CEFET/RJ
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO

ALGORITMO FX-LMS COM COMPENSAÇÃO DE VIÉS

Lucas Paiva Rodrigues da Silva

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador(a): Diego Barreto Haddad
Coorientador(a): Amaro Azevedo de Lima

Rio de Janeiro,
Fevereiro 2021

ALGORITMO FX-LMS COM COMPENSAÇÃO DE VIÉS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Lucas Paiva Rodrigues da Silva

Banca Examinadora:

Diego Barreto Haddad

Presidente, Professor D.Sc. Diego Barreto Haddad (CEFET/RJ) (Orientador(a))



Professor Ph.D. Amaro Azevedo de Lima (CEFET/RJ) (Coorientador(a))



Professora D.Sc. Milena Faria Pinto (CEFET/RJ)

Tadeu Nagashima Ferreira

Professor D.Sc. Tadeu Nagashima Ferreira (UFF)

Rio de Janeiro,

Fevereiro 2021

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

S586 Silva, Lucas Paiva Rodrigues da
Algoritmo FX-LMS com compensação de viés / Lucas Paiva Rodrigues da Silva — 2020.
76f : il. (algumas color.) , enc.

Dissertação (Mestrado) Centro Federal de Educação Tecnológica Celso Suckow da Fonseca , 2020.
Bibliografia : f. 69-76
Orientador: Diego Barreto Haddad
Coorientador: Amaro Azevedo de Lima

1.Controle de ruído. 2. Processamento de sinais – Técnicas digitais. 3. Algorítmos. 4. Teoria dos sinais (Telecomunicações). I. Haddad, Diego Barreto.(Orient.). II. Lima, Amaro Azevedo de (Coorient.) Título.

CDD 621.38224

Elaborada pela bibliotecária Tania Mello – CRB/7 nº 5507/04

AGRADECIMENTOS

Agradeço a minha família por todo o amparo e recursos disponibilizados para a minha formação.

Aos colegas e amigos que fiz durante o curso e também aos de longa data, que me ajudaram a suportar os percalços e os momentos difíceis com mais leveza e alegria.

A todos os professores que me instruíram e contribuíram para a minha jornada.

Aos membros da banca por aceitarem o convite e auxiliarem no aperfeiçoamento desta dissertação.

Ao meu Coorientador Amaro, que sempre me recebeu bem e me deu suporte.

E um agradecimento especial ao meu Orientador Diego Haddad, por toda ajuda e dedicação nesses últimos dois anos, pela disponibilidade de me atender até nos fins de semana e por toda paciência até mesmo nas dúvidas mais ingênuas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 pela bolsa de Mestrado.

RESUMO

Algoritmo FX-LMS com Compensação de Viés

Devido às desvantagens das técnicas passivas de cancelamento de ruído e vibrações e à emergência dos ambientes microprocessados capacitados para efetuar o processamento de sinais em tempo real, as variantes ativas destas técnicas vêm atraindo um interesse crescente. Cumpre a tais técnicas sintetizar artificialmente um sinal (ou diversos sinais) que, por meio do fenômeno da interferência destrutiva e da técnica de filtragem adaptativa, mitigue(m) ao máximo o distúrbio de interesse (seja este vibração, seja este ruído). Historicamente, apenas sensores situados no ponto de atuação do distúrbio são modelados na literatura como ruidosos. Entretanto, sabemos que o sinal de entrada do filtro adaptativo também pode ser corrompido por ruído de medição (o qual pode ser inclusive oriundo do processo de quantização). Este trabalho modela o impacto nocivo que tal ruído apresenta no estimador do algoritmo de filtragem adaptativa *Filtered-X Least Mean Square (FX-LMS)*, demonstrando a emergência de um viés assintótico capaz de induzir perda de desempenho. Tal impacto é confirmado por simulações. Para mitigar este problema, propõe-se um novo algoritmo *Bias Compensated Filtered-X Least Mean Square (BC-FX-LMS)*, o qual é capaz de compensar este viés. O novo algoritmo requer uma estimativa da variância do ruído de medição, porém continua sendo vantajoso, em alguns contextos, mesmo quando tal estimativa é demasiado imprecisa. Como resultado, o novo algoritmo foi capaz de obter, a depender do cenário de teste, ganhos no MSD em regime permanente superiores a 6 dB.

Palavras-chave: Controle Ativo de Ruído, FX-LMS, Compensação de Viés

ABSTRACT

Algoritmo FX-LMS com Compensação de Viés

Due to the disadvantages of passive noise and vibration cancellation techniques and the emergence of digital processors capable of performing real-time signal processing, the active variants of these techniques have been attracting increasing interest. Such techniques must artificially synthesize a signal (or several signals) that, by means of the phenomenon of destructive interference and the adaptive filtering technique, mitigate the maximum disturbance of interest (whether this vibration or this noise). Historically, only sensors located at the point of action of the disturbance are modeled in the literature as noisy. However, we know that the input signal of the adaptive filter can also be corrupted by measurement noise (which may even come from the quantization process). This work models the harmful impact that such noise has on the estimator of the FX-LMS adaptive filtering algorithm, demonstrating the emergence of an asymptotic bias capable of inducing loss of performance. This impact is confirmed by simulations. To mitigate this problem, a new algorithm BC-FX-LMS is proposed, which is able to compensate for this bias. The new algorithm requires an estimate of the variance of the measurement noise, but it remains advantageous in some contexts, even when such an estimate is too imprecise. As a result, the new algorithm was able to obtain, depending on the test scenario, gains in the MSD in a steady state greater than 6 dB.

Keywords: Active Noise Cancellation, FX-LMS, Bias-Compensated

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação de um sistema discreto no tempo. 12

Figura 2 – Exemplo de função de transferência (modelo 1 de [12]). 14

Figura 3 – Exemplo de transformação de um sinal $u(k)$, cujo resultado dá origem a um sinal $d(k)$. Os elementos do vetor $\boldsymbol{w}^{\star}$ contém os coeficientes da função de transferência (ou resposta ao impulso, cujo suporte é considerado finito) do sistema linear e invariante no tempo que mapeia a sequência $u(k)$ na sequência $d(k)$. 14

Figura 4 – Diagrama em blocos de uma tarefa de identificação adaptativa de sistemas. 15

Figura 5 – Diagrama em blocos do algoritmo FX-LMS. 18

Figura 6 – Ilustração da configuração analisada neste trabalho, na qual o sinal de entrada do filtro é corrompido por um ruído aditivo $\eta(k)$. 19

Figura 7 – Estrutura típica de um filtro adaptativo. 24

Figura 8 – Evolução do MSD (em dB) ao longo das iterações, para $\beta \in \{10^{-3}, 10^{-2,5}, 10^{-2}\}$. 30

Figura 9 – Evolução do MSE (em dB) em ao longo das iterações, para $\beta \in \{10^{-3}, 10^{-2,5}, 10^{-2}\}$. 31

Figura 10 – Diagrama em blocos do algoritmo LMS. 32

Figura 11 – Comparação da evolução do desvio médio simulado referente aos três primeiros coeficientes (linha azul sólida) com a evolução predita pela teoria (Equação (69), em linha preta tracejada). Os resultados simulados foram obtidos por meio de 10 ensaios independentes de Monte Carlo. 38

Figura 12 – Sistema de controle ativo de ruído [50]. 42

Figura 13 – Diagrama em blocos do algoritmo FX-LMS. 43

Figura 14 – Coeficientes do filtro $\boldsymbol{w}^{\star}$, responsável pelo acoplamento entre o ruído do primário e o sinal registrado no microfone de erro. 46
Figura 15 – As curvas sólidas representam valores médios (obtidos a partir de 100 ensaios independentes de Monte Carlo) dos coeficientes adaptativos estimados com o algoritmo FX-LMS, assumindo-se uma estimativa perfeita do caminho do secundário. As retas tracejadas indicam os respectivos valores de $\boldsymbol{w}_{\text{opt}}$, calculados de acordo com (93). 47
Figura 16 – Evolução do MSD (em dB) ao longo das iterações. A curva foi computada através de 100 ensaios independentes de Monte Carlo. 48
Figura 17 – Evolução do MSE (em dB) ao longo das iterações. A curva foi computada através de 100 ensaios independentes de Monte Carlo. 48
Figura 18 – Diagrama do modelo de esparsidade em blocos para geração da resposta ao impulso. 56
Figura 19 – Evolução média dos três primeiros coeficientes de desvio, com 10 ensaios independentes de Monte Carlo. O modelo teórico (127) é representado por linhas vermelhas tracejadas. 64
Figura 20 – Comportamento de estado estacionário dos três primeiros coeficientes de desvio com 10^5 testes independentes de Monte Carlo. As linhas vermelhas tracejadas horizontais são estimativas teóricas obtidas de (129). 65
Figura 21 – Evolução MSD do FX-LMS com algoritmos de entrada ruidosa (em azul) e BC-FX-LMS (em vermelho). 66
Figura 22 – MSD em estado estacionário do algoritmo BC-FX-LMS em relação a α, onde $\hat{\sigma}_{\eta}^2 = \alpha\sigma_{\eta}^2$. A linha azul tracejada é o MSD assintótico obtido pelo algoritmo FX-LMS com entrada ruidosa. 67

LISTA DE TABELAS

Tabela 1 – Complexidade computacional envolvida em uma atualização do algoritmo LMS. A complexidade de uma subtração é considerada equivalente à de uma adição. 28

Tabela 2 – Complexidade computacional envolvida em uma atualização do algoritmo FX-LMS. A complexidade de uma subtração é considerada equivalente à de uma adição. O cômputo de $e'(k)$ não implica operações aritméticas, já que ele é resultado da medida do resultado líquido da interferência física entre os sinais do primário e do secundário. 44

Tabela 3 – Contexto das principais referências que utilizam o princípio da compensação de viés. 52

Tabela 4 – Definições dos principais sinais utilizados pelos diferentes algoritmos desta dissertação. 62

Tabela 5 – Simulações divididas por cenários e seus respectivos parâmetros. 63

LISTA DE ABREVIATURAS E SIGLAS

ANC	Active Noise Control
BC-FX-LMS	Bias Compensated Filtered-X Least Mean Square
FIR	Finite Impulse Response
FX-LMS	Filtered-X Least Mean Square
LMS	Least Mean Square
MSD	Mean Square Deviation
MSE	Mean Square Error
NLMS	Normalized Least Mean Square
RLS	Recursive Least Square
SPU	Selective Partial Update

SUMÁRIO

1	**Introdução**	**10**
1.1	Objetivos deste Trabalho	18
1.2	Metodologia	20
1.3	Estrutura	21
1.4	Notação	22
1.5	Publicações associadas a esta dissertação	22
2	**Algoritmo Least Mean Squares**	**23**
2.1	Introdução	23
2.2	O Algoritmo Gradiente Descendente	24
2.3	LMS	27
2.4	Simulações	29
3	**Análise de Primeira Ordem do LMS**	**32**
3.1	Hipóteses Estatísticas	33
3.2	Análise de Primeira Ordem do LMS	34
3.2.1	Exemplo Ilustrativo	37
3.2.2	Desacoplamento dos desvios	38
3.3	Conclusões Parciais	41
4	**FX-LMS**	**42**
4.1	Solução Ótima do FX-LMS	44
4.2	Exemplo Ilustrativo	46
4.3	Conclusões Parciais	48
5	**Algoritmos Adaptativos com Compensação de Viés**	**49**
5.1	Introdução	49
5.2	Revisão Bibliográfica	50

5.3 Redução do Custo Computacional de uma Técnica de Compensação de Viés 52
5.4 Comentários Parciais 57

6 Algoritmo FX-LMS com Compensação de Viés 58
6.1 Introdução 58
6.2 FX-LMS com Entrada Ruidosa 59
6.3 Análise de Primeira Ordem do FX-LMS com Entrada Ruidosa 59
6.4 Algoritmo Proposto BC-FX-LMS 61

7 Resultados e Discussões 63

8 Conclusões e Trabalhos Futuros 68

Referências Bibliográficas 69

1- Introdução

Com a disseminação dos ambientes microprocessados, o processamento digital de sinais tem angariado crescente destaque nas últimas décadas. A necessidade de tratar e manipular sinais de natureza diversa, impulsionada pelos avanços científico e tecnológico, motiva cientistas e engenheiros a pesquisarem por aprimoramentos na área.

Um dos problemas cuja solução pode ser encontrada neste campo reside na atenuação de ruídos ou alteração de componentes espectrais indesejadas [1]. A poluição sonora é um problema ubíquo nas grandes cidades e nas indústrias, podendo impactar diretamente a saúde de seres humanos [2], modular a auto-percepção da saúde de um indivíduo (um poderoso preditor de mortalidade [3]), prejudicar o desempenho de estudantes e professores [4], sendo possivelmente até um fator de risco adicional para o desenvolvimento de doenças como o Alzheimer [5].

Este trabalho enfoca o problema de cancelamento de ruído e de vibrações[1], o qual pode ser abordado de duas formas: passiva ou ativa [6]. Enquanto a modalidade passiva se concentra sobre os materiais capazes de absorver as frequências mais críticas do sinal de perturbação, a modalidade ativa se vale do princípio da interferência destrutiva, pretendendo gerar um sinal artificial que reduza a intensidade de um sinal considerado indesejável ou mesmo danoso à saúde humana.

Os métodos passivos de cancelamento de ruído são os mais tradicionais, em termos históricos. Um avanço recente neste tipo de método consiste na adoção de impressão 3D de guias de onda, cuja geometria é projetada para atenuar sinais considerados perturbadores [7]. Um exemplo importante de aplicação de técnicas passivas de cancelamento de ruído acústico ocorre com o ruído oriundo da cavidade dos pneus dos automóveis, cujo espectro apresenta uma característica de “banda estreita”. A mitigação do caminho de transmissão entre o cubo de roda e o interior do carro não é trivial, já que as frequências de ressonância das cavidades variam com a velocidade, a carga e a temperatura, exigindo um projeto judicioso de ressonadores passivos (tais como os ressonadores de Helmholtz) [8].

[1] As técnicas desenvolvidas neste trabalho para o cancelamento de ruído são matematicamente equivalentes às técnicas congêneres destinadas ao cancelamento de vibrações; daí pode-se dizer que ambos os problemas serão contemplados.

Métodos passivos sabidamente costumam ser efetivos na supressão de ruído acústico somente em frequências mais elevadas [1]. Desafortunadamente, diversos sinais de ruído que ocorrem na indústria apresentam concentração de energia em baixas frequências [6]. Neste caso, o comprimento de onda do som é tão longo que técnicas passivas não são necessariamente mais competitivas em termos de custo $\times$ benefício, já que demandam material muito espesso ou muito pesado, como os empregados em silenciadores de carros [1].

Esta é uma das razões pela qual esta dissertação se concentra sobre técnicas *ativas* de cancelamento de ruído, as quais tipicamente mobilizam ferramentas de processamento digital de sinais. Tais técnicas normalmente demandam um processo de filtragem, responsável por extrair de um determinado sinal informações de interesse. O componente utilizado para a filtragem, por razões óbvias, é costumeiramente denominado de filtro [9].

Em aplicações tão diversas como sensoriamento remoto, gravação de músicas, biomedicina e metereologia é não raro necessário o registro, o processamento e, por vezes, a emissão de sinais (no caso de exames de ultrassom, por exemplo) [10]. Embora tais sinais comumente sejam contínuos no tempo, doravante assumiremos que eles são discretos, de sorte a serem passíveis de tratamento por um microprocessador. Cumpre notar ser possível, sob determinadas condições de fácil obtenção na prática, efetuar a digitalização de sinais analógicos (isto é, contínuos tanto no tempo quanto na amplitude) de modo a perder pouca informação relevante, mesmo quando tais sinais são reconstruídos no domínio analógico [11]. Os detalhes matemáticos de tal tradução (a qual, no caso de uma conversão analógico-digital, classicamente requer filtros *anti-aliasing*, amostragem e quantização) constituem uma área já madura. Assume-se doravante que seus conceitos essenciais são conhecidos pelo leitor; caso não seja este o caso, recomenda-se a leitura de Mitra [9] e Diniz et al. [11].

Sinais discretos podem ser interpretados como uma sequência, de sorte que $x(k)$ denota a amplitude de que uma sequência x (ou, equivalentemente, que um sinal x) assume no "instante"[2] k. Tais sinais sofrem transformações ao passarem por sistemas, os quais podem ser naturais (no caso da reverberação em uma sala ou dos múltiplos percursos em sistemas de comunicação) ou artificiais (como no caso de compressão de

[2]Cabe ressaltar que nem sempre um sinal armazena uma informação em função do tempo; no caso de imagens, por exemplo, a amplitude é dependente da posição, sendo por conseguinte dependente do espaço, e não do tempo.

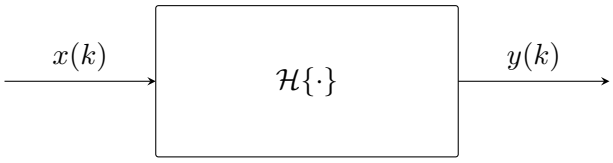


Figura 1 – Representação de um sistema discreto no tempo.

sinais de áudio no padrão conhecido como MP3). Em termos gerais, podemos descrever a modificação de um sinal *de entrada* $x(k)$ como a geração de um sinal de saída $y(k)$ [11]:

$$y(k) = \mathcal{H}\left\{x(k)\right\}, \tag{1}$$

onde o operador $\mathcal{H}\left\{\cdot\right\}$ representa um sistema discreto no tempo (o qual pode emular um sistema contínuo no tempo equivalente), representado na Figura 1. Assim, a transformação pela qual sofre a força aplicada a um freio de carro até se transformar na força aplicada a um determinado componente da roda pode ser modelada como uma modificação engendrada por um determinado sistema (ou operador) em particular.

Este trabalho assume que todos os sistemas físicos (isto é, naturais) envolvidos são lineares. Neste caso, o operador $\mathcal{H}\left\{\cdot\right\}$ da Equação (1) deve satisfazer o princípio da superposição, para quaisquer constantes $a \in \mathbb{R}$ e $b \in \mathbb{R}$ [11]:

$$\mathcal{H}\left\{ax_1(k) + bx_2(k)\right\} = a\mathcal{H}\left\{x_1(k)\right\} + b\mathcal{H}\left\{x_2(k)\right\}, \tag{2}$$

a qual deve ser também válida para qualquer escolha de sequências $x_1(k)$ e $x_2(k)$.

Uma outra característica que assumimos (e que é comumente atendida na prática) é a *invariância no tempo*, a qual matematicamente implica que [11]

$$\mathcal{H}\left\{x(k-k_0)\right\} = y(k-k_0), \tag{3}$$

se $\mathcal{H}\left\{x(k)\right\} = y(k)$. É importante enfatizar que a identidade (3) deve ser válida para todas as possíveis entradas $x(k)$ e para qualquer $k_0 \in \mathbb{Z}$. A grosso modo, ela significa que se atrasarmos uma entrada por k_0 amostras, a saída deverá também será equivalentemente atrasada.

Se definirmos $\delta(k)$ como o impulso unitário, de modo que

$$\delta(k) = \begin{cases} 1, & \text{para } k = 0 \\ 0, & \text{para } k \neq 0 \end{cases}, \tag{4}$$

podemos expressar uma sequência genérica $x(k)$ como uma soma ponderada de impulsos unitários:

$$x(k) = \sum_{n=-\infty}^{\infty} x(n)\delta(k-n), \tag{5}$$

de modo que a saída $y(k)$ resultante de um sistema *discreto linear e invariante no tempo* pode ser expressa por

$$y(k) = \mathcal{H}\left\{\sum_{n=-\infty}^{\infty} x(n)\delta(k-n)\right\} = \sum_{n=-\infty}^{\infty} \mathcal{H}\left\{x(n)\delta(k-n)\right\} = \sum_{n=-\infty}^{\infty} x(n)h(k-n), \tag{6}$$

onde as equações (2) e (3) foram utilizadas e

$$h(k) \triangleq \mathcal{H}\{\delta(k)\} \tag{7}$$

denota a *resposta ao impulso do sistema*. A equação (6) significa que a resposta a *qualquer* sequência entrada de um sistema linear e invariante no tempo cuja resposta ao impulso $h(k)$ seja conhecida pode ser computada por meio da seguinte operação de convolução discreta:

$$y(k) = \sum_{n=-\infty}^{\infty} x(n)h(k-n), \tag{8}$$

de modo que $h(k)$ é suficiente para caracterizar completamente tal tipo de sistema discreto [11].

A Figura 2 apresenta um exemplo de função de transferência medida numa rede digital. Como ela representa a resposta a um impulso unitário, podemos associar o índice k a um atraso na resposta e o coeficiente $h(k)$ ao ganho (possivelmente de magnitude inferior a um, quando o sistema é passivo) referente a este atraso em particular. No caso de uma função de transferência que reflita o fenômeno físico dos múltiplos percursos, temos que os coeficientes de atenuação referentes a atrasos maiores *tendem* a apresentar magnitudes menores, porque um atraso maior corresponde a um percurso mais longo do sinal, o que normalmente implica uma atenuação maior (derivada, por exemplo, de perdas por propagação).

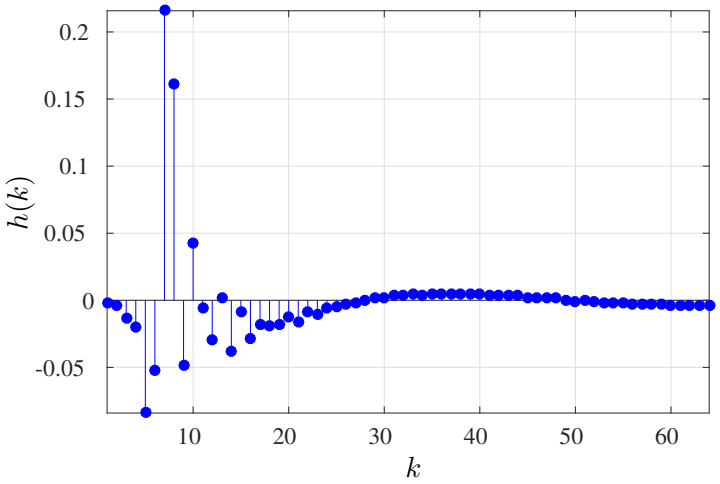


Figura 2 – Exemplo de função de transferência (modelo 1 de [12]).

Sinais de ruído ou de vibração costumam ser extremamente indesejáveis (e, como vimos, possivelmente danosos à saúde). Não raro, é possível efetuar a aquisição de um sinal de excitação $u(k)$ que contenha informação correlata ao distúrbio $d(k)$ que desejamos cancelar. Havendo tal correlação, podemos assumir que existe um mapeamento linear e invariante no tempo que, tendo como entrada $u(k)$, apresente como saída $d(k)$ (vide Figura 3). Como exemplo de tal configuração[3], suponha um repórter falando ao microfone num helicóptero em voo. Se pretendemos eliminar digitalmente o ruído oriundo das hélices do helicóptero, podemos posicionar um microfone próximo a estas hélices, e assumir que o ruído gravado neste microfone passa por uma transformação (para os nossos propósitos, linear e invariante no tempo). Se esta transformação puder ser identificada, seria possível eliminar *digitalmente* a componente ruidosa do registro da fala do repórter, a qual foi provocada por um acoplamento físico (e não por um ambiente microprocessado, no qual se efetuaria a "limpeza" do registro da voz).

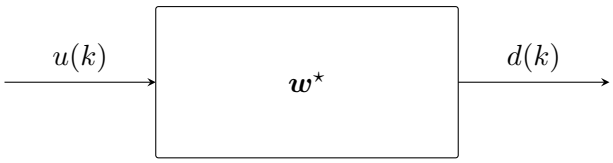


Figura 3 – Exemplo de transformação de um sinal $u(k)$, cujo resultado dá origem a um sinal $d(k)$. Os elementos do vetor $\boldsymbol{w}^\star$ contém os coeficientes da função de transferência (ou resposta ao impulso, cujo suporte é considerado finito) do sistema linear e invariante no tempo que mapeia a sequência $u(k)$ na sequência $d(k)$.

[3] Em termos técnicos, este exemplo se enquadraria no paradigma dos problemas de *identificação de sistemas*, estruturalmente mais simples do que as técnicas de cancelamento de vibração ou de ruído que são objeto deste trabalho. Optamos por este exemplo por motivos didáticos.

A Figura 4 ilustra o diagrama de blocos de um problema de identificação adaptativa de sistemas. Ela é adotada, por exemplo, em aplicações de cancelamento de eco acústico, ubíquas em telecomunicações. Em telefonia viva-voz, por exemplo, existe um acoplamento acústico entre alto-falante e o microfone. Tal acoplamento físico é indesejável, já que a captura pelo microfone do sinal emitido pelo alto-falante gera eco, o que prejudica a comunicação. Entretanto, um sistema digital de processamento tem na prática acesso ao sinal enviado à caixa de som (sinal $u(k)$ na Figura 4), bem como ao sinal gravado no microfone (equivalente à soma $d(k) + \nu(k)$ na Figura 4). Observe que nesta aplicação o vetor desconhecido $\boldsymbol{w}^\star$ armazena os coeficientes referentes à reverberação (ou função de transferência acústica) do ambiente. Tal sistema, supostamente linear, pode ser *artificialmente* emulado (ou seja, replicado num ambiente digital) por um filtro adaptativo $\boldsymbol{w}(k)$, o qual, caso treinado adequadamente, progressivamente se aproxima de $\boldsymbol{w}^{\text{opt}} = \boldsymbol{w}^\star$. Quando tal identificação é bem-sucedida, a saída do filtro tende a ser próxima da componente do sinal registrado no microfone derivada do alto-falante, e o erro (diferença entre o sinal registrado no microfone e a saída do filtro) pode ser utilizado como um sinal "limpo", ou seja, no qual inexiste o *feedback* da caixa de som no microfone. Tal eliminação de *feedback* é uma operação necessária em diversas aplicações de teleconferência, já que é responsável por eliminar o eco, o qual tem o potencial de inviabilizar uma comunicação.

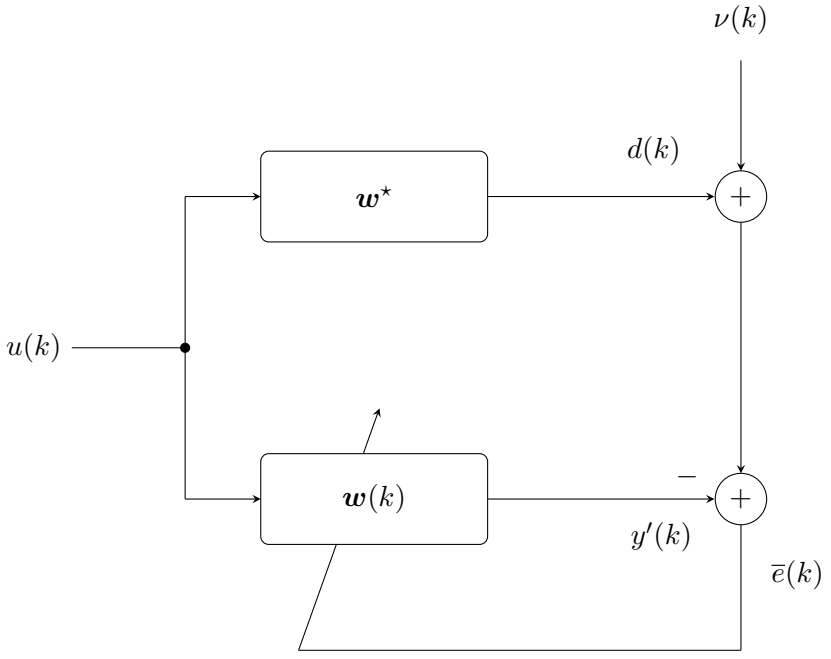


Figura 4 – Diagrama em blocos de uma tarefa de identificação adaptativa de sistemas.

Inicialmente, os coeficientes do filtro $\boldsymbol{w}(k)$ são inicializados como zeros, já que inexiste em geral qualquer informação *a priori* acerca dos coeficientes de $\boldsymbol{w}^{\star}$. Espera-se que, após um treinamento bem executado, a saída do filtro $y(k)$ se aproxime da soma

$$\overline{d}(k) \triangleq d(k) + \nu(k), \tag{9}$$

de modo que o sinal de erro

$$\overline{e}(k) \triangleq \overline{d}(k) - y'(k) \tag{10}$$

avalia indiretamente a qualidade da identificação. Uma magnitude baixa de $\overline{e}(k)$ é portanto um indicativo de que a qualidade da solução atual $\boldsymbol{w}(k)$ é boa. Embora idealmente uma aferição de qualidade mais apropriada seja a discrepância entre $d(k)$ e $y'(k)$, na prática $d(k)$ é corrompido pelo ruído aditivo $\nu(k)$, sendo portanto não observável.

O sinal de erro é uma grandeza observável, já que depende da diferença entre o sinal registrado no microfone $\overline{d}(k)$ e a saída do filtro adaptativo[4]

$$y'(k) \triangleq [\boldsymbol{w}(k)]^T \boldsymbol{u}(k), \tag{11}$$

onde

$$\boldsymbol{w}(k) \triangleq \begin{bmatrix} w_0(k) & w_1(k) & \dots & w_{N-1}(k) \end{bmatrix}^T \tag{12}$$

e

$$\boldsymbol{u}(k) \triangleq \begin{bmatrix} u(k) & u(k-1) & \dots & u(k-N+1) \end{bmatrix}^T. \tag{13}$$

Algoritmos de filtragem adaptativa (foco do próximo capítulo) tradicionalmente utilizam o sinal de erro para iterativamente atualizar o vetor de coeficientes $\boldsymbol{w}(k)$, de modo que

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \Delta\boldsymbol{w}(k), \tag{14}$$

onde a modificação $\Delta\boldsymbol{w}(k)$ é uma função do erro. Um algoritmo clássico para este tipo de aplicação é o LMS (do inglês *least mean squares*), o qual será detalhado no Capítulo 2. Observe que um erro nulo não deve implicar modificações em $\boldsymbol{w}(k)$, pois é um *indicativo* de que a identificação encontra-se com boa qualidade. Cumpre ressaltar que um erro eventualmente nulo não implica rigorosamente uma identificação perfeita, devido a *i*) existência do sinal de ruído $\nu(k)$, o qual contribui para o sinal de erro e não

[4] Todos os vetores ao longo desta dissertação são do tipo coluna.

depende da entrada do filtro (portanto, o filtro é incapaz de prevê-lo); *ii*) o sinal de entrada pode ser, num determinado instante, nulo; neste caso, a saída do filtro independe dos valores do filtro adaptativo; *iii*) o lugar geométrico referente à condição de erro nulo é um hiperplano com infinitas soluções possíveis $\boldsymbol{w}(k)$, pois a condição de erro nulo se traduz numa restrição linear, enquanto o filtro adaptativo pode apresentar na prática milhares de coeficientes (ou seja, variáveis na restrição).

Este trabalho, como vimos, enfoca o problema de cancelamento de ruído ou de vibrações. Tal problema exige um sistema adaptativo estruturalmente mais complexo do que no caso de identificação de sistemas. Seja a Figura 5, a qual apresenta o diagrama em blocos do algoritmo FX-LMS. A primeira diferença que vemos entre a Figura 5 e a Figura 4 é que existe um sistema $\boldsymbol{s}$ que modifica a saída do filtro $y'(k)$, de modo que a componente registrada no sensor (que seria o microfone na aplicação anteriormente descrita) oriunda do filtro seria

$$y'_{\mathrm{f}}(k) = \sum_{i=0}^{M-1} s_i y'(k-i), \tag{15}$$

onde s_i (para $i \in \{0, 1, \ldots, M-1\}$) denota o i-ésimo coeficiente da chamada planta secundária (ou, mais concisamente, do secundário), a qual apresenta comprimento $M \in \mathbb{N}$. A função de transferência do secundário deriva do fato de que na prática a saída do filtro passa por transformações físicas antes de atuar no sensor. Como exemplo, seja o caso em que um alto-falante deva cancelar um ruído existente num determinado ponto relativamente distante, no qual existe um microfone. Ora, o sinal gravado no microfone, assumindo por simplicidade que a única fonte sonora no ambiente seja o alto-falante, é uma versão filtrada do som emitido por este. Afinal, como vimos, há uma reverberação no ambiente e, mesmo na ausência de reverberação, existiria pelo menos um atraso, devido ao fato de a velocidade do som não ser infinita.

O algoritmo FX-LMS pode ser entendido como uma adaptação do LMS de modo a lidar com o fato de que agora (diferentemente do caso de identificação de sistemas), a componente do sinal gravado no sensor oriunda do filtro adaptativo depende de uma planta secundária $\boldsymbol{s}$. Sem entrar em maiores detalhes nesta introdução, o algoritmo FX-LMS compensa a modificação induzida pelo secundário *reproduzindo tal modificação no sinal de entrada* utilizado pelo algoritmo LMS, conforme pode ser visto na Figura 5. Observe que tal compensação é feita por uma estimativa $\hat{s}$ da função de transferência

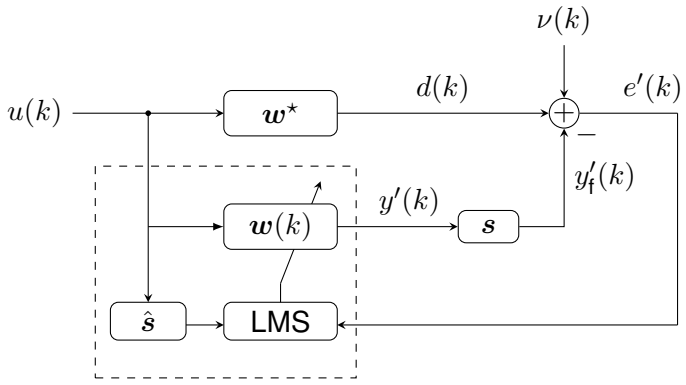


Figura 5 – Diagrama em blocos do algoritmo FX-LMS.

do secundário $\boldsymbol{s}$, cabendo a tal estimativa emular o efeito do secundário. Ademais, na configuração da Figura 5, a solução ótima para $\boldsymbol{w}(k)$ não é mais $\boldsymbol{w}^\star$, como era no caso de identificação de sistemas (representado pela Figura 4). Isso se deve ao fato de que $\boldsymbol{w}(k)$ encontra-se em cascata com secundário, de modo que $\boldsymbol{w}(k)$ também deve reverter o impacto do secundário (ou, equivalentemente, incorporar a uma transformação inversa à transformação engendrada pelo secundário). Assim, no caso de cancelamento de ruído, a solução ótima $\boldsymbol{w}^{\mathrm{opt}}$ para o vetor de coeficientes adaptativos não é mais $\boldsymbol{w}^\star$, diferentemente da configuração de identificação de sistemas.

Duas abordagens diferentes são costumeiramente empregadas para se estimar o secundário [13]. A primeira delas combina a injeção de um ruído aleatório branco no sistema de cancelamento ativo de ruído com algum algoritmo de identificação de sistemas [14]–[17], enquanto a segunda estima o secundário diretamente da saída do filtro de controle [18]. De todo modo, o método particular pelo qual o secundário é estimado é alheio ao escopo deste trabalho. Cumpre notar que quaisquer imperfeições nesta estimativa se refletirão na qualidade do processo de cancelamento de vibrações.

1.1- Objetivos deste Trabalho

A discussão pregressa mobilizou concisamente os conceitos necessários para a compreensão do principal objetivo deste trabalho. Ele deriva de uma constatação: embora na prática os sinais de ruído sejam ubíquos, na literatura pertinente assume-se a existência de ruído apenas no sensor, embora na prática este ruído (oriundo de

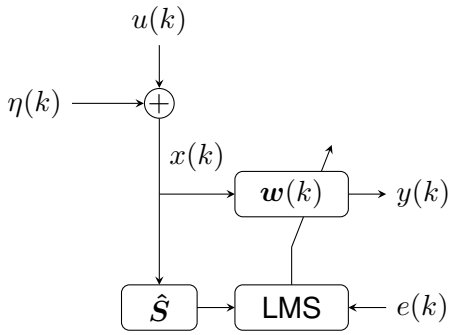


Figura 6 – Ilustração da configuração analisada neste trabalho, na qual o sinal de entrada do filtro é corrompido por um ruído aditivo $\eta(k)$.

interferências ou mesmo de quantização inerente ao processo de aquisição) também exista no sinal de excitação do filtro adaptativo. A incorporação deste sinal de ruído na estrutura do algoritmo é mostrada na Figura 6, a qual introduz os seguintes sinais:

$$x(k) \triangleq u(k) + \eta(k), \tag{16}$$

$$\boldsymbol{x}(k) \triangleq \begin{bmatrix} x(k) & x(k-1) & \dots & x(k-N-1) \end{bmatrix}^T, \tag{17}$$

$$y(k) \triangleq [\boldsymbol{w}(k)]^T \boldsymbol{x}(k), \tag{18}$$

$$y_{\mathrm{f}}(k) \triangleq \sum_{i=0}^{M-1} s_i y(k-i), \tag{19}$$

$$e(k) \triangleq \overline{d}(k) - y_{\mathrm{f}}(k). \tag{20}$$

Esta dissertação tem como foco a configuração de entrada apresentada na Figura 6, na qual o algoritmo adaptativo tem acesso a uma versão corrompida $x(k)$ do sinal de entrada $u(k)$, o qual encontra-se somado a um ruído $\eta(k)$. Dentre os diversos algoritmos contemplados nesta dissertação, os três principais são:

1. FX-LMS: trata-se do algoritmo adaptativo clássico de cancelamento de ruído;

2. N-FX-LMS: do inglês *Noisy FX-LMS*, consiste na adoção do algoritmo FX-LMS num contexto em que o sinal de entrada encontra-se corrompido por um ruído $\eta(k)$. Observe que quando inserimos no modelo estrutural do algoritmo FX-LMS um ruído aditivo na entrada, teremos como resultado o algoritmo N-FX-LMS. Demonstraremos neste trabalho que a existência do ruído $\eta(k)$ introduz um viés na estimativa do algoritmo N-FX-LMS;

3. BC-FX-LMS: o algoritmo FX-LMS com compensação de viés é *o algoritmo proposto*

nesta dissertação, o qual atenua o impacto na compensação de viés oriunda do ruído $\eta(k)$, responsável por corromper a entrada do filtro adaptativo.

Esta dissertação apresenta como objetivos principais:

1. Modelar o impacto do ruído $\eta(k)$ no processo de estimativa do algoritmo N-FX-LMS;
2. Propor uma variante do algoritmo N-FX-LMS (intitulada BC-FX-LMS) capaz de mitigar o impacto da presença de $\eta(k)$.

Nesta dissertação, é demonstrado que o viés introduzido por $\eta(k)$ é dependente da variância $\sigma_\eta^2(k)$ do ruído $\eta(k)$. Ademais, verificaremos que nossa variante proposta (isto é, o algoritmo BC-FX-LMS) requer uma estimativa da variância σ_η^2 do sinal de ruído $\eta(k)$. Logo, o *objetivo secundário* desta dissertação consiste em avaliar a sensibilidade do algoritmo proposto perante imprecisões na estimativa de tal variância.

1.2- Metodologia

Tanto a modelagem do impacto do ruído $\eta(k)$ quanto a proposição de uma variante do algoritmo FX-LMS dependem do desenvolvimento de uma análise estocástica que preveja o comportamento de aprendizado do algoritmo. Devido à dificuldade envolvida na análise do FX-LMS (bem como de seus derivados), este trabalho se concentrará sobre uma análise de primeira ordem (isto é, que utiliza apenas estatísticas de primeira ordem dos coeficientes adaptativos) deste algoritmo. Tal escolha metodológica implica em uma restrição: devido ao fato de nos limitarmos a estatísticas de primeira ordem (e portanto não podermos inferir nada a respeito da variância dos coeficientes adaptativos, já que esta é uma estatística de segunda ordem), não será possível obter adequadamente uma região teórica para o passo de aprendizagem que garanta estabilidade do algoritmo.

Também para efeito de simplificação na complexidade das equações, serão utilizadas algumas hipóteses de independência estatística, a saber:

- Hipótese da independência (HI): também conhecida como heurística da independência [19], ela preconiza independência estatística entre as amostras do sinal de excitação $x(k)$ e os coeficientes adaptativos;

- Hipótese concernente aos sinais de ruído (HR): esta hipótese assume que os sinais de ruído $\eta(k)$ e $\nu(k)$ são estatisticamente independentes entre si, assim como dos demais processos estocásticos.

Estas hipóteses são ubíquas na literatura e serão descritas em maior detalhe no momento apropriado.

A análise proposta dá origem a estimativas teóricas *médias* de desempenho do algoritmo. Para aferir a qualidade destas estimativas, os algoritmos serão implementados num ambiente simulado. Cada execução do algoritmo gera uma curva de aprendizado ruidosa, a qual depende da realização em particular dos processos estocásticos envolvidos. Tal curva ruidosa não pode ser comparada com os resultados oriundos da análise teórica, já que estes se referem ao comportamento *médio* do algoritmo, e não ao relativo a uma execução deste em particular. Por essa razão, diversas curvas empíricas (isto é, simuladas) são executadas de modo independente um determinado número de vezes. Após o cômputo da média destas curvas simuladas, teremos uma curva média empírica, a qual pode ser cotejada com o modelo teórico. Este procedimento de média das curvas simuladas é conhecido como *média de Monte Carlo*.

1.3- Estrutura

Esta dissertação é estruturada do seguinte modo: o Capítulo 2 apresenta o algoritmo LMS, considerado o padrão na área de filtragem adaptativa. O Capítulo 3 descreve algumas técnicas de análise deste algoritmo. O Capítulo 4 introduz o algoritmo FX-LMS. O Capítulo 5, apresenta a técnica de compensação de viés. O capítulo 6 é a inovação deste trabalho, ele introduz o algoritmo N-FX-LMS num contexto no qual o sinal de entrada é corrompido por ruído e modela o impacto deste ruído na qualidade do estimador. Ademais, este capítulo propõe um método de compensação de viés do algoritmo N-FX-LMS e avalia sua sensibilidade com relação à falta de acurácia com relação à variância do ruído $\eta(k)$. O Capítulo 7 apresenta as simulações feitas e discute os resultados. O Capítulo 8 encerra este trabalho, apresentando suas conclusões e trabalhos futuros considerados promissores.

1.4- Notação

Essa dissertação trata de diferentes algoritmos, como por exemplo o LMS, o FX-LMS, o N-FX-LMS e o BC-FX-LMS. Cada um destes algoritmos apresenta sinais cuja definição matemática é ligeiramente diferente, o que exige um cuidado trabalho de notação para evitar ambiguidades.

Ao longo desta dissertação, escalares serão escritos em minúsculas, vetores em letras minúsculas e em negrito e matrizes em letras maiúsculas e em negrito. O operador $(\cdot)^T$ denota transposição, enquanto $\mathbb{E}[\cdot]$ denota o operador de valor esperado. O símbolo $\nabla_{\boldsymbol{w}} x$ denota o gradiente de x com relação ao vetor $\boldsymbol{w}$ e $\|\boldsymbol{x}\|^2 \triangleq \boldsymbol{x}^T\boldsymbol{x}$ significa a norma quadrática do vetor $\boldsymbol{x}$.

1.5- Publicações associadas a esta dissertação

Esta dissertação foi responsável por gerar, direta ou indiretamente, dois artigos.

- *Publicação em revista internacional*

  Lucas P. R. da Silva, J. V. G. de Souza, Jefferson C. de Paula, Amaro A. de Lima, Diego B. Haddad. "*Bias-compensated FX-LMS algorithm*", *Electronics Letters*, vol. 56, n. 24, pp. 1340-1343, Nov. 2020.

- *Publicação em congresso*

  Rodrigo M. S. Pimenta, Leonardo C. Resende, **Lucas P. R. da Silva**, Newton N. Siqueira, Jurair R. P. Junior, Mariane R. Petraglia, Diego B. Haddad. "*A Block-Sparsity Induced NLMS Algorithm with Bias Compensation*", SBrT (Simpósio Brasileiro de Telecomunicações e Processamento de Sinais), Novembro, 2020.

2- Algoritmo Least Mean Squares

2.1- Introdução

Em ocasiões onde as especificações de processamento não são totalmente conhecidas, filtros cujos parâmetros são fixos não costumam obter um desempenho adequado, sendo nestes casos indicada a adoção de técnicas de processamento *adaptativo* de sinais, entre as quais não raro se destacam os *filtros adaptativos* [20] [21].

Um filtro adaptativo consiste em um dispositivo não linear que utiliza um algoritmo recursivo para atualizar seus coeficientes a cada iteração. A despeito de sua estrutura de processamento poder ser linear quando seus coeficientes são fixos, o fato de estes se alterarem dinamicamente implica que a saída do filtro não respeita o princípio do superposição (vide (2)), de modo que, em termos técnicos, os algoritmos de filtragem adaptativa efetuam uma transformação não linear nos dados.

A atualização engendrada por algoritmos de filtragem adaptativa é normalmente orientada pela minimização de uma função custo, a qual normalmente se vale de um sinal de erro, o qual reflete uma discrepância entre o sinal de referência e o sinal desejado (assumindo um contexto supervisionado) [22] [23]. Tais algoritmos possuem diversas aplicações, dentre as quais podemos destacar [24] [25]:

- Identificação de Sistemas
- Predição de Sinais
- Cancelamento de Eco Acústico
- Equalização de Canais
- Controle Ativo de Ruído

A Figura 7 apresenta uma estrutura clássica de um filtro adaptativo, onde $u(k)$ é o sinal de entrada, $y'(k)$ é a saída do filtro adaptativo, $d'(k)$ é o sinal desejado (ou de referência) e $\overline{e}(k)$ é o sinal de erro [21].

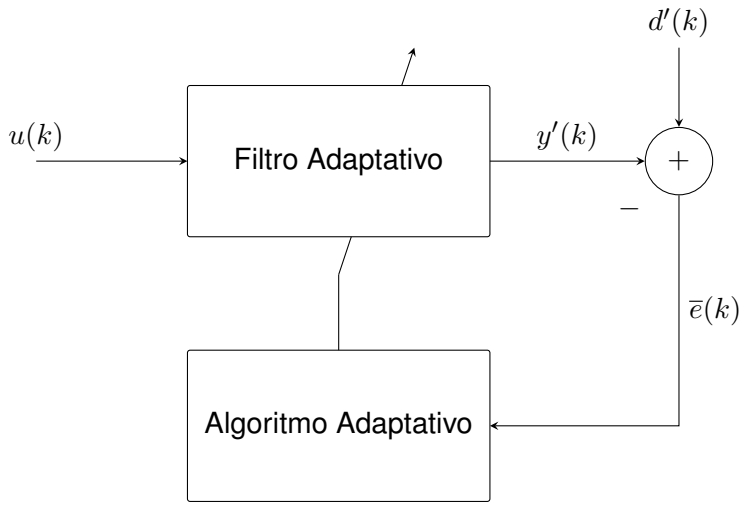


Figura 7 – Estrutura típica de um filtro adaptativo.

A estrutura utilizada no presente trabalho é a de um filtro transversal, também conhecido como de resposta ao impulso finita (Finite Impulse Response (FIR)), cujos coeficientes, na k-ésima iteração, encontram-se armazenados no vetor $\boldsymbol{w}(k)$. Assim,

$$y'(k) \triangleq [\boldsymbol{w}(k)]^T \boldsymbol{u}(k), \tag{21}$$

onde

$$\boldsymbol{u}(k) \triangleq \begin{bmatrix} u(k) & u(k-1) & \dots & u(k-N+1) \end{bmatrix}. \tag{22}$$

Nas últimas décadas, foram desenvolvidos diversos algoritmos adaptativos, sendo os mais tradicionais o *Least Mean Square (LMS)*, o *Normalized Least Mean Square (NLMS)* e o *Recursive Least Square (RLS)* [26] [27]. O LMS é um dos mais populares entre eles, em especial por sua baixa complexidade computacional e robustez [21]. Este capítulo se concentra sobre o algoritmo LMS, num contexto de identificação de sistemas.

2.2- O Algoritmo Gradiente Descendente

O LMS foi proposto em 1960 por *Bernard Widrow* [28]. Ele pode ser considerado uma aproximação estocástica do algoritmo gradiente descendente, descrito nesta seção.

Assumindo um contexto supervisionado, o LMS se vale do sinal de erro definido

por

$$\overline{e}(k) \triangleq d'(k) - y'(k), \tag{23}$$

de modo que o quadrado do erro, *para um filtro fixo* $\boldsymbol{w}$, pode ser expresso por

$$\overline{e}^2(k) = \left[d'(k)\right]^2 - 2d'(k)\left[\boldsymbol{u}(k)\right]^T \boldsymbol{w} + \boldsymbol{w}^T \left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\right] \boldsymbol{w}. \tag{24}$$

Aplicando-se o valor esperado em (24) e assumindo estacionariedade dos sinais envolvidos, obtemos o seguinte erro quadrático médio (Mean Square Error (MSE)) para a configuração da Figura 7:

$$\text{MSE} \triangleq \mathbb{E}\left[\overline{e}^2(k)\right] = \sigma_{d'}^2 - 2\left[\boldsymbol{p}'\right]^T \boldsymbol{w} + \boldsymbol{w}^T \boldsymbol{R}_u \boldsymbol{w}, \tag{25}$$

onde $\sigma_{d'}^2 \triangleq \mathbb{E}\left[d'(k)\right]^2$, $\boldsymbol{p}' \triangleq \mathbb{E}\left\{d'(k)\left[\boldsymbol{u}(k)\right]\right\}$ e $\boldsymbol{R}_u \triangleq \mathbb{E}\left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\right]$ denota a *matriz de autocorrelação* do sinal de entrada $u(k)$.

O MSE (Equação (25)) pode ser minimizado igualando-se ao vetor nulo seu gradiente com relação a $\boldsymbol{w}$, dando origem a

$$-2\boldsymbol{p}' + 2\boldsymbol{R}_u \boldsymbol{w} = \boldsymbol{0} \tag{26}$$

$$\Rightarrow \boldsymbol{w}^{\text{opt}} = \boldsymbol{R}_u^{-1} \boldsymbol{p}', \tag{27}$$

onde a solução ótima $\boldsymbol{w}^{\text{opt}}$ é conhecida como *solução de Wiener* [22].

Note que se o filtro adaptativo tiver o mesmo comprimento de $\boldsymbol{w}^{\text{opt}}$ e admitirmos que o sinal de referência é gerado consoante o seguinte modelo afim nos parâmetros:

$$d'(k) = \left[\boldsymbol{w}^\star\right] \boldsymbol{u}(k) + \nu(k), \tag{28}$$

então, utilizando-se (27), temos

$$\boldsymbol{w}^{\text{opt}} = \left\{\mathbb{E}\left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\right]\right\}^{-1} \left\{\mathbb{E}\left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\right] \boldsymbol{w}^\star + \mathbb{E}\left[\nu(k)\boldsymbol{u}(k)\right]\right\}, \tag{29}$$

a qual pode ser simplificada caso recorramos à hipótese HR, a qual assume que o ruído de média zero $\nu(k)$ é estatisticamente independente das demais variáveis aleatórias, dando origem a

$$\boldsymbol{w}^{\text{opt}} = \boldsymbol{w}^\star, \tag{30}$$

significando que, no caso de identificação de sistemas apresentado na Figura 18, a solução ótima é aquela na qual o filtro adaptativo coincide com $\boldsymbol{w}^{\star}$.

Infelizmente, a solução de Wiener (27), além de ser custosa computacionalmente (observe que ela exige uma inversão matricial), assume implicitamente que a relação entrada e saída é fixa. Portanto, ela é incapaz de rastrear modificações nesta relação.

Seu custo computacional pode ser mitigado caso recorramos a um método iterativo baseado em gradiente, de modo que a nova estimativa $\boldsymbol{w}(k+1)$ seja oriunda de uma modificação da estimativa $\boldsymbol{w}(k)$ do seguinte modo:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) - \beta \nabla_{\boldsymbol{w}(k)} \left\{ \mathsf{MSE}[\boldsymbol{w}(k)] \right\}, \tag{31}$$

onde $\beta \in \mathbb{R}_{+}$ é um passo (ou fator de aprendizagem), o qual é responsável por recolher uma fração do gradiente $\nabla_{\boldsymbol{w}(k)} \left\{ \mathsf{MSE}[\boldsymbol{w}(k)] \right\}$ (pois estamos mais interessados na sua direção do que em sua magnitude) e o sinal de negativo significa que estamos procurando minimizar o erro quadrático médio. Como o gradiente aponta para a direção de subida mais íngreme, cumpre inverter o seu sinal para que o método efetivamente minimize a função custo (o MSE).

Utilizando-se (26), a equação recursiva pode ser escrita como:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) - 2\beta \boldsymbol{R}_u \boldsymbol{w}(k) + 2\beta \boldsymbol{p}', \tag{32}$$

equação que descreve o algoritmo gradiente descendente ou algoritmo de descida mais íngreme. Este algoritmo, a depender da escolha de β, faz com que $\boldsymbol{w}(k)$ progressivamente se aproxime de $\boldsymbol{w}^{\mathsf{opt}}$, sem exigir inversões de matrizes.

Entretanto, o algoritmo de descida mais íngreme ainda conserva desvantagens importantes, tal como a necessidade de se estimar $\boldsymbol{R}_u$ e $\boldsymbol{p}'$ ao longo de uma sequência de dados disponíveis (o que tem um custo computacional não desprezível) e de assumir que a solução ótima é estacionária ao longo de toda a base de dados. O algoritmo LMS, visto na próxima seção, mitiga estas desvantagens.

2.3- LMS

O algoritmo LMS pode ser interpretado como uma implementação estocástica do algoritmo de descida mais íngreme, a qual se vale de duas aproximações:

$$\boldsymbol{R}_u \approx \boldsymbol{u}(k)\boldsymbol{u}^T(k) \tag{33}$$

e

$$\boldsymbol{p}' \approx d'(k)\boldsymbol{u}(k), \tag{34}$$

de modo que a atualização do algoritmo de descida mais íngreme (32) se converte em[1]

$$\begin{aligned} \boldsymbol{w}(k+1) &= \boldsymbol{w}(k) - \beta\boldsymbol{u}(k)\boldsymbol{u}^T(k)\boldsymbol{w}(k) + \beta d'(k)\boldsymbol{u}(k) & (35)\\ &= \boldsymbol{w}(k) + \beta\left[d'(k) - y'(k)\right]\boldsymbol{u}(k), & (36) \end{aligned}$$

o que dá origem à equação de atualização do LMS:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta\overline{e}(k)\boldsymbol{u}(k), \tag{37}$$

na qual o fator de aprendizagem β tipicamente é empregado para ajustar a relação entre velocidade de convergência, estabilidade e desempenho em regime permanente [29]. O algoritmo LMS pode ser considerado um caso particular do algoritmo *backpropagation* quando a estrutura da rede neural é linear.

Existem outras formas de se derivar o algoritmo LMS, capazes de elucidar algum aspecto do processo de aprendizagem deste algoritmo. Uma derivação mais simples do que a pregressa utiliza o gradiente do erro quadrático instantâneo:

$$\begin{aligned} \boldsymbol{w}(k+1) &= \boldsymbol{w}(k) - \beta\nabla_{\boldsymbol{w}(k)}\left[\frac{1}{2}\overline{e}^2(k)\right] = \boldsymbol{w}(k) - \beta\overline{e}(k)\nabla_{\boldsymbol{w}(k)}\left[\overline{e}(k)\right]\\ &= \boldsymbol{w}(k) - \beta\overline{e}(k)\nabla_{\boldsymbol{w}(k)}\left[d'(k) - \boldsymbol{w}^T(k)\boldsymbol{u}(\boldsymbol{k})\right] = \boldsymbol{w}(k) + \beta\overline{e}(k)\boldsymbol{u}(k). \quad (38) \end{aligned}$$

Outra derivação o compreende como o solucionador do seguinte problema deter-

[1] Observe que um fator multiplicativo de 2 foi retirado da equação, de modo a torná-la compatível com a definição mais comum do LMS. Como o fator 2 está multiplicado pelo parâmetro ajustável β, pode-se supor que foi por ele absorvido.

minístico e local (já que se utiliza apenas de informação acessível no instante k) [30]:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}\left[\boldsymbol{w}(k+1)\right] \triangleq \frac{1}{2}\|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2 \text{ sujeito a } e_{\mathsf{p}}(k) = (1 - \beta\|\boldsymbol{u}(k)\|^2)\overline{e}(k), \quad (39)$$

onde $e_{\mathsf{p}}(k)$ denota o erro *a posteriori*, ou seja o erro medido no par de dados $\{d'(k), \boldsymbol{u}(k)\}$ *após* a atualização ser efetuada neste mesmo par de dados, ou:

$$e_{\mathsf{p}}(k) \triangleq d'(k) - \boldsymbol{w}^T(k+1)\boldsymbol{u}(k). \quad (40)$$

Utilizando-se a técnica dos multiplicadores de Lagrange, o problema de otimização (39) pode ser reescrito como o seguinte problema equivalente *sem restrições*:

$$\min_{\boldsymbol{w}(k+1)} \frac{1}{2}\|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2 + \lambda\left[e_{\mathsf{p}}(k) - (1 - \beta\|\boldsymbol{u}(k)\|^2)\overline{e}(k)\right], \quad (41)$$

onde $\lambda \in \mathbb{R}$ é o multiplicador de Lagrange. A minimização (39) é efetuada zerando-se seu gradiente com relação a $\boldsymbol{w}(k+1)$:

$$\boldsymbol{w}(k+1) - \boldsymbol{w}(k) - \lambda\boldsymbol{u}(k) = \boldsymbol{0}, \quad (42)$$

ou:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \lambda\boldsymbol{u}(k). \quad (43)$$

Aplicando-se (43) na restrição de (39), encontramos:

$$\lambda = \beta\overline{e}(k). \quad (44)$$

Operação	Número de adições	Número de multiplicações
Cálculo de $y'(k) = [\boldsymbol{w}(k)]^T\boldsymbol{u}(k)$	$N-1$	N
Cálculo de $\overline{e}(k)$	1	0
Cálculo de $\beta\overline{e}(k)$	0	1
Cálculo de $\beta\overline{e}(k)\boldsymbol{u}(k)$	0	N
Cálculo de $\boldsymbol{w}(k+1)$	N	0
Total	$2N$	$2N+1$

Tabela 1 – Complexidade computacional envolvida em uma atualização do algoritmo LMS. A complexidade de uma subtração é considerada equivalente à de uma adição.

Aplicando-se (44) em (43), temos o LMS, cuja complexidade computacional é proporcional ao tamanho N do filtro, conforme explicitado na Tabela 1. O pseudo-código do LMS é apresentado no Algoritmo 1.

Algorithm 1 Pseudo-código do Algoritmo LMS

Entrada: β

Dados: $\{u(k), d'(k)\}$

Resultado: $\boldsymbol{w}(K+1)$

Inicialização: $\boldsymbol{w}(0) = \mathbf{0}$, $k = 0$

while $k \leq K$ **do**

$y'(k) \leftarrow [\boldsymbol{w}(k)]^T \boldsymbol{u}(k)$

$\overline{e}(k) \leftarrow d'(k) - y'(k)$

$\boldsymbol{w}(k+1) \leftarrow \boldsymbol{w}(k) + \beta \boldsymbol{u}(k)\overline{e}(k)$

$k \leftarrow k + 1$

end

2.4- Simulações

Por razões didáticas, um experimento foi implementado com o LMS, utilizando como função de transferência a resposta ao impulso apresentada na Figura 2, com $N = 64$. O ruído aditivo $\nu(k)$ é gaussiano, branco e de variância $\sigma_\nu^2 = 10^{-6}$. O sinal de entrada consiste na realização de um processo aleatório branco, gaussiano e de variância unitária. Cada curva empírica foi calculada como a média de 1000 ensaios independentes de Monte Carlo.

Como métricas de aferição de desempenho, classicamente se utiliza o MSE ou a distorção quadrática média (Mean Square Deviation (MSD)), definidos como:

$$\text{MSE}(k) \triangleq \mathbb{E}\left[e^2(k)\right], \tag{45}$$

$$\text{MSD}(k) \triangleq \mathbb{E}\left[\|\boldsymbol{w}^{\text{opt}} - \boldsymbol{w}(k)\|^2\right]. \tag{46}$$

Normalmente, tais métricas de aferição de desempenho são apresentadas em dB:

$$\text{MSE}(k)\ \text{(dB)} \triangleq 10\log_{10}\left\{\mathbb{E}\left[e^2(k)\right]\right\}, \tag{47}$$

$$\text{MSD}(k)\ \text{(dB)} \triangleq 10\log_{10}\left\{\mathbb{E}\left[\|\boldsymbol{w}^{\text{opt}} - \boldsymbol{w}(k)\|^2\right]\right\}. \tag{48}$$

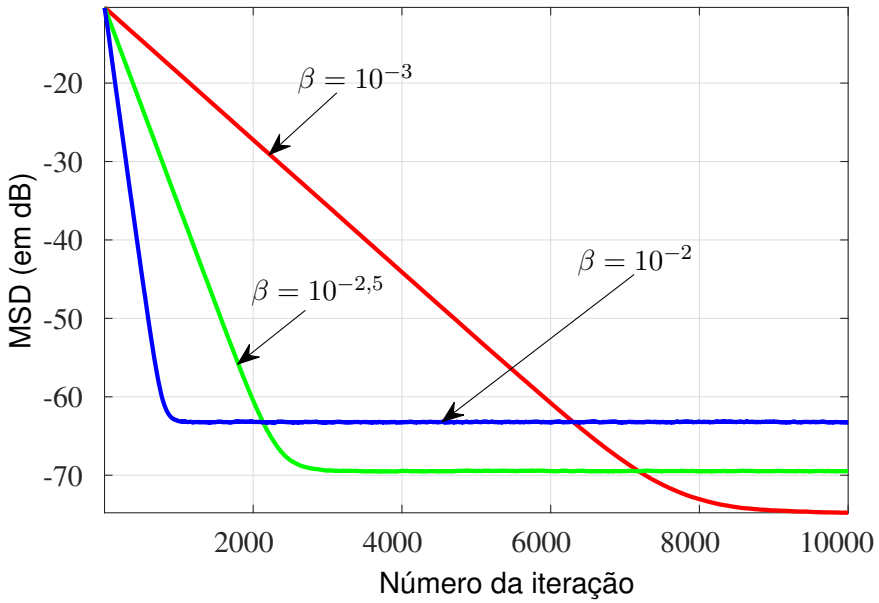


Figura 8 – Evolução do MSD (em dB) ao longo das iterações, para $\beta \in \{10^{-3}, 10^{-2,5}, 10^{-2}\}$.

A Figura 8 apresenta a evolução do MSD para a configuração descrita, para três diferentes valores de β[2]. Observe que valores maiores de β implicam uma taxa de convergência mais elevadas, porém implicam um desempenho pior em regime permanente, porque aumenta a variância dos coeficientes adaptativos em torno da solução ótima. Tal compromisso é incontornável quando se adota um β fixo, podendo ser mitigado quando se permite que β varie ao longo das iterações [31], [32]. Observações similares podem ser feitas com relação ao MSE, cuja curvas de evolução são apresentadas na Figura 9.

Por fim, cabe ressaltar que existe um valor máximo de β para o qual o algoritmo converge; caso este limiar seja violado, o algoritmo pode divergir. A divergência é uma situação catastrófica, na qual tanto o MSD quanto o MSE crescem indefinidamente. Este

[2]No decorrer desta dissertação foram utilizados diferentes valores para determinados parâmetros como : β, número de iterações e médias de Monte Carlo. A escolha destes valores está associada, caso seja utilizado um baixo valor de β o algoritmo leva mais tempo para atingir o regime permanente, logo será necessário um maior número de iterações, assim como um alto valor de β converge de maneira mais rápida porém de maneira mais ruidosa, portanto se faz necessário valores maiores de média de Monte Carlo para se obter uma curva suavizada.

valor máximo de β, para o caso do LMS, depende principalmente de características estatísticas do sinal e de N, o comprimento do filtro adaptativo.

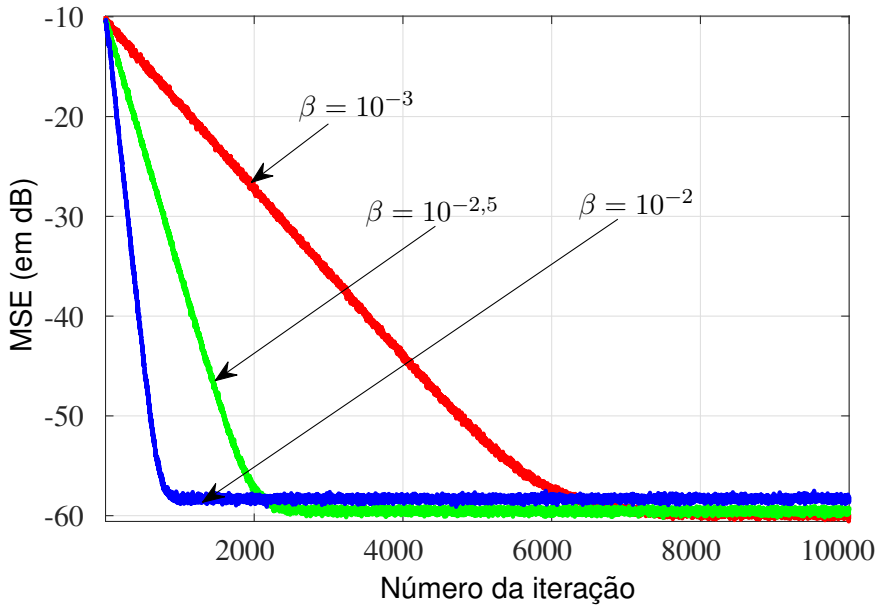


Figura 9 – Evolução do MSE (em dB) em ao longo das iterações, para $\beta \in \{10^{-3}, 10^{-2,5}, 10^{-2}\}$.

3- Análise de Primeira Ordem do LMS

As propriedades de aprendizado dos diversos algoritmos de filtragem adaptativa são dependentes de características estatísticas dos sinais envolvidos, bem como de parâmetros ajustáveis pelo projetista [33], [34].

Para que tais algoritmos possam ser implementados, cabe fornecer ao projetista garantias concernentes quanto à tendenciosidade do algoritmo, à robustez, ao desempenho em regime permanente, à estabilidade, à taxa de convergência e à capacidade de rastreamento [35]. Tais garantias são normalmente oferecidas por modelos estocásticos mais diversos, os quais são capazes de prever as habilidades de aprendizado destes algoritmos [36]–[38].

A grosso modo, podemos dividir as análises estocásticas dos algoritmos de filtragem adaptativa em duas modalidades: análises de estatísticas de primeira ordem e de segunda ordem. As análises de primeira ordem são importantes principalmente porque permitem prever se o algoritmo apresenta ou não tendenciosidade em regime permanente. Neste trabalho, estamos essencialmente interessados em efetuar análises de primeira ordem.

Neste capítulo, a técnica de análise de primeira ordem será concisamente introduzida, para situar o leitor dentro deste vasto campo. Como o algoritmo LMS é o mais intensamente estudado (já que é o mais clássico dos algoritmos de filtragem adaptativa), este capítulo contemplará a análise deste algoritmo, cujo diagrama em blocos (para o caso de identificação de sistemas) é ilustrado pela Figura 10.

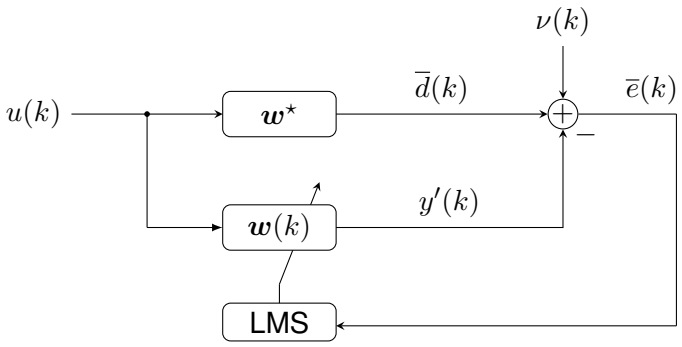


Figura 10 – Diagrama em blocos do algoritmo LMS.

Em geral, as análises estocásticas de algoritmos de filtragem adaptativa se valem, como um primeiro passo, de determinadas hipóteses acerca dos sinais que serão processados pelo filtro adaptativo. A próxima etapa normalmente consiste na construção de equações a diferenças que relacionam recursivamente determinadas variáveis de interesse no instante $k+1$ a valores que estas mesmas variáveis assumiram no instante k. Após a construção destas recursões, costumeiramente aplica-se o operador de valor esperado $\mathbb{E}[\cdot]$, o qual leva em consideração todas as possíveis realizações dos processos estocásticos envolvidos. A aplicação de tal valor esperado é costumeiramente modulada por *hipóteses estatísticas*, as quais apresentam o potencial de simplificar sobremaneira a complexidade das equações resultantes. Por fim, o resultado é analisado para avaliação, entre outras propriedades, da taxa de convergência, da estabilidade ou da tendenciosidade assintótica do algoritmo.

3.1- Hipóteses Estatísticas

A hipótese mais utilizada na simplificação dos cálculos dos modelos estocásticos é conhecida como *hipótese da independência* (HI), também recentemente denominada *heurística da independência* [19]. Popular no âmbito das aproximações estocásticas [39], ela assume que os vetores $\boldsymbol{u}(k_1)$ e $\boldsymbol{u}(k_2)$ são independentes se $k_1 \neq k_2$. Embora conveniente para simplificação dos modelos, tal hipótese é claramente violada na prática, já que a estrutura transversal de processamento induz uma *coerência determinística* entre vetores de entrada consecutivos[1]. Por exemplo, para o caso $N=6$, os vetores $\boldsymbol{u}(k-2)$, $\boldsymbol{u}(k-1)$ e $\boldsymbol{u}(k)$ são

$$\begin{aligned}
\boldsymbol{u}(k-2) &= [\ \underline{u(k-2)}\ \ \underline{u(k-3)}\ \ \underline{u(k-4)}\ \ \underline{u(k-5)}\ \ \underline{u(k-6)}\ \ u(k-7)\ ], \\
\boldsymbol{u}(k-1) &= [\ \underline{u(k-1)}\ \ \underline{u(k-2)}\ \ \underline{u(k-3)}\ \ \underline{u(k-4)}\ \ \underline{u(k-5)}\ \ \underline{u(k-6)}\ ], \\
\boldsymbol{u}(k) &= [\ u(k)\ \ \underline{u(k-1)}\ \ \underline{u(k-2)}\ \ \underline{u(k-3)}\ \ \underline{u(k-4)}\ \ \underline{u(k-5)}\ ],
\end{aligned}$$

onde se pode observar a emergência de diversas amostras de idêntico valor (assinaladas com sublinhado), o que claramente viola a plausibilidade física da hipótese da independência.

[1] Em contextos nos quais a estrutura transversal atua na dimensão espacial (isto é, ao longo de diversos sensores) e não na temporal e os sinais de entrada são iid (isto é, independentes e identicamente distribuídos), a hipótese da independência não é violada. Tal é o caso de um cenário de comunicação síncrona multiusuário [40].

Não obstante o fato de não ser factível, a hipótese da independência frequentemente consiste numa aproximação adequada para descrever o processo de aprendizagem, notadamente quando o passo de aprendizagem β assume valores pequenos [34], porque neste caso o estimador adaptativo atua como um filtro passa baixas [41]. Assim, os coeficientes adaptativos $w_i(k)$ apresentam uma dinâmica muito mais lenta do que a do sinal de entrada[2], podendo-se assumir que são praticamente independentes das amostras mais recentes do sinal de entrada.

Outra hipótese que nos é de interesse é relativa ao ruído $\nu(k)$:

Hipótese referente ao ruído (HR): as amostras do ruído aditivo $\nu(k)$, cuja média é zero, são estatisticamente independentes entre si, assim como estatisticamente independentes do sinal de entrada. A adoção da HR é ubíqua em artigos de modelagem de algoritmos de filtragem adaptativa [42]–[46]. Três referências que se notabilizam por *não* adotarem-na são [34], [47], [48].

3.2- Análise de Primeira Ordem do LMS

Seja a equação de atualização do LMS, reproduzida aqui por conveniência:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta\overline{e}(k)\boldsymbol{u}(k), \tag{49}$$

onde

$$\overline{e}(k) = \overline{d}(k) - y'(k). \tag{50}$$

Como vimos, o sinal de referência $\overline{d}(k)$ é gerado por meio do modelo

$$\overline{d}(k) = [\boldsymbol{w}^{\star}]^T \boldsymbol{u}(k) + \nu(k), \tag{51}$$

com a saída do filtro $y'(k)$ sendo computada via

$$y'(k) = [\boldsymbol{w}(k)]^T \boldsymbol{u}(k). \tag{52}$$

[2]Cabe notar que quanto menor o valor de β, menos drásticas serão as modificações nos coeficientes adaptativos.

Seja o vetor de desvio definido por[3]

$$\tilde{\boldsymbol{w}}(k) \triangleq \boldsymbol{w}^{\text{opt}} - \boldsymbol{w}(k) = \boldsymbol{w}^{\star} - \boldsymbol{w}(k). \tag{53}$$

Com o objetivo de se converter a recursão (49) numa recursão em termos do desvio, apliquemos (51), (52) e (53) em (50), obtendo:

$$\overline{e}(k) = \left[\tilde{\boldsymbol{w}}(k)\right]^T \boldsymbol{u}(k) + \nu(k). \tag{54}$$

Aplicando-se (54) em (49), obtemos:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{u}(k)\boldsymbol{u}^T(k)\tilde{\boldsymbol{w}}(k) + \beta\nu(k)\boldsymbol{u}(k). \tag{55}$$

Multiplicando-se ambos os lados de (55) por -1 e somando-se $\boldsymbol{w}^{\star}$ em ambos os lados, chegamos a

$$\tilde{\boldsymbol{w}}(k+1) = \tilde{\boldsymbol{w}}(k) - \beta \boldsymbol{u}(k)\boldsymbol{u}^T(k)\tilde{\boldsymbol{w}}(k) - \beta\nu(k)\boldsymbol{u}(k), \tag{56}$$

a qual é uma equação determinística e exata, válida para qualquer realização em particular do processo de aprendizagem.

Para que possamos efetuar uma análise em termos de comportamento médio, cumpre aplicar em (56) o operador de valor esperado:

$$\mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] - \beta\mathbb{E}\left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\tilde{\boldsymbol{w}}(k)\right] - \beta\mathbb{E}\left[\nu(k)\boldsymbol{u}(k)\right]. \tag{57}$$

Usando-se HR, temos que $\mathbb{E}\left[\nu(k)\boldsymbol{u}(k)\right] = \boldsymbol{0}$. Utilizando-se adicionalmente HI encontramos

$$\mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] - \beta\boldsymbol{R}_u\mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right], \tag{58}$$

onde

$$\boldsymbol{R}_u \triangleq \mathbb{E}\left[\boldsymbol{u}(k)\boldsymbol{u}^T(k)\right] \tag{59}$$

é a matriz de autocorrelação do sinal de entrada $u(k)$.

Se supusermos que o algoritmo apresenta um comportamento de aprendizagem

[3] Como estamos considerando o algoritmo LMS num contexto de identificação de sistemas, temos que $\boldsymbol{w}^{\text{opt}} = \boldsymbol{w}^{\star}$. No caso do FX-LMS, como tal igualdade não se mantém, não podemos escrever $\tilde{\boldsymbol{w}}(k) = \boldsymbol{w}^{\star} - \boldsymbol{w}(k)$.

convergente, podemos escrever:

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right], \tag{60}$$

situação na qual (58) é equivalente a

$$\lim_{k\to\infty} \boldsymbol{R}_u \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \boldsymbol{0}, \tag{61}$$

o que implica, caso assumamos inversibilidade de $\boldsymbol{R}_u$, que

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \boldsymbol{0}, \tag{62}$$

o que configura a propriedade desejável de estimação assintótica não enviesada.

Caso queiramos compreender melhor o transiente do processo de aprendizagem, importa reescrever como (58)

$$\mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \left(\boldsymbol{I}_N - \beta \boldsymbol{R}_u\right) \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right], \tag{63}$$

onde $\boldsymbol{I}_N$ denota a matriz identidade de dimensões $N \times N$. A recursão apresentada na equação (65) é uma recursão vetorial, a qual pode ser reescrita de modo escalar:

$$\mathbb{E}\left[\tilde{w}_i(k+1)\right] = \left(1 - \beta r_{u,ii}\right) \mathbb{E}\left[\tilde{w}_i(k)\right] - \beta \sum_{j=0, j\neq i}^{N-1} r_{u,ii} \mathbb{E}\left[\tilde{w}_j(k)\right], \tag{64}$$

para $i \in \{0, 1, \ldots, N-1\}$, com $r_{u,ij}$ denotando o elemento de $\boldsymbol{R}_u$ situado na posição (i, j).

A Equação (64) mostra que há um acoplamento entre os diferentes desvios, de modo que o valor de $\mathbb{E}\left[\tilde{w}_i(k+1)\right]$ depende de $\{\mathbb{E}\left[\tilde{w}_j(k+1)\right]\}$, para $j \in \{0, 1, \ldots, N-1\}$. Como geralmente, devido à falta de informação *a priori*, o filtro adaptativo contém coeficientes nulos (isto é, $w_i(k) = 0$, para $i \in \{0, 1, \ldots, N-1\}$), a inicialização da recursão (64) normalmente é

$$\tilde{\boldsymbol{w}}(0) = \boldsymbol{w}^\star - \boldsymbol{w}(0) = \boldsymbol{w}^\star. \tag{65}$$

3.2.1 Exemplo Ilustrativo

Considere o caso $N = 3$, no qual o sinal de entrada $u(k)$ é obtido por meio de um filtro de média móvel (MA, do inglês *moving average*):

$$u(k) = b_0\overline{u}(k) + b_1\overline{u}(k-1), \tag{66}$$

onde $b_i \in \mathbb{R}$, para $i \in \{0, 1\}$ e $\overline{u}(k) \in \mathbb{R}$ é um processo aleatório branco, gaussiano de média nula e de variância $\sigma^2_{\overline{u}}$.

Sob esta configuração, a matriz de autocorrelação pode ser escrita como

$$\boldsymbol{R}_u = \begin{bmatrix} \mathbb{E}\left[\overline{u}^2(k)\right] & \mathbb{E}\left[\overline{u}(k)\overline{u}(k-1)\right] & \mathbb{E}\left[\overline{u}(k)\overline{u}(k-2)\right] \\ \mathbb{E}\left[\overline{u}(k-1)\overline{u}(k)\right] & \mathbb{E}\left[\overline{u}^2(k-1)\right] & \mathbb{E}\left[\overline{u}(k-1)\overline{u}(k-2)\right] \\ \mathbb{E}\left[\overline{u}(k-2)\overline{u}(k)\right] & \mathbb{E}\left[\overline{u}(k-2)\overline{u}(k-1)\right] & \mathbb{E}\left[\overline{u}^2(k-2)\right] \end{bmatrix}, \tag{67}$$

cujos elementos são descritos como

$$\begin{aligned} \mathbb{E}\left[\overline{u}^2(k-i)\right] &= b_0^2\mathbb{E}\left[\overline{u}^2(k-i-1)\right] + 2b_0b_1\mathbb{E}\left[\overline{u}(k-i)\overline{u}(k-i-1)\right] + b_1^2\mathbb{E}\left[\overline{u}^2(k-i-2)\right] \\ &= \left(b_0^2 + b_1^2\right)\sigma^2_{\overline{u}}, \\ \mathbb{E}\left[\overline{u}(k)\overline{u}(k-1)\right] &= b_0b_1\mathbb{E}\left[\overline{u}^2(k-1)\right] = b_0b_1\sigma^2_{\overline{u}}, \\ \mathbb{E}\left[\overline{u}(k)\overline{u}(k-2)\right] &= 0, \end{aligned}$$

de modo que $\boldsymbol{R}_u$, nesta configuração, é

$$\boldsymbol{R}_u = \sigma^2_{\overline{u}}\begin{bmatrix} b_0^2 + b_1^2 & b_0b_1 & 0 \\ b_0b_1 & b_0^2 + b_1^2 & b_0b_1 \\ 0 & b_0b_1 & b_0^2 + b_1^2 \end{bmatrix}. \tag{68}$$

Logo, na configuração deste exemplo, a equação (65) se transmuta em

$$\begin{bmatrix} \mathbb{E}\left[\tilde{w}_0(k+1)\right] \\ \mathbb{E}\left[\tilde{w}_1(k+1)\right] \\ \mathbb{E}\left[\tilde{w}_2(k+1)\right] \end{bmatrix} = \begin{bmatrix} 1 - \beta\sigma^2_{\overline{u}}\left(b_0^2 + b_1^2\right) & -\beta b_0b_1 & 0 \\ -\beta\sigma^2_{\overline{u}}b_0b_1 & 1 - \beta\sigma^2_{\overline{u}}\left(b_0^2 + b_1^2\right) & -\beta\sigma^2_{\overline{u}}b_0b_1 \\ 0 & -\beta\sigma^2_{\overline{u}}b_0b_1 & 1 - \beta\sigma^2_{\overline{u}}(b_0^2 + b_1^2) \end{bmatrix} \begin{bmatrix} \mathbb{E}\left[\tilde{w}_0(k)\right] \\ \mathbb{E}\left[\tilde{w}_1(k)\right] \\ \mathbb{E}\left[\tilde{w}_2(k)\right] \end{bmatrix}. \tag{69}$$

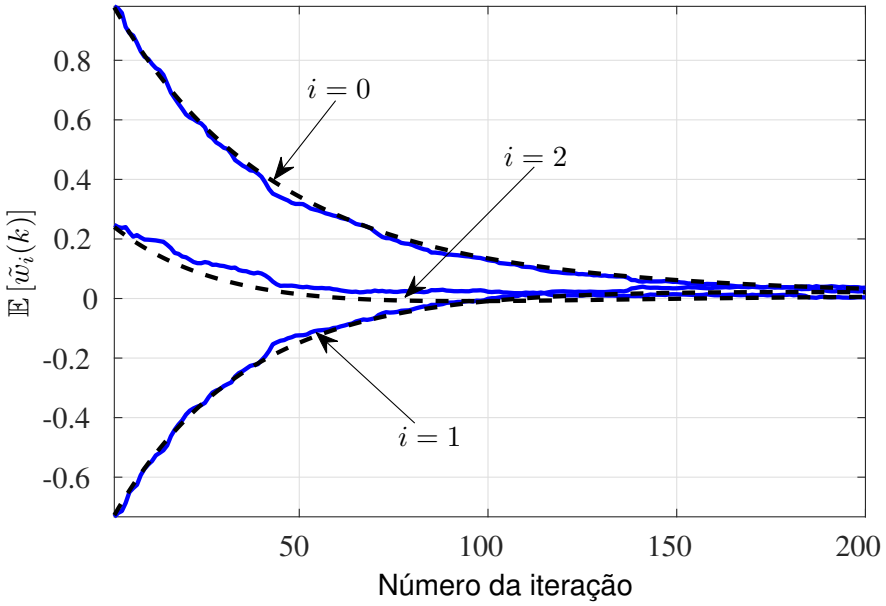


Figura 11 – Comparação da evolução do desvio médio simulado referente aos três primeiros coeficientes (linha azul sólida) com a evolução predita pela teoria (Equação (69), em linha preta tracejada). Os resultados simulados foram obtidos por meio de 10 ensaios independentes de Monte Carlo.

A Figura 11 compara as evoluções teórica e simulada dos desvios médios, para o caso em que $\boldsymbol{w}^\star = \begin{bmatrix} 1 & -0,75 & 0,25 \end{bmatrix}^T$, $\beta = 10^{-2}$, $\sigma_\nu^2 = 10^{-1}$, $b_0 = 1$, $b_1 = -0,8$, $\sigma_{\overline{u}}^2 = 1$, com os sinais $\overline{u}(k)$ e $\nu(k)$ sendo processos brancos, gaussianos e de média zero. A curva empírica se torna mais suave à medida que o número de ensaios independentes de Monte Carlo aumenta. Do mesmo modo, o aumento de σ_ν^2, embora não tenha impacto direto na curva média teórica dos desvios, pode tornar a curva simulada menos suave, exigindo mais ensaios de Monte Carlo para garantir aderência da curva empírica à teórica.

3.2.2 Desacoplamento dos desvios

A Equação (64) apresenta um inconveniente, cada desvio médio no instante $k+1$ depende, em geral, de um conjunto de desvios médios no instante pregresso k. Felizmente, este problema pode ser contornado, conforme veremos a seguir.

É possível se decompor a matriz de autocorrelação $\boldsymbol{R}_u$ do seguinte modo:

$$\boldsymbol{R}_u = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T, \tag{70}$$

onde $\boldsymbol{\Lambda}_u$ é uma matriz diagonal a qual armazena, no i-ésimo elemento de sua diagonal principal, o i-ésimo autovalor $\lambda_i \in \mathbb{R}_+$ de $\boldsymbol{R}$ e $\boldsymbol{Q} \in \mathbb{R}^{N \times N}$ é uma matriz que contém em cada coluna um autovetor, e para a qual vale a relação:

$$\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_N. \tag{71}$$

Assim, utilizando-se a configuração descrita no Exemplo Ilustrativo 1, a matriz

$$\boldsymbol{R}_u = \begin{bmatrix} 1,64 & -0,8 & 0 \\ -0,8 & 1,64 & -0,8 \\ 0 & -0,8 & 1,64, \end{bmatrix} \tag{72}$$

pode ser reescrita, aplicando-se a decomposição (70), como:

$$\boldsymbol{R}_u = \underbrace{\begin{bmatrix} 0,5 & -0,7071 & -0,5 \\ 0,7071 & 0 & 0,7071 \\ 0,5 & 0,7071 & -0,5 \end{bmatrix}}_{=\boldsymbol{Q}} \underbrace{\begin{bmatrix} 0,5086 & 0 & 0 \\ 0 & 1,64 & 0 \\ 0 & 0 & 2,7714 \end{bmatrix}}_{=\boldsymbol{\Lambda}} \begin{bmatrix} 0,5 & -0,7071 & -0,5 \\ 0,7071 & 0 & 0,7071 \\ 0,5 & 0,7071 & -0,5 \end{bmatrix}^T. \tag{73}$$

Seja o *vetor de desvio rotacionado* $\boldsymbol{v}(k)$, definido como

$$\boldsymbol{v}(k) \triangleq \boldsymbol{Q}^T\tilde{\boldsymbol{w}}(k). \tag{74}$$

Multiplicando-se ambos os lados de (65) por $\boldsymbol{Q}^T$, obtemos

$$\mathbb{E}\left[\tilde{\boldsymbol{v}}(k+1)\right] = \boldsymbol{Q}^T\left(\boldsymbol{I}_N - \beta\boldsymbol{R}_u\right)\mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right]. \tag{75}$$

Como $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}_N$ (vide (71)), podemos reescrever (75) como

$$\mathbb{E}\left[\boldsymbol{v}(k+1)\right] = \boldsymbol{Q}^T\left(\boldsymbol{I}_N - \beta\boldsymbol{R}_u\right)\boldsymbol{Q}\boldsymbol{Q}^T\mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \boldsymbol{Q}^T\left(\boldsymbol{I}_N - \beta\boldsymbol{R}_u\right)\boldsymbol{Q}\mathbb{E}\left[\boldsymbol{v}(k)\right]. \tag{76}$$

Utilizando-se (70), (76) leva a

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{v}(k+1)\right] &= \left(\boldsymbol{Q}^T\boldsymbol{Q} - \beta\boldsymbol{Q}^T\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T\boldsymbol{Q}\right)\mathbb{E}\left[\boldsymbol{v}(k)\right] \\ &= \left(\boldsymbol{I}_N - \beta\boldsymbol{\Lambda}\right)\mathbb{E}\left[\boldsymbol{v}(k)\right], \end{aligned} \tag{77}$$

de modo que a recursão do i-ésimo elemento do vetor de desvio rotacionado $v_i(k)$ não depende de outros elementos de $\boldsymbol{v}(k)$:

$$\mathbb{E}\left[v_i(k+1)\right] = (1-\beta\lambda_i)\mathbb{E}\left[v_i(k)\right], \tag{78}$$

de modo que

$$\mathbb{E}\left[v_i(k)\right] = (1-\beta\lambda_i)^k\mathbb{E}\left[v_i(0)\right]. \tag{79}$$

Devido a (74), temos que o algoritmo converge na média apenas quando

$$\lim_{k\to\infty}\mathbb{E}\left[v_i(k)\right] = 0.$$

Por meio de (79), vemos que isso somente ocorre quando

$$-1 < 1-\beta\lambda_i < 1, \tag{80}$$

a qual pode ser transposta numa região de estabilidade para o parâmetro ajustável β:

$$0 < \beta < \frac{2}{\lambda_i}, \tag{81}$$

a qual deve ser satisfeita $\forall i \in \{0, 1, \ldots, N-1\}$, ou, usando o intervalo mais restrito de todos:

$$0 < \beta < \frac{2}{\lambda_{\mathsf{max}}}, \tag{82}$$

onde λ_{max} é o maior autovalor da matriz $\boldsymbol{R}_u$. Como na prática λ_{max} não é conhecido, ele pode ser conservadoramente estimado por meio de

$$\lambda_{\mathsf{max}} < \sum_{i=0}^{N-1}\lambda_i = \mathsf{Tr}\left[\boldsymbol{R}_u\right] = N\sigma_u^2, \tag{83}$$

onde $\mathsf{Tr}[\boldsymbol{A}]$ denota o traço da matriz $\boldsymbol{A}$.

3.3- Conclusões Parciais

Este capítulo descreveu um modo clássico de análise de primeira ordem do algoritmo LMS. Os conceitos aqui trazidos serão mobilizados para caracterizar a tendenciosidade do algoritmo N-FX-LMS, bem como para se obter o algoritmo BC-FX-LMS, que é uma inovação desta dissertação. Antes de tratarmos destas variantes do FX-LMS, iremos no próximo capítulo descrever o algoritmo FX-LMS, o qual terminará de fornecer a base conceitual necessária para o entendimento das soluções propostas.

4- FX-LMS

O algoritmo FX-LMS é objeto de diversas análises teóricas na literatura [49]–[53]. Uma análise que particularmente nos interessa é a determinação de sua solução ótima (no sentido de minimização do erro quadrático médio), a qual será descrita na próxima seção. Uma das principais áreas de aplicação do FX-LMS é o *Active Noise Control (ANC)* ou Controle Ativo de Ruído que é uma das técnicas mais eficazes de redução de ruído. Sistemas de controle passivo como, por exemplo, absorvedores acústicos apresentam resultados satisfatórios com relação às altas frequências mas, em se tratando de frequências baixas, a solução costuma ser onerosa em termos financeiros e quantitativos [54].

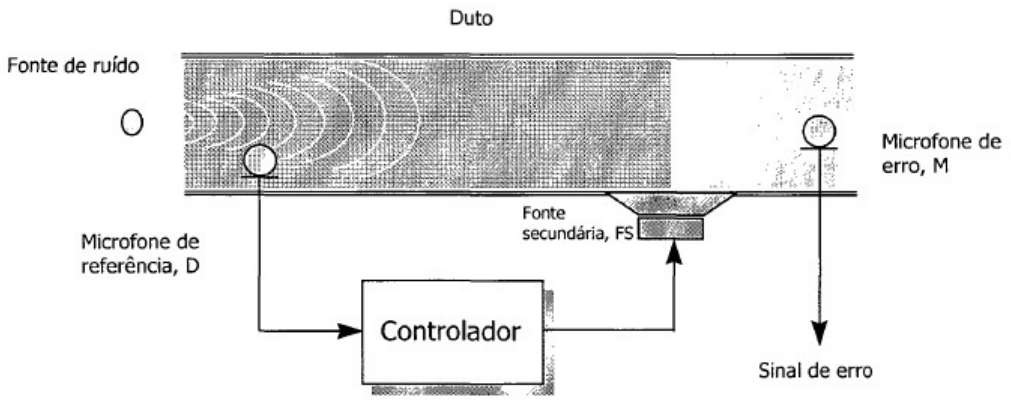


Figura 12 – Sistema de controle ativo de ruído [50].

Na Fig. 12 podemos ver um exemplo da aplicação de controle ativo de ruído através de um sistema de controle onde um sinal ruidoso é propagado através de um duto e capturado por um microfone. Cumpre ao sistema de cancelamento de ruído a geração de um ”campo antirruído” (*antinoise field*) responsável por cancelar, mediante o concurso do fenômeno conhecido como *interferência destrutiva*, o campo de ruído induzido pelo primário [55]. Caso o caminho entre o atuador secundário e o microfone de erro não introduza distorções, o algoritmo LMS pode ser empregado. Desafortunadamente, raramente é este o caso em aplicações de cancelamento de ruído ou de vibrações; assim, o algoritmo FX-LMS passa a ser uma opção competitiva como escolha de controlador [53].

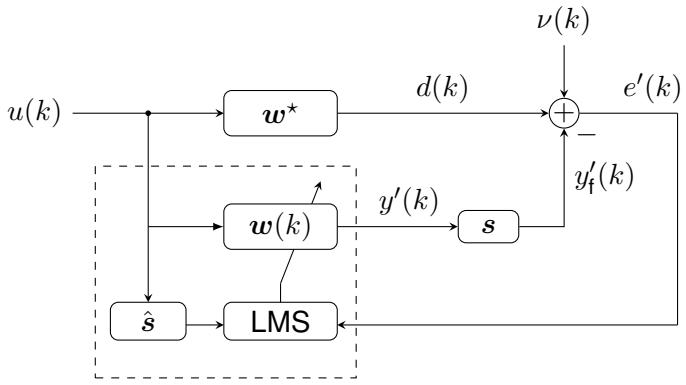


Figura 13 – Diagrama em blocos do algoritmo FX-LMS.

Conforme visto no Capítulo 1, seja $u(k)$ o sinal registrado no microfone de referência e $e'(k)$ o sinal obtido pelo microfone de erro. O algoritmo FX-LMS pode ser mais bem compreendido pelo diagrama de blocos da Fig. 13, com a saída do filtro adaptativo $y'(k) \in \mathbb{R}$ sendo calculada por

$$y'(k) = \boldsymbol{w}^T(k)\boldsymbol{u}(k), \tag{84}$$

onde

$$\boldsymbol{u}(k) \triangleq \begin{bmatrix} u(k) & u(k-1) & u(k-2) & \dots & u(k-N+1) \end{bmatrix}^T. \tag{85}$$

Já o sinal registrado no microfone de erro $y'_{\mathsf{f}}(k) \in \mathbb{R}$ é uma versão distorcida da saída do filtro $y'(k)$, já que é obtida *após* a passagem pelo caminho secundário $S(z)$:

$$y'_{\mathsf{f}}(k) = \sum_{m=0}^{M-1} s_m y'(k-m) = \sum_{m=0}^{M-1} s_m \boldsymbol{w}^T(k-m)\boldsymbol{u}(k-m), \tag{86}$$

onde assume-se que o caminho secundário apresenta comprimento M. Cabe enfatizar que não se leva em consideração um eventual *feedback* que $y'(k)$ tenha em $u(k)$ [51]; tal impacto é considerado desprezível.

A equação de atualização do algoritmo FX-LMS pode ser descrita como [56]

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{u}_{\mathsf{f}}(k) e'(k), \tag{87}$$

onde

$$\boldsymbol{u}_{\mathsf{f}}(k) \triangleq \sum_{m=0}^{M-1} \hat{s}_m \boldsymbol{u}(k), \tag{88}$$

com $\hat{s}_m \in \mathbb{R}$ denotando a estimativa de s_m, para $m \in \{0, 1, \ldots, M-1\}$, e

$$e'(k) = d(k) + \nu(k) - y'_{\mathsf{f}}(k). \tag{89}$$

Observe que assume-se que o comprimento da estimativa do caminho secundário coincide com o comprimento real deste caminho. Ademais, note que a equação (87) implementa uma filtragem de sucessivos vetores de entrada $\{\boldsymbol{u}(k), \boldsymbol{u}(k-1), \ldots, \boldsymbol{u}(k-M+1)\}$ que emula digitalmente o impacto que a planta secundária apresenta na saída do filtro. A Tabela 2 apresenta a complexidade computacional requerida por iteração do algoritmo FX-LMS. Cabe ressaltar que o algoritmo apresenta um custo computacional proporcional ao tamanho N do filtro adaptativo.

Operação	Número de adições	Número de multiplicações
Cálculo de $y'(k) = [\boldsymbol{w}(k)]^T \boldsymbol{u}(k)$	$N-1$	N
Cálculo de $e'(k)$	0	0
Cálculo de $\beta e'(k)$	0	1
Cálculo de $\boldsymbol{u}_{\mathsf{f}}(k)$	$(M-1)N$	MN
Cálculo de $\beta\overline{e}(k)\boldsymbol{u}_{\mathsf{f}}(k)$	0	N
Cálculo de $\boldsymbol{w}(k+1)$	N	0
Total	$(M+1)N-1$	$(M+2)N+1$

Tabela 2 – Complexidade computacional envolvida em uma atualização do algoritmo FX-LMS. A complexidade de uma subtração é considerada equivalente à de uma adição. O cômputo de $e'(k)$ não implica operações aritméticas, já que ele é resultado da medida do resultado líquido da interferência física entre os sinais do primário e do secundário.

4.1- Solução Ótima do FX-LMS

A solução ótima do algoritmo FX-LMS (no sentido quadrático médio), minimiza o erro quadrático médio (MSE), o qual pode ser escrito como

$$\xi(k) \triangleq \mathbb{E}[(e'(k))^2], \tag{90}$$

onde $\mathbb{E}[\cdot]$ denota o operador de valor esperado.

Tendo em vista a simplificação da análise, considere a seguinte hipótese estocástica:

Hipótese Referente ao Ruído (HR). O ruído aditivo de média zero $\nu(k)$ é branco, estacionário e estatisticamente independente das variáveis aleatórias restantes.

Observe que HR é uma hipótese estatística onipresente, embora existam métodos para evitá-la. Usando HR, (89) e assumindo um vetor fixo $\boldsymbol{w} \in \mathbb{R}^N$ leva a

$$\begin{aligned} \xi[\boldsymbol{w}] &= \mathbb{E}\left[d^2(k)\right] - 2\sum_{i=0}^{M-1} s_i \boldsymbol{p}_i^T \boldsymbol{w} + \mathbb{E}\left[\nu^2(k)\right] \\ &\quad + \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \boldsymbol{w}^T \boldsymbol{R}_{j-i} \boldsymbol{w}, \end{aligned} \tag{91}$$

onde $\boldsymbol{p}_i \triangleq \mathbb{E}\left[d(k)\boldsymbol{u}(k-i)\right]$ e $\boldsymbol{R}_{j-i} \triangleq \mathbb{E}\left[\boldsymbol{u}(k-i)\boldsymbol{u}^T(k-j)\right]$.

Usando manipulações diretas, (91) pode ser reescrita como

$$\xi[\boldsymbol{w}] = \mathbb{E}[d^2(k)] - 2\tilde{\boldsymbol{p}}^T \boldsymbol{w} + \boldsymbol{w}^T \tilde{\boldsymbol{R}} \boldsymbol{w} + \sigma_\nu^2, \tag{92}$$

onde $\tilde{\boldsymbol{p}} \triangleq \sum_{i=0}^{M-1} s_i \boldsymbol{p}_i$ e $\tilde{\boldsymbol{R}} \triangleq \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \boldsymbol{R}_{j-i}$.

A partir de (92), pode ser provado que a solução ótima para $\boldsymbol{w}$ que minimiza $\xi(k)$ pode ser encontrada igualando-se o gradiente de $\xi[\boldsymbol{w}]$ em relação a $\boldsymbol{w}$ a zero, o que leva a:

$$\nabla_{\boldsymbol{w}} \xi[\boldsymbol{w}] = -2\tilde{\boldsymbol{p}} + 2\tilde{\boldsymbol{R}}\boldsymbol{w} = \boldsymbol{0} \Rightarrow \boldsymbol{w}^{\text{opt}} = \left[\tilde{\boldsymbol{R}}\right]^{-1} \tilde{\boldsymbol{p}}, \tag{93}$$

um resultado que será usado na análise de primeira ordem do FX-LMS com entrada ruidosa. É digno de nota que $\boldsymbol{w}^{\text{opt}}$ não coincide com $\boldsymbol{w}^\star$ em geral, em contraste com o que acontece em tarefas de identificação de sistemas. Cabe ressaltar que *não* se espera que o algoritmo FX-LMS, mesmo num contexto estacionário, convirja para a solução ótima, porque infelizmente a função de transferência do secundário é desconhecida, precisando ser estimada. Desafortunadamente, imprecisões nesta estimativa degradam o desempenho do FX-LMS.

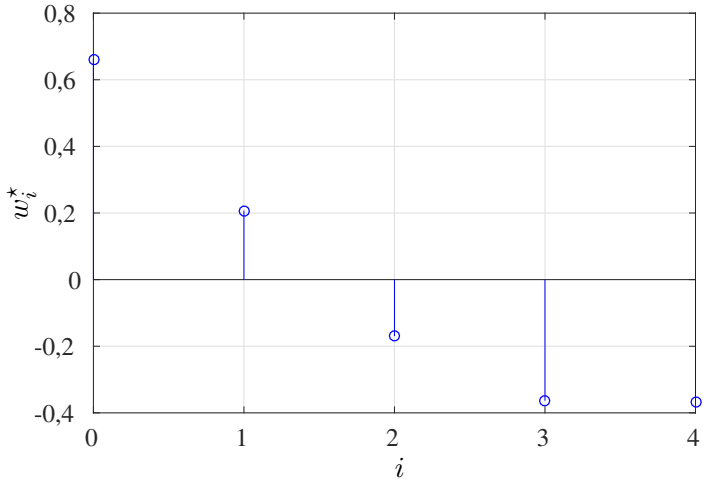


Figura 14 – Coeficientes do filtro $\boldsymbol{w}^\star$, responsável pelo acoplamento entre o ruído do primário e o sinal registrado no microfone de erro.

4.2- Exemplo Ilustrativo

De sorte a ilustrar alguns conceitos importantes referentes ao aprendizado do algoritmo FX-LMS, seja um exemplo no qual $\boldsymbol{w}^\star \in \mathbb{R}^5$, com

$$w_i^\star = \cos(0,2\pi(i+1))\exp\left[-0,2(i+1)\right], \tag{94}$$

para $i \in \{0,1,\ldots,4\}$. A Fig. 14 apresenta os coeficientes de $\boldsymbol{w}^\star$. O filtro secundário apresenta a função de transferência $S(z) = 1 + 0,8z^{-1}$, e o filtro adaptativo apresenta comprimento $N = 5$. A medição do erro é corrompida por um ruído $\nu(k)$ branco, gaussiano, de média zero e variância 10^{-6}. O sinal de entrada $u(k)$ é obtido mediante o colorimento de um sinal branco, gaussiano, de média zero e variância unitária por um filtro de colorimento $B(z) = 1 - 0,2z^{-1}$.

Para esta configuração, o algoritmo FX-LMS foi implementado com $N = 5$ e $\beta = 10^{-2}$. Devido à ausência de informação *a priori*, os coeficientes do filtro adaptativo são inicializados como zeros. A Fig. 15 apresenta a evolução média dos coeficientes adaptativos, assumindo-se uma perfeita estimação da planta do secundário. Pode-se observar que os coeficientes convergem para a solução ótima (93), o que é condizente com a característica não tendenciosa do estimador implementado pelo algoritmo FX-LMS [50].

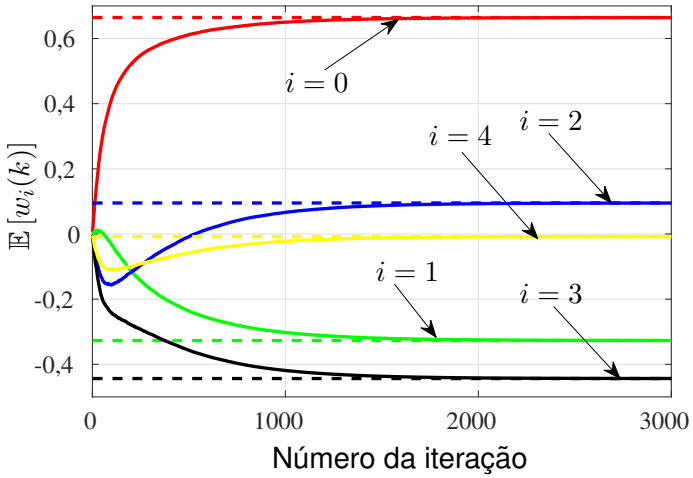


Figura 15 – As curvas sólidas representam valores médios (obtidos a partir de 100 ensaios independentes de Monte Carlo) dos coeficientes adaptativos estimados com o algoritmo FX-LMS, assumindo-se uma estimativa perfeita do caminho do secundário. As retas tracejadas indicam os respectivos valores de $\boldsymbol{w}_{\text{opt}}$, calculados de acordo com (93).

A Fig. 16 apresenta a evolução do MSD (em dB), ao longo das iterações, o qual é computado, como vimos, por meio de

$$\text{MSD}(k) \text{ (em dB)} = 10\log_{10}\left\{\mathbb{E}\left[\|\boldsymbol{w}_{\text{opt}} - \boldsymbol{w}(k)\|^2\right]\right\}. \tag{95}$$

Observa-se inicialmente um alto valor de MSD, associado ao fato de o filtro adaptativo ainda não ter sido treinado. Com a evolução do procedimento de aprendizado, o algoritmo vai, pouco a pouco, atenuando a energia do sinal de erro, conforme desejado. Esta etapa do processo de aprendizagem é conhecida como transiente. Após um determinado número de iterações, o processo de aprendizagem se estabiliza (assumindo-se um contexto estacionário), dando origem a um patamar de desempenho constante, normalmente associado ao regime permanente ou à região de operação assintótica do algoritmo. Tal patamar é altamente dependente de parâmetros ajustáveis do filtro, tais como β e N. Considerações similares podem ser efetuadas acerca da evolução do MSE (em dB), apresentado na Fig. 17.

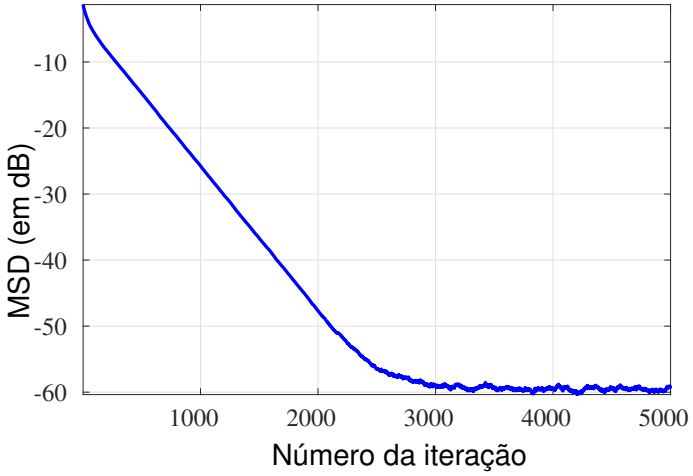


Figura 16 – Evolução do MSD (em dB) ao longo das iterações. A curva foi computada através de 100 ensaios independentes de Monte Carlo.

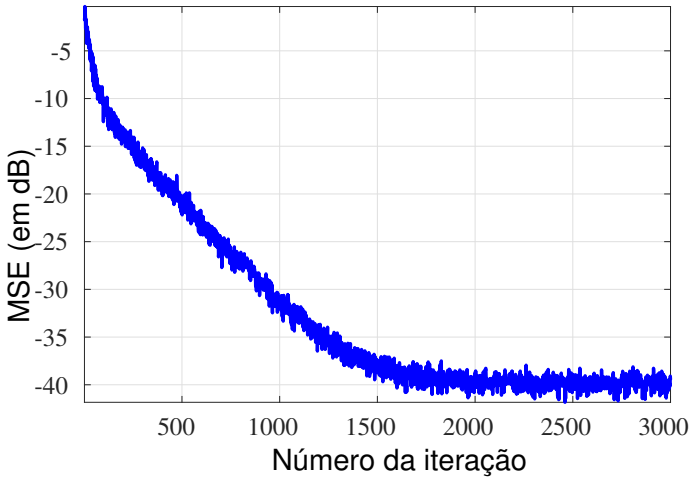


Figura 17 – Evolução do MSE (em dB) ao longo das iterações. A curva foi computada através de 100 ensaios independentes de Monte Carlo.

4.3- Conclusões Parciais

Este capítulo apresentou o algoritmo FX-LMS clássico, bem como sua solução ótima. Algumas simulações foram utilizadas para ilustrar alguns pontos importantes, como a delimitação das regiões de transiente e de regime permanente.

Classicamente, os artigos que analisam o algoritmo FX-LMS assumem que apenas existe ruído no sensor de erro. Porém, na prática, a aquisição do sinal no sensor de referência também pode ser um processo ruidoso, por questões relativas à quantização inerente à conversão analógico-digital ou mesmo devido a interferências.

5- Algoritmos Adaptativos com Compensação de Viés

5.1- Introdução

O artigo [57] foi o primeiro a propor uma técnica de filtragem adaptativa para compensar o viés induzido por ruído aditivo na entrada mediante um custo computacional apropriado. Este artigo deu origem a uma série de publicações, as quais utilizam a compensação de viés em diversos algoritmos e configurações, normalmente enfocando o problema de identificação de sistemas. Cabe ressaltar que historicamente existe uma abordagem concorrente, a qual se vale do método *Total Least Squares* [58], a qual entretanto apresenta aplicabilidade limitada devido ao seu alto custo computacional.

As equações de atualização que definem os diferentes algoritmos com compensação de viés podem ser obtidas por diferentes métodos. Por razões didáticas, iremos aqui descrever o método cujo entendimento é mais fácil, aplicando-o ao algoritmo mais simples de filtragem adaptativa: o LMS. Quando este algoritmo opera com uma entrada ruidosa, sua equação de atualização pode ser descrita como:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{x}(k) e(k), \tag{96}$$

onde o erro $e(k)$ pode ser descrito como

$$e(k) = [\boldsymbol{w}^{\star}]^T u(k) + \nu(k) - \boldsymbol{w}^T(k)\boldsymbol{x}(k) = \tilde{\boldsymbol{w}}^T(k)\boldsymbol{u}(k) + \nu(k) - \boldsymbol{w}^T(k)\boldsymbol{\eta}(k). \tag{97}$$

Aplicando-se (97) em (96) e utilizando (53), encontramos:

$$\tilde{\boldsymbol{w}}(k+1) = \tilde{\boldsymbol{w}}(k) - \beta \boldsymbol{x}(k)\boldsymbol{u}^T(k)\tilde{\boldsymbol{w}}(k) - \beta \boldsymbol{x}(k)\nu(k) + \beta \boldsymbol{x}(k)\boldsymbol{\eta}^T(k)\boldsymbol{w}(k). \tag{98}$$

Aplicando-se o operador de valor esperado em (98), HR e a hipótese da independência, temos que:

$$\mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \left[\boldsymbol{I}_N - \beta \boldsymbol{R}_u\right] \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] + \beta \sigma_{\eta}^2 \boldsymbol{w}(k). \tag{99}$$

Assumindo que houve convergência, em regime permanente temos que

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] \tag{100}$$

e conclui-se de (99) a emergência de um viés em regime permanente, descrito por:

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \lim_{k\to\infty} \sigma_\eta^2 \boldsymbol{R}_u^{-1} \mathbb{E}\left[\boldsymbol{w}(k)\right]. \tag{101}$$

Como este viés assintótico pode prejudicar o desempenho do algoritmo LMS na presença de ruído aditivo na entrada, pode-se reduzi-lo por meio da inserção de um termo de compensação de viés, de modo que a equação do algoritmo BC-LMS (LMS com compensação de viés) pode ser descrita como

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{x}(k) e(k) + \boldsymbol{\psi}_{\text{BC-LMS}}(k), \tag{102}$$

onde $\boldsymbol{\psi}_{\text{BC-LMS}}(k) \in \mathbb{R}^N$ denota o vetor de compensação de viés introduzido pelo algoritmo.

Refazendo-se toda a derivação que deu origem a (101), percebe-se que a seguinte escolha de $\boldsymbol{\psi}_{\text{BC-LMS}}(k)$ é passível de anular a polarização em regime permanente:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta \boldsymbol{x}(k) e(k) + \overbrace{\beta \hat{\sigma}_\eta^2 \boldsymbol{w}(k)}^{=\boldsymbol{\psi}_{\text{BC-LMS}}(k)}, \tag{103}$$

onde $\hat{\sigma}_\eta^2$ é a estimativa da variância σ_η^2 de $\eta(k)$.

5.2- Revisão Bibliográfica

Embora os primeiros artigos sobre compensação de viés correlatos aos abordados nessa dissertação sejam de 2013, existe já uma ampla literatura sobre o assunto, em geral enfatizando o caso de identificação de sistemas. Nesta seção, iremos tratar concisamente das principais destas referências. Em comum entre elas, está a exploração das propriedades estatísticas do ruído presente nas entradas e nos sinais de erro [59].

A referência [60] adapta a técnica de compensação de viés para o contexto de algoritmos adaptativos difusos. As variâncias dos ruídos aditivos nos sinais de entrada

são estimadas em tempo real, com a estabilidade e a convergência dos algoritmos sendo caracterizadas por fórmulas teóricas fechadas.

Sabe-se que diversas funções de transferência na prática são esparsas, significando que sua energia é concentrada em poucos coeficientes [61]–[63]. Até o início dos anos 2000, os algoritmos adaptativos eram agnósticos a este respeito. Entretanto, observou-se que a incorporação no mecanismo de aprendizado dos algoritmos adaptativos desta informação referente à esparsidade pode aumentar significativamente a taxa de convergência destes algoritmos. Esta constatação deu origem à família de algoritmos ditos *conscientes de esparsidade* [64], [65]. Em geral, tais algoritmos podem utilizar o princípio proporcional (ou seja, concedem aos coeficientes de maior magnitude uma maior energia de atualização) ou adicionam um termo de regularização por norma. A referência [66] propõe a adoção da técnica de compensação de viés em conjunto com uma penalização por norma[1] ℓ_0 na equação de atualização, de modo a conjugar a compensação de viés com a consciência de esparsidade. A referência [67] adota uma estratégia similar, sendo a principal diferença com relação a [66] o emprego da adaptação proporcional, ao invés de efetuar uma regularização por norma. Assumindo que a esparsidade ocorre em blocos, a referência [68] aplica a técnica de compensação de viés, e permite reduzir o seu custo computacional por meio da estratégia de atualização seletiva[2].

Em aplicações com funções de transferência longas (como cancelamento de eco acústico), a taxa de convergência de algoritmos de filtragem adaptativa pode se tornar excessivamente lenta, violando os requisitos de algumas aplicações. Uma forma de contornar este problema consiste na adoção de técnicas de processamento em subbandas, as quais descorrelacionam o sinal de entrada ao longo das subbandas, aumentando a taxa de convergência [69]–[72]. Neste contexto, a referência [73] se destaca por combinar a compensação de viés com técnicas de processamento adaptativo em subbandas. O artigo também propõe um novo método para estimação da variância do ruído na entrada, a qual não requer antecipadamente a razão das variâncias da entrada e da saída.

Um problema comum em identificação de sistemas ocorre quando o ruído envolvido tem natureza impulsiva. Neste caso, as caudas da distribuição em probabilidade do sinal de ruído são mais significativas dos que as caudas de uma distribuição gaussi-

[1] Na realidade, a norma ℓ_0 é uma pseudo-norma, já que não adere às propriedades clássicas de uma norma.

[2] Como veremos, esta é uma contribuição desta dissertação, embora não seja a principal delas.

ana. Quando isso ocorre, funções custo quadráticas não são as mais adequadas. Uma forma de compensar este problema consiste em utilizar funções custo mais robustas, baseando-se em conceitos tais como o da *máxima correntropia* [62]. A referência [74] notabilizou-se por conjugar um aprendizado orientado pelo critério de máxima correntropia com a compensação de viés.

A partir da constatação de que diversos pares de entrada $\{\boldsymbol{x}(k), d(k)\}$ não oferecem informação relevante, a filosofia *Set-Membership*, quando aplicada a algoritmos de filtragem adaptativa, adota uma espécie de censura aos dados [75]–[77]. Assim, apenas os dados mais informativos são utilizados para adaptação, reduzindo o custo computacional do filtro adaptativo. A referência [78] é de especial interesse no nosso contexto, porque foi a primeira a conjugar a filosofia *Set-membership* com a técnica de compensação de viés.

Cumpre notar que as diversas referências aqui citadas notabilizam-se por combinar a técnica de compensação de viés com alguma estratégia específica de filtragem adaptativa. A Tabela 3 explicita o contexto no qual tais referências estão envolvidas.

Referência(s)	Contexto
[60]	Adaptação difusa
[66]–[68]	Algoritmos conscientes da esparsidade
[73]	Filtragem adaptativa em subbandas
[74]	Ruído impulsivo
[78]	Filosofia *Set-Membership*

Tabela 3 – Contexto das principais referências que utilizam o princípio da compensação de viés.

5.3- Redução do Custo Computacional de uma Técnica de Compensação de Viés

Uma das contribuições desta dissertação reside na redução do custo computacional de um algoritmo de compensação de viés consciente da esparsidade. Tal contribuição deu origem à publicação[3] [68].

[3]Cabe ressaltar que a proposta do algoritmo original deve-se a Rodrigo Pimenta, cabendo ao autor desta dissertação tão somente a aplicação do princípio da atualização seletiva para redução do custo

O algoritmo em questão é o BC-BS-NLMS (do inglês *Block-Sparsity Induced NLMS Algorithm with Bias Compensation*). Tal algoritmo insere um termo de regularização de norma mista $\ell_{2,0}$. Tal regularização é aplicada a uma partição de M' subconjuntos do vetor adaptativo $\boldsymbol{w}(k)$, de modo que $\boldsymbol{w}(k)$ é decomposto em M' blocos $\boldsymbol{w}_i(k)$, para $i \in \{0, 1, \ldots, M'-1\}$,

$$\boldsymbol{w}(k) = \begin{bmatrix} \boldsymbol{w}_0^T(k) & \boldsymbol{w}_1^T(k) & \ldots & \boldsymbol{w}_{M'-1}^T(k) \end{bmatrix}^T, \tag{104}$$

onde o comprimento L' de cada bloco é dado por $L' = N/M'$, com[4] $L' \in \mathbb{N}$.

A motivação do algoritmo BC-BS-NLMS se dá em contextos nos quais os coeficientes de maior magnitude estão concentrados (ao invés de estarem distribuídos de modo aleatório), de modo que se espera que a norma mista $\ell_{2,0}$ não apresente valores elevados na prática [79]:

$$\|\boldsymbol{w}(k)\|_{2,0} \triangleq \left\| \begin{bmatrix} \|\boldsymbol{w}_0(k)\|_2 \\ \|\boldsymbol{w}_1(k)\|_2 \\ \vdots \\ \|\boldsymbol{w}_{M'-1}(k)\|_2 \end{bmatrix} \right\|_0, \tag{105}$$

onde a pseudo-norma ℓ_0 é normalmente aproximada, de modo a tornar a matemática tratável [80]. Observe que (105) avalia a esparsidade de um vetor em $\mathbb{R}^{M'}$, o qual contém as energias de M' blocos. Ou seja, a técnica BS incorpora *a priori* uma hipótese mais forte do que os algoritmos clássicos conscientes da esparsidade, almejando aumentar a taxa de convergência quando tal hipótese for efetivamente verdadeira.

A derivação do algoritmo BC-BS-NLMS começa pela resolução do seguinte problema de otimização:

$$\min_{\boldsymbol{w}(k+1)} \mathcal{F}_{\mathsf{BS}}[\boldsymbol{w}(k+1)] \triangleq \frac{1}{2}\|\boldsymbol{w}(k+1) - \boldsymbol{w}(k)\|^2 + \overline{\gamma}\|\boldsymbol{w}(k+1)\|_{2,0}$$

$$\text{s.t. } e_{\mathsf{p}}(k) = (1-\beta)e(k), \tag{106}$$

onde $\overline{\gamma} \in \mathbb{R}_+$ é um parâmetro ajustável que controla a intensidade da regularização.

Após a resolução de (106) e a introdução de um termo para combater o viés

computacional do algoritmo.

[4] Por simplicidade, assume-se que a razão N/M é um inteiro.

assintótico[5], o algoritmo BC-BS-NLMS apresenta a seguinte equação de atualização:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \gamma \boldsymbol{g}[\boldsymbol{w}(k)] + \beta \boldsymbol{x}(k)e(k) + \beta \hat{\sigma}_{\eta}^{2} \boldsymbol{w}(k), \tag{107}$$

onde o i-ésimo elemento de $\boldsymbol{g}[\boldsymbol{w}]$ é definido por

$$g_i(k) \triangleq \begin{cases} 2\rho^2 w_i(k) - \frac{2\rho w_i(k)}{\|\boldsymbol{w}_{\lfloor i/L \rfloor}\|_2}, & 0 < \|\boldsymbol{w}_{\lfloor i/L \rfloor}\|_2 < 1/\rho \\ 0, & \text{caso contrário.} \end{cases}, \tag{108}$$

onde $\rho \in \mathbb{R}_+$ é um parâmetro ao arbítrio do projetista que influencia a aproximação da norma ℓ_0 [81]. Note em (108) que o termo de atração para zero $g_i(k)$ só é ativo quando a energia do bloco respectivo é pequena.

Uma das contribuições desta dissertação consiste em modificar a equação de atualização (107) de modo a nela inserir o princípio da atualização parcial seletiva [82], almejando introduzir um compromisso entre taxa de convergência e custo computacional.

Como vimos, aplicações que requerem filtros adaptativos longos podem demandar uma complexidade computacional extremamente elevada. Nestes casos críticos, a adoção de um método de atualização parcial seletiva (do inglês *Selective Partial Update (SPU)* é capaz de reduzir tal complexidade ao atualizar apenas uma fração dos coeficientes adaptativos a cada iteração. Tal vantagem, todavia, tem um preço: a redução da taxa de convergência.

Para motivar o algoritmo *proposto* SPU-BC-BS-NLMS, considere a seguinte partição do vetor do sinal de entrada em M' blocos de igual comprimento[6]:

$$\boldsymbol{x}(k) \triangleq \begin{bmatrix} \boldsymbol{x}_0^T(k) & \boldsymbol{x}_1^T(k) & \dots & \boldsymbol{x}_{M'-1}^T(k) \end{bmatrix}^T, \tag{109}$$

onde assume-se que a cada iteração o algoritmo atualiza B blocos do vetor de pesos adaptativos (veja (104)).

O método SPU pode ser obtido pela solução resultante do seguinte problema de otimização com restrições:

$$\min_{\boldsymbol{w}_i(k+1)} \frac{1}{2}\|\boldsymbol{w}_i(k+1) - \boldsymbol{w}_i(k)\|^2 + \gamma \|\boldsymbol{w}_i(k+1)\|_{2,0}$$

[5] A derivação é relativamente longa e, por não ser contribuição do autor desta dissertação, não será aqui detalhada.

[6] Note que não é necessário que o número de partições da estratégia SPU seja o mesmo do método BS (do inglês *Block-Induced Sparsity*). Tal igualdade foi imposta apenas para simplificar as equações.

$$\text{s.t. } e_{\mathsf{p}}(k) = (1-\beta)e(k), \tag{110}$$

o qual, utilizando a metodologia já apresentada, leva a

$$\boldsymbol{w}_i(k+1) = \boldsymbol{w}_i(k) + \gamma \boldsymbol{g}[\boldsymbol{w}_i(k)] + \beta \frac{\boldsymbol{x}_i(k)e(k)}{\|\boldsymbol{x}_i(k)\|^2}, \tag{111}$$

enquanto que os demais blocos de $\boldsymbol{w}(k)$ continuam inalterados. Observe entretanto que (111) não indica qual bloco deve ser atualizado. Para efetuar tal escolha, o seguinte critério foi adotado:

$$\begin{aligned} i &= \arg \min_{1\leq j\leq M} \|\boldsymbol{w}_j(k+1) - \boldsymbol{w}_j(k)\|^2 \\ &\approx \arg \min_{1\leq j\leq M} \left[\frac{\beta^2 e^2(k)}{\|\boldsymbol{x}_i(k)\|^2}\right] = \arg \max_{1\leq j\leq M} \|\boldsymbol{x}_j(k)\|^2, \end{aligned} \tag{112}$$

uma aproximação que despreza o componente $\gamma \boldsymbol{g}[\boldsymbol{w}_i(k)]$ em (111), de sorte a reduzir o esforço computacional.

A atualização de apenas um bloco por iteração pode ser uma restrição muito forte na prática, porque pode reduzir em demasia a taxa de convergência. Considere que o projetista pretenda atualizar $B \in \mathbb{N}$ blocos a cada iteração, cujos índices pertencem ao conjunto $\mathcal{I}_B = \{i_0, i_1, \ldots, i_{B-1}\}$, o qual é um subconjunto de $S = \{0, 1, \ldots, M'-1\}$. O caso $B = 0$ não tem relevância prática, pois implicaria ausência de atualização. O caso $B = M'$ degeneraria o algoritmo SPU-BC-BS-NLMS no algoritmo BC-BS-NLMS, já que implicaria atualização de todos os coeficientes em cada iteração. Assim, normalmente espera-se que $1 \leq B < M'$.

Para tratar deste caso menos restritivo, podemos generalizar (110), o que leva a

$$\min_{\boldsymbol{w}_{\mathcal{I}_B}(k+1)} \frac{1}{2}\|\boldsymbol{w}_{\mathcal{I}_B}(k+1) - \boldsymbol{w}_{\mathcal{I}_B}(k)\|^2 + \gamma \|\boldsymbol{w}_{\mathcal{I}_B}(k+1)\|_{2,0}$$

$$\text{s.t. } e_{\mathsf{p}}(k) = (1-\beta)e(k), \tag{113}$$

cuja solução (seguindo os mesmos passos que deram origem a (111)) é

$$\boldsymbol{w}_{\mathcal{I}_B}(k+1) = \boldsymbol{w}_{\mathcal{I}_B}(k) + \gamma \boldsymbol{g}[\boldsymbol{w}_{\mathcal{I}_B}(k)] + \beta \frac{\boldsymbol{x}_{\mathcal{I}_B}(k)e(k)}{\|\boldsymbol{x}_{\mathcal{I}_B}(k)\|^2}, \tag{114}$$

onde

$$\mathcal{I}_B = \arg \ \max_{J_B \in S} \sum_{j\in J_B} \|\boldsymbol{x}_j(k)\|^2 \tag{115}$$

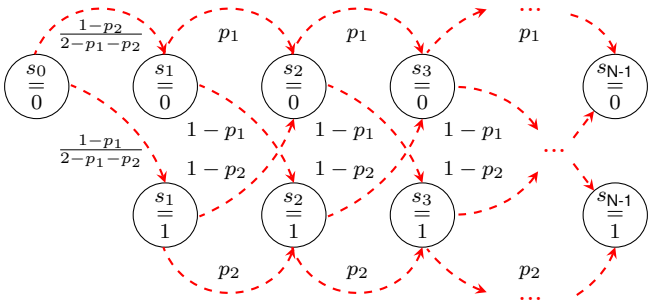


Figura 18 – Diagrama do modelo de esparsidade em blocos para geração da resposta ao impulso.

consiste num critério de seleção que requer $N + BL + 2$ multiplicações. A Eq. (114) pode ser generalizada para obter o método proposto SPU-BC-BS-NLMS:

$$\begin{aligned} \boldsymbol{w}_{\mathcal{I}_B}(k+1) &= \boldsymbol{w}_{\mathcal{I}_B}(k) + \gamma \boldsymbol{g}[\boldsymbol{w}_{\mathcal{I}_B}(k)] + \beta \frac{\boldsymbol{x}_{\mathcal{I}_B}(k) e(k)}{\|\boldsymbol{x}_{\mathcal{I}_B}(k)\|^2} \\ &\quad + \beta \hat{\sigma}_\eta^2 \frac{\boldsymbol{w}_{\mathcal{I}_B}(k)}{\|\boldsymbol{x}_{\mathcal{I}_B}(k)\|^2}. \end{aligned} \tag{116}$$

Para exemplificar o compromisso introduzido pela técnica SPU aplicada ao algoritmo BC-BS-NLMS, seja um cenário cujas funções de transferência a serem estimadas são amostradas de um modelo de Markov-Gauss (M-G) $\mathcal{M}(N, p_1, p_2, \sigma_s^2)$, o qual dá origem a uma ampla faixa de sistemas com esparsidade em bloco [79]. Com este modelo, a resposta ao impulso ideal $\boldsymbol{w}^\star$ é computada a partir de dois passos. No primeiro deles, um processo de Markov de primeira ordem é responsável por gerar o conjunto de índices dos coeficientes nulos e dos não nulos (veja Fig. 18), de acordo com a seguinte regra:

$$P\{s_j = 0 | s_{j-1} = 0\} = p_1, \tag{117}$$

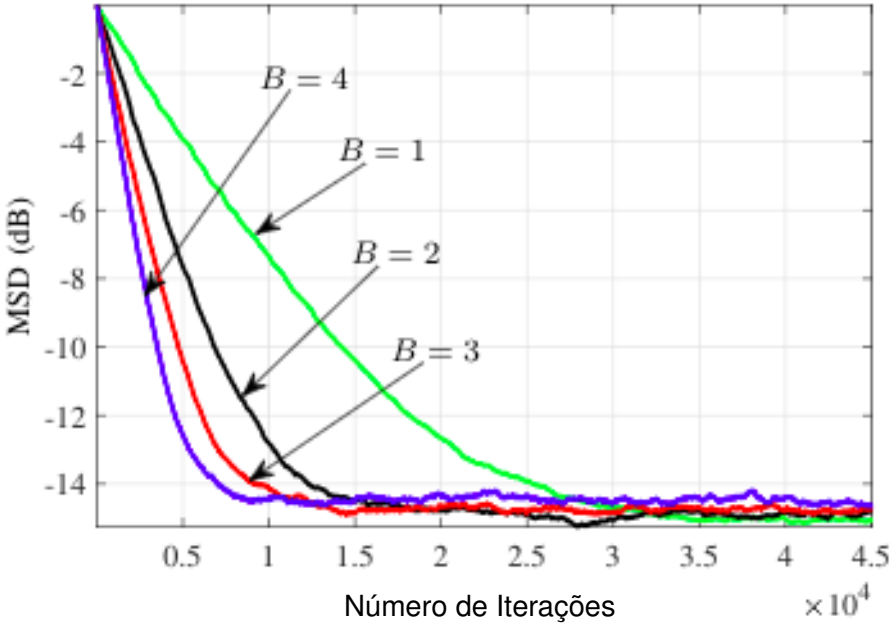
$$P\{s_j \neq 0 | s_{j-1} \neq 0\} = p_2, \tag{118}$$

onde p_1 e p_2 são parâmetros ajustáveis. Note que $(1 - p_2)$ deve ser bem maior do que $(1 - p_1)$ para garantir uma resposta ao impulso esparsa em blocos [79].

Após a determinação do conjunto de índices de coeficientes não nulos, as amplitudes dos coeficientes não nulos são amostrados a partir de uma distribuição gaussiana de média zero e variância σ_s^2. Matematicamente, pode-se escrever:

$$w_i^\star = \begin{cases} 0, & \text{se } s_i = 0 \\ r, & \text{se } s_i = 1 \end{cases}, \tag{119}$$

onde r é uma variável aleatória gaussiana com variância σ_s^2. No resultado explicitado a seguir, a configuração $(p_1; p_2; \sigma_s^2) = (0,999; 0,9; 1)$ foi adotada.



A Figura 5.3 apresenta o MSD como uma função do número de iterações para valores de $B \in 1, 2, 3, 4$ e $M' = 4$. Os seguintes parâmetros foram escolhidos: $\beta_{\text{BCBSNLMS}} = 0,4$, $N = 800$, $\rho = 1$, $\sigma_\nu^2 = 2 \times 10^{-2}$, $\sigma_x^2 = 1$, $\hat{\sigma}_\eta^2 = \sigma_\eta^2 = 10^{-1}$. Todos os resultados foram obtidos a partir de 100 ensaios indepedentes de Monte Carlo. Quando $B = 4$, o algoritmo se degenera no BC-BS-NLMS. Ademais, cumpre ressaltar que as versões SPU permitem trocar taxa de convergência por complexidade computacional.

5.4- Comentários Parciais

Este capítulo forneceu uma visão panorâmica da técnica de compensação de viés, normalmente aplicada ao problema de identificação de sistemas. Uma contribuição desta dissertação para este tipo de problema foi descrita, a qual reduz a complexidade computacional de uma técnica de compensação de viés por meio de uma estratégia de atualização seletiva. O próximo capítulo detalha a principal contribuição desta dissertação, a qual aplica a técnica de compensação de viés ao algoritmo FX-LMS, o qual é utilizado para problemas de cancelamento ativo de ruído.

6- Algoritmo FX-LMS com Compensação de Viés

6.1- Introdução

O controle de ruído ativo é um campo em expansão que requer uma síntese adequada de perturbações secundárias. Infelizmente, a maioria dos esquemas de cancelamento de ruído não leva em consideração o fato de que o sinal de entrada que aciona o filtro adaptativo pode ser ruidoso [57], [83]. Neste capítulo, é teoricamente demonstrado que as perturbações de ruído nos dados de excitação degradam o desempenho do algoritmo FX-LMS. Além disso, um método que compensa tal problema é desenvolvido e uma análise estocástica de primeira ordem do algoritmo resultante é realizada. Os resultados que serão apresentados futuramente revelam que o esquema proposto supera o algoritmo FX-LMS clássico, mesmo quando a variância do ruído aditivo na entrada não é estimada com precisão.

Os métodos tradicionais de cancelamento de ruído baseados em tecnologia de controle passivo, como a absorção de som, às vezes não atendem aos requisitos rigorosos de configurações práticas [84]. Esse fato explica o crescente interesse que as técnicas de controle ativo de ruído (ANC) têm recebido como uma alternativa poderosa para abordar a poluição sonora ou a rejeição de vibração [85].

O algoritmo FX-LMS é um dos esquemas adaptativos mais populares para ANC, utilizado quando os efeitos do caminho secundário S não podem ser ignorados. Ele emprega uma linha de retardo de derivação transversal que aplica os princípios de superposição e interferência destrutiva com a finalidade de obter a mitigação da perturbação desejada (ou seja, vibração ou ruído).

A arquitetura de controle retroalimentada do FX-LMS se baseia nos dados de excitação $u(k)$ que fornecem informações sobre as perturbações que devem ser mitigadas no sensor de erro. A figura 5 mostra que existe um canal físico representado por S no caminho do erro (ou seja, entre a saída do filtro e a resposta desejada disponível). Além disso, o sinal de entrada de canal único $u(k)$ deve ser filtrado por $\hat{S}$, um modelo do filtro secundário. Observe que, como esse modelo é uma estimativa de S, pode ser impreciso.

6.2- FX-LMS com Entrada Ruidosa

Considere a seguir que a aquisição de $u(k)$ é um processo ruidoso (ver Fig. 6), de modo que o sinal de entrada do esquema adaptativo é dado por

$$x(k) = u(k) + \eta(k), \tag{120}$$

onde $\eta(k)$ é um ruído branco estacionário de média zero com variância σ_η^2.

Em (120), as seguintes definições são necessárias

$$y(k) \triangleq \boldsymbol{w}^T(k)\boldsymbol{x}(k), \tag{121}$$

$$y_{\mathsf{f}}(k) \triangleq \sum_{i=0}^{M-1} s_i y(k-i) = \sum_{i=0}^{M-1} s_i \boldsymbol{x}^T(k-i)\boldsymbol{w}(k-i), \tag{122}$$

$$e(k) \triangleq d(k) - y_{\mathsf{f}}(k) + \nu(k), \tag{123}$$

que resulta na equação de atualização do FX-LMS *com entrada ruidosa*:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta e(k)\boldsymbol{x}_{\mathsf{f}}(k), \tag{124}$$

onde $\boldsymbol{x}_{\mathsf{f}}(k) \triangleq \sum_{i=0}^{M-1} \hat{s}_i \boldsymbol{u}(k-i) + \sum_{i=0}^{M-1} \hat{s}_i \boldsymbol{\eta}(k-i)$.

6.3- Análise de Primeira Ordem do FX-LMS com Entrada Ruidosa

Nesta seção, uma análise estocástica de primeira ordem do FX-LMS com entrada ruidosa é realizada, a qual demonstra a emergência de um viés assintótico.

Considere $\tilde{\boldsymbol{w}}(k) \triangleq \boldsymbol{w}_{\mathsf{opt}} - \boldsymbol{w}(k)$ como vetor de desvio. A partir (124), é possível estabelecer a seguinte recursão para o desvio:

$$\tilde{\boldsymbol{w}}(k+1) = \tilde{\boldsymbol{w}}(k) - \beta e(k)\boldsymbol{x}_{\mathsf{f}}(k), \tag{125}$$

onde o erro pode ser escrito em função dos desvios:

$$\begin{aligned} e(k) &= \sum_{i=0}^{M-1} s_i \boldsymbol{u}^T(k-i)\tilde{\boldsymbol{w}}(k-i) - \sum_{i=0}^{M-1} s_i \boldsymbol{u}^T(k-i)\boldsymbol{w}_{\text{opt}} \\ &\quad + \sum_{i=0}^{M-1} s_i \boldsymbol{\eta}^T(k-i)\tilde{\boldsymbol{w}}(k-i) - \sum_{i=0}^{M-1} s_i \boldsymbol{\eta}^T(k-i)\boldsymbol{w}_{\text{opt}} \\ &\quad + d(k) - \nu(k). \end{aligned} \tag{126}$$

A fim de efetuar as simplificações necessárias manipularmos matematicamente, consideramos a hipótese da independência, dando origem, após a aplicação do operador de valor esperado, a:

$$\begin{aligned} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k+1)\right] = \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] &- \beta \sum_{i=0}^{M-1} \hat{s}_i \boldsymbol{p}_i \\ &+ \beta \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} \hat{s}_i s_j \boldsymbol{R}_{j-i}\left[\boldsymbol{w}_{\text{opt}} - \mathbb{E}[\tilde{\boldsymbol{w}}(k-j)]\right] \\ &+ \beta \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} \hat{s}_i s_j \boldsymbol{R}^{\eta}_{j-i}\left[\boldsymbol{w}_{\text{opt}} - \mathbb{E}[\tilde{\boldsymbol{w}}(k-j)]\right], \end{aligned} \tag{127}$$

onde $\boldsymbol{R}^{\eta}_{j-i} \triangleq \mathbb{E}\left[\boldsymbol{\eta}(k-i)\boldsymbol{\eta}(k-j)\right]$.

Embora apropriada para análise transitória de primeira ordem, a recursão (127) também pode auxiliar o entendimento do ponto de operação em estado estacionário. Supondo que o algoritmo converge, ou seja,

$$\lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k)\right] = \lim_{k\to\infty} \mathbb{E}\left[\tilde{\boldsymbol{w}}(k-i)\right], \tag{128}$$

e que a estimativa do caminho secundário é perfeita (*ou seja*, $\hat{s}_i = s_i$ para $i \in \{0, 1, \ldots, M-1\}$), pode ser mostrado que

$$\lim_{k\to\infty} \mathbb{E}[\tilde{\boldsymbol{w}}(k)] = \underbrace{\left[\tilde{\boldsymbol{R}} + \tilde{\boldsymbol{R}}^{\eta}\right]^{-1}}_{\triangleq \boldsymbol{\Gamma}} \tilde{\boldsymbol{R}}^{\eta} \boldsymbol{w}_{\text{opt}}, \tag{129}$$

onde $\tilde{\boldsymbol{R}}^{\eta} \triangleq \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \boldsymbol{R}^{\eta}_{j-i}$.

A equação (129) mostra que, sob as premissas adotadas, o procedimento de estimação do algoritmo FX-LMS com entrada ruidosa é assintoticamente enviesado, uma vez que $\lim_{k\to\infty} \mathbb{E}[\tilde{\boldsymbol{w}}(k)] \neq \boldsymbol{0}$. Na seção seguinte, um procedimento para mitigar tal pro-

blema é teoricamente derivado, o que dá origem ao algoritmo FX-LMS com compensação de viés (BC-FX-LMS) proposto.

6.4- Algoritmo Proposto BC-FX-LMS

O algoritmo BC-FX-LMS proposto atualiza $\boldsymbol{w}(k)$ utilizando

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta e(k)\boldsymbol{x}_{\mathsf{f}}(k) + \boldsymbol{\psi}(k), \tag{130}$$

onde o vetor $\boldsymbol{\psi}(k)$ é um termo novo e adicional de correção, introduzido com o propósito de mitigar o viés. Usando os mesmos passos que levaram a (129), o seguinte resultado pode ser estabelecido:

$$\lim_{k\to\infty} \mathbb{E}[\tilde{\boldsymbol{w}}(k)] = \boldsymbol{\Gamma}\left[\sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \boldsymbol{R}^{\eta}_{j-i}\mathbb{E}\left[\boldsymbol{w}(k-i)\right] - \frac{\boldsymbol{\psi}(k)}{\beta}\right], \tag{131}$$

de modo que o valor esperado do termo de correção $\boldsymbol{\psi}(k)$, se um estimador não enviesado for desejado, deve satisfazer a seguinte restrição:

$$\mathbb{E}\left[\boldsymbol{\psi}(k)\right] = \beta\sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \boldsymbol{R}^{\eta}_{i-j}\mathbb{E}\left[\boldsymbol{w}(k-i)\right], \tag{132}$$

que pode ser imposta usando os pesos adaptativos (que são *observáveis*) escolhendo:

$$\boldsymbol{\psi}(k) = \beta\sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \hat{\boldsymbol{R}}^{\eta}_{j-i}\boldsymbol{w}(k-i), \tag{133}$$

onde $\hat{\boldsymbol{R}}^{\eta}_{j-i}$ é uma estimativa de $\boldsymbol{R}^{\eta}_{j-i}$. A aplicação de (133) em (130) é suficiente para especificar completamente o algoritmo BC-FX-LMS proposto.

Assumindo que o ruído $\eta(k)$ é de média zero e estatisticamente independente das demais variáveis aleatórias (assim como $\nu(k)$), e que $\hat{\sigma}^2_{\eta}$ é uma estimativa de sua atual variância σ^2_{η}, o (n,m)-elemento $r_{j-i}(m,n)$ das matrizes $\hat{\boldsymbol{R}}^{\eta}_{j-i}$ pode ser aproximado como

$$r_{j-i}(m,n) = \begin{cases} 0, & \text{se } j-i+n-m \neq 0 \\ \hat{\sigma}^2_{\eta}, & \text{caso contrário} \end{cases}. \tag{134}$$

Assim, utilizando-se o vetor $\boldsymbol{\psi}(k)$ definido em (133), obtemos o algoritmo BC-FX-LMS proposto:

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \beta e(k)\boldsymbol{x}_{\mathsf{f}}(k) + \beta \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} s_i s_j \hat{\boldsymbol{R}}_{j-i}^{\eta}\boldsymbol{w}(k-i). \tag{135}$$

Observe que o novo termo demandado pelo algoritmo proposto exige uma estimativa da variância de $\eta(k)$; tal estimativa é alheia ao escopo de tal trabalho, pois requer uma análise de segunda ordem. Iremos verificar, por meio de simulações, a sensibilidade do desempenho do algoritmo perante imprecisões desta estimativa.

Assumindo-se estacionariedade dos sinais envolvidos, temos que as matrizes

$$\boldsymbol{\Phi}(i) \triangleq \beta \sum_{j=0}^{M-1} s_j \hat{\boldsymbol{R}}_{j-i}^{\eta} \tag{136}$$

são invariantes no tempo, de modo que o termo adicional de correção de viés $\boldsymbol{\psi}(k)$ pode ser expresso por

$$\boldsymbol{\psi}(k) = \sum_{i=0}^{M-1} \boldsymbol{\Phi}(i)\boldsymbol{w}(k-i), \tag{137}$$

o que demanda, no máximo, MN^2 multiplicações e $M(N-1)$ adições (por iteração) a mais do que o algoritmo FX-LMS.

De modo a facilitar a consulta do leitor, a Tabela 4 apresenta as definições dos símbolos mais importantes dos algoritmos mencionados nesta dissertação.

Símbolo	Definição	Algoritmos
$d(k)$	$d(k) \triangleq [\boldsymbol{w}^{\star}]^T \boldsymbol{u}(k)$	LMS, FX-LMS
$\overline{d}(k)$	$\overline{d}(k) \triangleq d(k) + \nu(k) = [\boldsymbol{w}^{\star}]^T \boldsymbol{u}(k) + \nu(k)$	LMS, FX-LMS
$y'(k)$	$y'(k) = [\boldsymbol{w}(k)]^T \boldsymbol{u}(k)$	LMS, FX-LMS
$\overline{e}(k)$	$\overline{e}(k) \triangleq \overline{d}(k) - y'(k)$	LMS, FX-LMS
$y'_{\mathsf{f}}(k)$	$y'_{\mathsf{f}}(k) \triangleq \sum_{i=0}^{M-1} s_i y'(k-i)$	FX-LMS
$x(k)$	$x(k) = u(k) + \eta(k)$	N-FX-LMS, BC-FX-LMS
$y(k)$	$y(k) = [\boldsymbol{w}(k)]^T \boldsymbol{x}(k)$	N-FX-LMS, BC-FX-LMS
$y_{\mathsf{f}}(k)$	$y_{\mathsf{f}}(k) \triangleq \sum_{i=0}^{M-1} s_i y(k-i)$	N-FX-LMS, BC-FX-LMS
$e(k)$	$e(k) \triangleq \overline{d}(k) - y_{\mathsf{f}}(k)$	N-FX-LMS, BC-FX-LMS

Tabela 4 – Definições dos principais sinais utilizados pelos diferentes algoritmos desta dissertação.

7- Resultados e Discussões

Nas simulações seguintes foram utilizados 4 tipos de cenários diferentes para as figuras 19, 20 e 21, cada qual com diferentes parâmetros que foram alterados a fim de averiguar o comportamento do algoritmo proposto em diversos ambientes. Alguns parâmetros foram mantidos constantes como a função de transferência do filtro secundário $S(z) = 1 + 0,8z^{-1}$. O sinal $u(k)$ é obtido filtrando um ruído gaussiano branco de variância unitária de média zero com a função de transferência $B(z) = 1 - 0,2z^{-1}$. Entretanto, no cenário 3, o filtro de colorimento foi removido com o intuito de verificar o desempenho com um ruído branco. Os sinais de ruído $\nu(k)$ and $\eta(k)$ são amostrados a partir de processos gaussianos brancos de média zero, suas variâncias foram mantidas constantes nos três primeiros cenários, já no cenário 4 foram utilizados valores de maior grandeza a fim de verificar a atuação num patamar mais ruidoso.

O sinal de referência é gerado de acordo com o seguinte modelo de parâmetros afins:

$$d(k) = [\boldsymbol{w}^{\star}]^{T} \boldsymbol{x}(k), \tag{138}$$

onde $\boldsymbol{w}^{\star} \in \mathbb{R}^{L}$ e o i-ésimo coeficiente de $\boldsymbol{w}^{\star}$ é

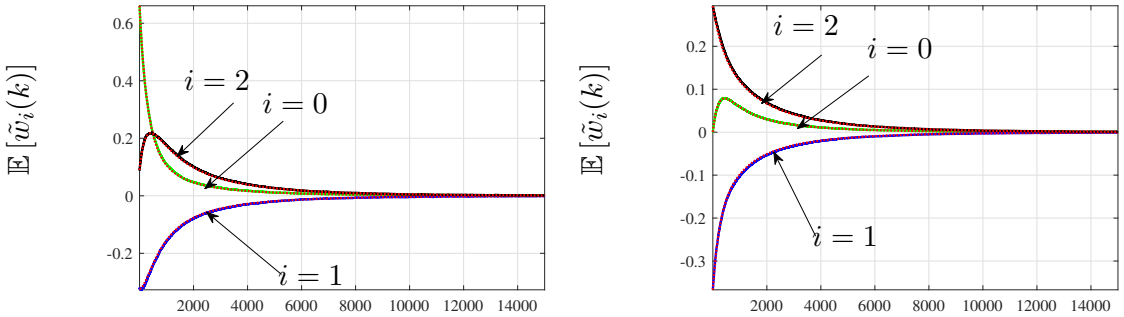
$$w_i^{\star} = \cos\left[\gamma\pi(i+1)\right] \exp[-\gamma(i+1)], \text{ para } 0 \leq i \leq L-1. \tag{139}$$

A Tabela 5 informa os parâmetros escolhidos em cada cenário. A escolha destes se deu através do possível grau de impacto que os mesmos podem causar no sistema. Portanto, os parâmetros cuja influência foi considerada mais impactante foram os alternados.

Simulações	L, N	γ	$B(z)$	σ_{ν}^2	σ_{η}^2
Cenário 1	20	$0,2$	$1 - 0,2z^{-1}$	10^{-6}	10^{-3}
Cenário 2	100	$0,5$	$1 - 0,2z^{-1}$	10^{-6}	10^{-3}
Cenário 3	20	1	1	10^{-6}	10^{-3}
Cenário 4	20	$0,2$	$1 - 0,2z^{-1}$	10^{-1}	10^{-1}

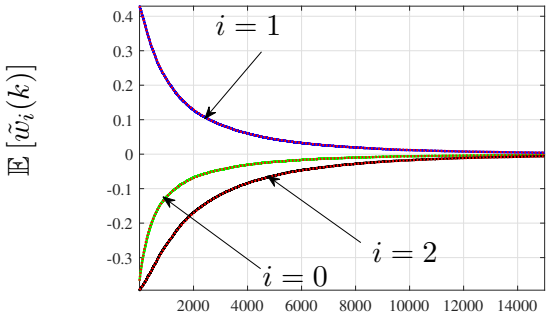
Tabela 5 – Simulações divididas por cenários e seus respectivos parâmetros.

Tendo em vista a análise da resposta do algoritmo para um filtro com dimensões diferentes foram variados os parâmetros L e N que denotam os comprimentos de $\boldsymbol{w}^\star$ e $\boldsymbol{w}(k)$ respectivamente, assim como o fator γ que modifica a função de transferência (139) com o objetivo de demonstrar o viés em funções dissemelhantes.

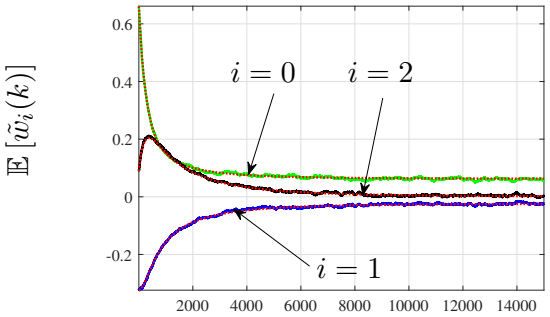


(a) Cenário 1

(b) Cenário 2



(c) Cenário 3



(d) Cenário 4

Figura 19 – Evolução média dos três primeiros coeficientes de desvio, com 10 ensaios independentes de Monte Carlo. O modelo teórico (127) é representado por linhas vermelhas tracejadas.

A Fig. 19 permite comparar a evolução empírica dos três primeiros coeficientes de desvio com o modelo teórico proposto (127). Pode-se notar que este modelo descreve com precisão o comportamento médio dos coeficientes de desvio, mesmo quando o caminho secundário é modelado de forma imprecisa e quando variamos γ nos cenários 2 e 3. Já no cenário 4 notamos uma maior instabilidade nas curvas por conta do aumento

na variância dos ruídos. A Fig. 20 descreve o comportamento médio assintótico dos três primeiros coeficientes de desvio do FX-LMS com entrada ruidosa (isto é, o N-FX-LMS), que é bem aproximado por (129). Esta figura confirma o fato de que o algoritmo FX-LMS é assintoticamente enviesado quando sua entrada é corrompida por ruído aditivo. Um movimento browniano do desvio médio é observado devido ao uso de um número finito de ensaios independentes de Monte Carlo. Nas figuras 19 e 20 foram utilizados os seguintes parâmetros: $\beta = 2,5\times10^{-3}$ e para a planta do secundário estimada $\hat{S}(z) = 1,02+0,77z^{-1}$.

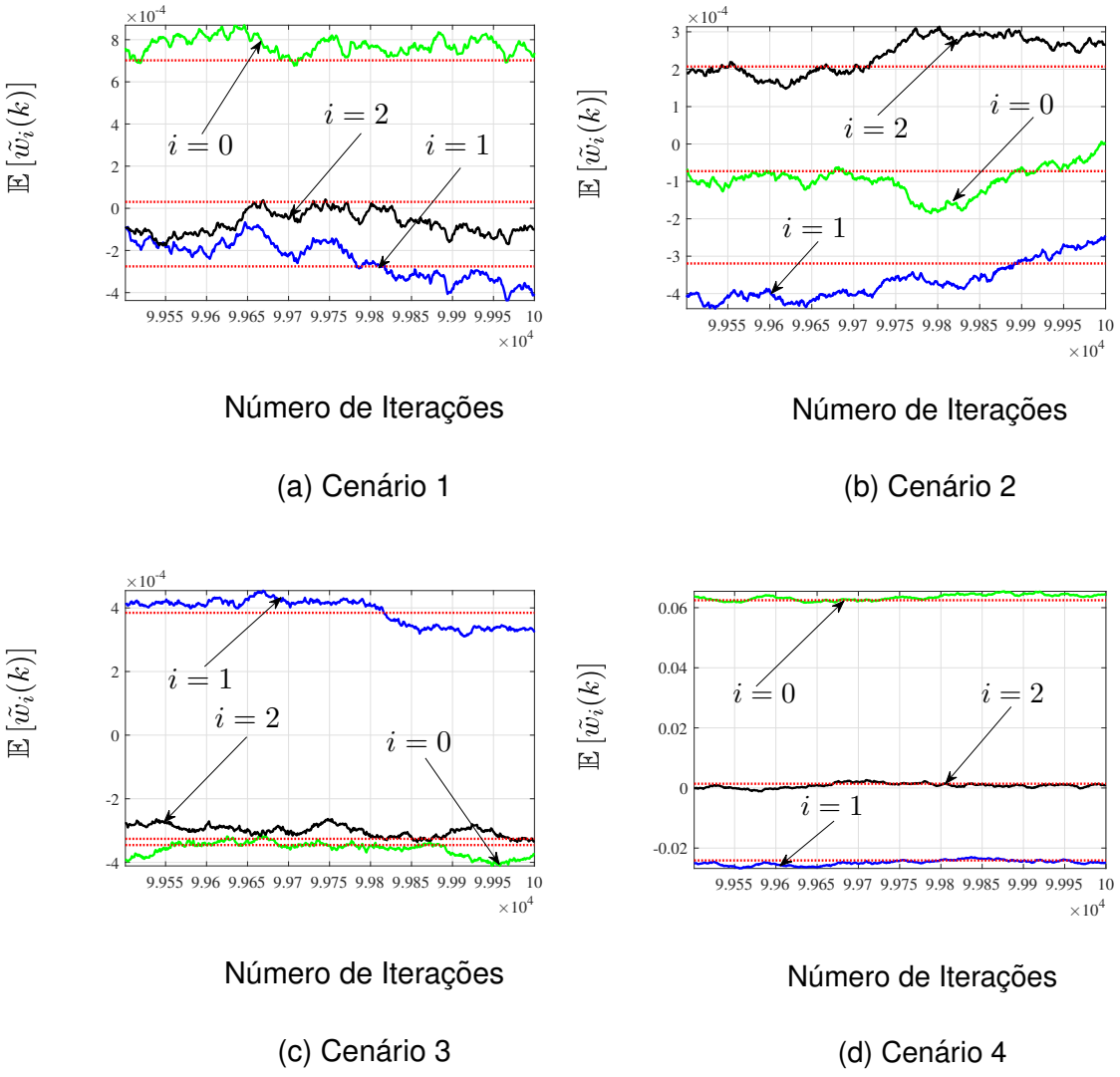


Figura 20 – Comportamento de estado estacionário dos três primeiros coeficientes de desvio com 10^5 testes independentes de Monte Carlo. As linhas vermelhas tracejadas horizontais são estimativas teóricas obtidas de (129).

A fim de avaliar o desempenho quadrático médio, considere a distorção quadrada

média (MSD) definida como

$$\text{MSD}(k) \triangleq \mathbb{E}\left[\|\boldsymbol{w}_{\text{opt}} - \boldsymbol{w}(k)\|^2\right] .- \tag{140}$$

A capacidade do BC-FX-LMS proposto para abordar a tendência de estado estacionário pode ser avaliada pela Fig. 21, onde $\beta = 2 \times 10^{-3}$ e $\hat{S}(z) = S(z)$. Observe que apesar de não apresentar ganhos substanciais nos três primeiros cenários, o algoritmo proposto é capaz de obter uma melhoria de 6 dB no MSD de estado estacionário no cenário mais ruidoso (cenário 4), o que pode ser atribuído ao recurso de cancelamento de polarização do algoritmo BC-FX-LMS proposto.

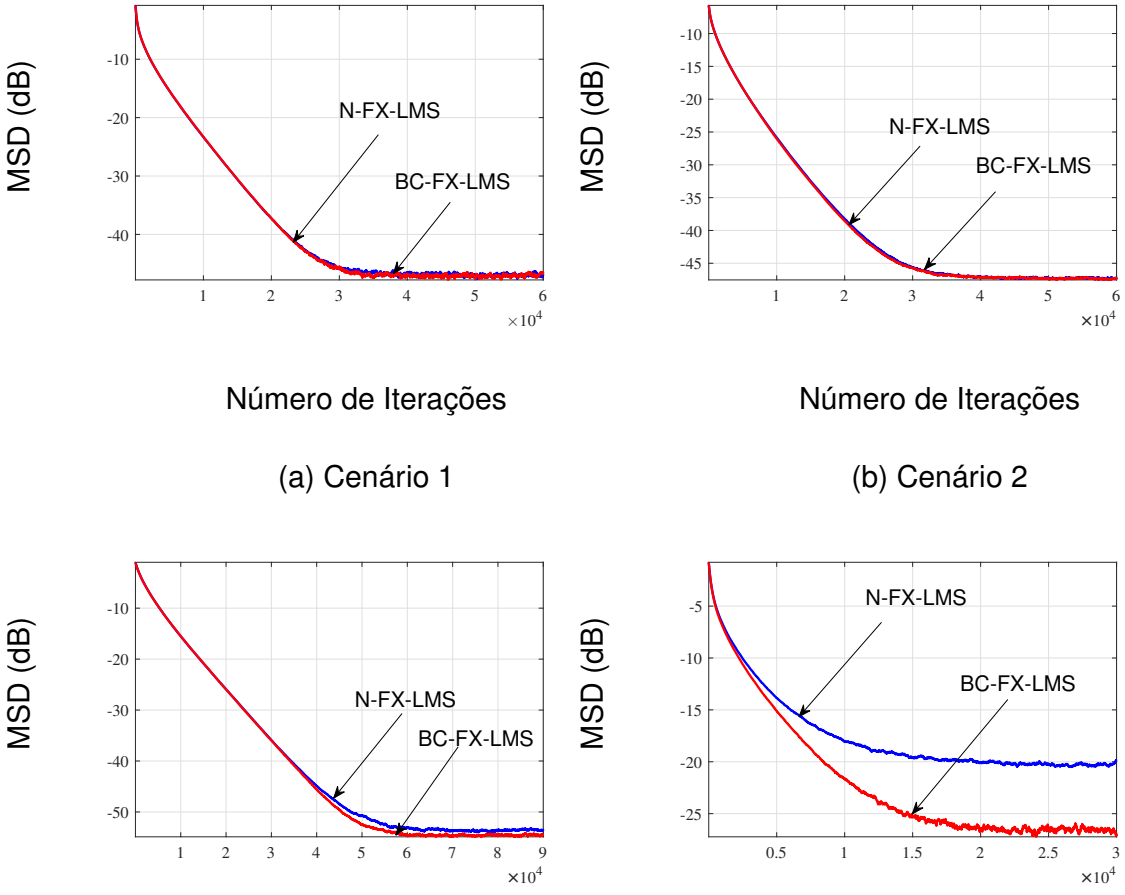


Figura 21 – Evolução MSD do FX-LMS com algoritmos de entrada ruidosa (em azul) e BC-FX-LMS (em vermelho).

Um ponto fraco do método proposto é que ele requer a estimativa de σ_η^2. Como na prática sua estimativa $\hat{\sigma}_\eta^2$ pode estar incorreta, é importante avaliar a sensibilidade do desempenho do BC-FX-LMS comparado a tal estimativa. Considere a seguir que $\hat{\sigma}_\eta^2 = \alpha\sigma_\eta^2$, de forma que a variância de η seja estimada corretamente apenas quando $\alpha = 1$. A Fig. 22 mostra o desempenho em regime permanente do algoritmo BC-FX-LMS em relação a α, usando o mesmo valor de β e $\hat{S}(z) = S(z)$ como na Fig. 21. Nesta figura foram utilizados os parâmetros do cenário 4. Pode-se notar que o algoritmo desenvolvido é vantajoso na configuração considerada, mesmo para grandes erros de estimativa (como $\alpha = 1.7$ ou $\alpha = 0.2$), portanto, uma estimativa precisa de σ_η^2 não é crucial para a obtenção de melhoria de desempenho com o algoritmo proposto.

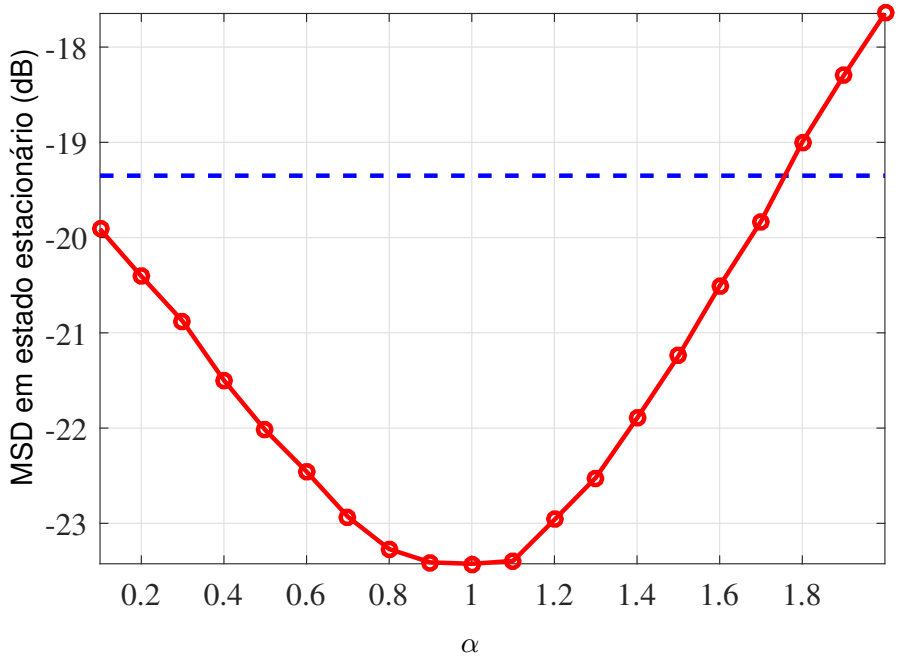


Figura 22 – MSD em estado estacionário do algoritmo BC-FX-LMS em relação a α, onde $\hat{\sigma}_\eta^2 = \alpha\sigma_\eta^2$. A linha azul tracejada é o MSD assintótico obtido pelo algoritmo FX-LMS com entrada ruidosa.

8- Conclusões e Trabalhos Futuros

Este trabalho insere-se no campo do cancelamento ativo de ruído ou de vibrações. O algoritmo FX-LMS é o algoritmo clássico para este problema, quando há ausência de *feedback* entre o sinal sintetizado e o sinal de excitação do filtro adaptativo.

Após efetuar uma varredura bibliográfica na área, foi observado que a literatura desconsidera a possibilidade de o sinal de excitação se encontrar corrompido por ruído aditivo. Tal ruído pode ocorrer numa ampla variedade de contextos, tal como emergir do próprio processo de conversão analógico/digital. A primeira contribuição deste trabalho foi investigar teoricamente o impacto que tal ruído apresenta no estimador do algoritmo FX-LMS. Como resultado, deduziu-se o advento de um viés assintótico, o que significa que este algoritmo encontra-se polarizado em regime permanente, quando o sinal de excitação encontra-se imerso em ruído.

Como segunda contribuição deste trabalho, projetou-se um algoritmo novo, baseado no FX-LMS, com a inclusão de um termo de correção ($\boldsymbol{\psi}(k)$) projetado para mitigar o problema da polarização. Tal termo de correção depende do fator de aprendizagem, de estimativas do secundário, de estimativas da variância do ruído aditivo $\eta(k)$ (supostamente branco) e dos coeficientes adaptativos. Quando as estimativas que este termo adicional requer são perfeitas, ocorre a eliminação do viés. Verificou-se, por fim, que mesmo estimativas de σ_η^2 bastante imprecisas podem, ainda, resultar em melhorias de desempenho, tornando vantajoso o algoritmo BC-FX-LMS proposto.

Como possível desdobramento futuro derivado deste trabalho, elencamos o desenvolvimento de uma análise estocástica de segunda ordem (capaz de prever, entre outras coisas, a evolução do MSE e do MSD) do novo algoritmo proposto, assim como sua extensão para algoritmos de cancelamento de *feedback* acústico, muito utilizados em aparelhos de auxílio à audição (veja, por exemplo, [86]). Uma outra projeção futura seria utilizar alguma métrica para estimar com precisão o canal secundário.

Referências Bibliográficas

[1] C. Chang e K. Shyu, “Active noise cancellation with a fuzzy adaptive filtered-X algorithm”, *IEE Proceedings - Circuits, Devices and Systems*, v. 150, n. 5, pp. 416–422, 2003.

[2] J. Khan, M. Ketzel, K. Kakosimos, M. Sørensen e S. S. Jensen, “Road traffic air and noise pollution exposure assessment – A review of tools and techniques”, *Science of The Total Environment*, v. 634, pp. 661–676, 2018.

[3] J. O. Klompmaker, N. A. Janssen, L. D. Bloemsma, U. Gehring, A. H. Wijga, C. van den Brink, E. Lebret, B. Brunekreef e G. Hoek, “Residential surrounding green, air pollution, traffic noise and self-perceived general health”, *Environmental Research*, v. 179, pp. 1–9, 2019.

[4] T. B. de Souza, K. C. Alberto e S. A. Barbosa, “Evaluation of noise pollution related to human perception in a university campus in Brazil”, *Applied Acoustics*, v. 157, pp. 1–11, 2020.

[5] A. Oudin, “Short review: Air pollution, noise and lack of greenness as risk factors for Alzheimer’s disease- epidemiologic and experimental evidence”, *Neurochemistry International*, v. 134, pp. 1–6, 2020.

[6] H. Jang, Y. M. Kim, J. Chung e K. S. Lee, “KTX’s interior noise reduction performance comparison for each section using multichannel active noise control”, em *2012 12th International Conference on Control, Automation and Systems*, 2012, pp. 1265–1270.

[7] A. Arjunan, “Targeted sound attenuation capacity of 3D printed noise cancelling waveguides”, *Applied Acoustics*, v. 151, pp. 30–44, 2019.

[8] D. J. O’Boy, “Automotive wheel and tyre design for suppression of acoustic cavity noise through the incorporation of passive resonators”, *Journal of Sound and Vibration*, v. 467, pp. 1–14, 2020.

[9] S. Mitra, *Digital Signal Processing: A Computer-based Approach*, sér. McGraw-Hill series in electrical and computer engineering. McGraw-Hill, 2011.

[10] J. Chen, C. Lu, C. Hsia e Y. Chen, “A low noise amplifier employing noise canceling technique for ultrasound system applications”, em *2012 International Symposium on Instrumentation Measurement, Sensor Network and Automation (IMSNA)*, vol. 1, 2012, pp. 35–38.

[11] P. Diniz, E. da Silva e S. Netto, *Processamento Digital de Sinais - 2.ed.: Projeto e Análise de Sistemas*. Bookman Editora, 2014.

[12] I.-T. S. G. 15, “Digital Network Echo Cancellers (Recommendation)”, ITU-T, rel. técn. G.168, 2004.

[13] A. Carini e S. Malatini, “Optimal Variable Step-Size NLMS Algorithms With Auxiliary Noise Power Scheduling for Feedforward Active Noise Control”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 16, n. 8, pp. 1383–1395, 2008.

[14] S. M. Kuo e D. Vijayan, “A secondary path modeling technique for active noise control systems”, *IEEE Transactions on Speech and Audio Processing*, v. 5, n. 4, pp. 374–377, 1997.

[15] Ming Zhang, Hui Lan e Wee Ser, “Cross-updated active noise control system with online secondary path modeling”, *IEEE Transactions on Speech and Audio Processing*, v. 9, n. 5, pp. 598–602, 2001.

[16] ——, “A robust online secondary path modeling method with auxiliary noise power scheduling strategy and norm constraint manipulation”, *IEEE Transactions on Speech and Audio Processing*, v. 11, n. 1, pp. 45–53, 2003.

[17] M. T. Akhtar, M. Abe e M. Kawamata, “A new variable step size LMS algorithm-based method for improved online secondary path modeling in active noise control systems”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 2, pp. 720–726, 2006.

[18] S. M. Kuo e D. R. Morgan, *Active noise control systems*. Wiley, New York, 1996, vol. 4.

[19] V. Solo, “LMS: Past, Present and Future”, em *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 7740–7744.

[20] P. Diniz e et al., *Signal Processing Theory and Machine Learning*. Elsevier, 2014, vol. 1, pp. 619–761.

[21] P. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, sér. Kluwer international series in engineering and computer science. Springer US, 2008.

[22] S. Haykin, *Adaptive Filter Theory*, sér. Prentice-Hall information and system sciences series. Prentice Hall, 1996.

[23] E. C. Menguc e N. Acir, "Lyapunov Stability Theory Based Adaptive Filter Algorithm for Noisy Measurements", em *2013 UKSim 15th International Conference on Computer Modelling and Simulation*, IEEE, 2013.

[24] L. Tan e J. Jiang, "Adaptive Filters and Applications", em *Digital Signal Processing*, Elsevier, 2019, pp. 421–474.

[25] M. Lazaro-Gredilla, L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal e J. Arenas-Garcia, "Adaptively Biasing the Weights of Adaptive Filters", *IEEE Transactions on Signal Processing*, v. 58, n. 7, pp. 3890–3895, 2010.

[26] M. V. S. Lima, T. N. Ferreira, W. A. Martins e P. S. R. Diniz, "Sparsity-Aware Data-Selective Adaptive Filters", *IEEE Transactions on Signal Processing*, v. 62, n. 17, pp. 4557–4572, 2014.

[27] S. A. Ghauri e M. F. Sohail, "System identification using LMS, NLMS and RLS", em *2013 IEEE Student Conference on Research and Developement*, IEEE, 2013.

[28] B. Widrow e M. E. Hoff, "Adaptive Switching Circuits", em *1960 IRE WESCON Convention Record, Part 4*, New York: IRE, 1960, pp. 96–104.

[29] K. da S. Olinto, D. B. Haddad e M. R. Petraglia, "Transient analysis of l0-LMS and l0-NLMS algorithms", *Signal Processing*, v. 127, pp. 217–226, 2016.

[30] A. H. Sayed, *Adaptive Filters*. John Wiley & Sons, Inc., 2008.

[31] P. Lara, F. Igreja, L. D. T. J. Tarrataca, D. B. Haddad e M. R. Petraglia, "Exact Expectation Evaluation and Design of Variable Step-Size Adaptive Algorithms", *IEEE Signal Processing Letters*, v. 26, n. 1, pp. 74–78, 2019.

[32] L. C. Resende, D. B. Haddad e M. R. Petraglia, "A Variable Step-Size NLMS Algorithm with Adaptive Coefficient Vector Reusing", em *2018 IEEE International Conference on Electro/Information Technology (EIT)*, 2018, pp. 0181–0186.

[33] P. Lara, L. D. Tarrataca e D. B. Haddad, "Exact expectation analysis of the deficient-length LMS algorithm", *Signal Processing*, v. 162, pp. 54 –64, 2019.

[34] P. Lara, D. B. Haddad e L. Tarrataca, “Advances on the analysis of the LMS algorithm with a colored measurement noise”, *Signal, Image and Video Processing*, 2019.

[35] J. Homer, R. R. Bitmead e I. Mareels, “Quantifying the effects of dimension on the convergence rate of the LMS adaptive FIR estimator”, *IEEE Transactions on Signal Processing*, v. 46, n. 10, pp. 2611–2615, 1998.

[36] M. T. M. Silva e V. H. Nascimento, “Improving the Tracking Capability of Adaptive Filters via Convex Combination”, *IEEE Transactions on Signal Processing*, v. 56, n. 7, pp. 3137–3149, 2008.

[37] J. D. Griesbach, M. R. Lightner e D. M. Etter, “Mean square error analysis of nonuniform subband adaptive filters for system modeling”, *IEEE Transactions on Signal Processing*, v. 53, n. 2, pp. 550–563, 2005.

[38] S. Gazor, “Prediction in LMS-type adaptive algorithms for smoothly time varying environments”, *IEEE Transactions on Signal Processing*, v. 47, n. 6, pp. 1735–1739, 1999.

[39] B. Widrow, J. M. McCool, M. G. Larimore e C. R. Johnson, “Stationary and nonstationary learning characteristics of the LMS adaptive filter”, *Proceedings of the IEEE*, v. 64, n. 8, pp. 1151–1162, 1976.

[40] A. Kapur, M. K. Varanasi e D. Das, “Noncoherent MMSE multiuser receivers and their blind adaptive implementations”, *IEEE Transactions on Communications*, v. 50, n. 3, pp. 503–513, 2002.

[41] K. T. Wagner e M. I. Doroslovacki, “Towards analytical convergence analysis of proportionate-type nlms algorithms”, em *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 141, 2008, pp. 3825–3828.

[42] J. Wang e J. Luo, “Performance analysis of an improved variable tap-length LMS algorithm”, em *2009 International Conference on Image Analysis and Signal Processing*, 2009, pp. 377–380.

[43] E. Eweda, “Tracking analysis of the normalized LMS algorithm without the independence and small step size assumptions”, em *2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, vol. 9, 2009, pp. 129–134.

[44] S. C. Douglas, “Mean-square analysis of the multiple-error and block LMS adaptive algorithms”, em *Proceedings of ICASSP ’94. IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, 1994, pp. 429–432.

[45] N. J. Bershad e J. C. M. Bermudez, “Stochastic analysis of the LMS algorithm for non-stationary white Gaussian inputs”, em *2011 IEEE Statistical Signal Processing Workshop (SSP)*, vol. 11, 2011, pp. 57–60.

[46] J. Apolinario, M. L. R. Campos e P. S. R. Diniz, “Convergence analysis of the binormalized data-reusing LMS algorithm”, *IEEE Transactions on Signal Processing*, v. 48, n. 11, pp. 3235–3242, 2000.

[47] P. Lara, K. d. S. Olinto, F. R. Petraglia e D. B. Haddad, “Exact analysis of the least-mean-square algorithm with coloured measurement noise”, *Electronics Letters*, v. 54, n. 24, pp. 1401–1403, 2018.

[48] H. J. Butterweck, “A wave theory of long LMS adaptive filters”, em *2000 10th European Signal Processing Conference*, 2000, pp. 1–4.

[49] S. Miyoshi e Y. Kajikawa, “Theoretical discussion of the filtered-X LMS algorithm based on statistical mechanical analysis”, em *2012 IEEE Statistical Signal Processing Workshop (SSP)*, 2012, pp. 341–344.

[50] O. J. Tobias, “Análise estatistica do comportamento do algoritmo LMS filtrado”, tese de dout., Universidade Federal de Santa Catarina, Curso de Pós-Graduação em Engenharia Elétrica, Florianópolis, 1999.

[51] A. K. Wang e Wei Ren, “Convergence analysis of the multi-variable filtered-X LMS algorithm with application to active noise control”, *IEEE Transactions on Signal Processing*, v. 47, n. 4, pp. 1166–1169, 1999.

[52] Dayong Zhou, V. E. DeBrunner, L. DeBrunner e Yunhua Wang, “Geometric analysis of filtered-X LMS algorithms”, em *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, 2005, pp. 127–132.

[53] F. Yang, J. Guo e J. Yang, “Stochastic Analysis of the Filtered-x LMS Algorithm for Active Noise Control”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 28, n. 1, pp. 2252–2266, 2020.

[54] S. Elliott e P. Nelson, “Active noise control”, *IEEE Signal Processing Magazine*, v. 10, n. 4, pp. 12–35, 1993.

[55] Y. Hinamoto e H. Sakai, “A Filtered-X LMS Algorithm for Sinusoidal Reference Signals—Effects of Frequency Mismatch”, *IEEE Signal Processing Letters*, v. 14, n. 4, pp. 259–262, 2007.

[56] O. J. Tobias e R. Seara, “Mean weight behavior of the FXAFA LMS algorithm”, *IEEE Transactions on Signal Processing*, v. 54, n. 2, pp. 801–804, 2006.

[57] B. Kang, J. Yoo e P. Park, “Bias-compensated normalised LMS algorithm with noisy input”, *Electronics Letters*, v. 49, n. 8, pp. 538–539, 2013.

[58] D. Feng e W. X. Zheng, “Fast Approximate Inverse Power Iteration Algorithm for Adaptive Total Least-Squares FIR Filtering”, *IEEE Transactions on Signal Processing*, v. 54, n. 10, pp. 4032–4039, 2006.

[59] S. M. Jung e P. Park, “Stabilization of a Bias-Compensated Normalized Least-Mean-Square Algorithm for Noisy Inputs”, *IEEE Transactions on Signal Processing*, v. 65, n. 11, pp. 2949–2961, 2017.

[60] R. Abdolee e B. Champagne, “Diffusion LMS Strategies in Sensor Networks With Noisy Input Data”, *IEEE/ACM Transactions on Networking*, v. 24, n. 1, pp. 3–14, 2016.

[61] R. A. d. Prado, F. d. R. Henriques e D. B. Haddad, “Sparsity-Aware Distributed Adaptive Filtering Algorithms for Nonlinear System Identification”, em *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–8.

[62] D. B. Haddad, M. R. Petraglia e A. Petraglia, “A unified approach for sparsity-aware and maximum correntropy adaptive filters”, em *2016 24th European Signal Processing Conference (EUSIPCO)*, 2016, pp. 170–174.

[63] M. R. Petraglia e D. B. Haddad, “New adaptive algorithms for identification of sparse impulse responses — Analysis and comparisons”, em *2010 7th International Symposium on Wireless Communication Systems*, 2010, pp. 384–388.

[64] T. N. Ferreira, M. V. S. Lima, P. S. R. Diniz e W. A. Martins, “Low-complexity proportionate algorithms with sparsity-promoting penalties”, em *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2016, pp. 253–256.

[65] M. V. S. Lima, T. N. Ferreira, W. A. Martins e P. S. R. Diniz, “Sparsity-Aware Data-Selective Adaptive Filters”, *IEEE Transactions on Signal Processing*, v. 62, n. 17, pp. 4557–4572, 2014.

[66] J. Yoo, J. Shin e P. Park, “An Improved NLMS Algorithm in Sparse Systems Against Noisy Input Signals”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, v. 62, n. 3, pp. 271–275, 2015.

[67] W. Ma, D. Zheng, X. Tong, Z. Zhang e B. Chen, “Proportionate NLMS With Unbiasedness Criterion for Sparse System Identification in the Presence of Input and Output Noises”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, v. 65, n. 11, pp. 1808–1812, 2018.

[68] R. M. S. Pimenta, L. C. Resende, L. P. R. da Silva, N. N. Siqueira, J. R. P. Junior, M. R. Petraglia e D. B. Haddad, “A Block-Sparsity Induced NLMS Algorithm with Bias Compensation”, Florianópolis: SBrT 2020, 2020, pp. 1–5.

[69] M. R. Petraglia, E. L. Marques e D. B. Haddad, “Low-complexity affine projection subband algorithm for robust adaptive filtering in impulsive noise”, em *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016, pp. 1–5.

[70] M. R. Petraglia, D. B. Haddad e E. L. Marques, “Affine Projection Subband Adaptive Filter With Low Computational Complexity”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, v. 63, n. 10, pp. 989–993, 2016.

[71] ——, “Normalized Subband Adaptive Filtering Algorithm With Reduced Computational Complexity”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, v. 62, n. 12, pp. 1164–1168, 2015.

[72] P. P. S. Xavier, M. R. Petraglia e D. B. Haddad, “A non-uniform NSAF-SF adaptive algorithm”, *Electronics Letters*, v. 1, n. 1, pp. 1–3, 2021.

[73] Z. Zheng e H. Zhao, “Bias-Compensated Normalized Subband Adaptive Filter Algorithm”, *IEEE Signal Processing Letters*, v. 23, n. 6, pp. 809–813, 2016.

[74] W. Ma, D. Zheng, Y. Li, Z. Zhang e B. Chen, “Bias-compensated normalized maximum correntropy criterion algorithm for system identification with noisy input”, *Signal Processing*, v. 152, pp. 160 –164, 2018.

[75] T. N. Ferreira, W. A. Martins, M. V. S. Lima e P. S. R. Diniz, “Convex Combination of Constraint Vectors for Set-membership Affine Projection Algorithms”, em *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4858–4862.

[76] P. S. R. Diniz, “On Data-Selective Adaptive Filtering”, *IEEE Transactions on Signal Processing*, v. 66, n. 16, pp. 4239–4252, 2018.

[77] P. S. R. Diniz e H. Yazdanpanah, “Data censoring with set-membership algorithms”, em *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2017, pp. 121–125.

[78] Z. Zheng, Z. Liu e L. Lu, “Bias-compensated robust set-membership NLMS algorithm against impulsive noises and noisy inputs”, *Electronics Letters*, v. 53, n. 16, pp. 1100–1102, 2017.

[79] S. Jiang e Y. Gu, “Block-Sparsity-Induced Adaptive Filter for Multi-Clustering System Identification”, *IEEE Transactions on Signal Processing*, v. 63, n. 20, pp. 5318–5330, 2015.

[80] G. Su, J. Jin, Y. Gu e J. Wang, “Performance Analysis of ℓ_0-Norm Constraint Least Mean Square Algorithm”, *IEEE Transactions on Signal Processing*, v. 60, n. 5, pp. 2223–2235, 2012.

[81] Y. Gu, J. Jin e S. Mei, “ℓ_0-Norm Constraint LMS Algorithm for Sparse System Identification”, *IEEE Signal Processing Letters*, v. 16, n. 9, pp. 774–777, 2009.

[82] T. Aboulnasr e K. Mayyas, “Complexity reduction of the NLMS algorithm via selective coefficient update”, *IEEE Transactions on Signal Processing*, v. 47, n. 5, pp. 1421–1424, 1999.

[83] H. Zhao e Z. Zheng, “Bias-compensated affine-projection-like algorithms with noisy input”, *Electronics Letters*, v. 52, n. 9, pp. 712–714, 2016.

[84] H. R. Pota e A. G. Kelkar, “On perfect acoustic noise cancelling controllers”, em *Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No.00CH36334)*, vol. 5, 2000, pp. 3018–3022.

[85] M. A. Sahib e S. Streif, “Design of an active noise controller for reduction of tire/road interaction noise in environmentally friendly vehicles”, em *2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, 2017, pp. 59–62.

[86] Y. Maluenda e J. C. M. Bermudez, “Transient Mean-Square Analysis of Prediction Error Method-Based Adaptive Feedback Cancellation in Hearing Aids”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 20, n. 1, pp. 261–275, 2012.