

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA ELÉTRICA
MESTRADO EM ENGENHARIA ELÉTRICA**

Thales Wulfert Cabral

**Localização de fontes harmônicas utilizando técnicas de processamento
estatístico de sinais**

Juiz de Fora

2020

Thales Wulfert Cabral

Localização de fontes harmônicas utilizando técnicas de processamento estatístico de sinais

Dissertação de Mestrado em Engenharia Elétrica apresentada na Universidade Federal de Juiz de Fora, na área de concentração de Sistema Eletrônicos, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. D.Sc. Leandro Rodrigues Manso Silva

Coorientador: Prof. D.Sc. Carlos Augusto Duque

Juiz de Fora

2020

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Cabral, Thales Wulfert.

Localização de fontes harmônicas utilizando técnicas de processamento estatístico de sinais / Thales Wulfert Cabral. -- 2020. 128 f.

Orientador: Leandro Rodrigues Manso Silva

Coorientador: Carlos Augusto Duque

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Engenharia Elétrica, 2020.

1. Análise de Componentes Independentes. 2. Carga não linear. 3. Fonte harmônica de corrente. 4. Localização de fontes. I. Silva, Leandro Rodrigues Manso, orient. II. Duque, Carlos Augusto, coorient. III. Título.

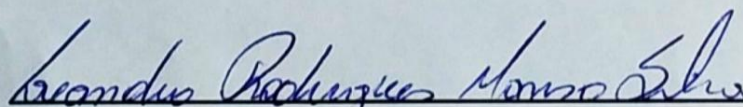
Thales Wulfert Cabral

Localização de Fontes Harmônicas utilizando Técnicas de Processamento Estatístico de Sinais

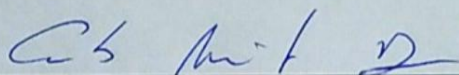
Dissertação de Mestrado em Engenharia Elétrica apresentada na Universidade Federal de Juiz de Fora, na área de concentração de Sistema Eletrônicos, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Aprovada em: 28 de Fevereiro de 2020

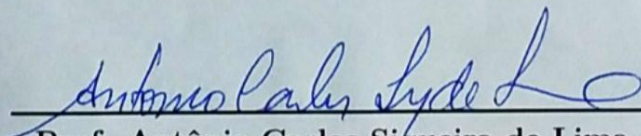
BANCA EXAMINADORA



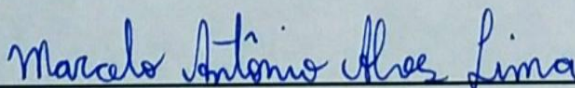
Prof. Leandro Rodrigues Manso Silva, D.Sc.
Universidade Federal de Juiz de Fora, UFJF



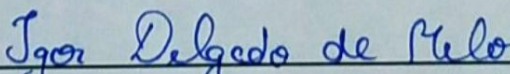
Prof. Carlos Augusto Duque, D.Sc.
Universidade Federal de Juiz de Fora, UFJF



Prof. Antônio Carlos Siqueira de Lima, D.Sc.
Universidade Federal do Rio de Janeiro, UFRJ



Prof. Marcelo Antônio Alves Lima, D.Sc.
Universidade Federal de Juiz de Fora, UFJF



Prof. Igor Delgado de Melo, D.Sc.
Universidade Federal de Juiz de Fora, UFJF

À minha família e amigos.

AGRADECIMENTOS

Agradeço a Deus, por sempre ter me proporcionado força, especialmente, nesta etapa.

À toda minha família, sobretudo ao meus pais, Sérgio e Wania, que sempre me apoiaram incondicionalmente, e também, à minha irmã e seu marido, Thalita e João, pela presença e incentivo.

Ao professor Leandro Manso, pela oportunidade de orientação, confiança em meu trabalho, paciência durante todo o processo, compreensão, e especialmente, amizade. Ao professor Marcelo Lima, por todos os ensinamentos, pela confiança depositada em mim, incentivos, exemplo e sobretudo um grande amigo. Ao professor Igor Delgado, por não poupar esforços em ajudar, pela boa convivência e pela amizade construída. Ao professor Carlos Duque pelas significativas contribuições, atenção e apoio.

Agradeço a banca, pelo aceite em avaliar o presente trabalho e contribuir com sugestões e comentários.

Ao CNPq, CAPES, FAPEMIG e à Universidade Federal de Juiz de Fora, pelo fomento à pesquisa.

Aos colegas do laboratório LAPTEL, pela convivência e amizade. Aos amigos formados em Juiz de Fora, durante todos os anos aqui decorridos.

Agradeço à oportunidade de aprendizado com participação do projeto de pesquisa “Estudo de Metodologias para a Conexão de Usinas Eólicas de Indústrias *Offshore* de Grande e Pequeno Porte - Aspectos Elétricos e Energéticos” em parceria com a Petrobrás.

Por fim, sou grato a todos os colaboradores da Universidade Federal de Juiz de Fora.

RESUMO

A pesquisa em qualidade de energia elétrica (QEE) é um campo da engenharia elétrica responsável por fornecer melhorias no fornecimento e consumo de energia. O termo QEE trabalha com a prevenção e mitigação de distúrbios elétricos. Cargas não lineares conectadas aos Sistemas de Energia Elétrica (SEPs) promovem mudanças indesejadas nos sinais elétricos e contribuem para a geração de distúrbios. Este trabalho apresenta uma metodologia para detectar fontes harmônicas de corrente como uma possível contribuição, identificando a localização das fontes harmônicas. Neste documento, o problema de estimação da fonte é tratado como um paradigma de separação cega de fontes, do inglês *Blind Source Separation* (BSS). Essa filosofia assume que as fontes não são conhecidas. Conseqüentemente, a medição de tensão fasorial é o único parâmetro acessível para monitoramento da rede. Além disso, a premissa inicial assume que a topologia da rede elétrica é desconhecida. Neste caso, situações com grandes volumes de dados devem ser consideradas. Assim, a Análise de Componentes Principais, do inglês *Principal Component Analysis* (PCA), é implementada como uma ferramenta para reduzir a dimensionalidade e auxiliar na complexidade computacional na etapa de estimação. A Análise de Componentes Independentes, do inglês *Independent Component Analysis* (ICA), é uma ferramenta de BSS e é responsável pela estimação dos sinais, neste caso, os perfis que caracterizam as fontes. Com os perfis estimados, a entropia conjunta é usada para determinar os prováveis locais onde as fontes harmônicas de corrente estão conectadas. Nesta etapa, o valor mínimo de entropia conjunta é usado como parâmetro de decisão para detecção. A eficácia da metodologia é verificada através do uso de simulações computacionais e sistemas de teste, tal como o IEEE 14 barras e o IEEE 33 barras.

Palavras-chave: Análise de Componentes Independentes. Carga não linear. Fonte harmônica de corrente. Localização de fontes.

ABSTRACT

Research in Electric Power Quality (QEE) is a field of electrical engineering responsible for providing improvements in energy supply and consumption. The term QEE works with the prevention and mitigation of electrical disturbances. Non-linear loads connected to the Electric Power Systems (SEPs) promote undesirable changes in electrical signals and contribute to the disturbance generation. This work presents a methodology for detection of harmonic current sources as a possible contribution, identifying the locations of the harmonic sources. In this document, the source estimation problem is treated as a Blind Source Separation (BSS) paradigm. This philosophy assumes that the sources are not known. Consequently, the voltage measurement is the only available parameter for network monitoring. Additionally, the initial premise assumes the power grid topology is unknown. In this case, situations with large volumes of data must be considered. Thus, Principal Component Analysis (PCA) is implemented as a tool to reduce dimensionality and aid in computational complexity for the estimation step. Independent Component Analysis (ICA) is a BSS tool and it is responsible for signals estimation, in this case, the profiles that characterize the sources. With the estimated profiles, the joint entropy is used to determine the probable locations where the harmonic current sources are connected. In this step, the minimum value of joint entropy is used as a decision parameter for detection. The methodology effectiveness is verified through the use of computer simulations and test systems, such as the IEEE 14 bus and the IEEE 33 bus.

Key-words: Independent Component Analysis. Non-linear load. Harmonic current source. Location of sources.

LISTA DE DIAGRAMAS

Diagrama 1 – Etapas essenciais à metodologia	48
Diagrama 2 – Processo de redução de dimensionalidade via PCA	52
Diagrama 3 – Diagrama da estimação de fontes harmônicas de corrente	55
Diagrama 4 – Sistema elétrico modelo	63
Diagrama 5 – Aferição das misturas	63
Diagrama 6 – Processo de localização de fontes harmônicas de corrente	64

LISTA DE FIGURAS

Figura 1 – Neurônio biológico simplificado	23
Figura 2 – Neurônio artificial	23
Figura 3 – Rede MLP	24
Figura 4 – Diferentes exemplos de margens e superfícies de separação de SVMs . .	25
Figura 5 – Problema de BSS	32
Figura 6 – Misturas e sinais originais	33
Figura 7 – Estimativas e sinais originais	34
Figura 8 – Redundância na distribuições de dados	41
Figura 9 – Níveis de incerteza	44
Figura 10 – Intervalo de confiança	54
Figura 11 – Correspondências via coeficientes de correlação	62
Figura 12 – Desempenho da entropia conjunta na localização das fontes harmônicas de corrente no sistema teste	65
Figura 13 – Perfil lento eólico	70
Figura 14 – Sistema IEEE 14 barras	73
Figura 15 – Perfis injetados para $ \rho = 37\%$	74
Figura 16 – Etapa de estimação para o caso de $ \rho = 37\%$	76
Figura 17 – Correspondência via $ \rho $ para o primeiro caso	76
Figura 18 – Resultados de localização dos harmônicos para o caso de $ \rho = 37\%$.	77
Figura 19 – Mapa para as localizações para o caso de $ \rho = 37\%$	77
Figura 20 – Perfis injetados para $ \rho = 75\%$	78
Figura 21 – Correspondência via $ \rho $ para o segundo caso	80
Figura 22 – Etapa de estimação para o caso de $ \rho = 75\%$	81
Figura 23 – Resultados de localização dos harmônicos para o caso de $ \rho = 75\%$.	81
Figura 24 – Mapa para as localizações para o caso de $ \rho = 75\%$	81
Figura 25 – Perfis injetados com $ \rho = 80\%$	82
Figura 26 – Correspondência via $ \rho $ para os perfis residenciais	85
Figura 27 – Resultado de estimação para s_2 e \hat{s}_2	85
Figura 28 – Resultados de localização dos harmônicos para o caso de $ \rho = 80\%$.	86
Figura 29 – Mapa para as localizações para o caso de $ \rho = 80\%$	86
Figura 30 – Perfis injetados para o caso de $ \rho = 98\%$	87
Figura 31 – Resultados de localização dos harmônicos para o caso de $ \rho = 98\%$.	90
Figura 32 – Mapa para as localizações para o caso de $ \rho = 98\%$	90
Figura 33 – Histogramas das variações rápidas para o caso 1	94
Figura 34 – Localização das fontes harmônicas de corrente para o intervalo = $[-0,0002; 0,0002]$	95

Figura 35 – Localização das fontes harmônicas de corrente para o intervalo = $[-0,002; 0,002]$	96
Figura 36 – Localização das fontes harmônicas de corrente para o intervalo = $[-0,02; 0,02]$	96
Figura 37 – Histogramas das variações rápidas para o caso 2	97
Figura 38 – Localização das fontes harmônicas de corrente para o intervalo = $[0; 0,0002]$	98
Figura 39 – Localização das fontes harmônicas de corrente para o intervalo = $[0; 0,002]$	98
Figura 40 – Localização das fontes harmônicas de corrente para o intervalo = $[0; 0,002]$	98
Figura 41 – Histogramas das variações rápidas para o caso 3	99
Figura 42 – Localização das fontes harmônicas de corrente para VAR = 0,0002	100
Figura 43 – Localização das fontes harmônicas de corrente para VAR = 0,002	100
Figura 44 – Localização das fontes harmônicas de corrente para VAR = 0,02	101
Figura 45 – Resultados de entropia conjunta para o caso 1	102
Figura 46 – Mapa de probabilidade para o caso 1	102
Figura 47 – Resultados de entropia conjunta para o caso 2	103
Figura 48 – Mapa de probabilidade para o caso 2	103
Figura 49 – Resultados de entropia conjunta para o caso 3	104
Figura 50 – Mapa de probabilidade para o caso 3	104
Figura 51 – Sistema IEEE 33 barras	105
Figura 52 – Perfis utilizados	106
Figura 53 – Correspondência via coeficiente de correlação para as três estimativas	108
Figura 54 – Resultados de localização das fontes harmônicas do sistema IEEE 33 barras	108
Figura 55 – Mapa para as localizações do teste com o IEEE 33 barras	109
Figura 56 – Mapeamento de uma variável aleatória.	124
Figura 57 – Representação do intervalo infinitesimal.	125

LISTA DE GRÁFICOS

Gráfico 1 – <i>Tracking</i> via filtro de Kalman	28
Gráfico 2 – Perfil lento residencial	68
Gráfico 3 – Perfil lento rural	68
Gráfico 4 – Perfil lento industrial	69
Gráfico 5 – Perfil lento eólico	69
Gráfico 6 – THD de tensão ao longo do dia para o caso de $ \rho = 37\%$	75
Gráfico 7 – THD de tensão no horário de pico para o caso de $ \rho = 37\%$	75
Gráfico 8 – THD de tensão ao longo do dia para o caso de $ \rho = 75\%$	79
Gráfico 9 – THD de tensão no horário de pico para o caso de $ \rho = 75\%$	79
Gráfico 10 – THD de tensão ao longo do dia para o caso de $ \rho = 80\%$	83
Gráfico 11 – THD de tensão no horário de pico para o caso de $ \rho = 80\%$	83
Gráfico 12 – Intervalo de confiança para o 5º harmônico	84
Gráfico 13 – Intervalo de confiança para o 7º harmônico	84
Gráfico 14 – Intervalo de confiança para o 11º harmônico	84
Gráfico 15 – THD de tensão ao longo do dia para o caso de $ \rho = 98\%$	88
Gráfico 16 – THD de tensão no horário de pico para o caso de $ \rho = 98\%$	88
Gráfico 17 – Correspondência via $ \rho $ para o 5º harmônico do quarto caso	89
Gráfico 18 – Correspondência via $ \rho $ para os harmônicos do quarto caso	89
Gráfico 19 – Limiar Estabelecido para o 5º Harmônico	91
Gráfico 20 – Limiar Estabelecido para o 7º Harmônico	91
Gráfico 21 – Limiar Estabelecido para o 11º Harmônico	92
Gráfico 22 – Correspondência via $ \rho $ para o 5º harmônico ($N = 42$)	92
Gráfico 23 – Correspondência via $ \rho $ para os harmônicos ($N = 42$)	93
Gráfico 24 – Correspondência via coeficiente de correlação ($N = 11$)	93
Gráfico 25 – THD de tensão no horário de pico (19 horas) para o caso 1	101
Gráfico 26 – THD de tensão no horário de pico (19 horas) para o caso 2	103
Gráfico 27 – THD de tensão no horário de pico (19 horas) para o caso 3	104
Gráfico 28 – THD de tensão ao longo do dia para o IEEE 33 barras.	107
Gráfico 29 – THD de tensão no horário de pico para o IEEE 33 barras	107

LISTA DE QUADROS

Quadro 1 – Classificação das cargas por finalidade	17
Quadro 2 – Perfis e espectros utilizados nas cargas para $ \rho = 37\%$	73
Quadro 3 – Perfis e espectros utilizados nas cargas para $ \rho = 75\%$	78
Quadro 4 – Perfis e espectros utilizados nas cargas para $ \rho = 80\%$	82
Quadro 5 – Perfis e espectros utilizados nas cargas para $ \rho = 98\%$	87
Quadro 6 – Modelo da variação rápida utilizada no caso 1	94
Quadro 7 – Modelo da variação rápida utilizada no caso 3	99
Quadro 8 – Perfis e espectros utilizados nas cargas do sistema IEEE 33 barras . .	105

LISTA DE TABELAS

Tabela 1 – Espectro harmônico das cargas não lineares	72
Tabela 2 – Espectro nas barras com carga não linear para $ \rho = 37\%$. . .	74
Tabela 3 – Correspondência via $ \rho $ para os três intervalos utilizados no caso 1	95
Tabela 4 – Correspondência via $ \rho $ para os três intervalos ($a = 0$)	97
Tabela 5 – Correspondência via $ \rho $ para os três intervalos de variância do caso 3	100
Tabela 6 – Coeficiente de Correlação entre os perfis com o sistema IEEE 33 bar- ras	106

LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
ANN	<i>Artificial Neural Network</i>
BD	<i>Big Data</i>
BSS	<i>Blind Source Separation</i>
CCN	<i>Cascade Correlation Network</i>
CP	Componente Principal
EOS	Estatística de Ordem Superior
GD	Geração Distribuída
HSE	Estimação de Estado Harmônico (do inglês, <i>Harmonic State Estimation</i>)
HSI	<i>Harmonic State Identification</i>
ICA	<i>Independent Component Analysis</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
ILH	Índice de Localização Harmônica
JADE	<i>Joint Approximate Diagonalization of Eigen-matrices</i>
KF	Filtro de Kalman (do inglês, <i>Kalman Filter</i>)
KLT	<i>Karhunen-Loève Transform</i>
LLSME	<i>Linear Least Mean Squares Estimator</i>
LQE	<i>Linear Quadratic Estimator</i>
MLP	<i>Multilayer Perceptron</i>
MSE	<i>Mean Squared Error</i>
ONS	Operador Nacional do Sistema Elétrico
PCA	<i>Principal Component Analysis</i>
PNN	<i>Probabilistic Neural Network</i>
QEE	Qualidade de Energia Elétrica
RBF	<i>Radial Basis Function</i>
RCT	Reator Controlado a Tiristor
RF	<i>Random Forest</i>
RNA	Rede Neural Artificial
SEP	Sistema Elétrico de potência
SOA	<i>Seeker Optimization Algorithm</i>
SOBI	<i>Second Order Blind Identification</i>
SVM	<i>Support Vector Machine</i>
THD	Distorção Harmônica Total (do inglês, <i>Total Harmonic Distortion</i>)
WLS	<i>Weighted Least-Squares</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	IDENTIFICAÇÃO DO PROBLEMA	16
1.2	MOTIVAÇÃO	18
1.3	OBJETIVOS	18
1.4	PRODUÇÕES BIBLIOGRÁFICAS	19
1.5	ESTRUTURA DO DOCUMENTO	19
2	FUNDAMENTOS TEÓRICOS	20
2.1	REVISÃO SOBRE MÉTODOS DE DETECÇÃO DE FONTES HARMÔNICAS	20
2.1.1	Métodos Baseados em Problemas de Otimização	20
2.1.2	Métodos Baseados em Inteligência Computacional	22
2.1.3	Métodos Baseados em Rastreamento	27
2.1.4	Métodos Baseados em Separação Cega	29
2.2	FUNDAMENTOS DA METODOLOGIA DE ESTIMAÇÃO	31
2.2.1	Separação Cega de Fontes	31
2.2.1.1	<i>Observando Misturas Desconhecidas</i>	32
2.2.1.2	<i>Separação de Fontes Baseada na Independência</i>	34
2.2.2	Análise de Componentes Independentes	35
2.2.2.1	<i>Modelo Trivial da Análise de Componentes Independentes</i>	35
2.2.2.2	<i>Modelo Generativo da Análise de Componentes Independentes</i>	36
2.2.2.3	<i>Restrições de Análise de Componentes Independentes</i>	37
2.2.2.4	<i>Ambiguidades da Análise de Componentes Independentes</i>	37
2.2.3	Análise de Componentes Principais	38
2.2.3.1	<i>Definição da Análise de Componentes Principais</i>	39
2.2.3.2	<i>Observações da Análise de Componentes Principais na Redução de Dimensionalidade</i>	40
2.2.4	Branqueamento	42
2.3	FUNDAMENTOS DA METODOLOGIA DE LOCALIZAÇÃO	43
2.3.1	A Origem da Entropia e sua Relação com Níveis de Incerteza	43
2.3.2	Entropia	44
2.3.2.1	<i>Entropia Conjunta</i>	45
2.3.2.2	<i>Entropia Condicional</i>	46
2.4	CONSIDERAÇÕES FINAIS DO CAPÍTULO	46
3	METODOLOGIA	48
3.1	MODELO DE ICA ADOTADO E ADEQUAÇÕES	48
3.2	REDUÇÃO DE DIMENSIONALIDADE	51
3.3	FILTRAGEM	52

3.4	ALGORITMOS DE BSS	56
3.4.1	Second Order Blind Identification (SOBI)	56
3.4.2	FastICA	57
3.4.3	RobustICA	59
3.4.4	Joint Approximate Diagonalization of Eigen-matrices (JADE) .	60
3.5	ESTIMAÇÃO DE FONTES HARMÔNICAS DE CORRENTE EM UM SISTEMA TESTE	61
3.6	LOCALIZAÇÃO DE FONTES HARMÔNICAS DE CORRENTE . . .	62
3.7	CONSIDERAÇÕES FINAIS DO CAPÍTULO	66
4	RESULTADOS	67
4.1	MODELOS UTILIZADOS NA BASE DE DADOS	67
4.1.1	Curvas de Carga e Perfis de Carga	67
4.1.2	Independência Estatística do Perfil de Carga	69
4.1.3	Injeção Harmônica	71
4.2	RESULTADOS PARA O SISTEMA 14 BARRAS	73
4.2.1	Perfis lentos com $ \rho = 37\%$ de correlação.	73
4.2.2	Perfis lentos com $ \rho = 75\%$ de correlação.	78
4.2.3	Perfis lentos com $ \rho = 80\%$ de correlação.	82
4.2.4	Perfis lentos com $ \rho = 98\%$ de correlação.	87
4.2.5	Janela Fora do Intervalo de Confiança	91
4.2.6	Alterando a Distribuição de Probabilidade da Variação Rápida.	93
4.2.7	Variando as Posições das Fontes	101
4.3	RESULTADOS PARA O SISTEMA IEEE 33 BARRAS	105
4.4	CONSIDERAÇÕES FINAIS DO CAPÍTULO	110
5	CONCLUSÕES	111
5.1	TRABALHOS FUTUROS	112
	REFERÊNCIAS	114
	APÊNDICE A – NOÇÕES DE PROBABILIDADE	122
	APÊNDICE B – FLUXO DE POTÊNCIA	127

1 INTRODUÇÃO

A pesquisa em Qualidade de Energia Elétrica (QEE) representa um ramo da engenharia elétrica incumbido de proporcionar melhorias tanto na forma como a energia é fornecida quanto no modo em que a mesma é consumida. Neste aspecto, o termo QEE está intrinsecamente relacionado ao fornecimento de energia quando é necessário lidar com a prevenção, ou a posteriori, mitigação de distúrbios elétricos.

O acrônimo QEE pode ser visto como um termo cuja finalidade reside em sintetizar conceitos capazes de descrever a qualidade das formas de onda de tensão e corrente de um Sistema Elétrico de Potência (SEP). Na literatura, existem diversas definições para a QEE. Para o Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) QEE é definida como sendo o conceito de fornecer e estabelecer a alimentação de um equipamento elétrico sensível de forma adequada ao seu funcionamento (MENDES *et al.*, 2015). Entretanto, o termo possui o sentido de qualidade de serviço efetuado pelas concessionárias, considerando a confiabilidade, fornecimento de informação e a qualidade da energia fornecida (ARRILLAGA; BOLLEN; WATSON, 2000). Porém, adotando um significado mais abrangente, o termo pode ser a “combinação entre a qualidade de tensão e a qualidade da corrente. Em que uma tensão ideal é uma tensão senoidal com amplitude e frequências constantes, onde os mesmos possuem valores nominais”(BOLLEN; GU, 2006).

O interesse em pesquisas relacionadas à QEE surge da necessidade de tornar os sistemas mais estáveis e, por conseguinte, mais confiáveis. Quando a QEE é insuficiente, inconveniências podem ocorrer: a instabilidade, diminuição da vida útil dos dispositivos elétricos e mau funcionamento dos mesmos. Entretanto, é desejável que o SEP ofereça aos consumidores finais tensões senoidais com frequência, amplitude e fase constantes.

1.1 IDENTIFICAÇÃO DO PROBLEMA

A inserção, na rede elétrica, de equipamentos oriundos da eletrônica de potência, obrigou a QEE a lidar com os desafios das cargas elétricas e seus distúrbios. Com a popularização desses dispositivos, diversas camadas de consumo contam com a presença dos mesmos:

- Nas residências, com o emprego de grandes números de fontes chaveadas, como televisores, aparelhos de som, DVDs, dentre outros. A presença de lâmpadas fluorescentes com reatores eletrônicos (ARSENEAU; OUELLETTE, 1993; JR *et al.*, 1997; SAXENA; VERMA; SINGH, 2010) e mais recentemente as lâmpadas LED com seus conversores;
- Nos sistemas de Geração Distribuída (GD), em que utiliza-se de inversores, que por sua vez geram problemas de QEE e também apresentam sensibilidade aos distúrbios

de QEE no SEP (SAXENA; VERMA; SINGH, 2010);

- No setor industrial, através da utilização de dispositivos de eletrônica de potência para o acionamento dos retificadores, das fontes chaveadas, dos motores, dos fornos a arco (GRAY; HAYDOCK, 1995; PILVELAIT; ORTMEYER; GRIZER, 1992; HOLTZ, 1994).

Devido a essa vasta capilaridade é possível classificar as cargas em três campos de aplicação, conforme o Quadro 1.

Quadro 1 – Classificação das cargas por finalidade

Industriais	Comerciais	Residenciais
Caldeiras	Computadores	Chuveiros
Fornos	Fornos	Lâmpadas
Conversores	Iluminação	Micro-ondas
Transformadores	Sistema de Refrigeração	Televisores

Fonte: Elaborado pelo autor (2020).

Contudo, essa classificação é uma divisão simplista e que apenas qualifica a carga por sua utilidade, sendo ineficaz em representar as características intrínsecas e essenciais das cargas, como por exemplo, os elementos elétricos que as compõem e seus efeitos.

Com a finalidade de compreender melhor o comportamento das cargas, é necessário efetuar uma classificação mais técnica, capaz de descrever o comportamento elétrico. Dessa forma, semelhante ao que ocorre na física, fenômenos são descritos através de equações matemáticas as quais traduzem os significados dos eventos observados. Assim, as cargas podem conter em sua essência duas naturezas: a linearidade e a não linearidade.

Do ponto de vista matemático, cargas lineares são regidas por uma relação linear trivial e seu comportamento é bem definido, carregado de alta previsibilidade. Em contraponto, as cargas não lineares não possuem um comportamento tão trivial e previsível, seus modelos matemáticos são mais elaborados. As cargas não lineares exibem uma relação não linear entre a tensão em seus terminais e a corrente drenada ou injetada, dependendo do ângulo de fase.

Em relação à QEE, as cargas não lineares geram correntes harmônicas que, por sua vez, distorcem as tensões da rede, através da impedância por onde elas circulam. As cargas não lineares mais comuns são os retificadores, equipamentos eletrônicos ou inversores. Também os conversores eletrônicos que conectam as fontes alternativas de energia ao sistema (e.g., solar e eólica), produzem distorções harmônicas.

Para efeito de estudo em qualidade de energia, no que diz respeito a distorção harmônica, estas cargas não lineares e as fontes de geração distribuída podem ser modeladas como fontes de corrente harmônicas inseridas no SEP (LIMA, 2013).

Neste contexto, a presença de fontes harmônicas de corrente no SEP é uma questão delicada, pois esse distúrbio pode causar: danos à rede elétrica devido ao aumento de tensão e corrente via ressonância (MONTEIRO, 2018); aquecimento de componentes (ZEBARDAST; MOKHTARI, 2017) e deterioração do funcionamento de equipamentos (ALMEIDA, 2011);

1.2 MOTIVAÇÃO

Devido à grande popularização dessas interfaces não lineares e sua vasta utilização, distúrbios elétricos tornaram-se recorrentes nas redes elétricas. Como esses fenômenos são capazes de causar desequilíbrios no SEP e prejudicar a QEE, é necessário estudar e fomentar o desenvolvimento de metodologias que sejam capazes de detectar essas fontes de distúrbios em meio ao sistema interligado.

Existe uma gama de distúrbios elétricos e com as mais variadas características. Dentre os fenômenos de curta duração, é possível citar o afundamento, a elevação de tensão e a interrupção momentânea (KUSKO; THOMPSON, 2007). Ao passo que para os eventos de longa duração pode-se elencar a subtensão e sobretensão (KUSKO; THOMPSON, 2007), por exemplo. Entretanto, ainda existem alguns fenômenos que são capazes de distorcer, incisivamente, as formas de onda num SEP. De acordo com a classificação estabelecida pela forma de onda (KUSKO; THOMPSON, 2007), esses fenômenos podem ser: o ruído, cortes de tensão, inter-harmônicos e harmônicos. Esse último, o harmônico, foi escolhido para o estudo efetuado Neste trabalho devido a sua proeminente recorrência no SEP.

Logo, o presente trabalho possui em seu cerne a detecção de fontes harmônicas de corrente em SEP. Ou seja, se o distúrbio está excursionando pelo sistema, o objetivo é localizar as fontes responsáveis pelo mesmo.

1.3 OBJETIVOS

Os objetivos deste trabalho estão relacionados ao contexto de QEE atrelado a técnicas de processamento de sinais para detecção e localização das fontes de corrente harmônicas do ponto de vista estático. O foco principal é, portanto, o desenvolvimento de algoritmos de processamento de sinais. Em termos de objetivos parciais pode-se elencar:

- Revisão bibliográfica das principais técnicas de detecção de fontes harmônicas de corrente, bem como, estimação de fontes harmônicas, localização de fontes harmônicas e, quando possível, ambas as etapas de forma concomitante;
- Apresentar um método de detecção de fontes harmônicas de corrente via técnicas de processamentos de sinais;
- Apresentar a localização das fontes harmônicas;
- Avaliar a acurácia das metodologias apresentadas.

1.4 PRODUÇÕES BIBLIOGRÁFICAS

Contribuições bibliográficas produzidas e publicadas durante o mestrado, cujas duas primeiras contribuições estão relacionadas com esta dissertação:

- CABRAL, T. W. *et al.* Estimaco e localizaco de fontes de correntes harmnicas em circuitos eltricos utilizando a anlise de componentes independentes. In: CBA. **Congresso Brasileiro de Automtica**. Joo Pessoa. 2018;
- CABRAL, T. W. *et al.* Deteco de fontes harmnicas de corrente em sistemas eltricos de potncia via anlise de componentes independentes. In: CBQEE. **Conferncia Brasileira sobre Qualidade da Energia Eltrica**. So Caetano do Sul. 2019;
- OLIVEIRA, D. R. *et al.* Comparaco de tcnicas de alta resoluo espectral para estimaco de componentes harmnicos e inter-harmnicos no-estacionrios. In: CBQEE. **Conferncia Brasileira sobre Qualidade da Energia Eltrica**. So Caetano do Sul. 2019;
- DIAS, F. M. *et al.* Extrao de sinal de ECG fetal comprimido utilizando sistema de inferncia Neuro-Fuzzy adaptativo. In: SBrT. **Simpsio Brasileiro de Telecomunicaes e Processamento de Sinais**. Petrpolis. 2019;

Contribuies bibliogrficas produzidas e submetidas durante o mestrado, cuja primeira contribuio esta relacionada com esta dissertaco:

- MESQUITA, A. S *et al.* Uma proposio de filtragem de dados para melhoria da estimaco complexa de fontes harmnicas em sistemas eltricos de potncia via anlise de componentes independentes. In: SBSE. **Simpsio Brasileiro de Sistemas Eltricos**. Santo Andr. 2020;
- CABRAL, T. W *et al.* Localizaco de fontes harmnicas em sistemas eltricos atravs de inteligncia computacional e estatstica de ordem superior. In: SBSE. **Simpsio Brasileiro de Sistemas Eltricos**. Santo Andr. 2020.

1.5 ESTRUTURA DO DOCUMENTO

No prximo captulo, Captulo 2, os principais trabalhos relacionados  temtica dessa dissertaco so elencados. Alm disso, a base terica  tratada para o entendimento do presente trabalho. No Captulo 3, a metodologia  discutida apresentando as etapas primordiais ao trabalho, estimaco e localizaco, bem como as tcnicas complementares a essas. No Captulo 4, a eficcia da metodologia  testada atravs de simulaes computacionais que utilizam os sistemas de transmisso e de distribuo. Finalmente, no Captulo 5, as concluses obtidas deste trabalho so apresentadas e propostas de trabalhos futuros so sugeridas.

2 FUNDAMENTOS TEÓRICOS

Um levantamento acerca da literatura que compreende pesquisas correlatas a este trabalho é apresentado neste capítulo, o qual será composto por duas seções majoritárias com suas respectivas motivações. A função primordial é elencar e promover o entendimento dos conceitos essenciais que compõem o presente trabalho.

Na primeira seção, os trabalhos que propõem métodos correlatos à pesquisa desenvolvida serão abordados. A motivação reside na demonstração do estado da arte de detecção de fontes harmônicas através dos trabalhos já consolidados com as mais diversas metodologias presentes na literatura.

Na segunda seção, um aprofundamento das técnicas utilizadas e combinadas Neste trabalho será apresentado, visando mostrar os conceitos essenciais das mesmas. Essa motivação é fundamentada em clarificar o modelo e ferramentas utilizadas, para que seja possível obter um fluxo linear e contínuo de raciocínio, quando a metodologia for apresentada.

2.1 REVISÃO SOBRE MÉTODOS DE DETECÇÃO DE FONTES HARMÔNICAS

Na revisão bibliográfica a seguir, os trabalhos de identificação e localização de fontes harmônicas foram organizados em quatro grupos: (i) métodos baseados em problemas de otimização; (ii) métodos baseados em inteligência computacional; (iii) métodos baseados em rastreamento; e (iv) métodos baseados em separação cega.

2.1.1 Métodos Baseados em Problemas de Otimização

Um termo altamente utilizado na detecção de fontes harmônicas é Estimção de Estados Harmônicos, do inglês *Harmonic State Estimation* (HSE). A técnica de HSE foi introduzida inicialmente em 1989 (HEYDT, 1989), com base na minimização da diferença quadrática ponderada entre os valores harmônicos medidos e os estimados. A partir disso, muitos trabalhos foram desenvolvidos e adotaram HSE como acrônimo preferencial quando se trata de detecção de fontes harmônicas. Adicionalmente, o termo *Harmonic State Identification* (HSI) também pode ser encontrado na literatura, no âmbito da detecção de fontes harmônicas.

Os métodos baseados em HSE necessitam do conhecimento total dos parâmetros do sistema para diferentes frequências harmônicas (DU *et al.*, 1999). Além disso, o HSE necessita de muitas medições harmônicas de tensão, potências ativas e reativas. Em adição, técnicas de alocação de medidores são utilizadas para otimizar as técnicas de HSE e, em geral, precisam de algum conhecimento prévio sobre a existência das fontes harmônicas e seus locais.

O modelo linear de medição para HSE pode ser considerado como:

$$\mathbf{z} = \mathbf{H}\mathbf{v} + \mathbf{e}, \quad (2.1)$$

onde: \mathbf{z} é o vetor de medição harmônica, \mathbf{H} é a matriz das funções que relacionam medições e estados harmônicos, \mathbf{v} é o vetor de estados harmônicos e \mathbf{e} é o vetor de erro.

O conceito de estimação de estados refere-se à incumbência de estimar os estados do sistema (GURSOY; NIEBUR, 2008), ou seja, as tensões harmônicas do barramento via equações super determinadas do SEP. Em posse da estimação de estados, as cargas harmônicas podem ser determinadas, visto que, o fluxo de carga harmônica pode ser estimado. Tradicionalmente, o problema de HSE é tratado como um problema de mínimos quadrados ponderados, do inglês *Weighted Least-Squares* (WLS), cuja função objetivo a ser minimizada pode ser representada conforme (2.2):

$$J(\mathbf{v}) = [\mathbf{z} - \mathbf{H}\mathbf{v}]^T \mathbf{K} [\mathbf{z} - \mathbf{H}\mathbf{v}], \quad (2.2)$$

em que: \mathbf{K} é a matriz diagonal de pesos, em que, seus elementos representam a confiabilidade e precisão das medidas. Para a solução do problema de otimização definido em (2.2), alguns importantes pontos devem ser considerados:

1. A matriz \mathbf{H} que representa a topologia do sistema é fornecida, as impedâncias de ramificação são conhecidas com precisão;
2. Séries temporais de medição não são necessárias, apenas um instante no tempo é necessário para a estimação;
3. Suposições acerca das propriedades estatísticas sobre a carga não são feitas;
4. As tensões harmônicas dos barramentos são estimadas.

Em 2004, um método baseado em otimização foi proposto para determinação precisa de múltiplas fontes harmônicas em sistemas de potência (KUMAR; DAS; SHARMA, 2004). A metodologia consiste em dois estágios: um estágio inicial para determinar os possíveis barramentos candidatos que portam as fontes harmônicas, via WLS; e um segundo estágio que utiliza a norma euclidiana que serve como um refinamento determinando, de fato, o barramento cuja fonte harmônica está conectada. A relação básica para a localização das fontes é dada por (2.3):

$$[I_h] = [Y_h][V_h], \quad (2.3)$$

em que $[I_h]$ é o vetor de corrente de injeção no barramento, $[Y_h]$ a matriz de admitância e $[V_h]$ o vetor de tensão do barramento, todos de ordem harmônica h . A relação (2.3) faz a ponte entre as correntes injetadas nos barramentos e as tensões via a matriz de admitância. Os autores consideram que $I_h \neq 0$ para os locais onde as cargas não lineares

não estão conectadas. Neste trabalho, preocupa-se em obter uma estimativa bem acurada para I_h . Caso isso ocorra, os barramentos cujas fontes harmônicas estão presentes podem ser determinados precisamente através da norma L_2 . Felizmente, os autores possuíam os dados de Y_h devido à informações do sistema e também do conhecimento de V_h . Os autores validam seu método apresentando os resultados no sistema IEEE 30 barras. Os resultados verificam que: para o estágio inicial, as menores correntes indicam as conexões das fontes harmônicas aos barramentos reais, a menos de uma conexão a mais que o esperado. Assim, o segundo estágio é utilizado através da norma L_2 refinando a indicação das conexões das fontes harmônicas de corrente nos devidos barramentos esperados.

Em 2012, um método de HSE foi proposto e baseado em WLS (KETABI; SHEIBANI; NOSRATABADI, 2012). Neste caso, o HSE faz estimativas das parte reais e imaginárias de cada harmônica minimizando o erro entre valores medidos e estimados. Entretanto, a novidade é a proposta da utilização de um algoritmo de otimização de buscador, do inglês *Seeker Optimization Algorithm* (SOA), para alocação ótima de medidores de qualidade de energia para aprimoramento de HSE. A alocação é definida minimizando a soma ponderada do máximo dos erros relativos obtidos do HSE para as ordens harmônicas existentes. Para validação, quatro cargas foram conectadas individualmente nos barramentos 4, 5, 6 e 7, do sistema IEEE 14 barras. Os resultados apresentaram-se satisfatórios quanto à estimação de estado e na alocação ideal dos medidores.

Em 2019, uma metodologia de HSE propõe que os problemas de otimização sejam formulados individualmente para cada ordem harmônica, com o objetivo de estimar perfis diários de carga e componentes harmônicos de fontes predominantes (MELO *et al.*, 2019). O HSE é proposto para sistemas de distribuição de energia trifásica desequilibrada, através de um número limitado de medidores fasoriais e alocados estrategicamente. O método é validado num sistema de distribuição IEEE 69 barras. Os resultados mostram um desempenho robusto para a estimação de potência ativa e reativa, num perfil diário de carga. Além disso, as estimações de magnitude e fase para os harmônicos mostram-se altamente satisfatórias.

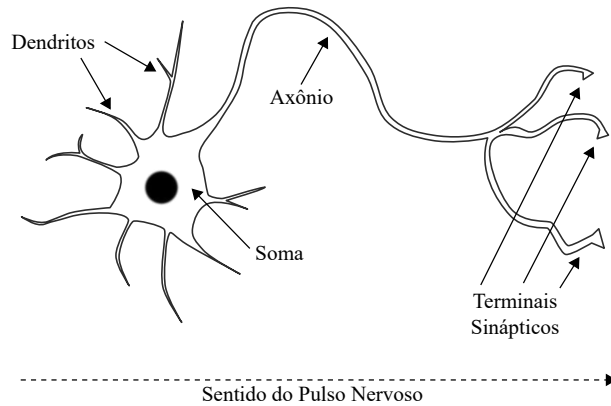
2.1.2 Métodos Baseados em Inteligência Computacional

Na década de 90 surgem as proposições baseadas em Redes Neurais Artificiais (RNAs) (HARTANA; RICHARDS, 1990; HARTANA; RICHARDS, 1993), do inglês *Artificial Neural Networks* (ANN), para o auxílio na identificação de fontes harmônicas nos SEPs.

A unidade primal de processamento de uma RNA é o neurônio artificial (LIMA, 2013), uma estrutura inspirada no neurônio biológico, cuja representação encontra-se na Figura 1. O neurônio biológico é composto essencialmente por soma, axônio e dendritos. A soma é o centro dos processos metabólicos. Os dendritos são prolongamentos encarregados da recepção de impulsos nervosos oriundos de neurônios vizinhos e o axônio é responsável

pela condução do impulso elétrico entre as células (FONTANA; MARIM, 2009), os terminais sinápticos conectam o axônio aos dendritos do neurônio vizinho.

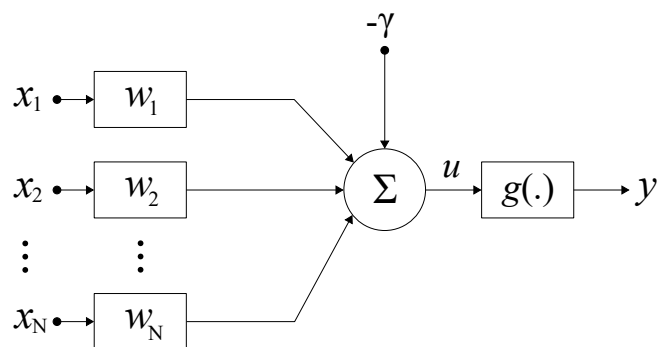
Figura 1 – Neurônio biológico simplificado



Fonte: Elaborado pelo autor (2020).

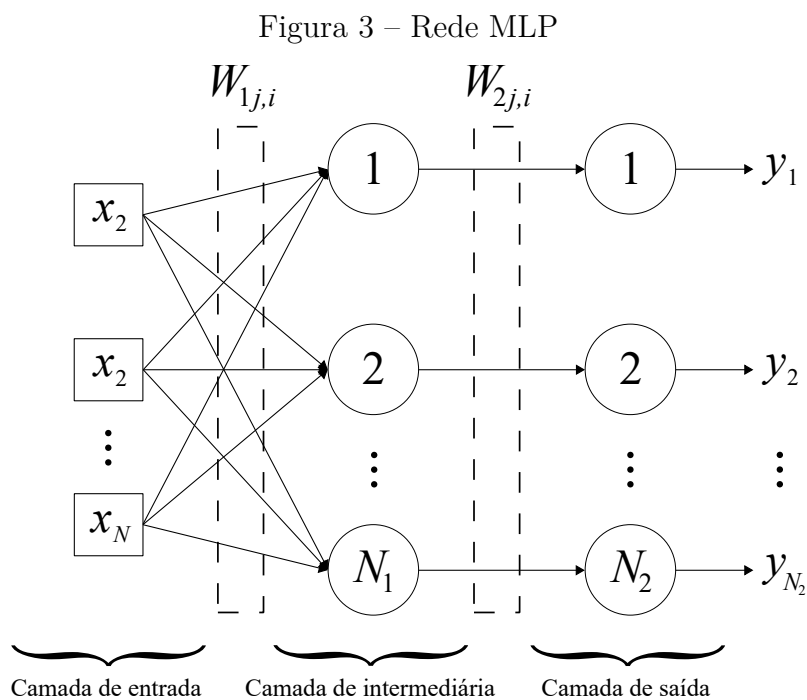
Inspirado no neurônio biológico, surge o modelo matemático, representado pela Figura 2. Um neurônio artificial calcula a soma ponderada de várias entradas, aplica uma função e gera um resultado. Conforme a Figura 2, as variáveis x_1, x_2, \dots, x_N são sinais de entrada equivalentes aos pulsos elétricos captados pelos dendritos, os pesos w_1, w_2, \dots, w_N são análogos às conexões sinápticas, tais pesos quantificam cada entrada de cada neurônio artificial, γ é o limiar de ativação (bias) e que apropria o patamar o qual u será adequado para um valor de disparo, que se conecta à saída do neurônio. Geralmente a saída no neurônio matemático y é limitada em valores contidos no intervalo $[0, 1]$, ou $[-1, 1]$, cujo mesmo é determinado pela função de ativação $g(\cdot)$. As funções de ativação geralmente são as funções da família parcialmente diferenciável e totalmente diferenciável, exemplos da primeira família são funções degrau e rampa simétrica (SILVA; SPATTI; FLAUZINO, 2012). Já as funções gaussiana e linear são exemplos da segunda família (SILVA; SPATTI; FLAUZINO, 2012).

Figura 2 – Neurônio artificial



Fonte: Elaborado pelo autor (2020).

Um exemplo típico de RNA para reconhecimento estrutural de padrões é a rede MLP (SILVA, 2012), do inglês *Multilayer Perceptron*. A arquitetura de uma RNA pode ser diversificada contendo camada única ou sendo multicamada, conforme a Figura 3. Neste tipo de estrutura, o modelo contém: a camada de entrada, a camada intermediária contendo os neurônios ocultos ou simplesmente denominada de camada oculta, responsável por tornar a rede capaz de extrair estatísticas de ordem elevada, e finalmente a camada de saída, cuja finalidade é produzir os resultados finais. Neste caso, os pesos sinápticos são armazenados em matrizes, conforme $W_{1j,i}$ $W_{2j,i}$. As redes de camada única diferem-se da estrutura da rede MLP por apresentarem uma arquitetura de RNA com apenas as camadas de entrada e saída.



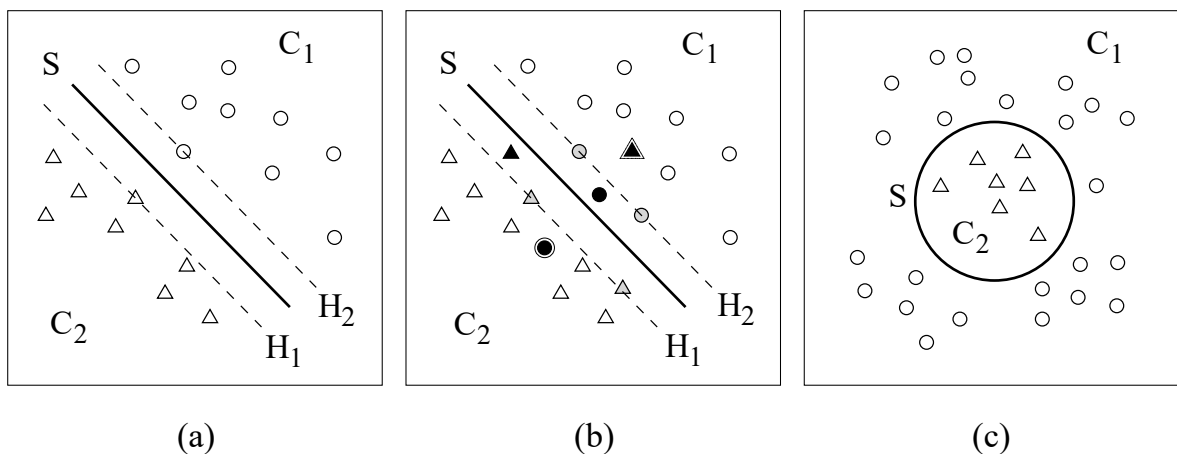
Fonte: Elaborado pelo autor (2020).

Em 2005, pesquisadores apresentaram uma rede de correlação em cascata, do inglês *Cascade Correlation Network* (CCN), para a identificação das fontes harmônicas. Esse tipo de RNA apresenta uma estrutura em torre (LIN *et al.*, 2005). Neste caso, aborda-se apenas o contexto de localização. Esse trabalho sugere algo interessante: é proposta uma abordagem de alocação de medidores com base na Distorção Harmônica de Total (DHT), do inglês *Total Harmonic Distortion* (THD), de tensão. Para essa alocação, inicialmente observa-se quais barramentos os valores da DHT são infringidos, definindo-se assim barramentos chave. Isso justifica fortemente a proposta de localização de fontes harmônicas, pois a DHT é um indicativo de que algo não está operando em condições normais de regime num sistema. Logo, a partir dessa anomalia pode-se pensar em soluções mitigatórias. Outro ponto importante é que como locais chave são definidos para a alocação dos medidores,

nem todos os barramentos possuem medidores sendo uma simplificação e vantagem de implementação. Os resultados apresentam altas taxas de acerto na localização. Entretanto, como trata-se de um método supervisionado, ainda necessita de treinamento.

Ainda em 2015, uma visão mais específica da detecção de fontes harmônicas é proposta (SRINIVASAN; NG; LIEW, 2005). O trabalho está interessado em distinguir as “assinaturas harmônicas” de cada dispositivo não linear conectado numa instalação elétrica. Ferramentas baseadas em RNAs como MLP, rede de função base radial, do inglês *Radial Basis Function* (RBF), e também, Máquina de Vetor de Suporte, do inglês *Support Vector Machine* (SVM), são utilizadas para a classificação das assinaturas. Neste contexto, as SVMs são alternativas de classificadores supervisionados às RNAs. As SVMs geralmente são utilizadas para classificar dados de duas classes (dicotomias). Porém, podem ser utilizadas para classificações m -dimensionais. Esse tipo de classificador pode assumir margens de decisão rígidas como a Figura 4 (a), suaves conforme a Figura 4 (b) e até mesmo superfícies não lineares de acordo com a Figura 4 (c). Quando as classes, por exemplo C_1 e C_2 , são linearmente separáveis opta-se por margens rígidas (Hiperplanos: H_1 e H_2 , por exemplo), mas em aplicações reais nem sempre essa condição é aplicável devido a uma série de condições como *outliers*, ruídos ou a natureza do problema. Neste caso, as margens (Hiperplanos: H_1 e H_2) do classificador linear são suavizadas. Assim, alguns dos dados podem permanecer na região entre os hiperplanos de separação, o que pode ocasionar erros de classificação. Já para problemas cuja separação não seja linear, a adequação a ser aplicada é a utilização de curvas de decisão, ou superfície de separação não linear, como a superfície S da Figura 4 (c).

Figura 4 – Diferentes exemplos de margens e superfícies de separação de SVMs



Fonte: Elaborado pelo autor (2020).

As classes, por exemplo num caso de localização de macro contexto, poderiam ser as posições que fontes harmônicas específicas estão conectadas em barramentos de grandes redes. Logo, pensando num caso de alta dimensionalidade (grande topologia),

é necessário utilizar um número de combinações muito grande para o treinamento. No caso do contexto reduzido, o objetivo reside em detectar as assinaturas harmônicas que já estão conectadas numa rede específica de topologia pequena (SRINIVASAN; NG; LIEW, 2005). Assim, deseja-se identificar através das assinaturas quais tipos de equipamentos estão presentes na instalação elétrica.

Semelhante ao trabalho citado anteriormente, surge uma proposição em 2010 que se incumbe a identificar cargas não lineares numa instalação residencial (FERNANDES; SILVA; OLESKOVICZ, 2010). Entretanto, o trabalho difere-se da referência anterior pois propõe-se a demonstrar que a informação de fase é irrelevante para a identificação de cargas não lineares. Altas taxas de acerto são obtidas para a diferenciação entre perfis que correspondem a lâmpada incandescente, fluorescente, fluorescente compacta, ventilador, computador pessoal e monitor. Essa utilização de classificadores para a identificação de cargas não lineares é um *approach* moderno e pode ser visto ainda em 2018 (KHALID *et al.*, 2018).

Ainda em 2018 e voltando para o macro contexto, redes com grandes topologias, uma metodologia baseada em floresta aleatória, do inglês *Random Forest* (RF), é proposta para a identificação de fontes de distúrbios de qualidade de energia (FENG *et al.*, 2018). Os autores utilizaram uma região local da China, não informada, que continha ferrovias de alta velocidade, energia eólica e energia fotovoltaica. Foi apresentado que a metodologia é capaz de separar as fontes de perturbações citadas com uma altíssima acurácia. Entretanto, utilizou-se 90% dos dados para treinamento. Esse tipo de ferramenta é altamente aplicável num contexto de análise de *Big Data* (BD). Porém, é preciso ter parcimônia na escolha da porcentagem de treinamento.

Em 2019, uma metodologia que mescla Rede Neural Probabilística, do inglês *Probabilistic Neural Network* (PNN), com lógica *Fuzzy*, foi proposta para a detecção de fontes harmônicas (MORADIFAR; FOROUD; FIROUZJAH, 2019). Os autores propõem um novo índice chamado de Índice de Localização Harmônica (ILH), baseado em lógica *Fuzzy*. Com a finalidade de determinar a localização e o nível relativo de fontes harmônicas, o índice considera parte real e imaginária das potências harmônicas, juntamente com a impedância da rede. Foi salientado que a DHT de tensão não é um índice confiável para a determinar a localização das cargas não lineares. Os resultados demonstram que a acurácia é muito elevada para a distinção entre os tipos de cargas. Quatro tipos foram utilizadas: um retificador de 6 pulsos, outro retificador de 12 pulsos, um motor CC e um Reator Controlado a Tiristor (RCT), do inglês *Thyristor Controlled Reactor* (TCR). Também, dois sistemas foram usados: o IEEE 18 barras e o IEEE 69 barras. O ILH demonstrou ser muito efetivo para localizar as fontes harmônicas, especialmente no sistema IEEE 69 barras.

2.1.3 Métodos Baseados em Rastreamento

Algumas aplicações cujo HSE baseia-se em rastreamento utilizam o conhecido filtro de Kalman, do inglês *Kalman Filter* (KF), que é responsável por estimar os estados de um processo. Como esta ferramenta possui natureza estocástica, pode ser implementada em situações de incerteza e ruído. O Filtro de Kalman também é conhecido por *Linear Least Mean Squares Estimator* (LLMSE), pois minimiza o erro de estimação do quadrado médio em sistemas estocásticos lineares via sensores lineares ruidosos (GREWAL; ANDREWS, 2014). Também pode ser chamado de *Linear Quadratic Estimator* (LQE), porque minimiza uma função quadrática de erro de estimativa em alguns tipos de sistemas lineares dinâmicos. Sua utilização consiste basicamente em etapas de predição e correção (BROWN, 1983), tais etapas serão representadas a seguir.

Quando se deseja obter o valor de uma variável aleatória dentro de um processo como (2.4):

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{w}_k, \quad (2.4)$$

em que: \mathbf{x}_k é o vetor de estados na amostra k , $\mathbf{\Phi}$ é a matriz que excursiona o processo do estado k para o estado $k + 1$ e \mathbf{w}_k é o ruído branco de covariância conhecida. As observações acerca da variável de interesse podem ser formuladas conforme (2.5):

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad (2.5)$$

onde: \mathbf{z}_k é a medida atual de \mathbf{x}_k na amostra k , \mathbf{H} é a matriz ideal sem ruído que conecta o vetor de estado e o vetor de medição e \mathbf{v}_k é o erro de medida. Para que a minimização do MSE gere um filtro ideal, é necessário modelar os erros do sistema como distribuições gaussianas. Assim, as covariâncias dos ruídos modelados são adotadas como estacionárias no tempo, conforme (2.6) e (2.7).

$$\mathbf{Q} = E[\mathbf{w}_k \mathbf{w}_k^T] \quad (2.6)$$

$$\mathbf{R} = E[\mathbf{v}_k \mathbf{v}_k^T] \quad (2.7)$$

Neste caso, o erro médio quadrático pode ser equivalente a (2.8).

$$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^T] = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] \quad (2.8)$$

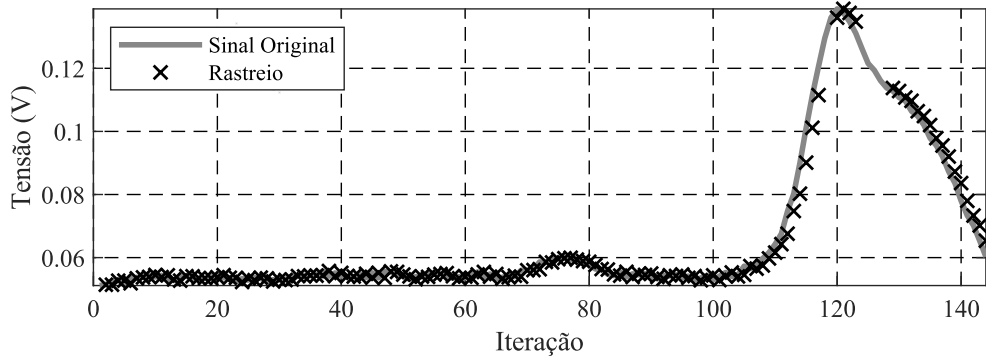
Para combinar a estimativa antiga com os dados de medição, a equação (2.9) é utilizada. Assim, o processo de atualização da estimativa é efetuado.

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-), \quad (2.9)$$

cuja $\hat{\mathbf{x}}_k^-$ é a estimativa anterior de $\hat{\mathbf{x}}_k$ e foi obtida com o conhecimento do sistema, \mathbf{K}_k é conhecido como o ganho de Kalman. Ainda é necessário atualizar a matriz de covariância

do erro e projetá-la, bem como a estimativa, em $k + 1$ para fechar o *loop* de Kalman. Devido a essa possibilidade de atualização de estimativas, equação (2.9), o Filtro de Kalman é muito utilizado para o rastreo de componentes harmônicos, abordagem conhecida como *tracking*, como no Gráfico 1. Neste caso, cada iteração busca encontrar a posição original de excursão do sinal. O resultado contido no Gráfico 1, apresenta um rastreo de uma tensão harmônica.

Gráfico 1 – *Tracking* via filtro de Kalman



Fonte: Elaborado pelo autor (2020).

Em 1991, o filtro de Kalman é usado como método de estimação da magnitude da tensão de fase de diferentes ordens harmônicas (BEIDES; HEYDT, 1991). Neste caso, a estimação possui um viés supracitado de rastreo de parâmetros harmônicos. Os autores validam seus resultados de estimação num sistema IEEE 14 barras modificado e utilizam um ciclo diário de potência real do sistema, denominado de *load cycle*. Vale salientar que o objetivo de localização de fontes harmônicas não é atacado no trabalho. Entretanto, bons resultados de estimação são encontrados para os espectros harmônicos de ordem 5, 7, 11 e 13.

Em 1996, o filtro de Kalman é implementado para determinar a estimação das injeções harmônicas de corrente nos barramentos e a alocação ótima de medidores (MA; GIRGIS, 1996). Neste trabalho, a proposta de alocação ótima de medidores é salientada, visto que os autores aproveitam a própria formulação do filtro de Kalman para isso. Essa alocação ótima ocorre quando o traço da matriz, ou soma dos elementos diagonais da matriz (THACKER; LACEY, 1998), de covariância do erro \mathbf{P}_k , é mínimo. Os resultados mostram uma dependência na qualidade da estimação dos perfis harmônicos em relação a alocação ótima.

Em 2015, pesquisadores implementam um filtro de Kalman adaptativo para a estimação das fontes harmônicas e rastreamento das mesmas (YU; WATSON; ARRILLAGA, 2005). A novidade salientada é que o sistema é modelado de forma a ser um modelo de estado que não depende do conhecimento preciso da matriz de covariância do ruído, necessário para as utilizações anteriores do filtro de Kalman.

Em 2016, um algoritmo alternativo ao filtro de Kalman é desenvolvido com o propósito de localizar fontes harmônicas em sistemas de distribuição (UJILE; DING, 2016). O método realiza estimações harmônicas para combinações de nós suspeitos, então, o erro da estimativa é utilizado para confirmar a existência das fontes. A proposição cobre a presença de fontes únicas e múltiplas fontes no sistema. Na presença de fonte única, os autores estabeleceram um limiar cujo *Mean Squared Error* (MSE) deve conter valores inferiores para que a combinação de nós que indique a existência da fonte harmônica. Os resultados comprovaram que a fonte está localizada num barramento específico ao mesmo tempo que todos os valores de MSE, computados em todas as medidas, são inferiores ao limiar. Para localizar múltiplas fontes, o sistema foi dividido em dois subsistemas. Para ambos subsistemas, as combinações de valores de MSE conseguiram identificar as fontes em seus respectivos barramentos de origem.

2.1.4 Métodos Baseados em Separação Cega

Métodos baseados na Separação Cega são muito úteis devido ao fato de não necessitarem do conhecimento de muitos parâmetros da rede, visto que, a metodologia convencional de estado harmônico necessita de um número redundante de medições harmônicas.

Em 2008, a técnica da Análise de Componentes Independentes, do inglês *Independent Component Analysis* (ICA), é utilizada para estimar as fontes harmônicas (GURSOY; NIEBUR, 2008). Neste caso, a matriz de medição é desconhecida, logo, o conhecimento da topologia do sistema é irrelevante para o método. Se \mathbf{H} for conhecido e medições suficientes estiverem disponíveis, não é necessário identificar a matriz de medição (ICA), pois as correntes harmônicas podem ser determinadas a partir da estimativa das tensões harmônicas. A metodologia pode ser resumizada conforme:

1. Se a matriz de mistura não for quadrada, aplica-se a Análise de Componentes Principais, do inglês *Principal Component Analysis* (PCA), para redução de medidas;
2. Aplica-se uma filtragem linear no vetor de medidas;
3. Efetua-se a centralização e branqueamento dos dados;
4. Utiliza-se o algoritmo FastICA complexo nos componentes de variação rápida (variação horária do perfil);
5. Obtém-se as estimações de correntes para a frequência harmônica desejada;
6. Para as frequências de interesse repete-se do passo II) ao V);
7. Efetua-se a reordenação e recuperação da escala via série histórica. O trabalho apresenta resultados de estimação satisfatórios para o sistema IEEE 30 barras.

Em 2010, um método baseado em ICA e informação mútua é desenvolvido com o propósito de determinar a localização de múltiplas fontes num sistema de distribuição (FARHOODNEA; MOHAMED; SHAREEF, 2010). A técnica da ICA é utilizada para estimar os perfis das correntes injetadas no sistema pelas fontes harmônicas e a informação mútua é responsável pela etapa de localização, indicando qual barramento a fonte provavelmente está conectada. A metodologia segue quatro procedimentos essenciais: I) Aferição das tensões em todos os barramentos; II) Reconstrução dos vestígios das correntes das fontes harmônicas via algoritmo FastICA; III) Cálculo da informação mútua entre os vestígios estimados e as tensões dos barramentos e IV) Afirmar a localização das fontes harmônicas através da informação mútua. As fontes harmônicas de corrente foram conectadas nos barramentos 15, 25 e 33 do sistema IEEE 34 barras radial. A fidelidade do perfil estimado em relação ao original é testada via coeficiente de correlação e MSE. Os maiores valores de informação mútua correspondem aos barramentos cujas fontes estão presentes. O artigo não explicita se as estimações são complexas, ou seja fasores, ou se apenas a magnitude está sendo estimada. Entretanto, o algoritmo FastICA não é um algoritmo tradicionalmente adaptado para estimações complexas, e quando a faz, os autores informam a devida adequação (KARIMZADEH; ESMAEILI; HOSSEINIAN, 2015).

Em 2011, a estimação de fontes harmônicas é efetuada via ICA complexa (PULIMERA; RAJAN, 2011), com modelos de misturas constantes em segmentos. A utilização da sobreposição de segmentos fornece informação para ordenar os sucessivos segmentos estimados das fontes, via sobreposição da última amostra de um segmento à primeira amostra do segmento sucessivo. No artigo, os autores mostram que o coeficiente de correlação entre as fontes estimadas no segmento sobreposto é alto. Esse alto valor entre segmentos sucessivos conduz à conclusão de que as estimativas são oriundas do mesmo tipo de fonte, ou seja, é possível recuperar o ordenamento. A determinação da escala também é resolvida através dos segmentos sucessivos. Por fim, os autores utilizam o sistema IEEE 14 barras para a validação da metodologia.

Em 2016, a identificação de fontes harmônicas é proposta via estimação e localização (ZANG *et al.*, 2016). A etapa de estimação é composta pela ICA, mais especificamente por um algoritmo denominado de *Fast Kernel Entropy Optimization Independent Component Analysis* (FKEO-ICA). As correntes harmônicas injetadas são estimadas pelo FKEO-ICA sem o conhecimento das impedâncias harmônicas, e depois disso, o mínimo valor de entropia condicional é utilizado para determinar o barramento cujas fontes harmônicas de corrente estão conectadas. Neste caso, não há indícios de estimação complexa das fontes harmônicas de corrente. Além disso, as cargas não lineares estão apenas nos barramentos no final dos alimentadores laterais. Em reais situações, as principais fontes harmônicas podem ser localizadas no meio ou distribuídas por todo o sistema.

2.2 FUNDAMENTOS DA METODOLOGIA DE ESTIMAÇÃO

Inicia-se a fundamentação através da apresentação dos conceitos básicos de Separação Cega de Fontes, ressaltando o modelo para esse tipo de problema. Após esse entendimento, a técnica da ICA é apresentada como solução para o problema anteriormente delineado. Por conseguinte, também são apresentados seus modelos, restrições e ambiguidades. Além disso, técnicas complementares que aprimoram a ICA são elencadas, por exemplo, a PCA, utilizada no presente trabalho.

2.2.1 Separação Cega de Fontes

O processo de extração de fontes subjacentes em conjuntos de observações obtidas via sensores é denominado separação de fontes (O'GRADY; PEARLMUTTER; RICKARD, 2005). Conforme a quantidade de informação acessível sobre o processo de mistura e sobre as fontes, o problema pode ser dividido em separação cega e separação semi-cega. Logo, a separação cega de fontes, do inglês *Blind Source Separation* (BSS), é um termo oriundo da separação de fontes. Ao levar em consideração o número de fontes (n) e sensores (m), o problema ainda pode ser classificado como: problema sub-determinado ($m < n$), problema determinado ($m = n$) e problema sobre-determinado ($m > n$).

Trivialmente, a mais simples descrição matemática do problema de BSS leva em conta uma mistura linear instantânea e livre de ruído. Assim, adota-se a existência de n fontes desconhecidas que podem ser representadas pelo vetor de fontes $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$, ao passo que, as m misturas observadas estão contidas no vetor $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$. Também, adota-se que o processo de mistura é desconhecido, linear e instantâneo. Tal processo é representado pela matriz de misturas \mathbf{A} , cujas dimensões devem ser $m \times n$. Logo, o problema de BSS pode ser equacionado conforme (2.10). E do ponto de vista dos coeficientes constituintes da matriz de misturas, o mesmo problema é descrito pela equação (2.11).

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (2.10)$$

$$\mathbf{x}(t) = \sum_{i=1}^n a_{ij}s_i(t) \quad (2.11)$$

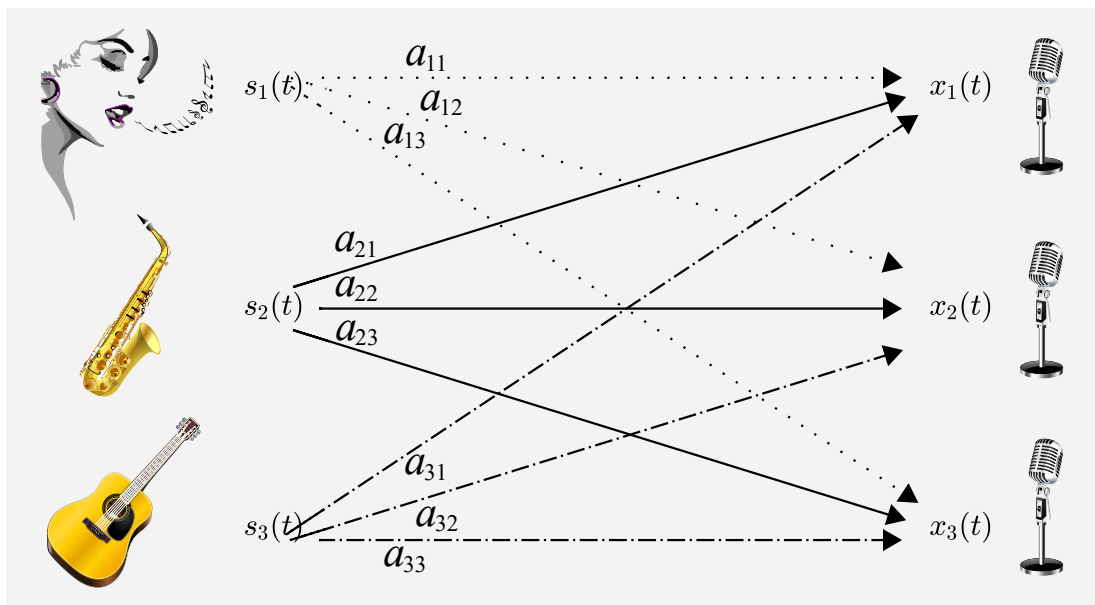
O objetivo da BSS é: recuperar $\mathbf{s}(t)$ unicamente através das misturas observáveis $\mathbf{x}(t)$. Neste contexto, a palavra cega se refere ao fato de que não é possível ter conhecimento de como os sinais de interesse foram gerados e como foram misturados. Contudo, mesmo com a dificuldade de identificação, atendendo as restrições necessárias, a BSS pode apresentar resultados satisfatórios para determinadas aplicações, como por exemplo: remoção de imperfeições em sinais de encefalografia (JUNG *et al.*, 2000), sinais de voz (MURATA; IKEDA; ZIEHE, 2001), monitoramento cardíaco (POH; MCDUFF; PICARD, 2010), ressonância magnética (WU *et al.*, 2017), dentre outros. Além disso, do início de 2019 até

julho do mesmo ano, 1.741 artigos que contém a frase “*blind source separation*” foram publicados na base de dados da *Science Direct* da editora *Elsevier*.

2.2.1.1 Observando Misturas Desconhecidas

Na literatura, o problema do *cocktail party* (CHERRY, 1953; YOST, 1997; AL-SAEGH, 2019) talvez seja o exemplo mais utilizado para elucidar o problema de BSS. Considere três pessoas distintas (fontes) vocalizando ao mesmo tempo, entre si, num mesmo espaço físico. Ao mesmo tempo, existem três microfones (medidores) dispostos em locais distintos recebendo esses sinais sonoros. Obviamente, cada microfone irá aferir misturas desses sinais sonoros vocalizados por essas pessoas. Adaptando esse problema para uma situação abrangente onde existem não apenas pessoas, mas também instrumentos musicais por exemplo, é possível ilustrar o problema de BSS através da Figura 5.

Figura 5 – Problema de BSS



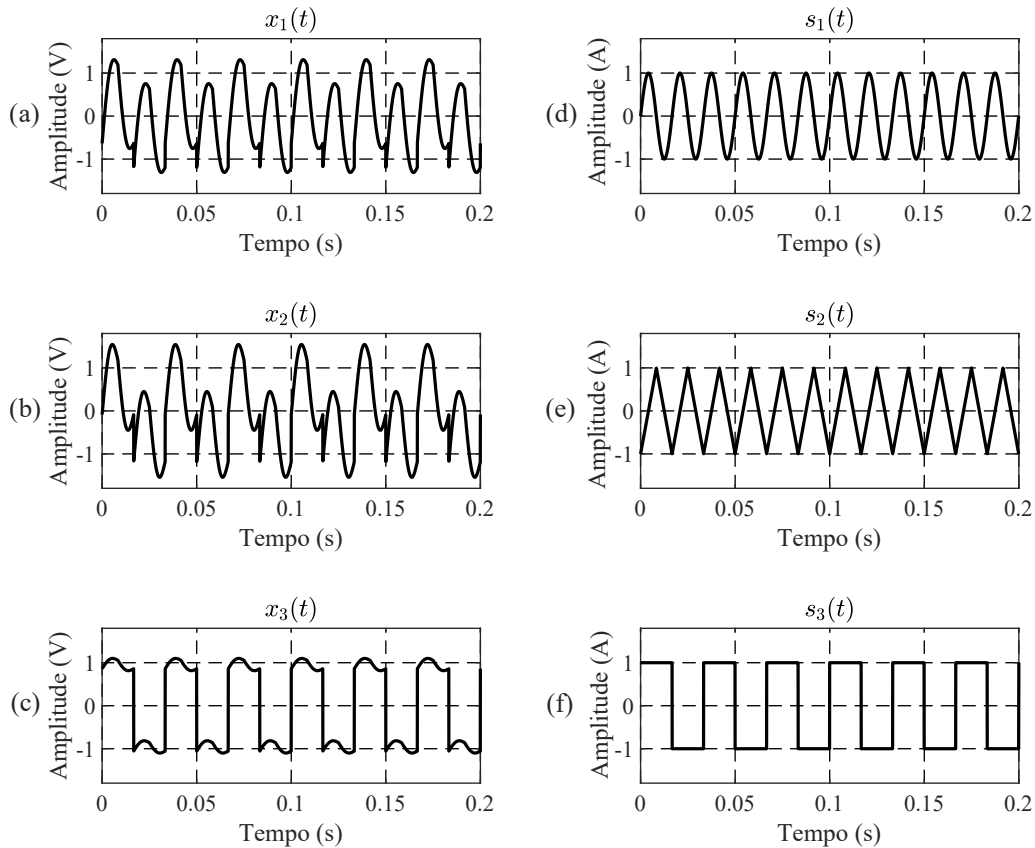
Fonte: Elaborado pelo autor (2020).

Seguindo o raciocínio supracitado, é possível imaginar um modelo matemático trivial que contém três fontes e misturas. Neste caso, considera-se $x_1(t)$, $x_2(t)$ e $x_3(t)$ os sinais observados e os sinais originais considerados como $s_1(t)$, $s_2(t)$ e $s_3(t)$. Assim, cria-se um sistema cujos coeficientes são dependentes da distância existente entre as fontes e os sensores, podendo ser representado pela equação (2.12) e pela Figura 6.

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix} \quad (2.12)$$

Neste caso, a Figura 6 (a) é a mistura $x_1(t)$, a Figura 6 (b) é a mistura $x_2(t)$, a Figura 6 (c) é a mistura $x_3(t)$, a Figura 6 (d) é o sinal senoidal $s_1(t)$ de 60 Hz, a Figura 6 (e) é o sinal triangular $s_2(t)$ de 60 Hz e a Figura 6 (f) é a onda quadrada $s_3(t)$ de 30 Hz.

Figura 6 – Misturas e sinais originais



Fonte: Elaborado pelo autor (2020).

Os coeficientes a_{ij} contidos na equação (2.12), são constantes que caracterizam as ponderações das misturas. Esses coeficientes são assumidos não conhecidos, pois não é possível determinar seus valores sem conhecer totalmente as propriedades do sistema de mistura. Conhecer essas propriedades em sua totalidade seria algo extremamente complexo. Outro fato é que os sinais de origem também são desconhecidos, posto que não são observáveis de maneira direta, assim, são denominados “variáveis latentes”.

Felizmente, pode-se considerar que os coeficientes a_{ij} da mistura são constituídos de tal forma que \mathbf{A} seja não singular, ou seja, ser inversível. Além disso, existe a matriz \mathbf{W} composta pelos coeficientes w_{ij} que é capaz de “extraír” as fontes s_i . Essa alternativa inicial está representada pela equação (2.13).

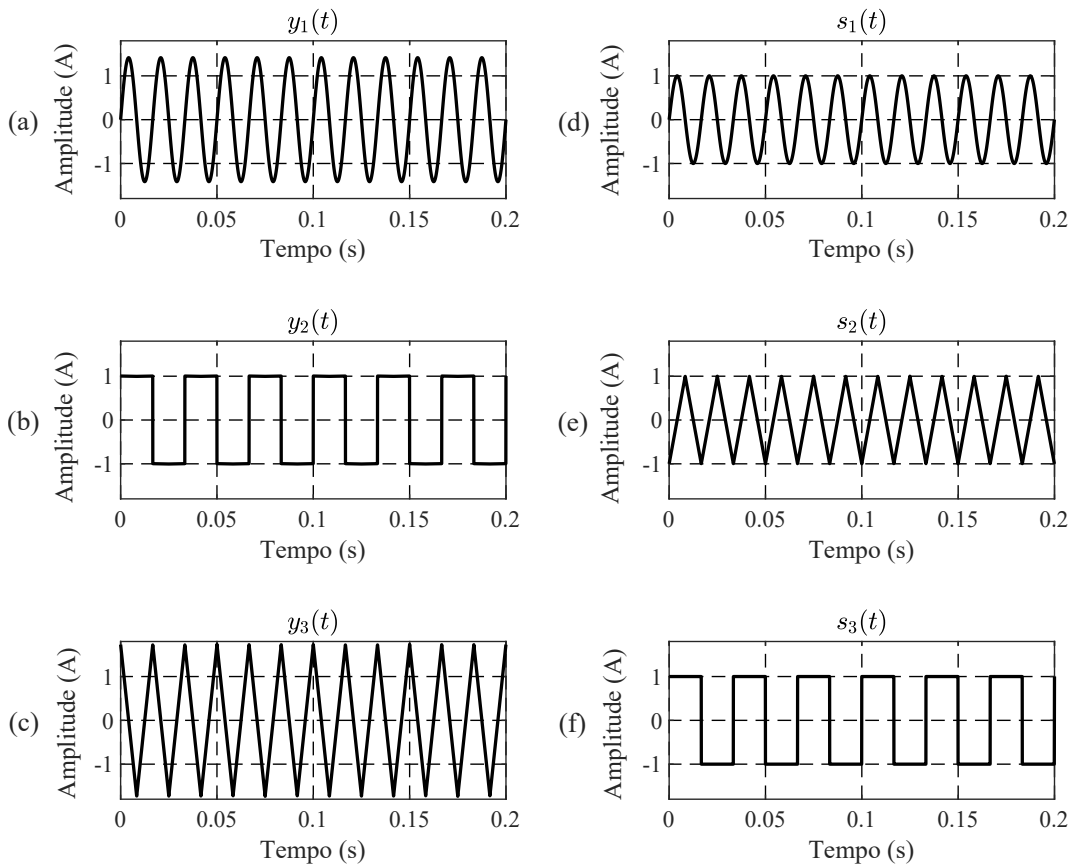
$$\begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \\ \hat{s}_3(t) \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad (2.13)$$

2.2.1.2 Separação de Fontes Baseada na Independência

Neste caso, deseja-se estimar os coeficientes w_{ij} . Como o observador possui acesso apenas à misturas, uma possível solução é encontrada considerando apenas a independência estatística das fontes. Basta que os sinais sejam estatisticamente independentes, não gaussianos, e além disso, pelo menos n misturas ($n \leq m$) têm que ser linearmente independentes. Se os sinais $y_1(t)$, $y_2(t)$ e $y_3(t)$ cumprirem esses requisitos, então eles serão semelhantes, em perfil, aos sinais originais $s_1(t)$, $s_2(t)$ e $s_3(t)$, conforme mostra a Figura 7. Neste caso, a Figura 7 (a) é Estimativa $y_1(t)$ do sinal senoidal, a Figura 7 (b) é estimativa $y_2(t)$ da onda quadrada, a Figura 7 (c) é estimativa $y_3(t)$ do sinal triangular, a Figura 7 (d) é o sinal senoidal $s_1(t)$ de 60 Hz, a Figura 7 (e) é o sinal triangular $s_2(t)$ de 60 Hz e a Figura 7 (f) é a onda quadrada $s_3(t)$ de 30 Hz. Logo, usando a informação de independência estatística, pode-se estimar a matriz de coeficientes w_{ij} , equação (2.14), com base nos sinais misturados (HYVÄRINEN; KARHUNEN; OJA, 2001).

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad (2.14)$$

Figura 7 – Estimativas e sinais originais



Fonte: Elaborado pelo autor (2020).

Ao comparar a Figura 7 (a) e a Figura 7 (d), nota-se que a escala em amplitude do sinal não é mantida. Além disso, quando se compara a Figura 7 (c) e a Figura 7 (e), percebe-se uma inversão de fase. Finalmente, fica perceptível que a ordem das fontes não é mantida para as estimativas, visto que, a Figura 7 (e) é a segunda fonte cuja forma é triangular, entretanto, a segunda estimativa, Figura 7 (b), é uma onda quadrada. Essas são as características das ambiguidades e que serão discutidas mais adiante.

2.2.2 Análise de Componentes Independentes

Herauld e Jutten foram os responsáveis, inicialmente, pela formulação da Análise de Componentes Independentes (HERAULT; JUTTEN, 1986). Os autores buscavam resolver o problema de BSS adotando uma mistura linear, instantânea e um problema determinado. A possibilidade do problema BSS ser solucionado via independência era investigada. Posteriormente, utilizou-se uma regra similar a Hebbian (MORRIS, 1999) que utilizava funções não-lineares, cuja finalidade era testar a independência entre pares de produtos. E então, quando a independência fosse encontrada um produto desaparecia. Entretanto, somente em 1991 uma base teórica sólida fora desenvolvida e o termo “Análise de Componentes Independentes”, do inglês *Independent Component Analysis* (ICA), surgia pela primeira vez (COMON; JUTTEN; HERAULT, 1991; JUTTEN; HERAULT, 1991).

Após a consolidação do modelo da ICA, a técnica tornou-se uma das preferidas para possível solução dos problemas de BSS. A técnica é um método amplamente aplicado (HYVÄRINEN; KARHUNEN; OJA, 2001), visto que, em suma, é uma técnica não paramétrica para estimar sinais individuais a partir de misturas, cujo grande trunfo consiste na consideração de que diferentes processos físicos geram sinais não-correlacionados.

Neste aspecto, a ICA ganhou tanto destaque e importância que diversas áreas do conhecimento também optaram por sua utilização. A técnica é empregada em identificação de anomalias em registros eletroencefalográficos e magnetoencefalográficos (VIGÁRIO *et al.*, 2000), reconhecimento facial (BARTLETT; MOVELLAN; SEJNOWSKI, 2002), eletromiogramas (DAI; HU, 2019), monitoramento do tráfego aéreo (JIANG *et al.*, 2019), detecção de distúrbios elétricos (RAY; MOHANTY; PANIGRAHI, 2019), dentre outros. Neste contexto, do início de 2019 até julho do mesmo ano, 78.749 artigos que contém a frase “*independent component analysis*” foram publicados na base de dados da *Science Direct* da editora *Elsevier*.

2.2.2.1 Modelo Trivial da Análise de Componentes Independentes

Para definir ICA é preciso usar o modelo de “variáveis latentes” (HYVÄRINEN; KARHUNEN; OJA, 2001). Isso significa que apenas o vetor aleatório de misturas $\mathbf{x} = [x_1, \dots, x_n]^T$ é observado. Além disso, o vetor aleatório de fontes $\mathbf{s} = [s_1, \dots, s_n]^T$ é constituído por elementos estatisticamente independentes entre si. Assim, adota-se uma modelagem em que os elementos do vetor de misturas \mathbf{x} sejam combinações lineares dos

elementos que constituem o vetor aleatório de fontes \mathbf{s} . Logo, para uma representação compacta, utiliza-se a famigerada equação (2.15):

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{s} \quad (2.15)$$

A matriz \mathbf{A} é a chamada matriz de mistura, cujas dimensões são $M \times N$. A técnica busca por uma transformação linear, equação (2.16), que minimiza a dependência estatística entre os elementos do vetor aleatório $\mathbf{y} = [y_1, \dots, y_n]^T$ (COMON, 2004; LIMA, 2013).

$$\mathbf{y} = \mathbf{W} \cdot \mathbf{x} \quad (2.16)$$

Assim, a ICA encontra a matriz \mathbf{W} (matriz de separação). Essa matriz é responsável por promover a determinação das estimativas das fontes originais, visto que, \mathbf{W} é a inversa correspondente de \mathbf{A} . Esses sinais recuperados são chamados de estimativas porque a ICA não preserva suas amplitudes originais, fases originais e a ordem original das fontes.

2.2.2.2 Modelo Gerativo da Análise de Componentes Independentes

O modelo gerativo é muito semelhante ao modelo trivial a menos de um detalhe, Neste caso, considera-se um ruído nas observações. Assim, as misturas são geradas num espaço de origem e excursionam até os sensores dependentes, cujas fontes originais continuam independentes. Sejam L fontes e M medidores, mantém-se a estrutura do vetor aleatório de fontes $\mathbf{s} = [s_1, \dots, s_L]^T$ e do vetor aleatório de misturas $\mathbf{x} = [x_1, \dots, x_M]^T$, cujo ruído pode ser adicionado pelos sensores antes de produzir as sequências M . Logo, o modelo gerativo pode ser descrito pelo mapeamento de L componentes independentes para M sensores, conforme (2.17):

$$\mathbf{x}(\mathbf{t}) = \mathbf{f}(\mathbf{s}(\mathbf{t})) + \mathbf{n}(\mathbf{t}), \quad (2.17)$$

onde \mathbf{f} é uma função de mapeamento determinístico e \mathbf{n} é um ruído aditivo. A ICA busca estimar o mapeamento, estatísticas de ruído e \mathbf{s} , dado \mathbf{x} . Neste caso, costuma-se adotar o mapeamento linear e instantâneo devido à simplicidade matemática. Então, de maneira semelhante ao modelo trivial, a notação compacta pode ser definida por (2.18).

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}, \quad (2.18)$$

em que, \mathbf{X} é uma matriz de dados $M \times T$, \mathbf{A} é a matriz de misturas $M \times L$, \mathbf{S} a matriz de fontes ou componentes independentes $L \times T$ e \mathbf{N} a matriz de ruído $M \times T$. Neste aspecto, a ICA pode ser vista como um modelo de densidade de dados, cuja densidade $p(\mathbf{x})$ é considerada uma transformação linear de rotação, de uma densidade de dados desconhecida $p(\mathbf{s})$. Assim, a ICA é usada para recuperar a densidade de origem, e Neste caso, a ICA modela a densidade de dados como uma densidade de origem fatorada.

2.2.2.3 Restrições de Análise de Componentes Independentes

Para que o modelo básico da ICA possa ser estimado, será necessário satisfazer alguns pressupostos e restrições (HYVÄRINEN; KARHUNEN; OJA, 2001).

1. Os componentes independentes são assumidos estatisticamente independentes:

Esse é o princípio pelo qual a técnica da ICA reside. Não é necessário muito mais do que esse pressuposto para verificar se o modelo pode ser estimado. Por esse fato, a ICA torna-se um método tão poderoso com aplicações em diversas áreas.

Variáveis aleatórias y_1, y_2, \dots, y_n são ditas independentes se a informação sobre o valor de y_i não diz informação sobre o valor de y_j para $i \neq j$;

2. Apenas uma componente independente pode apresentar distribuição gaussiana:

Pode-se dizer que as distribuições gaussianas são “muito simples”. Os cumulantes de ordem superior são zero para distribuições gaussianas, mas tais informações de ordem superior são essenciais para a estimativa do modelo ICA. Logo a ICA torna-se impraticável se as variáveis observadas tiverem distribuições gaussianas, podendo no máximo uma fonte possuir distribuição gaussiana;

3. A matriz de mistura \mathbf{A} deve ser quadrada:

Isso significa que o número de componentes independentes é igual ao número de misturas observadas. Genericamente, depois de estimar a matriz \mathbf{A} , pode-se calcular sua inversa, \mathbf{B} , e obter os componentes independentes por:

$$\mathbf{s} = \mathbf{B} \cdot \mathbf{x} \quad (2.19)$$

2.2.2.4 Ambiguidades da Análise de Componentes Independentes

No modelo básico da ICA, notam-se as seguintes ambiguidades ou indeterminações:

1. Não é possível determinar as variâncias (energias) das componentes independentes.

Pela razão de que, sendo \mathbf{s} e \mathbf{A} ambos desconhecidos, qualquer multiplicador escalar em uma das fontes s_i sempre pode ser cancelado dividindo a coluna correspondente a_i de \mathbf{A} pelo mesmo escalar. Portanto, é estabelecida a seguinte relação:

$$\mathbf{x} = \sum_i \left(\frac{1}{\alpha_i} \cdot a_i \right) \cdot (s_i \cdot \alpha_i) \quad (2.20)$$

Em que α_i é o multiplicador;

2. Não é possível determinar a ordem dos componentes independentes.

A razão é a de que, com s e \mathbf{A} desconhecidos, pode-se alterar livremente a ordem dos termos na soma da equação (2.15) e chamar qualquer um dos componentes independentes de primeiro;

3. É possível ocorrer inversão da fase.

Neste caso, a estimativa pode apresentar a forma de onda da fonte original, entretanto, com fase invertida.

2.2.3 Análise de Componentes Principais

Os primórdios da Análise de Componentes Principais surgiram em 1901, através de formulações matemáticas (PEARSON, 1901). Entretanto, somente em 1933, as propriedades estatísticas foram investigadas (HOTELLING, 1933). Pearson buscava encontrar projeções que se adequavam a um conjunto de pontos num espaço k -dimensional, enquanto Hotelling investigava um reduzido conjunto fundamental de variáveis independentes cuja função era a determinação dos valores das k -variáveis originais. A sigla comumente usada para a técnica é PCA, cuja origem advém do inglês *Principal Component Analysis* (PCA). Entretanto, a PCA também é conhecida como *Karhunen-Loève Transform* (KLT), ou ainda, por transformada de Hotelling (VRANIC; SAUPE; RICHTER, 2001).

Muito tempo após Pearson e Hotelling, a usabilidade da PCA expandiu-se em diversas áreas do conhecimento: botânica (MAIONE; JR; BARBOSA, 2019), sismologia (NISHITSUJI *et al.*, 2019), triagem de infecções (DITTA *et al.*, 2019), sustentabilidade ecológica (ONAT; KUCUKVAR; AFSHAR, 2019), dentre outros. A prova disso é que, do início de 2019 até julho do mesmo ano, 28.193 artigos que apresentam a frase “*principal component analysis*” foram publicados na base de dados da *Science Direct* da editora *Elsevier*.

Além disso, o emprego da PCA contém finalidades distintas conforme o viés da aplicação, por exemplo, em sistemas de classificação a PCA pode contribuir para a extração de parâmetros (SHAO *et al.*, 2014) e redução linear da dimensionalidade dos dados (KAMBHATLA; LEEN, 1997). Essa última pode resultar em diminuição do esforço computacional.

A PCA é largamente utilizada para a redução da dimensão dos dados devido a sua simplicidade de compreensão, implementação e por ser considerada uma técnica não supervisionada (MAO, 2005), o que de fato é algo extremamente simples e vantajoso, visto que, não depende de um banco de dados prévio de treinamento porque trabalha apenas com a matriz de recursos por amostras.

A técnica preserva a informação principal que descreve o conjunto de dados de interesse e utiliza a correlação para mensurar a redundância existente entre os componentes. A redundância detectada pode ser eliminada, e conseqüentemente, a eliminação reverbera

na diminuição da demanda computacional. Pois, quanto menor o volume de dados a ser analisado, menor o esforço. E o novo conjunto de dados após a PCA é composto pelos de Componentes Principais (CP).

2.2.3.1 Definição da Análise de Componentes Principais

Dado o vetor aleatório $\mathbf{x} = [x_1, \dots, x_n]^T$, a priori, o modelo generativo não é considerado, ou seja, é como se os elementos de \mathbf{x} fossem medidas, por exemplo, de valores de um sinal em instantes diferentes de tempo (HYVÄRINEN; KARHUNEN; OJA, 2001).

Então, realiza-se um processo de centralização. Tal processo utiliza o operador de esperança estatística (estimado) aplicado aos dados de interesse e subtrai esse resultado dos dados originais conforme a equação (2.21).

$$\mathbf{x} = \mathbf{x} - E\{\mathbf{x}\} \quad (2.21)$$

Após a centralização, os dados atuais devem apresentar $E\{\mathbf{x}\} = 0$. Então, \mathbf{x} é transformado linearmente em um novo vetor \mathbf{y} com m elementos, com dimensão reduzida ($m < n$), cuja redundância consequente da correlação é eliminada. Essa transformação linear promove um transporte das variáveis originais para um sistema de coordenadas cujos elementos de \mathbf{x} , nas atuais coordenadas, sejam descorrelacionados e as variâncias sejam maximizadas. Assim, o primeiro eixo carrega a variância máxima, o segundo eixo carrega a variância máxima na direção ortogonal ao primeiro eixo, e assim por diante até o k -ésimo componente (HYVÄRINEN; KARHUNEN; OJA, 2001).

Seguindo o viés da maximização da variância (HYVÄRINEN; KARHUNEN; OJA, 2001), adota-se a combinação linear descrita em (2.22) dos elementos $x_1 \dots x_n$ do vetor \mathbf{x} , onde os termos $w_{11} \dots w_{n1}$ são coeficientes escalares e elementos de \mathbf{w}_1 . O operador $(\cdot)^T$ indica transposição.

$$y_1 = \sum_{k=1}^n w_{k1} x_k = \mathbf{w}_1^T \mathbf{x}, \quad (2.22)$$

y_1 é denominado o primeiro PC de \mathbf{x} . Em posse disso, a PCA busca o vetor de coeficientes \mathbf{w}_1 , cuja finalidade seja maximizar a variância de y_1 , de acordo com (2.23).

$$J_1^{PCA}(\mathbf{w}_1) = E\{y_1^2\} = E\{(\mathbf{w}_1^T \mathbf{x})^2\} = \mathbf{w}_1^T E\{\mathbf{x}\mathbf{x}^T\} \mathbf{w}_1 = \mathbf{w}_1^T \mathbf{C}_x \mathbf{w}_1 \text{ e } \|\mathbf{w}_1\| = 1, \quad (2.23)$$

onde $\|\cdot\|$ é o operador de norma euclidiana definido conforme (2.24) e \mathbf{C}_x é a matriz de covariância definida por (2.25). Além disso, adota-se a restrição de que a norma de \mathbf{w}_1 seja constante, em geral, igual a 1 (LIMA, 2013).

$$\|\mathbf{w}_1\| = (\mathbf{w}_1^T \mathbf{w}_1)^{\frac{1}{2}} \quad (2.24)$$

$$\mathbf{C}_x = E\{\mathbf{x}\mathbf{x}^T\} \quad (2.25)$$

O resultado da PCA é dado de acordo com os autovetores $\mathbf{e}_1, \dots, \mathbf{e}_n$ da matriz \mathbf{C}_x . O ordenamento dos autovetores é feito conforme os autovalores $\lambda_1, \dots, \lambda_n$ satisfazendo $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Logo, a maximização de (2.23) é encontrada via (2.26).

$$\mathbf{w}_1 = \mathbf{e}_1 \quad (2.26)$$

Após isso, o primeiro CP de \mathbf{x} é representado por (2.27):

$$y_1 = \mathbf{e}_1^T \mathbf{x} \quad (2.27)$$

Generalizando para k componentes, conforme $\mathbf{w}_k = \mathbf{e}_k$, têm-se (2.28), o k -ésimo componente.

$$y_k = \mathbf{e}_k^T \mathbf{x} \quad (2.28)$$

Após o resultado, é possível escolher os CPs mais valiosos para a compressão de dados. As variâncias dos CPs correspondem aos autovalores de \mathbf{C}_x . Caso o CP possua um autovalor pequeno, como as médias dos CPs são nulas, o CP também será aproximadamente nulo. Assim, o conjunto original de dados pode ser aproximado por (2.29).

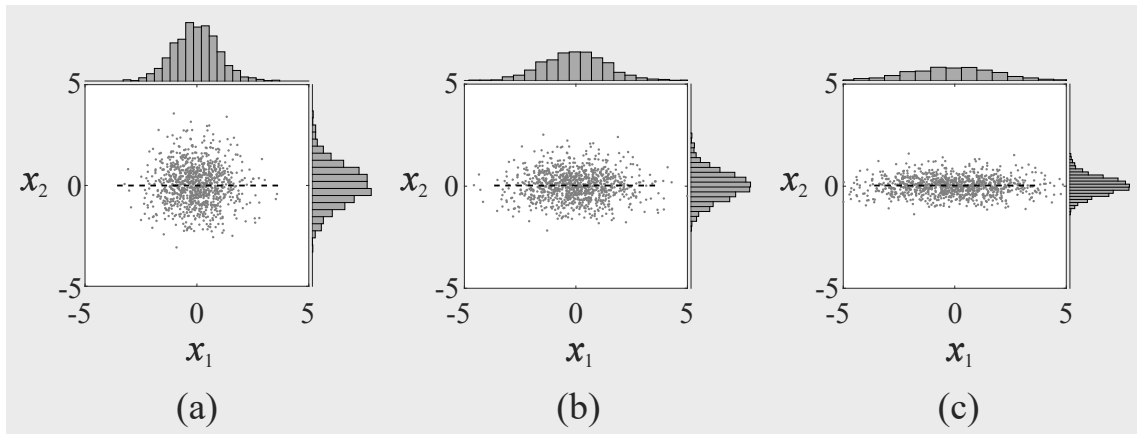
$$\hat{\mathbf{x}} = \sum_{i=1}^m y_i \mathbf{e}_i, \quad (m < n) \quad (2.29)$$

onde o conjunto de vetores de bases ortonormais reside num subespaço m -dimensional.

2.2.3.2 Observações da Análise de Componentes Principais na Redução de Dimensionalidade

Quando a finalidade da PCA é compressão, é necessário compreender a forma com que a redundância atua no conjunto de dados. A Figura 8 apresenta duas medidas (x_1, x_2) em três situações diferentes: (a) baixa redundância, o que implica em dados desconcorrelacionados e assim a partir de x_1 não se pode prever algo sobre x_2 e vice-versa; (b) redundância moderada entre x_1 e x_2 ; (c) dados com alta redundância, implicando em alta correlação. A alta correlação pode ser justificada se, em dados mensurados por sensores, por exemplo, houver proximidade entre ambos.

Figura 8 – Redundância na distribuições de dados



Fonte: Elaborado pelo autor (2020).

Observando a Figura 8 (c), percebe-se que não seria necessário armazenar as duas variáveis x_1 e x_2 . Utilizando uma linha de melhor ajuste é possível calcular x_1 a partir de x_2 , ou vice-versa. Ou seja, existe uma correlação entre x_1 e x_2 . Assim, no exemplo dos sensores, apenas um sensor poderia ser utilizado para a coleta dos dados. Esse conceito inicializa a ideia de compressão dimensional.

Formalizada a ideia de compressão, pode-se elencar algumas questões importantes que devem ser consideradas quando deseja-se aplicar a PCA:

1. Deve existir correlação:

Os elementos do conjunto de dados de interesse devem ser correlacionados, caso contrário, se forem independentes, não será possível efetuar a compressão dos mesmos (HYVÄRINEN; KARHUNEN; OJA, 2001);

2. Média e variância:

As distribuições de probabilidade dos dados de interesse devem ser inteiramente descritas através de média e variância, como as distribuições exponenciais, gaussianas e uniformes. Isso assegura que a matriz de covariância caracteriza inteiramente as redundâncias (WOLD; ESBENSEN; GELADI, 1987);

3. A variabilidade é importante:

A premissa também assume que os dados possuem alta SNR. Por conseguinte, os CPs com variâncias altas associadas apresentam dinâmica importante, ao passo que os CPs cujas variâncias são menores possivelmente caracterizam apenas ruído (WOLD; ESBENSEN; GELADI, 1987);

4. Linearidade:

O conceito de linearidade classifica o problema como uma mudança de base (WOLD; ESBENSEN; GELADI, 1987);

5. Os CPs são ortogonais:

Essa premissa torna a PCA solúvel através de decomposição via álgebra linear (WOLD; ESBENSEN; GELADI, 1987).

2.2.4 Branqueamento

O jargão “branco” é oriundo do espectro de energia do ruído branco ser constante em todas as frequências (HYVÄRINEN; KARHUNEN; OJA, 2001). É uma técnica simples quando deseja-se descorrelacionar medidas de interesse.

Um vetor $\mathbf{z} = [z_1, \dots, z_n]^T$ de média zero é dito branco quando seus elementos são descorrelacionados e de variância unitária $E\{z_i z_j\} = \sigma_{ij}$. Neste caso, a matriz de covariância de \mathbf{z} é $E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{I}$. Além disso, segundo a literatura, o branqueamento simplifica o problema para a ICA (HYVÄRINEN; KARHUNEN; OJA, 2001).

O branqueamento pode ser efetuado via operação linear. Assim, considere um vetor aleatório $\mathbf{x} = [x_1, \dots, x_n]^T$ de n elementos. Então é necessário encontrar uma transformação linear \mathbf{V} , como (2.30), em que \mathbf{z} seja branco.

$$\mathbf{z} = \mathbf{V}\mathbf{x} \quad (2.30)$$

Neste caso, \mathbf{V} é denominada matriz de branqueamento. Tal matriz é determinada via (2.31), onde \mathbf{D} é uma matriz diagonal constituída pelos autovalores de $\mathbf{C}_\mathbf{x} = E\{\mathbf{x}\mathbf{x}^T\}$. Além disso, $\mathbf{E} = (\mathbf{e}_1 \dots \mathbf{e}_n)$, onde as colunas são constituídas pelos autovetores de norma unitária de $\mathbf{C}_\mathbf{x}$.

$$\mathbf{V} = \mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T \quad (2.31)$$

Adotando o modelo trivial (2.32), cujas fontes são descorrelacionadas, a independência estatística deve manter-se, conforme (2.33), resultando na matriz identidade.

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (2.32)$$

$$E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{V}E\{\mathbf{x}\mathbf{x}^T\}\mathbf{V}^T = \mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T\mathbf{E}\mathbf{D}\mathbf{E}^T\mathbf{E}\mathbf{D}^{-\frac{1}{2}} = \mathbf{I} \quad (2.33)$$

Assim, em duas dimensões uma matriz ortogonal pode ser encontrada apenas por um ângulo, e assim, a matriz ortogonal apresenta aproximadamente metade da quantidade de parâmetros da matriz original (LIMA, 2013). Logo, reduz fortemente o problema para a ICA.

2.3 FUNDAMENTOS DA METODOLOGIA DE LOCALIZAÇÃO

Após a apresentação das técnicas inerentes à etapa de estimação, elucidam-se os conceitos necessários para a compreensão da etapa de localização. Inicialmente, uma breve revisão é feita acerca da origem da entropia e sua relação com níveis de incerteza. Após a apresentação mencionada, os conceitos probabilísticos necessários para o entendimento da entropia conjunta são sumarizados. Por fim, as entropias são definidas formalmente.

2.3.1 A Origem da Entropia e sua Relação com Níveis de Incerteza

A concepção do conceito de entropia propiciou um grande marco na ciência, impactando diretamente a mecânica estatística, a teoria da termodinâmica, áreas correlatas à filosofia, economia, computação e também possui papel central na teoria da informação. Muitos autores trabalharam na generalização da entropia na teoria da informação como Rényi, Havrda, Charvat e Daróczy (RÉNYI *et al.*, 1961; HAVRDA; CHARVÁT, 1967; DARÓCZY, 1970). Já Constantino Tsallis propôs uma generalização para a mecânica estatística (TSALLIS, 1988).

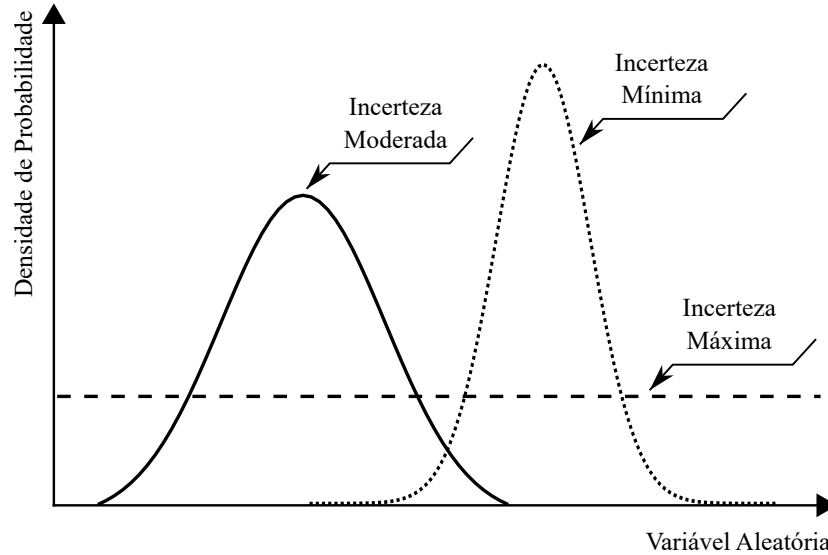
Na metade do século 19, no contexto da termodinâmica, surge pela primeira vez o conceito de entropia. Neste contexto, a entropia é interpretada como medida de irreversibilidade dos processos físicos. Um processo irreversível não pode ser operado de maneira inversa, isto é, uma parcela da energia despendida em forma de trabalho é convertida em calor, que por sua vez, é perdido. Assim, essa parcela de energia perdida impossibilita a capacidade do sistema gerar trabalho definitivamente com essa parcela de energia perdida. Na termodinâmica, a geração de calor é associada a um aumento da entropia. Em geral, processos naturais são irreversíveis.

Na década de 70, Boltzmann formulou inicialmente a teoria estatística da mecânica, que por sua vez, a conceitos microscópicos associou-se a entropia (uma variável macroscópica). A mecânica estatística propõe-se a determinar propriedades macroscópicas via informações microscópicas. É Neste contexto que surge a ideia de que quanto maior a desordem, maior a entropia.

Na teoria da informação, o conceito de entropia foi motivado pela intenção em decifrar mensagens criptografadas de guerra. Em 1948, Claude Shannon inicia o conceito de entropia da informação (SHANNON, 1948). Assim, surge a interpretação de “grau de incerteza existente antes de efetuar uma escolha”. E mais especificamente, o conceito faz referência à incerteza de uma distribuição de probabilidade. Existem três interpretações para a incerteza. Inicialmente, a incerteza determinística, cujos estados de um sistema são desconhecidos. A incerteza entrópica, cujos estados possíveis são conhecidos, entretanto, as chances de ocorrência de cada um deles é desconhecida. E por fim, a incerteza probabilística, cujos estados são conhecidos e as distribuições de probabilidade que os governam também as são (MATTOS; VEIGA, 2002). A Figura 9 apresenta a relação entre as distribuições de

probabilidade e a incerteza. Onde, a entropia busca quantificar a incerteza probabilística. Quando tratamos de níveis de incerteza, Figura 9, significa que um nível de incerteza alto corresponde ao caos da informação, isto é, não há uma direção em que se possa afirmar algo sobre um determinado experimento.

Figura 9 – Níveis de incerteza



Fonte: Elaborado pelo autor (2020).

Neste aspecto, Shannon formulou uma medida para computar o grau de incerteza de uma distribuição de probabilidade: a equação (2.34).

$$S(\mathbf{p}) = - \sum_{i=1}^N \ln p_i, \quad (2.34)$$

onde S é a entropia de Shannon, $\mathbf{p}^T = [p_1, \dots, p_N]$ é a distribuição de probabilidade e T é o operador de transposição (MATTOS; VEIGA, 2002).

2.3.2 Entropia

Segundo a literatura, a entropia é uma medida de incerteza de uma variável aleatória (COVER; THOMAS, 2012). Assim considere X como uma variável aleatória discreta com conjunto \mathcal{X} e função de massa (ou densidade, para o caso contínuo) de probabilidade $p(x) = Pr\{X = x\}$, $x \in \mathcal{X}$. Denota-se a função de massa de probabilidade por $p(x)$ em vez de $p_X(x)$ por conveniência. Assim, $p(x)$ e $p(y)$ referem-se a diferentes variáveis aleatórias e são diferentes funções de massa de probabilidade, $p_X(x)$ e $p_Y(y)$.

A entropia $H(X)$ de uma variável aleatória discreta X é definida por:

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log p(x) \quad (2.35)$$

É possível escrever $H(p)$ para a quantidade acima, se temos que a entropia de um lance, não viciado, de moeda é de 1 bit. Utiliza-se que $0 \log 0 = 0$, que é justificado pela continuidade desde $x \log x \rightarrow 0$ como $x \rightarrow 0$. Portanto, adicionando termos de probabilidade zero não altera-se a entropia. Porém, se a base logarítmica for b , denota-se a entropia por $H_b(X)$. Note que a entropia é uma função de distribuição de X , não depende dos valores assumidos pela variável aleatória X , mas somente das suas probabilidades. Se $X \sim p(x)$, então a esperança E de uma variável aleatória $g(X)$ é escrita por:

$$E_p g(X) = \sum_{x \in \mathcal{X}} g(x) p(x) \quad (2.36)$$

Ou simplesmente $Eg(X)$, quando a função de massa de probabilidade está previamente definida pelo contexto. Observando pela ótica da esperança, para $g(X) = \log \frac{1}{p(x)}$, encontra-se que a entropia de X também pode ser interpretada como um valor de $\log \frac{1}{p(x)}$, onde X é desenhada de acordo com a f.m.p.¹ $p(x)$, logo:

$$H(X) = E_p \log \frac{1}{p(X)} \quad (2.37)$$

Essa definição de entropia está relacionada com a definição de entropia na termodinâmica. Outro modo de compreender o conceito de entropia (H) da distribuição de probabilidades, é considerar que essa entropia é uma média ponderada das entropias dos eventos dessa distribuição (h_i). A entropia do evento i é dada por $h_i = \log(1/p_i)$. Tomando o exemplo de lances de moedas honestas, a entropia dos eventos poderia ser 1 bit, assim como a entropia do evento coroa, da mesma forma que a média ponderada, com probabilidade 0,5 para cada um dos eventos, também seria de 1 bit.

Além disso, seguem algumas consequências da definição de entropia (COVER; THOMAS, 2012):

1. $H(X) \geq 0$;
2. $H_b(X) = (\log_b a)H_a(X)$

A segunda consequência decorre apenas da mudança de base logarítmica.

2.3.2.1 Entropia Conjunta

Na seção anterior foi definida a entropia de uma variável aleatória. Agora será ampliada a definição para um par de variáveis aleatórias. Não há algo novo na definição porque (X, Y) podem ser consideradas uma variável aleatória de um vetor único.

A entropia conjunta $H(X, Y)$ de um par de variáveis aleatórias discretas (X, Y) com distribuição conjunta $p(x, y)$, conforme a referência (COVER; THOMAS, 2012), é dada por:

¹ Função Massa de Probabilidade (f.m.p.) definida no Apêndice A

$$H(X, Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y) \quad (2.38)$$

Que pode ser expressada também como:

$$H(X, Y) = -E \log p(X, Y) \quad (2.39)$$

2.3.2.2 Entropia Condicional

A entropia condicional é expressa de acordo com as outras entropias e distribuições conjuntas. Logo, se $(X, Y) \sim p(x, y)$, a entropia condicional $H(Y|X)$ é definida como:

$$\begin{aligned} H(X|Y) &= \sum_{x \in \mathcal{X}} p(x) H(Y|X = x) \\ &= - \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log p(y|x) \\ &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y|x) \\ &= -E \log p(Y|X) \end{aligned} \quad (2.40)$$

A relação entre entropia conjunta e condicional pode ser estabelecida, resumidamente, conforme (2.41):

$$H(X, Y) = H(X) + H(Y|X) \quad (2.41)$$

É importante salientar que $H(Y|X) \neq H(X|Y)$, porém, $H(X) - H(X|Y) = H(Y) - H(Y|X)$.

Finalmente, é válido frisar que a entropia conjunta, a qual é utilizada Neste trabalho, trabalha com a níveis de incerteza. Assim, quanto menor a incerteza entre as medidas estudadas, por conseguinte, maior é a dependência estatística entre essas medidas. Logo, o valor mínimo de entropia conjunta corresponde à menor incerteza ou a maior dependência estatística entre as estimativas e as tensões medidas.

2.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo, a priori, apresentou-se uma revisão do estado da arte acerca da temática de detecção de fontes harmônicas de corrente. Muitos trabalhos tratam apenas da estimação dessas fontes, outros apenas da localização, e em alguns casos, ambas as etapas compõem uma metodologia de detecção. Assim, a revisão contida nos fundamentos teóricos busca englobar todas as vertentes. Entretanto, é válido enfatizar que o presente capítulo apresentou conceitos e definições acerca das técnicas que promovem a detecção, ou localização, de fontes harmônicas de corrente, visto que, o problema em si não foi abordado.

Alguns trabalhos notórios para essa dissertação são interessantes de serem citados: I) Embora Lin et al. (2005) tenha sua metodologia caracterizada como supervisionada, pois utiliza uma rede de correlação em cascata, os autores utilizam THD como parâmetro para a poluição harmônica; II) O trabalho proposto por Gursoy (2007) utiliza ferramentas de soma importância para a BSS e é utilizado com base para a etapa de estimação desta dissertação; III) Finalmente, o trabalho proposto por Zang et al. (2016) faz utilização da entropia, neste caso entropia condicional, cuja aplicação serve como base para a etapa de localização de fontes harmônicas de corrente.

Após a revisão da literatura, são conferidos todos os fundamentos empregados para a estimação das fontes harmônicas de corrente, tratando principalmente do contexto de BSS, em que o trabalho está inserido, bem como a utilização da ferramenta de redução de dimensionalidade, a PCA, e também a ICA, a ferramenta primordial de estimação. Por conseguinte, as ferramentas estatísticas oriundas da teoria da informação são descritas formalmente, inclusive, elencando alguns momentos históricos e sua relação probabilística com os níveis de incerteza, e assim, finalizando os fundamentos necessários para o entendimento para a etapa de localização utilizada neste trabalho.

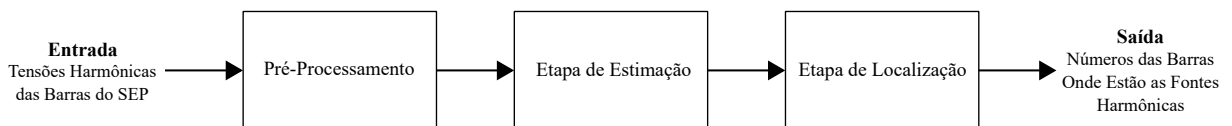
3 METODOLOGIA

Neste capítulo será apresentada a metodologia desenvolvida para a detecção de fontes harmônicas de corrente. Todos os conceitos e ferramentas utilizadas para o desenvolvimento do método serão elencados. Como o presente trabalho não possui foco majoritariamente na modelagem dos parâmetros e modelos da área de potência, os mesmos serão citados brevemente e carregarão a descrição necessária para o entendimento de sua utilização e aplicação nesta dissertação.

O Diagrama 1 apresenta as etapas essenciais que compõem a metodologia como um todo. Inicialmente é necessário aferir as tensões harmônicas do SEP. Após essa medição, efetua-se o pré-processamento composto por: PCA e filtragem. A PCA efetua a redução de dimensionalidade nas tensões medidas. Aos componentes principais determinados, os CPs, efetua-se uma filtragem. Assim, a parcela referente a variação rápida do sinal é utilizada pelo algoritmo de ICA na etapa de estimação, com a finalidade de promover a estimação dos perfis lentos das fontes harmônicas de corrente. Finalmente, tais perfis são utilizados na etapa final de localização.

Inicialmente, neste capítulo, apresenta-se o modelo de ICA utilizado e suas equivalências, por exemplo, as adequações do modelo de misturas às equações de eletricidade. Então, o modelo de misturas é inserido na perspectiva da etapa de redução de dimensionalidade. Posteriormente, os componentes oriundos da redução de dimensionalidade serão tratados nos aspectos do processo de filtragem. Assim, encerra-se a etapa de estimação com os algoritmos testados. Após a estimação, a entropia conjunta é tratada na etapa de localização, bem como, a abordagem utilizada conhecida por “um contra todos”, ilustrando o desempenho dos algoritmos na etapa de localização.

Diagrama 1 – Etapas essenciais à metodologia



Fonte: Elaborado pelo autor (2020).

3.1 MODELO DE ICA ADOTADO E ADEQUAÇÕES

Inicia-se a modelagem da metodologia com o desejo de detectar fontes harmônicas de corrente para SEPs que possuem topologias não especificadas. Para tal, adota-se o modelo (3.1) que relaciona tensões e correntes através das impedâncias da rede. É válido salientar que a equação (3.1) está relacionada a uma ordem harmônica específica, e assim,

um índice h é adicionado ao vetor de tensões medidas.

$$\begin{aligned}
 V_1^h[n] &= Z_{11}^h I_1^h[n] + Z_{12}^h I_2^h[n] + \cdots + Z_{1N}^h I_N^h[n] \\
 V_2^h[n] &= Z_{21}^h I_1^h[n] + Z_{22}^h I_2^h[n] + \cdots + Z_{2N}^h I_N^h[n] \\
 &\vdots \\
 V_M^h[n] &= Z_{M1}^h I_1^h[n] + Z_{M2}^h I_2^h[n] + \cdots + Z_{MN}^h I_N^h[n]
 \end{aligned} \tag{3.1}$$

onde:

$I_i^h[n]$: fontes harmônicas de corrente ($i = 1, 2, 3, \dots, N$);

$V_k^h[n]$: tensões medidas ($k = 1, 2, 3, \dots, M$);

Z_{ki}^h : impedância de ordem harmônica h do trecho da rede entre a fonte harmônica i e de notação compacta conforme (3.2), com um vetor de fontes harmônicas de corrente \mathbf{I} , um vetor de tensões \mathbf{V} e uma matriz de impedâncias \mathbf{Z} .

$$\mathbf{V}^h = \mathbf{Z}^h \mathbf{I}^h \tag{3.2}$$

Conforme (3.1), para a adequação ao modelo de ICA adotado, considera-se a premissa da existência de N fontes e a utilização de M medidores. Por conseguinte, o modelo de mistura linear da ICA, ou generativo (seção 2.2.2.2), deve ser utilizado e pode ser minuciosamente expandido conforme (3.3):

$$\begin{aligned}
 x_1[n] &= a_{11}s_1[n] + a_{12}s_2[n] + \cdots + a_{1N}s_N[n] + \xi_1[n] \\
 x_2[n] &= a_{21}s_1[n] + a_{22}s_2[n] + \cdots + a_{2N}s_N[n] + \xi_2[n] \\
 &\vdots \\
 x_M[n] &= a_{M1}s_1[n] + a_{M2}s_2[n] + \cdots + a_{MN}s_N[n] + \xi_M[n]
 \end{aligned} \tag{3.3}$$

em que:

$\mathbf{s}[n] = \{s_1[n], \dots, s_N[n]\}^T$: vetor de fontes com dimensão N ;

$\mathbf{x}[n] = \{x_1[n], \dots, x_M[n]\}^T$: vetor de misturas com tamanho M ;

a_{ij} : coeficientes reais da matriz de misturas;

$\xi_i[n]$: ruído aditivo;

n : índice da amostra, cujo $n \in \mathbb{Z}^+$.

Entretanto, o ruído pode ser negligenciado conforme (GURSOY, 2007), semelhante ao modelo trivial (seção 2.2.2.1), o que resulta em (3.4):

$$\begin{aligned}
 x_1[n] &= a_{11}s_1[n] + a_{12}s_2[n] + \cdots + a_{1N}s_N[n] \\
 x_2[n] &= a_{21}s_1[n] + a_{22}s_2[n] + \cdots + a_{2N}s_N[n] \\
 &\vdots \\
 x_M[n] &= a_{M1}s_1[n] + a_{M2}s_2[n] + \cdots + a_{MN}s_N[n]
 \end{aligned} \tag{3.4}$$

e de notação compacta conforme (3.5):

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{3.5}$$

Nessa etapa, algumas equivalências podem consideradas:

$$\mathbf{I}^h \iff \mathbf{s} \tag{3.6}$$

$$\mathbf{V}^h \iff \mathbf{x} \tag{3.7}$$

$$\mathbf{Z}^h \iff \mathbf{A} \tag{3.8}$$

Ao aplicar a ICA, obtém-se uma matriz de separação \mathbf{W} responsável por determinar estimativas das fontes harmônicas de corrente ($\hat{\mathbf{s}}$). Essa ação está na forma compacta (3.9):

$$\hat{\mathbf{s}} = \mathbf{W}\mathbf{x}, \tag{3.9}$$

o que, quando expandido, é equivalente a (3.10):

$$\begin{aligned}
 \hat{s}_1[n] &= w_{11}x_1[n] + w_{12}x_2[n] + \cdots + w_{1M}x_M[n] \\
 \hat{s}_2[n] &= w_{21}x_1[n] + w_{22}x_2[n] + \cdots + w_{2M}x_M[n] \\
 &\vdots \\
 \hat{s}_N[n] &= w_{N1}x_1[n] + w_{N2}x_2[n] + \cdots + w_{NM}x_M[n]
 \end{aligned} \tag{3.10}$$

em que:

$\hat{\mathbf{s}}[n] = \{\hat{s}_1[n], \dots, \hat{s}_N[n]\}^T$: vetor de estimativas das fontes com tamanho N ;

$\mathbf{x}[n] = \{x_1[n], \dots, x_M[n]\}^T$: vetor de misturas com tamanho M ;

w_{ij} : coeficientes da matriz de separação \mathbf{W} ;

n : índice da amostra de valor inteiro positivo.

além disso, sabe-se que $\mathbf{A}^{-1} \approx \mathbf{W}$. Logo, existem as correspondências estabelecidas abaixo:

$\hat{\mathbf{s}}$: corresponde de forma semelhante a [\mathbf{I}];

\mathbf{W} : corresponde de forma semelhante a [\mathbf{Y}].

3.2 REDUÇÃO DE DIMENSIONALIDADE

Pensando em situações cujos sistemas sejam formados por uma grande topologia, onde ocorra a situação de sobre determinação (seção 2.2.1), onde existam mais medidores conectados à rede do que fontes, é sugerida a utilização de uma ferramenta de redução de dimensionalidade.

A redução de dimensionalidade é justificável, pois extrapolando para sistemas cuja topologia seja relativamente extensa, e em situações de sobre determinação, existe um forte indício de que informações redundantes estejam sendo aferidas. Além disso, outro fato é que a redução de dimensionalidade é fortemente recomendada para sistemas contextualizados em BD, onde uma quantidade muito grande de medidas devem ser processadas a complexidade computacional torna-se, também, uma preocupação.

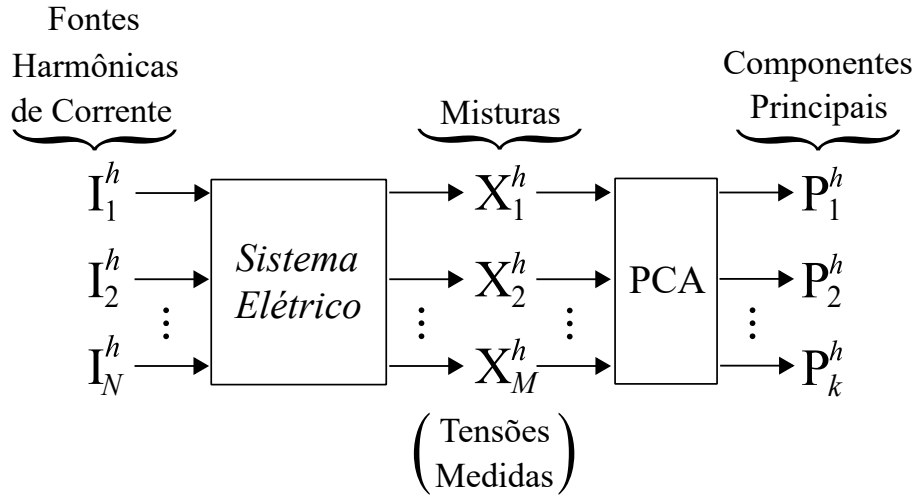
Neste caso, utiliza-se a PCA devido a sua facilidade de implementação e entendimento intuitivo (seção 2.2.3.1). O modelo de misturas permanece o mesmo conforme a equação (3.5) ou $\mathbf{x} = \mathbf{A}\mathbf{s}$. Lembrando que nesta etapa não há estimação, apenas uma transformação é realizada Neste modelo de misturas, ou seja, uma transformação é realizada na parcela conhecida do modelo, que é \mathbf{x} (tensões medidas). Assim, vislumbrando a diminuição da complexidade e conseqüentemente o esforço computacional, uma etapa de pré-processamento é inserida via PCA.

Como visto na seção (2.2.3.1), a PCA é uma transformação linear e pode ser utilizada sem prejudicar o modelo da ICA, pois ambas as técnicas adotam um modelo de misturas lineares. Então, a transformação da PCA, representada pela matriz Φ , é aplicada nas tensões medidas de forma que os dados originais sejam excursionados para um espaço cuja variância seja privilegiada.

$$\bar{\mathbf{P}} = \mathbf{A}\mathbf{s}\Phi = \mathbf{x}\Phi \quad (3.11)$$

Assim, um novo conjunto de dados é gerado, representado por $\bar{\mathbf{P}}$ em (3.11). A partir desse novo conjunto de dados faz-se o ranqueamento dos autovalores de $\bar{\mathbf{P}}$ a partir dos autovetores de $\bar{\mathbf{P}}$, conforme a seção (2.2.3.1). Por fim, têm-se o conjunto resultante de CPs representado por \mathbf{P} (tal que $\mathbf{P} \in \bar{\mathbf{P}}$). A etapa de compressão pode ser representada pelo Diagrama 2, onde as misturas \mathbf{x} são representadas pelas tensões de ordem harmônica h e aferidas em M barras e os k CPs resultantes representados por $\mathbf{P}_1^h, \dots, \mathbf{P}_k^h$.

Diagrama 2 – Processo de redução de dimensionalidade via PCA



Fonte: Elaborado pelo autor (2020).

Após a redução de dimensionalidade, com a finalidade de aprimorar os resultados da etapa de estimação, utiliza-se um processo de filtragem nos CPs resultantes.

3.3 FILTRAGEM

Segundo (GURSOY, 2007), as variações rápidas e lentas, ou os perfis rápidos e lentos, podem ser separados via filtragem. A filtragem é importante pois o modelo de ICA é proposto para a separação de fontes estatisticamente independentes. Logo, é interessante obter separadamente as variações rápidas. Conforme (HYVÄRINEN, 1998), a filtragem linear não prejudica o modelo trivial de ICA, pois trata-se de um modelo linear de misturas. A transformação ocasionada pela filtragem pode ser representada por Ψ e aplicada conforme se segue:

$$\mathbf{x}_{slow} = \mathbf{x}\Psi, \quad (3.12)$$

como $\mathbf{x} = \mathbf{A}\mathbf{s}$, logo:

$$\mathbf{x}_{slow} = \mathbf{x}\Psi = \mathbf{A}\mathbf{s}\Psi = \mathbf{A}\mathbf{s}_{slow}, \quad (3.13)$$

ao passo que, a variação rápida pode ser obtida conforme:

$$\mathbf{x}_{fast} = \mathbf{x} - \mathbf{x}_{slow} = \mathbf{A}\mathbf{s} - \mathbf{A}\mathbf{s}_{slow}, \quad (3.14)$$

ou seja,

$$\mathbf{x}_{fast} = \mathbf{A}\mathbf{s}_{fast}, \quad \text{pois } \mathbf{s} - \mathbf{s}_{slow} = \mathbf{s}_{fast}, \quad (3.15)$$

Após a filtragem linear, os componentes de maior variabilidade e independência correspondem ao perfil rápido. Assumindo que a flutuação rápida é independente e não possui distribuição gaussiana, a ICA pode ser aplicada. Porém, como existe uma etapa anterior de redução de dimensionalidade (PCA), a filtragem é aplicada posteriormente nos CPs oriundos dos resultados obtidos da PCA. Novamente, o modelo misturas lineares para a ICA não é prejudicado porque, conforme apresentado na seção (2.2.3.1), a PCA é uma transformação linear da mesma forma que a filtragem também apresenta essa característica.

Neste processo, opta-se por um filtro média móvel, visto que, trata-se de uma filtragem linear. Neste caso, o filtro de média móvel apresenta na saída o perfil lento porque é caracterizado como um passa-baixas. Além disso, situações em que deseja-se eliminar algum nível de ruído, a utilização dos filtros média móvel é uma solução comumente utilizada. Outro fato é que, o filtro média móvel possui baixa complexidade computacional. Tal filtro é constituído pela média de um conjunto de valores de interesse. Matematicamente, conforme (MITRA; KUO, 2006), o filtro média móvel é definido pela equação (3.16):

$$\mathbf{P}_{sl}[n] = \frac{1}{L+1} \sum_{i=0}^L \mathbf{P}[n-i], \quad (3.16)$$

em que $\mathbf{P}_{sl}[n]$ é o sinal filtrado, $\mathbf{P}[n-i]$ é o conjunto de valores que serão somados de forma deslizante, n é o índice da amostra atual, $L+1$ é o número de amostras utilizadas no processo de filtragem e $sl = slow$.

Neste caso, deseja-se maximizar a independência na medida. Assim, o processo de filtragem fica encarregado de separar a variação rápida da variação lenta dos CPs, essa “transformação” é representada novamente por Ψ , conforme (3.17):

$$\mathbf{P}_{sl} = \mathbf{P}\Psi \quad (3.17)$$

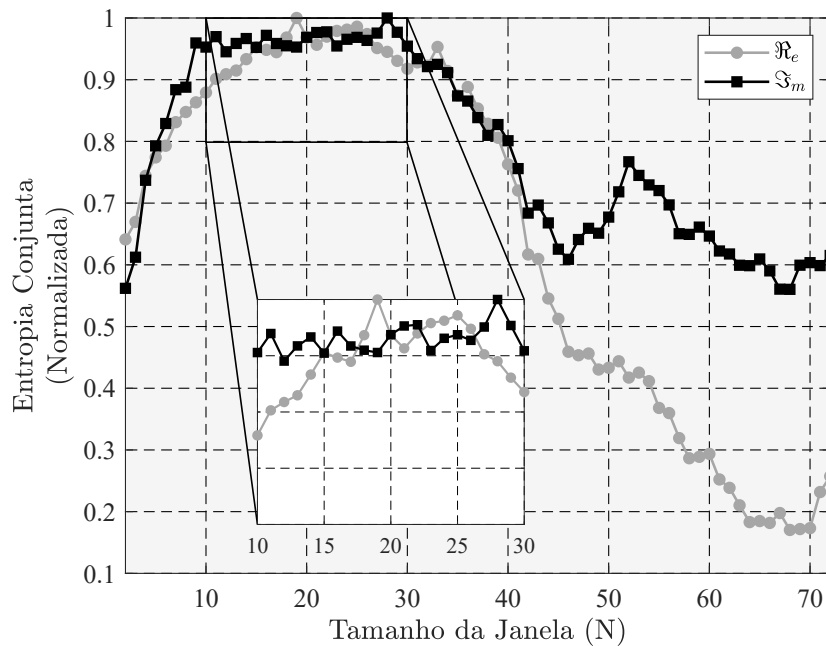
$$\mathbf{P}_{fs} = \mathbf{P}(1 - \Psi) = \mathbf{P} - \mathbf{P}\Psi = \mathbf{P} - \mathbf{P}_{sl}, \quad (3.18)$$

onde \mathbf{P}_{fs} representa a variação rápida contida nos CPs.

De fato a parcela de maior interesse dos componentes principais é \mathbf{P}_{fs} , pois carrega a maior maximização da independência entre os componentes. Essa maximização depende do tamanho da janela utilizada no projeto do filtro. Sabendo disso, utilizou-se a entropia conjunta com a finalidade de inferir qual o intervalo cujo o tamanho da janela favoreça a melhor maximização de independência entre os CPs. Isso se justifica porque a ICA admite a premissa de fontes independentes, logo esse “intervalo de confiança” ajudará a determinar um bom tamanho de janela, que facilite o processo de separação das fontes pela ICA.

Para um sistema de teste IEEE 14 barras, com duas fontes harmônicas de corrente que inserem apenas o 5º harmônico, a abordagem supracitada foi aplicada. O resultado obtido indicou que o melhor intervalo de maximização de independência, entre os componentes, está entre os tamanhos de 10 e 30 amostras, conforme a Figura 10. Neste caso, uma janela de $N = 11$ amostras é escolhida pois está contida dentro do intervalo de confiança e é uma janela de tamanho relativamente pequeno, o que confere uma complexidade computacional reduzida. Porém, do ponto de vista estatístico, qualquer janela contida Neste intervalo favorecerá a independência estatística.

Figura 10 – Intervalo de confiança



Fonte: Elaborado pelo autor (2020).

Além disso, é importante salientar que como os componentes pertencem ao domínio complexo a entropia deve ser aplicada tanto na parte real quanto na parte imaginária dos componentes, como ilustrado na Figura 10.

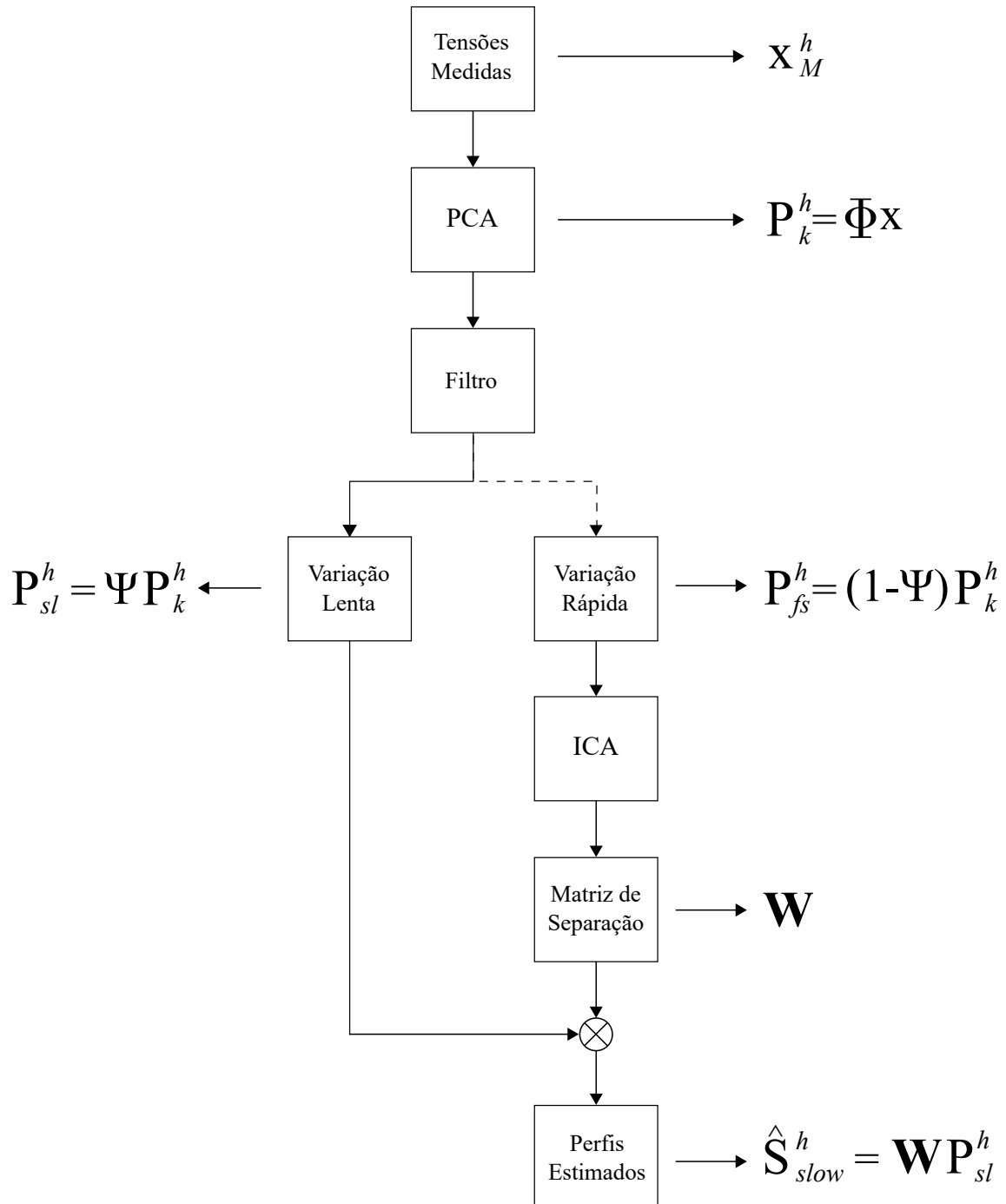
Como o perfil lento carrega a variação principal da carga, sua parcela é mais interessante para as análises.

O processo de estimação do perfil lento pode ser descrito com os seguintes passos: I) após a PCA, componentes complexos são determinados (CPs); II) então efetua-se a filtragem, com o objetivo de separar a variação rápida da variação lenta dos CPs; III) a variação rápida dos CPs é conduzida ao algoritmo de ICA que determina a matriz de separação \mathbf{W} ; IV) finalmente, utiliza-se a variação lenta dos CPs juntamente com a matriz \mathbf{W} para a estimação dos perfis lentos, conforme a equação (3.19).

$$\hat{\mathbf{s}}_{slow} = \mathbf{W}\mathbf{P}_{sl} \quad (3.19)$$

O processo de estimação de fontes harmônicas de corrente pode ser sumarizado conforme o Diagrama 3 e é baseado no trabalho desenvolvido pela referência (GURSOY, 2007).

Diagrama 3 – Diagrama da estimação de fontes harmônicas de corrente



Fonte: Elaborado pelo autor (2020).

3.4 ALGORITMOS DE BSS

Nesta seção serão apresentados de forma resumida os algoritmos de ICA utilizados no presente trabalho. A motivação em testar tais algoritmos reside em detectar qual algoritmo possui o melhor desempenho para esse tipo de aplicação e sua influência na etapa de localização.

Os algoritmos de BSS podem ser das mais variadas naturezas, entretanto, o objetivo dos mesmos reside em determinar parâmetros que possibilitem a separação das fontes, ou separá-las propriamente. Dentre os mesmos, encontram-se as mais variadas formas para alcançar esse objetivo em comum. Existem algoritmos baseados em estatísticas de segunda ordem, por exemplo, o algoritmo SOBI, e algoritmos baseados em Estatística de Ordem Superior (EOS), como o RobustICA.

Os algoritmos baseados em EOS, em geral, focam-se na maximização da não gaussianidade das fontes. Ou seja, estão inseridos na restrição de não gaussianidade. Segundo (LIMA, 2009), o conceito da maximização da não gaussianidade pode ser sintetizado no ajuste do sistema separador para que cada uma das densidades de probabilidade correspondentes as suas estimativas, sejam o mais distante possível de uma variável gaussiana. Os algoritmos baseados em EOS, são verdadeiramente considerados algoritmos de ICA.

A maximização da não gaussianidade baseia-se no princípio do Teorema do Limite Central. Assume-se que as misturas devem ser mais gaussianas do que as fontes (LIMA, 2009). Para ilustrar essa ideia, assumamos que deseja-se separar uma única fonte e para isso utiliza-se o vetor de separação \mathbf{w}_1 .

$$y_1 = \mathbf{w}_1^T \mathbf{x} = \mathbf{q}^T \mathbf{s}, \quad (3.20)$$

cuja estimativa y_1 seja a equação (3.20), onde $\mathbf{q} = \mathbf{w}_1^T \mathbf{A}$. Como y_1 é uma combinação linear de \mathbf{s} , logo, é necessário que apenas um elemento do vetor \mathbf{q} seja não nulo para que se obtenha uma estimativa da fonte. Assim, quando \mathbf{q} possuir somente um elemento não nulo, y_1 será o menos gaussiano possível e estará mais próximo da distribuição de uma das fontes de \mathbf{s} . Portanto, a maximização da não gaussianidade é realizada através da determinação de um vetor \mathbf{w}_1 capaz de maximizar a não gaussianidade de $\mathbf{w}_1^T \mathbf{x}$.

Nos tópicos subsequentes serão apresentados os algoritmos de BSS e suas características, quando necessário, as abordagens para a maximização da não gaussianidade.

3.4.1 Second Order Blind Identification (SOBI)

O algoritmo SOBI foi proposto por Belouchrani em (BELOUCHRANI *et al.*, 1997), sendo o mesmo baseado em estatísticas de segunda ordem. Tal algoritmo consiste resumidamente, conforme (CICHOCKI; AMARI, 2002), nas seguintes etapas:

1. Branqueamento das misturas, conforme (3.21):

$$\bar{\mathbf{x}} = \mathbf{Q}\mathbf{x}; \quad (3.21)$$

onde \mathbf{Q} é a matriz de branqueamento.

2. Estimar o conjunto de matrizes de covariância, conforme (3.22):

$$\hat{\mathbf{C}}_{\bar{\mathbf{x}}} = \mathbf{Q}\mathbf{C}_{\mathbf{x}}\mathbf{Q}^T \quad (3.22)$$

3. Efetuar a diagonalização, de acordo com (3.23):

$$\mathbf{U}\mathbf{D}_i\mathbf{U}^T = \hat{\mathbf{C}}_{\bar{\mathbf{x}}} \quad (3.23)$$

o qual significa estimar a matriz ortogonal \mathbf{U} . Ambas as matrizes \mathbf{U} e \mathbf{D}_i estão definidas em (CICHOCKI; AMARI, 2002);

4. Por fim, estimar as fontes via (3.24):

$$\mathbf{s} = \mathbf{U}^T\mathbf{Q}\mathbf{x} \quad (3.24)$$

O SOBI é muito utilizado em abordagens *single-channel* para BSS. Como baseia-se em estatísticas de segunda ordem e não de ordem superior, não pode ser tratado como um algoritmo de ICA. Entretanto, o objetivo em testá-lo neste trabalho é verificar seu desempenho, pois é um algoritmo que não possui a restrição de que as fontes não possuam distribuição gaussiana.

3.4.2 FastICA

Segundo (ZARZOSO; COMON, 2010) o algoritmo FastICA foi inicialmente desenvolvido apenas para sinais de valor real. Entretanto, existe sua extensão para fontes complexas. Sua extensão para fontes de valor complexo mostra manter a propriedade de convergência global cúbica de sua contraparte real. Além disso, diferentemente dos algoritmos baseados em gradiente, o FastICA não necessita da decisão a priori de nenhum parâmetro de taxa de aprendizagem, tornando-o mais simples de utilizar.

O FastICA pode ser utilizado via negentropia ou kurtosis, pois é um algoritmo que baseia-se na maximização da não gaussianidade das fontes. Neste caso, será tratada a abordagem via negentropia.

A negentropia é uma medida de não gaussianidade em que, dado um conjunto cujas variáveis aleatórias que possuam a mesma variância, a gaussiana é a variável cuja entropia é a maior. Assim, sugere-se a busca da minimização da gaussianidade via minimização

das entropias marginais das estimativas (LIMA, 2009). Dada uma variável aleatória y , a negentropia da mesma pode ser obtida conforme (3.25):

$$J_{Neg}(y) = H(y_{gs}) - H(y), \quad (3.25)$$

em que $H(\cdot)$ é a entropia e y_{gs} corresponde a uma variável aleatória de distribuição gaussiana cuja média e variância são iguais às de y . Além disso, a maximização da negentropia proporciona a maximização da não gaussianidade. Devido a sua determinação ser complicada matematicamente, utilizam-se aproximações como a não polinomial definida em (3.26), para uma variável aleatória y :

$$J(y) = \alpha(E\{G(y)\} - E\{G(v)\})^2, \quad (3.26)$$

cujo α é uma constante, v é uma variável aleatória gaussiana de média zero e variância unitária e $G(\cdot)$ uma função não linear e não quadrática.

Como o FastICA baseia-se na aproximação (3.26), assuma que seja desejável estimar apenas um componente, ou equivalentemente, ajustar uma única linha de \mathbf{W} (equivalente a \mathbf{w}_i^T). Assim, uma estimativa corresponderá a $y_i = \mathbf{w}_i^T x$. Logo, o algoritmo pode ser sumarizado conforme:

1. Inicializar aleatoriamente \mathbf{w}_i ;
2. Executar as iterações via (3.27):

$$\mathbf{w}_i \leftarrow E\{\mathbf{x}\dot{G}(\mathbf{w}_i^T \mathbf{x})\} - E\{\ddot{G}(\mathbf{w}_i^T \mathbf{x})\}\mathbf{w}_i; \quad (3.27)$$

3. Aplicar a normalização abaixo:

$$\mathbf{w}_i \leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|; \quad (3.28)$$

4. Voltar ao passo 2 caso não convirja.

E para a estimação das fontes seguintes, o processo se repete.

A convergência é cúbica, ou pelo menos quadrática, com base na suposição do modelo de dados da ICA. Isso faz oposição com outros algoritmos de ICA, pois os que são baseados em métodos de gradiente descendente apresentam convergência linear. Assim o FastICA apresenta uma rápida convergência. Além disso, pode-se estimar as componentes independentes uma a uma. Isso diminui o custo computacional em situações onde apenas algumas componentes independentes precisam ser estimadas.

3.4.3 RobustICA

Em (ZARZOSO; COMON, 2010) é proposto um novo método para a deflação ICA, chamado RobustICA. Esse método é baseado na função de contraste geral, a Kurtosis, que é otimizada por uma técnica computacionalmente eficiente oriunda de um passo de tamanho ótimo (coeficiente de adaptação). Qualquer componente independente com a Kurtosis não nula, pode ser extraído dessa maneira e sem a necessidade de suposições relativas a tipos específicos de fontes como reais, ou complexas, circular, não circular, sub-gaussiana ou super gaussiana.

A kurtosis, ou cumulante de quarta ordem, é definida conforme (3.29):

$$\mathcal{K} = \frac{E\{|y^4|\} - 2E^2\{|y^2|\} - |E\{y^2\}|^2}{E^2\{|y^2|\}} \quad (3.29)$$

onde E é o operador de esperança.

O RobustICA, em geral, possui as seguintes etapas:

1. Branqueamento dos dados obtendo a matriz branca das observações \mathbf{B} (nem sempre se efetua o branqueamento);
2. Calcular os coeficientes do polinômio de passo ótimo, para a kurtosis o polinômio de passo ideal é dado por (3.30):

$$p(\mathbf{w}_i) = \sum_{i=0}^4 \alpha_i \mu^i, \quad (3.30)$$

3. Extrair as raízes do polinômio de passo ótimo (PRESS *et al.*, 2007);
4. Selecionar a raiz que leva ao máximo absoluto de contraste ao longo da direção de busca:

$$\mu_{otm} = \arg \max_i |\mathcal{K}(\mathbf{w} + \mu_i \mathbf{g})| \quad (3.31)$$

5. Passo de atualização:

$$\mathbf{w} \leftarrow \mathbf{w} + \mu_{otm} \mathbf{g}, \quad (3.32)$$

onde \mathbf{g} é o gradiente de contraste de $\mathcal{K}(\cdot)$;

6. Normalização:

$$\mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\|; \quad (3.33)$$

Um maior detalhamento do algoritmo pode ser encontrado em (ZARZOSO; COMON, 2010) e (MEINECKE; HARMELING; MÜLLER, 2004).

A generalidade do contraste de Kurtosis garante que os sinais de valor real e complexo podem ser tratados exatamente pelo mesmo algoritmo sem qualquer modificação. Os tipos de sinais fonte podem estar presentes simultaneamente em uma dada mistura e fontes complexas não precisam ser circulares. Os coeficientes da matriz de mistura podem ser reais ou complexos, independentemente do tipo de fonte (ZARZOSO; COMON, 2010).

Ao contrário da maioria dos métodos ICA, o pré-processamento através do branqueamento não é necessário, de modo que as limitações de desempenho podem ser evitadas. A extração sequencial (deflação) pode ser realizada, por exemplo, por regressão linear. Essa característica pode ser especialmente benéfica em cenários mal condicionados, como o caso convolutivo (ZARZOSO; COMON, 2010).

O método mostra uma velocidade de convergência muito alta, medida em termos de qualidade de extração de fonte versus número de operações. Na faixa de dois sinais de valor real, o algoritmo converge em uma única iteração, mesmo sem pré-branqueamento (ZARZOSO; COMON, 2010).

3.4.4 Joint Approximate Diagonalization of Eigen-matrices (JADE)

Conforme (CARDOSO; SOULOUMIAC, 1993; HONG; KIM, 2015), o algoritmo *Joint Approximate Diagonalization of Eigen-matrices* (JADE) pode ser resumido nos seguintes passos:

1. Efetuar o branqueamento nos dados originais, conforme (3.34)

$$\bar{\mathbf{x}} = \mathbf{Q}\mathbf{x}, \quad (3.34)$$

onde \mathbf{Q} é a matriz de branqueamento;

2. Calcular os cumulantes de quarta ordem dos sinais transformados ($\bar{\mathbf{x}}$) e os n pares mais significativos de autovalores $N^e = \{\lambda_l, M_r | 1 \leq r \leq n\}$;
3. Efetuar a diagonalização do conjunto N^e pela matriz unitária \mathbf{U} ;
4. Estimar \mathbf{A} por $\hat{\mathbf{A}} = \mathbf{Q}\#\mathbf{U}$.

onde $\#$ é o operador de pseudo inversa.

Segundo (GURSOY, 2007), o algoritmo JADE trabalha na maximização da função objetivo de diagonalização dos elementos:

$$\text{maximizar } J(\mathbf{W}) = \sum_{i,j=1}^N |\text{diag}(\mathbf{W}\mathbf{M}_{ij}\mathbf{W}^T)|^2 \quad (3.35)$$

em que: neste caso \mathbf{W} é a matriz que diagonaliza todos os elementos da matriz de covariância \mathbf{C}_x cujo \mathbf{x} representa as misturas, \mathbf{M}_{ij} é a matriz definida por (3.36) e $|\text{diag}(\cdot)|^2$ é a soma dos quadrados dos elementos diagonais.

$$\mathbf{M}_{ij} = \text{Cum}(\mathbf{x}(ij)), \quad (3.36)$$

onde: $i, j = 1, 2, \dots, N$, pois \mathbf{M}_{ij} é uma matriz quadrada de dimensões $N \times N$.

3.5 ESTIMAÇÃO DE FONTES HARMÔNICAS DE CORRENTE EM UM SISTEMA TESTE

Para o mesmo sistema IEEE 14 barras utilizado anteriormente, na determinação do intervalo ótimo para o tamanho da janela do filtro, aplicou-se cada um dos algoritmos apresentados anteriormente: JADE, RobustICA, FastICA e SOBI.

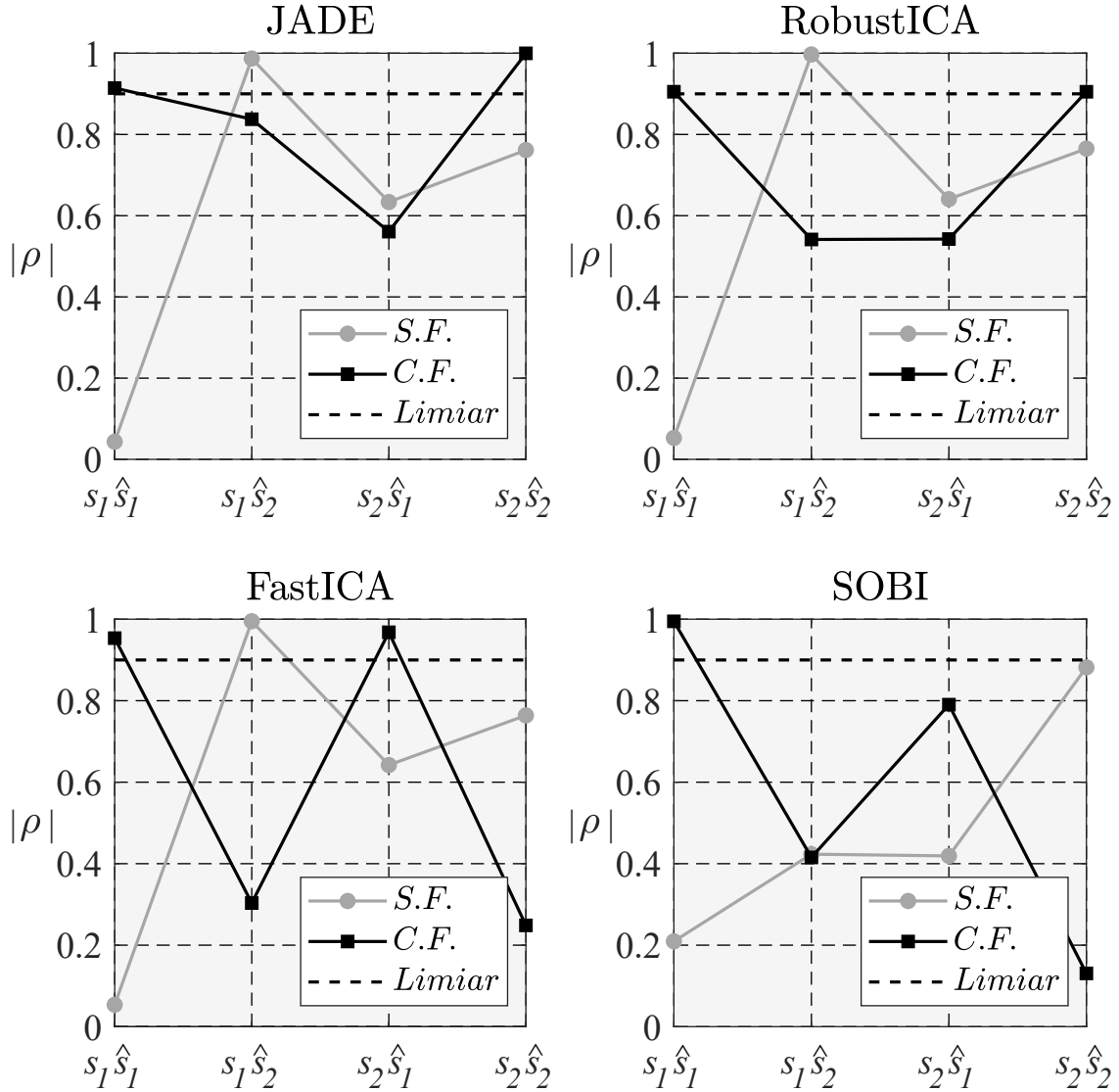
Neste tópico, o objetivo reside em avaliar o desempenho dos algoritmos na presença e na ausência de filtragem. Além disso, para determinar uma estimativa confiável das fontes harmônicas de corrente, definiu-se um limiar que corresponde a um coeficiente de correlação de 0.9, alta correlação, entre os sinais da fonte e estimativa como indicativo de uma estimação boa e aceitável para a fonte harmônica. A equação (3.37), define o cálculo do coeficiente de correlação para duas variáveis aleatórias, Neste caso, a fonte harmônica de corrente (\mathbf{s}) e estimativa ($\hat{\mathbf{s}}$):

$$\rho = \frac{\text{cov}(\mathbf{s}, \hat{\mathbf{s}})}{\sigma_s \sigma_{\hat{s}}} \quad (3.37)$$

Nota-se a superioridade dos desempenhos com o filtro (*C.F.*) em relação a ausência, sem o filtro (*S.F.*), conforme a Figura 11, e além disso, apenas com a filtragem é possível obter duas estimativas de alta correlação para os casos aceitáveis (RobustICA e JADE). Isso ocorre devido à maximização da independência entre as fontes devido à filtragem. Outro ponto importante é acerca do desempenho entre algoritmos, com a utilização da filtragem. Pode-se afirmar que o SOBI é capaz de estimar apenas uma única fonte (correspondência $s_1 \hat{s}_1$), apresentando um dos piores desempenhos. O FastICA também é inadequado para utilização devido a sua ambiguidade de estimação, isto é, uma mesma estimativa (\hat{s}_1) apresenta correspondência para duas fontes distintas (s_1 e s_2). Os desempenhos aceitáveis são obtidos via RobustICA e JADE, cuja estimação apresenta valores acima do limiar, e além disso, não ocorre dupla correspondência de uma única estimativa para duas fontes. Entretanto, ao observar mais atentamente, o desempenho do algoritmo JADE ainda é superior ao desempenho do RobustICA (segundo melhor resultado), apresentando um valor muito próximo de 1, para a correspondência $s_2 \hat{s}_2$, o que não ocorre para o RobustICA (*C.F.*). Logo, opta-se pelo algoritmo JADE por apresentar o melhor desempenho, na

presença da filtragem, entre todos os algoritmos testados. É válido salientar que esses valores são obtidos para uma análise na presença apenas do 5º harmônico.

Figura 11 – Correspondências via coeficientes de correlação

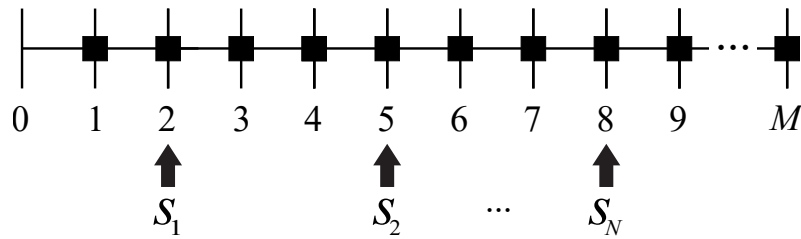


Fonte: Elaborado pelo autor (2020).

3.6 LOCALIZAÇÃO DE FONTES HARMÔNICAS DE CORRENTE

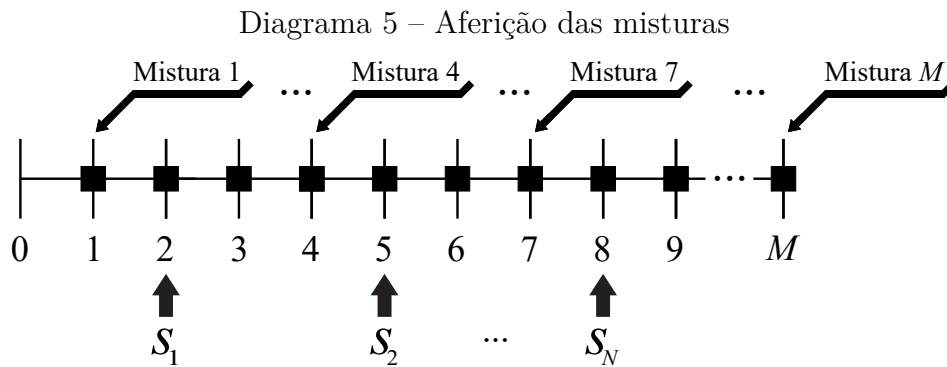
Imagine um sistema cuja topologia não seja conhecida, composto por N fontes harmônicas de corrente (s) e M medidores (\blacksquare), conforme é exposto na metodologia de estimação e é representado pelo Diagrama 4. O modelo representado pelo Diagrama 4 apresenta $M + 1$ barramentos, pois não faz sentido aplicar a metodologia de localização em barramentos de geração.

Diagrama 4 – Sistema elétrico modelo



Fonte: Elaborado pelo autor (2020).

Aferindo as tensões nos M barramentos, obtém-se as misturas, de acordo com o modelo de ICA e conforme o Diagrama 5.



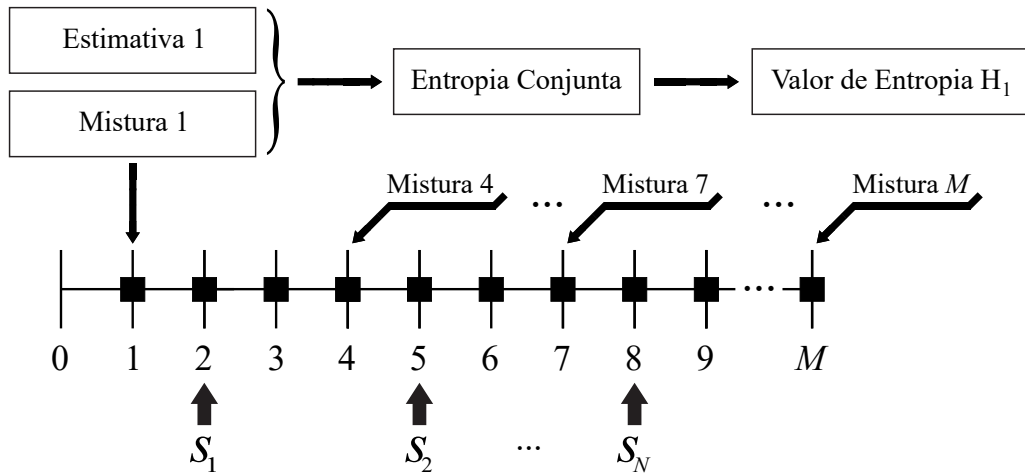
Fonte: Elaborado pelo autor (2020).

Em posse das estimativas das fontes harmônicas de corrente e medições de tensão, as partes reais e imaginárias de ambas são convertidas em vetores com valores de magnitudes. Por conseguinte, aplica-se a entropia conjunta entre os pares: magnitude da estimativa individual e magnitude da tensão aferida individual, conforme o Diagrama 6 e a equação (3.38):

$$H_{1,1}^h(X_1, \hat{S}_1) = - \sum_{l=1}^L \sum_{l=1}^L p(\mathbf{x}_{1,l}^h, \hat{\mathbf{s}}_{1,l}^h) \log p(\mathbf{x}_{1,l}^h, \hat{\mathbf{s}}_{1,l}^h) \quad (3.38)$$

cujos parâmetros são: $\hat{\mathbf{s}}_{1,l}^h$ o vetor de estimativa 1 e definido pela variável aleatória \hat{S}_1^h , $\mathbf{x}_{1,l}^h$ o vetor de mistura 1 e definido pela variável aleatória X_1^h , L é a amostra final dos vetores e h a ordem harmônica para qual está sendo feita análise.

Diagrama 6 – Processo de localização de fontes harmônicas de corrente



Fonte: Elaborado pelo autor (2020).

O processo é o mesmo para um outro vetor de estimativa, repetindo-se. Generalizando para N fontes e M medições de tensão, a metodologia de localização de fontes pode ser resumida para todo um SEP conforme a matriz (3.39):

$$\mathbf{H}^h = \begin{bmatrix} H_{1,1}^h & H_{1,2}^h & \cdots & H_{1,k}^h & \cdots & H_{1,m}^h \\ H_{2,1}^h & H_{2,2}^h & \cdots & H_{2,k}^h & \cdots & H_{2,m}^h \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ H_{k,1}^h & H_{k,2}^h & \cdots & H_{k,k}^h & \cdots & H_{k,m}^h \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ H_{n,1}^h & H_{n,2}^h & \cdots & H_{n,k}^h & \cdots & H_{n,m}^h \end{bmatrix} \quad (3.39)$$

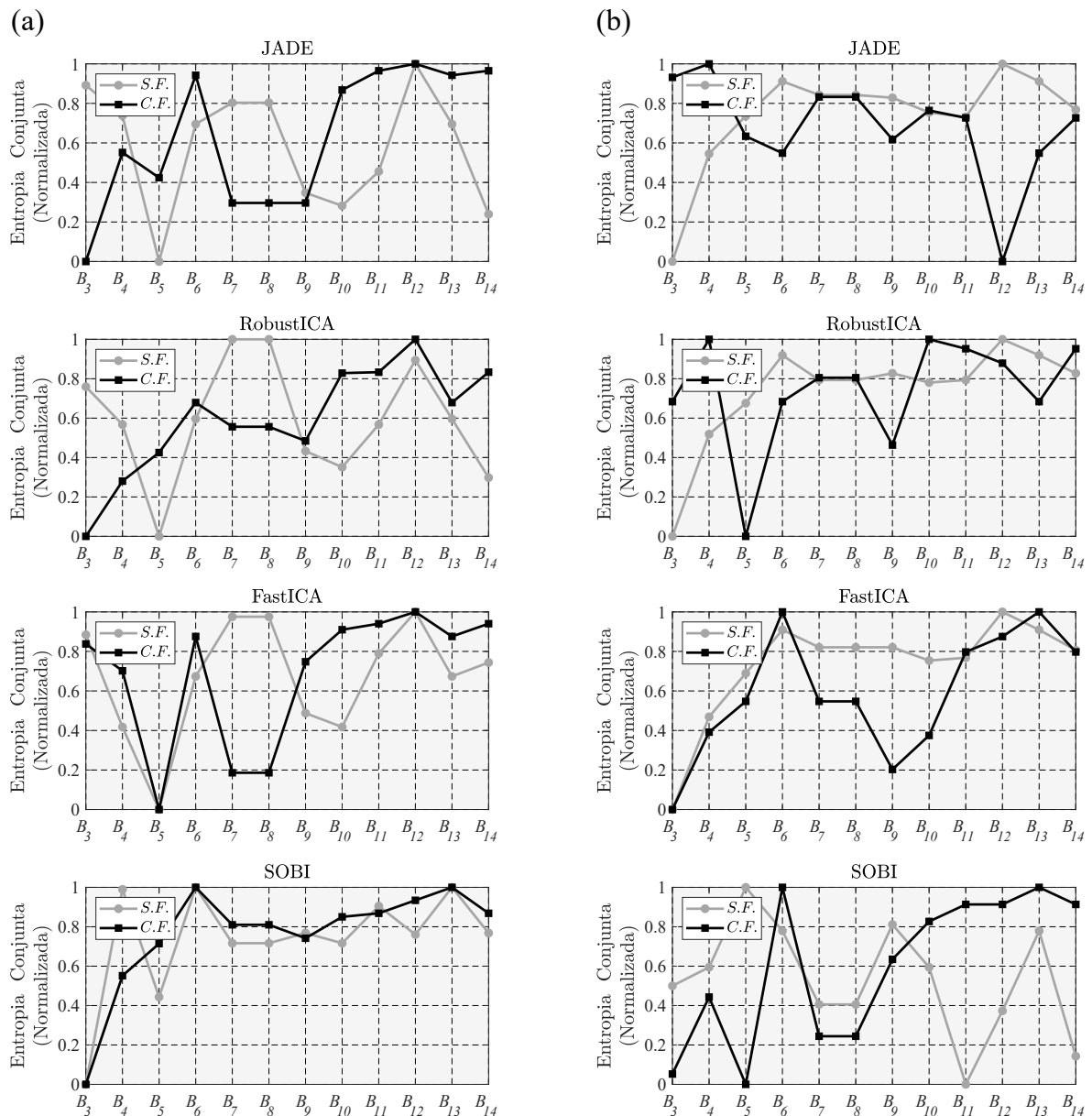
onde: \mathbf{H}^h é uma matriz composta por todos os valores computados de entropia conjunta na ordem harmônica de interesse h ; n é o número de estimativas e m é o número de medidores.

Essa maior probabilidade é apresentada pelo barramento cuja a entropia conjunta apresenta o valor mínimo. Isso é possível pois a entropia está relacionada a níveis de incerteza, por conseguinte, quanto menor for seu valor, menor será o nível de incerteza, ou será a maior dependência estatística entre as variáveis analisadas.

Diferentemente de (ZANG *et al.*, 2016), que utiliza a entropia condicional, a proposição do presente trabalho utiliza a entropia conjunta por ser mais generalista. A entropia condicional, assim como a probabilidade condicional, depende de uma condição a priori, a existência de uma fonte harmônica de corrente no sistema, premissa que não é necessária quando se utiliza a entropia conjunta.

Para o mesmo sistema IEEE 14 barras utilizado na seção (3.5), aplicou-se a entropia entre as estimativas e as tensões nos barramentos de interesse, produzindo o comparativo ilustrado na Figura 12. Neste caso, excetuou-se o primeiro barramento e também o segundo barramento, pois eram barramentos de geração. As cargas não lineares, modeladas como RCTs (FANDI, 1998), foram conectadas no barramento três (B_3) e no barramento doze (B_{12}).

Figura 12 – Desempenho da entropia conjunta na localização das fontes harmônicas de corrente no sistema teste



Fonte: Elaborado pelo autor (2020).

Ao observar os resultados ilustrados para a Figura 12 (a), estimativa 1 (fonte em B_3), e Figura 12 (b), estimativa 2 (fonte em B_{12}), é perceptível a necessidade da filtragem. Em todas as situações cujo filtro não foi utilizado, ocorreu erro na indicação do barramento

de conexão de uma das fontes. Esse resultado é esperado, visto que, perfis altamente correlacionados foram usados, da ordem de 80% de correlação entre os perfis lentos.

Outro ponto é em relação aos algoritmos. Os algoritmos SOBI e FastICA já apresentavam resultados ruins de estimação, então, resultados ruins de localização também eram esperados. O algoritmo RobustICA também não foi capaz de favorecer a etapa de localização e se comparado ao algoritmo JADE, não foi capaz de localizar efetivamente as fontes. Logo, optou-se pelo algoritmo JADE, pois foi o único algoritmo capaz de favorecer a etapa de localização, de modo que ambas as fontes fossem detectadas em suas respectivas barras de conexão, e também, foi o algoritmo que apresentou os valores mais altos para os coeficientes de correlação na etapa de estimação.

Seguindo o raciocínio apresentado pela matriz (3.39), é possível obter o conjunto de valores de entropias para o sistema IEEE 14 barras, de acordo com os pares de estimativas e conforme (3.41).

$$\mathbf{H}^h = \left[\begin{array}{c|c|c|c|c|c} H_{1,3}^h & H_{1,4}^h & \cdots & H_{1,12}^h & \cdots & H_{1,14}^h \\ \hline H_{2,3}^h & H_{2,4}^h & \cdots & H_{2,12}^h & \cdots & H_{2,14}^h \end{array} \right] \quad (3.40)$$

Assim, os mínimos obtidos estão contidos em (3.41):

$$H_{min}^h = \{H_{1,3}^h, H_{2,12}^h\} \quad (3.41)$$

3.7 CONSIDERAÇÕES FINAIS DO CAPÍTULO

O presente capítulo apresentou em detalhes as ferramentas que formam a metodologia de localização de fontes harmônicas. Elencou-se a etapa de estimação, bem como os algoritmos testados e a necessidade da utilização de um filtro na independência estatística. Além disso, os algoritmos foram testados no aspecto da qualidade de estimação, via coeficiente de correlação, provando o desempenho superior do algoritmo JADE na estimação das fontes, neste caso, apenas com a inserção do 5º harmônico no sistema. Ainda Neste contexto, a entropia conjunta foi discutida no ponto de vista dos algoritmos, para finalmente, definir o algoritmo de ICA mais apropriado para a metodologia, ainda com a injeção do 5º harmônico apenas. Finalmente, o processo de localização é formalizado através da indicação de mínimos, ou seja, os valores mínimos de entropia conjunta equivalem à maior probabilidade de uma fonte harmônica de corrente estar conectada ao barramento.

4 RESULTADOS

Para este capítulo, foram efetuadas simulações via *software* MATLAB com o propósito de validar a metodologia apresentada. Neste âmbito, a metodologia foi aplicada em dois SEPs, um IEEE 14 barras (VARIZ *et al.*, 2008) e um IEEE 33 barras (MELO, 2018). A utilização desses dois sistemas reside em verificar a robustez do método em sistemas de naturezas distintas, visto que, o IEEE 14 barras é um sistema de transmissão, ao passo que, o sistema IEEE 33 barras é um sistema de distribuição.

Para os testes com IEEE 14 barras, níveis diferentes de coeficiente de correlação foram utilizados, desde baixa correlação até o cenário de altíssima correlação, buscando verificar a necessidade e o ponto o qual a filtragem torna-se uma ferramenta essencial à metodologia. Neste aspecto, um ensaio mais acurado quanto ao intervalo de confiança é efetuado. Além disso, testes com modelos diferentes para a variação rápida foram efetuados como verificação da robustez do trabalho. Finalmente, testes trocando as posições das fontes harmônicas de corrente foram realizados para verificar o desempenho da metodologia em situações diversas.

Para o teste com o sistema IEEE 33 barras, três fontes harmônicas de corrente foram utilizadas como um cenário de *stress* para a metodologia, visto que, quanto maior a quantidade de fontes harmônicas de corrente, maior é a dificuldade em separá-las.

4.1 MODELOS UTILIZADOS NA BASE DE DADOS

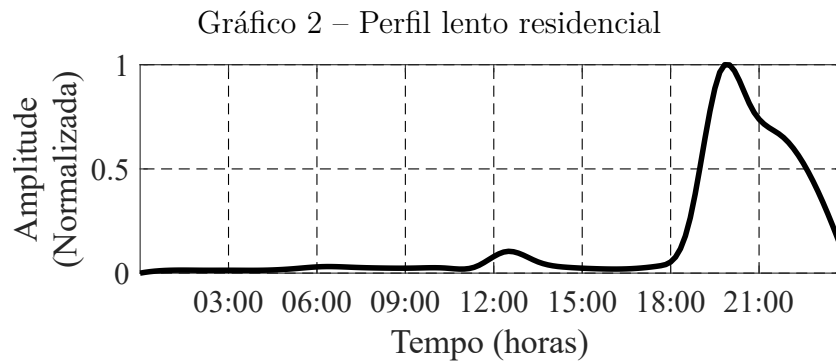
É válido salientar que a modelagem não é o foco deste trabalho, entretanto, é necessária para promover as condições necessárias para que a eficácia da metodologia seja comprovada.

4.1.1 Curvas de Carga e Perfis de Carga

Para a Agência Nacional de Energia Elétrica (ANEEL), a curva de carga é definida como registro horário, em um período diário, das demandas de capacidade, podendo ser, excepcionalmente para um período semanal, mensal ou anual (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, 2012). Saber o comportamento de uma curva de carga possibilita uma classificação mais acurada dos consumidores, possibilitando inferir o comportamento de consumo dos mesmos num intervalo de tempo finito. Além disso, as curvas de cargas podem ser endereçadas a diferentes naturezas. Existem curvas de cargas que podem caracterizar um consumidor como: rural, residencial, comercial, industrial, serviço público ou iluminação pública (QUEIROZ, 2011). Em geral, as curvas de carga são apresentadas em termos de potência. Entretanto, análises semelhantes podem ser extraídas via correntes harmônicas ao invés de se utilizar potências complexas, visto que, a potência complexa é uma função da corrente (GURSOY, 2007). O presente trabalho opta pela representação

em perfis de corrente pois não prejudica a análise, conforme supracitado, e torna mais intuitivo o entendimento da metodologia de injeção de fontes harmônicas de corrente. Neste caso, os modelos que utilizam corrente como unidade serão chamados de perfis de carga, para diferenciá-los da curva de carga que utiliza potência como unidade.

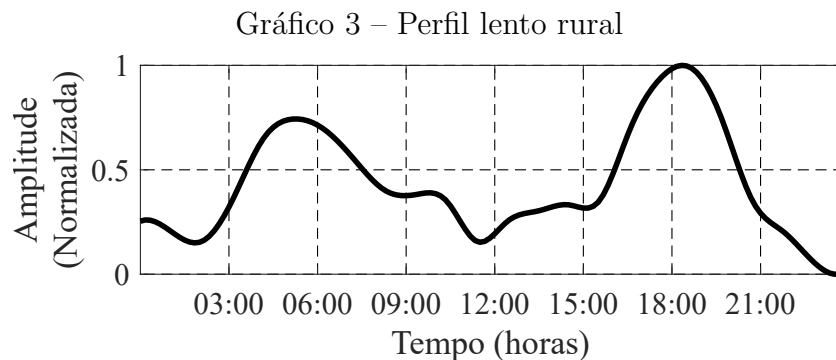
Um exemplo de perfil de carga é o modelo residencial (FRANCISQUINI, 2006; QUEIROZ, 2011) e é, guardadas as devidas proporções, representado pelo Gráfico 2. Neste ponto as variações diárias ainda não serão consideradas por motivos didáticos, mas as mesmas serão discutidas mais adiante.



Fonte: Elaborado pelo autor (2020).

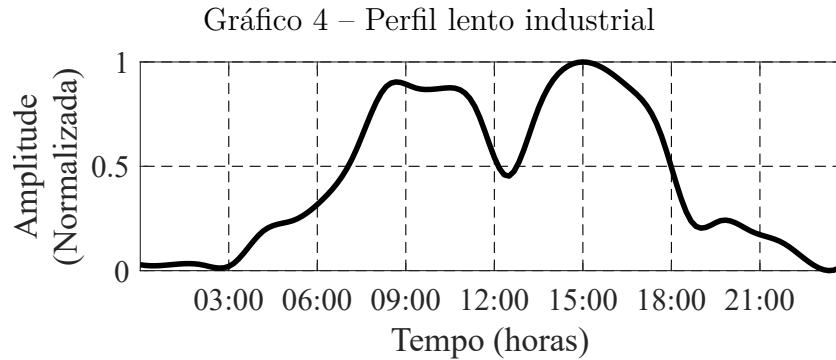
Um consumidor residencial pode ser caracterizado por uma demanda suave e quase constante durante o dia, e por fim, um aumento no consumo entre 18 e 21 horas (FRANCISQUINI, 2006). Este aumento da demanda é ocasionado pelo uso intenso do chuveiro elétrico e equipamentos característicos residenciais (FRANCISQUINI, 2006).

Da mesma forma, isto é, guardadas as devidas proporções, é possível elencar o consumidor rural (MARQUES, 2014), conforme o Gráfico 3. Nessa classe, há uma elevação do consumo no começo da manhã e um pico em torno das 18 horas. Este tipo de consumidor é caracterizado pelas atividades agrícolas e pecuárias.



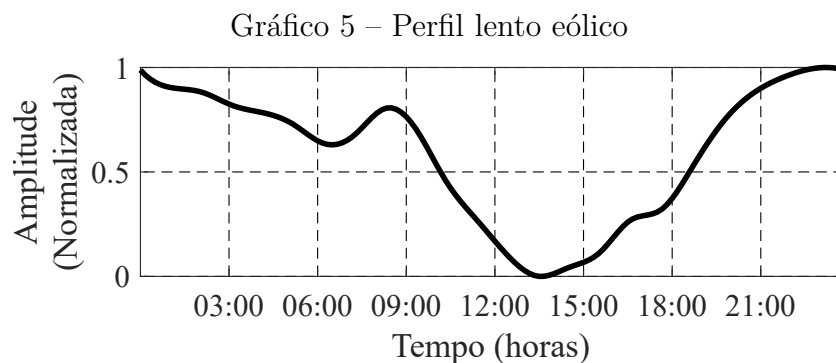
Fonte: Elaborado pelo autor (2020).

Outro perfil utilizado é o industrial (MELO *et al.*, 2018; FRANCISQUINI, 2006). Neste caso, consumo elevado no período da manhã próximo de 8 até 11 horas, e também, aproximadamente no período da tarde entre 14 e 17 horas (FRANCISQUINI, 2006). Tal perfil está apresentado pelo Gráfico 4.



Fonte: Elaborado pelo autor (2020).

Finalmente, um perfil que representa a atividade eólica está ilustrado no Gráfico 5. Tal perfil segue o comportamento para o subsistema do Nordeste (OPERADOR NACIONAL DO SISTEMA ELÉTRICO, 2018), presente no Boletim Mensal de Geração Eólica de Agosto de 2018 do Operador Nacional do Sistema Elétrico (ONS). A atividade eólica apresenta maior intensidade no começo do dia e ao final do mesmo, para essa região. Embora o perfil eólico seja um perfil de geração, ele é apresentado porque os geradores eólicos podem injetar correntes harmônicas no SEP.



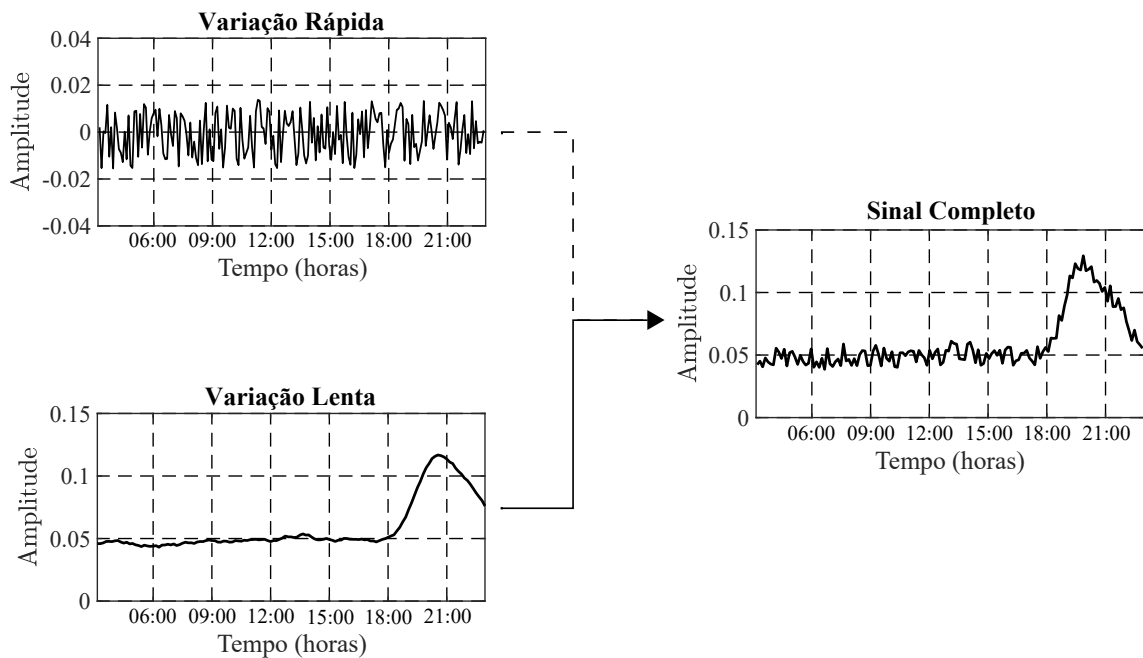
Fonte: Elaborado pelo autor (2020).

4.1.2 Independência Estatística do Perfil de Carga

A demanda de potência de uma carga não é constante ao longo do dia, semana ou por um período maior de observação. Temperatura, dia da semana, período do dia, são parâmetros que podem influenciar a demanda de potência de uma carga. Essas adversidades são responsáveis por criarem a variação lenta do perfil, denominada de *slow variation*.

A variação lenta é responsável por promover flutuações de período prolongado, por exemplo, períodos de horas. Em contrapartida, flutuações abruptas de segundos ou até mesmo minutos, são denominadas de variações rápidas do perfil, ou *fast variation*. Essas variações são difíceis de prever e são modeladas como flutuações aleatórias. Neste caso, é sugerido uma modelagem com variância em torno de 0,002 para uma distribuição laplaciana (GURSOY, 2007). Além disso, estudos de monitoramento de longo prazo de tensões e correntes harmônicas foram efetuados no Reino Unido (DUGGAN; MORRISON, 1991; MORRISON; DUGGAN, 1988; MORRISON; DUGGAN, 1993), que por conseguinte, apresentaram análises em termos de flutuações rápidas e lentas dos perfis. E também, os sinais podem ser tratados como a soma de uma parcela determinística e outra aleatória (BAGHZOUZ *et al.*, 1998; CAVALLINI; MONTANARI; CACCIARI, 1995), conforme a Figura 13:

Figura 13 – Perfil lento eólico



Fonte: Elaborado pelo autor (2020).

A independência estatística reside nas parcelas aleatórias do perfil, ou flutuações rápidas. Além disso, do ponto de vista prático, a independência estatística tem a ver com as particularidades pontuais de cada fonte, mesmo as fontes sendo de um mesmo tipo. Matematicamente, isso significa que, por exemplo para duas fontes harmônicas de corrente (y_1 e y_2), a distribuição conjunta de probabilidade ($p_{y_1 y_2}$) entre ambas é equivalente à multiplicação das distribuições individuais de probabilidade (p_{y_1} e p_{y_2}), conforme (4.1). Generalizando para N fontes harmônicas de corrente, obtém-se (4.2).

$$p_{y_1 y_2}(y_1, y_2) = p_{y_1}(y_1)p_{y_2}(y_2) \quad (4.1)$$

$$p_{y_1 y_2 \dots y_n}(y_1, y_2, \dots, y_n) = p_{y_1}(y_1)p_{y_2}(y_2), \dots, p_{y_n}(y_n) \quad (4.2)$$

4.1.3 Injeção Harmônica

Inicialmente, a fonte harmônica de corrente é modelada em concordância com a seção 4.1.2, ou seja, flutuações lentas e rápidas são somadas de forma a compor um perfil de fonte harmônica de corrente que represente os esforços ao longo de um dia e esforços ao longo de horas, respectivamente. Então, para um perfil lento Γ , adiciona-se uma flutuação rápida γ , de acordo com (4.3):

$$\Gamma[n] + \gamma[n] = T[n], \quad (4.3)$$

onde T é o perfil completo de variação diária e n é o índice da amostra, ou seja, se cada amostra n está sendo computada a cada hora, em cada uma dessas amostras é gerada uma variação aleatória de acordo com um tipo de distribuição fixo. Neste caso, a modelagem da flutuação rápida pode ser definida por uma distribuição uniforme, por exemplo.

Logo, para N fontes harmônicas e suas respectivas variações, obtém-se (4.4):

$$\begin{array}{c} T_1[n] \\ T_2[n] \\ \vdots \\ T_N[n] \end{array} \quad (4.4)$$

Entretanto, ainda é necessário anexar espectros harmônicos aos perfis de carga estabelecidos. Assim, suponha que determinado tipo de carga α_1 , por exemplo um RCT, possua um espectro harmônico com magnitude e fase, para uma ordem harmônica específica h , definidos por (4.5):

$$A_1(h) = |\alpha_1|^h + \angle\alpha_1^h \quad (4.5)$$

Para caracterizar os espectros de frequência gerados pela carga α_1 ao perfil da mesma, é necessário efetuar uma multiplicação entre perfil e as ordens harmônicas desejadas a serem injetadas na rede. Além do mais, como o espectro harmônico é de natureza complexa, é necessário fazer uma consideração ao perfil de forma que seja possível essa multiplicação. Para tal, optou-se por considerar a parte real equivalente à parte imaginária para a variação lenta, ou seja, as partes reais e imaginárias diferem-se na parcela equivalente à variação rápida. Logo, através de (4.5) e (4.6), é possível estabelecer (4.7):

$$T_{\alpha_1} = T'_{\alpha_1} + jT''_{\alpha_1} \quad (4.6)$$

onde: T_{α_1} é um perfil de carga para uma fonte harmônica de corrente qualquer, composto pela parte real T'_{α_1} e imaginária T''_{α_1} , associadas ao tipo de carga α_1 .

$$T_{A_1} = T_{\alpha_1} \cdot A_1(h), \quad (4.7)$$

onde: T_{A_1} é um perfil de carga para uma fonte harmônica de corrente qualquer, cujo A_1 representa o índice que identifica o conjunto de espectros associados à mesma.

Finalmente, os perfis harmônicos cuja ordem harmônica deseja-se injetar são escolhidos para se inserir na rede. Neste trabalho, os espectros utilizados estão presentes na Tabela 1. Os espectros do HVDC e do RCT são baseados no trabalho proposto por Gursoy (2007). Além disso, utiliza-se um retificador e um inversor (GALHARDO; PINHO, 2008). Finalmente, utiliza-se um espectro que representa atividade eólica (PEREIRA, 2019).

Tabela 1 – Espectro harmônico das cargas não lineares

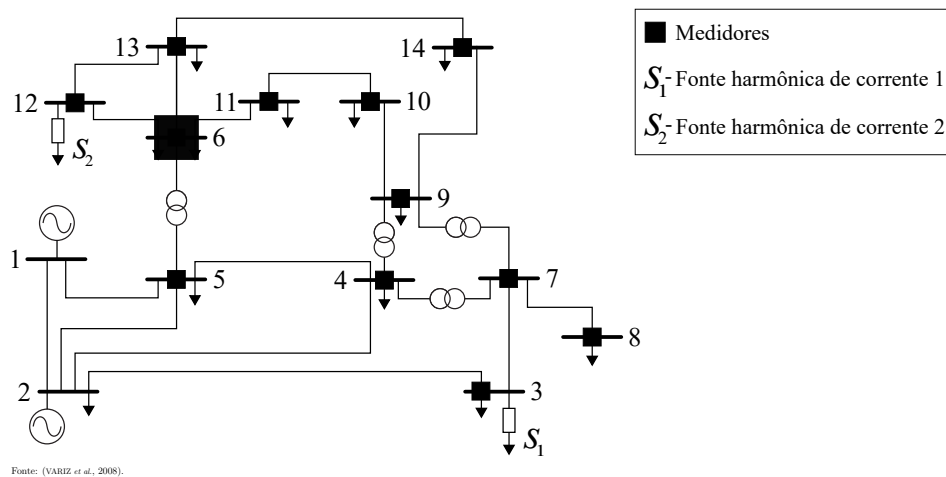
HVDC		
h	Mag. (%)	Fase $\theta(^{\circ})$
1	100,00	-49,56
5	19,41	-67,77
7	13,09	11,90
11	07,58	-7,13
RCT		
h	Mag. (%)	Fase $\theta(^{\circ})$
1	100,00	46,92
5	07,02	-124,40
7	02,50	-29,87
11	01,36	23,75
Retificador		
h	Mag. (%)	Fase $\theta(^{\circ})$
1	100,00	-100,00
5	32,00	-146,00
7	08,00	-37,00
11	05,90	152,00
Inversor		
h	Mag. (%)	Fase $\theta(^{\circ})$
1	100,00	56,00
5	07,02	-56,00
7	02,50	-130,00
11	01,36	83,00
Eólico		
h	Mag. (%)	Fase $\theta(^{\circ})$
1	100,00	0,00
5	00,09	89,78
7	00,07	89,51
11	00,02	87,43

Fonte: Elaborada pelo autor (2020).

4.2 RESULTADOS PARA O SISTEMA 14 BARRAS

Os resultados seguintes contemplam as análises com o sistema IEEE 14 barras adaptado (VARIZ *et al.*, 2008). Tal sistema é composto por 14 barras, 4 transformadores e 2 geradores, os quais, estão dispostos nas barras 1 e 2, conforme a Figura 14. Embora, a representação de todos os componentes da rede seja a de elementos trifásicos, para a metodologia é necessário apenas a tensão de uma das fases. Neste caso, como o sistema é equilibrado, para a aplicação do método qualquer uma das fases bastaria, logo, utiliza-se a fase A. Mais informações sobre o referido sistema podem ser encontradas no trabalho proposto em Variz *et al.* (2008).

Figura 14 – Sistema IEEE 14 barras



Logo, as seções a seguir irão abordar os testes efetuados com o sistema IEEE 14 barras:

4.2.1 Perfis lentos com $|\rho| = 37\%$ de correlação.

No presente ensaio, um perfil residencial (fonte 1) (FRANCISQUINI, 2006; QUEIROZ, 2011), e um industrial (fonte 2) (MELO *et al.*, 2018; FRANCISQUINI, 2006), foram utilizados, ambos de acordo com o Quadro 2. A esses perfis lentos, foram adicionados perfis rápidos modelados como distribuições normais, cuja ocorrência está contida no intervalo $[0, 0,02]$.

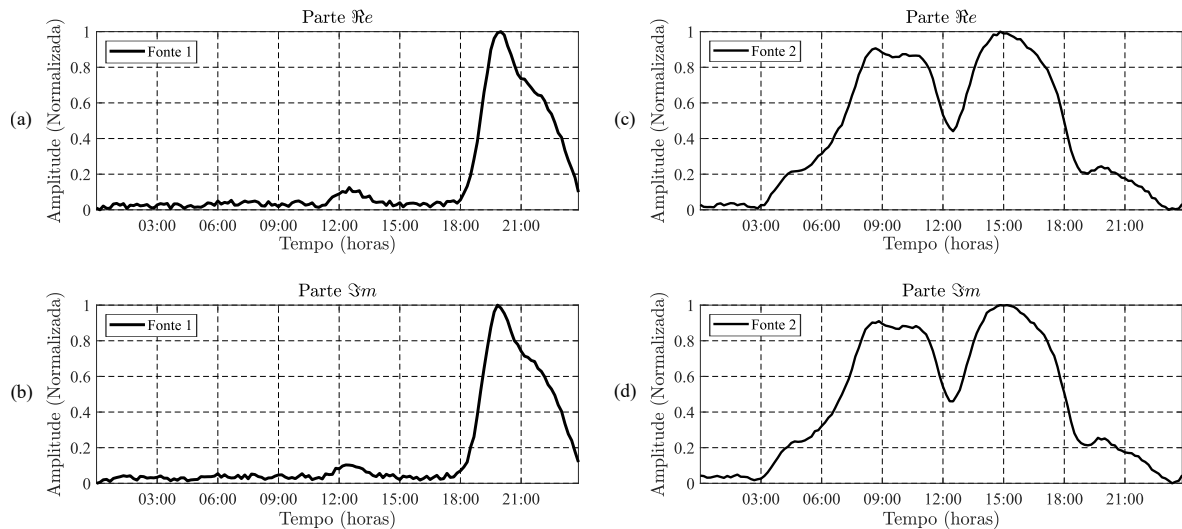
Quadro 2 – Perfis e espectros utilizados nas cargas para $|\rho| = 37\%$

Fonte 1		Fonte 2	
Perfil	Espectro	Perfil	Espectro
Residencial	Retificador	Industrial	Inversor

Fonte: Elaborado pelo autor (2020).

A fonte 1 foi conectada ao barramento 3, ao passo que, a fonte 2 ao barramento 12. A Figura 15 apresenta as partes reais e imaginárias que foram injetadas no sistema. Neste caso, a Figura 15 (a) é a parte real do perfil residencial, a Figura 15 (b) é a parte imaginária do perfil residencial, a Figura 15 (c) é parte real do perfil industrial e a Figura 15 (d) é parte imaginária do perfil industrial.

Figura 15 – Perfis injetados para $|\rho| = 37\%$



Fonte: Elaborado pelo autor (2020).

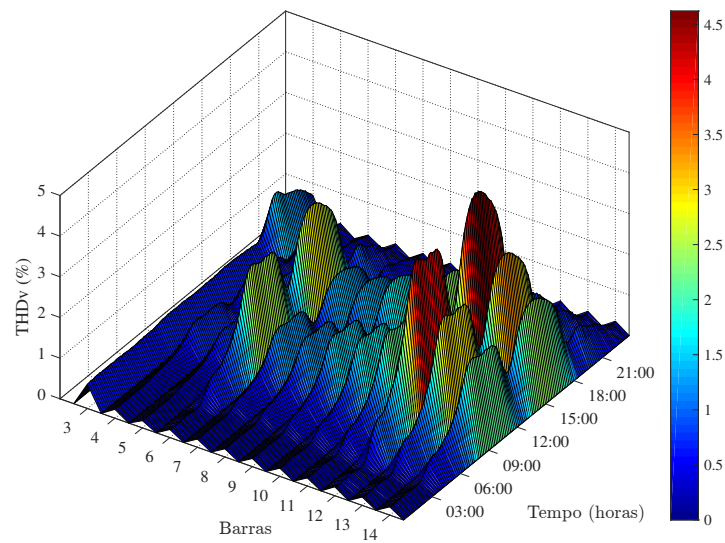
A Tabela 2 apresenta os espectros nas barras com fonte harmônica de corrente, após a execução do fluxo de potência harmônico, em que P representa o valor de pico e M o valor médio.

Tabela 2 – Espectro nas barras com carga não linear para $|\rho| = 37\%$

h	Barra 3 (S_1)		Barra 12 (S_2)	
	Mag. $_P$ (%)	Mag. $_M$ (%)	Mag. $_P$ (%)	Mag. $_M$ (%)
1	100,0000	100,0000	100,0000	100,0000
5	1,8413	1,3755	20,8531	17,0731
7	0,3061	0,0811	14,0371	11,3648
11	0,1609	0,1593	1,1684	0,9801

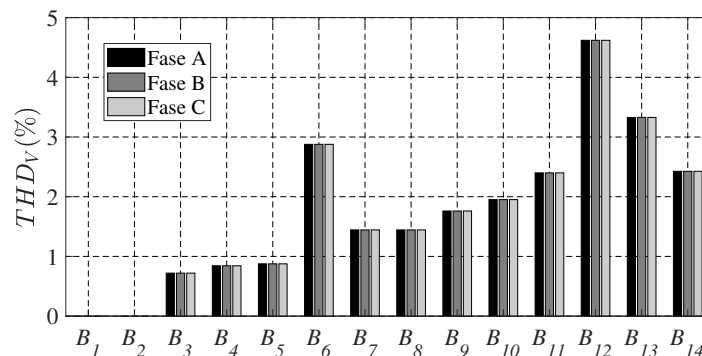
Fonte: Elaborada pelo autor (2020).

Como a metodologia utiliza tensões harmônicas, o caminho mais lógico e intuitivo seria medir a Distorção Harmônica Total (DHT), do inglês *Total Harmonic Distortion* (THD), de Tensão. Neste caso, a sigla THD será utilizada e acrescida da letra v, para esse teste e os subsequentes, indicando o THD de tensão. Neste aspecto, a Figura 6 apresenta a evolução do THDv ao longo do dia.

Gráfico 6 – THD de tensão ao longo do dia para o caso de $|\rho| = 37\%$ 

Fonte: Elaborado pelo autor (2020).

Ao observar o Gráfico 6, são perceptíveis os máximos de distorção nos horários de pico, por exemplo, entre 9 e 12 horas, e também, 15 e 18 horas, para o 12º barramento. Além disso, existe a violação de 3% do THD de tensão (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, 2016), para o referido barramento, o que justificaria uma medição para a verificação de uma carga poluente conectada no referido barramento. Essa análise está exemplificada na Figura 7 e salienta que a região de maior distorção é a que comporta a fonte harmônica 2 (conectada ao barramento 12), no horário de pico de 15 horas. Além disso, essa região sofre grande interferência de forma que o limite de 3% seja extrapolado nos barramentos vizinhos como o 13. Entretanto, nada pode se afirmar em relação à fonte harmônica 1, conectada ao barramento 3.

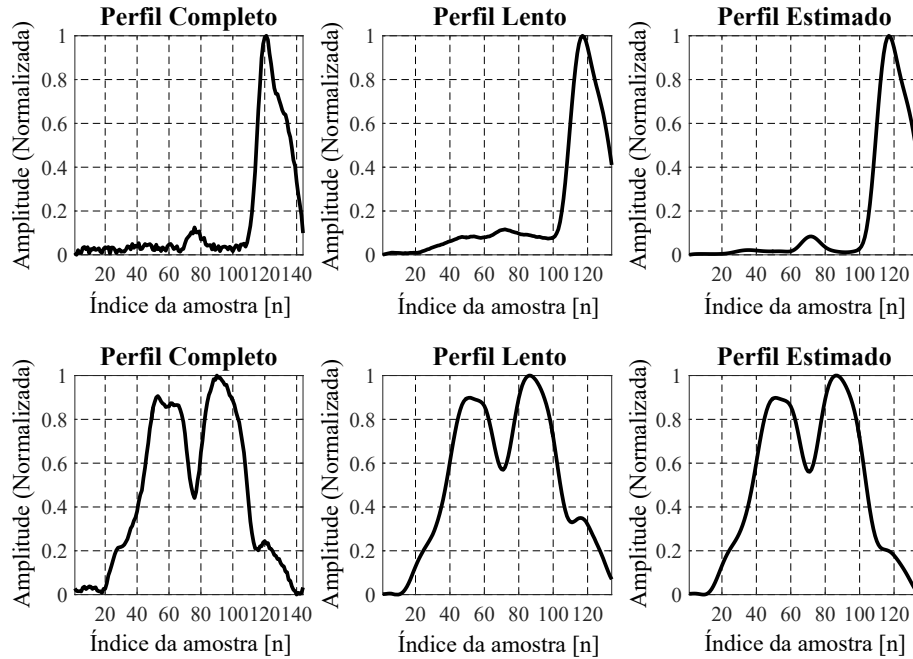
Gráfico 7 – THD de tensão no horário de pico para o caso de $|\rho| = 37\%$ 

Fonte: Elaborado pelo autor (2020).

Como não é possível detectar a fonte da barra 3 apenas verificando o THDv, tanto

ao longo do dia quanto pontualmente, aplica-se a metodologia de detecção. A Figura 16 apresenta o que realmente é estimado na primeira etapa da metodologia via ICA, salientando que a etapa de estimação busca encontrar a forma do perfil lento.

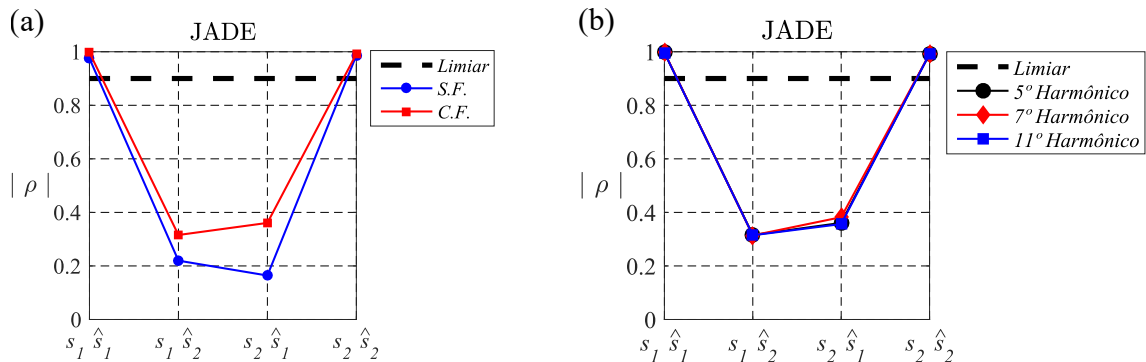
Figura 16 – Etapa de estimação para o caso de $|\rho| = 37\%$



Fonte: Elaborado pelo autor (2020).

A Figura 17 (a) mostra a correspondência entre as fontes (s) e estimativas (\hat{s}), para a variação lenta. Os altos valores de $|\rho|$ entre s_1, \hat{s}_1 e s_2, \hat{s}_2 , acima do limiar de 0,9 (90%), indicam que a ICA foi capaz de estimar bem os perfis desejados. Para as ordens harmônicas 5, 7 e 11, observa-se que o comportamento das estimativas é recorrente, conforme a Figura 17 (b), ou seja, de acordo com os fundamentos teóricos, não importa a ordem harmônica, pois trata-se de um método estatístico de estimação.

Figura 17 – Correspondência via $|\rho|$ para o primeiro caso

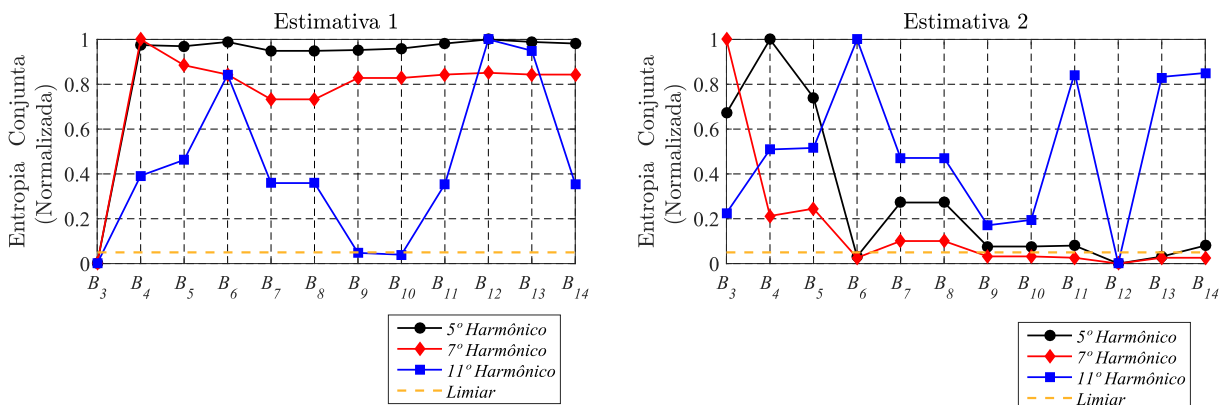


Fonte: Elaborado pelo autor (2020).

Além disso, como os perfis são pouco correlacionados, o impacto da filtragem é muito pequeno, visto que, há pouca dependência a ser minimizada pela filtragem. Porém, mantém-se a filtragem conforme a metodologia.

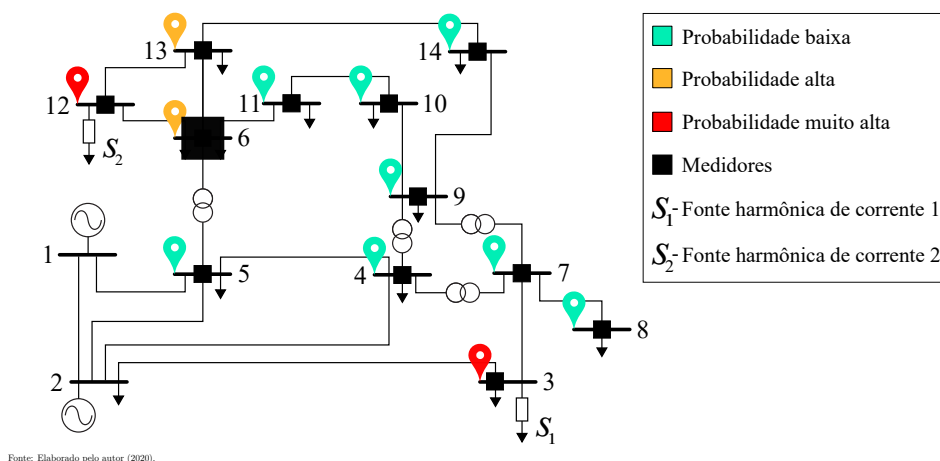
Após a estimação, obtém-se os resultados de entropia conjunta, conforme a Figura 18. Para uma análise mais acurada, um limiar de 5% é estabelecido. Então, quando para três ordens harmônicas distintas as entropias estão abaixo desse limiar, a probabilidade de conexão da fonte é tratada como muito alta. Quando apenas para duas ordens harmônicas as entropias estão abaixo do limiar, essa probabilidade é considerada alta. Quando a entropia para um harmônico está abaixo desse limiar, ou para os três harmônicos a entropia está acima, a probabilidade é considerada baixa, conforme a Figura 19. Essa interpretação se reflete no resultado 18, onde para a estimativa 1, as entropias correspondentes para as três ordens harmônicas estão abaixo do limiar apenas na barra 3. E também, a mesma observação é percebida na barra 12.

Figura 18 – Resultados de localização dos harmônicos para o caso de $|\rho| = 37\%$



Fonte: Elaborado pelo autor (2020).

Figura 19 – Mapa para as localizações para o caso de $|\rho| = 37\%$



Fonte: Elaborado pelo autor (2020).

4.2.2 Perfis lentos com $|\rho| = 75\%$ de correlação.

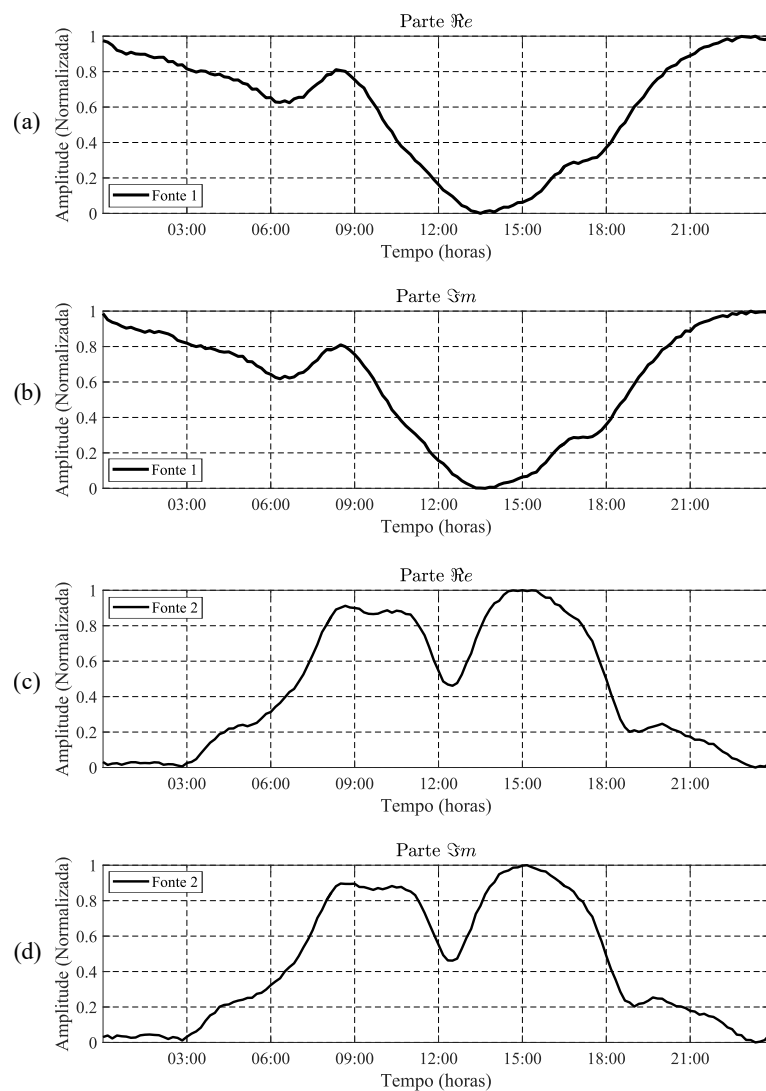
Para este ensaio, um perfil eólico (fonte 1) (OPERADOR NACIONAL DO SISTEMA ELÉTRICO, 2018), e um industrial (fonte 2) (MELO *et al.*, 2018; FRANCISQUINI, 2006), foram utilizados, ambos de acordo com o Quadro 3, para a variação lenta. A esses perfis lentos, foram adicionados perfis rápidos modelados como distribuições normais, novamente, com intervalo $[0; 0,02]$. A fonte 1 foi conectada ao barramento 3 e a fonte 2 ao barramento 12. A Figura 20 apresenta os perfis que foram injetados.

Quadro 3 – Perfis e espectros utilizados nas cargas para $|\rho| = 75\%$

Fonte 1		Fonte 2	
Perfil	Espectro	Perfil	Espectro
Eólico	Eólico	Industrial	HDVC

Fonte: Elaborado pelo autor (2020).

Figura 20 – Perfis injetados para $|\rho| = 75\%$

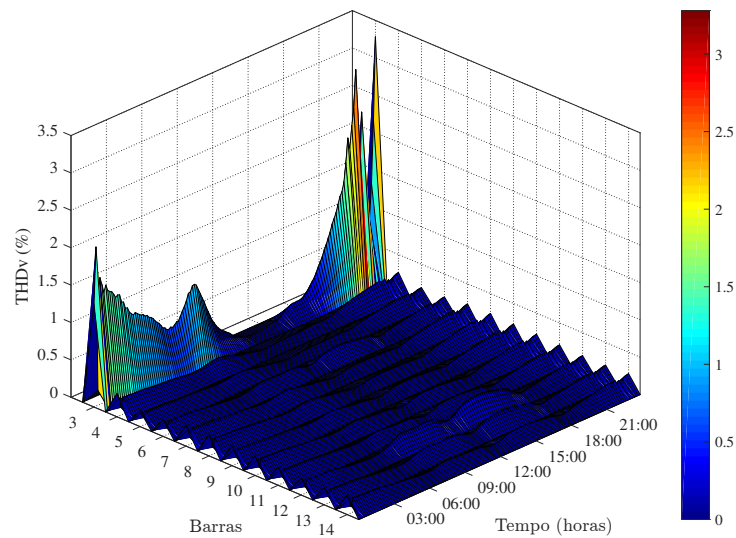


Fonte: Elaborado pelo autor (2020).

A Figura 20 (a) é parte real do perfil eólico, a Figura 20 (b) é a parte imaginária do perfil eólico, a Figura 20 (c) é a parte real do perfil industrial e a Figura 20 (d) é a parte imaginária do perfil eólico.

O Gráfico 8 apresenta a evolução do THDv ao longo do dia com a utilização dos perfis distintos. Neste caso, a violação de 3% é observada apenas na barra 3, onde a fonte 1 está conectada.

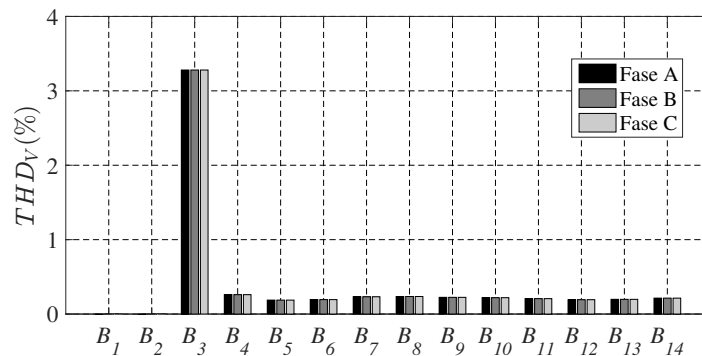
Gráfico 8 – THD de tensão ao longo do dia para o caso de $|\rho| = 75\%$



Fonte: Elaborado pelo autor (2020).

Verificando o THDv às 00 : 00 horas, horário de pico, fica evidente que a segunda fonte harmônica de corrente (barra 12) não pode ser detectada com uma simples medição de distorção harmônica.

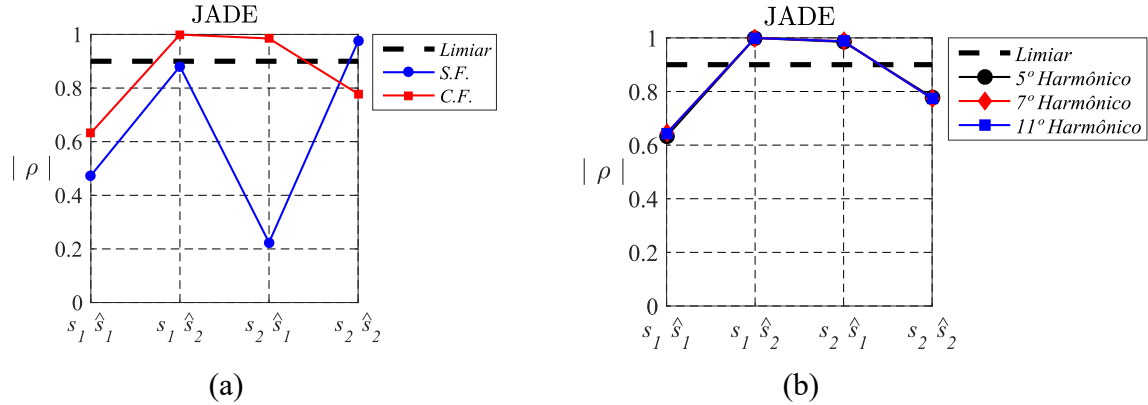
Gráfico 9 – THD de tensão no horário de pico para o caso de $|\rho| = 75\%$



Fonte: Elaborado pelo autor (2020).

O resultado apresentado na Figura 21 (a) demonstra a necessidade da filtragem para perfis onde a correlação começa a ficar elevada.

Figura 21 – Correspondência via $|\rho|$ para o segundo caso



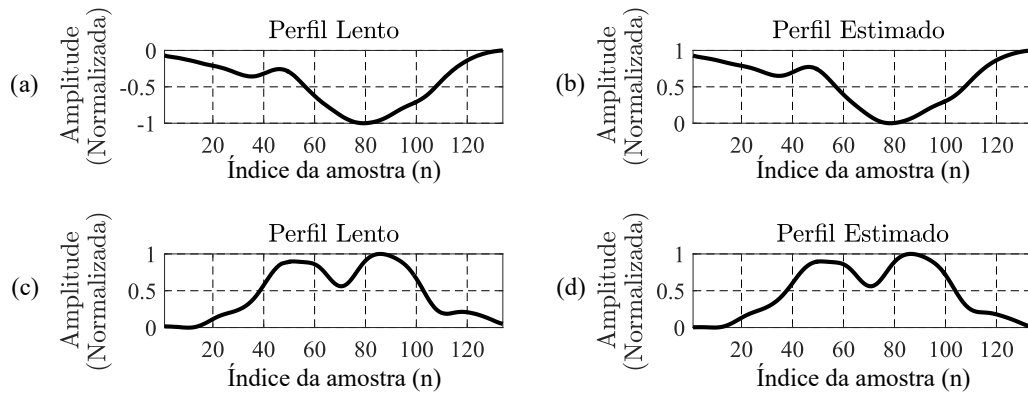
Fonte: Elaborado pelo autor (2020).

Neste caso, quando não se utiliza a filtragem (*S.F.*), existem duas correspondências. A primeira correspondência é obtida entre s_1 e \hat{s}_2 , abaixo do limiar. A segunda entre s_2 e \hat{s}_2 , acima do limiar. Porém, esse é um caso de dupla correspondência, ou seja, uma mesma estimativa apresenta similaridade para as duas fontes harmônicas. Logo, quando o filtro não é utilizado a separabilidade entre as fontes fica comprometida. Em contrapartida, quando a filtragem é utilizada (*C.F.*), os resultados de correspondência via $|\rho|$ apresentam valores altíssimos e a dupla correspondência se desfaz. Neste caso, a estimativa 2 recupera com sucesso a fonte 1 e a estimativa 1 recupera com sucesso a fonte 2. A questão da filtragem no âmbito do intervalo de confiança será tratada com uma maior atenção no próximo ensaio, em que, a correlação entre os perfis é da ordem de 80%, ou seja, uma situação ainda mais difícil para o processo de estimação. No presente caso, opta-se por manter a janela de $N = 11$, posto que, ainda está contida no intervalo de confiança, conforme os fundamentos teóricos.

Além disso, para as três ordens harmônicas estudados, os resultados apresentam a mesma tendência de comportamento. Conforme a Figura 21 (b), é válido salientar a permutação de ordens das estimativas, isto é, a estimativa 1 (\hat{s}_1) apresenta correspondência para a fonte 2 (s_2), ao mesmo tempo que, a estimativa (\hat{s}_2) apresenta correspondência para a fonte 1 (s_1).

A Figura 22 ilustra os resultados oriundos da etapa de estimação, em que, a Figura 22 (a) é o perfil lento da fonte 1, a Figura 22 (b) é o perfil lento estimado da fonte 1, a Figura 22 (c) é o perfil lento da fonte 2 e a Figura 22 (d) é o perfil lento estimado da fonte 2. Como as correspondências de $|\rho|$ foram altíssimas, é natural que os perfis estimados sejam muito parecidos com os originais.

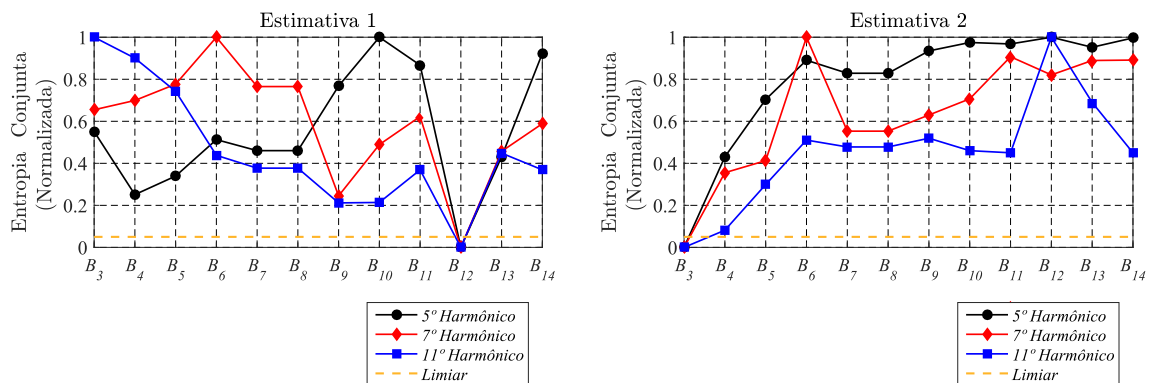
Figura 22 – Etapa de estimação para o caso de $|\rho| = 75\%$



Fonte: Elaborado pelo autor (2020)

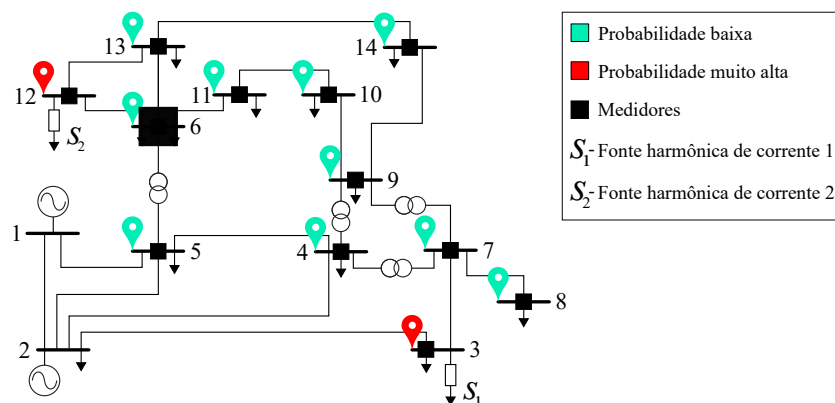
A inversão da ordem entre as estimativas reverbera nos resultados de localização, de acordo com a Figura 23. Além disso, via a abordagem do limiar para o mapa de probabilidade, a Figura 24 mostra que a entropia conjunta foi capaz de identificar precisamente os barramentos cujas fontes estão conectadas.

Figura 23 – Resultados de localização dos harmônicos para o caso de $|\rho| = 75\%$



Fonte: Elaborado pelo autor (2020).

Figura 24 – Mapa para as localizações para o caso de $|\rho| = 75\%$



Fonte: Elaborado pelo autor (2020).

4.2.3 Perfis lentos com $|\rho| = 80\%$ de correlação.

Para o ensaio com $|\rho| = 80\%$ entre os perfis lentos, utilizou-se perfis residenciais (MELO *et al.*, 2018; FRANCISQUINI, 2006), ambos como base para a variação lenta novamente. As fontes foram mantidas conectadas nos barramentos anteriores: fonte 1 no barramento 3, e a fonte 2 no barramento 12.

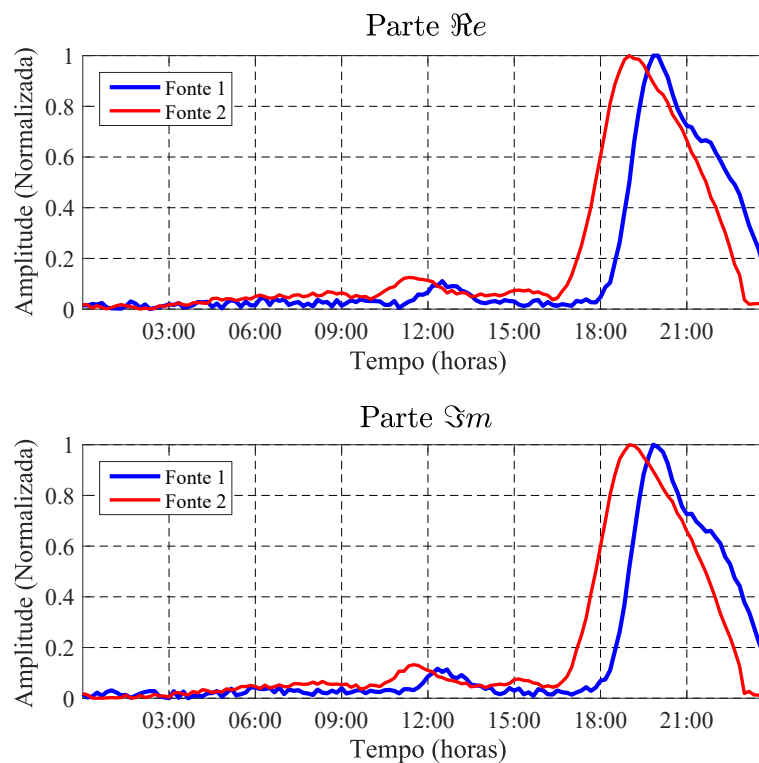
Quadro 4 – Perfis e espectros utilizados nas cargas para $|\rho| = 80\%$

Fonte 1		Fonte 2	
Perfil	Espectro	Perfil	Espectro
Residencial	Retificador	Residencial (mod.)	HDVC

Fonte: Elaborado pelo autor (2020).

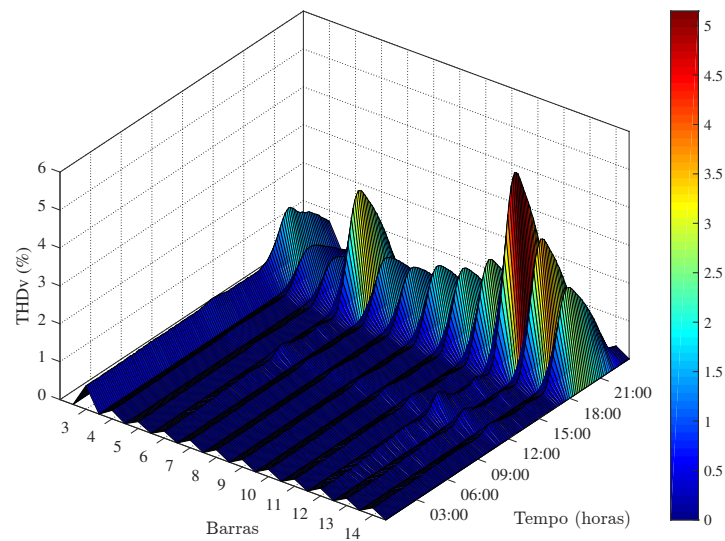
Neste caso, a fonte 2 possui um deslocamento de 6 amostras, o que corresponde à uma hora de adiantamento em relação ao perfil residencial da fonte 1. A Figura 25 apresenta as partes reais e imaginárias que foram injetadas no sistema.

Figura 25 – Perfis injetados com $|\rho| = 80\%$



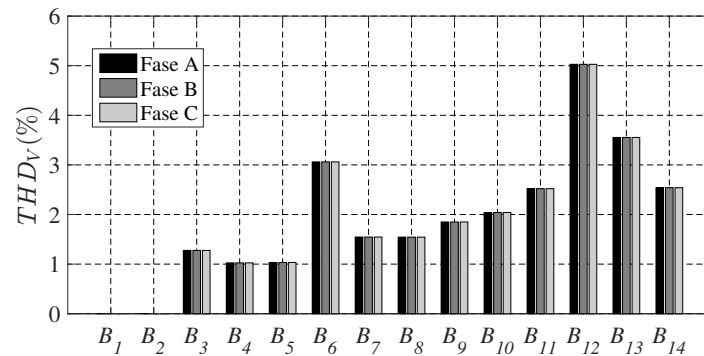
Fonte: Elaborado pelo autor (2020).

O Gráfico 10 apresenta a evolução do THDv ao longo do dia com a utilização dos perfis residenciais. Neste caso, a violação de 3% é observada nas barras 6, 12 e 13. Em relação ao caso anterior, o presente ensaio abrange uma maior região de poluição harmônica para a vizinhança da barra 12.

Gráfico 10 – THD de tensão ao longo do dia para o caso de $|\rho| = 80\%$ 

Fonte: Elaborado pelo autor (2020).

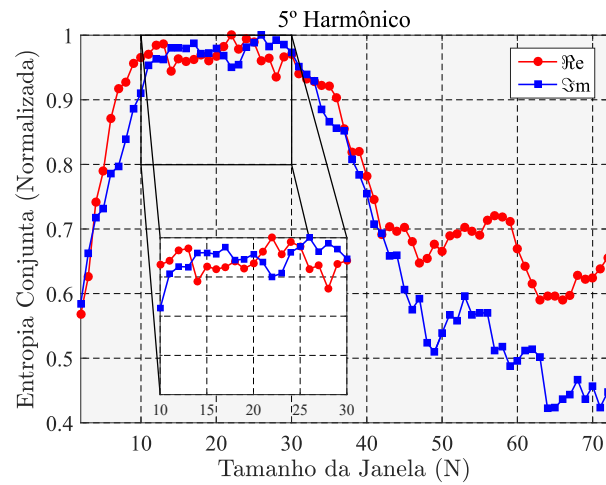
Observado o THDv no horário de pico, 19 horas, essa poluição harmônica nos barramentos 6, 12 e 13 fica mais evidente. Além disso, mesmo sem a violação de 3% as barras vizinhas sofrem com a distorção, por exemplo, a barra 11.

Gráfico 11 – THD de tensão no horário de pico para o caso de $|\rho| = 80\%$ 

Fonte: Elaborado pelo autor (2020).

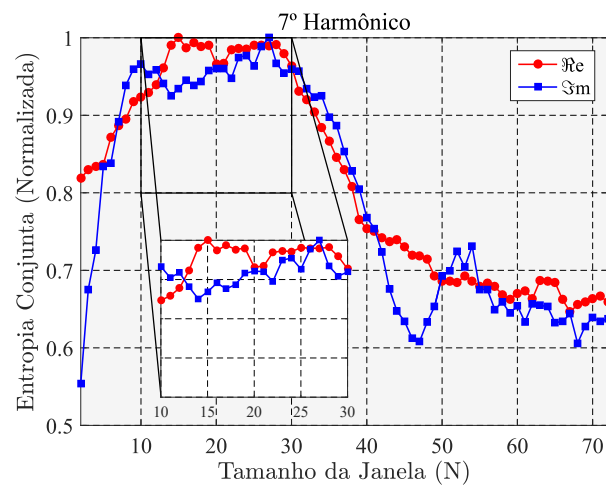
Entretanto, nada se pode inferir sobre uma possível localização da outra fonte. Poder-se-ia até mesmo julgar que a barra 6 possui carga não-linear conectada na mesma ao invés da barra 3. Assim, inicia-se a etapa de estimação. Porém, um $|\rho| = 80\%$ implica em correlação alta. Assim, é necessário maximizar a independência entre os componentes principais, para que se obtenha uma estimação de qualidade via ICA. Logo, um intervalo de confiança é estabelecido para o tamanho da janela utilizada pelo filtro. Os Gráficos 12, 13 e 14, mostram esse intervalo para o 5º, 7º e 11º harmônicos.

Gráfico 12 – Intervalo de confiança para o 5º harmônico



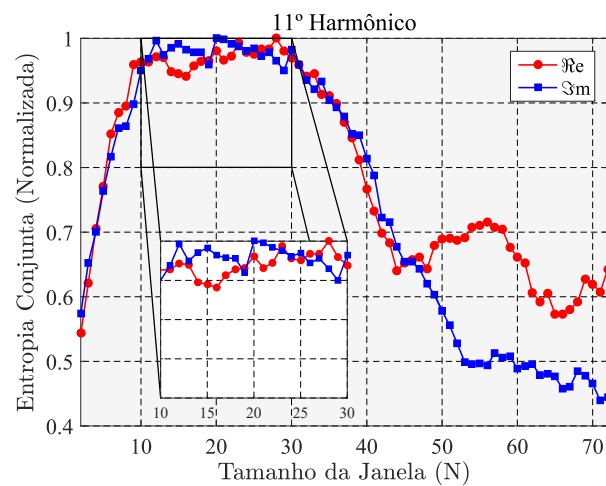
Fonte: Elaborado pelo autor (2020).

Gráfico 13 – Intervalo de confiança para o 7º harmônico



Fonte: Elaborado pelo autor (2020).

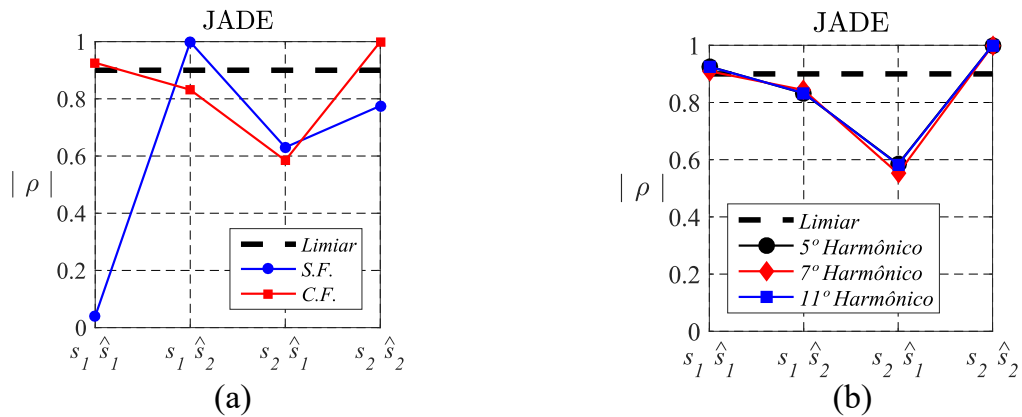
Gráfico 14 – Intervalo de confiança para o 11º harmônico



Fonte: Elaborado pelo autor (2020).

A janela $N = 11$ é mantida para os três harmônicos, pois se encontra dentro do intervalo de confiança, que maximiza a independência entre os CPs, ou seja, valores máximos de entropia conjunta para as partes reais e imaginárias. Então, aplica-se o algoritmo JADE, para a estimação das fontes. As estimativas são confrontadas com os perfis lentos originais, o que resulta na correspondência apresentada pela Figura 26.

Figura 26 – Correspondência via $|\rho|$ para os perfis residenciais

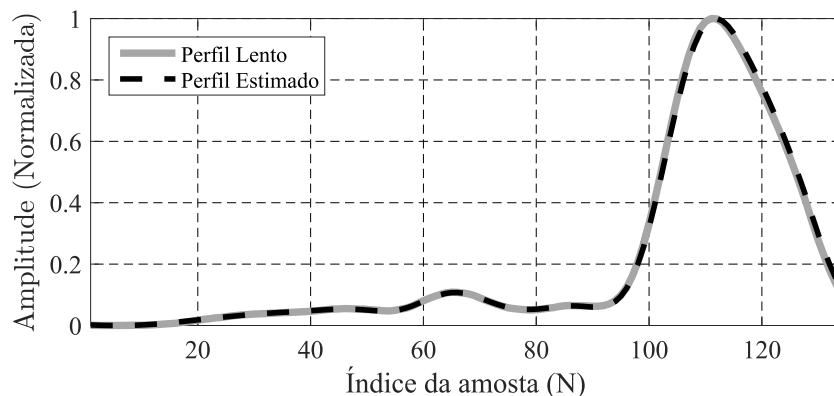


Fonte: Elaborado pelo autor (2020).

Para esse caso, $|\rho| = 80\%$, a filtragem impacta significativamente. Conforme a Figura 26 (a), dois componentes são estimados com a filtragem, acima do limiar de 0,9 e sem a presença de dupla correspondência. Em contrapartida, sem a filtragem apenas um componente é determinado acima do limiar de 0,9. Além disso, como esperado e de acordo com a Figura 26 (b), a estimação para as outras ordens harmônicas também se mantém com boa qualidade.

A Figura 27 mostra o resultado de estimação para a fonte 2.

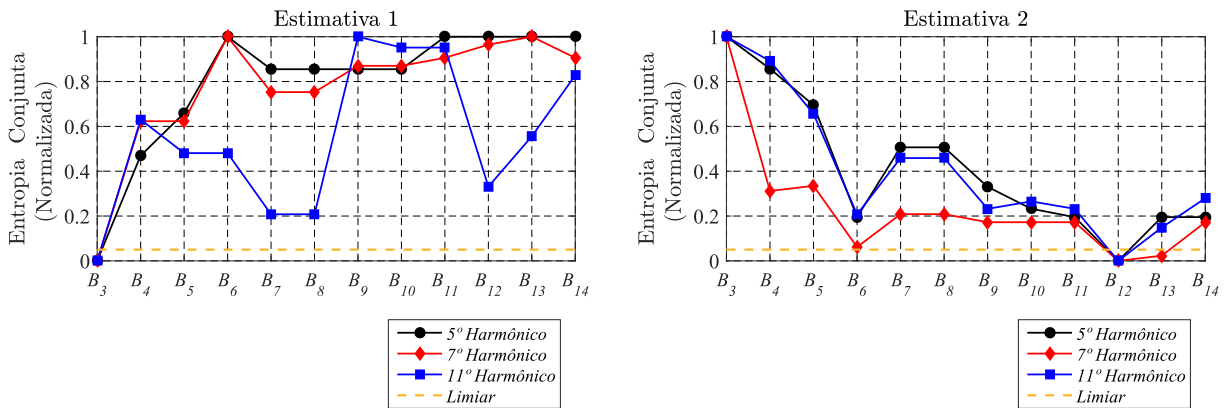
Figura 27 – Resultado de estimação para s_2 e \hat{s}_2



Fonte: Elaborado pelo autor (2020).

Aplicando a entropia conjunta, conforme a Figura 28, é possível determinar as posições das fontes harmônicas de corrente. Neste caso, para todas as ordens harmônicas, o mínimo indica a posição exata das fontes, B_3 e B_{12} . Neste teste, como não há inversão da ordem das fontes harmônicas no processo de estimação, não existe a presença desse efeito na etapa de localização.

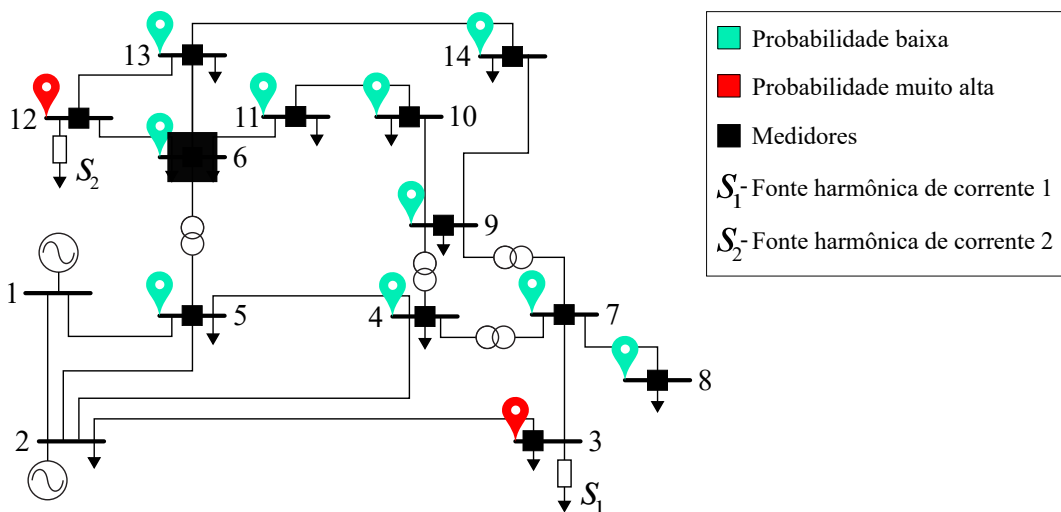
Figura 28 – Resultados de localização dos harmônicos para o caso de $|\rho| = 80\%$



Fonte: Elaborado pelo autor (2020).

Seguindo a abordagem dos 5%, não existe dois harmônicos abaixo desse limiar, apenas três. Logo, ao representar as probabilidades de conexão das fontes, através da Figura 29, somente as posições de conexão das fontes são ilustradas na cor vermelha.

Figura 29 – Mapa para as localizações para o caso de $|\rho| = 80\%$



Fonte: Elaborado pelo autor (2020).

4.2.4 Perfis lentos com $|\rho| = 98\%$ de correlação.

Para o ensaio com $|\rho| = 98\%$ entre os perfis lentos, utilizou-se um perfil industrial (fonte 1) e outro também industrial (fonte 2) (MELO *et al.*, 2018; FRANCISQUINI, 2006), ambos de acordo com o Quadro 5.

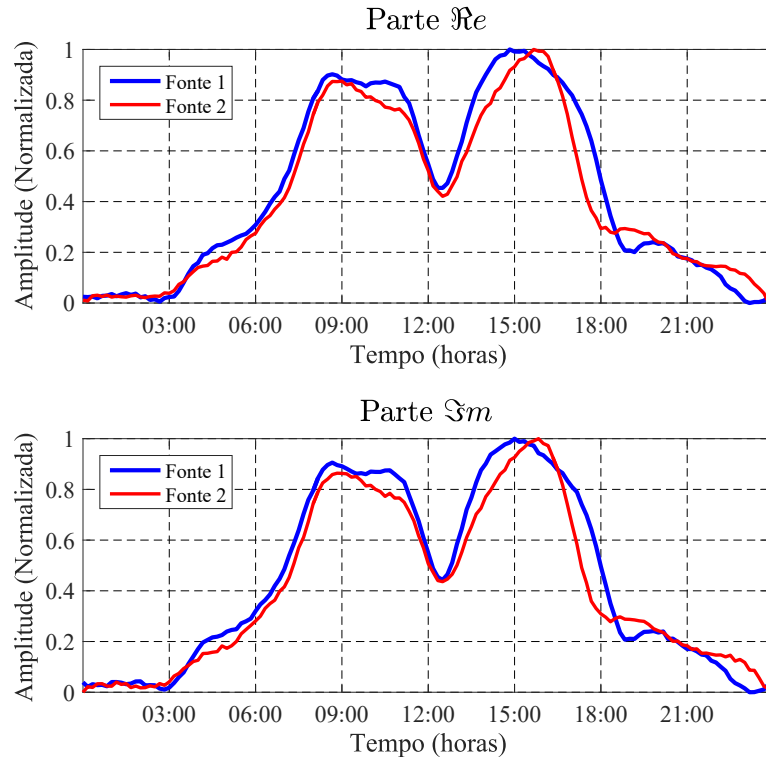
Quadro 5 – Perfis e espectros utilizados nas cargas para $|\rho| = 98\%$

Fonte 1		Fonte 2	
Perfil	Espectro	Perfil	Espectro
Industrial	HVDC	Industrial (mod.)	RTC

Fonte: Elaborado pelo autor (2020).

As fontes foram conectadas nos barramentos de forma que a fonte 1 fique no barramento 12 e a fonte 2 no barramento 3. Neste caso, a fonte 2 possui uma leve modificação de formato do perfil lento em relação ao perfil industrial da fonte 1, promovendo essa diferença de 2% no coeficiente de correlação do perfil lento. A Figura 30 apresenta as partes reais e imaginárias que foram utilizadas.

Figura 30 – Perfis injetados para o caso de $|\rho| = 98\%$

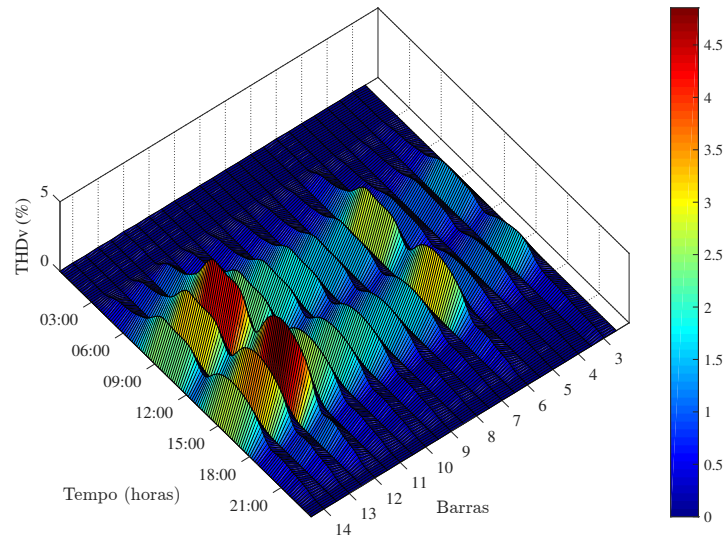


Fonte: Elaborado pelo autor (2020).

O Gráfico 15 apresenta a evolução do THDv ao longo do dia com a utilização dos perfis industriais. A violação dos 3% é observada nas barras 6, 12 e 13, novamente, com a

predominância do efeito na barra 12. Além disso, avaliando o barramento 3 ao longo do dia, nenhum efeito exacerbado é visivelmente perceptível.

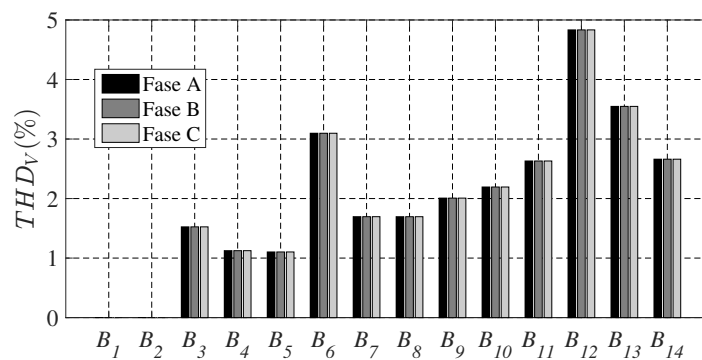
Gráfico 15 – THD de tensão ao longo do dia para o caso de $|\rho| = 98\%$



Fonte: Elaborado pelo autor (2020).

Conforme o Gráfico 16, verifica-se o THDv no horário de pico de poluição harmônica, às 15 horas. Essa verificação salienta que existe uma região de perturbações correspondente à região noroeste do sistema IEEE 14 barras, onde estão contidas as barras 6, 12 e 13. Elevações do THDv nas barras 11 e 14 também são perceptíveis. A hipótese mais natural e intuitiva seria indicar a existência de uma fonte harmônica de corrente apenas na região noroeste do sistema.

Gráfico 16 – THD de tensão no horário de pico para o caso de $|\rho| = 98\%$

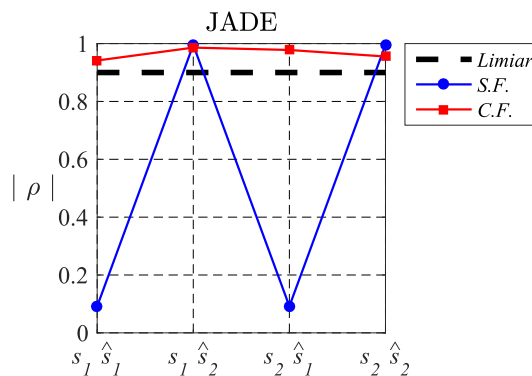


Fonte: Elaborado pelo autor (2020).

De acordo com o resultado obtido na seção (4.2.5), para uma boa estimação de perfis altamente correlacionados é indispensável que a janela do filtro esteja contida no intervalo de confiança. Neste caso, a janela de $N = 11$ ainda contempla essa condição para os três harmônicos. Logo, utiliza-se $N = 11$.

Após a definição da janela, aplica-se a metodologia de redução de dimensionalidade, posteriormente a filtragem e, em seguida, a ICA via o algoritmo JADE, resultando na correspondência apresentada pelo Gráfico 17. Ao observar o resultado sem filtro (*S.F.*), ainda no Gráfico 17, um questionamento pode surgir: como esse resultado é ruim se existem dois valores de altíssima correlação? Porém, Neste caso, ocorre uma dupla correspondência, isto é, a estimativa 2 (\hat{s}_2) apresenta correspondência tanto para a fonte 1 (s_1) quanto para a fonte 2 (s_2). Logo, sem a filtragem não há separabilidade entre as fontes.

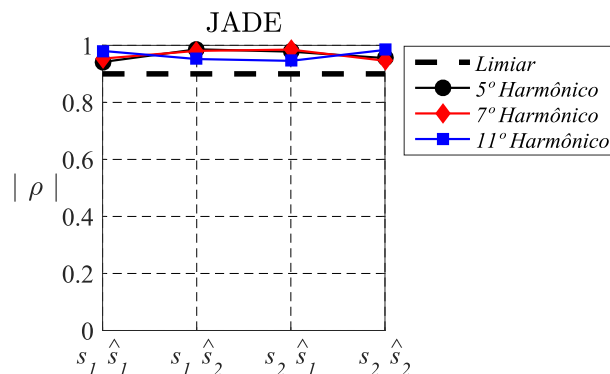
Gráfico 17 – Correspondência via $|\rho|$ para o 5º harmônico do quarto caso



Fonte: Elaborado pelo autor (2020).

Conforme a Figura 17, existem valores altos de correspondência para todas as estimativas para a abordagem com filtro. Isso pode ser compreendido devido à alta correlação entre os perfis injetados no sistema ($|\rho| = 98\%$). Entretanto, existem correspondências mais elevadas que outras, como $s_1\hat{s}_2$ e $s_2\hat{s}_1$. Este resultado implica na inversão da ordem das estimativas em relação às fontes, uma ambiguidade característica da ICA. Este comportamento se mantém para o 5º e 7º harmônicos, entretanto, a inversão da ordem não ocorre para o 11º harmônico, conforme o Gráfico 18.

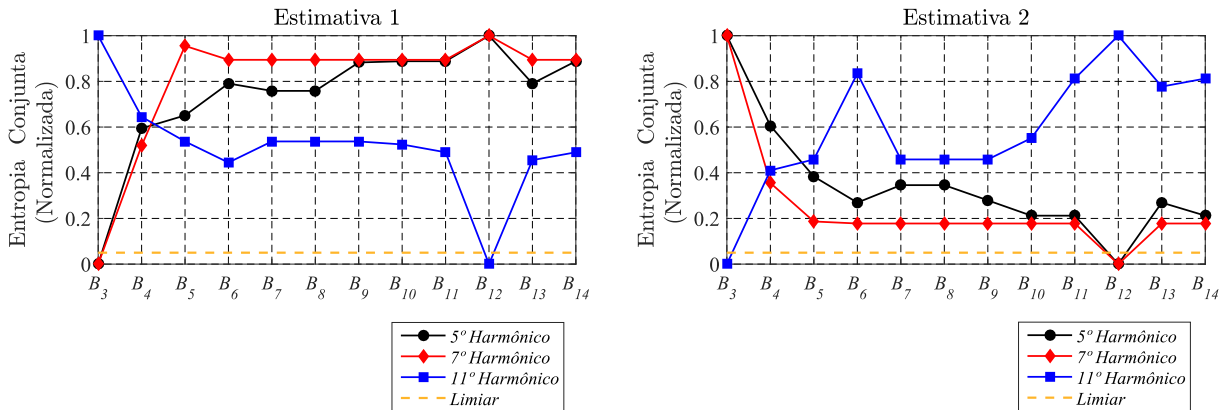
Gráfico 18 – Correspondência via $|\rho|$ para os harmônicos do quarto caso



Fonte: Elaborado pelo autor (2020).

A Figura 31 apresenta os resultados para a entropia conjunta, e inicialmente, é perceptível o efeito da inversão da ordem das fontes para o 11º harmônico no processo de localização.

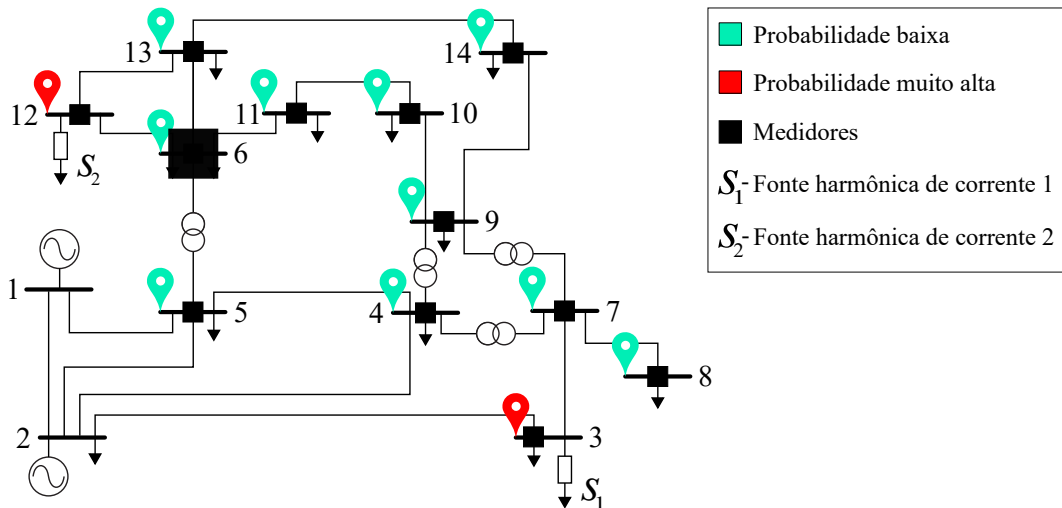
Figura 31 – Resultados de localização dos harmônicos para o caso de $|\rho| = 98\%$



Fonte: Elaborado pelo autor (2020).

O mapa de probabilidade está representado na Figura 32. Neste teste, a inversão da ordem das fontes não prejudica o limiar, visto que, ao observar os resultados como um todo, três ordens harmônicas estão abaixo do limiar para a barra 3, ao mesmo tempo que outras três ordens harmônicas estão abaixo do limiar de 5% para a barra 12.

Figura 32 – Mapa para as localizações para o caso de $|\rho| = 98\%$

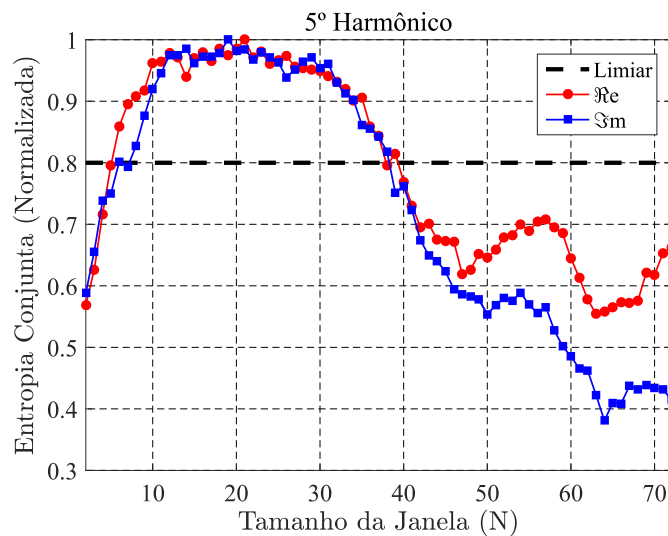


Fonte: Elaborado pelo autor (2020).

4.2.5 Janela Fora do Intervalo de Confiança

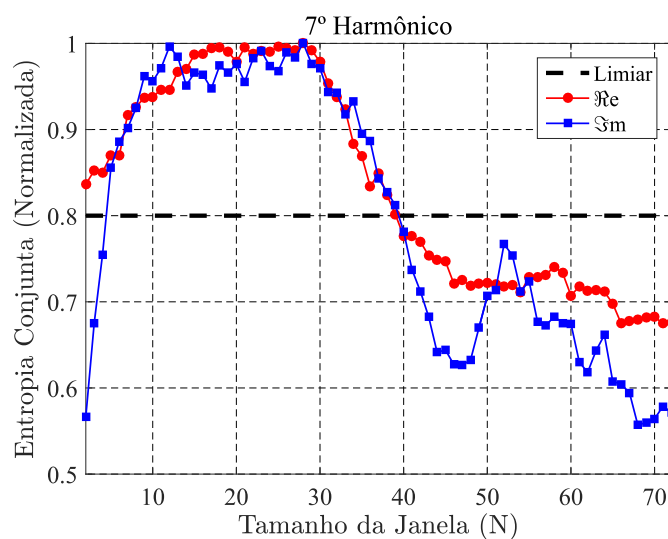
O propósito deste ensaio é verificar se existe um limiar para a escolha da janela do filtro. Para a entropia conjunta entre os pares reais e imaginários dos CPs, foi estabelecido um limiar de 80% e uma janela (N) abaixo desse limiar foi escolhida para as três ordens harmônicas. Os Gráficos 19, 20 e 21, apresentam a situação supracitada para as ordens harmônicas estudadas.

Gráfico 19 – Limiar Estabelecido para o 5º Harmônico



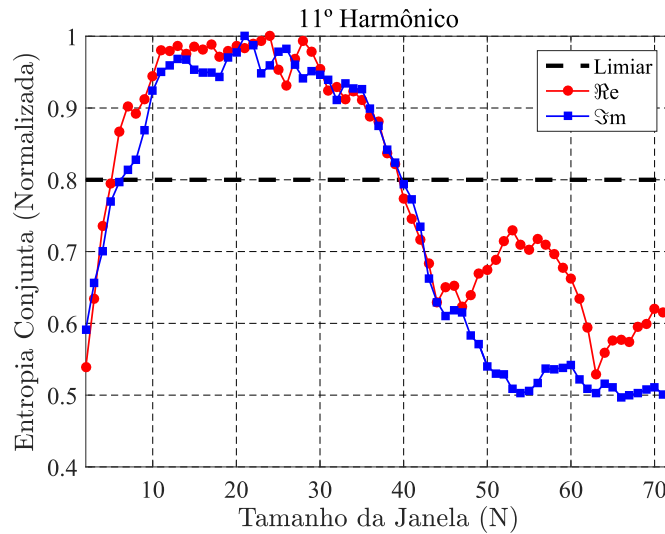
Fonte: Elaborado pelo autor (2020).

Gráfico 20 – Limiar Estabelecido para o 7º Harmônico



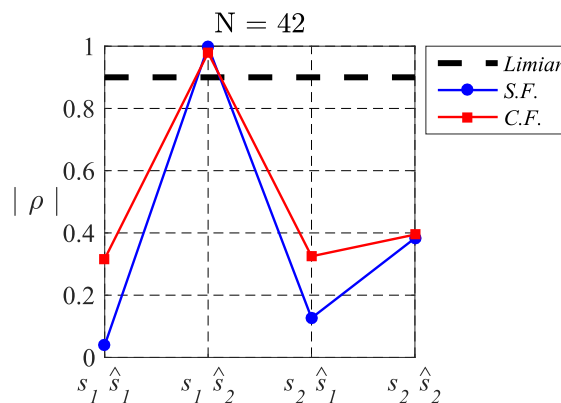
Fonte: Elaborado pelo autor (2020).

Gráfico 21 – Limiar Estabelecido para o 11º Harmônico

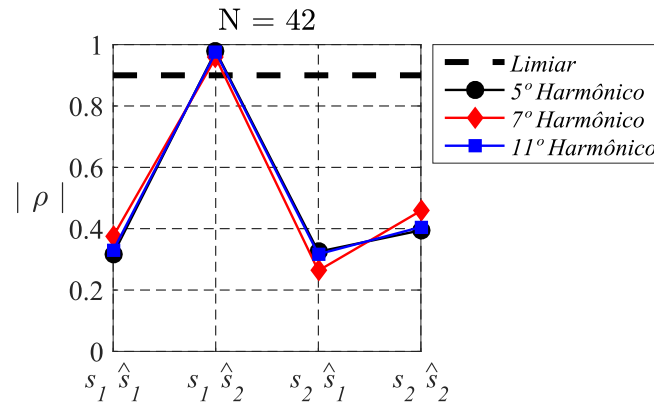


Fonte: Elaborado pelo autor (2020).

Assim, observou-se que quando uma janela de $N = 42$ que está abaixo do limiar é utilizada, os resultados de estimação se comportam da mesma forma que uma situação de não filtragem. O Gráfico 22 apresenta essa situação para o 5º harmônico. Além disso, de acordo com o Gráfico 23, esse resultado se repete para as três ordens harmônicas. Outro fato interessante entre os Gráficos 22 e 23, é que uma estimativa apresenta altíssima correlação. Neste caso, pode até ter ocorrido estimação da fonte 1 (s_1) via estimativa (\hat{s}_2), mas do ponto de vista de BSS não ocorreu separação entre as fontes a partir da mistura utilizada.

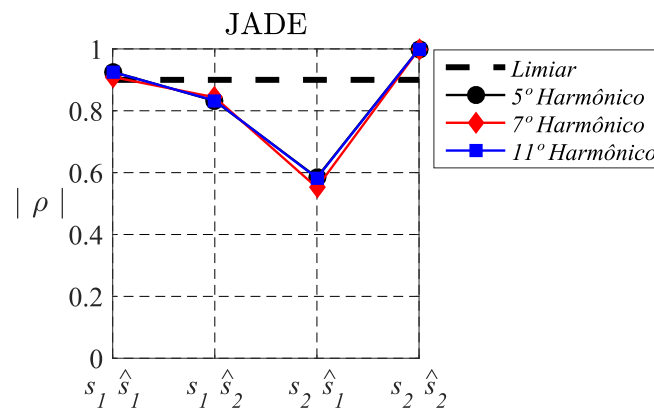
Gráfico 22 – Correspondência via $|\rho|$ para o 5º harmônico ($N = 42$)

Fonte: Elaborado pelo autor (2020).

Gráfico 23 – Correspondência via $|\rho|$ para os harmônicos ($N = 42$)

Fonte: Elaborado pelo autor (2020).

Porém, com uma janela adequada, dentro do intervalo de confiança ($N = 11$), conforme o Gráfico 24, os resultados apresentam-se bons para as estimativas obtidas.

Gráfico 24 – Correspondência via coeficiente de correlação ($N = 11$)

Fonte: Elaborado pelo autor (2020).

4.2.6 Alterando a Distribuição de Probabilidade da Variação Rápida.

O propósito deste ensaio é testar a robustez do método na presença de modelos diferentes de distribuições na variação rápida. Além disso, com o intuito de formalização, utiliza-se a nomenclatura d.d.s para a distribuição do sinal em todos os ensaios e intervalos. Outro ponto é que a frequência trata-se simplesmente da ocorrência da amostra dentro da amplitude apresentada graficamente, não há normalização ou conversão em porcentagem.

- **Caso 1:** Para o presente ensaio, três intervalos de ocorrência foram utilizados, obedecendo uma distribuição uniforme, conforme a f.d.p. do Quadro 6, e cada um desses intervalos compreende uma simulação.

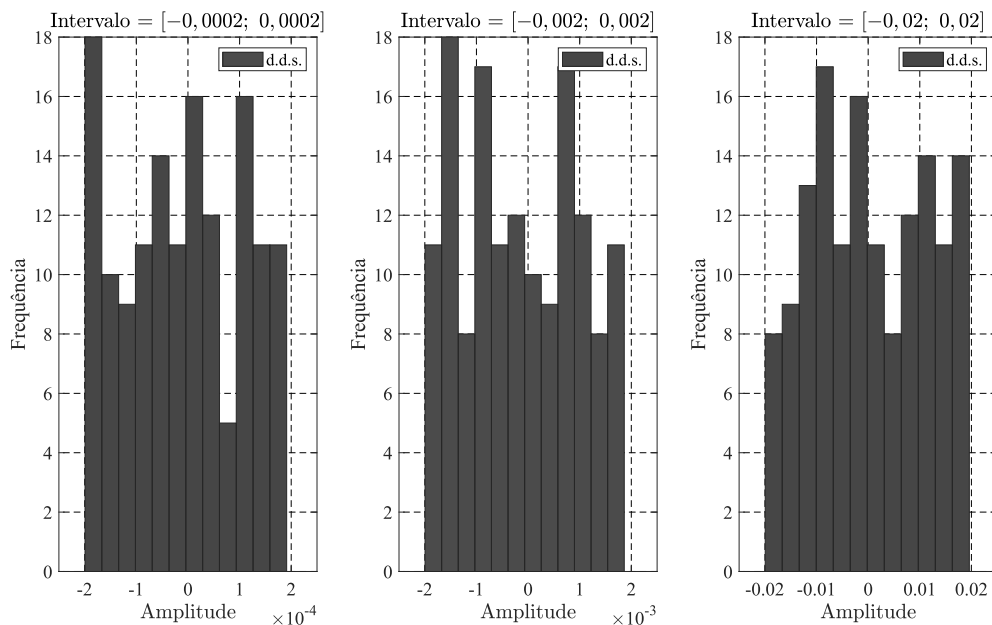
Quadro 6 – Modelo da variação rápida utilizada no caso 1

Função Densidade de Probabilidade		
$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário.} \end{cases}$		
Intervalos	a	b
Primeiro	-0,0002	0,0002
Segundo	-0,0020	0,0020
Terceiro	-0,0200	0,0200

Fonte: Elaborado pelo autor (2020).

Conforme esse intervalo se estende, a volatilidade das ocorrências aumenta, isto é, o intervalo de ocorrência abrange valores maiores. Isso se verifica conforme a Figura 33, que apresenta as amplitudes que compõem o histograma das distribuições utilizadas. É válido salientar que o sinal, como um todo, apresenta apenas 144 amostras. Além disso, quanto maior a amplitude dessa variação rápida, maior sua interferência no comportamento da variação do perfil completo.

Figura 33 – Histogramas das variações rápidas para o caso 1



Fonte: Elaborado pelo autor (2020).

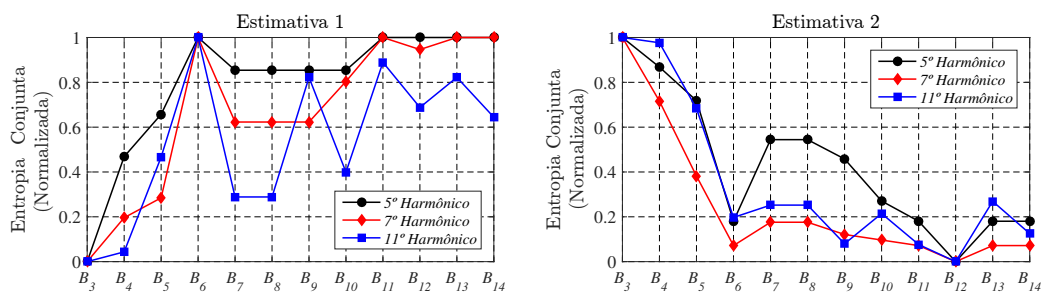
A Tabela 3 apresenta resultados que, mesmo com a progressão da grandeza do intervalo de ocorrência, as estimativas apresentam valores altos de $|\rho|$ em relação às suas fontes correspondentes (s_1, \hat{s}_1 e s_2, \hat{s}_2). Além disso, as Figuras 34, 35 e 36, mostram que não há um impacto negativo no processo de localização, pois em todos os casos a entropia foi capaz de indicar os locais das fontes harmônicas de corrente.

Tabela 3 – Correspondência via $|\rho|$ para os três intervalos utilizados no caso 1

Intervalo = $[-0,0002; 0,0002]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9228	0,8322	0,5777	0,9996
7	0,9228	0,8319	0,5775	0,9996
11	0,9228	0,8322	0,5776	0,9996
Intervalo = $[-0,0020; 0,0020]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9235	0,8314	0,5790	0,9996
7	0,9249	0,8293	0,5821	0,9994
11	0,9238	0,8311	0,5797	0,9995
Intervalo = $[-0,0200; 0,0200]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9128	0,8363	0,5577	0,9998
7	0,9274	0,8200	0,5918	0,9933
11	0,9181	0,8342	0,5685	0,9996

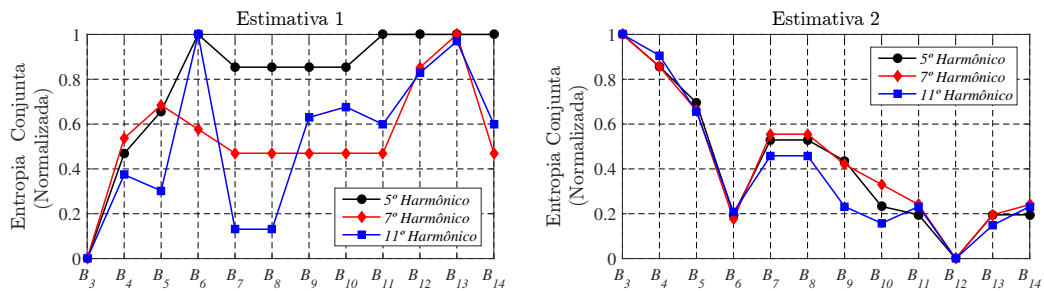
Fonte: Elaborada pelo autor (2020).

Figura 34 – Localização das fontes harmônicas de corrente para o intervalo $[-0,0002; 0,0002]$



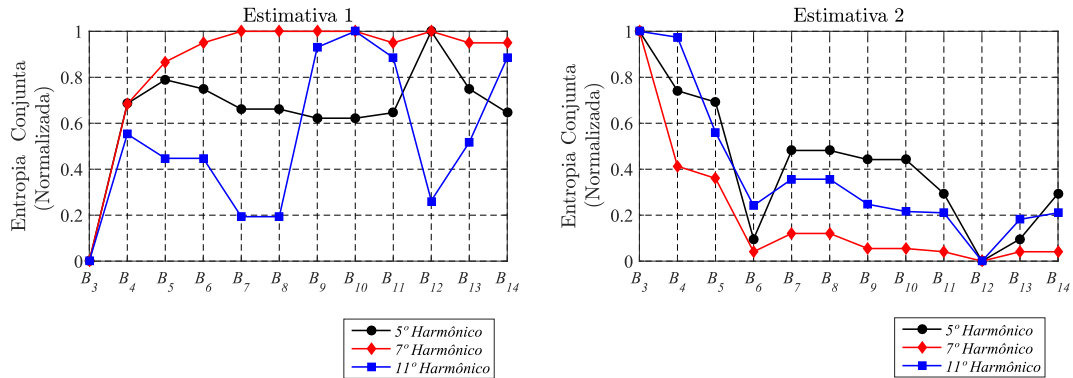
Fonte: Elaborado pelo autor (2020).

Figura 35 – Localização das fontes harmônicas de corrente para o intervalo $= [-0,002; 0,002]$



Fonte: Elaborado pelo autor (2020).

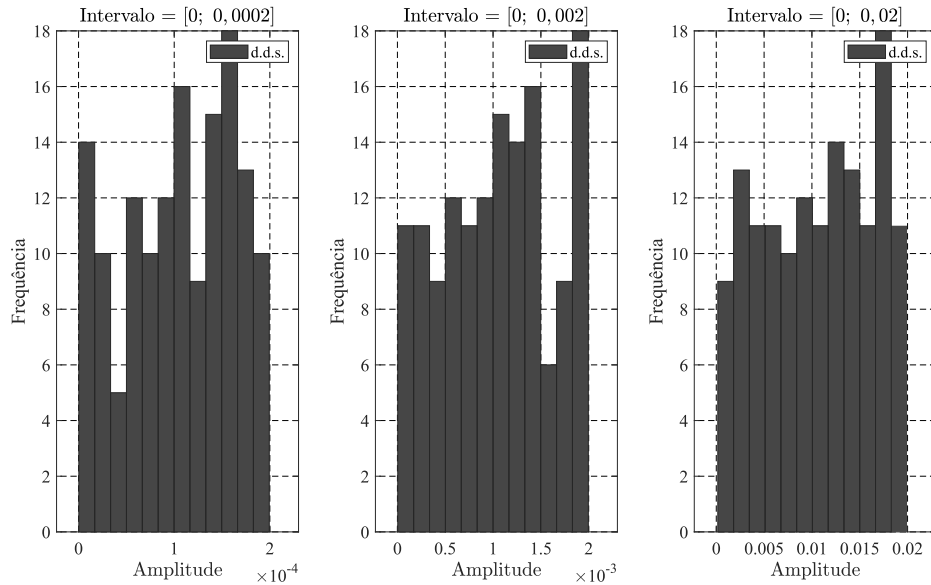
Figura 36 – Localização das fontes harmônicas de corrente para o intervalo $= [-0,02; 0,02]$



Fonte: Elaborado pelo autor (2020).

- Caso 2:** Para este ensaio, as distribuições da variação rápida obedecem a f.d.p. do caso anterior, porém, com os intervalos apresentados pela Figura 33, em que $a = 0$ na Tabela 3. Partindo do primeiro intervalo, o segundo é modificado na ordem de dez vezes e o terceiro em dez vezes em relação ao segundo, conforme o ensaio anterior. Essa variação de grandeza é importante, pois testa a robustez do método. Verificando as amplitudes apresentadas pelos histogramas da Figura 37, é válido salientar que para todos os testes cujo foco não reside em estudar as distribuições da variação rápida, utilizam, como padrão, a distribuição uniforme que é regida pelo terceiro intervalo.

Figura 37 – Histogramas das variações rápidas para o caso 2



Fonte: Elaborado pelo autor (2020).

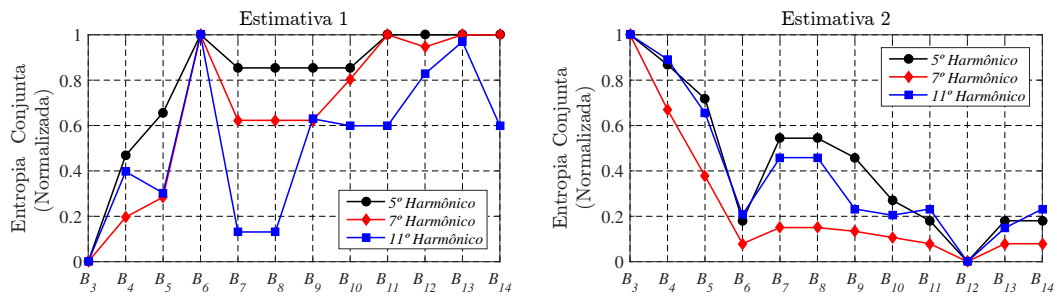
Ao verificar os resultados de correspondência via coeficiente de correlação, Tabela 4, em comparação com os resultados de localização, Figuras 38, 39 e 40, observa-se que a entropia não encontrou obstáculos independentemente da grandeza do intervalo. Além disso, os resultados para os intervalos de $[0; 0,0002]$ e $[0; 0,002]$ são muito semelhantes, apresentando a capacidade da entropia conjunta em localizar as fontes para todos os componentes harmônicos.

Tabela 4 – Correspondência via $|\rho|$ para os três intervalos ($a = 0$)

Intervalo = $[0; 0,0002]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9229	0,8321	0,5779	0,9996
7	0,9229	0,8322	0,5778	0,9996
11	0,9229	0,8321	0,5779	0,9996
Intervalo = $[0; 0,0020]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9231	0,8322	0,5782	0,9996
7	0,9225	0,8327	0,5769	0,9996
11	0,9230	0,8322	0,5781	0,9996
Intervalo = $[0; 0,0200]$				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9225	0,8320	0,5771	0,9996
7	0,9192	0,8443	0,5712	0,9990
11	0,9224	0,8332	0,5768	0,9996

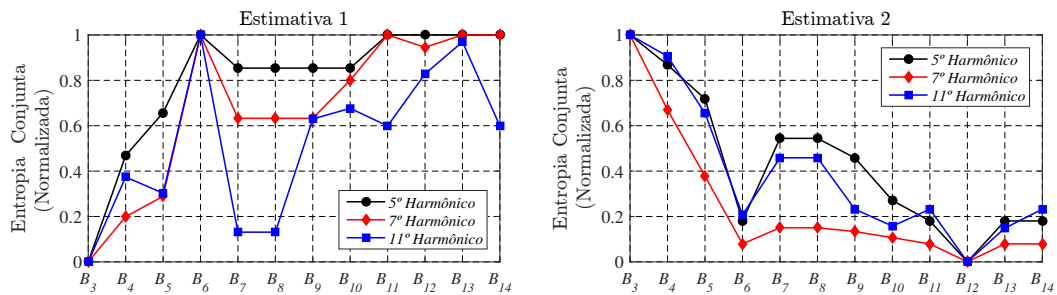
Fonte: Elaborada pelo autor (2020).

Figura 38 – Localização das fontes harmônicas de corrente para o intervalo $= [0; 0,0002]$



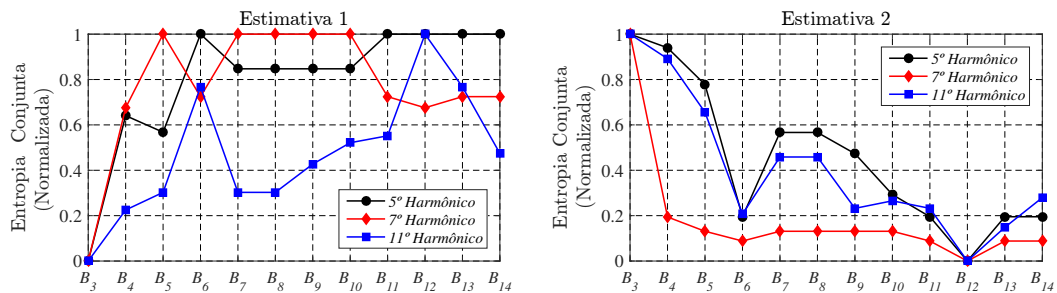
Fonte: Elaborado pelo autor (2020).

Figura 39 – Localização das fontes harmônicas de corrente para o intervalo $= [0; 0,002]$



Fonte: Elaborado pelo autor (2020).

Figura 40 – Localização das fontes harmônicas de corrente para o intervalo $= [0; 0,002]$



Fonte: Elaborado pelo autor (2020).

- **Caso 3:** Finalmente, foram utilizadas distribuições laplacianas (KOTZ; KOZUBOWSKI; PODGORSKI, 2012), conforme a f.d.p. do Quadro 7. Além disso, as variações da ordem de grandeza também foram mantidas para estudo da eficácia da metodologia.

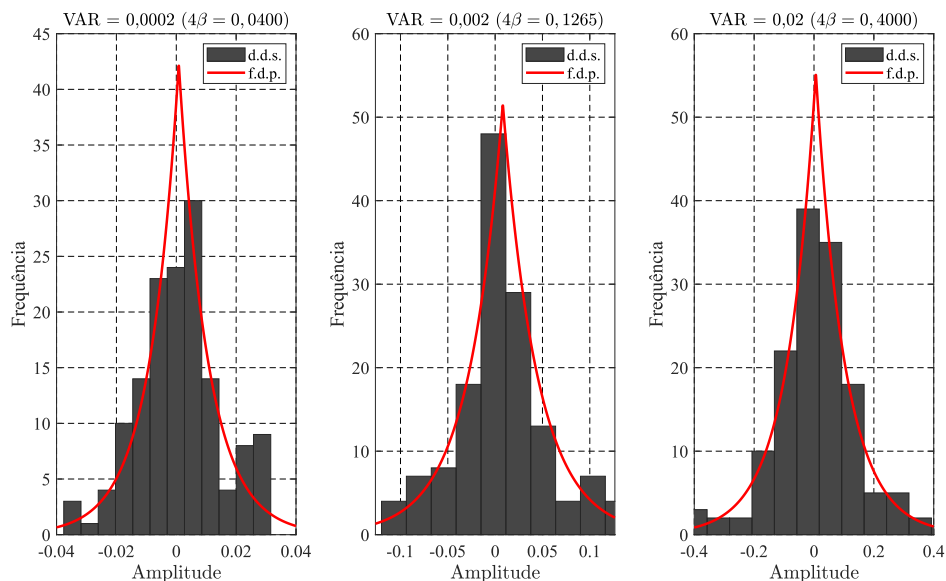
Quadro 7 – Modelo da variação rápida utilizada no caso 3

Função Densidade de Probabilidade	
$f_X(x) = \frac{1}{2\beta} e^{- x /\beta}, \text{ VAR} = 2\beta^2$	
Variâncias (VAR)	
Primeiro valor	0,0002
Segundo valor	0,0020
Terceiro valor	0,0200

Fonte: Elaborado pelo autor (2020).

As amplitudes obtidas são apresentadas conforme a Figura 41, Neste caso, a grandeza β , ou diversidade, é utilizada para ilustrar o intervalo de ocorrência. Neste ensaio, é importante salientar que o modelo de variação rápida utilizada pela referência da metodologia de estimação (GURSOY, 2007), é $\text{VAR} = 0,002$, Neste caso, a variância intermediária do ensaio.

Figura 41 – Histogramas das variações rápidas para o caso 3



Fonte: Elaborado pelo autor (2020).

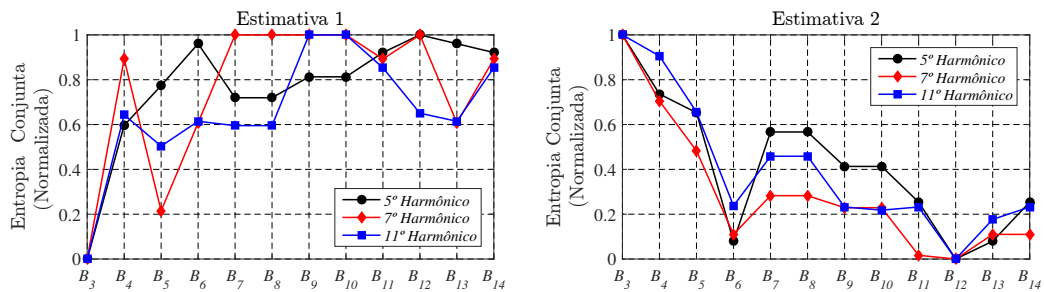
Para as variâncias de 0,0002 e 0,002, tanto a estimação, Tabela 5, quanto a localização, Figuras 42 e 43, se comportam bem e encontram o local de conexão de ambas as fontes harmônicas de corrente. Porém, um infortúnio ocorre quando $\text{VAR} = 0,02$ é utilizada, em que, a metodologia é incapaz de estimar e localizar ambas as fontes para o 7º harmônico. Uma possível justificativa é o valor elevado para a variância, visto que, $\text{VAR} = 0,02$ corresponde a $4\beta = 0,4$, o que é muito elevado se comparado a unidade, onde encontra-se o valor de escala da variação lenta.

Tabela 5 – Correspondência via $|\rho|$ para os três intervalos de variância do caso 3

VAR = 0,0002				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9327	0,8332	0,6029	0,9996
7	0,9016	0,8262	0,5526	0,9954
11	0,9345	0,8317	0,6059	0,9995
VAR = 0,0020				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9164	0,8277	0,5750	0,9989
7	0,9336	0,8265	0,7014	0,9047
11	0,9241	0,8320	0,5869	0,9985
VAR = 0,0200				
Harmônico	$s_1\hat{s}_1$	$s_1\hat{s}_2$	$s_2\hat{s}_1$	$s_2\hat{s}_2$
5	0,9504	0,7708	0,7532	0,9875
7	0,7379	0,2365	0,7657	0,3406
11	0,9829	0,7455	0,8216	0,9696

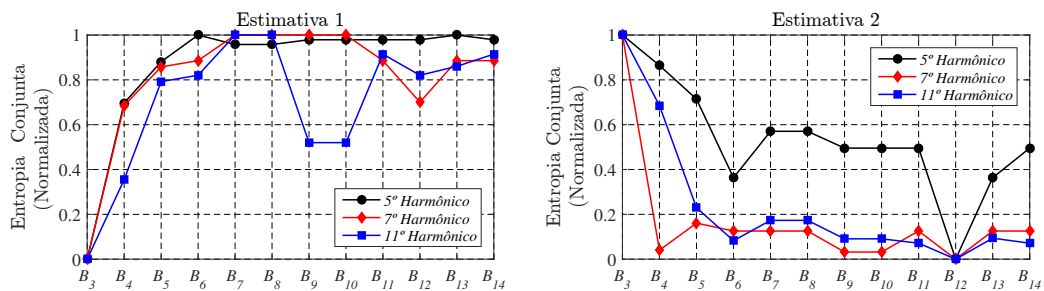
Fonte: Elaborada pelo autor (2020).

Figura 42 – Localização das fontes harmônicas de corrente para VAR = 0,0002



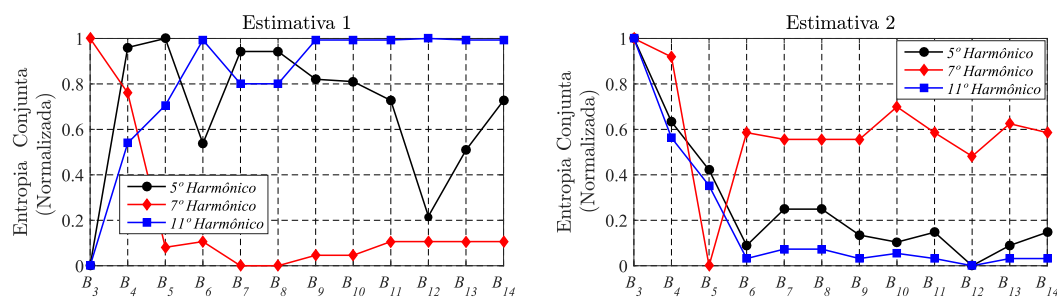
Fonte: Elaborado pelo autor (2020).

Figura 43 – Localização das fontes harmônicas de corrente para VAR = 0,002



Fonte: Elaborado pelo autor (2020).

Figura 44 – Localização das fontes harmônicas de corrente para VAR = 0,02



Fonte: Elaborado pelo autor (2020).

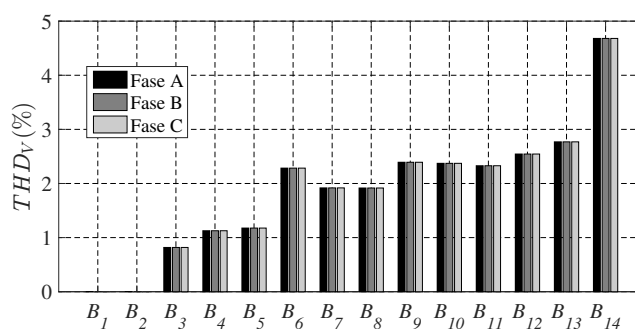
Conclui-se que o trabalho é robusto para variações rápidas diversas, perdendo performance somente para o modelo laplaciano com altíssima variância (VAR = 0,02). Além disso, contempla ótimos resultados quando colocado na mesma situação do trabalho proposto por Gursoy (2007), que é a referência base para a etapa de estimação.

4.2.7 Variando as Posições das Fontes

Nesta seção, as posições das fontes foram variadas como forma de avaliar a robustez do método.

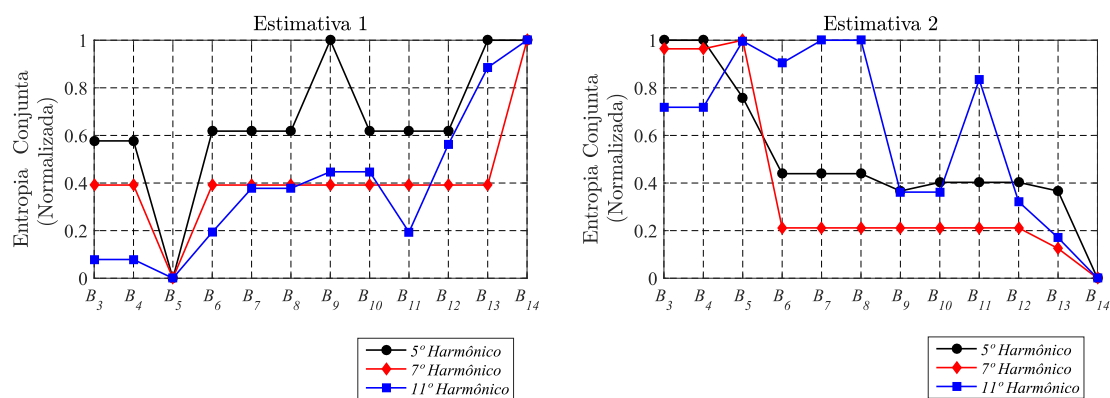
- **Caso 1:** Neste ensaio, a fonte harmônica de corrente 1 foi conectada à barra 5, ao passo que a fonte 2 foi conectada à barra 14. Observando o THDv no horário de pico, Gráfico 25, nota-se que o THDv se eleva apenas na região cuja fonte 2 está conectada. Ao aplicar a metodologia, detecta-se a fonte harmônica de corrente 2 na barra 14 e a fonte 1 na barra 5, antes imperceptível via análise do Gráfico 25. Correlacionando a topologia do sistema e os resultados de entropia, Figura 45, é possível produzir um mapa único com as possíveis localizações das fontes, conforme a Figura 46.

Gráfico 25 – THD de tensão no horário de pico (19 horas) para o caso 1



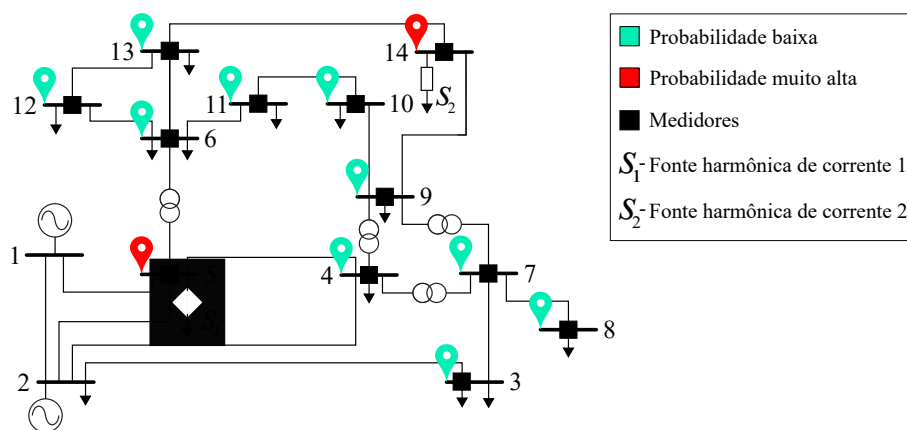
Fonte: Elaborado pelo autor (2020).

Figura 45 – Resultados de entropia conjunta para o caso 1



Fonte: Elaborado pelo autor (2020).

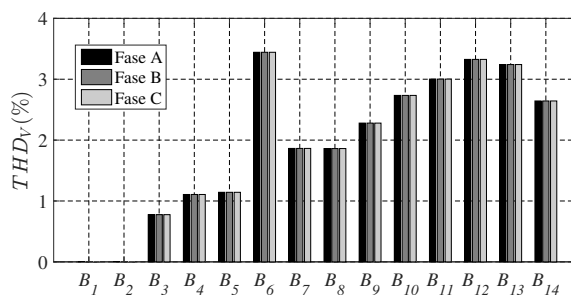
Figura 46 – Mapa de probabilidade para o caso 1



Fonte: Elaborado pelo autor (2020).

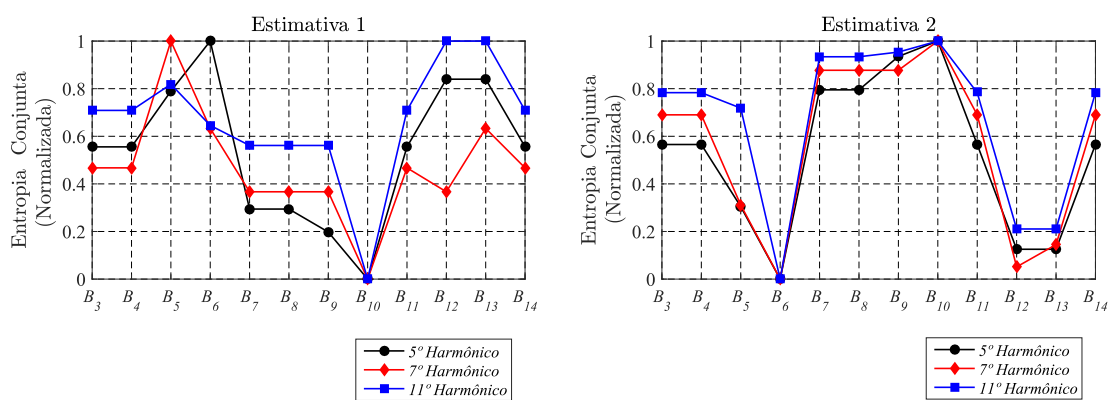
- **Caso 2:** Para o segundo ensaio, a fonte harmônica de corrente 1 foi conectada à barra 10 e a fonte 2 foi conectada à barra 6. Neste caso, a distância entre as fontes foi drasticamente reduzida, sendo separadas apenas pela barra 11. Infortunadamente, era esperado que o THDv das barras 6 e 10 fossem os maiores. Porém, o Gráfico 26 apresenta os maiores valores além de 3% para as barras 6, 12 e 13. É um resultado que fortalece a utilização da metodologia no lugar de uma análise simples de THDv. A Figura 47 identifica os mínimos de entropia com as posições corretas das fontes, mesmo com ambas muito próximas do ponto de vista de topologia, conforme o mapa da Figura 48.

Gráfico 26 – THD de tensão no horário de pico (19 horas) para o caso 2



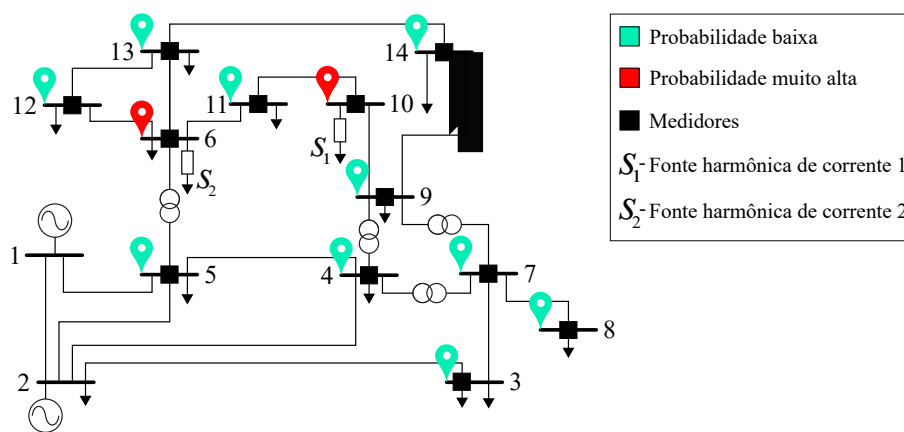
Fonte: Elaborado pelo autor (2020).

Figura 47 – Resultados de entropia conjunta para o caso 2



Fonte: Elaborado pelo autor (2020).

Figura 48 – Mapa para de probabilidade para o caso 2

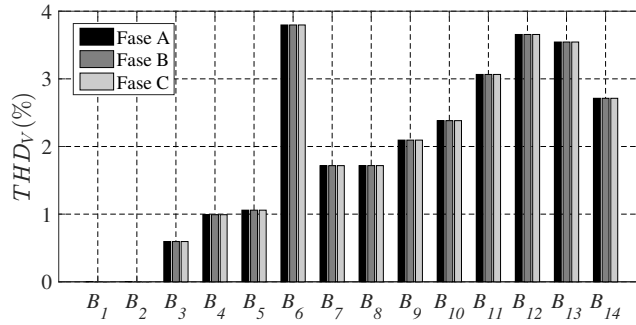


Fonte: Elaborado pelo autor (2020).

- **Caso 3:** Do caso anterior, mantém-se a fonte 2 no barramento 6. Porém, a fonte 1 é movida para a barra 5, situada abaixo de um transformador. As fontes estão separadas apenas por um único transformador. O Gráfico 27 mostra que o 3% de THDv foi superado em toda a região de vizinhança cuja fonte 2 está conectada. Mas

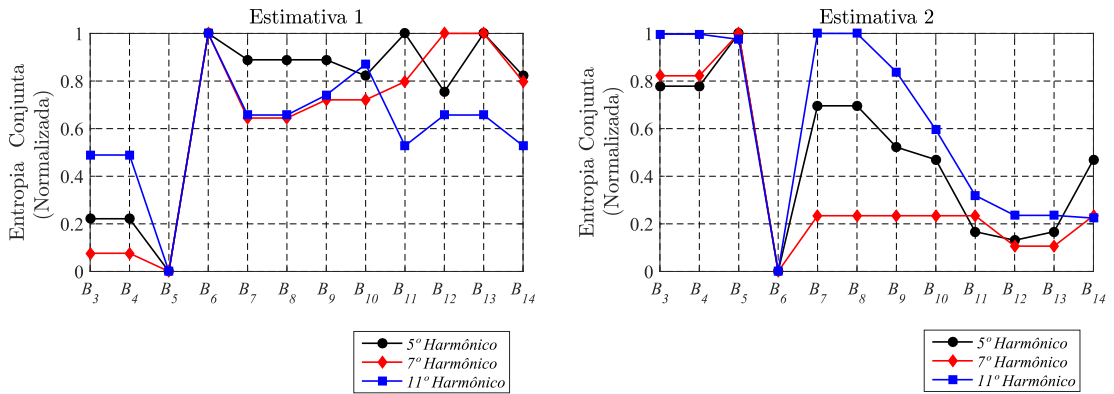
no barramento 5, conforme o THD_v, não há indícios de uma fonte harmônica de corrente. Uma sugestão para essa situação, seria a presença do transformador que possa estar “filtrando” a poluição causada pela fonte 1. Ao utilizar a metodologia, as posições das fontes ficam evidentes, conforme os resultados expostos na Figura 49, bem como o mapa da Figura 50.

Gráfico 27 – THD de tensão no horário de pico (19 horas) para o caso 3



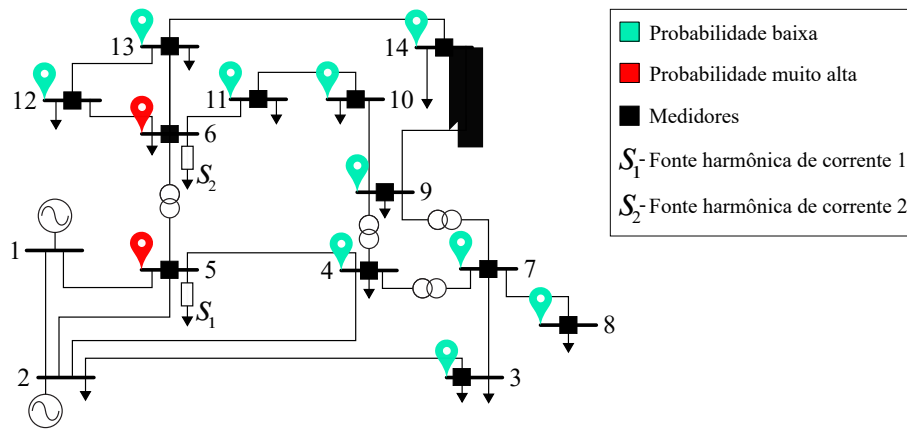
Fonte: Elaborado pelo autor (2020).

Figura 49 – Resultados de entropia conjunta para o caso 3



Fonte: Elaborado pelo autor (2020).

Figura 50 – Mapa de probabilidade para o caso 3

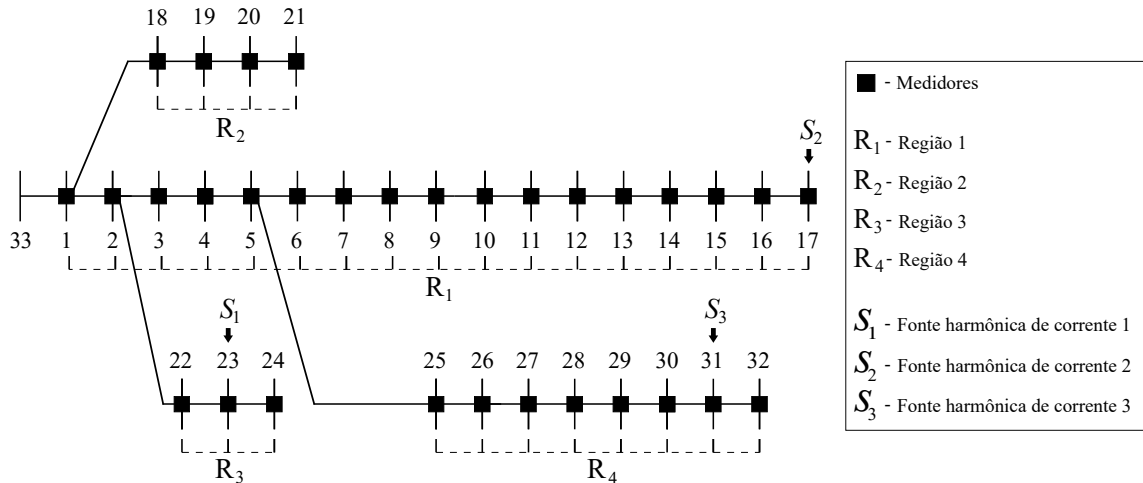


Fonte: Elaborado pelo autor (2020).

4.3 RESULTADOS PARA O SISTEMA IEEE 33 BARRAS

O sistema IEEE 33 barras utilizado é oriundo do trabalho proposto por Melo (2018), possui topologia radial e é desequilibrado. O sistema possui um alimentador principal da barra 33 à barra 17, um alimentador lateral da barra 1 até barra 21, outro alimentado lateral da barra 2 ao barramento 24 e, por fim, o último alimentador lateral da barra 5 ao barramento 32 (MELO, 2018). Neste ensaio, quatro regiões são definidas: R_1 , R_2 , R_3 e R_4 . Essas regiões contemplam os ramos desse sistema com a finalidade de compreender melhor os resultados obtidos da entropia conjunta, que serão discutidos no final desta seção. Todas essas descrições referentes à topologia estão ilustradas na Figura 51. Ademais, os espectros utilizados para as cargas não lineares foram novamente baseados na Tabela 1.

Figura 51 – Sistema IEEE 33 barras



Fonte: Elaborado pelo autor (2020).

Além disso, este teste utiliza combinações entre perfis e espectros harmônicos, para as fontes harmônicas de corrente, conforme o Quadro 8.

Quadro 8 – Perfil e espectros utilizados nas cargas do sistema IEEE 33 barras

Fonte 1		Fonte 2		Fonte 3	
Perfil	Espectro	Perfil	Espectro	Perfil	Espectro
Residencial	Retificador	Rural	Inversor	Industrial	HVDC

Fonte: Elaborado pelo autor (2020).

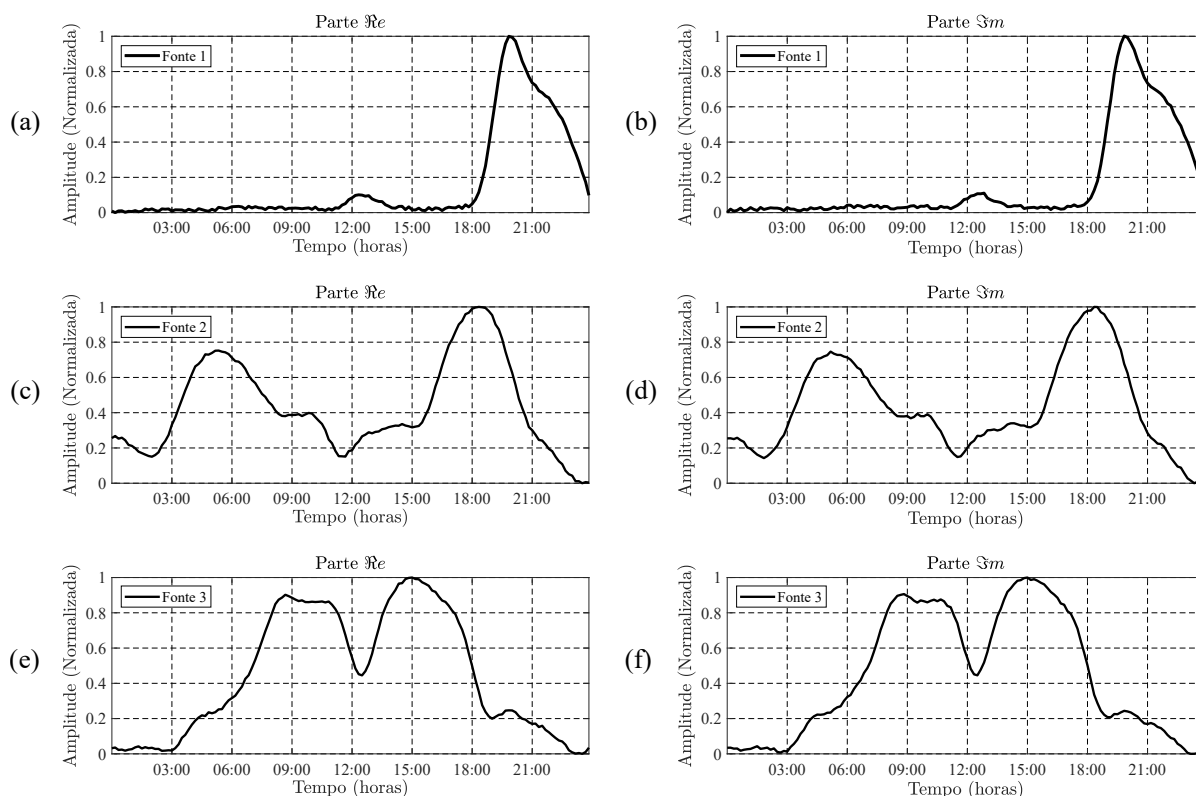
Para este ensaio, utilizou-se um perfil residencial (fonte 1) (FRANCISQUINI, 2006; QUEIROZ, 2011), outro de um consumidor rural (fonte 2) (MARQUES, 2014) e um perfil industrial (fonte 3) (MARQUES, 2014). Essas fontes foram conectadas nas barras 23, 17 e 31, respectivamente. Os coeficientes de correlação entre as fontes utilizadas estão presentes na Tabela 6. Finalmente, a Figura 52 apresenta as partes reais e imaginárias que foram injetadas no sistema, da seguinte forma: a Figura 52 (a) é a parte real do perfil residencial, a Figura 52 (b) é a parte imaginária do perfil residencial, a Figura 52 (c) é a parte real do perfil rural, a Figura 52 (d) é a parte imaginária do perfil rural, a Figura 52 (e) é a parte real do perfil industrial e a Figura 52 (f) é a parte imaginária do perfil industrial.

Tabela 6 – Coeficiente de Correlação entre os perfis com o sistema IEEE 33 barras

s_1s_2	s_1s_3	s_2s_3
0,83 %	37,75 %	12,29 %

Fonte: Elaborada pelo autor (2020).

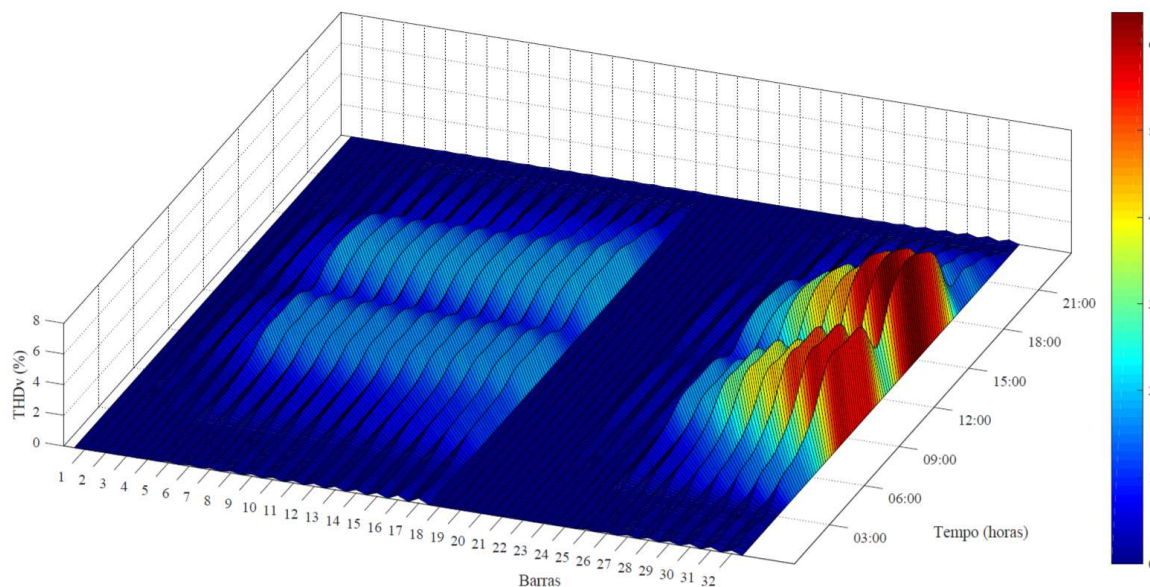
Figura 52 – Perfis utilizados



Fonte: Elaborado pelo autor (2020).

O Gráfico 28 apresenta a evolução do THDv ao longo do dia com a utilização dos perfis. Neste caso, a violação da porcentagem de 5% é observada nas barras 30, 31 e 32.

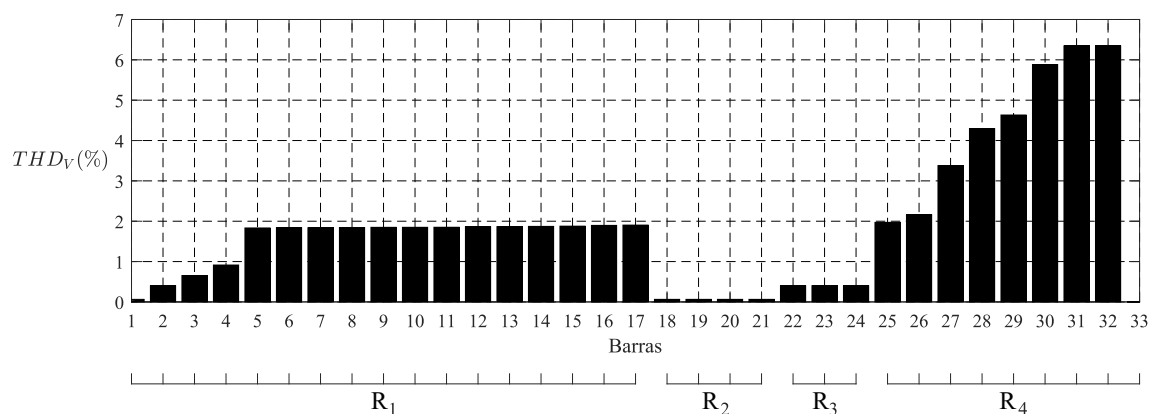
Gráfico 28 – THD de tensão ao longo do dia para o IEEE 33 barras.



Fonte: Elaborado pelo autor (2020).

Como o sistema é desequilibrado, a referência para o ensaio foi a fase A, visto que, o desequilíbrio entre as fases não promoveu uma diferença muito significativa entre as fases. Assim, o THD de tensão observado, Gráfico 29, corresponde ao THDv da fase A. Ao observar o THDv, percebe-se um conjunto de valores constantes entre os barramentos 5 e 17, outro conjunto entre as barras 22 e 24, e outro entre 31 e 32. É válido salientar que esses conjuntos de valores constantes contemplam os barramentos de conexão das cargas, entretanto, não há uma discrepância de valores que indique exatamente a barra de conexão das mesmas.

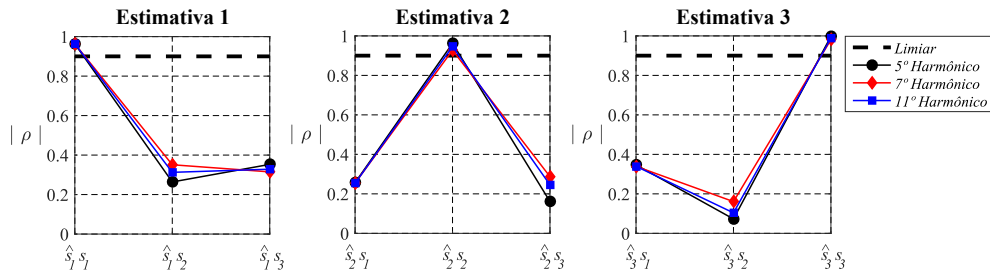
Gráfico 29 – THD de tensão no horário de pico para o IEEE 33 barras



Fonte: Elaborado pelo autor (2020).

Como as fontes possuem perfis diferentes, ou seja, são pouco correlacionados, a filtragem é capaz de melhorar levemente a estimação, mas não necessita de um intervalo de confiança para determinar a janela do filtro. Assim, uma janela de $N = 3$ mostrou-se mais que suficiente para uma estimação de qualidade, conforme a Figura 53, entre todas as três ordens harmônicas.

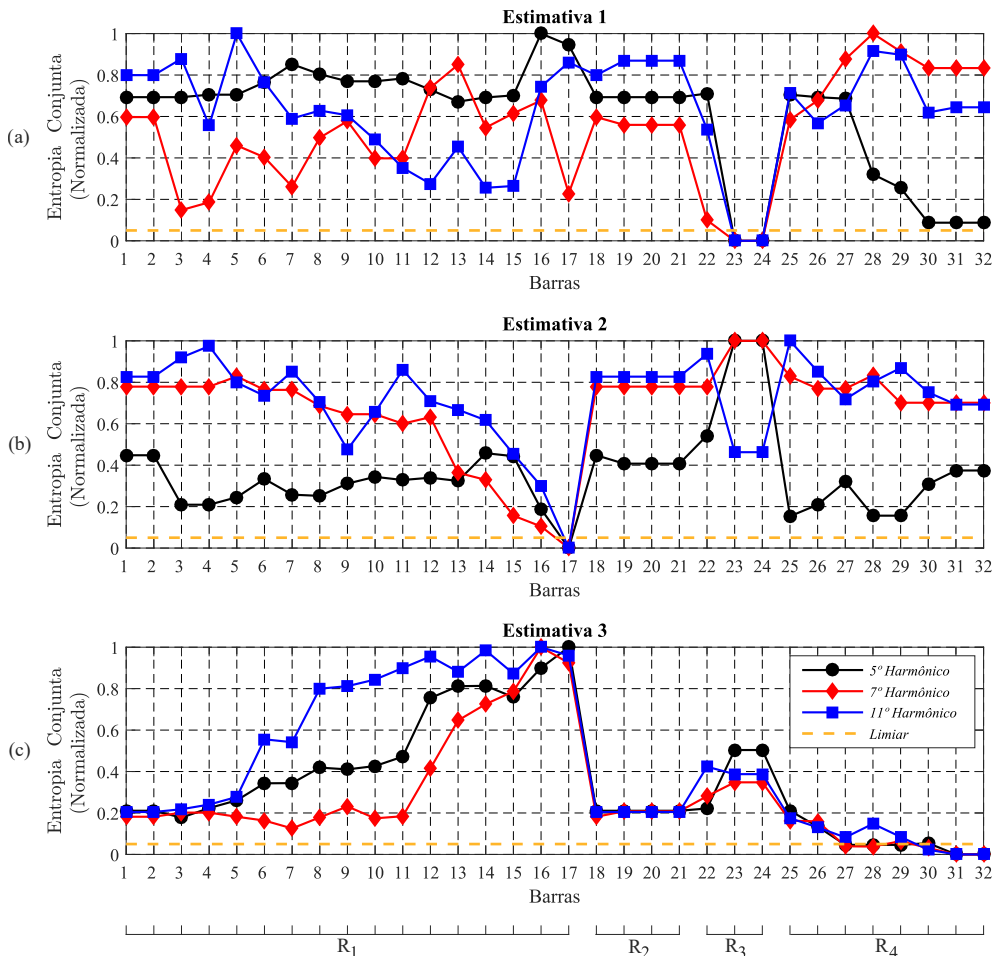
Figura 53 – Correspondência via coeficiente de correlação para as três estimativas



Fonte: Elaborado pelo autor (2020).

O resultado de entropia conjunta para as estimativas está ilustrado na Figura 54.

Figura 54 – Resultados de localização das fontes harmônicas do sistema IEEE 33 barras



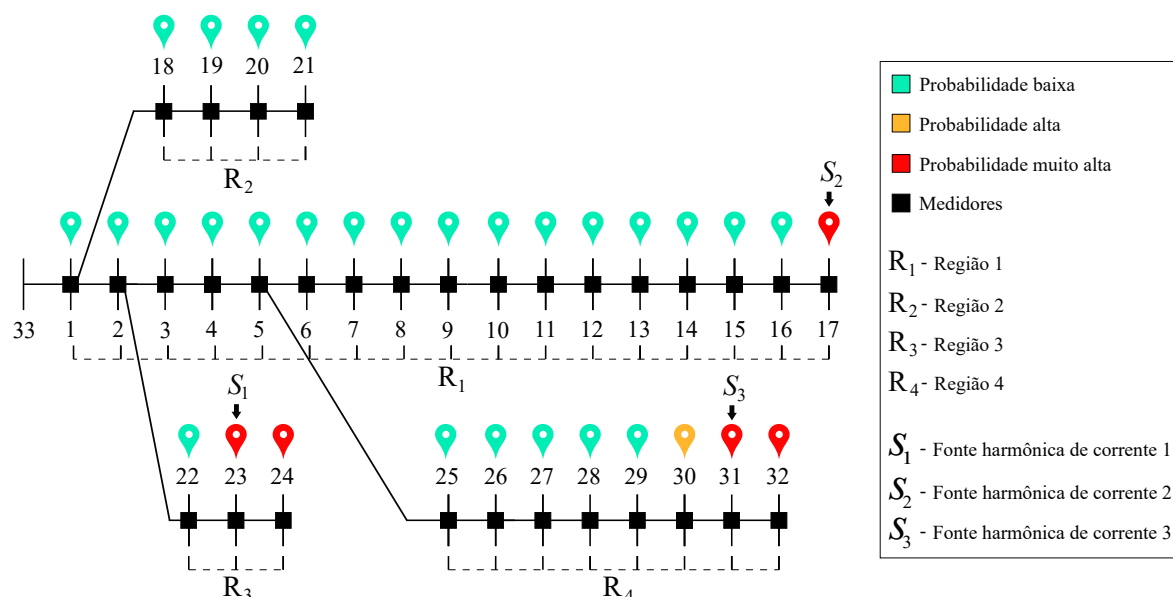
Fonte: Elaborado pelo autor (2020).

Conforme a Figura 54 (a), os mínimos são indicados nas barras 23 e 24. Entretanto, a barra 24 é uma barra de final de ramal. Tal resultado está ilustrado na Figura 55, através do mapa de probabilidades.

Outro resultado de entropia conjunta, para a estimativa 2, referente ao 5º, 7º e 11º harmônico, é representado através da Figura 54 (b). O mapa para esse resultado está ilustrado na Figura 55. Neste caso, a fonte harmônica de corrente já está na barra de fim de ramal. Este caso é uma exemplo de situação semelhante ao que o trabalho proposto por Zang et al. (2016) apresenta, isto é, com fontes harmônicas de corrente conectadas em barras de final de ramal. Além disso, utiliza-se a entropia condicional, outra metodologia, para os resultados de localização. E no caso do presente trabalho, a entropia conjunta apresenta os mesmos resultados satisfatórios que a referência na etapa de localização nas mesmas condições, isto é, a fonte harmônica de corrente conectada na barra de final de ramal.

O resultado de entropia conjunta para a estimativa 3, referente aos três harmônicos estudados, está ilustrado na Figura 54 (c). Como a fonte harmônica de corrente não está conectada numa barra de fim de ramal, novamente, a entropia conjunta indica duas localidades, as barras 31 e 32, conforme a Figura 55. Além disso, esse foi o único caso em que duas ordens harmônicas estiveram abaixo do limiar resultando na cor laranja para a barra 30.

Figura 55 – Mapa para as localizações do teste com o IEEE 33 barras



localização via o primeiro mínimo. Pois o mínimo acontece na barra de conexão e continua ocorrendo até a barra de final de ramal.

4.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo, foram apresentados os resultados obtidos com a aplicação da metodologia de localização de fontes harmônicas de corrente nos SEPs. Para isso, dois sistemas foram utilizados. Inicialmente, utilizou-se um sistema IEEE 14 barras de transmissão, e por fim, um sistema IEEE 33 barras de distribuição. Para o sistemas IEEE 14 barras, a metodologia foi testada no âmbito de várias correlações diferentes, entre os perfis lentos, o que comprovou a necessidade de um processo de filtragem para perfis altamente correlacionados. Além disso, a proposição de um limiar (80%) para a entropia conjunta no intervalo de confiança com os CPs promoveu a percepção de que quando a janela está fora do mesmo, a dependência estatística não é minimizada, ocasionando a obtenção de resultados ruins para perfis altamente correlacionados. Em uma situação dessas, os resultados assemelham-se a não utilização da filtragem, tal consequência impacta diretamente nas estimativas e, por conseguinte, na localização. Outro resultado importante obtido foi quando o modelo da distribuição de probabilidade do perfil rápido foi alterado. Neste aspecto, os resultados salientaram a robustez do método para a maioria dos casos, inclusive em situações equivalentes às da referência base. Finalmente, para o IEEE 14 barras, as posições das fontes foram trocadas para outras localidades da topologia do sistema. Em tais testes o método foi capaz de acertar com precisão as barras cujas fontes estavam conectadas, mesmo nas localidades mais diversas. Como último teste, o sistema IEEE 33 barras foi utilizado com o emprego de três fontes harmônicas de corrente. Essa situação insere um nível maior de dificuldade para separação das fontes. Entretanto, foi possível obter coeficientes de correlação altos para as três estimativas o que confirma com um bom resultado de estimação, ao passo que, a localização necessitou de uma maior interpretação para os resultados de entropia conjunta. Porém, mesmo nessa situação, ainda é possível inferir sobre a posição das fontes harmônicas de corrente.

5 CONCLUSÕES

O presente trabalho apresentou uma metodologia relacionada à QEE cuja finalidade reside em determinar as posições de fontes poluidoras, sobretudo, fontes harmônicas. A partir de técnicas cuja filosofia advém da pesquisa em BSS, isto é, parte-se da premissa de que as fontes harmônicas não são conhecidas, bem como o SEP, e assim, desenvolve-se uma ferramenta de detecção. Esse é um modelo de misturas, em que, somente as misturas são conhecidas, o que corresponde às tensões. Pensando em uma situação em que a complexidade computacional seja elevada, utiliza-se a PCA, uma técnica multivariada cujo objetivo é a redução de dimensionalidade dos dados. A ICA fica responsável por estimar os perfis das fontes harmônicas de corrente a partir da redução de dimensionalidade, oriunda da PCA. Além disso, para promover condições estatisticamente favoráveis à ICA, antes da aplicação da mesma, um processo de filtragem é estabelecido, em que, busca-se maximizar a independência estatística entre os CPs. Neste contexto, o trabalho está interessado na estimação dos perfis lentos. Então, com base na entropia conjunta, uma técnica da área da teoria da informação, buscam-se o valores de mínimo, que correspondem a maior probabilidade de conexão das fontes.

Para verificar o desempenho da metodologia apresentada, dois sistemas com características topológicas distintas foram utilizados. Inicialmente, um sistema de transmissão, o IEEE 14 barras, que possui aplicabilidade interessante, visto que, os medidores ficariam em subestações. Em outro ponto de vista, um sistema de distribuição também foi testado. Neste caso, sua utilização é válida porque há o interesse em verificar a técnica em situações onde a dimensão do sistema é maior, então, empregou-se o IEEE 33 barras.

Para o sistema IEEE 14 barras, vários testes foram efetuados. A metodologia foi verificada quanto à evolução do coeficiente de correlação, e Neste aspecto, concluiu-se que para perfis lentos com alta correlação a filtragem é indispensável. Um intervalo de confiança foi determinado para que a independência estatística entre os CPs fosse maximizada. Notou-se também que a janela dentro desse intervalo pode ser determinada a partir de um limiar. Ainda no mesmo sistema, os testes que apresentaram modelagens diferentes para a variação rápida confirmaram a robustez do método, além disso, resultados tão bons quanto a referência base, da etapa de estimação, foram obtidos. Finalmente, as fontes harmônicas de corrente foram conectadas em barras de diferentes localizações da topologia do SEP. Neste caso, a entropia conjunta mostrou-se eficaz na indicação das posições mais variadas em que fontes estavam conectadas.

Para o sistemas IEEE 33 barras, três fontes foram empregadas. Essa situação é complexa do ponto de vista de separabilidade entre as fontes. Quanto menor o número de fontes mais fácil é a separação entre elas. Entretanto, mesmo com um número maior dessas fontes os coeficientes de correlação confirmaram que o método é robusto para estimar

os perfis lentos, mesmo nessa situação adversa. Em comparação com a referência que utiliza a entropia condicional, uma das fontes foi conectada em uma barra de final de ramal e obteve resultados tão bons quanto. Entretanto, com a finalidade de testar mais assiduamente o desempenho do método, as outras fontes foram conectadas em barras que não fossem as de fim de ramal. Neste caso, conclui-se que é possível determinar a posição das fontes harmônicas de corrente através do primeiro mínimo de entropia.

Para os dois sistemas, um limiar de 5% foi proposto para que fosse possível combinar os resultados de entropia conjunta para as três ordens harmônicas analisadas. A partir desse limiar, um mapa de probabilidades é desenvolvido para apresentar as mais prováveis posições de conexão dessas fontes no sistema. No contexto de localização, a entropia conjunta apresentou um desempenho relevante se comparada às ferramentas semelhantes presentes na literatura, porém, sua característica primordial é ser mais generalista, sendo uma vantagem em relação as demais entropias e metodologias de localização.

Além disso, foi concluído que observar o THD de tensão, ao longo do dia e pontualmente, não é o suficiente para determinar: tanto a quantidade de fontes harmônicas quanto a localização das mesmas. Sendo necessário implementar um método que possa determinar a posição dessas cargas.

Finalmente, pode-se concluir que as principais contribuições deste trabalho são: a união da metodologia de estimação baseada no trabalho proposto por Gursoy (2007) com métodos estatísticos (entropia conjunta), para determinar a localização das fontes harmônicas de corrente; a determinação de um intervalo de confiança para a etapa de estimação dos perfis lentos, o que não ocorre no trabalho proposto por Gursoy (2007); a utilização de uma ferramenta mais generalista na obtenção das posições das fontes harmônicas (entropia conjunta), se comparado ao trabalho proposto por Zang et al. (2016); a análise dos resultados em situações mais adversas em relação ao trabalho proposto por Zang et al. (2016), com posições de conexão além das de final de ramal e um limiar (5%) que possibilite refinar os resultados de localização.

5.1 TRABALHOS FUTUROS

Como possíveis temas para trabalhos futuros e melhorias, as possíveis investigações são sugeridas:

- Avaliar o desempenho da metodologia com um número de amostras mais significativo, visto que, foram utilizadas somente 144 amostras e métodos estatísticos, como a ICA, tendem a melhorar seu desempenho com um número maior de informação;
- Avaliar a influência do perfil lento dentro do perfil rápido, visto que, a filtragem é não ideal e não elimina 100% o perfil lento do perfil rápido;

- Estudar outras abordagens de filtragem;
- Avaliar a presença de filtros conectados no sistema;
- Avaliar o desempenho da metodologia em sistemas diferentes, por exemplo o IEEE 69 barras, para verificar se o comportamento da ferramenta de localização se mantém em sistemas de distribuição distintos;
- Avaliar outras ferramentas estatísticas na etapa de localização;
- Avaliar a alocação de medidores, bem como, a redução dos mesmos;
- Recuperar as escalas das fontes;
- Implementar versões online das etapas algorítmicas visando a estimação e a localização dinâmica das fontes;
- Quantificar a probabilidade da fonte estar conectada no ramal através de mais rodadas de teste;
- Realizar mais simulações para verificar o comportamento do THD de tensão, com o intuito de documentar sua ineficácia em detectar todas as fontes;
- Relacionar os resultados entre os dois tipos de sistema, 14 e 33 barras;
- Verificar a metodologia em sistemas multi alimentados;
- Utilizar algoritmos online para atacar o problema de situações em que utilizam-se fasores dinâmicos.

REFERÊNCIAS

- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Curvas de Carga**. Brasília, 2012. Disponível em: <https://www.aneel.gov.br/>. Acesso em: 10 fev. 2020.
- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional**. Brasília, 2016. Disponível em: <https://www.aneel.gov.br/modulo-8>. Acesso em: 10 fev. 2020.
- AL-SAEGH, A. Comparison of complex-valued independent component analysis algorithms for eeg data. **Iraqi Journal for Electrical And Electronic Engineering**, Basrah University, v. 15, n. 1, p. 1–12, 2019.
- ALMEIDA, C. F. M. **Fontes distribuídas de harmônicos em sistemas elétricos de potência**. Tese (Doutorado) — Universidade de São Paulo, São Paulo, 2011.
- ARRILLAGA, J.; BOLLEN, M. H.; WATSON, N. R. Power quality following deregulation. **Proceedings of the IEEE**, IEEE, v. 88, n. 2, p. 246–261, 2000.
- ARSENEAU, R.; OUELLETTE, M. The effects of supply harmonics on the performance of compact fluorescent lamps. **IEEE Transactions on Power Delivery**, IEEE, v. 8, n. 2, p. 473–479, 1993.
- BAGHZOUZ, Y. *et al.* Time-varying harmonics. i. characterizing measured data. **IEEE Transactions on Power Delivery**, IEEE, v. 13, n. 3, p. 938–944, 1998.
- BARTLETT, M. S.; MOVELLAN, J. R.; SEJNOWSKI, T. J. Face recognition by independent component analysis. **IEEE Transactions on neural networks**, IEEE, v. 13, n. 6, p. 1450–1464, 2002.
- BEIDES, H. M.; HEYDT, G. T. Dynamic state estimation of power system harmonics using kalman filter methodology. **IEEE Transactions on Power Delivery**, IEEE, v. 6, n. 4, p. 1663–1670, 1991.
- BELOUHRANI, A. *et al.* A blind source separation technique using second-order statistics. **IEEE Transactions on signal processing**, IEEE, v. 45, n. 2, p. 434–444, 1997.
- BOLLEN, M. H.; GU, I. **Signal processing of power quality disturbances**. New York: John Wiley & Sons, 2006. v. 30.
- BROWN, R. G. **Introduction to random signal analysis and Kalman filtering**. New York: Wiley, 1983. v. 8.
- CARDOSO, J.-F.; SOULOUMIAC, A. Blind beamforming for non-gaussian signals. In: IET. **IEE Proceedings F - Radar and Signal Processing**. Paris, 1993. v. 140, n. 6, p. 362–370.
- CAVALLINI, A.; MONTANARI, G.; CACCIARI, M. Stochastic evaluation of harmonics at network buses. **IEEE Transactions on Power Delivery**, IEEE, v. 10, n. 3, p. 1606–1613, 1995.

- CHERRY, E. C. Some experiments on the recognition of speech, with one and with two ears. **The Journal of the acoustical society of America**, ASA, v. 25, n. 5, p. 975–979, 1953.
- CICHOCKI, A.; AMARI, S. **Adaptive blind signal and image processing: learning algorithms and applications**. Chichester: John Wiley & Sons, 2002.
- COMON, P. Blind identification and source separation in 2/spl times/3 under-determined mixtures. **IEEE Transactions on Signal Processing**, v. 52, p. 11–22, 2004.
- COMON, P.; JUTTEN, C.; HERAULT, J. Blind separation of sources, part ii: Problems statement. **Signal processing**, Elsevier, v. 24, n. 1, p. 11–20, 1991.
- COVER, T. M.; THOMAS, J. A. **Elements of information theory**. 2. ed. New Jersey: John Wiley & Sons, 2012.
- DAI, C.; HU, X. Independent component analysis based algorithms for high-density electromyogram decomposition: Experimental evaluation of upper extremity muscles. **Computers in biology and medicine**, Elsevier, v. 108, p. 42–48, 2019.
- DARÓCZY, Z. Generalized information functions. **Information and control**, Elsevier, v. 16, n. 1, p. 36–51, 1970.
- DITTA, A. *et al.* Principal components analysis of raman spectral data for screening of hepatitis c infection. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, Elsevier, p. 117173, 2019.
- DU, Z. P. *et al.* Identification of harmonic sources of power systems using state estimation. **IEE Proceedings-Generation, Transmission and Distribution**, IET, v. 146, n. 1, p. 7–12, 1999.
- DUGGAN, E.; MORRISON, R. Probabilistic methods applied to harmonic distortion. In: IET. **Third International Conference on Probabilistic Methods Applied to Electric Power Systems**. London, 1991. p. 43–46.
- FANDI, M. **Reator Controlado a Tiristores: Modelagem para análises sob condições de operação em regimes permanente e transitório**. Dissertação (Mestrado) — Universidade Federal de Uberlândia, Uberlândia, 1998.
- FARHOODNEA, M.; MOHAMED, A.; SHAREEF, H. A new method for determining multiple harmonic source locations in a power distribution system. In: IEEE. **2010 IEEE International Conference on Power and Energy**. Kuala Lumpur, 2010. p. 146–150.
- FENG, D. *et al.* Identification of disturbance sources based on random forest model. In: IEEE. **2018 International Conference on Power System Technology**. Guangzhou, 2018. p. 3370–3375.
- FERNANDES, R. A.; SILVA, I. N.; OLESKOVICZ, M. Identificação de cargas lineares e não-lineares em sistemas elétricos residenciais usando técnicas para seleção de atributos e redes neurais artificiais. **SBA-Controle & Automação**, v. 21, n. 4, p. 389–405, 2010.
- FONTANA, D. R.; MARIM, L. Sistema de autenticação/identificação pessoal biométrica através da palma da mão com o auxílio de redes neurais artificiais. In: ITA. **140 Encontro de Iniciação Científica e Pós-Graduação do ITA**. São José dos Campos, 2009.

FRANCISQUINI, A. A. **Estimação de Curvas de Carga em Pontos de Consumo e em Transformadores de Distribuição**. Tese (Doutorado) — Universidade Estadual Paulista, 2006.

GALHARDO, M.; PINHO, J. Mutual influence between harmonics and nonlinear loads. **IEEE Latin America Transactions**, v. 6, n. 7, p. 608–617, 2008.

GRAY, W.; HAYDOCK, F. Industrial power quality considerations when installing adjustable speed drive systems. In: IEEE. **Cement Industry Technical Conference, 1995. XXXVII Conference Record., 1995 IEEE**. San Juan, 1995. p. 17–33.

GREWAL, M. S.; ANDREWS, A. P. **Kalman filtering: Theory and practice with matlab**. New York: John Wiley & Sons, 2014.

GURSOY, E. **Independent component analysis for harmonic source identification in electric power systems**. Tese (Doutorado) — Drexel University, 2007.

GURSOY, E.; NIEBUR, D. Harmonic load identification using complex independent component analysis. **IEEE Transactions on Power Delivery**, IEEE, v. 24, n. 1, p. 285–292, 2008.

HARTANA, R.; RICHARDS, G. Harmonic source monitoring and identification using neural networks. **IEEE Transactions on Power Systems**, IEEE, v. 5, n. 4, p. 1098–1104, 1990.

HARTANA, R. K.; RICHARDS, G. G. Constrained neural network-based identification of harmonic sources. **IEEE transactions on industry applications**, IEEE, v. 29, n. 1, p. 202–208, 1993.

HAVRDA, J.; CHARVÁT, F. Quantification method of classification processes. concept of structural α -entropy. **Kybernetika**, Institute of Information Theory and Automation AS CR, v. 3, n. 1, p. 30–35, 1967.

HERAULT, J.; JUTTEN, C. Space or time adaptive signal processing by neural network models. In: AIP. **AIP conference proceedings**. New York, 1986. v. 151, n. 1, p. 206–211.

HEYDT, G. T. Identification of harmonic sources by a state estimation technique. **IEEE Transactions on Power Delivery**, IEEE, v. 4, n. 1, p. 569–576, 1989.

HOLTZ, J. Pulsewidth modulation for electronic power conversion. **Proceedings of the IEEE**, IEEE, v. 82, n. 8, p. 1194–1214, 1994.

HONG, T. S.; KIM, J. Fpga implementation of jade ica algorithm. In: IEEE. **2015 International SoC Design Conference**. Gyungju, 2015. p. 31–32.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of educational psychology**, Warwick & York, v. 24, n. 6, p. 417, 1933.

HYVÄRINEN, A. Independent component analysis for time-dependent stochastic processes. In: SPRINGER. **International Conference on Artificial Neural Networks**. London, 1998. p. 135–140.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. **Independent component analysis**. New York: John Wiley & Sons, 2001. v. 46.

JIANG, X. *et al.* A complex network analysis approach for identifying air traffic congestion based on independent component analysis. **Physica A**, Elsevier, v. 523, p. 364–381, 2019.

JR, Â. S. *et al.* O impacto do uso de microcomputadores na qualidade da energia elétrica. In: SBQEE. **Seminário Brasileiro sobre Qualidade da Energia Elétrica**. São Lourenço, 1997.

JUNG, T.-P. *et al.* Removing electroencephalographic artifacts by blind source separation. **Psychophysiology**, Cambridge University Press, v. 37, n. 2, p. 163–178, 2000.

JUTTEN, C.; HERAULT, J. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. **Signal processing**, Elsevier, v. 24, n. 1, p. 1–10, 1991.

KAMBHATLA, N.; LEEN, T. K. Dimension reduction by local principal component analysis. **Neural computation**, MIT Press, v. 9, n. 7, p. 1493–1516, 1997.

KARIMZADEH, F.; ESMAEILI, S.; HOSSEINIAN, S. H. A novel method for noninvasive estimation of utility harmonic impedance based on complex independent component analysis. **IEEE Transactions on Power Delivery**, IEEE, v. 30, n. 4, p. 1843–1852, 2015.

KETABI, A.; SHEIBANI, M. R.; NOSRATABADI, S. M. Power quality meters placement using seeker optimization algorithm for harmonic state estimation. **International Journal of Electrical Power & Energy Systems**, Elsevier, v. 43, n. 1, p. 141–149, 2012.

KHALID, K. *et al.* Performance comparison of artificial intelligence techniques for non-intrusive electrical load monitoring. **Bulletin of Electrical Engineering and Informatics**, v. 7, n. 2, p. 143–152, 2018.

KOTZ, S.; KOZUBOWSKI, T.; PODGORSKI, K. **The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance**. 1. ed. New York: Springer Science & Business Media, 2012.

KUMAR, A.; DAS, B.; SHARMA, J. Determination of location of multiple harmonic sources in a power system. **International journal of electrical power & energy systems**, Elsevier, v. 26, n. 1, p. 73–78, 2004.

KUSKO, A.; THOMPSON, M. T. **Power quality in electrical systems**. New York: McGraw-Hill, 2007. v. 23.

LEON-GARCIA, A. **Probability, statistics, and random processes for electrical engineering**. 3. ed. Upper Saddle River: Pearson Education, 2017.

LIMA, M. A. A. **Análise de componentes independentes aplicada em distúrbios de qualidade de energia**. Dissertação (Mestrado) — Universidade Federal de Juiz de Fora, Juiz de Fora, 2009.

LIMA, M. A. A. **Uma Nova Metodologia para Análise de Qualidade da Energia Elétrica sob Condições de Ocorrência de Múltiplos Distúrbios**. Tese (Doutorado) — Universidade de São Paulo, 2013.

LIN, W.-M. *et al.* Multiple harmonic source detection and equipment identification with cascade correlation network. **IEEE Transactions on Power Delivery**, IEEE, v. 20, n. 3, p. 2166–2173, 2005.

MA, H.; GIRGIS, A. A. Identification and tracking of harmonic sources in a power system using a kalman filter. **IEEE Transactions on Power Delivery**, IEEE, v. 11, n. 3, p. 1659–1665, 1996.

MAIONE, C.; JR, F. B.; BARBOSA, R. M. Predicting the botanical and geographical origin of honey with multivariate data analysis and machine learning techniques: A review. **Computers and electronics in agriculture**, Elsevier, v. 157, p. 436–446, 2019.

MAO, K. Identifying critical variables of principal components for unsupervised feature selection. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 35, n. 2, p. 339–344, 2005.

MARQUES, M. d. S. **Metodologia para modelagem de curvas típicas de demanda elétrica utilizando redes neurais artificiais considerando variáveis climáticas**. Dissertação (Mestrado) — Universidade Federal do Pampa, 2014.

MATTOS, R. S. d.; VEIGA, Á. Otimização de entropia: implementação computacional dos princípios maxent e minxent. **Pesquisa Operacional**, SciELO Brasil, v. 22, n. 1, p. 37–59, 2002.

MEINECKE, F. C.; HARMELING, S.; MÜLLER, K.-R. Robust ica for super-gaussian sources. In: SPRINGER. **International Conference on Independent Component Analysis and Signal Separation**. Berlin, 2004. p. 217–224.

MELO, I. D. *et al.* Harmonic state estimation for distribution systems based on optimization models considering daily load profiles. **Electric Power Systems Research**, Elsevier, v. 170, p. 303–316, 2019.

MELO, I. D. d. **Estimação estática de estados harmônicos em redes trifásicas de distribuição monitoradas por PMUs**. Tese (Doutorado) — Universidade Federal de Juiz de Fora, Juiz de Fora, 2018.

MELO, L. S. *et al.* Study of the behavior of an electric power distribution system under the influence of photovoltaic generation. In: IEEE. **2018 Simposio Brasileiro de Sistemas Eletricos**. Niterói, 2018. p. 1–6.

MENDES, T. M. *et al.* Monitoramento multidimensional de qualidade de energia elétrica para smart grids. In: CBQEE. **Conferência Brasileira sobre Qualidade da Energia Elétrica**. Campina Grande, 2015. p. 1–4.

MITRA, S. K.; KUO, Y. **Digital signal processing: a computer-based approach**. New York: McGraw-Hill, 2006. v. 2.

MONTEIRO, H. L. M. **Método de estimação de impedância utilizando a injeção de pequenos sinais**. Tese (Doutorado) — Universidade Federal de Juiz de Fora, Juiz de Fora, 2018.

MORADIFAR, A.; FOROUD, A. A.; FIROUZJAH, K. G. Comprehensive identification of multiple harmonic sources using fuzzy logic and adjusted probabilistic neural network. **Neural Computing and Applications**, Springer, v. 31, n. 1, p. 543–556, 2019.

MORRIS, R. M. Do hebb: The organization of behavior. **Brain research bulletin**, Elsevier Science, v. 50, n. 5-6, p. 473, 1999.

MORRISON, R.; DUGGAN, E. Converting data into information (power system harmonic analysis). In: IET. **IEE Colloquium on Electronics in Power Systems Measurement**. London, 1988. p. 3–1.

MORRISON, R.; DUGGAN, E. Long term monitoring of power systems harmonics. In: IET. **IEE Colloquium on Three Phase LV Industrial Supplies**. London, 1993. p. 2–1.

MURATA, N.; IKEDA, S.; ZIEHE, A. An approach to blind source separation based on temporal structure of speech signals. **Neurocomputing**, Elsevier, v. 41, n. 1-4, p. 1–24, 2001.

NISHITSUJI, Y. *et al.* Tectonic tremor characterized by principal-component analysis in the vicinity of central chile and argentina. **Journal of South American Earth Sciences**, Elsevier, p. 102178, 2019.

O'GRADY, P. D.; PEARLMUTTER, B. A.; RICKARD, S. T. Survey of sparse and non-sparse methods in source separation. **International Journal of Imaging Systems and Technology**, Wiley Online Library, v. 15, n. 1, p. 18–33, 2005.

ONAT, N. C.; KUCUKVAR, M.; AFSHAR, S. Eco-efficiency of electric vehicles in the united states: a life cycle assessment based principal component analysis. **Journal of cleaner production**, Elsevier, v. 212, p. 515–526, 2019.

OPERADOR NACIONAL DO SISTEMA ELÉTRICO. **Boletim Mensal de Geração Eólica de Agosto de 2018**. Brasília, 2018. Disponível em: <https://www.ons.org.br/>. Acesso em: 10 fev. 2020.

PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, Taylor Francis, v. 2, n. 11, p. 559–572, 1901.

PEEBLES, P. Z. **Probability, random variables, and random signal principles**. New York: McGraw-Hill, 2001. v. 3.

PEREIRA, J. G. **Impacto da Modelagem da Rede em Função da Frequência na Análise Harmônica de Sistemas com Geração Eólica**. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

PILVELAIT, B.; ORTMEYER, T.; GRIZER, M. Harmonic evaluation of inductor location in a variable speed drive. In: IEEE. **ICHPS V International Conference on Harmonics in Power Systems**. Atlanta, 1992. p. 267–271.

POH, M.-Z.; MCDUFF, D. J.; PICARD, R. W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. **Optics express**, Optical Society of America, v. 18, n. 10, p. 10762–10774, 2010.

PRESS, W. H. *et al.* **Numerical recipes 3rd edition**: The art of scientific computing. 3. ed. New York: Cambridge university press, 2007.

PULIMERA, K.; RAJAN, P. K. Independent component analysis for harmonic source estimation from piecewise constant parameter mixed measurements. In: **IEEE. 2011 IEEE 43rd Southeastern Symposium on System Theory**. Auburn, 2011. p. 81–86.

QUEIROZ, P. V. S. d. **Mensuração do consumo de energia elétrica: algoritmo para detecção de potenciais usuários da termoacumulação como alternativa para o deslocamento de carga**. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2011.

RAY, P. K.; MOHANTY, A.; PANIGRAHI, T. Power quality analysis in solar pv integrated microgrid using independent component analysis and support vector machine. **Optik**, Elsevier, v. 180, p. 691–698, 2019.

RÉNYI, A. *et al.* On measures of entropy and information. **Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability**, p. 531–551, 1961.

SAXENA, D.; VERMA, K.; SINGH, S. Power quality event classification: an overview e key issues. **International journal of engineering, science e technology**, MultiCraft Limited, v. 2, n. 3, p. 186–199, 2010.

SHANNON, C. E. A mathematical theory of communication. **Bell system technical journal**, Wiley Online Library, v. 27, n. 3, p. 379–423, 1948.

SHAO, R. *et al.* The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform. **Measurement**, Elsevier, v. 54, p. 118–132, 2014.

SILVA, I. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. **Revista Controle & Automação**, v. 23, n. 5, p. 33–111, 2012.

SILVA, L. R. M. **Inteligência computacional aplicada à modelagem de cargas não-lineares e estimação de contribuição harmônica**. Dissertação (Mestrado) — Universidade Federal de Juiz de Fora, Juiz de Fora, 2012.

SRINIVASAN, D.; NG, W.; LIEW, A. Neural-network-based signature recognition for harmonic source identification. **IEEE Transactions on Power Delivery**, IEEE, v. 21, n. 1, p. 398–405, 2005.

THACKER, N.; LACEY, A. Tutorial: The kalman filter. **Imaging Science and Biomedical Engineering Division, Medical School, University of Manchester**, Citeseer, p. 61, 1998.

TSALLIS, C. Possible generalization of boltzmann-gibbs statistics. **Journal of statistical physics**, Springer, v. 52, n. 1-2, p. 479–487, 1988.

UJILE, A.; DING, Z. A dynamic approach to identification of multiple harmonic sources in power distribution systems. **International Journal of Electrical Power & Energy Systems**, Elsevier, v. 81, p. 175–183, 2016.

VARIZ, A. M. **Cálculo do Fluxo de Harmônicas em Sistemas Trifásicos Utilizando o Método de Injeção de Correntes**. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

VARIZ, A. M. *et al.* Cálculo do fluxo de harmônicos em sistemas de potência trifásicos utilizando o método de injeção de correntes com solução iterativa. **Sba: Controle & Automação Sociedade Brasileira de Automatica**, SciELO Brasil, v. 19, n. 2, p. 178–198, 2008.

VIGÁRIO, R. *et al.* Independent component approach to the analysis of eeg and meg recordings. **IEEE transactions on biomedical engineering**, IEEE, v. 47, n. 5, p. 589–593, 2000.

VRANIC, D. V.; SAUPE, D.; RICHTER, J. Tools for 3d-object retrieval: Karhunen-loeve transform and spherical harmonics. **2001 IEEE Fourth Workshop on Multimedia Signal Processing (Cat. No. 01TH8564)**, p. 293–298, 2001.

WOLD, S.; ESBENSEN, K.; GELADI, P. Principal component analysis. **Chemometrics and intelligent laboratory systems**, Elsevier, v. 2, n. 1-3, p. 37–52, 1987.

WU, X. *et al.* Blind source separation of functional mri scans of the human brain based on canonical correlation analysis. **Neurocomputing**, Elsevier, v. 269, p. 220–225, 2017.

YATES, R. D.; GOODMAN, D. J. **Probability and stochastic processes**. 3. ed. New York: John Wiley & Sons, 2014.

YOST, W. A. The cocktail party problem: Forty years later. **Binaural and spatial hearing in real and virtual environments**, p. 329–347, 1997.

YU, K. K.; WATSON, N.; ARRILLAGA, J. An adaptive kalman filter for dynamic harmonic state estimation and harmonic injection tracking. **IEEE Transactions on power delivery**, IEEE, v. 20, n. 2, p. 1577–1584, 2005.

ZANG, T. *et al.* Harmonic source localization approach based on fast kernel entropy optimization ica and minimum conditional entropy. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 18, n. 6, p. 214, 2016.

ZARZOSO, V.; COMON, P. Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size. **IEEE Transactions on Neural Networks**, IEEE, v. 21, n. 2, p. 248–261, 2010.

ZEBARDAST, A.; MOKHTARI, H. New method for assessing the utility harmonic impedance based on fuzzy logic. **IET Generation, Transmission & Distribution**, IET, v. 11, n. 10, p. 2448–2456, 2017.

APÊNDICE A – NOÇÕES DE PROBABILIDADE

Os tópicos seguintes são responsáveis por elencar todos os conceitos primordiais para o entendimento das entropias.

A.1 PROBABILIDADE DE EVENTOS

O conceito de entropia conjunta advém da ideia de probabilidade conjunta. Assim, é necessário passar pelos conceitos probabilísticos que os tangem.

A noção de probabilidade exprime uma ideia de frequência de eventos. No lançamento de dados, por exemplo, em 5 lançamentos quantas vezes a face com o número 9 ocorreu? Ou seja, o interesse reside na chance de ocorrência de um evento específico.

Na teoria de conjuntos, um espaço amostral S (conjunto universal) com resultados s (elementos de S), cujos eventos são denotados por A (conjuntos de elementos), implica que a probabilidade de cada evento é representada por $P[A]$ (YATES; GOODMAN, 2014). Neste caso, se a ideia de frequência de ocorrência for seguida, a probabilidade $P[A]$ pode ser definida por (A.1) e pela seguinte definição: dado um espaço amostral S (universo) e o evento A , a probabilidade $P[A]$ é:

$$P[A] = \frac{\text{número de elementos de } A}{\text{número de elementos de } S} \quad (\text{A.1})$$

Logo, a medida de probabilidade $P[\cdot]$ é uma função que mapeia evento em num espaço amostral para números reais de tal forma que atenda os seguintes axiomas:

1. **Axioma I:** Para qualquer evento A , $P[A] \geq 0$;
2. **Axioma II:** $P[S] = 1$;
3. **Axioma III:** Para eventos mutuamente exclusivos A_1, A_2, \dots , segue (A.2):

$$P[A_1 \cup A_2 \cup \dots] = P[A_1] + P[A_2] + \dots \quad (\text{A.2})$$

Assim, se $P[A] \approx 1$ significa que certamente A ocorrerá. Se $P[A] \approx 0$ é muito provável que A não ocorrerá quando o experimento ocorre. Caso $P[A] \approx 1/2$ pouco se sabe que A ocorre ou não.

A.1.1 Probabilidade Conjunta

É um cálculo que descreve a probabilidade de ocorrência conjunta de eventos. A probabilidade conjunta de dois eventos A e B é denominada $P[A \cap B]$ (PEEBLES, 2001), conforme (A.3):

$$P[A \cap B] = P[A] + P[B] - P[A \cup B] \quad (\text{A.3})$$

Caso os eventos sejam mutuamente exclusivos $A \cap B = \emptyset$, então $P[A \cap B] = P[\emptyset] = 0$.

A.1.2 Probabilidade Condicional

A probabilidade condicional $P[A|B]$ descreve a ocorrência de A quando se têm o conhecimento de que B (um outro evento) ocorreu, porém, não sabemos precisamente o resultado (YATES; GOODMAN, 2014). Logo, a probabilidade condicional pode ser formulada conforme (A.4) e pela definição seguinte: a probabilidade condicional de A dada a ocorrência do evento B é:

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad (\text{A.4})$$

No jargão estatístico para $P[A|B]$ diz-se “a probabilidade de A dado B ”.

A.1.3 Independência Estatística

Os eventos A e B são independentes se (e se somente se):

$$P[AB] = P[A]P[B] \quad (\text{A.5})$$

Se as probabilidades $P[A] \neq 0$ e $P[B] \neq 0$, a independência se reflete também na probabilidade condicional dos eventos:

$$P[A|B] = P[A], \quad (\text{A.6})$$

ou

$$P[B|A] = P[B] \quad (\text{A.7})$$

Neste caso, conforme (A.6), significa que B não altera as informações sobre A . O mesmo raciocínio é estabelecido para (A.7).

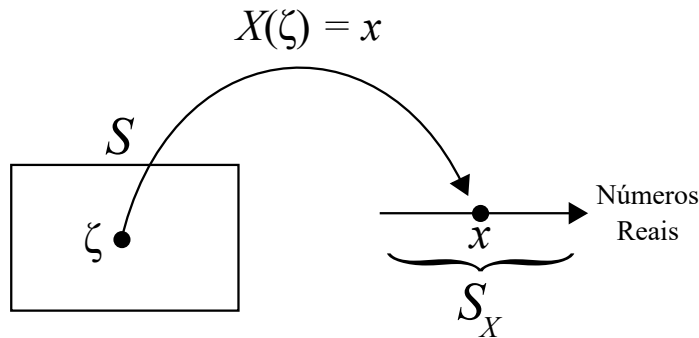
Outro fato é que, eventos independentes e disjuntos não são sinônimos. Para eventos disjuntos $P[AB] = 0$, ou seja, não há resultados em comum. Entretanto, a exceção ocorre quando para eventos independentes $P[A] = 0$ ou $P[B] = 0$.

A.2 VARIÁVEIS ALEATÓRIAS E FUNÇÕES DE PROBABILIDADE

Tratando-se de experimentos aleatórios, muitas vezes o resultado propriamente dito não é o foco de interesse. Neste caso, o interesse reside em alguma medida que diz respeito sobre o resultado. Em lançamentos de dados, por exemplo, talvez o interesse esteja em computar quantas vezes a face que contém o número 6 foi sorteada, não em um ordenamento específico de faces. Para essa experiência aleatória, uma medição pode atribuir um valor numérico ao resultado da experiência randômica. Por conseguinte, resultados aleatórios resultam em medições cujos valores são aleatórios.

Essa ideia está ilustrada na Figura 56, onde o espaço amostral S é o domínio da variável aleatória, o conjunto S_X de todos os valores assumidos por X é o intervalo da variável aleatória, assumindo um subconjunto do conjunto de todos os números reais. Neste caso, uma variável aleatória atribui um número $X(\zeta)$ a cada resultado ζ no espaço amostral S num dado experimento aleatório (LEON-GARCIA, 2017).

Figura 56 – Mapeamento de uma variável aleatória.



Fonte: Elaborado pelo autor (2020).

Uma variável aleatória genérica é definida formalmente como se segue: dado um experimento aleatório com espaço amostral S e a classe de eventos F . Uma variável aleatória X é uma função do espaço amostral S para R com a propriedade que o conjunto $A_b = \{\zeta : X(\zeta) \leq b\}$ está em F para todo b em R (LEON-GARCIA, 2017).

A entropia é uma medida que envolve conceitos das funções de probabilidade, em especial, da função massa de probabilidade. Para sua compreensão, é necessário compreender os conceitos interdependentes das funções: distribuição cumulativa e densidade de probabilidade.

1. Função Distribuição Cumulativa:

Uma função distribuição cumulativa (*f.d.c.*) de uma variável aleatória X é definida como a probabilidade do evento $\{X \leq x\}$ e conforme (A.8):

$$F_X(x) \triangleq P[X \leq x], \quad -\infty < x < \infty, \quad (\text{A.8})$$

a equação (A.8) simplesmente define a probabilidade de uma variável aleatória X tomar um valor no intervalo $(-\infty, \infty]$. Assim, o evento $\{X \leq x\}$ e a probabilidade associada a ele alteram-se conforme x se altera.

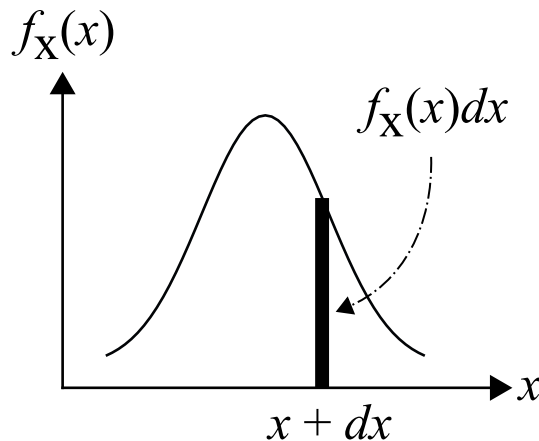
2. Função Densidade de Probabilidade:

Após a *f.d.c.* ser definida, é possível elaborar a formulação da função densidade de probabilidade. Logo, uma função densidade de probabilidade (*f.d.p.*) de uma variável aleatória X , caso exista, é por definição a derivada de $F_X(x)$, conforme (A.9):

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (\text{A.9})$$

Essa função representa a probabilidade de intervalos de largura infinitesimal conforme a Figura 57.

Figura 57 – Representação do intervalo infinitesimal.



Fonte: Elaborado pelo autor (2020).

Tomando $f_X(x)$ como uma função não negativa. A probabilidade de um intervalo $[a, b]$ é definida conforme (A.10) e tomando larguras infinitesimais:

$$P[a \leq x \leq b] = \int_a^b f_X(x) dx \quad (\text{A.10})$$

3. Função Massa de Probabilidade:

A função massa de probabilidade (*f.m.p.*) de X é o conjunto de probabilidades $p_X(x_k) = P[X = x_k]$ dos elementos em $S_X = [x_0, x_1, \dots, x_n]$. Além disso, a fdc de uma variável aleatória discreta pode ser definida como uma soma ponderada pela fmp em de funções degrau unitário, conforme (A.11):

$$F_X(x) = \sum_k p_X(x_k) u(x - x_k), \quad (\text{A.11})$$

cujo $p_X(x_k) = P[X = x_k]$ fornece a magnitude dos saltos da (*f.d.c.*).

Mas esse é o ponto de vista discreto, por parte do domínio contínuo temos a *f.d.p.*.

APÊNDICE B – FLUXO DE POTÊNCIA

Embora o fluxo de potência não seja o foco dessa dissertação, ele é parte importante para a execução da metodologia desenvolvida. Assim, é interessante que algumas considerações sejam refletidas acerca do mesmo.

Um fluxo de potência harmônico é um método de análise de harmônicos nos SEPs (MELO, 2018). É uma ferramenta potente que mimetiza o comportamento de um sistema atuando num ponto de operação (VARIZ *et al.*, 2008). Através de iterações, tal ferramenta, modifica a matriz jacobiana para que grandezas elétricas sejam determinadas a cada iteração.

Em geral, as modelagens partem de formulações matriciais intuitivas (VARIZ, 2006). Neste contexto, um sistema matricial trivial para a ordem harmônica específica h é dado por (B.1):

$$[I]^h = [Y]^h [V]^h, \quad (\text{B.1})$$

em que h representa a ordem harmônica de interesse, $[I]^h$ é o vetor com as correntes harmônicas de ordem h , $[Y]^h$ é a matriz admitância nodal para a ordem harmônica h , $[V]^h$ é o vetor de variáveis de estado com as componentes de tensão de ordem harmônica h . Assim, para um sistema com m -barras a matriz de admitância nodal na ordem harmônica específica h é representada por (B.2):

$$[Y]^h = \begin{bmatrix} Y_{11}^h & Y_{12}^h & \cdots & Y_{1k}^h & \cdots & Y_{1m}^h \\ Y_{21}^h & Y_{22}^h & \cdots & Y_{2k}^h & \cdots & Y_{2m}^h \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ Y_{k1}^h & Y_{k2}^h & \cdots & Y_{kk}^h & \cdots & Y_{km}^h \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ Y_{m1}^h & Y_{m2}^h & \cdots & Y_{mk}^h & \cdots & Y_{mm}^h \end{bmatrix} \quad (\text{B.2})$$

onde Y_{kk}^h é a admitância própria da k -ésima barra na frequência harmônica específica de ordem h e Y_{km}^h é a admitância mútua entre as barras k e m , também na frequência harmônica específica de ordem h . Então, sempre que se deseja estudar uma frequência, obtém-se um vetor de injeção de corrente, um vetor de tensão nodal e uma matriz de admitância nodal.

Partindo para uma simplificação, o problema descrito em (B.1), pode ser equivalente a um sistema de soluções triviais, ou seja (B.3):

$$\mathbf{A}x = \mathbf{b}, \quad (\text{B.3})$$

cujas matriz \mathbf{A} é assumida como quadrada, de componentes complexos ou reais e que admite inversa, x é o vetor objetivo, ou que se deseja solucionar (desconhecido) e \mathbf{b} é o vetor composto por variáveis complexas ou reais (conhecido).

Existem várias propostas para solucionar (B.3), por exemplo as famílias dos métodos: de Inversão explícita; diretos e iterativos ou indiretos (VARIZ, 2006).

Nessa dissertação, utiliza-se o Método de Injeção de Corrente Trifásico Harmônico (MICTH), desenvolvido por Variz (2006), que é resolvido através de métodos iterativos. O MICTH é um solução interessante por apresentar poucas iterações para convergência (MELO, 2018).

Neste contexto, o problema a ser resolvido pelo MICTH está brevemente descrito pela representação matricial (B.4):

$$\begin{bmatrix} [\Delta I]^0 \\ [\Delta I]^1 \\ [\Delta I]^2 \\ \vdots \\ [\Delta I]^h \end{bmatrix} = \begin{bmatrix} [J]^{00} & [J]^{01} & [J]^{02} & \dots & [J]^{0h} \\ [J]^{10} & [J]^{11} & [J]^{12} & \dots & [J]^{1h} \\ [J]^{20} & [J]^{21} & [J]^{22} & \dots & [J]^{2h} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ [J]^{h0} & [J]^{h1} & [J]^{h2} & \dots & [J]^{hh} \end{bmatrix} \begin{bmatrix} [\Delta V]^0 \\ [\Delta V]^1 \\ [\Delta V]^2 \\ \vdots \\ [\Delta V]^h \end{bmatrix} \quad (\text{B.4})$$

em que: $[\Delta I]^h$ é o resíduo de corrente trifásica de ordem harmônica h , $[J]^{hh}$ é a matriz jacobiana de ordem harmônica h e $[\Delta V]^h$ são os incrementos de tensões nodais na ordem harmônica h . Ademais, a Jacobiana e a matriz de admitâncias da rede são muito próximas (MELO, 2018).

Além disso, para a utilização do algoritmo é necessário inserir os seguintes dados de entrada para o fluxo de potência, dentre os mais importantes encontram-se: frequência fundamental (Neste caso 60 Hz); número de componentes harmônicos para serem simulados; topologia de construção das cargas lineares utilizadas, série ou paralelo; dados de injeção de corrente das cargas não lineares.

Ao passo que, após a solução do problema, o fluxo retorna dados de saída como: perfil de tensão para frequência fundamental e suas harmônicas; injeções de corrente, potências ativas e reativas para cada harmônica.