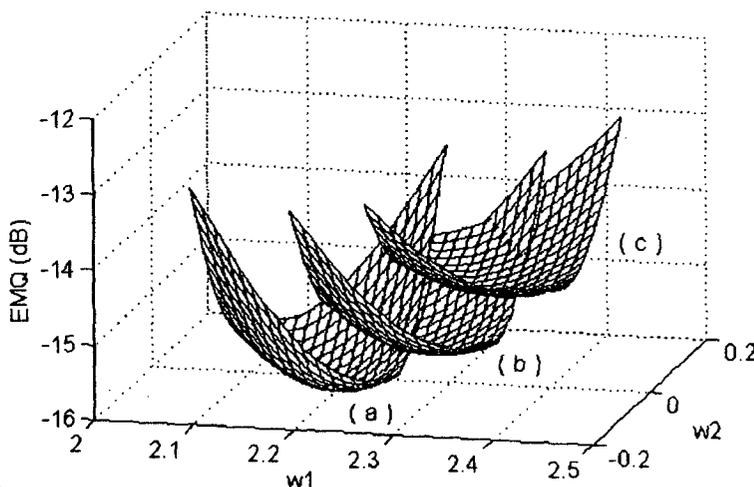


Figura 7.2 – Superfícies de desempenho (equação (3.13)) para: (a) $\sigma^2 = 5 \cdot 10^{10}$; (b) $\sigma^2 = 50$ e (c) $\sigma^2 = 30$.

A partir da figura 7.2 pode-se verificar as alterações sofridas pela superfície de desempenho em regiões próximas de seu mínimo. O aumento da não-linearidade provoca um aumento do valor do mínimo erro médio quadrático, assim como um deslocamento da posição do ponto de mínimo.

Analisando-se ambas as figuras 7.1 e 7.2, pode-se verificar que a superfície de desempenho é uma função *bem comportada*, isto é, aparentemente não apresentando mínimos locais. As alterações sofridas pelo sistema, decorrentes da não-linearidade, ocorrem de maneira suave e contínua, mantendo a característica de unimodalidade do caso linear.

O aumento da característica não-linear do sistema, ou seja, a diminuição do limiar de saturação, acarreta um achatamento assimétrico da superfície de desempenho. Esse achatamento não se restringe apenas às altas potências, mas afeta também o erro médio quadrático (EMQ) no ponto de mínimo. De uma maneira informal, mas bastante elucidativa, poderíamos dizer que a superfície de desempenho é “comprimida tanto por cima como por baixo” e de forma assimétrica.

As observações aqui realizadas são válidas para os três tipos de modelagem utilizadas para o caminho secundário.

No caso em que o número de coeficientes é maior que dois, podemos entender a influência da não-linearidade não apenas como um fator limitador da potência do sinal de controle, mas também como um fator de perturbação do procedimento de otimização utilizado. À medida que essa influência não-linear se torna mais significativa, maiores são as alterações dos níveis possíveis de cancelamento.

7.1.2 Pontos Estacionários

Nesta seção são apresentados exemplos ilustrativos que demonstram a validade dos modelos teóricos em regime permanente para o ponto de mínimo da função objetivo.

Suponha os seguintes parâmetros para o sistema apresentado na figura 3.1:
 $S_2 = [0,5505 \ 0,4954 \ 0,4404 \ 0,3853 \ 0,3305]^T$; $S_1 = [0,9759 \ 0,0976 \ 0,1952]^T$;
 $W^o = [0,6288 \ 0,5031 \ 0,4192 \ 0,3354 \ 0,2515]^T$; sinal de excitação correlacionado
de potência unitária e dispersão de autovalores da matriz de correlação do sinal de
excitação igual a 16,24; $\sigma_z^2 = 10^{-4}$.

Tabela 7.1 – Resultados teóricos segundo as equações (3.41) e (3.60)

σ^2	EMQ (dB)	Coefficientes
10^{10}	-25,8	$[1,158 \ -0,237 \ -0,258 \ 0,051 \ 0,053]^T$
100	-25,7	$[1,164 \ -0,238 \ -0,259 \ 0,051 \ 0,053]^T$
10	-22,6	$[1,218 \ -0,249 \ -0,271 \ 0,054 \ 0,056]^T$
5	-18,5	$[1,272 \ -0,26 \ -0,283 \ 0,056 \ 0,058]^T$
2	-11,6	$[1,387 \ -0,284 \ -0,309 \ 0,061 \ 0,063]^T$
1	-6,7	$[1,482 \ -0,301 \ -0,33 \ 0,065 \ 0,068]^T$
0,1	1,3	$[1,619 \ -0,331 \ -0,360 \ 0,071 \ 0,074]^T$

Tabela 7.2 – Resultados de busca numérica utilizando o método *simplex*

σ^2	EMQ (dB)	Coefficientes
10^{10}	-25,8	$[1,158 \ -0,237 \ -0,258 \ 0,051 \ 0,053]^T$
100	-25,7	$[1,164 \ -0,238 \ -0,259 \ 0,051 \ 0,053]^T$
10	-22,6	$[1,218 \ -0,249 \ -0,267 \ 0,052 \ 0,055]^T$
5	-18,6	$[1,273 \ -0,258 \ -0,27 \ 0,052 \ 0,054]^T$
2	-12	$[1,407 \ -0,268 \ -0,270 \ 0,0767 \ 0,029]^T$
1	-7,4	$[1,537 \ -0,239 \ -0,217 \ 0,061 \ 0,008]^T$
0,1	1	$[2,189 \ -0,13 \ -0,073 \ 0,167 \ -0,056]^T$

Nas tabelas 7.1 e 7.2 encontram-se os valores teóricos (equação (3.41) e (3.60)) e os obtidos através de busca numérica (método *simplex* - Nelder-Mead) para o ponto e valor de mínimo da superfície de desempenho para diferentes valores de σ^2 .

Comparando-se as tabelas 7.1 e 7.2 pode-se verificar que para valores grandes de σ^2 (caso mais comum em sistemas ANC reais), os modelos teóricos possuem uma boa concordância com os valores obtidos através da busca numérica. Note-se que as equações (3.41) e (3.60) são apenas aproximações em decorrência da grande complexidade envolvida em sua obtenção.

Diversas simulações demonstraram que o mínimo valor de σ^2 para o qual os modelos são válidos depende do tipo e da duração das respostas dos sistemas envolvidos no caminho secundário. Entretanto, a estimativa do máximo cancelamento demonstrou ser bastante robusta mesmo para σ^2 muito pequenos.

O procedimento de busca numérica utilizado apresentou problemas de convergência para σ^2 muito pequenos, dependendo do ponto de inicialização. Esse fato pode ser esperado como consequência das regiões de paralisação. Entretanto, para os diversos sistemas analisados, para uma grande variedade de parâmetros, não foram encontrados mínimos locais, e a análise numérica da superfície de desempenho concorda com a análise teórica no sentido da existência de apenas um mínimo global (demonstrada para os casos Wiener e sem memória).

Embora a equação (3.41) indique como consequência da presença da não-linearidade no sistema apenas o escalamento da solução ótima para o caso linear, sabe-se que essa propriedade é apenas uma aproximação para σ^2 pequeno.

Através das tabelas 7.1 ou 7.2 pode-se verificar que, à medida que σ^2 diminui, as modificações dos coeficientes não ocorrem de forma proporcional. Ou seja, a equação (3.33) seria uma aproximação mais adequada (se fosse possível sua solução), em vista de

que a influência da não-linearidade é mais adequadamente modelada como uma matriz de transformação sobre o vetor de coeficientes.

Uma outra forma de se verificar essa propriedade é a análise do caso particular em que se utilizam apenas dois coeficientes. Supondo-se um sistema com os mesmos parâmetros utilizados para a construção das tabelas 7.1 e 7.2, mas com uma dispersão de autovalores igual a 5,29 e $W^o = [0,8944 \quad 0,4472]^T$, pode-se traçar gráficos das posições dos coeficientes ótimos, para diferentes valores de σ^2 (figura 7.3), a partir do procedimento numérico anteriormente descrito.

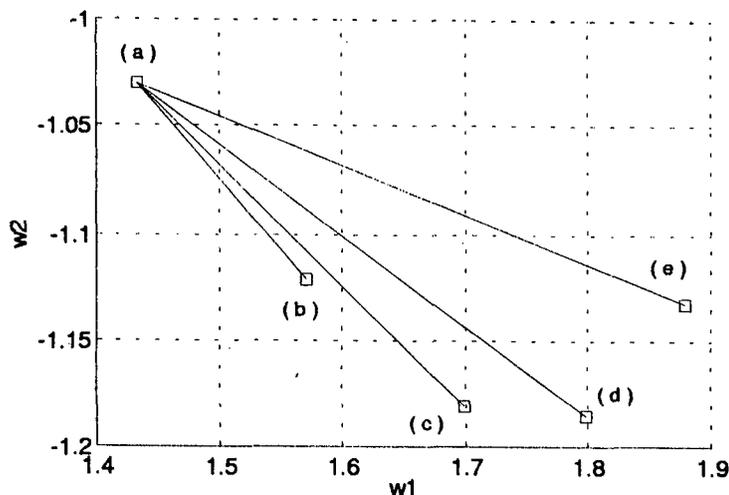


Figura 7.3 – Posição dos pontos de mínimo para: (a) $\sigma^2 = 10^{10}$; (b) $\sigma^2 = 5$; (c) $\sigma^2 = 2$; (d) $\sigma^2 = 1$ e (e) $\sigma^2 = 0,5$. Os segmentos de reta unem o conjunto ótimo de coeficientes para o caso linear ($\sigma^2 \rightarrow \infty$) com os demais pontos de mínimo para os diferentes σ^2 .

Na figura 7.3 são traçados segmentos de reta que unem o conjunto ótimo de coeficientes para o caso linear ($\sigma^2 \rightarrow \infty$) com os demais pontos de mínimo para os diversos σ^2 analisados. A variação angular dessas retas decorre da existência de memória no sistema S_2 . Os eixos horizontal e vertical correspondem ao primeiro e segundo coeficientes, respectivamente. À medida que σ^2 diminui, é alterada a distância entre a posição do caso em questão e o ponto de mínimo para o caso linear. Entretanto, esse

distanciamento não é proporcional para todos os coeficientes. Conforme σ^2 diminui percebe-se claramente uma variação angular dos segmentos de reta apresentados na figura 7.3.

No caso aproximado (σ^2 pequeno, equação (3.41)), o escalamento (idêntico para todos os coeficientes), pode ser visualizado na figura 7.4 em função de β^2 (equação (3.39)).

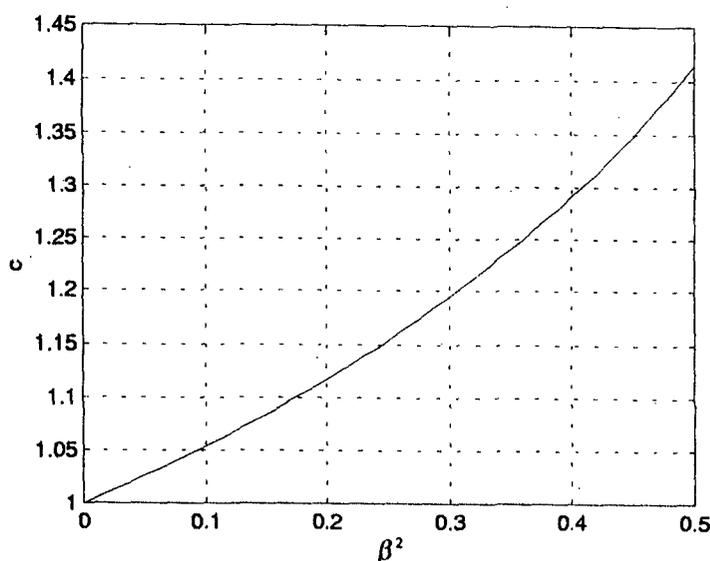


Figura 7.4 – Fator de amplificação dos coeficientes para o sistema Wiener-Hammerstein. Solução aproximada (equação (3.41)).

A equação (3.41) fornece uma aproximação para o ponto de mínimo da superfície de desempenho para valores pequenos de β^2 (σ^2 elevado).

Em sistemas que não possuem memória na saída da não-linearidade (em S_2) é possível determinar uma solução exata para o ponto de mínimo da superfície de desempenho (equações (3.46) e (3.47)). Supondo-se os mesmos parâmetros do exemplo anterior e $S_2 = [1 \ 0 \ \dots \ 0]^T$, pode-se verificar que a posição dos mínimos encontra-se sempre na mesma direção (figura 7.5).

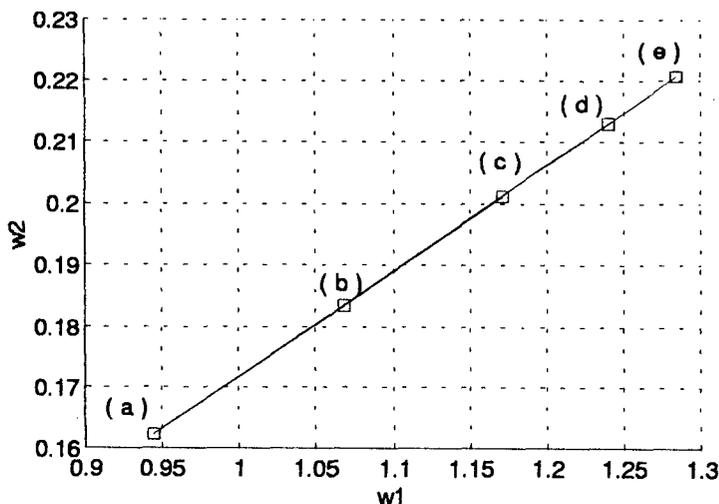


Figura 7.5 - Posição dos pontos de mínimo (equação (3.46)) para: (a) $\sigma^2 = 10^{10}$; (b) $\sigma^2 = 5$; (c) $\sigma^2 = 2$; (d) $\sigma^2 = 1$ e (e) $\sigma^2 = 0,5$.

A figura 7.5 demonstra que a influência da não-linearidade para sistemas que não possuam memória na saída da não-linearidade pode ser caracterizada por um escalamento simples dos coeficientes em função de β_w^2 ou β_{SM}^2 (equações (3.44) ou (3.48)). Esse fator de amplificação, único para todos os coeficientes, pode ser visualizado na figura 7.6.

Na figura 7.6, a linha contínua é o fator de amplificação teórico para o ponto de mínimo da função objetivo, segundo (3.46) e (3.47). Os asteriscos correspondem ao fator de ganho determinado através do procedimento de busca numérica. O sistema analisado possui uma planta W° com 5 coeficientes, S_1 com 10 coeficientes, S_2 não existe, e a dispersão de autovalores da matriz de correlação do sinal de excitação igual a 16,24.

Confrontações entre os resultados do procedimento de busca numérica e modelos teóricos demonstraram excelente concordância, para todo β^2 , na inexistência de memória após a saída da não-linearidade.

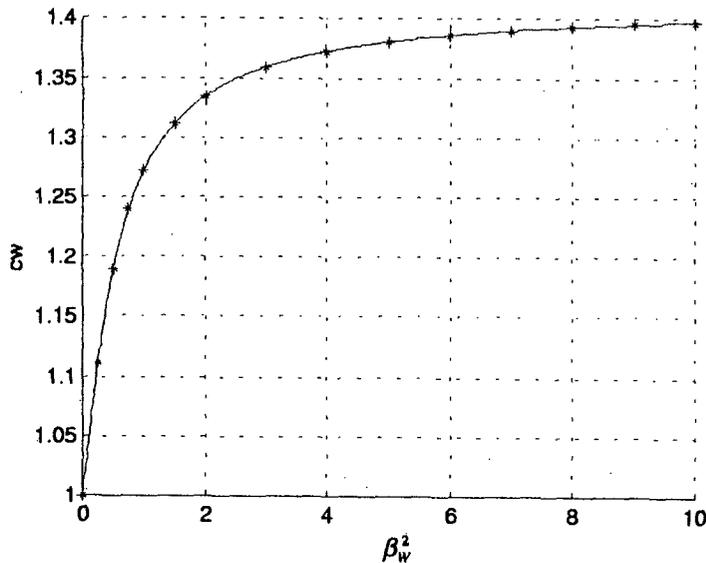


Figura 7.6 – Fator de amplificação dos coeficientes para sistemas sem memória após a não-linearidade. A linha contínua representa o modelo teórico (fator escalar das equações (3.46) e (3.47)), e os asteriscos são resultados obtidos por busca numérica.

7.1.3 Erro em Excesso Devido à Não-Linearidade

O cálculo do erro médio quadrático em excesso relativo ao caso linear (perda no máximo nível possível de cancelamento em decorrência exclusiva da não-linearidade) depende dos parâmetros do sistema. Embora cada situação específica (WH, Wiener e sem memória) possa ser avaliada através das equações pertinentes, a análise de um caso simplificado permite um primeiro *sentimento* acerca do erro em excesso para as diferentes formas de implementação.

A figura 7.7 apresenta o erro médio quadrático em excesso relativo ao caso linear ($EMQE_{LIN} = \xi_{SM MIN} - \sigma_z^2$) para a situação em que o caminho secundário não possui memória, e a potência do sinal desejado ($d(n)$) é unitária (equação (3.62)).

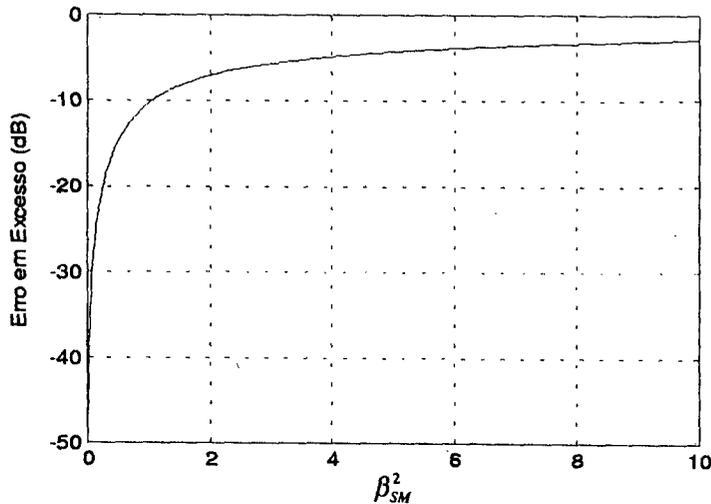


Figura 7.7 – Erro em excesso ($EMQE_{LIN} = \xi_{SM MIN} - \sigma_z^2$) em função do grau de não-linearidade para um sistema sem memória e potência unitária do sinal desejado.

Pode-se notar que uma pequena influência não-linear altera significativamente o nível de cancelamento. Portanto, verifica-se a importância da consideração das características não-lineares do sistema ANC pelo projetista.

7.2 Comportamento em Regime Transitório

Nesta seção são abordadas diversas características e propriedades das configurações estudadas com relação ao comportamento do algoritmo adaptativo durante o regime transitório, ou seja, durante o período de adaptação.

7.2.1 Coeficientes e Erro Médio Quadrático

Seja um sistema ANC cujo caminho secundário é modelado por um sistema Wiener-Hammerstein (figura 4.1), descrito pelos seguintes parâmetros: $\sigma_z^2 = 10^{-4}$; sinal de excitação: ruído correlacionado com variância unitária e dispersão dos autovalores igual a 5,1; $\mathbf{S}_1 = [0,7809 \ 0,6247]^T$; $\mathbf{S}_2 = [0,8111 \ 0,4867 \ 0,3244]^T$; $\mathbf{S}^T \mathbf{S} = 1$; $\hat{\mathbf{S}}^T \hat{\mathbf{S}} = 1$; $\hat{\mathbf{S}} = [0,7692 \ 0,5385 \ 0,3077 \ 0,1538]^T$; $\mu_{max} = 0,1$ (máximo passo de convergência

para o sistema linear determinado de forma empírica); $\mathbf{W}^{ot} \mathbf{W}^o = 1$; $\mathbf{W}^o = [0,7756 \quad 0,5171 \quad -0,3620]^T$; $\eta^2 = \{0,0001 \quad 0,2 \quad 0,5\}$ e $\mu = \{0,001 \quad 0,02\}$.

A influência da não-linearidade ao longo do processo de adaptação pode ser avaliada qualitativamente pelos histogramas da figura 7.8. Nesta, são comparados o comportamento da função não-linear com o histograma de amplitudes do sinal aplicado em sua entrada ($y_f(n)$ na figura 4.1). O valor $\eta^2 = 0,0001$ é considerado como característico de um sistema quase linear, enquanto que $\eta^2 = 0,2$ e $\eta^2 = 0,5$ referem-se a sistemas com elevado grau de não-linearidade. Esses gráficos mostram que os sinais empregados apresentam excursão suficiente para que os efeitos não-lineares afetem o comportamento do algoritmo adaptativo.

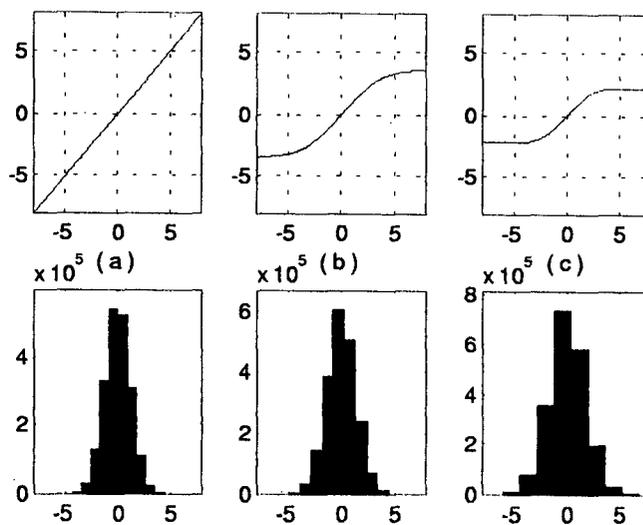


Figura 7.8 – Histogramas do sinal $y_f(n)$ (figura 4.1) e característica não-linear de $g(y_f(n))$ para $\mu = \mu_{\max}/5$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$.

As figuras 7.9 até 7.12 apresentam gráficos do processo de adaptação referentes ao comportamento do erro médio quadrático e coeficientes para três valores do grau de não-linearidade $\eta^2 = \{0,0001 \quad 0,2 \quad 0,5\}$. O passo de convergência é considerado pequeno $\mu = \mu_{\max}/100 = 0,001$ e são realizadas 500 promediações para os resultados simulados. A

linha contínua é o modelo teórico obtido através das equações (4.11), (5.11), (6.7) e (6.26), e a curva espiculada, o resultado da simulação. De modo a facilitar a visualização é traçada 1 a cada 100 amostras.

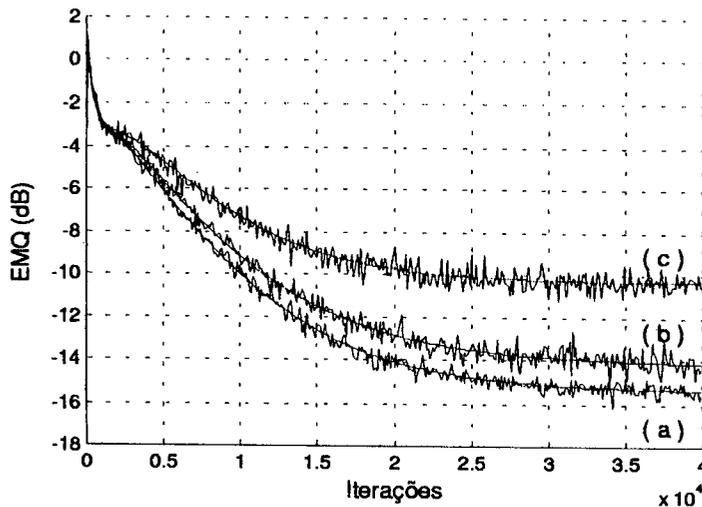


Figura 7.9 – Erro médio quadrático para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (5.11), (6.7) e (6.26). Curva espiculada: simulação (média de 500 realizações).

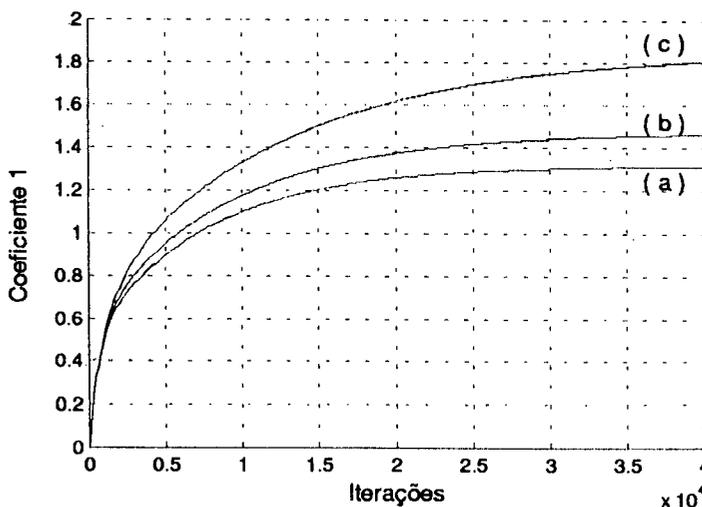


Figura 7.10 – Comportamento médio do primeiro coeficiente para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (6.7) e (6.26). Curva espiculada: simulação (média de 500 realizações).

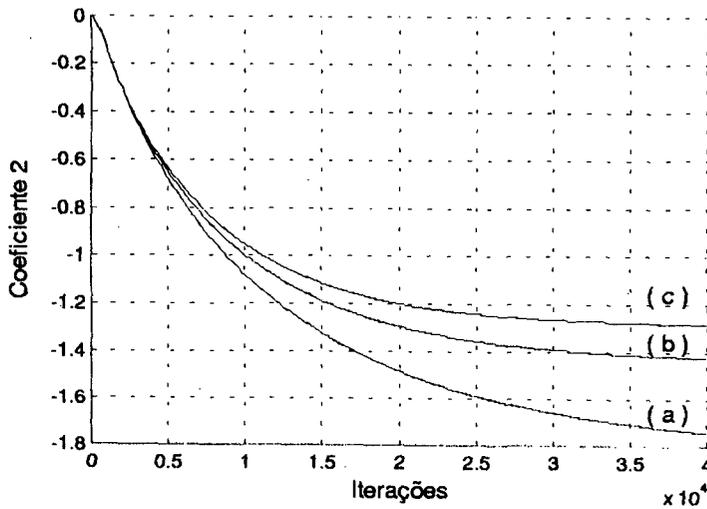


Figura 7.11 – Comportamento médio do segundo coeficiente para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (6.7) e (6.26). Curva espiculada: simulação (média de 500 realizações).

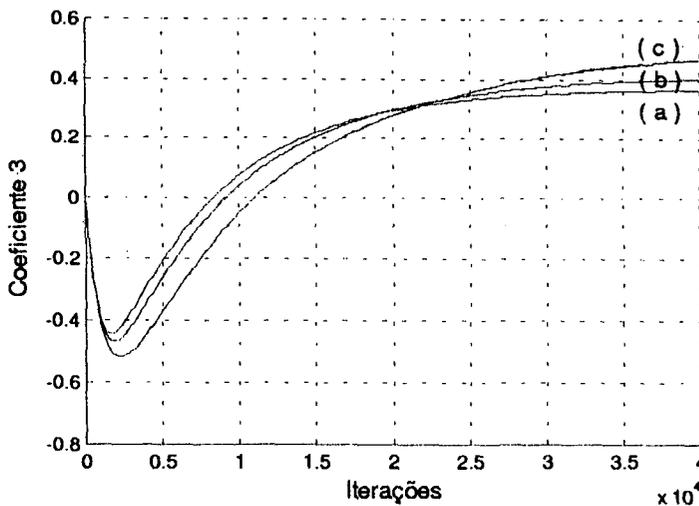


Figura 7.12 – Comportamento médio do terceiro coeficiente para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (6.7) e (6.26). Curva espiculada: simulação (média de 500 realizações).

Através das figuras 7.9 a 7.12 pode-se verificar que o modelo teórico composto pelas equações (4.11), (5.11), (6.7) e (6.26) é capaz de prever adequadamente o comportamento do sistema para diferentes graus de não-linearidade ao longo de todo o período de adaptação. No caso dos coeficientes, o modelo teórico está superposto à simulação.

Como previsto pelo modelo, à medida que o grau de não-linearidade aumenta, aumenta o erro médio quadrático em regime permanente. Da mesma forma, os coeficientes em regime permanente são amplificados.

De forma a verificar a validade dos modelos para o caso de um passo de convergência elevado, utilizaram-se os mesmos parâmetros descritos no teste anterior, porém com $\mu = \mu_{\max}/5 = 0,02$ e 5000 promediações. Os resultados são apresentados nas figuras 7.13 a 7.16, onde é traçada uma a cada 10 amostras.

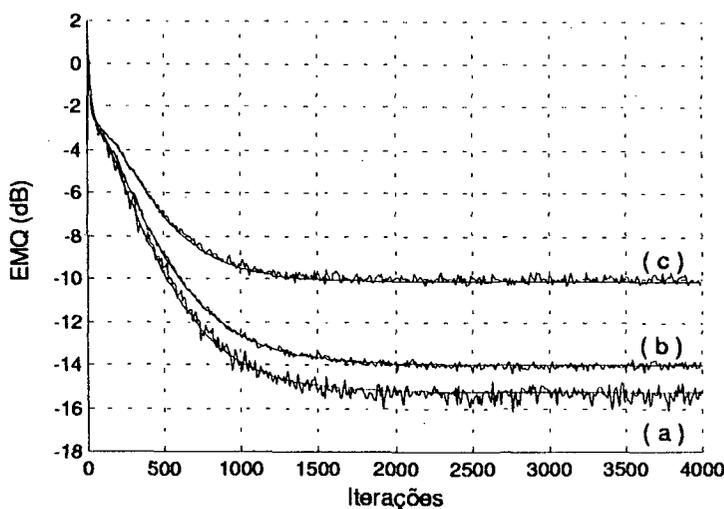


Figura 7.13 – Erro médio quadrático para $\mu = \mu_{\max}/5$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (5.11), (6.7) e (6.26). Curva espiculada: simulação (média de 5000 realizações). É desenhada uma a cada 10 amostras.

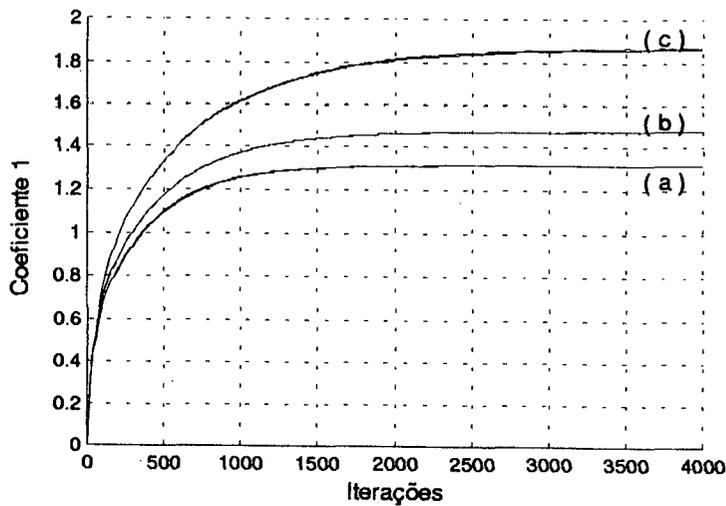


Figura 7.14 – Comportamento médio do primeiro coeficiente para $\mu = \mu_{\max}/5$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (6.7) e (6.26). Curva espiculada: simulação (média de 5000 realizações). É desenhada uma a cada 10 amostras.

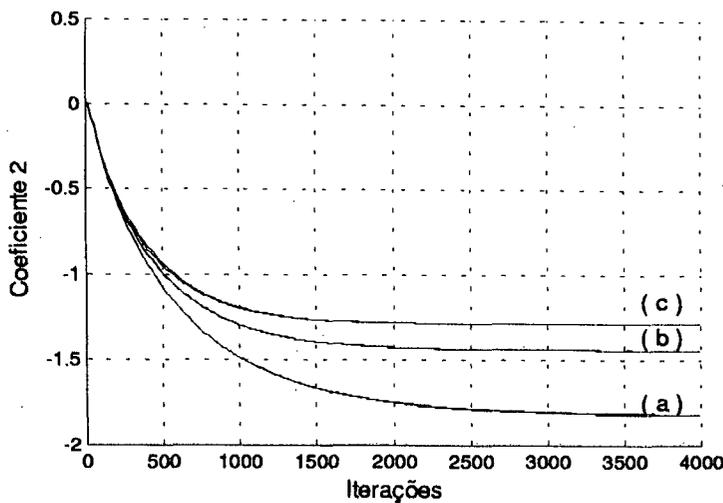


Figura 7.15 – Comportamento médio do segundo coeficiente para $\mu = \mu_{\max}/5$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (6.7) e (6.26). Curva espiculada: simulação (média de 5000 realizações). É desenhada uma a cada 10 amostras.

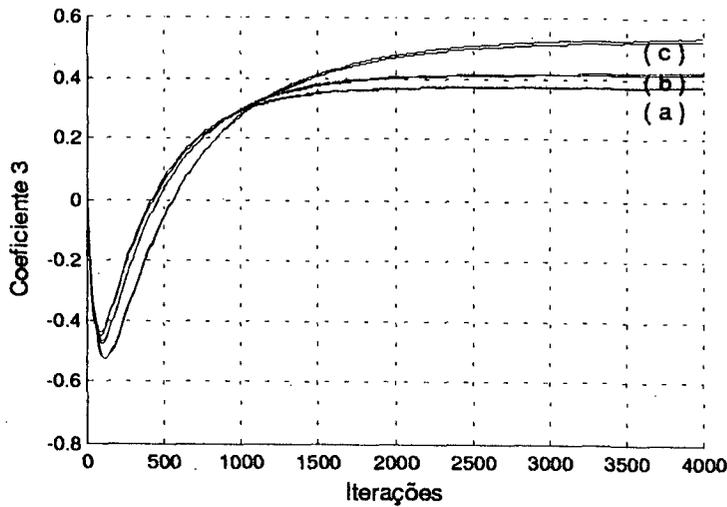


Figura 7.16 – Comportamento médio do terceiro coeficiente para $\mu = \mu_{\max}/5$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,2$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico obtido através das equações (4.11), (5.11), (6.7) e (6.26). Curva espiculada: simulação (média de 5000 realizações). É desenhada uma a cada 10 amostras.

As figuras 7.13 até 7.16 demonstram que, embora os modelos analíticos compostos pelas equações (4.11), (6.7) e (6.26) tenham sido construídos através da suposição de que o passo de convergência deva ser suficientemente pequeno (de modo que o processo de adaptação ocorra de maneira lenta), neste exemplo, essa restrição não compromete os resultados para μ elevado.

7.2.2 Efeitos dos Momentos de Segunda Ordem

Em decorrência da elevada carga computacional necessária para a utilização dos modelos de segunda ordem pelo modelo analítico, é de grande interesse verificar seu impacto sobre os resultados obtidos. Na figura 7.17 é apresentado um recorte da figura 7.13, onde também é traçado o modelo teórico utilizando a seguinte aproximação para os momentos de segunda ordem: $E\{\mathbf{W}(n-k)\mathbf{W}^T(n-l)\} \cong E\{\mathbf{W}(n-k)\} \cdot E\{\mathbf{W}^T(n-l)\}$. Ou seja, aproximam-se os momentos de segunda ordem pelo produto dos momentos de primeira ordem (modelo simplificado). Nessa figura é desenhada uma a cada 10 amostras.

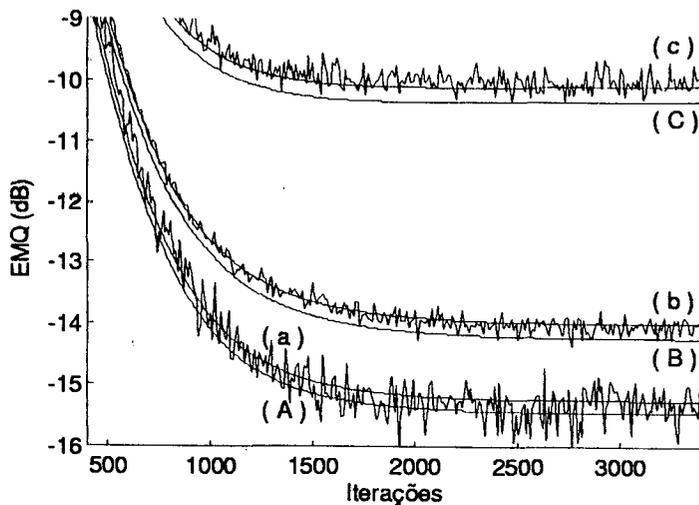


Figura 7.17 – Erro médio quadrático para: (a), (b) e (c) modelo completo (equações (4.11), (5.11), (6.7) e (6.26)) e (A), (B) e (C) modelo simplificado (equações (4.11), (5.11) e $E\{\mathbf{W}(n-k)\mathbf{W}^T(n-l)\} \equiv E\{\mathbf{W}(n-k)\} \cdot E\{\mathbf{W}^T(n-l)\}$) para $\mu = \mu_{\max}/5$. (a) e (A) $\eta^2 = 0,0001$; (b) e (B) $\eta^2 = 0,2$ e (c) e (C) $\eta^2 = 0,5$. As linhas espiculadas representam os resultados de simulação.

Na figura 7.17, o modelo simplificado encontra-se um pouco abaixo do modelo completo para todos os graus de não-linearidade testados. Pode-se concluir que, nesse caso em especial, a influência dos momentos de segunda ordem é pouco significativa. Entretanto, isso nem sempre ocorre.

Como contra-exemplo à baixa significância da influência dos momentos de segunda ordem no comportamento transitório do algoritmo, é apresentado o seguinte exemplo.

Suponha-se o seguinte sistema: $\hat{\mathbf{S}} = \mathbf{S}_1 = \mathbf{S}_2 = [1]$; $\sigma_z^2 = 10^{-6}$; $\mu = 0,05$; sinal de excitação gaussiano correlacionado de média zero e com dispersão de autovalores de $\chi = 18,3$; $\eta^2 = 0,0001$; 200 realizações e 13 coeficientes para \mathbf{W}^o .

A figura 7.18 apresenta o EMQ simulado (linha espiculada) e os modelos teórico completo (equações (4.16), (5.14) e (6.41)) e teórico simplificado (equações (4.16), (5.14) e $E\{\mathbf{W}(n)\mathbf{W}^T(n)} \cong E\{\mathbf{W}(n)} \cdot E\{\mathbf{W}^T(n)\}$).

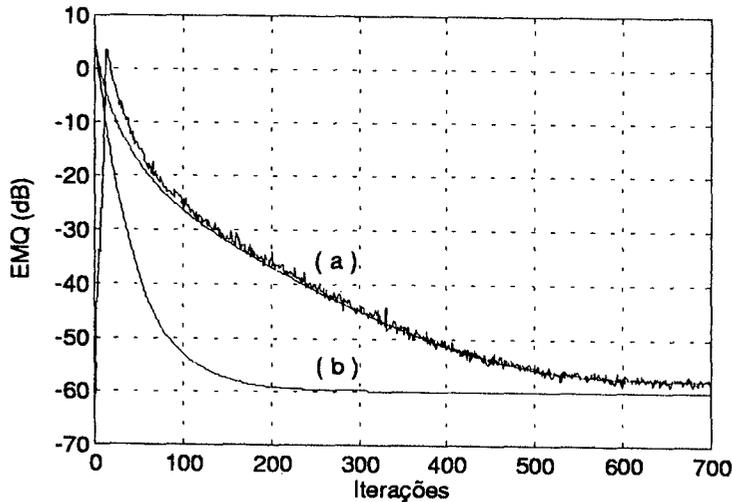


Figura 7.18 – Erro médio quadrático para $\eta^2 = 0,0001$. (a) modelo completo (equações (4.16), (5.14) e (6.41)) e (b) modelo simplificado (equações (4.16), (5.14) e $E\{\mathbf{W}(n)\mathbf{W}^T(n)} \cong E\{\mathbf{W}(n)} \cdot E\{\mathbf{W}^T(n)\}$).

Através desse resultado verifica-se que em determinadas aplicações os efeitos dos momentos de segunda ordem são muito significativos para a descrição do comportamento transitório do algoritmo, justificando, assim, sua utilização.

7.2.3 Efeitos do Passo de Convergência

Seja o seguinte sistema: $\sigma_z^2 = 10^{-6}$; excitação ruído branco gaussiano com variância $\sigma_x^2 = 1$; $\mathbf{S}_1 = [0,9356 \ 0,2807 \ 0,1871 \ 0,0936 \ 0,0468]^T$; $\mathbf{S}_2 = [1]^T$; $\mathbf{S}_1^T \mathbf{S}_1 = 1$; $\hat{\mathbf{S}}^T \hat{\mathbf{S}} = 1$; $\mathbf{W}^{oT} \mathbf{W}^o = 1$; $\hat{\mathbf{S}} = [0,8922 \ 0,3965 \ 0,1487 \ 0,1487 \ 0,0496]^T$; $\mu_{\max} = 0,05$ (máximo passo de convergência para o FXLMS convencional determinado de forma empírica); $\boldsymbol{\mu} = \{\mu_{\max}/10 \ \mu_{\max}/100\}$; $\boldsymbol{\eta}^2 = \{0,0001 \ 0,1 \ 0,5\}$ e 15 coeficientes para \mathbf{W}^o .

As figuras 7.19 e 7.20 referem-se ao erro médio quadrático e ao comportamento médio do terceiro coeficiente segundo as equações (4.13), (5.13), (6.37) e (6.38) para o sistema descrito com $\mu = \mu_{\max}/10$. Os demais coeficientes apresentam comportamento semelhante.

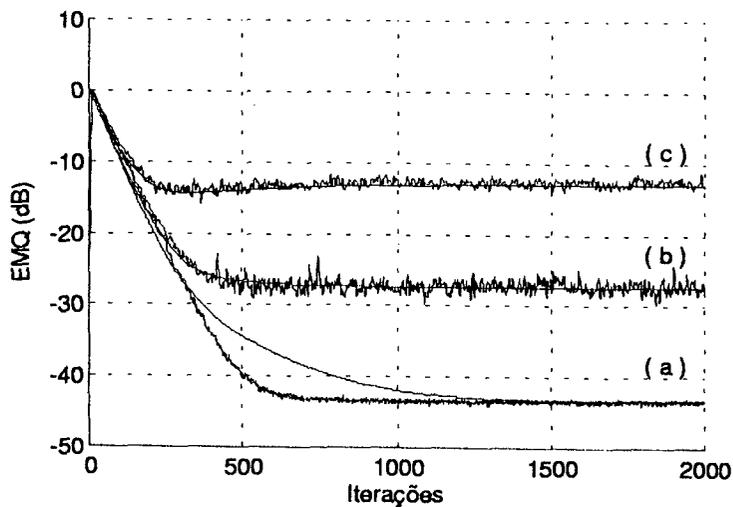


Figura 7.19 – Erro médio quadrático para $\mu = \mu_{\max}/10$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,1$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.13), (5.13), (6.37) e (6.38). Curva espiculada: simulação (média de 1000 realizações).

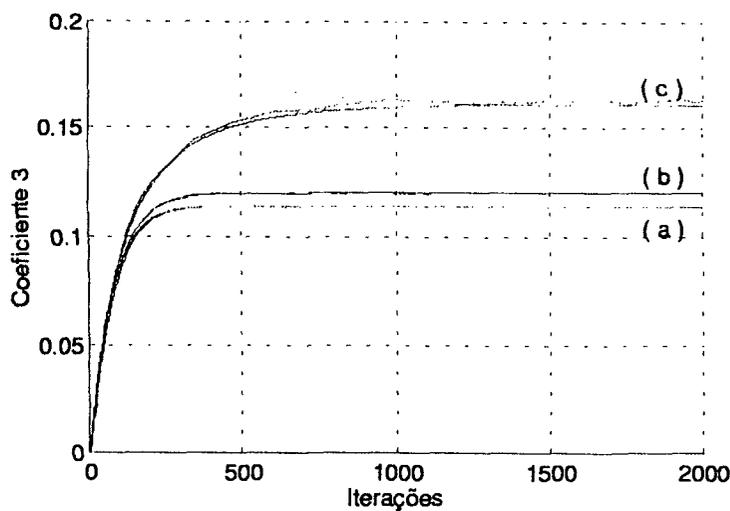


Figura 7.20 – Comportamento médio do primeiro coeficiente para $\mu = \mu_{\max}/10$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,1$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.13), (6.37) e (6.38). Curva espiculada: simulação (média de 1000 realizações).

Na figura 7.19 pode-se observar que para $\eta^2 = 0,0001$ o modelo teórico utilizando as equações (4.13), (5.13), (6.37) e (6.38) não consegue prever corretamente o comportamento do sistema em questão. Entretanto, o comportamento médio dos coeficientes possui uma adequada concordância.

O desajuste apresentado pode ser compreendido através das figuras 7.21 e 7.22, onde é utilizado o mesmo conjunto de parâmetros das figuras 7.19 e 7.20, porém com um passo de convergência dez vezes menor. Para melhor visualização é desenhada uma a cada 50 amostras.

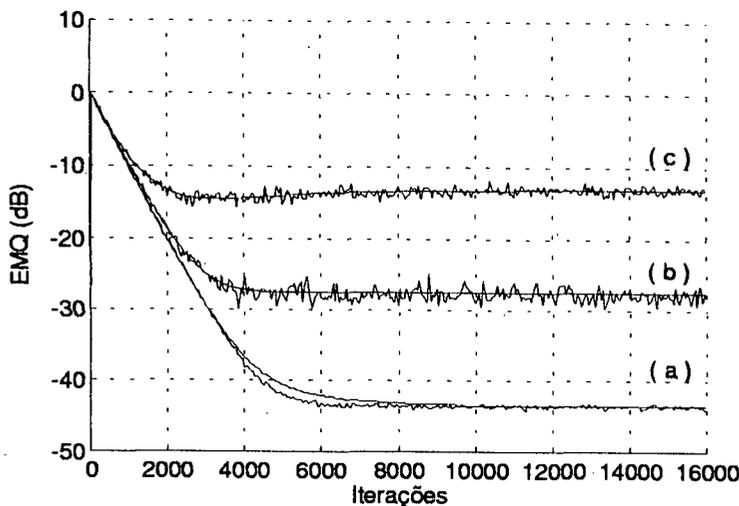


Figura 7.21 – Erro médio quadrático para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,1$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.13), (5.13), (6.37) e (6.38). Curva espiculada: simulação (média de 1000 realizações).

A partir dos resultados obtidos, verificamos que o modelo teórico composto pelas equações (4.13), (5.13), (6.37) e (6.38) reproduz tanto melhor o comportamento do sistema quanto menor for o passo de convergência. Esse fato não constitui uma grande limitação, em vista de que, em aplicações práticas de passo fixo, o projetista, interessado no máximo cancelamento possível, geralmente faz uso de passos pequenos.

Diversas simulações demonstram que os modelos desenvolvidos são bastante robustos com relação à escolha do passo de convergência. Entretanto, essa característica depende dos parâmetros envolvidos.

Os resultados obtidos indicam que, quanto maior for o grau de não-linearidade do sistema, mais robusto será o modelo teórico, para passos de convergência elevados.

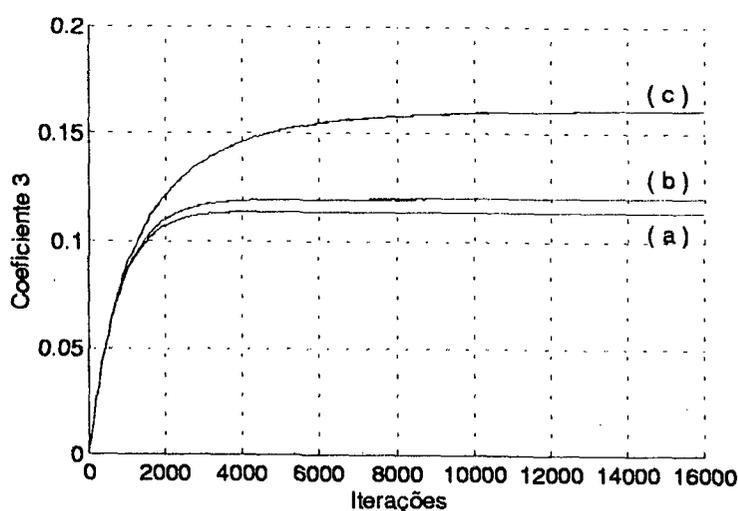


Figura 7.22 – Comportamento médio do primeiro coeficiente para $\mu = \mu_{\max}/100$. (a) $\eta^2 = 0,0001$; (b) $\eta^2 = 0,1$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.13), (6.37) e (6.38). Curva espiculada: simulação (média de 1000 realizações).

7.2.4 Efeitos do Grau de Correlação e da Duração da Resposta da Planta

Imagine-se o seguinte sistema: $\sigma_z^2 = 10^{-6}$; excitação ruído gaussiano correlacionado com variância unitária, média zero e dispersão de autovalores de 32,22; $\hat{S} = S_1 = S_2 = [1]^T$; $W^{oT} W^o = 1$; $\mu_{\max} = 0,033$ (máximo passo de convergência para o FXLMS convencional determinado de forma empírica) e 30 coeficientes para W^o (de $w_0^o = 0,252$ até $w_{29}^o = 0,103$ em passos de $-0,005$) e $\mu = \mu_{\max}/2$.

As figuras 7.23 e 7.24 apresentam confrontações entre modelos analíticos (equações (4.16), (5.14) e (6.41)) e simulações para o erro médio quadrático e comportamento médio do décimo coeficiente para diferentes graus de não-linearidade. É desenhada uma a cada cinco amostras. As curvas simuladas foram adiantadas em 30 amostras para compensar o tempo necessário para o preenchimento do vetor de dados, visto que o modelo teórico pressupõe uma distribuição gaussiana para o vetor $\mathbf{X}(n)$ desde $n = 0$.

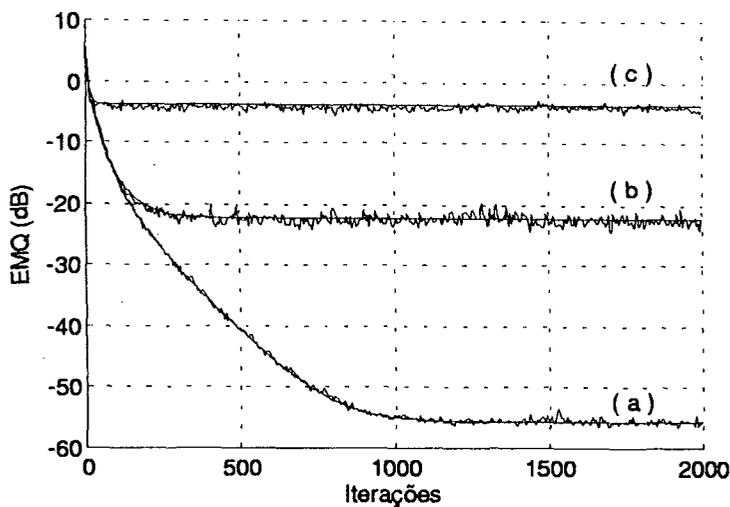


Figura 7.23 – Erro médio quadrático para $\mu = \mu_{\max}/2$. (a) $\eta^2 = 0,0005$; (b) $\eta^2 = 0,05$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.16), (5.14) e (6.41). Curva espiculada: simulação (média de 500 realizações). As curvas simuladas foram adiantadas em 30 amostras para compensar o tempo necessário para o preenchimento do vetor de dados.

Esse exemplo demonstra a capacidade do modelo teórico em prever o comportamento do filtro adaptativo em condições extremamente desfavoráveis: passo de convergência elevado, vários atrasos no caminho primário e alta correlação do sinal de excitação.

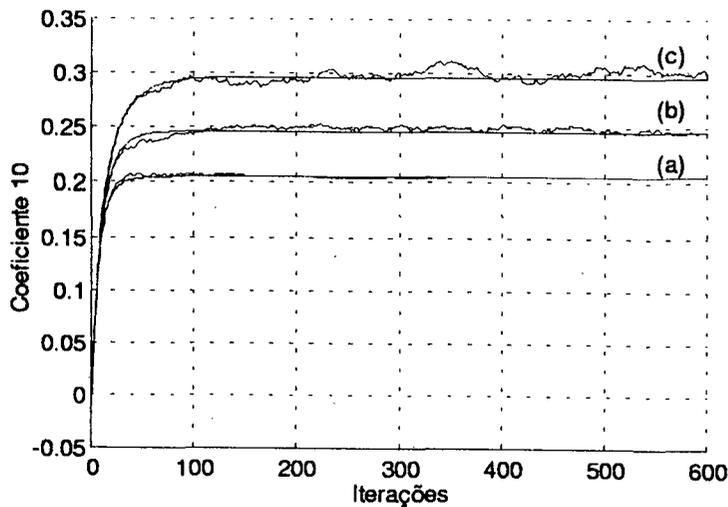


Figura 7.24 – Comportamento médio do décimo coeficiente para $\mu = \mu_{\max}/2$. (a) $\eta^2 = 0,0005$; (b) $\eta^2 = 0,3$ e (c) $\eta^2 = 0,5$. Linha contínua: modelo teórico segundo as equações (4.16) e (6.41). Curva espiculada: simulação (média de 500 realizações). As curvas simuladas foram adiantadas em 30 amostras para compensar o tempo necessário para o preenchimento do vetor de dados.

7.3 Comportamento em Regime Permanente

Nos exemplos apresentados nesta seção, supõe-se a perfeita estimação do caminho secundário. Em decorrência disso obtém-se a seguinte simplificação: $\eta^2 = \beta^2$. Esse caso é especialmente importante, pois prevê o melhor desempenho possível do algoritmo.

A análise da influência de erros na estimativa do caminho secundário pode ser realizada de forma independente do grau de não-linearidade e é encontrada em TOBIAS (1999). Os diversos exemplos apresentados nas seções anteriores também ilustram a qualidade dos novos modelos no caso de estimação imperfeita.

Diversas simulações demonstram que os resultados obtidos a partir de casos mais simples, como para o caso particular do caminho secundário sem memória, possibilitam o *sentimento* pelo projetista das características aqui analisadas, sem perda de generalidade.

7.3.1 Polarização dos Coeficientes

Uma figura de mérito adequada para medir a influência da não-linearidade sobre a identificação dos parâmetros *ótimos* do problema de minimização do erro médio quadrático é a polarização dos coeficientes. Essa grandeza relaciona os coeficientes em regime permanente obtidos ao final do processo de adaptação com a posição do mínimo da superfície de desempenho ($POL_i = w_{RPi} / \tilde{w}_i$).

Como demonstrado através dos modelos analíticos, a influência da não-linearidade sobre os coeficientes adaptativos, assim como sobre a posição do mínimo, é apenas o de um escalar multiplicando a solução do caso linear. Portanto, podemos avaliar a polarização em termos de um único parâmetro que irá influenciar todos os coeficientes.

No caso do sistema cujo caminho secundário é modelado por um sistema Wiener-Hammerstein, a posição teórica do mínimo (equação (3.41)) coincide com o regime permanente do filtro adaptativo quando β^2 ($\hat{S} = S_1 * S_2$) é suficientemente pequeno (equação (4.24)). Nesse caso não há polarização dos coeficientes. À medida que β^2 aumenta, este resultado não é mais confiável.

Entretanto, na inexistência de memória na saída da função não-linear, os resultados teóricos são válidos para todo β_w^2 , e a polarização (multiplicativa) dos coeficientes ($POL = k_w / c_w$) é obtida a partir de (3.46) e (4.25) para $\hat{S} = S$ e $\eta_w^2 = \beta_w^2$:

$$POL = \frac{k_w}{c_w} = \left(\frac{2\beta_w^2}{(1 - \beta_w^2)(1 - 4\beta_w^4)^{1/2} - 2\beta_w^4 + 3\beta_w^2 - 1} \right)^{1/2} \quad (7.1)$$

A figura 7.25 representa a polarização dos coeficientes para esse caso.

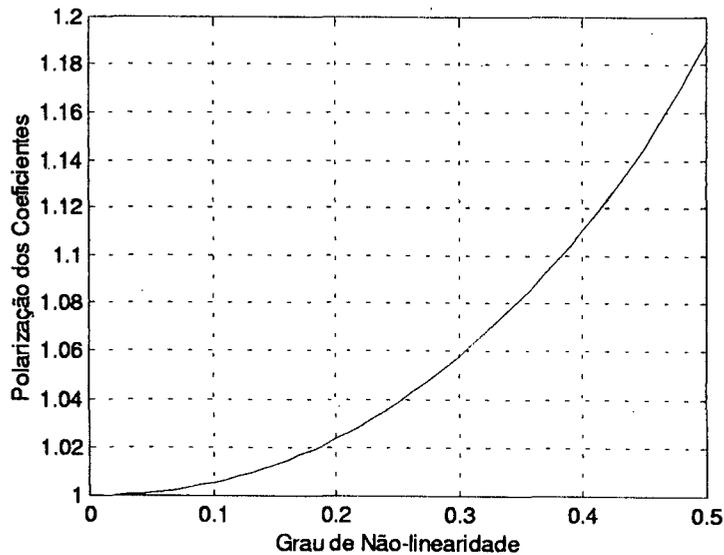


Figura 7.25 – Polarização multiplicativa dos coeficientes $w_{wRPi}/\tilde{w}_{wi}, \forall i, \beta^2$ supondo a inexistência de memória após a não-linearidade (ausência de S_2) e a correta estimativa do caminho secundário ($\hat{S} = S$).

Noté-se que à medida que $\beta_w^2 \rightarrow 1$ os coeficientes crescem sem limite. Para $\beta_w^2 \geq 1$ não existem pontos estacionários, entretanto o vetor de coeficientes mantém-se na direção da solução ótima (relação de proporcionalidade entre os coeficientes).

Esses resultados são válidos também para o sistema sem memória.

7.3.2 Erro Médio Quadrático em Excesso

O erro médio quadrático em excesso (EMQE) do algoritmo LMS corresponde à degradação do máximo cancelamento possível, quando o algoritmo LMS é utilizado para adaptar o vetor de coeficientes. Ou seja, o EMQE é dado pela diferença entre o EMQ em regime permanente e o mínimo EMQ da superfície de desempenho para um determinado grau de não-linearidade.

A figura 7.26 apresenta o gráfico do EMQE do algoritmo LMS para diferentes graus de não-linearidade, quando o caminho secundário não possui memória (equações (3.62) e (5.20)).

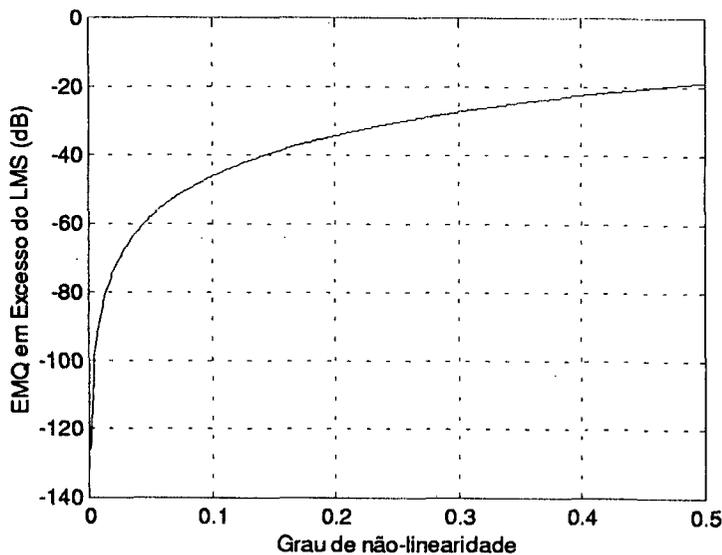


Figura 7.26 – Erro médio quadrático em excesso do LMS ($EMQE_{LMS} = \xi_{SM RP} - \xi_{SM MIN}$) em função do grau de não-linearidade (equações (3.62) e (5.20)).

Verifica-se que mesmo sistemas com grau de não-linearidade reduzido sofrem uma significativa degradação de desempenho.

7.3.3 Desajuste em Regime Permanente

O desajuste pode ser obtido normalizando-se o EMQE pelo valor do mínimo erro médio quadrático da superfície de desempenho ($M = EMQE/\xi_{MIN}$).

Na figura 7.27 encontra-se o gráfico do desajuste quando o caminho secundário não possui memória (equações (3.62) e (5.20)), supondo-se potência unitária do sinal desejado.

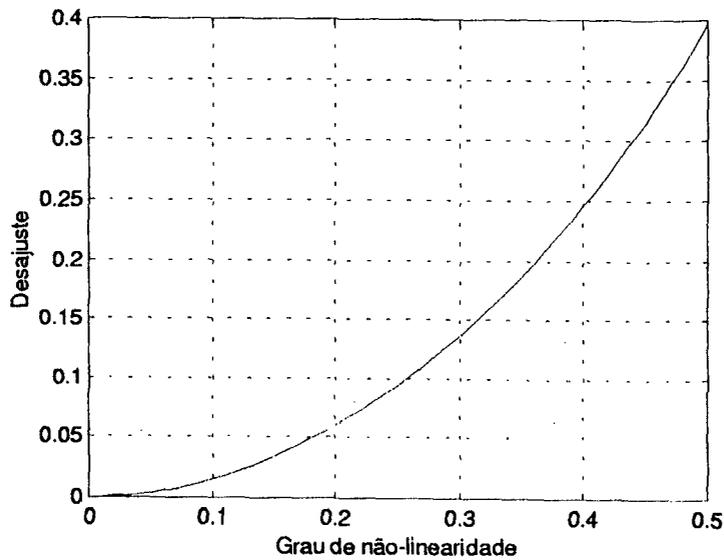


Figura 7.27 – Desajuste do erro médio quadrático ($M = EMQE_{LMS} / \xi_{SM MIN}$) em função do grau de não-linearidade (equações (3.62) e (5.20)).

7.3.4 Estabilidade

Para todas as topologias estudadas, a estabilidade do algoritmo adaptativo é regida por um novo critério de estabilidade em função do grau de não-linearidade: $\eta^2 < 1$.

Essa nova condição não invalida a dependência do algoritmo com o passo de convergência, mas vem a constituir-se em mais uma restrição.

Na figura 7.28 são apresentadas trajetórias dos coeficientes (linhas espiculadas) sobre curvas de nível (equação (3.18)) representando a superfície de desempenho (linhas contínuas). A partir dessas pode-se verificar que para $\eta^2 > 1$, embora ainda exista um mínimo para a superfície de desempenho, o algoritmo adaptativo não converge.

Os parâmetros de simulação utilizados na figura 7.28 foram: $\sigma_z^2 = 10^{-6}$; $\mathbf{W}^o = [0,707 \ 0,707]^T$; $\hat{\mathbf{S}} = \mathbf{S}_1 = \mathbf{S}_2 = [1]$, e sinal de excitação gaussiano correlacionado com dispersão de autovalores igual a 24.

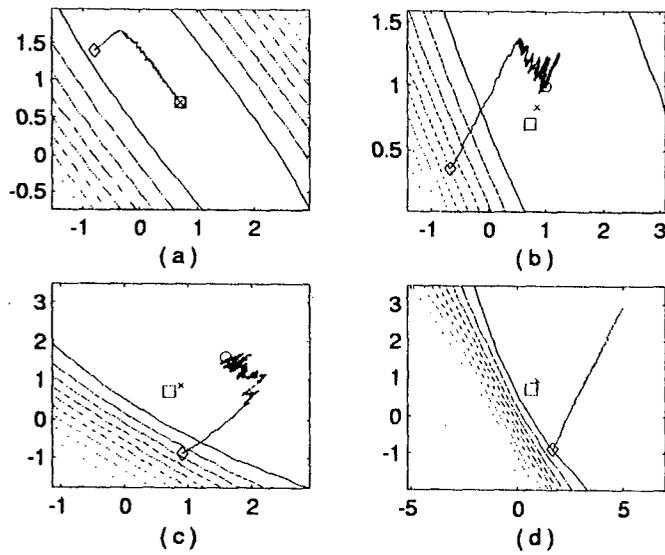


Figura 7.28 – Contornos da superfície de desempenho e trajetória dos coeficientes. \square : W^0 , \circ : W_{RP} , \diamond : inicialização dos coeficientes. Linhas espiculadas: trajetória dos coeficientes. Linhas contínuas: curvas de nível da superfície de desempenho (equação (3.18)). (a) $\eta^2 = 10^{-5}$; (b) $\eta^2 = 0,5$; (c) $\eta^2 = 0,8$ e (d) $\eta^2 = 2$

8. DISCUSSÃO E CONCLUSÕES

Neste trabalho foi realizada a análise estatística do algoritmo LMS Filtrado (FXLMS) quando o caminho secundário (sistema que conecta a saída do filtro adaptativo ao ponto de cancelamento) possui propriedades não-lineares. A estrutura mais geral analisada (e da qual se derivam as demais estruturas como casos particulares) corresponde ao caso em que o caminho secundário é modelado por um sistema Wiener-Hammerstein (WH). Esse sistema é composto por dois modelos dinâmicos lineares interpostos por uma função estática não-linear. A função não-linear em questão modela efeitos de saturação. A topologia estudada é representativa de uma série de aplicações do algoritmo FXLMS. Em especial, em aplicações de controle ativo de ruído acústico e vibrações, onde a função não-linear representa os efeitos de saturação dos amplificadores de potência e transdutores associados (alto-falantes e atuadores piezoelétricos). Os sistemas lineares associados representam os processos de filtragem decorrentes da propagação do sinal no meio acústico-estrutural e na parte elétrica do sistema de processamento.

Primeiramente foram estudadas as propriedades da superfície de desempenho do sistema proposto. Este estudo resultou na determinação de modelos matemáticos do erro médio quadrático em função do vetor de coeficientes de controle. A partir do modelo teórico foi demonstrado que o aumento da não-linearidade do sistema acarreta distorções significativas da superfície de desempenho. Essas distorções incluem o aparecimento de regiões de baixo gradiente que podem ocasionar *paralisia* em algoritmos de otimização.

Foram determinadas expressões analíticas para os pontos estacionários das superfícies de desempenho relacionadas às diversas modelagens do caminho secundário. Com exceção de uma configuração, foi possível determinar a solução exata para o ponto de gradiente nulo dessas superfícies. Nesses casos também foi possível demonstrar que esse ponto corresponde ao mínimo da superfície do erro médio quadrático. Para a situação particular em que inexitem operações de filtragem na saída da função não-linear, foi demonstrado que o ponto de mínimo da superfície de desempenho corresponde a uma

versão escalada do ponto de mínimo do caso linear. Como consequência, a não-linearidade causa uma polarização do vetor ótimo de coeficientes.

Para o caso geral (sistema Wiener-Hammerstein), apenas uma solução aproximada foi encontrada devido à complexidade matemática do problema de determinação dos pontos estacionários. Entretanto, essa solução aproximada modela adequadamente o comportamento do sistema para graus de não-linearidade reduzidos, o caso mais comum nas aplicações práticas visadas.

Utilizando-se os resultados anteriores foi possível determinar o máximo cancelamento possível para o problema de estimação média quadrática não-linear em questão. Os resultados demonstram que o aumento da não-linearidade acarreta uma diminuição do máximo cancelamento possível. Essa diminuição é quantificada, sendo expressiva mesmo em sistemas com reduzido grau de não-linearidade. Em decorrência desses resultados, verifica-se a importância de uma modelagem não-linear mesmo para o estudo de sistemas de cancelamento com características quase lineares.

Em um segundo momento, supondo-se a utilização do algoritmo FXLMS para a determinação do vetor de coeficientes de controle, foram derivadas equações recursivas não-lineares determinísticas para o comportamento médio dos coeficientes, momentos de segunda ordem e erro médio quadrático para sinais de entrada gaussianos. Estas equações permitem prever importantes características do processo de adaptação e cancelamento. Em decorrência da complexidade matemática envolvida, diversas considerações simplificadoras foram realizadas. Entretanto, inúmeras simulações demonstraram a validade das mesmas nas mais diversas situações.

Foi definida a figura de mérito denominada de grau de não-linearidade, que quantifica o nível de influência da não-linearidade sobre o comportamento do algoritmo adaptativo. Um método para a sua determinação em sistemas reais é sugerido.

Foi também estudado o comportamento do algoritmo em regime permanente, tendo sido demonstrado o profundo impacto dos efeitos não-lineares sobre o máximo nível de cancelamento que pode ser obtido. A par disso, condições de estabilidade foram obtidas em função do grau de não-linearidade do sistema.

Modelos simplificados foram determinados de tal forma que, para diferentes configurações do caminho secundário e características dos sinais envolvidos, a carga computacional fosse minimizada. Entretanto, mesmo os casos mais simples ainda mantêm informações relevantes sobre a influência da não-linearidade, de tal modo que podem ser utilizados para uma compreensão geral sobre os mecanismos envolvidos no processo de adaptação em função da não-linearidade.

Os modelos analíticos demonstraram uma excelente concordância com simulações de Monte Carlo e procedimentos numéricos, mesmo quando os parâmetros do sistema se afastam das considerações de análise.

8.1 Propostas para Continuação do Trabalho

Dada a abrangência deste trabalho, diversos caminhos podem ser sugeridos para sua continuação, entre esses:

- **Análise de estabilidade** - Diversos experimentos apontam para um aumento dos limites de estabilidade do algoritmo adaptativo em relação ao seu passo de convergência à medida que aumenta o grau de não-linearidade. Além disso, ainda não foram definidos limites de estabilidade para o passo de convergência quando o modelo do FXLMS convencional leva em consideração as correlações entre os vetores de dados observados atrasados (TOBIAS, 1999).
- **Desenvolvimento de novos algoritmos** - Visto que o algoritmo FXLMS não é capaz de produzir o vetor de controle ótimo no caso não-linear, faz-se

necessária a determinação de novos algoritmos capazes de atingir o máximo cancelamento possível em um problema de estimação média quadrática não-linear (vide COSTA e BERMUDEZ (2001) no ANEXO 10).

- **Implementação de um sistema ANC** - A implementação prática de um sistema de controle ativo de som permitiria a verificação das propriedades determinadas ao longo deste trabalho.
- **Análise de sinais de banda estreita** - Um grande número de sistemas ANC faz uso de sinais de banda estreita, periódicos e/ou determinísticos. A metodologia utilizada neste trabalho aplica-se somente a sinais aleatórios gaussianos. Portanto, novas técnicas de análise devem ser pesquisadas.
- **Análise de outros algoritmos adaptativos** - Diversos algoritmos são utilizados em aplicações ANC; portanto estes também estão sujeitos à influência das características não-lineares e deveriam ser analisados quanto às conseqüências em seu desempenho.
- **Sistemas multicanal** - Muitas aplicações ANC necessitam de vários canais de controle para a obtenção de controle global em determinados ambientes fechados. Efeitos não-lineares influenciam seu desempenho e devem ser avaliados (vide COSTA *et al.* (2000c) no ANEXO 10).

ANEXO 1 – DETERMINAÇÃO DE $E\{g(y_1)Y_2\}$

Este anexo apresenta a avaliação da expressão $E\{g(y_1)Y_2\}$ utilizando-se o Teorema de Price Modificado (PAWULA, 1967). O resultado final concorda com o obtido por SHYNK e BERSHAD (1991).

A função não-linear $g(\cdot)$ é definida como

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (\text{a.1})$$

Supõe-se que y_1 e y_2 sejam variáveis aleatórias gaussianas correlacionadas de média zero.

O Teorema de Price Modificado exige que as variáveis envolvidas possuam variância unitária. Em decorrência disso, é necessária a utilização de variáveis normalizadas z_1 e z_2 relacionadas à y_1 e y_2 da seguinte forma:

$$\begin{cases} y_1 = \sigma_{y_1} z_1 \\ y_2 = \sigma_{y_2} z_2 \end{cases} \quad (\text{a.2})$$

onde z_1 e z_2 possuem variâncias unitárias, e σ_{y_1} e σ_{y_2} são os desvios padrões das variáveis y_1 e y_2 , respectivamente.

O teorema de Price modificado consiste em uma seqüência de quatro passos, apresentados a seguir.

1. Determinação de $G(z)$

Substituindo as variáveis normalizadas (a.2) em $E\{g(y_1)y_2\}$, obtém-se:

$$E\{g(y_1)y_2\} = E\{g(\sigma_{y_1}z_1)\sigma_{y_2}z_2\} \quad (\text{a.3})$$

A partir de (a.3) pode-se definir a função $G(z_1, z_2)$ como

$$G(z_1, z_2) = g(\sigma_{y_1}z_1)\sigma_{y_2}z_2 \quad (\text{a.4})$$

2. Determinação de $h[G(z)]$

O termo $h[G(z_1, z_2)]$ é obtido através de

$$h[G(z_1, z_2)] = \sum_{i < j} \rho_{i,j} \frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} \quad (\text{a.5})$$

Para avaliar (a.5) é preciso calcular a derivada parcial de (a.4), resultando em:

$$\frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} = \sigma_{y_1} \sigma_{y_2} e^{-\frac{(\sigma_{y_1}z_1)^2}{2\sigma^2}} \quad (\text{a.6})$$

onde a derivada de $g(y)$ é

$$g'(y) = e^{-\frac{y^2}{2\sigma^2}} dy \quad (\text{a.7})$$

Substituindo (a.6) em (a.5), tem-se

$$h[G(z_1, z_2)] = \rho \sigma_{y_1} \sigma_{y_2} e^{-\frac{(\sigma_{y_1}z_1)^2}{2\sigma^2}} \quad (\text{a.8})$$

onde $\rho = E\{z_1 z_2\}$ é a covariância entre z_1 e z_2 .

3. Determinação de $E_\alpha \{h [G(z)]\}$

Calculando-se o valor esperado de $h [G(z_1, z_2)]$ para $E\{z_1, z_2\} = \alpha\rho$, tem-se que

$$E_\alpha \{h [G(z_1, z_2)]\} = \int_{-\infty}^{\infty} \rho \sigma_{y_1} \sigma_{y_2} e^{-\frac{(\sigma_{y_1} z_1)^2}{2\sigma^2}} f_{z_1}(z_1) dz_1 \quad (\text{a.9})$$

onde a função densidade de probabilidade $f_{z_1}(z_1)$ é dada por (PAPOULIS, 1991):

$$f_{z_1}(z_1) = \frac{1}{\sigma_{z_1} \sqrt{2\pi}} e^{-\frac{z_1^2}{2\sigma_{z_1}^2}} \quad (\text{a.10})$$

porém, devido à normalização, tem-se que $\sigma_{z_1}^2 = 1$. Procedendo-se à integração em (a.9), chega-se a (ZWILLINGER, 1996):

$$E_\alpha \{h [G(z_1, z_2)]\} = \frac{\rho \sigma_{y_1} \sigma_{y_2} \sigma}{\sqrt{\sigma_{y_1}^2 + \sigma^2}} \quad (\text{a.11})$$

4. Determinação de R_1

Segundo PAWULA (1967),

$$E \{g(\sigma_{y_1} z_1) \sigma_{y_2} z_2\} = R_1 \quad (\text{a.12})$$

mas R_1 pode ser calculado através de

$$R_1 = E_o \{G(z_1, z_2)\} + \int_0^1 E_\alpha \{h [G(z_1, z_2)]\} d\alpha \quad (\text{a.13})$$

Para $\alpha = 0$, z_1 e z_2 são descorrelacionados e média zero, portanto

$$E_o \{G(z_1, z_2)\} = 0 \quad (\text{a.14})$$

O segundo termo de R_1 é dado por:

$$\int_0^1 E_\alpha \{h [G(z_1, z_2)]\} d\alpha = \frac{\rho \sigma_{y_1} \sigma_{y_2} \sigma}{\sqrt{\sigma_{y_1}^2 + \sigma^2}} \quad (\text{a.15})$$

Substituindo-se (a.14) e (a.15) em (a.13) e voltando-se à notação inicial em função de y_1 e y_2 , obtém-se

$$E\{g(y_1) y_2\} = \frac{\rho \sigma_{y_1} \sigma_{y_2}}{\left(\frac{\sigma_{y_1}^2}{\sigma^2} + 1\right)^{1/2}} \quad (\text{a.16})$$

onde

$$\begin{cases} \rho_{1,2} = E\{z_1 z_2\} = \frac{1}{\sigma_{y_1} \sigma_{y_2}} E\{y_1 y_2\} \\ \sigma_{y_1}^2 = E\{y_1^2\} \\ \sigma_{y_2}^2 = E\{y_2^2\} \end{cases} \quad (\text{a.17})$$

e σ^2 é o coeficiente de não-linearidade da função $g(\bullet)$. Substituindo-se as relações (a.17) em (a.16), chega-se finalmente em

$$E\{g(y_1) y_2\} = \frac{E\{y_1 y_2\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} \quad (\text{a.18})$$

Para o caso em que a variável y_2 é um vetor, tem-se:

$$E\{g(y_1) \mathbf{Y}_2\} = \frac{E\{y_1 \mathbf{Y}_2\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} \quad (\text{a.19})$$

ANEXO 2 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)\}$

Este anexo apresenta a avaliação da expressão $E\{g(y_1)g(y_2)\}$ utilizando-se o Teorema de Price Modificado (PAWULA, 1967). O resultado final concorda com o obtido por BERSHAD *et al.* (1999).

A função não-linear $g(\cdot)$ é definida como

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (\text{b.1})$$

Supõe-se que y_1 e y_2 sejam variáveis aleatórias gaussianas correlacionadas de média zero.

O Teorema de Price Modificado exige que as variáveis envolvidas possuam variância unitária. Em decorrência disso, é necessária a utilização de variáveis normalizadas z_1 e z_2 relacionadas a y_1 e y_2 da seguinte forma:

$$\begin{cases} y_1 = \sigma_{y_1} z_1 \\ y_2 = \sigma_{y_2} z_2 \end{cases} \quad (\text{b.2})$$

onde z_1 e z_2 possuem variâncias unitárias, e σ_{y_1} e σ_{y_2} são os desvios padrões das variáveis y_1 e y_2 , respectivamente.

O Teorema Modificado de Price consiste em uma seqüência de quatro passos, apresentados a seguir.

1. Determinação de $G(z)$

Substituindo as variáveis normalizadas (b.2) em $E\{g(y_1)g(y_2)\}$, tem-se:

$$E\{g(y_1)g(y_2)\} = E\{g(\sigma_{y_1}z_1)g(\sigma_{y_2}z_2)\} \quad (\text{b.3})$$

A partir de (b.3) pode-se definir a função $G(z_1, z_2)$ como

$$G(z_1, z_2) = g(\sigma_{y_1}z_1)g(\sigma_{y_2}z_2) \quad (\text{b.4})$$

2. Determinação de $h[G(z)]$

O termo $h[G(z_1, z_2)]$ é obtido através de

$$h[G(z_1, z_2)] = \sum_{i < j} \rho_{i,j} \frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} \quad (\text{b.5})$$

Para avaliar (b.5) é necessário calcular a derivada parcial de (b.4), resultando (ZWILLINGER, 1996):

$$\frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} = \sigma_{y_1} \sigma_{y_2} e^{-\frac{(\sigma_{y_1}z_1)^2}{2\sigma^2}} e^{-\frac{(\sigma_{y_2}z_2)^2}{2\sigma^2}} \quad (\text{b.6})$$

onde a derivada de $g(y)$ é

$$g'(y) = e^{-\frac{y^2}{2\sigma^2}} dy \quad (\text{b.7})$$

Substituindo (b.6) em (b.5), obtém-se

$$h[G(z_1, z_2)] = \rho \sigma_{y_1} \sigma_{y_2} e^{-\frac{(\sigma_{y_1}z_1)^2}{2\sigma^2}} e^{-\frac{(\sigma_{y_2}z_2)^2}{2\sigma^2}} \quad (\text{b.8})$$

onde $\rho = E\{z_1 z_2\}$ é a covariância entre z_1 e z_2 .

3. Determinação de $E_\alpha \{h[G(z)]\}$

Calculando-se o valor esperado de $h[G(z_1, z_2)]$ para $E\{z_1 z_2\} = \alpha\rho$, tem-se

$$E_\alpha \{h[G(z_1, z_2)]\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\rho \sigma_{y_1} \sigma_{y_2}}{2\pi |\mathbf{K}_\alpha|^{1/2}} e^{-\frac{(\sigma_{y_1} z_1)^2}{2\sigma^2}} e^{-\frac{(\sigma_{y_2} z_2)^2}{2\sigma^2}} e^{-\frac{1}{2} z^T \mathbf{K}_\alpha^{-1} z} dz_1 dz_2 \quad (\text{b.9})$$

onde $\mathbf{z} = [z_1 \ z_2]^T$. A função densidade de probabilidade conjunta $f_z(\mathbf{z})$ é (MILLER, 1964):

$$f_z(\mathbf{z}) = \frac{1}{2\pi |\mathbf{K}_\alpha|^{1/2}} e^{-\frac{1}{2} z^T \mathbf{K}_\alpha^{-1} z} \quad (\text{b.10})$$

onde

$$\mathbf{K}_\alpha = \begin{bmatrix} 1 & \alpha\rho \\ \alpha\rho & 1 \end{bmatrix} \quad (\text{b.11})$$

é a matriz de correlação entre z_1 e z_2 . Definindo-se a matriz \mathbf{Q}_α^{-1} como

$$\mathbf{Q}_\alpha^{-1} = \frac{1}{|\mathbf{K}_\alpha|} \begin{bmatrix} 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| & -\alpha\rho \\ -\alpha\rho & 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \quad (\text{b.12})$$

pode-se expressar (b.9), da seguinte forma:

$$E_\alpha \{h[G(z_1, z_2)]\} = \rho \sigma_{y_1} \sigma_{y_2} \frac{|\mathbf{Q}_\alpha|^{1/2}}{|\mathbf{K}_\alpha|^{1/2}} \left\{ \frac{1}{2\pi |\mathbf{Q}_\alpha|^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^T \mathbf{Q}_\alpha^{-1} z} dz \right\} \quad (\text{b.13})$$

Porém, segundo MILLER (1964),

$$\frac{1}{2\pi |\mathbf{Q}_\alpha|^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2} z^T \mathbf{Q}_\alpha^{-1} z} dz = 1 \quad (\text{b.14})$$

e portanto

$$E_\alpha \{h[G(z_1, z_2)]\} = \rho \sigma_{y_1} \sigma_{y_2} \frac{|\mathbf{Q}_\alpha|^{1/2}}{|\mathbf{K}_\alpha|^{1/2}} \quad (\text{b.15})$$

A matriz \mathbf{Q}_α pode ser obtida intuitivamente a partir de (b.12), como

$$\mathbf{Q}_\alpha = \frac{|\mathbf{K}_\alpha|}{\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2} \begin{bmatrix} 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| & \alpha \rho \\ \alpha \rho & 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \quad (\text{b.16})$$

Calculando-se os determinantes de \mathbf{K}_α e \mathbf{Q}_α ,

$$\begin{cases} |\mathbf{K}_\alpha| = 1 - \alpha^2 \rho^2 \\ |\mathbf{Q}_\alpha| = \frac{|\mathbf{K}_\alpha|^2}{\left[\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2\right]} \end{cases} \quad (\text{b.17})$$

e substituindo-se esses resultados em (b.15), obtém-se:

$$E_\alpha \{h[G(z_1, z_2)]\} = \frac{\sigma^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2\right]^{1/2}} \cdot \frac{\rho}{|\rho|} \quad (\text{b.18})$$

onde $|\rho| = \sqrt{\rho^2}$

4. Determinação de R_1

Segundo PAWULA (1967),

$$E\{g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2)\} = R_1 \quad (\text{b.19})$$

e R_1 pode ser calculado através de

$$R_1 = E_o\{G(z_1, z_2)\} + \int_0^1 E_\alpha\{h[G(z_1, z_2)]\} d\alpha \quad (\text{b.20})$$

Para $\alpha = 0$, z_1 e z_2 são não-correlacionados, portanto

$$E_o\{G(z_1, z_2)\} = E\{g(\sigma_{y_1} z_1)\} E\{g(\sigma_{y_2} z_2)\} = 0 \quad (\text{b.21})$$

visto que a função erro não altera a média de seu argumento e $E\{z_2\} = 0$. O segundo termo de R_1 é avaliado da seguinte forma:

$$\int_0^1 E_\alpha\{h[G(z_1, z_2)]\} d\alpha = \frac{\rho}{|\rho|} \cdot \int_0^1 \frac{\sigma^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{1/2}} d\alpha \quad (\text{b.22})$$

Procedendo-se à integração (ZWILLINGER, 1996) de (b.22) e substituindo-se o resultado, conjuntamente com (b.21), em (b.20), chega-se a

$$R_1 = \sigma^2 \cdot \frac{\rho}{|\rho|} \cdot \arcsen\left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{\sqrt{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}} \right) \quad (\text{b.23})$$

como $E\{g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2)\} = R_1$ e $\text{senal}(\rho) = \rho/|\rho|$, então:

$$E\{g(y_1) g(y_2)\} = \sigma^2 \arcsen\left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{\sqrt{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}} \right) \cdot \text{senal}(\rho) \quad (\text{b.24})$$

onde

$$\begin{cases} \rho = E\{z_1 z_2\} = \frac{1}{\sigma_{y_1} \sigma_{y_2}} E\{y_1 y_2\} \\ \sigma_{y_1}^2 = E\{y_1^2\} \\ \sigma_{y_2}^2 = E\{y_2^2\} \end{cases} \quad (\text{b.25})$$

Substituindo-se as relações (b.25) em (b.24) e levando-se em consideração o fato de o arco-seno ser uma função ímpar, obtém-se finalmente

$$E\{g(y_1)g(y_2)\} = \sigma^2 \arcsen\left(\frac{E\{y_1 y_2\}}{\sqrt{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2)}}\right) \quad (\text{b.26})$$

onde σ^2 é o coeficiente de não-linearidade da função $g(\cdot)$.

ANEXO 3 – DETERMINAÇÃO DE $E\{\mathbf{Y}_1\mathbf{Y}_2^T\mathbf{W}_1\mathbf{W}_2^T\mathbf{Y}_3\mathbf{Y}_4^T\}$

O seguinte desenvolvimento trata da simplificação de $E\{\mathbf{Y}_1\mathbf{Y}_2^T\mathbf{W}_1\mathbf{W}_2^T\mathbf{Y}_3\mathbf{Y}_4^T\}$ em termos de valores esperados de no máximo duas variáveis aleatórias. O resultado final concorda com o obtido por HAYKIN (1991) e TOBIAS (1999).

Supõe-se que \mathbf{Y}_1 , \mathbf{Y}_2 , \mathbf{Y}_3 e \mathbf{Y}_4 sejam vetores de variáveis aleatórias gaussianas correlacionadas e de média zero ($\mathbf{Y}_i = [y_{i1} \ y_{i2} \ \dots \ y_{iN}]^T$ para $i = 1, 2, 3, 4$) e \mathbf{W}_1 e \mathbf{W}_2 vetores de elementos constantes ($\mathbf{W}_i = [w_{i1} \ w_{i2} \ \dots \ w_{iN}]^T$ para $i = 1, 2$).

Arbitrando-se uma matriz \mathbf{Q} como

$$\mathbf{Q} = \mathbf{Y}_1\mathbf{Y}_2^T\mathbf{W}_1\mathbf{W}_2^T\mathbf{Y}_3\mathbf{Y}_4^T \quad (\text{c.1})$$

e expressando-se (c.1) sob a forma escalar (ZWILLINGER, 1996), tem-se

$$q_{i,j} = \sum_{l=1}^N \sum_{p=1}^N (\mathbf{Y}_1\mathbf{Y}_2^T)_{i,l} (\mathbf{W}_1\mathbf{W}_2^T)_{l,p} (\mathbf{Y}_3\mathbf{Y}_4^T)_{p,j} \quad (\text{c.2})$$

onde “ i ” representa a i -ésima linha e “ j ” a j -ésima coluna da matriz quadrada \mathbf{Q} . Os elementos de (c.2) podem ser representados mais explicitamente pela seguinte forma:

$$\begin{cases} (\mathbf{Y}_1\mathbf{Y}_2^T)_{i,l} = y_{1i}y_{2l} \\ (\mathbf{W}_1\mathbf{W}_2^T)_{l,p} = w_{1l}w_{2p} \\ (\mathbf{Y}_3\mathbf{Y}_4^T)_{p,j} = y_{3p}y_{4j} \end{cases} \quad (\text{c.3})$$

Substituindo-se (c.3) em (c.2) e tomando-se seu valor esperado,

$$E\{q_{i,j}\} = \sum_{l=1}^N \sum_{p=1}^N E\{y_{li}y_{2l}y_{3p}y_{4j}\} w_{1l}w_{2p} \quad (\text{c.4})$$

Para a simplificação desse termo, utiliza-se o teorema de fatoração de variáveis gaussianas (HAYKIN, 1991). Esse teorema estabelece a seguinte relação de valores esperados para variáveis reais gaussianas:

$$E\{x_1x_2x_3x_4\} = E\{x_1x_2\}E\{x_3x_4\} + E\{x_1x_3\}E\{x_2x_4\} + E\{x_1x_4\}E\{x_2x_3\} \quad (\text{c.5})$$

Substituindo-se (c.5) em (c.4) e rearranjando-se os termos

$$\begin{aligned} E\{q_{i,j}\} &= \sum_{l=1}^N \sum_{p=1}^N E\{y_{li}y_{2l}\} w_{1l}w_{2p} E\{y_{3p}y_{4j}\} \\ &+ \sum_{p=1}^N \sum_{l=1}^N E\{y_{li}y_{3p}\} w_{2p}w_{1l} E\{y_{2l}y_{4j}\} \\ &+ E\{y_{li}y_{4j}\} E\left\{ \sum_{l=1}^N w_{1l}y_{2l} \sum_{p=1}^N w_{2p}y_{3p} \right\} \end{aligned} \quad (\text{c.6})$$

Voltando à forma vetorial de \mathbf{Q} , obtém-se, finalmente,

$$\begin{aligned} E\{\mathbf{Y}_1\mathbf{Y}_2^T\mathbf{W}_1\mathbf{W}_2^T\mathbf{Y}_3\mathbf{Y}_4^T\} &= E\{\mathbf{Y}_1\mathbf{Y}_2^T\}\mathbf{W}_1\mathbf{W}_2^T E\{\mathbf{Y}_3\mathbf{Y}_4^T\} + E\{\mathbf{Y}_1\mathbf{Y}_3^T\}\mathbf{W}_2\mathbf{W}_1^T E\{\mathbf{Y}_2\mathbf{Y}_4^T\} \\ &+ \mathbf{W}_1^T E\{\mathbf{Y}_2\mathbf{Y}_3^T\}\mathbf{W}_2 \cdot E\{\mathbf{Y}_1\mathbf{Y}_4^T\} \end{aligned} \quad (\text{c.7})$$

ANEXO 4 – DETERMINAÇÃO DE $E\{g(y_1)y_2\mathbf{Y}_3\mathbf{Y}_4^T\}$

O seguinte desenvolvimento trata da simplificação de $E\{g(y_1)y_2\mathbf{Y}_3\mathbf{Y}_4^T\}$ em termos de valores esperados de no máximo duas variáveis aleatórias. Para isso, utiliza-se o mesmo enfoque de BERSHAD *et al.* (1999).

Supõe-se que y_1 , y_2 , \mathbf{Y}_3 e \mathbf{Y}_4 sejam variáveis aleatórias gaussianas correlacionadas de média zero, e $g(\bullet)$ uma operação não-linear sem memória sobre seu argumento e que preserve a média do mesmo.

Inicialmente é realizada a expansão das variáveis y_i (para $i = 1, 2, 3, 4$) em uma série ortonormal de quatro dimensões em torno de w_i ,

$$\begin{cases} y_1 = k_o w_1 \\ y_2 = a_1 w_1 + a_2 w_2 \\ \mathbf{Y}_3 = \mathbf{B}_1 w_1 + \mathbf{B}_2 w_2 + \mathbf{B}_3 w_3 \\ \mathbf{Y}_4 = \mathbf{C}_1 w_1 + \mathbf{C}_2 w_2 + \mathbf{C}_3 w_3 + \mathbf{C}_4 w_4 \end{cases} \quad (\text{d.1})$$

onde os w_i (para $i = 1, 2, 3, 4$) configuram uma base ortonormal para os y_i . Visto que os y_i são média zero, é razoável supor que os w_i sejam também média zero. Portanto

$$\begin{cases} E\{w_i w_j\} = 0 & \text{para } i \neq j \\ E\{w_i w_j\} = 1 & \text{para } i = j \end{cases} \quad (\text{d.2})$$

e os valores esperados dos produtos dos y_i podem ser obtidos através das seguintes relações:

$$\begin{cases} E\{y_1 y_2\} = k_o a_1 \\ E\{y_1 \mathbf{Y}_3\} = k_o \mathbf{B}_1 \\ E\{y_2 \mathbf{Y}_3\} = a_1 \mathbf{B}_1 + a_2 \mathbf{B}_2 \\ \dots \end{cases} \quad (\text{d.3})$$

onde $k_o^2 = E\{y_i^2\}$. Substituindo-se (d.1) em $E\{g(y_i)y_2Y_3Y_4^T\}$, tem-se

$$E\left\{g(k_o w_1) \cdot [a_1 w_1 + a_2 w_2] \cdot [B_1 w_1 + B_2 w_2 + B_3 w_3] \cdot [C w_1 + C w_2 + C_3 w_3 + C_4 w_4]^T\right\} \quad (d.4)$$

mas

$$\begin{aligned} & [a_1 w_1 + a_2 w_2] \cdot [B_1 w_1 + B_2 w_2 + B_3 w_3] \cdot [C_1 w_1 + C_2 w_2 + C_3 w_3 + C_4 w_4]^T = \\ & = (a_1 B_1 C_1^T) w_1^3 + (a_2 B_2 C_2^T) w_2^3 + (a_1 B_2 C_2^T + a_2 B_1 C_2^T + a_2 B_2 C_1^T) w_1 w_2^2 \\ & + (a_1 B_3 C_3^T) w_1 w_3^2 + (a_2 B_2 C_3^T) w_2 w_3^2 + (a_1 B_2 C_1^T + a_1 B_1 C_2^T + a_2 B_1 C_1^T) w_1^2 w_2 \\ & + (a_1 B_3 C_1^T + a_1 B_1 C_3^T) w_1^2 w_3 + (a_1 B_1 C_4^T) w_1^2 w_4 + (a_2 B_3 C_2^T + a_2 B_2 C_3^T) w_2^2 w_3 \\ & + (a_2 B_2 C_4^T) w_2^2 w_4 + (a_2 B_3 C_1^T + a_1 B_3 C_2 + a_1 B_2 C_3^T + a_2 B_1 C_3^T) w_1 w_2 w_3 \\ & + (a_1 B_2 C_4^T + a_2 B_1 C_4) w_1 w_2 w_4 + (a_1 B_3 C_4^T) w_1 w_3 w_4 + (a_2 B_3 C_4^T) w_2 w_3 w_4 \end{aligned} \quad (d.5)$$

A equação (d.5) é substituída em (d.4) e são realizadas as seguintes considerações:

- os coeficientes k_o , a_i , B_i e C_i são constantes;
- visto que os w_i são ortonormais, pode-se separar o valor esperado das multiplicações de w_i de diferentes índices pela multiplicação dos valores esperados de w_i de mesmos índices: $E\{f(w_i, w_j)\}_{i \neq j} = E\{f_1(w_i)\}E\{f_2(w_j)\}$;
- como a média dos w_i é zero, então todos os valores esperados obtidos através do item anterior que possuírem w_i elevado a uma potência ímpar serão nulos: $E\{w_i\} = E\{w_i^3\} = \dots = 0$;
- como a não-linearidade não influi no valor médio da variável, então: $E\{g(k_o w_1)\} = 0$.

Como resultado, a equação (d.4) reduz-se:

$$\begin{aligned}
E\{g(y_1)y_2\mathbf{Y}_3\mathbf{Y}_4^T\} &= a_1\mathbf{B}_1\mathbf{C}_1^T E\{w_1^3g(k_0w_1)\} \\
&+ [a_1\mathbf{B}_2\mathbf{C}_2^T + a_2\mathbf{B}_1\mathbf{C}_2^T + a_2\mathbf{B}_2\mathbf{C}_1^T] E\{w_1g(k_0w_1)\}E\{w_2^2\} \\
&+ a_1\mathbf{B}_3\mathbf{C}_3^T E\{w_1g(k_0w_1)\} E\{w_3^2\}
\end{aligned} \tag{d.6}$$

mas $E\{w_2^2\} = E\{w_3^2\} = 1$, então:

$$\begin{aligned}
E\{g(y_1)y_2\mathbf{Y}_3\mathbf{Y}_4^T\} &= a_1\mathbf{B}_1\mathbf{C}_1^T E\{w_1^3g(k_0w_1)\} \\
&+ [a_1\mathbf{B}_2\mathbf{C}_2^T + a_2\mathbf{B}_1\mathbf{C}_2^T + a_2\mathbf{B}_2\mathbf{C}_1^T + a_1\mathbf{B}_3\mathbf{C}_3^T] E\{w_1g(k_0w_1)\}
\end{aligned} \tag{d.7}$$

Consideremos:

$$\begin{cases}
E\{y_1y_2\} = k_0a_1 \\
E\{y_1\mathbf{Y}_3\} = k_0\mathbf{B}_1 \\
E\{y_1\mathbf{Y}_4^T\} = k_0\mathbf{C}_1^T \\
E\{y_2\mathbf{Y}_3\} = a_1\mathbf{B}_1 + a_2\mathbf{B}_2 \\
E\{y_2\mathbf{Y}_4^T\} = a_1\mathbf{C}_1^T + a_2\mathbf{C}_2^T \\
E\{\mathbf{Y}_3\mathbf{Y}_4^T\} = \mathbf{B}_1\mathbf{C}_1^T + \mathbf{B}_2\mathbf{C}_2^T + \mathbf{B}_3\mathbf{C}_3^T
\end{cases} \tag{d.8}$$

Após um pouco de manipulação algébrica, pode-se obter as seguintes expressões:

$$a_1\mathbf{B}_1\mathbf{C}_1^T = \frac{1}{k_0^3} E\{y_1y_2\}E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\} \tag{d.9}$$

e

$$\begin{aligned}
(a_1\mathbf{B}_2\mathbf{C}_2^T + a_1\mathbf{B}_3\mathbf{C}_3^T + a_2\mathbf{B}_1\mathbf{C}_2^T + a_2\mathbf{B}_2\mathbf{C}_1^T) &= -\frac{3}{k_0^3} E\{y_1y_2\} E\{y_1\mathbf{Y}_3\} E\{y_1\mathbf{Y}_4^T\} \\
&+ \frac{1}{k_0} [E\{y_1y_2\}E\{\mathbf{Y}_3\mathbf{Y}_4^T\} + E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]
\end{aligned} \tag{d.10}$$

Substituindo esses resultados em (d.7) e sabendo-se que $k_0^2 = E\{y_1^2\}$, obtém-se,

finalmente:

$$\begin{aligned}
E\{g(y_1)y_2\mathbf{Y}_3\mathbf{Y}_4^T\} &= E\{y_1y_2\}E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\} \left[\frac{E\{g(y_1)y_1^3\}}{E\{y_1^2\}^3} - 3\frac{E\{g(y_1)y_1\}}{E\{y_1^2\}^2} \right] \\
&+ \left[E\{y_1y_2\}E\{\mathbf{Y}_3\mathbf{Y}_4^T\} + E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\} \right] \frac{E\{g(y_1)y_1\}}{E\{y_1^2\}}
\end{aligned}
\tag{d.11}$$

ANEXO 5 – DETERMINAÇÃO DE $E\{g(y)y^3\}$

Este apêndice refere-se ao cálculo da expressão $E\{g(y)y^3\}$ utilizando-se apenas técnicas de integração.

A função não-linear $g(\bullet)$ é definida como

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (\text{e.1})$$

Supõe-se que y seja uma variável aleatória gaussiana de média zero, cuja função densidade de probabilidade seja expressa por (PAPOULIS, 1991):

$$f_y(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_y^2}} \quad (\text{e.2})$$

O valor esperado de $g(y)y^3$ pode ser obtido integrando-se essa função de forma ponderada por sua função densidade de probabilidade (THERRIEN, 1991):

$$E\{g(y)y^3\} = \int_{-\infty}^{\infty} g(y)y^3 \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_y^2}} dy \quad (\text{e.3})$$

Realizando-se a seguinte troca de variáveis:

$$\begin{cases} u = g(y) \rightarrow du = e^{-\frac{y^2}{2\sigma_y^2}} dy \\ dv = y^3 e^{-\frac{y^2}{2\sigma_y^2}} dy \rightarrow v = -\sigma_y^2 e^{-\frac{y^2}{2\sigma_y^2}} (y^2 + 2\sigma_y^2) \end{cases} \quad (\text{e.4})$$

Pode-se agora resolver a equação (e.3) integrando-se por partes (ZWILLINGER, 1996), ou seja,

$$\int_A^B u dv = uv \Big|_A^B - \int_A^B v du \quad (\text{e.5})$$

por conseguinte,

$$E\{g(y)y^3\} = \lim_{A \rightarrow \infty} \frac{1}{\sigma_y \sqrt{2\pi}} \left\{ -g(y)\sigma_y^2 e^{-\frac{y^2}{2\sigma_y^2}} (y^2 + 2\sigma_y^2) \Big|_{-A}^A + \int_{-A}^A \sigma_y^2 e^{-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)y^2} (y^2 + 2\sigma_y^2) dy \right\} \quad (\text{e.6})$$

Visto que o primeiro termo de (e.6) é uma função ímpar, pode-se escrever a seguinte igualdade:

$$\lim_{A \rightarrow \infty} \left(\frac{\sigma_y}{\sqrt{2\pi}} g(y) e^{-\frac{y^2}{2\sigma_y^2}} (y^2 + 2\sigma_y^2) \Big|_{-A}^A \right) = \frac{\sqrt{2}\sigma_y}{\sqrt{\pi}} \lim_{y \rightarrow \infty} \left(g(y) e^{-\frac{y^2}{2\sigma_y^2}} (y^2 + 2\sigma_y^2) \right) \quad (\text{e.7})$$

Entretanto,

$$\lim_{y \rightarrow \infty} g(y) = \frac{\sigma \sqrt{\pi}}{\sqrt{2}} \quad ; \quad \lim_{y \rightarrow \infty} e^{-\frac{y^2}{2\sigma_y^2}} = 0 \quad ; \quad \lim_{y \rightarrow \infty} y^2 e^{-\frac{y^2}{2\sigma_y^2}} = 0 \quad (\text{e.8})$$

Substituindo-se (e.8) em (e.7), chega-se a

$$\lim_{A \rightarrow \infty} \left(\frac{\sigma_y}{\sqrt{2\pi}} g(y) \sigma_y^2 e^{-\frac{y^2}{2\sigma_y^2}} (y^2 + 2\sigma_y^2) \Big|_{-A}^A \right) = 0 \quad (\text{e.9})$$

Substituindo-se (e.9) em (e.6):

$$E\{g(y)y^3\} = \frac{\sigma_y}{\sqrt{2\pi}} \left\{ 2 \int_0^{\infty} y^2 e^{-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)y^2} dy + 4 \sigma_y^2 \int_0^{\infty} e^{-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)y^2} dy \right\} \quad (\text{e.10})$$

Porém, a primeira integral de (e.10) pode ser descrita por (ZWILLINGER, 1996):

$$\int_0^{\infty} y^2 e^{-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)y^2} dy = \sqrt{\frac{\pi}{2}} \frac{1}{\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)^{3/2}} \quad (\text{e.11})$$

O segundo termo da equação (e.10) também pode ser obtido em ZWILLINGER (1996) e é dado por

$$\int_0^{\infty} e^{-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)y^2} dy = \sqrt{\frac{\pi}{2}} \frac{1}{\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_y^2}\right)^{1/2}} \quad (\text{e.12})$$

Substituindo os resultados (e.11) e (e.12) em (e.10) e sabendo-se que $\sigma_y^2 = E\{y^2\}$, obtém-se, finalmente:

$$E\{g(y)y^3\} = \frac{E\{y^2\}^2}{\left(\frac{E\{y^2\}}{\sigma^2} + 1\right)^{3/2}} + 2 \frac{E\{y^2\}^2}{\left(\frac{E\{y^2\}}{\sigma^2} + 1\right)^{1/2}} \quad (\text{e.13})$$

onde σ^2 é o coeficiente de não-linearidade da função $g(\cdot)$.

ANEXO 6 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)Y_3Y_4^T\}$

O seguinte desenvolvimento trata da simplificação de $E\{g(y_1)g(y_2)Y_3Y_4^T\}$ em termos de valores esperados de no máximo duas variáveis aleatórias. Para isso, utiliza-se o mesmo enfoque de BERSHAD *et al.* (1999).

Supõe-se que $y_1, y_2, Y_3 = [y_{31} \ y_{32} \ \dots \ y_{3N}]^T$ e $Y_4 = [y_{41} \ y_{42} \ \dots \ y_{4N}]^T$ sejam variáveis aleatórias gaussianas correlacionadas de média zero. Letras minúsculas indicam escalares; e maiúsculas em negrito, vetores coluna; $g(\cdot)$ representa uma operação não-linear sem memória.

Os vetores colunas Y_i são formados por N variáveis aleatórias pertencentes ao mesmo espaço vetorial i .

Inicialmente é realizada a expansão das variáveis y_i (para $i = 1, 2, 3, 4$) em uma série ortonormal de quatro dimensões em torno de w_i :

$$\begin{cases} y_1 = k_0 w_1 \\ y_2 = a_1 w_1 + a_2 w_2 \\ Y_3 = B_1 w_1 + B_2 w_2 + B_3 w_3 \\ Y_4 = C_1 w_1 + C_2 w_2 + C_3 w_3 + C_4 w_4 \end{cases} \quad (\text{f.1})$$

onde os w_i (para $i = 1, 2, 3, 4$) configuram uma base ortonormal para os y_i . Os coeficientes k_0 e a_i (para $i = 1, 2$) são escalares e os coeficientes B_i e C_i (para $i = 1, 2, 3, 4$) são vetores coluna. Visto que os y_i são média zero, é razoável supor que os w_i sejam também média zero. Portanto

$$\begin{cases} E\{w_i w_j\} = 0 & \text{para } i \neq j \\ E\{w_i w_j\} = 1 & \text{para } i = j \end{cases} \quad (\text{f.2})$$

Substituindo-se as expressões de (f.1) em $E\{g(y_1)g(y_2)Y_3Y_4^T\}$, tem-se

$$E\{g(k_0 w_1) \cdot g(a_1 w_1 + a_2 w_2) \cdot [B_1 w_1 + B_2 w_2 + B_3 w_3] \cdot [C_1^T w_1 + C_2^T w_2 + C_3^T w_3 + C_4^T w_4]\} \quad (\text{f.3})$$

mas

$$\begin{aligned} [B_1 w_1 + B_2 w_2 + B_3 w_3] \cdot [C_1 w_1 + C_2 w_2 + C_3 w_3]^T &= \\ &= B_1 C_1^T w_1^2 + B_2 C_2^T w_2^2 + B_3 C_3^T w_3^2 + (B_1 C_2^T + B_2 C_1^T) w_1 w_2 \\ &+ (B_1 C_3^T + B_3 C_1^T) w_1 w_3 + (B_2 C_3^T + B_3 C_2^T) w_2 w_3 + B_1 C_4^T w_1 w_4 \\ &+ B_2 C_4^T w_2 w_4 + B_3 C_4^T w_3 w_4 \end{aligned} \quad (\text{f.4})$$

A equação (f.4) é substituída em (f.3) e são realizadas as seguintes considerações:

- os coeficientes k_0 , a_i , B_i e C_i são constantes;
- visto que os w_i são ortonormais, pode-se separar o valor esperado das multiplicações de w_i de diferentes índices pela multiplicação dos valores esperados de w_i de mesmos índices: $E\{f(w_i, w_j)\}_{i \neq j} = E\{f_1(w_i)\}E\{f_2(w_j)\}$;
- como a média dos w_i é zero, então: $E\{w_i\} = 0$.

Sabendo-se que $E\{w_3^2\} = 1$, então

$$\begin{aligned} E\{g(y_1)g(y_2)Y_3Y_4^T\} &= T_1 + T_2 + T_3 + T_4 \\ &= B_1 C_1^T E\{g(k_0 w_1)g(a_1 w_1 + a_2 w_2)w_1^2\} \\ &+ B_2 C_2^T E\{g(k_0 w_1)g(a_1 w_1 + a_2 w_2)w_2^2\} \\ &+ B_3 C_3^T E\{g(k_0 w_1)g(a_1 w_1 + a_2 w_2)\} \\ &+ (B_1 C_2^T + B_2 C_1^T) E\{g(k_0 w_1)g(a_1 w_1 + a_2 w_2)w_1 w_2\} \end{aligned} \quad (\text{f.5})$$

Expressando-se as variáveis w_1 e w_2 da seguinte forma

$$\begin{cases} w_1 = \frac{1}{k_o} y_1 \\ w_2 = \frac{1}{a_2} y_2 - \frac{a_1}{a_2} w_1 \end{cases} \quad (\text{f.6})$$

podemos substituir (f.6) em (f.5) e obter uma expressão em função das variáveis aleatórias y_1 e y_2 e dos coeficientes k_o , a_1 , \mathbf{B}_i e \mathbf{C}_i . Entretanto, isso é possível somente se o coeficiente a_2 for diferente de zero ($a_2 = E\{y_1^2\}E\{y_2^2\} - E\{y_1 y_2\}^2 \neq 0$). Portanto, a seguir será desenvolvido o equacionamento supondo-se $a_2 \neq 0$, e em seqüência a solução para o caso especial em que $a_2 = 0$.

Substituindo-se (f.1) e (f.6) em (f.5) chega-se a

$$\begin{aligned} E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\} &= \mathbf{B}_3\mathbf{C}_3^T E\{g(y_1)g(y_2)\} \\ &+ \left[\frac{\mathbf{B}_1\mathbf{C}_1^T}{k_o^2} - \frac{a_1}{k_o^2 a_2} (\mathbf{B}_1\mathbf{C}_2^T + \mathbf{B}_2\mathbf{C}_1^T) + \frac{a_1^2}{k_o^2 a_2^2} \mathbf{B}_2\mathbf{C}_2^T \right] E\{g(y_1)g(y_2)y_1^2\} \\ &+ \frac{\mathbf{B}_2\mathbf{C}_2^T}{a_2^2} E\{g(y_1)g(y_2)y_2^2\} \\ &+ \left[\frac{1}{k_o a_2} (\mathbf{B}_1\mathbf{C}_2^T + \mathbf{B}_2\mathbf{C}_1^T) - 2 \frac{a_1}{k_o a_2^2} \mathbf{B}_2\mathbf{C}_2^T \right] E\{g(y_1)g(y_2)y_1 y_2\} \end{aligned} \quad (\text{f.7})$$

Para avaliar os coeficientes dos valores esperados não-lineares de (f.7) é necessário primeiramente obter as seguintes relações:

$$\begin{cases} E\{y_1^2\} = k_o^2 & E\{y_2^2\} = a_1^2 + a_2^2 \\ E\{y_1 y_2\} = k_o a_1 & E\{y_1 \mathbf{Y}_3\} = k_o \mathbf{B}_1 \\ E\{y_1 \mathbf{Y}_4\} = k_o \mathbf{C}_1 & E\{y_2 \mathbf{Y}_3\} = a_1 \mathbf{B}_1 + a_2 \mathbf{B}_2 \\ E\{y_2 \mathbf{Y}_4\} = a_1 \mathbf{C}_1 + a_2 \mathbf{C}_2 & E\{\mathbf{Y}_3 \mathbf{Y}_4^T\} = \mathbf{B}_1 \mathbf{C}_1^T + \mathbf{B}_2 \mathbf{C}_2^T + \mathbf{B}_3 \mathbf{C}_3^T \end{cases} \quad (\text{f.8})$$

Substituindo-se (f.8) em (f.7), consegue-se chegar a

$$\begin{aligned}
E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\} &= \left[E\{\mathbf{Y}_3\mathbf{Y}_4^T\} \right. \\
&\quad \frac{E\{y_1^2\}E\{y_2\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2^2\}E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}}{E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2} \\
&\quad \left. + \frac{E\{y_1y_2\}[E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2} \right] E\{g(y_1)g(y_2)\} \\
&\quad + \left[\frac{E\{y_2^2\}^2 E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right. \\
&\quad \frac{E\{y_2^2\}E\{y_1y_2\}[E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \\
&\quad \left. + \frac{E\{y_1y_2\}^2 E\{y_2\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\}}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right] E\{g(y_1)g(y_2)y_1^2\} \\
&\quad + \left[\frac{E\{y_1^2\}^2 E\{y_2\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\}}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right. \\
&\quad \frac{E\{y_1^2\}E\{y_1y_2\}[E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \\
&\quad \left. + \frac{E\{y_1y_2\}^2 E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right] E\{g(y_1)g(y_2)y_2^2\} \\
&\quad + \left[\frac{E\{y_1^2\}E\{y_2^2\} [E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right. \\
&\quad \left. - 2 \frac{E\{y_1y_2\}[E\{y_1^2\}E\{y_2\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2^2\}E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right. \\
&\quad \left. + \frac{E\{y_1y_2\}^2 [E\{y_1\mathbf{Y}_3\}E\{y_2\mathbf{Y}_4^T\} + E\{y_2\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}]}{(E\{y_1^2\}E\{y_2^2\} - E\{y_1y_2\}^2)^2} \right] E\{g(y_1)g(y_2)y_1y_2\}
\end{aligned}$$

Como descrito anteriormente, e como pode ser constatado através do denominador de (f.9), esse resultado é válido apenas se $a_2 \neq 0$ (equação (f.6)). Da mesma forma, supõe-se que $E\{y_1^2\}$ e $E\{y_2^2\}$ são diferentes de zero.

O caso em que $a_2 = 0$ deve ser desenvolvido separadamente. Ou seja, retornando-se à equação (f.5), tem-se que, para $a_2 = 0$:

$$E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\} = \mathbf{B}_1\mathbf{C}_1^T E\{g(k_0 w_1)g(a_1 w_1)w_1^2\} + [\mathbf{B}_2\mathbf{C}_2^T + \mathbf{B}_3\mathbf{C}_3^T] E\{g(k_0 w_1)g(a_1 w_1)\} \quad (\text{f.10})$$

Utilizando-se (f.1) e $a_2 = 0$ tem-se que

$$E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\} = \frac{E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}}{E\{y_1^2\}^2} E\{g(y_1)g(y_2)y_1^2\} + \left[E\{\mathbf{Y}_3\mathbf{Y}_4^T\} - \frac{E\{y_1\mathbf{Y}_3\}E\{y_1\mathbf{Y}_4^T\}}{E\{y_1^2\}} \right] E\{g(y_1)g(y_2)\} \quad (\text{f.11})$$

Finalizando, a solução de $E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\}$ constitui-se por

$$E\{g(y_1)g(y_2)\mathbf{Y}_3\mathbf{Y}_4^T\} = \begin{cases} 0 & \text{se } E\{y_1^2\} = 0 \text{ ou } E\{y_2^2\} = 0 \\ \text{Eq. (f.10)} & \text{se } E\{y_1^2\}E\{y_2^2\} - E\{y_1 y_2\}^2 \neq 0 \\ \text{Eq. (f.12)} & \text{se } E\{y_1^2\}E\{y_2^2\} - E\{y_1 y_2\}^2 = 0 \end{cases} \quad (\text{f.12})$$

ANEXO 7 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)y_1^2\}$

Este apêndice refere-se ao cálculo da expressão $E\{g(y_1)g(y_2)y_1^2\}$ utilizando-se o Teorema de Price Modificado (PAWULA, 1967), onde

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \tag{g.1}$$

Supõe-se que y_1 e y_2 sejam variáveis aleatórias gaussianas correlacionadas de média zero.

O Teorema de Price Modificado exige que as variáveis envolvidas possuam variância unitária. Em decorrência disso é necessária a utilização de variáveis normalizadas z_1 e z_2 relacionadas à y_1 e y_2 da seguinte forma:

$$\begin{cases} y_1 = \sigma_{y_1} z_1 \\ y_2 = \sigma_{y_2} z_2 \end{cases} \tag{g.2}$$

onde z_1 e z_2 possuem variâncias unitárias e σ_{y_1} e σ_{y_2} são os desvios padrões das variáveis y_1 e y_2 , respectivamente.

O Teorema de Price Modificado consiste em uma seqüência de quatro passos, apresentados a seguir.

1. Determinação de $G(z)$

Substituindo as variáveis normalizadas (g.2) em $E\{g(y_1)g(y_2)y_1^2\}$, tem-se

$$E\{g(y_1)g(y_2)y_1^2\} = E\{g(\sigma_{y_1} z_1)g(\sigma_{y_2} z_2)\sigma_{y_1}^2 z_1^2\} \tag{g.3}$$

A partir de (g.3) podemos definir a função $G(z_1, z_2)$ como

$$G(z_1, z_2) = \sigma_{y_1}^2 z_1^2 g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) \quad (\text{g.4})$$

2. Determinação de $h[G(z)]$

O termo $h[G(z_1, z_2)]$ é obtido através de

$$h[G(z_1, z_2)] = \sum_{i < j} \rho_{i,j} \frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} \quad (\text{g.5})$$

Para avaliar (g.5) precisamos calcular a derivada parcial de (g.4):

$$\frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} = 2\sigma_{y_1}^2 \sigma_{y_2} z_1 g(\sigma_{y_1} z_1) e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} + \sigma_{y_1}^3 \sigma_{y_2} z_1^2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} \quad (\text{g.6})$$

onde a derivada de $g(y)$ é:

$$g'(y) = e^{-\frac{y^2}{2\sigma^2}} dy \quad (\text{g.7})$$

Substituindo (g.6) em (g.5), tem-se

$$h[G(z_1, z_2)] = 2\rho\sigma_{y_1}^2 \sigma_{y_2} z_1 g(\sigma_{y_1} z_1) e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} + \rho\sigma_{y_1}^3 \sigma_{y_2} z_1^2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} \quad (\text{g.8})$$

onde $\rho = E\{z_1 z_2\}$ é a covariância entre z_1 e z_2 .

3. Determinação de $E_\alpha\{h[G(z)]\}$

Calculando-se o valor esperado de $h[G(z_1, z_2)]$ para $E\{z_1 z_2\} = \alpha\rho$, tem-se

$$\begin{aligned}
 E_{\alpha} \{h [G(z_1, z_2)]\} &= 2\rho\sigma_{y_1}^2\sigma_{y_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 g(\sigma_{y_1} z_1) e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} f_z(\mathbf{z}) dz_1 dz_2 \\
 &+ \rho\sigma_{y_1}^3\sigma_{y_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1^2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} f_z(\mathbf{z}) dz_1 dz_2
 \end{aligned}
 \tag{g.9}$$

onde a função densidade de probabilidade conjunta $f_z(\mathbf{z})$ é (MILLER, 1964)

$$f_z(\mathbf{z}) = \frac{1}{2\pi |\mathbf{K}_{\alpha}|^{1/2}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{K}_{\alpha}^{-1} \mathbf{z}}
 \tag{g.10}$$

onde $\mathbf{z} = [z_1 \quad z_2]^T$ e:

$$\mathbf{K}_{\alpha} = \begin{bmatrix} 1 & \alpha\rho \\ \alpha\rho & 1 \end{bmatrix}
 \tag{g.11}$$

é a matriz de correlação entre z_1 e z_2 . Substituindo (g.10) em (g.9),

$$\begin{aligned}
 E_{\alpha} \{h [G(z_1, z_2)]\} &= T_1 + T_2 \\
 &= \frac{2\rho\sigma_{y_1}^2\sigma_{y_2}}{2\pi |\mathbf{K}_{\alpha}|^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 g(\sigma_{y_1} z_1) e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{K}_{\alpha}^{-1} \mathbf{z}} dz_1 dz_2 \\
 &+ \frac{\rho\sigma_{y_1}^3\sigma_{y_2}}{2\pi |\mathbf{K}_{\alpha}|^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1^2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{K}_{\alpha}^{-1} \mathbf{z}} dz_1 dz_2
 \end{aligned}
 \tag{g.12}$$

onde a inversa de \mathbf{K}_{α} pode ser determinada intuitivamente, a partir de (g.11), como

$$\mathbf{K}_{\alpha}^{-1} = \frac{1}{|\mathbf{K}_{\alpha}|} \begin{bmatrix} 1 & -\alpha\rho \\ -\alpha\rho & 1 \end{bmatrix}
 \tag{g.13}$$

Definindo-se a matriz \mathbf{L}_{α} e sua inversa \mathbf{L}_{α}^{-1} como

$$\left\{ \begin{array}{l} \mathbf{L}_\alpha = \frac{1}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \begin{bmatrix} 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| & \alpha\rho \\ \alpha\rho & 1 \end{bmatrix} \\ \mathbf{L}_\alpha^{-1} = \frac{1}{|\mathbf{K}_\alpha|} \begin{bmatrix} 1 & -\alpha\rho \\ -\alpha\rho & 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \\ |\mathbf{L}_\alpha| = \frac{|\mathbf{K}_\alpha|}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \end{array} \right. \quad (\text{g.14})$$

então o primeiro termo de (g.12) pode ser escrito como

$$T_1 = \frac{2\rho\sigma_{y_1}\sigma_{y_2}}{\left(\frac{\sigma_{y_2}^2}{\sigma^2} + 1\right)^{1/2}} E\left\{g(\sigma_{y_1}z_1)\sigma_{y_1}z_1\right\} \Big|_{\mathbf{L}_\alpha} \quad (\text{g.15})$$

onde $E\{\cdot\} \Big|_{\mathbf{L}_\alpha}$ significa a operação de valor esperado do argumento em termos da função densidade de probabilidade descrita por \mathbf{L}_α .

Utilizando o resultado obtido no ANEXO 1 e fazendo $y_1 = y_2 = y$, tem-se que

$$E\{g(y)y\} = \frac{E\{y^2\}}{\left(\frac{1}{\sigma^2}E\{y^2\} + 1\right)^{1/2}} \quad (\text{g.16})$$

Substituindo-se (g.16) em (g.15) obtém-se

$$T_1 = \frac{2\rho\sigma_{y_1}\sigma_{y_2}}{\left(\frac{\sigma_{y_2}^2}{\sigma^2} + 1\right)^{1/2}} \cdot \frac{E\{\sigma_{y_1}^2 z_1^2\} \Big|_{\mathbf{L}_\alpha}}{\left(\frac{1}{\sigma^2}E\{\sigma_{y_1}^2 z_1^2\} \Big|_{\mathbf{L}_\alpha} + 1\right)^{1/2}} \quad (\text{g.17})$$

mas $E\{\sigma_{y_1}^2 z_1^2\}_{L_\alpha} = \sigma_{y_1}^2 E\{z_1^2\}_{L_\alpha}$. O significado físico deste valor esperado é a variância da variável aleatória z_1 para uma função densidade de probabilidade descrita por

$$f_z(\mathbf{z}) = \frac{1}{2\pi|\mathbf{L}_\alpha|^{1/2}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{L}_\alpha^{-1} \mathbf{z}} \tag{g.18}$$

A avaliação desse termo não exige cálculos, visto que seu valor é o elemento da primeira linha e primeira coluna da matriz de correlação \mathbf{L}_α , portanto

$$E\{z_1^2\}_{L_\alpha} = \frac{1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \tag{g.19}$$

Substituindo (g.19) e $|\mathbf{K}_\alpha| = 1 - \alpha^2 \rho^2$ em (g.17), chega-se a

$$T_1 = \frac{2\sigma_{y_1}^2 \sigma^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} \cdot \frac{\rho}{|\rho|} - \frac{2\sigma_{y_1}^2 \sigma_{y_2}^2}{\frac{\sigma_{y_2}^2}{\sigma^2} + 1} \cdot \frac{\alpha^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} \cdot \frac{\rho^3}{|\rho|} \tag{g.20}$$

onde: $|\rho| = \sqrt{\rho^2}$

Analisando-se o segundo termo (T2) de (g.12), pode-se escrevê-lo da seguinte forma:

$$T_2 = \rho \sigma_{y_1}^3 \sigma_{y_2} \frac{|\mathbf{Q}_\alpha|^{1/2}}{|\mathbf{K}_\alpha|^{1/2}} E\{z_1^2\}_{Q_\alpha} \tag{g.21}$$

onde $E\{\cdot\}_{Q_\alpha}$ significa a operação de valor esperado do argumento em termos da função densidade de probabilidade descrita por

$$f_z(\mathbf{z}) = \frac{1}{2\pi|\mathbf{Q}_\alpha|^{1/2}} e^{-\frac{1}{2}\mathbf{z}^T \mathbf{Q}_\alpha^{-1} \mathbf{z}} \tag{g.22}$$

onde

$$\left\{ \begin{aligned} \mathbf{Q}_\alpha &= \frac{|\mathbf{K}_\alpha|}{\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2} \begin{bmatrix} 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| & \alpha\rho \\ \alpha\rho & 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \\ \mathbf{Q}_\alpha^{-1} &= \frac{1}{|\mathbf{K}_\alpha|} \begin{bmatrix} 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| & -\alpha\rho \\ -\alpha\rho & 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \\ |\mathbf{Q}_\alpha| &= \frac{|\mathbf{K}_\alpha|^2}{\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2} \end{aligned} \right. \tag{g.23}$$

O valor esperado em (g.21) pode ser interpretado como a variância da variável gaussiana z_1 segundo a função densidade de probabilidade dada por (g.22). Essa informação pode ser obtida diretamente da matriz de correlação \mathbf{Q}_α :

$$E\{z_1^2\} \Big|_{\mathbf{Q}_\alpha} = \frac{|\mathbf{K}_\alpha| \left(\frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| + 1 \right)}{\left(\frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| + 1 \right) \left(\frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| + 1 \right) - \alpha^2 \rho^2} \tag{g.24}$$

Substituindo-se (g.23) e (g.24) em (g.21), chega-se a

$$T_2 = \frac{\sigma^6 \frac{\rho}{|\rho|^3} \left(\frac{\sigma_{y_2}^2}{\sigma^2} + 1 \right)}{\sigma_{y_2}^2 \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} \frac{\sigma^4 \alpha^2 \frac{\rho^3}{|\rho|^3}}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} \tag{g.25}$$

Juntando os termos T_1 (equação (g.20)) e T_2 (equação (25)) em (g.12), chega-se a

$$\begin{aligned}
 E_{\alpha} \{h [G(z_1, z_2)]\} &= \frac{2\sigma_{y_1}^2 \sigma^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} \cdot \frac{\rho}{|\rho|} \\
 &+ \frac{2\sigma_{y_1}^2 \sigma_{y_2}^2}{\frac{\sigma_{y_2}^2}{\sigma^2} + 1} \cdot \frac{\alpha^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} \cdot \frac{\rho^3}{|\rho|} \\
 &+ \frac{\sigma^6 \left(\frac{\sigma_{y_2}^2}{\sigma^2} + 1 \right)}{\sigma_{y_2}^2 \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} \cdot \frac{\rho}{|\rho|^3} \\
 &+ \frac{\sigma^4 \alpha^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} \cdot \frac{\rho^3}{|\rho|^3}
 \end{aligned} \tag{g.26}$$

4. Determinação de R_1

Segundo PAWULA (1967),

$$E\{g(\sigma_{y_1} z_1)g(\sigma_{y_2} z_2) \sigma_{y_1}^2 z_1^2\} = R_1 \tag{g.27}$$

e R_1 pode ser calculado através de

$$R_1 = E_o \{G(z_1, z_2)\} + \int_0^1 E_{\alpha} \{h [G(z_1, z_2)]\} d\alpha \tag{g.28}$$

Para $\alpha = 0$, z_1 e z_2 são descorrelacionados, portanto

$$E_o \{G(z_1, z_2)\} = E\{g(y_1)y_1^2\} \cdot E\{g(y_2)\} = 0 \tag{g.29}$$

visto que a função erro não altera a média de um argumento média zero e $E\{y_2\}=0$.

Portanto

$$\begin{aligned}
 R_1 = & 2\sigma_{y_1}^2 \sigma^2 \cdot \frac{\rho}{|\rho|} \cdot \int_0^1 \frac{1}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} d\alpha \\
 & - \frac{2\sigma_{y_1}^2 \sigma_{y_2}^2}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \cdot \frac{\rho^3}{|\rho|} \cdot \int_0^1 \frac{\alpha^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{1/2}} d\alpha \\
 & + \frac{\sigma^6 \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} \right)}{\sigma_{y_2}^2 \rho^2} \cdot \frac{\rho}{|\rho|^3} \cdot \int_0^1 \frac{1}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} d\alpha \\
 & - \sigma^4 \cdot \frac{\rho}{|\rho|} \cdot \int_0^1 \frac{\alpha^2}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - \alpha^2 \right]^{3/2}} d\alpha
 \end{aligned} \tag{g.30}$$

As quatro integrais em R_1 podem ser facilmente avaliadas através de uma tabela de integrais, de tal forma que (ZWILLINGER, 1996)

$$T_1 = 2\sigma_{y_1}^2 \sigma^2 \cdot \frac{\rho}{|\rho|} \cdot \arcsen \left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) \tag{g.31}$$

$$\begin{aligned}
 T_2 = & \frac{1}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \left[\sigma_{y_1}^2 \sigma_{y_2}^2 \cdot \frac{\rho^3}{|\rho|} \cdot \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - 1 \right]^{1/2} \right. \\
 & \left. - (\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2) \cdot \frac{\rho}{|\rho|} \cdot \arcsen \left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) \right] \tag{g.32}
 \end{aligned}$$

$$T_3 = \frac{\sigma^6}{\sigma_{y_2}^2} \cdot \frac{\rho}{|\rho|^3} \cdot \frac{\left(1 + \frac{\sigma_{y_2}^2}{\sigma^2}\right)}{\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - 1 \right]^{1/2}} \quad (\text{g.33})$$

e

$$T_4 = \sigma^4 \cdot \frac{\rho}{|\rho|} \cdot \arcsen \left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) - \frac{\sigma^4}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - 1 \right]^{1/2}} \cdot \frac{\rho}{|\rho|} \quad (\text{g.34})$$

Substituindo-se os resultados (g.31), (g.32), (g.33) e (g.34) em (g.30) e sabendo-se que $\text{ sinal}(\rho) = \rho/|\rho|$, então

$$E\{g(y_1)g(y_2)y_1^2\} = \left[2\sigma_{y_1}^2 \sigma^2 - \frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} + \sigma^4 \right] \cdot \arcsen \left(\frac{|\rho \sigma_{y_1} \sigma_{y_2}|}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) \cdot \text{ sinal}(\rho) + \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \left(\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - 1 \right) - \frac{\sigma^4 (\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} + \frac{\sigma^6 \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2}\right)}{\sigma_{y_2}^2 \rho^2} \right] \cdot \frac{\text{ sinal}(\rho)}{\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} \left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\sigma_{y_1}^2 \sigma_{y_2}^2 \rho^2} - 1 \right]^{1/2}} \quad (\text{g.35})$$

onde

$$\begin{cases} \rho = E\{z_1 z_2\} = \frac{1}{\sigma_{y_1} \sigma_{y_2}} E\{y_1 y_2\} \\ \sigma_{y_1}^2 = E\{y_1^2\} \\ \sigma_{y_2}^2 = E\{y_2^2\} \end{cases} \quad (\text{g.36})$$

Substituindo-se as relações (g.36) em (g.35) e levando-se em consideração que o arco-seno é uma função ímpar, chega-se finalmente a

$$\begin{aligned} E\{g(y_1)g(y_2)y_1^2\} &= \sigma^2 E\{y_1^2\} \arcsen \left(\frac{E\{y_1 y_2\}}{(E\{y_1^2\} + \sigma^2)^{1/2} (E\{y_2^2\} + \sigma^2)^{1/2}} \right) \\ &+ \frac{\sigma^2 E\{y_1^2\} E\{y_1 y_2\} (E\{y_1^2\} + \sigma^2) (E\{y_2^2\} + \sigma^2)}{(E\{y_1^2\} + \sigma^2) (E\{y_2^2\} + \sigma^2) \sqrt{(E\{y_1^2\} + \sigma^2) (E\{y_2^2\} + \sigma^2) - E\{y_1 y_2\}^2}} \\ &+ \frac{\sigma^2 E\{y_1^2\} E\{y_1 y_2\} [-E\{y_1 y_2\}^2 + \sigma^2 (E\{y_2^2\} + \sigma^2)] - \sigma^4 E\{y_1 y_2\}^3}{(E\{y_1^2\} + \sigma^2) (E\{y_2^2\} + \sigma^2) \sqrt{(E\{y_1^2\} + \sigma^2) (E\{y_2^2\} + \sigma^2) - E\{y_1 y_2\}^2}} \end{aligned} \quad (\text{g.37})$$

onde σ^2 é o coeficiente de não-linearidade da função $g(\cdot)$.

ANEXO 8 – DETERMINAÇÃO DE $E\{g(y_1)g(y_2)y_1y_2\}$

Este apêndice refere-se ao cálculo da expressão $E\{g(y_1)g(y_2)y_1y_2\}$ utilizando-se o Teorema de Price Modificado (PAWULA, 1967) onde

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \quad (\text{h.1})$$

Supõe-se que y_1 e y_2 sejam variáveis aleatórias gaussianas correlacionadas de média zero.

O Teorema de Price Modificado exige que as variáveis envolvidas possuam variância unitária. Em decorrência disso, é necessária a utilização de variáveis normalizadas z_1 e z_2 relacionadas à y_1 e y_2 da seguinte forma:

$$\begin{cases} y_1 = \sigma_{y_1} z_1 \\ y_2 = \sigma_{y_2} z_2 \end{cases} \quad (\text{h.2})$$

onde z_1 e z_2 possuem variâncias unitárias, e σ_{y_1} e σ_{y_2} são os desvios padrões das variáveis y_1 e y_2 , respectivamente.

O Teorema de Price Modificado consiste em uma seqüência de quatro passos, apresentados a seguir.

1. Determinação de $G(z)$

Substituindo as variáveis normalizadas (h.2) em $E\{g(y_1)g(y_2)y_1y_2\}$, tem-se

$$E\{g(y_1)g(y_2)y_1y_2\} = E\{g(\sigma_{y_1} z_1)g(\sigma_{y_2} z_2)\sigma_{y_1} z_1 \sigma_{y_2} z_2\} \quad (\text{h.3})$$

A partir de (3.3) podemos definir a função $G(z_1, z_2)$, como

$$G(z_1, z_2) = \sigma_{y_1} \sigma_{y_2} z_1 z_2 g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) \quad (\text{h.4})$$

2. Determinação de $h[G(z)]$

O termo $h[G(z_1, z_2)]$ é obtido através de

$$h[G(z_1, z_2)] = \sum_{i < j} \rho_{i,j} \frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} \quad (\text{h.5})$$

Para avaliar (h.5) precisamos calcular a derivada parcial de (h.4)

$$\begin{aligned} \frac{\partial^2 G(z_1, z_2)}{\partial z_1 \partial z_2} &= \sigma_{y_1} \sigma_{y_2} g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) + \sigma_{y_1}^2 \sigma_{y_2}^2 z_1 z_2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} \\ &+ \sigma_{y_1} \sigma_{y_2}^2 z_2 e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} g(\sigma_{y_1} z_1) + \sigma_{y_1}^2 \sigma_{y_2} z_1 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} g(\sigma_{y_2} z_2) \end{aligned} \quad (\text{h.6})$$

onde a derivada de $g(y)$ é

$$g'(y) = e^{-\frac{y^2}{2\sigma^2}} dy \quad (\text{h.7})$$

Substituindo (h.6) em (h.5), tem-se

$$\begin{aligned} h[G(z_1, z_2)] &= \rho \sigma_{y_1} \sigma_{y_2} g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) + \rho \sigma_{y_1}^2 \sigma_{y_2}^2 z_1 z_2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} \\ &+ \rho \sigma_{y_1} \sigma_{y_2}^2 z_2 e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} g(\sigma_{y_1} z_1) + \rho \sigma_{y_1}^2 \sigma_{y_2} z_1 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} g(\sigma_{y_2} z_2) \end{aligned} \quad (\text{h.8})$$

onde $\rho = E\{z_1 z_2\}$ é a covariância entre z_1 e z_2 .

3. Determinação de $E_{\alpha} \{h [G(\mathbf{z})]\}$

Calculando-se o valor esperado de $h [G(z_1, z_2)]$ para $E\{z_1 z_2\} = \alpha\rho$, tem-se

$$\begin{aligned}
 E_{\alpha} \{h [G(z_1, z_2)]\} &= T_1 + T_2 + T_3 + T_4 \\
 &= \rho\sigma_{y_1}\sigma_{y_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) f_{\mathbf{z}}(\mathbf{z}) dz_1 dz_2 \\
 &+ \rho\sigma_{y_1}^2\sigma_{y_2}^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 z_2 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} f_{\mathbf{z}}(\mathbf{z}) dz_1 dz_2 \\
 &+ \rho\sigma_{y_1}\sigma_{y_2}^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_2 e^{-\frac{\sigma_{y_2}^2 z_2^2}{2\sigma^2}} g(\sigma_{y_1} z_1) f_{\mathbf{z}}(\mathbf{z}) dz_1 dz_2 \\
 &+ \rho\sigma_{y_1}^2\sigma_{y_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_1 e^{-\frac{\sigma_{y_1}^2 z_1^2}{2\sigma^2}} g(\sigma_{y_2} z_2) f_{\mathbf{z}}(\mathbf{z}) dz_1 dz_2
 \end{aligned} \tag{h.9}$$

onde a função densidade de probabilidade conjunta $f_{\mathbf{z}}(\mathbf{z})$ é dada por (MILLER, 1964)

$$f_{\mathbf{z}}(\mathbf{z}) = \frac{1}{2\pi |\mathbf{K}_{\alpha}|^{1/2}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{K}_{\alpha}^{-1} \mathbf{z}} \tag{h.10}$$

onde $\mathbf{z} = [z_1 \quad z_2]^T$ e:

$$\mathbf{K}_{\alpha} = \begin{bmatrix} 1 & \alpha\rho \\ \alpha\rho & 1 \end{bmatrix} \tag{h.11}$$

é a matriz de correlação entre z_1 e z_2 .

Analisando o primeiro termo de (h.9), podemos expressá-lo da seguinte forma:

$$T_1 = \rho\sigma_{y_1}\sigma_{y_2} E \left\{ g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) \right\} \Big|_{\mathbf{K}_{\alpha}} \tag{h.12}$$

onde $E\{\bullet\}_{\mathbf{K}_\alpha}$ significa a operação de valor esperado do argumento em termos da função densidade de probabilidade descrita por \mathbf{K}_α .

O valor esperado em (h.12) já foi avaliado no ANEXO 2:

$$E\{g(y_1)g(y_2)\} = \sigma^2 \arcsen \left(\frac{E\{y_1 y_2\}}{\sqrt{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2)}} \right) \quad (\text{h.13})$$

Substituindo (h.13) em (h.12), tem-se

$$T_1 = \rho \sigma_{y_1} \sigma_{y_2} \sigma^2 \arcsen \left(\frac{\sigma_{y_1} \sigma_{y_2} E\{z_1 z_2\}_{\mathbf{K}_\alpha}}{(\sigma_{y_1}^2 E\{z_1^2\}_{\mathbf{K}_\alpha} + \sigma^2)^{1/2} (\sigma_{y_2}^2 E\{z_2^2\}_{\mathbf{K}_\alpha} + \sigma^2)^{1/2}} \right) \quad (\text{h.14})$$

As correlações presentes em (h.14) podem ser obtidas através de (h.11), de tal modo que:

$$T_1 = \rho \sigma_{y_1} \sigma_{y_2} \sigma^2 \arcsen \left(\frac{\rho \sigma_{y_1} \sigma_{y_2} \alpha}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) \quad (\text{h.15})$$

A segunda integral de (h.9) pode ser expressa da seguinte forma:

$$T_2 = \rho \sigma_{y_1}^2 \sigma_{y_2}^2 \frac{|\mathbf{Q}_\alpha|^{1/2}}{|\mathbf{K}_\alpha|^{1/2}} E\{z_1 z_2\}_{\mathbf{Q}_\alpha} \quad (\text{h.16})$$

onde $E\{\bullet\}_{\mathbf{Q}_\alpha}$ significa a operação de valor esperado do argumento em termos da função densidade de probabilidade descrita por

$$f_z(\mathbf{z}) = \frac{1}{2\pi |\mathbf{Q}_\alpha|^{1/2}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{Q}_\alpha^{-1} \mathbf{z}} \quad (\text{h.17})$$

onde

$$\left\{ \begin{aligned}
 \mathbf{Q}_\alpha &= \frac{|\mathbf{K}_\alpha|}{\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2} \begin{bmatrix} 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| & \alpha\rho \\ \alpha\rho & 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \\
 \mathbf{Q}_\alpha^{-1} &= \frac{1}{|\mathbf{K}_\alpha|} \begin{bmatrix} 1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha| & -\alpha\rho \\ -\alpha\rho & 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha| \end{bmatrix} \\
 |\mathbf{Q}_\alpha| &= \frac{|\mathbf{K}_\alpha|^2}{\left(1 + \frac{\sigma_{y_1}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) \left(1 + \frac{\sigma_{y_2}^2}{\sigma^2} |\mathbf{K}_\alpha|\right) - \alpha^2 \rho^2}
 \end{aligned} \right. \tag{h.18}$$

O valor esperado em (h.16) pode ser interpretado como a correlação entre as variáveis z_1 e z_2 segundo a função densidade de probabilidade dada por (h.17). Essa informação pode ser obtida diretamente da matriz de correlação \mathbf{Q}_α . Substituindo-se esse resultado em (h.16) chega-se a

$$T_2 = \frac{\sigma^6}{\sigma_{y_1} \sigma_{y_2}} \cdot \frac{\alpha}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{3/2}} \cdot \frac{1}{|\rho|} \tag{h.19}$$

onde $|\rho| = \sqrt{\rho^2}$.

O terceiro termo de (h.9) pode ser expresso como

$$T_3 = \rho \sigma_{y_1} \sigma_{y_2} \frac{|\mathbf{L}_\alpha|^{1/2}}{|\mathbf{K}_\alpha|^{1/2}} E \left\{ g(\sigma_{y_1} z_1) \sigma_{y_2} z_2 \right\} \Big|_{\mathbf{L}_\alpha} \tag{h.20}$$

onde $E\{\cdot\}_{|L_\alpha}$ significa a operação de valor esperado do argumento em termos da função densidade de probabilidade descrita por

$$f_z(\mathbf{z}) = \frac{1}{2\pi|L_\alpha|^{1/2}} e^{-\frac{1}{2}\mathbf{z}^T L_\alpha^{-1} \mathbf{z}} \tag{h.21}$$

onde

$$\begin{cases} L_\alpha = \frac{1}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \begin{bmatrix} 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |K_\alpha| & \alpha\rho \\ \alpha\rho & 1 \end{bmatrix} \\ L_\alpha^{-1} = \frac{1}{|K_\alpha|} \begin{bmatrix} 1 & -\alpha\rho \\ -\alpha\rho & 1 + \frac{\sigma_{y_2}^2}{\sigma^2} |K_\alpha| \end{bmatrix} \\ |L_\alpha| = \frac{|K_\alpha|}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \end{cases} \tag{h.22}$$

O valor esperado em (h.20) já foi avaliado no ANEXO 2 e é expresso por:

$$E\{g(y_1) y_2\} = \frac{E\{y_1 y_2\}}{\left(\frac{1}{\sigma^2} E\{y_1^2\} + 1\right)^{1/2}} \tag{h.23}$$

Substituindo-se (h.23) em (h.20), chega-se a:

$$T_3 = \rho \sigma_{y_1} \sigma_{y_2} \frac{|L_\alpha|^{1/2} \sigma_{y_1} \sigma_{y_2} E\{z_1 z_2\}_{|L_\alpha}}{|K_\alpha|^{1/2} \left(\frac{\sigma_{y_1}^2}{\sigma^2} E\{z_1^2\}_{|L_\alpha} + 1\right)^{1/2}} \tag{h.24}$$

As esperanças em (h.24) podem ser determinadas por simples inspeção da matriz L_α . Substituindo-as em (h.24), chega-se a

$$T_3 = \frac{\sigma_{y_1} \sigma_{y_2} \sigma^2}{1 + \frac{\sigma_{y_2}^2}{\sigma^2}} \cdot \frac{\alpha}{\left[\frac{1}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} (\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2) - \alpha^2 \right]^{1/2}} \cdot |\rho| \quad (\text{h.25})$$

O método de resolução para o quarto termo de (h.9) é idêntico ao utilizado para a obtenção de (h.25), portanto

$$T_4 = \frac{\sigma_{y_1} \sigma_{y_2} \sigma^2}{1 + \frac{\sigma_{y_1}^2}{\sigma^2}} \cdot \frac{\alpha}{\left[\frac{1}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} (\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2) - \alpha^2 \right]^{1/2}} \cdot |\rho| \quad (\text{h.26})$$

Substituindo os resultados de T_1 , T_2 , T_3 e T_4 em (h.9), obtém-se

$$\begin{aligned} E_\alpha \{ h [G(z_1, z_2)] \} &= \rho \sigma_{y_1} \sigma_{y_2} \sigma^2 \arcsen \left(\frac{\rho \sigma_{y_1} \sigma_{y_2} \alpha}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) \\ &+ \frac{\sigma^6}{\sigma_{y_1} \sigma_{y_2}} \cdot \frac{\alpha}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{3/2}} \cdot \frac{1}{|\rho|} \\ &+ \sigma_{y_1} \sigma_{y_2} \sigma^4 \frac{\sigma_{y_1}^2 + 2\sigma^2 + \sigma_{y_2}^2}{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)} \cdot \frac{\alpha}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{1/2}} \cdot |\rho| \end{aligned} \quad (\text{h.27})$$

4. Determinação de R_1

Segundo PAWULA (1967),

$$E \{ g(\sigma_{y_1} z_1) g(\sigma_{y_2} z_2) \sigma_{y_1} z_1 \sigma_{y_2} z_2 \} = R_1 \quad (\text{h.28})$$

e R_1 pode ser calculado através de

$$R_1 = E_o \{G(z_1, z_2)\} + \int_0^1 E_\alpha \{h[G(z_1, z_2)]\} d\alpha \quad (\text{h.29})$$

Para $\alpha = 0$, z_1 e z_2 são descorrelacionados, portanto

$$E_o \{G(z_1, z_2)\} = E \{g(\sigma_{y_1} z_1) \sigma_{y_1} z_1\} E \{g(\sigma_{y_2} z_2) \sigma_{y_2} z_2\} \quad (\text{h.30})$$

Segundo o ANEXO 2, para $y_1 = y_2 = y$, tem-se que

$$E \{g(y) y\} = \frac{E \{y^2\}}{\left(\frac{1}{\sigma^2} E \{y^2\} + 1\right)^{1/2}} \quad (\text{h.31})$$

e portanto:

$$E_o \{G(z_1, z_2)\} = \frac{\sigma_{y_1}^2 \sigma_{y_2}^2}{\left(\frac{1}{\sigma^2} \sigma_{y_1}^2 + 1\right)^{1/2} \left(\frac{1}{\sigma^2} \sigma_{y_2}^2 + 1\right)^{1/2}} \quad (\text{h.32})$$

Substituindo-se (h.32) e (h.27) em (h.29), chega-se a

$$\begin{aligned} R_1 &= \frac{\sigma^2 \sigma_{y_1}^2 \sigma_{y_2}^2}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \\ &+ \rho \sigma_{y_1} \sigma_{y_2} \sigma^2 \int_0^1 \arcsen \left(\frac{\rho \sigma_{y_1} \sigma_{y_2} \alpha}{(\sigma_{y_1}^2 + \sigma^2)^{1/2} (\sigma_{y_2}^2 + \sigma^2)^{1/2}} \right) d\alpha \\ &+ \frac{\sigma^6}{|\rho| \sigma_{y_1} \sigma_{y_2}} \int_0^1 \frac{\alpha}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{3/2}} d\alpha \\ &+ |\rho| \sigma_{y_1} \sigma_{y_2} \sigma^4 \frac{\sigma_{y_1}^2 + 2\sigma^2 + \sigma_{y_2}^2}{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)} \int_0^1 \frac{\alpha}{\left[\frac{(\sigma_{y_1}^2 + \sigma^2)(\sigma_{y_2}^2 + \sigma^2)}{\rho^2 \sigma_{y_1}^2 \sigma_{y_2}^2} - \alpha^2 \right]^{1/2}} d\alpha \end{aligned} \quad (\text{h.33})$$

Resolvendo as integrais em (h.33) e sabendo-se que

$$\begin{cases} \rho = E\{z_1 z_2\} = \frac{1}{\sigma_{y_1} \sigma_{y_2}} E\{y_1 y_2\} \\ \sigma_{y_1}^2 = E\{y_1^2\} \\ \sigma_{y_2}^2 = E\{y_2^2\} \end{cases} \quad (\text{h.34})$$

chega-se finalmente a

$$\begin{aligned} E\{g(y_1)g(y_2)y_1 y_2\} &= \sigma^2 E\{y_1 y_2\} \arcsen \left(\frac{E\{y_1 y_2\}}{\sqrt{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2)}} \right) \\ &+ \frac{\sigma^2 E\{y_1^2\} E\{y_2^2\} [(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2) - E\{y_1 y_2\}^2] + \sigma^6 E\{y_1 y_2\}^2}{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2) \sqrt{(E\{y_1^2\} + \sigma^2)(E\{y_2^2\} + \sigma^2) - E\{y_1 y_2\}^2}} \end{aligned} \quad (\text{h.35})$$

ANEXO 9 – DETERMINAÇÃO DE η^2

Neste anexo, será apresentado um procedimento para a estimação do grau de não-linearidade em um sistema real, supondo-se que o caminho secundário seja modelado por um sistema Wiener-Hammerstein.

1. Definição do Processo de Estimação

Para estimar-se o grau de não-linearidade, pode-se utilizar o seguinte artifício:

$$\eta^2 = \frac{E\{\hat{y}_f^2(n)\}}{E\{\hat{y}_f^2(n)\} + \hat{\sigma}^2} \tag{i.1}$$

onde $E\{\hat{y}_f^2(n)\}$ é a potência estimada na entrada da não-linearidade (figura i.1), e $\hat{\sigma}^2$ é a estimativa do parâmetro que descreve a função não-linear.

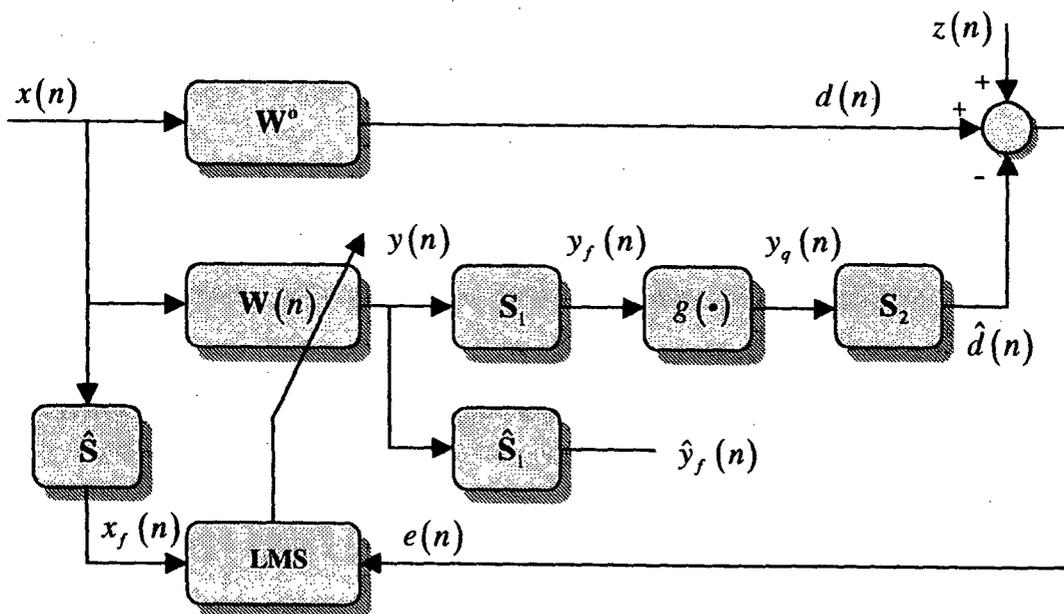


Figura i.1 - Estimação do grau de não-linearidade

Note-se que o sistema S_1 caracteriza apenas o sistema elétrico (amplificador de potência, parte linear do alto-falante, conversor D/A) e portanto podemos supor que o mesmo é constante durante todo o processo de adaptação. Em um sistema em que as propriedades do meio de propagação variem instante a instante, pode-se supor que apenas o sistema S_2 (atrasos de propagação no meio, microfone, filtros analógicos, conversor A/D) varia. Portanto, mesmo em implementações onde é realizada a estimação do caminho acústico secundário em tempo real podemos supor que S_1 e sua estimativa sejam aproximadamente constantes.

Se $\hat{S}_1 \cong S$ e $\hat{\sigma}^2 \cong \sigma^2$ então podemos dizer que é possível realizar a estimação do grau de não-linearidade em regime permanente.

O sistema S_1 pode ser identificado em um procedimento *a-priori* utilizando-se um sinal branco de baixa potência (assim a influência da não-linearidade será desprezível) com o sensor de erro posicionado próximo à saída do atuador (alto-falante, transdutor piezoelétrico). O parâmetro σ^2 também pode ser identificado *a-priori* verificando-se a máxima excursão possível na saída do transdutor.

2. Verificação

Segundo a figura i.1, em regime permanente o sinal de controle na entrada da função não-linear possui o seguinte equacionamento:

$$y_f(n) = \sum_{j=0}^{M_f-1} s_{1j} \mathbf{W}_{RP}^T \mathbf{X}(n-j) \quad (\text{i.2})$$

Elevando-se ao quadrado e tomando-se seu valor esperado, chega-se a

$$E\{y_f^2(n)\} = \mathbf{W}_{RP}^T \mathbf{R}_{S_1 S_1} \mathbf{W}_{RP} \quad (\text{i.3})$$

onde:

$$\mathbf{R}_{S_1 S_1} = \sum_{j=0}^{M_1-1} \sum_{i=0}^{M_1-1} s_{1j} s_{1i} \mathbf{R}_{i-j} \quad (\text{i.4})$$

A equação (4.24) que descreve os coeficientes em regime permanente é dada por:

$$\mathbf{W}_{RP} = \frac{1}{\sqrt{1-\eta^2}} \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (\text{i.5})$$

Sabendo-se que:

$$\eta^2 = \frac{1}{\sigma^2} \mathbf{W}^{oT} \tilde{\mathbf{R}}_{\hat{S}}^T \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-T} \mathbf{R}_{S_1 S_1} \tilde{\mathbf{R}}_{S_2 \hat{S}_1}^{-1} \tilde{\mathbf{R}}_{\hat{S}} \mathbf{W}^o \quad (\text{i.6})$$

e substituindo-se (i.5) e (i.6) em (i.3), chega-se a

$$\eta^2 = \frac{E\{y_f^2(n)\}}{E\{y_f^2(n)\} + \sigma^2} \quad (\text{i.7})$$

A partir de (i.7) pode-se constatar que necessariamente o grau de não-linearidade deve ser menor que a unidade na existência de convergência dos coeficientes.

ANEXO 10 - PUBLICAÇÕES

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 1999a. Statistical Analysis of the LMS Algorithm with a Zero Memory Nonlinearity After the Adaptive Filter. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING – ICASSP (Mar. 1999: Phoenix, EUA). *Proceedings*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 1999b. Statistical Analysis of the FXLMS Algorithm with a Nonlinearity in the Secondary Path. In: IEEE INTERNATIONAL CONFERENCE ON CIRCUITS AND SYSTEMS – ISCAS (May 1999: Orlando, EUA). *Proceedings*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 2000a. The Performance Surface in Nonlinear Mean Square Estimation: Application to the Active Noise Control Problem. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING – ICASSP (June 2000: Istambul, Turquia). *Proceedings*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 2000b. Nonlinear Secondary-Path Effects on the Transient Behavior of The Multiple-Error FXLMS Algorithm. In: IEEE INTERNATIONAL CONFERENCE ON CIRCUITS AND SYSTEMS – ISCAS (May 2000: Genebra, Suíça). *Proceedings*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 2000c. The Performance Surface in Nonlinear Mean Square Estimation: Application to Active Noise Control Problems with Correlated Signals. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA – CBA (13.: set. 2000 : Florianópolis, Santa Catarina). *Anais*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; 2001. A True Stochastic Gradient Adaptive Algorithm for Applications Using Nonlinear Actuators. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING – ICASSP (May 2001: Salt Lake City, EUA). *Proceedings*. CD-ROM.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.. The Performance Surface in Nonlinear Mean Square Estimation: Application to Active Noise Control Problems with Correlated Signals. *SBA Controle & Automação*, artigo convidado.

COSTA, M.H.; BERMUDEZ, J.C.M.; BERSHAD, N.J.. Stochastic Analysis of the LMS Algorithm with a Saturation Nonlinearity Following the Adaptive Filter Output. *IEEE Transactions on Signal Processing*, aceito para publicação.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIRRE, L.A.; 2000. *Introdução à Identificação de Sistemas: Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Belo Horizonte: Editora UFMG.
- ALEXANDER, S.T.; 1987. Transient Weight Misadjustment Properties for the Finite Precision LMS Algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-35, n. 9 (Sept.), p. 1250-1258.
- BAEK, K.H.; ELLIOT, S.J.; 1995. Natural Algorithms for Choosing Source Locations in Active Control Systems. *Journal of Sound and Vibration*, v. 186, n. 2, p. 245-267.
- BENDAT J.S.; PIERSOL, A.G.; 1986. *Random Data: Analysis and Measurement Procedures*. 2.ed. John Wiley & Sons.
- BERMUDEZ, J.C.M; BERSHAD, N.J.; 1996a. A Nonlinear Analytical Model for the Quantized LMS Algorithm - The Arbitrary Step Size Case. *IEEE Transactions on Signal Processing*, v. 44, n. 5 (May), p. 1175-1183.
- _____; _____. 1996b. Transient and Tracking Performance Analysis of the Quantized LMS Algorithm for Time-Varying System Identification. *IEEE Transactions on Signal Processing*, v. 44, n. 8 (Aug.), p. 1990-1997.
- BERNHARD, R.J.; 1995. Active Control of Road Noise Inside Automobiles. In: ACTIVE 95 (July 1995: Newport Beach, USA). *Proceedings*. USA. p. 21-32.
- _____; DAVIES, P.; KURTH, S.W.; 1997. Effects of Nonlinearities on System Identification in Active Noise Control Systems. In: NOISE-CON 97 (June 1997: Pennsylvania, USA). *Proceedings*, USA, p. 231-236.
- BERSHAD, N.J; 1986. On the Optimum Data Nonlinearity in LMS Adaptation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-34, n. 1 (Feb.), p. 69-76.
- _____; 1988. On Error-Saturation in LMS Adaptation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 36, n. 4 (Apr.), p. 440-452.
- _____; 1990. On Weight Update Saturation Nonlinearities in LMS Adaptation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 38 (Apr.), p. 623-630.
- _____, N.J.; SHYNK, J.J.; FEINTUCH, P.L.; 1993a. Statistical Analysis of the Single-Layer Backpropagation Algorithm: Part I - Mean Weight Behavior. *IEEE Transactions on Signal Processing*, v. 41, n. 2 (Feb.), p. 573-582.

- _____; _____; _____; 1993b. Statistical Analysis of the Single-Layer Backpropagation Algorithm: Part II - MSE and Classification Performance. *IEEE Transactions on Signal Processing*, v. 41, n. 2 (Feb.), p. 583-591.
- _____; CUBAUD, N.; SHYNK, J.J.; 1996. Stochastic Convergence Analysis of the Single-Layer Backpropagation Algorithm for Noisy Input Data. *IEEE Transactions on Signal Processing*, v. 44, n. 5 (May), p. 1315-1319.
- _____; BERMUDEZ, J.C.M.; 1996a. New Insights on the Transient and Steady-State Behavior of the Quantized LMS Algorithm. *IEEE Transactions on Signal Processing*, v. 44, n. 10 (Oct.), p. 2623-2625.
- _____; _____; 1996b. A Nonlinear Analytical Model for the Quantized LMS Algorithm – The Power-of-Two Step Size Case. *IEEE Transactions on Signal Processing*, v. 44, n. 11 (Nov.), p. 1-5.
- _____; IBNKAHLA, M.; CASTANIÉ, F.; 1997. Statistical Analysis of a Two-Layer Backpropagation Algorithm Used for Modeling Nonlinear Memoryless Channel. *IEEE Transactions on Signal Processing*, v. 45, n. 3 (Mar.), p. 747-756.
- _____; CELKA, P.; VESIN, J.M.; 1999. Stochastic Analysis of Gradient Adaptive Identification of Nonlinear Systems with Memory for Gaussian data and Noisy Input and Output Measurements. *IEEE Transactions on Signal Processing*, v. 47, n. 3 (Mar.), p. 675-689.
- _____; BOUCHIRED, S.; CASTANIE, F.; 2000. Stochastic Analysis of Adaptive Gradient Identification of Wiener-Hammerstein Systems for Gaussian Inputs. *IEEE Transactions on Signal Processing*, v. 48, n. 2 (Feb.), p. 557-560.
- BJARNASON, E.; 1995. Analysis of the Filtered-X LMS Algorithm. *IEEE Transactions on Speech and Audio Processing*, n. 3, v. 6, p. 504-514.
- BULLMORE, A.J.; NELSON, P.A.; CURTIS, A.R.D. et al.; 1987. The Active Minimization of Harmonic Enclosed Sound Fields, Part II: A Computer Simulation. *Journal of Sound and Vibration*, v. 117, n. 1, p. 15-33.
- BURGESS, J.C.; 1981. Active Adaptive Sound Control in a Duct: A Computer Simulation. *Journal of the Acoustical Society of America*, v. 70, n. 3 (Sept.), p. 715-726.
- BUSSGANG, J.J.; 1952. Cross Correlation Functions of Amplitude Distorted Gaussian Signals. Technical Report 216. Res. Lab. Electron., M.I.T., Cambridge.
- CARAISCOS, C.; LIU, B.; 1984. A Roundoff Error Analysis of the LMS Adaptive Algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-32 (Feb.), n. 1, p. 34-41.

- CHEN, G.; SONE, T.; ABE, M.; 1996. Effects of Multiple Secondary Paths on Convergence Properties in Active Noise Control Systems with LMS Algorithm. *Journal of Sound and Vibration*, v. 195, n. 1, p. 217-228.
- CHEN, C.K.; CHIUUEH T.D.; CHEN, J.H.; 1999. Active Cancellation System of Acoustic Noise in MR Imaging. *IEEE Transactions on Biomedical Engineering*, v. 46, n. 2 (Feb.), p. 186-191.
- DOUGLAS, S.C.; MENG, T.H.-Y.; 1994a. Stochastic Gradient Adaptation under General Error Criteria. *IEEE Transactions on Signal Processing*, v. 42, n. 6 (June), p. 1335-1351.
- _____; _____; 1994b. Normalized Data Nonlinearities for LMS Adaptation. *IEEE Transactions on Signal Processing*, v. 42, n. 6 (June), p. 1353-1365.
- _____; 1999. Fast Implementations of the Filtered-X LMS and LMS Algorithms for Multichannel Active Noise Control. *IEEE Transactions on Speech and Audio Processing*, v. 7, n. 4 (July), p. 454-465.
- DUTTWEILER, D.L.; 1982. Adaptive Filter with Nonlinearities in the Correlation Multiplier. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-30, n. 4 (Aug.), p. 578-586.
- ELLIOT, S.J.; STOTHERS, I.M.; NELSON, P.A.; 1987a. A Multiple Error LMS Algorithm and Its Application to the Active Control of Sound and Vibration. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-35, n. 10 (Oct.), p. 1423-1434.
- _____; CURTIS, A.R.D.; BULLMORE, A.J. et al.; 1987b. The Active Minimization of Harmonic Enclosed Sound Fields, Part III: Experimental Verification. *Journal of Sound and Vibration*, v. 117, n. 1, p. 35-38.
- _____; STOTHERS, I.M.; NELSON, P.A. et al.; 1988. The Active Control of Engine Noise Inside Cars. In: INTER-NOISE 88. *Proceedings*. p. 987-990.
- _____; BOUCHER, C.C.; _____; 1992. The Behavior of a Multiple Channel Active Control System. *IEEE Transactions on Signal Processing*, v. 40, n. 5 (May), p. 1041-1052.
- _____; NELSON, P.A.; 1993. Active Noise Control. *IEEE Signal Processing Magazine*, (Oct.), p. 12-35.
- ERIKSSON, L.J.; ALLIE, M.C.; GREINER, R.A.; 1987. The Selection and Application of an IIR Adaptive Filter for Use in Active Sound Attenuation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. ASSP-35, n. 4 (Apr.), p. 433-437.

- _____; ZUROSKI, M.T.; 1997. From Passive to Active: A Family of Silencing Possibilities. In: NOISE-CON 97 (June 1997: Pennsylvania, USA). *Proceedings. USA.* p. 325-336.
- FÜLLER, C.R.; FLOTOW, A.H.; 1995. Active Control of Sound and Vibration. *IEEE Control Systems*, (Dec.), p. 9-19.
- GITLIN, R.D.; MAZO, J.E.; TAYLOR, M.G.; 1973. On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters. *IEEE Transactions on Circuit Theory*, v. CT-20 (Mar.), p. 125-136.
- GLENTIS, G.O.; BERBERIDIS, K.; THEORIDIS, S.; 1999. Efficient Least Squares Adaptive Algorithms for FIR Transversal Filtering. *IEEE Signal Processing Magazine*, v. 16, n. 4 (July), p. 13-41.
- GUICKING, D.; 1990. On the Invention of Active Noise Control. *Journal of the Acoustical Society of America*, v. 87, p. 2251-2254.
- GUO, J.; PAN, J.; 1995. Analysis of Active Noise Control in Free Space. In: ACTIVE 95 (July 1995: Newport Beach, USA). *Proceedings. USA.* p. 649-660.
- HANSEN, C.; SIMPSON, M.T.; WANGLER, C.T.; 1996. Application of Genetic Algorithms to Active Noise and Vibration Control. In: INTERNATIONAL CONGRESS ON SOUND AND VIBRATION (4.: 1996: St. Petersburg, Russia). *Proceedings. Russia.* p. 371-388.
- _____; 1997. Active Noise Control – From Laboratory to Industrial Implementation. In: NOISE-CON 97 (June 1997: Pennsylvania, USA). *Proceedings. USA.* p. 3-38.
- HAYKIN, S.; 1991. *Adaptive Filter Theory*. 2.ed. Prentice-Hall.
- _____; 1999. *Neural Networks – A Comprehensive Foundation*. NY: Macmillian.
- HEINLE, F.; RABENSTEIN, R.; STENGER, A.; 1998. A Measurement Method for the Linear and Nonlinear Properties of Electro-Acoustic Transmission Systems. *Signal Processing*, v. 64, n. 1, p. 49-60.
- HORN, R.A.; JOHNSON, C.R.; 1999. *Matrix Analysis*. Cambridge University Press.
- KIM, Y.W.; KIM, K.D.; LEE, Y.H.; 1997. Improvement of Active Noise Control in Vehicle Interior Using RLS Algorithm. In: INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING APPLICATIONS AND TECHNOLOGY. *Proceedings.*
- KLIPPEL, W.; 1990. Dynamic Measurements and Interpretation of the Nonlinear Parameters of Electrodynamical Loudspeakers. *Journal of Audio Engineering Society*, v. 38, n. 12 (Dec.), p. 944-955.

- _____; 1992. Nonlinear Large-Signal Behavior of Electrodynamic Loudspeakers at Low Frequencies. *Journal of Audio Engineering Society*, v. 40, n. 6 (June), p. 483-496.
- KUO S.M.; MORGAN, D.R.; 1996. *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: John Wiley.
- _____; PANAH, I.; CHUNG K.M. et al.; 1996. *Design of Active Noise Control Systems with the TMS320 Family*. Texas Instruments.
- _____; MORGAN D.R.; 1999. Active Noise Control: A Tutorial Review. *Proceedings of IEEE*, v. 87, n. 6 (June), p. 941-973.
- LEITCH, R.R.; TOKHI, M.O.; 1987. Active Noise Control Systems. *IEE Proceedings – Part A*, v. 134, n. 6 (Jun), p. 525-546.
- LIAVAS, A.P.; REGALIA, P.A.; 1998. Acoustic Echo Cancellation: Do IIR Models Offer Better Modeling Capabilities than Their FIR Counterparts. *IEEE Transactions on Signal Processing*, v. 46, n. 9 (Sept.), p. 2499-2504.
- LY, K.; 1997. Simulation of an Active Vibration Control System for Aerospace Application. In: INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING APPLICATIONS AND TECHNOLOGY. *Proceedings*.
- MANOLAKIS, D.G.; INGLE, V.K.; KOGON, S.M.; 2000. *Statistical and Adaptive Signal Processing – Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. McGraw-Hill.
- MASSARANI, P.M.; 1990. *Controle Ativo de Ruído em Dutos*. Rio de Janeiro. Dissertação de Mestrado – COPPE, Universidade Federal do Rio de Janeiro.
- MAZO, J.E.; 1979. On the Independence Theory of Equalizer Convergence. *The Bell System Technical Journal*, v. 58, n. 5 (May), p. 963-993.
- MELO, P.L.; WERNECK, M.M.; GIANELLA-NETO, A.; 1998. Linear Servo-Controlled Pressure Generator for Forced Oscillation Measurements. *Medical & Biological Engineering & Computing*, v. 36, p. 11-16.
- _____; _____.; _____.; 2000. Effect of Generator Nonlinearities on the Accuracy of Respiratory Impedance Measurements by Forced Oscillation. *Medical & Biological Engineering & Computing*, v. 38, p. 102-108.
- MILLER, K.; 1964. *Multidimensional Gaussian Distributions*. John Wiley and Sons.
- MITRA, S.K.; KAISER, J.F.; 1993. *Handbook for Digital Signal Processing*. John Wiley & Sons.

- MORGAN, D.R.; 1991. A Hierarchy of Performance Analysis Techniques for Adaptive Active Control of Sound and Vibration. *Journal of the Acoustical Society of America*, v. 85, n. 5 (May), p. 2362-2369.
- NELSON, P.A.; CURTIS, A.R.D.; ELLIOT, S.J. et al.; 1987. The Active Minimization of Harmonic Enclosed Sound Fields, Part I: Theory. *Journal of Sound and Vibration*, v. 117, n. 1, p. 1-13.
- _____; ELLIOT S.J.; 1995. *Active Control of Sound*. Londres: Academic Press.
- NOLLET, B.S.; JONES, D.L.; 1997. Nonlinear Echo Cancellation for Hands-Free Speakerphones. In: NONLINEAR SIGNAL AND IMAGE PROCESSING (Sept. 1997: Michigan, USA). *Proceedings*. USA.
- OPPENHEIM, A.; SCHAFFER R.W.; 1989. *Discrete Time Signal Processing*. Prentice Hall.
- OSÓRIO, P.O.; NÓBREGA, M.V.; 1995. Controle Ativo de Ruído de Banda Larga em Dutos. *SBA Controle & Automação*, v. 6, n. 2 (maio), p. 70-78.
- PAPOULIS, A.; 1991. *Probability, Random Variables and Stochastic Processes*. 3. ed. McGraw-Hill.
- PAWULA, R.F.; 1967. A Modified Version of Price's Theorem. *IEEE Transactions on Information Theory*, v. IT-13, n. 2 (Apr.), p. 285-288.
- PLA, F.G.; SOMMERFELDT, S.D.; HEEDEN, R.A.; 1995. Active Control of Noise in Magnetic Resonance Imaging. In: ACTIVE 95 (July 1995: Newport Beach, USA). *Proceedings*. USA. p. 573-582.
- PRICE, R.; 1958. A Useful Theorem for Nonlinear Devices Having Gaussian Inputs. *IRE Transactions on Information Theory*, v. IT-4 (June), p. 69-72.
- ROWE, H.E.; 1982. Memoryless Nonlinearities With Gaussian Inputs: Elementary Results. *The Bell System Technical Journal*, v. 61, n. 7 (Sept.), p. 1519-1525.
- SCHETZEN, M.; 1989. *The Volterra and Wiener Theories of Nonlinear Systems*. Krieger Publishing Company.
- SCOTT, J.; KELLY, J.; LEEMBRUGGEN, G.; 1996. New Method of Characterizing Driver Nonlinearity. *Journal of Audio Engineering Society*, v. 44, n. 4 (Apr.), p. 258-265.
- SHYNK, J.J.; BERSHAD, N.J.; 1991. Steady-State Analysis of a Single-Layer Perceptron Based on a System Identification Model with Bias Terms. *IEEE Transactions on Circuits and Systems*, v. 38, n. 9 (Sept.), p.1030-1042.

- SNYDER, S.D.; HANSEN, C.H.; 1992. Design Considerations for Active Noise Control Systems Implementing the Multiple Input, Multiple Output LMS Algorithm. *Journal of Sound and Vibration*, v. 159, n. 1, p. 157-174.
- _____; _____; 1994. The Effect of Transfer Function Estimation Errors on the Filtered-X LMS Algorithm. *IEEE Transactions on Signal Processing*, (Apr.), p. 950-953.
- STENGER, A.; RABENSTEIN, R.; 1998. An Acoustic Echo Canceller with Compensation of Nonlinearities. In: EUROPEAN SIGNAL PROCESSING CONFERENCE (Sept. 1998: Rode Island, Greece). *Proceedings*. Greece, p. 969-972.
- _____; KELLERMAN, W.; 2000. Adaptation of a Memoryless Preprocessor for Nonlinear Acoustic Echo Cancelling. *Signal Processing*, n. 80, p. 1747-1760.
- STEVENS J.C.; AHUJA, K.K.; 1991. Recent Advances in Active Noise Control. *American Institute of Aeronautics and Astronautics Journal*, v. 29, n. 7 (July), p. 1058-1065.
- SUTTON, T.J.; ELLIOT, S.J.; 1995. Active Attenuation of Periodic Vibration in Nonlinear Systems Using an Adaptive Harmonic Controller. *Journal of Vibration and Acoustics*, v. 117 (July), p. 355-362.
- SUYKENS, J.; VANDEWALLE, J.; GINDERDEUREN, J.; 1995. Feedback Linearization in Electrodynamic Loudspeakers. *Journal of the Audio Engineering Society*, v. 43, n. 9, p. 690-694.
- TAO, G.; KOKOTOVIC, P.V.; 1996. *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. John Wiley & Sons.
- THERRIEN, C.W.; 1991. *Discrete Random Signals and Statistical Signal Processing*. Prentice-Hall.
- TOBIAS, O.J.; BERMUDEZ, J.C.M.; BERSHAD, N.J. et al.; 1998. Mean Weight Behavior of the Filtered-X LMS Algorithm. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (Sept. 1998: Seattle, USA). *Proceedings*. USA. p. 3545-3548.
- _____; 1999. *Análise Estatística do Algoritmo LMS Filtrado*, Florianópolis. Tese de Doutorado – PPGEEL, Universidade Federal de Santa Catarina.
- _____; BERMUDEZ, J.C.M.; BERSHAD, N.J.; 2000. Mean Weight Behavior of the Filtered-X LMS Algorithm. *IEEE Transactions on Signal Processing*, v. 48, n. 4 (Apr.), p. 1061-1075.

- TOKHI, M.O.; LEITCH, R.R.; 1991. The Robust Design of Active Noise Control Systems Based on Relative Stability Measures. *Journal of the Acoustical Society of America*, v. 90, n. 1 (July), p. 334-345.
- VIPPERMAN, J.S.; BURDISSO, R.A.; 1995. Adaptive Feedforward Control of Non-Minimum Phase Structural Systems. *Journal of Sound and Vibration*, v. 183, n. 3, p. 369-382.
- WANG, A.K.; REN, W.; 1999. Convergence Analysis of the Multi-Variable Filtered-X LMS Algorithm with Application to Active Noise Control. *IEEE Transactions on Signal Processing*, v. 47, n. 4 (Apr.), p. 1166-1169.
- WIDROW, B.; HOFF JR, M.; 1960. Adaptive Switching Circuits. In: IRE WESCON. *Proceedings*. p. 96-104.
- _____; GLOVER JR, J.R.; MCCOOL, J.M. et al.; 1975. Adaptive Noise Cancelling: Principles and Applications. *Proceedings of IEEE*, v. 63, n. 12 (Dec.), p. 1692-1716.
- _____; SHUR, D.; SHAFFER, S.; 1981. On Adaptive Inverse Control. In: ASILOMAR CONFERENCE (15.). *Proceedings*. p. 185-189.
- _____; STEARNS S.D.; 1985. *Adaptive Signal Processing*. Prentice-Hall.
- WOODARD, S.E.; NAGCHAUDHURI, A.; 1998. Application of Least Mean Square Algorithms to Spacecraft Vibration Compensation. *The Journal of the Astronautical Sciences*, v. 46, n. 1 (Jan.), p. 83-90.
- ZHANG, M.; LAN, H.; SER, W.; 2000. An Improved Secondary Path Modeling Method for Active Noise Control Systems. *IEEE Signal Processing Letters*, v. 7, n. 4 (Apr.), p. 73-75.
- ZWILLINGER, D.; 1996. *Standard Mathematical Tables and Formulae*. 30.ed. CRC.