FEDERAL UNIVERSITY OF ESPÍRITO SANTO
TECHNOLOGICAL CENTER
POSTGRADUATE PROGRAM IN ELECTRICAL ENGINEERING

Ph.D. Thesis

Signal Processing on Graphs Methodology for Evaluating the Load Current Variability in Feeders with High Integration of Distributed Generation

Mariana Altoé Mendes

Vitória

2023

Mariana Altoé Mendes

Signal Processing on Graphs Methodology for Evaluating the Load Current Variability in Feeders with High Integration of Distributed Generation

Ph. D. Thesis presented to the Postgraduate Program in Electrical Engineering of the Federal University of Espírito Santo as a requirement to obtain the degree of Doctor/PhD in Electrical Engineering.

Advisor: Prof. D.Sc. Oureste Elias Batista
Co-advisor: Prof. D.Sc. Marcia Helena Moreira Paiva

Vitória

2023

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

---

Altoé Mendes, Mariana, 1992-
A469s Signal Processing on Graphs Methodology for Evaluating the Load Current Variability in Feeders with High Integration of Distributed Generation / Mariana Altoé Mendes. - 2023.
130 f. : il.

Orientador: Oureste Elias Batista.
Coorientadora: Marcia Helena Moreira Paiva.
Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Geração distribuída de energia elétrica. 2. Sistemas de energia elétrica. 3. Teoria dos grafos. I. Elias Batista, Oureste. II. Moreira Paiva, Marcia Helena. III. Universidade Federal do Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

Mariana Altoé Mendes

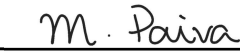
Signal Processing on Graphs Methodology for Evaluating the Load Current Variability in Feeders with High Integration of Distributed Generation

Ph. D. Thesis presented to the Postgraduate Program in Electrical Engineering of the Federal University of Espírito Santo as a requirement to obtain the degree of Doctor/PhD in Electrical Engineering.

Approved on February 2nd, 2023.

Examination board:

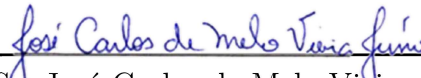
Prof. D.Sc. Oureste Elias Batista
Federal University of Espírito Santo, UFES, Brazil
Advisor



Prof. D.Sc. Marcia Helena Moreira Paiva
Federal University of Espírito Santo, UFES, Brazil
Co-advisor

Prof. D.Sc. Augusto César Rueda Medina
Federal University of Espírito Santo, UFES, Brazil
Internal member

Prof. D.Sc. Alcebíades Dal Col Júnior
Federal University of Espírito Santo, UFES, Brazil
Internal member



Prof. D.Sc. José Carlos de Melo Vieira Júnior
University of São Paulo, USP, Brazil
External member

Prof. D.Sc. Leonardo Silva de Lima
Federal University of Paraná, UFPR, Brazil
External member

To my family, especially to my grandparents.

ACKNOWLEDGEMENTS

There are a number of people to whom I am greatly indebted and helped in the development of my research.

I dedicate my thesis to my mother, who always did everything so that I could have the best life possible and supported me in my determination to find and realize my potential and to my father (*in memoriam*) who, in the years we spent together, taught me what love is. I'm sure you both are proud of me. I also dedicate my research to all my family, especially my grandparents, who were great supporters of my study.

I am dedicating my work to my fiancé Gilberto Avance Filho who with great patience never doubted my ability, never left my side and was a great encourager and life partner.

I would like to thank my advisors D.Sc Oureste Elias Batista and D.Sc Marcia Helena Moreira Paiva who gave me technical and emotional support in difficult times, helped me in the development of the work and pushed me to grow as a professional. And also thanks the Examination board for all the contributions made to the research.

I give special thanks to my Ph.D friends Luiz Tonini, Luann Queiroz and Murillo Vargas for their complicity, help in the development of my research and for all the knowledge we built together. You were essential for my work. I thank all my friends for their emotional support and patience during my long years of absence and my work colleages that taught me so much and helped me to enhance my knowledge.

I would also like to thank all the other professors that crossed my life and the ones that taught me. You encouraged me to always seek more knowledge. I especially thanks professors D.Sc Pablo Rodrigues Muniz and Dr. Domingos Savio Lyrio Simonetti who, together with my advisors, were great supporters of my academic life.

A special thanks to my work colleagues at ArcelorMittal Tubarão who help me to grow professionally every day, especially to my manager Paulo Loss who has always been solicitous for my academic needs.

Most of all thanks to God who continues to make the impossible possible.

This thesis was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001.

ABSTRACT

The emergence of new elements as distributed generators in transmission and distribution networks is a challenge for supplying energy with quality, reliability and continuity. Although there are researches in the area, the impacts of distributed power generation (DG) in these systems are complex and involves many variables and due to this, power systems studies are essential to analyze the feeder behavior in this new scenario ensuring a good compliance with electricity quality levels. This thesis utilizes a graph-theory based model and proposes a novel method that associates concepts of power flow and graph signals to identify in distribution feeders with DG the cases in which the load current varies more for a steady-state analysis. The graphs were performed with Signal Processing on Graphs approach, where the nodes represent the feeder buses and the graph signal is a parameter calculated based on the power flow algebraic formulation. To fit and validate the methodology, the results are related with the current values obtained by solving a power flow problem, using MatLab/Simulink data. Results were presented for 13 and 34 bus electrical grid systems. The Spearman's and Pearson's rank-order correlation showed a good agreement between the results of the graph analysis and the Simulink data. The method proposes an alternative analytic way to identify the topological position of DG in a feeder that most impact in the current variation of the substation bus.

Keywords: Distributed Generation; Graph Theory; Graph Signal Processing; Power System; Simulation.

RESUMO

O surgimento de novos elementos como os geradores distribuídos nas redes de transmissão e distribuição é um desafio para o fornecimento de energia elétrica com qualidade, confiabiliade e continuidade. Apesar de existirem pesquisas na área, os impactos da geração distribuída (GD) nesses sistemas são complexos e envolvem diferentes variáveis. Devido a isso, estudos envolvendo os sistemas elétricos de potência são essenciais para analisar o comportamento dos alimentadores nesses novos cenários, garantindo a qualidade da energia elétrica. O trabalho dessa tese utiliza um modelo baseado na teoria dos grafos e propõe um novo método que associa conceitos de fluxo de potência e sinais em grafos para identificar nos alimentadores de distribuição com GD os casos de maior variabilidade da corrente de carregamento em uma análise em regime permanente. O processamento de sinais em grafos foi utilizado para analisar o sistema, onde os nós representam as barras do alimentador e o sinal do grafo é um parâmetro calculado com base no fluxo de potência algébrico. Para adequar e validar a metodologia, os resultados foram comparados com os valores de corrente de carregamento obtidos pelo cálculo do fluxo de potência no MatLab/Simulink. Os resultados apresentados mostram dados dos sistemas com 13 e 34 barras. Os coeficientes de correlação de Pearson e Spearman demonstraram uma boa relação entre os resultados na análise de grafos e os dados do Simulink. O método propõe uma alternativa analítica para identificar a posição topológica da GD no alimentador que mais impacta na variação da corrente no barramento da subestação.

Palavras-chave: Geração Distribuída; Processamento de Sinais em Grafos; Simulação; Sistemas Elétricos de Potência; Teoria de Grafos.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

13NTF	13-Node Test Feeder
13NTFM	13-Node Test Feeder Modified
13NTFMA	13-Node Test Feeder Modified A
13NTFMB	13-Node Test Feeder Modified B
13NTFMC	13-Node Test Feeder Modified C
34NTFM	34-Node Test Feeder Modified
ANEEL	National Agency of Electrical Energy
CNA	Complex Network Analysis
DG	Distributed Generation
EPA	Environmental Protection Agency
GFT	Graph Fourier Transform
IEEE	Institute of Electrical and Electronics Engineers
IRENA	International Renewable Energy Agency
PV	Solar Photovoltaic
PVDG	Photovoltaic distributed generations
SPG	Signal Processing on Graphs

LIST OF SYMBOLS

α	N×1 vector with one non-zero element (1) in the i^{th} position
α_s	2N×1 vector containing two non-zero elements (tan Δ_{vs} and -1) at the s^{th} and $(N+s)^{th}$ position
τ	Generic Current Value
λ	Eigenvalue
$\lambda_{P1}, \lambda_{P2}$	Eigenvalues of 2N×2N J_{Pi} matrix
$\lambda_{Pa1}, \lambda_{Pa2}$	Eigenvalues of (2N-1)×(2N-1) J_{Pi} matrix
$\lambda_{Q1}, \lambda_{Q2}$	Eigenvalues of 2N×2N J_{Qi} matrix
$\lambda_{Qa1}, \lambda_{Qa2}$	Eigenvalues of (2N-1)×(2N-1) J_{Qi} matrix
$\mathcal{C}_k$	Cluster
$\mathcal{J}(\mathcal{C}_k)$	Squared error between μ_k and the points in the cluster $\mathcal{C}_\|$
$\mathcal{J}(\mathcal{C})$	Minimum squared error over all K clusters
$\mathcal{M}$	Matrix similar to identity
$\mathcal{N}$	Number of objects to be clustered
$\mathcal{P}$	Order of a squared matrix
γ	N×1 arbitrary vector
$\vec{v}$	Eigenvector
θ	Impedance Angle
$\mathbb{R}$	Real set number
δ	Minimum degree of a graph
δ_i	Voltage Angle of a bus i
δ_v	Reference Voltage Angle
Δ_g	Maximum degree of a graph
Δ	Load current variation

Ψ	Quantity of PVDG
ξ	Frequency function
Υ	One of the 512 studied cases
$\sigma(LGFT)$	Spectrum of frequency
σ	Number of geodesics
$\mathcal{E}$	Finite set of edges
$\mathcal{G}$	Graph
$\mathcal{V}$	Non-empty set of vertices
A	Generic matrix
A_Q	Quadratic matrix
$^{\mathrm{T}}$	Transpose of a matrix
$A(\mathcal{G})$	Adjacency matrix
b	i^{th} column of B
B	Susceptance
c	Intersection points of the hypersurface with the principal axis
c_B	Betweenneess centrality
c_C	Closeness centrality
c_D	Degree centrality
c_E	Eigenvector centrality
c_P	Intersection points of the hypersurface with the principal axis considering the real power injection or graph signal
c_Q	Intersection points of the hypersurface with the principal axis considering the reactive power injection
C_{IGFT}	Importance coefficient of the graph signal from the GFT
D	Diagonal matrix
$D(\mathcal{G})$	Diagonal matrix of $\mathcal{G}$
e	N×1 vector of the real part of V

e_i	Canonical basis vector
f	N×1 vector of the imaginary part of V
$\hat{f}$	Graph Fourier Transform of a function f
g	i^{th} column of G
G	Conductance
H	Invertible orthogonal matrix
i	Specific bus
$\dot{I}$	Vector of current injection into buses
Id	Identity matrix
j	Unit imaginary number
J_i	Real symmetric matrix describing the network structure
K	Number of clusters
$L(\mathcal{G})$	Laplacian matrix
L_{GFT}	Laplacian of graphs with some modifications of GSP
m	2N×1 vector of zeroes except for a "1" at position $(N+s)$
M	J matrix of the quadratic form
N	Number of buses in the system
NSI_Q	Notion of smoothness indes using quadratic form
O	N×N null matrix
P	Real power or vector of net real power injection
P_{di}	Real demand of the load at bus i
P_{gi}	Real generated power at bus i
P_{load}	Real load power on a 13NTFM node
P_{node}	Real power on a 13NTFM node
P_{PVDG}	Real power of the PVDG on a 13NTFM node
Pol	Characteristic polynomial

Q	reactive power or vector of net reactive power injection
Q_{di}	Reactive demand of the load at bus i
Q_{gi}	Reactive generated power at bus i
$Q(x)$	Quadratic Form
s	Reference bus number
S	complex power or N×1 vector of complex bus powers
v	Normalized eigenvalue
V	Voltage
$\dot{V}$	Vector of voltage injection into buses
x_r	2N×1 vector of the real and imaginary parts of V
$\tilde{X}$	Normalized value
W	Watt
Y	Admitance matrix
Z	Impedance matrix
z_i	Any specified injection at bus i

CONTENTS

1 INTRODUCTION AND MOTIVATION

1.1 Introduction and Motivation

The electric power transmission and distribution systems have been considerably modified in recent years due to the growth in electricity demand, change in consumption profile and climate changes. Thus, they stand out as complex systems and new elements as distributed generators started to integrate the power grid. Although it is not mandatory, in practice, their integration is associated with the use of renewable energy sources, which provides environmental benefits [1, 2].

In general, the solar photovoltaic (PV) systems are the most popular distributed generators type and the Sun is a feasible generation source capable of reducing the emission of air pollutants [3]. In Brazil, the PV systems represent more than 97.6% of consumer units with generation [52].

The insertion of PV distributed generators can provide advantages for the electrical system such as improve the voltage profile, reduce the electrical losses, improve the energy quality and system efficiency, and may postpone the need for investments in distribution and transmission lines [5]. Therefore, despite the technical, environmental, social and economic advantages, when there is no planning and study of the system, the arbitrary introduction of new generating sources in the distribution network, initially designed to operate in an unidirectional way, can modify electrical parameters and the power flow in a negative way, causing technical problems and malfunctions. That is, the connection of PV distributed generators in the power networks could change the feeder load and short-circuit values, change the voltage profile, cause voltage issues by exceeding service limits [6], and modify voltage stability [7, 8]. In addition, due to the possible bidirectional behavior caused by the insertion of distributed generation (DG), some grid equipment such as protection devices could operate inappropriately or maloperate.

As the integration of DG units has been happening quickly and in some cases unrestrained, it is essential that extensive studies and researches with simulations are developed to predict the network behaviors in these new scenarios and avoid technical issues, ensuring the conformity of the electrical system. Some researches evaluate the impacts that distributed generation (DG) can cause on the power network and there are also papers that suggests how these effects can be mitigated [9–11]. These studies, however, represent feeders and DG models in computational simulation tools where, depending on the model, the size of the system and the software used, can be time-consuming and computationally costly [7, 8, 12]. Furthermore, this thesis is focused in the feeders with high quantity of DG units connected in the power network, which is not a common scenario in

the literature [7, 8]. So, the use of some traditional power flow simulation software may be computationally and timing costly considering a high quantity of DG units, high DG penetration level or high number of feeder scenarios.

Considering that electric power systems have become more complex due to the advent of DG, some mathematical tools and complex network analysis techniques and methodologies, as graph theory, have been used to analyze and better understand the power system behavior [13, 14]. In the graph theory field, there are tools such as the Signal Processing on Graphs (SPG) that incorporate concepts of traditional signal processing and graphs, allowing an analysis in different domains and supporting the use of weighted nodes, which is unusual in the classic graph theory. As the conventional electrical power flow calculation assign values to the system buses, the SPG is a potential technique for the analysis of electrical parameters, as load current values.

That said, this thesis aims to evaluate the load current variability in the substation bus, in order to obtain a global current variation of the feeder. The load current variation is related to the power of DG units that are connected on the feeder and its topological position. Thus, the search for the situations where the load current varies more implies in the identification of the buses where the distributed generators are connected. For that, it is proposed a novel method that associates concepts of power flow algebraic formulation and graph signals to identify in the studied cases the variation in the load current for a steady-state analysis without the need to model the feeder in a software and simulate the power flow, which is usually computationally costly. The methodology is useful to identify the topological locations for the photovoltaic distributed generators (PVDG) that provide greater and smaller current variations in the substation. It was chosen to analyse the load current variation values in the substation bus aiming to obtain a global current variation and based in the fact that all electrical substations have a protection and/or measuring device in this bus. As the load current may vary, this change can cause a maloperation of the protection devices. All the distributed feeders studied in this thesis are balanced.

The proposed method fills a gap in the literature regarding the significant amount of simulations that a researcher and electric utility need to do to study a power system with high integration of PVDG considering different topologies for connecting the generators. This thesis methodology filter out among the proposed scenarios the cases in which the load current in the substation varies more and less. In this way, in addition to identifying which generators were connected in these cases of extreme load current variation, it also reduces the number of scenarios in the feeder analysis, focusing only on the filtered cases, which are usually the ones of greatest interest in the literature and electric power utilities. To fit the methodology, the results are related with the current values obtained by solving a power flow problem, allowing an analysis that do not rely on time-consuming and computationally costly simulations.

1.2 State-of-the-Art

The interest in the distribution generation has increased in the last decades due to the environmental, economic, social and technical advantages. The trending of decarbonisation and decentralization of the electrical energy encourages the installation of generating units all over the electrical grids. With this intense change in distribution and transmission systems, the researches about the technical impacts of the integration of DG units on the power grid has grown in industries, energy utilities and universities.

In the literature, there are papers that present mathematical models and optimization-based algorithms focused in determining the optimal expansion planning of distributed generation units and how to accommodate a high penetration of renewable DG in distribution networks [15–20]. Other studies aim to identify the best sizing and siting for the generators using tools such as second order conic programming problem, fuzzy-logic or optimization methods [21–23]. It is important to note that, the best sizing and siting depends of the problem formulation. The research may be focused in the solution that improve the energy quality based on region rules, improve the voltage stability or voltage profile, reduce the power losses or even allocate the DG units based in maximizing the present value profit gained by the distribution company via procuring power from DGs and the market at a minimum expense [21, 24].

There are also papers that shows the contribution of the DG units for a network reconfiguration in distribution systems, helping to reduce voltage sags and drops trough an metaheuristic optimization algorithm [25] and studies that aim to identify how the DG can affect in the voltage profile [6, 17]. A lot of researches using genetic algorithm, neural network, non-linear programming optimization or simulations in electrical traditional software as MATLAB or OpenDSS analyses the load and short-circuit current variation, aiming to evaluate how the protection equipments and philosophy can be impacted [26–30].

The graph theory has also been used to study electrical systems, as the electric power networks have been studied as a typical example of complex networks. In general, it is used to identify topological characteristics, determine the most important node (based in specific notion of importance), the most critical to the maintenance of the system nodes, identify contingencies involving line outages, voltage stability assessment and the most loaded lines [14, 31–34].

But in recent years the features and tools of graph theory have also been used to analyse the electrical systems considering the new scenarios with distributed generation and smart-grids. For example, the graph theory has been used to evaluate the adequacy of the current electrical infrastructure for a decentralized generation [35], to remove the nodes to measure the disruption of the system under failure scenarios considering a system with DG [13], to propose improvement in power electronics functionalities [36, 37] and to

determine the stages of selectivity of relay protection based on an algorithm for adaptive current protection [38]. There are also papers that propose a network-topology-based method to solve the load-flow problem of radial distribution networks [39].

The SPG also proved to be an interesting tool for analyzing complex networks, helping to identify important nodes, allowing the weighting of nodes, including using complex numbers for it. Considering the electrical power grids, this mathematical approach was applied to analyze the inter-related smart grid measurement data for security and reliability analyses, to detect false data injection attacks in smart grids and to characterize and classify cyber attacks, to detect and identify topology change in electrical networks and to optimize power flow calculation [40–45]. Although not yet demonstrated in electrical networks, the SPG is also used in researches aiming to analyze and classify the entire network based on the weights of the nodes. That is, when the weights of the nodes vary, the SPG results in a different value. This means that the same feeder with different node weights will result in a distinct Graph Fourier Transform (GFT). This possibility is very useful and convenient for electrical systems where different scenarios can be predicted, especially when distributed generators are inserted into the grid [46, 47].

So, it is clear that there are some researches that propose new methods for power flow calculation, some of then using graph theory, and there are SPG methodologies being applied in grids to measure the importance of a node or to weight and compare some specific grid configuration. This thesis uses concepts of a load flow algebraic calculations to weight the graph nodes, considering the connection of distributed generation in some nodes, and then, integrates the results with the SPG to weight different electrical grid configurations, which represent the distinct scenarios that can exist in a electrical grid with DG.

The proposed method identifies the scenarios with the highest and lowest load current variation, which are the cases with the greatest impact on the electrical power grid, filtering all scenarios. In this way, it provides a gain in the reduction of cases that need to be analyzed by researchers or energy concessionaires, reducing the complexity of the problem and optimizing the simulations. Considering that the insertion of DG in electrical power grids around the world is rising, the reduction in the number of simulations required for electric planning and operation in the electric utilities is extremely important.

1.3 Problem Statement, Hypothesis, and Objectives

1.3.1 Problem Statement

Will a methodology capable of inferring the load current variation in a distribution feeder dominated by DG ensuring accuracy without using traditional electrical computer simulations?

1.3.2 Hypothesis

Mathematical and graph tools can be used as parameter for load flow analysis in distribution feeders and allow a simpler and faster load current variation evaluating than the traditional computer simulations.

1.3.3 Thesis Objectives

1.3.3.1 General objectives

This thesis has the objective to propose a new faster and efficient methodology for load current variation analysis in distribution feeders dominated by DG using graph concepts and graph signal processing, defining as weights values derived from calculations related to the power flow algebraic formulation.

1.3.3.2 Specific objectives

- Study and describe the State-of-art related to methodologies that allows a load current variation analysis in a feeder;
- Model the problem based on load flow algebraic formulation, graph theory and signal processing on graph and determine the distinct scenarios with DG;
- Develop a mathematical expression for the importance coefficient of the graph signal suited of relate the load current variation in the substation bus for each one of the proposed cases;
- Validate the methodology by comparing the results of the developed equation with the MATLAB/Simulink load current values for different feeders.

1.4 Contributions

The main contributions of this thesis are:

- A methodology based on SPG that applies the power flow algebraic formulation to weight the graph nodes, capable to determine the topological positions of the DG units in a feeder that results in higher load current variation in the substation;
- The development of a method less computationally costly than the simulations using the traditional power flow software;
- Possibility of study different feeders with a massive number of DGs, allowing the analysis of a large number of different cases to identify which of them most impact the load current variation.

1.5 Related Papers

Since the beginning of the doctorate, in December of 2018, and the beginning of the research development, due to the relevance of the topic, several articles were accepted or are under review in conferences and journals.

1.5.1 Journal papers

- **M. A. Mendes**, M. H. M. Paiva and O. E. Batista, "Signal processing on graphs for estimating load current variability in feeders with high integration of distributed generation", Sustainable Energy, Grids and Networks, Mar. 2023, doi: https://doi.org/10.1016/j.segan.2023.101032.
- **M. A. Mendes**, M. C. Vargas, O. E. Batista, and D. S. L. Simonetti, "Load Currents Behavior in Distribution Feeders Dominated by Photovoltaic Distributed Generation," Electric Power Systems Research, vol. 201, p. 107532, Dec. 2021, doi: 10.1016/j.epsr.2021.107532.

1.5.2 Conference papers

- L. G. O. Queiroz, **M. A. Mendes** and O. E. Batista, "Simplified Dynamic PV Generator Model for Analysis of Voltage and Current Variation in Feeder with High DG Integration," 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 2021, pp. 1-6, doi: 10.1109/ICECCME52200.2021.9590840.
- P. T. Barbosa, L. G. O. Queiroz, **M. A. Mendes**, O. E. Batista and J. F. Fardin, " Analysis of Current Variation in a Distribution System with High Penetration of Distributed Photovoltaic Generation" In: XVI Brazilian Power Electronics Conference (COBEP), 2021, João Pessoa. Proceedings of XVI COBEP, 2021.
- L. G. O. Queiroz **et al.**, "Variação de Tensão em Rede de Distribuição de Energia Elétrica com Alta Penetração de Geração Distribuída Fotovoltaica," in Anais do XXIII Congresso Brasileiro de Automática, 2020, vol. 2, no. 1, pp. 1–8, doi: 10.48011/asba.v2i1.1698.
- C. M. de Lima, **M. A. Mendes** and O. E. Batista, "Estudo da Estabilidade de Tensão em Sistemas de Distribuição com Alta Penetração de Geração Distribuída," in Anais do XXIII Congresso Brasileiro de Automática, 2020, vol. 2, no. 1, pp. 1–6, doi: 10.48011/asba.v2i1.1132.
- **M. A. Mendes**, M. C. Vargas, C. J. Espíndula, L. G. O. Queiroz, L. G. R. Tonini, D. S. L. Simonetti, and O. E. Batista, "Didatic Overcurrent Protection Coordination for

the IEEE 34-Node Radial Test Feeder," in Anais do Simpósio Brasileiro de Sistemas Elétricos 2020, 2020, vol. 1, no. 1, pp. 1–6, doi: 10.48011/sbse.v1i1.2185.

- D. C. Pompermayer, **M. A. Mendes**, M. Dimanski and A. C. R. Medina,"Reactive Power Control on a Consumer Unit with High Integration of PV Generation. in Anais do Simpósio Brasileiro de Sistemas Elétricos 2020, 2020, vol. 1, no. 1, pp. 1–6, doi: 10.48011/sbse.v1i1.2263.

- **M. A. Mendes**, M. C. Vargas, O. E. Batista, Y. Yang, and F. Blaabjerg, "Simplified Single-phase PV Generator Model for Distribution Feeders with High Penetration of Power Electronics-based Systems," in 2019 IEEE 15th Brazilian Power Electronics Conference and 5th IEEE Southern Power Electronics Conference, COBEP/SPEC 2019, 2019, pp. 1–7, doi: 10.1109/COBEP/SPEC44138.2019.9065417.

- L. F. S. Azeredo, L. G. R. Tonini, **M. A. Mendes**, M. C. Vargas, O. E. Batista, and C. J. Espindula, "Voltage Regulator Behavior on Power Distribution Grids with High Integration of PVDG," in 2019 IEEE 15th Brazilian Power Electronics Conference and 5th IEEE Southern Power Electronics Conference, COBEP/SPEC 2019, 2019, pp. 1–6, doi: 10.1109/COBEP/SPEC44138.2019.9065898.

- D. C. Pompermayer, C. Marim, L. G. O. Queiroz, L. G. R. Tonini, **M. A. Mendes**, M. C. Vargas, and O. E. Batista, "Extra Reactive Power Analysis on a Distribution Grid with High Integration of PV Generation," in 2019 IEEE 15th Brazilian Power Electronics Conference and 5th IEEE Southern Power Electronics Conference, COBEP/SPEC 2019, 2019, pp. 1–6, doi: 10.1109/COBEP/SPEC44138.2019.9065632.

- M. C. Vargas, **M. A. Mendes**, and O. E. Batista, "Fault Current Analysis on Distribution Feeders with High Integration of Small Scale PV Generation," in 2019 IEEE Power & Energy Society General Meeting (PESGM), 2019, pp. 1–5, doi:10.1109/PESGM40551.2019.8973786.

- L. G. O. Queiroz, E. V. S. Freitas, **M. A. Mendes**, L. G. R. Tonini and O. E. Batista, "Geração distribuída fotovoltaica e seu impacto na redução de emissões de CO_2: estudo de caso de uma usina fotovoltaica situada no semiárido potiguar" in Anais do Congresso Internacional de Meio Ambiente e Sociedade e III Congresso Internacional da Diversidade do Semiárido, 2019, Campina Grande.

- M. C. Vargas, **M. A. Mendes**, L. G. R. Tonini, and O. E. Batista, "Grid Support of Small-scale PV Generators with Reactive Power Injection in Distribution Systems," in 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), 2019, pp. 1–6, doi: 10.1109/ISGT-LA.2019.8895424.

- L. G. R. Tonini, O. E. Batista, A. C. Rueda-Medina, **M. A. Mendes**, and M. C. Vargas, "Power Flow Analysis on Networks with High Integration of Distributed

Photovoltaic Generation," in 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), 2019, pp. 1–6, doi: 10.1109/ISGTLA.2019.8895415.

1.6 Outline of Thesis

This thesis is divided in 7 chapters, as follows. In chapter 1, the introduction and motivation of the research is presented, specifying the objectives, research contribution and articles that have already been published or are under review.

In chapters 2, 3 and 4 a literature review is presented to explain the historical context and what is already known about load flow algebraic approach, graph theory and Signal Processing on Graphs. The most important concepts of the thesis are discussed and some examples of the theory are presented.

The methods used to developed the analysis of the problem are described in chapter 5, where the feeder used is described, and the problem modelling is characterized. Chapter 6 discusses results obtained from the simulations and analyzes the data. In chapter 7, the conclusion is presented highlighting the benefits and shortcomings of this work approach, and suggesting future topics for research.

2 DISTRIBUTED GENERATION

2.1 General Concepts

There is not a unique definition for DG. The United States Environmental Protection Agency (EPA) defines it as:

> "Distributed generation refers to a variety of technologies that generate electricity at or near where it will be used, such as solar panels and combined heat and power. Distributed generation may serve a single structure, such as a home or business, or it may be part of a microgrid (a smaller grid that is also tied into the larger electricity delivery system), such as at a major industrial facility, a military base, or a large college campus [48]."

Another official, but not necessarily exclusive, definition is the one given by the Brazilian National Agency of Electrical Energy (ANEEL) that divides the DG in two concepts: distributed micro-generation and distributed mini-generation. The main difference is that the micro-generation includes electric power generating plants with installed power less than or equal to 75 kW and the mini-generation consist of bigger plants, with installed power greater than 75 kW and less than or equal to 5MW. For both definitions, the plants must use co-generation (as ANEEL regulations), or renewable sources of electricity, connected to the distribution grids through consumer units installation [49]. It is important to highlight that the definition of micro- and mini-generation do not restrict the distributed generation. There is distributed generation with installed power greater than 5MW, such as small thermoelectric power stations and other generators connected to the distribution or transmission power grid.

Some other authors present their own definitions but in general, all the interpretations describe the DG as a way to generate electricity, not necessarily using renewable energy sources, close to consumer units. Although not mandatory, most units with DGs do use renewable and/or not pollutant sources [1, 2, 48, 50].

In European and North American countries, the distributed generators have been part of the electricity grid for many years [51]. In Brazil the use of distributed generators has been growing in last two decades years, but its effective integration in the distribution network is recent compared to other countries. The first registration of a consumer unit with mini or micro DG in ANEEL system occurred only in 2009 [52]. Before that, some Brazilian government programs (such as Proinfa - *Programa de Incentivo às Fontes Alternativas de Energia Elétrica*, Incentive Program for Alternative Electric Energy Sources) already promoted the use of distributed generation, but through larger scale plants, such as wind farms and small thermoelectric plants [53].

Due to the low environment impact, reduced noise, facility to install in the residence roofs, low maintenance as it has no rotating parts, decreasing price of technology and great potential to meet the world's growing energy demand, the PVDG represents the vast majority of DG worldwide. The PVDG increased 22% (+131 TWh) in 2019 and in an optimist scenario, it can reach over 149 GW in 2022. This growing results in an estimation that distributed solar photovoltaic capacity growth will reach 220 GW by 2026 [54, 55]. According to IRENA (International Renewable Energy Agency), the global installed capacity of solar generation in 2020, 600 GW, is about 12 times greater than in 2010 [56]. In Brazil, the PVDG systems represents more than 97.6% of consumer units with distributed generators [4].

2.2 Challenges and Concerns

The use of DG can provide advantages for the electrical system such as improved voltage profile, reduced electrical losses, enhanced system reliability and security, improved energy quality and system efficiency, and may postpone the need for investments in distribution and transmission lines. Indices have been proposed and studied to quantitatively evaluate these technical benefits [20, 28, 57].

Some countries have established specific legislation for DG, usually aiming to accelerate administrative procedures for authorisation and connection to the electricity distribution and transmission networks, urging simplified authorisation procedures [57].

Nevertheless, for some countries and electrical utilities, there are no regulations that attempts to regulate the total amount of DG power allowed in a distribution feeder. Although, there is a general interest in increasing the amount of distributed generating units in the world. Similarly, no concerns with the topological position of generators on the grid is normally presented in the regulations imposed by governments and electricity utilities. This lack of rules and also planning strategies culminates into distribution network operators often connecting DG units in a "fit and forget" structure [58].

Thus, when there is no planning and study of the system, the arbitrary introduction of new generating sources in the distribution network, initially designed to operate in a unidirectional way, can modify electrical parameters and the power flow in a negative way, causing technical problems and malfunctions [7, 8, 12, 26, 28].

More specifically, the insertion of distributed generators on the electric power grid may cause load current variation all over the feeder. In most of cases, it is interesting to have a small electric current passing through the distribution lines as this may reduce power losses. Although less common, it is important to highlight that depending on the feeder behaviour, topology, power of the DG and topological position of the generators, insertion of DG may increase discreetly the value of load currents [28]. The fault currents values

are also affected due to the fault current behaviour of PV generators. The fault and load currents changes may cause protection coordination problems as miscoordination, damage to equipment, saturation of the current transformers and fuse blowing [12, 26, 27, 59].

The voltage profile can also be changed, modifying voltage stability and causing voltage issues as the nominal voltage limits at point of common coupling are regulated. So it may cause voltage issues by exceeding service limits. Moreover, so far, in Brazil and in some countries there is no obligation for DG to support the grid injecting or absorbing reactive power during steady-state conditions. However, both Brazilian and international standards and codes expect more participative grid support from the generating units, although it is not an obligation [27, 60, 61]. Thus, considering that the insertion of DG can result in feeders with a greater amount of reactive loads (due to the fact that it is not yet mandatory the injection/absorption of reactive power) and as some systems have been pushed to their operational limits, there may be voltage instability even with the voltage profile at conforming values [6, 62].

Therefore, is essential to investigate the power system electrical variations to predict the possible modification of the grid through the new emerging scenarios from the massive insertion of distributed generators in the electrical networks. The technical problems arising from the insertion of DGs in the network are a challenge for the future of the electrical system, since each year the number of consumer units with generators in the world grows. So, several studies try to define the high or optimal amount of PVDG that may be connected in power grid without affecting the system operating conditions [63]. Despite the efforts of the academic community to study and try to find methodologies and equipment capable of mitigating the DG effects, there are knowledge gaps to be filled with regard to the analysis of the electrical system in the new scenarios.

2.3 High Integration of Distributed Generation

As mentioned above, the recent interest in DG and the financial support provided by governments have provided a boost in the distributed generators installation. Studies in this area, which aim to find the weaknesses and mitigate the possible problems of the new structure of the electrical power system, usually address distribution and/or transmission networks with high integration of distributed generators.

In the literature, the definition of feeders dominated by DG is not unique and it is known that the PV penetration level in a power system depends upon multiple variables as load type, solar irradiance level, PVDG concentration along the feeders, power voltage (low, medium or high), power grid parameters and topology, among others [54, 63]. A few studies indicate that the penetration level is related to the ratio of the houses with PV systems to the total houses in feeders under study, other affirm that it is related to the

transformer capacity, feeder maximum load versus PV peak capacity or it is the annual energy from PV systems to the total energy consumption ratio. Furthermore, other ratios that are normally considered by the studies to determine the concept of high integration of DG are: the ratio of installed PV peak capacity to the feeder maximum load and the ratio of the actual PV output generation to the actual real power load of the power grid [54].

Considering this information, some researches show that a small percentage, approximately 15%, of PVDG integration (considering one of the methods presented above) violates some voltage or current limits and other studies indicate that a high PVDG penetration percentage, at about 500% w.r.t feeder minimum load do not trigger any voltage violation [54]. Once again, this reinforces the difficulty of establishing a single definition of high integration of PVDG. In this work, the DG penetration level is defined as the ratio between the sum of the installed PV system power generation and the total power of the feeder (sum of all loads).

3 POWER FLOW

3.1 Definition

The power flow calculation of an electrical system aims to obtain the operating state of an electrical network in the steady-state. This means that it is possible to compute the voltage magnitude and angle at each system bus and also the real (P) and reactive (Q) power flow through all equipments interconnecting the buses.

Electric power flow simulation software and mathematical methods as Gauss-Seidel, Newton-Raphson and backward/forward to calculate the power flow are necessary to guarantee reliable real-time operation, expansion and operation planning [64, 65]. The power flow results can also be used to analyze the grid behaviour for different scenarios, for example, with more or less load, to study the impacts of the insertion of DGs and assess the viability of the power system operation in emergency conditions.

Normally it is necessary to use iterative numerical methods to find a solution for the power flow, as the problem results in non linear algebraic equations. Although there are several studies that aim to simplify calculations, which often do not converge due to the non-linearity of the system, the classic most well-known and widespread mathematical methods for power flow calculations are Gauss-Seidel, Newton-Raphson and backward/forward sweep method [64–68].

The backward/forward sweep method for load flow calculation is an iterative technique normally used in radial systems. The forward sweep method will determine a computed source voltage. Using Kirchhoff's voltage law the values of voltage and current of the subsequent buses of the feeder, in ascending order, are calculated, explaining the method's name. The backward sweep uses the values of the voltage and the currents specified from the forward sweep and it will do the sweep in the inverse order, starting with the bus in the end of the feeder, coming to the feeder source. A second iteration of the forward method will then start, now with a new value for the source voltage, specified in the backward sweep process. The sweep method is continued until the difference between the calculated and specified values are smaller than the defined error [68, 69].

Despite the Gauss-Seidel method is mathematically simpler, in most of the times it takes more iterations to converge. Moreover, in some situations this method may diverge. So, usually, this method is not the chosen one [65, 66].

The Newton-Raphson method is based on expansion of Taylor series. It is mathematically superior to Gauss-Seidel and it is a method less likely to diverge. For larger systems, the Newton-Raphson is more efficient and the number of iterations do not depend

of the system size [70]. It is possible to demonstrate that this method has a quadratic convergence rate, that is, the norm of the residual vector decreases with the square of the number of iterations. This means that the method converges quickly in the vicinity of the solution. Convergence is strongly dependent on the initial condition of the problem, which is usually given by the voltage and angles in the load buses [71].

For the power flow calculation, admittance matrix Y, or impedance matrix Z, of the system should be known. Due to some mathematical facilities, normally, the admittance matrix is used. The (N $\times$ N) Y matrix, where N is the number of buses of the electrical system, relates nodal electrical voltages to currents injected into the system through generators, according to Kirchoff's law as (3.1)

$$\dot{I} = Y\dot{V}, \tag{3.1}$$

$$\begin{bmatrix} \dot{I}_1 \\ \dot{I}_2 \\ \vdots \\ \dot{I}_{\mathrm{N}} \end{bmatrix} = \begin{bmatrix} \overline{Y}_{11} & \overline{Y}_{12} & \cdots & \overline{Y}_{1\mathrm{N}} \\ \overline{Y}_{21} & \overline{Y}_{22} & \cdots & \overline{Y}_{2\mathrm{N}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{Y}_{\mathrm{N}1} & \overline{Y}_{\mathrm{N}2} & \cdots & \overline{Y}_{\mathrm{NN}} \end{bmatrix} \cdot \begin{bmatrix} \dot{V}_1 \\ \dot{V}_2 \\ \vdots \\ \dot{V}_{\mathrm{N}} \end{bmatrix}, \tag{3.2}$$

where $\dot{I}$ represents the vector of current injections into the buses, Y the admittance matrix and $\dot{V}$ the vector of the voltage injections into the buses. Each element of Y matrix can be described by (3.3) [65]

$$Y_{\mathrm{ij}} = |Y_{\mathrm{ij}}| \, \underline{/\theta_{\mathrm{ij}}} = |Y_{\mathrm{ij}}| \cdot \cos\theta_{\mathrm{ij}} + j|Y_{\mathrm{ij}}| \cdot \sin\theta_{\mathrm{ij}} = G_{\mathrm{ij}} + jB_{\mathrm{ij}}, \tag{3.3}$$

where G is known as conductance, B susceptance, both are symmetric matrices, θ is the impedance angle and j depicts the complex number defined by $j^2 = -1$.

The voltage at a specific bus i and the net current injected into the network at bus i can be defined as (3.4) and (3.5) respectively, where δ_v is the voltage angle.

$$V_{\mathrm{i}} = |V_{\mathrm{i}}| \, \underline{/\delta_{\mathrm{i}}} = |V_{\mathrm{i}}|(\cos\delta_{\mathrm{vi}} + j\sin\delta_{\mathrm{vi}}) \tag{3.4}$$

$$I_{\mathrm{i}} = Y_{i1}V_1 + Y_{i2}V_2 + \cdots + Y_{iN}V_N = \sum_{n=1}^{N} Y_{in}V_n. \tag{3.5}$$

Considering i the bus under analysis, as mentioned before, P_i and Q_i denote the real and reactive power, respectively, entering the system at bus i [65]. Thus, the complex

conjugated of the electric power injected at bus i is given by (3.6) [65].

$$P_i - jQ_i = V_i^* \sum_{n=1}^{N} Y_{in} V_n. \tag{3.6}$$

By substituting (3.3) and (3.4) into (3.6), the complex conjugated of the electric power injected at a bus i can be rewritten as (3.7) [65]

$$P_i - jQ_i = \sum_{n=1}^{N} |Y_{in} V_i V_n| \, \underline{/\theta_{in} + \delta_{vn} - \delta_{vi}}. \tag{3.7}$$

By expanding (3.7) into real and imaginary parts, it is possible to obtain the real, depicted by (3.8), and reactive, defined as (3.9), power in polar form entering the electric network at a specific bus i.

$$P_i = \sum_{n=1}^{N} |Y_{in} V_i V_n| \ \cos(\theta_{in} + \delta_{vn} - \delta_{vi}), \tag{3.8}$$

$$Q_i = -\sum_{n=1}^{N} |Y_{in} V_i V_n| \ \sin(\theta_{in} + \delta_{vn} - \delta_{vi}). \tag{3.9}$$

For each bus of the system, four quantities are defined with respect to the voltage and the power injection:

- V_i: voltage magnitude of the bus i
- δ_{vi}: voltage angle of the bus i
- P_i: net real power magnitude of the bus i
- Q_i: net reactive power magnitude of the bus i

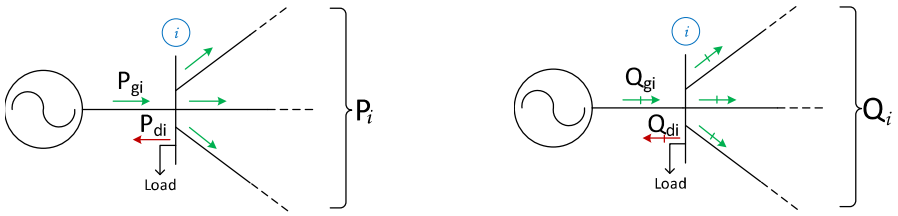
It is important to notice that the net real and reactive power, in fact, can be depicted as (3.10) and (3.11)

$$P_i = P_{gi} - P_{di}, \tag{3.10}$$

$$Q_i = Q_{gi} - Q_{di}. \tag{3.11}$$

where P_{gi} and Q_{gi} denote the real and reactive generated power at bus i and P_{di} and Q_{di} denote the real and reactive power demand of the load at that bus as shown in Figure 1 [65]. If the bus has no load or generation, the correspondent term is set equal to zero.

Figure 1 – Real and reactive power notation for a specific bus i



Source: Adapted from [65]

The calculated real and reactive power are compared to the scheduled values. If there is a difference higher than the allowed error, the calculation method must continue [65].

Two of the four quantities are specified and the others must be calculated. For each bus, there are two equations. For electric systems with N buses, there are $2N$ equations. Normally, in power-flow studies, it is necessary to identify the type of the buses in the network. With this information, one can specify the known quantities [65].

Each bus can be categorized into 3 types:

- Swing bus or Slack bus or $V\delta$ bus: there is only one bus of this type in the entire system. This bus is normally numbered as "Bus 1", it is the angular reference for the system and it is a bus to supply the system losses. For this type of bus, the knowing variables are voltage angle and magnitude - V and δ_v - and these values are typically $1 \angle 0^{o}$. The power flow for this kind of bus computes P and Q [66].
- Load bus or PQ bus: in this type of bus there is no voltage control. Most of the system buses are of this type. The input values for a load bus are P and Q and the load flow calculates V and δ_v [66].
- Voltage controlled bus or PV bus: in this type of bus there are control devices that allow P and V to be set at specified values as synchronous generators and compensators. The power flow program computes the Q and δ_v values [65, 66].

A summary of the power flow problem is presented at Table 1, where i denotes the bus under analysis, N is the number of buses of the electric system and N_g is the number of generation buses.

The functions of real and reactive power, presented in (3.8) and (3.9) are nonlinear functions of the state variables $|V_i|$ and δ_{vi}. So, as already mentioned, power-flow calculations usually employ iterative techniques such as the Gauss-Seidel and Newton-Raphson procedures [65].

Table 1 – Power flow problem summary

Bus type	**Number of buses**	**Quantities specified**	**Quantities calculated**	**Number of available equations**
Reference (Slack, Swing, Vδ_v)	1	$\|V_i\|$, δ_i	P_i, Q_i	0
PQ	N - N_g - 1	P_i, Q_i	$\|V_i\|$, δ_{vi}	2(N - N_g - 1)
PV	N_g	P_i, $\|V_i\|$	Q_i, δ_{vi}	N_g

These mathematical procedures require a considerable computational effort and in some cases (for example, very unbalanced electric power systems, a distribution or transmission line with very high impedance or a bus with very small power load in relation to other values) the method can diverge, making it impossible to calculate the power flow. Thus, as it is very important to know power, voltage and current levels in an electric system, for a lot of reasons already mentioned, specially to simulate the grid scenarios with DG units, using non-iterative methods for power flow calculations can make the process faster and simpler, avoiding divergences. An alternative way to calculate the power flow using a non-iterative method is presented in Section 3.3.

3.2 Mathematical Approach

3.2.1 Diagonalizable Matrix: Linear Algebra

In this chapter, some concepts of linear algebra will be discussed. The whole theory was extracted from [72–77]. The theory presented is fundamental to understand the methodology applied in this thesis.

3.2.1.1 Orthogonal Matrix

The transpose A^T of matrix A is obtained by interchanging rows and columns of A. Thus, for example, the transpose matrix of A_1 (3.12) is determined by (3.13)

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \tag{3.12}$$

$${A_1}^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \tag{3.13}$$

The inverse of a matrix (A^{-1}) is obtained when the multiplication of the matrix by its inverse is equal do the identity matrix. The identity matrix of order $\mathcal{P}$ is a matrix of

$(\mathcal{P} \times \mathcal{P})$ dimension with all elements equal to zero except the elements of the principal diagonal, which are equal to one.

$$A \cdot A^{-1} = Id \tag{3.14}$$

where Id is identity matrix of order $(n \times n)$.

A matrix $A \in \mathbb{R}^n$ is orthogonal when $A \cdot A^T = Id$. This means that the transpose of a matrix A is equal to the inverse of matrix A ($A^{\mathrm{T}} = A^{-1}$) and for this type of matrix, the inverse can be obtained by transposition. Considering that a matrix always commutes with its inverse, the formulation $A^T \cdot A = Id$ is also valid [76].

For the square matrices, all the relationship presented by (3.15) are equivalent and can be used to describe an orthogonal matrix.

$$A \cdot A^{\mathrm{T}} = A^{\mathrm{T}} \cdot A = Id; \quad A^{\mathrm{T}} = A^{-1} \tag{3.15}$$

These matrices occur in a lot of applications and appear as the transition matrices when moving from one orthonormal basis to another.

An important observation that can be made is that an orthogonal matrix has orthonormal lines and columns.

3.2.1.2 Orthogonally Diagonalizable Matrices

Many matrices that appear in mathematical and engineering problems are symmetric. A symmetric matrix A is necessarily square and meets the following criteria

$$A^{\mathrm{T}} = A \tag{3.16}$$

The symmetric matrix A is orthogonally diagonalizable if it is possible to find an invertible orthogonal matrix H such that

$$H^{\mathrm{T}} \cdot A \cdot H = D \tag{3.17}$$

where D is a diagonal matrix. Multiplying both sides of equation (3.17) by H in the left and H^T in the right

$$H \cdot H^{\mathrm{T}} \cdot A \cdot H \cdot H^{\mathrm{T}} = H \cdot D \cdot H^{\mathrm{T}} \tag{3.18}$$

By applying equation (3.15) into (3.18), knowing that H is orthogonal, it is possible to obtain

$$Id \cdot A \cdot Id = H \cdot D \cdot H^{\mathrm{T}} \Longleftrightarrow A = H \cdot D \cdot H^{\mathrm{T}} \tag{3.19}$$

To prove that A is symmetric, just transpose both sides of the equation (3.19)

$$A^{\mathrm{T}} = (H \cdot D \cdot H^{\mathrm{T}})^{\mathrm{T}} \tag{3.20}$$

From the properties of transpose

$$(A \cdot B)^{\mathrm{T}} = B^{\mathrm{T}} \cdot A^{\mathrm{T}} \tag{3.21}$$

Likewise, for more terms

$$(A \cdot B \cdot C)^{\mathrm{T}} = C^{\mathrm{T}} \cdot B^{\mathrm{T}} \cdot A^{\mathrm{T}} \tag{3.22}$$

Using equation (3.22) in (3.20) and as D is a diagonal matrix $D^T = D$

$$A^{\mathrm{T}} = (H^{\mathrm{T}})^{\mathrm{T}} \cdot D^{\mathrm{T}} \cdot H^{\mathrm{T}} = H \cdot D \cdot H^{\mathrm{T}} = A \tag{3.23}$$

So, according to [72], any symmetric matrix is orthogonally diagonalizable. Besides that, as stated in equations (3.16) and (3.18) from a symmetric matrix A, it is possible to calculate a diagonal matrix D where its elements are eigenvalues of A. The matrix H is the eigenvector matrix of A, where each column of H is an eigenvector of the eigenvalue associated with the same column in D.

Spectral Theorem: Let $A \in \mathbb{R}^{n \times n}$. Then A is symmetric if and only if A is orthogonally diagonalizable. If A is symmetric and $A \in \mathbb{R}$ (i) A has a set of real eigenvalues [76].

3.2.1.3 Eigenvalues and Eigenvectors: Eigenfunctions of a Matrix

The eigenfunctions of a matrix are characteristics that can be employed in the study of a lot of scientific applications, being very useful for dynamic systems and to understand better their stability. A lot of eigenfuctions applications comes from the study of systems of differential equations [76]. Some generic applications are implementation of algorithms for pattern identification, algorithms for data processing and analysis, and reducing dimensions of data to decrease the training time and errors of the modelling process [78, 79]. A simple example of eigenfunctions application is the explanation of biologic phenomena [80]. It can also be used for power systems analysis, which is the scope of this thesis.

The applications of eigenfunctions for power systems studies are vast. Some examples can be cited [71]:

- Damping and control of electromechanical oscillations in the face of minor disturbances
- Sub-synchronous resonance
- Harmonic Behaviour
- Analysis of voltage from the QV sensitivity matrix;
- Control conflicts;
- Power flow studies.

When calculated, the eigenvalues and eigenvectors of a matrix help to understand the system behaviour.

The eigenvalues of a matrix can be defined as (3.24) [81]

$$A_Q \cdot \vec{v} = \lambda \cdot \vec{v} \tag{3.24}$$

where A_Q represents a symmetric square matrix ($n \times n$ dimension), $\vec{v} \neq 0$ is an ($n \times 1$ dimension) vector, eigenvector and λ is the eigenvalue of A_Q related to $\vec{v}$ [76, 81].

On the other hand, if $\vec{v}$ is an eigenvector of A_Q with associated eigenvalue λ, the $e^{\lambda t}\vec{v}$ is a solution of a homogeneous system. Each eigenvector of A_Q gives rise to a solution of the system and if A_Q has enough eigenvectors, then every solution of the homogeneous system is a linear combination of the simpler solution [76].

The number of eigenvalues (n) depends of the matrix dimension ($n \times n$). To obtain λ values it is necessary to find the roots of the characteristic polynomial equation of matrix A_Q. This means [76, 81]:

Theorem 1: *λ is the eigenvalue of A_Q if and only if*

$$det(\lambda \cdot Id - A_Q) = 0 \tag{3.25}$$

where Id is the identity matrix of ($n \times n$) dimension. For each eigenvalue λ a corresponding eigenvector can be found by solving $(\lambda \cdot \ Id - A_Q)\vec{v} = 0$. It is possible to see that $det(\lambda \cdot Id - A_Q)$ is a polynomial in λ of degree n. Is it called *characteristic polynomial of* A_Q.

To obtain the eigenvectors of a matrix, first, it is necessary to find the eigenvalues. It is important to say that it is possible to have eigenvalues equal to zero but, the eigenvector are by definition nonzero.

Considering equation (3.24), the eigenvector is the column vector $\vec{v}$ [$(n \times 1)$ dimension] associated to the eigenvalue λ. Different λ may generate different eigenvector. Distinct matrices may have the same λ values, but it do not necessarily means that the matrices have the same characteristics. It is important to note that the eigenvalues of symmetric matrices are real numbers.

Theorem 2: *Similar matrices have the same eigenvalue* [76].

To prove that two matrices are similar it is sufficient to show that their characteristic polynomial are the same [76].

In this thesis, to standardize, the set of eigenvectors are defined as $v_1, v_2, \cdots, v_n$ and they are related, respectively to their eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$.

3.2.2 Quadratic Forms: Linear Algebra

The quadratic forms are real functions of several variables in which each term is the square of some variable or the product of two variables. These functions describe behaviour in a variety of applications, including the vibrations of mechanical systems, Geometry, Statistics and Electrical Engineering, for example, design, optimization and signal processing [72].

Quadratic forms of $\mathbb{R}^{\mathrm{N}}$ are functions of the form described by (3.26)

$$a_1 {x_1}^2 + a_2 {x_2}^2 + a_3 {x_3}^2 + \cdots + a_{\mathrm{n}} {x_{\mathrm{n}}}^2 + (\text{all possible terms } a_k x_i x_j \text{ for } x_i \neq x_j) \tag{3.26}$$

The $a_k x_i x_j$ terms are called mixed terms and it is usual to combine terms involving $x_i x_j$ with terms involving $x_j x_i$. Thus, an arbitrary quadratic form of $\mathbb{R}^3$ can be written as

$$a_1 {x_1}^2 + a_2 {x_2}^2 + a_3 {x_3}^2 + 2a_4 x_1 x_2 + 2a_5 x_1 x_3 + 2a_6 x_2 x_3 \tag{3.27}$$

The equation (3.27) can be depicted in the matrix form as

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot \begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x^{\mathrm{T}} A_Q x \tag{3.28}$$

The matrix A_Q in equation (3.28) is symmetric: its diagonal entries are the coefficients of the squared terms, as shown in equation (3.27), and its entries outside the

diagonal are half the coefficients of the mixed terms. In general, if A_Q is a symmetric $(n \times n)$ matrix and x is the column vector $(n \times 1)$ of the variables, then the function of equation (3.29) is the quadratic form depicted in (3.29) [72]

$$Q(x) = x^{\mathrm{T}} A_Q x \tag{3.29}$$

3.2.2.1 Change of Variable

When necessary, a change of variable can be made to simplify the quadratic form, as shown in (3.30).

$$x = Hy \tag{3.30}$$

With this, the variables $x_1, x_2, x_3, \cdots, x_n$ are expressed in terms of new variables $y_1, y_2, y_3, \cdots, y_n$. Considering that H is orthogonal and invertible, (3.30) represents an orthogonal change of variable.

Using (3.30) and (3.17) in (3.29)

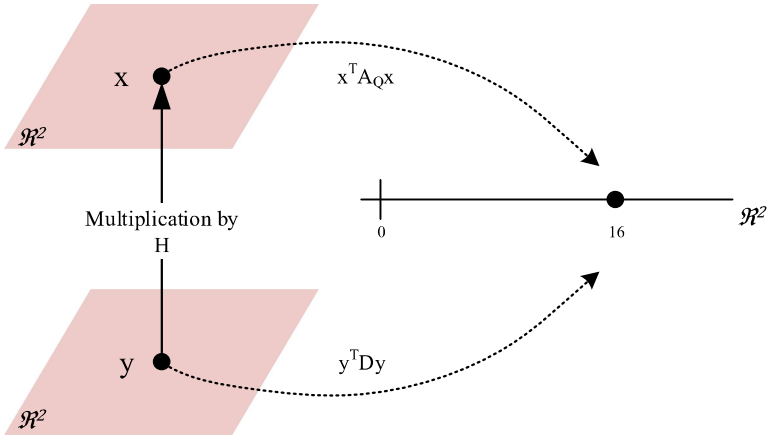
$$(Hy)^{\mathrm{T}} A_Q (Hy) = y^{\mathrm{T}} H^{\mathrm{T}} A_Q H y = y^{\mathrm{T}} (H^{\mathrm{T}} A_Q H) y = y^{\mathrm{T}} D y = Q(y) \tag{3.31}$$

As D is a diagonal matrix with the eigenvalues of A_Q in the main diagonal, the change of variables represents x in the eigenbase associated to the eigenvalues. The quadratic form of (3.31) can be represented in matrix form

$$\begin{aligned} x^{\mathrm{T}} A_{\mathrm{Q}} x = y^{\mathrm{T}} D y = \begin{bmatrix} y_1 & y_2 & \cdots & y_{\mathrm{n}} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{\mathrm{n}} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\mathrm{n}} \end{bmatrix} \\ = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_{\mathrm{n}} {y_{\mathrm{n}}}^2 \end{aligned} \tag{3.32}$$

With this, it is possible to represent (3.29), depicted by the matrix A_Q, as *Q(x)*, described by matrix *D*. This change simplifies the equation by using a diagonal matrix. Furthermore, representing *Q(x)* as a function of variable y means finding a new coordinate axis for the system, as shown in Figure 2. This is know as Principal Axis Theorem.

Figure 2 – Random example of change of variable in $x^T A_Q x$



Source: Author

3.2.3 Principal Axis Theorem

According to [72], if there is a symmetric matrix A_Q with dimension $n \times n$, then there is an orthogonal change in variables that transforms the quadratic form $x^T A_Q x$ into the quadratic form $y^T Dy$ without the mixed terms.

If A_Q is a diagonal matrix, D, for example, the graphic of the corresponding quadratic form will be in canonical position as shown in Figure 3. If A_Q is not a diagonal matrix, the graphic will be a rotation of the graphic in canonical position, Figure 4. So, finding the main axes, determined by the eigenvectors of A_Q means to determine a new coordinate system in relation to which the graphic is in a canonical position [74].

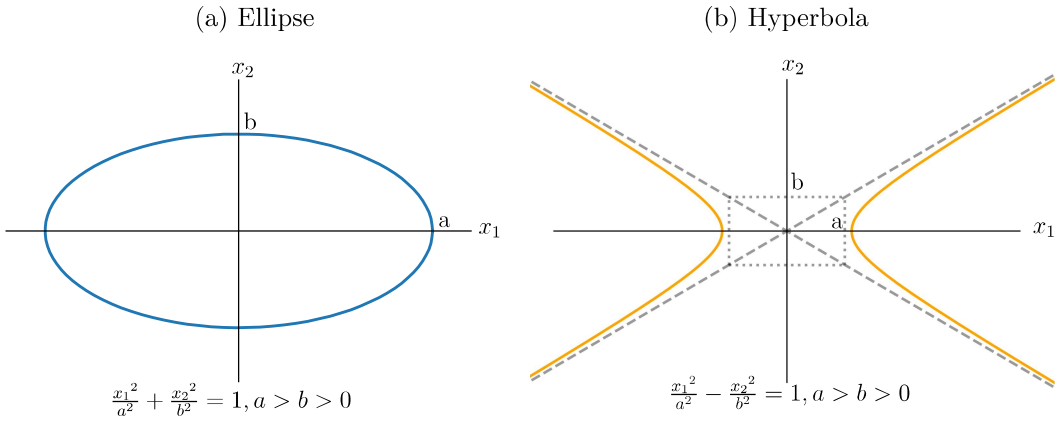
As already mentioned, specifically if matrix H diagonals A_Q orthogonally, then changing variables shown in (3.30) transforms the quadratic form $x^T A_Q x$ into the quadratic form presented in (3.32), in which λ_1,λ_2,$\cdots$,λ_n are the eigenvalues of A_Q associated to the eigenvectors that compose the columns of matrix H.

3.3 Load Flow Algebraic Formulation

Considering the possible difficulties of using the traditional iterative methods to calculate the power flow, in this section a methodology based on Quadratic Forms for power flow calculation will be presented [75].

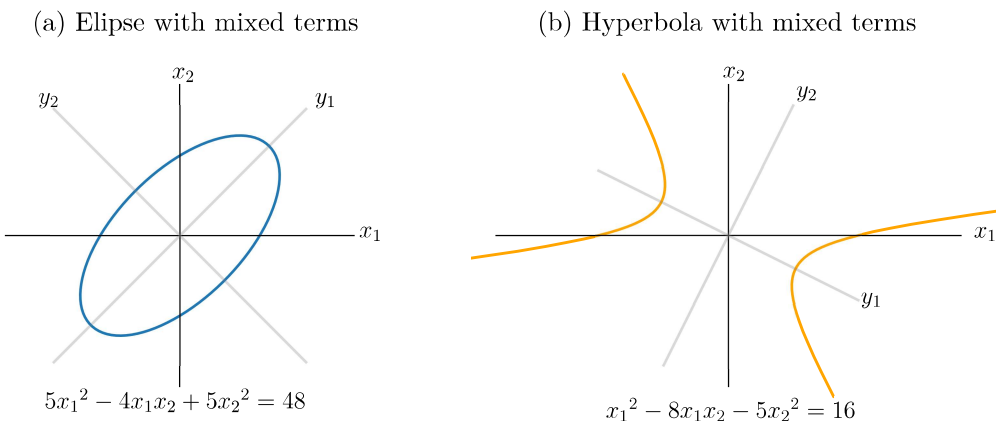
It is know that the load flow equations may be formulated in rectangular coordinates, which is a more convenient representation for analytical studies, and polar coordinates,

Figure 3 – Ellipse and hyperbola in canonical position, with no mixed terms



Source: Adapted from [74]

Figure 4 – Ellipse and hyperbola with mixed terms, graph rotation in relation to canonical position



Source: Adapted from [74]

more interesting for numerical calculations [75]. As the scope of this thesis is to study analytically the behavior of electrical systems, the rectangular coordinates will be used.

So, using the rectangular formulation to describe the power electrical system load flow equations, it is possible to formulate it in a specific way, including an equation for the specification on the reference bus. With this manipulation, the power flow equations become fully quadratic, allowing the use of some mathematical properties [75].

3.3.1 (2N) Equations

In a power system of N-buses, it is possible to specify P (real power), Q (reactive power) or V^2 (quadratic voltage), for each bus i, by means of z_i. These quantities can be expressed as (3.33) [75].

$$z_{\mathrm{i}} = {x_{\mathrm{r}}}^{\mathrm{T}} J_{\mathrm{i}} x_{\mathrm{r}} \tag{3.33}$$

where x_r is a (2N×1) vector that represents the bus voltage in rectangular coordinates, originated from (3.34), as shown in (3.35). In this equation, e denotes the real components and f the imaginary voltage component at the bus.

$$V = e + jf \tag{3.34}$$

$$x_{\mathrm{r}} = \begin{bmatrix} e \\ f \end{bmatrix} \tag{3.35}$$

The J_i term is a (2N×2N) constant real symmetric matrix that is determined by the type of injection (P, Q or V^2) and the admittance matrix [75].

As said in (3.1), shown in Section 3.1, the complex current injected in a bus is represented by (3.36)

$$I = YV \tag{3.36}$$

where I is a (N×1) matrix of currents, Y is the (N×N) admittance matrix and V is the (N×1) voltage matrix.

In a power system, however, usually the known values are the complex power (S), not the current. So, it is possible to relate current and power by (3.37)

$$S = VI^* \tag{3.37}$$

Knowing that all terms of (3.37) are matrices, and based on the matrix multiplication definition, it is necessary to transform the matrix V in a diagonal matrix *diag(V)* with (N×N) dimension. So, according to this and using (3.36), (3.37) can be rewritten as shown in (3.38)

$$\begin{aligned} S &= diag(V)I^* \\ S &= diag(V)(YV)^* \\ S &= P + jQ \end{aligned} \tag{3.38}$$

where S is a (N×1) complex matrix of bus powers and P and Q are (N×1) vectors of real and imaginary power injections, respectively.

Adding (3.3) and (3.34) in (3.38) it is possible to obtain (3.39) and (3.40), which denotes the real and imaginary parts of S, respectively.

$$P = diag(e) \cdot [Ge - Bf] + diag(f) \cdot [Gf + Be] \tag{3.39}$$

$$Q = diag(e) \cdot [-Gf - Be] + diag(f) \cdot [Ge - Bf] \tag{3.40}$$

where *diag(e)* and *diag(f)* are (N×N) matrices whose diagonal elements are equal to e and f.

The equations presented until now depict voltage, current and power matrices that describes the entire system. If the interest of the analysis is in determining values of a specific bus, some mathematical manipulations should be done, as shown below.

For a real constant vector, $\gamma_{(\mathrm{N}\times 1)}$, considering $z_{(\mathrm{N}\times 1)}$ the relation presented in (3.41) is true

$$\gamma^{\mathrm{T}} z = \sum_{i=1}^{N} \gamma_{\mathrm{i}} z_{\mathrm{i}}. \tag{3.41}$$

As previously mentioned, z can denote P, Q or V^2 and knowing the relation described by (3.33), it is possible to substitute (3.39) and (3.40) in (3.41), developing, respectively, (3.42) and (3.43).

$$\gamma^{\mathrm{T}} P = \sum_{i=1}^{N} \gamma_{\mathrm{i}} P_{\mathrm{i}} = \gamma^{\mathrm{T}} diag(e) \cdot [Ge - Bf] + \gamma^{\mathrm{T}} diag(f) \cdot [Gf + Be] \tag{3.42}$$

$$\gamma^{\mathrm{T}} Q = \sum_{i=1}^{N} \gamma_{\mathrm{i}} Q_{\mathrm{i}} = \gamma^{\mathrm{T}} diag(e) \cdot [-Gf - Be] + \gamma^{\mathrm{T}} diag(f) \cdot [Ge - Bf] \tag{3.43}$$

For a better understanding, (3.42) will be developed in (3.44) and the same procedure can be done for Q and V^2

$$\gamma^{\mathrm{T}} P = \gamma^{\mathrm{T}} diag(e) Ge - \gamma^{\mathrm{T}} diag(e) Bf + \gamma^{\mathrm{T}} diag(f) Gf + \gamma^{\mathrm{T}} diag(f) Be \tag{3.44}$$

Equation (3.45) presents a relation that can be added to (3.44) [75].

$$\gamma^{\mathrm{T}} diag(x) = x^{\mathrm{T}} diag(\gamma) \tag{3.45}$$

According to (3.45) and (3.35), (3.44) can be rewritten and put in matrix format as (3.46)

$$\begin{aligned}\gamma^{\mathrm{T}}P = e^{\mathrm{T}}diag(\gamma)Ge - e^{\mathrm{T}}diag(\gamma)Bf + f^{\mathrm{T}}diag(\gamma)Gf + f^{\mathrm{T}}diag(\gamma)Be \\ \gamma^{\mathrm{T}}P = \begin{bmatrix} e & f \end{bmatrix} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} e \\ f \end{bmatrix} \\ = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot x_{\mathrm{r}} \\ = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot \begin{bmatrix} diag(\gamma)G & -diag(\gamma)B \\ diag(\gamma)B & diag(\gamma)G \end{bmatrix} \cdot x_{\mathrm{r}} \\ = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot M \cdot x_{\mathrm{r}} \end{aligned} \tag{3.46}$$

where, for this equation, $A = diag(\gamma)G, B = -diag(\gamma)B, C = diag(\gamma)B$ and $D = diag(\gamma)G$.

But it is noticed that (3.46) represents a bilinear form and that the J matrix of the quadratic form (represented by M in (3.46)), according to (3.33), is not symmetric. So, some more mathematical manipulations are necessary for the equation to represent a quadratic form.

Due to its constructive form and mathematical properties, the relation presented in (3.47) is true.

$$\gamma^{\mathrm{T}}P = P^{\mathrm{T}}\gamma \tag{3.47}$$

Equation (3.48) shows one of the matrices properties. If there is a matrix $M_{(\mathrm{n}\times\mathrm{n})}$, not necessarily symmetric, the matrix T is symmetric.

$$T = M + M^{\mathrm{T}} \tag{3.48}$$

A term ${x_{\mathrm{r}}}^{\mathrm{T}} \cdot M^{\mathrm{T}} \cdot x_{\mathrm{r}}$ will be added to both sides of (3.46), as shown in (3.49)

$${x_{\mathrm{r}}}^{\mathrm{T}} \cdot M^{\mathrm{T}} \cdot x_{\mathrm{r}} + \gamma^{\mathrm{T}}P = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot M \cdot x_{\mathrm{r}} + {x_{\mathrm{r}}}^{\mathrm{T}} \cdot M^{\mathrm{T}} \cdot x_{\mathrm{r}} \tag{3.49}$$

According to (3.48), (3.49) can be rewritten as (3.50)

$${x_{\mathrm{r}}}^{\mathrm{T}} \cdot M^{\mathrm{T}} \cdot x_{\mathrm{r}} + \gamma^{\mathrm{T}}P = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot T \cdot x_{\mathrm{r}} \tag{3.50}$$

From (3.22) and (3.47), (3.51) is obtained.

$$(\gamma^{\mathrm{T}}P)^{\mathrm{T}} = (x_{\mathrm{r}}{}^{\mathrm{T}} \cdot M \cdot x_{\mathrm{r}})^{\mathrm{T}}$$
$$P^{\mathrm{T}}\gamma = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot M^{\mathrm{T}} \cdot x_{\mathrm{r}} = \gamma^{\mathrm{T}}P \tag{3.51}$$

Equation (3.50) is modified according to (3.51) and is shown in (3.52)

$$2\gamma^{\mathrm{T}}P = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot T \cdot x_{\mathrm{r}}$$
$$\gamma^{\mathrm{T}}P = \frac{1}{2}x_{\mathrm{r}}{}^{\mathrm{T}} \cdot T \cdot x_{\mathrm{r}} \tag{3.52}$$

.

With the definition of T matrix, made in (3.48), and according to transpose properties, presented in (3.21), the equation (3.52) is rewritten as (3.53)

$$\gamma^{\mathrm{T}}P = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot \frac{1}{2} \cdot \begin{bmatrix} diag(\gamma)G + Gdiag(\gamma) & Bdiag(\gamma) - diag(\gamma)B \\ -Bdiag(\gamma) + diag(\gamma)B & diag(\gamma)G + Gdiag(\gamma) \end{bmatrix} \cdot x_{\mathrm{r}} = z_{\mathrm{i}} \tag{3.53}$$

As mentioned before, the same process can be done for Q, resulting in (3.54)

$$\gamma^{\mathrm{T}}Q = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot \frac{1}{2} \cdot \begin{bmatrix} -diag(\gamma)B - Bdiag(\gamma) & Gdiag(\gamma) - diag(\gamma)G \\ -Gdiag(\gamma) + diag(\gamma)G & -diag(\gamma)B - Bdiag(\gamma) \end{bmatrix} \cdot x_{\mathrm{r}} = z_{\mathrm{i}} \tag{3.54}$$

where *diag(γ)* is a diagonal matrix with (N×N) dimension.

If γ vector consists of zeros, except one position i, the values of $\gamma^T P$ and $\gamma^T Q$ are equal to P_i and Q_i, respectively. That means that (3.53) and (3.54) will depict the values at a specific bus i [75]. This particular γ vector will be defined as α and it will be used to simplify (3.53) and (3.54), transforming than into (3.55) and (3.56)

$$\alpha^{\mathrm{T}}P = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot \frac{1}{2} \cdot \begin{bmatrix} \alpha g^{\mathrm{T}} + g\alpha^{\mathrm{T}} & -\alpha b^{\mathrm{T}} + b\alpha^{\mathrm{T}} \\ \alpha b^{\mathrm{T}} - b\alpha^{\mathrm{T}} & \alpha g^{\mathrm{T}} + g\alpha^{\mathrm{T}} \end{bmatrix} \cdot x_{\mathrm{r}} \tag{3.55}$$

$$\alpha^{\mathrm{T}}Q = x_{\mathrm{r}}{}^{\mathrm{T}} \cdot \frac{1}{2} \cdot \begin{bmatrix} -\alpha b^{\mathrm{T}} - b\alpha^{\mathrm{T}} & -\alpha g^{\mathrm{T}} + g\alpha^{\mathrm{T}} \\ \alpha g^{\mathrm{T}} - g\alpha^{\mathrm{T}} & -\alpha b^{\mathrm{T}} - b\alpha^{\mathrm{T}} \end{bmatrix} \cdot x_{\mathrm{r}} \tag{3.56}$$

where g and b are the i^{th} column of G and B, respectively.

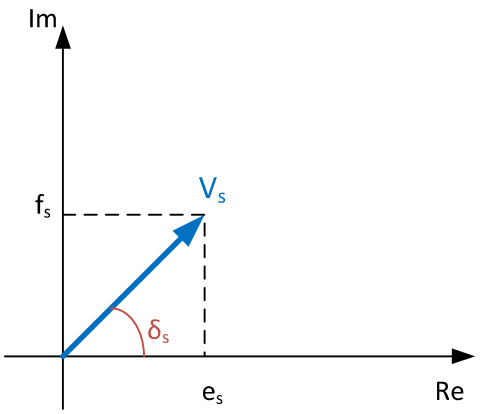
Following the same methodology, the quadratic form for V^2 can be described as (3.57)

$$\alpha^{\mathrm{T}} V^2 = {x_{\mathrm{r}}}^{\mathrm{T}} \cdot \begin{bmatrix} diag(\alpha) & O \\ O & diag(\alpha) \end{bmatrix} \cdot x_{\mathrm{r}} \tag{3.57}$$

where O is a (N×N) null matrix.

In power system load flow analysis, normally the reference bus (swing or slack bus) voltage angle are specified. This is necessary to provide a reference for the other angles. According to Figure 5, the relation between the reference bus voltage components are determined by (3.58) where δ_s is the reference angle and s is the reference bus.

Figure 5 – Relation of the Real and Imaginary components of bus voltage V_s



Source: Author

$$\begin{aligned} f_{\mathrm{s}} &= e_{\mathrm{s}} \cdot tan(\delta_{\mathrm{s}}) \\ e_{\mathrm{s}} \cdot tan(\delta_{\mathrm{s}}) - f_{\mathrm{s}} &= 0 \end{aligned} \tag{3.58}$$

The equation (3.58) can still be rewritten as (3.59) considering α_s a (2N×1) vector of zeros except for the s^{th} and $(\mathrm{N} + \mathrm{s})^{\mathrm{th}}$ positions which are equal to $\tan(\delta_s)$ and -1, respectively. Equation (3.59) describes an hyperplane through the origin.

$${\alpha_{\mathrm{s}}}^{\mathrm{T}} x_{\mathrm{r}} = 0 \tag{3.59}$$

For each one of the PQ and PV buses of the power system analyzed there are two quadratic forms describing these injections (that depends of the bus type). At the slack or swing bus there is one quadratic equation describing the voltage magnitude squared and

at reference bus there is an additional equation that specifies informations about (3.58) and (3.59) [75].

The injection equations form (2N×1) non-linear equations and with (3.58) and (3.59) compose 2N load flow equations, which characterise the static behaviour of the power system [75].

The study of the eigenvalues and eigenvectors of the J matrix, matrix that define the power flow injections, is important to analyze some characteristics of the quadratic forms. Due to the structure of the matrix J, it is possible to find the eigenvalues and eigenvectors using non-iterative means, regardless of the system size [75].

When the J matrix represents a power injection, only four eigenvalues are non-zero (and consequently, there are only four eigenvectors associated). For a squared voltage injection, there are two non-zero eigenvalues and two eigenvectors [75].

This analysis is described in Sections (3.3.1.1) and (3.3.1.2) below.

3.3.1.1 Eigenvalues and Eigenvectors of J_{P_i} and J_{Q_i}

In section (3.2.1.3) the eigenvalues were defined in (3.24). Starting from this equation, to find the eigenvalues and eigenvectors of matrix J, (3.60) can be used

$$J \cdot x_r = \lambda \cdot x_r \tag{3.60}$$

When J defines a real power injection, it can be represented as shown in (3.55). So, knowing the definition of x_r, presented in (3.35), (3.60) can be developed as (3.61)

$$\frac{1}{2} \cdot \begin{bmatrix} \alpha g^T + g\alpha^T & -\alpha b^T + b\alpha^T \\ \alpha b^T - b\alpha^T & \alpha g^T + g\alpha^T \end{bmatrix} \cdot \begin{bmatrix} e \\ f \end{bmatrix} = \lambda \cdot \begin{bmatrix} e \\ f \end{bmatrix} \tag{3.61}$$

From here, a series of substitutions and mathematical manipulations are made to find eigenvalues and eigenvectors in a non traditional way, which was not used in this thesis. More information on the mathematical step-by-step can be found in Appendix B of [75].

The mathematical procedures conclude that in this case, the matrix that defines the quadratic form for real power injection has two identical pairs of eigenvalues, represented by λ_{P_1}, λ_{P_2}, λ_{P_1}, λ_{P_2}.

The same procedure can be done for the matrix J defining the reactive power injection. This matrix also has two identical pairs of eigenvalues, represented by λ_{Q_1}, λ_{Q_2}, λ_{Q_1}, λ_{Q_2}.

The eigenvalues and eigenvectors can be found mathematically, using the methodology described by [75], for example, or using a software, which was done in this thesis.

3.3.1.2 Eigenvalues and Eigenvectors of $J_{V_i}{}^2$

The J matrix that describes V^2 at bus i has two non-zero eigenvalues, and both of them are equal to 1 [75]. The two eigenvectors associated to the eigenvalues are given by (3.62), where 1 is in the i^{th} position, and (3.63) where 1 is in the $(N+i)^{th}$ position.

$$e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.62}$$

$$e_{N+i} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.63}$$

3.3.2 (2N-1) Equations

When a bus is defined to be the reference bus, its angle serves as a reference for the other buses. With this, it is possible to reduce the number of variables unknown by one, as said in Section 3.3.1. Usually the slack bus is taken as the reference bus [65, 66, 75] and with this consideration, two unknowns can be eliminated for the mathematical analysis.

This results in (2N-2) equations that are no longer fully quadratic but they have a linear term x that specifies the reference bus. Thus, some mathematical manipulations are necessary to ensure that it is still possible to represent the power flow in quadratic forms.

The first step is use (3.58) and (3.59) to eliminate f_s of the reference bus, that represents the imaginary part of voltage, as (3.34). This produces a number of (2N-1)

purely quadratic equations with (2N-1) unknowns, implying a reduction of the matrix J dimensions from (2N)×(2N) to (2N-1)×(2N-1). To make this reduction successfully, one more time some mathematical manipulations are necessary. First, a new variable d is defined, as shown in (3.64).

$$\begin{aligned} d &= \mathcal{M} \cdot x_{\mathrm{r}} \\ &or \\ \mathcal{M}^{-1} \cdot d &= x_{\mathrm{r}} \end{aligned} \tag{3.64}$$

where $\mathcal{M}$ is a matrix similar to the identity matrix Id, but its $(\mathrm{N+s})^{\mathrm{th}}$ line describes the initial conditions of the reference bus, shown in (3.58) and (3.59). So, $\mathcal{M}$ can be defined as (3.65).

$$\mathcal{M} = Id - e_{N+s} \cdot {e_{N+s}}^{T} + e_{N+s} \cdot {\alpha_{\mathrm{s}}}^{\mathrm{T}} \tag{3.65}$$

With the information presented by (3.64) and (3.65), (3.33) can be rewritten as (3.66), where J_i has dimension (2N×2N) and can be defined as (3.55), (3.56) or (3.57) and $J_{s_i} = (\mathcal{M}^{-1})^{T} \cdot J_i \cdot \mathcal{M}^{-1}$.

$$\begin{aligned} z_{\mathrm{i}} &= (\mathcal{M}^{-1} \cdot d)^{\mathrm{T}} \cdot J_{\mathrm{i}} \cdot (\mathcal{M}^{-1} \cdot d) \\ z_{\mathrm{i}} &= d^{\mathrm{T}} \cdot (\mathcal{M}^{-1})^{\mathrm{T}} \cdot J_{\mathrm{i}} \cdot \mathcal{M}^{-1} \cdot d \\ z_{\mathrm{i}} &= d^{\mathrm{T}} \cdot J_{\mathrm{s}_i} \cdot d \end{aligned} \tag{3.66}$$

According to (3.64) and (3.65), the $(\mathrm{N+s})^{\mathrm{th}}$ element of matrix d ($e_s \cdot tan\ (\delta_s)$ - f_s), holds reference bus information and, according to (3.58), is equal to zero. Due to this, it can be eliminated. The reduced (2N-1)×(2N-1) matrix J_i can be obtained by deleting the $(\mathrm{N+s})^{\mathrm{th}}$ row and column of J_{s_i}.

The reduced matrix J_i is symmetric, which allows the quadratic form representation. Its s^{th} row and column have elements with the term $tan\ (\delta_s)$. So, the reduced matrix varies according to the chosen reference bus. More information about this can be found on Appendix C of [75].

It is important to say that, in most of the cases, the angle chosen for the reference bus is zero ($\delta_s = 0$). This simplifies the mathematical procedure, since $tan\ (0^o) = 0$. So, it is possible to obtain the reduced (2N-1)×(2N-1) J_i matrix only by simply deleting the $(\mathrm{N+s})^{\mathrm{th}}$ row and column of the (2N×2N) matrix.

For the same reasons as explained in Section (3.3.1) it is important to study and analyze the eigenvalues and eigenvectors when the slack bus is taken as reference bus. The

difference here, as already mentioned, is that the matrices that defines the injections are (2N-1)×(2N-1).

When matrix J_i describes real and reactive power, the matrix has, one more time, two pairs of eigenvalues. One of these pairs is identical to the values found for the (2N×2N) matrix.

3.3.2.1 Eigenvalues and Eigenvectors of J_{P_i} and J_{Q_i}

The first pair of eigenvalues of J_{P_i}, λ_{P_1} and λ_{P_2} for real power and J_{Q_i}, λ_{Q_1} and λ_{Q_2} for reactive power is defined in Section 3.3.1.1. As mentioned above, a software can be used to find the second pair.

3.3.2.2 Eigenvalues and Eigenvectors of $J_{V_i}{}^2$

The eigenvalues and eigenvectors of the matrix J_i that describes the squared voltages are the same obtained in Section 3.3.1.2.

3.4 Quadratic Surfaces Defined by J_i

The real and reactive power injections, denoted by J_i matrix are represented graphically by hypersurfaces. The four eigenvectors associated to the eigenvalues obtained from (2N-1)×(2N-1) J_i matrix represent a set of principal axes of this particular surface [75].

The extreme points of this hypersurface are obtained by determining their intercepts with the principal axes. This can be done by first expressing J_i in its canonical form, as shown in (3.67).

$$z_\mathrm{i} = \sum_{k=1}^{4} \lambda_\mathrm{k} (e_\mathrm{k}{}^\mathrm{T} x)^2 \tag{3.67}$$

where z_i is the power injection, λ_k is the eigenvalue and v_k is the normalised eigenvector of J_i.

The intersection points are defined by c, and can be represented by (3.68), where z_i and λ must be of the same sign.

$$c = \pm\sqrt{\frac{z_\mathrm{i}}{\lambda}} \tag{3.68}$$

So, c is a mathematical parameter that holds important information about the power flow in a bus. A c value can be found for each one of the feeder buses, and it assigns a value for the bus according to the power flow characteristics, that is, according to the

electrical characteristics of the entire network. Considering that z_{i} is the power injection, and can be P or Q, it is possible to define two parameters of interest: c_P and c_Q which represents the c value related to the real power and reactive power, respectively.

The load flow equations can be solved by finding the intersection of 2N quadratic hypersurfaces and one hyperplane in a 2N dimensional space or by finding the intersection of (2N-1) quadratic hypersurfaces and one hyperplane in a (2N-1) dimensional space. More information about this load flow analysis can be found at [75].

This chapter presented a mathematical formulation based in [75] that can be used as an alternative way to calculate the power flow. In the course of the method it is possible to obtain the c value depicted by (3.68), which assigns a value to each bus based on the power flow of the network under analysis. The c value is a mathematical representation that graphically represents the intersection of the hypersurfaces and their principal axes, which is an alternative mathematical way of solve the load flow equations.

This c value will be used in the development of the methodology proposed by this thesis, shown in chapter 5, as a weight for the buses of the studied electrical feeder.

4 GRAPH THEORY

In the last decades, a new field of research know as “Complex Network Analysis” (CNA) emerged. This tool is useful to study and better understand the properties of large networks that behave as complex systems [35]. The CNA has been used in different fields, not only in technology. It can be used to study biology and chemistry phenomena, linguistics and social sciences behaviors, from telephone call patterns to computer networks and web to virus spreading to logistics and also inter-banking systems [35].

Some examples illustrate the wide range of applications and level of analysis. An interesting one shows the Google’s PageRank, a procedure that considers the web as a complex network and analyzes a huge number of documents from the Web aiming to “answer keyword queries by returning documents that appear relevant to the query” [82]. So, for this case, each web document has a measure of relevance assigned.

An electric power grid, either power transmission or distribution is a perfect example of a complex network, specially considering the recent incorporation of information and communication technologies and also the growing of DG that makes the complexity of the grid increase [13, 31]. Thus, statistical tools, techniques and methodologies of CNA have been used to identify critical buses, mitigate the effects of a cyber attack, prevent network failures and contingencies and avoid blackouts, ensuring system reliability [13, 31, 35, 83].

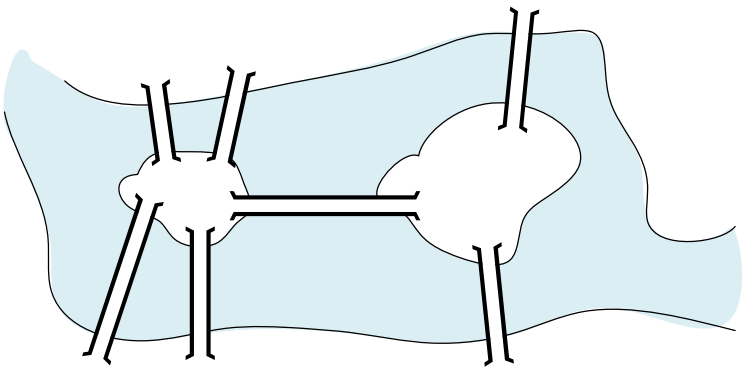
An interesting way to study some characteristics of a complex network is using principles of graph theory. That is a very suitable tool when it is desired to find patterns and it can provide a framework to analyze connections in different types of network and other practical applications. Hence, it has been widely used for electrical networks [13, 14]. Through graph theory, it is possible to use topological metrics of centrality, metrics that use electrical parameters such as impedance and admittance and power flow metrics to analyze power systems parameters.

The first register of the use of graph theory is related to the famous Königsberg bridge problem, a mathematical problem discussed in Königsberg city, Prussia, solved by Leonard Euler [84, 85]. The problem consisted of a river which surrounded two central islands connected by a bridge. The two islands were connected to the mainland portions of the city by six bridges, as shown in Figure 6. The challenge was to find a route starting from any of the river banks and reaching the same bank by crossing each bridge just once [86].

The *network* term utilized in this chapter refers to the informal concept describing an object composed of elements and interactions or connections between these elements [82]. For example, in a power grid, the buses of the system can be treated as nodes and

transmission or distribution lines represent the connections between these nodes, the edges.

Figure 6 – Bridges of Königsberg



Source: Author

4.1 Definitions

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is an abstract object formed by a finite and non-empty set $\mathcal{V}$ of *vertices* (nodes), and a set of finite $\mathcal{E}$ of *edges* (links) that connect pairs of vertices [82, 87]. A set of n vertices can be represented by $\mathcal{V}(\mathcal{G}) = \{v_1, v_2, \cdots, v_{\mathrm{n}}\}$ and a set of m edges is denoted by $\mathcal{E}(\mathcal{G}) = \{e_1, e_2, \cdots, e_{\mathrm{m}}\}$. If two vertices are connected by an edge, they are adjacent and can be called neighbors.

Graphs can be undirected or directed. In an undirected graph, the vertices are connected but there is no distinction between the source and destination vertices. In a directed graph, it is necessary to identify the order of the endvertex of an edge. An edge with source $u \in \mathcal{V}$ and destination $v \in \mathcal{V}$ is represented by an ordered pair (u, v). For undirected graphs, (u, v) and (v, u) are the same. Depending on the application, it is useful to associate numerical values (weights) with the edges and/or vertices of a graph, characterising an weighted graph. A graph is called simple or strict when there is no loops or parallel edges, this means there is no edge joining a vertex to itself [82].

Some other important definitions about graph are [82, 83, 87, 88]:

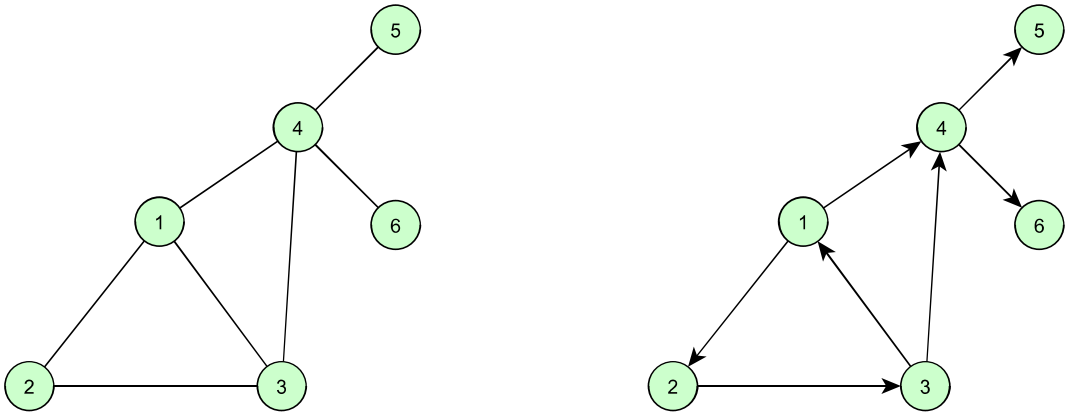
- *Walk*: a walk from v_0 to v_k in a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is an alternating sequence v_0, e_1, v_1, e_2, $\cdots$, $v_{k\text{-}1}$, $e_{k\text{-}1}$, v_k of vertices and edges. In a walk there may be repetition of edges and vertices. The length of the walk is defined as the number of edges on the walk;
- *Path*: is a walk with no repetition of edges and vertices;
- *Geodesic*: is the shortest path that links two nodes;
- *Cicle*: is a path that starts and ends at the same node;

- *Distance*: is the number of edges of a geodesic between two specific nodes u and v. If no such path exists, then the distance is set equal to infinite;
- *Acyclic Graph*: is a graph with no cycles;
- *Diameter*: is the length of the longest geodesic between any two nodes of $\mathcal{G}$.

The degree of a vertex v, denoted as $d(v)$, is the number of nodes adjacent to v in $\mathcal{G}$. The minimum degree of a vertex of $\mathcal{G}$ is depicted by $\delta(\mathcal{G})$ and the maximum degree, $\Delta_g(\mathcal{G})$ [82]. If there is at least one path between every pair (u, v) of nodes in $\mathcal{G}$, the $\mathcal{G}$ is called a connected graph [82,83]. Figure 7 shows two simple graphs; an undirected graph with $n = 6$ vertices and $m = 7$ edges is depicted in Figure 7a and the same topology but represented by a directed graph is shown in Figure 7b, where there are in-degree and out-degree. Some information as the degree and adjacent vertices for each node are presented in Tables 2 and 3.

Figure 7 – Undirected and Directed Simple Graph

(a) An undirected graph $\mathcal{G}_1$ (b) A directed graph $\mathcal{G}_2$



Source: Author

Table 2 – Degree and Adjacent vertices for the simple graph presented by Figure 7a information

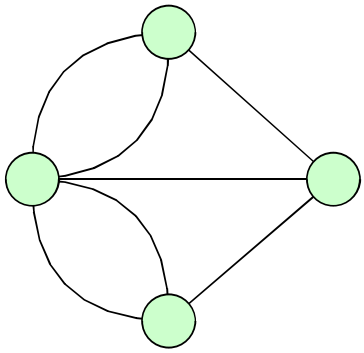
Node	**Degree**	**Adjacent vertices**
1	3	{2, 3, 4}
2	2	{1, 3}
3	3	{1, 2, 4}
4	4	{1, 3, 5, 6}
5	1	{4}
6	1	{4}

Table 3 – Degree and Adjacent vertices for the simple graph presented by Figure 7b information

Node	**In-Degree**	**Out-Degree**	**Adjacent vertices**
1	1	2	{2, 3, 4}
2	1	1	{1, 3}
3	1	2	{1, 2, 4}
4	2	2	{1, 3, 5, 6}
5	1	-	{4}
6	1	-	{4}

With this informations, now the Königsberg bridge problem can be represented as a graph. So Figure 6 can be transformed into Figure 8. In a simple graph, a pair of vertices cannot be linked to more than one edge. When the graph has this property, one can say that it is a multigraph. The graph of the Königsberg bridge problem is, therefore, a multigraph.

Figure 8 – Graph of Bridges of Königsberg



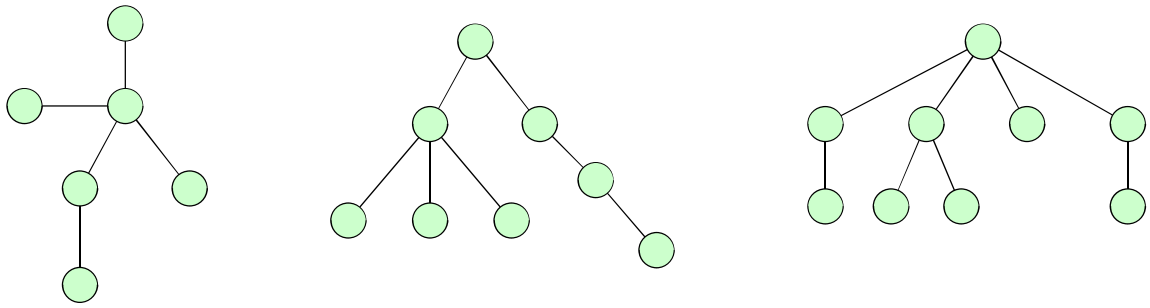
Source: Author

Besides being widely used in computactional applications, as the scope of this thesis if study the electrical power distribution systems, which are, in most of the cases, radial, the definition of "tree" is important. A connected acyclic graph is a tree. Some examples of trees are presented in Figure 9.

4.2 Matrix Representation

The graphs can be represented by matrices. The adjacency matrix of a simple undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n vertices is a $(n \times n)$ matrix represented by $A(\mathcal{G})$. The elements of $A(\mathcal{G})$ are defined as (4.1) [82]. In a simple way the $A(\mathcal{G})$ matrix shows the vertices that are connected in an undirected graph.

Figure 9 – Examples of trees



Source: Author

$$a_{ij} = \begin{cases} 1, & \text{if } (v_i, v_j) \in \mathcal{E} \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

If $\mathcal{G}$ is an undirected graph, $A(\mathcal{G})$ is symmetric. An example of adjacency matrix is shown below. This matrix represents the graph of Figure 7a.

$$A(\mathcal{G}_1) = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The sum of each line of the $A(\mathcal{G}_1)$ matrix presented represents the degree of the vertex as shown in Table 2. This relation can be verified for undirected graphs.

For comparison, the adjacency matrix of Figure 7b is also shown. By analysing the two matrices $A(\mathcal{G}_1)$ and $A(\mathcal{G}_2)$, it is possible to identify the differences in $A(\mathcal{G})$ between a directed and an undirected graph.

$$A(\mathcal{G}_2) = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

If the graph analysed is a weighted graph, the "1" of the $A(\mathcal{G})$ matrix will be replaced by the specific edge weight that links the vertex i to vertex j.

Another important matrix is the diagonal matrix, *D(G)*. In this matrix, which is a diagonal matrix, as its names indicates, the non-zero term of each i^{th} line depicts the degree of the i^{th} vertex. For Figure 7a, *D(G)* is given by:

$$D(\mathcal{G}_1) = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The Laplacian matrix of an undirected graph $\mathcal{G}$ is an $n \times n$ symmetric positive-semidefinite matrix defined as (4.2)

$$L(\mathcal{G}) = D(\mathcal{G}) - A(\mathcal{G}) \tag{4.2}$$

where $D(\mathcal{G})$ is the diagonal matrix and $A(\mathcal{G})$ is the degree matrix, as already defined.

The matrix representation allows specific studies of the graph and for some applications makes the analysis simpler.

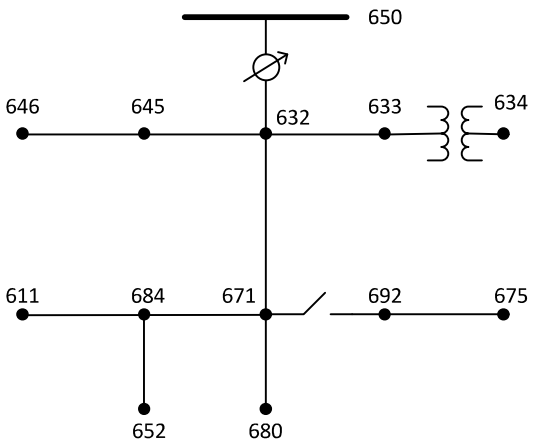
4.3 Representing an electrical power grid as a simple graph

As mentioned before, the scope of this thesis is to study the distribution power network. An electrical system can be represented by a graph where the buses are modelled as the vertices and the distribution lines are the edges. A simple example of a power system representation is shown in Figure 10. Figure 10a depicts an electric feeder that has, although not represented clearly by this figure, load nodes, transformers and a voltage regulator and the node 650 is the substation (the only generator of the system). The transmission system has some typical characteristics as balanced phases and bidirectional flow, which means that, according to the demand, the power can flow in any direction on the transmission lines. The distribution systems are normally radial and unbalanced. So, according to the system characteristic, the graph of the electrical system can be modelled. Figure 10b shows the representation of the 13-Node Test Feeder (13NTF) Graph. For this case, as the system describes a distribution system, the graph is directed. It is possible to observe that this graph represents a tree. The graph vertices were labelled numbered according to the respective buses number. The edges could have been represented as weighted edges. The weights can be defined according to several criteria, such as a length or admittance.

For some applications, it will be required to work with a weighted graph - the vertex and/or the edge is weighted. For others, the weight can be neglected. Even for a distribution system, that usually is radial, there may be the need to the graph be undirected for a specific study. It is clear that the graph model depends on the system, the application and the parameters that will be studied.

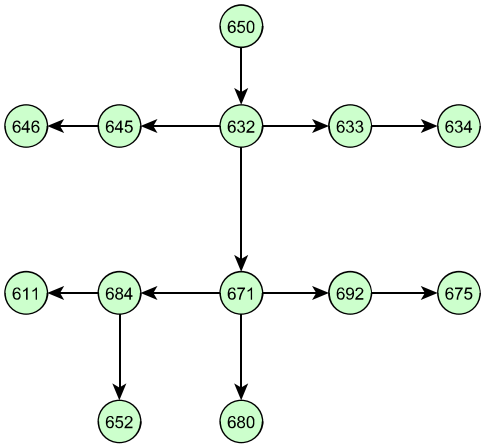
Figure 10 – 13 Node IEEE Test Feeder

(a) 13 Node Test Feeder



Source: [89]

(b) 13 Node Test Feeder Graph



Source: Author

4.4 Centrality Measures

Centrality measures are indices used to quantify the importance of the nodes and edges of a graph. As the name indicates it is a measure of how central, based in some

specific criteria, the vertex and edge are compared to the others in a complex network [82, 83].

There are several measures of centrality defined and largely used in the academic community [82, 83, 86]. Some authors use to modify the classic measures and transform them into new measures to meet a specific application. This is due to the fact that not every centrality measure is suitable to every application. Besides, the same index may represent different meanings according to the network. For example, for a social network or an electric system network, a specific centrality index may suggest different interpretations of the edge/vertex importance [82].

The choice of the centrality measure used, therefore, will depend a lot on the complex network analyzed, the criterion of interest and how that result will be interpreted. In an electrical system, it is interesting to determine the most central nodes to analyse the system vulnerability, preventing failures, blackouts and choosing the best bus to connect the DG [3, 83, 90].

As it is one of the most important parameters for the graphs, the centrality measures will be defined. In this thesis, only the more classic centrality measures will be presented bellow.

4.4.1 Degree

The Degree Centrality of a vertex i, $c_D(i)$, is the most simple centrality measure. It represents the connections of a specific node to the rest of the network [3, 82, 91].

For a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n vertices, the mathematical formulation for the Degree Centrality of a vertex i is presented by (4.3)

$$c_D(i) = \frac{d(i)}{n-1} \tag{4.3}$$

where $d(i)$ is the degree of vertex i and $(n-1)$ is a normalization factor.

To exemplify, Table 4 shows the Degree Centrality for each vertex of Figure 7a.

4.4.2 Betweenness

The Betweenness Centrality, first proposed by [91], is one of the most used indices for centrality measure and it indicates the frequency in which a vertex fall within the shortest paths (geodesics) of all pairs of nodes. This measure reflects the influence of a vertex within a graph, how the node is communicable with the rest of the graph [82, 83, 92].

Table 4 – Degree Centrality for Figure 7a

Node	**Degree**	**Degree Centrality**
1	3	$c_D(1) = \frac{3}{5} = 0.6$
2	2	$c_D(2) = \frac{2}{5} = 0.4$
3	3	$c_D(3) = \frac{3}{5} = 0.6$
4	4	$c_D(4) = \frac{4}{5} = 0.8$
5	1	$c_D(5) = \frac{1}{5} = 0.2$
6	1	$c_D(6) = \frac{1}{5} = 0.2$

The Betweenness Centrality of a vertex $i \in \mathcal{V}$, $c_B(i)$, is computed as (4.4)

$$c_B(i) = \sum_{s \neq i \neq t \in \mathcal{V}} \frac{\sigma_{st}(i)}{\sigma_{st}} \tag{4.4}$$

where σ_{st} is the number of shortest paths from s to t and $\sigma_{st}(i)$ is the number of shortest paths passing through i.

To illustrate, the Betweenness Centrality for vertex 1 of Figure 7a is calculated as bellow:

$$\begin{aligned} c_B(1) = \frac{\sigma_{2,3}(1)}{\sigma_{2,3}} + \frac{\sigma_{2,4}(1)}{\sigma_{2,4}} + \frac{\sigma_{2,5}(1)}{\sigma_{2,5}} + \frac{\sigma_{2,6}(1)}{\sigma_{2,6}} + \frac{\sigma_{3,4}(1)}{\sigma_{3,4}} + \frac{\sigma_{3,5}(1)}{\sigma_{3,5}} + \frac{\sigma_{3,6}(1)}{\sigma_{3,6}} + \\ + \frac{\sigma_{4,5}(1)}{\sigma_{4,5}} + \frac{\sigma_{4,6}(1)}{\sigma_{4,6}} + \frac{\sigma_{5,6}(1)}{\sigma_{5,6}} \\ = \frac{0}{1} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1} = 1.5 \end{aligned}$$

Some authors normalize the $c_B(i)$ by $\frac{2}{(n-1)(n-2)}$ for undirected graphs, and $\frac{1}{(n-1)(n-2)}$ for directed graphs, where n is the number of nodes in $\mathcal{G}$ [3].

The Betweenness Centrality for each vertex of Figure 7a is shown in Table 5.

4.4.3 Closeness

The Closeness centrality, $c_c(i)$ can be used when it is important to identify how fast an information can spread from one node to every other nodes of a network. It is the

Table 5 – Betweenness Centrality for Figure 7a

Node	**Betweenness Centrality**	**Normalized Betweenness Centrality**
1	1.5	0.15
2	0	0
3	1.5	0.15
4	7	0.7
5	0	0
6	0	0

best position in the graph to guarantee that the distance to all other vertices are minimal [3, 82, 92].

Mathematically, according to [91], the closseness centrality is computed by

$$c_c(i) = \frac{1}{\sum\limits_{t \in V \setminus i} dist(i,t)}, \tag{4.5}$$

where $dist(i,t)$ is the distance between vertices i and t. So the Closeness centrality increases with decreasing total distance of vertex i related to the other vertices.

As the same way as in Betweenness Centrality, some authors normalize the $c_c(i)$ by multiplying it by $(n-1)$ where n is the number of vertices of the graph [3].

The Closeness Centrality rounded values for each vertex of Figure 7a is shown in Table 6.

Table 6 – Closeness Centrality for Figure 7a

Node	**Closeness Centrality**	**Normalized Closeness Centrality**
1	0.14	0.71
2	0.1	0.5
3	0.14	0.71
4	0.17	0.83
5	0.1	0.5
6	0.1	0.5

4.4.4 Eigenvector

The Eigenvector Centrality is a spectral centrality measure that depends on spectral properties of graph matrices. It was introduced in 1972 by Phillip Bonacich [93, 94] and it

is based on eigenvalues and their respective eigenvectors. For this measure, the relevance of a node depends on the importance of its neighbors [95].

The eigenvector index is calculated by finding the eigenvector denoted by v_{n-1} associated with $\lambda_{max} = \lambda_{n-1}$ [31, 34, 93].

The eigenvector centrality measure for each node i of a graph can be calculated by

$$c_e(i) = \frac{1}{\lambda_{max}} \sum_{j=i}^{n} |Ax_j| \tag{4.6}$$

Once again, considering the Figure 7a, the eigenvector centrality for each vertex is presented in Table 7.

Table 7 – Eigenvector Centrality for Figure 7a

Node	**Eigenvector Centrality**
1	0.51
2	0.38
3	0.51
4	0.51
5	0.19
6	0.19

As already mentioned, if the graph is undirected, the adjacency matrix will be symmetric. Symmetric matrices results in real eigenvalues.

In Table 8, it is possible to compare the results for the presented centrality measures. As the example shown in Figure 7a is simple, with no weights and with a small number of nodes, the results for the rank in the different centrality measures were very similar. The most central nodes for each centrality is highlighted in red and underlined in Table 8. Despite that, it is possible to observe some difference between the values and the rank of the nodes. This reinforces, once again, the need for a good modelling and delimitation of the problem.

4.5 Signal Processing on Graphs (SPG)

All the centrality measures presented in Section 4.4 above are widely used in the literature and consolidated in different areas of knowledge. Nevertheless, as stated previously, each centrality measure must be used according to the analyzed network and the expected data output. An interesting observation point is that the most traditional centrality measures are focused on ranking nodes based on edge weights or number of edges. It is not usual to weight the nodes when modelling the network and complex numbers are usually not allowed as weights.

Table 8 – Centrality Measure Comparison for Figure 7a

Node	Degree Centrality	Normalized Betweenness Centrality	Normalized Closeness Centrality	Eingenvector Centrality
1	0.6	0.15	0.71	0.51
2	0.4	0	0.5	0.38
3	0.6	0.15	0.71	0.51
4	0.8	0.7	0.83	0.51
5	0.2	0	0.5	0.19
6	0.2	0	0.5	0.19

Considering that the electrical load flow analysis quantities are complex numbers and as this thesis is interested on weighting the nodes using the "c" value present in the Section 3.4, equation (3.68), an alternative approach for graph analysis will be used for measure of importance, the Signal Processing on Graphs (SPG).

The SPG is a suitable tool to represent, model, process and analyze the complex networks, where discrete signals are defined on the nodes of a weighted (or not) graph. A lot of researches in the literature extended the classical concepts of discrete signal processing theory from time signals, images or data to signals defined over the nodes of a graph. The use of graphs in data processing including operations such as filtering, denoising, inpainting, and compressing graph signals encouraged researchers to extract information, statistically or visually, in an efficient way [96, 97].

Typically in a graph, weights are defined for the edges, indicating the strength of bonding between nodes. With this information, it is possible to calculate the centrality measures for each node, which can be interpreted as a finite collection of samples. Depending on the network or problem being studied, weights can be given specifically for each node, which can also be considered a finite collection of samples. For both the situations mentioned, a term can be used for the samples indicating the weight or strength of each node: graph signal, which is a fundamental and important concept for understanding and applying the theory of graph signal processing [46, 96].

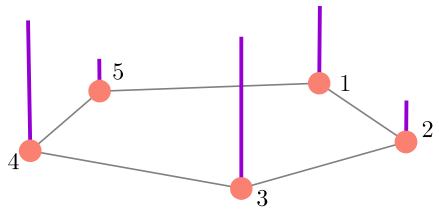
So, as that SPG appears as a way to solve data processing problems, it also comes as an extra option for measuring the importance of a node in a complex network and merges algebraic and spectral graph theoretic concepts with computational analysis to process and analyze the signals on graphs [46, 96].

Figure 11 presents a generic graph with random graph signal depicted in purple bars. The height of each purple bar represents the signal value, corresponding to the graph signal at that node. Although not represented in Figure 11, the weight values can be negative or complex numbers [96].

The relation between the classical signal processing and SPG comes from the fact

that the signal on a graph with n vertices and a classical discrete-time signal with n samples can be viewed as vectors in $\mathbb{R}^n$ [96]. The n-dimensional vector cited is represented by $f = [f(1)\ f(2)\ f(3)\ \ldots\ f(n)]^{\mathrm{T}}$, where $f(i)$ depicts the graph signal at i^{th} node and $n = |\mathcal{V}|$ is the graph total number of nodes [46]. For the graph represented by Figure 11 the n-dimensional vector that presents the graph signal is $f = [1.5\ \ 1.2\ \ 1.8\ \ 1.9\ \ 1.25]^{\mathrm{T}}$ (random values).

Figure 11 – Graph signal on the vertices of a undirected graph



Source: Author

Some mathematical tools as classical discrete signal processing, wavelet, time-frequency, curvelet and other transforms have been used in the analysis of SPG. However, the spectral analysis has a prominent role in the signal processing of graph field and has been utilized as a tool to define frequency spectra for graph Fourier transforms. So, the Fourier transform is a key tool for analyzing these signals and to transform them to other domain, performing an harmonic analysis of graph signals. The Graph Fourier Transform (GFT) is used in most research in the field of SPG [46, 96, 97].

4.5.1 Graph Fourier Transform (GFT)

Knowing that some mathematical problems can be difficult to solve directly, the use of transforms appears as a way to simplify the solution. The data can be transformed, resolved and then, if necessary, the inverse transform can be used.

In the literature, the representation of a signal in the time domain is common in the daily situations. But some analysis are much simpler if the data are in the frequency domain, which is possible to be achieved, in the classical mathematical approach, by using Fourier Transform.

The classical Fourier transform is the expansion of a function f in terms of the complex exponentials and is defined by (4.7). From a time function $f(t)$, the corresponding frequency function $\hat{f}(\xi)$ is obtained. The inverse transform can also be applied, from $\hat{f}(\xi)$ to find $f(t)$. The result is usually a complex number [96].

$$\hat{f}(\xi) := \langle f, e^{j2\pi\xi t} \rangle = \int_{\mathbb{R}} f(t) e^{-j2\pi\xi t} dt \tag{4.7}$$

where e is the Euler number and $e^{-j2\pi\xi t} = cos(2\pi\xi t) - j \cdot sin(2\pi\xi t)$. As already mentioned, j is the complex number.

In the traditional Fourier analysis, the complex exponentials of (4.7) are the eigenfunctions of the one-dimensional Laplace operator, defined by (4.8) and the eigenvalues $\{(2\pi\xi)^2\}_{\xi \in \mathbb{R}}$ carry the notion of frequency to ξ. This means that for ξ values close to zero, indicating low frequencies, the complex exponential functions are smooth and oscillate slowly. When ξ is far from zero, high frequencies, the functions oscillate faster [96].

$$-\Delta(e^{j2\pi\xi t}) = \frac{\partial^2}{\partial t^2} e^{\{j2\pi\xi t\}} = (2\pi\xi)^2 e^{\{j2\pi\xi t\}} \tag{4.8}$$

In the graph domain, the graph Laplacian eigenfunctions produce an analogous notion of frequency. Considering a connected graph $\mathcal{G}$, the Laplacian eigenvector u_0 associated with the eigenvalue 0, λ_0, is constant and equal to $\frac{1}{\sqrt{n}}$, where n is the total number of vertices in $\mathcal{V}$, as defined in Section 4.1 [96].

As stated before, the graph Laplacian eigenvectors associated with low frequencies λ_l vary slowly across the graph, which means that when two nodes are connected by an edge with a high weight, their eigenvectors tends to be similar. The eigenvectors associated with larger eigenvalues oscillate more rapidly and then, tends to have distinct values on vertices connected by an edge with high weight [96].

Starting from classical theory, the GFT $\hat{f}$ of any function $f \in \mathbb{R}^n$ on the vertices of $\mathcal{G}$ is depicted by (4.9) [96]:

$$\hat{f}(\lambda_l) := \langle f, u_l \rangle = \sum_{i=1}^{n} f(i) u_l{}^*(i) \tag{4.9}$$

where u_l is the eigenvector associated to the eigenvalue λ_l, the symbol * depicts the conjugate transpose and f is the graph signal. The inverse GFT can also be obtained [96].

The graph Laplacian L was already defined in Section 4.1. In the GFT approach, however, it is very common to work with the Laplacian with a different definition. First, some definitions will be made to help the understanding [98].

Definition 1: *Considering an weighted undirected graph $\mathcal{G} = (\mathcal{V}; \mathcal{E})$ with $|\mathcal{V}| = n$, the weighted adjacency matrix is the symmetric matrix $W = \{w_{ij}\}$ for i, j = 1,2,3, ..., n, where w_{ij} is the weight of edge $\{v_i, v_j\}$. If $w_{ij} = 0$ the vertices v_i and v_j are not connected through an edge* [98].

Definition 2: *The degree of a node $v_i \in \mathcal{V}$ is defined by the sum of the elements*

of the i^{th} row of W, as (4.10)*:*

$$d_i = \sum_{j=1}^{n} w_{ij} \tag{4.10}$$

The degree matrix D_w is defined as the diagonal matrix with degrees d_1,d_2,d_3,...,d_n [98].

With the two definitions above, it is possible to define the unnormalized weighted graph Laplacian matrix that will be used in the GFT analysis as (4.11):

$$L_{GFT} = D_w - W \tag{4.11}$$

where D_w is the degree matrix and W is the adjacency matrix determined by **Definition 1**.

It is important to reinforce that there are some mathematical properties associated with the Laplacian matrix. Furthermore, to be considered a "Laplacian matrix L", the matrix must satisfies some properties. The L_{GFT} presented by (4.11) meets these requirements [98].

The L_{GFT} matrix is a real symmetric matrix, which implies in a complete set of orthonormal eigenvectors denoted by $\{u_l\}_{l\,=\,0,\,1,\,2,\,\ldots\,,\,n\,-\,1}$, that can also be represented as $[u_0\ u_1\ u_2\ \ldots\ \ u_{n\,-\,1}]$. This set of eigenvectors is associated to a set of real non-negative eigenvalues $\{\lambda_l\}_{l\,=\,0,\,1,\,2,\,\ldots\,,\,n\,-\,1}$. So, the eigenvalues of L_{GFT} in ascending order are $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} := \lambda_{max}$. This set of eigenvalues is the spectrum of the graph which provides the notion of frequency, as already cited, and the entire spectrum can be presented by (4.12) [46, 96] .

$$\sigma(L_{GFT}) := \{\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_{N-1}\} \tag{4.12}$$

Thus, the GFT allows the representation of the same information in two domains: the vertex domain and the graph spectral domain. An interesting advantage of using GFT is the possibility of working with a graph signal vector composed of complex numbers, that is a limitation of the traditional graph analysis. Furthermore, a difficulty presented in the beginning of Section 4.5, to weight the nodes in the graph modelling process, is easily bypassed when working with the SPG.

To represent the hole spectrum for a specific graph signal, (4.7) can be rewritten considering its version of the Fourier Transform for Graphs as (4.13):

$$\hat{f} = U^T f \tag{4.13}$$

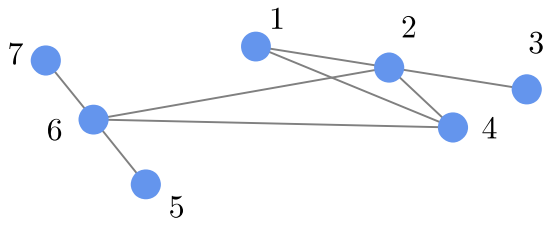
where $U = [u_0\ u_1\ u_2 \ldots\ u_{n-1}]$.

In the transformed vector, $\hat{f} = [\hat{f}(\lambda_0)\ \hat{f}(\lambda_1)\ \hat{f}(\lambda_2)\ \ldots\ \hat{f}(\lambda_{n-1})]^T$, $\hat{f}(\lambda_l) = \langle f, u_l \rangle$ is the GFT coefficient corresponding to the eigenvalue λ_l.

To illustrate the idea of graph signal spectrum, another graph is presented in Figure 12. The definition of a graph signal depends on the graph modelling and the parameters of the system that one are studying. A traditional importance signal commonly used in the literature is the inverse of the cost to reach from an individual node to the reference node [46]. Although this parameter is not related to the scope of this thesis, and the graph signal utilized in this research will be explained in the next chapters, the inverse of the cost to reach a node from a reference node will be used to represent the graph signal (node weights/importance) only for didactic purposes.

Figure 12 presents a sample undirected graph and in order to simplify the problem the graph is unweighted (edges unweighted and the numbers are the labels of the vertices). The value of the importance signal of a specific reference node r will be defined as $f_r = [f_r(1)\ f_r(2)\ f_r(3)\ \ldots\ f_r(R)]^{\mathrm{T}}$, where $f_r(i)$ is the inverse of the sum of weights in the geodesic from node i to node r [46]. The signal will be normalized and the sum of signal values is equal to one. At the reference node, the signal is unitary $f_r(r) = 1$.

Figure 12 – Generic graph with sample graph signals



Source: Author

The importance signals considering “1” as reference node are presented in Table 9. The same procedure is done considering the other nodes as reference. It is possible to observe that as the cost to reach the reference node decreases, the value of the importance signal increases [46].

After calculating the importance signal vector by considering each one of the nodes as the reference node, one can plot the spectrum importance graphic, as shown in Figure 13. It is possible to observe that each subplot depicts the importance spectrum of the graph considering a specific node as reference, which demonstrates that, depending on the values chosen as graph signal, the spectrum result can be completely different. Similar examples to better understand the spectrum can be found in [46, 47, 96].

In Figure 13, the importance spectra of nodes 2 (Figure 13b) and 6 (Figure 13f) have large high frequencies (absolute values) components compared to the other nodes. On

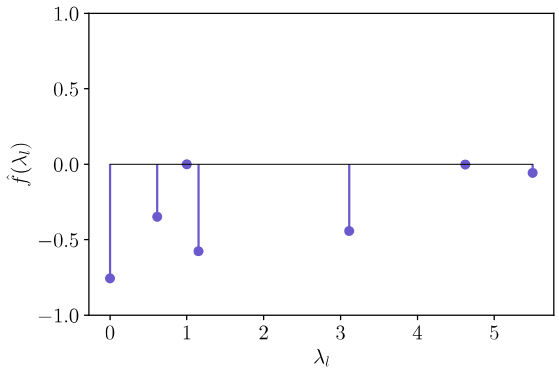
Table 9 – Importance signal values for node 1 of Figure 12 as reference

Node	**Shortest Path**	**Cost**	**Cost^{-1}**	**Important Signal Value**
1	-	0	-	1
2	2 - 1	1	1/1	0.30
3	3 - 2 - 1	2	1/2	0.15
4	4 - 1	1	1/1	0.30
5	5 - 6 - 2 - 1	3	1/3	0.075
6	6 - 4 - 1	2	1/2	0.1
7	7 - 6 - 2 - 1	3	1/3	0.075

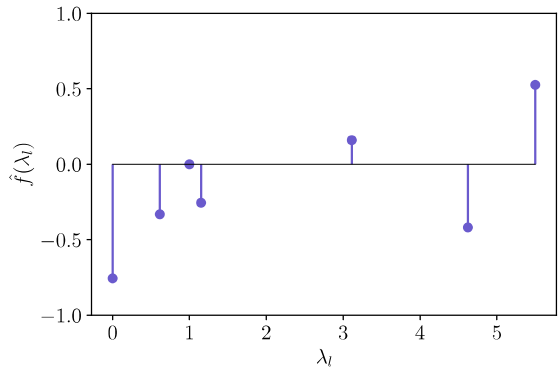
the other hand, nodes 3, 5 and 7 (Figures 13c, 13e and 13g, respectively) presents small absolute values for high frequencies components. As already mentioned, the importance information is encoded in the high frequency components of the importance spectrum. If a node is considered central to a network, its importance signal is non-smooth, that means, variations in the importance signal are high. This statement allows one to measure the importance of a node based on the graph importance spectrum [46].

Figure 13 – Importance Spectra of the graph illustrated in Figure 12

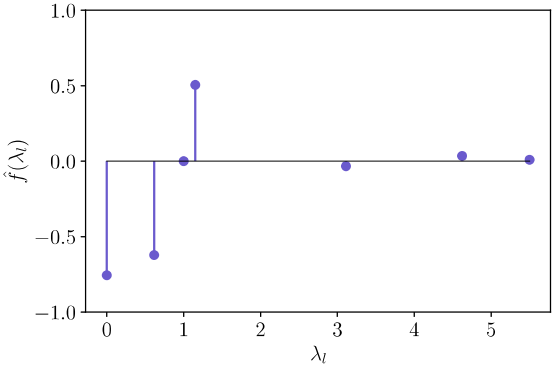
(a) Importance spectrum (node 1 as reference)



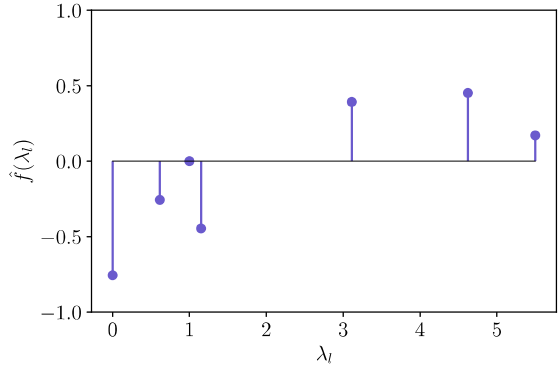
(b) Importance spectrum (node 2 as reference)



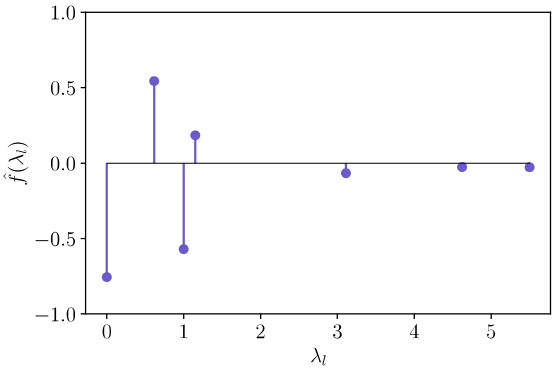
(c) Importance spectrum (node 3 as reference)



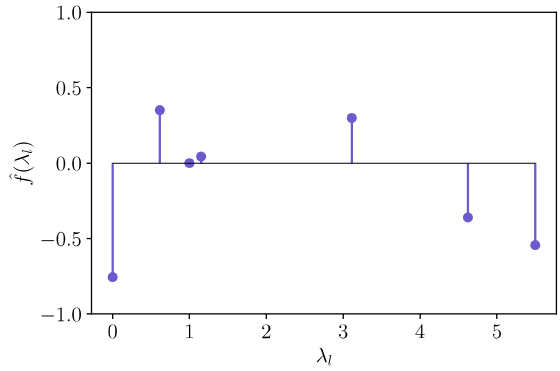
(d) Importance spectrum (node 4 as reference)



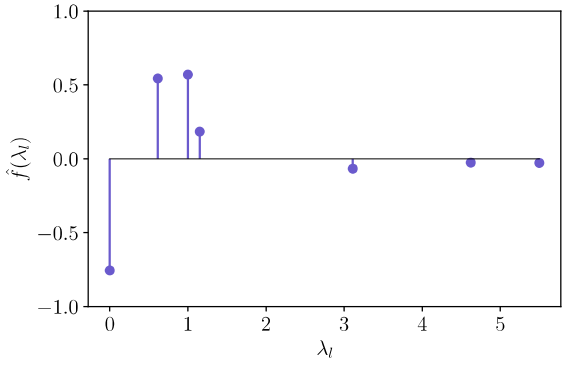
(e) Importance spectrum (node 5 as reference)



(f) Importance spectrum (node 6 as reference)



(g) Importance spectrum (node 7 as reference)



Source: Author

Thus, in a simplified way, the signal f can be represented in the vertex domain, as shown in Figure 11 or it may be represented as $\hat{f}$ in the graph spectral domain, as shown in Figure 12.

Although the spectrum graphic analysis is an interesting way to qualitatively analyse the signal information, sometimes only this information is not enough. From this difficulty, coefficients of importance signal can be used to transform the graphical information into numbers.

There are some mathematical expressions that provide the notion of smoothness with respect to a node of the graph [96]. For a global notion of smoothness, which is the interest of this thesis, *graph quadratic Laplacian form* can be used, as defined in Section 3.2.2. In this case, the importance or notion of smoothness index using quadratic forms, NSI_Q is defined by (4.14)

$$NSI_Q = f^T L_{GFT} f \tag{4.14}$$

where f is the graph signal and L_{GFT} the weighted Laplacian matrix as defined in (4.11).

So, to help the understanding of Figure 13, the frequency graphics were transformed into numbers, using (4.14). The results are presented in Table 10 and reiterates the result discussed above, where the higher values represent the most important nodes.

Table 10 – NSI_Q values for Figure 12

Node	**NSI_Q**
1	1.08
2	2.55
3	0.54
4	1.85
5	0.56
6	2.58
7	0.57

Although very used in literature, the *graph quadratic Laplacian form* presented by (4.14) are not the only possible mathematical expression to measure the global importance of a graph signal. Other mathematical equations can be used to define the graph signal importance and summarize the frequency graphic data [46, 47].

5 METHODOLOGY

In this chapter the methodology used to obtain the load current variability in a feeder dominated by DG is shown. The complete methodology was developed based on a trivial two buses feeder and all the mathematics presented use literal equations to generalize the method. The load flow using algebraic forms is developed and the complete procedure for the graph modelling is also presented, using SPG to generate a mathematical expression that provides the notion of smoothness for each analysed case.

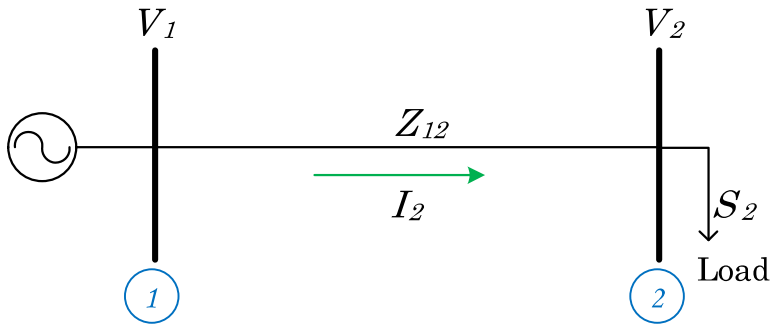
All the studied matrices of this chapter and the following ones are symmetric (the graphs modelled are undirected). This thesis is focused in the analysis of the load current variation values in the substation bus, aiming to obtain a global current variation and based in the fact that all electrical substations have a protection and/or measuring device in this bus. All the distributed feeders studied in this thesis are balanced.

The computer used for the analysis is configured with an Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz processor and 16 GB RAM running on Windows 10 operating system.

5.1 Distribution Feeder and Power Flow Calculation

Figure 14 presents a trivial two buses feeder used to illustrate the proposed methodology, where $V_1 = 1\underline{/0^\circ}$ pu, $I_2 = I_{12}$, $Z_{12} = c + jd$ and $S_2 = P_2 + jQ_2$.

Figure 14 – Trivial two buses feeder



Source: Author

Firstly, it is essential to obtain the values of the electrical quantities of the feeder. To obtain the current values for each bus, the power flow was calculated using the backward/forward sweep method. For larger feeders, the power flow was calculated through

simulations performed in MATLAB/Simulink. This methodology step is only needed during the development of the method. Once validated, it is possible to identify the maximum and minimum current variations for feeders dominated by DG without the need to calculate the power flow.

For the sweep method, first the V_2 voltage (${V_2}^0$)† was estimated to be 1 pu and I_2 was calculated for the first iteration (${I_2}^1$), as showed by (5.1). After that, it is possible to calculate a new value for V_2 (${V_2}^1$), as (5.2). As mentioned in chapter 3, the power flow methods are generally iterative methods, so the procedure most be repeated, as shown below, until the expected error is obtained.

$${I_2}^1 = \left(\frac{S_2}{{V_2}^0}\right)^* = \frac{{S_2}^*}{1} \tag{5.1}$$

$${V_2}^1 = V_1 - Z_{12} \cdot {I_2}^1 = 1 - Z_{12} \cdot {S_2}^* \tag{5.2}$$

$${I_2}^2 = \left(\frac{S_2}{{V_2}^1}\right)^* = \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*} \tag{5.3}$$

$${V_2}^2 = V_1 - Z_{12} \cdot {I_2}^2 = 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*} \tag{5.4}$$

$${I_2}^3 = \left(\frac{S_2}{{V_2}^2}\right)^* = \frac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}} \tag{5.5}$$

$${V_2}^3 = V_1 - Z_{12} \cdot {I_2}^3 = 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}} \tag{5.6}$$

$${I_2}^4 = \left(\frac{S_2}{{V_2}^3}\right)^* = \frac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}}} \tag{5.7}$$

$${V_2}^4 = V_1 - Z_{12} \cdot {I_2}^4 = 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot \dfrac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}}} \tag{5.8}$$

†In the power flow calculation, the superscript numbers indicate the iteration number and not an exponential.

$$I_2{}^5 = \left(\frac{S_2}{V_2{}^4}\right)^* = \frac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*}}}}} \tag{5.9}$$

$$V_2{}^5 = V_1 - Z_{12} \cdot I_2{}^5 = 1 - Z_{12} \cdot \frac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*}}}}} \tag{5.10}$$

This procedure must continue, until the last iteration presents values for I_2 and V_2 according to the specified error. As this is a literal calculation and no value of error is known for this equation, it is necessary to do infinite iterations to obtain the values of I_2 and V_2.

It is possible to observe that the V_2 equations behave as an infinite continued fraction. So, considering $\varphi = Z_{12} \cdot S_2{}^*$, a substitution is made in the voltage equations presented above to facilitate the identification of the term that is repeated. From (5.11) one can see that the term that eventually repeat from some point forwards is $1 - \frac{\varphi}{1-\varphi}$. Thus, the voltage value V_2 in the infinite iteration can be found as (5.12). The calculation is shown in the Annex.

$$\begin{aligned}
V_2{}^1 &= 1 - Z_{12} \cdot S_2{}^* = 1 - \varphi \\
V_2{}^2 &= 1 - Z_{12} \cdot \frac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*} = 1 - \frac{\varphi}{1 - \varphi} \\
V_2{}^3 &= 1 - Z_{12} \cdot \frac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*}} = 1 - \frac{\varphi}{1 - \dfrac{\varphi}{1 - \varphi}} \\
V_2{}^4 &= 1 - Z_{12} \cdot \frac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*}}} = 1 - \frac{\varphi}{1 - \dfrac{\varphi}{1 - \dfrac{\varphi}{1 - \varphi}}} \\
V_2{}^5 &= 1 - Z_{12} \cdot \frac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot \dfrac{S_2{}^*}{1 - Z_{12} \cdot S_2{}^*}}}} = 1 - \frac{\varphi}{1 - \dfrac{\varphi}{1 - \dfrac{\varphi}{1 - \dfrac{\varphi}{1 - \varphi}}}}
\end{aligned} \tag{5.11}$$

$$V_2 = \frac{1 \pm \sqrt{1 - 4 \cdot Z_{12} \cdot S_2{}^*}}{2} \tag{5.12}$$

As the current and voltage are related, the current I_2 value in the infinite iteration is represented by (5.13)

$$I_2 = \left(\frac{S_2}{V_2}\right)^* = \frac{2 \cdot {S_2}^*}{1 \pm \sqrt{1 - 4 \cdot Z_{12} \cdot {S_2}^*}} \tag{5.13}$$

In this thesis, the DG penetration level is defined as the ratio between the sum of the installed PV system power generation and the total power of the feeder. In the feeder of Figure 14, there is only one load bus. As one of the main advantages of DG is to generate electric energy near to the consumer, the PVDG was installed on this bus. The PVDG is represented as current sources [12]. For this circuit, it is considered two cases of analysis, the reference case, where there is no distributed generator connected to the feeder and a scenario where the DG is connected at bus 2.

As the scope of this thesis is to evaluate the load current variation in a feeder with the insertion of DG units, the analysis was focused on the load current in steady-state for the substation bus for each one of the scenarios. It is important to say that the electric current was chosen to be the variable of interest in this thesis because its alteration can cause difficulties and malfunctions/maloperation in the protection devices. Despite this, it is known that voltage and real and reactive power can also be changed in networks with high penetration of DG. So, they are also quantities that deserve attention and study as they can cause problems for electrical systems. The choice of the substation bus as a reference for the analysis of the load current variability is due to the fact that this bus concentrates all the current variations of the feeder and it is a topological location that will have a protection device in almost all situations of real cases.

As already mentioned, it is know that the power S delivered in a bus is represented by (3.10) and (3.11). In this thesis, for each distributed generator only two operational conditions were considered: on, generating 100% of the real power of the load or off [28, 89]. So, when there is DG connected in a bus, $S = -Q_d j$ ($S^* = Q_d j$) and when there is no DG, $S = -P_d - Q_d j$ ($S^* = -P_d + Q_d j$). So, for the 2 cases presented, the values of V and I are depicted in Tables 11 and 12.

Table 11 – Power Flow values for Bus 1 of Figure 14

Cases	**Bus 1**		
	V_1	I_1	S_1
1 (Reference case)	$1\angle 0°$	$\frac{2\cdot(-P_{2d}+Q_{2d}j)}{1+\sqrt{1-4\cdot Z_{12}\cdot(-P_{2d}+Q_{2d}j)}}$	0
2 (DG connected at bus 2)	$1\angle 0°$	$\frac{2\cdot(Q_{2d}j)}{1+\sqrt{1-4\cdot Z_{12}\cdot(Q_{2d}j)}}$	0

Source: Author

Table 12 – Power Flow values for Bus 2 of Figure 14

Cases	**Bus 2**		
	V_2	I_2	S_2
1 (Reference case)	$\frac{1+\sqrt{1-4\cdot Z_{12}\cdot(-P_{2d}+Q_{2d}j)}}{2}$	$\frac{2\cdot(-P_{2d}+Q_{2d}j)}{1+\sqrt{1-4\cdot Z_{12}\cdot(-P_{2d}+Q_{2d}j)}}$	$-P_{2d}-Q_{2d}j$
2 (DG connected at bus 2)	$\frac{1+\sqrt{1-4\cdot Z_{12}\cdot(Q_{2d}j)}}{2}$	$\frac{2\cdot(Q_{2d}j)}{1+\sqrt{1-4\cdot Z_{12}\cdot(Q_{2d}j)}}$	$-Q_{2d}j$

Source: Author

5.2 Graph Modelling and SPG

Beyond the electrical analysis, an electrical system can be represented by a graph where the buses are represented as the vertices and the distribution lines are the edges, as shown in Figure 15.

Figure 15 – Two Buses (Figure 14) graph representation

Source: Author

The graph is modelled according to the system characteristic. For some applications, it will be indispensable to work with a weighted graph - the vertex and/or the edge are weighted. For others, the weight can be neglected [88]. The weights can be defined according to several criteria, such as a length or admittance. It is clear that the graph model depends on the system, the application and the parameters that will be studied.

As the PVDG connection creates a bidirectional behaviour on the grid, in the proposed methodology the feeders should be modelled as an undirected graph. The objective of this methodology is to simplify and facilitate the analysis of an electric network. Thus, all edges were considered with the same weight, equal to one. Typically, the distribution power feeders are radial systems, so it will be modelled as a tree.

The modelling of the research problem predicts the need to weight the nodes, indicating values/weights for the feeder buses. The use of SPG allows weights to be assigned to graph nodes, which is a very important characteristic of the proposed methodology. The load flow algebraic formulation in Quadratic Forms presented in Section 3.3, which culminated in the development of the c index (3.68) showed in Section 3.4, assigns to each node a weight that is associated with the power flow.

Thus, c is a parameter that holds important information about the power flow in a bus. A c value can be found for each one of the feeder buses, and it assigns a value for the bus according to the power flow characteristics, that is, according to the electrical characteristics of the entire network. Considering that z_{i} is the power injection and can be P or Q, it is possible to define two parameters of interest: c_P and c_Q which represents the c value related to the real power and reactive power, respectively.

For the nodes with loads (in the example case only bus 2), a c_P can be obtained, based on the power injection. In this way, according to the topological characteristics of the feeder, loads, impedances and power flow, a value is assigned to each node. The c index was calculated for each node generating a graph signal, as mentioned in Section 3.4. If the node has no load, c value is 0.

The PVDG model used in this thesis considers that the generators inject only real power on the grid [12]. When the generator is connected on a bus, the real power on the node is equal to 0, $P_{node} = P_{PVDG} - P_{load}$, as the power of the generator is considered equal to the power of the load [28]. Thus, there will be no change in the value of c_Q of the buses, as the reactive power in each bus does not change. For that reason, only c_P values will be considered.

The procedure explained in Section 3.4 is described below, step-by-step to model the example system (Figure 14).

As already mentioned, the literal values of electrical quantities of Figure 14 are $V_1 = 1\underline{/0^\circ}$ pu, $I_2 = I_{21}$ and $S_2 = P_2 + jQ_2$. The admittance is $Y_{12} = a + bj$ ($Y_{12} = \frac{1}{Z_{12}}$). The admittance matrix Y for this circuit is depicted by (5.14)

$$Y_{bus} = \begin{bmatrix} a+bj & -a-bj \\ -a-bj & a+bj \end{bmatrix} \tag{5.14}$$

The J_{P2}, as shown in Section 3.4 can be represented by (5.15). As Bus 1 has no load, according to (3.68) $c_P = 0$. So, there is no need to calculate J_{P1}.

$$J_{P2} = \begin{bmatrix} 0 & -\frac{a}{2} & -\frac{b}{2} \\ -\frac{a}{2} & a & 0 \\ -\frac{b}{2} & 0 & a \end{bmatrix} \tag{5.15}$$

The eigenvalues of J_{P2} must be calculated. For this, it can be used a traditional calculation or using the equation presented in [75]. According to [75], no matter the size of the feeder/circuit, the J_P matrix will always have 4 eigenvalues. For this example, the 4 eigenvalues are:

$$\lambda_{p1} = \frac{a + \sqrt{2 \cdot a^2 + b^2}}{2} \tag{5.16}$$

$$\lambda_{p2} = \frac{a - \sqrt{2 \cdot a^2 + b^2}}{2} \tag{5.17}$$

$$\lambda_{pa1} = a \tag{5.18}$$

$$\lambda_{pa2} = 0 \tag{5.19}$$

The c_P index is calculated for each one of the nodes in all of the simulated cases and it will be used as the graph signal for that node in that case. So, to calculate the c_P value (3.68) only $\lambda < 0$ values are considered as the S values are always negative, according to Tables 11 and 12. The c_P value is:

$$c_P = \frac{P_2}{\frac{a - \sqrt{2 \cdot a^2 + b^2}}{2}} \tag{5.20}$$

As two different scenarios were analysed, two graph signals were generated, one for each case, defined as $f_\Upsilon = [f_\Upsilon(1)\ f_\Upsilon(2)]^T$, where Υ represent the case in question and "(1)" and "(2)" represents buses 1 and 2 respectively. For each case, the graph signals can be described as:

$$f_1 = [f_1(1)\ f_1(2)]^T = [c_{P1}(1)\ c_{P2}(2)]^T = \left[0\ \frac{P_2}{\frac{a-\sqrt{2\cdot a^2+b^2}}{2}}\right]^T \tag{5.21}$$

$$f_2 = [c_{P2}(1)\ c_{P2}(2)]^T = [0\ 0]^T \tag{5.22}$$

The c_P value for bus 1 are equal to zero because there is no load on it. For f_2, the $c_{P2}(2)$ are equal to zero because the distributed generation power are equal to the load power (3.10).

The graph signal for the two scenarios is then treated mathematically by the GFT, as shown in Section 4.5.1, to extract a number that represents all the graph nodes spectrum frequency into a number. In this way, two different values are generated, given the response of the system to the different electrical characteristics imposed by the variation of DG combinations in the feeder.

According to Section 4.2, considering the graph of Figure 15, the degree (D), adjacency (A) and Laplacian (L) matrices are:

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.23}$$

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{5.24}$$

$$L = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{5.25}$$

The eigenvalues λ of matrix L are $\lambda_1 = 0$ and $\lambda_2 = 2$. The eigenvector associated to λ_1 is $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and the eigenvector associated to λ_2 is $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. Then, the $\hat{f}_\Upsilon$ values can be calculated using (4.13) as shown in (5.26) and (5.27).

$$\hat{f}_1 = \left[-\frac{\sqrt{2}}{2} \cdot \frac{P_2}{\frac{a-\sqrt{2 \cdot a^2+b^2}}{2}} \qquad \frac{\sqrt{2}}{2} \cdot \frac{P_2}{\frac{a-\sqrt{2 \cdot a^2+b^2}}{2}} \right] \tag{5.26}$$

$$\hat{f}_2 = [0\ 0] \tag{5.27}$$

As mentioned before, there are different mathematical formulations to measure the importance of a graph signal and it needs to be chosen according on the graph model and data under analysis. That is, it is not mandatory to use the graph quadratic Laplacian form (NSI_Q), and other formulations can be used and developed aiming to fit the problem in analysis. In this thesis, the proposed equation was developed from other equations already used and consolidated in the literature [46, 47]. In this thesis, the expression has been modified empirically through optimization methods to meet the specifications of the research problem. The importance coefficient of the graph signal from the GFT (C_{IGFT}) is depicted by (5.28), and it was tested and validated in the power feeders shown in chapter 6.

$$C_{IGFT} = \sum_{l=0}^{N-1} 0.6^{19\lambda_l} \cdot |\hat{f}(\lambda_l)| \tag{5.28}$$

where N is the number of buses, λ_l are the eigenvalues of L_{GFT}, as defined in (4.11) and $\hat{f}(\lambda_l)$ is the importance of the signal calculated as (4.13). This equation was applied in all studied feeders, as shown in chapter 6.

The weight of all edges of the graph were considered equal to one. The C_{IGFT} values for cases 1 and 2 are calculated in (5.29) and (5.30), where it is possible to observe that for the case with DG connected at bus 2, the C_{IGFT} reduces to zero.

$$C_{IGFT(1)} = 0.6^{19\lambda_1} \cdot |\hat{f}_1(\lambda_1)| + 0.6^{19\lambda_2} \cdot |\hat{f}_1(\lambda_2)| = -\frac{\sqrt{2}}{2} \cdot \frac{P_2}{\frac{a-\sqrt{2 \cdot a^2+b^2}}{2}} \tag{5.29}$$

$$C_{IGFT(2)} = 0.6^{19\lambda_1} \cdot |\hat{f}_2(\lambda_1)| + 0.6^{19\lambda_2} \cdot |\hat{f}_2(\lambda_2)| = 0 \tag{5.30}$$

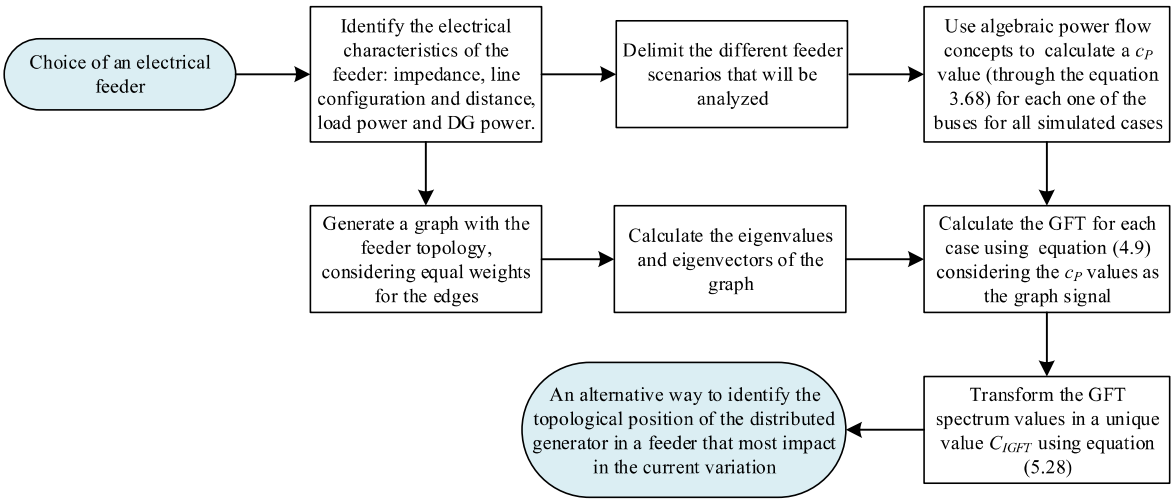
To fit the C_{IGFT}, the results of (5.29) and (5.30) are related with the current values obtained by power flow, Tables 11 and 12. It is necessary to calculate the variation of the load current (Δ) in the substation bus (Bus 1) as shown in (5.31), where *Reference case value* is the current value for the case where there is no DG connected to the feeder. Thus, the developing of (5.28) was done by comparing the C_{IGFT} with the load current variation, calculated by the backward/forward sweep method. Once the C_{IGFT} equation has been validated, it is not necessary to recalculate the power flow values to find the cases of greater and smaller current variation of the studied feeder.

The C_{IGFT} is a parameter that associates concepts of graphs and load flow algebraic formulation and does not indicates the load current absolute values, but the load current variation in relation to the reference case.

$$\Delta = Scenario\ current\ value - Reference\ case\ value \tag{5.31}$$

The steps of the presented methodology can be seen in Figure 16, indicating how the method can be replicated for other feeders.

Figure 16 – Methodology Flowchart



Source: Author

Thus, the methodology is based in other mathematical branches that, integrated allows a load current variation analysis. Since the electrical feeders can be described as

nodes and edges connected, the graph theory is used to model the grid. The insertion of DG units in an electrical power grid can change the power flow. Thus, to ponder the generators connected along the grid, the c index obtained from the load flow algebraic formulation is used, due to the fact that besides to considering the power at that node, it manages to weight the bus considering the entire power flow of the network. The use of SPG method is due to three points: the SPG allows to ponder the buses instead of the edges, the SPG admits the use of complex numbers to weight the buses and the SPG considers vectors of importance graph signals, that consist in an analysis that rely on the weights of the all buses, making it possible to analyze and compare the different DG combination cases proposed in this thesis.

6 RESULTS AND DISCUSSION

In this chapter it will be presented and discussed the results of the methodology proposed in chapter 5. To validate the method, it is applied to four different distribution feeders.

6.1 13-Node Test Feeder (13NTFM)

The methodology was firstly applied in a variation of the well-known in the literature feeder 13 Node Test Feeder (13NTF), showed first in Figure 10a. The original 13NTF has 13 buses, operates at 4.16 kV and is relatively highly loaded with a regulator at the substation bus, different types of distribution lines, shunt capacitors, a transformer in a lateral branch, has unbalanced loading and single-phase, phase-phase and three-phase branches [89]. The feeder used in this thesis is based on the 13NTF, but it is modified (13NTFM). The differences that distinguish the original 13NTF and the 13NTFM are: the voltage regulator was removed, considering that DG usually improves the voltage profile and the breaker between nodes 671 and 692 is considered always closed [6, 28]. The choice to leave the circuit breaker always closed is to represent a larger and integrated distribution network, not allowing a load to be isolated. Furthermore, the always-closed circuit breaker configure the system as a connected graph.

Moreover, all the feeder branches are considered three-phase. To do this modification, the branches impedances were replaced by typical impedance values that already exist in the 13NTF. For the loads, the total real and reactive power of each node were divided by 3, in order to guarantee a balanced feeder. The 13NTFM single-line is presented in Figure 17 and the information about the lines configuration and load data is depicted in Table 13 and Table 14. More details, specially about the line configuration, can be found at [89].

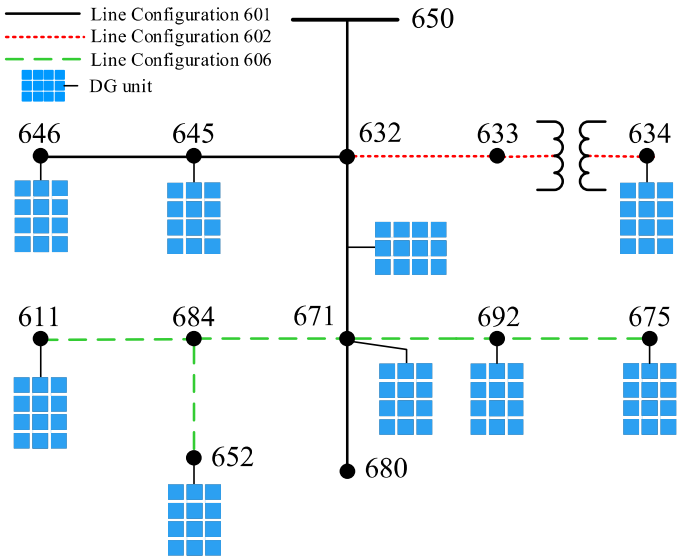
Table 13 – Basic information of Line Configuration Data of the 13NTFM system

Configuration	**Phasing**	**Spacing ID**	**Type of Line**
601	B A C N	500	Overhead Line
602	C A B N	500	Overhead Line
606	A B C N	515	Underground Line

Source: Adapted from [89]

As already mentioned, in this thesis the DG penetration level is defined as the ratio between the sum of the installed PV system power generation and the total power of

Figure 17 – 13NTFM



Source: Adapted from [89]

Table 14 – Load Data of the 13NTFM system

Node	**Type of Load**	**Load Model**	**Real power in each phase (kW)**	**Reactive Power in each phase (kVAr)**
632 - 671	Distributed	Y-PQ	$200/3$	$116/3$
634	Spot	Y-PQ	$400/3$	$290/3$
645	Spot	Y-PQ	$170/3$	$125/3$
646	Spot	D-Z	$230/3$	$132/3$
671	Spot	D-PQ	385	220
611	Spot	Y-I	$170/3$	$80/3$
652	Spot	Y-Z	$128/3$	$86/3$
692	Spot	D-I	$170/3$	$151/3$
675	Spot	Y-PQ	281	154

Source: Adapted from [89]

the feeder. In the 13NTFM there are nine load buses, spotted and distributed. As one of the main advantages of DG is to generate electric energy near to the consumer, the PVDG were installed on the nodes with loads, as shown in Figure 17. Considering all possible combinations for the nine generators, including the reference case, where there is no distributed generator connected to the feeder, 512 scenarios were created representing real feasible different operational cases, as shown in (6.1)

$$1 + \sum_{\Psi=1}^{9} \binom{9}{\Psi} = 512 \tag{6.1}$$

where Ψ is the quantity of PVDG.

The feeder is considered dominated by DG units because different locations are established for the generators, spreading the DG impact across the feeder. In order to reduce the simulation time, but still guaranteeing reliable results, all PVDG were represented as current sources [12]. The analysis was focused on the load current in steady-state for the substation bus (node 650 of Figure 17) for each one of the 512 scenarios.

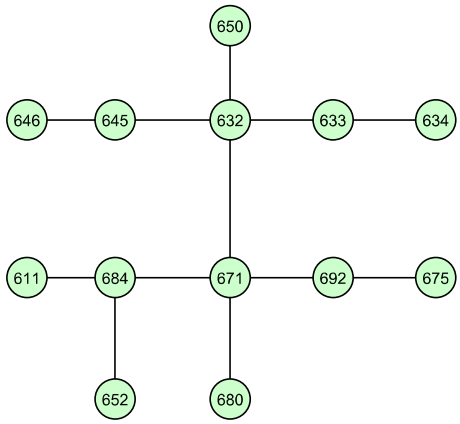
The load flow was performed in MATLAB/Simulink and the load current value in node 650 obtained in each case was subtracted from the current reference case, generating a load current variation (Δ) as (5.31). The Δ values found by (5.31) were resized and normalized, as shown generically by (6.2), where τ depicts a generic value. The treated normalized values for the current variation were used in the analysis. This step in fundamental only in the validation of the method. Once validated, it is not necessary to recalculate the current values to identify the cases with higher and smaller load current variation.

$$\tilde{\tau} = \frac{\tau - min(\tau)}{max(\tau) - min(\tau)} \tag{6.2}$$

So, for each one of the 512 cases depicted by (6.1), the current values were found. In the reference case where no PVDG is connected to the 13NTFM, the current at bus 650 is 966.87 A, which is also the highest value found in all simulations. The minimum value found in the simulations was 856.55 A for the case with all nine DG units were connected. So, it is possible to observe that for the extreme cases, the current may decrease about 11%.

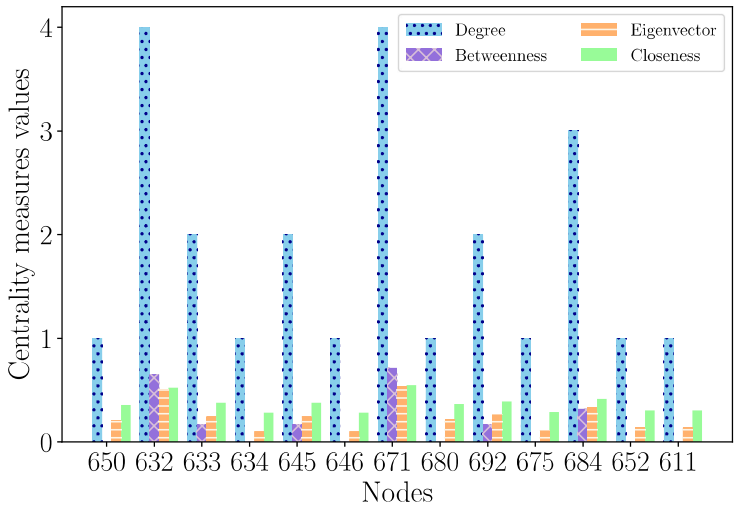
The graph of the 13NTFM is undirected and unweighed and is represented by Figure 18. The Degree, Betweenness, Eigenvector and Closeness Centrality measures were calculated to obtain the most important nodes based on the graph topology, Figure 19. Nodes 632, 671 and 684 are the most central nodes for the centrality measures under study. For some terminal nodes, due to the topology of the graph, the Betweenness Centrality is zero. This result is based only on the topological characteristics of the graph and do not faithfully represents the model studied in this thesis, as each node has unique electrical characteristics.

Figure 18 – 13NTFM graph representation



Source: Author

Figure 19 – Centrality measures of 13NTFM: Degree, Betweenness, Eigenvector and Closeness



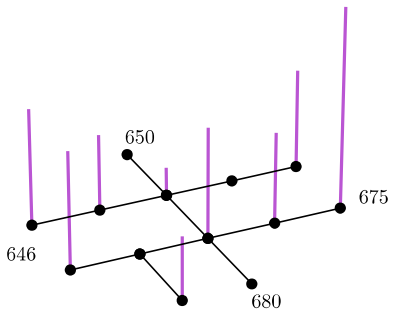
The PVDG model used in this thesis considers that the generators inject real power on the grid and for that reason, only c_P values will be considered [12]. So, the c_P index is calculated for each one of the 13 nodes in all of the simulated cases and it will be used as the graph signal for that node in that case. As 512 different scenarios were analysed, 512 graph signals were generated, one for each case, defined as $f_\Upsilon = [f_\Upsilon(650)\ f_\Upsilon(632)\ f_\Upsilon(633)\ f_\Upsilon(634)\ f_\Upsilon(645)\ f_\Upsilon(646)\ f_\Upsilon(671)\ f_\Upsilon(680)\ f_\Upsilon(684)\ f_\Upsilon(611)\ f_\Upsilon(652)\ f_\Upsilon(692)\ f_\Upsilon(675)]^T$, where Υ represent the case in question.

The absolute values of the generated *graph signal* (c_P) were then obtained. Some of the simulated graph signal cases are shown in Figure 20. The cases shown were chosen randomly. The purple bars depict the absolute value of the c_P calculated for each node in that case. The dotted circles represent the buses where PV generators are connected. As

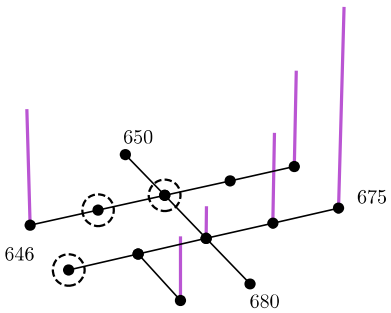
c_p depends on P_{node} value of the node, the nodes that have a value for f different from zero are the nodes where no distributed generator is connected. In other words, if there is a generator in a node, its c_P is zero.

Figure 20 – Graph signal c_P of some cases of 13NTFM (Figure 18)

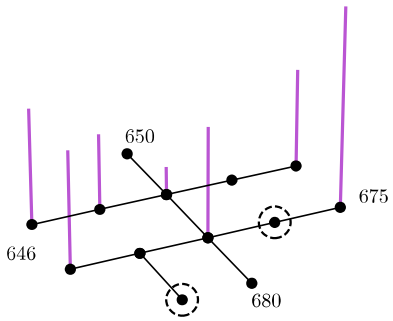
(a) Graph signal of reference case - Case #512



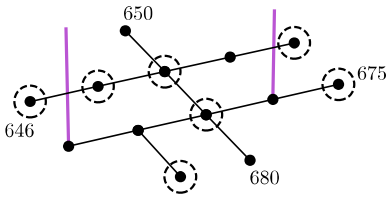
(b) Graph signal of Case #53



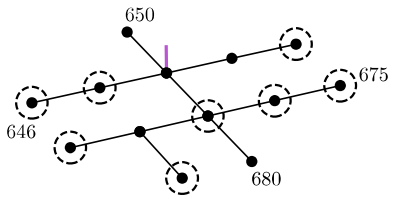
(c) Graph signal of Case #192



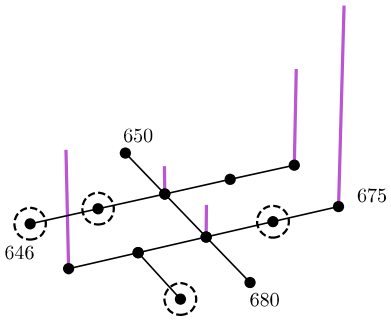
(d) Graph signal of Case #351



(e) Graph signal of Case #510



(f) Graph signal of Case #220



Source: Author

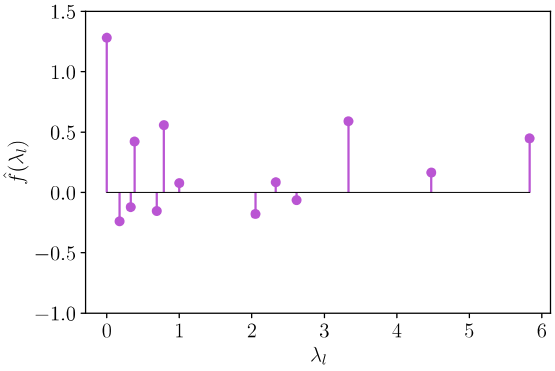
With the values obtained for the graph signal, shown in Figure 20, the GFT can be calculated as depicted by (4.13). Figure 21 shows the spectrum of frequency of 13NTFM. It is possible to observe that, depending on the graph signal, which are different from each one of the cases, the importance spectrum presents a diverse behavior. As mentioned in the

chapter 5, it is possible to transform the spectrum graphic into a number, facilitating the interpretation of the results. This is done by the importance signal, the C_{IGFT} index (5.28), that depicts the importance information concealed in the high frequency components of the importance spectrum presented in Figure 21. In this way, 512 different values of importance signal are generated, given the response of the system to the different electrical characteristics imposed by the variation of DG combinations in the feeder.

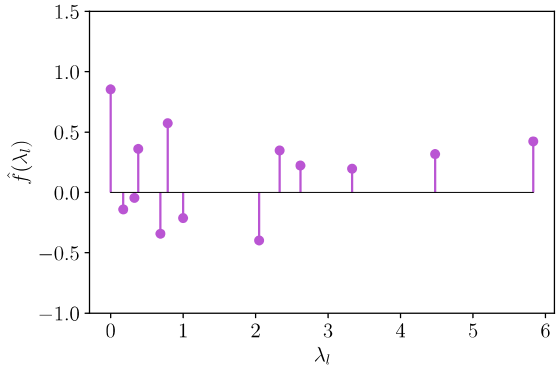
Within the cases shown in the Figure 21, the cases with higher absolute values of $\hat{f}(\lambda_\ell)$, information concealed in the higher frequencies, are the cases #512, #53, #192 and #220 (Figures 21a, 21b, 21c and 21f, respectively). The same cases are the ones that returned higher C_{IGFT} values, as expected. Table 15 depicts the C_{IGFT} values for the cases shown in Figure 21.

Figure 21 – Importance spectrum of 13NTFM

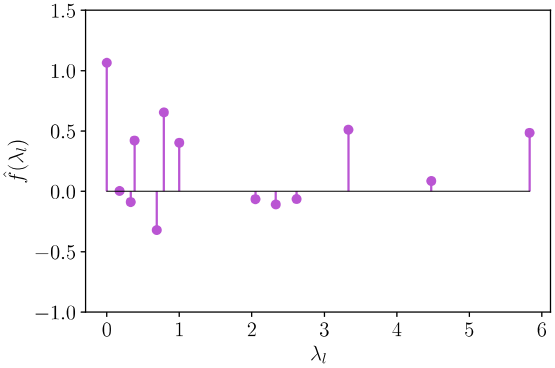
(a) Spectrum of reference case - Case #512



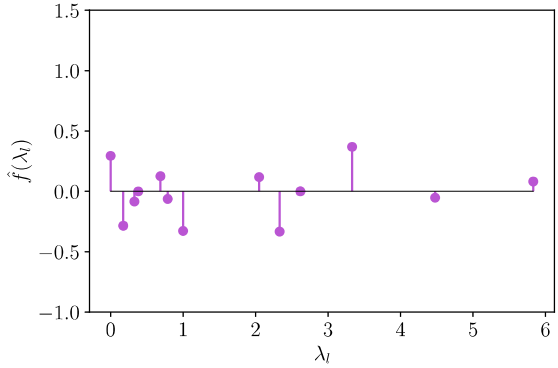
(b) Spectrum of Case #53



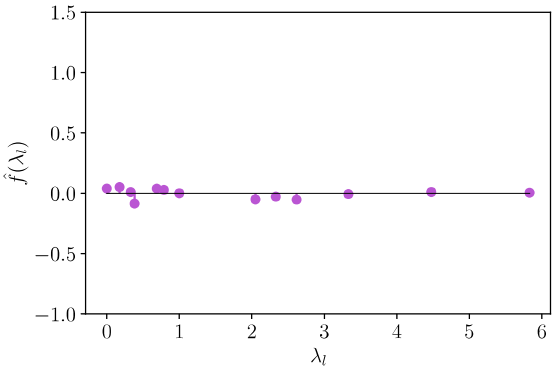
(c) Spectrum of Case #192



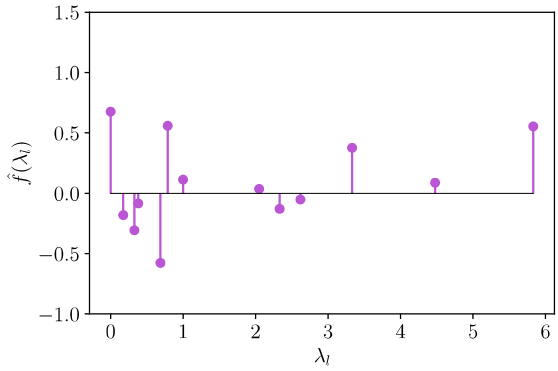
(d) Spectrum of Case #351



(e) Spectrum of Case #510



(f) Spectrum of Case #220



Source: Author

Table 15 – C_{IGFT} for 13NTFM

# Case	C_{IGFT}
512	1.34
53	0.89
192	1.08
351	0.35
510	0.05
220	0.72

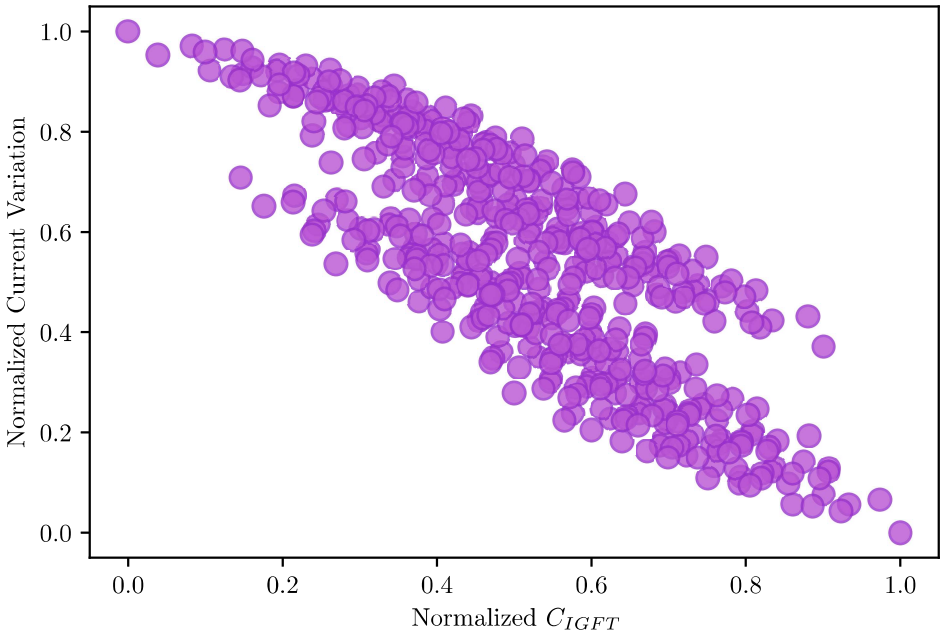
Thus, the GFT extracts information from the frequency contents from signals defined on a graph, which normally are not visible on the spatial and graph domains. Then the advantage of using the GFT is the capability to capture the importance signal globally, gathering characteristics of the graph topology. Associated with this intrinsic characteristic of the method, the values chosen as nodes input signals/weights were also values that gather the characteristics of an electrical network. The use of these two global tools strengthens the method that aims to find a value to represent the entire network topology in each of the simulated cases, in this case reported by C_{IGFT} in (5.28).

To verify the relationship between load current simulation data and the C_{IGFT} values, to validate the methodology, Spearman's and Pearson's rank correlation analyses were made. The results shown a correlation of -0.81 for Pearson's rank correlation and -0.79 for Spearmans's rank correlation, indicating a good relation for the data. Thus, it is possible to identify a relationship between the load current variation calculated by Simulink power flow and C_{IGFT}, which allows to affirm, for this feeder, that small values of C_{IGFT} will result in higher current variation and higher values of C_{IGFT} implies small current variations, which means higher absolute values of electric current.

The results correlating the load current variation at node 650 performed by Simulink and the proposed measure C_{IGFT} calculated are shown in Figure 22. The values were resized and normalized, as (6.2), to a range between 0 and 1. The data present a negative direct correlation and each circle illustrate one of the 512 cases. In Figure 22, the higher values of normalized current variation, y axis, are related to the smaller values of C_{IGFT} with higher current variation and vice versa. The most critical cases in this analysis are the extreme bands, cases where the current variation is minimum or maximum, and may impact more in the energy quality or in the proper functioning of equipment, especially protection ones.

With the support of the Figure 22, it is possible to approximate the value of the current variation knowing the C_{IGFT}, which is relatively simple and computationally inexpensive to calculate. Although computer simulations have evolved considerably in recent decades, the need to simulation hundreds or thousands of different cases are very computationally costly. For this study case, the electrical simulation performed in MATLAB/Simulink took approximately 19 minutes to perform the 512 distinct cases. It is important to say that, for the validation and fitting of the method, it was necessary simulate all the 512 proposed cases. Thus, the developing of this methodology were computationally costly, but this step do not need to be redone for future simulations of the feeders considered in this thesis.

Figure 22 – Normalized relation between current variation values in bus 650 and C_{IGFT} - 13NTFM



Source: Author

The system studied has a reduced number of buses when compared to real electrical distribution systems in the world. As the number of buses of a electrical grid increases, the number of possible combination cases according to the method proposed in this thesis grows, increasing the complexity of the computer simulations and making, in many cases, the power flow simulation using traditional software infeasible according to the methodology proposed in this thesis (DG in all buses with load, varying for all combinations). Just out of curiosity, the IEEE 34-Node Test Feeder (34NTF), another well-known feeder used in the literature composed of 34 buses, would present, for the methodology used here, more than 33 million cases of combinations of DG. This amount of cases makes the simulation in Simulink infeasible.

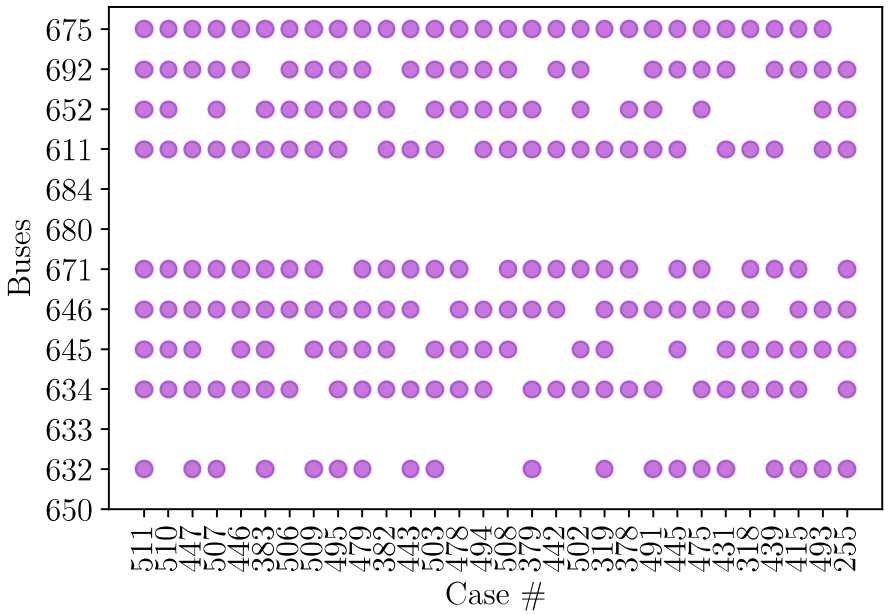
Then, it is feasible to identify which cases presented a certain range of current variation. For example, the maximum and minimum values of Normalized Current Variation presented in Figure 22 are an interesting information for the researchers and electrical utilities, because knowing the DG combinations that causes the smallest and largest current variations allows an analysis of that network topology or even cause the installation of DG to be prohibited or restricted in certain feeder buses.

By focusing on the extreme data, higher and smaller current and C_{IGFT}, it is possible to observe that the greater the current variation (current reduction) the lower the C_{IGFT} found in the Figure 22. The 30 cases with smaller C_{IGFT} are shown in Figure 23, which presents the generators that were connected in the cases of smaller values of C_{IGFT}.

The vertical axis represents the buses and the horizontal axis shows the case number. Thus, for each one of the represented cases, the circles represent a DG connected in the bus shown in vertical axis. The cases, on the x-axis, are represented in ascending order of the values of C_{IGFT}. Once again, largest current variations means small absolute values of current.

For example, the case # 511 is the one where all the 9 DGs are connected and represents the case with smaller C_{IGFT} value and the higher current variation (smaller current absolute value), as expected. It is possible to notice that, for the cases shown in Figure 23, almost all of them present DG connected at buses 671 and 675. These two nodes are nodes with important load values, as Table 14. The node 671 is in a central topological position, as shown in Figure 19, and nodes 671 and 675 presents higher values for c_p, which indicates that they have important influence in the load current variation. From the topological point of view, terminal nodes have a small importance for the graph analysis and that is the reason why node 675 does not present high values for the centrality measures in Figure 19. However, due to its load and power flow influence, it has an extremely relevant influence on the feeder.

Figure 23 – Cases number and the buses with DG connected in the 30 smaller values of C_{IGFT} of 13NTFM



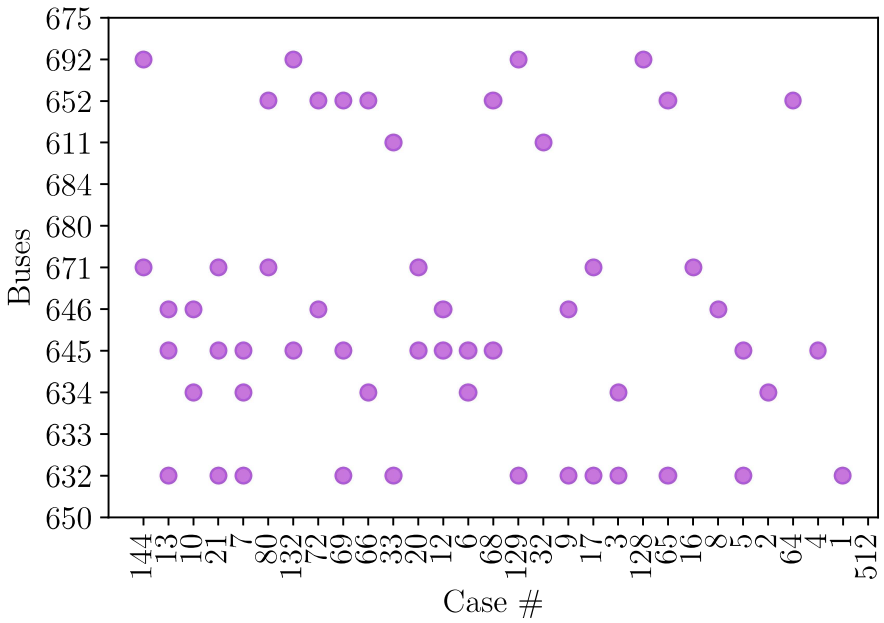
Source: Author

Despite the case of high current variation (# 511) being expected (all DGs connected), the other cases that presented larger current variation values did not presented any regular behavior. That is, it is not possible to automatically infer, without calculations,

which topological DG configuration would have the greatest impact on the load current variation in the substation bus. This reinforces the importance of this methodology.

The same procedure was done for the higher normalized C_{IGFT} values, which resulted in the smaller current variation (higher absolute values) presented by Figure 22. The zoom in the 30 higher values are shown in Figure 24. In this group analysis there is no a specific DG connected in all cases.

Figure 24 – Cases number and the buses with DG connected in the 30 higher values of C_{IGFT} of 13NTFM



Source: Author

When comparing Figures 23 and 24, it is interesting to notice the presence of some specific DG units in the cases with the small C_{IGFT} values. For example, as already mentioned, the DG units in nodes 671 and 675 are present in most of the cases and, on the other hand, no case presented DG in node 675 in Figure 24. Then the presence of DG at nodes 671 and 675 may be an indication that there will be a greater absolute reduction in the current value.

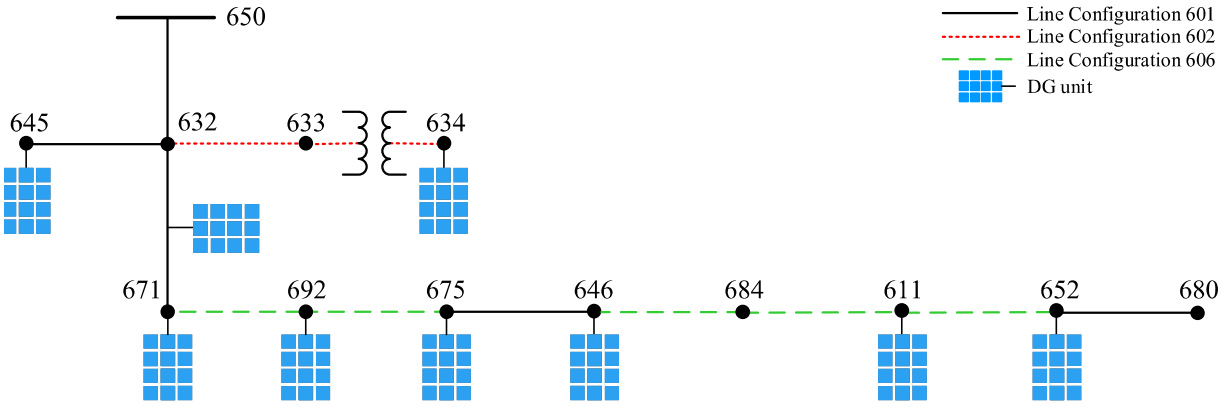
Thus, the method proposed by this thesis was able to identify the cases of greater and lesser load current variation through the C_{IGFT} index. This allows researchers and electric power utilities to use the methodology to filter the cases of greatest interest, that is, cases in which there is a greater variation in the charging current at the substation bus. To validate the method, the methodology will be tested on other feeders in the sections below.

6.2 13-Node Test Feeder Modified A (13NTFMA)

The 13NTFMA has the same characteristics of load powers and lines configuration of 13NTFM but in a different topology, connecting the buses in a distinct sequence. The feeder topology was chosen randomly, aiming to configure a different feeder. The system characteristics can be found in Tables 13 and 14 and all the branches are three-phase.

To evaluate the methodology in other circuit, a feeder originated from the 13NTFM were created and named 13NTFMA, as shown by Figure 25.

Figure 25 – 13NTFMA



Source: Author

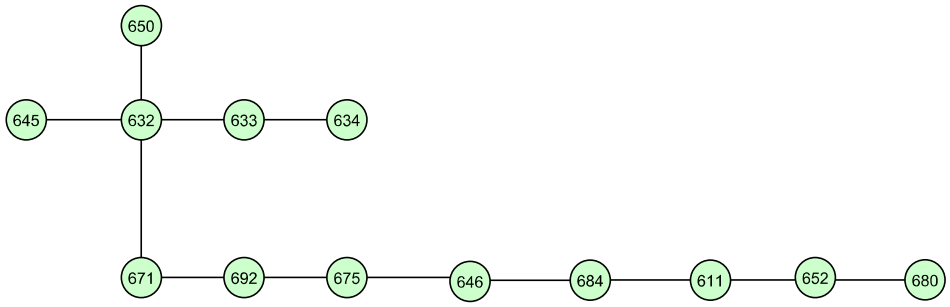
The same methodology utilized in the 13NTFM was applied for the 13NTFMA: 512 scenarios were studied, according to (6.1), only two operational condition were considered for the DGs, on and off, and the current variation were analysed in the substation bus (node 650). The Δ in bus 650 for each scenario were calculates as (5.31) and (6.2), using the values of the Simulink data. For the reference case, the current at bus 650 is 917.38 A and in the case with the higher current variation, the current decay 42.41%. This values were obtained from Simulink simulations.

The undirected graph for the 13NTFMA was created, as Figure 26, and the edges weights are all equal to one.

The centrality measures were calculated to indicate the most important nodes in the graph analysis. The results are shown in Figure 27 and it is possible to observe that for this topology, the most important nodes vary according to the chosen centrality measure. For the Degree, Betweenness ans Eigenvector Centrality, node 632 are the most important. For the Closeness, 671 are the most important. Once again, it is important to note that this results are based on topological characteristics of the graph and may not represent

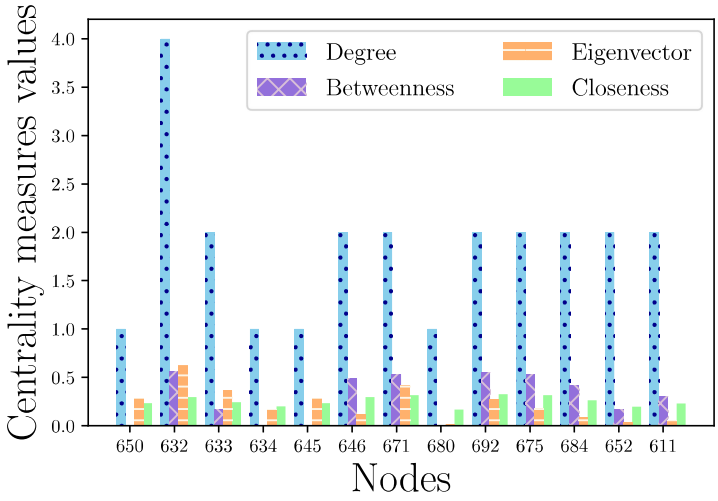
the methodology of this thesis. Figure 27 is used to compare the results obtained by this methodology and the traditional centrality measures used in the graph analysis.

Figure 26 – 13NTFMA graph representation



Source: Author

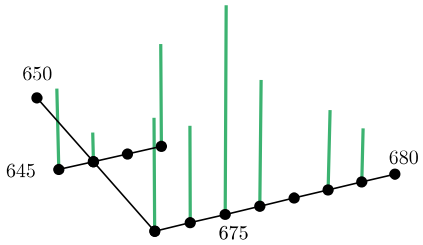
Figure 27 – Centrality measures of 13NTFMA: Degree, Betweenness, Eigenvector and Closeness



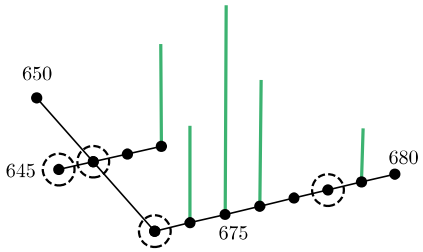
The node weights were obtained by the c_P index developed in Section 3.4. To illustrate the graph signals obtained, some of the 512 cases were shown in Figure 28. The dotted circles represents a DG unit connected in that bus, generating 100% of the bus real load and the green bars depict the absolute values of c_P.

Figure 28 – Graph signal c_P of some cases of 13NTFMA (Figure 26)

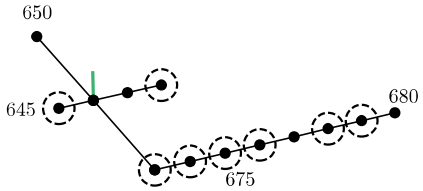
(a) Graph signal of reference case - Case #512



(b) Graph signal of Case #53



(c) Graph signal of Case #510

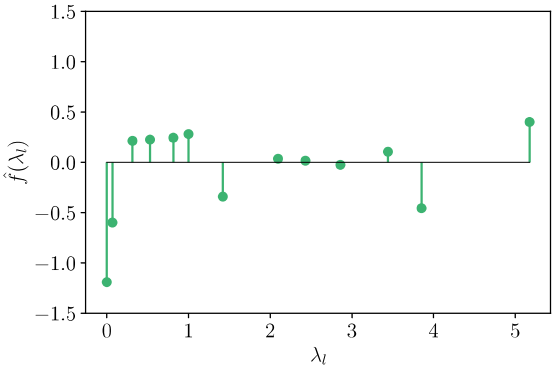


Source: Author

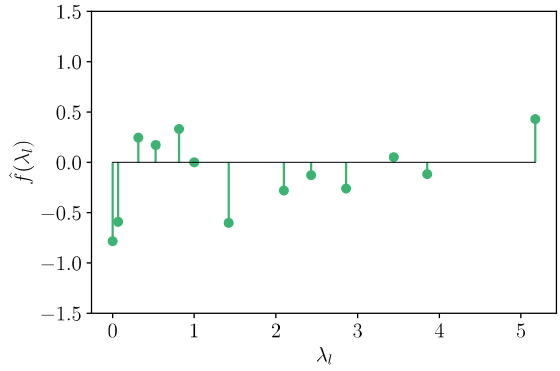
The graph signals of Figure 28 were used in the GFT, generating the spectrum of frequency for this feeder, as shown by Figure 29. The spectrum frequency graphic vary according to the graph signal, as showed above. Each case has a different graph signal that generates a distinct spectrum. In Figure 29, the importance information is concealed in the higher frequencies components. To facilitate understanding of the spectrum, the importance coefficient of the graph signal C_{IGFT} is calculated by (5.28), translating the importance spectrum graphic into a number as show by Table 16. The results confirm that the importance information are intrinsic to the highest frequencies.

Figure 29 – Importance spectrum 13NTFMA

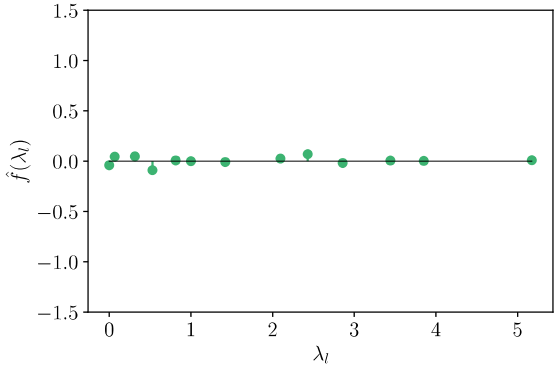
(a) Spectrum of reference case - Case #512



(b) Spectrum of Case #53



(c) Spectrum of Case #510



Source: Author

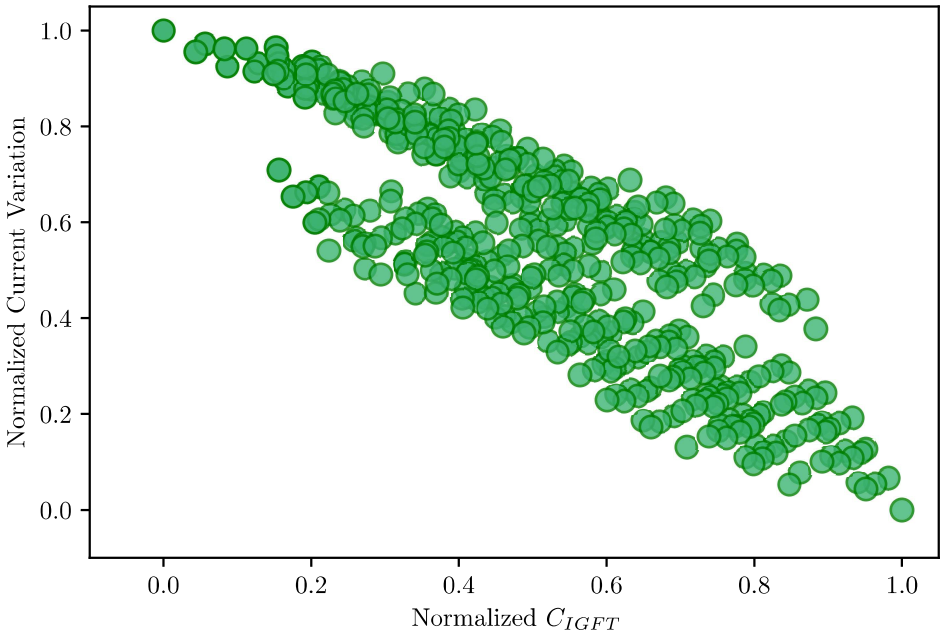
Table 16 – C_{IGFT} for 13NTFMA

# Case	C_{IGFT}
512	1.51
53	1.17
510	0.07

The values of the current variation at node 650 and the C_{IGFT} values were related using Pearson's and Spearman's rank correlation, resulting in correlation values of and -0.82 and -0.79 respectively, indicating a good data relation and a negative direct correlation. As the values are negative, smaller values of C_{IGFT} results in higher current variation and vice versa.

This results can be verified in Figure 30, which shows the values resized and normalized. The higher values of normalized current variation, what means smaller absolute current values, are related to the smaller values of C_{IGFT}. The opposite is also true.

Figure 30 – Normalized relation between current variation values in bus 650 and C_{IGFT} - 13NTFMA

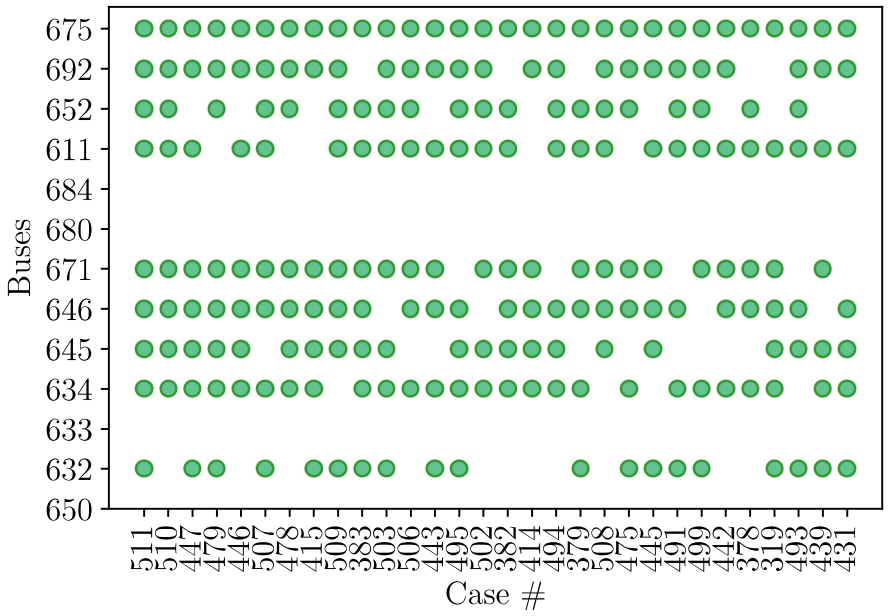


Source: Author

The higher and smaller values of C_{IGFT}, and consequently current variation, are the most important as they indicates the extreme cases. So, it is possible to identify the scenarios with higher and smaller current variation by knowing the C_{IGFT}. Figure 31 shows the 30 smaller values of C_{IGFT}, which represents the cases with higher current variation at substation bus identifying on which buses there were generators connected. The case number is shown in the abscissa axis. Figure 32 shows the 30 higher values of C_{IGFT} and in the same way, shows which DGs were connected.

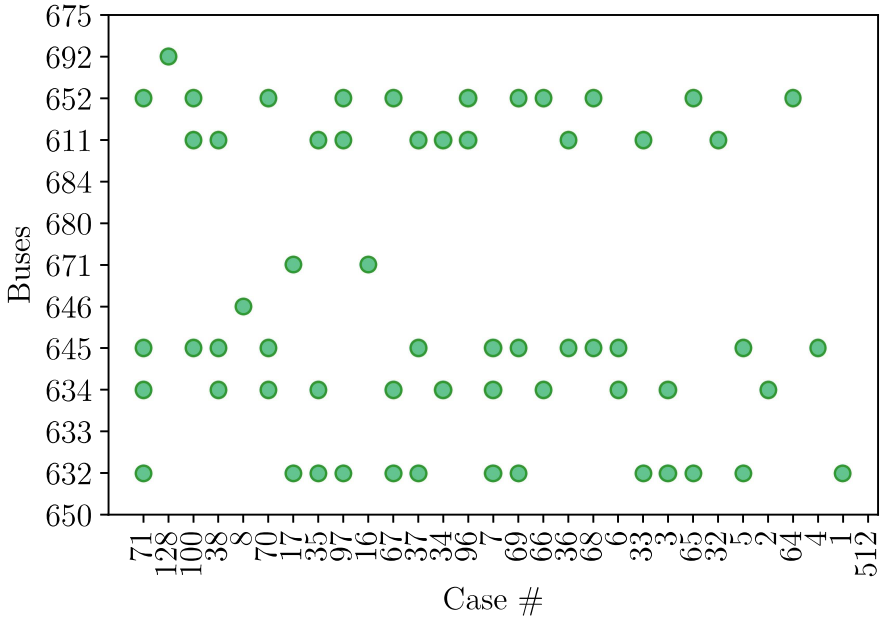
Figure 31 depicts the 30 smaller values of C_{IGFT} in ascending order in the abscissas axis. Each green circle represent that there is a DG unit connected in that bus, showed in the ordinate axis. The same analysis was done considering the 30 cases with current variation, shown in Figure 32. Comparing the data obtained in this two figures and the results of Figure 27 it is possible to observe that, the buses identified as the most important by the centrality measures, 632 and 671 appear in most cases of Figure 31, but not in all of then. The node 675, moreover, appears in all of the 30 smaller values of C_{IGFT}, what makes it the most important node.

Figure 31 – Cases number and the buses with DG connected in the 30 smaller values of C_{IGFT} of 13NTFMA



Source: Author

Figure 32 – Cases number and the buses with DG connected in the 30 higher values of C_{IGFT} of 13NTFMA



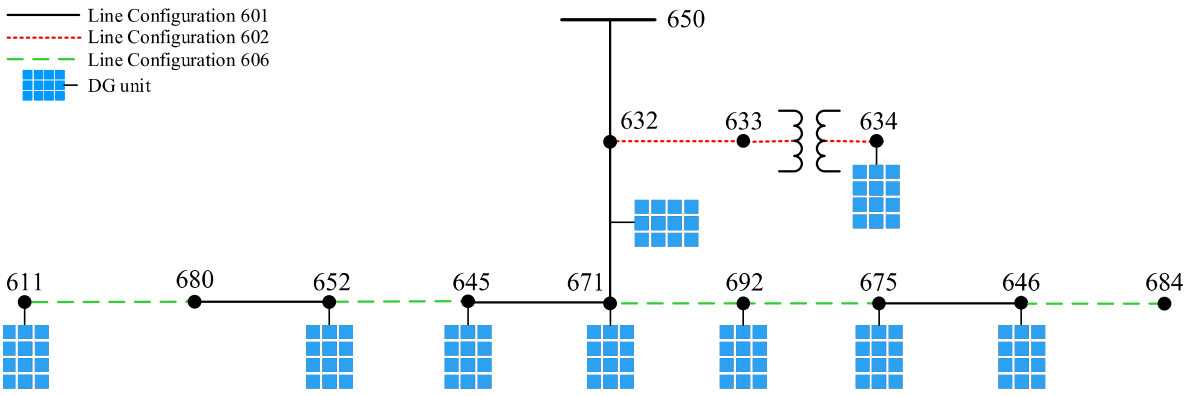
Source: Author

6.3 13-Node Test Feeder Modified B (13NTFMB)

The 13NTFMB is another modified version of the 13NTFM, shown in Figure 17. The modifications done in the 13NTFM was randomly chosen, for the unique purpose of generating a topologically different feeder. As the feeder 13NTFMA, the 13NTFMB has the same characteristics of load powers and lines configuration, but arranged in distinct topological positions. So, the system characteristics can be found in Tables 13 and 14.

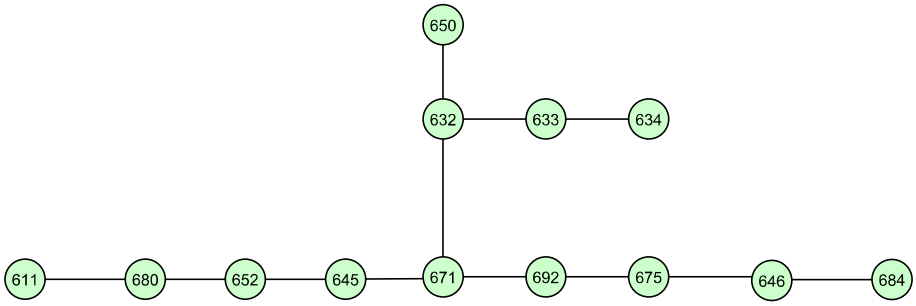
As the exactly same procedure done in Sections 6.1 and 6.2 was done for the 13NTFMB, the results will be presented more succinctly in this section. Figure 33 depicts the one-line diagram for the 13NTFMB and Figure 34 shows the graph representation of the feeder. For this topology, the results of the Simulink load flow simulation presented a current value of 912.7 A in the substation bus for the reference case, where there is no DG unit connected presented. The case with the higher current variation, which means the smaller absolute value for the current at bus 650 presented a current variation of 42.6%.

Figure 33 – 13NTFMB



Source: Author

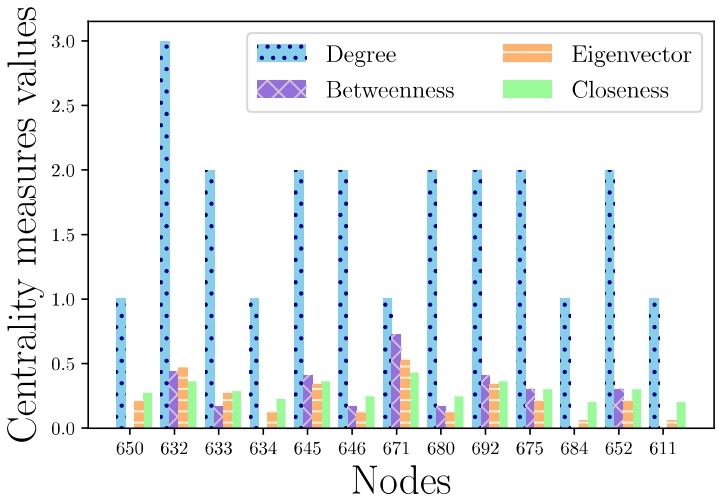
Figure 34 – 13NTFMB graph representation



Source: Author

The Degree, Betweenness, Eigenvector and Closeness centrality measures were again calculated aiming to compare the results of the graph topological analysis and the results of the methodology proposed by this thesis. The most important node considering Degree centrality is 632. For the Betweenness, Eigenvector and Closeness, node 671 is the most important, as can be seen in Figure 35

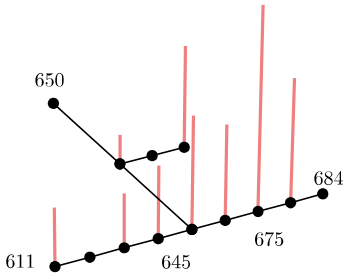
Figure 35 – Centrality measures of 13NTFMB: Degree, Betweenness, Eigenvector and Closeness



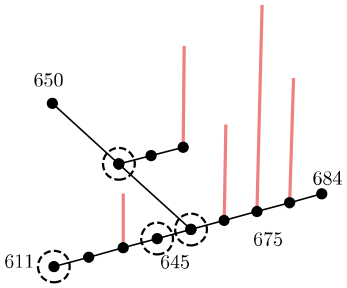
The node weights for the graph were obtained by the c_P index developed in Section 3.4. The graph signals obtained for some of the 512 cases were shown in Figure 36. The dotted circles represents a DG unit connected in that bus, generating 100% of the bus real load.

Figure 36 – Graph signal c_P of some cases of 13NTFMB (Figure 34)

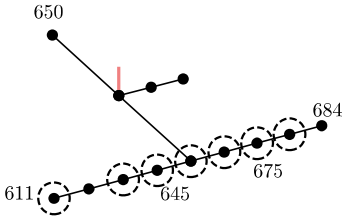
(a) Graph signal of reference case - Case #512



(b) Graph signal of Case #53



(c) Graph signal of Case #510

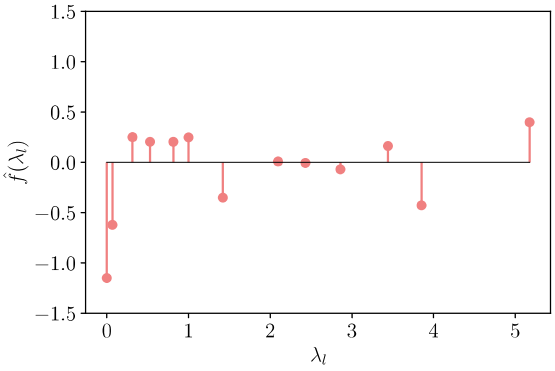


Source: Author

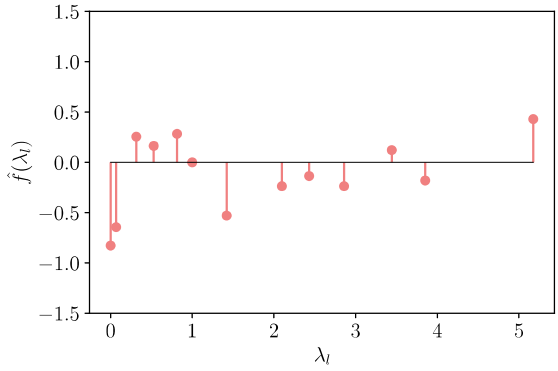
The spectrum of frequency for the 13NTFMB feeder was generating using GFT, as shown by Figure 37. The spectrum frequency graphic is different for the presented cases due the fact that it varies according to the graph signal. The C_{IGFT} index, importance coefficient of the graph signal, calculated by (5.28), interpret the hole importance spectrum graphic into a number and is shown by Table 17. As the importance information of the spectrum frequency are hidden in the higher frequencies, the cases with higher frequencies resulted in higher C_{IGFT}.

Figure 37 – Importance spectrum of 13NTFMB

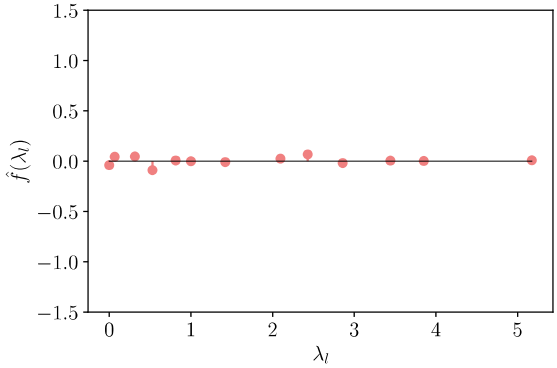
(a) Spectrum of reference case - Case #512



(b) Spectrum of Case #53



(c) Spectrum of Case #510



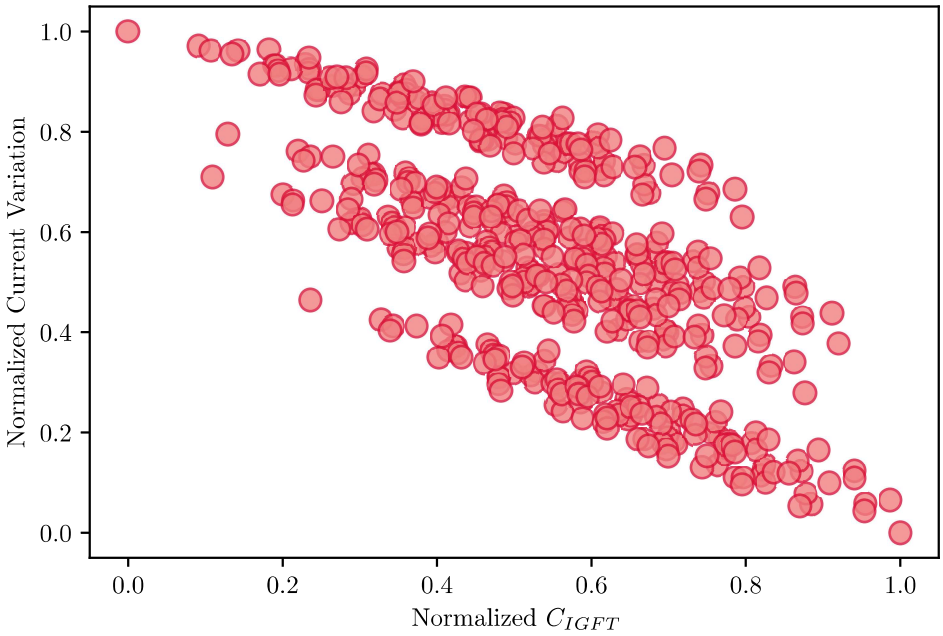
Source: Author

Table 17 – C_{IGFT} for 13NTFMB

# Case	C_{IGFT}
512	1.59
53	1.16
510	0.21

The current variation at node 650 and the C_{IGFT} values were resized and normalized and were related by Figure 38. The Pearson's rank correlation showed a correlation of -0.702 and the Spearman's -0.686. For this feeder, it is notice that the correlation analysis was mathematically worse, although it is still possible to extract the important information of the extreme current variation cases. The data present a negative direct correlation and each red circle illustrates one of the 512 cases.

Figure 38 – Normalized relation between current variation values in bus 650 and C_{IGFT} - 13NTFMB



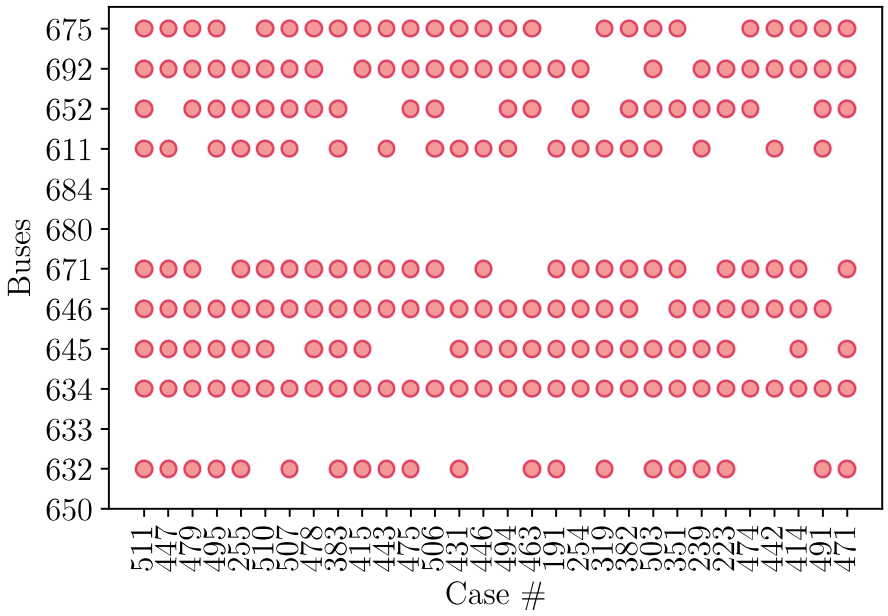
Source: Author

To identify the cases with higher and smaller values of current variation, the values of higher and smaller C_{IGFT} were analysed. Figure 39 and Figure 40 shows the 30 smaller and higher values of C_{IGFT}, respectively. The case number is shown in the abscissa axis and in the ordinate axis it is possible to observe the generators that were connected for each case.

As in the 13NTFMA, for the 13NTFMB the nodes 632 and 671 are the most important ones according to the centrality measures showed in Figure 35. Once again, in the 30 smaller values of C_{IGFT} for the 13NTFMB, the generators 632 and 671 are connect in the most of the cases, but not in all of then. The DG 634 are the one that are connected in all 30 smaller cases, indicating that for this feeder, this node is the most important one for the load current variation.

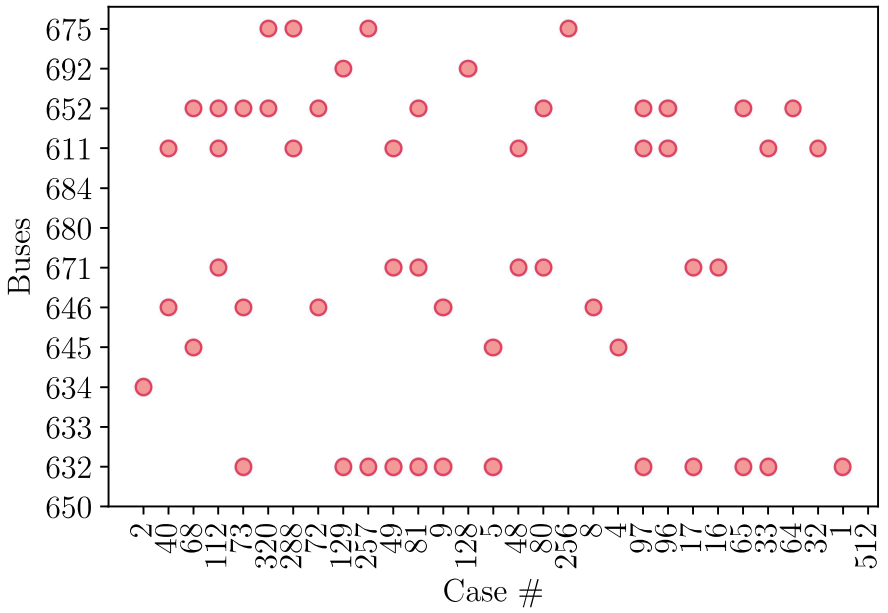
As each feeder has a different topology, it is expected distinct behaviours for the correlation data. Nevertheless, as already mentioned, it was possible to identify the cases of greatest interest, where the C_{IGFT} index is inversely related to the current variation at node 650.

Figure 39 – Cases number and the buses with DG connected in the 30 smaller values of C_{IGFT} of 13NTFMB



Source: Author

Figure 40 – Cases number and the buses with DG connected in the 30 higher values of C_{IGFT} of 13NTFMB

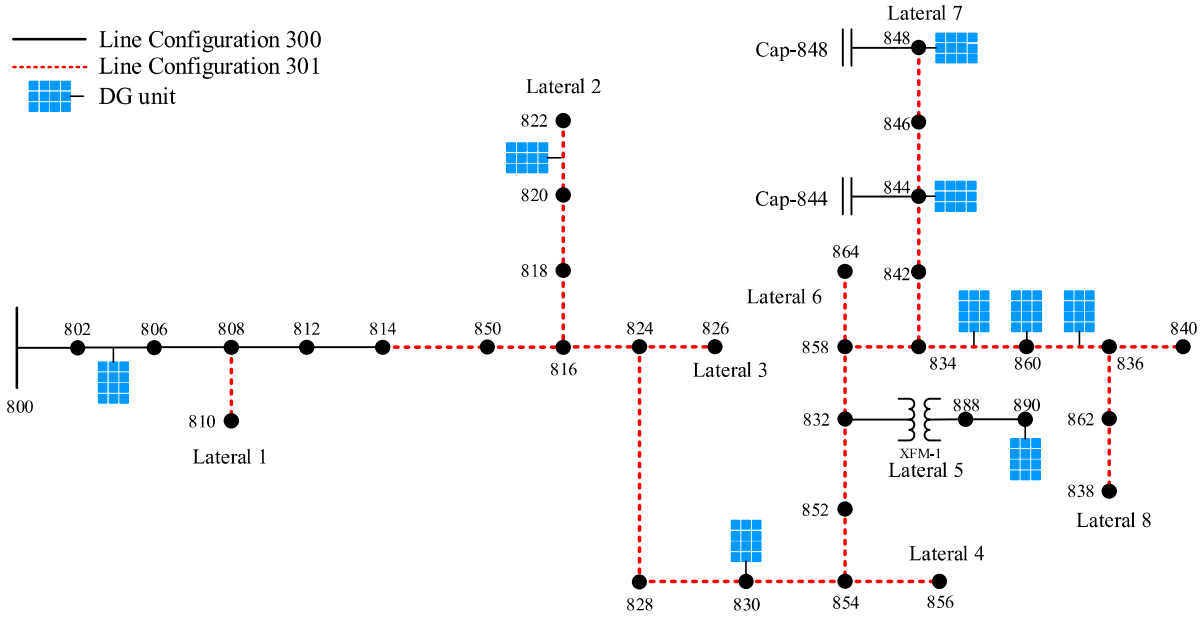


Source: Author

6.4 34-Node Test Feeder Modified (34NTFM)

To test the methodology on a feeder with more buses, the 34-Node Test Feeder (34NTF), also a circuit created by IEEE [89] and largely used in the literature was modified and used in this thesis. This feeder is an actual feeder located in Arizona (United States of America), with a nominal voltage of 24.9 kV. In its original form, presented by IEEE [89], it is characterized by long and lightly loaded, two in-line regulators, an in-line step-down transformer for a 4.16 kV brach, unbalanced loading, and shunt capacitors. The modifications of the 34NTF creates the 34NTFM, and were made to enable the application of the method contemplates the use of all three-phase feeder branches (this includes the use of some specific line configurations original of the 34NTF), as done with the previous feeders tested, and balanced loads. The 34NTFM is depicted by Figure 41. The regulators were once again removed to simplify the study and based in the fact that the DG tends to improve the voltade profile. The principal characteristics of the feeder lines are shown in Tables 18 and 19 and more details can be found in [89].

Figure 41 – 34NTFM



Source: Adapted from [89]

Table 18 – Basic information of Line Configuration Data of the 34NTFM system

Configuration	Phasing	Spacing ID	Type of Line
300	B A C N	500	Overhead Line
301	B A C N	500	Overhead Line

Source: Adapted from [89]

Table 19 – Load Data of the 34NTFM system

Node	Type of Load	Load Model	Real power in each phase (kW)	Reactive Power in each phase (kVAr)
802 - 806	Distributed	Y-PQ	55/3	29/3
808 - 810	Distributed	Y-I	16/3	8/3
818 - 820	Distributed	Y-Z	34/3	17/3
820 - 822	Distributed	Y-PQ	135/3	70/3
816 - 824	Distributed	D-I	5/3	2/3
824 - 826	Distributed	Y-I	40/3	20/3
824 - 828	Distributed	Y-PQ	4/3	2/3
828 - 830	Distributed	Y-PQ	7/3	3/3
854 - 856	Distributed	Y-PQ	4/3	2/3
832 - 858	Distributed	D-Z	15/3	7/3
858 - 864	Distributed	Y-PQ	2/3	1/3
858 - 834	Distributed	D-PQ	32/3	17/3
834 - 860	Distributed	D-Z	146/3	73/3
860 - 836	Distributed	D-PQ	82/3	43/3
836 - 840	Distributed	D-I	40/3	20/3
862 - 838	Distributed	Y-PQ	28/3	14/3
842 - 844	Distributed	Y-PQ	9/3	5/3
844 - 846	Distributed	Y-PQ	45/3	23/3
846 - 848	Distributed	Y-PQ	23/3	11/3
860	Spot	Y-PQ	20	16
840	Spot	Y-I	9	7
844	Spot	Y-Z	135	105
848	Spot	D-PQ	20	16
890	Spot	D-I	150	75
830	Spot	D-Z	45/3	20/3

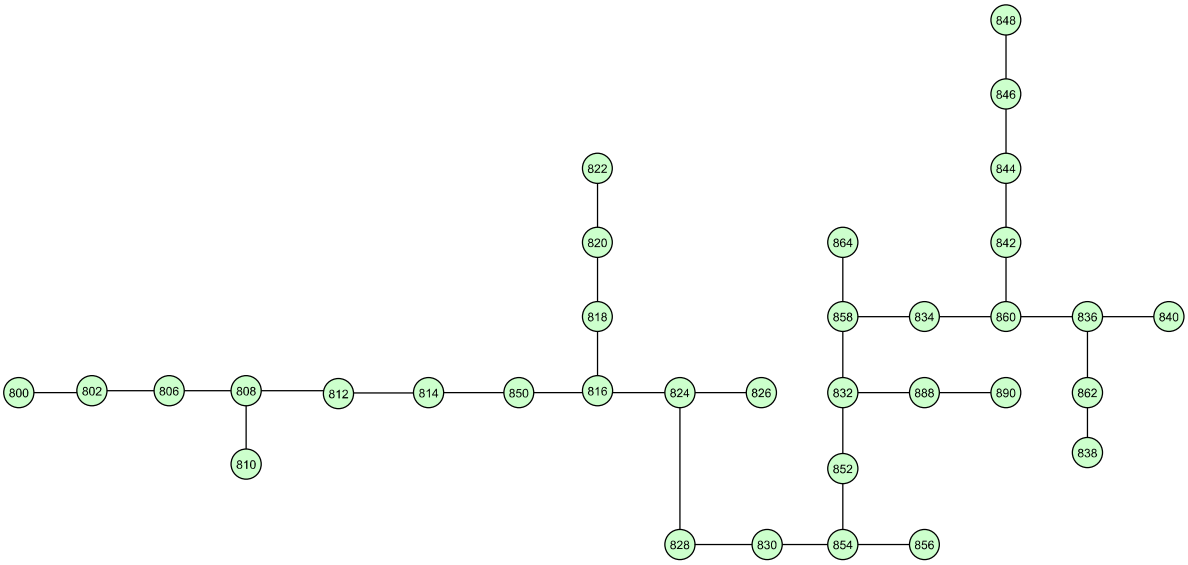
Source: Adapted from [89]

All the steps of the methodology were applied in the 34NTFM. For this feeder, the load current variation was analysed in the bus 800, which is the substation bus. As this feeders has 34 buses and 25 loads, the application of the proposed method for the DG topological position, presented in (6.1), for all the 25 loads will result in more than 33 millions of different cases, making the simulations in Simulink, an important step to the validation process, impossible. So, it was chosen to position the DG only in 9 loads, the 9 greater power values. This choice was an author decision to equalize the analyzes in this thesis but does not necessarily represent the cases where there will be smaller and larger current variation in the feeder (considering the total 33 millions cases). This criterion generates 512 cases, making the simulation feasible. The topological position and DG powers are represented in Table 19 and Figure 41. In the reference case, where there is no DG unit connected in the feeder, the current at substation bus 800 is 28.57 A. For the case with the higher current variation, the load current at node 800 is equal to 25.34 A, a

variation of 11.3%.

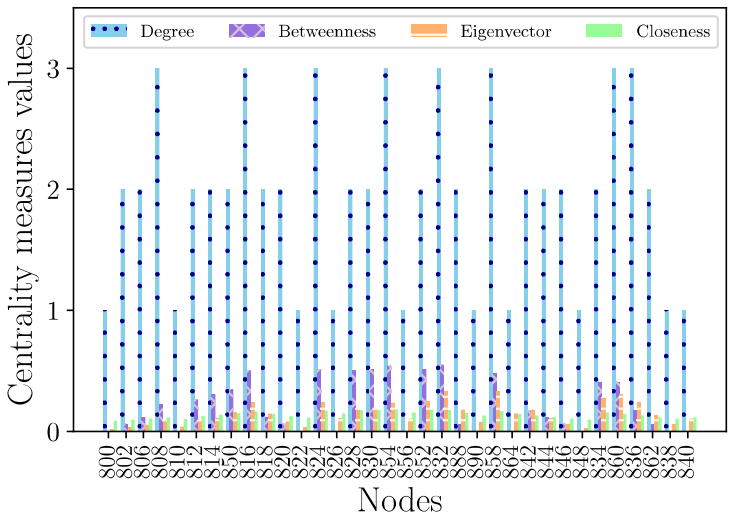
The graph that represents the 34NTFM, considering no weight for the edges are depicted by Figure 42. The centrality measures were, once again, calculated to identify the most important nodes for the graph classic analysis and the results are shown in Figure 43. For the Degree centrality measure, the nodes 808, 816, 824, 854, 832, 858, 860 and 836 are the most important. For the Betweenness and Closeness, the node 854 is the most important and the node 832 are the most important for the Eigenvector centrality measure.

Figure 42 – 34NTFM graph representation



Source: Author

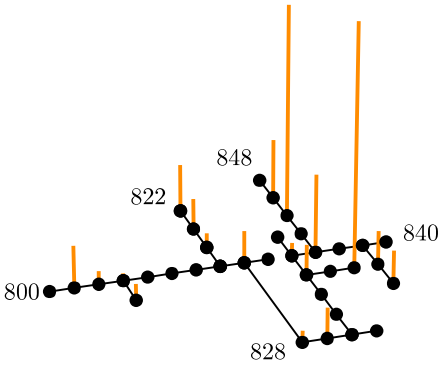
Figure 43 – Centrality measures of 34NTFM: Degree, Betweenness, Eigenvector and Closeness



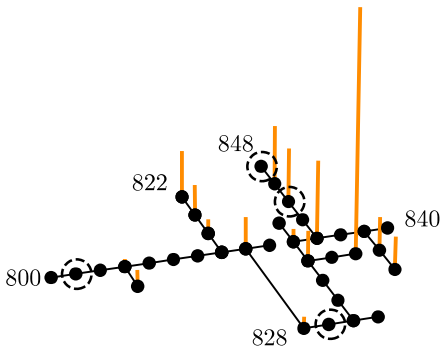
The graph signal c_P of some of the 512 cases are shown in Figure 44 where, again, the dotted circles represents a DG unit connected in that bus and the orange bars the absolute value of c_P. The importance spectrum of the respective cases are presented by Figure 45 and the C_{IGFT} for this cases, calculated using (5.28), are depicted by Table 20 below.

Figure 44 – Graph signal c_P of some cases of 34NTFM (Figure 42)

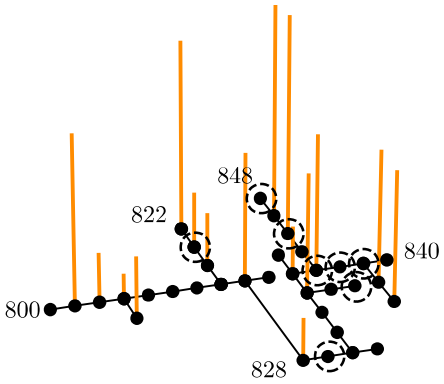
(a) Graph signal of reference case - Case #512



(b) Graph signal of Case #53



(c) Graph signal of Case #510

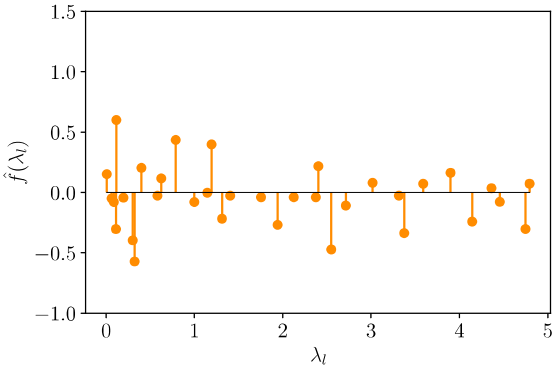


Source: Author

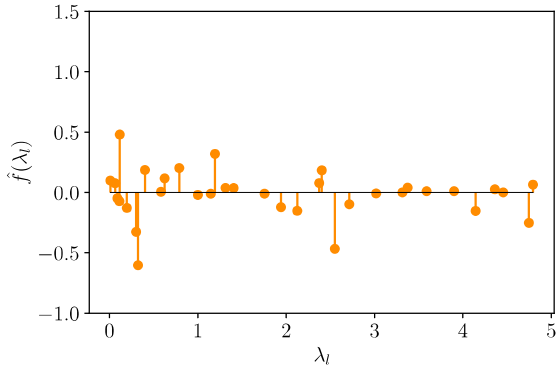
It is important to highlight that, for this feeder, the number of buses with PV generators is less than the number of buses with load and exist some distributed load as shown by Table 19. In this thesis, the distributed loads are represented as two loads: half of the total load are represented in the bus in the beginning of the distribution line and the other half is represented in the end of the line. This information explains the fact that, for this feeder, the c_P values may be different from zero even if the DG is connected in the bus (the c_P comes from the distributed load).

Figure 45 – Importance spectrum of 34NTFM

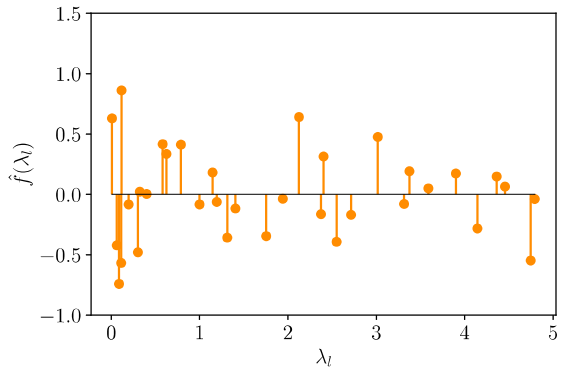
(a) Spectrum of reference case - Case #512



(b) Spectrum of Case #53



(c) Spectrum of Case #510



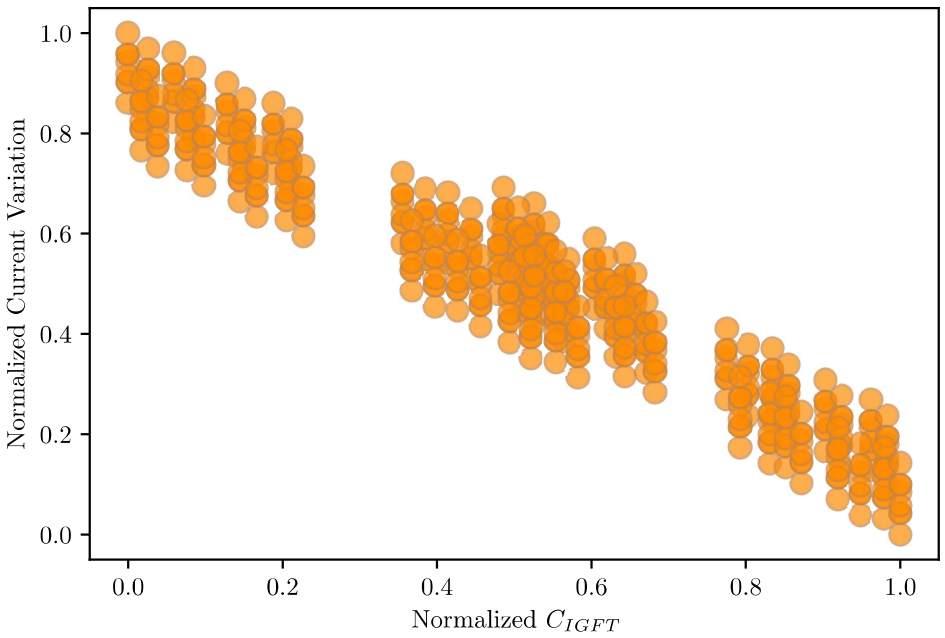
Source: Author

Table 20 – C_{IGFT} for 34NTFM

# Case	C_{IGFT}
512	30.54
53	23.81
510	13.34

The relationship between the C_{IGFT} calculated and the load current variation are presented in Figure 46. The values are normalized and the smaller values of normalized C_{IGFT} indicates greater current variation, which means smaller absolute values in practice and vice versa. The Pearson's rank correlation for this feeder is -0.96 and the Spearman's rank correlation is -0.95, indicating a very good correlation between data. It is possible to observe, that the data correlation, once again, does not behave exactly like the others presented above (Figures 22, 30 and 38). This is totally expected, since the feeders are completely different. It is possible to identify the cases with higher and smaller current variation using the C_{IGFT}, though.

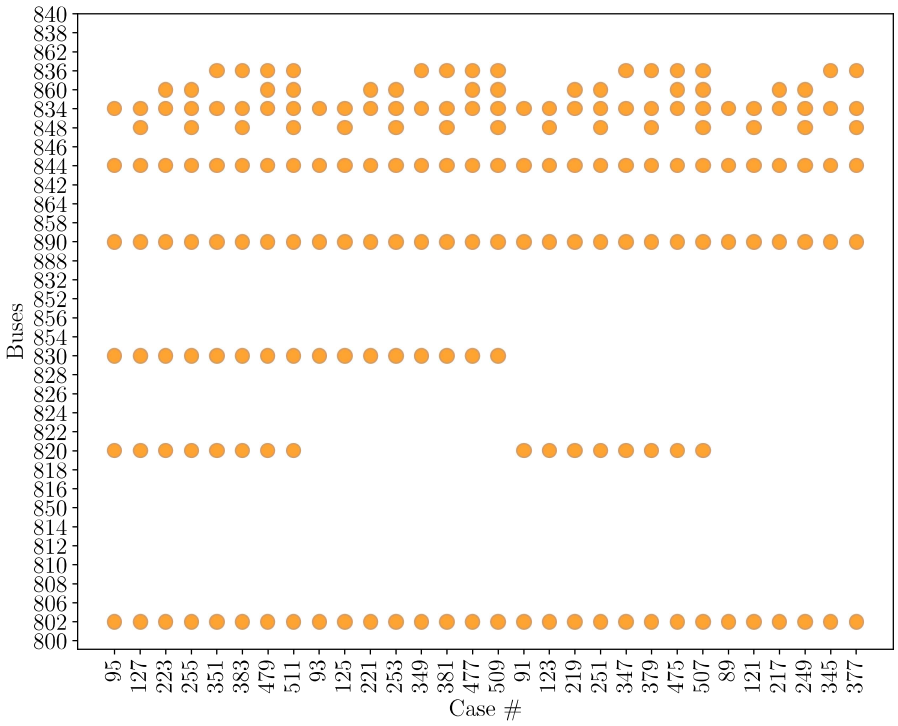
Figure 46 – Normalized relation between current variation values in bus 800 and C_{IGFT}



Source: Author

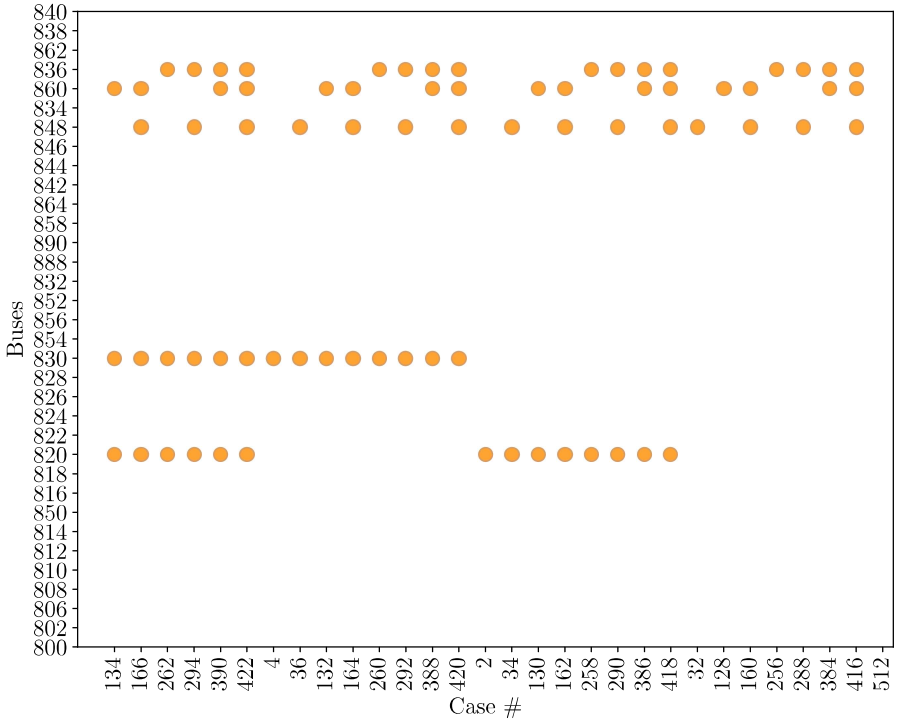
An analysis on the data of 30 highest and lowest value of C_{IGFT} was done, once again aiming to identify which DG unit was generating in the cases that presented greater and smaller load current variation. The results are depicted in Figures 47 and 48. For this feeder, the relation of the data of Figures 47 and 48 with Figure 43 are more complex, since the centrality measures resulted in different values according to the centrality used. However, it is possible to notice that, 3 buses with DG units appear in all 30 smaller C_{IGFT} of Figure 47: 802, 844 and 890. None of these buses were pointed as the most important ones by the centrality measures.

Figure 47 – Cases number and the buses with DG connected in the 30 smaller values of C_{IGFT} of 34NTFM



Source: Author

Figure 48 – Cases number and the buses with DG connected in the 30 higher values of C_{IGFT} of 34NTFM



Source: Author

6.5 Common Discussion

To evaluate the degree of accuracy of the proposed method for each feeder, an analysis of the extreme data, what means higher and smaller values of C_{IGFT}, are compiled in Table 21 which shows the percentage of correct cases found using C_{IGFT}, for different number of samples, for all the studied feeders. Table 21 is a comparison of the smaller and higher current variation cases found by the proposed method and the cases found by the MATLAB/Simulink. It is possible to see in Table 21 that depending on the number of samples analysed, the accuracy may vary. For the five smaller and five greater samples, the two first lines of Table 21, all the feeders presented an accuracy of 60%, this means that the C_{IGFT} was able to identify correctly 60% of the cases with higher and smaller current variation. Considering 200 samples, which is 39% of the total samples studied in this research, the C_{IGFT} was capable to correctly identify more than 70% of the cases with extreme current variation for all studied feeder.

Table 21 – Percentage of correct cases found using C_{IGFT}, for different number of samples, for all the studied feeders

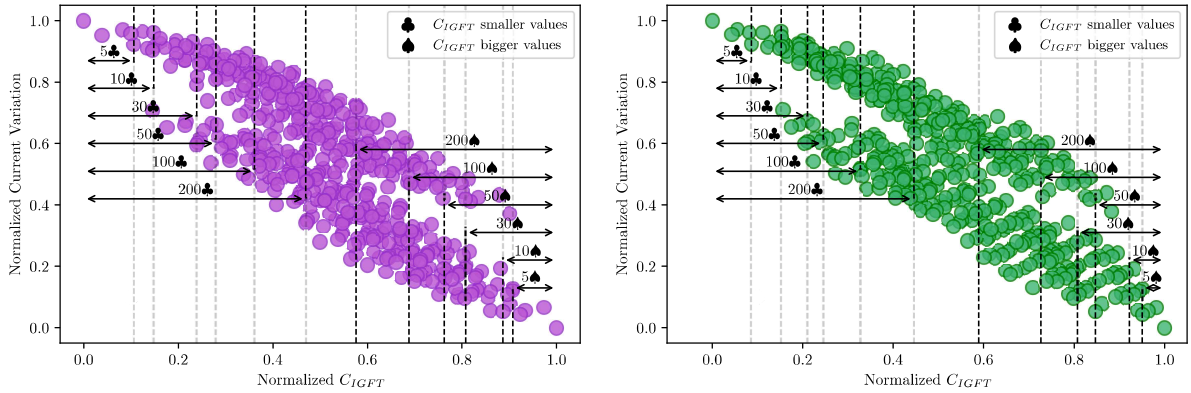
Number of analysed samples	**‡ C_{IGFT}**	**13NTFM**	**13NTFMA**	**13NTFMB**	**34NTFM**
		Percentage of correct cases			
5 ‡ samples or 0.98% of the total samples	**Smaller**	60	60	60	60
	Higher	60	60	60	60
10 ‡ samples or 1.95% of the total samples	**Smaller**	60	70	70	40
	Higher	60	50	50	40
30 ‡ samples or 5.86% of the total samples	**Smaller**	70	73.3	63.3	60
	Higher	60	56.7	60	60
50 ‡ samples or 9.76% of the total samples	**Smaller**	72	76	58	70
	Higher	68	62	56	70
100 ‡ samples or 19.5% of the total samples	**Smaller**	68	71	50	88
	Higher	68	71	50	88
200 ‡ samples or 39% of the total samples	**Smaller**	73.5	70.5	71.5	83
	Higher	72.5	72	70	84.5

Figure 49 presents the 5, 10, 30, 50, 100 and 200 smaller and higher samples considering the C_{IGFT} values of the four feeders analysed in this thesis. Figure 49a presents the 5, 10, 30, 50, 100 and 200 smaller and higher samples considering the C_{IGFT} values of 13NTFM and it is possible to assume that, for the 5 smaller C_{IGFT} values, the current variation is greater than 0.92. For the 10, 30, 50, 100 and 200 smaller C_{IGFT} values shown in Figure 49a , the current variation is greater than 0.71, 0.59, 0.59, 0.48 and 0.34, respectively. For the 5 and 10 higher values of C_{IGFT} the current variation is up to 0.12 and 0.37, respectively. The current variation is up to 0.48 for the 30 higher values of C_{IGFT}, up to 0.50 for the 50 higher C_{IGFT}, up to 0.55 for the 100 higher values of C_{IGFT} and up to 0.71 for the 200 higher values of C_{IGFT}. The same analysis was done for the other feeders, as showed in Figures 49b, 49c and 49d. The results presented in Figure 49d do not show the same groups of analysis just aiming to facilitate the visualization, since the values found for the samples were very close from each other, but the same 6 groups for the smaller and higher values of C_{IGFT} were analysed in the 34NTFM.

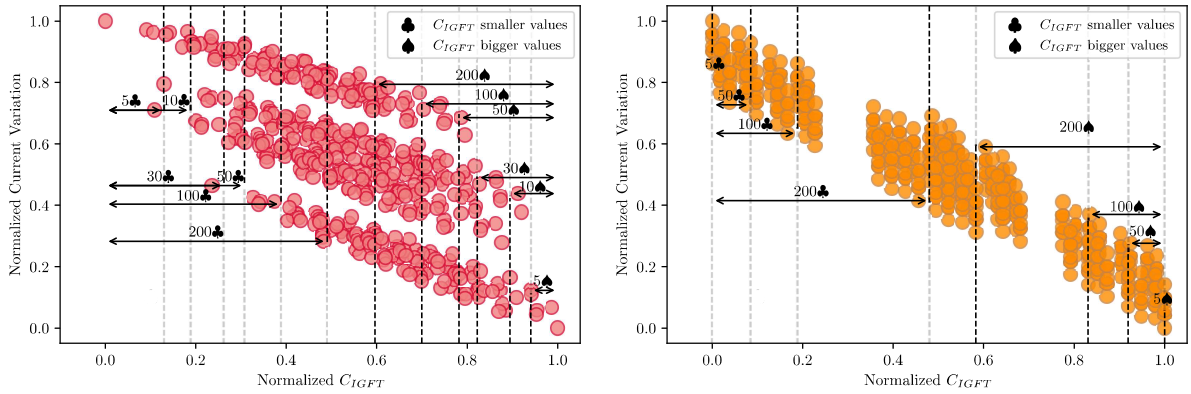
Thus, Figure 49a helps the understanding of Table 21 data, considering that results presented are complementary. The combination of Table 21 and Figure 49 allows to determine the accuracy of the proposed method and the percentage of current variation based on a group of analysis. They help to identify the cases were the DG units will impact the most in the load current variation and permit to predict the approximate variation of the load current only based in the C_{IGFT} values.

Figure 49 – Smaller and bigger values of C_{IGFT}

(a) 13NTFM Smaller and bigger values of C_{IGFT} (b) 13NTFMA Smaller and bigger values of C_{IGFT}



(c) 13NTFMB Smaller and bigger values of C_{IGFT} (d) 34NTFM Smaller and bigger values of C_{IGFT}



Source: Author

It is important to highlight that, for the analysed feeders, the proposed method presented different nodes as the most important considering the load current variation in comparison with the calculated centrality measures. This information is important because it reinforces that the traditional centrality measures, which basically consider the topology of the graph, are not enough to indicate the important nodes when it comes to the load current variation in feeders with high integration of DG.

Although computer simulations have evolved considerably in recent decades, the need to simulation hundreds or thousands of different cases are very computationally costly. For this thesis, the electrical simulation performed in MATLAB/Simulink for the feeders with 13 nodes took approximately 17.8 minutes, in average, to perform the 512 distinct cases. For the feeders with 13 nodes and 48 minutes for the feeder with 34 buses. The whole algorithm to perform the proposed method and proposed system to find the C_{IGFT} takes 10 minutes to run for the 13NTFM, 13NTFMA and 13NTFMB and 23 minutes to run the code for the 34NTFM. These values represent a reduction of 43.8% and 52%, respectively, in the simulation time. The most time-consuming part is modelling the feeder and calculating c_p for all cases. If the *graph signals* (c_p) are known, the proposed

methodology reduced the simulation time in at about 99.12%. All data are related to these proposed systems simulated. Table 22 depicts the average time for the proposed method and for the Simulink simulations.

Table 22 – Time processing - Proposed Methodology and Simulink data according to the feeder buses quantity

Feeder	**Simulink**	**Proposed Method**	
		Without calculating c_P	**Calculating c_P**
13 buses	17.8 minutes	10 seconds	10 minutes
34 buses	48 minutes	10 seconds	23 minutes

So, use a method that are faster and less computationally costly is interesting, especially considering feeders with a large number of buses, where the simulation time in a power flow software tends to be high. Despite that, more important than saving simulation time is the possibility of reducing the number of cases under analysis among the total number of cases. Considering that there are power flow software more efficient than MATLAB, which simulate feeders with hundreds of buses, use the C_{IGFT} to filter the cases with small and higher current variation are very useful for the literature and electric utilities. Thus, with the proposed method it is possible to identify the most critical cases, the DG units that are connected in these cases and, if it is necessary, simulate only these specific states in a power flow software, reducing the problem dimension.

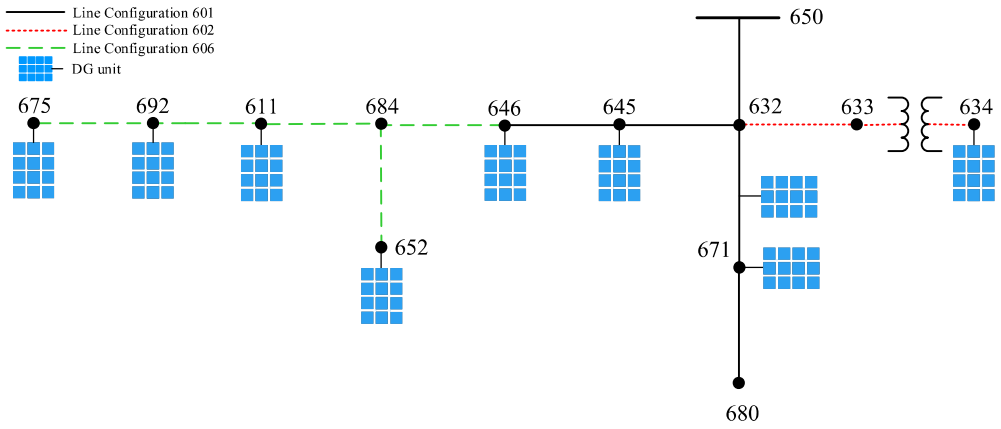
Thus, testing four different feeders, it was possible to validate the proposed methodology. For each feeder presented was observed that the results are not identical but the method behaves well and allows the identification of the cases of greater and smaller current variation in the substation bus, using only the C_{IGFT} index, which is the main core of this thesis. It is important to highlight that there may be infinite distribution feeders with different electric characteristics. Depending on the feeder size or topology, adaptations in the coefficient of importance signal may be necessary, for example, changing the constants or using other equations for calculating the importance coefficient already consolidated in the literature.

6.6 Case Study: 13-Node Test Feeder Modified C (13NTFMC)

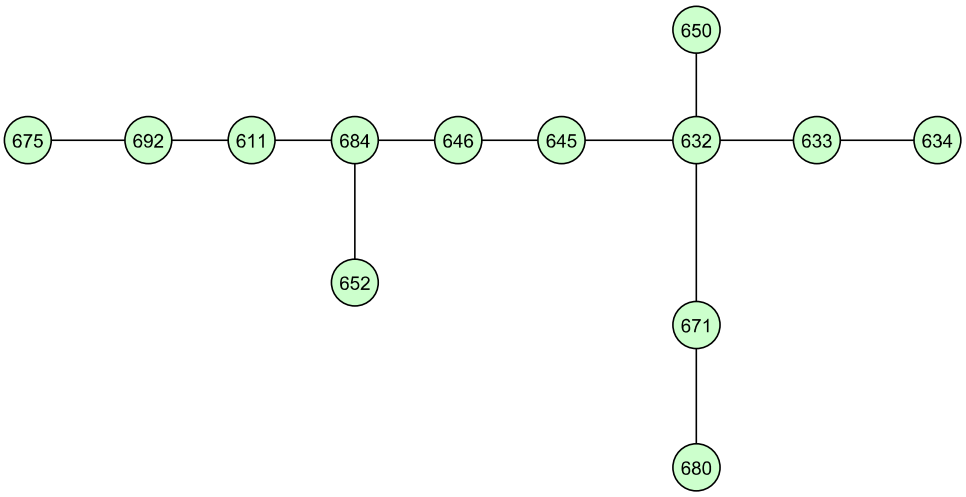
In this section, after validated the methodology, a new configuration of the 13NTFM is proposed: the 13NTFMC. It is presented in Figure 50, where is shown the electrical feeder and the 13NTFMC graph. It is important to reinforce that the topology of this feeder was, once again, chosen randomly.

Figure 50 – 13NTFMC

(a) 13NTFMC electrical feeder



(b) 13NTFMC graph



Source: Author

Different from the other feeders shown above, where the objective was to validate the proposed methodology, in this study case, the C_{IGFT} is used directly to find out the cases of interest, which are the situations where the load current variation at substation are higher and smaller.

The complete method presented in chapter 5 was applied and, knowing the C_{IGFT} for each one of the 512 cases it is possible to estimate the configurations were the load current varies more. The data obtained in Table 21 was used as a basis for choosing the number of cases to be analyzed, aiming to ensure a good accuracy in the results. In this way, the 50 highest and lowest values of C_{IGFT} will be used to identify the cases of highest and lowest load current variation.

Thus, the cases of fifty small values of C_{IGFT} and the cases of fifty bigger values of C_{IGFT} are presented in Figure 51 sorted in ascending order, illustrating the buses where

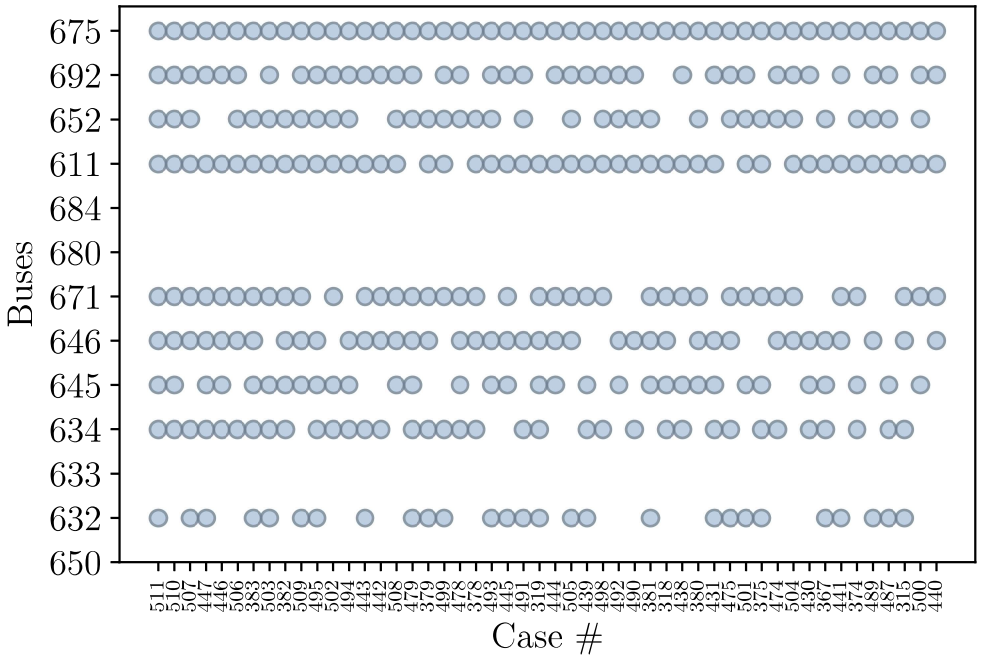
there is DG units connected. Figure 51a shows the 50 smaller values of C_{IGFT} which implies in the 50 higher current variations and Figure 51b shows the opposite, the 50 higher values of C_{IGFT} which implies in the 50 smaller current variations.

In Figure 51a, it is possible to notice that most of the 50 smaller cases presented have generators connected at nodes 611, 671 and 675. This is an indication that the cases with DG connected to these nodes tends to have a higher current variation at the substation bus. Thus, when the DG are connected at nodes 611, 671 and 675, power flow software simulations should be necessary for further analysis.

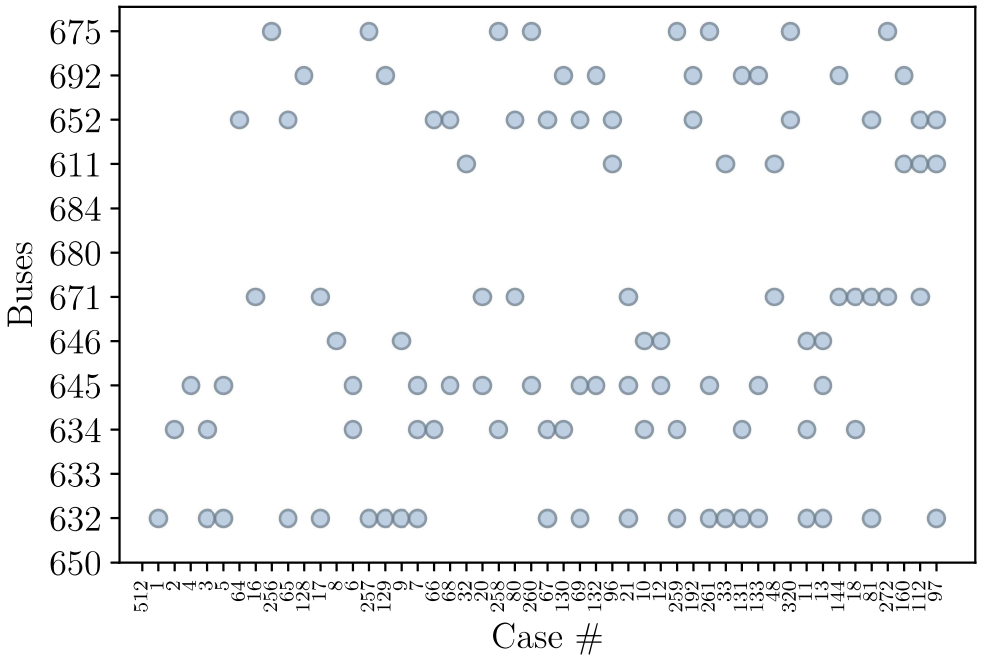
With the information presented in Figure 51, the utilities or the researchers can identify the most critical cases and simulate only this configurations in a load flow software, if necessary. This provides a reduction in the number of cases that need to be studied, reducing the time and complexity of simulations in power flow software.

Figure 51 – 13NTFMC

(a) Cases number and the buses with DG connected in the 50 smaller values of C_{IGFT} of 13NTFMC



(b) Cases number and the buses with DG connected in the 50 higher values of C_{IGFT} of 13NTFMC



Source: Author

7 CONCLUSION AND FUTURE WORK

7.1 Conclusion

The high integration of photovoltaic distributed generation is happening quickly worldwide and as the insertion of new generating sources can change the power flow in a distribution feeder, studies are essential to analyse the feeder behavior in this new scenario, ensuring a good indicator of voltage and current levels.

A novel method that associates concepts of power flow and graph signals is presented in this thesis and proposes an original approach from SPG in electrical feeders, according to their importance in the grid power flow and the power of the load/DG in that node. The methodology integrates mathematical fields to allow the identification of the topological DG positions that influence the most in the load current variation. Because the electrical feeders can be described as nodes and edges connected, the graph theory is used to model the grid. The insertion of DG units in an electrical power grid can change the power flow. So, to ponder the generators connected along the grid, the c index obtained from the load flow algebraic formulation is used, since besides to considering the power at that node, it manages to weight the bus considering the entire power flow of the network. The use of SPG method allows to ponder the buses instead of the edges, admits the use of complex numbers to weight the buses and considers vectors of importance graph signals, that consist in an analysis that rely on the weights of the all buses, making it possible to analyze and compare the different DG combination cases proposed in this thesis.

In the tested systems, three variations of the 13-Node Test Feeder (13NTFM, 13NTFMA and 13NTFMB) and a modification of the 34-Node Test Feeder (34NTFM), 512 combinations of feeder load and distributed generators power were analysed, creating 512 different C_{IGFT} values for each feeder. The C_{IGFT} values represents the load current variation in the substation bus in relation to the reference case. The Spearman's and Pearson's rank-order correlation were calculated to fit and validate the methodology. The results showed a good agreement between the C_{IGFT} values calculated and the Simulink current variation data. The methodology allows to identify the cases with higher and smaller current variation only through the C_{IGFT} values with a good accuracy. The use of this method to analyse the current variation reduced the computational time by 50% when compared to the electrical simulations performed in MATLAB/Simulink.

More important than saving simulation time is reducing the number of cases under analysis among the total number of cases. The C_{IGFT} can be used to filter the cases with small and higher current variation. Thus, the proposed method allows the identification of the most critical cases and the DG units that are connected in these cases. With that, it

is possible to reduce the problem dimension, which is very useful for the literature and for the electric utilities.

The accuracy and assertiveness of the proposed method varies according to the tested feeder and the number of cases under analysis. For the 13-Node feeders, considering up to the 30 largest and smaller (up to 5.86% of the total 512 analysed cases) values of C_{IGFT}, the methodology correctly identified about 60% of the cases of greater and smaller current variation compared to the result found by Simulink. For the same analysis, the method correctly identified at about 53.33% of the cases for the 34NTFM. Over 50 samples or 9.76% of the samples, the methodology was able to correctly identify 66.7% of the smaller and higher cases of current variation for 13-Node feeders and 80.6% for the 34NTFM.

The methodology proposed by this thesis allows to identify the most critical cases of current variation using a mathematical model, with no need to use a power flow software simulation. Knowing that each C_{IGFT} value is associated to a different case, the method also determine which distributed generator is connected in those cases, supporting the determination of the nodes that should not have a distributed generator connected. With this, if necessary, only the cases of interest identified by C_{IGFT} should be studied, reducing the number of simulations and saving time and computational processing.

7.2 Future Work

The next steps to continue and improve this methodology research are:

- Develop the modelling used for SPG, aiming to find an equation for the importance coefficient that fits better the load current variation values simulated;
- Expand the methodology for other feeders;
- Expand the methodology considering load and demand curve variation for the feeders;
- Expand the methodology considering the intermittence of solar generation;
- Implement the proposed methodology for unbalanced electrical system with single-phase, phase-phase and three-phase branches;
- Implement the proposed methodology considering the effect of the voltage regulators;
- Expand the methodology considering the load current variation in all buses;
- Investigate the applicability to analyse the bus voltage variation;
- Compare the methodology results with other load flow software.

BIBLIOGRAPHY

[1] SIWAKOTI, Y. P.; BLAABJERG, F.; LOH, P. C. High Step-Up Trans-Inverse (Tx-1) DC–DC Converter for the Distributed Generation System. *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, v. 63, p. 4278 – 4291, 2019. Cited 2 times at pages 18 and 26.

[2] VASCONCELOS, P. G. S.; GONZÁLEZ, M. O. A. Photovoltaic solar energy : Conceptual framework. *Renewable and Sustainable Energy Reviews*, v. 74, n. February, p. 590–601, 2017. Cited 2 times at pages 18 and 26.

[3] WANG, Z.; SCAGLIONE, A.; THOMAS, R. J. Electrical centrality measures for power grids. *Control and Optimization Methods for Electric Smart Grids*, IEEE, n. 2009, p. 239–255, 2012. Cited 4 times at pages 18, 59, 60, and 61.

[4] Agência Nacional de Energia Elétrica. *Unidades Consumidoras com Geração Distribuída*. 2021. Disponível em: <http://www.aneel.gov.br/outorgas/geracao/-/asset_publisher/mJhnKIi7qcJG/content/registro-de-central-geradora-de-capacidade-reduzida/655808?inheritRedirect=false&redirect=http%3A%2F%2Fwww.aneel.gov.br%2Foutorgas%2Fgeracao%3Fp_p_id%3D101_INSTANCE_mJhnKIi7qc>. Cited on page 27.

[5] MUTTAQI, K. M. et al. An algebraic approach for determination of DG parameters to support voltage profiles in radial distribution networks. *IEEE Transactions on Smart Grid*, v. 5, n. 3, p. 1351–1360, 2014. ISSN 19493053. Cited on page 18.

[6] VARGAS, M. C.; MENDES, M. A.; BATISTA, O. E. Impacts of high PV penetration on voltage profile of distribution feeders under brazilian electricity regulation. In: *13th IEEE International Conference on Industry Applications*. [S.l.: s.n.], 2018. ISBN 9781538679951. Cited 4 times at pages 18, 20, 28, and 81.

[7] JUANUWATTANAKUL, P.; MASOUM, M. A. Voltage stability enhancement for unbalanced multiphase distribution networks. *IEEE Power and Energy Society General Meeting*, IEEE, p. 1–6, 2011. ISSN 19449925. Cited 3 times at pages 18, 19, and 27.

[8] SA'ED, J. A. et al. Reassessment of voltage stability for distribution networks in presence of DG. In: *International Conference on Environment and Electrical Engineering*. [S.l.: s.n.], 2016. ISBN 9781509023196. Cited 3 times at pages 18, 19, and 27.

[9] BARKER, P. P.; De Mello, R. W. Determining the impact of distributed generation on power systems: Parte 1 - Radial Distribution Systems. *Power Engineering Society Summer Meeting, 2000. IEEE*, v. 3, n. c, p. 1645–1656 vol. 3, 2000. Cited on page 18.

[10] BLAABJERG, F. et al. Distributed power-generation systems and protection. *Proceedings of the IEEE*, v. 105, n. 7, p. 1311–1331, 2017. Cited on page 18.

[11] MESKIN, M.; DOMIJAN, A.; GRINBERG, I. Impact of distributed generation on the protection systems of distribution networks: analysis and remedies – review paper. *IET Generation, Transmission & Distribution*, v. 14, n. 24, p. 5944–5960, 2020. Disponível em: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-gtd.2019.1652>. Cited on page 18.

[12] MENDES, M. A. et al. Simplified Single-phase PV Generator Model for Distribution Feeders With High Penetration of Power Electronics-based Systems. In: *2019 IEEE 15th Brazilian Power Electronics Conference and 5th Southern Power Electronics Conference (COBEP/SPEC)*. Santos: [s.n.], 2019. Cited 7 times at pages 18, 27, 28, 74, 76, 83, and 84.

[13] CAMPION, M.; RANGANATHAN, P. Identification of critical buses based on betweenness-centrality in a smart grid. In: *IEEE Electrical Power and Energy Conference*. [S.l.: s.n.], 2017. p. 1–5. ISBN 9781538608173. Cited 3 times at pages 19, 20, and 52.

[14] COELHO, E. P. R. et al. A New Approach for Contingency Analysis Based on Centrality Measures. *IEEE SYSTEMS JOURNAL*, v. 13, 2019. Cited 3 times at pages 19, 20, and 52.

[15] JOOSHAKI, M. et al. A model for stochastic planning of distribution network and autonomous dg units. *IEEE Transactions on Industrial Informatics*, v. 16, n. 6, p. 3685–3696, 2020. Cited on page 20.

[16] SHAABAN, M. F.; EL-SAADANY, E. F. Accommodating high penetrations of pevs and renewable dg considering uncertainties in distribution systems. *IEEE Transactions on Power Systems*, v. 29, n. 1, p. 259–270, 2014. Cited on page 20.

[17] MACHADO, I.; BORBA, B.; MACIEL, R. Modeling distributed pv market and its impacts on distribution system: A brazilian case study. *IEEE Latin America Transactions*, v. 14, n. 11, p. 4520–4526, 2016. Cited on page 20.

[18] GE, L. et al. Optimal integrated energy system planning with dg uncertainty affine model and carbon emissions charges. *IEEE Transactions on Sustainable Energy*, v. 13, n. 2, p. 905–918, 2022. Cited on page 20.

[19] GöZEL, T.; HOCAOGLU, M. H. An analytical method for the sizing and siting of distributed generators in radial systems. *Electric Power Systems Research*, v. 79, p. 912–918, 6 2009. ISSN 03787796. Cited on page 20.

[20] HERNáNDEZ, J.; RUIZ-RODRIGUEZ, F.; JURADO, F. Technical impact of photovoltaic-distributed generation on radial distribution systems: Stochastic simulations for a feeder in spain. *International Journal of Electrical Power & Energy Systems*, v. 50, p. 25–32, 2013. ISSN 0142-0615. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0142061513000707>. Cited 2 times at pages 20 and 27.

[21] ABRI, R. S. A.; EL-SAADANY, E. F.; ATWA, Y. M. Optimal placement and sizing method to improve the voltage stability margin in a distribution system using distributed generation. *IEEE Transactions on Power Systems*, v. 28, n. 1, p. 326–334, 2013. Cited on page 20.

[22] ERDINÇ, O. et al. Comprehensive optimization model for sizing and siting of dg units, ev charging stations, and energy storage systems. *IEEE Transactions on Smart Grid*, v. 9, n. 4, p. 3871–3882, 2018. Cited on page 20.

[23] BARIN, A. et al. Methodology for placement of dispersed generation systems by analyzing its impacts in distribution networks. *IEEE Latin America Transactions*, v. 10, n. 2, p. 1544–1549, 2012. Cited on page 20.

[24] HEJAZI, H. A. et al. Independent distributed generation planning to profit both utility and dg investors. *IEEE Transactions on Power Systems*, v. 28, n. 2, p. 1170–1178, 2013. Cited on page 20.

[25] ASRARI, A.; WU, T.; LOTFIFARD, S. The impacts of distributed energy sources on distribution network reconfiguration. *IEEE Transactions on Energy Conversion*, v. 31, n. 2, p. 606–613, 2016. Cited on page 20.

[26] MATOS, S. et al. Protection philosophy for distribution grids with high penetration of distributed generation. *Electric Power Systems Research*, v. 196, p. 107203, 2021. ISSN 0378-7796. Disponível em: <https://www.sciencedirect.com/science/article/pii/S037877962100184X>. Cited 3 times at pages 20, 27, and 28.

[27] VARGAS, M. C.; MENDES, M. A.; BATISTA, O. E. Fault Current Analysis on Distribution Feeders with High Integration of Small Scale PV Generation. In: *IEEE Power and Energy Society General Meeting*. [S.l.: s.n.], 2019. ISBN 9781728119816. ISSN 19449933. Cited 2 times at pages 20 and 28.

[28] MENDES, M. A. et al. Load Currents Behavior in Distribution Feeders Dominated by Photovoltaic Distributed Generation. *Electric Power Systems Research*, v. 201, n. August, 2021. ISSN 03787796. Cited 5 times at pages 20, 27, 74, 76, and 81.

[29] ZHAN, H. et al. Relay protection coordination integrated optimal placement and sizing of distributed generation sources in distribution networks. *IEEE Transactions on Smart Grid*, v. 7, n. 1, p. 55–65, 2016. Cited on page 20.

[30] ZEINELDIN, H. H. et al. A protection coordination index for evaluating distributed generation impacts on protection for meshed distribution systems. *IEEE Transactions on Smart Grid*, v. 4, n. 3, p. 1523–1532, 2013. Cited on page 20.

[31] ERNSTER, T. A.; SRIVASTAVA, A. K. Power system vulnerability analysis - Towards validation of centrality measures. In: *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*. [S.l.]: IEEE, 2012. p. 1–6. ISBN 9781467319348. ISSN 21608555. Cited 3 times at pages 20, 52, and 62.

[32] ADEBAYO, I.; JIMOH, A. A.; YUSUFF, A. Voltage stability assessment and identification of important nodes in power transmission network through network response structural characteristics. *IET Generation, Transmission and Distribution*, Institution of Engineering and Technology, v. 11, p. 1398–1408, 4 2017. ISSN 17518687. Cited on page 20.

[33] NASIRUZZAMAN, A. B.; POTA, H. R. Bus dependency matrix of electrical power systems. *International Journal of Electrical Power and Energy Systems*, v. 56, p. 33–41, 2014. ISSN 01420615. Cited on page 20.

[34] CARO-RUIZ, C.; MOJICA-NAVA, E. Centrality measures for voltage instability analysis in power networks. In: *2015 IEEE 2nd Colombian Conference on Automatic Control, CCAC 2015 - Conference Proceedings*. [S.l.]: IEEE, 2015. ISBN 9781467393058. Cited 2 times at pages 20 and 62.

[35] PAGANI, G. A.; AIELLO, M. The Power Grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V.,

v. 392, n. 11, p. 2688–2700, 2013. ISSN 03784371. Disponível em: <http://dx.doi.org/10.1016/j.physa.2013.01.023>. Cited 2 times at pages 20 and 52.

[36] CAMARGO, R. S. et al. A novel cascaded multilevel converter topology based on three-phase cells with model predictive control. In: *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)*. [S.l.: s.n.], 2020. p. 1161–1166. Cited on page 20.

[37] PONTES, G. P. et al. A control method for dg based on automatic current slope gain. In: *2019 IEEE Applied Power Electronics Conference and Exposition (APEC)*. [S.l.: s.n.], 2019. p. 3269–3273. Cited on page 20.

[38] SMIRNOV, A. I.; SHKLYARSKIY, J. E. Adaptive current protection in electrical complexes with small power plants. In: . [S.l.]: EDP Sciences, 2021. v. 266. ISSN 22671242. Cited on page 21.

[39] ABUL'WAFA, A. R. A network-topology-based load flow for radial distribution networks with composite and exponential load. *Electric Power Systems Research*, Elsevier, v. 91, p. 37–43, 10 2012. ISSN 0378-7796. Cited on page 21.

[40] RAMAKRISHNA, R.; SCAGLIONE, A. Grid-graph signal processing (grid-gsp): A graph signal processing framework for the power grid. *IEEE Transactions on Signal Processing*, v. 69, p. 2725–2739, 2021. Cited on page 21.

[41] DRAYER, E.; ROUTTENBERG, T. Detection of false data injection attacks in smart grids based on graph signal processing. *IEEE Systems Journal*, v. 14, n. 2, p. 1886–1896, 2020. Cited on page 21.

[42] HASNAT, M. A.; RAHNAMAY-NAEINI, M. A graph signal processing framework for detecting and locating cyber and physical stresses in smart grids. *IEEE Transactions on Smart Grid*, v. 13, n. 5, p. 3688–3699, 2022. Cited on page 21.

[43] HASNAT, M. A.; RAHNAMAY-NAEINI, M. Characterization and classification of cyber attacks in smart grids using local smoothness of graph signals. In: *2021 North American Power Symposium (NAPS)*. [S.l.: s.n.], 2021. p. 01–06. Cited on page 21.

[44] ANDERSON, O.; YU, N. Detect and identify topology change in power distribution systems using graph signal processing. In: *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*. [S.l.: s.n.], 2021. p. 01–06. Cited on page 21.

[45] OWERKO, D.; GAMA, F.; RIBEIRO, A. Optimal power flow using graph neural networks. In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2020. p. 5930–5934. Cited on page 21.

[46] SINGH, R.; CHAKRABORTY, A.; MANOJ, B. S. Gft centrality: A new node importance measure for complex networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 487, p. 185–195, 12 2017. ISSN 03784371. Cited 8 times at pages 21, 63, 64, 66, 67, 68, 70, and 78.

[47] TSENG, C. C.; LEE, S. L. Graph fourier transform centrality for taipei metro system. In: . [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2020. p. 201–204. ISBN 9781728193625. Cited 4 times at pages 21, 67, 70, and 78.

[48] United States Environmental Protection Agency (EPA). *Distributed Generation of Electricity and its Environmental Impacts*. 2021. Disponível em: <https://www.epa.gov/energy/distributed-generation-electricity-and-its-environmental-impacts>. Cited on page 26.

[49] Agência Nacional de Energia Elétrica (ANEEL). *Resolução Normativa Nº 482*. 2012. Disponível em: <http://www.aneel.gov.br/cedoc/bren2012482.pdf>. Cited on page 26.

[50] PEPERMANS, G. et al. Distributed generation: Definition, benefits and issues. *Energy Policy*, v. 33, n. 6, p. 787–798, 2005. ISSN 03014215. Cited on page 26.

[51] HUBER, M.; DIMKOVA, D.; HAMACHER, T. Integration of wind and solar power in Europe: Assessment of flexibility requirements. *Energy*, v. 69, 2014. Cited on page 26.

[52] ELéTRICA, A. N. de E. *Geração Distribuída*. 2023. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiY2VmMmUwN2QtYWFiOS00ZDE3LWI3NDMtZDk0NGI4MGU2NTkxIiwidCI6IjQwZDZmO
Cited 2 times at pages 18 and 26.

[53] ELETROBRáS. *Proinfa*. 2023. Disponível em: <https://eletrobras.com/en/Paginas/Proinfa.aspx>. Cited on page 26.

[54] AZIZ, T.; KETJOY, N. PV penetration limits in low voltage networks and voltage variations. *IEEE Access*, v. 5, p. 16784–16792, 2017. Cited 3 times at pages 27, 28, and 29.

[55] (IEA), I. E. A. *Solar*. 2021. Disponível em: <https://www.iea.org/fuels-and-technologies/solar>. Cited on page 27.

[56] International Renewable Energy Agency (IRENA). *Solar Energy*. 2021. Disponível em: <https://irena.org/solar>. Cited on page 27.

[57] COLMENAR-SANTOS, A. et al. Technical challenges for the optimum penetration of grid-connected photovoltaic systems: Spain as a case study. *Renewable Energy*, v. 145, p. 2296–2305, 2020. ISSN 0960-1481. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0960148119311383>. Cited on page 27.

[58] KEANE, A. et al. State-of-the-art techniques and challenges ahead for distributed generation planning and optimization. *IEEE Transactions on Power Systems*, v. 28, n. 2, p. 1493–1502, 2013. Cited on page 27.

[59] CHO, N.; YUN, S.; JUNG, J. Determining the reverse fault current by the type of transformer and distributed generation in distribution system during the single-line to ground fault. *Renewable and Sustainable Energy Reviews*, v. 109, p. 102–115, 2019. ISSN 1364-0321. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1364032119302485>. Cited on page 28.

[60] Institute of Electrical and Electronics Engineers. *IEEE 1547 Standard for Interconnecting Distributed Resources with Electric Power Systems*. [S.l.: s.n.], 2003. ISBN 9780738168234. Cited on page 28.

[61] Associação Brasileira de Normas Técnicas. *ABNT NBR 16149 - Sistemas fotovoltaicos (FV) – Características da interface de conexão com a rede elétrica de distribuição*. 2013. Cited on page 28.

[62] CHOU, H. M.; BUTLER-PURRY, K. L. Investigation of voltage stability in unbalanced distribution systems with DG using three-phase current injection based CPF. In: *IEEE Power and Energy Society General Meeting*. [S.l.]: IEEE, 2014. ISBN 9781479964154. ISSN 19449933. Cited on page 28.

[63] STECANELLA, P. A. J. et al. Statistical analysis of photovoltaic distributed generation penetration impacts on a utility containing hundreds of feeders. *IEEE Access*, v. 8, p. 175009–175019, 2020. Cited on page 28.

[64] PANDEY, A. et al. Robust power flow and three-phase power flow analyses. *IEEE Transactions on Power Systems*, v. 34, n. 1, p. 616–626, 2019. Cited on page 30.

[65] GRAINGER, J. J.; Stevenson Jr., W. D. *Power System Analysis*. [S.l.: s.n.], 1994. v. 53. 1689–1699 p. ISSN 0-07-061293-5. ISBN 9788578110796. Cited 5 times at pages 30, 31, 32, 33, and 48.

[66] GLOVER, J. D.; SARMA, M. S.; OVERBYE, T. J. *Power System Analysis and Design*. 5. ed. [S.l.]: Cengage Learning, 2012. ISBN 9781111425777. Cited 3 times at pages 30, 33, and 48.

[67] LI, Z.; YU, J.; WU, Q. H. Approximate linear power flow using logarithmic transform of voltage magnitudes with reactive power and transmission loss consideration. *IEEE Transactions on Power Systems*, v. 33, n. 4, p. 4593–4603, 2018. Cited on page 30.

[68] KERSTING, W. H. *Distribution System Modeling and Analysis*. [S.l.]: CRC Press LLC, 2002. 1-329 p. ISBN 0-8493-0812-7. Cited on page 30.

[69] TONINI, L. G. R.; FERRAZ, R. S. F.; BATISTA, O. E. Load flow and short-circuit methods for grids dominated by inverter-based distributed generation. *Energies*, MDPI, v. 15, 7 2022. ISSN 19961073. Cited on page 30.

[70] TINNEY, W. F.; HART, C. E. Power flow solution by newton's method. *IEEE Transactions on Power Apparatus and Systems*, PAS-86, n. 11, p. 1449–1460, 1967. Cited on page 31.

[71] HENRIQUES, R. M. *Utilização de Autovalores e Autovetores no Problema de Fluxo de Potência para Determinação de Áreas de Controle de Tensão*. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2009. Cited 2 times at pages 31 and 37.

[72] ANTON, H.; RORRES, C. *Álgebra Linear com Aplicações*. 10ª. ed. [S.l.]: Bookman, 2012. Cited 5 times at pages 34, 36, 38, 39, and 40.

[73] STRANG, G. *Introdução à Álgebra Linear*. 4ª. ed. [S.l.: s.n.], 2013. Cited on page 34.

[74] LAY, D. C.; LAY, S. R.; MCDONALD, J. J. *Álgebra linear e suas aplicações*. 5ª. ed. [S.l.]: LTC, 2018. Cited 3 times at pages 34, 40, and 41.

[75] TOUSSAINT, M. *An Analytical Study of the Load Flow Problem*. Tese (Doutorado) — McGill University, 1994. Cited 12 times at pages 34, 40, 41, 42, 43, 45, 47, 48, 49, 50, 51, and 76.

[76] WATKINS, D. S. *Fundamentals of Matrix Computations*. Second edition. [S.l.: s.n.], 2002. ISBN 0-471-21394-2. Cited 5 times at pages 34, 35, 36, 37, and 38.

[77] LIESEN, J.; MEHRMANN, V. *Springer Undergraduate Mathematics Series Linear Algebra*. [s.n.], 2015. ISSN 1615-2085. ISBN 978-3-319-24344-3. Disponível em: <http://www.springer.com/series/3423>. Cited on page 34.

[78] PENG, D.; YI, Z. Dynamics of generalized PCA and MCA learning algorithms. *IEEE Transactions on Neural Networks*, v. 18, n. 6, p. 1777–1784, 2007. ISSN 10459227. Cited on page 36.

[79] GENG, Z.; CHEN, J.; HAN, Y. Energy Efficiency Prediction Based on PCA-FRBF Model: A Case Study of Ethylene Industries. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, IEEE, v. 47, n. 8, p. 1763–1773, 2017. ISSN 21682232. Cited on page 36.

[80] HENDRY, A.; KINNISON, M. *Microevolution Rate, Pattern, Process*. [S.l.]: Springer Science, 2001. ISBN 978-94-010-3889-8. Cited on page 36.

[81] BOLDRINI, J. L. et al. *Algebra Linear*. 3. ed.. ed. São Paulo: [s.n.], 1980. Cited on page 37.

[82] BRANDES, U.; ERLEBACH, T. *Network Analysis - Methodological Foundations*. [S.l.]: Springer, 2005. v. 3663. 482 p. ISSN 03029743. ISBN 3540249796. Cited 6 times at pages 52, 53, 54, 55, 59, and 61.

[83] COELHO, E. P. et al. A complex network analysis of the Brazilian Power Test System. *2015 IEEE PES Innovative Smart Grid Technologies Latin America, ISGT LATAM 2015*, p. 113–118, 2016. Cited 4 times at pages 52, 53, 54, and 59.

[84] ESTRADA, E. *The Structure of Complex Networks: Theory and Applications*. Oxford University Press, 2011. ISBN 9780199591756. Disponível em: <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199591756.001.0001/acprof-9780199591756>. Cited on page 52.

[85] BARABÁSI, A.-L.; PÓSFAI, M. *Network science*. Cambridge: Cambridge University Press, 2016. Cited on page 52.

[86] NETTO, P. O. B.; JURKIEWICZ, S. *Grafos: introdução e prática*. 2. ed. São Paulo: [s.n.], 2017. Cited 2 times at pages 52 and 59.

[87] BENJAMIN, A.; CHARTRAND, G.; ZHANG, P. *The Fascinating World of Graph Theory*. New Jersey: [s.n.], 2015. ISBN 9780691175638. Cited on page 53.

[88] COELHO, E. P. R. *Análise de Vulnerabilidade em Smart Grid Utilizando Métricas de Centralidade em Grafos*. 92 p. Tese (Doutorado) — Federal University of Espírito Santo, 2019. Cited 2 times at pages 53 and 75.

[89] KERSTING, W. H. *Radial distribution test feeders*. 2001. 908–912 p. Disponível em: <https://ewh.ieee.org/soc/pes/dsacom/testfeeders/>. Cited 6 times at pages 58, 74, 81, 82, 104, and 105.

[90] CHEN, G. et al. An improved model for structural vulnerability analysis of power networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 388, n. 19, p. 4259–4266, 2009. ISSN 03784371. Disponível em: <http://dx.doi.org/10.1016/j.physa.2009.06.041>. Cited on page 59.

[91] FREEMAN, L. C. A set of measures of centrality based on betweenness. *Sociometry*, [American Sociological Association, Sage Publications, Inc.], v. 40, n. 1, p. 35–41, 1977. ISSN 00380431. Disponível em: <http://www.jstor.org/stable/3033543>. Cited 2 times at pages 59 and 61.

[92] SEGARRA, S.; RIBEIRO, A. Stability and continuity of centrality measures in weighted graphs. *IEEE Transactions on Signal Processing*, IEEE, v. 64, n. 3, p. 543–555, 2016. ISSN 1053587X. Cited 2 times at pages 59 and 61.

[93] NEWMAN, M. E. A measure of betweenness centrality based on random walks. *Social Networks*, v. 27, n. 1, p. 39–54, 2005. ISSN 03788733. Cited 2 times at pages 61 and 62.

[94] BONACICH, P. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, v. 2, p. 113–120, 1972. Cited on page 61.

[95] PERRA, N.; FORTUNATO, S. Spectral centrality measures in complex networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, v. 78, n. 3, p. 1–11, 2008. ISSN 15393755. Cited on page 62.

[96] SHUMAN, D. I. et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. 10 2012. Disponível em: <http://arxiv.org/abs/1211.0053http://dx.doi.org/10.1109/MSP.2012.2235192>. Cited 6 times at pages 63, 64, 65, 66, 67, and 70.

[97] SARDELLITTI, S.; BARBAROSSA, S.; LORENZO, P. D. On the graph fourier transform for directed graphs. *IEEE Journal on Selected Topics in Signal Processing*, Institute of Electrical and Electronics Engineers Inc., v. 11, p. 796–811, 9 2017. ISSN 19324553. Cited 2 times at pages 63 and 64.

[98] LEVADA, A. L. M.; HADDAD, M. F. C. Entropic laplacian eigenmaps for unsupervised metric learning. In: *2021 34th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. [S.l.: s.n.], 2021. p. 307–314. Cited 2 times at pages 65 and 66.

ANNEX

Infinite continued fraction

An simple example is given to show the resolution of an infinite continued fraction. This method should be replicated for the other infinite continued fractions.

Definition: A real number $\mathbb{R}$ can be represented as a continued fraction $a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cdots}}}$ where a_0 is an integer and the other numbers $a_1, a_2, \ldots, b_1, b_2, b_3 \ldots$ are positive integers.

Example: Calculating the infinite continued fraction $x = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}$

$$x = 1 + \frac{1}{\boxed{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}}$$

The red square in the equation above is then substituted by x

$$x = 1 + \frac{1}{x}$$

$$x^2 = x + 1$$

After this mathematical manipulations, to found the value of x, one just need to calculate a quadratic equation. This procedure was done in chapter 5 (5.12).

$$
\begin{aligned}
{V_2}^1 &= 1 - Z_{12} \cdot {S_2}^* = 1 - \varphi \\
{V_2}^2 &= 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*} = 1 - \frac{\varphi}{1 - \varphi} \\
{V_2}^3 &= 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}} = 1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \varphi}} \\
{V_2}^4 &= 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}}} = 1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \varphi}}} \\
{V_2}^5 &= 1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot \frac{{S_2}^*}{1 - Z_{12} \cdot {S_2}^*}}}} = 1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \frac{\varphi}{1 - \varphi}}}}
\end{aligned}
\tag{7.1}
$$