

Parametrização das Distribuições da
Estatística de Teste GID sob as
Hipóteses \mathcal{H}_0 e \mathcal{H}_1 via Redes Neurais
Artificiais

Alan Lima Lemes

Setembro/ 2019

Parametrização das Distribuições da Estatística de Teste GID sob as Hipóteses \mathcal{H}_0 e \mathcal{H}_1 via Redes Neurais Artificiais

ALAN LIMA LEMES

Dissertação apresentada ao Instituto Nacional de Telecomunicações, como parte dos requisitos para obtenção do Título de Mestre em Telecomunicações.

ORIENTADOR: Prof. Dr. Dayan Adionel Guimarães.

COORIENTADOR: Prof. Dr. Yvo Marcelo Chiaradia Masselli

L552p Lemes, Alan Lima.
Parametrização das Distribuições da Estatística de Teste GID sob as Hipóteses H0 e H1 via Redes Neurais Artificiais. / Alan Lima Lemes. – Santa Rita do Sapucaí, 2019.
89 p.
Orientador: Prof. Dr. Dayan Adionel Guimarães.
Coorientador: Prof. Dr. Yvo Marcelo Chiaradia Masselli.
Dissertação de Mestrado em Telecomunicações – Instituto Nacional de Telecomunicações – INATEL.
Inclui bibliografia e anexo.
1. Rádio Cognitivo 2. GID 3. Sensoriamento Espectral Cooperativo 4. Teste de Aderência. 5. Redes Neurais Artificiais. 6. Mestrado em Telecomunicações. I. Guimarães, Dayan Adionel. II. Masselli, Yvo Marcelo Chiaradia III. Instituto Nacional de Telecomunicações – INATEL. IV. Título.

CDU 621.39

FOLHA DE APROVAÇÃO

Dissertação defendida e aprovada em ____ / ____ / ____,
pela comissão julgadora:

Prof. Dr. Dayan Adionel Guimarães
INATEL

Prof. Dr. Carlos Alberto Ynoguti
INATEL

Prof. Dr. João Paulo Reus Rodrigues Leite
UNIFEI

Coordenador do Curso de Mestrado
Prof. Dr. José Marcos Câmara Brito

*“Lembre-se que as pessoas podem
tirar tudo de você, menos o seu
conhecimento.”*

Albert Einstein

Aos meus pais, Augusto e Ana.

Agradecimentos

Agradeço, primeiramente, a Deus pelo dom da vida, dando-me saúde, capacidade, inteligência e perseverança para trilhar minha caminhada.

Aos meus pais, Augusto e Ana, pelo incentivo e apoio em todas as minhas decisões. Muito obrigado por todas as oportunidades e por sempre acreditarem em mim.

Aos Professores Dayan e Yvo, por ter me orientado ao longo deste trabalho. Muito obrigado por todos os ensinamentos técnicos e comportamentais e, principalmente, pelo grande exemplo de profissionalismo.

A todos os professores e funcionários, pelos ensinamentos transmitidos que muito contribuíram com minha formação profissional.

Aos colegas do laboratório Lambda, em especial ao Lucas, pelos ensinamentos, convivência e experiências trocadas.

Ao INATEL pela excelência no ensino em telecomunicações.

Serei eternamente grato a todos que me acompanharam nesta jornada. Concluo a realização deste sonho com a certeza de que não chegaria tão longe sozinho.

Índice

Lista de Figuras	xi
Lista de Tabelas	xiii
Lista de Abreviaturas e Siglas	xv
Lista de Símbolos	xvii
Resumo	xix
Abstract	xxi
1 Introdução	1
1.1 Definição do problema	3
1.2 Trabalhos relacionados	4
1.3 Estrutura e contribuições da dissertação	7
1.4 Publicações	9
2 Sensoriamento Espectral	11
2.1 Introdução ao sensoriamento espectral	12
2.1.1 Teste hipótese binário	14
2.2 Sensoriamento espectral cooperativo centralizado utilizando fusão de dados	15
2.3 Métricas de avaliação para sensoriamento	16
2.3.1 Avaliação por meio de curvas ROC	16
2.3.2 Área sob a curva	18
2.4 Principais técnicas de sensoriamento espectral	18
2.4.1 Detecção por filtro casado	18
2.4.2 Detecção de energia	20
2.4.3 Detecção por autovalores	20
2.5 Detector robusto inspirado no coeficiente de Gini	21

3	Redes Neurais Artificiais	25
3.1	Introdução às redes neurais	25
3.2	Topologias comuns de redes e suas aplicações	27
3.2.1	Rede <i>Feedforward</i> de camada única	27
3.2.2	Rede <i>Feedforward</i> de multicamadas	27
3.2.3	Rede recorrentes	28
3.2.4	Rede convolucionais	28
3.3	Processo de aprendizagem	29
3.3.1	Algoritmos de treinamento	30
3.4	Funções de ativação	31
4	Parametrização das Distribuições da Estatística Teste GID sob \mathcal{H}_0 e \mathcal{H}_1 via ANN	33
4.1	Metodologia	34
4.2	CDFs candidatas sob hipótese \mathcal{H}_0 e \mathcal{H}_1	35
4.2.1	Distribuições de <i>Stable</i> e <i>Generalized Extreme Value</i>	38
4.3	Verificação gráfica das CDFs	39
4.4	Desenvolvimento da rede neural	39
4.4.1	Obtenção dos dados	40
4.4.2	Treinamento da rede	42
4.4.3	Avaliação do treinamento	45
5	Avaliação da Parametrização das Distribuições da Estatística Teste GID	47
5.1	Método de avaliação teórico	47
5.2	ROCs teóricas e empíricas	48
5.3	Capacidade de interpolação e extrapolação	50
5.4	Área sob a curva ROC	50
6	Conclusões e trabalhos futuros	53
6.1	Conclusões	53
6.2	Proposição para trabalhos futuros	54
	Referências Bibliográficas	55
A	Como utilizar as funções <code>NnH0.m</code> e <code>NnH1.m</code>	61

Lista de Figuras

2.1	Compartilhamento espectral entre rede primária e secundária.	13
2.2	Distribuições hipotéticas $f_0(t)$ e $f_1(t)$	14
2.3	Exemplos hipotéticos de curvas ROC.	17
2.4	Diagrama em blocos para detecção por filtro casado.	19
3.1	Topologia de rede neural <i>feedforward</i> generalizada.	27
4.1	CDFs empíricas e teóricas sob hipótese \mathcal{H}_0	40
4.2	CDFs empíricas e teóricas sob hipótese \mathcal{H}_1	41
4.3	Histogramas dos erros de mapeamento sob \mathcal{H}_0	46
4.4	Histogramas dos erros de mapeamento sob \mathcal{H}_1	46
5.1	Desempenho teórico e simulado do GID, para parâmetros presentes no treinamento da ANN.	49
5.2	Desempenho teórico e simulado do GID, para parâmetros não presentes no treinamento da ANN.	51
5.3	Histograma do erro entre AUC teóricas e empíricas.	52

Lista de Tabelas

4.1	Lista de Distribuições do MATLAB, versão R2018a	35
4.2	Classificação da distribuições sob \mathcal{H}_0	37
4.3	Classificação da distribuições sob \mathcal{H}_1	37
4.4	Estatísticas dos parâmetros das distribuições <i>Stable</i> e <i>GEV</i>	41
4.5	MSE da ANN sob \mathcal{H}_0	44
4.6	MSE da ANN sob \mathcal{H}_1	45

Lista de Abreviaturas e Siglas

5G	Quinta Geração das Redes de Comunicação
AI	Inteligência Artificial (<i>Artificial Intelligence</i>)
ANN	Rede Neural Artificial (<i>Artificial Neural Network</i>)
ASK	Chaveamento na Amplitude (<i>Amplitude Shift Keying</i>)
AUC	Área Sob a Curva (<i>Area Under the Curve</i>)
AWGN	Ruído Aditivo Gaussiano Branco (<i>Additive White Gaussian Noise</i>)
BER	Taxa de Erro de Bit (<i>Bit Error Rate</i>)
CDF	Função Densidade Cumulativa (<i>Cumulative Distribution Function</i>)
CFAR	Taxa de Falso Alarme Constante (<i>Constant False Alarm Rate</i>)
CNN	Redes Neurais Convolucionais (<i>Convolutional Neural Network</i>)
CR	Rádio Cognitivo (<i>Cognitive Radio</i>)
DARPA	Agência de Projetos de Pesquisa Avançada da Defesa (<i>Defence Advanced Research Projects Agency</i>)
ERD	Detecção pela Razão de Autovalores (<i>Eigenvalue Ratio Detection</i>)
FC	Centro de Fusão (<i>Fusion Center</i>)
FCC	Comissão Federal de Comunicação (<i>Federal Communications Commission</i>)
FFT	Transformada Rápida de Fourier (<i>Fast Fourier Transformer</i>)
GEV	<i>Generalized Extreme Value</i>
GID	Detector Inspirado no Índice de Gini (<i>Gini Index Detector</i>)
GLRT	Teste da Razão de Verossimilhança Generalizado (<i>Generalized Likelihood Ratio Test</i>)
GoF	Análise de Aderência (<i>Goodness-of-Fit</i>)
IoT	Internet das Coisas (<i>Internet of Things</i>)
LRT	Teste da Razão de Verossimilhança (<i>Likelihood Ratio Test</i>)
MED	Detecção pelo Máximo Autovalor (<i>Maximum Eigenvalue Detection</i>)
MKSS	Média da Estatística de Kolmogorov-Smirnov (<i>Mean Kolmogorov-Smirnov Statistic</i>)
MLP	Perceptrons por Multicamadas (<i>Multi Layer Perceptrons</i>)
MMED	Detecção pela Razão entre o Máximo e o Mínimo Autovalores (<i>Maximum-Minimum Eigenvalue Detection</i>)
MSE	Erro Quadrático Médio (<i>Mean Squared Error</i>)
PDF	Função Densidade de Probabilidade (<i>Probability Density Function</i>)

PU	Usuários Primários (<i>Primary Users</i>)
QoS	Qualidade de Serviço (<i>Quality of Service</i>)
RLRT	Teste de Roy de Máxima Raiz (<i>Roy's Largest Root Test</i>)
RNN	Redes Neurais Recorrentes (<i>Recurrent Neural Network</i>)
ROC	Curva Característica de Operação do Receptor (<i>Receiver Operating Characteristic</i>)
SD	Desvio Padrão (<i>Standard Deviation</i>)
SNR	Razão Sinal-Ruído (<i>Signal-to-Noise Ratio</i>)
SU	Usuários Secundários (<i>Secondary Users</i>)
WPAN	Redes Pessoais Sem Fio (<i>Wireless Personal Area Networks</i>)

Lista de Símbolos

\mathbf{A}	Matriz de amostra com elementos com distribuição de Rice
$\phi(v)$	Função de ativação
α	Parâmetro da distribuição de <i>Stable</i> que determina o formato
λ_m	Autovalores da matriz de covariância estimada
β	Parâmetro da distribuição de <i>Stable</i> que determina o formato
$\mathbb{C}^{x \times y}$	Dimensão de uma matriz complexa com x linhas e y colunas
\dagger	Conjugado transposto
\mathbf{R}_Y	Matriz de covariância
$\hat{\mathbf{R}}_Y$	Matriz de covariância estimada
δ	Parâmetro da distribuição de <i>Stable</i> que determina a localização
$\text{diag}(\cdot)$	Operador diagonal
E	Função erro
ϵ	Taxa de aprendizagem da rede neural
$\exp(\cdot)$	Operador exponencial
$\varphi(t)$	Função característica
$F_{\mathbb{X}}^{-1}(\cdot)$	Função distribuição cumulativa inversa de uma variável aleatória \mathbb{X}
$f_{\mathbb{X}}(\cdot)$	Função densidade de probabilidade de uma variável aleatória \mathbb{X}
$F_{\mathbb{X}}(\cdot)$	Função distribuição cumulativa de uma variável aleatória \mathbb{X}
γ	Parâmetro da distribuição de <i>Stable</i> que determina a escala
\mathbf{H}	Matriz de ganho complexo do canal
h	Ganho ou atenuação do canal
\mathcal{H}_1	Teste de hipóteses dado a hipótese 1
$h([n])$	Vetor com resposta ao impulso do filtro casado no instante n
\mathcal{H}_0	Teste de hipóteses dado a hipótese 0
a	Parâmetro de inclinação da função logística
X_n	n -ésima entrada da rede neural
k	Parâmetro da distribuição <i>GEV</i> que determina o formato
K	Fator de Rice
λ	Limiar de decisão
$\log(\cdot)$	Função logarítmica
m	Número de transmissores secundários
$\mathbb{E}(\cdot)$	Operador esperança matemática
μ	Parâmetro da distribuição <i>GEV</i> que determina a localização
n	Número de amostras colhidas por cada usuário secundário
NnH0	Função para determinar os parâmetro da distribuição de <i>Stable</i> via rede neural artificial
NnH1	Função para determinar os parâmetro da distribuição de <i>GEV</i> via

	rede neural artificial
\mathcal{N}	Distribuição normal
Y_n	n -ésima saída da rede neural
p	Número de transmissores primários
$\frac{\partial Y}{\partial X}$	Derivada parcial da função Y em relação a variável X
P_d	Probabilidade de detecção
$W_{n,i}$	Peso sináptico do n -ésimo neurônio na i -ésima camada oculta
P_{fa}	Probabilidade de falso alarme
$\Pr[\cdot]$	Probabilidade de ocorrência do evento
$\mathbb{R}^{x \times y}$	Dimensão de uma matriz real com x linhas e y colunas
\mathbf{s}	Vetor com potência do sinal recebido
s_{avg}	Potência média do sinal recebido
$\text{sgn}(\cdot)$	Função sinal
σ	Parâmetro da distribuição GEV que determina a escala
$s([n])$	Vetor com amostras do sinal transmitido
net_n	Soma dos pesos sináptico dos n neurônios
$s(t)$	Sinal transmitido
T	Estatística de teste
$\tan(\cdot)$	Função tangente
T_{ED}	Estatística do teste de detecção de energia
T_{GID}	Estatística do teste GID
T_{GLRT}	Estatística do teste da razão de verossimilhança generalizado
T_{MED}	Estatística do teste de detecção pelo máximo autovalor
T_{MMED}	Estatística do teste de detecção pela razão entre o máximo e o mínimo autovalores
\mathcal{U}	Distribuição uniforme
\mathbf{V}	Matriz de amostra de ruído aditivo gaussiano branco
$v_{(n)}$	Valor da entrada do n -ésimo neurônio
σ^2	Variância da distribuição
$w([n])$	Vetor com amostras de ruído aditivo gaussiano branco
$w(t)$	Ruído aditivo gaussiano branco
\mathbf{X}	Matriz de amostra dos sinais transmitidos
\mathbf{Y}	Matriz das amostras no centro de fusão
$y([n])$	Vetor com amostras do sinal recebido
$y(t)$	Sinal recebido

Resumo

Lemes, A.L. Parametrização das Distribuições da Estatística de Teste GID sob as Hipóteses \mathcal{H}_0 e \mathcal{H}_1 via Redes Neurais Artificiais. Dissertação de mestrado. Instituto Nacional de Telecomunicações. Santa Rita do Sapucaí, 2019.

A escassez espectral é um grande limitador para o avanço dos sistemas de comunicação sem fio sobretudo quando tais sistemas devem prover elevada taxa de transmissão de dados e suportar grande densidade de conexão, que é o que se espera da quinta geração das redes de telecomunicações. A tecnologia de rádio cognitivo permite utilizar de maneira oportunista e eficiente as faixas que por ventura estejam subutilizadas no espectro eletromagnético e, portanto, podem ser uma solução para o problema supracitado. Para determinar as bandas espectrais livres os rádios cognitivos utilizam uma técnica denominada sensoriamento espectral. Muitas técnicas de sensoriamento foram propostas na literatura, porém realizar a avaliação de desempenho de tais técnicas e relacioná-las com os parâmetros sistêmicos não é uma tarefa trivial. Recentemente foi proposto o teste GID (*Gini index detector*) para sensoriamento espectral cooperativo centrizado em sistemas de rádio cognitivo. Suas principais características são a baixa complexidade e a robustez frente a potências de sinal recebido e de ruído desiguais e variantes no tempo. Nesta dissertação apresentam-se os procedimentos e os resultados da análise de aderência da estatística de teste GID a diversas distribuições de probabilidade. É demonstrado que a distribuição *Stable* caracteriza adequadamente a estatística sob a hipótese \mathcal{H}_0 , enquanto a distribuição *Generalized Extreme Value* melhor se aplica a \mathcal{H}_1 . Duas redes neurais artificiais são em seguida desenvolvidas para estabelecer o mapeamento entre os parâmetros sistêmicos e os parâmetros que caracterizam tais distribuições, permitindo que cálculos teóricos do desempenho e do limiar de decisão do sensoriamento espectral sejam realizados.

Palavras-chave: rádio cognitivo, GID, sensoriamento espectral cooperativo, teste de aderência, redes neurais artificiais.

Abstract

Lemes, A.L. Parametrização das Distribuições da Estatística de Teste GID sob as Hipóteses \mathcal{H}_0 e \mathcal{H}_1 via Redes Neurais Artificiais. Dissertação de mestrado. Instituto Nacional de Telecomunicações. Santa Rita do Sapucaí, 2019.

Spectral shortage is a major constraint to the advancement of wireless communication systems especially when such systems must provide a high data rate and support high connection density, which is expected from the fifth generation of telecommunications networks. Cognitive radio technology allows opportunistic and efficient use of bands that may be underutilized in the electromagnetic spectrum and, therefore, may be a solution to the aforementioned problem. To determine free spectral bands, cognitive radios use a technique called spectral sensing. Many sensing techniques have been proposed in the literature, but performing the performance evaluation of such techniques and relating them to the systemic parameters is not a trivial task. Recently the GID (Gini index detector) test was proposed for centralized cooperative spectrum sensing on cognitive radio systems. Its main features are the low computational complexity, the robustness against unequal and dynamical noise and received signal powers. In this dissertation the procedures and the results of the goodness-of-fit of the GID test statistic are presented to diverse distributions of probability. It is demonstrated that the *Stable* distribution adequately characterizes the statistic under hipotese \mathcal{H}_0 , while the *Generalized Extreme Value* distribution best applies to \mathcal{H}_1 . Two artificial neural networks are then developed to establish the mapping between the systemic parameters and the parameters that characterize these distributions, allowing theoretical calculations of the performance and the decision threshold of spectral sensing are performed.

Keywords: cognitive radio, GID, cooperative spectrum sensing, goodness-of-fit, artificial neural networks,

Capítulo 1

Introdução

A procura por faixas de frequências livres no espectro eletromagnético tem aumentado com o desenvolvimento de novas tecnologias e serviços de comunicação sem fio. No entanto, percebe-se certa escassez desse recurso, principalmente causada pela política de alocação espectral vigente conhecida como alocação fixa das bandas de operação. Para cada tipo de serviço é destinada uma banda de frequência e sua utilização é normalmente vinculada ao pagamento de uma licença de uso. A escassez espectral pode ser agravada substancialmente com a internet das coisas (*Internet of Things*, IoT) [1] e com a quinta geração das redes de comunicação (5G) [2]. Para se ter uma ideia, no ano de 2020 espera-se que haja entre 50 e 100 bilhões de dispositivos conectados em rede, a maioria deles utilizando comunicação sem fio [3].

É possível atingir maior eficiência espectral, mantendo a qualidade de serviço (*Quality of Service*, QoS) desejada, através da utilização de modulações mais eficientes e implementação de códigos corretores de erros mais robustos. No entanto, o avanço da tecnologia nesta direção já foi tamanho que ganhos expressivos são mais difíceis de serem alcançados hoje em dia. Para atender a demanda por maiores taxas de comunicação e conexões simultâneas uma solução é o aumento da banda alocada, o que também esbarra na limitação de espectro citada anteriormente.

Apesar da aparente escassez de espectro, um estudo feito pela comissão federal de comunicação (*Federal Communications Commission*, FCC) dos Estados Unidos constatou que, embora a maior parte das faixas de frequência já esteja alocada, uma grande parte delas possui diferentes graus de ociosidade a depender da

hora e localização geográfica, podendo estar consideravelmente subutilizadas [4].

Nesse contexto surge o conceito de rádio cognitivo (*Cognitive Radio*, CR) [5]. Trata-se de uma tecnologia revolucionária que permite explorar a subutilização do espectro, promovendo um compartilhamento deste entre os usuários primários (*Primary Users*, PU), aqueles que possuem licença de operação em uma determinada banda de frequência, e os usuários secundários (*Secondary Users*, SU), aqueles que não possuem licença de operação. Um CR que faça parte de uma rede secundária é capaz de realizar acesso dinâmico e oportunista ao espectro de frequências, tomando por base o conhecimento das variáveis que tem influência na adaptação de seu funcionamento ao ambiente ou rede primária em que se insere. Portanto, os rádios cognitivos representam uma das mais promissoras soluções para o problema de escassez e congestionamento do espectro de frequências em sistemas de comunicação sem fio.

A técnica empregada pelos CRs para auxiliar a rede secundária a identificar as faixas ociosas é denominada de sensoriamento espectral [4,6,7]. O que se pretende como resultado do sensoriamento espectral é o acesso pelos SUs as faixas primárias ociosas do espectro de forma oportunista, sem comprometer o funcionamento da rede primária.

O sensoriamento espectral não é o único atributo de um CR. Além dele, o CR deve possuir inteligência suficiente para coletar informações sobre o ambiente ao seu redor, contextualizá-las para transformá-las em aprendizado e reconfigurar dinamicamente seus parâmetros de operação para atingir o desempenho ótimo e se adaptar ao ambiente que o cerca de maneira confiável e eficiente. Por exemplo, o CR deve ser capaz de aprender sobre as formas de ondas utilizadas, protocolos de transmissão, posicionamento geográfico e consumo energético para otimizar a utilização dos recursos, mantendo a QoS satisfatória ao usuário [8]. O processo de aprendizado de uma máquina (*machine learning*) está intimamente ligado a tais atributos e à escolha da tecnologia de inteligência artificial adequada. Para que o rádio adquira capacidade cognitiva e tome decisões de maneira autônoma, respeitando as políticas de regulamentação, a utilização de tal tecnologia torna-se-á imprescindível.

Esse capítulo será estruturado da seguinte forma: a Seção 1.1 apresenta a principal motivação para o desenvolvimento deste trabalho e o que se pretende resolver com a metodologia aplicada; a Seção 1.2 faz a revisão da literatura citando trabalhos desenvolvidos que se relacionam com o que será apresentado

nesta dissertação; a Seção 1.3 expõe os objetivos, contribuições e estrutura da dissertação; e, por fim, a Seção 1.4 lista as publicações de artigos sobre o tema em questão.

1.1 Definição do problema

O desempenho do sensoriamento espectral é comumente medido por meio da probabilidade de falso alarme, P_{fa} , e da probabilidade de detecção, P_d . A primeira é a probabilidade de se tomar uma decisão global em favor da presença do sinal primário, dado que o canal sensoriado se encontra vago (hipótese \mathcal{H}_0). A segunda é a probabilidade de se decidir em favor da presença do sinal primário dado que, de fato, o canal sensoriado está ocupado (hipótese \mathcal{H}_1) [9]. Matematicamente,

$$P_{fa} = \Pr[T > \lambda | \mathcal{H}_0] = \int_{\lambda}^{\infty} f(t | \mathcal{H}_0) dt = 1 - F(\lambda | \mathcal{H}_0), \quad (1.1)$$

$$P_d = \Pr[T > \lambda | \mathcal{H}_1] = \int_{\lambda}^{\infty} f(t | \mathcal{H}_1) dt = 1 - F(\lambda | \mathcal{H}_1), \quad (1.2)$$

Em que $\Pr[\cdot]$ denota a probabilidade de ocorrência do evento em questão, T é a estatística de teste computada de acordo com a técnica de sensoriamento escolhida, λ é o limiar de decisão, $f(t | \mathcal{H}_0)$ e $F(t | \mathcal{H}_1)$ são, respectivamente, a função densidade de probabilidade (*Probability Density Function*, PDF) e a função densidade cumulativa (*Cumulative Distribution Function*, CDF) de T sob a hipótese \mathcal{H}_0 , e $f(t | \mathcal{H}_1)$ e $F(t | \mathcal{H}_1)$ são respectivamente a PDF e a CDF de T sob \mathcal{H}_1 .

É desejado que P_d seja alta para que se proteja a rede primária de interferências que seriam causadas por transmissões secundárias em bandas erroneamente consideradas desocupadas. É também desejado que P_{fa} seja baixa, elevando assim a probabilidade de uso oportunista do espectro, por consequência elevando a vazão de dados da rede secundária.

Infelizmente os objetivos de elevar a P_d e reduzir a P_{fa} são concorrentes, significando que o aumento da primeira (reduzindo o limiar de decisão) também causa o aumento da segunda; e a redução da segunda (aumentando o limiar de decisão) causa também a redução da primeira. Na prática adota-se uma solução de compromisso tipicamente regida por norma. Por exemplo, no padrão IEEE

802.22 exige-se $P_d > 0,9$ e $P_{fa} < 0,1$ [10].

Por meio de (1.1) e (1.2) se pode notar que, para calcular as métricas de desempenho de uma técnica de sensoriamento espectral, há que se conhecer $f(t|\mathcal{H}_0)$ e $f(t|\mathcal{H}_1)$ ou, equivalentemente, $F(t|\mathcal{H}_0)$ e $F(t|\mathcal{H}_1)$. Esse conhecimento também se faz necessário para computar o limiar de decisão λ que é calculado para que se atinja a P_{fa} alvo, sendo a P_d governada pela razão sinal-ruído (*Signal-to-Noise Ratio*, SNR) nas entradas dos receptores dos SUs. Então, tem-se que

$$\lambda = F^{-1}(1 - P_{fa}|\mathcal{H}_0). \quad (1.3)$$

sendo $F^{-1}(\cdot)$ a função inversa de $F(\lambda|\mathcal{H}_0)$, ou seja, λ é o valor t tal que $F(t|\mathcal{H}_0) = 1 - P_{fa}$.

No entanto, o conhecimento das PDFs e CDFs próprias de cada estatística de teste tem grande valia se puder ser relacionado com os parâmetros sistêmicos do sensoriamento espectral. Somente assim será possível avaliar previamente, na fase de projeto, as situações (combinações de parâmetros sistêmicos) em que o teste em questão será efetivo e o CR poderá tomar uma decisão confiável baseada na resposta do sensoriamento realizado.

No cenário de CR entende-se por parâmetros sistêmicos, no mínimo, o número de PUs, números de SUs, quantidade de amostras do sinal recebido por cada SU e a relação sinal ruído.

Tendo em vista o que foi exposto, é necessário encontrar as distribuições de probabilidade que caracterizam de maneira satisfatória as hipóteses \mathcal{H}_0 e \mathcal{H}_1 e, posteriormente, relacionar tais distribuições com os parâmetros sistêmicos que influenciam na formação da estatística de teste do sensoriamento espectral.

1.2 Trabalhos relacionados

Várias técnicas de sensoriamento espectral já foram propostas, destacando-se a detecção de energia, a detecção por filtro-casado, a detecção por propriedades ciclo-estacionárias do sinal [6, 9], várias técnicas baseadas nos autovalores da matriz de covariância do sinal recebido [11], e outras baseadas no teorema dos círculos de Gerschgorin [12, 13]. Pode-se classificar as técnicas de sensoriamento

espectral quanto à necessidade de conhecer informações *a priori* sobre o sinal recebido ou sobre o ruído. A detecção de energia e algumas técnicas baseadas em autovalores são consideradas semi-cegas, já que precisam conhecer a potência do ruído na entrada dos receptores dos terminais secundários. Algumas técnicas também são ditas semi-cegas por necessitarem conhecer características do sinal primário, como é o caso da detecção por filtro-casado e da detecção por propriedades ciclo-estacionárias. Quando não é necessário o conhecimento prévio sobre o sinal primário e sobre a potência de ruído, a técnica de sensoriamento é denominada cega.

Em termos da busca pelas distribuições de estatísticas de teste diversas e correspondentes expressões de cálculo das métricas de desempenho, pode-se citar o detector de energia, para o qual o desenvolvimento analítico não é complexo por demais [14], até análises de extrema complexidade, como é o caso de alguns detectores baseados nos autovalores da matriz de covariância do sinal recebido. Como exemplos, em [15] utiliza-se o método dos momentos para parametrizar a distribuição Beta para a estatística de teste denominada de razão de Hadamard, sob a hipótese \mathcal{H}_1 . O método dos momentos é também aplicado em [16] para caracterizar com a distribuição Beta a estatística denominada de teste de esfericidade. Em [17] demonstra-se que a distribuição Gamma, devidamente parametrizada, pode ser utilizada para avaliar o desempenho da técnica baseada em volume. Há outras técnicas de sensoriamento cujo desenvolvimento analítico em busca das distribuições da estatística de teste se torna tão complexo, ou até mesmo intratável matematicamente, que lança-se mão de técnicas empíricas. Por exemplo, em [18] determina-se empiricamente que a PDF Log Pearson 3P pode ser utilizada para caracterizar a estatística de teste da técnica de máximo autovalor sob a hipótese \mathcal{H}_1 .

Algumas tecnologias de inteligência artificial, sobretudo redes neurais artificiais, foram introduzidas em rádios cognitivos com objetivo de melhorar o desempenho desses equipamentos. É possível também encontrar na literatura essa tecnologia aplicada exclusivamente à solução de escassez espectral ou como melhoramento de técnicas de sensoriamento, como pode-se notar nos trabalhos listados abaixo.

Em [19], uma rede neural artificial (*Artificial Neural Network*, ANN) com alimentação direta (*feedforward*) e duas camadas ocultas (*hidden layers*), utilizando algoritmos de retro-propagação (*back propagation*), foi treinada para estabelecer a relação entre parâmetros sistêmicos e o desempenho ideal de um sistema de

comunicação. A SNR, o número de quadros recebidos com sucesso e com erro, entre outras métricas, foram relacionados com a taxa de transmissão, com a confiabilidade e com o atraso, formando uma rede com quatro parâmetros de entradas e três de saída.

A identificação de qual transmissor está ativo, utilizando as características do seu sinal (*RF fingerprint*), é uma tarefa utilizada em redes pessoais sem fio (*Wireless Personal Area Networks*, WPAN). Em [20] demonstrou-se que redes neurais convolucionais, as mesmas utilizadas em processamento de imagem e reconhecimento de fala, também possuem bom desempenho nessa tarefa.

As redes neurais também já foram introduzidas especificamente ao contexto da tecnologia de rádios cognitivos. Como exemplo, em [21] é feita uma revisão sobre vários trabalhos de pesquisa que abordaram esse tema. Um deles [22] utiliza uma ANN para estabelecer uma relação entre a SNR, o código corretor de erro a taxa de erro de bit (*Bit Error Rate*, BER) requerida pelo sistema de comunicação, gerando uma base de conhecimento que auxilia na escolha da modulação para manter a taxa de erro requerida sob certa SNR. Outro trabalho [23] citado em [21] propõe a utilização de algumas características do sinal para identificar o tipo de modulação, via ANN, detectando assim a presença do sinal primário. Em [24], uma ANN é combinada com a detecção de energia e com o teste da razão de verossimilhança (*Likelihood Ratio Test*, LRT), formando um esquema de sensoriamento espectral híbrido que traz melhorias em relação a detectores de energia.

Uma abordagem mais ampla da inteligência artificial (*Artificial Intelligence*, AI) está sendo aplicada em uma das iniciativas mais recentes para resolver o problema da escassez de espectro: Pesquisadores do grupo *spectrum collaboration challenge* da agência de projetos de pesquisa avançada da defesa (*Defence Advanced Research Projects Agency*, DARPA) dos Estados Unidos estão projetando novos rádios que usam inteligência artificial para aprender como o espectro pode ser compartilhado entre os concorrentes para aumentar a taxa de transferência geral de dados [25].

Em [26] o autor implementa um experimento prático utilizando o método de detecção de energia em conjunto com uma ANN treinada. O sinal transmitido e gerado por uma *Raspberry Pi* e irradiado por um transmissor que utiliza chaveamento na amplitude (*Amplitude Shift Keying*, ASK) na frequência de 433 MHz, o receptor é composto por um rádio definido por software. A decisão sobre a

ocupação do canal é tomada por uma rotina implementada no MATLAB porém, ao invés de se comparar a energia do sinal recebido com um limiar de decisão, este mesmo sinal é digitalizado, filtrado e convertido para domínio da frequência utilizando transformada rápida de fourier (*Fast Fourier Transformer*, FFT). O resultado da FFT é aplicado à ANN que toma a decisão sobre a ocupação do canal, melhorando assim a eficiência do tradicional detector de energia, principalmente em condições de SNR precárias.

De acordo com o exposto, fica claro que é necessário estabelecer um método para análise de desempenho das técnicas de sensoriamento espectral. A chave para essa análise é encontrar uma distribuição de probabilidade que caracteriza tal técnica. Em nenhuma das pesquisas citadas anteriormente se identificou o uso da análise de aderência como alternativa para a determinação das distribuições de estatísticas de teste, talvez pela consequente dificuldade de se mapear, em seguida, os parâmetros das distribuições nos parâmetros sistêmicos.

A necessidade de implementar no rádio capacidades cognitivas de aprendizado e tomada de decisões fizeram com que tecnologias de AI fossem introduzidas no cenário de CRs, porém também não se identificou a aplicação de redes neurais artificiais como alternativa à parametrização das distribuições de estatísticas de teste em função de variáveis sistêmicas.

Neste trabalho é utilizado o teste de aderência para encontrar a distribuição que mais se adequa a técnica de sensoriamento espectral sob análise e posteriormente uma rede neural é treinada para estabelecer uma relação entre os parâmetros das distribuição e os parâmetros sistêmicos.

1.3 Estrutura e contribuições da dissertação

Em [27] foi proposta uma nova estatística de teste para sensoriamento espectral a partir do detector inspirado no índice de gini (*Gini Index Detector*, GID) o qual se mostrou bastante robusto nos cenários de potências de sinal e de ruído desiguais nos SUs, bem como variantes no tempo, situação essa comum na prática. Este trabalho complementa a avaliação feita em [27] estabelecendo as distribuições que caracterizam o teste GID e relaciona os parâmetros de tais distribuições de probabilidade com os parâmetros sistêmicos.

Vale ressaltar que o estabelecimento da relação entre os parâmetros sistêmicos

e os parâmetros das distribuições da estatística de teste nem sempre é trivial, muitas vezes necessitando de métodos analíticos consideravelmente complexos. Há também aqueles casos em que a própria formação da estatística de teste torna intratável a análise matemática de suas distribuições sob as hipóteses \mathcal{H}_1 (principalmente) e \mathcal{H}_0 . Este é o caso do GID, o que motivou a utilização de uma ANN para mapear os parâmetros sistêmicos nos parâmetros das distribuições.

Em vista do exposto, o objetivo desta dissertação de mestrado é:

- Apresentar os procedimentos e os resultados da análise de aderência (*Goodness-of-Fit*, GoF) da estatística de teste do GID a diversas distribuições de probabilidade, sob as hipóteses de ausência e de presença do sinal sensoriado
- Fornecer expressões fechadas para cálculo de P_{fa} , P_d e λ ;
- Estabelecer um conjunto de treinamento para rede neural artificial;
- Relacionar parâmetros das distribuições selecionadas com os parâmetros sistêmicos por meio de uma rede neural artificial;
- Permitir que as expressões mencionadas sejam operadas de acordo com as especificações do sistema;
- Realizar a verificação da acurácia das expressões de cálculo de P_d , P_{fa} e λ ;
- Estabelecer um método que pode ser aplicado em outros testes de sensoriamento espectral.

Para atingirmos esses objetivos, esta dissertação apresenta no **Capítulo 2** uma introdução aos fundamentos de sensoriamento espectral. Em seguida é dado um enfoque especial no sensoriamento cooperativo centralizado por fusão de dados. Além disso, as métricas de avaliação de desempenho e algumas técnicas de sensoriamento são apresentadas. Por fim o teste GID é explicado. O objetivo deste capítulo é fornecer uma base teórica para o entendimento dos capítulos seguintes.

No **Capítulo 3** os conceitos de redes neurais artificiais são introduzidos, as principais topologias de rede e suas aplicações são apresentados em seguida. Este capítulo se encerra como uma breve explicação sobre o processo de aprendizagem e das funções de ativação comumente utilizadas. Esses conceitos são a base para compreensão do restante da dissertação.

No **Capítulo 4** descreve-se o método empírico utilizado para análise de aderência da estatística de teste GID a diversas distribuições de probabilidade. A obtenção dos dados de treinamento, a escolha do tipo de rede neural e os resultados do treinamento são apresentados em seguida.

O **Capítulo 5** traz alguns resultados numéricos para verificação da acurácia das expressões de cálculo de P_d e P_{fa} bem como uma verificação da precisão do treinamento das redes desenvolvidas para mapeamento dos parâmetros das distribuições nos parâmetros sistêmicos.

Por fim, o **Capítulo 6** conclui este trabalho com as considerações finais. São ainda abordados os aspectos de possíveis trabalhos futuros para a continuidade do trabalho de pesquisa abordado nesta dissertação.

1.4 Publicações

Os seguintes artigos foram produzidos como resultado das pesquisas relacionadas a este trabalho:

1. A. L. Lemes and D.A Guimarães, “**Estimação das Distribuições da Estatística de Teste GRCR sob as Hipóteses \mathcal{H}_0 e \mathcal{H}_1** ”, in *36th Simpósio Brasileiro De Telecomunicações e Processamento de Sinais (SBrT)*, Campina Grande, Set. 2018.
2. A. L. Lemes, D. A. Guimarães, and Y. M. C. Masselli, “**Artificial Neural Networks Applied to the System-to-Distribution Parameter Mapping for the Gini Index Detector Test Statistic**”, in *ELSEVIER Engineering Applications of Artificial Intelligence Journal*. Submetido em Jul. 2019.

Capítulo 2

Sensoriamento Espectral

Conforme mencionado no Capítulo 1 o sensoriamento espectral é a técnica empregada pelo CR para detecção de lacunas espectrais da rede primária, porém esta forma de encontrar faixas com frequência livres não é a única. A fim de determinar quais bandas do espectro estão disponíveis para uso secundário, há três métodos que merecem destaque [28]: registro em banco de dados, sinais balizadores e sensoriamento espectral.

No método de registro em banco de dados os usuários primários armazenam em um servidor centralizado as informações relevantes de seus estados, como, por exemplo, as posições em que se encontram, as potências irradiadas e os tempos estimados de conexão. Através de uma conexão com a Internet, os usuários secundários acessam tais informações, analisando-as a fim de determinar lacunas espectrais e, por consequência, oportunidades de transmissão [29].

No método de sinais balizadores, em vez de serem armazenadas em um banco de dados, as informações relevantes dos usuários primários são transmitidas na interface aérea. Basta aos usuários secundários monitorarem um canal de controle padronizado a fim de determinar quais porções do espectro eletromagnético podem ser utilizadas [29].

No método de sensoriamento espectral, para determinar as lacunas espectrais os usuários secundários monitoram diretamente a banda de interesse, através de técnicas específicas, dispensando a necessidade da participação dos usuários de sistemas legados. Isto torna esta técnica mais atrativo que os métodos de registro em banco de dados e de sinais balizadores [29].

Este capítulo se inicia com uma breve introdução ao sensoriamento espectral, em seguida, na Seção 2.1 demonstra-se que o resultado do sensoriamento pode ser interpretado como um teste de hipótese binário. Um destaque maior é dado ao sensoriamento centralizado por fusão de dados, também são apresentadas as métricas de avaliação de desempenho, finalizando com a explicação de algumas técnicas de sensoriamento.

2.1 Introdução ao sensoriamento espectral

O sensoriamento pode ser realizado de maneira independente por cada SU ou de forma cooperativa. A Figura 2.1 apresenta um cenário de compartilhamento espectral entre uma rede de rádios cognitivos e uma rede primária. Caso o sensoriamento seja realizado de maneira independente, cada rádio cognitivo colhe amostras do sinal recebido do sistema primário e, usando uma das várias técnicas existentes decide pelo estado de ocupação do canal através de um teste de hipóteses. Este tipo de sensoriamento sofre grande influência do canal de comunicação entre os transmissores primários e os rádios cognitivos: suponha que um determinado rádio cognitivo esteja em uma situação de sombra de cobertura (*shadowing*) e que não esteja ao alcance do sinal de um transmissor primário. Ele pode erroneamente decidir utilizar um canal que de fato pode estar sendo ocupado pelo sistema primário, podendo causar neste uma forte interferência [29].

Outro problema comum neste cenário é a incerteza do terminal: na Figura 2.1 percebe-se que o rádio cognitivo CR3 está fora do alcance da rede primária. Portanto, ele pode decidir usar um canal que está na verdade ocupado, interferindo no funcionamento da rede primária.

Os problemas do sombreamento e da incerteza do terminal podem ser drasticamente minimizados com o uso do sensoriamento cooperativo. Nele, vários rádios cognitivos monitoram o canal de interesse e reportam suas medidas ou suas decisões individuais a um centro de fusão (*Fusion Center*, FC), que pode ser outro CR ou a estação base da rede secundária. A decisão final sobre o estado de ocupação de um canal é tomada no FC e é difundida em seguida aos CRs que, a partir deste momento, disputarão acesso ao canal. A difusão de informações entre o centro de fusão e os rádios cognitivos é feita através de um canal de controle dedicado. Esta forma de cooperação configura o que se conhece como sensoriamento espectral cooperativo com decisão centralizada. Nos casos em que

os rádios cognitivos envolvidos no processo reportam ao centro de fusão suas decisões individuais, cabe ao centro de fusão a combinação lógica das decisões individuais para posterior decisão final.

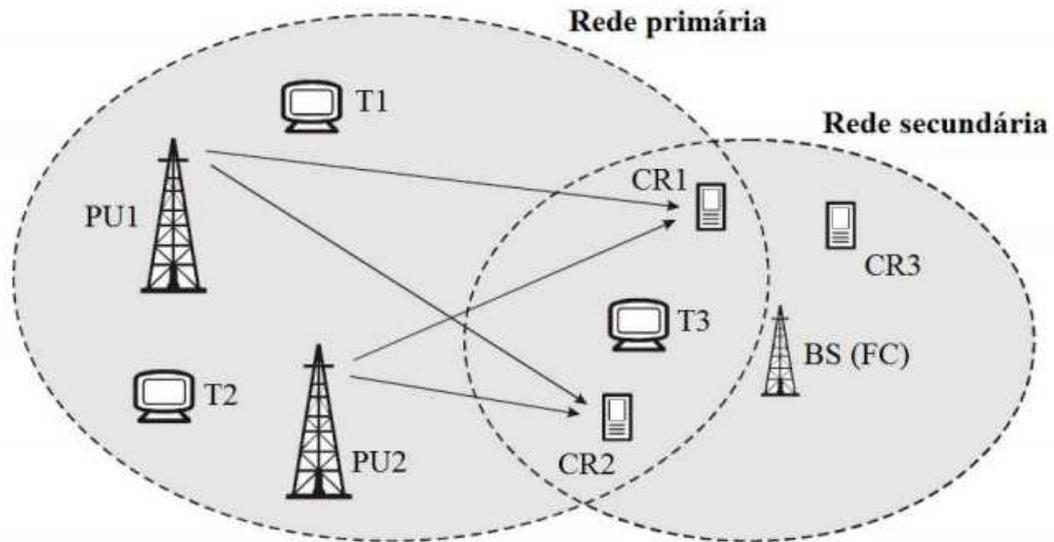


Figura 2.1: Compartilhamento espectral entre rede primária e secundária.

As técnicas de sensoriamento espectral devem ser capazes de identificar as transmissões dos usuários primários dentro de certo limite de tempo que é dependente da aplicação almejada na rede secundária e do tipo de sistema primário. Nesse processo há dois parâmetros temporais de suma importância: o tempo de detecção do canal e o período de sensoriamento. O tempo de detecção representa uma relação de compromisso entre a confiabilidade do teste de hipótese em questão (canal ocupado ou livre) e o número de oportunidades identificadas. Tem-se mais certeza sobre o resultado da classificação quanto à ocupação de um determinado canal quando se tem mais tempo para o sensoriamento [30]. À medida que esse tempo vai sendo reduzido, piora-se a qualidade da classificação. Por outro lado, tem-se mais tempo para que outras bandas sejam sensoriadas.

A periodicidade do sensoriamento espectral, ou seja, quantas vezes ele é executado em um intervalo de tempo, tem impacto direto na interferência nos usuários primários. Uma vez encontrado um canal livre, um usuário secundário inicia sua transmissão. Porém, o processo de sensoriamento deve ser executado frequentemente a fim de identificar um possível retorno de atividade do usuário primário. Se o tempo entre um sensoriamento e outro é grande, corre-se o risco do retorno

do usuário primário não ser percebido, resultando em transmissão simultânea do usuário primário e do usuário secundário, caracterizando assim um cenário de interferência.

2.1.1 Teste hipótese binário

Como resultado do sensoriamento do espectro, várias técnicas de detecção podem ser empregadas para auxiliar na decisão final sobre a presença ou ausência do sinal primário. Porém, independentemente do tipo do sinal em teste, ou da estratégia de detecção utilizada, a presença ou ausência do PU pode ser formulada como um teste de hipóteses binário, dado por [31]:

$$y(t) = \begin{cases} w(t) & : \mathcal{H}_0 \\ hs(t) + w(t) & : \mathcal{H}_1 \end{cases}, \quad (2.1)$$

em que $y(t)$ é o sinal recebido pela rede secundária, h representa o ganho ou atenuação do canal entre um usuário primário e o CR, $s(t)$ é o sinal transmitido pela rede primária e $w(t)$ representa o ruído aditivo gaussiano branco (*Additive White Gaussian Noise*, AWGN) [32]. As hipóteses \mathcal{H}_0 e \mathcal{H}_1 denotam a ausência e a presença do sinal primário, respectivamente.

Após encontrar a distribuição de probabilidade que melhor representa o teste sob as hipóteses \mathcal{H}_0 e \mathcal{H}_1 o desempenho da técnica adotada pode ser medido em função de P_{fa} e P_d utilizando as equações (1.1) e (1.2) que, graficamente estão representadas na Figura 2.2.

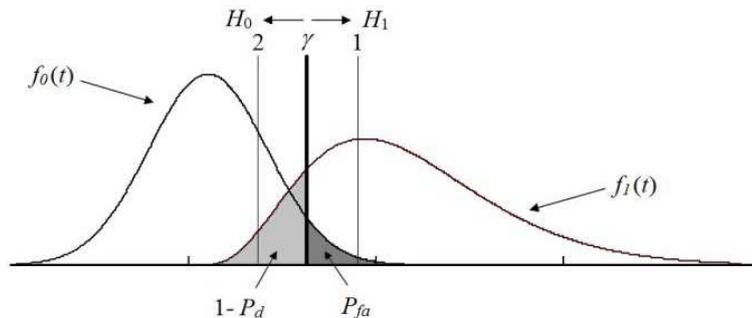


Figura 2.2: Distribuições hipotéticas $f_0(t)$ e $f_1(t)$.

na Figura 2.2 $f_0(t)$ representa a PDF sob hipótese \mathcal{H}_0 e $f_1(t)$ representa a PDF sob hipótese \mathcal{H}_1 .

2.2 Sensoriamento espectral cooperativo centralizado utilizando fusão de dados

No processo de sensoriamento, a presença de desvanecimento temporal no canal dificulta consideravelmente a obtenção de estimativas confiáveis porém, dificilmente todos os rádios que compõem a rede secundária cognitiva estarão sujeitos simultaneamente a esse problema. Tais deficiências podem ser mitigadas se vários SUs forem utilizados para sensoriamento espectral, de modo que seja possível obter uma estimativa sobre a ocupação do canal através da fusão dos dados enviados por esses SUs. O fato de cada rádio cognitivo estar sujeito a desvanecimentos diferentes gera uma diversidade espacial que pode ser explorada usando técnicas de detecção que tornam o resultado do processo de sensoriamento mais confiável [33].

Considerando um modelo de sensoriamento espectral cooperativo em canal discreto e sem memória, com ruído AWGN e desvanecimento plano e lento, ainda admitindo que haja m receptores secundários ou apenas um único receptor secundário que possua m antenas e que cada receptor ou cada antena colete n amostras do sinal recebido dos p transmissores primários durante o intervalo de sensoriamento. Essas amostras são dispostas em uma matriz $\mathbf{Y} \in \mathbb{C}^{m \times n}$. Tal matriz, sob as hipóteses \mathcal{H}_0 e \mathcal{H}_1 é dada por:

$$\mathbf{Y} = \begin{cases} \mathbf{V} & : \mathcal{H}_0 \\ \mathbf{H}\mathbf{X} + \mathbf{V} & : \mathcal{H}_1 \end{cases}, \quad (2.2)$$

sendo que as amostra dos sinais transmitidos são arranjadas na matriz $\mathbf{X} \in \mathbb{C}^{p \times n}$ e as amostras de ruído AWGN de média nula e variância σ^2 são dispostas na matriz $\mathbf{V} \in \mathbb{C}^{m \times n}$. A matriz $\mathbf{H} \in \mathbb{C}^{m \times p}$ contém os elementos h_{ij} , que representam o ganho complexo do canal de sensoriamento entre o j -ésimo transmissor primário e o i -ésimo receptor secundário, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, p$.

É comum considerar que os ganhos desse canal são constantes durante o intervalo de sensoriamento isso modela um canal com desvanecimento lento. É importante notar que se todos os elementos h_{ij} da matriz \mathbf{H} forem iguais a 1 o modelo da equação (2.2) descarta o efeito do canal e passa a representar um canal de sensoriamento apenas AWGN.

Na fusão de dados, a i -ésima linha da matriz \mathbf{Y} gerada no i -ésimo SU é enviada

ao centro de fusão, onde \mathbf{Y} é formada. A partir daí, usando alguma técnica de sensoriamento específica, o centro de fusão processa as amostra de \mathbf{Y} para dar origem à estatística de teste desejada a então aplicar o teste de hipóteses e decidir sobre a ocupação do canal. Essa decisão é então enviada para todos os SUs que compõem a rede secundária através do canal de controle.

2.3 Métricas de avaliação para sensoriamento

O desempenho do sensoriamento espectral é comumente avaliado através da probabilidade de detecção (P_d) e da probabilidade de falso alarme (P_{fa}). Essas métricas estão definidas no Capítulo 1, Seção 1.1. De forma a aglutinar as principais métricas de avaliação de desempenho nesta seção, a definição matemática de P_d e P_{fa} são apresentadas novamente a seguir.

$$P_d = \Pr\{T > \lambda \mid \mathcal{H}_1\} , \quad (2.3)$$

$$P_{fa} = \Pr\{T > \lambda \mid \mathcal{H}_0\} , \quad (2.4)$$

É também usual se especificar uma probabilidade de falso alarme constante, mais conhecida como taxa de falso alarme constante (*Constant False Alarm Rate*, CFAR), garantindo-se que a probabilidade de detecção fique acima de um valor mínimo aceitável, que também recomendado por norma assim como os valores de P_d e P_{fa} .

Além de P_d e P_{fa} é comum também utilizar outros métodos para realizar a avaliação de desempenho do sensoriamento.

2.3.1 Avaliação por meio de curvas ROC

As probabilidades de detecção (P_d) e de falso alarme (P_{fa}) de uma técnica de sensoriamento espectral são comumente avaliadas por meio da curva característica de operação do receptor (*Receiver Operating Characteristic*, ROC), na qual P_d é expressa em função de P_{fa} variando-se o limiar de decisão λ . Cada ponto sobre uma curva ROC corresponde a um valor de limiar de decisão λ , cuja faixa

de variação corresponde à interseção dos suportes das PDFs empíricas (ou dos histogramas) da estatística de teste sob \mathcal{H}_0 e \mathcal{H}_1 .

Um exemplo de curva ROC é mostrada na Figura 2.3 note que a curva 'a' é a que retrata pior desempenho, seguida da curva 'b' e por fim da curva 'c'. Nota-se isso fixando P_{fa} e verificando que P_d é menor para a curva 'a' do que para as curvas 'b' e 'c'.

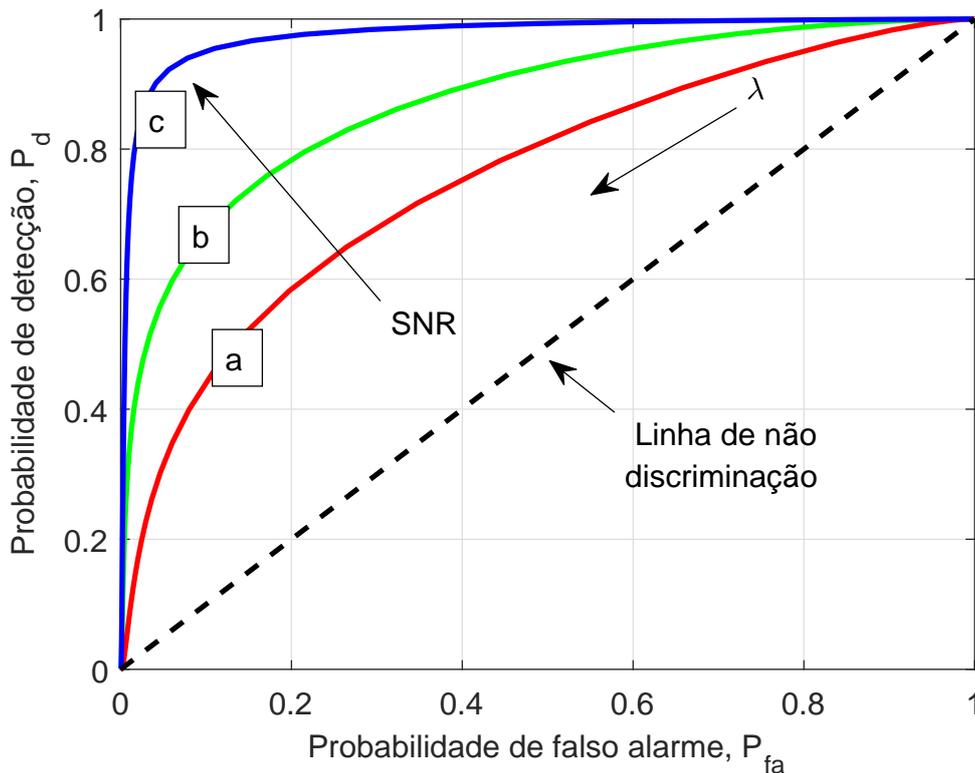


Figura 2.3: Exemplos hipotéticos de curvas ROC.

Em geral, o aumento da SNR é a medida tomada para que o desempenho de uma dada técnica de sensoriamento seja melhorado. No entanto, essa melhoria pode advir de alterações em outros parâmetros sistêmicos, como o aumento do número de amostras colhidas pelo SU ou o aumento do número de SUs em cooperação. Diferentes técnicas de sensoriamento também podem apresentar desempenhos diferentes sob as mesmas condições sistêmicas.

Uma decisão aleatória sobre o estado de ocupação da banda sensoriada levaria a um ponto ao longo da linha identificada como linha de não-discriminação. Curvas ROC abaixo dessa linha correspondem a desempenhos inúteis, exceto que se saiba que ali se encontram, caso em que o desempenho se reverteria em útil

apenas invertendo-se as decisões tomadas.

2.3.2 Área sob a curva

Outra métrica de desempenho também utilizada é a área sob a curva (*Area Under the Curve*, AUC) ROC que definida como:

$$AUC = \int_0^1 P_d(T, \lambda) dP_{fa}(\lambda) \quad (2.5)$$

onde P_d é a probabilidade de detecção e P_{fa} a probabilidade de falso alarme do teste em questão e λ é o limiar de decisão.

O valor da AUC varia entre 0 e 1. Caso seja menor que 0,5 significa que a ROC está abaixo da linha de não-discriminação.

A AUC é particularmente útil quando se deseja aglutinar P_d e P_{fa} em uma única métrica, o que se torna atrativo se, por exemplo, curvas ROC estão se cruzando ou quando se deseja reduzir a quantidade de valores de medidas de desempenho expressos em uma tabela, lista ou gráfico, por restrições de espaço.

2.4 Principais técnicas de sensoriamento espectral

2.4.1 Detecção por filtro casado

O dispositivo para a detecção ótima de um sinal em um canal sob a influência do ruído AWGN envolve o uso de um filtro linear invariante no tempo conhecido como filtro casado [34]. O filtro casado é capaz de maximizar a SNR do sinal recebido no momento da decisão. Por isso a utilização dessa técnica é bastante atrativa em um sistema de rádios cognitivos. Em sua implementação é preciso projetar a resposta ao impulso do filtro de forma casada com o sinal transmitido, o que significa que é necessário conhecer o formato dos pulsos transmitidos pelos PUs. Em função dessa necessidade, essa estratégia de detecção é considerada ótima desde que o sinal transmitido pela rede primária seja conhecido pela rede secundária [35]. O detector por filtro casado é ilustrado na Figura 2.4.

O sinal de entrada, $y(t)$, do filtro é formado pelo sinal transmitido, $s(t)$, pela

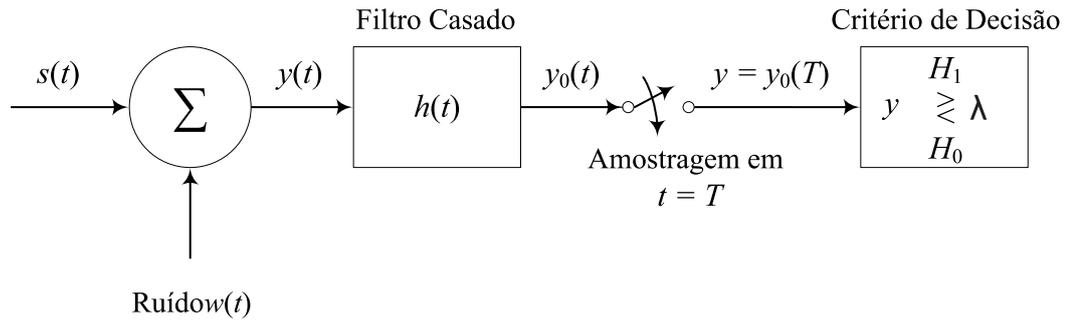


Figura 2.4: Diagrama em blocos para detecção por filtro casado.

rede primária contaminado pelo ruído, $w(t)$. O sinal de saída, $y_0(t)$, é amostrado a cada $t = T$ segundos e os valores de amostras extraídos, $y_0(T) = y$, são comparados com o limiar, λ , segundo um critério de decisão definido. Neste caso decide-se em favor da presença do sinal primário se os valores das amostras forem maiores que o limiar de decisão ($y > \lambda$), caso contrário, se os valores das amostras forem menores que o limiar ($y < \lambda$) decide-se em favor do canal sensoriado estar desocupado.

Sendo $y[n]$ as amostras extraídas no sensoriamento dentro do intervalo discreto de observação, $n = 0, 1, \dots, N - 1$, o sinal recebido por um CR pode ser descrito por meio da seguinte expressão:

$$y[n] = \begin{cases} w[n] & : \mathcal{H}_0 \\ h[n]s[n] + w[n] & : \mathcal{H}_1 \end{cases}, \quad (2.6)$$

em que o sinal $y[n]$ é entregue à entrada do filtro casado, que possui resposta ao impulso igual a $h[n] = s[N - 1 - n]$, $n = 0, 1, \dots, N - 1$. A decisão sobre a ocupação do canal é alcançada amostrando a saída do filtro $y_0[n]$ no tempo discreto $t = N - 1$ e comparando com o limiar de decisão λ . A resposta ótima do filtro pode então ser obtida conforme a expressão apresentada em [36], dado por

$$y = \sum_{n=0}^{N-1} y[n]s[n] \begin{matrix} \mathcal{H}_1 \\ \geq \lambda \\ \mathcal{H}_0 \end{matrix}. \quad (2.7)$$

A equação 2.7 corresponde à operação de correlação entre as amostras do sinal recebido e as amostras referentes ao sinal transmitido pela rede primária durante o período de análise.

2.4.2 Detecção de energia

Por ser uma técnica de detecção não coerente, ao contrário das técnicas de detecção por filtro casado, a estratégia baseada em detecção de energia não requer nenhum conhecimento *a priori* do sinal primário. Neste método a decisão final a respeito do estado de ocupação do canal é obtida apenas com base na quantidade de energia contida na faixa do espectro sensoriado. O teste estatístico consiste em comparar o valor de energia medido na faixa em questão com o limiar de decisão para determinar a situação do estado de ocupação do canal. Sendo assim, o canal é considerado ocupado quando o resultado da estatística de teste é maior que o limiar, caso contrário o canal é dado como livre, podendo ser aproveitado para efetuar transmissões secundárias.

A estatística de teste do detector de energia pode ser definida como o valor médio da energia contida em N amostras extraídas do sinal $y[n]$ no canal alvo do sensoriamento [37]. Logo, pode-se escrever que

$$T_{\text{ED}} = \frac{1}{N} \sum_{n=0}^{N-1} |y[n]|^2 \begin{matrix} \mathcal{H}_1 \\ \geq \lambda \\ \mathcal{H}_0 \end{matrix} \quad (2.8)$$

2.4.3 Detecção por autovalores

Na técnica de detecção baseada em autovalores não há necessidade de se conhecer as características do sinal primário para alcançar uma decisão final a respeito do estado de ocupação do canal em teste [38]. Seu emprego baseia-se, em primeiro lugar, na formação de uma matriz cujos elementos são as amostras coletadas do sinal sensoriado dentro da banda de interesse. A partir dessa matriz de amostras é gerada a matriz de covariância do sinal recebido, de onde são calculados os autovalores e encaminhados, via canal de controle, ao FC para as devidas combinações, no caso do esquema de fusão de autovalores.

A matriz de covariância é formada a partir da matriz \mathbf{Y} que por sua vez é obtida no centro de fusão de acordo com o exposto na Seção 2.2. O próximo passo para encontrar a matriz de covariância a partir de \mathbf{Y} é utilizar a expressão $\mathbf{R}_Y = E[\mathbf{Y}\mathbf{Y}^\dagger]$, em que \dagger representa o conjugado transposto. No entanto, na prática é comum o uso da estimativa de máxima verossimilhança para obter \mathbf{R}_Y . Portanto, a equação para o cálculo da matriz de covariância estimada será dada

por

$$\hat{\mathbf{R}}_Y = \frac{1}{n} \mathbf{Y} \mathbf{Y}^\dagger. \quad (2.9)$$

Agora os autovalores $\{\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{m-1} \geq \lambda_m\}$ de $\hat{\mathbf{R}}_Y$ são usados para gerar as estatísticas de teste de hipóteses.

As principais estatísticas de teste no sensoriamento por autovalores são o teste da razão de verossimilhança generalizado (*Generalized Likelihood Ratio Test*, GLRT), a detecção pela razão entre o máximo e o mínimo autovalores (*Maximum-Minimum Eigenvalue Detection*, MMED), também conhecida como detecção pela razão de autovalores (*Eigenvalue Ratio Detection*, ERD), e a detecção pelo máximo autovalor (*Maximum Eigenvalue Detection*, MED), também conhecida como teste de roy de máxima raiz (*Roy's Largest Root Test*, RLRT) [11]. Os detectores GLRT e MMED são considerados cegos, enquanto o MED é considerado semi-cego devido ao fato de necessitar conhecer a potência de ruído nas entradas dos receptores secundários.

As estatísticas de teste GLRT, MMED e MED são computadas no FC de acordo com as equações 2.10, 2.11 e 2.12.

$$T_{\text{GLRT}} = \frac{\lambda_1}{\frac{1}{m} \sum_{i=1}^m \lambda_i}. \quad (2.10)$$

$$T_{\text{MMED}} = \frac{\lambda_1}{\lambda_m} \quad (2.11)$$

$$T_{\text{MED}} = \frac{\lambda_1}{\sigma^2}. \quad (2.12)$$

2.5 Detector robusto inspirado no coeficiente de Gini

O índice de Gini é uma medida de dispersão estatística amplamente utilizada nas ciências econômicas e sociais para medir as desigualdades. Esse mesmo índice

serviu de inspiração para construção de uma nova estatística de teste denominada GID utilizada para sensoriamento espectral cooperativo [27].

Em [27] foi proposta a estatística que possui as seguintes características: i) robusto nos cenários de potências de sinal e de ruído desiguais nos SUs, bem como variantes no tempo; ii) foi desenvolvido para ser aplicado em canais com linha de visada, embora também opere sem visada, mas com desempenho inferior; iii) supera o desempenho da maioria dos detectores robustos; iv) é um dos detectores mais simples já propostos. Além disso, o GID tem a importante propriedade de taxa de falso alarme constante e possui baixa complexidade, pois necessita apenas do cômputo da matriz de covariância do sinal recebido para gerar sua estatística de teste. Essas características o tornam uma opção para dispositivos que possuem pouca capacidade de processamento e necessitam de alta eficiência energética, como é o caso de dispositivos IoT e certos terminais de comunicação móvel

Para formação da estatística do teste GID é considerado o sensoriamento espectral cooperativo centralizado com fusão de dados. Como esse teste é baseado na matriz de covariância a formação da matriz \mathbf{Y} segue o que foi exposto na Seção 2.2.

Porém a matriz \mathbf{H} sofre uma variação pois, de forma a modelar um canal com desvanecimento Rice plano e lento, em (2.2) tem-se $\mathbf{H} = \mathbf{G}\mathbf{A}$, em que $\mathbf{A} \in \mathbb{C}^{m \times p}$ tem elementos $\alpha_{ij} \sim \mathcal{CN}[\sqrt{K/(2K+2)}, 1-(K+1)]$, sendo K o fator de Rice que estabelece a razão entre a potência de componente de multipercurso dominante e a potência dos demais componentes [27]. A matriz $\mathbf{G} \in \mathbb{R}^{m \times m}$ é dada por $\mathbf{G} = \text{diag}(\sqrt{\mathbf{s}/s_{\text{avg}}})$, em que o operador $\text{diag}(\cdot)$ gera uma matriz diagonal cuja diagonal principal é formada pelo vetor no argumento, $\mathbf{s} = [s_1, s_2, \dots, s_m]^T$ é o vetor que contém as potências dos sinais recebidos pelos SUs e $s_{\text{avg}} = \frac{1}{m} \sum_{i=1}^m s_i$ é a potência média destes sinais.

De posse de \mathbf{Y} , o FC computa a matriz de covariância amostral de acordo com (2.9) e a partir de \mathbf{R}_Y a estatística GID é formada de acordo a equação

$$T_{GID} = \frac{2(m^2 - m) \sum_{i=1}^{m^2} |r_i|}{\sum_{i=1}^{m^2} \sum_{j=1}^{m^2} |r_i - r_j|}, \quad (2.13)$$

em que r_i é vetor \mathbf{r} é formado pelo empilhamento de todas as colunas da matriz \mathbf{R} . A decisão é global sobre o estado de ocupação da banda sensoriada é então

tomada a favor de \mathcal{H}_1 se $T_{GID} > \lambda$, lembrando que λ é o limiar de decisão; se $T_{GID} \leq \lambda$, decide-se por \mathcal{H}_0 .

Capítulo 3

Redes Neurais Artificiais

As técnicas baseadas em conhecimento, em especial redes neurais artificiais, podem ser aplicadas em diversas áreas de pesquisa como: neurociência, matemática, estatística, física, ciência da computação e engenharia. Na grande maioria das vezes são utilizadas em aplicações que necessitam modelar e analisar uma série de dados, reconhecer padrões, processamento de sinais e controle pois possuem a importante propriedade de aprender a partir de uma serie de dados e utilizar esse aprendizado para tomar decisões futuras.

Na seção seguinte é apresentada uma breve introdução sobre ANNs, seguida de suas principais topologias de rede. Nas Seções 3.3 e 3.4 são discutidos, respectivamente, o processo de aprendizagem e as funções de ativação.

3.1 Introdução às redes neurais

Redes neurais artificiais são técnicas computacionais inspiradas na estrutura cerebral biológica de seres inteligentes com o objetivo de reproduzir o comportamento dos mesmos [39, 40]. Compostas por neurônios artificiais conectados, de forma a trocarem informações ponderadas entre si. A função do neurônio é receber informações externas e processá-las. O resultado é propagado para outros neurônios através das múltiplas conexões. A resposta do neurônio às informações de entrada dependem da função de ativação. As mais utilizadas são as de limiar de decisão abrupta (degrau), linear e as com limite de decisão mais suave, essas conhecidas como sigmoides por possuírem a aparência da letra "S". Toda ANN deve ser treinada para que os pesos sinápticos sejam otimizados. O método do

conhecimento *a priori* é comumente utilizado. A ANN recebe um conjunto de dados de entrada para os quais as saídas já são conhecidas e, através de algoritmos de treinamento, ajusta os pesos sinápticos das interconexões entre os neurônios com objetivo de reduzir o erro entre as saídas esperadas e o valor desejado. Após a finalização do treino, um conjunto de dados ainda não conhecido pela ANN, é utilizado para avaliar a eficácia do processo de treinamento [39, 40].

A utilização de redes neurais oferecem algumas capacidades aos sistemas nos quais são implementadas [39]. No contexto de rádio cognitivos e mapeamento dos parâmetros das distribuições de probabilidade em função dos parâmetros sistêmicos as principais capacidades estão listadas a seguir:

Mapeamento das saídas de um sistema em função das entradas: A capacidade de aprender através de exemplos e de generalizar a informação aprendida está associada à capacidade da rede de aprender através de um conjunto de exemplos e, posteriormente, dar respostas coerentes para dados não conhecidos. As ANNs são capazes de extrair informações não apresentadas de forma explícitas através de exemplos e com isso mapear as saídas em função das entradas após o ajuste dos pesos sinápticos das interconexões entre os neurônios [39].

Adaptabilidade: Redes neurais tem a capacidade de adaptar os seus pesos sinápticos e medida que o ambiente ao seu redor se altera. Isso significa que, a partir de uma mudança no ambiente, a rede pode ser rapidamente retreinada para se adaptar a tais mudanças. Isso é útil para ambientes não estacionários. Nesse caso a ANN pode ser projetada para alterar seus pesos sinápticos em tempo real.

É importante ressaltar que inteligência artificial e redes neurais artificiais são dois campos distintos de pesquisa dentro da ciência da computação. No entanto eles se relacionam entre si do ponto de vista que as ANNs são um ponto de partida para se atingir o que se espera de uma inteligência artificial. A AI é um vasto campo de pesquisa que tem como objetivo criar máquinas inteligentes, algo que já foi alcançado dependendo de como se define inteligência. A inteligência se adquire através de processos de aprendizagem. Nos seres humanos experimentamos situações e aprendemos com elas. Por esse ponto de vista através de ANNs que foram treinadas sob certas situações é possível afirmar que a máquina aprendeu como lidar com tal situação e por isso possui conhecimento sobre ela.

3.2 Topologias comuns de redes e suas aplicações

A maneira como que os neurônios estão interconectados define o que se denomina como topologia da rede neural. Algumas dessas topologias se tornaram bem difundidas e cada uma delas se adequa melhor para soluções de determinados problemas.

3.2.1 Rede *Feedforward* de camada única

É considerada a forma simples de uma rede em camadas, composta por uma camada de entrada (*input layer*) conectada diretamente a camada de saída (*output layer*) que contem os neurônios responsáveis por computar a informação recebida e fornecer os respectivos resultados. A camada de entrada não executa qualquer processamento de informação. Na Figura 3.1 está representada uma topologia do tipo *Feedforward* genérica. No caso da rede de uma única camada não existem as camadas intermediárias também conhecidas como *hidden layers*.

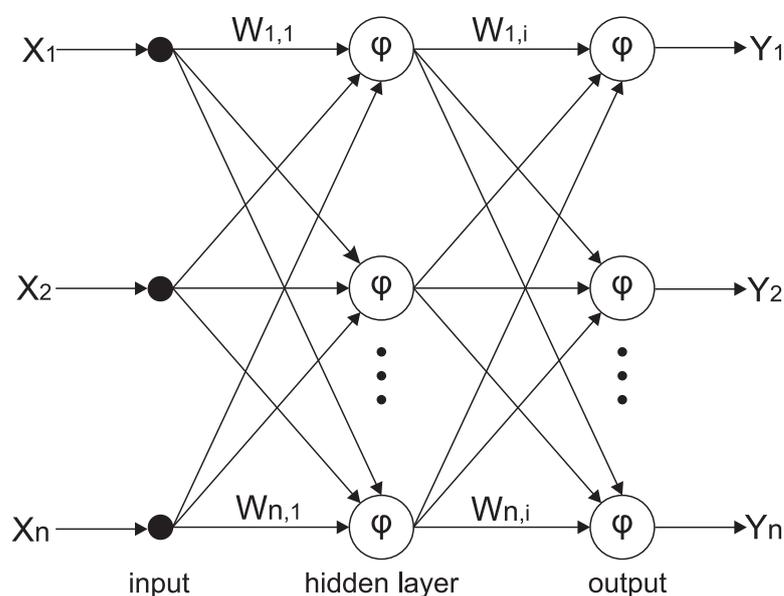


Figura 3.1: Topologia de rede neural *feedforward* generalizada.

3.2.2 Rede *Feedforward* de multicamadas

Essa rede se distingue da anterior pois possui uma ou mais camadas intermediárias que são utilizadas como nós de processamento dos dados de entrada.

À medida que o número de camada intermediárias aumenta, o poder de processamento da rede também cresce pois o número de neurônios e de interconexões sinápticas entre eles também se torna maior. Essa topologia é também chamada de redes de perceptrons por multicamadas (*Multi Layer Perceptrons*, MLP). As MLPs são frequentemente aplicadas a problemas de aprendizagem supervisionados: treinam um conjunto de pares entrada-saída e aprendem a modelar a correlação (ou dependências) entre essas entradas e saídas.

Para uma topologia do tipo MLP a equação de saída é a formula geral da interpolação linear [41]. Considerando o número adequado de camadas intermediárias, de neurônios em cada uma delas, a correta parametrização e treinamento é possível utilizar a ANN para mapear as saídas da rede de acordo com as entradas. De fato, apenas como exemplo, se simplificarmos a rede neural genérica da Figura 3.1 para uma entrada, uma saída, uma *Hidden Layer* e considerarmos uma função de ativação do tipo $v_{(\phi)}$, a saída é definida por $Y = X \times v_{(\phi)}$. Após o ajuste dos pesos sinápticos essa estrutura de rede neural deverá ser capaz de estimar qualquer função não linear contínua admitindo uma determinada precisão

3.2.3 Rede recorrentes

As redes neurais recorrentes (*Recurrent Neural Network*, RNN) são topologias utilizadas geralmente para reconhecimento e predições de padrões. São muito parecidas com as redes *feedforward*, porém possuem conexões de realimentação entre as camadas e portanto, a informação pode circular dentro da própria rede. Essa realimentação faz com que a decisão tomada pela rede no instante b_t tenha dependência do resultado obtido no instante b_{t-1} , o que a torna útil, por exemplo, para prever a próxima palavra em um frase. Essa tarefa se torna mais previsível se tivermos informação sobre o que foi dito anteriormente [42].

3.2.4 Rede convolucionais

As redes neurais convolucionais (*Convolutional Neural Network*, CNN) são geralmente aplicadas em reconhecimento de imagens e são uma variação das redes MLP. Uma CNN consiste em múltiplas partes com funções diferentes. Inicialmente é comum aplicar sobre o dado de entrada camadas ditas de convolução. Uma camada de convolução é composta por diversos neurônios, cada um responsá-

vel por aplicar um filtro em um pedaço específico da imagem. Podemos imaginar cada neurônio sendo conectado a um conjunto de *pixels* da camada anterior e, a cada uma dessas conexões se atribui um peso. A combinação das entradas de um neurônio, utilizando os pesos respectivos de cada uma de suas conexões, produz uma saída que é passada para a camada seguinte. Os pesos atribuídos as conexões de um neurônio podem ser interpretados como uma matriz que representa o filtro de uma convolução de imagens no domínio espacial.

Nas CNNs apenas um subconjunto de entradas é conectado a cada neurônio. As redes neurais convolucionais realizam análise de campos receptivos locais. Os neurônios da mesma camada são agrupados em mapas. Um mapa é produzido pelo agrupamento das saídas de neurônios que juntos cobrem uma parte da imagem que tenha sido processado por um filtro em comum [43].

3.3 Processo de aprendizagem

A característica mais importante das redes neurais é a capacidade de aprender a partir do ambiente e melhorar seu desempenho através desse processo. De maneira geral, a rede adquire mais conhecimento à medida que são utilizados diferentes situações no processo de aprendizagem. Há dois métodos que usualmente são aplicados para o processo de aprendizado, são eles:

Aprendizado supervisionado ou por associação: Nesse processo as redes são treinadas utilizando um conjunto de entradas e saídas pré determinadas. Para cada combinação de entrada é fornecida uma saída alvo e a rede ajusta os pesos sinápticos com objetivo de atingir essas saídas com menor erro possível. Esses pares de entrada-saída são geralmente fornecidos por agentes externos chamados de professores que são, na maioria das vezes, os desenvolvedores dos sistemas [40].

Aprendizado não supervisionado ou auto-organização: Nesse processo cada saída da rede responde a um conjunto de parâmetros de entrada. Nesse método a rede deve ser capaz de descobrir as características das estatísticas do conjunto de entrada. Diferentemente do aprendizado supervisionado, não há um conjunto de treinamento que relaciona entrada-saída utilizado para fornecer a rede um conhecimento prévio. Ou seja, a rede deve ser capaz de desenvolver sua própria representação do conhecimento através das análises das estatísticas do conjunto de entrada.

Independentemente da maneira de como se adquire o conhecimento este se dá através do ajuste dos pesos sinápticos e tal ajuste é realizado utilizando algoritmos de treinamento. Cada algoritmo utiliza uma técnica para atingir o menor erro.

3.3.1 Algoritmos de treinamento

O algoritmo de retropropagação (*Backpropagation*) é um dos métodos mais utilizados para calcular os gradientes da função de erro em relação aos parâmetros da rede neural. Ele funciona computando os gradientes na camada de saída e usando esses gradientes para calcular os gradientes na camada anterior e assim por diante. É a maneira mais utilizada para otimizar os pesos sinápticos de uma rede neural. Com base no gradiente local são feitos os ajustes de pesos na referida camada. A ideia básica é aplicar a regra da cadeia para computar a influência dos pesos sinápticos da função erro E , em geral, minimiza-se o erro quadrático médio, ou seja:

$$\frac{\partial E}{\partial w_{ni}} = \frac{\partial E}{\partial Y_i} \frac{\partial Y_i}{\partial net_n} \frac{\partial net_n}{\partial w_{ni}} \quad (3.1)$$

onde w_{ni} é o peso da conexão do n -ésimo neurônio da i -ésima camada oculta, Y_i é a saída do i -ésimo neurônio, e net_n é a soma dos pesos das entradas dos n neurônios. Para minimizar o erro é utilizado o método do gradiente descendente.

A taxa com que o erro converge para o valor esperado depende da taxa de aprendizagem ϵ . Quanto menor o valor de ϵ mais iterações serão necessárias para se atingir o erro desejado. Se ϵ for muito grande, oscilações podem ocorrer no treinamento e os valores ótimos dos pesos sinápticos não serão atingidos [44].

O Algoritmo de retropropagação é estável, porém para se atingir um bom desempenho é preciso fazer com que a taxa de aprendizagem seja pequena o que o torna lento. Para solucionar o problema da lentidão do processo de treinamento algoritmos como o de Newton e Gauss-Newton foram desenvolvidos porém são instáveis [45].

Como uma solução para o problema de lentidão e instabilidade foi desenvolvido o algoritmo de Levenberg-Marquardt que utiliza as matrizes Hessiana e Jacobiana da função erro para atingir o valor ótimo dos pesos sinápticos. Durante o processo de treinamento o algoritmo de Levenberg-Marquardt alterna entre o

algoritmo de retropropagação e o Gauss-Newton fazendo com que essa solução se torne rápida e estável e por isso é utilizado para treinamento de redes neurais com um número médio de amostras [45]. Além de desses, existem outros algoritmos baseados no método do gradiente e Jacobiano como: gradiente descendente com momento, Fletcher-Powell, Ploak-Ribiére e gradiente conjugado em escala, este último utilizado em redes treinadas para reconhecimento de padrão [46].

3.4 Funções de ativação

As funções de ativação, geralmente representadas por $\phi(v)$, definem como serão as respostas dos neurônios aos estímulos de entrada. A escolha de tais funções depende de qual é o propósito da rede, em linhas gerais as mais utilizadas são a função degrau e as sigmóides.

A função degrau, também conhecida como limiar de decisão abrupta, é definida conforme a expressão,

$$\phi(v) = \begin{cases} 1 & \text{se } v_n \geq 0 \\ 0 & \text{se } v_n \leq 0 \end{cases}, \quad (3.2)$$

Onde v_n é o valor da entrada aplicada ao neurônio.

A função sigmóide, cujo gráfico é em forma de "S", é a forma mais comum de função de ativação usada na construção de redes neurais. É definido como uma função estritamente crescente que exhibe um equilíbrio entre o comportamento linear e o não-linear. Um exemplo da função sigmóide é a função logística definida como:

$$\phi(v) = \frac{1}{1 + \exp(-av)}, \quad (3.3)$$

Onde a é o parâmetro de inclinação da função. Portanto, essa função assume valores contínuos que podem variar de 0 a 1.

Outra função sigmóide largamente utilizada é a tangente hiperbólica, essa função também possui comportamento suave porém sua variação estende de -1 a 1.

Capítulo 4

Parametrização das Distribuições da Estatística Teste GID sob \mathcal{H}_0 e \mathcal{H}_1 via ANN

Encontrar as distribuições de probabilidade que representem satisfatoriamente uma técnica de sensoriamento espectral é o primeiro passo para se realizar a avaliação de desempenho da mesma. Para que a análise de desempenho se torne útil a de se estabelecer uma relação entre o comportamento da distribuição de probabilidade escolhida e os parâmetros sistêmicos que influenciam na formação da estatística de teste utilizada no sensoriamento. Para algumas estatísticas de teste o modelamento matemático analítico de tal parametrização é demasiadamente complexo ou até mesmo intratável, como é o caso do teste GID. Neste capítulo serão apresentados os métodos utilizados para a caracterização das estatísticas de teste do detector GID sob distribuições de probabilidade contínuas já conhecidas bem como o processos utilizados para a obtenção de dados, treinamento e avaliação do treinamento de uma ANN desenvolvida para estabelecer um relação entre os parâmetros das distribuições e os parâmetros sistêmicos. A determinação das CDFs está descrita na Seção 4.2, uma validação gráfica, que facilita o entendimento da escolha das CDFs e apresentada na Seção 4.3 e, por fim, os processos para desenvolvimento da ANN estão descritos na Seção 4.4.

4.1 Metodologia

Inicialmente foram geradas amostras com vários valores da estatística de teste T_{GID} definida em (2.13), sob as hipóteses \mathcal{H}_0 e \mathcal{H}_1 . As amostras do sinal transmitido pelos p usuários primários, que são dispostas na matriz $\mathbf{X} \in \mathbb{C}^{p \times n}$, tem distribuição Gaussiana de média nula e variância dependente da SNR, representando as variações de envoltória tipicamente presentes em sinais modulados e filtrados. Além disso, de modo a levar em conta o cenário mais realista no qual as potências de ruído e dos sinais recebidos nos SUs não são iguais entre si e são variantes no tempo, admitiu-se que ambas as potências são aleatoriamente determinadas para cada SU ao longo do tempo. Especificamente fez-se $\sigma_i^2 \sim \mathcal{U}[0.05\sigma_{\text{avg}}^2, 1.95\sigma_{\text{avg}}^2]$ e $p_i \sim \mathcal{U}[0.05s_{\text{avg}}, 1.95s_{\text{avg}}]$ a cada realização do sensoriamento espectral.

Tais amostras foram submetidas à estimação de máxima verossimilhança dos parâmetros de todas as distribuições contínuas disponíveis no software MATLAB, versão 2018a, as quais estão listadas na Tabela 4.1, com exceção da distribuição Beta (no MATLAB ela está restrita ao intervalo $[0, 1]$, não se aplicando à estatística de teste do GID, a qual pode estar contida em um intervalo diferente deste). Em seguida, como métricas do grau de aderência entre cada distribuição candidata e cada distribuição empírica gerada a partir dos valores de T_{GID} , calculou-se o erro quadrático médio (*Mean Squared Error*, MSE) entre elas, e também a média da estatística de kolmogorov-smirnov (*Mean Kolmogorov-Smirnov Statistic*, MKSS) [47]. Finalmente, as distribuições candidatas foram ranqueadas a partir da ordenação destas duas métricas.

A adoção da métrica MSE adicionalmente à métrica MKSS deu-se em razão do variado e às vezes pequeno número de ocorrências em que cada uma das distribuições candidatas foi aceita no teste de Kolmogorov-Smirnov, sob o típico nível de significância de 0,05. Sendo assim, o simples fato de uma determinada distribuição ser ou não aceita como representativa da correspondente amostra da estatística de teste T_{GID} definida em (2.13) não é, como esperado, suficiente para realizar a classificação supramencionada. Como demonstrado por meio das análises apresentadas nas Seções 4.2 e 4.3, este procedimento se mostrou adequado ao problema em questão e preciso o suficiente para que fossem determinadas as distribuições de T_{GID} nas hipóteses \mathcal{H}_0 e \mathcal{H}_1 .

Tabela 4.1: Lista de Distribuições do MATLAB, versão R2018a

	Distribuição	Nº de parâmetros
1	Beta	2
2	BirnbaumSaunders	2
3	Burr	3
4	Exponential	1
5	ExtremeValue	2
6	Gamma	2
7	Generalized Extreme Value	3
8	Generalized Pareto	3
9	Half Normal	2
10	Inverse Gaussian	2
11	Kernel	1
12	Logistic	2
13	Loglogistic	2
14	Lognormal	2
15	Nakagami	2
16	Normal	2
17	Rayleigh	2
18	Rician	2
19	Stable	4
20	t Location Scale	3
21	Weibull	2

4.2 CDFs candidatas sob hipótese \mathcal{H}_0 e \mathcal{H}_1

Para a obtenção das distribuições da estatística de teste definida em (2.13), foram geradas 70000 valores ¹ de T_{GID} para cada uma das hipóteses (\mathcal{H}_0 e \mathcal{H}_1) e para 70 combinações aleatórias dos parâmetros sistêmicos: $p = \{1, 2, \dots, 4\}$ PUs; $m = \{2, 3, \dots, 7\}$ SUs; $n = \{10, 20, \dots, 100\}$ amostras; $\text{SNR} = \{-15, -14, \dots, -5\}$ dB e $K = \{2, 3, \dots, 12\}$. Para cada grupo de 70000 valores da estatística de teste T_{GID} foi executado o teste Kolmogorov-Smirnov de aderência entre a CDF empírica e todas as distribuições da Tabela 4.1, calculando-se e armazenado-se o erro quadrático e a estatística de Kolmogorov-Smirnov entre elas. Por fim foram computados os valores médios dos valores armazenados, gerando as métricas MSE

¹Para a realização das pesquisas associadas ao presente trabalho foram, em diversos momentos, escolhidos dados para análise com números de amostras que se mostraram suficientes para que fosse solucionado o problema descrito na Seção 1.1. Ressalta-se que tais números poderiam ter sido menores se a seleção dos dados fosse amparada, por exemplo, em alguma análise de intervalo de confiança ou de projeto de experimentos. O mesmo vale para a definição do número de eventos de Monte Carlo nas simulações cujos resultados estão apresentados no Capítulo 5.

e MKSS. O **Algoritmo 1** mostra os passos desse procedimento, os quais foram aplicados da mesma maneira a \mathcal{H}_0 e \mathcal{H}_1 .

Algoritmo 1 Teste de aderência via MSE e MKSS

 $Runs \leftarrow 70000, Loops \leftarrow 70, \Sigma_{MSE} \leftarrow 0, \Sigma_{KSS} \leftarrow 0$
for $i = 1 : 1 : Loops$ **do**

Gerar combinação dos parâmetros sistêmicos

for $i = 1 : 1 : Runs$ **do**

 | Gerar T_j de acordo com (2.13)

end

 Gerar CDF empírica (CDF_e) do conjunto $\{T_j\}$

 Usando $\{T_j\}$, estimar parâmetros das CDFs da Tabela 4.1

 Gerar CDF teórica (CDF_t) para cada CDF da Tabela 4.1

 Computar KSS entre a CDF_e e as CDF_t

 Computar MSE entre a CDF_e e as CDF_t
 $\Sigma_{MSE} \leftarrow \Sigma_{MSE} + MSE$ para cada CDF_t
 $\Sigma_{KSS} \leftarrow \Sigma_{KSS} + KSS$ para cada CDF_t
end
 $MSE \leftarrow \Sigma_{MSE} / Loops$ para cada CDF_t
 $MKSS \leftarrow \Sigma_{KSS} / Loops$ para cada CDF_t

Como resultado, nas Tabelas 4.2 e 4.3 estão classificadas as 10 distribuições que apresentaram menores MSEs e menores MKSSs para as hipóteses \mathcal{H}_0 e \mathcal{H}_1 , respectivamente. Para ambas hipóteses e ambas as métricas, a distribuição melhor classificada foi a *Kernel*. Este resultado era esperado, pois a *Kernel* é uma distribuição não-paramétrica (portanto não aplicável ao problema em voga) definida por uma função núcleo e por um parâmetro responsável por controlar a suavidade da densidade de probabilidade obtida [48]. Na distribuição *Kernel* cria-se uma função para representar a densidade de probabilidade usando os dados da amostra sob análise, de forma análoga à construção de um histograma normalizado. No entanto, ao passo que tal histograma distribui os valores de frequência relativa em compartimentos discretos, uma distribuição *Kernel* soma funções núcleo para cada frequência relativa de forma a produzir uma função contínua e suave. Como consequência, é esperado que uma distribuição *Kernel* represente de forma fiel as amostras analisadas, principalmente quando o número de valores em cada amostra é elevado, como é o caso neste trabalho, no qual este número é 70000.

Por outro lado, a inclusão da distribuição *Kernel* na presente análise permite estabelecer uma referência de métrica, de forma que se possa melhor avaliar o

Tabela 4.2: Classificação da distribuições sob \mathcal{H}_0

Distribuição	MSE (ranque)	MKSS (ranque)
Kernel	$1,68 \times 10^{-6}$ (1)	$3,14 \times 10^{-3}$ (1)
Stable	$6,09 \times 10^{-5}$ (2)	$1,24 \times 10^{-2}$ (2)
Burr	$1,33 \times 10^{-4}$ (3)	$1,86 \times 10^{-2}$ (3)
Log-logistic	$1,74 \times 10^{-4}$ (4)	$2,26 \times 10^{-2}$ (4)
t location-scale	$1,98 \times 10^{-4}$ (5)	$2,68 \times 10^{-2}$ (6)
General extreme value	$2,23 \times 10^{-4}$ (6)	$2,43 \times 10^{-2}$ (5)
Logistic	$2,53 \times 10^{-4}$ (7)	$2,93 \times 10^{-2}$ (9)
Log-normal	$4,09 \times 10^{-4}$ (8)	$2,89 \times 10^{-2}$ (7)
Birnbaum-Saunders	$4,22 \times 10^{-4}$ (9)	$2,92 \times 10^{-2}$ (8)
Gamma	$5,31 \times 10^{-4}$ (10)	$3,26 \times 10^{-2}$ (10)

Tabela 4.3: Classificação da distribuições sob \mathcal{H}_1

Distribuição	MSE (ranque)	MKSS (ranque)
Kernel	$1,99 \times 10^{-6}$ (1)	$3,42 \times 10^{-3}$ (1)
Generalized extreme value	$4,38 \times 10^{-5}$ (2)	$1,02 \times 10^{-2}$ (2)
Stable	$1,37 \times 10^{-4}$ (3)	$1,93 \times 10^{-2}$ (3)
Burr	$1,73 \times 10^{-4}$ (4)	$2,21 \times 10^{-2}$ (4)
Log-logistic	$4,16 \times 10^{-4}$ (5)	$3,41 \times 10^{-2}$ (5)
t location-scale	$6,96 \times 10^{-4}$ (6)	$4,80 \times 10^{-2}$ (6)
Logistic	$7,27 \times 10^{-4}$ (7)	$4,91 \times 10^{-2}$ (9)
Log-normal	$9,50 \times 10^{-4}$ (8)	$4,86 \times 10^{-2}$ (7)
Birnbaum-Saunders	$9,75 \times 10^{-4}$ (9)	$4,90 \times 10^{-2}$ (8)
Gamma	$1,25 \times 10^{-3}$ (10)	$5,56 \times 10^{-2}$ (10)

grau de aderência das demais distribuições candidatas. Por exemplo, na hipótese \mathcal{H}_0 pode-se observar na Tabela 4.2 que a segunda distribuição melhor ranqueada tanto em termos de MSE quanto de MKSS é a distribuição *Stable*, resultando em um MSE (resp. MKSS) aproximadamente 36 (resp. 4) vezes maior que a distribuição *Kernel*. Já na hipótese \mathcal{H}_1 , pode-se observar na Tabela 4.3 que a segunda distribuição melhor ranqueada tanto em termos de MSE quanto de MKSS é a distribuição *Generalized Extreme Value (GEV)*, a qual resultou em um MSE (resp. MKSS) aproximadamente 22 (resp. 3) vezes maior que a distribuição *Kernel*.

Em suma, deste ponto em diante foi adotada a distribuição *Stable* para caracterizar a estatística de teste T_{GID} definida em (2.13) sob a hipótese \mathcal{H}_0 . Já sob a hipótese \mathcal{H}_1 adotou-se a distribuição *GEV*.

4.2.1 Distribuições de *Stable* e *Generalized Extreme Value*

Não se pode expressar analiticamente a PDF e a CDF *Stable*, exceto para certos valores de seus parâmetros [49]. Uma variável aleatória X com distribuição *Stable* tem como PDF a transformada de Fourier da sua função característica $\varphi(t)$, ou seja

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) e^{-jxt} dt, \quad (4.1)$$

em que

$$\varphi(t) = \exp [jt\delta - |\gamma t|^\alpha (1 - j\beta \operatorname{sgn}(t)\Phi)], \quad (4.2)$$

sendo $\operatorname{sgn}(t)$ o sinal de t e

$$\Phi = \begin{cases} \tan\left(\frac{\pi\alpha}{2}\right) & \alpha \neq 1 \\ -\frac{2}{\pi} \log|t| & \alpha = 1 \end{cases}. \quad (4.3)$$

Os parâmetros α ($0 < \alpha \leq 2$) e β ($-1 \leq \beta \leq 1$) determinam o formato da distribuição, γ ($0 < \gamma < \infty$) é um fator de escala e δ ($-\infty < \delta < \infty$) corresponde à sua localização.

A CDF *Stable* pode ser obtida por integração direta, na maior parte das vezes solucionada numericamente, ou seja,

$$F(x) = \int_{-\infty}^x f(z) dz. \quad (4.4)$$

Embora a distribuição *Stable* tenha um tratamento analítico complexo, ela já consta do rol de funções cujo tratamento numérico se tornou maduro e bem aceito o suficiente para que expressões que a contenham possam ser consideradas como fechadas (*closed form*). Como exemplo, a PDF, a CDF e a CDF inversa da distribuição *Stable* estão incorporadas no MATLAB, podendo ser inseridas normalmente em cálculos, como acontece com a função erro complementar e outras que só podem ser operadas numericamente.

Já a distribuição GEV possui tratamento analítico mais simples que a *Stable*.

Uma variável aleatória com distribuição GEV tem PDF

$$f(x) = \frac{1}{\sigma} t(x)^{k+1} e^{-t(x)}, \quad (4.5)$$

em que

$$t(x) = \begin{cases} [1 + k(\frac{x-\mu}{\sigma})]^{-1/k} & \text{se } k \neq 0 \\ e^{-\frac{(x-\mu)}{\sigma}} & \text{se } k = 0, \end{cases} \quad (4.6)$$

em que μ ($-\infty < \mu < \infty$) é o parâmetro de localização, σ ($\sigma > 0$) refere-se à escala e k ($-\infty < k < \infty$) determina o formato da distribuição.

A CDF GEV é simplesmente

$$F(x) = e^{-t(x)}. \quad (4.7)$$

4.3 Verificação gráfica das CDFs

Como se pode concluir a partir das Figuras 4.1 e 4.2, de fato as distribuições *Stable* e GEV podem representar de forma satisfatória a estatística de teste T_{GID} definida em (2.13) sob as hipóteses \mathcal{H}_0 e \mathcal{H}_1 , respectivamente. Para a construção dessas figuras foram utilizados 3 conjuntos de parâmetros sistêmicos distintos. Pode-se verificar que as CDFs teóricas estão praticamente sobrepostas às correspondentes CDFs empíricas em todos os casos.

4.4 Desenvolvimento da rede neural

Nesta seção são descritos os procedimentos para geração dos conjuntos de treinamento, validação e testes. Os mesmos serão utilizados no processo de aprendizagem da ANN. Ao fim deste, a ANN deverá ser capaz de realizar o mapeamento entre os parâmetros sistêmicos do sensoriamento espectral e aqueles relacionados às CDFs que representam a estatística de teste GID sob \mathcal{H}_0 e \mathcal{H}_1 .

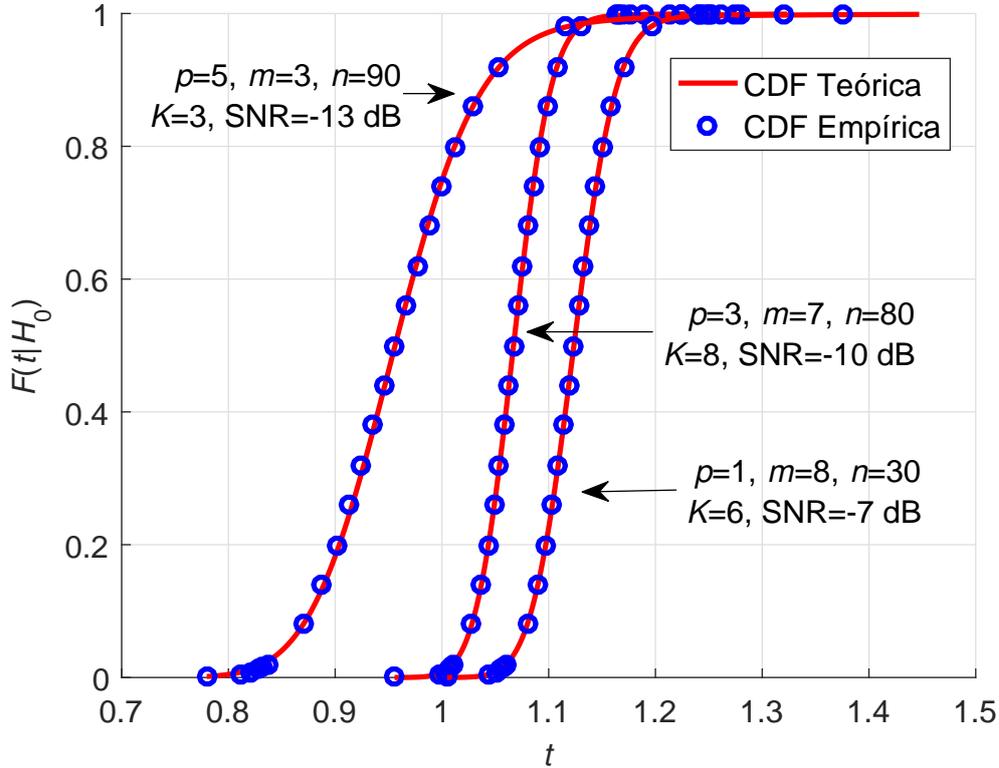


Figura 4.1: CDFs empíricas e teóricas sob hipótese \mathcal{H}_0

4.4.1 Obtenção dos dados

Para geração dos dados foi desenvolvida uma rotina no software MATLAB, de acordo com o **Algoritmo 2**. As combinações de parâmetros sistêmicos de entrada da ANN foram $p = \{1, 2, 3, 4\}$ PUs, $m = \{2, 3, 4, 5, 6, 7\}$ SUs, $n = \{10, 20, \dots, 100\}$ amostras por SU, $\text{SNR} = \{-15, -14, \dots, -5\}$ dB, e fator de Rice $K = \{2, 3, \dots, 12\}$, resultando em um conjunto de 29040 valores. Para cada combinação, os parâmetros das CDFs sob \mathcal{H}_0 e \mathcal{H}_1 foram estimados sob o critério de máxima verossimilhança, utilizando a função `fitdist` do MATLAB, a partir de 50000 amostras da estatística de teste T_{GID} definida em (2.13).

A Tabela 4.4 mostra algumas estatísticas das estimativas dos parâmetros das distribuições, no conjunto completo de parâmetros sistêmicos simulados. Tais estatísticas permitirão, mais adiante, a avaliação da precisão da ANN na tarefa de mapear os parâmetros sistêmicos nos parâmetros das distribuições *Stable* e *GEV*. Esta avaliação é realizada comparando-se a variabilidade de cada parâmetro de

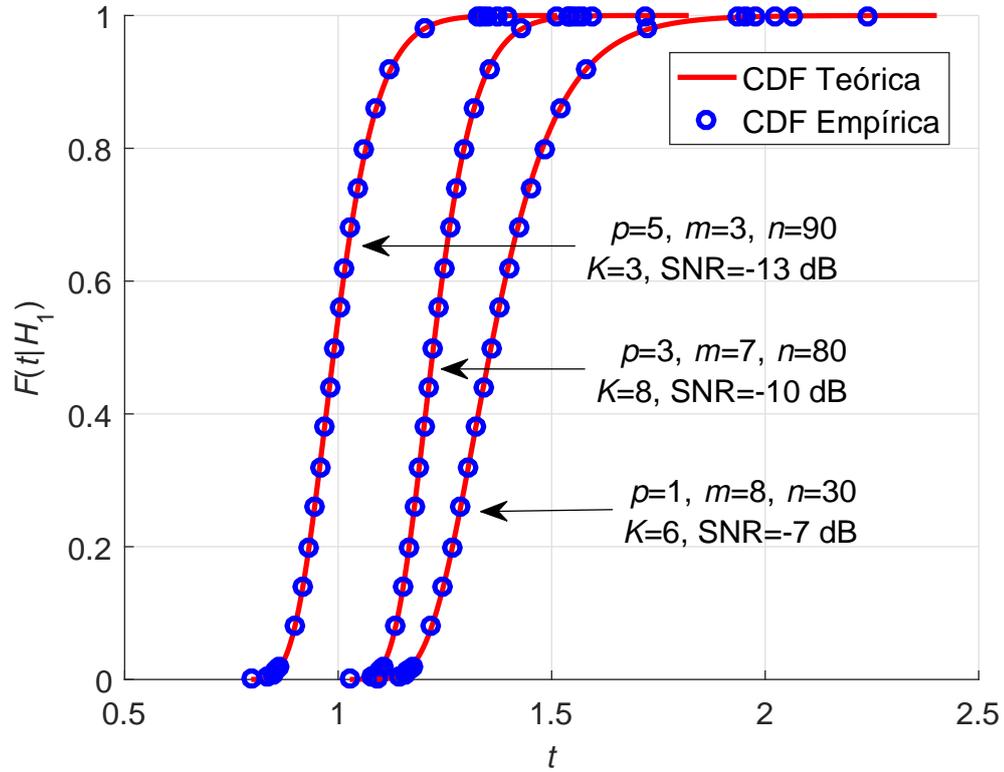


Figura 4.2: CDFs empíricas e teóricas sob hipótese \mathcal{H}_1 .

uma distribuição, para uma dada combinação de parâmetros sistêmicos, com a variabilidade máxima daquele parâmetro no conjunto completo de parâmetros sistêmicos de entrada.

Tabela 4.4: Estatísticas dos parâmetros das distribuições *Stable* e GEV

	Máximo	Mínimo	Média	Desvio padrão
α	1,97410535	1,28317473	1,83739486	0,12514507
β	1	1	1	0
γ	0,09752231	0,01875923	0,04150517	0,02172807
δ	1,14610722	0,88813109	1,01183088	0,06887544
k	0,34484909	-0,12527059	-0,01009917	0,07374319
σ	0,24991361	0,03108020	0,09902625	0,04310061
μ	1,66034555	0,87178048	1,11753791	0,13899918

É interessante notar na Tabela 4.4 que, para os parâmetros sistêmicos aqui considerados, o parâmetro β da distribuição de *Stable* é sempre 1. Além disso, observa-se que, na média, os valores de α são altos, o que faz com que β não

Algoritmo 2 Geração dos dados de treinamento da ANN

```

Loop ← 1 ; Runs ← 50000
for p = 1 : 1 : 4 do
  m ← 2; n ← 10 ; K ← 2; SNR ← -15
  for m = 2 : 1 : 7 do
    n ← 10 ; K ← 2; SNR ← -15
    for SNR = -15 : 1 : -5 do
      n ← 10 ; K ← 2
      for n = 10 : 10 : 100 do
        K ← 2
        for K = 2 : 1 : 12 do
          for i = 1 : 1 : Runs do
            Gerar  $T_i$  de acordo
            com (2.13), sob  $\mathcal{H}_0$  e  $\mathcal{H}_1$ 
          end
          Estimar  $\alpha, \beta, \gamma$  e  $\delta$  da PDF Stable sob  $\mathcal{H}_0$ 
          Estimar  $\mu, \sigma$  e  $k$  da PDF GEV sob  $\mathcal{H}_1$ 
          Inputs(loop, :) ← [p m n K SNR]
          Outputs(loop, :) ← [ $\alpha \beta \gamma \delta \mu \sigma k$ ]
          Loop ← Loop + 1
        end
      end
    end
  end
end
end
end
end
end

```

influencie de maneira significativa no formato da PDF [49]. Como consequência, seguramente se pode adotar $\beta = 1$, independentemente do parâmetro sistêmico de entrada.

Em suma, haverá uma ANN para mapear os parâmetros sistêmicos de entrada $\{p, m, n, K, SNR\}$ nos parâmetros de saída $\{k, \sigma, \mu\}$ para a distribuição GEV, e outra para mapear os primeiros nos parâmetros $\{\alpha, \gamma, \delta\}$ para a distribuição *Stable*. Como $\beta = 1$, não é necessário ser definido como um parâmetro de saída da segunda ANN. Portanto, os dados gerados serão utilizado em duas ANNs, ambas com cinco variáveis de entrada e três de saída.

4.4.2 Treinamento da rede

Conforme apresentado no capítulo 3 Seção 3.2 a ANN que melhor se enquadra para esse trabalho é a do tipo MLP. Para esse tipo de ANN, a equação de

saída é a fórmula geral da interpolação linear [41]. Assim, após o processo de aprendizagem tal ANN é capaz de estimar qualquer função não-linear contínua com determinada precisão. Adicionalmente, devido à sua capacidade de interpolação, permite estimar parâmetros das CDFs com valores das variáveis sistêmicas intermediários aos utilizados no treinamento.

A função de ativação utilizada nos neurônios da ANN em questão é do tipo tangente hiperbólica. Em [50], o autor cita várias referências que mostram que qualquer função não-linear pode ser aproximada por uma rede do tipo MLP que utiliza tal função de ativação. Tal autor também menciona que há várias maneiras de se estimar o número necessário de neurônios e de camadas intermediárias (*hidden layers*) mas que, em geral, com apenas uma camada intermediária é possível estimar uma função não-linear sem erros. Isto é possível se o número de neurônios excede em 1 o número de elementos do conjunto de dados de treinamento. Entretanto, a complexidade computacional da ANN aumenta à medida que o número de neurônios cresce [51]. Neste trabalho, o número de neurônios adotado foi escolhido em função do erro produzido pelas ANNs, o que se explica com mais detalhes adiante.

Para o processo de aprendizagem dividiu-se o conjunto de dados em três partes. A primeira, contendo cerca de 70 % dos dados gerados foi usada para o ajuste dos pesos sinápticos (treinamento). Já o restante foi usado nos processos de validação (20%) e testes (10%). O algoritmo utilizado no processo de treinamento foi o Levenberg Marquardt [52]. O número de épocas foi definido em função de uma grande quantidade de testes realizados anteriormente com quantidades de épocas muito maiores. Observou-se que a partir de 2000 épocas de treinamento o erro quadrático médio não reduzia. Assim, afim de evitar o *overfitting*, limitou-se ao número máximo de 2000 épocas. Trata-se de uma quantidade de épocas relativamente pequena, que viabiliza o retreinamento da ANN sempre que necessário, em um curtíssimo espaço de tempo. Além do número máximo de épocas, em todos os testes realizados, o treinamento foi encerrado sempre que o erro de validação permanecesse crescente durante 15 épocas. Nestes casos, o último conjunto de pesos sinápticos, com erro mínimo, era definido como ótimo.

A ANN adotada foi simulada utilizando-se o pacote de redes neurais do MATLAB (*neural network toolbox*, mais recentemente renomeado para *deep learning toolbox*) e foi configurada de acordo com os seguintes parâmetros:

- Função de ativação: tangente hiperbólica;

- Avaliação de desempenho: *mean squared error*, MSE;
- MSE desejado: 0;
- Número máximo de épocas para treinamento: 2000;
- Taxa de aprendizagem: 0,05;
- Porcentagem dos dados de treinamento utilizado para validação: 20 %;
- Porcentagem dos dados de treinamento utilizado para teste: 10 %.

O treinamento foi realizado a partir do conjunto gerado de acordo com **Algoritmo 2**. Como resultado, as Tabelas 4.5 e 4.6 apresentam a média dos MSEs entre os valores dos parâmetros de saída de cada uma das ANNs em relação aos parâmetros reais (estimados por máxima verossimilhança) das distribuições *Stable* e GEV, respectivamente, em função do número de neurônios da camada intermediária. Em caráter apenas informativo, o número de épocas utilizadas no treinamento é também apresentado nestas tabelas, para cada número de neurônios utilizado. Analisando tais tabelas verifica-se que o aumento do número de neurônios faz com que o MSE diminua, com o esperado. No entanto, sob a hipótese \mathcal{H}_0 nota-se que a partir de 15 neurônios não se tem redução significativa do MSE. Sob \mathcal{H}_1 , o mesmo acontece quando o número de neurônios se aproxima de 30. Portanto, foram utilizados 15 neurônios na camada intermediária da ANN que estima os parâmetros da distribuição *Stable* e 30 para a ANN que estima os parâmetros da GEV.

Tabela 4.5: MSE da ANN sob \mathcal{H}_0

Nº de neurônios	MSE	Nº de épocas
5	$9,95 \times 10^{-6}$	331
10	$6,40 \times 10^{-6}$	672
15	$6,38 \times 10^{-6}$	206
20	$6,37 \times 10^{-6}$	119
25	$6,32 \times 10^{-6}$	167

Tabela 4.6: MSE da ANN sob \mathcal{H}_1

Nº de neurônios	MSE	Nº de épocas
5	$1,10 \times 10^{-4}$	615
10	$2,63 \times 10^{-5}$	1338
15	$1,14 \times 10^{-5}$	687
20	$8,08 \times 10^{-6}$	898
25	$6,12 \times 10^{-6}$	878
30	$4,65 \times 10^{-6}$	711
35	$4,18 \times 10^{-6}$	1087

As ANNs resultantes foram convertidas em funções do MATLAB, as quais estão descritas no **Apêndice A**. Especificamente,

$$\text{NnH0}(p \ m \ \text{SNR} \ n \ K) \quad (4.8)$$

retorna os parâmetros α , γ e δ da distribuição *Stable* que caracteriza a estatística de teste GID sob \mathcal{H}_0 (lembrando que $\beta = 1$ sempre), em função dos parâmetros sistêmicos p , m , SNR e K . Analogamente,

$$\text{NnH1}(p \ m \ \text{SNR} \ n \ K) \quad (4.9)$$

retorna os parâmetros k , σ e μ da distribuição GEV que caracteriza a estatística de teste GID sob \mathcal{H}_1 .

4.4.3 Avaliação do treinamento

Para avaliação de desempenho do treinamento das ANNs resultantes, o mesmo conjunto de variáveis sistêmicas utilizado na geração dos dados de treinamento foi aplicado como entrada para as funções supracitadas. Após obter as respostas das ANNs para cada uma das 29040 combinações de parâmetros de entrada, foram calculados os erros entre tais respostas e os parâmetros reais obtidos por estimação de máxima verossimilhança. As Figuras 4.3 e 4.4 apresentam os histogramas dos erros de mapeamento para as hipóteses \mathcal{H}_0 e \mathcal{H}_1 , respectivamente. Comparando as dispersões destes histogramas, medidas pelo desvio padrão (*Standard*

Deviation, SD), com os desvios padrão dados na Tabela 4.4 para cada parâmetro, conclui-se que os erros de estimação produzidos pelas ANNs são de fato pequenos. Além disso, nota-se que os erros estão simetricamente dispostos em torno de zero, o que significa que o mapeamento realizado pelas ANNs não produz resultados polarizados. Estas conclusões são ratificadas no Capítulo 5, por meio da comparação do desempenho simulado do sensoriamento espectral com aquele obtido teoricamente, a partir dos parâmetros gerados pelas ANNs.

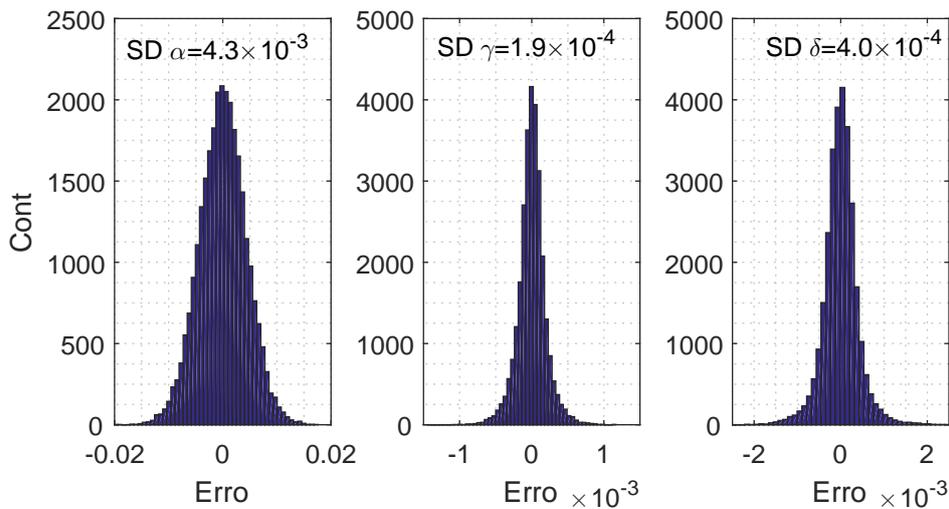


Figura 4.3: Histogramas dos erros de mapeamento sob \mathcal{H}_0

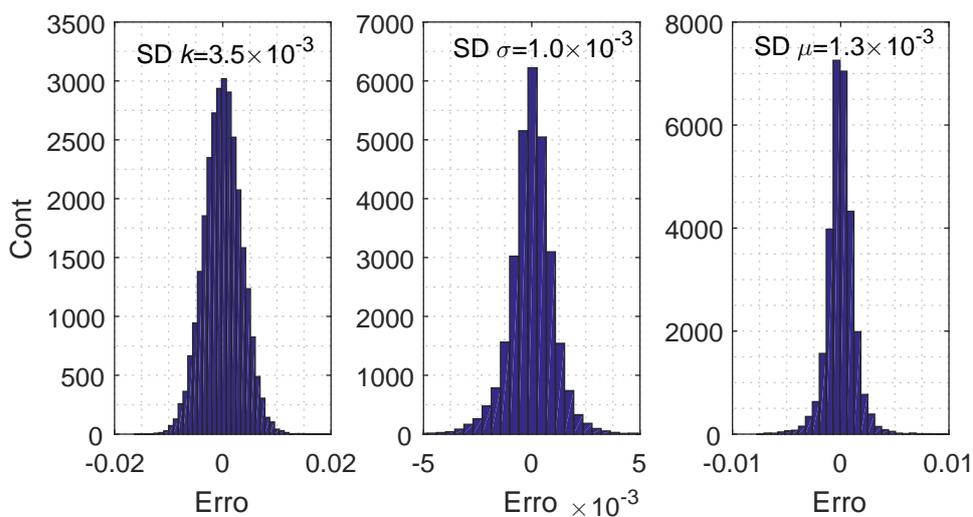


Figura 4.4: Histogramas dos erros de mapeamento sob \mathcal{H}_1

Capítulo 5

Avaliação da Parametrização das Distribuições da Estatística Teste GID

Até o momento verificou-se a acurácia da rede neural em mapear os parâmetros das distribuições utilizando como método de comparação as CDFs teóricas e empíricas sobre as hipóteses \mathcal{H}_0 e \mathcal{H}_1 e as estatísticas de tais parâmetros. Nesse capítulo será feita uma avaliação aplicada ao desempenho teórico e simulado das estatística de teste GID em um cenário de sensoriamento espectral.

5.1 Método de avaliação teórico

Em termos da análise de desempenho teórico, com base nas equações (1.1), (1.2) e (1.3) tem-se

$$P_{\text{fa}} = 1 - F_S(\lambda, \alpha, 1, \gamma, \delta), \quad (5.1)$$

em que $F_S(\lambda, \alpha, 1, \gamma, \delta)$ é o valor da CDF *Stable* parametrizada por $\alpha, \beta = 1, \gamma, \delta$, no ponto $t = \lambda$,

$$P_{\text{d}} = 1 - F_G(\lambda, k, \sigma, \mu), \quad (5.2)$$

sendo $F_G(\lambda, k, \sigma, \mu)$ o valor da CDF GEV parametrizada por λ, k, σ, μ , no ponto $t = \lambda$, com o limiar de decisão determinado de acordo com a P_{fa} alvo, por meio de

$$\lambda = F_S^{-1}(1 - P_{fa}, \alpha, 1, \gamma, \delta), \quad (5.3)$$

em que $F_S^{-1}(1 - P_{fa}, \alpha, 1, \gamma, \delta)$ é o valor da função *Stable* inversa no ponto $1 - P_{fa}$.

Estes cálculos podem ser realizados numericamente, sendo que no MATLAB são respectivamente computados por meio de

$$P_{fa} = 1 - \text{cdf}('Stable', \lambda, \alpha, 1, \gamma, \delta), \quad (5.4)$$

$$P_d = 1 - \text{cdf}('GeneralizedExtremeValue', \lambda, k, \sigma, \mu), \quad (5.5)$$

$$\lambda = \text{icdf}('Stable', 1 - P_{fa}, \alpha, 1, \gamma, \delta). \quad (5.6)$$

Os resultados obtidos através dos cálculos teóricos serão utilizados posteriormente como comparação com os resultados obtidos via simulação de Monte Carlo.

5.2 ROCs teóricas e empíricas

Para os resultados de simulação subsequentemente apresentados, cada ponto sobre uma curva ROC foi gerado a partir de 50000 eventos de Monte Carlo.

Para ilustrar o grau de precisão do mapeamento de parâmetros realizado pelas redes neurais, a Figura 5.1 mostra o desempenho teórico e simulado do detector GID para três conjuntos de parâmetros sistêmicos arbitrariamente escolhidos: $\{p = 4, m = 3, n = 50, K = 12, \text{SNR} = -5 \text{ dB}\}$, $\{p = 3, m = 5, n = 80, K = 3, \text{SNR} = -11 \text{ dB}\}$ e $\{p = 1, m = 4, n = 20, K = 7, \text{SNR} = -10 \text{ dB}\}$. Os correspondentes parâmetros das distribuições, mapeados pelas ANNs são: $\{\alpha \approx 1,8011, \beta = 1, \gamma \approx 0,0487, \delta \approx 0,9637, k \approx -0,0466, \sigma \approx 0,1888, \mu \approx 1,2856\}$, $\{\alpha \approx 1,9166, \beta = 1, \gamma \approx 0,0262, \delta \approx 1,0242, k \approx -0,0420, \sigma \approx$

0,0573, $\mu \approx 1,0825$) e $\{\alpha \approx 1,7426, \beta = 1, \gamma \approx 0,0425, \delta \approx 1,0344, k \approx 0,0710, \sigma \approx 0,0844, \mu \approx 1,0864\}$. Pode-se perceber na Figura 5.1 que as ROCs empíricas estão praticamente sobrepostas às correspondentes ROCs teóricas, demonstrando que a distribuição *Stable* de fato representa de forma satisfatória a estatística de teste GID sob a hipótese \mathcal{H}_0 , bem como a GEV representa de forma satisfatória a estatística de teste GID sob a hipótese \mathcal{H}_1 . Adicionalmente, conclui-se que as redes neurais aqui desenvolvidas estimam os parâmetros de tais distribuições com suficiente precisão, permitindo que os cálculos teóricos de desempenho sejam igualmente precisos.

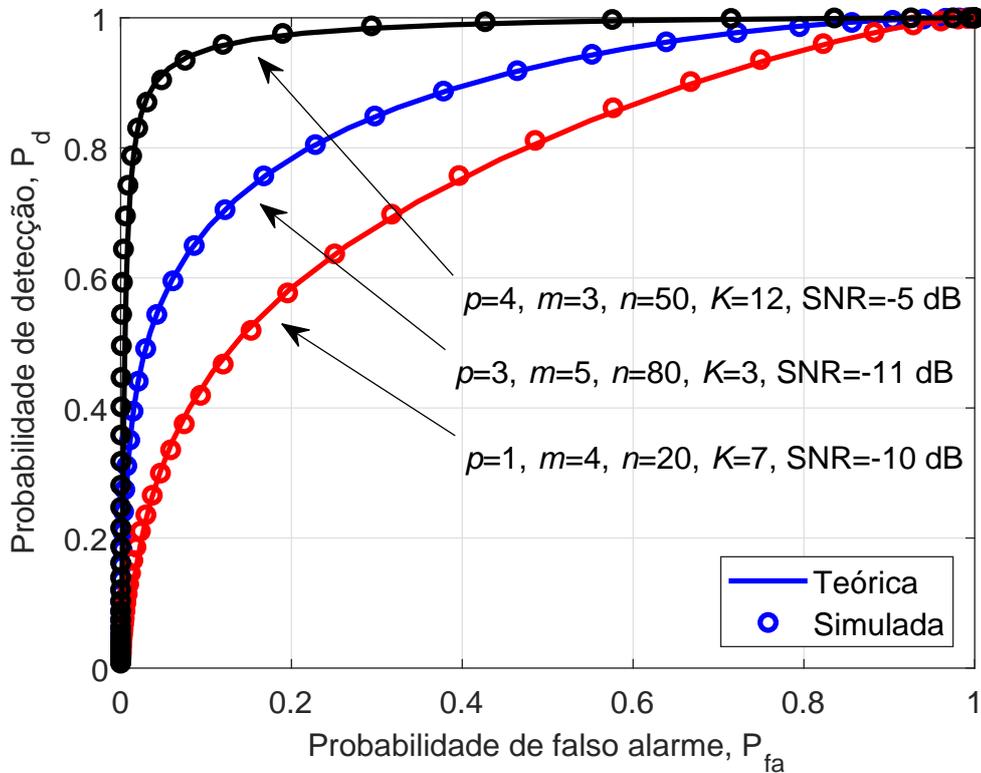


Figura 5.1: Desempenho teórico e simulado do GID, para parâmetros presentes no treinamento da ANN.

Apenas como ilustração de cálculo, no caso da curva mais acima na Figura 5.1, para $P_{fa} = 0,1$ tem-se $\lambda = \text{icdf}(\text{'Stable'}, 0.9, 1.8011, 1, 0.0487, 0.9637) \approx 1,0719$ e, como consequência, $P_d = 1 - \text{cdf}(\text{'GeneralizedExtremeValue'}, 1.0719, -0.0466, 0.1888, 1.2856) \approx 0,9509$. Para a curva mais abaixo, ainda mantendo $P_{fa} = 0,1$ como referência, tem-se $\lambda = \text{icdf}(\text{'Stable'}, 0.9, 1.7426, 1, 0.0425, 1.0344) \approx 1,1351$ e $P_d = 1 - \text{cdf}(\text{'GeneralizedExtremeValue'}, 1.1351, 0.0710, 0.0844,$

1.0864) $\approx 0,4414$ valores estes em consonância com aqueles observados na referida figura.

5.3 Capacidade de interpolação e extrapolação

Como a ANN do tipo MLP é uma interpoladora por construção, é possível mapear parâmetros sistêmicos diferentes daqueles utilizados no treinamento. Como exemplo, a Figura 5.2 apresenta ROCs teóricas e simuladas, considerando as seguintes combinações de parâmetros sistêmicos: $\{p = 2, m = 4, n = 25, K = 7,7, \text{SNR} = -10,3 \text{ dB}\}$ e $\{p = 3, m = 6, n = 64, K = 4,8, \text{SNR} = -12,5 \text{ dB}\}$. Tais parâmetros foram mapeados pelas ANNs nos seguintes parâmetros das distribuições *Stable* e GEV: $\{\alpha \approx 1,9804, \beta = 1, \gamma \approx 0,0175, \delta \approx 1,0486, k \approx -0,1046, \sigma \approx 0,0577, \mu \approx 1,1591\}$ e $\{\alpha \approx 1,7679, \beta = 1, \gamma \approx 0,0407, \delta \approx 1,0276, k \approx 0,0479, \sigma \approx 0,0838, \mu \approx 1,0811\}$, respectivamente.

Constatou-se ainda que em situações de bom desempenho do sensoriamento, por exemplo $P_d > 0,8$ e $P_{fa} < 0,2$, é possível utilizar as ANNs desenvolvidas para extrapolar resultados, como se pode observar na Figura 5.2 no caso dos parâmetros sistêmicos $\{p = 4, m = 7, n = 150, K = 8, \text{SNR} = -11,5 \text{ dB}\}$, os quais foram mapeados nos seguintes parâmetros das distribuições *Stable* e GEV: $\{\alpha \approx 1,9264, \beta = 1, \gamma \approx 0,0234, \delta \approx 1,0553, k \approx -0,0493, \sigma \approx 0,0467, \mu \approx 1,0965\}$. Nota-se que a curva teórica se encontra praticamente sobreposta aos correspondente resultados obtidos por simulação.

5.4 Área sob a curva ROC

Então, por fim, com o objetivo de identificar algum possível resultado insatisfatório de mapeamento entre parâmetros, todos aqueles utilizados no treinamento as ANNs a partir dos 29040 conjuntos de parâmetros sistêmicos foram utilizados para gerar parâmetros das distribuições *Stable* e GEV estimados por máxima verossimilhança pelo MATLAB, e também parâmetros mapeados pelas ANNs desenvolvidas. Cada par de parâmetros estimados e mapeados foi utilizado em uma simulação de Monte Carlo, gerando o correspondente par de ROCs do GID. O histograma dos erros entre as AUCs de todos os pares de ROCs está apresentado na Figura 5.3. Esse histograma apresenta média $\approx 3,39 \times 10^{-5}$ e desvio

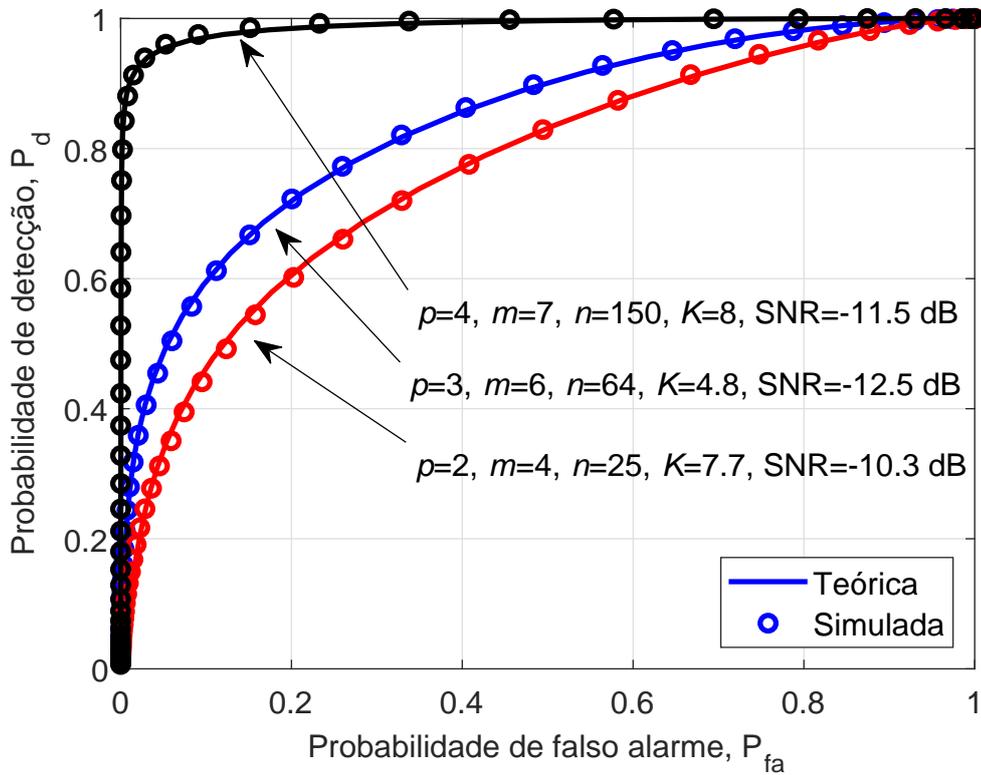


Figura 5.2: Desempenho teórico e simulado do GID, para parâmetros não presentes no treinamento da ANN.

padrão $\approx 0,0028$. Levando em conta os limites mencionados na Seção 2.3 de uma AUC, conclui-se que, de um modo geral, os erros de mapeamento produzidos pelas ANNs são consideravelmente pequenos, praticamente imperceptíveis entre duas ROCs em um gráfico de desempenho.

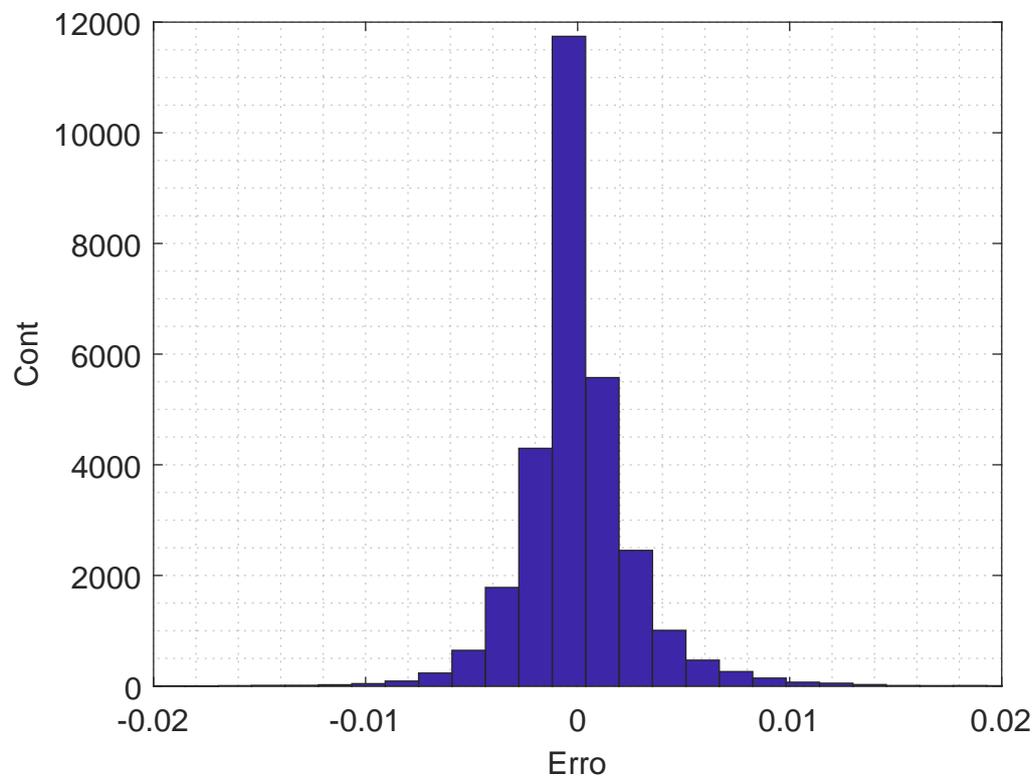


Figura 5.3: Histograma do erro entre AUC teóricas e empíricas.

Capítulo 6

Conclusões e trabalhos futuros

6.1 Conclusões

A estatística de teste GID, inspirada no índice de Gini, é um dos mais recentes testes proposto para sensoriamento espectral aplicado em rádio cognitivo. Suas características o tornam forte candidato para ser utilizado em dispositivos IoT já que a ubiquidade de tais dispositivos certamente irá esbarrar no problema de escassez espectral. A tecnologia de rádio cognitivo é considerada um pilar fundamental para que a rede de comunicação móvel de quinta geração emergja nos próximos anos. Os CRs serão utilizados principalmente para compartilhamento das faixas de frequência [2, 53, 54] o que faz com que estudos relacionado ao sensoriamento espectral sejam sempre interessantes.

Nessa dissertação foram apresentados os procedimentos e os resultados da análise de aderência da estatística de teste GID a diversas distribuições de probabilidade. Demonstrou-se que a distribuição *Stable* caracteriza com adequada precisão a estatística de teste sob a hipótese \mathcal{H}_0 , enquanto a distribuição *Generalized Extreme Value* melhor se aplica a \mathcal{H}_1 . Redes neurais artificiais foram desenvolvidas para estabelecer o mapeamento entre os parâmetros sistêmicos e os parâmetros que caracterizam tais distribuições, permitindo cálculos teóricos do desempenho e do limiar de decisão do sensoriamento espectral. Estes cálculos não somente podem auxiliar em pesquisas relacionadas ao GID, mas também como suporte à parametrização deste detector em aplicações práticas, sem a necessidade de demoradas simulações computacionais.

Grande será a contribuição dessa dissertação pois os procedimentos e resul-

tados aqui apresentados encorajam sua aplicação em outras estatísticas de teste cujo desenvolvimento analítico para obtenção das suas distribuições sob \mathcal{H}_0 ou \mathcal{H}_1 seja muito complexo ou até mesmo intratável.

6.2 Proposição para trabalhos futuros

Existem algumas sugestões para futuros trabalhos com objetivo de dar continuidade a esta dissertação. A primeira sugestão seria o estudo mais abrangente utilizando uma combinação maior de parâmetros sistêmicos, sobretudo aumentando números de usuários secundários afim de verificar o desempenho do detector GID em cenários com grande densidade de conexão, que é o que se espera da rede 5G.

Um segundo trabalho poderia ser voltado para a confecção de um protótipo. As funções que descrevem as redes neurais desenvolvidas poderiam ser exportadas para linguagens mais abrangentes como C e Java, ou até mesmo implementá-las em FPGAs (*field programmable gate arrays*) ou DSPs (*digital signal processors*) para execução prática em, por exemplo, técnicas adaptativas de operação do sensoriamento espectral com o detector do GID.

Outra sugestão é utilizar os procedimentos apresentados para obtenção das distribuições sob \mathcal{H}_0 ou \mathcal{H}_1 de outras técnicas de sensoriamento espectral e, posteriormente, investigar se redes neurais artificiais também podem ser utilizadas para o mapeamento de tais distribuições nos parâmetros sistêmicos.

Por fim, devido a flexibilidade apresentada pelas ANNs, é possível investigar a utilização de redes neurais dentro de diferentes cenário de rádios cognitivos. Um tema interessante é desenvolver uma ANN para sensoriamento espectral dinâmico onde a técnica escolhida para gerar as hipóteses \mathcal{H}_0 ou \mathcal{H}_1 é definida em função dos parâmetros sistêmicos.

Referências Bibliográficas

- [1] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [2] F. Hu, B. Chen, and K. Zhu, “Full spectrum sharing in cognitive radio networks toward 5g: A survey,” *IEEE Access*, vol. 6, pp. 15 754–15 776, 2018.
- [3] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, “Context aware computing for the internet of things: A survey,” *IEEE communications surveys & tutorials*, vol. 16, no. 1, pp. 414–454, 2014.
- [4] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, “Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey,” *Computer networks*, vol. 50, no. 13, pp. 2127–2159, 2006.
- [5] J. Mitola, G. Q. Maguire *et al.*, “Cognitive radio: making software radios more personal,” *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.
- [6] T. Yucek and H. Arslan, “A survey of spectrum sensing algorithms for cognitive radio applications,” *IEEE communications surveys & tutorials*, vol. 11, no. 1, pp. 116–130, 2009.
- [7] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, “Spectrum sensing for cognitive radio: State-of-the-art and recent advances,” *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, 2012.
- [8] B. Wang and L. KJ Ray, “Advances in cognitive radio networks: A survey,” *IEEE Journal of selected topics in signal processing*, vol. 5, no. 1, pp. 5–23, 2011.
- [9] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, “Cooperative spectrum sensing in cognitive radio networks: A survey,” *Physical communication*, vol. 4, no. 1, pp. 40–62, 2011.

-
- [10] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "Ieee 802.22: the first worldwide wireless standard based on cognitive radios," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*. Ieee, 2005, pp. 328–337.
- [11] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise level," in *2011 IEEE International Conference on Communications (ICC)*. IEEE, 2011, pp. 1–5.
- [12] R. Li, L. Huang, Y. Shi, and H.-C. So, "Gerschgorin disk-based robust spectrum sensing for cognitive radio," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 7278–7282.
- [13] D. A. Guimarães, "Robust test statistic for cooperative spectrum sensing based on the gerschgorin circle theorem," *IEEE Access*, vol. 6, pp. 2445–2456, 2018.
- [14] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE transactions on communications*, vol. 55, no. 1, pp. 21–24, 2007.
- [15] L. Huang, Y. Xiao, H. C. So, and J. Fang, "Accurate performance analysis of hadamard ratio test for robust spectrum sensing," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 750–758, 2015.
- [16] L. Wei and O. Tirkkonen, "Spectrum sensing in the presence of multiple primary users," *IEEE Transactions on Communications*, vol. 60, no. 5, pp. 1268–1277, 2012.
- [17] L. Huang, C. Qian, Y. Xiao, and K. Q. Zhang, "Performance analysis of volume-based spectrum sensing for cognitive radio," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 317–330, 2015.
- [18] J. S. Neto, D. A. Guimarães, and R. A. A. de Souza, "Estimação empírica da distribuição da estatística de teste para o sensoriamento espectral por máximo autovalor sob a hipótese H_1 ," in *XXX Simpósio Brasileiro de Telecomunicações, Brasília, DF, Setembro 2012*.
- [19] N. Baldo and M. Zorzi, "Learning and adaptation in cognitive radios using neural networks," in *2008 5th IEEE Consumer Communications and Networking Conference*. IEEE, 2008, pp. 998–1003.

- [20] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, “Deep learning for rf device fingerprinting in cognitive communication networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.
- [21] S. Pattanayak, P. Venkateswaran, and R. Nandi, “Artificial neural networks for cognitive radio: a preliminary survey,” in *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*. IEEE, 2012, pp. 1–4.
- [22] X. Dong, Y. Li, C. Wu, and Y. Cai, “A learner based on neural network for cognitive radio,” in *2010 IEEE 12th International Conference on Communication Technology*. IEEE, 2010, pp. 893–896.
- [23] J. J. Popoola and R. van Olst, “Application of neural network for sensing primary radio signals in a cognitive radio environment,” in *IEEE Africon’11*. IEEE, 2011, pp. 1–6.
- [24] M. R. Vyas, D. K. Patel, and M. Lopez-Benitez, “Artificial neural network based hybrid spectrum sensing scheme for cognitive radio,” in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.
- [25] P. Tilghman, “Will rule the airwaves: A darpa grand challenge seeks autonomous radios to manage the wireless spectrum,” *IEEE Spectrum*, vol. 56, no. 6, pp. 28–33, June 2019.
- [26] M. Saber, K. A. Hatim, E. R. Abdessamad, and S. Rachid, “Energy detection and artificial neural networks for spectrum sensing based on real world signals,” *Conference: Workshop on Advancing Technology for Humanitarian Sciences*, 07 2018.
- [27] D. A. Guimarães, “Gini index inspired robust detector for spectrum sensing over ricean channels,” *Electronics Letters*, 2018.
- [28] A. Ghasemi and E. S. Sousa, “Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs,” *IEEE Communications magazine*, vol. 46, no. 4, pp. 32–39, 2008.
- [29] J. S. Neto and D. A. Guimarães, “Sensoriamento espectral cooperativo baseado em autovalores para rádios cognitivos,” *Revista Telecomunicações*, vol. 14, no. 1, pp. 1–9, 2012.

-
- [30] C. Cordeiro, K. Challapali *et al.*, “Spectrum agile radios: utilization and sensing architectures,” in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*. IEEE, 2005, pp. 160–169.
- [31] S. Ciftci and M. Torlak, “A comparison of energy detectability models for spectrum sensing,” in *IEEE GLOBECOM 2008-2008 IEEE Global Telecommunications Conference*. IEEE, 2008, pp. 1–5.
- [32] C.-Y. Chen, Y.-H. Chou, H.-C. Chao, and C.-H. Lo, “Toward secure centralized spectrum sensing by utilizing geographical information,” in *2010 5th International Conference on Future Information Technology*. IEEE, 2010, pp. 1–6.
- [33] W. Zhang and K. B. Letaief, “Cooperative spectrum sensing with transmit and relay diversity in cognitive radio networks-[transaction letters],” *IEEE Transactions on wireless communications*, vol. 7, no. 12, pp. 4761–4766, 2008.
- [34] S. Haykin and M. Moher, *Sistemas de Comunicação-5*. Bookman Editora, 2011.
- [35] K. B. Letaief and W. Zhang, “Cooperative communications for cognitive radio networks,” *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, 2009.
- [36] R. Tandra and A. Sahai, “Fundamental limits on detection in low snr under noise uncertainty,” in *2005 International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 1. IEEE, 2005, pp. 464–469.
- [37] B. Wang and K. R. Liu, “Advances in cognitive radio networks: A survey,” *IEEE Journal of selected topics in signal processing*, vol. 5, no. 1, pp. 5–23, 2010.
- [38] A. Kortun, T. Ratnarajah, M. Sellathurai, C. Zhong, and C. B. Papadias, “On the performance of eigenvalue-based cooperative spectrum sensing for cognitive radio,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 49–55, 2010.
- [39] S. Haykin and K. Elektroingenieur, *Neural networks and learning machines*. Pearson education Upper Saddle River, 2009, vol. 3.
- [40] B. Kröse and P. van der Smagt, “An introduction to neural networks,” 1993.

-
- [41] H.-X. Li and E. S. Lee, “Interpolation functions of feedforward neural networks,” *Computers & Mathematics with Applications*, vol. 46, no. 12, pp. 1861–1874, 2003.
- [42] G. Balázs, “Cascade-correlation neural networks: A survey.”
- [43] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *European conference on computer vision*. Springer, 2014, pp. 818–833.
- [44] M. Riedmiller and H. Braun, “A direct adaptive method for faster backpropagation learning: The rprop algorithm,” in *Proceedings of the IEEE international conference on neural networks*, vol. 1993. San Francisco, 1993, pp. 586–591.
- [45] H. Yu and B. M. Wilamowski, “Levenberg-marquardt training,” *Industrial electronics handbook*, vol. 5, no. 12, p. 1, 2011.
- [46] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.
- [47] L. Sachs, *Applied statistics: a handbook of techniques*. Springer Science & Business Media, 2012.
- [48] Z. Qin, W. Li, and X. Xiong, “Estimating wind speed probability distribution using kernel density method,” *Electric Power Systems Research*, vol. 81, no. 12, pp. 2139–2146, 2011.
- [49] J. Nolan, *Stable distributions: models for heavy-tailed data*. Birkhauser New York, 2003.
- [50] G. A. Anastassiou, “Multivariate sigmoidal neural network approximation,” *Neural Networks*, vol. 24, no. 4, pp. 378–386, 2011.
- [51] B. Llanas and F. J. Sainz, “Constructive approximate interpolation by neural networks,” *Journal of Computational and Applied Mathematics*, vol. 188, no. 2, pp. 283–308, 2006.
- [52] M. T. Hagan and M. B. Menhaj, “Training feedforward networks with the marquardt algorithm,” *IEEE transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, 1994.

- [53] W. Zhang, C.-X. Wang, X. Ge, and Y. Chen, “Enhanced 5g cognitive radio networks based on spectrum sharing and spectrum aggregation,” *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6304–6316, 2018.
- [54] X. Hong, J. Wang, C.-X. Wang, and J. Shi, “Cognitive radio in 5g: a perspective on energy-spectral efficiency trade-off,” *IEEE Communications Magazine*, vol. 52, no. 7, pp. 46–53, 2014.

Apêndice A

Como utilizar as funções `NnH0.m` e `NnH1.m`

`NnH0.m` e `NnH1.m` são funções do MATLAB. Para utilizá-las é preciso que ambas estejam salvas no mesmo diretório que a rotina que as utiliza. Essas funções não dependem da tollbox de machine learning do MATLAB para serem executadas.

Sintaxe:

`ParamH0 = NnH0(p m SNR n K);`

`ParamH0` será um vetor coluna de três posições contendo os valores de α , γ e δ , respectivamente. Apenas como lembrete β é sempre igual a 1

`ParamH1 = NnH1(p m SNR n K);`

`ParamH1` será um vetor coluna com três posições contendo os valores de k , σ e μ respectivamente.

