

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
DEPARTAMENTO DE COMUNICAÇÕES

**TÉCNICAS DE PROCESSAMENTO DE
SINAIS APLICADAS À TRANSMISSÃO DE
DADOS VIA REDE ELÉTRICA E AO
MONITORAMENTO DA QUALIDADE DE
ENERGIA**

Autor

Moisés Vidal Ribeiro

Orientador

Prof. Dr. João Marcos Travassos Romano

Banca Examinadora:

Prof. Dr. Carlos Augusto Duque (Faculdade de Engenharia/UFJF)

Prof. Dr. Jacques Szczupak (DEE/PUC-RJ)

Prof. Dr. Luís Geraldo Pedroso Meloni (FEEC/UNICAMP)

Prof. Dr. Sigmar Maurer Deckmann (FEEC/UNICAMP)

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Campinas, Abril de 2005

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

R354t

Ribeiro, Moisés Vidal

Técnicas de processamento de sinais aplicadas à transmissão de dados via rede elétrica e ao monitoramento da qualidade de energia / Moisés Vidal Ribeiro. --Campinas, SP: [s.n.], 2005.

Orientador: João Marcos Travassos Romano.
Tese (doutorado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Algoritmos difusos. 2. Redes neurais (Computação). 3. Sistemas de transmissão de dados. 4. Processamento de sinais – Técnicas digitais. 5. Compressão de dados (Telecomunicações). 6. Reconhecimento de padrões. 7. Detecção de sinais. 8. Sistemas de energia elétrica. I. Romano, João Marcos Travassos. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Titulo em Inglês: Signal processing techniques for power line communication and power quality applications

Palavras-chave em Inglês: Fuzzy algorithms, Neural networks, Data transmission systems, Signal processing – Digital techniques, Data compression (Telecommunication), Pattern recognition, Signal detection e Electric power systems

Área de concentração: Telecomunicações

Titulação: Doutor em Engenharia Elétrica

Banca examinadora: Carlos Augusto Duque, Jacques Szczupak, Luís Geraldo Pedroso Meloni e Sigmar Maurer Deckmann

Data da defesa: 04/04/2005

Resumo

A presente tese tem por objetivo propor e discutir o uso de algumas técnicas de processamento de sinais e de inteligência computacional para a melhoria da transmissão digital de dados via redes elétricas e da análise da qualidade da energia elétrica em sistemas de potência. No que tange à transmissão de dados via rede elétrica, novas técnicas são introduzidas para solucionar os problemas de cancelamento de ruídos impulsivos e equalização de canais de comunicação. Para a melhoria do monitoramento da qualidade da energia elétrica, propõem-se novas técnicas para a análise espectral das componentes fundamental e harmônicas, e para a detecção, a classificação e a compressão de distúrbios. As várias técnicas apresentadas no presente trabalho são fundamentadas no princípio de dividir e conquistar, largamente utilizado em diversas áreas do conhecimento. A aplicação adequada desse princípio através de técnicas de processamento de sinais e de inteligência computacional nos permitiram fornecer análises mais precisas dos problemas estudados e propor novas soluções para os mesmos. Os resultados numéricos obtidos nas simulações computacionais confirmam a relevância das técnicas propostas.

Abstract

This thesis is aimed at proposing and discussing the use of signal processing and computational intelligence techniques to improve digital communications through power line channels and a more precise power quality analysis of power systems. Regarding power line communication applications, advanced techniques for impulse noise mitigation and channel equalization are introduced. For power quality monitoring applications, novel techniques are proposed for spectral analysis of power line signals and for detection, classification and compression of disturbance events. The techniques proposed are developed on the light of the divider and conquer principle. The appropriate application of such principle, by means of signal processing and computational intelligence techniques, enable us to offer a more precise analysis of the problems investigated and novel solutions for them. By introducing a set of signal processing techniques along with some computational intelligence ones, this contribution succeeds in offering improvements for all the problems investigated. Numerical results obtained by computational simulations verify such improvement and confirm the relevance of the techniques proposed.

*À minha esposa Patrícia e
ao meu filhinho Gabriel*

Ao meu pai Isaías Ribeiro que me conduziu ao verdadeiro caminho.

in memoriam

O princípio da sabedoria é: Adquire a sabedoria; sim, com tudo o que possuis, adquire o entendimento. Estima-a, e ela te exaltarás; se a abraçares, ela te honrará; dará à tua cabeça um diadema de graça e uma coroa de glória te entregará.

Provérbios (Pv) 4:7-9, Bíblia Sagrada

Agradecimentos

A Deus por ter-me mostrado o quão sublime é ser alcançado pela Sua infinita misericórdia, amor e salvação através de Jesus Cristo, o nosso Redentor.

Ao professor João Marcos Travassos Romano pela orientação e apoio durante esta gratificante e prazerosa jornada.

A minha amada esposa Patrícia pelo carinho, atenção e companheirismo em todos os momentos. Sem ela esta pequena contribuição não seria possível.

A minha mãe que nunca mediu esforços para me ajudar e me apoiar, mesmo à distância. Seu exemplo me ofereceu a educação e os princípios que me fizeram chegar aqui.

Aos meus sogros pelo carinho, atenção e ajuda, enquanto eu me debruçava na tese.

Aos meus irmãos e familiares que sempre me apoiaram.

Aos professores Carlos Augusto Duque e Sigmar M. Deckmann pelas frutíferas discussões sobre qualidade de energia.

Ao prof. Fernando J. Von Zuben pelas riquíssimas discussões sobre inteligência computacional.

Ao prof. J. M. Mendel por ter-me encorajado e motivado a trabalhar com sistemas fuzzy do tipo 2.

Ao prof. Sanjit K. Mitra que me orientou durante os 6 meses que trabalhei na Universidade da Califórnia, Santa Barbara.

Aos amigos do Laboratório de Processamento de Sinais para Comunicações - DSPCom pelos 5 anos e 6 meses de convívio e aprendizado.

Aos alunos Gustavo Mendes e Dayyan Shayani que contribuíram com esta tese através de seus trabalhos de iniciação científica sob minha co-orientação.

Aos amigos Dom Fabione, Maurício, Japa, Edson, Jayme, Kenji, Andrézinho, Adriana, Massakit, Chicão, Cícero, Dona Benedita, Edvaldo, Clodoaldo, Jim, Elder, Jaqueline, Lívio, Relman, Gabriel, Rosário, Marcelo, Élcio e Mylene pela amizade e pelos bons momentos vividos durante este período.

Aos funcionários da FEEC, em especial à Lucia, Noêmia, Eloisa, Celi, Gislaine, Mazé, Seu Hélio, Washington, Juraci, Nelson e dona Nicinha, pela gentileza e apoio indispensável.

Ao povo brasileiro que financiou este trabalho, através da FAPESP e CAPES, na expectativa de que o mesmo contribua para o desenvolvimento científico e tecnológico do nosso país.

Sumário

Resumo e Abstract	iii
Dedicatória	vii
Agradecimentos	xiii
Lista de Figuras	xxiii
Lista de Tabelas	xxxix
1 Introdução	1
2 Sistemas Fuzzy do Tipo 2	5
2.1 Breve Introdução sobre Sistemas Fuzzy do Tipo 2	7

2.2	Algoritmos fuzzy do Tipo 2 Não Normalizados	23
2.2.1	Algoritmo Fuzzy Não Normalizado I	24
2.2.2	Algoritmo Fuzzy Não Normalizado II	32
2.3	Algoritmos fuzzy do Tipo 2 Normalizados	40
2.3.1	Algoritmo Fuzzy do Tipo 2 Adaptativo I	41
2.3.2	Algoritmo Fuzzy do Tipo 2 Adaptativo II	46
2.3.3	Estratégias para Atualização	53
2.3.3.1	Estratégia Baseada em Informações de 1ª Ordem: <i>Delta Rule Delta</i>	56
2.3.3.2	Estratégia Baseada em informações de 1ª Ordem: <i>Local Lipschitz Estimation</i>	58
2.4	Classificador Fuzzy do tipo 2	61
2.5	Sumário	68
3	Transmissão de Dados Via Rede Elétrica	69
3.1	Modelo dos Canais PLC <i>Outdoor</i>	75
3.2	Modelo do Ruído Aditivo	79
3.3	Técnica OFDM/DMT	83

<i>SUMÁRIO</i>	xix
3.4 Sumário	92
4 Redução da Presença de Ruídos Impulsivos	93
4.1 Inteligência Computacional	100
4.2 Resultados Numéricos	109
4.3 Sumário	118
5 Equalização de Canais PLC	121
5.1 Equalização de canais PLC	123
5.2 Relações entre o Equalizador Fuzzy do Tipo 2 e o Bayesiano	132
5.3 Resultados Numéricos	138
5.3.1 Estratégias de Atualização do Passo de Adaptação	138
5.3.2 Resultados com os Algoritmos Fuzzy do Tipo 2	143
5.4 Sumário	148
6 Monitoramento da QE	149
6.1 Motivação para o Monitoramento da QE	150
6.2 Solução para o Monitoramento da QE	154
6.3 Sumário	160

7	Análise Espectral dos Distúrbios	163
7.1	Domínio do Operador Delta	175
7.2	Filtro Notch Adaptativo	180
7.3	Transformada WDFT	181
7.4	Resultados Numéricos	185
7.5	Sumário	191
8	Detecção de Distúrbios	193
8.1	Técnica para a Extração de Parâmetros	198
8.2	Algoritmo de Detecção	201
8.3	Resultados Numéricos	203
8.4	Sumário	212
9	Classificação de Distúrbios	215
9.1	Classificador Proposto	219
9.2	Técnica para a Extração de Parâmetros	225
9.3	Classificadores ML e Fuzzy do Tipo 2	230
9.3.1	Classificador ML	231

<i>SUMÁRIO</i>	xxi
9.4 Resultados Numéricos	233
9.5 Sumário	238
10 Compressão de Distúrbios	241
10.1 EDCM	242
10.1.1 Estimação da Amplitude e da Fase	248
10.1.2 Estimação da Frequência	249
10.1.3 Resultados Numéricos	251
10.2 FHTCM	263
10.2.1 Análise das Componentes Harmônicas	267
10.2.2 Estratégia para a Compressão	270
10.2.3 Critério MDL	272
10.2.4 Resultados Numéricos	276
10.2.4.1 Distúrbio I	277
10.2.4.2 Distúrbio II	279
10.2.4.3 Distúrbio III	279
10.2.4.4 Distúrbio IV	279

10.2.4.5 Distúrbio V	280
10.3 Sumário	283
11 Conclusões e Perspectivas Futuras	285
12 Referências Bibliográficas	289
Referências	289
A Produção Científica	319
A.1 Artigos em Revistas	319
A.2 Artigos em Congressos	320
A.3 Premiações	324
A.4 Patentes	324

Lista de Figuras

2.1	Estrutura de um sistema fuzzy do tipo 1.	8
2.2	Função Gaussiana de pertinência de um sistema fuzzy do tipo 1. . . .	9
2.3	Funções de pertinência para a modelar a incerteza no valor da temperatura quando o sistema fuzzy é do tipo 1.	9
2.4	Representação de um conjunto fuzzy do tipo 2.	15
2.5	Função de pertinência secundária do conjunto fuzzy do tipo 2, mostrado na Fig. 2.4.	16
2.6	Estrutura de um sistema fuzzy do tipo 2.	16
2.7	Representação de um conjunto fuzzy do tipo 2 intervalado.	19
2.8	Função de pertinência secundária do conjunto fuzzy do tipo 2 intervalado, mostrado na Fig. 2.7.	19

2.9	Funções de pertinência usadas para modelar as incertezas relacionadas à medida temperatura, quando o sistema fuzzy aplicado é o do tipo 2.	20
3.1	Módulo da Resposta em Freqüência de três canais PLC de redes de baixa tensão.	78
3.2	Ruídos aditivos em canais PLC <i>outdoor</i>	80
3.3	Ruído aditivo de acordo com o modelo proposto por Zimmermann e Dostert.	82
3.4	Modelo banda base do sistema de transmissão multi-portadoras OFDM.	85
3.5	Modelo banda base do sistema de transmissão multi-portadoras DMT.	85
3.6	Canais paralelos obtidos com o sistema multi-portadora OFDM/DMT.	87
4.1	Modelo banda base do sistema de transmissão multi-portadoras OFDM modificado.	98
4.2	Modelo banda base do sistema de transmissão multi-portadoras DMT modificado.	99
4.3	Resposta em freqüência do canal PLC 1.	110
4.4	Resposta em freqüência do canal PLC 2.	111
4.5	(a) saída do canal PLC $\tilde{y}(n)$, (b) saída do canal PLC corrompida com ruído aditivo $y(n)$ e (c) saída do bloco CRI $\hat{\tilde{y}}(n)$, referentes ao canal PLC 1.	112

4.6	(a) saída do canal PLC $\tilde{y}(n)$, (b) saída do canal PLC corrompida com ruído aditivo $y(n)$ e (c) saída do bloco CRI $\hat{y}(n)$, referentes ao canal PLC 2.	112
4.7	Desempenho dos sistemas DMT e DMT modificado em função de L_w para $L_b = 16$, quando o canal PLC é o 1.	117
4.8	Desempenho dos sistemas DMT e DMT modificado em função de L_w para $L_b = 16$, quando o canal PLC é o 2.	118
4.9	Desempenho dos sistemas DMT e DMT modificado em função de L_b quando $L_w = 88$ para o canal PLC 1.	119
4.10	Desempenho dos sistemas DMT e DMT modificado em função de L_b quando $L_w = 104$ para o canal PLC 2.	119
5.1	Modelo banda base para sistemas de comunicação com equalizadores transversais.	124
5.2	Modelo banda base para sistemas de comunicação com equalizadores que possuem realimentação decidida.	124
5.3	Esquema de um equalizador transversal (TE).	126
5.4	Esquema do equalizador com realimentação decidida (DFE).	127
5.5	Resposta impulsiva do canal PLC utilizado nas simulações. O comprimento do canal foi truncado, resultando em $L_h = 50$	140
5.6	Exemplo do ruído aditivo em canais PLC quando $\sigma_x^2 = 0$ dB, $\sigma_{imp}^2 = 0$ dB, $\sigma_{bkg}^2 = -20$ dB e número de amostra igual a $N = 10^5$	141

5.7	Curva de convergência dos equalizadores fuzzy do tipo 1 com estrutura transversal.	142
5.8	Curva de convergência dos equalizadores fuzzy do tipo 1 com estrutura DFE.	142
5.9	Desempenho dos equalizadores fuzzy do tipo 1 com a estrutura do DFE quando existe erro de propagação no mesmo.	144
5.10	Desempenho dos equalizadores fuzzy do tipo 1 com a estrutura DFE quando não existe erro de propagação no mesmo.	144
5.11	Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2. . . .	146
5.12	Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2 <i>nonsingleton</i>	147
5.13	Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2 <i>singleton</i>	147
6.1	Estrutura em bloco do software PQView desenvolvido pela EPRI (EPRI, 2002).	153
6.2	Estrutura básica de um sistema completo de monitoramento da qualidade da energia elétrica.	155
6.3	Diagrama de blocos do MEDCC.	156
6.4	Diagrama de blocos da solução hierárquica baseada em dois níveis. . .	160
6.5	Diagrama de blocos do princípio dividir e conquistar.	160

7.1	Técnica NF-WDFT I.	167
7.2	Técnica NF-WDFT II.	167
7.3	Relação entre os valores da frequência do sinal $\{f(n)\}$ nos domínios da DFT e WDFT.	172
7.4	Relação entre os valores da fase do sinal $\{f(n)\}$ nos domínios da DFT e WDFT.	172
7.5	Relação entre os valores da amplitude de $\{f(n)\}$ nos domínios da DFT e WDFT.	173
7.6	Desempenho da rede MLP. (a) e (b) são os resultados de treinamento e teste referentes à A_0 . (c) e (d) são os resultados de treinamento e teste referentes à θ_0 . (e) e (f) são os resultados de treinamento e teste referentes à ω_0	176
7.7	Resposta em frequência do filtro Notch de 2^a ordem nos domínios z e do operador Δ . Os coeficientes do filtro foram quantizados com 12 bits.	179
7.8	Comportamento da WDFT para valores diferentes de $ \alpha $, quando $\varphi = \pi/8$ e $N = 16$	184
8.1	Técnica para a detecção de distúrbios.	198
8.2	Fluxograma do algoritmo desenvolvido para a detecção de ocorrência dos distúrbios e de seus instantes inicial e final correspondentes.	203

8.3	Transitório com duas variações de curta duração, $f_s = 192 \times 60$ amostras/s.	204
8.4	Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.3(b), para diferentes valores de N	205
8.5	Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.3(b), para diferentes valores de N	206
8.6	Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.	206
8.7	Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.6(b), para diferentes valores de N	207
8.8	Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.6(b), para diferentes valores de N	208
8.9	Sinal sem distúrbio, $f_s = 256 \times 60$ amostras/s.	208
8.10	Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.9(b), para diferentes valores de N	209
8.11	Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.9(b), para diferentes valores de N	210
8.12	Intervalo de detecção de distúrbios para o sinal com harmônicas.	212
8.13	Intervalo de detecção de distúrbios para o sinal com duplo <i>sag</i>	213
8.14	Sinal sem a ocorrência de distúrbios.	213

9.1	Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.	218
9.2	Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.	218
9.3	Sinal com sub-tensão e com transitórios de curta duração, $f_s = 256 \times$ 60 amostras/s.	218
9.4	Diagrama de bloco do classificador padrão de distúrbios.	220
9.5	Diagrama de bloco da técnica proposta para a classificação de múltiplos distúrbios em sistemas de potência.	221
9.6	Diagrama de bloco do classificador $\mathbf{j}(n)$, $j = f, h, u$	222
10.1	<i>Standard Data Compression Method</i> (SDCM).	243
10.2	<i>Enhanced Data Compression Method</i> (EDCM).	245
10.3	Cabeçalho do EDCM.	247
10.4	Diagrama de blocos do i -ésimo filtro Notch.	250
10.5	Convergência do parâmetro a_{01} do filtro notch adaptativo.	253
10.6	Convergência do parâmetro $g_c(n)$ do filtro de Kalman.	253
10.7	Convergência do parâmetro $g_s(n)$ do filtro de Kalman.	254
10.8	Distúrbio I: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$	255

10.9	Distúrbio I: comparação de desempenho entre a SDCM e a EDCM.	255
10.10	Distúrbio II: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$	257
10.11	Distúrbio II: comparação de desempenho entre a SDCM e a EDCM.	257
10.12	Distúrbio III: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$	258
10.13	Distúrbio III: comparação de desempenho entre a SDCM e a EDCM.	259
10.14	Banco de filtro de dois canais com filtros de análise e síntese quantizados.	261
10.15	Técnica NF-WDFT I.	267
10.16	Diagrama de blocos da FHTCM.	268
10.17	Sinal sintético I: análise de desempenho.	278
10.18	Sinal sintético II: análise de desempenho.	280
10.19	Sinal sintético III: análise de desempenho.	281
10.20	Sinal real I: análise de desempenho.	282
10.21	Sinal real II: análise de desempenho.	283

Lista de Tabelas

2.1	Custo computacional das técnicas DRD e LLE.	61
3.1	Parâmetros da PSD do ruído colorido de fundo.	80
4.1	Canal PL 1: Valores de SNR (dB), na entrada do bloco TEQ, para diferentes valores de PSD (dB/Hz) do ruído impulsivo.	115
4.2	Canal PL 2: Valores de SNR (dB), na entrada do bloco TEQ, para diferentes valores de PSD (dB/Hz) do ruído impulsivo.	115
4.3	Canal PL 1: Taxas de bit em Mbps, obtidas com as técnicas DMT e DMT modificada, para diferentes valores de PSD(dB/Hz) dos ruídos impulsivos.	116
4.4	Canal PL 2: Taxas de bit em Mbps, obtidas com as técnicas DMT e DMT modificada, para diferentes valores de PSD(dB/Hz) dos ruídos impulsivos.	116

7.1	Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 57,430$ Hz.	187
7.2	Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 60,110$ Hz.	187
7.3	Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 62,890$ Hz.	188
7.4	Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 57,00$ Hz.	189
7.5	Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 60,01$ Hz.	189
7.6	Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 62,50$ Hz.	190
7.7	DHT em (%) obtidas com as técnicas NF-WDFT, DFT-WDFT e FFT, a partir dos resultados das tabelas 7.4-7.6	190
7.8	Estimação da fase das componentes harmônicas com a técnica NF-WDFT.	191
8.1	Taxa em (%) de detecção, obtida com a técnica de detecção proposta, em função de diferentes conjuntos de parâmetros extraídos do sinal monitorado e de diferentes números de amostras.	211
9.1	Tipos de distúrbios relacionados com o sinal $\{f(n)\}$	223

9.2	Taxa em (%) de classificação, obtida com a técnica de classificação proposta, para distúrbios em $\mathbf{f}(n)$. Note que C é a abreviatura de ciclos da componente fundamental.	235
9.3	Taxa em (%) de classificação, obtida com a técnica de classificação proposta, para a detecção de distúrbios em $\mathbf{h}(n)$	236
9.4	Taxa em (%) de classificação, obtida com a técnica de classificação proposta, dos distúrbios em $\mathbf{u}(n)$	237
9.5	Taxa em (%) de classificação, obtida com a técnica de classificação proposta, dos múltiplos distúrbios em $\mathbf{x}(n)$ quando $N = 1024$	238
10.1	Distúrbio I: SNR e taxa de compressão (TC) para diferentes valores em (%) de T	256
10.2	Distúrbio II: SNR e taxa de compressão (TC) para diferentes valores em (%) de T	256
10.3	Distúrbio III: SNR e taxa de compressão (TC) para diferentes valores em (%) de T	259
10.4	Desempenho da EDCM e da SDCM em função dos valores do parâmetro $FHIR$	260
10.5	Influência da precisão finita nos resultados obtidos com a SDCM e a EDCM.	263
10.6	Influência da precisão finita nos resultados obtidos com a SDCM e a EDCM.	263

Abreviaturas

ADC:	<i>Analogic to Digital Converter</i> – Conversor Analógico Digital
AM:	<i>Amplitude Modulation</i>
ANF:	<i>Adaptive Notch Filter</i> – Filtro Notch Adaptativo
ASIC:	<i>Application-Specific Integrated Circuit</i>
AWGN:	<i>Additive White Gaussian Noise</i>
BER:	<i>Bit Error Rate</i> – Taxa de Bits Errados
BP:	<i>BackPropagation</i>
BPSK:	<i>Binary Pulse Shift Keying</i>
CA:	Corrente Alternada
CC:	Corrente Contínua
CDMA:	<i>Code Division Multiple Access</i>
CP:	<i>Cyclic Prefix</i> - Prefixo Cíclico
CRI:	Cancelamento de Ruídos Impulsivos
DAC:	<i>Digital Analogic Converter</i> – Conversor Digital Analógico
dB:	Decibel
DFE:	<i>Decision Feedback Equalizer</i> – Equalizador com Decisão Realimentada
DFT:	<i>Discrete Fourier Transform</i>

DHT:	Distorção Harmônica Total
DMT:	<i>Discrete Multitone Transceiver</i>
DRD:	<i>Delta Rule Delta</i>
DSL:	<i>Digital Subscriber Line</i>
DSOMCS:	<i>Diagonal Second-Order Mixed Cumulant Slice</i>
DSP:	<i>Digital Signal Processing</i>
EDCM:	<i>Enhanced Data Compression Method</i>
EM:	<i>Expectation Maximization</i>
EO2:	Estatísticas (ou Momentos) de Segunda Ordem
EOS:	Estatísticas (ou Momentos) de Ordem Superior
EQM:	Erro Quadrático Médio
FCC:	<i>Federal Communications Commission</i>
FFT:	<i>Fast Fourier Transform</i>
FHIR:	<i>Fundamental to Harmonic and Inter-Harmonic Ratio</i>
FHR:	<i>Fundamental to Harmonic Ratio</i>
FHTCM:	<i>Fundamental, Harmonic and Transient Separation Compression Method</i>
FIR:	<i>Finite Impulse Response</i> – Resposta Finita ao Impulso
FPGA:	<i>Flexible Programmable Gate Array</i>
GGD:	<i>Generalized Gaussian Distribution</i>
HDTV:	<i>High Definition Television</i>
ICC:	Interferência Co-Canal
ICI:	<i>Inter Carrier Interference</i>
IDFT:	<i>Inverse Discrete Fourier Transform</i>
IIR:	<i>Infinite Impulse Response</i> – Resposta infinita ao Impulso
IIS:	Interferência Inter-Simbólica
LLE:	<i>Local Lipschitz Estimation</i>

LMS:	<i>Least-Mean-Square</i>
LZW:	<i>Lempel-Ziv-Welch</i>
MAP:	<i>Maximum a posteriori</i>
MDL:	<i>Minimum Description Length</i>
MEDCC:	<i>Method for Estimation, Detection, Compression, and Classification</i>
MIMO:	<i>Multiple-Input Multiple-Output</i> – Várias Entradas e Várias Saídas
ML:	<i>Maximum Likelihood</i>
MLP:	<i>Multilayer Perceptron</i>
MLSE:	<i>Maximum Likelihood Sequence Estimation</i>
MMSE:	<i>Minimum Mean Squared Error</i> – Erro Quadrático Médio Mínimo
MSE:	<i>Mean Squared Error</i> – Erro Quadrático Médio
NDFT:	<i>Nonuniform Discrete Fourier Transform</i>
NF:	<i>Notch Filter</i> – Filtro Notch
NS:	<i>NonSingleton</i>
OFDM:	<i>Orthogonal Frequency Division Multiplexing</i>
PAM:	<i>Pulse Amplitude Modulation</i>
PLC:	<i>Power Line Communication</i>
PLL:	<i>Phase Locked Loop</i>
PSD:	<i>Power Spectral Density</i> – Densidade Espectral de Potência
QAM:	<i>Quadrature Amplitude Modulation</i> – Modulação de Amplitude em Quadratura
QE:	Qualidade de Energia
RBF:	<i>Radial Basis Function</i>
RLS:	<i>Recursive Least-Squares</i>
RN:	Rede Neural
S:	<i>Singleton</i>
SD:	<i>Steepest Descent</i>

SDCM:	<i>Standard Data Compression Method</i>
SER:	<i>Symbol Error Rate</i> – Taxa de Símbolos Errados
SIMO:	<i>Single-Input Multiple-Output</i> – Uma Entrada e Várias Saídas
SNR:	<i>Signal to Noise Ratio</i> – <i>Relação Sinal Ruído</i>
SVD:	<i>Singular Vector Decomposition</i>
TC:	Taxa de Compressão
TE:	<i>Transversal Equalizer</i> – <i>Equalizador Transvesal</i>
TEQ:	<i>Time Domain Equalizer</i>
TGTV:	<i>Teste Generalizado da Taxa de Verossimilhança</i>
TTV:	<i>Teste da Taxa de Verossimilhança</i>
TW:	Transformada Wavelet
TWD:	Transformada Wavelet Discreta
TWP:	Transformada Wavelet Packet
UTQ:	<i>Uniform Threshold Quantization</i>
WDFT:	<i>Warped Discrete Fourier Transform</i>

1

Introdução

A presente tese se debruça em duas frentes distintas de pesquisa: a transmissão digital de dados via redes elétricas e o monitoramento da qualidade de energia (QE). O aspecto comum no tratamento dessas duas áreas de pesquisa são as técnicas utilizadas nas soluções propostas, as quais são baseadas em técnicas de processamento de sinais e de inteligência computacional.

No que tange à transmissão de dados via rede elétrica, novas técnicas para o cancelamento de ruídos impulsivos e para a equalização de canais PLC (*Power line Communication*) foram pesquisadas e desenvolvidas. Na área de monitoramento

da QE foram investigadas novas técnicas para os problemas de análise espectral, detecção, classificação e compressão de distúrbios.

As técnicas propostas para os diferentes problemas abordados fundamentam-se no princípio dividir e conquistar, que sugere a decomposição de um problema analisado em vários sub-problemas. Como tais sub-problemas apresentam menor grau de complexidade, as soluções para os mesmos são mais simples. Os resultados numéricos, obtidos por meio de simulações computacionais, são animadores e apontam a adequação das técnicas desenvolvidas para os diversos problemas abordados, abrindo perspectivas futuras de trabalhos.

No intuito de melhor apresentar nossas contribuições para a solução dos vários problemas abordados, esta tese foi dividida em vários capítulos. No *Capítulo 2* é brevemente discutido o sistema fuzzy do tipo 2 e são introduzidos algoritmos baseados neste sistema fuzzy. No *Capítulo 3* discute-se o uso das redes elétricas para a transmissão digital de dados, baseada em sistemas multiportadoras, e apresenta-se o modelo de canais PLC *outdoor* e seu modelo de ruído aditivo correspondente. No *Capítulo 4* apresenta-se o problema da presença de ruídos impulsivos em sistemas OFDM/DMT (*Discrete Multitone Transceiver/ Orthogonal Frequency Division Multiplexing*) e, em seguida, introduz-se técnicas baseadas em inteligência computacional para cancelar a presença dos ruídos impulsivos na saída de canais PLC. Uma contribuição para o problema de equalização de canais PLC é fornecida no *Capítulo 5*. O foco principal desse capítulo é a investigação de alguns algoritmos fuzzy do tipo 2, propostos no *Capítulo 2*, para a equalização de canais PLC. No *Capítulo 6* discute-se o problema do monitoramento da qualidade de energia elétrica e, a seguir, introduz-se o uso de técnicas de processamento de sinais e de inteligência computacional para auxiliar a análise dos distúrbios que ocorrem nas redes de distribuição e de transmissão de energia elétrica. O *Capítulo 7* apresenta o problema da análise espectral em sistemas de potência e, a seguir, propõe

uma técnica baseada na warped-DFT (*Warped Discrete Fourier Transform*), para a estimação dos parâmetros das componentes fundamental e harmônicas. O *Capítulo 8* discute a detecção de distúrbios e propõe uma nova técnica para a detecção de distúrbios em redes elétricas. Para a classificação de distúrbios, uma nova técnica que possibilita não apenas a classificação de distúrbios isolados, mas também de múltiplos distúrbios em sistemas de potência é proposta no *Capítulo 9*. No *Capítulo 10* duas novas técnicas para a compressão de distúrbios em sistemas de potência são introduzidas.

Finalmente, o *Capítulo 11* apresenta as conclusões e observações finais da presente tese e lista possíveis investigações futuras. O apêndice reúne a produção científica, resultante das investigações realizadas durante o doutorado.

2

Sistemas Fuzzy do Tipo 2

Recentemente, os sistemas fuzzy do tipo 2 têm sido investigados para a solução de problemas envolvendo elevados níveis de incertezas. Teóricos argumentam que os algoritmos baseados nestes sistemas fuzzy são capazes de levar tais incertezas em suas formulações e, conseqüentemente, apresentam melhores resultados que aqueles obtidos com algoritmos fundamentados em sistemas fuzzy do tipo 1. No entanto, os sistemas fuzzy do tipo 2 são mais difíceis de serem entendidos e utilizados que os sistemas fuzzy do tipo 1. Isso se deve às seguintes razões: *i*) a natureza tri-dimensional dos conjuntos fuzzy do tipo 2 é de difícil visualização e *ii*) a complexidade computacional demandada no projeto e no uso dos sistemas fuzzy

do tipo 2 é extremamente elevada.

Os sistemas fuzzy do tipo 2 têm sido recentemente utilizados com sucesso nas seguintes aplicações: classificação de seqüências de vídeo codificadas, eliminação da interferência co-canal em canais de comunicação não lineares e variantes no tempo, controle de admissão de conexão, equalização de canais não lineares com desvanecimento rápido, extração de conhecimento, predição de séries temporais, aproximação de funções, pre-processamento de imagens, resolução de equações fuzzy, logística para transporte, classificação de voz, controle, etc. Nessas aplicações, os sistemas fuzzy do tipo 2 mostraram desempenhos superiores àqueles obtidos com outras técnicas, pois em tais aplicações os níveis de incertezas são bastante elevados.

De um modo geral, os sistemas fuzzy do tipo 2 são aplicados nas seguintes situações (Mendel & John, 2002): *i*) o sistema de geração de dados é variante no tempo e a descrição matemática desta variação temporal não é conhecida, *ii*) o ruído de medição é não estacionário e a representação matemática do mesmo é desconhecida, *iii*) os parâmetros em aplicações de reconhecimento de padrões apresentam não estacionaridade cuja representação matemática é desconhecida, *iv*) o conhecimento é obtido a partir de um grupo de especialistas, e, finalmente, *v*) os termos lingüísticos apresentam domínios não mensuráveis.

Neste contexto, objetivando investigar o uso dos sistemas fuzzy do tipo 2 para solucionar alguns dos problemas tratados nesta tese, neste capítulo introduzimos algoritmos fuzzy do tipo 2, treinados por métodos de otimização baseados em informações de 1ª e 2ª ordens. O uso de tais algoritmos para a redução da presença de ruídos impulsivos em receptores de sistemas multi-portadoras OFDM/DMT, a equalização de canais PLC e a classificação de distúrbios em redes elétricas são questões discutidas nos capítulos 4, 5 e 9, respectivamente. Tais problemas se enquadram no escopo de aplicação dos sistemas fuzzy do tipo 2.

Objetivando justificar o uso de sistemas fuzzy do tipo 2 para essas questões, a Seção 2.1 traz uma breve introdução sobre sistemas fuzzy do tipo 2 a partir dos sistemas fuzzy do tipo 1. A seguir, a Seção 2.2 apresenta os algoritmos fuzzy do tipo 2 aplicados para o cancelamento de ruídos impulsivos em sistemas DMT/OFDM no capítulo 4 e a Seção 2.3 introduz os algoritmos fuzzy do tipo 2 utilizados para a equalização de canais PLC no capítulo 5. Finalmente, a Seção 2.4 introduz o uso de um algoritmo fuzzy do tipo 2 para a classificação de distúrbios em redes elétricas e mostra a relação do mesmo com o classificador Bayesiano baseado no critério ML e o classificador fuzzy do tipo 1, quando algumas restrições são consideradas.

2.1 Breve Introdução sobre Sistemas Fuzzy do Tipo 2

A Fig. 2.1 mostra a estrutura do sistema fuzzy do tipo 1. Na Fig. 2.1, o termo *crisp* refere-se a um conjunto \mathcal{A} que tem a função de pertinência dada por

$$\mathcal{A} \Rightarrow \mu_{\mathcal{A}}(x) = \begin{cases} 1, & \text{se } x \in \mathcal{A} \\ 0, & \text{se } x \notin \mathcal{A} \end{cases} . \quad (2.1)$$

Quando uma entrada é aplicada neste sistema fuzzy, a mesma é fuzzificada e, em seguida, a máquina de inferência calcula o conjunto fuzzy de saída do tipo 1 correspondente a cada regra. A seguir, o defuzzificador calcula o valor *crisp* da saída do sistema fuzzy. Esse processo de obtenção da saída do sistema fuzzy do tipo 1 é expresso da seguinte maneira:

Seja um conjunto fuzzy do tipo 1 com p entradas e uma saída, o fuzzificador do tipo *singleton*, a defuzzificação pelo centro dos conjuntos (*center of sets*) e as M

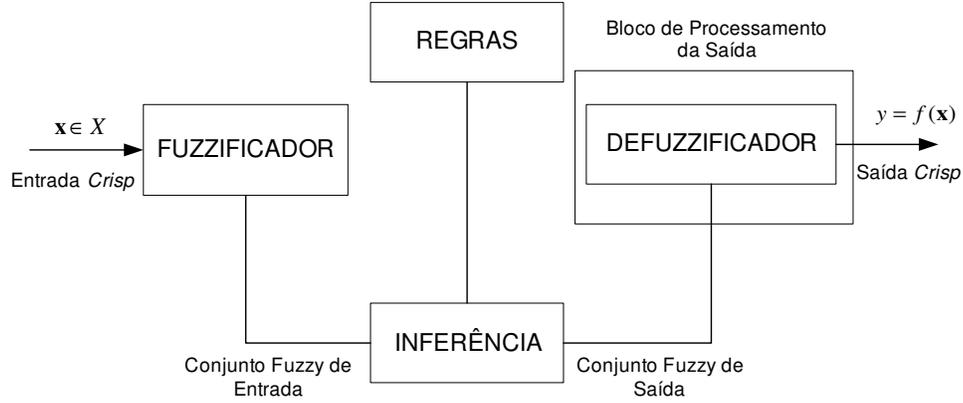


Figura 2.1: Estrutura de um sistema fuzzy do tipo 1.

regras fuzzy dada por (Mendel, 2001, 1995)

$$R^l : \text{Se } x_1 \text{ é } F_1^l \text{ e } \dots \text{ e } x_p \text{ é } F_p^l, \text{ Então } y^l \text{ é } G^l \quad l = 1, \dots, M, \quad (2.2)$$

onde

$$F_i^l = \left\{ \left(x, \mu_{F_i^l}(x) \right) \mid x \in X, \mu_{F_i^l}(x) \in [0, 1] \right\} \quad (2.3)$$

é o i -ésimo conjunto fuzzy do tipo 1 da l -ésima regra fuzzy cuja função de pertinência dada por $\mu_{F_i^l}(x_i)$, a expressão lingüística " Se x_1 é F_1^l e ... e x_p é F_p^l " é o antecedente da l -ésima regra fuzzy, a sentença lingüística " então y^l é G^l " é o conseqüente da l -ésima regra fuzzy e G^l um conjunto fuzzy do tipo 1. A Fig. 2.2 exemplifica uma função Gaussiana de pertinência utilizada no antecedente de uma regra do sistema fuzzy do tipo 1, enquanto a Fig. 2.3 mostra as funções de pertinência do tipo 1 utilizada para modelar a imprecisão no valor de uma variável, por exemplo a temperatura ambiente.

Assumindo que a fuzzificação é *singleton*, então quando o vetor de entrada $\mathbf{x} = [x_1 \cdots x_p]^T$ é aplicado ao sistema fuzzy do tipo 1, o grau de ativação da l -ésima regra é expresso por

$$\mu_{F_1^l}(x_1) \star \mu_{F_2^l}(x_2) \star \dots \star \mu_{F_p^l}(x_p) = \mathcal{T}_{i=1}^p \mu_{F_i^l}(x_i), \quad (2.4)$$

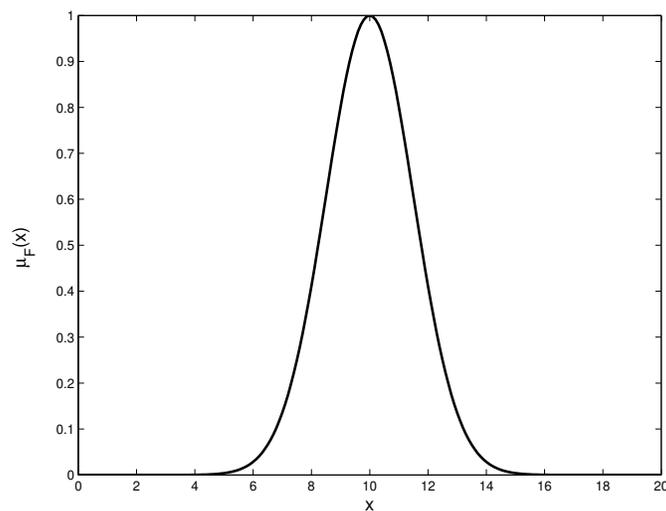


Figura 2.2: Função Gaussiana de pertinência de um sistema fuzzy do tipo 1.

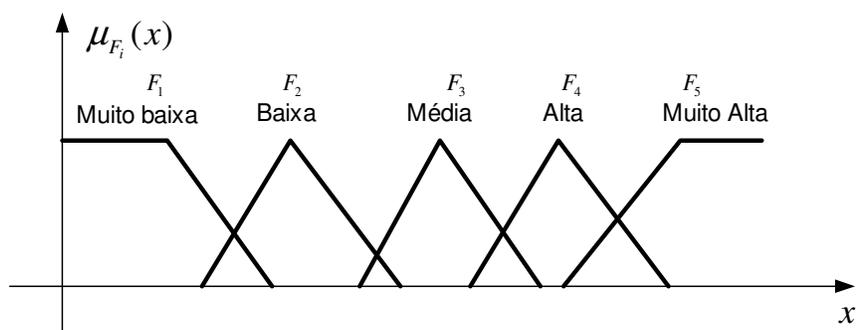


Figura 2.3: Funções de pertinência para a modelar a incerteza no valor da temperatura quando o sistema fuzzy é do tipo 1.

onde \star e \mathcal{T} indicam as t -norm escolhidas. As t -norms usualmente empregadas em sistemas fuzzy do tipo 1 para aplicações de engenharia são os operadores produto e mínimo. Considerando que o defuzzificador é baseado no centro dos conjuntos, então a saída *crisp* do sistema fuzzy do tipo 1 e *singleton* é dada por

$$y_{\text{cos}} = \frac{\sum_{l=1}^M c_{G^l} \mathcal{T}_{i=1}^p \mu_{F_i^l}(x_i)}{\sum_{l=1}^M \mathcal{T}_{i=1}^p \mu_{F_i^l}(x_i)}, \quad (2.5)$$

onde y_{cos} é a saída do sistema fuzzy do tipo 1 e *singleton* quando o método de defuzzificação é o centro dos conjuntos (*center of sets*). Note que na defuzzificação pelo centro dos conjuntos, o conjunto fuzzy do conseqüente da l -ésima regra é substituído por um valor c_{G^l} situado no centro deste conjunto fuzzy. Tal valor tem uma amplitude igual ao nível de ativação do antecedente da regra fuzzy dado por $\mathcal{T}_{i=1}^p \mu_{F_i^l}(x_i)$. Finalmente, a saída y_{cos} fornece o valor do centro do conjunto fuzzy do tipo 1 constituído pelos valores c_{G^l} , $l = 1, \dots, M$. Usando-se outros métodos de defuzzificação diferentes saídas serão obtidas.

Os sistemas fuzzy do tipo 1, baseados nos fuzzificadores *singleton*, têm sido amplamente utilizado devido à simplicidade e baixo custo computacional destes fuzzificadores. Entretanto, este tipo de fuzzificador pode não ser adequado em casos onde há a presença de ruídos nos dados de treinamento ou nos dados que serão futuramente processados pelos sistemas fuzzy. Para considerar a presença de ruídos nas entradas dos sistemas fuzzy do tipo 1, em (Mouzouris & Mendel, 1997) foram desenvolvidos fuzzificadores *nonsingleton* que assumem que os dados na entrada dos sistemas fuzzy do tipo 1 apresentam incertezas e, conseqüentemente, tais dados têm associados à eles funções de pertinência que assumem valores no intervalo $[0, 1]$. Substituindo o fuzzificador *singleton* pelo fuzzificador *nonsingleton* no sistema fuzzy do tipo 1, discutido acima, tem-se que a saída do mesmo é dada por

$$y_{\text{cos}} = \frac{\sum_{l=1}^M c_{G^l} \mathcal{T}_{i=1}^p \mu_{F_i^l}(x_{i,\text{max}}^l)}{\sum_{l=1}^M \mathcal{T}_{i=1}^p \mu_{F_i^l}(x_{i,\text{max}}^l)}, \quad (2.6)$$

onde

$$x_{i,max}^l = \sup_{x_{i,max}^l \in X_i} \mu_{X_i}(x_i) \star \mu_{F_i^l}(x_i), \quad i = 0, \dots, p-1 \quad (2.7)$$

e $\mu_{X_i}(x_i)$ e $\mu_{F_i^l}(x_i)$ são, respectivamente, as funções de pertinência do sinal na entrada do sistema fuzzy e da regra da função de pertinência.

Apesar dos sistemas fuzzy do tipo 1, discutidos acima, serem largamente empregados em diversas áreas (Pedrycz & Gomide, 1998; Mendel, 2000, 1995), sabe-se que os mesmos não consideram as incertezas que ocorrem nos antecedentes e conseqüentes das regras. Desta forma, o uso de tais sistemas fuzzy é limitado a solucionar problemas nos quais o grau de incerteza é baixo. Os dois exemplos dados a seguir ilustram muito bem o problema da incerteza nos antecedentes e conseqüentes das regras fuzzy (Liang, Karnik, & Mendel, 2000):

i) se especialistas diferentes fornecem respostas variadas para uma mesma questão, então as regras fuzzy usadas para representar uma pergunta, apesar de terem antecedentes iguais (perguntas iguais), apresentam conseqüentes diferentes (respostas diferentes). As respostas para as questões baseadas em regras fuzzy geram, assim, conseqüentes incertos;

ii) sabendo-se que as palavras têm significados variados para pessoas diferentes e supondo-se que tais palavras estão associadas às funções de pertinência, provavelmente conjuntos diferentes de parâmetros destas funções de pertinência serão fornecidos por diferentes especialistas. Isso resulta em incertezas na descrição dos antecedentes dos sistemas fuzzy e, conseqüentemente, incertezas adicionais nos conseqüentes aparecem.

Em sistemas fuzzy, as incertezas presentes nas regras fuzzy podem ocorrer devido às incertezas lingüísticas e/ou numéricas associadas aos dados e informações usadas

na construção destas regras. A literatura mostra que tais tipos de incertezas somente podem ser tratadas com sistemas fuzzy do tipo 2 (Karnik, Mendel, & Liang, 1999). Os sistemas fuzzy do tipo 2 foram introduzidos em (Zadeh, 1975). A formulação dos sistemas fuzzy do tipo 2 se fundamenta na conceituação dos conjuntos fuzzy do tipo 2 que, de acordo com (Zadeh, 1975), representa uma extensão do conceito de conjunto fuzzy do tipo 1 expresso por (2.3).

Um conjunto fuzzy do tipo 2, simbolizado por \tilde{A} , é caracterizado pela função de pertinência do tipo 2, $\mu_{\tilde{A}}(x, u)$, onde $x \in X$ e $u \in J_x \subseteq [0, 1]$, ou seja,

$$\begin{aligned} \tilde{A} &= \{((x, u), \mu_{\tilde{A}}(x, u)) \mid \forall x \in X, \forall u \in J_x \subseteq [0, 1]\} \\ &0 \leq \mu_{\tilde{A}}(x, u) \leq 1 \end{aligned} \quad (2.8)$$

O conjunto fuzzy do tipo 2 também pode ser representado por

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x, u)/(x, u), \quad J_x \subseteq [0, 1], \quad (2.9)$$

onde \int representa a operação de união em sistemas fuzzy do tipo 2 no domínio contínuo, $X \in \mathfrak{R}$. O conjunto fuzzy do tipo 2 é caracterizado pela reunião de várias funções de pertinência do tipo 1 que são associadas a todos os elementos do conjunto fuzzy do tipo 2.

Em sistemas fuzzy do tipo 2 as seguintes definições são consideradas (Mendel, 2001):

i) Para cada valor $x=x'$, tem-se que o plano em duas dimensões, cujos eixos são u e $\mu_{\tilde{A}}(x', u)$, é denominado fatia vertical de $\mu_{\tilde{A}}(x, u)$. A partir desta definição, define-se que a função secundária de pertinência de \tilde{A} é uma fatia vertical de $\mu_{\tilde{A}}(x, u)$ dada por

$$\mu_{\tilde{A}}(x = x', u) \equiv \mu_{\tilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/(u), \quad (2.10)$$

onde $J_{x'} \subseteq [0, 1]$ e $0 \leq f_{x'}(u) \leq 1$. Observe que $\mu_{\tilde{A}}(x')$ é um conjunto fuzzy do tipo 1 que representa a função secundária de pertinência. Baseando-se em (2.10), pode-se representar \tilde{A} por

$$\tilde{A} = \int_{x \in X} \mu_{\tilde{A}}(x)/x = \int_{x \in X} \left[\int_{u \in J_x} f_x(u)/u \right] /x, \quad J_x \subseteq [0, 1]; \quad (2.11)$$

ii) o domínio da função secundária de pertinência é chamado de pertinência primária de x . Assim sendo, J_x é a pertinência primária de x , onde $J_x \subseteq [0, 1]$ para $\forall x \in X$;

iii) para o conjunto dos números discretos, o conjunto fuzzy do tipo 2 é expresso por

$$\tilde{A} = \sum_{x \in X} \left[\sum_{u \in J_x} f_x(u)/u \right] /x = \sum_{i=1}^N \left[\sum_{u \in J_{x_i}} f_{x_i}(u)/u \right] /x_i, \quad (2.12)$$

onde \sum é a operação de união em sistemas fuzzy do tipo 2 no domínio discreto;

iv) a amplitude da função secundária de pertinência é chamada grade secundária. Note que $f_x(u)$ é a grade secundária de \tilde{A} ;

v) a operação união entre dois conjuntos fuzzy do tipo 2 é dada por

$$\mu_{\tilde{A} \sqcup \tilde{B}}(x) \equiv \mu_{\tilde{A}} \sqcup \mu_{\tilde{B}} = \int_{u \in U} \int_{w \in W} [f_x(u) \star g_x(w)]/u \vee w, \quad (2.13)$$

onde \sqcup é o operador *join*, \vee é o operador máximo.

vi) a operação intersecção entre dois conjuntos fuzzy do tipo 2 é expressa por

$$\mu_{\tilde{A} \sqcap \tilde{B}}(x) \equiv \mu_{\tilde{A}} \sqcap \mu_{\tilde{B}} = \int_{u \in U} \int_{w \in W} [f_x(u) \star g_x(w)]/u \wedge w, \quad (2.14)$$

onde \sqcap é o operador *meet*, \wedge é o operador mínimo ou produto.

vii) a operação complemento para o conjunto fuzzy do tipo 2 é expressa por

$$\overline{\tilde{A}} \Leftrightarrow \mu_{\overline{\tilde{A}}}(x) \equiv \neg \mu_{\tilde{A}}(x) = \int_{u \in J_x^u} f_x(u)/(1-u), \quad (2.15)$$

onde \neg é o operador negação.

viii) a incerteza na pertinência primária de um conjunto fuzzy do tipo 2, \tilde{A} , consiste de uma região limitada denominada *footprint of uncertainty* (FOU). A FOU é a união de todas as pertinências primárias do conjunto fuzzy do tipo 2, ou seja,

$$FOU(\tilde{A}) = \cup_{x \in X} J_x. \quad (2.16)$$

ix) as funções *upper* e *lower* de pertinência são funções (de pertinência) do tipo 1 que limitam a FOU de um conjunto fuzzy do tipo 2 \tilde{A} . Enquanto a função *upper* de pertinência, expressa por

$$\overline{\mu}_{\tilde{A}}(x) \equiv \overline{FOU(\tilde{A})}, \quad (2.17)$$

é associada ao limite superior da FOU(\tilde{A}), $\forall x \in X$, a função *lower* de pertinência é expressa por

$$\underline{\mu}_{\tilde{A}}(x) \equiv \underline{FOU(\tilde{A})}. \quad (2.18)$$

Esta última é associada ao limite inferior da FOU(\tilde{A}), $\forall x \in X$.

A Fig. 2.4 ilustra um conjunto fuzzy do tipo 2 e Gaussiano, pois a função secundária de pertinência deste conjunto, mostrada na Fig. 2.5, é dada por uma função Gaussiana cujos parâmetros são dados por $\mathcal{N}(0, 0, 01)$. Conforme pode ser observado, a intensidade da região sombreada é modelada pela função secundária de pertinência. Pode-se notar que as regiões escuras são representadas por valores da função secundária de pertinência quando $x \in [0, 8]$. Uma região plana de incerteza ocorre quando $x \in [4.0, 5.5]$, pois a pertinência primária só pode assumir valores no intervalo $[0, 1]$. O domínio da grade de pertinência para $x' = 3$ é também mostrado

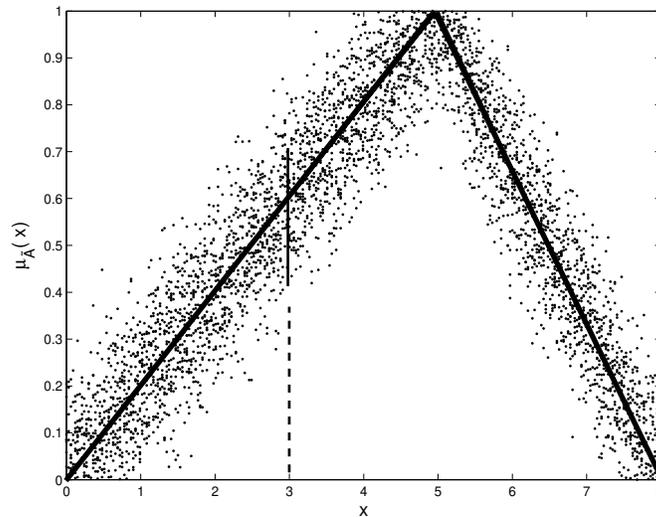


Figura 2.4: Representação de um conjunto fuzzy do tipo 2.

na fig. 2.5. Conforme pode ser observado, a grade de pertinência é um conjunto fuzzy do tipo 1 que assume valores no intervalo $[0,1]$.

A Fig. 2.6 mostra a estrutura de um sistema fuzzy do tipo 2 *singleton* que é bastante similar à estrutura do sistema fuzzy do tipo 1 e *singleton* mostrada na Fig. 2.1. No entanto, deve-se ressaltar que para o sistema fuzzy do tipo 2, os antecedentes e os consequentes são conjuntos fuzzy do tipo 2, ou seja, a saída de cada regra é um conjunto fuzzy do tipo 2. O bloco de processamento da saída do sistema fuzzy do tipo 2 realiza as operações de redução de tipo e defuzzificação. Baseando-se em funções estendidas das técnicas de defuzzificação dos sistemas fuzzy do tipo 1, os conjuntos fuzzy do tipo 2, gerados pela máquina de inferência, são reduzidos em conjuntos fuzzy do tipo 1. Tal processo, denominado redução de tipo (*type reduction*), foi proposto como um estágio anterior ao processo de defuzzificação. Finalmente, o defuzzificador calcula a saída *crisp* do sistema fuzzy do tipo 2, a partir dos conjuntos fuzzy do tipo 1 obtidos na saída do bloco redutor de tipo (Mendel, 2001).

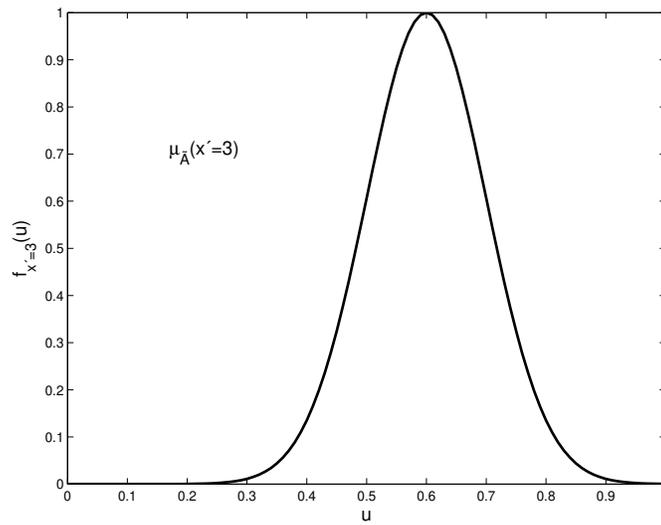


Figura 2.5: Função de pertinência secundária do conjunto fuzzy do tipo 2, mostrado na Fig. 2.4.

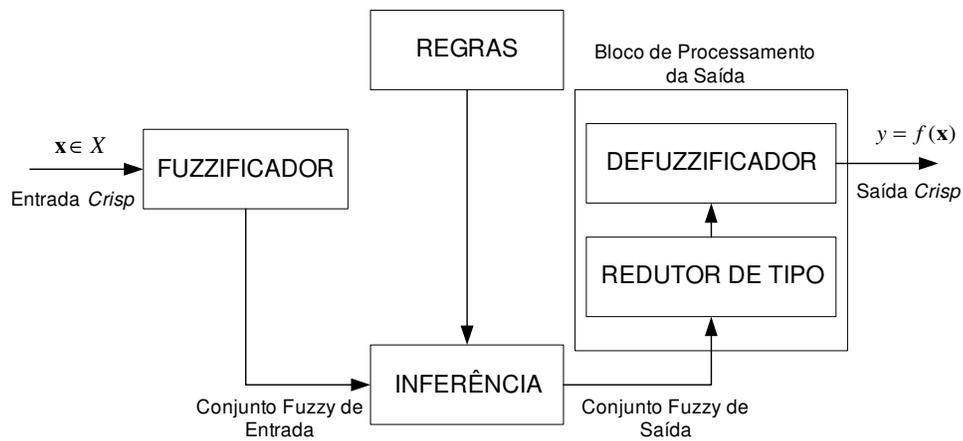


Figura 2.6: Estrutura de um sistema fuzzy do tipo 2.

Supondo-se que o fuzzificador é *nonsingleton*, os sistemas fuzzy do tipo 2 são denominados sistemas fuzzy do tipo 2 e *nonsingleton*. Tais sistemas são adequados para os problemas onde, além da existência das incertezas nos antecedentes e conseqüentes da regras, os dados de treinamento e/ou os sinais futuros a serem aplicados aos sistemas fuzzy são corrompidos pela presença de ruídos aditivos.

O sistema fuzzy do tipo 2 e *singleton* é expresso da seguinte maneira:

Seja um sistema fuzzy do tipo 2, a fuzzificação *singleton*, o processo *type reduction* baseado no centroidos conjuntos fuzzy do tipo 2 (Mendel, 2001) e a l -ésima regra dada por

$$R^l : \text{Se } x_1 \text{ é } \tilde{F}_1^l \text{ e } \dots \text{ e } x_p \text{ é } \tilde{F}_p^l, \text{ Então } \theta^l \text{ é } \tilde{G}^l \quad l = 1, \dots, M, \quad (2.19)$$

então, o nível de ativação da l -ésima regra é dado por

$$\mu_{\tilde{F}_1^l}(x_1) \sqcap \mu_{\tilde{F}_2^l}(x_2) \sqcap \dots \sqcap \mu_{\tilde{F}_p^l}(x_p) = \sqcap_{i=1}^p \mu_{\tilde{F}_i^l}(x_i). \quad (2.20)$$

Após o cálculo do centro do conseqüente, o redutor de tipo é aplicado para reduzir o conjunto fuzzy do tipo 2, dado por (2.20), para um conjunto fuzzy do tipo 1. A saída do redutor de tipo é dada por

$$Y_{\cos}(\mathbf{x}) = \int_{\theta_1} \dots \int_{\theta_M} \int_{f_1} \dots \int_{f_M} T_{l=1}^M \mu_{\tilde{C}^l}(\theta_l) * T_{l=1}^M \mu_{\tilde{F}^l}(f_l) \left/ \frac{\sum_{i=1}^M \theta_i f_i}{\sum_{i=1}^M f_i} \right., \quad (2.21)$$

onde \mathcal{T} e \star indicam as *t-norms*, $\mu_{\tilde{C}^l}$ é a função de pertinência do centro da l -ésima regra, $\theta_l \in \tilde{C}^l = C_{\tilde{C}^l}$ é o centro do conjunto fuzzy do tipo 2 do conseqüente da l -ésima regra, $f_l \in \tilde{F}^l = \sqcap_{i=1}^p \mu_{\tilde{F}_i^l}(x_i)$ é o grau de ativação associado ao conjunto fuzzy do tipo 2 do l -ésimo conseqüente. Finalmente, a saída *crisp* do sistema fuzzy do tipo 2 é obtida após o defuzzificador calcular o centro da expressão (2.21).

O sistema fuzzy do tipo 2, descrito acima, é bastante complexo, pois o cálculo da expressão (2.21) requer altíssima carga computacional. No entanto, supondo-se

que a grade secundária é igual 1, ou seja $f_x(u) = 1$ $x \in X$ e $u \in J_x$, tem-se que a complexidade computacional para o cálculo da saída do sistema fuzzy do tipo 2 diminui consideravelmente. Neste caso, o sistema fuzzy é denominado sistema fuzzy do tipo 2 e intervalado. A entrada do defuzzificador desse sistema fuzzy do tipo 2 e intervalado é dado por (Liang & Mendel, 2000b)

$$Y_{\text{cos}}(\mathbf{x}) = \int_{\theta_1} \dots \int_{\theta_M} \int_{f_1} \dots \int_{f_M} 1 \left/ \frac{\sum_{l=1}^M \theta_l f_l}{\sum_{l=1}^M f_l} \right. = [y^l, y^r], \quad (2.22)$$

onde $f^l \in [\underline{f}^l, \overline{f}^l]$,

$$\underline{f}^l = \prod \underline{\mu}_{\tilde{F}_2^l}(x_2) \prod \dots \prod \underline{\mu}_{\tilde{F}_p^l}(x_p) = \prod_{i=1}^p \underline{\mu}_{\tilde{F}_i^l}(x_i), \quad (2.23)$$

$$\overline{f}^l = \prod \overline{\mu}_{\tilde{F}_2^l}(x_2) \prod \dots \prod \overline{\mu}_{\tilde{F}_p^l}(x_p) = \prod_{i=1}^p \overline{\mu}_{\tilde{F}_i^l}(x_i). \quad (2.24)$$

Finalmente, a saída *crisp* obtida pelo processo de defuzzificação do sistema fuzzy do tipo 2 e intervalado é dada por

$$y = \frac{y^l + y^r}{2}. \quad (2.25)$$

Em (Liang et al., 2000) foi proposto um procedimento para o cálculo da saída do sistema fuzzy do tipo 2 e intervalado.

A Fig. 2.7 ilustra um conjunto fuzzy do tipo 2 intervalado, enquanto a Fig. 2.8 mostra a função de pertinência deste conjunto fuzzy. Conforme é observado, a grade de pertinência é constante e igual a 1.

Os sistemas fuzzy do tipo 2 e intervalados são bastante útil quando não há informações acerca das funções secundárias de pertinência. Esta restrição nos conjuntos fuzzy do tipo 2 é similar à suposição de que um conjunto de dados apresenta distribuição uniforme quando a teoria das probabilidades é aplicada. A Fig. 2.9 mostra funções de pertinência do tipo 2 usadas para modelar a incerteza nas regras fuzzy aplicadas para fuzzificar os valores da temperatura.

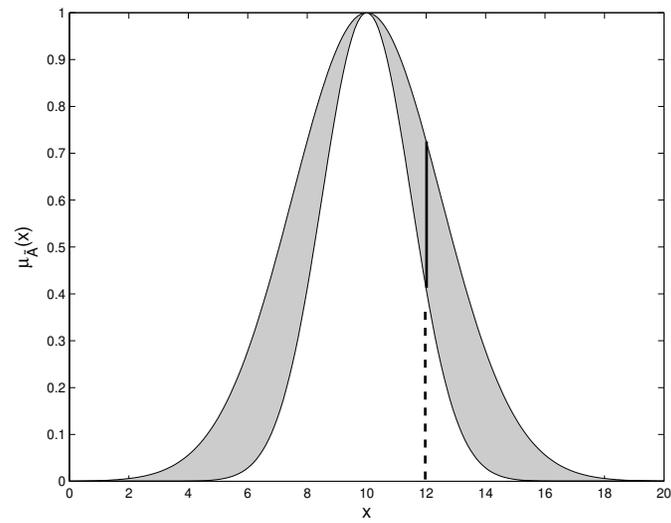


Figura 2.7: Representação de um conjunto fuzzy do tipo 2 intervalado.

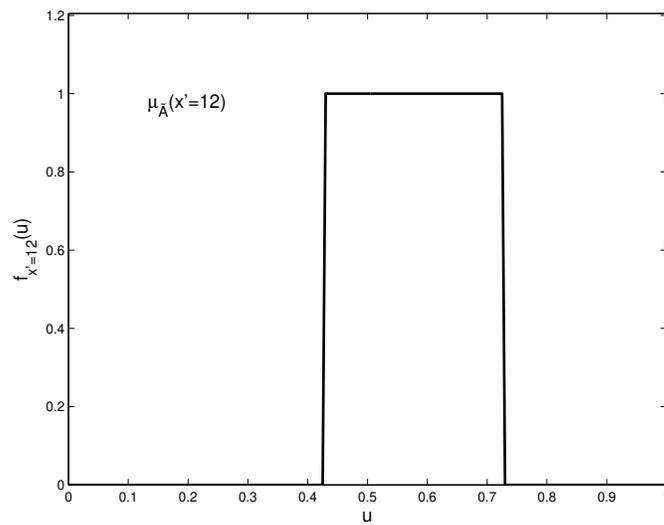


Figura 2.8: Função de pertinência secundária do conjunto fuzzy do tipo 2 intervalado, mostrado na Fig. 2.7.

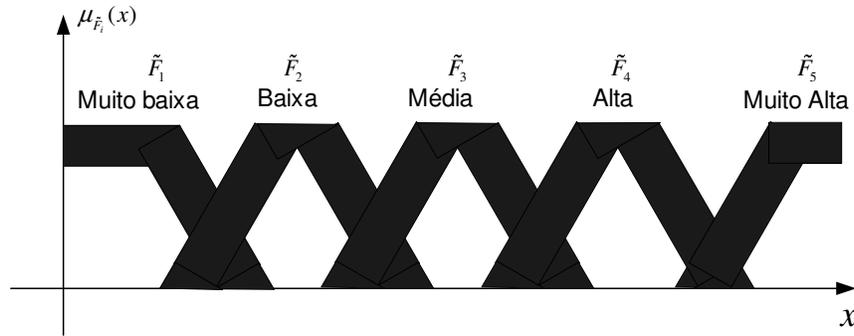


Figura 2.9: Funções de pertinência usadas para modelar as incertezas relacionadas à medida temperatura, quando o sistema fuzzy aplicado é o do tipo 2.

Outros comentários interessantes sobre os sistemas fuzzy do tipo 2, visando mostrar algumas similaridade como os sistemas baseados em modelos estatísticos, são reunidos a seguir: *i*) da mesma forma que as incertezas geradas por ruídos podem ter seus efeitos analisados com a média e a variância, as incertezas lingüísticas e aquelas geradas por ruídos podem ter seus efeitos analisados com os sistemas fuzzy do tipo 2, *ii*) da mesma forma que a variância fornece uma medida de dispersão em torno da média, a saída do redutor de tipo do sistema fuzzy do tipo 2 também fornece uma medida de dispersão em torno da saída do defuzzificador, e finalmente, *iii*) da mesma forma que a variância aumenta à medida que a incerteza do ruído aumenta, o conjunto fuzzy gerado pelo redutor de tipo também têm seu domínio expandido à medida que as incertezas lingüísticas e/ou aquelas geradas por ruídos aumentam.

Apesar do ganho de desempenho em termos de custo computacional obtido com o uso dos sistemas fuzzy do tipo 2 e intervalado, observa-se que o cálculo da saída destes sistemas é ainda complexo e a atualização dos parâmetros dos mesmos demanda alta carga computacional, conforme é discutido em (Mendel, 2004). Dadas essas limitações dos sistemas fuzzy do tipo 2, em (Liang et al., 2000) foi proposto a seguinte regra, baseando-se em sistema fuzzy Mandami, para modelar as regras dos

sistemas fuzzy do tipo 2 e intervalado

$$R^k : \text{Se } x_1 \text{ é } \tilde{F}_1^k \text{ e } \dots \text{ e } x_p \text{ é } \tilde{F}_p^k, \text{ Então } \theta \text{ é } G^k \quad k = 1, \dots, M \quad (2.26)$$

onde G^k é um conjunto fuzzy do tipo 1, cuja função de pertinência é dada por

$$\mu_{G^k}(\theta) = \begin{cases} 1, & \text{se } \theta = \theta_k \\ 0, & \text{caso contrário} \end{cases} . \quad (2.27)$$

Fazendo uso de (2.26), tem-se que a saída do sistema *fuzzy* do tipo 2 e intervalado é expressa por

$$y_n(\mathbf{x}) = \frac{y_n^l + y_n^r}{2} = \frac{\sum_{k=1}^M \theta_k \left(\underline{f}^k(\mathbf{x}) + \overline{f}^k(\mathbf{x}) \right)}{\sum_{k=1}^M \left(\underline{f}^k(\mathbf{x}) + \overline{f}^k(\mathbf{x}) \right)} \quad (2.28)$$

onde

$$\frac{y_n^l}{2} = \frac{\sum_{k=1}^M \theta_k \underline{f}^k(\mathbf{x})}{\sum_{k=1}^M \left(\underline{f}^k(\mathbf{x}) + \overline{f}^k(\mathbf{x}) \right)} \quad (2.29)$$

e

$$\frac{y_n^r}{2} = \frac{\sum_{k=1}^M \theta_k \overline{f}^k(\mathbf{x})}{\sum_{k=1}^M \left(\underline{f}^k(\mathbf{x}) + \overline{f}^k(\mathbf{x}) \right)} . \quad (2.30)$$

Supondo-se que o sistema fuzzy do tipo 2 é intervalado não é normalizado, tem-se que a saída do mesmo é dada por

$$y_{\tilde{n}}(\mathbf{x}) = \frac{y_{\tilde{n}}^l + y_{\tilde{n}}^r}{2} = \sum_{k=1}^M \theta_k \left(\underline{f}^k(\mathbf{x}) + \overline{f}^k(\mathbf{x}) \right) \quad (2.31)$$

onde

$$\frac{y_{\tilde{n}}^l}{2} = \sum_{k=1}^M \theta_k \underline{f}^k(\mathbf{x}) \quad (2.32)$$

e

$$\frac{y_{\tilde{n}}^r}{2} = \sum_{k=1}^M \theta_k \overline{f}^k(\mathbf{x}) . \quad (2.33)$$

Os sistemas fuzzy do tipo 2 não normalizado são aqui propostos, pois o uso dos métodos de otimização baseados em informação de 2ª ordem é bastante complicada

para a obtenção dos parâmetros das funções de pertinência dos sistemas fuzzy do tipo 2 e intervaldos, cuja saída é dada por (2.28).

Na dedução das expressões (2.28) e (2.31) supôs-se que os conseqüentes das regras fuzzy são independentes de seus antecedentes correspondentes. Tal restrição permite que técnicas de otimização, baseadas no método *backpropagation*, sejam aplicadas em (2.28) e (2.31), uma vez que o procedimento para cálculo da saída dos sistemas fuzzy do tipo 2, introduzido em (Liang et al., 2000; Mendel, 2001), não é aqui adotado. Esta restrição facilita o processo de obtenção e ajuste dos parâmetros do algoritmo baseado em sistemas fuzzy do tipo 2.

Objetivando explorar as expressões (2.28) e (2.31) para a obtenção de algoritmos fuzzy adaptativos cujos treinamentos são baseados em técnicas de otimização que fazem uso de informações de 1ª e 2ª ordens, a Seção 2.2 introduz dois novos algoritmos fuzzy treinados pelo método de otimização baseado em informações de 2ª ordem denominado método de otimização baseado em informações de 2ª ordem denominado método Gradiente Conjugado, escalonado e modificado (GCEM) (Santos & Zuben, 2000), o qual é reproduzido na Seção 4.1. A motivação para uso deste método reside no fato de que o mesmo faz uso do operador diferencial dado por (Pearlmutter, 1994)

$$\mathfrak{R}_{\mathbf{d}}\{f(\mathbf{w}(n))\} \equiv \frac{\partial}{\partial \alpha} f(\mathbf{w}(n) + \alpha \mathbf{d}(n))|_{\alpha=0}, \quad (2.34)$$

onde $\mathbf{d}(n)$ é um vetor direcional, $f(\mathbf{w}(n))$ representa o vetor gradiente do vetor de parâmetros $\mathbf{w}(n)$, $0 < \alpha \ll 1$ é um número real e $\mathfrak{R}_{\mathbf{d}}\{f(\mathbf{w}(n))\} = \mathbf{H}(\mathbf{w}(n))\mathbf{w}(n)$ é o produto entre a matriz Hessiana e um vetor direcional $\mathbf{d}(n)$. Tal operador garante que a complexidade do método GCEM aumenta linearmente com o aumento do número de parâmetros, uma vez que a expressão $\mathbf{H}(\mathbf{w}(n))\mathbf{w}(n)$ é calculada sem a necessidade da o cálculo explícito da matriz Hessiana $\mathbf{H}(\mathbf{w}(n))$.

A seguir, a Seção 2.3 introduz outros algoritmos fuzzy adaptativos baseados na

expressão (2.28) e em informações de 1ª ordem. Finalmente, a Seção 2.4 introduz o classificador fuzzy do tipo 2, baseado na expressão (2.28), para a classificação de distúrbios que ocorrem em redes elétricas. Nesta Seção, ainda é mostrado que, sob algumas restrições, tal classificador fuzzy do tipo 2 reduz-se ao classificador baseado no critério de máxima verossimilhança (*Maximum likelihood* - ML) e ao classificador fuzzy do tipo 1, deduzidos em (Wei & Mendel, 2000, 1999).

2.2 Algoritmos fuzzy do Tipo 2, Intervalados, *Nonsingleton* e Adaptativos

Nesta Seção algoritmos fuzzy do tipo 2, intervalados e *nonsingleton* baseados nas expressões (2.31) e (2.28) são introduzidos. Para o desenvolvimento de tais algoritmos, supôs-se que a função de pertinência dos antecedentes das regras fuzzy é modelada pela função Gaussiana e a função de pertinência das entradas é Gaussiana ou exponencial. Assume-se que as *t-norms* são implementada pelo operador produto.

A motivação por trás do uso das funções Gaussiana e exponencial para modelar as incertezas das regras e das entradas dos sistemas fuzzy do tipo 2 reside no fato de que o uso das mesmas já é bastante estabelecido na literatura para aplicações na área de engenharia. De fato, a função Gaussiana é amplamente utilizada em sistemas fuzzy do tipo 1 e 2, enquanto a função exponencial é utilizada apenas em sistema fuzzy do tipo 1 (Mendel, 2000, 2001; Pedrycz & Gomide, 1998).

Na Seção 2.2.1, novos algoritmos fuzzy do tipo 2 *nonsingleton*, intervalado, não normalizado e com função de pertinência Gaussiana para as regras fuzzy e para os sinais nas entradas são introduzidos. A Seção 2.2.2 apresenta os algoritmos fuzzy do tipo 2, *nonsingleton*, intervalado, não normalizado e com função de

pertinência Gaussiana para modelar as incertezas associadas às regras fuzzy e a função de pertinência exponencial para modelar os sinais as incertezas nas entradas dos mesmos.

2.2.1 Algoritmo Fuzzy Não Normalizado I

Supondo-se que o sistema fuzzy do tipo 2 é intervalado e *nonsingleton*, as incertezas associadas às suas entradas e das premissas dos antecedentes de suas regras sejam modeladas pela função Gaussiana e baseando-se na expressão (2.31), nesta Seção deduz-se um algoritmo fuzzy do tipo 2, *nonsingleton* e não normalizado que permite o uso de métodos de otimização baseados em informações de 2ª ordem, tal como o método GCEM reproduzido na Seção 4.1, da seguinte forma.

Sejam

$$\bar{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \bar{m}_{y_i}(n)}{\bar{\sigma}_{y_i}(n)} \right)^2 \right] \quad (2.35)$$

e

$$\underline{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \underline{m}_{y_i}(n)}{\underline{\sigma}_{y_i}(n)} \right)^2 \right] \quad (2.36)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das funções de pertinência do tipo 2 que modelam as incertezas nas entradas do sistema fuzzy do tipo 2, intervalado e *nonsingleton*,

$$\bar{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \bar{m}_{\tilde{F}_i^l}(n)}{\bar{\sigma}_{\tilde{F}_i^l}(n)} \right)^2 \right] \quad (2.37)$$

e

$$\underline{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \underline{m}_{\tilde{F}_i^l}(n)}{\underline{\sigma}_{\tilde{F}_i^l}(n)} \right)^2 \right], \quad (2.38)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das funções de pertinência do tipo 2 que modelam os antecedentes das premissas das regras do

sistemas fuzzy do tipo 2, intervalado e *nonsingleton*. Note que $y_i(n)$ é a i -ésima amostra do vetor $\mathbf{y}(n) = [y_0(n) \cdots y_{L-1}(n)]^T$ de entrada.

Então, tem-se que o supremo entre as funções *upper* de pertinência e o supremo entre as funções *lower* de pertinência que permitem relacionar as incertezas presentes nas entradas com as incertezas presentes nas premissas das regras fuzzy são dados por

$$\bar{\mu}_{\tilde{Q}_i^l}(\bar{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \bar{\mu}_{\tilde{Y}_i}(y_i(n)) \star \bar{\mu}_{\tilde{F}_i^l}(y_i(n)), \quad (2.39)$$

e

$$\underline{\mu}_{\tilde{Q}_i^l}(\underline{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \underline{\mu}_{\tilde{Y}_i}(y_i(n)) \star \underline{\mu}_{\tilde{F}_i^l}(y_i(n)), \quad (2.40)$$

onde o símbolo \star é uma t -norm implementada pelo operador produto, $i = 0, \dots, L-1$ e $l = 1, \dots, M$. Para o nosso sistema fuzzy do tipo 2, as expressões (2.39) e (2.40) alcançam os seus máximos nas seguintes condições (Mendel, 2001; Mouzouris & Mendel, 1997):

$$\bar{y}_{i,\max}^l(n) = \frac{\bar{\sigma}_{y_i}^2(n) \bar{m}_{\tilde{F}_i^l}(n) + \bar{\sigma}_{\tilde{F}_i^l}^2(n) \bar{m}_{y_i}(n)}{\bar{\sigma}_{y_i}^2(n) + \bar{\sigma}_{\tilde{F}_i^l}^2(n)} \quad (2.41)$$

e

$$\underline{y}_{i,\max}^l(n) = \frac{\underline{\sigma}_{y_i}^2(n) \underline{m}_{\tilde{F}_i^l}(n) + \underline{\sigma}_{\tilde{F}_i^l}^2(n) \underline{m}_{y_i}(n)}{\underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_{\tilde{F}_i^l}^2(n)}. \quad (2.42)$$

Em (Mendel, 2001) foi mostrado que o uso das expressões (2.41) e (2.42) para a definição as FOU das premissas das regras fuzzy de um sistema fuzzy do tipo 2 e *nonsingleton* é um procedimento bastante complexo e dependente dos parâmetros das funções de pertinências expressas por (2.35)-(2.38). Tal procedimento acarreta o uso de várias condições para definir de forma apropriada as FOU das premissas das regras fuzzy. A nossa escolha pela extensão direta dos resultados obtidos quando o sistema fuzzy é do tipo 1 e *nonsingleton* para os sistemas fuzzy do tipo 2 e *nonsingleton*, vide expressões (2.41) e (2.42), e não pelo uso do procedimento

proposto por Mendel (Mendel, 2001), significa a imposição de algumas restrições nas FOUs que determinam as regiões de incertezas das premissas das regras fuzzy. Tais suposições, que também são utilizadas nos algoritmos fuzzy do tipo 2 propostos nas seções 2.2.2, 2.3.1 e 2.3.2, facilitam o processo de atualização dos parâmetros dos algoritmos fuzzy do tipo 2. Essa vantagem é mostrada de forma bastante clara nas expressões que definem as saídas dos algoritmos fuzzy propostos e de seus gradientes correspondentes deduzidos nesta Seção e nas seções acima citadas.

Supondo-se nas expressões (2.41) e (2.42) que $m_{y_i(n)} = y_i(n)$, vide consideração feita em (Mendel, 2001) para sistemas fuzzy do tipo 1, tem-se que a saída do algoritmo fuzzy do tipo 2, *nonsingleton* e intervalado é dada por

$$\begin{aligned} f_{ns,2}(\mathbf{y}(n)) &= \sum_{l=1}^M \theta_l(n) \left(\bar{f}_l(\mathbf{y}(n)) + \underline{f}_l(\mathbf{y}(n)) \right) \\ &= \sum_{l=1}^M \theta_l(n) \left[\prod_{i=0}^{L-1} \exp \left(-\frac{1}{2} \frac{(y_i(n) - \underline{m}_{\bar{F}_i^l}(n))^2}{\underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_{\bar{F}_i^l}^2(n)} \right) + \prod_{i=0}^{L-1} \exp \left(-\frac{1}{2} \frac{(y_i(n) - \bar{m}_{\bar{F}_i^l}(n))^2}{\bar{\sigma}_{y_i}^2(n) + \bar{\sigma}_{\bar{F}_i^l}^2(n)} \right) \right], \end{aligned} \quad (2.43)$$

onde $\theta_l(n)$, $\underline{\sigma}_{y_i}(n)$, $\underline{m}_{\bar{F}_i^l}(n)$, $\underline{\sigma}_{\bar{F}_i^l}(n)$, $\bar{\sigma}_{y_i}(n)$, $\bar{m}_{\bar{F}_i^l}(n)$ e $\bar{\sigma}_{\bar{F}_i^l}(n)$ são os parâmetros do algoritmo fuzzy do tipo 2, L é o comprimento do vetor de entrada $\mathbf{y}(n)$ e M é o número de regras.

Agora seja a função custo dada por

$$J(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} J_c(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} \mathbf{e}^2(n) = \frac{1}{2} \sum_{n=0}^{N-1} [y_d(n) - f_{ns,2}(\mathbf{y}(n))]^2, \quad (2.44)$$

onde $y_d(n)$ é a saída desejada, $f_{ns,2}(\mathbf{y}(n))$ é a saída do sistema fuzzy do tipo 2, $\{y_d(n), f_{ns,2}(\mathbf{y}(n))\}_{n=0}^{N-1}$ são os N pares entrada saída utilizados durante o treinamento do algoritmo fuzzy do tipo 2, $\mathbf{w}(n) = [\boldsymbol{\theta}^T(n) \ \bar{\boldsymbol{\sigma}}^T(n) \ \bar{\mathbf{m}}^T(n) \ \bar{\boldsymbol{\sigma}}_{y_i}^T(n) \ \underline{\boldsymbol{\sigma}}^T(n) \ \underline{\mathbf{m}}^T(n) \ \underline{\boldsymbol{\sigma}}_{y_i}^T(n)]^T$ é o vetor constituído de todos os parâmetros do algoritmo fuzzy, $\boldsymbol{\theta}(n) = [\theta_1(n) \cdots \theta_M(n)]^T$ é o vetor de pesos que pondera a influência de cada regra na saída do algoritmo fuzzy, $\bar{\mathbf{m}}(n) = [\bar{\mathbf{m}}_1^T(n) \cdots \bar{\mathbf{m}}_M^T(n)]^T = [\bar{m}_{\bar{F}_0^1}(n) \cdots \bar{m}_{\bar{F}_{L-1}^1}(n) \cdots \bar{m}_{\bar{F}_0^M}(n) \cdots \bar{m}_{\bar{F}_{L-1}^M}(n)]^T$

é o vetor constituído dos valores $\bar{m}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$,
 $\bar{\sigma}(n) = [\bar{\sigma}_1^T(n) \cdots \bar{\sigma}_M^T(n)]^T = [\bar{\sigma}_{\tilde{F}_0^1}(n) \cdots \bar{\sigma}_{\tilde{F}_{L-1}^1}(n) \cdots \bar{\sigma}_{\tilde{F}_0^M}(n) \cdots \bar{\sigma}_{\tilde{F}_{L-1}^M}(n)]^T$
é o vetor constituído dos valores $\bar{\sigma}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$,
 $\underline{\mathbf{m}}(n) = [\underline{\mathbf{m}}_1^T(n) \cdots \underline{\mathbf{m}}_M^T(n)]^T = [\underline{m}_{\tilde{F}_0^1}(n) \cdots \underline{m}_{\tilde{F}_{L-1}^1}(n) \cdots \underline{m}_{\tilde{F}_0^M}(n) \cdots \underline{m}_{\tilde{F}_{L-1}^M}(n)]^T$
é o vetor constituído dos valores $\underline{m}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$,
 $\underline{\sigma}(n) = [\underline{\sigma}_1^T(n) \cdots \underline{\sigma}_M^T(n)]^T = [\underline{\sigma}_{\tilde{F}_0^1}(n) \cdots \underline{\sigma}_{\tilde{F}_{L-1}^1}(n) \cdots \underline{\sigma}_{\tilde{F}_0^M}(n) \cdots \underline{\sigma}_{\tilde{F}_{L-1}^M}(n)]^T$ é
o vetor constituído dos valores $\underline{\sigma}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$, $\bar{\sigma}_{y_i}(n)$ é o
parâmetro da função de pertinência $\bar{\mu}_{\tilde{Y}_i}$, $\underline{\sigma}_{y_i}(n)$ é o parâmetro da função de
pertinência $\underline{\mu}_{\tilde{Y}_i}$. Então pode-se escrever que o vetor

$$\nabla J_c(\mathbf{w}(n)) = \left[\nabla_{\boldsymbol{\theta}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\bar{\boldsymbol{\sigma}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\mathbf{m}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\bar{\sigma}_{y_i}(n)} J_c(\mathbf{w}(n)) \right. \\ \left. \nabla_{\underline{\sigma}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\mathbf{m}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\sigma}_{y_i}(n)} J_c(\mathbf{w}(n)) \right]^T, \quad (2.45)$$

é o vetor gradiente do vetor de parâmetros $\mathbf{w}(n)$.

Supondo-se que as incertezas associadas aos ruídos aditivos são iguais para todas as entradas do sistema fuzzy, tem-se $\underline{\sigma}_{y_i}(n) = \underline{\sigma}_{y_1}(n)$ e $\bar{\sigma}_{y_i}(n) = \bar{\sigma}_{y_1}(n)$ para $i = 0, \dots, L-1$ e levando-se em consideração que o operador de Hadarmard é dado por

$$\mathbf{r}(n) \odot \mathbf{s}(n) = [r_0(n)s_0(n) \cdots r_{P-1}(n)s_{P-1}(n)]^T, \quad (2.46)$$

e $\mathbf{1}_L = [1 \cdots 1]^T$ é um vetor coluna constituído por L elementos iguais a 1, tem-se que os vetores gradientes dos parâmetros do algoritmo fuzzy do tipo 2, cuja saída é expressa por (2.43), são os seguintes:

$$\nabla_{\boldsymbol{\theta}(n)} J_c(\mathbf{w}(n)) = e(n)\mathbf{z}(n), \quad (2.47)$$

onde $\mathbf{z}(n) = \left[\bar{f}_1(\mathbf{y}(n)) + \underline{f}_1(\mathbf{y}(n)) \cdots \bar{f}_M(\mathbf{y}(n)) + \underline{f}_M(\mathbf{y}(n)) \right]^T$,

$$\nabla_{\underline{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right), \quad (2.48)$$

$$\nabla_{\underline{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right) \quad (2.49)$$

$$\nabla_{\bar{\sigma}_l(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right)^2 \odot \bar{\sigma}_l(n), \quad (2.50)$$

$$\nabla_{\underline{\sigma}_l(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right)^2 \odot \underline{\sigma}_l(n), \quad (2.51)$$

$$\nabla_{\bar{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_1(n) \exp(-\bar{u}_1^1(n)) \left(\frac{y_1(n) - \bar{m}_{F_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n)} \right)^2 \bar{\sigma}_{y_1}(n) \quad (2.52)$$

e

$$\nabla_{\underline{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_1(n) \exp(-\underline{u}_1^1(n)) \left(\frac{y_1(n) - \underline{m}_{F_1^1}(n)}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n)} \right)^2 \underline{\sigma}_{y_1}(n), \quad (2.53)$$

onde $l = 1 \dots M$, $\bar{\mathbf{u}}_l(n) = \left[\frac{1}{2} \frac{(y_0(n) - \bar{m}_{F_0^l}(n))^2}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_0^l}^2(n)} \dots \frac{1}{2} \frac{(y_{L-1}(n) - \bar{m}_{F_{L-1}^l}(n))^2}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_{L-1}^l}^2(n)} \right]^T$ e $\underline{\mathbf{u}}_l(n) = \left[\frac{1}{2} \frac{(y_0(n) - \underline{m}_{F_0^l}(n))^2}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_0^l}^2(n)} \dots \frac{1}{2} \frac{(y_{L-1}(n) - \underline{m}_{F_{L-1}^l}(n))^2}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_{L-1}^l}^2(n)} \right]^T$.

Objetivando a aplicação do método de otimização GCEM para o treinamento do algoritmo fuzzy do tipo 2 cuja saída é dada por (2.43), o cálculo do produto entre a matriz Hessiana e um vetor direcional, $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$, é necessária. Tal produto é obtido ao se aplicar o operador diferencial expresso por (2.34) em (2.47)-(2.53). Tal procedimento permite a obtenção das seguintes expressões:

$$\Re_{\mathbf{d}} \{ \nabla_{\boldsymbol{\theta}(n)} J_c(\mathbf{w}(n)) \} = \Re_{\mathbf{d}} \{ e(n) \} \mathbf{z}(n) + e(n) \Re_{\mathbf{d}} \{ \mathbf{z}(n) \}, \quad (2.54)$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\bar{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) \} &= -\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right) \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\bar{\mathbf{u}}_l(n)) \} \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left[(\mathfrak{R}_d \{ \mathbf{y}(n) \} - \mathfrak{R}_d \{ \bar{\mathbf{m}}_l(n) \}) \odot \frac{1}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} + \right. \\
&\quad \left. (\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)) \odot \mathfrak{R}_d \left\{ \frac{1}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right\} \right], \tag{2.55}
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\underline{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) \} &= -\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right) \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\underline{\mathbf{u}}_l(n)) \} \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left[(\mathfrak{R}_d \{ \mathbf{y}(n) \} - \mathfrak{R}_d \{ \underline{\mathbf{m}}_l(n) \}) \odot \frac{1}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} + \right. \\
&\quad \left. (\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)) \odot \mathfrak{R}_d \left\{ \frac{1}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right\} \right], \tag{2.56}
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\bar{\sigma}_l(n)} J_c(\mathbf{w}(n)) \} &= \\
&- \mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right)^2 \odot \bar{\sigma}_l(n) \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right)^2 \odot \bar{\sigma}_l(n) \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\bar{\mathbf{u}}_l(n)) \} \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right)^2 \odot \bar{\sigma}_l(n) \\
&- e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left[(\mathfrak{R}_d \{ (\mathbf{y}(n) - \bar{\mathbf{m}}_l(n))^2 \}) \odot \frac{1}{(\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n))^2} + \right. \\
&\quad \left. (\mathbf{y}(n) - \bar{\mathbf{m}}_l(n))^2 \odot \mathfrak{R}_d \left\{ \frac{1}{(\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n))^2} \right\} \right] \\
&- e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \bar{\mathbf{m}}_l(n)}{\mathbf{1}_L \bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_l^2(n)} \right)^2 \odot \mathfrak{R}_d \{ \bar{\sigma}_l(n) \}, \tag{2.57}
\end{aligned}$$

$$\begin{aligned}
& \mathfrak{R}_d \{ \nabla_{\underline{\sigma}_l(n)} J_c(\mathbf{w}(n)) \} = \\
& -\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_l^2(n)} \right)^2 \odot \underline{\sigma}_l(n) \\
& -e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_l^2(n)} \right)^2 \odot \underline{\sigma}_l(n) \\
& -e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\underline{\mathbf{u}}_l(n)) \} \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_l^2(n)} \right)^2 \odot \underline{\sigma}_l(n) \\
& -e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left[\left(\mathfrak{R}_d \{ (\mathbf{y}(n) - \underline{\mathbf{m}}_l(n))^2 \} \right) \odot \frac{1}{(\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n))^2} + \right. \\
& \qquad \qquad \qquad \left. (\mathbf{y}(n) - \underline{\mathbf{m}}_l(n))^2 \odot \mathfrak{R}_d \left\{ \frac{1}{(\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n))^2} \right\} \right] \\
& -e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\mathbf{y}(n) - \underline{\mathbf{m}}_l(n)}{\mathbf{1}_L \underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_l^2(n)} \right)^2 \odot \mathfrak{R}_d \{ \underline{\sigma}_l(n) \},
\end{aligned} \tag{2.58}$$

$$\begin{aligned}
& \mathfrak{R}_d \{ \nabla_{\bar{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \} = \\
& -\mathfrak{R}_d \{ e(n) \} \theta_1(n) \exp(-\bar{u}_1(n)) \left(\frac{y_1(n) - \bar{m}_{F_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n)} \right)^2 \bar{\sigma}_{y_1}(n) \\
& -e(n) \mathfrak{R}_d \{ \theta_1(n) \} \exp(-\bar{u}_1(n)) \left(\frac{y_1(n) - \bar{m}_{F_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n)} \right)^2 \bar{\sigma}_{y_1}(n) \\
& -e(n) \theta_1(n) \mathfrak{R}_d \{ \exp(-\bar{u}_1(n)) \} \left(\frac{y_1(n) - \bar{m}_{F_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n)} \right)^2 \bar{\sigma}_{y_1}(n) \\
& -e(n) \theta_1(n) \exp(-\bar{u}_1(n)) \left[\left(\mathfrak{R}_d \left\{ \left(y_1(n) - \bar{m}_{F_1^1}(n) \right)^2 \right\} \right) \frac{1}{(\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n))^2} \right. \\
& \qquad \qquad \qquad \left. + (y_1(n) - \bar{m}_{F_1^1}(n))^2 \mathfrak{R}_d \left\{ \frac{1}{(\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n))^2} \right\} \right] \\
& -e(n) \theta_1(n) \exp(-\bar{u}_1(n)) \left(\frac{y_1(n) - \bar{m}_{F_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{F_1^1}^2(n)} \right)^2 \mathfrak{R}_d \{ \bar{\sigma}_{y_1}(n) \}
\end{aligned} \tag{2.59}$$

e

$$\begin{aligned}
\mathfrak{R}_d \left\{ \nabla_{\underline{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \right\} = & \\
& -\mathfrak{R}_d \{e(n)\} \theta_1(n) \exp(-\underline{u}_1(n)) \left(\frac{y_1(n) - \underline{m}_{F_1^1}(n)}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n)} \right)^2 \underline{\sigma}_{y_1}(n) \\
& -e(n) \mathfrak{R}_d \{ \theta_1(n) \} \exp(-\underline{u}_1(n)) \left(\frac{y_1(n) - \underline{m}_{F_1^1}(n)}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n)} \right)^2 \underline{\sigma}_{y_1}(n) \\
& -e(n) \theta_1(n) \mathfrak{R}_d \{ \exp(-\underline{u}_1(n)) \} \left(\frac{y_1(n) - \underline{m}_{F_1^1}(n)}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n)} \right)^2 \underline{\sigma}_{y_1}(n) \\
& -e(n) \theta_l(n) \exp(-\underline{u}_1(n)) \left[\left(\mathfrak{R}_d \left\{ \left(y_1(n) - \underline{m}_{F_1^1}(n) \right)^2 \right\} \right) \frac{1}{\left(\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n) \right)^2} \right. \\
& \quad \left. + \left(y_1(n) - \underline{m}_{F_1^1}(n) \right)^2 \mathfrak{R}_d \left\{ \frac{1}{\left(\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n) \right)^2} \right\} \right] \\
& -e(n) \theta_1(n) \exp(-\underline{u}_1(n)) \left(\frac{y_1(n) - \underline{m}_{F_1^1}(n)}{\underline{\sigma}_{y_1}^2(n) + \underline{\sigma}_{F_1^1}^2(n)} \right)^2 \mathfrak{R}_d \{ \underline{\sigma}_{y_1}(n) \}.
\end{aligned} \tag{2.60}$$

Finalmente, o produto $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ é expresso por

$$\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n) = \mathfrak{R}_d \{ \nabla J_c(\mathbf{w}(n)) \} = \begin{bmatrix} \mathfrak{R}_d \{ \nabla_{\underline{\theta}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\mathbf{m}}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\sigma}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\mathbf{m}}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\sigma}(n)} J_c(\mathbf{w}(n)) \} \\ \mathfrak{R}_d \{ \nabla_{\underline{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \} \end{bmatrix}. \tag{2.61}$$

A aplicação do método GCEM para o treinamento do algoritmo fuzzy do tipo 2, cuja saída é dada por (2.43), é realizada a partir das equações dos gradientes dos parâmetros do algoritmo fuzzy dadas por (2.47)-(2.53) e do produto $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ dado por (2.61).

2.2.2 Algoritmo Fuzzy Não Normalizado II

Considerando-se que a função Gaussiana modela as incertezas dos antecedentes das regras fuzzy e a função exponencial modela as incertezas presentes nas entradas do sistema fuzzy do tipo 2 e intervalado, pode-se deduzir um algoritmo fuzzy do tipo 2, intervalado, *nonsingleton* e não normalizado da seguinte maneira:

Sejam

$$\bar{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp\left(-\bar{\beta}_{y_i}(n) \left|y_i(n) - \bar{m}_{\tilde{F}_i^l}(n)\right|\right) \quad (2.62)$$

e

$$\underline{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp\left(-\underline{\beta}_{y_i}(n) \left|y_i(n) - \underline{m}_{\tilde{F}_i^l}(n)\right|\right) \quad (2.63)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das funções de pertinência do tipo 2 que modelam as incertezas nas entradas de um sistema fuzzy do tipo 2, intervalado e *nonsingleton*,

$$\bar{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp\left[-\frac{1}{2} \left(\frac{y_i(n) - \bar{m}_{\tilde{F}_i^l}(n)}{\bar{\sigma}_{\tilde{F}_i^l}(n)}\right)^2\right] \quad (2.64)$$

e

$$\underline{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp\left[-\frac{1}{2} \left(\frac{y_i(n) - \underline{m}_{\tilde{F}_i^l}(n)}{\underline{\sigma}_{\tilde{F}_i^l}(n)}\right)^2\right], \quad (2.65)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das funções de pertinência do tipo 2 que modelam as incertezas nas premissas dos antecedentes das regras fuzzy do sistema fuzzy do tipo 2, intervalado e *nonsingleton*. Então, pode-se relacionar as incertezas presentes nas entradas com as incertezas presentes nas premissas dos antecedentes das regras com as expressões dadas por

$$\bar{\mu}_{\tilde{Q}_i^l}(\bar{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \bar{\mu}_{\tilde{Y}_i}(y_i(n)) \star \bar{\mu}_{\tilde{F}_i^l}(y_i(n)) \quad (2.66)$$

como sendo o supremo entre as funções *upper* de pertinência e por

$$\underline{\mu}_{\tilde{Q}_i^l}(\underline{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \underline{\mu}_{\tilde{Y}_i}(y_i(n)) \star \underline{\mu}_{\tilde{F}_i^l}(y_i(n)) \quad (2.67)$$

como sendo o supremo entre as funões *lower* de pertinência. Note que $i = 0, \dots, L-1$ e $l = 1, \dots, M$, Em (2.66) e (2.67), o smbolo \star  uma *t-norm* implementada pelo operador produto e as amostras $y_i(n)$, $i = 0, \dots, L-1$, formam o vetor de entrada do algoritmo fuzzy. Extendendo-se os resultados obtidos em sistemas fuzzy do tipo 1 (Wei & Mendel, 1999) para os sistemas fuzzy do tipo 2, as expresses (2.66)-(2.67) alcanam seus mximos correspondentes nas seguintes situaes

$$\bar{y}_{i,max}^l(n) = \begin{cases} \bar{m}_{\tilde{F}_i^l}(n) - \bar{\beta}_{y_i}(n)\bar{\sigma}_{\tilde{F}_i^l}^2(n), & \text{se } \bar{m}_{\tilde{F}_{my_i}^l}(n) \leq \bar{m}_{\tilde{F}_i^l}(n) - \bar{\beta}_{y_i}(n)\bar{\sigma}_{\tilde{F}_i^l}^2(n) \\ \bar{m}_{\tilde{F}_i^l}(n) + \bar{\beta}_{y_i}(n)\bar{\sigma}_{\tilde{F}_i^l}^2(n), & \text{se } \bar{m}_{\tilde{F}_{my_i}^l}(n) > \bar{m}_{\tilde{F}_i^l}(n) + \bar{\beta}_{y_i}(n)\bar{\sigma}_{\tilde{F}_i^l}^2(n) \\ \bar{m}_{\tilde{F}_{my_i}^l}(n), & \text{caso contrario} \end{cases} \quad (2.68)$$

e

$$\underline{y}_{i,max}^l(n) = \begin{cases} \underline{m}_{\tilde{F}_i^l}(n) - \underline{\beta}_{y_i}(n)\underline{\sigma}_{\tilde{F}_i^l}^2(n), & \text{se } \underline{m}_{\tilde{F}_{my_i}^l}(n) \leq \underline{m}_{\tilde{F}_i^l}(n) - \underline{\beta}_{y_i}(n)\underline{\sigma}_{\tilde{F}_i^l}^2(n) \\ \underline{m}_{\tilde{F}_i^l}(n) + \underline{\beta}_{y_i}(n)\underline{\sigma}_{\tilde{F}_i^l}^2(n), & \text{se } \underline{m}_{\tilde{F}_{my_i}^l}(n) > \underline{m}_{\tilde{F}_i^l}(n) + \underline{\beta}_{y_i}(n)\underline{\sigma}_{\tilde{F}_i^l}^2(n) \\ \underline{m}_{\tilde{F}_{my_i}^l}(n), & \text{caso contrario} \end{cases} \quad (2.69)$$

onde $\bar{y}_{i,max}^l(n)$ e $\underline{y}_{i,max}^l(n)$ so os valores que fornecem o supremo. A justificativa para a escolha da extenso direta dos resultados obtidos em sistemas fuzzy do tipo 1 para sistemas fuzzy do tipo 2  a mesma usada no algoritmo fuzzy do tipo 2, deduzido na Seo 2.2.1.

Agora, supondo-se que as variaveis $y_i(n)$, $i = 0, \dots, L-1$, apresentem os mesmos nveis de incertezas, ou seja, $\bar{\beta}_{y_i} = \bar{\beta}_{y_1}$ e $\underline{\beta}_{y_i} = \underline{\beta}_{y_1}$ (vide resultados para sistemas fuzzy do tipo 1 em (Wei & Mendel, 2000)) e aqui extendida para sistemas fuzzy do tipo 2. Ento, tem-se a sada do algoritmo baseado no sistema fuzzy do tipo 2,

intervalado e *nonsingleton* expressa por

$$\begin{aligned}
f_{ns,2}(\mathbf{y}(n)) &= \sum_{l=1}^M \theta_l \left[\bar{f}_l(\mathbf{y}(n)) + \underline{f}_l(\mathbf{y}(n)) \right] \\
&= \sum_{l=1}^M \theta_l(n) \left[\prod_{i=0}^{L-1} \exp \left(-\underline{\beta}_{y_1}(n) \left| \underline{y}_{i,max}^l(n) - \underline{m}_{\bar{F}_{m_{y_i}}^l}(n) \right| - \frac{1}{2} \frac{\left(\underline{y}_{i,max}^l(n) - \underline{m}_{\bar{F}_i^l}(n) \right)^2}{\underline{\sigma}_{\bar{F}_i^l}^2(n)} \right) \right. \\
&\quad \left. + \prod_{i=0}^{L-1} \exp \left(-\bar{\beta}_{y_1}(n) \left| \bar{y}_{i,max}^l(n) - \bar{m}_{\bar{F}_{m_{y_i}}^l}(n) \right| - \frac{1}{2} \frac{\left(\bar{y}_{i,max}^l(n) - \bar{m}_{\bar{F}_i^l}(n) \right)^2}{\bar{\sigma}_{\bar{F}_i^l}^2(n)} \right) \right], \tag{2.70}
\end{aligned}$$

onde M é o número de regras e $\theta_l(n)$, $\underline{\beta}_{y_1}(n)$, $\underline{m}_{\bar{F}_i^l}(n)$, $\underline{\sigma}_{\bar{F}_i^l}(n)$, $\underline{m}_{\bar{F}_{m_{y_i}}^l}(n)$, $\bar{\beta}_{y_1}(n)$, $\bar{m}_{\bar{F}_i^l}(n)$, $\bar{\sigma}_{\bar{F}_i^l}(n)$ e $\bar{m}_{\bar{F}_{m_{y_i}}^l}(n)$ são os parâmetros do algoritmo fuzzy do tipo 2.

Assumindo-se a função custo a ser minimizada dada por

$$J(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} J_c(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} e^2(n) = \frac{1}{2} \sum_{n=0}^{N-1} [y_d(n) - f_{ns,2}(\mathbf{y}(n))]^2, \tag{2.71}$$

onde onde $y_d(n)$ é a saída desejada, $f_{ns,2}(\mathbf{y}(n))$ é a saída do sistema fuzzy do tipo 2, $\{y_d(n), f_{ns,2}(\mathbf{y}(n))\}_{n=0}^{N-1}$ são os N pares entrada saída utilizados durante o treinamento do algoritmo fuzzy do tipo 2, $\mathbf{w}(n) = \left[\boldsymbol{\theta}^T(n) \ \bar{\boldsymbol{\sigma}}^T(n) \ \bar{\mathbf{m}}^T(n) \ \bar{\mathbf{m}}_{F_{m_y}}^T(n) \ \bar{\beta}_{y_1}(n) \ \boldsymbol{\sigma}^T(n) \ \underline{\mathbf{m}}^T(n) \ \underline{\mathbf{m}}_{F_{m_y}}^T(n) \ \underline{\beta}_{y_1}(n) \right]^T$ é o vetor constituído de todos os parâmetros do algoritmo fuzzy, $y_d(n)$ é o sinal desejado, $\boldsymbol{\theta}(n) = [\theta_1(n) \cdots \theta_M(n)]^T$ é o vetor de parâmetros que pondera a influência de cada regra na saída do algoritmo fuzzy, $\bar{\mathbf{m}}(n) = [\bar{\mathbf{m}}_1^T(n) \cdots \bar{\mathbf{m}}_M^T(n)]^T = [\bar{m}_{\bar{F}_0^1}(n) \cdots \bar{m}_{\bar{F}_{L-1}^1}(n) \cdots \bar{m}_{\bar{F}_0^M}(n) \cdots \bar{m}_{\bar{F}_{L-1}^M}(n)]^T$ é o vetor constituído dos valores $\bar{m}_{\bar{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\bar{F}_i^l}$, $\bar{\mathbf{m}}_{\bar{F}_{m_y}}(n) = [\bar{\mathbf{m}}_{\bar{F}_{m_y}^1}^T(n) \cdots \bar{\mathbf{m}}_{\bar{F}_{m_y}^M}^T(n)]^T = [\bar{m}_{\bar{F}_{m_{y_0}}^l}(n) \cdots \bar{m}_{\bar{F}_{m_{y_{L-1}}}^l}(n) \cdots \bar{m}_{\bar{F}_{m_{y_0}}^M}(n) \cdots \bar{m}_{\bar{F}_{m_{y_{L-1}}}^M}(n)]^T$ é o vetor constituído das médias das funções exponenciais que modelam as incertezas do vetor de entrada, $\bar{\boldsymbol{\sigma}}(n) = [\bar{\boldsymbol{\sigma}}_1^T(n) \cdots \bar{\boldsymbol{\sigma}}_M^T(n)]^T = [\bar{\sigma}_{\bar{F}_0^1}(n) \cdots \bar{\sigma}_{\bar{F}_{L-1}^1}(n) \cdots \bar{\sigma}_{\bar{F}_0^M}(n) \cdots \bar{\sigma}_{\bar{F}_{L-1}^M}(n)]^T$ é o vetor constituído dos valores $\bar{\sigma}_{\bar{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\bar{F}_i^l}$, $\underline{\mathbf{m}}(n) = [\underline{\mathbf{m}}_1^T(n) \cdots \underline{\mathbf{m}}_M^T(n)]^T =$

$\left[\underline{m}_{\tilde{F}_0^1}(n) \cdots \underline{m}_{\tilde{F}_{L-1}^1}(n) \cdots \underline{m}_{\tilde{F}_0^M}(n) \cdots \underline{m}_{\tilde{F}_{L-1}^M}(n) \right]^T$ é o vetor constituído dos valores $\underline{m}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$, $\underline{\mathbf{m}}_{\tilde{F}_{m_y}}(n) = \left[\underline{\mathbf{m}}_{\tilde{F}_{m_y}^1}^T(n) \cdots \underline{\mathbf{m}}_{\tilde{F}_{m_y}^M}^T(n) \right]^T = \left[\underline{m}_{\tilde{F}_{m_y}^1}(n) \cdots \underline{m}_{\tilde{F}_{m_y}^{L-1}}(n) \cdots \underline{m}_{\tilde{F}_{m_y}^M}(n) \cdots \underline{m}_{\tilde{F}_{m_y}^{L-1}}(n) \right]^T$ é o vetor constituído das médias das funções exponenciais que modelam as incertezas do vetor de entrada, $\underline{\sigma}(n) = \left[\underline{\sigma}_1^T(n) \cdots \underline{\sigma}_M^T(n) \right]^T = \left[\underline{\sigma}_{\tilde{F}_0^1}(n) \cdots \underline{\sigma}_{\tilde{F}_{L-1}^1}(n) \cdots \underline{\sigma}_{\tilde{F}_0^M}(n) \cdots \underline{\sigma}_{\tilde{F}_{L-1}^M}(n) \right]^T$ é o vetor constituído dos valores $\bar{\sigma}_{\tilde{F}_i^l}$ das funções de pertinência $\bar{\mu}_{\tilde{F}_i^l}$. Então

$$\begin{aligned} \nabla J_c(\mathbf{w}(n)) = & \left[\nabla_{\boldsymbol{\theta}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\bar{\boldsymbol{\sigma}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\bar{\mathbf{m}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\bar{\mathbf{m}}_{\tilde{F}_{m_y}}} J_c^T(\mathbf{w}(n)) \right. \\ & \left. \nabla_{\bar{\beta}_{y_1}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\boldsymbol{\sigma}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\mathbf{m}}(n)} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\mathbf{m}}_{\tilde{F}_{m_y}}} J_c^T(\mathbf{w}(n)) \quad \nabla_{\underline{\beta}_{y_1}(n)} J_c^T(\mathbf{w}(n)) \right]^T, \end{aligned} \quad (2.72)$$

Os vetores gradientes dos parâmetros do algoritmo fuzzy do tipo 2, cuja saída é expressa por (2.70), são listados a seguir:

$$\nabla_{\boldsymbol{\theta}(n)} J_c(\mathbf{w}(n)) = e(n) \mathbf{z}(n), \quad (2.73)$$

onde $\mathbf{z}(n) = \left[\bar{f}_1(\mathbf{y}(n)) + \underline{f}_1(\mathbf{y}(n)) \cdots \bar{f}_M(\mathbf{y}(n)) + \underline{f}_M(\mathbf{y}(n)) \right]^T$,

$$\nabla_{\bar{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) = -e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n)}{\bar{\sigma}_l^2(n)} \right), \quad (2.74)$$

$$\nabla_{\underline{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) = -e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\underline{y}_{max}^l(n) - \underline{\mathbf{m}}_l(n)}{\underline{\sigma}_l^2(n)} \right), \quad (2.75)$$

$$\nabla_{\bar{\sigma}_l(n)} J_c(\mathbf{w}(n)) = -e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \frac{(\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2}{\bar{\sigma}_l^3(n)}, \quad (2.76)$$

$$\nabla_{\underline{\sigma}_l(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_l(n) \odot \exp(-\underline{\mathbf{u}}^l(n)) \odot \frac{(\underline{y}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2}{\underline{\sigma}_l^3(n)}, \quad (2.77)$$

$$\nabla_{\overline{\mathbf{m}}_{F_{m_y}^l}}(n) J_c(\mathbf{w}(n)) = e(n)\theta_l(n) \odot \exp(-\overline{\mathbf{u}}_l(n)) \odot \overline{\beta}_{y_1}(n) \overline{\Psi}_l(n), \quad (2.78)$$

onde

$$\overline{\Psi}_l(n) = \begin{bmatrix} \overline{\mathcal{K}}_{l,0}(n) \\ \vdots \\ \overline{\mathcal{K}}_{l,L-1}(n) \end{bmatrix} \quad (2.79)$$

e

$$\overline{\mathcal{K}}_{l,i}(n) = \begin{cases} 1, & \text{se } \overline{y}_{i,max}^l(n) - \overline{m}_{F_{m_{y_i}}^l}(n) \geq 0 \\ -1, & \text{caso contrário} \end{cases}, i = 0, \dots, L-1, \quad (2.80)$$

$$\nabla_{\underline{\mathbf{m}}_{F_{m_y}^l}}(n) J_c(\mathbf{w}(n)) = e(n)\theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \underline{\beta}_{y_1}(n) \underline{\Psi}_l(n), \quad (2.81)$$

$$\underline{\Psi}_l(n) = \begin{bmatrix} \underline{\mathcal{K}}_{l,0}(n) \\ \vdots \\ \underline{\mathcal{K}}_{l,L-1}(n) \end{bmatrix} \quad (2.82)$$

e

$$\underline{\mathcal{K}}_{l,i}(n) = \begin{cases} 1, & \text{se } \underline{y}_{i,max}^l(n) - \underline{m}_{F_{m_{y_i}}^l}(n) \geq 0 \\ -1, & \text{caso contrário} \end{cases}, i = 0, \dots, L-1, \quad (2.83)$$

e $l = 1 \cdots M$,

$$\nabla_{\overline{\beta}_{y_1}(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_1(n) \exp(-\overline{u}_1^1(n)) |\overline{y}_{1,max}^1(n) - y_1(n)|, \quad (2.84)$$

e

$$\nabla_{\underline{\beta}_{y_1}(n)} J_c(\mathbf{w}(n)) = -e(n)\theta_1(n) \exp(-\underline{u}_1^l(n)) \left| \underline{y}_{1,max}^l(n) - y_1(n) \right|, \quad (2.85)$$

$$\begin{aligned} \text{onde } \bar{\mathbf{u}}^l(n) &= [\bar{u}_0^l(n) \cdots \bar{u}_{L-1}^l(n)]^T, & \mathbf{u}^l(n) &= [\underline{u}_0^l(n) \cdots \underline{u}_{L-1}^l(n)]^T, \\ \bar{u}_i^l(n) &= \bar{\beta}_{y_1}(n) |\bar{y}_{i,max}^l(n) - y_i(n)| + \frac{1}{2} \frac{(\bar{y}_{i,max}^l(n) - \bar{m}_{F_i^l}(n))^2}{\bar{\sigma}_{F_i^l}^2(n)}, & \underline{u}_i^l(n) &= \\ \beta_{y_1}(n) |\underline{y}_{i,max}^l(n) - y_i(n)| + \frac{1}{2} \frac{(y_{i,max}^l(n) - m_{F_i^l}(n))^2}{\sigma_{F_i^l}^2(n)}, & \bar{y}_{max}^l(n) &= \\ [\bar{y}_{0,max}^l(n) \cdots \bar{y}_{L-1,max}^l(n)]^T & \text{e } \underline{\mathbf{y}}_{max}^l(n) = [\underline{y}_{0,max}^l(n) \cdots \underline{y}_{L-1,max}^l(n)]^T. \end{aligned}$$

Para a aplicação do método de otimização GCEM para o treinamento do algoritmo fuzzy, cuja saída é dada por (2.70), faz-se uso do mesmo procedimento discutido na Seção 2.2.1 para a obtenção do produto da matriz Hessiana por um vetor direcional $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$. Desta maneira, aplicando-se o operador diferencial expresso por (2.34) em (2.73)-(2.85), tem-se

$$\Re_{\mathbf{d}} \{ \nabla_{\theta(n)} J_c(\mathbf{w}(n)) \} = \Re_{\mathbf{d}} \{ e(n) \} \mathbf{z}(n) + e(n) \Re_{\mathbf{d}} \{ \mathbf{z}(n) \}, \quad (2.86)$$

$$\begin{aligned} \Re_{\mathbf{d}} \{ \nabla_{\bar{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) \} &= -\Re_{\mathbf{d}} \{ e(n) \} \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n)}{\bar{\sigma}_l^2(n)} \right) \\ &- e(n) \Re_{\mathbf{d}} \{ \theta_l(n) \} \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left(\frac{\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n)}{\bar{\sigma}_l^2(n)} \right) \\ &- e(n) \theta_l(n) \odot \Re_{\mathbf{d}} \{ \exp(-\bar{\mathbf{u}}_l(n)) \} \odot \left(\frac{\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n)}{\bar{\sigma}_l^2(n)} \right) \\ &- e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left[\left(\Re_{\mathbf{d}} \{ \bar{y}_{max}^l(n) \} - \Re_{\mathbf{d}} \{ \bar{\mathbf{m}}_l(n) \} \right) \odot \frac{1}{\bar{\sigma}_l^2(n)} + \right. \\ &\quad \left. (\bar{y}_{max}^l(n) - \bar{\mathbf{m}}_l(n)) \odot \Re_{\mathbf{d}} \left\{ \frac{1}{\bar{\sigma}_l^2(n)} \right\} \right], \end{aligned} \quad (2.87)$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\underline{\mathbf{m}}_l(n)} J_c(\mathbf{w}(n)) \} &= -\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n)}{\underline{\sigma}_l^2(n)} \right) \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left(\frac{\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n)}{\underline{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\underline{\mathbf{u}}_l(n)) \} \odot \left(\frac{\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n)}{\underline{\sigma}_l^2(n)} \right) \\
&- e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left[\left(\mathfrak{R}_d \left\{ \underline{\mathbf{y}}_{max}^l(n) \right\} - \mathfrak{R}_d \left\{ \underline{\mathbf{m}}_l(n) \right\} \right) \odot \frac{1}{\underline{\sigma}_l^2(n)} + \right. \\
&\quad \left. \left(\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n) \right) \odot \mathfrak{R}_d \left\{ \frac{1}{\underline{\sigma}_l^2(n)} \right\} \right], \tag{2.88}
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\bar{\sigma}_l(n)} J_c(\mathbf{w}(n)) \} &= \\
&-\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \frac{(\bar{\mathbf{y}}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2}{\bar{\sigma}_l^3(n)} \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \frac{(\bar{\mathbf{y}}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2}{\bar{\sigma}_l^3(n)} \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\bar{\mathbf{u}}_l(n)) \} \odot \frac{(\bar{\mathbf{y}}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2}{\bar{\sigma}_l^3(n)} \\
&- e(n) \theta_l(n) \odot \exp(-\bar{\mathbf{u}}_l(n)) \odot \left[\left(\mathfrak{R}_d \left\{ (\bar{\mathbf{y}}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2 \right\} \right) \odot \frac{1}{\bar{\sigma}_l^3(n)} + \right. \\
&\quad \left. (\bar{\mathbf{y}}_{max}^l(n) - \bar{\mathbf{m}}_l(n))^2 \odot \mathfrak{R}_d \left\{ \frac{1}{\bar{\sigma}_l^3(n)} \right\} \right], \tag{2.89}
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_d \{ \nabla_{\underline{\sigma}_l(n)} J_c(\mathbf{w}(n)) \} &= \\
&-\mathfrak{R}_d \{ e(n) \} \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \frac{(\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2}{\underline{\sigma}_l^3(n)} \\
&- e(n) \mathfrak{R}_d \{ \theta_l(n) \} \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \frac{(\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2}{\underline{\sigma}_l^3(n)} \\
&- e(n) \theta_l(n) \odot \mathfrak{R}_d \{ \exp(-\underline{\mathbf{u}}_l(n)) \} \odot \frac{(\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2}{\underline{\sigma}_l^3(n)} \\
&- e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \left[\left(\mathfrak{R}_d \left\{ (\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2 \right\} \right) \odot \frac{1}{\underline{\sigma}_l^3(n)} + \right. \\
&\quad \left. (\underline{\mathbf{y}}_{max}^l(n) - \underline{\mathbf{m}}_l(n))^2 \odot \mathfrak{R}_d \left\{ \frac{1}{\underline{\sigma}_l^3(n)} \right\} \right], \tag{2.90}
\end{aligned}$$

$$\begin{aligned}
\mathfrak{R}_d \left\{ \nabla_{\overline{m}_{F_{m_y}^l}}(n) J_c(\mathbf{w}(n)) \right\} &= \mathfrak{R}_d \{e(n)\} \theta_l(n) \odot \exp(-\overline{\mathbf{u}}_l(n)) \odot \overline{\beta}_{y_1}(n) \overline{\Psi}_l(n) \\
&\quad + e(n) \mathfrak{R}_d \{\theta_l(n)\} \odot \exp(-\overline{\mathbf{u}}_l(n)) \odot \overline{\beta}_{y_1}(n) \overline{\Psi}_l(n) \\
&\quad + e(n) \theta_l(n) \odot \mathfrak{R}_d \{\exp(-\overline{\mathbf{u}}_l(n))\} \odot \overline{\beta}_{y_1}(n) \overline{\Psi}_l(n) \quad , \\
&\quad + e(n) \theta_l(n) \odot \exp(-\overline{\mathbf{u}}_l(n)) \odot \mathfrak{R}_d \{\overline{\beta}_{y_1}(n)\} \overline{\Psi}_l(n) \\
&\quad + e(n) \theta_l(n) \odot \exp(-\overline{\mathbf{u}}_l(n)) \odot \overline{\beta}_{y_1}(n) \mathfrak{R}_d \{\overline{\Psi}_l(n)\}
\end{aligned} \tag{2.91}$$

$$\begin{aligned}
\mathfrak{R}_d \left\{ \nabla_{\underline{m}_{F_{m_y}^l}}(n) J_c(\mathbf{w}(n)) \right\} &= \mathfrak{R}_d \{e(n)\} \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \underline{\beta}_{y_1}(n) \underline{\Psi}_l(n) \\
&\quad + e(n) \mathfrak{R}_d \{\theta_l(n)\} \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \underline{\beta}_{y_1}(n) \underline{\Psi}_l(n) \\
&\quad + e(n) \theta_l(n) \odot \mathfrak{R}_d \{\exp(-\underline{\mathbf{u}}_l(n))\} \odot \underline{\beta}_{y_1}(n) \underline{\Psi}_l(n) \quad , \\
&\quad + e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \mathfrak{R}_d \{\underline{\beta}_{y_1}(n)\} \underline{\Psi}_l(n) \\
&\quad + e(n) \theta_l(n) \odot \exp(-\underline{\mathbf{u}}_l(n)) \odot \underline{\beta}_{y_1}(n) \mathfrak{R}_d \{\underline{\Psi}_l(n)\}
\end{aligned} \tag{2.92}$$

$$\begin{aligned}
\mathfrak{R}_d \left\{ \nabla_{\overline{\beta}_{y_1}(n)} J_c(\mathbf{w}(n)) \right\} &= -\mathfrak{R}_d \{e(n)\} \theta_1(n) \exp(-\overline{u}_1(n)) \left| \overline{y}_{1,max}^1(n) - \overline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \mathfrak{R}_d \{\theta_1(n)\} \exp(-\overline{u}_1(n)) \left| \overline{y}_{1,max}^1(n) - \overline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \theta_1(n) \mathfrak{R}_d \{\exp(-\overline{u}_1(n))\} \left| \overline{y}_{1,max}^1(n) - \overline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \theta_1(n) \exp(-\overline{u}_1(n)) \mathfrak{R}_d \left\{ \left| \overline{y}_{1,max}^1(n) - \overline{m}_{F_{m_{y_1}}^1}(n) \right| \right\}
\end{aligned} \tag{2.93}$$

e

$$\begin{aligned}
\mathfrak{R}_d \left\{ \nabla_{\underline{\beta}_{y_1}(n)} J_c(\mathbf{w}(n)) \right\} &= -\mathfrak{R}_d \{e(n)\} \theta_1(n) \exp(-\underline{u}_1(n)) \left| \underline{y}_{1,max}^1(n) - \underline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \mathfrak{R}_d \{\theta_1(n)\} \exp(-\underline{u}_1(n)) \left| \underline{y}_{1,max}^1(n) - \underline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \theta_1(n) \mathfrak{R}_d \{\exp(-\underline{u}_1(n))\} \left| \underline{y}_{1,max}^1(n) - \underline{m}_{F_{m_{y_1}}^1}(n) \right| \\
&\quad - e(n) \theta_1(n) \exp(-\underline{u}_1(n)) \mathfrak{R}_d \left\{ \left| \underline{y}_{1,max}^1(n) - \underline{m}_{F_{m_{y_1}}^1}(n) \right| \right\} .
\end{aligned} \tag{2.94}$$

Finalmente, tem-se que o produto entre a matriz Hessiana $\mathbf{H}(\mathbf{w}(n))$ e um vetor direcional $\mathbf{d}(n)$ é expresso por

$$\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n) = \Re_{\mathbf{d}}\{\nabla J_c(\mathbf{w}(n))\} = \begin{bmatrix} \Re_{\mathbf{d}}\{\nabla_{\boldsymbol{\theta}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\mathbf{m}}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\boldsymbol{\sigma}}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\left\{\nabla_{\bar{\mathbf{m}}_{F_{m_y}}(n)} J_c(\mathbf{w}(n))\right\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\beta}_{y_1}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\mathbf{m}}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\boldsymbol{\sigma}}(n)} J_c(\mathbf{w}(n))\} \\ \Re_{\mathbf{d}}\left\{\nabla_{\bar{\mathbf{m}}_{F_{m_y}}(n)} J_c(\mathbf{w}(n))\right\} \\ \Re_{\mathbf{d}}\{\nabla_{\bar{\beta}_{y_1}(n)} J_c(\mathbf{w}(n))\} \end{bmatrix} \quad (2.95)$$

O método GCEM pode ser aplicado para o treinamento do algoritmo fuzzy do tipo 2, cuja saída é dada por (2.70), a partir das equações dos gradientes dos parâmetros do algoritmo fuzzy dadas por (2.73)-(2.85) e do produto $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ dado por (2.95).

2.3 Algoritmos fuzzy do Tipo 2, Intervalados, *Nonsingleton*, Normalizados e Adaptativos

Baseando-se na expressão (2.28), os algoritmos fuzzy adaptativos utilizados para a equalização de canais PLC no capítulo 5 são introduzidos nesta Seção. Além disso, é discutido o uso de técnicas de atualização do passo de adaptação dos parâmetros de tais algoritmos para que os mesmos possam apresentar velocidades de convergência maiores.

Assim na Seção 2.3.1 descreve-se, primeiramente, o algoritmo fuzzy do tipo 1,

nonsingleton e adaptativo, proposto por Mendel e seus alunos (Wang & Mendel, 1993; Mendel, 2001), e, em seguida, baseando-se na expressão (2.28), deriva-se um algoritmo fuzzy do tipo 2, *nonsingleton*, intervalado e adaptativo que pode ser visto como uma extensão do algoritmo fuzzy do tipo 1, *nonsingleton* e adaptativo proposto por Mendel e seus alunos. A seguir, na Seção 2.3.2 introduz-se o algoritmo fuzzy do tipo 1, *nonsingleton* e adaptativo, que faz uso das função Gaussiana para modelar as incertezas nas suas regras e a função exponencial para modelar as incertezas nas entradas. Em seguida, baseando-se em (2.28), um novo algoritmo fuzzy do tipo 2, intervalado, *nonsingleton* e adaptativo é introduzido. Pode-se observar que este algoritmo fuzzy do tipo 2 é uma extensão do algoritmo fuzzy do tipo 1 deduzido nesta Seção. O uso da função exponencial para modelar a incerteza na entrada do sistema fuzzy foi baseado em (Wei & Mendel, 2000). Finalmente, na Seção 2.3.3 discute-se o uso de duas técnicas para a atualização do passo de adaptação dos algoritmos fuzzy que promovem o aumento da velocidade de convergência dos mesmos.

2.3.1 Algoritmo Fuzzy do Tipo 2 Adaptativo I

Nesta Seção um algoritmo fuzzy do tipo 2, intervalado, *nonsingleton* e adaptativo, derivado a partir da expressão (2.28), é descrito. Para a descrição deste algoritmo, inicialmente, um algoritmo fuzzy do tipo 1 e *nonsingleton* proposto por Mendel e seus alunos é reproduzido e, na seqüência, o algoritmo fuzzy do tipo 2 proposto é descrito. Tal algoritmo fuzzy do tipo 2 e adaptativo foi inspirado em (Wei & Mendel, 2000). A motivação para a apresentação do algoritmo proposto desta maneira reside no fato de que o mesmo é uma extensão direta do algoritmo fuzzy do tipo 1 reproduzido abaixo:

Sejam

$$\mu_{Y_i}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - m_{y_i}(n)}{\sigma_{y_i}(n)} \right)^2 \right] \quad (2.96)$$

e

$$\mu_{F_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - m_{F_i^l}(n)}{\sigma_{F_i^l}(n)} \right)^2 \right] \quad (2.97)$$

as funções de pertinência que modelam as incertezas nas entradas do algoritmo fuzzy e as incertezas associadas às regras do sistema fuzzy, tem-se que o sistema fuzzy é do tipo 1 e *nonsingleton*. Conseqüentemente, a ativação da i -ésima premissa do antecedente da l -ésima regra fuzzy é dada por

$$\mu_{Q_i^l}(y_{i,\max}^l(n)) = \sup_{y_i(n) \in Y_i} \mu_{Y_i}(y_i(n)) \star \mu_{F_i^l}(y_i(n)), \quad (2.98)$$

onde \star é uma t -norm implementada pelo operador produto, $y_i(n)$ a i -ésima amostra do vetor de entrada $\mathbf{y}(n) = [y_0(n) \cdots y_{L-1}(n)]^T$. O valor da expressão (2.98) é atingido na seguinte situação (Mendel, 2001):

$$y_{i,\max}^l(n) = \frac{\sigma_{y_i}^2(n)m_{F_i^l}(n) + \sigma_{F_i^l}^2(n)m_{y_i}(n)}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)}. \quad (2.99)$$

Fazendo $m_{y_i(n)} = y_i(n)$ (Mendel, 2001), a saída do algoritmo fuzzy *nonsingleton* do tipo 1 é dada por

$$f_{ns,1}(\mathbf{y}(n)) = \frac{\sum_{l=1}^M \theta_l(n) \prod_{i=0}^{L-1} \mu_{Q_i^l}(y_{i,\max}^l(n))}{\sum_{l=1}^M \prod_{i=0}^{L-1} \mu_{Q_i^l}(y_{i,\max}^l(n))} \quad (2.100)$$

e

$$\mu_{Q_i^l}(y_{i,\max}^l(n)) = \exp \left[-\frac{1}{2} \frac{\left(y_i(n) - m_{F_i^l}(n) \right)^2}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)} \right]. \quad (2.101)$$

De acordo com (Mendel, 2001), as expressões usadas para a atualização dos parâmetros deste algoritmo fuzzy são as seguintes:

Seja a função custo dada por

$$J_c(\mathbf{w}(n)) = \frac{1}{2} [y_d(n) - f_{ns,1}(\mathbf{y}(n))]^2 = \frac{1}{2} e^2(n), \quad n = 0, \dots, N-1, \quad (2.102)$$

onde $y_d(n)$ é a saída desejada, então, as equações de atualização dos parâmetros do algoritmo fuzzy, cuja saída é dada por (2.100), a partir do método *backpropagation* (BP) (Rumelhart, Hinton, & Williams, 1986), são dadas por

$$\theta_l(n+1) = \theta_l(n) - \alpha_{\theta_l}(n) \nabla_{\theta_l(n)} J_c(\mathbf{w}(n)) = \theta_l(n) + \alpha_{\theta_l}(n) e(n) \frac{a^l(n)}{b(n)} \quad (2.103)$$

$$\begin{aligned} m_{F_i^l}(n+1) &= m_{F_i^l}(n) - \alpha_{m_{F_i^l}}(n) \nabla_{m_{F_i^l}(n)} J_c(\mathbf{w}(n)) \\ &= m_{F_i^l}(n) + \alpha_{m_{F_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) \left(\frac{y_i(n) - m_{F_i^l}(n)}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)} \right), \end{aligned} \quad (2.104)$$

$$\begin{aligned} \sigma_{F_i^l}(n) &= \sigma_{F_i^l}(n) - \alpha_{\sigma_{F_i^l}}(n) \nabla_{\sigma_{F_i^l}(n)} J_c(\mathbf{w}(n)) \\ &= \sigma_{F_i^l}(n) + \alpha_{\sigma_{F_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) \left(\frac{y_i(n) - m_{F_i^l}(n)}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)} \right)^2 \sigma_{F_i^l}(n) \end{aligned} \quad (2.105)$$

e

$$\begin{aligned} \sigma_{y_1}(n) &= \sigma_{y_1}(n) - \alpha_{\sigma_{y_1}}(n) \nabla_{\sigma_{y_1}(n)} J_c(\mathbf{w}(n)) \\ &= \sigma_{y_1}(n) + \alpha_{\sigma_{y_1}}(n) e(n) \left[\frac{\theta_1(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^1(n) \left(\frac{y_1(n) - m_{F_1^1}(n)}{\sigma_{y_1}^2(n) + \sigma_{F_1^1}^2(n)} \right)^2 \sigma_{y_1}(n), \end{aligned} \quad (2.106)$$

onde $l = 1, \dots, M$, $i = 0, \dots, L-1$. Normalmente, considera-se $y_i(n)$ com o mesmo nível de incerteza, ou seja, $\sigma_{y_i}(n) = \sigma_{y_1}(n)$. $\alpha_{\theta_l}(n)$, $\alpha_{m_{F_i^l}}(n)$, $\alpha_{\sigma_{F_i^l}}(n)$ e $\alpha_{\sigma_{y_1}}(n)$ são os passos de adaptação dos parâmetros do algoritmo fuzzy e

$$a^l(n) = \prod_{i=0}^{L-1} \exp \left[-\frac{1}{2} \frac{\left(y_i(n) - m_{F_i^l}(n) \right)^2}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)} \right] \quad (2.107)$$

e

$$b(n) = \sum_{l=1}^M a^l(n) = \sum_{l=1}^M \prod_{i=0}^{L-1} \exp \left[-\frac{1}{2} \frac{\left(y_i(n) - m_{F_i^l}(n) \right)^2}{\sigma_{y_i}^2(n) + \sigma_{F_i^l}^2(n)} \right]. \quad (2.108)$$

Abaixo deduz-se um algoritmo novo quando várias restrições são impostas ao sistema fuzzy do tipo 2 e intervalado. Conforme ficará claro no equacionamento do mesmo, as expressões que fornecem a saída e permitem a atualização deste novo algoritmo fuzzy do tipo 2 podem ser vistas como versões extendidas do algoritmo fuzzy do tipo 1, acima descrito.

A dedução deste novo algoritmo fuzzy do tipo 2 é baseada não só nas restrições impostas para a obtenção da expressão (2.28), mas também nas expressões (2.35)-(2.38) e nas restrições aplicadas para a obtenção dos supremos das expressões (2.39)-(2.40) para a proposição do algoritmo fuzzy do tipo 2, introduzido na Seção 2.2.1.

Com base nestas considerações, a saída deste novo algoritmo fuzzy que é do tipo 2, intervalado, *nonsingleton*, normalizado e adaptativo é dada pelas expressões

$$f_{ns,2}(\mathbf{y}(n)) = \frac{\sum_{l=1}^M \theta_l a_{ns,2}^l(n)}{b_{ns,2}(n)}, \quad (2.109)$$

$$a_{ns,2}^l(n) = \bar{a}_{ns,2}^l(n) + \underline{a}_{ns,2}^l(n), \quad (2.110)$$

$$\bar{a}_{ns,2}^l(n) = \prod_{i=0}^{L-1} \exp \left[-\frac{1}{2} \frac{\left(y_i(n) - \bar{m}_{F_i^l}(n) \right)^2}{\sigma_{y_1}^2(n) + \sigma_{F_i^l}^2(n)} \right], \quad (2.111)$$

$$\underline{a}_{ns,2}^l(n) = \prod_{i=0}^{L-1} \exp \left[-\frac{1}{2} \frac{\left(y_i(n) - \underline{m}_{F_i^l}(n) \right)^2}{\sigma_{y_1}^2(n) + \sigma_{F_i^l}^2(n)} \right], \quad (2.112)$$

e

$$b_{ns,2}(n) = \sum_{l=1}^M a_{ns,2}^l(n). \quad (2.113)$$

Supondo-se que a função custo a ser minimizada durante o processo de atualização dos parâmetros deste algoritmo fuzzy do tipo 2 seja dada por

$$J_c(\mathbf{w}(n)) = \frac{1}{2} [y_d(n) - f_{ns,2}(\mathbf{y}(n))]^2 = \frac{1}{2} e^2(n), \quad n = 0, \dots, N - 1, \quad (2.114)$$

onde $y_d(n)$ é a saída desejada, e aplicando-se o método BP, tem-se que as equações de atualização dos parâmetros deste algoritmo são as seguintes:

$$\theta_l(n+1) = \theta_l(n) - \alpha_{\theta_l}(n) \nabla_{\theta_l(n)} J_c(\mathbf{w}(n)) = \theta_l(n) + \alpha_{\theta_l}(n) e(n) \frac{a_{ns,2}^l(n)}{b_{ns,2}(n)} \quad (2.115)$$

$$\begin{aligned} \underline{m}_{\hat{F}_i^l}(n+1) &= \underline{m}_{\hat{F}_i^l}(n) - \alpha_{\underline{m}_{\hat{F}_i^l}}(n) \nabla_{\underline{m}_{\hat{F}_i^l}(n)} J_c(\mathbf{w}(n)) \\ &= \underline{m}_{\hat{F}_i^l}(n) + \alpha_{\underline{m}_{\hat{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \underline{a}_{ns,2}^l(n) \left(\frac{y_i(n) - \underline{m}_{\hat{F}_i^l}(n)}{\sigma^2(n) + \sigma_{\hat{F}_i^l}^2(n)} \right), \end{aligned} \quad (2.116)$$

$$\begin{aligned} \overline{m}_{\hat{F}_i^l}(n+1) &= \overline{m}_{\hat{F}_i^l}(n) - \alpha_{\overline{m}_{\hat{F}_i^l}}(n) \nabla_{\overline{m}_{\hat{F}_i^l}(n)} J_c(\mathbf{w}(n)) \\ &= \overline{m}_{\hat{F}_i^l}(n) + \alpha_{\overline{m}_{\hat{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \overline{a}_{ns,2}^l(n) \left(\frac{y_i(n) - \overline{m}_{\hat{F}_i^l}(n)}{\sigma^2(n) + \sigma_{\hat{F}_i^l}^2(n)} \right), \end{aligned} \quad (2.117)$$

$$\begin{aligned} \underline{\sigma}_{\hat{F}_i^l}(n+1) &= \underline{\sigma}_{\hat{F}_i^l}(n) - \alpha_{\underline{\sigma}_{\hat{F}_i^l}}(n) \nabla_{\underline{\sigma}_{\hat{F}_i^l}^2(n)} J_c(\mathbf{w}(n)) \\ &= \underline{\sigma}_{\hat{F}_i^l}(n) + \alpha_{\underline{\sigma}_{\hat{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \underline{a}_{ns,2}^l(n) \left(\frac{y_i(n) - \underline{m}_{\hat{F}_i^l}(n)}{\sigma^2(n) + \sigma_{\hat{F}_i^l}^2(n)} \right)^2 \underline{\sigma}_{\hat{F}_i^l}(n) \end{aligned} \quad (2.118)$$

$$\begin{aligned} \overline{\sigma}_{\hat{F}_i^l}(n+1) &= \overline{\sigma}_{\hat{F}_i^l}(n) - \alpha_{\overline{\sigma}_{\hat{F}_i^l}}(n) \nabla_{\overline{\sigma}_{\hat{F}_i^l}^2(n)} J_c(\mathbf{w}(n)) \\ &= \overline{\sigma}_{\hat{F}_i^l}(n) + \alpha_{\overline{\sigma}_{\hat{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \overline{a}_{ns,2}^l(n) \left(\frac{y_i(n) - \overline{m}_{\hat{F}_i^l}(n)}{\sigma^2(n) + \sigma_{\hat{F}_i^l}^2(n)} \right)^2 \overline{\sigma}_{\hat{F}_i^l}(n) \end{aligned} \quad (2.119)$$

$$\begin{aligned} \underline{\sigma}_{y_1}(n) &= \underline{\sigma}_{y_1}(n) - \alpha_{\underline{\sigma}_{y_1}}(n) \nabla_{\underline{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \\ &= \underline{\sigma}_{y_1}(n) + \alpha_{\underline{\sigma}_{y_1}}(n) e(n) \left[\frac{\theta_1(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \underline{a}_{ns,2}^1(n) \left(\frac{y_1(n) - \underline{m}_{\hat{F}_1^1}(n)}{\sigma_{y_1}^2(n) + \sigma_{\hat{F}_1^1}^2(n)} \right)^2 \underline{\sigma}_{y_1}(n), \end{aligned} \quad (2.120)$$

e

$$\begin{aligned} \bar{\sigma}_{y_1}(n) &= \bar{\sigma}_{y_1}(n) - \alpha_{\bar{\sigma}_{y_1}}(n) \nabla_{\bar{\sigma}_{y_1}(n)} J_c(\mathbf{w}(n)) \\ &= \bar{\sigma}_{y_1}(n) + \alpha_{\bar{\sigma}_{y_1}}(n) e(n) \left[\frac{\theta_1(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] a_{ns,2}^1(n) \left(\frac{y_1(n) - \bar{m}_{\bar{F}_1^1}(n)}{\bar{\sigma}_{y_1}^2(n) + \bar{\sigma}_{\bar{F}_1^1}^2(n)} \right)^2 \bar{\sigma}_{y_1}(n), \end{aligned} \quad (2.121)$$

onde $l = 1, \dots, M$, $i = 0, \dots, L-1$. $\underline{\alpha}_{\theta_l}(n)$, $\bar{\alpha}_{\theta_l}(n)$, $\underline{\alpha}_{m_{\bar{F}_i^l}}(n)$, $\bar{\alpha}_{m_{\bar{F}_i^l}}(n)$, $\underline{\alpha}_{\sigma_{\bar{F}_i^l}}(n)$, $\bar{\alpha}_{\sigma_{\bar{F}_i^l}}(n)$, $\underline{\alpha}_{\sigma_{y_1}}(n)$ e $\bar{\alpha}_{\sigma_{y_1}}(n)$ são os passos de adaptação dos parâmetros do algoritmo fuzzy do tipo 2, cujas entradas apresentam incertezas modeladas pela função Gaussiana de pertinência.

2.3.2 Algoritmo Fuzzy do Tipo 2 Adaptativo II

A literatura mostra que os sistemas fuzzy *nonsingleton* têm sido timidamente aplicados na solução de problemas de engenharia, apesar de apresentarem melhores desempenhos do que os sistemas fuzzy *singleton*, para aplicações em que os dados de treinamento e/ou as futuras entradas do sistema fuzzy são corrompidos pela presença de ruídos aditivos. Tal uso limitado dos sistemas fuzzy *nonsingleton* deve-se ao fato de que a obtenção da expressão que relaciona as funções de pertinência das entradas e das premissas dos antecedentes das regras fuzzy, geralmente, não apresentam uma expressão fechada. No entanto, sabe-se que esta relação pode ser explicitada quando as funções de pertinência que modelam as incertezas nas entradas do sistema fuzzy e as incertezas nas premissas das regras fuzzy são funções Gaussianas, conforme mostrado em (Mouzouris & Mendel, 1997).

Sabe-se que o uso de uma função Gaussiana para modelar as incertezas geradas pela presença de ruídos aditivos nas entradas dos sistemas fuzzy representa a modelagem *worst case*, pois, normalmente, estes ruídos aditivos não são estacionários e não são gaussianos. Baseando-se nesta observação, em (Wei & Mendel, 2000) foi proposto o uso da função exponencial para modelar as incertezas geradas

pela presença de ruídos aditivos nas entradas dos sistemas fuzzy do tipo 1. Tal proposição permitiu a obtenção de resultados bastante satisfatórios para o problema de reconhecimento do tipo de técnica de modulação empregada na transmissão digital de dados através de um canal de comunicação. Apesar desta melhoria, nenhum avanço foi proposto visando a aplicação desta modelagem para a proposição de algoritmos fuzzy do tipo 1 e do tipo 2 adaptativos.

Objetivando propor algoritmos fuzzy adaptativos que levem em consideração a função exponencial como função de pertinência para a modelagem das incertezas nas entradas dos sistemas fuzzy adaptativos e a função Gaussiana para modelar as premissas dos antecedentes das regras fuzzy, nesta Seção, primeiramente, deduz-se um algoritmo fuzzy do tipo 1 e *nonsingleton* que leva em consideração estas funções de pertinência. A seguir, baseando-se nestas mesmas funções de pertinência, um novo algoritmo fuzzy do tipo 2, intervalado e *nonsingleton* é introduzido.

Sejam

$$\mu_{Y_i}(y_i(n)) = \exp \left[-\beta_{y_i}(n) \left| y_i(n) - m_{F_{m_{y_i}}^l}(n) \right| \right], \quad (2.122)$$

$$\mu_{F_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - m_{F_i^l}(n)}{\sigma_{F_i^l}(n)} \right)^2 \right] \quad (2.123)$$

as funções de pertinência que modelam as incertezas presentes nas entradas do sistema fuzzy do tipo 1 e nas premissas das regras fuzzy, respectivamente, então a expressão

$$\mu_{Q_i^l}(y_{i,\max}^l(n)) = \sup_{y_i(n) \in Y_i} \mu_{Y_i}(y_i(n)) \star \mu_{F_i^l}(y_i(n)), \quad (2.124)$$

onde \star é uma *t-norm* implementada pelo operador produto e $y_i(n)$ constitui o vetor de entrada $\mathbf{y}(n) = [y_0(n) \cdots y_{L-1}(n)]^T$, fornece a relação entre as funções de pertinência dadas por (2.122) e (2.123). De acordo com Wei & Mendel (Wei & Mendel, 1999), a expressão (2.124) alcança o seu máximo nas seguintes situações:

$$y_{i,max}^l(n) = \begin{cases} m_{F_i^l}(n) - \beta_{y_i}(n)\sigma_{F_i^l}^2(n), & \text{se } m_{y_i}(n) \leq m_{F_i^l}(n) - \beta_{y_i}(n)\sigma_{F_i^l}^2(n) \\ m_{F_i^l}(n) + \beta_{y_i}(n)\sigma_{F_i^l}^2(n), & \text{se } m_{y_i}(n) > m_{F_i^l}(n) + \beta_{y_i}(n)\sigma_{F_i^l}^2(n) \\ m_{F_{m_{y_i}}^l}(n), & \text{caso contrário} \end{cases}, \quad (2.125)$$

onde $y_{i,max}^l(n)$ é o valor que fornece o supremo.

Supondo-se que as variáveis $y_i(n)$, $i = 0, \dots, L - 1$ apresentam o mesmo nível de incerteza, $\beta_{y_i}(n) = \beta(n)$, tem-se que a saída do algoritmo fuzzy do tipo 1 e *nonsingleton* é dada por

$$f_{ns,1}(\mathbf{y}(n)) = \frac{\sum_{l=1}^M \theta_l \prod_{i=0}^{L-1} \mu_{Q_i^l}(y_{i,max}^l(n))}{\sum_{l=1}^M \prod_{i=0}^{L-1} \mu_{Q_i^l}(y_{i,max}^l(n))}, \quad (2.126)$$

$$\mu_{Q_i^l}(y_{i,max}^l(n)) = \exp \left[-\beta(n) |y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n)| - \frac{1}{2} \left(\frac{y_{i,max}^l(n) - m_{F_i^l}(n)}{\sigma_{F_i^l}(n)} \right)^2 \right], \quad (2.127)$$

e

$$|y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n)| = \begin{cases} y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n), & \text{se } y_{i,max}^l(n) \geq m_{F_{m_{y_i}}^l}(n) \\ - \left(y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n) \right), & \text{se } y_{i,max}^l(n) < m_{F_{m_{y_i}}^l}(n) \end{cases}. \quad (2.128)$$

Assumindo-se que a função custo a ser minimizada é dada por

$$J_c(\mathbf{w}(n)) = \frac{1}{2} [y_d(n) - f_{ns,1}(\mathbf{y}(n))]^2 = \frac{1}{2} e^2(n), \quad n = 0, \dots, N - 1, \quad (2.129)$$

onde $y_d(n)$ é a saída desejada, e, baseando-se no método *backpropagation* (Rumelhart et al., 1986) para a atualização dos parâmetros deste algoritmo fuzzy do tipo 1 e *nonsingleton*, tem-se que as expressões para a atualização dos parâmetros deste algoritmo fuzzy do tipo 1 são dadas por:

$$\theta_l(n+1) = \theta_l(n) - \alpha_{\theta_l}(n) \nabla_{\theta_l(n)} J_c(\mathbf{w}(n)) = \theta_l(n) + \alpha_{\theta_l(n)} e(n) \frac{a^l(n)}{b(n)}, \quad (2.130)$$

$$\begin{aligned}
m_{F_i^l}(n+1) &= m_{F_i^l}(n) - \alpha_{m_{F_i^l}}(n) \nabla_{m_{F_i^l}(n)} J_c(\mathbf{w}(n)) \\
&= m_{F_i^l}(n) + \alpha_{m_{F_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) \left(\frac{y_{i,max}^l(n) - m_{F_i^l}(n)}{\sigma_{F_i^l}^2(n)} \right), \tag{2.131}
\end{aligned}$$

$$\begin{aligned}
\sigma_{F_i^l}(n+1) &= \sigma_{F_i^l}(n) - \alpha_{\sigma_{F_i^l}}(n) \nabla_{\sigma_{F_i^l}(n)} J_c(\mathbf{w}(n)) \\
&= \sigma_{F_i^l}(n) + \alpha_{\sigma_{F_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) \frac{(y_{i,max}^l(n) - m_{F_i^l}(n))^2}{\sigma_{F_i^l}^3(n)} \sigma_{F_i^l}(n), \tag{2.132}
\end{aligned}$$

$$\begin{aligned}
\beta(n+1) &= \beta(n) - \alpha_{\beta}(n) \nabla_{\beta}(n) J_c(\mathbf{w}(n)) \\
&= \beta(n) + \alpha_{\beta}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) |y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n)|, \tag{2.133}
\end{aligned}$$

$$\begin{aligned}
m_{F_{m_{y_i}}^l}(n+1) &= m_{F_{m_{y_i}}^l}(n) - \alpha_{m_{F_{m_{y_i}}^l}}(n) \nabla_{m_{F_{m_{y_i}}^l}(n)} J_c(\mathbf{w}(n)) \\
&= m_{F_{m_{y_i}}^l}(n) - \alpha_{m_{F_{m_{y_i}}^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,1}(\mathbf{y}(n))}{b(n)} \right] a^l(n) \beta(n) \mathcal{K}_{l,i}(n), \tag{2.134}
\end{aligned}$$

e

$$\mathcal{K}_{l,i}(n) = \begin{cases} 1, & \text{se } y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n) \geq 0 \\ -1, & \text{caso contrário} \end{cases}, \tag{2.135}$$

onde $l = 1, \dots, M$ e $i = 0, \dots, L-1$. $\alpha_{\theta_l}(n)$, $\alpha_{m_{F_i^l}}(n)$, $\alpha_{\sigma_{F_i^l}}(n)$, $\alpha_{\beta}(n)$ e $\alpha_{m_{F_{m_{y_i}}^l}}(n)$ são os passos de adaptação dos parâmetros do algoritmo fuzzy do tipo 1 cuja saída é dada por

$$f_{ns,1}(\mathbf{y}(n)) = \frac{\sum_{l=1}^M \theta_l(n) a^l(n)}{b(n)}, \tag{2.136}$$

$$a^l(n) = \prod_{i=0}^{L-1} \exp \left[-\beta(n) |y_{i,max}^l(n) - m_{F_{m_{y_i}}^l}(n)| - \frac{1}{2} \left(\frac{y_{i,max}^l(n) - m_{F_i^l}(n)}{\sigma_{F_i^l}(n)} \right)^2 \right], \tag{2.137}$$

e

$$b(n) = \sum_{l=1}^M a^l(n). \quad (2.138)$$

A seguir, introduz-se um novo algoritmo fuzzy do tipo 2 que leva em consideração as mesmas funções de pertinências usadas para a obtenção do algoritmo fuzzy do tipo 1 adaptativo, deduzido acima. As expressões que fornecem a saída deste novo algoritmo e que permitem a atualização dos parâmetros do mesmo podem ser vistas como versões extendidas do algoritmo fuzzy do tipo 1 descrito anteriormente cuja saída é dada (2.126). Tal observação é similar àquela feita para o algoritmo fuzzy do tipo 2, intervalado e *nonsingleton*, introduzido na Seção 2.3.1.

A dedução deste novo algoritmo fuzzy do tipo 2 é baseada nas restrições impostas para a obtenção da expressão (2.28), nas expressões (2.62)-(2.65) e nas restrições aplicadas para a obtenção dos supremos das expressões (2.66)-(2.67) que fundamentam a proposição do algoritmo fuzzy do tipo 2, introduzido na Seção 2.2.2.

Baseando-se nessas restrições e considerações, a saída do algoritmo fuzzy do tipo 2, *nonsingleton*, intervalado e adaptativo, aqui proposto, é dada por

$$f_{ns,2}(\mathbf{y}(n)) = \frac{\sum_{l=1}^M \theta_l a_{ns,2}^l(n)}{b_{ns,2}(n)}, \quad (2.139)$$

onde

$$a_{ns,2}^l(n) = \bar{a}_{ns,2}^l(n) + \underline{a}_{ns,2}^l(n), \quad (2.140)$$

$$\bar{a}_{ns,2}^l(n) = \prod_{i=0}^{L-1} \exp \left[-\bar{\beta}(n) |\bar{y}_{i,max}^l(n) - \bar{m}_{\bar{F}_i^l}(n)| - \frac{1}{2} \frac{(\bar{y}_{i,max}^l(n) - \bar{m}_{\bar{F}_i^l}(n))^2}{\bar{\sigma}_{\bar{F}_i^l}^2(n)} \right], \quad (2.141)$$

$$\underline{a}_{ns,2}^l(n) = \prod_{i=0}^{L-1} \exp \left[-\underline{\beta}(n) |\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l m_{y_i}}(n)| - \frac{1}{2} \frac{(\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l m_{y_i}}(n))^2}{\underline{\sigma}_{\tilde{F}_i^l}^2(n)} \right], \quad (2.142)$$

$$|\bar{y}_{i,max}^l(n) - \bar{m}_{\tilde{F}_i^l m_{y_i}}(n)| = \begin{cases} \bar{y}_{i,max}^l(n) - \bar{m}_{\tilde{F}_i^l m_{y_i}}(n), & \text{se } \bar{y}_{i,max}^l(n) \geq \bar{m}_{\tilde{F}_i^l m_{y_i}}(n) \\ -(\bar{y}_{i,max}^l(n) - \bar{m}_{\tilde{F}_i^l m_{y_i}}(n)), & \text{se } \bar{y}_{i,max}^l(n) < \bar{m}_{\tilde{F}_i^l m_{y_i}}(n) \end{cases}, \quad (2.143)$$

$$|\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l m_{y_i}}(n)| = \begin{cases} \underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l m_{y_i}}(n), & \text{se } \underline{y}_{i,max}^l(n) \geq \underline{m}_{\tilde{F}_i^l m_{y_i}}(n) \\ -(\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l m_{y_i}}(n)), & \text{se } \underline{y}_{i,max}^l(n) < \underline{m}_{\tilde{F}_i^l m_{y_i}}(n) \end{cases} \quad (2.144)$$

e

$$b_{ns,2}(n) = \sum_{l=1}^M a_{ns,2}^l(n), \quad (2.145)$$

onde os valores de $\bar{y}_{i,max}^l(n)$ e $\underline{y}_{i,max}^l(n)$ são dados pelas expressões (2.68) e (2.69).

Considerando-se o uso do método *backpropagation* e que a função custo a ser minimizada é dada por

$$J = \frac{1}{2} [y_d(n) - f_{ns,2}(\mathbf{y}(n))]^2 = \frac{1}{2} e^2(n), \quad n = 0, \dots, N-1, \quad (2.146)$$

onde N é o número de conjuntos de treinamento, obtém-se as seguintes equações para a atualização dos parâmetros deste algoritmo fuzzy do tipo 2 e adaptativo:

$$\theta_l(n+1) = \theta_l(n) - \alpha_{\theta_l}(n) \nabla_{\theta_l(n)} J_c(\mathbf{w}(n)) = \theta_l(n) + \alpha_{\theta_l}(n) e(n) \frac{a_{ns,2}^l(n)}{b_{ns,2}(n)} \quad (2.147)$$

$$\begin{aligned} \underline{m}_{\tilde{F}_i^l}(n+1) &= \underline{m}_{\tilde{F}_i^l}(n) - \alpha_{\underline{m}_{\tilde{F}_i^l}}(n) \nabla_{\underline{m}_{\tilde{F}_i^l}(n)} J_c(\mathbf{w}(n)) \\ &= \underline{m}_{\tilde{F}_i^l}(n) + \alpha_{\underline{m}_{\tilde{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \underline{a}_{ns,2}^l(n) \left(\frac{\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l}(n)}{\underline{\sigma}_{\tilde{F}_i^l}^2(n)} \right), \end{aligned} \quad (2.148)$$

$$\begin{aligned}
\overline{m}_{\tilde{F}_i^l}(n+1) &= \overline{m}_{\tilde{F}_i^l}(n) - \alpha_{\overline{m}_{\tilde{F}_i^l}}(n) \nabla_{\overline{m}_{\tilde{F}_i^l}} J_c(\mathbf{w}(n)) \\
&= \overline{m}_{\tilde{F}_i^l}(n) + \alpha_{\overline{m}_{\tilde{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \overline{a}_{ns,2}^l(n) \left(\frac{\overline{y}_{i,max}^l(n) - \overline{m}_{\tilde{F}_i^l}(n)}{\overline{\sigma}_{\tilde{F}_i^l}^2(n)} \right), \quad (2.149)
\end{aligned}$$

$$\begin{aligned}
\underline{\sigma}_{\tilde{F}_i^l}(n+1) &= \underline{\sigma}_{\tilde{F}_i^l}(n) - \alpha_{\underline{\sigma}_{\tilde{F}_i^l}}(n) \nabla_{\underline{\sigma}_{\tilde{F}_i^l}} J_c(\mathbf{w}(n)) \\
&= \underline{\sigma}_{\tilde{F}_i^l}(n) + \alpha_{\underline{\sigma}_{\tilde{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \underline{a}_{ns,2}^l(n) \frac{\left(\underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_i^l}(n) \right)^2}{\underline{\sigma}_{\tilde{F}_i^l}^3(n)} \underline{\sigma}_{\tilde{F}_i^l}(n), \quad (2.150)
\end{aligned}$$

$$\begin{aligned}
\overline{\sigma}_{\tilde{F}_i^l}(n+1) &= \overline{\sigma}_{\tilde{F}_i^l}(n) - \alpha_{\overline{\sigma}_{\tilde{F}_i^l}}(n) \nabla_{\overline{\sigma}_{\tilde{F}_i^l}} J_c(\mathbf{w}(n)) \\
&= \overline{\sigma}_{\tilde{F}_i^l}(n) + \alpha_{\overline{\sigma}_{\tilde{F}_i^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}(n)} \right] \overline{a}_{ns,2}^l(n) \frac{\left(\overline{y}_{i,max}^l(n) - \overline{m}_{\tilde{F}_i^l}(n) \right)^2}{\overline{\sigma}_{\tilde{F}_i^l}^3(n)} \overline{\sigma}_{\tilde{F}_i^l}(n), \quad (2.151)
\end{aligned}$$

$$\begin{aligned}
\underline{\beta}(n+1) &= \underline{\beta}(n) - \alpha_{\underline{\beta}}(n) \nabla_{\underline{\beta}} J_c(\mathbf{w}(n)) \\
&= \underline{\beta}(n) + \alpha_{\underline{\beta}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}^l(n)} \right] \underline{a}_{ns,2}^l(n) \left| \underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_{m_{y_i}}^l}(n) \right|, \quad (2.152)
\end{aligned}$$

$$\begin{aligned}
\overline{\beta}(n+1) &= \overline{\beta}(n) - \alpha_{\overline{\beta}}(n) \nabla_{\overline{\beta}} J_c(\mathbf{w}(n)) \\
&= \overline{\beta}(n) + \alpha_{\overline{\beta}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b_{ns,2}^l(n)} \right] \overline{a}_{ns,2}^l(n) \left| \overline{y}_{i,max}^l(n) - \overline{m}_{\tilde{F}_{m_{y_i}}^l}(n) \right|. \quad (2.153)
\end{aligned}$$

$$\begin{aligned}
\underline{m}_{\tilde{F}_{m_{y_i}}^l}(n+1) &= \underline{m}_{\tilde{F}_{m_{y_i}}^l}(n) - \alpha_{\underline{m}_{\tilde{F}_{m_{y_i}}^l}}(n) \nabla_{\underline{m}_{\tilde{F}_{m_{y_i}}^l}} J_c(\mathbf{w}(n)) \\
&= \underline{m}_{\tilde{F}_{m_{y_i}}^l}(n) - \alpha_{\underline{m}_{\tilde{F}_{m_{y_i}}^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b(n)} \right] \underline{a}_{ns,2}^l(n) \underline{\beta}(n) \underline{\mathcal{K}}_{l,i}(n), \quad (2.154)
\end{aligned}$$

$$\underline{\mathcal{K}}_{l,i}(n) = \begin{cases} 1, & \text{se } \underline{y}_{i,max}^l(n) - \underline{m}_{\tilde{F}_{m_{y_i}}^l}(n) \geq 0 \\ -1, & \text{caso contrário} \end{cases}, \quad (2.155)$$

e, finalmente,

$$\begin{aligned} \overline{m}_{\tilde{F}_{m_{y_i}}^l}(n+1) &= \overline{m}_{\tilde{F}_{m_{y_i}}^l}(n) - \alpha_{\overline{m}_{\tilde{F}_{m_{y_i}}^l}}(n) \nabla_{\overline{m}_{\tilde{F}_{m_{y_i}}^l}}(n) J_c(\mathbf{w}(n)) \\ &= \overline{m}_{\tilde{F}_{m_{y_i}}^l}(n) - \alpha_{\overline{m}_{\tilde{F}_{m_{y_i}}^l}}(n) e(n) \left[\frac{\theta_l(n) - f_{ns,2}(\mathbf{y}(n))}{b(n)} \right] \overline{a}_{ns,2}^l(n) \overline{\beta}(n) \overline{\mathcal{K}}_{l,i}(n), \end{aligned} \quad (2.156)$$

e

$$\overline{\mathcal{K}}_{l,i}(n) = \begin{cases} 1, & \text{se } \overline{y}_{i,max}^l(n) - \overline{m}_{\tilde{F}_{m_{y_i}}^l}(n) \geq 0 \\ -1, & \text{caso contrário} \end{cases}. \quad (2.157)$$

onde $l = 1, \dots, M$, $i = 0, \dots, L - 1$.

2.3.3 Estratégias para a Atualização do Passo de Adaptação

Recentemente, o uso dos sistemas fuzzy do tipo 1 e 2 para minimizar a interferência inter-simbólica em canais de comunicação tem sido investigado (Liang & Mendel, 2000a, 2000c; Lee, 1994; Patra & Mulgrew, 1998a; Wang & Mendel, 1993), destacando-se as versões fuzzy dos tipos 1 e 2 e *nonsingleton* que apresentam bons resultados alcançados. Entretanto, a convergência lenta quando o método *backpropagation* (Rumelhart et al., 1986) é implementado para o treinamento dos sistemas fuzzy, vide (Wang & Mendel, 1993), faz com que o projeto dos algoritmos fuzzy adaptativos para aplicações em tempo real seja limitado. As razões para a baixa velocidade de convergência dos algoritmos fuzzy adaptativos são as seguintes: *i*) o sentido contrário do vetor gradiente pode não estar apontado para o mínimo da função custo utilizada como critério de minimização e *ii*) o uso do passo de adaptação constante pode não ser adequado para todas as regiões da superfície da função custo.

Visando levar as duas razões acima em consideração, algumas estratégias para a atualização do passo de adaptação são, aqui, discutidas para determinar o passo de adaptação mais adequado para os algoritmos fuzzy adaptativos. A escolha das técnicas aqui analisadas justifica-se pelo fato de que as mesmas são baseadas em informações de 1ª ordem e, conseqüentemente, permitem que a convergência mais rápida dos algoritmos fuzzy seja alcançada, ao custo de um aumento linear da complexidade computacional.

De um modo geral, pode-se afirmar que o objetivo do treinamento supervisionado do algoritmo fuzzy adaptativo é minimizar uma função custo expressa por $J_c(\mathbf{w}(n))$, onde $\mathbf{w}(n)$ é o vetor constituído por todos os parâmetros do algoritmo fuzzy. No entanto, a atualização dos parâmetros dos algoritmos fuzzy adaptativos é uma tarefa bastante difícil, posto que o número de parâmetros é, normalmente, elevado e a função custo a ser minimizada, $J_c(\mathbf{w}(n))$, possui vários mínimos locais, platos e regiões íngrimes.

Em (Wang & Mendel, 1993) foi mostrado que o método *backpropagation* (BP) (Rumelhart et al., 1986), baseado em informação de 1ª ordem, pode ser utilizado para a atualização dos parâmetros do algoritmo fuzzy adaptativo. No entanto, é sabido que o método BP apresenta baixa taxa de convergência. Ainda em (Wang & Mendel, 1993) foi introduzido o algoritmo fuzzy adaptativo treinado pelo método RLS (*Recursive Least-Square*). Apesar do mesmo apresentar desempenho superior àquele obtido com o método BP baseado em informação de 1ª ordem, a complexidade computacional é excessivamente alta e, portanto, a implementação do algoritmo fuzzy adaptativo, com treinamento baseado no método RLS, não é recomendada para as aplicações em tempo real.

Uma técnica de treinamento baseada em SVD (*Singular Vector Decomposition*) foi proposta por Mendel (Mendel, 2001). Novamente, a complexidade computacional

necessária para atualizar os parâmetros do algoritmo fuzzy foi salientada. O uso do método GCEM (Santos & Zuben, 2000), aplicado na Seção 2.2 para o treinamento de algoritmos fuzzy adaptativos, também apresenta elevada carga computacional, apesar desta ser menor do que aquelas observadas em outros métodos de otimização baseados em informações de 2ª ordem. Tal característica do método GCEM inibe o uso do mesmo quando limitações de capacidade de processamento são impostas no projeto do algoritmo fuzzy a ser desenvolvido para resolver um problema específico.

Recentes contribuições sugerem o uso do método BP com técnicas que permitam a atualização do passo de adaptação de cada um dos parâmetros do algoritmo (Ortega & Rheinboldt, 1970; Magoulas & Vrahatis, 1999; Jacobs, 1988; Magoulas & Vrahatis, 2000; Vrahatis, Magoulas, & Plagianakos, 2000; Silva & Almeida, 1990; Karayiannis & Venetsanopoulos, 1992; Vrahatis et al., 2000; Vogl, Mangis, Rigler, Zink, & Alkon, 1988; Riedmiller, 1994; Fahlman, 1988). De fato, é sabido que o uso do método BP com passo de adaptação fixo não é uma solução adequada, pois a superfície da função custo não é convexa em algoritmos não lineares, tais como os algoritmos fuzzy adaptativos.

Objetivando analisar o uso de tais técnicas em algoritmos fuzzy adaptativos, as seções 2.3.3.2 e 2.3.3.1 discutem o uso de duas estratégias denominadas *Delta Rule Delta* (DRD) (Jacobs, 1988) e *Local Lipschitz Estimation* (LLE) (Magoulas & Vrahatis, 1999) para a atualização do passo de adaptação dos parâmetros dos algoritmos fuzzy adaptativos.

A seguinte formulação é adotada a seguir: $\mathbf{w}(n) = [w_i(n) \cdots w_{L_1-1}(n)]^T$ é o vetor que contém todos os parâmetros do algoritmo fuzzy adaptativo na n -ésima iteração, \mathbf{w}^* é o vetor de parâmetros ótimos,

$$J_c(\mathbf{w}(n)) = \frac{1}{2} [y_a(n) - \hat{y}(n-d)]^2 \quad (2.158)$$

é a função custo a ser minimizada, $y(n)$ é a saída do algoritmo e $\hat{y}(n-d)$ é a saída

desejada atrasada de d amostras, $\partial J_c(\mathbf{w}(n))$ é a derivada parcial de $J_c(\mathbf{w}(n))$ em relação à j -ésima variável $w_j(n)$ do vetor $\mathbf{w}(n)$, $\nabla J_c(\mathbf{w}(n))$ é o vetor gradiente.

2.3.3.1 Estratégia Baseada em Informações de 1ª Ordem: *Delta Rule Delta*

De um modo geral, pode-se afirmar que a técnica *steepest descent* (SD), baseada no método BP, é eficiente para a obtenção dos parâmetros que minimizam $J_c(\mathbf{w}(n))$. No entanto, a velocidade de convergência do SD é bastante lenta em função da topologia da superfície de $J_c(\mathbf{w}(n))$, devido às seguintes razões.

Primeiramente, a amplitude de $\partial J_c(\mathbf{w}(n))$ pode ser tal que o uso de um passo de adaptação específico irá produzir pequena redução em $J_c(\mathbf{w}(n))$. Isso pode acontecer em duas situações: 1) a superfície de $J_c(\mathbf{w}(n))$ é plana na direção do j -ésimo parâmetro, ou seja $0 < |\partial J_c(\mathbf{w}(n))| \ll 1$ e, conseqüentemente, a atualização do j -ésimo parâmetro em cada iteração é insignificante, resultando num número considerável de iterações para se obter uma minimização considerável de $J_c(\mathbf{w}(n))$ e 2) se a superfície de $J_c(\mathbf{w}(n))$ é inclinada na direção do j -ésimo parâmetro, então $\partial J_c(\mathbf{w}(n)) \gg 1$, resultando no não alcance do mínimo na direção do j -ésimo parâmetro.

A segunda justificativa para a velocidade de convergência baixa da técnica SD é que a direção de $-\nabla J_c(\mathbf{w}(n))$ não aponta na direção de $J_c(\mathbf{w}^*)$, pois os contornos do ponto $\mathbf{w}(n)$ são, normalmente, elípticos e não circulares. No caso do contorno circular $-\nabla J_c(\mathbf{w}(n))$ aponta na direção do mínimo. Baseando-se na teoria de filtragem adaptativa (Haykin, 1996), a seguinte justificativa é apresentada para esta afirmação.

Seja $\mathbf{R}_y = E\{\mathbf{y}(n)\mathbf{y}^T(n)\}$ a matriz de autocorrelação do vetor $\mathbf{y}(n)$ aleatório.

Então, a superfície de erro possui um único \mathbf{w}^* e os autovalores de \mathbf{R}_y determinam a forma da superfície de $J_c(\mathbf{w}(n))$ (Haykin, 1996; Widrow & Hoff, 1985). Note que \mathbf{R}_y pode ser expresso por

$$\mathbf{R}_y = \mathbf{Q}\Lambda\mathbf{Q}^{-1}, \quad (2.159)$$

sendo Λ a matriz diagonal cujos componentes são os autovalores de \mathbf{R}_y e \mathbf{Q} é a matriz cujas colunas representam o conjunto de autovetores associados aos autovalores em Λ , os autovalores numa ordem ascendentes são denominados $\lambda_1, \dots, \lambda_N$. Os autovetores que definem os eixos principais da superfície de $J_c(\mathbf{w}(n))$ fornecem maiores deslocamentos nas direções dos autovetores associados aos maiores autovalores e o contrário ocorre para as direções associadas aos menores autovalores. Pode-se constatar que se $\lambda_1 = \dots = \lambda_N$, então o contorno formado pelos autovetores de \mathbf{R}_y é circular e $-\nabla J_c(\mathbf{w}(n))$ aponta na direção do mínimo de $J_c(\mathbf{w}(n))$. Por outro lado, se $\lambda_1 \neq \dots \neq \lambda_N$ o contorno formado pelos autovetores de \mathbf{R}_y é elíptico e $-\nabla J_c(\mathbf{w}(n))$ não aponta na direção do mínimo de $J_c(\mathbf{w}(n))$. Note ainda que se um mesmo passo de adaptação for adotado para todos os parâmetros, isso significa que pequenos e/ou grandes deslocamentos na superfície de $J_c(\mathbf{w}(n))$ serão observados nas direções dos autovetores associados aos autovalores com menores e/ou maiores e amplitudes, respectivamente. No entanto, o fato das direções dos autovetores associados aos menores autovalores representarem as regiões planas da superfície de $J_c(\mathbf{w}(n))$ significa que passos de adaptação maiores devem ser utilizados de forma a se avançar nessas direções. Por outro lado, as direções dos autovetores associados aos maiores autovalores representam superfícies inclinadas de $J_c(\mathbf{w}(n))$.

Objetivando levar tais análises em consideração para a formulação de técnicas de atualização dos parâmetros de redes neurais, Jacobs (Jacobs, 1988), baseado em várias contribuições (Barto & Sutton, 1981; Kesten, 1958; Saridis, 1970; Sutton, 1986), propôs uma regra heurística para a atualização do passo de adaptação dos parâmetros do vetor $\mathbf{w}(n)$. Tal regra, denominada de *delta-rule-delta* (DRD) faz

uso da seguinte regra para a atualização do vetor $\mathbf{w}(n)$

$$\mathbf{w}(n+1) = \mathbf{w}(n) - (1 - \alpha) \text{diag}\{\mu_0(n) \dots \mu_{L_1-1}(n)\} \nabla J_c(\mathbf{w}(n)) + \alpha \Delta \mathbf{w}(n)(n) \quad (2.160)$$

$$\mu_j(n+1) = \begin{cases} \mu_j(n) + K_1, & \text{se } \bar{\nabla}_j J_c(\mathbf{w}(n-1)) \nabla_j J_c(\mathbf{w}(n)) > 0, \\ \mu_j(n) - K_2 \mu_j(n), & \text{se } \bar{\nabla}_j J_c(\mathbf{w}(n-1)) \nabla_j J_c(\mathbf{w}(n)) < 0, \\ 0, & \text{caso contrário,} \end{cases} \quad (2.161)$$

sendo $j = 0, \dots, L_1 - 1$, K_1 é o incremento no passo de adaptação, $\alpha \in [0, 1]$, K_2 é o valor para decremento do passo de adaptação, $\Delta \mathbf{w}(n) = \mathbf{w}(n) - \mathbf{w}(n-1)$ e

$$\bar{\nabla}_j J_c(\mathbf{w}(n)) = (1 - \rho_j) \nabla_j J_c(\mathbf{w}(n)) + \rho_j \bar{\nabla}_j J_c(\mathbf{w}(n-1)) \quad (2.162)$$

é a média exponencial de $\nabla_j J_c(\mathbf{w}(m))$, $m = \dots, n-2, n-1, n$ e $\rho_j \in [0, 1]$.

De acordo com a equação de atualização da técnica DRD, vide (2.160)-(2.161), se $\nabla_j J_c(\mathbf{w}(n))$ e $\bar{\nabla}_j J_c(\mathbf{w}(n-1))$ apresentam o mesmo sinal, então $\mu_j(n+1) = \mu_j(n) + K_1$. Por outro lado, se $\nabla_j J_c(\mathbf{w}(n))$ e $\bar{\nabla}_j J_c(\mathbf{w}(n-1))$ têm sinais diferentes, então $\mu_j(n+1) = \mu_j(n) - K_2 \mu_j(n)$, ou seja, o passo de adaptação é diminuído para um valor proporcional de $\mu_j(n)$. A técnica DRD aumenta linearmente e diminui exponencialmente o passo de adaptação. O aumento linear do passo de adaptação evita que o mesmo aumente muito rapidamente, enquanto a diminuição exponencial garante que o passo de adaptação diminua rapidamente e seja sempre positivo.

2.3.3.2 Estratégia Baseada em informações de 1ª Ordem: *Local Lipschitz Estimation*

Analisando-se a sensibilidade do processo de minimização da função custo $J_c(\mathbf{w}(n))$, pode-se observar que, para vizinhanças suficientemente próximas de \mathbf{w}^* , os eixos do contorno em torno de \mathbf{w}^* são dados pelos autovetores de $\nabla^2 J_c(\mathbf{w}^*)$, enquanto o comprimento dos eixos serão inversamente proporcional às raízes

quadradas dos correspondentes autovalores. Como a superfície de $J_c(\mathbf{w}(n))$ é não convexa e existem diferentes mínimos locais, o contorno gerado a partir dos autovetores de $\nabla^2 J_c(\mathbf{w}^*)$ é elíptico. Conforme discutido em 2.3.3.1, uma variação na direção do autovetor correspondente ao autovalor de maior amplitude irá causar a maior variação em $J_c(\mathbf{w}^*)$, enquanto uma variação no autovetor correspondente ao menor autovalor gerará a mínima variação em $J_c(\mathbf{w}^*)$. Baseando-se nessa análise, a Seção 2.3.3.1 demonstrou que o passo de adaptação na direção do j -ésimo parâmetro de $\mathbf{w}(n)$ não é necessariamente a melhor escolha para a atualização do k -ésimo parâmetro de $\mathbf{w}(n)$, $k \neq j$, $k = 0, \dots, L_1 - 1$. Além disso, esse mesmo passo de adaptação não é adequado para diferentes regiões da superfície de $J_c(\mathbf{w}(n))$. Desta forma, a questão fundamental para a atualização dos parâmetros do algoritmo fuzzy adaptativo é encontrar o passo de adaptação adequado para os parâmetros que se encontram em regiões planas e inclinadas da superfície de $J_c(\mathbf{w}(n))$. Observa-se que o passo de adaptação deve ser grande quando o parâmetro encontra-se numa região plana e pequeno quando o parâmetro encontra-se numa região com inclinação da superfície de $J_c(\mathbf{w}(n))$.

Uma solução para tal problema é a atualização do passo de adaptação a partir da estimativa local da constante de Lipschitz (*Local Lipschitz Estimation* - LLE) (Armijo, 1966) nas direções de cada um dos parâmetros (Magoulas & Vrahatis, 1999). A técnica baseada na LLE explora as direções dos vetores gradientes e da morfologia da superfície de $J_c(\mathbf{w}(n))$ no ponto corrente para dinamicamente atualizar o passo de adaptação de cada parâmetro. A inversa da constante Lipschitz pode ser utilizada para a obtenção da atualização ótima do passo de adaptação na n -ésima iteração. Dado que as características topológicas da superfície de $J_c(\mathbf{w}(n))$ e do valor da constante de Lipschitz não são conhecidas, *a priori*, usa-se a norma infinita para a estimação da constante Lipschitz local com a seguinte expressão

$$\lambda(n) = \frac{\max_{0 \leq j \leq L_1 - 1} |\partial_j J_c(\mathbf{w}(n)) - \partial_j J_c(\mathbf{w}(n-1))|}{\max_{0 \leq j \leq L_1 - 1} |w_j(n) - w_j(n-1)|}. \quad (2.163)$$

Para se levar em consideração a topologia da superfície de $J_c(\mathbf{w}(n))$ na atualização dinâmica de cada um dos parâmetros dos algoritmos fuzzy adaptativos, introduzidos nas seções 2.3.1 e 2.3.2, a LLE é obtida nas direções de todos os L_1 parâmetros, ou seja,

$$\lambda_j(n) = \frac{|\partial_j J_c(\mathbf{w}(n)) - \partial_j J_c(\mathbf{w}(n-1))|}{|w_j(n) - w_j(n-1)|}, \quad j = 0, \dots, L_1 - 1. \quad (2.164)$$

Baseando-se em (2.164), tem-se que a expressão de atualização dos parâmetros dos algoritmos fuzzy adaptativos é dada por

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \gamma(n) \text{diag} \left\{ \lambda_0^{-1}(n) \quad \dots \quad \lambda_{L_1-1}^{-1}(n) \right\} \nabla J(\mathbf{w}(n)). \quad (2.165)$$

A seguinte regra deve ser utilizada para atualizar o coeficiente de relaxamento $\gamma(n)$:

Se a inequação dada por

$$J_c(\mathbf{w}(n+1)) - J_c(\mathbf{w}(n)) \leq -\frac{1}{2}\gamma(n) \left\| \text{diag} \left\{ \lambda_0^{-1}(n), \quad \dots, \lambda_{L_1-1}^{-1}(n) \right\} \nabla J(\mathbf{w}(n)) \right\|^2. \quad (2.166)$$

é verdadeira, então

$$m = m - 1 \text{ e}$$

$$\gamma(n+1) = \frac{\gamma(0)}{q^{m-1}}$$

Senão $m = m + 1$ e

$$\gamma(n+1) = \frac{\gamma(0)}{q^{m-1}}.$$

Finalmente, o custo computacional das técnicas DRD e LLE são mostrados na Tab. 2.1. Nessa tabela, $Ca(BP)$, $Cs(BP)$, $Cm(BP)$ e $Cd(BP)$ são os custos

computacionais devidos às operações de adição, subtração, multiplicação, divisão e comparação quando a técnica BP é utilizada para a atualização dos parâmetros dos algoritmos fuzzy. Note que apenas as operações aritméticas em ponto flutuante foram consideradas.

Operação	<i>DRD</i>	<i>LLE</i>
Adição	$Ca(BP) + 3L_1$	$2 * Ca(BP)$
Subtração	$Cs(BP) + L_1$	$2 * Cs(BP) + 3L_1$
Multiplicação	$Cm(BP) + 5L_1$	$2 * Cm(BP) + 3L_1$
Divisão	-	$2L_1$
Comparação	L_1	L_1

Tabela 2.1: Custo computacional das técnicas DRD e LLE.

Resultados numéricos que comprovem o ganho de desempenho, em termos de velocidade de convergência, promovido pelas técnicas DRD e LLE em aplicações de equalização de canais PLC são mostrados na Seção 5.3.1 do capítulo 5.

2.4 Classificador Fuzzy do tipo 2

A formalização do uso dos sistemas fuzzy do tipo 1 para a aplicação de reconhecimento de padrões foi estabelecida em (Wei & Mendel, 1999). No entanto, sabe-se que os sistemas fuzzy do tipo 1 é um caso particular dos sistemas fuzzy do tipo 2 (Mendel, 2001). Ao explorar esta característica dos sistemas fuzzy do tipo 2 e levando-se em consideração a expressão (2.28), deduz-se, a seguir, um classificador fuzzy do tipo 2, o qual é usado nesta tese para a classificação de distúrbios em redes elétricas. Conforme será mostrado no final da Seção, este classificador fuzzy do tipo 2 com algumas restrições, reduz-se ao classificador fuzzy do tipo 1 introduzido em (Wei & Mendel, 1999).

Assumindo-se que as entradas dos classificadores de padrões propostos para classificar sinais nos vetores $\mathbf{x}(n) = [x_0(n) \cdots x_{N-1}(n)]^T$ sejam dados por $\mathbf{q}_n \in \mathcal{S}^x$, os quais são extraídos dos vetores $\mathbf{x}(n)$, onde \mathcal{S}^x é uma matriz constituída pelos vetores de parâmetros extraídos dos vetores $\mathbf{x}(n)$ e $\mathbf{q}_n = [q_{n,1}, \dots, q_{n,L}]^T$. Seja \mathbf{E}_i , $i = 1, \dots, C$, a i -ésima classe de distúrbios, tal que $\mathcal{S}^x = \bigcup_{i=1}^C \mathbf{E}_i$, ou seja \mathcal{S}^x é particionado em C regiões de classificação, definidas por $\Gamma_1, \Gamma_2, \dots, \Gamma_C$. Note que à região Γ_i está associada a classe \mathbf{E}_i de distúrbio. A função característica associada à região Γ_i é dada por

$$\mu_{\Gamma_i}(\mathbf{q}_n) = \begin{cases} 1, & \text{se } \mathbf{q}_n \in \Gamma_i, \\ 0, & \text{caso contrário} \end{cases}, \quad i = 1, \dots, C. \quad (2.167)$$

A expressão (2.167) implementa o classificador do tipo *hard decision* (Wei & Mendel, 1999), o qual é deduzido na Seção 9.3.1.

Supondo-se que o conjunto de classes $\mathbf{V} = \{\mathbf{E}_1, \dots, \mathbf{E}_C\}$ forma um universo de discurso sobre o qual os conjuntos fuzzy são definidos para representar o conceito de classes vagas, tem-se que a extensão da função característica dada por (2.167), para que a sua saída seja um conjunto fuzzy do tipo 1 e *singleton*, ou seja, $\mu_{\Gamma_i}(\mathbf{q}_n) \in [0, 1]$, é dada por

$$\mu_B(\mathbf{q}_n) = \frac{\mu_{\Gamma_i}(\mathbf{q}_n)}{\sum_{i=1}^C \mu_{\Gamma_i}(\mathbf{q}_n)}. \quad (2.168)$$

Note que $\mu_B(\mathbf{q}_n)$ é uma função característica que implementa um classificador baseado no critério *soft decision* (Wei & Mendel, 1999) e B é um conjunto fuzzy do tipo 1.

Baseando-se os resultados obtidos em (Wei & Mendel, 1999) para o conjunto fuzzy do tipo 1 e nas restrições utilizadas para a obtenção dos algoritmos fuzzy

introduzidos nas seções 2.2 e 2.3, deduz-se, a seguir, um classificador fuzzy do tipo 2. A motivação principal para a introdução do mesmo reside no fato de que quando existem incertezas relacionadas às regras fuzzy que definem as regiões das classes de distúrbios, o classificador fuzzy do tipo 1 e *singleton* não é capaz de levar tais informações em consideração e, conseqüentemente, o seu desempenho é limitado. Uma característica marcante do classificador fuzzy do tipo 2 deduzido abaixo é que o mesmo sob algumas restrições reduz-se ao classificador do tipo 1 e *singleton* e ao classificador ML propostos em (Wei & Mendel, 1999) e (Wei & Mendel, 2000), respectivamente.

Para o classificador fuzzy do tipo 2, a seguinte regra fuzzy do tipo 2 é aplicada para a i -ésima classe de distúrbio

$$R^{l,i} : \text{Se } q_{n,1} \text{ é } \tilde{F}_1^{l,i} \text{ Se } q_{n,2} \text{ é } \tilde{F}_2^{l,i} \text{ e } \dots \text{ e } q_{n,L} \text{ é } \tilde{F}_L^{l,i}, \text{ então } y^{l,i} \text{ é } G^{l,i}, \quad l = 1, \dots, M_i \quad (2.169)$$

onde $\tilde{F}_k^{l,i}$ e $G_k^{l,i}$ são, respectivamente, funções de pertinência do tipo 2 e do tipo 1, $l = 1, \dots, M_i$, M_i é o número de regras associadas à i -ésima classe, $i = 1, \dots, C$ são as classe de sinais. Cada regra em (2.169) é interpretada como uma implicação fuzzy do tipo 2. Conseqüentemente, (2.169) pode se dada por

$$R^{l,i} : \tilde{F}_1^{l,i} \times \dots \times \tilde{F}_L^{l,i} \rightarrow G^{l,i} = \tilde{A}^{l,i} \rightarrow G^{l,i}, \quad l = 1, \dots, M_i. \quad (2.170)$$

Se a regra $R^{l,i}$ é descrita pela função de pertinência

$$\mu_{R^{l,i}}(\mathbf{q}_n, y^{l,i}) = \mu_{\tilde{A}^{l,i} \rightarrow G^{l,i}}(\mathbf{q}_n, y^{l,i}) = \left[\prod_{l=1}^L \mu_{\tilde{F}_l^{l,i}}(q_{n,k}) \right] \sqcap \mu_{G^{l,i}}(y^{l,i}), \quad (2.171)$$

onde \sqcap é o operador *meet* associado ao cálculo da intersecção entre os conjuntos fuzzy do tipo 2 (Mizumoto & Tanaka, 1976), então

$$\begin{aligned} \mu_{\tilde{B}^{l,i}}(y^{l,i}) &= \sqcup_{\mathbf{q}_n} \left[\mu_{\tilde{A}^{l,i}}(\mathbf{q}_n) \sqcap \mu_{R^{l,i}}(\mathbf{q}_n, y^{l,i}) \right] = \sqcup_{\mathbf{q}_n} \left\{ \left[\prod_{l=1}^L \mu_{\tilde{Q}_l^{l,i}}(q_{n,k}) \right] \sqcap \mu_{G^{l,i}}(y^{l,i}) \right\} \\ &= \mu_{G^{l,i}}(y^{l,i}) \sqcap \left\{ \sqcup_{\mathbf{q}_n} \left[\prod_{l=1}^L \mu_{\tilde{Q}_l^{l,i}}(q_{n,k}) \right] \right\}, \end{aligned} \quad (2.172)$$

onde \sqcup é o operador *join* associado ao cálculo da união entre os conjuntos fuzzy do tipo 2 (Mizumoto & Tanaka, 1976),

$$\mu_{\tilde{A}_{\mathbf{q}_n}}(\mathbf{q}_n) = \prod_{l=1}^L \mu_{\tilde{X}_l^{l,i}}(q_{n,k}), \quad (2.173)$$

e

$$\mu_{\tilde{Q}_1^{l,i}}(q_{n,k}) = \mu_{\tilde{X}_l^{l,i}}(q_{n,k}) \prod \mu_{\tilde{F}_l^{l,i}}(q_{n,k}). \quad (2.174)$$

Finalmente, escreve-se

$$\mu_{\tilde{B}^{l,i}}(y^{l,i}) = \mu_{G^{l,i}}(y^{l,i}) \prod F^{l,i}(\mathbf{q}_n), \quad (2.175)$$

onde

$$F^{l,i}(\mathbf{q}_n) = \left\{ \sqcup_{\mathbf{q}_n} \left[\prod_{l=1}^L \mu_{\tilde{Q}_l^{l,i}}(q_{n,k}) \right] \right\}, \quad (2.176)$$

fornece o grau de pertinência do vetor \mathbf{q}_n à k -ésima regra *fuzzy*. Tem-se ainda que

$$\mu_{\tilde{B}^{l,i}}(y^{l,i}) = y^{l,i} \prod F^{l,i}(\mathbf{q}_n). \quad (2.177)$$

Supondo-se que o sistema fuzzy do tipo 2 é intervalado (Mendel, 2001), pode-se definir $F^{l,i}(\mathbf{q}_n)$ com a seguinte expressão

$$F^{l,i}(\mathbf{q}_n) = \left[\underline{f}^{l,i}(\mathbf{q}_n), \overline{f}^{l,i}(\mathbf{q}_n) \right], \quad (2.178)$$

onde

$$\begin{aligned} \underline{f}^{l,i}(\mathbf{q}_n) &= \sup_{\mathbf{q}_n} \int_{q_{n,1}} \dots \int_{q_{n,L}} \underline{\mu}_{\tilde{Q}_1^{l,i}}(q_{n,k}) \star \dots \star \underline{\mu}_{\tilde{Q}_l^{l,i}}(q_{n,k}) / \mathbf{q}_n, \\ &= \mathcal{T}_{l=1}^L \underline{f}^{l,i}(q_{n,k}) = \mathcal{T}_{l=1}^L \underline{\mu}_{\tilde{Q}_1^{l,i}}(q_{n,k,\max}) \end{aligned}, \quad (2.179)$$

$$\begin{aligned} \overline{f}^{l,i}(\mathbf{q}_n) &= \sup_{\mathbf{q}_n} \int_{q_{n,1}} \dots \int_{q_{n,L}} \overline{\mu}_{\tilde{Q}_1^{l,i}}(q_{n,k}) \star \dots \star \overline{\mu}_{\tilde{Q}_l^{l,i}}(q_{n,k}) / \mathbf{q}_n, \\ &= \mathcal{T}_{l=1}^L \overline{f}^{l,i}(q_{n,k}) = \mathcal{T}_{l=1}^L \overline{\mu}_{\tilde{Q}_1^{l,i}}(q_{n,k,\max}) \end{aligned}, \quad (2.180)$$

$$\underline{f}^{l,i}(q_{n,k}) = \sup_{q_{n,k}} \underline{\mu}_{\tilde{Q}_l^{l,i}}(q_{n,k,\max}) \quad (2.181)$$

e

$$\overline{f}^{l,i}(q_{n,k}) = \sup_{q_{n,k}} \overline{\mu}_{\tilde{Q}_l^{l,i}}(q_{n,k,\max}). \quad (2.182)$$

Os operadores $\underline{(\cdot)}$ e $\overline{(\cdot)}$ definem a região FOU, \mathcal{T} e \star são t -norm implementadas pelo operador produto.

Agora, seja

$$\mu_{\tilde{X}_k^{l,i}}(q_{n,k}) = \begin{cases} 1/1, & q_{n,k} = q_{n,k} \\ 1, 0, & q_{n,k} \neq q_{n,k} \end{cases}, \quad (2.183)$$

ou seja, o sistema fuzzy do tipo 2 e intervalado é *singleton*, então,

$$\underline{f}^{l,i}(\mathbf{q}_n) = \mathcal{T}_{l=1}^L \underline{\mu}_{\tilde{F}_k^{l,i}}(q_{n,k}) \quad (2.184)$$

e

$$\overline{f}^{l,i}(\mathbf{q}_n) = \mathcal{T}_{l=1}^L \overline{\mu}_{\tilde{F}_k^{l,i}}(q_{n,k}), \quad (2.185)$$

então, baseando-se em (2.28), tem-se que a saída do classificador fuzzy do tipo 2, intervalado e *singleton* da i -ésima classe de sinal é dada por

$$\mu_{\tilde{B}^i}(\mathbf{q}_n) = \frac{\mu_{\tilde{\Gamma}_i}(\mathbf{q}_n)}{\sum_{i=1}^C \mu_{\tilde{\Gamma}_i}(\mathbf{q}_n)} = \frac{\sum_{l=1}^{M_i} \underline{f}^{l,i}(\mathbf{q}_n) + \overline{f}^{l,i}(\mathbf{q}_n)}{\sum_{i=1}^C \sum_{l=1}^{M_i} \underline{f}^{l,i}(\mathbf{q}_n) + \overline{f}^{l,i}(\mathbf{q}_n)} \quad (2.186)$$

onde $\tilde{\Gamma}_i(\mathbf{q}_n)$ é um conjunto fuzzy do tipo 2, intervalado e *singleton* associado à classe de distúrbio \mathbf{E}_i . A classificação do sinal $\mathbf{x}(n)$ com pertencente à i -ésima classe de sinal é dada por

$$H_i = \arg \max_i \mu_{\tilde{B}^i}(\mathbf{q}_n) \quad i = 1, \dots, C \quad (2.187)$$

Agora, supondo-se que as t -norm são implementadas pelo operador produto e as funções que definem as FOUS das funções de pertinência das premissas dos antecedentes das regras sejam dadas por

$$\underline{\mu}_{\tilde{F}_l^{l,i}}(q_{n,k}) = \exp \left[-\frac{1}{2} \left(\frac{q_{n,k} - \underline{m}_l}{\underline{\sigma}_l} \right)^2 \right] \quad (2.188)$$

e

$$\overline{\mu}_{\tilde{F}_l^{l,i}}(q_{n,k}) = \exp \left[-\frac{1}{2} \left(\frac{q_{n,k} - \overline{m}_l}{\overline{\sigma}_l} \right)^2 \right]. \quad (2.189)$$

Então, tem-se que a função característica do da i -ésima classe de distúrbio é expressa por

$$\mu_{\tilde{B}^i}(\mathbf{q}_n) = \frac{\sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \underline{\mathbf{m}}^{l,i})^T \underline{\Sigma}_{\underline{\sigma}^{l,i}}^{-1} (\mathbf{q}_n^f - \underline{\mathbf{m}}^{l,i}) \right] + \exp \left[-\frac{1}{2} (\mathbf{q}_n - \overline{\mathbf{m}}^{l,i})^T \overline{\Sigma}_{\overline{\sigma}^{l,i}}^{-1} (\mathbf{q}_n^f - \overline{\mathbf{m}}^{l,i}) \right]}{\sum_{i=1}^C \sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \underline{\mathbf{m}}^{l,i})^T \underline{\Sigma}_{\underline{\sigma}^{l,i}}^{-1} (\mathbf{q}_n^f - \underline{\mathbf{m}}^{l,i}) \right] + \exp \left[-\frac{1}{2} (\mathbf{q}_n - \overline{\mathbf{m}}^{l,i})^T \overline{\Sigma}_{\overline{\sigma}^{l,i}}^{-1} (\mathbf{q}_n^f - \overline{\mathbf{m}}^{l,i}) \right]}, \quad (2.190)$$

onde $\underline{\mathbf{m}}^{l,i} = [\underline{m}_1^{l,i} \cdots \underline{m}_L^{l,i}]^T$, $\overline{\mathbf{m}}^{l,i} = [\overline{m}_1^{l,i} \cdots \overline{m}_L^{l,i}]^T$, $\underline{\Sigma}_{\underline{\sigma}^{l,i}} = \mathbf{I} \underline{\sigma}^{l,i}$, $\overline{\Sigma}_{\overline{\sigma}^{l,i}} = \mathbf{I} \overline{\sigma}^{l,i}$, $\underline{\sigma}^{l,i} = [\underline{\sigma}_1^{l,i} \cdots \underline{\sigma}_L^{l,i}]^T$, $\overline{\sigma}^{l,i} = [\overline{\sigma}_1^{l,i} \cdots \overline{\sigma}_L^{l,i}]^T$ e \mathbf{I} é a matriz identidade.

Caso não existam incertezas nas funções de pertinência, tem-se que a função característica do classificador fuzzy do tipo 2 é dada por

$$\mu_{\tilde{B}^i}(\mathbf{q}_n) = \frac{\sum_{l=1}^{M_i} f^{l,i}(\mathbf{q}_n)}{\sum_{i=1}^C \sum_{l=1}^{M_i} f^{l,i}(\mathbf{q}_n)} = \mu_{B^i}(\mathbf{q}_n), \quad (2.191)$$

pois

$$\underline{f}^{l,i}(\mathbf{q}_n) = \overline{f}^{l,i}(\mathbf{q}_n) = f^{l,i}(\mathbf{q}_n) = \mathcal{T}_{l=1}^L \mu_{F_k^{l,i}}(q_{n,k}). \quad (2.192)$$

Supondo-se que a função de pertinência do tipo 1, $\mu_{F_k^{l,i}}(q_{n,k})$, seja modelada pela função

$$\mu_{F_k^{l,i}}(q_{n,k}) = \exp \left[-\frac{1}{2} \left(\frac{q_{n,k} - m_k^{l,i}}{\sigma_k^{l,i}} \right)^2 \right], \quad (2.193)$$

então a saída do classificador fuzzy do tipo 1 e *singleton* é dado por

$$\mu_{\tilde{B}^i}(\mathbf{q}_n) = \frac{\sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \mathbf{m}^{l,i})^T \Sigma_{\sigma^{l,i}}^{-1} (\mathbf{q}_n^f - \mathbf{m}^{l,i}) \right]}{\sum_{i=1}^C \sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \mathbf{m}^{l,i})^T \Sigma_{\sigma^{l,i}}^{-1} (\mathbf{q}_n^f - \mathbf{m}^{l,i}) \right]}, \quad (2.194)$$

onde $\mathbf{m}^{l,i} = [m_1^{l,i} \cdots m_L^{l,i}]^T$, $\Sigma_{\sigma^{l,i}} = \mathbf{I}\sigma^{l,i}$ e $\sigma^{l,i} = [\sigma_1^{l,i} \cdots \sigma_L^{l,i}]^T$.

Agora, supondo-se que o classificador seja baseado na teoria de Bayes, no critério ML, e que a probabilidade à *posteriori* seja Gaussiana e $\Sigma_{\sigma^{l,i}} = \mathbf{I}\sigma^{l,i}$, então tem-se que a função característica do mesmo é expressa por

$$\mu_{Bayesi}(\mathbf{q}_n) = \frac{\sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \mathbf{m}^{l,i})^T \Sigma_{\sigma^{l,i}}^{-1} (\mathbf{q}_n^f - \mathbf{m}^{l,i}) \right]}{\sum_{i=1}^C \sum_{l=1}^{M_i} \exp \left[-\frac{1}{2} (\mathbf{q}_n - \mathbf{m}^{l,i})^T \Sigma_{\sigma^{l,i}}^{-1} (\mathbf{q}_n^f - \mathbf{m}^{l,i}) \right]}. \quad (2.195)$$

Em outras palavras, o classificador fuzzy do tipo 2 e *singleton* cuja função característica é dada por (2.186) reduz-se ao classificador fuzzy do tipo 1 e ao classificador Bayesiano quando algumas considerações nas funções de pertinência das premissas dos antecedentes das regras são feitas. Isso significa que tal classificador fuzzy do tipo 2 reduz-se ao ótimo (de acordo com a teoria de Bayes e o critério ML), desde que os vetores \mathbf{q}_n satisfaçam as restrições impostas pela teoria de Bayes e o critério ML. Para o caso mais geral, onde existam elevados níveis de incertezas, o sistema fuzzy do tipo 2 apresentará melhores desempenhos que o classificador Bayesiano, uma vez que o primeiro leva em consideração tais incertezas em sua formulação.

É importante salientar que se outras funções (ex. funções exponencial, de cauchy, ou outra função que se enquadre dentro da definição de função de pertinência em sistemas fuzzy e da definição de funções de densidade de probabilidade em sistemas probabilísticos) são utilizadas para definir as FOU das funções de pertinência do classificador fuzzy do tipo 2, cuja função característica é dada por (2.186), forem consideradas, tal classificador fuzzy do tipo 2 reduz-se a um classificador Bayesiano baseado no critério ML.

2.5 Sumário

Neste capítulo, uma breve introdução sobre os sistemas fuzzy do tipo 2 foi apresentada. Em seguida, dois algoritmos fuzzy do tipo 2 treinados pelo método GCEM foram introduzidos. Além disso, algoritmos fuzzy adaptativos, treinados pelos métodos LLE e DRD foram propostos. Finalmente, um classificador fuzzy do tipo 2 foi introduzido. Este classificador apresenta maior flexibilidade para trabalhar com a classificação de sistemas ou sinais que apresentem elevados níveis de incertezas e reduz-se ao classificador fuzzy do tipo 1 (Wei & Mendel, 2000) e ao classificador ótimo baseado na teoria de Bayes e no critério ML (Wei & Mendel, 1999) quando algumas restrições são feitas nos parâmetros do classificador proposto.

3

Transmissão de Dados Via Rede Elétrica

A crescente demanda por acesso rápido à internet e por novos serviços, o estágio atual de desenvolvimento de técnicas de processamento de sinais associado à processadores cada vez mais rápidos e baratos, assim como a desregulamentação do mercado de telecomunicações têm incentivado o uso de diversos tipos de meios para a transmissão banda larga e banda estreita de dados. Palavras tais como *anytime*, *anyone* e *anywhere* são, atualmente, amplamente empregadas para caracterizar o estágio atual de desenvolvimento do setor de telecomunicações. É sabido que o aumento e o alcance maior dos sistemas de telecomunicações passam primeiramente pela investigação de diferentes meios de comunicação de dados,

tendo como parâmetros de análise o desempenho, o custo, a facilidade de uso e a adequação para a aplicação em questão. Por estes motivos, especial atenção tem sido, recentemente, direcionada às redes elétricas como meio de transmissão de dados.

A investigação recente do uso das redes elétricas como meio de transmissão de dados é devido às seguintes vantagens: *i*) essas redes estão presentes em cerca de 95 % dos domicílios residenciais, industriais e comerciais (Ascom, 2001); *ii*) o baixo custo de implementação da infra-estrutura de comunicação (Dostert, 2001; Langfeld, 2001), pois a mesma já se encontra instalada; *iii*) a capacidade teórica dos canais PLC ultrapassa 500 Mbps (Zimmermann & Dostert, 1999; Sanderson, 2000); *iv*) o aumento explosivo da capacidade de processamento dos modernos DSP e FPGA e a diminuição constante do custo dos mesmos; *v*) o desenvolvimento de novas técnicas de transmissão digital de dados (Ascom, 2001; Dostert, 2001); *vi*) a possibilidade das empresas de distribuição de energia se tornarem empresas provedoras de acesso às redes de comunicação de informações e o oferecimento de seus inerentes serviços (Ascom, 2001; Dostert, 2001; Pavlidou, Vinck, Yazdani, & Honary, 2003).

O uso das redes elétricas para a transmissão de dados iniciou-se em 1838. Alguns avanços foram obtidos até o final da década de 70 do século XX. No entanto, somente após o início da década de 80 do século XX foi possível a implementação de técnicas de processamento de sinais para a modulação digital e a codificação de canais em modems que permitiram taxas em torno de 14,4 kbps. Nesse mesmo período, a Europa estabeleceu o padrão CELENEC que opera numa faixa de frequência de até 500 kHz e permite taxas de transmissão de dados de até 144 kbps para distâncias inferiores a 500 m e limita a potência média dos sinais transmitidos em 5 mW.

Ao longo da década de 90 do século XX algumas empresas propuseram o uso da faixa entre 1 MHz e 30 MHz para a transmissão banda larga de dados

(algumas dezenas de Mbps) (DS2, 2004; Intellon, 2004; Echelon, 2004; Ascom, 2003). Atualmente, investigações na faixa entre 1 MHz e 60 MHz são conduzidas com este propósito. Atualmente, enquanto a Europa está realizando um estudo técnico profundo para viabilizar as aplicações PLC em banda larga (OPERA, 2004), relatórios da FCC (*Federal Communications Commission*) dos EUA apontam para uma nova regulamentação para viabilizar o uso da tecnologia PLC em banda larga (FCC, 2004) e sua coexistência com outros tipos de tecnologias, já existentes, na mesma banda de frequência.

Informações mais recentes relatam o desenvolvimento da nova versão do padrão HomePlug versão 2 (Intellon, 2004) para a transmissão de dados via canais *indoor* (Lin, Latchman, Lee, & Katar, 2002; Lin, Latchman, Newman, & Katar, 2003), levando-se em conta, principalmente, as demandas atuais de países de primeiro mundo, as quais serão, no futuro próximo ou distante, a dos países em desenvolvimento e sub-desenvolvidos, necessariamente nessa ordem. Por exemplo, a nova versão do padrão HomePlug levará em conta aplicações de alta velocidade com taxas de pico e média iguais a 200 Mbps e 120 Mbps, tais como HDTV (*High Definition Television*), vídeo-conferência e outras aplicações multimídia, etc. As taxas de transmissão de dados da nova versão do padrão HomePlug será superior a dos atuais *chipsets* disponíveis no mercado, tais como o *chipset* DS2 (DS2, 2004) com taxas médias de 45 Mbps e o *chipset* da Intellon que oferece taxas médias de 14 Mbps. Este último implementa o padrão HomePlug 1.0 (Intellon, 2004).

É interessante ressaltar ainda que os canais PLC estão presentes em pelo menos 95 % das instalações residenciais, industriais e comerciais. Analisando-se apenas os canais PLC em ambiente prediais, estimativas recentes apontam capacidades de canais PLC superiores à 600 Mbps. Note que as atuais soluções possibilitam taxas médias de no máximo 45 Mbps.

Investigações sobre as redes elétricas de baixa, média e alta tensões (*last miles*) também estão sendo realizadas (Abad et al., 2003; Jee, Edison, Rao, & Cern, 2003; Lin et al., 2002, 2003). Para esses ambientes nenhuma padronização foi definida, visto que para as aplicações *last miles*, que caracterizam tais ambientes, outras tecnologias já estão estabelecidas, tais como *cable modem*, xDSL (*Digital Subscriber Line*), fibra óptica e comunicação sem fio. É importante enfatizar que os conceitos *last miles* ou *outdoor* estão associados ao acesso até o ponto de entrada do prédio ou da residência do consumidor, enquanto que o conceito *last meters*, *intrabuilding* ou *indoor* é empregado para definir a comunicação de dados nas instalações elétricas prediais, encontradas em ambientes comerciais, industriais e residenciais (Ascom, 2001).

Sabe-se que para as aplicações PLC nas redes elétricas de média e alta tensão, as taxas não são, ainda, tão altas como aquelas obtidas por meio de outras tecnologias, pois as distâncias consideradas para a transmissão de dados são grandes e as atenuações em redes elétricas aumentam consideravelmente quando a distância entre o receptor e o transmissor aumenta.

Em artigo recente a expressão *horrible channel* (Biglieri, 2003) foi usada para caracterizar os canais PLC. De fato, as investigações mostram que os canais PLC apresentam características similares aos canais de comunicação sem fio. De um modo geral, as seguintes características têm sido apresentadas como as principais responsáveis por essa caracterização dada aos canais PLC (Canete, Cortes, Diez, & Entrambasaguas, 2002, 2003; ?, ?): *i*) as variações entre as diferentes redes elétricas; *ii*) as diferentes condições físicas e características topológicas das redes elétricas; *iii*) as atenuações proporcionais à distância e à frequência; *iv*) as ocorrências de ruídos impulsivos com alta potência; *v*) as variações das impedâncias vistas pelo transmissor e pelo receptor; *vi*) as interferências eletromagnéticas geradas por outros serviços que operam na mesma faixa de frequência e, finalmente, *vii*) as variações

periódicas das respostas impulsivas dos canais PLC.

As técnicas de processamento de sinais são aplicadas para minimizar tais efeitos nocivos à transmissão de dados através de canais PLC. Dentre as técnicas recentemente investigadas, pode-se citar (Ribeiro, Lopes, Duque, & Romano, 2005; Ribeiro, 2003; Loiola, Ribeiro, & Romano, 2004; Langfeld, 2001; Biglieri, 2003; Day & Poor, 2003; Re, Fantacci, Morosi, & Serravalle, 2003; Baig & Gohar, 2003): *i*) técnicas mono-portadoras, multi-portadoras e de espalhamento espectral; *ii*) técnicas de codificação de canais; *iii*) técnicas de equalização; *iv*) técnicas de estimação de canais; *v*) técnicas de acesso ao meio; e *vi*) técnicas de cancelamento de ruídos impulsivos; *vii*) técnicas de gerenciamento de espectro. Algumas dessas técnicas foram implementadas nos *chipsets* DS2 e da Intellon. Além disso, melhorias das mesmas são investigadas com vistas à implementação de futuros *chipsets*.

É importante salientar que no Brasil as investigações dos canais PLC como meio de transmissão de dados em banda larga e banda estreita seguem duas linhas principais:

1) a análise de sistemas PLC desenvolvidos na Europa e nos EUA (Aranha, 2003; Romano, 2003). O uso de tais sistemas não é condicionado a nenhum tipo de pesquisa e/ou desenvolvimento que possibilite melhorias. Esses sistemas têm sido implementados em campos de prova para que as empresas do setor de energia tenham uma idéia da potencialidade real da tecnologia PLC. É de se esperar, no entanto, que os sistemas PLC desenvolvidos nos EUA e Europa apresentem desempenhos bastante diferentes, quando implementados no Brasil.

2) a análise dos sistemas PLC nas universidades (Batista, 2001; Rapp, 2002): Análises computacionais de desempenho de algoritmos no nível da camada física têm sido o principal foco destas investigações. Alguns trabalhos práticos têm sido realizados, mas limitam-se à montagem de soluções já prontas.

Pode-se afirmar que a tecnologia PLC possibilita desenvolver soluções que podem ser introduzidas em diferentes aplicações, gerando, como conseqüência, inovações tecnológicas bastante interessantes. Por exemplo, um robô de uma linha de produção que faz uso do cabo de alimentação para transmissão de dados de controle. É

As questões discutidas acima são bastante pertinentes quando se pretende desenvolver projetos de P&D, tendo como foco não apenas na produção científica na forma de artigos publicados em reuniões de caráter científico e em revista especializadas, mas também no desenvolvimento de novas técnicas e soluções que sejam absorvidas pelo setor produtivo.

No entanto, o desenvolvimento de tais soluções demanda primeiramente o conhecimento do estado da arte da tecnologia PLC (Pavlidou et al., 2003). Este conhecimento pode ser colocado da seguinte forma: *i*) Como se implementa um sistema completo de transmissão e recepção ponto a ponto e multi-ponto? *ii*) Quais são as técnicas atuais mais adequadas para lidar com os problemas que caracterizam as aplicações *indoor* e *outdoor* em banda larga ou banda estreita ? *iii*) Como desenvolver ou melhorar os sistemas de repetição de sinais? e, finalmente, *iv*) Como garantir o acesso de vários usuários a um mesmo meio?

O uso ótimo das redes elétricas de distribuição de energia para a transmissão de dados passa, primeiramente, pelo estudo das características inerentes a esses meios de transmissão. Tal estudo traz à luz as reais potencialidades e os principais problemas encontrados nas redes elétricas que dificultam a transmissão de dados a altas taxas e com baixa probabilidade de erro. Obtém-se, assim, o conhecimento necessário para o desenvolvimento de uma modelagem adequada dos canais PLC e dos ruídos presentes neste meio. Dentre os modelos de canais PLC encontradas na literatura, a seção 3.1 descreve o modelo de multi-percursos dos sinais transmitidos, o mais atual e coerente modelo para canais PLC *outdoor*. A seção 3.2 descreve

as características principais dos diversos tipos de ruídos presentes nos canais PLC. Finalmente, a seção 3.3 apresenta a técnica multi-portadora OFDM/DMT empregada em aplicações PLC banda larga.

3.1 Modelo dos Canais PLC *Outdoor*

De um modo geral, as redes de distribuição de energia diferem consideravelmente umas das outras, em termos de estrutura, topologia e propriedades físicas quando comparada aos convencionais meio *wireline*, tais como fio telefônico, par trançado e coaxial. Uma das características predominante e diferenciadora é o compartilhamento de um mesmo meio físico de comunicação por vários usuários e a presença de várias fontes de ruídos. Além disso, os canais PLC podem ser caracterizados como um meio termo entre um tradicional canal de telefonia fixa e um canal de comunicação sem fio. De fato, os mesmos apresentam desvanecimentos seletivos em frequência e severas atenuações do sinal em função da distância, além da presença de ruídos com nível de potência suficiente para corromper rajadas de transmissão de dados.

Do ponto de vista da transmissão banda larga via canais PLC, o enfoque tem sido a caracterização das respostas em frequência na faixa entre 500 kHz e 30 MHz. Nesse espectro de frequência disponível, investigações recentes apontam a faixa de frequência de 500 kHz a 10 MHz para aplicações *last miles* e a faixa entre 10 MHz e 30 MHz para aplicações *last meters* ou *intrabuilding* (Langfeld, 2001). A busca de uma modelagem consistente e representativa dos canais PLC para aplicações banda larga é baseada nas duas estratégias descritas a seguir (Zimmermann & Dostert, 1999, 2002a).

Estratégia *Bottom up*: nesse caso o comportamento das redes elétricas é descrito pelas matrizes de admitância e de impedância do circuito elétrico (Barnes, 1998; Esmailian, Kschischang, & Gulak, 2000; Bostoen & Wiel, 2000; Esmailian, Gulak, & Kschischang, 2000). Esse tipo de modelagem requer o conhecimento detalhado dos componentes da rede elétrica para determinar os elementos das matrizes. De um modo geral, essa modelagem requer a manipulação de uma quantidade razoavelmente elevada de parâmetros para que o modelo da rede elétrica tenha um nível de precisão razoável. É, normalmente, adotado para canais *indoor*, posto que os mesmos não apresentam elevado número de parâmetros;

Estratégia *Top down*: os métodos baseados nesta estratégia consideram os canais PLC como uma "caixa preta" e descrevem as características do canal em termos de sua função de transferência (Dostert, 2001; Langfeld, 2001; Zimmermann & Dostert, 1999, 2002a; Dostert, 1998; Waldeck, Busser, & Dostert, 1998; Tanaka, 1988). Esse tipo de modelo é baseado nas influências dos efeitos físicos do canal PLC no sinal transmitido. A modelagem dos canais PLC em termos da função de transferência possibilita uma representação simples e independente da topologia da rede elétrica (Zimmermann & Dostert, 1999, 2002a). Esta estratégia é adotada para a modelagem de canais PLC *outdoor*, pois tais canais são complexos e apresentam elevado número de parâmetros.

Levando-se em consideração a estratégia *top down*, alguns modelos de multi-propagação dos sinais em redes elétricas foram introduzidos para aplicações banda larga em ambientes *last miles* (Dostert, 2001; Zimmermann & Dostert, 1999, 2002a; Tanaka, 1988). O modelo de multi-propagação proposto em (Zimmermann & Dostert, 2002a) é o mais recente e atual modelo utilizado e por isso é resumidamente descrito abaixo, pois o mesmo é utilizado para a obtenção dos resultados mostrados

nos capítulos 4 e 5.

A resposta em frequência do canal PLC, obtida a partir da modelagem multi-percursos dos sinais transmitidos, é dada por (Zimmermann & Dostert, 2002a)

$$H(f) = \sum_{i=1}^N G_i(f, d_i) A(f, d_i) \exp(-j2\pi f \tau_i), \quad (3.1)$$

$$G_i(f, d_i) = |g_i(f)| e^{\varphi_{g_i}(f)}, \quad (3.2)$$

$$A(f, d_i) = \exp[-(a_0 + a_1 f^k) d_i], \quad (3.3)$$

e

$$\tau_i = \frac{d_i \sqrt{\epsilon_r}}{c_0}, \quad (3.4)$$

onde i representa o i -ésimo percurso de multi-propagação, a_0 e a_1 são os parâmetros de atenuação do i -ésimo percurso, k é o expoente do fator de atenuação, d_i é o comprimento do i -ésimo percurso, τ_i é o atraso devido ao i -ésimo percurso, c_0 é a velocidade da luz e, finalmente ϵ_r é a constante dielétrica do cabo. $A(f, d_i)$ é a atenuação gerada em função da distância no i -ésimo multi-percurso. A expressão (3.2), $G_i(f, d_i)$, representa o fator de ponderação no i -ésimo percurso. Tal valor é dado pelo produto dos fatores de transmissão e reflexão no i -ésimo percurso. O módulo da atenuação no i -ésimo percurso tem um limitante superior dado por

$$|G_i(f, d_i)| \leq 1. \quad (3.5)$$

Para percursos com número elevado de reflexões e transmissões $|G_i(f, d_i)| \rightarrow 0$ e, conseqüentemente, sua influência na modelagem dos canais PLC é insignificante. Desta forma, o número de percursos utilizados na modelagem dos canais PLC deve levar em consideração os percursos cujos fatores de ponderação são dominantes. Em

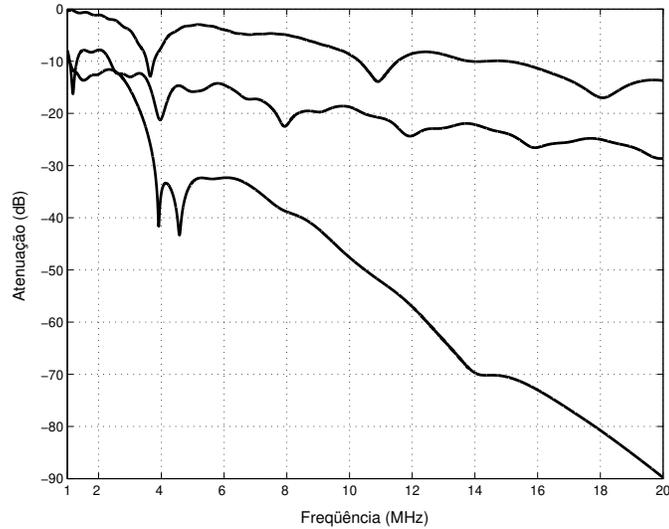


Figura 3.1: Módulo da Resposta em Freqüência de três canais PLC de redes de baixa tensão.

termos práticos, longos percursos são caracterizados por um número muito grande de transmissões e reflexões e, conseqüentemente, exibem grandes atenuações.

A Fig. 3.1 mostra as respostas em freqüência de alguns canais PLC, introduzidos em (Zimmermann & Dostert, 2002a) e (Langfeld, 2001), de redes de distribuição de baixa tensão. Conforme observado, os canais PLC variam consideravelmente, sendo dependentes das cargas conectadas, da distância entre o transmissor e o receptor e da faixa de freqüência utilizada.

3.2 Caracterização do Ruído Aditivo em Canais PLC *Outdoor*

Além da modelagem apropriada dos canais PLC, o desenvolvimento de soluções para a transmissão banda larga e confiável em ambientes *last miles* é dependente da caracterização dos ruídos presentes nestes canais, os quais são aditivos e de natureza não Gaussiana (Dostert, 2001; Zimmermann & Dostert, 2002b). Alguns estudos sobre os principais tipos de ruídos presentes em canais PLC foram realizados (Dostert, 2001; Tanaka, 1988; Brwon, 1998; Burr & Brwon, 1998; Zimmermann & Dostert, 2000a). Em (Dostert, 2001; Zimmermann & Dostert, 2002b) foi apresentada uma classificação dos ruídos presentes em canais PLC, a partir da proposta apresentada em (Hooijen, 1998) para canais PLC residenciais utilizados em aplicações banda estreita. Uma vez que tal modelo de ruído aditivo é o mais completo já proposto na literatura, o mesmo é descrito abaixo e é utilizado na simulações mostradas nos capítulos 4 e 5.

De acordo com (Dostert, 2001; Zimmermann & Dostert, 2002b), os ruídos presentes em canais PLC podem ser classificados conforme é mostrado na Fig. 3.2. O ruído aditivo na saída dos canais PLC é constituído de várias componentes e pode ser expresso por

$$v(t) = v_{bkgr}(t) + v_{nb}(t) + v_{pa}(t) + v_{ps}(t) + v_{imp}(t). \quad (3.6)$$

Cada uma das componentes do ruído aditivo reunida no lado direito de (3.6) é descrita a seguir.

Ruído colorido de fundo, $v_{bkgr}(t)$, tem potência espectral relativamente baixa e variante com a frequência. É normalmente causado pelo somatório de diferentes fontes de ruídos de baixa potência. Sua PSD (*Power Spectral Density*) varia com o tempo em termos de minuto ou até mesmo em termos de hora. Nessa contribuição,

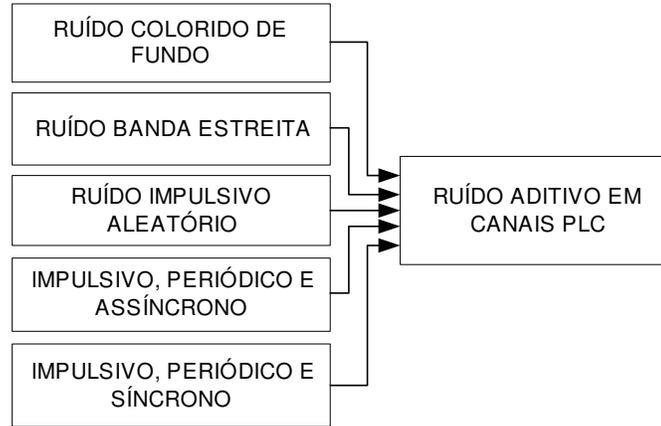


Figura 3.2: Ruídos aditivos em canais PLC *outdoor*.

adotou-se a PSD dada por

$$S_{n_{bkg.}}(f) = a + b|f|^c \text{ dBm/Hz}, \quad (3.7)$$

onde os parâmetros a , b e c usados para modelar o ruído colorido nos canais PLC *outdoor* são dados na Tab. 3.1, vide (Esmailian et al., 2000). Note que o modelo desenvolvido para canais *indoor*, (3.7), foi adotado para canais *outdoor*.

Parâmetros	a	b	c
Melhor Caso	-140	38,75	-0,720
Pior Caso	-145	53,23	-0,337

Tabela 3.1: Parâmetros da PSD do ruído colorido de fundo.

Ruídos em banda estreita, $v_{nb}(t)$, são caracterizados por sinais senoidais modulados em amplitude e causados, principalmente, pela indução dos sinais de rádio nos fios da rede elétrica na banda de frequência de interesse, ou seja, até 30 MHz. A amplitude desse ruído é dependente do período do dia. Por exemplo, durante o período da manhã esse tipo de ruído é bastante observado, pois as condições de propagação das ondas curtas de rádio são as melhores. Por outro

lado, durante a noite, a sua potência é bastante baixa. O modelo

$$n_{nb}(t) = \sum_{i=1}^{L_{nb}} A_i(t) \sin(\omega_i t + \theta_i) \quad (3.8)$$

(Zimmermann & Dostert, 2002b) é bastante razoável para modelar esse tipo de ruído. Note que a amplitude da i -ésima fonte geradora de ruído em banda estreita, $A_i(t)$, pode ser considerada constante ou modulada de forma a se obter a melhor representação dos sinais AM (*Amplitude Modulation*) de rádio. Investigações nas redes elétricas sugerem que as amplitudes dos ruídos em banda estreita variam lentamente durante um longo período de tempo e, conseqüentemente, uma modelagem com parâmetros fixos é recomendada sem perda de generalidade (Zimmermann & Dostert, 2002b).

Ruídos impulsivos periódicos e assíncronos à frequência fundamental, $v_{pa}(t)$, têm taxa de repetição, na maioria dos casos, entre 50 kHz e 200 kHz. Apresenta espectro com linhas discretas cujo espaçamento espectral é dado pela taxa de repetição dos mesmos. Eles são causados, principalmente, pelos chaveamentos dos geradores de energia.

Ruídos impulsivos periódicos e sincronizados com a frequência fundamental, $v_{ps}(t)$, têm uma taxa de repetição de 60 Hz e 120 Hz e são sincronizados com a frequência da componente fundamental da rede elétrica. Eles são de curta duração (alguns microsegundos) e têm densidade espectral de potência decrescente com a frequência. Eles são principalmente, causados pelo chaveamento dos retificadores de tensão dos conversores eletrônicos. São variantes no tempo.

Ruídos impulsivos assíncronos, $v_{imp}(t)$, são causados, principalmente, pelos transitórios gerados por chaveamentos nas redes elétricas. Eles têm tempo de duração na ordem de alguns microsegundos até alguns milisegundos e apresentam tempos aleatórios de chegada. A PSD dos ruídos impulsivos assíncronos pode alcançar 50 dB acima da PSD dos ruídos de fundo. São altamente variantes no tempo

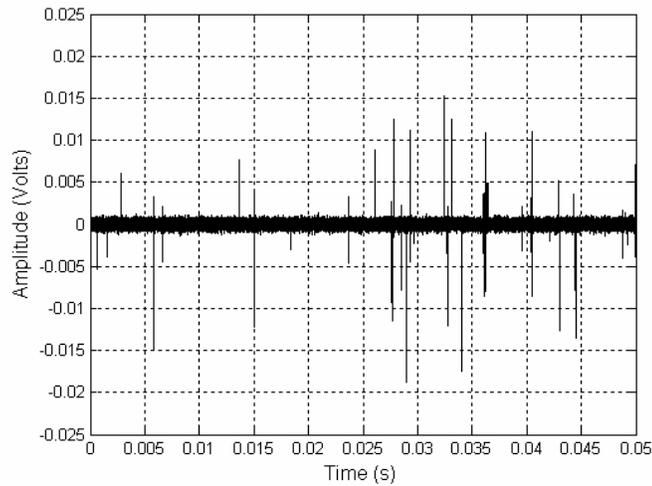


Figura 3.3: Ruído aditivo de acordo com o modelo proposto por Zimmermann e Dostert.

e, juntamente com o ruído anterior, podem causar erros em alguns bits ou rajadas de bits transmitidos. Uma descrição matemática desse tipo de ruído baseada na cadeia de Markov Particionada (Fritchman, 1967) foi elaborada por (Zimmermann & Dostert, 2002b, 2000b). Este tipo de modelagem separa os n estados de ruídos em dois grupos distintos: um grupo de estados sem distúrbios e um outro grupo de estados com distúrbios.

A Fig. 3.3 mostra o ruído aditivo constituído pelo ruído de fundo e pelos ruídos impulsivos, que foram gerados a partir do modelo discutido acima, na faixa de frequência entre 2 e 3 MHz. As PSDs do ruído de fundo e dos ruídos impulsivos são, respectivamente, iguais a -130 dBV²/Hz e -110 dBV²/Hz. Conforme pode ser observado, a amplitude máxima dos ruídos impulsivos é menor que 20 mV.

3.3 Técnica OFDM/DMT

A técnica OFDM (*Orthogonal Frequency Division Multiplexing*) é uma técnica multi-portadora que divide a largura de banda disponível em várias sub-bandas através do emprego de sub-portadoras ortogonais. A origem da técnica OFDM remonta os trabalhos (Chang, 1966; Satzberg, 1967; Weinstein & Ebert, 1971; Peled & Ruiz, 1980). O nome OFDM é usado para designar esse tipo de sistema multi-portadora em aplicações sem fio, enquanto o nome DMT (*Discrete multitone Transceiver*) é empregado para aplicações com fio. Ambas as técnicas OFDM e DMT implementam as mesmas sub-portadoras ortogonais para dividir a largura de banda do canal em várias sub-bandas. A diferença entre essas técnicas reside, principalmente, no fato de que a técnica DMT faz uso de algoritmos de otimização para a alocação ótima dos bits nas sub-bandas, enquanto a técnica OFDM aloca um mesmo número de bits em todas as sub-bandas. Além disso, a técnica DMT é utilizada em canais bidirecionais que variam lentamente com o tempo, tais como cable TV, canal telefônico e certas classes de canais PLC. Por outro lado, a técnica OFDM é usada em aplicações unidirecionais cujos canais variam rapidamente com o tempo, tais como os canais de comunicação sem fio. Informações detalhadas sobre OFDM e DMT podem ser encontradas em (Bingham, 1990, 2000; Starr, Cioffi, & Silverman, 1998; Hanzo, Münster, Choi, & Keller, 2003; Hanzo, Wong, & Yee, 2002; Bahai, Satzberg, & Ergen, 2004; Hanzo, Webb, & Keller, 2000).

A principal vantagem das técnicas multiportadoras OFDM/DMT consiste no fato de que a mesma minimiza a ocorrência da interferência inter-simbólica, uma vez que o tempo de duração da forma de onda do sinal que representa o símbolo OFDM é muito maior que o tempo de duração da resposta impulsiva do canal de comunicação. Diferentemente dos sistemas SISO (*Single Input Single Output*) que transmitem uma seqüência de N símbolos com uma única portadora e ocupa toda a banda disponível, W , por um período $T = 1/W$, nos sistemas MIMO (*Multiple*

Input Multiple Output) a largura de banda é utilizada para transmitir paralelamente N símbolos, sendo que cada um dos símbolos ocupa a largura de banda igual a W/N por um período igual a N/W . Note que ambas as abordagens usam a mesma largura de banda disponível.

Para eliminar completamente a presença da IIS com a técnica OFDM/DMT, basta utilizar um período de guarda entre os símbolos OFDM consecutivos, cujo comprimento é maior que o comprimento da resposta impulsiva do canal de comunicação. Além de reduzir a IIS, a interferência entre as sub-portadoras (*inter-carrier interference* ICI) pode ser removida se o período de guarda é constituído de uma cópia da parte final do símbolo OFDM. Neste caso, o período de guarda, também chamado de prefixo cíclico (*Cyclic Prefix* - CP), é removido no receptor (Peled & Ruiz, 1980). Dentre as várias soluções propostas para a obtenção das N sub-portadoras ortogonais para a modulação e demodulação dos dados em sistemas multiportadoras, os sistemas OFDM/DMT baseados na DFT (*Discrete Fourier Transform*) (Weinstein & Ebert, 1971) são os que apresentam o menor custo computacional e são assintoticamente ótimos, em termos de desempenho, quando $N \rightarrow \infty$. A DFT elimina a necessidade do uso de um banco de osciladores, os quais são normalmente difíceis de serem sintonizados para gerar as sub-portadoras. A recente popularidade da OFDM deve-se, em parte, ao fato de que a DFT pode ser eficientemente implementada com algoritmos rápidos, tal como a FFT (*Fast Fourier Transform*) que demanda $\log N$ operações por sub-portadora.

As Figs. 3.4 e 3.5 mostram os modelos banda base dos sistemas multi-portadoras OFDM e DMT. Note que o bloco TEQ do sistema DMT implementa um equalizador para encurtar o comprimento da resposta impulsiva do canal de comunicação (Starr et al., 1998). A formulação do sistema OFDM será apresentada a seguir, posto que a mesma é válida para o sistema DMT com pouquíssimas modificações.

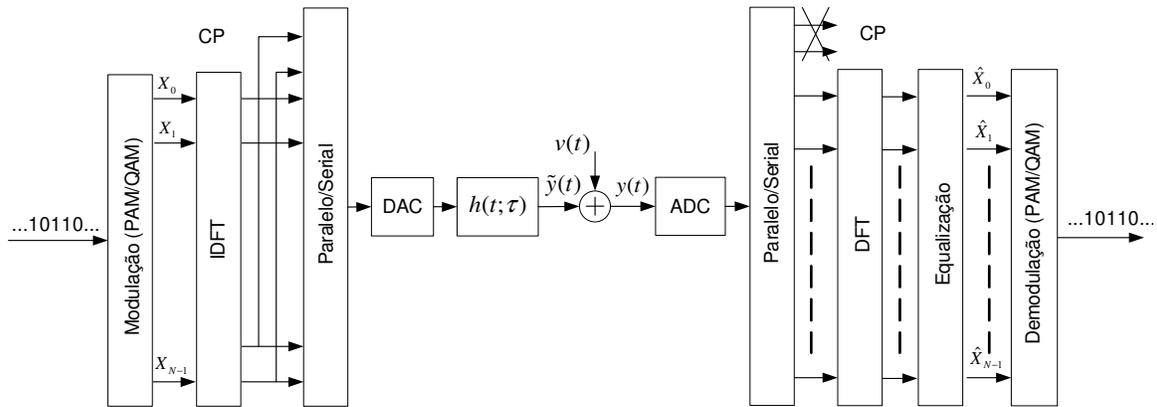


Figura 3.4: Modelo banda base do sistema de transmissão multi-portadoras OFDM.

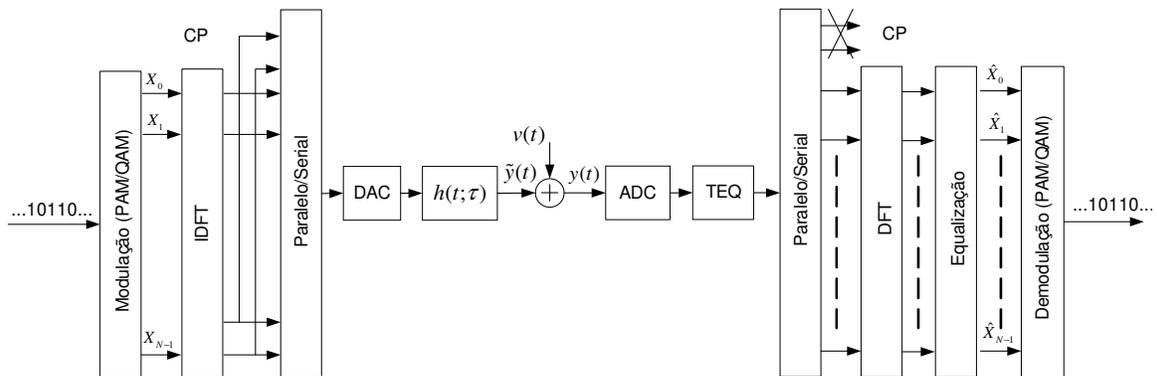


Figura 3.5: Modelo banda base do sistema de transmissão multi-portadoras DMT.

O sinal recebido na banda base do sistema OFDM é expresso por

$$y(t) = \tilde{y}(t) + v(t) = x(t) * h(t) + v(t), \quad (3.9)$$

onde $h(t)$ é a resposta impulsiva do canal de comunicação que é suposto ser invariante no tempo e $v(t)$ é o ruído aditivo. No caso dos canais PLC, o ruído aditivo é dado por

$$v(t) = v_{bkgr}(t) + v_{nb}(t) + v_{pa}(t) + v_{ps}(t) + v_{imp}(t) \quad (3.10)$$

onde $v_{bkgr}(t)$ é o ruído de fundo aditivo, $v_{nb}(t)$ é o ruído banda-estreita, $v_{pa}(t)$ é o ruído impulsivo periódico e síncrono à frequência da fundamental, $v_{ps}(t)$ é o ruído impulsivo periódico e assíncrono à frequência da componente fundamental e $v_{imp}(t)$ é o ruído impulsivo.

O sinal recebido pelo receptor é demodulado com a DFT para produzir os N símbolos na frequência, os quais são dados por

$$Y_k = \tilde{Y}_k + V_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y(n/f_s) e^{j \frac{-2\pi nk}{N}}, \quad 0 \leq k \leq N-1 \quad (3.11)$$

onde k é o índice da sub-portadora e $f_s = 1/T$,

$$\tilde{Y}_k = H_k X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [x(n/f_s) * h(n/f_s)] e^{j \frac{-2\pi nk}{N}} \quad (3.12)$$

e

$$V_k = V_{bkgr, k} + V_{nb, k} + V_{pa, k} + V_{ps, k} + V_{imp, k}, \quad (3.13)$$

sendo

$$V_{bkgr, k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [v_{bkgr}(n/f_s)] e^{j \frac{-2\pi nk}{N}}, \quad (3.14)$$

$$V_{nb, k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [v_{nb}(n/f_s)] e^{j \frac{-2\pi nk}{N}}, \quad (3.15)$$

$$V_{pa, k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [v_{pa}(n/f_s)] e^{j \frac{-2\pi nk}{N}}, \quad (3.16)$$

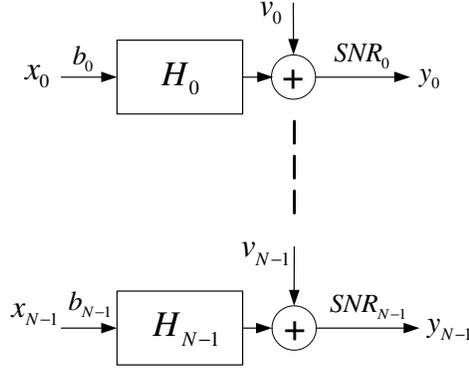


Figura 3.6: Canais paralelos obtidos com o sistema multi-portadora OFDM/DMT.

$$V_{ps, k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [v_{ps}(n/f_s)] e^{j\frac{-2\pi nk}{N}} \quad (3.17)$$

e

$$V_{imp, k} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} [v_{imp}(n/f_s)] e^{j\frac{-2\pi nk}{N}}. \quad (3.18)$$

Note que

$$H_k = \int_0^{L_h/f_s} h(t) e^{-j\frac{2\pi k f_s}{N} t} dt = H \left(k \frac{f_s}{N} \right) \quad (3.19)$$

é a atenuação complexa na k sub-portadora do canal PLC, $f_k = k f_s / N$, L_h é o comprimento do canal PLC. A Fig. 3.6 mostra os N sub-canais independentes e paralelos no sistema OFDM.

A SNR na k -ésima sub-portadora é dada por

$$SNR_k = \frac{\mathcal{S}_k |H_k|^2}{\mathcal{S}_{V_k}} = \frac{\mathcal{S}_k |H_k|^2}{\mathcal{S}_{V_{bkg, k}} + \mathcal{S}_{V_{nb, k}} + \mathcal{S}_{V_{ps, k}} + \mathcal{S}_{V_{pa, k}} + \mathcal{S}_{V_{imp, k}}}, \quad (3.20)$$

onde \mathcal{S}_k é a PSD do sinal transmitido na k -ésima sub-banda, $\mathcal{S}_{V_{bkg, k}}$, $\mathcal{S}_{V_{pa, k}}$, $\mathcal{S}_{V_{ps, k}}$ e $\mathcal{S}_{V_{imp, k}}$ são as PSDs das componentes $V_{bkg, k}$, $V_{pa, k}$, $V_{ps, k}$ e $V_{imp, k}$ do ruído aditivo na k -ésima sub-banda.

No sistema multi-portadora, a k -ésima sub-portadora transmite b_k bits/símbolo

ou \bar{b}_k bits/dimensão. O número total de bits transmitidos pelo sistema multiportadora é dado por

$$b = \sum_{k=0}^{N-1} M\bar{b}_k, \quad (3.21)$$

onde M é o número de dimensões,

$$\bar{b}_k = \frac{1}{2} \log_2 \left(1 + \frac{SNR_k}{\Gamma} \right) \quad (3.22)$$

e Γ é um valor que garante que o sistema irá trabalhar num certo com uma probabilidade de erro P_e .

A taxa de transmissão é expressa por

$$R = \frac{b}{T} = \sum_{k=0}^{N-1} R_k, \quad (3.23)$$

sendo $R_k = M\bar{b}_k/T$.

Para o conjunto de N sub-canais paralelos, o número médio de bits/símbolo por sub-canal é expresso por

$$\bar{b}_m = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{2} \log_2 \left(1 + \frac{SNR_k}{\Gamma} \right). \quad (3.24)$$

Considerando a introdução de margem, γ_m , para garantir o mesmo desempenho do sistema se a potência do sinal recebido cai γ_m dBs e do ganho de codificação, γ_c , obtido com o uso de técnicas de codificação de canais para garantir o funcionamento do sistema dentro da região de desempenho definida pelos parâmetros de desempenho, a relação (3.24) é dada por

$$\bar{b}_m = \frac{1}{N} \sum_{k=0}^{N-1} \bar{b}_k = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{2} \log_2 \left(1 + \frac{SNR_k \gamma_c}{\Gamma \gamma_m} \right). \quad (3.25)$$

Por definição, a margem, γ_m , fornece o valor que a expressão (3.20) pode ser reduzida ou aumentada (neste caso para valores negativos da margem em dB) e, ainda, manter a probabilidade de erro, P_e , abaixo do valor máximo permitido para um esquema de

transmissão em sub-canal AWGN (*Additive White Gaussian Noise*) com um certo número de bits/dim, um certo esquema de codificação de canal e valor P_e dado pelo *gap*, Γ . Note que o *gap*, Γ , é uma aproximação que especifica o número de bits que pode ser transmitido com as modulações PAM (*Pulse Amplitude Modulation*) ou QAM (*Quadrature Amplitude Modulation*) com uma certa P_e (probabilidade de erro de bit). Por exemplo, para garantir $P_e = 10^{-6}$ deve-se considerar $\Gamma = 8,8$ dB e para $P_e = 10^{-7}$ usa-se $\Gamma = 9,8$ dB.

É importante salientar que se $L_{CP} \geq L_h - 1$, onde L_{CP} é o comprimento do prefixo cíclico e L_h é o comprimento da resposta impulsiva do canal no domínio discreto, então a equalização do canal PLC é simples, posto que cada sub-canal é somente afetado por um ganho complexo. A formulação apresentada abaixo fundamenta esta característica do sistema OFDM/DMT.

Sejam $\mathbf{x} = [x_0 \cdots x_{N-1}]^T$ um vetor de símbolos PAM ou QAM, $\mathbf{h} = [h_0 \cdots h_{L_h-1}]^T$ a resposta impulsiva do canal PLC no domínio discreto, $\mathbf{v} = [v_0 \cdots v_{N-1}]^T$ o vetor de ruído na saída do canal, $L_{CP} = L_h$ o comprimento do prefixo cíclico,

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \quad (3.26)$$

e

$$\mathbf{Q}^* = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^{-1} & W_N^{-2} & \cdots & W_N^{-(N-1)} \\ 1 & W_N^{-2} & W_N^{-4} & \cdots & W_N^{-2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{-(N-1)} & W_N^{-2(N-1)} & \cdots & W_N^{-(N-1)(N-1)} \end{bmatrix}, \quad (3.27)$$

onde

$$W_N^k = e^{-j\frac{2\pi k}{N}}. \quad (3.28)$$

Então a saída do canal PLC é dada por

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_2 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{L_h-1} & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & h_{L_h-2} & h_{L_h-1} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_{L_h-1} \\ h_{L_h-1} & 0 & \cdots & 0 & h_0 & \cdots & h_{L_h-2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_1 & \cdots & h_{L_h-1} & 0 & \cdots & 0 & h_0 \end{bmatrix} \begin{bmatrix} x_{-v+1} \\ \vdots \\ x_{N-1} \end{bmatrix} + \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix} \quad (3.29)$$

ou por

$$\mathbf{y} = \tilde{\mathbf{y}} + \mathbf{v} = \mathcal{H}\mathbf{x} + \mathbf{v}, \quad (3.30)$$

onde $x_{-k} = x_{N-k}$, $k = 1, \dots, L_h - 1$.

Como a matriz de convolução do canal \mathcal{H} é circular, ou seja,

$$\mathcal{H} = \mathbf{Q}^* \Lambda \mathbf{Q}, \quad (3.31)$$

onde $\Lambda = \text{diag}\{\lambda_0, \dots, \lambda_{N-1}\}$, $\lambda_k = H_k(kf_s/N)$ e

$$\mathbf{x} = \mathbf{Q}^* \mathbf{X}, \quad (3.32)$$

então

$$\begin{aligned} \mathbf{Y} &= \mathbf{Q}\mathbf{y} = \mathbf{Q}(\mathbf{Q}^* \Lambda \mathbf{Q})\mathbf{Q}^* \mathbf{X} + \mathbf{Q}\mathbf{v} \\ &= \Lambda \mathbf{X} + \mathbf{V} \end{aligned}, \quad (3.33)$$

ou

$$\mathbf{Y} = \mathbf{\Lambda}\mathbf{X} + \mathbf{V} = \begin{bmatrix} H_0 & 0 & \cdots & 0 \\ 0 & H_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & H_{N-1} \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{bmatrix} + \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{N-1} \end{bmatrix} = \begin{bmatrix} X_0 H_0 \\ X_1 H_1 \\ \vdots \\ X_{N-1} H_{N-1} \end{bmatrix} + \begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{N-1} \end{bmatrix} \quad (3.34)$$

sendo $\mathbf{Y} = [Y_0 \cdots Y_{N-1}]^T$, $\mathbf{V} = [V_0 \cdots V_{N-1}]^T$. Note que \mathbf{X} é o vetor formado pelos símbolos obtidos na saída do bloco modulação, \mathbf{x} é o vetor obtido após a aplicação da IDFT ao vetor \mathbf{X} . Em (3.34) fica claro que se a matriz de convolução do canal, \mathcal{H} , é circular, então a matriz $\mathbf{\Lambda} = \text{diag} \left\{ H_0 \ \cdots \ H_{N-1} \right\}$ é constituída pelas amostras obtidas com a expressão (3.19). A estimativa do símbolo X_i , $i = 0, \dots, N-1$ é, então, dada por

$$\hat{\mathbf{X}} = \mathbf{Y} \odot \begin{bmatrix} 1/H_0 \\ \vdots \\ 1/H_{N-1} \end{bmatrix} = \begin{bmatrix} X_0 + V_0/H_0 \\ X_1 + V_1/H_1 \\ \vdots \\ X_{N-1} + V_{N-1}/H_{N-1} \end{bmatrix}, \quad (3.35)$$

onde \odot é o operador de Hadarmard definido por

$$\mathbf{p} \odot \mathbf{q} = \left[p_0 q_0 \ \cdots \ p_{L-1} q_{L-1} \right]^T. \quad (3.36)$$

Note que, no caso da aplicação DMT, o vetor $\mathbf{X} = [X_0 \cdots X_{N-1}]^T$ é dado por

$$X_i = \begin{cases} X_1(i) & , i = 1, \dots, N/2 - 1 \\ \Re(X_1(N)) & , i = 0 \\ \Im(X_1(N)) & , i = N/2 \\ X_1^*(N-i) & , i = N/2 + 1, \dots, N/2 - 1 \end{cases}. \quad (3.37)$$

De acordo com (3.35) a estimação dos símbolos transmitidos é degradada quando a potência do ruído aditivo é alta e/ou quando $H_i \rightarrow 0$, $i = 0, \dots, N-1$. O problema

que surge quando $H_i \rightarrow 0$, $i = 0, \dots, N - 1$ já tem sido, há algum tempo, tratado na literatura. No entanto, o problema do ruído aditivo constituído de componentes impulsivas de elevada potência ainda não foi solucionado. Assim, o capítulo 4, que segue, tem como foco apresentar o uso de técnicas de inteligência computacional para minimizar a potência de ruídos impulsivos na saída de canais PLC, uma vez que a presença de ruídos impulsivos gera elevadas taxas de erro na transmissão de dados via canais PLC.

Finalmente, um simulador de canais PLC para a obtenção dos resultados numéricos reunidos nos capítulos 4 e 5 é implementado. Esse simulador faz uso dos sistemas multi-portadora DMT/OFDM, descrito acima, e mono-portadora (Proakis, 2000). Os modelos de canais PLC e de ruídos aditivos, utilizados nesse simulador, são os modelos de canais PLC e do ruído aditivo para ambientes *outdoor* discutidos nas seções 3.1 e 3.2, e os modelos de canais PLC e do ruído aditivo para ambientes *indoor* que foram implementados em (Mendes, 2004).

3.4 Sumário

Esse capítulo discutiu modelos de canais e ruídos impulsivos para aplicações PLC *outdoor* que serão utilizados nos capítulos 4 e 5. Além disso, uma discussão sucinta sobre o sistema multi-portadora OFDM/DMT foi apresentada e brevemente discutido o problema causado pelos ruídos impulsivos em sistema multi-portadora OFDM/DMT. Finalmente, foi comentado a elaboração de um simulador de canais PLC para aplicações banda larga em ambientes *outdoor* e *indoor*.

4

Redução da Presença de Ruídos Impulsivos na Saída de Canais de Comunicação

Nesse capítulo as definições são consideradas: i) a saída do canal de comunicação é dada por

$$y(n) = \tilde{y}(n) + v(n) = \sum_{i=0}^{L_h-1} h(i)x(n-i) + v(n), \quad (4.1)$$

onde $x(n)$ são as amostras dos símbolos OFDM transmitidos, $\tilde{y}(n)$ são amostras livre de ruído na saída do canal, $v(n)$ são as amostras discretas do ruído aditivo em canais PLC, vide a expressão (3.9) e $\{h(n)\}_{n=0}^{L_h-1}$ é a versão discreta da resposta impulsiva do canal PLC, cuja resposta em frequência é dada por (3.1); *ii*) $\mathbf{w}(n) = [w(n) \cdots w(n - L_w + 1)]^T$ é a resposta impulsiva do equalizador implementado no bloco TEQ do sistema DMT para diminuir o comprimento da resposta impulsiva do canal PLC; *iii*) $\mathbf{H} = [H_0 \cdots H_{N-1}]^T$ e $\mathbf{W} = [W_0 \cdots W_{N-1}]^T$ são os vetores constituídos pelos valores da resposta em frequência do canal PLC e do equalizador obtidas, respectivamente, com $H_k = \sum_{n=0}^{N-1} h(n)e^{-j(2\pi/N)kn}$, $k = 0, \dots, N - 1$ e $W_k = \sum_{i=0}^{N-1} w(n)e^{-j(2\pi/N)kn}$, $k = 0, \dots, N - 1$; *iv*) $\mathbf{c}(n) = [c(n) \cdots c(n - L_c + 1)]^T$ é a resposta impulsiva resultante entre a convolução das respostas impulsivas do canal PLC e do equalizador; *v*) uma vez que o equalizador não é ideal tem-se que $\mathbf{c}(n) = \mathbf{c}^{Sinal}(n) + \mathbf{c}^{IIS}(n)$, onde $\mathbf{c}^{Sinal}(n) = [\mathbf{0}_{L_c}^T c(n - L) \cdots c(n - L - L_b + 1) \mathbf{0}_{L_c - L - L_b}^T]^T$ é a resposta impulsiva da componente de $\mathbf{c}(n)$ que não gera interferência intersimbólica, $\mathbf{c}^{IIS}(n) = [c(n - L) \cdots c(n - L + 1) \mathbf{0}_{L_b}^T c(n - L - L_b + 1) \cdots c(n - L_c + 1)]^T$. Note que $\mathbf{c}^{Sinal}(n) = [\mathbf{0}_{L_c}^T c(n - L) \cdots c(n - L - L_b + 1) \mathbf{0}_{L_c - L - L_b}^T]^T$ é a componente da resposta impulsiva $\mathbf{c}(n)$ que gera interferência intersimbólica e $\mathbf{0}_M$ é um vetor constituído de M zeros; *vi*) As k -ésimas amostras C_k^{Sinal} e C_k^{IIS} do vetor $\mathbf{C} = [C_0 \cdots C_{N-1}]^T = [C_0^{Sinal} + C_0^{IIS} \cdots C_{N-1}^{Sinal} + C_{N-1}^{IIS}]^T$ são obtidas a partir das seguintes respostas em frequência

$$C_k^{Sinal} = \sum_{n=0}^{N-1} c^{Sinal}(n)e^{-j(2\pi/N)kn}, \quad n = 0, \dots, N - 1, \quad (4.2)$$

e

$$C_k^{IIS} = \sum_{n=0}^{N-1} c^{IIS}(n)e^{-j(2\pi/N)kn}, \quad n = 0, \dots, N - 1; \quad (4.3)$$

vii) $v(n) = v_{bkgr}(n) + v_{nb}(n) + v_{pa}(n) + v_{ps}(n) + v_{imp}(n)$ é a versão discreta do ruído de fundo em canais PLC expresso por (3.6), as respostas em frequência das seqüências $\{v(n)\}_{n=0}^{N-1}$, $\{v_{bkgr}(n)\}_{n=0}^{N-1}$, $\{v_{nb}(n)\}_{n=0}^{N-1}$, $\{v_{pa}(n)\}_{n=0}^{N-1}$, $\{v_{ps}(n)\}_{n=0}^{N-1}$ e $\{v_{imp}(n)\}_{n=0}^{N-1}$ são expressas, respectivamente, pelas expressões (3.13)-(3.18), *viii*) as PSDs na k -ésima sub-banda dos sinais $x(n)$, $v(n)$, $v_{bkgr}(n)$, $v_{nb}(n)$, $v_{pa}(n)$, $v_{ps}(n)$ e $v_{imp}(n)$ são

representadas por \mathcal{S}_k , $\mathcal{S}_{bkg,k}$, $\mathcal{S}_{nb,k}$, $\mathcal{S}_{pa,k}$, $\mathcal{S}_{ps,k}$ e $\mathcal{S}_{imp,k}$, respectivamente.

De um modo geral, os sistemas multi-portadoras OFDM/DMT são menos sensíveis à presença de ruídos impulsivos que os sistemas mono-portadoras. Essa característica deve-se ao fato de que os símbolos OFDM são mais longos que os símbolos dos sistemas mono-portadoras e, conseqüentemente, a potência dos ruídos impulsivos em sistemas OFDM/DMT é distribuída entre as várias sub-bandas. Entretanto, resultados recentes evidenciam que os ruídos impulsivos podem prejudicar o desempenho dos sistemas multi-portadoras. De fato, as atenuações severas em algumas sub-bandas degradam o desempenho médio dos sistemas multi-portadoras e requerem o uso de técnicas poderosas de codificação de canal para que a confiabilidade do sistema seja garantida.

Na presença de ruídos impulsivos na saída dos canais PLC, a não utilização de técnicas sofisticadas para a codificação de canais resulta na ocorrência de rajadas de erros ou erros isolados. Tais erros afetam o desempenho dos decodificadores de canais, posto que os mesmos dependem da seqüência de símbolos recebida para a correta decodificação do canal. Desta forma, o uso de decodificadores robustos à presença de ruídos impulsivos é necessário para minimizar o impacto dos erros que ocorrem no processo de decodificação do canal.

Os efeitos dos ruídos impulsivos em sistemas multi-portadoras foram analisados em (Giles, 1995; Ghosh, 1996; Tan, Chew, Tjhung, & Wu, 2003) e várias soluções para minimizar a presença dos mesmos foram desenvolvidas desde então. Dentre elas, destaca-se a solução desenvolvida para identificar a ocorrência de picos na saída dos canais PLC e reduzir a presença dos ruídos impulsivos através de dois tipos de funções não-lineares (Giles, 1995). A primeira função não-linear é dada por

$$\hat{y}(n) = \begin{cases} y(n), & \text{se } |y(n)| < \mathcal{A} \\ \mathcal{A} \exp \{ \angle y(n) \}, & \text{caso contrário} \end{cases}, \quad n = 0, \dots, N - 1, \quad (4.4)$$

onde \mathcal{A} é o módulo da amplitude máxima permitida na saída do canal e $\hat{y}(n)$ é a estimativa da saída do canal livre de ruído. Conforme observado, o limitador dado por (4.4) reduz os efeitos decorrentes da presença de ruídos impulsivos com amplitudes elevadas. A outra função não-linear proposta para limitar a presença de ruídos impulsivos na saída de canais de comunicação é expressa por

$$\hat{y}(n) = \begin{cases} y(n), & \text{se } |y(n)| < \mathcal{A} \\ 0, & \text{caso contrário} \end{cases}, \quad n = 0, \dots, N - 1. \quad (4.5)$$

Em (Zhidkov, 2003) foi proposto uma técnica para estimar o ruído impulsivo e realizar a subtração do mesmo no domínio da frequência. No entanto, conforme observado pelo autor deste trabalho, o desempenho desta técnica é degradado severamente na presença de ruídos impulsivos com amplitudes elevadas.

Além das técnicas descritas acima, o uso de técnicas de codificação de canais têm sido investigado recentemente de maneira intensa, objetivando-se a minimização da presença de ruídos impulsivos em sistemas multi-portadoras (Haring & Vinck, 2003, 2004; Toumpakaris, Cioffi, & Gardan, 2004; Faber, Scholand, & Jung, 2003; Haring & Vinck, 2003; Day & Poor, 2003; Dai & Poor, 2003; Zhang & Blum, 2001; He, Tjhung, Chew, & Ko, 2000; Stopler & Zamir, 2001; Wang & Chen, 2000; Modestino, Sargrad, & Bollen, 1988; Moulin, Ouzzif, Zeddou, & Gauthier, 2003; Baig & Gohar, 2003). No entanto, dois fatores impedem o uso exclusivo de tais técnicas para o cancelamento de ruídos impulsivos. Primeiro, a ocorrência de ruídos impulsivos demanda o uso de técnica de entrelaçamento com comprimento muito grande, o que acarreta o aumento do atraso fim a fim e, conseqüentemente, dificulta aplicações em tempo real. Um segundo fator a ser considerado, conforme discutido anteriormente, é o fato de que os erros gerados nos decodificadores pelos ruídos impulsivos deterioram o desempenho dos mesmos.

Com o intuito de reduzir a presença de ruídos impulsivos em sistemas multi-portadoras aplicados à canais PLC, as Figs. 4.1-4.2 apresentam os modelos de

sistemas OFDM/DMT modificados, os quais são propostos para o cancelamento dos ruídos impulsivos em sistemas multi-portadoras (Ribeiro et al., 2005; Romano et al., 2003). Nesse modelo, o bloco CRI (Cancelamento de Ruído Impulsivo) implementa um algoritmo não-linear baseado em inteligência computacional para reduzir a presença dos ruídos impulsivos a partir dos N_t pares de treinamento dados por $\{\tilde{y}(n), y_d(n)\}_{n=0}^{N_t-1}$, onde N_t é o número de pares de treinamento e $y_d(n)$ é a seqüência desejada. Para realizar tal tarefa, o algoritmo não-linear é, primeiramente, submetido à um treinamento supervisionado. A seqüência desejada, utilizada durante o período de treinamento do algoritmo não-linear, pode ser obtida de duas formas distintas: *i*) através da geração da seqüência $\{\tilde{y}(n)\}$, saída do canal livre de ruído, a partir da estimativa da resposta impulsiva do canal e do conhecimento, pelo receptor, da seqüência transmitida ou *ii*) através da estimação dos estados na saída do canal PLC. Supondo-se que a seqüência $\{\tilde{y}(n)\}$ seja conhecida, a seguinte questão é feita: Qual tipo de técnica pode ser utilizada para minimizar a presença de ruídos impulsivos dentro da abordagem proposta acima?

Objetivando apontar direções que conduzam à solução do problema em questão, a seção 4.1 discute o uso de técnicas baseadas em inteligência computacional (redes neurais e sistemas fuzzy) para minimizar a presença de ruídos impulsivos em sistemas OFDM/DMT. As motivações para o uso dessas técnicas são as seguintes: *i*) os ruídos impulsivos podem ser vistos como componentes não-lineares na saída do canal PLC; *ii*) as técnicas lineares são incapazes de mudar as características originais dos ruídos impulsivos, (ex. regularidade) (Zhang, 2001) e, finalmente, *iii*) as técnicas lineares não são adequadas para reduzir a presença dos ruídos impulsivos quando o mesmo é banda larga e não estacionário (Zhang, 2001). Note que as técnicas discutidas na seção 4.1 são técnicas heurísticas que demandam custo computacional elevado, no entanto, o uso das mesmas abre uma nova perspectiva para o problema de cancelamento de ruídos impulsivos não apenas em sistemas multi-portadoras, mas também em sistemas mono-portadoras.

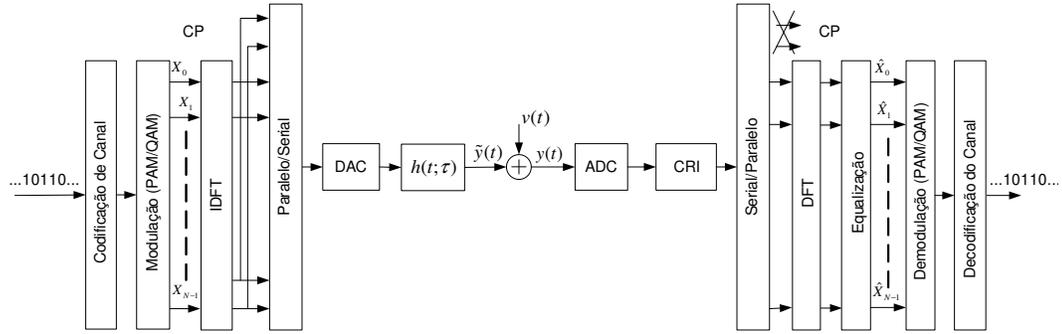


Figura 4.1: Modelo banda base do sistema de transmissão multi-portadoras OFDM modificado.

É importante salientar que o uso das técnicas propostas para o cancelamento de ruídos impulsivos não exclui o uso de técnicas de codificação de canais. De fato, as técnicas propostas são projetadas para minimizar as condições desfavoráveis que os sistemas de codificação de canais irão encontrar em aplicações reais, nas quais a ocorrência de ruídos impulsivos é comum. Como consequência do uso das técnicas propostas, a ocorrência de erros nos decodificadores é minimizada e o comprimento do entrelaçador é menor, o que contribui de forma decisiva para a viabilização das aplicações em tempo real.

Note ainda que o uso da técnica proposta é bastante interessante em sistemas DMT, uma vez que as técnicas de equalização empregadas para encurtar a resposta impulsiva dos canais de comunicação não são projetadas levando-se em conta a presença de ruídos impulsivos. Assim sendo, o desempenho de tais técnicas é deteriorado consideravelmente à medida que os ruídos impulsivos ocorrem na saída dos canais PLC.

Para os sistemas multi-portadoras OFDM/DMT, as SNRs nos sub-canais são

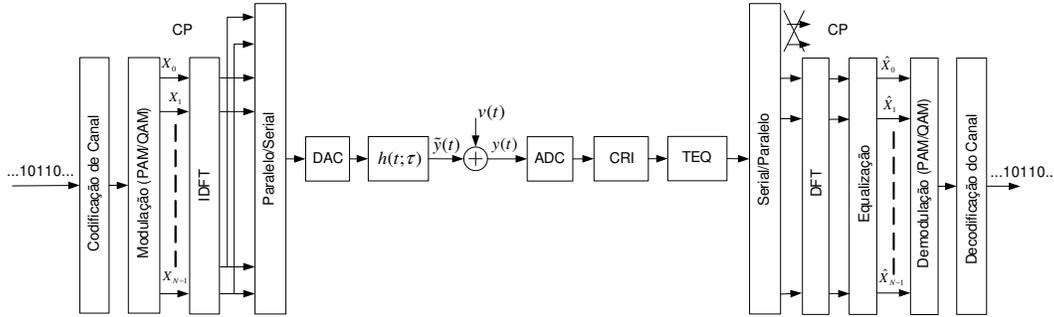


Figura 4.2: Modelo banda base do sistema de transmissão multi-portadoras DMT modificado.

dadas por

$$SNR_k = \begin{cases} \frac{S_k |C_k^{Sinal}|^2}{S_{V,k} |W_k|^2 + S_k |C_k^{IIS}|^2}, & \text{se } DMT \\ \frac{S_k |H_k|^2}{S_{V,k}}, & \text{se } OFDM \end{cases}, \quad k = 0, \dots, N - 1. \quad (4.6)$$

Como todos os termos do lado direito da expressão (4.6) são positivos, fica claro que técnicas de processamento não-linear de sinais para cancelar os ruídos impulsivos podem ser aplicadas. Se, além disso, a técnica de processamento não-linear não afetar as outras componentes, um aumento da taxa de transmissão é observada, uma vez que as novas SNR_k são expressas por

$$SNR_k = \begin{cases} \frac{S_k |C_k^{Sinal}|^2}{S_{bkgr,k} |W_k|^2 + S_k |C_k^{IIS}|^2}, & \text{se } DMT \\ \frac{S_k |H_k|^2}{S_{bkgr,k}}, & \text{se } OFDM \end{cases}, \quad k = 0, \dots, N - 1, \quad (4.7)$$

e, conseqüentemente, para o sistema DMT tem-se

$$\frac{S_k |C_k^{Sinal}|^2}{S_{bkgr,k} |W_k|^2 + S_k |C_k^{IIS}|^2} \geq \frac{S_k |C_k^{Sinal}|^2}{S_{V,k} |W_k|^2 + S_k |C_k^{IIS}|^2}, \quad k = 0, \dots, N - 1, \quad (4.8)$$

e para o sistema OFDM a seguinte desigualdade é válida

$$\frac{S_k |H_k|^2}{S_{bkgr,k}} \geq \frac{S_k |H_k|^2}{S_{V,k}}, \quad k = 0, \dots, N - 1. \quad (4.9)$$

A minimização completa da presença dos ruídos impulsivos, conforme explicitada pela expressão (4.7), é uma tarefa praticamente impossível de ser realizada. No entanto, resultados interessantes são obtidos com as técnicas discutidas na seção 4.1.

4.1 Inteligência Computacional para Cancelamento de Ruídos Impulsivos

As redes MLP (*Multilayer Perceptron*) (Haykin, 1999) são amplamente utilizadas em várias aplicações, pois as mesmas são capazes de aproximar funções não-lineares e complexas. Levando-se em consideração tal característica, pode-se conjecturar que as mesmas são boas candidatas para a redução dos ruídos impulsivos em canais PLC quando os sistemas OFDM/DMT são utilizados para a transmissão banda larga de dados, uma vez que essas redes foram aplicadas com sucesso para reduzir a presença de ruídos impulsivos em sistemas CDMA (*Code Division Multiple Access*) (T. C. Chuah, 2001). No entanto, as melhorias observadas em sistemas CDMA (T. C. Chuah, 2001) quando a rede MLP é treinada pelo método *backpropagation* não se repetem em sistemas OFDM/DMT, uma vez que a convergência do método *backpropagation* é muito lenta.

Dada a baixa velocidade de convergência da rede MLP, na presente seção é formalizado o uso do método GCEM de treinamento, baseado em informações de 2ª ordem, para o treinamento da rede MLP. Além disso, são apresentadas as redes RBF (*Radial Basis Function*) (Mulgrew, 1996) e wavelet chapéu mexicano (Mallat, 2001) treinadas pelo método de treinamento utilizado para treinar a rede MLP, visando avaliar qual das técnicas apresenta o melhor desempenho. A motivação para o uso das redes RBF e wavelet é que as mesmas são capazes, assim como a rede MLP é, de

aproximar funções não-lineares e complexas e podem apresentar melhores resultados que as redes MLP.

As redes MLP, RBF e Wavelet com uma única camada escondida podem ser definidas por

$$\mathbf{s}(n) = \mathbf{A}^T(n) \begin{bmatrix} \mathbf{y}(n) \\ 1 \end{bmatrix}, \quad (4.10)$$

$$\mathbf{z}(n) = \mathbf{f}(\mathbf{s}(n)) = \left[f(s_0(n)) \quad \cdots \quad f(s_{L-1}(n)) \right]^T, \quad (4.11)$$

$$\mathbf{x}(n) = \left[x_0(n) \quad \cdots \quad x_{M-1}(n) \right]^T = \mathbf{B}^T(n) \begin{bmatrix} \mathbf{z}(n) \\ 1 \end{bmatrix} \quad (4.12)$$

e

$$f(s_i(n)) = \begin{cases} \tanh(s_i(n)), & \text{Se Rede MLP} \\ \exp \left\{ -\frac{(s_i(n))^2}{2} \right\}, & \text{Se Rede RBF} \\ (s_i^2(n) - 1) \exp \left\{ -\frac{s_i^2(n)}{2} \right\}, & \text{Se Rede Wavelet} \end{cases} \quad i = 0, \dots, L-1, \quad (4.13)$$

onde $\mathbf{x}(n) = [y_0(n) \cdots y_{K-1}(n)]^T$ é o vetor de entrada, $\mathbf{z}(n) = [z_0(n) \cdots z_{L-1}(n)]^T$ é o vetor de saída da camada escondida ou intermediária, $\mathbf{x}(n) = [x_0(n) \cdots x_{M-1}(n)]^T$ é o vetor de saída da rede, $\mathbf{A}(n) \in \mathfrak{R}^{(K+1) \times L}$ é a matriz de pesos entre as camadas de entrada e intermediária, $\mathbf{B}(n) \in \mathfrak{R}^{(L+1) \times M}$ é a matriz de pesos entre as camadas intermediária e de saída e $f(\cdot)$ é a função de ativação da rede.

Agora, sejam $\mathbf{a}(n)$ e $\mathbf{b}(n)$ os vetores colunas formados pelas linhas da matrizes $\mathbf{A}(n)$ e $\mathbf{B}(n)$. Então, o vetor, $\mathbf{w}(n)$, que contém todos os pesos da rede, a função custo, $J_c(\mathbf{w}(n))$ a ser minimizada durante o treinamento, e o vetor gradiente

$\nabla J_c(\mathbf{w}(n))$ são dados por

$$\mathbf{w}(n) = \left[\mathbf{a}^T(n) \quad \mathbf{b}^T(n) \right]^T, \quad (4.14)$$

$$J_c(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} J(\mathbf{w}(n)) = \frac{1}{2} \sum_{n=0}^{N-1} \mathbf{e}^2(n) = \frac{1}{2} \sum_{n=0}^{N-1} [\mathbf{x}(n) - \mathbf{x}_d(n)]^2 \quad (4.15)$$

e

$$\nabla J(\mathbf{w}(n)) = \left[\nabla J_{\mathbf{a}}^T(\mathbf{w}(n)) \quad \nabla J_{\mathbf{b}}^T(\mathbf{w}(n)) \right]^T, \quad (4.16)$$

onde $\mathbf{x}_d(n) = [x_{d,0}(n) \cdots x_{d,M-1}(n)]^T$ é o vetor desejado na saída da rede, $\mathbf{e}(n) = [x_0(n) - x_{d,0}(n) \cdots x_{M-1}(n) - x_{d,M-1}(n)]^T$ é o vetor erro observado na saída da rede, $\nabla J_{\mathbf{a}}(\mathbf{w}(n))$ é o vetor gradiente de $\mathbf{a}(n)$, $\nabla J_{\mathbf{b}}(\mathbf{w}(n))$ é o vetor gradiente de $\mathbf{b}(n)$.

A escolha adequada de um método de otimização para o treinamento da redes MLP (Príncipe, Euliano, & Lefebvre, 2000), RBF (Haykin, 1999; Mulgrew, 1996) e Wavelet (Mallat, 2001) exerce um papel fundamental para a obtenção de soluções que sejam eficientes e simples para o problema de cancelamento de ruídos impulsivos. De fato, o método de otimização deve ser capaz de garantir a rápida convergência da rede para um mínimo local e demandar baixo custo computacional.

Dentre os algoritmos encontrados na literatura para o treinamento da rede, pode-se destacar o método *back-propagation* (Rumelhart et al., 1986), baseado em informações de 1ª ordem, que produz razoável desempenho se os dados são representativos do problema e se o dimensionamento *a priori* da estrutura da rede é feita de maneira apropriada (Haykin, 1999; Theodoridis & Koutroumbas, 1999). Entretanto, o método *back-propagation* apresenta baixo desempenho devido à baixa velocidade de convergência. De fato, as simulações realizadas mostraram que para o problema de cancelamento de ruídos impulsivos em sistemas OFDM, o método *back-propagation* não permite a obtenção de uma rede capaz de minimizar a presença de ruídos impulsivos.

Por outro lado, o uso de métodos de otimização baseados em informação de 2ª ordem apresentam velocidade de convergência superior àquela obtida com os métodos de 1ª ordem, uma vez que os mesmos fazem uso da informação contida na matriz Hessiana para atualizar os pesos da rede. A desvantagem dos métodos de 2ª ordem reside no fato de que os mesmos apresentam uma complexidade computacional elevada por causa da necessidade de se calcular a matriz Hessiana. Tal problema se torna significativo se a rede tem um número elevado de pesos. No entanto, esse custo computacional pode ser consideravelmente reduzido se a multiplicação exata e rápida do produto da matriz Hessiana por um vetor (Pearlmutter, 1994), sem o cálculo explícito da matriz Hessiana, é considerado. Objetivando explorar esta abordagem, o método GCEM (Santos & Zuben, 2000) é utilizado para treinar as redes MLP, RBF e Wavelet, pois o mesmo faz uso da técnica desenvolvida em (Pearlmutter, 1994) para a multiplicação exata e rápida do produto da matriz Hessiana por um vetor, garantindo que a complexidade seja linear com o número de parâmetros da rede. Outra motivação para o uso do método GCEM é que o mesmo faz uso de uma regra de atualização dos pesos baseada nos critérios de Polak-Ribiere e Fletcher-Reeves. Tal regra permite ao método GCEM apresentar maior velocidade de convergência que os métodos anteriores, conforme mostrado em (Santos & Zuben, 2000). Além disso, esse método foi utilizado com sucesso em (Ribeiro, Barbedo, Romano, & Lopes, 2005).

A implementação do método GCEM para treinar a rede dada por (4.10)-(4.13) é baseada no cálculo do vetor gradiente $\nabla J_w(\mathbf{w}(n))$ e da matriz Hessiana $\mathbf{H}(\mathbf{w}(n))$, vide (Santos & Zuben, 2000). Entretanto, é sabido que o cálculo da matriz Hessiana demanda, a cada iteração, elevado custo computacional quando o número de pesos da rede é grande. Objetivando minimizar tal problema, em (Pearlmutter, 1994) foi proposto uma forma aproximada para se calcular o produto da matriz Hessiana por um vetor direcional $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$. O uso desta forma aproximada na implementação do método GCEM para o treinamento das redes MLP, RBF e

Wavelet (chapéu mexicano) reduz consideravelmente a complexidade computacional durante o treinamento sem que haja perda significativa de desempenho. Para o cálculo de $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ sem o cálculo explícito da matriz Hessiana, deve-se utilizar o operador diferencial dado por (Pearlmutter, 1994)

$$\mathfrak{R}_{\mathbf{d}}\{f(\mathbf{w}(n))\} \equiv \frac{\partial}{\partial \alpha} f(\mathbf{w}(n) + \alpha \mathbf{d}(n))|_{\alpha=0}, \quad (4.17)$$

onde $\mathbf{d}(n)$ é um vetor direcional e $0 < \alpha \ll 1$, $\alpha \in \mathfrak{R}$.

Agora sejam

$$\nabla_{\mathbf{A}} J(\mathbf{w}(n)) = \frac{\partial J_c(\mathbf{w}(n))}{\partial \mathbf{A}(n)} = \begin{bmatrix} \frac{\partial J_c(\mathbf{w}(n))}{\partial a_{1,1}(n)} & \cdots & \frac{\partial J_c(\mathbf{w}(n))}{\partial a_{1,L}(n)} \\ \vdots & \ddots & \vdots \\ \frac{\partial J_c(\mathbf{w}(n))}{\partial a_{(K+1),1}(n)} & \cdots & \frac{\partial J_c(\mathbf{w}(n))}{\partial a_{(K+1),L}(n)} \end{bmatrix}, \quad (4.18)$$

$$\nabla_{\mathbf{A}} J(\mathbf{w}(n)) = \begin{bmatrix} \mathbf{y}(n) \\ 1 \end{bmatrix} \frac{\partial J^T(\mathbf{w}(n))}{\partial \mathbf{s}(n)}, \quad (4.19)$$

$$\frac{\partial J(\mathbf{w}(n))}{\partial \mathbf{s}(n)} = \left(\mathbf{B}(n) \odot \dot{\mathbf{f}}(\mathbf{s}(n)) \right) \mathbf{e}(n), \quad (4.20)$$

$$\nabla_{\mathbf{B}} J(\mathbf{w}(n)) = \frac{\partial J(\mathbf{w}(n))}{\partial \mathbf{B}(n)} = \begin{bmatrix} \frac{\partial J(\mathbf{w}(n))}{\partial b_{1,1}(n)} & \cdots & \frac{\partial J_c(\mathbf{w}(n))}{\partial b_{1,M}(n)} \\ \vdots & \ddots & \vdots \\ \frac{\partial J(\mathbf{w}(n))}{\partial b_{(L+1),1}(n)} & \cdots & \frac{\partial J_c(\mathbf{w}(n))}{\partial b_{(L+1),M}(n)} \end{bmatrix} \quad (4.21)$$

e

$$\nabla_{\mathbf{B}} J(\mathbf{w}(n)) = \begin{bmatrix} \mathbf{z}(n) \\ 1 \end{bmatrix} \mathbf{e}^T(n), \quad (4.22)$$

onde o operador de Hadarmard é dado por

$$\begin{aligned}
\mathbf{R}(n) \odot \mathbf{u}(n) &= \begin{bmatrix} r_{0,0}(n) & \cdots & r_{0,Q}(n) \\ \vdots & \ddots & \vdots \\ r_{P-1,0}(n) & \cdots & r_{P-1,Q-1}(n) \end{bmatrix} \odot \begin{bmatrix} u_0(n) \\ \vdots \\ u_{P-1}(n) \end{bmatrix} \\
&= \begin{bmatrix} r_{0,0}(n)s_0(n) & \cdots & r_{0,Q-1}(n)u_0(n) \\ \vdots & \ddots & \vdots \\ r_{P-1,0}(n)s_{P-1}(n) & \cdots & r_{P-1,Q-1}(n)u_{P-1}(n) \end{bmatrix}. \tag{4.23}
\end{aligned}$$

para o caso que $\mathbf{R}(n)$ é uma matriz e $\mathbf{u}(n)$ é um vetor. Para o caso de dois vetores $\mathbf{r}(n)$ e $\mathbf{u}(n)$, o operador de Hadarmard é dado por

$$\mathbf{r}(n) \odot \mathbf{u}(n) = [r_0(n)u_0(n) \cdots r_{P-1}(n)u_{P-1}(n)]^T. \tag{4.24}$$

O cálculo direto do produto $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ é obtido com a aplicação do operador diferencial nas expressões (4.18)-(4.22), ou seja,

$$\mathfrak{R}_d\{\mathbf{s}(n)\} = \nabla_{\mathbf{A}^T} J(\mathbf{w}(n)) \begin{bmatrix} \mathbf{y}(n) \\ 1 \end{bmatrix} + \mathbf{A}^T \begin{bmatrix} \mathfrak{R}_d\{\mathbf{y}(n)\} \\ 0 \end{bmatrix}, \tag{4.25}$$

$$\mathfrak{R}_d\{\mathbf{z}(n)\} = \dot{\mathbf{f}}(\mathbf{s}(n)) \odot \mathfrak{R}_d\{\mathbf{s}(n)\}, \tag{4.26}$$

$$\mathfrak{R}_d\{\mathbf{x}(n)\} = \mathbf{B}^T(n) \begin{bmatrix} \mathfrak{R}_d\{\mathbf{z}(n)\} \\ 0 \end{bmatrix} + \nabla_{\mathbf{B}^T} J(\mathbf{w}(n)) \begin{bmatrix} \mathbf{z}(n) \\ 1 \end{bmatrix}, \tag{4.27}$$

$$\mathfrak{R}_d\{\dot{\mathbf{f}}(n)\} = \ddot{\mathbf{f}}(n) \odot \mathfrak{R}_d\{\mathbf{s}(n)\}, \tag{4.28}$$

$$\mathfrak{R}_d\{\nabla_{\mathbf{A}}J(\mathbf{w}(n))\} = \begin{bmatrix} \mathfrak{R}_d\{\mathbf{y}(n)\} \\ 0 \end{bmatrix} \frac{\partial \nabla J^T(\mathbf{w}(n))}{\partial \mathbf{z}(n)} + \begin{bmatrix} \mathbf{y}(n) \\ 1 \end{bmatrix} \mathfrak{R}_d\left\{\frac{\partial \nabla J^T(\mathbf{w}(n))}{\partial \mathbf{z}(n)}\right\} \quad (4.29)$$

$$\mathfrak{R}_d\left\{\frac{\partial \nabla J(\mathbf{w}(n))}{\partial \mathbf{s}(n)}\right\} = \left(\mathbf{B}(n) \odot \dot{\mathbf{f}}(n)\right) \mathfrak{R}_d\{\mathbf{e}(n)\} + \left(\nabla_{\mathbf{B}}J(\mathbf{w}(n)) \odot \dot{\mathbf{f}}(n) + \mathbf{B}(n) \odot \mathfrak{R}_d\{\dot{\mathbf{f}}(n)\}\right) \mathbf{e}(n), \quad (4.30)$$

$$\mathfrak{R}_d\{\nabla_{\mathbf{B}}J(\mathbf{w}(n))\} = \begin{bmatrix} \mathbf{z}(n) \\ 1 \end{bmatrix} \mathfrak{R}_d\{\mathbf{e}^T(n)\} + \begin{bmatrix} \mathfrak{R}_d\{\mathbf{z}(n)\} \\ 0 \end{bmatrix} \mathbf{e}^T(n), \quad (4.31)$$

e, finalmente,

$$\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n) = \mathfrak{R}_d\{\nabla J(\mathbf{w}(n))\} = \begin{bmatrix} \mathfrak{R}_d\{\nabla_{\mathbf{a}}J(\mathbf{w}(n))\} \\ \mathfrak{R}_d\{\nabla_{\mathbf{b}}J(\mathbf{w}(n))\} \end{bmatrix} \quad (4.32)$$

O procedimento baseado no método GCEM, para o treinamento das redes MLP, RBF e Wavelet é descrito a seguir:

Sejam $\mathbf{g}(n) = \nabla J(\mathbf{w}(n))$, N_w é o número total de pesos, n é a n -ésima iteração, $\beta_{PR}(n)$ o ganho $\beta(n)$ obtido com a expressão de Polak-Ribiere e $\beta_{FR}(n)$ o ganho $\beta(n)$ fornecido pela expressão de Fletcher-Reeves (Santos & Zuben, 2000), então o algoritmo do método GCEM para treinamento das redes MLP, RBF e Wavelet é dado a seguir.

Inicialização: inicialize o vetor $\mathbf{w}(n)$ com valores aleatórios, especifique o número máximo de épocas e faça $\xi \ll 1$ e $Q = -1$

Passo 1: Calcule $\mathbf{g}(0) = \nabla J_c(\mathbf{w}(0))$ e faça $\mathbf{d}(0) = -\mathbf{g}(0)$

Passo 2: Faça $Q = Q + 1$ e $n = -1$

Passo 3: Faça $n = n + 1$

Passo 4: Para $n = 0, \dots, N - 1$

1. Faça $\mathbf{d}(n) = -\mathbf{g}(n) + \beta(n - 1)\mathbf{d}(n)$
2. Calcule $\mathbf{w}(n + 1) = \mathbf{w}(n) + \alpha(n)\mathbf{d}(n)$, sendo

$$\alpha(n) = \frac{-\mathbf{g}^T(n)\mathbf{d}(n)}{\mathbf{d}^T(n)\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)} \quad (4.33)$$

e $\mathbf{H}(\mathbf{w}(n))\mathbf{d}(n)$ é dado por (4.32)

3. Faça $\mathbf{g}(n + 1) = \nabla J_c(\mathbf{w}(n + 1))$
4. Calcule

$$\beta_{PR}(n) = \frac{\mathbf{g}^T(n+1)(\mathbf{g}(n+1) - \mathbf{g}(n))}{\mathbf{g}^T(n)\mathbf{g}(n)}$$

$$\beta_{FR}(n) = \frac{\mathbf{g}^T(n+1)\mathbf{g}(n+1)}{\mathbf{g}^T(n)\mathbf{g}(n)}$$
5. Atualize o ganho $\beta(n)$ e o vetor direcional $\mathbf{d}(n)$ de acordo com a seguinte regra:

Se $n = N - 1$

$$\beta(n) = 0$$

$$\mathbf{d}(n + 1) = -\mathbf{g}(n + 1)$$

$$n = 0$$

Senão

Se $\beta_{PR}(n) > 0$ e $\beta_{PR}(n) < \beta_{FR}(n)$

$$\beta(n) = \beta_{PR}(n)$$

$$\mathbf{d}(n + 1) = -\mathbf{g}(n + 1) + \beta(n)\mathbf{d}(n)$$

Se $-\mathbf{g}^T(n+1)\mathbf{d}(n+1) < 0$

$$\beta(n) = 0$$

$$\mathbf{d}(n+1) = -\mathbf{g}(n+1)$$

$$n = 0$$

Senão

$$n = n + 1$$

Senão Se $\beta_{PR}(n) > \beta_{FR}(n)$

$$\beta(n) = \beta_{FR}(n)$$

$$\mathbf{d}(n+1) = -\mathbf{g}(n+1) + \beta(n)\mathbf{d}(n)$$

Se $-\mathbf{g}^T(n+1)\mathbf{d}(n+1) < 0$

$$\beta(n) = 0$$

$$\mathbf{d}(n+1) = -\mathbf{g}(n+1)$$

$$n = 0$$

Senão

$$n = n + 1$$

Senão

$$\beta(n) = 0$$

$$\mathbf{d}(n+1) = -\mathbf{g}(n+1)$$

$$n = 0$$

Passo 5: Se $\nabla J_c(\mathbf{w}(n)) > \xi$ ou $Q \leq Q_{max}$, então $\mathbf{w}(0) = \mathbf{w}(N-1)$, caso contrário volte ao **Passo 2**

Finalmente, para fazer pequenos ajustes em tempo real dos pesos da rede levando-se em consideração pequenas modificações nas características dos ruídos impulsivos, o uso da técnica LLE, vide seção 2.3.3.2, pode ser periodicamente utilizada, conforme sugerido em (Ribeiro et al., 2005). De fato, a estimação local da

Lipschitz permite a convergência rápida da rede para um novo estado sem demandar o uso do método GCEM.

A seção de resultados irá mostrar os resultados obtidos quando as redes MLP, RBF e wavelet e os algoritmos fuzzy do tipo 2 não normalizados, propostos na seção 2.2, são implementados no bloco CRI do sistema OFDM/DMT modificado visando o cancelamento de ruídos impulsivos.

4.2 Resultados Numéricos

Nesta seção o desempenho das técnicas propostas para o cancelamento de ruídos impulsivos em sistemas de transmissão multi-portadora DMT, vide (Ribeiro et al., 2005), é analisado através de simulações computacionais. O parâmetro principal para validar o desempenho das técnicas propostas é a taxa de bit expressa por (Acker, Leus, Moonen, Wiel, & Pollet, 2001)

$$TaxaBit = \left(\sum_{k=1}^N b_k \right) \frac{f_s}{N + L_v} \quad (4.34)$$

e

$$b_k = \left\lfloor \log_2(1 + 10^{((SNR_k - \Gamma - \gamma_m + \gamma_c)/10)}) \right\rfloor, \quad (4.35)$$

onde b_k é o número de bits alocados para a k -ésima sub-banda, L_v é o comprimento do prefixo cíclico que aqui é assumido ser igual a L_b , $f_s = 9,6$ MHz é a frequência de amostragem, N é o número de sub-bandas, Γ é o *gap*, γ_m é a margem e γ_c é o ganho de codificação. Para a obtenção dos resultados numéricos mostrados a seguir, as seguintes considerações foram feitas: *i*) o ruído de fundo tem PSD igual a -120 dB/Hz; *ii*) a PSD do ruído impulsivo varia entre -90 dB/Hz e -105 dB/Hz; *iii*) a PSD dos símbolos OFDM transmitidos é igual à -70 dB/Hz; *iv*) o comprimento do símbolo OFDM é igual a 2048; os símbolos OFDM de treinamento apresentam

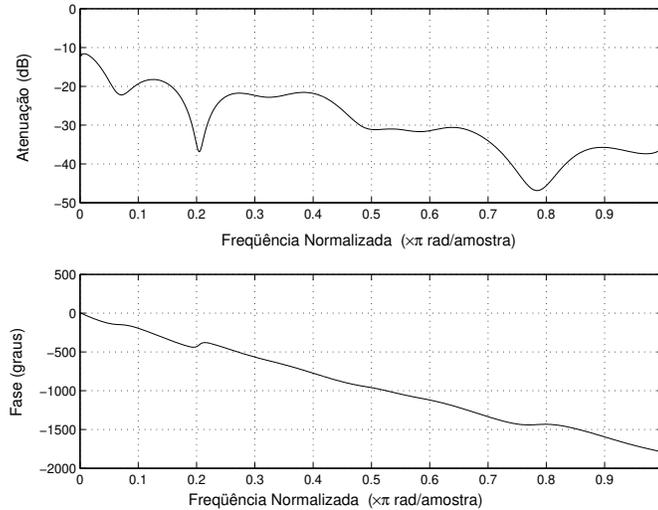


Figura 4.3: Resposta em frequência do canal PLC 1.

espectro plano em todas as sub-bandas e foram gerados de acordo com (Bingham & Putten, 1997); *v*) o número de símbolos OFDM empregados para o treinamento da técnica de cancelamento de ruídos impulsivos e para o TEQ (*time domain equalizer*) são iguais a 50 e 350, respectivamente; *vi*) a técnica TEQ utilizada nas simulações foi a min-ISI (Arslan, Evans, & Kiaei, 2001) e *vii*) a SNR_k é dada por (4.7) para o sistema OFDM/DMT e por (4.8) para o sistema OFDM/DMT modificado. As redes neurais MLP, RBF e Wavelet têm 1 entrada, 18 neurônios, na única camada escondida, e uma saída. Os algoritmos fuzzy proposto na seção 2.2, por sua vez, têm 1 entrada, 18 regras e uma saída. A banda de frequência utilizada é entre 0,2 e 5 MHz. As respostas em frequência dos canais PLC, vide (Zimmermann & Dostert, 2002a; Langfeld, 2001), utilizadas nestas simulações são mostradas nas Figs. 4.3-4.4, aqui denominados de canais PLC 1 e PLC 2.

As Figs. 4.5-4.6 mostram o sinal na saída do canal livre de ruído (a), o sinal na saída do canal livre de ruído corrompida pela presença do ruído aditivo (b) o sinal na saída do bloco CRI que implementa a rede MLP para cancelar a presença dos

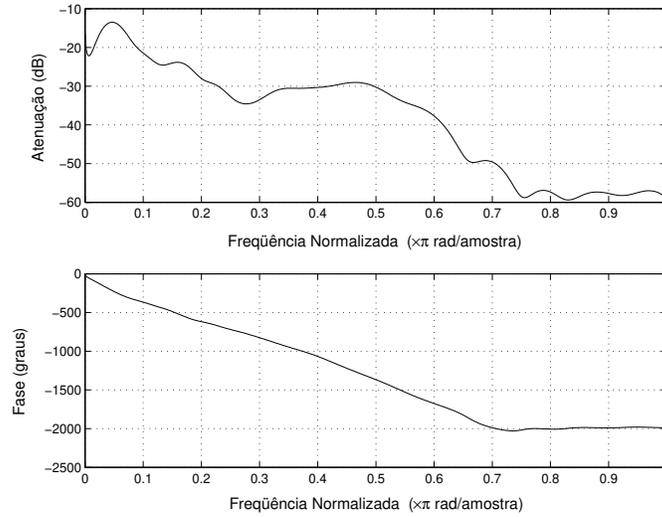


Figura 4.4: Resposta em frequência do canal PLC 2.

ruídos impulsivos. Objetivando facilitar a visualização, os sinais apresentados nas Figs. 4.5-4.6 foram submetidos a ganhos. Para obter esses resultados, considerou-se que a PSD do sinal transmitido é igual a -70 dB/Hz, a PSD do ruído de fundo é igual a -120 dB/Hz e a PSD do ruído impulsivo é igual a -90 dB/Hz. Além disso, $L_b = 16$ e $L_w = 48$ foram escolhidos para o canal PLC 1 e $L_b = 16$ e $L_w = 56$ foram utilizados para o canal PLC 2. Conforme pode ser observado via inspeção visual, o uso da técnica de cancelamento de ruídos impulsivos reduz consideravelmente a presença de ruídos impulsivos na saída do canal PLC.

Conforme discutido na parte introdutória deste capítulo, a redução da presença dos ruídos impulsivos em canais de comunicação resulta no aumento da SNR nos sub-canais. Resultados numéricos que confirmam tal comportamento são mostrados nas Tabs. 4.1-4.2 para diferentes valores da PSD dos ruídos impulsivos. Os valores SNR, mostrado nessas duas tabelas, são dados por

$$SNR = 10 \log_{10} \left(\frac{\sigma_y^2}{\sigma_{v_\eta}^2} \right), \quad (4.36)$$

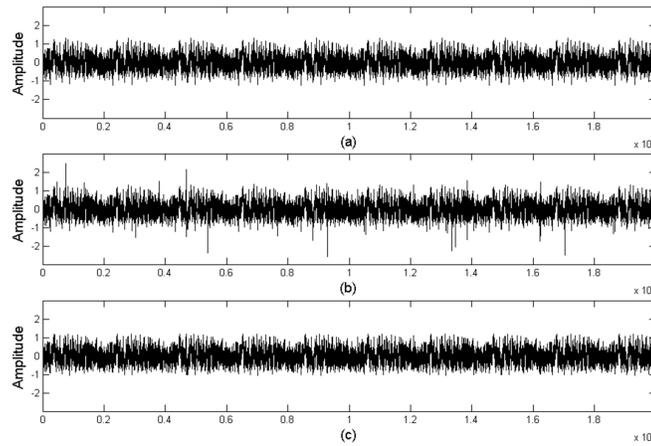


Figura 4.5: (a) saída do canal PLC $\tilde{y}(n)$, (b) saída do canal PLC corrompida com ruído aditivo $y(n)$ e (c) saída do bloco CRI $\hat{\tilde{y}}(n)$, referentes ao canal PLC 1.

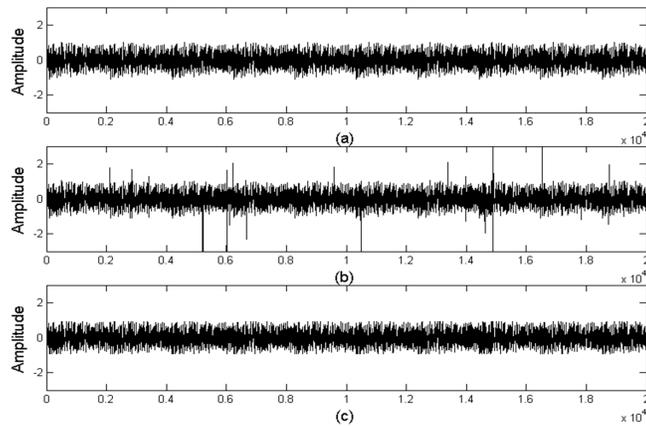


Figura 4.6: (a) saída do canal PLC $\tilde{y}(n)$, (b) saída do canal PLC corrompida com ruído aditivo $y(n)$ e (c) saída do bloco CRI $\hat{\tilde{y}}(n)$, referentes ao canal PLC 2.

onde σ_y^2 é a potência do sinal na saída do canal livre de ruídos e $\sigma_{v_\eta}^2$ é a potência do ruído aditivo na entrada do bloco TEQ. As colunas 2,...,9 das Tabs. 4.1-4.2 trazem os seguintes resultados:

Coluna 2 (SAÍDA): mostra a SNR na saída do canal PLC;

Coluna 3 (Hard): mostra a SNR obtida após a aplicação da técnica dada pela expressão (4.4);

Coluna 4 (Zero): mostra a SNR obtida após a aplicação da técnica dada pela expressão (4.5);

Coluna 5 (Wav): mostra a SNR, obtida após a aplicação da técnica baseada na rede Wavelet proposta para a redução de ruídos impulsivos em sistema OFDM/DMT, vide seção 4.1;

Coluna 6 (RBF): mostra a SNR, obtida após a aplicação da técnica baseada na rede RBF proposta para a redução de ruídos impulsivos em sistema OFDM/DMT, vide seção 4.1;

Coluna 7 (MLP): mostra a SNR, obtida após a aplicação da técnica baseada na rede MLP proposta para a redução de ruídos impulsivos em sistema OFDM/DMT, vide seção 4.1;

Coluna 8 (FUZZY I): mostra a SNR, obtida após a aplicação da técnica baseada no algoritmo Fuzzy I proposta para a redução de ruídos impulsivos em sistema OFDM/DMT, vide seção 2.2.1;

Coluna 9 (FUZZY II): mostra a SNR, obtida após a aplicação da técnica baseada

no algoritmo Fuzzy II proposta para a redução de ruídos impulsivos em sistema OFDM/DMT, vide seção 2.2.2;

Conforme observado nas Tabs. 4.1-4.2, o uso da técnica proposta mostra-se eficaz para reduzir a presença dos ruídos impulsivos em canais PLC. De acordo com os resultados numéricos reunidos, ganhos superiores a 1.8 dB são obtidos com às técnicas propostas em relação as técnicas baseadas nas expressões (4.4) e (4.5).

Dentre as redes neurais analisadas, observa-se que a rede wavelet apresentou o pior resultado, enquanto a rede MLP apresentou um resultado levemente superior ao obtido com a rede RBF. Comparando-se o desempenho das técnicas fuzzy do tipo 2 propostas para a redução de ruídos impulsivos, nota-se que o algoritmo fuzzy do tipo 2 *nonsingleton*, cuja incerteza do ruído aditivo é modelada com uma função exponencial apresenta resultados superiores àqueles obtidos com o algoritmo fuzzy do tipo 2 cuja incerteza do ruído aditivo é modelada com uma função Gaussiana. De fato, conforme mostrado na seção 2.4, o algoritmo fuzzy do tipo 2, cuja incerteza do ruído aditivo é modelada com uma função Gaussiana, reduz-se para um algoritmo fuzzy do tipo 2 *singleton* quando a variância das funções de pertinência das regras e do ruído aditivo é analisada conjuntamente.

Para mostrar que a redução da presença dos ruídos impulsivos na saída dos canais PLC aumenta o desempenho do sistema DMT, as Tabs. 4.3-4.4 ilustram as taxas de bit, em Mbps, obtidas para ambos os canais PLC quando $\Gamma = 9.8$ dB para garantir $P_e = 10^{-7}$, $\gamma_m = 6$ dB, $\gamma_c = 3$ dB, $L_b = 16$, $L_w = 40$ para o canal PLC 1 e $L_w = 48$ para o canal PLC 2. Baseado no fato de que a SNR na entrada dos equalizadores TEQ é maior, vide resultados mostrados nas Tabs. 4.3-4.4, é de se esperar que tal resultado promova um aumento da taxa de bit considerando a melhoria da SNR na entrada do bloco TEQ. Esse resultado evidencia, em conformidade com a teoria, que o aumento da SNR na entrada do equalizador

PSD	SAÍDA	Hard	Zero	Wav	RBF	MLP	FUZZY I	FUZZY II
-90,0	14,22	16,36	17,00	16,95	18,21	18,26	18,83	18,99
-92,5	17,14	18,51	18,00	18,58	19,64	19,66	20,01	20,15
-95,0	18,89	19,75	18,51	19,62	20,32	20,45	20,72	20,81
-97,5	21,12	21,41	19,22	21,28	21,48	21,62	21,89	21,98
-100,0	22,79	22,90	19,85	22,79	22,98	23,01	23,38	23,51
-102,5	24,22	24,27	20,40	24,16	24,29	24,29	24,65	24,76
-105,0	25,17	25,18	20,83	25,07	25,19	25,19	25,23	25,29

Tabela 4.1: Canal PL 1: Valores de SNR (dB), na entrada do bloco TEQ, para diferentes valores de PSD (dB/Hz) do ruído impulsivo.

PSD	SAÍDA	Hard	Zero	Wav	RBF	MLP	FUZZY I	FUZZY II
-90,0	12,29	16,15	17,10	16,62	18,07	18,11	18,51	19,01
-92,5	14,88	17,42	17,55	17,39	18,66	18,71	19,64	19,88
-95,0	18,97	18,36	18,09	19,51	20,22	20,24	20,98	21,21
-97,5	20,45	19,06	18,55	20,16	21,56	21,55	21,73	22,00
-100,0	21,03	20,86	19,11	20,62	21,86	21,97	22,03	22,37
-102,5	21,87	22,02	19,69	21,92	22,06	22,06	22,55	22,66
-105,0	22,95	23,01	20,22	22,94	23,02	23,04	23,06	23,09

Tabela 4.2: Canal PL 2: Valores de SNR (dB), na entrada do bloco TEQ, para diferentes valores de PSD (dB/Hz) do ruído impulsivo.

promove um aumento efetivo da taxa de bit. Dentre as técnicas propostas para cancelamento de ruídos impulsivos, observa-se que a rede neural baseada em wavelet apresentou o menor ganho, enquanto o algoritmo fuzzy do tipo 2 discutido na seção 2.2.2 apresentou o melhor desempenho.

Objetivando analisar o comportamento da técnica TEQ dentro do sistema DMT modificado, as Figs. 4.7-4.8 mostram o desempenho do sistema DMT modificado em

PSD	SAÍDA	Hard	Zero	Wav	RBF	MLP	FUZZY I	FUZZY II
-90,0	32,02	34,15	33,41	34,72	35,57	35,82	36,13	36,21
-92,5	34,93	35,09	34,00	35,96	36,70	36,67	37,73	37,86
-95,0	36,07	36,72	34,30	36,67	37,21	37,33	37,78	38,01
-97,5	37,54	37,88	34,70	37,32	37,96	38,07	38,86	39,78
-100,0	38,71	38,05	34,96	38,73	38,83	38,85	39,41	39,93
-102,5	39,76	39,08	35,23	39,52	39,58	39,58	39,88	40,08
-105,0	41,76	41,07	35,46	41,67	41,77	41,77	41,82	41,84

Tabela 4.3: Canal PL 1: Taxas de bit em Mbps, obtidas com as técnicas DMT e DMT modificada, para diferentes valores de PSD(dB/Hz) dos ruídos impulsivos.

PSD	SAÍDA	Hard	Zero	Wav	RBF	MLP	FUZZY I	FUZZY II
-90,0	23,13	24,64	25,49	26,13	26,79	26,81	27,01	27,13
-92,5	24,94	25,64	26,49	26,52	27,08	27,12	27,56	27,77
-95,0	25,07	26,58	25,51	26,81	27,45	27,46	27,68	27,86
-97,5	26,18	26,14	21,43	27,18	27,72	27,76	27,94	28,15
-100,0	27,64	26,80	26,78	27,75	27,84	27,89	28,01	28,13
-102,5	28,26	27,30	27,15	28,30	28,30	28,32	28,54	28,60
-105,0	28,59	28,62	28,53	28,65	28,61	28,63	28,72	28,74

Tabela 4.4: Canal PL 2: Taxas de bit em Mbps, obtidas com as técnicas DMT e DMT modificada, para diferentes valores de PSD(dB/Hz) dos ruídos impulsivos.

função do comprimento da resposta impulsiva do equalizador TEQ, L_w , que varia entre 8 e 128. As seguintes considerações foram feitas para a obtenção dos resultados numéricos reunidos nas Figs. 4.7-4.8: $\Gamma = 9.8$ dB para garantir $P_e = 10^{-7}$, $\gamma_m = 6$ dB, $\gamma_c = 3$ dB, $L_b = 16$, a PSD do sinal transmitido igual a -70 dB/Hz, a PSD do ruído de fundo igual a -120 dB/Hz e a PSD dos ruídos impulsivos igual a -90 dB/Hz. Conforme observado, o uso da técnica de cancelamento de ruídos impulsivos permite

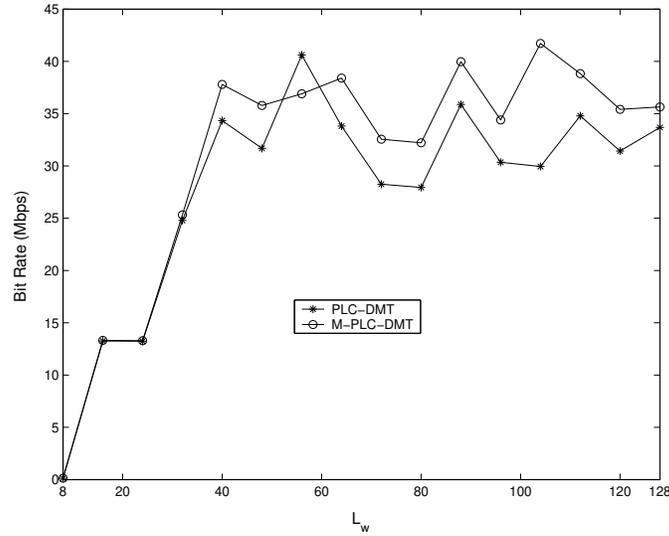


Figura 4.7: Desempenho dos sistemas DMT e DMT modificado em função de L_w para $L_b = 16$, quando o canal PLC é o 1.

que o equalizador TEQ tenha, em sua entrada, o sinal dado por

$$y(n) = \tilde{y}(n) + v_\eta(n), \quad (4.37)$$

onde $\tilde{y}(n)$ é a saída do canal PLC e $v_\eta(n)$ é o ruído aditivo resultante após o uso da técnica não-linear para o cancelamento das componentes impulsivas do ruído aditivo. Pode-se constatar que $\sigma_{v_\eta}^2 \leq \sigma_v^2$. Como a potência do ruído aditivo na entrada do bloco TEQ é menor no sistema proposto devido a redução da presença dos ruídos impulsivos, tem-se que a técnica TEQ pode ser aplicada para encurtar a resposta impulsiva do canal PLC em condições menos severas e, conseqüentemente, melhores desempenhos são obtidos para uma ampla gama de valores de L_w (vide os resultados mostrados nas Figs. 4.7 e 4.8, obtidos quando a rede MLP foi implementada no bloco CRI).

Finalmente, para avaliar a influência do uso da técnica de cancelamento de ruídos impulsivos na obtenção da resposta impulsiva combinada entre o canal PLC e o

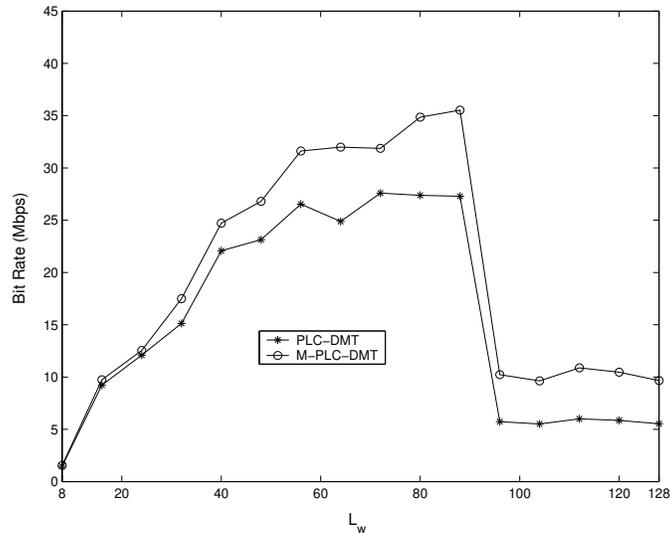


Figura 4.8: Desempenho dos sistemas DMT e DMT modificado em função de L_w para $L_b = 16$, quando o canal PLC é o 2.

equalizador TEQ, algumas simulações foram realizadas. As Figs. 4.9-4.10 mostram o desempenho dos sistemas DMT e DMT modificada em função do comprimento da resposta impulsiva $c(n)$ que varia entre 4 e 128, quando a rede MLP é implementada no bloco CRI. Conforme é observado, o uso da técnica de cancelamento de ruídos impulsivos permite que melhores desempenhos sejam obtidos em ambos os canais PLC analisados quando o comprimento da resposta combinada varia.

4.3 Sumário

Neste capítulo introduziu-se uma modificação no sistema DMT/OFDM que leva em consideração o cancelamento de ruídos impulsivos na saída dos canais PLC. Para cancelar ou reduzir a presença dos ruídos impulsivos nos canais PLC foi proposto o uso das redes MLP, RBF, Wavelet e dos sistemas fuzzy do tipo 2 *nonsingleton*.

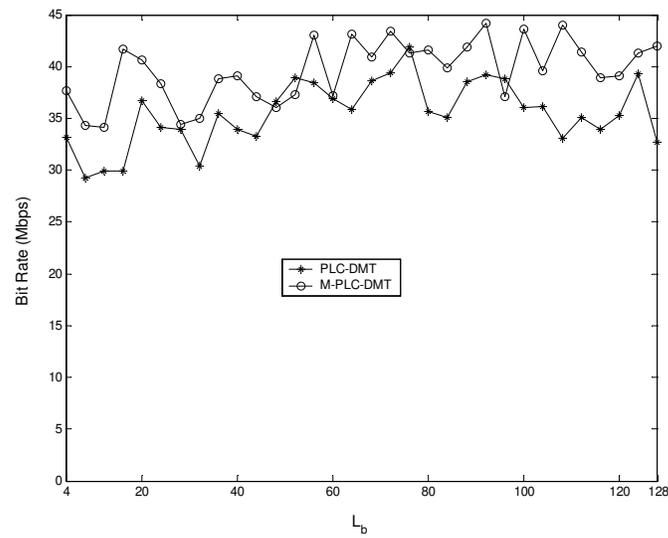


Figura 4.9: Desempenho dos sistemas DMT e DMT modificado em função de L_b quando $L_w = 88$ para o canal PLC 1.

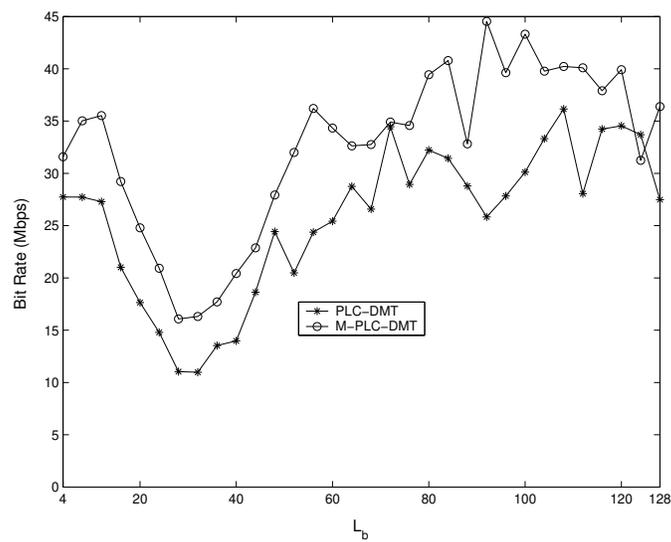


Figura 4.10: Desempenho dos sistemas DMT e DMT modificado em função de L_b quando $L_w = 104$ para o canal PLC 2.

Dentre as técnicas investigadas, constatou-se que os algoritmos fuzzy do tipo 2 *nonsingleton* apresentam os melhores desempenhos, enquanto o pior desempenho foi obtido com a rede wavelet. Dentre as redes analisadas, pode-se observar que as redes wavelet, RBF e MLP, necessariamente nessa ordem, promovem o aumento do desempenho dos sistemas DMT/OFDM.

5

Equalização de Canais PLC baseada em Sistemas Fuzzy

Recentemente, as técnicas baseadas em inteligência computacional (Ribeiro, 2005, 2003; Chen, Gibson, Cowan, & Grant, 1990; Mendel, 2000; Patra & Mulgrew, 1998a; Ibnkahla, 2000; Reyneri, 1999; Liang & Mendel, 2000c; Engelbrecht, 2002; Patra, Pal, Chatterji, & Panda, 1999; Mulgrew, 1996) têm sido amplamente utilizadas como soluções não-lineares para a equalização canais de comunicação. Dentre as técnicas investigadas, destacam-se aquelas baseadas em sistemas fuzzy. As questões de investigação relacionadas ao uso dos sistemas fuzzy para a equalização

de canais (Wang & Mendel, 1993; Patra & Mulgrew, 1998a, 1998b; Lee, 1994, 1994; Sarwal & Srinath, 1995; Lee, 1996; Liang & Mendel, 2000a; Patra & Mulgrew, 2000; Ribeiro, 2003; Dualibe, Jespers, & Verleysen, 2001) são as seguintes: *i*) a equalização de canais invariantes e variantes no tempo, *ii*) a minimização dos ruídos na saída do canal, com distribuição diferente da Gaussiana, *iii*) o cancelamento da ICC, *iv*) o desenvolvimento de novos equalizadores baseados nas estruturas transversais e com realimentação decidida para sinais unidimensionais e bidimensionais e *v*) o desenvolvimento de equalizadores fuzzy adaptativos. Mais recentemente, o uso de equalizadores fuzzy numa estrutura de turbo equalização (Loiola et al., 2004), de estratégias diferentes para a adaptação dos parâmetros dos equalizadores fuzzy (Ribeiro, 2005), de sistemas fuzzy do tipo 2 (Liang & Mendel, 2000c, 2000a) e da equalização cega, baseada em sistemas fuzzy do tipo 1, (Ferrari et al., 2003) tem sido investigado.

Fundamentado no fato de que os sistemas fuzzy do tipo 2 são técnicas heurísticas adequadas para a solução de problemas envolvendo elevados níveis de incertezas, neste capítulo é analisado o uso dos algoritmos fuzzy do tipo 2, propostos na seção 2.3, para a equalização de canais PLC. Tal análise fundamenta-se no fato de que os algoritmos baseados em sistemas fuzzy do tipo 2 são capazes de lidar com as incertezas presentes nestes canais, tais como a presença de ruídos impulsivos e não estacionários e, ainda, no fato de que tais canais são variantes no tempo.

Com esse objetivo, a equalização de canais de comunicação é brevemente discutida na seção 5.1. Na seqüência, a seção 5.2 mostra que o algoritmo fuzzy do tipo 2, proposto na seção 2.3.1, reduz-se ao equalizador Bayesiano baseado no critério ML. Finalmente, Resultados computacionais, mostrados na seção 5.3, ilustram que, apesar do aumento do carga computacional, estes equalizadores fuzzy do tipo 2 oferecem mais vantagens se comparados com seus antecessores, quando são aplicados para à equalização de canais PLC.

5.1 Equalização de canais PLC

Sejam os modelos em banda base e no domínio discreto de um sistema de comunicação digital mostrados nas Fig. 5.1 e 5.2, então o sinal na saída do canal de comunicação é dado por (Proakis, 2000; Qureshi, 1985)

$$y(n) = \tilde{y}(n) + v(n) = \sum_{i=0}^{L_h-1} h(i)x(n-i) + v(n) \quad (5.1)$$

sendo $\tilde{y}(n)$ a saída do canal livre de ruído, $\{h(n)\}_{n=0}^{L_h-1}$ a resposta impulsiva do canal PLC, L_h o comprimento da resposta impulsiva do canal PLC, $v(n)$ a n -ésima amostra do ruído aditivo formado, de acordo com a expressão (3.8), pela combinação dos diversos ruídos presentes nas saídas dos canais PLC, $x(n)$ o n -ésimo símbolo transmitido, $\tilde{y}(n)$ a n -ésima amostra na saída do canal livre do ruído aditivo. Nas Figs. 5.1 e 5.2, o bloco $Q(\cdot)$ implementa uma função de decisão. Neste capítulo é assumido que $x(n)$ assume valores equiprováveis, independentes e apresenta as seguintes estatísticas

$$\begin{aligned} E\{x(n)\} &= 0 \\ E\{(x(n_1)x(n_2))\} &= \sigma_x^2 \delta(n_1 - n_2) \end{aligned} \quad (5.2)$$

onde

$$\delta(n_1 - n_2) = \begin{cases} 1, & \text{se } n_1 = n_2 \\ 0, & \text{caso contrário} \end{cases} \quad (5.3)$$

e $E\{\cdot\}$ é o operador esperança.

A reconstrução dos símbolos transmitidos a partir da minimização da interferência intersimbólica (IIS) na saída do canal PLC é realizada por algoritmos denominados equalizadores de canais de comunicação. Tais equalizadores podem ser agrupados em dois grupos distintos: *i*) os equalizadores supervisionados e *ii*) os equalizadores não supervisionados (Haykin, 1996; Qureshi, 1985). No que tange a equalização supervisionada, aqui considerada, duas formas distintas de equalização se destacam:

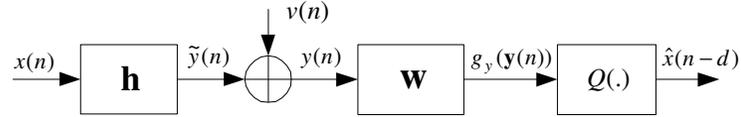


Figura 5.1: Modelo banda base para sistemas de comunicação com equalizadores transversais.

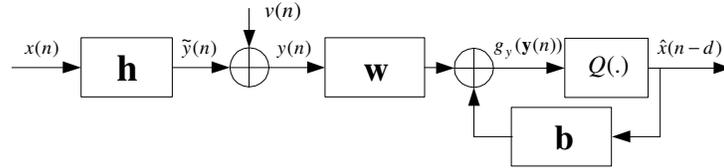


Figura 5.2: Modelo banda base para sistemas de comunicação com equalizadores que possuem realimentação decidida.

i) a equalização baseada na estimação da seqüência transmitida. Neste caso, o equalizador possui memória infinita e é denominado MLSE (*Maximum Likelihood Sequence Estimation*) (Forney, 1972). Conhecendo-se a resposta impulsiva do canal de comunicação, o MLSE pode ser implementado com o algoritmo de Viterbi (Jr., 1973). O MLSE, baseado no algoritmo de Viterbi, determina a seqüência $\{\hat{x}(1), \hat{x}(2), \dots\}$ de símbolos transmitidos ao minimizar a função custo dada por

$$J_{custo} = \sum_{n=1}^{\infty} \left[y(n) - \sum_{i=0}^{L_h-1} h(i)\hat{x}(n-i) \right]^2, \quad (5.4)$$

onde $\hat{x}(n)$ é o n -ésimo símbolo decidido. Deve-se ainda ressaltar que o algoritmo de Viterbi apresenta uma complexidade computacional proibitiva para a maioria dos canais de comunicação, pois ele demanda LMV operações para detectar uma seqüência de L símbolos que assumem M possíveis valores na presença de V inteferentes.

Sabe-se que o MLSE atinge o seu desempenho ótimo quando o canal de

comunicação é invariante no tempo e a estimação do mesmo é perfeita. Na prática, a estimação do canal não é exata e, portanto, há perda de desempenho do MLSE. A literatura mostra que se o canal de comunicação é invariante, é possível estimar com grande precisão a sua resposta impulsiva e, conseqüentemente, a perda de desempenho é desprezível para estes canais. No entanto, o desempenho do MLSE é bastante degradado quando os canais de comunicação são variantes no tempo (Chen, McLaughlin, Mulgrew, & Grant, 1995). De fato, a estimação de canais variantes no tempo é bastante problemática, posto que os coeficientes das respostas impulsivas destes canais variam ao longo do tempo.

ii) a equalização baseada na estimação símbolo a símbolo da seqüência transmitida. A decisão de um equalizador símbolo a símbolo pode ser considerado como uma função de um vetor contendo as amostras passadas na saída do canal. Esta função de decisão pode ser linear e, conseqüentemente, o equalizador resultante é dito ser do tipo linear. Os equalizadores lineares são amplamente conhecidos por sua baixa complexidade computacional. No entanto, é sabido que os mesmos são inadequados para a equalização de canais com nullos espectrais, uma vez que eles geram *noise enhancement*. Além disso, eles não são adequados para a equalização de canais seletivos em freqüência (Proakis, 2000).

Por outro lado, se a função de decisão é não-linear, então o equalizador é não linear. A função de decisão ótima é não-linear e é obtida através do uso do critério *Maximum a Posteriori Probability* (MAP) (Abend & Fritchman, 1970), derivado da teoria de Bayes (Duda & Hart, 1973; Trees, 1968/1971), é utilizado. Estes equalizadores símbolo a símbolo ótimo é denominado equalizadores Bayesianos. Fazendo-se algumas restrições nesses equalizadores, obtém-se variações dos mesmos com diferentes relações de desempenho versus complexidade computacional em comparação ao equalizador símbolo a símbolo ótimo (Chen, McLaughlin, Mulgrew, & Grant, 1996; Chen, Mulgrew, & Hanzo, 2000; Chen, Mulgrew, & McLaughlin,

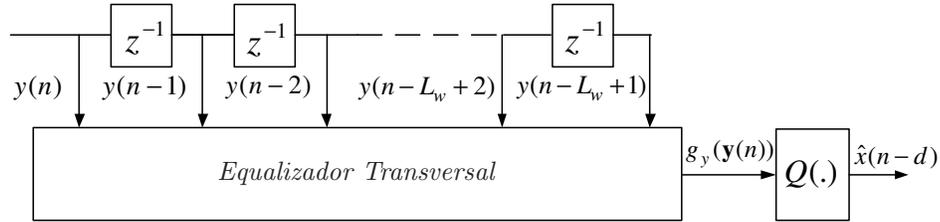


Figura 5.3: Esquema de um equalizador transversal (TE).

1993; Chen et al., 1995; Geordiadis & Mulgrew, 2001; Montalvão, Dorizzi, & Mota, 2001; Chen & Wu, 2004; Lee & Hanzo, 2002). A principal vantagem desses equalizadores advém do fato de que os mesmos podem ter desempenhos próximos daqueles alcançados com o MLSE demandando menor complexidade computacional. Tais equalizadores geram o limitante superior de desempenho para os equalizadores símbolo a símbolo e podem ser implementados de forma aproximada por meio de soluções lineares e não lineares. Além de serem aplicados à equalização supervisionada de canais de comunicação, os equalizadores Bayesianos têm sido, recentemente, desenvolvidos para a equalização não supervisionada de canais (José-Revuelta & Sueiro, 2004; Lin & Yamashita, 2001)

As duas estruturas principais normalmente utilizadas para a implementação dos equalizadores símbolo a símbolo são a estrutura transversal (*Transversal Equalizer* - TE), vide a Fig. 5.3 e a estrutura com decisão realimentada (*Decision Feedback Equalizer* - DFE) (Austin, 1967), vide a Fig. 5.4.

A seguinte formulação é aplicada para os equalizadores símbolos a símbolos:

Sejam um alfabeto dos símbolos transmitidos dado por $S := \{x(n) = s_k | k = 1, \dots, N_S\}$, o i -ésimo vetor de símbolos transmitidos dado por $\mathbf{x}_i(n) = [x(n) \cdots x(n - L_h - L_w + 1)]^T$, o i -ésimo vetor $\tilde{\mathbf{y}}_i(n) = [\tilde{y}(n) \cdots y(n - L_w + 1)]^T$ formado pelas amostras livre do ruído aditivo na saída do canal, o vetor $\mathbf{v}_i(n) = [v(n) \cdots v(n -$

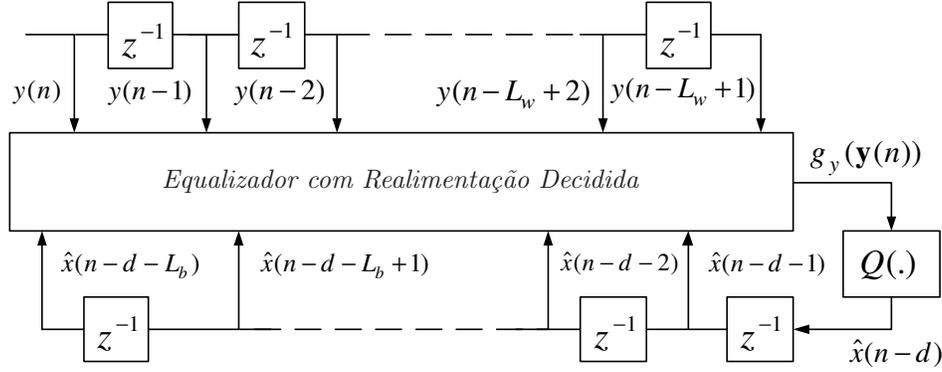


Figura 5.4: Esquema do equalizador com realimentação decidida (DFE).

$L_w + 1)]^T$ constituído por amostras do ruído aditivo e a matriz de convolução do canal de comunicação dada por

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{L_h-1} & \cdots & 0 \\ 0 & h_0 & \cdots & h_{L_h-1} & \cdots & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & h_0 & \cdots & h_{L_h-2} & h_{L_h-1} \end{bmatrix}, \quad (5.5)$$

então o i -ésimo vetor de entrada do equalizador transversal, cuja dimensão é L_w , é dado por

$$\mathbf{y}_i(n) = \tilde{\mathbf{y}}_i(n) + \mathbf{v}_i(n) = \mathbf{H}\mathbf{x}_i(n) + \mathbf{v}_i(n). \quad (5.6)$$

Agora, seja $N_{\mathbf{x}_i} = (N_S)^{L_h+L_w}$ o número de possíveis vetores $\mathbf{x}_i(n)$, $1 \leq i \leq N_{\mathbf{x}_i}$, na entrada do canal de comunicação, os quais constituem o espaço de estados

$$\mathcal{X} = \{\mathbf{x}_i(n) | 1 \leq i \leq N_{\mathbf{x}_i}\}, \quad (5.7)$$

$$\mathcal{Y} = \{\tilde{\mathbf{y}}_i(n) | 1 \leq i \leq N_{\tilde{\mathbf{y}}_i}\} \quad (5.8)$$

o espaço de estados formado pelos vetores livres de ruídos, $\tilde{\mathbf{y}}_i(n)$, na saída do canal PLC e assumindo-se que o equalizador Bayesiano, tem-se que a estimativa da n -ésima amostra do sinal transmitido atrasada de d amostras, $\hat{x}(n-d) = s_k$, está associada ao subespaço de estados

$$\tilde{\mathcal{Y}}^k = \{\tilde{\mathbf{y}}_{k_i}(n), i = 1, \dots, N_{\tilde{\mathcal{Y}}^k} | \hat{x}(n-d) = s_k\}, \quad k = 1, \dots, N_S, \quad (5.9)$$

onde $\tilde{\mathbf{y}}_{k_i}(n)$, $i = 1, \dots, N_{\tilde{\mathcal{Y}}^k}$ são os vetores de estado na saída do canal que constituem o k -ésimo subespaço de estados $\tilde{\mathcal{Y}}^k$, $N_{\tilde{\mathcal{Y}}^k}$ é o número de vetores de estados na saída do canal de comunicação que constitui o subespaço de estados $\tilde{\mathcal{Y}}^k$. O espaço de estados $\tilde{\mathcal{Y}}$ formado por todos os vetores de estados na saída do canal de comunicação é dividido em N_S subespaço de estados, $\tilde{\mathcal{Y}}^k$, $k = 1, \dots, N_S$. Conseqüentemente, pode-se escrever $\tilde{\mathcal{Y}} = \cup_{k=1}^{N_S} \tilde{\mathcal{Y}}^k$. Cada um dos subespaço de estados está associado a um símbolo $\hat{x}(n-d) = s_k$ transmitido. O número de vetores no k -ésimo subespaço de estados, $\tilde{\mathcal{Y}}^k$, é dado por $N_{\tilde{\mathcal{Y}}^k} = \frac{(N_S)^{L_h + L_w}}{N_S}$. Note que o atraso de equalização, d , pode ser estimado de acordo com (Chng & Chen, 2005).

Supondo-se que as distribuições de probabilidades do ruído aditivo e dos símbolos transmitidos são dadas, pode-se aplicar o critério MAP para a dedução do equalizador ótimo. Para o caso em que assume-se que os símbolos transmitidos são equiprováveis, tem-se que o critério MAP reduz-se ao critério *maximum likelihood* (ML). Neste caso, os equalizadores Bayesianos são ótimos de acordo com o critério ML. Supondo-se que os símbolos transmitidos não sejam equiprováveis e tem função distribuição de probabilidade não conhecidas, o equalizador Bayesiano baseado no critério ML fornece os melhores resultados.

Os equalizadores transversais e Bayesianos podem ser aproximados pelos equalizadores transversais lineares ou não lineares. Aproximando-se os equalizadores transversais e Bayesianos pelo filtro transversal e linear, cuja estrutura é mostrada

na Fig. 5.3, tem-se que a saída do mesmo é dada por

$$\hat{x}(n-d) = Q(g_y(\mathbf{y}(n))). \quad (5.10)$$

onde

$$g_y(\mathbf{y}(n)) = \mathbf{w}^T \mathbf{y}(n) = \sum_{i=0}^{L_w-1} w_i y(n-i), \quad (5.11)$$

$$\mathbf{y}(n) = [y(n) \quad \cdots \quad y(n-L_w+1)]^T, \quad (5.12)$$

$Q(\cdot)$ é a função de decisão e $\mathbf{w} = [w_0 \cdots w_{L_w-1}]^T$ é o vetor de coeficientes do filtro transversal que, geralmente, é obtido de forma adaptativa. O desenvolvimento e a utilização de diferentes critérios para a atualização do vetor \mathbf{w} permite que diferentes níveis de desempenhos sejam obtidos com os equalizadores transversais e lineares.

Apesar dos bons resultados obtidos com os equalizadores transversais, os equalizadores baseados na estrutura transversal de decisão com realimentação (DFE) (vide Fig. 5.4) apresentam desempenhos melhores. De forma simplificada, pode-se dizer que os equalizadores DFE são obtidos por meio da expansão da entrada do equalizador transversal para incluir o vetor de símbolos

$$\hat{\mathbf{x}}(n-d) = [\hat{x}(n-d-1) \quad \cdots \quad \hat{x}(n-d-L_b)]^T, \quad (5.13)$$

onde L_b é a ordem do filtro de realimentação, constituído dos símbolos $\hat{x}(n-d-i)$, $i = 1, \dots, L_b$ detectados anteriormente. Enquanto a saída dos equalizadores DFE (Austin, 1967; Qureshi, 1985) é dada por

$$g_y(\mathbf{y}(n)) = \mathbf{w}^T \mathbf{y}(n) + \mathbf{b}^T \hat{\mathbf{x}}(n-d) = \sum_{i=0}^{L_w-1} w_i y(n-i) + \sum_{i=0}^{L_b-1} b_i \hat{x}(n-d-i-1), \quad (5.14)$$

a saída decidida dos mesmos é dada por

$$\hat{x}(n-d) = Q(g_y(\mathbf{y}(n))), \quad (5.15)$$

sendo $\mathbf{b} = [b_0 \quad \cdots \quad b_{L_b-1}]^T$ o vetor de coeficiente do filtro de realimentação e $Q(\cdot)$ uma função de decisão. Devido ao fato de que o vetor $\hat{\mathbf{x}}(n-d)$ é constituído de

amostras obtidas na saída da função $Q(\cdot)$ não linear, o equalizador DFE é dito ser um equalizador não linear.

Agora, seja

$$L_b = L_h + L_w - 1 - d \quad (5.16)$$

o comprimento do vetor de realimentação \mathbf{b} , $N_b = (N_s)^{L_b}$ o número de estado do vetor \mathbf{b} e $\mathbf{x}_{\mathbf{b}j}$, $1 \leq j \leq N_b$ os estados de \mathbf{b} , então o número de estados de $\tilde{\mathcal{Y}}^k$ é agora dado por $N_{\tilde{\mathcal{Y}}^k} = \frac{(N_s)^{L_h+L_w}}{(N_s)^{L_h+L_w-1-d}} = (N_s)^{d+1}$. Note que o espaço de estados $\tilde{\mathcal{Y}}^k$ gerados na saída do canal é agora dado por

$$\tilde{\mathcal{Y}}^k = \{\tilde{\mathbf{y}}_{k_i}(n), i = 1, \dots, N_{\tilde{\mathcal{Y}}^k} | \hat{\mathbf{x}}(n-d) = s_k \cap \hat{\mathbf{x}}(n-d-1) = \mathbf{x}_{\mathbf{b}j}\}, \quad (5.17)$$

sendo $k = 1, \dots, N_s$, $j = 1, \dots, N_b$ e $\hat{\mathbf{x}}(n-d-1) = [\hat{x}(n-d-1) \cdots \hat{x}(n-d-L_b)]^T$. Conforme demonstrado em (Chen et al., 1993), tal solução reduz consideravelmente a complexidade computacional demandada pelo equalizador Bayesiano transversal, pois L_w é bem menor na estrutura DFE.

Daqui em diante adotaremos a seguinte consideração:

$$L = \begin{cases} L_w, & \text{se estrutura TE} \\ L_w + L_b, & \text{se estrutura DFE} \end{cases} \quad (5.18)$$

Para o caso de $L = L_w$, tem-se o vetor

$$\mathbf{y}(n) = [y(n) \cdots y(n-L_w+1)]^T \quad (5.19)$$

como entrada do equalizador e para $L = L_w + L_b$, o vetor na entrada do equalizador é dado por

$$\mathbf{y}(n) = [y(n) \cdots y(n-L_w+1) \hat{x}(n-d-1) \cdots \hat{x}(n-d-L_b)]^T. \quad (5.20)$$

Uma consideração importante a ser feita nas expressões mostradas acima é que o canal é considerado invariante no tempo. No entanto, se os canais de comunicação

forem variantes no tempo, os estados na saída dos mesmos apresentam dispersões, além daquela provocada pela presença dos ruídos aditivos nessas saídas. Para os canais variantes no tempo é evidente que os estados na saída dos mesmos são regiões $R_{\tilde{\mathbf{y}}_i}$ e não pontos no espaço de L dimensões. Conseqüentemente, se $R_{\tilde{\mathbf{y}}_i} \cap R_{\tilde{\mathbf{y}}_j} \neq \{0\}$, $i \neq j$, então a equalização do canal de comunicação, sem a adição de ruído, é comprometida. De fato, a sobreposição dessas regiões devido às incertezas nos estados dos canais de comunicação compromete de forma significativa a equalização dos canais de comunicação, posto que não é possível precisar, com baixa probabilidade de erro, a qual símbolo transmitido, $x(n) = s_k$, $k = 1, \dots, N_S$, está associada a região $R_{\tilde{\mathbf{y}}_i}$, observada na saída do canal de comunicação. Adicionando-se a este problema a presença de ruídos aditivos com componentes não estacionárias, tal como em canais PLC, pode-se afirmar que a equalização desses canais é bastante complicada. De fato a ocorrência de incertezas geradas pelos canais de comunicação variantes no tempo e pelos ruídos aditivos com características não estacionárias resultam em incertezas que degradam consideravelmente o desempenho dos equalizadores.

O fato dos estados na saída dos canais de comunicação, tais como os canais PLC, serem regiões e não pontos do espaço vetorial $\tilde{\mathcal{Y}}$, conduz a conclusão de que o uso de técnicas de equalização baseadas em sistemas fuzzy do tipo 2 oferecem ganhos consideráveis em termos de desempenho, conforme apontado em (Liang & Mendel, 2000a, 2000c). De fato, como destacado no capítulo 2, os sistemas fuzzy do tipo 2 são adequados para resolver problemas cujos níveis de incertezas são elevados, tal como o problema colocado acima.

No capítulo 2 desta tese, novos algoritmos fuzzy do do tipo 2, intervalado e *nonsingleton* foram introduzidos, os quais podem ser aplicados para a equalização de canais PLC. Com este objetivo, a seção 5.2 mostra que o algoritmo fuzzy do tipo 2, intervalado e *nonsingleton*, proposto na seção 2.3, pode, sob algumas restrições,

ser reduzido ao equalizador Bayesiano baseado no critério ML quando o ruído aditivo é tem distribuição Gaussiana.

5.2 Relações entre os Equalizadores Fuzzy do Tipo 2 e Bayesiano baseado no critério ML

Para mostrar que o equalizador fuzzy do tipo 2, intervalado, *nonsingleton*, deduzido na seção 2.3, reduz-se ao equalizador Bayesiano, consideramos que os símbolos transmitidos $x(n) \in S = \{+1, -1\}$, a resposta impulsiva do canal é invariante no tempo e dada por $\{h(n)\}_{n=0}^{L_h-1}$, a validade das restrições explicitadas por (5.2) e supondo-se que equalizador tem estrutura transversal. Note que as deduções mostradas abaixo podem ser facilmente extendidas para outros tipos de modulação (M-PAM, M-QAM, M-PSK, etc) e/ou quando o equalizador tem estrutura DFE.

A seguir, primeiramente, deduzimos o equalizador Bayesiano ótimo quando o ruído aditivo na saída do canal de comunicação é Gaussiano. Em seguida, deduzimos o equalizador fuzzy do tipo 1 ótimo quando funções de pertinência dos dados nas entradas e das regras são funções gaussianas. Além disso, mostraremos que tal equalizador fuzzy reduz-se ao equalizador Bayesiano ótimo quando algumas restrições são consideradas. Finalmente, baseando-se na expressão (2.31) e na seção 2.3.1, o equalizador fuzzy do tipo 2, intervalado, *nonsingleton* e ótimo é deduzido e é mostrado que o mesmo com algumas algumas restrições reduz-se ao equalizador Bayesiano ótimo e ao equalizador fuzzy do tipo 1 e *nonsingleton*, descritos acima.

Supondo-se que a saída do canal livre de ruído do canal PLC seja corrompido pela presença de ruído aditivo *i.i.d.*, independente do sinal $x(n)$ e que as amostras deste ruído tenham distribuição Gaussiana $\mathcal{N}(0, \sigma_v^2)$, então, baseando-se na teoria

de Bayes, tem-se que a saída do equalizador Bayesiano ótimo e normalizado é dado por

$$f_b(\mathbf{y}(n)) = \frac{\left\{ \sum_{\mathbf{y}_i \in \mathcal{Y}_d^+} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_i\|^2}{2\sigma_v^2}\right) - \sum_{\mathbf{y}_j \in \mathcal{Y}_d^-} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_j\|^2}{2\sigma_v^2}\right) \right\}}{\sum_{\mathbf{y}_k \in \mathcal{Y}_d} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_k\|^2}{2\sigma_v^2}\right)}, \quad (5.21)$$

onde $\mathcal{Y}_d = \mathcal{Y}_d^+ \cup \mathcal{Y}_d^-$ é o espaço de estados constituído por todos os vetores \mathbf{y}_i , cuja dimensão é L_w , obtidos com a expressão (5.6), $\mathcal{Y}_d^+ = \{\mathbf{y}_i(n), i = 1, \dots, 2^{L_h-1} | x(n-d) = +1\}$ e $\mathcal{Y}_d^- = \{\mathbf{y}_j(n), j = 1, \dots, 2^{L_h-1} | x(n-d) = -1\}$. Note que disconsideramos os termos que multiplicam o lado direito da expressão (5.21) para simplificar a mesma. A superfície de separação entre os vetores de estados \mathcal{Y}_d^+ e \mathcal{Y}_d^- é dada por

$$f_b(\mathbf{y}(n)) = 0. \quad (5.22)$$

Agora, seja o sistema fuzzy do tipo 1 e *nonsingleton* (Mouzouris & Mendel, 1997), então, supondo-se que as funções Gaussianas sejam aplicadas para modelar as premissas dos antecedentes das regras fuzzy e as incertezas nas entradas do sistema fuzzy, as *t-norms* são implementadas pelo operador produto, o método de defuzzificação é pelas alturas e a contribuição das regras associadas a $\mathcal{Y}_d^+ = \{\mathbf{y}_i(n), i = 1, \dots, 2^{L_h-1} | x(n-d) = +1\}$ e $\mathcal{Y}_d^- = \{\mathbf{y}_j(n), j = 1, \dots, 2^{L_h-1} | x(n-d) = -1\}$ sejam iguais à $1/2^{L_h-1}$ e $-1/2^{L_h-1}$, respectivamente. Então, a saída do equalizador fuzzy do tipo 1 e *nonsingleton* ótimo é dada por

$$f_{ns1}(\mathbf{y}(n)) = \frac{\left\{ \sum_{\mathbf{y}_i \in \mathcal{Y}_d^+} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_i\|^2}{2(\sigma_r^2 + \sigma_v^2)}\right) - \sum_{\mathbf{y}_j \in \mathcal{Y}_d^-} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_j\|^2}{2(\sigma_r^2 + \sigma_v^2)}\right) \right\}}{\sum_{\mathbf{y}_k \in \mathcal{Y}_d} \exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_k\|^2}{2(\sigma_r^2 + \sigma_v^2)}\right)}, \quad (5.23)$$

onde σ_r^2 é a variância da função de pertinência que modela as premissas dos antecedentes das regras fuzzy. Para a dedução da expressão (5.23) consideramos que todas as entradas do equalizador fuzzy do tipo 1 e *nonsingleton* e as premissas

dos antecedentes das regras fuzzy apresentam os mesmos níveis de incertezas, os quais são dados por σ_r^2 e σ_v^2 . Note ainda que os vetores \mathbf{y}_i , \mathbf{y}_j e \mathbf{y}_k substituem os vetores \mathbf{m}_i , \mathbf{m}_j e \mathbf{m}_k , os quais correspondem às médias das funções de pertinência das regras fuzzy alocadas para os estados em \mathcal{Y}_d^+ , \mathcal{Y}_d^- e \mathcal{Y}_d .

A superfície de separação gerada pelo classificador fuzzy do tipo 1 e *nonsingleton* para separar \mathcal{Y}_d^+ de \mathcal{Y}_d^- é dada por

$$f_{ns1}(\mathbf{y}(n)) = 0. \quad (5.24)$$

Pode-se observar que

$$\frac{f_b(\mathbf{y}(n))}{f_{ns1}(\mathbf{y}(n))} = 1 \quad (5.25)$$

desde que $\sigma_r^2 = 0$, ou seja, o equalizador fuzzy do tipo 1 e *nonsingleton* ótimo, cuja saída é dada por (5.23), torna-se igual ao equalizador Bayesiano normalizado quando a incerteza na regra fuzzy é igual a zero.

Baseando-se na expressão (2.31) e na seção 2.3.1, tem-se que a saída do equalizador fuzzy do tipo 2, intervalado, *nonsingleton* e ótimo é obtida da seguinte maneira

Sejam

$$\bar{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \bar{m}_{y_i}(n)}{\bar{\sigma}_{y_i}(n)} \right)^2 \right] \quad (5.26)$$

e

$$\underline{\mu}_{\tilde{Y}_i}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \underline{m}_{y_i}(n)}{\underline{\sigma}_{y_i}(n)} \right)^2 \right] \quad (5.27)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das incertezas nas entradas de um sistema fuzzy do tipo 2, intervalado e *nonsingleton*,

$$\bar{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \bar{m}_{\tilde{F}_i^l}(n)}{\bar{\sigma}_{\tilde{F}_i^l}(n)} \right)^2 \right] \quad (5.28)$$

e

$$\underline{\mu}_{\tilde{F}_i^l}(y_i(n)) = \exp \left[-\frac{1}{2} \left(\frac{y_i(n) - \underline{m}_{\tilde{F}_i^l}(n)}{\underline{\sigma}_{\tilde{F}_i^l}(n)} \right)^2 \right], \quad (5.29)$$

as funções *upper* e *lower* de pertinência que definem as FOUs das incertezas nas premissas dos antecedentes das regras fuzzy do sistemas fuzzy do tipo 2, intervalado e *nonsingleton*. Então, tem-se que o supremo entre as funções *upper* de pertinência e o supremo entre as funções *lower* de pertinência que permitem relacionar as incertezas presentes nas entradas com as incertezas presentes nas premissas das regras fuzzy são dados por

$$\bar{\mu}_{\tilde{Q}_i^l}(\bar{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \bar{\mu}_{\tilde{Y}_i}(y_i(n)) \star \bar{\mu}_{\tilde{F}_i^l}(y_i(n)), \quad (5.30)$$

e

$$\underline{\mu}_{\tilde{Q}_i^l}(\underline{y}_{i,\max}^l(n)) = \sup_{y_i(n) \in \tilde{Y}_i} \underline{\mu}_{\tilde{Y}_i}(y_i(n)) \star \underline{\mu}_{\tilde{F}_i^l}(y_i(n)), \quad (5.31)$$

onde o símbolo \star é uma *t-norm* implementada pelo operador produto, $y_i(n)$ é a i -ésima amostra do vetor $\mathbf{y}(n) = [y_0(n) \cdots y_{L-1}(n)]^T$ de entrada. Para o nosso sistema fuzzy do tipo 2, as expressões (5.30) e (5.31) alcançam os seus máximos nas seguintes condições (Mendel, 2001; Mouzouris & Mendel, 1997):

$$\bar{y}_{i,\max}^l(n) = \frac{\bar{\sigma}_{y_i}^2(n) \bar{m}_{\tilde{F}_i^l}(n) + \bar{\sigma}_{\tilde{F}_i^l}^2(n) \bar{m}_{y_i}(n)}{\bar{\sigma}_{y_i}^2(n) + \bar{\sigma}_{\tilde{F}_i^l}^2(n)} \quad (5.32)$$

e

$$\underline{y}_{i,\max}^l(n) = \frac{\underline{\sigma}_{y_i}^2(n) \underline{m}_{\tilde{F}_i^l}(n) + \underline{\sigma}_{\tilde{F}_i^l}^2(n) \underline{m}_{y_i}(n)}{\underline{\sigma}_{y_i}^2(n) + \underline{\sigma}_{\tilde{F}_i^l}^2(n)}. \quad (5.33)$$

A saída do equalizador fuzzy do tipo 2, intervalado, *nonsingleton* e ótimo é dada por

$$f_{ns2}(\mathbf{y}(n)) = f_{ns2,\{1\}}(\mathbf{y}(n)) - f_{ns2,\{-1\}}(\mathbf{y}(n)) \quad (5.34)$$

onde

$$f_{ns2,\{1\}}(\mathbf{y}(n)) = \frac{\underline{f}_{ns2,\{1\}}(\mathbf{y}(n)) + \bar{f}_{ns2,\{1\}}(\mathbf{y}(n))}{\underline{f}_{ns2}(\mathbf{y}(n)) + \bar{f}_{ns2}(\mathbf{y}(n))}, \quad (5.35)$$

$$f_{ns2,\{-1\}}(\mathbf{y}(n)) = \frac{\underline{f}_{ns2,\{-1\}}(\mathbf{y}(n)) + \bar{f}_{ns2,\{-1\}}(\mathbf{y}(n))}{\underline{f}_{ns2}(\mathbf{y}(n)) + \bar{f}_{ns2}(\mathbf{y}(n))}, \quad (5.36)$$

$$\underline{f}_{ns2,\{1\}}(\mathbf{y}(n)) + \bar{f}_{ns2,\{1\}}(\mathbf{y}(n)) = \sum_{\tilde{\mathbf{y}}_i, \bar{\mathbf{y}}_i \in \mathcal{Y}_d^+} \left[\exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_i\|^2}{2(\underline{\sigma}_r^2 + \underline{\sigma}_v^2)}\right) + \exp\left(\frac{-\|\mathbf{y}(n) - \bar{\mathbf{y}}_i\|^2}{2(\bar{\sigma}_r^2 + \bar{\sigma}_v^2)}\right) \right], \quad (5.37)$$

$$\underline{f}_{ns2,\{-1\}}(\mathbf{y}(n)) + \bar{f}_{ns2,\{-1\}}(\mathbf{y}(n)) = \sum_{\tilde{\mathbf{y}}_j, \bar{\mathbf{y}}_j \in \mathcal{Y}_d^-} \left[\exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_j\|^2}{2(\underline{\sigma}_r^2 + \underline{\sigma}_v^2)}\right) + \exp\left(\frac{-\|\mathbf{y}(n) - \bar{\mathbf{y}}_j\|^2}{2(\bar{\sigma}_r^2 + \bar{\sigma}_v^2)}\right) \right], \quad (5.38)$$

e

$$\underline{f}_{ns2}(\mathbf{y}(n)) + \bar{f}_{ns2}(\mathbf{y}(n)) = \sum_{\tilde{\mathbf{y}}_k, \bar{\mathbf{y}}_k \in \mathcal{Y}_d} \left[\exp\left(\frac{-\|\mathbf{y}(n) - \tilde{\mathbf{y}}_k\|^2}{2(\underline{\sigma}_r^2 + \underline{\sigma}_v^2)}\right) + \exp\left(\frac{-\|\mathbf{y}(n) - \bar{\mathbf{y}}_k\|^2}{2(\bar{\sigma}_r^2 + \bar{\sigma}_v^2)}\right) \right]. \quad (5.39)$$

Note que os vetores $\tilde{\mathbf{y}}_k$ e $\bar{\mathbf{y}}_k$ representam as incertezas associadas aos centro do k -ésimo estado na saída do canal de comunicação.

A superfície de separação entre os vetores de estados \mathcal{Y}_d^+ e \mathcal{Y}_d^- gerada pelo classificador fuzzy do tipo 2, intervalado e *nonsingleton*, deduzido acima, é dada por

$$f_{ns2}(\mathbf{y}(n)) = 0. \quad (5.40)$$

Conforme pode ser observado nas expressões (5.41)-(5.39), o equalizador fuzzy do tipo 2, intervalado, *nonsingleton* e ótimo é adequado para os problemas de

equalização quando existem incertezas nas funções de pertinências que modelam as incertezas presentes no ruído aditivo e nas premissas das regras fuzzy. Supondo-se que tais incertezas não existem, tem-se que $\underline{\sigma}_r = \overline{\sigma}_r = \sigma_r$, $\underline{\sigma}_v = \overline{\sigma}_v = \sigma_v$, $\underline{\tilde{\mathbf{y}}}_k = \overline{\tilde{\mathbf{y}}}_k = \tilde{\mathbf{y}}_k$ e, conseqüentemente, tem-se que

$$f_{ns2}(\mathbf{y}(n)) = f_{ns1}(\mathbf{y}(n)), \quad (5.41)$$

onde $f_{ns1}(\mathbf{y}(n))$ é dado por (5.23). Em outras palavras, pode-se dizer que o equalizador fuzzy do tipo 2, intervalado, *nonsingleton* e ótimo baseado no algoritmo fuzzy proposto na seção 2.3.1 pode ser reduzido a um equalizador fuzzy do tipo 2 e *nonsingleton* que sob algumas restrições reduz-se ao equalizador Bayesiano baseado no critério ML.

É importante salientar que as relações mostradas acima são, dadas algumas restrições, possíveis de serem explícitas, pois as funções de pertinência que modelam as incertezas nas entradas dos equalizadores fuzzy e as incertezas nas premissas das regras fuzzy são Gaussianas. Para o caso em que tais incertezas não sejam concomitantemente Gaussianas não é possível mostrar analiticamente relações semelhantes às apresentadas acima. De fato, a expressão que relaciona as incertezas nas entradas com as incertezas nas regras não produz, a princípio, nenhuma resultado que permite uma comparação com equalizadores Bayesianos e fuzzy do tipo 1, quando outros tipos de funções de pertinência são consideradas. Como exemplo, pode-se citar os algoritmos fuzzy do tipo 1 e *nonsingleton* e fuzzy do tipo 2, intervalado e *nonsingleton*, propostos na seção 2.3.2, que fazem uso da função exponencial para modelar as incertezas nas entradas e da função Gaussiana para modelar as premissas dos antecedentes das regras fuzzy. Para estes algoritmos não é possível mostrar relações com os equalizadores Bayesianos.

Substituindo o equalizador transversal pelo equalizador com realimentação decidida (DFE), equações equivalentes às deduzidas acima são obtidas. Para o caso em que os sinais sejam bidimensionais, basta considerar a função de pertinência

sendo bidimensional, conforme foi proposto em (Wei & Mendel, 2000) e proceder de forma idêntica para deduzir os equalizadores ótimos.

5.3 Resultados Numéricos

A seção 5.3.1 apresenta resultados que confirmam a eficiência dos métodos BP, DRD e LLE para a atualização dos parâmetros dos algoritmos fuzzy propostos na seção 2.3 e na seção 5.3.2 alguns resultados são mostrados para ilustrar o fato de que os algoritmos fuzzy do tipo 2, propostos na seção 2.3, apresentam desempenho superior quando comparados com os algoritmos baseados em sistemas fuzzy do tipo 1. Os resultados computacionais mostrados nesta seção foram obtidos com os algoritmos implementados com a linguagem C.

5.3.1 Estratégias de Atualização do Passo de Adaptação

Visando verificar o desempenho dos métodos LLE, DRD e BP de otimização, os quais são baseados em informações de 1ª ordem, quando são aplicados para a atualização dos parâmetros dos algoritmos fuzzy do tipo 1 utilizados para a equalização de canais PLC, nesta seção comparamos a convergência dos equalizadores fuzzy-S-LMS-DRD, fuzzy-S-LMS-LLE, fuzzy-S-DFE-DRD, fuzzy-S-DFE-LLE, fuzzy-NS-LMS-DRD, fuzzy-NS-LMS-LLE, fuzzy-NS-DFE-DRD e fuzzy-NS-DFE-LLE com aquelas obtidas com os equalizadores fuzzy-S-LMS, fuzzy-S-DFE, fuzzy-NS-LMS e fuzzy-NS-DFE (Lee, 1994, 1994; Sarwal & Srinath, 1995; Lee, 1996; Wang & Mendel, 1993; Ribeiro, 2003). Para esses equalizadores fuzzy, *S* e *NS* significam *singleton* e *nonsingleton*, *LMS* significa o uso do método BP para a atualização, *LMS-DRD* e *LMS-LLE* referem-se ao uso dos métodos DRD

e LLE, respectivamente, para a atualização dos parâmetros dos equalizadores. Os equalizadores fuzzy que têm no seu nome a sigla DFE referem-se ao equalizadores com realimentação decidida, cuja estrutura é mostrada na Fig. 5.4. Os equalizadores listados acima, cujo nome não contém a sigla DFE são os equalizadores fuzzy transversais, cuja estrutura é ilustrada na Fig. 5.3. Os equalizadores fuzzy *nonsingleton*, implementados nesta seção, são aqueles que consideram a função Gaussiana para modelar as incertezas nas entradas do equalizador fuzzy. Note que tais incertezas são geradas pela presença dos ruídos aditivos em canais PLC *outdoor*.

Para a obtenção dos resultados mostrados nesta seção, as seguintes considerações foram feitas: *i*) o canal PLC e a modelagem do ruído impulsivo são aqueles discutidos nas seções 3.1 e 3.2 do capítulo 3; *ii*) o canais PLC foi normalizados; *iii*) a faixa de frequência utilizada é entre 1 MHz e 2.5 MHz; *iv*) a potência dos símbolos BPSK transmitidos e dos ruídos impulsivos são iguais a $\sigma_x^2 = 0$ dB e $\sigma_{imp}^2 = 0$ dB, respectivamente; *v*) a potência do ruído de fundo σ_{bkg}^2 varia entre -2.5 dB e -20 dB; *vi*) L_w , L_b , M e d foram fixados em 15, 8, 100 e 0, respectivamente; *vii*) o passo de adaptação para a atualização dos equalizadores fuzzy-S-LMS, fuzzy-S-DFE, fuzzy-NS-LMS e fuzzy-NS-DFE foi fixado em 0,001; *viii*) para o método DRD adotou-se $\alpha \in [0, 1, 0, 4]$, $K_1 \in [0, 0001, 0, 001]$, $K_2 \in [0, 6, 1, 0]$ e o passo de atualização inicial igual a 0,03; *ix*) para a técnica LLE adotou-se $\lambda_0 = 4$, $m = 1$ e $q = 1, 038$, *x*) a saída dos equalizadores fuzzy do tipo 1 e *nonsingleton* e das equações de atualização de seus parâmetros são dadas pelas expressões (2.100)-(2.108), *xi*) os equalizadores fuzzy do tipo 1 e *singleton* são obtidos quando as incertezas nos dados de entrada é desconsiderada. Neste caso, as saídas de tais equalizadores e as equações de atualização dos parâmetros dos mesmos podem ser derivadas a partir de (2.100)-(2.108) (vide (Mendel, 2001)), finalmente, *xii*) os parâmetros dos equalizadores foram igualmente inicializados.

Enquanto a Fig. 5.5 mostra o canal PLC utilizado nas simulações, a Fig. 5.6

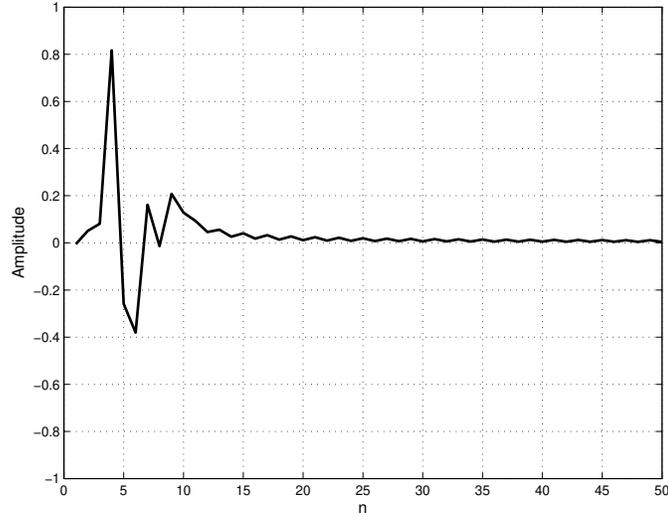


Figura 5.5: Resposta impulsiva do canal PLC utilizado nas simulações. O comprimento do canal foi truncado, resultando em $L_h = 50$.

mostra uma seqüência gerada a partir do modelo do ruído aditivo em canais PLC, descrito na seção 3.1, que é utilizado nas simulações.

As Figs. 5.7 e 5.8 mostram a convergência dos equalizadores analisados em função dos valores MSE, obtidos na saída dos mesmos, quando $\sigma_x^2 = 0$ dB, $\sigma_{imp}^2 = 0$ dB e $\sigma_{bkg}^2 = -20$ dB. Conforme é observado, os equalizadores fuzzy *nonsingleton* apresentam velocidade de convergência maior do que as suas versões fuzzy *singleton*. Tal ganho de desempenho já era esperado, pois a saída do canal PLC é corrompida por ruído e os sistemas fuzzy *nonsingleton* foram desenvolvidos para lidar com tal problema (vide (Mouzouris & Mendel, 1997)).

Dentre os métodos de otimização analisados, o método LLE foi o que apresentou melhor desempenho em termos de velocidade convergência. Esse resultado é bastante interessante, posto que o método LLE requer o ajuste heurístico de poucos parâmetros. Já o método DRD apresenta uma inicialização mais difícil, pois são

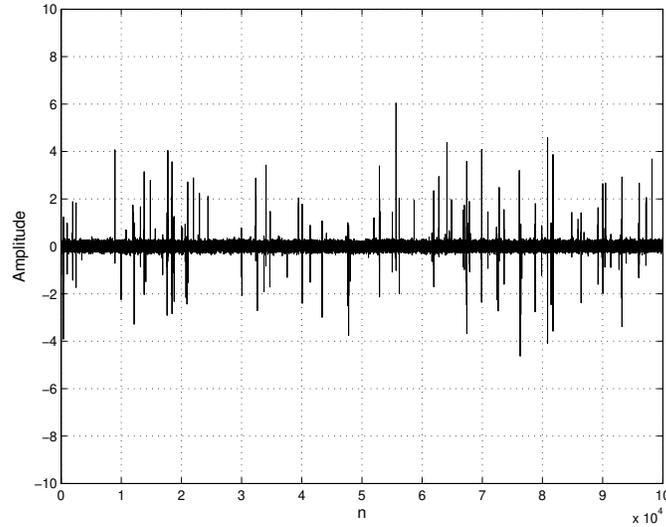


Figura 5.6: Exemplo do ruído aditivo em canais PLC quando $\sigma_x^2 = 0$ dB, $\sigma_{imp}^2 = 0$ dB, $\sigma_{bkg}^2 = -20$ dB e número de amostra igual a $N = 10^5$.

vários os parâmetros a serem inicializados de forma heurística. Tal característica do método DRD faz com que o uso do mesmo seja bastante complicado.

É importante salientar que, embora a complexidade computacional por iteração seja maior quando os métodos DRD e LLE são utilizados para a atualização dos parâmetros dos equalizadores fuzzy, vide Tab. 2.1, o ganho, em termos de velocidade de convergência, é representativo quando o método LLE é utilizado. Isso se deve ao fato de que os equalizadores treinados com tal técnica permitem que a convergência seja alcançada com menor número de iterações.

De forma a mostrar o desempenho dos equalizadores fuzzy do tipo 1 listados acima, as Figs. 5.9 e 5.10 ilustram os desempenhos dos equalizadores fuzzy-S-DFE-LLE, fuzzy-NS-DFE-LLE, DFE e Bayesiano (Chen et al., 1993; Duda & Hart, 1973) com e sem erro de propagação. A relação sinal ruído (*signal to noise ratio* - SNR)

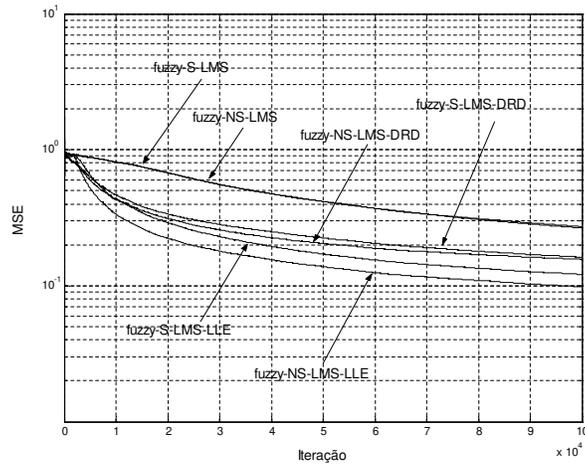


Figura 5.7: Curva de convergência dos equalizadores fuzzy do tipo 1 com estrutura transversal.

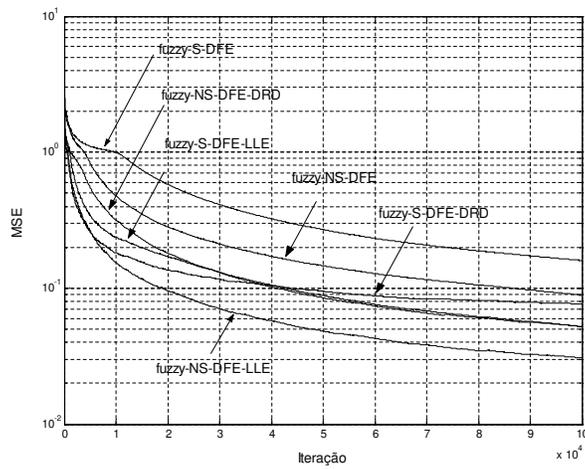


Figura 5.8: Curva de convergência dos equalizadores fuzzy do tipo 1 com estrutura DFE.

mostrada nos gráficos é dada por

$$SNR(dB) = 10 \log_{10} \left(\frac{\sigma_y^2}{\sigma_{imp}^2 + \sigma_{bkgr}^2} \right). \quad (5.42)$$

O número de iterações utilizado nas simulações variaram de 2×10^6 à 10^7 .

É importante ressaltar que apenas os resultados obtidos com os equalizadores fuzzy-DFE-S-LLE e fuzzy-DFE-NS-LLE foram considerados, pois os mesmos apresentam maior velocidade de convergência e seus respectivos competidores convergem para o mesmo mínimo local em iguais condições de simulação. A diferença entre eles reside exatamente no fato de que o método LLE permite que o equalizador fuzzy convirja mais rapidamente.

Sabe-se que em situações em que os estados na saída do canal e a distribuição de probabilidade do ruído aditivo é conhecida, os equalizadores fuzzy do tipo 1 apresentam desempenhos similares àqueles obtidos com os equalizadores Bayesianos. No entanto, conforme é observado nas Figs. 5.9 e 5.10, tais equalizadores apresentam desempenhos levemente inferiores àqueles obtidos com os equalizadores DFE. Tais resultados devem-se ao fato de que os equalizadores fuzzy analisados são soluções sub-ótimas, pois o número de regras utilizado no mesmo é menor que o número necessário para garantir que tais equalizadores considerem todos os vetores de estados na saída do canal PLC em questão.

5.3.2 Resultados com os Algoritmos Fuzzy do Tipo 2

Nesta seção o desempenho dos equalizadores fuzzy do tipo 1 e do tipo 2 é comparado. As considerações de simulações aqui utilizadas para a obtenção dos resultados são as mesmas da seção anterior.

Embora, a utilização dos sistemas fuzzy do tipo 2 seja apropriada para a solução

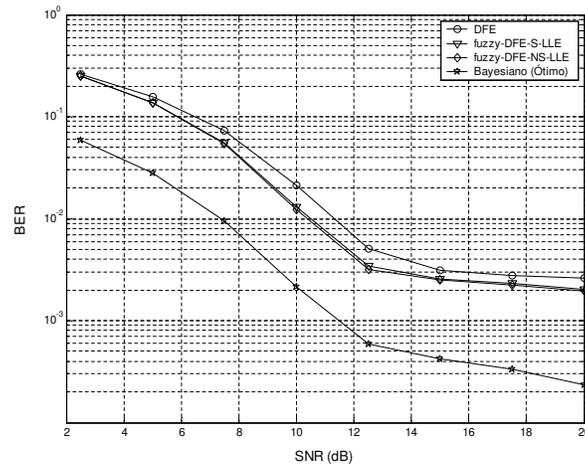


Figura 5.9: Desempenho dos equalizadores fuzzy do tipo 1 com a estrutura do DFE quando existe erro de propagação no mesmo.

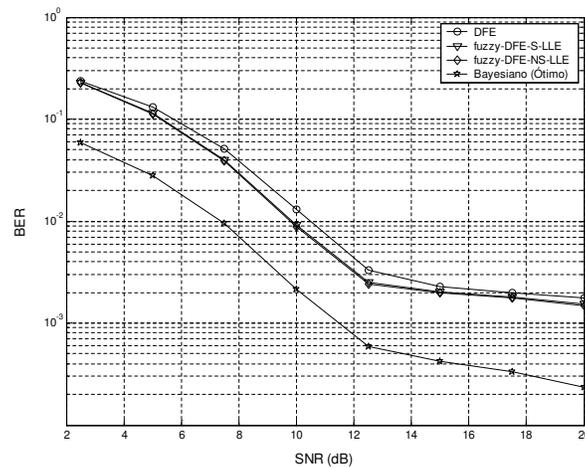


Figura 5.10: Desempenho dos equalizadores fuzzy do tipo 1 com a estrutura DFE quando não existe erro de propagação no mesmo.

de problemas que geram regras fuzzy com incertezas, os resultados mostrados na Fig. 5.11 evidenciam que tais equalizadores oferecem algumas vantagens quando aplicados aos problemas que não produzem incertezas nas regras fuzzy. Na Fig. 5.11, os equalizadores fuzzy2-S-DFE e fuzzy2-NS-DFE referem-se aos equalizadores fuzzy do tipo 2, intervalado, *singleton* e *nonsingleton* implementados dentro da estrutura DFE. Note que o equalizador fuzzy2-S-DFE é uma versão simplificada do equalizador fuzzy2-NS-DFE, quando as incertezas associadas às entradas dos equalizadores são iguais a zero. Os equalizadores fuzzy-S-DFE e fuzzy-NS-DFE são equalizadores fuzzy do tipo 1, *singleton* ou *nonsingleton*.

O método de otimização utilizado para a atualização dos parâmetros destes equalizadores fuzzy do tipo 1 e 2 é o BP. Objetivando garantir condições iguais de simulação desses equalizadores, os seguintes critérios são adotados na inicialização de seus parâmetros: *i*) os parâmetros dos equalizadores fuzzy do tipo 2, intervalado e *singleton* foram obtidos a partir de perturbações aplicadas nos parâmetros dos equalizadores fuzzy do tipo 1 e *ii*) os parâmetros dos equalizadores fuzzy do tipo 2, intervalado e *nonsingleton*, são os mesmos usados no equalizador fuzzy do tipo 2 *singleton*. Os parâmetros que modelam as incertezas nas entradas destes equalizadores fuzzy do tipo 2 foram obtidos a partir de uma perturbação gerada no parâmetro equivalente utilizado no equalizador fuzzy do tipo 1 e *nonsingleton*. O mesmo passo de adaptação $\alpha = 0,001$ foi empregado para a atualização dos parâmetros dos algoritmos fuzzy analisados, objetivando realizar uma comparação não polarizada entre as diferentes os equalizadores fuzzy do tipo 1 e do tipo 2 analisados numericamente nesta seção.

Conforme pode ser observado na Fig. 5.11-5.13, os equalizadores fuzzy do tipo 2 e intervalados, propostos na seção 2.3.1, apresentam maiores velocidade de convergência do que suas versões fuzzy do tipo 1. Nestas figuras, os termos fuzzy-S-DFE e fuzzy-NS-DFE referem-se aos algoritmos fuzzy analisados na seção

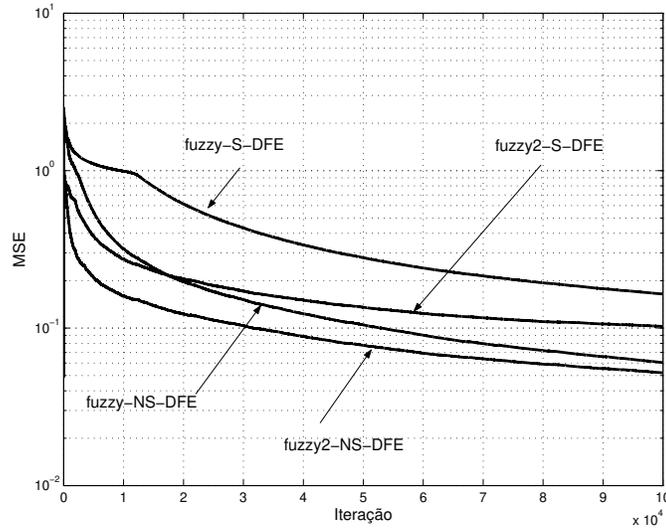


Figura 5.11: Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2.

anterior e os termos fuzzy2-S-DFE e fuzzy2-NS-DFE são os algoritmos fuzzy do tipo 2 deduzidos na seção 2.3.1.

É interessante observar que tal resultado é possível, pois, apesar de não haver incertezas no canal PLC, o número maior de parâmetros oferecido pelo equalizador fuzzy do tipo 2 contribui de forma decisiva para a minimização do MSE. Além disso, pode-se conjecturar as seguintes justificativas para este desempenho dos equalizadores fuzzy do tipo 2 e intervalado para estas simulações: *i*) o ruído aditivo na saída dos canais PLC apresenta componentes impulsivas de alta potência e não são gaussianos, vide Fig. 3.3 e, apesar dos equalizadores fuzzy *nonsingleton* analisados modelarem as incertezas nas entradas do mesmo com a função Gaussiana, as regras fuzzy permitem um grau maior de liberdade para lidar com tais componentes impulsivas; *ii*) o número de regras utilizadas para a equalização deste canal PLC é pequeno e, conseqüentemente, o problema de saturação dos equalizadores fuzzy do tipo 2 não existe.

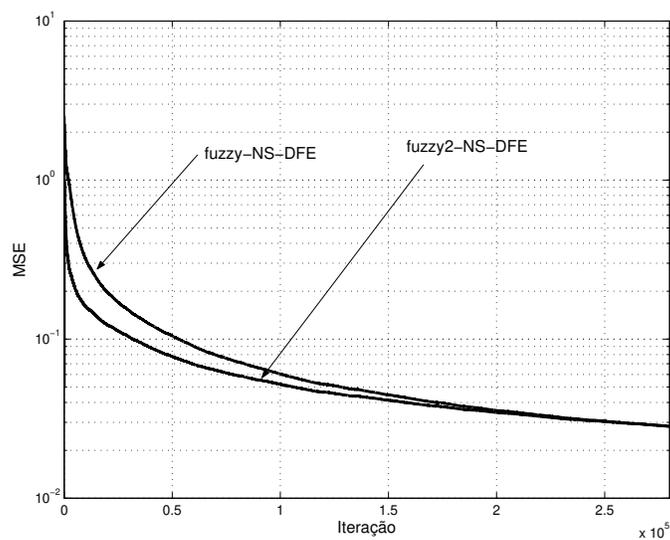


Figura 5.12: Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2 *nonsingleton*.

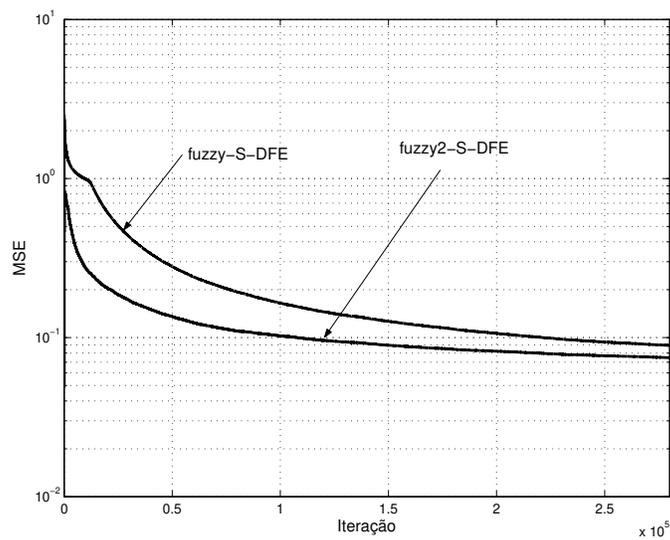


Figura 5.13: Convergência dos equalizadores fuzzy DFE do tipo 1 e do tipo 2 *singleton*.

Na seção 5.2 foi mostrado que se não existem incertezas nas regras fuzzy, os equalizadores fuzzy do tipo 2 e intervalado reduzem-se ao equalizador fuzzy do tipo 1. Tal resultado é bastante interessante, pois significa que na ausência de incertezas no canal PLC (ausência de incertezas nas funções de pertinência que modelam as premissas dos antecedentes das regras fuzzy), os equalizadores fuzzy do tipo 2 convergem para o resultado obtido com os equalizadores fuzzy do tipo 1. Tais resultados são, claramente, tipificados nas Figs. 5.12 e 5.13. Nessas figuras, observa-se que as curvas mostradas são as mesmas ilustradas na Fig. 5.11, com a consideração de que o número de iterações é igual a $2,8 \times 10^6$.

5.4 Sumário

O presente capítulo discutiu a aplicação dos algoritmos fuzzy do tipo 2, introduzidos na seção 2.3, para a equalização de canais PLC. Apesar da maior complexidade computacional demandada por esses algoritmos, os resultados numéricos mostram que eles atingem desempenhos superiores àqueles obtidos com o uso de seus antecessores tanto em termos de velocidade de convergência, quanto em termos de desempenho aferido através da BER.

6

Monitoramento da Qualidade de Energia Elétrica

Nos capítulos anteriores discutiu-se o uso das redes elétricas para a transmissão de dados e apresentou-se algumas técnicas para melhorar a confiabilidade dos sistemas de comunicação que fazem uso das redes elétricas como meio de transmissão de dados. Além disso, comentou-se que a tecnologia PLC é uma solução bastante promissora para as empresas de transmissão e distribuição de energia, uma vez que as mesmas podem fazer uso de tal tecnologia para aplicações de supervisão, gerenciamento e controle dos sistemas de potência e de transmissão banda larga de

dados.

Sabe-se que para as aplicações de supervisão, gerenciamento e controle, o monitoramento da QE exerce um papel fundamental, pois o mesmo fornece uma análise qualitativa e quantitativa da energia transmitida e distribuída pelas empresas de transmissão e distribuição de energia.

Visando propor algumas técnicas que auxiliem a análise da QE, a partir deste capítulo discute-se o uso de técnicas de processamento de sinais e de inteligência computacional para a realização de algumas tarefas de monitoramento da QE. Com esse objetivo, a seção 6.1 discute algumas motivações para o monitoramento da QE e a seção 6.2 descreve o diagrama de blocos da solução constituída por alguns algoritmos baseados em técnicas de processamento de sinais e inteligência computacional a serem empregados para analisar a ocorrência de distúrbios e comenta os dois conceitos que inspiraram o desenvolvimento destes algoritmos.

6.1 Motivação para o Monitoramento da QE

A desregulamentação dos setores de geração, transmissão e distribuição de energia elétrica promovida no mundo, a partir do final do século passado, introduziu os fatores competição e qualidade no mercado de energia elétrica. Nesse novo cenário, diferentes empresas, reunidas em seis diferentes grupos: geração, produção independente, transmissão, distribuição, operadores independentes e negociadores de energia (Abdel-Galil, El-Saadany, & Salama, 2001; Loehr, Casazza, & John, 1999; Arrillaga, Bollen, & Watson, 2000) regulam e competem entre si por uma fatia do mercado de energia. As empresas, movidas pela necessidade de atendimento e superação das expectativas dos consumidores industriais, comerciais e residenciais e pela competitividade realizam investimentos significativos, visando a expansão e a

modernização dos sistemas de potência.

Atualmente, alguns segmentos industriais e de serviços necessitam do fornecimento ininterrupto de energia, ou seja, durante 24 horas por dia e 7 dias por semana. Pode-se citar, como exemplo, o setor financeiro, as indústrias automatizadas e as operadoras de redes de comunicação.

Tal fornecimento é melhor planejado e desenvolvido quando um completo sistema de monitoramento da QE é realizado continuamente em tempo real (Khan, 2001; McGrannaghan, 2001; Abdel-Galil et al., 2001; Arrillaga et al., 2000). Tal medida possibilita a obtenção de informações suficientes para as tomadas de decisões de caráter preditivo, preventivo, corretivo e estratégico que garantam a qualidade e a confiabilidade do sistema de energia. Desta forma, o uso de um sistema completo de monitoramento em tempo real, de transmissão de dados para uma central de processamento de informações e, ainda, de análise destes dados é uma das questões em aberto e mais importantes na concepção atual de um sistema de monitoramento integrado da qualidade da energia elétrica (Khan, 2001; McGrannaghan, 2001; Abdel-Galil et al., 2001).

Para o monitoramento da QE, o desenvolvimento de sofisticadas técnicas de processamento de sinais e de inteligência computacional têm sido investigado, com destaque para as seguintes técnicas: *i*) técnica para a análise espectral do sinal monitorado (fundamental, harmônicas e inter-harmônicas), *ii*) técnica para a detecção da ocorrência de distúrbios, *iii*) técnica para a classificação dos distúrbios, *iv*) técnica para a compressão dos distúrbios para fins de armazenamento e/ou transmissão via canais de comunicação, *v*) técnica para a identificação das fontes geradoras de distúrbios no sistema de potência, *vi*) técnica para a localização das fontes geradoras de distúrbios, *vii*) técnicas para a alocação ótima de equipamento de monitoramento no sistema de potência, *viii*) técnica para a extração de índices

representativos para avaliar o nível de qualidade da energia elétrica e, finalmente, *ix*) técnicas para a transmissão dos dados de monitoramento para uma central de processamento e análise.

Para implementar tais técnicas, duas abordagens para o monitoramento da QE são empregadas: *i*) a abordagem descentralizada na qual os algoritmos de análise são implementados nos equipamentos de monitoramento. Neste caso, usa-se um canal de comunicação com baixa taxa de transmissão para a transmissão das informações obtidas pelos equipamentos de monitoramento à central de processamento. Os equipamentos utilizados na abordagem descentralizada apresentam nível elevado de sofisticação, pois implementam todos os algoritmos necessários para a análise dos distúrbios que ocorrem nas redes elétricas; *ii*) na abordagem centralizada, os algoritmos para a análise dos distúrbios são implementados na central de processamento. A abordagem centralizada demanda o uso de canais de comunicação com altas taxas de transmissão, pois a forma de onda de todos os distúrbios detectados pelos equipamentos de monitoramento devem ser enviados para a central de processamento. Além disso, essa abordagem requer o uso de equipamentos de monitoramento simples e uma central com capacidade de processamento elevada.

Como exemplo de sistema de monitoramento descentralizado, pode-se citar o projeto DPQ da EPRI (EPRI, 2002), cuja estrutura pode ser visualizada na Fig. 6.1. O software PQview abrange todos os processos de aquisição, extração de parâmetros e informações, classificação e reconhecimento, interpretação das informações, apresentação de relatórios e armazenamento das informações.

Baseando-se no projeto DPQ e em idéias e concepções discutidas em (McGrannaghan, 2001; Bollen, 2000; Abdel-Galil et al., 2001; Arrillaga et al., 2000), a Fig. 6.2 mostra uma solução descentralizada para o problema do monitoramento da qualidade da energia elétrica (Khan, 2001). Conforme pode ser observado, as

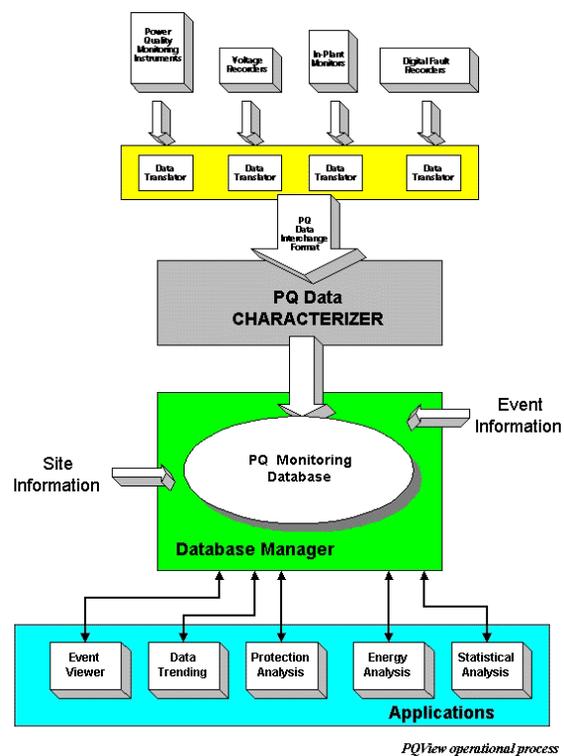


Figura 6.1: Estrutura em bloco do software PQView desenvolvido pela EPRI (EPRI, 2002).

subestações e alguns clientes têm sistemas de monitoramento conectados por meio de um canal de comunicação (canal PLC, canal telefônico, canal sem fio ou canal óptico) à central de processamento para a análise e o armazenamento dos dados, visando prover informações para a análise do comportamento das redes elétricas monitoradas e a tomada de decisões. Neste esquema descentralizado de monitoramento da QE, os canais PLC trabalham conjuntamente com outros meios de comunicação, uma vez que é sabido que dependendo das características de uma comunicação ponto-à-ponto, diferentes tecnologias apresentam diferentes desempenhos. Note ainda que os sistemas de transmissão de dados via rede elétrica podem ser utilizado para a transmissão de dados nas redes de baixa e média tensões.

A estrutura básica de um sistema completo de monitoramento pode ser dividida em três partes principais: *i*) os equipamentos de monitoramento com capacidade de processamento suficiente para analisar os distúrbios monitorados e interface de comunicação, *ii*) o canal de comunicação disponível e *iii*) a central de processamento para reunir os dados dos equipamentos de monitoramento e realizar a análise das redes elétricas monitoradas.

6.2 Solução Proposta para o Monitoramento da Qualidade da Energia

Objetivando aplicar técnicas de processamento de sinais que possam contribuir para o monitoramento da QE, a Fig. 6.3 apresenta o diagrama de blocos de um conjunto de técnicas desenvolvidas que realizam as seguintes tarefas: a estimação de parâmetros espectrais, a detecção, a compressão e a classificação de distúrbios em rede elétrica. A reunião e o uso combinado destas técnicas de processamento de sinais é denominado (*Method for estimation, detection, compression, and*

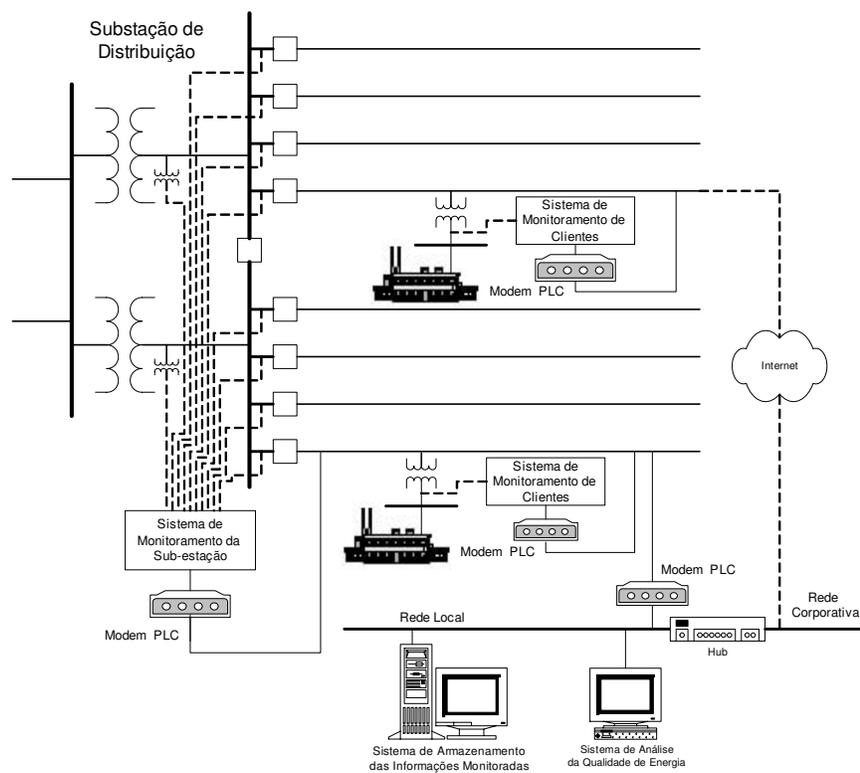


Figura 6.2: Estrutura básica de um sistema completo de monitoramento da qualidade da energia elétrica.

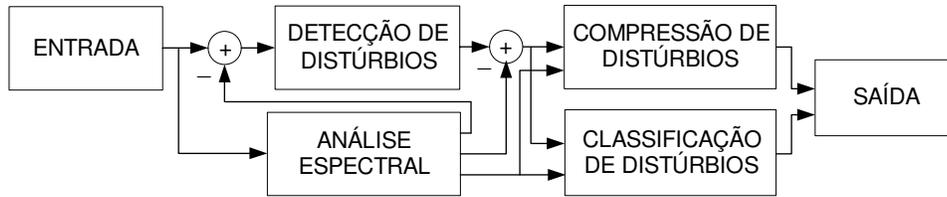


Figura 6.3: Diagrama de blocos do MEDCC.

classification - MEDCC). A proposta MEDCC, quando inserida em um equipamento de monitoramento, permite o processamento descentralizado dos distúrbios em redes elétricas, pois os equipamentos instalados realizam as tarefas de processamento intensivo dos distúrbios, enquanto as unidades centrais de análise apenas avaliam as informações recebidas dos equipamentos de monitoramento, os quais foram instalados em pontos estratégicos.

O MEDCC é dividido em 6 blocos principais que executam tarefas diferentes, as quais são requeridas pelos modernos equipamentos de monitoramento. A descrição da funcionalidade de cada bloco é dada a seguir.

O Bloco de Entrada armazena o quadro de sinais a ser submetido às técnicas de detecção, de classificação e de compressão de distúrbios e para a estimação de parâmetros espectrais;

O Bloco de Detecção de Distúrbios implementa um novo algoritmo baseado em técnicas de processamento de sinais e inteligência computacional para a detecção de distúrbios e a estimação dos instantes iniciais e finais de ocorrência dos mesmos. O uso dessa técnica é pertinente, pois possibilita o uso racional dos algoritmos de compressão e de classificação de distúrbios. A técnica de detecção proposta é discutida no Capítulo 8;

O Bloco de Análise Espectral faz uso de uma nova técnica para a estimação

de parâmetros senoidais, chamada Notch filter - warped DFT (NF-WDFT), desenvolvida especificamente para a análise espectral dos sinais em sistemas de potência (estimação dos parâmetros das componentes fundamental e harmônicas). Tal técnica é discutida no Capítulo 7;

O Bloco Compressão de Distúrbios aplica uma nova técnica de codificação de fonte baseada em transformada para a compressão dos distúrbios monitorados. O Capítulo 10 apresenta duas técnicas propostas para esse fim: *i) enhanced data compression method* (EDCM), discutida na seção 10.1, e *ii) fundamental, harmonic, and transient separation compression method* (FHTCM), apresentada na seção 10.2;

O Bloco de Classificação de Distúrbios refere-se à implementação de uma nova técnica para a classificação de distúrbios em sistemas de potência. A técnica proposta para a classificação de distúrbios é apresentada no Capítulo 9;

O Bloco de Saída armazena a forma de onda do sinal codificado, os parâmetros estimados, os índices de qualidade obtidos a partir das técnicas de classificação, de detecção e de análise espectral (parâmetros das componentes fundamental e harmônicas, tipos de distúrbios, forma de onda codificada dos distúrbios, instantes inicial e final de ocorrência dos distúrbios).

Para introduzir as técnicas de análise espectral, de detecção, de classificação e de compressão, as quais são discutidas nos Capítulos 7, 8, 9 e 10, as seguintes definições são necessárias.

Os sinais obtidos a partir dos sistemas de aquisição de dados (circuito de condicionamento e conversor analógico digital, *Analog to Digital Converter - ADC*) instalados em linhas de transmissão ou de distribuição de energia para fins de monitoramento são dados por

$$x(n) = x(t)|_{t=nT_s} := f(n) + h(n) + i(n) + t(n) + v(n), \quad n = 0, \dots, N - 1, \quad (6.1)$$

onde T_s é o período de amostragem e os sinais $\{f(n)\}$, $\{h(n)\}$, $\{i(n)\}$, $\{t(n)\}$ e $\{v(n)\}$ representam, respectivamente, as componentes fundamental e harmônicas, as componentes inter-harmônicas, os transitórios e o ruído de fundo. Cada uma dessas componentes é definida por

$$f(n) := A_0(n) \cos(2\pi f_0(n)n + \theta_0(n)), \quad (6.2)$$

$$h(n) := \sum_{m=1}^M h_m(n), \quad (6.3)$$

$$i(n) := \sum_{j=1}^J i_j(n), \quad (6.4)$$

$$t(n) := \sum_{k=1}^K t_k(n), \quad (6.5)$$

e a seqüência $\{v(n)\}_{n=0}^{N-1}$ é constituída de variáveis aleatórias independentes e distribuídas igualmente (*i.i.d.*) com distribuição normal $\mathcal{N}(0, \sigma_v^2)$ e independentes das seqüências $\{f(n)\}_{n=0}^{N-1}$, $\{h(n)\}_{n=0}^{N-1}$, $\{i(n)\}_{n=0}^{N-1}$ e $\{t(n)\}_{n=0}^{N-1}$.

Em (6.2), $A_0(n)$, $f_0(n)$ e $\theta_0(n)$ são as amostras da amplitude, da freqüência e da fase da componente fundamental. Em (6.3) e (6.4), $h_m(n)$ e $i_j(n)$ são, respectivamente, a m -ésima componente harmônica e a j -ésima componente inter-harmônica, as quais são definidas por

$$h_m(n) := A_m(n) \cos(2\pi m f_0(n)n + \theta_m(n)), \quad (6.6)$$

e

$$i_j(n) := A_{I,j}(n) \cos(2\pi f_{I,j}(n)n + \theta_{I,j}(n)). \quad (6.7)$$

Em (6.6), $A_m(n)$ e $\theta_m(n)$ definem a amplitude e a fase da m -ésima componente harmônica e em (6.7), $A_{I,j}(n)$, $f_{I,j}(n)$ e $\theta_{I,j}(n)$ representam a amplitude, a

frequência e a fase da j -ésima componente inter-harmônica. O sinal $\{t_k(n)\}$, em (6.5), representa o k -ésimo transitório que é constituído das componentes inter-harmônicas, spikes, notches, etc.

Note que as definições dadas por (6.1)-(6.7) são representativas do sinal monitorado em linhas de transmissão e/ou de distribuição de energia, uma vez que considera os parâmetros dos sinais variantes no tempo. A variabilidade de cada parâmetro em relação ao tempo depende das características topológicas da rede de distribuição e das cargas conectadas ao sistema de distribuição em análise. De forma geral, é sabido que $\frac{d}{dt}\phi_m(t) = \frac{d}{dt}(\omega_m(t)t + \theta_m(t))$ é bastante baixa e $\frac{d}{dt}A_m(t)$, $\frac{d}{dt}\theta_0(t)$, $\frac{d}{dt}\phi_{I,j}(t) = \frac{d}{dt}(\omega_{I,j}(t)t + \theta_{I,j})$, $\frac{d}{dt}A_{I,j}(t)$ e $\frac{d}{dt}\theta_{I,j}(t)$ variam de acordo com o tipo de distúrbio observado na rede elétrica.

A modelagem dada por (6.1)-(6.7) evidencia o sinal monitorado como o somatório de várias contribuições distintas associadas a um grande número de parâmetros. Dessa forma, a análise direta do lado direito de (6.1) representa um problema com grande complexidade, já que tal sinal é constituído pela mistura de vários sinais. Visto que a análise de sinais ou sistemas complexos, normalmente, demandam o uso de soluções complexas, nessa seção propõe-se a utilização de duas idéias conceituais para a análise do sinal $\{x(n)\}$, discutidas a seguir.

Solução hierárquica baseada em dois níveis (Effros, 1998) é uma modelagem para a solução de problemas complexos. Este modelo faz uso do primeiro nível para agrupar as classes de sinais ou sistemas com característica comuns e, a seguir, o segundo nível é utilizado para analisar separadamente os grupos de sinais ou sistemas com características comuns. A Fig. 6.4 mostra, graficamente, o conceito proposto para o problema de classificação. O nível 1 realiza a classificação ou o agrupamento dos sistemas ou dos sinais em vários grupos e, na seqüência, a classificação do sinal é realizada no nível 2.

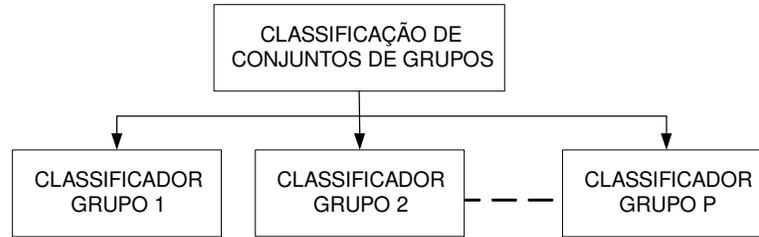


Figura 6.4: Diagrama de blocos da solução hierárquica baseada em dois níveis.

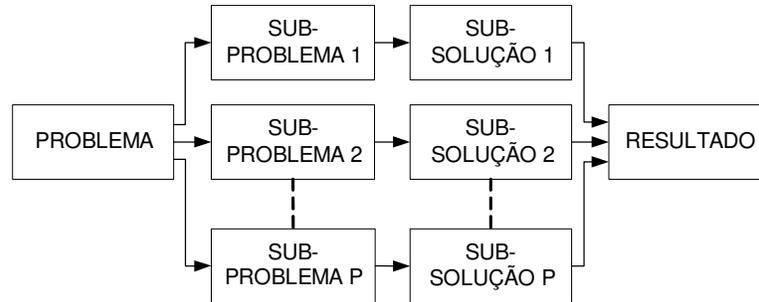


Figura 6.5: Diagrama de blocos do princípio dividir e conquistar.

Princípio dividir e conquistar (Haykin & Li, 1995) postula que para resolver um problema complexo, basta dividir o mesmo em vários problemas simples e projetar soluções para cada um deles. Nesse caso, as soluções para cada um dos sub-problemas é, normalmente, simples. A Fig. 6.5 mostra o diagrama em blocos do princípio dividir e conquistar.

6.3 Sumário

Neste capítulo foi discutido as motivações para o monitoramento da QE e foi apresentado uma solução para o monitoramento descentralizado da QE, levando-se em consideração as questões de análise espectral, detecção, classificação e compressão dos distúrbios monitorados. Além disso, foram descritos os princípios, baseados nos

quais fundamentam-se as técnicas propostas nos Capítulos 7, 8, 9 e 10.

7

Análise Espectral dos Distúrbios

O problema de identificação da frequência de sinais estacionários e não-estacionários constituídos de um ou múltiplos sinais senoidais, em diferentes áreas da engenharia, é amplamente investigado. Uma rápida pesquisa bibliográfica sobre tal questão revela inúmeras contribuições (Cohen, 1989; Boashash, 1992a).

Para a análise dos sistemas de potência, não apenas são necessárias as estimativas das frequências da fundamental, das harmônicas e das inter-harmônicas, mas também as estimativas de suas amplitudes e fases. De fato, a obtenção de tais parâmetros permite a análise precisa do comportamento do sistema para fins de

monitoramento, controle, supervisão e etc. No entanto, é sabido que as técnicas atualmente empregadas para esse fim apresentam limitações de desempenho devido as restrições de resolução, complexidade computacional, precisão e necessidade do uso de longas seqüências.

Objetivando minimizar os problemas acima relacionados para a análise das componentes fundamental e harmônicas, nesta seção apresenta-se uma nova técnica para a estimação dos parâmetros senoidais, que explora as características inerentes das componentes fundamental e harmônicas do sinal dado por (6.1).

De forma geral, pode-se dizer que a estimação dos parâmetros das componentes fundamental e das harmônicas do sinal $\{x(n)\}$ é uma tarefa que se torna cada vez mais complexa a medida que é influenciada pela presença das outras componentes harmônicas (Boashash, 1992b), das inter-harmônicas, dos transitórios e do ruído de fundo. Assim sendo, pode-se estabelecer que a estimação da m -ésima componente fundamental harmônica, presente no sinal $\{x(n)\}$, depende da relação dada por

$$SNR_m := 10 \log_{10} \left(\frac{S_m}{\eta_m} \right) \quad (7.1)$$

onde

$$S_m := \frac{\sum_{n=0}^{N-1} (h_m(n))^2}{N}, \quad (7.2)$$

$$\eta_m := \frac{\sum_{n=0}^{N-1} \left(\sum_{\substack{i=0 \\ i \neq m}}^M (h_i(n))^2 + \sum_{j=1}^J (i_j(n))^2 + \sum_{k=1}^K (t_k(n))^2 \right)}{N} + \sigma_v^2 \quad (7.3)$$

e

$$h_m(n) = \begin{cases} f(n), & \text{se } m = 0 \\ h_m(n), & \text{se } m > 0, \end{cases} \quad (7.4)$$

onde $f(n)$ e $h(n)$ são dadas por (6.2) e (6.3), respectivamente

A maioria das técnicas de estimação de parâmetros senoidais, tais como filtros notch, FFT, DFT-WDFT, algoritmo RLS, algoritmo LMS, estimadores baseados na

distribuição tempo-freqüência do sinal, no PLL (*Phase Locked Loop*) e no filtro de Kalman, no critério de máxima verossimilhança (ML) assume (6.1) em sua entrada e, conseqüentemente, observa a SNR dada por (7.2) (Mitra, 2001; Franz, Mitra, & Doblinger, 2003; Boashash, 1992a; Dash, Pradhan, & Panda, 1999; Hsieh & Hung, 1996; Gharieb, 2000; K.-Ghartemani, Karimi, & Iravani, 2004; James, Anderson, & Williamson, 1994; Cohen, 1989; K.-Ghartemani & Iravani, 2003; Liu, 2000; Shayani, 2004). Sabe-se que, das técnicas listadas acima, as técnicas baseadas no critério ML são as soluções ótimas para a estimação de parâmetros senoidais. No entanto, a complexidade computacional demandada pelas técnicas baseadas no critério ML impossibilita o uso das mesmas em aplicações práticas (Boashash, 1992a; Rife & Boorstyn, 1974).

O requerimento mínimo para uma estimativa confiável da amplitude, da freqüência e da fase do m -ésimo sinal senoidal é a *unbiasedness* quando $\eta_m = 0$. Em outras palavras, se a m -ésima componente senoidal não é corrompida por nenhum tipo de ruído e seus parâmetros $A_m(n)$, $\omega_m(n)$ e $\theta_m(n)$ são constantes, então,

$$\hat{A}_m(n) = A_m(n), \quad (7.5)$$

$$\hat{\omega}_m(n) = \omega_m(n), \quad (7.6)$$

e

$$\hat{\theta}_m(n) = \theta_m(n) \quad (7.7)$$

são constantes em condições de estado permanente. O desempenho das técnicas aplicadas para estimar os parâmetros da m -ésima componente senoidal em sistemas de potência é degradado pela ocorrência das outras componentes harmonicamente relacionadas com a mesma, das componentes inter-harmônicas, dos sinais transitórios e do ruído de fundo, conforme explicitado em (7.2)-(7.3). No entanto, o valor SNR dado por (7.1) pode ser melhorado se considera-se uma solução que reduz ou elimina a presença das componentes harmônicas, inter-harmônicas e

dos transitórios. Nessas condições, o novo valor da SNR na entrada do estimador de parâmetros senoidais é dado por

$$SNR_{novo,m} := \frac{\sum_{n=0}^{N-1} (h_m(n))^2}{N\sigma_v^2} \geq SNR_m, \quad m = 0, \dots, M, \quad (7.8)$$

uma vez que apenas o ruído de fundo corrompe o sinal $\{h_m(n)\}$.

Objetivando implementar (7.8) e ao mesmo tempo prover uma solução robusta e confiável para a estimação dos parâmetros senoidais em sistemas de potência, mesmo quando baixas taxas de amostragem são consideradas, neste capítulo propõe-se a técnica *Notch Filter Warped DFT* (NF-WDFT) (Ribeiro, Mitra, & Romano, 2004c). As Figs. 7.1 e 7.2 mostram os diagramas de blocos das versões I e II da técnica NF-WDFT. De um modo geral, pode-se dizer que essa técnica é a combinação das seguintes técnicas: filtros notch não adaptativos (Hirano, Nishimura, & Mitra, 1974), filtros notch adaptativos (Regalia, 1995), baseados em EOS (*Estatísticas de Ordem Superior*) (Gharieb, 2000), operador delta (Feuer & Goodwin, 1996), rede neural MLP treinada (Haykin, 1996) por um algoritmo baseado em informações de 2ª ordem e a *warped DFT* (Makur & Mitra, 2001; Bagchi & Mitra, 1999).

A NF-WDFT faz uso da mesma idéia aplicada em (Rife & Boorstyn, 1974), ou seja, primeiramente garante-se a estimação grosseira da frequência da componente fundamental e, a seguir, aplica-se um processo iterativo para obter uma estimativa mais precisa da frequência da mesma.

A técnica NF-WDFT pode ser vista como a extensão e/ou a generalização da técnica DFT-WDFT (Franz et al., 2003), recentemente desenvolvida para a estimação da frequência de uma única componente senoidal corrompida por ruído aditivo. De fato, a técnica NF-WDFT é capaz de estimar não apenas a frequência de uma componente senoidal, mas também de várias outras e, além disso, fornece estimativas das amplitudes e fases de todas as componentes fundamental e harmônicas.

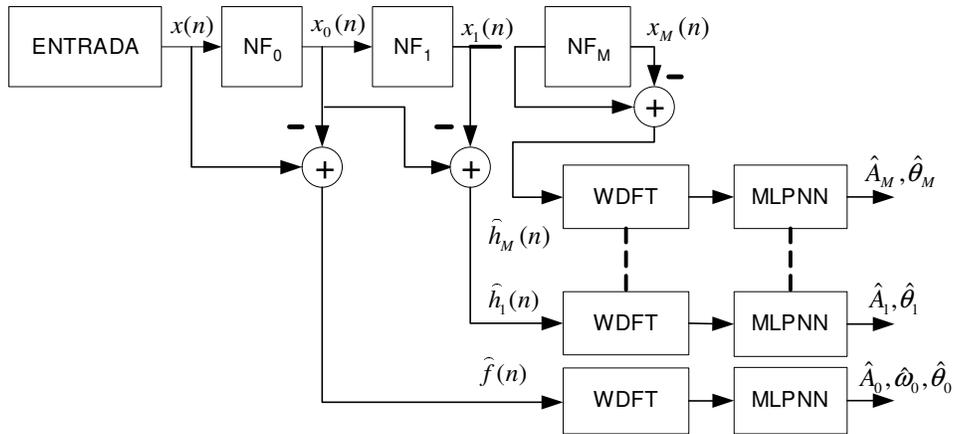


Figura 7.1: Técnica NF-WDFT I.

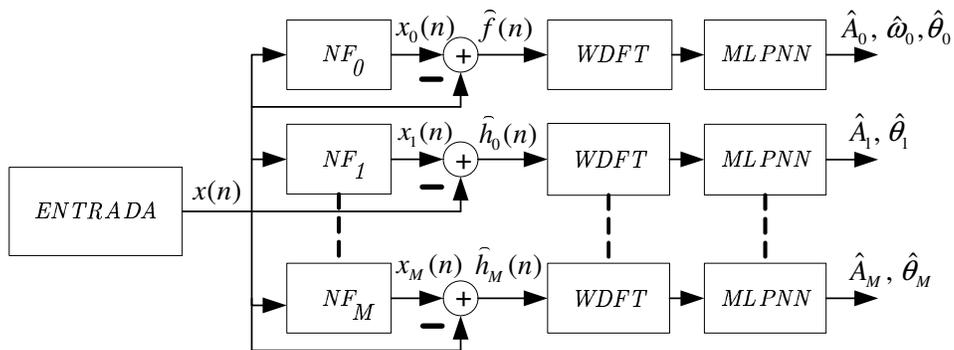


Figura 7.2: Técnica NF-WDFT II.

Conforme pode ser observado nas Figs. 7.1 e 7.2, o i -ésimo bloco NF implementa um filtro notch de 2^a ordem, cuja função de transferência é dada por

$$H_m(z) = \frac{1 + a_m z^{-1} + z^{-2}}{1 + \rho_m a_m z^{-1} + \rho_m^2 z^{-2}}, \quad (7.9)$$

onde

$$a_m = -2 \cos \omega_m \quad (7.10)$$

e $m = 0, \dots, M$. Note que ω_m é a frequência angular dada por

$$\omega_m = \begin{cases} \omega_0 = 2\pi f_0/f_s, & \text{se } m = 0 \\ m\omega_0, & \text{se } m > 0 \end{cases}, \quad (7.11)$$

onde $f_s = 1/T_s$ é a taxa de amostragem e $0 \ll \rho_m < 1$. É importante ressaltar que a implementação com precisão finita do filtro notch pode torna-se instável se ρ_m é próximo de 1 e poucos bits são considerados na quantização dos seus coeficientes. No entanto, esta implementação é necessária, pois M é normalmente grande, podendo chegar a 60, e isso demanda a implementação de muitos filtros notch. De forma a minimizar os efeitos inerentes à implementação em ponto fixo no domínio z (tempo discreto), a seção 7.1 discute o uso do filtro notch no domínio do operador delta (Kauraniemi, Laakso, Hartimo, & Ovaska, 1998; Goodwin, Middleton, & Poor, 1992; Li & Gevers, 1993; Feuer & Goodwin, 1996), já que neste domínio o filtro notch é mais robusto às influências da precisão finita.

Seja o sinal dado por (6.1) na saída do bloco *Entrada* na Fig. 7.1, então, para a versão I da técnica NF-WDFT, tem-se que para $m = 0, \dots, M$

$$x_m(n) \cong \begin{cases} x(n + d_0) - |H_0(e^{j\omega_0(n)})| A_0(n) \cos(n\omega_0(n) + \theta_0(n) + \Delta\theta_0(n)), & \text{se } m = 0 \\ x_{m-1}(n + d_{m-1}) - \sum_{i=0}^m |H_i(e^{j\omega_i(n)})| A_i(n) \cos(n\omega_i(n) + \theta_i(n) + \Delta\theta_i(n)), & \\ & \text{se } m > 0 \end{cases}, \quad (7.12)$$

$$\begin{aligned} \widehat{h}_m(n) &= \begin{cases} x(n + d_m) - x_m(n), & \text{se } m = 0 \\ x_{m-1}(n + d_m) - x_m(n), & \text{se } m > 0 \end{cases}, \\ &= (1 - |H_m(e^{j\omega_m(n)})|) A_m(n) \cos(n\omega_m(n) + \theta_m(n) + \Delta\theta_m(n)) + v_m(n) \end{aligned} \quad (7.13)$$

$$\Delta\theta_m(n) = \sum_{i=0}^m \angle H_i(e^{j\omega_i(n)}) \quad (7.14)$$

e

$$\sigma_{v,m}^2 = \sigma_v^2 \left(1 - \prod_{i=0}^m |H_i(e^{j\omega_i(n)})|^2\right). \quad (7.15)$$

Para a versão II do método NF-WDFFT, tem-se

$$\begin{aligned} \widehat{h}_m(n) &= \begin{cases} \widehat{f}(n), & \text{se } m = 0 \\ \widehat{h}_m(n), & \text{se } m > 0 \end{cases}, \\ &= (1 - |H_m(e^{j\omega_m(n)})|) A_m(n) \cos(n\omega_m(n) + \theta_m(n) + \Delta\theta_m(n)) + v_m(n) \end{aligned} \quad (7.16)$$

$$\Delta\theta_m(n) = \angle H_m(e^{j\omega_m(n)}) \quad (7.17)$$

e

$$\sigma_{m,v}^2(n) = \sigma_v^2 (1 - |H_m(e^{j\omega_m(n)})|^2). \quad (7.18)$$

Nas equações acima, $x_m(n)$ representa a saída do m -ésimo block NF que implementa um filtro notch de 2^a ordem; d_m é o atraso inserido pelo m -ésimo filtro notch; $\widehat{h}_m(n)$ é uma estimativa de (7.4); $v_m(n)$ é o ruído colorido de fundo, cuja potência é dada por $\sigma_{m,v}^2$. Note que $\sigma_{m,v}^2 < \sigma_v^2$ devido às características do filtro notch; $\Delta\theta_m(n)$ é a distorção de fase introduzida pelos filtros notch implementados

nos blocos $NF_{0,\dots,M}$, se a técnica NF-WDFT versão I é considerada, vide (7.14), ou pelo m -ésimo filtro notch quando se considera a técnica NF-WDFT tipo II, vide (7.17). O sinal $\widehat{h}_m(n)$ apresenta uma SNR dada por

$$\begin{aligned} SNR_{novo,m} &= \\ & \frac{1}{N\sigma_{v,m}^2} \sum_{n=0}^{N-1} [(1 - |H_m(e^{j\omega_m(n)})|) A_m(n) \cos(n\omega_m(n) + \theta_m(n) + \Delta\theta_m(n))]^2 . \\ & \geq SNR_m, \quad m = 0, \dots, M \end{aligned} \quad (7.19)$$

onde

$$H_m(e^{j\omega_m(n)}) = \begin{cases} \prod_{i=0}^m H_i(e^{j\omega_i(n)}), & \text{se versão I,} \\ H_m(e^{j\omega_m(n)}), & \text{se versão II.} \end{cases} \quad (7.20)$$

A equação (7.19) mostra que as entradas dos blocos $WDFT_{0,\dots,M}$ apresentam *SNRs* superiores àquelas obtidas com a expressão (7.1). Conseqüentemente, melhores estimativas dos parâmetros senoidais podem ser obtidas.

Na técnica NF-WDFT, realiza-se primeiro a estimação dos parâmetros da componente fundamental, por causa da relação harmônica explicitada em (6.3). Assim sendo, primeiramente uma estimativa grosseira, $\tilde{a}_0(n)$, do coeficiente $a_0(n)$ de $H_0(z)$ (vide as expressões (7.9) e (7.10)) é realizada usando-se o filtro notch adaptativo baseado em EOS (Gharieb, 2000). A descrição do filtro notch baseado em EOS é dada na seção 7.2. Note que o filtro notch, aqui considerado, pode ser substituído pelo filtro de Kalman, por um PLL ou por um outro estimador que forneça uma estimação inicial da freqüência da componente fundamental. A seguir, o valor $\tilde{a}_0(n)$ é utilizado como argumento do parâmetro α da transformada *warped*-DFT (WDFT), implementada nos blocos $WDFT_0$, detalhes da transformada WDFT são encontrados na seção 7.3. $N/4$ vetores de base da WDFT são aplicados ao sinal (7.13) ou (7.16). A estimativa de ω_0 no domínio da WDFT é descrita a seguir.

Sejam $X[k], k = 1, \dots, N - 2$, os coeficientes no domínio da WDFT e $X[k - 1]$ e $X[k + 1]$ os coeficientes mais próximos de $X[k]$ que apresenta a maior amplitude.

Então os parâmetros da série de Taylor de 2ª ordem, utilizada para aproximar a curva que contém $|X[k-1]|$, $|X[k]|$ e $|X[k+1]|$, é dada por

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \omega_{w,0}^2(k-1) & \omega_{w,0}(k-1) & 1 \\ \omega_{w,0}^2(k) & \omega_{w,0}(k) & 1 \\ \omega_{w,0}^2(k+1) & \omega_{w,0}(k+1) & 1 \end{bmatrix}^{-1} \begin{bmatrix} |X[k-1]| \\ |X[k]| \\ |X[k+1]| \end{bmatrix}, \quad (7.21)$$

onde $\omega_{w,0}(k-1)$, $\omega_{w,0}(k)$ e $\omega_{w,0}(k+1)$ são os valores das frequências angulares no domínio da WDFT correspondentes as amostras $X[k-1]$, $X[k]$ e $X[k+1]$, respectivamente. As estimativas da frequência angular $\hat{\omega}_{w,0}$ e da amplitude $\hat{A}_{w,0}$ do sinal $\{f(n)\}$, no domínio da WDFT, podem ser dadas por

$$\hat{\omega}_{w,0} = -\frac{b}{2a} \quad (7.22)$$

e

$$\hat{A}_{w,0} = \frac{-b^2 + 4ac}{2a} \quad (7.23)$$

respectivamente. Já a fase $\hat{\theta}_{w,0}$ do sinal $\{f(n)\}$ no domínio da WDFT é dada por uma aproximação linear entre as fases dos coeficientes $X[k-1]$ e $X[k]$ quando $\hat{\omega}_{w,0} \in [\omega_{w,0}(k-1), \omega_{w,0}(k)]$ ou entre as fases dos coeficientes $X[k]$ e $X[k+1]$ quando $\hat{\omega}_{w,0} \in [\omega_{w,0}(k), \omega_{w,0}(k+1)]$

Sabe-se que os valores $\hat{A}_{w,0}$, $\hat{\omega}_{w,0}$ e $\hat{\theta}_{w,0}$ não correspondem aos valores reais dos parâmetros do sinal $\{f(n)\}$ (Franz et al., 2003). Objetivando mostrar esses comportamentos, as Figs. 7.3-7.5 ilustram as relações não lineares entre $\hat{A}_{w,0}$ e A_0 , $\hat{\theta}_{w,0}$ e θ_0 , e $\hat{\omega}_{w,0}$ e ω_0 . Os valores mostrados nessas figuras foram normalizados para facilitar a visualização.

A Fig. 7.3 foi obtida a partir do sinal

$$x(n) = \cos(\omega_0 n + \theta_0), \quad (7.24)$$

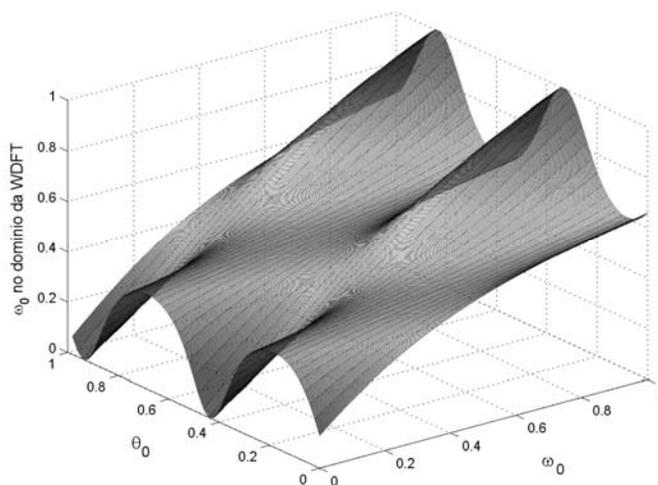


Figura 7.3: Relação entre os valores da frequência do sinal $\{f(n)\}$ nos domínios da DFT e WDFT.

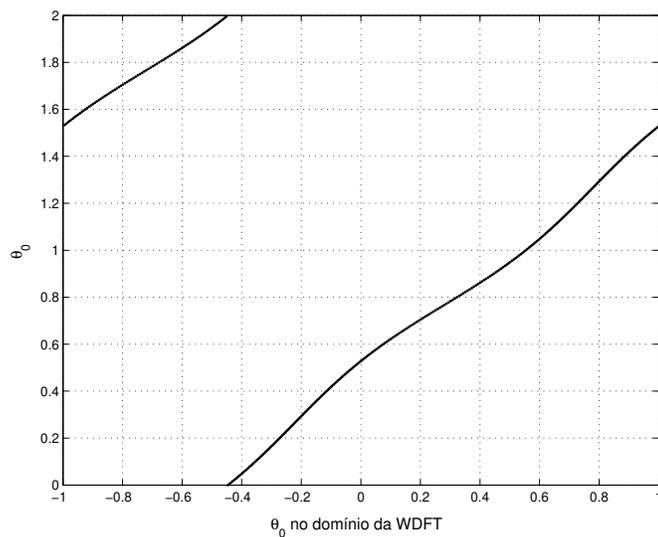


Figura 7.4: Relação entre os valores da fase do sinal $\{f(n)\}$ nos domínios da DFT e WDFT.

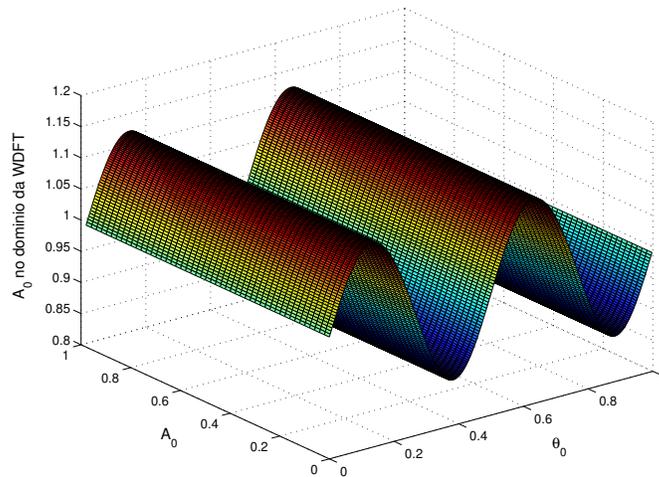


Figura 7.5: Relação entre os valores da amplitude de $\{f(n)\}$ nos domínios da DFT e WDFT.

submetido à WDFT e ao processo de aproximação descrito por (7.21)-(7.23). Para a obtenção da superfície mostrada na Fig. 7.3, foram considerados $f_0 \in [55, 65]$, $f_s = 60 \times 64$ amostras/seg, $\theta_0 \in [0, 2\pi)$ e $N = 64$. A escolha de $f_0 \in [55, 65]$ deve-se ao fato de que a frequência da componente fundamental no Brasil é igual a 60 Hz.

A curva mostrada na Fig. 7.4 explicita a relação entre as fases da componente senoidal nos domínios da WDFT e da DFT. Ela foi obtida levando-se em conta as seguintes considerações: $f_s = 60 \times 64$ amostras/seg, $\theta_0 \in [0, 2\pi)$ e $N = 64$. Note que na Fig. 7.4 $\hat{\theta}_{w,0} \in [-\pi, \pi)$.

A superfície mostrada na Fig. 7.5 ilustra a relação entre A_0 e $\hat{A}_{w,0}$ quando $\theta_0 \in [0, 2\pi)$. Esse gráfico mostra os valores normalizados das amplitudes para $A_0 \in (0, 1]$. Para a obtenção desses resultados $f_s = 60 \times 64$ Hz, $\theta \in [0, 2\pi)$ e $N = 64$.

As relações mostradas nas Figs. 7.3-7.5 podem ser representadas por

$$\hat{\omega}_0 := g_\omega(\hat{\omega}_{w,0}, \hat{\theta}_0), \quad (7.25)$$

$$\hat{\theta}_0 := g_\theta(\hat{\theta}_{w,0}) + \Delta\theta_0, \quad (7.26)$$

e

$$\hat{A}_0 := g_A(\hat{A}_{w,0}, \hat{\theta}_0). \quad (7.27)$$

Onde $g_\omega(\cdot)$, $g_\theta(\cdot)$ e $g_A(\cdot)$ são relações que mapeiam os valores das frequências angulares, fases e amplitudes entre os domínios da WDFT e DFT.

Substituindo-se o índice 0 por m , $m = 1, \dots, M$ em (7.26) e (7.27), obtém-se as estimativas das amplitudes e das fases das componentes harmônicas. Note que a estimativa da frequência angular da m -ésima componente harmônica é dada por $m\hat{\omega}_0$.

Em nosso trabalho, a rede MLP foi utilizada para implementar as relações g_A , g_ω e g_θ . Durante as simulações e testes, observou-se que o desempenho da rede MLP apresentou resultados satisfatórios quando a mesma foi aplicada para a estimação da frequência e da amplitude. No entanto, resultados significativos foram obtidos quando a rede MLP foi aplicada para a estimação da fase do sinal senoidal. De fato, as superfícies mostradas nas Figs. 7.3 e 7.5 são bastantes complexas, o que de certa forma degrada o desempenho da rede MLP. Objetivando superar tal problema, considerou-se a partição do intervalo $[0, 2\pi)$ em vários subintervalos de tal forma que várias redes MLP são utilizadas. O número de subintervalos de $[0, 2\pi)$, aqui sugerido, é $Q = 360 \times 2^\kappa$, $\kappa = 1, 2, \dots$. Para o q -ésimo intervalo, duas redes MLP são utilizadas para estimar a amplitude e a fase da senóide analisada. O uso desta estratégia se justifica, pois é possível o uso da rede MLP com poucos neurônios para realizar o mapeamento em cada subintervalo, apesar do considerável espaço

de memória necessária para armazenar os pesos de todas as redes neurais quando $\kappa \gg 1$.

A Fig. 7.6 mostra o desempenho de uma rede MLP com uma camada escondida para realizar o mapeamento entre os valores das amplitudes, das fases e das frequências nos domínios da WDFFT e da DFT. As seguintes considerações foram utilizadas nas simulações: $\kappa = 2$, os números de neurônios utilizados nas camadas intermediárias das redes MLP para a estimação dos parâmetros A_0 , ω_0 e θ_0 , a partir dos valores $\hat{A}_{w,0}$, $\hat{\omega}_{w,0}$ e $\hat{\theta}_{w,0}$, são iguais a 2, 4 e 6, respectivamente. Conforme pode ser observado na Fig. 7.6, os valores MSE, obtidos com os dados de teste, são menores que -55 dB, -60 dB e -70 dB para a estimação de A_0 , θ_0 e ω_0 .

Finalmente, após a estimativa de ω_0 ou f_0 , os parâmetros dos filtros notch implementados nos blocos NF_m , $m = 1, \dots, M$, são determinados e o procedimento descrito acima é empregado para a estimação das fases e das amplitudes das componentes harmônicas.

7.1 Filtro Notch no Domínio do Operador Delta

A minimização dos efeitos inerentes à precisão finita em filtros digitais foi bastante investigada. As principais investigações foram direcionadas para minimizar o erro de *roundoff* na saída dos filtros, quando as estruturas em espaço de estados é usada para implementar os filtros digitais (Hwang, 1977; Mullis & Roberts, 1976; Jackson, Lindgren, & Kim, 1979), mas também para a análise de redes que minimizam a sensibilidade de algoritmos de processamento de sinais (Thiele, 1986) e para o desenvolvimento de métodos que minimizam os efeitos da quantização dos sinais (*residue feedback*, *error spectrum shaping* e *noise shaping*) (Thong & Liu, 1977; Higgins & Jr., 1982; Laakso & Hartimo, 1992).

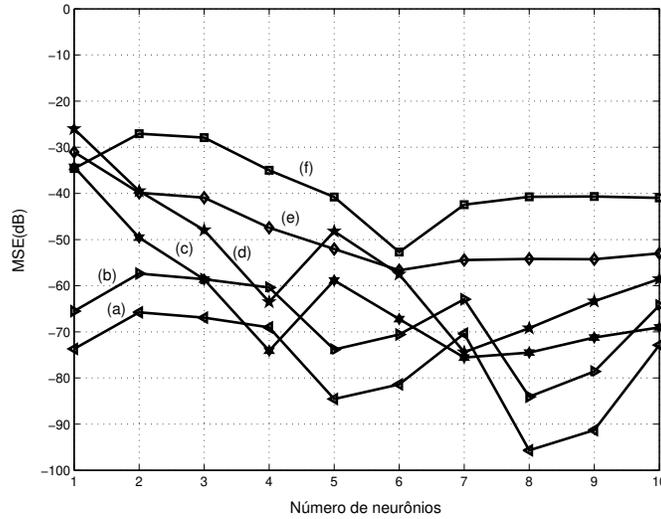


Figura 7.6: Desempenho da rede MLP. (a) e (b) são os resultados de treinamento e teste referentes à A_0 . (c) e (d) são os resultados de treinamento e teste referentes à θ_0 . (e) e (f) são os resultados de treinamento e teste referentes à ω_0 .

Normalmente, a quantização dos coeficientes dos filtros digitais influencia o comportamento do mesmo. Um caso patológico ocorre quando f_s é muito maior que a largura de banda de interesse. Nesse caso, os efeitos da precisão finita se tornam mais significativos (Goodwin et al., 1992). Por exemplo, os pólos de um filtro passa-baixa e banda estreita, onde $0 < \Delta\omega \ll \pi$, são localizados perto do ponto $z = 1$ no plano z , torna o filtro digital bastante sensível aos efeitos da quantização de seus coeficientes e leva o mesmo a ter um comportamento instável. No entanto, conforme já discutido em (Goodwin et al., 1992; Kauraniemi et al., 1998; Li & Gevers, 1993; Goodwin & Middleton, 1986), o projeto de filtros digitais no domínio do operador δ , permite a minimização dos efeitos de quantização e *roundoff* quando esses filtros são passa-baixa ou banda estreita com a restrição de que $2\pi f_s \gg \Delta\omega$, sendo $\Delta\omega$ a banda do filtro digital.

Objetivando discutir o uso do operador δ para a implementação dos filtros notch,

a definição formal do operador δ é dada a seguir.

Seja o domínio de possíveis instantes t dado por

$$\Omega(\Delta) = \begin{cases} \mathfrak{R}, & \Delta = 0 \\ \{\dots, -2\Delta, -\Delta, 0, \Delta, 2\Delta, \dots\}, & \Delta \neq 0, \end{cases} \quad (7.28)$$

onde Δ representa o período de amostragem no tempo discreto. \mathfrak{R} representa o conjunto de números reais. Note que se $\Delta = 0$ temos o tempo contínuo. Como o interesse é normalmente voltado para os valores positivos do tempo contínuo ou discreto, pode-se definir

$$\begin{aligned} \Omega^+(\Delta) &= \{x \in \Omega(\Delta), \quad x \geq 0 \\ &= \begin{cases} \mathfrak{R}^+ \cup \{0\}, & \Delta = 0 \\ \{0, \Delta, 2\Delta, \dots\}, & \Delta \neq 0 \end{cases} . \end{aligned} \quad (7.29)$$

Agora, seja a função $X(t)$ um mapeamento entre $\Omega^+(\Delta)$ e números reais ou complexos, por exemplo $X : \Omega^+(\Delta) \rightarrow \mathbb{C}$, e o operador de deslocamento direto z dado por

$$zX(\cdot) := Y(\cdot) \quad (7.30)$$

e

$$Y(t) := X(t + \Delta), \quad \forall t \in \Omega^+(\Delta). \quad (7.31)$$

Então o operador δ é definido por

$$\delta(\Delta) = \begin{cases} \frac{d}{dt}(\cdot), & \Delta = 0 \\ (z - 1)/\Delta, & \Delta \neq 0 \end{cases} . \quad (7.32)$$

Aplicando-se o operador δ ao sinal $X(t)$, temos que

$$\delta(\Delta) = \begin{cases} \frac{d}{dt}X(t), & \Delta = 0 \\ (X(t + 1) - X(t))/\Delta, & \Delta \neq 0 \end{cases} . \quad (7.33)$$

As motivações para o uso do operador δ são as seguintes (Feuer & Goodwin, 1996; Kauraniemi et al., 1998): *i*) o operador δ não apresenta laços livres de atrasos (*delay-free loops*), é causal e tem uma forma recursiva para a sua implementação, *ii*) o operador δ é dado em função de z e o recíproco é verdadeiro, *iii*) $\lim_{\Delta \rightarrow 0^+} \delta(\Delta) = \delta(0)$.

A implementação da i -ésima seção de 2^a ordem de um filtro notch (Hirano et al., 1974) no domínio do operador δ (Feuer & Goodwin, 1996; Goodwin et al., 1992) é dado por

$$H_i(\delta) = H_i(z) \Big|_{z=1+\Delta\delta} = \frac{1 + \alpha_{i,1}\delta^{-1} + \alpha_{i,2}\delta^{-2}}{1 + \beta_{i,1}\delta^{-1} + \beta_{i,2}\delta^{-2}} \quad (7.34)$$

onde

$$\alpha_{i,1} = \frac{2}{\Delta}(1 - \cos \omega_i), \quad (7.35)$$

$$\alpha_{i,2} = \frac{2}{\Delta^2}(1 - \cos \omega_i), \quad (7.36)$$

$$\beta_{i,1} = \frac{2}{\Delta}(1 - \rho_i \cos \omega_i), \quad (7.37)$$

$$\beta_{i,2} = \frac{1 + \rho_i^2 - 2\rho_i \cos \omega_i}{\Delta^2}, \quad (7.38)$$

$i = 0, \dots, M$ e $\Delta \in [0, \infty)$. Apesar da definição do parâmetro Δ dada acima, aqui o mesmo é considerado como um parâmetro a ser especificado de forma que os erros de *roundoff* sejam reduzidos.

A implementação do filtro notch baseada em (7.34) apresenta melhor desempenho que a implementação baseada na expressão (7.9) quando os coeficientes do filtro são

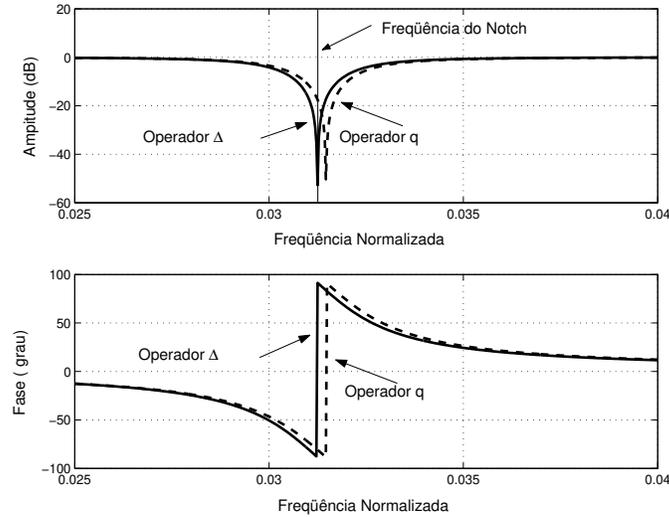


Figura 7.7: Resposta em frequência do filtro Notch de 2ª ordem nos domínios z e do operador Δ . Os coeficientes do filtro foram quantizados com 12 bits.

implementados com precisão finita (Feuer & Goodwin, 1996; Goodwin et al., 1992; Kauraniemi et al., 1998). De fato, a escolha correta do parâmetro Δ resulta em ruído de *roundoff* baixo na saída do filtro notch no domínio do operador δ .

A Fig. 7.7 mostra as respostas em frequência do filtro notch de 2ª ordem nos domínios dos operadores z e δ . Para obter essas respostas, utilizou-se 12 bits para a quantização dos coeficientes do filtro notch, $\Delta = 2^{-5}$, $\rho = 0.995$, $f_s = 64 \times 60$ Hz, $f_0 = 60$ Hz e a estrutura do filtro notch de 2ª ordem é II transposta, pois a mesma apresenta o melhor desempenho dentre todas as estruturas (Kauraniemi et al., 1998). Conforme pode ser observado na Fig. 7.7, o domínio do operador δ garante maior robustez frente aos efeitos da precisão finita do que o domínio do operador z . Note que as respostas mostradas na Fig. 7.7, o ganho observado com o filtro notch no domínio do operador δ em relação ao filtro notch no domínio do operador z é de cerca de 35 dB na frequência do notch. De fato, o uso da precisão finita faz com que o filtro notch implementado no domínio z seja polarizado.

7.2 Filtro Notch Adaptativo baseado em EOS

A principal vantagem relacionada ao uso de EOS (Mendel, 1991) para a atualização dos coeficientes do filtro notch adaptativo é a sua inerente insensibilidade ao ruído branco e Gaussiano e o melhor desempenho obtido na presença de ruído colorido em relação ao uso de E02 (Gharieb, 2000) para a atualização do filtro notch adaptativo. As características das EOS são relevantes visto que as componentes harmônicas, inter-harmônicas e transitórias podem ser vistas como ruído colorido. Outra informação bastante relevante é que na presença de ruído colorido, o desempenho dos filtros notch adaptativos, baseado em EO2, é ruim, pois informações importantes do ruído colorido e do sinal estão contidas nas EOS (Mendel, 1991).

De acordo com (Gharieb, 2000), para estimar a frequência da componente fundamental, deve-se calcular a *diagonal second-order mixed cumulant slice* (DSOMCS) do sinal $\{x(n)\}$ que é dado por

$$\begin{aligned} \hat{c}_x(\tau)^{(L)} &= \hat{c}_x(\tau, \tau, \tau)^{(L)} \\ &\approx \frac{1}{L} \left(\sum_{n=0}^{L-1} x(n)x^3(n+\tau) - 3 \sum_{n=0}^{L-1} x(n)x(n+\tau) \sum_{n=0}^{L-1} x^2(n+\tau) \right), \end{aligned} \quad (7.39)$$

onde L é o comprimento da seqüência $\{x(n)\}$, $\tau = 0, \dots, P-1$ e $2 \leq P \leq L-1$. Em seguida deve-se utilizar a solução dos quadrados mínimos (*least squares solution*) às equações do filtro *forward* de predição que relacionam o parâmetro a_0 do filtro notch adaptativo à estimativa do DSOMCS, ou seja,

$$\hat{a}_0 = -2 \cos \hat{\omega}_0 = -(\Phi^T \Phi)^{-1} \Phi^T \Theta, \quad (7.40)$$

onde $\Theta = [\hat{c}_x(2)^{(L)} + \hat{c}_x(0)^{(L)} \quad \dots \quad \hat{c}_x(P-1)^{(L)} + \hat{c}_x(P-3)^{(L)}]^T$ e $\Phi = [\hat{c}_x(1)^{(L)} \quad \dots \quad \hat{c}_x(P-2)^{(L)}]^T$. Note que o resultado obtido pela expressão (7.40)

é uma estimativa grosseira da frequência da componente fundamental. No entanto, é importante observar que esta solução apresenta baixo custo computacional e o seu resultado serve de parâmetro, baseado no qual a WDFT combinada com a rede MLP, estima, com maior precisão, a frequência da componente fundamental.

7.3 Transformada Warped-DFT (WDFT)

Considere uma seqüência, $x(n)$, $n = 0, \dots, N - 1$, cuja transformada z é dada por

$$X(z) = \sum_{n=0}^{N-1} x(n)z^{-n}. \quad (7.41)$$

As amostras $X[k]$, $k = 0, \dots, N - 1$, da DFT da seqüência $\{x(n)\}$ é definida pelas amostras uniformemente espaçadas de $X(z)$ sobre o ciclo de raio unitário (Mitra, 2001) dadas por

$$X[k] = X(z)|_{z=e^{j2\pi k/N}} = \sum_{n=0}^{N-1} x(n)e^{-j(2\pi/N)kn}, \quad k = 0, \dots, N - 1, \quad (7.42)$$

e a inversa da DFT (IDFT) é dada por

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X[k]e^{j(2\pi/N)kn}, \quad n = 0, \dots, N - 1. \quad (7.43)$$

Para a análise espectral do sinal $\{x(n)\}$, a DFT fornece uma resolução espectral constante, dada por $2\pi/N$, na faixa espectral igual a $[0, 2\pi)$ (Mitra, 2001; Bagchi & Mitra, 1999). A DFT é amplamente utilizada, visto que a mesma pode ser eficientemente calculada usando-se algoritmos rápidos, tais como as várias versões da FFT.

Uma possível generalização da DFT foi introduzida em (Bagchi & Mitra, 1999). Esta generalização denominada *Nonuniform discrete Fourier transform* (NDFT) é

dada por (Bagchi & Mitra, 1999)

$$X_{NDFT}[k] = \sum_{n=0}^{N-1} x(n)z_k^{-n}, \quad 0 \leq k \leq N-1, \quad (7.44)$$

onde z_k , $k = 0, \dots, N-1$, são pontos distintos no plano z . Para calcular (7.44) sobre o círculo de raio unitário basta substituir z_k por $e^{j\Omega_k}$, $k = 0, \dots, N-1$, sendo Ω_k a frequência normalizada em *rad/amostras* que gera amostras não uniformemente espaçadas. Usando notação matricial, (7.44) pode ser dada por

$$\mathbf{X} = \mathbf{D}\mathbf{x}, \quad (7.45)$$

onde $\mathbf{x} = [x(0), \dots, x(N-1)]^T$ é o vetor de amostras do sinal, $\mathbf{X} = [X[0], \dots, X[N-1]]^T$ é o vetor de amostras das frequências Ω_k , $k = 0, \dots, N-1$, e \mathbf{D} é a matriz de Vandermode dada por

$$\mathbf{D} = \left\{ \begin{array}{ccccc} 1 & e^{-j\Omega_0} & e^{-j2\Omega_0} & \dots & e^{-j(N-1)\Omega_0} \\ 1 & e^{-j\Omega_1} & e^{-j2\Omega_1} & \dots & e^{-j(N-1)\Omega_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\Omega_{N-1}} & e^{-j2\Omega_{N-1}} & \dots & e^{-j(N-1)\Omega_{N-1}} \end{array} \right\}. \quad (7.46)$$

Para valores distintos de Ω_k a inversa da NDFT existe e é dada por

$$\mathbf{x} = \mathbf{D}^{-1}\mathbf{X}. \quad (7.47)$$

A warped DFT (WDFT) é um caso particular da NDFT (Makur & Mitra, 2001). A WDFT é obtida a partir das amostras espaçadas de $\tilde{X}(z)$ sobre o círculo de raio unitário. $\tilde{X}(z)$ é obtido pela substituição de z^{-1} em $X(z)$ pela função do filtro passado. Desta forma, $\tilde{X}(z)$ é dado por

$$\tilde{X}(z) = \sum_{n=0}^{N-1} x(n)B(z)^n. \quad (7.48)$$

Dado $\tilde{X}(z)$, o mapeamento dos valores na frequência são *warped* e, conseqüentemente, as amostras uniformemente espaçadas de $\tilde{X}(z)$ sobre o círculo de raio unitário correspondem as amostras não uniformemente espaçadas de $X(z)$.

A função $B(z)$ é um filtro passa tudo de 1ª ordem dado por

$$B(z) = \frac{\alpha^* + z^{-1}}{1 + \alpha z^{-1}}, \quad (7.49)$$

onde $\alpha = |\alpha|e^{j\varphi}$ é o parâmetro complexo que determina o *warping*. Note que $|\alpha|$ determina o nível de concentração das funções de bases da WDFT em torno do valor da frequência angular dada por φ . O mapeamento da frequência é dado por

$$\omega_w = \omega + 2 \arctan \left(\frac{|\alpha| \sin(\varphi - \omega)}{1 + |\alpha| \cos(\varphi - \omega)} \right), \quad -\pi \leq \omega \leq \pi, \quad (7.50)$$

onde Ω_w é a frequência *warped* em rad/s. Quando o parâmetro de *warping* é real, o mapeamento da frequência é dada por

$$\omega_w = 2 \arctan \left(\frac{(1 - \alpha)}{(1 + \alpha)} \tan \left(\frac{\omega}{2} \right) \right). \quad (7.51)$$

Finalmente, dado (7.49), tem-se que

$$X_{WDFT}[k] = \sum_{n=0}^{N-1} x(n) \left(\frac{\alpha^* + e^{-j\frac{2\pi}{N}k}}{1 + \alpha e^{-j\frac{2\pi}{N}k}} \right)^n, \quad k = 0, \dots, N-1. \quad (7.52)$$

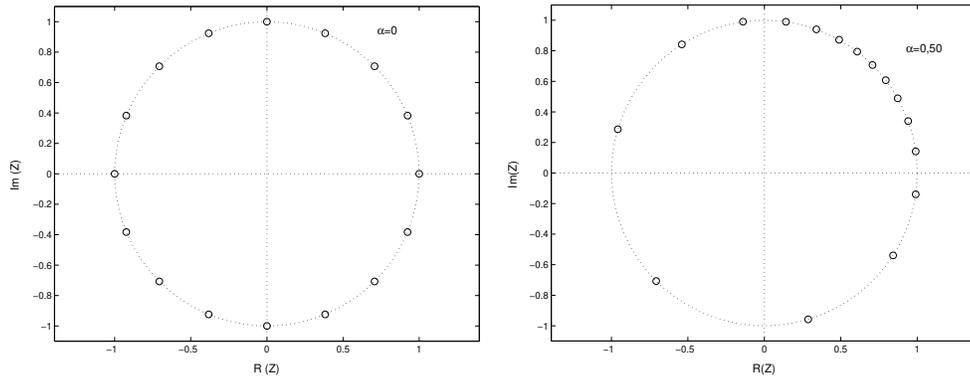
Observe que se $\alpha = 0$ em (7.52) então $X_{WDFT}[k] = X[k]$. Apesar da existência das equações (7.50) e (7.51) para se calcular a frequência angular da componente senoidal, é possível utilizar a expressão

$$\omega = \arg \left(\frac{\alpha^* + e^{-j(2\pi/N)k}}{1 + \alpha e^{-j(2\pi/N)k}} \Big|_{k=1} \right) \quad (7.53)$$

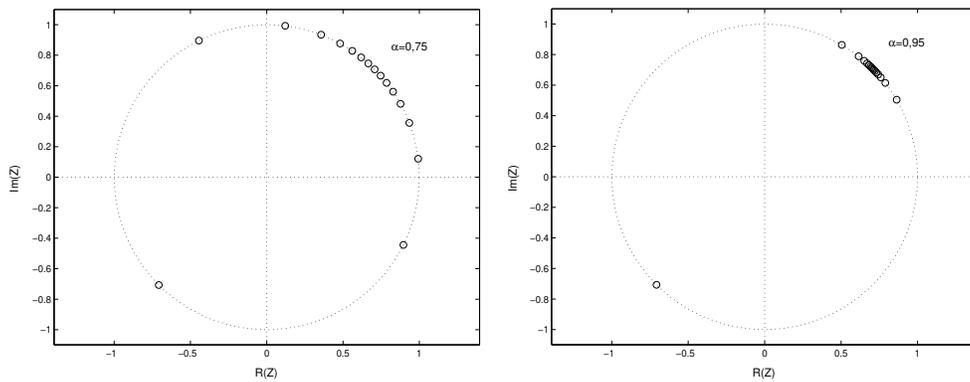
com o mesmo objetivo. A vantagem ao se usar (7.53) é a baixa complexidade computacional quando comparada com as expressões (7.51) e (7.52).

A Fig. 7.8 mostra o comportamento da WDFT para diferentes valores de $|\alpha|$. Conforme pode ser observado na Fig. 7.8(a), quando $|\alpha| = 0$ a WDFT se reduz a DFT. As Figs. 7.8(a)-7.8(d) ainda revelam que quando $|\alpha| \rightarrow 1$ a concentração de

zeros próximo do valor de fase de φ aumente consideravelmente. Como consequência, pode-se dizer que a WDFT possibilita amostrar componentes espectrais que residam nos intervalos $[i2\pi/N, (i+1)2\pi/N]$, $i = 1, \dots, N-1$. Tal propriedade é bastante interessante do ponto de vista da análise espectral do sinal $\{x(n)\}$.



(a) Localização dos zeros da WDFT quando $\alpha = 0,00$. (b) Localização dos zeros da WDFT quando $\alpha = 0,50$.



(c) Localização dos zeros da WDFT quando $\alpha = 0,75$. (d) Localização dos zeros da WDFT quando $\alpha = 0,95$.

Figura 7.8: Comportamento da WDFT para valores diferentes de $|\alpha|$, quando $\varphi = \pi/8$ e $N = 16$.

Note que a resolução da DFT depende de $2\pi/N$, ver (7.42) e, conseqüentemente, um aumento da resolução espectral requer um aumento do valor de N . Entretanto,

o aumento de N demanda maior complexidade computacional e atraso. Nesse contexto, a WDFT aparece como uma solução bastante interessante para a análise espectral de $\{x(n)\}$, posto que dado o valor do parâmetro α em (7.49) é possível concentrar um grande número de amostras espectrais em torno de um único ponto. Desta forma, mesmo utilizando-se uma seqüência com reduzido número de amostras do sinal $\{x(n)\}$ é possível obter elevadas resoluções espectrais.

Várias propriedades da DFT, tais como linearidade, independência aos deslocamentos temporal e espectral, simetria e etc se mantém para a WDFT. Entretanto, as propriedades de periodicidade no domínio do tempo, convolução cíclica e o teorema de Parseval não são válidas na WDFT (Makur & Mitra, 2001; Franz et al., 2003).

Analisando-se a complexidade computacional da WDFT, pode-se observar que a implementação da WDFT baseada no algoritmo de Goertzel demanda $2(N+2)$ multiplicações reais e $4(N+1)$ adições reais para o cálculo de um único coeficiente da WDFT. Assim sendo, a WDFT com N amostras demanda $2(N+2)N$ multiplicações reais e $4(N+1)N$ adições reais (Makur & Mitra, 2001; Franz et al., 2003).

7.4 Resultados Numéricos

Nesta seção o desempenho da técnica NF-WDFT é comparado com os desempenhos das técnicas DFT-WDFT (Franz et al., 2003) e FFT (*Fast Fourier Transform*). Os resultados apresentados a seguir representam algumas das simulações que foram realizadas. Os sinais utilizados para analisar o desempenho das técnicas foram gerados de acordo com (Moore & Português, 2003).

As tabelas 7.1-7.3 mostram os resultados obtidos com o NF-WDFT e o DFT-

WDFT (DFT com 64 pontos) para estimar a frequência da componente fundamental do sinal

$$\begin{aligned} x(n) &= f(n) + h(n) + v(n) = \\ &= A_0 \cos(n\omega_0 + \theta_0) + \sum_{i=1}^7 A_{2i+1} \cos(n\omega_{2i+1} + \theta_i) + v(n) , \end{aligned} \quad (7.54)$$

onde θ_i , $i = 0, \dots, 7$ são variáveis aleatórias com distribuição uniforme no intervalo $[0, 2\pi)$, ω_i é obtido a partir de (7.11) e $v(n)$ é uma variável aleatória *i.i.d.* e distribuição Gaussiana $\mathcal{N}(0, \sigma_v^2)$. Os valores de A_i , $i = 0, \dots, 7$ são iguais à 1,000, 0,757, 0,426, 0,145, 0,025, 0,009, 0,058 e 0,007, respectivamente. As frequências de amostragem são $f_s = 64 \times 60$ Hz quando as técnicas DFT-WDFT e NF-WDFT são aplicadas e $f_s = 128 \times 60$ Hz quando a técnica FFT é aplicada. Os valores de f_0 são iguais a 57,430 Hz, 60,110 Hz e 63,890 Hz. 5/4 ciclos da fundamental foram utilizados para a estimação da frequência da componente fundamental.

Os resultados mostrados nas tabelas 7.1-7.3 estão relacionados à diferentes valores de SNR, dado por (7.1), que variam em função dos valores de σ_v^2 considerados. Os resultados apresentados foram obtidos com 5000 simulações Monte Carlo (Jeruchim, Balaban, & Shanmugan, 2000).

Para o uso da técnica NF-WDFT foram considerados $\rho = 0.9985$, $|\alpha| = 0.9$ e $\angle\alpha$ é a estimativa da frequência obtida com a WDFT. A estrutura da rede MLP é $2 \times 6 \times 1$. Os filtros notchs foram implementados com 12 bits e $\Delta = 2^{-5}$.

Conforme pode ser observado, a técnica NF-WDFT obtém melhores desempenhos que a técnica DFT-WDFT. Ainda pode ser observado que quando σ_v^2 é alto, a técnica NF-WDFT apresenta bons resultados. De fato, o pré-processamento realizado pelos filtros notches provê a SNR dada por (7.8) em vez da SNR dada por (7.1) e, conseqüentemente, o processo de estimação das amplitudes é facilitado.

Apesar de não mostrado aqui, os desempenhos das técnicas DFT-WDFT e NF-WDFT são melhores quando f_s aumenta.

σ_v^2	<i>NF – WDFT</i>	<i>DFT – WDFT</i>
$-5dB$	57,355	54,662
$-7.5dB$	57,432	54,677
$-10dB$	57,429	54,631
$-12.5dB$	57,432	54,631
$-15dB$	57,431	54,661
$-20dB$	57,430	54,650

Tabela 7.1: Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 57,430$ Hz.

σ_v^2	<i>NF – WDFT</i>	<i>DFT – WDFT</i>
$-5dB$	60,113	57,220
$-7.5dB$	60,108	57,249
$-10dB$	60,112	57,200
$-12.5dB$	60,112	57,208
$-15dB$	60,111	57,195
$-20dB$	60,110	57,209

Tabela 7.2: Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 60,110$ Hz.

Para analisar o desempenho das técnicas NF-WDFT, DFT-WDFT e FFT na presença das componentes harmônicas, o sinal dado por (7.54) foi considerado, sendo $\sigma_v^2 = -20$ dB. 5/4 ciclos da componente fundamental foram utilizados pelo NF-WDFT e DFT-WDFT quando $f_s = 60 \times 64$ Hz. Devido as limitações de desempenho da FFT, a mesma foi implementada com 128 e 256 amostras e $f_s = 60 \times 128$ Hz. A rede MLP utilizada para estimar as amplitudes das componentes harmônicas

σ_v^2	NF – WDFT	DFT – WDFT
$-5dB$	62,898	59,861
$-7.5dB$	62,892	59,864
$-10dB$	62,902	59,815
$-12.5dB$	62,891	59,850
$-15dB$	62,889	59,847
$-20dB$	62,890	59,839

Tabela 7.3: Desempenho das técnicas NF-WDFT e DFT-WDFT para a estimação da frequência, quando $f_0 = 62,890$ Hz.

tem estrutura $2 \times 5 \times 1$. As tabelas 7.4-7.6 mostram o desempenho das técnicas NF-WDFT, DFT-WDFT e FFT (com 128 e 256 amostras) quando $f_0 = 57,00$ Hz, $f_0 = 60,01$ Hz e $f_0 = 62,50$ Hz. Esses resultados foram obtidos a partir da realização de 5000 simulações Monte Carlo. Conforme pode ser observado nas Tabs. 7.4-7.6, a técnica NF-WDFT apresenta os melhores resultados, mas demanda maior complexidade computacional. Além disso, pode-se notar que o desempenho da técnica NF-WDFT é superior quando as amplitudes das componentes harmônicas apresentam as seguintes características $0 < A_0 \ll 1$.

Uma análise importante, baseada nos resultados mostrados nas tabelas 7.4-7.6, para avaliar o quão representativas são as estimativas obtidas com as técnicas avaliadas, é o cálculo da *Distorção Harmônica Total* (DHT). Os valores calculados da DHT, a partir dos resultados mostrados nas tabelas 7.4-7.6 são reunidos na tabela 7.7. Note que a coluna 1 informa os valores ideais da DHT. Conforme é observado nesta tabela, o melhor desempenho da técnica NF-WDFT para a estimação das amplitudes das componentes harmônicas reflete diretamente no cálculo da DHT com melhor precisão.

Para mostra o desempenho da técnica NF-WDFT, quando aplicada à estimação

<i>Amplitude</i>	<i>NF – WDFT</i>	<i>DFT – WDFT</i>	<i>DFT – 128</i>	<i>DFT – 256</i>
1 ^a 1,0000	1,0008	1,0146	1,0007	1,0004
3 ^a 0,3030	0,3009	0,3596	0,3063	0,3037
5 ^a 0,3100	0,3073	0,3404	0,3132	0,3111
7 ^a 0,1070	0,1065	0,1306	0,1134	0,1104
9 ^a 0,0340	0,0356	0,0791	0,0580	0,0469
11 ^a 0,0550	0,0554	0,1078	0,0720	0,0627
13 ^a 0,0550	0,0552	0,1069	0,0716	0,0627
15 ^a 0,0090	0,0064	0,0847	0,0503	0,0357

Tabela 7.4: Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 57,00$ Hz.

<i>Amplitude</i>	<i>NF – WDFT</i>	<i>DFT – WDFT</i>	<i>DFT – 128</i>	<i>DFT – 256</i>
1 ^a 1,0000	1,0006	1,0101	1,0013	1,0001
3 ^a 0,3030	0,3011	0,2966	0,3052	0,3047
5 ^a 0,3100	0,3075	0,3361	0,3124	0,3115
7 ^a 0,1070	0,1063	0,1669	0,1146	0,1104
9 ^a 0,0340	0,0357	0,1126	0,0586	0,0470
11 ^a 0,0550	0,0556	0,0483	0,0707	0,0627
13 ^a 0,0550	0,0553	0,1243	0,0707	0,0625
15 ^a 0,0090	0,0064	0,0962	0,0502	0,0363

Tabela 7.5: Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 60,01$ Hz.

das amplitudes das componentes harmônicas, foi utilizado o sinal de corrente gerado pelo disco flexível de um microcomputador, de acordo com (Moore & Português, 2003). Esse sinal é também representado por (7.54), sendo $\sigma_v^2 = -40$ dB. Os resultados mostrados na tabela 7.8 foram obtidos com 4000 simulações Monte Carlo.

<i>Amplitude</i>	<i>NF – WDFT</i>	<i>DFT – WDFT</i>	<i>DFT – 128</i>	<i>DFT – 256</i>
1 ^a 1,0000	1,0008	1,0110	1,0005	1,0005
3 ^a 0,3030	0,3009	0,3281	0,3051	0,3043
5 ^a 0,3100	0,3073	0,3373	0,3120	0,3112
7 ^a 0,1070	0,1065	0,0629	0,1145	0,1109
9 ^a 0,0340	0,0356	0,1132	0,0581	0,0473
11 ^a 0,0550	0,0554	0,0530	0,0704	0,0629
13 ^a 0,0550	0,0552	0,0623	0,0708	0,0627
15 ^a 0,0090	0,0064	0,0464	0,0500	0,0360

Tabela 7.6: Desempenho das técnicas NF-WDFT, DFT-WDFT e FFT para a estimação de amplitudes, quando $f_0 = 62,50$ Hz.

f_0 (Hz)	<i>DHT, Ideal</i>	<i>NF – WDFT</i>	<i>DFT – WDFT</i>	<i>DFT – 128</i>	<i>DFT – 256</i>
57,00	45,46	45,10	53,87	46,97	46,08
60,01	45,46	45,13	49,16	46,82	46,14
62,50	45,46	45,14	51,30	46,83	46,19

Tabela 7.7: DHT em (%) obtidas com as técnicas NF-WDFT, DFT-WDFT e FFT, a partir dos resultados das tabelas 7.4-7.6

A estrutura da rede MLP é $1 \times 3 \times 1$. Conforme pode ser observado na Tab. 7.8, a técnica NF-WDFT obtém bons resultados quando aplicada para a estimação das fases das componentes harmônicas.

Baseando-se nos resultados mostrados acima, fica claro a relevância da aplicação da técnica NF-WDFT para sistemas de potência, uma vez que a mesma foi desenvolvida em função das características intrínsecas das componentes harmônicas do sinais de potência. Note ainda que a técnica NF-WDFT demanda custo computacional superior a FFT e a DFT-WDFT, pois faz uso de várias outras técnicas que requerem maior esforço computacional.

<i>Harm.</i>	θ	$\hat{\theta}$
1 ^a	0°	0,04°
3 ^a	181°	181,04°
5 ^a	356°	356,05°
7 ^a	163°	162,97°
9 ^a	243°	243,03°
11 ^a	351°	350,94°
13 ^a	181°	181,03°
15 ^a	20°	20,05°

Tabela 7.8: Estimação da fase das componentes harmônicas com a técnica NF-WDFT.

7.5 Sumário

Este capítulo apresentou uma nova técnica para a análise espectral das componentes fundamental e harmônicas em sistemas de potência. A principal inovação introduzida pela técnica proposta é o uso da transformada *warped* DFT que permite o aumento da resolução espectral, mesmo que baixas taxas de amostragem sejam utilizadas na aquisição do sinal $\{x(n)\}$. Conforme observado, melhores resultados são obtidos com a técnica proposta quando menores taxas de amostragem do sinal monitorado são utilizadas. Apesar da melhoria de desempenho observada com a técnica proposta, pode-se observar que a mesma demanda maior custo computacional do que as outras técnicas aqui analisadas.

8

Detecção de Distúrbios

A detecção de transitórios surge em várias áreas de aplicação, tais como telemetria, controle de tráfego aéreo, sistemas de potência, biomédia, telecomunicações e etc. Dentre as principais motivações para o desenvolvimento de técnicas de detecção de distúrbios em sistemas de potência, estão as seguintes: *i*) a crescente sensibilidade dos equipamentos eletro-eletrônicos devida ao aumento do uso de cargas de natureza não-linear em plantas industriais, comerciais e residenciais. É interessante notar que esses equipamentos, ao mesmo tempo que geram distúrbios, têm seu desempenho alterado, devido à ocorrência de distúrbios na rede elétrica; *ii*) o monitoramento contínuo da tensão e da corrente de forma que medidas preditivas,

preventivas e corretivas sejam tomadas para minimizar os impactos dos distúrbios nas cargas conectadas à rede elétrica; *iii*) a correta detecção da presença dos distúrbios facilita o emprego das técnicas de classificação, de identificação e de compressão de sinais, posto que o uso das mesmas deve ocorrer quando os distúrbios são detectados e, finalmente, *iv*) a minimização do custo computacional, da latência e do problema de erros de deslocamento (Trees, 1968/1971; McDonough & Whalen, 1995).

A detecção da ocorrência de transitórios na janela de observação do sinal $\{x(n)\}$, $n = 0, \dots, N - 1$, pode ser formulada como uma decisão entre duas hipóteses (Trees, 1968/1971; McDonough & Whalen, 1995):

$$\begin{aligned} H_0 : \mathbf{x}(n) &= \mathbf{f}(n) + \mathbf{v}(n) \\ H_1 : \mathbf{x}(n) &= \mathbf{f}(n) + \mathbf{u}(n) + \mathbf{v}(n) \end{aligned} \quad , \quad (8.1)$$

onde $\mathbf{x}(n)$ é o vetor cujas amostras são dadas por (6.1), $\mathbf{u}(n) = \mathbf{h}(n) + \mathbf{i}(n) + \mathbf{t}(n)$ e $\mathbf{v}(n)$ é o ruído aditivo. Os vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$, $\mathbf{i}(n)$ e $\mathbf{t}(n)$ representam as componentes fundamental, harmônicas, inter-harmônicas e transitórios. Sendo L a duração do transitório e supondo $N > L$, pode-se escrever que $\mathbf{x}(n) = [x(n) \dots x(n - N + 1)]^T = \mathbf{f}(n) + \mathbf{u}(n) + \mathbf{v}(n)$, $\mathbf{u}(n) = [\mathbf{0}_d^T, \mathbf{u}_L^T(n), \mathbf{0}_{N-L-d}^T]^T$ é o vetor sinal, $\mathbf{u}_L(n) = [u(n + d + 1) \dots u(n - N + L + d)]^T$ é o vetor constituído por amostras do sinal transitório, $\mathbf{v}(n) = [v(n) \dots v(n - N + 1)]^T$ é o vetor constituído por amostras do ruído de fundo. $\mathbf{0}_m$ é o vetor de comprimento m cujos elementos são iguais a zero. $d + 1$ e $d + L$ são as amostras que determinam o início e o fim do distúrbio. Logo, o intervalo de ocorrência do distúrbio pode ser representado pela seguinte função

$$\pi(n) = \mu(n + d + 1) - \mu(n + d + L), \quad (8.2)$$

onde μ é a função degrau (Mitra, 2001).

Conforme observado em (8.1), a detecção de distúrbios tem várias variáveis,

as quais dependem do tipo da aplicação em questão. Por exemplo, o instante inicial de ocorrência do transitório d pode ou não ser conhecido; a duração do transitório L pode ou não ser conhecida; a forma de onda do sinal $\{u_L(n)\}$ pode ser conhecida, parcialmente conhecida ou completamente desconhecida (McDonough & Whalen, 1995; Fishler & Messer, 1993; Colonnese & Scarano, 1999). Nesse contexto, H_0 é uma simples hipótese e H_1 é uma hipótese composta. Se o ruído $\{v(n)\}$ é um ruído aditivo cujas amostras são *i.i.d.* (independente e identicamente distribuído) e têm distribuição Gaussiana, o teste ótimo da taxa de verosimilhança (TTV) ou o teste ótimo generalizado da taxa de verosimilhança (TGTV) assumem a forma de um filtro casado ou correlador (Trees, 1968/1971; Giannakis & Tsatsanis, 1990; McDonough & Whalen, 1995), entretanto, tais soluções demandam altíssima carga computacional. Note que, se as amostras do ruído $\{v(n)\}$ apresentam distribuições diferentes da distribuição Gaussiana, a implementação do TGTV é bastante complicada (Trees, 1968/1971; Fishler & Messer, 1993; McDonough & Whalen, 1995).

No contexto dos sistemas de potência, diversas técnicas de detecção de transitórios ou de distúrbios estão sendo desenvolvidas (Duque, Ribeiro, Ramos, & Szczupak, 2005; Poisson, Rioual, & Meunier, 2000; Ferrero & Salicone, 2004; Yang & Liao, 2001; Karimi, Mokhtari, & Iravani, 2000; Mokhtari, K.-Ghartemani, & Iravani, 2002; Ece & Gerek, 2004; Angrisani, Daponte, D'Apuzzo, & Pietrosanto, 1999; Lu & Huang, 2004; Abdel-Galil, El-Saadany, & Salama, 2003; Gu & Styvaktakis, 2004; Dash & Chilukuri, 2004; Murphy, 1996; Zhang, Liu, & Malik, 2003; Gaouda, Salama, Sultan, & Chikhani, 1999; Santoso, Powers, & Hofmann, 1996; Castaldo, Gallo, Landi, & Testa, 2004; Ramos, Ribeiro, Romano, & Duque, 2002; Ribeiro, Romano, & Deckmann, 2003). De fato, a correta detecção dos instantes inicial e final da ocorrência dos distúrbios nas linhas de transmissão ou de distribuição é de grande valia para a análise futura do tipo de problema, bem como da fonte causadora dos distúrbios.

É importante salientar que a grande maioria dessas técnicas faz uso da *Transformada Wavelet* (TW). De fato, a TW, quando combinada com outras técnicas, tem possibilitado o desenvolvimento de soluções novas e eficientes para o problema da detecção de distúrbios. Entretanto, quando se dimensiona uma solução computacional adequada para o monitoramento descentralizado da QE, o uso de tais soluções demandam um custo computacional expressivo. Por outro lado, em (Duque et al., 2005; Ramos et al., 2002) foi apresentada uma nova técnica de detecção que demanda baixo custo computacional e faz uso do princípio dividir e conquistar e o da inovação, utilizado na formulação do filtro de Kalman (Anderson & Moore, 1979), para gerar um solução com baixo custo computacional. No entanto, a técnica de detecção proposta em (Duque et al., 2005; Ramos et al., 2002) leva em consideração a estimação da variância do ruído de fundo, o que demanda o uso de outra técnica para que tal parâmetro seja calculado de forma bastante precisa. Além disso, essa técnica é aplicada a cada nova amostra do sinal monitorado e uma janela de 2 ciclos da componente fundamental é considerada para a detecção da ocorrência de distúrbios. Note ainda que o sinal $\{u(n)\} = \{h(n)\} + \{i(n)\} + \{t(n)\} + \{v(n)\}$, o qual é utilizado na técnica proposta em (Duque et al., 2005; Ramos et al., 2002), não é um ruído branco e Gaussiano e, assim sendo, analisar apenas as EO2 (estatísticas de ordem 2) não é uma solução adequada para este tipo de problema, pois as EOS (estatísticas de ordem superior) existem e contém informações representativas do sinal $\{u(n)\}$.

Visando minimizar os problemas acima descritos, a Fig. 8.1 apresenta o diagrama de blocos de uma nova técnica para a detecção de distúrbios (Ribeiro & Romano, 2005a). A descrição de cada bloco da técnica de detecção proposta é dada a seguir.

O Bloco NF_0 é o bloco NF_0 da técnica NF-WDFT, discutida no capítulo 7. Note que a saída do bloco NF_0 é dada por (7.12), ou seja,

$$u(n) := x(n + d_0) - |H_0(e^{j\omega_0(n)})| A_0(n) \cos(n\omega_0(n) + \theta_0(n) + \Delta\theta_0(n)). \quad (8.3)$$

O Bloco Extração de Parâmetros implementa os algoritmos de processamento de sinais que extraem informações do sinal expresso por (8.3). Os parâmetros extraídos são utilizados por um outro algoritmo que decide se no sinal dado por (8.1) existe ou não a presença de distúrbios.

O Bloco de Detecção implementa o algoritmo que verifica a ocorrência de distúrbios a partir do vetor de parâmetros extraídos do sinal monitorado. Além disso, este bloco implementa o algoritmo que detecta os instantes inicial e final de ocorrência do distúrbio na seqüência analisada.

O Bloco Saída faz uso de um número binário para informar a ocorrência de distúrbio no sinal $\{x(n)\}$ e armazena os instantes inicial e final de ocorrência do distúrbio.

A técnica de detecção de distúrbio proposta é inspirada no conceito *Solução hierárquica baseada em dois níveis*, descrito no início do capítulo 6 pois, primeiramente, o algoritmo implementado no bloco detecção realiza a detecção do distúrbio ou transitório e, a seguir, um outro algoritmo determina os instantes inicial e final de ocorrência do distúrbio. Essa estratégia é bastante interessante por dois motivos: *i)* a detecção do distúrbio é verificada em cada vetor monitorado, ou seja, dada uma seqüência de comprimento N qualquer, cujos elementos são obtidos a partir das amostras dadas por (8.3), parâmetros são extraídos e a presença de distúrbios é detectada; *ii)* ao ser detectada a ocorrência do distúrbios na janela de sinal, técnicas que explorem a não estacionaridade dos distúrbios podem ser desenvolvidas e aplicadas para a obtenção de estimações mais precisas dos instantes inicial e final de ocorrência do distúrbio na janela analisada.

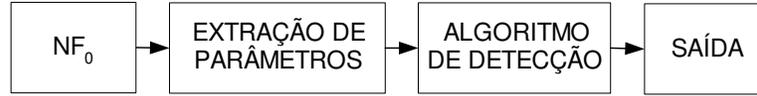


Figura 8.1: Técnica para a detecção de distúrbios.

As seções que seguem objetivam descrever a técnica de detecção de ocorrência de distúrbios aqui proposta. A seção 8.1 destaca os parâmetros que são extraídos do sinal dado por (8.3), a seção 8.2 discute a técnica utilizada para implementar a decisão baseada em (8.1) e a seção 8.3 apresenta os resultados obtidos com a solução proposta.

8.1 Técnica para a Extração de Parâmetros

A literatura mostra que as EOS são utilizadas com sucesso em processos não-Gaussianos e em sistemas não-lineares. Vale a pena citar o uso de EOS para a detecção e a classificação de sinais, a identificação de sistemas (Nikias & Mendel, 1999, 1987; Nikias & Petropulu, 1993; Mendel, 1991), e a detecção de transitórios (Colonnese & Scarano, 1999; Fishler & Messer, 1993; Nemer, Goubran, & Mahmoud, 2001; Gürçan, Çetin, & Ansari, 1997). Uma definição sucinta sobre as EOS, aqui utilizada, é dada a seguir.

Seja $\{u(n)\}$ uma seqüência cujos momentos até a ordem p existam, então a p -ésima função momento é dada por

$$m_p(\tau_1, \tau_2, \dots, \tau_{p-1}) = E\{u(n)u(n + \tau_1)\dots u(n + \tau_{p-1})\} \quad (8.4)$$

Note que $m_p(\tau_1, \tau_2, \dots, \tau_{p-1})$ depende das variáveis $\tau_1, \tau_2, \dots, \tau_{p-1}$ que assumem os valores $\pm 1, \pm 2, \pm 3, \dots$.

Agora, suponha que $\{u(n)\}$ seja média zero, então os seus cumulantes de terceira

e de quarta ordem são dadas por

$$C_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2) \quad (8.5)$$

e

$$C_4(\tau_1, \tau_2, \tau_3) = m_4(\tau_1, \tau_2, \tau_3) - m_2(\tau_1)m_2(\tau_3 - \tau_2) - m_2(\tau_2)m_2(\tau_3 - \tau_1) - m_2(\tau_3)m_2(\tau_2 - \tau_1), \quad (8.6)$$

respectivamente. Fazendo as variáveis $\tau_1 = \tau_2 = \tau_3 = 0$ em (8.5)-(8.6), tem-se que a *Skewness* e o *Kurtosis* são dados por

$$C_3^u(0, 0) = E\{u^3(n)\} \quad (8.7)$$

e

$$C_4^u(0, 0, 0) = E\{u^4(n)\} - 3[E\{u^2(n)\}]^2, \quad (8.8)$$

respectivamente.

Se o sinal $\{u(n)\}$ é finito, as estimativas normalizadas da *Skewness* e do *Kurtosis* são dadas por (Nikias & Mendel, 1999), (Mendel, 1991)

$$\gamma_3 = \bar{C}_3^u(0, 0) = \frac{C_3(0, 0)}{[C_2(0)]^{\frac{3}{2}}} = \frac{E\{u^3(n)\}}{[E\{u^2(n)\}]^{\frac{3}{2}}} \quad (8.9)$$

e

$$\bar{C}_4^u(0, 0, 0) = \frac{C_4(0, 0, 0)}{[C_2(0)]^2} = \frac{E\{u^4(n)\}}{[E\{u^2(n)\}]^2} - 3, \quad (8.10)$$

respectivamente, onde

$$C_2^u(0) = E\{u^2(n)\}. \quad (8.11)$$

Agora, supondo que o termo $H_0(e^{j\omega_0(n)})$ em (8.3) é ideal, tem-se que o sinal na saída do bloco NF_0 é dado por

$$u(n) = h(n) + i(n) + t(n) + v_0(n), \quad n = 0, \dots, N - 1, \quad (8.12)$$

sendo os sinais $\{h(n)\}$, $\{i(n)\}$ e $\{t(n)\}$ dados por (6.3)-(6.5) e o sinal $\{v_0(n)\}$ um ruído colorido de fundo gerado pela filtragem do sinal $\{x(n)\}$ realizada pelo filtro notch implementado no bloco NF_0 , vide capítulo 7. A expressão (8.12) evidencia que o sinal $\{u(n)\}$ não tem distribuição Gaussiana, ou seja, os cumulantes de ordem maior que 2 são diferentes de zero e, portanto, revelam a ocorrência de distúrbios no sinal $\{u(n)\}$. Por outro lado, se o sinal $\{x(n)\}$ não apresentar nenhum distúrbio, o sinal $\{u(n)\}$ observado na saída do bloco NF_0 é dado por

$$u(n) = v_0(n), \quad n = 0, \dots, N - 1. \quad (8.13)$$

Devido as características do filtro notch, o ruído colorido de fundo em (8.13) apresenta cumulantes de ordem maior que 2 aproximadamente iguais a zero. Tal informação é bastante valiosa, pois a mesma indica a não ocorrência de distúrbios no sinal $\{x(n)\}$.

Para fins de detecção, a *Skewness* dada por (8.9) e o *Kurtosis* normalizado e modificado

$$\gamma_4 = \overline{C}_4^u(0, 0, 0) = \frac{E\{u^4(n)\}}{[E\{u^2(n)\}]^2}, \quad (8.14)$$

são extraídas do sinal $\{u(n)\}$ para caracterizar a ocorrência ou não de distúrbios no sinal $\{x(n)\}$. Além da extração dos parâmetros *Skewness* e *Kurtosis* normalizados do sinal, também foram extraídos os parâmetros de 3ª ordem do módulo do sinal (Gürçan et al., 1997) e a estatística de 4ª ordem (Maranda & Fawcett, 1990), os quais são expressos por

$$S^{(3)} = \frac{1}{N} \sum_{n=0}^{N-1} |u(n)|^3 \quad (8.15)$$

e

$$S^{(4)} = \frac{1}{N} \sum_{n=0}^{N-1} u^4(n), \quad (8.16)$$

respectivamente.

O vetor de parâmetros extraído do sinal (8.3) para fins de detecção de distúrbios

é dado por

$$\mathbf{v} = [\gamma_3 \ \gamma_4 \ S^{(3)} \ S^{(4)}]^T. \quad (8.17)$$

Os dados numéricos mostrados na seção 8.3 evidenciam que a detecção de distúrbios com alta taxa de acerto pode ser obtida a partir dos seguintes vetores de parâmetros extraídos do sinal:

$$\mathbf{v}_1 = [\gamma_3 \ \gamma_4 \ S^{(3)} \ S^{(4)}]^T. \quad (8.18)$$

$$\mathbf{v}_4 = [\gamma_4 \ S^{(3)}]^T. \quad (8.19)$$

8.2 Algoritmo de Detecção

A literatura mostra que o teste generalizado da taxa de verossimilhança (TGTV) é a solução ótima para a detecção de distúrbios, desde que o ruído aditivo tenha distribuição Gaussiana (Trees, 1968/1971; Fishler & Messer, 1993; Colonnese & Scarano, 1999; Giannakis & Tsatsanis, 1990; McDonough & Whalen, 1995). Note que o TGTV consiste de um esquema de detecção e estimação, no qual a taxa de verossimilhança é calculada continuamente para cada amostra do sinal monitorado e o seu valor máximo é comparado com um valor limiar. O TGTV apresenta uma alta carga computacional, pois demanda altas taxas de amostragem do sinal para que a perda de desempenho, devida aos erros gerados por deslocamentos, seja evitada.

Para o processo dado por (8.1), tem-se que a solução ótima para a detecção é o processador quadrático (Trees, 1968/1971), definido pela taxa de verossimilhança (TV) dada por

$$TV = \sum_{k=0}^{K-1} \frac{c_k^2 d_k}{\sigma_v^2 + d_k} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (8.20)$$

onde K é o número de autovalores da matriz de autocorrelação de $\mathbf{u}(n) = [u(n) \cdots u(n - N + 1)]^T$, c_k é o coeficiente resultante do produto interno entre o processo observado e o k -ésimo autovetor da matriz de autocorrelação de $\mathbf{u}(n)$, d_k é o k -ésimo autovalor e η é o limiar para comparação. O valor η pode ser obtido usando os critérios de Bayes ou Neyman-Pearson (Trees, 1968/1971).

A detecção também pode ser vista como um problema de classificação de padrões e, portanto, várias técnicas de reconhecimento de padrões (Theodoridis & Koutroumbas, 1999), tais como os classificadores baseados na teoria de Bayes, os classificadores lineares e os classificadores não lineares, podem ser aplicados.

O fato dos distúrbios em sistemas de potência terem características não estacionárias, cujos parâmetros de modelagem são de difícil estimação, nos levou a optar pelo uso de um classificador não linear baseado na rede MLP para emular o critério de decisão dado por (8.20).

A Fig. 8.2 mostra o fluxograma da solução utilizada para a classificação de distúrbios. Primeiramente, a detecção da ocorrência de distúrbios é realizada pelo bloco *Detecção Ocorrência*. O bloco *Detecção Instantes* detecta os instantes inicial e final de ocorrência dos distúrbios baseando-se no seguinte procedimento:

Passo 1: gera-se as janelas de sinais dadas por $\mathbf{u}(n - i) = [u(n - i) \cdots u(n - i + N + 1)]^T$, $i = N - 1, \dots, 0, \dots, -N + 2$ e extrai-se os seus respectivos vetores de parâmetros;

Passo 2: o início do distúrbio é dado pelo primeiro vetor $\mathbf{u}(n - i)$ em que o algoritmo de detecção acusar a ocorrência do distúrbio;

Passo 3: o fim do distúrbio é dado pelo primeiro vetor $\mathbf{u}(n - j)$, $j > i$ em que o algoritmo de detecção não acusar a ocorrência de distúrbio.

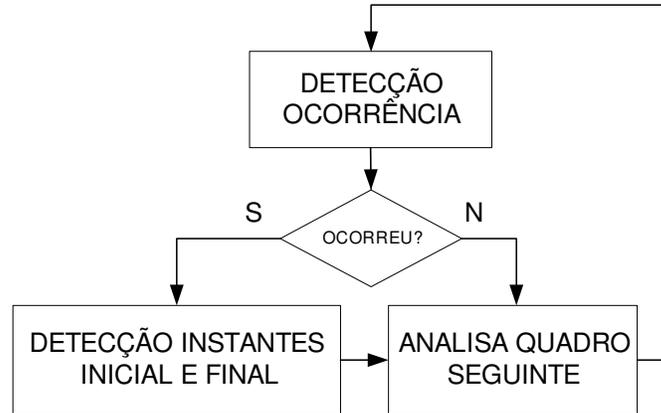


Figura 8.2: Fluxograma do algoritmo desenvolvido para a detecção de ocorrência dos distúrbios e de seus instantes inicial e final correspondentes.

Note que os blocos *Detecção Ocorrência* e *Detecção Instantes* fazem uso do vetor de parâmetros dado por (8.19) como entrada da rede MLP com estrutura $2 \times 2 \times 1$. A seção de resultados 8.3 justifica a escolha do vetor \mathbf{v}_4 como sendo o vetor de parâmetros adequados para detectar não só a ocorrência dos distúrbios, mas também os instantes inicial e final dos mesmos.

8.3 Resultados Numéricos

De forma a mostrar quais dos parâmetros do vetor $\mathbf{v}(n)$, vide (8.17), são adequados para a detecção de distúrbios, várias simulações computacionais foram realizadas e alguns dos resultados obtidos são aqui apresentados.

Os parâmetros γ_3 , γ_4 , $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ foram extraídos para diferentes valores de N , de forma a avaliar sua influência no processo de detecção de distúrbios. Os sinais escolhidos foram os seguintes: 1) um sinal transitório com duas variações de curta duração e $f_s = 192 \times 60$ amostras/s, 2) um sinal com transitórios de curta

duração e poluído com harmônicas cuja $f_s = 256 \times 60$ amostras/s e, finalmente, 3) um sinal sem distúrbio cuja $f_s = 256 \times 60$ amostras/s.

As Figs. 8.3(a)-8.3(b) mostram o sinal $\{x(n)\}$ e a sua componente correspondente $\{u(n)\}$, a qual é obtida na saída do bloco NF_0 . As Figs. 8.4 e 8.5 mostram o comportamento dos parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ e dos parâmetros γ_3 e γ_4 , quando N assume os valores 32, 64, 128 e 256. Esses parâmetros são representativos para caracterizar a ocorrência de distúrbios na seqüência analisada, desde que N , comprimento da seqüência, seja grande. De fato, uma análise das regiões de decisão referentes às duas hipóteses dadas por (8.1) para $N \rightarrow 0$, mostra que todos os parâmetros analisados tornam-se inexpressivos para caracterizar a ocorrência de distúrbios.

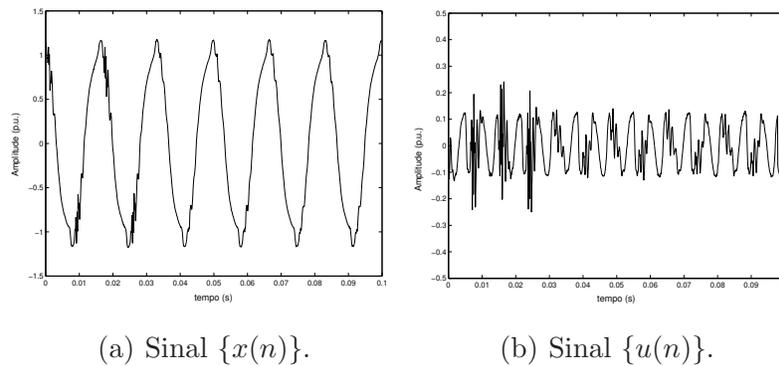


Figura 8.3: Transitório com duas variações de curta duração, $f_s = 192 \times 60$ amostras/s.

As Figs. 8.6(a) e 8.6(b) mostra o sinal $\{x(n)\}$ e a sua componente correspondente $\{u(n)\}$, a qual é obtida na saída do bloco NF_0 , enquanto as Figs. 8.7 e 8.8 mostram os parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ e os parâmetros γ_3 e γ_4 , para os valores de $N = 32, 64, 128, 256$, respectivamente. Note que durante a ocorrência de distúrbios, o sinal $\{u(n)\}$ apresenta EOS e, conseqüentemente, estas informações são relevantes para verificar a existência de distúrbios.

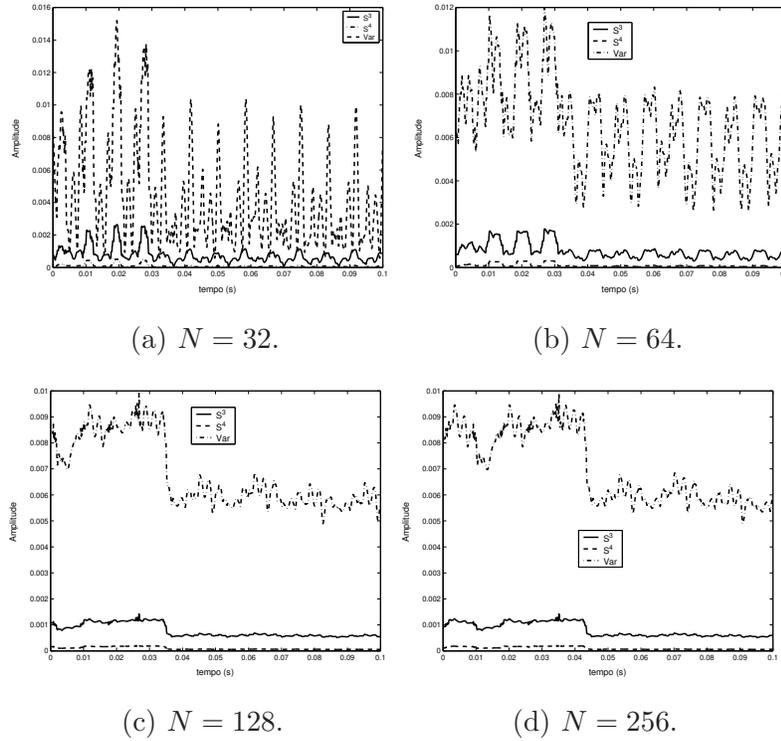


Figura 8.4: Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.3(b), para diferentes valores de N .

Finalmente, as Figs. 8.9(a) e 8.9(b) mostram o sinal $\{x(n)\}$ e a sua componente correspondente $\{u(n)\}$, a qual é obtida na saída do bloco NF_0 . Já as Figs. 8.10 e 8.11 mostram os parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ e os parâmetros γ_3 e γ_4 , para os valores de N iguais a 32, 64, 128 e 256.

Nesse contexto, a seguinte questão surge: Qual conjunto de parâmetros é representativo para caracterizar a ocorrência de distúrbios em sistemas de potência? Objetivando a responder esta questão, simulações foram conduzidas e os resultados desta simulação são apresentados na Tab. 8.1. Note que o parâmetro $Var(\mathbf{u}(n))$ não foi considerado, pois os resultados revelaram que o uso desse parâmetro gera baixo desempenho da técnica de detecção de distúrbios. Na Tab. 8.1, a primeira coluna

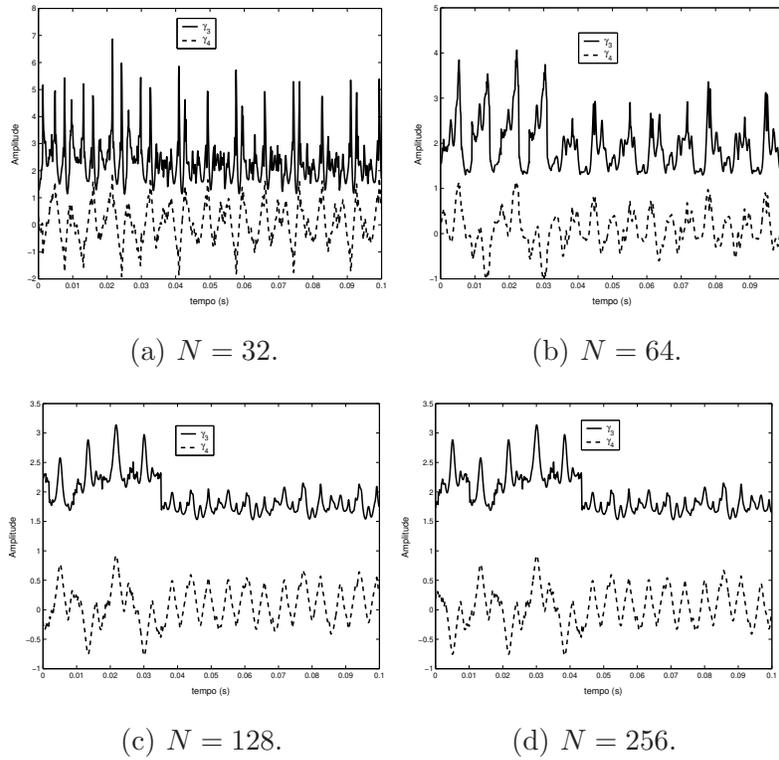


Figura 8.5: Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.3(b), para diferentes valores de N .

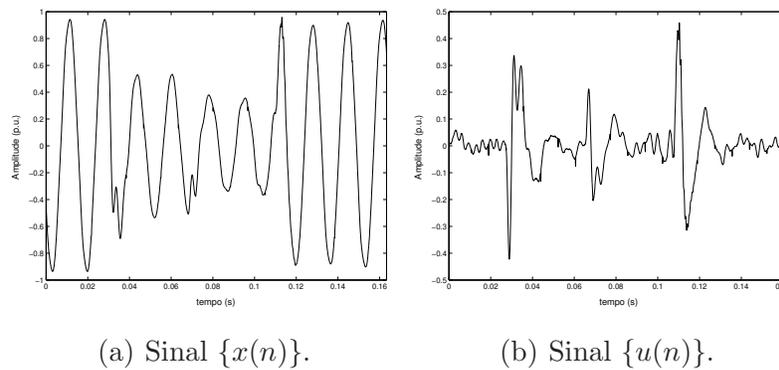


Figura 8.6: Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.

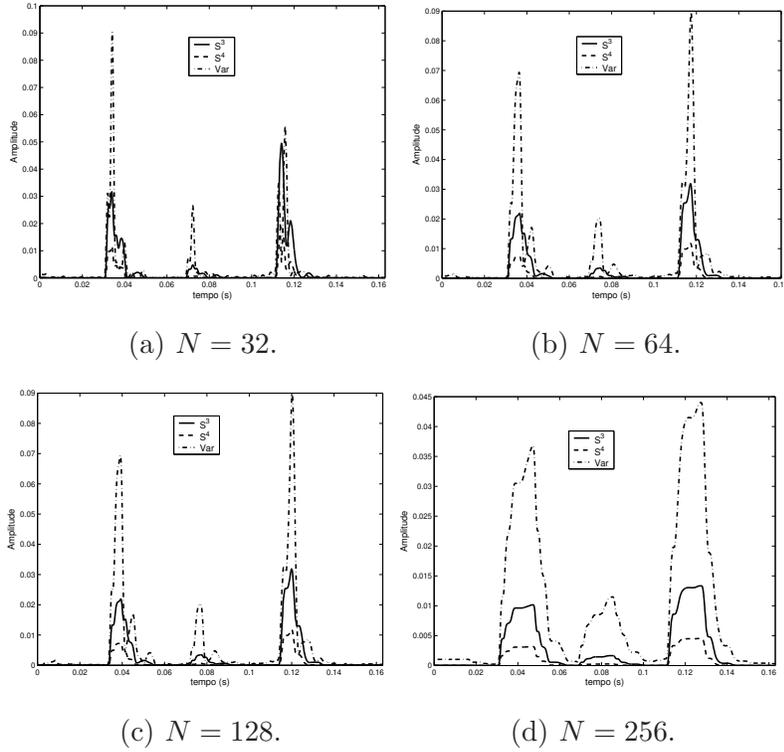


Figura 8.7: Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.6(b), para diferentes valores de N .

informa se os resultados são referentes aos dados de treinamento ou de teste assim como informa o número de ciclos da fundamental considerado para a extração de parâmetros e a detecção em bloco dos distúrbios. As colunas 2, 3, 4, 5 e 6 mostram os valores de detecção obtidos (em %) para diferentes conjuntos de parâmetros, \mathbf{v}_i , $i = 1, \dots, 5$, os quais são dados por

$$\mathbf{v}_1 = [\gamma_3 \ \gamma_4 \ S^{(3)} \ S^{(4)}]^T, \quad (8.21)$$

$$\mathbf{v}_2 = [S^{(3)} \ S^{(4)}]^T, \quad (8.22)$$

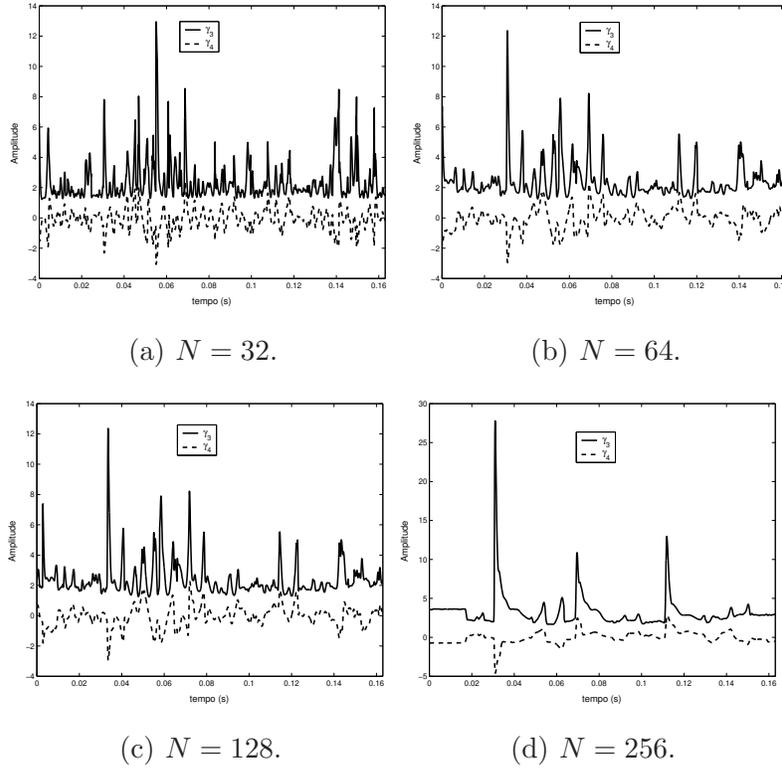


Figura 8.8: Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.6(b), para diferentes valores de N .

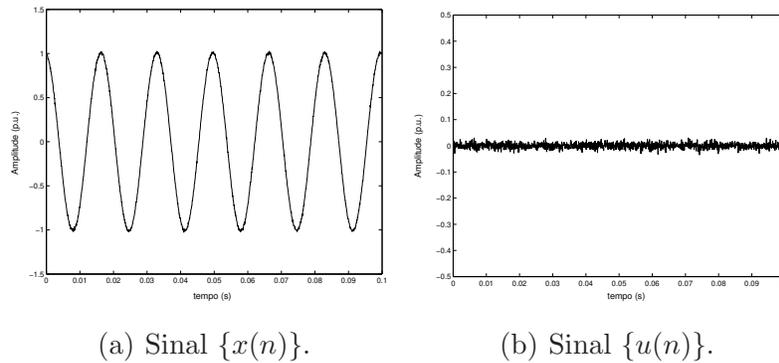


Figura 8.9: Sinal sem distúrbio, $f_s = 256 \times 60$ amostras/s.

$$\mathbf{v}_3 = [\gamma_3 \ \gamma_4]^T, \quad (8.23)$$

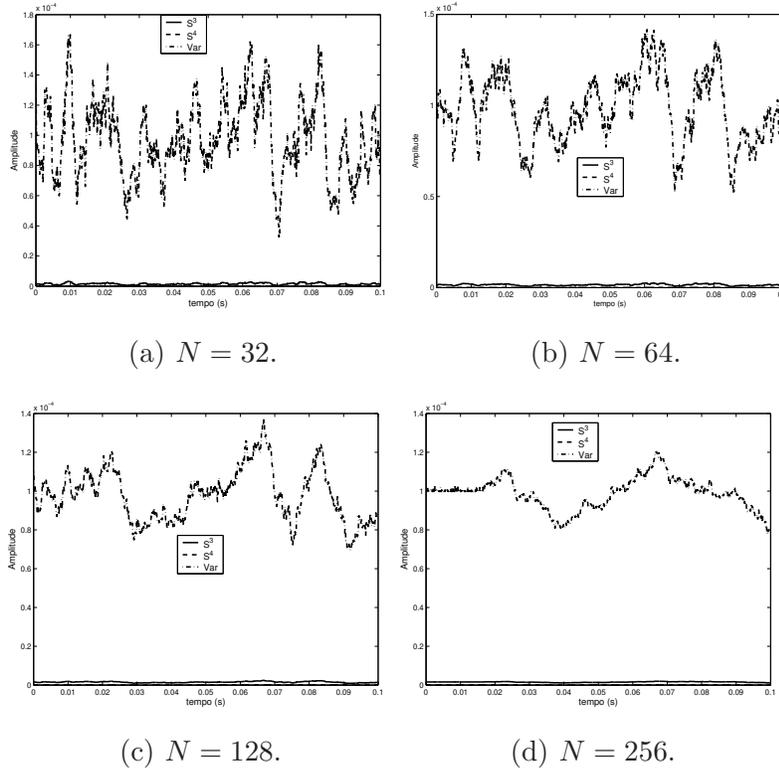


Figura 8.10: Parâmetros $S^{(3)}$, $S^{(4)}$ e $Var(\mathbf{u}(n))$ extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.9(b), para diferentes valores de N .

$$\mathbf{v}_4 = [\gamma_4 S^{(3)}]^T, \quad (8.24)$$

e

$$\mathbf{v}_5 = [\gamma_3 S^{(4)}]^T. \quad (8.25)$$

Para a obtenção dos resultados numéricos mostrados na Tab. 8.1, considerou-se o seguinte: *i*) 1610 quadros de sinais com e sem distúrbios, os quais representam *sags*, *swell*, *overvoltage*, *undervoltage*, transitório oscilatório amortecido de baixa e alta frequência, *notchs*, *spikes*, harmônicas e faltas, *ii*) os sinais com distúrbios são dados reais obtidos do IEEE, da CPFL e do grupo de pesquisa em QE da UFJF,

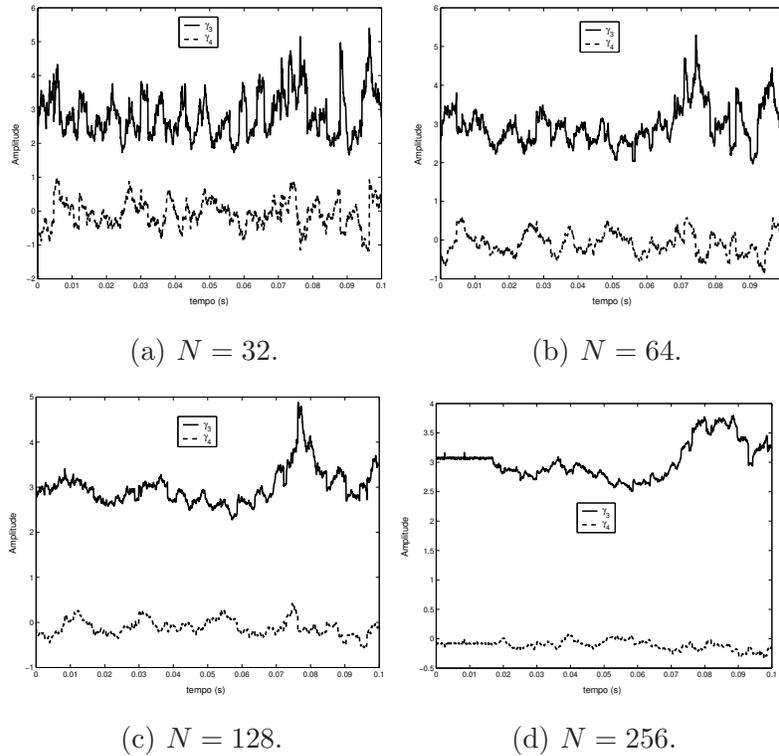


Figura 8.11: Parâmetros γ_3 e γ_4 extraídos do sinal $\{u(n)\}$, mostrado na Fig. 8.9(b), para diferentes valores de N .

iii) os dados foram igualmente divididos entre dados de treinamento e de teste, *iv)* a rede MLP utilizada possui a estrutura $p \times 2 \times 1$, onde $p = 2, 4$. Durante o treinamento 10000 épocas foram utilizadas. As redes MLP foram igualmente inicializadas, *v)* os parâmetros extraídos dos sinais foram normalizados, *vi)* duas taxas amostragem foram consideradas: $f_s = 60 \times 256$ amostras/s (dados CEMIG e IEEE) e $f_s = 60 \times 192$ amostras/s (dados CPFL) e *vii)* os sinais sem a ocorrência de distúrbios foram gerados sinteticamente com o MatLab.

A Tab. 8.1 mostra que a técnica proposta é capaz de detectar a ocorrência da maioria dos distúrbios, desde que 1 ciclo da fundamental seja considerado e \mathbf{v}_1 ou \mathbf{v}_4 seja considerado como o vetor de parâmetros na entrada da rede MLP. De fato,

as estimativas dos parâmetros γ_3 e γ_4 , $S^{(3)}$ e $S^{(4)}$ tendem aos valores reais de γ_3 e γ_4 , $S^{(3)}$ e $S^{(4)}$ desde que $N \rightarrow \infty$. No entanto, observou-se que para valores de $N > 200$, as estimativas dos parâmetros γ_3 e γ_4 , $S^{(3)}$ e $S^{(4)}$ são suficientes para garantir que a rede MLP implemente a expressão dada por (8.1). Na Tab. 8.1, a primeira coluna informa se o resultado é de treinamento ou de teste e o comprimento da seqüência considerada, enquanto as outras colunas informam o desempenho em termos de taxa de classificação (%) quando os vetores \mathbf{v}_i , $i = 1, \dots, 5$ são utilizados para a extração de parâmetros dos distúrbios monitorados.

<i>Fase/L</i>	\mathbf{v}_1	\mathbf{v}_2	\mathbf{v}_3	\mathbf{v}_4	\mathbf{v}_5
<i>Trein., 1C</i>	100,00	99,15	99,28	100,00	99,04
<i>Teste, 1C</i>	99,04	99,04	99,04	99,93	98,57
<i>Trein., 1/2C</i>	95,55	95,55	66,11	94,26	92,04
<i>Teste, 1/2C</i>	95,00	94,44	64,26	94,06	91,29
<i>Trein., 1/4C</i>	83,85	68,79	58,07	63,01	55,66
<i>Teste, 1/4C</i>	79,15	68,32	55,42	58,55	51,08
<i>Trein., 1/8C</i>	57,02	57,13	54,72	57,16	50,00
<i>Teste, 1/8C</i>	56,72	56,28	54,18	56,04	59,23

Tabela 8.1: Taxa em (%) de detecção, obtida com a técnica de detecção proposta, em função de diferentes conjuntos de parâmetros extraídos do sinal monitorado e de diferentes números de amostras.

De forma a ilustrar se a técnica proposta para a detecção de distúrbios é adequada para detectar os instantes inicial e final de ocorrência dos distúrbios, algumas simulações sobre os atrasos foram realizadas, pois os melhores resultados são obtidos para 1 ciclo da componente fundamental. As Figs. 8.12-8.14 mostram os sinais analisados e suas componentes correspondentes $\{u(n)\}$ e os instantes iniciais e finais de ocorrência dos distúrbios. A técnica de detecção proposta mostrou-se capaz de detectar com bastante precisão a ocorrência dos distúrbios. De fato, o

uso combinado dos parâmetros γ_4 e $S^{(3)}$ e da rede MLP é uma solução bastante adequada para detectar os instantes inicial e final de ocorrência dos distúrbios. Por outro lado, a Fig. 8.12 mostra que não houve nenhuma detecção de transitório no sinal.

Para obter os resultados mostrados nas Figs. 8.12-8.14, considerou-se que o vetor de parâmetros extraído do sinal é $\mathbf{v} = [\gamma_4 S^{(3)}]^T$ e 1 ciclo da componente fundamental.

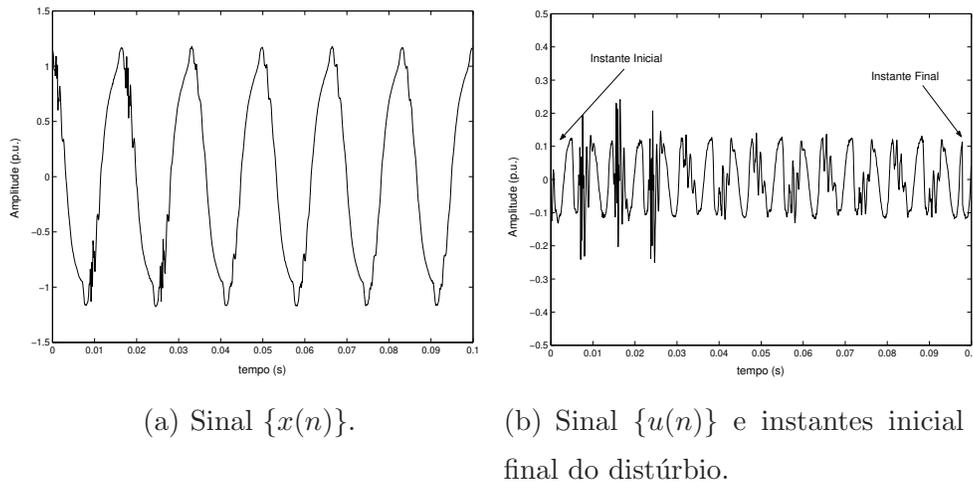


Figura 8.12: Intervalo de detecção de distúrbios para o sinal com harmônicas.

8.4 Sumário

O presente capítulo apresentou uma nova proposta para a detecção de distúrbios em sistemas de potência. Conforme argumentado ao longo das seções deste capítulo, o princípio dividir e conquistar, o conceito solução hierárquica baseada em dois níveis, o uso de estatísticas de ordem superior e de técnicas de inteligência computacional mostraram-se adequados para a detecção de distúrbios em sistemas

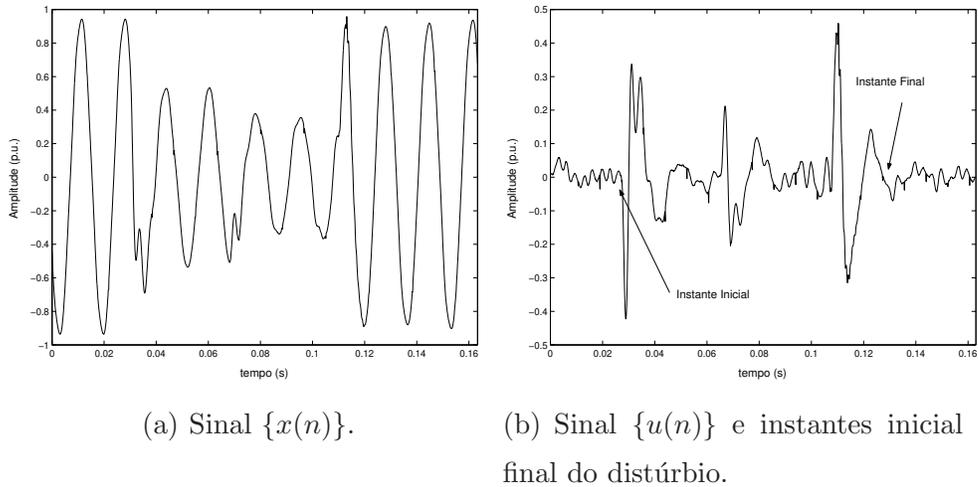


Figura 8.13: Intervalo de detecção de distúrbios para o sinal com duplo *sag*.

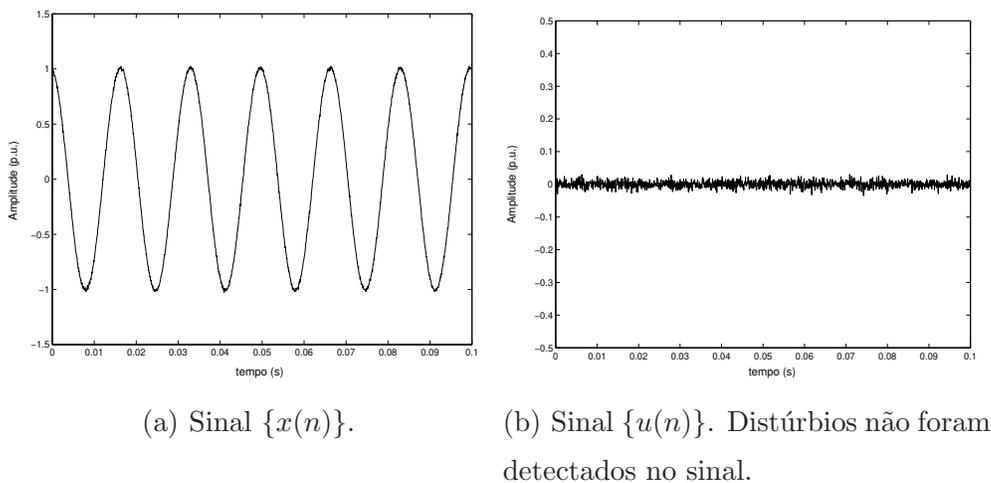


Figura 8.14: Sinal sem a ocorrência de distúrbios.

de potência, quando o tamanho da seqüência analisada é maior que 128 amostras ou 1/2 ciclo da fundamental.

9

Classificação de Distúrbios

O aumento do número de cargas de natureza não linear (sistemas microprocessados, fontes chaveadas, etc) em instalações residenciais, comerciais e industriais provoca o aumento significativo da ocorrência de distúrbios em sistemas de potência e o do número de falhas e, conseqüentemente, a diminuição da vida útil dos equipamentos conectados às redes elétricas. Além disso, destaca-se ainda a interligação dos sistemas de potência, o aumento do número de geradores alternativos de energia, o não compartilhamento de informações entre as empresas do setor elétrico, dentre outras (Arrillaga et al., 2000; Bollen, 2000; Dugan, McGranaghan, & Beaty, 1996), como contribuintes para a degradação da QE. Neste

contexto, a aplicação de técnicas de reconhecimento de padrões e de extração de parâmetros (Theodoridis & Koutroumbas, 1999) para o monitoramento da QE, com destaque para a classificação de distúrbios e a identificação de fontes geradoras de tais distúrbios são relevantes e têm, recentemente, recebido atenção de pesquisadores.

Dentre as principais técnicas de reconhecimento de padrões aplicadas em sistemas de potência, destacam-se os classificadores baseados na teoria de Bayes, os classificadores lineares, os classificadores não lineares (sistemas fuzzy, redes neurais e etc), *template matching* e os classificadores dependentes do contexto (Theodoridis & Koutroumbas, 1999; Mendel, 2001; Príncipe et al., 2000; Haykin, 1999; Pedrycz & Gomide, 1998).

Para as aplicações de monitoramento da QE, a classificação de distúrbios pode ser decomposta em duas tarefas distintas: *i*) a extração de um conjunto reduzido e representativo de parâmetros dos distúrbios monitorados que minimize as intersecções entre as regiões de classificações associadas aos diferentes distúrbios e *ii*) o uso de técnicas adequadas para a classificação dos distúrbios a partir do vetor de parâmetros extraídos do sinal monitorado. A combinação ótima de técnicas de extração de parâmetros e de reconhecimento de padrões garantem resultados significativos para a classificação de distúrbios.

O uso combinado das técnicas de reconhecimento de padrões e de extração de parâmetros para a classificação de distúrbios em redes elétricas foi investigado nas seguintes contribuições: (Zhu, Tao, & Lo, 2004; Wang & Marnishev, 2004; Azam, Tu, Pattipati, & Karanam, 2004; Wang, Rowe, & Marnishev, 2004; Gaing, 2004; Youssef, Abdel-Galil, El-Saadany, & Salama, 2004; Chung, Powers, Grady, & Bhatt, 2004; Abdel-Galil, Kamel, Youssef, El-Saadany, & Salama, 2004; Ibrahim & Morcos, 2003; Lee & Dash, 2003; Dash, Panigrahi, Sahoo, & Panda, 2003; Huang, Negnevitsky, & Nguyen, 2002; Gaouda, Kanoun, Salama, & Chikhani, 2002; Kezunovic & Liao,

2002; Lee, Won, Jeong, & Nam, 2002; Hoang & Nguyen, 2002; Wijayakulasooriya, Putrus, & Minns, 2002; Styvaktakis, Bollen, & Gu, 2002; Santoso, Grady, Powers, Lamooore, & Bhatt, 2000; Kezunovic & Liao, 2001; Dash, Jena, & Salama, 1999; Gaouda et al., 1999; Huang, Hsieh, & Huang, 1998; Lee & Nam, 1998; Lee, Lee, Kim, & Nam, 1997; Ghosh & Lubkeman, 1995). Nesses trabalhos, as taxas de classificação obtidas para a classificação de eventos isolados é entre 90 % e 98.5 %.

Conforme já ressaltado, nos sistemas de potência existem as ocorrências de vários distúrbios, provocados por diferentes fontes. No entanto, todas as técnicas citadas acima foram desenvolvidas para classificar apenas distúrbios isolados. Os sinais mostrados nas Figs. 9.1-9.3 ilustram muito bem esse problema. A Fig. 9.1(a) mostra o sinal $\{x(n)\}$, enquanto as Figs. 9.1(b)-9.1(c) mostram os sinais $\{f(n)\}$ e $\{u(n)\}$, os quais foram obtidos após a extração da componente fundamental do sinal $\{x(n)\}$. Note que o sinal $\{u(n)\}$ é constituído da 3^a harmônica e de alguns transitórios caracterizados por chaveamentos capacitivos entre outros transitórios, a princípio, não identificáveis. As Figs. 9.2(a)-9.2(c) mostram os sinais $\{x(n)\}$, $\{f(n)\}$ e $\{u(n)\}$ de um outro distúrbio observado nas redes elétricas, com características similares ao distúrbio mostrado na Fig. 9.1. A Fig. 9.2(c) mostra a ocorrência de vários distúrbios (harmônicas e transitórios oscilatórios e amortecidos). A Fig. 9.3 mostra uma variação de tensão de curta duração e a ocorrência de alguns distúrbios. Note que após a ocorrência de dois transitórios oscilatórios e amortecidos, um novo distúrbio na tensão aparece. Analisando-se outros distúrbios em sistemas de potência, pode-se constatar que, normalmente, existe a ocorrência simultânea de vários distúrbios. No entanto, as técnicas desenvolvidas, até o presente momento, centram-se na classificação de eventos isolados.

O uso das técnicas listadas anteriormente poderá gerar resultados contraditórios e não representativos sobre os tipos de distúrbios que ocorreram nas redes elétricas monitoradas e, conseqüentemente, o uso de uma solução para a identificação da

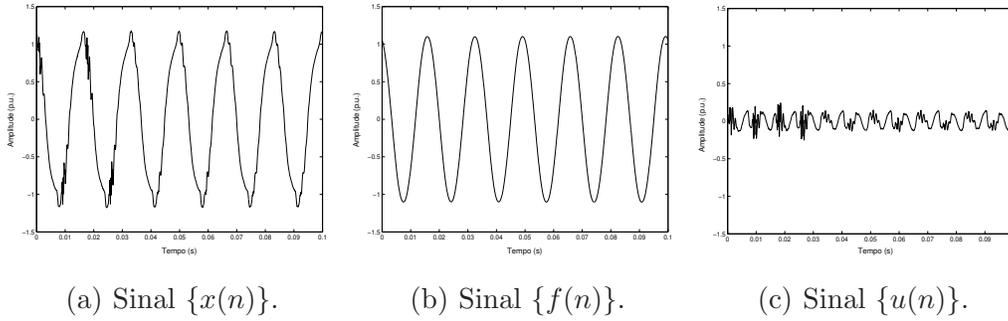


Figura 9.1: Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.

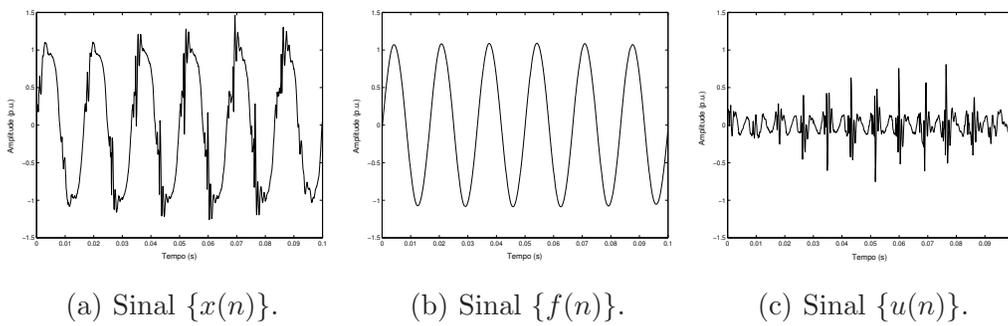


Figura 9.2: Sinal com transitórios de curta duração e poluído com harmônicas, $f_s = 256 \times 60$ amostras/s.

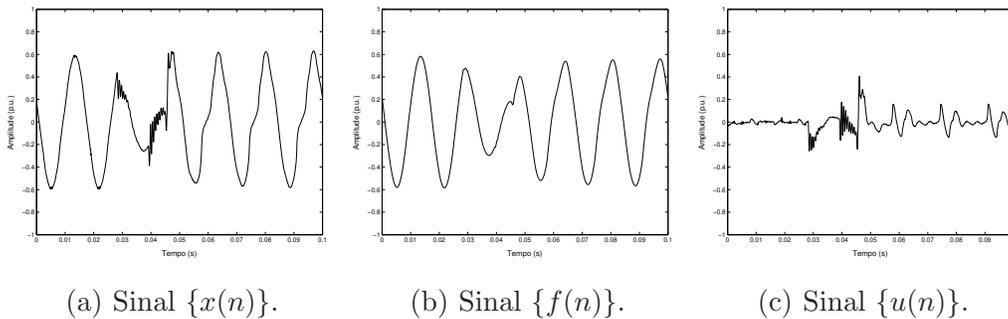


Figura 9.3: Sinal com sub-tensão e com transitórios de curta duração, $f_s = 256 \times 60$ amostras/s.

fonte geradora de distúrbio, tal como proposto em (Azam et al., 2004), é bastante comprometido. De fato, a classificação de distúrbios do sinal $\{x(n)\}$ constituído de vários distúrbios é um problema ainda em aberto e, portanto, demanda novas investigações para que a caracterização destes distúrbios seja a mais correta possível. Isso resultará em maior confiabilidade na identificação das fontes geradoras de distúrbios, dentre outras coisas.

Objetivando propor uma técnica capaz de não apenas classificar a ocorrência isolada de distúrbios, mas também a ocorrência conjunta dos mesmos, a seção 9.1 apresenta o classificador proposto para classificar os distúrbios que ocorram isoladamente ou conjuntamente no sinal $\{x(n)\}$, a seção 9.2 discute as técnicas de extração de parâmetros utilizadas, a seção 9.3 apresenta a estratégia de classificação e, finalmente, a seção 9.4 apresenta os resultados obtidos com a técnica proposta.

9.1 Classificador Proposto

Para descrever o classificador de distúrbios, as seguintes definições são necessárias: o vetor $\mathbf{x}(n) = [x(n) \ \cdots \ x(n - N + 1)]^T$ é constituído de amostras do sinal dado por (6.1), o vetor $\mathbf{f}(n) = [f(n) \ \cdots \ f(n - N + 1)]^T$ é constituído de estimativas do sinal dado por (6.2), o vetor $\mathbf{h}(n) = [h(n) \ \cdots \ h(n - N + 1)]^T$ é formado por estimativas do sinal dado por (6.3) e o vetor $\mathbf{u}(n) = \mathbf{i}(n) + \mathbf{t}(n) + \mathbf{v}(n) = [u(n) \ \cdots \ u(n - N + 1)]^T$ é constituído por estimativas dos sinais expressos por (6.4) e (6.5) e pelo ruído de fundo.

Pode-se afirmar que as técnicas de classificação de distúrbios encontradas na literatura seguem a estrutura mostrada na Fig. 9.4, ou seja, realizam a classificação dos distúrbios a partir da extração de parâmetros diretamente do sinal $\{x(n)\}$. A estrutura para a classificação dos distúrbios ilustrada pelo diagrama de bloco



Figura 9.4: Diagrama de bloco do classificador padrão de distúrbios.

mostrado na Fig. 9.4 é denominada de estrutura padrão e \mathbf{p}_x é o vetor de parâmetros extraído do vetor $\mathbf{x}(n)$. A extração de parâmetros representativos diretamente do sinal $\{x(n)\}$ para fins de classificação é bastante prejudicada, pois o mesmo é constituído de várias componentes, tais como fundamental, harmônicas, inter-harmônicas e transitórios. Conseqüentemente, não apenas a classificação de um único transitório torna-se uma tarefa difícil, mas também a classificação de múltiplos transitórios é prejudicada, conforme é apontado em (Wang & Marnishev, 2004; Wang et al., 2004). De fato, as regiões de classificação geradas a partir dos vetores de parâmetros \mathbf{p}_x extraídos dos sinais $\{x(n)\}$ são de difícil determinação, uma vez que existem sobreposições entre as regiões que definem a ocorrência isolada ou conjunta dos vários distúrbios observados no sinal $\{x(n)\}$. Tal problema resulta em baixo desempenho do classificador tanto para a classificação de um único distúrbio quanto para a classificação de vários distúrbios (vide a lista de trabalhos apresentados anteriormente).

Objetivando minimizar esse problema, a Fig. 9.5 mostra o diagrama de bloco do classificador proposto, baseado no princípio dividir e conquistar, para a classificação de distúrbios isolados ou de vários distúrbios no sinal $\{x(n)\}$. Os blocos Extração $\mathbf{f}(n)$, Extração $\mathbf{h}(n)$ e Extração $\mathbf{u}(n)$ implementam os algoritmos que decompõem o sinal $\mathbf{x}(n)$ nas componentes $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$, respectivamente. As componentes $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$ são obtidas com a técnica NF-WDFT, discutida no capítulo 7. Usando-se a versão I da técnica NF-WDFT, tem-se

$$\mathbf{f}(n) = \widehat{\mathbf{f}}(n), \quad (9.1)$$

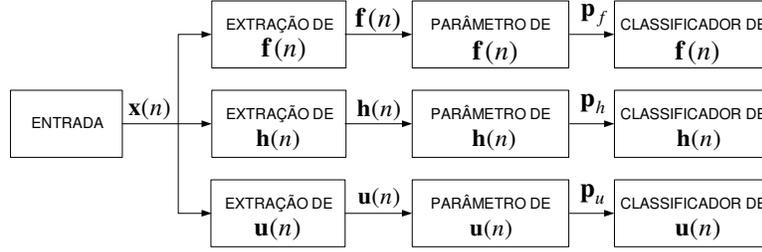


Figura 9.5: Diagrama de bloco da técnica proposta para a classificação de múltiplos distúrbios em sistemas de potência.

$$\mathbf{h}(n) = \sum_{m=0}^M \widehat{\mathbf{h}}_m(n), \quad (9.2)$$

e

$$\mathbf{u}(n) = \mathbf{x}_M(n), \quad (9.3)$$

onde $\widehat{\mathbf{f}}(n) = [\widehat{f}(n) \ \cdots \ \widehat{f}(n - N + 1)]^T$ é o vetor constituído de amostras obtidas na entrada do bloco WDFT_0 da técnica NF-WDFT I, $\widehat{\mathbf{h}}_m(n) = [\widehat{h}_m(n) \ \cdots \ \widehat{h}_m(n - N + 1)]^T$ é o vetor constituído de amostras obtidas na entrada do bloco WDFT_m da técnica NF-WDFT I e $\mathbf{x}_M(n) = [x_M(n) \ \cdots \ x_M(n - N + 1)]^T$ é o vetor constituído das amostras geradas na saída do bloco NF_M da técnica NF-WDFT I (vide capítulo 7).

A seguir, os algoritmos de processamento de sinais, implementados nos blocos Parâmetros de $\mathbf{f}(n)$, Parâmetros de $\mathbf{h}(n)$ e Parâmetros de $\mathbf{u}(n)$, extraem parâmetros representativos dos sinais $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$. Finalmente, os blocos Classificador de $\mathbf{f}(n)$, Classificador de $\mathbf{h}(n)$ e Classificador de $\mathbf{u}(n)$ aplicam os algoritmos que classificam os distúrbios relacionados aos sinais $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$. Na Fig. 9.5, \mathbf{p}_f , \mathbf{p}_h e \mathbf{p}_u são os vetores de parâmetros extraídos dos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$, respectivamente.

Uma vez que os vetores de parâmetros \mathbf{p}_f , \mathbf{p}_h e \mathbf{p}_u tenham sido extraídos dos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$, os blocos Classificador de $\mathbf{j}(n)$, $j = f, h, u$ fazem uso do

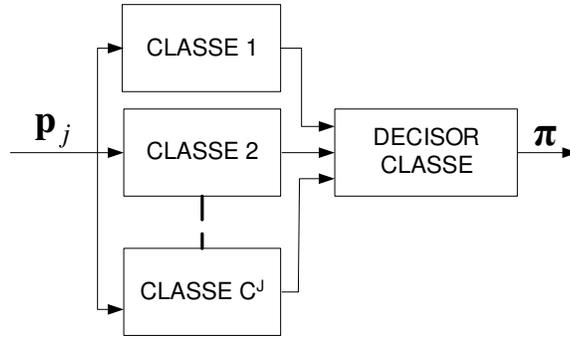


Figura 9.6: Diagrama de bloco do classificador $\mathbf{j}(n)$, $j = f, h, u$.

esquema mostrado na Fig. 9.6 para detectar a ocorrência dos distúrbios associados ao sinal $\mathbf{j}(n)$, $j = f, h, u$. No esquema mostrado na Fig. 9.6, os blocos Classe i , $i = 1, \dots, C^j$ implementam os algoritmos que classificam individualmente os distúrbios e, finalmente, o bloco Decisor Classe implementa o algoritmo para realizar a decisão final sobre a ocorrência das diferentes classes de distúrbios no sinal $\{x(n)\}$. Observe que os algoritmos de classificação alocado para cada tipo de distúrbio implementa a expressão dada por

$$\begin{aligned} H_0 : \mathbf{j}(n) &= \mathbf{v}_{c^j}(n) \\ H_1 : \mathbf{j}(n) &= \mathbf{j}_{c^j}(n), \quad c^j = 1, \dots, C^j \end{aligned} \quad (9.4)$$

onde $\mathbf{v}_{c^j}(n)$ representa o distúrbio que não pertence à classe c^j enquanto $\mathbf{j}_{c^j}(n)$ é o distúrbio associado à classe c^j .

A separação do sinal $\{x(n)\}$ nos sinais $\{f(n)\}$, $\{h(n)\}$ e $\{u(n)\}$, promovida pela técnica de classificação proposta, é justificável pelos seguintes argumentos:

i) os distúrbios na componente fundamental $\{f(n)\}$ são relacionados intrinsecamente aos parâmetros $\theta_0(n)$, $\omega_0(n)$ e $A_0(n)$. Analisando-se o comportamento de $A_0(n)$ pode-se detectar a ocorrência dos seguintes eventos: 1) *sag*, 2) *swell*, 3) *undervoltage*, 4) *overvoltage*, 5) interrupção e 6) normal (Bollen, 2000; Dugan et al., 1996). As características destes distúrbios são listadas na Tab.

9.1. Na Tab. 9.1, $T \leq \Delta T_0 < 9T$ e $9T \leq \Delta T_1$, sendo $T = 16,667$ ms. Note que os distúrbios citados acima ocorrem apenas no sinal $\{f(n)\} = f(t)|_{t=nT_s}$, $n = 0, 1, 2, \dots$

Tipo	Parâmetro de Controle	Equação
Interrupção	$0, 9 \leq \alpha < 1, 0$ e ΔT_0	$f(t) = A_0(1 - \alpha)\sin(\omega_0 t)$
<i>Sag</i>	$0, 1 \leq \alpha < 0, 9$ e ΔT_0	$f(t) = A_0(1 - \alpha)\sin(\omega_0 t)$
Undervoltage	$0, 1 \leq \alpha < 0, 9$ e ΔT_1	$f(t) = A_0(1 - \alpha)\sin(\omega_0 t)$
Normal	$0, 9 \leq \alpha < 1, 1$	$f(t) = A_0(1 - \alpha)\sin(\omega_0 t)$
<i>Swell</i>	$0, 1 \leq \alpha < 0, 8$ e ΔT_0	$f(t) = A_0(1 + \alpha)\sin(\omega_0 t)$
<i>Overvoltage</i>	$0, 1 \leq \alpha < 0, 8$ e ΔT_1	$f(t) = A_0(1 + \alpha)\sin(\omega_0 t)$

Tabela 9.1: Tipos de distúrbios relacionados com o sinal $\{f(n)\}$.

ii) A análise do sinal $\{h(n)\}$ para detectar a ocorrência das componentes harmônicas é direta. Para tal detecção deve-se considerar (Duque et al., 2005)

$$\mathbf{x}(n) = \mathbf{f}(n) + \mathbf{u}(n), \quad (9.5)$$

$$\mathbf{u}(n) = \mathbf{h}(n) + \mathbf{v}(n), \quad (9.6)$$

e a DHT ser dada por

$$DHT = \sqrt{\frac{\sum_{i=1}^M |h_i|^2}{|h_0|^2}}, \quad (9.7)$$

onde h_i , $i = 0, 1, \dots, M$, é o valor RMS (*root mean square*) das componentes fundamental e harmônicas. Então

$$h_{rms} = \sqrt{\sum_{i=0}^M |h_i|^2} = h_0 \sqrt{1 + DHT^2} \quad (9.8)$$

ou

$$h_{rms} = \sqrt{a_0 \cdots a_{N-1}}, \quad (9.9)$$

onde $a_0 = |f(n)|^2 + |h(n)|^2$ e $a_{N-1} = |f(n - N + 1)|^2 + |h(n - N + 1)|^2$. Agora, suponha que $\sigma_v^2 \ll \|\mathbf{h}(n)\|^2$, então

$$\|\mathbf{h}(n)\|^2 = \sum_{i=1}^M h_i^2 \approx \|\mathbf{u}(n)\|^2, \quad (9.10)$$

onde $\|\mathbf{h}(n)\|^2 = |h(n)|^2 + \dots + |h(n - N + 1)|^2$. Logo, a estimativa instantânea da DHT é dada por

$$DHT(n) = \sqrt{\frac{2\|\mathbf{h}(n)\|^2}{A_0^2(n)}}, \quad (9.11)$$

onde $A_0^2(n)$ é a estimativa instantânea da amplitude da componente fundamental e $\|\mathbf{h}(n)\|^2$ é dado por (9.10). Finalmente, a DHT pode ser obtida a partir de

$$DHT = E\{DHT(n)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} DHT(n). \quad (9.12)$$

Note que a DHT calculada com a expressão (9.12) não necessita do cálculo das amplitudes das componentes harmônicas. Para o caso em que a técnica NF-WDFFT é utilizada para estimar os parâmetros das componentes fundamental e harmônicas, o cálculo da DHT é direto. A estratégia para identificar a ocorrência de harmônicas é a seguinte:

- Passo 1: *Se* $DHT > \mathcal{K}$, então existe a presença de harmônicas no vetor $\mathbf{x}(n)$.
- Passo 2: *Caso Contrário*, a presença de harmônicas no vetor $\mathbf{x}(n)$ não é detectada.

A estratégia acima é válida desde que $\sigma_v^2 = 0$. Essa restrição torna tal estratégia não confiável em aplicações reais onde $\sigma_v^2 \neq 0$.

iii) no sinal $\{u(n)\}$, observa-se a ocorrência de transitórios oscilatórios amortecidos de baixa e alta frequências, *notchs*, *spikes* e transitórios devidos às mudanças entre os estados da rede (ex. transição entre as condições normal e *sag*). Desta forma, o classificador aplicado ao sinal $\{u(n)\}$ deverá ser capaz de detectar a ocorrência desses transitórios.

A discussão sobre a extração de parâmetros dos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$ para fins de classificação de distúrbios é dada na seção 9.2.

9.2 Técnica para a Extração de Parâmetros

A extração de um conjunto representativo de parâmetros para fins de classificação é uma questão bastante complexa, visto que a mesma é dependente da aplicação alvo. Tal questão, normalmente, é analisada conjuntamente com a seleção de um classificador adequado para a aplicação.

Dado o grande número de técnicas disponíveis para a extração de parâmetros dos sinais, o número de parâmetros disponíveis para o sistema de classificação é bastante grande. No entanto, no projeto de um classificador, geralmente, um número reduzido de parâmetros é considerado por várias razões, dentre as quais, destacam-se:

i) a complexidade computacional devido às limitações de custo computacional e de processamento em tempo real. Essa questão pode ser exemplificada da seguinte maneira: embora dois parâmetros, quando analisados separadamente, possam conter informações relevantes do sinal para fins de classificação, o mesmo não se aplica a uma análise conjunta dos mesmos se houver um valor elevado de correlação mútua entre eles. Isso significa que o uso de um conjunto grande de parâmetros oferece reduzidos ganhos ou pode até piorar o desempenho e, além disso, demanda maior complexidade computacional;

ii) a propriedade de generalização do classificador. De fato, se $K \gg L$, onde K é o número de vetores de treinamento e L é o número de parâmetros do classificador, então a capacidade de generalização do classificador será melhor. Note que se $P \rightarrow \infty$, onde P é o número de parâmetros do vetor de treinamento, então, necessariamente, $L \rightarrow \infty$. Como, geralmente, o número dos vetores de treinamento é finito, uma boa prática é utilizar o valor mínimo possível para P e, conseqüentemente, classificadores com boa capacidade de generalização são obtidos. De fato, sabe-se que classificadores com grande número de parâmetros tende a se

adaptar apenas aos detalhes específicos dos conjuntos de dados de treinamento.

Dadas as informações acima, surge a seguinte questão: Como selecionar parâmetros representativos dos sinais $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$, considerando que P seja mínimo e os P parâmetros sejam capazes de garantir a intersecção mínima entre classes diferentes?

Diferentes técnicas foram propostas para responder a essa pergunta. Dentre as quais pode-se citar, o teste estatístico de hipóteses, as medidas de separação de classes, as matrizes *scatter*, *Fisher's discriminat ratio*, a transformação linear, etc (Theodoridis & Koutroumbas, 1999; Devroye, Györfi, & Lugosi, 1996).

Recentemente, em aplicações de classificação de distúrbios para o monitoramento da QE, a transformação linear (transformadas Wavelet, Wavelet contínua, Wavelet packet, *Short-Time Fourier Transform*, DFT, Wavelet, Lapped, S, etc) (Bagchi & Mitra, 1999; Vitterli & Kovacevic, 1995; Malvar, 1992; Mallat, 2001; Burrus, Gopinath, & Guo, 1997; Daubechies, 1992) tem sido largamente empregada, vide (Zhu et al., 2004; Azam et al., 2004; Gaing, 2004; Youssef et al., 2004; Chung et al., 2004; Abdel-Galil et al., 2004; Lee & Dash, 2003; Dash et al., 2003; Huang et al., 2002; Gaouda et al., 2002; Kezunovic & Liao, 2002; Lee et al., 2002; Hoang & Nguyen, 2002; Wijayakulasooriya et al., 2002; Santoso et al., 2000; Dash et al., 1999; Gaouda et al., 1999; Huang et al., 1998; Lee & Nam, 1998; Lee et al., 1997; Ghosh & Lubkeman, 1995). De fato, o uso de transformadas lineares tem se mostrado uma ferramenta poderosa para a extração de parâmetros representativos do sinal $\{x(n)\}$. No entanto, sabe-se que o uso de tais transformadas lineares demandam um custo computacional, além do custo computacional requerido para a extração de parâmetros no domínio da transformada.

Diferentemente das técnicas acima citadas, que classificam os distúrbios a partir do vetor $\mathbf{x}(n)$, a técnica de classificação proposta (Ribeiro & Romano, 2005b) faz

uso de diferentes técnicas de processamento para a separação do vetor $\mathbf{x}(n)$ em três vetores: $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$. Conforme discutido na seção 9.2, os vetores $\mathbf{f}(n)$ e $\mathbf{h}(n)$ são constituídos de sinais senoidais e a classificação de distúrbios nos mesmos é em função da potência dessas componentes senoidais. No entanto, o uso do valor RMS do vetor $\mathbf{f}(n)$ não é efetivo para caracterizar os distúrbios relacionados à componente fundamental que ocorrem no sinal $\{x(n)\}$, visto que o valor RMS não é imune à presença de ruído de fundo uma vez que apenas lida com estatísticas de ordem 2. Isso sugere que o uso de parâmetros baseados em estatísticas de ordem superior (EOS) oferece vantagens significativas no que tange a classificação dos distúrbios no vetor $\mathbf{f}(n)$. Por exemplo, os cumulantes de ordem superior são imunes à presença de ruído Gaussiano, são capazes de fornecer informações representativas do sinal analisado e, conseqüentemente, facilitam o processo de classificação dos distúrbios associados ao vetor $\mathbf{f}(n)$, de acordo com a Tab. 9.1. As EOS escolhidas na presente tese para a classificação de distúrbios nos vetores $\mathbf{f}(n)$ são a *Skewness* normalizada, *Kurtosis* normalizada e a estimativa da amplitude, \hat{A}_0 , da componente fundamental. Então, o vetor de parâmetros extraídos é dado por $\mathbf{p}_f = [\hat{A}_0 \ \bar{C}_3^f(0, 0) \ \bar{C}_4^f(0, 0, 0)]^T$, onde \hat{A}_0 é a estimativa da amplitude da componente fundamental fornecida pela técnica NF-WDFFT I, introduzida no capítulo 7,

$$\bar{C}_3^f(0, 0) = \frac{C_3^f(0, 0)}{[C_2^f(0)]^{\frac{3}{2}}} = \frac{E\{f^3(n)\}}{[E\{f^2(n)\}]^{\frac{3}{2}}} \quad (9.13)$$

é a *Skewness* do sinal $\{f(n)\}$ e

$$\bar{C}_4^f(0, 0, 0) = \frac{C_4^f(0, 0, 0)}{[C_2^f(0)]^2} = \frac{E\{f^4(n)\}}{[E\{f^2(n)\}]^2} - 3, \quad (9.14)$$

é o *kurtosis* do do sinal $\{f(n)\}$ (Mendel, 1991; Nikias & Petropulu, 1993). Note que

$$C_2^f(0) = E\{f^2(n)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} f^2(n) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{A}_0^2(n) \cos^2(\omega_0(n) + \theta_0(n)). \quad (9.15)$$

$$C_3^f(0,0) = E\{f^3(n)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} f^3(n) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{A}_0^3(n) \cos^3(\omega_0(n) + \theta_0(n)). \quad (9.16)$$

$$C_4^f(0,0,0) = E\{f^4(n)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} f^4(n) = \frac{1}{N} \sum_{n=0}^{N-1} \hat{A}_0^4(n) \cos^4(\omega_0(n) + \theta_0(n)). \quad (9.17)$$

Nas expressões (9.13)-(9.17) foi considerado que $E\{f(n)\} = 0$ e $E\{. \}$ é o operador esperança.

Para a classificação de distúrbios no vetor $\mathbf{h}(n)$, pode-se adotar o procedimento proposto em (Duque et al., 2005). No entanto, é sabido que as componentes harmônicas podem ter seus parâmetros (amplitude, frequência e fase) modelados como variáveis aleatórias. Assim sendo, verificar a ocorrência de componentes harmônicas baseando-se em momentos de ordem 2 é uma solução que não leva em consideração informações representativas do vetor $\mathbf{h}(n)$ e, conseqüentemente, perda de desempenho no processo de detecção de harmônicas em sistema de potência é prejudicado. Neste contexto, nesta seção propomos a extração de parâmetros baseados em EOS, os quais fornecem informações qualitativas do distúrbio e permitem que a detecção da ocorrência de harmônicas no sinal $\mathbf{h}(n)$ seja confiável a partir de 1/2 ciclo da componente fundamental, vide a seção de resultados. Para a detecção da presença de harmônicas no vetor $\mathbf{h}(n)$, o vetor de parâmetros $\mathbf{p}_h = [\bar{C}_3^h(0,0) \ \bar{C}_4^h(0,0,0)]^T$ é extraído. Note que

$$C_2^h(0) = E\{h^2(n)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} h^2(n) = \frac{1}{N} \sum_{n=0}^{N-1} \left[\sum_{m=1}^M \hat{A}_m(n) \cos(\omega_m(n) + \theta_m(n)) \right]^2, \quad (9.18)$$

$$\bar{C}_3^h(0, 0) = \frac{C_3^h(0, 0)}{[C_2^h(0)]^{\frac{3}{2}}} = \frac{E\{h^3(n)\}}{[E\{h^2(n)\}]^{\frac{3}{2}}} \approx \frac{\sum_{n=0}^{N-1} h^3(n)}{[E\{h^2(n)\}]^{\frac{3}{2}}} = \frac{\frac{1}{N} \sum_{n=0}^{N-1} [\sum_{m=1}^M \hat{A}_m(n) \cos(\omega_m(n) + \theta_m(n))]^3}{[C_2^h(0)]^{\frac{3}{2}}} \quad (9.19)$$

e

$$\bar{C}_4^h(0, 0, 0) = \frac{C_4^h(0, 0, 0)}{[C_2^h(0)]^2} = \frac{E\{h^4(n)\}}{[E\{h^2(n)\}]^2} - 3 \approx \frac{\frac{1}{N} \sum_{n=0}^{N-1} [\sum_{m=1}^M \hat{A}_m(n) \cos(\omega_m(n) + \theta_m(n))]^4}{[C_2^h(0)]^2} - 3. \quad (9.20)$$

Nas expressões (9.18)-(9.20) foi considerado que $E\{h(n)\} = 0$ e $E\{.\}$ é o operador esperança. Observe ainda que $C_2^h(0)$, $C_3^h(0, 0)$, $C_4^h(0, 0, 0)$ são os cumulantes de ordem 2, 3 e 4 quando os atrasos são iguais a zero, vide (Mendel, 1991; Nikias & Petropulu, 1993).

Dentre as componentes $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$ do vetor $\mathbf{x}(n)$, a componente $\mathbf{u}(n)$ é a que reúne diferentes classes de distúrbios, tais como notches, spikes, transitórios oscilatórios amortecidos, etc. Tais transitórios ocorrem de forma aleatória na rede elétrica. Neste contexto, a extração de parâmetros, baseados em EOS, do vetor $\mathbf{u}(n)$ para fins de classificação apresenta-se como bastante pertinente e promissora abordagem, pois as EOS permitem extrair informações representativas dos distúrbios analisados e, conseqüentemente, oferecem melhores condições para que o processo de classificação dos distúrbios seja realizado com melhor desempenho. Para a classificação dos distúrbios no vetor $\mathbf{u}(n)$, os parâmetros $\bar{C}_3^u(i) = \bar{C}_3^u(i, i)$, $i = 0, 1, 2, 3$ e $\bar{C}_4^u(i) = \bar{C}_4^u(i, i, i)$, $i = 0, 1, 2, 3$ foram extraídos para constituir o vetor de parâmetros $\mathbf{p}_u = [\bar{C}_3^u(0), \bar{C}_3^u(1), \bar{C}_3^u(2), \bar{C}_3^u(3), \bar{C}_4^u(0), \bar{C}_4^u(1), \bar{C}_4^u(2), \bar{C}_4^u(3)]^T$, onde $\bar{C}_3^u(i)$ é o valor normalizado de

$$C_3^u(i) = C_3^u(i, i, i) = E\{u(n)u^2(n+i)\} \approx \frac{1}{N} \sum_{n=0}^{N-1} u(n)u^2(n+i) \quad (9.21)$$

e $\bar{C}_4^u(i)$ é o valor normalizado de

$$\begin{aligned} C_4^u(i) &= C_4^u(i, i, i) = E\{u(n)u^3(n+i)\} - 3[E\{u(n)u(n+i)\}]^2 \\ &\approx \frac{1}{N} \sum_{n=0}^{N-1} u(n)u^3(n+i) - 3 \left[\frac{1}{N} \sum_{n=0}^{N-1} u(n)u(n+i) \right]^2 \end{aligned} \quad (9.22)$$

Nas expressões (9.21)-(9.22) foi considerado que $E\{u(n)\} = 0$ e $E\{.\}$ é o operador esperança. Observe ainda que $C_3^h(i, i)$, $C_4^h(i, i, i)$ são os cumulantes de ordem 3 e 4 quando os atrasos são iguais, vide (Mendel, 1991; Nikias & Petropulu, 1993).

Fazendo uso da técnica matriz *scatter* e do *Fisher's discriminant ratio* (Theodoridis & Koutroumbas, 1999), observa-se que os parâmetros listados acima são representativos para gerar regiões de classificação que minimizam a taxa de erro de classificação, vide resultados obtidos na seção 9.4.

9.3 Classificadores baseados no Critério Máxima Verossimilhança e em Sistemas Fuzzy do Tipo 2

Dentre as várias técnicas de reconhecimento de padrão encontradas na literatura, optou-se pela investigação teórica do classificador baseado no critério de máxima verossimilhança e do classificador baseado em sistemas nebulosos. Nesse sentido, a seção 9.3.1 deduz o classificador baseado no critério ML. A dedução do classificador fuzzy do tipo 2 foi realizada na seção 2.4. Devido à complexidade desses dois classificadores, na seção 9.4 alguns resultados sub-ótimos são apresentados com o classificador baseado na rede MLP, deduzido na seção 4.1.

9.3.1 Classificador de Distúrbios Baseado no Critério de Máxima Verossimilhança

Sejam \mathbf{E}_i^j , $i = 1, \dots, C^j$ e $j = f, h, u$, as C^j , $j = f, h, u$, classes de distúrbios associadas aos n -ésimos vetores de parâmetros $\mathbf{q}_n^j = [q_{n,1}^j, \dots, q_{n,L_j}^j]^T$, $j = f, h, u$, extraídos dos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$, respectivamente. Considere ainda \mathbf{E}_i^j , $i = 1, \dots, C^j$, como sendo um espaço vetorial dado por

$$\mathbf{E}_i^j = \left[\boldsymbol{\mu}_{\mathbf{E}_{i,1}^j} \boldsymbol{\mu}_{\mathbf{E}_{i,2}^j} \cdots \boldsymbol{\mu}_{\mathbf{E}_{i,M^j}^j} \right]^T, \quad i = 1, \dots, C^j, \quad j = f, h, u. \quad (9.23)$$

Note que os vetores $\boldsymbol{\mu}_{\mathbf{E}_{i,k}^j}$ definem a classe \mathbf{E}_i^j e M^j representa o número desses vetores na i -ésima classe do j -ésimo sinal.

Baseando-se na teoria de Bayes (Trees, 1968/1971; Theodoridis & Koutroumbas, 1999; Wei & Mendel, 2000), a classificação dos distúrbios observados nos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$ atinge a taxa mínima de erro, desde que as probabilidades, *a posteriori*, máximas sejam encontradas em $P(\mathbf{E}_i^f | \mathbf{q}_n^f)$, $i = 1, \dots, C^f$, $P(\mathbf{E}_i^h | \mathbf{q}_n^h)$, $i = 1, \dots, C^h$ e $P(\mathbf{E}_i^u | \mathbf{q}_n^u)$, $i = 1, \dots, C^u$. Se a ocorrência dos distúrbios nos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$ são equiprováveis, então os classificadores ótimos para os distúrbios que ocorrem nestes vetores é o classificador de máxima-verossimilhança (*maximum-likelihood* - ML). Note que o classificador ML acha a probabilidade condicional máxima dentre as probabilidades condicionais $P(\mathbf{q}_n^f | \mathbf{E}_i^f)$, $i = 1, \dots, C^f$, $P(\mathbf{q}_n^h | \mathbf{E}_i^h)$, $i = 1, \dots, C^h$ e $P(\mathbf{q}_n^u | \mathbf{E}_i^u)$, $i = 1, \dots, C^u$ para a classificação de distúrbios nos vetores $\mathbf{f}(n)$, $\mathbf{h}(n)$ e $\mathbf{u}(n)$.

Assim sendo, baseando-se na teoria de detecção (Trees, 1968/1971), a classificação dos distúrbios pode ser considerada como testes de hipóteses acerca das classes C^j , $j = f, h, u$, representadas por

- H_i^f : a classe é \mathbf{E}_i^f , $i = 1, \dots, C^f$, para o vetor $\mathbf{f}(n)$,

- H_i^h : a classe é \mathbf{E}_i^h , $i = 1, \dots, C^h$, para o vetor $\mathbf{h}(n)$,
- H_i^u : a classe é \mathbf{E}_i^u , $i = 1, \dots, C^u$, para o vetor $\mathbf{u}(n)$.

A técnica ML escolhe as hipóteses cujas funções de verossimilhança ou a função logaritmo da verossimilhança sejam maximizadas, ou seja, (Wei & Mendel, 2000)

$$H_i^{f,*} = \arg \max_{H_i^f} \ln L \left(H_i^f | \mathbf{q}_n^f \right) = \arg \max_{H_i^f} \ln p \left(\mathbf{q}_n^f | H_i^f \right), \quad (9.24)$$

$$H_i^{h,*} = \arg \max_{H_i^h} \ln L \left(H_i^h | \mathbf{q}_n^h \right) = \arg \max_{H_i^h} \ln p \left(\mathbf{q}_n^h | H_i^h \right) \quad (9.25)$$

e

$$H_i^{u,*} = \arg \max_{H_i^u} \ln L \left(H_i^u | \mathbf{q}_n^u \right) = \arg \max_{H_i^u} \ln p \left(\mathbf{q}_n^u | H_i^u \right). \quad (9.26)$$

Baseando-se na distribuição Gaussiana, tem-se que as densidades conjuntas de probabilidade dos vetores \mathbf{q}_n^f , \mathbf{q}_n^h e \mathbf{q}_n^u condicionadas à H_i^f , H_i^h e H_i^u , respectivamente, são dadas por

$$p(\mathbf{q}_n^f | H_i^f) = \sum_{k_f=1}^{M^f} P(\mathbf{q}_n^f) \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^f - \boldsymbol{\mu}_{\mathbf{E}_{i,k_f}^f} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_f}^f}^{-1} \left(\mathbf{q}_n^f - \boldsymbol{\mu}_{\mathbf{E}_{i,k_f}^f} \right) \right], \quad (9.27)$$

$$p(\mathbf{q}_n^h | H_i^h) = \sum_{k_h=1}^{M^h} P(\mathbf{q}_n^h) \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^h - \boldsymbol{\mu}_{\mathbf{E}_{i,k_h}^h} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_h}^h}^{-1} \left(\mathbf{q}_n^h - \boldsymbol{\mu}_{\mathbf{E}_{i,k_h}^h} \right) \right] \quad (9.28)$$

e

$$p(\mathbf{q}_n^u | H_i^u) = \sum_{k_u=1}^{M^u} P(\mathbf{q}_n^u) \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^u - \boldsymbol{\mu}_{\mathbf{E}_{i,k_u}^u} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_u}^u}^{-1} \left(\mathbf{q}_n^u - \boldsymbol{\mu}_{\mathbf{E}_{i,k_u}^u} \right) \right]. \quad (9.29)$$

Note que $P(\mathbf{q}_n^f) = \frac{1}{M^f}$, $P(\mathbf{q}_n^h) = \frac{1}{M^h}$ e $P(\mathbf{q}_n^u) = \frac{1}{M^u}$ são as probabilidades, *a priori*. $\boldsymbol{\mu}_{\mathbf{E}_{i,k_f}^f}$, $\boldsymbol{\mu}_{\mathbf{E}_{i,k_h}^h}$ e $\boldsymbol{\mu}_{\mathbf{E}_{i,k_u}^u}$ são os vetores que representam as classes \mathbf{E}_i^f , \mathbf{E}_i^h e \mathbf{E}_i^u , respectivamente. $\boldsymbol{\Sigma}_{\mathbf{E}_{i,k_f}^f} = \mathbf{I}\sigma_{i,k_f}^f$, $\boldsymbol{\Sigma}_{\mathbf{E}_{i,k_h}^h} = \mathbf{I}\sigma_{i,k_h}^h$ e $\boldsymbol{\Sigma}_{\mathbf{E}_{i,k_u}^u} = \mathbf{I}\sigma_{i,k_u}^u$ são as matrizes de covariância dos vetores que representam as classes \mathbf{E}_i^f , \mathbf{E}_i^h e \mathbf{E}_i^u . Note que \mathbf{I} é a matriz identidade. Logo, pode-se escrever

$$l(H_i^f|\mathbf{q}_n^f) = \ln L(H_i^f|\mathbf{q}_n^f) = \ln p(\mathbf{q}_n^f|H_i^f) = \ln \left\{ \frac{1}{M^f} \sum_{k_f=1}^{M^f} \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^f - \boldsymbol{\mu}_{\mathbf{E}_{i,k_f}^f} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_f}^f}^{-1} \left(\mathbf{q}_n^f - \boldsymbol{\mu}_{\mathbf{E}_{i,k_f}^f} \right) \right] \right\}, \quad (9.30)$$

$$l(H_i^h|\mathbf{q}_n^h) = \ln L(H_i^h|\mathbf{q}_n^h) = \ln p(\mathbf{q}_n^h|H_i^h) = \ln \left\{ \frac{1}{M^h} \sum_{k_h=1}^{M^h} \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^h - \boldsymbol{\mu}_{\mathbf{E}_{i,k_h}^h} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_h}^h}^{-1} \left(\mathbf{q}_n^h - \boldsymbol{\mu}_{\mathbf{E}_{i,k_h}^h} \right) \right] \right\} \quad (9.31)$$

e

$$l(H_i^u|\mathbf{q}_n^u) = \ln L(H_i^u|\mathbf{q}_n^u) = \ln p(\mathbf{q}_n^u|H_i^u) = \ln \left\{ \frac{1}{M^u} \sum_{k_u=1}^{M^u} \exp \left[-\frac{1}{2} \left(\mathbf{q}_n^u - \boldsymbol{\mu}_{\mathbf{E}_{i,k_u}^u} \right)^T \boldsymbol{\Sigma}_{\mathbf{E}_{i,k_u}^u}^{-1} \left(\mathbf{q}_n^u - \boldsymbol{\mu}_{\mathbf{E}_{i,k_u}^u} \right) \right] \right\}. \quad (9.32)$$

Logo, as probabilidades condicionais de erro são dadas por

$$P(\text{erro}^j|H_i^j) = 1 - P[l(H_i^j|\mathbf{p}_n^j) < l(H_k^j|\mathbf{p}_n^j), \forall k \neq i], \quad i = 1, \dots, C^j, \quad j = f, h, u, \quad (9.33)$$

e, conseqüentemente,

$$P(\text{erro}^j) = \frac{1}{C^j} \sum_{i=1}^{M^j} P(\text{erro}|H_i^j), \quad j = f, h, u. \quad (9.34)$$

9.4 Resultados Numéricos

Para verificar se o uso da técnica proposta para a classificação de distúrbios em sistema de potência é efetiva, algumas simulações foram feitas. A Tab. 9.2 mostra

os resultados de treinamento e teste para a identificação da ocorrência de distúrbios no vetor $\mathbf{f}(n)$. Os resultados foram obtidos com as seguintes considerações: 1) $N = 128, 256, 512, 768, 1024$, ou seja, vetores $\mathbf{f}(n)$, contendo 1/2, 1, 2, 3 e 4 ciclos da componente fundamental foram analisados; 2) $\mathbf{p}_f = [\overline{C}_3^f(0, 0) \overline{C}_4^f(0, 0, 0) \hat{A}_0]$; 3) 4000 amostras de distúrbios gerados com o software Matlab e igualmente divididas entre conjuntos de teste e de treinamento; 4) o uso de uma rede MLP com estrutura $3 \times 2 \times 1$ para a classificação de cada um tipo de evento; 5) 2500 épocas foram utilizadas no treinamento da rede MLP, 6) os pesos das redes MLP foram igualmente inicializados, 7) $f_s = 256 \times 60$ Hz e 32 bits para quantização das amplitudes; 8) $\sigma_v^2 = -40$ dB; 9) Os distúrbios no vetor $\mathbf{f}(n)$, os quais são relacionados às interrupções, *sags*, condição normal de operação e *swell*, foram gerados de acordo com a Tab. 9.1.

Conforme pode ser observado na Tab. 9.2, a técnica para a classificação de distúrbios no sinal $\{f(n)\}$ é bastante efetiva, pois é capaz de classificar cerca de 99,94 % dos distúrbios observados. De fato, a análise isolada do vetor $\mathbf{f}(n)$ permite que técnicas adequadas e com baixo custo computacional sejam empregadas apropriadamente para a classificação de distúrbios e, conseqüentemente, soluções simples e efetivas para o problema de classificação de distúrbios são obtidas.

Para avaliar numericamente o desempenho do classificador proposto, baseado na rede MLP para a detecção do distúrbio harmônica, as seguintes considerações foram feitas: 1) $N = 128, 256, 512, 768, 1024$, ou seja, vetores $\mathbf{h}(n)$ contendo 1/2, 1, 2, 3 e 4 ciclos da componente fundamental foram analisados; 2) $\mathbf{p}_h = [\overline{C}_3^h(0, 0) \overline{C}_4^h(0, 0, 0)]$; 3) 2000 amostras de distúrbios gerados com o Matlab e igualmente divididas entre conjuntos de teste e de treinamento; 4) o uso de uma rede MLP com 2 entradas, dois neurônios na camada escondida e uma saída, para a detecção do distúrbios harmônicas; 5) 500 épocas foram utilizadas no treinamento da rede MLP, 6) os pesos das redes MLP foram igualmente inicializados, 7) $f_s = 256 \times 60$ Hz e 32 bits; 8) $\sigma_v^2 = -40$ dB; 9) as componentes harmônicas $\omega_i, i = 3, 5, 7, 9, 11, 13, 15$ têm

Resultado / Tipo Distúrbio	Interrup.	Sag	Normal	Swell
<i>Trein.</i> , 1/2 <i>C</i>	98,65	100,00	99,85	100,00
<i>Teste</i> , 1/2 <i>C</i>	98,40	99,95	99,60	99,95
<i>Trein.</i> , 1 <i>C</i>	100,00	100,00	100,00	100,00
<i>Teste</i> , 1 <i>C</i>	99,95	99,95	99,95	99,90
<i>Trein.</i> , 2 <i>C</i>	99,85	100,00	99,10	100,00
<i>Teste</i> , 2 <i>C</i>	99,75	99,95	98,75	99,95
<i>Trein.</i> , 3 <i>C</i>	98,85	100,00	98,85	100,00
<i>Teste</i> , 3 <i>C</i>	98,55	99,95	98,55	99,95
<i>Trein.</i> , 4 <i>C</i>	98,80	100,00	98,75	100,00
<i>Teste</i> , 4 <i>C</i>	98,75	99,95	98,70	99,95

Tabela 9.2: Taxa em (%) de classificação, obtida com a técnica de classificação proposta, para distúrbios em $\mathbf{f}(n)$. Note que C é a abreviatura de ciclos da componente fundamental.

parâmetros A_i e θ_i modelados como variáveis aleatórias com distribuição uniforme nos intervalos $(0, 005, 0, 4]$ e $(0, 2\pi]$, respectivamente.

A Tab. 9.3 mostra os resultados de treinamento e de teste obtidos para diferentes valores de N . Obteve-se bons resultados para a detecção de harmônicas com a técnica proposta. Destaca-se ainda a baixa complexidade da rede MLP, que tem 3 entradas, 2 neurônios e uma saída. A obtenção dos resultados de classificação iguais a 100%, mostrados na Tab. 9.3, não é nenhuma novidade. Outras técnicas obtém tais desempenhos para $N \geq 1024$ (4 ciclos da fundamental) (Wang & Mamishev, 2004, 2004; Abdel-Galil et al., 2004; Youssef et al., 2004). A novidade é a obtenção desses resultados quando $N = 128, 256, 512, 768$, ou seja, para tamanhos de janelas de análise iguais a 1/2, 1, 2 e 3 ciclos da fundamental quando $f_s = 256 \times 60$ amostras/seg.

Tipo Distúrbio	Harmônicas
<i>Trein., 1/2C</i>	100,00
<i>Teste, 1/2C</i>	100,00
<i>Trein., 1C</i>	100,00
<i>Teste, 1C</i>	100,00
<i>Trein., 2C</i>	100,00
<i>Teste, 2C</i>	100,00
<i>Trein., 3C</i>	100,00
<i>Teste, 3C</i>	100,00
<i>Trein., 4C</i>	100,00
<i>Teste, 4C</i>	100,00

Tabela 9.3: Taxa em (%) de classificação, obtida com a técnica de classificação proposta, para a detecção de distúrbios em $\mathbf{h}(n)$.

Finalmente, para verificar a eficácia do classificador proposto para os distúrbios relacionados ao vetor $\mathbf{u}(n)$, as seguintes considerações foram feitas: 1) $N = 1024$, ou seja, vetores $\mathbf{u}(n)$ contendo 4 ciclos da componente fundamental foram analisados; 2) $\mathbf{p}_u = [\overline{C}_3^u(0, 0) \dots \overline{C}_3^u(3, 3) \overline{C}_4^u(0, 0, 0) \dots \overline{C}_4^u(3, 3, 3)]$; 3) 400 amostras de distúrbios geradas sinteticamente, as quais foram igualmente divididas entre conjuntos de teste e de treinamento; 4) o uso de uma rede MLP com estrutura $8 \times 4 \times 1$ para a detecção dos distúrbios *spikes*, *notchs* e *oscilação amortecidas*; 5) 3500 épocas foram utilizadas no treinamento da rede MLP, 6) os pesos das redes MLP foram igualmente inicializados, 7) $f_s = 256 \times 60$ Hz e 32 bits para a quantização da amplitude e 8) $\sigma_v^2 = -40$ dB. Conforme pode ser observado, o uso da abordagem proposta provê resultados bastante significativos para o problema de detecção de distúrbios em sistemas de potência.

Finalmente, a Tab. 9.5 mostra os resultados obtidos quando a técnica de

Resultados / Tipo Distúrbio	Notchs	Spikes	Oscil. Amort.
<i>Trein., 4C</i>	100,00	100,00	100,00
<i>Teste, 4C</i>	100,00	99,00	99,50

Tabela 9.4: Taxa em (%) de classificação, obtida com a técnica de classificação proposta, dos distúrbios em $\mathbf{u}(n)$.

classificação proposta e a técnica introduzida em (Wang & Mamishev, 2004; Wang et al., 2004) foram aplicadas aos distúrbios gerados com $N = 1024$. Note que a última linha da Tab. 9.5 corresponde a ocorrência de múltiplos distúrbios. A técnica proposta apresenta um resultado levemente superior à técnica proposta em (Wang et al., 2004) se as mesmas são utilizadas para a classificação de distúrbios que ocorram isoladamente. No entanto, a técnica proposta na presente tese apresenta resultados superiores se há a ocorrência de múltiplos distúrbios (vide resultados da última linha da Tab. 9.5). De fato, o uso adequado do princípio dividir e conquistar permite que a técnica proposta seja adequada não apenas para a classificação de distúrbios que ocorram isoladamente, mas também para a classificação de múltiplos distúrbios.

Uma questão não analisada aqui é o uso de banco de dados reais para validar os resultados obtidos. No entanto, é importante salientar que os sinais sintéticos foram gerados de acordo com (Abdel-Galil et al., 2004; Youssef et al., 2004; Moore & Português, 2003), o que, de certa forma, garante a confiabilidade dos mesmos. Além disso, a comparação da técnica proposta com a técnica introduzida em (Wang & Mamishev, 2004; Wang et al., 2004) é um indicador bastante confiável dos resultados obtidos com a técnica proposta.

Tipo Distúrbio / Técnica	Técnica Proposta	(Wang et al., 2004)
Swell	99,95	98,85
Normal	99,70	99,25
Sags	99,95	100,00
Interrupção	99,90	98,75
Spikes	99,00	98,50
Harmônicas	100,00	100,00
Oscilat. Amort.	99,50	99,50
Notchs	100,00	99,50
Mix. Múltiplos Eventos	99,45	92,32

Tabela 9.5: Taxa em (%) de classificação, obtida com a técnica de classificação proposta, dos múltiplos distúrbios em $\mathbf{x}(n)$ quando $N = 1024$.

9.5 Sumário

Inspirado no princípio dividir e conquistar, o presente capítulo introduziu uma nova técnica para a classificação de distúrbios em sistemas de potência. Baseando-se nessa nova abordagem, propôs-se um método novo para a classificação de distúrbios isolados e/ou de múltiplos distúrbios. A técnica proposta, basicamente, faz uso de uma técnica de pré-processamento para a obtenção das componentes principais dos sinais monitorados (fundamental, harmônicas e transitórios). A seguir, são extraídos EOS dessas componentes, de forma que a classificação dos distúrbios seja realizada com a ocorrência mínima de intersecções entre as classes de distúrbios associadas às componentes distintas. Finalmente, a partir do vetor de parâmetros extraído das componentes, o uso de técnicas baseadas em inteligência computacional é proposto para detectar a ocorrência dos diferentes tipos de distúrbios encontrados nas redes elétricas. Dentre as técnicas de inteligência computacional, optou-se pelo classificador baseado na rede MLP para a classificação dos distúrbios. Resultados

numéricos foram apresentados para verificar o desempenho da técnica proposta para a classificação dos mesmos. Tais resultados validam o uso da técnica proposta para a classificação de distúrbios isolados e de múltiplos distúrbios em redes elétricas.

10

Compressão de Distúrbios

Atualmente, os equipamentos de monitoramento de QE incluem diferentes técnicas de processamento de sinais com objetivos diversos, dentre os quais pode-se citar a compressão dos distúrbios monitorados. O desenvolvimento de técnicas de codificação de fonte que promovam o armazenamento eficiente das formas de ondas dos distúrbios monitorados e contribuam para o uso eficiente dos canais de comunicações surge como uma área de pesquisa com nítida aplicação e demanda em sistemas de potência (Mehta & Russel, 1989; Santoso, Powers, & Grady, 1997).

É importante salientar que a compressão dos distúrbios é um problema de

interesse relativamente recente devido, principalmente, ao uso crescente das cargas de natureza eletrônica e não-linear com componentes espectrais acima de 1 MHz e a proliferação de distúrbios em sistemas de potência. Nesse contexto, a demanda por equipamentos de monitoramento instalados em vários pontos do sistema e sua transmissão ou armazenamento demanda bastante memória ou largura de banda. Por exemplo, considerando-se 1000 equipamentos de monitoramento cujo A/D é de 16 bits e a taxa de amostragem igual a 15360 amostras/s, uma largura de banda igual a 1,03Gbps é necessária para transferir os sinais de tensão e de corrente das três fases monitoradas pelos equipamentos de monitoramento para uma central de processamento (Khan, 2001). Objetivando diminuir tal necessidade, as seções 10.1 e 10.2 descrevem duas novas técnicas para a compressão de distúrbios em sistemas de potência. Conforme ficará claro ao longo do trabalho, a técnica proposta na seção 10.2 é a generalização da técnica proposta na seção 10.1.

10.1 *Enhanced Data Compression Method* (EDCM)

A eficiência das transformada Wavelet (TW) e transformada Wavelet Packet (TWP) (Burrus et al., 1997; Daubechies, 1992; Mallat, 2001, 1989; Vitterli & Kovacevic, 1995; Vaidyanathan, 1995) aplicadas à compressão de dados em sistemas de potência tem sido bastante analisada (Santoso et al., 1997; Littler & Morrow, 1999; Chung, Powers, Grady, & Bhatt, 1999; Hamid & kawasaki, 2002; Duque, Ribeiro, Braga, & Almeida, 2000; Chung et al., 2004; Dash et al., 2003; Mehta & Russel, 1989; Hsieh & Huang, 2003; Gerek & Ece, 2004; Panda, Dash, Pradhan, & Meher, 2002; Meher, Pradhan, & Panda, 2004; Chung et al., 1999). No entanto, tais soluções não levam totalmente em consideração a propriedade de representação esparsa da TW para a análise de transitórios e sua não adequação para a análise de

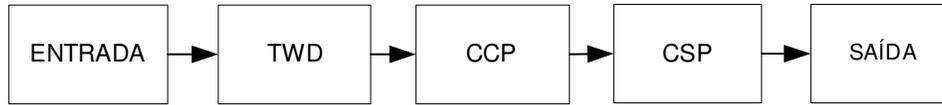


Figura 10.1: *Standard Data Compression Method* (SDCM).

sinais senoidais em sistemas de potência, conforme discutido em (Ribeiro, Duque, & Romano, 2004, 2001). De forma geral, todas as técnicas citadas acima aplicam a TW e, na seqüência, técnicas de compressão com e sem perda (Gersho & Gray, 1999; Gray & Neuhoff, 1998; Berger & Gibson, 1998) para a compressão de distúrbios. Tal abordagem é ilustrada na Fig. 10.1. O distúrbio monitorado é submetido à TW ou à TWP e, a seguir, utilizam-se as técnicas de quantização escalar, seguida da técnica de codificação por entropia.

Uma nova abordagem proposta para a compressão de sinais de potência foi proposta em (Hsieh, Huang, & Huang, 1998). Os autores de (Hsieh et al., 1998) consideraram a separação do sinal monitorado em componentes fundamental e transitórios. No entanto, nenhuma solução computacional para a separação dos dois tipos de sinais (transitórios e sinais senoidais) foi introduzida por eles. Além disso, nenhuma justificativa foi apresentada para tal abordagem. Note que tal abordagem foi bem sucedida quando aplicada à compressão de sinais de voz (Shlomot, Cuperman, & Gersho, 2001).

A partir do sinal monitorado dado por (6.1), pode-se definir o sinal $\{f(n)\}$ como sendo a parte estacionária do sinal $\{x(n)\}$ e $\{u(n)\} = \{h(n)\} + \{t(n)\} + \{i(n)\} + \{v(n)\}$ como a componente não-estacionária do sinal (constituída das harmônicas, inter-harmônicas, ruído de fundo e dos diversos transitórios). Essa definição é baseada nas seguintes suposições: *i*) os parâmetros da componente fundamental quase não variam com o tempo, e *ii*) a ocorrência das componentes harmônicas e inter-harmônicas, dos transitórios e do ruído aditivo é aleatória e os parâmetros

destas componentes senoidais variam de acordo com a ocorrência das mesmas.

Dadas as restrições acima, percebe-se que apenas três parâmetros (amplitude, fase e frequência) são suficientes para caracterizar a componente fundamental. O fato da componente fundamental ser vista como um sinal estacionário significa que a mesma pode se obtida a partir de

$$a_0 \cong \langle x(n), \psi_0^*(n) \rangle = \sum_{n=0}^{N-1} x(n) \cos(\omega_0 n), \quad (10.1)$$

$$b_0 \cong \langle x(n), \phi_0^*(n) \rangle = \sum_{n=0}^{N-1} x(n) \sin(\omega_0 n), \quad (10.2)$$

$$A_0 \cong \sqrt{a_0^2 + b_0^2} \quad (10.3)$$

e

$$\theta_0 \cong \tan^{-1} \left(\frac{b_0}{a_0} \right). \quad (10.4)$$

Assim sendo, o uso de técnicas adequadas para a estimação desses parâmetros (amplitude, fase e frequência) oferece nítida vantagem na compressão da componente fundamental. De fato, como o sinal $\{u(n)\}$ representa a componente não-estacionária, isso significa que o mesmo pode ser submetido à compressão por transformação baseada na TW ou TWP, enquanto a componente fundamental é armazenada através de seus parâmetros. Em (Hsieh et al., 1998) essa idéia foi aplicada, no entanto, não foi apresentada nenhuma solução computacional que decomponha o sinal $\{x(n)\}$ nos sinais $\{f(n)\}$ (componente fundamental) e $\{u(n)\} = \{x(n)\} - \{f(n)\}$. Além disso, eles não fizeram uso das técnicas de quantização e de codificação por entropia (Gersho & Gray, 1999).

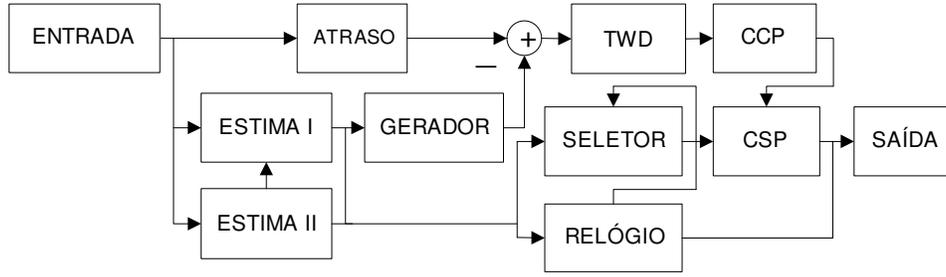


Figura 10.2: *Enhanced Data Compression Method (EDCM)*.

Antes de seguir adiante é importante mostrar que a versão discreta da TW aplicada ao sinal $\{u(n)\}$ é dada por (Mallat, 2001)

$$DWT(m, n) = 2^{-m/2} \sum_{m=0}^J \sum_{n=0}^{N-1} u(n) \Psi \left(\frac{t - n2^m}{2^m} \right) \quad (10.5)$$

e

$$\Psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \Psi \left(\frac{t - nb_0 a_0^m}{a_0^m} \right) \quad (10.6)$$

Em (10.6) $a_0 = 2$ e $b_0 = 1$.

Objetivando implementar a idéia acima discutida, a Fig. 10.2 mostra a solução *Enhanced Data Compression Method (EDCM)*, proposta em (Ribeiro et al., 2004, 2001), a qual representa a extensão dos resultados iniciais apresentados em (Duque et al., 2000).

A descrição de cada bloco da proposta EDCM é dada abaixo.

Entrada: armazena o sinal amostrado;

Atraso: atrasa o sinal monitorado de forma a sincronizá-lo com o sinal gerado no bloco de geração de senóide. Tal atraso garante que o sinal erro ou a componente não-estacionária será gerada corretamente;

Estima I: realiza a estimação da amplitude e da fase da componente fundamental, a partir do Filtro de Kalman. Conforme mostrado em (Liu, 2000), o filtro de Kalman é, apesar de sua inerente complexidade computacional, um estimador de amplitude e fase com excelente desempenho;

Estima II: neste bloco, o filtro notch adaptativo (Regalia, 1995; Romano & Bellanger, 1988) é utilizado para estimar a frequência da componente fundamental. Segundo (Regalia, 1995; Romano & Bellanger, 1988), esta solução apresenta um resultado bastante satisfatório quando se analisa, conjuntamente, o desempenho e a complexidade computacional;

Gerador: considerando que a geração de um sinal senoidal a partir da implementação direta de sua equação apresenta um custo computacional elevado, este bloco implementa um filtro IIR de segunda ordem (Cheng, 2000) de forma que a reprodução da componente fundamental é feita com baixo custo computacional;

Seletor: seleciona as estimativas da amplitude, frequência e fase de acordo com as informações presentes no bloco Relógio;

Relógio: gera os instantes em que os parâmetros da componente fundamental devem ser selecionados;

TWD : aplica a transformada Wavelet discreta (TWD) no sinal erro, gerado após a subtração da componente fundamental. O sinal erro é dado por $\{u(n)\}$.

CCP: aplica um limiar nos coeficientes obtidos após a aplicação da TWD. Os coeficientes que apresentam amplitudes menores que o limiar adotado são descartados (Donoho, 1995). Após este procedimento, o sinal resultante é submetido ao processo de quantização escalar . O descarte dos coeficientes desprezíveis é

EA	EFA	EFR	IE	DADOS
----	-----	-----	----	-------

Figura 10.3: Cabeçalho do EDCM.

realizado pelo *hard threshold* dado por

$$\rho_y(y(n)) = \begin{cases} y(n), & \text{se } |y(n)| > T \\ 0, & \text{se } |y(n)| \leq T \end{cases}, \quad (10.7)$$

onde $y(n)$, $n = 0, \dots, N - 1$ é o sinal no domínio da TWD e T é o limiar a ser escolhido. A escolha de T depende da distribuição de probabilidade do ruído e dos coeficientes no domínio da TW (Donoho & Johnstone, 1994; Donoho, 1995; Mallat, 2001).

CSP: aplica o algoritmo LZW , Lempel-Ziv-Welch, (Gersho & Gray, 1999) para realizar a compressão sem perda.

SAÍDA: armazena os dados obtidos após a aplicação da compressão sem perda.

A Fig. 10.3 mostra a estrutura do cabeçalho proposta para armazenar o sinal monitorado antes da aplicação do algoritmo LZW. As seções do cabeçalho são descritas a seguir:

AE: armazena a amplitude da componente fundamental estimada pelo filtro de Kalman;

EFA: armazena a fase da componente fundamental estimada pelo filtro de Kalman;

EFR: armazena a frequência da componente fundamental estimada pelo filtro notch adaptativo;

IE: armazena o instante inicial do evento armazenado;

DADOS: armazena o sinal obtido após a aplicação da técnica de codificação por entropia no sinal $\{u(n)\}$ quantizado.

Para avaliar o desempenho da técnica proposta em comparação com as técnicas tradicionais, propõe-se o índice FHIR (*fundamental to harmonic and inter-harmonic ratio*) definido por

$$FHIR := 10 \log_{10} \left[\frac{\sum_{n=0}^{N-1} f^2(n)}{\sum_{m=1}^M \left(\sum_{i=0}^{N-1} h_m^2(n) \right) + \sum_{j=1}^J \left(\sum_{i=0}^{N-1} i_j^2(n) \right)} \right]. \quad (10.8)$$

Conforme será visto na seção de resultados, este índice é representativo para avaliar, comparativamente, o desempenho da EDCM com o desempenho da outra técnica de compressão de sinais de potência, aqui denominada *standard data compression method* (SDCM) (Santoso et al., 1997).

10.1.1 Estimação da Amplitude e da Fase da Componente Fundamental

A estimação conjunta da fase e da amplitude de sinais senoidais pode ser obtida utilizando-se o filtro de Kalman (Kalman, 1960; Liu, 2000). A implementação direta do algoritmo de Kalman apresenta complexidade $O(Q^2)$, onde Q é o número de estados a serem estimados. O desempenho do filtro de Kalman é influenciado pela SNR presente nos vetores de dados de entrada do algoritmo. A implementação do

filtro de Kalman é, *a priori*, bastante simples, desde que as equações de estado do processo analisado sejam corretamente definidas.

Supondo que o parâmetro $f_0 = \frac{\omega_0}{2\pi} f_s$ da componente fundamental seja conhecido, então pode-se escrever que

$$x(n) = \begin{bmatrix} \cos(\omega_0 n) & \sin(\omega_0 n) \end{bmatrix} \begin{bmatrix} A_0 \cos \theta_0 \\ A_0 \sin \theta_0 \end{bmatrix} + u(n), \quad (10.9)$$

$$\begin{bmatrix} g_c(n) \\ g_s(n) \end{bmatrix} = \begin{bmatrix} A_0 \cos \theta_0 \\ A_0 \sin \theta_0 \end{bmatrix} \quad (10.10)$$

e

$$v(n) = x(n) - A_0 \cos(\omega_0 n + \theta_0) \quad (10.11)$$

onde $g_c(n)$ e $g_s(n)$ são os estados do processo a serem estimados. ω_0 é a frequência angular da componente fundamental. As amostras $v(n)$, $n = 0, \dots, N - 1$, são modeladas como variáveis aleatórias *i.i.d.* com distribuição Gaussiana $\mathcal{N}(0, \sigma_v^2)$. Assim sendo, as equações de estado na $(n+1)$ -ésima iteração são dadas por

$$\begin{bmatrix} g_c(n+1) \\ g_s(n+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g_c(n) \\ g_s(n) \end{bmatrix} + q(n), \quad (10.12)$$

onde $q(n)$ é modelada como uma variável aleatória de distribuição Gaussiana $\mathcal{N}(0, \sigma_q^2)$. Utilizando-se (10.12) na formulação do filtro de Kalman, estimativas instantâneas da amplitude e da fase do componente fundamental são obtidas.

10.1.2 Estimação da Frequência da Componente Fundamental

O desempenho do filtro de Kalman aplicado para a estimação da amplitude e da fase de um sinal senoidal é dependente da estimativa da frequência do sinal

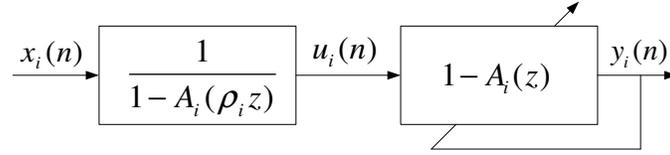


Figura 10.4: Diagrama de blocos do i -ésimo filtro Notch.

senoidal, conforme explicitado por (10.9). De forma a realizar a estimativa da frequência de um sinal senoidal com baixo custo computacional, utilizou-se o filtro notch adaptativo. A função de transferência de P seções de 2^a ordem do filtro notch adaptativo é dada por

$$H(z) = \prod_{i=0}^{P-1} \frac{1 - a_{i1}z^{-1} - a_{i2}z^{-2}}{1 - \rho_i a_{i1}z^{-1} - \rho_i^2 a_{i2}z^{-2}} = \prod_{i=0}^{P-1} \frac{1 - A_i(z)}{1 - A_i(\rho_i z)} \quad (10.13)$$

onde P é o número de seções de 2^a ordem. Uma única seção do filtro notch adaptativo é mostrada na Fig. 10.4.

O filtro notch adaptativo consiste, basicamente, na adaptação dos parâmetros da parte não recursiva de cada seção de 2^a ordem, uma vez que tal parte assume o papel de um filtro transversal de erro de predição. Os algoritmos adaptativos *Recursive Least-Square* (RLS) e *Least Mean Square* (LMS) (Haykin, 1996) podem ser diretamente aplicados para atualizar cada seção de 2^a ordem do filtro notch adaptativo. As equações de atualização dos parâmetros do i -ésimo filtro adaptativo de 2^a ordem são as seguintes:

$$u_i(n+1) = x_i(n+1) + \sum_{k=1}^2 a_{ik}(n) \rho_i^k u_m(n+1-k), \quad (10.14)$$

$$y_i(n+1) = u_i(n+1) - \sum_{k=1}^2 a_{ik}(n) \rho_i^k u_i(n+1-k) \quad (10.15)$$

e

$$\begin{bmatrix} a_{i1}(n+1) \\ a_{i2}(n+1) \end{bmatrix} = \begin{bmatrix} a_{i1}(n) \\ a_{i2}(n) \end{bmatrix} + G_i(n)y_i(n+1), \quad (10.16)$$

onde $x_i(n)$ e $y_i(n) = x_{i+1}(n)$, $i = 0, \dots, P - 2$ representam a entrada e a saída da i -ésima seção do filtro notch (Regalia, 1995; Romano & Bellanger, 1988). Para cada seção de 2ª ordem, o ganho de adaptação chamado ganho de Kalman, $G_m(n)$, é atualizado pelo algoritmo RLS. O algoritmo RLS foi escolhido neste trabalho, pois apresenta melhor desempenho em termos de velocidade de convergência do que o algoritmo LMS (Haykin, 1996). Apesar de apresentar maior complexidade computacional quando comparado com o algoritmo LMS, pode-se observar que, para a aplicação em questão, o custo computacional é bastante reduzido, visto que apenas dois parâmetros são estimados. Finalmente, a estimativa da frequência é obtida a partir da equação não-linear dada por

$$\hat{f}_i = \cos^{-1} \left(\frac{a_{i1}}{2} \right) \frac{f_s}{2\pi}, \quad (10.17)$$

onde f_s é a frequência de amostragem do sinal. De fato, tal expressão corresponde a fase da raiz complexa do termo $[1 - A_i(z)]$, $i = 0, \dots, P - 1$, que assume módulo unitário após a convergência do filtro notch. Para a estimação da frequência, fica evidente que o uso de uma tabela para armazenar os valores do termo $\cos^{-1} \left(\frac{a_{i1}}{2} \right) \frac{f_s}{2\pi}$ contribui para reduzir a complexidade computacional

10.1.3 Resultados Numéricos

Objetivando analisar o desempenho da presente proposta, três tipos de sinais são considerados: 1) o sinal distorcido por *notches* impulsivos gerados pelos retificadores dos conversores AC/DC. Esse tipo de distúrbio tem componentes espectrais em diversas faixas de frequência; 2) o sinal transitório gerado pelo chaveamento capacitivo na rede. Tal fenômeno exemplifica um distúrbio que ocorre numa limitada

faixa de freqüência; 3) o sinal distorcido pela presença de algumas componentes harmônicas, tal como é dado pela expressão

$$\begin{aligned} x(n) = f(n) + h_3(n) + h_5(n) + h_7(n) + \Phi_1(n) + v(n) = \cos(\omega_0 n + \pi/4) + \\ + 0,3 \cos(3\omega_0 n + \pi/3) + 0,1 \cos(5\omega_0 n + \pi/6) + 0,1 \cos(7\omega_0 n + \pi/4) + \Phi_1(n) + v(n), \end{aligned} \quad (10.18)$$

onde $\Phi_1(n)$ representa o distúrbio gerado por um chaveamento capacitivo na rede.

Para a geração dos sinais, acima listados, foram consideradas $f_s = 60$ kHz e 12 bits para a quantização da amplitude. A *wavelet* utilizada para ilustrar a eficácia da presente proposta foi a Daubechies 4 (Daubechies, 1992). A compressão foi realizada em quadros contendo 5 ciclos da componente fundamental. Os parâmetros utilizados para avaliar o desempenho da presente proposta são o erro médio quadrático (EQM) ou (*mean square error* - MSE) e a relação sinal ruído (*signal to noise ratio* - SNR) dados por

$$MSE(dB) = 10 \log_{10} \left(\frac{1}{N} \sum_{n=0}^{N-1} (x(n) - \hat{x}(n))^2 \right) \quad (10.19)$$

e

$$SNR(dB) = 10 \log_{10} \left(\frac{\sum_{n=0}^{N-1} (x(n))^2}{\sum_{n=0}^{N-1} (x(n) - \hat{x}(n))^2} \right), \quad (10.20)$$

respectivamente. Em (10.19)-(10.20), $\{x(n)\}$ e $\{\hat{x}(n)\}$ são, respectivamente, os sinais original e reconstruído. Além disso, a taxa de compressão é também utilizada como parâmetro de análise.

A Fig. 10.5 mostra a convergência do coeficiente a_{01} do filtro notch adaptativo para os três tipos de distúrbios analisados. Conforme observado na Fig. 10.5, obtém-se um bom desempenho com o filtro notch adaptativo. As Figs. 10.6 e 10.7, por sua vez mostram a convergência dos parâmetros $g_c(n)$ e $g_s(n)$, a partir dos quais a amplitude e a fase da componente fundamental são estimadas. Observe que MSE menores que -20 dB são obtidos para os diferentes casos considerados. Tais resultados garantem uma boa estimativa da fase e da amplitude.

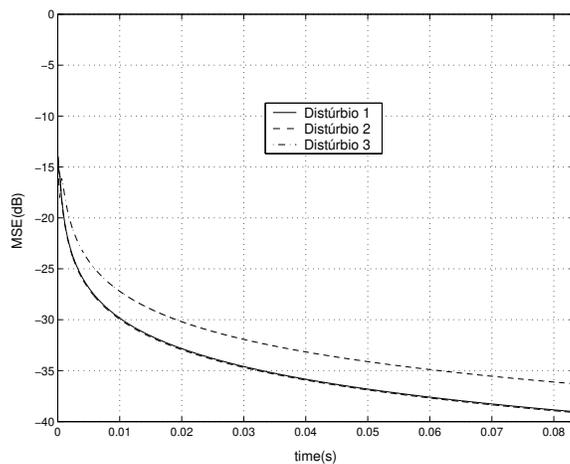


Figura 10.5: Convergência do parâmetro a_{01} do filtro notch adaptativo.

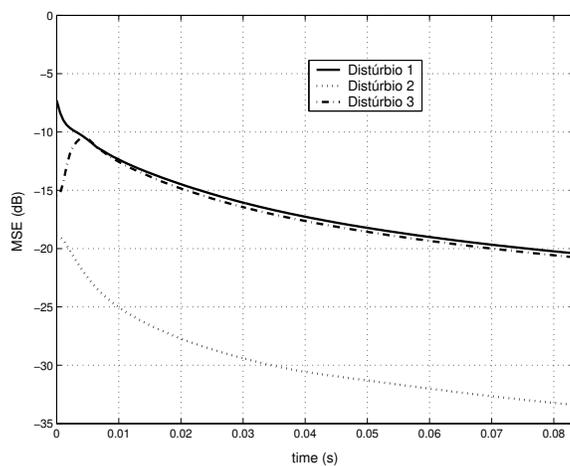


Figura 10.6: Convergência do parâmetro $g_c(n)$ do filtro de Kalman.

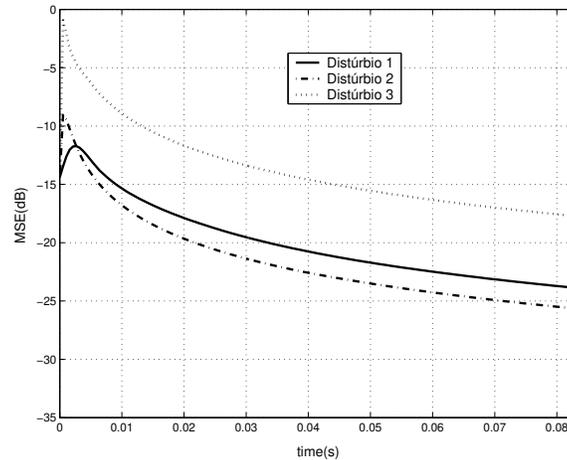


Figura 10.7: Convergência do parâmetro $g_s(n)$ do filtro de Kalman.

A Fig. 10.8 ilustra o distúrbio I, $\{x(n)\}$, e o sinal erro gerado $\{u(n)\}$. Observe que o sinal erro $\{u(n)\}$ representa com bastante precisão a componente não-estacionária do sinal monitorado. Aplicando-se a EDCM e a SDCM para diferentes valores de limiares, dados por (10.7), e levando-se em consideração vários tipos de wavelet Daubechies, é possível obter as superfícies de SNR, mostrada na Fig. 10.9. A EDCM apresenta resultados melhores, em termos de SNR, do que a SDCM. As diversas simulações computacionais feitas evidenciam que quando $FHIR \rightarrow \infty$, melhores resultados são obtidos com a EDCM. Para verificar os desempenhos das técnicas EDCM e SDCM quando aplicadas para a compressão do distúrbio I, a Tabela 10.1 reúne os melhores resultados obtidos em termos de taxa de compressão, a partir das superfícies mostradas na Fig. 10.9. Observe que na tabela 10.1, a segunda linha refere-se ao número de níveis de decomposição da TW que apresentou os melhores resultados quando ambas as técnicas foram aplicadas.

De forma a dar prosseguimento à análise de desempenho de ambas as técnicas, EDCM e SDCM, a Fig. 10.10 mostra o distúrbio II e o sinal erro correspondente gerado. Já a Fig. 10.11 ilustra as superfícies de SNR obtidas com ambas as

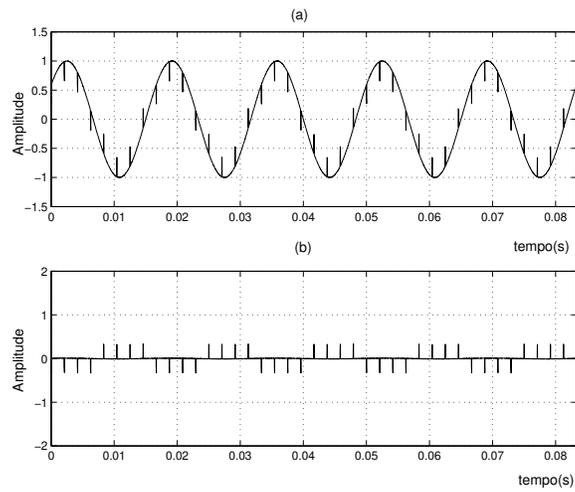


Figura 10.8: Distúrbio I: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$.

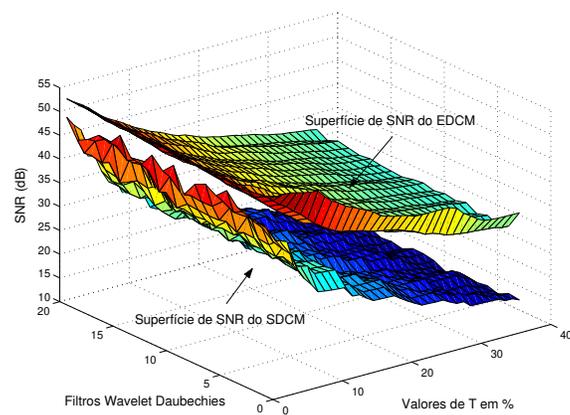


Figura 10.9: Distúrbio I: comparação de desempenho entre a SDCM e a EDCM.

EDCM			SDCM		
TWD com 1 níveis de decomposição			TWD com 3 níveis de decomposição		
$T(\%)$	SNR	TC	$T(\%)$	SNR	TC
2	52,82	3,63:1	2	53,63	3,71:1
6	41,76	16,45:1	6	41,69	7,02:1
10	41,23	17,61:1	10	33,63	7,53:1
14	37,90	17,86:1	14	29,25	8,29:1
18	35,63	18,03:1	18	26,28	9,19:1
22	35,17	18,01:1	22	23,78	9,97:1

Tabela 10.1: Distúrbio I: SNR e taxa de compressão (TC) para diferentes valores em (%) de T .

técnicas. Observe que, para este sinal, numa região com SNR elevada, a SDCM apresenta melhores resultados, em termos de SNR . Para esse caso, as melhores taxas de compressão são agrupadas na Tabela 10.2. Conforme pode ser observado, o desempenho da EDCM é superior àquele obtido com a SDCM.

EDCM			SDCM		
TWD com 6 níveis de decomposição			TWD com 6 níveis de decomposição		
$T(\%)$	SNR	TC	$T(\%)$	SNR	TC
2	46,38	25,51:1	2	50,91	3,86:1
6	35,20	39,06:1	6	37,90	23,15:1
10	32,16	42,13:1	10	30,92	25,34:1
14	30,89	44,12:1	14	26,39	26,22:1
18	28,76	45,18:1	18	25,36	27,37:1
22	26,35	48,07:1	22	22,16	27,98:1

Tabela 10.2: Distúrbio II: SNR e taxa de compressão (TC) para diferentes valores em (%) de T .

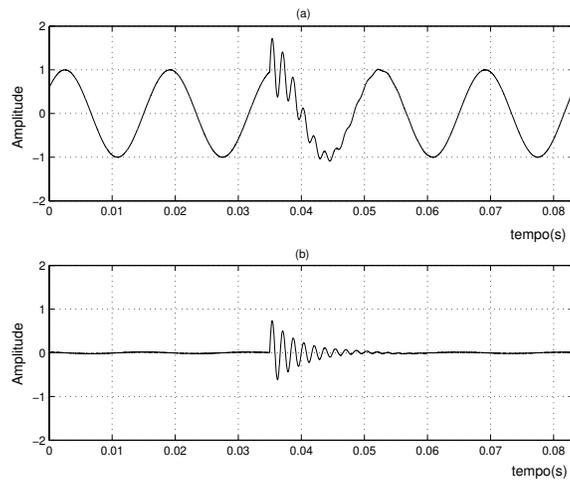


Figura 10.10: Distúrbio II: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$.

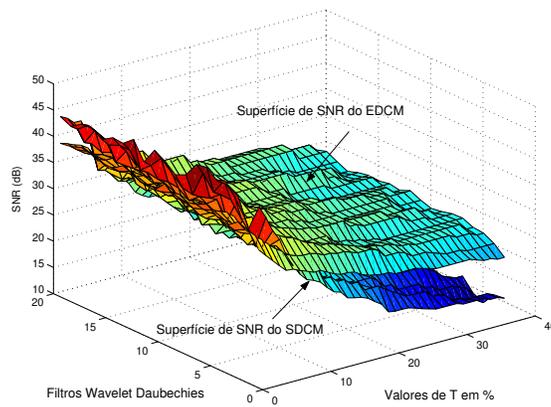


Figura 10.11: Distúrbio II: comparação de desempenho entre a SDCM e a EDCM.

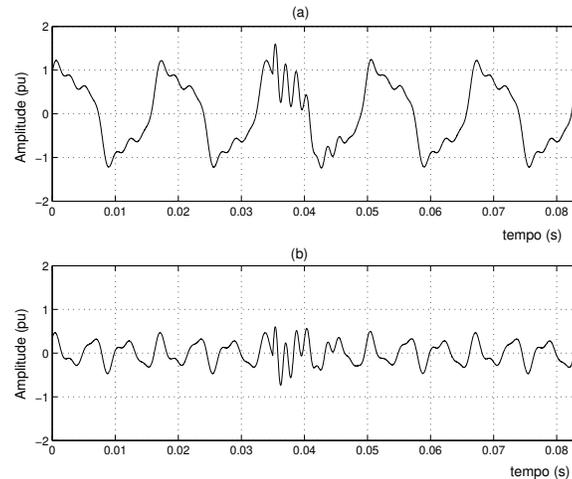


Figura 10.12: Distúrbio III: (a) sinal gerado $\{x(n)\}$ e (b) transitório $\{u(n)\}$.

O distúrbio III e o sinal erro são mostrados na Fig. 10.12. Por outro lado, a Fig. 10.13 mostra os mapas de SNR obtidos com a EDCM e a SDCM. A Tabela 10.3 mostra as taxas de compressão obtidas quando ambas as técnicas são aplicadas ao distúrbio III. Apesar da diferença de desempenho ser menor entre as propostas, a EDCM ainda apresenta melhor desempenho.

Para analisar o comportamento de ambas as técnicas em função do parâmetro $FHIR$, que indica a magnitude da presença das componentes harmônicas e inter-harmônicas, a tabela 10.4 mostra o desempenho de ambas as técnicas quando $T = 0,06 \max(\{u(n)\})$ e $FHIR$ assume vários valores. Conforme observado, quando $FHIR \rightarrow 0$, o desempenho da SDCM é melhor. Por outro lado, a EDCM apresenta melhores resultados quando $FHIR \rightarrow \infty$. Esse resultado sugere que para sinais cuja presença de componentes harmônicas e inter-harmônicas é representativa, as mesmas devem ser consideradas como componentes estacionárias e subtraídas do sinal. A seção 10.2 trata de uma nova técnica de compressão que leva essa questão em consideração.

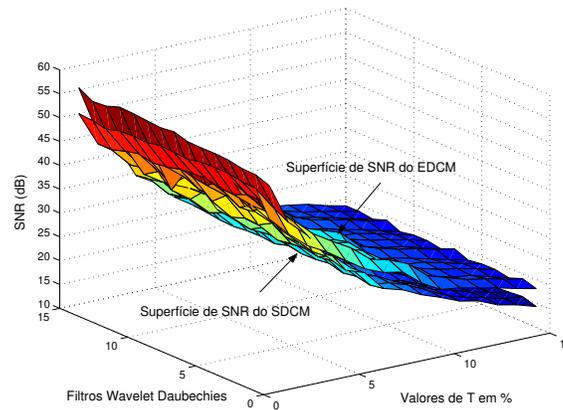


Figura 10.13: Distúrbio III: comparação de desempenho entre a SDCM e a EDCM.

EDCM			SDCM		
TWD com 1 níveis de decomposição			TWD com 3 níveis de decomposição		
$T(\%)$	SNR	TC	$T(\%)$	SNR	TC
2	47,02	21,23:1	2	51,56	5,78:1
6	35,09	25,54:1	6	37,29	21,60:1
10	30,16	32,44:1	10	30,48	26,62:1
14	26,69	33,75:1	14	27,29	28,42:1
18	26,41	33,90:1	18	25,97	28,42:1
22	23,97	34,21:1	22	22,93	29,07:1

Tabela 10.3: Distúrbio III: SNR e taxa de compressão (TC) para diferentes valores em (%) de T .

Finalmente, para completar a análise da EDCM e da SDCM, algumas análises numéricas foram realizadas sobre a influência da precisão finita no comportamento do sinal monitorado e dos bancos de filtros que implementam a TWD e a inversa da TWD. Para esta análise, considerou-se que a transformada z do sinal monitorado, $\{x(n)\}$, após a aplicação do processo de quantização é dado por

EDCM			SDCM		
TWD com 6 níveis de decomposição			TWD com 6 níveis de decomposição		
$FHIR(dB)$	SNR	TC	$FHIR(dB)$	SNR	TC
0,00	35,56	15,01:1	0,00	40,78	17,22:1
3,87	37,35	14,85:1	3,87	40,45	17,22:1
9,59	35,09	25,54:1	9,59	37,29	21,60:1
15,23	35,20	25,53:1	21,60	36,79	21,01:1
∞	41,76	16,45:1	∞	41,69	7,02:1

Tabela 10.4: Desempenho da EDCM e da SDCM em função dos valores do parâmetro $FHIR$

$$\tilde{X}(z) = \sum_{n=0}^{M-1} \tilde{x}(n)z^{-n} = \sum_{n=0}^{M-1} [x(n) + q(n)]z^{-n} = X(z) + \Delta X(z), \quad (10.21)$$

onde $\{\tilde{x}(n)\}$ é o sinal quantizado, $X(z)$ é a transformada z do sinal $\{x(n)\}$ e $\Delta X(z)$ é a transformada z do sinal diferença gerado pela quantização do sinal $\{x(n)\}$ e . Supondo que o erro de quantização é modelado como ruído branco, Gaussiano e aditivo, tem-se que o MSE, devido ao processo de quantização, é dado por (Mitra, 2001; Gray & Neuhoff, 1998)

$$MSE_{\tilde{X}} = -6R - 10.8dB, \quad (10.22)$$

onde R é o número de bits utilizados na quantização do sinal $\{x(n)\}$.

Considerando a quantização dos coeficientes de um banco de filtro de dois canais, cuja parte de análise implementa a TWD e a parte de síntese implementa a inversa da TWD, vide Fig. 10.14, as seguintes relações são obtidas

$$\tilde{H}(z) = \sum_{n=0}^{N-1} \tilde{h}(n)z^{-n} = H(z) + \Delta H(z) \quad (10.23)$$

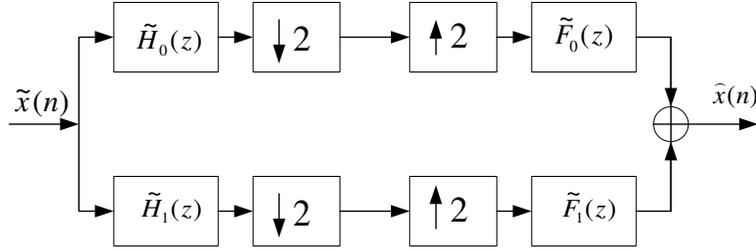


Figura 10.14: Banco de filtro de dois canais com filtros de análise e síntese quantizados.

$$\hat{X}(z) = \tilde{T}(z)\tilde{X}(z) + \tilde{A}(z)\tilde{X}(-z) \quad (10.24)$$

$$\tilde{T}(z) = T(z) + \Delta T(z) \quad (10.25)$$

$$\tilde{A}(z) = A(z) + \Delta A(z) \quad (10.26)$$

$$A(z) = \frac{1}{2} [H_0(-z)F_0(z) + H_1(-z)F_1(z)] \quad (10.27)$$

$$T(z) = \frac{1}{2} [H_0(z)F_0(z) + H_1(z)F_1(z)] \quad (10.28)$$

$$\begin{aligned} \Delta T(z) = \frac{1}{2} \{ & H_0(z)\Delta F_0(z) + \Delta H_0(z)[F_0(z) + \Delta F_0(z)] + \\ & + H_1(z)\Delta F_1(z) + \Delta H_1(z)[F_1(z) + \Delta F_1(z)] \} \end{aligned} \quad (10.29)$$

$$\begin{aligned} \Delta A(z) = \frac{1}{2} \{ & H_0(-z)\Delta F_0(z) + \Delta H_0(-z)[F_0(z) + \Delta F_0(z)] \\ & + H_1(-z)\Delta F_1(z) + \Delta H_1(-z)[F_1(z) + \Delta F_1(z)] \} \end{aligned} \quad (10.30)$$

onde $\hat{X}(z)$ é o sinal $\tilde{X}(z)$ reconstruído. Para que a análise seja representativa é necessário representar o erro de *round-off* (Mitra, 2001), que resulta da incapacidade de representação do resultado de uma operação aritmética devido à precisão finita utilizada. No caso do banco de filtros mostrados na Fig. 10.14, o erro de *round-off* é dado por

$$e_r(n) = e_a(n) + e_s(n), \quad (10.31)$$

onde $e_a(n)$ e $e_s(n)$ são, respectivamente, os erros de *round-off* (Mitra, Hirano, & Sakaguchi, 1974) gerados pelos bancos de análise e de síntese. Assim sendo, o erro total devido a precisão finita é, então, dado por

$$e_t(n) = e_q(n) + e_r(n) = e_q(n) + e_a(n) + e_s(n), \quad (10.32)$$

onde $e_q(n)$ é o erro de quantização. Logo, o sinal $\{x(n)\}$ é dado por

$$\bar{x}(n) \cong \hat{x}(n) + \Delta x(n) + e_t(n) \quad (10.33)$$

A Tab. 10.5 apresenta alguns resultados relacionados à influência da precisão finita, quando o transitório I, vide Fig. 10.8, é submetido às SDCM e EDCM. A EDCM apresenta os melhores resultados. Na tabela 10.5, a precisão 12/8 bits significa que o sinal monitorado foi quantizado com 12 bits e os filtros de análise e síntese com 8 bits. Já a tabela 10.6, mostra os resultados obtidos quando ambas as soluções foram aplicadas ao distúrbio II, vide Fig. 10.10. Novamente, o resultado obtido com a proposta EDCM foi superior em termos da taxa de compressão. Nas simulações realizadas para a obtenção dos valores mostrados nas tabelas 10.5 e 10.6 foram considerados apenas 4 ciclos do sinal $\{x(n)\}$.

SDCM			EDCM		
TWD com 1 níveis de decomposição			TWD com 3 níveis de decomposição		
<i>Bits</i>	<i>SNR</i>	<i>TC</i>	<i>Bits</i>	<i>SNR</i>	<i>TC</i>
12/8	35,37	18,34:1	12/8	35,08	31,19:1
12	34,81	13,39:1	12	35,45	23,81:1
16	34,74	11,10:1	16	35,37	19,75:1
32	34,71	11,09:1	32	35,35	19,81:1

Tabela 10.5: Influência da precisão finita nos resultados obtidos com a SDCM e a EDCM.

SDCM			EDCM		
TWD com 1 níveis de decomposição			TWD com 3 níveis de decomposição		
<i>Bits</i>	<i>SNR</i>	<i>TC</i>	<i>Bits</i>	<i>SNR</i>	<i>TC</i>
12/8	35,08	25,49:1	12/8	35,68	36,57:1
12	35,90	12,11:1	12	35,78	24,57:1
16	36,60	10,21:1	16	35,85	19,44:1
32	35,82	11,25:1	32	35,90	19,50:1

Tabela 10.6: Influência da precisão finita nos resultados obtidos com a SDCM e a EDCM.

10.2 *Fundamental, Harmonic and Transient Compression Method (FHTCM)*

A EDCM, discutida na seção 10.1, faz uso de diferentes técnicas de processamento de sinais para a compressão de sinais de potência. No entanto, modificações podem ser introduzidas de forma a melhorar o desempenho da mesma. De fato, conforme observado na tabela 10.4, a presença de componentes harmônicas de elevada potência no sinal monitorado deteriora o desempenho da EDCM. Uma outra questão bastante

relevante é a existência de um número consideravelmente grande de *wavelets* e de possíveis bases de representação. Além disso, a questão da quantização por sub-banda ainda não foi muito bem tratada por nenhuma das contribuições citadas no início da seção 10.1.

Considerando essas questões para a formalização de um novo método para compressão de sinais de potência, a presente seção introduz a *fundamental, harmonic, and transient separation compression method* (FHTCM), proposta em (Ribeiro, Mitra, & Romano, 2004a). As principais características dessa técnica são as seguintes: *i*) a FHTCM generaliza as idéias propostas na EDCM ao fazer uso da técnica NW-WDFT, vide Capítulo 7, (Ribeiro, Mitra, & Romano, 2004b) para a detecção, a estimação e a subtração das componentes fundamental e harmônicas do sinal monitorado; *ii*) a FHTCM introduz um novo critério MDL (*minimum description Length*) (Rissanen, 1978; A. Barron & Yu, 1998) que leva em consideração um dicionário de wavelets, as possíveis bases de representação do sinal monitorado e a quantização escalar dos coeficientes em cada sub-banda. Uma versão parcial da FHTCM foi apresentada em (Ribeiro et al., 2004c).

A formulação a ser apresentada, a seguir, é baseada na definição do sinal monitorado dada por (6.1)-(6.7). A seguinte definição será utilizada para descrever o FHTCM:

Sejam

$$s(n) := f(n) + h(n) \tag{10.34}$$

e

$$u(n) := i(n) + t(n), \tag{10.35}$$

então (6.1) é dado por

$$x(n) = s(n) + u(n) + v(n). \tag{10.36}$$

As componentes inter-harmônicas foram definidas como sinais transitórios, pois apresentam potência baixa em relação as componentes fundamental e harmônicas, as quais ocorrem de forma aleatória no sinal monitorado. Dada tal suposição, objetivava-se encontrar a *wavelet*, a decomposição em sub-bandas e a quantização escalar que minimizem o número de bits necessários para representar o sinal monitorado, dado que um certo nível de distorção, inserido pelo processo de codificação, seja especificado *a priori*.

A idéia por trás do objetivo proposto é o princípio *dividir e conquistar*, discutido no início do capítulo 6. De fato, devido as características do sinal (6.1), o mesmo pode ser decomposto em sinais determinísticos (10.34) e sinais transitórios (10.35). Conseqüentemente, o uso de diferentes técnicas de compressão de sinais, que possam levar em consideração as características dos sinais determinísticos e dos sinais transitórios, podem ser exploradas para a obtenção de melhores desempenhos. Por exemplo, diferentemente dos sinais de voz, música, imagem e vídeo, as funções senoidais podem ser eficientemente armazenadas com a determinação e quantização de seus parâmetros (amplitude, fase e frequência).

Conforme já discutido na seção 10.1, as componentes transitórias do sinal monitorado podem ser bem representadas pela TW. Essa solução é bastante interessante, pois não apenas o sinal monitorado pode ser codificado com poucos bits, mas também os parâmetros das componentes harmônicas podem ser utilizadas para calcular os índices QE. Note que nenhuma das técnicas de compressão desenvolvidas até a presente data apresentam tal vantagem (Hamid & Kawasaki, 2002; Santoso et al., 1997; Hsieh & Huang, 2003; Littler & Morrow, 1999; Gerek & Ece, 2004; Panda et al., 2002; Hsieh & Huang, 2003; Dash et al., 2003; Meher et al., 2004; Ribeiro et al., 2004; Chung et al., 1999; Duque et al., 2000).

Pode-se dizer que a separação do sinal $\{s(n)\}$ de $\{x(n)\}$ é equivalente a estimação

e a subtração da componente $\{s(n)\}$ do sinal $\{x(n)\}$. Note que de forma geral, $\sigma_s^2 \gg \sigma_u^2$, onde σ_s^2 é a variância de $\{s(n)\}$ e σ_u^2 é a variância de $\{u(n)\}$. Neste caso, a distorção devida a compressão do sinal $\{x(n)\}$ depende, principalmente, do erro de estimação dos parâmetros senoidais. Assim sendo, o desenvolvimento de uma técnica eficiente para a detecção e a estimação dos parâmetros das componentes fundamental e harmônicas se faz necessário.

Para a detecção e a estimação dos sinais senoidais várias técnicas foram propostas para uso em sistemas de potência, dentre as quais pode-se citar a DFT, o phase-locked loop (PLL), o filtro de Kalman e o filtro notch adaptativo (Ribeiro et al., 2004; Mitra, 2001; Gharieb, 2000; K.-Ghartemani & Iravani, 2003; Testa, Gallo, & Langella, 2004; Routray, Pradhan, & Rao, 2002; Liu, 2000). No entanto, devido às razões apresentadas no capítulo 7, a técnica NF-WDFT é aqui adotada para a estimação dos parâmetros das componentes fundamental e das harmônicas, pelas seguintes razões: *i*) a técnica NF-WDFT não necessita do uso de longas seqüências para a estimação dos parâmetros das componentes fundamental e harmônicas e *ii*) a técnica NF-WDFT oferece estimativas mais precisas dos parâmetros dos sinais senoidais que a técnica DFT, amplamente utilizada para a análise espectral do sinais de sistemas de potência.

A Fig. 10.15 mostra o diagrama de blocos da técnica NF-WDFT versão I, reproduzido aqui por questão de conveniência. Os parâmetros das componentes fundamental e harmônicas são estimados a partir do vetor de amostras do sinal monitorado $\{x(0), \dots, x(N-1)\}$. Note que os blocos NF_m ($0 \leq m \leq M$) representam os filtros notch que eliminam a m -ésima componente harmônica cuja frequência é dada por mf_0 .

Como os distúrbios ocorrem em intervalos irregulares e ocupam uma ampla faixa do espectro, o FHTCM é baseado num esquema de compressão adaptativa. De

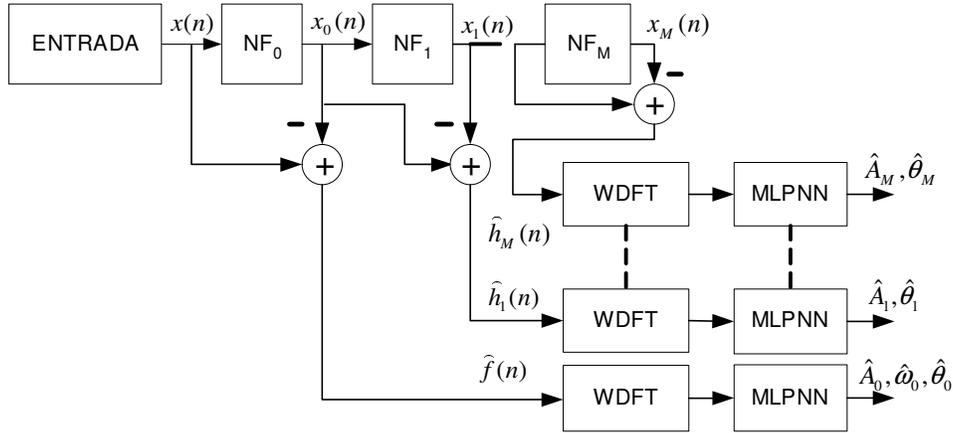


Figura 10.15: Técnica NF-WDFT I.

fato, o FHTCM faz uso das características do sinal $\{x(n)\}$ durante o processo de codificação para que a compressão ótima seja obtida de acordo com o critério MDL (*minimum description length*) proposto na seção 10.2.3.

Para discutir os detalhes da FHTCM, as seções 10.2.1, 10.2.2, 10.2.3 discutem o critério adotado para a subtração das componentes harmônicas do sinal $\{x(n)\}$, a estratégia aplicada para a compressão e remoção de ruído (*denoising*) e o critério MDL, respectivamente.

10.2.1 Análise das Componentes Harmônicas

Na Fig. 10.16, O bloco **K** é responsável pela detecção da presença das componentes harmônicas. Para executar essa tarefa, os parâmetros estimados das componentes fundamental e harmônicas, que foram previamente estimados pela técnica NF-WDFT, são utilizados. Baseadas nessas informações, as estimativas das formas de onda das componentes fundamental e harmônicas são geradas de acordo com os algoritmos propostos em (Mitra, 2001; Cheng, 2000) e, finalmente,

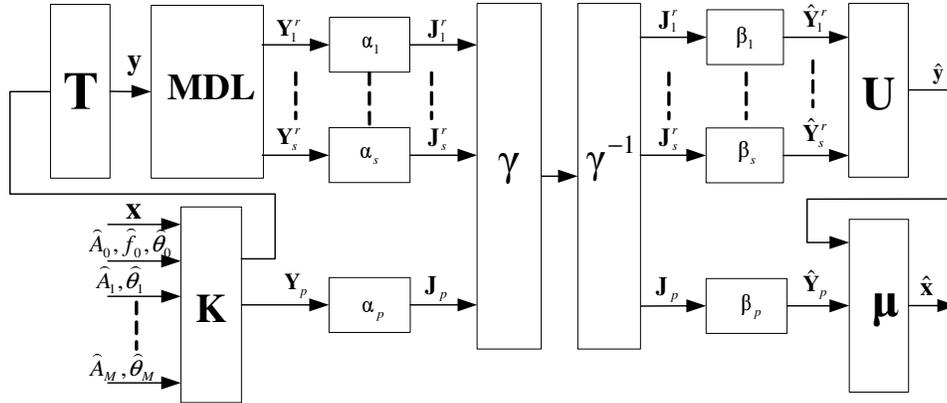


Figura 10.16: Diagrama de blocos da FHTCM.

as estimativas das componentes fundamental e harmônicas são subtraídas do sinal $\{x(n)\}$, de forma que apenas o sinal $\{u(n)\}$, adicionado de sinal erro $\{e(n)\}$ gerado pelo processo de subtração, é submetido à codificação por transformada baseada na TW. O critério usado para decidir quais as componentes harmônicas devem ser subtraídas do sinal $\{x(n)\}$ é dado pelo parâmetro *fundamental to harmonic ratio* (FHR), uma simplificação do parâmetro FHIR dado por (10.8). O parâmetro *FHR* associado à m -ésima componente harmônica, FHR_m , é definido por

$$FHR_m := 10 \log_{10} \left[\frac{\sum_{n=0}^{N-1} \{\tilde{f}(n)\}^2}{\sum_{l=2}^m \sum_{n=0}^{N-1} \{\tilde{h}_l(n)\}^2} \right], \quad (10.37)$$

onde $\tilde{f}(n)$ e $\tilde{h}_l(n)$ representam as estimativas das componentes $f(n)$ e $h_l(n)$, respectivamente.

Baseando-se em simulações numéricas, desempenhos satisfatórios foram obtidos quando os seguintes passos foram considerados para o uso de (10.37):

Passo 1

Faça $m = 2$. Ir para o passo 2.

Passo 2

Se $FHR_m \geq T_{FHR}$, onde T_{FHR} é o limiar escolhido. Então as m primeiras componentes harmônicas são extraídas do sinal $\{x(n)\}$ e o sinal $\{u(n)\}$ resultante é submetido à técnica de compressão por transformada. O algoritmo é finalizado

Senão, ir para o passo 3

Passo 3

Faça $m = m + 1$.

Se $m > M$, então o algoritmo é finalizado

Senão ir para o Passo 2.

Nas simulações, observou-se que para T_{FHR} igual a 40, resultados satisfatórios são obtidos. Note que o valor 40 é bastante conservador, no entanto, o mesmo funciona muito bem para uma gama muito de grande de distúrbios nas redes de transmissão e distribuição de energia.

O bloco MDL implementa o critério MDL discutido na seção 10.2.3.

Na Fig. 10.16, \mathbf{Y}_p representa o vetor que contém os parâmetros estimados das componentes fundamental e harmônicas, obtidas com a técnica NF-WDFT, α_p e β_p representam a quantização escalar e sua correspondente inversa, as quais são aplicadas ao vetor de parâmetros \mathbf{Y}_p .

10.2.2 Estratégia para a Compressão dos Transitórios e a Redução do Ruído de Fundo

Na Fig. 10.16, o bloco **T** executa a TWD. Note que é suposto o uso de um dicionário de wavelet dado por $\Psi := \{\Psi_r | r = 1, \dots, R\}$, onde R é o número de wavelets no dicionário.

Sejam $\hat{\mathbf{s}} = (\hat{s}_1, \dots, \hat{s}_N)^T$ a aproximação de $\mathbf{s} = (s_1, \dots, s_N)^T$, onde $s_n = s(n-1)$ e \hat{s}_n é dado por

$$\hat{s}_n := \sum_{m=1}^M \hat{A}_m \cos\left(2\pi m \hat{f}_1(n-1) + \hat{\theta}_m\right), 1 \leq n \leq N, . \quad (10.38)$$

Em (10.38), \hat{f}_1 , \hat{A}_m e $\hat{\theta}_m$ são os valores quantizados de \tilde{f}_1 , \tilde{A}_m e $\tilde{\theta}_m$, respectivamente. M é o número de componentes senoidais obtidas com o algoritmo implementado no bloco **K**.

Agora, seja $\mathbf{x} := (x_1, \dots, x_N)^T$ o vetor que representa a seqüência $\{x(n)\}$, onde $x_n = x(n-1)$. Definindo $\mathbf{y} := \mathbf{x} - \hat{\mathbf{s}}$ e $\mathbf{e} := \mathbf{s} - \hat{\mathbf{s}}$, então

$$\mathbf{y} = \mathbf{u} + \mathbf{v} + \mathbf{e}. \quad (10.39)$$

Logo, os coeficientes do sinal \mathbf{y} no domínio da TWD são representados por Ψ_r . Para a decomposição obtida com a TWD baseada na r -ésima wavelet do dicionário, temos que

$$\mathbf{Y}^r = \mathbf{W}_r \mathbf{y}, \quad (10.40)$$

$$= \mathbf{U}^r + \mathbf{V}^r + \mathbf{E}^r, \quad (10.41)$$

onde $\mathbf{W}_r \in \mathbf{R}^{N \times N}$ é a matriz de transformação associada à wavelet Ψ_r contida no dicionário de wavelets, $\mathbf{U}^r = \mathbf{W}_r \mathbf{u}$, $\mathbf{V}^r = \mathbf{W}_r \mathbf{v}$ e $\mathbf{E}^r = \mathbf{W}_r \mathbf{e}$.

Agora, seja $\mathbf{Y}_s^r := (Y_{s,1}^r, \dots, Y_{s,N_{r,s}}^r)^T$, $s = 1, \dots, s_d$ o vetor que representa o sinal na s -ésima sub-banda obtida com a d -ésima decomposição gerada com a TWD, onde s_d é o número de sub-bandas da d -ésima decomposição e $1 \leq s \leq s_d$. Note que $\sum_{s=1}^{s_d} N_{r,s} = N$, $\forall r = 1, 2, \dots, k$ e, conseqüentemente, $\mathbf{Y}^r = (\mathbf{Y}_1^r, \dots, \mathbf{Y}_{s_d}^r)^T$.

O bloco α_s , $1 \leq s \leq s_d$ é responsável pela quantização dos coeficientes na s -ésima sub-banda. Normalmente, os métodos de compressão por transformada fazem uso da técnica *uniform threshold quantization* (UTQ) (He & Mitra, 2001) e a *dead-zone* corresponde ao processo de *thresholding* que implementa o *denoising* (Chang, Yu, & Vitterli, 2000; Donoho, 1995; Donoho & Johnstone, 1994; Vidakovic, 1995; Donoho & Johnstone, 1995). No FHTCM, o UTQ é utilizado para quantizar os coeficientes no domínio da TWD.

Para $Y_{s,i}^r$, $1 \leq i \leq N_{r,s}$, $\alpha_s : \mathbf{R} \mapsto \mathbf{J}_s^r$, o UTQ realiza o mapeamento pelo qual uma partição de \mathbf{R} é mapeada num dado alfabeto representado por \mathbf{J}_s^r . Note que \mathbf{R} é particionado em células dadas por $S_j = \{Y \in \mathbf{R} | \alpha_s(Y) = j\}$, $j \in \mathbf{J}_s^r$. Dadas estas definições, temos que α_s é definido por

$$\alpha_s(\Delta_s, T_s; Y) = \begin{cases} 0, & \text{Se } |Y| \leq T_s \\ \left\lfloor \frac{Y - T_s}{\Delta_s} \right\rfloor, & \text{Se } Y > T_s \\ \left\lfloor \frac{Y + T_s}{\Delta_s} \right\rfloor, & \text{Se } Y < -T_s \end{cases}, \quad (10.42)$$

onde Δ_s é o *step size* and T_s é o *dead-zone threshold* da s -ésima sub-banda. Agora, seja q_s^r o número de intervalos de quantização do lado positivo de Y , então o número

total de intervalos de quantização é dado por $(2q_s^r + 1)$, pois existe simetria na definição do *dead-zone* do UTQ. No FHTCM foi considerado que $T_s = \hat{\sigma}_v^2$, onde $\hat{\sigma}_v^2$ é a estimativa do variância do ruído de fundo $\{v(n)\}$.

O dequantizador $\beta_s : \mathbf{J}_s^r \mapsto \mathbf{R}$ é especificado pelo dicionário de palavras definido por $C_s^r = \{\hat{Y}_{s,j}^r\}_{j \in \mathbf{J}_{s^r}}$. O valor ótimo de $\hat{Y}_{s,j}^r$ é o centróide da região de quantização que depende da distribuição de probabilidade dos coeficientes, dado por $Y_{s,j}^r$, no domínio da TWD, .

Os blocos γ and γ^{-1} representam o codificador e decodificador de entropia. O projeto dos codificadores e decodificadores de entropia está fora do escopo desta seção. No entanto, sabe-se que existem diferentes técnicas de codificação por entropia, dentre as quais podemos citar o algoritmo LZW (Lempel-Ziv-Welch), a codificação de Huffman e a codificação aritmética (Gersho & Gray, 1999), utilizadas em aplicações de monitoramento da QE (Ribeiro et al., 2004; Hsieh & Huang, 2003).

O bloco \mathbf{U} executa a inversa da TWD dada por \mathbf{W}_r para reconstruir \mathbf{y} . Observe que a reconstrução de \mathbf{y} é dada por

$$\hat{\mathbf{y}} = \mathbf{W}_r^T \left(\hat{\mathbf{Y}}_1^r, \dots, \hat{\mathbf{Y}}_{s_d}^r \right)^T. \quad (10.43)$$

Finalmente, o bloco $\boldsymbol{\mu}$ reconstrói o sinal monitorado que é dado por

$$\hat{\mathbf{x}} = \hat{\mathbf{s}} + \hat{\mathbf{y}}. \quad (10.44)$$

10.2.3 Critério MDL

O critério MDL proposto por Rissanen (Rissanen, 1978) diz que a melhor representação de um sinal qualquer é aquela que minimiza os números de bits utilizados para representar o sinal e os parâmetros desta representação. Várias

contribuições mostram que esse critério fornece resultados relevantes em aplicações de compressão de sinais e remoção de ruídos aditivos (A. Barron & Yu, 1998; Krim & Schick, 1999; Hansen & Yu, 2000; Chang et al., 2000; Krim & Schick, 1999; Krim, Tucker, Mallat, & Donoho, 1999).

Para a compressão de distúrbios em sistemas de potência, uma versão do critério MDL proposto por Saito (Saito, 1994) foi aplicada para sinais de potência em (?). No entanto, tal contribuição não levou em consideração a quantização e a extração das componentes harmônicas do sinal $\{x(n)\}$. Visando levar em consideração esses dois aspectos, um critério MDL derivado de (Chang et al., 2000) para a compressão e *denoising* de sinais de potência é proposto nessa seção. Tal critério faz uso das seguintes considerações em sua formulação: *i*) a existência de um dicionário de wavelets, *ii*) a decomposição adaptativa, *iii*) a quantização escalar em subbandas, baseada no $UTQ(\Delta, T; \cdot)$, *iv*) a reconstrução dos coeficientes quantizados, considerando-se a distribuição Gaussiana generalizada (*Generalized Gaussian Distribution* - GGD) com a distribuição que descreve os coeficientes nas subbandas. Além disso, o ruído de fundo, $v(n)$, é modelado como uma variável aleatória *i.i.d.*, cuja distribuição é Gaussiana σ^2 , $\mathcal{N}(0, \sigma^2)$. Os coeficientes no domínio da transformada do sinal $\{v(n)\}$, que são representados por $\{V^r(n)\}$, também são *i.i.d.* com distribuição Gaussiana $\mathcal{N}(0, \sigma^2)$, de tal forma que $Y_{s,i}^r | U_{s,i}^r \sim \mathcal{N}(U_{s,i}, \sigma^2)$.

Supondo-se que o sinal $\{s(n)\}$ é perfeitamente estimado e quantizado com um número adequado de bits, então $\sigma_e^2 \ll \sigma^2$, sendo σ_e^2 a variância do erro de estimação do sinal $\{s(n)\}$, então a expressão (10.41) é dado por

$$\mathbf{Y}^r \simeq \mathbf{U}^r + \mathbf{V}^r. \quad (10.45)$$

A complexidade necessária para descrever \mathbf{Y}^r e os seus parâmetros é expressa

pela soma dos comprimentos, em número de bits, dos seguintes termos: *i*) de dois números inteiros para armazenar a wavelet escolhida e a estrutura de decomposição (r, d) , *ii*) de $2s_d$ números reais para armazenar os valores do *threshold* e *step size* (T_s^r, Δ_s^r) para as s_d sub-bandas, *iii*) de s_d números inteiros para armazenar os intervalos de quantização q_s^r e *iv*) dos índices dos coeficientes quantizados no domínio da TWD.

Baseando-se nas considerações acima, o princípio MDL implementado para encontrar a representação de \mathbf{U}^r que minimiza o número de bits necessário para representar o sinal $\{u(n)\}$ e a presença de ruído aditivo é dado por

$$L(\mathbf{y}, \hat{\mathbf{U}}, \boldsymbol{\theta}, r, d) = \sum_{s=1}^{s_d} \left\{ L(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d) + L(\mathbf{Y}_s^r | \hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d) \right\}, \quad (10.46)$$

onde $\boldsymbol{\theta} := (T_s^r, \Delta_s^r, q_s^r)$, $L(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$ é o comprimento do código necessário para representar $(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$ e $L(\mathbf{Y}_s^r | \hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$ e o comprimento de código utilizado para representar \mathbf{Y}_s^r , supondo-se $(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$.

Para facilitar o cálculo de (10.46), a restrição do comprimento do código ser inteiro é aliviada, de forma que o comprimento dado por (10.46) é real. Então, $L(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$ em (10.46) é dado por

$$L(\hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d) = L(r, d) + L(\hat{\boldsymbol{\theta}} | r, d) + L(\hat{\mathbf{U}}_s^r | \boldsymbol{\theta}, r, d). \quad (10.47)$$

Note que em (10.47), o termo $L(r, d)$ pode ser ignorado no processo de minimização pois é constante. Note ainda que $L(\boldsymbol{\theta} | r, d)$ é dado por

$$L(\boldsymbol{\theta} | r, d) = L(q_s^r, T_s^r, \Delta_s^r | r, d) \quad (10.48)$$

$$= L(T_s^r, \Delta_s^r | r, d) + L(q_s^r | T_s^r, \Delta_s^r, r, d) \quad (10.49)$$

$$= s_d(2b_f + b_i), \quad (10.50)$$

onde b_f e b_i representam os números de bits necessários para codificar um número real e um número inteiro, respectivamente. Logo, $L(\hat{\mathbf{U}}_s^r | \boldsymbol{\theta}, r, d)$ em (10.47) é dado por

$$L(\hat{\mathbf{U}}_s^r | \boldsymbol{\theta}, r, d) = - \sum_{k=-q_s}^{q_s} P_{r,s,k} \log P_{r,s,k}, \quad (10.51)$$

sendo

$$P_{r,s,k} \simeq \frac{N_{r,s,k}}{N_{r,s}}, \quad (10.52)$$

$N_{r,s,k}$ o número de ocorrências de $\hat{Y}_{s,k}^r$ e $N_{r,s} = \sum_{k=-q_s}^{q_s} N_{r,s,k}$.

Agora pode-se dizer que $L(\mathbf{Y}_s^r | \hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d)$ em (10.46) é dado por

$$L(\mathbf{Y}_s^r | \hat{\mathbf{U}}_s^r, \boldsymbol{\theta}, r, d) = - \sum_{i=1}^{N_s} \log p(Y_{s,i}^r | \hat{U}_{s,i}^r) \quad (10.53)$$

$$= \frac{1}{2\sigma^2 \log 2} \sum_{n=1}^{N_{r,s}} \left(Y_{s,i}^r - \hat{U}_{s,i}^r \right)^2 + \frac{1}{2} \log (2\pi\sigma^{2N_{r,s}}), \quad (10.54)$$

onde $\hat{U}_{s,i}^r \in C_s^r$ representa a quantização do coeficiente $U_{s,i}^r$. O segundo termo de (10.54) é igual a uma constante e, então, pode ser ignorado durante o processo de minimização.

Baseando-se nas deduções acima, o critério MDL proposto para selecionar o modelo que fornece o menor comprimento de código para descrever o sinal $\{x(n)\}$ é dado por

$$\begin{aligned}
\text{MDL}(\mathbf{y}, \hat{\mathbf{U}}, \boldsymbol{\theta}, r, d) &= \frac{1}{2\sigma_v^2 \log 2} \sum_{s=1}^{s_d} \sum_{i=1}^{N_{r,s}} \left(Y_{s,i}^r - \hat{U}_{s,i}^r \right)^2 \\
&\quad - \sum_{s=1}^{s_d} \frac{1}{N_{r,s}} \sum_{k=-q_s}^{q_s} N_{r,s,k} \log \frac{N_{r,s,k}}{N_{r,s}} + s_d(2b_f + b_i).
\end{aligned} \tag{10.55}$$

A implementação do critério dado por (10.55) é baseada na estratégia *bottom-up* (Ramchandran, Vitterli, & Herley, 1996), pois tal estratégia, assumindo que cada uma das possíveis sub-bandas do sinal é associada a um custo dado por (10.46), faz uso do seguinte critério para a escolha das sub-bandas que contribuem para minimizar (10.55): se a soma dos custos das duas sub-bandas *children* é menor que o custo da sub-banda *parent*, então as duas sub-bandas *children* são selecionadas. Caso contrário, a sub-banda *parent* é selecionada e as duas sub-bandas *children* são descartadas.

Sabendo-se que o critério proposto por Saito (Saito, 1994) não leva em consideração a quantização do sinal e que Chang não leva em consideração a busca pela wavelet e a decomposição em sub-bandas que contribuam para minimizar o valor do critério proposto em (Chang et al., 2000), pode-se dizer que o critério proposto acima é mais geral que ambos. De fato, o critério proposto leva em consideração em sua formulação a quantização do sinal, a busca pela wavelet e a decomposição que minimizem o comprimento do sinal.

10.2.4 Resultados Numéricos

Para analisar o desempenho do FHTCM quando aplicado para a compressão de sinais de potência, esta seção analisa 5 tipos de distúrbios em sistemas de potência. Os três primeiros casos foram sinteticamente gerados para representar alguns típicos

distúrbios em sistemas de potência, considerando-se $f_s = 60 \times 512$ Hz e 16 bits para a quantização da amplitude. Os dois últimos distúrbios foram obtidos da base de dados do IEEE PES Working Group P1433 Power Quality website (PES, 2003).

Os seguintes parâmetros foram utilizados nas simulações mostradas: *i*) o número de bits utilizados para a quantização dos parâmetros estimados é 24, *ii*) O número máximo de estágios de decomposição é 4, *iii*) o dicionário de wavelet é constituída das wavelets Symlet, Daubechies e Meyer, *iv*) os números de bits usados para quantizar o *dead-zone threshold* e o *step size* do UTQ são iguais a 24, *v*) o número de bits necessários para codificar o número de regiões de quantização em cada subbanda é 8 e *vi*) o número de bits utilizado para codificar a decomposição selecionada é 6.

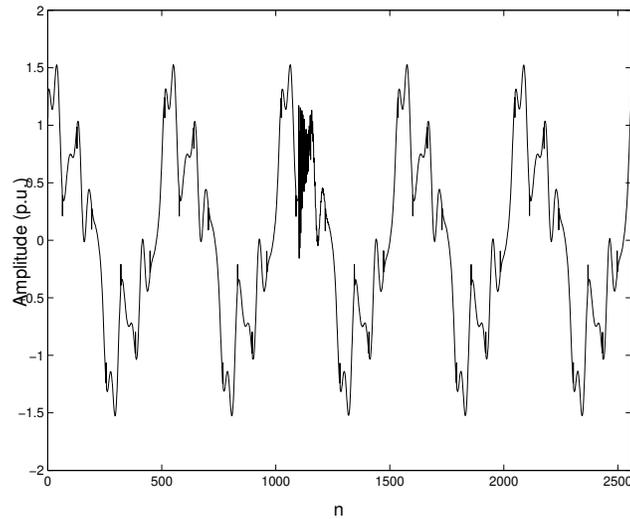
Para verificar o desempenho do FHTCM, todos os distúrbios foram submetido a outros 4 métodos de compressão de sinais de potência: SDCM-MDL (SDCM + critério MDL proposto), EDCM-MDL (EDCM + critério MDL proposto), SAITO (Saito, 1994; ?, ?), EDCM (Ribeiro et al., 2004) e SDCM (Ribeiro et al., 2004).

10.2.4.1 Distúrbio I

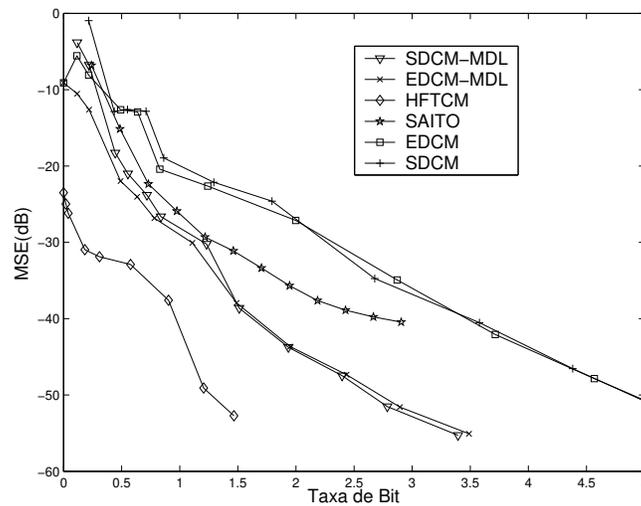
A Fig. 10.17(a) mostra o distúrbio gerado de acordo com (Moore & Português, 2003), o qual apresenta um número significativo de componentes harmônicas. Já a Fig. 10.17(b) compara o desempenho dos 6 métodos de compressão de sinais em termos dos valores MSE e das taxas de bit. A potência do ruído de fundo é dada por $\sigma_v^2 = -50$ dB. Observando a Fig. 10.17(b), fica evidente que o FHTCM apresenta resultados superiores aos alcançados pelos outros métodos. A estimação e a subtração de todas as componentes harmônicas combinadas com o critério MDL são os principais responsáveis por esse desempenho superior.

Comparando as curvas de MSE do SDCM-MDL e EDCM-MDL com o SDCM e

EDCM, pode-se notar que o critério MDL melhora o desempenho do EDCM e do SDCM. Esse é um resultado muito interessante, pois demonstra que o critério MDL proposto é efetivo não apenas no contexto do FHTCM, mas também quando usado com outras técnicas.



(a) Sinal analisado.



(b) Desempenho dos algoritmos de compressão.

Figura 10.17: Sinal sintético I: análise de desempenho.

10.2.4.2 Distúrbio II

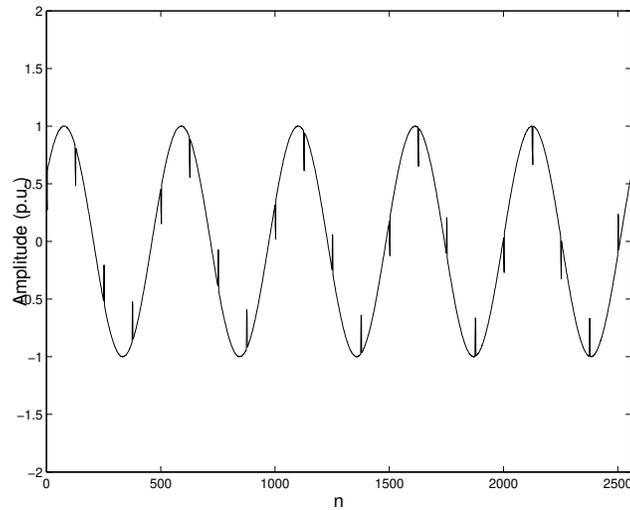
A Fig. 10.18(a) mostra um sinal corrompido por *notchs* impulsivos e ruído de fundo cuja $\sigma_v^2 = -50$ dB. A análise é apropriada, pois notch impulsivos são bastantes comuns devidos aos conversores CA-CC. A Fig. 10.18(b) mostra os valores MSE gerados pelos seis métodos de compressão para diferentes taxas de bits. As curvas de MSE mostram resultados similares aos ilustrados na Fig. 10.17(b), quando se compara o desempenho dos seis métodos.

10.2.4.3 Distúrbio III

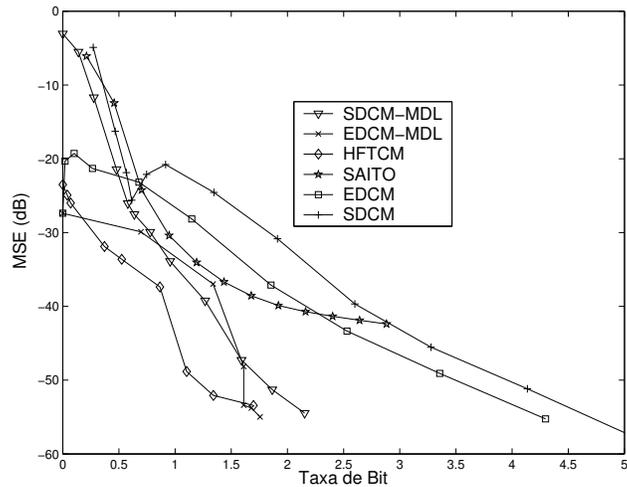
A Fig. 10.19(a) mostra um sinal contendo distúrbio em banda estreita (chaveamento capacitivo) e em banda larga (notchs impulsivos) e ruído de fundo com $\sigma_v^2 = -50$ dB. A Fig. 10.19(b) mostra as curvas de MSE para os vários valores de taxas de bits dos seis métodos de compressão. Como as componentes harmônicas não são observadas nesse distúrbio, o desempenho do EDCM-MDL e FHTCM são bastante similares.

10.2.4.4 Distúrbio IV

Enquanto a Fig. 10.20(a) mostra o distúrbio contido no banco de dados disponível no site (PES, 2003), a Fig. 10.20(b) mostra a curva MSE para as várias taxas de bits geradas pelos 6 métodos de compressão. Conforme pode ser observado, o FHTCM novamente apresenta melhores resultados que os outros 5 métodos.



(a) Sinal analisado.

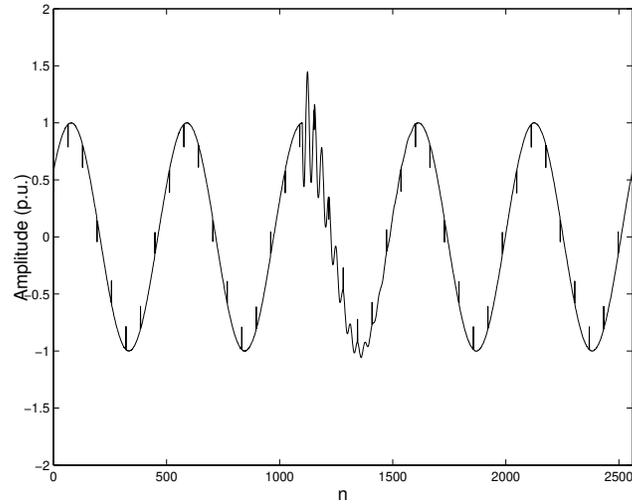


(b) Desempenho dos algoritmos de compressão.

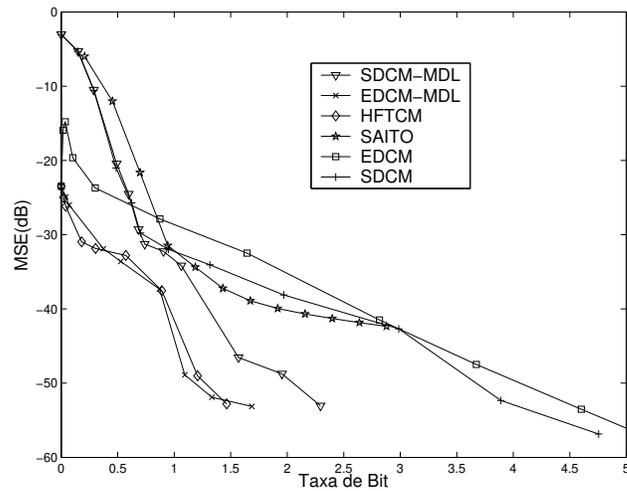
Figura 10.18: Sinal sintético II: análise de desempenho.

10.2.4.5 Distúrbio V

Finalmente, a Fig. 10.21(a) mostra outro distúrbio obtido em (PES, 2003). A Fig. 10.21(b) mostra as curvas de valores MSE produzidas pelos 6 algoritmos



(a) Sinal analisado.

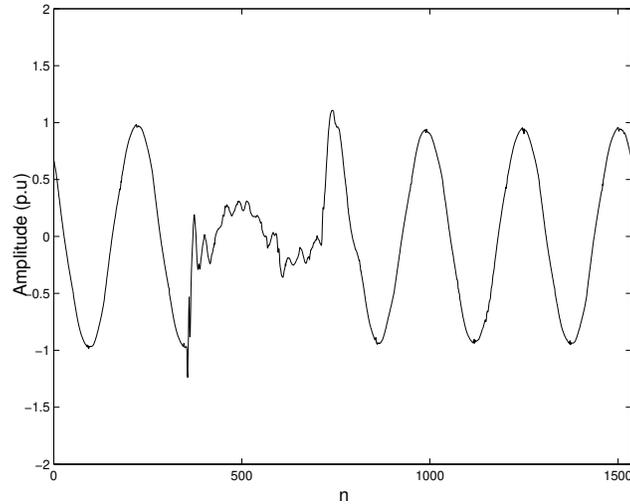


(b) Desempenho dos algoritmos de compressão.

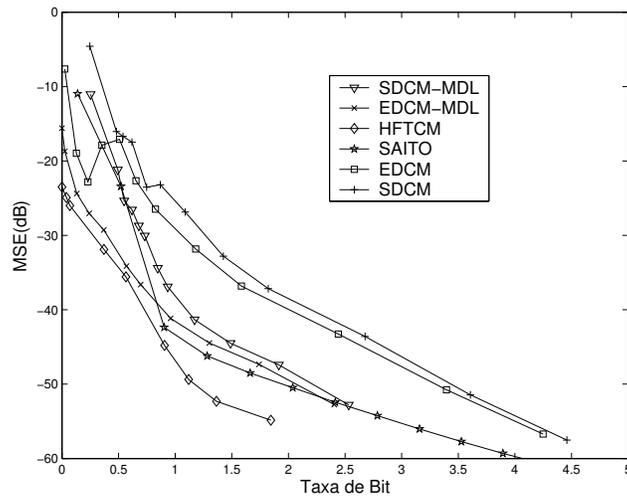
Figura 10.19: Sinal sintético III: análise de desempenho.

de compressão para diferentes taxas de bits. É interessante notar que o FHTCM apresenta melhores resultados que os outros métodos aqui considerados.

Baseando-se nos resultados de simulações, pode-se observar que o método de



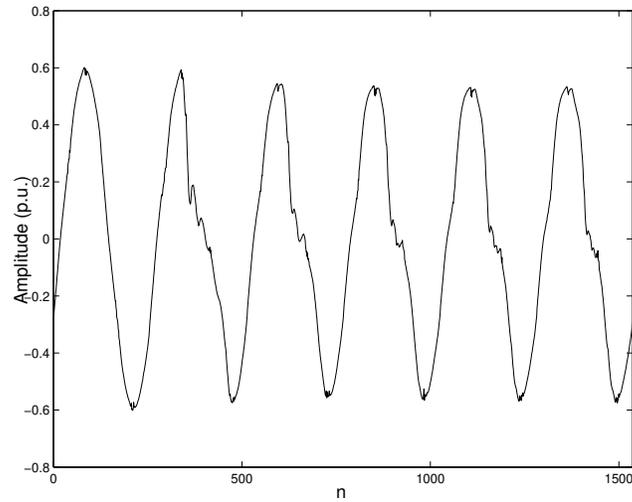
(a) Sinal analisado.



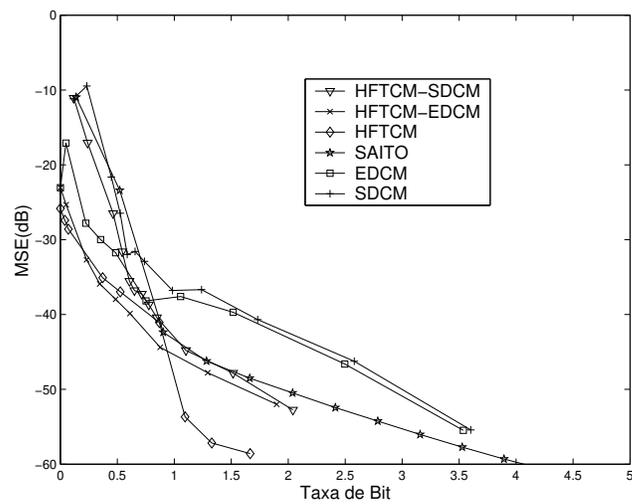
(b) Desempenho dos algoritmos de compressão.

Figura 10.20: Sinal real I: análise de desempenho.

compressão proposto apresenta desempenho superior aos outros 5 métodos aqui analisados. Ganhos consideráveis são observados quando o sinal é constituído de várias componentes harmônicas e as mesmas são previamente estimadas e subtraídas do sinal $\{x(n)\}$.



(a) Sinal analisado.



(b) Desempenho dos algoritmos de compressão.

Figura 10.21: Sinal real II: análise de desempenho.

10.3 Sumário

O presente capítulo discutiu duas novas técnicas para a compressão de distúrbios em sistema de potência, EDCM e HFTCM. Dentre as técnicas propostas, vale

a pena ressaltar que o desempenho geral do FHTCM é parcialmente dependente do dicionário de wavelets considerado. De fato, se um maior número de wavelets é considerado, menores erros de compressão são obtidos, pois dentro de um universo maior de wavelets é possível encontrar uma wavelet mais adequada ao sinal monitorado. No entanto, a complexidade aumenta proporcionalmente ao número de wavelets no dicionário. Além disso, deve-se observar que o número de wavelets no dicionário deve ser cuidadosamente selecionado de forma que possa representar ótimas decomposições dos vários tipos de distúrbios em sistemas de potência. De modo geral, pode-se dizer que o problema de construção do dicionário é bastante relacionado com a classificação dos distúrbios e requer maiores investigações.

11

Conclusões e Perspectivas Futuras

A presente tese discutiu o uso de técnicas de processamento de sinais e de inteligência computacional em duas áreas distintas: a transmissão de dados via rede elétrica e o monitoramento da QE.

Duas contribuições foram apresentadas na área de transmissão de dados via redes elétricas, a saber: *i*) a introdução novos algoritmos fuzzy do tipo 2 para o cancelamento de ruídos impulsivos em sistemas multi-portadora OFDM/DMT e *ii*) a apresentação novos algoritmos fuzzy do tipo 1 e 2 adaptativos para a equalização de canais PLC, que tipicamente apresentam a presença de ruídos impulsivos em

suas saídas. Apesar da complexidade computacional das técnicas propostas para o cancelamento de ruídos impulsivos e para a equalização de canais PLC, os resultados obtidos apontam direções de investigação relevantes não apenas para canais PLC, mas também em outras aplicações onde a complexidade computacional não seja um fator limitante.

Na área de monitoramento da QE, novas técnicas *i)* para a análise espectral (fundamental e harmônicas) dos sinais monitorados, *ii)* para a detecção de distúrbios e os instantes iniciais de início e fim de ocorrência dos mesmos, *iii)* para a classificação de distúrbios isolados e múltiplos distúrbios, e, finalmente, *iv)* para a compressão de distúrbios. Os resultados obtidos nas simulações computacionais comprovam o quão efetivo e adequado é o uso do conceito dividir e conquistar para monitoramento da QE, abordados nesta tese.

As possíveis questões futuras de pesquisa, resultantes do trabalho desenvolvido, são listadas a seguir:

O cancelamento de ruídos impulsivos: Trabalhos futuros nesta linha de pesquisa estão relacionados à análise da técnica proposta quando o estimador do canal não é ideal, a obtenção de parâmetros que otimizem o uso combinado das técnicas de cancelamento de ruídos impulsivos com as técnicas de codificação de canais, a análise num contexto MIMO (*wireline* e *wireless*) e, finalmente, a análise da técnica proposta em canais de comunicação variantes no tempo. Uma outra proposta interessante é o uso do critério MDL, discutido na seção 10.2, para a subtração de ruídos na saída do canal PLC.

A equalização de canais de comunicação: Desdobramentos nesta linha apontam a aplicação dos algoritmos propostos em outros tipos de canais de comunicação. Além disso, o uso dos equalizadores fuzzy propostos dentro da abordagem da equalização turbo de canais (Loiola et al., 2004) é uma questão

de investigação bastante promissora

A análise espectral de sinais de potência: apesar do desempenho obtido com a técnica NF-WDFT alguns comentários merecem destaque: *i*) O uso de outras técnicas que possam realizar o mapeamento executado pela rede MLP pode ser investigado, objetivando-se melhorar o desempenho e/ou diminuir a complexidade computacional; *ii*) a investigação da WDFT em ponto fixo é uma questão em aberto que merece investigação, uma vez que o custo computacional da solução pode ser reduzido; *iii*) a busca por relações explícitas entre os parâmetros (frequência angular, fase e amplitude) nos domínios da DFT e WDFT, visando melhorar ou eliminar o processo de mapeamento baseado na rede MLP; *iv*) finalmente, a extensão da técnica para a estimação das componentes inter-harmônicas. Existem ainda algumas questões relacionadas à técnica NF-WDFT passíveis de investigações, cujos resultados tendem a oferecer uma solução com diferentes relações de desempenho e complexidade computacional.

A detecção de distúrbios em sistemas de potência: apesar dos resultados alcançados serem bastantes efetivos, os seguintes comentários merecem destaque: *i*) os conjuntos de dados de treinamento e de teste foram obtidos manualmente, ou seja, o autor desta tese, através de inspeção visual, definiu os instantes iniciais e finais de ocorrência dos distúrbios. Assim sendo, o uso de outros bancos de dados é recomendado para validar o uso da técnica proposta, visto que os resultados são bastante expressivos quando $N = 256$; *ii*) apesar do uso combinado dos parâmetros $S^{(3)}$ e $\bar{\gamma}_4$ como parâmetros de entrada da rede MLP apresentar resultados bastante satisfatórios, o uso de outras técnicas de reconhecimento de padrões (Theodoridis & Koutroumbas, 1999; Mendel, 2001; Haykin, 1999; Pedrycz & Gomide, 1998; Príncipe et al., 2000) pode ser analisado e investigado, visando melhorar o desempenho ou diminuir ainda mais a complexidade computacional; e *iii*) a análise de outros

parâmetros pode ser investigada, tendo como motivação principal verificar se é possível obter bons resultados quando $N < 256$.

A classificação de distúrbios em sistemas de potência: Os próximos passos nesta área é analisar o uso dos algoritmos para o cancelamento de ruídos impulsivos propostos no capítulo 4, levando-se em consideração a complexidade e o desempenho. Finalmente, o desenvolvimento de técnicas que possam identificar a fonte geradora do(s) distúrbio(s), a partir dos resultados obtidos com o classificador de múltiplos distúrbios proposto na seção 9. Essa é uma questão em aberto e que merece investigação.

A compressão de distúrbios em sistemas de potência: A questão da quantização vetorial e da codificação por entropia carecem de investigação no contexto da FHTCM. Dentre as técnicas de codificação por entropia, o uso de sistemas caóticos aparece com uma linha de pesquisa bastante promissora, posto que recentes resultados apontam que as técnicas de codificação por entropia, baseada em sistemas caóticos, obtém desempenhos similares aos obtidos com os códigos aritméticos (Luca, Serbanescu, Azou, & Burel, 2004; Schweizer & Schimming, 2001a, 2001b). Além disso, a investigação dos valores máximos de distorções inseridos no processo de codificação dos sinais que não prejudiquem o cálculo dos índices de qualidade é uma questão de grande importância e que está sendo investigada.

12

Referências Bibliográficas

Referências

- Abad, J., Badenes, A., Blasco, J., Carreras, J., Dominguez, V., Gomes, C., Iranzo, S., Riveiro, J. C., Ruiz, D., Torres, L. M., & Comabella, J. (2003, Abr.). Extending the power line LAN up to the neighborhood transformer. *IEEE Comm. Magazine*, 41(4), 64-70.
- A. Barron, J. R., & Yu, B. (1998, Out.). The minimum description length principle in coding and modeling. *IEEE Trans. on Information Theory*, 44(6), 2743-

2760.

- Abdel-Galil, T. K., El-Saadany, E. F., & Salama, M. M. A. (2001). Effect of new deregulation policy on power quality monitoring and mitigation techniques. In *Proc. of IEEE/PES Transmission and Distribution Conference and Exposition* (Vol. 1, p. 554-560).
- Abdel-Galil, T. K., El-Saadany, E. F., & Salama, M. M. A. (2003). Power quality event detection using Adaline. *Electric Power Systems Research*, 64, 137-144.
- Abdel-Galil, T. K., Kamel, M., Youssef, A. M., El-Saadany, E. F., & Salama, M. M. (2004, Out.). Power quality disturbance classification using the inductive inference approach. *IEEE Trans. on Power Delivery*, 19(4), 1818-1818.
- Abend, K., & Fritchman, B. D. (1970). Statistical detection for communication channels with intersymbol interference. *Proc. of the IEEE*, 58(5), 779-785.
- Acker, K. V., Leus, G., Moonen, M., Wiel, O. van de, & Pollet, T. (2001, Jan.). Per tone equalization for DMT-based systems. *IEEE Trans. Commun.*, 49(1), 109-119.
- Anderson, B. D. O., & Moore, J. B. (1979). *Optimal Filtering*. Prentice-Hall.
- Angrisani, L., Daponte, P., D'Apuzzo, M., & Pietrosanto, A. (1999). A VXI power quality analyser implementing a wavelet transform-based measurement procedure. *Measurement*, 26, 91-102.
- Aranha, A. B. (2003). Análise de transmissão de dados em banda larga via equipamentos powerline para rede elétrica secundária. In *XVII Seminário Nacional de Produção e Transmissão de Energia Elétrica - XVIIISNPTEE*. Uberlândia.
- Armijo, L. (1966). Minimization of functions having Lipschitz-continuous first partial derivatives. *Pacific Journal of Mathematics*, 16, 1-3.
- Arrillaga, J., Bollen, M. H. J., & Watson, N. R. (2000, Fev.). Power quality following deregulation. *Proc. of the IEEE*, 88(2), 246-261.
- Arslan, G., Evans, B. L., & Kiaei, S. (2001, Dec). Equalization for discrete multitone transceivers to maximize bit rate. *IEEE Trans. on Signal Processing*, 49(12),

3123-3135.

- Ascom. (2001, Junho). *Nothing is as powerful as an idea for which the time is ripe: welcome to the second discovery of electricity*. www.ascom.com.
- Ascom. (2003). *Ascom powerline communications aeg*. www.ascom.com.
- Austin, M. (1967, Ago.). *Decision feedback equalization for digital communication over dispersive models* (Technical Report No. 461). MIT Research Laboratory of electronics.
- Azam, M. S., Tu, F., Pattipati, K. R., & Karanam, R. (2004, Julho). A dependency model-based approach for identifying and evaluating power quality problems. *IEEE Trans. on Power delivery*, 19(3), 1154-1166.
- Bagchi, S., & Mitra, S. K. (1999). *The nonuniform discrete Fourier transform and its applications in signal processing*. Boston: Kluwer Academic Publishers.
- Bahai, A. R. S., Satzberg, B. R., & Ergen, M. (2004). *Multi-carrier digital communications; Theory and applications of OFDM* (2 ed.). Springer-Verlag Telos.
- Baig, S., & Gohar, N. D. (2003, Abr.). A discrete multitone transceiver at the heart of the PHY layer of an in-home power line communication local-area network. *IEEE Comm. Magazine*, 41(4), 48-53.
- Barnes, J. (1998, Mar.). A physical multi-path model for power distribution network propagation. In *Proc. of 2nd International Symposium on Power-line Communication and its Applications* (p. 76-89). Tokyo.
- Barto, A. G., & Sutton, R. S. (1981). *Goal seeking components for adaptive intelligence: An initial assessment* (Technical Report No. AFWAL-TR-81-1070). Ohio: Air Force Wright Aeronautical laboratories/ Avionics Laboratory.
- Batista, D. P. (2001). *Transmissão de dados em redes elétricas*. Tese de Mestrado, Universidade Federal de Juiz de Fora (Brasil).
- Berger, T., & Gibson, J. D. (1998, Out.). Lossy source coding. *IEEE Trans. on Information Theory*, 44(6), 2693-2723.

- Biglieri, E. (2003, Maio). Coding and modulation for a horrible channel. *IEEE Comm. Magazine*, 41(5), 92-98.
- Bingham, J. A., & Putten, F. V. D. (1997, Set.). *T1.413 issue 2: Standards project for interfaces relating to carrier to customer connection of asymmetrical digital subscriber line (adsl) equipment* (Relatório Técnico No. T1E1.4/97-007R6). ANSI.
- Bingham, J. A. C. (1990, Maio). Multicarrier modulation for data transmission: An idea whose time has come. *IEEE Comm. Magazine*, 28(5), 5-14.
- Bingham, J. A. C. (2000). *ADSL, VDSL, and Multicarrier Modulation*. NJ: John Wiley & Sons.
- Boashash, B. (1992a, Abr.). Estimating and interpreting the instantaneous frequency of a signal - Part 2: Algorithms and applications. *Proceedings of the IEEE*, 80(4), 540-568.
- Boashash, B. (1992b, Abr.). estimating and interpreting the instantaneous frequency of a signal - Part 1: Fundamentals. *Proceedings of the IEEE*, 80(4), 520-538.
- Bollen, M. H. J. (2000). *Understanding Power Quality Problems - Voltage Sags and Interruptions*. NJ: IEEE Press.
- Bostoen, T., & Wiel, O. V. (2000, Fev.). Modelling the low-voltage power distribution network in the frequency band from 0.5 MHz to 30 MHz for broadband powerline communications (PLC). In *Proc. of International Zurich Symposium on Broadband Communications* (p. 171-178).
- Brwon, P. A. (1998). Some keys factors influencing data transmission rates in the power line environment when utilizing carrier frequencies above 1 MHz. In *Proc. of IEEE 2nd International Symposium on Spread Spectrum Techniques and Applications* (Vol. 2, p. 67-75).
- Burr, A. G., & Brwon, P. A. (1998). HF broadcast interference on low voltage mains distribution networks. In *Proc. of 2nd International Symposium on Power-Line Communications and its Applications* (Vol. 2, p. 253-262).
- Burrus, C. S., Gopinath, R. A., & Guo, H. (1997). *Introduction to wavelets and*

- wavelets transforms: A primer* (1 ed.). Prentice Hall.
- Canete, F. J., Cortes, J. A., Diez, L., & Entrambasaguas, J. T. (2002, Fev.). Broadband modelling of indoor power-line channels. *IEEE Trans. on Consumer Electronics*, 48(1), 175-183.
- Canete, F. J., Cortes, J. A., Diez, L., & Entrambasaguas, J. T. (2003, Abr.). modeling and evaluation of the indoor power line transmission medium. *IEEE Comm. Magazine*, 45, 41-47.
- Castaldo, D., Gallo, D., Landi, C., & Testa, A. (2004, Out.). A digital instrumentation for nonstationary disturbance analysis in power lines. *IEEE Trans. on Instrumentation and Measurement*, 53(5), 1353-1361.
- Chang, R. W. (1966, Dez.). Synthesis of band-limited orthogonal signals for mutichannel data transmission. *Bell system Tech. J.*, 45, 1775-1796.
- Chang, S. G., Yu, B., & Vitterli, M. (2000, Set.). Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. on Image Processing*, 9(9), 1538-1546.
- Chen, R., & Wu, W.-R. (2004, Maio). Adaptive asymptotic bayesian equalization using a signal spece partitioning technique. *IEEE Trans. on Signal Processing*, 52(5), 1376-1386.
- Chen, S., Gibson, G. J., Cowan, C. F. N., & Grant, P. M. (1990). Adaptive equalization of finite nonlinear channels using multilayer perceptrons. *Signal Processing*, 20, 107-119.
- Chen, S., McLaughlin, S., Mulgrew, B., & Grant, P. M. (1995, Maio). Adaptive bayesian decision feedback equalizer for dispersive mobile radio channels. *IEEE Trans. on Communications*, 43(5), 1937-1946.
- Chen, S., McLaughlin, S., Mulgrew, B., & Grant, P. M. (1996, Ago.). Bayesian decision feedback equaliser for overcoming co-channel interference. *IEE Proc. Commun.*, 143(4), 219-225.
- Chen, S., Mulgrew, B., & Hanzo, L. (2000, Dez.). Asymptotic Bayesian Decision Feedback equalizer using a set of hyperplanes. *IEEE Trans. on Signal*

- Processing*, 48(12), 3493-3500.
- Chen, S., Mulgrew, B., & McLaughlin, S. (1993, Set.). Adaptive bayesian equalizer with decision feedback. *IEEE Trans. on Signal Processing*, 41(9), 2918-2927.
- Cheng, Y. T. (2000, Nov.). TMS320C62x algorithm: sine wave generation. In *Texas Instruments, Rel. Téc. SPRA708*.
- Chng, E., & Chen, S. (2005). Determining the optimal decision delay parameter for the linear equalizer. *Int. J. Automation and Computing*. (a ser publicado)
- Chung, J., Powers, E. J., Grady, W. M., & Bhatt, S. C. (1999). Variable rate power disturbance signal processing using embedded zerotree wavelet transform coding. In *Proc. IEEE Power Eng. Soc. Winter Meeting* (Vol. 2, p. 1305-1309).
- Chung, J., Powers, E. J., Grady, W. M., & Bhatt, S. C. (2004, Jan.). Power disturbance classifier using a rule-based method and wavelet packet-based Hidden Markov Model. *IEEE Trans. on Power Delivery*, 17(1), 233-241.
- Cohen, L. (1989, Abr.). Time-frequency distributions - A review. *Proceedings of the IEEE*, 77(7), 942-981.
- Colonnese, S., & Scarano, G. (1999, Fev.). Transient signal detection using higher order moments. *IEEE Trans. on Signal Processing*, 47(2), 515-520.
- Dai, H., & Poor, H. V. (2003, Out.). Crosstalk mitigation in DMT VDSL with impulse noise. *IEEE Trans. on Circuits and Systems - I: Fundamental Theory and Applications*, 48(10), 1205-1213.
- Dash, P., Pradhan, A. K., & Panda, G. (1999, Jul.). Frequency estimation of distorted power system signals using extended complex Kalman filter. *IEEE Trans. on Power Delivery*, 14(3), 761-766.
- Dash, P. K., & Chilukuri, M. V. (2004). Hybrid S-transform and Kalman filtering approach for detection and measurement of short duration disturbances in power networks. *IEEE Trans. on Instrumentation and Measurement*, 53(2), 588-596.
- Dash, P. K., Jena, P. K., & Salama, M. M. A. (1999). Power quality monitoring

- using an integrated Fourier linear combiner and fuzzy expert system. *Electrical Power and Energy Systems*, 21, 497-506.
- Dash, P. K., Panigrahi, B. K., Sahoo, D. K., & Panda, G. (2003, Abr.). Power quality disturbance data compression, detection, and classification using integrated spline wavelet and S-Transform. *IEEE Trans. on Power Delivery*, 18(2), 595-600.
- Daubechies, I. (1992). *Ten lectures on wavelet*. SIAM, Philadelphia.
- Day, H., & Poor, H. V. (2003, Maio). Advanced signal processing for power line communications. *IEEE Communications Magazine*, 100-107.
- Devroye, P. A., Györfi, L., & Lugosi, G. (1996). *A Probabilistic theory of pattern recognition*. San Diego: Springer-Verlag.
- Donoho, D. L. (1995, Apr.). De-noising by soft-thresholding. *IEEE Trans. on Information Theory*, 17(3), 613-627.
- Donoho, D. L., & Johnstone, I. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81, 425-455.
- Donoho, D. L., & Johnstone, L. M. (1995, Dez.). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432), 1200-1224.
- Dostert, K. (1998, Abr.). Power lines as high speed data transmission channels—modelling the physical limits. In *Proc. of IEEE 2nd International Symposium on Spread Spectrum Techniques and Applications* (Vol. 2, p. 585-589).
- Dostert, K. (2001). *Power Line Communications*. Prentice Hall.
- DS2. (2004). *Ds2*. www.ds2.es.
- Dualibe, C., Jespers, P., & Verleysen, M. (2001). Embedded fuzzy control for automatic channel equalization after digital transmissions. In *IEEE International Symposium on Circuits and Systems* (Vol. 2, p. 173-176).
- Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. John Wiley and Sons.
- Dugan, R. C., McGranaghan, M. F., & Beaty, H. W. (1996). *Electrical Power*

Systems Quality. McGraw-Hill.

- Duque, C. A., Ribeiro, M. V., Braga, H. A. C., & Almeida, S. Q. (2000). Compressão de dados para análise de qualidade de energia elétrica utilizando transformada de wavelet e filtro de Kalman. In *Proc. Automatic Brazilian Congress* (p. 207-211).
- Duque, C. A., Ribeiro, M. V., Ramos, F. R., & Szczupak, J. (2005). Power quality event detection based on the principle of divide to conquer and innovation concept. *IEEE Trans. on Power Delivery*. (aceito para publicação)
- Ece, D. G., & Gerek, Ö. N. (2004, Ago.). Power quality event detection using joint 2-D-Wavelet subspaces. *IEEE Transactions on Instrumentation and Measurement*, 53(4), 1040-1046.
- Echelon. (2004). *Echelon*. www.echelon.com.
- Effros, M. (1998, Nov.). Optimal modeling for complex system design (data compression). *IEEE Signal Processing Magazine*, 15(6), 51-73.
- Engelbrecht, A. P. (2002). *Computational intelligence : An introduction*. John Wiley & Sons.
- EPRI, E. P. R. I. (2002). *Pqview software*. www.pqview.com.
- Esmailian, T., Gulak, P. G., & Kschischang, F. R. (2000). A Discrete multitone power line communications system. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP* (p. 2953-2956).
- Esmailian, T., Kschischang, F. R., & Gulak, P. G. (2000). Characteristics of in-building power lines at high frequencies and their channel capacity. In *Proc. of 4th International Symposium on Power-line Communication and its Applications* (p. 52-59).
- Faber, T., Scholand, T., & Jung, P. (2003, Jul.). Turbo decoding in impulsive noise environments. *IEEE Electronics Letters*, 39(14), 1069-1071.
- Fahlman, S. E. (1988). *An empirical study of learning speed in back-propagation networks* (Technical Report No. Computer Science Technical Report). Carnegie-Mellon University.

- FCC. (2004). *Federal communications commission fcc*. www.fcc.org.
- Ferrari, R., Panazio, C. M., Attux, R. R. F., Cavalcante, C. C., Castro, L. N. de, Zuben, F. J. V., & Romano, J. M. T. (2003). Unsupervised channel equalization using fuzzy prediction-error filters. In *IEEE Neural Networks for Signal Processing* (p. 1680-1686).
- Ferrero, A., & Salicone, S. (2004, Maio). An easy VI program to detect transient disturbances in the supply voltage. In *Proc. 21st IEEE Instrumentation and Measurement Technology Conference* (Vol. 2, p. 1384-1387).
- Feuer, A., & Goodwin, G. C. (1996). *Sampling in digital signal processing and control*. Boston: Birkhauser.
- Fishler, M., & Messer, H. (1993, Jun.). Passive transient signal detection using information in the likelihood-ratio-test. *IEEE Trans. on Signal Processing*, 41(6), 2177-2193.
- Forney, G. D. (1972, Maio). Maximum likelihood sequence estimator of digital sequences in the presence of intersymbol interference. *IEEE Trans. on Information Theory, IT-18*, 363-378.
- Franz, S., Mitra, S. K., & Doblinger, G. (2003). Frequency estimation using warped discrete Fourier transform. *Signal Processing*, 83(2), 1661-1671.
- Fritchman, B. D. (1967, Abr.). A binary channel characterization using partitioned Markov chain. *IEEE Trans. on Information Theory, IT-13*(2), 221-227.
- Gaing, Z.-L. (2004, Out.). Wavelet-based neural network for power disturbance recognition and classification. *IEEE Trans. on Power Delivery*, 19(4), 1560-1568.
- Gaouda, A. M., Kanoun, S. H., Salama, M. M. A., & Chikhani, A. Y. (2002, Maio). Wavelet-based signal processing for disturbance classification and measurement. *IEE Proc.-Gener. Transm. Distrib.*, 149(3), 310-318.
- Gaouda, A. M., Salama, M. M., Sultan, M. R., & Chikhani, A. Y. (1999, Out.). Power quality detection and classification using wavelet-multiresolution signal decomposition. *IEEE Trans. on Power Delivery*, 14(4), 1469-1476.

- Geordiadis, A. T., & Mulgrew, B. (2001). Adaptive Bayesian decision feedback equaliser for alpha-stable noise environments. *Signal Processing*, 81, 1601-1623.
- Gerek, O. N., & Ece, D. G. (2004, Abril). 2-D analysis and compression of power quality event data. *IEEE Trans. on Power Delivery*, 19(2), 791-798.
- Gersho, A., & Gray, R. M. (1999). *Vector quantization and signal compression*. Boston: Kluwer Academic Publishers.
- Gharieb, R. R. (2000, Abril). Higher order statistics based IIR notch filtering scheme for enhancing sinusoids in coloured noise. *IEE Proc.-Vis. Image Signal Processing*, 147(2), 115-121.
- Ghosh, A. K., & Lubkeman, D. L. (1995, Jan.). The classification of power system disturbance waveforms using a neural network approach. *IEEE Trans. on Power Delivery*, 10(1), 109-115.
- Ghosh, M. (1996, Feb.). Analysis of the effect of impulse noise on Multicarrier and single carrier QAM systems. *IEEE Trans. on Communications*, 44(2), 145-147.
- Giannakis, G. B., & Tsatsanis, M. K. (1990, Jul.). Signal detection and classification using matched filtering and higher order statistics. *IEEE Trans. on Acoust., Speech, Signal Processing*, 38(7), 1284-1296.
- Giles, T. C. (1995). *On the design of HF radio modems*. Tese de Doutorado, University of Adelaide.
- Goodwin, G. C., & Middleton, R. H. (1986, Nov.). Improved finite word length characteristics in digital control using delta operators. *IEEE Trans. on Automatic Control*, AC-31, 1015-1021.
- Goodwin, G. C., Middleton, R. H., & Poor, V. (1992, Jan.). High-speed digital signal processing and control. *Proceedings of the IEEE*, 80(2), 240-259.
- Gray, R. M., & Neuhoff, D. L. (1998, Out.). Quantization. *IEEE Trans. on Information Theory*, 44(6), 2325-2383.
- Gu, I. Y.-H., & Styvaktakis, E. (2004). Bridge the gap: signal processing for power

- quality applications. *Electric Power Systems Research*, 66, 83-96.
- Gürçan, M. N., Çetin, Y. Y. A. E., & Ansari, R. (1997, Ago.). Detection of microclacifications in mammograms using higher order statistics. *IEEE Signal Processing Letters*, 4(8), 213-216.
- Hamid, E. Y., & kawasaki, Z. I. (2002, Abr.). Wavelet-based data compression of power system disturbances using the minimum description length criterion. *IEEE Trans. on Power Delivery*, 17(2), 460-466.
- Hansen, M., & Yu, B. (2000, Ago.). Wavelet thresholding via MDL for natural images. *IEEE Trans. on Information Theory*, 46(5), 1778-1788.
- Hanzo, L., Münster, M., Choi, B. J., & Keller, T. (2003). *OFDM and MC-CDMA for broadband multi-user communications, WLANs and Broadcasting* (1 ed.). John Wiley & Sons.
- Hanzo, L., Webb, W., & Keller, T. (2000). *Single- and multi-carrier quadrature amplitude modulation: principles and applications for personal communications, WLANs and broadcasting* (1 ed.). John Wiley & Sons.
- Hanzo, L., Wong, C. H., & Yee, M. S. (2002). *Adaptive wireless transceivers: turbo-coded, turbo-equalized and space-time coded TDMA, CDMA, and OFDM systems* (1 ed.). John Wiley & Sons.
- Haring, J., & Vinck, J. H. (2003, Maio). Iterative decoding of codes over complex numbers for impulsive noise channels. *IEEE Trans. on Information Theory*, 49(5), 1251-1260.
- Haring, J., & Vinck, J. H. (2004, Maio). Coding and signal space diversity for a class of fading and impulsive noise channels. *IEEE Trans. on Information Theory*, 50(5), 887-895.
- Haykin, S. (1996). *Adaptive filter theory*. NJ: Englewood Cliffs, Prentice Hall.
- Haykin, S. (1999). *Neural networks. A comprehensive foundation* (2 ed.). NJ: Englewood Cliffs, Prentice Hall.
- Haykin, S., & Li, L. (1995, Feb.). Nonlinear adaptive prediction of nonstationary signals. *IEEE Trans. on Signal Processing*, 43(2), 526-535.

- He, T., Tjhung, T. T., Chew, Y. H., & Ko, C. C. (2000, Mar.). BER performance of Reed-Solomon coded baseband data transmission in impulse noise channels. *Electronics Letters*, *36*(2), 146-147.
- He, Z., & Mitra, S. K. (2001, Dez.). A unified rate-distortion analysis framework for transform coding. *IEEE Trans. on Circuits and Systems for Video Technology*, *11*(12), 1221-1236.
- Higgins, W. E., & Jr., D. C. M. (1982, Dez.). Noise reduction strategies for digital filters: Error spectrum shaping versus the optimal linear state-space formulation. *IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-30*, 963-973.
- Hirano, K., Nishimura, S., & Mitra, S. K. (1974, Jul.). Design of digital notch filters. *IEEE Trans. on Communications*, *22*(7), 964-970.
- Hoang, T. A., & Nguyen, D. T. (2002, Ago.). Improving training of radial basis function network for classification of power quality disturbances. *Electronics Letters*, *38*(17), 976-977.
- Hooijen, O. (1998, Ago.). A channel model for the residential power circuit used as a digital communications médium. *IEEE Trans. on Electromagnetic Compatibility*, *40*(4), 331-336.
- Hsieh, C., Huang, S., & Huang, C. (1998). Data reduction of power quality disturbances - a wavelet transform approach. *Electric Power Systems Research*, *47*, 79-86.
- Hsieh, C.-T., & Huang, S.-J. (2003, Jan.). Disturbance data compression of a power system using the Huffman coding approach with wavelet transform enhancement. *IEE Proc. Generation, Transmission and Distribution*, *150*(1), 7-14.
- Hsieh, G.-C., & Hung, J. C. (1996, Jun.). Phase-locked loop techniques - A Survey. *IEEE Trans. on Industrial Electronics*, *43*(6), 609-615.
- Huang, J., Negnevitsky, M., & Nguyen, D. T. (2002, Abr.). A neuro-fuzzy classifier for recognition of power quality disturbances. *IEEE Trans. on Power Delivery*,

- 17(2), 609-616.
- Huang, S.-J., Hsieh, C.-T., & Huang, C.-L. (1998). application of wavelet to classify power system disturbances. *Electric Power Systems Research*, 47, 87-93.
- Hwang, S. Y. (1977, Ago.). Minimum uncorrelated unit noise in state-space digital filtering. *IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-25*, 273-281.
- Ibnkahla, M. (2000). Applications of neural networks to digital communications - a survey. *Signal Processing*, 80(7), 1185-1215.
- Ibrahim, T. R. A., & Morcos, M. M. (2003, Jul.). A power quality perspective to system operational diagnosis using fuzzy logic and adaptive techniques. *IEEE Trans. on Power Delivery*, 18(3), 903-909.
- Intellon. (2004). *Intellon*. www.intellon.com/.
- Jackson, L. B., Lindgren, A. G., & Kim, Y. (1979, Mar.). Optimal synthesis of second-order state-space structures for digital filters. *IEEE Trans. on Circuits Systems, CAS-26*, 143-156.
- Jacobs, R. A. (1988, Set.). Increased rates of convergence through learning rate adaptation. *Neural Networks*, 1, 295-307.
- James, B., Anderson, D. O., & Williamson, R. C. (1994, Jun.). Conditional mean and maximum likelihood approaches to multiharmonic frequency estimation. *IEEE Trans. on Signal Processing*, 42(6), 1366-1375.
- Jee, G., Edison, C., Rao, R. D., & Cern, Y. (2003, Maio). Demonstration of the technical viability of PLC systems on medium- and low-voltages lines in the United States. *IEEE Comm. Magazine*, 41(5), 108-112.
- Jeruchim, M. C., Balaban, P., & Shanmugan, K. S. (2000). *Simulation of communication systems. Modeling, methodology, and techniques* (2 ed.). NY: Kluwer Academic/Plenum Publishers.
- José-Revuelta, L. M. S., & Sueiro, J. cid. (2004). Robust joint channel and noise estimation in Bayesian blind equalizers. *Signal Processing*, 84, 535-548.
- Jr., G. D. F. (1973, Março). The Viterbi algorithms. *Proc. of the IEEE*, 61, 268-278.

- Kalman, R. E. (1960, Set.). A new approach to linear filtering and prediction problems. *IEEE Trans. of the ASME - Journal of Basic Engineering*, 82(D), 35-45.
- Karayiannis, N. B., & Venetsanopoulos, A. N. (1992). Fast learning algorithms for neural networks. *IEEE Trans. on Circuits and Systems - II: analog and digital signal processing*, 39(7), 453-474.
- Karimi, M., Mokhtari, H., & Iravani, M. R. (2000, Oct.). Wavelet based on-line disturbance detection for power quality applications. *IEEE Trans. on Power Delivery*, 15(4), 1212-1220.
- Karnik, N. N., Mendel, J. M., & Liang, Q. (1999, Dez.). Type-2 fuzzy logic systems. *IEEE Trans. on Fuzzy Systems*, 7(6), 643-658.
- Kauraniemi, J., Laakso, T. I., Hartimo, I., & Ovaska, S. J. (1998, Jan.). Delta operator realizations of direct-form IIR filters. *IEEE Trans. on Circuits and Systems - II: Analog and Digital Signal Processing*, 45(1), 41-52.
- Kesten, H. (1958). Accelerated stochastic approximation. *Annals of Mathematical Statistics*, 29, 41-59.
- Kezunovic, M., & Liao, Y. (2001). A new method for classification and characterization of voltage sags. *Electric Power Systems Research*, 58, 27-35.
- Kezunovic, M., & Liao, Y. (2002, Abr.). A novel software implementation concept for power quality study. *IEEE Trans. on Power Delivery*, 17(2), 544-549.
- K.-Ghartemani, M., & Iravani, M. R. (2003, Out.). A signal processing module for power system applications. *IEEE Trans. on Power Delivery*, 18(4), 1118-1126.
- K.-Ghartemani, M., Karimi, H., & Iravani, M. R. (2004, Abr.). A magnitude/phase-locked loop system based on estimation of frequency and in-phase/quadrature-phase amplitudes. *IEEE Trans. on Industrial Electronics*, 51(2), 511-517.
- Khan, A. K. (2001, Abr.). Monitoring power for the future. *Power Engineering Journal*, 15(2), 81-85.
- Krim, H., & Schick, I. C. (1999, Abr.). Minimax description length for signal denoising and optimized representation. *IEEE Trans. on Information Theory*,

- 45(3), 898-908.
- Krim, H., Tucker, D., Mallat, S., & Donoho, D. (1999). On denoising and best signal representation. *IEEE Trans. on Information Theory*, 45(7), 2225-2238.
- Laakso, T. I., & Hartimo, I. (1992, Maio). Noise reduction in recursive digital filters using high order error feedback. *IEEE Trans. on Signal Processing*, 40, 1096-1107.
- Langfeld, P. J. (2001). The Capacity of typical power line reference channels and strategies for system design. In *Proc. 5th International Symposium on Power-line Communication and its Applications* (p. 271-278).
- Lee, C. H., & Nam, S. W. (1998, Maio). Efficient feature vector extraction for automatic classification of power quality disturbances. *Electronics Letters*, 34(11), 1059-1061.
- Lee, I. W. C., & Dash, P. K. (2003, Ago.). S-Transform-based intelligent system for classification of power quality disturbance signals. *IEEE Trans. on Industrial Electronics*, 50(4), 800-805.
- Lee, J. S., Lee, C. H., Kim, J. O., & Nam, S. W. (1997, Ago.). Classification of power quality disturbances using orthogonal polynomial approximation and bispectra. *Electronics Letters*, 33(18), 1523-1524.
- Lee, J. Y., Won, Y. J., Jeong, J.-M., & Nam, S. W. (2002, Jul.). Classification of power disturbances using feature extraction in time-frequency plane. *Electronic Letters*, 38(15), 833-835.
- Lee, K. (1994, Maio). Fuzzy adaptive decision feedback equaliser. *Electronics Letters*, 30(10), 749-751.
- Lee, K. Y. (1994). Complex RLS fuzzy adaptive filter and its application to channel equalisation. *Electronics Letters*, 30(19), 1572-1574.
- Lee, K. Y. (1996, Fev.). Complex fuzzy adaptive filter with LMS algorithm. *IEEE Trans. on Signal Processing*, 44(2), 424-427.
- Lee, M. S., & Hanzo, L. (2002, Maio). A wide-band radial basis function decision feedback equalizer-assisted burst-burst adaptive modem. *IEEE Trans. on*

- Communications*, 50(5), 693-697.
- Li, G., & Gevers, M. (1993, Fev.). Roundoff noise minimization using delta-operator realizations. *IEEE Trans. on Signal Processing*, 41(2), 629-637.
- Liang, Q., Karnik, N. N., & Mendel, J. M. (2000, Ago.). Connection admission control in ATM networks using survey-based type-2 fuzzy logic systems. *IEEE Trans. on Syst., Man, and Cyber - Part C: Applications and Reviews*, 30(3), 329-339.
- Liang, Q., & Mendel, J. M. (2000a, Out.). Equalization of nonlinear time-varying channels using type-2 fuzzy adaptive filters. *IEEE Trans. on Fuzzy Systems*, 8(5), 551-563.
- Liang, Q., & Mendel, J. M. (2000b, Out.). Interval type-2 fuzzy logic systems: Theory and design. *IEEE Trans. on Fuzzy Systems*, 8(5), 535-550.
- Liang, Q., & Mendel, J. M. (2000c, Dez.). Overcoming time-varying co-channel interference using type-2 Fuzzy adaptive filter. *IEEE Trans. on Circuits and Systems*, 47(12), 1419-1428.
- Lin, H., & Yamashita, K. (2001). Blind equalization using parallel bayesian decision feedback equalizer. *Mathematics and Computers in Simulation*, 56, 247-257.
- Lin, Y.-J., Latchman, H. A., Lee, M., & Katar, S. (2002, Dez.). A power line communication network infrastructure for the smart home. *IEEE Wireless Communications*, 9(6), 104-111.
- Lin, Y.-J., Latchman, H. A., Newman, R. E., & Katar, S. (2003, Abr.). A comparative performance study of wireless and power line networks. *IEEE Comm. Magazine*, 41(4), 54-63.
- Littler, T. B., & Morrow, D. J. (1999). Wavelets for the analysis and compression of power system disturbances. *IEEE Trans. on Power Delivery*, 14(2), 358-364.
- Liu, S. (2000). An adaptive Kalman filter for dynamic estimation of harmonic signals. In *Proc. 9th IEEE International Conference on Harmonics and Quality of Power*.
- Loehr, G., Casazza, C., & John, A. (1999). *The evolution of electric power*

- transmission under deregulation: selected readings*. IEEE Press.
- Loiola, M. B., Ribeiro, M. V., & Romano, J. M. T. (2004). A turbo equalizer using fuzzy filters. In *Proc. IEEE workshop on Machine Learning for Signal Processing* (p. num. pg. a confirmar).
- Lu, C., & Huang, S.-J. (2004). An application of B-spline wavelet transform for notch detection enhancement. *IEEE Trans. on Power Delivery*, 19(3), 1419-1425.
- Luca, M. B., Serbanescu, A., Azou, S., & Burel, G. (2004, June). A new compression method using a chaotic symbolic approach. In *IEEE-Communications*.
- Magoulas, G. D., & Vrahatis, M. N. (1999). Improving the convergence of the backpropagation algorithm using learning rate adaptation methods. *Neural Computation*, 11, 1769-1796.
- Magoulas, G. D., & Vrahatis, M. N. (2000). A class of adaptive rate algorithms derived by one-dimensional subminimization methods. *Neural, Parallel & Scientific Computations*, 8, 147-168.
- Makur, A., & Mitra, S. K. (2001, Set.). Warped discrete-Fourier transform: theory and applications. *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, 48(9), 1086-1093.
- Mallat, S. (2001). *A wavelet tour of signal processing* (2 ed.). San Diego: Academic Press.
- Mallat, S. G. (1989, Set.). Multiresolution Approximation and Wavelet Orthonormal Bases of L2. *Trans. of the American Mathematical Society*, 315(9), 69-87.
- Malvar, H. S. (1992). *Signal processing with Lapped transforms*. Philadelphia: Artech House Norwood.
- Maranda, B. H., & Fawcett, J. A. (1990, Abr.). The performance analysis of a fourth order detector. In *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.* (p. 3-6).
- McDonough, R. N., & Whalen, A. D. (1995). *Detection of signals in noise*. London: Academic Press.

- McGrannaghan, M. (2001, Out.). Trends in power quality monitoring. *IEEE Power Engineering Review*, 21(10), 3-9.
- Meher, S. K., Pradhan, A. K., & Panda, G. (2004). An integrated data compression scheme for power quality events using spline wavelet and neural networks. *Electric Power Systems Research*, 69, 213-220.
- Mehta, K., & Russel, B. D. (1989, Jul.). Data compression for digital data from power systems disturbances: Requirements and technique evaluation. *IEEE Trans. on Power Delivery*, 4(3), 1683-1688.
- Mendel, J. M. (1991, Mar.). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications,. *Proc. of the IEEE*, 79(3), 278-305.
- Mendel, J. M. (1995, Mar.). Fuzzy logic systems for engineering: A tutorial. *Proc. of the IEEE*, 83(3), 345-377.
- Mendel, J. M. (2000). Uncertainty, fuzzy logic, and signal processing. *Signal Processing*, 80, 913-933.
- Mendel, J. M. (2001). *Uncertain rule-based fuzzy logic systems. Introduction and new directions*. NJ: Prentice Hall.
- Mendel, J. M. (2004, Fev.). Computing derivatives in interval type-2 fuzzy logic systems. *IEEE Trans. on Fuzzy Systems*, 12(1), 84-98.
- Mendel, J. M., & John, R. I. B. (2002, Abr.). Type-2 fuzzy sets made simple. *IEEE Trans. on Fuzzy Systems*, 10(2), 117-127.
- Mendes, G. D. (2004). *Desenvolvimento de um simulador de canais de comunicações digitais via rede elétrica (plc) para ambientes prediais*. UNICAMP. (Primeiro e Segundo Relatório Científico, Bolsa FAPESP de Iniciação Científica. Co-orientação: Moisés V. Ribeiro)
- Mitra, S. K. (2001). *Digital signal processing - A computer based approach* (2^a ed.). NJ: McGraw-Hill.
- Mitra, S. K., Hirano, K., & Sakaguchi, H. (1974, Out.). A simple method of computing the input quantization and multiplication roundoff errors in a

- digital filter. *IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-22*(5), 326-970.
- Mizumoto, M., & Tanaka, K. (1976). Some properties of fuzzy sets of type-2. *Information and Control*, 31(4), 312-340.
- Modestino, J. W., Sargrad, D. H., & Bollen, R. E. (1988, Maio). Use of coding to combat impulse noise on digital subscriber line. *IEEE Trans. on Communications*, 36(5), 529-537.
- Mokhtari, H., K.-Ghartemani, M., & Iravani, M. R. (2002, Jan.). Experimental performance evaluation of a wavelet-based on-line voltage detection method for power quality applications. *IEEE Trans. on Power Delivery*, 17(1), 161-172.
- Montalvão, J., Dorizzi, B., & Mota, J. C. M. (2001). Why use Bayesian equalization based on finite data blocks? *Signal Processing*, 81, 137-147.
- Moore, P. J., & Português, I. E. (2003, Out.). The influence of personal computer processing mode on line current harmonics. *IEEE Trans. on Power Delivery*, 18(4), 1363-1587.
- Moulin, F., Ouzzif, M., Zeddami, A., & Gauthier, F. (2003, Nov.). Discrete-multitone-based ADSL and VDSL systems performance analysis in an impulse noise environment. *IEE Proc.-Sci. Meas. Technol*, 150(6), 273-278.
- Mouzouris, G. C., & Mendel, J. M. (1997, Fev.). Nonsingleton fuzzy logic systems: Theory and design. *IEEE Trans. on Fuzzy Systems*, 5(1), 56-71.
- Mulgrew, B. (1996, Mar.). Applying radial basis functions. *IEEE Signal Processing Magazine*, 13, 50-65.
- Mullis, C. T., & Roberts, R. A. (1976, Set.). Synthesis of minimum roundoff noise fixed point digital filters. *IEEE Trans. on Circuits Systems, CAS-23*, 551-562.
- Murphy, R. J. (1996, Jan.). Disturbance recorders trigger detection and protection. *IEEE Computer Applications in Power*, 24-28.
- Nemer, E., Goubran, R., & Mahmoud, S. (2001, Mar.). Robust voice activity detection using higher-order statistics in the LPC residual domain. *IEEE*

- Trans. on Speech and Audio Processing*, 9(3), 217-231.
- Nikias, C., & Mendel, J. (1987, Jul.). Bispectrum estimation: A digital signal processing framework. *Proc. of the IEEE*, 75(7), 869-891.
- Nikias, C., & Mendel, J. (1999, Jan.). Signal processing with higher-order statistics. *IEEE Trans. on Signal Processing*, 41(1), 10-38.
- Nikias, C. L., & Petropulu, A. P. (1993). *Higher-order spectra analysis - A nonlinear signal processing framework*. NJ: Englewood Cliffs, Prentice Hall.
- OPERA. (2004). *Open plc european research alliance opera*. www.ist-opera.org.
- Ortega, J. M., & Rheinboldt, W. C. (1970). *Iterative solution of nonlinear equations in several variables*. NY: Academic Press.
- Panda, G., Dash, P. K., Pradhan, A. K., & Meher, S. K. (2002, Abril). Data compression of power quality events using the Slantlet transform. *IEEE Trans. on Power Delivery*, 17(2), 662-667.
- Patra, J. C., Pal, R. N., Chatterji, B. N., & Panda, G. (1999). Nonlinear channel equalization for QAM signal constellation using artificial neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 29(2), 262-271.
- Patra, S. K., & Mulgrew, B. (1998a). Efficient architecture for bayesian equalization using fuzzy filters. *IEEE Trans. on Circuits and Systems - II, Analog and Digital Signal Processing*, 45(7), 812-820.
- Patra, S. K., & Mulgrew, B. (1998b). Fuzzy implementation of a Bayesian equaliser in the presence of intersymbol and co-channel interference. *IEE Proc.-Commun.*, 145(5), 323-330.
- Patra, S. K., & Mulgrew, B. (2000). Fuzzy techniques for adaptive nonlinear equalization. *Signal Processing*, 80(6), 985-1000.
- Pavlidou, N., Vinck, A. J. H., Yazdani, J., & Honary, B. (2003, Abr.). Power line communications: State of art and future trends. *IEEE Comm. Magazine*, 41(4), 34-40.
- Pearlmutter, B. A. (1994). Fast exact multiplication by the Hessian. *Neural*

- Computation*, 6(1), 147-160.
- Pedrycz, W., & Gomide, F. (1998). *An introduction to fuzzy sets: Analysis and design*. MIT Press.
- Peled, A., & Ruiz, A. (1980). Frequency domain data transmission using reduced computational complexity algorithms. In *IEEE Int. Confe. Acoust., Speech, Signal Processing* (p. 964-967). Denver.
- PES, I. (2003). *Ieee pes working group p1433*.
<http://grouper.ieee.org/groups/1433/>.
- Poisson, O., Rioual, P., & Meunier, M. (2000, July). Detection and measurement of power quality disturbances using wavelet transform. *IEEE Trans. on Power Delivery*, 15(3), 1039-1044.
- Príncipe, J. C., Euliano, N. R., & Lefebvre, W. C. (2000). *Neural and adaptive systems. Fundamentals through simulations*. John Wiley & Sons.
- Proakis, J. G. (2000). *Digital communications* (4 ed.). McGraw-Hill.
- Qureshi, S. (1985, Set.). Adaptive equalization. *Proc. of the IEEE*, 73, 1349-1387.
- Ramchandran, K., Vitterli, M., & Herley, C. (1996, Abr.). Wavelets, subband coding, and best bases. *Proceedings on the IEEE*, 84(4), 541-560.
- Ramos, F. R., Ribeiro, M. V., Romano, J. M. T., & Duque, C. A. (2002). On signal processing approach for event detection and compression applied to power quality evaluation. In *Proc. IEEE 10th International Conference on Harmonic and Quality of Power* (p. 133-138).
- Rapp, T. (2002). *Remote load control via powerline and embedded webserver*. Tese de Mestrado, Universidade Federal de Minas Gerais (Brasil).
- Re, E. D., Fantacci, R., Morosi, S., & Serravalle, R. (2003, Out.). Comparison of CDMA and OFDM techniques for downstream power-line communications on low voltage grids. *IEEE Trans on Power Delivery*, 18(4), 1104-1109.
- Regalia, P. A. (1995). *Adaptive IIR Filtering in Signal Processing and Control* (4 ed.). Marcel Dekker Inc., NY.
- Reyneri, L. M. (1999, Julho). Unification of neural and wavelet networks and fuzzy

- systems. *IEEE Trans. on Neural Networks*, 10(4), 801-814.
- Ribeiro, M. V. (2003). On Fuzzy-DFE-LMS and Fuzzy-DFE-RLS algorithms to equalize power line channels. In *IEEE International Symposium on Industrial Electronics* (p. 1001-1006).
- Ribeiro, M. V. (2005). Learning rate updating methods applied to adaptive fuzzy equalizers for broadband power line communications. *EURASIP Journal on Applied Signal Processing, Special Issue on Nonlinear and Image Processing*. (aceito para publicação)
- Ribeiro, M. V., Barbedo, J. G. A., Romano, J. M. T., & Lopes, A. (2005). Fourier-Lapped-Multilayer Perceptron (FLMLP) Method for Speech Quality Assessment. *EURASIP Journal on Applied Signal Processing, Special Issue on Anthropomorphic Processing of Audio and Speech*. (aceito para publicação)
- Ribeiro, M. V., Duque, C. A., & Romano, J. M. T. (2001). An enhanced data compression method for applications in power quality analysis. In *Proc. 27th Annual IEEE Industrial Electronics Society Conference* (Vol. 2, p. 676-681).
- Ribeiro, M. V., Duque, C. A., & Romano, J. M. T. (2004, Abril). An improved method for signal processing and compression in power quality evaluation. *IEEE Trans. on Power Delivery*, 19(2), 464-471.
- Ribeiro, M. V., Lopes, R. R., Duque, C. A., & Romano, J. M. T. (2005). Impulse noise mitigation for improved bit-rate in PLC-DMT. *IEEE Trans. on Power Delivery*. (aceito para publicação)
- Ribeiro, M. V., Mitra, S. K., & Romano, J. M. T. (2004a). A New Method for Decentralized Power Quality Monitoring. Part II: Compression of Disturbance Events. *submitted to IEEE Trans. on Power Delivery*.
- Ribeiro, M. V., Mitra, S. K., & Romano, J. M. T. (2004b). A new method for the decentralized power quality monitoring. Part I: Spectral analysis of the monitored signals. *IEEE Trans. on Power Delivery*. (submetido)
- Ribeiro, M. V., Mitra, S. K., & Romano, J. M. T. (2004c). A novel approach for spectral analysis of monitored power systems. In *Proc. IEEE 11th*

International Conference on Harmonic and Quality of Power.

- Ribeiro, M. V., & Romano, J. M. T. (2005a). A new method for the decentralized power quality monitoring. Part III: Detection of disturbance events. *IEEE Trans. on Power Delivery*. (a ser submetido)
- Ribeiro, M. V., & Romano, J. M. T. (2005b). A new method for the decentralized power quality monitoring. Part IV: Classification of disturbance events. *IEEE Trans. on Power Delivery*. (a ser submetido)
- Ribeiro, M. V., Romano, J. M. T., & Deckmann, S. M. (2003). Adaptive filtering, wavelet and Lapped transforms for power quality problem detection and identification. In *Proc. IEEE International Symposium on Industrial Electronics* (p. 301-306).
- Riedmiller, M. (1994). Advanced supervised learning in multi-layer perceptrons - From backpropagation to adaptive learning algorithms. *International Journal of Computer Standards and Interfaces - Special Issue on Neural Networks*, 16, 265-278.
- Rife, D. c., & Boorstyn, R. R. (1974, Abr.). Single tone parameter estimation from discrete-time observation. *IEEE Trans. on Information Theory*, 20, 591-598.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 44, 467-471.
- Romano, J. M. T., & Bellanger, M. (1988, Set.). Fast least squares adaptive notch filtering. *IEEE Trans. on Acoustics, Speech and Signal Processing, ASSP-19(9)*, 1536-1540.
- Romano, J. M. T., Filho, J. P., Tavares, M. C. D., Ribeiro, M. V., Loiola, M. B., & Mendes, G. D. (2003). Caracterização de canais PLC (Power line Communications) e desenvolvimento de técnicas para equalização e codificação. In *Workshop de Pesquisa e Desenvolvimento UNICAMP-CPFL* (p. 33-35). Campinas.
- Romano, R. B. (2003). Avaliação preliminar de desempenho de um sistema PLC de banda larga. In *XVII Seminário Nacional de Produção e Transmissão de Energia Elétrica - XVIISNPTEE*. Uberlândia.

- Routray, A., Pradhan, A. K., & Rao, K. P. (2002, Junho). A novel Kalman filter for frequency estimation of distorted signals in power systems. *IEEE Trans. on Instrumentation and Measurement*, 51(3), 469-479.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Parallel distributed processing explorations in the microstructure of cognition. In D. E. Rumelhart & J. L. McClelland (Eds.), (Vol. 1, chap. Learning internal representation by error propagation). MIT Press.
- Saito, N. (1994). Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion. *Wavelets in Geophysics E. Foufoula-Georgiyo and P. Kumar, Eds. New York Academic*, 299-324.
- Sanderson, W. (2000). Broadband communications over a rural power distribution circuit. In *Proc. IEEE Southeastcon* (p. 497-504).
- Santos, E. P., & Zuben, F. J. V. (2000). Efficient second-order learning algorithm for discrete-time recurrent neural networks. In L. R. Medsker & L. Jain (Eds.), (p. 47-75). Boca Raton, FL: CRC Press.
- Santoso, S., Grady, W. M., Powers, E. J., Lamoore, J., & Bhatt, S. C. (2000, Jan.). Characterization of distribution power quality events with Fourier and wavelet transforms. *IEEE Trans. on Power Delivery*, 15(1), 247-254.
- Santoso, S., Powers, E. J., & Grady, W. M. (1997, Julho). Power quality disturbance data compression using wavelet transform methods. *IEEE Trans. on Power Delivery*, 12(3), 1250-1257.
- Santoso, S., Powers, E. J., & Hofmann, P. (1996, Abr.). Power quality assessment via wavelet transform analysis. *IEEE Trans. on Power Delivery*, 11(2), 924-930.
- Saridis, G. N. (1970). Learning applied to successive approximation algorithms. *IEEE Trans. on Systems Science and Cybernetics*, SSC-6, 97-103.
- Sarwal, P., & Srinath, M. (1995, Maio). A fuzzy logic system for channel equalization. *IEEE Trans. on Fuzzy Systems*, 3(2), 246-249.
- Satzberg, B. R. (1967, Dez.). Performance of an efficient parallel data transmission

- system. *IEEE Trans. on Communications, COM-15*, 805-811.
- Schweizer, J., & Schimming, T. (2001a, Mar.). Symbolic dynamics for processing chaotic signal. I. Noise reduction of chaotic sequences. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(11), 1269-1282.
- Schweizer, J., & Schimming, T. (2001b, Mar.). Symbolic dynamics for processing chaotic signal. II. Communication and coding Circuits and Systems I: Fundamental Theory and Applications. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(11), 1283-1295.
- Shayani, D. A. (2004). *Técnicas de processamento adaptativo de sinais aplicadas ao monitoramento da qualidade da energia elétrica via canais plc (powerline communication)*. co-orientação: Moisés v. ribeiro. UNICAMP. (Primeiro e Segundo Relatório Científico, Bolsa FAPESP de Iniciação Científica)
- Shlomot, E., Cuperman, V., & Gersho, A. (2001, Set.). Hybrid coding: combined harmonic and waveform coding of speech at 4 kb/s. *IEEE Trans. on Speech and Audio Processing*, 9(6), 632-646.
- Silva, F., & Almeida, L. (1990). Acceleration techniques for the back-propagation algorithm. *Lecture Notes in computer Science*, 412, 110-119.
- Starr, T., Cioffi, J. M., & Silverman, P. J. (1998). *Understanding digital subscriber line technology*. NJ: Prentice Hall.
- Stopler, D., & Zamir, R. (2001, Mar.). Capacity and error probability in single-tone and multitone multiple access over an impulsive noise. *IEEE Trans. on Communications*, 49(3), 506-517.
- Styvaktakis, E., Bollen, M. H. J., & Gu, I. Y. H. (2002, Abril). Expert system for classification and analysis of power system events. *IEEE Trans. on Power Delivery*, 17(2), 423-428.
- Sutton, R. S. (1986). Two problems with backpropagation and other steepest-descent learning procedures for networks. In *The Eighth Annual Conference of the Cognitive Science Society* (p. 823-831).

- Tan, S. Y., Chew, Y. H., Tjhung, T. T., & Wu, M. W. (2003, Fev.). A two-dimensional impulse noise model and error probability of QAM in DSL. *IEEE Trans. on Communications*, 51(2), 129-134.
- Tanaka, M. (1988, Maio). High frequency noise power spectrum, impedance and transmission loss of power line in Japan on intrabuilding power line communications. *IEEE Trans. on Consumer Electronics*, 34(2), 321-326.
- T. C. Chuah, O. R. H., B. S. Sharif. (2001, Mai.). Robust adaptive spread-spectrum receiver with neural-net preprocessing in non-Gaussian noise. *IEEE Trans on Neural Networks*, 12(3), 546-558.
- Testa, A., Gallo, D., & Langella, R. (2004, Jan.). On the Processing of harmonics and interharmonics: using Hanning window in standard framework. *IEEE Trans. on Power Delivery*, 19(1), 28-34.
- Theodoridis, S., & Koutroumbas, K. (1999). *Pattern recognition*. San Diego: Academic Press.
- Thiele, L. (1986, Maio). On the sensitivity of linear state-space systems. *IEEE Trans. on Circuits Systems, CAS-33*, 502-510.
- Thong, T., & Liu, B. (1977, Abr.). Error spectrum shaping in narrow-band recursive filters. *IEEE Trans. on Acoustics, Speech, and Signal Processing, ASSP-30*, 200-203.
- Toumpakaris, D., Cioffi, J. M., & Gardan, D. (2004, Nov.). Reduced-delay protection of DSL systems against nonstationary disturbances. *IEEE Trans. on Communications*, 52(11), 1927-1938.
- Trees, H. L. V. (1968/1971). *Detection, estimation and modulation theory, Parts I and III*. NY: Springer-Verlag.
- Vaidyanathan, P. P. (1995). *Multirate systems and filter banks*. NJ: Englewood Cliffs, Prentice Hall.
- Vidakovic, B. (1995, Mar.). Nonlinear wavelet shrinkage with Bayes rules and Bayes factors. *Journal of the American Statistical Association*, 93(441), 173-179.
- Vitterli, M., & Kovacevic, J. (1995). *Wavelets and subband Coding*. Englewood

- Cliffs, NJ: Prentice Hall.
- Vogl, T. P., Mangis, J. K., Rigler, J. K., Zink, W. T., & Alkon, D. L. (1988). Accelerating the convergence of the backpropagation method. *Biological Cybernetics*, 59, 25-263.
- Vrahatis, M. N., Magoulas, G. D., & Plagianakos, V. P. (2000). Globally convergent modification of the quickprop method. *Neural Processing Lett.*, 12, 159-169.
- Waldeck, T., Busser, M., & Dostert, K. (1998). Telecommunication applications over the low voltage power distribution grid. In *Proc. of IEEE 5th International Symposium on Spread Spectrum Techniques and Applications* (Vol. 1, p. 73-77).
- Wang, L. X., & Mendel, J. M. (1993). Fuzzy adaptive filters, with application to nonlinear channel equalization. *IEEE Trans. on Fuzzy Systems*, 1, 161-170.
- Wang, M., & Mamishev, A. V. (2004, Julho). Classification of power quality events using optimal time-frequency representations - Part 1: Theory. *IEEE Trans. on Power Delivery*, 19(3), 1488-1495.
- Wang, M., Rowe, G. I., & Mamishev, A. V. (2004, Julho). Classification of power quality events using optimal time-frequency representations - Part 2: Application. *IEEE Trans. on Power Delivery*, 19(3), 1496-1503.
- Wang, X., & Chen, R. (2000, Jul.). Adaptive bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise. *IEEE Trans. on Signal Processing*, 47(7), 2013-2028.
- Wei, W., & Mendel, J. M. (1999). Maximum-likelihood classification for digital amplitude-phase modulations. *IEEE Trans. on Communications*, 48(2), 189-193.
- Wei, W., & Mendel, J. M. (2000). A fuzzy logic method for modulation classification in onideal environments. *IEEE Trans. on Fuzzy Systems*, 7(3), 353-344.
- Weinstein, S. B., & Ebert, P. M. (1971, Out.). Data transmission by frequency-division multiplexing using the discrete Fourier transform. *IEEE Trans. on Communications*, COM-19, 628-634.

- Widrow, B., & Hoff, M. E. (1985). *Adaptive signal processing*. NJ: Prentice Hall.
- Wijayakulasooriya, J. V., Putrus, G. A., & Minns, P. D. (2002, Jan.). Electric power quality disturbance classification using self-adapting artificial neural networks. *IEE Proc.-Gener. Transm. Distrib.*, 149(1), 98-101.
- Yang, H.-T., & Liao, C.-C. (2001, July). A de-noising scheme for enhancing wavelet-based power quality monitoring system. *IEEE Trans. on Power Delivery*, 16(3), 353-360.
- Youssef, A. M., Abdel-Galil, T. K., El-Saadany, E. F., & Salama, M. M. A. (2004, Jan.). Disturbance classification utilizing dynamic time warping classifier. *IEEE Trans. on Power delivery*, 19(1), 272-278.
- Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning - 1. *Inf. Sci.*, 8, 199-249.
- Zhang, H., Liu, P., & Malik, O. P. (2003, Set.). Detection and classification of power quality disturbances in noisy conditions. *IEE Proc.-Gener. Transm. Distrib.*, 150(5), 567-572.
- Zhang, X.-P. (2001, Maio). Thresholding neural network for adaptive noise reduction. *IEEE Trans. on Neural Networks*, 12(3), 567-584.
- Zhang, Y., & Blum, R. S. (2001, Março). Iterative decoding of codes over complex numbers for impulsive noise channels. *IEEE Trans. on Communications*, 49(3), 397-400.
- Zhidkov, S. V. (2003, Nov.). Impulsive noise suppression in OFDM based communication systems. *IEEE Trans. on Consumer Electronics*, 49(4), 944-948.
- Zhu, T. X., Tao, K., & Lo, K. L. (2004, Out.). Wavelet-based fuzzy reasoning approach to power-quality disturbance recognition. *IEEE Trans. on Power Delivery*, 19(4), 1928-1935.
- Zimmermann, M., & Dostert, K. (1999). A multi-path signal propagation model for the power line channel in the high frequency range. In *Proc. 3th International Symposium on Power-line Communication and its Applications* (p. 45-51).

Lancast, UK.

- Zimmermann, M., & Dostert, K. (2000a). An analysis of the broadband noise scenario in powerline networks. In *Proc. of 4th International Symposium on Power-Line Communications and its Applications* (p. 131-138).
- Zimmermann, M., & Dostert, K. (2000b, Jan.). The low voltage power distribution network as last miles access network - signal propagation and noise scenario in the HF-range. *AEÜ International Journal on Electronics and Communications*, 54(1), 13-22.
- Zimmermann, M., & Dostert, K. (2002a, Abr.). A Multipath model for the powerline channel. *IEEE Trans. on Communications*, 50(4), 553-559.
- Zimmermann, M., & Dostert, K. (2002b, Maio). Analysis and modeling of impulse noise in broad-band powerline communications. *IEEE Trans. on Electromagnetic Compatibility*, 44(1), 249-258.



Produção Científica

A.1 Artigos em Revistas

M. V. Ribeiro, R. Lopes, J. M. T. Romano, C. A. Duque, “Impulse Noise Mitigation based on Computational Intelligence for Improved Bit Rate in PLC-DMT,” *IEEE Trans. Power Delivery*, 2005. (aceito para publicação)

M. V. Ribeiro, J. G. A. Barbedo, J. M. T. Romano, A. Lopes, “Fourier-Lapped-Multilayer Perceptron (FLMLP) Method for Speech Quality Assessment,”

EURASIP Journal on Applied Signal Processing, Special Issue on Anthropomorphic Processing of Audio and Speech, 2005. (aceito para publicação)

C. A. Duque, M. V. Ribeiro, F. R. Ramos, J. Szczupak, “Power Quality Event Detection Based on the Principle of Divide to Conquer and Innovation Concept,” *IEEE Trans. on Power Delivery*, 2005. (aceito para publicação)

M. V. Ribeiro, “Learning Rate Updating Methods Applied to Adaptive Fuzzy Equalizers for Broadband Power Line Communications,” *EURASIP Journal on Applied Signal Processing, Special Issue on Nonlinear Signal and Image Processing*, vol. 16, pgs. 2592 - 2599, 2004.

M. V. Ribeiro, J. M. T. Romano, C. A. Duque, “An Improved Method for Signal Processing and Compression in Power Quality Evaluation,” *IEEE Trans. on Power Delivery*, vol. 19, no. 2, pgs. 464-471, Apr. 2004.

M. V. Ribeiro, “Packets Reconstruction Proposals applied to Waveform Coding Techniques for VoIP Solution - Real Time Implementation,” *IEEE Latin America Transactions*, vol. 2, no. 1, Mar. 2004.

M. V. Ribeiro, “Packets Reconstruction Proposals applied to Waveform Coding Techniques for VoIP Solution,” *IEEE Latin America Transactions*, vol. 1, no. 1, Oct. 2003.

A.2 Artigos em Congressos

A. S. Cerqueira, C. A. Duque, M. V. Ribeiro e R. M. Trindade, “Redes Neurais na Classificação de Eventos de Qualidade de Energia Elétrica: Uma Abordagem Comparativa,” *VI Seminário Brasileiro sobre a Qualidade da Energia Elétrica*,

Belém-PA, Brasil, 2005. (aceito para publicação)

A. S. Cerqueira, C. A. Duque, R. M. Trindade, M. V. Ribeiro, “Digital System for Detection and Classification of Electrical Events,” *IEEE International Symposium on Circuits and Systems*, Kobe, Japão, 2005. (aceito para publicação)

M. V. Ribeiro, S. K. Mitra, J. M. T. Romano, “Digital Signal Processing Techniques for Compression of Power Quality Disturbance Events: New Results,” *IEEE PES T&D 2004 Latin America*, São Paulo, Brasil, 2004.

M. V. Ribeiro, S. K. Mitra, J. M. T. Romano, “A Novel Approach for Spectral Analysis of Monitored Power Systems,” *IEEE 11th International Conference on Harmonic and Quality of Power (ICHQP' 04)*, Lake Placid, Nova Iorque, EUA, 2004.

M. B. Loiola, M. V. Ribeiro, J. M. T. Romano, “A Turbo Equalizer Using Fuzzy Filters,” *IEEE International Workshop on Machine Learning for Signal Processing (MLSP'04)*, São Luís, Brasil, 2004.

J. M. T. Romano, J. Pissolato Filho, M. C. D. Tavares, M. V. Ribeiro, M. B. Loyola, G. D. Mendes, “Caracterização de Canais PLC (Power line Communications) e Desenvolvimento de Técnicas para Equalização e Codificação,” *Workshop de Pesquisa e Desenvolvimento UNICAMP-CPFL*, Campinas, Brasil, pgs. 33-35, 2003.

J. M. T. Romano, S. M. Deckmann, M. V. Ribeiro, D. A. Shayani, “Uma Solução Integrada, baseada em Técnicas de Processamento de Sinais, para o Monitoramento da Qualidade da Energia Elétrica,” *Workshop de Pesquisa e Desenvolvimento UNICAMP-CPFL*, Campinas, Brasil, pp. 42-44, 2003.

M. V. Ribeiro, J. M. T. Romano, C. A. Duque, “An Improved Method for

Signal Processing and Compression in Power Quality Evaluation,” *2003 IEEE Power Engineering Society General Meeting*, Toronto, Canada, pp. 13-17.

M. V. Ribeiro, C. A. Duque, J. M. T. Romano, “Implementação Computacional de um Simulador para Transmissão broadband via Canais PLC,” *XVII SNPTEE Seminário Nacional de Produção e Transmissão de Energia Elétrica*, Uberlândia, Brasil, 2003.

M. V. Ribeiro, C. A. Duque, J. M. T. Romano , “Técnicas de Processamento de Sinais Aplicadas ao Monitoramento Descentralizado da Qualidade da Energia Elétrica,” *XVII SNPTEE Seminário Nacional de Produção e Transmissão de Energia Elétrica*, Uberlândia, Brasil, 2003.

M. V. Ribeiro, “On Fuzzy-DFE-LMS and Fuzzy-DFE-RLS Algorithms to Equalize Power Line Channels,” *IEEE International Symposium on Industrial Electronics*, Rio de Janeiro, Brazil, pp. 1001-1006, 2003.

M. V. Ribeiro, J. M. T. Romano, S. M. Deckmann, “Adaptive Filtering, Wavelet and Lapped Transforms for Power Quality Problem Detection and Identification,” *IEEE International Symposium on Industrial Electronics*, Rio de Janeiro, Brasil, pgs. 301-306, 2003.

J. G. A. Barbedo, M. V. Ribeiro, A. Lopes, J. M. T. Romano, “Estimation of the Subjective Quality of Speech Signals using the Kohonen Self-Organizing Maps,” *IEEE ITS International Telecommunication Symposium*, Natal, Brasil, pgs. 834-839, 2002.

M. V. Ribeiro, C. A. Duque, “The Word Length Influence on Waveform Coding Techniques based on Wavelet Transform applied to Disturbance Compression,” *IEEE 10th International Conference on Harmonic and Quality of Power*, Rio de

Janeiro, Brasil, pgs. 139-143, 2002.

F. R. Ramos, M. V. Ribeiro, J. M. T. Romano, C. A. Duque, “On Signal Processing Approach for Event Detection and Compression applied to Power Quality Evaluation,” *IEEE 10th International Conference on Harmonic and Quality of Power*, Rio de Janeiro, Brasil, pgs. 133-138, 2002.

J. G. A. Barbedo, M. V. Ribeiro, F. J. Von Zuben, A. Lopes, J. M. T. Romano, “Application of Kohonen Self-Organizing Maps to Improve the Performance of Objective Methods for Speech Quality Assessment,” *EUSIPCO XI European Signal Processing Conference*, Toulouse, França, pgs. 519-522, 2002.

J. B. Destro-Filho, M. V. Ribeiro, “A Software Technique for Fault-tolerant Adaptive Circuits in Spatial Applications,” *IEEE LATW Latin-American Testing Workshop*, Montevideo, Uruguai, pgs. 26-31, 2000.

M. V. Ribeiro, C. A. Duque, J. M. T. Romano, “Enhanced Data Compression Method Applied to Power Quality Applications,” *IEEE IECON The 27th Annual Conference of the Industrial Electronics Society - '01*, Denver, EUA, pgs. 676-681, 2001.

C. Junqueira, M. V. Ribeiro, J. M. T. Romano, J. B. D. Filho, “Uma Solução Adaptativa Combinada Aplicada ao Cancelamento de Sinais Interferentes na recepção GPS,” *XIX SBrT Telecommunication Brazilian Symposium*, Fortaleza, Brasil, 2001.

C. Junqueira, M. V. Ribeiro, J. M. T. Romano, J. B. D. Filho, C. Lima, “A Hybrid Algorithm Solution for GPS Antenna Array,” *KIS International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Banff, Canadá, 2001.

A.3 Premiações

Student Award do *2003 IEEE International Symposium on Industrial Electronics (ISIE'03)*. Artigos: “On Fuzzy-DFE-LMS and Fuzzy-DFE-RLS Algorithms to Equalize Power Line Channels” e “Adaptive Filtering, Wavelet and Lapped Transforms for Power Quality Problem Detection and Identification”.

Student Award do *27th Annual Conference of the 2001 IEEE Industrial Electronics Society (IECON'01)*. Artigo: “Enhanced Data Compression Method Applied to Power Quality Applications”.

A.4 Patentes

A. Lopes, J. M. T. Romano, M. V. Ribeiro, J. G. A. Barbedo, C. A. M. Lima, “Método FL-PMC (Fourier Lapped - Perceptron Multicamadas) para Estimação da Qualidade de Sinais de Voz,” Pedido de Patente INP: PI 0204932-5, 06/11/2002.