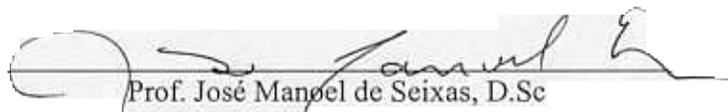


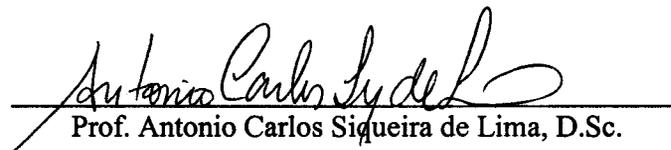
CARACTERIZAÇÃO DOS SINAIS DE DESCARGAS PARCIAIS EM
EQUIPAMENTOS DE ALTA TENSÃO A PARTIR DE MODELOS
EXPERIMENTAIS

Walter Martin Huamán Cuenca

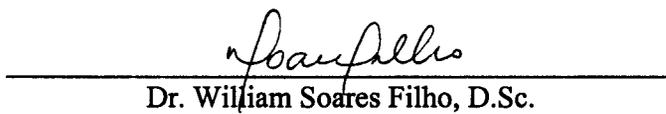
TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM
ENGENHARIA ELÉTRICA.

Aprovado por:


Prof. José Manoel de Seixas, D.Sc


Prof. Antonio Carlos Siqueira de Lima, D.Sc.


Prof. Luiz Pereira Calôba, Dr.Ing


Dr. William Soares Filho, D.Sc.


Prof. Luiz Cera Zanetta Júnior, D.Sc.

Rio de Janeiro, RJ – BRASIL.

Junho de 2005

HUAMÁN CUENCA, WALTER MARTIN

Caracterização dos Sinais de Descargas
Parciais em Equipamentos de Alta Tensão a partir
de Modelos Experimentais [Rio de Janeiro] 2005
XIV, 140 p. 29,7 cm (COPPE/UFRJ, D.Sc.,
Engenharia Elétrica, 2005)

Tese - Universidade Federal do Rio de
Janeiro, COPPE

1. Caracterização de Sinais 2. Descargas Parciais
3. Modelos experimentais 4. ICA/PCA e IA 5.
Eliminação de ruídos 6. Gerador e Transformador

I. COPPE/UFRJ II. Título (série)

A meus queridos pais, Jorge e Gabriela, a minha mais profunda gratidão pelo grande esforço em brindarem uma valiosa formação profissional e uma verdadeira lição de vida e a minha amada família Karyn e Adriana.

A mis queridos padres, Jorge y Gabriela mi más profunda gratitud por el grande esfuerzo que me brindaron, una valiosa formación profesional y una verdadera lección de vida y a mi amada familia Karyn y Adriana que siempre estuvieron a mi lado.

AGRADECIMENTOS

Ao Prof. Sandoval Carneiro Jr, orientador deste trabalho, por suas valiosas sugestões e críticas ao longo da execução deste trabalho.

Ao Prof. José Manoel de Seixas, também orientador, pela confiança em mim depositada e pelo grande aprendizado durante todo este tempo de convívio, além de demonstrar um constante apoio e incentivo no encaminhamento no mundo das publicações de artigos e, desenvolvimento deste trabalho.

Agradeço a Alain Fraçois S. Levy, pela colaboração constante no desenvolvimento dos modelos experimentais e na execução das medições.

Ao Professor Carlos Portela, pelo ensino, pela atenção e pelos esclarecimentos das ambigüidades em descargas parciais.

Agradeço, em geral, à COPPE pela oportunidade de me aceitarem como estudante de doutorado e pela grande experiência no ensino. Também agradeço ao CEPTEL/DIE pela valiosa ajuda e aceitação como bolsista oferecendo suporte financeiro, instalações de informática e laboratórios para a realização deste trabalho, com muita satisfação.

Agradeço ao coordenador da área de conhecimentos de eletro-técnica, pela habilitação e permissão de efetuar este trabalho, desenvolvido dentro do projeto (Sistema Inteligente de Diagnóstico Projeto CEPTEL 2094).

Aos colegas que tiveram sua compreensão e paciência em me suportar.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

CARACTERIZAÇÃO DOS SINAIS DE DESCARGA PARCIAIS EM EQUIPAMENTOS DE ALTA TENSÃO

Walter Martin Huamán Cuenca

Junho/2005

Orientadores: Sandoval Carneiro Junior

José Manoel de Seixas

Programa: Engenharia Elétrica

Nesta tese são realizados estudos de aplicação de métodos e técnicas de processamento de sinais, viáveis para caracterizar as fontes de sinais de descargas parciais (DPs). Destaca-se a importância da detecção e medição dos sinais de DPs e o problema de ruídos sobre estes, uma vez que o ruído é um dos principais obstáculos na caracterização destes sinais, assim como nos sistemas de medição, monitoração e diagnóstico de equipamentos em alta tensão. Além dos métodos tradicionais de processamento de sinais, emprega-se a análise de componentes independentes (ICA), propondo-se um modelo de sistema integrado para explorar e caracterizar fontes de sinais de DPs. Para a realização deste trabalho, utiliza-se um banco de dados contendo observações proveniente de medições experimentais dos padrões de defeitos de DPs mais representativos em transformadores e geradores.

Vale ressaltar que a pesquisa além de estar orientada para sistemas de medição pode também ser aplicada para os processos de monitoração, diagnóstico de equipamentos e para fins de purificação e compactação (qualidade de informação) em certos bancos de dados.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

CHARACTERIZATION OF THE SIGNALS OF PARTIAL DISCHARGE IN HIGH VOLTAGE EQUIPMENTS

Walter Martin Huamán Cuenca

June/2005

Advisors: Sandoval Carneiro Junior

José Manoel de Seixas

Department: Electric Engineering

In this thesis are realized studies of application of methods and techniques in the processing of signal, viable to characterize the sources of signals of partial discharge (DPs). Stand out the importance of the detection and measurement of the signals of DPs and the problem of noise upon those. The noise is the principal obstacles in characterization, so as in measurement systems, monitoring and diagnostic of the equipments in high voltage. Besides the traditional methods, for the processing of the signals, the independent components analyzes (ICA) is used, in order to investigate and, for propose a model of integrated system to explorer and to characterize DPs signals sources. In order to realize this works, advantage is taken of a data bank including observations derived of the experimental measurements of the fault patterns in DPs more representatives in power transformer and hydro generators.

The investigations, be important for measurement systems, can also be applied for monitoring, equipment diagnosis and purpose to compact data banks (quality information).

Índice

AGRADECIMENTOS	iv
Índice	vii
Lista de Figuras	x
Lista de Tabelas	xii
Capítulo 1 Introdução	1
1.1 Visão Global e Motivação	2
1.2 Objetivo	4
1.3 Contribuições.....	5
1.4 Escopo da Tese	5
Capítulo 2 Revisão Literária e Conceitos Básicos	8
2.1 Revisão Literária e Inovações.....	9
2.2 Revisão Teórica	13
2.3 Mecanismos das Descargas Parciais	13
2.3.1 Ionização dos átomos de um material isolante	13
2.3.2 Algumas considerações sobre o Campo Elétrico	17
2.3.3 O campo elétrico e o meio em que ele se encontra	17
2.3.4 O campo elétrico e a geometria dos eletrodos	20
2.4 Tipos das Descargas Parciais.....	21
2.5 Conceito de Descargas Parciais.....	23
2.6 Detecção e Medição de Descargas Parciais.....	25
2.7 Causas e Conseqüências de Descargas Parciais	27
2.8 Circuitos de Medição e Ensaio de DPs	28
2.9 Interferências na Detecção de DPs e Formalização do Problema	33
Capítulo 3 Modelos Experimentais.....	35
3.1 Detecção de Descargas Parciais	35
3.2 Medição de Descargas Parciais Através de Modelos	36
3.3 Modelagem de Defeitos Típicos.....	37
3.3.1 Célula Teste de Ensaio	37
3.3.2 Barras de Prova.....	38
3.4 Calibração do Circuito de Ensaio	40
Capítulo 4 Pesquisas Anteriores e Classificadores Inteligentes em DPs	43
4.1 Princípios de Reconhecimento de Padrões	43
4.2 Classificadores e Reconhecimento de Padrões Inteligentes	44
4.2.1 Redes Neurais Artificiais.....	44
4.2.2 Híbridos Neuro-Fuzzy (MANFIS)	46
4.2.3 Simulated Annealing em DPs.....	47
4.2.4 Simulated Annealing e MANFIS	49

4.2.5	Algoritmos Genéticos e MANFIS em DPs.....	49
	Montagem de um Algoritmo Genético.....	50
	Implementação de AGs no MANFIS	50
Capítulo 5	Caracterizadores de Sinais e Tratamento de Ruídos	51
5.1	Pré-Processamento de Dados em DPs.....	52
5.2	Compactação de Dados por Fatores Principais	52
5.3	Análise de Componentes Principais	53
5.4	Introdução a Separação Cega de Fontes.....	54
5.5	Caracterizadores de Sinais em Fontes de DPs.....	56
5.6	Fundamentos - Técnicas de Caracterização.....	56
5.6.1	PCA Linear	56
	Fundamentos da Análise de Componentes Principais.....	57
5.6.2	PCA por Redes Neurais e PCA Não Linear	60
5.6.3	Filtros Casados	61
	Detecção de ruídos impulsivos por MF.....	61
5.6.4	Análise de Componentes Independentes.....	63
5.6.5	Separação Cega de Fontes	64
	Identificabilidade e ambigüidade do modelo ICA.....	66
	Indeterminações no modelo ICA.....	67
	Funções Objetivo em ICA	67
5.6.6	Algoritmos ICA	71
	Algoritmo do Infomax	71
	Algoritmo do FastICA	72
	Algoritmo do JADE.....	73
5.6.7	Extensão de Algoritmos ICA – Noisy_ICA	74
5.6.8	ICA para dados contaminados por ruído	74
	Ruído de Sensor frente a Ruído de Fonte.....	74
	Remoção do Bias.....	76
	Estimação das componentes independentes por funções shrinkage.....	77
Capítulo 6	Caracterizadores: Implementação e Resultados	79
6.1	PCA para Identificar Sinais de DPs em Transformadores de Potência.....	79
6.1.1	Componentes Principais	81
6.1.2	Mapeamento dos Componentes Principais.....	83
6.1.3	Classificação Neural.....	85
6.1.4	Conclusões Referentes a PCA	88
6.2	Caracterizador por Filtros casados em Identificação DPs.....	88
6.2.1	Sistema Discriminador Híbrido de Sinais	88
6.2.2	Resultados do Sistema Caracterizador Híbrido.....	91
6.2.3	Conclusões referentes ao sistema MF/ANN.....	94
6.3	Caracterizador ICA para Extração Cega de Ruídos	95
6.3.1	Representação como um Problema BSS	96
6.3.2	Descrição do ICA não-Holonômico	97
6.3.3	Separação de Fontes e Ruídos	100
6.3.4	Relação SNR e Crosstalks	100
6.3.5	Seqüência experimental.....	100
6.3.6	O ICA como Extrator de Ruídos	102
6.3.7	Resultados Referentes ao ICA.....	103
6.4	Caracterizador de Fontes Ruidosas de DPs por ICA Shrinkage.....	106

6.4.1	Descrição do ICA Shrinkage	107
	Modelos de densidade esparsa.....	107
	Modelo duplo laplaciano para remoção de ruídos em DPs	112
Capítulo 7	Conclusões e Análise dos Resultados	114
7.1	Escopo para Trabalhos Futuros	116
Apêndice A	Conceitos Básicos de Estatística para Métodos de Separação de Fontes.....	117
A1.1	Introdução	117
A1.2	momentos e cumulantes.....	117
A1.2.1	Momentos	118
A1.2.2	Cumulantes	119
A1.2.3	Distribuições de probabilidade multivariável. Cumulantes cruzados...	121
A1.3	Funções de densidade de probabilidade de sinais aleatórios.....	124
A1.3.1	Sinais com densidade de probabilidade uniforme.	124
A1.3.2	Sinal com densidade de probabilidade gaussiana (normal).....	125
A1.3.3	Função de densidade de probabilidade gamma.	127
A1.3.5	Função densidade de Probabilidade	131
Apêndice B	Publicações em Congressos	133
B1.1	Artigos em Congressos.....	133
	Referências Bibliográficas	135

Lista de Figuras

Figura 2.1 – Átomo de um material dielétrico a) Átomo não polarizado. b) Forças no átomo quando se aplica \vec{E} [1].....	15
Figura 2.2 – Processo de avalanche de elétrons iniciado a partir de um eletrodo negativo a) Início. b) Formação de um par de elétrons – Impacto de um elétron com um átomo neutro libera um elétron adicional e deixa um íon positivo para trás. c) Multiplicação – Os elétrons se movem criando íons positivos quando se multiplicam.....	15
Figura 2.3 – Fronteira entre dois meios dielétricos diferentes.	19
Figura 2.4 – Eletrodo ponta-plano. [1]	20
Figura 2.5 – Eletrodos: duas pontas paralelas com resistores e um eletrodo plano [1]..	21
Figura 2.7 – Descargas em materiais isolantes sólidos (treeing channels): Trajetória de propagação de uma treeing channels em volta de uma barreira de base de mica.	23
Figura 2.8 – Sinais de DPs: (a) Pulsos intermitentes medidos em ultra-alta frequência, (b) Queda de potencial por DPs em cavidades, (c) e (d) mapeamento típicos de DPs gerados por modelos experimentais.....	24
Figura 2.9 – Circuitos básicos de medição de DPs [16]. a) Impedância Z de medição em série com o capacitor C de acoplamento. b) Impedância de medição Z_m em série com o equipamento sob ensaio. c) Circuito de ensaio equilibrado ou balanceado.	29
Figura 2.10 – Esquema de medição de DPs para transformadores no campo.....	31
Figura 2.11 – Localização de descargas por sensores de campo elétrico.....	32
Figura 2.12 – Corte transversal e longitudinal de barras estatóricas de gerador, zonas possíveis de formação de DPs.	33
Figura 2.13 – Classificação ruídos em detecção de sinais DPs [32], [30].....	34
Figura 3.1 – Circuito básico de medição de DPs em geradores: Estratégia praticada no enrolamento estatórico da fase C no gerador U10 na usina Itaipu [35], onde AT (alta tensão), NC (cubículo de neutro), P (acoplador de 9 nF) e T (acoplador de 150 nF-fase e 160 nF-neutro).	35
Figura 3.2 – Enrolamento estatórico, onde cada fase é constituído por um conjunto de barras.	36
Figura 3.3 – Mapas digitais de DPs obtidos pelo ICM: (a) DPs em 3-D $f(\phi, q, N)$ e (b) é Mapa 2-D $H(\phi, q)$	37
Figura 3.4 – Células de ensaio para medições de DPs reais. a) Ponta flutuante b) Ponta plana c) Ponta terra.....	38
Figura 3.5 – Modelos experimentais por barras de prova para reproduzir defeitos típicos no enrolamento estatórico de geradores.	39
Figura 3.6 – Montagem de barra de prova e dos capacitores de acoplamento.	40
Figura 3.7 – Diagrama do circuito de ensaio e calibração [17], [43], C_k é capacitor de acoplamento.....	41
Figura 4.1 – Rede neural artificial.....	45
Figura 4.2 - Sistema Híbrido Neuro-Fuzzy MANFIS.....	47
Figura 4.3 – Simulated annealing estocástico de uma função de energia [40].....	48
Figura 5.1 – Banco de filtros casados para detecção de ruídos pulsantes [72].....	63

Figura 5.2 – Modelo básico BSS sem ruído. Sinais não observáveis: S, observações: X, sinais fonte estimadas: y.....	64
Figura 5.3 – Modelo básico do BSS instantâneo linear com componente de ruído [53]	65
Figura 6.1 – Amostras de mapas característicos dos padrões FL, PP e PT originais (a esquerda) e reduzidos (à direita).	80
Figura 6.2 - As primeiras componentes apresentam as maiores energias $E[C^2]$. maiores energias $E[C^2]$	82
Figura 6.3 - As primeiras componentes apresentam as maiores porcentagens das variâncias.....	83
Figura 6.4 - Os nove primeiros componentes apresentam maiores porcentagens das variâncias.....	84
Figura 6.5 - Componentes principais do PC10 ao PC18.	85
Figura 6.6 - Esquema do classificador baseados em CPs e RN.	85
Figura 6.7 - Eficiência média de classificação em função do número de componentes principais.	87
Figura 6.8 - Filtro casado linear.	89
Figura 6.9 - Sistema discriminador híbrido: MF+ANN	90
Figura 6.10 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é FL.....	92
Figura 6.11 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é PP	93
Figura 6.12 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é PT.	94
Figura 6.13 – Fontes de DPs provenientes das medições no gerador U10 da usina Itaipu (conforme Tabela 5.6).	96
Figura 6.14 – Espaço homogêneo das matrizes, dB é o espaço de tangentes [85].....	99
Figura 6.15 – Amostra de fontes com 20% de ruído.....	101
Figura 6.16 – Fontes estimadas com SNR 13,40dB pelo ICA sem pré-branqueamento.	102
Figura 6.17 – Fontes estimadas com SNR 5dB pelo ICA sem pré-branqueamento.	104
Figura 6.18 – Fontes estimadas com SNR 13,40dB pelo ICA sem pré-branqueamento.	104
Figura 6.19 – Esquerdo: Observações com 40%-6,29dB de ruído, Direito: Fontes estimadas pelo ICA não-holonômico com SNR= 6,29dB.....	105
Figura 6.21 – Dado coletado de DPs: histograma com 256x256 canais (bins).....	108
Figura 6.22 – Modelo duplo gaussiano	108
Figura 6.23 – Modelo laplaciano para Noisy_ICA em dados DPs.....	109
Figura 6.24 – Modelo gaussiano para resíduos em dados DPs	109
Figura 6.25 – Modelos de componentes esparsos. Linha sólida: densidade laplaciana. Linha traçada: uma típica densidade moderadamente esparsa (super-gaussiana). Linha de traçada e pontilhada: uma densidade típica fortemente esparsa (super-gaussiana). Linha pontilhada: densidade gaussiana para referencia [80], [76].	110
Figura 6.26 – Funções shrinkage. Linha da cor vermelha: correspondente à densidade laplaciana. Linha da cor azul: shrinkage correspondente à densidade moderadamente esparsa (supergaussiana), linha da cor preta: shrinkage correspondente à densidade fortemente esparsa (supergaussiana). Linha cor amarela: Linhas de referencia $x=y$, $x=0$, [76].	111
Figura 6.27 – Família de Funções shrinkage.....	112
Figura 6.28 – Modelo de função proposto para remoção de ruídos	113
Figura 6.29 – fonte estimada pelo caracterizador ICA Shrinkage.....	113

Lista de Tabelas

Tabela 2.1 – Permissividade de meios dielétricos	18
Tabela 2.2 – Rigidez dielétrica de diversos materiais dielétricos comuns	20
Tabela 6.1 – Conjuntos de caracterização e validação	81
Tabela 6.2 - Eficiência de classificação para cada padrão usando-se cinco componentes principais	86
Tabela 6.3 - Eficiência de classificação para cada padrão considerando-se sete componentes principais	87
Tabela 6.4 - Eficiência parcial alcançada por cada padrão na saída da rede neural.....	93
Tabela 6.5 - Comparação de desempenhos entre as técnicas de reconhecimento padrão (PCA+ANN) e o discriminador híbrido (MF+ANN).....	94
Tabela 6.6 - Fontes de Sinais de DPs	95
Tabela 6.7 - Parâmetros Estatísticos e desempenho por Crosstalks	103
Figura 6.20 – Ganho obtido pelo caracterizador 30,43dB : Direita: Fontes originais Esquerda: Fontes estimadas com 6,29DB de ruído. Entre estas duas fontes tiveram ganhos que podem observar-se visualmente.	105
Tabela 6.8 – Quadro comparativo de três Modelos ICA.....	106

Acrônimos e Símbolos

DPs	Descargas Parciais ou Sinais de Descarga Parcial
DP	Descarga Parcial
PT	Eletrodo de Ponta a Terra - Célula
PP	Eletrodo de Ponta Plana - Célula
PF	Eletrodo de Ponta Flutuante – Célula
RF	Rádio Interferência
RLC	Impedância de Detecção Resistiva, Indutiva e Capacitiva.
RC	Resistiva e Capacitiva de Banda Larga de Frequências
SIG	Subestações Isoladas a Gás
∇V	Diferença de potencial
q_0	Amplitude de Carga elétrica
ICM	Nome de Fabricação do Instrumento de Medição de Descargas Parciais
RNA ou ANN	Rede neural artificial (“Artificial Neural Network”)
GHA	Algoritmo hebbiano generalizado (“Generalized Hebbian Algorithm”)
ICA	Análise de Componentes Independentes (“Independent Component Analysis”)
BSS	Separação Cega de Fontes (Blind Separation Source)
HOS/HOSA	Estatística de ordens superiores ou altas ordens (“High order Statistics”)
MF	Filtro Casado (“Matched Filter”)
ME	Máxima Entropia (“Maximum Entropy”)
MI	Informação Mútua (“Mutual Information”)
ML	Máxima Semelhança (“Maximum Likelihood”)
PCA	Análise de Componentes Principais (“Principal Components Analysis”)
FA	Análise de Fatores (“Factors Analysis”)
RMS	Erro quadrático Médio (“Root Mean Square”)
SNR	Relação Sinal/Ruído (“Signal to Noise Rate”)
MANFIS	Sistema de Inferência Fuzzy de Mandami
AG ou GA	Algoritmo Genético
SA	<i>Simulated Annealing</i>
σ	Desvio padrão
Φ	Função de contraste
μ	Valor médio
$\mu(x,y)$	Função de pertinência fuzzy
$N(\mu, \sigma^2)$	Distribuição gaussiana de media μ y variância σ^2
$\eta(m)$	Ganho de aprendizagem no instante m
Φ^0	Função de contraste ortogonal
σ^2	Variância
λ_j	Valores próprios ou Autovalores
$\kappa_{y_i,k}$	Cumulante (auto-cumulante) de ordem k da variável aleatória y_i

$\langle s_i \rangle$	valor médio da variável aleatória s_i
A	Matriz de mistura
B	Matriz de reconstrução das fontes a partir das observações
$C_{1\dots k}[\mathbf{y}] = \text{Cum}[y_1, \dots, y_k] = \kappa_{y, 1\dots k}$	Cumulante de ordem k das variáveis aleatórias y_1, \dots, y_k
D	Matriz diagonal
$D(f_1//f_2; \mathbf{B})$	Divergência de Kullback-Leibler entre as funções de distribuição f_1 y f_2 expressa em função de B
$E\{x\}$	Esperança matemática da variável x
$\mathbf{e} = (e_i)^T$	Vetor de sinais observadas
\mathbf{A}^{-1}	Inversa de uma matriz de A
\mathbf{A}^T	Matriz transposta de A ;
\mathbf{s}^T	vetor transposto de s
F_s	Freqüência de amostragem
$f_{si}(s_{si})$	Função de densidade de probabilidade do sinal aleatório s_i
$H(\mathbf{X})$	Entropia do sinal X
$h(\mathbf{X})$	Entropia diferencial de X
$I(\mathbf{X}; \mathbf{Y})$	Informação mútua entre X e Y
I	Matriz diagonal unitária
$L(\bullet)$	Função logarítmica de máxima semelhança
m	Número de amostra, passo de iteração $t = mT_s$
M	Número de amostras
$\mathbf{n} = (n_i(t))^T$	Vetor de ruído
n_c	Número de classes
P	Matriz de prestações
p	Número de fontes
R	Matriz de correlação
$\mathbf{s} = (s_i)^T$	Vetor de fontes o sinais originais
$\text{máx}(M)$	Índice de valor máximo
T_s	Período de amostragem
$U(s_m, s_M)$	Distribuição uniforme de faixa $[s_m, s_M]$
v	vetor sinal de ruído
V	Matriz de transformação de branqueio
W	Matriz de pesos de uma rede neural ou matriz que estima a matriz de mistura
A , dependendo do contexto.	
$\mathbf{y} = (y_i)^T$	Vetor de sinais recuperados ou estimados
Z	Matriz de permutação
z	Vetor de sinais branqueados

Capítulo 1

Introdução

Nesta tese, são realizados estudos de aplicação de alguns métodos e técnicas, baseados em processamento de sinais e inteligência artificial, com o propósito de se conseguir características dos fenômenos de Descargas Parciais (DPs) mais próximas da sua natureza.

A importância que envolve os fenômenos de DPs em equipamentos de alta tensão motiva os fabricantes e, especialmente, as concessionárias do setor elétrico que lidam com estes equipamentos a prestar uma maior atenção em sua prevenção, preservação e também sua administração do tempo de vida útil. Assim, detecção e medição de DPs, atualmente, é uma das ferramentas mais utilizadas, tanto em laboratórios como no campo, para detectar defeitos que possam ter sido originados na fabricação e/ou serviço desses equipamentos. A maioria dos casos de detecção e/ou medição dos sinais de DPs estão corrompidos por outros fenômenos parasitas acoplados, geralmente sinais de ruído que contaminam o sinal desejado. Muitas vezes o ruído se apresenta com maior amplitudes e elevadas taxas de acumulação, com respeito aos sinais de DPs, o que dificulta a detecção dos sinais desejados e conseqüentemente sua caracterização.

Diversos estudos para tratamento e filtragem de ruídos vêm utilizando convencionais filtros analógicos até os mais sofisticados com tecnologia digital. Mesmo assim, os ruídos acoplados são difíceis de serem minimizados ou removidos, tal como acontece nos casos em que não se conhece nenhum de seus parâmetros.

Por outro lado, há necessidade de garantir a robustez no funcionamento e na qualidade do serviço de energia elétrica. Tanto que novos estudos para medição, detecção e interpretação de DPs têm sido empreendidos por muitas fontes de pesquisa. Conseqüentemente, diversos avanços na tecnologia digital em hardware e software estão sendo adotados para se buscar novas estratégias que conduzam a minimizar e/ou eliminar os ruídos parasita, sendo esta tese um destes empreendimentos.

A tese desenvolve uma estratégia na tentativa de avaliar o desempenho de algumas ferramentas que ajudam a caracterizar os sinais de DPs, especialmente focalizando sua possível integração para os sistemas de medição no campo ou laboratório. As fontes de informação, utilizadas nesta pesquisa, são provenientes de alguns equipamentos de alta tensão, tais como geradores, transformadores e modelos experimentais representativos de fontes de DPs.

A seguir, descreve-se o roteiro da tese apresentando uma visão global do marco situacional e motivação do trabalho de pesquisa, assim como os objetivos, contribuições e o escopo geral que compreende os capítulos da dissertação.

1.1 Visão Global e Motivação

A indústria de potência, neste novo cenário de economia de mercado, traz fortes tendências, incertezas e novas funções que são impostas com o intuito de gerar maiores utilidades elétricas, provocando mudanças obrigatórias em sua infra-estrutura de comunicação e informação. Expandir redes de serviços, como diagnóstico e medições remotas, em tempo real, e monitoração, leva à necessidade de se ter maior largura de banda nas redes de comunicação e em certos instrumentos de medição, para torná-los mais eficientes e persuasivos. Novas aplicações de controle e proteção remotas em tempo real também reforçam esta necessidade.

O atual cenário do mercado brasileiro, envolvendo desregulamentação e privatização, trouxe novos desafios visando maior utilidade e drástica redução nos custos de manutenção dos equipamentos. Os níveis de manutenção se tornaram indicadores de reestruturação e qualidade gerencial nas empresas concessionárias de energia elétrica. A redução da manutenção a níveis ótimos é o desafio que as empresas têm que buscar continuamente. Para este propósito, estender o tempo de serviço (vida útil) e o tempo entre manutenções dos equipamentos de alta tensão (geradores, transformadores, etc.) são imposições para novas condições de serviço.

Conseqüentemente, essas novas condições, somadas a um possível crescimento da demanda, fazem com que os equipamentos trabalhem com sobrecarga por longos períodos de tempo. Além disso, projetos de equipamentos, atrelados a normas deficientes impostas pelos fabricantes fazem com que estes não sejam adequados para condições críticas de funcionamento. Por conseguinte, estes e outros fatores podem conduzir a um aumento dos riscos de falha, degradando os componentes isolantes,

causando a perda de suas propriedades elétricas e mecânicas e, contribuindo para uma possível formação de DPs e perdas econômicas.

Neste cenário de mudanças, cada vez mais o uso de técnicas adequadas de medição, monitoração e diagnóstico do estado operativo dos sistemas e equipamentos elétricos, com a finalidade de melhorar sua confiabilidade e reduzir o número de paradas não programadas e eventuais faz-se necessário.

Atualmente, as empresas de energia elétrica estão prestando cada vez mais atenção na preservação de seus equipamentos mais importantes, especialmente transformadores de potência, disjuntores e geradores, para garantir um serviço contínuo de energia elétrica, ainda com uma tímida inversão no monitoramento e automação de suas subestações.

DPs podem ser sintetizadas como sinais elétricos pulsantes incompletos, intermitentes e rápidos, da ordem dos nanosegundos, que ocorrem pela proximidade entre duas partes condutoras de eletricidade e o meio isolante, através do efeito de ionização em cavidades gasosas, no interior dos materiais isolantes e nas interfaces condutor-isolante ou isolante-isolante [1].

Medir DPs é uma das ferramentas mais importantes para o diagnóstico de equipamentos de alta tensão como hidrogeradores, transformadores e disjuntores. Os sinais de DPs ajudam a revelar o envelhecimento prematuro dos materiais isolantes, defeitos de fabricação, assim como prevenir saídas de serviço não programadas. DPs vêm sendo pesquisadas por várias décadas, formando-se linhas de pesquisa em medição, detecção e localização. A realização de uma medição de DPs em equipamentos de alta tensão não só é dificultada pela acessibilidade e complexidade dos circuitos internos, como também por diversas fontes de ruído poluentes, que se acoplam ao sinal e ao sistema de medição. Instituições de pesquisa e fabricantes de equipamentos de medição de DPs trouxeram avanços significativos na eliminação de alguns tipos de ruídos, através de técnicas de processamento digital; porém, estes avanços tecnológicos estão conduzindo para a diversificação que, se por um lado abre novas formas possíveis de ver o fenômeno de DPs. Por outro, inibe o estabelecimento de procedimentos padrão com relação à identificação de fontes características de DPs. Se bem que, as entidades normativas estão buscando a unificação de métodos, procedimentos e consolidação destes.

O problema de fontes de ruídos parasitas em medições de sinais de DPs é um dos principais obstáculos na caracterização destes sinais, assim como em outros sistemas de medição, monitoração e diagnóstico de equipamentos de alta tensão. Estes ruídos, geralmente, são tratados por ferramentas tradicionais e/ou convencionais como filtros e injeção de sinais invertidos [2]. Estas ferramentas, no entanto, são projetadas apenas para aplicações específicas de filtragem. E, no entanto, uma tarefa difícil fazer uma modelagem completa da remoção dos ruídos associados a um equipamento, utilizando unicamente filtros convencionais.

Devido à necessidade de melhorar a interpretação de medições de DPs, esta tese enfatiza o tratamento dos ruídos, na etapa de pós-processamento do sistema de medição, utilizando ferramentas emergentes e até algumas já consolidadas em vários campos, tal como inteligência artificial.

O ruído pode ser extraído dos sinais-fontes de forma aproximada. Utilizam-se sistemas artificiais e inteligentes (Redes Neurais, *Simulated Annealing*, Algoritmos Genéticos etc.), métodos ICA/BSS (Análise de Componentes Independentes e separação cega de sinal), PCA (Análise de Componentes Principais), estatísticas de alta ordem (HOSA) entre outros, com a finalidade de encontrar características que, muitas vezes, não são observáveis ou exploradas pelos métodos tradicionais.

É evidente que estes métodos podem ser complementares e cooperativos, a partir de uma análise de componentes independentes (ICA). Este método é uma das ferramentas em expansão por seu potencial de aplicabilidade em diversos campos de pesquisa, tais como processamento de sinais, imagens, telecomunicações, reconhecimento de voz, biomédica, separação cega de fontes etc. Em particular, investiga-se aqui a extração cega de ruídos em sinais de descarga parcial.

O ICA é um método de processamento de sinal para extrair fontes independentes a partir de misturas de fontes desconhecidas, adquiridas por medição (observações). O ICA busca direções coordenadas entre componentes mais independentes e não necessariamente ortogonais que representem as observações.

1.2 Objetivo

Investigar técnicas convencionais e não convencionais que ajudem a identificar informações relevantes dos sinais de descarga parcial. Explorar técnicas que se complementem, aproveitando suas vantagens individuais, com a finalidade de obter técnicas híbridas e robustas, caracterizadoras dos sinais de DPs. Caracterizar e

classificar fontes de DPs, removendo (parcial, mas suficientemente) ruídos das fontes dos sinais úteis utilizando a análise de componentes independentes. Reduzir convenientemente a dimensão dos dados para melhorar a eficácia e desempenho dos sistemas de reconhecimento de padrões de DPs. Desvendar possível estrutura oculta de fontes de DPs a partir das observações coletadas, separando fontes que possuem características diferentes e estatisticamente independentes entre si.

1.3 Contribuições

O trabalho de pesquisa realizado no desenvolvimento desta tese ensejou cinco publicações científicas (publicadas em diversos simpósios e congressos importantes) aplicadas à interpretação, discriminação e qualidade do sinal de DPs, através de sua caracterização para sistemas de detecção, monitoração assim como, em diagnóstico de equipamentos.

O trabalho busca novos métodos e modelos de compactação e redução da dimensionalidade aplicáveis aos sinais de DPs [3], visando a melhora do desempenho dos sistemas de reconhecimento de padrões e de diagnóstico de equipamentos [4].

Nesta pesquisa é mostrado que os métodos ICA podem, de forma aproximada, fornecer uma solução satisfatória para caracterização dos sinais de DPs [5], procedentes de fontes de defeitos em sistemas de isolamento, possíveis de serem encontrados nos equipamentos citados.

1.4 Escopo da Tese

O trabalho de pesquisa compreende os seguintes capítulos, os quais são introduzidos e brevemente delineados da seguinte forma:

Capítulo 1 - Introdução

Neste segmento faz-se uma descrição geral do tema da tese, destacando a importância da detecção e medição de DPs. Outro importante assunto é o problema de ruídos sob sinais de DPs sendo estes um dos principais obstáculos na caracterização destes sinais, assim como nos sistemas de medição, monitoração e diagnóstico de equipamentos em alta tensão. Além dos métodos tradicionais de processamento de sinais, existem novos métodos como a análise de componentes independentes, que será abordada nesta pesquisa, enfatizando a caracterização de fontes.

Capítulo 2 - Revisão Literária e Conceitos Básicos

Na primeira parte deste capítulo são resumidos os fatos literários mais importantes que conduziram os esforços de investigações no campo de DPs, desde os princípios dos

anos de 1920 até os dias atuais, destacando-se as novas tendências em sistemas de medição e detecção dos sinais DP. Na segunda parte, são sintetizados os fundamentos básicos utilizados no desenvolvimento da tese. Aborda-se o fenômeno de DPs, suas causas e conseqüências nos equipamentos estudados. A matemática e estatística básicas envolvidas nos processamentos dos sinais de DPs são apresentadas.

Capítulo 3 - Modelos Experimentais

Descrevem-se exclusivamente os métodos de medição e detecção de DPs aplicados na coleta das observações, utilizados no desenvolvimento da pesquisa segundo normas IEC-60270 [6] e IEEE P1434 [7], dos quais o método elétrico foi considerado, abordando especificamente o método da carga aparente. Também são descritos os modelos experimentais para representação de possíveis fontes de defeitos em transformadores de potência, denominados células testes e para alguns defeitos em enrolamentos estatóricos, baseados em barras envelhecidas.

Capítulo 4 - Pesquisas Anteriores e Classificadores Inteligentes em DPs

Neste capítulo, realiza-se um resumo das técnicas aplicáveis já desenvolvidas em pesquisas anteriores, das quais algumas serão utilizadas para avaliar o desempenho dos caracterizadores dos sinais de DPs.

Capítulo 5 - Caracterizadores de Sinais e Tratamento de Ruídos

Descrevem-se neste capítulo as principais técnicas e métodos aplicáveis e analisados nesta tese, enfatizando a análise de componentes principais, separação cega de fontes e a análise de componentes independentes. Além disso, se faz uma extensão dos métodos para remoção de ruídos.

Capítulo 6 - Caracterizadores: Implementação e Resultados

Neste capítulo se realiza a implementação e análise do desempenho dos caracterizadores dos sinais de DPs. Também uma possível integração destes em um sistema robusto caracterizador de sinais. Além disso, os caracterizadores podem ser aplicáveis para outros casos de medição, monitoração on-line e serem parte de um diagnóstico integrado de equipamentos de subestações de energia elétrica.

Capítulo 7 - Conclusões e Análise dos Resultados

A característica de ocorrência de forma incerta e imprecisa dos fenômenos de DPs torna o uso de métodos tais como o proposto nesta tese adequado e de extrema importância no que diz respeito a ser mais uma eficaz ferramenta para a informação em sistemas de diagnósticos de equipamentos de grande porte.

Seria interessante, como continuação deste trabalho, testar os modelos apresentados para outros equipamentos, bem como incrementar não só a quantidade de modelos com a finalidade de refinar a associação de modelos a fontes reais, mas também reunir um maior número de informações medidas no campo e no laboratório. Outro ponto importante seria realizar as correlações respectivas entre resultados laboratoriais e resultados encontrados no campo.

Capítulo 2

Revisão Literária e Conceitos Básicos

A seguinte seção apresenta um compêndio das pesquisas relacionadas aos fenômenos das descargas parciais (DPs). Esta revisão abrange investigações realizadas por diversas técnicas para detecção e medição dos sinais de DPs, assim como suas características e mecanismos relativos aos métodos elétricos.

Descargas corona ou parciais é um termo cotidianamente utilizado desde o passado. Sua base consolidou-se no começo dos anos de 1920, onde, já era considerado como um campo de pesquisa desenvolvido. Sua preeminente importância ferramental tem crescido para assistir na qualidade e no desempenho característico dos equipamentos de alta tensão. Dessa maneira, as descargas parciais têm se tornado responsáveis pelo crescimento das atividades de investigação relacionadas aos mecanismos físicos e químicos dos materiais isolantes nos equipamentos. Dessa forma, as diversas técnicas de detecção e medição foram surgindo e colocadas em prática. Anos mais tarde, o nível de esforços de investigação no campo de descarga parcial se diversificou substancialmente. Algumas destas tendências relacionam o tipo de equipamento de alta tensão com o tipo e comportamento de descarga.

Fazendo uma rápida revisão enumerativa na literatura, pode-se encontrar estudos relacionados à natureza e forma de descarga, sensibilidade de detecção, degradação da isolação exposta aos efeitos de DPs, quantidade de pulsos armazenados num determinado espaço de tempo, transferência de carga aparente, taxa de repetição de pulsos de descarga, perdas de energia, distribuição de pulsos de descarga, épocas da descarga relativa a uma fase da tensão referencial, intervalos de tempo de separação entre pulsos e o reconhecimento de padrões de pulsos em termos das fontes de descarga (fonte de defeito).

Talvez em nenhuma parte das questões expostas acima se encontre as diferentes tendências em estudos de DPs e procedimentos de teste. Desta forma, a obtenção do “desejado procedimento” de um teste, fácil de seguir, ainda não foi possível, seguindo a

evolução cronológica nos últimos cinquenta anos, além daqueles que, são aplicáveis a isolamentos sólidos em cabos de distribuição de energia [8]. Pois um dos primeiros estudos de DPs foi dado em isolamentos sólidos em cabos de distribuição de energia que, graças aos avanços dos estudos em linhas de transmissão e sua simples geometria, têm sido progressivos os resultados que facilitam a interpretação das medições de DPs. Já nos componentes de alta tensão, tal como em buchas e capacitores com baixa capacitância concentrada, há algumas dificuldades de interpretação de DPs, e em capacitores de alta capacitância as dificuldades aumentam devido a fatores complexos envolvidos na medição.

A detecção de DPs e a precisão na medição em transformadores se tornam apreciavelmente mais complicada. Apresenta um comportamento mais complexo que as linhas de transmissão, devido ao acoplamento e ao efeito de ressonância existente entre os enrolamentos. Dificuldades semelhantes, encontradas nos transformadores, são deparadas em máquinas giratórias (ex: hidrogeradores), principalmente na interpretação e calibração, onde, as grandezas dos pulsos detectados podem variar extremamente, estendendo-se desde os níveis mais baixos, geralmente intrínsecos a descargas internas dentro dos isolamentos das barras estatóricas, até os níveis supremamente altos, geralmente associados às descargas tipo em ranhuras conhecidas como *slot*.

Contudo, apesar dos esforços das entidades normativas e grupos de desenvolvimento relacionados às DPs em conseguir um procedimento de teste fácil a seguir, o setor ainda se carece de regras sólidas e claras. Como exemplo, a questão de calibração em medições de DPs ainda não foi resolvida e, sem dúvida, é controversa a decisão quanto a calibração ser um pré-requisito para medição em máquinas giratórias.

2.1 Revisão Literária e Inovações

Os componentes elétricos mais investigados por descargas parciais são os cabos de energia e equipamentos de alta tensão, onde uns dos primeiros métodos de detecção de DPs foi a análise por oscilógrafos, caracterizando-se para descargas tipo pulsante (Bartnikas [8]- estudos por Tykociner em 1933). Porém, podem ocorrer descargas não pulsantes que emitem calor ou luz e descargas pseudoluminescentes. A descarga pulsante pode ser facilmente detectada pela maioria de detectores convencionais de pulsos de DPs, evidenciando que são indicadores efetivos da presença de DPs. Porém, nem sempre esses detectores de pulsos podem indicar o nível total de intensidade presente nos pulsos de DPs. Apesar de existirem detectores do tipo ponte para DPs, os

quais têm correspondido às descargas pulsantes e não pulsantes (Bartnikas: 1979, IEC-270: 1981), mas sua baixa sensibilidade intrínseca impede sua implementação em grande escala na área de DPs.

O limite do tempo de subida das descargas tipo pulsante, pode ser estabelecido teoricamente. Historicamente as medições rápidas do tempo de subida (medido por osciloscópio) precisam ser armazenadas de maneira experimental, induzindo a uma relação que diminui monotonicamente com a capacidade e a largura de banda dos osciloscópios utilizados. Osciloscópios com largura de banda até os GigaHertz (GHz) podem ser encontrados na atualidade, com estes podem-se medir tempos de subida menores que 1 nano segundo. Dessa forma, pode-se detectar pulsos de DPs a frequências de até 1 GHz, devendo-se lembrar que o conteúdo de energia dos pulsos das DPs é uma função decrescente da frequência.

Detectores convencionais, encontrados comercialmente [9] para usos rotineiros em cabos, capacitores e transformadores são do tipo de banda estreita e são projetados para operar na banda aproximada de 30 até 400 kHz. Estes são dispositivos integradores de carga e podem ser calibrados diretamente para fornecer a transferência de carga associada com pulsos de descarga detectados de acordo com as normas ASTM D1868 [8]. Detectores com larguras superiores, ou banda larga, são utilizados em pesquisas, nos quais uma reprodução exata (confiável) da forma de pulsos de DPs é de suprema importância. Além disso, para melhorar a resolução de pulso, larguras de banda estendidas são utilizadas para tarefas de localização no campo envolvendo descargas em cabos (~20 MHz), máquinas girantes (800 kHz a 1 GHz) e barramentos em gás comprimido (~1 GHz).

Os primeiros sistemas de detecção para DPs empregavam instrumentação analógica, apresentando um desempenho adequado tanto para o início de descargas (*inception*) quanto para as medições da tensão de extinção. O padrão de pulso da DP era mostrado por meio de oscilógrafos na base do tempo-frequência e calibrado numa escala arbitrária. A transferência de carga geralmente associada com os pulsos de carga discreta podia ser estimada visualmente e a fase aproximada relacionando os pulsos e a tensão aplicada era conhecida pelo observador que realizava os testes.

O aparecimento da contagem de pulsos de DP por cristal controlado nos anos de 1950, permitiu a contagem de pulsos de DPs por unidade de tempo e, através disso, a determinação da densidade de pulsos das DPs de padrões de descarga (Bartnikas 1963). Subseqüentemente, apareceram analisadores diferenciais de amplitude dos pulsos e

analísadores de amplitude de pulsos de canal simples [8]. No começo dos anos de 1960s, os conversores *led* analógico/digital introduziram comercialmente a baixo custo os analisadores multicanais confiáveis para análise de distribuição da amplitude dos pulsos de DPs (Bartnikas 1969). A área sob o intervalo do pulso e sua distribuição em fase foi desenvolvida nos anos de 1970s, sendo estendidos na prática para aplicações em máquinas girantes [8]. Estes desenvolvimentos foram seguidos de perto pela introdução de técnicas computacionais para medição de pulsos de DP [10].

Os avanços dos computadores PC, na década de 1980, e seu extensivo uso nos anos de 1990, alteraram rapidamente a aproximação na área de análise da distribuição de pulsos de DPs [11]. Os sistemas de medição foram transferidos da instrumentação baseada em hardware para técnicas via software [8], [12], [13]. Estes adventos conduziram para investigações em reconhecimento de padrões e classificadores, envolvendo redes neurais e lógica fuzzy de DPs [14], [13], [15]. Os primeiros estudos indicavam que a amplitude de um pulso de descarga e sua fase de ocorrência são influenciadas fortemente pela ocorrência de pulso ou pulsos precedentes, efeito de memória ou reverberação. Dito efeito, considerado como processo não Marcoviano, foi rigorosamente analisado usando aproximação estocástica por Van Brunt [16], [17], [18], cuja finalidade foi esclarecer a natureza estatística condicional dos mecanismos das descargas. Os resultados obtidos colocam alguns questões sérias quanto à efetividade de classificação e reconhecimento de padrões de DPs, bem como, algum tratamento estatístico de tais dados que torne mais tratável a sua interpretação. Apesar disso, deve ser observado que o tratamento estatístico de padrões de DP, amplitude/fase [19] ou forma de pulso (*shape*) [47], [20], tem rendido alguns resultados práticos e interessantes.

Ainda na década dos anos de 1990, viu-se a introdução dos circuitos digitais de resposta rápida para aplicações de medição de DPs [14], [13], [21]. Enquanto a utilização das técnicas digitais em detecção de pulsos de DPs, medição e aquisição têm crescido notavelmente, detectores de DPs comerciais têm mantido separadas opções de medição analógica e digital. A propósito, o pico do pulso de descarga geralmente é determinado pelo sistema digital, e não será o mesmo se for determinada a verdadeira magnitude em tempo real, pelo circuito analógico. Por um lado, há forte dependência da taxa de amostragem, da largura de banda e da capacidade de armazenamento dos circuitos digitais [92]. Por outro lado, os componentes analógicos são alvos do efeito da

temperatura e saturação que atenuam a verdadeira magnitude de um pulso de DP. Além disso, há o ruído e outros fenômenos intrínsecos [92], [93], [94], [95].

Dessa forma, o circuito analógico normalmente precede ao sistema de aquisição digital como amplificador e capturador da verdadeira forma do pulso [22]. Assim a maioria dos circuitos sensores de DPs tem uma configuração híbrida analógico-digital [23]. Na última década, a variedade dos circuitos digitais avaliáveis e em uso pelas medições de DPs tem incitado publicações de artigos em revistas e eventos técnicos [].

No final da década de 1990, apareceram outros estudos inovadores, tal como os sistemas de medição com analisador de resolução estendida [19], [23], com histogramas sofisticados envolvendo duas ou três grandezas relativas às descargas em tempo real. O detector de amplitude de pulsos no domínio da fase útil foi introduzido em osciloscópios digitais rápidos para detecção DPs. E também um sistema de medição baseado em dois detectores paralelos a fim de obter uma distribuição sincronizada para intervalos de tempo entre pulsos de DPs e sua posição em fase [23].

Depois da virada do milênio, outros estudos têm sido colocados em prática, tal como a importância da largura de banda para analisadores de espectros e filtros, estudos de acopladores capacitivos maiores que 80 pC e estudos híbridos de resposta em frequência do material dielétrico com faixa de frequência do sinal de DP. Contudo, os estudos no domínio da frequência [96] parecem ser intermináveis, visto que os sinais de DPs dependem também da frequência e têm uma estreita ligação com o circuito de ensaio [24].

Além da tradicional detecção de DPs *off-line*, a monitoração de DP *on-line* tem amadurecido e está sendo cada vez mais adotada pelas concessionárias. Entretanto, os sensores de acoplamentos ainda têm dificuldades a acessos restritos ao equipamento e se instalados têm uma faixa de exploração muito curta. Além disso, os ruídos são mais intensos quando o equipamento trabalha com carga, sendo que, este é também influenciado por interferências dos equipamentos vizinhos. Para contornar estes inconvenientes, estudos diversos em filtros com tecnologias recentes e técnicas avançadas em processamento de sinais tanto com hardware e software têm surgido cada vez mais. Entre estas técnicas tem-se a separação cega de sinais e a eliminação de ruídos em fontes adquiridas por análise de componentes independentes, que este trabalho de tese apresenta.

2.2 Revisão Teórica

Nas primeiras experiências praticadas, a detecção de DPs (aproximadamente há quatro décadas) era feita através do teste de tensão de rádio interferência (TRI). Já na década de noventa as técnicas de medição foram melhoradas pelas novas tecnologias no processamento de sinais. Começaram a surgir instrumentos mais eficientes, porém estes ainda são genéricos e os mais especializados costumam ser custosos. Além disso, o ruído ainda é pouco tratável e de difícil eliminação. Mesmo assim, a detecção de DPs vem sendo cada vez mais utilizadas em motores, geradores, transformadores, pára-raios, interruptores e até em bancos de capacitores.

Para que se possa melhor entender o complexo fenômeno das DPs, é importante conhecer alguns princípios físicos, químicos e elétricos envolvidos. Neste capítulo são abordados sucintamente os mecanismos da ionização dos átomos de um material isolante, conceitos importantes sobre campo elétrico, características, causas e consequências das DPs. Além de entendê-los, há necessidade de estudá-los para descobrir outras características baseadas em alguns de seus princípios e nas suas observações coletadas.

2.3 Mecanismos das Descargas Parciais

2.3.1 Ionização dos átomos de um material isolante

Para compreender o fenômeno de descargas parciais é preciso entender como são ionizados os átomos de um material dielétrico [1], [25]. O processo de ionização e avalanche (início de descargas) de cargas satisfaz as relações de Paschen e Townsend [25]. Porém, em ensaios práticos de DPs, as “impurezas” e “imperfeições internas” no dielétrico, distorcem as equações de Townsend, fazendo com que a tensão de *inception* possa ocorrer em tensões menores do que a tensão nominal.

Para casos práticos e de simples ionização, deve-se estabelecer a relação entre a diferença de potencial entre dois pontos a e b (V_{ab}) e o campo elétrico (\vec{E}). Essa relação está implícita na Equação (2.1), onde se considera um dielétrico com comportamento linear, isotrópico e sistema homogêneo (a permissividade ϵ do meio é estática ou é considerada constante para baixas frequências).

$$\nabla V = \nabla \phi = -\vec{E} \quad (2.1)$$

Admitindo que o campo seja uniforme, as linhas de campo são paralelas e a diferença de potencial por unidade de comprimento é constante. Logo, as linhas equipotenciais, que são ortogonais às linhas de campo, estarão espaçadas

uniformemente e quanto maior a diferença de potencial entre dois pontos, mais intenso é o campo. Uma formulação mais completa da Equação (2.1) incluindo o efeito do potencial vetor de campo pode ser encontrada em [26]. A Equação (2.2) representa sistemas não-homogêneos, onde o potencial escalar $\nabla\phi$ corresponde a cargas estáticas ou fenômenos lentos, $\frac{\partial\mathbf{A}}{\partial t}$ é o componente para cargas dinâmicas ou fenômenos rápidos e \mathbf{A}^* é o potencial vetor de uma fonte totalmente externa à região considerada (assume tipicamente valor nulo). Assim, os potenciais das cargas são representados por ϕ (escalar) e \mathbf{A} (vetorial) [26]. Além disso, ε tem dependência da pressão, temperatura e frequência [11].

$$\vec{\mathbf{E}} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} - \frac{1}{\varepsilon}\nabla\times\mathbf{A}^* \quad (2.2)$$

Em relação ao caso simples e prático, quando um campo elétrico é aplicado a um condutor, os elétrons da camada externa de seus átomos são facilmente “arrancados” e migram prontamente de átomo para átomo. Porém, os elétrons de um material dielétrico estão bem “presos” perto de suas posições de equilíbrio e não podem ser “arrancados” facilmente. Quando um campo elétrico é aplicado a um material dielétrico, este é polarizado, ou seja, há um deslocamento de cargas positivas e negativas da sua posição de equilíbrio, embora não haja migração de carga neste caso. Em geral, este fato faz com que os materiais dielétricos (resinas, vidro, mica etc.) se comportem como bons isolantes.

Os átomos individuais na presença de um campo elétrico têm o elétron negativamente carregado, sujeito a uma força no sentido contrário ao campo $\vec{\mathbf{F}} = q\vec{\mathbf{E}}$ e o núcleo positivamente carregado, sujeito a uma força no mesmo sentido do campo. Essas forças distorcem a estrutura do átomo de maneira que o núcleo fica à direita do seu centro efetivo. Os átomos se comportam então como dipolos, ou seja, podem ser considerados como duas cargas pontuais $+Q$ e $-Q$, representando respectivamente o núcleo e a carga eletrônica. Diz-se, então, que o material dielétrico está polarizado.

A Figura 2.1 ilustra um átomo de um material dielétrico e o deslocamento do elétron com relação ao núcleo quando na presença de um campo elétrico.

Como foi mostrado pela Equação (2.1), se a diferença de tensão entre os eletrodos é aumentada, o campo elétrico irá aumentar até um ponto onde as forças externas exercidas no elétron serão maiores que as forças internas e, o elétron será então

“arrancado” da sua órbita. O átomo estará então ionizado, ou seja, se tornará um íon positivo.

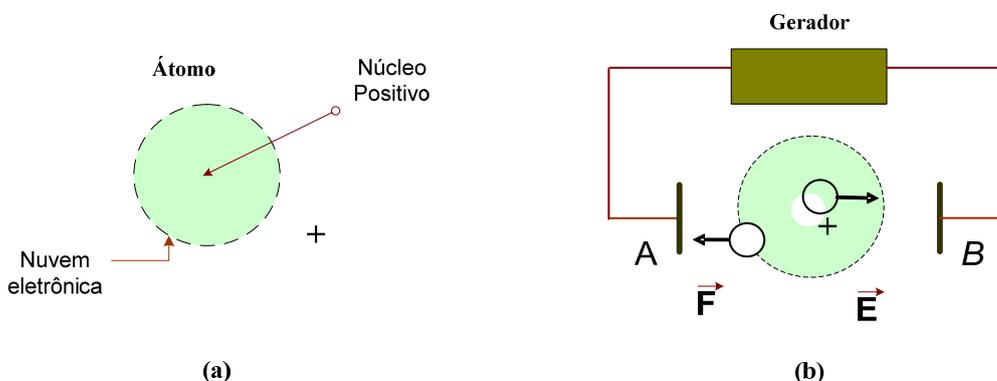


Figura 2.1 – Átomo de um material dielétrico a) Átomo não polarizado. b) Forças no átomo quando se aplica \vec{E} [1].

O processo de ionização principal, no entanto, é a ionização por colisão entre um elétron e um átomo ou molécula neutra. Se o elétron livre está submetido a um campo elétrico, ele é acelerado e colide com átomos de nitrogênio, oxigênio e outros gases presentes. A velocidade do elétron está diretamente ligada à intensidade do campo elétrico. Se o campo elétrico não for muito intenso, estas colisões serão elásticas, similares a colisões entre bolas de bilhar, e não haverá transferência de energia. Se por outro lado a intensidade do campo elétrico exceder um valor crítico, qualquer elétron livre presente nesse campo irá adquirir uma velocidade suficiente para tornar a colisão com uma molécula de ar inelástica, o que significa que o elétron terá energia para deslocar outro elétron de sua órbita e ionizar o átomo.

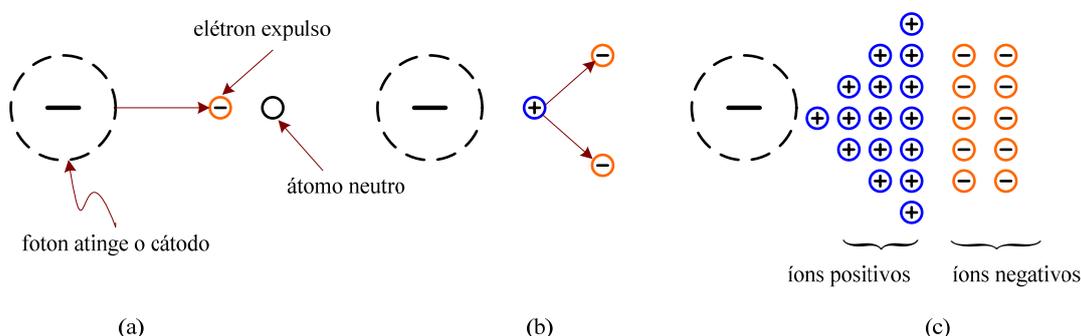


Figura 2.2 – Processo de avalanche de elétrons iniciado a partir de um eletrodo negativo a) Início. b) Formação de um par de elétrons – Impacto de um elétron com um átomo neutro libera um elétron adicional e deixa um íon positivo para trás. c) Multiplicação – Os elétrons se movem criando íons positivos quando se multiplicam.

O elétron inicial – que perdeu a maior parte da sua velocidade na colisão – e o elétron expelido da molécula de ar – que também tem uma velocidade baixa – é acelerado pelo campo elétrico, e na próxima colisão cada elétron é capaz de ionizar uma molécula de ar. Após a segunda colisão, existem quatro elétrons capazes de ionizar outros átomos e assim sucessivamente, com o número de elétrons dobrando a cada colisão. Este processo é conhecido como avalanche de elétrons, sendo sempre iniciado com um elétron livre submetido a um campo elétrico intenso. A Figura 2.2 ilustra a seqüência da ionização dos átomos.

Os íons positivos deixados para trás no processo da avalanche de elétrons se movem na direção do eletrodo negativo. No entanto, eles se movem muito vagorosamente, devido à sua massa, que é aproximadamente cinqüenta mil vezes a massa do elétron [1], [25]. Tendo uma carga positiva, esses íons atraem elétrons que estão vagando e, quando algum elétron livre consegue ser capturado, forma-se outra molécula neutra de ar. O nível de energia de uma molécula neutra é menor do que o do íon positivo correspondente, logo, quando um elétron livre é capturado, um quantum de energia é emitido pela molécula. Esse quantum de energia é exatamente igual em magnitude à energia que inicialmente foi requerida para deslocar o elétron inicial fora de sua molécula. Uma onda eletromagnética é irradiada e, para moléculas de ar como oxigênio ou nitrogênio, essa radiação está dentro da faixa visível de luz. Portanto, um observador pode ver essa radiação como uma luz violeta clara, que ocorre principalmente na recombinação de íons de nitrogênio com elétrons livres. Se não houver mais a fonte de energia, o processo de ionização cessará; mas a recombinação continuará até que não haja mais elétrons livres nem íons positivos.

Os elétrons e íons positivos provenientes deste processo são suficientes para conduzir corrente entre os eletrodos e absorver uma quantidade relativamente grande de potência da fonte, fazendo com que uma centelha surja entre os eletrodos. Quando ocorre o centelhamento, diz-se que há uma ruptura do material dielétrico. Como essa centelha normalmente não atravessa completamente o material dielétrico entre os eletrodos, é chamada de descarga parcial.

A avalanche de elétrons descrita na Figura 2.2 possui certo número de elétrons por segundo, que pode variar de centenas de elétrons por segundo até 10^{22} elétrons por segundo em um período típico de 100 nano segundos. Para quantificar melhor a carga, utiliza-se como unidade o Coulomb, que é equivalente a uma carga de $6,2 \times 10^{18}$ elétrons. Como um Ampère é definido como um fluxo de carga de um Coulomb por

segundo, a corrente da avalanche eletrônica pode variar de 10^{-17} A até alguns milhares de Ampères.

A energia da descarga, no entanto, é extremamente pequena quando se tenta medir a amplitude da tensão de um pulso de descarga.

Se a descarga ocorrer no ar, em torno de um elemento condutor, é denominada de efeito corona, assim como “*streamer*” ou “descarga autosustentada”.

O movimento de elétrons para o eletrodo positivo (A na Figura 2.1.b) e o movimento de íons positivos para o eletrodo negativo (B na Figura 2.1.b) significa um fluxo de elétrons através do gerador entre os dois terminais.

Se uma impedância for introduzida entre o gerador e os terminais, a diferença de potencial não será mais constante e irá decrescer linearmente com o aumento da corrente. Ajustando o valor dessa impedância, pode-se variar a taxa de dissipação de energia na descarga ou a potência consumida.

2.3.2 Algumas considerações sobre o Campo Elétrico

Como foi visto anteriormente, o campo elétrico tem um papel fundamental no surgimento das descargas parciais, pois influencia diretamente o processo de ionização dos elétrons de um material dielétrico. Podem ocorrer descargas nos pontos do material dielétrico onde houver maior intensidade de campo elétrico, o que torna muito importante um estudo mais detalhado sobre as circunstâncias em que isso ocorre.

A intensidade do campo elétrico está relacionada basicamente com o valor de tensão, com o meio em que ele se encontra e com a geometria do conjunto eletrodos/dielétrico em que ele atua. Como descrito na Equação (2.1), o campo elétrico é diretamente proporcional à tensão aplicada, ou seja, aumenta para valores maiores de tensão. As outras duas situações serão vistas a seguir.

2.3.3 O campo elétrico e o meio em que ele se encontra

Para compreender como o campo varia devido ao meio em que ele se encontra, é preciso analisar as relações de fronteira entre dois meios dielétricos diferentes, pois o campo elétrico pode variar abruptamente, tanto em intensidade como em direção. Essa análise é feita em duas partes, considerando, em primeiro lugar, a relação entre os campos tangentes à fronteira, e em segundo lugar, os campos normais a ela.

Estas relações de fronteira levam em consideração uma característica muito importante nos materiais dielétricos: a permissividade ou constante dielétrica (ϵ). Como a permissividade de um dielétrico é sempre maior do que a permissividade do vácuo,

2.3 Mecanismos das Descargas Parciais

muitas vezes é conveniente usar a permissividade relativa do dielétrico, isto é, a razão entre sua permissividade e a do vácuo. Assim,

$$\varepsilon_r = \frac{\varepsilon}{\varepsilon_0} \quad (2.3)$$

onde, ε_r é a permissividade relativa do dielétrico, ε é a permissividade do dielétrico e ε_0 é permissividade do vácuo = $8,85 \text{ pF m}^{-1}$.

Enquanto que ε e ε_0 são expressos em farads por metro (F.m^{-1}), a permissividade relativa ε_r é uma razão adimensional e é o valor dado geralmente em tabelas. A permissividade relativa de alguns meios está apresentada na Tabela 2.1, onde os valores referem-se a campos estáticos (ou em baixas frequências) e são aproximados, exceto para o vácuo ou ar. Observe que ε_r para o ar está tão próximo da unidade que, na maioria dos casos, podemos considerar o ar equivalente ao vácuo [27].

Tabela 2.1 – Permissividade de meios dielétricos

Meio	Permissividade relativa ε_r
Vácuo	1
Ar (pressão atmosférica)	1,0006
Espuma de poliestireno (Styrofoam)	1,03
Parafina	2,1
Madeira compensada	2,1
Poliestireno	2,7
Âmbar	3,0
Borracha	3
Plexiglas	3,4
Solo arenoso seco	3,4
Nylon (sólido)	3,8
Enxofre	4
Quartzo	5
Baquelita	5
Fórmica	6
Vidro com maior composição de chumbo (Lead Glass)	6
Mica	6
Mármore	8
Cristal (Flint glass)	10
Amônia (líquida)	22
Glicerina	50
Água (destilada)	81
Titanato de bário (BaTiO_3)	1200
Titanato de bário e estrôncio ($2\text{BaTiO}_3: 1\text{SrTiO}_3$)	10000
Titanato e zirconato de bário ($4\text{BaTiO}_3: 1\text{BaZrO}_3$)	13000
Estanho e titanato de bário ($9\text{BaTiO}_3: 1\text{BaSnO}_3$)	20000

A permissividade relativa é um indicador do nível de isolamento elétrico de um material dielétrico. Quanto maior a permissividade relativa do material, melhor a

suportabilidade de tensão desse material, considerando um campo elétrico uniforme e uma temperatura de 20 °C. Para faixas de frequência muito amplas, os materiais podem apresentar variações da permissividade ϵ caracterizadas por componente real ϵ' e imaginária ϵ'' , os quais dependem da frequência e podem apresentar diversos modos ressonantes causados pelos diversos modos de vibração [26], [27].

Considerando dois meios dielétricos de permissividade ϵ_1 e ϵ_2 separados por uma fronteira plana, e supondo que os dois meios são isolantes perfeitos, tem-se que as componentes tangenciais do campo elétrico são iguais nos dois lados da fronteira, como mostram a Figura 2.3. Isso significa que o campo elétrico tangencial é contínuo através dessa fronteira.

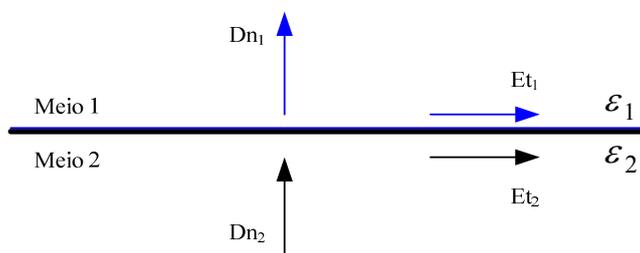


Figura 2.3 – Fronteira entre dois meios dielétricos diferentes.

Para tratar das componentes normais, será utilizada a densidade de fluxo, cuja componente normal é contínua através da fronteira desprovida de cargas entre dois dielétricos, como pode ser visto nas equações (2.4) e (2.5).

$$\mathbf{D}n_1 = \mathbf{D}n_2 \quad (2.4)$$

$$\epsilon_1 \mathbf{E}_{n_1} = \epsilon_2 \mathbf{E}_{n_2} \quad (2.5)$$

De acordo com a Equação (2.5), verificamos que os campos elétricos normais à fronteira estão inversamente relacionados com as permissividades relativas ϵ_1 e ϵ_2 . Se o meio 1 tiver uma permissividade relativa menor que a do meio 2, o campo elétrico no meio 1 será maior que no meio 2.

Esse tipo de situação ocorre, por exemplo, com pequenas cavidades de ar no interior de materiais dielétricos sólidos ou bolhas de gás no interior de materiais dielétricos líquidos. Como o ar tem uma permissividade-relativa menor que a do material dielétrico (sólido ou líquido), o campo elétrico no ar é maior, o que favorece a ionização dos elétrons e conseqüentemente a ocorrência de descargas parciais. A produção de descargas também se dá no ar pela ionização das moléculas que se encontram nos pontos de maior gradiente de potencial.

A intensidade máxima de campo que um dielétrico pode suportar sem que ocorram centelhamentos é a rigidez dielétrica. A Tabela 2.2 relaciona os valores de rigidez dielétrica de vários materiais dielétricos. A rigidez dielétrica refere-se a um campo uniforme e os materiais estão listados em ordem crescente de rigidez [27].

Tabela 2.2 – Rigidez dielétrica de diversos materiais dielétricos comuns

Material	Rigidez Dielétrica (MV m ⁻¹)
Ar (pressão atmosférica)	3
Óleo (mineral)	15
Papel (impregnado)	15
Poliestireno	20
Borracha (dura)	21
Baquelita	25
Vidro (placa)	30
Parafina	30
Quartzo (fundido)	30
Mica	200

2.3.4 O campo elétrico e a geometria dos eletrodos

A geometria dos eletrodos influencia diretamente a distribuição da intensidade do campo elétrico. O gradiente de tensão aumenta quanto menor for a área do eletrodo. Se o formato do eletrodo for uma ponta, o gradiente de tensão é alto. Se, para o mesmo nível de tensão aplicada, o eletrodo tiver um formato plano, o gradiente de tensão é bem menor. A Figura 2.4 ilustra um exemplo de eletrodo ponta-plano tendo um gás como dielétrico.

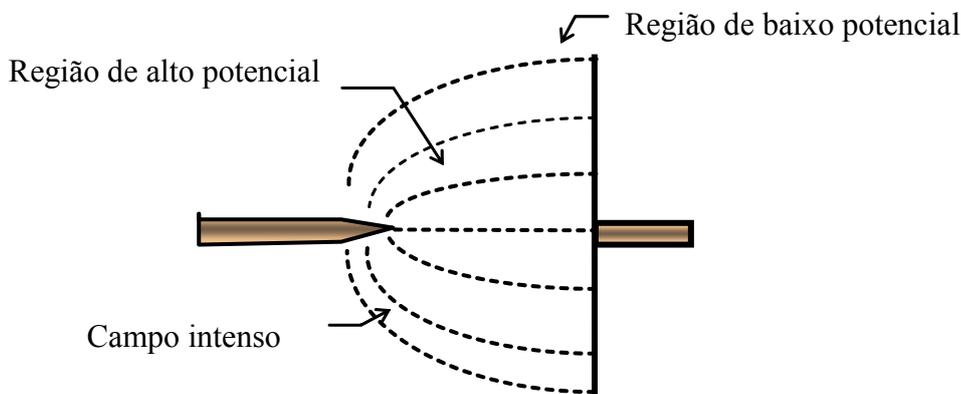


Figura 2.4 – Eletrodo ponta-plano. [1]

Perto da ponta, como o gradiente de potencial é maior, a intensidade do campo elétrico é maior, como foi mostrado pela Equação (2.1). O elétron que está perto da ponta pode então ser acelerado até adquirir energia cinética suficiente para causar uma ionização. Perto do plano, a intensidade do campo elétrico é menor e o elétron pode não adquirir energia cinética suficiente para causar uma ionização. Logo, a descarga é

limitada a um volume perto da ponta. O volume de gás restante entre o plano e a descarga funciona como uma excelente resistência, limitando o centelhamento.

A estrutura de eletrodos cilíndricos, onde o cilindro interno é bem menor que o cilindro externo, funciona exatamente da mesma maneira, sendo a descarga adjacente ao cilindro interno, que tem o maior gradiente de potencial.

Um outro exemplo de eletrodos seria o de duas agulhas (separadas de aproximadamente 0,5 cm) e um plano. Se a tensão é ajustada para fornecer uma descarga pequena perto da ponta, as duas agulhas irão centelhar. Se, no entanto, a tensão é aumentada para fornecer descargas maiores, somente uma agulha irá centelhar. Para fazer com que as duas agulhas sofram as descargas, é necessário adicionar uma resistência série em cada agulha, como mostra a Figura 2.5 . Quando as descargas são pequenas, as impedâncias presentes no volume de gás não ionizado entre as pontas e o plano atuam, limitando a corrente.

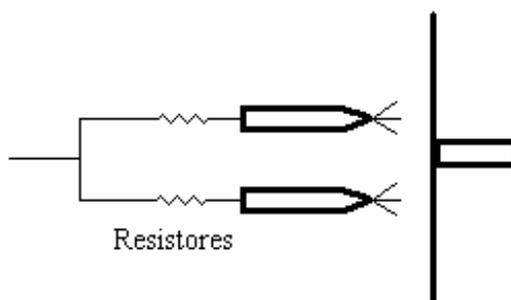


Figura 2.5 – Eletrodos: duas pontas paralelas com resistores e um eletrodo plano [1].

Se o campo for uniforme (linhas de campo paralelas em toda parte) e se for aumentado gradualmente, ocorrerão centelhas no ar quase que imediatamente após um valor crítico do campo ter sido ultrapassado; mas se o campo não for uniforme (linhas de campo divergentes) poderá ocorrer inicialmente uma descarga, e depois, com o aumento de \vec{E} , uma descarga ininterrupta.

2.4 Tipos das Descargas Parciais

Descargas parciais podem ocorrer em qualquer ponto do material dielétrico; na junção de dois materiais dielétricos diferentes ou adjacentes ao eletrodo, seguidamente em vários pontos do dielétrico e até mesmo sem eletrodos, em uma cavidade do material dielétrico.

A idéia errônea de que eletrodos metálicos são necessários decorre da sua larga utilização em demonstrações de laboratório, seja através de eletrodos ponta-plano ou um

2.4 Tipos das Descargas Parciais

condutor e um plano. Descargas parciais podem ocorrer entre dois isolantes agindo como eletrodos; essas descargas sem eletrodos são problemáticas em alguns veículos espaciais.

O termo descargas parciais envolve um amplo grupo de fenômenos de descarga tal como descargas internas, superficiais e descargas corona [19]. Descargas internas podem ocorrer em cavidades em dielétricos sólidos ou bolhas em dielétricos líquidos, descargas superficiais podem ocorrer na interface de isolamentos ou condutores e descargas corona são relativos a descargas em dielétricos gasosos se estão presentes elevados campos elétricos não-homogêneos. Por último o impacto das descargas em dielétricos sólidos podem formar-se canais de descarga chamados *treeing*. A Figura 2.6 mostra algumas descargas típicas baseadas em arranjos isolantes simples.

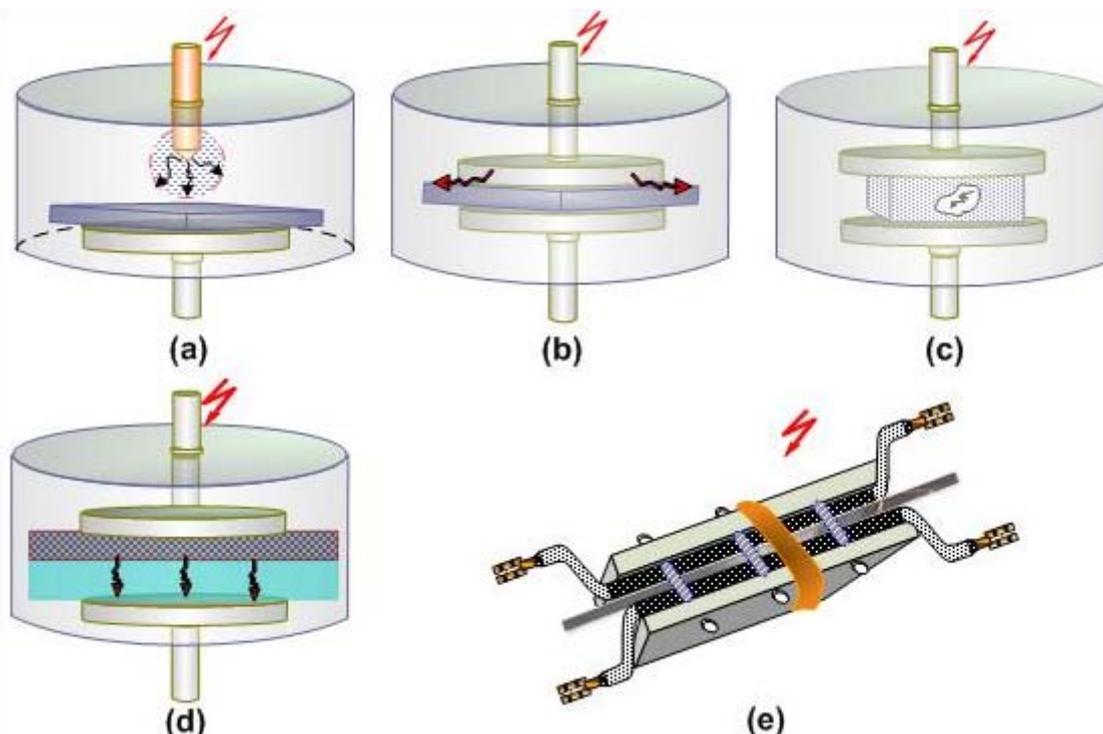


Figura 2.6 – Tipos de DPs em arranjos básicos a) Descargas tipo corona. b) Descarga na superfície. c) Descargas em material laminado. d) Descargas internas. e) Descargas entre espiras em enrolamentos de máquinas elétricas girantes (turn-to-turn)

Descargas tipo efeito corona (Figura 2.6 .a) ocorrem em eletrodos tipo ponta ou condutores finos conectados a um alto potencial ou à terra, particularmente em instalações isoladas com ar/gás ou com materiais isolantes líquidos. Descargas na superfície (Figura 2.6 .b) e descargas em materiais laminados (Figura 2.6 .c) ocorrem nas interfaces entre materiais isolantes diferentes, como por exemplo, entre gás/sólido. Descargas em pequenas cavidades de ar (Figura 2.6 .d) ocorrem em líquidos tanto quanto

em materiais isolantes sólidos. Descargas tipo *Slot* ocorrem em máquinas girantes: os modelos se baseiam em barras de geradores, nas quais possuem ranhuras para reprodução do defeito (**Figura 2.6 .e**) [29]. Em materiais isolantes sólidos podem ocorrer também as descargas contínuas que deterioram parcialmente o material isolante geralmente denominado descargas *treeing* ou *treeing channels* (ver Figura 2.7).

Os *treeing channels* se propagam em volta do isolamento sólido similar aos encontrados em barras de gerador. Estudos experimentais advertem que este processo pode ser reduzido por pesquisa de parâmetros elétricos, térmicos e carregamento mecânico [30], [31]. Porém, muitos destes estudos se tornam inviáveis quando aplicado a equipamento real, devido a restrições de acesso e limitações de implementação. Assim estes estudos têm sido realizados em ambientes e laboratórios especiais com a finalidade de determinar o comportamento do material sob condições adversas aplicadas.

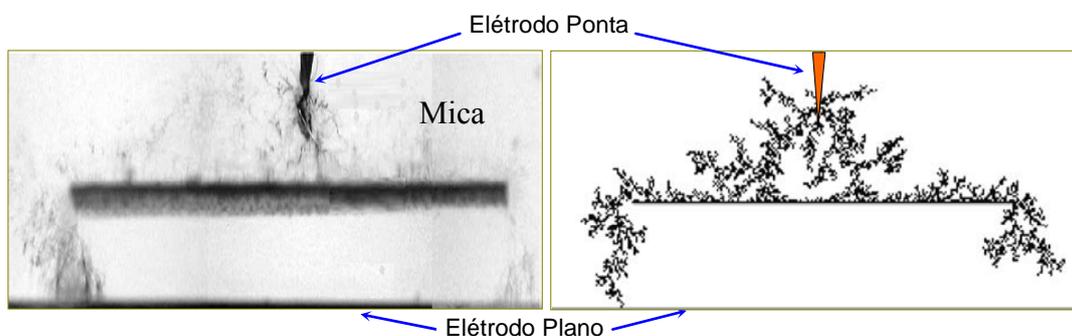


Figura 2.7 – Descargas em materiais isolantes sólidos (*treeing channels*): Trajetória de propagação de uma *treeing channels* em volta de uma barreira de base de mica.

Em aplicações práticas, é feita uma distinção entre “DP interna” e “DP externa”. Descargas externas são definidas como processos que ocorrem fora do equipamento, por exemplo, em armaduras; e descargas internas são definidas como processos que ocorrem dentro de sistemas fechados. Essa distinção é apropriada para os sistemas de medição utilizados, pois descargas externas (em linhas de alimentação ou eletrodos) podem ser distintas de descargas internas ocorrendo no objeto a ser testado.

Descargas parciais são de curta duração em relação ao período da sinusóide de tensão aplicada (da ordem de nano segundos), são repetitivas, têm sua frente muito íngreme e muito acentuada, podendo ser consideradas como uma função impulso.

2.5 Conceito de Descargas Parciais

As DPs são sucessões de descargas elétricas incompletas, rápidas e intermitentes da ordem de nano segundos (ver Figura 2.8.a), que ocorrem pela proximidade entre duas partes condutoras de eletricidade e um meio isolante, pelo efeito de ionização em

2.5 Conceito de Descargas Parciais

cavidades gasosas no interior dos materiais isolantes sólidos e líquidos. Tais descargas também podem ocorrer na zona de contato de isolantes sólidos (dielétricos) e superfícies metálicas submetidas à tensão elétrica [32]. Esta definição coloca em evidência algumas características notáveis do fenômeno físico das DPs, como:

- Realizam-se rapidamente, considerando a sua curta duração em relação ao semi-período da tensão referencial aplicada (60 Hz).
- São fenômenos intermitentes, determinando sobre os circuitos elétricos impulsos como fontes muito íngremes e discretas no tempo.
- Caracterizam-se por elevadas frequências espectrais (da ordem de megahertz).
- O ambiente onde se desenvolvem as disrupções é de natureza gasosa. O espaço onde se realizam as DPs restringe-se a uma zona limitada do dielétrico, não chegando a curto-circuitar os dois eletrodos, pois quando tal se verifica ocorre a disrupção do dielétrico e, portanto a sua deterioração.

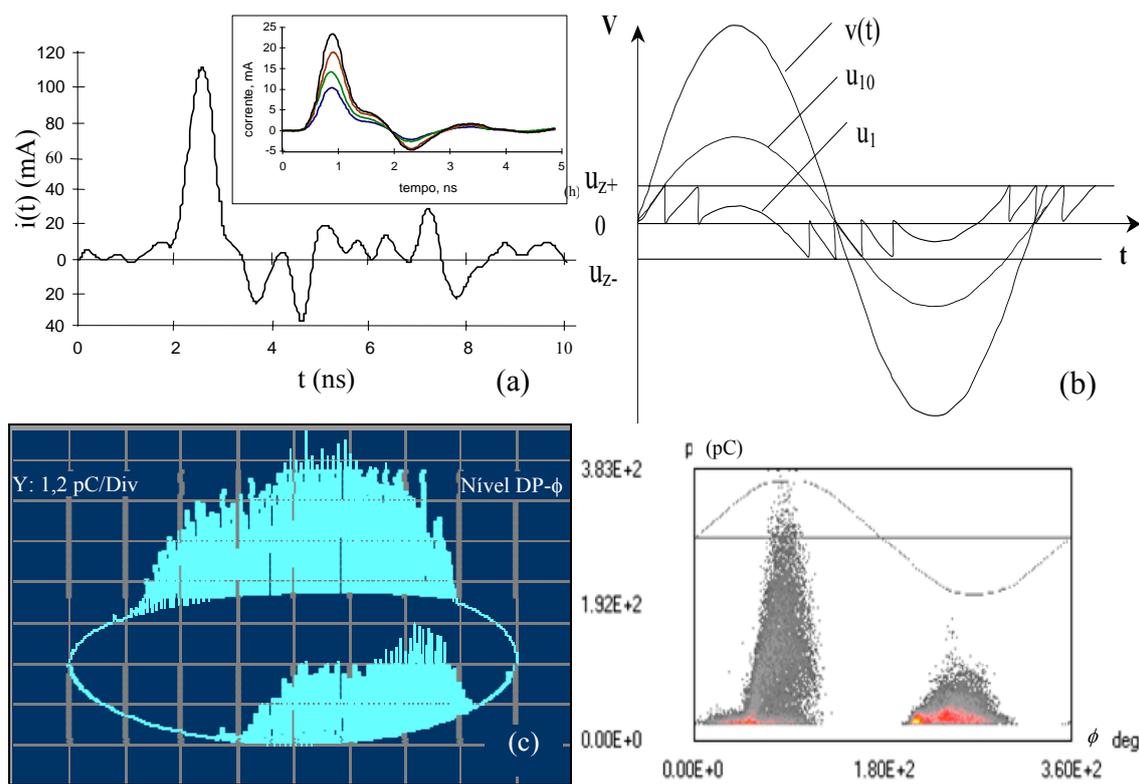


Figura 2.8 – Sinais de DPs: (a) Pulsos intermitentes medidos em ultra-alta frequência, (b) Queda de potencial por DPs em cavidades, (c) e (d) mapeamento típicos de DPs gerados por modelos experimentais.

A Figura 2.8.b) mostra um modelo típico de medição e simulação de DPs em cavidades gasosas e dielétricos sólidos, onde $V(t)$, U_{10} e U_1 são os valores máximos das tensões de referência, da cavidade e de descarga, respectivamente e, U_{z+} e U_{z-} são níveis de tensão de início de descarga em cavidades gasosas, maiores detalhes da Figura

2.8b encontra-se em [1], [19]. As Figura 2.8.c e d são mapas digitais dos pulsos máximos de DPs, obtidos por instrumentos de medição de DPs sob um modelo experimental.

Como foi exposto acima, a ocorrência de DPs é favorecida em cavidades gasosas no interior de materiais isolantes. Essas cavidades são as responsáveis pelas falhas mais comuns no material dielétrico, e são geralmente introduzidas durante as várias etapas de fabricação dos materiais isolantes e outras solicitações durante o seu envelhecimento. Elas podem ocorrer também devido a uma aplicação inadequada de vácuo durante o ciclo de impregnação, como é o caso em sistemas de papéis isolantes impregnados a óleo usados em cabos, capacitores e transformadores. Nos geradores, particularmente nas barras estatóricas, elas ocorrem entre a camada de verniz e um condutor elementar, entre camada de pintura semicondutora e a isolação principal (efeito ranhura).

2.6 Detecção e Medição de Descargas Parciais

Existem várias técnicas de detecção de DPs: elétrica, acústica, óptica, química etc. Porém, a maioria destas são usadas apenas para complementar a detecção elétrica. A técnica acústica de medição por ultra-som está ganhando maior interesse, no aspecto de localização, triangulação de pulsos e reflexões acústicas, procurar as ocorrências das descargas no equipamento, sendo razoável para experiências de laboratório, mas no campo ainda é alvo fácil de ruídos e interferências, visto que, os sinais acústicos podem sofrer forte atenuação nos casos como DPs próximas dos enrolamentos. Por outro lado, se têm os modelos experimentais baseadas no método elétrico e nos conhecimentos *a priori*, com a finalidade também de localizar fontes de descargas, por meio de semelhanças físico/químicas, além de detectá-las. A seguir são listadas algumas dessas técnicas [1], [33]:

Método óptico ou visual – DP pode ser observada visualmente em ambientes escuros, após acostumar os olhos à escuridão. O observador pode utilizar binóculos de grande alcance, se necessário. Alternativamente, registros fotográficos ultra-rápidos podem ser feitos e, para fins especiais, são usados intensificadores de imagem.

Método químico – A presença de DPs em óleo ou equipamentos com isolantes a gás pode ser detectada, em alguns casos, pela análise da decomposição de produtos dissolvidos no óleo ou no gás. Esses produtos se acumulam durante operações prolongadas. A análise química pode também ser usada para estimar a degradação do isolamento causada pela ocorrência de DPs.

Método acústico – Este procedimento é baseado no ruído audível ou ultra-sônico gerado pelas DPs, isto é, ruído no ar ou vibrações em materiais adjacentes à fonte de descarga. A técnica acústica consiste em utilizar sensores piezelétricos ou transdutores, que podem ser conectados dentro ou fora do equipamento, de preferência em ambientes com baixo nível de ruído. Medidas acústicas feitas com microfones ou outros transdutores, em conjunto com amplificadores e instrumentos de registro adequados, podem ser úteis para localizar a descarga no equipamento. Diversos sensores, com alta sensibilidade para faixas de frequência acima da faixa audível, são usados para localizar descargas corona no ar. Transdutores também podem ser usados para localizar descargas em subestações isoladas a gás (SIGs) ou equipamentos imersos em óleo, como transformadores. Esta técnica é a mais adequada para localizar fisicamente a ocorrência de descargas e estão sendo utilizadas em diversos equipamentos tal como grandes transformadores.

Método da tensão de rádio interferência – O teste de tensão de rádio interferência (TRI) é baseado no fato de que DPs geram ondas eletromagnéticas na forma de interferência estática. O receptor é calibrado em micro-volts para medir o valor do sinal de entrada. Este método não é adequado para localizar fisicamente, dentro do equipamento, o local onde ocorrem descargas e sim para quantificar o valor das descargas. A medição de TRI é usada tanto em equipamentos de alta tensão como em linhas de transmissão. Nos equipamentos, os sinais de descarga são detectados através de uma resistência, enquanto que em linhas de transmissão, a interferência em rádio-frequência (RF) gerada pelas descargas, é detectada usando-se uma antena.

Método elétrico - Essa técnica faz com que o instrumento de detecção de descargas se torne parte do circuito elétrico, incluindo o equipamento onde as descargas estão ocorrendo. Neste método as descargas são medidas normalmente em pico Coulombs. Esse procedimento de medição é o mais utilizado em geradores, transformadores, cabos de alta tensão, capacitores, entre outros. Uma impedância de detecção RLC (resistiva, indutiva e capacitiva) ou RC (resistiva e capacitiva) pode ser usada, sendo a RLC para um modo de detecção de banda estreita de frequência e a RC para um modo de detecção de banda larga de frequência.

O método elétrico é, sem dúvida, o mais utilizado para quantificar DPs. As normas IEC 60270 e IEEE P1434 [6], [7] são entidades que tentam unificar e consolidar procedimentos para medição de DPs e fornecem várias alternativas de circuitos de ensaio e terminologia apropriada para realizar medição de DPs. Cabe ressaltar que ainda

não se têm procedimentos confiáveis e fáceis para situações bem específicas se aplicados num determinado modelo ou equipamento. O presente trabalho de tese, põe em evidência, a necessidade de ter novas ferramentas e faz uma introdução na busca de outras características das DPs que ajudem a desvendar particularidades a partir de dados coletados.

2.7 Causas e Conseqüências de Descargas Parciais

Como principais causas de DPs em isolantes, pode-se citar a inclusão de bolhas de ar em materiais sólidos e materiais isolantes de qualidade inferior. O surgimento de DPs também é facilitado pelo tratamento e impregnação inadequadas na fabricação do material dielétrico. Condutores dobrados inadequadamente, objetos isolantes estranhos (ex. restos de pinturas), cunhas frouxas ou soltas, pontas e rebarbas, colocação inadequada do isolante, dimensionamento inadequado da isolação e das blindagens e campos elétricos mal distribuídos, também contribuem para a ocorrência de DPs.

Os efeitos decorrentes de DPs em materiais dielétricos podem variar desde meramente indesejáveis a totalmente desastrosos. Dependendo da intensidade do campo elétrico, a ionização pode se dar em todo o percurso do material dielétrico, sendo denominada, nesse caso, de descarga disruptiva. DPs podem produzir luz, ruído audível e ozônio. Além disso, podem dar origem a outros efeitos, como [32]:

- Elevação da temperatura do material dielétrico.
- Perda de potência e perda de energia nas cavidades do material isolante.
- Geração de radiação ultravioleta.
- Agentes oxidantes como oxigênio e ozônio.
- Erosões nas cavidades de gás existentes em materiais dielétricos pelo choque mecânico entre elétrons e moléculas da parede da cavidade. Alguns materiais diferem nas suas taxas de erosão devido às descargas; materiais a base de polímeros sofrem mais erosão que mica, vidro e outros produtos cerâmicos.
- Geração de ácido oxálico em cavidades de polietileno e outros materiais isolantes.
- Absorção ou geração de gás.
- Interferência em comunicações de rádio e TV.

Em materiais isolantes líquidos (óleo isolante) as DPs podem gerar gases inflamáveis especialmente acetileno e metano causando a diminuição das propriedades

dielétricas e assim a vida útil deste material, obrigando a fazer substituições desses isolantes com mais frequência. Já em materiais isolantes sólidos, estas são mais severas, devido a sua difícil ou impossível substituição. Nesses isolantes, havendo descargas, poderá ocorrer erosão contínua ou a abertura de fissuras ou rachaduras que com o tempo poderão levar o isolamento à falha total.

Como pode ser observado, além de serem altamente prejudiciais ao material isolante, as DPs também têm influência negativa em sistemas adjacentes, como recepções de rádio e TV. Logo, a ocorrência de DPs em materiais isolantes de equipamentos elétricos de alta tensão deve ser analisada cuidadosamente, podendo ser um indicador qualitativo da vida útil esperada dos isolamentos assim como, do estado operativo de equipamentos.

2.8 Circuitos de Medição e Ensaio de DPs

Os circuitos utilizados para medir DPs podem ser deduzidos de acordo às indicações da norma IEC 60270. Nestas, existem três circuitos básicos de medição de DPs. Todos eles são baseados na detecção de uma queda de tensão sobre uma impedância de medição, provocada por impulsos de corrente no circuito exterior à amostra. A impedância de medição Z_m consiste normalmente de um resistor de medição, um indutor de alguns mH para dissipar as correntes a 60 Hz e uma capacitância parasita inerente ao circuito.

Para fazer testes de DPs em outras faixas de frequência, não previstas pelos fabricantes e/ou normas, deve-se levar em conta a variação dos capacitores com a frequência, assim como os erros introduzidos pelos transdutores (sensores) na medição de DPs. Além disso, deve-se levar em consideração as impurezas contidas nos dielétricos e defeitos de fabricação, pois estes podem trazer resultados ruins e conclusões erradas.

Os circuitos de medição ilustram a maneira como a impedância de medição pode ser inserida no circuito de ensaio, de acordo com a Figura 2.9. C_t representa o objeto sob teste, o qual na maioria dos casos é representado de maneira parcial por uma capacitância C_t . O calibrador conectado em paralelo com C_t serve para indicar a definição real da magnitude da DP, a carga de impulso aparente, a qual produz no objeto submetido a teste uma pequena queda de potencial, simulando o fenômeno interno ao ter-se DPs no isolamento. Esta queda momentânea do potencial pode produzir um impulso de corrente $i(t)$ unicamente quando se tem um capacitor de

acoplamento C em paralelo com C_t suficientemente grande e que a impedância de medição Z_m seja muito pequena.

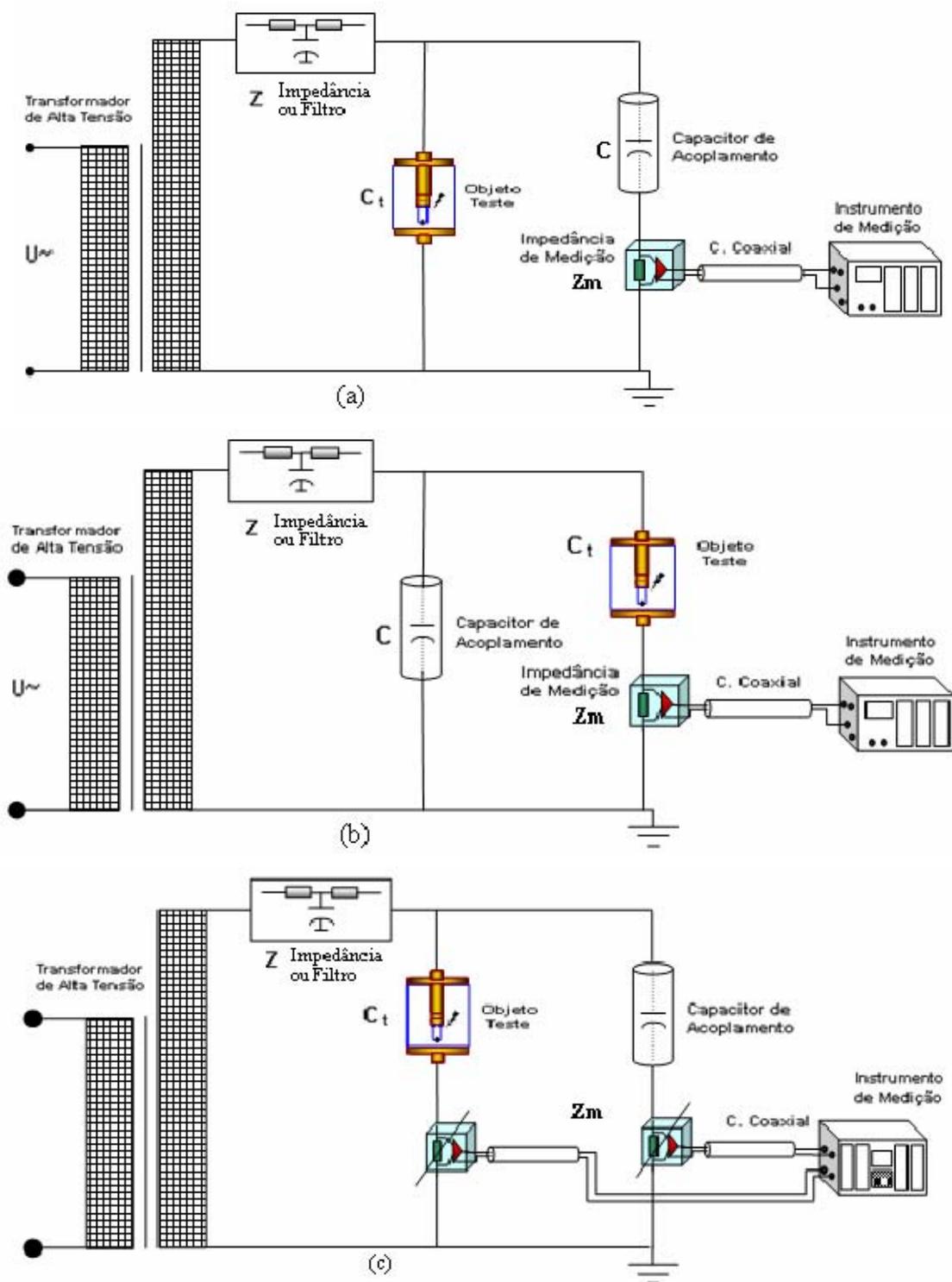


Figura 2.9 – Circuitos básicos de medição de DPs [16]. a) Impedância Z de medição em série com o capacitor C de acoplamento. b) Impedância de medição Z_m em série com o equipamento sob ensaio. c) Circuito de ensaio equilibrado ou balanceado.

As Figura 2.9.a e b, ministram medição direta de implementação simples e rápida. Porém apresentam desvantagens sendo sensíveis às interferências, inclusive os de

baixas frequências. O circuito de medição por ponte indicado pela Figura 2.9.c suprime grande parte das interferências, mas a dificuldade está na calibração e no balanço e sincronização do equipamento de medição tipo ponte.

De acordo com normas internacionais a medição de carga é definida em pico-Coulomb (pC). Os equipamentos de medição de DPs efetuam uma integração das correntes que na maioria funcionam baseadas na semi-integração de valores, por meio de amplificadores de banda larga e banda estreita.

O circuito de medição correspondente ao da Figura 2.9.a é utilizado nos casos em que o objeto sob ensaio possui uma extremidade aterrada. Esse tipo de configuração protege o equipamento de medição de falhas advindas do objeto sob ensaio. No entanto, o circuito da Figura 2.9.b, com a impedância da medição em série com o objeto sob ensaio, possui uma maior sensibilidade de medição. Esse tipo de configuração é utilizado nos casos em que o lado de baixa tensão do objeto sob ensaio fica isolado do referencial terra. O circuito de medição da Figura 2.9.c tem o lado de baixa tensão do objeto sob ensaio e do capacitor de acoplamento isolado do referencial terra através das impedâncias de medição Z_m e Z_m' . Segundo [6], este circuito apresenta vantagens no que se refere aos problemas de interferências externas.

Dentre essas possibilidades oferecidas pela norma IEC 60270 para configurações de circuitos de medição, o circuito mais utilizado é aquele com a impedância de medição em série com o capacitor de acoplamento, correspondente à Figura 2.9.a. Nesse caso, o capacitor de acoplamento evita que a tensão à frequência industrial passe para a impedância de medição, passando a constituir um caminho preferencial para os pulsos das correntes correspondente às DPs.

Deve ficar claro, no entanto, que a escolha do tipo de circuito de medição deve estar associada às características elétricas dos equipamentos sob ensaio. Por exemplo, para equipamentos com capacitância elevada, como os capacitores de potência, o circuito que melhor se aplica para medição de DPs é o circuito de ensaio balanceado, correspondente à Figura 2.9.c.

Medições de sinais de DPs em transformadores de potência têm sido pesquisadas por várias décadas, formando-se várias linhas de pesquisa em detecção, assim como localização. Ambos requerem várias condições, tais como baixo nível de ruído, adequada calibração e sensibilidade da instrumentação de medição. A realização de uma medição de DPs em transformadores é dificultada pela acessibilidade nos circuitos internos e a complexidade destes circuitos, altamente indutivos.

No transformador, a fim de solucionar esses problemas, a conexão ao sistema de medição é feita através da bucha, que é um tap capacitivo. A Figura 2.10 ilustra o circuito para medição de DPs acoplado no braço de uma bucha do transformador.

As medições, realizadas para possíveis fontes de DPs em transformadores de potência, através de células testes foram feitas de acordo com o circuito da Figura 2.8.a onde os pulsos de corrente provocados pelas DPs produzem um sinal nos terminais da impedância de medição Z_m . Para pulsos de curta duração, o sinal produz um pulso de tensão, cujo valor de crista é proporcional à carga aparente da amostra, que é normalmente medida em pico-Coulomb.

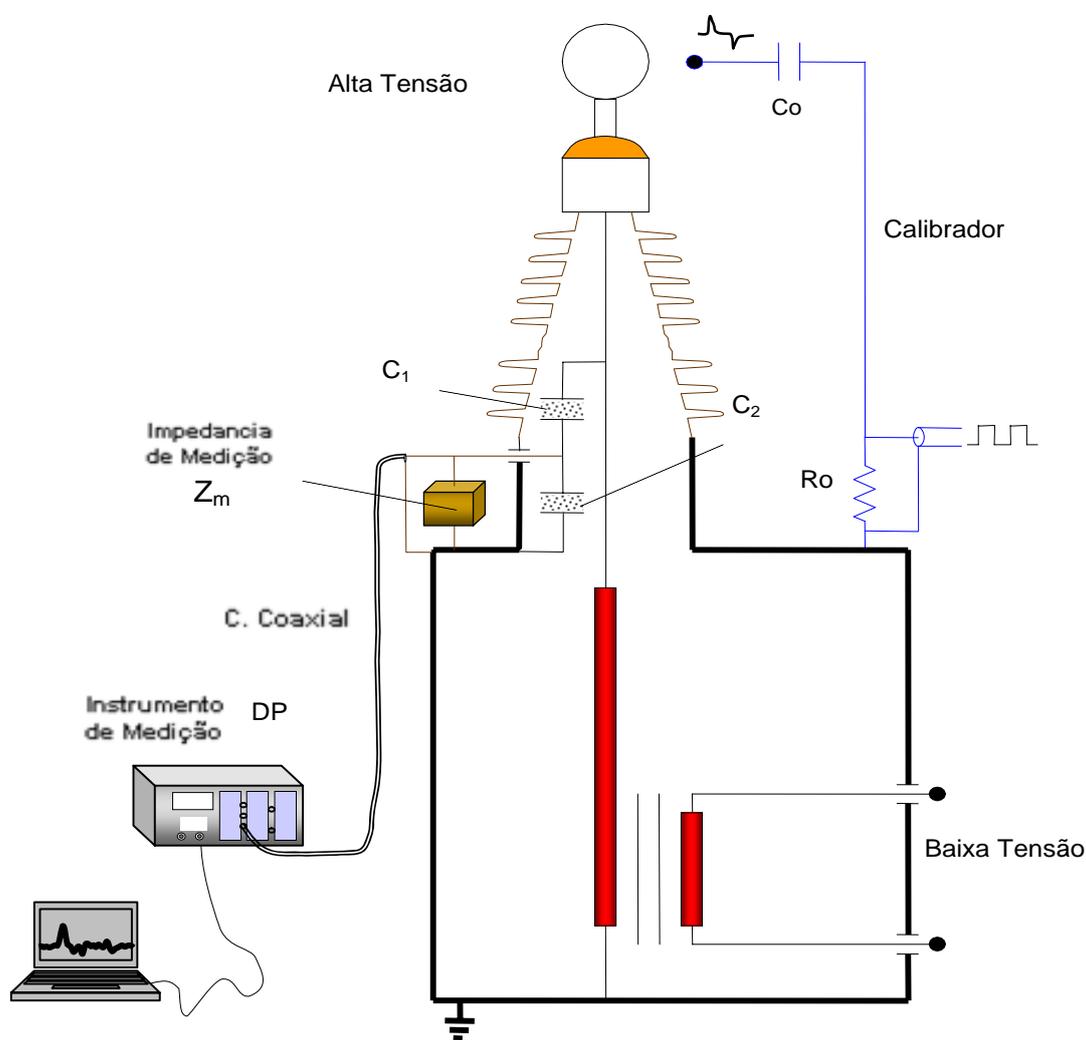


Figura 2.10 – Esquema de medição de DPs para transformadores no campo.

Os pulsos de carga aparente, individuais, podem ser coletados por um amplificador e apresentados em um instrumento de medição adequado. A amplitude dos pulsos detectados pode, então, ser medida através de comparações com pulsos de

calibração. O valor C (capacitor de acoplamento) é de aproximadamente $1nF$ e a impedância de medição normalmente se encontra na faixa de 1000 até 2000 Ohms. O filtro está representado por dois indutores e uma capacitância, podendo ter configurações diferentes, desde que sejam filtros de alta frequência. Um filtro típico é um indutor de $3mH$.

Vale ressaltar que, em casos gerais, além de instalar capacitores de acoplamento de alta tensão, também é possível utilizar alternativas como transformadores de corrente de alta frequência ou bobinas de Rogowski e sensores de campo indutivo/capacitivo. Alguns tipos de acopladores são instalados durante uma interrupção do serviço pelas atividades da manutenção e, por outro lado, devido à necessidade de ter melhores pontos de medição alguns fabricantes de máquinas têm se esforçado em desenvolver barras estatóricas especiais com micro sensores de medição. Assim, torna-se viável obter melhores condicionamentos para detectar ocorrências de sinais de DPs off-line e por monitoração on-line. Porém, muitos acopladores e estratégias de medição ainda são vulneráveis e sensíveis a interferências.

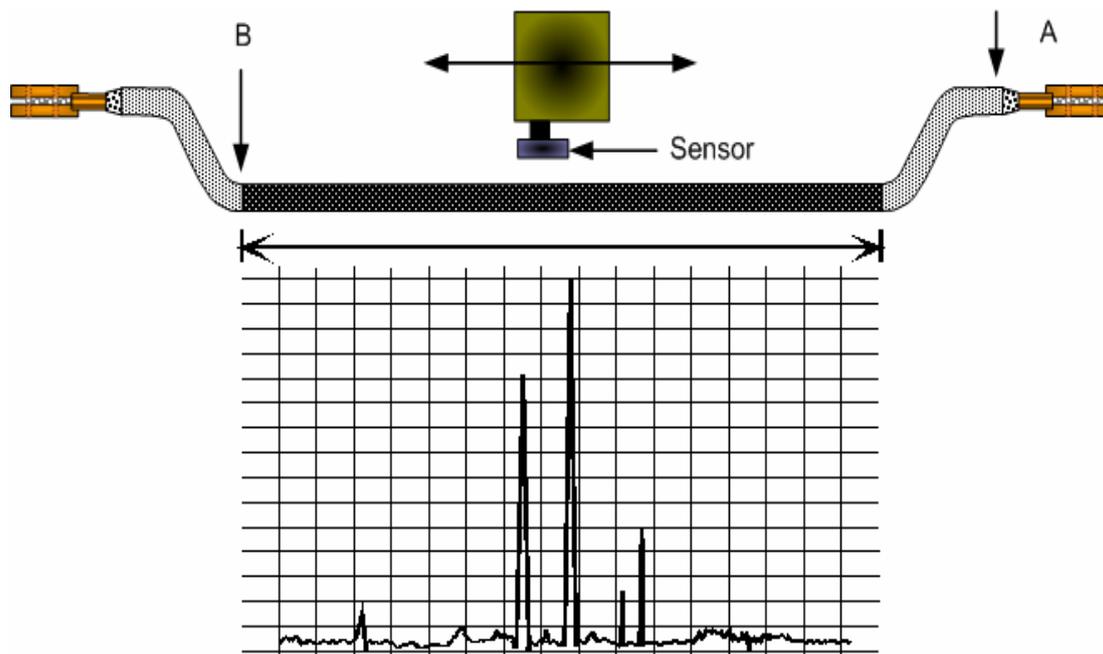


Figura 2.11 – Localização de descargas por sensores de campo elétrico.

A Figura 2.11 mostra um sistema de detecção de descargas utilizando sensores indutivos de campo os quais são comparados com sensores capacitivos e a impedância de medição num círculo característico. O sensor percorre desde o braço do ponto A até o ponto B [34].

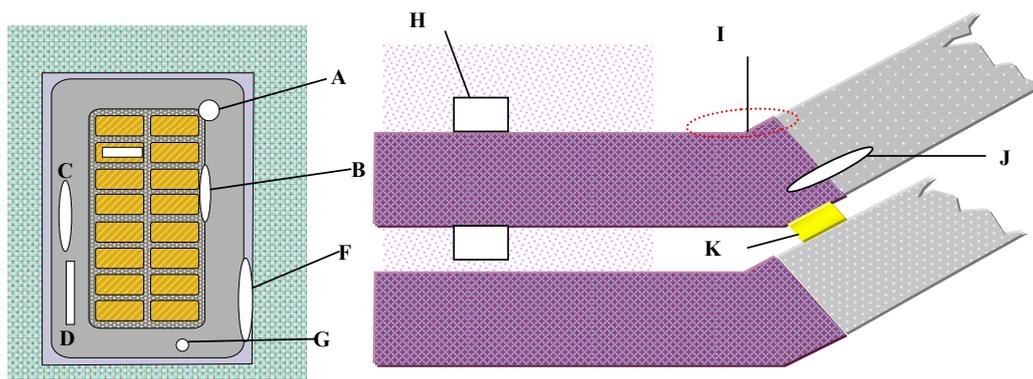


Figura 2.12 – Corte transversal e longitudinal de barras estatóricas de gerador, zonas possíveis de formação de DPs.

A Figura 2.12 mostra um corte transversal e longitudinal de barras estatóricas de gerador, na qual se indicam zonas possíveis de formação de DPs: A até G indicam defeitos típicos em barras, localizadas em ranhuras estatóricas (slot). H indica possível defeito no canal de refrigeração e I, J e K defeitos na cabeça de bobina [30], [90], [88], [89].

2.9 Interferências na Detecção de DPs e Formalização do Problema

As interferências presentes no processo de medição, os quais foram descritas em pesquisas anteriores como em [32], mostra-se na Figura 2.13. Outras formas usuais de classificação mais simples são chamadas de ruídos conduzidos e irradiados (acoplantes) e ruídos externos e internos [15], [24]. Por conseguinte, os problemas devido às interferências no processo de medição têm sido pesquisados e tratados para serem eliminados, como foi introduzido na Secção 2.1. Atualmente existem diversas tecnologias para contornar e reduzir as influências das perturbações em detecção de DPs.

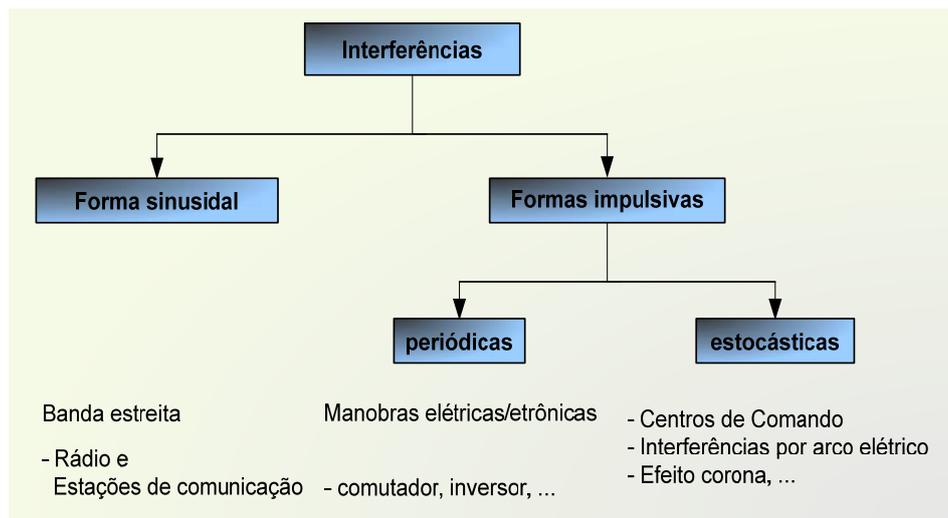


Figura 2.13 – Classificação ruídos em detecção de sinais DPs [32], [30].

Contudo, a informação coletada das medições dificulta a interpretação e identificação de DPs, pois, a coleção de dados ainda apresenta ruído de fundo. Assim, essa coleção terá uma baixa confiabilidade para se diagnosticar o equipamento em questão.

Pesquisas apontam uma outra fonte de interferências ou de fontes espúrias (*outliers*) introduzidas pelos conversores Analógico-Digital (A/D) dos instrumentos de medição e o efeito de acoplamento (*crostalks*) entre canais de medição.

Finalmente, a supressão de sinais de ruídos em dados coletados é o propósito desta dissertação.

Capítulo 3

Modelos Experimentais

As técnicas de detecção das DPs são baseadas em vários fenômenos produzidos pela descarga tal como luminescência, calor, ruído (ultra-som e audível), transformações químicas e pulsos de descargas elétricas. Estas técnicas são descritas amplamente nas referências [32]. Nesta parte da tese resumem-se unicamente as que foram utilizadas no desenvolvimento da pesquisa, enfatizando o método elétrico segundo as normas IEC-60270 e IEEE-P1434. Apresentam-se também os modelos experimentais para representação de fontes de defeitos, possíveis de serem encontrados em transformadores de potência, denominadas células testes. Para alguns defeitos em enrolamentos estáticos de geradores os modelos são baseados em barras envelhecidas.

3.1 Detecção de Descargas Parciais

A técnica adotada, para o projeto de tese, foi o método elétrico, tanto para detecção como para medição de DPs, visto que o Centro de Pesquisas de Energia Elétrica – CEPTEL trabalha há duas décadas com esse método. Por conseguinte, o CEPTEL forneceu gentilmente suas instalações de laboratório e proporcionou a informação necessária, assim como, os dados coletados de hidrogeradores.

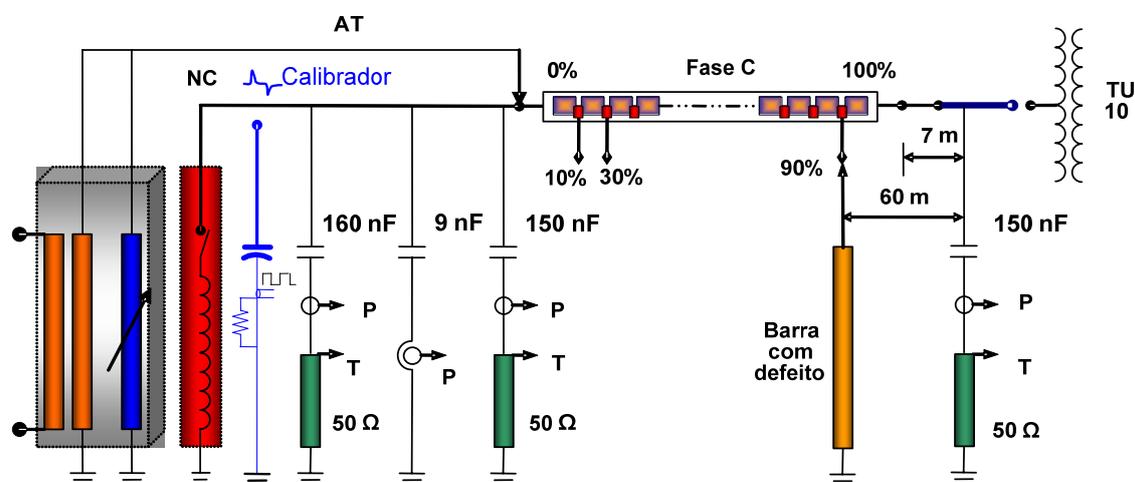


Figura 3.1 – Circuito básico de medição de DPs em geradores: Estratégia praticada no enrolamento estático da fase C no gerador U10 na usina Itaipu [35], onde AT (alta tensão), NC (cubículo de neutro), P (acoplador de 9 nF) e T (acoplador de 150 nF-fase e 160 nF-neutro).

3.2 Medição de Descargas Parciais Através de Modelos

A Figura 3.1 mostra o arranjo para medição de DPs em geradores, particularmente em bobinas estatóricas. Foram implementadas diversas combinações do posicionamento dos capacitores (para alta tensão) de acoplamento. Também se realizam a alocação de barras com defeito (chamados também modelos experimentais) em posições diferentes do enrolamento, níveis de tensão aplicada e variadas faixas de frequência de medição. Outros complementos e arranjos são mostrados na Figura 3.2, onde cada fase do estator é constituída por um conjunto de espiras que por sua vez são compostas por barras. Para efeitos de detecção de DPs faz-se um percurso exclusivamente por espiras cujos terminais estejam disponíveis para medição.

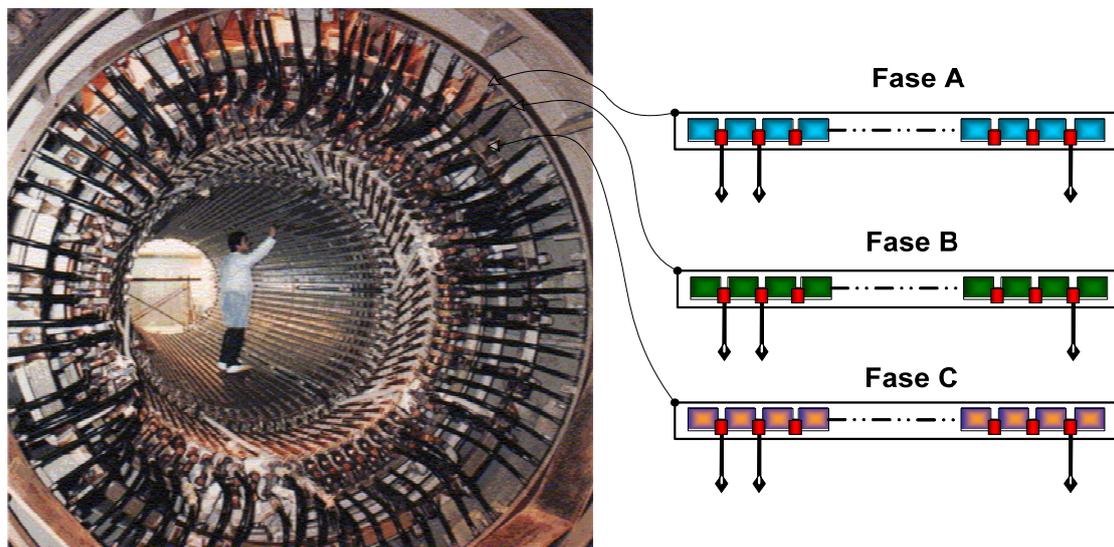


Figura 3.2 – Enrolamento estatórico, onde cada fase é constituído por um conjunto de barras.

3.2 Medição de Descargas Parciais Através de Modelos

A detecção de DPs é um dos meios altamente solicitados para testar a confiabilidade dos equipamentos de alta tensão, como geradores e transformadores. Medição de DPs é uma das mais importantes ferramentas para detectar imperfeições nos isolamentos, sendo indicador de formação de cavidades gasosas na isolação dos equipamentos. Porém esta é dificultada por interferências e ruídos externos e internos no equipamento além da complexidade de seus circuitos, altamente indutivos e/ou capacitivos. Estas características afetam os sinais de DPs, contribuindo na sua deformação e atenuação. Os ruídos se propagam facilmente através dos equipamentos por condução e/ou irradiação tendo, às vezes, amplitudes maiores que os sinais de DPs [32].

Na atualidade a instrumentação de medição para DPs tende a ser mais sofisticada. Baseados nos processamentos de sinais digitais estes instrumentos permitem um melhor controle de filtros e *threads*, diminuem alguns casos de ruídos e permitem ter maior

capacidade de registro dos sinais e melhor controle de variáveis por meio de ajustes dos parâmetros de medição.

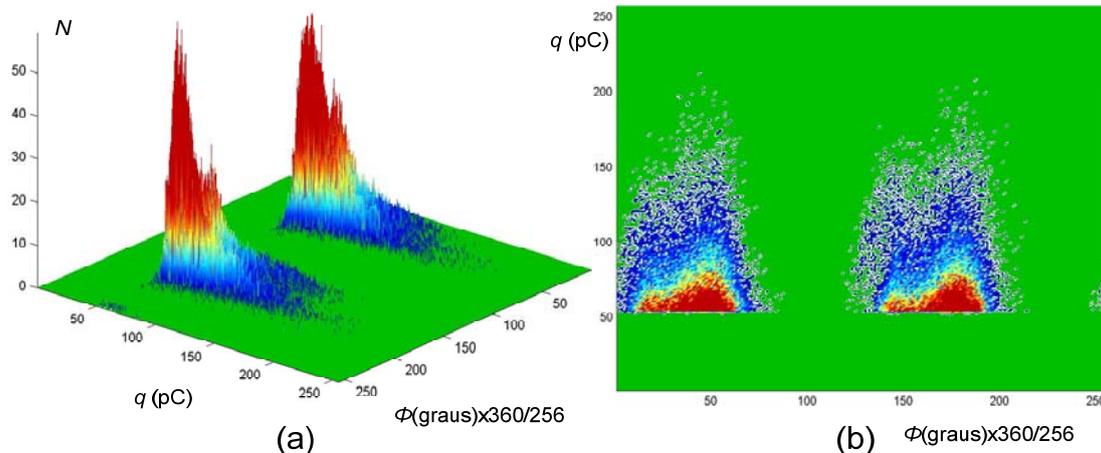


Figura 3.3 – Mapas digitais de DPs obtidos pelo ICM: (a) DPs em 3-D $f(\phi, q, N)$ e (b) é Mapa 2-D $H(\phi, q)$.

O instrumento de medição ICM, da Power Diagnostix Systems, utilizado para os ensaios praticados nesta tese, emprega a técnica de medição de carga aparente, especificada na norma IEC 60270 [6]. Os procedimentos de medição e calibração do equipamento estão descritos em [36], [37]. O ICM tem a particularidade de relacionar o ângulo de fase, a amplitude e taxa de repetição (ϕ, q, N) em forma de mapas de assinaturas das DPs dispostas em 2-D (bidimensional), onde as descargas são referenciadas e distribuídas angularmente sob a tensão aplicada, as amplitudes (q) são picos máximos dos pulsos de descarga atingida em janelas ajustáveis ($5 \mu s$ a $6 ms$) de registro em tempos tais como de $20 \mu s$ (usualmente) e a taxa de repetição (N) foi codificada em cores normalizadas (ver Figura 3.3 a e b).

Os mapas digitais (ϕ, q, N) são também chamados de PRPHA – *Phase Resolved Pulse Height Analyzer* – segundo [19].

3.3 Modelagem de Defeitos Típicos

3.3.1 Célula Teste de Ensaio

As medições realizadas neste projeto foram feitas, na sua maioria, utilizando células experimentais de ensaio, desenvolvidas no CEPEL. Essas células são inseridas no circuito de medição no lugar do objeto sob ensaio, considerando que o fenômeno de DPs que ocorre nas células ocorreria em um equipamento real. Isto é, cada célula está associada fisicamente e quimicamente (conhecimento a priori) a uma ou mais partes internas do equipamento [32].

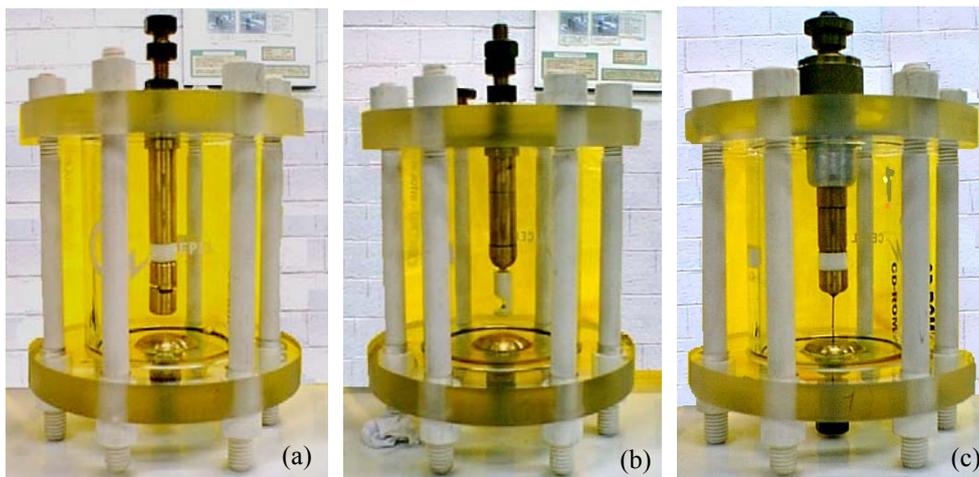


Figura 3.4 – Células de ensaio para medições de DPs reais. a) Ponta flutuante b) Ponta plana c) Ponta terra.

As células foram construídas a partir de uma configuração teórica e livre de influências externas (interferências bloqueadas) para fins de pesquisa, sendo de forma cilíndrica, de vidro e de material acrílico translúcido. Para geração de pulsos de corrente e maior concentração de carga, foram utilizados eletrodos de agulha de aço de 1 a 2 mm de diâmetro e raio de curvatura de ponta de 50 μm . A distância entre os eletrodos de descarga é de 5 a 15 mm.

As células de ensaio (ver Figura 3.4) foram configuradas para isolar possíveis defeitos que possam ocorrer num transformador de potência, sendo denominados como em [32]:

Célula ponta flutuante – Esta configuração representa descargas induzidas envolvendo o enrolamento de alta tensão ou de baixa tensão do transformador e descargas entre camadas de bobinas.

Célula ponta plana - Descargas no eletrodo de alta tensão representam descargas envolvendo saídas em alta tensão, algum objeto condutor flutuante ou na bucha do transformador, mau contato nos conectores de alta tensão ou nos enrolamentos do transformador.

Célula ponta terra – Esta célula representa descargas entre a bucha e o tanque aterrado do transformador ou entre a bobina de alta ou baixa tensão e o núcleo. Também pode envolver grampos de fixação e conectores.

3.3.2 Barras de Prova

Foram preparadas algumas amostras de barras de aproximadamente um metro de comprimento cada, nas quais foram provocados defeitos característicos e conhecidos

com a finalidade de reproduzirem DPs localizadas. Alguns destes são mostrados na Figura 3.5 e denominados como:

Barra 1 (slot) – descargas de ranhura

Barra 2 (corona) – corona na cabeça de bobina (*end-winding*)

Barra 3 (delaminação) – delaminação (internal void of the ground wall insulation)

Algumas variantes foram realizadas conseguindo até oito fontes de sinais de DPs, as quais serão descritos na seção de processamento de sinais de DPs.

As barras de prova representam defeitos reduzidos (Ver Figura 3.5), visto que estes possuem localização *a priori* e são fontes de descarga que representam o fenômeno com a melhor relação-sinal-ruído possível. Isto é, algumas fontes de ruído típicos em medição foram filtradas e/ou eliminadas por ajuste e calibração dos acopladores de medição. No entanto, as verdadeiras magnitudes de DPs ainda não são consolidadas, já que ainda existem outras fontes de ruídos, conduzidas e acoplantes, que podem ser inseridas pela superposição de pulsos e até pelos conversores analógico/digital do instrumento medição.

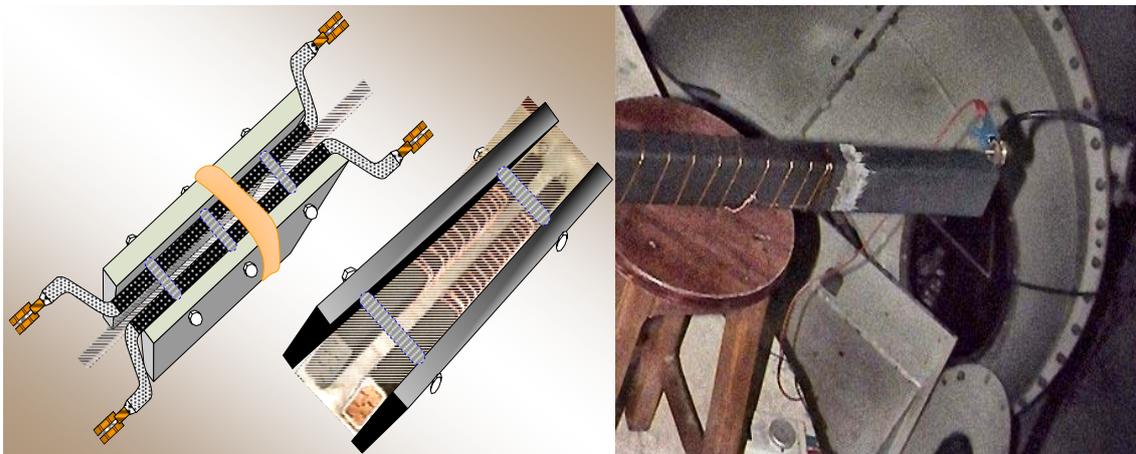


Figura 3.5 – Modelos experimentais por barras de prova para reproduzir defeitos típicos no enrolamento estático de geradores.

A montagem do circuito de medição de DPs *off-line* em geradores é realizada quando o gerador é retirado do serviço e após o desligamento do sistema de resfriamento. Segundo a Figura 3.6 a montagem é feita no campo. Antes de efetuar os ensaios de medição é realizada uma rigorosa busca da melhor relação sinal ruído auxiliado pela calibração dos acopladores de medição e de um analisador de espectros. Assim, é possível ter a melhor resposta em frequência do circuito de medição e do material isolante envolvido.

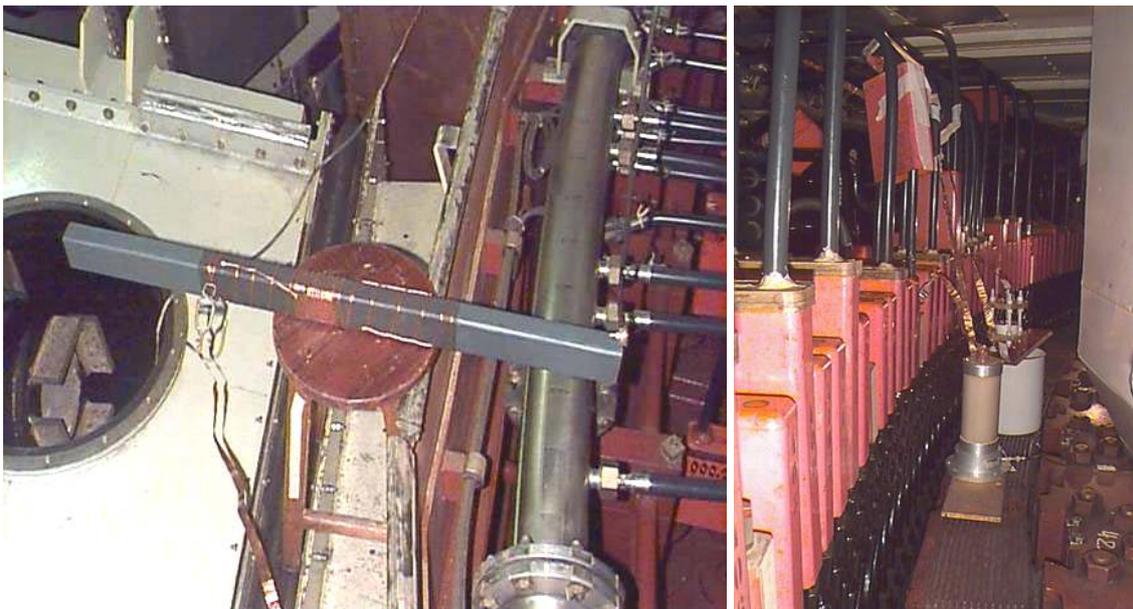


Figura 3.6 – Montagem de barra de prova e dos capacitores de acoplamento.

3.4 Calibração do Circuito de Ensaio

O sistema de medição de DPs detecta a carga medida (q_m), que, dependendo do valor do capacitor de acoplamento, é somente uma fração da carga aparente. Como medida para ajustar as diferenças entre a carga medida e a carga aparente, é feita uma calibração no circuito de ensaio. O processo de calibração, que é essencial para uma medição correta de DPs, deve ser repetido para cada configuração do circuito de ensaio ou quando for efetuada alguma mudança nos parâmetros desse circuito.

O gerador de calibração (GC) pode ser implementado como um elemento integrado ao instrumento de medição ou como uma unidade independente operada a bateria. No caso mais simples, o GC é um gerador de onda quadrada com amplitude ajustável V_0 e emite pulsos de calibração de carga q_0 com uma magnitude definida através de um capacitor C_0 , que é muito menor que a capacitância do objeto sob ensaio (C_t). A carga q_0 é definida pela Equação (3.1).

$$q_0 = V_0 C_0 \quad (3.1)$$

Os pulsos de calibração devem ser sincronizados com a frequência da fonte de tensão para que se obtenha uma posição estacionária dos mesmos.

O tempo de subida deve estar na faixa de 50 a 100 ns para assegurar que os pulsos tenham um espectro de frequência similar aos pulsos de DP originais. O circuito de calibração pode ser visto na Figura 3.7 .

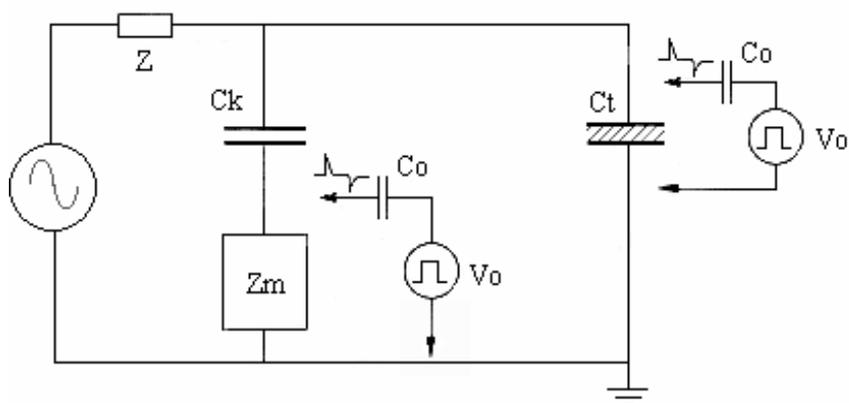


Figura 3.7 – Diagrama do circuito de ensaio e calibração [17], [43], Ck é capacitor de acoplamento.

No método da calibração direta, os pulsos com a carga q_0 são inseridos no objeto sob ensaio, ou seja, no circuito de alta tensão. A leitura do instrumento de medição deve ser calibrada para o valor de q_0 para que durante a medição de DPs, os valores sejam exatamente iguais à carga aparente q . É importante remover o gerador de calibração do circuito de ensaio antes que a fonte de alta tensão seja ligada ao circuito.

No método de calibração indireta, os pulsos de calibração são inseridos na impedância de medição Z_m . Como não há a necessidade de tirar o gerador de calibração do circuito, este tipo de calibração pode ser feito durante o ensaio de alta tensão e isso permite uma visão simultânea dos pulsos de calibração e dos sinais de descarga que estão sendo medidos.

Para assegurar uma calibração correta, recomendado pela norma IEC-60270, deve ser feita a calibração direta no início da medição, para qualquer configuração do circuito de ensaio. Na prática, é feita a calibração direta, injetando uma carga q_0 diretamente nos terminais do objeto sob ensaio.

Calibrações descritas na norma IEC-60270 de descarga parcial não é aplicável diretamente a geradores, visto que um enrolamento estatórico íntegro não pode ser tratado como uma capacitância concentrada.

Quando os pulsos de descarga parcial se propagam através do enrolamento estatórico, estes podem sofrer forte atenuação, particularmente em altas frequências causando sua ressonância. Os pulsos de DPs ocorrem realmente em alguma parte do enrolamento estatórico, porém são medidos e calibrados nos terminais da máquina. Portanto, os pulsos de DPs podem ter amplitudes e formas de ondas diferentes dos pulsos originais. Como consequência é difícil precisar o levantamento da chamada "condição de isolamento estatórica", usando apenas as amplitudes de DPs.

3.4 Calibração do Circuito de Ensaio

Os medidores de DPs encontrados comercialmente têm variadas larguras de banda e diferentes centros de frequência. Conseqüentemente produzem diferentes saídas de descarga parcial para um mesmo evento de medição de DPs.

Contudo, ciente destas observações, as medições são realizadas em circuitos independentes do gerador para fins de ajuste e calibração do circuito de ensaio e aparelhagem. Ou seja, nesta etapa é realizada a verificação do efeito das capacitâncias tanto nos acopladores de medição quanto no objeto alvo de ensaio. Após isso, é realizado um conjunto de medições para cada defeito típico.

Defeitos típicos ao longo da fase C do enrolamento estático do gerador podem ser encontrados na referência [35], [90].

Capítulo 4

Pesquisas Anteriores e Classificadores Inteligentes em DPs

Com o intuito de atender ao problema de interferências introduzidas, no contexto das DPs na Seção 2.9, neste Capítulo da dissertação se realiza uma breve revisão dos classificadores inteligentes introduzidos em pesquisas anteriores, assim como em [32], [38] e em alguns recentes artigos que foram desenvolvidos no andamento da tese [3], [39]. Deste modo, alguns destes resultados serão utilizados para avaliar o desempenho dos caracterizadores dos sinais de DPs que serão descritos no Capítulo 5.

4.1 Princípios de Reconhecimento de Padrões

O tratamento da informação depende da natureza, quantidade e qualidade dos dados, que em sua maioria são incertos. Uma das técnicas desenvolvidas a fim de entender os problemas de incertezas e ambigüidades, buscando a forma mais próxima possível da realidade, é a identificação ou descobrimento de padrões de dados que guardem certas relações comparativas de igualdade ou similaridade entre eles. Assim, fazer reconhecimento de padrões significa realizar pesquisa de estrutura de dados [40].

As metodologias utilizadas para esquemas de reconhecimento, em DPs, incluem classificadores, aproximações estatísticas (estimadores e probabilidade), teoria dos conjuntos fuzzy, perceptrons (rede neural) e outras variantes, e técnicas de compactação de dados.

Os dados coletados geralmente podem ser transformados em valores numéricos, os quais podem ser dispostos em vetores padrão ou em pares relacionais (similaridades e proximidades); de modo que, sejam quantificados, qualificados ou caracterizados por um reconhecedor de padrões, por exemplo.

As medições de similaridade definem uma medida da distância geométrica ou física dos dados. Estas medidas têm sido exploradas em diversos algoritmos que enriquecem técnicas e/ou métodos classificadores, discriminadores e reconhecedores de

padrões: caracterizadores estatísticos de sinais (ex.: filtros, PCA, ICA), Redes Neurais (ex.: NLPCA, ICA) [5], [41], Híbridos Neuro-Fuzzy (ManFis) [38] e Lógica Fuzzy, onde os dados são interpretados como variáveis lingüísticas.

Por conseguinte, na atualidade existem variadas técnicas para processamento e análise da informação dos sinais de DPs. No entanto, nem todas são aplicáveis para o mesmo fenômeno [32]. Logo, uma adequada análise característica em DPs dependerá de quanto da sua natureza se conhece, além da quantidade e qualidade dos dados coletados.

Em muitos casos, os dados apresentam natureza incerta, o que dificulta sua análise e exige melhores algoritmos, mais robustos e flexíveis, para processar a informação que, na maioria dos casos, é altamente complexo, proveniente de sinais ultra-rápidos e intermitentes.

Uma coleção de dados obtidos por sistemas de medição está sujeita às incertezas como ruídos térmicos em componentes eletrônicos e nos transdutores que têm a tarefa de fazer a conversão físico-química em elétrica. Ou seja, uma boa aquisição de dados depende também dos equipamentos de medição, de sua calibração e sensibilidade.

4.2 Classificadores e Reconhecimento de Padrões Inteligentes

As Redes Neurais Artificiais (RNA) foram utilizadas para reconhecer padrões de DPs, tendo desempenhos extraordinários frente às técnicas holonômicas e estatísticas. Quando hibridizadas com um Sistema de Inferência Neuro-Fuzzy adaptativo de Mamdani (MANFIS) [42], formam um classificador de padrões de DPs eficiente. Por outro, lado Simulated Annealing (SA), qualificado como um bom otimizador estável, foi hibridizado com o MANFIS e, este por sua vez foi hibridizado com um Algoritmo Genético (AG) [39]. Dessa forma, os benefícios de cada técnica foram aproveitados para trazer melhores desempenhos na caracterização dos sinais de DPs.

4.2.1 Redes Neurais Artificiais

Redes Neurais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência [43], [44]. As Redes Neurais Artificiais (RNA) representam uma ferramenta poderosa e já com plena aplicação em inúmeras áreas. Elas têm a capacidade de realizar mapeamentos complexos e de aprender uma representação sem o conhecimento explícito do algoritmo que as rege.

Uma Rede Neural Artificial é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Essas unidades, geralmente, são conectadas por canais de comunicação que estão associados a um determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma Rede Neural Artificial vem das interações entre as unidades de processamento da rede [43].

A maioria dos modelos de RNAs possui alguma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados. Em outras palavras, elas aprendem através de exemplos.

Arquiteturas neurais são tipicamente organizadas em camadas, como *Multilayer Perceptron* (MLP), de acordo com a Figura 4.1, com unidades que podem estar conectadas às unidades da camada posterior. Porém, a conexão de uma unidade de processamento com outra que pertence a uma camada diferente pode dar-se. Este último é uma arquitetura artesanal construída por otimizador artificial, tal como AGs, por exemplo.

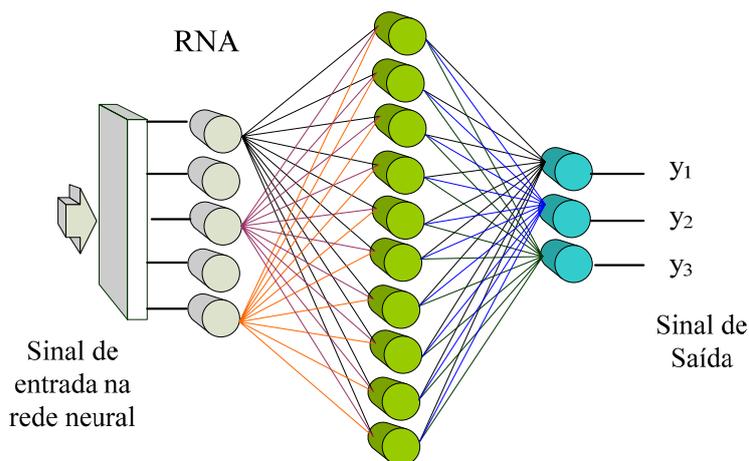


Figura 4.1 – Rede neural artificial

Uma das RNA utilizadas para identificação de classes de fontes de DPs é simples, baseada na retropropagação do erro de saída e uma combinação de funções linear e tangente hiperbólica, organizadas em uma rede de três camadas. Esta rede é amplamente descrita na referência [32].

Para facilitar a nomenclatura típica adota-se “rede neuronal” como rede neural artificial e unidades de processamento como neurônios.

4.2.2 Híbridos Neuro-Fuzzy (MANFIS)

Há várias formas de se definir sistemas híbridos. De uma forma simples, sistemas híbridos são aqueles que utilizam mais de uma técnica de identificação de sistemas para a solução de um problema [32]. Existem vários sistemas híbridos, sendo abordados nesta pesquisa apenas o Sistema Neuro-Fuzzy de Mandami. Este modelo pode apresentar grandes vantagens em relação às técnicas de identificação de sistemas individuais (por exemplo, redes neurais, sistemas fuzzy), pois tem a capacidade de interpretação, aprendizado, estimação de parâmetros e generalização.

Neuro-Fuzzy combina a capacidade de aprendizado das Redes Neurais Artificiais (RNA) [38] com o poder de interpretação lingüístico dos Sistemas de Inferência Fuzzy (SIF) [43]. Tal fato deve-se principalmente à união dos seguintes fatores: a) aplicabilidade dos algoritmos padrões de aprendizado, desenvolvidos para redes neurais, b) possibilidade de promover a integração de conhecimentos e c) possibilidade de extração de conhecimentos, sob o formato de regras fuzzy, a partir de um conjunto de dados. Além disso, pode-se incorporar informação adicional do especialista através de variáveis lingüísticas.

A idéia básica de um Sistema Neuro-Fuzzy é implementar um SIF numa arquitetura paralela distribuída de tal forma que os paradigmas de aprendizado comuns às RNA's possam ser aproveitados nesta arquitetura híbrida [39], [45]. A pesquisa na área dos Sistemas Neuro-Fuzzy busca, portanto, a maximização das propriedades desejáveis das RNA e SIF e a minimização de suas limitações/deficiências.

O MANFIS (Mamdani Adaptive Neuro-Fuzzy Inference System) [38], mostrado na Figura 4.2, foi implementado para fazer o reconhecimento de DPs. Este é um Sistema Neuro-Fuzzy Híbrido, onde foi implementado um sistema fuzzy numa estrutura de rede neural, baseado nas regras de inferência de Mamdani. Esse modelo possibilita uma grande redução do esforço computacional através de um pré-processamento fuzzy dos dados de entrada. Após este estágio, é possível determinar os conseqüentes de todas as regras fuzzy, diretamente extraídas do banco de dados. Portanto, o MANFIS é baseado num sistema de inferência fuzzy e o processo adaptativo é iniciado após todas as condições fuzzy estarem completamente estabelecidas [32] [38].

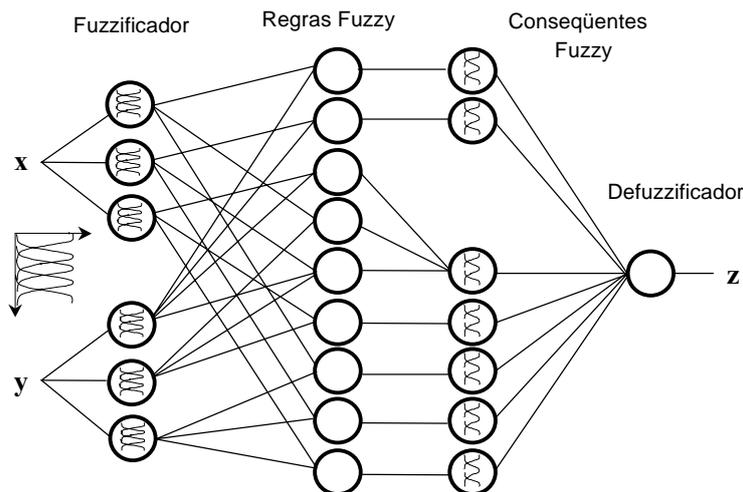


Figura 4.2 - Sistema Híbrido Neuro-Fuzzy MANFIS

A Função de ativação “bellshape” é caracterizada pela função de pertinência $\mu(g)$, conforme equação:

$$\mu(x, a, b, c) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}} \quad (4.1)$$

onde o valor da função $\mu(g)$ é definido na faixa de $\{0-1\}$, c é o centro da função “bellshape”, a é o ponto de cruzamento, b define a função *steepness* ($b/2a$ é a inclinação da função bellshape em $x = c - a$) e x representa alguma variável de entrada no nó.

De acordo com a Figura 4.2, as *bellshapes* das camadas fuzzificador e de conseqüentes determinam a função de saída do defuzzificador na (variável z) com um erro de saída. Os parâmetros a , b , c e cz variam a cada iteração até chegarem a um equilíbrio e não se deslocarem mais [38].

Outra característica importante do MANFIS é que nem todos os neurônios de uma camada estão conectados e estas *sinapses* servem exclusivamente de conexão unitária, ou seja, os pesos de cada conexão não são atualizados pela retropropagação do erro. Só os parâmetros de cada *Bellshape* são atualizados a cada iteração do algoritmo, motivo pelo qual se denomina processo adaptativo.

4.2.3 Simulated Annealing em DPs

Simulated Annealing é uma ferramenta poderosa para problemas de otimização, utilizada em casos de enorme complexidade. Foi introduzida por Metropolis [40], [46], como um método de se determinar propriedades físico-químicas relacionadas a um grupo de átomos em transição para o equilíbrio térmico. Mais a frente Kirkpatrick [40]

tornou explícito o conceito de esquema de resfriamento. Já Geman e Geman [47], [39] implementaram limite inferior para a velocidade de queda de temperatura na condição de garantir a convergência do algoritmo para um mínimo global.

O processo de Metropolis é constituído basicamente em duas etapas. Na primeira procede-se à elevação inicial da temperatura a um estado de energia máxima e, na segunda, verifica-se o seu abaixamento sucessivo e suficientemente lento para que as partículas do sistema se combinem de forma a atingirem o estado de energia mínima, isto é, de tal forma que seja atingido o equilíbrio térmico.

A simulação da evolução das soluções é baseada em técnicas de Monte Carlo e na geração de estados sucessivos. Supõe-se, como ponto de partida, um estado possuindo energia E_i a partir do qual, recorrendo a um mecanismo apropriado, é gerado um outro estado com energia E_j . Se a diferença de energias $E_j - E_i$ for inferior ou igual a zero o novo estado é aceite como estado atual. Se a condição anterior não se verificar, o novo estado poderá ser ainda aceite com uma probabilidade função da diferença de energias, entre estados sucessivos da temperatura do banho. A expressão (4.2) apresenta o processo de calculo dessa probabilidade:

$$p(\gamma) = e^{\frac{-(E_j - E_i)}{K_B T}} \quad (4.2)$$

onde, T é a Temperatura e K_B é a constante de Boltzmann.

A Figura 4.3 mostra o processo de resfriamento utilizando aleatoriedade, governada por um parâmetro de controle ou temperatura $T(k)$ para buscar um espaço discreto para um mínimo de uma função de energia $E(k)$.

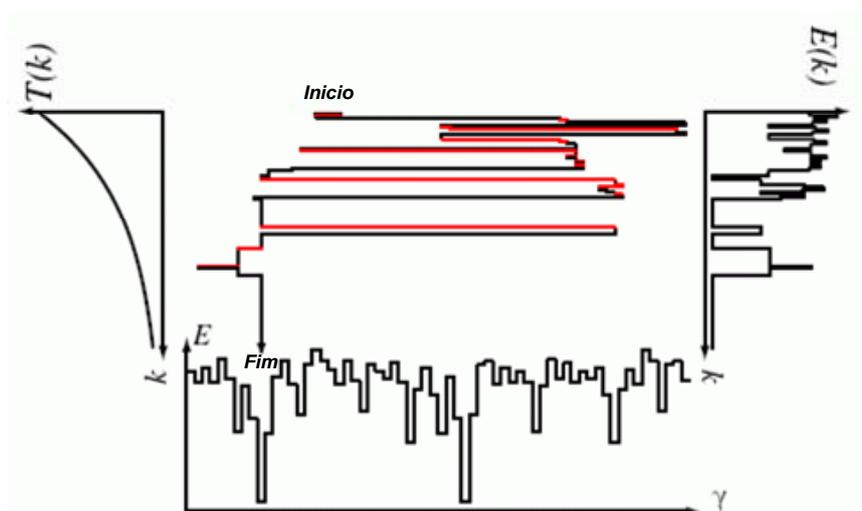


Figura 4.3 – Simulated annealing estocástico de uma função de energia [40]

4.2.4 Simulated Annealing e MANFIS

O MANFIS, com três conjuntos *fuzzy*, obteve bons resultados na aproximação da função (verificar mais a frente), mas não chegou a apresentar o melhor resultado. Pretende-se melhorar este resultado utilizando *Simulated Annealing*.

O MANFIS, com maior número de conjuntos *fuzzy*, cai no problema tradicional do custo computacional para resolver este problema. Neste caso, implementou-se com apenas três conjuntos *fuzzy* por variável x e y , respectivamente, e com nove regras adaptativas a saída do sistema, variável z . Alguns patamares da função não tiveram boa aproximação, devido à resolução dos conjuntos *fuzzy*. Um dos propósitos deste trabalho consiste em resgatar o resultado achado por qualquer técnica usada e otimizá-la adotando uma função de minimização como objetivo.

Ferramentas comuns usam gradiente descendente para achar o mínimo de suas funções-objetivo (valor esperado) e têm grande chance de ficar presas e convergir para mínimos locais. *Simulated Annealing* assegura a convergência para o mínimo global, desde que se faça uma boa escolha e assunções das condições iniciais e taxas de resfriamento adequadas.

4.2.5 Algoritmos Genéticos e MANFIS em DPs

Devido à necessidade de se procurar as melhores respostas, diversas ferramentas foram desenvolvidas com o objetivo de otimizar resultados ou encontrar respostas com ótimo global. Diversos métodos de busca têm sido desenvolvidos baseados em processos naturais e físicos. Entre os mais usuais estão os Algoritmos genéticos e *Simulated Annealing*.

Algoritmos genéticos (AGs) são métodos estocásticos de busca que emulam teorias evolucionárias biológicas para resolver problemas de otimização. Diversos trabalhos sobre *Simulated Annealing* citam vantagens desse método sobre os AGs, e vice-versa. É evidente que certos algoritmos adaptam-se melhor do que outros a determinados problemas, e isso contribui para acalorar essa discussão.

Os AGs se caracterizam por sua flexibilidade e facilidade de paralelização e, as vantagens atribuídas aos algoritmos de *Simulated Annealing* são sua melhor fundamentação analítica – que permite um maior controle sobre as características do algoritmo– e, principalmente, a existência de uma prova de convergência em probabilidade. Entretanto, os AGs são a mais utilizadas para administrar problemas de

descontinuidade, não diferenciável, multimodais e não convexos. AGs podem conviver com estes problemas [40]. [48].

Os AGs se adaptam mais naturalmente aos problemas de otimização combinatória, devido a sua natureza intrínseca discreta. Já os algoritmos por Simulated Annealing se adequam com menor esforço em problemas contínuos [46], [47].

Montagem de um Algoritmo Genético

- Representação genética das soluções viáveis do problema.
- Determinação de uma população inicial de cromossomos.
- Definição da função de avaliação dos cromossomos.
- Definição dos operadores genéticos eficazes na reprodução de novos cromossomos.
- Definição dos parâmetros que compreendem o tamanho da população, critérios evolutivos, estagnação, parada, etc.

Implementação de AGs no MANFIS

O MANFIS adapta seus parâmetros (a , b e c), Equação (4.1), usando gradiente descendente na condição de garantir um mínimo valor para seu desempenho. Esta estrutura faz com que o MANFIS dependa das derivadas parciais de seus parâmetros que podem cair em ótimos locais. Assim, para contornar estes inconvenientes, os AGs podem atuar como um otimizador natural, eliminando o efeito das derivadas e das possíveis descontinuidades das funções características.

Capítulo 5

Caracterizadores de Sinais e Tratamento de Ruídos

Os ruídos e a grande dimensão dos dados coletados são um forte obstáculo para a implementação eficiente de técnicas de classificação, discriminação e reconhecimento de padrões em DPs. Na maioria dos casos, é necessário um pré-processamento, baseado em técnicas de compactação da dimensão dos dados de DPs. Para tal procedimento, utiliza-se nesta dissertação a Análise de Componentes Principais (PCA) de Karhunen-Loève [40], [49], [50]. O PCA permite resgatar as informações mais relevantes e diminuir o esforço computacional das técnicas aplicadas. Vale lembrar que, existem outras técnicas PCAs para a mesma finalidade, sendo que algumas destas também serão alvos de pesquisa para o desenvolvimento desta tese.

Para caracterizar os sinais, reduzindo os ruídos em DPs, empregam-se Filtros Casados (MF-Matched Filter), [52] e classificador neural, como uma técnica de processamento de sinais híbridos. Os MFs são baseados na técnica de filtragem ótima linear, maximizando a relação sinal-ruído, onde o sinal é discriminado através de sua média espelhada (geralmente para sinais determinísticos) [51] e o classificador neural, caracterizado por sua robustez, realiza a discriminação das flutuações do sinal (DPs como sinal estocástico) [118].

Outra maneira de lidar com os obstáculos mencionados acima é a introdução de técnicas mais especializadas, baseadas na separação cega de fontes (BSS) e na análise de componentes independentes (ICA) [120].

O que é proposto neste Capítulo é a integração destas técnicas num sistema caracterizador. Assim, pretende-se caracterizar os sinais fontes, na fase de pós-deteção de DPs, isto é, a partir da informação crua (dados coletados) para produzir dados consolidados.

Descrevem-se a seguir as principais técnicas e métodos aplicados e analisados nesta pesquisa, enfatizando a análise de componentes principais, separação cega de fontes e a análise de componentes independentes.

5.1 Pré-Processamento de Dados em DPs

O banco de dados coletado pelo equipamento de medição é composto por sinais de descargas parciais provenientes dos modelos de ensaio: três células de ensaio e barras de prova. O equipamento de medição de descargas parciais ICM fornece os dados de medição em código binário na forma de mapa bidimensional distribuído, de acordo com códigos de cores (ver Figura 3.3). Estes dados são convenientemente convertidos em bases decimais e normalizados de acordo com a maior amplitude alcançada por um pulso.

O problema inerente de um sistema de classificação é o tamanho do padrão representado pela densidade de distribuição ângulo-amplitude, que é uma matriz inicial que tem 2^{16} elementos, a qual não é apropriada para alimentar um classificador ou discriminador em um sistema de reconhecimento de padrões. Os dados devem ser compactados e convertidos em uma representação conveniente para seu reconhecimento, processo este denominado seleção de característica [40]. A matriz primitiva de dimensão 256×256 (ver Figura 3.3b) é a representação de um padrão de DP, pode ser disposta, sem perdas importantes de generalidades, em uma matriz reduzida por várias modalidades de resolução. Entre elas pode-se utilizar, por exemplo, janelas de 8×8 das 256^2 para gerar uma nova matriz de 64×64 ($256/4 \equiv 64 \times 64$). Outras reduções de escala, tais como 32×32 ($256/8 \equiv 32 \times 32$), 16×16 ($256/16 \equiv 16 \times 16$) ou 8×8 ($256/32 \equiv 8 \times 8$) [8] também podem ser utilizadas desde que as perdas não sejam significativas.

A modalidade empregada para a seleção característica, nesta dissertação, é a matriz reduzida 32×32 , pois a resolução de 1024 amostras é a redução mínima suficiente para utilizar em algoritmos e técnicas. Após essa seleção de característica, é aplicada análise de componentes principais (PCA), para extrair informações relevantes, processo conhecido como extração característica [73].

5.2 Compactação de Dados por Fatores Principais

PCA é talvez uma das mais bem conhecidas técnicas em análise multivariável e em mineração de dados. Ela foi introduzida por Pearson, que a utilizou num contexto biológico e, logo desenvolvida por Hotelling com trabalhos feitos em psicometria. PCA foi também desenvolvida por Karhunen no contexto da teoria de probabilidade e foi subsequentemente generalizada por Loève. Recentemente, eficientes e poderosos algoritmos adaptativos têm sido desenvolvidos por PCA, SVD e MCA e suas extensões

[40], [44], [53]. Ao mesmo tempo, algoritmos como PCA Robusto (RPCA) [53], PCA Probabilístico (PPCA) [40], Kernel PCA (KPCA) [49], [54], PCA não-linear (NLPCA) [41] e Fator de Análise não-linear (NFA) [55], [49], [61] enriquecem o vasto grupo de técnicas, as quais podem ser encontradas em diversas em aplicações [56] [57], [58], [59].

As técnicas citadas acima, além das aplicações tradicionais, também são utilizadas com frequência como métodos de pré-branqueamento de sinais de fontes ruidosas. Adicionalmente, a separação cega de fontes (BSS) e a análise de componentes independentes (ICA) frequentemente aplicam uma análise PCA ou FA como uma etapa de pré-processamento, devido às seguintes propriedades:

- Ajuda a filtrar o possível ruído gaussiano contido nos sinais fonte.
- Permite descorrelacionar os sinais fonte com o propósito de branqueá-las.
- Permite compactar a informação, no sentido de otimizar o erro médio quadrático. Transforma um espaço de dados de uma determinada dimensão para um espaço característico de menor dimensão.
- Permite realizar a separação de fontes com um PCA não linear.
- Permite identificar grupos de objetos ou *outliers*.

5.3 Análise de Componentes Principais

A Análise de Componentes Principais é uma técnica útil para simplificar a visualização de agrupamentos ocultos num sistema ortogonal, que representa, de forma reduzida, o espaço completo das observações [38], [44]. As direções principais, expressos pelos componentes são tipicamente em número muito menor que a dimensão original dos dados, o que permite uma compactação significativa destes com preservação da informação original principal. Isso significa uma diminuição do esforço computacional, já que permite utilizar arquiteturas menores das técnicas caracterizadoras. Assim, a implementação destas poderá ser feita com dados para os quais fenômenos menos relevantes não tenham uma influência marcante no processo de caracterização do sinal [5].

No capítulo anterior, se fez referência às redes neurais como classificador e reconhecedor de padrões. Além dessas, as redes neurais podem ter outras finalidades e particularidades, como os algoritmos de aprendizado não-supervisionado. Estes se organizam em certo sentido, tal que, eles podem detectar ou extrair: características úteis, regularidades, correlacionar ou separar dados/sinais ou, descorrelacionar alguns sinais

com pequeno ou sem conhecimento a priori do resultado desejado [60]. No entanto, as regras de aprendizado hebbiana normalizado e antihebbiana são, em particular, meramente variantes de um algoritmo básico não-supervisionado. Algoritmos de aprendizado por componentes principais (PCA), decomposição pelo valor singular (SVD) e análise de componentes minoritários (MCA) pertence a esta classe de regras não supervisionadas. Nesta parte da tese faremos uma revisão do PCA utilizado como compactador de dados das descargas parciais.

5.4 Introdução a Separação Cega de Fontes

Recentemente, Separação Cega de Fontes (BSS) por Análise de Componentes Independentes (ICA) tem recebido bastante atenção devido ao seu potencial e aplicabilidade para emprego em diversas áreas onde muitas técnicas tradicionais não conseguem ter bom desempenho devido à suas restrições inerentes.

Fazendo uma divisão arbitrária, pode-se dizer que existem duas tendências de pesquisa em ICA. Por um lado, o estudo de separação de fontes misturadas, observadas num *array* de sensores tem sido um difícil e clássico problema de processamento de sinais.

Um dos primeiros estudos nesta área foi dado por Jutten, Herault e Guerin (1988) onde eles propuseram um algoritmo adaptativo numa arquitetura de simples retroalimentação [60]. A regra de aprendizado foi baseada numa aproximação *neuromimetic* e, foi habilitada para separar de forma simultânea, fontes desconhecidas e independentes. Esta aproximação tem sido explicada em diversos desenvolvimentos por Jutten e Herault (1991), Comon (1991), Karhunen e Joutsensalo (1993), Cichocki e Moczyski (1992) e outros. Mais a frente, Comon (1994) [62] introduz o conceito de análise de componentes independentes e propõe funções custo, relacionadas à minimização da informação mútua entre os sensores [53].

Por outro lado, paralelamente ao estudo de separação cega de fontes tem sido estudadas regras de aprendizado não-supervisionado em teoria da informação por Linsker (1992), Becker e Hinton (1992) e outros. A idéia foi maximizar a informação mútua entre a entrada e saída de uma rede neural. Esta aproximação já tinha sido estudada por Barlow (1961), relacionando a maximização da informação com a redução da redundância, estratégia proposta como uma codificação em neurônios. Cada neurônio deve codificar características que são estatisticamente independentes com os outros

neurônios. Esta idéia foi utilizada para a notação do código fatorial que foi explorada para estratégia de processamento visual por Attik (1992).

Nadal e Parga (1994) mostraram que no caso de baixo nível de ruído, a maximização da informação entrada/saída de um processador neural implica que a distribuição dos sinais de saída foi fatorada. Roth e Baram (1996) e Bell e Sejnowski (1995) deduziram independentemente a regra de aprendizado pelo método do gradiente estocástico para critérios de maximização e separação cega de fontes. Eles aplicaram para previsão de dados e series de tempo respectivamente.

Bell e Sejnowski (1995) foram um dos primeiros a explicar o problema de separação cega de fontes, baseado na teoria da informação, aplicada para separar e decompor a convolução de fontes por métodos adaptativos. Este método é mais plausível de uma perspectiva de processamento neural que as funções custo, baseadas em cumulantes proposto por Comon. Um método similar não neural para separação cega de fontes foi proposto por Cardoso e Laheld (1996).

Outros algoritmos baseados em diversas aproximações têm sido propostos: a estimação da máxima semelhança aproximada proposta por Gaeta e Lacoume (1990), maximização da negentropia aproximada por Girolami e Fyfe (1996), algoritmo PCA não linear desenvolvido por Karhunen e Joutsensalo (1994) e Oja (1995). Lee, Girolami e Sejnowski (1997) propõem um modelo unificado para o problema de separação cega de fontes, explicando a relação entre os algoritmos. A otimização da regra de aprendizado foi proposta por Amari (1997) [53], [63] utilizando o conceito de gradiente natural ou também chamado de gradiente relativo, proposto por Cardoso e Laheld (1996).

O algoritmo proposto por Bell e Sejnowski (1995) foi confiável para fontes super gaussianas. Para contornar esta limitação outras técnicas tem sido desenvolvidas, habilitadas para separar simultaneamente as fontes sub-gaussianas e super-gaussianas, Perlmutter e Parra (1996).

Novos algoritmos foram introduzidos por Amari e Akuzawa (2001), baseados na teoria Reimanniana, restrições não-holonômicas e geometria diferencial. Eles desenvolveram um ICA, que não requer pré-processamento, isto é, o algoritmo é robusto frente a ruídos de natureza gaussiana, exonerando as estatísticas de segunda ordem e explorando apenas cumulantes de altas ordens, visto que o ruído gaussiano se define com estatísticas de segunda ordem.

5.5 Caracterizadores de Sinais em Fontes de DPs

O ruídos inerentes aos equipamentos de alta tensão mencionados nos capítulos anteriores e os encontrados nos próprios componentes eletrônicos do instrumento de medição, como nos cabos e transdutores, são influentes no processo de aquisição dos sinais de DPs. Geralmente eles são tratados por implementação de filtros convencionais, anexados ao circuito de medição (hardware) e também os sintonizáveis introduzidos via software. Destes há atualmente produtos comerciais, porém custosos e alguns não são práticos [9], [91].

Outro enfoque para tratamento de ruídos é a remoção dos mesmos na fase de pós-deteção das DPs nos dados coletados, que é material de estudo desta pesquisa.

Para desenvolver um caracterizador híbrido MF-RNA é importante conhecer os princípios dos filtros ótimos que têm seus fundamentos na teoria de MF e RNA, descrita no Capítulo 4, a qual se emprega para qualificar o desempenho do filtro.

Outro caracterizador/classificador de DPs é PCA e RNA. A técnica de PCA é resumida a seguir.

Também a técnica de Separação Cega de Fontes (BSS), por meio da Análise de Componentes Independentes (ICA) é aplicada com os seguintes algoritmos: ICA não-holonômico, JADE, Infomax e FastICA, seus fundamentos serão descritos a seguir.

5.6 Fundamentos - Técnicas de Caracterização

5.6.1 PCA Linear

O PCA de Karhunen-Loève [49], [53] é uma das ferramentas clássicas para análise de dados com muitas variáveis. Além do PCA, outra técnica semelhante é a Análise de Fatores (FA) [64], [65]. Ambas são técnicas estatísticas aplicadas a um conjunto de dados para descobrir quais subconjuntos de variáveis, dentro do conjunto, formam grupos coerentes que são ortogonais entre si. A diferença entre PCA e FA se dá quando é realizada a computação de matriz (ex.: covariância). PCA assume que toda variância é comum para todos os fatores, enquanto que FA assume que há uma única variância. O nível de única variância se dá pelo modelo de FA escolhido. Conseqüentemente PCA é considerado como um sistema fechado, contrariamente ao FA, considerado como sistema aberto [65].

Fundamentos da Análise de Componentes Principais

Suponha um vetor \mathbf{X} com dimensão m . Assuma que este vetor tenha média zero, ou seja, $E[\mathbf{X}] = 0$, onde E é o operador estatístico de esperança. O vetor \mathbf{q} é unitário, também de dimensão m , no qual o vetor \mathbf{X} será projetado. Essa projeção é definida pelo produto interno de \mathbf{X} e \mathbf{q} , conforme Equação (5.1) a seguir,

$$\mathbf{a} = \mathbf{X}^T \mathbf{q} = \mathbf{q}^T \mathbf{X} \quad (5.1)$$

sujeito à condição:

$$\|\mathbf{q}\| = (\mathbf{q}^T \mathbf{q})^{\frac{1}{2}} = 1, \quad (5.2)$$

sob a suposição que o vetor \mathbf{X} tem média zero, segue-se que o valor médio da projeção \mathbf{A} é zero também, ou seja, $E[\mathbf{A}] = \mathbf{q}^T E[\mathbf{X}] = 0$.

A variância de \mathbf{A} é conseqüentemente a mesma de seu valor médio quadrático:

$$\begin{aligned} \sigma^2 &= E[\mathbf{A}^2] = E[(\mathbf{q}^T \mathbf{X})(\mathbf{X}^T \mathbf{q})] \\ &= \mathbf{q}^T E[\mathbf{X}\mathbf{X}^T] \mathbf{q} \\ \sigma^2 &= \mathbf{q}^T \mathbf{R} \mathbf{q} \end{aligned} \quad (5.3)$$

a matriz $\mathbf{R}_{m \times m}$ é a matriz de auto-correlação do vetor \mathbf{X} , formalmente definida como o valor esperado do produto externo do vetor \mathbf{X} com ele mesmo:

$$\mathbf{R} = E[\mathbf{X}\mathbf{X}^T], \quad (5.4)$$

observa-se que a matriz \mathbf{R} é simétrica, ou seja:

$$\mathbf{R}^T = \mathbf{R}, \quad (5.5)$$

desta propriedade, segue-se que \mathbf{a} e \mathbf{b} são vetores quaisquer $m \times 1$, logo:

$$\mathbf{a}^T \mathbf{R} \mathbf{b} = \mathbf{b}^T \mathbf{R} \mathbf{a} \quad (5.6)$$

Da Equação (5.3) vemos que a variância σ^2 , da projeção \mathbf{A} , é uma função do vetor unitário \mathbf{q} ,

$$\psi(\mathbf{q}) = \sigma^2 = \mathbf{q}^T \mathbf{R} \mathbf{q} \quad (5.7)$$

A próxima questão a ser considerada é encontrar os vetores unitários \mathbf{q} nos quais $\psi(\mathbf{q})$ tem valores extremos e estacionários (máximos e mínimos locais), sujeitos à condição da norma Euclidiana de q . A solução deste problema está na estrutura da matriz de correlação \mathbf{R} . Se \mathbf{q} é um vetor unitário de tal forma que a variância teste $\psi(q)$ tenha um valor extremo, logo para uma pequena variação $\delta \mathbf{q}$ do vetor \mathbf{q} , busca-se um $\delta \mathbf{q}$ de primeiro ordem,

$$\psi(\mathbf{q} + \delta\mathbf{q}) = \psi(\mathbf{q}) \quad (5.8)$$

Da definição da variância $\psi(\mathbf{q})$ dada na Equação (5.7), temos:

$$\psi(\mathbf{q} + \delta\mathbf{q}) = (\mathbf{q} + \delta\mathbf{q})^T \mathbf{R}(\mathbf{q} + \delta\mathbf{q}) = \mathbf{q}^T \mathbf{R}\mathbf{q} + 2(\delta\mathbf{q})^T \mathbf{R}\mathbf{q} + (\delta\mathbf{q})^T \mathbf{R}\delta\mathbf{q}$$

onde, Equação 5.6 é usada e o termo de segunda ordem $(\delta\mathbf{q})^T \mathbf{R}\delta\mathbf{q}$ foi ignorado, finalmente a Equação 5.7 foi utilizada, resultando:

$$\psi(\mathbf{q} + \delta\mathbf{q}) = \mathbf{q}^T \mathbf{R}\mathbf{q} + 2(\delta\mathbf{q})^T \mathbf{R}\mathbf{q} = \psi(\mathbf{q}) + 2(\delta\mathbf{q})^T \mathbf{R}\mathbf{q} \quad (5.9)$$

O uso das equações (5.8) e (5.9) implicam que,

$$(\delta\mathbf{q})^T \mathbf{R}\mathbf{q} = 0 \quad (5.10)$$

Qualquer variação $\delta\mathbf{q}$ de \mathbf{q} não será admissível, ou melhor, é restrito o uso de perturbações em que a norma Euclidiana do vetor $\mathbf{q} + \delta\mathbf{q}$ seja igual a um, ou seja, $\|\mathbf{q} + \delta\mathbf{q}\| = 1$, ou equivalentemente, $(\mathbf{q} + \delta\mathbf{q})^T (\mathbf{q} + \delta\mathbf{q}) = 1$. Assim, tomando a Equação 5.2, imprescindível para restringir uma primeira ordem em $\delta\mathbf{q}$,

$$(\delta\mathbf{q})^T \mathbf{q} = 0 \quad (5.11)$$

A Equação 5.11, implica que a variação $\delta\mathbf{q}$ deve ser ortogonal a \mathbf{q} e por tanto unicamente uma mudança na direção de \mathbf{q} é permitida.

Por convenção, os elementos dos vetores unitários \mathbf{q} são de menor dimensão no sentido físico. Se combinássemos as equações (5.10) e (5.11), seria preciso introduzir um fator de escala λ na equação subsequente com as mesmas dimensões da matriz de correlação \mathbf{R} . Assim, temos:

$$(\delta\mathbf{q})^T (\mathbf{R}\mathbf{q} - \lambda\mathbf{q}) = 0 \quad (5.12)$$

Para que a condição da Equação (5.12) seja mantida é suficiente e necessário que:

$$\mathbf{R}\mathbf{q} = \lambda\mathbf{q} \quad (5.13)$$

A Equação (5.13) representa a equação característica, onde a correlação é o núcleo da transformação similar da forma contínua [49]. A partir desta equação, podemos obter os autovalores que representam os valores próprios da matriz (λ) e suas projeções ortogonais dadas pelos autovalores (ql). Assim,

$$\mathbf{R}\mathbf{q}_j = \lambda_j \mathbf{q}_j \quad j = 1, 2, \dots, m \quad (5.14)$$

Os autovalores são arranjados em ordem decrescente,

$$\lambda_1 > \lambda_2 > \dots > \lambda_m \quad (5.15)$$

sendo $\lambda_1 = \lambda_{\text{máxima}}$. Os autovalores utilizados para construir a matriz $\mathbf{Q}_{m \times m}$ são:

$$\mathbf{Q} = [q_1, q_2, \dots, q_j, \dots, q_m] \quad (5.16)$$

Podemos combinar as m equações representadas na Equação (5.14) numa simples equação:

$$\mathbf{RQ} = \mathbf{Q}\mathbf{\Lambda} \quad (5.17)$$

Onde $\mathbf{\Lambda}$ é uma matriz diagonal definida pelos autovalores da matriz \mathbf{R} :

$$\mathbf{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_j, \dots, \lambda_m] \quad (5.18)$$

A matriz \mathbf{Q} é uma matriz ortogonal no sentido de que os vetores coluna (ex.: autovetores de \mathbf{R}) satisfazem as condições de ortogonalidade:

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1, & j=i \\ 0, & j \neq i \end{cases} \quad (5.19)$$

A Equação (5.19) requer autovalores distintos. Equivalentemente, podemos escrever $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$, da qual deduzimos que o inverso da matriz \mathbf{Q} é igual a sua transposta, conforme Equação (5.20) abaixo:

$$\mathbf{Q}^T = \mathbf{Q}^{-1} \quad (5.20)$$

Isso significa que a Equação (5.17) pode ser reescrita em uma forma conhecida como a transformação ortogonal de similaridade,

$$\mathbf{Q}^T \mathbf{RQ} = \mathbf{\Lambda} \quad (5.21)$$

Ou na forma expandida:

$$\mathbf{q}_j^T \mathbf{Rq}_k = \begin{cases} \lambda_j, & k=j \\ 0, & k \neq j \end{cases} \quad (5.22)$$

A transformação ortogonal de similaridade (unitária) da Equação (5.21) transforma a matriz de correlação \mathbf{R} em uma matriz diagonal de autovalores. A matriz de correlação \mathbf{R} pode ser expressa em termos de seus autovalores e autovetores como:

$$\mathbf{R} = \sum_{i=1}^m \lambda_i \mathbf{q}_i \mathbf{q}_i^T \quad (5.23)$$

Resumem-se dois pontos importantes encontrados na estrutura de PCA, conforme abaixo:

- Os autovetores da matriz de auto-correlação \mathbf{R} pertencentes ao vetor \mathbf{X} definem os vetores unitário \mathbf{q}_j , representando a principal direção ao longo da qual a variância $\psi(q_j)$ tem seus valores extremos.

- Os autovalores definem os valores extremos da variância $\psi(q_j)$

Resumindo, dado um vetor x , é possível reduzir a dimensionalidade do mesmo através dos sistemas de equações (5.24) abaixo, onde o vetor a_j é conhecido como o vetor projeção dos componentes principais:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{bmatrix} = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_l^T \end{bmatrix} x, \quad l \leq m \quad (5.24)$$

5.6.2 PCA por Redes Neurais e PCA Não Linear

As redes neurais utilizadas para a extração de componentes principais foram introduzidas por Oja (1982) [53], por meio de um simples neurônio linear e utilizando a regra Hebbiana e modelo iterativo. Subseqüentemente, surgiram estudos utilizando a mesma estrutura para formar bancos de filtragem como os filtros PCA de Foldiak (1988) [49] e foi implementada a regra Hebbiana generalizada numa rede neural para extrair os primeiros k componentes principais. Entretanto, em 1985, Baldi e Hornik [41] expõem seu trabalho para extração supervisionada de componentes. Eles utilizaram redes neurais de duas camadas com neurônios lineares. Na década de 90 surgiram vários estudos, como os de Diamantaras e Kung (1996), batizando seu algoritmo adaptativo, para extração de componentes, de APEX, o qual se caracteriza por ter dois tipos de conexões *feedforward* e lateral [49]. Segundo Edwar C Malthouse [66], este tipo de conexão apresenta inconvenientes quando a conexão lateral é suficientemente grande e se torna impraticável em dimensões elevadas [66], [67].

A análise de componentes principais não lineares (NLPCA), utilizando redes neurais autoassociativas, foi introduzida por Kramer (1991) [68] na área de engenharia química, apresentando uma estrutura de cinco camadas, desacoplável para calcular os componentes não lineares. A rede possui uma camada gargalo de nós para reduzir a dimensão das variáveis de entrada. A segunda e quarta camada da rede tem funções de ativação não linear. Porém, este tipo de PCA apresenta limitações, tal como uma baixa eficiência quando modela não linearidades entre variáveis. Isto é, esta técnica não pode modelar curvas ou superfícies que se interceptam e modelar parametrizações que tenham descontinuidades.

Outro tipo PCA é o Kernel PCA (KPCA) [54], que dilata por fatores exponenciais a matriz de covariância, entre outros modelos cujas variáveis latentes podem ser

alargadas. A implementação deste algoritmo se torna inviável quando aplicado em problemas de grandes dimensões, pois demandaria um custo encontrar os parâmetros necessários para sua funcionalidade.

5.6.3 Filtros Casados

Existem várias formas para executar o processo de detecção dos sinais de DPs. Além das técnicas tradicionais em filtragem de ruídos de fundo e dos recentes filtros *wavelets* [69], [70], para supressão de ruídos, tem-se utilizado parâmetros estatísticos. Entre as técnicas que utilizam estes parâmetros tem-se a média de pulsos utilizados em algoritmos de semi-integração dos sinais de DPs [6] e a máxima amplitude utilizada nos detectores de picos dos instrumentos de medição (ex.:ICM) [37]. Porém, estas técnicas não são robustas, visto que, há vestígios e possibilidades da existência de sinais de ruído com energia similar ou própria até maiores que os sinais de DPs. Assim o sinal de ruído é confundido com o sinal de DP, apesar do auxílio de filtros [36], [37].

Para contornar os problemas citados acima, é introduzido como caracterizador de sinais de DPs o Filtro Casado (MF). Este filtro é baseado na teoria dos sistemas lineares de filtragem ótima [71] e [51], e sua característica principal é que, utiliza informações estatísticas e leva em consideração a forma (*shape*) do sinal na detecção do mesmo. E com a ajuda de um limiar adaptativo, baseado em redes neurais, pode-se estruturar um caracterizador de DPs em dados coletados e, pode-se estender para detecção de sinais após consolidar o aprendizado da rede neural. Nos seguintes parágrafos introduzem-se alguns conceitos importantes de MF.

Detecção de ruídos impulsivos por MF

Para a detecção de um sinal tipo pulso $n(t)$ observado em um sinal $s(t)$ adicional, o sinal e pulso, pode ser modelado como:

$$v[t] = b[t]n[t] + s[t] \quad (5.25)$$

onde $[.]$ denota para o tempo discreto, $b[t]$ é um indicador binário que sinaliza a presença ou ausência de um ruído impulsivo. Utilizando o modelo da Equação (5.25), a detecção de um sinal de ruído impulsivo pode ser considerada como a estimação da condição do estado binário ($\hat{b}[t]$), do indicador de ruído $b[t]$. Os ruídos impulsivos podem ser detectados utilizando um modelo de ruído num filtro casado.

O produto interno de dois vetores de sinais fornece uma medida de similaridade dos sinais. Já que a filtragem é basicamente uma operação de produto interno, isso

indica que, a saída de um filtro deve fornecer uma medida de similaridade do filtro de entrada e sua resposta impulsional. O método clássico para a detecção de um sinal consiste na utilização de um filtro cuja resposta impulsional é casada com a forma do sinal a ser detectado. O princípio de um filtro MF, para a detecção de um pulso $n[t]$, é baseado na maximização da amplitude de saída do filtro, quando a entrada contém o pulso $n[t]$. O MF para a detecção de um pulso $n[t]$, observado num sinal $x[t]$ de fundo é definido como:

$$H(f) = K \frac{N^*(f)}{P_{xx}(f)} \quad (5.26)$$

onde $P_{xx}(f)$ é a potência espectral de $x[t]$ e $N^*(f)$ é o complexo conjugado do espectro do ruído impulsivo. Quando o sinal $x[t]$ é um sinal decorrelacionado com média igual a zero e variância σ_x^2 , o MF para a detecção de ruídos impulsivos $s[t]$ se torna:

$$H(f) = \frac{K}{\sigma_x^2} N^*(f) \quad (5.27)$$

A resposta impulsional do MF correspondente à Equação (5.27) é dada por

$$h[t] = Cn[T-t] \quad (5.28)$$

onde C é o fator de escala dado por $C=K/\sigma_x^2$. A saída de MF é denotada por $z[t]$. A resposta à entrada de ruído pulsante, o filtro de saída é dado pela relação de convolução.

$$z[t] = Cn[-t]*n[t] \quad (5.29)$$

onde o ‘*’ denota convolução. No domínio da frequência a Equação (5.29) se torna em $Z(f) = N(f)H(f) = c|N(f)|^2$. A saída, $z[t]$, do MF é passado por um limiar não-linear, onde é chaveada (decisão binária) a presença ou ausência de um ruído impulsivo.

$$\hat{b}[t] = \begin{cases} 1 & \text{se } |z[t]| \geq \text{Limiar} \\ 0 & \text{Outros} \end{cases} \quad (5.30)$$

Na Equação (5.30), quando a saída do MF excede um limiar, o detector sinaliza a presença do sinal na entrada. A Figura 5.1 mostra um detector de pulsos composto de M diferentes MFs. Ou seja, o detector de sinais em presença ou ausência de um ruído pulsante. Se um pulso está presente, então a informação adicional fornece um tipo de pulso, a máxima correlação cruzada entre a entrada e o modelo de ruído impulsivo, e um atraso pode ser utilizado para alinhar o sinal de entrada (ruído) e o modelo do ruído.

Esta informação pode ser utilizada para subtrair o ruído impulsivo a partir do sinal ruidoso.

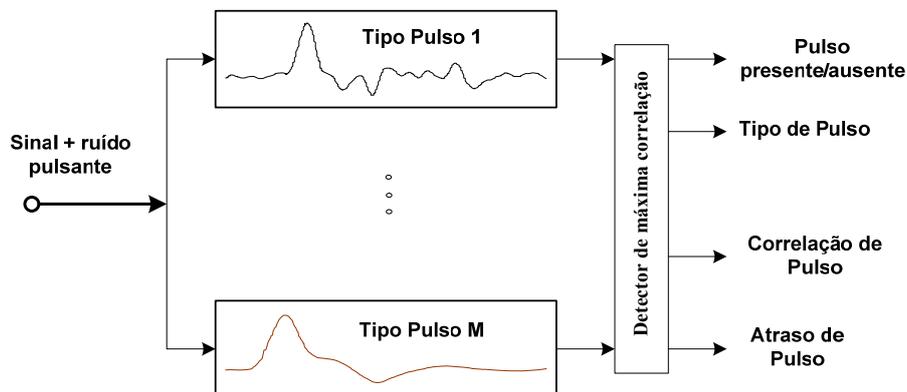


Figura 5.1 – Banco de filtros casados para detecção de ruídos pulsantes [72]

5.6.4 Análise de Componentes Independentes

Separação Cega de Fontes (BSS) e a Análise de Componentes Independentes (ICA) são técnicas emergentes que têm por objetivo recuperar os sinais ou “fontes” não observadas a partir dos dados coletados denominado de “observações” ou misturas observadas, explorando apenas algumas assunções, tal como a independência mútua entre os sinais. O adjetivo Cego (*Blind*) coloca em ênfase o fato que: i) as fontes de sinais não são observadas e ii) não há informação disponível com respeito à mistura. Esta carência, de haver um conhecimento a priori com respeito à mistura, é compensada pela informação estatística de ordem elevada. As vulnerabilidades destas assunções fazem da BSS uma aproximação útil para processamento dos sinais de DPs.

O ICA é baseado na teoria estatística, tendo como função de contraste as estatísticas de ordens superiores (cumulantes), Kurtosis e funções critério. Um dos métodos importantes de ICA, o ICA não-holonômico, é baseado na análise de eventos extremos nas bases de matrizes reais (*manifold*).

A divergência de Kullback-Leibler, conhecida como Entropia Relativa, tem uso na separação cega de fontes, servindo como função-objetivo que ajuda a procurar a independência estatística entre os dados.

Nesta parte da tese descrevem-se brevemente as técnicas de separação cega de fontes e a teoria básica da análise de componentes independentes. Posteriormente se descreve, de maneira sucinta, três dos mais importantes métodos ICA: Infomax, FastICA e JADE. Estes foram os algoritmos ICA utilizados, no Capítulo 6, como métodos comparativos para caracterizar os sinais DPs. Finalmente, descrevem-se

algoritmos de separação cega quando as fontes de sinais de DPs estão contaminadas por sinais fontes de ruídos.

5.6.5 Separação Cega de Fontes

O modelo simples BSS, segundo a Figura 5.2, assume a existência de n sinais independentes $s_1(t), \dots, s_n(t)$ e a observação das tais misturas $x_1(t), \dots, x_m(t)$. Estas misturas são de tipo linear e instantâneo, isto é, $x_i(t) = \sum_{j=1}^n a_{ij}s_j(t)$ para cada $i = 1, \dots, m$. Esta pode ser representada na forma compacta pela equação de mistura,

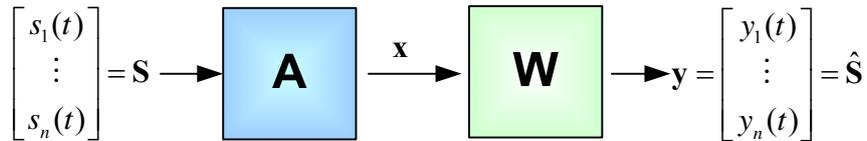


Figura 5.2 – Modelo básico BSS sem ruído. Sinais não observáveis: **S**, observações: **X**, sinais fonte estimadas: **y**

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) \quad (5.31)$$

onde o índice “ t ” foi substituído por “ k ”, de forma que represente um índice generalizado do sinal contínuo (t) ou discreto $[\cdot]$. Assim, $\mathbf{s}(k) = [s_1(k), \dots, s_n(k)]^T$ é uma coleção de sinais fonte, $\mathbf{x}(k)$ os m sinais observados e a “matriz de mistura” $\mathbf{A}_{n \times m}$ contém os $n \times m$ coeficientes da mistura. O alvo do BSS consiste em recuperar os vetores fonte $\mathbf{s}(k)$ utilizando apenas os dados observados $\mathbf{x}(k)$, a assunção de independência entre fontes do vetor de entrada $\mathbf{s}(k)$ e, possivelmente, algum conhecimento *a priori* com respeito à distribuição de probabilidade das entradas. O BSS pode ser formulado como a estimação da “matriz de separação” $\mathbf{W}_{m \times n}$, sendo a saída

$$\mathbf{y}(k) = \mathbf{W}\mathbf{x}(k), \quad (5.32)$$

uma estimativa dos sinais fonte.

O modelo básico do BSS pode ser estendido considerando-se, por exemplo, mais sensores que fontes e misturas ruidosas. Este, conforme a Figura 5.3, considera a componente $\mathbf{v}(k)$ como a coleção de sinais ruído que têm assunções de independência com o vetor de entrada $\mathbf{s}(k)$. As relações entre fontes, observações e ruídos se apresentam na Equação (5.33).

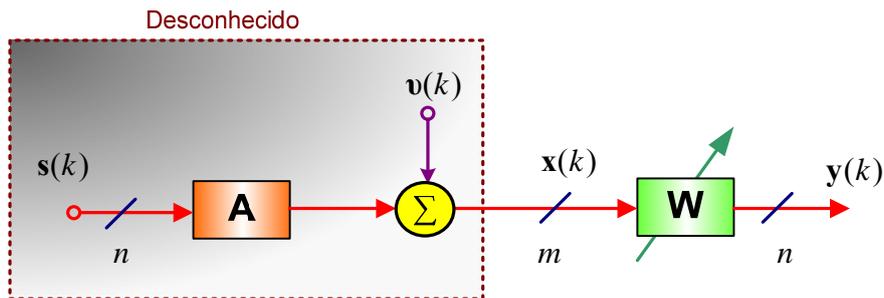


Figura 5.3 – Modelo básico do BSS instantâneo linear com componente de ruído [53]

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{v}(k), \quad \mathbf{y}(k+1) = \mathbf{W}_k \mathbf{x}(k+1) \quad (5.33)$$

Vários princípios têm sido desenvolvidos em estatísticas, computação neural e processamentos de sinais para resolver o problema BSS.

Os métodos ICA, básicos, podem ser divididos em duas principais categorias [73].

- *Métodos de segunda ordem.* Estes métodos tentam buscar o modelo de separação “matriz de separação” utilizando apenas a informação contida na matriz de covariância do vetor de dados \mathbf{x} . As técnicas fundamentais mais conhecidas e empregadas nesta categoria são PCA [49] e FA [64]. Pode-se dizer que, a finalidade dos métodos de segunda ordem é buscar uma representação “confiável” dos dados no sentido de que o erro de reconstrução dos dados, erro médio quadrático, atinja seu valor mínimo possível.

- *Métodos de ordem superior.* Neste caso, é também considerado que a informação das distribuições de probabilidade relativas às fontes não estejam contidas na matriz de covariância. A distribuição de \mathbf{x} não deve ser assumida como gaussiana, visto que, se esse fosse o caso, toda a informação relativa a \mathbf{x} estaria contida na matriz de covariância. Isso é somente verdadeiro para variáveis gaussianas que tenham média zero. Dessa forma a utilização das estatísticas de ordens superiores pode não ser útil. Exemplos deste tipo de métodos são *projection pursuit* [55], [61], [74], [75] e *redundancy reduction* (introduzido por Barlow 1961).

Para estimar o modelo de dados do ICA, o procedimento comum é a formulação de uma “função-objetivo” F com o intuito de maximizá-la ou minimizá-la de acordo com o enfoque de projeto. Assim, um algoritmo ICA pode ser decomposto em duas partes: A definição de sua função-objetivo ou “função contraste” F e o algoritmo de otimização utilizado para maximizar/minimizar F . As propriedades de um método ICA dependem dessas duas partes. As propriedades estatísticas (ex.: consistência, variância

assintótica, robustez) dependem da escolha da função-objetivo, enquanto as propriedades algorítmicas (ex.: velocidade de convergência, requerimentos de memória, etc.) dependem do algoritmo de otimização. Funções de contraste típicos em ICA e algoritmos são descritos na secção 5.7.

Identificabilidade e ambigüidade do modelo ICA

ICA, em princípio, explora a diversidade espacial, considerando que sensores distintos recebam diferentes misturas das fontes originais. Desse modo, a aproximação ICA para separação de fontes procura pela estrutura ótima dos sinais através da informação dos sensores e não por meio do tempo. Uma consequência de ignorar a estrutura temporal dos sinais observados é que a informação contida nos dados seja inteiramente representada pela distribuição amostral do vetor de observações \mathbf{x} . Então, o BSS se torna um problema de identificação das distribuições de probabilidades das observações $P(\mathbf{x} = \mathbf{W}\mathbf{s})$, dada uma distribuição amostral das fontes. Assim, o modelo ICA estatístico tem dois componentes: A matriz de misturas \mathbf{A} e as distribuições de probabilidade das fontes \mathbf{s} . Para a identificabilidade do modelo ICA sem ruído, é possível considerar algumas restrições:

- *Independência mútua das fontes.* Se uma das fontes $i=1, \dots, n$ tem uma função de densidade de probabilidade (pdf), denotada $q_i(\cdot)$, a assunção de independência tem uma simples expressão matemática: a pdf $q(\mathbf{s})$ do vetor fonte \mathbf{s} é

$$q(\mathbf{s}) = q_1(s_1) \times \dots \times q_n(s_n) = \prod_{i=1}^n q_i(s_i), \quad (5.34)$$

isto é, ela é produto das densidades (marginal) de todas as fontes.

- *Todas as componentes independentes \mathbf{s}_i , com a possível exceção de uma componente, devem ser não gaussianas.* Para variáveis aleatórias gaussianas simples não-correlacionadas isto implica independência e assim qualquer representação decorrelacionada pode dar componentes independentes. Porém, se mais de uma das componentes \mathbf{s}_i é gaussiana, ainda assim é possível identificar os componentes independentes não gaussianas, bem como as colunas correspondentes da matriz de mistura.

- O número de misturas lineares observadas \mathbf{m} deve ser igual ou maior que o número de componentes independentes \mathbf{n} , isto é, $\mathbf{m} \geq \mathbf{n}$. Esta restrição não é totalmente necessária e pode ser superada utilizando ICA com bases supercompletas. [53].

- *As colunas da matriz \mathbf{A} devem ser de posto completo.* Isto é, estas colunas devem ser linearmente independentes, de modo que a matriz \mathbf{A} seja invertível.

Se \mathbf{x} e \mathbf{s} são interpretados como processos estocásticos, surgem restrições adicionais, o que inclui a necessidade do processo ser estacionário no sentido estrito. Outras restrições, como a ergodicidade, são também requeridas [50], [52]. Se o processo é estatisticamente independente (i.i.d) sobre o tempo, aqueles requerimentos são cumpridos. Dessa forma, pode-se considerar o processo estocástico como uma variável randômica.

Indeterminações no modelo ICA

- *Não é possível determinar as variâncias (energias) das componentes independentes.* A razão é que, o efeito de qualquer constante que multiplica um componente pode ser cancelado pela divisão da correspondente coluna da matriz \mathbf{A} pela mesma constante.

- *Devido à indeterminação anterior, não é possível ter uma ordem para as componentes independentes.* No entanto, podem-se utilizar as normas das colunas da matriz de mistura, que forneçam as contribuições das componentes independentes para as variâncias das observações, para ordenar \mathbf{s}_i segundo a norma descendente da correspondente coluna de \mathbf{A} .

Funções Objetivo em ICA

Podem-se diferenciar dois tipos de funções-objetivo, dependendo de como são estimados os componentes independentes. A função de contraste *Multi-unit* estima todas as componentes independentes ao mesmo tempo, enquanto que, a função de contraste *One-unit* permite a estimação de uma simples componente independente (procedimentos que são repetidos para buscarem vários componentes). Entre estas se podem citar a Negentropia e os cumulantes de ordens elevadas.

- *Máxima Semelhança e infomax.* O princípio da máxima semelhança (ML) conduz a vários contrastes que são expressos via a convergência de Kullback-Leibler, definida para duas distribuições de probabilidade $f(s)$ e $g(s)$ como:

$$\mathbf{K}(f|g) \triangleq \int_s f(s) \log \left(\frac{f(s)}{g(s)} \right) ds, \quad (5.35)$$

e que pode ser entendido como uma métrica estatística de distância que quantifica a proximidade das duas distribuições.

A função de contraste logaritmo da máxima semelhança (LML ou LMS), em modelos ICA sem ruído, pode ser formulada como:

$$\phi_{ML}[\mathbf{y}] = \frac{1}{T} \sum_{t=1}^T \log p(x(t) | A, q) \quad (5.36)$$

onde $\mathbf{x}(t)$ é o vetor de observações na realização t , \mathbf{A} é a matriz de separação e q é a distribuição do vetor fonte \mathbf{s} . Simples cálculos mostram que [76], [77]:

$$\phi_{ML}[\mathbf{y}] \xrightarrow{T \rightarrow \infty} \mathbf{K}[\mathbf{y}|\mathbf{s}] + \text{constante} \quad (5.37)$$

E assim, o princípio ML tem associado uma função de contraste $\phi_{ML} = \mathbf{K}[\mathbf{y}|\mathbf{s}]$, isto é, ML tenta buscar uma matriz \mathbf{A} tal que a distribuição de $\mathbf{A}^{-1}\mathbf{x}$ seja a mais próximo possível (no sentido da divergência de Kullback) para a distribuição em hipótese das fontes.

O princípio do infomax propõe uma função contraste que maximize a entropia das componentes independentes [63]:

$$\phi_{IM}[\mathbf{y}] \triangleq -\mathbf{H}[g(\mathbf{y})] \quad (5.38)$$

onde $\mathbf{H}[\cdot]$ denota a entropia de Shannon [73], [76]. No caso, pode ser mostrado que $\phi_{IM}[\mathbf{y}] = \phi_{ML}[\mathbf{y}]$ e o infomax é equivalente ao critério de máxima semelhança [63].

- *Informação Mútua.* A simples aproximação de semelhança descrita acima é baseada em uma hipótese pré-determinada referente à distribuição dos sensores. Espera-se bons resultados da ML unicamente se as distribuições em hipótese não se diferenciam muito das verdadeiras. Para problemas supercompletos, deve-se minimizar a divergência $\mathbf{K}[\mathbf{y}|\mathbf{s}]$ não só com respeito a \mathbf{A} (via a distribuição de $\mathbf{y} = (\mathbf{A}^{-1}\mathbf{x})$) mas também com respeito à distribuição de \mathbf{s} . Se nós denotamos $\tilde{\mathbf{y}}$ como um vetor randômico de entradas independentes e cada entrada distribuído como a correspondente entrada de \mathbf{y} então:

$$\mathbf{K}[\mathbf{y}|\mathbf{s}] = \mathbf{K}[\mathbf{y}|\tilde{\mathbf{y}}] + \mathbf{K}[\tilde{\mathbf{y}}|\mathbf{y}], \quad (5.39)$$

para qualquer vetor \mathbf{s} com entradas independentes. Logo, a tarefa de minimização pode ser concluída pela minimização dos termos do lado direito da Equação (5.39). O primeiro termo é independente de \mathbf{s} e, desse modo, a minimização em \mathbf{s} é equivalente a minimizar $\mathbf{K}[\tilde{\mathbf{y}}|\mathbf{s}]$, que é feito simplesmente fixando-se $\mathbf{s} = \mathbf{y}$ para que ajustes de

$\mathbf{K}[\tilde{\mathbf{y}}|\mathbf{s}] = 0$ assim que $\min_{\mathbf{s}} \mathbf{K}[\mathbf{y}|\mathbf{s}] = \mathbf{K}[\mathbf{y}|\tilde{\mathbf{y}}]$ finalmente considerando que $\min_{(\mathbf{s}, \mathbf{y})} \mathbf{K}[\mathbf{y}|\mathbf{s}] = \min_{\mathbf{y}} \mathbf{K}[\mathbf{y}|\tilde{\mathbf{y}}]$, isto é, deve-se minimizar a função contraste:

$$\phi_{MI}[\mathbf{y}] \triangleq \mathbf{K}[\mathbf{y}|\tilde{\mathbf{y}}] \quad (5.40)$$

A divergência de Kullback $\mathbf{K}[\mathbf{y}|\tilde{\mathbf{y}}]$ entre a distribuição e a distribuição próxima com entradas independentes é tradicionalmente chamada de informação mútua e pode também ser expressa como [73], [77], [78], [79]:

$$MI(y_1, \dots, y_n) = \sum_{i=1}^n H[y_i] - \mathbf{H}[\mathbf{y}], \quad (5.41)$$

onde $\mathbf{H}[\cdot]$ é a entropia de Shannon. É fácil de se verificar que a informação mútua satisfaz $\phi_{MI}[\mathbf{y}] \geq 0$ com igualdade se e só se \mathbf{y} é distribuído como $\tilde{\mathbf{y}}$, isto é, se as entradas de \mathbf{y} são independentes. Assim, a informação mútua pode ser entendida como uma medida quantitativa de independência associada ao princípio da máxima semelhança. O problema principal com a informação mútua é que ela é difícil de se estimar, visto que ela é baseada na entropia, que demanda estimar as funções de densidade das observações y_i .

- *Negentropia*. A negentropia é definida como uma *one-unit* de contraste como [73]:

$$J(\mathbf{y}) = H(\mathbf{y}_{gaussiano}) - H(\mathbf{y}), \quad (5.42)$$

onde $\mathbf{y}_{gaussiano}$ é um vetor aleatório gaussiano com matriz de covariância igual a das observações \mathbf{y} . Por definição $J(\mathbf{y}) \geq 0$, com $J(\mathbf{y}) = 0$ se e só se \mathbf{y} tem uma distribuição gaussiana. Mostra-se que se as misturas y_i são não correlacionadas, a informação mútua pode ser expressa como [73]:

$$MI(y_1, \dots, y_n) = J(\mathbf{y}) - \sum_{i=1}^n J(y_i), \quad (5.43)$$

e podem-se buscar as direções de máxima negentropia, ou seja, direções onde os elementos da soma $J(y_i)$ são maximizados. Isto se equivale a buscar uma representação em que a informação mútua é minimizada. Desafortunadamente, as restrições feitas com respeito à informação mútua são aqui também válidas. A estimação da negentropia é difícil e assim ela não é muito prática como função de contraste.

- *Aproximações ordens superiores.* A principal desvantagem das funções de contraste, derivadas a partir da aproximação da ML é que elas requerem a estimação das distribuições de probabilidade. Uma possível solução para este problema é a utilização das estatísticas de ordens elevadas para definir funções de contraste que são simples aproximações daquelas derivadas do critério ML. A forma mais fácil de expressar a informação de ordem elevada é utilizar cumulantes. Para observações randômicas de média zero y_i, y_j, y_k, y_l , o cumulante de segunda ordem pode ser expresso como [73]

$$C_{ij}[\mathbf{y}] = E[y_i y_j], \quad (5.44)$$

e o cumulante de quarta ordem:

$$C_{ijkl}[\mathbf{y}] = E[y_i y_j y_k y_l] - E[y_i y_j]E[y_k y_l] - E[y_i y_k]E[y_j y_l] - E[y_i y_l]E[y_j y_k] \quad (5.45)$$

Uma medida próxima de incompatibilidade entre a distribuição da saída e a distribuição da fonte pode ser definida a partir da incompatibilidade quadrática dos cumulantes.

$$\begin{aligned} \phi_2[\mathbf{y}] &\triangleq \sum_{ij} (C_{ij}[\mathbf{y}] - C_{ij}[\mathbf{s}])^2 \\ \phi_4[\mathbf{y}] &\triangleq \sum_{ijkl} (C_{ijkl}[\mathbf{y}] - C_{ijkl}[\mathbf{s}])^2 \end{aligned}$$

Utilizando ϕ_2 e ϕ_4 , se \mathbf{s} e \mathbf{y} estão simetricamente distribuídas com distribuições suficientemente próximas à normal, então é possível aproximar a divergência de Kullback por

$$\mathbf{K}[\mathbf{y}|\mathbf{s}] \approx \phi_{24}[\mathbf{y}] \triangleq \frac{1}{48} (12\phi_2[\mathbf{y}] - \phi_4[\mathbf{y}]) \quad (5.46)$$

A *kurtosis* é também utilizada em algoritmos ICA como medida de não-gaussianidade das estimações das componentes independentes em y_i . Ela pode ser definida utilizando cumulantes como:

$$k_i \triangleq C_{iiii} = E[y_i^4] - 3(E[y_i^2])^2 \quad (5.47)$$

Também produzido a partir de cumulantes é a função de contraste do algoritmo JADE, que é baseado em um subconjunto de cumulantes cruzados:

$$\phi_{JADE}[\mathbf{y}] \triangleq \sum_{ijkl \neq ijkk} C_{ijkl}^2[\mathbf{y}] \quad (5.48)$$

Uma revisão dos mais importantes contrastes de elevadas ordens para ICA pode ser encontrada [73], [76], [80].

5.6.6 Algoritmos ICA

Depois de escolher uma apropriada função de contraste, faz necessário um método prático ou algoritmo para sua implementação. Nesta secção da dissertação, explicam-se brevemente três algoritmos bem conhecidos: Infomax, FastICA e JADE.

Alguns passos de pré-processamento são comuns para os algoritmos ICA:

- *Centralização*. A média dos dados é subtraída a partir dos dados reais de tal forma que a média dos dados centrados seja zero. Isto é, $\mathbf{x}_c = \mathbf{x} - E[\mathbf{x}]$. Após a estimação da matriz de misturas \mathbf{A} , a média é restituída para os dados.

- *Branqueamento ou sphering*. Uma transformação linear é aplicada aos dados \mathbf{x} , de modo que a matriz de covariância dos dados transformados \mathbf{x}_w seja igual à identidade, $E[\mathbf{x}_w \mathbf{x}_w^T] = \mathbf{I}$. Esta transformação sempre é possível, utilizando FA ou PCA/SVD, que decompõe os autovalores da matriz de covariância $E[\mathbf{x} \mathbf{x}^T] = \mathbf{E} \mathbf{D} \mathbf{E}^T$ para transformar os dados observados de acordo com:

$$\mathbf{x}_w = \mathbf{E} \mathbf{D}^{1/2} \mathbf{E}^T \mathbf{x}, \quad (5.49)$$

onde \mathbf{E} é a matriz ortogonal dos autovalores da matriz de covariância dos dados, \mathbf{D} é a matriz diagonal associada aos autovalores $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ e $\mathbf{D}^{1/2} = \text{diag}(d_1^{-1/2}, \dots, d_m^{-1/2})$.

- *Redução da dimensionalidade*. Quando os dados são branqueados, nesse passo, também pode ser realizada a compactação dos dados, descartando-se os autovalores da matriz de covariância que sejam muito pequenos e não possuam características inerentes. Isto é feito pela PCA. Assim, reduzindo a dimensão dos dados, pode-se ajudar na supressão dos ruídos e evitar um sobreaprendizado do algoritmo ICA.

Algoritmo do Infomax

Um dos primeiros algoritmos desenvolvidos para ICA é o chamado algoritmo Infomax, baseado na maximização da entropia da rede, que é, sobre algumas condições, equivalente à maximização da semelhança. Geralmente, estes algoritmos são baseados no gradiente ascendente da função objetivo. O algoritmo Infomax, original de Bell e Sejnowski [63], [81], utiliza um gradiente estocástico que produz atualização dos pesos da forma:

$$\Delta \mathbf{B} \propto [\mathbf{B}^T]^{-1} - 2 \tanh(\mathbf{B} \mathbf{x}) \mathbf{x}^T \quad (5.50)$$

Esta função trabalha para a estimação da maior componente independente super-Gaussiana, porém, para componentes sub-Gaussianas, outras funções devem ser utilizadas [82]. O principal inconveniente do gradiente estocástico é sua lenta convergência.

Para impor a velocidade de convergência e simplificar o algoritmo, pode-se utilizar o método do gradiente natural de Amari [53], [57]. Este produz um algoritmo da forma:

$$\Delta \mathbf{B} \propto \left(\mathbf{I} - 2 \tanh(\mathbf{B}\mathbf{x})(\mathbf{B}\mathbf{x})^T \right) \mathbf{B} \quad (5.51)$$

Depois desta modificação, o algoritmo não necessita de branqueamento.

Algoritmo do FastICA

Algoritmos adaptativos semelhantes ao Infomax podem ser problemáticos quando são utilizados em um ambiente onde a adaptação não é necessária. A convergência destes algoritmos é usualmente baixa e depende crucialmente da escolha da taxa de aprendizado em cada passo do processo de treinamento ICA. Para contornar o problema, podem ser utilizados algoritmos em batelada (batch), baseados na iteração de ponto-fixo [78].

O algoritmo FastICA é um destes algoritmos de ponto-fixo. Ele foi originalmente introduzido utilizando a kurtosis e foi generalizado mais tarde [88] por funções de contrastes. Para dados branqueados, o algoritmo FastICA *one-unit* possui a seguinte forma:

$$\mathbf{w}(k) = E \left[\mathbf{a}g \left(\mathbf{a}(k-1)^T \mathbf{x} \right) \right] - E \left[g' \left(\mathbf{a}(k-1)^T \mathbf{x} \right) \right] \mathbf{a}(k-1), \quad (5.52)$$

onde o vetor peso é normalizado para a norma unitária após cada iteração e a função g é a derivada da função G utilizada na função geral de contraste dado por

$$J_G(y) = \left| E[G(y)] - E_\nu[G(\nu)] \right|^p \quad (5.53)$$

onde ν é uma variável gaussiana generalizada, y é assumido ser normalizado para uma variância unitária e o expoente p é tipicamente 1 e 2.

A velocidade de convergência dos algoritmos FastICA, ICA de ponto-fixo, é superior aos algoritmos adaptativos, tais como o Infomax. Fatores de velocidade de 10 e 100 são usualmente observados [82]. Outra vantagem do FastICA é que ele ainda pode estimar sem problemas componentes independentes subgaussianas e supergaussianas.

Além disso, ele é um algoritmo que pode ser otimizado para funções de contraste *multi-unit* ou *one-unit*.

Algoritmo do JADE

O JADE [81] é baseado na otimização de Jacobi [73] de uma função de contraste ortonormal, ao contrário à otimização pelos algoritmos por gradientes. O JADE pode ser resumido nos seguintes passos:

- *Inicialização*. Estima uma matriz de branqueamento \mathbf{W} e fixa $\mathbf{y} = \mathbf{W}\mathbf{x}$.
- *Forma estatística*. Estima um máximo conjunto $\{\hat{Q}_i^Z\}$ de matrizes cumulantes.

Dado um vetor randômico $X_{n \times 1}$ e uma matriz $M_{n \times n}$, a matriz cumulante associada $Q_x(M)$ é a matriz $n \times n$ definida em componentes pares por:

$$[Q_x(M)]_{ij} \triangleq \sum_{k,l=1}^n C_{ijkl} M_{kl}, \quad (5.54)$$

quando a técnica de Jacobi é utilizada.

- *Otimize uma função contraste ortogonal*. Busca a matriz de rotação \hat{V} , de maneira que as matrizes cumulantes sejam os mais diagonais possíveis, tal que, resolva $\hat{V} = \operatorname{argmin} \sum_i \operatorname{Off}(V^T \hat{Q}_i^Z V)$ sendo $\operatorname{Off}(\mathbf{F})$ a soma dos quadrados dos elementos fora da diagonal de uma matriz \mathbf{F} , isto é, $\operatorname{Off}(\mathbf{F}) \triangleq \sum_{i \neq j} (f_{ij})^2$.

- *Separado*. Estima \mathbf{A} como $\hat{A} = \hat{V}\hat{W}^{-1}$ e/ou estima os componentes como $S = \hat{A}^{-1}X = \hat{V}^T Z$.

O JADE tem mostrado um desempenho muito eficiente em dimensões menores. Porém, em grandes dimensões, os requerimentos de memória podem ser proibitivos, porque as matrizes cumulantes devem ser guardadas em memória, o que requer $O(m^4)$ unidades de memória. Outra desvantagem é que o JADE, como todos os outros algoritmos Jacobi, tende a ser muito complicado para se codificar, requerendo sofisticadas manipulações matriciais.

5.6.7 Extensão de Algoritmos ICA – Noisy_ICA

Depois de revisar meramente alguns algoritmos tradicionais dos métodos ICA, observa-se que eles têm bom desempenho quando os sinais não têm contaminação por ruído. Ou seja, estes algoritmos estão definidos na forma básica do BSS, onde assumem-se fontes livres de ruído e, se há ruídos, estes são bem comportados e/ou conhecidos. Assim, pode-se fazer um pré-processamento dos dados e/ou fazer alterações das funções de contraste ou conhecer alguns parâmetros do ruído.

Contudo, faz necessário um algoritmo que contemple dados reais, como é o caso dos sinais das DPs. Nesta parte do documento, introduzem-se algoritmos ICA recentes que lidam com dados reais, quando o ruído se encontra presente na fonte, nos sensores ou nos modelos imprecisos.

5.6.8 ICA para dados contaminados por ruído

As bases do modelo ICA básico é adaptado para os casos onde o ruído está presente. Considera-se que o ruído é assumido como componente adicional, sendo que, essa assunção é utilizada em análise de fatores e em processamento de sinais [2], [53], [73] e permite apresentar uma formulação do modelo ruidoso. Segundo a Equação (5.33) tem-se que:

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{v}(k) \quad (5.55)$$

onde $\mathbf{v}(k) = [\nu_1(k), \dots, \nu_m(k)]$ é o vetor de ruído que, em alguns casos típicos, pode ser assumido como:

- O ruído é independente dos componentes independentes.
- O ruído é gaussiano.
- A matriz da covariância do ruído Σ em muitos casos é assumida como a matriz $\sigma^2\mathbf{I}$. Em alguns casos, o ruído é assumido como conhecido.

A identificabilidade da matriz de misturas em um modelo noisy_ICA é garantida sobre as mesmas restrições do ICA básico, independência e não-gaussianidade. No entanto, as realizações dos componentes independentes s_i podem-se tornar de difícil identificação, visto que, eles não podem ser separados completamente dos ruídos.

Ruído de Sensor frente a Ruído de Fonte

Nos casos típicos em que a covariância do ruído é assumida como $\sigma^2\mathbf{I}$, o ruído na Equação (5.55), pode ser considerado como “ruído devido ao sensor”. Isto é, as variáveis do ruído são adicionadas de forma separada em cada sensor, compondo a

variável observada x_i . Porém, no caso em que o ruído é adicionado aos componentes independentes, chamado de “ruído devido à fonte”, o modelo é levemente diferente do antecessor,

$$\mathbf{x} = \mathbf{A}(\mathbf{s} + \mathbf{v}), \quad (5.56)$$

sendo também a covariância do ruído diagonal. Na realidade, se podem considerar os componentes independentes ruidosos como $\tilde{s}_i = s_i + v_i$ e o novo modelo será:

$$\mathbf{x} = \mathbf{A}(\tilde{\mathbf{s}}) \quad (5.57)$$

Observa-se que este é justamente o modelo básico ICA, com os componentes independentes modificados. Assim podem-se aproveitar as suposições do ICA básico que ainda são válidos, visto que os componentes independentes de \tilde{s}_i são não-gaussianas e independentes. Portanto, pode-se estimar o modelo da Equação (5.57) pelo ICA básico. Dessa maneira, podem-se estimar a matriz de mistura e os componentes independentes ruidosos.

Para se estimar os componentes independentes originais a partir das fontes ruidosas identificadas, é necessário ter suposições adicionais. Assumindo-se que a covariância do ruído tem a forma $\Sigma = \mathbf{A}\mathbf{A}^T\sigma^2$, o vetor ruído pode ser transformado para $\tilde{\mathbf{v}} = \mathbf{A}^{-1}\mathbf{v}$, que pode ser chamado ruído equivalente devido à fonte. Assim a Equação (5.55) torna-se em:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{A}\tilde{\mathbf{v}} = \mathbf{A}(\mathbf{s} + \tilde{\mathbf{v}}) \quad (5.58)$$

Desse modo que a covariância de $\tilde{\mathbf{v}}$ seja $\sigma^2\mathbf{I}$ e, assim as componentes transformadas $\mathbf{s} + \tilde{\mathbf{v}}$ serão independentes. Finalmente, a matriz \mathbf{A} poderá ser estimada pelo ICA básico.

Existem poucos métodos para estimar um noisy_ICA. A estimação de modelo sem ruídos parece ser uma tarefa desafiante para ele mesmo e, assim, o ruído é usualmente esquecido para fins de se obter resultados simples e tratáveis.

Na seguinte seção se faz um resumo de algumas técnicas para estimação de matriz de mistura e dos componentes independentes. Entre estas, a remoção do Bias e os métodos por cumulantes de altas ordens, já que, estas técnicas podem ter aplicação no noisy_ICA

Remoção do Bias

As técnicas para remoção do bias podem ser utilizadas em um ICA modificado, de modo que o bias devido a ruído seja removido ou, ao menos, reduzido.

Segundo [73], utiliza-se $\mathbf{v} = \mathbf{A}\mathbf{s}$ para dados livre de ruídos, de modo que sejam procuradas as projeções em $\mathbf{w}^T \mathbf{v}$ para se obter estimativas consistentes dos componentes independentes quando se tem uma boa medida de não-gaussianidade, visto que, o modelo é imune aos ruídos gaussianos. Assim, esta aproximação pode ser utilizada no noisy_ICA, na forma, $\mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{v} + \mathbf{w}^T \mathbf{v}$. A idéia consiste em medir a não-gaussianidade de $\mathbf{w}^T \mathbf{v}$ a partir dos dados observados $\mathbf{w}^T \mathbf{x}$, a fim de que as medidas não sejam afetadas pelo ruído $\mathbf{w}^T \mathbf{v}$.

Kurtosis para remoção de bias. Sabe-se que a kurtosis de um sinal de ruído gaussiano tem valor zero. Então, a kurtosis pode ser uma boa medida de não-gaussianidade. Logo $kur(\mathbf{w}^T \mathbf{x}) = kur(\mathbf{w}^T \mathbf{v}) + \underbrace{kur(\mathbf{w}^T \mathbf{v})}_0$. Porém a kurtosis tem algumas desvantagens como: é sensível a *outliers* ou intrusos, depende da dimensão das observações e do tipo de simetria da função estatística, quando se pretende estimá-la [73].

Medidas de não gaussianidade geral. Outra técnica, melhor que a kurtosis, para este fim é:

$$\mathbf{J}_G(\mathbf{w}^T \mathbf{v}) = \left[E\{\mathbf{G}(\mathbf{w}^T \mathbf{v})\} - E\{\mathbf{G}(\mathcal{G})\} \right]^2 \quad (5.59)$$

onde a função \mathbf{G} é uma função quadrática regular e v é uma variável gaussiana. Denota-se z como a variável randômica e n como uma variável gaussiana de variância σ^2 . Segundo [73], a idéia é expressar uma relação entre $E\{\mathbf{G}(z)\}$ e $E\{\mathbf{G}(z+n)\}$ em simples termos algébricos. Porém, esta relação é complicada e só é calculada utilizando integração numérica. No entanto, para certas escolhas de \mathbf{G} esta relação se torna simples [73]. A idéia básica é escolher \mathbf{G} como função de densidade de uma variável randômica gaussiana ou uma função parecida à função gaussiana, chamada de momentos gaussianos, sendo estes não polinomiais, isto é a função de densidade gaussiana com variância c^2 :

$$\varphi_c(x) = \frac{1}{c} \varphi\left(\frac{x}{c}\right) = \frac{1}{\sqrt{2\pi c}} \exp\left(-\frac{x^2}{2c^2}\right) \quad (5.60)$$

onde $\varphi_c^k(x)$ é a k-ésima derivada de $\varphi_c(x)$.

Máxima semelhança. Esta aproximação é voltada para a estimação da matriz de mistura com dados ruidosos, maximizando a semelhança conjunta e as realizações dos componentes principais:

$$\log \mathbf{L}(\mathbf{A}, \mathbf{s}(1), \dots, \mathbf{s}(T)) = -\sum_{t=1}^T \left[\frac{1}{2} \|\mathbf{A}\mathbf{s}(t) - \mathbf{x}(t)\|_{\Sigma^{-1}}^2 + \sum_{i=1}^n f_i(s_i(t)) \right] + C \quad (5.61)$$

onde $\mathbf{A}\mathbf{s}(t) - \mathbf{x}(t) \triangleq \mathbf{m}$, então $\|\mathbf{m}\|_{\Sigma^{-1}}^2$ é definido como $\mathbf{m}^T \Sigma^{-1} \mathbf{m}$, $\mathbf{s}(t)$ são realizações de componentes independentes, C é uma constante e f_i são os logaritmos das funções de densidade das componentes independentes. Porém, esta aproximação é computacionalmente custosa [73].

Estimação das componentes independentes por funções shrinkage

Estimar a matriz de mistura não é suficiente para o modelo noisy_ICA. Faz-se necessário também estimar os componentes principais. Para este fim, a partir da Equação (5.55), temos:

$$\mathbf{W}\mathbf{x} = \mathbf{s} + \mathbf{W}\nu \quad (5.62)$$

onde unicamente se tomam as estimativas ruidosas dos componentes independentes. Isto é, para obter estimativas das componentes independentes \hat{s}_i , que são de alguma forma ótimos ou contém um mínimo de ruídos, pode ser realizada uma simples aproximação para este problema, utilizando o estimador de máxima a posteriori (MAP), onde se toma as \hat{s}_i cujos valores maximizam a semelhança conjunta (Equação 5.61) que é chamado estimador de máxima semelhança (ML). Tomando o gradiente da Equação (5.60) com respeito a $s(t)$ e igualando-a a zero, tem-se:

$$\hat{\mathbf{A}}^T \Sigma^{-1} \hat{\mathbf{A}} \hat{\mathbf{s}}(t) - \hat{\mathbf{A}}^T \Sigma^{-1} \hat{\mathbf{A}} \hat{\mathbf{x}}(t) + f'(\hat{s}(t)) = 0 \quad (5.63)$$

onde a derivada do logaritmo da densidade f' é aplicada separadamente em cada componente do vetor $\hat{\mathbf{s}}(t)$.

De fato este método conduz à generalização não linear do clássico Filtro Wiener [92]. Uma aproximação alternativa é utilizar os componentes independentes na estrutura do tempo para remoção de ruídos.

Contudo, \hat{s} não é uma tarefa fácil de resolver, quando utilizado otimização numérica. Porém, um caso simples e especial pode ser adotado, quando se assume a covariância do ruído como $\Sigma = \mathbf{A}\mathbf{A}^T\sigma^2$. Logo, a Equação (5.63) será equivalente à Equação (5.58), caso com ruído na fonte. Substituindo a covariância do ruído em 5.63 temos:

$$\begin{aligned} \hat{s}(t) - \hat{A}^{-1}x(t) + \sigma^2 f'(\hat{s}(t)) = 0 &\rightarrow \hat{A}^{-1}x(t) = \hat{s}(t) + \sigma^2 f'(\hat{s}(t)) \\ \hat{s} = g(\hat{A}^{-1}x) &\end{aligned} \quad (5.64)$$

Para um componente escalar da função g é obtido invertendo a relação

$$g^{-1}(u) = u + \sigma^2 f'(u) \quad (5.65)$$

Assim, o MAP é obtido invertendo uma função f' ou função contraste da densidade de probabilidade de s . Para variáveis não-gaussianas a função contraste é não-linear $g(u)$.

Capítulo 6

Caracterizadores: Implementação e Resultados

Neste capítulo, se faz a implementação dos caracterizadores dos sinais DPs e eliminação de ruídos. O desempenho dos métodos é mostrado sobre um conjunto de fontes constituídas pelos modelos experimentais descrito no Capítulo 3. Além disso, as técnicas caracterizadoras foram adaptadas para tratamento dos sinais de DPs. Assim, é possível obter uma melhor adequação entre Informação e Técnica com a finalidade de levar em consideração as informações *a priori* inerentes às DPs e aos modelos experimentais.

Os caracterizadores podem ser aplicáveis para outros casos de medição, para monitoração on-line e fazerem parte de um diagnóstico integrado de equipamentos de subestações de energia elétrica, uma vez que, estes caracterizadores não são específicos para sinais de DPs, podendo processar sinais aleatórios, instantâneos ou dependentes do tempo.

6.1 PCA para Identificar Sinais de DPs em Transformadores de Potência

Nesta seção se faz uma análise de pré-processamento dos sinais de DPs, que são possíveis de serem encontradas em transformadores de potência. Os dados são baseados em medições de DPs provenientes de três células experimentais (descritos no Capítulo 3), as quais sintetizam fontes de sinais DPs em diversas partes do equipamento.

O método de pré-processamento usa PCA com o propósito de reduzir a dimensão dos dados originais, além de reduzir as incertezas da informação, obtendo, como resultado, uma melhora no desempenho dos sistemas de reconhecimento de padrões. As perdas de informação produzidas pelo PCA são intrínsecas para os sistemas de reconhecimento de padrões [49], [3].

As células FL, PP e PT geram os mapas característicos (amostras binárias com resolução inicial de 256^2) em primeira instância, com ajuda do instrumento de medição de DPs, sendo em seguida convenientemente reduzidas e convertidas para a base

6.1 PCA para Identificar Sinais de DPs em Transformadores de Potência

decimal formando uma matriz $[R]$ de 32×32 . Esta é representada na forma de vetor amostra com 1024 elementos, de maneira que um conjunto destes foi escolhido e representado como $X(t)$, cuja média foi compensada para que atinja um valor nulo (zero) [3].

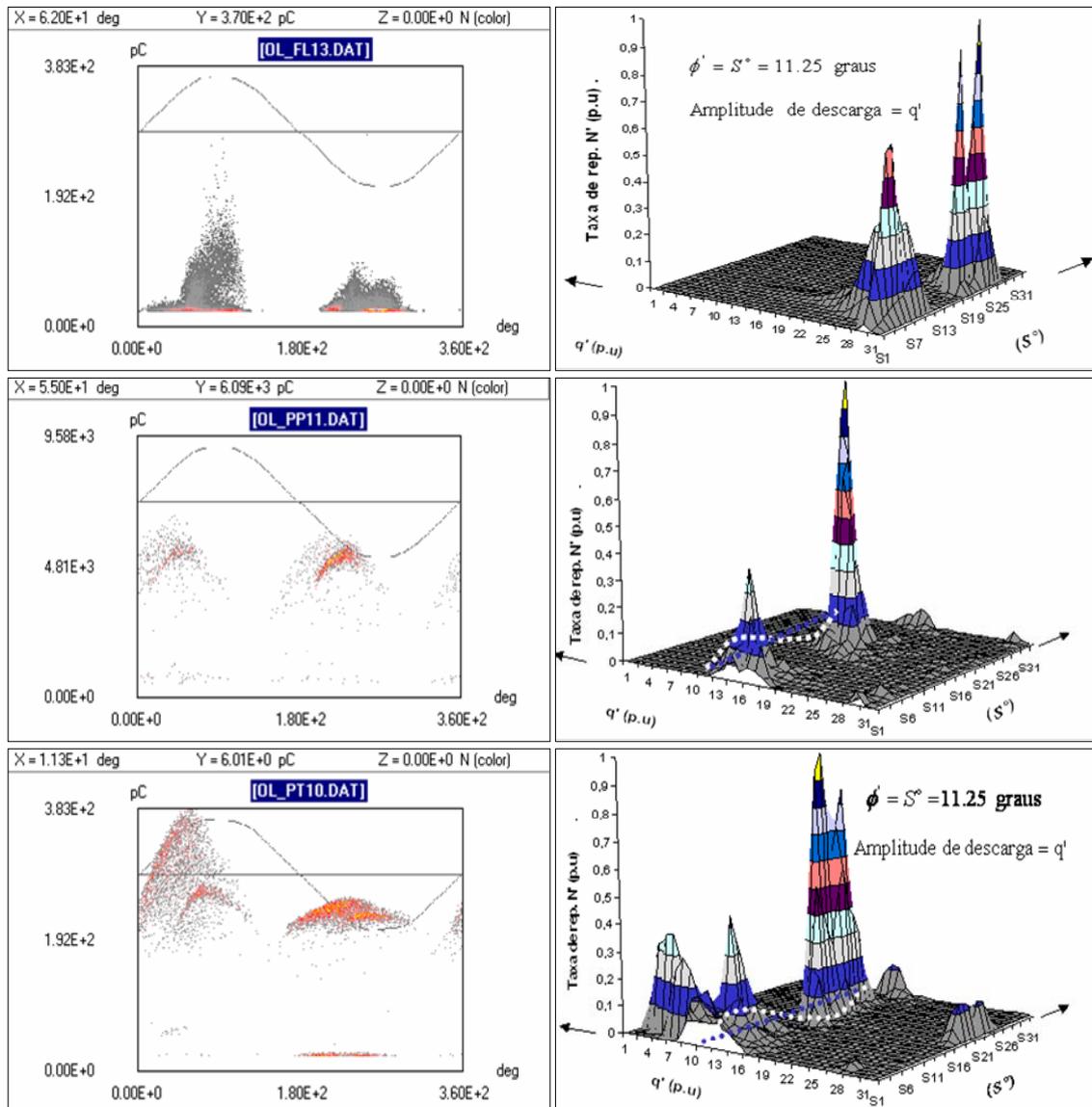


Figura 6.1 – Amostras de mapas característicos dos padrões FL, PP e PT originais (a esquerda) e reduzidos (à direita).

Um classificador neural faz a validação do PCA e mede seu desempenho. Um conjunto de dados disponíveis foi dividido de forma a compor os conjuntos de caracterização e validação de acordo com a Tabela 6.1. O critério adotado para fazer a escolha do conjunto de validação obedece a correlação cruzada, balanço de energia dos sinais e, principalmente, o conhecimento do especialista (técnico experiente em DPs), tem maior peso nesta escolha. A Figura 6.1 mostra mapas característicos primitivos e suas respectivas reduções onde a taxa de repetição N é reproduzível facilmente em 3-D.

A figura mostra os três padrões antes e depois de uma redução em resolução, explicado na seção 5.1.

Tabela 6.1 – Conjuntos de caracterização e validação

Conjunto de dados	Conjunto de Caracterização	Conjunto de Validação	Total amostras
Padrão teste FL	20	15	35
Padrão teste PP	26	29	55
Padrão teste PT	15	10	25

O padrão PP apresentou eventos com alto nível de ruídos e baixa taxa de descargas, escolhendo-se apenas 26 mais representativos para caracterização.

O padrão PT tem poucas amostras devido à dificuldade de medir sinais com baixa densidade de descargas e pelo modelo físico da célula teste sujeita às descargas francas.

Considerando $X(t)$ como processo estocástico de dimensão M igual a 1024, de valor esperado nulo e a matriz de correlação $\mathbf{R} = E[\mathbf{X}'\mathbf{X}]$, a equação característica [73], $\mathbf{R}q_i = \lambda_i q_i$ representa a análise de componentes principais, cujas direções (q_i) são as dos autovetores de \mathbf{R} , que têm λ_i s para autovalores.

6.1.1 Componentes Principais

A análise de componentes principais foi desenvolvida no ambiente de programação MatLab V6.5, gerando-se 1024 componentes principais (vetores ortonormais) com a mesma dimensão dos padrões e, 1024 autovalores (diagonal da matriz de correlação), que são equivalentes às variâncias apresentadas em cada componente. Foram empregados 61 vetores eventos (conjunto de caracterização) para gerar a matriz de covariância. Isto implica que, do total de componentes principais e autovalores, apenas as 60 primeiras direções são significativas, devido à perda de um grau de liberdade quando a média é compensada para zero.

Para verificar a ortogonalidade das componentes, usou-se o produto interno entre todos os vetores componentes. Obteve-se uma matriz de produtos internos com elementos da diagonal unitários (representando um ângulo de 90 graus) e fora da diagonal, valores nulos, ou seja, uma completa ortogonalidade entre os componentes.

A reconstrução do sinal original pode ser obtida a partir das projeções deste nos componentes principais extraídos. Um dos objetivos desta análise é reduzir a dimensão

do espaço original de entrada, projetando-o neste subespaço de menor dimensão. Para isso, procura-se uma representação efetiva dos sinais, descartando projeções que contenham variância reduzida e retendo apenas termos que tenham variância significativa. Admite-se cometer um erro de aproximação, que pode ser calculado pela diferença entre os sinais original e aproximado, formando-se o vetor erro a partir dos autovalores não considerados na aproximação. Nas Figura 6.2 e 6.3, os componentes são ordenados através dos autovalores, observando-se que uns poucos componentes são responsáveis por uma fração significativa da energia dos sinais.

]

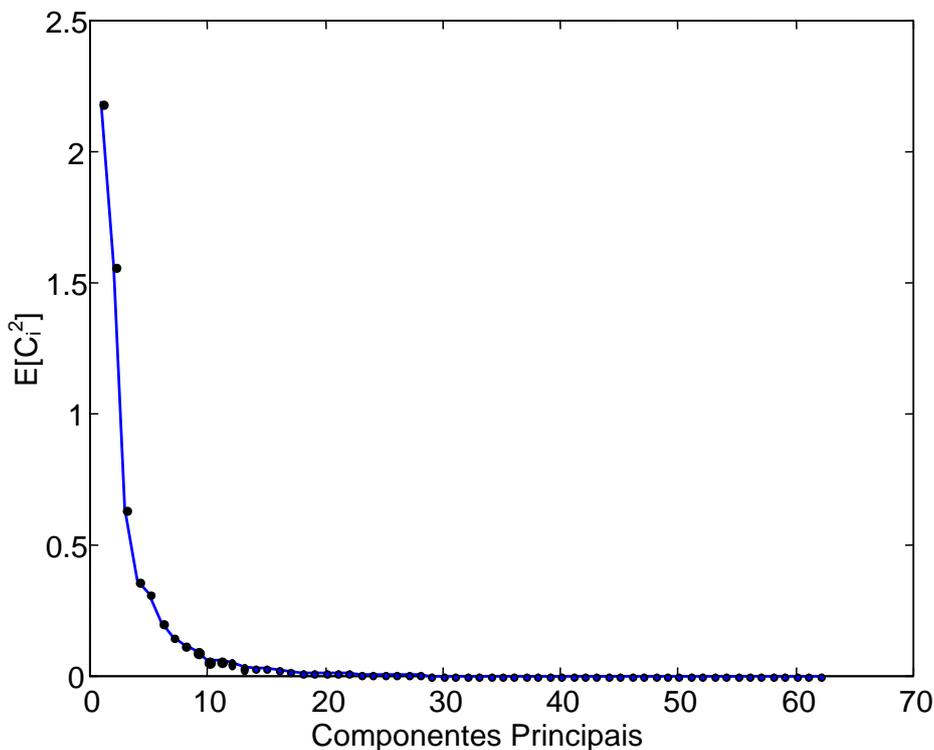


Figura 6.2 - As primeiras componentes apresentam as maiores energias $E[C^2]$. maiores energias $E[C^2]$.

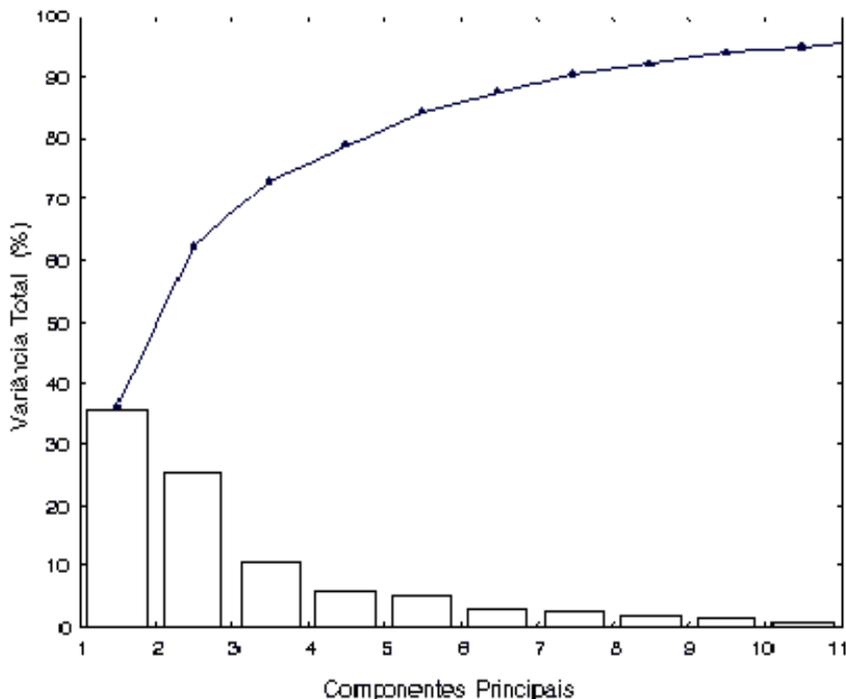


Figura 6.3 - As primeiras componentes apresentam as maiores porcentagens das variâncias

A Figura 6.3 indica que o primeiro componente representa 35% da variância total, e com os sete maiores componentes consegue-se explicar aproximadamente 92 % da variância total. Isto quer dizer que, com sete componentes, tem-se uma boa representação dos dados de DPs.

6.1.2 Mapeamento dos Componentes Principais

Podemos observar a estrutura dos componentes principais, de forma a identificar se esta estrutura é capaz de determinar as características relevantes de um dado padrão. Uma vez que cada autovetor representa um componente principal com módulo unitário, determinando assim a projeção de um dado vetor de entrada nesse componente, a observação destes componentes pode descobrir informações aparentemente ocultas. A análise de componentes principais não determina toda a informação intrínseca dos dados, apenas nos informa o que é mais representativo.

A Figura 6.4 mostra os nove primeiros componentes principais (PCs) da matriz de covariância. Observa-se que cada um destes componentes é diferente tanto em forma como em amplitude. Os primeiros três componentes praticamente têm características próprias de cada tipo de padrão. Associando os mapas característicos originais, mostrados na Figura 6.1, com os primeiros componentes principais, conclui-se que o primeiro componente tem a forma do padrão FL, enquanto que PC2 e PC7 apresentam

6.1 PCA para Identificar Sinais de DPs em Transformadores de Potência

quatro distribuições bem diferentes das outras e guardam características mais relevantes do padrão PP (ver Figura 6.1). O terceiro componente, por sua vez, tem características do padrão PT, reconhecido pelos dois picos negativos e as duas distribuições desenvolvidas na região média do plano horizontal. Nos outros componentes, não se percebe características associativas a algum padrão, apresentando diferentes amplitudes e formas. Entretanto, o classificador poderá extrair características sutis destas estruturas, revelando cada padrão de DPs com eficiência.

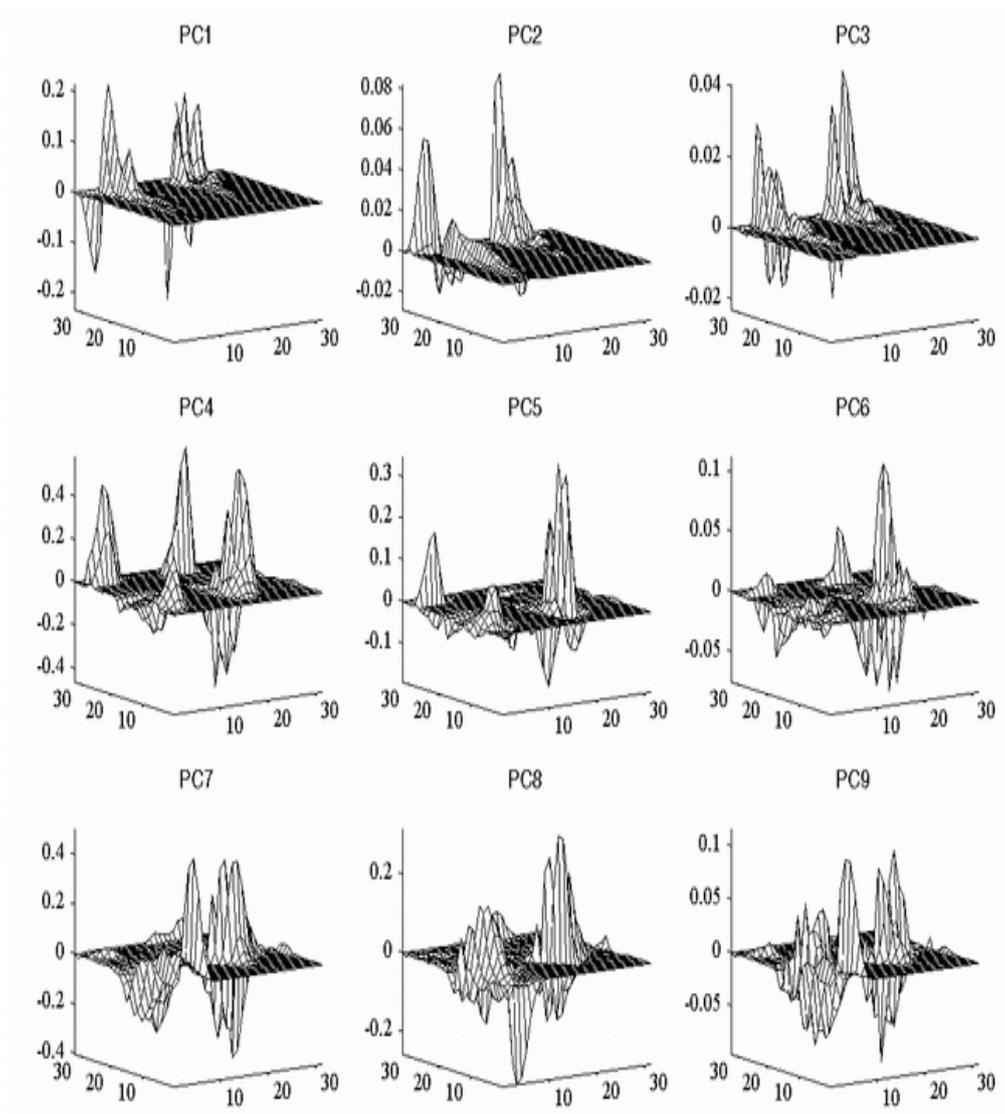


Figura 6.4 - Os nove primeiros componentes apresentam maiores porcentagens das variâncias.

A Figura 6.5 mostra os nove componentes principais seguintes onde pode apreciar-se que PC10-PC12 apresentam semelhanças do padrão PP. As outras não podem ser identificadas visualmente. Também se observa que as amplitudes destas projeções decrescem rapidamente (fator aproximado de 1/10).

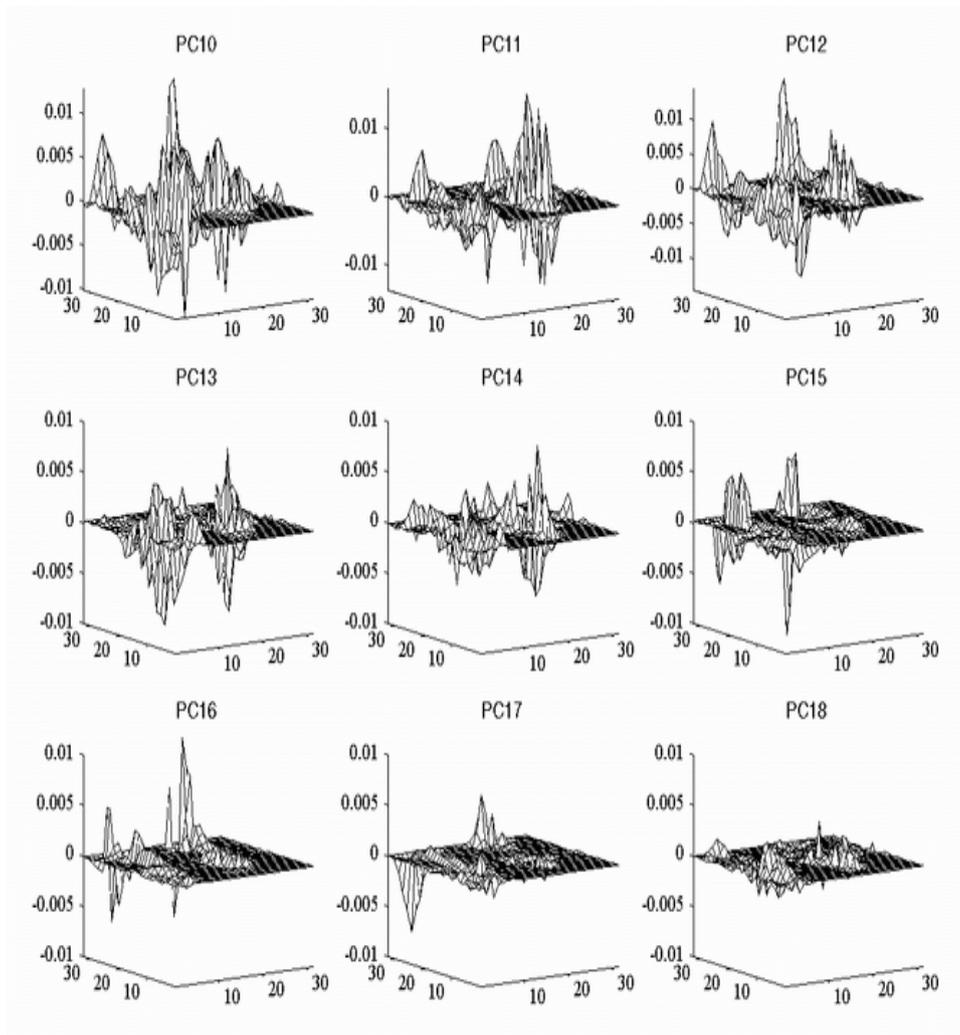


Figura 6.5 - Componentes principais do PC10 ao PC18.

6.1.3 Classificação Neural

Com base nos componentes principais extraídos, desenvolveu-se um sistema classificador neural de DPs.

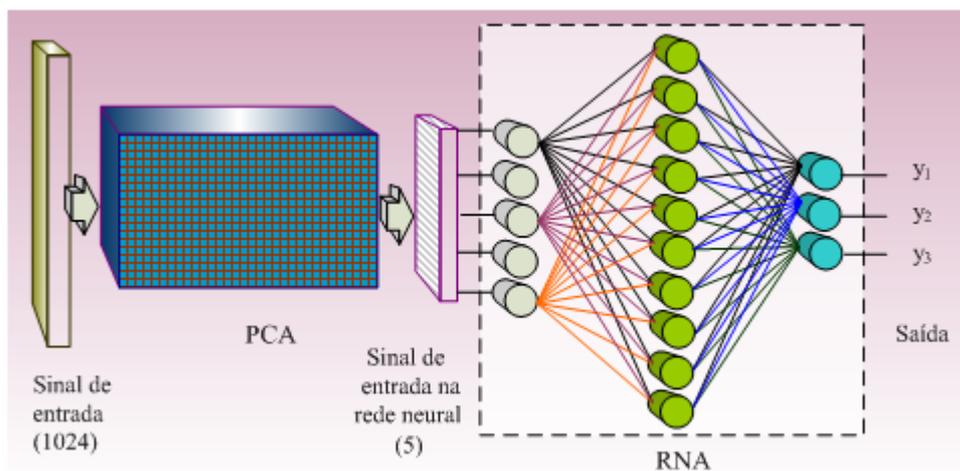


Figura 6.6 - Esquema do classificador baseado em CPs e RN.

A Figura 6.6 apresenta a topologia utilizada. Observa-se que um vetor original é projetado (inicialmente, em cinco componentes principais) e estas projeções são processadas pela rede neural, que apresenta cinco nós na primeira camada, dez neurônios na camada oculta e três neurônios na camada de saída. Desta maneira, cada neurônio de saída se encontra associado a uma classe de DPs. O critério de máxima probabilidade [52], [51] foi utilizado para identificar se um dado evento de entrada pertence a uma dada classe. O desempenho do classificador, para o conjunto de teste, pode ser visto na Tabela 6.2.

Tabela 6.2 - Eficiência de classificação para cada padrão usando-se cinco componentes principais

Padrões	Saída FL	Saída PP	Saída PT
Padrão teste FL (15 amostras)	67%	33%	0%
Padrão teste PP (29 amostras)	0	100%	0
Padrão teste PT (10 amostras)	0	40%	60%

Observa-se que o padrão FL apresenta 67% de eficiência (eficiência relativa), relativo às 15 amostras desse padrão, o padrão PP foi reconhecido em 100%, relativo às suas 29 amostras, enquanto que o padrão PT foi reconhecido em 60 %, relativo às 10 amostras de teste. Este resultado pode significar que um maior número de componentes principais deva ser considerado. O padrão FL apresenta cinco amostras ambíguas, que foram associadas ao padrão PP.

Na Tabela 6.3, se observa que a rede neural teve bastante sucesso na sua tentativa de classificação usando as projeções dos sete primeiros componentes. Neste caso, o padrão FL foi identificado em 93%, o padrão PP com 100% de eficiência, já demonstrada com cinco componentes, e o padrão PT com 70% de eficiência. Analisando a eficiência global da rede neural chega-se a 93% de eficiência total, relativa às cinquenta e quatro amostras de teste.

Em uma outra análise a eficiência total pode ser calculada como a média geométrica das eficiências parciais (utilizadas acima). Neste caso, a eficiência média da rede neural equivale a 87 % com 54 dados para teste.

Tabela 6.3 - Eficiência de classificação para cada padrão considerando-se sete componentes principais

Padrões	Saída FL	Saída PP	Saída PT
Padrão teste FL (15 amostras)	93%	7%	0
Padrão teste PP (29 amostras)	0	100%	0
Padrão teste PT (10 amostras)	0	30%	70%

Na medida em que aumentamos o número de componentes principais a serem utilizadas no sistema de reconhecimento não se consegue melhor desempenho. Observa-se, para nove componentes, uma eficiência total de 91%. Acrescentando-se mais componentes, a eficiência total começa a diminuir, assim como a eficiência média. Usando mais de 32 componentes a eficiência tem uma ligeira recuperação.

A Figura 6.7 mostra as eficiências na saída da rede neural classificadora, onde o número de elementos de entrada na rede é igual ao número de componentes utilizados na projeção dos dados de treino e de teste, máximo de até dezesseis neurônios na camada escondida e três neurônios na camada classificadora. O tipo de treinamento é simples, sendo retropropagado o erro médio quadrático. Portanto, para gerar cada ponto do gráfico foram realizados vários treinamentos partindo-se de condições iniciais diferentes e a quantidade de neurônios na camada escondida é variável (5 até 16).

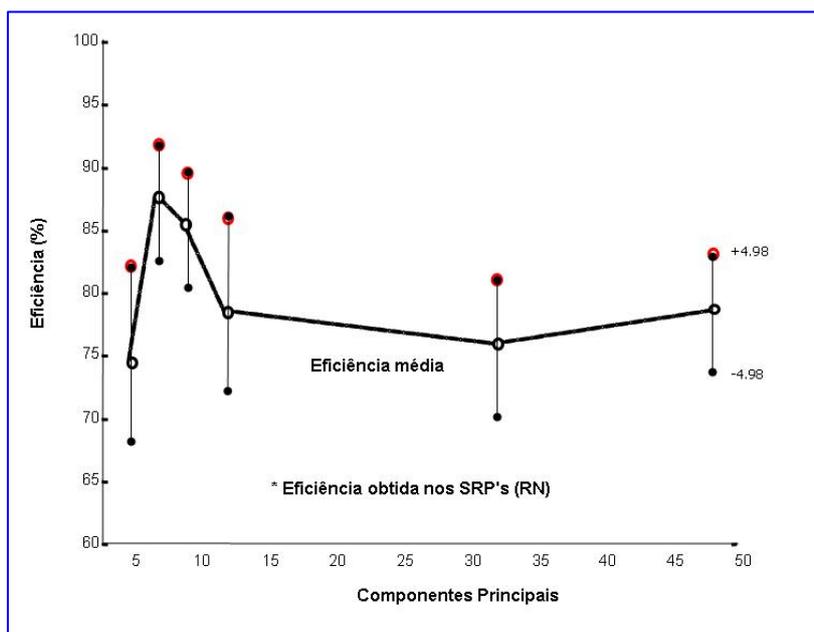


Figura 6.7 - Eficiência média de classificação em função do número de componentes principais.

Estes resultados revelam uma melhora significativa em relação aos obtidos em [32], onde os sinais foram tratados com suas dimensões originais, além de reduzir o custo computacional dos algoritmos de SRPs.

A eficiência anteriormente alcançada no reconhecimento dos padrões [32] foi de 83%, usando redes neurais simples. Enquanto a utilização do PCA atingiu uma eficiência de 100% no padrão PP (Tabela 6.3) e eficiência total de 93%.

6.1.4 Conclusões Referentes a PCA

O pré-processamento pela análise por componentes principais é uma ferramenta bastante útil e eficaz, que ajuda a revelar a informação mais importante que está aparentemente oculta. A eficiência total relativa ao conjunto de teste pode indicar qual a quantidade de componentes que serão necessários para ter uma boa representação, reduzindo drasticamente a dimensão do espaço original de entrada.

Utilizando-se apenas sete componentes principais se obtém bons resultados na classificação de padrões de DPs, obtendo-se uma eficiência total de 93%. Deve-se destacar que, para os dados analisados, os sete primeiros componentes principais correspondem a 92% das variâncias totais destes dados.

A análise por componentes principais feita sobre os dados de DPs, provenientes dos eletrodos imersos em óleo isolante de transformadores de potência, mostrou que estes tipos de dados podem ser significativamente compactados, melhorando os resultados obtidos em [32] e o desempenho dos SRPs.

6.2 Caracterizador por Filtros casados em Identificação DPs

Nesta seção, se constrói um sistema caracterizador híbrido para identificar DPs. Os dados são os mesmos provenientes das três células experimentais (descritos no Capítulo 3) e da Tabela 6.1.

6.2.1 Sistema Discriminador Híbrido de Sinais

Este sistema foi implementado na base de filtros casados (MF) e redes neurais artificiais (ANN), com a finalidade de se obter um discriminador cooperativo e robusto. Na primeira etapa, foram implementados filtros casados baseados na teoria dos sistemas lineares de filtragem ótima [49] e [52]. Na segunda etapa foi implementada uma ANN do tipo multicamadas e sem realimentação, treinada por algoritmo de *backpropagation* [49]. A escolha deste tipo de rede baseou-se em comparações realizadas em [32].

Os filtros casados são importantes na teoria de decisão, e usualmente são considerados dois casos: filtros casados com ruído aditivo branco e filtro casado generalizado para ruído aditivo colorido. No caso de ruído colorido, o MF inclui um filtro branqueador na entrada, que torna branco o ruído à sua saída. A figura de mérito da relação sinal-ruído R_0 é usada para avaliar o desempenho de tal sistema discriminador. Supõe-se que o sinal a detectar seja determinístico, $s(t)$ (o sinal de DPs foi aproximado como determinístico), e que se encontra mascarado pela presença de ruído do tipo aditivo, $n(t)$. As seguintes considerações norteiam o projeto dos filtros casados, assumindo o caso mais genérico de ruído colorido:

- Para o filtro branqueador de entrada, os sinais de entrada e saída são, respectivamente (ver Figura 6.8):

$$v(t) = s(t) + n(t), \quad (6.1)$$

$$v_0(t) = s_0(t) + n_0(t), \quad (6.2)$$

Como o MF inclui um filtro branqueador, $n_0(t)$ é um ruído branco. Além disso, neste caso de ruído colorido, o filtro deverá estar casado para $s_0(t)$, que passe a ser o sinal desejado.

- O sistema é linear e invariante no tempo, com resposta ao impulso $h(T-t)$, sendo T a duração do sinal desejado.

- O critério de otimização será:

$$R_0 = \frac{|s_0(t_1)|^2}{E\{n_0^2(t_1)\}} \quad (6.3)$$

A relação sinal-ruído da saída R_0 atingirá um máximo em algum instante t_1 . Onde $|s_0(t_1)|^2$ é a potência do sinal de saída e $E\{n_0^2(t)\}$ o valor médio quadrado de $n_0(t)$, usado como potência do ruído de saída. O sistema linear $h(t)$ que realiza a maximização do R_0 é chamado de "FILTRO CASADO", tendo uma resposta ao impulso (Figura 6.8).



Figura 6.8 - Filtro casado linear.

$$h[t] = s_0[T-t] \quad (6.4)$$

Ou seja, a resposta ao impulso para o filtro casado é a imagem espelhada do sinal desejado, deslocada da sua duração, por razões de causalidade [32]. Sendo o sinal de DPs de natureza estocástica, o MF tomará como um sinal determinístico e não cuidará das flutuações das descargas.

Neste trabalho, assumimos a hipótese simplificadora de que os sinais desejados sejam conhecidos e não aleatórios, visto que as flutuações dos sinais de DPs são reduzidas. Isto facilita o projeto de MF, que, então, segue a Equação (6.4). Assumimos ainda que tais sinais de DPs tenham sido distorcidos por ruídos no processo de coleta e armazenamento. Assume-se, inicialmente, que o ruído existente em DPs seja branco, o que elimina o projeto do filtro branqueador e faz $s_0(t) = s(t)$.

O filtro casado, geralmente, é usado para sinais no tempo. As medições de DPs, provenientes das células teste, não são amostradas no tempo, mas, pode-se considerar que, em cada cela da matriz $[\mathbf{R}]$, são armazenados sinais de DPs com três características fundamentais: ângulo de fase da tensão, amplitude e taxa de repetição (ϕ, q, N) expressas numericamente. Estas podem ser rearranjadas, sem considerar a relação de vizinhança, na forma de vetor, sem perder a referência ϕ , representando o sinal a detectar.

A aproximação do sinal determinístico é obtida calculando-se os sinais médios para as descargas do tipo FL, PP e PT para o conjunto de treinamento. Para detectar cada padrão de DP, projetam-se três filtros casados, um para cada classe de descarga parcial [32]. Em outras palavras, um sinal *array* (1024 componentes) será casado com a taxa de acumulação (N) média depositada em cada célula da matriz $[\mathbf{R}]$.

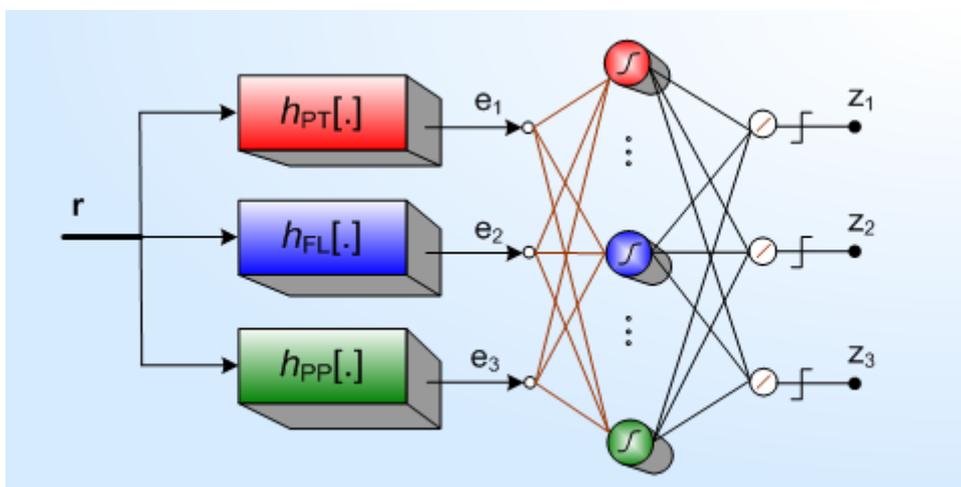


Figura 6.9 - Sistema discriminador híbrido: MF+ANN

O sistema discriminador usa três filtros casados, que funcionam como pré-processamento da informação, e uma rede neural, que realiza a tarefa final de discriminação. Esta rede neural é constituída de três neurônios na primeira camada e , dez neurônios na camada intermediária e três neurônios na camada de saída z (Ver Figura 6.9).

Os filtros têm resposta ao impulso $h_{FL}[\cdot]$ para eventos r de eletrodo FL, $h_{PP}[\cdot]$ para PP e $h_{PT}[\cdot]$ para PT. As respostas impulsivas são obtidas pela imagem espelhada do vetor médio de energia (com 1024 amostras cada) depositada em cada célula.

Em termos de normalização dos sinais de entrada, deve-se atender às exigências dos MFs, para os quais os sinais para os padrões de entrada devem estar normalizados em energia, e, por outro lado, deve-se atender aos requerimentos da ANN, que possui funções não lineares como ganho dos neurônios, que podem produzir saturação para os padrões de entrada. Neste sentido, devemos adotar um procedimento de normalização adequado e simples.

Sabe-se que a energia do sinal de padrões pode ser calculada de:

$$\mathbf{E} = \int_0^T [s(t)]^2 dt, \text{ ou } \mathbf{E} = \int_0^T s(t)h(T-t)dt \quad (6.5)$$

onde $h(t) = s(T-t)$, $0 < t < T$. Assumindo um sinal de entrada de energia \mathbf{E} e uma energia unitária à saída do filtro, o fator de normalização para cada padrão será:

$$K = \frac{1}{\sqrt{\mathbf{E}}} \quad (6.6)$$

onde, para padrões discretos:

$$\mathbf{E} = \sum_{i=1}^N s_0^2[i] \quad (6.7)$$

Desta maneira, a normalização em energia à saída do MF se adapta também à faixa dinâmica do classificador neural.

6.2.2 Resultados do Sistema Caracterizador Híbrido

O Sistema Discriminador avalia os sinais pré-processados para as células de teste, identificando a classe de padrão de DPs. Primeiramente, os eventos de teste são processados no nível máximo dos filtros casados quando $t = T$. A Figura 6.10 mostra as saídas dos filtros h_{FL} , h_{PP} e h_{PT} , quando os sinais de entrada são do tipo FL. Observa-

se que as respostas são notavelmente diferenciadas por cada filtro, de tal modo que um separador linear (limiar) pode ser suficiente para discriminar os eventos FL dos demais.

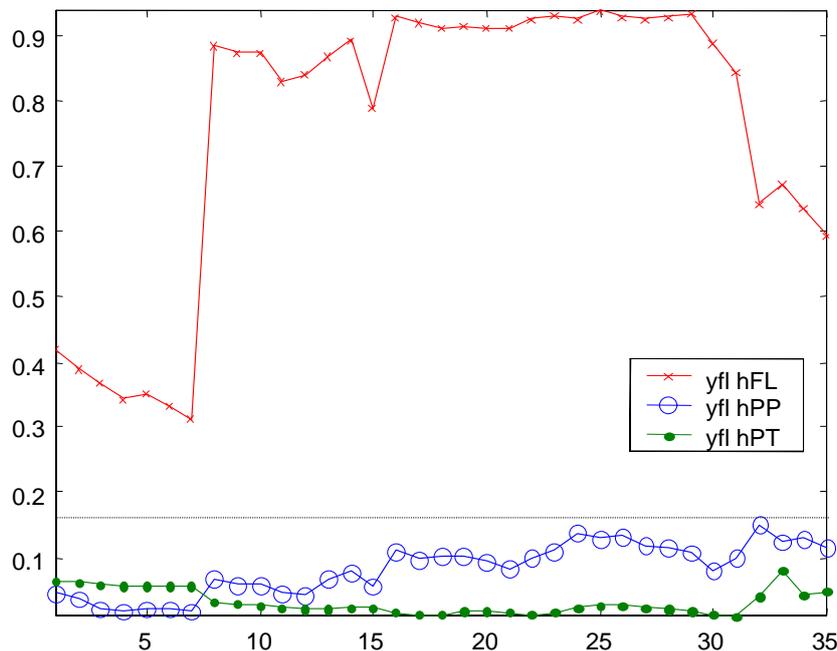


Figura 6.10 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é FL.

A Figura 6.11 e Figura 6.12 mostram as saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando os sinais de entrada são do tipo PP e PT, respectivamente. Nestes casos não podem ser implementados facilmente limiares lineares. Uma opção é subdividir em vários limiares, mas estes tornam complicado o problema de classificação de padrões de DPs, podendo resultar numa classificação ineficiente. Portanto, uma combinação neural das saídas dos filtros pode explorar, de forma não linear, as correlações existentes para as respostas casadas, otimizando o desempenho das classificações.

A Tabela 6.4 mostra as eficiências alcançadas pela rede neural, a partir das amostras instantâneas das saídas dos MFs, no instante $t=T$ (1024 amostras). Observa-se que, para os padrões FL e PP, atinge-se 100 % de eficiência, relativo aos 15 e 29 eventos disponíveis para estes, enquanto que o padrão PT foi reconhecido em 70 % dos casos, com relação aos 10 eventos de teste. Analisando a eficiência global do sistema caracterizador híbrido, chega-se a 97 % de eficiência total, relativa às 54 amostras de teste.

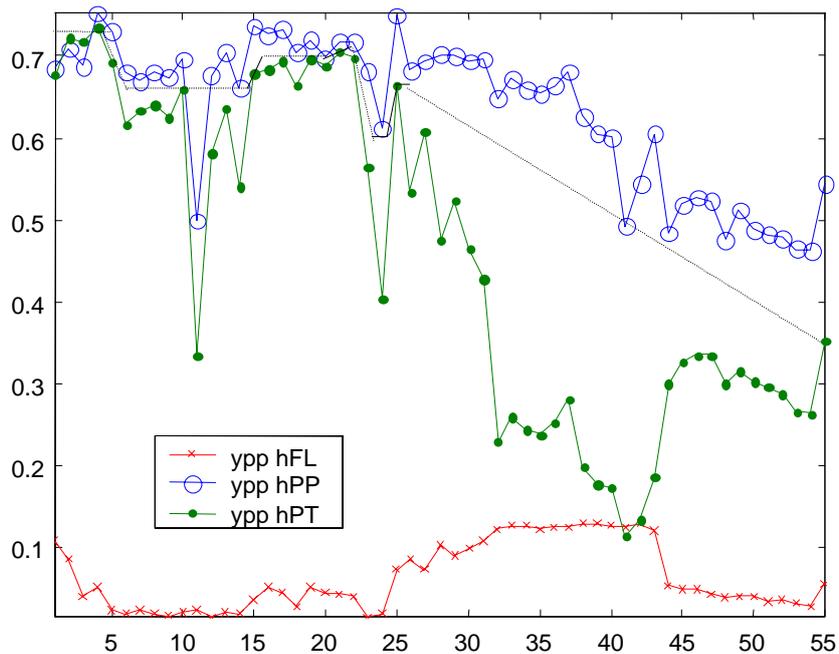


Figura 6.11 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é PP

Tabela 6.4 - Eficiência parcial alcançada por cada padrão na saída da rede neural.

Padrões	Saída FL	Saída PP	Saída PT
Teste FL (15)	100 %	0 %	0 %
Teste PP (29)	0 %	100 %	0 %
Teste PT (10)	0 %	30 %	70 %

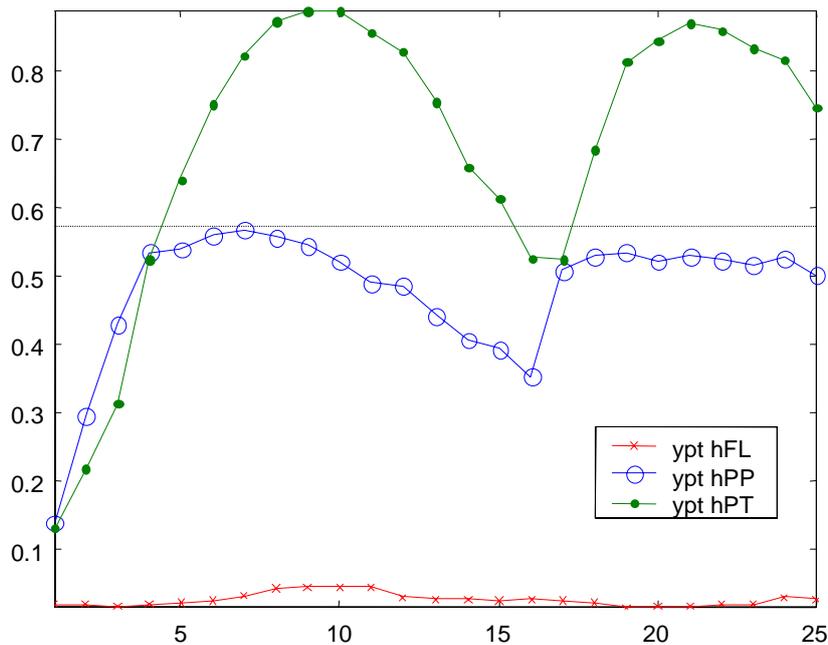


Figura 6.12 - Saídas dos filtros h_{FL} , h_{PP} e h_{PT} quando o sinal de entrada é PT.

Tabela 6.5 - Comparação de desempenhos entre as técnicas de reconhecimento padrão (PCA+ANN) e o discriminador híbrido (MF+ANN)

Padrões	PCA+ANN	MF+ANN
FL (15)	93 %	100 %
PP (29)	100 %	100 %
PT (10)	70 %	70 %
Total	93 %	97 %

Na Tabela 6.5, são mostrados os resultados das duas técnicas apresentadas para o tratamento dos sinais de DPs. Observa-se que o sistema (MF+ANN) identificou uma amostra a mais dos padrões de DPs, em relação ao sistema de reconhecimento de padrões (PCA+ANN). As diferenças nas eficiências totais (97 % para 93%) observadas não são conclusivas, haja vista as incertezas estatísticas presentes no número reduzido de amostras experimentais. Portanto, admite-se que ambas as técnicas tiveram desempenhos compatíveis.

6.2.3 Conclusões referentes ao sistema MF/ANN

Este sistema discriminador pode ser utilizado para avaliar a qualidade e caracterização dos modelos experimentais em DPs. Com base na amostra experimental recolhida, uma eficiência de 97 % na identificação das DPs foi obtida. Este discriminador híbrido provou, desta maneira, ter um desempenho mais compacto,

tratando o sinal de DPs em duas etapas: determinístico (branqueamento pelo MF) e estocástico (tratamento de flutuações pela ANN) resultando em estatísticas do sinal desejado.

6.3 Caracterizador ICA para Extração Cega de Ruídos

Nesta seção, realiza-se a remoção de ruídos em sinais fonte de DPs usando os recentes e consagrados conceitos BSS/ICA [80], [83], [73].

É mostrado que os métodos ICA podem ser aplicados na caracterização dos sinais de DPs, fornecendo uma solução satisfatória no que se refere a separação e/ou eliminação de ruídos dos sinais fonte. Estes sinais são fontes de defeitos procedentes de enrolamentos e barra estatórica envelhecida, entre outros, utilizados na representação de defeitos em sistemas de isolamento. Esses defeitos são possíveis de serem encontrados em hidrogeradores e geradores. No Capítulo 3 se mostram os modelos de barras e, a fontes avaliadas se descrevem na Tabela 6.6.

O Centro de Pesquisas de Energia Elétrica (CEPEL) analisou oito fontes de sinais de DP provenientes dos ensaios e medições rotineiras em geradores da usina Itaipu [35]. Destas, quatro são fontes características (Tabela 6.6'*) e as outras representam são diversas combinações no posicionamento dos capacitores de acoplamento (ver Figura 6.13.), nível de tensão aplicada, faixas de frequência de medição e posicionamento das barras defeituosas ao longo da fase C.

Tabela 6.6 - Fontes de Sinais de DPs

Fonte	Defeito (em 10,5kV)
S1	Sistema: capacitor de acoplamento/ barra. 150nF, conj.siemens + 9nF + barra + 2uF
S2	Barra 1 (slot)*
S3	Barra 2 (corona)*
S4	Barra 2 (corona na cabeça da bobina)
S5	Barra 3 (delaminação)*
S6	tiristor fase B*
S7	tiristor fase C
S8	tiristor neutro

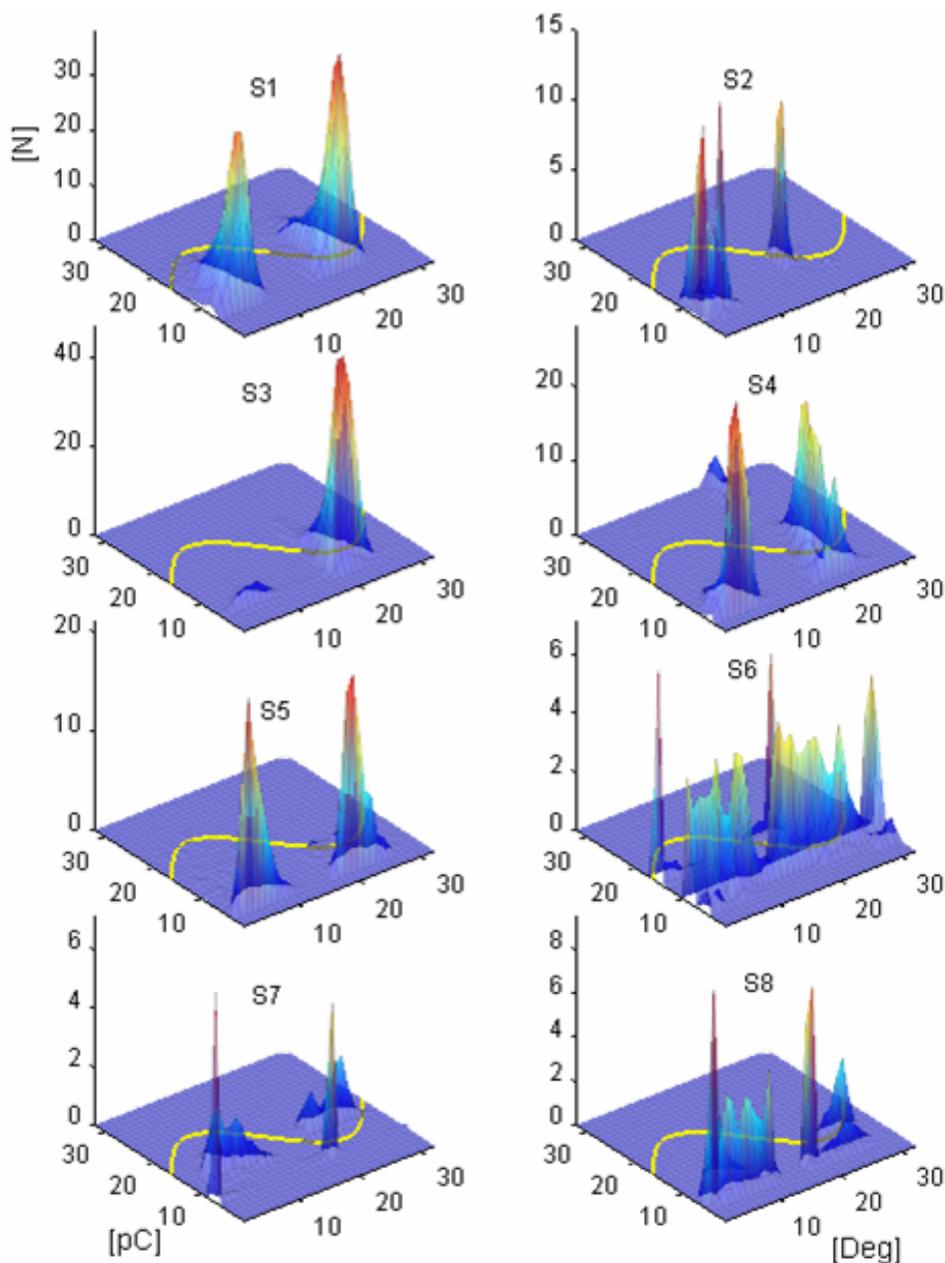


Figura 6.13 – Fontes de DPs provenientes das medições no gerador U10 da usina Itaipu (conforme Tabela 5.6).

6.3.1 Representação como um Problema BSS

Cada matriz reduzida $[\mathbf{R}]_{32 \times 32}$ pode ser disposta na forma de um vetor amostra $\mathbf{x}(k)$, de tal forma que:

$$\mathbf{x}(k) = [x_1(k), \dots, x_N(k)]^T \quad (6.8)$$

onde k indica a k -ésima amostra tal que,

$$x_i(k) = \sum_{j=1}^N a_{ij}s_j(k) + \eta_i(k) \text{ ou } \mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \boldsymbol{\eta}(k) \quad (6.9)$$

Neste caso, \mathbf{A} é uma matriz escalar $N \times M$ não singular, desconhecida, mas invariante a transformações, $\boldsymbol{\eta}(k)$ é o vetor ruído, que é assumido de média nula e variância v_{ij} , e $\mathbf{s}(k)$ é o vetor de M fontes de sinais independentes,

$$\mathbf{s}(k) = [s_1(k), \dots, s_M(k)]^T \quad (6.10)$$

O objetivo do ICA é estimar a matriz \mathbf{A} ou sua inversa expressa como:

$$\mathbf{y}(k+1) = \mathbf{W}_k \mathbf{x}(k+1), \quad \mathbf{W}_k = \mathbf{A}_k^{-1} + \varepsilon \mathbf{A}^{-1} \mathbf{A}^{-T} \mathbf{A}^{-1} \quad (6.11)$$

onde ε é um valor real (ex. $\varepsilon = v_{ij}$).

O ICA com restrições não-holonômicas, desenvolvido por Amari et al. [58] e Akuzawa [56] e [84], foi empregado em todos os ensaios laboratoriais desta tese. Akuzawa mostra que o ICA, baseado em métodos de Newton, é robusto sob ruídos gaussianos. Neste caso ele funciona melhor do que os métodos JADE [81] e FastICA [79]. Uma de suas características é a estabilidade global resultante da introdução de duas deformações, sendo uma delas chamada deformação homotópica tipo ponto fixo na matriz de atualização [56]. Outra característica é que o método não requer pré-branqueamento, ou seja, poupa a metade dos graus de liberdade que naturalmente são utilizados nos métodos que precisam de branqueamento.

6.3.2 Descrição do ICA não-Holonômico

Este método utiliza cumulantes de quarta ordem de variáveis estocásticas, onde a média amostral denotada por $\langle \rangle$ e $\langle \rangle_c$ são cumulantes estimados a partir dos dados observados. Os cumulantes de quarta ordem $\langle y_1^3 y_2^1 \rangle_c$, $\langle y_1^2 y_2 y_3 \rangle_c$ e $\langle y_1^2 y_2^2 \rangle_c$ são chamados de tipo_ (3,1), tipo_ (2,1,1) e tipo_ (2,2) respectivamente.

O ICA baseado em quase-newton num espaço de dimensão P pode ser interpretado de duas maneiras:

1- A minimização de uma função custo baseada em sua expansão de segunda ordem.

2- A determinação de um ponto onde funções-objetivo Q_p para $1 \leq i \leq P$ se desvanece simultaneamente, a partir das suas expansões de primeira ordem.

A segunda interpretação será escolhida como função-objetivo, a fim de evitar estatísticas de segunda ordem.

Assume-se que haja K amostras dos sinais de DP de dimensão N , podendo ser denotados na forma geral por $\{(y^{(0)})_{ik} \mid 1 \leq i \leq N, 1 \leq k \leq K\}$. O alinhamento destas foi realizado, tal que, $\sum_{k=1}^K (y^{(0)})_{ik} = 0, \forall i$. O algoritmo de Azukawa é baseado na construção de uma seqüência de vetores $y^{(0)}, y^{(1)}, y^{(2)}, \dots$, que converge para um ponto ótimo $y^{(\infty)}$, onde cada componente torna-se mutuamente independente. O fluxo desta seqüência é especificado pelas matrizes $N \times N$ $\{\Delta^{(s)}; s=0,1,2,\dots\}$, que descreve a quantidade de passos, então a Equação (6.11) pode ser escrito como:

$$y^{(s+1)} = (\exp \Delta^{(s)}) y^{(s)} \quad (6.12)$$

para $s=0$ tem-se $y_k^{(0)}$ é o k -ésimo dado das observações. Então para o fluxo de seqüência no n -ésimo passo o $y^{(n)}$ dependerá dos $\{\Delta^{(k)}\}_s$ e será da forma $y^{(n)} = e^{\Delta^{(n-1)}} e^{\Delta^{(n-2)}} \dots L e^{\Delta^{(0)}} y^{(0)}$.

A forma vetorial para uma amostra com dimensão N é:

$$y = B y^{(0)} \quad (6.13)$$

onde $B \in GL(N, R)$ é um subconjunto do grupo linear de matrizes reais e não singular $N \times N$. Este também pode ser útil para sistemas ou problemas que necessitem de otimização que se dá através de B , assumindo p critérios, $\{Q_p \mid \mathfrak{R}^K \times \mathfrak{R}^K \rightarrow \mathfrak{R} \mid 1 \leq p \leq P\}$ até que se torne invariante sob multiplicação no processo de atualização do fluxo de seqüência. O $\exp(\Delta^{(s)})$ é um subconjunto (*coset*) de B e,

$$Q_p(y_i, y_j) = \langle y_i^p y_j^{4-p} \rangle_c \quad (6.14)$$

é função de cumulantes.

O gradiente não-holonômico é escolhido e deformado por ξ tal como:

$$V = \begin{pmatrix} K_i & (3-\xi)R_{ij} \\ (3-\xi)R_{ij} & K_j \\ 2Q_{ij} & 2Q_{ji} \end{pmatrix} \quad (6.15)$$

onde $K_i = \langle y_i^4 \rangle_c$, $K_j = \langle y_j^4 \rangle_c$ e $R_{ij} = \langle y_i^2 y_j^2 \rangle_c$

Para resolver (6.12), a quantidade de passos é determinada à maneira de Newton, tal que os Δ_{ij} e Δ_{ji} sejam determinados pela seguinte regra de atualização:

$$\begin{pmatrix} \Delta_{ji} \\ \Delta_{ij} \end{pmatrix} = -(V^T V)^{-1} V^T \begin{pmatrix} Q_{ij} \\ Q_{ji} \\ R_{ij} \end{pmatrix} \quad (6.16)$$

onde a matriz de transformação $(V^T V)^{-1} V^T$ é interpretada como uma inversa generalizada de V [56].

A Figura 6.14 mostra uma visão dos gradientes de B no espaço homogêneo da geometria diferencial, onde o gradiente baseado em restrições não holonômicas é equivalente ao gradiente natural com considerações geométricas. Este deve lidar com mudanças rápidas e/ou intermitentes nas magnitudes das fontes de sinal.

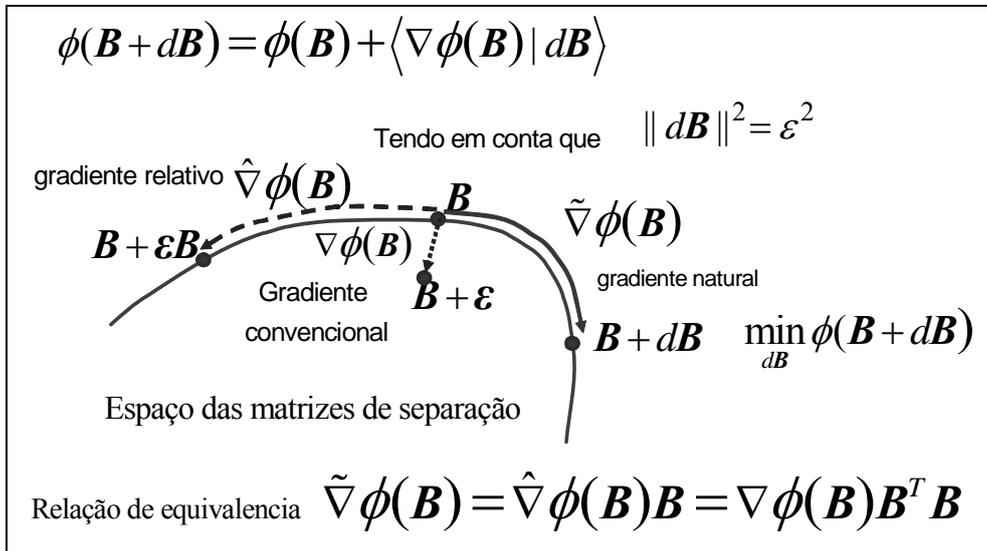


Figura 6.14 – Espaço homogêneo das matrizes, $d\mathbf{B}$ é o espaço de tangentes [85].

Esta figura apresenta um ponto, no espaço de parâmetros, B e outro $B + dB$, que representa uma pequena variação com respeito a B . A direção pontilhada define a direção de mínima distância euclidiana, desempenhado pelo gradiente convencional $\nabla \phi(B): B \rightarrow B + \varepsilon$. E a direção contínua define a direção geodésica, através do espaço curvo das matrizes inversíveis, regidos pelo gradiente natural $\tilde{\nabla} \phi(B)$ ou gradiente relativo $\hat{\nabla} \phi(B)$. Isto é, dentro do conjunto de direções (espaço de tangentes) que partem do ponto B , aquela que considera a estrutura do espaço de parâmetros é a que permite avançar mais rápido. Logo ao se definir um gradiente, para introduzi-lo como direção de máxima descida, num algoritmo de aprendizado deve-se considerar a métrica (medida

de distancia), pois esta caracteriza o espaço. Nesta figura se mostra uma relação de equivalência entre os gradientes, detalhes poder ser encontrados em [85].

6.3.3 Separação de Fontes e Ruídos

A BSS e ICA são as novas ferramentas aplicadas aos sinais de DP com a finalidade de conseguir extrair características ocultas pelos sinais de ruído, as quais podem ajudar a descobrir e descrever uma fonte real de DP. O resultado esperado, pós-tratamento iterativo BSS-ICA sob os sinais observados de DP, é que se consigam as possíveis fontes estimadas independentes estatisticamente. O ruído aditivo, no caso assumido normal, desvanece pelo fato de que o tratamento ICA elimina fontes gaussianas.

6.3.4 Relação SNR e Crosstalks

Duas medições são utilizadas para avaliar o algoritmo ICA utiliza a relação sinal-ruído (SNR) e crosstalks.

A SNR é um importante parâmetro utilizado para determinar a qualidade da informação. Também é um dos índices mais importantes para avaliar o desempenho dos algoritmos ICA. Ou seja, o SNR é uma medida de longo prazo para verificar a qualidade do sinal reconstruído.

Os crosstalks são interferências ou acoplamentos mútuos entre observações, misturados até serem pré-processados numa batelada de observações. Uma medição deste parâmetro indica o ruído remanescente, que é indicado pelo parâmetro SNR.

$$SNR = -10 \log\left(\frac{V_\eta}{V_x}\right), \quad CT_{1 \leq N \leq 8} = SNR(\mathbf{B} * \mathbf{A}k), \quad CT_{medio} = \sum_i |CT_N| \quad (6.17)$$

$$\mathbf{A}k = \mathbf{A}_{original} * \text{diag}(\text{Cum4}(S))^{1/4}$$

6.3.5 Seqüência experimental

Os dados coletados na medição de DP representam os sinais observados, compostos pelo sinal fonte (desconhecido) e o ruído aditivo (natural). Com o intuito de estimar fontes mais próximas de sua natureza, foram praticados os seguintes passos:

- Introduz-se um ruído artificial do tipo gaussiano nas fontes de DPs, com 1% de amplitudes com respeito ao sinal de DPs.
- Observa-se que o ICA consegue limpar o ruído artificial introduzida, estimando as fontes de entrada.

- Acrescenta-se gradualmente o nível do ruído até o limite de 60% do sinal de DPs.

- Verificar se o ICA limpa o ruído artificial para cada nível percentual acrescentado.

- Em seguida aplica-se a teoria cega (BSS) nas fontes, vistas agora como observações, ou seja, sem acréscimo de ruído artificial. Neste estágio é avaliada a capacidade do ICA em remover ruídos gaussianos, naturalmente achados nas fontes originais. Desta forma, consegue-se estimar as fontes sem ruído gaussiano e, por conseguinte mais próximas da sua natureza.

A Figura 6.15 mostra as oito fontes de sinais com 20% de ruído branco, observa-se que, a maioria destes estão submersos pelo efeito do ruído, tornando-se difícil o reconhecimento de pelo menos uma fonte original (que é geralmente desconhecida).

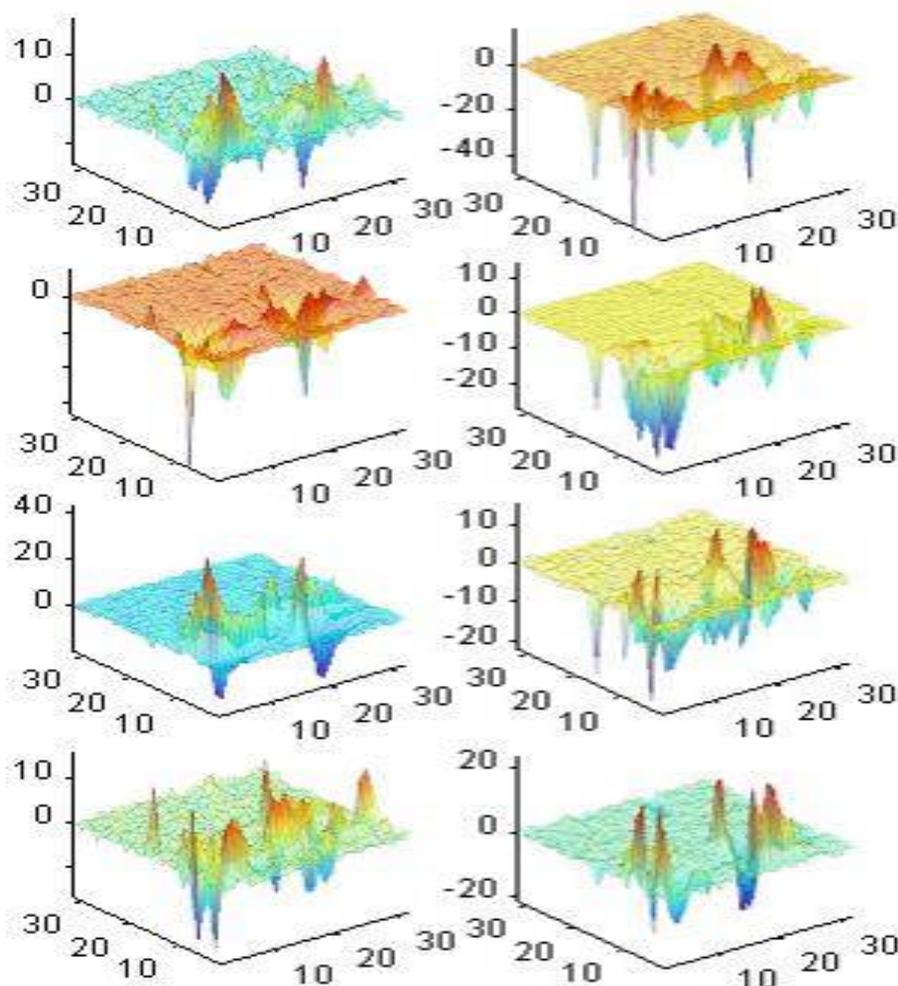


Figura 6.15 – Amostra de fontes com 20% de ruído.

6.3.6 O ICA como Extrator de Ruídos

Como se pode observar na Figura 6.16, os ruídos das observações foram significativamente reduzidos pelas características próprias do ICA, já que o algoritmo ICA não usa estatística de segunda ordem. Assim, obtém-se uma estimativa boa das fontes originais. Para qualquer ICA é impossível manter uma ordem das fontes, uma vez que as permutações das matrizes são livres, assim como as escalas [73]. Estas são as desvantagens dos métodos ICA que lidam com fontes desconhecidas, além de se desconhecer a maneira como estas foram combinadas.

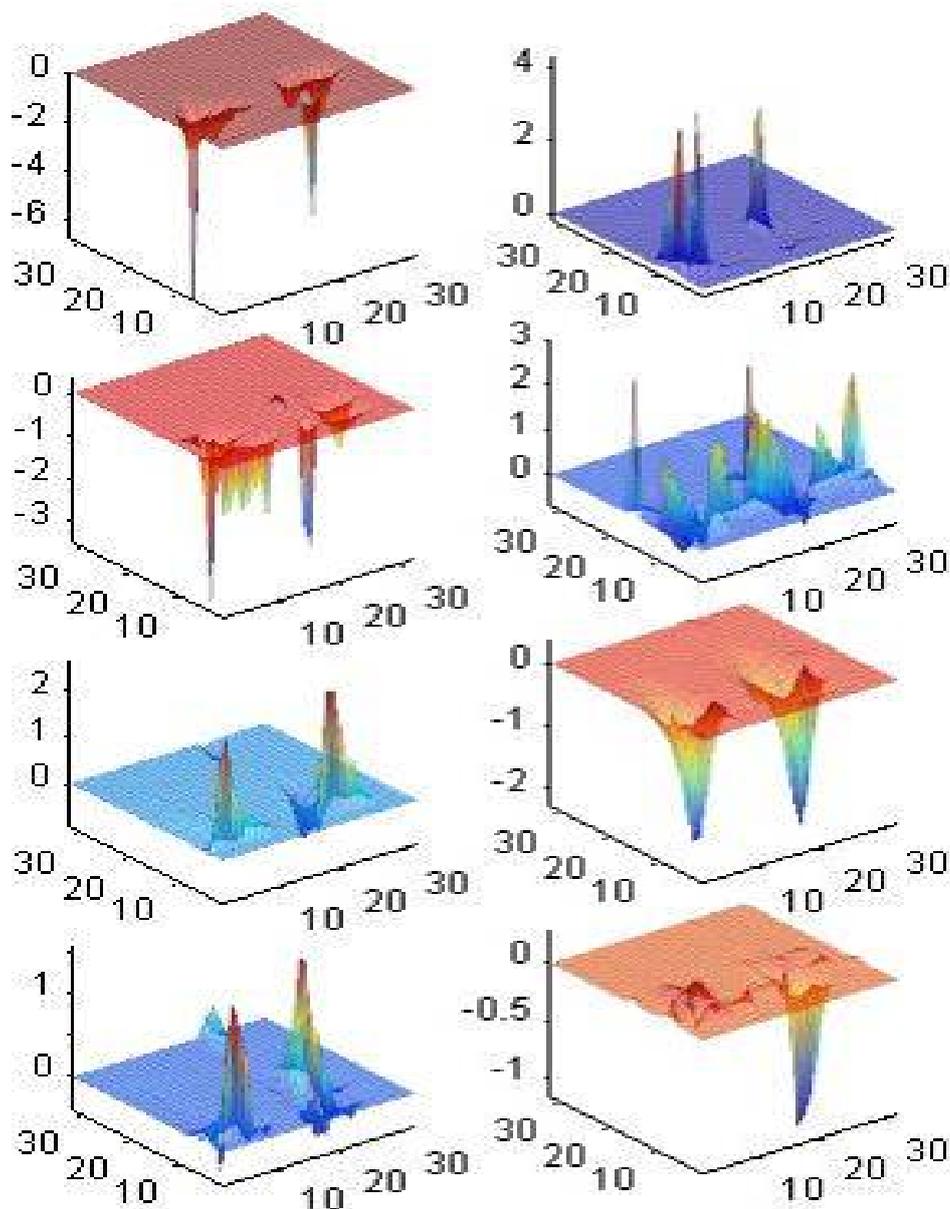


Figura 6.16 – Fontes estimadas com SNR 13,40dB pelo ICA sem pré-branqueamento.

O desempenho do ICA não-holonômico foi avaliado em função do crosstalk médio, a máxima variância das observações e o maior valor do quarto cumulante. O

índice *crosstalk* (Equação 6.17) indica a informação mútua remanescente (ver Tabela 6.7). Assim, a matriz estimada **B** é mostrada nas Figuras 6.16 e 6.17 para o SNR 5dB e 13,40dB respectivamente.

Este método ICA se diferencia do outros métodos ICA convencionais. Enquanto que no ICA convencional é necessário obter a matriz identidade, resultado do produto interno entre a matriz inicial **A** e a matriz estimada **B**, para garantir boa separação das fontes; o ICA não-holonômico não apresenta uma matriz identidade no produto interno, e sim uma matriz de restrições não-holonômicas $dX = dBB^{-1}$, $(dX)_{ii} = 0$ em razão do gradiente natural utilizado.

Tabela 6.7 - Parâmetros Estatísticos e desempenho por Crosstalks

Crosstalk Por VAR.	Crosstalk Por CUM.	Max. VAR.	Max.4. CUM.	SNR (dB)	Ruído (%)	 Ganho Apx. (dB)
2,020	2,150	0,041	1,03 {3}	Inf	--	5,031
0,288	0,288	0,043	1,86 {6}	41,13	1	8,550
0,296	0,293	0,202	7,87 {7}	20,12	10	21,247
0,321	0,328	0,104	5,15 {8}	16,84	15	22,083
0,310	0,314	0,196	4,14 {5}	13,40	20	21,312
0,503	0,505	0,134	2,67 {7}	7,33	30	24,443
0,426	0,426	0,134	2,67 {7}	6,29	40	30,430
0,610	0,603	0,054	0,30 {6}	5,01	60	27,26

(*) Medições preliminares por cumulantes de quarta ordem foram realizadas nas fontes, onde uma delas apresentava o menor valor cumulante, isto é o Cum4(S) é igual a 22,215dB.

6.3.7 Resultados Referentes ao ICA

O método ICA pode ser utilizado na separação de fontes ruído dos sinais de DP, apresentando resultados melhores do que os métodos tradicionais utilizados na filtragem de sinais de DP. Este método não requer pré-branqueamento, diminui o número de operações em N^2 e não é necessário tratar operações matriciais gigantescas, precisando

apenas calcular uma matriz inversa generalizada de 3×2 que concentra os critérios-cumulantes como condição do desvanecimento destes quando converge o algoritmo.

A mais importante vantagem deste método, no processamento de sinais de DP, é sua robustez sob ruídos gaussianos. Caso o ruído seja colorido um pré-tratamento pode ser implementado.

Nos casos em que as observações foram pré-branqueadas, haverá a necessidade de pós-processar via pós-análise de fatores (semelhante ao PCA) [84], devido ao fato que parte do ruído foi pré-processado, tornando-se ruído gaussiano multidimensional com correlação cruzada. A Figura 6.17 e 6.18 mostra a matriz estimada B para SNR 5dB e 13,4dB respectivamente. Observa-se que as amplitudes diminuem para condições de menor energia quando SNR cresce.

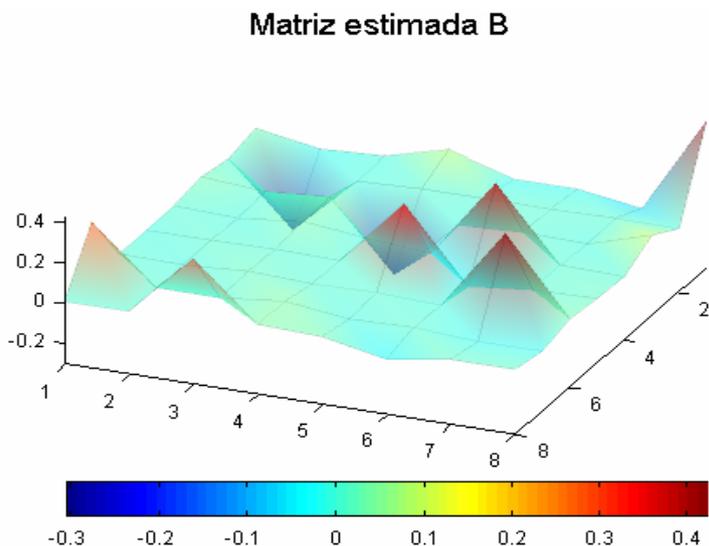


Figura 6.17 – Fontes estimadas com SNR 5dB pelo ICA sem pré-branqueamento.

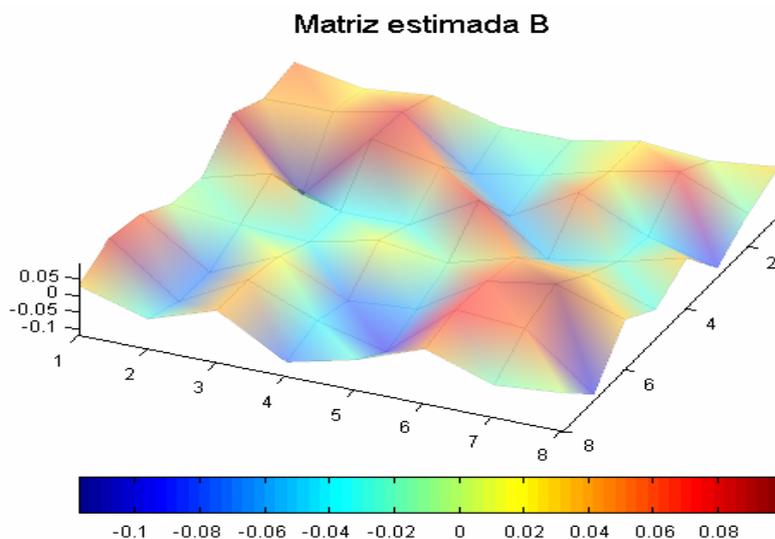


Figura 6.18 – Fontes estimadas com SNR 13,40dB pelo ICA sem pré-branqueamento.

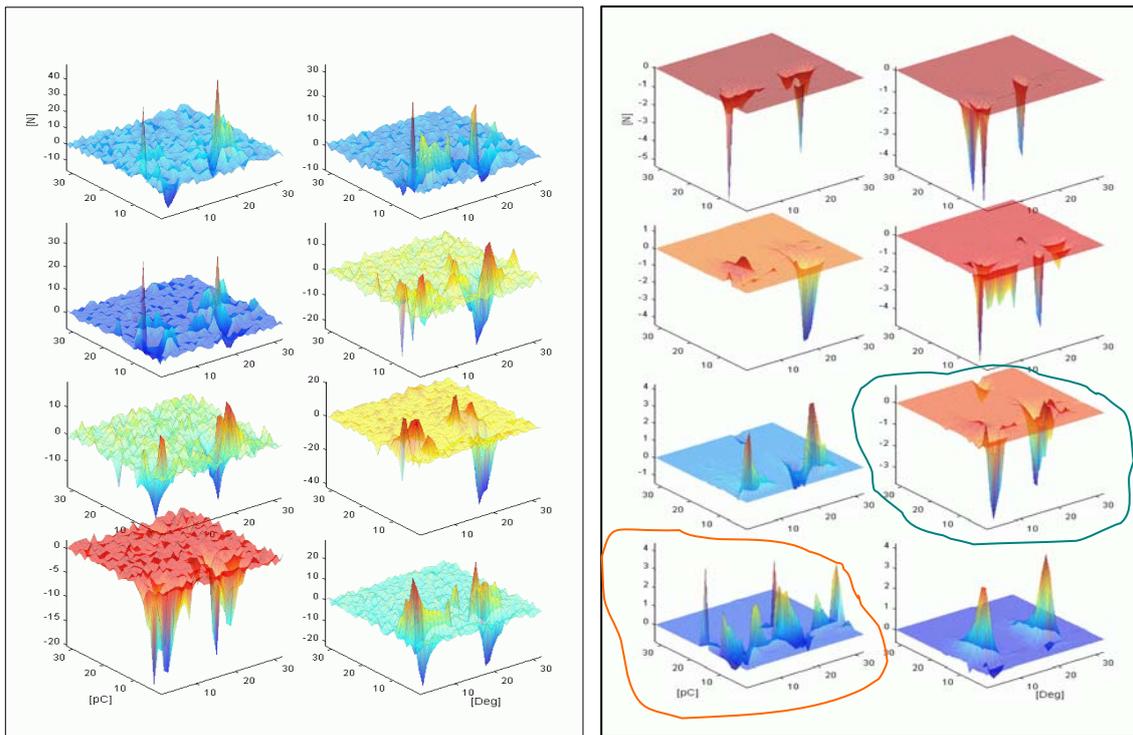


Figura 6.19 – Esquerdo: Observações com 40%-6,29dB de ruído, Direito: Fontes estimadas pelo ICA não-holonômico com SNR= 6,29dB

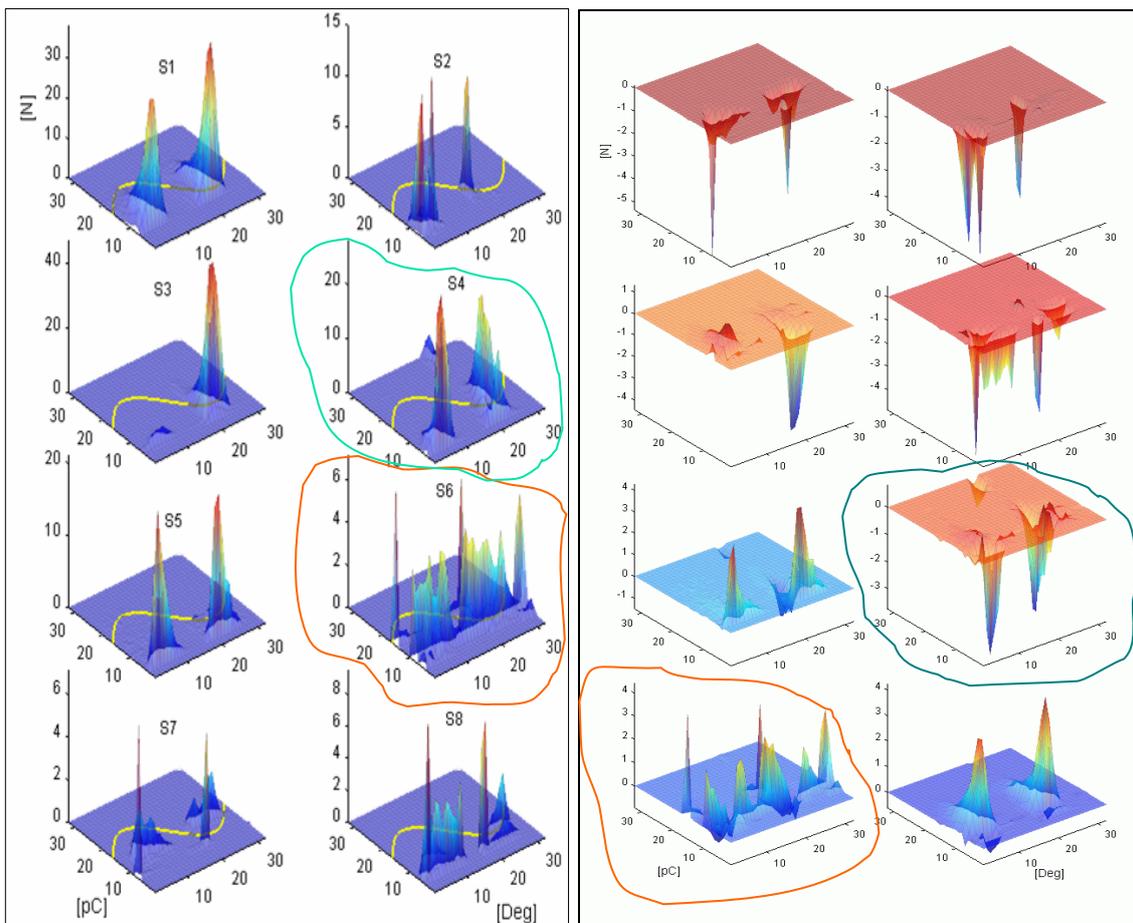


Figura 6.20 – Ganho obtido pelo caracterizador 30,43dB : Direita: Fontes originais Esquerda: Fontes estimadas com 6,29DB de ruído. Entre estas duas fontes tiveram ganhos que podem observar-se visualmente.

A Figura 6.19 e Figura 6.20 mostra o desempenho realizado pelo caracterizador. Foi injetado ruído gaussiano artificial de 6.29dB nas observações, observa-se que o ICA consegue separar e estimar as oito fontes, conseguindo obter um ganho de 30,43dB (ver Tabela 6.7). As fontes originais e as fontes estimadas foram comparadas e mostradas na Figura 6.20. Pode-se identificar visualmente a recuperação de todas as fontes. Além disso, pode-se ver até mesmo uma melhoria em suas topologias (estruturas), pelo menos em duas fontes sinalizadas por curvas fechadas. Tais fontes obtiveram os maiores ganhos, onde alguns componentes originais foram eliminados e acredita-se que eram componentes de ruído, visto que a anulação dos cumulantes de quarta ordem, por otimização, foi o critério da função contraste do ICA (eeA).

Tabela 6.8 – Quadro comparativo de três Modelos ICA

	FastICA	JADE	eeA-Akuzawa
Erro (DB)	2,3133	2,2442	1,2283
Soma Cum4(Y)	25,6500	26,2193	0,2812
Soma Cum4(S)	22,2148	22,2148	22,2148

A Tabela 6.8 mostra o desempenho de três modelos ICA. Destaca-se que o FastICA e o JADE precisam de pré-processamento, “branqueamento”, enquanto que o eeA não utiliza estatísticas de segunda ordem e pode trabalhar diretamente com os dados.

O quadro acima mostra o somatório dos cumulantes de quarta ordem da entrada Cum4(S) e da saída Cum4(Y). Observa-se que, o FastICA e o JADE têm desempenhos parecidos e, o eeA apresentou menor erro (1,2283 dB). O eeA trabalha com gradiente natural e distância geodésica com restrições não-holonômicas, a função contraste é uma matriz de 3x3 de cumulantes de quarta ordem. Uma desvantagem deste algoritmo é a sensibilidade a *outliers*- (dados intrusos) [84], [73].

6.4 Caracterizador de Fontes Ruidosas de DPs por ICA Shrinkage

Nesta seção, se realiza uma aproximação para o problema de remoção de ruídos em sinais fonte de DPs, usando o ICA estendido, Noisy_ICA,

É mostrado que os modelos Shrinkage de funções podem ser aplicados na caracterização dos sinais de DPs, desde que se faça uma boa escolha das funções contraste ou *function score* e, pelo menos, conheça-se a variância do ruído. Caso não seja conhecido o ruído, estes precisam ser estimados pelos estimadores MAP e ML [60], [73], [77], [80].

Os dados são fontes de defeitos procedentes de enrolamentos e barra estática envelhecida, utilizadas na representação de defeitos em sistemas de isolamento, possíveis de serem encontrados em hidrogeradores e geradores e outros variantes de defeitos encontrados no próprio gerador fonte (ver Capítulo 3). Para realização deste caracterizador se tomaram as fontes, as quais, para este modelo caracterizador, foram consideradas como observações ruidosas.

6.4.1 Descrição do ICA Shrinkage

Na secção 5.5.8, foram descritas as equações mais importantes para implementar o Noisy_ICA. A estimação dos componentes principais será realizada por meio das funções shrinkage, que, por sua vez, utilizam os critério de codificação esparsa em conjunção com o princípio da máxima semelhança para remover ruídos dos dados DPs.

Cabe ressaltar que na Equação (5.65) a inversão solicitada pode ser impossível, analiticamente. Segundo Hyvärinen et. al. [60], [73], [76], [80], [83], existem algumas funções que têm grande praticidade na implementação deste método, conhecido como codificação shrinkage (SCS).

Modelos de densidade esparsa

Dados esparsos são característicos nas medições de DPs, visto que o depósito dos pulsos das descargas se concentram em uma área em volta de grupo de ângulos de referência da tensão aplicada aos ensaios de DPs. Estas descargas podem ser observadas na Figura 6.21, mostra também a ocorrência de duas variantes no eixo das projeções das amplitudes da descarga e no eixo das projeções angulares, além das quantidades de descarga armazenada no eixo vertical. Estas variantes não têm forma gaussiana, sendo que, estas possuem caldas alongadas por amplitudes e levemente na referência angular. Contudo, esta figura mostra um verdadeiro histograma das DPs. Porém, observa-se também a existência de outros pontos espúrios, que são apontados como supostos intrusos *outliers* ou ruídos, dependendo da amplitude.

A Figura 6.22 mostra duas formas gaussianas simétricas. Observa-se que, fazer um modelo de funções compostas gaussianas, não seria coerente com o mapa das descargas, se estaria perdendo tanto na informação quanto na forma. Além disso, o modelo ICA não seria adequando para funções gaussianas, visto que, é fundamental que as funções tenham a característica de não gaussianidade.

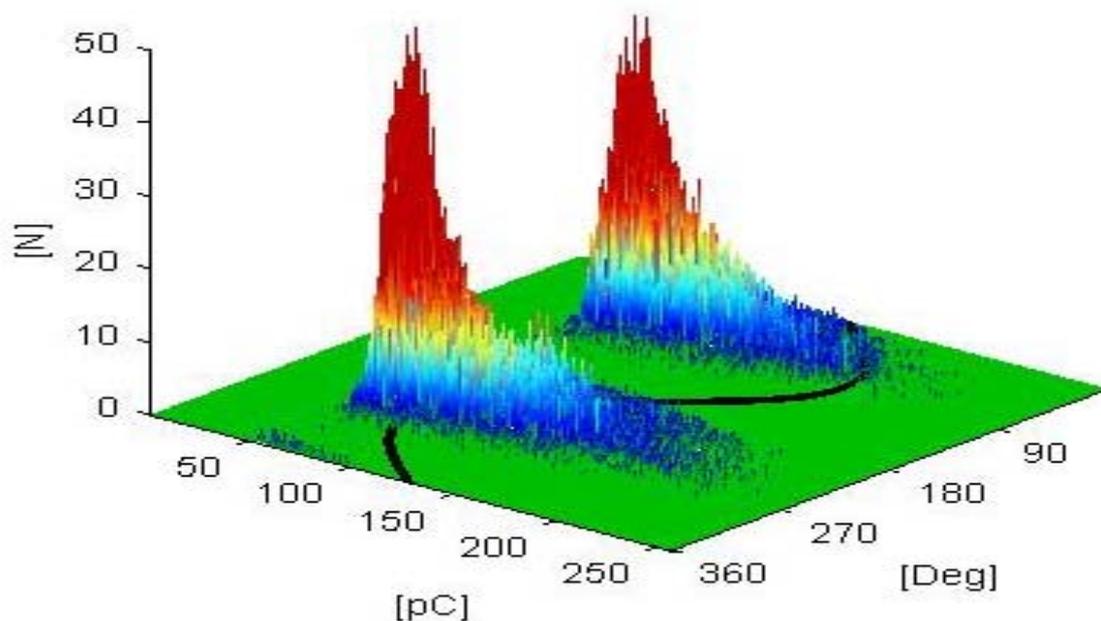


Figura 6.21 – Dado coletado de DPs: histograma com 256x256 canais (bins)

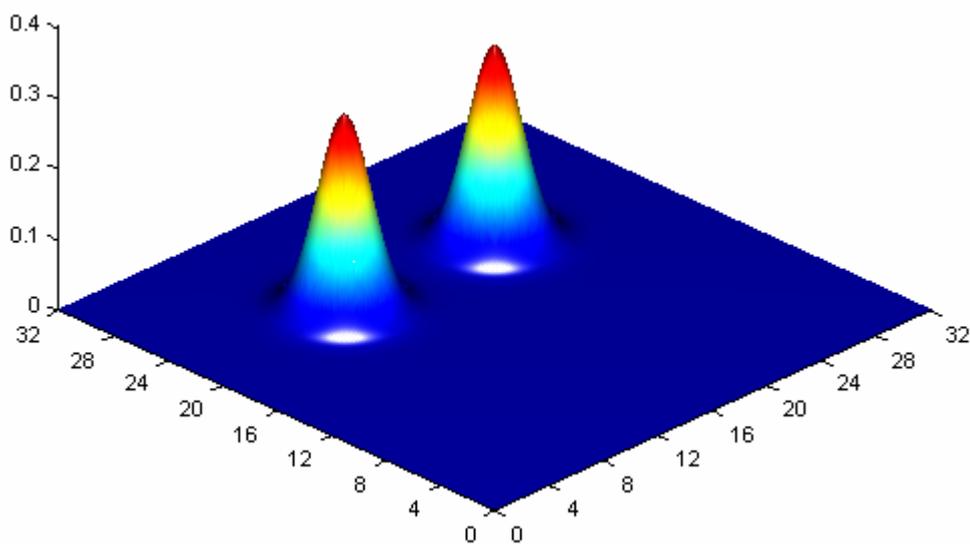


Figura 6.22 – Modelo duplo gaussiano

Dados esparsos são caracterizados por serem suavemente ou fortemente esparsos. Logo, serão mostrados alguns modelos destas funções, os quais são dados em [73], [76]. Nota-se que, a densidade laplaciana ou duplo exponencial é uma das funções utilizadas como função pdf dos componentes s . A Figura 6.23 mostra a função laplaciana na

forma multidimensional e simétrica para expressar a estimação não paramétrica das DPs bivariantes frente a ruídos gaussianos relativos. Talvez seja preciso formar um composto de funções densidade para modelar DPs tridimensional.

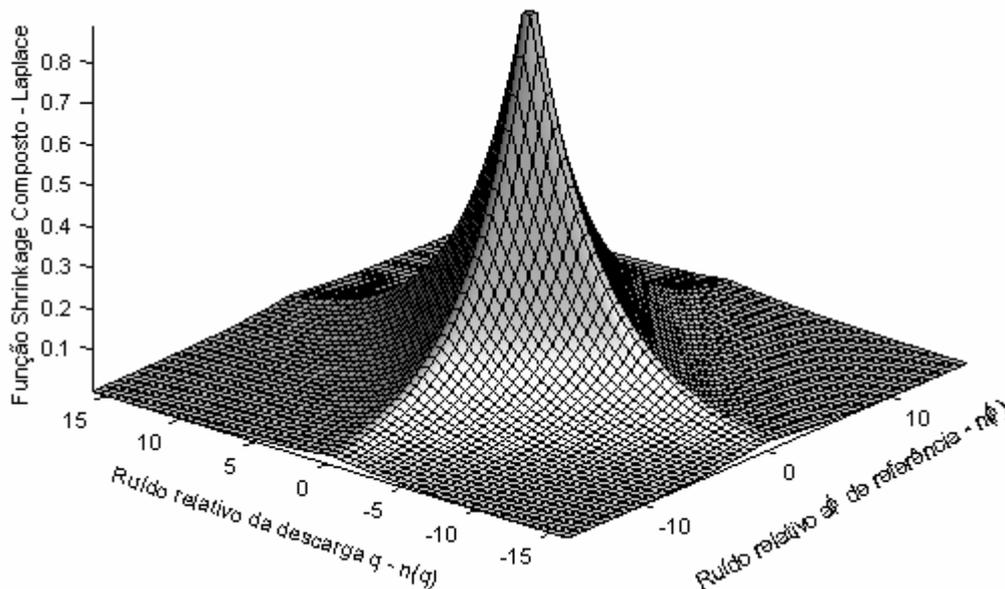


Figura 6.23 – Modelo laplaciano para Noisy_ICA em dados DPs

A Figura 6.24 mostra pictoricamente o histograma de ruídos gaussianos introduzidos nas fontes das DPs, sendo que, segundo o Capítulo 3, a calibração dos circuitos e do instrumento de medição, encontram este tipo de ruído, para reduzir tal efeito, filtros de linha são instalados ou filtros adaptativos via hardware/software.

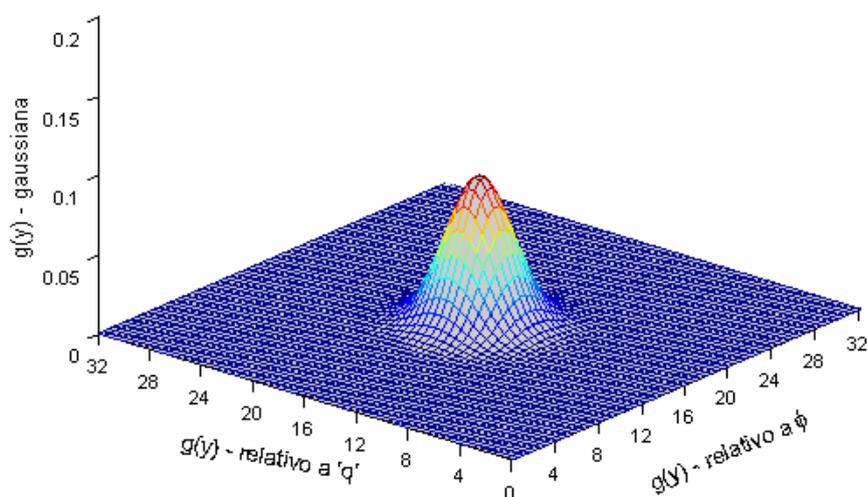


Figura 6.24 – Modelo gaussiano para resíduos em dados DPs
Densidade esparsa suave:

$$p(s) = C \exp\left(\frac{-as^2}{2} - b|s|\right) \quad (6.18)$$

Densidade laplace:

$$p(s) = \frac{1}{2d} \exp\left(\frac{-\sqrt{2}s}{d}|s|\right) \quad (6.19)$$

Densidade laplace esparsa forte:

$$p(s) = \frac{1}{2d} \frac{(\alpha + 2)[\alpha(\alpha + 1)/2]^{(\alpha/2+1)}}{\left[\sqrt{\alpha(\alpha + 1)/2} + |s/d|\right]^{(\alpha+3)}} \quad (6.20)$$

onde a , b e d são os parâmetros que determinam a natureza das distribuições e devem ser estimados a partir das observações. A seguir, mostra-se o gráfico das equações 6.18-6.20 na Figura 6.25.

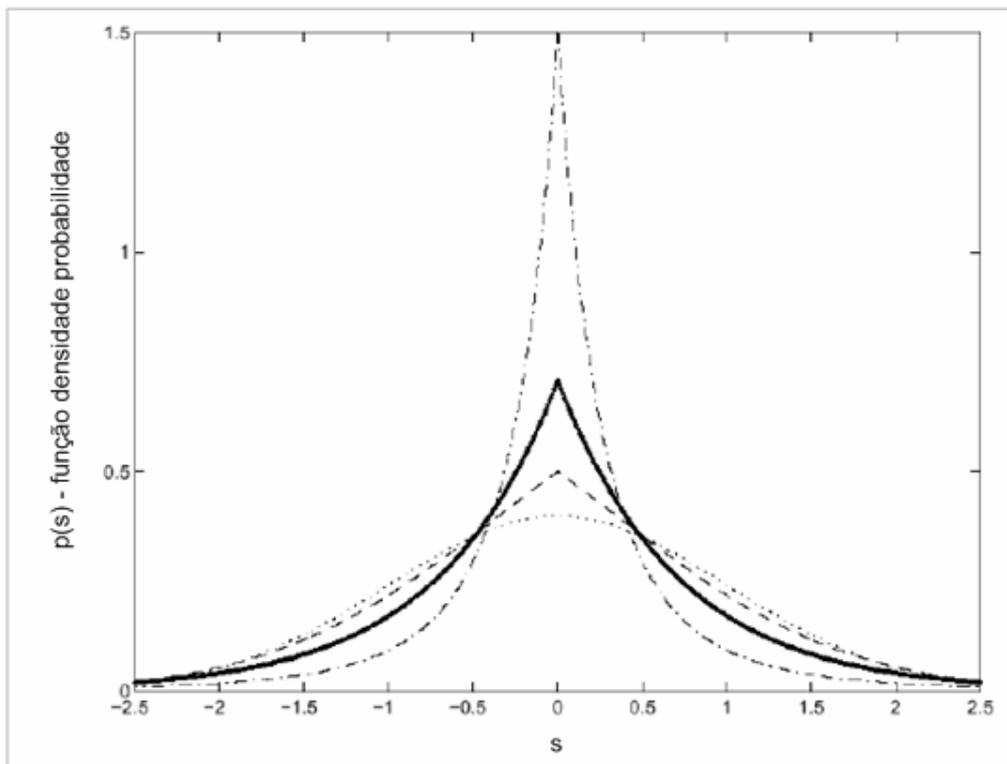


Figura 6.25 – Modelos de componentes esparsos. Linha sólida: densidade laplaciana. Linha traçada: uma típica densidade moderadamente esparsa (super-gaussiana). Linha de traçada e pontilhada: uma densidade típica fortemente esparsa (super-gaussiana). Linha pontilhada: densidade gaussiana para referencia [80], [76].

O modelo do chamado função shrinkage $g(u)$ correspondente às distribuições acima se tornam funções de contraste, alguns chamam de função ou ativação:

- Para densidades suavemente esparsas;

$$g(u) = \frac{1}{1+\sigma^2} \text{sign}(u) \max(0, |u| - b\sigma^2) \quad (6.21)$$

$$b = \frac{2p_s(0)E\{s^2\} - E\{s\}}{E\{s^2\} - [E\{s\}]^2} \quad (6.22)$$

$$a = \frac{1}{E\{s^2\}} [1 - E\{|s|\}b] \quad (6.23)$$

- Para densidades fortemente esparsas;

$$g(u) = \frac{1}{2d} \text{sign}(u) \max\left(0, \frac{|u| - ad}{2} + \frac{1}{2} \sqrt{(|u| + ad)^2 - 4\sigma^2(\alpha + 3)}\right) \quad (6.24)$$

$$d = \sqrt{E\{s^2\}} \quad (6.25)$$

$$\alpha = \frac{2 - k + \sqrt{k(k+4)}}{2k - 1} \quad (6.26)$$

- Para densidades laplacianas

$$g(u) = \text{sign}(u) \max(0, |u| - 2\sigma^2 / d) \quad (6.27)$$

onde σ^2 é a variância do ruído em todos os casos (ex.: $\sigma^2=0,3$).

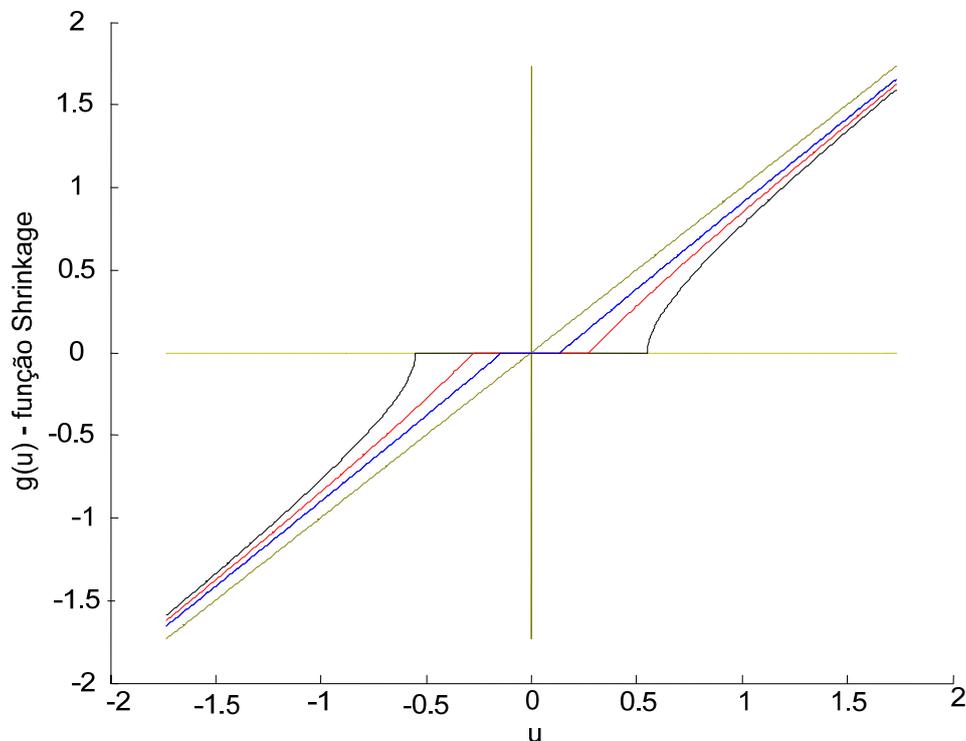


Figura 6.26 – Funções shrinkage. Linha da cor vermelha: correspondente à densidade laplaciana. Linha da cor azul: shrinkage correspondente à densidade moderadamente esparsa (supergaussiana), linha da cor preta: shrinkage correspondente à densidade fortemente esparsa (supergaussiana). Linhas cor amarela: Linhas de referencia $x=y$, $x=0$, [76].

A Figura 6.27 mostra pictoricamente uma família de funções *shrinkage* $g(u)$ tal como foi visto na Figura anterior (Figura 6.26), segundo as Equações 6.21, 6.24 e 6.27.

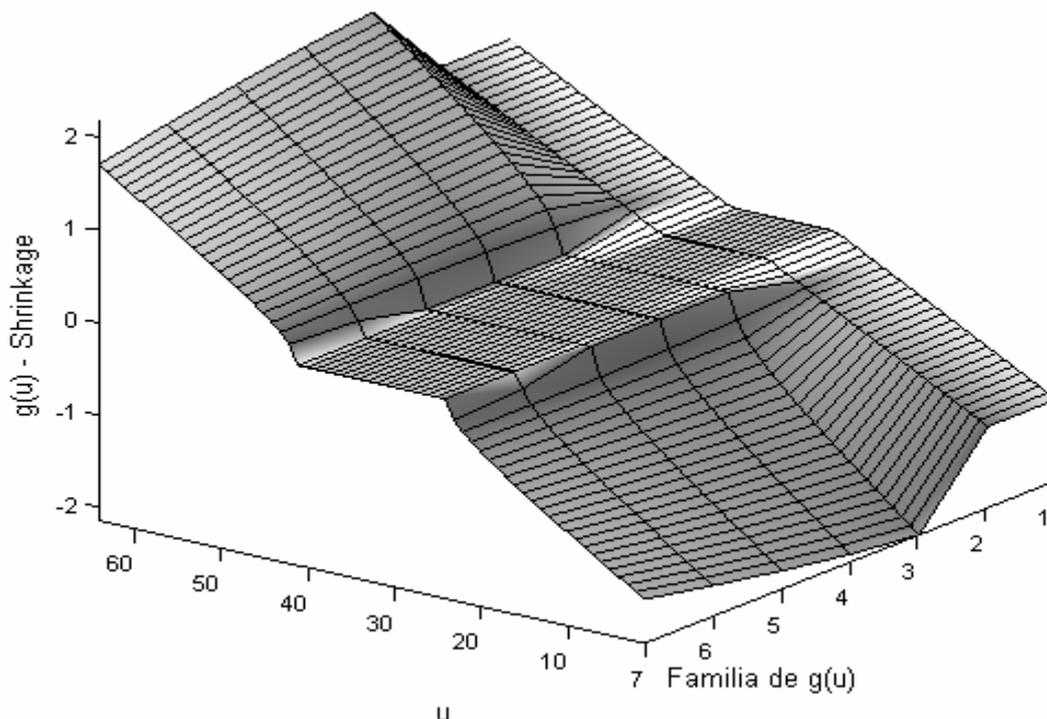


Figura 6.27 – Família de Funções shrinkage

Modelo duplo laplaciano para remoção de ruídos em DPs

Segundo a Figura 6.21, é realizada uma escolha de funções de densidade, parece que a função laplaciana é mais prática que as outras revisadas na secção anterior. O princípio ML é utilizado para este fim, remoção de ruídos gaussianos em variáveis aleatórias não-gaussianas, especificamente variáveis aleatórias super-gaussianas. Logo, é necessário estimar a matriz de mistura \mathbf{A} e a matriz de separação \mathbf{w} , sabendo que, $\mathbf{w}^T \mathbf{w} = \mathbf{I}$ e $\mathbf{w}^T = \mathbf{A}$. Além disso, é necessário conhecer um parâmetro do ruído, tipicamente sua variância. Existem varias formas de estimar o ruído, sendo um deles o uso do resíduo do PCA quando se faz compactação de dados. De forma similar o análise de fatores FA pode servir par estimar o nível do ruído utilizando a relação que existe entre a dimensão $m \leq \frac{1}{2} \{2n+1 - \sqrt{8n+1}\}$ [56]. Além disso, têm-se outros estimadores métricos baseados na medida de não-gaussianidade.

Para este caso usou-se o critério de quarto cumulante para estimar a matriz \mathbf{A} , o nível do ruído estimado pelo FA foi aproximadamente o 5% [56].

Os resultados são apresentados na forma visual. A fonte original, mostrada na Figura 6.21 é utilizada para mostrar os resultados. Utilizou-se o duplo Laplace (Figura 6.28) como função contraste. O logaritmo desta função foi tomado, assim como sua derivada com respeito à fonte estimada s (realizada no MAPLE8). Após realizar as estimativas, observou-se que, houve uma melhora na resolução do mapa de DPs, sendo que alguns pontos espúrios com baixa amplitude foram eliminados (Figura 6.29).

Finalmente, este método, mesmo com várias assunções mostrou-se eficiente para os propósitos de um caracterizador de sinais de DPs.

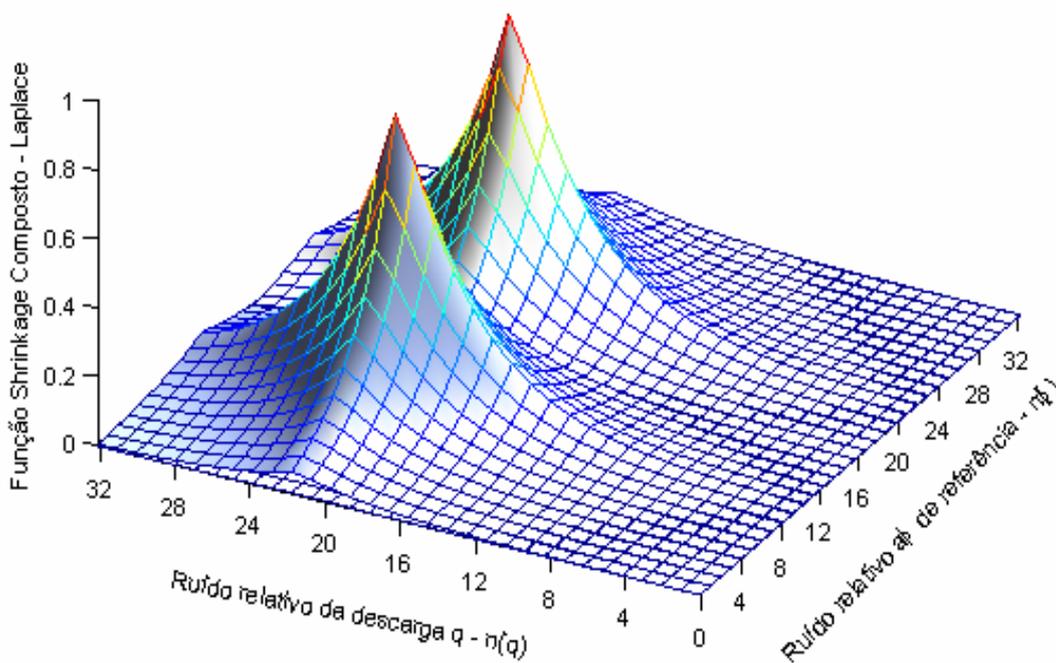


Figura 6.28 – Modelo de função proposto para remoção de ruídos

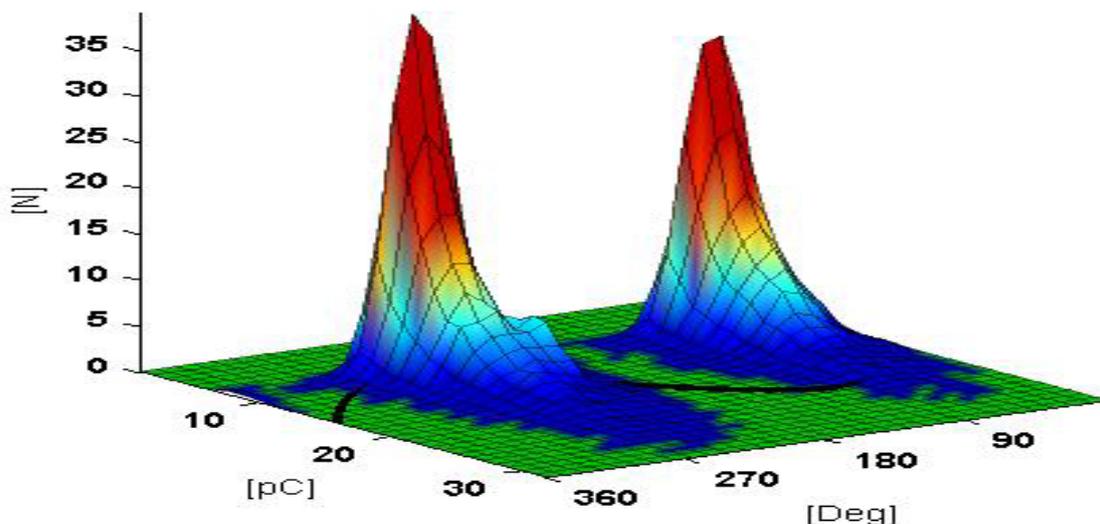


Figura 6.29 – fonte estimada pelo caracterizador ICA Shrinkage

Capítulo 7

Conclusões e Análise dos Resultados

O fenômeno de descarga parcial é pesquisado por mais de sete décadas, sua preeminente importância ferramental tem crescido para assistir na qualidade e, no desempenho de vários equipamentos em alta tensão. Assim, detectar e medir DPs numa etapa mais prematura possível e buscar sua caracterização, nesses equipamentos, têm sido uma busca incessante. Porém, os ruídos-parasitas vêm sendo um dos principais obstáculos. Devido à necessidade de melhorar a interpretação de medições de DPs, a tese propôs modelos experimentais (células teste) correlacionados de forma *a priori*, com locais críticos e de grande chance de ocorrerem DPs, em transformadores de potência. E também a construção de barras experimentais para correlacionar defeitos típicos em bobinas estatóricas de máquinas girantes, como os hidrogeradores (descritas no Capítulo 3). Além disso, para a redução ou supressão de ruídos gaussianos, foram tratados os dados na etapa de pós-processamento, empregando-se técnicas emergentes em processamento de sinais (ICA), baseadas na separação cega de fontes ruidosas (descritas no Capítulo 5 e 6).

A coleta de dados de medições de DPs, tanto em células experimentais quanto em barras estatóricas, foi obtida através da utilização do equipamento de medição de DP ICM-System da Power Diagnostic. Este equipamento gera mapas digitais contendo informações de fase, carga e taxa de repetição das amplitudes máximas das descargas, compondo uma dimensão de 256x256. Tornando-se necessário uma etapa de pré-processamento para a redução da dimensionalidade. Isso se faz necessário para reduzir o custo computacional dos algoritmos codificados para classificação e caracterização de DPs.

Uma das primeiras análises foi realizada por meio de uma de associação, baseada na análise de componentes principais (aplicada para a compactação ou a redução da dimensionalidade) e redes neurais (como classificador universal). A compactação dimensional dos dados contribuiu para melhorar a eficácia e desempenho dos sistemas

de reconhecimento de padrões de DPs. A aplicação desta associação, sobre aos dados de DPs, obteve bastante sucesso nos seus resultados, sendo portanto sua aplicabilidade recomendável.

Outra análise de dados foi realizada através de filtros ótimos, baseada na teoria de filtros casados. Esta associada com redes neurais apresentou uma boa eficiência na caracterização-classificação de sinais de DPs, devido as vantagens trazidas por suas características individuais. O filtro casado tem a especialidade de levar informação da forma do sinal. O MF apesar de ser para sinais determinísticos, foi utilizado para discriminar sinais de DPs, cujas características revelam ser de natureza estocástica e a rede neural foi utilizada para discriminar as flutuações não tratadas pelo MF Logo este híbrido resultou ser um bom classificador de sinais de DPs.

O processamento de separação de ruídos das fontes de DPs utilizando a técnica de Separação Cega de Fontes (BSS) por meio da Análise de Componentes Independentes (ICA) teve desempenho muito satisfatório, visto que foram recuperadas 100% das fontes avaliadas e muitas delas tiveram um ganho ao reduzir seus componentes ruidosos. Após pesquisa dos métodos ICA, que lidam com ruídos, concluiu-se que, as técnicas: ICA não-holonômico, JADE, Infomax e FastICA, podem ajudar na extração dos ruídos, sendo a primeira mais explorada devido a sua robustez frente a ruídos gaussianos além de ter critérios estatísticos e geométricos.

Extensões ICA foram utilizadas para remoção de ruídos e estimação de fontes de sinais expressos por seus componentes independentes. Estas técnicas provaram sua aplicabilidade nos casos onde os dados têm contaminação por ruídos. E sob certas condições como assunção de funções esparsa e forma do ruído (ex.: gaussiano), consegue-se removê-lo da fonte.

Além disso, estas técnicas podem ser utilizadas para compactação de bancos de dados que armazenam este tipo de informação. Os quais servem de consultas aos agentes de informação tal como sistemas de diagnósticos de equipamentos.

A remoção total ou até parcial dos ruídos ajuda esclarecer algumas dúvidas e ambigüidades antigas do conhecimento de fontes de sinais DPs. Por exemplo, encontrar a forma do padrão do modelo experimental mais próxima a sua natureza e possível correlação entre o modelo e local crítico (descrito no Capítulo 3).

Os sistemas de diagnóstico de equipamentos on-line podem alcançar uma melhor eficácia quando faz uso de dados reduzidos dimensionalmente, e de banco de dados

contendo melhor qualidade de informação, produto do tratamento de ruídos pelos sistemas caracterizadores.

7.1 Escopo para Trabalhos Futuros

Realizar uma ampliação dos novos modelos experimentais a fim de que se obtenham mais fontes refinadas para melhorar a representação das descargas localizadas nos equipamentos, desde que os modelos tenham correlação com dados reais encontrados no campo.

Uma proposta futura é a implementação do sistema caracterizador em programas computacionais para propósitos de tratamento de sinais de DPs no campo.

Outra proposta é o melhoramento dos algoritmos empregados na pesquisa para fins de serem utilizados em hardware via DSPs. Assim, seria possível obter um melhor desempenho dos caracterizadores. Visto que, os DSPs são altamente competentes e dedicados no processamento de sinais rápidos.

Modelar outros tipos de ruído a fim de otimizar os caracterizadores descritos. Ruídos coloridos podem ser branquiados por filtros casados ou estimados seus parâmetros por estimadores de máxima semelhança, por exemplo.

Pesquisar outras técnicas de pré-processamento, como o NLPCA, NFA, Kernel PCA, a fim de melhorar a compactação e/ou o desempenho dos métodos ICA que utilizem pré-processamento. Com estes se espera modelar com maior eficiência as componentes principais que garantam uma melhor representação dos dados originais.

Realizar uma extensão de pesquisa para outros métodos ICA Noisy e filtros Wavelet Shrinkage, devido a sua acelerada evolução em diversos campos e aplicações.

Também podem ser aplicáveis para outros casos de medição, monitoração on-line e serem parte de um diagnóstico integrado de equipamentos de subestações de energia elétrica.

Outro campo de aplicação é a purificação e, compactação de bancos de dados que contêm este tipo de informação para a melhoria da qualidade de informação.

Apêndice A Conceitos Básicos de Estatística para Métodos de Separação de Fontes

Neste apêndice se incluem alguns conceitos básicos da estatística que ajudarão a compreender melhor os conteúdos da presente memória, visto que, com estes se tem uma referência próxima e na notação que se usa na tese. Descrevem-se algumas funções de densidade de probabilidade que se citam nos Capítulos 5 e 6, porém podem aplicar-se para outros tipos de sinais que não sejam descargas parciais. Para maiores informações podem ser consultadas, por exemplo, as referências bibliográficas [86], [50], [52].

A1.1 Introdução

Na análise de um sinal aleatório, s , é de grande interesse o estudo de sua função de distribuição, $f(s)$ ($s_m \leq s \leq s_M$), visto que, está proporcional a probabilidade de que se gere um valor determinado, s , do sinal dentro da faixa em que esteja definida o sinal, $s \in [s_m, s_M]$.

Na prática pode aproximar-se a distribuição de probabilidade de um sinal a partir de sua distribuição de frequências (número de vezes que se repetem os elementos de cada uma das classes em que se subdivide o faixa de valores da variável aleatória) sem necessidade de dividir o número de elementos de cada classe por o número total de amostras, NM .

Um conjunto de NM observações de um sinal, $s(1), s(2), \dots, s(NM)$, pode ser descrito, além da função de distribuição, por meio de umas medidas descritivas que especificam, entre outras propriedades a localidade ou tendência central (média, mediana, moda, etc.), a dispersão ou variação (faixa, desvio média, desvio padrão, variância, etc.), a assimetria (“skewness”) e a agudeza (“kurtosis”).

A1.2 momentos e cumulantes

A1.2.1 Momentos

Os momentos de uma variável aleatória podem utilizar-se para descrever sua distribuição de probabilidade [86], [50]. O k-ésimo momento, μ'_k , da variável aleatória s centrado no zero (0) se define como:

$$\mu'_k(s) = E\{s^k\} = \int_{-\infty}^{\infty} s^k \cdot f(s) \cdot ds \quad k = 1, 2, \dots \quad (A1.1)$$

onde $E\{\cdot\}$ é o valor esperado ou a esperança matemática:

$$E\{s\} = \int_{-\infty}^{\infty} s \cdot f(s) \cdot ds \quad (A1.2)$$

O k-ésimo momento, μ_k , da variável s em volta de sua média, $\mu' = \langle s \rangle$, se define como:

$$\mu_k(s) = E\{(s - \mu)^k\} = \int_{-\infty}^{\infty} (s - \mu)^k \cdot f(s) \cdot ds \quad (A1.3)$$

Os momentos mais utilizados são a média $\mu = \mu'$, e a variância, $\sigma^2 = \mu_2$, onde σ é a desvio padrão. Se c é uma constante e s uma variável aleatória com média μ e variância σ^2 , se verificam as seguintes propriedades [86]:

$$E\{c\} = c \quad (A1.4)$$

$$E\{c \cdot s\} = c \cdot E\{s\} \quad (A1.5)$$

$$\sigma^2 = E\{(s - \mu)^2\} = E\{s^2\} - \mu^2 \quad (A1.6)$$

$$\sigma^2_{(c+s)} = \sigma^2 \quad (A1.7)$$

$$\sigma^2_{c \cdot s} = c^2 \cdot \sigma^2 \quad (A1.8)$$

Os momentos de uma variável aleatória podem ser estimados diretamente, a partir das definições (A1.1) e (A1.3) ou a partir da função geradora de momentos, $m(t)$, daquela variável que se define como:

$$m(t) = E\{e^{ts}\} \quad (A1.9)$$

Se $m(t)$ existe, então, os momentos centrados na referência zero (0) podem ser calculados com a expressão:

$$\mu'_k = \left. \frac{d^k m(t)}{dt^k} \right]_{t=0} \quad (A1.10)$$

Sobre certas condições, que costumam cumprir a maioria das funções de distribuição, pode-se dizer que, se duas variáveis aleatórias, s_1, s_2 , têm seus respectivos momentos iguais,

$$\mu_k(s_1) = \mu_k(s_2) \quad k = 1, 2, \dots, \quad (\text{A1.11})$$

então suas distribuições de probabilidade, $f1(s_1)$ e $f2(s_2)$, serão idênticas. Por conseguinte, se elas têm iguais os três ou quatro primeiros momentos, logo estas podem ser consideradas como idênticas.

A1.2.2 Cumulantes

Os cumulantes, κ_k , são outras constantes que descrevem e especificam propriedades e circunstâncias de uma distribuição e são más úteis, desde um ponto de vista teórico, que os momentos. Os cumulantes de ordem k podem ser calculados a partir dos momentos de ordem igual ou inferior a k . Os cumulantes de ordem 1 até 4, centrados na origem “zero” e em função dos momentos são:

$$\begin{aligned} \kappa_1 &= \mu_1' \\ \kappa_2 &= \mu_2' - \mu_1'^2 \\ \kappa_3 &= \mu_3' - 3\mu_2' \cdot \mu_1' + 2\mu_1'^3 \\ \kappa_4 &= \mu_4' - 4\mu_3' \cdot \mu_1' - 3\mu_2'^2 + 12\mu_2' \mu_1'^2 - 6\mu_1'^4 \end{aligned} \quad (\text{A1.12})$$

Se as variáveis aleatórias estão centradas em sua média, é dizer, $\mu_1' = 0$, as expressões dos cumulantes se simplificam:

$$\begin{aligned} \kappa_1 &= 0 \\ \kappa_2 &= E\{s^2\} = \mu_2 \\ \kappa_3 &= E\{s^3\} = \mu_3 \\ \kappa_4 &= E\{s^4\} - 3E\{\kappa^2\} = \mu_4 - 3\mu_2^2 \end{aligned} \quad (\text{A1.13})$$

Para uma variável gaussiana (Secção. A1.3.2) os cumulantes de ordem superior a 2 são todos nulos; é dizer, $\kappa_k = 0$ ($\forall k > 2$).

A variância σ^2 , cumulante de s de ordem 2, $\sigma^2 = \kappa_2$, caracteriza a potencia o dispersão de s .

Centrando na média e normalizando o sinal aleatório pode-se efetuar o seguinte cambio de variável:

$$\tilde{s} = \frac{s - \mu}{\sqrt{\kappa_2}} = \frac{s - \mu}{\sigma} \quad (\text{A1.14})$$

Obtém-se a variável ou *sinal estandardizado* \tilde{s} , é dizer, todo sinal estandardizado terá média 0 e variância 1. Os cumulantes estandardizados, $\tilde{\kappa}_k$, são os cumulantes do sinal estandardizado.

O cumulante estandardizado de ordem 3, $\tilde{\kappa}_3$, mede a assimetria (“skewness”) da função de distribuição, visto que é nula se a função de densidade de probabilidade de s tem um eixo de simetria. Se s é centrada, o fator de assimetria (γ_1) resulta ser:

$$\gamma_1 = \tilde{\kappa}_3 = \frac{E\{s^3\}}{E\{s^2\}^{3/2}} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}} \quad (\text{A1.15})$$

Se a distribuição é simétrica: $\gamma_1=0$ (já que toda distribuição de probabilidade simétrica tem os momentos impares nulos, e em particular $\mu_3=0$). Usualmente se $\gamma_1>0$ quiere dizer que, a cauda superior da distribuição é a más baixa (média > mediana > moda) e se $\gamma_1<0$ a cauda inferior é a más baixa (média < mediana < moda).

O cumulante estandardizado de ordem 4, $\tilde{\kappa}_4$, é o fator de agudeza ou kurtosis (γ_2). Se o sinal está centrado na média:

$$\gamma_2 = \tilde{\kappa}_4 = \frac{E\{s^4\}}{E\{s^2\}^2} - 3 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\kappa_4}{\kappa_2^2} \quad (\text{A1.16})$$

O sinal da agudeza caracteriza o comportamento no infinito das densidades de probabilidade mono-modais (possuem apenas um modo). Se $\gamma_2=0$ se diz que a distribuição é meso-aguda ou *mesokurtic* (ex.: gaussiana). Se a agudeza é negativa ($\gamma_2<0$) se dize que a distribuição é sub-gaussiana ou *platykurtic*, já que tende para zero

Analogamente podem-se dizer:

$$f(s_1, s_2) = f(s_2/s_1) \cdot f(s_2) \quad (\text{A1.19})$$

Verifica-se que:

$$f(s_1/s_2) \leq f(s_1); f(s_2/s_1) \leq f(s_2); \quad (\text{A1.20})$$

é dizer, a probabilidade de que se produza s_1 , tendo-se produzido s_2 , é menor ou igual à probabilidade de que se produza s_1 isoladamente (ou a probabilidade de que se produza s_2 tendo-se produzido s_1 é menor o igual que a probabilidade de que se produza s_2). A igualdade se obtém no caso em que a geração de s_1 não dependa de s_2 (ou que a geração de s_2 no dependa de s_1); é dizer, quando s_1 e s_2 sejam estatisticamente independentes.

De (A1.18), (A1.19) e (A1.20) deduz-se que:

$$f(s_1, s_2) \leq f(s_1) \cdot f(s_2) \quad (\text{A1.21})$$

Verificando-se a igualdade quando s_1 e s_2 sejam estatisticamente independentes. Neste caso a igualdade é óbvia já que indica que a probabilidade da geração de um par (s_1, s_2) depende de que se produza o valor de s_1 e o valor de s_2 , isoladamente.

Em general, p sinais s_1, \dots, s_p são estatisticamente independentes se e unicamente se:

$$f(s_1, \dots, s_p) = \prod_{i=1}^p f(s_i) \quad (\text{A1.22})$$

As mesmas definições e conclusões podem ser aplicadas em variáveis aleatórias contínuas, sem ter mais que substituir em PA2.1 e em (A1.17) as somatórias por integrais.

Suponha-se uma variável aleatória p -dimensional, $s^T = (s_1, \dots, s_p)$ está centrada em sua média, é dizer $E\{s_i\} = 0$ para todo i , do mesmo modo que em no caso escalar (A1.9).

Logo, pode-se definir uma função característica conjunta de p variáveis aleatórias s_i , como:

$$m(\mathbf{v}) \triangleq E \left\{ e^{j \sum_i v_i s_i} \right\} = E \left\{ e^{j \mathbf{v}^T \mathbf{x}} \right\} \quad (\text{A1.23})$$

onde $\mathbf{v}^T = (v_i)$. Se as componentes s_i do vetor aleatório s admitem uma densidade de probabilidade conjunta $f(s)$ então a função geradora de momentos de s vem dada por a transformada de Fourier de aquela densidade:

$$m(\mathbf{v}) = \int_{\mathbb{R}^p} e^{j \mathbf{v}^T \mathbf{u}} \cdot f_s(\mathbf{u}) \cdot d\mathbf{u} \quad (\text{A1.24})$$

Os cumulantes de ordem 2, o covariâncias, são magnitudes com dos subscritos que se podem organizar em uma matriz de covariância cujos elementos são da forma:

$$\begin{aligned} Cum[s_i, s_j] &= \kappa_{s_{ij}} = \kappa_{s,2} = E\{\mu_1(s_1)\mu_1(s_2)\} = \\ &= E\{s_1, s_2\} - \mu_1(s_1)\mu_1(s_2) = \mu_1(s_1, s_2) - \mu_1(s_1)\mu_1(s_2) \end{aligned} \quad (\text{A1.25})$$

A expressão anterior indica diversas formas de representar os cumulantes (neste caso de ordem 2). Acostuma-se utilizar uma ou outra forma dependendo do contexto.

Os cumulantes de ordem 3 são:

$$Cum[s_1, s_2, s_3] = \kappa_{s,3} \quad (\text{A1.26})$$

Os cumulantes de ordem 4 são:

$$\begin{aligned} Cum[s_1, s_2, s_3, s_4] &= \kappa_{s,4} = E\{\mu_1(s_1)\mu_1(s_2)\mu_1(s_3)\mu_1(s_4)\} - \\ &\quad - E\{\mu_1(s_1)\mu_1(s_2)\}E\{\mu_1(s_3)\mu_1(s_4)\} - \\ &\quad - E\{\mu_1(s_1)\mu_1(s_3)\}E\{\mu_1(s_2)\mu_1(s_4)\} - \\ &\quad - E\{\mu_1(s_1)\mu_1(s_4)\}E\{\mu_1(s_2)\mu_1(s_3)\} - \end{aligned} \quad (\text{A1.27})$$

Em general, os cumulantes de ordem r se relacionam com os momentos segundo a fórmula de Leonov e Shirayev [86]:

$$Cum[s_1, \dots, s_r] = \sum (-1)^{k-1} (k-1)! E \left\{ \prod_{i \in v_1} s_i \right\} \cdot E \left\{ \prod_{j \in v_2} s_j \right\} \cdots E \left\{ \prod_{k \in v_p} s_k \right\} \quad (A1.28)$$

onde a somatória se estende para todos os conjuntos $\{v_1, \dots, v_p\}$, $1 \leq p \leq r$ que formam uma partição de $\{1, \dots, r\}$, e k é o número de elementos que compõem a partição.

Denomina-se ordem da estatística ao número de índices, r , utilizados para representar os cumulantes [86].

Uma propriedade de grande importância, utilizada em diversos métodos de separação cega de sinais, indica que, vários sinais aleatórios são estatisticamente independentes, se e somente se, seus cumulantes cruzados de qualquer ordem são nulos. É conveniente notar que, tornar nulos os cumulantes cruzados de ordem 2 (covariâncias) é condição necessária mais não suficiente (salvo no caso de distribuições gaussianas) para a independência estatística, visto que, os cumulantes de ordem superior também devem ser nulos. Quando as covariâncias entre sinais são nulas, se diz que os sinais estão descorrelacionados.

A1.3 Funções de densidade de probabilidade de sinais aleatórios.

A1.3.1 Sinais com densidade de probabilidade uniforme.

Um sinal aleatório se diz que tem uma densidade de probabilidade uniforme, na faixa $[s_m, s_M]$, se todo valor do sinal compreendido nessa faixa tem igual probabilidade de produzir-se (Figura A1.1).

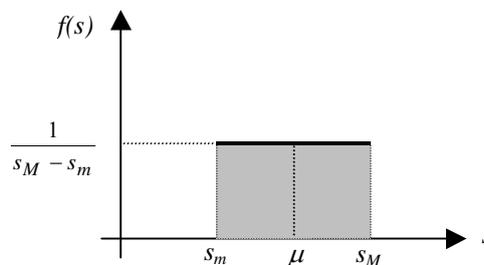


Figura A1. 1. Função de densidade de probabilidade uniforme.

Tendo em conta que a área total baixo a curva de probabilidade (área sombreada da Figura A1.1) tem que ser a unidade, se verifica que:

$$\begin{cases} f(s) = \frac{1}{s_M - s_m} & \text{se } s_m \leq s \leq s_M \\ f(s) = 0 & \text{em outro caso} \end{cases} \quad (\text{A1.29})$$

A função geradora de momentos de uma distribuição de probabilidade uniforme é:

$$m(t) = \frac{e^{ts_M} - e^{ts_m}}{t(s_M - s_m)} \quad (\text{A1.30})$$

Os momentos de ordem par são:

$$\mu_k = \frac{1}{k+1} \left(\frac{s_M - s_m}{2} \right)^k \quad (\text{A1.31})$$

Assim, a média e a variância são respectivamente:

$$\begin{aligned} \mu &= \frac{s_M + s_m}{2} \\ \sigma^2 &= \frac{(s_M - s_m)^2}{12} \end{aligned} \quad (\text{A1.32})$$

e o fator de agudeza (kurtosis):

$$\gamma_2 = -\frac{6}{5} \quad (\text{A1.33})$$

Por analogia com a função de distribuição gaussiana se diz que, uma função de distribuição é super-uniforme se $f(s_m) > \mu$ e $f(s_M) > \mu$ sub-uniforme se $f(s_m) < \mu$ e $f(s_M) < \mu$.

A1.3.2 Sinal com densidade de probabilidade gaussiana (normal).

A função de distribuição gaussiana também se denomina *distribuição normal* devido a que é o modelo ao qual se ajustam muitas distribuições que aparecem em

diversos campos científicos e técnicos. Por exemplo, a distribuição de erros relativos em medidas experimentais corresponde a uma densidade de probabilidade gaussiana.

A função de distribuição gaussiana é dada por:

$$f(s) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(s-\mu)^2/(2\sigma^2)} \stackrel{def}{=} N(\mu, \sigma) \quad -\infty < s < \infty \quad (\text{A1.34})$$

A função geradora de os momentos é:

$$m(t) = e^{\mu t + (t^2 \sigma^2 / 2)} \quad (\text{A1.35})$$

Devido à simetria desta função (ver Figura A1.2), os momentos de ordem impar são nulos e os de ordem par são dados por:

$$\mu_{2r} = \sigma^{2r} \cdot \frac{(2 \cdot r)!}{r! \cdot 2^r} \quad r = 1, 2, 3, \dots \quad (\text{A1.36})$$

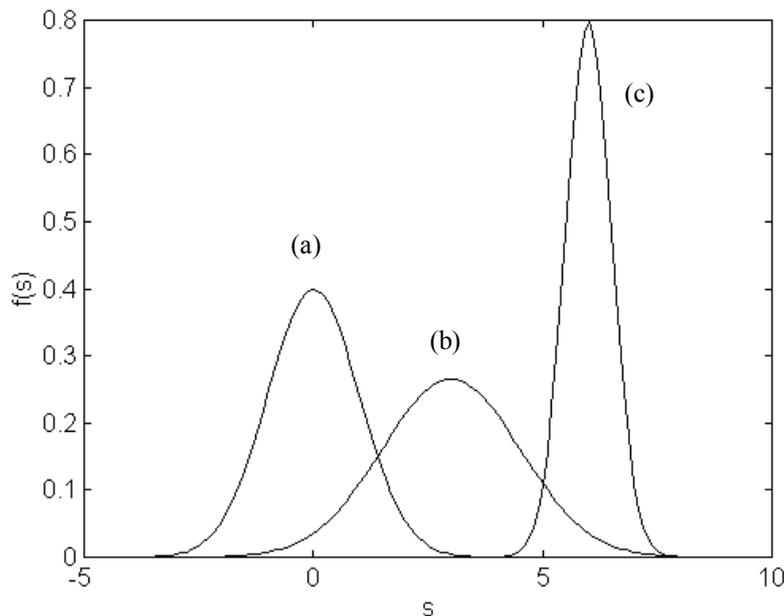


Figura A1. 2. Distribuições de probabilidade gaussianas com:

(a) $\mu=0, \sigma=1$; (b) $\mu=3, \sigma=1.5$; (c) $\mu=6, \sigma=0.5$.

A1.3.3 Função de densidade de probabilidade gamma.

Esta função se define como:

$$f(s) = (s - \gamma)^{\alpha-1} \cdot \frac{e^{-(s-\gamma)/\beta}}{\beta^\alpha \cdot \Gamma(\alpha)} \quad \text{com } \alpha > 0, \beta > 0, s > \gamma \quad (\text{A1.37})$$

onde $\Gamma(\alpha)$ é a função gamma:

$$\Gamma(\alpha) = \int_0^\infty s^{\alpha-1} \cdot e^{-s} \cdot ds \quad (\text{A1.38})$$

O parâmetro α é o fator de forma e β é o fator de dispersão. Usualmente se considera $\gamma = 0$, com o qual (A1.25) se converte em:

$$\begin{cases} f(s) = \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} \cdot s^{\alpha-1} \cdot e^{-s/\beta} & \text{com } \alpha > 0, \beta > 0, 0 \leq s \leq \infty \\ f(s) = 0 & \text{com } s < 0 \end{cases} \quad (\text{A1.39})$$

Na Figura A1.3 se mostram algumas funções de probabilidade gamma.

A função geradora de momentos resulta ser:

$$m(t) = (1 - \beta \cdot t)^{-\alpha} \quad (\text{A1.40})$$

De forma que, os momentos respeito à origem são:

$$\begin{aligned} \mu &= \mu_1' = \alpha \cdot \beta \quad (\text{media}) \\ \mu_2' &= \alpha \cdot (\alpha + 1) \cdot \beta^2 \\ \mu_3' &= \alpha \cdot (\alpha + 1) \cdot (\alpha + 2) \cdot \beta^3 \\ \mu_4' &= \alpha \cdot (\alpha + 1) \cdot (\alpha + 2) \cdot (\alpha + 3) \beta^4 \end{aligned} \quad (\text{A1.41})$$

os momentos centrais resultam ser [86]:

$$\begin{aligned}\sigma^2 &= \mu_2 = \alpha \cdot \beta^2 \quad (\text{variância}) \\ \mu_3 &= 2 \cdot \alpha \cdot \beta^3 \\ \mu_4 &= 3 \cdot \alpha \cdot (\alpha + 2) \beta^4\end{aligned}\tag{A1.42}$$

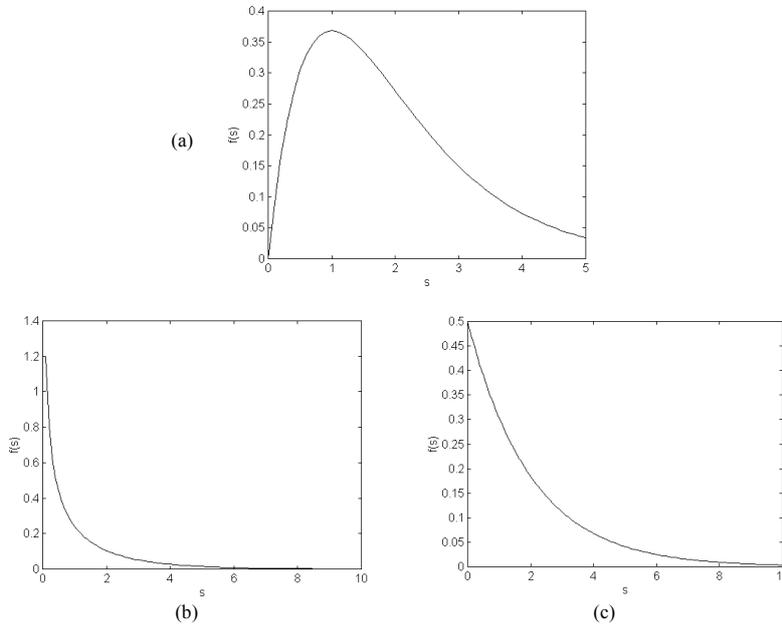


Figura A1. 3. Distribuições de probabilidade gamma com:
 (a) $\beta=1, \alpha=2$; (b) $\beta=2, \alpha=0.5$; (c) $\beta=2, \alpha=1$ (exponencial).

A função de densidade de probabilidade gamma tem sido de útil em descargas parciais para modelar formas de pulsos ultra-rápidos [10]. Outro exemplo, no campo da fonologia, já que os sinais de voz se aproximam notavelmente para um tipo destas funções que se obtém fazendo $\alpha = \frac{1}{2}$ [86], tendo em conta que $\Gamma(1/2)=\pi^{1/2}$, a expressão (A1.27) se torna em:

$$f(s) = \left(\frac{1}{\beta \cdot \pi \cdot s} \right)^{1/2} \cdot e^{-s/\beta} \quad 0 \leq s < \infty \tag{A1.43}$$

A densidade de probabilidade de o sinal de voz, a diferencia da dada por (A1.31, Figura A1.3b) é simétrica (bilateral), segundo se mostra na Figura A1.4, com a finalidade de seguir-se conservando a área baixo a função de densidade de probabilidade ($-\infty < s < \infty$) seja a unidade, (A1.31) deve tornar-se em:

$$f(s) = \frac{1}{2} \cdot \left(\frac{1}{\beta \cdot \pi \cdot |s|} \right)^{1/2} \cdot e^{-|s|/\beta} \quad -\infty \leq s < \infty \quad (\text{A1.44})$$

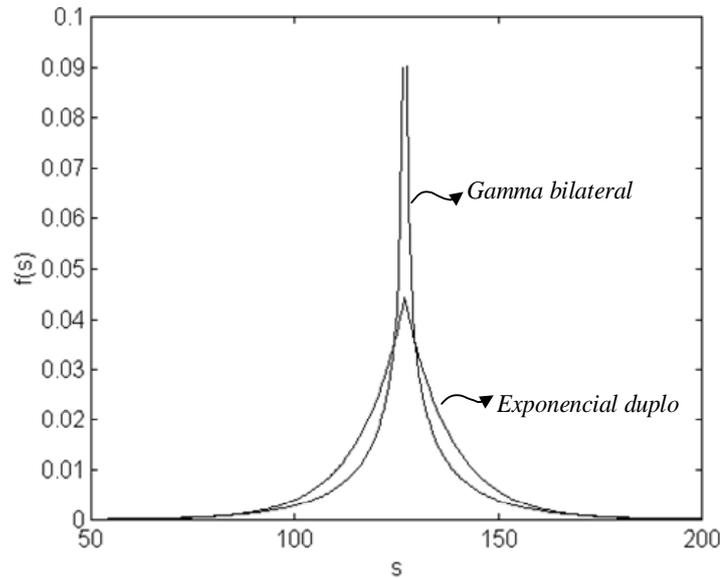


Figura A1. 4. Densidade de probabilidade gamma bilateral e duplo exponencial com $\mu=127$ e $\sigma=16$.

A variância da função (A1.32) resulta ser:

$$\sigma^2 = \frac{3}{4} \cdot \beta^2 \quad (\text{A1.45})$$

onde a densidade de probabilidade gamma bilateral em função da média (μ) e da variância (σ^2) é:

$$f(s) = \left(\frac{\sqrt{3}}{8 \cdot \pi \cdot \sigma \cdot |s - \mu|} \right)^{1/2} \cdot e^{-(\sqrt{3} \cdot |s - \mu| / (2\sigma))} \quad -\infty \leq s < \infty \quad (\text{A1.46})$$

A expressão anterior (A1.34) coincide com a dada por Rabiner em [86].

Duas propriedades de interesse são as seguintes: Se s_1 e s_2 são duas variáveis aleatórias independentes, com função da distribuição gamma e com valores paramétricos α : α_1 e α_2 , e valores de β comuns: $\beta = \beta_1 = \beta_2$, então se verifica:

$s_1 + s_2$ é uma distribuição gamma com igual β e com $\alpha = \alpha_1 + \alpha_2$.

$\frac{s_1}{s_1 + s_2}$ y $\frac{s_2}{s_1 + s_2}$ tem distribuição beta com parâmetros α_1 e α_2 .

A1.3.4 Função densidade de probabilidade exponencial e duplo exponencial (ou de Laplace).

Se na expressão (A1.27) da função de probabilidade gamma se faz $\alpha=1$ se obtém a função de probabilidade exponencial (Figura A1.3c):

$$\begin{cases} f(s) = \frac{1}{\beta} \cdot e^{-s/\beta} & \text{si } s \geq 0 \\ f(s) = 0 & \text{si } s < 0 \end{cases} \quad (\text{A1.47})$$

cuja média, variância, assimetria, fator de agudeza e cumulantes resultam ser respectivamente:

$$\mu = \beta; \quad \sigma = \beta; \quad \gamma_1 = 2; \quad \gamma_2 = 6; \quad \kappa_r = (r-1)! \sigma^r \quad (\text{A1.48})$$

Os sinais de voz também se acostuma aproximar com uma distribuição de probabilidade duplo exponencial (também denominada distribuição de Laplace), consistente em fazer a função exponencial (A1.35) simétrica:

$$f(s) = \frac{1}{2\beta} \cdot e^{-|s|/\beta} \quad (\text{A1.49})$$

onde $\mu=0$ e $\sigma = \sqrt{2} \cdot \beta$, então (A1.37) pode expressar-se como:

$$f(s) = \frac{1}{\sqrt{2}\sigma} \cdot e^{-\sqrt{2} \cdot |s-\mu|/\sigma} \quad (\text{A1.50})$$

que coincide com a expressão dada por Rabiner em [59].

Uma forma normalizada desta distribuição se obtém fazendo $\mu=0$ e $\sigma = \sqrt{2}$, então (A1.38) se converte na primeira lei do erro de Poisson:

$$f(s) = \frac{1}{2} e^{-|s|} \quad (\text{A1.51})$$

O k-ésimo momento central resulta ser:

$$\begin{aligned} \mu_k(x) &= 0 \quad (\text{si } k \text{ es impar}) \\ \mu_k(x) &= k! \quad (\text{si } k \text{ es par}) \end{aligned} \tag{A1.52}$$

e o fator de agudeza:

$$\gamma_2 = 3 \tag{A1.53}$$

logo, a distribuição resulta ser super-gaussiana.

Na Figura A1.4 se mostra uma função de probabilidade duplo exponencial superposta com uma função gamma bilateral. A densidade de probabilidade de um sinal de voz se encontra entre as curvas [116].

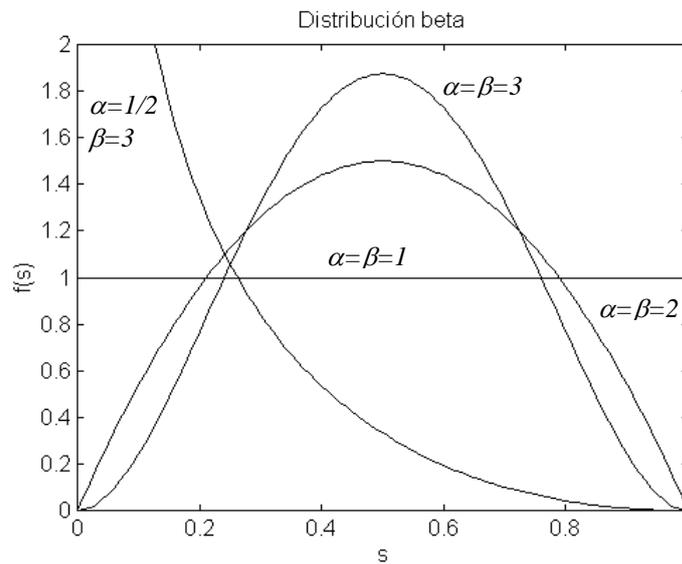


Figura A1. 5. Funções de probabilidade beta.

A1.3.5 Função densidade de Probabilidade

A função de densidade de probabilidade para um sinal aleatória tipo beta é dada por:

$$f(s) = \frac{1}{B(\alpha, \beta)} \cdot \frac{(s - s_m)^{\alpha-1} (s_M - s)^{\beta-1}}{(s_M - s_m)^{\alpha+\beta-1}} \quad \begin{matrix} s_m \leq s \leq s_M \\ \alpha > 0, \beta > 0 \end{matrix} \tag{A1.54}$$

onde:

$$B(\alpha, \beta) = \int_0^1 s^{\alpha-1} (1-s)^{\beta-1} \cdot ds = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)} \tag{A1.55}$$

Caso em que se faz a troca de variável:

$$x = \frac{s - s_m}{s_M - s_m} \quad (\text{A1.56})$$

se obtém a forma original, isto é:

$$f(x) = \frac{1}{B(x, \beta)} \cdot x^{\alpha-1} (1-x)^{\beta-1} \quad 0 \leq x \leq 1 \quad (\text{A1.57})$$

Se $\alpha = \beta$ obtém-se uma família de densidades de probabilidade beta, simétricas (Figura A1.5).

Se o sinal está estandardizada segundo (A1.36), o momento r-ésimo com respeito a 0 é:

$$\mu'_r = \frac{\alpha^{[r]}}{[\alpha + \beta]^{[r]}} \quad (\text{A1.58})$$

onde $e[r] = e(e+1)\dots(e+r-1)$ é o fatorial ascendente.

Em particular, a média (μ) e a variância (σ^2) resultam ser:

$$\begin{aligned} \mu &= \frac{\alpha}{\alpha + \beta} \\ \sigma^2 &= \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 \cdot (\alpha + \beta + 1)} \end{aligned} \quad (\text{A1.59})$$

A função de densidade beta proporciona um modelo adequado para sinais aleatórios, s , que têm limites inferior e superior fixos e finitos $[s_m, s_M]$.

Apêndice B Publicações em Congressos

Vários artigos, relacionados ao tema desta tese, foram aceitos para publicação e para apresentação em congressos e simpósios importantes. A seguir, realiza-se uma descrição de cada uma destes trabalhos.

B1.1 Artigos em Congressos

- Título:** Análise de Componentes Independentes para Extração Cega de Ruídos dos Sinais de Descargas Parciais em Equipamentos de Alta Tensão [5].
Autores: Walter M. Huamán Cuenca, Alain François Levy, José Manoel de Seixas, Sandoval Carneiro Jr
Congresso: XVII SNPTEE, Uberlândia, Minas Gerais, 2003.
Resumo: No presente trabalho é apresentado uma aproximação para o problema de remoção de ruídos em sinais fonte de descargas parciais (DP), usando os recentes e consagrados conceitos de separação cega de fontes (BSS) e a análise de componentes independentes (ICA).
- Título:** Análise de Componentes Independentes para Extração Cega de Ruídos Gaussianos dos Sinais de Descargas Parciais em Equipamentos de Alta Tensão [87].
Autores: Walter M. Huamán Cuenca, Alain François Levy, José Manoel de Seixas, Sandoval Carneiro Jr
Congresso: Proceedings of the VI Brazilian Conference on Neural Networks, São Paulo – SP, Brazil, June 2–5, 2003.
Resumo: *In this paper is presented an approach for the problem of noise removal from partial discharges (PD) signals, using recent and well-established concepts of Blind Source Separation (BSS) and Independent Component Analysis (ICA). It is showed that the ICA methods can be applied to PD signals characterization, arriving to a satisfactory solution for separation and/or elimination of noise from original signals. The original signals can be some defects coming from aged stator bars windings or from several other kinds of defects coming from insulated systems inside power transformers and hydrogenerators, among others.*
- Título:** Sistema Caracterizador Híbrido para Identificar Descargas Parciais em Transformadores de Potência [4].
Autores: Walter M. Cuenca, Alain F. Levy, José M. Seixas
Congresso: Proceedings of the V Brazilian Conference on Neural Networks, Rio de Janeiro - RJ, Brazil, April 2-5, 2001.

Resumo: *This partial discharge detection and localization in power transformers is addressed by the use of a simplified experimental model (for partial discharge generation) and a hybrid signal processing technique. The experimental model is obtained using test cells for which partial discharge signals are generated. Experimental data were collected in laboratory by extensive measurement on those test cells. The detection efficiency for partial discharges is shown to be above 97 %, using optimal lineal filtering and neural processing.*

4. **Título:** Otimização por Algoritmos Genéticos e Simulated Annealing após Tratamento de Dados por Sistemas Neuro-Fuzzy [39].
Autores: Walter M. Cuenca, Sandoval Carneiro Jr, Luis A. Cheim, Luis P. Calôba
Congresso: Proceedings of the V Brazilian Conference on Neural Networks, Rio de Janeiro - RJ, Brazil, April 2-5, 2001.
Resumo: *This A great part of the scientific applications and of the engineering problems, they can be considered as search and optimization problems. Due to need to obtain better results in these problems, several tools were developed looking for the optimal global. Search methods have been developed based on natural and physical processes. The most usual are: Genetic Algorithm (AG) and Simulated Annealing (SA). In this paper applies AG and SA separately, with the objective of reinforcing the performance of the Neuro-Fuzzy system in a function approximation problem, whose improvement and efficiency shown the results.*

5. **Título:** Análise de Componentes Principais para Identificar Descargas Parciais em Transformadores de Potência [3].
Autores: Walter M. Cuenca, José M. Seixas, Alain F. Levy
Congresso: VI Simpósio Brasileiro de Redes Neurais SBC/SIG - RJ, Brasil, Novembro 22-25, 2000.
Resumo: *This work presents a preprocessing analysis of partial discharges data in power transformers. These data were measured starting from experimental models. The pre-processing method uses the principal component analysis (PCA), aiming at reducing the dimension of the original data space, using only seven principal components, and classification efficiency better than 93% was achieved.*

Referências Bibliográficas

- [1] Bartnicas R., *Corona Discharge Processes in Voids*. In: R. Bartnicas, E. McMahon, Engineering Dielectrics: Corona Measurement and Interpretation Vol. I, Chapter 2, Philadelphia, ASTM, 1979.
- [2] Harmeling S. et. al., "Analyzing ICA by Injecting Noise ", *harmeli@fist.fhg.de*.
- [3] Walter M., Jose M. Seixas, Alain F. "Análise de Componentes Principais para Identificar Descargas Parciais em Transformadores de Potência". In: *VI Brazilian Symposium on Neural Networks*, Aplicações Cuenca46, Rio de Janeiro, Nov. 2000.
- [4] Walter M. Cuenca, Alain F. Levy, José M. Seixas, "Sistema Caracterizador Híbrido para Identificar Descargas Parciais em Transformadores de Potência", *Proceedings of the V Brazilian Conference on Neural Networks*, pp. 169-174, PUC, Rio de Janeiro - RJ, Brazil, April 2-5, 2001.
- [5] Walter M. Cuenca, Alain F. Levy, José M. Seixas e Sandoval Carneiro Jr, "Análise de Componentes Independentes para Extração Cega de Ruídos dos Sinais de Descargas Parciais em Equipamentos de Alta Tensão", *XVII SNPTEE*, 2003.
- [6] IEC Publication-60270, High-voltage Test Technique-Partial Discharge Measurement. Third Edition, 2000.
- [7] IEEE P1434, *Guide to the Measurement of Partial Discharges in Rotating Machinery*, Electrical Machinery Committee of the IEEE Power Society, 1997.
- [8] Bartnikas R., "Partial Discharge Their Mechanism, Detection and Measurement", *IEEE Trans. Elec. Insul.* Vol. 9, no. 5, pp. 763-808, Oct., 2002
- [9] Steiner J. P., Partial discharge- Part IV:Commercial PD testing, *IEEE Elec. Insul. Magazine*, Vol.7, No.1, pp. 20–33, 1991.
- [10] Okamoto T. and Tanaka T., Novel Partial Discharge Measurement Computer-Aided Measurement Systems, *IEEE Trans. Electr. Insul.* Vol. 21, No. 6, 1015–1016, 1986.
- [11] Malik N. H. and Al-rainy A. A., Statically Variation of dc corona pulse amplitudes in point-to-plane air gaps, *IEEE Trans. Electr. Insul.* Vol. 22, No. 6, 825–829, 1987.
- [12] Glahn P., von and Van Brunt R. J., Continuous Recording and Stochastic Analysis of PD, *IEEE Trans. Diel. Elec. Insul.* Vol.2, No.4, pp. 590–601, 1995.
- [13] Gulski E., Meijer S., Rutgers W. R., and Brooks R., "Recognition of PD in Sf 6 Insulation Using Digital Data Processing", In: *Conf. Elec. Insul. Diel. Phenomena*, IEEE, New York, pp. 577–580, 1996.
- [14] Gulski E., "Digital Analysis of Partial Discharge", *IEEE Trans. Diel. and Elect. Insul.*, Vol. 2, pp 822-837, 1995.

- [15] Krivda A., Automated recognition of partial discharges, *IEEE Trans. Dielect. Insul.* Vol. 2 No. 5, 796-821, 1995.
- [16] Van Brunt R. J. and Kulkarni S. V., “Method for Measuring the Stochastic Properties of Corona and Partial-Discharge Pulses”, *Rev. Sci. Instrum.* Vol.60, No. 9, pp. 3012–3023, 1989.
- [17] Van Brunt R. J. and Leep D. “Characterization of Point-Plane Corona Pulses in SF 6”, *J. Appl. Phys.* Vol. 52, No.11, pp. 6588–6600, 1981.
- [18] Van Brunt R. J. and Kulkarni S. V., “Stochastic Properties of Trichel-pulse Corona: A non-Markovian Random Point Process”, *Phys. Rev. A* 42 No.8, pp. 4908–4932, 1990.
- [19] Kumar Senthil S., Development of a phase Resolved Pulse Height Analyser and its Application to Partial Discharge Measurements, MSc. thesis, February 1997.
- [20] Van Brunt R. J. and Cernyar E. W. “System for Measuring Conditional Amplitude, Phase, or Time Distributions of Pulsating Phenomena”, *J. Res. Natl. Inst. Stand. Technol.* Vol. 97, No.6, pp. 635–672, 1992.
- [21] Hucker T., Glahn P. von, Kranz H. -G., and Okamoto T., “A Standardised Computer Data File Format for Storage, Transport, and Off-line Processing of Partial Discharge Data”, In: *Ninth International Symposium on High Voltage Engineering, Inst. High Voltage Eng. Graz, Austria*, pp. 5613, 1995.
- [22] Kumar Senthil S. et al., Limitations and Errors in Partial Discharge Measurements Scope for Automation, 12th International Symposium on High Voltage Engineering, Bangalor, India, 6-66:1047-1050, August 2001
- [23] Kumar Senthil S., Development of Novel Techniques for Partial Discharge Measurement, Analysis and Representation, DP thesis, February 2002.
- [24] Kumar Senthil S. et al., Influence of Detectors on Measurements of Partial Discharge. In: *12th International Symposium on High Voltage Engineering, Bangalor, India*, 6-66:1047-1050, August 2001
- [25] Engineering Dielectrics. *Electrical Insulation Liquids*, Vol.III, ASTM, USA,1994.
- [26] Stratton J. A., *Electromagnetic Theory*, McGraw Hill book Company, NY, 1941.
- [27] Kraus, John D., *Eletromagnetismo*, Guanabara Dois, 1978
- [28] Bartnikas R., “Use of Multichannel Analyzer for Corona Pulse-Height Distribution Measurements on Cables and Other Electrical Apparatus”, *IEEE Trans. Instr. Meas.*, Vol. 22, pp. 403–407, 1973.
- [29] Raja K., Devaux F., Lelaidier S., "Recognition of Discharges sources Using UHF PD Signatures", *IEEE Elec. Insul. Magazine*, vol. 18, No.5,pp. 8–14, Sep/Oct. 2002.
- [30] Pascoli Ger, “Monitoring System als Instrume Modeone Instandhaltungsstrategien der Energietechnik“. In: *IT'S, TIME*, Fev., 2001.
- [31] Volgelsang R., Farr T., “ Ageing of Micacious Insulation Materials Failure Mechanisms of the Electrical Insulation of Rotating Machines”
- [32] Huamán Cuenca, W. *Aplicação de Sistemas inteligentes no Reconhecimento de Padrões de DPs em Transformadores de Potência*. Tese de Mestrado, Programa de Engenharia Elétrica, COPPE/UFRJ, Rio de Janeiro, RJ, 1998.

- [33] Bartnikas R, “A Commentary on Partial Discharge Measurement and Detection”, *IEEE Trans. Elec. Insul.*, Vol.22, No.5, pp. 629–653, 1987.
- [34] Lemke E., Schmiegel P., “Progress in PD Prove Measuring Technique”, *Proc. of 7th. Desden*, pp. 72.0, 1991.
- [35] Latini M., et al. "Medição de descargas Parciais em Hidrogerador da Usina de Itaipu - Uma Avaliação Experimental das Limitações, Dificuldades de Medição e Interpretação de resultados". In: *XV SNPTEE, Foz do Iguaçu - Paraná*, 1999.
- [36] Levy, A. et al., "Avaliação do Efeito de Propagação de Memória no Desempenho de Materiais Elétricos - Medição de DP em Ultra-Alta Frequência ao Ar Ambiente e Óleo Isolante", *Relatório Técnico ACET-095/97*, Fev. 1997.
- [37] Levy, A. et al., "Utilização do Sistema de Descargas Parciais ICM System da Power Diagnostix Systems", *Procedimento Específico CEPEL*, No. 3530-3007-070, jan. 1997.
- [38] Cheim, L, Cuenca, W and Varricchio, S. "A simple Mamdani Adaptive Neuro-Fuzzy Inference System (MANFIS)", *manuscrito Submetido à IEEE trans. on Fuzzy System*, 1998.
- [39] Walter M. Cuenca, Sandoval Carneiro Jr, Luis A. Cheim, Luis P. Calôba, "Otimização por Algoritmos Genéticos e Simulated Annealing após Tratamento de Dados por Sistemas Neuro-Fuzzy", *Proceedings of the V Brazilian Conference on Neural Networks*, pp. 175-180, – PUC, Rio de Janeiro – RJ, Brazil, April 2-5, 2001.
- [40] Duda, R, Hart, P and Stork, D. *Pattern Classification*, 2nd. ed. 1997.
- [41] Soares F. W., Classificação do Ruído por Navios Usando Redes Neurais, Tese Doutorado, COPPE/UFRJ 2001.
- [42] Jang R., Sun C.T., Muzitani E., *Neuro-Fuzzy and Soft Computing*, Prentice Hall, USA 1997.
- [43] Chin-Teng Lin, C. S. George Lee: *Neural Fuzzy Systems*, Prentice Hall, Upper Saddle River, NJ 07458, 1996.
- [44] Cichocki, A., & Unbehauen, R. "Robust Neural Networks With On-Line Learning for Blind Identification and Blind Separation of Sources". *IEEE Transactions on Circuits and Systems*, 43 (11), 894–906, 1996.
- [45] Nauck D., Klawonn F., Kruse R., *Neuro-Fuzzy Systems*, John Wiley Sons, Ltd, New York, 1997.
- [46] Pereira J., et al., “Identificação de Estratégias de Exploração de redes de Distribuição de Energia Elétrica Utilizado Simulated Annealing”. In: *4o. Encontro Luso-Afro-Brasileiro de Planejamento e Exploração de Redes de Energia*. ELAB'99.
- [47] Mendonça Paulo R., *Novos Algoritmos de Simulated Annealing*. Tese, COPPE UFRJ, No.192, 1997.
- [48] Tanumaro, J., "Motivação, Fundamentos e Aplicações de Algoritmos Genéticos". In: *Segundo Congresso Brasileiro de Redes Neurais*, 1995.
- [49] Haykin Simon, *Neural Network: A Comprehensive Foundation*, 2da. Edition ISBN 0-02-352761-7 Prentice Hall, 1999.

- [50] Papoulis, A., *Probability, Random variables, and Stochastic Processes*. McGraw-Hill, 3rd ed., New York: McGraw-Hill, 1991.
- [51] Whalen A., *Detection of Signal in Noise*, Academic Press, Inc. New York 1971.
- [52] Peyton Pebbles, *Probability, Random Variables, and Random Signal Principles*. ISBN 0-07-112782-8, 3rd Ed, McGraw-Hill 1993.
- [53] Cichocki, A., & Amari S, *Adaptative Blind Signal and Image Processing: Learing Algorithms and Aplications*. John Wiley & Sons, Ltd. New york, 2002.
- [54] Rosipal R., Girolami M. Trejo J. L., Cichocki A., “Kernel PCA for Feature Extraction and De-Noiseing in Nonlinear Regression”, *Nerual Comput & Applic*, Vol.10, pp. 231-243, Springer-Verlag, 2001.
- [55] Friedman, J. “Exploratory Projection Pursuit”. *Journal of the American Statistical Association*, 82 (397), 249–266, 1987.
- [56] Akuzawa T., "Extended Quasi-Newton Method for the ICA". In: *Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation*, pp. 521-525, 2000.
- [57] Amari, S.-I., Cichocki, A., & Yang, H., “A New Learning Algorithm for Blind Source Separation”, *Advances in Neural Information Processing Systems*, Vol. 8. MIT Press: Cambridge, MA, pp. 757–763, 1996.
- [58] Amari S., T-P. Chen, and A. Cichocki, "Non-holonomic Constraints in Learning Algorithms for Blind Source Separation", *preprint* 1997.
- [59] Bell, A., & Sejnowski, T. “An Information-Maximization Approach To Blind Separation And Blind Deconvolution”. *Neural Computation*, Vol. 7, pp. 1129–1159, 1995.
- [60] Hyvärinen, A., “Independent Component Analysis in the Presence of Gaussian Noise by Maximizing Joint Likelihood”. *Neurocomputing*, 22, 49–67 1998a.
- [61] Friedman, J. H., & Tukey, J. W. “A Projection Pursuit Algorithm for Exploratory Data Analysis”, *IEEE Transactions of Computers*, c-23 (9), 881–890, 1974.
- [62] Comon, P., “Independent Component Analysis -A New Concept?” *Signal Processing*, 36, 287–314, 1994.
- [63] Cardoso, J.-F., “Infomax and Maximum Likelihood for Source Separation”. *IEEE Letters on Signal Processing*, Vol. 4, pp. 112–114, 1997.
- [64] Attias H., “Independent factor Analysis”, *Neural Computation*, in press, 1-34.
- [65] Ikeda Shiro, “Factor Analysis Preprocessing for ICA”, shiro@brain.riken.go.jp
- [66] Malthouse C. E.,”Limitations of Nonlinear PCA as Performed with Generic Neural Networks”, *IEEE Tans. Neural Netwoks*, Vol. 9, No. 1, January 1998.
- [67] William W. H., “Nonlinear Parincipal Component Analysis by Neural Networks”, <http://www.ocgy.ubc.ca/projects/clim.pred> (visitado em 2001).
- [68] Kramer M. A., “Nonlinear Principal Component Analysis Using Autoassociative Neural Networks”, *AICe Journal*, Vol 37, N0. 2, pp. 233-243, 1991.
- [69] Donoho, D. L., Johnstone, I. M., Kerkyacharian, G., & Picard, D., “Wavelet shrinkage: asymptopia?”, *Journal of the Royal Statistical Society, Series B*, 57, 301–337, 1995.

- [70] Mallat, S. G., “A Theory For Multiresolution Signal Decomposition: The Wavelet Representation”, *IEEE Transactions on PAMI*, Vol.11, pp. 674–693, 1989.
- [71] Shanmugan K.S., AM Breipoh, *Random Signal Detection, Estimate and Data Analysis*, John Wiley & Sons, 1988.
- [72] Vaseghi S. V., *Advanced Digital Signal Processing and Noise Reduction*, John Wiley & Sons, Ltd, New York, 2000.
- [73] Hyvärinen A., Karhunen J. & Oja E., *Independent Component Analysis*, John Wiley & Sons. Inc.2001
- [74] Hyvärinen, A.,“New Approximations of Differential Entropy for Independent Component Analysis and Projection Pursuit”, *Advances in Neural Information Processing Systems*, Vol. 10. Cambridge, MA: MIT Press (pp. 273–279), 1998b.
- [75] Jones, M., & Sibson, R. “What is projection pursuit?” *Journal of the Royal Statistical Society*, Series A, 150, 1–36, 1987.
- [76] Hyvärinen, A., Hoyer, P. O., & Oja, E. *Image Denoising by Sparse Code Shrinkage*. In S. Haykin & B. Kosko (Eds.), New York: IEEE Press, 2000.
- [77] Hyvärinen, A.,“The Fixed-Point Algorithm and Maximum Likelihood Estimation for Independent Component Analysis”. *Neural Processing Letters*, 10 (1), 1–5, 1999b.
- [78] Hyvärinen, A.,“Fast and Robust Fixed-Point Algorithms For Independent Component Analysis”. *IEEE Transactions on Neural Networks*, 10 (3), 626–634, 1999a.
- [79] Hyvärinen A. "A Fast Fixed-Point Algorithm for Independent Component Analysis," *Neural Computation*, Vol. 9, pp. 1483-1492, 1997.
- [80] Hyvärinen, A.,”Sparse Code Shrinkage: Denoising of Nongaussian Data by Maximum Likelihood Estimation”, *Neural Computation*, Vol.11, No. 7, 1739–1768, 1999d.
- [81] Cardoso J-F. and Souloumiac A., "Blind Beamforming for Non Gaussian Signals." In: *IEE Proceedings –F*, Vol. 140, pp. 362-370, 1993.
- [82] Gómez Herrero G., *Mismatch Negativity Detection in EEG recordings using Wavelets*, Master of Science Thesis, Tampere University of Technology, April, 2003.
- [83] Hyvärinen, A.,“Gaussian Moments for Noisy Independent Component Analysis”. *IEEE Signal Processing Letters*, Vol. 6, No. 6, 145–147, 1999c.
- [84] Akuzawa T. and Murata N., "Multiplicative Non-holonomic/Newton like algorithm", *Chaos, Solitons & Fractals*, vol. 12, pp. 785-793, 2001.
- [85] Murillo Fuentes J., *Aplicación de Técnicas de Separación Ciega de Fuentes al Procesado de Señal y Comunicaciones*, Tesis Doctoral, Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Leganés, 2001.
- [86] Beatriz P. C. Nuevos Algoritmos para Separación Ciega de Fuentes Utilizando Métodos Geométricos, Tesis Doctoral, Granada, España, 1999.
- [87] Walter M. Cuenca, Alain F. Levy, José M. Seixas e Sandoval Carneiro Jr, "Análise de Componentes Independentes para Extração Cega de Ruídos

- Gaussianos dos Sinais de Descargas Parciais em Equipamentos de Alta Tensão", *Proceedings of the VI Brazilian Conference on Neural Networks*, pp. 133-138, São Paulo - SP, Brazil, June 2-5, 2003.
- [88] Kauhold M., et. al., "Interface Phenomena in Stator Winding Insulation – Changes in Design, Diagnosis, and Service Experience", *IEEE Elec. Insul. Magazine*, Vol. 18, No.2, pp.27-35, March/April. 2002.
- [89] Kheirmand Akbar, Leijon Mats, "advances in Online Monitoring and localization of Partial Discharge in Large Rotating Machines", *IEEE Trans. On Energy Convert*, Vol. 19, no. 1, pp.53-59, March 2004.
- [90] Gleizer H., *Utilização de Redes Neurais para a Análise do Envelhecimento de Barras Estáticas de Hidrogeradores*, Tese de Doutorado, Programa de Engenharia Metalúrgica e de Materiais, COPPE/UFRJ, Rio de Janeiro, RJ, 2000.
- [91] Tettex Instruments, Partial Discharges Measuring Techniques
- [92] Boss S.A., Stone G.C., "Fundamental Limitations in the Measurement of Corona and Partial Discharge", *IEEE Trans. Elec. Insul.*, Vol. EI-17, pp. 143-150, 1982. in Proc. 2nd Int. Conf. Gas Discharges, IEE, London, pp. 236–238, 1972.
- [93] Champion R. L. and Doverspike L. D., *Electron Detachment Processes, in Electron-Molecule Interactions and Their Applications*, L. G. Christophorou, ed., Academic Press, Orlando, pp. 619–681, 1984.
- [94] Cavallini A., Contin A, Puletti F., "A New Approach to Diagnosis of Solid Insulation Systems Based on PD Signal Inference", *IEEE Electrical Insulation Magazine*, Vol. 19, n. 32, pp. 23-30, April 2003.
- [95] Cavallini A., Contin A, Puletti F., "Advanced PD Inference in On-Field Measurements. Part 1: Noise Rejection", *IEEE Trans. on Dielectrics and Electrical Insulation*, Vol. 10, n. 2, pp. 216-224, April 2003.
- [96] Satish L., Nazneen B., "Wavelet Denoising of Partial Discharge Signals Buried in Excessive Noise and Interference", *IEEE Trans. Dielect. Elec. Insul.*, Vol.10, No. 2, April 2003.