

RENATO CANDIDO

**COMBINAÇÃO AFIM DE
ALGORITMOS ADAPTATIVOS**

Dissertação apresentada à Escola
Politécnica da Universidade de
São Paulo para obtenção do
Título de Mestre em Engenharia
Elétrica.

São Paulo
2009

RENATO CANDIDO

**COMBINAÇÃO AFIM DE
ALGORITMOS ADAPTATIVOS**

Dissertação apresentada à Escola
Politécnica da Universidade de
São Paulo para obtenção do
Título de Mestre em Engenharia
Elétrica.

Área de Concentração:
Sistemas Eletrônicos

Orientador:
Prof. Dr. Magno T. M. Silva

São Paulo
2009

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 4 de maio de 2009.

Renato Candido

Prof. Dr. Magno Teófilo Madeira da Silva

Candido, Renato

Combinação afim de algoritmos adaptativos / R. Candido. – São Paulo, 2009.
119 p.

Edição Revisada

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo.
Departamento de Engenharia de Sistemas Eletrônicos.

1. Filtros elétricos adaptativos 2. Algoritmos (Combinação) 3. Análise de algoritmos I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II. t.

Agradecimentos

Ao amigo e orientador, Prof. Magno T. M. Silva por toda a atenção dedicada durante os anos do mestrado. As inúmeras reuniões, conversas e idéias foram fundamentais para o desenvolvimento deste trabalho. Além disso, gostaria de agradecer pelos conselhos e oportunidades oferecidas que ajudaram no desenvolvimento de minha carreira.

Ao Prof. Vítor H. Nascimento pelas idéias, sugestões e pelas conversas técnicas muito descontraindo.

Ao Prof. Phillip M. S. Burt e ao Prof. José Carlos M. Bermudez pelas sugestões dadas no exame de qualificação.

Aos demais professores que contribuíram na realização deste trabalho, em especial à Profa. Maria D. S. Miranda e ao Prof. Márcio Eisencraft.

À minha família, em especial aos meus pais Maria e Luiz e à minha irmã Fabiana pelo apoio, motivação, afeto e dedicação incondicional.

À minha namorada Lúcia pelo apoio, afeto, carinho e paciência mesmo nos momentos mais difíceis.

Aos amigos Eduardo Rezende e Ricardo Martins pelo companheirismo e incentivo a iniciar o mestrado.

Aos colegas do Laboratório de Processamento de Sinais, em especial ao amigo João Mendes pelas inúmeras conversas e ótima convivência.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento deste trabalho.

“O bom senso é a coisa mais bem distribuída do mundo: todos pensamos tê-lo em tal medida que até os mais difíceis de se contentar nas outras coisas não costumam desejar mais bom senso do que têm.”

René Descartes

Resumo

A combinação de algoritmos tem despertado interesse para melhorar o desempenho de filtros adaptativos. Esse método consiste em combinar linearmente as saídas de dois filtros operando em paralelo com passos de adaptação diferentes para se obter um filtro com convergência rápida e um erro quadrático médio em excesso (EMSE - *excess mean squared error*) reduzido. Nesse contexto, foi proposta a combinação afim de dois algoritmos LMS (*least-mean square*), cujo parâmetro de mistura não fica restrito ao intervalo $[0, 1]$ e por isso é considerada como uma generalização da combinação convexa. Neste trabalho, a combinação afim de dois algoritmos LMS é estendida para os algoritmos supervisionados NLMS (*normalized LMS*) e RLS (*recursive least squares*) e também para equalização autodidata, usando o CMA (*constant modulus algorithm*). Foi feita uma análise em regime da combinação afim desses algoritmos de forma unificada, considerando entrada branca ou colorida e ambientes estacionários ou não-estacionários. Através dessa análise, verificou-se que a combinação afim de dois algoritmos da mesma família pode apresentar uma redução de EMSE de até 3 dB em relação ao EMSE de seus filtros componentes e conseqüentemente ao EMSE da combinação convexa. Para garantir que a estimativa combinada seja pelo menos tão boa quanto a do melhor filtro componente, foram propostos e analisados três novos algoritmos para adaptação do parâmetro de mistura. Utilizando resultados da análise desses algoritmos em conjunto com os resultados da análise de transitório de filtros adaptativos, analisou-se o comportamento transitório da combinação afim. Através de simulações, observou-se uma boa concordância entre os resultados analíticos e os de simulação. No caso de equalização autodidata, também foi proposta uma combinação de dois equalizadores CMA com inicializações diferentes. Verificou-se através de simulações que em alguns casos a combinação afim é capaz de evitar a convergência para mínimos locais da função custo do módulo constante.

Abstract

In order to improve the performance of adaptive filters, the combination of algorithms is receiving much attention in the literature. This method combines linearly the outputs of two filters operating in parallel with different step-sizes to obtain an adaptive filter with fast convergence and reduced excess mean squared error (EMSE). In this context, it was proposed an affine combination of two least-mean square (LMS) filters, whose mixing parameter is not restricted to the interval $[0, 1]$. Hence, the affine combination is a generalization of the convex combination. In this work, the affine combination of two LMS algorithms is extended to the supervised algorithms NLMS (normalized LMS) and RLS (recursive least squares), and also to blind equalization, using the constant modulus algorithm (CMA). A steady-state analysis of the affine combination of the considered algorithms is presented in a unified manner, assuming white or colored inputs, and stationary or nonstationary environments. Through the analysis, it was observed that the affine combination of two algorithms of the same family can provide a 3 dB EMSE gain in relation to its best component filter and consequently in relation to the convex combination. To ensure that the combined estimate is at least as good as the best of the component filters, three new algorithms to adapt the mixing parameter were proposed and analyzed. Using the analysis results of these algorithms in conjunction with the results of the transient analysis of adaptive filters, the transient behavior of the affine combination was analyzed. Through simulations, a good agreement between analytical and experimental results was always observed. In the blind equalization case, a combination of two CMA equalizers with different initializations was also proposed. The simulation results suggest that the affine combination can avoid local minima of the constant modulus cost function.

Sumário

1	Introdução e formulação do problema	1
1.1	Uma breve introdução à filtragem adaptativa	1
1.2	Combinação de filtros adaptativos	4
1.2.1	Combinação Convexa	7
1.2.2	Combinação afim	11
1.3	Objetivos	13
1.4	Justificativa	13
1.5	Contribuições da dissertação	14
1.6	Organização da dissertação	15
2	Análise em regime das combinações	16
2.1	Funções custo e algoritmos	16
2.2	Modelos utilizados	18
2.3	O parâmetro de mistura ótimo em regime	20
2.4	O EMSE ótimo da combinação em regime	21
2.5	Expressões para o EMSE em regime de filtros adaptativos	22
2.6	Resultados em ambiente estacionário	24
2.6.1	Possibilidade de ganho maior que 3 dB	27
2.7	Resultados em ambiente não-estacionário	29
2.8	Resultados de Simulação	31
2.9	Conclusões	44
3	Análise do transitório das combinações e propostas para adaptação do parâmetro de mistura	47
3.1	O EMSE da combinação	47
3.2	Análise da adaptação do algoritmo η -LMS	49
3.2.1	Momento de primeira ordem	50
3.2.2	Momento de segunda ordem	52
3.3	Expressões para o EMSE de filtros adaptativos	54
3.4	Resultados de Simulação	55
3.5	Propostas para adaptação do parâmetro de mistura	61

3.5.1	O algoritmo <i>signed regressor</i>	62
3.5.2	Algoritmos normalizados	66
3.5.3	Resumo dos algoritmos propostos	73
3.5.4	Resultados de simulação	73
3.6	Conclusões	76
4	Estratégias para inicialização do CMA	78
4.1	Combinações de CMA com diferentes inicializações	78
4.2	A transferência de coeficientes	84
4.3	Comportamento em equalizadores com mais coeficientes	87
4.4	Conclusões	89
5	Conclusões e trabalhos futuros	90
	Referências Bibliográficas	98
	Apêndices	99
A	Utilização do <i>random-walk model</i> em equalização	99
B	Expressão do EMSE cruzado em regime para o NLMS	102
C	Expressão do EMSE cruzado em regime para o RLS	105
C.1	Expressão para o termo $\mathcal{B}_1$	108
D	Expressões da matriz de covariância do vetor de erro dos coeficientes	112
D.1	LMS	114
D.2	NLMS	114
D.3	RLS	116
D.4	CMA	118

Lista de Figuras

1.1	Filtro adaptativo do tipo FIR.	2
1.2	Equalização autodidata.	3
1.3	Combinação de filtros adaptativos transversais para (a) filtragem supervisionada e (b) equalização autodidata.	5
1.4	Exemplo de combinação convexa de dois filtros LMS. Configuração de identificação de sistemas, $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$ e $\mu_2 = 0,01$. Entrada colorida (modelo autoregressivo com pólo em 0,8) e solução ótima dada por um vetor aleatório de média nula e variância unitária. Média de 50 realizações com $\mu_\alpha = 100$, $\alpha^+ = 4$ e $\epsilon_C = 0,1$ usando o parâmetro de mistura universal definido em (1.26).	9
1.5	Comportamento da combinação convexa de dois algoritmos LMS no caso não-estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 2 \times 10^{-2}$, $\mu_2 = 2 \times 10^{-3}$. Entrada colorida (modelo autoregressivo com pólo em 0,8) com potência $1/M$	11
2.1	Interpretação geométrica da combinação afim.	25
2.2	Comparação entre $\zeta(\infty)$ e $\zeta_i(\infty)$, $i = 1, 2$	26
2.3	Esquema da combinação afim de duas combinações afim. (a) Caso supervisionado. (b) Caso autodidata.	27
2.4	Curvas de NSD para as combinações afim (NSD_a) e convexa (NSD_c) de dois algoritmos CMA com passos de adaptação distintos. $M = 4$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Sinal de entrada com constelação $[-3, -1, 1, 3]$, $\bar{\gamma} = 6,8$, $\xi = 190,24$, $\sigma_\beta^2 = 28,8$	31
2.5	Curvas de NSD para as combinações afim (NSD_a) e convexa (NSD_c) de dois algoritmos CMA com passos de adaptação próximos. $M = 4$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,9$. Sinal de entrada com constelação $[-3, -1, 1, 3]$, $\bar{\gamma} = 6,8$, $\xi = 190,24$, $\sigma_\beta^2 = 28,8$	32

- 2.6 Combinação afim no caso supervisionado estacionário. Valores teóricos (linha cheia) e experimentais ($\bigcirc$) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1,2$ para 50 realizações com $M = 7$, $\sigma_v^2 = 0,01$ e diversos valores de δ_{XX} . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_{o,1}(0)$. (a) μ_1 -LMS com μ_2 -LMS, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$, $\mu_\eta = 1$. (b) $\tilde{\mu}_1$ -NLMS com $\tilde{\mu}_2$ -NLMS, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ e $\mu_\eta = 0,1$ para $\delta_{NN} = 0,1$, $\delta_{NN} = 0,3$ e $\delta_{NN} = 0,5$; $\mu_\eta = 0,5$ para $\delta_{NN} = 0,7$ e $\mu_\eta = 3$ para $\delta_{NN} = 0,9$. (c) λ_1 -RLS com λ_2 -RLS, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 1$ 33
- 2.7 Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS com $\tilde{\mu}_1 = 0,1$ e $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ no caso supervisionado estacionário. Valores teóricos (linha cheia) e experimentais ($\bigcirc$) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1,2$ para 50 realizações com $\sigma_v^2 = 0,01$ e diversos valores de δ_{NN} . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_{o,1}(0)$. Em azul, $M = 30$, $\mu_\eta = 6$ e solução ótima formada por 30 coeficientes aleatórios de média 0 e variância 1. Em cinza, resultados obtidos na Figura 2.6. 33
- 2.8 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 2,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 2,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados. 34
- 2.9 Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,5$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,1$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 5000$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 5000$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,5$. Valores teóricos tracejados. 35
- 2.10 Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,1$ e $\epsilon_R = 1 \times 10^{-3}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 1500$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 1500$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados. 35
- 2.11 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,9$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados. 36
- 2.12 Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,5$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,9$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,5$. Valores teóricos tracejados. 36

- 2.13 Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,9$ e $\epsilon_R = 1 \times 10^{-3}$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados. 37
- 2.14 Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$ até $n = 6 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 6 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados. 38
- 2.15 Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,9$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição e canal $\mathbf{h}_1$. Média de 500 realizações com $\mu_\eta = 0,1$. Valores teóricos tracejados. 38
- 2.16 Resultados da combinação afim de dois algoritmos com $\delta_{XX} = 0,999$ e parâmetro $\eta(n)$ constante e igual a $\bar{\eta}_o(\infty)$. Média de 500 realizações. No caso supervisionado, solução ótima dada por $\mathbf{w}_{o,1}(0)$, entrada colorida (modelo AR com pólo em 0,8) com potência $1/M$ e $\sigma_v^2 = 0,01$; No caso autodidata, canal $\mathbf{h}_1$ e sinal de entrada 4-PAM. (a) μ_1 -LMS e μ_2 -LMS com $\mu_1 = 0,05$. (b) $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS com $\tilde{\mu}_1 = 0,5$ e $\epsilon_N = 1 \times 10^{-5}$. (c) λ_1 -RLS e λ_2 -RLS com $\lambda_1 = 0,9$ e $\epsilon_R = 1 \times 10^{-3}$. (d) μ_1 -CMA e μ_2 -CMA com $\mu_1 = 1 \times 10^{-3}$ 39
- 2.17 Coeficiente $c_{k,1}(n)$ do vetor $\mathbf{E}\{\mathbf{V}^T \tilde{\mathbf{w}}_1(n)\}$ com menor velocidade de convergência na combinação de dois algoritmos LMS mostrada na Figura 2.16-(a). 40
- 2.18 Combinação afim das combinações A_1 (μ_1 -CMA e μ_2 -CMA) e A_2 (μ_1 -CMA e μ_2 -CMA) no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC,1}\mu_1$ e $\delta_{CC,1} = 0,95$, $\mu_3 = 0,9405 \times 10^{-3}$, $\mu_4 = \delta_{CC,2}\mu_3$ e $\delta_{CC,2} = 0,95$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição e canal $\mathbf{h}_1$. Média de 500 realizações com $\mu_{\eta,A_1} = \mu_{\eta,A_2} = \mu_\eta = 0,1$. Valores teóricos tracejados. 41
- 2.19 Combinação afim no caso supervisionado não-estacionário. Valores teóricos (linha cheia) e experimentais ($\bigcirc, \square, \nabla$) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1,2$ para 50 realizações com $M = 7$, $\sigma_v^2 = 0,01$ e diversos valores de δ_{XY} e diversos valores de σ_q^2 . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_1(0)$. (a) μ_1 -LMS com μ_2 -LMS, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$, $\mu_\eta = 1$. (b) $\tilde{\mu}_1$ -NLMS com $\tilde{\mu}_2$ -NLMS, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ e $\mu_\eta = 1$. (c) λ_1 -RLS com λ_2 -RLS, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 1$. (d) λ_1 -RLS com μ_2 -LMS, $\lambda_1 = 0,9$, $\mu_2 = \delta_{RL}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 0,5$. 42

- 2.20 Curvas teóricas de NSD da combinação afim de λ_1 -RLS e μ_2 -LMS. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,99$, $\mu_2 = 1 \times 10^{-3}$ e $\mathbf{Q} \propto \mathbf{R}^{-1}$. Entrada colorida (modelo AR com pólo em 0,8). 43
- 2.21 Combinações afim e convexa no caso autodata não-estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$, $\delta_{CC} = 0,1$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$. Valores teóricos (linhas) e experimentais ($\times$) do NSD dos filtros componentes, da combinação afim (NSD_a, $\mu_\eta = 0,0075$) e da combinação convexa (NSD_c, $\mu_\alpha = 0,075$, $\alpha^+ = 4$). Média de 50 realizações. 44
- 3.1 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados. 57
- 3.2 Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,1$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados. 57
- 3.3 Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,99$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,1$ e $\epsilon_R = 0,02$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 2$. Valores teóricos tracejados. 58
- 3.4 Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetor de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$ até $n = 4 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 4 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados. 58
- 3.5 Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetor de coeficientes inicializados em $0,9\mathbf{w}_{zf}$, sendo $\mathbf{w}_{zf}$ assumido como o valor médio dos coeficientes em regime. Canal $\mathbf{h}_1$ até $n = 2,5 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 2,5 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados. 59

- 3.6 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados. 60
- 3.7 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,1$. Valores teóricos tracejados. 61
- 3.8 Comparação entre o valor experimental de $E\{\sqrt{p(n)}\}$ e o valor obtido com a aproximação $\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}$ para diferentes valores de ν , com o sinal $[y_1(n) - y_2(n)]$ gaussiano com variância $\Delta\zeta_1(n) + \Delta\zeta_2(n) = 2 \times 10^{-3}$ 71
- 3.9 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo SR. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_{\eta,s} = 0,025$. Valores teóricos tracejados. 74
- 3.10 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo PN. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\tilde{\mu}_\eta = 0,003$, $\nu = 0,99$ e $\epsilon = 1 \times 10^{-3}$. Valores teóricos tracejados. 75
- 3.11 Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo NSR. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\tilde{\mu}_{\eta,s} = 0,003$, $\nu = 0,99$ e $\epsilon = 0,1$. Valores teóricos tracejados. 75
- 3.12 Comportamento do parâmetro de mistura obtido com os diferentes tipos de adaptação considerados. Comparação entre os resultados mostrados nas Figuras 3.6, 3.9, 3.10 e 3.11. 76
- 4.1 Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 0,2$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ -0,4]$ e $\mathbf{w}_2^T(0) = [1 \ 0]$. Canal IIR definido por (4.1). Média de 10^3 realizações. 81
- 4.2 Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 412 realizações de um total de 10^3 82

- 4.3 Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 588 realizações de um total de 10^3 83
- 4.4 Combinação afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes e usando a transferência de coeficientes dos filtros componentes para a combinação com $\varrho = 0,7$. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 10^3 realizações. 85
- 4.5 Combinação convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes e usando a transferência de coeficientes dos filtros componentes para a combinação com $\varrho = 0,7$. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 10^3 realizações. 86
- 4.6 Combinação afim (PN com $\tilde{\mu}_\eta = 5 \times 10^{-4}$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 11$ e $\mu_1 = \mu_2 = 0,0025$. Filtro i_1 -CMA inicializado com um vetor pino com um elemento unitário na quarta posição e i_2 -CMA inicializado com um vetor pino com um elemento unitário na décima primeira posição. Canal $\mathbf{h}_3$. Média de 10^3 realizações. 88
- 4.7 Combinação afim (PN com $\tilde{\mu}_\eta = 5 \times 10^{-4}$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 4$ com subamostragem de fator $S = 2$ e $\mu_1 = \mu_2 = 0,0016$. Vetor de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0 \ -1 \ -0,5 \ 0]$ e $\mathbf{w}_2^T(0) = [0 \ 0 \ 1 \ 0,8]$. Canal $\mathbf{h}_1$ com SNR de 25 dB. Média de 10^3 realizações. 88

Lista de Tabelas

1.1	Parâmetros da Equação (1.1) para os algoritmos LMS, NLMS, RLS e CMA.	3
2.1	Funções custo e gradientes dos algoritmos utilizados.	17
2.2	Expressões analíticas para o EMSE dos algoritmos considerados.	23
2.3	Expressões analíticas para o EMSE cruzado dos algoritmos considerados.	23
2.4	Expressões analíticas mais precisas para o EMSE do RLS.	24
2.5	Expressões analíticas para $\bar{\eta}_o(\infty)$ e $\zeta(\infty)$ no caso estacionário.	24
2.6	Expressões analíticas para q_{12} e $\zeta(\infty)$ para as situações (i) e (ii) no caso não-estacionário.	30
3.1	Expressões para o cálculo de $\mathbf{S}_{ij}(n)$ no caso estacionário.	55
3.2	Resumo dos algoritmos para adaptação de $\eta(n)$ e respectivos intervalos para escolha do passo de adaptação	73
5.1	Tópicos abordados na dissertação.	91
D.1	Parâmetros das expressões (D.3) e (D.4) para os algoritmos considerados.	112

Lista de Abreviaturas

A seguir são listadas as principais abreviações usadas na dissertação. No caso de siglas consagradas na literatura internacional, optou-se por manter as mesmas em inglês.

AR	Autoregressivo
AWGN	<i>Additive White Gaussian Noise</i> (Ruído aditivo branco e gaussiano)
CMA	<i>Constant Modulus Algorithm</i> (Algoritmo do módulo constante)
CVS-LMS	<i>Combination of Variable Step Least-Mean Square</i>
EMSE	<i>Excess Mean Squared Error</i> (Erro médio quadrático em excesso)
f.d.p.	Função densidade de probabilidade
FIR	<i>Finite Impulse Response</i> (Resposta ao pulso unitário finita)
HOS	<i>Higher Order Statistics</i> (Estatísticas de ordem superior)
i.i.d.	Independente e identicamente distribuído
IIR	<i>Infinite Impulse Response</i> (Resposta ao pulso unitário infinita)
LMS	<i>Least-Mean Square</i>
NLMS	<i>Normalized Least-Mean Square</i>
NSD	<i>Normalized Square Deviation</i>
NSR	<i>Normalized Signed Regressor</i>
PAM	<i>Pulse Amplitude Modulation</i>
PN	<i>Power Normalized</i>
RLS	<i>Recursive Least Squares</i>
SNR	<i>Signal-to-Noise Ratio</i> (Relação sinal-ruído)
SR	<i>Signed Regressor</i>

Lista de Símbolos

Nesta dissertação, matrizes são indicadas por letras maiúsculas em negrito, por exemplo **R**. Vetores coluna são indicados usando-se letras minúsculas em negrito, por exemplo **u**(*n*) e **q**(*n*). Escalares são representados por letras maiúsculas ou minúsculas em itálico, por exemplo *M*, *μ* e *y*(*n*). A seguir são listados os principais símbolos utilizados.

Símbolos gerais

$\mathbb{R}$	conjunto dos números reais
$\mathbb{N}$	conjunto dos números naturais
$(\cdot)^T$	transposição de vetores ou matrizes
$f(\infty)$	limite da função $f(n)$ com $n \rightarrow \infty$
$\hat{}$	valor estimado
$E\{\cdot\}$	operador esperança matemática
z^{-1}	operador de atraso, $z^{-1}y(n) = y(n-1)$
$ x $	módulo do escalar x
$\ \mathbf{x}\ $	norma euclideana ou l_2 do vetor $\mathbf{x}$
$\text{Tr}(\mathbf{A})$	traço (soma dos elementos da diagonal) da matriz $\mathbf{A}$
σ_x^2	variância do sinal x
$\text{sign}(\cdot)$	função sinal
$\text{erf}(\cdot)$	função erro
$\text{sgm}(\cdot)$	função sigmoideal
$\mathbf{I}$	matriz identidade
n	instante de tempo
M	número de coeficientes do filtro
S	taxa de sobreamostragem do equalizador
$X(z)$	transformada Z da sequência $x(n)$
τ_d	atraso em número de amostras

a	sinal transmitido
$r, \bar{\gamma}, \xi$ e σ_β^2	constantes que dependem de estatísticas do sinal transmitido
$\hat{a}$	sinal recuperado após equalização
u	sinal de entrada do filtro
$\mathbf{u}$	vetor regressor de entrada do filtro
$\mathbf{R}$	matriz de autocorrelação do sinal de entrada do filtro
ϑ_i	i -ésimo autovalor da matriz $\mathbf{R}$
$\mathbf{\Lambda}$	matriz diagonal dos autovalores de $\mathbf{R}$
$\mathbf{V}$	matriz de autovetores de $\mathbf{R}$
σ_u^2	potência do sinal de entrada do filtro
α_u	esperança do inverso da norma ao quadrado do vetor $\mathbf{u}$
γ_u	constante que depende de estatísticas do sinal de entrada do filtro
d	sinal desejado
v	ruído de medida
e	sinal de erro
e_d	sinal de erro de decisão
σ_v^2	potência do ruído de medida
$\mathbf{w}$	vetor de coeficientes do filtro
$\mathbf{w}_o$	vetor de coeficientes ótimos
$\hat{\mathbf{w}}_{zf}$	estimativa da solução de <i>zero-forcing</i>
$\mathbf{q}$	perturbação do vetor $\mathbf{w}_o$ num ambiente não estacionário
$\mathbf{Q}$	matriz de autocorrelação de $\mathbf{q}$
$\mathbf{h}$	vetor de coeficientes do canal
J	função custo
ζ_o	EMSE ótimo

Filtros componentes

$\mathbf{w}_i$	vetor de coeficientes do filtro i
$\tilde{\mathbf{w}}_i$	vetor de erro dos coeficientes do filtro i
$\mu_i, \tilde{\mu}_i$	passos de adaptação
λ_i	fator de esquecimento do RLS
$\tilde{\lambda}_i$	complemento de 1 do fator de esquecimento do RLS, $\tilde{\lambda}_i = 1 - \lambda_i$

y_i	sinal de saída do filtro i
e_i	sinal de erro do filtro i
$e_{a,i}$	sinal de erro <i>a priori</i> do filtro i
ζ_i ou ζ_{ii}	EMSE do filtro i
$\zeta_{ij}, i \neq j$	EMSE cruzado entre os filtros i e j
$\Delta\zeta_i$	diferença entre o EMSE do filtro i e o EMSE cruzado dos filtros combinados
NSD_i	relação entre o EMSE do filtro i e o EMSE ótimo
$\hat{\mathbf{R}}_i$	estimativa da matriz de autocorrelação no filtro i
ϵ_N, ϵ_R	constantes positivas pequenas

Combinação de filtros

$\mathbf{w}_{eq}$	vetor de coeficientes equivalente da combinação
y	sinal de saída da combinação
e_a	erro <i>a priori</i> da combinação
η	parâmetro de mistura
η_o	parâmetro de mistura ótimo
$\bar{\eta}_o$	valor médio do parâmetro de mistura ótimo
$\tilde{\eta}$	erro do parâmetro de mistura em relação ao ótimo
σ_η^2	variância do parâmetro de mistura
η_u	parâmetro de mistura universal da combinação convexa
α	sinal auxiliar para o cálculo do parâmetro de mistura da combinação convexa
α^+ e ϵ_C	constantes para cálculo do parâmetro de mistura universal na combinação convexa
μ_η e μ_α	passos de adaptação das combinações afim e convexa
$\mu_{\eta,s}, \tilde{\mu}_\eta$ e $\tilde{\mu}_{\eta,s}$	passos de adaptação dos algoritmos normalizados da combinação afim
ν	fator de esquecimento usado nos algoritmos PN e NSR
ϱ	constante que determina a proporção de transferência de coeficientes dos filtros componentes para a combinação
ζ	EMSE da combinação
NSD	relação entre o EMSE da combinação e o EMSE ótimo

Capítulo 1

Introdução e formulação do problema

Neste capítulo são apresentados o contexto, as contribuições e a organização do trabalho. Inicialmente é feita uma introdução à filtragem adaptativa e são apresentadas as combinações convexas e afim. Em seguida, são descritas algumas propriedades básicas dessas combinações e são mostrados alguns resultados existentes na literatura. Por fim, são apresentados os objetivos, justificativa, contribuições e a organização da dissertação.

1.1 Uma breve introdução à filtragem adaptativa

Filtros adaptativos são utilizados em diversas aplicações como equalização de canais de comunicação, identificação de sistemas, cancelamento de eco acústico, controle ativo de ruído, entre outras [Haykin, 2002]. Um esquema de um filtro adaptativo com resposta impulsiva de duração finita (FIR - *finite impulse response*) de M coeficientes é mostrado na Figura 1.1. O sinal de saída $y(n)$ é obtido através do produto interno entre o vetor regressor de entrada

$$\mathbf{u}(n) = [u(n) \quad u(n-1) \quad \cdots \quad u(n-M+1)]^T$$

e o vetor de coeficientes do filtro

$$\mathbf{w}(n-1) = [w_0(n-1) \quad w_1(n-1) \quad \cdots \quad w_{M-1}(n-1)]^T,$$

ou seja, $y(n) = \mathbf{u}^T(n)\mathbf{w}(n-1)$, sendo que $(\cdot)^T$ indica a transposição de um vetor. A saída $y(n)$ representa a estimativa do sinal desejado $d(n)$. O vetor de coeficientes é então atualizado utilizando-se o erro $e(n) = d(n) - y(n)$.

Algoritmos do tipo quase-Newton baseados na função custo do erro quadrático médio atualizam o vetor de coeficientes do filtro da forma

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \rho(n)\mathbf{M}(n)\mathbf{u}(n)e(n), \quad (1.1)$$

sendo $\rho(n)$ um passo de adaptação e $\mathbf{M}(n)$ uma matriz não-singular de dimensão $M \times M$.

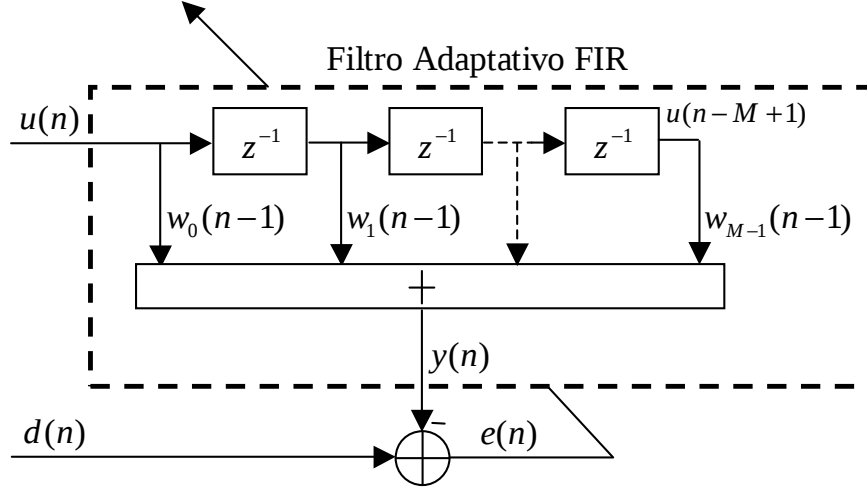


Figura 1.1: Filtro adaptativo do tipo FIR.

A partir de escolhas adequadas para $\rho(n)$ e $\mathbf{M}(n)$ chega-se às equações de atualização dos coeficientes do filtro para diversos algoritmos. Com $\mathbf{M}(n) = \mathbf{I}$, é possível obter o algoritmo LMS (*least-mean square*) fazendo $\rho(n) = \mu$ ou o NLMS (*normalized LMS*), fazendo

$$\rho(n) = \frac{\tilde{\mu}}{\epsilon_N + \|\mathbf{u}(n)\|^2}, \quad (1.2)$$

sendo $\mathbf{I}$ a matriz identidade, μ e $\tilde{\mu}$ os passos de adaptação do LMS e NLMS respectivamente e ϵ_N um fator de regularização. Para obter o RLS (*recursive least squares*), $\rho(n) = 1$ e $\mathbf{M}(n)$ deve ser uma estimativa da inversa da matriz de autocorrelação do sinal de entrada, definida como

$$\mathbf{R} = \mathbf{E}\{\mathbf{u}(n)\mathbf{u}^T(n)\}, \quad (1.3)$$

em que $\mathbf{E}\{\cdot\}$ representa o operador de esperança matemática. O cálculo da estimativa de $\mathbf{M}^{-1}(n) = \hat{\mathbf{R}}(n)$ é feito usando um fator de esquecimento λ , ou seja,

$$\mathbf{M}^{-1}(n) = \hat{\mathbf{R}}(n) \triangleq \lambda \hat{\mathbf{R}}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n), \quad (1.4)$$

com a inicialização dada geralmente por

$$\hat{\mathbf{R}}(0) = \epsilon_R \mathbf{I}, \quad (1.5)$$

sendo ϵ_R uma constante positiva de pequeno valor. A matriz $\mathbf{M}(n)$ é então obtida aplicando-se o lema da inversão matricial em (1.4) [Sayed, 2003, pág. 246].

Existem aplicações em que o sinal desejado é desconhecido, como em equalização auto-didata, conforme esquematizado na Figura 1.2. Nessa aplicação, o objetivo é obter uma estimativa $\hat{a}(n - \tau_d)$ do sinal transmitido $a(n)$ a partir do sinal recebido $u(n)$. Idealmente,

$\hat{a}(n - \tau_d)$ e $a(n)$ devem ser relacionados apenas por um atraso de τ_d amostras.

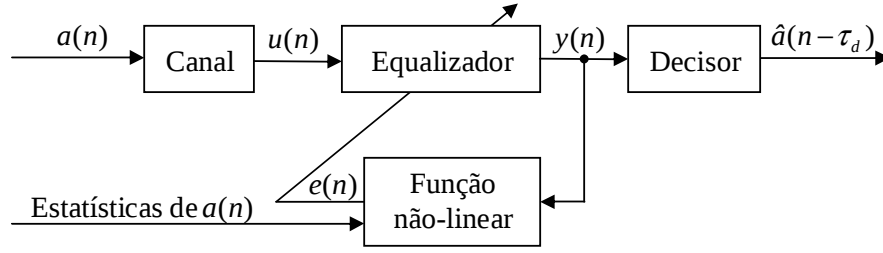


Figura 1.2: Equalização autodidata.

Nesse caso, $d(n)$ pode ser estimado através de estatísticas de ordem superior (HOS - *high-order statistics*) do sinal transmitido. A função custo do módulo constante é uma das mais utilizadas na literatura e sua minimização dá origem ao CMA (*constant modulus algorithm*), proposto em [Godard, 1980]. A equação de atualização do CMA pode ser escrita na forma (1.1) com $\rho = \mu$, $\mathbf{M}(n) = \mathbf{I}$ e $e(n) = [r - y^2(n)]y(n)$, sendo r uma constante que depende de estatísticas de ordem superior do sinal transmitido, definida como

$$r = \frac{\mathbb{E}\{a^4(n)\}}{\mathbb{E}\{a^2(n)\}}. \quad (1.6)$$

Na Tabela 1.1, é mostrado um resumo dos parâmetros utilizados na Equação (1.1) para os algoritmos considerados.

Tabela 1.1: Parâmetros da Equação (1.1) para os algoritmos LMS, NLMS, RLS e CMA.

Algoritmo.	$\rho(n)$	$\mathbf{M}^{-1}(n)$	$e(n)$
LMS	μ	$\mathbf{I}$	$d(n) - y(n)$
NLMS	$\tilde{\mu} / (\epsilon_N + \ \mathbf{u}(n)\ ^2)$		
RLS	1		
CMA	μ	$\hat{\mathbf{R}}(n) = \lambda \hat{\mathbf{R}}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n)$	$[r - y^2(n)]y(n)$

Embora seja comum utilizar a mesma notação para os passos de adaptação do LMS e do CMA, os intervalos dos passos que asseguram a estabilidade e a convergência desses algoritmos são diferentes. No caso do LMS, o resultado é bem conhecido na literatura [Haykin, 2002, Sayed, 2003, Diniz, 2008], enquanto para o CMA, há alguns resultados restritos, como por exemplo os resultados de [Silva e Nascimento, 2008a], onde é considerada a hipótese do algoritmo ter sido inicializado próximo do vetor de coeficientes ótimos.

O desempenho de um filtro adaptativo pode ser avaliado por diversos fatores como precisão da solução em regime, velocidade de convergência, capacidade de *tracking*, custo computacional, robustez numérica, estabilidade, etc. [Haykin, 2002, Sayed, 2003, Diniz, 2008]. O estudo

de uma solução que apresente um bom compromisso entre esses fatores é um problema de amplo interesse.

1.2 Combinação de filtros adaptativos

Uma das maiores dificuldades do projeto de filtros adaptativos é a escolha dos parâmetros fixos do filtro, como o passo de adaptação para os algoritmos do tipo LMS ou o fator de esquecimento para algoritmos do tipo RLS. Na literatura, há diversas soluções que propõem o uso do passo ou fator de esquecimento variáveis [Kwong e Johnston, 1992, Aboulnasr e Mayyas, 1997, Zhuang, 1998, Bilcu et al., 2002, Costa e Bermudez, 2006]. No caso do LMS, por exemplo, o passo deve ser grande o suficiente para garantir uma rápida convergência e após a convergência inicial deve ter seu valor diminuído de modo a obter uma precisão maior na solução em regime.

Uma outra solução que está sendo atualmente muito estudada é a combinação de dois ou mais filtros adaptativos [Arenas-García e Figueiras-Vidal, 2005, Arenas-García et al., 2005a,b, Arenas-García e Figueiras-Vidal, 2006, Arenas-García et al., 2006a,b, Silva e Nascimento, 2008b, Bershad et al., 2008]. Esse é um método relativamente simples e proporciona um desempenho global melhor ou igual ao de cada filtro individual operando independentemente.

A idéia da combinação de diversos filtros adaptativos independentes para se obter um melhor desempenho do que cada filtro individual não é nova. Ela foi proposta inicialmente em [Andersson, 1985] e posteriormente melhorada em [Niedźwiecki, 1990, 1992]. No entanto, o método de [Arenas-García et al., 2006a] está recebendo mais atenção devido à sua relativa simplicidade e à prova, dadas certas condições, de que a combinação é universal, i.e., considerando entradas estacionárias, a estimativa combinada é pelo menos tão boa quanto a do melhor filtro componente a cada instante.

O esquema da combinação de dois filtros adaptativos é mostrado na Figura 1.3. Na Figura 1.3-(a), é considerada a filtragem supervisionada que pode ser usada para diferentes aplicações como identificação de sistemas, equalização adaptativa, cancelamento de eco acústico, etc. [Haykin, 2002, Sayed, 2003]. O sinal de saída global $y(n)$ é obtido a partir da combinação das saídas $y_1(n)$ e $y_2(n)$ dos filtros individuais, ou seja,

$$y(n) = \eta(n)y_1(n) + [1 - \eta(n)]y_2(n), \quad (1.7)$$

sendo $\eta(n)$ o parâmetro de mistura da combinação. Os vetores de coeficientes de cada filtro $\mathbf{w}_1(n-1)$ e $\mathbf{w}_2(n-1)$ são adaptados com seus respectivos erros

$$e_i(n) = d(n) - y_i(n), \quad i = 1, 2, \quad (1.8)$$

sendo $d(n)$ a resposta desejada. O erro global da combinação é calculado como

$$e(n) = d(n) - y(n), \quad (1.9)$$

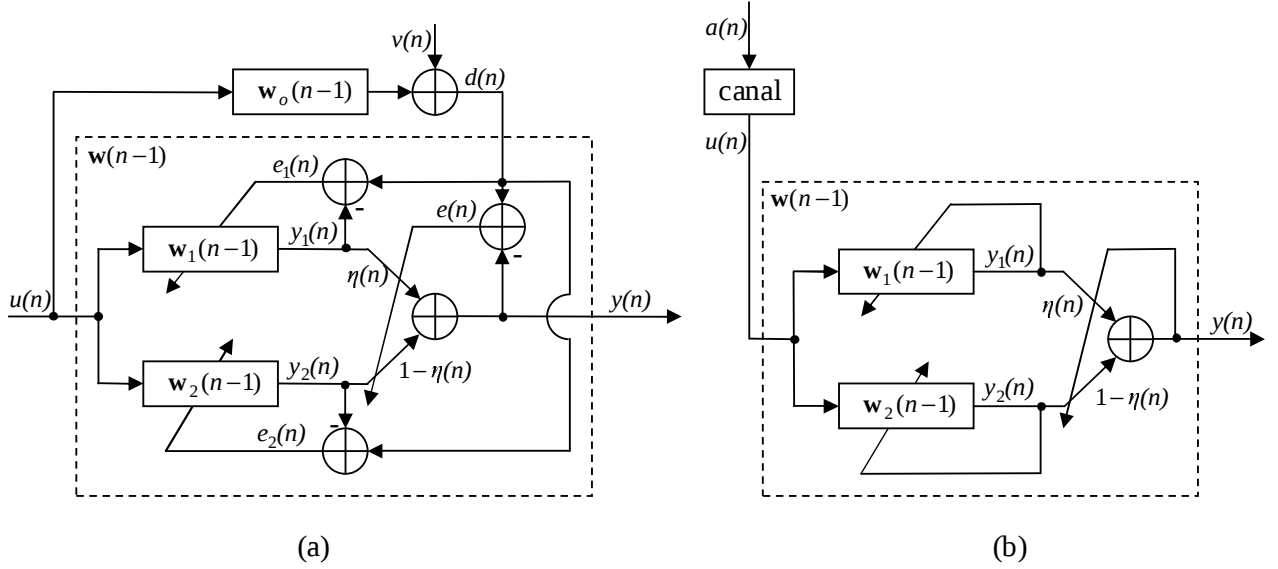


Figura 1.3: Combinação de filtros adaptativos transversais para (a) filtragem supervisionada e (b) equalização autodidata.

que pode ser escrito, usando (1.7) e (1.8), como

$$e(n) = \eta(n)e_1(n) + [1 - \eta(n)]e_2(n). \quad (1.10)$$

Nesse caso, assume-se um modelo de regressão linear para $d(n)$, isto é,

$$d(n) = \mathbf{u}^T(n)\mathbf{w}_o(n-1) + v(n), \quad (1.11)$$

sendo $\mathbf{w}_o(n-1)$ o vetor de coeficientes ótimos e $v(n)$ um processo aleatório de média zero e variância $\sigma_v^2 = E\{v^2(n)\}$, não-correlacionado com $\mathbf{u}(n)$ [Sayed, 2003]. As seqüências $\{u(n)\}$ e $\{v(n)\}$ são assumidas estacionárias. No caso estacionário, $\mathbf{w}_o(n)$ é uma constante e no caso não-estacionário é comum assumir o *random-walk model* [Sayed, 2003], ou seja

$$\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n), \quad (1.12)$$

sendo $\mathbf{q}(n)$ um vetor i.i.d. (independente e identicamente distribuído) com matriz de autocorrelação $\mathbf{Q} = E\{\mathbf{q}(n)\mathbf{q}^T(n)\}$ positiva-definida, independente das condições iniciais $\{\mathbf{w}_o(0), \mathbf{w}(0), \eta(0)\}$ e de $\{\mathbf{u}(l), d(l)\}$ para todo l [Sayed, 2003, Sec. 7.4].

Na Figura 1.3-(b), é mostrado um sistema de comunicação simplificado com uma combinação de dois equalizadores autodidatas. Nessa aplicação, o sinal $a(n)$, assumido i.i.d. e

não-gaussiano, é transmitido através de um canal desconhecido, cujo modelo é constituído por um filtro FIR e ruído branco aditivo gaussiano (AWGN - *additive white Gaussian noise*). A partir do sinal recebido $u(n)$ e de propriedades estatísticas do sinal transmitido, o equalizador autodidata deve minimizar os efeitos do canal e recuperar o sinal $a(n)$ para algum atraso τ_d . Na combinação de algoritmos baseados na função custo do módulo constante, o erro de cada equalizador componente é dado por [Godard, 1980]

$$e_i(n) = [r - y_i^2(n)]y_i(n), \quad i = 1, 2, \quad (1.13)$$

enquanto o erro global é

$$e(n) = [r - y^2(n)]y(n). \quad (1.14)$$

Nesse caso, a solução ótima $\mathbf{w}_o(n)$ em muitos casos é desconhecida, o que inviabiliza o uso do *random-walk model* segundo (1.12) para o caso não-estacionário. Uma forma de contornar o problema é aplicar a perturbação causada pelo vetor $\mathbf{q}(n)$ no vetor de coeficientes do filtro, conforme mostrado no Apêndice A.

Como no caso supervisionado, a combinação é dada por (1.7). Entretanto, como no CMA os erros $e_1(n)$, $e_2(n)$ e $e(n)$ são calculados a partir de uma função não-linear de $y_1(n)$, $y_2(n)$ e $y(n)$, respectivamente, uma equação similar a (1.10) é uma aproximação que depende de algumas hipóteses, entre elas a inicialização do vetor de coeficientes do equalizador próximo da solução ótima de *zero-forcing* [Silva e Nascimento, 2008b, Nascimento e Silva, 2008].

O desempenho da combinação e dos filtros componentes pode ser medido, dentre outras formas, através do erro quadrático médio em excesso (EMSE - *excess mean squared error*). O EMSE é definido como a diferença entre o erro quadrático médio e seu valor mínimo que se obtém com o vetor de coeficientes ótimos. O EMSE de um filtro i é calculado como

$$\zeta_i(n) \triangleq \mathbb{E}\{e_{a,i}^2(n)\}, \quad (1.15)$$

em que $e_{a,i}(n)$ é o erro *a priori* do filtro i , definido como,

$$e_{a,i}(n) \triangleq \mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1), \quad (1.16)$$

sendo

$$\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o(n) - \mathbf{w}_i(n) \quad (1.17)$$

o vetor de erro de coeficientes do filtro i . Analogamente, define-se o EMSE da combinação como

$$\zeta(n) \triangleq \mathbb{E}\{e_a^2(n)\}, \quad (1.18)$$

sendo $e_a(n)$ o erro *a priori* da combinação. Para a medida de desempenho da combinação,

também é necessário definir o EMSE cruzado

$$\zeta_{12}(n) \triangleq E\{e_{a,1}(n)e_{a,2}(n)\}. \quad (1.19)$$

O valor em regime do EMSE de um filtro i é definido como

$$\zeta_i(\infty) \triangleq \lim_{n \rightarrow \infty} \zeta_i(n) = \lim_{n \rightarrow \infty} E\{e_{a,i}^2(n)\} \quad (1.20)$$

e o valor em regime do EMSE cruzado é dado por

$$\zeta_{12}(\infty) \triangleq \lim_{n \rightarrow \infty} \zeta_{12}(n) = \lim_{n \rightarrow \infty} E\{e_{a,1}(n)e_{a,1}(n)\}. \quad (1.21)$$

O primeiro tipo de combinação a receber maior atenção foi a combinação convexa. Nesse esquema, o parâmetro de mistura $\eta(n)$ fica limitado ao intervalo $[0, 1]$ [Arenas-García e Figueiras-Vidal, 2005, Arenas-García et al., 2005a,b, Arenas-García e Figueiras-Vidal, 2006, Arenas-García et al., 2006a,b, Silva e Nascimento, 2008b]. Entretanto, foi proposta em [Bershad et al., 2008] uma combinação afim de dois filtros LMS, que pode ser interpretada como uma generalização da combinação convexa de [Arenas-García et al., 2006a]. Nesse tipo de combinação, o valor de $\eta(n)$ não fica restrito ao intervalo $[0, 1]$, podendo inclusive adquirir um valor negativo. Esses dois tipos de combinação são detalhados a seguir.

1.2.1 Combinação Convexa

Na combinação convexa, para manter o parâmetro $\eta(n)$ no intervalo $[0, 1]$, pode-se utilizar uma variável auxiliar $\alpha(n-1)$ e uma função sigmoideal [Arenas-García et al., 2006b,a], isto é,

$$\eta(n) = \text{sgm}(\alpha(n-1)) = \frac{1}{1 + e^{-\alpha(n-1)}}, \quad (1.22)$$

com o parâmetro $\alpha(n)$ atualizado como

$$\alpha(n) = \alpha(n-1) + \mu_\alpha e_\eta(n) \eta(n) [1 - \eta(n)], \quad (1.23)$$

sendo

$$e_\eta(n) = e(n)[y_1(n) - y_2(n)] \quad (1.24)$$

e $e(n)$ calculado segundo (1.9) e (1.14) para o caso supervisionado e autodidata, respectivamente. A Equação (1.23) foi obtida em [Arenas-García et al., 2006a,b] e em [Arenas-García e Figueiras-Vidal, 2006], usando o método do gradiente estocástico para minimizar a função custo instantânea do erro quadrático médio no caso supervisionado e do módulo constante no caso autodidata.

Um problema desse esquema é que a atualização de $\alpha(n)$ pára quando $\eta(n)$ fica próximo

de 0 ou 1. Para evitar isso, Arenas-García et al. [2006a] sugeriram que $\alpha(n)$ fique restrito (por simples saturação) ao intervalo $[-\alpha^+, \alpha^+]$. Assim, um mínimo nível de atualização é sempre garantido. Em [Arenas-García et al., 2006a], foi proposta ainda uma variação dessa combinação onde a universalidade é garantida. A idéia é arredondar o valor do parâmetro de mistura para 0 ou 1 a partir de determinados valores de $\alpha(n)$, obtendo-se assim um parâmetro de mistura $\eta_u(n)$ que garante a universalidade da combinação. A saída da combinação universal $y_u(n)$ é calculada de maneira semelhante, ou seja,

$$y_u(n) = \eta_u(n)y_1(n) + [1 - \eta_u(n)]y_2(n), \quad (1.25)$$

e o parâmetro de mistura universal $\eta_u(n)$ é calculado como

$$\eta_u(n) = \begin{cases} 1, & \alpha(n) \geq \alpha^+ - \epsilon_C \\ \eta(n), & -\alpha^+ + \epsilon_C < \alpha(n) < \alpha^+ - \epsilon_C \\ 0, & \alpha(n) \leq -\alpha^+ + \epsilon_C \end{cases}, \quad (1.26)$$

sendo ϵ_C uma constante positiva de pequeno valor.

A combinação convexa de dois algoritmos do tipo LMS ou CMA com passos de adaptação μ_1 e μ_2 , sendo $\mu_1 > \mu_2$, tem uma interpretação simples. No início da convergência, $\eta(n) \rightarrow 1$ e a combinação se aproxima do filtro com passo μ_1 , que converge mais rapidamente. Em regime, $\eta(n) \rightarrow 0$ e a combinação se aproxima do filtro com passo μ_2 , que por ser mais lento, faz um “ajuste fino” a fim de atingir um erro de estimação menor. Na Figura 1.4 é mostrado um exemplo desse comportamento para uma combinação convexa de dois filtros LMS com passos de adaptação $\mu_1 = 0,1$ e $\mu_2 = 0,01$. Nesse exemplo, foi feita uma média de 50 realizações usando o algoritmo que garante a universalidade na configuração de identificação de sistemas descrita em [Arenas-García et al., 2006a], com $\mu_\alpha = 100$, $\alpha^+ = 4$ e $\epsilon_C = 0,1$. Para facilitar a visualização, as curvas foram filtradas por um filtro de média-móvel de 128 coeficientes. Pode-se notar que no início da convergência, a combinação tem um desempenho igual ao do filtro μ_1 -LMS, com $\eta(n) \approx 1$ e em regime, o desempenho é igual ao do filtro μ_2 -LMS com $\eta(n) \approx 0$. Há situações em que $0 < \eta(n) < 1$ e a combinação pode apresentar um desempenho melhor do que o de cada filtro individual [Arenas-García et al., 2006a, Bershad et al., 2008]. Na Figura 1.4, esse comportamento pode ser observado na região de transição, quando $\eta(n)$ varia de 1 para 0 e o desempenho da combinação é melhor que o desempenho dos filtros componentes durante algumas iterações. Entretanto, no caso não-estacionário, esse comportamento também pode ocorrer em regime [Arenas-García et al., 2006a, Silva e Nascimento, 2008b].

Um dos problemas da combinação convexa é a escolha do passo μ_α para a adaptação do parâmetro de mistura da combinação. Esse problema ocorre principalmente quando a relação sinal-ruído (SNR - *signal-to-noise ratio*) do sinal de entrada é variável com o tempo. Para facilitar a escolha do passo de adaptação da combinação e melhorar o seu desempenho em cenários como esse, Azpicueta-Ruiz et al. [2008] propuseram um esquema normalizado de

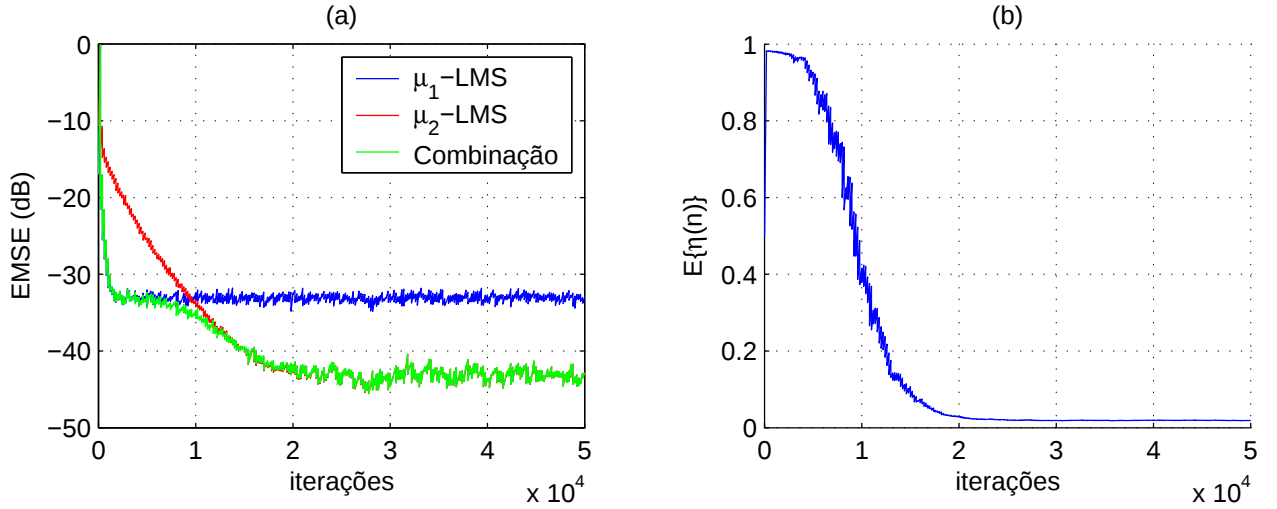


Figura 1.4: Exemplo de combinação convexa de dois filtros LMS. Configuração de identificação de sistemas, $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$ e $\mu_2 = 0,01$. Entrada colorida (modelo autoregressivo com pólo em 0,8) e solução ótima dada por um vetor aleatório de média nula e variância unitária. Média de 50 realizações com $\mu_\alpha = 100$, $\alpha^+ = 4$ e $\epsilon_C = 0,1$ usando o parâmetro de mistura universal definido em (1.26).

adaptação do parâmetro de mistura. Nesse esquema, o parâmetro $\alpha(n)$ da combinação é adaptado por

$$\alpha(n) = \alpha(n-1) + \frac{1}{p(n)} \underbrace{\mu_\alpha \eta(n)[1 - \eta(n)]}_{\text{passo de adaptação}} e(n) \underbrace{[y_1(n) - y_2(n)]}_{\text{signal de entrada}} \quad (1.27)$$

com

$$p(n) = \nu p(n-1) + (1 - \nu)[y_1(n) - y_2(n)]^2, \quad (1.28)$$

sendo $0 < \nu < 1$, um fator de esquecimento.

A adaptação de $\alpha(n)$, feita originalmente através do gradiente estocástico (Equação (1.23)), pode ser interpretada como um filtro LMS de apenas um coeficiente $\alpha(n)$ com sinal de entrada $[y_1(n) - y_2(n)]$. A normalização consiste em dividir o passo de adaptação pela norma ao quadrado do vetor regressor de entrada. No entanto, como o filtro tem apenas um coeficiente, a norma ao quadrado do vetor regressor de entrada é dada por $|y_1(n) - y_2(n)|^2$, que tem uma variação muito grande e pode fazer com que o algoritmo não funcione corretamente. Por isso, foi utilizada uma estimativa desse valor, dada por $p(n)$ e baseada numa janela exponencial. Nessa estimativa, o parâmetro ν funciona como um fator de esquecimento e pode ser facilmente ajustado, por exemplo, $\nu = 0,9$ [Azpicueta-Ruiz et al., 2008]. Com essa adaptação, o passo μ_α fica mais fácil de ser ajustado pois se torna independente da potência do sinal de entrada.

No caso não-estacionário, para determinados valores de $\text{Tr}(\mathbf{Q})$, σ_v^2 e $\text{Tr}(\mathbf{R})$, em que $\text{Tr}(\mathbf{A})$ se refere ao traço de uma matriz $\mathbf{A}$, existe um valor ótimo do parâmetro de adaptação (μ para algoritmos do tipo LMS ou λ para algoritmos do tipo RLS) que minimiza o EMSE em regime

[Sayed, 2003]. Nesse caso, a comparação do desempenho da combinação convexa com o de seus filtros componentes pode ser feita através de uma medida normalizada denominada NSD (*normalized square deviation*). O NSD é definido como a razão entre o EMSE em regime de um filtro componente ou da combinação e o EMSE em regime obtido com o parâmetro de adaptação ótimo denotado por $\zeta_o(\infty)$, ou seja, para cada filtro componente, o NSD é dado por

$$\text{NSD}_i(\infty) = \frac{\zeta_i(\infty)}{\zeta_o(\infty)}, \quad i = 1, 2 \quad (1.29)$$

e para a combinação

$$\text{NSD}(\infty) = \frac{\zeta(\infty)}{\zeta_o(\infty)}. \quad (1.30)$$

Define-se também o NSD cruzado

$$\text{NSD}_{12}(\infty) = \frac{\zeta_{12}(\infty)}{\zeta_o(\infty)}. \quad (1.31)$$

Na Figura 1.5 são mostradas as curvas de NSD dos filtros componentes ($\text{NSD}_1(\infty)$ e $\text{NSD}_2(\infty)$), do cruzado ($\text{NSD}_{12}(\infty)$) e da combinação convexa ($\text{NSD}(\infty)$) em função de $\text{Tr}(\mathbf{Q})$ para a combinação de dois filtros LMS como em [Arenas-García et al., 2006a]. É possível notar que quando $\text{Tr}(\mathbf{Q})$ está entre os valores q_1 e q_2 , $q_i = \mu_i^2 \sigma_v^2 \text{Tr}(\mathbf{R})$, $i = 1, 2$, a combinação apresenta um NSD e conseqüentemente um EMSE menor do que o dos filtros componentes. Nesse intervalo de $\text{Tr}(\mathbf{Q})$, é possível notar também que $0 < E\{\eta(\infty)\} < 1$. Um outro aspecto importante que foi observado em [Arenas-García et al., 2006a] é que quando $\text{Tr}(\mathbf{Q}) = q_i$, $i = 1$ ou 2 , o valor de $E\{\eta(\infty)\}$ é 1 ou 0 , respectivamente. Com base nesse resultado, Arenas-García et al. [2006a] propuseram uma combinação de algoritmos de passo variável (CVS-LMS - *combination of variable step LMS*) que altera o passo de adaptação dos filtros componentes. Mudando o passo, os valores de q_1 e q_2 são alterados de forma a manter $q_2 < \text{Tr}(\mathbf{Q}) < q_1$, o que faz com que a combinação possa apresentar uma capacidade de *tracking* superior em relação à dos filtros componentes.

Uma extensão da combinação convexa para o caso autodidata foi feita em [Arenas-García e Figueiras-Vidal, 2006]. Nesse trabalho, foram combinados dois algoritmos CMA com passos de adaptação diferentes, inicializados com o mesmo vetor de coeficientes. Os resultados obtidos foram semelhantes aos do caso supervisionado, possibilitando uma maior velocidade de convergência e um EMSE reduzido em regime. Cabe observar que para o caso do CMA, ainda pode ser estudada a combinação de filtros com passos de adaptação iguais mas vetores de inicialização diferentes. Sabe-se que dependendo da forma como é feita a inicialização do CMA, este pode convergir para um mínimo local ou global [Johnson Jr. et al., 1998]. Assim, combinar algoritmos com inicializações diferentes pode ser vantajoso com o objetivo de evitar a convergência para um mínimo local.

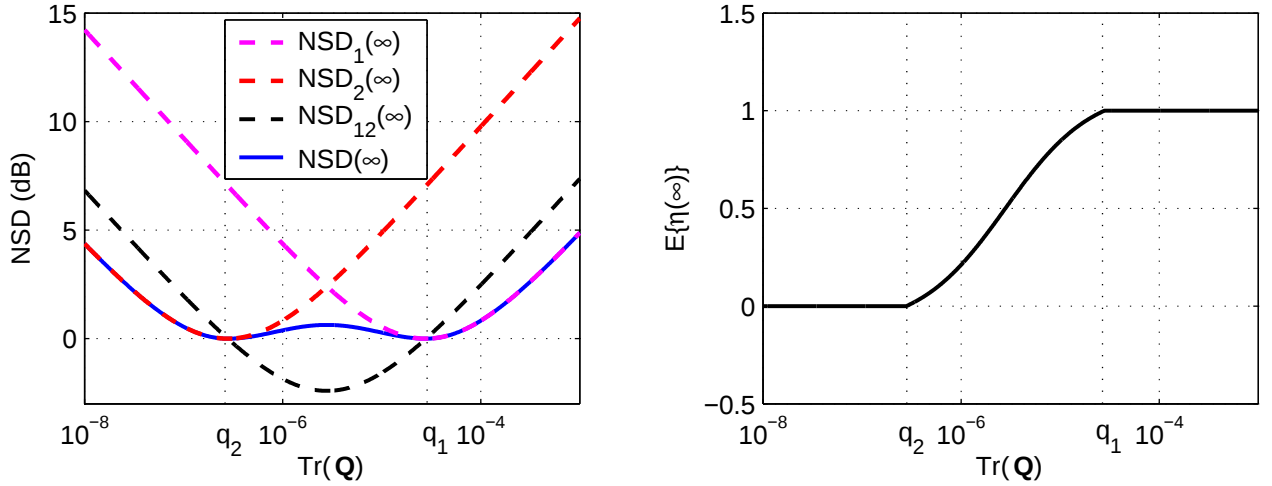


Figura 1.5: Comportamento da combinação convexa de dois algoritmos LMS no caso não-estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 2 \times 10^{-2}$, $\mu_2 = 2 \times 10^{-3}$. Entrada colorida (modelo autoregressivo com pólo em 0,8) com potência $1/M$.

1.2.2 Combinação afim

Uma generalização da combinação convexa de dois filtros adaptados com o algoritmo LMS foi proposta em [Bershad et al., 2008]. Nesse artigo, o parâmetro de combinação $\eta(n)$ é escolhido de forma ótima para minimizar o MSE instantâneo, não ficando restrito ao intervalo $[0, 1]$. Dessa forma, a saída global $y(n)$ é uma combinação afim¹ das saídas dos filtros individuais ($y_1(n)$ e $y_2(n)$) e a combinação convexa é um caso particular. Em Bershad et al. [2008] foi feito um estudo do combinador ótimo não-realizável para ambientes estacionários onde se concluiu que o valor do parâmetro de mistura ótimo instantâneo deve ser

$$\eta_o(n) = \frac{[\mathbf{w}_o - \mathbf{w}_2(n)]^T \mathbf{R} [\mathbf{w}_1(n) - \mathbf{w}_2(n)]}{[\mathbf{w}_1(n) - \mathbf{w}_2(n)]^T \mathbf{R} [\mathbf{w}_1(n) - \mathbf{w}_2(n)]}. \quad (1.32)$$

Considerando o passo de adaptação $\mu_2 = \delta \mu_1$, $0 < \delta < 1$ e assumindo algumas hipóteses equivalentes a escolher μ_1 para máxima velocidade de convergência, com entrada branca e gaussiana, chegou-se à conclusão que o valor em regime do parâmetro de mistura ótimo é dado por

$$\bar{\eta}_o(\infty) \triangleq \lim_{n \rightarrow \infty} E\{\eta_o(n)\} \approx \frac{\delta}{2(\delta - 1)}. \quad (1.33)$$

Como $0 < \delta < 1$, o parâmetro de mistura ótimo em regime é sempre negativo.

A Equação (1.32) depende da matriz de autocorrelação do sinal de entrada e do vetor de coeficientes ótimos e por isso não é possível utilizá-la para o cálculo do parâmetro de mistura numa aplicação prática. Por isso, foram então propostos dois novos esquemas para adaptação de $\eta(n)$. O primeiro usa o método do gradiente estocástico e tem um desempenho próximo

¹Uma combinação afim dos vetores $\mathbf{x}_1 \dots \mathbf{x}_L$ é a combinação linear $\sum_{i=1}^L c_i \mathbf{x}_i$ com $\sum_{i=1}^L c_i = 1$.

do ótimo não-realizável, apresentando a mesma velocidade de convergência e comportamento em regime. Nesse esquema, para que o parâmetro $\eta(n)$ possa atingir qualquer valor, não se utiliza uma função sigmoïdal como em (1.22) e sua atualização é feita diretamente como

$$\eta(n+1) = \eta(n) + \mu_\eta e_\eta(n), \quad (1.34)$$

com $e_\eta(n)$ calculado segundo a Equação (1.24). O algoritmo (1.34), que é chamado de η -LMS neste trabalho, foi obtido em [Bershad et al., 2008] usando o método do gradiente estocástico, de maneira análoga à utilizada na combinação convexa. A diferença na adaptação das duas combinações está no termo $\eta(n)[1-\eta(n)]$, que aparece na combinação convexa devido à função sigmoïdal. Um problema que foi observado em [Bershad et al., 2008] é que (1.34) pode divergir durante o início da adaptação quando é utilizado um passo de adaptação relativamente grande ($\mu_\eta \geq 1$). Para evitar esse problema, uma solução encontrada em [Bershad et al., 2008] foi limitar o valor de $\eta(n)$ em 1 para todo n .

Na segunda forma de adaptação do parâmetro de mistura proposta, são consideradas médias temporais dos erros instantâneos

$$\hat{e}_i(n) = \frac{1}{K} \sum_{m=n-K+1}^n e_i^2(m), \quad i=1,2, \quad (1.35)$$

em que K é o tamanho da janela de média móvel. O parâmetro de mistura $\eta(n)$ é dado por

$$\eta(n) = 1 - \kappa_\eta \operatorname{erf} \left(\frac{\hat{e}_1(n)}{\hat{e}_2(n)} \right), \quad (1.36)$$

com

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2/2} dt. \quad (1.37)$$

Nesse esquema de adaptação, $\eta(n)$ pode variar no intervalo $[1-\kappa_\eta, 1]$. Para que seja possível obter o valor do parâmetro de mistura ótimo em regime, considera-se

$$\kappa_\eta = 1 - \frac{\delta}{2(\delta - 1)}. \quad (1.38)$$

Foram feitas inúmeras simulações e os resultados experimentais obtidos com as duas formas de adaptação correspondem ao resultado teórico previsto por (1.32). Cabe notar que na segunda forma de adaptação, κ_η foi calculado para uma determinada situação de operação do filtro considerando ambiente estacionário. Para outras situações, o valor de κ_η deve ser alterado de forma a permitir que $\eta(n)$ atinja o valor de $\bar{\eta}_o(\infty)$. Na primeira forma de adaptação, é necessário apenas o ajuste de um passo de adaptação, o que pode ser mais interessante para aplicações práticas. Porém, através de simulações é possível verificar que a variância de $\eta(n)$ é grande quando a adaptação é feita através de (1.34). Diante disso, é interessante considerar outras formas de adaptação para o parâmetro de mistura e para tanto uma análise se faz

necessária. Uma alternativa é usar uma extensão do esquema de adaptação normalizado de [Azpicueta-Ruiz et al., 2008] para a combinação afim.

Os resultados de [Bershad et al., 2008] abrem novas perspectivas para diferentes combinações de outros algoritmos adaptativos e podem ser estendidos para outras combinações de algoritmos.

1.3 Objetivos

Nesta dissertação, o esquema proposto em [Bershad et al., 2008] é estendido para outros tipos de combinação afim envolvendo os algoritmos LMS, NLMS, RLS e CMA. São feitas análises em regime e do transitório dessas combinações considerando entrada branca ou colorida e ambientes estacionários ou não-estacionários. Para simplificar a análise, são considerados sinais reais.

São analisadas as combinações de dois filtros que utilizam o mesmo algoritmo (LMS, NLMS e RLS para o caso supervisionado e CMA para o caso autodidata) e também combinações de algoritmos diferentes para melhorar a capacidade de *tracking* do filtro global [Silva e Nascimento, 2008b] como, por exemplo, a combinação dos algoritmos LMS e RLS. Nesse contexto, são estudadas algumas formas alternativas de adaptação do parâmetro de mistura para melhorar a universalidade da combinação. Um outro aspecto investigado é o comportamento da combinação afim de dois algoritmos CMA com inicializações diferentes já que este é um fator determinante para o bom desempenho desse algoritmo.

1.4 Justificativa

O projeto de filtros adaptativos envolve a escolha de parâmetros fixos para manter um compromisso entre a velocidade de convergência e o nível do erro em regime do filtro. Uma das alternativas para facilitar o projeto é a combinação de algoritmos. Além disso, a combinação de algoritmos de diferentes famílias (LMS e RLS, por exemplo) permite que seja possível obter um melhor desempenho em ambientes não-estacionários já que dependendo do tipo de não-estacionariedade do ambiente, pode ser interessante usar uma determinada família de algoritmos ou outra. Nesse cenário, um estudo da combinação afim é algo de interesse já que esta permite obter o valor do parâmetro de mistura ótimo na maioria dos casos. Isso faz com que seja possível obter um desempenho superior ao dos filtros componentes em algumas situações em que a combinação convexa se comporta de maneira igual ao melhor filtro componente. Um outro ponto a ser observado é que o parâmetro de mistura da combinação afim tem um maior grau de liberdade em relação ao da combinação convexa. Isso pode ser vantajoso em alguns casos como, por exemplo, quando se deseja combinar algoritmos autodidatas com diferentes inicializações para evitar a convergência para mínimos locais.

1.5 Contribuições da dissertação

As principais contribuições deste trabalho são enumeradas a seguir.

1. Análise em regime da combinação afim.

Foi feita uma análise do valor em regime do parâmetro de mistura e EMSE ótimos das combinações considerando ambientes estacionário ou não-estacionário com o objetivo de comparar o desempenho das combinações convexa e afim. Verificou-se que em algumas situações, o valor do parâmetro de mistura ótimo é negativo e é possível obter um ganho de EMSE de até 3dB, o que motiva o uso da combinação afim.

2. Análise de transitório da combinação afim.

Para obter um modelo do transitório do EMSE da combinação, foi feita uma análise considerando a adaptação através do algoritmo η -LMS. Através da análise é possível prever algumas situações em que o desempenho obtido com essa adaptação é inferior ao universal, o que motivou propostas de formas alternativas de adaptação do parâmetro de mistura.

3. Algoritmo *signed regressor* para adaptação do parâmetro de mistura.

Foi proposta a adaptação do parâmetro de mistura da combinação afim considerando uma normalização instantânea parcial, originando o algoritmo *signed regressor*. A análise de transitório foi estendida para esse algoritmo e verificou-se que seu desempenho é mais próximo ao universal em algumas situações.

4. Algoritmos normalizados para adaptação do parâmetro de mistura.

Para facilitar a escolha do passo de adaptação do parâmetro de mistura das combinações, foram propostos dois algoritmos normalizados baseados no algoritmo η -LMS e no *signed regressor*, originando os algoritmos *power normalized* e *normalized signed regressor*, respectivamente. A análise de transitório foi estendida a esses algoritmos e verificou-se que essa normalização também torna o desempenho da combinação mais próximo do universal em alguns casos.

5. Aplicação da combinação afim na combinação de equalizadores CMA com inicializações diferentes.

Foi proposta uma aplicação da combinação afim para equalizadores CMA com inicializações diferentes. Foram feitas algumas simulações considerando um cenário simples e foi possível verificar que a combinação afim pode, em alguns casos, evitar a convergência para mínimos locais.

Como resultado dessas contribuições, foram publicados os seguintes artigos:

- R. Candido, M. T. M. Silva e V. H. Nascimento. Uma combinação afim de dois equalizadores autodidatas adaptados com o CMA. In *Simpósio Brasileiro de Telecomunicações (SBrT'08)*, 2008.
- R. Candido, M. T. M. Silva e V. H. Nascimento. Affine combinations of adaptive filters. In *Proc. of 42nd Asilomar Conference on Signals, Systems and Computers*, 2008.
- R. Candido, M. T. M. Silva e V. H. Nascimento. On combinations of CMA equalizers. In *Proc. of International Conference on Acoustics Systems and Signal Processing (ICASSP'09)*, 2009.

1.6 Organização da dissertação

A dissertação está estruturada em 5 capítulos. No Capítulo 2, a combinação afim é estendida para os algoritmos supervisionados NLMS e RLS e para equalização autodidata com o CMA. É feita uma análise do valor em regime do parâmetro de mistura e do EMSE da combinação, sendo verificadas algumas propriedades da combinação afim. Por fim, são apresentados resultados de simulação para validar a análise.

No Capítulo 3, é feita uma análise de transitório do EMSE da combinação afim considerando a adaptação através do algoritmo η -LMS. Em seguida, são mostrados alguns resultados de simulação para validar a análise. Nesses resultados, é mostrada uma situação em que a combinação afim apresenta um desempenho inferior ao universal e, para melhorar o desempenho, são propostas e analisadas três formas alternativas de adaptação do parâmetro de mistura. Por fim, são mostrados resultados de simulação referentes a essas formas alternativas de adaptação.

No Capítulo 4, é proposta a combinação de equalizadores CMA com inicializações diferentes. Considerando um cenário simples, é feita uma comparação entre o comportamento da combinação afim e convexa através de simulações e é mostrada uma situação em que a combinação afim é capaz de evitar a convergência dos filtros componentes para um mínimo local.

Finalmente, no Capítulo 5, são apresentadas as conclusões do trabalho e propostas para trabalhos futuros.

Capítulo 2

Análise em regime das combinações

Neste capítulo, a combinação afim de dois algoritmos LMS de Bershad et al. [2008] é estendida para os algoritmos NLMS, RLS e CMA. A partir desse resultado, é feita uma análise em regime para obter expressões analíticas para o valor do parâmetro de mistura e para o EMSE ótimos das combinações afim desses algoritmos. Através dessa análise, são verificadas algumas propriedades da combinação afim em regime.

A extensão da combinação afim para o caso de dois filtros NLMS ou dois filtros RLS é direta pois assim como o LMS, a adaptação do parâmetro de mistura da combinação desses algoritmos é baseada na função custo do erro quadrático médio. Para o CMA, entretanto, a função custo utilizada é a do módulo constante, sendo necessário assumir algumas hipóteses para permitir a extensão e análise da combinação. A seguir são descritas as funções custo e os modelos utilizados para o caso supervisionado e autodidata de forma a permitir uma análise unificada.

2.1 Funções custo e algoritmos

A adaptação do parâmetro de mistura $\eta(n)$ através do gradiente pode ser escrita na forma

$$\eta(n+1) = \eta(n) - \mu_\eta \frac{\partial J_i}{\partial \eta(n)}, \quad (2.1)$$

em que J_i representa uma determinada função custo a ser minimizada. Para algoritmos supervisionados (LMS, NLMS e RLS), a função custo utilizada é o erro quadrático médio. Para o CMA, utiliza-se a função custo do módulo constante [Godard, 1980]. As funções custo e os gradientes para esses algoritmos são mostrados na Tabela 2.1. De modo geral, o algoritmo estocástico de atualização do parâmetro de mistura é dado por

$$\eta(n+1) = \eta(n) + \mu_\eta e(n) [y_1(n) - y_2(n)]. \quad (2.2)$$

Tabela 2.1: Funções custo e gradientes dos algoritmos utilizados.

Algoritmo	J_i	$\frac{\partial J_i}{\partial \eta(n)}$
LMS, NLMS e RLS	$E\{e^2(n)\}$	$-2E\{e(n) [y_1(n) - y_2(n)]\}$
CMA	$E\{[r - y^2(n)]^2\}$	$-4E\{e(n) [y_1(n) - y_2(n)]\}$

Cabe notar que as constantes 2 e 4 que aparecem nos gradientes dos algoritmos supervisionados e do CMA podem ser desprezadas considerando um ajuste no passo de adaptação μ_η . Além disso, $e(n)$ é diferente para o caso supervisionado (Equação (1.9), pág. 5) e para o CMA (Equação (1.14), pág. 6). Assim como observado para o caso da combinação de algoritmos LMS em [Bershad et al., 2008], o algoritmo (2.2) pode divergir durante as iterações iniciais e uma forma de contornar o problema é limitar o valor máximo do parâmetro de mistura em 1 para todo n .

No caso do CMA, embora o algoritmo (2.2) possa ser usado para adaptar $\eta(n)$, verificou-se por simulações que ele nem sempre garante o comportamento universal da combinação. Dessa forma, uma outra alternativa para adaptar esse parâmetro é utilizar um algoritmo do gradiente estocástico que minimiza o erro de decisão, ou seja,

$$J_D = E\{e_d^2(n)\}, \quad (2.3)$$

sendo

$$e_d(n) = \hat{a}(n - \tau_d) - y(n). \quad (2.4)$$

Nesse caso, o gradiente é

$$\frac{\partial J_D}{\partial \eta(n)} = -2E\{e_d(n) [y_1(n) - y_2(n)]\} \quad (2.5)$$

e a adaptação do parâmetro de mistura é dada por

$$\eta(n+1) = \eta(n) + \mu_\eta e_d(n) [y_1(n) - y_2(n)], \quad (2.6)$$

limitando $\eta(n)$ máximo em 1 para evitar a divergência. Através de simulações é possível verificar que esse algoritmo é capaz de fazer com que o comportamento da combinação seja próximo ao universal mesmo na presença de ruído. Como $\hat{a}(n - \tau_d)$ faz o papel de sinal desejado no caso de equalização, os algoritmos (2.2) e (2.6) são chamados neste trabalho de η -LMS.

2.2 Modelos utilizados

No caso supervisionado, substituindo (1.11) em (1.9), o erro de cada filtro componente pode ser escrito como

$$e_i(n) = e_{a,i}(n) + v(n), \quad i = 1, 2. \quad (2.7)$$

Analogamente, o erro da combinação pode ser escrito como

$$e(n) = e_a(n) + v(n). \quad (2.8)$$

Usando (2.7) e (2.8) em (1.7) (pág. 4), pode-se escrever o erro *a priori* da combinação como

$$e_a(n) = \eta(n)e_{a,1}(n) + [1 - \eta(n)]e_{a,2}(n). \quad (2.9)$$

Para se obter um modelo unificado para os casos supervisionado e autodidata, com o erro *a priori* da combinação em função do erro *a priori* dos filtros componentes, é usual assumir algumas hipóteses para o caso autodidata:

- R1. $E\{a^k(n)\} = 0$, $k = 2m + 1$, $m \in \mathbb{N}$. Em outras palavras, $a(n)$ é sub-gaussiana e a constelação é simétrica, como é o caso da maioria das constelações usadas em comunicações digitais [Sayed, 2003];
- R2. A relação sinal-ruído na entrada é alta o suficiente para que $\hat{a}_i(n - \tau_d) \approx a(n - \tau_d) \approx \mathbf{u}^T(n)\mathbf{w}_o(n - 1)$, $i = 1, 2$. Entretanto, devido à variação do canal e ao ruído do gradiente, o vetor de coeficientes $\mathbf{w}_i(n - 1)$, $i = 1, 2$ não é igual a $\mathbf{w}_o(n - 1)$, mesmo em regime. Para tornar essa hipótese mais realista, é essencial que os equalizadores sejam implementados utilizando sobreamostragem, o que possibilita a equalização perfeita de um canal FIR na ausência de ruído [Treichler et al., 1996].

A saída $y_i(n)$ do equalizador i pode ser escrita como

$$y_i(n) = \mathbf{u}^T(n)\mathbf{w}_i(n - 1) = \mathbf{u}^T(n) [\mathbf{w}_o(n - 1) - \tilde{\mathbf{w}}_i(n - 1)], \quad i = 1, 2, \quad (2.10)$$

que pode ser aproximada, usando R2, por

$$y_i(n) \approx a(n - \tau_d) - e_{a,i}(n), \quad i = 1, 2. \quad (2.11)$$

Essa aproximação tem sido amplamente utilizada na análise em regime do CMA [Yousef e Sayed, 1999, Mai e Sayed, 2000, Silva e Miranda, 2004, Silva e Nascimento, 2008a,b]. Nesse caso, o erro *a priori* dos filtros componentes é dado por

$$e_{a,i}(n) \approx a_i(n - \tau_d) - y_i(n) \approx \hat{a}_i(n - \tau_d) - y_i(n) = e_{d,i}(n), \quad (2.12)$$

em que $e_{d,i}(n)$ é o erro de decisão do equalizador i .

Usando a aproximação (2.11), o erro de um filtro componente da combinação de dois algoritmos CMA, dado por (1.13) (pág. 6), pode ser reescrito como

$$e_i(n) = \gamma(n)e_{a,i}(n) + \beta(n) + s_i(n), \quad (2.13)$$

sendo

$$\gamma(n) = 3a^2(n - \tau_d) - r, \quad (2.14)$$

$$\beta(n) = r a(n - \tau_d) - a^3(n - \tau_d) \quad (2.15)$$

e

$$s_i(n) = -3a(n - \tau_d)e_{a,i}^2(n) + e_{a,i}^3(n). \quad (2.16)$$

Se $e_{a,i}(n)$ for razoavelmente pequeno em regime, termos que dependem de combinações de ordem superior de $e_{a,i}(n)$ podem ser desprezados em (2.13), isto é, $\lim_{n \rightarrow \infty} s_i(n) \approx 0$, o que leva à aproximação

$$e_i(n) \approx \gamma(n)e_{a,i}(n) + \beta(n), \quad i = 1, 2. \quad (2.17)$$

Analogamente, o erro da combinação pode ser aproximado por

$$e(n) \approx \gamma(n)e_a(n) + \beta(n). \quad (2.18)$$

As Equações (2.17) e (2.18) para o caso autodidata são análogas às Equações (2.7) e (2.8) do caso supervisionado. No caso autodidata, $\beta(n)$ é uma variável i.i.d que faz o papel de um ruído de medida com $E\{\beta(n)\} = 0$ devido à hipótese R1 e variância

$$\sigma_\beta^2 \triangleq E\{\beta^2(n)\} = E\{a^6(n) - r^2 a^2(n)\}. \quad (2.19)$$

Cabe observar que $\beta(n)$ é identicamente nulo para constelações de módulo constante [Silva e Nascimento, 2008b].

Analogamente, $\gamma(n)$ também é i.i.d. e seus momentos de primeira e segunda ordens são dados respectivamente por

$$\bar{\gamma} \triangleq E\{\gamma(n)\} = 3E\{a^2(n)\} - r \quad (2.20)$$

e

$$\xi \triangleq E\{\gamma^2(n)\} = 3r E\{a^2(n)\} + r^2. \quad (2.21)$$

Substituindo (2.11) em (1.7) (pág. 4), pode-se escrever

$$e_a(n) \approx \eta(n)e_{a,1}(n) + [1 - \eta(n)]e_{a,2}(n), \quad (2.22)$$

que, pressupondo a hipótese R2, é válida em todo instante e coincide com a Equação (2.9) obtida para o caso supervisionado.

Da mesma forma que no caso supervisionado, substituindo (2.17) e (2.18) em (2.22), chega-se à expressão para o erro da combinação

$$e(n) \approx \eta(n)e_1(n) + [1 - \eta(n)]e_2(n). \quad (2.23)$$

Cabe observar que nesse caso a Equação (2.23) vale em todo instante se R2 e os modelos (2.17) e (2.18) forem satisfeitos. É usual, no entanto, considerar sua validade em regime [Silva e Nascimento, 2008b].

De modo unificado, os erros para o caso supervisionado ou autodidata satisfazem

$$e(n) \approx \kappa(n)e_a(n) + \varphi(n), \quad (2.24)$$

com $\kappa(n) = 1$ e $\varphi(n) = v(n)$ para o caso supervisionado e $\kappa(n) = \gamma(n)$ e $\varphi(n) = \beta(n)$ para o caso do CMA. Em ambos os casos, $E\{\varphi(n)\} = 0$ e $\kappa(n)$ e $\varphi(n)$ são assumidos independentes de $\tilde{\mathbf{w}}(n-1)$ em regime. O erro *a priori* e o erro da combinação podem ser aproximados em ambos os casos por (2.22) e (2.23), respectivamente [Silva e Nascimento, 2008b].

2.3 O parâmetro de mistura ótimo em regime

Expressões para o valor médio do parâmetro de mistura ótimo definido como

$$\bar{\eta}_o(n) \triangleq E\{\eta_o(n)\} \quad (2.25)$$

e seu valor em regime

$$\bar{\eta}_o(\infty) \triangleq \lim_{n \rightarrow \infty} \bar{\eta}_o(n), \quad (2.26)$$

podem ser obtidas igualando a zero a derivada da função custo em relação a $\eta(n)$ e calculando o limite para $n \rightarrow \infty$. No caso supervisionado, deve-se igualar a zero o gradiente mostrado na Tabela 2.1 e no caso do CMA, o gradiente (2.5). De modo unificado, chega-se a

$$E\{e_g(n) [y_1(n) - y_2(n)]\} = 0, \quad (2.27)$$

em que

$$e_g(n) = e(n) = e_a(n) + v(n) \quad (2.28)$$

para o caso supervisionado e

$$e_g(n) = e_d(n) \quad (2.29)$$

para o caso do CMA. Como $e_d(n) \approx e_a(n)$ no caso do CMA, de modo geral, pode-se escrever

$$e_g(n) \approx e_a(n) + r(n), \quad (2.30)$$

em que $r(n) = v(n)$ no caso supervisionado e $r(n) = 0$ no caso do CMA.

Substituindo (1.9) (pág. 5) em (2.27) para o caso supervisionado ou (2.11) em (2.27) para o CMA, chega-se a

$$E\{e_g(n) [e_{a,2}(n) - e_{a,1}(n)]\} = 0. \quad (2.31)$$

Substituindo (2.22) em (2.30) e o resultado em (2.31) obtém-se

$$\begin{aligned} E\{\eta_o(n)[e_{a,2}(n) - e_{a,1}(n)]^2\} &= E\{[e_{a,2}^2(n) - e_{a,1}(n)e_{a,2}(n)]\} \\ &\quad + E\{r(n)[e_{a,2}(n) - e_{a,1}(n)]\}. \end{aligned} \quad (2.32)$$

Para continuar é necessário assumir que

R3. para ambos os casos, $\eta_o(n)$ é independente de $\kappa(n)$ e $e_{a,i}(n)$, $i = 1, 2$ em regime. Essa hipótese requer basicamente que o parâmetro de mistura ótimo seja independente do erro *a priori* em regime e, no caso autodidata, seja independente do sinal transmitido.

Usando o fato de que $r(n)$ é independente de $e_{a,i}(n)$, $i = 1, 2$ e $E\{r(n)\} = 0$, verifica-se que o último termo do lado direito de (2.32) é nulo. Calculando o limite para $n \rightarrow \infty$ de ambos os lados de (2.32) e usando R3, chega-se a

$$\lim_{n \rightarrow \infty} E\{\eta_o(n)\} \approx \frac{\lim_{n \rightarrow \infty} E\{e_{a,2}^2(n) - e_{a,1}(n)e_{a,2}(n)\}}{\lim_{n \rightarrow \infty} E\{[e_{a,2}(n) - e_{a,1}(n)]^2\}}, \quad (2.33)$$

que pode ser reescrita como

$$\bar{\eta}_o(\infty) \approx \frac{\zeta_2(\infty) - \zeta_{12}(\infty)}{\zeta_1(\infty) + \zeta_2(\infty) - 2\zeta_{12}(\infty)}. \quad (2.34)$$

A Equação (2.34) também pode ser escrita na forma

$$\bar{\eta}_o(\infty) \approx \frac{\Delta\zeta_2(\infty)}{\Delta\zeta_1(\infty) + \Delta\zeta_2(\infty)}, \quad (2.35)$$

com

$$\Delta\zeta_i(\infty) = \zeta_i(\infty) - \zeta_{12}(\infty). \quad (2.36)$$

2.4 O EMSE ótimo da combinação em regime

Elevando (2.22) ao quadrado com $\eta(n) = \eta_o(n)$, tomando a esperança de ambos os lados da equação resultante, calculando o limite para $n \rightarrow \infty$ e usando R4, obtém-se

$$\begin{aligned} \lim_{n \rightarrow \infty} E\{e_a^2(n)\} &= \lim_{n \rightarrow \infty} E\{\eta_o^2(n)\}E\{e_{a,1}^2(n)\} + \lim_{n \rightarrow \infty} E\{[1 - \eta_o(n)]^2\}E\{e_{a,2}^2(n)\} \\ &\quad + \lim_{n \rightarrow \infty} 2E\{\eta_o(n)[1 - \eta_o(n)]\}E\{e_{a,1}(n)e_{a,2}(n)\}. \end{aligned} \quad (2.37)$$

Para prosseguir assume-se que

R4. a variância de $\eta_o(n)$ é suficientemente pequena em regime para que $\lim_{n \rightarrow \infty} E\{\eta_o^2(n)\} \approx \bar{\eta}_o^2(\infty)$.

Usando R4, obtém-se a seguinte estimativa para o EMSE ótimo da combinação em regime

$$\zeta(\infty) \approx \bar{\eta}_o^2(\infty)\zeta_1(\infty) + [1 - \bar{\eta}_o(\infty)]^2\zeta_2(\infty) + 2\bar{\eta}_o(\infty)[1 - \bar{\eta}_o(\infty)]\zeta_{12}(\infty). \quad (2.38)$$

Substituindo a expressão para $\bar{\eta}_o(\infty)$ dada por (2.34) em (2.38), obtém-se

$$\zeta(\infty) \approx \zeta_2(\infty) - \bar{\eta}_o(\infty)(\zeta_2(\infty) - \zeta_{12}(\infty)), \quad (2.39)$$

que pode ser reescrita como

$$\zeta(\infty) \approx \zeta_2(\infty) - \frac{\Delta\zeta_2^2(\infty)}{\Delta\zeta_1(\infty) + \Delta\zeta_2(\infty)}. \quad (2.40)$$

A precisão das expressões (2.35) e (2.40) depende da precisão das expressões utilizadas para $\zeta_i(\infty)$, $i = 1, 2$ e $\zeta_{12}(\infty)$. Essas expressões foram obtidas também em [Arenas-García et al., 2006a, Eqs.(29) e (33)] para a combinação convexa de dois algoritmos LMS. A diferença é que na combinação convexa, $\eta(n)$ e conseqüentemente $\bar{\eta}_o(\infty)$ estão restritos ao intervalo $[0, 1]$.

2.5 Expressões para o EMSE em regime de filtros adaptativos

Existem diversos trabalhos sobre a análise de *tracking* de algoritmos adaptativos na literatura (por exemplo, Sayed [2003], Eleftheriou e Falconer [1986], Slock [1993], Eweda [1994], Haykin [2002], Yousef e Sayed [2001] e referências contidas). Na Tabela 2.2, são mostradas expressões analíticas para o EMSE do LMS [Sayed, 2003, Haykin, 2002], NLMS [Sayed, 2003, Slock, 1993], RLS [Sayed, 2003] e CMA [Silva e Nascimento, 2008a]. Nessa tabela, μ_i representa o passo de adaptação para o algoritmo LMS ou CMA, $\tilde{\mu}_i$ representa o passo de adaptação do NLMS, λ_i representa o fator de esquecimento do RLS e $\alpha_u \triangleq E\{\|\mathbf{u}(n)\|^{-2}\}$. Para $u(n)$ gaussiano e M suficientemente grande, α_u pode ser estimado como $1/[\sigma_u^2(M - 2)]$, em que $\sigma_u^2 = E\{u^2(n)\}$ é a potência do sinal de entrada [Costa e Bermudez, 2002].

Tabela 2.2: Expressões analíticas para o EMSE dos algoritmos considerados.

Algoritmo	$\zeta_i(\infty)$
LMS	$\frac{\mu_i \sigma_v^2 \text{Tr}(\mathbf{R}) + \mu_i^{-1} \text{Tr}(\mathbf{Q})}{2 - \mu_i \text{Tr}(\mathbf{R})}$
NLMS	$\frac{\text{Tr}(\mathbf{R}) [\tilde{\mu}_i \sigma_v^2 \alpha_u + \tilde{\mu}_i^{-1} \text{Tr}(\mathbf{Q})]}{2 - \tilde{\mu}_i}$
RLS	$\frac{\sigma_v^2 (1 - \lambda_i) M + (1 - \lambda_i)^{-1} \text{Tr}(\mathbf{Q}\mathbf{R})}{2}$
CMA	$\frac{\mu_i \sigma_\beta^2 \text{Tr}(\mathbf{R}) + \mu_i^{-1} \text{Tr}(\mathbf{Q})}{2\bar{\gamma} - \mu_i \text{Tr}(\mathbf{R})\xi}$

O EMSE cruzado $\zeta_{12}(\infty)$ para a combinação de dois algoritmos LMS foi obtido em [Arenas-García et al., 2006a] usando a conservação de energia. Usando o método tradicional¹, expressões de $\zeta_{12}(\infty)$ para a combinação de dois algoritmos RLS, um algoritmo RLS com um LMS e dois algoritmos CMA foram obtidas em [Silva e Nascimento, 2008b]. A extensão da análise de [Arenas-García et al., 2006a] para obter uma expressão de $\zeta_{12}(\infty)$ para a combinação de dois algoritmos NLMS é direta e encontra-se no Apêndice B. Esses resultados são mostrados na Tabela 2.3. Nessa tabela, $\Sigma \triangleq [(1 - \lambda_1)\mathbf{I} + \mu_2\mathbf{R}]^{-1}\mathbf{R}$.

Tabela 2.3: Expressões analíticas para o EMSE cruzado dos algoritmos considerados.

Combinação	$\zeta_{12}(\infty)$
μ_1 -LMS e μ_2 -LMS	$\frac{\mu_1 \mu_2 \sigma_v^2 \text{Tr}(\mathbf{R}) + \text{Tr}(\mathbf{Q})}{\mu_1 + \mu_2 - \mu_1 \mu_2 \text{Tr}(\mathbf{R})}$
$\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS	$\frac{\text{Tr}(\mathbf{R}) [\tilde{\mu}_1 \tilde{\mu}_2 \sigma_v^2 \alpha_u + \text{Tr}(\mathbf{Q})]}{\tilde{\mu}_1 + \tilde{\mu}_2 - \tilde{\mu}_1 \tilde{\mu}_2}$
λ_1 -RLS e λ_2 -RLS	$\frac{(1 - \lambda_1)(1 - \lambda_2) M \sigma_v^2 + \text{Tr}(\mathbf{Q}\mathbf{R})}{(1 - \lambda_1) + (1 - \lambda_2)}$
λ_1 -RLS e μ_2 -LMS	$\mu_2 \tilde{\lambda}_1 \sigma_v^2 \text{Tr}(\Sigma) + \text{Tr}(\mathbf{Q}\Sigma)$
μ_1 -CMA e μ_2 -CMA	$\frac{\mu_1 \mu_2 \sigma_\beta^2 \text{Tr}(\mathbf{R}) + \text{Tr}(\mathbf{Q})}{\bar{\gamma}(\mu_1 + \mu_2) - \mu_1 \mu_2 \text{Tr}(\mathbf{R})\xi}$

Para a combinação de dois algoritmos RLS, uma outra expressão de $\zeta_i(\infty)$ foi obtida em [Eleftheriou e Falconer, 1986]. Através de simulações pode-se notar que essa expressão é mais precisa do que a mostrada na Tabela 2.2, principalmente quando se usa fatores de esquecimento menores. No Apêndice C, o resultado de [Eleftheriou e Falconer, 1986] é revisitado e estendido para se obter uma outra expressão para $\zeta_{12}(\infty)$ usando hipóteses semelhantes. Esses resultados

¹No método tradicional obtém-se um expressão recursiva para a matriz de autocorrelação do vetor de erro de coeficientes.

são mostrados na Tabela 2.4, onde $\tilde{\lambda}_i = 1 - \lambda_i$ e $\gamma_u = \sigma_{u^2}^2 / \sigma_u^2$, em que σ_x^2 representa a variância de x . Para $u(n)$ gaussiano, $\gamma_u = 2$.

Tabela 2.4: Expressões analíticas mais precisas para o EMSE do RLS.

$\zeta_i(\infty)$	$\frac{1 - \lambda_i}{1 + \lambda_i} \left[1 + \frac{1 - \lambda_i}{1 + \lambda_i} \gamma_u \right] M \sigma_v^2 + \frac{\text{Tr}(\mathbf{QR})}{1 - \lambda_i^2}$
$\zeta_{12}(\infty)$	$\frac{\tilde{\lambda}_1 \tilde{\lambda}_2 \left[1 + \frac{\tilde{\lambda}_1 \tilde{\lambda}_2}{1 - \lambda_1 \lambda_2} \gamma_u \right] M \sigma_v^2 + \text{Tr}(\mathbf{QR})}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_1 \tilde{\lambda}_2}$

2.6 Resultados em ambiente estacionário

Substituindo as expressões das Tabelas 2.2 e 2.3 com $\mathbf{Q} = 0$ nas Equações (2.34) e (2.39) podem-se obter expressões analíticas para $\bar{\eta}_o(\infty)$ e $\zeta(\infty)$ num ambiente estacionário. Alguns desses resultados são mostrados na Tabela 2.5. Nessa tabela, $\delta_{LL} = \delta_{CC} \triangleq \mu_2 / \mu_1$ representam a razão entre os passos de adaptação dos filtros LMS ou CMA. Analogamente, $\delta_{NN} \triangleq \tilde{\mu}_2 / \tilde{\mu}_1$ e $\delta_{RR} \triangleq (1 - \lambda_2) / (1 - \lambda_1)$. Em todos os casos, $0 < \delta_{XX} < 1$, onde δ_{XX} representa, de forma geral, a relação entre passos de adaptação ou fatores de esquecimento.

Tabela 2.5: Expressões analíticas para $\bar{\eta}_o(\infty)$ e $\zeta(\infty)$ no caso estacionário.

Combinação	$\bar{\eta}_o(\infty)$	$\zeta(\infty)$
μ_1 -LMS e μ_2 -LMS	$\frac{\delta_{LL}[2 - \mu_1 \text{Tr}(\mathbf{R})]}{2(\delta_{LL} - 1)}$	$\frac{1}{2} \left[\frac{\mu_2 \sigma_v^2 \text{Tr}(\mathbf{R})}{\delta_{LL} + 1 - \mu_2 \text{Tr}(\mathbf{R})} \right]$
$\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS	$\frac{\delta_{NN}[2 - \tilde{\mu}_1]}{2(\delta_{NN} - 1)}$	$\frac{1}{2} \left[\frac{\text{Tr}(\mathbf{R}) \tilde{\mu}_2 \sigma_v^2 \alpha_u}{\delta_{NN} + 1 - \tilde{\mu}_2} \right]$
λ_1 -RLS e λ_2 -RLS	$\frac{\delta_{RR}}{\delta_{RR} - 1}$	$\frac{1}{2} \left[\frac{(1 - \lambda_2) M \sigma_v^2}{\delta_{RR} + 1} \right]$
μ_1 -CMA e μ_2 -CMA	$\frac{\delta_{CC}[2 - \mu_1 \text{Tr}(\mathbf{R}) \xi(\bar{\gamma})^{-1}]}{2(\delta_{CC} - 1)}$	$\frac{1}{2} \left[\frac{\mu_2 \sigma_\beta^2 \text{Tr}(\mathbf{R})}{(\delta_{CC} + 1) \bar{\gamma} - \mu_2 \text{Tr}(\mathbf{R}) \xi} \right]$

Algumas propriedades interessantes podem ser observadas a partir desses resultados:

- i) $\bar{\eta}_o(\infty)$ é negativo para todas as combinações consideradas na tabela. Cabe observar que para assegurar a estabilidade do filtro μ_1 -LMS, μ_1 é escolhido no intervalo $0 < \mu_1 < 2 / \text{Tr}(\mathbf{R})$ [Diniz, 2008]. Analogamente para o CMA, $0 < \mu_1 < 2\bar{\gamma} / 3 \text{Tr}(\mathbf{R}) \xi$ [Nascimento e Silva, 2008, Eq. (14)] e para o NLMS, $0 < \tilde{\mu}_1 < 2$;
- ii) Quando $\delta_{XX} \approx 1$, $\zeta(\infty) \approx \zeta_2(\infty) / 2$ para todas as combinações mostradas na tabela. Como $\zeta_2(\infty) < \zeta_1(\infty)$ para todas as combinações, a combinação afim tem um EMSE 3 dB menor do que o menor EMSE dos filtros componentes. Nessa situação, $\bar{\eta}_o(\infty) \rightarrow -\infty$.

A propriedade i) foi também observada em [Bershad et al., 2008] para a combinação de dois algoritmos LMS considerando entrada branca gaussiana e hipóteses equivalentes à escolha do passo μ_1 para máxima velocidade de convergência. Se considerado $\mu_1 = 1/\text{Tr}(\mathbf{R})$ para máxima velocidade de convergência na expressão de $\bar{\eta}_o$ para a combinação de dois filtros LMS da Tabela 2.5, o valor de parâmetro de mistura ótimo obtido será

$$\bar{\eta}_o(\infty) = \frac{\delta_{LL}}{2(\delta_{LL} - 1)}, \quad (2.41)$$

que coincide com o resultado obtido em [Bershad et al., 2008].

Uma explicação para o fato de $|\bar{\eta}_o(\infty)| \rightarrow \infty$ quando $\delta_{XX} \approx 1$ pode ser obtida usando as Equações (2.23) e (2.24) para escrever o erro da combinação como

$$e(n) \approx e_2(n) + \eta(n)\kappa(n)[e_{a,1}(n) - e_{a,2}(n)]. \quad (2.42)$$

Usando as Equações (1.16) e (1.17), obtém-se

$$e(n) \approx \underbrace{e_2(n)}_{d_\eta(n)} + \eta(n) \underbrace{\kappa(n)[\mathbf{w}_2(n) - \mathbf{w}_1(n)]^T \mathbf{u}(n)}_{-u_\eta(n)}. \quad (2.43)$$

Do ponto de vista do cálculo de $\eta(n)$, $d_\eta(n)$ representa o sinal que deve ser estimado e $u_\eta(n)$ representa o sinal de entrada. Se assumirmos que $\mathbf{w}_i$ varia lentamente em comparação com η , a Equação (2.43) tem uma interpretação geométrica simples como mostrado na Figura 2.1. A combinação afim busca o melhor vetor na linha $\mathbf{w}_2 + \eta(\mathbf{w}_1 - \mathbf{w}_2)$. Na Figura 2.1-(a), a melhor combinação linear de $\mathbf{w}_1$ e $\mathbf{w}_2$ é $\mathbf{w}$. Quando $\delta_{XX} \approx 1$, os vetores de coeficientes dos filtros ficam muito parecidos, ou seja, $\mathbf{w}_1 \approx \mathbf{w}_2$ (Figura 2.1-(b)). Nesse caso, $\eta(n)$ deve ser grande o suficiente para levar o vetor de coeficientes equivalente da combinação o mais próximo possível de $\mathbf{w}$. Portanto, se $\delta_{XX} \rightarrow 1$, $(\mathbf{w}_1 - \mathbf{w}_2) \rightarrow \mathbf{0}$ e $|\eta(n)| \rightarrow \infty$.

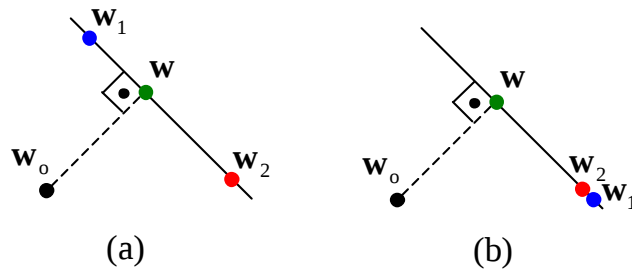


Figura 2.1: Interpretação geométrica da combinação afim.

Com as expressões da Tabela 2.4 para o caso da combinação de dois filtros RLS, fica difícil observar as propriedades i) e ii) analiticamente pois as expressões de $\bar{\eta}_o(\infty)$ e $\zeta(\infty)$ não podem ser simplificadas. Entretanto, as propriedades podem ser observadas graficamente na Seção 2.8 onde são mostradas as curvas teóricas usando essas expressões.

À medida que o valor de δ_{XX} é diminuído, o ganho de 3 dB em relação ao filtro 2 que ocorre quando $\delta_{XX} \approx 1$ também diminui. O mesmo ocorre com o valor do EMSE do filtro 2. Assim, a diminuição desse ganho não significa necessariamente que o EMSE da combinação aumenta. Esse resultado pode ser observado através dos gráficos de EMSE dos filtros componentes e da combinação. Considerando que para garantir a estabilidade dos filtros, $\mu_i \ll 2/\text{Tr}(\mathbf{R})$ para o LMS, $\tilde{\mu}_i \ll 2$ para o NLMS [Diniz, 2008] e $\mu_i \ll 2\bar{\gamma}/3\text{Tr}(\mathbf{R})\xi$ para o CMA [Nascimento e Silva, 2008, Eq. (14)], as expressões de EMSE dos filtros componentes no caso estacionário podem ser aproximadas por

$$\zeta_1(\infty) = \frac{K}{2} \quad (2.44)$$

e

$$\zeta_2(\infty) = \frac{\delta_{XX}K}{2}, \quad (2.45)$$

sendo K uma constante que depende do algoritmo usado e de parâmetros do cenário considerado. Para o LMS, por exemplo, $K = \mu_i \sigma_v^2 \text{Tr}(\mathbf{R})$. A expressão de EMSE da combinação pode ser aproximada por

$$\zeta(\infty) = \frac{1}{2} \frac{\delta_{XX}K}{\delta_{XX} + 1}. \quad (2.46)$$

Na Figura 2.2, estão representadas as curvas de $\zeta_i(\infty)$, $i = 1, 2$ e $\zeta(\infty)$ aproximadas pelas expressões (2.44), (2.45) e (2.46). Como pode ser observado, a situação do ganho de 3 dB que ocorre quando $\delta_{XX} \approx 1$ corresponde ao máximo EMSE da combinação. À medida que δ_{XX} é diminuído, o ganho em relação à $\zeta_2(\infty)$ diminui, porém $\zeta_2(\infty)$ também diminui, fazendo com que $\zeta(\infty)$, em termos absolutos, diminua. Quando δ_{XX} é pequeno, o ganho da combinação é praticamente inexistente, já que esta tem um desempenho muito próximo ao do filtro 2.

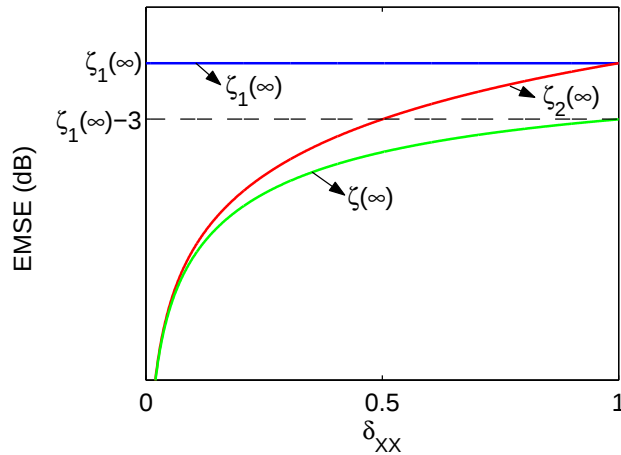


Figura 2.2: Comparação entre $\zeta(\infty)$ e $\zeta_i(\infty)$, $i = 1, 2$.

Apesar do ganho de EMSE proporcionado pela combinação afim quando $\delta_{XX} \rightarrow 1$, a combinação, nesse caso, apresenta uma velocidade de convergência muito baixa. Isso ocorre pois na adaptação de $\eta(n)$, é usado o termo $[y_1(n) - y_2(n)]$, que é muito pequeno caso os passos

de adaptação (ou fatores de esquecimento) sejam muito próximos. Essa propriedade pode ser utilizada de forma mais prática se o valor do parâmetro $\eta(n)$ for considerado constante e igual à $\bar{\eta}_o(\infty)$ mostrados na Tabela 2.5, para cada tipo de combinação considerado. Os valores dos parâmetros necessários para o cálculo de $\bar{\eta}_o(\infty)$ são facilmente obtidos já que $\text{Tr}(\mathbf{R}) = M\sigma_u^2$, sendo σ_u^2 a potência do sinal de entrada do filtro e no caso do CMA, σ_β^2 , $\bar{\gamma}$ e ξ são constantes que dependem de estatísticas do sinal transmitido. Assim, a saída combinação passa a ser calculada por

$$y(n) = \bar{\eta}_o(\infty)y_1(n) + [1 - \bar{\eta}_o(\infty)]y_2(n). \quad (2.47)$$

Como $\bar{\eta}_o(\infty)$ é uma constante, a combinação terá convergido assim que o filtro mais lento convergir. Como é necessário fazer $\delta_{XX} \approx 1$ para que o ganho de EMSE ocorra, as velocidades de convergência dos filtros componentes serão muito próximas e a velocidade de convergência da combinação será praticamente igual à dos filtros componentes. Assim, usando a combinação de filtros componentes com passos de adaptação (ou fatores de esquecimento) próximos e o parâmetro de mistura fixo, é possível obter um determinado nível de EMSE com uma velocidade de convergência um pouco maior que a oferecida por um filtro único, às custas da duplicação da complexidade computacional.

2.6.1 Possibilidade de ganho maior que 3 dB

Uma idéia para aumentar o ganho de EMSE que a combinação afim oferece, é combinar as saídas de duas combinações afim usando filtros componentes com passos de adaptação (ou fatores de esquecimento) próximos, como mostrado na Figura 2.3.

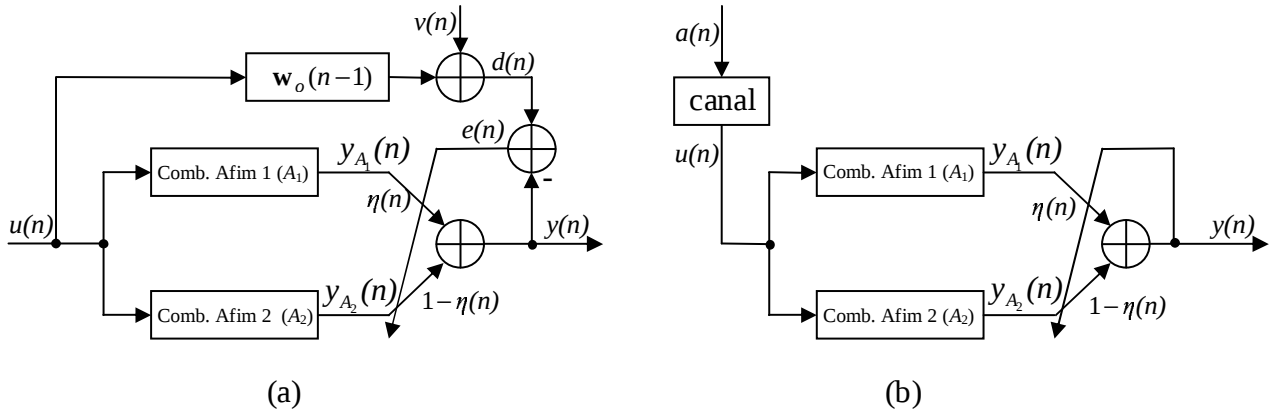


Figura 2.3: Esquema da combinação afim de duas combinações afim. (a) Caso supervisionado. (b) Caso autodidata.

Conforme o resultado obtido em (2.40), a combinação afim A_1 , composta pelos filtros 1 e 2, tem seu EMSE dado por

$$\zeta_{A_1}(\infty) \approx \zeta_2(\infty) - \frac{\Delta\zeta_2^2(\infty)}{\Delta\zeta_1(\infty) + \Delta\zeta_2(\infty)}. \quad (2.48)$$

Esse resultado pode ser estendido para o EMSE da combinação A_2 , composta pelos filtros 3 e 4, obtendo

$$\zeta_{A_2}(\infty) \approx \zeta_4(\infty) - \frac{\Delta\zeta_4^2(\infty)}{\Delta\zeta_3(\infty) + \Delta\zeta_4(\infty)}, \quad (2.49)$$

sendo $\Delta\zeta_3(\infty) \triangleq \zeta_3(\infty) - \zeta_{34}(\infty)$ e $\Delta\zeta_4(\infty) \triangleq \zeta_4(\infty) - \zeta_{34}(\infty)$. O cálculo de $\zeta_3(\infty)$, $\zeta_4(\infty)$ e $\zeta_{34}(\infty)$ pode ser feito estendendo os resultados obtidos para $\zeta_1(\infty)$, $\zeta_2(\infty)$ e $\zeta_{12}(\infty)$, mostrados nas Tabelas 2.2 e 2.3. O EMSE ótimo da combinação de A_1 e A_2 é dado por

$$\zeta(\infty) \approx \zeta_{A_2}(\infty) - \frac{\Delta\zeta_{A_2}^2(\infty)}{\Delta\zeta_{A_1}(\infty) + \Delta\zeta_{A_2}(\infty)}, \quad (2.50)$$

sendo $\Delta\zeta_{A_1}(\infty) \triangleq \zeta_{A_1}(\infty) - \zeta_{A_1A_2}(\infty)$ e $\Delta\zeta_{A_2}(\infty) \triangleq \zeta_{A_2}(\infty) - \zeta_{A_1A_2}(\infty)$. O cálculo de (2.50) requer uma expressão para o EMSE cruzado $\zeta_{A_1A_2}(\infty)$, definido por

$$\zeta_{A_1A_2}(\infty) = \lim_{n \rightarrow \infty} E \{e_{a,A_1}(n)e_{a,A_2}(n)\}. \quad (2.51)$$

Essa expressão pode ser obtida estendendo o resultado de (2.22) para as combinações A_1 e A_2 , ou seja,

$$e_{a,A_1}(n) \approx \eta_{A_1}(n)e_{a,1}(n) + [1 - \eta_{A_1}(n)]e_{a,2}(n) \quad (2.52)$$

e

$$e_{a,A_2}(n) \approx \eta_{A_2}(n)e_{a,3}(n) + [1 - \eta_{A_2}(n)]e_{a,4}(n). \quad (2.53)$$

Substituindo (2.52) e (2.53) em (2.51), obtém-se

$$\begin{aligned} \zeta_{A_1A_2}(\infty) \approx \lim_{n \rightarrow \infty} E \{ & \eta_{A_1}(n)\eta_{A_2}(n)e_{a,1}(n)e_{a,3}(n) + \eta_{A_1}(n)[1 - \eta_{A_2}(n)]e_{a,1}(n)e_{a,4}(n) \\ & + [1 - \eta_{A_1}(n)]\eta_{A_2}(n)e_{a,2}(n)e_{a,3}(n) \\ & + [1 - \eta_{A_1}(n)][1 - \eta_{A_2}(n)]e_{a,2}(n)e_{a,4}(n) \}. \end{aligned} \quad (2.54)$$

Calculando a esperança e o limite e considerando que as combinações são adaptadas com o parâmetro de mistura ótimo, chega-se a

$$\begin{aligned} \zeta_{A_1A_2}(\infty) \approx & \bar{\eta}_{o,A_1}(\infty)\bar{\eta}_{o,A_2}(\infty)\zeta_{13}(\infty) + \bar{\eta}_{o,A_1}(\infty)[1 - \bar{\eta}_{o,A_2}(\infty)]\zeta_{14}(\infty) \\ & + [1 - \bar{\eta}_{o,A_1}(\infty)]\bar{\eta}_{o,A_2}(\infty)\zeta_{23}(\infty) \\ & + [1 - \bar{\eta}_{o,A_1}(\infty)][1 - \bar{\eta}_{o,A_2}(\infty)]\zeta_{24}(\infty), \end{aligned} \quad (2.55)$$

sendo os termos $\zeta_{13}(\infty)$, $\zeta_{14}(\infty)$, $\zeta_{23}(\infty)$ e $\zeta_{24}(\infty)$ calculados através da extensão dos resultados da Tabela 2.3. Os valores dos parâmetros de mistura ótimos podem ser obtidos usando o resultado de (2.35), ou seja,

$$\bar{\eta}_{o,A_1}(\infty) \approx \frac{\Delta\zeta_2(\infty)}{\Delta\zeta_1(\infty) + \Delta\zeta_2(\infty)} \quad (2.56)$$

e

$$\bar{\eta}_{o,A_2}(\infty) \approx \frac{\Delta\zeta_4(\infty)}{\Delta\zeta_3(\infty) + \Delta\zeta_4(\infty)}. \quad (2.57)$$

O EMSE da combinação é obtido usando (2.56) e (2.57) para calcular (2.55) e usando esse resultado em (2.50). Em seguida, deve-se substituir as expressões de $\zeta_i(\infty)$ e $\zeta_{ij}(\infty)$, $i, j = 1, 2, 3, 4$ referentes ao algoritmo utilizado. Foi verificado, usando um programa de computação simbólica, que quando são considerados filtros de mesma família, com passos de adaptação (ou fatores de esquecimento) próximos, o EMSE $\zeta_{A_1, A_2}(\infty)$ é 3/8 menor que o EMSE dos filtros componentes, o que representa um ganho de 4,26 dB. Esse resultado indica que se esse esquema for estendido para mais filtros (8, 16, etc.), o ganho de EMSE obtido terá um limite, deixando de aumentar significativamente a partir de um determinado número de filtros componentes combinados. Entretanto, a complexidade da forma apresentada para o cálculo do EMSE final torna-se maior à medida que é aumentado o número de combinações. Assim, é necessário um estudo de uma forma mais simplificada de obter esses resultados.

2.7 Resultados em ambiente não-estacionário

Num ambiente não-estacionário, substituindo as expressões das Tabelas 2.2 e 2.3 nas Equações (2.34) e (2.39), pode-se notar que a maior redução do EMSE da combinação afim em relação ao EMSE dos filtros componentes também ocorre quando $\zeta_1(\infty) = \zeta_2(\infty)$. Essa condição pode ocorrer em duas situações:

- i) quando os filtros componentes tiverem passos de adaptação (ou fatores de esquecimento) distintos e $\text{Tr}(\mathbf{Q}) = q_{12}$, sendo q_{12} um escalar, dado pelas expressões mostradas na Tabela 2.6, para cada tipo de combinação considerado, ou
- ii) quando os filtros componentes tiverem passos de adaptação (ou fatores de esquecimento) muito próximos, ou seja, quando $\delta_{XX} \approx 1$.

Na Tabela 2.6, são mostradas expressões analíticas para q_{12} e $\zeta(\infty)$ da combinação nas duas situações descritas. É possível notar que a redução do EMSE da combinação em relação ao EMSE dos filtros componentes é limitada em 3 dB em ambas as situações. Na situação i), a máxima redução ocorre quando $\delta_{XX} \rightarrow 0$, ou seja, quando o passo de adaptação do filtro 1 é muito maior que o do filtro 2 no caso do LMS, CMA ou NLMS ou quando $1 - \lambda_1$ é muito maior que $1 - \lambda_2$ no caso do RLS. Já na situação ii), a máxima redução de EMSE ocorre quando o ambiente tende a ser estacionário, ou seja, quando $\mathbf{Q} \approx \mathbf{0}$.

Novamente, com as expressões da Tabela 2.4 para o caso da combinação de dois filtros RLS fica difícil observar esses resultados analiticamente. Entretanto, na Seção 2.8, esses resultados podem ser observados graficamente já que as curvas teóricas foram obtidas com essas expressões. Para o caso da combinação de um LMS com um RLS, é possível observar que há também um limite de redução de 3 dB do EMSE da combinação em relação ao EMSE

Tabela 2.6: Expressões analíticas para q_{12} e $\zeta(\infty)$ para as situações (i) e (ii) no caso não-estacionário.

Combinação	(i)		(ii)
	q_{12}	$\zeta(\infty)$	$\zeta(\infty)$
μ_1 -LMS e μ_2 -LMS	$\mu_1\mu_2\sigma_v^2\text{Tr}(\mathbf{R})$	$\frac{\zeta_2(\infty)}{2} + \frac{2\delta_{LL}\zeta_2(\infty)}{(1 + \delta_{LL})^2}$	$\frac{\zeta_2(\infty)}{2} + \frac{\sigma_v^2\text{Tr}(\mathbf{R})\text{Tr}(\mathbf{Q})}{2\zeta_2(\infty)}$
$\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS	$\tilde{\mu}_1\tilde{\mu}_2\sigma_v^2\alpha_u$	$\frac{\zeta_2(\infty)}{2} + \frac{2\delta_{NN}\zeta_2(\infty)}{(1 + \delta_{NN})^2}$	$\frac{\zeta_2(\infty)}{2} + \frac{\sigma_v^2[\text{Tr}(\mathbf{R})]^2\text{Tr}(\mathbf{Q})\alpha_u}{2\zeta_2(\infty)}$
λ_1 -RLS e λ_2 -RLS	$\tilde{\lambda}_1\tilde{\lambda}_2\sigma_v^2M$	$\frac{\zeta_2(\infty)}{2} + \frac{2\delta_{RR}\zeta_2(\infty)}{(1 + \delta_{RR})^2}$	$\frac{\zeta_2(\infty)}{2} + \frac{M\sigma_v^2\text{Tr}(\mathbf{Q}\mathbf{R})}{2\zeta_2(\infty)}$
μ_1 -CMA e μ_2 -CMA	$\mu_1\mu_2\sigma_\beta^2\text{Tr}(\mathbf{R})$	$\frac{\zeta_2(\infty)}{2} + \frac{2\delta_{CC}\zeta_2(\infty)}{(1 + \delta_{CC})^2}$	$\frac{\zeta_2(\infty)}{2} + \frac{\sigma_\beta^2\text{Tr}(\mathbf{R})\text{Tr}(\mathbf{Q})}{2\bar{\gamma}\zeta_2(\infty)}$

dos filtros componentes mas as condições em que esta redução ocorre são diferentes devido à distinção entre os algoritmos componentes da combinação. Esse resultado também é mostrado graficamente na Seção 2.8.

O comportamento da combinação no caso não-estacionário também pode ser estudado através das curvas de NSD. Na Figura 2.4-(a), são mostradas as curvas de NSD dos filtros componentes, da combinação afim ($\text{NSD}_a(\infty)$) e da combinação convexa ($\text{NSD}_c(\infty)$) para uma combinação de dois algoritmos CMA com passos de adaptação distintos (situação i). Na Figura 2.4-(b), são mostrados os valores de $\bar{\eta}_o(\infty)$ para a combinação afim e para a combinação convexa. O formato das curvas é semelhante para os outros tipos de combinação de algoritmos de mesma família. Conforme a análise feita através das expressões, a maior redução de EMSE ocorre quando $\text{Tr}(\mathbf{Q}) = q_{12}$ que faz com que $\zeta_1(\infty) = \zeta_2(\infty)$ e conseqüentemente $\text{NSD}_1(\infty) = \text{NSD}_2(\infty)$. Entretanto, conforme pode ser visto Figura 2.4-(a), apesar da máxima redução de EMSE, essa situação não é a mais favorável já que o NSD e conseqüentemente o EMSE da combinação não é mínimo. O EMSE mínimo da combinação ocorre quando $\text{Tr}(\mathbf{Q}) = q_1$ ou $\text{Tr}(\mathbf{Q}) = q_2$, representados na Figura 2.4-(a), o que faz com que respectivamente o filtro 1 ou o filtro 2 esteja ajustado com passo de adaptação (ou fator de esquecimento) ótimo e tenha desempenho ótimo. Nesse caso, a combinação tem EMSE mínimo, igual ao do filtro componente com desempenho ótimo. É possível observar ainda que nessa situação de passos de adaptação (ou fatores de esquecimento) distintos, o comportamento das combinações convexa e afim é muito parecido pois o valor de $\bar{\eta}_o(\infty)$ se encontra no intervalo $[0, 1]$ para $q_2 < \text{Tr}(\mathbf{Q}) < q_1$. Entretanto, quando $\text{Tr}(\mathbf{Q}) > q_1$ ou $\text{Tr}(\mathbf{Q}) < q_2$, $\bar{\eta}_o(\infty) > 1$ ou $\bar{\eta}_o(\infty) < 0$, respectivamente. Isso faz com que o desempenho da combinação afim seja um pouco superior ao da combinação convexa. No entanto, quanto mais distintos forem os passos de adaptação (ou fatores de esquecimento), menor a diferença entre as duas combinações.

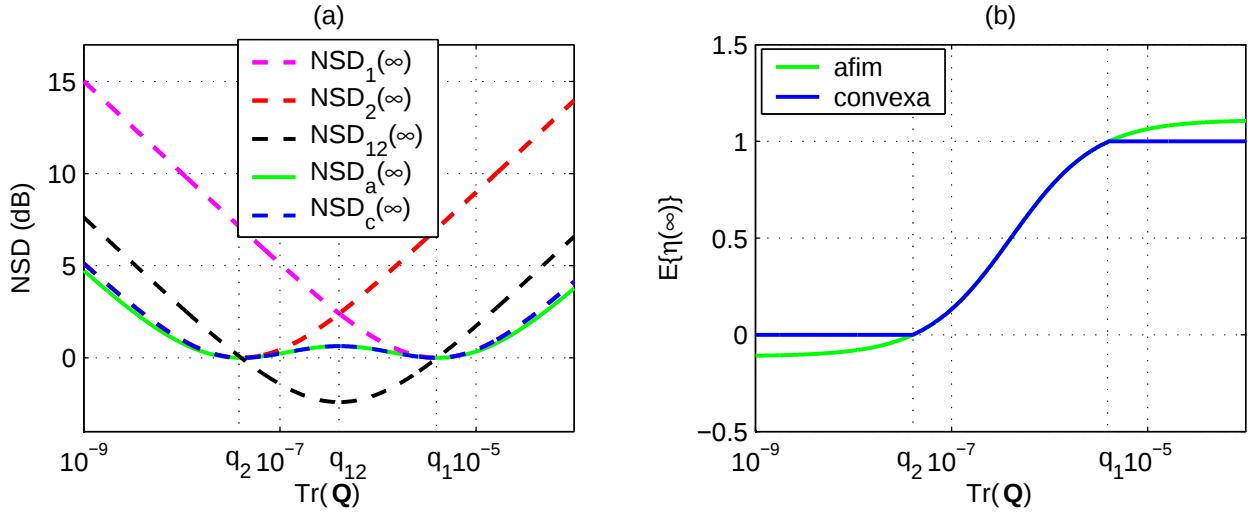


Figura 2.4: Curvas de NSD para as combinações afim (NSD_a) e convexa (NSD_c) de dois algoritmos CMA com passos de adaptação distintos. $M = 4$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Sinal de entrada com constelação $[-3, -1, 1, 3]$, $\bar{\gamma} = 6,8$, $\xi = 190,24$, $\sigma_\beta^2 = 28,8$.

No caso de passos de adaptação (ou fatores de esquecimento) próximos (situação ii), há uma diferença maior entre o desempenho da combinação afim e o da combinação convexa conforme pode ser visto na Figura 2.5. Nesse caso, a combinação convexa tem um desempenho igual ao dos filtros componentes enquanto a combinação afim apresenta um ganho de até 3 dB em relação aos filtros componentes quando $\text{Tr}(\mathbf{Q}) \approx 0$, conforme previsto pelas expressões analíticas. Há também um ganho da combinação afim quando $\text{Tr}(\mathbf{Q})$ é grande. Porém, na prática, os modelos de EMSE não funcionam nesse caso e o ganho é apenas uma conclusão teórica. É possível observar na Figura 2.5-(b), que o valor de $\bar{\eta}_o(\infty)$ da combinação afim chega a valores bem maiores que 1 e bem menores que 0, o que causa a grande diferença no desempenho das duas combinações.

2.8 Resultados de Simulação

Para verificar o comportamento da combinação afim e a validade da análise no caso supervisionado, foi considerada uma aplicação de identificação de sistemas e duas soluções ótimas formadas por $M = 7$ coeficientes aleatórios dados por

$$\mathbf{w}_{o,1}^T(0) = [0,90 \quad -0,54 \quad 0,21 \quad -0,03 \quad 0,78 \quad 0,52 \quad -0,09]$$

e

$$\mathbf{w}_{o,2}^T(0) = [0,25 \quad -0,47 \quad -0,38 \quad 0,05 \quad -0,18 \quad 0,79 \quad 0,15].$$

Nas simulações, quando utilizado sinal de entrada colorido, este foi gerado através de um filtro autoregressivo (AR) de primeira ordem com função de transferência $\sqrt{(1-a^2)}/(1-az^{-1})$,

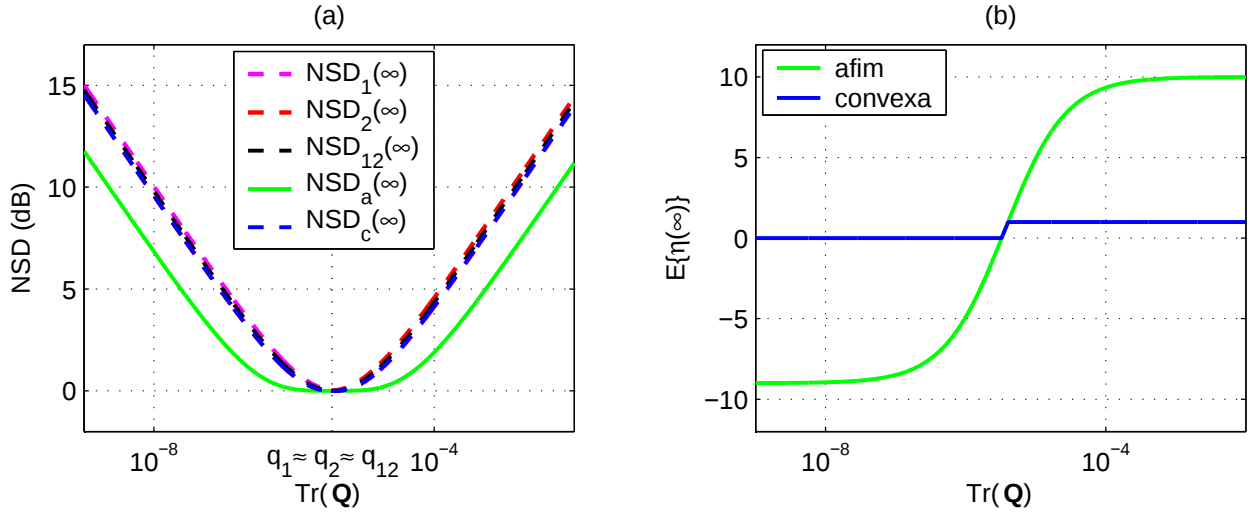


Figura 2.5: Curvas de NSD para as combinações afim (NSD_a) e convexa (NSD_c) de dois algoritmos CMA com passos de adaptação próximos. $M = 4$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,9$. Sinal de entrada com constelação $[-3, -1, 1, 3]$, $\bar{\gamma} = 6,8$, $\xi = 190,24$, $\sigma_\beta^2 = 28,8$.

com $a = 0,8$. Tanto para sinal de entrada colorido como para sinal de entrada branco, foi considerada variância $1/M$, o que leva a $\text{Tr}(\mathbf{R}) = 1$. Para formar o sinal desejado, é somado um sinal de ruído $v(n)$, com variância $\sigma_v^2 = 0,01$. Em todas as figuras do caso supervisionado, os valores teóricos do LMS e do NLMS foram calculados usando as expressões de $\zeta_i(\infty)$ e $\zeta_{12}(\infty)$ mostradas nas Tabelas 2.2 e 2.3 e para o RLS, foram usadas as expressões mostradas na Tabela 2.4.

Na Figura 2.6, são mostrados resultados para as combinações de algoritmos μ_1 -LMS com μ_2 -LMS, $\tilde{\mu}_1$ -NLMS com $\tilde{\mu}_2$ -NLMS e λ_1 -RLS com λ_2 -RLS para o caso estacionário com a solução ótima dada por $\mathbf{w}_{o,1}(0)$ e entrada colorida. As linhas cheias representam os valores teóricos. Para cada ponto, foi feita uma média de 50 realizações. Pode-se observar que nos três casos, quanto mais próximo de 1 for o valor de δ_{XX} , mais próximo de 0,5 é a razão $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1,2$, o que representa um ganho de 3 dB no EMSE da combinação em relação ao menor EMSE dos filtros componentes. A estimativa feita através da análise não é tão boa para o caso do NLMS porque as expressões de $\zeta_i(\infty)$ e $\zeta_{12}(\infty)$ para esse caso não são tão precisas quando é utilizado um número de coeficientes pequeno no filtro. Entretanto, quando é utilizado um maior número de coeficientes, a estimativa prevista pela análise tem uma concordância melhor com o resultado experimental conforme pode ser visto na Figura 2.7. Nessa figura, foi utilizada a combinação de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS com $M = 30$ coeficientes e solução ótima $\mathbf{w}_o$ formada por 30 coeficientes aleatórios com média nula e variância unitária. São mostrados também os resultados da Figura 2.6 para efeito de comparação. Pode-se notar que para um maior número de coeficientes, as hipóteses assumidas para obtenção de $\zeta_i(\infty)$ e $\zeta_{12}(\infty)$ são mais realistas e o resultado experimental coincide melhor com o resultado teórico.

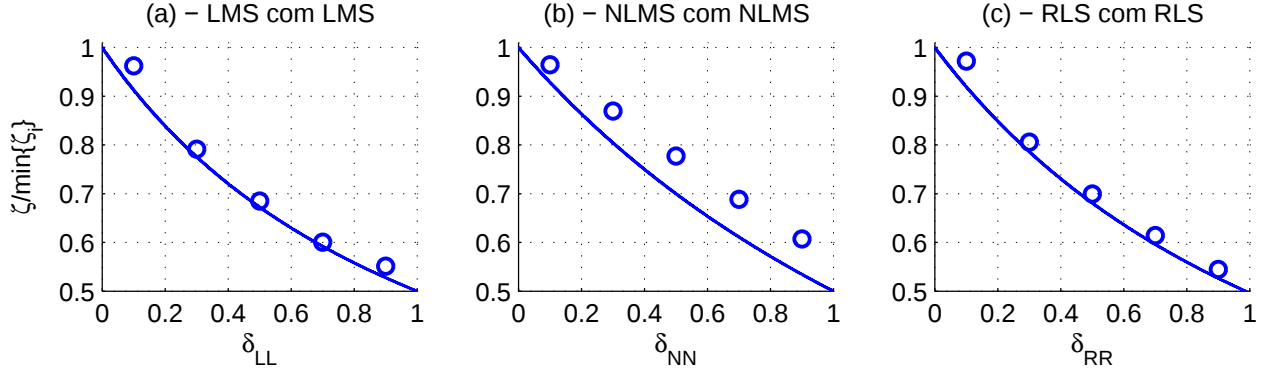


Figura 2.6: Combinação afim no caso supervisionado estacionário. Valores teóricos (linha cheia) e experimentais ($\circ$) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1, 2$ para 50 realizações com $M = 7$, $\sigma_v^2 = 0,01$ e diversos valores de δ_{XX} . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_{o,1}(0)$. (a) μ_1 -LMS com μ_2 -LMS, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$, $\mu_\eta = 1$. (b) $\tilde{\mu}_1$ -NLMS com $\tilde{\mu}_2$ -NLMS, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ e $\mu_\eta = 0,1$ para $\delta_{NN} = 0,1$, $\delta_{NN} = 0,3$ e $\delta_{NN} = 0,5$; $\mu_\eta = 0,5$ para $\delta_{NN} = 0,7$ e $\mu_\eta = 3$ para $\delta_{NN} = 0,9$. (c) λ_1 -RLS com λ_2 -RLS, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 1$.

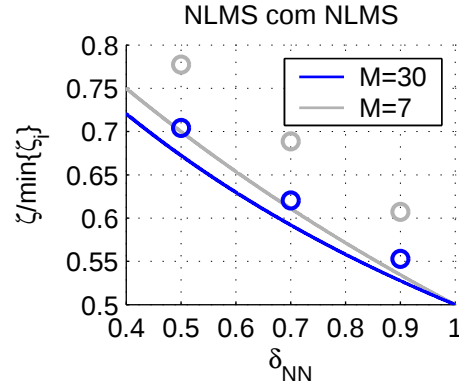


Figura 2.7: Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS com $\tilde{\mu}_1 = 0,1$ e $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ no caso supervisionado estacionário. Valores teóricos (linha cheia) e experimentais ($\circ$) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1, 2$ para 50 realizações com $\sigma_v^2 = 0,01$ e diversos valores de δ_{NN} . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_{o,1}(0)$. Em azul, $M = 30$, $\mu_\eta = 6$ e solução ótima formada por 30 coeficientes aleatórios de média 0 e variância 1. Em cinza, resultados obtidos na Figura 2.6.

Nas Figuras 2.8, 2.9 e 2.10, são mostrados o EMSE e o valor de $E\{\eta(n)\}$ ao longo das iterações no caso estacionário para as combinações usando dois algoritmos LMS, NLMS e RLS respectivamente. As linhas cheias representam os resultados experimentais e as linhas tracejadas representam os resultados teóricos. Nos três casos, foi considerado $\delta_{XX} = 0,1$ e foi feita uma média de 500 realizações. Inicialmente foi considerada a solução ótima $\mathbf{w}_{o,1}(0)$ com sinal de entrada branco e no meio da simulação alterou-se para $\mathbf{w}_{o,2}(0)$ e sinal de entrada colorido mantendo $\text{Tr}(\mathbf{R}) = 1$. Pode-se verificar que a análise faz uma boa estimativa dos valores de EMSE e $E\{\eta(n)\}$ em regime. Com relação à universalidade da combinação, pode-se notar que esta depende do ajuste do passo de adaptação μ_η . Nas Figuras 2.8 e 2.9, o

passo foi ajustado para a situação de entrada colorida, fazendo com que o comportamento da combinação fique próximo ao universal nessa condição e inferior ao universal na situação de entrada branca. Na Figura 2.10, o passo foi ajustado para a situação de entrada branca, fazendo com que o comportamento da combinação fique próximo ao universal nessa situação e inferior ao universal quando considerada entrada colorida.

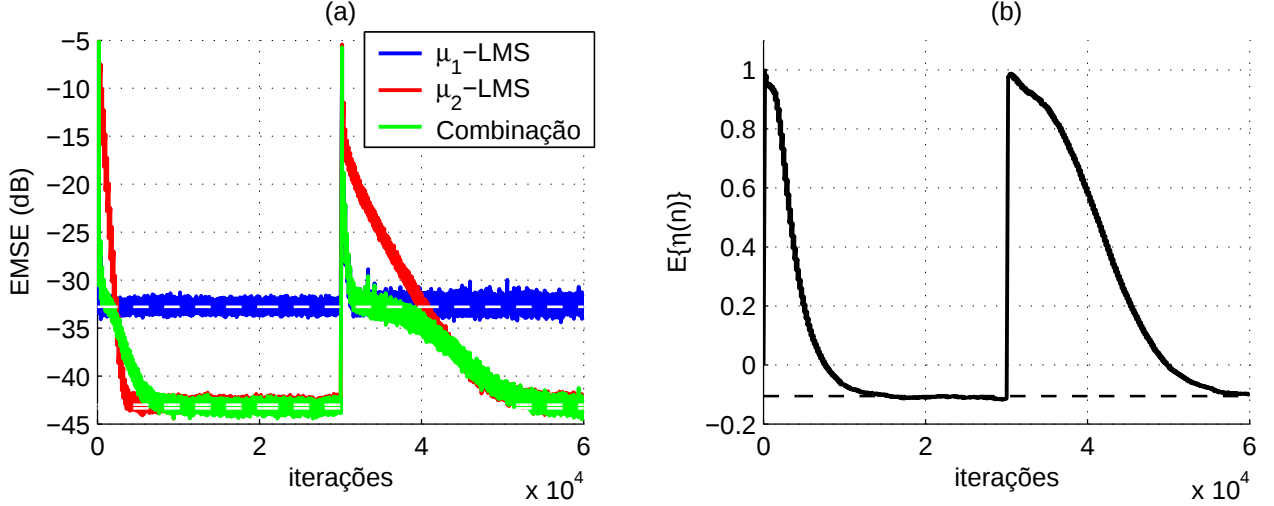


Figura 2.8: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 2,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 2,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados.

O comportamento das combinações de dois filtros LMS, dois filtros NLMS e dois filtros RLS quando $\delta_{XX} = 0,9$ no caso estacionário é mostrado respectivamente nas Figuras 2.11, 2.12 e 2.13. Nos três casos, foi feita uma média de 500 realizações com a solução ótima dada por $\mathbf{w}_{o,1}(0)$ com entrada colorida e utilizou-se um filtro de média móvel com 32 coeficientes. Pode-se notar que nos três casos ocorre um ganho de aproximadamente 3 dB e o valor de $\bar{\eta}_o(\infty)$ é negativo. Há uma boa concordância entre a análise e a simulação no caso das curvas de EMSE para a combinação com dois algoritmos LMS ou RLS. No caso do NLMS, a diferença entre análise e a simulação é um pouco maior devido à imprecisão que ocorre no modelo quando se utiliza um pequeno número de coeficientes nos filtros. O valor da análise para $\bar{\eta}_o(\infty)$ é menor do que o obtido na simulação nos três casos. Isso ocorre principalmente porque uma pequena variação no valor de δ_{XX} , quando $\delta_{XX} \approx 1$ nas expressões analíticas de $\bar{\eta}_o(\infty)$ causam uma grande variação no valor da estimativa de $\bar{\eta}_o(\infty)$. Pode-se notar ainda que apesar do ganho em relação aos filtros componentes, a velocidade de convergência da combinação é lenta já que a adaptação de $\eta(n)$ depende de $[y_1(n) - y_2(n)]$ (Equação (2.2)) que, nesse caso, é um valor muito pequeno já que os filtros têm passos (ou fatores de esquecimento) muito próximos.

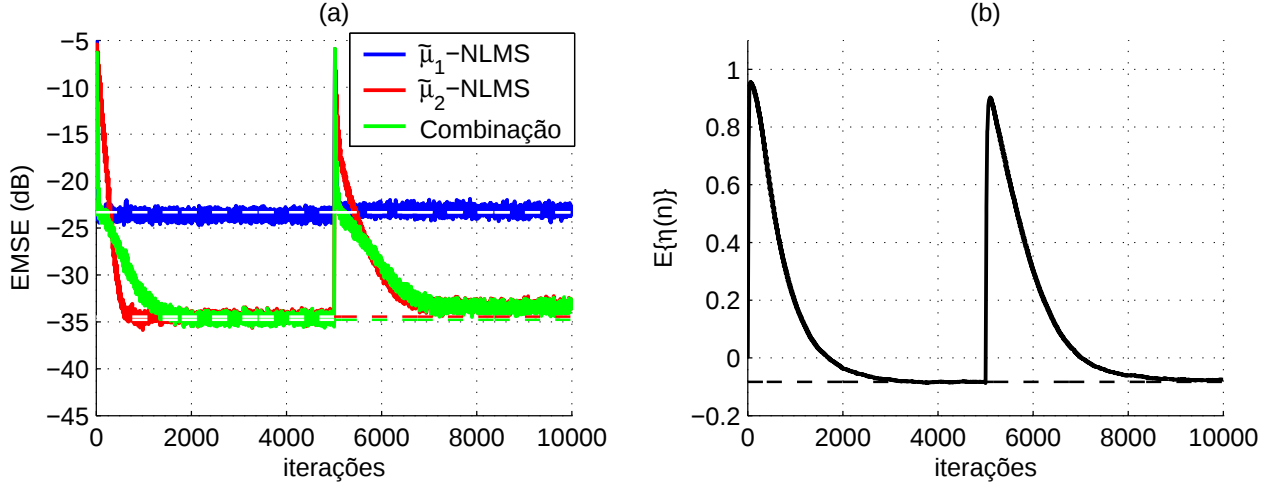


Figura 2.9: Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,5$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,1$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 5000$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 5000$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,5$. Valores teóricos tracejados.

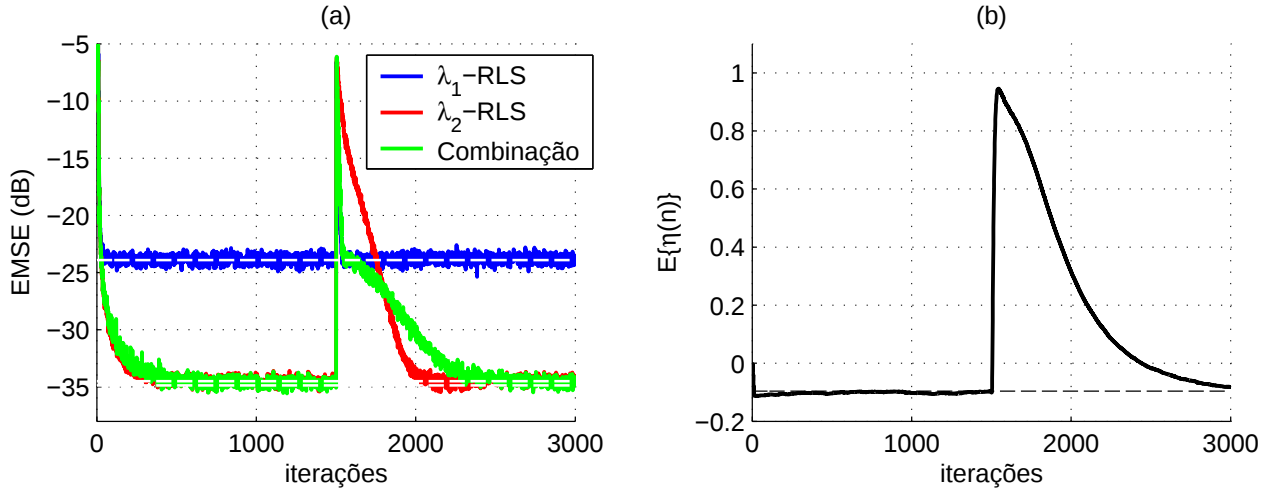


Figura 2.10: Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,1$ e $\epsilon_R = 1 \times 10^{-3}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}(0)$ até $n = 1500$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}(0)$ a partir de $n = 1500$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados.

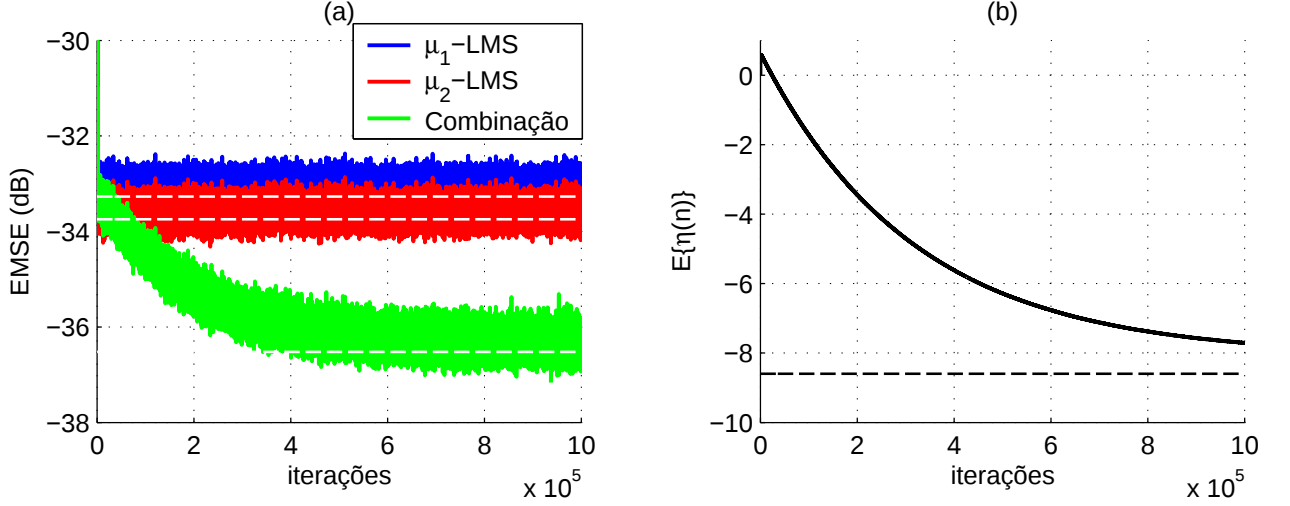


Figura 2.11: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,9$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados.

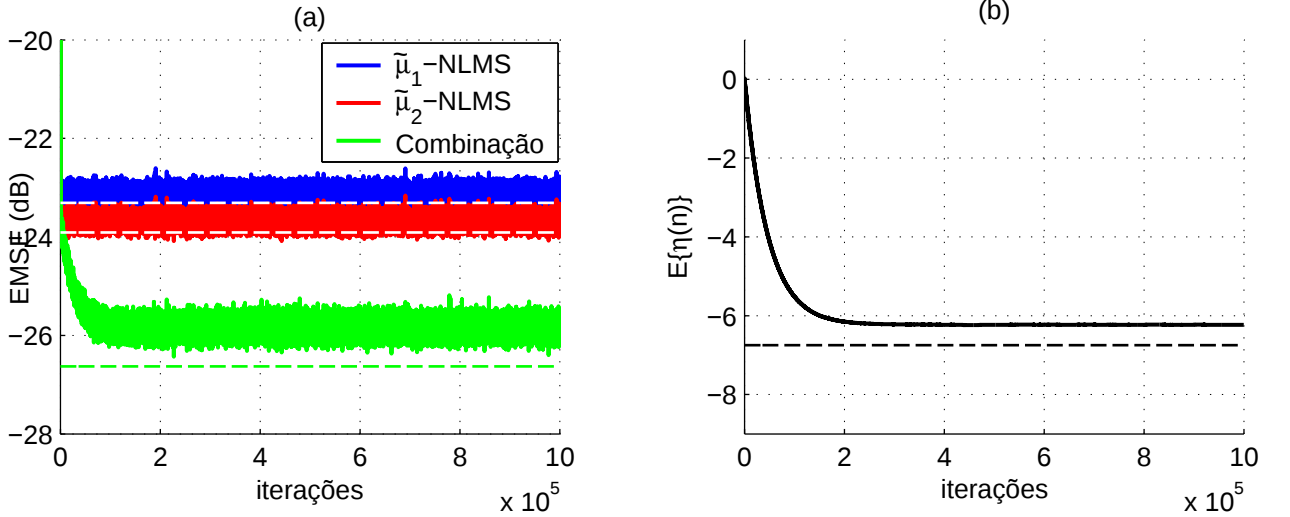


Figura 2.12: Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,5$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,9$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,5$. Valores teóricos tracejados.

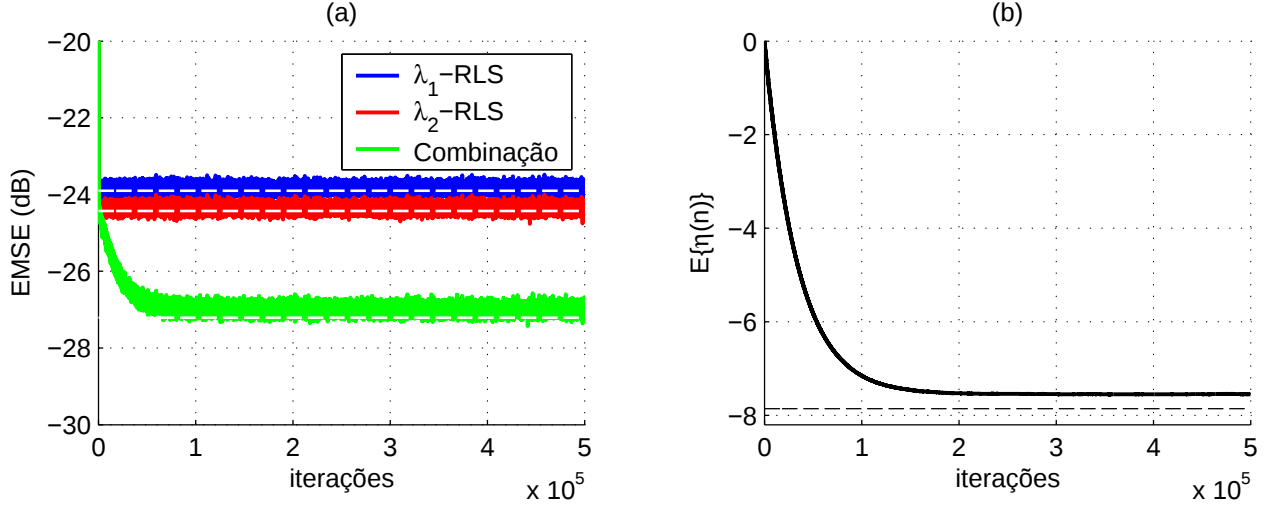


Figura 2.13: Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,9$ e $\epsilon_R = 1 \times 10^{-3}$. Entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,1}(0)$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 1$. Valores teóricos tracejados.

No caso autodidata, foram feitas simulações com a combinação afim de dois algoritmos CMA utilizando o algoritmo (2.6). Na combinação, foram usados dois equalizadores com $M = 4$ coeficientes cada, usando sobreamostragem de fator $S = 2$ [Treichler et al., 1996]. Os equalizadores foram inicializados com apenas um elemento não-nulo e unitário na segunda posição e o sinal de entrada considerado foi do tipo 4-PAM (*pulse amplitude modulation*). Nas simulações foram utilizados dois canais com coeficientes [Arenas-García e Figueiras-Vidal, 2006, Yousef e Sayed, 1999]

$$\mathbf{h}_1^T = [0,1 \quad 0,3 \quad 1,0 \quad -0,1 \quad 0,5 \quad 0,2]$$

e

$$\mathbf{h}_2^T = [0,25 \quad 0,64 \quad 0,80 \quad -0,55].$$

Na Figura 2.14, são mostrados o EMSE e o valor de $E\{\eta(n)\}$ ao longo das iterações no caso estacionário para a combinação de dois algoritmos CMA com $\mu_1 = 1 \times 10^3$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Foi feita uma média de 500 realizações e a curva de EMSE foi filtrada por um filtro de média móvel com 128 coeficientes para melhor visualização. Inicialmente foi usado o canal $\mathbf{h}_1$ e no meio da simulação, alterou-se para o canal $\mathbf{h}_2$. Pode-se observar que a análise faz uma boa estimativa do valor de $E\{\eta(n)\}$ em regime. No caso do EMSE, existe uma diferença entre os valores teóricos e experimentais. Entretanto a análise é capaz de prever o comportamento dos algoritmos razoavelmente bem, considerando que uma pequena diferença em dB é comum em modelos para algoritmos autodidatas devido às hipóteses que devem ser assumidas para possibilitar a análise.

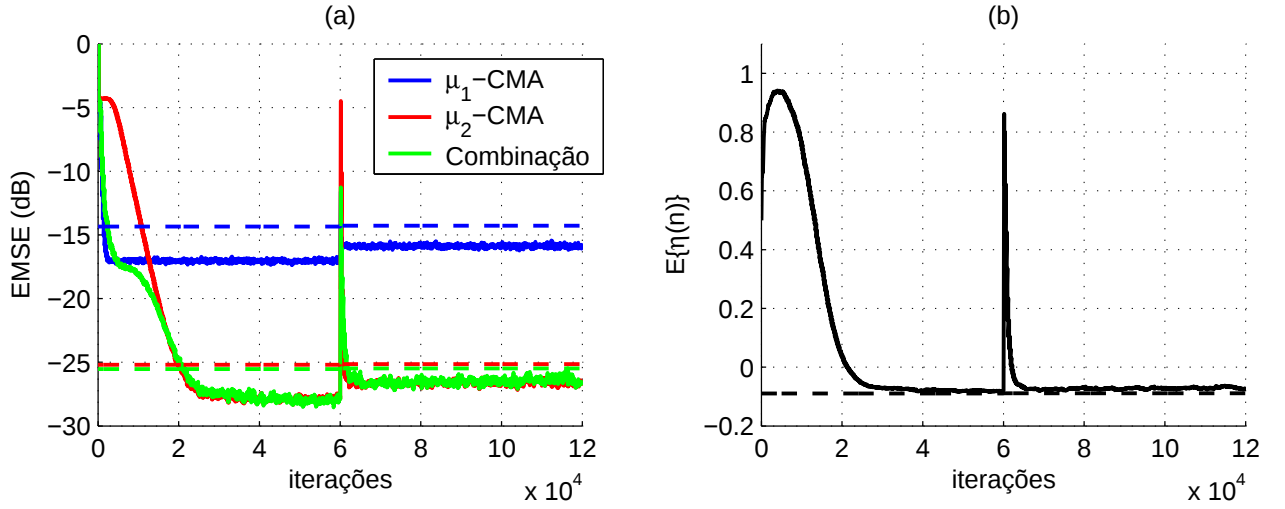


Figura 2.14: Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$ até $n = 6 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 6 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados.

Na Figura 2.15, são mostrados o EMSE e o valor de $E\{\eta(n)\}$ ao longo das iterações para o caso em que os passos de adaptação dos filtros componentes estão próximos ($\delta_{CC} = 0,9$). Nesse caso, utilizou-se o canal $\mathbf{h}_1$ e foram feitas 500 realizações. Além disso, as curvas de EMSE foram filtradas por um filtro de média-móvel com 128 coeficientes. Pode-se notar que o EMSE da combinação é cerca de 3 dB menor que o EMSE dos filtros componentes, conforme previsto pela análise. Como no caso supervisionado, a velocidade de convergência da combinação é muito lenta já que a adaptação de $\eta(n)$ depende de $[y_1(n) - y_2(n)]$ (Equação (2.6)).

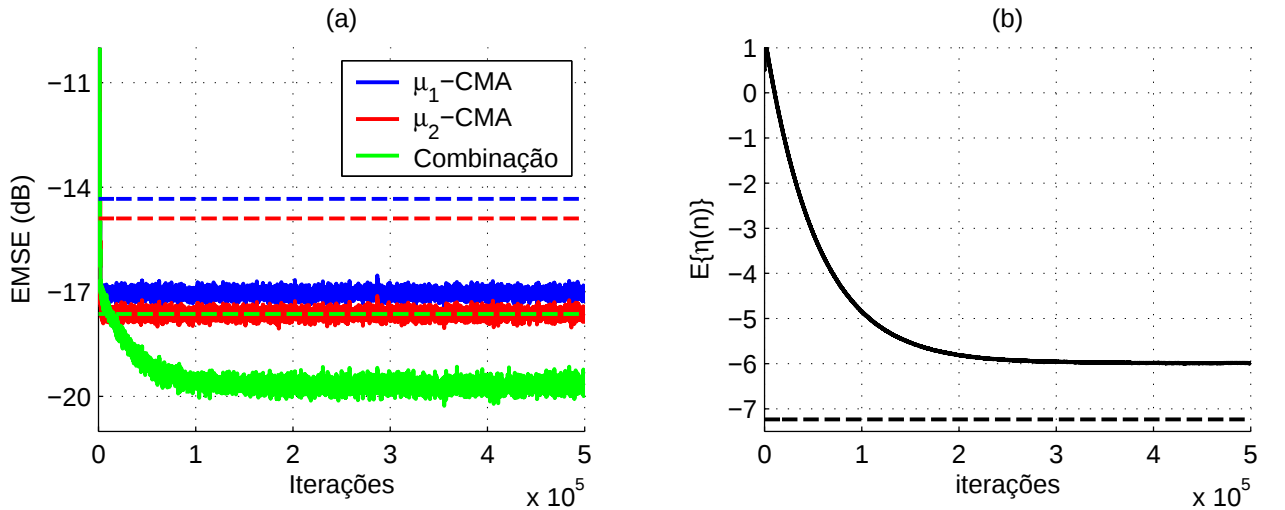


Figura 2.15: Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,9$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição e canal $\mathbf{h}_1$. Média de 500 realizações com $\mu_\eta = 0,1$. Valores teóricos tracejados.

Na Figura 2.16, são mostrados os resultados obtidos quando o parâmetro $\eta(n)$ é considerado constante e igual a $\bar{\eta}_o(\infty)$ mostrados na Tabela 2.5, com $\delta_{XX} = 0,999$. Para facilitar a visualização, utilizou-se um filtro de média móvel de 128 coeficientes. Pode-se notar que a combinação permite obter um ganho de aproximadamente 3 dB em relação aos filtros componentes.

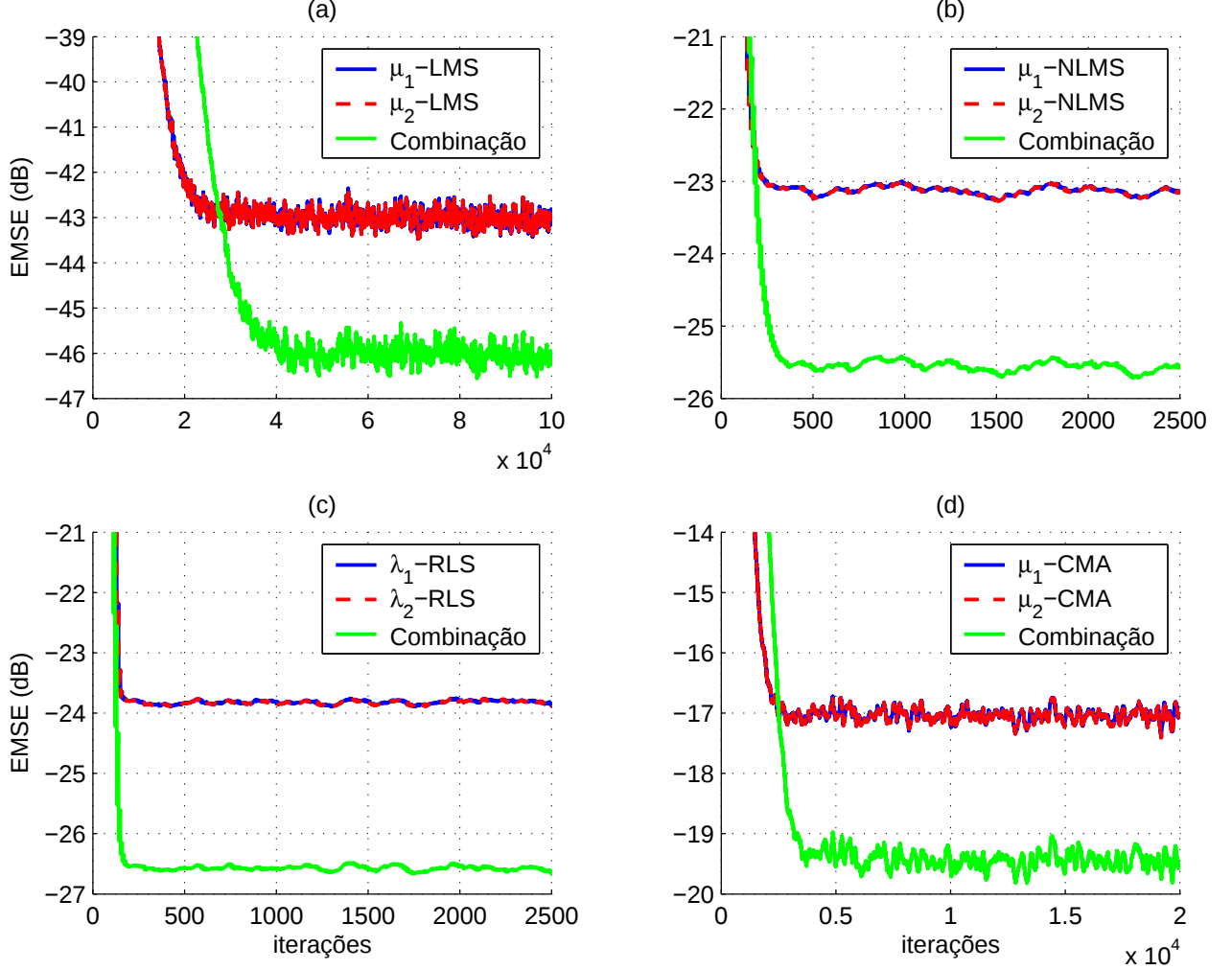


Figura 2.16: Resultados da combinação afim de dois algoritmos com $\delta_{XX} = 0,999$ e parâmetro $\eta(n)$ constante e igual a $\bar{\eta}_o(\infty)$. Média de 500 realizações. No caso supervisionado, solução ótima dada por $\mathbf{w}_{o,1}(0)$, entrada colorida (modelo AR com pólo em 0,8) com potência $1/M$ e $\sigma_v^2 = 0,01$; No caso autodidata, canal $\mathbf{h}_1$ e sinal de entrada 4-PAM. (a) μ_1 -LMS e μ_2 -LMS com $\mu_1 = 0,05$. (b) $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS com $\tilde{\mu}_1 = 0,5$ e $\epsilon_N = 1 \times 10^{-5}$. (c) λ_1 -RLS e λ_2 -RLS com $\lambda_1 = 0,9$ e $\epsilon_R = 1 \times 10^{-3}$. (d) μ_1 -CMA e μ_2 -CMA com $\mu_1 = 1 \times 10^{-3}$.

Com relação à velocidade de convergência, apesar do aparente atraso na convergência da combinação em relação à convergência dos filtros componentes, pode-se mostrar que tanto a combinação quanto os filtros componentes convergem com a mesma velocidade. Para o LMS, por exemplo, o valor esperado do vetor de erro dos coeficientes transformado de um filtro i

pode ser escrito na forma [Sayed, 2003]

$$\mathbf{E} \{ \mathbf{V}^T \tilde{\mathbf{w}}_i(n) \} \approx [\mathbf{I} - \mu_i \mathbf{\Lambda}]^n \mathbf{E} \{ \mathbf{V}^T \tilde{\mathbf{w}}_i(0) \}, \quad (2.58)$$

com as matrizes $\mathbf{\Lambda}$ e $\mathbf{V}$ obtidas a partir da decomposição da matriz de autocorrelação do sinal de entrada em autovalores e autovetores, ou seja, $\mathbf{R} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$, com $\mathbf{\Lambda}$ diagonal e $\mathbf{V}$ unitária. Assim, para cada coeficiente $c_{k,i}(n)$ de $\mathbf{E} \{ \mathbf{V}^T \tilde{\mathbf{w}}_i(n) \}$ pode-se escrever

$$c_{k,i}(n) \approx [1 - \mu_i \vartheta_k]^n c_{k,i}(0), \quad (2.59)$$

em que ϑ_k representa o autovalor de $\mathbf{R}$ referente a $c_{k,i}(n)$. A velocidade de convergência do filtro será determinada pelo coeficiente $c_{k,i}(n)$ com convergência mais lenta, ou seja, por $c_{k,i}(n)$ referente à ϑ_k mínimo. Na Figura 2.17, é mostrado o valor de $c_{k,1}(n)$ mais lento ao longo das iterações para a situação considerada na Figura 2.16-(a). Como pode ser observado, o coeficiente $c_{k,i}(n)$ com convergência mais lenta só converge totalmente por volta do instante $n = 4 \times 10^4$, o que justifica a convergência da combinação nesse mesmo instante, conforme pode ser visto na Figura 2.16-(a). Entretanto, vale notar que a influência do coeficiente $c_{k,i}(n)$ com convergência mais lenta não é significativa na convergência dos filtros componentes e passa a ser significativa na convergência da combinação.

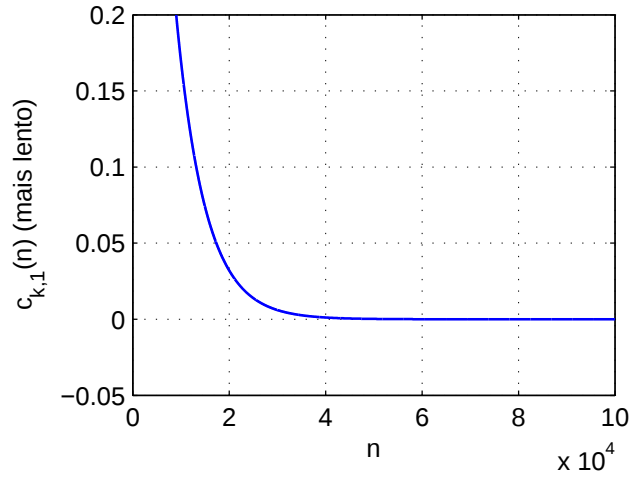


Figura 2.17: Coeficiente $c_{k,1}(n)$ do vetor $\mathbf{E} \{ \mathbf{V}^T \tilde{\mathbf{w}}_1(n) \}$ com menor velocidade de convergência na combinação de dois algoritmos LMS mostrada na Figura 2.16-(a).

Para verificar a possibilidade do ganho de EMSE maior que 3 dB no caso estacionário, foi considerada a combinação de duas combinações de equalizadores adaptados com o CMA. Para obter o ganho, os filtros componentes são adaptados com passos de adaptação próximos. Os resultados são mostrados na Figura 2.18, sendo os gráficos de EMSE dos filtros μ_2 -CMA e μ_3 -CMA omitidos para facilitar a visualização. É possível verificar que as combinações componentes obtêm um ganho de cerca de 3 dB em relação aos filtros componentes e a combinação global apresenta um ganho adicional, obtendo cerca de 4 dB em relação aos filtros compo-

nentes. Os resultados são coerentes com a análise, considerando a pequena diferença em dB, decorrente das hipóteses assumidas para a análise do CMA. A velocidade de convergência do parâmetro de mistura global é menor que a dos parâmetros de mistura das combinações componentes, o que motiva a utilização dos parâmetros de mistura fixos, com seus valores iguais aos valores analíticos ótimos em regime, como feito na Figura 2.16. Entretanto, no caso da combinação global, o valor analítico ótimo não é suficientemente preciso para que ocorra o ganho em relação ao EMSE das combinações componentes, sendo necessário fazer sua adaptação. Um comportamento interessante pode ser notado no início da convergência da combinação, onde esta é capaz de obter um nível de EMSE menor que o nível obtido em regime durante algumas iterações.

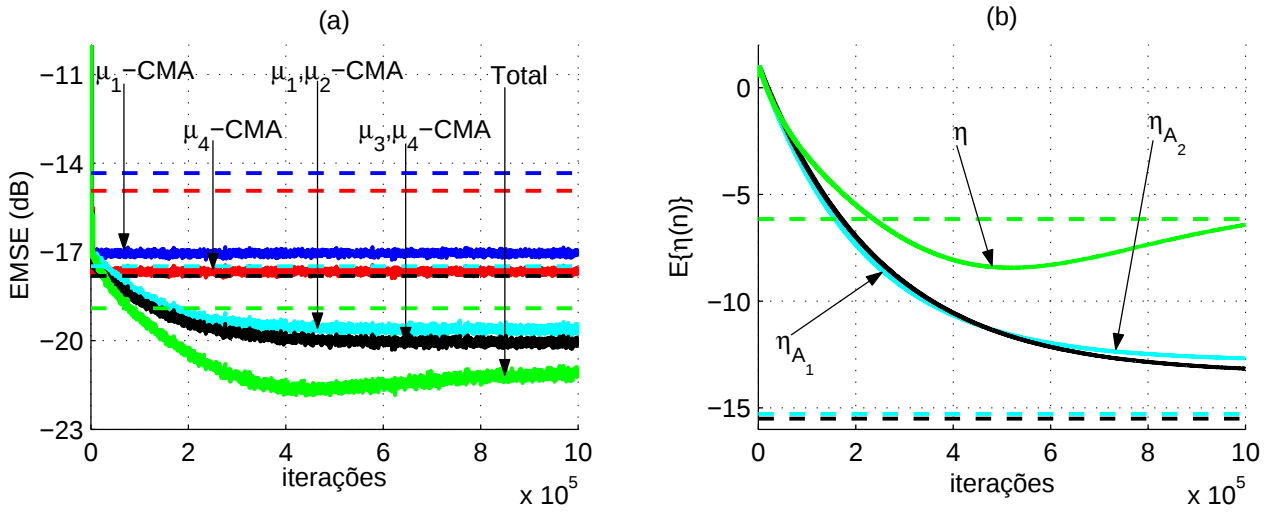


Figura 2.18: Combinação afim das combinações A_1 (μ_1 -CMA e μ_2 -CMA) e A_2 (μ_1 -CMA e μ_2 -CMA) no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC,1}\mu_1$ e $\delta_{CC,1} = 0,95$, $\mu_3 = 0,9405 \times 10^{-3}$, $\mu_4 = \delta_{CC,2}\mu_3$ e $\delta_{CC,2} = 0,95$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição e canal $\mathbf{h}_1$. Média de 500 realizações com $\mu_{\eta,A_1} = \mu_{\eta,A_2} = \mu_{\eta} = 0,1$. Valores teóricos tracejados.

Para verificar o comportamento das combinações e a validade da análise no caso não-estacionário, foram feitas simulações em diversos ambientes. Definindo $\delta_{RL} \triangleq \mu_2/(1 - \lambda_1)$ para a combinação λ_1 -RLS com μ_2 -LMS, na Figura 2.19, são mostrados os resultados de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1,2$ de combinações afim de dois filtros LMS, dois filtros NLMS, dois filtros RLS e um filtro RLS com um LMS para diversos valores de δ_{XY} e diversos valores de σ_q^2 . Os resultados experimentais para as combinações de dois filtros da mesma família coincidem com os resultados mostrados na Tabela 2.6, onde a maior redução de EMSE da combinação em relação ao EMSE dos filtros componentes ocorre quando $\text{Tr}(\mathbf{Q}) = q_{12}$ ou quando $\delta_{XX} \approx 1$. Também de acordo com a análise, quanto maior o valor de σ_q^2 e conseqüentemente $\text{Tr}(\mathbf{Q})$, menor o ganho da combinação em relação aos filtros componentes. Pode-se notar que existe um

ponto em que $\zeta(\infty) = \min\{\zeta_i(\infty)\}$, $i = 1, 2$. Nesse ponto, a combinação tem um desempenho igual ao desempenho do filtro componente com menor EMSE, que é adaptado com passo (ou fator de esquecimento) ótimo para o determinado valor de σ_q^2 , $\mathbf{R}$ e σ_v^2 . Apesar da combinação afim conseguir obter um ganho de EMSE em relação aos filtros componentes, o EMSE mínimo que pode ser obtido é limitado pelo EMSE obtido com um filtro adaptado com passo (ou fator de esquecimento) ótimo.

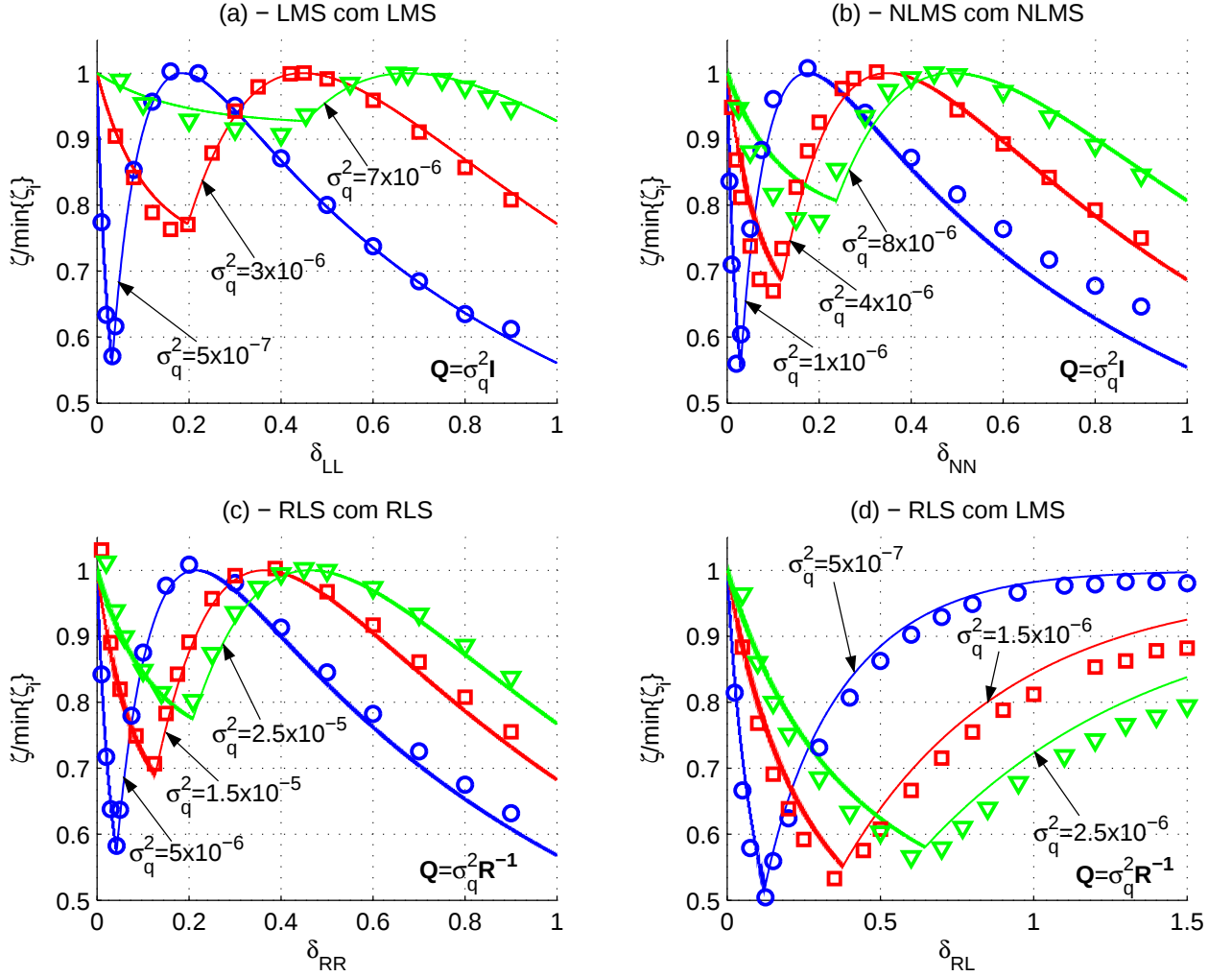


Figura 2.19: Combinação afim no caso supervisionado não-estacionário. Valores teóricos (linha cheia) e experimentais ($\circ$, $\square$, ∇) de $\zeta(\infty)/\min\{\zeta_i(\infty)\}$, $i = 1, 2$ para 50 realizações com $M = 7$, $\sigma_v^2 = 0,01$ e diversos valores de δ_{XY} e diversos valores de σ_q^2 . Entrada colorida (modelo AR com pólo em 0,8) de potência $1/M$ e solução ótima $\mathbf{w}_1(0)$. (a) μ_1 -LMS com μ_2 -LMS, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$, $\mu_\eta = 1$. (b) $\tilde{\mu}_1$ -NLMS com $\tilde{\mu}_2$ -NLMS, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\epsilon_N = 1 \times 10^{-5}$ e $\mu_\eta = 1$. (c) λ_1 -RLS com λ_2 -RLS, $\lambda_1 = 0,9$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 1$. (d) λ_1 -RLS com μ_2 -LMS, $\lambda_1 = 0,9$, $\mu_2 = \delta_{RL}(1 - \lambda_1)$, $\epsilon_R = 1 \times 10^{-3}$ e $\mu_\eta = 0,5$.

No caso da combinação de um filtro RLS com um LMS, fica difícil simplificar a expressão de EMSE para caracterizar o comportamento da combinação. Nesse contexto, as curvas teóricas do NSD mostradas na Figura 2.20 possibilitam uma outra forma de interpretação da

combinação. Pode-se notar que não existe uma simetria entre as curvas NSD_1 e NSD_2 como ocorre com a combinação de dois filtros de mesma família. Isso ocorre pois não existe um valor único de EMSE ótimo e o NSD deve ser calculado como

$$\text{NSD}_i(\infty) = \frac{\zeta_i(\infty)}{\min\{\zeta_{o,\text{LMS}}(\infty), \zeta_{o,\text{RLS}}(\infty)\}}, \quad (2.60)$$

sendo $\zeta_{o,\text{LMS}}$ e $\zeta_{o,\text{RLS}}$ os valores ótimos em regime do EMSE obtidos com os algoritmos LMS e RLS, respectivamente. Com isso, a redução de EMSE obtida quando o EMSE do filtro 1 é igual ao EMSE do filtro 2 depende também da diferença entre $\zeta_{o,\text{RLS}}(\infty)$ e $\zeta_{o,\text{LMS}}(\infty)$. Nos resultados mostrados na Figura 2.19-(d), a maior redução de EMSE obtida foi de aproximadamente 3 dB e para as três situações consideradas, ocorreu quando $\zeta_1(\infty) \approx \zeta_2(\infty)$.

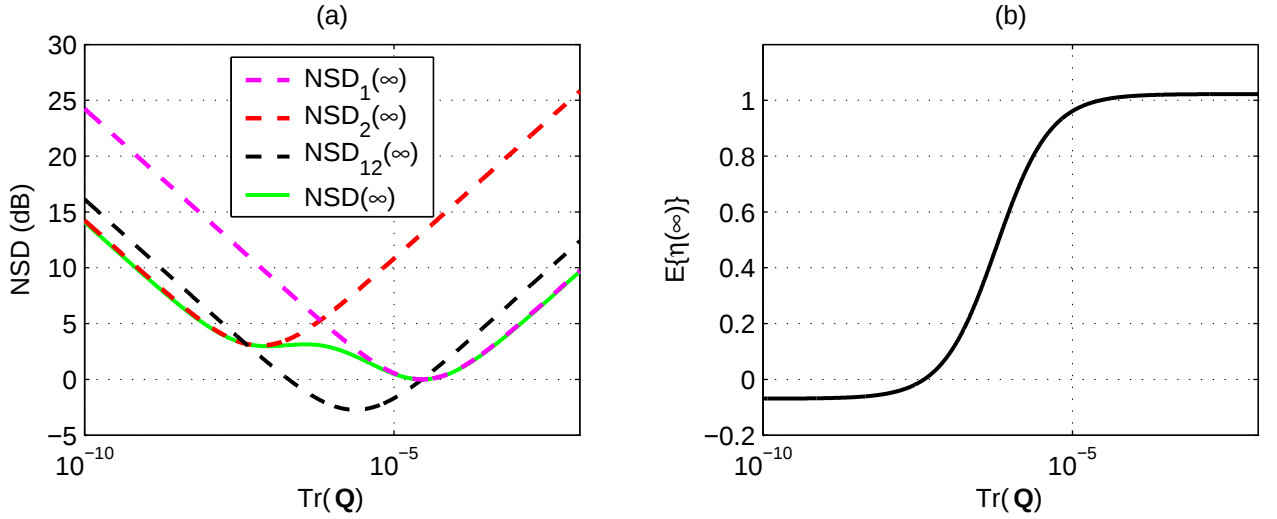


Figura 2.20: Curvas teóricas de NSD da combinação afim de λ_1 -RLS e μ_2 -LMS. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,99$, $\mu_2 = 1 \times 10^{-3}$ e $\mathbf{Q} \propto \mathbf{R}^{-1}$. Entrada colorida (modelo AR com pólo em 0,8).

Para o caso não-estacionário e autodidata, foram feitas simulações da combinação de dois algoritmos CMA para diversos valores de $\text{Tr}(\mathbf{Q})$ usando o canal $\mathbf{h}_1$. Na Figura 2.21, são mostrados os resultados do $\text{NSD}(\infty)$ para os filtros componentes ($\text{NSD}_1(\infty)$ e $\text{NSD}_2(\infty)$), o cruzado ($\text{NSD}_{12}(\infty)$) e para a combinação afim ($\text{NSD}_a(\infty)$). Além disso, para efeito de comparação, são também mostrados os resultados obtidos com a combinação convexa ($\text{NSD}_c(\infty)$) de [Arenas-García e Figueiras-Vidal, 2006]. Nesse caso, devido às diferenças entre os valores de EMSE teóricos e experimentais que ocorrem para o caso autodidata, os valores de ζ_o usados para o cálculo do NSD foram obtidos experimentalmente. Nessa figura, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Como $\delta_{CC} = 0,1$, a redução de EMSE que ocorre na combinação afim no intervalo $[q_2, q_1]$ é quase imperceptível para o nível de detalhe da figura. O desempenho das duas combinações é muito semelhante, apesar da restrição imposta a $\eta(n)$ na combinação convexa. Pode-se notar que o máximo ganho de EMSE da combinação em relação aos fil-

tros componentes é de aproximadamente 1,8 dB e ocorre para ambas as combinações quando $\text{Tr}(\mathbf{Q}) \approx \mu_1 \mu_2 \sigma_\beta^2 \text{Tr}(\mathbf{R}) \approx 4 \times 10^{-7}$, que faz com que $\zeta_1(\infty) \approx \zeta_2(\infty)$.

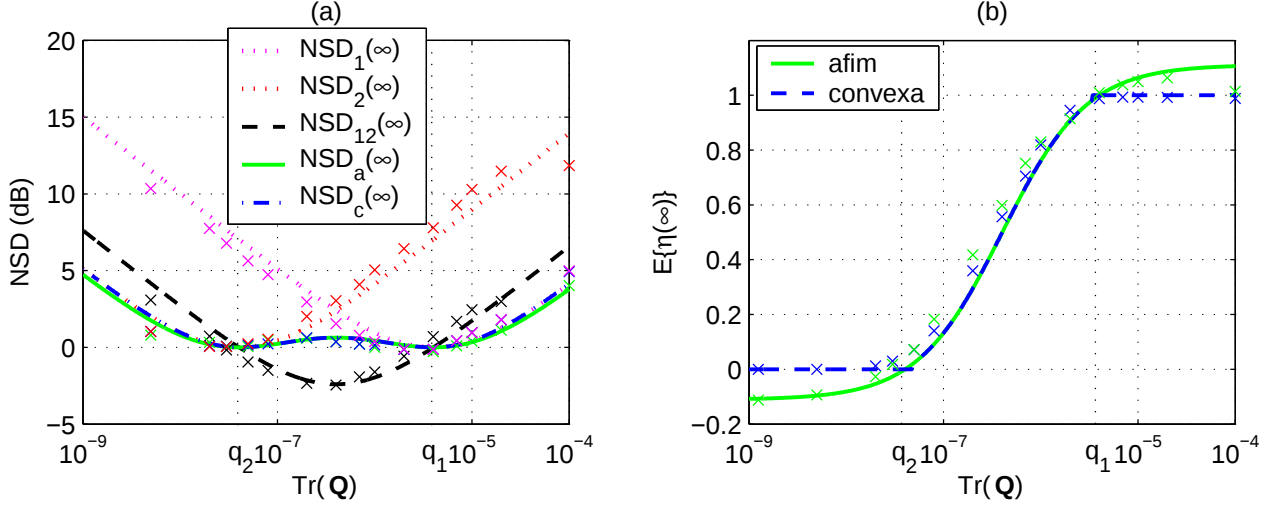


Figura 2.21: Combinações afim e convexa no caso autodidata não-estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-4}$, $\mu_2 = \delta_{CC}\mu_1$, $\delta_{CC} = 0,1$. Vetores de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$. Valores teóricos (linhas) e experimentais ($\times$) do NSD dos filtros componentes, da combinação afim (NSD_a , $\mu_\eta = 0,0075$) e da combinação convexa (NSD_c , $\mu_\alpha = 0,075$, $\alpha^+ = 4$). Média de 50 realizações.

2.9 Conclusões

Neste capítulo, a combinação afim de dois algoritmos LMS de Bershad et al. [2008] foi estendida para os algoritmos NLMS, RLS e CMA. Foi feita uma análise do parâmetro de mistura ótimo e do EMSE da combinação em regime obtendo-se expressões que dependem do EMSE dos filtros componentes em regime e do EMSE cruzado. Os resultados obtidos são semelhantes aos resultados existentes para a combinação convexa em Arenas-García et al. [2006a], exceto pela limitação do parâmetro de mistura ótimo ao intervalo $[0, 1]$ da combinação convexa. Usando resultados existentes na literatura para as expressões de EMSE dos filtros componentes e do EMSE cruzado, verificaram-se algumas propriedades da combinação afim.

No caso estacionário, o valor do parâmetro de mistura ótimo em regime é sempre negativo e a combinação ótima tem EMSE em regime sempre inferior ao menor EMSE em regime dos filtros componentes. Na prática, é possível verificar que a combinação afim tem comportamento próximo ao universal. Um resultado interessante é obtido quando são usados filtros componentes com passos de adaptação ou fatores de esquecimento muito próximos. Nesse caso, o valor do parâmetro de mistura ótimo em regime é um número negativo com módulo muito grande. Isso faz com que a combinação ótima seja capaz de obter um EMSE em regime igual

à metade do menor EMSE dentre os filtros componentes, ou seja, há um ganho de EMSE. A implementação prática dessa situação não é muito eficiente já que a adaptação do parâmetro de mistura envolve o termo $[y_1(n) - y_2(n)]$, o que faz com que a velocidade de convergência da combinação seja extremamente lenta já que as saídas dos filtros componentes são muito próximas. Uma alternativa para contornar esse problema é fixar o valor do parâmetro de mistura como sendo o valor do parâmetro de mistura ótimo em regime obtido analiticamente. Assim, a combinação consegue prover o ganho estimado pela análise com a mesma velocidade de convergência dos filtros componentes.

O ganho de EMSE pode ser aumentado, combinando as saídas de combinações de filtros componentes adaptados com passos (ou fatores de esquecimento) próximos. De acordo com a análise feita para o caso da combinação global de duas combinações componentes, é possível obter um ganho de EMSE de 4,26 dB em relação ao EMSE dos filtros componentes. Nesse caso, a velocidade de convergência do parâmetro de mistura global torna-se ainda mais lenta e a utilização do valor teórico como parâmetro de mistura fixo faz com que o ganho não ocorra, já que o valor teórico não é suficientemente preciso. Esse resultado indica que se considerado um esquema de combinação de combinações em formato de árvore, o ganho de EMSE obtido será aumentado a cada nível mas terá um limite. O cálculo desse limite e do número de filtros necessários para atingi-lo necessita do estudo de uma forma mais simples de análise, já que a análise apresentada aumenta muito em complexidade nesse caso.

No caso não-estacionário, verificou-se que quando os filtros componentes têm passos de adaptação (ou fatores de esquecimento) distintos, a combinação é capaz de prover um ganho de EMSE de até 3 dB em relação ao menor EMSE dentre os filtros componentes dependendo dos parâmetros do ambiente considerado. O ganho máximo sempre ocorre quando os níveis de EMSE dos filtros componentes são próximos. No caso não-estacionário, essa situação pode ocorrer também com passos de adaptação (ou fatores de esquecimento) distintos para um determinado conjunto de parâmetros do cenário de filtragem. Quanto maior a diferença entre os níveis de EMSE dos filtros componentes, menor o ganho de EMSE da combinação, sendo que o EMSE da combinação é sempre, na pior das hipóteses, igual ao menor EMSE dentre os filtros componentes. No caso de passos de adaptação (ou fatores de esquecimento) próximos, o resultado obtido é semelhante. Porém, nesse caso, o número de conjuntos de parâmetros do cenário de filtragem para os quais a combinação apresenta o ganho de EMSE é maior, já que os filtros componentes possuem desempenhos muito próximos.

Em comparação com a combinação convexa, a diferença da combinação afim é que a última é capaz de fazer com que o parâmetro de mistura seja negativo. Em ambiente estacionário, isso faz com que seja possível obter um ganho de EMSE de até 3 dB em relação ao menor EMSE dentre os filtros componentes, sendo que o ganho máximo ocorre quando os níveis de EMSE dos filtros componentes são muito próximos. Num ambiente não-estacionário, com os filtros componentes com passos de adaptação (ou fatores de esquecimento) distintos, o comportamento da combinação afim e da convexa é muito semelhante. Nesse caso, as duas

combinações são capazes de prover até 3 dB de ganho de EMSE. Isso ocorre porque no caso não-estacionário, o ganho ocorre numa situação em que o parâmetro de mistura ótimo não é negativo, fazendo com que seja possível obtê-lo também com a combinação convexa. Quando os passos de adaptação (ou fatores de esquecimento) são próximos, o comportamento é semelhante ao do caso estacionário. A combinação afim é capaz de proporcionar um ganho de EMSE de até 3 dB em relação ao menor EMSE dentre os filtros componentes enquanto a combinação convexa tem desempenho limitado pelo menor EMSE dos filtros componentes.

Capítulo 3

Análise do transitório das combinações e propostas para adaptação do parâmetro de mistura

O comportamento universal da combinação afim depende da adaptação do parâmetro de mistura. A adaptação desse parâmetro, por sua vez, depende da escolha apropriada do passo de adaptação μ_η . Essa escolha deve levar em conta algumas características do cenário de filtragem como a potência dos sinais de saída e os passos de adaptação dos filtros componentes, além da relação sinal-ruído. Com o objetivo de incluir esses fatores na escolha do passo, são propostas neste capítulo três alternativas para a adaptação do parâmetro de mistura.

Inicialmente é obtida uma expressão para o EMSE da combinação sem levar em conta o algoritmo de adaptação de $\eta(n)$. Em seguida, essa expressão é usada para analisar o algoritmo η -LMS. Para validar a análise, são mostrados alguns resultados de simulação e algumas condições em que o algoritmo η -LMS não apresenta um bom desempenho na adaptação do parâmetro de mistura. Para resolver o problema, é proposta uma normalização instantânea parcial na adaptação de $\eta(n)$, originando o algoritmo *signed regressor* (SR), que é analisado em seguida. Com base na análise do η -LMS e do SR, são propostos outros dois algoritmos normalizados a fim de facilitar a escolha do passo de adaptação da combinação. As análises desses algoritmos normalizados são obtidas da extensão das análises do η -LMS e SR. Por fim, são apresentados alguns resultados de simulação dos algoritmos propostos e as conclusões.

3.1 O EMSE da combinação

Como foi feito na Seção 2.2, o erro *a priori* da combinação afim pode ser escrito em função dos erros *a priori* dos filtros componentes, ou seja,

$$e_a(n) \approx \eta(n)e_{a,1}(n) + [1 - \eta(n)]e_{a,2}(n) = e_{a,2}(n) + \eta(n)[e_{a,1}(n) - e_{a,2}(n)]. \quad (3.1)$$

No caso supervisionado, essa expressão é válida em todo instante de tempo mas no caso auto-didata, é uma aproximação cuja validade em todo instante depende dos filtros componentes terem sido inicializados próximos da solução ótima de *zero-forcing*.

Elevando (3.1) ao quadrado e tomando a esperança de ambos os lados, obtém-se

$$\begin{aligned} E \{e_a^2(n)\} &\approx E \{e_{a,2}^2(n)\} + E \{ \eta^2(n) [e_{a,1}^2(n) + e_{a,2}^2(n) - 2e_{a,1}(n)e_{a,2}(n)] \} \\ &\quad + 2E \{ \eta(n) [e_{a,2}(n)e_{a,1}(n) - e_{a,2}^2(n)] \}. \end{aligned} \quad (3.2)$$

Para continuar, é necessário assumir as seguintes hipóteses:

T1. Na média, o parâmetro de mistura ótimo da combinação tende para seu valor ótimo em todo instante de tempo, ou seja, $E \{ \eta(n) \} \approx \bar{\eta}_o(n)$;

T2. $\eta(n)$ é independente de $e_{a,i}(n)$, $i = 1, 2$;

T3. $\eta_o(n)$ é independente de $e_{a,i}(n)$, $i = 1, 2$.

A hipótese T1 vale em regime, como será mostrado na Seção 3.2. No transitório, é uma hipótese razoável como pode ser observado através de simulações. As hipóteses T2 e T3 em geral não são verdadeiras uma vez que a adaptação do parâmetro $\eta(n)$ e seu valor ótimo $\bar{\eta}_o(n)$ dependem dos erros *a priori* $e_{a,i}(n)$, $i = 1, 2$. No entanto, elas fazem parte de um conjunto de hipóteses de independência amplamente usadas em filtragem adaptativa [Haykin, 2002]. Em [Mazo, 1979], por exemplo, é mostrado que as hipóteses baseadas na teoria da independência são satisfeitas pelo algoritmo LMS com passo de adaptação suficientemente pequeno. Para o caso da combinação afim, é possível verificar através de simulações que T2 e T3 são hipóteses razoáveis.

Usando T1 e T2, (3.2) pode ser reescrita como

$$\zeta(n) \approx \zeta_2(n) + E \{ \eta^2(n) \} [\Delta\zeta_1(n) + \Delta\zeta_2(n)] - 2\bar{\eta}_o(n)\Delta\zeta_2(n). \quad (3.3)$$

Usando a Equação (2.32) (pág. 21) e assumindo T3, obtém-se uma expressão para o parâmetro de mistura ótimo instantâneo dada por

$$\bar{\eta}_o(n) \approx \frac{\Delta\zeta_2(n)}{\Delta\zeta_1(n) + \Delta\zeta_2(n)}. \quad (3.4)$$

Assumindo que $\eta(n)$ seja igual ao valor do parâmetro de mistura ótimo pode-se obter uma expressão para o EMSE ótimo da combinação a partir de (3.3), ou seja,

$$\zeta_o(n) \approx \zeta_2(n) - \bar{\eta}_o(n)\Delta\zeta_2(n). \quad (3.5)$$

Na prática, no entanto, o valor do parâmetro de mistura em todo instante de tempo não

é igual ao seu valor ótimo. Usando T1, a variância de $\eta(n)$ vale

$$\sigma_\eta^2(n) = E\{\eta^2(n)\} - \bar{\eta}_o^2(n). \quad (3.6)$$

Assim, usando (3.6) e (3.4), (3.3) pode ser reescrita como

$$\zeta(n) \approx \zeta_2(n) + [\sigma_\eta^2(n) - \bar{\eta}_o^2(n)] [\Delta\zeta_1(n) + \Delta\zeta_2(n)]. \quad (3.7)$$

Vale notar que a expressão (3.7) fornece um resultado teórico para o EMSE da combinação em cada instante de tempo considerando a adaptação do parâmetro de mistura. A influência da adaptação de $\eta(n)$ aparece no termo $\sigma_\eta^2(n)$ que deve ser estimado para cada tipo de adaptação considerado. É necessário também usar modelos de $\zeta_i(n)$, $i = 1, 2$ e $\zeta_{12}(n)$. Para os algoritmos considerados, existem várias análises de transitório na literatura, em que são obtidas expressões teóricas para $\zeta_i(n)$, $i = 1, 2$ (por exemplo, [Farhang-Boroujeny, 1998], [Costa e Bermudez, 2002] e [Nascimento e Silva, 2008]). Expressões para o EMSE cruzado podem ser obtidas a partir da extensão desses resultados.

Uma expressão para a variância $\sigma_\eta^2(n)$ pode ser obtida através do erro do parâmetro de mistura em relação ao seu valor ótimo médio. Definindo

$$\tilde{\eta}(n) \triangleq \bar{\eta}_o(n) - \eta(n), \quad (3.8)$$

elevando ao quadrado e tomando a esperança de ambos os lados, obtém-se

$$E\{\tilde{\eta}^2(n)\} = \bar{\eta}_o^2(n) + E\{\eta^2(n)\} - 2\bar{\eta}_o(n)E\{\eta(n)\}. \quad (3.9)$$

Utilizando a hipótese T1, (3.9) pode ser reescrita como

$$E\{\tilde{\eta}^2(n)\} \approx E\{\eta^2(n)\} - \bar{\eta}_o^2(n) = \sigma_\eta^2(n). \quad (3.10)$$

A seguir, para validar a hipótese T1 em regime e obter uma expressão para $\sigma_\eta^2(n)$, serão feitas análises dos momentos de primeira e segunda ordens do parâmetro de mistura obtido usando o algoritmo η -LMS.

3.2 Análise da adaptação do algoritmo η -LMS

No Capítulo 2 a adaptação do parâmetro de mistura foi feita utilizando algoritmos do gradiente estocástico. Por conveniência, a forma de adaptação do parâmetro de mistura da combinação é repetida aqui. De maneira geral, a adaptação de $\eta(n)$ através do algoritmo η -LMS é dada por

$$\eta(n+1) = \eta(n) + \mu_\eta e_g(n)[y_1(n) - y_2(n)], \quad (3.11)$$

sendo

$$e_g(n) = e_a(n) + r(n), \quad (3.12)$$

com $r(n) = v(n)$ para o caso supervisionado e $r(n) = 0$ para o caso autodidata.

A seguir são feitas análises dos momentos de primeira e segunda ordens de (3.11).

3.2.1 Momento de primeira ordem

A expressão (3.7) permite obter uma estimativa para o EMSE da combinação desde que a adaptação de $\eta(n)$ considerada assegure a hipótese T1. A validade dessa hipótese em regime pode ser verificada através de uma análise de primeira ordem do erro do parâmetro de mistura em relação ao ótimo dado por (3.8). Assim, para obter uma expressão de $E\{\tilde{\eta}(n)\}$ para a adaptação feita com (3.11), subtrai-se ambos os lados de (3.11) de $\bar{\eta}_o(n)$, obtendo-se

$$\bar{\eta}_o(n) - \eta(n+1) = \bar{\eta}_o(n) - \eta(n) - \mu_\eta e_g(n)[y_1(n) - y_2(n)]. \quad (3.13)$$

Substituindo (2.22) (pág. 19) em (3.12), obtém-se

$$e_g(n) \approx \eta(n)e_{a,1}(n) + [1 - \eta(n)]e_{a,2}(n) + r(n). \quad (3.14)$$

Considerando que no caso supervisionado $y_i(n) = d(n) - e_i(n) = d(n) - e_{a,i}(n) - v(n)$ e no caso autodidata, $y_i(n) \approx a(n - \tau_d) - e_{a,i}(n)$, o termo $[y_1(n) - y_2(n)]$ pode ser aproximado em ambos os casos por

$$y_1(n) - y_2(n) \approx e_{a,2}(n) - e_{a,1}(n). \quad (3.15)$$

Note que no caso supervisionado, (3.15) não é uma aproximação e sim uma igualdade. Substituindo (3.14) e (3.15) em (3.13) chega-se a

$$\begin{aligned} \bar{\eta}_o(n) - \eta(n+1) &= \bar{\eta}_o(n) - \eta(n) + \mu_\eta \eta(n)[e_{a,2}(n) - e_{a,1}(n)]^2 \\ &\quad - \mu_\eta e_{a,2}(n)[e_{a,2}(n) - e_{a,1}(n)] - \mu_\eta r(n)[e_{a,2}(n) - e_{a,1}(n)]. \end{aligned} \quad (3.16)$$

Para continuar, é necessário assumir que

- T4. a variação de $\bar{\eta}_o(n)$ para $\bar{\eta}_o(n+1)$ é suficientemente pequena diante das outras parcelas de (3.16) para que seja possível usar a aproximação $\bar{\eta}_o(n) \approx \bar{\eta}_o(n+1)$. Em regime, essa hipótese é bastante realista já que $\eta(n)$ tem uma variação muito maior que $\bar{\eta}_o(n)$ mas durante o transitório, nem sempre é válida. Entretanto, para a adaptação considerada, pode-se observar através de simulações que essa aproximação permite a obtenção de um resultado teórico coerente com o resultado experimental.

Somando e subtraindo $\mu_\eta \bar{\eta}_o(n)[e_{a,2} - e_{a,1}]^2$ do lado direito de (3.16) e usando T4, obtém-se

$$\begin{aligned} \tilde{\eta}(n+1) = & \overbrace{[1 - \mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2]}^{\mathcal{A}} \tilde{\eta}(n) + \overbrace{\bar{\eta}_o(n)\mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2}^{\mathcal{B}} \\ & - \underbrace{\mu_\eta e_{a,2}(n)[e_{a,2}(n) - e_{a,1}(n)]}_c - \underbrace{\mu_\eta r(n)[e_{a,2}(n) - e_{a,1}(n)]}_d. \end{aligned} \quad (3.17)$$

Aplicando a esperança em ambos os lados de (3.17), usando T3 e considerando que $r(n)$ é independente de $e_{a,i}(n)$ com $E\{r(n)\} = 0$ para os casos supervisionado e autodidata, obtém-se

$$\begin{aligned} E\{\tilde{\eta}(n+1)\} = & E\{1 - \mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2\} E\{\tilde{\eta}(n)\} \\ & + \mu_\eta \bar{\eta}_o(n) E\{[e_{a,2}(n) - e_{a,1}(n)]^2\} - \mu_\eta E\{e_{a,2}(n)[e_{a,2}(n) - e_{a,1}(n)]\}. \end{aligned} \quad (3.18)$$

A validade de T1 em regime será garantida se $\lim_{n \rightarrow \infty} E\{\tilde{\eta}(n)\} = 0$. Essa condição ocorre se os dois últimos termos de (3.18) se anularem, ou seja,

$$\mu_\eta \bar{\eta}_o(n) E\{[e_{a,2}(n) - e_{a,1}(n)]^2\} - \mu_\eta E\{e_{a,2}(n)[e_{a,2}(n) - e_{a,1}(n)]\} = 0 \quad (3.19)$$

e se o fator que multiplica $\tilde{\eta}(n)$ for, em módulo, menor que um, ou seja,

$$|E\{1 - \mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2\}| < 1. \quad (3.20)$$

Usando (3.4) e lembrando que $\Delta\zeta_i(n) = \zeta_i(n) - \zeta_{12}(n)$, a primeira parcela do lado esquerdo de (3.19) pode ser reescrita como

$$\mu_\eta \bar{\eta}_o(n) E\{[e_{a,2}(n) - e_{a,1}(n)]^2\} = \mu_\eta \bar{\eta}_o(n) [\Delta\zeta_1(n) + \Delta\zeta_2(n)] = \mu_\eta \Delta\zeta_2(n) \quad (3.21)$$

e a segunda parcela como

$$\mu_\eta E\{e_{a,2}(n)[e_{a,2}(n) - e_{a,1}(n)]\} = \mu_\eta \Delta\zeta_2(n). \quad (3.22)$$

Como essas duas parcelas são iguais, verifica-se a igualdade de (3.19).

Para garantir (3.20), o passo de adaptação μ_η deve ser escolhido tal que

$$|E\{1 - \mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2\}| = |1 - \mu_\eta[\Delta\zeta_1(n) + \Delta\zeta_2(n)]| < 1, \quad (3.23)$$

ou seja,

$$0 < \mu_\eta < \frac{2}{\Delta\zeta_1(n) + \Delta\zeta_2(n)}, \quad (3.24)$$

e com isso a validade da hipótese T1 em regime é garantida.

3.2.2 Momento de segunda ordem

O modelo de EMSE da combinação dado pela expressão (3.7) requer ainda uma expressão para $\sigma_\eta^2(n)$, que pode ser obtida através de (3.10). Para tanto, a esperança do quadrado de (3.17) deve ser calculada, ou seja,

$$\begin{aligned} E \{ \tilde{\eta}^2(n+1) \} &= E \{ \mathcal{A}^2 \} + E \{ \mathcal{B}^2 \} + E \{ \mathcal{C}^2 \} + E \{ \mathcal{D}^2 \} \\ &\quad + E \{ 2\mathcal{AB} \} + E \{ 2\mathcal{AC} \} + E \{ 2\mathcal{AD} \} \\ &\quad + E \{ 2\mathcal{BC} \} + E \{ 2\mathcal{BD} \} + E \{ 2\mathcal{CD} \}. \end{aligned} \quad (3.25)$$

Como $r(n)$ é assumido independente dos outros termos e $E \{ r(n) \} = 0$, os termos $E \{ 2\mathcal{AD} \}$, $E \{ 2\mathcal{BD} \}$ e $E \{ 2\mathcal{CD} \}$ são nulos. Além disso, usando T1, T2 e T3, os termos $E \{ 2\mathcal{AB} \}$ e $E \{ 2\mathcal{AC} \}$ são aproximadamente nulos. Assim, (3.25) pode ser reescrita aproximadamente como

$$E \{ \tilde{\eta}^2(n+1) \} \approx E \{ \mathcal{A}^2 \} + E \{ \mathcal{B}^2 \} + E \{ \mathcal{C}^2 \} + E \{ \mathcal{D}^2 \} + E \{ 2\mathcal{BC} \}. \quad (3.26)$$

Para continuar, assume-se ainda que

T5. $e_{a,i}(n)$, $i = 1, 2$ são conjuntamente gaussianos, com média zero.

A hipótese de gaussianidade é freqüentemente usada em filtragem adaptativa para facilitar a análise de transitório mas não é estritamente realista. Mesmo no caso supervisionado de identificação de sistemas, com $u(n)$, $v(n)$ e $d(n)$ gaussianos, o sinal de erro $e_i(n)$ de um filtro componente não é gaussiano. Isso ocorre pois $e_i(n)$ tem uma relação não-linear com $\mathbf{u}(n)$ e $v(n)$ através do vetor de coeficientes $\mathbf{w}_i(n-1)$, que varia conforme sua adaptação. No entanto, se o passo de adaptação for suficientemente pequeno, a hipótese de gaussianidade é razoável [Sayed, 2003].

Usando T5, os seguintes momentos são válidos [Papoulis, 1991]

$$\begin{aligned} E \{ [e_{a,i}(n) - e_{a,j}(n)]^4 \} &= 3 [E \{ [e_{a,i}(n) - e_{a,j}(n)]^2 \}]^2 \\ &= 3 [\Delta\zeta_i(n) + \Delta\zeta_j(n)]^2 \end{aligned} \quad (3.27)$$

$$\begin{aligned} E \{ e_{a,i}^2(n) e_{a,j}^2(n) \} &= E \{ e_{a,i}^2(n) \} E \{ e_{a,j}^2(n) \} + 2 [E \{ e_{a,i}(n) e_{a,j}(n) \}]^2 \\ &= \zeta_i(n) \zeta_j(n) + 2\zeta_{ij}^2(n) \end{aligned} \quad (3.28)$$

$$\begin{aligned} E \{ e_{a,i}^3(n) e_{a,j}(n) \} &= 3 E \{ e_{a,i}^2(n) \} E \{ e_{a,i}(n) e_{a,j}(n) \} \\ &= 3 \zeta_i(n) \zeta_{ij}(n). \end{aligned} \quad (3.29)$$

Com essas hipóteses os termos de (3.26) são calculados a seguir:

Termo $E\{\mathcal{A}^2\}$: Usando T2 e T3, o termo $E\{\mathcal{A}^2\}$ vale aproximadamente

$$\begin{aligned} E\{\mathcal{A}^2\} &= E\left\{[1 - \mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2]^2 \tilde{\eta}^2(n)\right\} \\ &\approx E\left\{[1 + \mu_\eta^2[e_{a,2}(n) - e_{a,1}(n)]^4 - 2\mu_\eta[e_{a,2}(n) - e_{a,1}(n)]^2]\right\} E\{\tilde{\eta}^2(n)\}. \end{aligned} \quad (3.30)$$

Usando (3.27) em (3.30), chega-se a

$$E\{\mathcal{A}^2\} \approx [1 + 3\mu_\eta^2[\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2 - 2\mu_\eta[\Delta\zeta_1(n) + \Delta\zeta_2(n)]] \sigma_\eta^2(n). \quad (3.31)$$

Termo $E\{\mathcal{B}^2\}$: Usando novamente (3.27) para o cálculo do termo $E\{\mathcal{B}^2\}$, obtém-se

$$E\{\mathcal{B}^2\} = E\{\tilde{\eta}_o^2(n)\mu_\eta^2[e_{a,2}(n) - e_{a,1}(n)]^4\} \approx \tilde{\eta}_o^2(n)\mu_\eta^2 3[\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2 = 3\mu_\eta^2 \Delta\zeta_2^2(n). \quad (3.32)$$

Termo $E\{\mathcal{C}^2\}$: O termo $E\{\mathcal{C}^2\}$ é calculado como

$$\begin{aligned} E\{\mathcal{C}^2\} &= E\{\mu_\eta^2 e_{a,2}^2(n)[e_{a,2}(n) - e_{a,1}(n)]^2\} \\ &= \mu_\eta^2 E\{e_{a,2}^4(n) + e_{a,1}^2(n)e_{a,2}^2(n) - 2e_{a,1}(n)e_{a,2}^3(n)\}. \end{aligned} \quad (3.33)$$

Usando (3.28) e (3.29) em (3.33), chega-se a

$$E\{\mathcal{C}^2\} = \mu_\eta^2 [3\zeta_2^2(n) + \zeta_1(n)\zeta_2(n) + 2\zeta_{12}^2(n) - 6\zeta_2(n)\zeta_{12}(n)]. \quad (3.34)$$

Termo $E\{\mathcal{D}^2\}$: O termo $E\{\mathcal{D}^2\}$ é dado por

$$E\{\mathcal{D}^2\} = E\{\mu_\eta^2 r^2(n)[e_{a,2}(n) - e_{a,1}(n)]^2\} = \mu_\eta^2 \sigma_r^2 [\Delta\zeta_1(n) + \Delta\zeta_2(n)], \quad (3.35)$$

sendo $\sigma_r^2 = E\{r^2(n)\}$.

Termo $E\{2\mathcal{BC}\}$: O termo $E\{2\mathcal{BC}\}$ é calculado como

$$\begin{aligned} E\{2\mathcal{BC}\} &= E\{-2\tilde{\eta}_o(n)\mu_\eta^2[e_{a,2}(n) - e_{a,1}(n)]^3 e_{a,2}(n)\} \\ &= -2\mu_\eta^2 \tilde{\eta}_o(n) E\{e_{a,2}^4(n) - e_{a,1}^3(n)e_{a,2}(n) + 3e_{a,1}^2(n)e_{a,2}^2(n) - 3e_{a,2}^3(n)e_{a,1}(n)\} \\ &= -2\mu_\eta^2 \tilde{\eta}_o(n) [3\zeta_2^2(n) - 3\zeta_1(n)\zeta_{12}(n) + 3\zeta_1(n)\zeta_2(n) + 6\zeta_{12}^2(n) \\ &\quad - 9\zeta_2(n)\zeta_{12}(n)]. \end{aligned} \quad (3.36)$$

Usando (3.4) em (3.36), chega-se a

$$E\{2\mathcal{BC}\} = -6\mu_\eta^2 \Delta\zeta_2^2(n). \quad (3.37)$$

Usando esses resultados em (3.26), obtém-se

$$\begin{aligned}\sigma_\eta^2(n+1) = E\{\tilde{\eta}^2(n+1)\} \approx & \{1 + 3\mu_\eta^2[\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2 - 2\mu_\eta[\Delta\zeta_1(n) + \Delta\zeta_2(n)]\} \sigma_\eta^2(n) \\ & + \mu_\eta^2\sigma_r^2[\Delta\zeta_1(n) + \Delta\zeta_2(n)] \\ & + \mu_\eta^2[\zeta_1(n)\zeta_2(n) - \zeta_{12}^2(n)] .\end{aligned}\quad (3.38)$$

Com esse resultado e usando as expressões (3.7) e (3.4), é possível obter um modelo teórico para o EMSE da combinação em cada instante de tempo. Porém, são necessárias ainda expressões para $\zeta_i(n)$, $i = 1, 2$ e $\zeta_{12}(n)$ para os algoritmos utilizados nos filtros componentes. Na seção seguinte, essas expressões são apresentadas para as combinações de algoritmos de mesma família. A extensão dos resultados para algoritmos de famílias diferentes será tratada num trabalho futuro.

3.3 Expressões para o EMSE de filtros adaptativos

De modo geral, é necessário obter uma expressão para

$$\begin{aligned}\zeta_{ij}(n) = E\{e_{a,i}(n)e_{a,j}(n)\} &= E\{\mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1)\mathbf{u}^T(n)\tilde{\mathbf{w}}_j(n-1)\} = \\ &= E\{\text{Tr}(\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1)\tilde{\mathbf{w}}_j^T(n-1))\} , \quad i, j = 1, 2.\end{aligned}\quad (3.39)$$

Assumindo que $\mathbf{u}(n)$ é independente de $\tilde{\mathbf{w}}_k(n)$, $k = 1, 2$, obtém-se¹

$$\begin{aligned}\zeta_{ij}(n) = E\{e_{a,i}(n)e_{a,j}(n)\} &\approx \text{Tr}(\mathbf{R}E\{\tilde{\mathbf{w}}_i(n-1)\tilde{\mathbf{w}}_j^T(n-1)\}) \\ &\approx \text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1)) ,\end{aligned}\quad (3.40)$$

em que $\mathbf{S}_{ij}(n) \triangleq E\{\tilde{\mathbf{w}}_i(n)\tilde{\mathbf{w}}_j^T(n)\}$ é a matriz de covariância do vetor de erro dos coeficientes.

De (3.40) é possível calcular $\zeta_i(n)$, $i = 1, 2$ através de

$$\zeta_i(n) = \zeta_{ii}(n) = \text{Tr}(\mathbf{R}\mathbf{S}_{ii}(n-1)) , \quad i = 1, 2, \quad (3.41)$$

e $\zeta_{12}(n)$, através de

$$\zeta_{12}(n) = \text{Tr}(\mathbf{R}\mathbf{S}_{12}(n-1)) . \quad (3.42)$$

Na literatura existem diversos trabalhos sobre a análise do transitório de filtros adaptativos. Uma expressão de $\mathbf{S}_{ii}(n)$ para o LMS, por exemplo, foi obtida em [Farhang-Boroujeny, 1998], para o NLMS em [Costa e Bermudez, 2002] e para o CMA em [Nascimento e Silva, 2008]. Para o RLS, no entanto, as expressões para $\mathbf{S}_{ii}(n)$ que foram encontradas na literatura não levam em conta a inicialização do algoritmo e valem apenas para n suficientemente grande para que

¹Essa hipótese faz parte de um conjunto de hipóteses de independência amplamente utilizado em filtragem adaptativa [Haykin, 2002].

a inicialização tenha sido esquecida (por exemplo, [Haykin, 2002, Farhang-Boroujeny, 1998]). No Apêndice D, os resultados para o LMS, NLMS e CMA foram estendidos para o cálculo de $\mathbf{S}_{ij}(n)$ e, para o RLS, foi calculada uma aproximação para $\mathbf{S}_{ij}(n)$, $i, j = 1, 2$ levando em conta a inicialização do algoritmo. Os resultados são mostrados na Tabela 3.1, onde

$$\widehat{\mathbf{R}}_i^{-1}(n) = \mathbb{E} \left\{ \widehat{\mathbf{R}}_i^{-1}(n) \right\} = \left[\lambda_i^n \epsilon_{R,i} \mathbf{I} + \frac{1 - \lambda_i^n}{1 - \lambda_i} \mathbf{R} \right]^{-1}, \quad (3.43)$$

sendo λ_i o fator de esquecimento do filtro i e $\epsilon_{R,i}$ uma constante utilizada na inicialização da estimativa da matriz de autocorrelação, ou seja, $\widehat{\mathbf{R}}_i(0) = \epsilon_{R,i} \mathbf{I}$.

Tabela 3.1: Expressões para o cálculo de $\mathbf{S}_{ij}(n)$ no caso estacionário.

Algoritmo	$\mathbf{S}_{ij}(n)$
LMS	$\mathbf{S}_{ij}(n) \approx \mathbf{S}_{ij}(n-1) - \mu_i \mathbf{R} \mathbf{S}_{ij}(n-1) - \mu_j \mathbf{S}_{ij}(n-1) \mathbf{R}$ $+ \mu_i \mu_j [2 \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}]$
NLMS	$\mathbf{S}_{ij}(n) \approx \mathbf{S}_{ij}(n-1) - \frac{\tilde{\mu}_i}{\sigma_u^2(M-2)} \mathbf{R} \mathbf{S}_{ij}(n-1) - \frac{\tilde{\mu}_j}{\sigma_u^2(M-2)} \mathbf{S}_{ij}(n-1) \mathbf{R}$ $+ \frac{\tilde{\mu}_i \tilde{\mu}_j}{\sigma_u^4(M-2)(M-4)} [2 \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}]$
RLS	$\mathbf{S}_{ij}(n) \approx \mathbf{S}_{ij}(n-1) - \widehat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \mathbf{S}_{ij}(n-1) - \mathbf{S}_{ij}(n-1) \mathbf{R} \widehat{\mathbf{R}}_j^{-1}(n)$ $+ \widehat{\mathbf{R}}_i^{-1}(n) [2 \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}] \widehat{\mathbf{R}}_j^{-1}(n)$
CMA	$\mathbf{S}_{ij}(n) \approx \mathbf{S}_{ij}(n-1) - \mu_i \bar{\gamma} \mathbf{R} \mathbf{S}_{ij}(n-1) - \mu_j \bar{\gamma} \mathbf{S}_{ij}(n-1) \mathbf{R}$ $+ \mu_i \mu_j [2 \xi \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \xi \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_\beta^2 \mathbf{R}]$

3.4 Resultados de Simulação

Para verificar o comportamento da combinação afim e a validade da análise, foram feitas simulações considerando os algoritmos LMS, NLMS, RLS e CMA. Como nas simulações do Capítulo 2, no caso supervisionado, foi considerada uma aplicação de identificação de sistemas e duas soluções ótimas formadas por $M = 7$ coeficientes aleatórios dados por

$$\mathbf{w}_{o,1}^T = [0,90 \quad -0,54 \quad 0,21 \quad -0,03 \quad 0,78 \quad 0,52 \quad -0,09]$$

e

$$\mathbf{w}_{o,2}^T = [0,25 \quad -0,47 \quad -0,38 \quad 0,05 \quad -0,18 \quad 0,79 \quad 0,15].$$

O sinal de entrada colorido, quando usado, foi gerado através de um filtro autoregressivo de primeira ordem com função de transferência $\sqrt{(1-a^2)}/(1-az^{-1})$, com $a = 0,8$. Tanto para sinal de entrada colorido como para sinal de entrada branco, foi considerada variância

$1/M$, o que leva a $\text{Tr}(\mathbf{R}) = 1$. Para formar o sinal desejado, é somado um sinal de ruído branco e gaussiano $v(n)$, com variância $\sigma_v^2 = 0,01$. No caso autodidata, foi considerada a combinação de dois equalizadores com $M = 4$ coeficientes cada, usando sobreamostragem de fator $S = 2$ com sinal de entrada do tipo 4-PAM e dois canais com coeficientes [Arenas-García e Figueiras-Vidal, 2006, Yousef e Sayed, 1999]

$$\mathbf{h}_1^T = [0,1 \quad 0,3 \quad 1,0 \quad -0,1 \quad 0,5 \quad 0,2]$$

e

$$\mathbf{h}_2^T = [0,25 \quad 0,64 \quad 0,80 \quad -0,55].$$

Para facilitar a visualização dos resultados no caso autodidata, as curvas foram filtradas por um filtro de média-móvel de 128 coeficientes.

Nas Figuras 3.1 a 3.4, são mostrados, respectivamente, os resultados para a combinação de dois algoritmos LMS, dois NLMS, dois RLS e dois CMA com o parâmetro $\eta(n)$ adaptado através do algoritmo η -LMS e $\delta_{XX} = 0,1$. As simulações do caso supervisionado foram iniciadas com entrada branca e solução ótima dada por $\mathbf{w}_{o,1}$ e, durante a simulação, a solução ótima foi alterada para $\mathbf{w}_{o,2}$ com entrada colorida. No caso do CMA, a simulação foi iniciada usando o canal $\mathbf{h}_1$ e durante a simulação alterou-se para o canal $\mathbf{h}_2$. Ambos os filtros componentes foram inicializados com um vetor pino com um elemento unitário na segunda posição. Para a combinação de dois algoritmos LMS, pode-se verificar que o resultado teórico apresenta uma boa concordância com o resultado experimental. Para o NLMS e o RLS, os resultados não apresentam uma concordância tão boa quanto no caso do LMS. Isso ocorre porque as estimativas teóricas de $\zeta_i(n)$, $i = 1,2$ e $\zeta_{12}(n)$ para esses algoritmos não são tão boas quanto as estimativas desses parâmetros para o LMS quando é usado um número de coeficientes M pequeno. No caso do NLMS, é possível obter um modelo mais preciso usando o modelo do LMS mas considerando uma matriz de autocorrelação do sinal de entrada e a potência do ruído de medida modificados [Nascimento, 2001]. Porém, apesar de mais preciso, esses parâmetros modificados só podem ser calculados experimentalmente. No caso do RLS, o problema ocorre devido à aproximação do termo de quarta ordem que aparece na matriz de covariância do vetor de erro de coeficientes apresentada no Apêndice D.

No caso do CMA, a diferença entre a análise e o resultado experimental decorre principalmente devido às hipóteses que precisam ser assumidas para permitir a análise. Para tornar mais realista a hipótese R2 (pág. 18), na Figura 3.5, é mostrado o resultado para a mesma situação da Figura 3.4 mas com os vetores de coeficientes dos filtros componentes inicializados próximos da solução ótima de *zero-forcing* em $0,9\hat{\mathbf{w}}_{zf}$, onde a estimativa para a solução de *zero-forcing* $\hat{\mathbf{w}}_{zf}$ foi considerada igual a média do vetor de coeficientes dos filtros componentes em regime. É possível verificar que a estimativa dada pela análise nesse caso é mais coerente com o resultado experimental. No entanto, ainda há uma diferença, devido às demais hipóteses que são assumidas na análise.

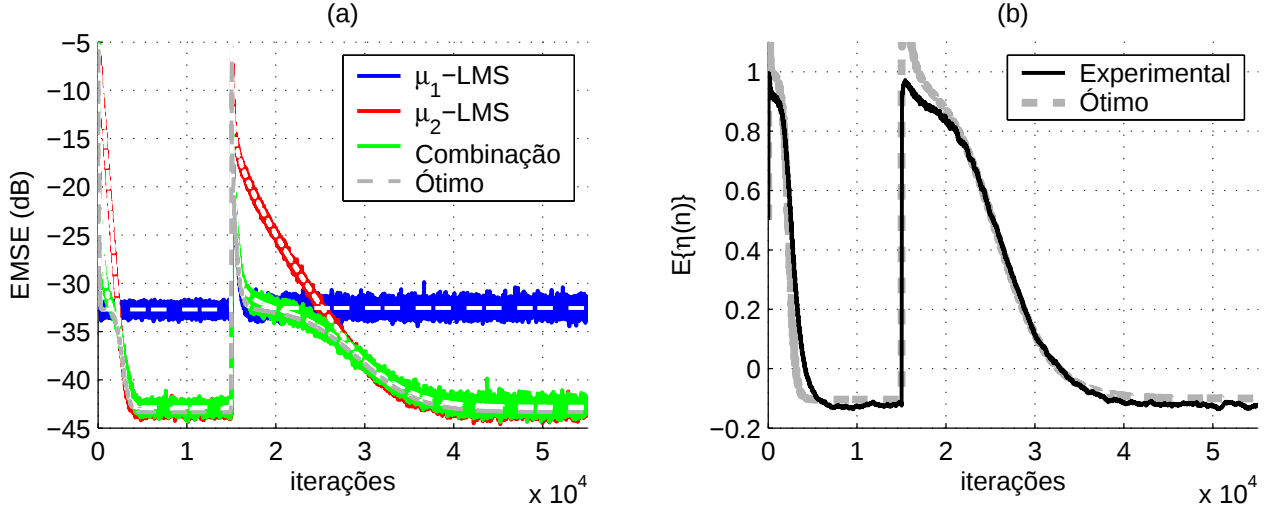


Figura 3.1: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,1$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados.

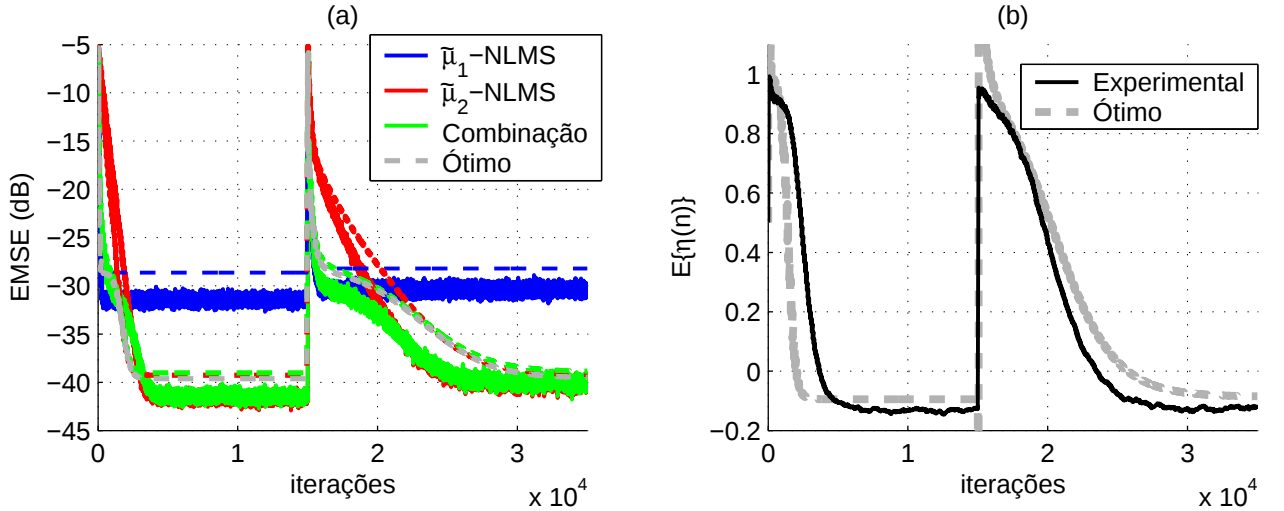


Figura 3.2: Combinação afim de $\tilde{\mu}_1$ -NLMS e $\tilde{\mu}_2$ -NLMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\tilde{\mu}_1 = 0,1$, $\tilde{\mu}_2 = \delta_{NN}\tilde{\mu}_1$, $\delta_{NN} = 0,1$ e $\epsilon_N = 1 \times 10^{-5}$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1,5 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1,5 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados.

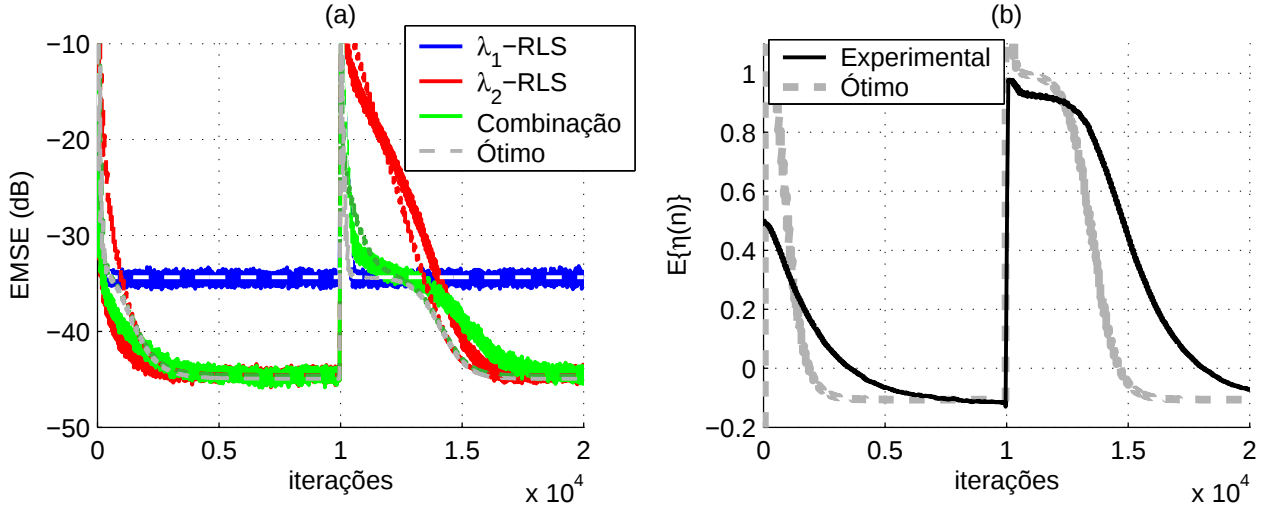


Figura 3.3: Combinação afim de λ_1 -RLS e λ_2 -RLS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\lambda_1 = 0,99$, $\lambda_2 = 1 - \delta_{RR}(1 - \lambda_1)$, $\delta_{RR} = 0,1$ e $\epsilon_R = 0,02$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ até $n = 1 \times 10^4$ e entrada colorida (modelo AR com pólo em 0,8) e solução ótima $\mathbf{w}_{o,2}$ a partir de $n = 1 \times 10^4$. Sinal de entrada com potência $1/M$. Média de 500 realizações com $\mu_\eta = 2$. Valores teóricos tracejados.

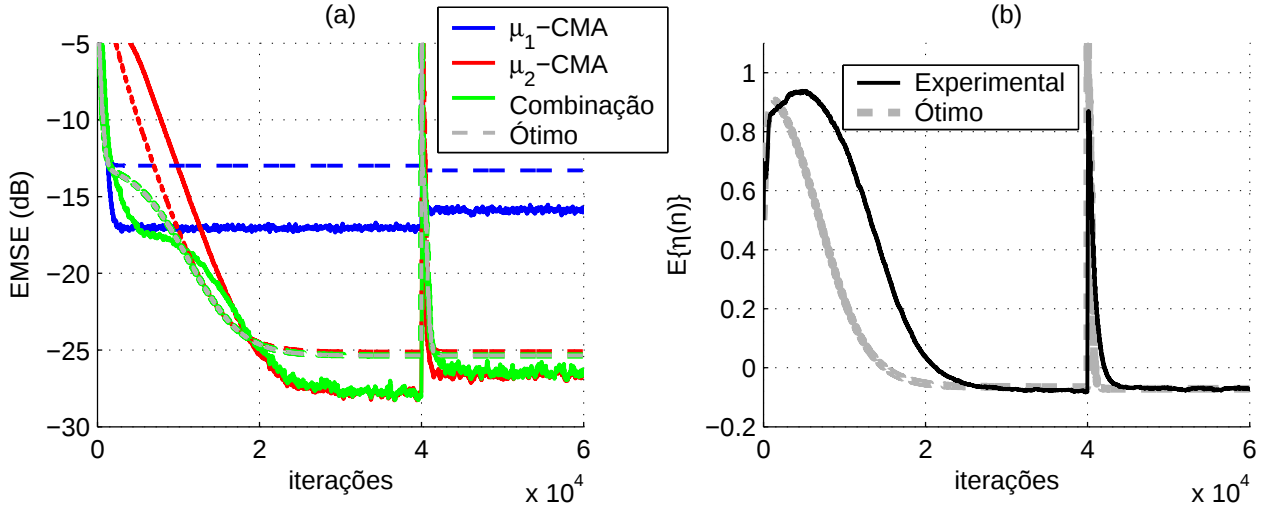


Figura 3.4: Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetor de coeficientes inicializados com um vetor pino com um elemento unitário na segunda posição. Canal $\mathbf{h}_1$ até $n = 4 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 4 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados.

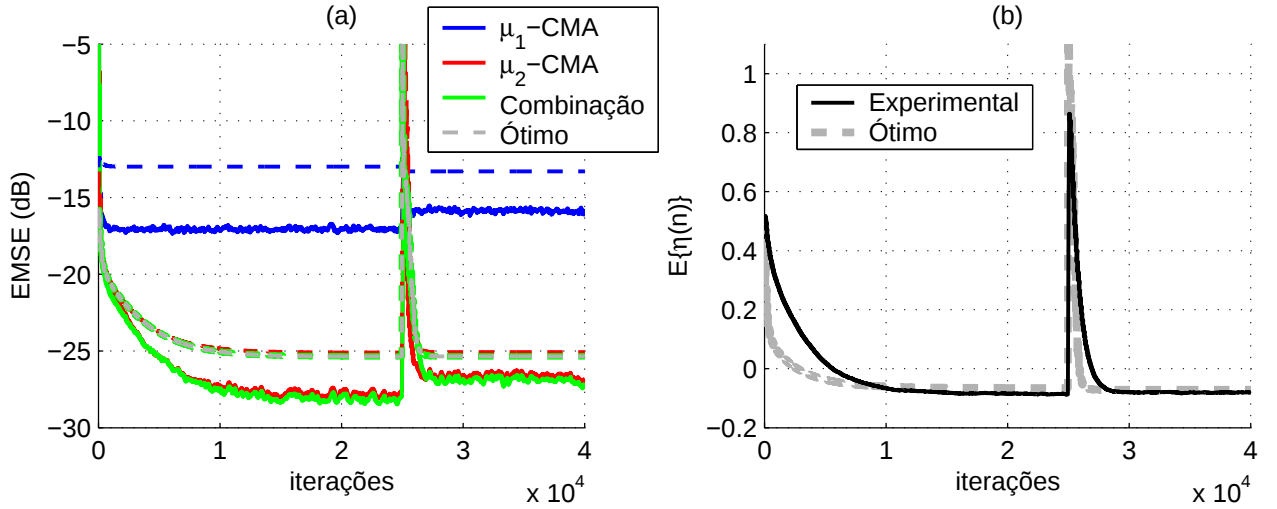


Figura 3.5: Combinação afim de μ_1 -CMA e μ_2 -CMA no caso estacionário. Sinal de entrada 4-PAM, $M = 4$ com sobreamostragem de fator $S = 2$, $\mu_1 = 1 \times 10^{-3}$, $\mu_2 = \delta_{CC}\mu_1$ e $\delta_{CC} = 0,1$. Vetor de coeficientes inicializados em $0,9\mathbf{w}_{zf}$, sendo $\mathbf{w}_{zf}$ assumido como o valor médio dos coeficientes em regime. Canal $\mathbf{h}_1$ até $n = 2,5 \times 10^4$ e canal $\mathbf{h}_2$ a partir de $n = 2,5 \times 10^4$. Média de 500 realizações com $\mu_\eta = 0,075$. Valores teóricos tracejados.

Cabe observar que em todos os casos, o valor teórico do parâmetro de mistura ótimo, em módulo, é muito grande nos instantes iniciais da simulação e quando é feita a troca de solução ótima ou canal. Isso ocorre porque nesses instantes os níveis de EMSE dos filtros componentes são muito próximos e nesse caso, de acordo com o resultado obtido no Capítulo 2, o valor absoluto do parâmetro de mistura ótimo deve ser grande para permitir um ganho na combinação. Entretanto, como a variação do gradiente ao redor desses instantes é muito grande, é necessário restringir o parâmetro de mistura tal que $\eta(n) < 1$ para evitar a divergência do algoritmo e isso faz com que o parâmetro de mistura não atinja seu valor ótimo nesses instantes.

A adaptação do parâmetro de mistura com o algoritmo η -LMS, nem sempre é capaz de fazer com que o comportamento da combinação esteja próximo do universal. Uma situação em que isso ocorre é mostrada nas Figuras 3.6 e 3.7. Em ambas as figuras, é considerada a combinação de dois filtros LMS com passos de adaptação $\mu_1 = 0,01$ e $\mu_2 = \delta_{LL}\mu_1$ com $\delta_{LL} = 0,1$ e solução ótima $\mathbf{w}_{o,1}$. Com esses passos de adaptação, a convergência dos filtros componentes torna-se mais lenta e a escolha do passo de adaptação da combinação para tornar o comportamento da combinação próximo do universal é dificultada.

Na Figura 3.6, é mostrado o resultado para $\mu_\eta = 3$. Com essa escolha, a adaptação do parâmetro de mistura é rápida o suficiente para permitir o chaveamento dos filtros por volta do instante $n = 3 \times 10^4$ e fazer com que a combinação tenha um comportamento próximo ao universal no chaveamento. Entretanto, nos instantes iniciais, quando os dois filtros componentes têm saídas muito distintas, pode-se verificar que $\mu_\eta = 3$ é um passo grande demais, fazendo com haja uma variação maior no valor do parâmetro $\eta(n)$ no início das iterações, conforme

pode ser notado na Figura 3.6-(c). Isso faz com que a combinação tenha um desempenho inferior ao universal.

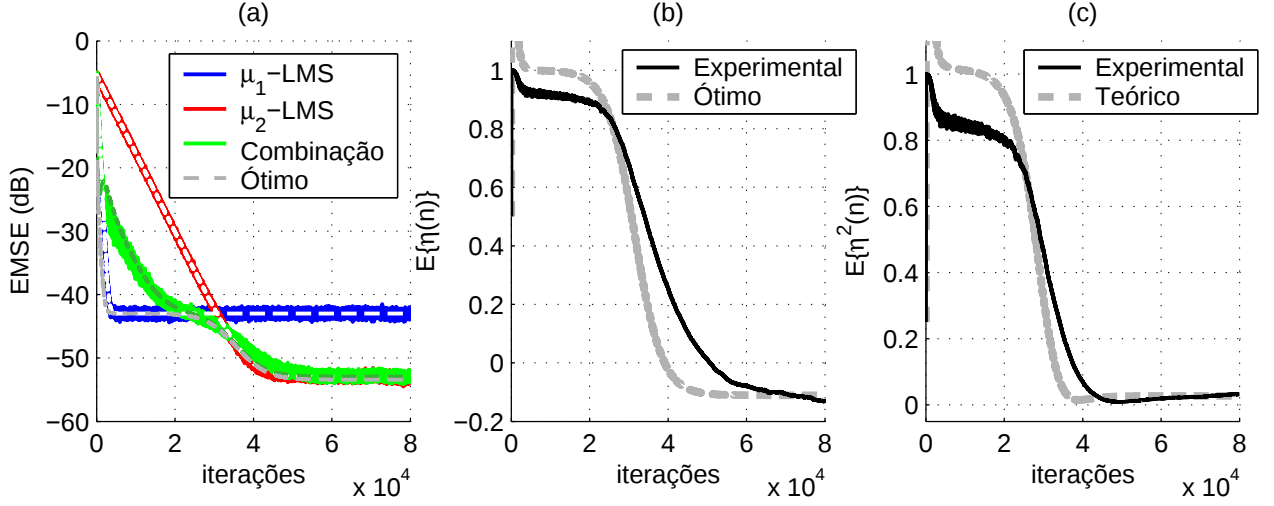


Figura 3.6: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_\eta = 3$. Valores teóricos tracejados.

Na Figura 3.7, é mostrado o resultado obtido quando o passo de adaptação da combinação é diminuído para $\mu_\eta = 0,1$. Nesse caso, devido à diminuição do passo, a combinação mostra um comportamento próximo ao universal durante os instantes iniciais mas apresenta um problema na velocidade do chaveamento entre os filtros componentes. Pode-se notar que o modelo teórico para o EMSE da combinação não corresponde ao resultado experimental. Enquanto o modelo teórico prevê o chaveamento entre os filtros componentes de forma adequada, o resultado experimental apresenta uma velocidade de convergência lenta nos instantes de chaveamento entre os filtros devido à utilização do passo de adaptação com valor pequeno. Essa diferença entre os resultados teórico e experimental é causada devido a utilização da hipótese T1, que pressupõe que a esperança matemática do parâmetro de mistura da combinação é igual à esperança matemática do parâmetro de mistura ótimo. Nesse caso, essa hipótese não é razoável, já que o passo de adaptação utilizado é muito pequeno, fazendo com que haja uma grande diferença entre o parâmetro de mistura da combinação e o parâmetro de mistura ótimo, principalmente durante o chaveamento, conforme pode ser observado na Figura 3.7-(b). Apesar da limitação do modelo nessa situação, o comportamento em situações mais práticas é coerente, pois em grande parte dos casos, o parâmetro de mistura converge, em média, para seu valor ótimo, considerando uma pequena diferença nos instantes de chaveamento. Sendo assim, o cálculo de uma expressão mais precisa para o EMSE da combinação, sem considerar a hipótese T1 será tratado num trabalho futuro.

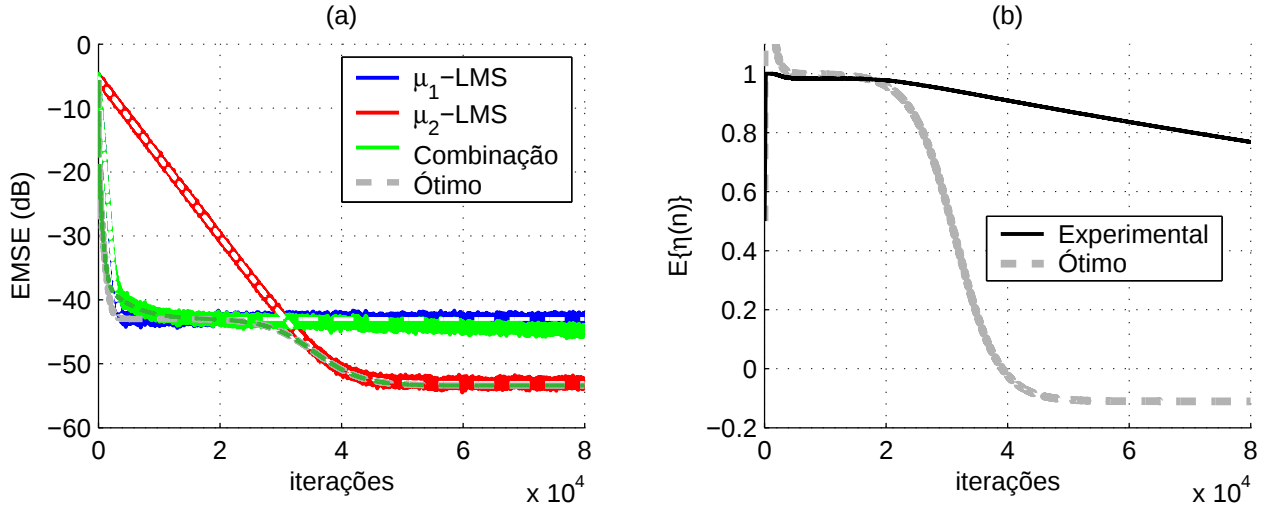


Figura 3.7: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_\eta = 0,1$. Valores teóricos tracejados.

Para fazer com que a combinação tenha um comportamento mais próximo do universal, outros tipos de adaptação podem ser usados para o parâmetro de mistura com o objetivo de conseguir acelerar a adaptação. A seguir, são propostas três alternativas para a adaptação do parâmetro de mistura.

3.5 Propostas para adaptação do parâmetro de mistura

A adaptação de $\eta(n)$ através do algoritmo η -LMS algumas vezes não é capaz de acompanhar as variações abruptas de $\bar{\eta}_o(n)$ e a combinação acaba tendo um desempenho inferior ao universal. Isso ocorre por exemplo, quando ambos os filtros componentes já convergiram e $[y_1(n) - y_2(n)]$ é um valor pequeno, como mostrado na Figura 3.7. Essa adaptação também pode causar divergência no início das iterações quando $[y_1(n) - y_2(n)]$ é um valor grande. Uma forma de contornar o problema da divergência é limitar o valor de $\eta(n)$ tal que $\eta(n) < 1$ [Bershad et al., 2008]. Esse método, no entanto, em algumas situações faz com que o comportamento da combinação seja inferior ao universal, como mostrado na Figura 3.7. A seguir, é proposta uma alternativa para acelerar a adaptação de $\eta(n)$ baseada numa normalização instantânea parcial, originando o algoritmo *signed regressor* (SR) para a adaptação do parâmetro de mistura. Em seguida, são propostos esquemas de normalização para os algoritmos η -LMS e SR, originando os algoritmos *power normalized* e *normalized signed regressor*.

3.5.1 O algoritmo *signed regressor*

Uma alternativa para acelerar a adaptação de $\eta(n)$ é usar algum tipo de normalização na adaptação. Com a normalização, a restrição de $\eta(n) < 1$ para evitar o problema da divergência pode não ser mais necessária e a combinação pode apresentar um comportamento mais próximo ao universal no início das iterações. Uma possibilidade é usar a normalização instantânea. Nesse caso, $[y_1(n) - y_2(n)]$ é o sinal equivalente ao sinal de entrada na adaptação de $\eta(n)$ e a normalização é obtida fazendo o passo de adaptação igual a

$$\mu_\eta \triangleq \frac{\mu_{\eta,s}}{[y_1(n) - y_2(n)]^2}, \quad (3.44)$$

sendo $\mu_{\eta,s}$ uma constante positiva. Porém, pelo fato do sinal $[y_1(n) - y_2(n)]$ variar muito, esse tipo de normalização leva a uma variação muito grande no valor de $\eta(n)$ e até mesmo à divergência [Bershad, 1986].

Outra possibilidade é usar uma normalização instantânea parcial, com o passo de adaptação

$$\mu_\eta \triangleq \frac{\mu_{\eta,s}}{|y_1(n) - y_2(n)|}. \quad (3.45)$$

Com isso, a equação de adaptação de $\eta(n)$ passa a ser dada por

$$\eta(n+1) = \eta(n) + \mu_{\eta,s} e_g(n) \frac{(y_1(n) - y_2(n))}{|y_1(n) - y_2(n)|}, \quad (3.46)$$

que pode ser aproximada por

$$\eta(n+1) = \eta(n) + \mu_{\eta,s} e_g(n) \text{sign}(y_1(n) - y_2(n)), \quad (3.47)$$

em que

$$\text{sign}(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases} \quad (3.48)$$

e $e_g(n)$ é definido em (2.30) (pág. 20). Essa normalização evita a grande variação do parâmetro de mistura nas iterações iniciais, fazendo com que a restrição imposta anteriormente para $\eta(n) \leq 1$ não seja mais necessária. Isso possibilita um melhor desempenho no início da convergência, conforme será mostrado através de simulações.

O algoritmo (3.47) é chamado de *signed regressor* e seu comportamento também pode ser estudado através do modelo (3.7). Para isso, é necessário calcular os momentos de primeira e segunda ordens a fim de verificar a validade da hipótese T1 em regime e obter uma expressão para $\sigma_\eta^2(n)$.

Momento de primeira ordem

Procedendo de maneira análoga à da Seção 3.2, subtrai-se ambos os lados de (3.47) de $\bar{\eta}_o(n)$, obtendo-se

$$\bar{\eta}_o(n) - \eta(n+1) = \bar{\eta}_o(n) - \eta(n) - \mu_{\eta,s} e_g(n) \text{sign}(y_1(n) - y_2(n)). \quad (3.49)$$

Substituindo (3.14) e (3.15) em (3.49), chega-se a

$$\begin{aligned} \bar{\eta}_o(n) - \eta(n+1) &= \bar{\eta}_o(n) - \eta(n) + \mu_{\eta,s} \eta(n) [e_{a,2}(n) - e_{a,1}(n)] \text{sign}(e_{a,2}(n) - e_{a,1}(n)) \\ &\quad - \mu_{\eta,s} e_{a,2}(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n)) \\ &\quad - \mu_{\eta,s} r(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n)). \end{aligned} \quad (3.50)$$

Considerando uma hipótese semelhante a T4 para (3.50), somando e subtraindo o termo $\bar{\eta}_o(n) \mu_{\eta,s} [e_{a,2} - e_{a,1}] \text{sign}(e_{a,2}(n) - e_{a,1}(n))$ do lado direito (3.50), obtém-se

$$\begin{aligned} \tilde{\eta}(n+1) &= [1 - \mu_{\eta,s} [e_{a,2}(n) - e_{a,1}(n)] \text{sign}(e_{a,2}(n) - e_{a,1}(n))] \tilde{\eta}(n) \\ &\quad + \mu_{\eta,s} \bar{\eta}_o(n) [e_{a,2}(n) - e_{a,1}(n)] \text{sign}(e_{a,2}(n) - e_{a,1}(n)) \\ &\quad - \mu_{\eta,s} e_{a,2}(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n)) - \mu_{\eta,s} r(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n)). \end{aligned} \quad (3.51)$$

Como $\text{sign}(x)x = |x|$, (3.51) pode ser reescrita como

$$\begin{aligned} \tilde{\eta}(n+1) &= \overbrace{[1 - \mu_{\eta,s} |e_{a,2}(n) - e_{a,1}(n)|] \tilde{\eta}(n)}^{\mathcal{A}} + \overbrace{\mu_{\eta,s} \bar{\eta}_o(n) |e_{a,2}(n) - e_{a,1}(n)|}^{\mathcal{B}} \\ &\quad - \underbrace{\mu_{\eta,s} e_{a,2}(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n))}_{\mathcal{C}} - \underbrace{\mu_{\eta,s} r(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n))}_{\mathcal{D}}. \end{aligned} \quad (3.52)$$

Aplicando a esperança de ambos os lados em (3.52), usando T3 e considerando que $r(n)$ é independente de $e_{a,i}(n)$ com $E\{r(n)\} = 0$, chega-se a

$$\begin{aligned} E\{\tilde{\eta}(n+1)\} &= E\{[1 - \mu_{\eta,s} |e_{a,2}(n) - e_{a,1}(n)|] \tilde{\eta}(n)\} \\ &\quad + \mu_{\eta,s} \bar{\eta}_o(n) E\{|e_{a,2}(n) - e_{a,1}(n)|\} \\ &\quad - \mu_{\eta,s} E\{e_{a,2}(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n))\}. \end{aligned} \quad (3.53)$$

A validade de T1 em regime será garantida se os dois últimos termos do lado direito de (3.53) se anularem, ou seja,

$$\mu_{\eta,s} \bar{\eta}_o(n) E\{|e_{a,2}(n) - e_{a,1}(n)|\} - \mu_{\eta,s} E\{e_{a,2}(n) \text{sign}(e_{a,2}(n) - e_{a,1}(n))\} = 0 \quad (3.54)$$

e se o termo que multiplica $E\{\tilde{\eta}(n)\}$ for, em módulo, menor que um, ou seja,

$$|E\{1 - \mu_{\eta,s}|e_{a,2}(n) - e_{a,1}(n)|\}| < 1. \quad (3.55)$$

Para prosseguir, é necessário assumir a hipótese T5. A motivação para isso é que a hipótese de gaussianidade permite expressar a esperança

$$E\{e_{a,2}(n)\text{sign}(e_{a,2}(n) - e_{a,1}(n))\}$$

em termos de

$$E\{e_{a,2}(n)(e_{a,2}(n) - e_{a,1}(n))\}.$$

Isso é possível devido a um caso especial do teorema de Price [Sayed, 2003, pág. 306]. Esse caso estabelece que para duas variáveis aleatórias gaussianas x e y com variâncias σ_x^2 e σ_y^2 , vale

$$E\{x \cdot \text{sign}(y)\} = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_y} E\{x \cdot y\}. \quad (3.56)$$

Além disso, se x é uma variável aleatória gaussiana, é possível demonstrar que [Papoulis, 1991]

$$E\{|x|\} = \sqrt{\frac{2}{\pi}} \sigma_x. \quad (3.57)$$

Aplicando esses resultados nos dois termos do lado esquerdo de (3.54), obtém-se

$$\mu_{\eta,s} \bar{\eta}_o(n) E\{|e_{a,2}(n) - e_{a,1}(n)|\} = \mu_{\eta,s} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \bar{\eta}_o(n) \quad (3.58)$$

e

$$\begin{aligned} \mu_{\eta,s} E\{e_{a,2}(n)\text{sign}(e_{a,2}(n) - e_{a,1}(n))\} &= \mu_{\eta,s} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}} [\zeta_2(n) - \zeta_1(n)] \\ &= \mu_{\eta,s} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \bar{\eta}_o(n). \end{aligned} \quad (3.59)$$

Usando a hipótese T5, como esses termos são iguais, a igualdade de (3.54) é garantida.

Usando (3.57) em (3.55) chega-se a um intervalo para o passo de adaptação

$$|E\{[1 - \mu_{\eta,s}|e_{a,2}(n) - e_{a,1}(n)|]\}| = \left|1 - \mu_{\eta,s} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}\right| < 1, \quad (3.60)$$

ou seja,

$$0 < \mu_{\eta,s} < \frac{\sqrt{2\pi}}{\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}}. \quad (3.61)$$

Assim, a escolha do passo $\mu_{\eta,s}$ no intervalo (3.61) garante a validade de T1 em regime. Cabe observar que com esse algoritmo, assim como com o η -LMS, a escolha do passo de adaptação depende de características do cenário de filtragem. Isso ocorre pois o intervalo para o passo de adaptação é limitado por um valor que envolve os termos $\Delta\zeta_1(n)$ e $\Delta\zeta_2(n)$.

Para usar a expressão (3.7) com o objetivo de se obter um modelo para o EMSE da combinação, é necessário ainda obter uma expressão para $\sigma_\eta^2(n)$. A seguir, é feito o cálculo do momento de segunda ordem para obtenção dessa expressão.

Momento de segunda ordem

A expressão para $\sigma_\eta^2(n)$ é obtida elevando (3.52) ao quadrado. Novamente, como $r(n)$ é assumido independente dos outros termos e $E\{r(n)\} = 0$, assumindo a hipótese T1, a esperança do quadrado de (3.52) pode ser aproximada por

$$E\{\tilde{\eta}^2(n+1)\} \approx E\{\mathcal{A}^2\} + E\{\mathcal{B}^2\} + E\{\mathcal{C}^2\} + E\{\mathcal{D}^2\} + E\{2\mathcal{BC}\}. \quad (3.62)$$

Os termos de (3.62) são calculados a seguir.

Termo $E\{\mathcal{A}^2\}$: Usando T2 e T3, o termo $E\{\mathcal{A}^2\}$ é dado por

$$\begin{aligned} E\{\mathcal{A}^2\} &= E\{[1 - \mu_{\eta,s}|e_{a,2}(n) - e_{a,1}(n)|]^2 \tilde{\eta}^2(n)\} \\ &\approx E\{[1 + \mu_{\eta,s}^2|e_{a,2}(n) - e_{a,1}(n)|^2 - 2\mu_{\eta,s}|e_{a,2}(n) - e_{a,1}(n)|]\} E\{\tilde{\eta}^2(n)\}. \end{aligned} \quad (3.63)$$

Usando (3.57) e considerando que $|x|^2 = x^2$ para $x \in \mathbb{R}$, chega-se a

$$E\{\mathcal{A}^2\} \approx \left[1 + \mu_{\eta,s}^2 [\Delta\zeta_1(n) + \Delta\zeta_2(n)] - 2\mu_{\eta,s} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \right] E\{\tilde{\eta}^2(n)\}. \quad (3.64)$$

Termo $E\{\mathcal{B}^2\}$: Lembrando que para uma variável x real, $|x|^2 = x^2$, o termo $E\{\mathcal{B}^2\}$ é calculado como

$$E\{\mathcal{B}^2\} = E\{\bar{\eta}_o^2(n) \mu_{\eta,s}^2 |e_{a,2}(n) - e_{a,1}(n)|^2\} \approx \bar{\eta}_o^2(n) \mu_{\eta,s}^2 [\Delta\zeta_1(n) + \Delta\zeta_2(n)]. \quad (3.65)$$

Termo $E\{\mathcal{C}^2\}$: O termo $E\{\mathcal{C}^2\}$ é dado por

$$E\{\mathcal{C}^2\} = E\{\mu_{\eta,s}^2 e_{a,2}^2(n) \text{sign}^2(e_{a,2}(n) - e_{a,1}(n))\}, \quad (3.66)$$

que, considerando que $\text{sign}^2(n) = 1$, pode ser reescrito como

$$E\{\mathcal{C}^2\} = \mu_{\eta,s}^2 \zeta_2(n). \quad (3.67)$$

Termo $E\{\mathcal{D}^2\}$: Usando novamente o fato que $\text{sign}^2(n) = 1$, o termo $E\{\mathcal{D}^2\}$ é calculado como

$$E\{\mathcal{D}^2\} = E\{\mu_{\eta,s}^2 r^2(n) \text{sign}^2(e_{a,2}(n) - e_{a,1}(n))\} = \mu_{\eta,s}^2 \sigma_r^2. \quad (3.68)$$

Termo $E\{2\mathcal{BC}\}$: Usando o fato que $|x|\text{sign}(x) = x$, obtém-se

$$\begin{aligned} E\{2\mathcal{BC}\} &= E\{-2\bar{\eta}_o(n)\mu_{\eta,s}^2 |e_{a,2}(n) - e_{a,1}(n)| \text{sign}(e_{a,2}(n) - e_{a,1}(n)) e_{a,2}(n)\} \\ &= -2\bar{\eta}_o(n)\mu_{\eta,s}^2 E\{[e_{a,2}(n) - e_{a,1}(n)] e_{a,2}(n)\} \\ &= -2\bar{\eta}_o(n)\mu_{\eta,s}^2 \Delta\zeta_2(n). \end{aligned} \quad (3.69)$$

Substituindo esses resultados em (3.62), chega-se a

$$\begin{aligned} \sigma_\eta^2(n+1) &= E\{\tilde{\eta}^2(n+1)\} \\ &\approx \left[1 + \mu_{\eta,s}^2 [\Delta\zeta_1(n) + \Delta\zeta_2(n)] - 2\mu_{\eta,s} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \right] \sigma_\eta^2(n) \\ &\quad - \mu_{\eta,s}^2 \frac{\Delta\zeta_2^2(n)}{\Delta\zeta_1(n) + \Delta\zeta_2(n)} + \mu_{\eta,s}^2 (\zeta_2(n) + \sigma_r^2). \end{aligned} \quad (3.70)$$

Substituindo (3.70) em (3.7) e usando os resultados para $\zeta_i(n)$, $i = 1, 2$ e $\zeta_{12}(n)$ mostrados na Seção 3.3, é possível obter o EMSE da combinação cujo parâmetro de mistura é adaptado com o algoritmo *signed regressor* em qualquer instante de tempo.

3.5.2 Algoritmos normalizados

O algoritmo η -LMS para adaptação do parâmetro de mistura, assim como o algoritmo SR faz com que seja necessária a estimação de parâmetros do cenário de filtragem para a escolha de seus passos de adaptação. Isso pode ser notado através dos intervalos de passo de adaptação (3.24) e (3.61) obtidos para o algoritmo η -LMS e para o SR, respectivamente. Em ambas as expressões, aparece o termo $[\Delta\zeta_1(n) + \Delta\zeta_2(n)]$, referente a potência do sinal $[e_{a,2}(n) - e_{a,1}(n)]$, que depende dos passos de adaptação dos filtros componentes, da potência do sinal de entrada, da ordem dos filtros componentes, da presença ou não de ruído, dentre outros fatores. Para facilitar a escolha do passo de adaptação da combinação (μ_η ou $\mu_{\eta,s}$), uma alternativa é usar algum tipo de normalização baseada na potência do sinal $[e_{a,2}(n) - e_{a,1}(n)]$, que pode ser obtida através do sinal $[y_1(n) - y_2(n)]$, conforme mostrado em (3.15). Entretanto, $[y_1(n) - y_2(n)]$ é um sinal que tem uma grande variação a cada instante, o que inviabiliza sua utilização direta num esquema de normalização [Bershad, 1986]. A fim de amenizar a variação de $[y_1(n) - y_2(n)]$, pode-se utilizar uma estimativa de seu valor ao quadrado, usando uma janela exponencial, ou seja,

$$p(n) = \nu p(n-1) + (1-\nu)[y_1(n) - y_2(n)]^2, \quad (3.71)$$

em que $0 \ll \nu < 1$ é um fator de esquecimento. Uma idéia semelhante foi usada na adaptação do parâmetro $\alpha(n)$ da combinação convexa em [Azpicueta-Ruiz et al., 2008] e é revisitada na Seção 1.2.1 (Eq. (1.27)). Com isso, $p(n)$ é capaz de fornecer uma estimativa da potência de $[y_1(n) - y_2(n)]$ com uma variação instantânea muito mais lenta.

A seguir, é considerada a utilização de $p(n)$ para a normalização do algoritmo η -LMS e do algoritmo SR.

Algoritmo *power normalized*

O intervalo para a escolha do passo μ_η do algoritmo η -LMS, obtido em (3.24), depende da potência do sinal $[y_1(n) - y_2(n)]$, dada por $[\Delta\zeta_1(n) + \Delta\zeta_2(n)]$. Assim, uma forma de tornar a escolha do passo independente da potência de $[y_1(n) - y_2(n)]$ é usar uma normalização com base na estimativa da potência, considerando o passo de adaptação variável

$$\mu_\eta(n) = \frac{\tilde{\mu}_\eta}{\epsilon + p(n)}. \quad (3.72)$$

Com isso, a adaptação de $\eta(n)$ se torna

$$\eta(n+1) = \eta(n) + \frac{\tilde{\mu}_\eta}{\epsilon + p(n)} e_g(n) [y_1(n) - y_2(n)], \quad (3.73)$$

sendo ϵ uma constante positiva de pequeno valor. O algoritmo (3.73) é chamado de *power normalized* (PN).

A análise do algoritmo PN envolve o cálculo dos momentos de primeira e segunda ordens para a obtenção de um intervalo para o passo $\tilde{\mu}_\eta$ e uma expressão para a variância $\sigma_\eta^2(n)$ a fim de utilizar (3.7) para estimar o EMSE da combinação. A diferença entre o algoritmo PN e o algoritmo η -LMS, analisado na Seção 3.2 é que o PN usa um passo de adaptação variável. A variação do passo $\mu_\eta(n)$, no entanto, é pequena, pois o sinal $p(n)$ varia lentamente devido à utilização da janela exponencial com fator de esquecimento $0 \ll \nu < 1$. Assim, considerando que a variação de $\mu_\eta(n)$ seja mais lenta que a dos sinais $e_{a,i}(n)$, $i = 1, 2$, assume-se a seguinte hipótese:

T6. $\mu_\eta(n)$ é independente de $e_{a,i}(n)$ $i = 1, 2$ em todo instante de tempo.

Com a hipótese T6, a análise do algoritmo PN é semelhante à análise do algoritmo η -LMS. A diferença entre as duas análises é que, enquanto para o algoritmo η -LMS aparecem termos envolvendo o passo constante μ_η^k , $k = 1, 2$, para o PN aparecerão termos envolvendo $E\{\mu_\eta^k(n)\}$, $k = 1, 2$, já que $\mu_\eta(n)$ agora é variável. Assim a análise do PN pode ser feita a partir de uma extensão direta dos resultados da Seção 3.2. Para isso, é necessário calcular os momentos de primeira e segunda ordens de $p(n)$.

Calculando a esperança de ambos os lados de (3.71), obtém-se

$$E \{p(n)\} = \nu E \{p(n-1)\} + (1-\nu) E \{[y_1(n) - y_2(n)]^2\}. \quad (3.74)$$

Como $p(n)$ varia lentamente, devido ao fator de esquecimento $\nu \approx 1$,

$$E \{p^k(n)\} \approx E \{p^k(n-1)\}, \quad k = 1, 2. \quad (3.75)$$

Usando a aproximação (3.75), (3.74) pode ser reescrita como

$$E \{p(n)\} = E \{[y_1(n) - y_2(n)]^2\}. \quad (3.76)$$

Como o termo $[y_1(n) - y_2(n)]$ é equivalente a $[e_{a,2}(n) - e_{a,1}(n)]$, chega-se a

$$E \{p(n)\} = \Delta\zeta_1(n) + \Delta\zeta_2(n). \quad (3.77)$$

O momento de segunda ordem é obtido elevando (3.71) ao quadrado e calculando a esperança de ambos os lados, obtendo-se

$$\begin{aligned} E \{p^2(n)\} &= \nu^2 E \{p^2(n-1)\} + (1-\nu)^2 E \{[y_1(n) - y_2(n)]^4\} \\ &\quad + 2\nu(1-\nu) E \{p(n-1)[y_1(n) - y_2(n)]^2\}. \end{aligned} \quad (3.78)$$

Usando (3.75) e considerando que $p(n)$ varia lentamente em comparação com $[y_1(n) - y_2(n)]^2$, o último termo do lado direito de (3.78) pode ser aproximado por

$$2\nu(1-\nu) E \{p(n-1)[y_1(n) - y_2(n)]^2\} \approx 2\nu(1-\nu) E \{p(n)\} E \{[y_1(n) - y_2(n)]^2\}. \quad (3.79)$$

Substituindo (3.79) em (3.78) e considerando que $(1-\nu)^2$ é um termo muito pequeno já que $\nu \approx 1$, chega-se a

$$(1+\nu) E \{p^2(n)\} \approx 2\nu E \{p(n)\} E \{[y_1(n) - y_2(n)]^2\}. \quad (3.80)$$

Considerando que $2\nu \approx 1 + \nu$ e substituindo (3.77) em (3.80), obtém-se

$$E \{p^2(n)\} \approx [\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2. \quad (3.81)$$

O intervalo do passo de adaptação $\tilde{\mu}_\eta$ pode ser obtido através de uma expressão equivalente a (3.23). Fazendo $\mu_\eta = \mu_\eta(n)$, obtém-se

$$|E \{1 - \mu_\eta(n) [e_{a,2}(n) - e_{a,1}(n)]^2\}| = |1 - E \{\mu_\eta(n)\} [\Delta\zeta_1(n) + \Delta\zeta_2(n)]| < 1, \quad (3.82)$$

ou seja,

$$0 < E \{ \mu_\eta(n) \} < \frac{2}{\Delta\zeta_1(n) + \Delta\zeta_2(n)}. \quad (3.83)$$

A esperança de $\mu_\eta(n)$ pode ser aproximada por [Sayed, 2003, pág. 304]

$$E \{ \mu_\eta(n) \} \approx \frac{\tilde{\mu}_\eta}{\epsilon + E \{ p(n) \}}. \quad (3.84)$$

Em geral, essa aproximação não é muito realista. No entanto, nesse caso é possível verificar através dos resultados de simulação que é uma aproximação razoável. Substituindo (3.77) em (3.84), obtém-se

$$E \{ \mu_\eta(n) \} \approx \frac{\tilde{\mu}_\eta}{\epsilon + \Delta\zeta_1(n) + \Delta\zeta_2(n)}. \quad (3.85)$$

Substituindo (3.85) em (3.83) e considerando que $\epsilon \ll \Delta\zeta_1(n) + \Delta\zeta_2(n)$, chega-se ao intervalo para o passo $\tilde{\mu}_\eta$ dado por

$$0 < \tilde{\mu}_\eta < 2. \quad (3.86)$$

Portanto, a escolha do passo de adaptação $\tilde{\mu}_\eta$ do algoritmo PN é facilitada, já que não depende de parâmetros do cenário de filtragem.

O comportamento do EMSE da combinação obtido adaptando o parâmetro de mistura com o PN também pode ser estimado através de (3.7). Para isso é necessário obter uma expressão para $\sigma_\eta^2(n)$ para esse algoritmo. No entanto, essa expressão pode ser obtida a partir da extensão do resultado obtido para o algoritmo η -LMS. Assumindo T6, uma expressão equivalente a (3.38) para o algoritmo PN é dada por

$$\begin{aligned} \sigma_\eta^2(n+1) \approx & \left\{ 1 + 3E \{ \mu_\eta^2(n) \} [\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2 \right. \\ & \left. - 2E \{ \mu_\eta(n) \} [\Delta\zeta_1(n) + \Delta\zeta_2(n)] \right\} \sigma_\eta^2(n) \\ & + E \{ \mu_\eta^2(n) \} \sigma_r^2 [\Delta\zeta_1(n) + \Delta\zeta_2(n)] \\ & + E \{ \mu_\eta^2(n) \} [\zeta_1(n)\zeta_2(n) - \zeta_{12}^2(n)]. \end{aligned} \quad (3.87)$$

O resultado para o termo $E \{ \mu_\eta(n) \}$ foi obtido em (3.85) e o momento de segunda ordem de $\mu_\eta(n)$, de maneira análoga, pode ser aproximado por

$$E \{ \mu_\eta^2(n) \} \approx \frac{\tilde{\mu}_\eta^2}{\epsilon^2 + 2\epsilon E \{ p(n) \} + E \{ p^2(n) \}}. \quad (3.88)$$

Substituindo (3.77) e (3.81) em (3.88), obtém-se

$$\begin{aligned} E \{ \mu_\eta^2(n) \} &\approx \frac{\tilde{\mu}_\eta^2}{\epsilon^2 + 2\epsilon[\Delta\zeta_1(n) + \Delta\zeta_2(n)] + [\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2} \\ &= \left[\frac{\tilde{\mu}_\eta}{\epsilon + \Delta\zeta_1(n) + \Delta\zeta_2(n)} \right]^2 \\ &= [E \{ \mu_\eta(n) \}]^2. \end{aligned} \quad (3.89)$$

Denotando

$$\bar{\mu}_\eta(n) \triangleq E \{ \mu_\eta(n) \} \quad (3.90)$$

e substituindo (3.85) e (3.89) em (3.87), obtém-se

$$\begin{aligned} \sigma_\eta^2(n+1) &\approx \left\{ 1 + 3\bar{\mu}_\eta^2(n)[\Delta\zeta_1(n) + \Delta\zeta_2(n)]^2 \right. \\ &\quad \left. - 2\bar{\mu}_\eta(n)[\Delta\zeta_1(n) + \Delta\zeta_2(n)] \right\} \sigma_\eta^2(n) \\ &\quad + \bar{\mu}_\eta^2(n) \sigma_r^2[\Delta\zeta_1(n) + \Delta\zeta_2(n)] \\ &\quad + \bar{\mu}_\eta^2(n) [\zeta_1(n)\zeta_2(n) - \zeta_{12}^2(n)]. \end{aligned} \quad (3.91)$$

Substituindo (3.91) em (3.7) e usando os resultados para $\zeta_i(n)$, $i = 1, 2$ e $\zeta_{12}(n)$ mostrados na Seção 3.3, é possível obter o EMSE da combinação adaptada com o algoritmo PN em qualquer instante de tempo. Uma vantagem desse algoritmo em relação ao η -LMS é que ele não necessita da restrição de $\eta(n) < 1$. Isso ocorre devido à normalização, que faz com que a variação de $\eta(n)$ obtida com a adaptação através do PN seja reduzida, mesmo no início das iterações.

Algoritmo *normalized signed regressor*

De maneira análoga, o intervalo para escolha do passo do algoritmo (3.47) depende raiz quadrada da potência de $[y_1(n) - y_2(n)]$. Assim, uma forma de tornar a escolha do passo de adaptação independente da potência é usar o passo variável

$$\mu_{\eta,s}(n) = \frac{\tilde{\mu}_{\eta,s}}{\epsilon + \sqrt{p(n)}}, \quad (3.92)$$

sendo $\eta(n)$ adaptado através de

$$\eta(n+1) = \eta(n) + \frac{\tilde{\mu}_{\eta,s}}{\epsilon + \sqrt{p(n)}} e_g(n) \text{sign}(y_1(n) - y_2(n)), \quad (3.93)$$

onde ϵ é uma constante positiva de pequeno valor. O algoritmo (3.93) é chamado de *normalized signed regressor* (NSR).

Assumindo uma hipótese semelhante a T6 para $\mu_{\eta,s}(n)$, a análise de (3.93) pode ser feita a

partir da extensão dos resultados obtidos para o SR, na Seção 3.5.1. Nesse caso, são necessários os momentos de primeira e segunda ordens de $\sqrt{p(n)}$.

O momento de segunda ordem de $\sqrt{p(n)}$ é dado pela Equação (3.77), já que $E\{\sqrt{p^2(n)}\} = E\{p(n)\}$. O cálculo do momento de primeira ordem, dado por $E\{\sqrt{p(n)}\}$ requer uma aproximação para a função distribuição de probabilidade de $\sqrt{p(n)}$. Entretanto, para valores de ν suficientemente próximos de 1, o momento de primeira ordem pode ser aproximado por

$$E\{\sqrt{p(n)}\} \approx \sqrt{E\{p(n)\}} = \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}, \quad (3.94)$$

conforme mostrado na Figura 3.8. Nessa figura, são comparados os resultados experimentais de $E\{\sqrt{p(n)}\}$ com os resultados obtidos usando a aproximação $\sqrt{E\{p(n)\}}$. Para a obtenção dos resultados, é utilizado um sinal $[y_1(n) - y_2(n)]$ gaussiano, de média nula e variância 2×10^{-3} . Pode-se notar que para valores de ν suficientemente próximos de 1, a aproximação tem um valor muito próximo do resultado experimental. Uma escolha razoável é fazer $\nu = 0,99$, obtendo uma boa aproximação e um bom comportamento da combinação, conforme será mostrado nos resultados de simulação. O resultado é semelhante para outros valores de variância de $[y_1(n) - y_2(n)]$.

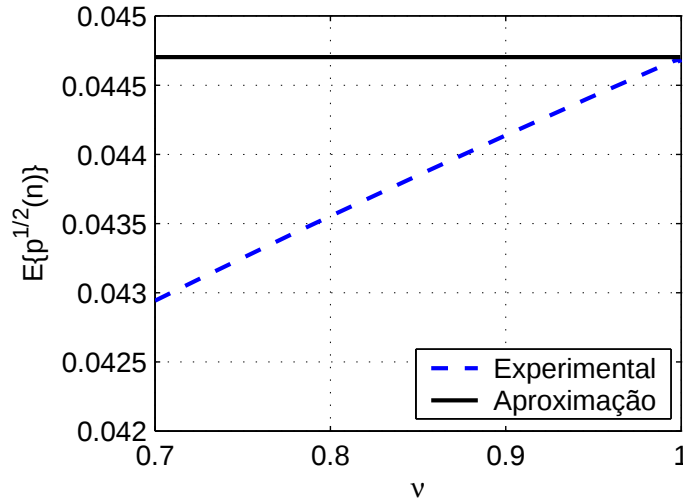


Figura 3.8: Comparação entre o valor experimental de $E\{\sqrt{p(n)}\}$ e o valor obtido com a aproximação $\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}$ para diferentes valores de ν , com o sinal $[y_1(n) - y_2(n)]$ gaussiano com variância $\Delta\zeta_1(n) + \Delta\zeta_2(n) = 2 \times 10^{-3}$.

O intervalo do passo de adaptação pode ser obtido usando o resultado de (3.60), obtendo-se

$$|E\{[1 - \mu_{\eta,s}(n)]e_{a,2}(n) - e_{a,1}(n)|\}| = \left|1 - E\{\mu_{\eta,s}(n)\} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}\right| < 1, \quad (3.95)$$

ou seja,

$$0 < E\{\mu_{\eta,s}(n)\} < \frac{\sqrt{2\pi}}{\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}}. \quad (3.96)$$

Assim como no caso do PN, a esperança de $\mu_{\eta,s}(n)$ pode ser aproximada por

$$E \{ \mu_{\eta,s}(n) \} \approx \frac{\tilde{\mu}_{\eta,s}}{\epsilon + E \left\{ \sqrt{p(n)} \right\}}. \quad (3.97)$$

Substituindo (3.94) em (3.97), obtém-se

$$E \{ \mu_{\eta,s}(n) \} \approx \frac{\tilde{\mu}_{\eta,s}}{\epsilon + \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}}. \quad (3.98)$$

Substituindo (3.98) em (3.96) e considerando que $\epsilon \ll \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}$ chega-se a

$$0 < \tilde{\mu}_{\eta,s} < \sqrt{2\pi}. \quad (3.99)$$

Assim como no PN, a escolha do passo de adaptação do algoritmo NSR também é facilitada, já que não depende de parâmetros do cenário de filtragem.

Para a análise do EMSE da combinação, uma expressão de $\sigma_\eta^2(n)$ pode ser obtida usando o resultado de (3.70), ou seja,

$$\begin{aligned} \sigma_\eta^2(n+1) \approx & \left[1 + E \{ \mu_{\eta,s}^2(n) \} [\Delta\zeta_1(n) + \Delta\zeta_2(n)] \right. \\ & - 2E \{ \mu_{\eta,s}(n) \} \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \left. \right] \sigma_\eta^2(n) \\ & - E \{ \mu_{\eta,s}^2(n) \} \frac{\Delta\zeta_2^2(n)}{\Delta\zeta_1(n) + \Delta\zeta_2(n)} + E \{ \mu_{\eta,s}^2(n) \} (\zeta_2(n) + \sigma_r^2). \end{aligned} \quad (3.100)$$

O momento de primeira ordem de $\mu_{\eta,s}(n)$ foi obtido em (3.98) e o momento de segunda ordem pode ser obtido usando uma aproximação semelhante, ou seja,

$$E \{ \mu_{\eta,s}^2(n) \} \approx \frac{\tilde{\mu}_{\eta,s}^2}{\epsilon^2 + 2\epsilon E \left\{ \sqrt{p(n)} \right\} + E \{ p(n) \}}. \quad (3.101)$$

Substituindo (3.77) e (3.94) em (3.101), obtém-se

$$\begin{aligned} E \{ \mu_{\eta,s}^2(n) \} & \approx \frac{\tilde{\mu}_{\eta,s}^2}{\epsilon^2 + 2\epsilon \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} + \Delta\zeta_1(n) + \Delta\zeta_2(n)} \\ & = \left[\frac{\tilde{\mu}_{\eta,s}}{\epsilon + \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}} \right]^2 \\ & = [E \{ \mu_{\eta,s}(n) \}]^2. \end{aligned} \quad (3.102)$$

Denotando

$$\bar{\mu}_{\eta,s}(n) \triangleq E \{ \mu_{\eta,s}(n) \} \quad (3.103)$$

e usando (3.98) e (3.102) em (3.100), chega-se a

$$\begin{aligned} \sigma_\eta^2(n+1) \approx & \left[1 + \bar{\mu}_{\eta,s}^2(n) [\Delta\zeta_1(n) + \Delta\zeta_2(n)] \right. \\ & \left. - 2\bar{\mu}_{\eta,s}(n) \sqrt{\frac{2}{\pi}} \sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)} \right] \sigma_\eta^2(n) \\ & - \bar{\mu}_{\eta,s}^2(n) \frac{\Delta\zeta_2^2(n)}{\Delta\zeta_1(n) + \Delta\zeta_2(n)} + \bar{\mu}_{\eta,s}^2(n) (\zeta_2(n) + \sigma_r^2). \end{aligned} \quad (3.104)$$

3.5.3 Resumo dos algoritmos propostos

A partir do algoritmo η -LMS, obtido através do uso do gradiente estocástico, foram propostas três formas de adaptação alternativas. A primeira delas é baseada numa normalização instantânea parcial do sinal $[y_1(n) - y_2(n)]$, originando o algoritmo *signed regressor*. Em seguida, com base no algoritmo η -LMS e no SR, foram propostos dois algoritmos normalizados, utilizando uma janela exponencial, fazendo com que os intervalos de escolha de seus passos de adaptação não dependam de parâmetros do cenário de filtragem. Com essa normalização, a partir do algoritmo η -LMS, obteve-se o *power normalized* e a partir do SR obteve-se o *normalized signed regressor*. Na Tabela 3.2, é mostrado um resumo dos algoritmos com as equações de atualização de $\eta(n)$ e os intervalos para escolha de seus passos de adaptação, sendo $p(n)$ calculado por (3.71).

Tabela 3.2: Resumo dos algoritmos para adaptação de $\eta(n)$ e respectivos intervalos para escolha do passo de adaptação

Alg.	$\eta(n+1)$	Passo de adaptação
η -LMS	$\eta(n) + \mu_\eta e_g(n) [y_1(n) - y_2(n)]$ Limitar $\eta(n) \leq 1$	$0 < \mu_\eta < \frac{2}{\Delta\zeta_1(n) + \Delta\zeta_2(n)}$
SR	$\eta(n) + \mu_{\eta,s} e_g(n) \text{sign}(y_1(n) - y_2(n))$	$0 < \mu_{\eta,s} < \frac{\sqrt{2\pi}}{\sqrt{\Delta\zeta_1(n) + \Delta\zeta_2(n)}}$
PN	$\eta(n) + \frac{\tilde{\mu}_\eta}{\epsilon + p(n)} e_g(n) [y_1(n) - y_2(n)]$	$0 < \tilde{\mu}_\eta < 2$
NSR	$\eta(n) + \frac{\tilde{\mu}_{\eta,s}}{\epsilon + \sqrt{p(n)}} e_g(n) \text{sign}(y_1(n) - y_2(n))$	$0 < \tilde{\mu}_{\eta,s} < \sqrt{2\pi}$

3.5.4 Resultados de simulação

O algoritmo *signed regressor*, através do uso da função $\text{sign}(y_1(n) - y_2(n))$ permite que se obtenha uma maior velocidade na adaptação de $\eta(n)$ durante o chaveamento entre os filtros componentes, com um passo suficientemente pequeno para que o desempenho da combinação

nos instantes iniciais não seja tão prejudicado. Na Figura 3.9, são mostrados os resultados para a combinação de dois filtros LMS com os mesmos parâmetros usados na Figura 3.6, mas usando o algoritmo *signed regressor*, com $\mu_{\eta,s} = 0,025$. Pode-se verificar que a combinação apresenta um desempenho próximo ao universal tanto nos instantes iniciais de convergência quanto nos instantes de chaveamento entre os filtros componentes. Esse comportamento é causado pelo uso da função $\text{sign}(y_1(n) - y_2(n))$, que faz uma espécie de normalização do sinal $[y_1(n) - y_2(n)]$ permitindo um ajuste único no passo de adaptação tanto nos instantes iniciais, quando $y_1(n) - y_2(n)$ é um valor grande, quanto nos instantes de chaveamento e regime, quando $[y_1(n) - y_2(n)]$ é um valor pequeno.

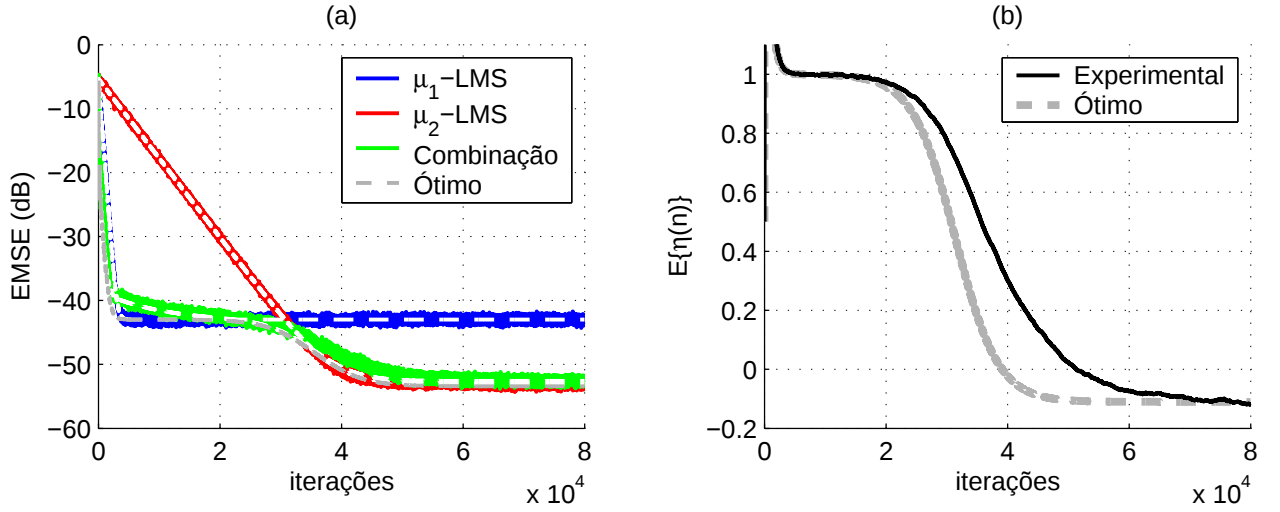


Figura 3.9: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo SR. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\mu_{\eta,s} = 0,025$. Valores teóricos tracejados.

Nas Figuras 3.10 e 3.11, são mostrados os resultados da combinação de dois filtros LMS usando os algoritmos PN e NSR, respectivamente. As condições das simulações são as mesmas da Figura 3.9. Para o algoritmo PN, $\tilde{\mu}_\eta = 0,003$, $\nu = 0,99$ e $\epsilon = 1 \times 10^{-3}$ e para o NSR, $\tilde{\mu}_{\eta,s} = 0,003$, $\nu = 0,99$ e $\epsilon = 0,1$. Pode-se notar que o desempenho da combinação em ambos os casos é próximo ao universal.

Na Figura 3.12, são reunidos os resultados do parâmetro de mistura obtidos nas Figuras 3.6, 3.9, 3.10 e 3.11. Pode-se notar que o algoritmo SR, o PN e o NSR têm um desempenho superior ao algoritmo η -LMS, principalmente nos instantes iniciais, onde são capazes de ficar mais próximos ao valor do parâmetro de mistura ótimo. É possível verificar ainda que o algoritmo NSR têm um desempenho ligeiramente superior aos algoritmos PN e do sinal durante o chaveamento entre os filtros componentes já que, como pode ser observado na figura, o algoritmo NSR é capaz de atingir o valor do parâmetro de mistura ótimo alguns instantes antes dos algoritmos PN e do sinal. Nas simulações realizadas, foi possível verificar que a

escolha dos passos de adaptação $\tilde{\mu}_\eta$ e $\tilde{\mu}_{\eta,s}$ e da constante ϵ para os algoritmos PN e NSR não é um problema. Os algoritmos conseguem manter um comportamento adequado para uma faixa de valores de $\tilde{\mu}_\eta$, $\tilde{\mu}_{\eta,s}$ e ϵ e apresentam um desempenho melhor que o η -LMS diante de variações no cenário de filtragem.

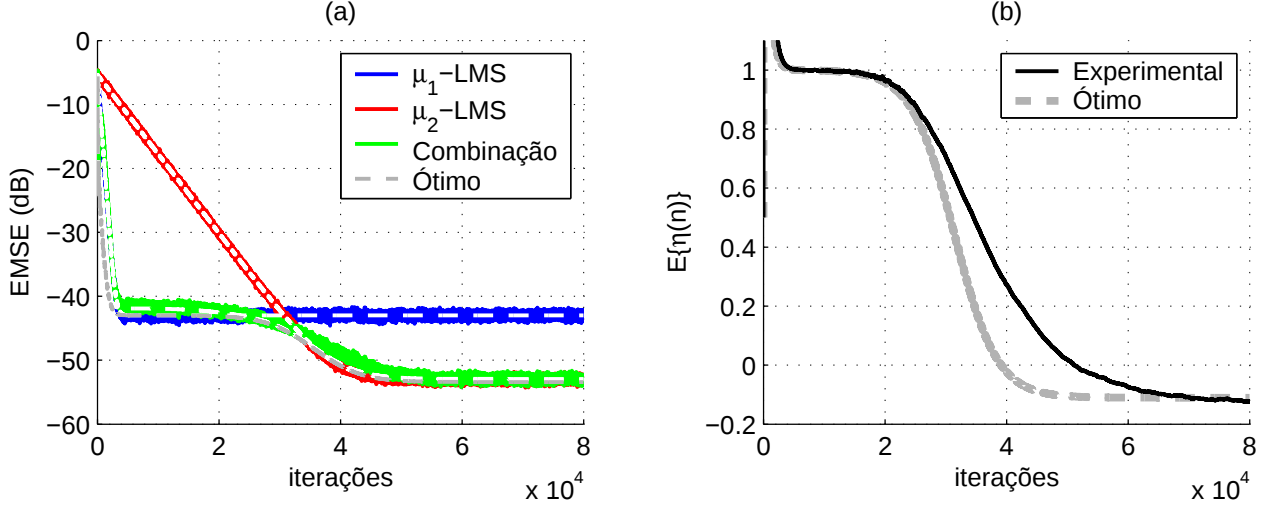


Figura 3.10: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo PN. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\tilde{\mu}_\eta = 0,003$, $\nu = 0,99$ e $\epsilon = 1 \times 10^{-3}$. Valores teóricos tracejados.

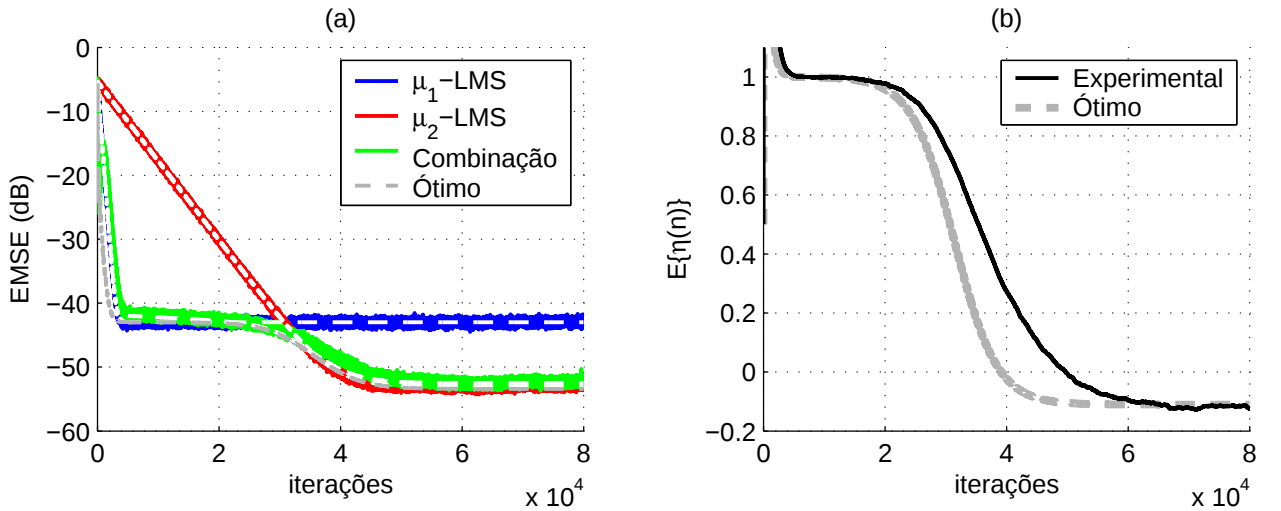


Figura 3.11: Combinação afim de μ_1 -LMS e μ_2 -LMS no caso estacionário. Adaptação com o algoritmo NSR. $M = 7$, $\sigma_v^2 = 0,01$, $\mu_1 = 0,01$, $\mu_2 = \delta_{LL}\mu_1$ e $\delta_{LL} = 0,1$. Entrada branca e solução ótima $\mathbf{w}_{o,1}$ com potência $1/M$. Média de 500 realizações com $\tilde{\mu}_{\eta,s} = 0,003$, $\nu = 0,99$ e $\epsilon = 0,1$. Valores teóricos tracejados.

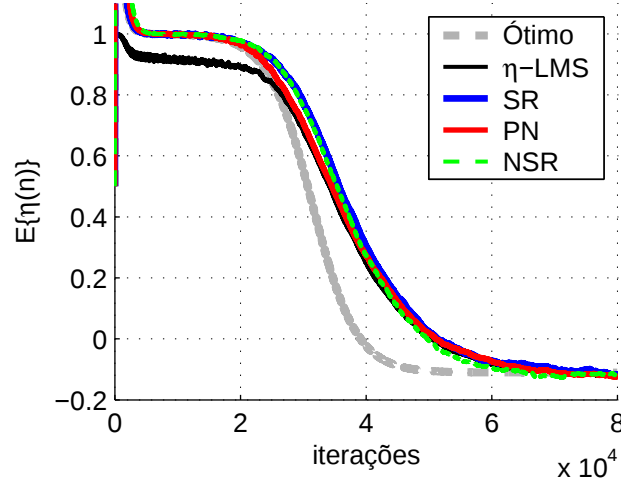


Figura 3.12: Comportamento do parâmetro de mistura obtido com os diferentes tipos de adaptação considerados. Comparação entre os resultados mostrados nas Figuras 3.6, 3.9, 3.10 e 3.11.

3.6 Conclusões

Neste capítulo, foi feita uma análise de transitório de um dos algoritmos de adaptação de $\eta(n)$ propostos em [Bershad et al., 2008]. A análise é capaz de mostrar que em alguns casos, a escolha do passo de adaptação μ_η para obter um comportamento da combinação próximo do universal nem sempre é simples. Em algumas situações, a combinação necessita de um passo grande o suficiente para garantir o comportamento próximo ao universal durante o chaveamento entre os filtros componentes, principalmente quando os níveis de EMSE dos filtros componentes são muito distintos. Entretanto, um passo de adaptação com valor muito alto faz com que a variação de $\eta(n)$ seja muito grande no início da convergência, o que faz com que a combinação tenha um comportamento inferior ao universal nos instantes iniciais. Em situações desse tipo, a escolha do passo de adaptação envolve o compromisso entre manter o comportamento da combinação mais próximo do universal nos instantes de chaveamento entre filtros e em regime ou nos instantes iniciais, à medida que são usados passos com valor grande ou pequeno respectivamente.

Com base nesse resultado, foi proposto o algoritmo SR. Esse algoritmo faz uma normalização instantânea parcial do passo de adaptação μ_η , através do módulo de $[y_1(n) - y_2(n)]$. Através da análise de transitório, pode-se notar que essa normalização é capaz de facilitar a escolha do passo de adaptação para manter o compromisso entre o comportamento próximo ao universal nos instantes iniciais e nos instantes de chaveamento entre filtros e em regime. Além disso, como a normalização faz uma espécie de ajuste no passo de adaptação de acordo com a situação dos filtros componentes, esse algoritmo não necessita da restrição de $\eta(n) < 1$ para evitar a divergência nos instantes iniciais. Isso faz com que seja possível obter um valor mais próximo do ótimo no início da convergência, que em geral é um valor um pouco maior

que 1.

Ainda foram propostos outros dois algoritmos, baseados no algoritmo η -LMS e no SR, denominados PN e NSR. Esses dois algoritmos usam um esquema de normalização baseado numa estimativa de $[y_1(n) - y_2(n)]^2$. A vantagem desse tipo de normalização é que o intervalo de escolha do passo de adaptação obtido, ao contrário do algoritmo usado inicialmente e do SR, não depende de parâmetros do cenário de filtragem, o que facilita a escolha do passo e melhora o comportamento da combinação diante de situações com parâmetros variáveis como a relação sinal-ruído, por exemplo. As análises feitas para o algoritmo usado inicialmente e para o SR foram estendidas para o PN e o NSR sendo obtida uma boa concordância entre os valores teóricos e os valores experimentais.

Os três algoritmos, SR, PN e NSR são capazes de fazer com que o comportamento da combinação seja mais próximo do universal nas situações em que o algoritmo η -LMS apresenta problemas e conforme os resultados de simulação obtidos, a diferença entre o valor esperado do parâmetro de mistura é muito pequena entre os três algoritmos.

Capítulo 4

Estratégias para inicialização do CMA

Neste capítulo, considera-se a combinação afim de equalizadores CMA com mesmo passo de adaptação mas com inicializações diferentes. Inicialmente, é feita uma comparação entre o comportamento da combinação afim e o da combinação convexa através de simulações simples nessa aplicação. Com base nesses resultados, é proposto o uso de uma transferência de coeficientes da combinação afim para os filtros componentes com o objetivo de melhorar o comportamento do esquema. Por fim, são apresentados alguns resultados preliminares na aplicação dessa combinação para equalizadores com mais coeficientes.

4.1 Combinações de CMA com diferentes inicializações

O CMA é um algoritmo que tem seu desempenho muito dependente da inicialização dos coeficientes [Johnson Jr. et al., 1998]. Assim, é interessante considerar a combinação de algoritmos CMA com inicializações diferentes para tentar evitar a convergência para um mínimo local. Nesse contexto, a combinação afim pode ser vantajosa em relação à combinação convexa, já que a primeira possibilita uma liberdade maior na variação do parâmetro de mistura.

Para comparar o comportamento das combinações afim e convexa, foi considerada a transmissão de um sinal binário (± 1), através do canal de comunicação [Ding e Li, 2001]

$$H(z) = \frac{1}{1 + 0,6z^{-1}}. \quad (4.1)$$

A escolha de usar um canal IIR (*infinite impulse response*) de primeira ordem, apesar de não ser muito realista, permite o cálculo da solução ótima com precisão e a representação gráfica da trajetória dos coeficientes dos filtros componentes e da combinação num plano, o que facilita a comparação entre o desempenho dos algoritmos e combinações. Nesse caso, existem duas soluções ótimas, dadas por

$$\mathbf{w}_{o,1}^T = [1 \quad 0,6] \quad (4.2)$$

e

$$\mathbf{w}_{o,2}^T = [-1 \quad -0,6] . \quad (4.3)$$

A solução ótima $\mathbf{w}_{o,2}$ recupera o sinal transmitido a menos de um fator multiplicativo de -1 mas é possível contornar o problema utilizando modulação diferencial. Lembrando que a saída de um filtro componente é calculada como

$$y_i(n) = \mathbf{u}^T(n) \mathbf{w}_i(n-1) \quad (4.4)$$

e a saída da combinação é calculada como

$$y(n) = \eta(n)y_1(n) + [1 - \eta(n)]y_2(n), \quad (4.5)$$

é possível obter uma expressão para os vetor de coeficientes equivalentes da combinação, tal que

$$y(n) = \mathbf{u}^T(n) \mathbf{w}_{eq}(n-1), \quad (4.6)$$

calculando

$$\mathbf{w}_{eq}(n) = \eta(n) \mathbf{w}_1(n) + [1 - \eta(n)] \mathbf{w}_2(n), \quad (4.7)$$

o que permite a visualização da trajetória dos coeficientes das combinações.

As simulações foram realizadas considerando uma relação sinal-ruído de 25 dB e fazendo a adaptação da combinação afim através do algoritmo PN. Analogamente, na combinação convexa, foi usado o método normalizado de [Azpicueta-Ruiz et al., 2008]. Além disso, na adaptação do parâmetro de mistura da combinação convexa também foi usado o erro de decisão $e_d(n)$ no lugar de $e(n)$, como feito para a combinação afim. Assim, o parâmetro $\alpha(n)$ da combinação convexa é adaptado através de

$$\alpha(n) = \alpha(n-1) + \frac{\tilde{\mu}_\alpha}{\epsilon + p(n)} \eta(n) [1 - \eta(n)] e_d(n), \quad (4.8)$$

sendo $\tilde{\mu}_\alpha$ o passo de adaptação e $p(n)$ calculado por (3.71).

Na Figura 4.1, são mostrados os resultados obtidos para a trajetória média dos coeficientes, EMSE e parâmetro de mistura das combinações afim e convexa, com os filtros componentes inicializados em $\mathbf{w}_1^T(0) = [0,05 \quad -0,4]$ e $\mathbf{w}_2^T(0) = [1 \quad 0]$, denominada “Situação 1” de inicialização. Na Figura 4.1-(a), também estão representadas as curvas de nível da função custo do CMA para o cenário considerado. Os pontos de mínimo global, obtidos com as soluções ótimas $\mathbf{w}_{o,1}$ e $\mathbf{w}_{o,2}$, estão identificados como MG e os pontos de mínimo local estão identificados como ML. Com essa inicialização, o comportamento das duas combinações foi semelhante nas 10^3 realizações consideradas. O filtro i_1 -CMA convergiu para um mínimo local, com EMSE de aproximadamente -5 dB e o filtro i_2 -CMA convergiu para um mínimo global, obtendo EMSE de aproximadamente -22 dB. Ambas as combinações acompanharam o filtro i_2 -CMA. Em re-

gime, a combinação convexa apresentou um desempenho igual ao do filtro i_2 -CMA enquanto a combinação afim apresentou um desempenho próximo, com cerca de -21 dB de EMSE. Isso ocorre pois a combinação convexa considera o uso do parâmetro de mistura universal $\eta_u(n)$ (Eq. (1.26) pág. 8), que é arredondado para 0 ou 1 à medida que $\eta(n)$ atinge uma certa proximidade de 0 ou 1. Essa diferença não é significativa na comparação, já que é possível definir um parâmetro de mistura universal de forma análoga para a combinação afim. No transitório, pode-se notar que os coeficientes da combinação afim fazem uma trajetória diferente da trajetória feita pelos coeficientes da combinação convexa. Isso faz com que a combinação afim seja capaz de apresentar uma pequena vantagem na velocidade de convergência. Esse comportamento é possível devido à inexistência da restrição do parâmetro de mistura, que adquire valor negativo no início da convergência conforme pode ser constatado na Figura 4.1-(c).

Nas Figuras 4.2 e 4.3, são mostrados os resultados considerando o mesmo cenário, mas inicializando os filtros componentes em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$, denominada “Situação 2” de inicialização. Nesse caso, em 412 das 10^3 realizações, o comportamento das combinações foi semelhante ao obtido com a inicialização na situação anterior, conforme pode ser visto na Figura 4.2. Entretanto, um comportamento interessante ocorreu nas outras 588 realizações conforme mostrado na Figura 4.3. Nessas realizações, ambos os filtros componentes convergiram para um mínimo local, obtendo EMSE em regime de aproximadamente -5 dB. No entanto, nos instantes iniciais da convergência do filtro i_2 -CMA, este fica próximo do mínimo global, chegando a obter um EMSE de cerca de -9 dB durante algumas poucas iterações. Diante desse cenário, a combinação convexa é capaz de acompanhar o melhor filtro, apresentando um desempenho próximo ao universal. A combinação afim, no entanto, apresenta uma grande variação na trajetória média dos coeficientes no início da convergência, fazendo com que fiquem próximos da solução ótima durante um número maior de iterações. Isso possibilita obter um EMSE menor que o EMSE dos filtros componentes, como pode ser visto na Figura 4.3-(b). Em regime, a combinação afim acaba retornando e convergindo para um mínimo local, porém parece fazer um esforço para sair dessa situação, já que $E\{\eta(n)\} \rightarrow -200$.

O comportamento da combinação afim diante de uma situação semelhante à mostrada na Figura 4.3 motiva o estudo de alguma forma para aproveitar o ganho de EMSE nos instantes iniciais com o objetivo de evitar a convergência para um mínimo local. Uma das formas de aproveitar esse comportamento, é utilizar um esquema de transferência de coeficientes da combinação para os filtros componentes, o que pode evitar a convergência dos filtros componentes e da combinação para um mínimo local.

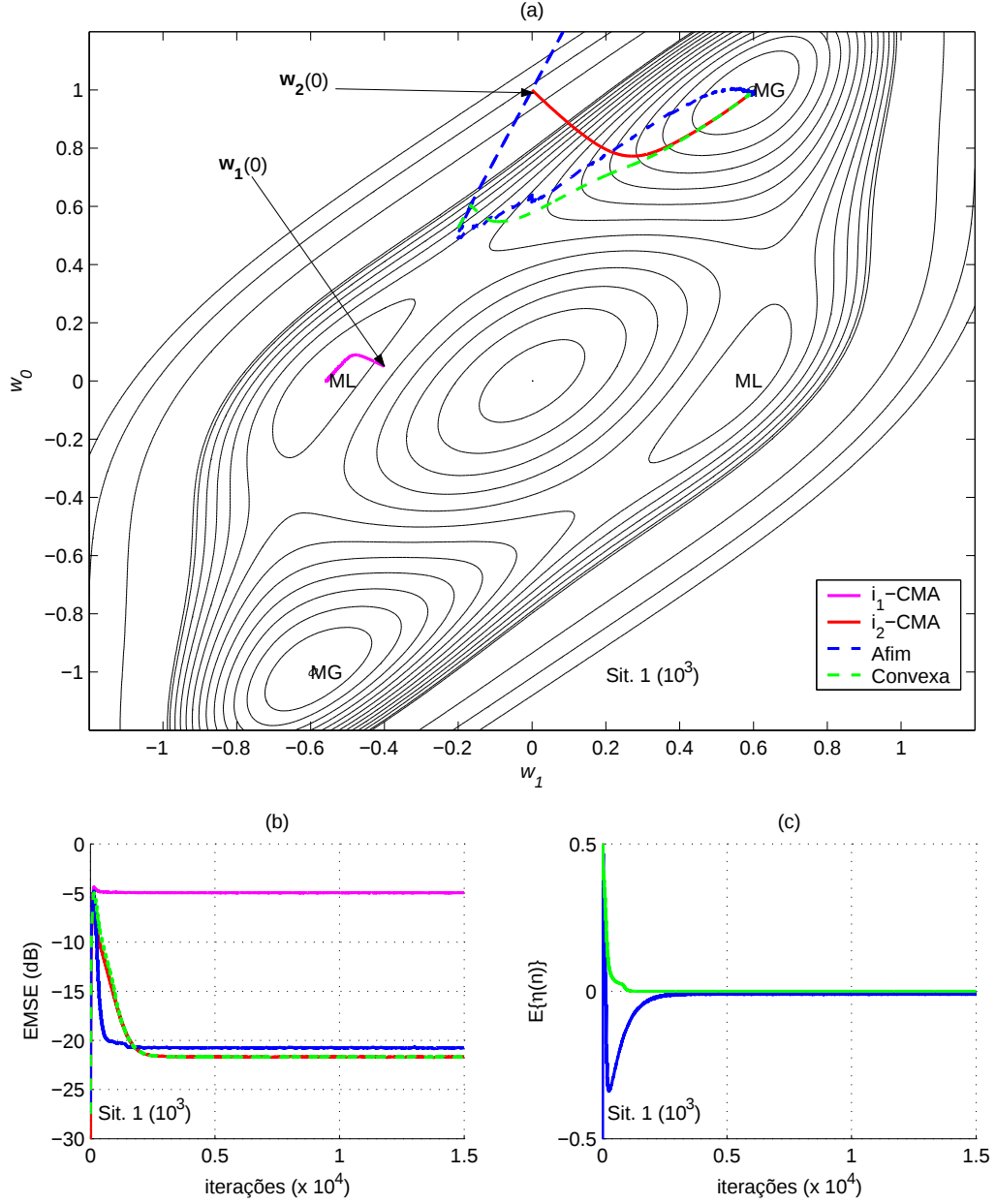


Figura 4.1: Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 0,2$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ -0,4]$ e $\mathbf{w}_2^T(0) = [1 \ 0]$. Canal IIR definido por (4.1). Média de 10^3 realizações.

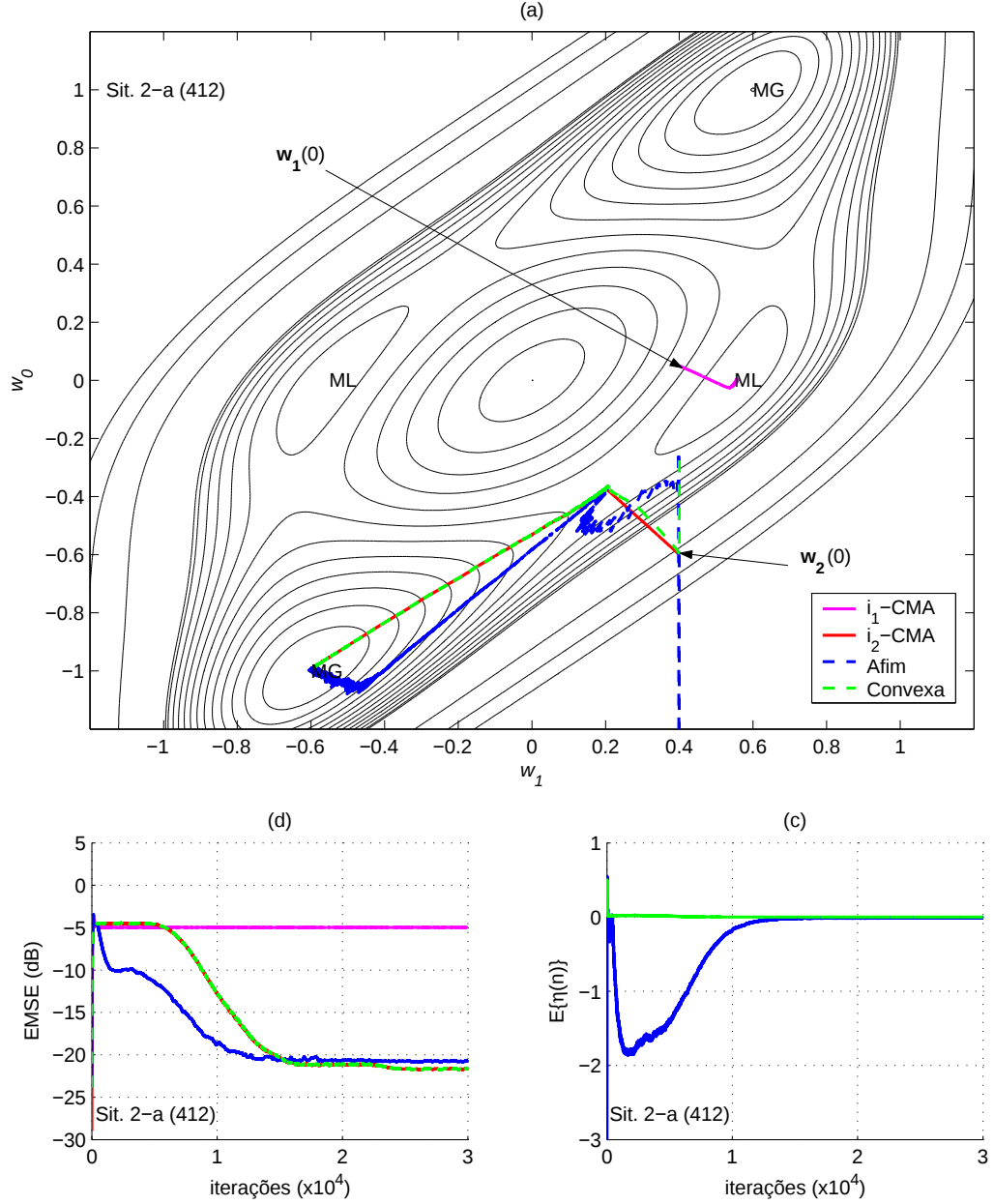


Figura 4.2: Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 412 realizações de um total de 10^3 .

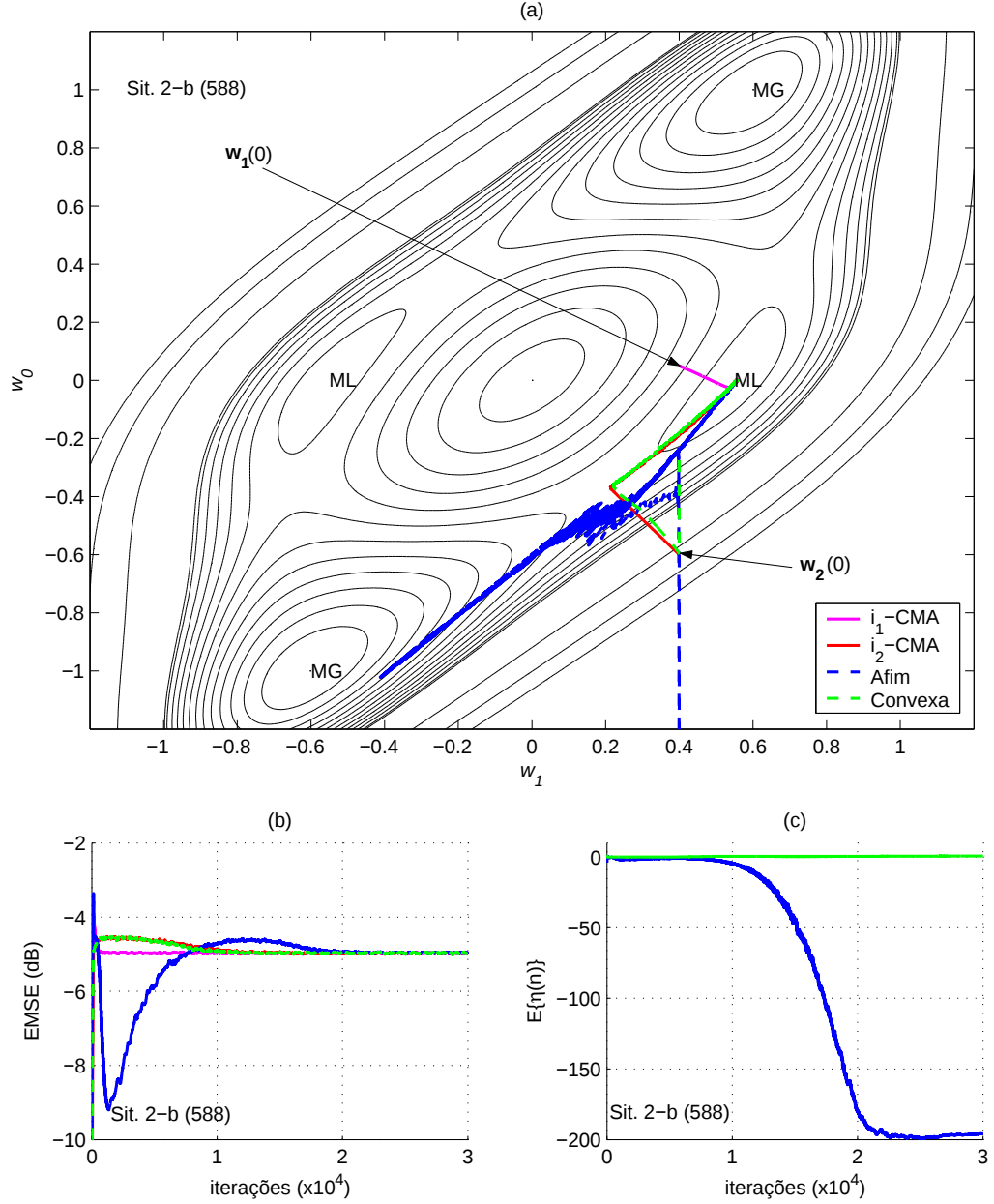


Figura 4.3: Combinações afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) e convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 588 realizações de um total de 10^3 .

4.2 A transferência de coeficientes

Em [Arenas-García et al., 2003], para melhorar o desempenho da combinação de mais de dois algoritmos LMS, foi proposta a adaptação dos coeficientes dos filtros componentes considerando a transferência de uma proporção dos coeficientes da combinação, ou seja,

$$\mathbf{w}_i(n) = \varrho[\mathbf{w}_i(n-1) + \mu_i e(n)\mathbf{u}(n)] + (1 - \varrho)\mathbf{w}_{eq}(n-1), \quad (4.9)$$

sendo $0 < \varrho < 1$ uma constante que determina a proporção da transferência de coeficientes. Essa idéia pode ser usada no caso em questão para aproveitar o comportamento da combinação afim durante os instantes iniciais da convergência.

Na Figura 4.4, são mostrados os resultados obtidos com a combinação afim usando a transferência de coeficientes no mesmo cenário considerado nas Figuras 4.2 e 4.3. Os filtros componentes foram inicializados na Situação 2 e os coeficientes da combinação afim foram transferidos na adaptação dos filtros componentes, segundo (4.9) com $\varrho = 0,7$. Como pode ser observado, nas 10^3 realizações, a transferência de coeficientes fez com que a trajetória dos coeficientes dos filtros componentes mudasse de direção, convergindo para o mínimo global apesar da inicialização desfavorável. Como a transferência de coeficientes é feita para ambos os filtros, o desempenho destes torna-se semelhante e o desempenho da combinação acaba ficando próximo ao dos filtros componentes, com $\eta(n) \approx 1$.

A utilização da transferência de coeficientes com a combinação convexa nesse caso, não oferece nenhuma vantagem, já que a combinação converge diretamente para o mínimo local. Esse resultado pode ser observado na Figura 4.5. Pode-se notar que a transferência de coeficientes altera o comportamento do filtro i_2 -CMA, fazendo com que este convirja para o mínimo local, apresentando um desempenho semelhante ao do filtro i_1 -CMA. Como os filtros componentes ficam com desempenhos semelhantes, a combinação convexa acaba obtendo esse mesmo desempenho, com $\eta(n) \approx 0$.

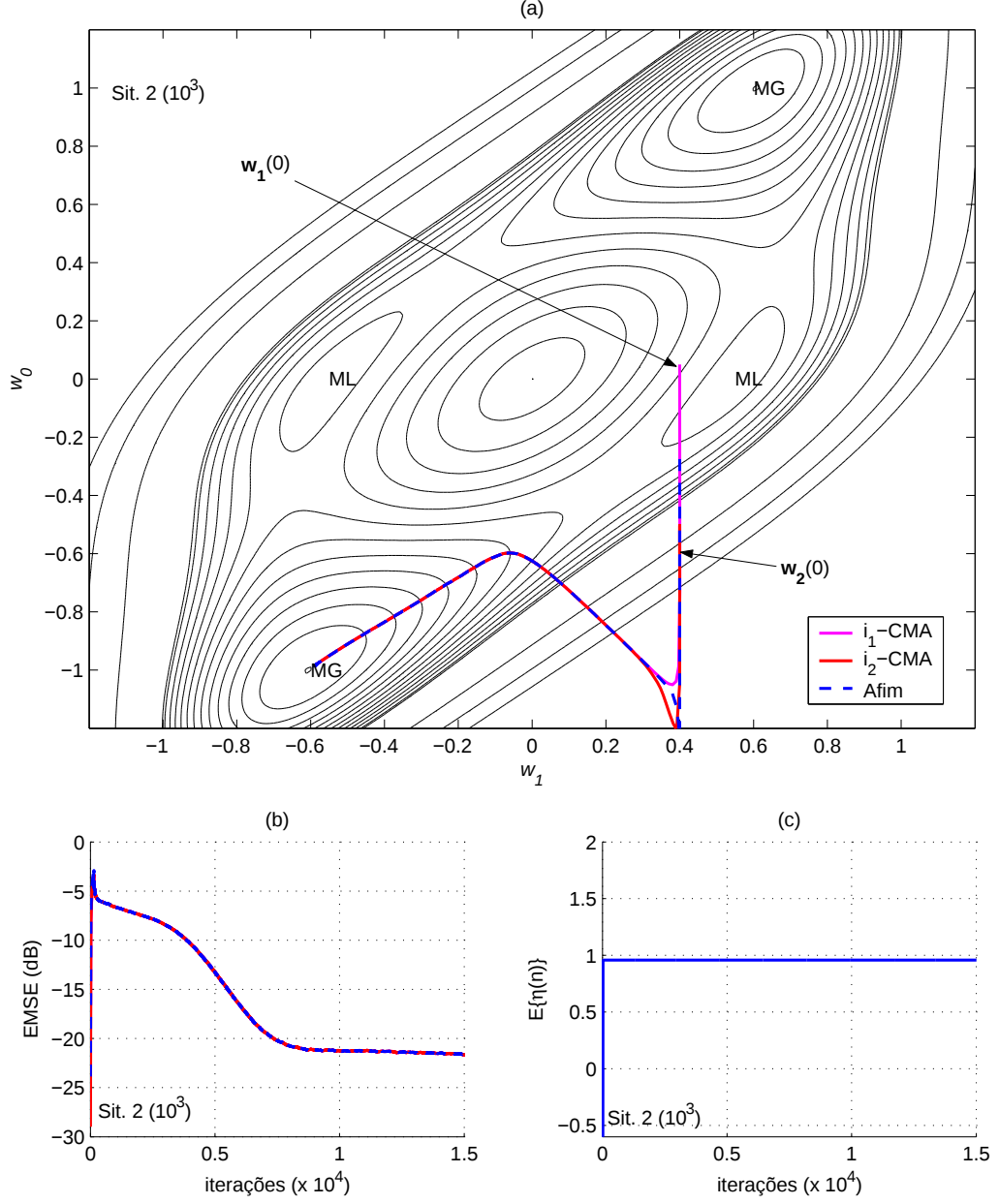


Figura 4.4: Combinação afim (PN com $\tilde{\mu}_\eta = 0,4$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes e usando a transferência de coeficientes dos filtros componentes para a combinação com $\varrho = 0,7$. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 10^3 realizações.

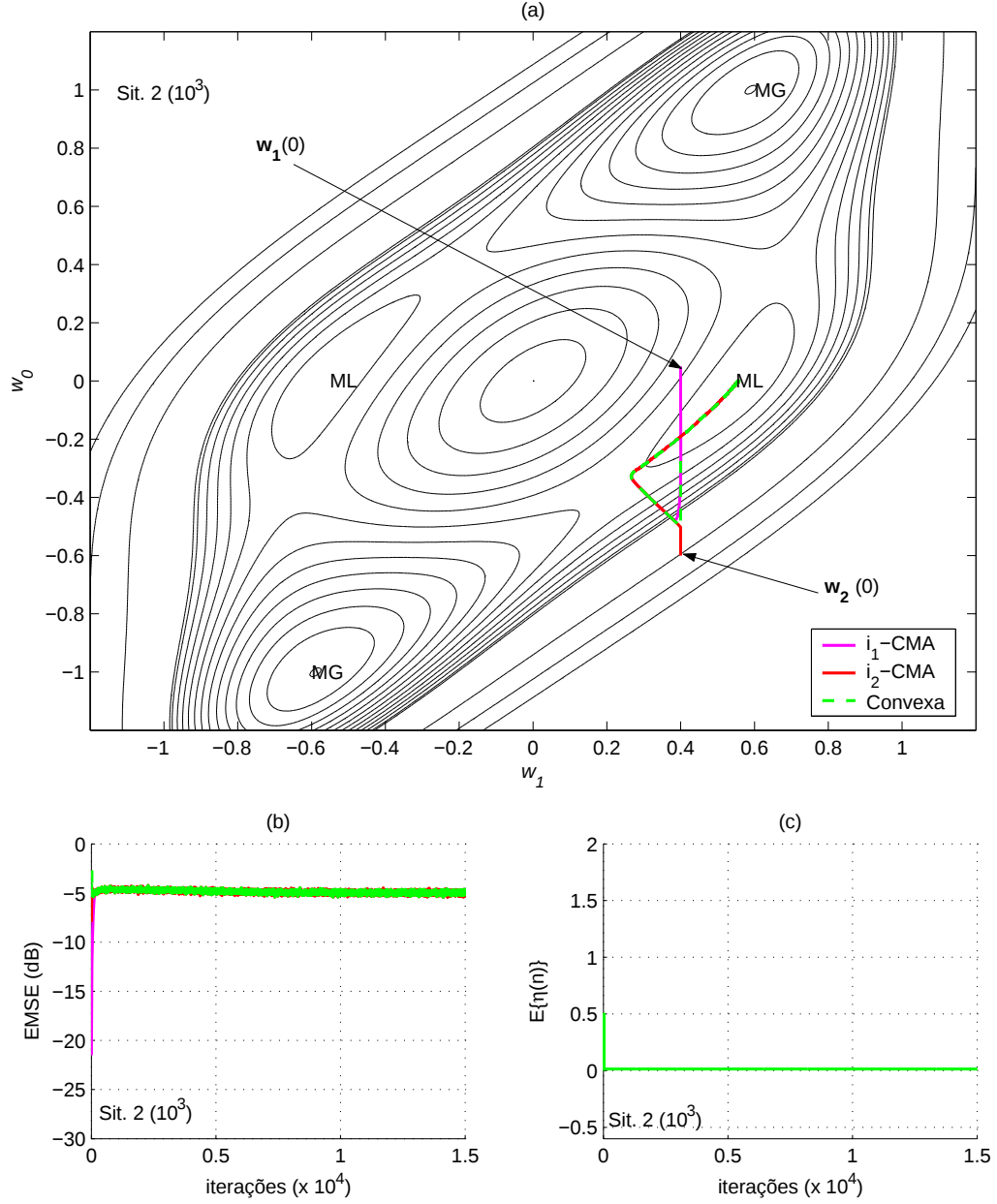


Figura 4.5: Combinação convexa (PN com $\tilde{\mu}_\alpha = 3$, $\alpha^+ = 4$, $\epsilon_C = 0,2$, $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes e usando a transferência de coeficientes dos filtros componentes para a combinação com $\rho = 0,7$. Sinal de entrada binário, $M = 2$, SNR de 25 dB e $\mu_1 = \mu_2 = 0,0016$. Vetores de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0,05 \ 0,4]$ e $\mathbf{w}_2^T(0) = [-0,6 \ 0,4]$. Canal IIR definido por (4.1). Média de 10^3 realizações.

4.3 Comportamento em equalizadores com mais coeficientes

Conforme os resultados apresentados, é possível melhorar o desempenho da combinação de equalizadores CMA usando a combinação afim de filtros com inicializações diferentes. Num caso mais realista, considerando equalizadores com maior número de coeficientes, fica mais difícil visualizar um comportamento semelhante ao obtido quando usados dois coeficientes apenas. Além da dificuldade de representação dos coeficientes dos filtros no caso de maior número de coeficientes, é difícil encontrar uma situação de inicialização que faça com que os filtros componentes apresentem um comportamento semelhante ao mostrado na Figura 4.3, onde um dos filtros componentes passa perto de um mínimo global e acaba convergindo para um mínimo local. Entretanto, através de simulações, verificou-se que o uso da combinação afim no caso de filtros com mais coeficientes e inicializações diferentes pode ser vantajoso para aproveitar o desempenho do filtro componente com a inicialização mais favorável.

Na Figura 4.6, são mostrados os resultados obtidos com a combinação afim usando o canal [Silva et al., 2002]

$$\mathbf{h}_3^T = [0,6364 \quad 0,5733 \quad 0,5160],$$

para filtros componentes com $M = 11$ coeficientes. Foi considerada entrada binária e os filtros componentes foram inicializados com um vetor pino na quarta posição para o filtro i_1 -CMA e na décima primeira posição para o filtro i_2 -CMA. É possível verificar que a inicialização do filtro i_1 -CMA é mais adequada que a do filtro i_2 -CMA e a combinação afim apresenta um desempenho próximo ao do filtro com a inicialização mais adequada. Ainda nesse cenário, foram feitas outras simulações considerando outras inicializações e a combinação, nos casos testados, sempre apresentou comportamento semelhante, acompanhando o filtro componente com inicialização mais adequada. Dentre as inicializações testadas, não foi observado nenhum caso em que a combinação afim fosse capaz de apresentar um desempenho superior ao dos filtros componentes, como mostrado nas simulações com dois coeficientes.

Na Figura 4.7, são mostrados os resultados obtidos usando o canal

$$\mathbf{h}_1^T = [0,1 \quad 0,3 \quad 1,0 \quad -0,1 \quad 0,5 \quad 0,2]$$

e relação sinal-ruído de 25 dB, para filtros componentes com $M = 4$ coeficientes e usando sobreamostragem de fator $S = 2$. Foi considerada entrada binária e os filtros componentes foram inicializados em $\mathbf{w}_1^T(0) = [0 \quad -1 \quad -0,5 \quad 0]$ e $\mathbf{w}_2^T(0) = [0 \quad 0 \quad 1 \quad 0,8]$. Devido à utilização de sobreamostragem nesse caso, os equalizadores tendem a obter o mesmo nível de EMSE independentemente de sua inicialização, variando apenas a velocidade de convergência. Como pode ser notado, o comportamento da combinação é semelhante ao observado na Figura 4.6, mas com os filtros componentes convergindo para um mesmo nível de EMSE.

Nesses casos, também foram feitas simulações para a combinação convexa e verificou-se

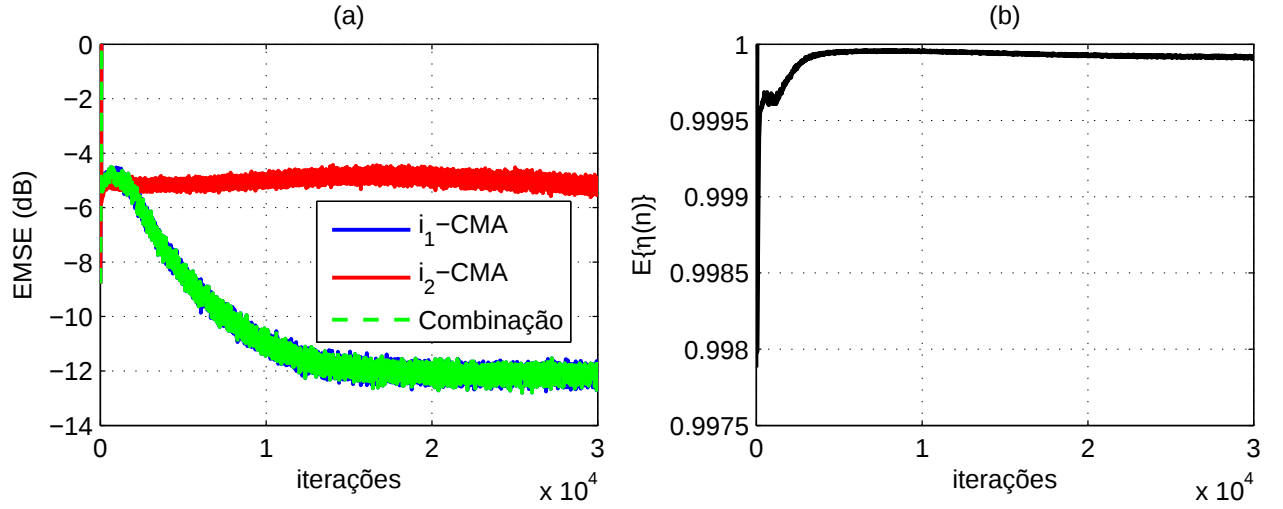


Figura 4.6: Combinação afim (PN com $\tilde{\mu}_\eta = 5 \times 10^{-4}$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 11$ e $\mu_1 = \mu_2 = 0,0025$. Filtro i_1 -CMA inicializado com um vetor pino com um elemento unitário na quarta posição e i_2 -CMA inicializado com um vetor pino com um elemento unitário na décima primeira posição. Canal $\mathbf{h}_3$. Média de 10^3 realizações.

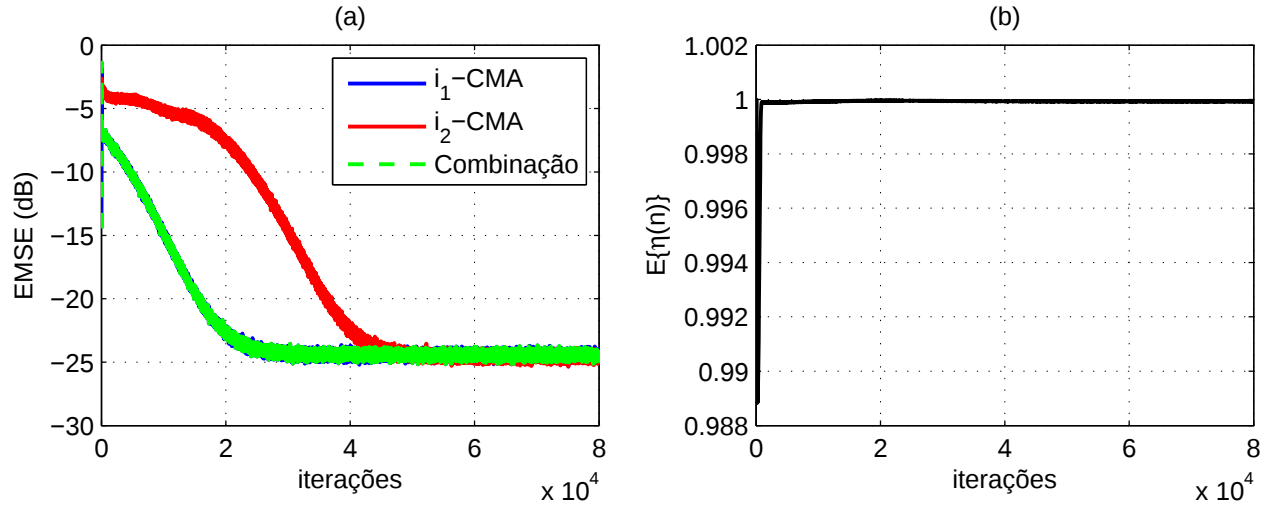


Figura 4.7: Combinação afim (PN com $\tilde{\mu}_\eta = 5 \times 10^{-4}$; $\nu = 0,9$ e $\epsilon = 5 \times 10^{-5}$) de i_1 -CMA e i_2 -CMA com inicializações diferentes. Sinal de entrada binário, $M = 4$ com sobreamostragem de fator $S = 2$ e $\mu_1 = \mu_2 = 0,0016$. Vetor de coeficientes dos filtros inicializados em $\mathbf{w}_1^T(0) = [0 \ -1 \ -0,5 \ 0]$ e $\mathbf{w}_2^T(0) = [0 \ 0 \ 1 \ 0,8]$. Canal $\mathbf{h}_1$ com SNR de 25 dB. Média de 10^3 realizações.

que o comportamento é semelhante, fazendo com que a combinação obtenha um desempenho próximo ao do filtro componente com inicialização mais favorável. Foram feitos também testes com a transferência de coeficientes, tanto para a combinação afim como para a combinação convexa e verificou-se que ela melhora o desempenho do filtro com inicialização menos favorável, fazendo com que ambos os filtros componentes e a combinação obtenham o mesmo desempenho do filtro com melhor inicialização.

4.4 Conclusões

Neste capítulo, foram feitos testes experimentais da combinação afim de dois equalizadores CMA com inicializações diferentes e os resultados foram comparados em relação aos obtidos com a combinação convexa. Considerando um cenário simples, com equalizadores com dois coeficientes, é possível verificar que ambas as combinações podem ser utilizadas com o objetivo de aproveitar o desempenho em regime do filtro componente com a inicialização mais favorável. Entretanto, a combinação afim apresenta uma vantagem devido à maior liberdade oferecida na adaptação do parâmetro de mistura, o que faz com que o vetor de coeficientes apresente uma grande variação no início da convergência, possibilitando uma velocidade de convergência um pouco maior que a obtida com a combinação convexa.

Foi mostrado ainda que a combinação afim, em alguns casos, pode ser usada para evitar a convergência para um mínimo local. A variação inicial do vetor de coeficientes da combinação pode ser aproveitada, usando a transferência de coeficientes, para ajudar os filtros componentes a convergirem na direção de um mínimo global. Entretanto, quando utilizada a transferência de coeficientes, o desempenho dos filtros componentes e o da combinação tendem a ficar muito próximos e nesse caso não ocorre ganho de EMSE na combinação afim. O motivo pelo qual o ganho deixa de ocorrer será tratado num trabalho futuro, podendo estar relacionado com a velocidade de convergência do parâmetro de mistura da combinação, que é extremamente baixa devido à semelhança no desempenho dos filtros componentes. Isso prejudica a eficiência do algoritmo em termos de custo computacional, já que após a convergência, o funcionamento do segundo filtro componente é redundante, pois seu desempenho passa a ser igual ao do primeiro.

Num cenário mais realista, considerando equalizadores mais longos, é possível verificar o funcionamento da combinação no sentido de aproveitar a inicialização mais favorável mas é difícil encontrar alguma situação onde a variação dos coeficientes da combinação afim possa trazer alguma vantagem como no caso de dois coeficientes. Essas características devem ser estudadas com mais profundidade à medida que a influência da inicialização no desempenho dos filtros componentes e os fatores que influenciam a variação da combinação afim são questões ainda em aberto.

Capítulo 5

Conclusões e trabalhos futuros

O projeto de filtros adaptativos envolve a escolha de parâmetros fixos, como o passo de adaptação para algoritmos do tipo LMS ou o fator de esquecimento para algoritmos do tipo RLS. A escolha desses parâmetros tem que levar em conta um compromisso entre velocidade de convergência do filtro e nível de erro em regime. A combinação de filtros surgiu como uma forma de melhorar esse compromisso à medida que se pode combinar um filtro de convergência rápida com um filtro com menor nível de erro em regime. Nesse contexto, a combinação convexa de dois algoritmos LMS de [Arenas-García et al., 2006a] se mostrou uma boa alternativa. Entretanto, através de uma análise para obter o parâmetro de mistura ótimo da combinação, constatou-se que em alguns casos ele pode ser um número negativo, o que não é possível de ser obtido com a combinação convexa. Foi proposta então, em [Bershad et al., 2008] a combinação afim de dois algoritmos LMS, uma generalização da combinação convexa de [Arenas-García et al., 2006a] e que permite que o parâmetro de mistura adquira valores negativos.

Nesta dissertação, a combinação afim foi estendida para os algoritmos NLMS, RLS e para o CMA, considerando entrada branca ou colorida e ambientes estacionários ou não-estacionários. Foi feita uma análise unificada em regime e do transitório da combinação afim dos algoritmos considerados. A partir dessa análise, verificou-se que a combinação afim oferece no mínimo o mesmo desempenho que a combinação convexa e, em alguns casos, apresenta algumas vantagens.

O ponto crítico da combinação afim é a maneira como é feita a adaptação do parâmetro de mistura. Na combinação convexa, é utilizada uma variável auxiliar que passa por uma função sigmoideal, mantendo o parâmetro de mistura no intervalo $[0, 1]$. Com isso é possível restringir a variável auxiliar a um determinado conjunto de valores máximo e mínimo, o que torna possível utilizar um passo de adaptação com valor elevado, melhorando a universalidade da combinação. No caso da combinação afim, entretanto, não há restrição. Isso faz com que a combinação possa apresentar alguns problemas devido à grande variação do gradiente utilizado na adaptação. Dentre os problemas, pode-se citar a divergência no início das iterações, a grande variação no valor do parâmetro de mistura e problemas com relação à universalidade da combinação. Para contornar esses problemas, foi feita uma análise do transitório da combinação e foram

propostas outras formas de adaptação do parâmetro de mistura.

Uma outra aplicação da combinação de algoritmos é a combinação de equalizadores CMA com inicializações diferentes. Como o desempenho do CMA depende da forma como o algoritmo foi inicializado, essa combinação é uma alternativa para evitar a convergência para um mínimo local.

A dissertação foi estruturada em cinco capítulos, sendo o Capítulo 2 relacionado com a análise em regime, o Capítulo 3 com a análise de transitório e o Capítulo 4 com a combinação de equalizadores CMA com inicializações diferentes. Na Tabela 5.1, são listados os principais tópicos abordados em cada capítulo e as conclusões são apresentadas a seguir.

Tabela 5.1: Tópicos abordados na dissertação.

Capítulo	Tópicos principais
2	- Propriedades da combinação afim
	- Implementação prática para obtenção do ganho de EMSE
	- Possibilidade de ganho de EMSE maior que 3 dB
3	- Algoritmo η -LMS
	- Algoritmo <i>signed regressor</i> (SR)
	- Algoritmo <i>power normalized</i> (PN)
	- Algoritmo <i>normalized signed regressor</i> (NSR)
4	- Combinações afim e convexa de equalizadores CMA de dois coeficientes com inicializações distintas
	- Transferência de coeficientes
	- Combinações de equalizadores com mais coeficientes

Análise em regime

Através da análise em regime, verificou-se as seguintes propriedades da combinação afim:

- o EMSE da combinação afim é sempre menor que o mínimo EMSE dos filtros componentes;
- quando o EMSE dos filtros componentes têm valores semelhantes, a combinação afim oferece um ganho de até 3 dB em relação ao mínimo EMSE dos filtros componentes. Essa situação pode ocorrer
 - no caso estacionário ou não estacionário, quando os passos de adaptação (ou fatores de esquecimento) dos filtros componentes são próximos. Nesse caso, o valor do parâmetro de mistura ótimo da combinação é um valor com módulo muito grande, menor que 0 ou maior que 1, na maioria dos casos;

- no caso não estacionário, com passos de adaptação distintos. Nesse caso, o valor do parâmetro de mistura ótimo é geralmente positivo, situado no intervalo $[0, 1]$.

A diferença entre o desempenho das combinações afim e convexa ocorre quando o valor do parâmetro de mistura ótimo se torna negativo ou quando o valor do parâmetro de mistura ótimo é maior que 1. Assim, quando os passos de adaptação (ou fatores de esquecimento) dos filtros componentes são próximos, a combinação afim é capaz de oferecer um ganho de EMSE de até 3 dB, enquanto a combinação convexa tem seu desempenho limitado pelo mínimo EMSE dos filtros componentes.

Em relação ao ganho obtido com a combinação, concluiu-se que:

- o ganho de EMSE da combinação não é uma situação ótima, apesar de ser mais favorável do que não obter ganho algum. No caso não-estacionário, o EMSE, em termos absolutos, é mínimo quando um dos filtros componentes é adaptado com passo (ou fator de esquecimento) ótimo, situação na qual a combinação afim ou convexa tem o desempenho igual ao do filtro componente ótimo, não oferecendo ganho algum. No caso estacionário, se fixados os parâmetros de um dos filtros, é possível verificar que o ganho de EMSE da combinação aumenta até o limite de 3 dB à medida que o EMSE do segundo filtro se aproxima do EMSE do primeiro filtro. Apesar do aumento do ganho, o EMSE da combinação em termos absolutos é menor à medida que o EMSE do segundo filtro é reduzido em relação ao EMSE do primeiro filtro. Nesse caso, deve-se levar em conta ainda o compromisso entre velocidade de convergência e nível de EMSE em regime pois à medida que o EMSE do segundo filtro é diminuído, sua velocidade de convergência também é diminuída, influenciando diretamente na velocidade de convergência da combinação;
- o uso da combinação afim com filtros componentes com passos de adaptação (ou fatores de esquecimento) próximos faz com que a combinação apresente uma velocidade de convergência extremamente lenta. Nesse caso, pode-se fixar o parâmetro de mistura da combinação, usando o valor analítico, sendo possível obter o ganho de EMSE com a mesma velocidade de convergência dos filtros componentes;
- o ganho oferecido pela combinação afim pode ser aumentado considerando uma combinação global de combinações componentes. Entretanto, verificou-se que o ganho oferecido por essa segunda combinação em relação ao EMSE das combinações componentes é aproximadamente igual a 1,26 dB, o que sugere que existe um limite de ganho de EMSE máximo que pode ser obtido se o esquema for estendido.

Análise de transitório

Com relação às formas de adaptação do parâmetro de mistura, concluiu-se que:

- em alguns casos, a adaptação da combinação afim com o algoritmo η -LMS não é capaz de fazer com que a combinação tenha um comportamento próximo ao universal;
- a utilização da normalização instantânea parcial na adaptação, com o algoritmo *signed regressor* se mostrou, nas simulações realizadas, uma boa opção para melhorar a universalidade da combinação;
- os algoritmos normalizados PN e NSR fazem com que a combinação afim obtenha um desempenho próximo ao obtido com o SR, sendo facilitada a escolha do passo de adaptação, que não depende de parâmetros do cenário de filtragem.

Combinação de equalizadores CMA com inicializações diferentes

Nesse contexto, concluiu-se que:

- a combinação afim é uma alternativa vantajosa em relação à combinação convexa. A maior liberdade na adaptação do parâmetro de mistura pode permitir que o vetor de coeficientes da combinação apresente uma variação maior, possibilitando sair de regiões de mínimos locais;
- através do uso da transferência de coeficientes, pode-se aumentar as chances de um filtro componente conseguir sair de uma região de mínimo local;
- numa aplicação mais realista, considerando filtros componentes com maior número de coeficientes, foram feitos testes preliminares e verificou-se que ambas as combinações são capazes de oferecer um desempenho próximo ao do filtro componente com melhor inicialização. Nesse caso, a transferência de coeficientes altera apenas o comportamento do filtro com a inicialização mais desfavorável, fazendo com que seu desempenho seja próximo ao do outro filtro componente e ao da combinação.

Trabalhos futuros

Diante das contribuições e conclusões expostas, existem algumas possibilidades de trabalhos futuros. Dentre elas, destacam-se:

- TF1 estudo sobre as causas do ganho de EMSE no caso de filtros componentes com EMSE próximos;
- TF2 cálculo do limite de ganho de EMSE no caso estacionário;
- TF3 estudo sobre como obter um ganho de EMSE da combinação no caso de equalizadores CMA com inicializações diferentes;

- TF4 análise e implementação da combinação afim de mais de dois filtros;
- TF5 extensão da análise de transitório da combinação afim para o caso de combinações de filtros de famílias diferentes;
- TF6 cálculo de uma expressão mais precisa para a variância do parâmetro de mistura, sem considerar a hipótese T1 (pág. 48);
- TF7 cálculo de uma expressão mais precisa para a matriz de covariância do vetor de erro de coeficientes do RLS;
- TF8 comparação do desempenho da combinação em relação ao desempenho de algoritmos de passo variável;
- TF9 extensão das análises para o caso complexo.

Cabe observar que o TF9 tem um interesse especial em equalização, onde são usados sinais complexos.

Referências Bibliográficas

- T. Aboulnasr e K. Mayyas. A robust variable step-size LMS-type algorithm: analysis and simulations. *IEEE Transactions on Signal Processing*, 45:631–639, Mar. 1997.
- P. Andersson. Adaptive forgetting in recursive identification through multiple models. *International Journal of Control*, 42:1175–1193, Nov. 1985.
- J. Arenas-García e A. R. Figueiras-Vidal. Adaptive combination of normalized filters for robust system identification. *Electronics Letters*, 41(15):874–875, Jul. 2005.
- J. Arenas-García e A. R. Figueiras-Vidal. Improved blind equalization via adaptive combination of constant modulus algorithms. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06)*, volume III, pages 756–759. IEEE, 2006.
- J. Arenas-García, M. Martínez-Ramón, V. Gómez-Verdejo e A. R. Figueiras-Vidal. Multiple plant identifier via adaptive LMS convex combination. In *Proceedings of IEEE International Workshop on Intelligent Signal Processing*, 2003.
- J. Arenas-García, A. R. Figueiras-Vidal e A. H. Sayed. Tracking properties of a convex combination of two adaptive filters. In *Proceedings of IEEE 13th Workshop on Statistical Signal Processing (SSP'05)*, pages 109–114. IEEE, 2005a.
- J. Arenas-García, V. Gómez-Verdejo e A. R. Figueiras-Vidal. New algorithms for improved adaptive convex combination of LMS transversal filters. *IEEE Transactions on Instrumentation and Measurement*, 54:2239–2249, Dec. 2005b.
- J. Arenas-García, A. R. Figueiras-Vidal e A. H. Sayed. Mean-square performance of a convex combination of two adaptive filters. *IEEE Transactions on Signal Processing*, 54:1078–1090, Mar. 2006a.
- J. Arenas-García, M. Martínez-Ramón, A. Navia-Vázquez e A. R. Figueiras-Vidal. Plant identification via adaptive combination of transversal filters. *Signal Processing*, 86:2430–2438, Sep. 2006b.

- L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal e J. Arenas-García. A normalized adaptation scheme for the convex combination of two adaptive filters. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08)*, 2008.
- N. J. Bershad. Analysis of the normalized LMS algorithm with gaussian inputs. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34:793–806, 1986.
- N. J. Bershad, J. C. M. Bermudez e J-Y. Tournet. An affine combination of two LMS adaptive filters - transient mean-square analysis. *IEEE Transactions on Signal Processing*, 56:1853–1864, May. 2008.
- R. C. Bilcu, P. Kuosmanen e K. Egiazarian. A transform domain LMS adaptive filter with variable step-size. *IEEE Signal Processing Letters*, 9:51–53, Feb. 2002.
- M. H. Costa e J. C. M. Bermudez. A robust variable step size algorithm for LMS adaptive filters. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06)*, 2006.
- M. H. Costa e J. C. M. Bermudez. An improved model for the normalized LMS algorithm with gaussian inputs and large number of coefficients. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'02)*, 2002.
- Z. Ding e Y. Li. *Blind Equalization and Identification*. Marcel Dekke, 2001.
- P. S. R. Diniz. *Adaptive Filtering: Algorithms and Practical Implementation*. Kluwer Academic Publishers, 2nd edition, 2008.
- E. Eleftheriou e D. D. Falconer. Tracking properties and steady-state performance of RLS adaptive filter algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-34:1097–1110, 1986.
- E. Eweda. Comparison of RLS, LMS and sign algorithms for tracking randomly time-varying channels. *IEEE Transactions on Signal Processing*, 42:2937–2944, Nov. 1994.
- B. Farhang-Boroujeny. *Adaptive Filters: Theory and Applications*. Wiley, 1998.
- D. N. Godard. Self-recovering equalization and carrier tracking in two dimensional data communication system. *IEEE Transactions on Communications*, 28:1867–1875, Nov. 1980.
- S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 4th edition, 2002.
- C. R. Johnson Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown e R. A. Casas. Blind equalization using the constant modulus criterion: a review. In *Proceedings of the IEEE*, 86:1927–1950, Oct. 1998.

- R. H. Kwong e E. W. Johnston. A variable step size LMS algorithm. *IEEE Transactions on Signal Processing*, 40:1633–1642, Jul. 1992.
- J. Mai e A. H. Sayed. A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers. *IEEE Transactions on Signal Processing*, 48:80–91, Jan. 2000.
- J. E. Mazo. On the independent theory of equalizer convergence. *Bell Syst. Tech. J.*, 58:963–993, May/Jun. 1979.
- V. H. Nascimento. The normalized LMS algorithm with dependent noise. In *Simpósio Brasileiro de Telecomunicações (SBrT'01)*, 2001.
- V. H. Nascimento e M. T. M. Silva. Stochastic stability for the Constant Modulus Algorithm (CMA). *Accepted for publication in the IEEE Transactions on Signal Processing*, 2008.
- M. Niedźwiecki. Identification of nonstationary stochastic systems using parallel estimation schemes. *IEEE Transactions on Automatic Control*, 35:329–334, Mar. 1990.
- M. Niedźwiecki. Multiple-model approach to finite memory adaptive filtering. *IEEE Transactions on Signal Processing*, 40:470–473, Feb. 1992.
- A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Companies, Inc., 1991.
- C. Samson e V. U. Reddy. Fixed point error analysis of the normalized ladder algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31:1177–1191, 1983.
- A. H. Sayed. *Fundamentals of Adaptive Filtering*. John Wiley & Sons, NJ, 2003.
- M. T. M. Silva e M. D. Miranda. Tracking issues of some blind equalization algorithms. *IEEE Signal Processing Letters*, 11:760–763, Sept. 2004.
- M. T. M. Silva e V. H. Nascimento. Tracking analysis of the Constant Modulus Algorithm. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08)*. IEEE, 2008a.
- M. T. M. Silva e V. H. Nascimento. Improving the tracking capability of adaptive filters via convex combination. *IEEE Transactions on Signal Processing*, 56:3137–3149, 2008b.
- M. T. M. Silva, M. Gerken e M. D. Miranda. An accelerated constant modulus algorithm. In *Proceedings of International Telecommunications Symposium*. SBrT, 2002.
- D. T. M. Slock. On the convergence behavior of the LMS and the normalized LMS algorithms. *IEEE Transactions on Signal Processing*, 41:2811–2825, Sep. 1993.

- J. R. Treichler, I. Fijalkow e C. R. Johnson Jr. Fractionally spaced equalizers. *IEEE Signal Processing Magazine*, 13:65–81, May 1996.
- N. R. Yousef e A. H. Sayed. A feedback analysis of the tracking performance of blind adaptive equalization algorithms. In *Proceedings of IEEE Conference on Decision and Control*, volume 1, pages 174–179, Dec. 1999.
- N. R. Yousef e A. H. Sayed. A unified approach to the steady-state and tracking analyses of adaptive filters. *IEEE Transactions on Signal Processing*, 49:314–324, Feb. 2001.
- W. Zhuang. RLS algorithm with variable forgetting factor for decision feedback equalizer over time-variant fading channels. *Wireless Personal Communications*, 8:15–29, Aug. 1998.

Apêndice A

Utilização do *random-walk model* em equalização

Na análise de algoritmos adaptativos no caso não-estacionário, é comum assumir o *random-walk model*, que considera uma solução ótima variante no tempo, segundo o modelo

$$\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n), \quad (\text{A.1})$$

sendo $\mathbf{q}(n)$ um vetor i.i.d. independente das condições iniciais $\{\mathbf{w}_o(0), \mathbf{w}(0)$ e de $\{\mathbf{u}(l), d(l)\}$ para todo l [Sayed, 2003, Sec. 7.4]. A aplicação desse modelo num cenário de equalização, entretanto, não é direta, já que a solução ótima $\mathbf{w}_o(n)$ geralmente não é conhecida. A seguir, será considerada uma forma equivalente de utilização do *random-walk model* numa aplicação de equalização, visando obter o mesmo EMSE, sem a necessidade do conhecimento da solução ótima.

Como definido no Capítulo 1, o EMSE $\zeta_i(n)$ de um filtro i é calculado como

$$\zeta_i(n) = \text{E} \{e_{a,i}^2(n)\}, \quad (\text{A.2})$$

sendo

$$e_{a,i}(n) = \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1). \quad (\text{A.3})$$

Lembrando que $\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o(n) - \mathbf{w}_i(n)$, (A.3) pode ser reescrita como

$$e_{a,i}(n) = \mathbf{u}^T(n) [\mathbf{w}_o(n-1) - \mathbf{w}_i(n-1)]. \quad (\text{A.4})$$

A partir de (A.1), é possível obter uma expressão para $\mathbf{w}_o(n)$ de forma iterativa, ou seja,

$$\begin{aligned}\mathbf{w}_o(1) &= \mathbf{w}_o(0) + \mathbf{q}(1) \\ \mathbf{w}_o(2) &= \mathbf{w}_o(0) + \mathbf{q}(1) + \mathbf{q}(2) \\ &\vdots \\ \mathbf{w}_o(n) &= \mathbf{w}_o(0) + \sum_{l=1}^n \mathbf{q}(l).\end{aligned}\tag{A.5}$$

De forma análoga, usando a expressão de atualização dos coeficientes do filtro

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \rho_i(n)\mathbf{M}_i(n)e(n)\mathbf{u}(n),\tag{A.6}$$

obtém-se

$$\begin{aligned}\mathbf{w}_i(1) &= \mathbf{w}_i(0) + \rho_i(1)\mathbf{M}_i(1)e(1)\mathbf{u}(1) \\ \mathbf{w}_i(2) &= \mathbf{w}_i(0) + \rho_i(1)\mathbf{M}_i(1)e(1)\mathbf{u}(1) + \rho_i(2)\mathbf{M}_i(2)e(2)\mathbf{u}(2) \\ &\vdots \\ \mathbf{w}_i(n) &= \mathbf{w}_i(0) + \sum_{l=1}^n \rho_i(l)\mathbf{M}_i(l)e(l)\mathbf{u}(l).\end{aligned}\tag{A.7}$$

Substituindo (A.5) e (A.7) em (A.4), obtém-se

$$e_{a,i}(n) = \mathbf{u}^T(n) \left\{ \mathbf{w}_o(0) + \sum_{l=1}^{n-1} \mathbf{q}(l) - \left[\mathbf{w}_i(0) + \sum_{l=1}^{n-1} \rho_i(l)\mathbf{M}_i(l)e(l)\mathbf{u}(l) \right] \right\},\tag{A.8}$$

que pode ser reescrita como

$$e_{a,i}(n) = \mathbf{u}^T(n) \left\{ \mathbf{w}_o(0) - \underbrace{\left[\mathbf{w}_i(0) + \sum_{l=1}^{n-1} \left(\rho_i(l)\mathbf{M}_i(l)e(l)\mathbf{u}(l) - \mathbf{q}(l) \right) \right]}_{\mathbf{w}_i^{(q)}(n-1)} \right\}.\tag{A.9}$$

É possível mostrar, de forma iterativa, que $\mathbf{w}_i^{(q)}(n)$ pode ser calculado através de

$$\mathbf{w}_i^{(q)}(n) = \mathbf{w}_i^{(q)}(n-1) + \rho_i(n)\mathbf{M}_i(n)e(n)\mathbf{u}(n) - \mathbf{q}(n),\tag{A.10}$$

com $\mathbf{w}_i^{(q)}(0) = \mathbf{w}_i(0)$.

Comparando (A.9) com (A.4), é possível notar que o $e_{a,i}(n)$ não será modificado se a solução ótima for considerada constante, ou seja, $\mathbf{w}_o(n-1) = \mathbf{w}_o(0)$ e se a adaptação dos

coeficientes do filtro for feita segundo (A.10), ou seja,

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \rho_i(n)\mathbf{M}_i(n)e(n)\mathbf{u}(n) - \mathbf{q}(n). \quad (\text{A.11})$$

Assim, é possível usar o *random-walk model* sem a necessidade do conhecimento da solução ótima $\mathbf{w}_o(n)$, considerando o vetor $\mathbf{q}(n)$ na adaptação dos coeficientes segundo (A.11), obtendo o mesmo erro *a priori* e, conseqüentemente, o mesmo EMSE.

Apêndice B

Expressão do EMSE cruzado em regime para o NLMS

Uma expressão para

$$\zeta_{12}(\infty) \triangleq \lim_{n \rightarrow \infty} E \{e_{a,1}(n)e_{a,2}(n)\} \quad (\text{B.1})$$

pode ser obtida a partir do cálculo de $E \{\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)\}$. No NLMS, o vetor de coeficientes do filtro i é adaptado através de

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \frac{\tilde{\mu}_i}{\epsilon_N + \|\mathbf{u}(n)\|^2} \mathbf{u}(n)e_i(n). \quad (\text{B.2})$$

Subtraindo ambos os lados de (B.2) de $\mathbf{w}_o(n)$ e usando o *random-walk model* (Eq. (1.12), pág. 5), obtém-se

$$\tilde{\mathbf{w}}_i(n) = \tilde{\mathbf{w}}_i(n-1) + \mathbf{q}(n) - \frac{\tilde{\mu}_i}{\epsilon_N + \|\mathbf{u}(n)\|^2} \mathbf{u}(n)e_i(n). \quad (\text{B.3})$$

Usando (B.3) para calcular $\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)$, obtém-se

$$\begin{aligned} \tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n) &= \tilde{\mathbf{w}}_1^T(n-1)\tilde{\mathbf{w}}_2(n-1) + \tilde{\mathbf{w}}_1^T(n-1)\mathbf{q}(n) - \tilde{\mathbf{w}}_1^T(n-1)\frac{\tilde{\mu}_2}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e_2(n) \\ &\quad + \mathbf{q}^T(n)\tilde{\mathbf{w}}_2(n-1) + \mathbf{q}^T(n)\mathbf{q}(n) - \mathbf{q}^T(n)\frac{\tilde{\mu}_2}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e_2(n) \\ &\quad - \frac{\tilde{\mu}_1}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}^T(n)e_1(n)\tilde{\mathbf{w}}_2(n-1) - \frac{\tilde{\mu}_1}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}^T(n)e_1(n)\mathbf{q}(n) \\ &\quad + \frac{\tilde{\mu}_1}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}^T(n)e_1(n)\frac{\tilde{\mu}_2}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e_2(n). \end{aligned} \quad (\text{B.4})$$

Calculando a esperança de ambos os lados de (B.4) e considerando que $\mathbf{q}(n)$ é independente de $\tilde{\mathbf{w}}_i(n)$, de $\mathbf{u}(n)$ e de $e_i(n)$ e que $E\{\mathbf{q}(n)\} = \mathbf{0}$, chega-se a

$$\begin{aligned} E\{\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)\} &= E\{\tilde{\mathbf{w}}_1^T(n-1)\tilde{\mathbf{w}}_2(n-1)\} \\ &\quad - E\left\{\tilde{\mathbf{w}}_1^T(n-1)\frac{\tilde{\mu}_2}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e_2(n)\right\} \\ &\quad + E\{\mathbf{q}^T(n)\mathbf{q}(n)\} \\ &\quad - E\left\{\frac{\tilde{\mu}_1}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}^T(n)e_1(n)\tilde{\mathbf{w}}_2(n-1)\right\} \\ &\quad + E\left\{\frac{\tilde{\mu}_1}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}^T(n)e_1(n)\frac{\tilde{\mu}_2}{\epsilon_N + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e_2(n)\right\}. \end{aligned} \quad (\text{B.5})$$

Considerando que $\epsilon_N \ll \|\mathbf{u}(n)\|^2$ e lembrando que

$$e_i(n) = e_{a,i}(n) + v(n), \quad (\text{B.6})$$

$$e_{a,i}(n) = \mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1) = \tilde{\mathbf{w}}_i^T(n-1)\mathbf{u}(n) \quad (\text{B.7})$$

e

$$\mathbf{u}^T(n)\mathbf{u}(n) = \|\mathbf{u}(n)\|^2, \quad (\text{B.8})$$

(B.5) pode ser reescrita como

$$\begin{aligned} E\{\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)\} &= E\{\tilde{\mathbf{w}}_1^T(n-1)\tilde{\mathbf{w}}_2(n-1)\} - E\left\{\frac{\tilde{\mu}_2}{\|\mathbf{u}(n)\|^2}e_{a,1}(n)e_{a,2}(n)\right\} \\ &\quad + E\{\mathbf{q}^T(n)\mathbf{q}(n)\} \\ &\quad - E\left\{\frac{\tilde{\mu}_1}{\|\mathbf{u}(n)\|^2}e_{a,1}(n)e_{a,2}(n)\right\} \\ &\quad + E\left\{\frac{\tilde{\mu}_1\tilde{\mu}_2}{\|\mathbf{u}(n)\|^2}e_{a,1}(n)e_{a,2}(n)\right\} \\ &\quad + E\left\{\frac{\tilde{\mu}_1\tilde{\mu}_2}{\|\mathbf{u}(n)\|^2}v^2(n)\right\} \\ &\quad - E\left\{[\tilde{\mu}_2e_{a,1}(n) + \tilde{\mu}_1e_{a,2}(n)]\frac{1}{\|\mathbf{u}(n)\|^2}v(n)\right\} \\ &\quad - E\left\{[e_{a,1}(n) + e_{a,2}(n)]\frac{\tilde{\mu}_1\tilde{\mu}_2}{\|\mathbf{u}(n)\|^2}v(n)\right\}. \end{aligned} \quad (\text{B.9})$$

Considerando que $v(n)$ é independente de $\mathbf{u}(n)$ e de $e_{a,i}(n)$ e que $E\{v(n)\} = 0$, conclui-se que as duas últimas parcelas do lado direito de (B.9) são nulas. Com isso, usando a aproximação

$$E\left\{\frac{1}{\|\mathbf{u}(n)\|^2}e_{a,1}(n)e_{a,2}(n)\right\} \approx \frac{E\{e_{a,1}(n)e_{a,2}(n)\}}{\text{Tr}(\mathbf{R})} = \frac{\zeta_{12}(n)}{\text{Tr}(\mathbf{R})}, \quad (\text{B.10})$$

lembrando que $E\{\|\mathbf{q}(n)\|^2\} = \text{Tr}(\mathbf{Q})$ e calculando o $\lim_{n \rightarrow \infty}$ de ambos os lados de (B.9), obtém-se

$$\begin{aligned} \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)\} &= \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_1^T(n-1)\tilde{\mathbf{w}}_2(n-1)\} + \tilde{\mu}_1\tilde{\mu}_2 E\left\{\frac{1}{\|\mathbf{u}(n)\|^2}\right\} \sigma_v^2 \\ &\quad + \frac{\zeta_{12}(\infty)}{\text{Tr}(\mathbf{R})} [\tilde{\mu}_1\tilde{\mu}_2 - \tilde{\mu}_1 - \tilde{\mu}_2] + \text{Tr}(\mathbf{Q}). \end{aligned} \quad (\text{B.11})$$

Em regime, vale a aproximação

$$\zeta_{12}(\infty) = \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_1^T(n)\tilde{\mathbf{w}}_2(n)\} \approx \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_1^T(n-1)\tilde{\mathbf{w}}_2(n-1)\}. \quad (\text{B.12})$$

Usando (B.12) em (B.11), obtém-se

$$\zeta_{12}(\infty) = \frac{\text{Tr}(\mathbf{R}) [\tilde{\mu}_1\tilde{\mu}_2\alpha_u\sigma_v^2 + \text{Tr}(\mathbf{Q})]}{\tilde{\mu}_1 + \tilde{\mu}_2 - \tilde{\mu}_1\tilde{\mu}_2}, \quad (\text{B.13})$$

em que $\alpha_u \triangleq E\{1/\|\mathbf{u}(n)\|^2\}$. Para M suficientemente grande e sinal de entrada gaussiano, α_u pode ser estimado como $1/[\sigma_u^2(M-2)]$.

Apêndice C

Expressão do EMSE cruzado em regime para o RLS

Em [Eleftheriou e Falconer, 1986], foi obtida uma expressão analítica para o EMSE em regime do RLS. A seguir, esse resultado é estendido para a obtenção do EMSE cruzado da combinação de dois algoritmos RLS.

Assumindo que $\mathbf{u}(n)$ é independente de $\tilde{\mathbf{w}}_i(n-1)$, $i = 1, 2$, o EMSE de um filtro componente ($\zeta_i(n) = \zeta_{ii}(n)$, $i = 1, 2$) ou o EMSE cruzado ($\zeta_{12}(n)$) pode ser obtido através da estimativa de

$$\zeta_{ij}(\infty) \triangleq \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n-1)\mathbf{R}\tilde{\mathbf{w}}_j(n-1)\}, i, j = 1, 2, \quad (\text{C.1})$$

sendo $\mathbf{R} \triangleq E\{\mathbf{u}(n)\mathbf{u}^T(n)\}$ a matriz de autocorrelação do sinal de entrada e $\tilde{\mathbf{w}}_i(n-1)$ o vetor de erro dos coeficientes do filtro i .

A equação de atualização dos coeficientes do filtro i com o algoritmo RLS é dada por

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \hat{\mathbf{R}}_i^{-1}(n)\mathbf{u}(n)e_i(n), \quad (\text{C.2})$$

sendo

$$\hat{\mathbf{R}}_i(n) = \lambda_i \hat{\mathbf{R}}_i(n-1) + \mathbf{u}(n)\mathbf{u}^T(n) \quad (\text{C.3})$$

e

$$e_i(n) = d(n) - \mathbf{u}^T(n)\mathbf{w}_i(n-1). \quad (\text{C.4})$$

Multiplicando (C.2) à esquerda por $\hat{\mathbf{R}}_i(n)$, obtém-se

$$\hat{\mathbf{R}}_i(n)\mathbf{w}_i(n) = \hat{\mathbf{R}}_i(n)\mathbf{w}_i(n-1) + \mathbf{u}(n)e_i(n). \quad (\text{C.5})$$

Substituindo (C.3) e (C.4) em (C.5), chega-se a

$$\widehat{\mathbf{R}}_i(n)\mathbf{w}_i(n) = \lambda_i \widehat{\mathbf{R}}_i(n-1)\mathbf{w}_i(n-1) + \mathbf{u}(n)d(n). \quad (\text{C.6})$$

Para prosseguir, é necessário assumir um modelo para o sinal desejado $d(n)$ e para a variação da solução ótima $\mathbf{w}_o(n)$. Assume-se o modelo de regressão linear para o sinal desejado, ou seja,

$$d(n) = \mathbf{u}^T(n)\mathbf{w}_o(n-1) + v(n), \quad (\text{C.7})$$

sendo $v(n)$ um ruído branco de média nula e variância σ_v^2 . Assume-se que solução ótima varia de acordo com o *random-walk model*, dado por

$$\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n), \quad (\text{C.8})$$

sendo $\mathbf{q}(n)$ um vetor aleatório de média nula, i.i.d., com matriz de covariância $\mathbf{Q} = \mathbb{E}\{\mathbf{q}(n)\mathbf{q}^T(n)\}$, independente de $\mathbf{u}(m)$ para todo $m \leq n$ e das condições iniciais $\mathbf{w}_o(0), \mathbf{w}_i(0)$.

Multiplicando (C.3) à direita por $\mathbf{w}_o(n)$ e usando (C.8), chega-se a

$$\begin{aligned} \widehat{\mathbf{R}}_i(n)\mathbf{w}_o(n) &= \lambda_i \widehat{\mathbf{R}}_i(n-1)\mathbf{w}_o(n-1) + \mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}_o(n-1) \\ &\quad + \lambda_i \widehat{\mathbf{R}}_i(n-1)\mathbf{q}(n) + \mathbf{u}(n)\mathbf{u}^T(n)\mathbf{q}(n). \end{aligned} \quad (\text{C.9})$$

Subtraindo (C.6) de (C.9) e substituindo $d(n)$ pelo modelo (C.7), obtém-se

$$\begin{aligned} \widehat{\mathbf{R}}_i(n)[\mathbf{w}_o(n) - \mathbf{w}_i(n)] &= \lambda_i \widehat{\mathbf{R}}_i(n-1)[\mathbf{w}_o(n-1) - \mathbf{w}_i(n-1)] - \mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}_o(n-1) \\ &\quad - \mathbf{u}(n)v(n) + \mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}_o(n-1) + \lambda_i \widehat{\mathbf{R}}_i(n-1)\mathbf{q}(n) \\ &\quad + \mathbf{u}(n)\mathbf{u}^T(n)\mathbf{q}(n). \end{aligned} \quad (\text{C.10})$$

Definindo o vetor de erro dos coeficientes $\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o(n) - \mathbf{w}_i(n)$, após algumas simplificações, (C.10) se reduz a

$$\widehat{\mathbf{R}}_i(n)\tilde{\mathbf{w}}_i(n) = \lambda_i \widehat{\mathbf{R}}_i(n-1)\tilde{\mathbf{w}}_i(n-1) - \mathbf{u}(n)v(n) + \underbrace{[\lambda_i \widehat{\mathbf{R}}_i(n-1) + \mathbf{u}(n)\mathbf{u}^T(n)]}_{\widehat{\mathbf{R}}_i(n)} \mathbf{q}(n). \quad (\text{C.11})$$

Multiplicando à esquerda por $\widehat{\mathbf{R}}_i^{-1}(n)$, chega-se a

$$\tilde{\mathbf{w}}_i(n) = \lambda_i \widehat{\mathbf{R}}_i^{-1}(n)\widehat{\mathbf{R}}_i(n-1)\tilde{\mathbf{w}}_i(n-1) - \widehat{\mathbf{R}}_i^{-1}(n)\mathbf{u}(n)v(n) + \mathbf{q}(n). \quad (\text{C.12})$$

Usando (C.12), é possível calcular $\lim_{n \rightarrow \infty} \mathbb{E}\{\tilde{\mathbf{w}}_i^T(n)\mathbf{R}\tilde{\mathbf{w}}_j(n)\}$, ou seja,

$$\begin{aligned}
& \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n) \mathbf{R} \tilde{\mathbf{w}}_j(n)\} = \\
& + \lambda_i \lambda_j \underbrace{\lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n-1) \hat{\mathbf{R}}_i(n-1) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \hat{\mathbf{R}}_j(n-1) \tilde{\mathbf{w}}_j(n-1)\}}_{\mathcal{A}} \\
& + \underbrace{\lim_{n \rightarrow \infty} E\{\mathbf{u}^T(n) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \mathbf{u}(n) v^2(n)\}}_{\mathcal{B}} \\
& + \underbrace{\lim_{n \rightarrow \infty} E\{\mathbf{q}^T(n) \mathbf{R} \mathbf{q}(n)\}}_{\mathcal{C}}. \tag{C.13}
\end{aligned}$$

Para simplificar, são apresentados em (C.13) apenas os termos não nulos. Como é assumido que $\mathbf{q}(n)$ é um vetor i.i.d. de média nula, independente das condições iniciais e $v(n)$ é independente de $\mathbf{u}(n)$ e $\tilde{\mathbf{w}}(n)$ e também tem média nula, termos que dependem de $E\{v(n)\}$ e $E\{\mathbf{q}(n)\}$ são nulos e por isso não foram considerados em (C.13).

Na sequência, aproximações para os termos $\mathcal{A}$, $\mathcal{B}$ e $\mathcal{C}$ são obtidas. Para isso, são necessárias algumas hipóteses.

Hipótese 1: Assume-se que em regime $\hat{\mathbf{R}}_i(n) \approx \hat{\mathbf{R}}_i(n-1)$, ou seja,

$$\hat{\mathbf{R}}_i^{-1}(n) \hat{\mathbf{R}}_i(n-1) \approx \mathbf{I}, \text{ quando } n \rightarrow \infty.$$

Usando a Hipótese 1, o termo $\mathcal{A}$ pode ser aproximado por

$$\mathcal{A} \approx \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n-1) \mathbf{R} \tilde{\mathbf{w}}_j(n-1)\} = \zeta_{ij}(\infty). \tag{C.14}$$

Cabe observar que em regime também vale a aproximação

$$\lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n) \mathbf{R} \tilde{\mathbf{w}}_j(n)\} \approx \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}_i^T(n-1) \mathbf{R} \tilde{\mathbf{w}}_j(n-1)\}. \tag{C.15}$$

O termo $\mathcal{C}$, por sua vez, é igual a $\text{Tr}(\mathbf{Q}\mathbf{R})$ sendo $\text{Tr}(\cdot)$ o traço de uma matriz. Usando essas aproximações, (C.13) pode ser reescrita como

$$(1 - \lambda_i \lambda_j) \zeta_{ij}(\infty) = \underbrace{\lim_{n \rightarrow \infty} E\{\mathbf{u}^T(n) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \mathbf{u}(n) v^2(n)\}}_{\mathcal{B}} + \text{Tr}(\mathbf{Q}\mathbf{R}). \tag{C.16}$$

Observando (C.16), ainda é necessário encontrar uma aproximação para o termo $\mathcal{B}$. Como $v(n)$ é independente de $\mathbf{u}(n)$, então

$$\mathcal{B} \approx \underbrace{\lim_{n \rightarrow \infty} E\{\mathbf{u}^T(n) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \mathbf{u}(n)\}}_{\mathcal{B}_1} \sigma_v^2. \tag{C.17}$$

Basta agora encontrar uma aproximação para $\mathcal{B}_1$, o que é feito em detalhes na Seção C.1. Usando o resultado dessa seção, o termo $\mathcal{B}$ pode ser aproximado por

$$\mathcal{B} \approx (1 - \lambda_i)(1 - \lambda_j) \left(1 + \frac{(1 - \lambda_i)(1 - \lambda_j)}{1 - \lambda_i \lambda_j} \gamma_u \right) M \sigma_v^2. \quad (\text{C.18})$$

Substituindo (C.18) em (C.16), chega-se a

$$\zeta_{ij} \approx \frac{(1 - \lambda_i)(1 - \lambda_j)}{(1 - \lambda_i \lambda_j)} \left(1 + \frac{(1 - \lambda_i)(1 - \lambda_j)}{1 - \lambda_i \lambda_j} \gamma_u \right) M \sigma_v^2 + \frac{1}{(1 - \lambda_i \lambda_j)} \text{Tr}(\mathbf{Q}\mathbf{R}), \quad (\text{C.19})$$

sendo $\gamma_u \triangleq \frac{\sigma_u^2}{[\sigma_u^2]^2}$. Se $u(n)$ for gaussiano, $\gamma_u = 2$.

Assim, fazendo $i = j$ obtém-se o resultado para o EMSE dos filtros componentes

$$\zeta_{ii}(\infty) \approx \frac{1 - \lambda_i}{1 + \lambda_i} \left(1 + \frac{1 - \lambda_i}{1 + \lambda_i} \gamma_u \right) M \sigma_v^2 + \frac{1}{1 - \lambda_i^2} \text{Tr}(\mathbf{Q}\mathbf{R}), \quad i = 1, 2 \quad (\text{C.20})$$

e fazendo $i = 1$ e $j = 2$, obtém-se o EMSE cruzado

$$\zeta_{12}(\infty) \approx \frac{(1 - \lambda_1)(1 - \lambda_2)}{(1 - \lambda_1 \lambda_2)} \left(1 + \frac{(1 - \lambda_1)(1 - \lambda_2)}{1 - \lambda_1 \lambda_2} \gamma_u \right) M \sigma_v^2 + \frac{1}{(1 - \lambda_1 \lambda_2)} \text{Tr}(\mathbf{Q}\mathbf{R}). \quad (\text{C.21})$$

C.1 Expressão para o termo $\mathcal{B}_1$

O objetivo aqui é encontrar uma aproximação para o termo

$$\mathcal{B}_1 = \lim_{n \rightarrow \infty} \text{E}\{\mathbf{u}^T(n) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \mathbf{u}(n)\}.$$

Tomando a esperança de ambos os lados de (C.3), obtém-se em regime a seguinte aproximação

$$\lim_{n \rightarrow \infty} \text{E}\{\hat{\mathbf{R}}_i(n)\} \approx \frac{1}{1 - \lambda_i} \mathbf{R}. \quad (\text{C.22})$$

Dessa forma, um bom modelo para $\hat{\mathbf{R}}_i(n)$ é

$$\hat{\mathbf{R}}_i(n) \approx \frac{1}{1 - \lambda_i} \mathbf{R} + \tilde{\mathbf{R}}_i(n), \quad (\text{C.23})$$

sendo $\tilde{\mathbf{R}}_i(n)$ uma matriz de perturbação hermitiana de média nula.

Usando o lema da inversão matricial¹ em (C.23), chega-se a

$$\widehat{\mathbf{R}}_i^{-1}(n) = (1 - \lambda_i)\mathbf{R}^{-1} - (1 - \lambda_i)^2\mathbf{R}^{-1} \left[\widetilde{\mathbf{R}}_i^{-1}(n) + (1 - \lambda_i)\mathbf{R}^{-1} \right]^{-1} \mathbf{R}^{-1}. \quad (\text{C.24})$$

Multiplicando (C.24) à direita por $\mathbf{R}$, obtém-se

$$\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R} = (1 - \lambda_i)\mathbf{I} - (1 - \lambda_i)^2\mathbf{R}^{-1} \left[\widetilde{\mathbf{R}}_i^{-1}(n) + (1 - \lambda_i)\mathbf{R}^{-1} \right]^{-1}. \quad (\text{C.25})$$

Lembrando da propriedade de matrizes $\mathbf{A}(\mathbf{B} + \mathbf{C})^{-1} = (\mathbf{B}\mathbf{A}^{-1} + \mathbf{C}\mathbf{A}^{-1})^{-1}$, (C.25) pode ser reescrita como

$$\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R} = (1 - \lambda_i)\mathbf{I} - (1 - \lambda_i)^2 \left[\widetilde{\mathbf{R}}_i^{-1}(n)\mathbf{R} + (1 - \lambda_i)\mathbf{I} \right]^{-1}. \quad (\text{C.26})$$

Usando novamente o lema da inversão matricial, é possível verificar que

$$\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R} = (1 - \lambda_i) \left[\mathbf{I} + (1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i^{-1}(n) \right]^{-1}. \quad (\text{C.27})$$

Hipótese 2: Assume-se que

$$\|(1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n)\| < 1,$$

sendo $\|\cdot\|$ algum tipo consistente de norma.

Utilizando a Hipótese 2, vale a aproximação

$$\left[(1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n) \right]^{-1} + \mathbf{I} \approx \left[(1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n) \right]^{-1}. \quad (\text{C.28})$$

Aplicando o lema da inversão matricial em $\left[\mathbf{I} + (1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n) \right]$ e a aproximação (C.28), chega-se a

$$\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R} \approx (1 - \lambda_i) \left[\mathbf{I} - (1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n) \right]. \quad (\text{C.29})$$

Definindo-se a matriz $\mathbf{F}_i(n) \triangleq (1 - \lambda_i)\mathbf{R}^{-1}\widetilde{\mathbf{R}}_i(n)$, (C.29) pode ser reescrita como

$$\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R} \approx (1 - \lambda_i) [\mathbf{I} - \mathbf{F}_i(n)]. \quad (\text{C.30})$$

Hipótese 3: Assume-se que $\mathbf{F}_i(n)$ é independente de $\mathbf{u}(n)$ e $E\{\mathbf{F}_i(n)\} = \mathbf{0}$.

Note que $\mathbf{F}_i(n)$ representa a flutuação do produto $\widehat{\mathbf{R}}_i^{-1}(n)\mathbf{R}$ em torno da identidade.

¹Seja $\mathbf{A}$ uma matriz tal que $\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^T$. Segundo o lema da inversão matricial, sua inversa pode ser calculada como [Sayed, 2003, pág. 67]

$$\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^T\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{B}.$$

Usando (C.30) e a Hipótese 3, o termo $\mathcal{B}_1$ pode ser reescrito como

$$\begin{aligned}\mathcal{B}_1 &= \lim_{n \rightarrow \infty} E\{\mathbf{u}^T(n) \hat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \hat{\mathbf{R}}_j^{-1}(n) \mathbf{u}(n)\} \approx \\ &\approx (1 - \lambda_i)(1 - \lambda_j) \lim_{n \rightarrow \infty} E\{\mathbf{u}^T(n) [\mathbf{I} - \mathbf{F}_i(n)] [\mathbf{I} - \mathbf{F}_j(n)] \mathbf{R}^{-1} \mathbf{u}(n)\} \approx \\ &\approx (1 - \lambda_i)(1 - \lambda_j) \lim_{n \rightarrow \infty} \text{Tr}(E\{[\mathbf{I} - \mathbf{F}_i(n)] [\mathbf{I} - \mathbf{F}_j(n)]\})\end{aligned}\quad (\text{C.31})$$

Note que aqui se utilizou o fato de que $E\{\mathbf{R}^{-1} \mathbf{u}(n) \mathbf{u}^T(n)\} = \mathbf{I}$. Como $E\{\mathbf{F}_i(n)\} = \mathbf{0}$,

$$\text{Tr}(E\{[\mathbf{I} - \mathbf{F}_i(n)] [\mathbf{I} - \mathbf{F}_j(n)]\}) = \text{Tr}(\mathbf{I} + E\{\mathbf{F}_i(n) \mathbf{F}_j(n)\}). \quad (\text{C.32})$$

Assim,

$$\mathcal{B}_1 \approx (1 - \lambda_i)(1 - \lambda_j) \lim_{n \rightarrow \infty} \text{Tr}(\mathbf{I} + E\{\mathbf{F}_i(n) \mathbf{F}_j(n)\}). \quad (\text{C.33})$$

Cabe observar que $E\{\mathbf{F}_i(n) \mathbf{F}_j(n)\} = (1 - \lambda_i)(1 - \lambda_j) E\{\mathbf{R}^{-1} \tilde{\mathbf{R}}_i(n) \mathbf{R}^{-1} \tilde{\mathbf{R}}_j(n)\}$ não é fácil de estimar e é dependente de estatísticas de quarta ordem do sinal de entrada. No entanto, em situações específicas, é possível calcular uma aproximação para essa matriz. Supondo que o sinal de entrada seja branco, pode-se assumir em regime que apenas os elementos da diagonal de $\mathbf{R}$, $\tilde{\mathbf{R}}_i(n)$ e $\tilde{\mathbf{R}}_j(n)$ são importantes para o cálculo do EMSE. Mesmo quando o sinal de entrada não for branco, a importância dos elementos da diagonal pode ser considerada como uma aproximação razoável na maior parte dos casos [Diniz, 2008].

Os elementos da diagonal principal de $E\{\mathbf{F}_i(n) \mathbf{F}_j(n)\}$ podem ser aproximados por

$$E\{f_{xx,i}(n) f_{xx,j}(n)\} = (1 - \lambda_i)(1 - \lambda_j) \frac{E\{\tilde{r}_{xx,i}(n) \tilde{r}_{xx,j}(n)\}}{(\sigma_u^2)^2}, \quad (\text{C.34})$$

sendo σ_u^2 a variância do sinal de entrada $u(n)$ e $\tilde{r}_{xx,i}(n)$ os elementos da diagonal principal de $\tilde{\mathbf{R}}_i(n)$. Usando (C.23) e (C.3), é possível obter uma recorrência para $\tilde{r}_{xx,i}(n)$, ou seja,

$$\tilde{r}_{xx,i}(n) = \lambda \tilde{r}_{xx,i}(n-1) + u^2(n-x) - r_{xx,i}, \quad (\text{C.35})$$

sendo r_{ii} os elementos da diagonal principal da matriz $\mathbf{R}$. Multiplicando por $\tilde{r}_{xx,j}(n)$ e tomando a esperança, chega-se a

$$E\{\tilde{r}_{xx,i}(n) \tilde{r}_{xx,j}(n)\} = \lambda_i \lambda_j E\{\tilde{r}_{xx,i}(n-1) \tilde{r}_{xx,j}(n-1)\} + E\{[u^2(n-x) - r_{xx}]^2\}. \quad (\text{C.36})$$

Cabe observar que $\tilde{r}_{xx,i}$ e $\tilde{r}_{xx,j}$ têm média nula e são assumidos independentes de $u(n)$. Em regime, obtém-se

$$E\{\tilde{r}_{xx,i}(n) \tilde{r}_{xx,j}(n)\} = \frac{1}{1 - \lambda_i \lambda_j} E\{[u^2(n-x) - r_{xx}]^2\} = \frac{1}{1 - \lambda_i \lambda_j} \sigma_{u^2}^2. \quad (\text{C.37})$$

Substituindo (C.37) em (C.34), chega-se a

$$E\{f_{xx,i}(n)f_{xx,j}(n)\} = \frac{(1-\lambda_i)(1-\lambda_j)}{1-\lambda_i\lambda_j} \underbrace{\frac{\sigma_u^2}{[\sigma_u^2]^2}}_{\gamma_u}. \quad (\text{C.38})$$

Finalmente, usando (C.38), $\mathcal{B}_1$ pode ser aproximado por

$$\mathcal{B}_1 \approx (1-\lambda_i)(1-\lambda_j) \left(1 + \frac{(1-\lambda_i)(1-\lambda_j)}{1-\lambda_i\lambda_j} \gamma_u \right) M. \quad (\text{C.39})$$

Apêndice D

Expressões da matriz de covariância do vetor de erro dos coeficientes

O objetivo é obter uma expressão analítica recorrente no caso estacionário para

$$\mathbf{S}_{ij}(n) = \mathbb{E} \{ \tilde{\mathbf{w}}_i(n) \tilde{\mathbf{w}}_j^T(n) \}, \quad (\text{D.1})$$

sendo

$$\tilde{\mathbf{w}}_i(n) = \mathbf{w}_o - \mathbf{w}_i(n) \quad (\text{D.2})$$

o vetor de erro dos coeficientes do filtro i e $\mathbf{w}_o$ o vetor de coeficientes ótimos.

De modo unificado, a equação de adaptação dos coeficientes é dada por

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + \rho_i(n) \mathbf{M}_i(n) e_i(n) \mathbf{u}(n), \quad (\text{D.3})$$

com $\rho_i(n)$, $\mathbf{M}_i(n)$ e $e(n)$ referentes aos algoritmos considerados mostrados na Tabela D.1.

O erro de um filtro componente pode ser escrito de forma análoga ao erro da combinação obtido na Equação (2.24) (pág. 20), ou seja,

$$e_i(n) \approx \kappa(n) e_{a,i}(n) + \varphi(n), \quad (\text{D.4})$$

sendo $\kappa(n)$ e $\varphi(n)$ definidos na Tabela D.1 para os algoritmos considerados.

Tabela D.1: Parâmetros das expressões (D.3) e (D.4) para os algoritmos considerados.

Alg.	$\rho_i(n)$	$\mathbf{M}_i^{-1}(n)$	$e_i(n)$	$\kappa(n)$	$\varphi(n)$
LMS	μ_i	$\mathbf{I}$	$d(n) - y_i(n)$	1	$v(n)$
NLMS	$\tilde{\mu}_i / (\epsilon_N + \ \mathbf{u}(n)\ ^2)$				
RLS	1	$\hat{\mathbf{R}}_i(n)$	$[r - y_i^2(n)] y_i(n)$	$\gamma(n)$	$\beta(n)$
CMA	μ_i	$\mathbf{I}$			

Subtraindo ambos os lados de (D.3) da solução ótima $\mathbf{w}_o$, obtém-se

$$\tilde{\mathbf{w}}_i(n) = \tilde{\mathbf{w}}_i(n-1) - \rho_i(n)\mathbf{M}_i(n)e_i(n)\mathbf{u}(n). \quad (\text{D.5})$$

Substituindo (D.4) em (D.5) e lembrando que $e_{a,i}(n) = \mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1)$, chega-se a

$$\tilde{\mathbf{w}}_i(n) = \underbrace{\tilde{\mathbf{w}}_i(n-1)}_{\mathbf{x}_{i,1}(n)} - \underbrace{\rho_i(n)\kappa(n)\mathbf{M}_i(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}_i(n-1)}_{\mathbf{x}_{i,2}(n)} - \underbrace{\rho_i(n)\varphi(n)\mathbf{M}_i(n)\mathbf{u}(n)}_{\mathbf{x}_{i,3}(n)}. \quad (\text{D.6})$$

Substituindo (D.6) em (D.1), obtém-se

$$\begin{aligned} \mathbf{S}_{ij}(n) = & \mathbb{E}\{\mathbf{x}_{i,1}(n)\mathbf{x}_{j,1}^T(n) + \mathbf{x}_{i,1}(n)\mathbf{x}_{j,2}^T(n) + \mathbf{x}_{i,1}(n)\mathbf{x}_{j,3}^T(n) \\ & + \mathbf{x}_{i,2}(n)\mathbf{x}_{j,1}^T(n) + \mathbf{x}_{i,2}(n)\mathbf{x}_{j,2}^T(n) + \mathbf{x}_{i,2}(n)\mathbf{x}_{j,3}^T(n) \\ & + \mathbf{x}_{i,3}(n)\mathbf{x}_{j,1}^T(n) + \mathbf{x}_{i,3}(n)\mathbf{x}_{j,2}^T(n) + \mathbf{x}_{i,3}(n)\mathbf{x}_{j,3}^T(n)\}. \end{aligned} \quad (\text{D.7})$$

Hipótese 1: Assume-se que $\tilde{\mathbf{w}}_k(n)$, $k = 1, 2$ é independente de $\mathbf{u}(n)$, $\kappa(n)$ e $\varphi(n)$ e que $\mathbb{E}\{\tilde{\mathbf{w}}_i(n)\} = 0$, $i = 1, 2$.

Hipótese 2: Assume-se que $\varphi(n)$ é independente de $\mathbf{u}(n)$ e $\kappa(n)$.

Usando as Hipóteses 1 e 2, verifica-se que os termos $\mathbb{E}\{\mathbf{x}_{i,k}(n)\mathbf{x}_{j,3}^T(n)\}$ e $\mathbb{E}\{\mathbf{x}_{i,3}(n)\mathbf{x}_{j,k}^T(n)\}$, $k = 1, 2$ são nulos. Assim, (D.7) se reduz a

$$\begin{aligned} \mathbf{S}_{ij}(n) \approx & \underbrace{\mathbb{E}\{\mathbf{x}_{i,1}(n)\mathbf{x}_{j,1}^T(n)\}}_{\mathcal{A}} + \underbrace{\mathbb{E}\{\mathbf{x}_{i,2}(n)\mathbf{x}_{j,2}^T(n)\}}_{\mathcal{B}} + \underbrace{\mathbb{E}\{\mathbf{x}_{i,3}(n)\mathbf{x}_{j,3}^T(n)\}}_{\mathcal{C}} \\ & + \underbrace{\mathbb{E}\{\mathbf{x}_{i,1}(n)\mathbf{x}_{j,2}^T(n)\}}_{\mathcal{D}} + \underbrace{\mathbb{E}\{\mathbf{x}_{i,2}(n)\mathbf{x}_{j,1}^T(n)\}}_{\mathcal{E}}. \end{aligned} \quad (\text{D.8})$$

O termo $\mathcal{A}$ é comum para todos os algoritmos considerados e é calculado como

$$\mathcal{A} = \mathbb{E}\{\mathbf{x}_{i,1}(n)\mathbf{x}_{j,1}^T(n)\} = \mathbb{E}\{\tilde{\mathbf{w}}_i(n-1)\tilde{\mathbf{w}}_j^T(n-1)\} = \mathbf{S}_{ij}(n-1). \quad (\text{D.9})$$

Os termos $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ e $\mathcal{E}$ dependem do algoritmo utilizado na adaptação de $\mathbf{w}(n)$. A seguir, esses termos são calculados para cada algoritmo considerado. No cálculo, $\mathbf{u}(n)$ é assumido gaussiano e com essa hipótese vale o resultado [Haykin, 2002]

$$\mathbb{E}\{\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{S}_{ij}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\} = 2\mathbf{R}\mathbf{S}_{ij}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1)), \quad (\text{D.10})$$

que é usado no cálculo dos termos a seguir.

D.1 LMS

Para o LMS, consideram-se os parâmetros $\rho_i(n)$, $\mathbf{M}_i(n)$, $\kappa(n)$ e φ da Tabela D.1. Dessa forma, os termos $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ e $\mathcal{E}$ são calculados a seguir.

Termo $\mathcal{B}$: O termo $\mathcal{B}$ é dado por

$$\mathcal{B} = \mathbb{E} \left\{ \mu_i \mu_j \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\}, \quad (\text{D.11})$$

que pode ser reescrito, usando a esperança condicional, como [Papoulis, 1991]

$$\begin{aligned} \mathcal{B} &= \mu_i \mu_j \mathbb{E} \left\{ \mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \middle| \mathbf{u}(n) \right\} \right\} \\ &= \mu_i \mu_j \mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \mathbf{S}_{ij}(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\}. \end{aligned} \quad (\text{D.12})$$

Usando o resultado de (D.10) em (D.12), obtém-se

$$\mathcal{B} = \mu_i \mu_j [2\mathbf{R}\mathbf{S}_{ij}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1))]. \quad (\text{D.13})$$

Termo $\mathcal{C}$: O termo $\mathcal{C}$ é dado por

$$\mathcal{C} = \mathbb{E} \left\{ \mu_i \mu_j v^2(n) \mathbf{u}(n) \mathbf{u}^T(n) \right\} = \mu_i \mu_j \sigma_v^2 \mathbf{R}. \quad (\text{D.14})$$

Termo $\mathcal{D}$: O termo $\mathcal{D}$ é dado por

$$\mathcal{D} = \mathbb{E} \left\{ -\mu_j \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \approx -\mu_j \mathbf{S}_{ij}(n-1) \mathbf{R}. \quad (\text{D.15})$$

Termo $\mathcal{E}$: O termo $\mathcal{E}$ é aproximado por

$$\mathcal{E} = \mathbb{E} \left\{ -\mu_i \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \right\} \approx -\mu_i \mathbf{R} \mathbf{S}_{ij}(n-1). \quad (\text{D.16})$$

Substituindo esses resultados e (D.9) em (D.8), obtém-se

$$\begin{aligned} \mathbf{S}_{ij}(n) &\approx \mathbf{S}_{ij}(n-1) - \mu_i \mathbf{R} \mathbf{S}_{ij}(n-1) - \mu_j \mathbf{S}_{ij}(n-1) \mathbf{R} \\ &\quad + \mu_i \mu_j [2\mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}]. \end{aligned} \quad (\text{D.17})$$

D.2 NLMS

No cálculo dos termos de $\mathbf{S}_{ij}(n)$ para o NLMS são usadas as aproximações

$$\mathbb{E} \left\{ \frac{\mathbf{u}(n) \mathbf{u}^T(n)}{\mathbf{u}^T(n) \mathbf{u}(n)} \right\} \approx \mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \right\} \mathbb{E} \left\{ \frac{1}{\mathbf{u}^T(n) \mathbf{u}(n)} \right\} \quad (\text{D.18})$$

e

$$\mathbb{E} \left\{ \frac{\mathbf{u}(n)\mathbf{u}^T(n)}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \right\} \approx \mathbb{E} \{ \mathbf{u}(n)\mathbf{u}^T(n) \} \mathbb{E} \left\{ \frac{1}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \right\}. \quad (\text{D.19})$$

Segundo [Costa e Bermudez, 2002], essas aproximações são razoáveis considerando que o número de coeficientes dos filtros componentes seja grande ($M > 30$) e são equivalentes às aproximações da teoria das médias de [Samson e Reddy, 1983].

Assumindo que $\mathbf{u}(n)$ é gaussiano, os seguintes resultados são válidos [Costa e Bermudez, 2002]

$$\mathbb{E} \left\{ \frac{1}{\mathbf{u}^T(n)\mathbf{u}(n)} \right\} = \frac{1}{\sigma_u^2(M-2)} \quad (\text{D.20})$$

e

$$\mathbb{E} \left\{ \frac{1}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \right\} = \frac{1}{\sigma_u^4(M-2)(M-4)}, \quad (\text{D.21})$$

sendo σ_u^2 a variância do sinal de entrada $u(n)$. Usando (D.20) e (D.21) em (D.18) e (D.19), obtém-se

$$\mathbb{E} \left\{ \frac{\mathbf{u}(n)\mathbf{u}^T(n)}{\mathbf{u}^T(n)\mathbf{u}(n)} \right\} \approx \frac{\mathbf{R}}{\sigma_u^2(M-2)} \quad (\text{D.22})$$

e

$$\mathbb{E} \left\{ \frac{\mathbf{u}(n)\mathbf{u}^T(n)}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \right\} \approx \frac{\mathbf{R}}{\sigma_u^4(M-2)(M-4)}. \quad (\text{D.23})$$

Com esses resultados, substituindo os parâmetros $\rho_i(n)$, $\mathbf{M}_i(n)$, $\kappa(n)$ e φ mostrados na Tabela D.1, considerando que a constante ϵ_N é suficientemente pequena e lembrando que $\|\mathbf{u}(n)\|^2 = \mathbf{u}^T(n)\mathbf{u}(n)$, os termos $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ e $\mathcal{E}$ são calculados a seguir.

Termo $\mathcal{B}$: O termo $\mathcal{B}$ é dado por

$$\mathcal{B} = \mathbb{E} \left\{ \frac{\tilde{\mu}_i \tilde{\mu}_j}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \mathbf{u}(n)\mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n)\mathbf{u}^T(n) \right\}. \quad (\text{D.24})$$

Usando a esperança condicional e considerando que $[\mathbf{u}^T(n)\mathbf{u}(n)]^2$ e o numerador de (D.24) são pouco correlacionados para M suficientemente grande, é possível fazer a aproximação

$$\mathcal{B} \approx \tilde{\mu}_i \tilde{\mu}_j \mathbb{E} \left\{ \frac{1}{[\mathbf{u}^T(n)\mathbf{u}(n)]^2} \right\} \mathbb{E} \{ \mathbb{E} \{ \mathbf{u}(n)\mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n)\mathbf{u}^T(n) | \mathbf{u}(n) \} \}.$$

Usando (D.21), chega-se a

$$\mathcal{B} \approx \frac{\tilde{\mu}_i \tilde{\mu}_j}{\sigma_u^4(M-2)(M-4)} \mathbb{E} \{ \mathbf{u}(n)\mathbf{u}^T(n) \mathbf{S}_{ij}(n-1) \mathbf{u}(n)\mathbf{u}^T(n) \}. \quad (\text{D.25})$$

Usando o resultado de (D.10) em (D.25), obtém-se

$$\mathcal{B} \approx \frac{\tilde{\mu}_i \tilde{\mu}_j}{\sigma_u^4(M-2)(M-4)} [2\mathbf{R}\mathbf{S}_{ij}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1))]. \quad (\text{D.26})$$

Termo $\mathcal{C}$: O termo $\mathcal{C}$ é dado por

$$\mathcal{C} = \mathbb{E} \left\{ \frac{\tilde{\mu}_i \tilde{\mu}_j}{[\mathbf{u}^T(n) \mathbf{u}(n)]^2} v^2(n) \mathbf{u}(n) \mathbf{u}^T(n) \right\} = \tilde{\mu}_i \tilde{\mu}_j \sigma_v^2 \mathbb{E} \left\{ \frac{\mathbf{u}(n) \mathbf{u}^T(n)}{[\mathbf{u}^T(n) \mathbf{u}(n)]^2} \right\}, \quad (\text{D.27})$$

que, usando (D.23), pode ser aproximado por

$$\mathcal{C} \approx \tilde{\mu}_i \tilde{\mu}_j \sigma_v^2 \frac{\mathbf{R}}{\sigma_u^4 (M-2)(M-4)}. \quad (\text{D.28})$$

Termo $\mathcal{D}$: O termo $\mathcal{D}$ é dado por

$$\begin{aligned} \mathcal{D} &= \mathbb{E} \left\{ -\frac{\tilde{\mu}_j}{\mathbf{u}^T(n) \mathbf{u}(n)} \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \\ &\approx -\tilde{\mu}_j \mathbf{S}_{ij}(n-1) \mathbb{E} \left\{ \frac{\mathbf{u}(n) \mathbf{u}^T(n)}{\mathbf{u}^T(n) \mathbf{u}(n)} \right\}, \end{aligned} \quad (\text{D.29})$$

que pode ser aproximado, usando (D.22), por

$$\mathcal{D} \approx -\frac{\tilde{\mu}_j}{\sigma_u^2 (M-2)} \mathbf{S}_{ij}(n-1) \mathbf{R}. \quad (\text{D.30})$$

Termo $\mathcal{E}$: O termo $\mathcal{E}$ é dado por

$$\begin{aligned} \mathcal{E} &= \mathbb{E} \left\{ -\frac{\tilde{\mu}_i}{\mathbf{u}^T(n) \mathbf{u}(n)} \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \right\} \\ &\approx -\tilde{\mu}_i \mathbb{E} \left\{ \frac{\mathbf{u}(n) \mathbf{u}^T(n)}{\mathbf{u}^T(n) \mathbf{u}(n)} \right\} \mathbf{S}_{ij}(n-1), \end{aligned} \quad (\text{D.31})$$

que pode ser aproximado, usando (D.22), por

$$\mathcal{E} \approx -\frac{\tilde{\mu}_i}{\sigma_u^2 (M-2)} \mathbf{R} \mathbf{S}_{ij}(n-1). \quad (\text{D.32})$$

Substituindo esses resultados e (D.9) em (D.8), obtém-se

$$\begin{aligned} \mathbf{S}_{ij}(n) &\approx \mathbf{S}_{ij}(n-1) - \frac{\tilde{\mu}_i}{\sigma_u^2 (M-2)} \mathbf{R} \mathbf{S}_{ij}(n-1) - \frac{\tilde{\mu}_j}{\sigma_u^2 (M-2)} \mathbf{S}_{ij}(n-1) \mathbf{R} \\ &\quad + \frac{\tilde{\mu}_i \tilde{\mu}_j}{\sigma_u^4 (M-2)(M-4)} [2\mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}]. \end{aligned} \quad (\text{D.33})$$

D.3 RLS

Assumindo que $\hat{\mathbf{R}}_i^{-1}(n)$ varia pouco em relação a $\tilde{\mathbf{w}}_i(n)$, $i = 1, 2$ quando $\lambda_i \approx 1$ e considerando os parâmetros $\rho_i(n)$, $\mathbf{M}_i(n)$, $\kappa(n)$ e φ mostrados na Tabela D.1, os termos $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ e $\mathcal{E}$ são calculados a seguir.

Termo $\mathcal{B}$: O termo $\mathcal{B}$ é dado por

$$\mathcal{B} = \text{E} \left\{ \widehat{\mathbf{R}}_i^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \widetilde{\mathbf{w}}_i(n-1) \widetilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \widehat{\mathbf{R}}_j^{-1}(n) \right\}, \quad (\text{D.34})$$

que pode ser reescrito, usando a esperança condicional, como [Papoulis, 1991]

$$\begin{aligned} \mathcal{B} &\approx \overline{\widehat{\mathbf{R}}_i^{-1}(n)} \text{E} \left\{ \text{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \widetilde{\mathbf{w}}_i(n-1) \widetilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \mid \mathbf{u}(n) \right\} \right\} \widehat{\mathbf{R}}_j^{-1}(n) \\ &\approx \overline{\widehat{\mathbf{R}}_i^{-1}(n)} \text{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \mathbf{S}_{ij}(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \widehat{\mathbf{R}}_j^{-1}(n), \end{aligned} \quad (\text{D.35})$$

sendo

$$\overline{\widehat{\mathbf{R}}_i^{-1}(n)} = \text{E} \left\{ \widehat{\mathbf{R}}_i^{-1}(n) \right\}. \quad (\text{D.36})$$

Usando o resultado de (D.10) em (D.35), obtém-se

$$\mathcal{B} \approx \overline{\widehat{\mathbf{R}}_i^{-1}(n)} [2\mathbf{R}\mathbf{S}_{ij}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1))] \widehat{\mathbf{R}}_j^{-1}(n). \quad (\text{D.37})$$

Termo $\mathcal{C}$: O termo $\mathcal{C}$ é dado por

$$\mathcal{C} = \text{E} \left\{ v^2(n) \widehat{\mathbf{R}}_i^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \widehat{\mathbf{R}}_j^{-1}(n) \right\} \approx \sigma_v^2 \overline{\widehat{\mathbf{R}}_i^{-1}(n)} \mathbf{R} \widehat{\mathbf{R}}_j^{-1}(n). \quad (\text{D.38})$$

Termo $\mathcal{D}$: O termo $\mathcal{D}$ é dado por

$$\mathcal{D} = \text{E} \left\{ -\widetilde{\mathbf{w}}_i(n-1) \widetilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \widehat{\mathbf{R}}_j^{-1}(n) \right\} \approx -\mathbf{S}_{ij}(n-1) \mathbf{R} \widehat{\mathbf{R}}_j^{-1}(n). \quad (\text{D.39})$$

Termo $\mathcal{E}$: O termo $\mathcal{E}$ é dado por

$$\mathcal{E} = \text{E} \left\{ -\widehat{\mathbf{R}}_i^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \widetilde{\mathbf{w}}_i(n-1) \widetilde{\mathbf{w}}_j^T(n-1) \right\} \approx -\overline{\widehat{\mathbf{R}}_i^{-1}(n)} \mathbf{R} \mathbf{S}_{ij}(n-1). \quad (\text{D.40})$$

É necessário, então, uma aproximação para $\overline{\widehat{\mathbf{R}}_i^{-1}(n)}$. Lembrando que a estimativa da matriz de autocorrelação no filtro i é inicializada como

$$\widehat{\mathbf{R}}_i(0) = \epsilon_R \mathbf{I} \quad (\text{D.41})$$

e atualizada por

$$\widehat{\mathbf{R}}_i(n) = \lambda_i \widehat{\mathbf{R}}_i(n-1) + \mathbf{u}(n) \mathbf{u}^T(n). \quad (\text{D.42})$$

É possível obter uma expressão para $\widehat{\mathbf{R}}_i(n)$ de forma iterativa, ou seja,

$$\begin{aligned}\widehat{\mathbf{R}}_i(0) &= \epsilon_R \mathbf{I} \\ \widehat{\mathbf{R}}_i(1) &= \lambda_i \epsilon_R \mathbf{I} + \mathbf{u}(1) \mathbf{u}^T(1) \\ \widehat{\mathbf{R}}_i(2) &= \lambda_i^2 \epsilon_R \mathbf{I} + \lambda_i \mathbf{u}(1) \mathbf{u}^T(1) + \mathbf{u}(2) \mathbf{u}^T(2) \\ &\vdots \\ \widehat{\mathbf{R}}_i(n) &= \lambda_i^n \epsilon_R \mathbf{I} + \sum_{l=1}^n \lambda_i^{n-l} \mathbf{u}(l) \mathbf{u}^T(l).\end{aligned}\tag{D.43}$$

Aplicando a esperança em (D.43), obtém-se

$$\mathbb{E} \left\{ \widehat{\mathbf{R}}_i(n) \right\} = \lambda_i^n \epsilon_R \mathbf{I} + \mathbf{R} \frac{1 - \lambda_i^n}{1 - \lambda_i}.\tag{D.44}$$

Para prosseguir, assume-se que $\mathbb{E} \{ \widehat{\mathbf{R}}_i^{-1}(n) \} \approx [\mathbb{E} \{ \widehat{\mathbf{R}}_i(n) \}]^{-1}$. Essa hipótese não é muito realista mas evita o cálculo da inversa de (D.43) e mostrou ser um resultado razoável para o transitório do EMSE nas simulações realizadas. O estudo de uma aproximação mais precisa será tratado num trabalho futuro. Com isso,

$$\widehat{\mathbf{R}}_i^{-1}(n) \approx \left[\mathbb{E} \left\{ \widehat{\mathbf{R}}_i(n) \right\} \right]^{-1} = \left[\lambda_i^n \epsilon_R \mathbf{I} + \mathbf{R} \frac{1 - \lambda_i^n}{1 - \lambda_i} \right]^{-1}.\tag{D.45}$$

Assim, usando esses resultados e (D.9) em (D.8), obtém-se

$$\begin{aligned}\mathbf{S}_{ij}(n) &\approx \mathbf{S}_{ij}(n-1) - \widehat{\mathbf{R}}_i^{-1}(n) \mathbf{R} \mathbf{S}_{ij}(n-1) - \mathbf{S}_{ij}(n-1) \mathbf{R} \widehat{\mathbf{R}}_j^{-1}(n) \\ &\quad + \widehat{\mathbf{R}}_i^{-1}(n) [2 \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_v^2 \mathbf{R}] \widehat{\mathbf{R}}_j^{-1}(n),\end{aligned}\tag{D.46}$$

em que $\widehat{\mathbf{R}}_i^{-1}(n)$ é dado por (D.45).

D.4 CMA

Para o CMA, considerando os parâmetros $\rho_i(n)$, $\mathbf{M}_i(n)$, $\kappa(n)$ e φ mostrados na Tabela D.1, os termos $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ e $\mathcal{E}$ são calculados a seguir.

Termo $\mathcal{B}$: O termo $\mathcal{B}$ é dado por

$$\mathcal{B} = \mathbb{E} \left\{ \mu_i \mu_j \gamma^2(n) \mathbf{u}(n) \mathbf{u}^T(n) \widetilde{\mathbf{w}}_i(n-1) \widetilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\},\tag{D.47}$$

que pode ser reescrito, usando a esperança condicional, como [Papoulis, 1991]

$$\begin{aligned}\mathcal{B} &\approx \mu_i \mu_j \xi \mathbb{E} \left\{ \mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \middle| \mathbf{u}(n) \right\} \right\} \\ &\approx \mu_i \mu_j \xi \mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \mathbf{S}_{ij}(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\},\end{aligned}\quad (\text{D.48})$$

com $\xi \triangleq \mathbb{E}\{\gamma^2(n)\}$. Usando o resultado de (D.10) em (D.48), obtém-se

$$\mathcal{B} \approx \mu_i \mu_j \xi [2\mathbf{R}\mathbf{S}_{ij}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}_{ij}(n-1))]. \quad (\text{D.49})$$

Termo $\mathcal{C}$: O termo $\mathcal{C}$ é dado por

$$\mathcal{C} = \mathbb{E} \left\{ \mu_i \mu_j \beta^2(n) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \approx \mu_i \mu_j \sigma_\beta^2 \mathbf{R}. \quad (\text{D.50})$$

Termo $\mathcal{D}$: O termo $\mathcal{D}$ é aproximado por

$$\mathcal{D} = \mathbb{E} \left\{ -\mu_j \gamma(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \approx -\mu_j \bar{\gamma} \mathbf{S}_{ij}(n-1) \mathbf{R}. \quad (\text{D.51})$$

Termo $\mathcal{E}$: O termo $\mathcal{E}$ é dado por

$$\mathcal{E} = \mathbb{E} \left\{ -\mu_i \gamma(n) \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}_i(n-1) \tilde{\mathbf{w}}_j^T(n-1) \right\} \approx -\mu_i \bar{\gamma} \mathbf{R} \mathbf{S}_{ij}(n-1). \quad (\text{D.52})$$

Substituindo esses resultados e (D.9) em (D.8), obtém-se

$$\begin{aligned}\mathbf{S}_{ij}(n) &\approx \mathbf{S}_{ij}(n-1) - \mu_i \bar{\gamma} \mathbf{R} \mathbf{S}_{ij}(n-1) - \mu_j \bar{\gamma} \mathbf{S}_{ij}(n-1) \mathbf{R} \\ &\quad + \mu_i \mu_j [2\xi \mathbf{R} \mathbf{S}_{ij}(n-1) \mathbf{R} + \xi \mathbf{R} \text{Tr}(\mathbf{R} \mathbf{S}_{ij}(n-1)) + \sigma_\beta^2 \mathbf{R}].\end{aligned}\quad (\text{D.53})$$