

Juan Lieber Marin

**UM NOVO MODELO DE COMPORTAMENTO  
ESTOCÁSTICO PARA O ALGORITMO DE PSEUDO  
PROJEÇÕES AFINS**

Dissertação submetida ao Programa  
de Pós-Graduação em Engenharia Elétrica  
para a obtenção do Grau de Mestre  
em Engenharia Elétrica.  
Orientador: Prof. José Carlos Moreira  
Bermudez, Ph.D

Florianópolis

2015

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Marin, Juan Lieber

Um novo modelo de comportamento estocástico para o algoritmo de pseudo projeções afins / Juan Lieber Marin ; orientador, José Carlos Moreira Bermudez - Florianópolis, SC, 2015.

108 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

1. Engenharia Elétrica. 2. Filtragem Adaptativa. 3. Algoritmo de Pseudo Projeções Afins. 4. Análise Estocástica. 5. Identificação de Sistemas. I. Bermudez, José Carlos Moreira. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

Juan Lieber Marin

**UM NOVO MODELO DE COMPORTAMENTO  
ESTOCÁSTICO PARA O ALGORITMO DE PSEUDO  
PROJEÇÕES AFINS**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Engenharia Elétrica”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica.

Florianópolis, 23 de outubro 2015.

---

Prof. Dr. Carlos Galup Montoro  
Coordenador  
Universidade Federal de Santa Catarina

**Banca Examinadora:**

---

Prof. José Carlos Moreira Bermudez, Ph.D  
Orientador

---

Prof. Dr. Márcio Holsbach Costa  
Universidade Federal de Santa Catarina

---

Prof. Dr. Sérgio José Melo de Almeida  
Universidade Católica de Pelotas

---

Prof. Dr. Leonardo Silva Resende  
Universidade Federal de Santa Catarina

---

Prof. Dr. Carlos Aurélio Faria da Rocha  
Universidade Federal de Santa Catarina

A todas as pessoas que de alguma forma  
torceram pelo meu sucesso.



## AGRADECIMENTOS

Ao professor José Carlos Moreira Bermudez pela orientação e empenho durante o desenvolvimento do trabalho.

Ao professor Márcio Holsbach Costa pela ajuda e disposição ao longo desse trabalho e pelas provocações feitas na hora do café.

Aos meus pais, Ieda e Pedro, por permitirem que eu chegasse até aqui.

À minha companheira, Caroline, que, apesar a distância, manteve-se ao meu lado sendo paciente nos momentos difíceis.

Aos professores do Programa de Pós-Graduação em Engenharia Elétrica pelo ambiente e aprendizado.

Aos funcionários da UFSC: Wilson Silva Costa e Marcelo Manoel Siqueira.

À CAPES pela oportunidade, incentivo e financiamento.

Aos colegas de laboratório, sem necessidade de listar razões.





Existem muitas hipóteses em ciência que estão erradas. Isso é perfeitamente aceitável, elas são a abertura para achar as que estão certas.

(Carl Sagan)



## RESUMO

Este trabalho apresenta um novo modelo de comportamento do algoritmo de Pseudo Projeções Afins (PAP) sob as seguintes condições: identificação de sistemas, passo de adaptação arbitrário, correção do problema de inicialização e ordem do algoritmo PAP menor que a ordem da entrada autoregressiva (AR). Para tanto, hipóteses simplificadoras utilizadas na literatura são reavaliadas. Complementarmente, uma generalização para o estimador de mínimos quadrados, utilizado durante o processo de adaptação do algoritmo PAP é proposta, considerando o caso onde o algoritmo opera com ordem menor do que o sinal AR. Como resultado dessas considerações, um novo modelo estocástico é obtido para o algoritmo PAP. As equações aqui obtidas corroboram com as simulações de Monte Carlo do algoritmo PAP.

**Palavras-chave:** Filtragem Adaptativa. Algoritmo de Pseudo Projeções Afins. Análise Estocástica. Identificação de Sistemas.



## ABSTRACT

This work presents a new model for the behaviour of the Pseudo Affine Projection Algorithm (PAP) under the following conditions: system identification, arbitrary step size, correction of initialization effect and projection order of PAP algorithm smaller than the autoregressive input (AR) order. For this purpose, the simplifying assumptions used in the literature are revalued. Complementary, a generalization to the least squares estimator, used during the adaptation process of the PAP algorithm, is proposed considering the case that the PAP algorithm operates with smaller order than the AR signal. As a result of those considerations, a new stochastic model is obtained for the PAP algorithm. The equations derived agree with the Monte Carlo simulation of the PAP algorithm.

**Keywords:** Adaptive Filtering. Pseudo Affine Projection Algorithm.

Stochastic Analysis. System Identification



## LISTA DE FIGURAS

Figura 1	Diagrama de blocos de um filtro adaptativo.....	28
Figura 2	Diagrama de identificação de sistemas .....	28
Figura 3	Diagrama para modelamento inverso.....	29
Figura 4	Diagrama para predição de sinal.....	29
Figura 5	Diagrama para cancelamento de interferência .....	30
Figura 6	Estrutura de um filtro transversal .....	31
Figura 7	Comparação entre os modelos propostos por Almeida et al. (2005) e Barcelos (2013) .....	49
Figura 8	Comparação entre diferentes versões do modelo de Almeida, Bermudez e Bershad (2009) .....	52
Figura 9	Comparação do modelo proposto por Almeida, Bermudez e Bershad (2009) com diferentes passos de adaptação .....	54
Figura 10	Comportamento instantâneo de $e_a(n)$ .....	64
Figura 11	Comparação do algoritmo PAP com a Equação (3.22)..	67
Figura 12	Comparação entre o modelo proposto e o modelo derivado por Almeida, Bermudez e Bershad (2009) da previsão do	

comportamento médio dos coeficientes .....	93
Figura 13 Comparação do EQM em excesso entre o modelo proposto e o modelo derivado por Almeida, Bermudez e Bershada (2009)	93
Figura 14 Comparação do comportamento médio dos coeficientes entre o modelo proposto e o modelo derivado por Almeida, Costa e Bermudez (2010).....	94
Figura 15 Comparação do modelo proposto e o algoritmo PAP ..	95
Figura 16 Comparação do EQM em excesso entre o modelo proposto e o modelo derivado por Almeida, Costa e Bermudez (2010)	96
Figura 17 Comparação do modelo proposto e o algoritmo PAP com vetores IID.....	97
Figura 18 Comparação do modelo proposto e o algoritmo PAP ...	98
Figura 19 Comparação do EMQE do modelo proposto e o algoritmo PAP .....	99
Figura 20 Comparação do modelo proposto e o algoritmo PAP ...	100
Figura 21 Comparação do EQME modelo proposto e o algoritmo PAP .....	100



## LISTA DE ABREVIATURAS E SIGLAS

LMS	<i>Least Mean-Square</i> .....	25
EQM	Erro Quadrático Médio .....	27
FIR	Resposta Finita ao Impulso .....	30
IIR	Resposta Infinita ao Impulso .....	30
NLMS	<i>Normalized Least Mean Square</i> .....	34
AP	<i>Affine Projections</i> .....	34
PAP	<i>Pseudo Affine Projections</i> .....	34
RLS	<i>Recursive Least Squares</i> .....	35



## LISTA DE SÍMBOLOS

$d(n)$	Sinal desejado.....	27
$u(n)$	Conjunto de observações.....	27
$e(n)$	Erro de estimação.....	27
$J(n)$	Erro Quadrático Médio.....	27
$y(n)$	Resposta do sistema desconhecido à entrada $u(n)$ .....	28
$N$	Número de coeficientes do filtro adaptativo.....	31
$\mathbf{u}(n)$	Vetor de entrada do filtro adaptativo.....	31
$\mathbf{p}(n)$	Vetor de correlação cruzada entre o $d(n)$ e $u(n)$ .....	32
$\mathbf{R}(n)$	Matriz de autocorrelação de $\mathbf{u}(n)$ .....	32
$\mathbf{w}_0$	Solução ótima do filtro adaptativo.....	32
$\mu$	Passo de adaptação.....	34
$\mathbf{w}(n)$	Vetor de coeficientes do filtro adaptativo.....	39
$\mathbf{U}_u(n)$	Matriz de regressor atual e dos $P$ passados de $\mathbf{u}(n)$ .....	40
$\mathbf{d}(n)$	Vetor formado pelas $P + 1$ amostras do sinal desejado.....	40

$\alpha$	Passo de adaptação do algoritmo AP .....	41
$\mathbf{e}_p(n)$	Vetor de erro de estimação modificado .....	41
$\hat{\mathbf{a}}(n)$	Estimador de mínimos quadrados de $\mathbf{u}(n)$ .....	42
$\phi(n)$	Erro de predição de $\mathbf{u}(n)$ .....	42
$r(n)$	Ruído aditivo ao sinal desejado .....	44
$z(n)$	Ruído de excitação do processo AR .....	45
$a_i$	i-ésimo coeficiente do processo AR .....	45
$z(n)$	Vetor de ruído de excitação do processo AR .....	45
$\mathbf{R}_\phi$	Matriz de covariância de $\phi(n)$ .....	47
$H$	Ordem do processo autoregressivo .....	47
$\mathbf{v}(n)$	Vetor de erro nos coeficientes .....	54
$r(n)$	Ruído aditivo da saída do sistema desconhecido .....	55
$\sigma_r^2$	Variância do ruído $r(n)$ .....	55
$\mathbf{r}(n-1)$	Vetor com $P$ amostras passadas de $r(n)$ .....	55
$\mathbf{q}(n)$	Vetor de correção do efeito de inicialização .....	56
$u_{-1}(n)$	Função degrau unitário .....	56

$\bar{\mathbf{U}}(n)$	Matriz dos $H - P$ vetores passados do sinal de entrada . . . .	57
$\mathbf{P}_\perp(n)$	Matriz de projeção no espaço complementar às colunas de $\mathbf{U}(n)$ . . . . .	57
$H$	Ordem do Processo AR . . . . .	59
$\mathbf{e}_a(n)$	Vetor dos $P$ erros de estimação passados . . . . .	61
$r_a(n)$	Ruído Filtrado . . . . .	65
$\mathbf{R}_{u_{P \times P}}$	Matriz de autocorrelação de ordem $P$ de $u(n)$ . . . . .	67
$\mathbf{p}_u$	Vetor de autorrelação de $u(n)$ . . . . .	68
$r_u(k)$	Sequência de correlação do sinal de entrada $u(n)$ . . . . .	68
$\mathbf{R}_\phi$	Matriz de correlação de $\phi(n)$ . . . . .	69
$\mathbf{R}_{\phi u}(n)$	Matriz de correlação entre $\phi(n)$ e $\mathbf{U}(n)\mathbf{a}_w$ . . . . .	70
$G$	Diferença entre $N$ e $P$ . . . . .	70
$\sigma_\phi^2$	Variância de $\phi(n)$ . . . . .	70
$\mathbf{K}(n)$	Matriz de correlação do vetor de erro nos coeficientes . . . . .	73
$\mathbf{R}_{\bar{u}}(n)$	Matriz de correlação de $\mathbf{U}(n)\mathbf{a}_w$ . . . . .	76



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	25
1.1	FILTRAGEM ADAPTATIVA SUPERVISIONADA .....	26
1.2	ESTRUTURAS DE FILTRAGEM .....	30
1.3	ALGORITMOS ADAPTATIVOS .....	32
1.4	OBJETIVOS DO TRABALHO .....	37
1.5	ESTRUTURA DO TRABALHO .....	38
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b> .....	39
2.1	LIMITAÇÕES DOS MODELOS EXISTENTES .....	48
2.2	O EFEITO DA INICIALIZAÇÃO DO ALGORITMO PAP	54
2.3	O EFEITO DA ORDEM INSUFICIENTE NO ALGO- RITMO PAP .....	57
2.4	RESULTADOS OBTIDOS NO TRABALHO .....	58
<b>3</b>	<b>ANÁLISE ESTOCÁSTICA DO ALGORITMO DE PSEUDO PROJEÇÕES AFINS</b> .....	59
3.1	CORREÇÃO DO PROBLEMA DE INICIALIZAÇÃO ...	60
3.2	COMPORTAMENTO MÉDIO DOS COEFICIENTES ...	68
3.3	ERRO QUADRÁTICO MÉDIO .....	71
3.4	MATRIZ DE COVARIÂNCIA DO VETOR DE ERRO NOS COEFICIENTES .....	78
<b>4</b>	<b>RESULTADOS</b> .....	91
4.1	AVALIAÇÃO DO MODELO PROPOSTO .....	92
<b>5</b>	<b>CONCLUSÃO</b> .....	103

5.1	SUGESTÃO DE TRABALHOS FUTUROS .....	104
	<b>REFERÊNCIAS .....</b>	<b>107</b>



## 1 INTRODUÇÃO

A filtragem adaptativa é aplicada em uma enorme gama de problemas de engenharia, como cancelamento de eco acústico, adaptação de antenas, equalização de canais, cancelamento de interferência e análise e síntese de voz (MANOLAKIS; INGLE; KOGON, 2000). O estudo na área foi motivado a partir do desenvolvimento do algoritmo *Least Mean-Square* (LMS) feito em (WIDROW; HOFF, 1960). Desde então, novos algoritmos adaptativos foram propostos, e o estudo teórico de seus respectivos desempenhos, visando a compreensão de seus comportamentos e auxílio ao projeto, tornou-se uma intensa linha de pesquisa.

O interesse no desenvolvimento desse ramo de estudo se deve ao fato de que um filtro adaptativo é capaz de operar com sinais e sistemas de natureza aleatória sem depender do conhecimento das estatísticas envolvidas, circunstância essa em que sistemas não adaptativos não obtêm um desempenho satisfatório. Um exemplo que ilustra tal situação é o de um sinal puramente senoidal com frequência variável de forma aleatória que contamina um sinal de interesse. No projeto de um filtro *notch* fixo, para eliminar tal interferência, observa-se um desempenho insatisfatório, uma vez que a frequência central do filtro é fixa. Ao substituir essa solução por um filtro *notch* adaptativo, este é capaz, sob certas condições, de ter sua frequência central alterada de acordo com a variação da frequência do sinal contaminante.

## 1.1 FILTRAGEM ADAPTATIVA SUPERVISIONADA

De acordo com Widrow e Stearns (1985), os sistemas adaptativos possuem todas ou algumas das características abaixo enumeradas:

1. Adaptação automática à medida em que ocorrem modificações do ambiente ou mudanças das necessidades do sistema (auto-otimização);
2. Podem ser treinados para desenvolver uma tarefa específica de filtragem ou decisão, ou seja, podem ser programados através de um processo de treinamento (autoprogramáveis);
3. Em decorrência do item anterior, não precisam de procedimentos elaborados de síntese;
4. Podem extrapolar o espaço de conhecimento e lidar com novas situações após o treinamento com um pequeno conjunto de padrões de entrada (autoaprendizado);
5. Até certo ponto podem reparar a si mesmos, já que são capazes de se adaptar em regiões próximas da ótima mesmo quando sujeitos a certos tipos de defeitos ou limitações;
6. Geralmente são mais complexos e difíceis de analisar que sistemas não adaptativos, mas oferecem a possibilidade de um desempenho substancialmente melhor quando as características do ambiente são desconhecidas ou variantes no tempo.

De uma forma geral, então, filtros adaptativos podem ser definidos como sistemas que variam ao longo do tempo visando satisfazer algum critério de desempenho. Na maioria das aplicações, tais filtros são concebidos para estimar alguma grandeza que está presente em um sinal de interesse  $d(n)$  a partir de observações de outro sinal  $u(n)$ , e o critério de adaptação do algoritmo adaptativo é baseado no erro de estimação  $e(n)$ . O diagrama de bloco que ilustra essa operação é mostrado na Figura 1.

Em geral, a função custo a ser minimizada é o erro quadrático médio (EQM). Matematicamente, o EQM é definido, para sinais reais, como:

$$J(n) = E\{e^2(n)\} = E\{(d(n) - y(n))^2\} \quad (1.1)$$

em que  $E\{\cdot\}$  corresponde ao operador valor esperado. A escolha do EQM como uma função custo a ser minimizada é justificada pela tratabilidade matemática e pela interpretação física que a mesma oferece; ademais, no caso de filtros com resposta ao impulso de duração finita, sua superfície de desempenho apresenta um ponto de mínimo global, caracterizando um hiperparabolóide.

Ainda conforme Haykin (2002), a maioria dos problemas de filtragem adaptativa recai em quatro casos básicos: identificação de sistemas, modelagem inversa, predição e cancelamento de interferência. No caso de identificação de sistemas, o objetivo do filtro adaptativo é estimar a resposta ao impulso de um sistema desconhecido. Para isso, tanto a planta desconhecida como o filtro adaptativo são excitados com o mesmo sinal de entrada. O diagrama de blocos associado a

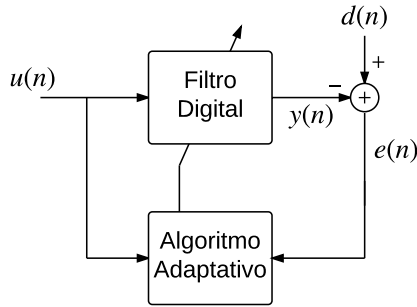


Figura 1 – Diagrama de blocos de um filtro adaptativo

essa aplicação encontra-se esquematizado na Figura 2, em que  $u(n)$  é o sinal de entrada aplicado no sistema desconhecido e ao filtro adaptativo,  $y(n)$  é a resposta do sistema desconhecido à entrada  $u(n)$ ,  $r(n)$  é um ruído aditivo responsável por modelar processos aleatórios que não podem ser obtidos através de uma combinação linear do sinal de entrada,  $d(n)$  é o sinal desejado e  $e(n)$  é o erro de estimação do sinal de saída do sistema desconhecido.

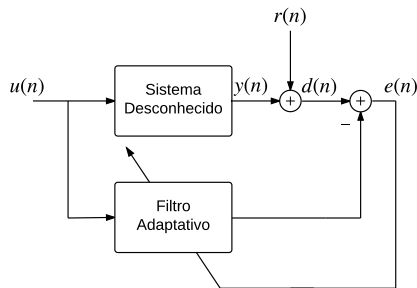


Figura 2 – Diagrama de identificação de sistemas

Na modelagem inversa, o objetivo é obter um sistema que realize a operação inversa à de um sistema desconhecido. Para tanto

o filtro adaptativo é colocado em série com o sistema desconhecido, como mostrado na Figura 3. Adicionalmente, o sinal desejado para este tipo de aplicação é uma versão atrasada do sinal de entrada. Essa abordagem pode ser utilizada em aplicações que visam, por exemplo, a equalização de um canal de comunicações para redução de interferência intersimbólica.

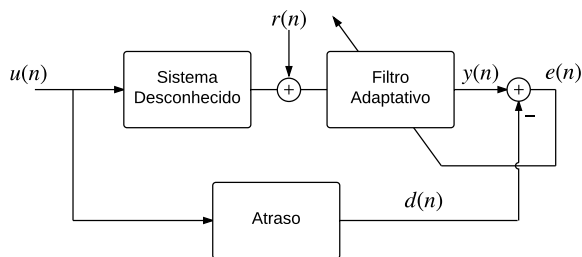


Figura 3 – Diagrama para modelamento inverso

A predição de sinal envolve a estimação de uma amostra de um sinal correlacionado temporalmente, a partir de observações passadas desse sinal. As observações são processadas pelo filtro adaptativo e o erro de predição é avaliado com relação à amostra que se deseja estimar, que constitui o sinal desejado conforme mostra a Figura 4.

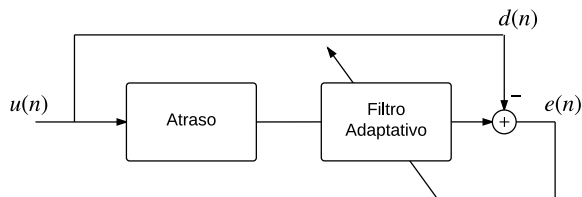


Figura 4 – Diagrama para predição de sinal

No caso de cancelamento de interferência, é assumido que o sinal desejado é composto por uma parcela de interesse e por outra parcela que corrompe o sinal. O objetivo dessa aplicação é reduzir ao máximo a parcela corruptora do sinal desejado. Para isso, um sinal de referência  $u(n)$  é aplicado ao filtro adaptativo e o erro de estimação é avaliado conforme mostrado na Figura 5.

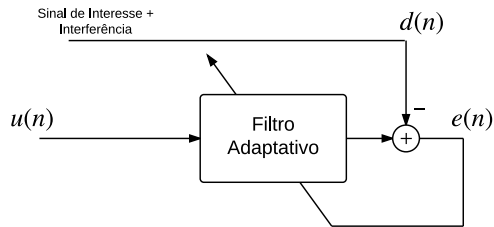


Figura 5 – Diagrama para cancelamento de interferência

## 1.2 ESTRUTURAS DE FILTRAGEM

As estruturas de implementação de filtros adaptativos podem ser classificadas, de maneira geral, em estruturas com resposta ao impulso de duração finita (FIR, do inglês *Finite Impulse Response*) e estruturas com resposta ao impulso de duração infinita (IIR, do inglês *Infinite Impulse Response*). Filtros FIR, por definição, possuem apenas zeros de transmissão e, conseqüentemente, são estáveis e podem possuir fase linear (PROAKIS; MANOLAKIS, 1996). Devido à sua estabilidade intrínseca, essa é a estrutura mais utilizada na implementação de filtros adaptativos em tempo real.

Já os filtros IIR, que possuem uma função de transferência ra-

cional, podem implementar sistemas com respostas ao impulso com espectros suaves com bem menos coeficientes do que as estruturas FIR (PROAKIS; MANOLAKIS, 1996).

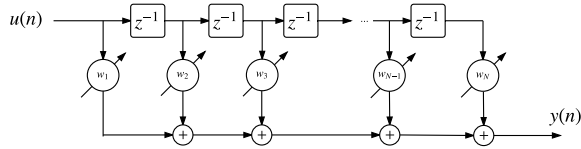


Figura 6 – Estrutura de um filtro transversal

Neste trabalho a estrutura utilizada para os filtros adaptativos será a estrutura transversal, ilustrada na Figura 6. Os blocos identificados pelo operador  $z^{-1}$  são responsáveis por obter amostras atrasadas do sinal de entrada  $u(n)$  e os elementos  $w_k$ ,  $k = 1, \dots, N$  representam os coeficientes do filtro adaptativo. O sinal  $y(n)$  representa a saída do filtro adaptativo.

Definindo o vetor:

$$\mathbf{u}(n) = \begin{bmatrix} u(n) & u(n-1) & \dots & u(n-N+1) \end{bmatrix}^T \quad (1.2)$$

como o vetor de sinal de entrada e ainda  $\mathbf{w}$  como o vetor de coeficientes do filtro digital mostrado na Figura 1, a Equação (1.1) pode ser escrita como:

$$\begin{aligned} J(n) &= E\{e^2(n)\} = E\{(d(n) - \mathbf{u}^T(n)\mathbf{w})^2\} \\ &= E\{d^2(n)\} - 2\mathbf{p}^T(n)\mathbf{w} + \mathbf{w}^T\mathbf{R}(n)\mathbf{w} \end{aligned} \quad (1.3)$$

em que  $\mathbf{p}(n) = E\{d(n)\mathbf{u}(n)\}$  é o vetor de correlação cruzada entre o sinal de entrada e o sinal desejado,  $\mathbf{R}(n) = E\{\mathbf{u}(n)\mathbf{u}^T(n)\}$  é a matriz de autocorrelação do vetor de entrada  $\mathbf{u}(n)$  e  $\mathbf{w}$  denota o vetor  $N \times 1$  de coeficientes do filtro transversal. Consoante a Haykin (2002), se os sinais  $d(n)$  e  $u(n)$  são estacionários,  $\mathbf{R}(n)$  e  $\mathbf{p}(n)$  são invariantes no tempo e a matriz  $\mathbf{R}$  assume uma estrutura *toeplitz* e é simétrica e semi-definida positiva.

Para minimizar o EQM apresentado pela Equação (1.3), pode-se utilizar o método do gradiente descendente - *steepest descent* - (HAYKIN, 2002; WIDROW; STEARNS, 1985). Para isso, é tomada a derivada do EQM em relação aos coeficientes do filtro adaptativo.

$$\nabla J(n) = \frac{\partial J(n)}{\partial \mathbf{w}} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w} \quad (1.4)$$

Igualando a Equação (1.4) a zero, obtém-se:

$$\mathbf{w}_0 = \mathbf{R}^{-1}\mathbf{p} \quad (1.5)$$

em que  $\mathbf{w}_0$  é a solução ótima para o filtro adaptativo FIR. O vetor  $\mathbf{w}_0$  é denominado solução de Wiener para sistemas estacionários.

### 1.3 ALGORITMOS ADAPTATIVOS

Vários algoritmos adaptativos foram desenvolvidos para operar em tempo real desde a concepção do algoritmo LMS, e algumas métricas são descritas por Haykin (2002) para avaliar seus respectivos desempenhos. São elas:



1. Taxa de convergência: define o número de iterações necessárias para que os coeficientes do algoritmo adaptativo se aproximem suficientemente de sua operação em regime permanente;
2. Desajuste: é uma medida quantitativa que avalia a diferença entre o EQM que o algoritmo alcança, após a convergência, e o EQM mínimo produzido pela solução de Wiener;
3. Rastreamento: define, em casos em que o algoritmo adaptativo opera em ambientes não estacionários, a capacidade do mesmo para acompanhar as variações das propriedades estatísticas dos sinais. Tal desempenho é influenciado pela taxa de convergência e pelas variações estatísticas em regime permanente;
4. Robustez: capacidade do algoritmo operar satisfatoriamente com sinais mal condicionados;
5. Complexidade computacional: número de operações matemáticas necessárias para o algoritmo adaptativo realizar uma iteração. Esse fator é determinante para aplicações em tempo real;
6. Estrutura: o filtro adaptativo pode ser implementado em diversas estruturas. A depender da estrutura, por exemplo, pode-se dizer se o algoritmo pode ser paralelizado ou não, indicando qual algoritmo pode ser melhor implementado em um tipo de *hardware* específico;
7. Estabilidade numérica: determina a estabilidade levando em consideração atributos de *hardware* como, por exemplo, níveis de

quantização de um processador dedicado. Um filtro adaptativo é dito numericamente estável se o mesmo é insensível a variação do número de bits usado em implementações digitais.

Na literatura, vários dos algoritmos recursivos propostos para a solução do problema de filtragem adaptativa são provenientes de dois métodos distintos: o método do gradiente descendente e o método dos Mínimos Quadrados (HAYKIN, 2002; WIDROW; STEARNS, 1985). Os algoritmos derivados a partir do método do gradiente descendente visam atualizar o vetor de coeficientes do filtro a cada iteração na direção de maior redução da função custo. Para tanto, a direção de atualização de tais algoritmos é tomada como o negativo do vetor gradiente da função custo, sendo este tomado com relação ao vetor de pesos do filtro adaptativo. Entre os algoritmos dessa família pode-se citar o LMS, o *Normalized Least Mean Squares* (NLMS), o *Affine Projections* (AP) e o *Pseudo Affine Projections* (PAP). Esses algoritmos são normalmente concebidos para minimizar o EQM que, para sistemas estacionários, possui como ponto de mínimo global a solução de Wiener dada pela Equação (1.5). As equações de atualização desses algoritmos são da forma:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha f(e(n)). \quad (1.6)$$

sendo  $\alpha$  o passo de adaptação que controla a velocidade de convergência e o erro em regime permanente e  $f(\cdot)$  uma função linear do erro de estimação.

Diversamente, algoritmos derivados da Teoria dos Mínimos Quadrados apresentam uma perspectiva determinística baseada na soma dos quadrados do erro de estimação. Pelo fato de não se basearem em aproximações estocásticas, algoritmos dessa família tendem a convergir mais rapidamente do que algoritmos provenientes do método do gradiente descendente para sinais estacionários e sistemas invariantes no tempo. Entretanto, possuem uma maior complexidade computacional e podem tornar-se numericamente instáveis. O algoritmo mais conhecido dessa família é o *Recursive Least Squares* (RLS) .

O algoritmo AP, por exemplo, possui um desempenho superior ao algoritmo LMS e complexidade computacional menor do que o algoritmo RLS (SAYED, 2003), tornando-o um algoritmo de grande interesse nas aplicações de engenharia. Por apresentar uma versão do erro de estimação vetorial sua análise é difícil de ser feita, a exceção do caso em que o algoritmo AP opere com passo unitário quando o vetor de erro de estimação torna-se um escalar (ALMEIDA et al., 2005). Sua versão simplificada, o algoritmo PAP, possui uma maior simplicidade de análise uma vez que, mesmo com passo de adaptação arbitrário, opera com o erro de estimação escalar (ALMEIDA; BERMUDEZ; BERSHAD, 2009). Por esse motivo, o algoritmo PAP é mais utilizado na prática do que o algoritmo AP, tornando necessário o estudo do comportamento estocástico do algoritmo PAP.

Alguns trabalhos na literatura propuseram modelos de comportamento estocástico para o algoritmo PAP. Os resultados desses trabalhos são resumidos a seguir:

- Em (ALMEIDA; BERMUDEZ; BERSHAD, 2009) obteve-se um modelo de comportamento estocástico sob o cenário de identificação de sistemas em que o sistema desconhecido é não estacionário. Nesse trabalho, o sinal de entrada é modelado como um processo AR de ordem igual ou inferior a ordem de projeção do algoritmo AP.
- Em (ALMEIDA; COSTA; BERMUDEZ, 2009) foi proposto um modelo de comportamento estocástico considerando planta estacionária com número de coeficientes maior do que o número de coeficientes do filtro adaptativo e ordem de projeção maior ou igual a ordem do processo AR de entrada.
- Em (ALMEIDA; COSTA; BERMUDEZ, 2010) foi elaborado um modelo de comportamento estocástico para o algoritmo PAP considerando o número de coeficientes do filtro adaptativo igual ao número de coeficientes do sistema a ser identificado, sendo este último estacionário. Como inovação com relação aos trabalhos anteriores, foi utilizado como sinal de entrada um processo autorregressivo de ordem superior a ordem de projeção do algoritmo PAP.
- Em (COSTA *et al.*, 2012; BARCELOS, 2013) foi estudado o efeito da inicialização do algoritmo PAP, utilizando passo unitário, ordem suficiente e planta estacionária.

Nota-se pelo conteúdo dos trabalhos existentes na literatura a necessidade da análise do efeito conjunto dos cenários expostos anterior-

mente, uma vez que cada trabalho analisou os efeitos de planta não estacionária, comprimento deficiente, ordem de projeção insuficiente e efeito de inicialização de forma isolada. Outro tipo de análise não abordada pelos trabalhos associados a modelagem do comportamento estocástico do algoritmo PAP é de assumir outros tipos de modelos de sinal de entrada a fim de tornar o modelo mais generalizado com relação ao sinal de entrada, uma vez que todos os trabalhos citados assumem um modelo autoregressivo para o sinal de entrada.

#### 1.4 OBJETIVOS DO TRABALHO

Diante do exposto, este trabalho concentra-se em:

- Estudar o problema de inicialização associado ao algoritmo PAP de forma teórica, apresentado por Costa et al. (2012);
- Avaliar algumas aproximações apresentadas por Almeida, Costa e Bermudez (2010), de forma a se obter uma forma mais simplificada para efeitos associados à ordem deficiente;
- Obter um modelo teórico para o comportamento do algoritmo PAP, utilizando uma metodologia que permite avaliar o regime transitório e permanente do EQM desse algoritmo, considerando os efeitos de inicialização, ordem deficiente e passo de adaptação não unitário.

## 1.5 ESTRUTURA DO TRABALHO

No Capítulo 2 apresenta-se o algoritmo AP e o algoritmo PAP, bem como uma revisão bibliográfica detalhada dos trabalhos associados a esses algoritmos. Nesse capítulo é feita também a dedução do algoritmo AP a partir de um sistema subdeterminado de equações. Nas seções seguintes é explorado o problema de inicialização do algoritmo PAP.

No Capítulo 3 é feita a análise de comportamento estocástico no cenário de identificação de sistemas, considerando passo de adaptação arbitrário, ordem de projeção deficiente e o efeito da inicialização. O resultado desse capítulo consiste em obter um conjunto de equações que descrevem o comportamento estocástico do algoritmo PAP.

O Capítulo 4 apresenta resultados de simulações computacionais que visam validar o conjunto de equações obtido no Capítulo 3. As simulações consistem em avaliar o comportamento médio de alguns coeficientes do filtro adaptativo e o EQM, além de comparar a acurácia da previsão com os modelos de comportamento estocásticos presentes na literatura para o algoritmo PAP.

Por fim, o Capítulo 5 apresenta as conclusões deste trabalho e as sugestões para trabalhos futuros.

Como notação, serão utilizadas letras minúsculas como variáveis escalares, como  $r(n)$  e  $e(n)$ , letras minúsculas em negrito, por exemplo  $\boldsymbol{\phi}(n)$  e  $\boldsymbol{w}(n)$ , como variáveis vetoriais e letras maiúsculas em negrito, a exemplo de  $\mathbf{R}(n)$ , como grandezas matriciais.

## 2 REVISÃO BIBLIOGRÁFICA

O algoritmo de projeções afins foi proposto por OZEKI e UMEDA (1984) para contornar o problema de baixa velocidade de convergência dos algoritmos LMS e NLMS para sinais correlacionados. A proposta do algoritmo teve como base um problema de otimização com restrições. Conforme a Figura 1, o erro de estimação pode ser descrito como:

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{u}(n) \quad (2.1)$$

em que  $\mathbf{w}(n)$  representa o vetor de coeficientes do filtro adaptativo. Deseja-se, então, determinar  $\mathbf{w}(n+1)$  que anule o erro de estimação. Dessa forma:

$$d(n) = \mathbf{w}^T(n+1)\mathbf{u}(n) \quad (2.2)$$

Assim, a Equação (2.2) forma um sistema de equações com  $N$  incógnitas e apenas uma Equação, caracterizando um sistema subdeterminado. Portanto, para determinar  $\mathbf{w}(n+1)$  é necessário adicionar um conjunto de restrições ao problema. A solução proposta por OZEKI e UMEDA (1984) consiste em utilizar  $P+1$  vetores de sinal de entrada, com  $N > P+1$ , com o objetivo de minimizar:

$$\|\mathbf{w}(n+1) - \mathbf{w}(n)\|^2 \quad (2.3)$$

sujeito às seguintes restrições:

$$\begin{aligned}
 \mathbf{u}^T(n)\mathbf{w}(n+1) &= d(n) \\
 \mathbf{u}^T(n-1)\mathbf{w}(n+1) &= d(n-1) \\
 &\vdots \\
 \mathbf{u}^T(n-P)\mathbf{w}(n+1) &= d(n-P)
 \end{aligned} \tag{2.4}$$

Definindo

$$\mathbf{U}_u(n) = \begin{bmatrix} \mathbf{u}(n) & \mathbf{u}(n-1) & \cdots & \mathbf{u}(n-P) \end{bmatrix} \tag{2.5}$$

como a matriz formada pelos  $P+1$  vetores de entrada e:

$$\mathbf{d}(n) = \begin{bmatrix} d(n) & d(n-1) & \cdots & d(n-P) \end{bmatrix}^T \tag{2.6}$$

como o vetor formado pelas  $P+1$  amostras do sinal desejado, a Equação (2.4) pode ser escrita como

$$\mathbf{d}(n) = \mathbf{U}_u^T(n)\mathbf{w}(n+1) \tag{2.7}$$

Desse modo, o vetor que contém  $P+1$  amostras do erro é dado por:

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{U}_u^T(n)\mathbf{w}(n) = \begin{bmatrix} e(n) & \cdots & e(n-P) \end{bmatrix}^T \tag{2.8}$$

Assim, o problema de minimização da norma  $l_2$  do vetor de erros pode ser resolvido utilizando o método dos multiplicadores de Lagrange, sob o conjunto de restrições apresentados pela Equação (2.7),



resultando em (OZEKI; UMEDA, 1984):

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{U}_u(n) [\mathbf{U}_u^T(n) \mathbf{U}_u(n)]^{-1} \mathbf{e}(n) \quad (2.9)$$

Com a finalidade de controlar a velocidade de convergência e o erro em regime permanente é possível incluir um passo de adaptação  $\alpha$ , resultando em:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha \mathbf{U}_u(n) [\mathbf{U}_u^T(n) \mathbf{U}_u(n)]^{-1} \mathbf{e}(n) \quad (2.10)$$

Nota-se, pela Equação (2.8), que a complexidade computacional do algoritmo AP é relativamente alta, dado que é necessário efetuar uma filtragem dos  $P+1$  sinais de entrada passados com os coeficientes  $\mathbf{w}(n+1)$  para compor um vetor de  $P+1$  erros de estimação modificado. Assim, definindo  $\mathbf{e}_p(n)$  como o vetor de erro modificado (GAY; TAVATHIA, 1995):

$$\mathbf{e}_p(n) = \mathbf{d}(n) - \mathbf{U}_u^T(n) \mathbf{w}(n+1) \quad (2.11)$$

é possível demonstrar que:

$$\mathbf{e}_p(n) = (1 - \alpha) \mathbf{e}(n) \quad (2.12)$$

Consequentemente, para passo unitário, o vetor  $\mathbf{e}_p(n)$  torna-se

nulo e o vetor de erro  $\mathbf{e}(n)$  é igual a:

$$\mathbf{e}(n) = \begin{bmatrix} d(n) - \mathbf{u}^T(n)\mathbf{w}(n) \\ d(n-1) - \mathbf{u}^T(n-1)\mathbf{w}(n) \\ \vdots \\ d(n-P) - \mathbf{u}^T(n-P)\mathbf{w}(n) \end{bmatrix} = \begin{bmatrix} e(n) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.13)$$

Assim, ao utilizar passo de adaptação unitário, a complexidade computacional do algoritmo AP é reduzida consideravelmente, uma vez que o erro de estimação torna-se um escalar, evitando a necessidade de filtrar os  $P$  vetores passados de sinal de entrada. Conseqüentemente, a Equação de atualização do algoritmo AP se torna:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{U}_u(n) [\mathbf{U}_u^T(n)\mathbf{U}_u(n)]^{-1} \begin{bmatrix} e(n) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.14)$$

É possível obter uma forma simplificada do algoritmo AP, com passo unitário, definindo um vetor  $\boldsymbol{\phi}(n)$  da forma:

$$\boldsymbol{\phi}(n) = \mathbf{u}(n) - \mathbf{U}(n)\hat{\mathbf{a}}(n) \quad (2.15)$$

em que:

$$\mathbf{U}(n) = \begin{bmatrix} \mathbf{u}(n-1) & \mathbf{u}(n-2) & \cdots & \mathbf{u}(n-P) \end{bmatrix} \quad (2.16)$$

e  $\hat{\mathbf{a}}(n)$  é o estimador ótimo de mínimos quadrados do preditor linear de

$u(n)$  dado por:

$$\hat{\mathbf{a}}(n) = [\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)\mathbf{u}(n) \quad (2.17)$$

Se o vetor de sinal de entrada é um processo autoregressivo de ordem  $P$ , pode-se assumir que  $\boldsymbol{\phi}(n)$  é ortogonal ao subespaço formado pelas colunas de  $\mathbf{U}(n)$  e é possível determinar a seguinte relação:

$$\mathbf{u}^T(n)\boldsymbol{\phi}(n) = \boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n) \quad (2.18)$$

Desse modo:

$$\mathbf{u}^T(n)[\mathbf{u}(n) - \mathbf{U}(n)\hat{\mathbf{a}}(n)] = \boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n) \quad (2.19)$$

$$\mathbf{U}^T(n)[\mathbf{u}(n) - \mathbf{U}(n)\hat{\mathbf{a}}(n)] = \mathbf{0}_{P \times 1} \quad (2.20)$$

Escrevendo na forma matricial:

$$\begin{bmatrix} \mathbf{u}^T(n) \\ \mathbf{U}^T(n) \end{bmatrix} \begin{bmatrix} \mathbf{u}(n) & \mathbf{U}(n) \end{bmatrix} \begin{bmatrix} 1 \\ -\hat{\mathbf{a}}(n) \end{bmatrix} = \boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n) \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.21)$$

Pré-multiplicando ambos os lados por  $\mathbf{U}_u(n)[\mathbf{U}_u^T(n)\mathbf{U}_u(n)]^{-1}$  chega-se a:

$$\boldsymbol{\phi}(n) = \mathbf{U}_u(n)[\mathbf{U}_u^T(n)\mathbf{U}_u(n)]^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n) \quad (2.22)$$

Multiplicando-se por  $e(n)$  em ambos os lados:

$$\mathbf{U}_u(n)[\mathbf{U}_u^T(n)\mathbf{U}_u(n)]^{-1} \begin{bmatrix} e(n) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} e(n) \quad (2.23)$$

Logo, é possível reescrever a Equação (2.10) da seguinte forma:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} e(n) \quad (2.24)$$

A versão do algoritmo AP com passo unitário apresentada pelas equações (2.15), (2.17) e (2.24) é apresentada em (RUPP, 1998). Naquele trabalho, o autor demonstra ainda que, ao utilizar passo unitário, o algoritmo AP com entrada AR de ordem  $P$  efetua um branqueamento do sinal de entrada, justificando o acréscimo na velocidade de convergência quando comparado aos algoritmos LMS e NLMS. Foi mostrado também que se o sinal desejado  $d(n)$  é acrescido de um ruído branco  $r(n)$ , o algoritmo AP enxerga uma versão filtrada de tal ruído com coeficientes  $\hat{\mathbf{a}}(n)$ . Como consequência, a versão filtrada do ruído observado pelo algoritmo AP possui potência  $(1 + \mathbf{a}^T \mathbf{a})\sigma_r^2$ , em que  $\mathbf{a}$  é o vetor de coeficientes do processo AR( $P$ ) de entrada e  $\sigma_r^2$  é a potência do ruído  $r(n)$ .

De forma independente, embora utilizando como justificativa a conclusão apresentada em (RUPP, 1998) sobre o comportamento do algoritmo AP com relação a  $r(n)$ , em (BOUTEILLE; SCARLAT; CORAZZA, 1999) foi proposto o algoritmo de Pseudo Projeções Afins de forma

a diminuir o erro residual causado pela filtragem do ruído  $r(n)$ . Na proposta do algoritmo AP com passo unitário o sinal de entrada  $u(n)$  é substituído por seu modelo AR de ordem  $P$  obtido pelo método dos mínimos quadrados. Isso leva à equação de adaptação (2.24) para qualquer sinal de entrada. Esse algoritmo corresponde ao algoritmo NLMS com vetor de entrada substituído pela sua projeção no espaço ortogonal complementar das colunas da matriz  $\mathbf{U}(n)$ . Reintroduzindo o passo de adaptação  $\mu$  leva a seguinte Equação de atualização:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} e(n) \quad (2.25)$$

Uma vez estabelecidas as propriedades do algoritmo AP, em (BERSHAD; LINEBARGER; MCLAUGHLIN, 2001) foi proposto um modelo de comportamento estocástico para o algoritmo AP assumindo que o sinal  $u(n)$  é um sinal autoregressivo com distribuição gaussiana com mesma ordem de projeções do algoritmo AP e utilizando passo de adaptação unitário. Assim, na forma escalar, o sinal  $u(n)$  é da forma:

$$u(n) = \sum_{i=1}^P a_i u(n-i) + z(n) \quad (2.26)$$

Considerando ainda, que o vetor de sinal de entrada é da forma da Equação (1.2), a Equação (2.26) pode ser escrita na forma:

$$\mathbf{u}(n) = \mathbf{U}(n)\mathbf{a} + \mathbf{z}(n) \quad (2.27)$$

em que  $\mathbf{z}(n)$  é um vetor de variáveis aleatórias gaussianas independentes

de média nula e variância  $\sigma_z^2$ . Assim, de (2.15) e (2.17):

$$\boldsymbol{\phi}(n) = \left[ \mathbf{I} - \mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n) \right] \mathbf{u}(n) \quad (2.28)$$

Utilizando o modelo de sinal de entrada dado pela Equação (2.27):

$$\boldsymbol{\phi}(n) = \left[ \mathbf{I} - \mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n) \right] \left[ \mathbf{U}(n)\mathbf{a} + \mathbf{z}(n) \right] \quad (2.29)$$

Como  $\left[ \mathbf{I} - \mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n) \right]$  é a matriz de projeção no espaço complementar às colunas de  $\mathbf{U}(n)$ , a projeção de  $\mathbf{U}(n)\mathbf{a}$  nesse espaço é nula. Assim:

$$\boldsymbol{\phi}(n) = \left[ \mathbf{I} - \mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n) \right] \mathbf{z}(n) \quad (2.30)$$

e ainda:

$$E\{\hat{\mathbf{a}}(n)\} = \mathbf{a} \quad (2.31)$$

Como consequência da Equação (2.30), foi demonstrado que a matriz de covariância do vetor  $\boldsymbol{\phi}(n)$  possui  $N - P$  autovalores não nulos e  $P$  autovalores nulos. Dessa forma, os resultados do trabalho feito por Bershada, Linebarger e McLaughlin (2001) permitiram um melhor entendimento das propriedades estatísticas do vetor  $\boldsymbol{\phi}(n)$ . Assim, o modelo proposto apresentou uma excelente acurácia na predição do comportamento médio do algoritmo AP.

Posteriormente, em (SHIN; SAYED, 2004), foi elaborado um modelo para o algoritmo AP com passo de adaptação arbitrário e qualquer modelo de sinal de entrada, baseado em considerações de conservação de energia. Entretanto, o modelo obtido depende da estimação numérica

de estatísticas de ordem elevada do sinal de entrada. A obtenção de modelos analíticos completos para casos especiais a partir da derivação feita por Shin e Sayed (2004) ainda é um trabalho em aberto.

Baseado em (BERSHAD; LINEBARGER; MCLAUGHLIN, 2001), o trabalho feito em (ALMEIDA et al., 2005) propôs um modelo de comportamento estocástico do algoritmo AP considerando passo unitário e entradas autoregressivas, podendo essas terem quaisquer distribuições probabilísticas. O modelo foi obtido em forma fechada, sem a necessidade de nenhuma estimação numérica das estatísticas do sinal de entrada. Além do modelo de comportamento médio em forma fechada, esse trabalho determinou também que a matriz de covariância de  $\boldsymbol{\phi}(n)$ , para um sinal de entrada autoregressivo da forma de (2.27) e com ordem igual à do algoritmo AP, seja definida por:

$$\mathbf{R}_\phi = E\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\} = \frac{N-P}{N}\sigma_z^2\mathbf{I} \quad (2.32)$$

Tal trabalho foi expandido posteriormente para o algoritmo PAP, conforme mostra o trabalho publicado em (ALMEIDA; BERMUDEZ; BERSHAD, 2009). Nesse, pode ser verificado que o modelo obtido recai exatamente sobre o trabalho feito por Almeida et al. (2005), quando aplica-se o passo unitário e planta estacionária.

Sob a justificativa de que a implementação do algoritmo AP com ordem elevada acarreta em um elevado custo computacional, em (ALMEIDA; COSTA; BERMUDEZ, 2010) foi proposto um modelo de comportamento estocástico do algoritmo PAP assumindo que a ordem do processo autoregressivo  $H$  é superior à ordem  $P$  do algoritmo. Como

consequência desse cenário, o vetor  $\phi(n)$  passa a possuir correlação temporal, fazendo com que o algoritmo PAP tenha uma menor velocidade de convergência quando comparada à velocidade do algoritmo operando em cenário de ordem suficiente.

Baseado no trabalho de Almeida et al. (2005), os trabalhos feitos em (COSTA et al., 2012; BARCELOS, 2013) investigaram o problema associado à inicialização do algoritmo AP com passo unitário. A derivação do modelo considera que o vetor  $\mathbf{w}(0)$  deve satisfazer um determinado conjunto de equações entre os instantes  $n = 0$  e  $n = P$ . Nesse trabalho foi verificado experimentalmente ainda que tal situação é mais evidente quando duas condições são atendidas simultaneamente: o sinal  $u(n)$  possui alta correlação temporal e o vetor de inicialização  $\mathbf{w}(0)$  está muito distante de  $\mathbf{w}_o$ . Uma explicação teórica para esse comportamento não foi proposta. A Figura 7 mostra a comparação entre o modelo proposto por Almeida et al. (2005) e Barcelos (2013). Para tanto, foram utilizados os seguintes parâmetros: sinal de entrada modelado como  $u(n) = 1.96u(n-1) - 0.9893u(n-2) + z(n)$  com  $\sigma_z^2 = 0.0575$ ,  $P = 2$ , planta desconhecida consistindo em uma janela de Hann normalizada com  $N = 64$  coeficientes e  $\sigma_r^2 = 10^{-6}$ .

## 2.1 LIMITAÇÕES DOS MODELOS EXISTENTES

A análise detalhada dos modelos existentes para o algoritmo PAP e a utilização extensiva desses modelos evidenciaram a necessidade de melhorias no trabalho de análise. Esta seção descreve as limitações encontradas e quais delas foram resolvidas neste trabalho.



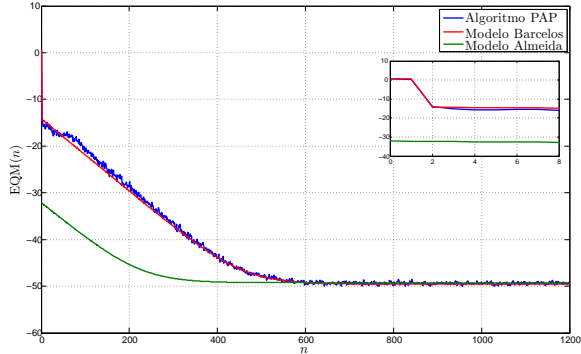


Figura 7 – Comparação entre os modelos propostos por Almeida et al. (2005) e Barcelos (2013)

Inicialmente, o modelo desenvolvido por Costa et al. (2012) e Barcelos (2013) considera os efeitos de inicialização apenas para o caso de passo unitário e para ordem suficiente (sinal de entrada autoregressivo com ordem igual à do algoritmo PAP). Embora o modelo desenvolvido com essas limitações permita inferências importantes a respeito do comportamento do algoritmo, a modelagem do comportamento desse algoritmo para passos não unitários é altamente desejável. Outro aspecto importante é que utilizações práticas do algoritmo PAP estão normalmente limitadas a valores reduzidos de  $P$  (tipicamente  $P \leq 3$ ) devido à complexidade de implementação. Por outro lado, sinais reais normalmente requerem modelagens autoregressivas de ordens bem mais elevadas do que isso. Assim, há grande interesse em modelar os efeitos da inicialização também para implementações de ordem insuficiente e com passo não unitário.

O trabalho desenvolvido por Almeida, Bermudez e Bershada (2009)

levou a um modelo analítico para o algoritmo PAP sem a restrição de passo unitário. No entanto, ainda considerou ordens iguais para o modelo AR do sinal de entrada e para o algoritmo. Além disso, o modelo em questão empresta a expressão aproximada do erro médio quadrático obtida pelos mesmos autores em Almeida et al. (2005) na análise do comportamento do algoritmo AP com passo unitário. Essa simplificação levou à eliminação de uma constante multiplicativa do erro quadrático médio em excesso, a qual pode ter um papel importante na acurácia do modelo dependendo dos coeficientes do processo AR do sinal de entrada. Para demonstrar essa simplificação, recorre-se à Equação (19) do referido trabalho:

$$e_a(n) = \frac{1}{\gamma} [\boldsymbol{\phi}^T(n) \mathbf{v}(n) - \mu \mathbf{ar}(n-1)] \quad (2.33)$$

em que:

$$\gamma = 1 - (1 - \mu) \sum_{i=1}^P a_i \quad (2.34)$$

O erro de estimação no trabalho feito por Almeida, Bermudez e Bershad (2009) é dado por:

$$e(n) = e_a(n) + r(n) \quad (2.35)$$

Elevando a Equação (2.35) ao quadrado, tem-se:

$$e^2(n) = e_a^2(n) + r^2(n) + 2e_a(n)r(n) \quad (2.36)$$

Elevando a Equação (2.33) ao quadrado e substituindo na Equação

(2.36), tem-se:

$$\begin{aligned}
e^2(n) &= r^2(n) + \frac{\mu^2}{\gamma^2} [\hat{\mathbf{a}}^T(n)\mathbf{r}(n-1)]^2 - 2\frac{\mu}{\gamma} \hat{\mathbf{a}}^T(n)\mathbf{r}(n-1)r(n) \\
&\quad + \frac{2}{\gamma} \boldsymbol{\phi}^T(n)\mathbf{v}(n)r(n) - 2\frac{\mu}{\gamma^2} \boldsymbol{\phi}^T(n)\mathbf{v}(n)\hat{\mathbf{a}}^T(n)\mathbf{r}(n-1) \\
&\quad + \frac{1}{\gamma^2} \boldsymbol{\phi}^T(n)\mathbf{v}(n)\boldsymbol{\phi}^T(n)\mathbf{v}(n)
\end{aligned} \tag{2.37}$$

Recorrendo à Equação (22) do trabalho feito por Almeida, Bermudez e Bershad (2009) que define a versão do ruído filtrada, tem-se:

$$r_a(n) = r(n) - \frac{\mu}{\gamma} \hat{\mathbf{a}}^T(n)\mathbf{r}(n-1) \tag{2.38}$$

Elevando a Equação (2.38) ao quadrado, obtém-se:

$$r_a^2(n) = r^2(n) + \frac{\mu^2}{\gamma^2} [\hat{\mathbf{a}}^T(n)\mathbf{r}(n-1)]^2 - 2\frac{\mu}{\gamma} \hat{\mathbf{a}}^T(n)\mathbf{r}(n-1)r(n) \tag{2.39}$$

Substituindo a Equação (2.39) na Equação (2.36), tem-se:

$$\begin{aligned}
e^2(n) &= r_a^2(n) + \frac{2}{\gamma} \boldsymbol{\phi}^T(n)\mathbf{v}(n)r(n) - 2\frac{\mu}{\gamma^2} \boldsymbol{\phi}^T(n)\mathbf{v}(n)\hat{\mathbf{a}}^T(n)\mathbf{r}(n-1) \\
&\quad + \frac{1}{\gamma^2} \boldsymbol{\phi}^T(n)\mathbf{v}(n)\boldsymbol{\phi}^T(n)\mathbf{v}(n)
\end{aligned} \tag{2.40}$$

Tomando o valor esperado da Equação (2.40) e considerando a hipótese de que o ruído aditivo  $r(n)$  é de média nula e independente de qualquer outro sinal, obtém-se:

$$E\{e^2(n)\} = E\{r_a^2(n)\} + \frac{1}{\gamma^2} \text{Tr}[\mathbf{R}_\phi \mathbf{K}(n)] \tag{2.41}$$

em que  $\text{Tr}[\cdot]$  é o operador de traço,  $\mathbf{K}(n)$  é a matriz de covariância do vetor de erro nos coeficientes  $\mathbf{v}(n)$  e  $\mathbf{R}_\phi$  é a matriz de covariância do vetor  $\boldsymbol{\phi}(n)$ . A Equação (2.41) está dissonante da Equação (28)

apresentada no trabalho feito por Almeida, Bermudez e Bershada (2009), que consiste em assumir  $\gamma = 1$ . Para ilustrar o efeito causado por essa simplificação, foi feita uma simulação computacional com os seguintes parâmetros: sinal de entrada modelado como  $u(n) = -0.9u(n-1) + z(n)$ , com  $\sigma_z^2 = 0.9687$ ,  $P = 2$ , sistema desconhecido consistindo em uma janela de Hann com 64 coeficientes, passo de adaptação  $\mu = 0.2$  e  $\sigma_r^2 = 10^{-6}$ . O resultado dessa simulação é mostrado na Figura 8.

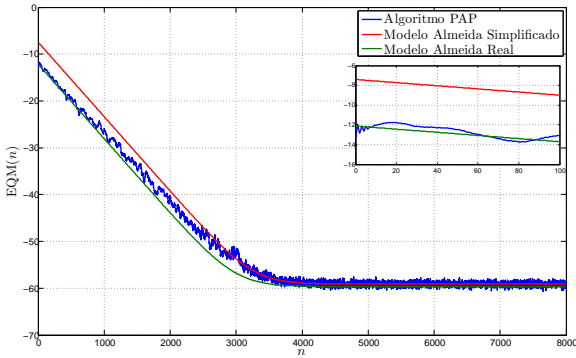


Figura 8 – Comparação entre diferentes versões do modelo de Almeida, Bermudez e Bershada (2009)

Ainda em relação ao modelo desenvolvido em Almeida, Bermudez e Bershada (2009), naquela análise é feita uma hipótese simplificadora de que o erro a priori varia lentamente na janela de observação do algoritmo (intervalo de  $P$  amostras), podendo ser considerado aproximadamente constante nesse intervalo. Da Equação (17) obtida por Almeida, Bermudez e Bershada (2009), considerando a planta estacionária, pode-se obter:

$$e_a(n) = \boldsymbol{\phi}^T(n)\mathbf{v}(n) + (1 - \mu)\hat{\mathbf{a}}^T(n)\mathbf{e}_a(n-1) + \mu\mathbf{r}(n-1) \quad (2.42)$$

Ao considerar a hipótese feita naquele trabalho, que consiste em dizer que as componentes do vetor  $\mathbf{e}_a(n-1)$  são aproximadamente constantes, obtém-se:

$$e_a(n) = \frac{1}{1 - (1 - \mu) \sum_{i=1}^P a_i} [\boldsymbol{\phi}^T(n) \mathbf{v}(n) - \mu \hat{\mathbf{a}}^T(n) \mathbf{r}(n-1)] \quad (2.43)$$

A passagem da Equação (2.42) para a Equação (2.43) elimina o efeito da filtragem do vetor  $\mathbf{e}_a(n-1)$  podendo, para algum exemplo, ser prejudicial na acurácia do modelo proposto por Almeida, Bermudez e Bershad (2009) durante o período transitório. Nota-se ainda que, na Equação (2.42) que o efeito da hipótese sobre o vetor  $\mathbf{e}_a(n-1)$  deve ser mais perceptível para passo de adaptação pequeno e irrelevante para passo de adaptação unitário. Para ilustrar o exposto, a Figura 9 mostra a comparação do modelo proposto por Almeida, Bermudez e Bershad (2009) em duas situações com o mesmo sinal de entrada:  $u(n) = -0.3u(n-1) - 0.5u(n-2)$  com  $\sigma_z^2 = 1.0323$ . Em ambas as comparações, foi utilizada a mesma resposta ao impulso utilizada no trabalho de Almeida, Bermudez e Bershad (2009) com  $N = 64$  coeficientes e  $P = 2$ . A diferença entre as simulações consiste na utilização de diferentes passos de adaptação: a primeira utiliza passo de adaptação  $\mu = 0.2$  e na segunda é utilizado passo de adaptação unitário.

Nota-se na Figura 9 que, ao utilizar passo de adaptação pequeno, a acurácia do modelo proposto por Almeida, Bermudez e Bershad (2009) é comprometida no regime transitório. Simulações em que os valores esperados foram estimados numericamente indicam que a hipótese

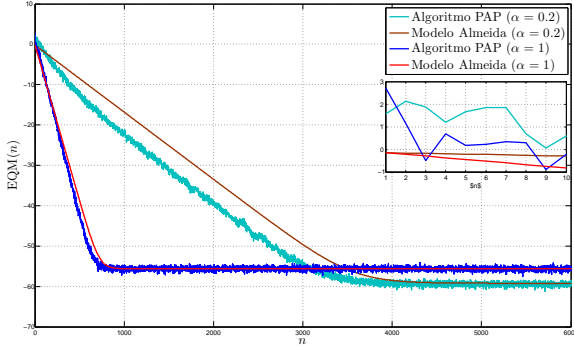


Figura 9 – Comparação do modelo proposto por Almeida, Bermudez e Bershad (2009) com diferentes passos de adaptação

feita sobre o vetor  $\mathbf{e}_a(n-1)$  tem uma grande influência na geração do descasamento observado entre a teoria e a simulação. Por outro lado, ao utilizar passo de adaptação unitário, nota-se uma boa acurácia no regime transitório do modelo proposto por Almeida, Bermudez e Bershad (2009) com relação ao comportamento do EQM do algoritmo PAP, situação essa que anula a filtragem do vetor  $\mathbf{e}_a(n-1)$ .

## 2.2 O EFEITO DA INICIALIZAÇÃO DO ALGORITMO PAP

O modelo desenvolvido por Costa et al. (2012) inclui o efeito da inicialização do algoritmo PAP com passo unitário e ordem de projeção suficiente. Nesta seção será mostrado, de forma resumida, o desenvolvimento feito naquele trabalho. Definindo  $\mathbf{v}(n)$  como o vetor de erro nos coeficientes do filtro adaptativo da forma

$$\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{w}_0 \quad (2.44)$$

é possível escrever a Equação (2.24) na forma:

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}e(n) \quad (2.45)$$

Considerando-se o cenário de identificação de sistemas, em que:

$$d(n) = \mathbf{u}^T(n)\mathbf{w}_0 + r(n) \quad (2.46)$$

e  $r(n)$  é um ruído aditivo de média nula e variância  $\sigma_r^2$ . O sinal de erro, então, pode ser escrito como:

$$e(n) = -\mathbf{u}^T(n)\mathbf{v}(n) + r(n) \quad (2.47)$$

e, assim, a Equação (2.45) se torna:

$$\mathbf{v}(n+1) = \mathbf{v}(n) - \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{u}^T(n)\mathbf{v}(n) + \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}r(n) \quad (2.48)$$

Fazendo a tranposta em (2.15) e substituindo na Equação (2.48):

$$\begin{aligned} \mathbf{v}(n+1) &= \mathbf{v}(n) - \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n) \\ &\quad - \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{U}^T(n)\mathbf{v}(n) + \frac{\boldsymbol{\phi}(n)r(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \end{aligned} \quad (2.49)$$

Assumindo que  $\boldsymbol{\phi}(n)$  é ortogonal às colunas de  $\mathbf{U}(n)$ , chega-se a duas propriedades (RUPP, 1998):

$$\mathbf{u}^T(n)\mathbf{v}(n+1) = r(n) \quad (2.50)$$

$$\mathbf{U}^T(n)\mathbf{v}(n+1) = \mathbf{U}^T(n)\mathbf{v}(n) \quad (2.51)$$

Combinando ambas as equações recursivamente chega-se a:

$$\mathbf{U}^T(n)\mathbf{v}(n) = \mathbf{r}(n-1) \quad (2.52)$$

em que:

$$\mathbf{r}(n-1) = \begin{bmatrix} r(n-1) & r(n-2) & \cdots & r(n-P) \end{bmatrix}^T \quad (2.53)$$

Se, na Equação (2.52), aplicar-se  $n = 0$ , tem-se  $\mathbf{U}^T(0)\mathbf{v}(0) = \mathbf{r}(-1)$ . Como, em geral, o vetor de inicialização  $\mathbf{w}(0)$  é arbitrário tal Equação pode não ser satisfeita, gerando assim o efeito da inicialização. Dessa forma, o trabalho proposto por Costa et al. (2012) adiciona um termo à Equação (2.52) para obter um novo modelo que descreva o comportamento do algoritmo quando o algoritmo opera sob as três condições descritas no trabalho. Baseando-se no descrito, a Equação (2.52) é modificada para:

$$\mathbf{U}^T(n)\mathbf{v}(n) = \mathbf{r}(n-1) + \mathbf{q}(n) \quad (2.54)$$

em que:

$$\mathbf{q}(n) = \mathbf{Q}(-n+1)[\mathbf{U}^T(n)\mathbf{v}(0) - \mathbf{r}(n-1)] \quad (2.55)$$

onde:

$$\mathbf{Q}(-n+1) = \begin{bmatrix} u_{-1}(-n+1) & 0 & \cdots & 0 \\ 0 & u_{-1}(-n+2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & u_{-1}(-n+P) \end{bmatrix} \quad (2.56)$$

é uma matriz  $P \times P$  e  $u_{-1}(n)$  é a função degrau unitário.



### 2.3 O EFEITO DA ORDEM INSUFICIENTE NO ALGORITMO PAP

Como consequência da implementação do algoritmo PAP com ordem deficiente, a matriz  $\mathbf{R}_\phi$  não é mais da forma mostrada pela Equação (2.32), conforme mostrado por Almeida, Costa e Bermudez (2010). Considerando, agora, que o vetor  $\mathbf{a}$  é:

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_P & a_{P+1} & \cdots & a_H \end{bmatrix}^T \quad (2.57)$$

e ainda definindo:

$$\mathbf{b} = \begin{bmatrix} a_1 & a_2 & \cdots & a_P \end{bmatrix}^T \quad (2.58)$$

$$\mathbf{c} = \begin{bmatrix} a_{P+1} & a_{P+2} & \cdots & a_H \end{bmatrix}^T \quad (2.59)$$

como subvetores do vetor  $\mathbf{a}$  e a matriz  $\bar{\mathbf{U}}(n)$  como:

$$\bar{\mathbf{U}}(n) = \begin{bmatrix} \mathbf{u}(n-P-1) & \mathbf{u}(n-P-2) & \cdots & \mathbf{u}(n-H) \end{bmatrix} \quad (2.60)$$

o trabalho de Almeida, Costa e Bermudez (2010) demonstra que, para o caso de ordem insuficiente, a matriz  $\mathbf{R}_\phi$  se torna:

$$\mathbf{R}_\phi = \frac{N-P}{N} \sigma_z^2 \mathbf{I} + \Delta \quad (2.61)$$

em que:

$$\Delta = E\{\mathbf{P}_\perp(n) \bar{\mathbf{U}}(n) \mathbf{c} \mathbf{c}^T \bar{\mathbf{U}}^T(n) \mathbf{P}_\perp^T(n)\} \quad (2.62)$$

em que  $\mathbf{P}_\perp(n)$  é a matriz de projeção no espaço complementar às colunas de  $\mathbf{U}(n)$ . Assim, a matriz  $\Delta$  é completamente definida pelas estatísticas do sinal de entrada.

Outra consequência é que não é mais possível assumir que  $\hat{\mathbf{a}}(n) \approx \mathbf{a}$ , uma vez que  $\hat{\mathbf{a}}(n)$ , dado por (2.17), estima apenas  $P$  coeficientes. No mesmo trabalho é mostrado que o estimador  $\hat{\mathbf{a}}(n)$  em (2.17) pode ser escrito como:

$$\hat{\mathbf{a}}(n) = \mathbf{b} + [\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)[\bar{\mathbf{U}}(n)\mathbf{c} + \mathbf{z}(n)] \quad (2.63)$$

Como conclusão, na situação em que o algoritmo PAP opera com ordem deficiente, o estimador  $\hat{\mathbf{a}}(n)$  torna-se polarizado com relação aos primeiros  $P$  coeficientes de  $\mathbf{a}$ .

## 2.4 RESULTADOS OBTIDOS NO TRABALHO

Diante do exposto, este trabalho concentrar-se-á em obter um modelo teórico para o comportamento médio do algoritmo PAP utilizando passo não unitário e considerando que o algoritmo opera em uma situação de ordem deficiente ( $P < H$ ). Complementarmente, será incluído no modelo, parcelas responsáveis por corrigir o problema de inicialização do algoritmo PAP sob o cenário exposto.

Como resultado deste trabalho, além de obter um modelo teórico de comportamento estocástico do algoritmo PAP, obteve-se uma forma generalizada para o resultado da estimativa feita por  $\hat{\mathbf{a}}(n)$  e uma nova hipótese foi feita sobre o vetor  $\mathbf{e}_a(n-1)$  e que leva a uma modelagem mais acurada para sinais de entrada com alta correlação temporal.

### 3 ANÁLISE ESTOCÁSTICA DO ALGORITMO DE PSEUDO PROJEÇÕES AFINS

Neste capítulo obter-se-á um modelo teórico de comportamento estocástico do algoritmo de Pseudo Projeções Afins considerando o cenário de ordem insuficiente ( $P < H$ ) utilizando passo não unitário e corrigindo o problema de inicialização. Para tanto, o vetor de sinal de entrada  $\mathbf{u}(n)$  será da forma dada pela Equação (2.27) e aqui serão assumidas as seguintes hipóteses:

1. O sinal autoregressivo  $u(n)$  é estacionário, gaussiano, com variância  $\sigma_u^2$ , média nula e de ordem  $H$ .
2. O ruído aditivo  $r(n)$  é estacionário, branco, gaussiano, com variância  $\sigma_r^2$ , possui média nula e é independente de qualquer outro sinal.
3. O sinal de excitação  $z(n)$  do processo AR é estacionário, branco, gaussiano, com variância  $\sigma_z^2$  e média nula.
4. Os vetores  $\mathbf{z}(n-k)$  e  $\mathbf{w}(n-p)$  são estatisticamente independentes para quaisquer  $k$  e  $p$ .
5. Os vetores  $\mathbf{z}(n-k)$  e  $\mathbf{z}(n-p)$  são estatisticamente independentes para quaisquer  $k \neq p$ .
6. Os vetores  $\phi(n)$  e  $\mathbf{w}(n)$  são considerados estatisticamente independentes. Essa hipótese equivale à teoria da independência apresentada em Haykin (2002).

7. O número de coeficientes do filtro adaptativo é tal qual  $N \gg P$ . Como consequência, conforme mostrado por Bershad, Linebarger e McLaughlin (2001), as flutuações em torno do vetor  $\hat{\mathbf{a}}(n)$  podem ser desprezadas e o mesmo pode ser aproximado por um vetor invariante no tempo.
8. As componentes do vetor  $\phi(n)$  possuem uma distribuição gaussiana de média nula.

### 3.1 CORREÇÃO DO PROBLEMA DE INICIALIZAÇÃO

Substituindo a Expressão (2.47) do erro  $e(n)$  na Equação (2.25), tem-se:

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \mu \frac{\phi(n)}{\phi^T(n)\phi(n)} [-\mathbf{u}^T(n)\mathbf{v}(n) + r(n)] \quad (3.1)$$

Para prosseguir com o equacionamento é preciso encontrar uma forma fechada para  $\mathbf{u}^T(n)\mathbf{v}(n)$ . Transpondo (2.15) e pós multiplicando por  $\mathbf{v}(n)$  ambos os lados da expressão resultante, tem-se:

$$\mathbf{u}^T(n)\mathbf{v}(n) = \phi^T(n)\mathbf{v}(n) + \hat{\mathbf{a}}^T(n)\mathbf{U}^T(n)\mathbf{v}(n) \quad (3.2)$$

Das definições (2.15)-(2.17) segue que  $\phi(n)$  é ortogonal às colunas da matriz  $\mathbf{U}(n)$ . Assim, pré-multiplicando a Equação (3.1) por  $\mathbf{u}^T(n)$ , tem-se:

$$\mathbf{u}^T(n)\mathbf{v}(n+1) = \mathbf{u}^T(n)\mathbf{v}(n) - \mu \frac{\mathbf{u}^T(n)\phi(n)\mathbf{u}^T(n)\mathbf{v}(n)}{\phi^T(n)\phi(n)} + \mu \frac{\mathbf{u}^T(n)\phi(n)r(n)}{\phi^T(n)\phi(n)} \quad (3.3)$$

Utilizando a definição dada por (2.15), chega-se a:

$$\mathbf{u}^T(n)\mathbf{v}(n+1) = (1 - \mu)\mathbf{u}^T(n)\mathbf{v}(n) + \mu r(n) \quad (3.4)$$

Pré-multiplicando a Equação (3.1) por  $\mathbf{U}^T(n)$  e utilizando o fato de que  $\mathbf{U}^T(n)\boldsymbol{\phi}(n) = \mathbf{0}_{P \times 1}$  obtém-se:

$$\mathbf{U}^T(n)\mathbf{v}(n+1) = \mathbf{U}^T(n)\mathbf{v}(n) \quad (3.5)$$

Combinando as equações (3.4) e (3.5) recursivamente, obtém-se:

$$\mathbf{U}^T(n)\mathbf{v}(n) = (1 - \mu)\mathbf{e}_a(n-1) + \mu r(n-1) \quad (3.6)$$

em que:

$$\mathbf{e}_a(n-1) = \begin{bmatrix} e_a(n-1) & e_a(n-2) & \cdots & e_a(n-P) \end{bmatrix}^T \quad (3.7)$$

e

$$e_a(n) = \mathbf{u}^T(n)\mathbf{v}(n) \quad (3.8)$$

Considerando agora o efeito da inicialização do algoritmo PAP, proceder-se-á de forma similar a Costa et al. (2012). Assim:

$$\mathbf{u}^T(n)\mathbf{v}(n+1) = \begin{cases} \mathbf{u}^T(n)[\mathbf{w}(0) - \mathbf{w}_o] & n \leq 0 \\ (1 - \mu)\mathbf{u}^T(n)\mathbf{v}(n) + \mu r(n) & n > 0 \end{cases} \quad (3.9)$$

$$\mathbf{U}^T(n)\mathbf{v}(n+1) = \begin{cases} \mathbf{U}^T(n)[\mathbf{w}(0) - \mathbf{w}_o] & n \leq 0 \\ \mathbf{U}^T(n)\mathbf{v}(n) & n > 0 \end{cases} \quad (3.10)$$

Das linhas da Equação (3.10), adotando o índice  $i$  para as colunas

de  $\mathbf{U}(n)$ , pode-se deduzir que:

$$\mathbf{u}^T(n-i)\mathbf{v}(n) = \begin{cases} \mathbf{u}^T(n-i)[\mathbf{w}(0) - \mathbf{w}_o] & n \leq 0 \\ \mathbf{u}^T(n-i)[\mathbf{w}(0) - \mathbf{w}_o] & 0 < n < i \\ (1 - \mu)\mathbf{u}^T(n-i)\mathbf{v}(n-i) + \mu\mathbf{r}(n-i) & n \geq i \end{cases} \quad (3.11)$$

A Equação (3.11) pode ser reescrita da seguinte maneira:

$$\mathbf{u}^T(n-i)\mathbf{v}(n) = (1 - \mu)e_a(n-i) + \mu\mathbf{r}(n-i) + q_i(n) \quad (3.12)$$

em que:

$$q_i(n) = [\mathbf{u}^T(n-i)\mathbf{v}(0) - (1 - \mu)\mathbf{u}^T(n-i)\mathbf{v}(n-i) - \mu\mathbf{r}(n-i)]u_{-1}(i-n) \quad (3.13)$$

em que  $u_{-1}(n)$  é a função degrau unitário. Na forma vetorial:

$$\mathbf{U}^T(n)\mathbf{v}(n) = (1 - \mu)\mathbf{e}_a(n-1) + \mu\mathbf{r}(n-1) + \mathbf{q}(n) \quad (3.14)$$

$$\mathbf{q}(n) = \mathbf{Q}(-n+1)[\mathbf{U}^T(n)\mathbf{v}(0) - (1 - \mu)\mathbf{e}_a(n-1) - \mu\mathbf{r}(n-1)] \quad (3.15)$$

Substituindo (3.15) em (3.14), tem-se:

$$\begin{aligned} \mathbf{U}^T(n)\mathbf{v}(n) &= (1 - \mu)\mathbf{e}_a(n-1) + \mu\mathbf{r}(n-1) \\ &+ \mathbf{Q}(-n+1)[\mathbf{U}^T(n)\mathbf{v}(0) - (1 - \mu)\mathbf{e}_a(n-1) - \mu\mathbf{r}(n-1)] \end{aligned} \quad (3.16)$$

Ao considerar passo de adaptação não unitário, os termos envolvendo o vetor  $\mathbf{e}_a(n-1)$  aparecem na Equação (3.30) no trabalho apresentado em (BARCELOS, 2013). Em (BARCELOS, 2013), em vir-

tude da utilização de passo unitário, não foi necessário efetuar uma hipótese sobre o vetor  $\mathbf{e}_a(n-1)$ , sendo isto um fator limitante para considerar passo não unitário na análise do comportamento estocástico do algoritmo PAP.

A hipótese sobre o vetor  $\mathbf{e}_a(n-1)$  apresentada por Almeida, Bermudez e Bershada (2009) consiste em assumir que as componentes deste vetor são aproximadamente iguais. Desta forma, conforme explorado anteriormente, o processo de filtragem das  $P$  componentes do erro *a priori* a cada iteração é eliminado do modelo apresentado por Almeida, Bermudez e Bershada (2009). Logo, uma aproximação para o comportamento das componentes do vetor  $\mathbf{e}_a(n-1)$  deve ser obtida para prosseguir com o equacionamento do comportamento do algoritmo PAP.

A fim de obter uma aproximação para o vetor  $\mathbf{e}_a(n-1)$ , que é composto por  $P$  amostras do erro *a priori* de estimação, é necessário conhecer como  $e_a(n)$  varia ao longo das iterações. A Figura 10 mostra o comportamento de  $e_a(n)$  nas seguintes condições: sinal AR da forma  $u(n) = -0.4u(n-1) + z(n)$ , com  $\sigma_z^2 = 0.8469$ ,  $P = 1$ , planta desconhecida consistindo em uma janela de Hann normalizada com  $N = 64$  coeficientes, passo de adaptação  $\mu = 0.8$  e o vetor de coeficientes do filtro adaptativo inicializado no vetor nulo.

Com o conhecimento, mostrado na Figura 10, de que o erro *a priori* de estimação cai exponencialmente com o tempo, optou-se por modelar as componentes do vetor  $\mathbf{e}_a(n-1)$  com um modelo de decaimento exponencial controlado por um autovalor da matriz de correlação

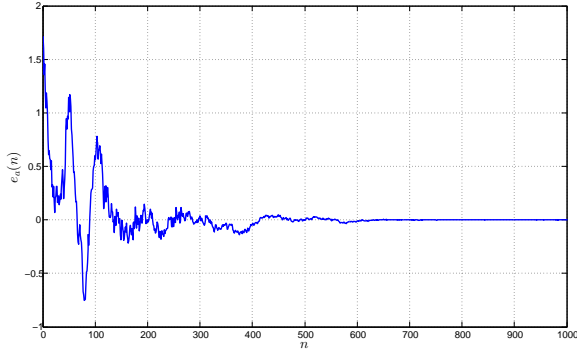


Figura 10 – Comportamento instantâneo de  $e_a(n)$

do sinal de entrada  $\mathbf{u}(n)$ . O autovalor escolhido para esta modelagem deve produzir o modo de convergência mais lento e oscilatório, uma vez que o sinal de erro é de média nula. Essa opção é feita baseada nos modos de convergência do algoritmo *steepest-descent*, conforme mostrado em (HAYKIN, 2002). É válido notar que o decaimento exponencial de  $e_a(n)$  é sempre observado quando o vetor de inicialização não é setado em  $\mathbf{w}_o$ . Deste modo, o vetor  $\mathbf{e}_a(n-1)$  se torna:

$$\mathbf{e}_a(n-1) = \begin{bmatrix} (1-\mu\lambda)^{n-1} & (1-\mu\lambda)^{n-2} & \dots & (1-\mu\lambda)^{n-P} \end{bmatrix}^T \quad (3.17)$$

em que  $\lambda$  é o autovalor escolhido da matriz de correlação do sinal de entrada  $x(n)$  que produza  $(1-\mu\lambda)$  o mais próximo de  $-1$ . Retomando a Equação (3.2), substituindo a Equação (3.16) e utilizando a Equação (3.17), tem-se:



$$\begin{aligned}
e_a(n) &= \boldsymbol{\phi}^T(n)\mathbf{v}(n) + (1 - \mu)\beta(n) + \mu\hat{\mathbf{a}}^T(n)\mathbf{r}(n - 1) \\
&+ \hat{\mathbf{a}}^T(n)\mathbf{Q}(-n + 1)\mathbf{U}^T(n)\mathbf{v}(0) - \mu\hat{\mathbf{a}}^T(n)\mathbf{Q}(-n + 1)\mathbf{r}(n - 1) \quad (3.18)
\end{aligned}$$

em que:

$$\beta(n) = \sum_{i=1}^P \hat{a}_i(n)(1 - \mu\lambda)^{n-i}[1 - u_{-1}(-n + i)] \quad (3.19)$$

Substituindo a Equação (3.18) na Equação (3.1), juntamente com a Equação (3.19), chega-se finalmente a:

$$\begin{aligned}
\mathbf{v}(n + 1) &= \mathbf{v}(n) - \mu \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n) - \mu(1 - \mu) \frac{\boldsymbol{\phi}(n)\beta(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
&- \mu^2 \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)\mathbf{r}(n - 1)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} + \mu \frac{\boldsymbol{\phi}(n)r(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
&- \mu \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)\mathbf{Q}(-n + 1)\mathbf{U}^T(n)\mathbf{v}(0)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
&+ \mu^2 \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)\mathbf{Q}(-n + 1)\mathbf{r}(n - 1)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \quad (3.20)
\end{aligned}$$

Definindo ainda o ruído filtrado  $r_a(n)$  como

$$r_a(n) = r(n) - \mu\hat{\mathbf{a}}^T(n)\mathbf{r}(n - 1) \quad (3.21)$$

obtém-se:

$$\begin{aligned}
 \mathbf{v}(n+1) = \mathbf{v}(n) &- \mu \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \mathbf{v}(n) \\
 &- \mu(1-\mu) \frac{\boldsymbol{\phi}(n)\beta(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
 &- \mu \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
 &+ \mu^2 \frac{\boldsymbol{\phi}(n)\hat{\mathbf{a}}^T(n)\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \\
 &+ \mu \frac{\boldsymbol{\phi}(n)r_a(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \tag{3.22}
 \end{aligned}$$

A fim de validar o a Equação (3.22), foi efetuada uma simulação computacional com os seguintes parâmetros: sinal de entrada  $AR$  da forma  $u(n) = 1.96u(n-1) - 0.9893u(n-2) + z(n)$ , com  $\sigma_z^2 = 0.0575$ ,  $\lambda = 2.3192$ ,  $N = 64$ ,  $P = 2$ , planta a estimar consistindo em uma janela de Hann normalizada e passo de adaptação  $\mu = 0.8$ . A escolha desses parâmetros se dá pelo fato de que tais parâmetros são utilizados em (BARCELOS, 2013) para explorar o efeito da inicialização. A Figura 11 avalia o comportamento do coeficiente de número 32 do vetor  $\mathbf{v}(n)$ .

Nota-se que durante o período transitório mostrado na Figura 11, não há uma perfeita acurácia da previsão feita pela Equação (3.22). Isso se deve ao fato da aproximação utilizada para o vetor  $\mathbf{e}_a(n-1)$  não ser a mais acurada. Assim, ainda há espaço para melhores aproximações das componentes do vetor  $\mathbf{e}_a(n-1)$ .

Em (ALMEIDA; COSTA; BERMUDEZ, 2010), as componentes de  $\hat{\mathbf{a}}(n)$  são determinadas a partir das equações (2.58)-(2.60). Assim, uti-

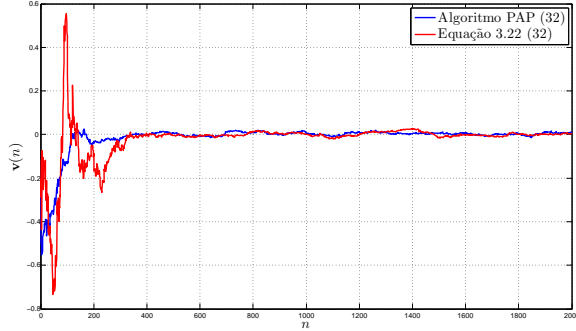


Figura 11 – Comparação do algoritmo PAP com a Equação (3.22)

lizando a definição dada pela Equação (2.27), o estimador  $\hat{\mathbf{a}}(n)$  se torna:

$$\hat{\mathbf{a}}(n) = \mathbf{b} + [\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)[\bar{\mathbf{U}}(n)\mathbf{c} + \mathbf{z}(n)] \quad (3.23)$$

Tomando o valor esperado da Equação (3.23), tem-se:

$$\hat{\mathbf{a}}(n) \approx E\{\hat{\mathbf{a}}(n)\} = \mathbf{b} + E\{[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)\bar{\mathbf{U}}(n)\mathbf{c}\} \quad (3.24)$$

Nota-se que a Equação (3.24) depende de uma estimação numérica de estatísticas de ordem elevada do sinal de entrada. Entretanto, conforme mostrado em Kay (1993), se um sistema da forma:

$$u(n) - \sum_{i=1}^P a_i u(n-i) = \phi(n) \quad (3.25)$$

é estável, o estimador  $\mathbf{a}_w$  de mínimo erro quadrático médio do vetor de  $P$  coeficientes  $\mathbf{a}$  é dado pela solução do sistema de equações normais:

$$\mathbf{R}_{u_P \times P} \mathbf{a}_w = -\mathbf{p}_u \quad (3.26)$$

em que  $\mathbf{R}_{u_P \times P}$  é a matriz de correlação do sinal de entrada  $u(n)$  de

ordem  $P$  e  $\mathbf{p}_u$  é definido como:

$$\mathbf{p}_u = \begin{bmatrix} r_u(1) & r_u(2) & \cdots & r_u(P) \end{bmatrix}^T \quad (3.27)$$

em que  $r_u(k)$  é a sequência de correlação do sinal de entrada  $u(n)$ . Dessa forma, assumindo que  $\mathbf{R}_{u_{P \times P}}$  é definida positiva, obtém-se:

$$\hat{\mathbf{a}}(n) \approx E\{\hat{\mathbf{a}}(n)\} = \mathbf{a}_w = -\mathbf{R}_{u_{P \times P}}^{-1} \mathbf{p}_u \quad (3.28)$$

A Equação (3.28) é conhecida como equação de *Yule-Walker* (KAY, 1993). O vetor  $\mathbf{a}_w$  pode ser considerado um bom estimador do vetor  $\mathbf{a}$  se o algoritmo PAP opera com ordem suficiente. A Equação (3.28) representa uma generalização dos resultados apresentados nos trabalhos anteriores que consideravam apenas o cenário de ordem suficiente. Retomando a Equação (3.19), tem-se:

$$\beta(n) = \sum_{\tilde{i}=1}^P a_{w_{\tilde{i}}} (1 - \mu\lambda)^{n-\tilde{i}} [1 - u_{-1}(-n + \tilde{i})] \quad (3.29)$$

### 3.2 COMPORTAMENTO MÉDIO DOS COEFICIENTES

Tomando o valor esperado da Equação (3.22), tem-se:

$$\begin{aligned}
E\{\mathbf{v}(n+1)\} &= E\{\mathbf{v}(n)\} - \mu \overbrace{E\left\{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n)\right\}}^1 \\
&\quad - \mu(1-\mu) \overbrace{E\left\{\frac{\boldsymbol{\phi}(n)\beta(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}}^2 + \mu \overbrace{E\left\{\frac{\boldsymbol{\phi}(n)r_a(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}}^3 \\
&\quad - \mu \overbrace{E\left\{\frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}}^4 \\
&\quad + \mu^2 \overbrace{E\left\{\frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}}^5
\end{aligned} \tag{3.30}$$

Sob a luz da Hipótese 6, é possível escrever o primeiro termo da Equação (3.30) como:

$$E\left\{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n)\right\} = E\left\{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}E\{\mathbf{v}(n)\} \tag{3.31}$$

E ainda, conforme mostrado por Almeida et al. (2005):

$$E\left\{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}E\{\mathbf{v}(n)\} \approx E\left\{\frac{1}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\}\mathbf{R}_\phi E\{\mathbf{v}(n)\} \tag{3.32}$$

em que  $\mathbf{R}_\phi$  é a matriz de correlação de  $\boldsymbol{\phi}(n)$ , dada pela Equação (2.61). Como  $\beta(n)$  é uma grandeza determinística, conforme mostrado pela Equação (3.19) e, como uma consequência da Hipótese 1, o segundo valor esperado da Equação (3.30) se torna:

$$E\left\{\frac{\boldsymbol{\phi}(n)\beta(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\} = \beta(n)E\left\{\frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\right\} = \mathbf{0}_{N \times 1} \tag{3.33}$$

Com base na Hipótese 2, o terceiro valor esperado se torna:

$$E \left\{ \frac{\boldsymbol{\phi}(n)r_a(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} = E \left\{ \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} E \{r_a(n)\} = \mathbf{0}_{N \times 1} \quad (3.34)$$

De maneira similar ao mostrado por Barcelos (2013), e uma vez que  $\mathbf{v}(0)$  é determinístico, o quarto valor esperado se torna:

$$E \left\{ \frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} = E \left\{ \frac{1}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} \mathbf{R}_{\phi u}(n) \mathbf{v}(0) \quad (3.35)$$

em que a matriz:

$$\mathbf{R}_{\phi u}(n) = E \left\{ \boldsymbol{\phi}(n)\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \right\} \quad (3.36)$$

é completamente definida pelas estatísticas do sinal de entrada.

De forma similar ao terceiro valor esperado, o último termo da Equação (3.30) se torna:

$$E \left\{ \frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} = \mathbf{0}_{N \times 1} \quad (3.37)$$

Ainda, conforme mostrado por Almeida et al. (2005):

$$E \left\{ \frac{1}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} \simeq \frac{1}{\sigma_\phi^2(G-2)} \quad (3.38)$$

em que:

$$G = N - P \quad (3.39)$$

Agrupando os resultados para os valores esperados da Equação (3.30), esta se torna:

$$E\{\mathbf{v}(n+1)\} = \left( \mathbf{I} - \frac{\mu \mathbf{R}_\phi}{\sigma_\phi^2(G-2)} \right) E\{\mathbf{v}(n)\} - \frac{\mu \mathbf{R}_{\phi u}(n) \mathbf{v}(0)}{\sigma_\phi^2(G-2)} \quad (3.40)$$

Em um cenário de ordem suficiente, a matriz  $\mathbf{\Delta}$  na Equação (3.36) é nula. Dessa forma, aplicando passo unitário e ordem suficiente a Equação (3.40) se torna:

$$E\{\mathbf{v}(n+1)\} = \left( 1 - \frac{1}{G-2} \right) E\{\mathbf{v}(n)\} - \frac{N}{G(G-2)\sigma_z^2} \mathbf{R}_{\phi u}(n) \mathbf{v}(0) \quad (3.41)$$

A Equação (3.41) é exatamente a Equação (4.11) do trabalho feito por Barcelos (2013). A Equação (3.40) é uma generalização da Equação que descreve o comportamento médio dos coeficientes em (BARCELOS, 2013).

### 3.3 ERRO QUADRÁTICO MÉDIO

Elevando a Equação (2.47) ao quadrado, tirando o valor esperado, tem-se:

$$E\{e^2(n)\} = E\{r^2(n)\} + E\{e_a^2(n)\} - 2E\{e_a(n)r(n)\} \quad (3.42)$$

Considerando a Hipótese 2, o terceiro termo da equação acima é nulo.

Logo:

$$E\{e^2(n)\} = E\{r^2(n)\} + E\{e_a^2(n)\} \quad (3.43)$$

Substituindo a Equação (3.18) na Equação acima, obtém-se

$$\begin{aligned}
E\{e^2(n)\} &= \overbrace{E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\boldsymbol{\phi}^T(n)\mathbf{v}(n)\}}^1 + (1-\mu) \overbrace{E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\beta(n)\}}^2 \\
&+ \mu \overbrace{E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w\mathbf{r}(n-1)\}}^3 \\
&+ \overbrace{E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\}}^4 \\
&- \mu \overbrace{E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)\}}^5 \\
&+ (1-\mu) \overbrace{E\{\beta(n)\boldsymbol{\phi}^T(n)\mathbf{v}(n)\}}^6 + (1-\mu)^2 \overbrace{E\{\beta^2(n)\}}^7 \\
&+ (1-\mu)\mu \overbrace{E\{\beta(n)\mathbf{a}_w^T\mathbf{r}(n-1)\}}^8 \\
&- (1-\mu) \overbrace{E\{\beta(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\}}^9 \\
&- (1-\mu) \overbrace{E\{\beta(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)\}}^{10} \\
&+ \mu \overbrace{E\{\mathbf{a}_w^T\mathbf{r}(n-1)\boldsymbol{\phi}^T(n)\mathbf{v}(n)\}}^{11} \\
&+ (1-\mu)\mu \overbrace{E\{\mathbf{a}_w^T\mathbf{r}(n-1)\beta(n)\}}^{12} + \mu^2 \overbrace{E\{[\mathbf{a}_w^T\mathbf{r}(n-1)]^2\}}^{13} \\
&+ \mu \overbrace{E\{\mathbf{a}_w^T\mathbf{r}(n-1)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\}}^{14} \\
&- \mu^2 \overbrace{E\{\mathbf{a}_w^T\mathbf{r}(n-1)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)\}}^{15} \\
&+ \overbrace{E\{\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\boldsymbol{\phi}^T(n)\mathbf{v}(n)\}}^{16} \\
&+ (1-\mu) \overbrace{E\{\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\beta(n)\}}^{17}
\end{aligned} \tag{3.44}$$



$$\begin{aligned}
& + \mu \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{a}_w^T \mathbf{r}(n-1)\}}^{18} \\
& + \overbrace{E\{[\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)]^2\}}^{19} \\
& - \mu \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \boldsymbol{\phi}^T(n) \mathbf{v}(n)\}}^{20} \\
& - (1-\mu)\mu \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \beta(n)\}}^{21} \\
& - \mu^2 \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \mathbf{a}_w^T \mathbf{r}(n-1)\}}^{22} \\
& - \mu^2 \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)\}}^{23} \\
& + \mu^2 \overbrace{E\{[\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)]^2\}}^{24} + \overbrace{E\{r^2(n)\}}^{25} \\
& - \mu \overbrace{E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)\}}^{26}
\end{aligned}$$

Como todos os termos da Equação (3.44) são escalares, pode-se escrever, para o primeiro termo:

$$E\{\boldsymbol{\phi}^T(n) \mathbf{v}(n) \boldsymbol{\phi}^T(n) \mathbf{v}(n)\} = E\{\text{Tr}[\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \mathbf{v}(n) \mathbf{v}^T(n)]\} \quad (3.45)$$

E ainda, considerando a Hipótese 6, tem-se:

$$E\{\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \mathbf{v}(n) \mathbf{v}^T(n)\} = \mathbf{R}_\phi \mathbf{K}(n) \quad (3.46)$$

em que:

$$\mathbf{K}(n) = E\{\mathbf{v}(n) \mathbf{v}^T(n)\} \quad (3.47)$$

Logo, o primeiro termo da Equação (3.44) se torna:

$$E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\boldsymbol{\phi}^T(n)\mathbf{v}(n)\} = \text{Tr} [\mathbf{R}_\phi \mathbf{K}(n)] \quad (3.48)$$

Como consequência da Hipótese 1, o vetor  $\boldsymbol{\phi}(n)$  possui média nula. Logo, utilizando a Hipótese 6, o segundo e o sexto valor esperado resulta em:

$$E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\beta(n)\} = E\{\boldsymbol{\phi}^T(n)\}E\{\mathbf{v}(n)\beta(n)\} = 0 \quad (3.49)$$

Considerando a Hipótese 2, o terceiro e o décimo primeiro valores esperados se tornam:

$$E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T \mathbf{r}(n-1)\} = E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T\}E\{\mathbf{r}(n-1)\} = 0 \quad (3.50)$$

De forma similar ao feito em Barcelos (2013), considerando a Hipótese 6 e a definição dada pela Equação (3.36), o quarto e o décimo sexto termos da Equação (3.30) são determinados por:

$$E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\} = E\{\mathbf{v}^T(n)\}\mathbf{R}_{\phi_u}(n)\mathbf{v}(0) \quad (3.51)$$

Sob a mesma argumentação utilizada para determinar o terceiro valor esperado, o quinto e o vigésimo valores esperados se tornam:

$$E\{\boldsymbol{\phi}^T(n)\mathbf{v}(n)\mathbf{a}_w^T \mathbf{Q}(-n+1)\}E\{\mathbf{r}(n-1)\} = 0 \quad (3.52)$$

Por se tratar de uma grandeza determinística, o sétimo valor esperado é determinado pelo seu valor instantâneo:

$$E\{\beta^2(n)\} = \beta^2(n) \quad (3.53)$$

Sob a luz da Hipótese 2, o oitavo e o décimo segundo valores esperados podem ser determinados como:

$$E\{\beta(n)\mathbf{a}_w^T \mathbf{r}(n-1)\} = E\{\beta(n)\mathbf{a}_w^T\}E\{\mathbf{r}(n-1)\} = 0 \quad (3.54)$$

Considerando a Hipótese 1 e usando a definição da matriz  $\mathbf{Q}(-n+1)$ , o nono e o décimo sétimo termo são nulos:

$$\begin{aligned} E\{\beta(n)\mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\} &= \beta(n)\mathbf{a}_w \mathbf{Q}(-n+1)E\{\mathbf{U}^T(n)\}\mathbf{v}(0) \\ &= 0 \end{aligned} \quad (3.55)$$

O décimo e o vigésimo primeiro termos da Equação (3.44), sob a Hipótese 2, se tornam:

$$E\{\beta(n)\mathbf{a}_w^T \mathbf{Q}(-n+1)\}E\{\mathbf{r}(n-1)\} = 0 \quad (3.56)$$

O décimo terceiro da Equação (3.44), considerando a Hipótese 2 e a Equação (3.28), é determinado por:

$$\begin{aligned} E\{[\mathbf{a}_w^T \mathbf{r}(n-1)]^2\} &= \mathbf{a}_w^T E\{\mathbf{r}(n-1)\mathbf{r}^T(n-1)\}\mathbf{a}_w \\ &= \mathbf{a}_w^T \mathbf{a}_w \sigma_r^2 \end{aligned} \quad (3.57)$$

Levando em consideração a Hipótese 2, o décimo quarto e o décimo oitavo termos são igualmente nulos.

$$E\{\mathbf{r}^T(n-1)\}E\{\mathbf{a}_w \mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\} = 0 \quad (3.58)$$

De forma similar ao décimo terceiro termo e ainda considerando que  $\mathbf{Q}^2(-n+1) = \mathbf{Q}(-n+1)$ , o décimo quinto, o vigésimo segundo e o vigésimo quarto termos da Equação (3.44) se tornam:

$$E\{[\mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{r}(n-1)]^2\} = \mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{a}_w \sigma_r^2 \quad (3.59)$$

De forma similar ao feito por Costa et al. (2012), o décimo nono é dado por:

$$E\{[\mathbf{a}_w^T \mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)]^2\} = \mathbf{v}^T(0)\mathbf{R}_{\tilde{u}}(n)\mathbf{v}(0) \quad (3.60)$$

em que a matriz  $\mathbf{R}_{\tilde{u}}(n)$  é definida como:

$$\mathbf{R}_{\tilde{u}}(n) = E\{\mathbf{U}(n)\mathbf{Q}(-n+1)\mathbf{a}_w \mathbf{a}_w^T \mathbf{Q}^T(-n+1)\mathbf{U}^T(n)\} \quad (3.61)$$

Considerando a Hipótese 2, o vigésimo quinto termo da Equação (3.44) é:

$$E\{r^2(n)\} = \sigma_r^2 \quad (3.62)$$

Por fim, considerando a Hipótese 2, o último termo da Equação (3.44) é nulo:

$$\begin{aligned} E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)\} = \\ E\{\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{a}_w^T \mathbf{Q}(-n+1)\} E\{\mathbf{r}(n-1)\} = 0; \end{aligned} \quad (3.63)$$

Retomando a Equação (3.44), e substituindo todos os termos calculados, a expressão resulta em:

$$\begin{aligned} E\{e^2(n)\} = (1 - \mu)^2 \beta^2(n) + \text{Tr}\{\mathbf{R}_\phi \mathbf{K}(n)\} \\ + (1 + \mu^2 \mathbf{a}_w^T \mathbf{a}_w) \sigma_r^2 + 2E\{\mathbf{v}^T(n)\} \mathbf{R}_{\phi u}(n) \mathbf{v}(0) \\ + \mathbf{v}^T(0) \mathbf{R}_{\tilde{u}}(n) \mathbf{v}(0) - \mu^2 \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{a}_w \end{aligned} \quad (3.64)$$

Pode-se observar que se a Equação (3.64) fosse particularizada para passo de adaptação unitário e para ordem suficiente, esta Equação se torna a Equação (5.35) de Barcelos (2013). Como seguimento, obter-se-á uma Equação em forma fechada para a matriz  $\mathbf{K}(n)$ .

### 3.4 MATRIZ DE COVARIÂNCIA DO VETOR DE ERRO NOS COEFICIENTES

Pós-multiplicando a Equação (3.22) por sua transposta, substituindo  $\hat{\mathbf{a}}(n)$  por  $\mathbf{a}_w$  e tomando o valor esperado condicionado em  $\mathbf{v}(n)$ <sup>1</sup>, tem-se:

$$\begin{aligned}
 E\{\mathbf{v}(n+1)\mathbf{v}^T(n+1)\} &= \overbrace{E\{\mathbf{v}(n)\mathbf{v}^T(n)\}}^1 - \mu E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n)\mathbf{v}^T(n)}^2}\right\} \\
 &\quad - \mu(1-\mu) E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\beta(n)\mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^3}\right\} + \mu E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)r_a(n)\mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^4}\right\} \\
 &\quad - \mu E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)\mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^5}\right\} \\
 &\quad + \mu^2 E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)\mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^6}\right\} - \mu E\left\{\overbrace{\mathbf{v}(n)\mathbf{v}^T(n)\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^7}\right\} \\
 &\quad + \mu^2 E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}(n)\mathbf{v}^T(n)\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^8}\right\} \\
 &\quad + \mu^2(1-\mu) E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)\beta(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}^T(n)\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^9}\right\} \\
 &\quad - \mu^2 E\left\{\overbrace{\frac{\boldsymbol{\phi}(n)r_a(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}\mathbf{v}^T(n)\frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)}}^{10}}\right\}
 \end{aligned}$$

<sup>1</sup>Por simplicidade, foi omitido dos valores esperados a expressão matemática do condicionamento em  $\mathbf{v}(n)$

$$\begin{aligned}
& \overbrace{+\mu^2 E \left\{ \frac{\phi(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)}{\phi^T(n) \phi(n)} \mathbf{v}^T(n) \frac{\phi(n) \phi^T(n)}{\phi^T(n) \phi(n)} \right\}}^{11} \\
& \overbrace{-\mu^3 E \left\{ \frac{\phi(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)}{\phi^T(n) \phi(n)} \mathbf{v}^T(n) \frac{\phi(n) \phi^T(n)}{\phi^T(n) \phi(n)} \right\}}^{12} \\
& -\mu(1-\mu) E \left\{ \mathbf{v}(n) \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} + \mu^2(1-\mu) E \left\{ \frac{\phi(n) \phi^T(n)}{\phi^T(n) \phi(n)} \mathbf{v}(n) \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} \\
& + \mu^2(1-\mu)^2 E \left\{ \frac{\phi(n) \beta(n)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} \\
& - \mu^2(1-\mu) E \left\{ \frac{\phi(n) r_a(n)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} \\
& + \mu^2(1-\mu) E \left\{ \frac{\phi(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} \\
& - \mu^3(1-\mu) E \left\{ \frac{\phi(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) \beta(n)}{\phi^T(n) \phi(n)} \right\} \\
& + \mu E \left\{ \mathbf{v}(n) \frac{\phi^T(n) r_a(n)}{\phi^T(n) \phi(n)} \right\} - \mu^2 E \left\{ \frac{\phi(n) \phi^T(n)}{\phi^T(n) \phi(n)} \mathbf{v}(n) \frac{\phi^T(n) r_a(n)}{\phi^T(n) \phi(n)} \right\} \\
& - \mu^2(1-\mu) E \left\{ \frac{\phi(n) \beta(n)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) r_a(n)}{\phi^T(n) \phi(n)} \right\} + \mu^2 E \left\{ \frac{\phi(n) r_a(n)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) r_a(n)}{\phi^T(n) \phi(n)} \right\} \\
& - \mu^2 E \left\{ \frac{\phi(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)}{\phi^T(n) \phi(n)} \frac{\phi^T(n) r_a(n)}{\phi^T(n) \phi(n)} \right\}
\end{aligned}$$

$$\begin{aligned}
& +\mu^3 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\boldsymbol{\phi}^T(n) r_a(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{24} \\
& -\mu E \overbrace{\left\{ \mathbf{v}(n) \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{25} \\
& +\mu^2 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \mathbf{v}(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{26} \\
& +\mu^2 (1-\mu) E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \beta(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{27} \\
& -\mu^2 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) r_a(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{28} \\
& +\mu^2 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{29} \\
& -\mu^3 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{30} \\
& +\mu^2 E \overbrace{\left\{ \mathbf{v}(n) \frac{\mathbf{r}^T(n-1) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{31} \\
& -\mu^3 E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \mathbf{v}(n) \frac{\mathbf{r}^T(n-1) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{32} \\
& -\mu^3 (1-\mu) E \overbrace{\left\{ \frac{\boldsymbol{\phi}(n) \beta(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \frac{\mathbf{r}^T(n-1) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\}}^{33}
\end{aligned}$$



$$\begin{aligned}
& + \mu^3 E \left\{ \overbrace{\frac{\phi(n)r_a(n)}{\phi^T(n)\phi(n)} \frac{\mathbf{r}^T(n-1)\mathbf{Q}^T(-n+1)\mathbf{a}_w\phi^T(n)}{\phi^T(n)\phi(n)}}^{34} \right\} \\
& - \mu^3 E \left\{ \overbrace{\frac{\phi(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)}{\phi^T(n)\phi(n)} \frac{\mathbf{r}^T(n-1)\mathbf{Q}^T(-n+1)\mathbf{a}_w\phi^T(n)}{\phi^T(n)\phi(n)}}^{35} \right\} \\
& + \mu^4 E \left\{ \overbrace{\frac{\phi(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{\phi^T(n)\phi(n)} \frac{\mathbf{r}^T(n-1)\mathbf{Q}^T(-n+1)\mathbf{a}_w\phi^T(n)}{\phi^T(n)\phi(n)}}^{36} \right\}
\end{aligned} \tag{3.65}$$

Considerando a Hipótese 6 e, de forma similar ao feito por Almeida et al. (2005), utilizando a Equação (3.38), o segundo e o sétimo termo da Equação (3.65) se tornam:

$$E \left\{ \frac{\phi(n)\phi^T(n)}{\phi^T(n)\phi(n)} \mathbf{v}(n)\mathbf{v}^T(n) \right\} \simeq \frac{1}{\sigma_\phi^2(G-2)} \mathbf{R}_\phi \mathbf{v}(n)\mathbf{v}^T(n) \tag{3.66}$$

Da mesma forma e usando adicionalmente a Hipótese 1, o terceiro e o décimo terceiro termo podem ser determinados:

$$E \left\{ \frac{\phi(n)\beta(n)\mathbf{v}^T(n)}{\phi^T(n)\phi(n)} \right\} \simeq \beta(n) E \left\{ \frac{\phi(n)}{\phi^T(n)\phi(n)} \right\} \mathbf{v}^T(n) = \mathbf{0}_{N \times N} \tag{3.67}$$

Devido às características estatísticas de  $r(n)$ , mostradas na Hipótese 2, o quarto e o décimo nono termos da Equação (3.65) são nulos.

$$E \left\{ \frac{\boldsymbol{\phi}(n) r_a(n) \mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \simeq E \{ r_a(n) \} E \left\{ \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \mathbf{v}^T(n) = \mathbf{0}_{N \times N} \quad (3.68)$$

Para calcular os valores esperados do quinto e do vigésimo quinto, proceder-se-á de forma similar a Barcelos (2013):

$$E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \simeq \frac{1}{\sigma_\phi^2(G-2)} \times \mathbf{R}_{\phi u}(n) \mathbf{v}(0) \mathbf{v}^T(n) \quad (3.69)$$

Sob a luz da Hipótese 2, o sexto e o trigésimo termos da Equação (3.65) são nulos.

$$E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \mathbf{v}^T(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \simeq E \{ \mathbf{r}^T(n-1) \} \times E \left\{ \frac{\mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \mathbf{v}^T(n) = \mathbf{0}_{N \times N} \quad (3.70)$$

Para determinar os demais valores esperados será considerado, de forma similar a Almeida et al. (2005) e Barcelos (2013), que:

$$E \left\{ \frac{1}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \simeq \frac{1}{\sigma_\phi^4 G(G+2)} \quad (3.71)$$

Para a calcular o oitavo valor esperado, usa-se a fatoração dos momentos gaussianos apresentada por Haykin (2002) juntamente com a Equação (3.71):

$$E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \mathbf{v}(n)\mathbf{v}^T(n) \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} \simeq \frac{1}{\sigma_\phi^4 G(G+2)} \times [\mathbf{R}_\phi \mathbf{v}^T(n)\mathbf{v}(n) + 2\mathbf{R}_\phi \mathbf{v}(n)\mathbf{v}^T(n)\mathbf{R}_\phi] \quad (3.72)$$

A expressão acima calculada leva a uma degradação do valor previsto do erro em regime (ALMEIDA, 2004). A opção por determinar o resultado deste valor esperado utilizando o procedimento descrito em (HAYKIN, 2002) se deu pela tratabilidade matemática, uma vez que não foi possível utilizar um procedimento similar ao feito em (ALMEIDA, 2004; ALMEIDA et al., 2005; BARCELOS, 2013) devido ao cenário de ordem insuficiente.

A fim de determinar o valor esperado dos nono e do décimo quarto termos da Equação (3.65), proceder-se-á da seguinte maneira:

$$\begin{aligned} [\boldsymbol{\phi}(n)\beta(n)\mathbf{v}^T(n)\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)]_{ij} &= \beta(n) \\ &\times \sum_{k=0}^{N-1} \phi(n-i+1)\phi(n-j+1)\phi(n-k)v_{k+1}(n) \end{aligned} \quad (3.73)$$

Tomando o valor esperado e utilizando a Hipótese 6, tem-se:

$$\begin{aligned} E\{[\boldsymbol{\phi}(n)\beta(n)\mathbf{v}^T(n)\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)]_{ij}\} &= \beta(n) \\ &\times \sum_{k=0}^{N-1} E\{\phi(n-i+1)\phi(n-j+1)\phi(n-k)\}v_{k+1}(n) \end{aligned} \quad (3.74)$$

Como o vetor  $\boldsymbol{\phi}(n)$  é um vetor de variáveis conjuntamente gaussianas e de média nula, o nono e o décimo quarto valores esperados são nulos, pois tal termo envolve um termo de ordem ímpar destas variáveis.

Considerando a Hipótese 2, o décimo e o vigésimo termos da Equação (3.65) são igualmente nulos.

$$\begin{aligned} E \left\{ \frac{\boldsymbol{\phi}(n)r_a(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \mathbf{v}^T(n) \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} &= E\{r_a(n)\} \\ \times E \left\{ \frac{\boldsymbol{\phi}(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \mathbf{v}^T(n) \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)}{\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)} \right\} &= \mathbf{0}_{N \times N} \end{aligned} \quad (3.75)$$

Utilizando a fatoração dos momentos gaussianos derivada no Anexo C em Barcelos (2013) dada por:

$$\begin{aligned} E\{\mathbf{y}_1\mathbf{y}_2^T \mathbf{w}_1 \mathbf{w}_2^T \mathbf{y}_3\mathbf{y}_4^T\} &= E\{\mathbf{y}_1\mathbf{y}_2^T\} \mathbf{w}_1 \mathbf{w}_2^T E\{\mathbf{y}_3\mathbf{y}_4^T\} \\ &+ E\{\mathbf{y}_1\mathbf{y}_3^T\} \mathbf{w}_2 \mathbf{w}_1^T E\{\mathbf{y}_2\mathbf{y}_4^T\} \\ &+ E\{\mathbf{y}_1\mathbf{y}_4^T\} \mathbf{w}_1^T E\{\mathbf{y}_2\mathbf{y}_3^T\} \mathbf{w}_2 \end{aligned} \quad (3.76)$$

e considerando  $\mathbf{v}(n)$  constante devido ao valor esperado condicionado neste vetor, define-se:  $\mathbf{y}_1 = \boldsymbol{\phi}(n)$ ,  $\mathbf{y}_2^T = \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n)$ ,  $\mathbf{w}_1 = \mathbf{v}(0)$ ,  $\mathbf{w}_2^T = \mathbf{v}^T(n)$  e  $\mathbf{y}_3^T = \mathbf{y}_4^T = \boldsymbol{\phi}^T(n)$ . Logo, o décimo primeiro e o vigésimo sexto valores esperados se tornam:

$$\begin{aligned}
& E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0) \mathbf{v}^T(n) \boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \\
& \simeq \frac{1}{\sigma_\phi^4 G(G+2)} [\mathbf{R}_{\phi u}(n) \mathbf{v}(0) E\{\mathbf{v}^T(n)\} \mathbf{R}_\phi + \\
& + \mathbf{R}_\phi \mathbf{v}(n) \mathbf{v}(0) \mathbf{R}_{\phi u}^T(n) + \mathbf{v}(0) \mathbf{R}_{\phi u}^T(n) \mathbf{v}(n) \mathbf{R}_\phi] \quad (3.77)
\end{aligned}$$

Sob a luz da Hipótese 2, o décimo segundo e o trigésimo segundo termos da Equação (3.65) são nulos.

$$\begin{aligned}
& E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \mathbf{v}^T(n) \boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = E\{\mathbf{r}^T(n-1)\} \\
& \times E \left\{ \frac{\mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}(n) \mathbf{v}^T(n) \boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = \mathbf{0}_{N \times N} \quad (3.78)
\end{aligned}$$

O fato de  $\beta(n)$  ser uma grandeza determinística, o décimo quinto valor esperado da Equação (3.65) é dado por:

$$E \left\{ \frac{\boldsymbol{\phi}(n) \beta(n) \boldsymbol{\phi}^T(n) \beta(n)}{\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)} \right\} \simeq \frac{\beta^2(n) \mathbf{R}_\phi}{\sigma_\phi^4 G(G+2)} \quad (3.79)$$

Novamente, devido às características estatísticas de  $r(n)$  explicitadas na Hipótese 2, o décimo sexto e o vigésimo primeiro termos são nulos.

$$E \left\{ \frac{\boldsymbol{\phi}(n) r_a(n) \boldsymbol{\phi}^T(n) \beta(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = E\{r_a(n)\} E \left\{ \frac{\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \beta(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = \mathbf{0}_{N \times N} \quad (3.80)$$

Para determinação do décimo sétimo termo e do vigésimo sétimo

termo da Equação (3.65), considerar-se-á que  $\boldsymbol{\phi}(n)$  e  $\mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)$  são conjuntamente gaussianos (BARCELOS, 2013). Assim, por se tratar de um momento de terceira ordem de uma variável aleatória gaussiana de média zero, estes termos são igualmente nulos.

De forma similar ao décimo sexto termo, o décimo oitavo e o trigésimo terceiro termos são nulos.

$$E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \boldsymbol{\phi}^T(n) \beta(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = E \{ \mathbf{r}^T(n-1) \}$$

$$\times E \left\{ \frac{\mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) \beta(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = \mathbf{0}_{N \times N} \quad (3.81)$$

Para determinar o valor esperado associado ao vigésimo segundo termo da Equação (3.65) é necessário utilizar a fatoração dos momentos gaussianos apresentada no Anexo B em (BARCELOS, 2013) que consiste em:

$$E \{ y_1 y_2 y_3^T y_4 \} = E \{ y_1 y_2 \} E \{ y_3^T y_4 \} + E \{ y_2 y_4 \} E \{ y_1 y_3^T \}$$

$$+ E \{ y_1 y_4 \} E \{ y_2 y_3^T \} \quad (3.82)$$

Escolhendo  $y_1 = y_4 = r_a(n)$  e  $y_2 = y_3 = \boldsymbol{\phi}(n)$ , utilizando a Equação (3.21), chega-se a:

$$E \left\{ \frac{\boldsymbol{\phi}(n) \boldsymbol{\phi}^T(n) r_a^2(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \simeq \frac{1}{\sigma_\phi^4 G(G+2)} (1 + \mu^2 \mathbf{a}_w^T \mathbf{a}_w) \sigma_r^2 \mathbf{R}_\phi \quad (3.83)$$

Considerando novamente as estatísticas de  $r(n)$  o vigésimo terceiro e o vigésimo oitavo termo são nulos.

$$\begin{aligned}
E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)r_a(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} &= E\{r_a(n)\} \\
\times E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{U}^T(n)\mathbf{v}(0)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} &= \mathbf{0}_{N \times N} \quad (3.84)
\end{aligned}$$

Para calcular o vigésimo quarto e o trigésimo quarto valor esperado da Equação (3.65), utilizar-se-á a definição de  $r_a(n)$  dada pela Equação (3.21). De forma similar ao feito em (BARCELOS, 2013):

$$\begin{aligned}
E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)r_a(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} &= E\{r(n)\} \\
\times E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} & \\
-\mu E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\mathbf{a}_w^T\mathbf{r}(n-1)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} & \quad (3.85)
\end{aligned}$$

O primeiro termo da equação acima é nulo. Entretanto, o segundo não o é, uma vez que envolve um momento de segunda ordem de  $r(n)$ . Assim, é possível utilizar a fatoração dos momentos gaussianos dada pela Equação (3.82). Dessa forma, utilizando  $\mathbf{y}_1 = \mathbf{a}_w^T\mathbf{r}(n-1)$ ,  $\mathbf{y}_4 = \mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)$  e  $\mathbf{y}_2 = \mathbf{y}_3 = \boldsymbol{\phi}(n)$ , chega-se a:

$$\begin{aligned}
E \left\{ \frac{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)r_a(n)\mathbf{a}_w^T\mathbf{Q}(-n+1)\mathbf{r}(n-1)}{[\boldsymbol{\phi}^T(n)\boldsymbol{\phi}(n)]^2} \right\} & \\
\simeq \frac{1}{\sigma_\phi^4 G(G+2)} \mathbf{R}_\phi \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{a}_w \sigma_r^2 & \quad (3.86)
\end{aligned}$$

Utilizando a fatoração dos momentos gaussianos apresentada

dada pela Equação (3.82) e definindo  $y_1 = y_4 = \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)$ ,  $\mathbf{y}_2 = \boldsymbol{\phi}(n)$  e  $\mathbf{y}_3^T = \boldsymbol{\phi}^T(n)$ , o vigésimo nono termo da Equação (3.65) se torna:

$$E \left\{ \frac{[\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{U}^T(n) \mathbf{v}(0)]^2}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \simeq \frac{1}{\sigma_\phi^4 G(G+2)} \\ \times [2\mathbf{R}_{\phi u}(n) \mathbf{v}(0) \mathbf{v}^T(0) \mathbf{R}_{\phi u}^T(n) + \mathbf{R}_\phi \mathbf{v}^T(0) \mathbf{R}_{\bar{u}}(n) \mathbf{v}(0)] \quad (3.87)$$

O trigésimo e o trigésimo quinto termo da Equação (3.65) são nulos devidos as estatísticas de  $r(n)$ .

$$E \left\{ \frac{\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1) \mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \\ = E \{ \mathbf{r}^T(n-1) \} E \left\{ \frac{\mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}(n) \mathbf{v}^T(0) \mathbf{U}(n) \mathbf{Q}^T(-n+1) \mathbf{a}_w \boldsymbol{\phi}^T(n)}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} \\ = \mathbf{0}_{N \times N} \quad (3.88)$$

Finalmente, o trigésimo sexto termo da Equação (3.65), sob a luz da Hipótese 2, se torna:

$$E \left\{ \frac{[\boldsymbol{\phi}(n) \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{r}(n-1)]^2}{[\boldsymbol{\phi}^T(n) \boldsymbol{\phi}(n)]^2} \right\} = \frac{1}{\sigma_\phi^4 G(G+2)} \mathbf{R}_\phi \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{a}_w \sigma_r^2 \quad (3.89)$$

Assim, agrupando todos os termos e tomando o valor esperando visando retirar o condicionamento em  $\mathbf{v}(n)$ , chega-se a uma equação



recursiva para  $\mathbf{K}(n)$ :

$$\begin{aligned}
\mathbf{K}(n+1) &= \mathbf{K}(n) - \\
&- \frac{\mu}{\sigma_\phi^2(G-2)} [\mathbf{R}_\phi \mathbf{K}(n) + \mathbf{K}(n) \mathbf{R}_\phi] \\
&+ \frac{\mu^2(1-\mu)^2}{\sigma_\phi^4 G(G+2)} \beta^2(n) \mathbf{R}_\phi + \frac{\mu^2}{\sigma_\phi^4 G(G+2)} (1 + \mu^2 \mathbf{a}_w^T \mathbf{a}_w) \sigma_r^2 \\
&+ \frac{\mu^2}{\sigma_\phi^4 G(G+2)} [\mathbf{R}_\phi \text{Tr} [\mathbf{R}_\phi \mathbf{K}(n)] + 2\mathbf{R}_\phi \mathbf{K}(n) \mathbf{R}_\phi] \\
&- \frac{\mu}{\sigma_\phi^2(G-2)} [\mathbf{R}_{\phi u}(n) \mathbf{v}(0) E\{\mathbf{v}^T(n)\} + E\{\mathbf{v}(n)\} \mathbf{v}^T(0) \mathbf{R}_{\phi u}^T(n)] \\
&+ \frac{2\mu^2}{\sigma_\phi^4 G(G+2)} [\mathbf{R}_{\phi u}(n) \mathbf{v}^T(0) E\{\mathbf{v}(n)\} \mathbf{R}_\phi + \mathbf{R}_\phi E\{\mathbf{v}(n)\} \mathbf{v}^T(0) \mathbf{R}_{\phi u}^T(n)] \\
&+ \frac{2\mu^2}{\sigma_\phi^4 G(G+2)} [\mathbf{v}^T(0) \mathbf{R}_{\phi u}^T(n) E\{\mathbf{v}(n)\} \mathbf{R}_\phi + \mathbf{R}_{\phi u}(n) \mathbf{v}(0) \mathbf{v}^T(0) \mathbf{R}_{\phi u}^T(n)] \\
&+ \frac{\mu^2}{\sigma_\phi^4 G(G+2)} \mathbf{R}_\phi \mathbf{v}^T(0) \mathbf{R}_{\tilde{u}}(n) \mathbf{v}(0) \\
&- \frac{\mu^4}{\sigma_\phi^4 G(G+2)} \mathbf{R}_\phi \mathbf{a}_w^T \mathbf{Q}(-n+1) \mathbf{a}_w \sigma_r^2 \tag{3.90}
\end{aligned}$$

De modo similar a Equação (3.64), ao aplicar-se a condição de passo de adaptação unitário e  $H \leq P$ , essa Equação recai na Equação (5.72) descrita por Barcelos (2013).

Por fim, as equações (3.40), (3.64) e (3.90) constituem um conjunto de equações que descrevem o comportamento médio do algoritmo PAP para passo não unitário e ordem  $H > P$ , dentro do cenário de identificação de sistemas, em que o sinal desejado  $d(n)$  é acrescido de um ruído branco  $r(n)$ .



## 4 RESULTADOS

Neste capítulo serão apresentados os resultados de simulações computacionais que visam comparar o desempenho do algoritmo PAP com as previsões fornecidas pelo modelo proposto neste trabalho e pelos modelos existentes na literatura. Para tanto, as seguintes condições de simulação são as mesmas para todos os exemplos usados neste capítulo:

- São eliminadas as primeiras 1000 amostras do sinal  $u(n)$  para evitar o transitório de inicialização do processo AR;
- A matriz  $\mathbf{U}(n)$  é totalmente preenchida antes da adaptação do algoritmo;
- A potência do ruído de excitação do processo AR,  $\sigma_z^2$ , é calculada para que a potência do sinal de entrada  $u(n)$  seja unitária;
- O conjunto de coeficientes ótimos  $\mathbf{w}_o$  consiste em uma janela de Hann normalizada pelo critério  $l_2$ ;
- O vetor de inicialização do algoritmo adaptativo  $\mathbf{w}(0)$  é sempre definido no vetor nulo  $\mathbf{0}_{N \times 1}$ ;
- A potência do ruído  $r(n)$  é  $\sigma_r^2 = 10^{-6}$ ;
- As simulações são feitas utilizando um conjunto de 600 realizações.

#### 4.1 AVALIAÇÃO DO MODELO PROPOSTO

Esta seção apresenta simulações que comparam as previsões do modelo proposto e dos modelos presentes na literatura com o comportamento estocástico do algoritmo PAP. Para todos os exemplos, o vetor  $\mathbf{a}_w$  dado pela Equação (3.28) é calculado conhecendo a função de autocorrelação de  $u(n)$ . Para os modelos de ordem suficiente, é substituído  $\mathbf{a}$  por  $\mathbf{a}_w$ .

- **Exemplo 1** (Ordem Suficiente)

A Figura 12 mostra o resultado de uma simulação computacional que comparara a previsões do modelo proposto neste trabalho e do modelo derivado por Almeida, Bermudez e Bershad (2009). Os parâmetros de simulação utilizados foram: sinal de entrada foi um processo AR da forma  $u(n) = 1.96u(n-1) - 0.9893u(n-2) + z(n)$ , com  $\sigma_z^2 = 0.0575$ , passo de adaptação  $\mu = 0.8$ , ordem do algoritmo PAP  $P = 2$ , planta com  $N = 64$  coeficientes e  $\lambda = 2.3192$ .

Nessa simulação, foi observado o comportamento do coeficiente de número 50. Nota-se ainda que, conforme esperado, o modelo derivado por Almeida, Bermudez e Bershad (2009) prevê de forma insatisfatória o comportamento médio deste coeficiente nas primeiras iterações devido ao fato do mesmo não considerar o efeito de inicialização do algoritmo PAP.

A Figura 13 mostra a comparação das previsões feitas do EQM em excesso pelo modelo proposto e pelo modelo derivado em Almeida, Bermudez e Bershad (2009) com relação ao algoritmo PAP. O conjunto

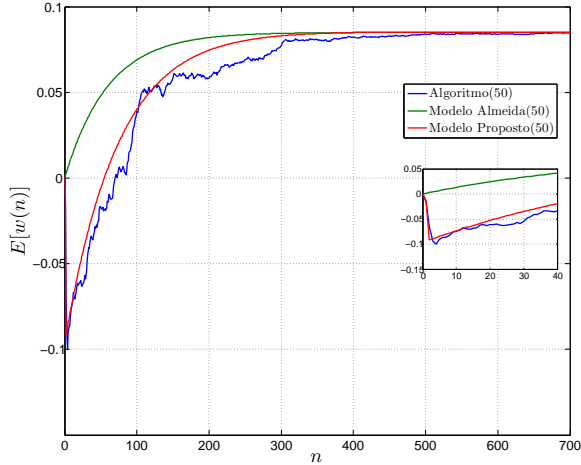


Figura 12 – Comparação entre o modelo proposto e o modelo derivado por Almeida, Bermudez e Bershad (2009) da previsão do comportamento médio dos coeficientes

de parâmetros utilizados é o mesmo do Exemplo 1. Essa escolha se

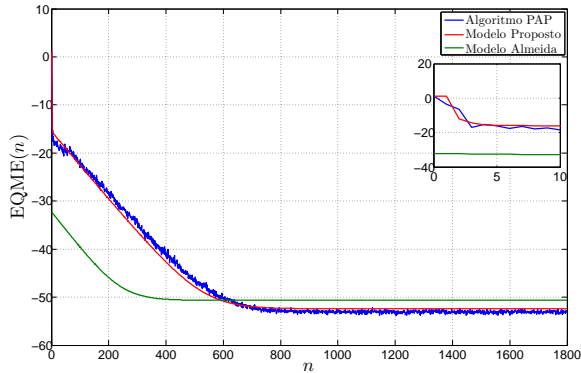


Figura 13 – Comparação do EQM em excesso entre o modelo proposto e o modelo derivado por Almeida, Bermudez e Bershad (2009)

deu pelo fato dos mesmos serem utilizados para ilustrar o problema de inicialização em (BARCELOS, 2013). Dessa forma, o modelo proposto

neste trabalho se comporta de forma satisfatória para casos em que o algoritmo PAP opera com ordem suficiente. Nota-se também uma discrepância de aproximadamente 30dB no modelo proposto por Almeida, Bermudez e Bershad (2009) e o EQM em excesso do algoritmo PAP na primeira iteração causada pelo efeito de inicialização do algoritmo.

- **Exemplo 2** (Ordem Insuficiente)

Esta simulação foi feita no intuito de comparar a previsão de comportamento feita pelo modelo proposto por este trabalho e o modelo derivado em Almeida, Costa e Bermudez (2010). Os parâmetros utilizados para esta simulação foram: sinal de entrada AR da forma  $u(n) = 0.2u(n-1) + 0.63u(n-2) + 0.036u(n-3) - 0.0324u(n-4) + z(n)$ , com  $\sigma_z^2 = 0.4270$ ,  $N = 64$ ,  $P = 1$ ,  $\mu = 0.2$  e  $\lambda = 9.9463$ . O coeficiente observado foi o de número 2, conforme mostra a Figura 14.

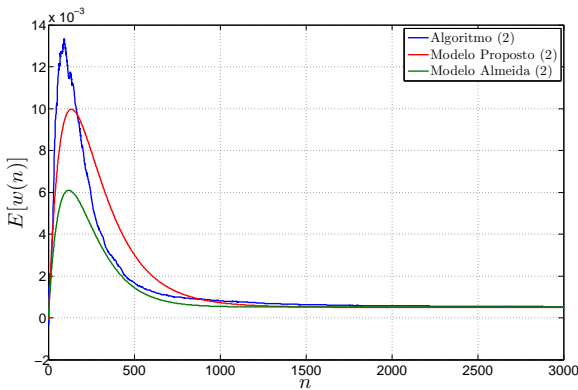


Figura 14 – Comparação do comportamento médio dos coeficientes entre o modelo proposto e o modelo derivado por Almeida, Costa e Bermudez (2010)

Nota-se no período transitório uma discrepância entre o modelo proposto e o comportamento real do segundo coeficiente. Isso ocorre porque a Teoria da Independência é violada, já que o vetor  $\phi(n)$ , no cenário onde  $P < H$ , é correlacionado com o vetor  $\phi(n-1)$ . Isso se agrava ainda mais conforme a correlação temporal do sinal  $u(n)$  sobe e a diferença  $H - P$  aumenta.

De forma a ilustrar o acima exposto, foi feita uma nova simulação onde o vetor  $\mathbf{u}(n)$  deixou de ser constituído por uma linha de atraso e passou a ser gerado a cada iteração, de acordo com os parâmetros utilizados no Exemplo 2. Dessa forma, todos os vetores gerados são independentes e identicamente distribuídos. Como consequência, a cada iteração é gerado um vetor  $\phi(n)$  que é independente dos vetores  $\phi(n-1)$  e  $\phi(n+1)$ , para todo  $n$ . O resultado dessa nova simulação é mostrado na Figura 15.

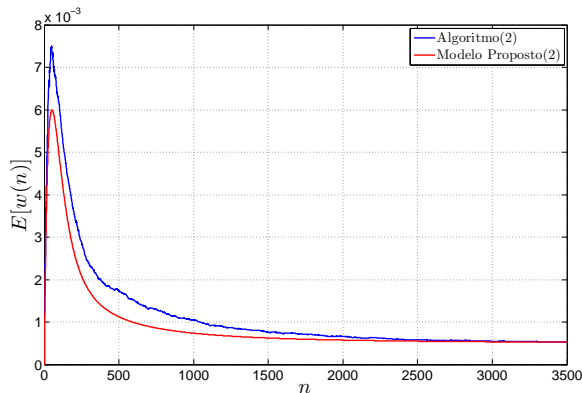


Figura 15 – Comparação do modelo proposto e o algoritmo PAP

Nota-se que a previsão do modelo proposto durante o transitório obtém uma melhora quando comparado a situação exposta pela Figura 14. Como as estatísticas de  $\mathbf{v}(n)$  são desconhecidas *a priori*, não é possível obter um modelo em forma fechada sem utilizar a Teoria da Independência.

Na Figura 16 é mostrado o EQM em excesso para ambos modelos. De forma similar ao exposto pela Figura 14, a previsão de ambos os modelos durante o período transitório possui uma discrepância significativa com relação ao EQM em excesso do algoritmo PAP, embora essa discrepância no modelo proposto por este trabalho apresenta-se de forma conservativa. A discrepância descrita é justificada pela violação da Teoria da Independência quando o algoritmo opera em um cenário de ordem deficiente.

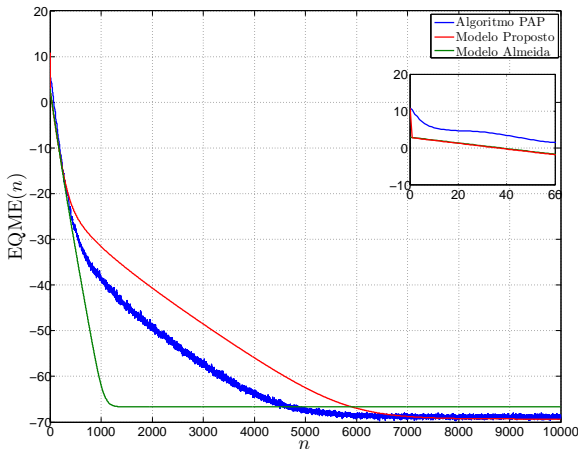


Figura 16 – Comparação do EQM em excesso entre o modelo proposto e o modelo derivado por Almeida, Costa e Bermudez (2010)



Para ilustrar que esse mesmo problema acontece na avaliação do EQM, uma simulação foi feita utilizando os mesmos parâmetros, mas gerando vetores do sinal de entrada  $\mathbf{u}(n)$  independentes e identicamente distribuídos. O resultado é mostrado na Figura 17. Nota-se, de forma similar a Figura 15, uma melhor previsão do modelo aqui proposto com relação ao algoritmo PAP durante o regime transitório, evidenciando que tal problema também ocorre quando uma forma fechada para a matriz  $\mathbf{K}(n)$  é obtida, já que para determinar o resultado de alguns valores esperados é necessário utilizar a Teoria da Independência.

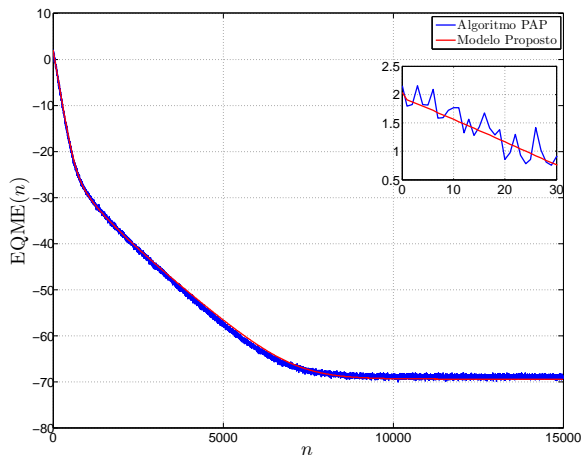


Figura 17 – Comparação do modelo proposto e o algoritmo PAP com vetores IID

Dessa forma, fica exposto que a violação da Teoria da Independência, quando o vetor de sinal de entrada  $\mathbf{u}(n)$  é constituído por uma linha de atraso, promove uma discrepância significativa entre o modelo proposto por este trabalho e o comportamento do EQM em excesso do algoritmo PAP durante o período transitório.

• **Exemplo 3** (Ordem Insuficiente)

A fim de avaliar a previsão do modelo com relação ao algoritmo PAP em um cenário em que a diferença  $H - P$  é pequena, uma nova simulação foi feita utilizando o processo AR(4) do Exemplo 2 utilizando os seguintes parâmetros:  $P = 3$ ,  $\mu = 0.6$ ,  $N = 32$  e  $\lambda = 2.7765$ . A Figura 18 mostra a avaliação do coeficiente de número 1 do modelo proposto, do modelo feito em (ALMEIDA; COSTA; BERMUDEZ, 2010) e o algoritmo PAP.

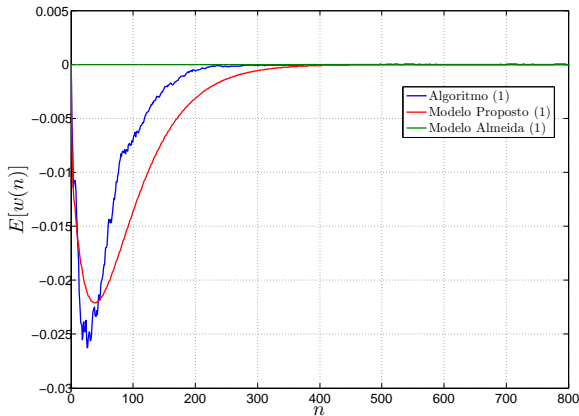


Figura 18 – Comparação do modelo proposto e o algoritmo PAP

Nota-se no comportamento do coeficiente de número 1 uma boa acurácia do modelo com relação ao algoritmo PAP. Verifica-se também uma robustez do modelo proposto à variação de parâmetros.

A Figura 19 mostra a acurácia do modelo proposto para o EQM em excesso, juntamente com o modelo proposto em (ALMEIDA; COSTA; BERMUDEZ, 2010).

Conforme esperado, o modelo proposto se comporta de maneira melhor do que o modelo proposto em (ALMEIDA; COSTA; BERMUDEZ,

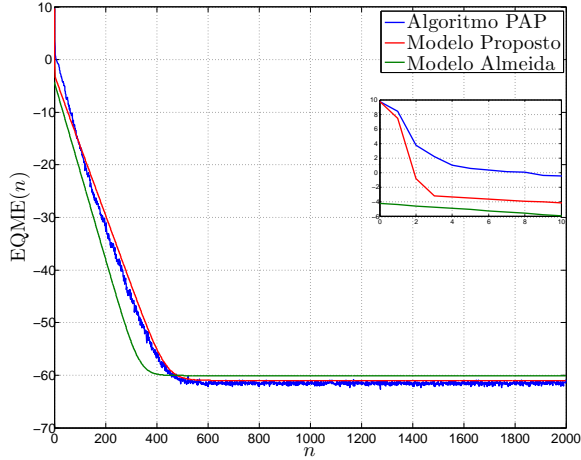


Figura 19 – Comparação do EMQE do modelo proposto e o algoritmo PAP

2010), tanto nas primeiras iterações, de aproximadamente 14dB na primeira iteração, quanto no erro em regime permanente.

- **Exemplo 4** (Ordem Insuficiente)

Com o intuito de explorar uma diferença de  $H - P$  elevada, outra simulação computacional foi feita utilizando os seguintes parâmetros: processo AR(10) com coeficientes  $a_i = [-0.9, 0.7, -0.6, 0.5, -0.45, 0.35, -0.3, 0.25, -0.2, 0.1]$ ,  $\sigma_z^2 = 0.5388$ ,  $\lambda = 2.4054$ ,  $P = 2$ ,  $\mu = 0.8$  e  $N = 32$ . Na Figura 20 foi observado o comportamento do coeficiente de número 25.

Verifica-se que o modelo consegue prever de forma satisfatória o comportamento do algoritmo PAP. Nota-se também que o modelo proposto em (ALMEIDA; COSTA; BERMUDEZ, 2010) tem uma leve discrepância com relação ao comportamento do algoritmo.

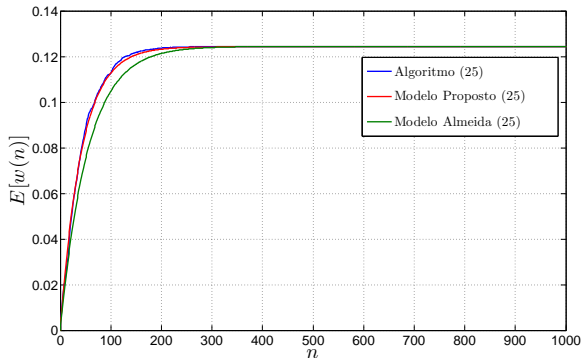


Figura 20 – Comparação do modelo proposto e o algoritmo PAP

A Figura 21 mostra a avaliação do EQM em excesso para o modelo proposto e o modelo derivado em (ALMEIDA; COSTA; BERMUDEZ, 2010).

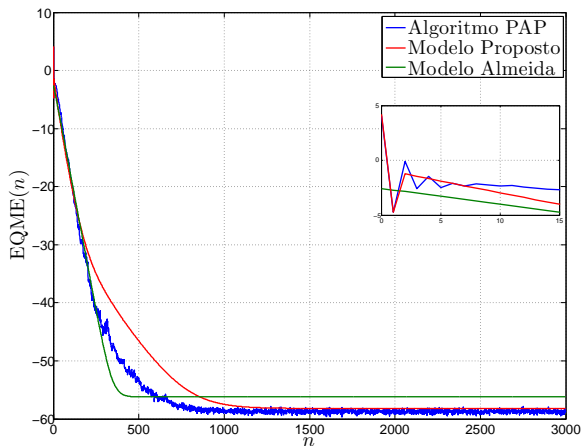


Figura 21 – Comparação do EQME modelo proposto e o algoritmo PAP

Nota-se que o modelo proposto prevê o comportamento do algoritmo PAP de forma satisfatória durante as primeiras iterações e o modelo derivado em (ALMEIDA; COSTA; BERMUDEZ, 2010) leva a um

erro de aproximadamente 4dB devido ao problema de inicialização. De forma similar ao exposto no exemplo 2, durante o regime transitório, existe uma discrepância da previsão do modelo devido a violação da Teoria da Independência.



## 5 CONCLUSÃO

Neste trabalho foi realizado o estudo do comportamento do algoritmo de Pseudo Projeções Afins, sob o cenário de identificação de sistemas e com sinal de entrada autoregressivo. Estudos anteriores desenvolveram modelos para a) ordem suficiente (ordem do algoritmo maior ou igual à ordem do processo AR do sinal de entrada), com passo qualquer mas sem considerar os efeitos da inicialização (ALMEIDA; BERMUDEZ; BERSHAD, 2009); b) o mesmo que (a) mas com ordem insuficiente (ALMEIDA; COSTA; BERMUDEZ, 2010); c) ordem suficiente, passo unitário e levando em conta os efeitos da inicialização (COSTA et al., 2012; BARCELOS, 2013). Este estudo considerou um passo de adaptação arbitrário, ordem insuficiente e os efeitos de inicialização. Este trabalho, portanto, deu origem a um modelo de comportamento que engloba todas as condições previstas nos modelos de Almeida, Bermudez e Bershada (2009), Almeida, Costa e Bermudez (2010) e Costa et al. (2012), Barcelos (2013).

A combinação de todas essas condições requerem algumas alterações significativas nas hipóteses empregadas na análise e, conseqüentemente, no equacionamento estatístico. Essas alterações incluíram uma nova hipótese sobre o comportamento do vetor de erros  $\mathbf{e}_a(n-1)$ , que permitiu considerar entradas AR altamente correlacionadas. Outra alteração introduzida foi uma nova aproximação para a estimativa do estimador de mínimos quadrados dos parâmetros do modelo AR do

sinal de entrada.

Sob a situação do algoritmo PAP operar com ordem deficiente, uma forma fechada foi obtida para a aproximação de  $\hat{\mathbf{a}}(n)$  por um vetor constante, simplificando as parcelas relativa a este vetor no modelo proposto por Almeida, Costa e Bermudez (2010).

O modelo resultante é capaz de prever o comportamento do algoritmo PAP sob condições não previstas pelos modelos existentes, sem perder acurácia nos casos já considerados pelos outros modelos.

## 5.1 SUGESTÃO DE TRABALHOS FUTUROS

Alguns pontos são sugeridos como continuação do estudo feito por este trabalho:

- **Ambiente não estacionário:** em geral, os filtros adaptativos possuem a capacidade de acompanhar variações do ponto ótimo da superfície de desempenho. Dessa forma, faz-se necessário a inclusão desse efeito no modelo estocástico obtido por este trabalho para que este seja capaz de prever o comportamento do algoritmo neste cenário (ALMEIDA; BERMUDEZ; BERSHAD, 2009);
- **Comprimento deficiente:** sob a justificativa de complexidade computacional ou por falta de informação a respeito da resposta a estimar, pode ser necessário que o algoritmo adaptativo opere com um número menor de coeficientes do que o ponto ótimo da superfície de desempenho (ALMEIDA; COSTA; BERMUDEZ, 2009). Assim, é importante incluir também esse efeito na obtenção de



modelos de comportamento estocástico;

- **Problema de inicialização:** conforme justificado em (COSTA et al., 2012; BARCELOS, 2013), as situações em que é necessário a utilização do modelo que corrige o problema de inicialização ainda é uma informação em aberto. Assim, uma abordagem quantitativa é necessária para justificar a utilização de um modelo com maior complexidade.



## REFERÊNCIAS

ALMEIDA, S. de; BERMUDEZ, J.; BERSHAD, N. A stochastic model for a pseudo affine projection algorithm. **Signal Processing, IEEE Transactions on**, v. 57, n. 1, p. 107–118, Jan 2009.

ALMEIDA, S. de et al. A statistical analysis of the affine projection algorithm for unity step size and autoregressive inputs. **Circuits and Systems I: Regular Papers, IEEE Transactions on**, v. 52, n. 7, p. 1394–1405, July 2005.

ALMEIDA, S. de; COSTA, M.; BERMUDEZ, J. A stochastic model for the deficient length pseudo affine projection adaptive algorithm. In: **Signal Processing Conference, 2009 17th European**. [S.l.: s.n.], 2009. p. 1715–1719.

ALMEIDA, S. de; COSTA, M.; BERMUDEZ, J. A theoretical model for the deficient order pseudo-affine projection algorithm. In: **Signal Processing Conference, 2010 18th European**. [S.l.: s.n.], 2010. p. 1718–1722.

ALMEIDA, S. J. M. de. **Análise estatística do comportamento de uma classe de algoritmos de projeções afins**. Tese (Doutorado) — Universidade Federal de Santa Catarina, Florianópolis, Fevereiro 2004.

BARCELOS, R. B. **O Efeito de inicizliação no Pseudo Algoritmo de Projeções Afins: Estudo e Análise Estocástica**. Dissertação (Mestrado) — Programa de Pós Graduação em Engenharia Elétrica, UFSC, Florianópolis, SC, Maio 2013.

BERSHAD, N. J.; LINEBARGER, D.; MCLAUGHLIN, S. A stochastic analysis of the affine projection algorithm for gaussian autoregressive inputs. **Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing**, p. 3837–3840, 2001.

BOUTEILLE, F.; SCARLAT, P.; CORAZZA, M. Pseudo affine projection algorithm: new solution for adaptive identification. In: **Eurospeech**. Budapest: [s.n.], 1999. p. 427–430.

COSTA, M. et al. New insights into the weight behaviour of the affine projection algorithm. In: **Signal Processing Conference**

- (**EUSIPCO**), **2012 Proceedings of the 20th European**. [S.l.: s.n.], 2012. p. 2610–2614.
- GAY, S.; TAVATHIA, S. The fast affine projection algorithm. In: **Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on**. [S.l.: s.n.], 1995. v. 5, p. 3023–3026.
- HAYKIN, S. **Adaptive Filter Theory**. 4. ed. Upper Saddle River, NJ, USA: Englewoods Cliffs: Prentice-Hall, 2002.
- KAY, S. M. **Fundamentals of Statistical Signal Processing: Estimation Theory**. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
- MANOLAKIS, D. G.; INGLE, V. K.; KOGON, S. M. **Statistical and Adaptive Signal Processing: spectral estimation, signal modeling, adaptive filtering and array processing**. New York, NY, USA: Mc-Graw-Hill, 2000.
- OZEKI, K.; UMEDA, T. An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. **Electron. Commun. Jpn**, v. 67-A, n. 5, p. 19–27, 1984.
- PROAKIS, J. D.; MANOLAKIS, D. G. **Digital Signal Processing: principles, algorithms and applications**. Upper Saddle River, NJ, USA: Prentice Hall, 1996.
- RUPP, M. A family of adaptive filter algorithms with decorrelating properties. **IEEE. Trans. Signal Processing**, v. 46, n. 3, p. 771–775, 1998.
- SAYED, A. L. **Fundamentals of Adaptive Filters**. 1. ed. Hoboken, NJ, USA: Wiley, 2003.
- SHIN, H.-C.; SAYED, A. Mean-square performance of a family of affine projection algorithms. **Signal Processing, IEEE Transactions on**, v. 52, n. 1, p. 90–102, Jan 2004.
- WIDROW, B.; HOFF, M. E. Adaptive switching circuits. In: **IRE WESCON Conv. Rec.** New York, NY, USA: [s.n.], 1960. v. 4, p. 96–104.
- WIDROW, S.; STEARNS, S. D. **Adaptive Signal Processing**. Upper Saddle River, NJ, USA: Englewoods Cliffs: Prentice-Hall, 1985.