Impact of Transmission Network Topology on Electrical Power Systems

Thesis by
Linqi Guo

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended May 8, 2019

Linqi Guo
ORCID: 0000-0001-5771-2752

ACKNOWLEDGEMENTS

I would like to express my very deep gratitude and appreciation to Professor Steven Low, who has guided and supported me through numerous tough moments over my PhD years. Steven is the best advisor I could have imagined. As a scholar, he always pursues hard-core research problems that can potentially reshape the field and foundationally revolutionize how the technology moves forward. As a teacher, he is very patient with his students, from the smallest technical details to the biggest picture concerning a research topic, no matter how trivial or how complicated the discussions can be. As a friend, he is willing to talk about and share his insights on anything in life (some of which are not even remotely related to research). He never discourages me from taking an unusual PhD path, yet is always there to back me up when I run into difficult situations. His enthusiasm and personality set an ethical model for all of his students to pursue, and it is my greatest fortune to have worked with him.

I am lucky enough to have also been mentored by Professor Adam Wierman and Dr. Neil Fromer during my studies. They taught me a lot in both research and life, and demonstrated how one can balance them well. This work is significantly improved by the new perspectives they suggested and the broader connections to other fields they noticed. Further, I am very thankful to Adam for his countless help on scientific writing and presentation, and to Neil for sharing his very unique experience to encourage me during one of my toughest moments. I feel indebted for their encouragement, support, and advice that extend way beyond research.

My gratitude also goes to Professor Venkat Chandrasekaran, John Doyle, and Babak Hassibi for serving on my candidacy and final exam committee. Their comments and suggestions greatly improved the quality of this work and I benefited a lot from our discussions. I also would like to thank my collaborators over the years, including: James Anderson, Janusz Bialek, Karl Erliksson, Enrique Mallada, Yorie Nakahira, Ashkan Panahi, John Pang, Shai Vardi, Christos Thrampoulidis, Anwar Walid, Changhong Zhao, and Alessandro Zocca, for all the hard yet fruitful and fun time we spent together solving numerous problems. They showed me how interesting tackling a challenging problem can be and helped me through multiple times when I felt lost.

My colleagues and friends from Netlab and RSRG at Caltech have made this journey very enjoyable. Many thanks to James Anderson, Navid Azizan-Ruhi, Desmond Cai, Niangjun Chen, Yue Chen, Krishnamurthy Dvijotham, Lingwen Gan, Huong Le, Zachary Lee, Tongxin Li, Chen Liang, Enrique Mallada, Yorie Nakahira, John Pang, Qiuyu Peng, Xiaoqi Ren, Yong Sheng So, Jialin Song, Yujie Tang, Shai Vardi, Zhaojian Wang, Lucien Werner, Pengcheng You, Changhong Zhao, Fengyu Zhou, Juba Ziani, Alessandro Zocca and many more, for all the enriching and delightful times together. Among them, I would like to express special thanks to Chen Liang for his constant help in making many simulations of this work possible. Christine Ortega has been an extremely helpful assistant for us, arranging so many things in the best possible way.

Last but not least, I want to thank my parents, Ms. Shujuan Liu and Mr. Jianming Guo, and my elder brother, Mr. Linjun Guo, for their endless and unconditional support. I owe the deepest thanks to my love, Lu Lydia Li, for believing in me more than everyone else, including myself.

ABSTRACT

Power system reliability is a crucial component in the development of sustainable infrastructure. Because of the intricate interactions among power system components, it is often difficult to make general inferences on how the transmission network topology impacts performance of the grid in different scenarios. This complexity poses significant challenges for researches in the modeling, control, and management of power systems.

In this work, we develop a theory that aims to address this challenge from both the fast-timescale and steady state aspects of power grids. Our analysis builds upon the transmission network Laplacian matrix, and reveals new properties of this well-studied concept in spectral graph theory that are specifically tailored to the power system context. A common theme of this work is the representation of certain physical quantities in terms of graphical structures, which allows us to establish algebraic results on power grid performance using purely topological information. This view is particularly powerful and often leads to surprisingly simple characterizations of complicated system behaviors. Depending on the timescale of the underlying problem, our results can be roughly categorized into the study of frequency regulation and the study of cascading failures.

Fast-timescale: Frequency Regulation. We first study how the transmission network impacts power system robustness against disturbances in transient phase. Towards this goal, we develop a framework based on the Laplacian spectrum that captures the interplay among network topology, system inertia, and generator/load damping. This framework shows that the impact of network topology in frequency regulation can be quantified through the network Laplacian eigenvalues, and that such eigenvalues fully determine the grid robustness against low frequency perturbations. Moreover, we can explicitly decompose the frequency signal along scaled Laplacian eigenvectors when damping-inertia ratios are uniform across the buses. The insights revealed by this framework explain why load-side participation in frequency regulation not only makes the system respond faster, but also helps lower the system nadir after a disturbance, providing useful guidelines in the controller design. We simulate an improved controller reverse engineered from our results on the IEEE 39-bus New England interconnection system, and illustrate its robustness against high frequency oscillations compared to both the conventional droop control and a recent controller design.

We then switch to a more combinatorial problem that seeks to characterize the controllability and observability of the power system in frequency regulation if only a subset of buses are equipped with controllers/sensors. Our results show that the controllability/observability of the system depends on two orthogonal conditions: (a) intrinsic structure of the system graph, and (b) algebraic coverage of buses with controllers/sensors. Condition (a) encodes information on graph symmetry and is shown to hold for almost all practical systems. Condition (b) captures how buses interact with each other through the network and can be verified using the eigenvectors of the graph Laplacian matrix. Based on this characterization, the optimal placement of controllers and sensors in the network can be formulated as a set cover problem. We demonstrate how our results identify the critical buses in real systems using a simulation in the IEEE 39-bus New England interconnection test system. In particular, for this testbed a single well chosen bus is capable of providing full controllability and observability.

Steady State: Cascading Failures. Cascading failures in power systems exhibit non-monotonic, non-local propagation patterns which make the analysis and mitigation of failures difficult. By studying the transmission network Laplacian matrix, we reveal two useful structures that make the analysis of this complex evolution more tractable: (a) In contrast to the lack of monotonicity in the physical system, there is a rich collection of monotonicity we can explore in the spectrum of the Laplacian matrix. This allows us to systematically design topological measures that are monotonic over the cascading event. (b) Power redistribution patterns are closely related to the distribution of different types of trees in the power network topology. Such graphical interpretation captures the Kirchhoff's Law in a precise way and naturally suggests that we can eliminate long-distance propagation of system disturbances by forming a tree-partition.

We then show that the tree-partition of transmission networks provides a precise analytical characterization of line failure localizability. Specifically, when a non-bridge line is tripped, the impact of this failure only propagates within well-defined components, which we refer to as cells, of the tree-partition defined by the bridges. In contrast, when a bridge line is tripped, the impact of this failure propagates globally across the network, affecting the power flow on all remaining transmission lines. This characterization suggests that it is possible to improve the system robustness by switching off certain transmission lines, so as to create more, smaller components in the tree-partition; thus spatially localizing line failures and making the grid less vulnerable to large-scale outages. We illustrate this approach using the IEEE 118-bus test system and demonstrate that switching off a negligible portion of transmission lines allows the impact of line failures to be significantly more localized without substantial changes in line congestion.

Unified Controller on Tree-partitions. Combining our results from both the fast-timescale and steady state behaviors of power grids, we propose a distributed control strategy that offers strong guarantees in both the mitigation and localization of cascading failures in power systems. This control strategy leverages a new controller design known as Unified Controller (UC) from frequency regulation literature, and revolves around the powerful properties that emerge when the management areas that UC operates over form a tree-partition. After an initial failure, the proposed strategy always prevents successive failures from happening, and regulates the system to the desired steady state where the impact of initial failures are localized as much as possible. For extreme failures that cannot be localized, the proposed framework has a configurable design that progressively involves and coordinates across more control areas for failure mitigation and, as a last resort, imposes minimal load shedding. We compare the proposed control framework with the classical Automatic Generation Control (AGC) on the IEEE 118-bus test system. Simulation results show that our novel control greatly improves the system robustness in terms of the $N-1$ security standard, and localizes the impact of initial failures in majority of the load profiles that are examined. Moreover, the proposed framework incurs significantly less load loss, if any, compared to AGC, in all of our case studies.

PUBLISHED CONTENT AND CONTRIBUTIONS

L. Guo, C. Liang, A. Zocca, S. Low, A. Wierman, "Less is More: Real-time Failure Localization in Power Systems", In: *arXiv preprint arXiv:1904.05461*, 2019. URL: `https://arxiv.org/abs/1904.05461`

L. Guo, C. Zhao, S. Low, "Cyber Network Design for Secondary Frequency Regulation: A Spectral Approach", In: *Power Systems Computation Conference (PSCC)*, Dublin, Ireland, 2018. DOI: 10.23919/PSCC.2018.8442814.

L. Guo, C. Zhao, S. Low, "Graph Laplacian Spectrum and Primary Frequency Regulation", In: *Decision and Control (CDC), IEEE 57th Conference on*, Miami, FL, 2018. DOI: 10.1109/CDC.2018.8619252.

L. Guo, C. Liang, A. Zocca, S. Low, A. Wierman, "Failure Localization in Power Systems via Tree Partitions", In: *Decision and Control (CDC), IEEE 57th Conference on*, Miami, FL, 2018. DOI: 10.1109/CDC.2018.8619562.

L. Guo, C. Liang, S. Low, "Monotonicity Properties and Spectral Characterization of Power Redistribution in Cascading Failures", In: *55th Annual Allerton Conference on Communications, Control, and Computing (Allerton)*, Allerton, IL, 2017. DOI: 10.1109/ALLERTON.2017.8262836.

L. Guo, S. Low, "Spectral Characterization of Controllability and Observability for Frequency Regulation Dynamics", In: *Decision and Control (CDC), IEEE 56th Conference on*, Melbourne, Australia, 2017. DOI: 10.1109/CDC.2017.8264611.

J. Pang, L. Guo, S. Low, "Optimal Load Control for Frequency Regulation under Limited Control Coverage", In: *SIGMETRICS Performance Evaluation Review*, New York, NY, 2017. DOI: 10.1145/3152042.3152071.

L. Guo established the foundational theory for this series of work, developed numerous applications of the spectral representation of power system, prepared and examined the simulations, and participated in the writing of the manuscripts.

TABLE OF CONTENTS

Acknowledgements iii
Abstract v
Published Content and Contributions viii
Table of Contents ix
List of Illustrations xi
List of Tables xiii
Chapter I: Introduction 1
- 1.1 Fast-Timescale: Frequency Regulation 1
- 1.2 Steady State: Cascading Failures 5
- 1.3 Unified Controller over Tree-partitions: Guaranteed Outage Mitigation 7

Chapter II: Preliminaries: Network Topology and Power Flows 9
- 2.1 Power Grid and Swing Dynamics 9
- 2.2 DC Power Flow 12
- 2.3 Laplacian Matrix 13
- 2.4 Graphical Interpretation of the Laplacian Inverse 14
- 2.5 Proofs 20

Chapter III: Spectral Decomposition and Frequency Regulation 23
- 3.1 System Model 24
- 3.2 Characterization of System Response 26
- 3.3 Interpretations 29
- 3.4 Controller Design for Load-Side Participation 34
- 3.5 Case Studies 35
- 3.6 Conclusion 38
- 3.7 Proofs 39

Chapter IV: Controllability and Observability under Limited Controller and Sensor Coverage 43
- 4.1 System Model 43
- 4.2 Controllability 45
- 4.3 Interpretations 48
- 4.4 Observability 50
- 4.5 Applications 51
- 4.6 Conclusion 53
- 4.7 Proofs 53

Chapter V: Monotonicity in Cascading Failures and Tree-partitions 59
- 5.1 System Model 60
- 5.2 Monotonicity in Cascading Failures 61
- 5.3 Line Outage Redistribution Factor 64
- 5.4 Tree-partitions of Power Grids 67
- 5.5 Guaranteed Localization 69

5.6 Conclusion 72
5.7 Proofs 72
Chapter VI: Failure Localization via Tree-partitions 75
6.1 Single Line Failure 76
6.2 Non-Bridge Failures are Localizable 77
6.3 Bridge Failures Propagate 81
6.4 Generalization to Multi-line Failure 85
6.5 Case Studies 89
6.6 Conclusion 92
6.7 Proofs 92
Chapter VII: Real-time Outage Mitigation 102
7.1 System Model 103
7.2 Connecting UC to Tree-partition 106
7.3 Proposed Control Strategy 107
7.4 Localizing Non-critical Failures 110
7.5 Controlling Critical Failures 112
7.6 Case Studies 114
7.7 Conclusion 117
7.8 Proofs 117
Bibliography 126

LIST OF ILLUSTRATIONS

Number *Page*

1.1 An example illustrating how a micro-grid is connected to the main grid. 3

1.2 An example network with symmetry. 4

1.3 The sequence of events, indexed by the circled numbers, that led to the Western US blackout in 1996. Adopted from [35]. 5

2.1 An example element in $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$, where circles correspond to elements in $\mathcal{N}_1$ and squares correspond to elements in $\mathcal{N}_2$. The two spanning trees containing $\mathcal{N}_1$ and $\mathcal{N}_2$ are highlighted as solid lines. . 15

2.2 An example element in $\mathcal{T}(\{i, w\}, \{j, z\})$. The spanning trees containing $\{i, w\}$ and $\{j, z\}$ are highlighted as solid lines. 18

3.1 Illustration of the gain $|\overline{\omega}_{i,j}(\mathrm{j}\sigma)|$ as a function of σ. 32

3.2 Line diagram of the IEEE 39-bus interconnection testbed. 36

3.3 Frequency trajectory at bus 30 when we add white Gaussian measurement noise of −20 dBW. 37

3.4 Frequency trajectory at bus 30 when we add a signal following the sine curve $0.2\sin(10\pi t)$ p.u. 37

3.5 Frequency trajectory at bus 36 under wind power output at bus 30. . . 38

4.1 Comparison of the system evolution with and without control at bus 35 after adding a step increase of 1 pu to the generation at bus 30. . . 53

5.1 A ring network with clockwise orientation. Edge e_1 can only spread "negative" impacts to other lines. 65

5.2 The construction of $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{P}$. 67

5.3 An illustration of the partial order $\geq$ over tree-partitions. The partition $\mathcal{P}^1 = \left\{\mathcal{N}_1^1, \mathcal{N}_2^2, \mathcal{N}_3^1, \mathcal{N}_4^1\right\}$ is finer than $\mathcal{P}^2 = \left\{\mathcal{N}_1^2, \mathcal{N}_2^2\right\}$. 69

5.4 (a) A double-ring network. G is the generator bus and L is the load bus. Arrows represent the original power flow. (b) The new network after removing an edge. Arrows represent the new power flow. 70

6.1 Non-zero entries of the $K_{e\hat{e}}$ matrix (as represented by the dark blocks) for a graph with tree-partition $\{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ and bridge set $\mathcal{E}_b$. The small blocks represent cells inside the regions. 76

6.2 (a) A butterfly network. (b) The block decomposition of the butterfly network into cells $\mathcal{C}_1$ and $\mathcal{C}_2$. 79

6.3 Influence graphs on the IEEE 118-bus network before and after switching off lines e_1, e_2 and e_3. Blue edges represent physical transmission lines and grey edges represent connections in the influence graph. 90

6.4 (a) Histogram of the normalized branch flow changes. (b) Cumulative distribution function of the positive normalized branch flow changes. Note that the curve intercepts the y-axis since 52.59% of the branch flows decrease. 91

6.5 The localized graph $\mathcal{G}_{\mathcal{N}_1}$. $\mathcal{N}_t$ is the imaginary bus containing e and $\mathcal{N}_{1,i}$'s are remaining imaginary buses. The power adjustments from the power balance rule $\mathcal{R}$ are shown near each participating bus in reaction to a power loss of M. 93

7.1 An illustration of the failure propagation model. 103

7.2 Flowchart of the events after an initial failure under the proposed control strategy. 108

7.3 One line diagram of the IEEE 118-bus test system with two control areas. Dashed blue lines are switched off when a tree-partition needs to be formed. 115

7.4 Number of vulnerable lines with respect to different levels of system congestion. 116

7.5 CCDF for load loss rate. 117

7.6 CCDF for generator response. 117

LIST OF TABLES

Number *Page*

2.1 Variables associated with buses and transmission lines in swing dynamics. 11

3.1 System performance in terms of network Laplacian eigenvalues, generator inertia, and damping ($\gamma := \delta/\mu$, $\Delta_i := |\gamma - 4\lambda_i|$). 33

Chapter 1

INTRODUCTION

Power system reliability is a crucial component in the development of sustainable infrastructure. Recent blackouts, especially the 2003 and 2012 blackouts in Northwestern U.S. [1] and India [2], demonstrated the devastating economic impact a grid failure can cause. In even worse cases, such as where facilities like hospitals are involved, blackouts pose direct threat to people's health and lives. Because of the intricate interactions among power system components, it is often difficult to make general inferences on how the transmission network topology impacts the robustness of power grids in different scenarios. Component outages, for instance, may cascade and propagate in a very complicated, non-local manner [9, 27, 35], exhibiting distinct patterns for different networks [62]. This complexity poses significant challenges for researches in the modeling, control, and management of power systems.

In this work, we develop a theory that aims to address this challenge from both the fast-timescale and steady state aspects of power grids. Our analysis builds upon the transmission network Laplacian matrix, and reveals new properties of this well-studied concept in spectral graph theory that are specifically tailored to the power system context. A common theme of this work is the representation of certain physical quantities in terms of graphical structures, which allows us to establish algebraic results on power grid performance using purely topological information. This view is particularly powerful and often leads to surprisingly simple characterizations of complicated system behaviors. Depending on the timescale of the underlying problem, our results can be roughly categorized into the study of frequency regulation and the study of cascading failures.

1.1 Fast-Timescale: Frequency Regulation

Frequency regulation balances the power generation and consumption in an electrical power network. Such control is governed by the swing dynamics and is traditionally implemented in generators through droop control, Automatic Generation Control, and Economic Dispatch [82, 83]. It has been widely realized that the increasing level of renewable penetration makes it harder to stabilize the system due to higher generation volatility and lower aggregate inertia. One popular approach to maintaining system stability in this new era is to integrate load-side participation

[4, 15, 17, 30, 38, 47, 48, 55, 61, 70], which not only helps stabilize the system in a more responsive and scalable fashion, but also improves the system transient behavior [81–83].

Related Work

With the goal of fully harnessing the benefits of load-side controllers, there has emerged a large body of work devoted to understanding how different system parameters and controller designs impact the grid transient performance. For instance, iDroop is proposed in [42] to improve dynamic performance of the power system through controlling power electronics or loads. Such controllers, however, can sometimes make the power system dynamics more sophisticated and uncertain, and hence make it harder to obtain a stability guarantee [53]. In [56], methods to determine the optimal placement of virtual inertia in power grids to accommodate loss of system stability are proposed and studied. There has also been work on characterizing the synchronization cost of the swing dynamics [21, 28, 51, 54, 68] that explicitly computes the response $\mathcal{H}_2$ norm in terms of system damping, inertia, resistive loss, line failures, etc. In certain cases, classical metrics studied in power engineering, such as nadir and maximum rate of change of frequency, can also be analytically derived [51].

Compared to the aforementioned system parameters, the role of transmission network topology on the transient stability of swing dynamics is less well understood. Indeed, it is usually hard to infer how a change to the network topology affects overall grid behavior and performance without detailed simulation and computation. For example, one can argue that the connectivity in the grid helps average the power demand imbalance over the network, and therefore higher connectivity should enhance system stability. On the other hand, one can also argue that higher connectivity means faster propagation of disturbances over the network, which should therefore decrease system stability. Both arguments seem plausible but they lead to (apparently) opposite conclusions (a corollary of our results in Section 3.2 will clarify this paradox). In fact, even the notion "connectivity" itself seems vague and is open to different interpretations.

Robustness against Disturbances

Our first goal is to clarify how exactly the transmission network topology is related to the system robustness against disturbances in transient phase. Such relation is often subtle and the insights developed from different types of applications do not always

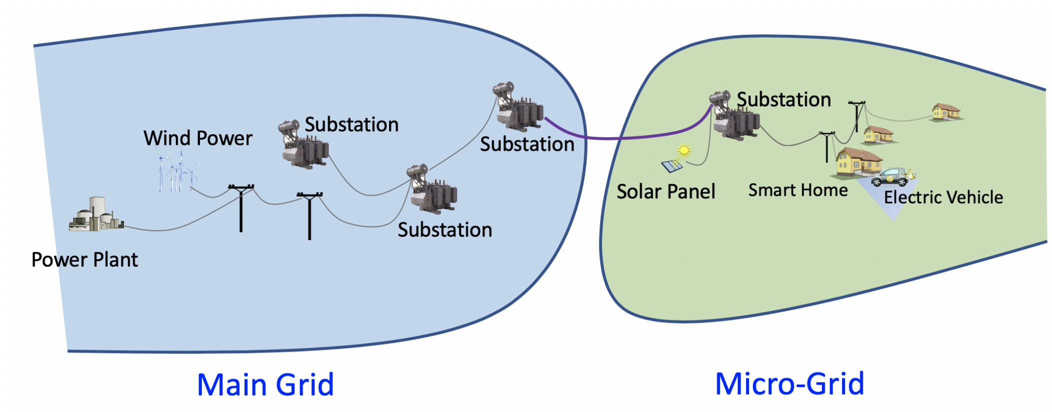


Figure 1.1: An example illustrating how a micro-grid is connected to the main grid.

agree with each other. For instance, having redundant transmission lines is believed to be a crucial part in maintaining the $N-1$ security of power grids [8, 12, 34], since such lines allow the power to flow through alternative paths if certain components are tripped from the system. In contrast, micro-grids are designed to be connected to major power systems via a single connection, as shown in Figure 1.1, which is considered to be helpful in isolating the micro-grid from the disruptions that occur in the main grid. It is natural to ask whether there is a way to reconcile the insights we gain from both scenarios and devise a common principle that applies in all cases.

Towards this goal, we develop a framework based on the transmission network Laplacian spectrum that captures the interplay among network topology, system inertia, and generator/load damping. It shows that the impact of network topology in frequency regulation can be quantified through the network Laplacian eigenvalues, and that such eigenvalues fully determine the grid robustness against low frequency perturbations. Moreover, we can explicitly decompose the frequency signal along scaled Laplacian eigenvectors when damping-inertia ratios are uniform across the buses. The insights revealed by this framework explain why load-side participation in frequency regulation not only makes the system respond faster, but also helps lower the system nadir after a perturbing event, providing useful guidelines in the controller design. We simulate an improved controller reverse engineered from our results on the IEEE 39-bus New England interconnection system, and illustrate its robustness against high frequency oscillations compared to both the conventional droop control and a recent controller design.

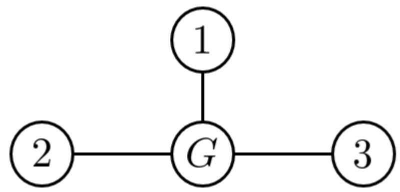


Figure 1.2: An example network with symmetry.

Controllability under Limited Controller Coverage

A second question we are interested in answering pertains the controllability and observability of power systems in frequency regulation when only a subset of buses are equipped with controllers and sensors. Indeed, it is almost always assumed in recent frequency regulation literature [25, 36, 37, 43, 65, 72, 73, 81, 82] that every bus in the transmission network comes with a controllable injection and proper sensors, which is not realistic for structure preserving models in power systems. For example, feasible placements of controllable loads such as electric vehicle charging stations [79] and aggregated households [70] are typically limited to certain geographical areas, and the penetration of such controllable loads takes investment and time. Moreover, the enormous amount of sensing devices needed for full load-side participation in a large scale network can be very costly [50, 79].

When only a subset of buses are controllable, it is less clear how much controllability we have over the system. As a motivating example, let us consider the highly symmetric network shown in Figure 1.2, where nodes 1, 2, and 3 are assumed to have the same load, damping, inertia, initial phase etc., and assume we can only control the node with label G. Then because of symmetry, no matter how we alter the mechanical power injection at node G, the power flows on all transmission lines would be the same, and therefore the system cannot be controllable. This, of course, is a highly contrived example; nevertheless, it is a manifestation of an intrinsic network property that leads to a loss of controllability.

More specifically, we show that the controllability/observability of the system depends on two orthogonal conditions: (a) intrinsic structure of the system graph, and (b) algebraic coverage of buses with controllers/sensors. Condition (a) encodes information on graph symmetry and is shown to hold for almost all practical systems. Condition (b) captures how buses interact with each other through the network and can be verified using the eigenvectors of the graph Laplacian matrix. Based on this characterization, the optimal placement of controllers and sensors in the network

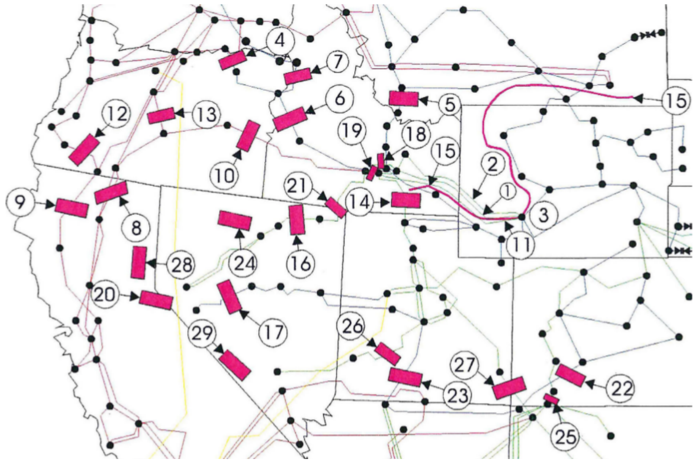


Figure 1.3: The sequence of events, indexed by the circled numbers, that led to the Western US blackout in 1996. Adopted from [35].

can be formulated as a set cover problem. We demonstrate how our results identify the critical buses in real systems using a simulation in the IEEE 39-bus New England interconnection test system. In particular, for this testbed, a single well chosen bus is capable of providing full controllability and observability.

1.2 Steady State: Cascading Failures

Cascading failures in power systems propagate non-locally, making their analysis and mitigation difficult. This fact is illustrated by the sequence of events leading to the 1996 Western US blackout (summarized in Figure 1.3) [35], in which successive failures happened hundreds of kilometers away from each other (e.g., from stage ③ to stage ④ and from stage ⑦ to stage ⑧). Non-local propagation makes it particularly challenging to design distributed controllers that reliably prevent and mitigate cascades in power systems. In fact, such control is widely considered impossible, even when centralized coordination is available [12, 34].

Related Work

Current industry practice for mitigating cascading failures mostly relies on simulation-based contingency analysis, which focuses on a small set of most likely initial failures [7]. Moreover, the size of the contingency set which is tested (and thus the level of security guaranteed) is often constrained by computational power, undermining its effectiveness in view of the enormous number of components in power networks. After a blackout event, a detailed study typically leads to a redesign of such contingency sets, potentially together with physical network upgrades and revision of system management policies and regulations [34].

The limitations of current practice have motivated a large body of literature to study and characterize analytical properties of cascading failures in power systems. This literature can be roughly categorized as follows: (a) applying Monte-Carlo methods to analytical models that account for the steady state power redistribution using DC [5, 9, 18, 78] or AC [49, 58, 63] flow models; (b) studying pure topological models built upon simplifying assumptions on the propagation dynamics (e.g., failures propagate to adjacent lines with high probability) and inferring component failure propagation patterns from graph-theoretic properties [16, 23, 39]; and (c) investigating simplified or statistical cascading failure dynamics [26, 35, 57, 74].

In each of these approaches, it is often difficult to make general inferences about failure patterns. For example, power flow over a specific branch can increase, decrease, and even reverse direction as cascading failure unfolds [45]. The failure of a line can cause another line that is arbitrarily far away to be tripped [9]. Load shedding instead of mitigating the cascading failure, can actually increase the congestion on certain lines [13].

Failure Localization

Our first goal in this context is to devise structural properties of the cascading process. By studying the transmission network Laplacian matrix, we reveal two useful structures that make the analysis of this complex evolution more tractable: (a) In contrast to the lack of monotonicity in the physical system, there is a rich collection of monotonicity we can explore in the spectrum of the Laplacian matrix. This allows us to systematically design topological measures that are monotonic over the cascading event. (b) Power redistribution patterns are closely related to the distribution of different types of trees in the power network topology. Such graphical interpretation captures the Kirchhoff's Law in a precise way, and naturally suggests that we can eliminate long-distance propagation of system disturbances by forming a tree-partition.

We then show that the tree-partition of transmission networks provides a precise analytical characterization of line failure localizability. Specifically, when a non-bridge line is tripped, the impact of this failure only propagates within well-defined components, which we refer to as cells, of the tree-partition defined by the bridges. In contrast, when a bridge line is tripped, the impact of this failure propagates globally across the network, affecting the power flow on all remaining transmission lines.

This characterization suggests that it is possible to improve the system robustness by switching off certain transmission lines, so as to create more, smaller components in the tree-partition; thus spatially localizing line failures and making the grid less vulnerable to large-scale outages. We illustrate this approach using the IEEE 118-bus test system and demonstrate that switching off a negligible portion of transmission lines allows the impact of line failures to be significantly more localized without substantial changes in line congestion.

1.3 Unified Controller over Tree-partitions: Guaranteed Outage Mitigation

Despite all its promising properties, tree-partition alone does not yield a fully satisfactory solution for mitigating and localizing failures, due to two main reasons: First, reducing redundancy to create a tree-partition may lead to single-point vulnerabilities, the failure of which has a global impact on the whole system and can potentially cause significant load loss. Second, information on unfolding cascading failures is not fed back into relevant controllers that could adjust the network topology (and in particular its tree-partition). Therefore, after an initial failure is triggered, the tree-partition based strategy only guarantees that any successive failure will occur in the same region as the initial failure, but does not prevent or stop successive failures. To overcome these drawbacks, there is a need for new control designs that can "close the loop" and respond actively and promptly to different failures.

Towards this goal, we combine our results from both the fast-timescale and steady state behaviors of power grids and propose a distributed control strategy that offers strong guarantees in both the mitigation and localization of cascading failures in power systems. This control strategy leverages a new controller design known as Unified Controller (UC) from frequency regulation literature, and revolves around the powerful properties that emerge when the management areas that UC operates over form a tree-partition. After an initial failure, the proposed strategy always prevents successive failures from happening, and regulates the system to the desired steady state where the impact of initial failures are localized as much as possible. For extreme failures that cannot be localized, the proposed framework has a configurable design that progressively involves and coordinates across more control areas for failure mitigation and, as a last resort, imposes minimal load shedding. We compare the proposed control framework with the classical Automatic Generation Control (AGC) on the IEEE 118-bus test system. Simulation results show that our novel control greatly improves the system robustness in terms of the $N-1$ security standard, and localizes the impact of initial failures in majority of the load profiles that are examined. Moreover, the proposed framework incurs significantly less load loss, if any, compared to AGC, in all of our case studies.

Chapter 2

PRELIMINARIES: NETWORK TOPOLOGY AND POWER FLOWS

In this chapter, we present the main power grid model considered throughout the thesis and establish its basic properties that will be used in later chapters. We also explain how our model is related to the Laplacian matrix of the transmission network and derive a new representation of DC power flow equations in terms of graph structures. In different applications, our model often needs to be augmented properly by adding elements specific to the problem under study, and such augmentation will be presented in the relevant chapters.

2.1 Power Grid and Swing Dynamics

Our studies revolve around both the fast-timescale aspect and the steady state behaviors of power grids. Although the system model takes different forms when we focus on different problems, these formulations are inherently connected and represent the same physical subject from different angles. In this section, we introduce the linearized swing dynamics which describe the power system behavior in fast-timescale, and explain the frequency synchronization structure of its equilibrium points.

We use the graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ to model the power transmission network, where $\mathcal{N} = \{1, \ldots, n\}$ is the set of buses and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ denotes the set of transmission lines. The terms bus/node/vertex and link/line/edge are used interchangeably. An edge in $\mathcal{E}$ is denoted either as e or (i, j). We further assign an arbitrary orientation over $\mathcal{E}$ so that if $(i, j) \in \mathcal{E}$ then $(j, i) \notin \mathcal{E}$. Let $n = |\mathcal{N}|, m = |\mathcal{E}|$ be the number of buses and transmission lines, respectively. The (node-edge) incidence matrix of $\mathcal{G}$ is a $n \times m$ matrix C defined as

$$C_{ie} = \begin{cases} 1 & \text{if node } i \text{ is the source of } e \\ -1 & \text{if node } i \text{ is the target of } e \\ 0 & \text{otherwise.} \end{cases}$$

The swing dynamics of power grids operate on the order of seconds and capture how generators react to disturbances by adjusting their rotating frequencies. The main control goal for a power system in this timescale is to minimize the system deviation from the last specified operating point. Such operating points often come from slow-timescale (on the order of ten minutes to an hour) optimization known as Economic Dispatch [8], and can be considered as unchanged when we focus on fast-timescale dynamics. It is thus natural to consider the system *deviations* when describing the system behaviors in this timescale. We take this approach here and refer interested readers to classical literatures (such as [8, 40]) for more details on other forms of swing dynamics and discussions therein.

We now describe the relevant physical quantities. For each bus $j \in \mathcal{N}$, denote its bus voltage phase angle deviation as θ_j, its frequency deviation as ω_j and its total injection deviation as p_j. The deviation p_j consists of a controllable part, denoted as d_j, and a uncontrollable part, denoted as r_j. The controllable part d_j represents mechanical power injection adjustment if j is a generator bus, and represents the aggregate change in controllable load if j is a load bus. Depending on the specific controller design, the value of d_j is often limited to an interval $[\underline{d}_j, \overline{d}_j]$ because of constraints like generator ratings or ramping constraints. The uncontrollable part r_j captures consumer load change or failure of infrastructure devices such as generator units. The inertia and damping constants of the bus j are denoted as $M_j > 0$ and $D_j \geq 0$ respectively. If j is a load bus, the sensitivity constant for frequency-sensitive load is also included in D_j (see [82]). For each transmission line $e \in \mathcal{E}$, we use f_e to denote its branch flow deviation and denote the line susceptance as B_e. In practical operations, the flow deviation f_e is also constrained to an interval $[\underline{f}_e, \overline{f}_e]$ because of line ratings. When f_e falls out of this interval, the line e overheats,, which over time can lead to line failure if no proper actions are taken.

With these notations, which are summarized in Table 2.1, the linearized swing dynamics are given by:

$$M_j\dot{\omega}_j = r_j - d_j - D_j\omega_j - \sum_{e\in\mathcal{E}} C_{je} f_e, \quad j \in \mathcal{N} \tag{2.1a}$$

$$\dot{f}_{ij} = B_{ij}(\omega_i - \omega_j), \qquad (i, j) \in \mathcal{E}. \tag{2.1b}$$

When the bus phase angles are relevant, we often extend (2.1) by adding

$$\dot{\theta}_j = \omega_j, \quad j \in \mathcal{N}. \tag{2.2}$$

The dynamics (2.1a) can be interpreted as the counterpart of Newton's Second Law in power systems, which says the generator rotating acceleration at bus j is proportional to the sum of power imbalance at j. The dynamics (2.1b) are essentially

$\theta := (\theta_j, j \in \mathcal{N})$	bus voltage angle deviations
$\omega := (\omega_j, j \in \mathcal{N})$	bus frequency deviations
$p := (p_j, j \in \mathcal{N})$	total injection deviations
$r := (r_j, j \in \mathcal{N})$	uncontrollable injection deviations
$d := (d_j, j \in \mathcal{N})$	mechanical power injection adjustment for generator buses; controllable load adjustment for load buses
$\overline{d}_j, \underline{d}_j, j \in \mathcal{N}$	upper and lower limits for the adjustable injection d_j
$D_j\omega_j, j \in \mathcal{N}$	deviation of frequency-sensitive injections
$f := (f_e, e \in \mathcal{E})$	branch flow deviations
$\overline{f}_e, \underline{f}_e, e \in \mathcal{E}$	upper and lower limits for branch flow deviations
$C \in \mathbb{R}^{\|\mathcal{N}\|\times\|\mathcal{E}\|}$	incidence matrix of $\mathcal{G}$: $C_{je} = 1$ if j is the source of e, $C_{je} = -1$ if j is the destination of e, and $C_{je} = 0$ otherwise
$B := \text{diag}(B_e, e \in \mathcal{E})$	branch flow linearization coefficients that depend on nominal state voltage magnitudes and reference phase angles

Table 2.1: Variables associated with buses and transmission lines in swing dynamics.

the linearized Ohm's Law applied to complex power phasors:

$$f_{ij} = B_{ij}(\theta_i - \theta_j), \quad (i, j) \in \mathcal{E}. \tag{2.3}$$

See [81, 82] for a more detailed justification and derivation of this dynamics.

Definition 2.1.1. *A state* $x^* := (\theta^*, \omega^*, d^*, f^*) \in \mathbb{R}^{3n+m}$ *is said to be an* **equilibrium** *of* (2.1) *(and* (2.2) *when phase angles are relevant) if the right hand sides of* (2.1) *are zero at* x^*.

Note that the x^* here contains θ^*, which only shows up in (2.2) but not in (2.1). The main reason for separating (2.2) from (2.1) is that the phase angles are usually of less concern in many problems we study so our results often pertain to the dynamics (2.1) only. Yet occasionally we also need to refer to the bus phase angles in the discussions, and thus we include θ^* as part of x^*. We emphasize that the right hand sides of (2.2) are not required to be zero at x^*, since the system (2.1) can converge to a state where the frequency deviations are nonzero.

Remark 2.1.2. *The d_j's in* (2.1a) *usually depend on the system states and may evolve by themselves in accordance with certain controller specific dynamics. The equilibrium defined above refers to the* closed-loop *equilibrium. It is thus possible to engineer the equilibrium of* (2.1) *by adopting different controller designs for d_j, which in turn can impact various system performance such as robustness against disturbances (see Chapter 3); controllability and observability (see Chapter 4); failure localization (see Chapter 6); and outage mitigation (see Chapter 7), etc.*

By (2.1b), at an equilibrium $x^* = (\theta^*, \omega^*, d^*, f^*)$ of (2.1), the frequencies synchronize among buses that are connected to each other:

$$\omega_i^* = \omega_j^*, \qquad (i, j) \in \mathcal{E}.$$

Therefore, as long as the graph $\mathcal{G}$ is connected, the whole grid attains a synchronized frequency, denoted as $\overline{\omega}^*$, across all buses. We can then sum (2.1a) over all buses $j \in \mathcal{N}$ to obtain

$$\begin{aligned} 0 &= \sum_{j\in\mathcal{N}} \left(r_j - d_j^* - D_j \overline{\omega}^* - \sum_{e\in\mathcal{E}} C_{je} f_e^* \right) \\ &= \sum_{j\in\mathcal{N}} \left(r_j - d_j^* \right) - \overline{\omega}^* \sum_{j\in\mathcal{N}} D_j - \sum_{e\in\mathcal{E}} f_e \sum_{j\in\mathcal{N}} C_{je} \\ &= \sum_{j\in\mathcal{N}} \left(r_j - d_j^* \right) - \overline{\omega}^* \sum_{j\in\mathcal{N}} D_j, \end{aligned}$$

where the last equation is because for $e = (i, j)$, $C_{ie} = 1$, $C_{je} = -1$ and $C_{ke} = 0$ for all other k. As a result, we know

$$\overline{\omega}^* = \frac{\sum_{j\in\mathcal{N}} \left(r_j - d_j^* \right)}{\sum_{j\in\mathcal{N}} D_j}. \tag{2.4}$$

In other words, at any equilibrium of (2.1), the synchronized frequency is proportional to the aggregate injection imbalance.

2.2 DC Power Flow

Practical power grids are usually equipped with frequency regulation controllers (such as the conventional Automatic Generation Control [8]) that aim to balance the power generation and consumption. For such systems, when the dynamics (2.1) converges to an equilibrium $x^* = (\theta^*, \omega^*, d^*, f^*)$, the aggregate injection imbalance $\sum_j (r_j - d_j^*)$ is zero. By (2.4), we then know $\overline{\omega}^* = 0$. This plugged into (2.1a) together with the linearized Ohm's law (2.3) implies that

$$\begin{cases} p_j^* = \sum_{e\in\mathcal{E}} C_{je} f_e^*, & j \in \mathcal{N} \\ f_{ij}^* = B_{ij}\left(\theta_i^* - \theta_j^*\right), & (i,j) \in \mathcal{E}, \end{cases}$$

where $p_j^* = r_j - d_j^*$. Dropping the stars from our notations and rewriting the above equation in matrix form, we then obtain the DC power flow equations:

$$\begin{cases} p = Cf \\ f = BC^T\theta. \end{cases} \tag{2.5}$$

The DC power flow equations determines steady state branch flows and phase angles from the injection p. Given a fixed p, the solutions to (2.5) are unique in f but generally not unique in θ, since the phase angles θ can be shifted by a constant value in each connected component of $\mathcal{G}$.

2.3 Laplacian Matrix

The DC power flow equations (2.5) imply that

$$p = CBC^T\theta.$$

That is, the steady state phase angle θ is related to the steady state injection p via the matrix CBC^T. This matrix is known as the Laplacian matrix of $\mathcal{G}$ and plays a central role in all of our analysis in later chapters. We denote $L = CBC^T$ and present below some of its basic properties.

For any $v \in \mathbb{R}^n$, it is easy to see that

$$v^T L v = \sum_{(i,j)\in\mathcal{E}} B_{ij}(v_i - v_j)^2 \geq 0. \tag{2.6}$$

This leads to the following well-known result:

Lemma 2.3.1. *L is positive semidefinite and hence diagonalizable.*

Moreover, equation (2.6) attains equality if and only if $v_i = v_j$ for any $(i,j) \in \mathcal{E}$. Thus the eigenspace of L corresponding to 0, which is equivalently the kernel of L, consists of vectors that take the same value on each of the connected component of $\mathcal{G}$. In particular, we can recover the following well-known result:

Lemma 2.3.2. *If $\mathcal{G}$ is connected, then the kernel of L is* $\operatorname{span}(\boldsymbol{1})$*, the set of vectors with uniform entries.*

This result tells us that the Laplacian matrix L for a connected graph has rank $n-1$, and hence contains a submatrix of size $(n-1)\times(n-1)$ that is invertible. A less obvious result, known as the Kirchhoff's Matrix Tree Theorem, states that the matrix obtained by removing the last row and the last column from L, which we denote as $\overline{L}$, is such a matrix (see [19] for more details):

Proposition 2.3.3. *For a connected graph $\mathcal{G}$, the determinant of $\overline{L}$ is given by*

$$\det(\overline{L}) = \sum_{E\in\mathcal{T}} \prod_{e\in E} B_e,$$

where $\mathcal{T}$ is the set of spanning trees on $\mathcal{G}$.

In particular, since the set of spanning trees $\mathcal{T}$ is non-empty for a connected graph and $B_e > 0$ for all e, we know that $\det(\overline{L}) \neq 0$ and hence $\overline{L}$ is invertible. The matrix $\overline{L}$ can be interpreted as the part of L after removing a reference bus (often known as the slack bus) from the system.

Define $A := \left(\overline{L}\right)^{-1}$. From the definition of A, it is not surprising that the elements of the matrix A encode information on the topology of $\mathcal{G}$ and thus carry graphical meanings. As we show in Section 2.4, all elements of A are in fact closely related to the tree distributions of $\mathcal{G}$, and suggest how the DC power flow equations (2.5) can be represented via the graph structure of $\mathcal{G}$.

Remark 2.3.4. *In certain applications, it is natural to assign weights to buses in $\mathcal{G}$ too. Denoting the weight for bus j as W_j, and putting $W = \mathrm{diag}(W_j, j\in\mathcal{N})$, we can define a scaled version of the Laplacian matrix given as*

$$\tilde{L} := W^{-1/2} L W^{-1/2}.$$

Such scaled Laplacian matrix appears in Chapters 3 and 4, and all results mentioned in this chapter for L can be generalized to $\tilde{L}$.

2.4 Graphical Interpretation of the Laplacian Inverse

In this section, we explain how the elements of the matrix A as defined in Section 2.3 are related to the tree distributions of the power network $\mathcal{G}$ and demonstrate how this relation reveals new perspectives on some well-studied quantities in power system analysis.

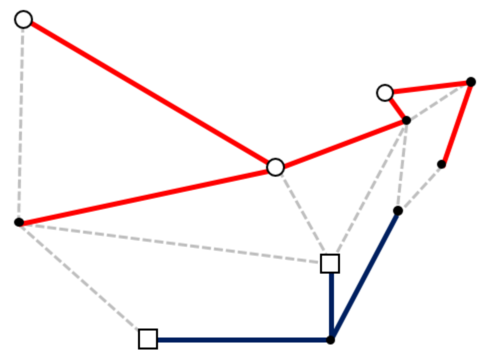


Figure 2.1: An example element in $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$, where circles correspond to elements in $\mathcal{N}_1$ and squares correspond to elements in $\mathcal{N}_2$. The two spanning trees containing $\mathcal{N}_1$ and $\mathcal{N}_2$ are highlighted as solid lines.

Spectral Representation

Recall from Proposition 2.3.3 that we know $\det(A)$ is determined by the spanning trees of $\mathcal{G}$. This is an example of how the tree structure of $\mathcal{G}$ is related to certain algebraic properties of the matrix A. We now present a finer-grained result that explicitly represents elements of A using the tree structure of $\mathcal{G}$. To do so, more notations are in order. Given a subset E of $\mathcal{E}$, we use $\mathcal{T}_E$ to denote the set of spanning trees of $\mathcal{G}$ with edges from E, which can be empty if E is too small. For two subsets $\mathcal{N}_1, \mathcal{N}_2$ of $\mathcal{N}$ (that do not need to be disjoint), we define $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$ to be the set of spanning forests of $\mathcal{G}$ consisting of exactly two trees that contain $\mathcal{N}_1$ and $\mathcal{N}_2$ respectively. See Figure 2.1 for an illustration of $\mathcal{T}(\mathcal{N}_1, \mathcal{N}_2)$. Given a set E of edges, we write

$$\chi(E) := \prod_{e \in E} B_e.$$

Then the celebrated All Minors Matrix Tree Theorem [19] applied to the matrix $\overline{L}$ implies:

Proposition 2.4.1. *The determinant of the matrix obtained by deleting the i-th row and j-th column of $\overline{L}$, denoted as $\overline{L}^{ij}$, is given by*

$$\det\left(\overline{L}^{ij}\right) = (-1)^{i+j} \sum_{E \in \mathcal{T}(\{i,j\},\{n\})} \chi(E).$$

This result leads to the following graphical interpretation of the elements of A:

Proposition 2.4.2. *For any $i, j \in \mathcal{N}$ such that $i, j \neq n$, we have*

$$A_{ij} = \frac{\sum_{E \in \mathcal{T}(\{i,j\},\{n\})} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)}.$$

Proof. Put A_j to be the j-th column of A. Note that $\overline{L}A_j = e_j$, where $e_j \in \mathbb{R}^{n-1}$ is the vector with 1 as its j-th component and 0 otherwise. Therefore by Cramer's rule, we have

$$A_{ij} = \frac{\det\left(\overline{L}_j^i\right)}{\det\left(\overline{L}\right)}, \tag{2.7}$$

where $\overline{L}_j^i$ is the matrix obtained by replacing the i-th column of $\overline{L}$ by e_j. Now by Proposition 2.4.1, we have

$$\det\left(\overline{L}_j^i\right) = (-1)^{i+j}\det\left(\overline{L}^{ij}\right) = \sum_{E\in\mathcal{T}(\{i,j\},\{n\})} \chi(E)$$

and by the Kirchhoff's Matrix Tree Theorem

$$\det\left(\overline{L}\right) = \sum_{E\in\mathcal{T}_{\mathcal{E}}} \chi(E).$$

The desired result then follows. □

In (2.7), the denominator is a common normalizer among all entries of A_{ij}, and the numerator is only related to the set $\mathcal{T}(\{i, j\}, \{n\})$. In other words, A_{ij} is proportional to the (weighted) number of trees that connect i to j, and can be interpreted as the "connection strength" between the buses i and j in $\mathcal{G}$. Moreover, since the matrix A fully determines the branch flow f from the system injection p, Proposition 2.4.2 tells us that power redistribution under DC power flow equations (2.5) can be described using the distribution of different types of trees in the transmission network. In particular, we can deduce analytical properties of DC power flows using purely graphical structures. As an example, the following corollary is an interesting implication of Proposition 2.4.2.

Corollary 2.4.3. *For all $i, j \in \mathcal{N}$ such that $i, j \neq n$, we have*

$$A_{ij} \geq 0,$$

where the equality holds if and only if every path from i to j contains n.

Proof. Since $\chi(E) \geq 0$ for all E, we clearly have $A_{ij} \geq 0$ and equality holds if and only if the set $\mathcal{T}(\{i, j\}, \{n\})$ is empty.

If every path from i to j contains n, then since any tree containing $\{i, j\}$ induces a path from i to j, we know this tree also contains n. As a result, $\mathcal{T}(\{i, j\}, \{n\}) = \emptyset$.

Conversely, if $\mathcal{T}(\{i,j\},\{n\}) = \emptyset$, then any path from i to j must contain n, since for any path from i to j not passing n, we can iteratively add edges that do not have n as an endpoint to obtain a spanning tree over the nodes set $\mathcal{N}\setminus\{n\}$. This tree together with the node n itself is an element of $\mathcal{T}(\{i,j\},\{n\})$.

We thus have shown that $\mathcal{T}(\{i,j\},\{n\})$ is empty if and only if every path from i to j contains n. This completes the proof. □

Generation Shift Sensitivity Factor

Given a pair of buses i, j and an edge $e = (w, z)$, if we shift an injection of amount Δp from i to j, the branch flow on e will change accordingly based on the DC power flow equations (2.5). Denote this change as Δf_e. The ratio

$$D_{ij,e} := \frac{\Delta f_e}{\Delta p}$$

is known as the generation shift sensitivity factor between the pair of buses i, j and the edge e [76]. When $\tilde{e} := (i, j) \in \mathcal{E}$ is an edge of the power network, we also write $D_{ij,e}$ as $D_{\tilde{e},e}$. Under DC power flow (2.5), $D_{ij,e}$ is fully determined by the matrix A, and can be computed as (see [76]):

$$D_{ij,e} = A_{iw} + A_{jz} - A_{iz} - A_{jw}.$$

This formula together with Proposition 2.4.2 implies the following result:

Corollary 2.4.4. *For $i, j \in \mathcal{N}$, $e = (w, z) \in \mathcal{N}$ such that $i, j, w, z \neq n$, we have*

$$D_{ij,e} = \frac{1}{\sum_{E\in\mathcal{T}_{\mathcal{E}}}\chi(E)}\left(\sum_{E\in\mathcal{T}(\{i,w\},\{j,z\})}\chi(E) - \sum_{E\in\mathcal{T}(\{i,z\},\{j,w\})}\chi(E)\right).$$

See Section 2.5 for its proof.

Despite the complexity of this formula, it carries clear graphical meaning as we now explain. In this formula, the sum is over the spanning forests $\mathcal{T}(\{i,w\},\{j,z\})$ and $\mathcal{T}(\{i,z\},\{j,w\})$. Each element in $\mathcal{T}(\{i,w\},\{j,z\})$, as illustrated in Figure 2.2, specifies a way to connect i to w and j to z through disjoint trees, and captures a possible path for i, j to "spread" impact to (w, z). Similarly, elements in $\mathcal{T}(\{i,z\},\{j,w\})$ captures possible paths for i, j to "spread" impact to (z, w), which counting orientation, contributes negatively. Therefore, Corollary 2.4.4 tells us that the impact of shifting generations from i to j propagates to the edge $e = (w, z)$ through all possible spanning trees that connect the endpoints i, j, w, z, counting orientation. The relative strength of the positive and negative impacts determines the sign of $D_{ij,e}$.

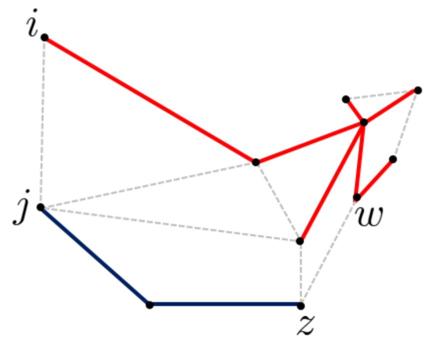


Figure 2.2: An example element in $\mathcal{T}(\{i, w\}, \{j, z\})$. The spanning trees containing $\{i, w\}$ and $\{j, z\}$ are highlighted as solid lines.

Effective Reactance

The Laplacian matrix L appears in circuit analysis as the admittance matrix (with a different weight), which explicitly relates the voltage and current vector in an electrical network [29]. In particular, given a network of resistors, it is shown in [29] that the effective resistance between two nodes i and j can be computed as

$$R_{ij} := L^{\dagger}_{ii} + L^{\dagger}_{jj} - L^{\dagger}_{ij} - L^{\dagger}_{ji}, \tag{2.8}$$

where $L^{\dagger}$ is the Penrose-Moore pseudoinverse of L. Following a similar calculation, we can show that (2.8) also gives the effective reactance between the buses i and j in a power system. That is, assuming we connect the buses i and j to an external probing circuit, when there is no other injection in the network, the power flow f_{ij} (from the external circuit) into bus i and out of bus j (into the external circuit) is given as

$$f_{ij} = \frac{\theta_i - \theta_j}{R_{ij}}.$$

Therefore the network can be equivalently reduced to a single line with reactance R_{ij}. If $(i, j) \in \mathcal{E}$, denoting $X_{ij} := 1/B_{ij}$ to be the reactance of (i, j), then physical intuition suggests that

$$R_{ij} < X_{ij},$$

as connections from the network should only decrease the overall reactance. We now show that $X_{ij} - R_{ij}$ also carries graphical meaning, proving its nonnegativity rigorously. To do so, we need the following relation between $L^{\dagger}$ and A, which is proved in Section 2.5.

Lemma 2.4.5. *For $(i, j) \in \mathcal{E}$ such that $i, j \neq n$, we have*

$$L^{\dagger}_{ii} + L^{\dagger}_{jj} - L^{\dagger}_{ij} - L^{\dagger}_{ji} = A_{ii} + A_{jj} - A_{ij} - A_{ji}.$$

In other words, we can replace the $L^\dagger$'s in the equation (2.8) by the matrix A, which allows us to apply Proposition 2.4.2 to obtain the following result:

Corollary 2.4.6. *For $(i, j) \in \mathcal{E}$ such that $i, j \neq n$, we have*

$$X_{ij} - R_{ij} = X_{ij} \cdot \frac{\sum_{E \in \mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)}.$$

In particular, we always have $X_{ij} \geq R_{ij}$, and the inequality is strict if the graph after removing (i, j) is connected.

Proof. By taking $w = i$, $z = j$ in Corollary 2.4.4 and applying Lemma 2.4.5, we see that

$$R_{ij} = A_{ii} + A_{jj} - 2A_{ij} = \frac{\sum_{E \in \mathcal{T}(\{i\},\{j\})} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)}.$$

For each forest in $\mathcal{T}(\{i\}, \{j\})$, we can add the edge (i, j) to form a spanning tree passing through (i, j). Conversely, each spanning tree passing through (i, j) produces a forest in $\mathcal{T}(\{i\}, \{j\})$ after removing the edge (i, j). This, by the definition of $\chi(E)$, implies

$$\sum_{E \in \mathcal{T}(\{i\},\{j\})} \chi(E) = X_{ij} \sum_{E \in \mathcal{T}'} \chi(E),$$

where $\mathcal{T}'$ denotes the set of all spanning trees passing through (i, j). As a result,

$$\begin{aligned} X_{ij} - R_{ij} &= X_{ij} \cdot \frac{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E) - \sum_{E \in \mathcal{T}'} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)} \\ &= X_{ij} \cdot \frac{\sum_{E \in \mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)}. \end{aligned}$$

□

From Corollary 2.4.6, we see that for an edge (i, j), the reduction ratio of its reactance coming from the network is precisely the (weighted) portion of spanning trees not passing through (i, j) among all spanning trees. Thus more connections from the network leads to more reduction in the effective reactance on (i, j), which agrees with our physical intuition.

We remark that this reactance reduction ratio is closely related to the spanning tree centrality measure [69]. Indeed, from the very definition of spanning tree centrality, we have

$$\frac{\sum_{E \in \mathcal{T}_{\mathcal{E}\setminus\{(i,j)\}}} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)} + c_{(i,j)} = 1,$$

where $c_{(i,j)}$ denotes the spanning tree centrality of (i,j). As a result

$$R_{ij} = X_{ij} c_{(i,j)},$$

or in other words, in a power redistribution setting, the spanning tree centrality of a transmission line precisely captures the ratio between its effective reactance R_{ij} and its physical line reactance X_{ij}.

2.5 Proofs

Proof of Corollary 2.4.4

By Proposition 2.4.2, we see that

$$\begin{aligned}
&\left(\sum_{E\in\mathcal{T}_{\mathcal{E}}} \chi(E)\right)\left(A_{iw} + A_{jz} - A_{iz} - A_{jw}\right) \\
&= \sum_{E\in\mathcal{T}(\{i,w\},\{n\})} \chi(E) + \sum_{E\in\mathcal{T}(\{j,z\},\{n\})} \chi(E) \\
&\quad - \sum_{E\in\mathcal{T}(\{j,w\},\{n\})} \chi(E) - \sum_{E\in\mathcal{T}(\{i,z\},\{n\})} \chi(E).
\end{aligned} \tag{2.9}$$

We can decompose the set $\mathcal{T}(\{i,w\},\{n\})$ based on the tree that the bus j belongs to. This leads to the identity

$$\mathcal{T}(\{i,w\},\{n\}) = \mathcal{T}(\{i,j,w\},\{n\}) \sqcup \mathcal{T}(\{i,w\},\{j,n\}),$$

where $\sqcup$ means disjoint union. Similarly we also have

$$\mathcal{T}(\{j,z\},\{n\}) = \mathcal{T}(\{i,j,z\},\{n\}) \sqcup \mathcal{T}(\{j,z\},\{i,n\}),$$

$$\mathcal{T}(\{i,z\},\{n\}) = \mathcal{T}(\{i,j,z\},\{n\}) \sqcup \mathcal{T}(\{i,z\},\{j,n\}),$$

$$\mathcal{T}(\{j,w\},\{n\}) = \mathcal{T}(\{i,j,w\},\{n\}) \sqcup \mathcal{T}(\{j,w\},\{i,n\}).$$

Plugging the above decompositions to (2.9) and canceling the common terms, we see

$$\begin{aligned}
&\left(\sum_{E\in\mathcal{T}_{\mathcal{E}}} \chi(E)\right)\left(A_{iw} + A_{jz} - A_{iz} - A_{jw}\right) \\
&= \sum_{E\in\mathcal{T}(\{i,w\},\{j,n\})} \chi(E) + \sum_{E\in\mathcal{T}(\{j,z\},\{i,n\})} \chi(E) \\
&\quad - \sum_{E\in\mathcal{T}(\{j,w\},\{i,n\})} \chi(E) - \sum_{E\in\mathcal{T}(\{i,z\},\{j,n\})} \chi(E).
\end{aligned} \tag{2.10}$$

Again we have the following set of identities

$$\mathcal{T}(\{i,w\},\{j,n\}) = \mathcal{T}(\{i,w\},\{j,z,n\}) \sqcup \mathcal{T}(\{i,w,z\},\{j,n\}),$$

$$\mathcal{T}(\{j,z\},\{i,n\}) = \mathcal{T}(\{j,z\},\{i,w,n\}) \sqcup \mathcal{T}(\{j,w,z\},\{i,n\}),$$

$$\mathcal{T}(\{j,w\},\{i,n\}) = \mathcal{T}(\{j,w\},\{i,z,n\}) \sqcup \mathcal{T}(\{j,w,z\},\{i,n\}),$$

$$\mathcal{T}(\{i,z\},\{j,n\}) = \mathcal{T}(\{i,z\},\{j,w,n\}) \sqcup \mathcal{T}(\{i,w,z\},\{j,n\}).$$

Plugging to (2.10) and canceling common terms, we see

$$\begin{aligned}
&\left(\sum_{E\in\mathcal{T}_{\mathcal{E}}} \chi(E)\right)\left(A_{iw} + A_{jz} - A_{iz} - A_{jw}\right) \\
&= \sum_{E\in\mathcal{T}(\{i,w\},\{j,z,n\})} \chi(E) + \sum_{E\in\mathcal{T}(\{j,z\},\{i,w,n\})} \chi(E) \\
&\quad - \sum_{E\in\mathcal{T}(\{j,w\},\{i,z,n\})} \chi(E) - \sum_{E\in\mathcal{T}(\{i,z\},\{j,w,n\})} \chi(E) \\
&= \sum_{E\in\mathcal{T}(\{i,w\},\{j,z\})} \chi(E) - \sum_{E\in\mathcal{T}(\{j,w\},\{i,z\})} \chi(E),
\end{aligned}$$

where the last equality follows from

$$\mathcal{T}(\{i,w\},\{j,z\}) = \mathcal{T}(\{i,w\},\{j,z,n\}) \sqcup \mathcal{T}(\{j,z\},\{i,w,n\})$$

and

$$\mathcal{T}(\{j,w\},\{i,z\}) = \mathcal{T}(\{j,w\},\{i,z,n\}) \sqcup \mathcal{T}(\{i,z\},\{j,w,n\}.$$

This completes the proof. □

Proof of Lemma 2.4.5

To prove this result, we derive explicit formulae for the branch flow vector f in terms of the power injection p. From $p = L\theta$ we know that whenever this equation is solvable, the solution θ is unique after quotienting away the kernel of L given as span $(\mathbf{1})$. Noting this is also the kernel of C^T, we see that $f = BC^T\theta$ is uniquely determined. Towards the goal of an explicit formula, we can proceed in two ways: The first way relies on the fact that $L^\dagger p$ always gives a feasible θ, and therefore

$$f = BC^T L^\dagger p. \tag{2.11}$$

The second way is to set the phase angle at the slack bus to zero, which implies that

$$\overline{\theta} = (\overline{L})^{-1}\overline{p} = A\overline{p},$$

where $\overline{\theta}$ and $\overline{p}$ are the vector of non-slack bus phase angles and injections. Denote as $\overline{C}$ the matrix obtained from C by removing the row corresponding to the slack bus. We then have

$$f = B\overline{C}^T\overline{\theta} = B\overline{C}^T A\overline{p}. \tag{2.12}$$

Now let $i, j \in \mathcal{N}$ with neither i nor j being the slack bus. Under the injection $p_i = -p_j = 1$, by equating the branch flow f_{ij} computed from (2.11) and (2.12), we obtain that

$$L^\dagger_{ii} + L^\dagger_{jj} - L^\dagger_{ij} - L^\dagger_{ji} = A_{ii} + A_{jj} - A_{ij} - A_{ji}.$$

□

Chapter 3

SPECTRAL DECOMPOSITION AND FREQUENCY REGULATION

In this chapter, we focus on the fast-timescale aspect of power grids and study how the system robustness against injection disturbances in transient state is related to the transmission network topology. Our approach relies on the decomposition of system response trajectory under the swing dynamics along scaled Laplacian spectrum, and captures the interplay among network topology, system inertia, and generator and load damping. The major results of this chapter can be summarized as follows: (a) We show that whether the system oscillates or not is determined by how strong the damping normalized by inertia is compared to network connectivity in the "corresponding" direction. (b) We prove that the power grid robustness against low frequency disturbance is mostly determined by network connectivity, while its robustness against high frequency disturbance is mostly determined by system inertia. (c) We demonstrate that although increasing system damping helps suppress disturbances, such benefits are mostly in the medium frequency band. (d) We devise a quantitative explanation for why load-side participation helps improve system behavior in the transient state, and demonstrate how our results suggest an improved controller design that can suppress input noise much more effectively.

To establish these results, we first show in Section 3.2 that the system response of swing dynamics can be decomposed along scaled Laplacian eigenvectors, in both the time and Laplace domain. Later in Section 3.3, we discuss how our results should be interpreted in a practical system; and in particular show that the transmission network topology determines the system robustness against low-frequency disturbances. In Section 3.4, we explain the benefits of load-side controllers using our framework and present a new controller that is specifically tailored to suppress high frequency oscillations. We then present a case study on the IEEE 39-bus New England interconnection testbed that confirms our analytical results.

3.1 System Model

Recall the linearized swing dynamics we introduced in Section 2.1 that describe the fast-timescale responses of a power system:

$$M_j \dot{\omega}_j = r_j - d_j - D_j \omega_j - \sum_{e \in \mathcal{E}} C_{je} f_e, \qquad j \in \mathcal{N} \tag{3.1a}$$

$$\dot{f}_{ij} = B_{ij}(\omega_i - \omega_j), \qquad (i,j) \in \mathcal{E}. \tag{3.1b}$$

Using x to denote the system state $x = [\omega; f]$, and putting $M = \text{diag}(M_j, j \in \mathcal{N})$, $D = \text{diag}(D_j, j \in \mathcal{N})$, $B = \text{diag}(B_e, e \in \mathcal{E})$, we can rewrite the system dynamics (3.1) in the state-space form

$$\dot{x} = \begin{bmatrix} -M^{-1}D & -M^{-1}C \\ BC^T & 0 \end{bmatrix} x + \begin{bmatrix} M^{-1} \\ 0 \end{bmatrix} p. \tag{3.2}$$

The matrix

$$A := \begin{bmatrix} -M^{-1}D & -M^{-1}C \\ BC^T & 0 \end{bmatrix}$$

is referred to as the system matrix in this chapter (note that A does not represent the inverse of a submatrix of L here). The system (3.2) can be interpreted as a multi-input-multi-output linear system with input p and output x. We emphasize that the variables $x = [\omega; f]$ denote deviations from their nominal values so that $x(t) = 0$ means the system is in its nominal state at time t.

In this chapter, we use a scaled version of the Laplacian matrix of $\mathcal{G}$ defined by $\tilde{L} = M^{-1/2} L M^{-1/2}$, which is explicitly given by

$$\tilde{L}_{ij} = \begin{cases} -\frac{B_{ij}}{\sqrt{M_i M_j}} & i \neq j, (i,j) \in \mathcal{E} \text{ or } (j,i) \in \mathcal{E} \\ \frac{1}{M_i} \sum_{j : j \in N(i)} B_{ij} & i = j \\ 0 & \text{otherwise.} \end{cases}$$

To simplify the notations, we drop the tilde and use L to denote the scaled Laplacian henceforth (in this chapter). Similar to the non-scaled version, the scaled Laplacian L is also positive semidefinite and thus diagonalizable. We denote its eigenvalues and corresponding orthonormal eigenvectors as $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ and $v_1, v_2, \cdots, v_n$. When the matrix L has repeated eigenvalues, for each repeated eigenvalue λ_i with multiplicity m_i, the corresponding eigenspace of L always has dimension m_i, hence an orthonormal basis consisting of eigenvectors of L exists (yet such bases are not unique). We assume one of the possible orthonormal bases is chosen and fixed throughout the chapter.

The eigenvalues of L measure the graph connectivity from an algebraic perspective, and larger Laplacian eigenvalues suggest stronger connectivity. To make such discussions more concrete, we define a partial order $\preceq$ over the set of all weighted graphs with vertex set $\mathcal{N}$ as follows: For two weighted graphs $\mathcal{G}_1 = (\mathcal{N}, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{N}, \mathcal{E}_2)$, we say $\mathcal{G}_1 \preceq \mathcal{G}_2$ if $\mathcal{E}_1 \subset \mathcal{E}_2$, and for any $e \in \mathcal{E}_1$, the weight of e in $\mathcal{G}_1$ is no larger than that in $\mathcal{G}_2$. It is routine to check that $\preceq$ defines a partial order.[1] A more interesting result is that the mapping from a graph to its Laplacian eigenvalues preserves this order.

Lemma 3.1.1. *Let L_1 and L_2 be the (scaled) Laplacian matrices of two weighted graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ with $\mathcal{G}_1 \preceq \mathcal{G}_2$. Let $0 = \lambda_1^1 \le \lambda_2^1 \le \cdots \le \lambda_n^1$ and $0 = \lambda_1^2 \le \lambda_2^2 \le \cdots \le \lambda_n^2$ be the eigenvalues of L_1 and L_2 respectively. Then*

$$\lambda_i^1 \le \lambda_i^2, i = 1, 2, \ldots, n.$$

Proof. This result follows from the fact that $L_2 - L_1$ is positive semidefinite, which is easy to check from our definition of $\preceq$. $\square$

In fact, we can devise better estimates on the relative orders of the eigenvalues λ_i^j. See Chapter 6 for more discussions and results therein. Throughout this chapter, whenever we compare two graphs in terms of their connectivity, we always refer to the partial order $\preceq$.

To make the analysis tractable, we further assume that the inertia and damping of the buses are proportional to its power ratings. That is, we assume there is a baseline inertia μ and damping δ such that for each generator j with power rating F_j, we have $M_j = F_j\mu$ and $D_j = F_j\delta$. This is a natural setting, as machines with high ratings are typically "heavy" and have more significant impact on the overall system dynamics. See [21, 51] for more details. Under such assumptions, the ratios D_j/M_j are independent of j, and therefore $M^{-1}D = \gamma I_n$ where $\gamma = \delta/\mu > 0$.

We will study both the transmission network Laplacian matrix L and Laplace domain properties of (3.2). To clear potential confusion, we agree that whenever the adjective Laplacian is used, we refer to quantities related to the Laplacian matrix L, while whenever the noun Laplace is used, we refer to notions about the Laplace transform

$$\mathcal{L}\left\{s(t)\right\}(\tau) := \int_0^\infty s(t)e^{-\tau t}dt$$

[1] We emphasize that this is not a complete order over all graphs. That is, not any pair of graphs with the same number of vertices are comparable through this order.

or notions defined in the Laplace domain.

3.2 Characterization of System Response

In this section, we give a complete characterization of the system response of (3.2) based on spectral decomposition in both time and Laplace domain.

Stability under Zero Input

We first determine the modes of the system (3.2). That is, we compute the eigenvalues of the system matrix A. Such eigenvalues indicate whether the system is stable, and if it is, how fast the system converges to an equilibrium state.

Theorem 3.2.1. *Let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of L with corresponding orthonormal eigenvectors $v_1, v_2, \ldots, v_n$. Then:*

1. *0 is an eigenvalue of A of multiplicity $m-n+1$, where m is the number of lines. The corresponding eigenvectors are of the form $[0; f]$ with $f \in \mathrm{kernel}(C)$;*
2. *$-\gamma$ is a simple eigenvalue of A with $\left[M^{-1/2}v_1; 0\right]$ as a corresponding eigenvector;*
3. *For $i = 2, 3, \ldots, n$, $\phi_{i,\pm} = \frac{-\gamma \pm \sqrt{\gamma^2 - 4\lambda_i}}{2}$ are eigenvalues of A. For any such $\phi_{i,\pm}$, an eigenvector is given by $\left[M^{-1/2}v_i; \phi_{i,\pm}^{-1} BC^T M^{-1/2} v_i\right]$.*

The proof of this Theorem is presented in Section 3.7. When $m - n + 1 = 0$ or equivalently when the network is a tree, item 1 of Theorem 3.2.1 is understood to mean that the system matrix A does not have 0 as an eigenvalue.

Assuming $\gamma^2 - 4\lambda_i \neq 0$ for all i, we get $2n - 1$ nonzero eigenvalues of A from item 2 and item 3 of Theorem 3.2.1, counting multiplicity, which together with the $m-n+1$ multiplicity from item 1 gives $m+n$ eigenvalues as well as $m+n$ linearly independent eigenvectors. Therefore we know A is always diagonalizable over the complex field $\mathbb{C}$, provided critical damping, that is $\gamma^2 - 4\lambda_i = 0$ for some i, does not occur. We assume this is the case in all the following derivations. When critical damping does occur, our results can be generalized using the standard Jordan decomposition.

Theorem 3.2.1 explicitly reveals the impact of the transmission network connectivity as captured by its Laplacian eigenvalues on the system (3.2) and tells us that the system mode shape is closely related to the corresponding Laplacian eigenvectors. In particular, we note that the real parts of $\phi_{i,\pm}$ are nonpositive, from which we deduce the following corollary:

Corollary 3.2.2. *The system* (3.2) *is marginally stable, with marginal stable states of the form* $[0; f]$ *with* $f \in \text{kernel}(C)$. *Therefore the system* (3.2) *is asymptotically stable on a tree.*

The kernel of C corresponds to the set of branch flow vectors f such that $\sum_{j \in N(i)} \tilde{f}_{ij} = 0$ for all $i \in \mathcal{N}$, where

$$\tilde{f}_{ij} := \begin{cases} f_{ij}, & (i,j) \in \mathcal{E} \\ -f_{ji}, & (j,i) \in \mathcal{E}. \end{cases}$$

They can be interpreted as flows that are balanced at all the buses (e.g., circulation flows on a loop) for which each bus i is neither a source node ($\sum_{j \in N(i)} \tilde{f}_{ij} > 0$) nor a sink node ($\sum_{j \in N(i)} \tilde{f}_{ij} < 0$). This corollary tells us that the only possible signals that can persist in (3.2) are the balanced branch flows. Of course, such marginally stable flows cannot exist in a real system because of losses in transmission lines (in which case our network dynamics (3.1b) is no longer accurate). Even if we take the simplified model (3.2), as long as the initial system branch flow does not belong to kernel(C), the system (3.2) under zero input $p = 0$ converges to the nominal state.

System Response to Step Input

Next we determine the system response to a step function. More precisely, we define $p(t) := r(t) - d(t)$ as the input function and compute the frequency trajectory $\omega(t)$ with $p(t)$ as input to (3.2), assuming $p(t)$ takes constant value p over time. The components p_j can be different over j. We put $p = \sum_i \hat{p}_i M^{1/2} v_i$ to be the decomposition of p along the scaled Lapalacian eigenvectors (note that the decomposition scaling $M^{1/2} v_i$ is different from the scaling $M^{-1/2} v_i$ in the following theorem statement).

Theorem 3.2.3. *Let* $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$ *be the eigenvalues of* L *with corresponding orthonormal eigenvectors* $v_1, v_2, \ldots, v_n$. *Assume:*

1. *The system* (3.2) *is initially at the nominal state* $x(0) = 0$
2. $\gamma^2 - 4\lambda_i \neq 0$ *for all* i.

Then

$$\omega(t) = \sum_{i=1}^{n} \frac{\hat{p}_i}{\sqrt{\gamma^2 - 4\lambda_i}} \left(e^{\phi_{i,+} t} - e^{\phi_{i,-} t} \right) M^{-1/2} v_i, \tag{3.3}$$

where

$$\phi_{i,+} := \frac{-\gamma + \sqrt{\gamma^2 - 4\lambda_i}}{2} \quad \phi_{i,-} := \frac{-\gamma - \sqrt{\gamma^2 - 4\lambda_i}}{2}.$$

See Section 3.7 for its proof.

We remark that all conditions in this theorem are for presentation simplicity and the frequency trajectory (3.3) can be generalized by adding correction terms to the case where neither condition is imposed. We opt not to doing so here as these terms lead to more tedious notations without revealing any new insights.

This result tells us that the frequency trajectory of (3.2) can be decomposed along scaled eigenvectors of the Laplacian matrix L. Moreover, we note that all $\phi_{i,\pm}$ have negative real parts except $\phi_{1,+} = 0$. Therefore the only term in (3.3) that persists is the term involving $\phi_{1,+}$ given as:

$$\frac{\hat{p}_1}{\sqrt{\gamma^2 - 4\lambda_1}} e^{\phi_{1,+}t} M^{-1/2} v_1 = \frac{\hat{p}_1}{\gamma} M^{-1/2} v_1.$$

Thus under the input $p = r - d$, the $\omega(t)$ signal converges to the steady state $\frac{\hat{p}_1}{\gamma} M^{-1/2} v_1$ exponentially fast. This allows us to recover the following result using a new argument.

Corollary 3.2.4. *Under step input p, the system* (3.2) *converges to a steady state with synchronized frequencies $\omega_i = \omega_j =: \overline{\omega}$. Moreover, $\overline{\omega} = 0$ if and only if the power injection is balanced $\sum_{i \in \mathcal{N}} p_i = 0$.*

Proof. It is easy to show

$$v_1 = \frac{M^{1/2}}{\sqrt{\sum_{j \in \mathcal{N}} M_j}} \mathbf{1}_n.$$

By Theorem 3.2.3, we know the steady state of (3.2) is $(\hat{p}_1/\gamma) M^{-1/2} v_1$, which then has all entries equal to the same value

$$\frac{\hat{p}_1}{\gamma \sqrt{\sum_{j \in \mathcal{N}} M_j}}.$$

Therefore $\omega_i = \omega_j =: \overline{\omega}$ for all $i, j \in \mathcal{N}$. From $p = \sum_i \hat{p}_i M^{1/2} v_i$ we see $\hat{p}_1 = (M^{-1/2} p)^T v_1 = p^T M^{-1/2} v_1$, and thus

$$\begin{aligned}\sum_{i \in \mathcal{N}} p_i = p^T \mathbf{1}_n &= \sqrt{\sum_{j \in \mathcal{N}} M_j} p^T M^{-1/2} v_1 = \sqrt{\sum_{j \in \mathcal{N}} M_j} \hat{p}_1 \\ &= \gamma \left(\sum_{j \in \mathcal{N}} M_j \right) \overline{\omega} = \left(\sum_{j \in \mathcal{N}} D_j \right) \overline{\omega}.\end{aligned}$$

Hence $\overline{\omega} = 0$ if and only if $\sum_{i \in \mathcal{N}} p_i = 0$. $\square$

Spectral Transfer Functions for Arbitrary Input

It is also informative to look at the system behavior of (3.2) from the Laplace domain. Instead of analyzing transfer functions from any input to any output as in the classical multi-input-multi-output system analysis, we take a slightly different approach such that the Laplacian matrix spectral information is preserved. More precisely, for a time-variant injection signal $p(t)$, we first decompose it into the spectral representation $p(t) = \sum_{i=1}^{n} \hat{p}_i(t) M^{1/2} v_i$. Now $\hat{p}_i(t)$ is a real-valued signal, and thus assuming enough regularity, we can rewrite $\hat{p}_i(t)$ as the integral of exponential signals $e^{\tau t}$ through inverse Laplace transform. It can be shown that when the input to system (3.2) takes the form $p(t) = e^{\tau t} M^{1/2} v_i$, the steady state frequency trajectory $\overline{\omega}(t)$ is given by $H_i(\tau) e^{\tau t} M^{-1/2} v_i$, where $H_i(\tau)$ is a complex-valued function of τ specifying the system gain and phase shift. We refer to the function $H_i(\tau)$ as the i-th spectral transfer function. Compared to classical transfer functions, the spectral version does not capture the relationship between any input-output pair, but in contrast captures the behavior of system (3.2) from a network perspective. Once the spectral transfer functions are known, we can compute the steady state trajectories for general input signal $p(t)$ through the following synthesis formula:

$$\overline{\omega}(t) = \sum_{i=1}^{n} \mathcal{L}^{-1} \left\{ H_i(\tau) \mathcal{L} \left\{ \hat{p}_i(t) \right\} (\tau) \right\} M^{-1/2} v_i.$$

Theorem 3.2.5. *For each i, assuming $\gamma^2 - 4\lambda_i \neq 0$, the i-th spectral transfer function is given by*

$$H_i(\tau) = \frac{\tau}{\tau^2 + \gamma\tau + \lambda_i}.$$

The proof of this result is presented in Section 3.7. We remark that a similar formula also shows up in [51] as the representative machine transfer function for swing dynamics.

3.3 Interpretations

In this section, we present a collection of intuition that can be devised from the results in Section 3.2. They are useful for making general inferences as well as for the controller design in Section 3.4.

Network Connectivity and System Stabilization

We first clarify how the network connectivity affects the system stability. Towards this goal, we rewrite (3.3) as

$$\omega(t) = \sum_{i=1}^{n} \hat{p}_i \hat{\omega}^i(t) M^{-1/2} v_i.$$

The signal $\hat{\omega}^i(t)$ captures the response of system (3.2) along $M^{-1/2}v_i$ to a step function input. By Theorem 3.2.1, we see that whether the system oscillates or not is determined by the signs of $\gamma^2 - 4\lambda_i$. For λ_i such that $\gamma^2 - 4\lambda_i > 0$, we have

$$\hat{\omega}^i(t) = \frac{1}{\sqrt{\gamma^2 - 4\lambda_i}} \left(e^{\phi_{i,+} t} - e^{\phi_{i,-} t} \right)$$

with $\phi_{i,\pm} \leq 0$. Thus the system is over-damped along $M^{-1/2}v_i$, and deviations along $M^{-1/2}v_i$ exponentially fade away without oscillation. The slower-decaying exponential has a decaying rate determined by $\phi_{i,+}$, which is a decreasing function in λ_i. Thus a larger λ_i implies faster decay. Intuitively, this tells us that when the system damping is strong with respect to its inertia, adding connectivity helps move more disturbances to the damping component so that disturbances can be absorbed sooner.

For γ such that $\gamma^2 - 4\lambda_i < 0$, we have

$$\hat{\omega}^i(t) = \frac{2}{\sqrt{4\lambda_i - \gamma^2}} e^{-\frac{\gamma}{2}t} \sin\left(\frac{\sqrt{4\lambda_i - \gamma^2}}{2} t \right).$$

Thus the system is under-damped along $M^{-1/2}v_i$ and oscillations do occur. We also note that larger values of λ_i lead to oscillations of higher frequency. This intuitively can be interpreted as the following: When the system damping is not strong enough compared to its inertia, adding connectivity causes the unabsorbed oscillations to propagate throughout the network faster, thus bringing disturbances to the already over-burdened damping components and making the system oscillate in a higher frequency.

We thus see that Theorem 3.2.1 and Theorem 3.2.3 precisely clarify our seemingly contradictory intuition on whether connectivity is beneficial to stabilization – it depends on how strong the system is damped compared to its inertia, i.e., how fast the system can dissipate energy.

Robustness to Disturbance

The impact of different system parameters in the Laplace domain can be understood from the spectral transfer functions H_i. Recall by Theorem 3.2.5 that for a signal of the form $p(t) = e^{\tau t} v_i$, the steady state output signal of (3.2) is

$$\overline{\omega}_i(t;\tau) = \frac{\tau e^{\tau t}}{\tau^2 + \gamma\tau + \lambda_i} M^{-1} v_i.$$

In particular, if we focus on the j-th component of $\overline{\omega}_i(t)$, which corresponds to the frequency trajectory of bus j, we have

$$\overline{\omega}_{i,j}(t;\tau) = \frac{\tau v_{i,j} e^{\tau t}}{M_j\tau^2 + D_j\tau + \lambda_i M_j}.$$

Under the proportional rating assumption mentioned in Section 3.1, one can show that $\lambda_i M_j = \overline{\lambda}_i$, where $\overline{\lambda}_i$ is the i-th Laplacian eigenvalue when the "heaviest" generator is normalized to have unit inertia $\max_{j\in\mathcal{N}} M_j = 1$ and can be interpreted as the pure topological part in the Laplacian eigenvalues λ_i. This allows us to compute

$$\left|\overline{\omega}_{i,j}(\mathrm{j}\sigma)\right| = \frac{|\sigma|}{\sqrt{M_j^2\sigma^4 + (D_j^2 - 2\overline{\lambda}_i M_j)\sigma^2 + \overline{\lambda}_i^2}} \tag{3.4}$$

and conclude the following (See Figure 3.1 for an illustration):

1. For high frequency signals, the gain can be approximated by $\left|\overline{\omega}_{i,j}(\mathrm{j}\sigma)\right| \approx \frac{1}{M_j\sigma}$ and therefore the key parameter to suppress such disturbance is the rotational inertia M_j;

2. For low frequency signals, the gain approximates to $\left|\overline{\omega}_{i,j}(\mathrm{j}\sigma)\right| \approx \frac{\sigma}{\overline{\lambda}_i}$ and hence low frequency disturbances are mostly suppressed by the network topology;

3. For any fixed frequency σ, $\left|\overline{\omega}_{i,j}(\mathrm{j}\sigma)\right|$ is decreasing in D_j. This means that a larger damping leads to smaller gains for all frequencies. Such decrease, however, is negligible for very large or very small σ, and therefore increasing D_j mostly helps the system suppress oscillations in the medium frequency band.

This tells us that a system with large μ value is generally more robust against measurement noise, as such noise is usually of high frequency. Further, it shows that in order to suppress fluctuations with high frequency (say from renewable sources), the only effective way is to increase the inertia constant μ. Adding damping level δ or connectivity λ_i, although helpful, would be much less fruitful.

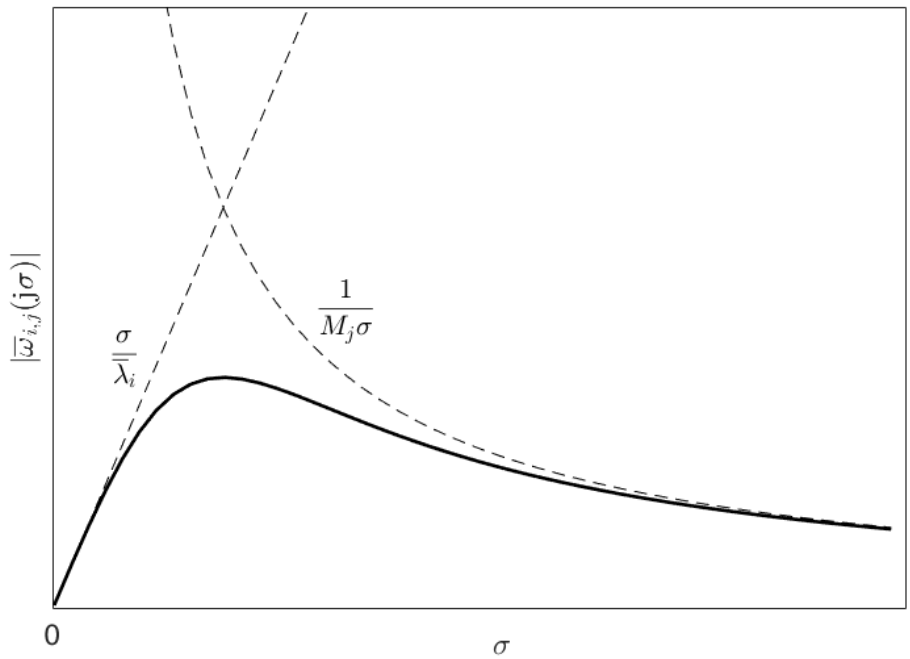


Figure 3.1: Illustration of the gain $|\overline{\omega}_{i,j}(\mathrm{j}\sigma)|$ as a function of σ.

Impact of Damping

As a by-product of our study, we can also examine how the system damping impacts the system performance. Towards this goal, we study two common metrics for $\hat{\omega}^i(t)$: (a) settling time, which is the time it takes $\hat{\omega}^i(t)$ to get within a certain range[2] around the steady state; and (b) nadir, which is defined to be the sup norm of $\hat{\omega}^i(t)$. Table 3.1 summarizes the formulae[3] for these metrics, and one can show using basic calculus that both the settling time and nadir are decreasing functions of γ, and thus decreasing in the damping constants D_j (provided that the inertia constants M_j are fixed).

This result, of course, does not generalize to $\omega(t)$ in a straightforward way because of the possibility of negative $\hat{p}_i$. Instead of focusing on $\omega(t)$ for a specific $p(t)$, we can look at all possible $\omega(t)$ and generalize our previous interpretations to the worst-case performance metric. To be concrete, let us take nadir as an example. By

[2]The range is specified as $[\omega_i^* - c, \omega_i^* + c]$, where ω^* is the equilibrium state and c is a constant.

[3]We define $\Delta_i = \left|\gamma^2 - 4\lambda_i\right|$ to simplify the formulae. The settling time formula is an upper bound as finding its exact value requires solving transcendental equations, which is generally hard.

Case	Settling Time	Nadir
$\gamma^2 > 4\lambda_i$	$\frac{1}{\gamma-\sqrt{\Delta_i}} \ln\left(\frac{1}{4c^2\Delta_i}\right)$	$\frac{1}{\sqrt{\Delta_i}}\left[\left(\frac{\gamma+\sqrt{\Delta_i}}{\gamma-\sqrt{\Delta_i}}\right)^{\frac{-\gamma+\sqrt{\Delta_i}}{2\sqrt{\Delta_i}}} - \left(\frac{\gamma+\sqrt{\Delta_i}}{\gamma-\sqrt{\Delta_i}}\right)^{\frac{-\gamma-\sqrt{\Delta_i}}{2\sqrt{\Delta_i}}}\right]$
$\gamma^2 < 4\lambda_i$	$\frac{1}{\gamma}\ln\left(\frac{4}{c^2\Delta_i}\right)$	$\frac{2}{\sqrt{\Delta_i}}\exp\left(-\frac{2\pi\gamma}{\sqrt{\Delta_i}}\right)$

Table 3.1: System performance in terms of network Laplacian eigenvalues, generator inertia, and damping ($\gamma := \delta/\mu, \Delta_i := |\gamma - 4\lambda_i|$).

(3.3), we see the nadir of frequency trajectory at bus j satisfies

$$\begin{aligned}
\left\|\omega_j(t)\right\|_\infty &\le M_j^{-1/2}\sum_{i=1}^n |\hat{p}_i| \left\|v_{i,j}\hat{\omega}^i(t)\right\|_\infty \\
&\le M_j^{-1/2}\sqrt{\sum_{i=1}^n |\hat{p}_i|^2}\sqrt{\sum_{i=1}^n \left\|v_{i,j}\hat{\omega}^i(t)\right\|_\infty^2} \\
&= M_j^{-1/2}\left\|M^{-1/2}p\right\|_2\sqrt{\sum_{i=1}^n \left\|v_{i,j}\hat{\omega}^i(t)\right\|_\infty^2} \\
&=: M_j^{-1/2}\left\|M^{-1/2}p\right\|_2 \|\omega\|_\infty^w .
\end{aligned}$$

It is easy to see that all the inequalities above can attain equalities. Therefore, among all input p with scaled unit energy $\left\|M^{-1/2}p\right\|_2 = 1$, the worst possible nadir is $M_j^{-1/2}\|\omega\|_\infty^w$, which is a decreasing function of δ from our previous discussions.

This worst-case nadir is a system level metric that is independent of the input. Although this metric does not predict the exact nadir for any specific input, it does reveal to what extent the system can tolerate disturbances of certain energy, which is a property that is intrinsic to the system itself. Moreover, for secure and robust operation of the grid, we need to make sure that the worst-case nadir is well-controlled. A similar argument can be also applied to the settling time for $\omega_j(t)$.

System Tradeoffs

When choosing system control parameters, there are usually tradeoffs among different performance goals, and we must balance different aspects to obtain a good design. A key tradeoff of this type revealed in our previous discussions is the tradeoff between having small network intrinsic frequency and improving system robustness against low frequency disturbance.

More specifically, it is easy to show that $\left|\overline{\omega}_{i,j}(\mathrm{j}\sigma)\right|$ is maximized at $\sigma_{i,j}^* = \sqrt{\frac{\lambda_i}{M_j}}$.

In other words, $\sigma^*_{i,j}$ can be interpreted as an intrinsic frequency of the network and oscillations around $\sigma^*_{i,j}$ are amplified at bus j through the transmission system. Typically high frequency oscillations should be suppressed, and thus we want smaller $\sigma^*_{i,j}$, which in turn leads to smaller connectivity $\overline{\lambda}_i$. On the other hand, we have shown that in order for the system (3.2) to be robust against low frequency noise (such as periodic load oscillations within a day), the transmission network should be designed with as large connectivity $\overline{\lambda}_i$ as possible. As a result, we cannot make the system (3.2) have small intrinsic frequency and be robust against low frequency noise at the same time.

3.4 Controller Design for Load-Side Participation

In this section, we discuss two implications of our results in Section 3.3 to load-side controller design.

Benefits of Load-side Participation

We adopt the controller design from [82] as an example to explain the benefits of load-side participation. We assume the system deviation is small so that the capacity bounds of load side controllers are not binding. In this setting, the control law of [82] simplifies to

$$d_j = K_p \omega_j, \tag{3.5}$$

which when plugged into (3.2) can be absorbed into the damping term $D_j\omega_j$. Therefore, the integration of controller (3.5) effectively increases the system damping level. Based on our discussions in Section 3.3, we conclude that load-side participation decreases both the settling time and nadir of (3.2). This means that with load-side participation, the system (3.2) is more responsive and its nadir under a disturbance is also better controlled.

Such benefits have been observed and confirmed in a series of work [43, 80–83] in their simulations. With our framework it is possible to analytically derive such results and quantify how beneficial the load-side integration can be when we use a certain system gain K_p. Moreover, it is observed in [80] that load-side participation also helps maintain system stability when the generator output fluctuates. Using our characterization in the Laplace domain, we see that such benefit comes from the improved system ability in suppressing oscillations of medium band frequency.

Proportional-Derivative (PD) Controller

Despite the many benefits of load-side controllers we have explained so far, one component still missing in (3.5) is that they only affect the system damping but cannot increase the system inertia. As mentioned in Section 3.3, the system inertia is the key parameter affecting the system robustness against high frequency oscillations. Nevertheless, a quick look at (3.2) suggests that in order to have a larger M_j, it suffices to add a derivative term in (3.5), which can be implemented through power electronics or invertors [42]:

$$d_j = K_p \omega_j + K_d \dot{\omega}_j. \tag{3.6}$$

Although it is a natural idea to generalize proportional controllers to PD controllers for performance tuning, we see that the need of this derivative term can actually be reversed engineered from our characterizations. Moreover, our framework reveals how the parameters K_p and K_d affect the system performance precisely, allowing us to optimize such gains subject to different design goals.

Using derivative terms in controller design is often problematic in practice due to the amplified noise in its measurement. However, we know from Section 3.3 that neither adding damping nor increasing network connectivity is particularly effective in suppressing disturbances in the high frequency regime. Thus in order to improve the grid stability under high frequency fluctuations, having certain components of the network that are able to measure the signal derivatives either explicitly or implicitly to provide the necessary inertia is inevitable.

3.5 Case Studies

In this section, we simulate the controller design (3.6) over the IEEE 39-bus New England interconnection system as shown in Figure 3.2, and compare its performance to that of (3.5) and the conventional droop control. There are 10 generators and 29 load nodes in the system, and we take the system parameters from the Matpower Simulation Package [85]. In contrast to our theoretical analysis, the simulation data does not satisfy the proportional rating assumption in Section 3.1. The droop control is implemented as the $D_j \omega_j$ term for generator buses and is deactivated for simulations with the controllers (3.5) and (3.6). We assume all the buses (including the generator buses) have load-side participation enabled and pick the controller gains K_p and K_d heterogeneously in proportional to the bus damping D_j.

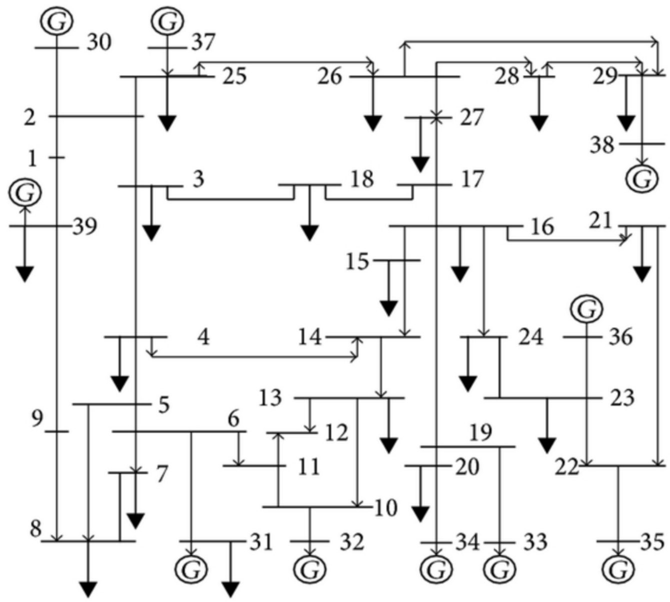


Figure 3.2: Line diagram of the IEEE 39-bus interconnection testbed.

Robustness against Measurement Noise

We first look at the controller performances against measurement noise. Towards this goal, we add a white Gaussian measurement noise of power -20 dBW to the frequency sensor at bus 30 and observe its frequency trajectory, which is shown in Figure 3.3. We can see that the controller (3.5) is less prone to measurement noise compared to the conventional droop control, because it increases the system damping level and therefore helps suppress the medium frequency part of the noise. However, its benefit in suppressing high frequency noise is limited, as one can see from its performance gap as compared with the controller (3.6). To more clearly see this distinction, we replace the measurement noise at bus 30 with the signal $0.2\sin(10\pi t)$ p.u. that contains only high frequency component and observe its trajectory. The result is shown in Figure 3.4. In this case, we see that controller (3.5) performs nearly the same as the conventional droop control, while the system under the improved controller (3.6) exhibits much smaller oscillation.

Wind Power Data

Next, we look at the performance of the controllers under real wind power generation data from [77]. We choose bus 30 to be the wind generator, whose output follows the profile given in [77] and look at the frequency trajectory at bus 36. The two buses are specifically chosen to be geographically far away so that the simulation

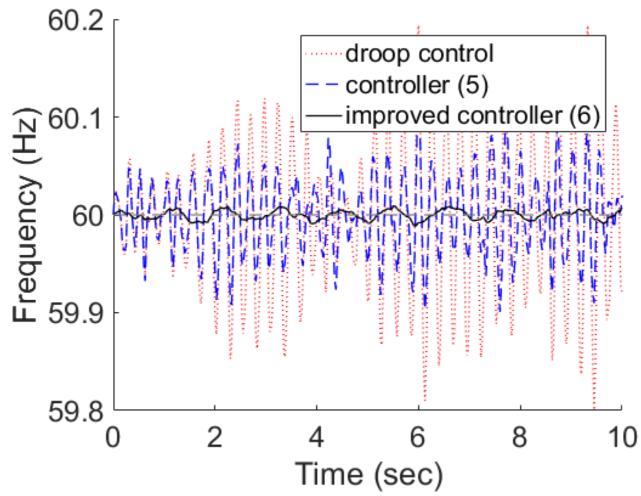


Figure 3.3: Frequency trajectory at bus 30 when we add white Gaussian measurement noise of −20 dBW.

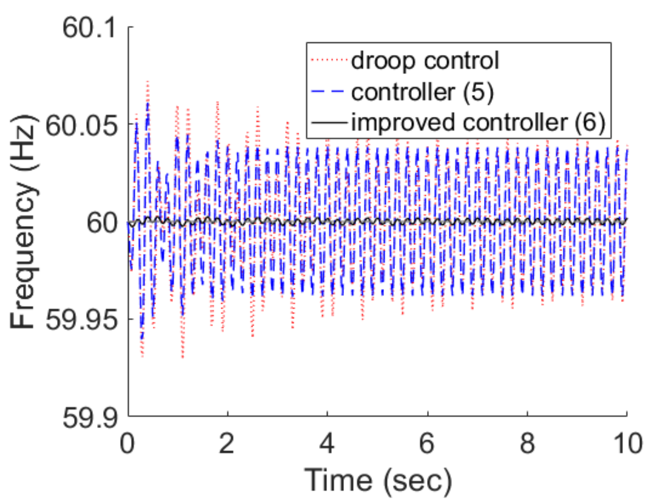


Figure 3.4: Frequency trajectory at bus 30 when we add a signal following the sine curve $0.2\sin(10\pi t)$ p.u.

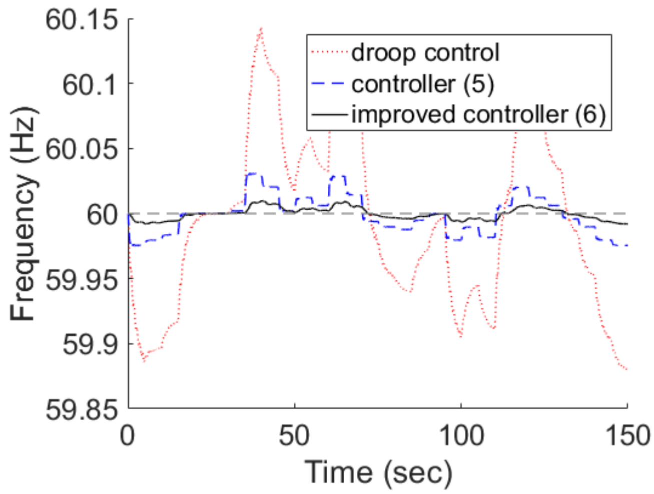


Figure 3.5: Frequency trajectory at bus 36 under wind power output at bus 30.

results reflect end user perception of such renewable penetration. The simulation result is shown in Figure 3.5. As one can see, compared to controller (3.5), the improved controller (3.6) incurs smaller frequency deviation at almost all times, and the resulting trajectory is smoother. This is because (3.6) filters away high frequency fluctuations in the generator profile. We expect such benefit to be more significant when the system aggregate load fluctuates more frequently because of increasing renewable penetration.

3.6 Conclusion

In this chapter, we proposed a framework that captures the interplay between transmission network topology and other system parameters. It leads to precise characterizations of how certain control parameters affect the system performance, and allows us to make general inferences without extensive simulation. We quantified the benefits of load-side participation within this framework, and explained how we can improve the controller design so that the system is more robust against high frequency oscillations.

3.7 Proofs

Proof of Theorem 3.2.1

By Schur complement, we can compute the characteristic polynomial of A as

$$\begin{aligned}
&\det(A - tI) \\
&= \det(-tI_m) \det\left(-\gamma I_n - tI_n - \frac{1}{t} M^{-1} C B C^T\right) \\
&= \det(-tI_m) \det\left(M^{-1/2}\right) \det\left(M^{1/2}\right) \\
&\quad \times \det\left(M^{1/2}(-\gamma I_n - tI_n - \frac{1}{t} M^{-1} C B C^T) M^{-1/2}\right) \\
&= (-1)^{m+n} t^{m-n} \det\left(L + (\gamma t + t^2) I_n\right).
\end{aligned}$$

All the above algebra is understood to be over the polynomial field generated from $\mathbb{R}[t]$, and thus we do not need to assume $t \neq 0$.

The term t^{m-n} contributes $m - n$ multiplicity to the eigenvalue 0 (in the case $\mathcal{G}$ is a tree, or equivalently $m = n - 1$, this is understood to mean that $t^{m-n} = t^{-1}$ cancels one multiplicity of 0). Let us now tackle the factor $\det\left(L + (\gamma t + t^2) I_n\right)$. It is easy to see that $\det\left(L + (\gamma t + t^2) I_n\right) = 0$ if and only if

$$t^2 + \gamma t + \lambda = 0$$

for some eigenvalue λ of L. Therefore the roots of $\det\left(L + (\gamma t + t^2) I_n\right)$ are given as

$$\phi = \frac{-\gamma \pm \sqrt{\gamma^2 - 4\lambda}}{2} \tag{3.7}$$

with λ traversing all eigenvalues of L. Among these roots, 0 appears exactly once, coming from the zero eigenvalue of L. Thus, altogether we know the eigenvalue of A consists of 0 with multiplicity $m - n + 1$ and non-zero roots of the form given by (3.7).

Next we determine the eigenvectors of the system matrix A. Let $\phi \neq 0$ be an eigenvalue corresponding to an eigenvector $[\omega; f]$. Then we have

$$-\gamma\omega - M^{-1} C f = \phi\omega \tag{3.8a}$$

$$B C^T \omega = \phi f. \tag{3.8b}$$

Substituting (3.8b) to (3.8a) and multipliying $M^{1/2}$ on both sides, we see that

$$L M^{1/2} \omega = -(\phi^2 + \gamma\phi) M^{1/2} \omega,$$

or in other words, $M^{1/2}\omega$ is an eigenvector of L affording $-(\phi^2 + \gamma\phi) = \lambda$. For any such ω, the corresponding f by (3.8b) is given by $f = \phi^{-1}BC^T\omega$. Moreover, we see that $\phi = -\gamma$ is a simple eigenvalue of A as the corresponding $\lambda = 0$ is a simple eigenvalue of L. Note that

$$\left\|B^{1/2}C^T M^{-1/2} v_1\right\|^2 = v_1^T L v_1 = 0$$

implies $BC^T M^{-1/2} v_1 = 0$, and therefore we see that

$$\left[M^{-1/2}v_1; -\frac{1}{\gamma}BC^T M^{-1/2} v_1\right] = \left[M^{-1/2}v_1; 0\right]$$

is an eigenvector of A affording $\phi = -\gamma$.

For $\phi = 0$, from (3.8b) we have $\omega = cI_n$ for some c. Plugging back to (3.8a), we have

$$cI_n = -\frac{1}{\gamma}M^{-1}Cf$$

and therefore $c\mathbf{1}^T M I_n = 0$, which implies $c = 0$. This then implies $f \in \text{kernel}(C)$ and $\omega = 0$. Therefore the eigenvectors corresponding to $\phi = 0$ are given by $[0; f]$ with $f \in \text{kernel}(C)$, which by dimension theorem has dimension $m - \text{rank}(C) = m - n + 1$. $\square$

Proof of Theorem 3.2.3

Recall we have shown in Section 3.2 that A is diagonalizable over the complex field $\mathbb{C}$, provided critical damping does not occur. Let $A = Q\Lambda Q^{-1}$ be an eigenvalue decomposition of A. Then we have $e^{At} = Qe^{\Lambda t}Q^{-1}$ for any $t \in \mathbb{R}$. Now the solution to the system (4.1) with a constant input p and nominal initial state is given as

$$x(t) = \int_0^t \left(e^{A(t-\tau)}\begin{bmatrix} M^{-1}p \\ 0 \end{bmatrix}\right) d\tau = Q\int_0^t e^{\Lambda(t-\tau)} d\tau Q^{-1}\begin{bmatrix} M^{-1}p \\ 0 \end{bmatrix}.$$

Write Λ in block diagonal form as

$$\Lambda = \begin{bmatrix} 0 & 0 \\ 0 & \Phi \end{bmatrix},$$

where Φ collect all nonzero eigenvalues of A. By Theorem 3.2.1, we can compute (3.9) to

$$\begin{aligned} x(t) &= Q\int_0^t \begin{bmatrix} (t-\tau)I_{m-n+1} & 0 \\ 0 & e^{\Phi(t-\tau)} \end{bmatrix} d\tau Q^{-1}\begin{bmatrix} M^{-1}p \\ 0 \end{bmatrix} \\ &= Q\begin{bmatrix} \frac{t^2}{2}I_{m-n+1} & 0 \\ 0 & \Phi^{-1}(e^{\Phi t} - I_{2n-1}) \end{bmatrix} Q^{-1}\begin{bmatrix} M^{-1}p \\ 0 \end{bmatrix}. \end{aligned} \tag{3.9}$$

Consider an eigen-pair (λ_i, v_i) of L. For $i = 1$, we have $\lambda_1 = 0$ and therefore by Theorem 3.2.1, we see $\left[M^{-1/2}v_1; 0\right]$ is an eigenvector of A affording $-\gamma$. For $i \geq 2$, we know

$$\phi_{i,+} := \frac{-\gamma + \sqrt{\gamma^2 - 4\lambda_i}}{2} \qquad \phi_{i,-} := \frac{-\gamma - \sqrt{\gamma^2 - 4\lambda_i}}{2}$$

are eigenvalues of A with corresponding eigenvectors $[M^{-1/2}v_i; \phi_{i,\pm}^{-1}BC^T M^{-1/2}v_i] =: z_{i,\pm}$. This allows us to decompose

$$\begin{aligned}
\begin{bmatrix} M^{-1/2}v_i \\ 0 \end{bmatrix} &= \frac{\sqrt{\gamma^2 - 4\lambda_i} - \gamma}{2\sqrt{\gamma^2 - 4\lambda_i}} \begin{bmatrix} M^{-1/2}v_i \\ \phi_{i,+}^{-1}BC^T M^{-1/2}v_i \end{bmatrix} \\
&\quad + \frac{\sqrt{\gamma^2 - 4\lambda_i} + \gamma}{2\sqrt{\gamma^2 - 4\lambda_i}} \begin{bmatrix} M^{-1/2}v_i \\ \phi_{i,-}^{-1}BC^T M^{-1/2}v_i \end{bmatrix} \\
&=: \lambda_{i,+}z_{i,+} + \lambda_{i,-}z_{i,-},
\end{aligned}$$

which then implies

$$\begin{aligned}
\begin{bmatrix} M^{-1}p \\ 0 \end{bmatrix} &= \sum_{i=1}^{n} \hat{p}_i \begin{bmatrix} M^{-1/2}v_i \\ 0 \end{bmatrix} \\
&= \hat{p}_1 \begin{bmatrix} M^{-1/2}v_1 \\ 0 \end{bmatrix} + \sum_{i=2}^{n} \lambda_{i,+}\hat{p}_i z_{i,+} + \sum_{i=2}^{n} \lambda_{i,-}\hat{p}_i z_{i,-}.
\end{aligned}$$

We emphasize that the input to system (3.2) is p, and $[M^{-1}p; 0]$ is the signal obtained by multiplying the input scaling matrix $[M^{-1}; 0]$ to p. By linearity, we can compute (3.9) as

$$\begin{aligned}
x(t) = &-\frac{\hat{p}_1}{\gamma}e^{-\gamma t} \begin{bmatrix} M^{-1/2}v_1 \\ 0 \end{bmatrix} + \sum_{i=2}^{n} \frac{\lambda_{i,+}\hat{p}_i}{\phi_{i,+}} e^{\phi_{i,+}t} z_{i,+} \\
&+ \sum_{i=2}^{n} \frac{\lambda_{i,-}\hat{p}_i}{\phi_{i,-}} e^{\phi_{i,-}t} z_{i,-} + \frac{\hat{p}_1}{\gamma} \begin{bmatrix} M^{-1/2}v_1 \\ 0 \end{bmatrix} \\
&- \sum_{i=2}^{n} \frac{\lambda_{i,+}\hat{p}_i}{\phi_{i,+}} z_{i,+} - \sum_{i=2}^{n} \frac{\lambda_{i,-}\hat{p}_i}{\phi_{i,-}} z_{i,-}.
\end{aligned}$$

One can check by direct computation that for $i \geq 2$,

$$\frac{\lambda_{i,+}}{\phi_{i,+}} + \frac{\lambda_{i,-}}{\phi_{i,-}} = 0.$$

Therefore, when restricting to the ω part in x, we have

$$M^{1/2}\omega(t) = \frac{\hat{p}_1}{\gamma}v_1 - \frac{\hat{p}_1}{\gamma}e^{-\gamma t}v_1 + \sum_{i=2}^{n} \frac{\lambda_{i,+}\hat{p}_i}{\phi_{i,+}} e^{\phi_{i,+}t} v_i + \sum_{i=2}^{n} \frac{\lambda_{i,-}\hat{p}_i}{\phi_{i,-}} e^{\phi_{i,-}t} v_i.$$

This, together with the fact

$$\frac{\lambda_{i,\pm}}{\phi_{i,\pm}} = \pm\frac{1}{\sqrt{\gamma^2 - 4\lambda_i}}$$

and the observation that

$$\frac{\hat{p}_1}{\gamma}v_1 - \frac{\hat{p}_1}{\gamma}e^{-\gamma t}v_1 = \frac{\hat{p}_1}{\sqrt{\gamma^2 - 4\lambda_1}}\left(e^{\phi_{1,+}t} - e^{\phi_{1,-}t}\right)v_1$$

completes the proof. □

Proof of Theorem 3.2.5

First consider $i \geq 2$. From the proof of Theorem 3.2.3, we know

$$\begin{bmatrix} M^{-1/2}v_i \\ 0 \end{bmatrix} = \lambda_{i,+}z_{i,+} + \lambda_{i,-}z_{i,-}. \tag{3.10}$$

This, together with the calculation in (3.9), implies that for input signal of the form $p_i(t)M^{1/2}v_i$, the system response of (3.2) is given as

$$x(t) = \lambda_{i,+}\int_0^t e^{\phi_{i,+}(t-\tau')}p_i(\tau')d\tau' z_{i,+} + \lambda_{i,-}\int_0^t e^{\phi_{i,-}(t-\tau')}p_i(\tau')d\tau' z_{i,-}.$$

For $p_i(t) = e^{\tau t}$, we then have

$$x(t) = \frac{\lambda_{i,+}}{\tau - \phi_{i,+}}\left(e^{\tau t} - e^{\phi_{i,+}t}\right)z_{i,+} + \frac{\lambda_{i,-}}{\tau - \phi_{i,-}}\left(e^{\tau t} - e^{\phi_{i,-}t}\right)z_{i,-}$$

and therefore when restricting to the frequency trajectory, we have

$$M^{1/2}\omega(t) = \frac{\lambda_{i,+}(\tau - \phi_{i,-}) + \lambda_{i,-}(\tau - \phi_{i,+})}{\tau^2 - (\phi_{i,+} + \phi_{i,-})\tau + \phi_{i,+}\phi_{i,-}}e^{\tau t}v_i - \left(\frac{\lambda_{i,+}e^{\phi_{i,+}t}}{\tau - \phi_{i,+}} + \frac{\lambda_{i,-}e^{\phi_{i,-}t}}{\tau - \phi_{i,-}}\right)v_i.$$

Noting $\lambda_{i,+} + \lambda_{i,-} = 1$, $\lambda_{i,+}\phi_{i,-} + \lambda_{i,-}\phi_{i,+} = 0$, $\phi_{i,+} + \phi_{i,-} = -\gamma$ and $\phi_{i,+}\phi_{i,-} = \lambda_i$, and dropping transient terms, we see

$$\overline{\omega}(t) = \frac{\tau}{\tau^2 + \gamma\tau + \lambda_i}e^{\tau t}M^{-1/2}v_i.$$

For $i = 1$, we do not need to decompose the signal as in (3.10), and a similar calculation leads to

$$\overline{\omega}(t) = \frac{e^{\tau t}M^{-1/2}v_1}{\tau + \gamma} = \frac{\tau}{\tau^2 + \gamma\tau + \lambda_1}e^{\tau t}M^{-1/2}v_1,$$

where the last equality is because $\lambda_1 = 0$. □

Chapter 4

CONTROLLABILITY AND OBSERVABILITY UNDER LIMITED CONTROLLER AND SENSOR COVERAGE

In this chapter, we switch our focus to a more combinatorial problem on the transient state of swing dynamics that aims to determine when the system is controllable/observable if we can only install controllers/sensors at a subset of the buses. We show that the controllability/observability of the swing dynamics in this setting is precisely characterized by two conditions: (a) intrinsic topological properties of the transmission network; and (b) algebraic coverage, which we define in Section 4.2, of buses with controllers/sensors. Condition (a) encodes information on graph symmetry, and is shown to hold for almost all practical systems. Condition (b) captures how buses interact with each other through the network, and can be verified using the eigenvectors of the graph Laplacian matrix.

The formal conditions we devise on the controllability of swing dynamics is presented in Section 4.2, and we explain the practical interpretations of these conditions in Section 4.3. The parallel results in system observability are given in Section 4.4. We present two applications of our characterizations in Section 4.5. The first application is more analytical, which reduces the problem of optimal placement for controllers and sensors to a set cover problem. The second application is an evaluation in the IEEE 39-bus New England interconnection test system, showing how a single well chosen critical bus based on our theory is capable of regulating the frequency of the whole grid.

4.1 System Model

In this section, we present the system model studied in this chapter, which is an extension of the swing dynamics (2.1) to include the limited coverage of controllers and sensors.

More specifically, we introduce three sets of buses as follows, and augment (2.1) to explicitly reflect how these sets impact the system dynamics.

1. **Controllers**. We assume that only a subset of the buses is equipped with controllers, which is captured by the corresponding d_j's. The set of buses with controllers is denoted as $\mathcal{U}$.

2. **Frequency Sensitive Component**. We assume that only a subset of buses has components that are sensitive to local frequency deviations. The injections from such components are captured by the corresponding $D_j\omega_j$'s. We do not allow direct control to such loads, and denote the set of buses with frequency sensitive loads as $\mathcal{F}$.

3. **Sensor**. We assume that only a subset of buses are equipped with components that measure the local frequency deviation ω_j. The set of such buses is denoted as $\mathcal{S}$.

With these notations, the swing dynamics (2.1) can be written as

$$\begin{aligned} M_j\dot{\omega}_j &= r_j - 1_{\mathcal{U}}(j)d_j - 1_{\mathcal{F}}(j)D_j\omega_j - \sum_{e\in\mathcal{E}} C_{je}f_e, && j \in \mathcal{N} \\ \dot{f}_{ij} &= B_{ij}(\omega_i - \omega_j), && (i,j) \in \mathcal{E}, \end{aligned}$$

and the system state is observed through

$$y_j = 1_{\mathcal{S}}(j)\omega_j, \qquad j \in \mathcal{N},$$

where $1_{\mathcal{U}}, 1_{\mathcal{F}}, 1_{\mathcal{S}}$ are the indicator functions of the subscript sets, e.g.,

$$1_{\mathcal{U}}(j) = \begin{cases} 1, & j \in \mathcal{U} \\ 0, & \text{otherwise.} \end{cases}$$

Now using x to denote the system state $x = [\omega; f]$, and putting F, U, S, M, D and B to be the diagonal matrices with $1_{\mathcal{F}}(j)$, $1_{\mathcal{U}}(j)$, $1_{\mathcal{S}}(j)$, M_j, D_j and B_{ij} as diagonal entries respectively, we can rewrite the system dynamics in the state-space form

$$\dot{x} = Ax - \begin{bmatrix} M^{-1}U \\ 0 \end{bmatrix} d + \begin{bmatrix} M^{-1} \\ 0 \end{bmatrix} r \tag{4.1a}$$

$$y = \begin{bmatrix} S & 0 \end{bmatrix} x, \tag{4.1b}$$

where

$$A = \begin{bmatrix} -M^{-1}FD & -M^{-1}C \\ BC^T & 0 \end{bmatrix}$$

and is referred to as the system matrix of (4.1) in the sequel.

In this chapter, we need the scaled graph Laplacian matrix defined as $\tilde{L} = M^{-1/2}LM^{-1/2}$, which is more explicitly given by

$$\tilde{L}_{ij} = \begin{cases} -\frac{B_{ij}}{\sqrt{M_iM_j}} & i \neq j, (i,j) \in \mathcal{E} \text{ or } (j,i) \in \mathcal{E} \\ \frac{1}{M_i}\sum_{j:j\in N(i)} B_{ij} & i = j \\ 0 & \text{otherwise,} \end{cases}$$

where $N(i)$ is the set of neighbors of i. To simplify the notations, we drop the tilde and use L to denote the scaled Laplacian $\tilde{L}$ henceforth in this chapter. Let $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of L, and put $V := \{v_1, v_2, \cdots, v_n\}$ to be an orthonormal set of its eigenvectors with v_s affording λ_s. The notation $\mathcal{N} = \{1, 2, \ldots, n\}$ is abused to also denote the index set of V. Whether $\mathcal{N}$ denotes the set of buses or denotes an index set for V will be clear from the context. The following property of the spectrum of L is particularly useful in this chapter:

Definition 4.1.1. *The matrix L is said to have a* **simple spectrum** *if all the eigenvalues of L are distinct.*

Throughout the analysis, we make the following assumption:

Sensitive Load: Frequency sensitive components only exist at buses with controllers. That is, we assume $\mathcal{F} \subset \mathcal{U}$.

4.2 Controllability

In this section, we analyze the state-space dynamics given in (4.1) and characterize its controllability using the spectrum of the scaled Laplacian matrix L.

Before presenting our characterization, we first clarify what we mean by the controllability of (4.1). The classical definition of controllability requires the whole state space $\mathbb{R}^{n+m}$ to be reachable from any initial point. This is too strong and is not suitable for our purpose. Indeed, from the branch flow dynamics

$$\dot{f} = BC^T \omega$$

we see that

$$B^{-1}\left(f(t) - f(0)\right) = \int_0^t C^T \omega(s) ds \in \text{range}(C^T).$$

If we assume the system is in the nominal state at time $t = 0$, that is $x(0) = [\omega(0); f(0)] = 0$, then we know $B^{-1} f(t) \in \text{range}(C^T)$ for any t. In other words, the scaled branch flow vector is confined in the range of C^T because of the system physics. This motivates the following definition:

Definition 4.2.1. *The dynamics* (4.1) *is said to be* **P-controllable** *or* **controllable in power system sense** *if for any $t > 0$, initial state $x(0) = [\omega(0); f(0)]$ and target state $x(t) = [\omega(t); f(t)]$ satisfying*

$$B^{-1}\left(f(t) - f(0)\right) \in \text{range}(C^T)$$

there exists a control u such that

$$x(t) = \phi(x(0), u, t),$$

where $\phi(x(0), u, t)$ is the system state at time t given initial state $x(0)$ and control input u.

Our first result generalizes the classical Kalman criteria for controllability testing to the context of P-controllability. It shows that to determine the system P-controllability, it suffices to form the controllability matrix with the scaled Laplacian matrix L (instead of the full system matrix A), and we can ignore the drift term r (even when it is time-variant) in (4.1a).

Proposition 4.2.2. *The dynamics* (4.1) *is P-controllable if and only if the matrix*

$$W = \begin{bmatrix} M^{-1/2}U & -LM^{-1/2}U & \cdots & (-L)^{n-1}M^{-1/2}U \end{bmatrix}$$

has full row rank.

The proof of this proposition is presented in Section 4.7. This result tells us that to decide the P-controllability of (4.1) amounts to computing the rank of W. Recall that $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$ are the eigenvalues of L, and $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal set of corresponding eigenvectors. Let Q be the matrix with v_j's as columns and let Λ be the diagonal matrix with λ_j's as diagonal entries, i.e., $L = Q\Lambda Q^T$. We introduce the concept of algebraic coverage.

Definition 4.2.3. *With all previous notations, the* **algebraic coverage** *of a bus $j \in \mathcal{N}$, denoted as* $\mathrm{cov}(j)$*, is defined to be the set*

$$\mathrm{cov}(j) := \left\{ s \in \mathcal{N} : v_{s,j} \neq 0 \right\},$$

where $v_{s,j}$ is the j-th entry of v_s.

Now we are ready to give our characterization for the P-controllability of (4.1).

Theorem 4.2.4. *With all the previous notations, the dynamics* (4.1) *is P-controllable if and only if*

1. *The scaled Laplacian matrix L has a simple spectrum; and*

2. *The algebraic coverage from controllers is full:*

$$\mathcal{N} = \bigcup_{j \in \mathcal{U}} \operatorname{cov}(j). \tag{4.2}$$

Proof. Recall U is the diagonal matrix encoding the placement of controllers. Let V be the Vandermonde matrix

$$V = \begin{bmatrix} 1 & -\lambda_1 & \lambda_1^2 & \cdots & (-\lambda_1)^{n-1} \\ 1 & -\lambda_2 & \lambda_2^2 & \cdots & (-\lambda_2)^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -\lambda_n & \lambda_n^2 & \cdots & (-\lambda_n)^{n-1} \end{bmatrix}$$

and

$$u_j = Q^T M^{-1/2} U_j, \quad \forall j \in \mathcal{N},$$

where U_j is the j-th column of U.

Since Q is orthogonal, we know $(-L)^k = Q(-\Lambda)^k Q^T$, and as a result,

$$W = Q \left[Q^T M^{-1/2} U \quad \cdots \quad (-\Lambda)^{n-1} Q^T M^{-1/2} U \right]. \tag{4.3}$$

For any integer p, q, we denote as $r(p, q)$ the unique number $r \in \{1, 2, \cdots, q\}$ such that $p = qk + r$ for some integer k. Define a permutation matrix $\Pi \in \mathbb{R}^{n^2 \times n^2}$ given by

$$\Pi_{ij} = \begin{cases} 1 & j = (r(i, n) - 1)n + \lfloor (i-1)/n \rfloor + 1 \\ 0 & \text{otherwise,} \end{cases}$$

Intuitively, multiplying Π on the right hand side of W collects columns corresponding to each u_j together. With such notations, we have

$$\begin{aligned} &\left[Q^T M^{-1/2} U \quad \cdots \quad (-\Lambda)^{n-1} Q^T M^{-1/2} U \right] \Pi \\ = &\left[\operatorname{diag}(u_1) \quad \cdots \quad \operatorname{diag}(u_n) \right] (I_n \otimes V), \end{aligned} \tag{4.4}$$

where $\otimes$ means the tensor multiplication and (4.4) can be checked by directly comparing each component. Noting both Q and Π are invertible, from (4.3) we know the rank of W is the same as the rank of (4.4). Therefore, by Proposition 4.2.2 we see the dynamics (4.1) is P-controllable if and only if (4.4) has full rank. It is well-known that the determinant of the Vandermonde matrix V is given by

$$\det(V) = \prod_{i<j} \left[(-\lambda_i) - (-\lambda_j) \right]$$

and therefore V is invertible if and only if the tuple $(\lambda_s : s \in \mathcal{N})$ has distinct entries. Also it is easy to see that $I_n \otimes V$ has full rank if and only if V has full rank; thus $I_n \otimes V$ is invertible if and only if L has simple spectrum.

Next, it can be checked that the nonzero rows of

$$\begin{bmatrix} \mathrm{diag}(u_1) & \mathrm{diag}(u_2) & \cdots & \mathrm{diag}(u_n) \end{bmatrix} \tag{4.5}$$

are independent because their nonzero entries appear in "orthogonal" positions. Therefore, (4.5) has full row rank if and only if all the rows have nonzero entries, or in other words,

$$\mathcal{N} = \bigcup_{j \in \mathcal{N}} \mathrm{supp}(u_j),$$

where $\mathrm{supp}(u_j)$ is the support of u_j.

From $u_j = Q^T M^{-1/2} U_j$ we can compute

$$u_{j,s} = \begin{cases} \frac{v_{s,j}}{\sqrt{M_j}} & j \in \mathcal{U} \\ 0 & j \notin \mathcal{U}, \end{cases}$$

from which we see

$$\mathrm{supp}(u_j) = \begin{cases} \left\{ s \in \mathcal{N} : v_{s,j} \neq 0 \right\} & j \in \mathcal{U} \\ \emptyset & j \notin \mathcal{U}. \end{cases}$$

But this implies

$$\bigcup_{j \in \mathcal{N}} \mathrm{supp}(u_j) = \bigcup_{j \in \mathcal{U}} \left\{ s \in \mathcal{N} : v_{s,j} \neq 0 \right\} = \bigcup_{j \in \mathcal{U}} \mathrm{cov}(j).$$

As a result, (4.4) has full rank if and only if L has simple spectrum and (4.2). This completes the proof. □

4.3 Interpretations

The characterization given in Theorem 4.2.4 is purely algebraic. In this section we explain the practical meanings of our controllability criteria.

L has a Simple Spectrum Almost Surely

Recall our mention in the introductory chapter that for the graph in Figure 1.2, the system (4.1) cannot be controllable because of the network symmetry. It turns out that L having a simple spectrum roughly means that the associated graph has few symmetries, and that this condition can be interpreted as a general criterion of whether the network topology is too symmetric that it loses controllability. Indeed, it is proven in [3] that if L has a simple spectrum, then any nontrivial automorphism of $\mathcal{G}$ has order two[1]. This specifically rules out star graphs with more than three nodes and symmetric weights (that is M_j's are the same for all $i \in \mathcal{N}$, and B_e's are the same for all $e \in \mathcal{E}$), including the graph in Figure 1.2, from having simple spectra. As another example, it can be shown that the Laplacian matrices associated with line graphs (under arbitrary B and M), which intuitively have only one symmetry, always have simple spectra [14]. For more results relating properties of the graph automorphism group to the spectrum of L, we refer the readers to [3, 14, 20].

This condition is much less restrictive than one would expect. In fact, one can check that the associated L matrix for all test cases coming with the Matpower 6.0 package [85] (including almost all IEEE and RTE testbeds) has simple spectra. We now establish a density result to explain such abundance of practical systems with simple spectra.

Consider a fixed transmission network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with line susceptance matrix B and inertia matrix M. Let

$$\Omega = \prod_{e \in \mathcal{E}} (-B_e, \infty)$$

be the space of feasible perturbations to B (so that we have positive line susceptances). We add a random perturbation $\omega \in \Omega$ drawn according to certain probability measure μ to the line susceptances, which can come from either measurement noise or manufacturing error, and consider the resulting scaled Laplacian matrix

$$L(\omega) = M^{-1/2} C (B + \mathrm{diag}(\omega)) C^T M^{-1/2}.$$

The following result shows that for a large family of perturbation distributions, $L(\omega)$ has simple spectrum almost surely.

Proposition 4.3.1. *Let $\mathcal{L}_m$ be the Lebesgue measure over $\mathbb{R}^m$. Assume the probability measure $\mu : \Omega \to [0, 1]$ is absolutely continuous with respect to $\mathcal{L}_m$. Let*

$$E := \{\omega \in \Omega : L(\omega) \text{ has simple spectrum}\}.$$

Then $\mu(E) = 1$.

[1]The result in [3] requires the assumption $M = I_m$. One can, however, prove similar results for general M by assigning M_j as node weights and requiring an automorphism to preserve both line and node weights.

The proof of this result is presented in Section 4.7. Note by the Radon-Nikodym Theorem [60], a probability measure is absolutely continuous with respect to the Lebesgue measure if and only if it affords a probability density function. Thus, for almost all practical probability models of such perturbation (e.g., truncated Gaussian noise with arbitrary covariance; bounded uniform distribution; truncated Laplace distribution), $L(\omega)$ has a simple spectrum almost surely. Similar perturbation results on M also hold.

Therefore, under mild assumptions on perturbations to the system parameters, the L matrix associated with a practical system almost always has a simple spectrum.

The Algebraic Coverage of Controllers should be Full

Intuitively, the algebraic coverage of a bus j reflects the set of eigenvectors of L (which are usually interpreted as the spectrum of the network graph $\mathcal{G}$ in spectral graph theory) that bus j can "interact" with. When the algebraic coverage of controllable buses is full, the control signals can interact with the entire spectrum of the network, and thus are able to drive the system to any state. As an illustration of this intuitive meaning of the algebraic coverage, we present an alternative interpretation for entries in the pseudo-inverse of L, and demonstrate that such interpretation is natural in certain scenarios. Fix two buses $i, j \in \mathcal{N}$. For each $s \in \mathcal{N}$, we put $\kappa_{ij}^s = v_{s,j} v_{s,k}$, which can be interpreted as the "mutual influence" between i and j through the spectrum s. We have $\kappa_{ij}^s \neq 0$ if and only if $s \in \mathrm{cov}(i) \cap \mathrm{cov}(j)$; and when this holds, s lies in the common coverage of i and j, and thus is a "bridging" spectrum. Recall that $L = Q \Lambda Q^T$ and therefore $L^\dagger = Q \Lambda^\dagger Q^T$, which then implies

$$L_{ij}^\dagger = \sum_{s \in \mathrm{cov}(i) \cap \mathrm{cov}(j), s \neq 1} \frac{\kappa_{ij}^s}{\lambda_s}.$$

In other words, $L_{ij}^\dagger$ can be interpreted as the weighted average of "mutual influence" between i and j over all "bridging" spectra.

4.4 Observability

In this section, we present our characterization of the observability of (4.1). The development in this section is parallel to Section 4.2, and thus we omit all proofs.

As in the case of controllability, the classical definition of observability is too strong. A more suitable notion of observability in our applications is given as follows:

Definition 4.4.1. *The dynamics* (4.1) *is said to be* **P-observable** *or* **observable in**

power system sense *if for any $t > 0$, an initial state $x(0) = [\omega(0); f(0)]$ such that*

$$B^{-1}\left(f(t) - f(0)\right) \in \mathrm{range}(C^T)$$

can be uniquely determined from the system input $d(s)$ and output $y(s)$ over $0 < s \le t$.

We can then give our characterization for the P-observability as follows:

Theorem 4.4.2. *The dynamics* (4.1) *is P-observable if and only if*

1. *The scaled Laplacian matrix L has simple spectrum; and*

2. *The algebraic coverage from sensors is full:*

$$\mathcal{N} = \bigcup_{j \in \mathcal{S}} \mathrm{cov}(j).$$

The second item in this criteria for observability again confirms our intuition that algebraic coverage encodes information on how buses interact with each other through the network.

4.5 Applications

In this section, we present two applications of our results. The first application is on the optimal placement of controllable loads/sensors so that controllability/observability of (4.1) is achieved. We show that this problem can be reduced to a set cover problem. The second application is over the IEEE 39-bus New England interconnection test system, where we demonstrate that a single critical bus chosen based on our theory is capable of regulating the frequency of the whole grid.

Optimal Placement of Controllers and Sensors

Given a power transmission network $\mathcal{G}$, if the associated L matrix does not have a simple spectrum, then by Theorem 4.2.4 and Theorem 4.4.2, such intrinsic deficiency of $\mathcal{G}$ forbids the dynamics (4.1) from being controllable/observable, no matter how many controllers or sensors we install. Fortunately, as Proposition 4.3.1 suggests, such deficiency usually does not occur for practical systems.

Now assume $\mathcal{G}$ has a simple spectrum. By Theorem 4.2.4, the dynamics (4.1) is P-controllable if and only if the union of algebraic coverage from controllable buses is full. Therefore the problem of choosing the minimum set of buses to place controllers such that (4.1) is P-controllable can be formulated as

$$\min_{J} \quad |J| \tag{4.6a}$$

$$\text{s.t.} \quad \bigcup_{j\in J} \mathrm{cov}(j) = \mathcal{N}. \tag{4.6b}$$

This is an instance of the well-studied set cover problem, one of Karp's 21 NP-complete problems [22]. Although Theorem 4.2.4 does not completely resolve (4.6), it shows that approximation algorithms devised for set cover problems can be readily applied to our setting to obtain placements with good quality.

A similar argument applied to the P-observability of (4.1) leads to the same optimization problem as (4.6). Therefore we are led to the following corollary, which is intuitive but non-trivial without Theorem 4.2.4 and Theorem 4.4.2.

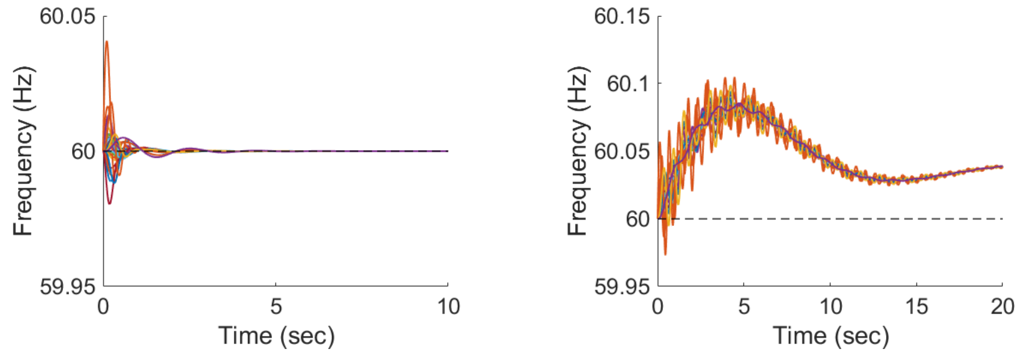
Corollary 4.5.1. *For the dynamics* (4.1)*, the collection of optimal placement sets of controllers and the collection of optimal placement sets of sensors are the same.*

This result tells us that, in practice, we should always install sensors at the buses with controllers, and vice versa.

Secondary Frequency Regulation with a Single Bus

We now demonstrate how our results can identify critical buses for controllability by evaluating the IEEE 39-bus New England interconnection test system, as shown in Figure 3.2. There are 10 generators and 29 load nodes in the system, and in contrast to our linearized model for theoretical study, the simulation adopts more realistic nonlinear dynamics.

One can check that the L matrix associated with this network has a simple spectrum (which is as expected according to Proposition 4.3.1), and that bus 35 has full algebraic coverage, i.e., all the eigenvectors v_s of L have nonzero entry at position 35. Therefore, Theorem 4.2.4 implies that even if we can only inject control at bus 35, the system is still P-controllable. Thus, we should be able to drive the whole system back to the nominal state after arbitrary disturbance. In order to verify this, we add a step increase of 1 pu to the generation at bus 30, and compare the system evolution with or without control at bus 35. In contrast to the standard control associated with the controllability Gramian, the control we adopt here utilizes only local frequency deviation. Details about the control scheme design can be found in [52]. The simulation results are shown in Figure 4.1.



(a) With control. (b) Without control.

Figure 4.1: Comparison of the system evolution with and without control at bus 35 after adding a step increase of 1 pu to the generation at bus 30.

As one can see from the figure, despite the geographical distance between the disturbance and the controllable node, the control scheme successfully drives the grid back to nominal state within 5 seconds. In contrast, when no control is posed, the bus frequencies still stabilize because of governor dynamics, but not to the nominal state. Moreover, the stabilization process takes considerably longer time. Such difference demonstrates that with a single bus 35 chosen based on our theory, frequency regulation over the grid can be achieved.

4.6 Conclusion

In this chapter, we developed full characterizations of the impact of limited controllers/sensors coverage over the controllability/observability for the swing dynamics. We presented two applications of our theoretical results: (a) an analytical application which reduces the problem of optimal placement of controllable loads and sensors to a set cover problem; (b) an evaluation over the IEEE 39-bus New England interconnection test system where secondary frequency control over the whole network can be achieved by a single critical bus chosen based on our theory.

4.7 Proofs

Proof of Proposition 4.2.2

Consider the vector space $Z := \mathbb{R}^n \times \mathrm{range}(BC^T) \subset \mathbb{R}^{n+m}$. Then from the definition we see that (4.1) is P-controllable if and only if the affine space

$$Z + [0; P(0)] = Z + x(0)$$

is reachable from $x(0)$ (the equality is because $[\omega(0); 0] \in Z$). Denote the set of Lebesgue integrable functions from $[0, t]$ to $\mathbb{R}^n$ as $I_{[0,t]}$. Now for any control input $d \in I_{[0,t]}$, the solution to (4.1a) is given by

$$\begin{aligned} x(t) = &-\int_0^t e^{(t-s)A} \begin{bmatrix} M^{-1}U \\ 0 \end{bmatrix} d(s)ds \\ &+ \int_0^t e^{(t-s)A} \begin{bmatrix} M^{-1} \\ 0 \end{bmatrix} r(s)ds + e^{tA}x(0). \end{aligned} \tag{4.7}$$

We will show Proposition 4.2.2 by inspecting each term in (4.7).

Lemma 4.7.1. *Let*

$$R_1 := \left\{ \int_0^t e^{(t-s)A} \begin{bmatrix} M^{-1}U \\ 0 \end{bmatrix} d(s)ds : d \in I_{[0,t]} \right\}$$

be the set of possible values of the first term in (4.7)*. Then R_1 is a subspace of Z and $R_1 = Z$ if and only if*

$$W = [M^{-1/2}U, -LM^{-1/2}U, \cdots, (-L)^{n-1}M^{-1/2}U]$$

has full row rank.

Proof. Since the set of buses with frequency sensitive components is contained in the set of buses with controllable loads, we can absorb the $D_j\omega_j$ term into d_j for all $j \in \mathcal{F}$ without affecting the system controllability. Therefore we can assume $F = 0$ in (4.1). Now, by induction we can compute $A^{(2k)}$ to be

$$\begin{bmatrix} M^{-1/2}(-L)^k M^{1/2} & 0 \\ 0 & (-BC^T M^{-1}C)^k \end{bmatrix}$$

and compute $A^{(2k+1)}$ to be

$$\begin{bmatrix} 0 & -M^{-1}C(-BC^T M^{-1}C)^k \\ BC^T M^{-1/2}(-L)^k M^{1/2} & 0 \end{bmatrix}.$$

Put $\tilde{B} := [M^{-1}U; 0]$. It is a classical result that R_1 is the same as the range of the controllability matrix of (4.1) given as

$$\tilde{W}' = \begin{bmatrix} \tilde{B} & A\tilde{B} & \cdots A^{n+m-1}\tilde{B} \end{bmatrix}.$$

Let

$$\tilde{W} = \begin{bmatrix} \tilde{B} & A\tilde{B} & \cdots A^{2(n+m-1)}\tilde{B} \end{bmatrix}.$$

Then, by the Cayley-Hamilton Theorem, we see that range$(\tilde{W}')$ = range$(\tilde{W})$. Multiplying $\tilde{B}$ to the powers of A and discarding the zero columns, we see the range of $\tilde{W}$ is equal to the range of

$$\begin{bmatrix} \tilde{M}\tilde{U} & 0 & -\tilde{M}L\tilde{U} & \cdots & 0 \\ 0 & BC^T\tilde{M}\tilde{U} & 0 & \cdots & BC^T\tilde{M}(-L)^{n+m-1}\tilde{U} \end{bmatrix}, \tag{4.8}$$

where $\tilde{M} := M^{-1/2}$ and $\tilde{U} := M^{-1/2}U$. Since BC^T is a common factor for the last m rows, we see the range of (4.8), and thus R_1, is a subspace of Z.

One can check that the dimension of Z is $2n-1$. Since R_1 is a subspace of Z, we know $R_1 = Z$ if and only if the rank of (4.8) is $2n-1$. Now define

$$W' = [\tilde{U}, -L\tilde{U}, L^2\tilde{U}, \cdots, (-L)^{n+m-1}\tilde{U}].$$

Since both $\tilde{M}$ and B are invertible, it is easy to see that the rank of (4.8) is given by rank(W') + rank(C^TW'). Moreover, from rank$(W') \le n$ and rank$(C^TW') \le$ rank$(C^T) \le n-1$, we know the matrix in (4.8) has rank $2n-1$ if and only if

$$\text{rank}(W') = n, \qquad \text{rank}(C^TW') = n-1.$$

Noting rank$(C^TW') = n-1$ is equivalent to having rank$(W') = n$, we thus see rank$(\tilde{W}) = 2n-1$ if and only if rank$(W') = n$. Finally by the Cayley-Hamilton Theorem, the range of W' is the same as the range of

$$W = [\tilde{U}, -L\tilde{U}, L^2\tilde{U}, \cdots, (-L)^{n-1}\tilde{U}].$$

In other words, we have shown that $R_1 = Z$ if and only if W has full row rank, and the desired result thus follows. □

Lemma 4.7.2. *For arbitrary $r \in I_{[0,t]}$, we have*

$$\int_0^t e^{(t-s)A}\begin{bmatrix} M^{-1} \\ 0 \end{bmatrix} r(s)ds \in Z.$$

Proof. The proof is similar to the part where we show $R_1 \subset Z$ in Lemma 4.7.1 and we omit the details here. □

Lemma 4.7.3. *For arbitrary $x(0) \in \mathbb{R}^{n+m}$, we have*

$$e^{tA}x(0) - x(0) \in Z.$$

Proof. It is easy to check that the convergence radius of $g(x) := \sum_{i=1}^{\infty} \frac{x^{i-1}}{i!}$ is infinite and thus the matrix series

$$g(A) := \sum_{i=1}^{\infty} \frac{A^{i-1}}{i!}$$

converges and is well-defined. Now note

$$e^{tA} - I = \sum_{i=1}^{\infty} \frac{A^i}{i!} = Ag(A),$$

thus $e^{tA}x(0) - x(0) = Ag(A)x(0) \in \mathrm{range}(A)$. It is easy to check $\mathrm{range}(A) \subset Z$, which then implies $e^{tA}x(0) - x(0) \in Z$. $\square$

Now put

$$R := \{\phi(x(0), d, t) : d \in I_{[0,t]}\}$$

to be the set of reachable states from $x(0)$ according to (4.7). From the above lemmas, we see that for any control d, we have

$$\phi(x(0), d, t) - x(0) \in Z$$

and therefore $R - x(0) \subset Z$. Moreover, since

$$\int_0^t e^{(t-s)A} \begin{bmatrix} M^{-1} \\ 0 \end{bmatrix} P^m(s)ds + e^{tA}x(0) - x(0) \in Z,$$

we know $R - x(0) = Z$ if and only if $R_1 = Z$, which in turn is equivalent to W having full row rank. As a result, we see (4.1) is P-controllable if and only if W has full row rank. $\square$

Proof of Proposition 4.3.1

Fix a network $\mathcal{G}$ and the associated B, M matrices. For any $\omega \in \Omega$, let

$$\chi(t;\omega) := \det\left(M^{-1/2}C(B + \mathrm{diag}(\omega))C^T M^{-1/2} - tI\right)$$

be the characteristic polynomial of the perturbed Laplacian matrix $L(\omega)$ and let $\mathrm{disc}(\omega) := \mathrm{disc}\left(\chi(t;\omega)\right)$ be the discriminant of χ. Recall $\mathrm{disc}(\omega)$ is a polynomial in the coefficients of χ and therefore is a polynomial of the entries in ω. Moreover, $\mathrm{disc}(\omega) = 0$ if and only if χ has multiple roots, or equivalently $L(\omega)$ does not have a simple spectrum. Put

$$E := \{\omega \in \Omega : L(\omega) \text{ has a simple spectrum}\}.$$

We then have $E = \Omega \backslash \mathrm{disc}^{-1}(0)$.

Lemma 4.7.4. *The polynomial* $\mathrm{disc}(\omega)$ *is not identically zero.*

Proof. To show $\mathrm{disc}(\omega)$ is not identically zero, it suffices to show that we can find $\omega \in \Omega$ such that $\mathrm{disc}(\omega) \neq 0$, which is equivalent to the existence of a link susceptance matrix B_0 such that $L_0 := M^{-1/2} C B_0 C^T M^{-1/2}$ has a simple spectrum. We use mathematical induction to prove the existence of such B_0. To facilitate the discussion, for any subgraph $\mathcal{G}' = (\mathcal{N}', \mathcal{E}')$ of $\mathcal{G}$ with line susceptance matrix B', we refer to the matrix

$$L' = M'^{-1/2} C' B' (C')^T M'^{-1/2}$$

as the Laplacian matrix of $\mathcal{G}'$, where M' is the submatrix of M corresponding to the nodes in $\mathcal{N}'$, and C' is the incidence matrix of $\mathcal{G}'$.

First we pick a random link e_1 in $\mathcal{E}$ and assign arbitrary susceptance to e_1. The subgraph $\mathcal{G}_1$ generated by e_1 has only one connected component (which is formed by e_1 itself), and therefore the Laplacian matrix L_1 of $\mathcal{G}_1$ has one zero eigenvalue and one nonzero eigenvalue, which are distinct.

Next, consider any subgraph $\mathcal{G}_k = (\mathcal{N}_k, \mathcal{E}_k)$ of $\mathcal{G}$ with $k < m$ many links and let $e_k = (i_k, j_k) \in \mathcal{E} \backslash \mathcal{E}_k$ be a link of $\mathcal{G}$ not in $\mathcal{G}_k$ yet at least one of i_k and j_k is in $\mathcal{N}_k$. Assume $\mathcal{G}_k$ has a simple spectrum. We claim that by choosing the susceptance for e_k properly, the graph $\mathcal{G}_{k+1}$ obtained by adjoining e_k (and possibly one of the vertices i_k, j_k) to $\mathcal{G}_k$ still has a simple spectrum. Indeed, for the case where both i_k and j_k are in $\mathcal{N}_k$, we know the Laplacian matrix L_{k+1} for $\mathcal{G}_{k+1}$ is given as

$$\begin{aligned} L_{k+1} &= L_k + \left(-\frac{B_{i_k j_k}}{\sqrt{M_{i_k} M_{j_k}}} (E_{i_k j_k} + E_{j_k i_k}) + \frac{B_{i_k j_k}}{M_{i_k}} E_{i_k i_k} + \frac{B_{i_k j_k}}{M_{j_k}} E_{j_k j_k} \right) \\ &=: L_k + \Delta L_k, \end{aligned} \tag{4.9}$$

where for any i, j, E_{ij} is the matrix with 1 at the intersection of i-th row and j-th column and 0 otherwise. Let δ_k be the minimum gap between the eigenvalues of L_k. Choose $B_{i_k j_k}$ small enough so that the spectral norm of ΔL_k is less than $\delta_k/2$. Then, by Weyl's inequality we know that each eigenvalue of L_k is perturbed by at most $\delta_k/2$ from adding ΔL_k. As a result, the eigenvalues of L_{k+1} are still distinct.

If i_k is not in $\mathcal{N}_k$, then $L_{k+1} = \overline{L}_k + \Delta L_k$, where $\overline{L}_k$ is the matrix obtained from L_k by appending a row and a column of zeros, and ΔL_k is the same as in (4.9). It is easy to see that $\overline{L}_k$ and L_k share the same nonzero eigenvalues, and $\overline{L}_k$ has two zero eigenvalues. Similar to the previous case, by choosing $B_{i_k j_k}$ small enough, we can ensure that the distinct nonzero eigenvalues of $\overline{L}_k$ after perturbation of ΔL_k

are still distinct and nonzero. Note that L_{k+1} has only one zero eigenvalue, thus by choosing $B_{i_k j_k}$ even smaller if necessary, the new nonzero eigenvalue coming from the perturbation of ΔL_k can be made arbitrarily small, and thus distinct from other eigenvalues. We have thus justified our claim.

Now, by induction we see that we can always pick the line susceptances properly so that the resulting L has a simple spectrum. This completes the proof. $\square$

It is well-known from algebraic geometry that the root set of a polynomial which is not identically zero has Lebesgue measure zero [31]. In particular, for the polynomial $\mathrm{disc}(\omega)$ which is not identically zero, we have $\mathcal{L}_m\left(\mathrm{disc}^{-1}(0)\right) = 0$. Since μ is absolutely continuous with respect to $\mathcal{L}_m$, we see $\mu\left(\mathrm{disc}^{-1}(0)\right) = 0$ or equivalently, L has a simple spectrum with probability 1 as $E = \Omega \backslash \mathrm{disc}^{-1}(0)$. $\square$

Chapter 5

MONOTONICITY IN CASCADING FAILURES AND TREE-PARTITIONS

Starting with this chapter, we switch our focus to the steady state aspect of power grids, and study how we can improve the system robustness against cascading failures. In this chapter, we study a generic cascading failure process in power system, and show that the Laplacian matrix of the transmission network captures two useful structures in this context: (a) In contrast to the lack of monotonicity in the physical domain, there is a rich collection of monotonicity we can explore in the spectrum of the Laplacian matrix. This allows us to systematically design topological measures that are monotonic over the cascading event. (b) Power redistribution patterns are closely related to the distribution of different types of trees in the power network topology. Such graphical interpretations capture the Kirchhoff's Law in a precise way using topological properties, and do not rely on any assumptions or simplifications on the failures propagation dynamics.

In Section 5.2, we study the evolution of Laplacian matrix eigenvalues as the cascading failure unfolds, and demonstrate how properties of such eigenvalue dynamics suggest monotonicity in physical power flows. Later in Section 5.3, we present the key result that relates the power redistribution to graphical structures as given in Theorem 5.3.1, which states that the distribution of different types of trees in the transmission network fully determines the power redistribution patterns after a transmission line is tripped. The new characterization reveals a Simple Loop Criteria that fully determines whether the failure of one line can impact another line.

Our graphical interpretations of power redistribution naturally suggests that we can eliminate long-distance propagation of system disturbances by forming a *tree-partition*, which we define in Section 5.4, of the transmission network. We discuss the uniqueness of tree-partitions for a general network, and show that the "finest" tree-partition can be computed in linear time. In Section 5.5, we demonstrate how the transmission network tree-partition localizes the impacts of line failure to the region in which the failure happens. The rigorous proof of such localization properties and how the proven properties can be leveraged to provide analytical guarantees for failure mitigation are presented in Chapter 6 and Chapter 7, respectively.

5.1 System Model

In this section, we present our cascading failure model and introduce the line outage distribution factor that is extensively used in the following chapters.

Given a power network, we describe the cascading failure process by keeping track of the set of failed lines at different stages, which is indexed by $\mathfrak{N} := \{1, 2, \cdots, N\}$. Each stage $\mathfrak{n} \in \mathfrak{N}$ corresponds to a topology $\mathcal{G}(\mathfrak{n}) := (\mathcal{N}, \mathcal{E}\backslash\mathcal{B}(\mathfrak{n}))$, where $\mathcal{B}(\mathfrak{n})$ is the set of all tripped lines at stage n and is naturally nested:

$$\mathcal{B}(\mathfrak{n}) \subset \mathcal{B}(\mathfrak{n}+1), \quad \forall \mathfrak{n} \in \mathfrak{N}.$$

We denote the Laplacian matrix of $\mathcal{G}(\mathfrak{n})$ by $L(\mathfrak{n})$. Within each stage $\mathfrak{n}$, the power flow redistributes over the network described by $\mathcal{G}(\mathfrak{n})$ according to the DC power flow model (2.5). After the system stabilizes, if all the branch flows are below the corresponding line ratings, then the new operating point is secure and the cascade stops. Otherwise, let $\mathcal{F}(\mathfrak{n})$ be the subset of lines whose branch flows exceed the corresponding line ratings. The lines in $\mathcal{F}(\mathfrak{n})$ are tripped in stage $\mathfrak{n}$, i.e., $\mathcal{B}(\mathfrak{n}+1) = \mathcal{B}(\mathfrak{n}) \cup \mathcal{F}(\mathfrak{n})$. This process repeats for stage $\mathfrak{n}+1$ and so on.

Next, we focus on a fixed stage $\mathfrak{n} \in \mathfrak{N}$, and describe how the failure of a line impacts the branch flows on remaining lines. To simplify the notation, we drop the stage index $\mathfrak{n}$ from symbols like $\mathcal{G}(\mathfrak{n})$ and simply write them as $\mathcal{G}$. Given a network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, when a line e is tripped from $\mathcal{G}$, if the newly formed graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E}\backslash\{e\})$ is still connected, then it is a well-known result (see [76], for instance) that the branch flow change on a line $\hat{e}$ is given as

$$\Delta f_{\hat{e}} = f_e \times K_{e\hat{e}},$$

where $K_{e\hat{e}}$ is the *line outage distribution factor* [76] from e to $\hat{e}$. This distribution factor is independent of the original power injection p and can be computed from the matrices B and C [76].

If the new graph $\mathcal{G}'$ is disconnected, then it is possible that the original injection p is no longer balanced in the connected components of $\mathcal{G}'$. Thus, to compute the new power flow, a certain power balance rule $\mathcal{R}$ needs to be applied. Several such rules have been proposed and evaluated in the literature based on load shedding or generator response [9, 11, 13, 62]. We do not specialize with respect to any such rule, and instead opt to identify the key properties of these rules that allow our results to hold. With this more abstract approach, we can characterize the power flow redistribution under a broad class of power balance rules.

5.2 Monotonicity in Cascading Failures

In this section, we present our results for monotonicity in cascading failure processes. Our characterization is related to known monotonicity results and suggests a systematic way to define monotonic topological metrics over a failure event.

Our approach focuses on the Laplacian spectrum of the system. In contrast to the lack of monotonicity in the physical system, when we look at the process from this spectral perspective, there is a rich set of monotonicity one can explore. They are built upon the following fundamental monotonicity result:

Theorem 5.2.1. *Let $\lambda_1(\mathfrak{n}) \leq \lambda_2(\mathfrak{n}) \leq \cdots \leq \lambda_n(\mathfrak{n})$ be the eigenvalues of $L(\mathfrak{n})$. Then $\lambda_i(\mathfrak{n})$ is a decreasing function in $\mathfrak{n}$ for each i. Moreover, for each stage $\mathfrak{n}$, as long as new lines are tripped at $\mathfrak{n}$, there exists i such that the decrease is strict:*

$$\lambda_i(\mathfrak{n}+1) < \lambda_i(\mathfrak{n}).$$

The Laplacian eigenvalues encode information on how well the graph is connected and how fast information can propagate in the network (see [24], for example). Therefore, this result shows that as the cascading failure process unfolds there is a decreasing level on the network connectivity and its "mixing ability". Although Theorem 5.2.1 is only related to the network topology evolution, we demonstrate in Corollary 5.2.6 that by applying such monotonicity properly it is possible to devise monotonic properties that are directly related to the power flow dynamics.

To prove Theorem 5.2.1, we first derive an eigenvalue interlacing result for generic weighted Laplacian matrices. Its special case, in which the graph is unweighted and only a single line is removed, is known in the literature [32].

Proposition 5.2.2. *Let $\mathcal{G}$ be a weighted graph with positive line weights $\{w_e\}$, and let $\mathcal{H}$ be a subgraph of $\mathcal{G}$ obtained by removing exactly s edges from $\mathcal{G}$. Denote $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ to be the eigenvalues of $L_{\mathcal{G}}$ and $L_{\mathcal{H}}$, respectively. Then for any $k = 1, 2, \ldots, n$, we have*

$$\mu_k \leq \lambda_k \tag{5.1}$$

and for $k = s+1, s+2, \ldots, n$, we have

$$\lambda_{k-s} \leq \mu_k. \tag{5.2}$$

The proof of this proposition is presented in Section 5.7. As an immediate corollary, we can deduce the following well-known result for $s = 1$:

Corollary 5.2.3. *With previous notations, when $\mathcal{H}$ is obtained by removing a single edge from $\mathcal{G}$, we have*

$$\mu_1 \leq \lambda_1 \leq \mu_2 \leq \cdots \leq \lambda_{n-1} \leq \mu_n \leq \lambda_n.$$

We now apply Proposition 5.2.2 to the transmission network Laplacian matrices $\mathcal{G}(\mathfrak{n})$. Note that in a cascading process described by the graph sequence $\{\mathcal{G}(\mathfrak{n})\}_{\mathfrak{n}\in\mathfrak{N}}$, $\mathcal{G}(\mathfrak{n}+1)$ is obtained from $\mathcal{G}(\mathfrak{n})$ by removing the tripped lines $\mathcal{F}(\mathfrak{n})$ incurred during stage $\mathfrak{n}$; therefore, we know that the functions $\lambda_i(\mathfrak{n})$ as defined in Theorem 5.2.1 are monotonically decreasing. To finish the proof of Theorem 5.2.1, it thus suffices to show that for each stage $\mathfrak{n}$ we can always find an i such that the decrease is strict.

Proof of Theorem 5.2.1. This is immediate after noting

$$\begin{aligned}\sum_i \lambda_i(\mathfrak{n}+1) = \operatorname{tr}(L(\mathfrak{n}+1)) &= \sum_{e\in\mathcal{E}(\mathfrak{n}+1)} B_e \\ &< \sum_{e\in\mathcal{E}(\mathfrak{n})} B_e = \operatorname{tr}(L(\mathfrak{n})) \\ &= \sum_i \lambda_i(\mathfrak{n}),\end{aligned}$$

where the inequality is strict because there are lines tripped at time $\mathfrak{n}$. □

Such monotonicity of Laplacian eigenvalues suggests that all metrics measuring the system from its spectrum should be monotonic as well. The most general result we can conclude along this line is the following:

Corollary 5.2.4. *Let $|||\cdot|||$ be a unitarily-invariant norm on the set of $n \times n$ matrices. Then $|||L(\mathfrak{n})|||$ is a decreasing function of $\mathfrak{n}$.*

Proof. This is an immediate result from the bijective correspondence between unitarily invariant norms on $n \times n$ matrices and symmetric gauge functions applied to the matrix singular values [10], because symmetric gauge functions are monotone in the vector components. □

Examples of unitarily-invariant norms include the spectral norm, nuclear norm, Frobenious norm, Schatten p-norms, and Ky-Fan k-norms, etc., each of which suggests a different way to measure the system monotonicity. For example, the monotonicity in nuclear norm recovers the fact that the sum of all link reactances decreases in a cascading failure process.

It is well-known from singular value decomposition that the nonzero eigenvalues of $L^\dagger(\mathfrak{n})$ are given as $1/\lambda_i(\mathfrak{n})$, with the same corresponding eigenvectors as $L(\mathfrak{n})$. Therefore, Theorem 5.2.1 implies that the nonzero eigenvalues of $L^\dagger(\mathfrak{n})$ are monotonically increasing. It is tempting to conclude from this fact that $v^T L^\dagger(\mathfrak{n}) v$ is monotonically increasing for a fixed $v \in \mathbb{R}^n$, but the situation becomes tricky after we notice that the eigenvectors of $L(\mathfrak{n})$ also evolve with $\mathfrak{n}$. Fortunately, we can still prove such monotonicity with careful algebra.

Proposition 5.2.5. *For any $v \in \mathbb{R}^n$, the function $V(\mathfrak{n}) := v^T L^\dagger(\mathfrak{n}) v$ is increasing in the stage index $\mathfrak{n}$.*

Proof. Without loss of generality, let us assume there is only a single edge $e = (i, j)$ tripped at stage $\mathfrak{n}$. The general case follows by tripping the lines one by one.

Under this assumption, by direct computation we have

$$L(\mathfrak{n} + 1) = L(\mathfrak{n}) - B_e C_e C_e^T,$$

where C_e is the column of C corresponding to e. It is shown in [6] that this rank one perturbation translates in its Moore-Penrose pseudoinverse to the equation

$$L^\dagger(\mathfrak{n} + 1) = L^\dagger(\mathfrak{n}) + \frac{1}{X_{ij} - R_{ij}} L^\dagger(\mathfrak{n}) C_e C_e^T L^\dagger(\mathfrak{n}), \tag{5.3}$$

where R_{ij} is the effective reactance between bus i and j defined in Section 2.4.4. Recall by Corollary 2.4.6 we always have $X_{ij} - R_{ij} > 0$ for directly connected i and j (as long as after removing e the network is still connected), we thus see the second term in (5.3) is positive semidefinite. The monotonicity of $V(\mathfrak{n})$ then follows. □

The network tension [41] at stage $\mathfrak{n}$ is defined to be $H(\mathfrak{n}) = f(\mathfrak{n})^T X(\mathfrak{n}) f(\mathfrak{n})$, which measures the aggregate load of the network and is shown to be an increasing function of $\mathfrak{n}$ in [41]. We now show this is a special case of our result.

Corollary 5.2.6. *$H(\mathfrak{n})$ is an increasing function in $\mathfrak{n}$.*

Proof. We can calculate that (for notation simplicity, we drop the stage index $\mathfrak{n}$)

$$\begin{aligned} f^T X f &= p^T L^\dagger C B X B C^T L^\dagger p \\ &= p^T L^\dagger L L^\dagger p \\ &= p^T L^\dagger p. \end{aligned}$$

By Proposition 5.2.5 we then know $H(\mathfrak{n})$ is monotonically increasing. □

The equation (5.3) not only shows the monotonicity of $H(\mathfrak{n})$, but also implies that the increment of $H(\mathfrak{n})$ at each $\mathfrak{n}$ is inversely proportional to the amount of reactance reduction of (i, j) from the network at time $\mathfrak{n}$.

5.3 Line Outage Redistribution Factor

As we discussed in Section 5.1, when a line e is tripped from a power network $\mathcal{G}$, the line outage distribution factor $K_{e\hat{e}}$ captures the ratio between the branch flow change over line $\hat{e}$ with respect to the original branch flow on e before it is tripped. Writing $e = (i, j), \hat{e} = (w, z)$ with $i, j, w, z \in \mathcal{N}$ not being the slack bus, the constant $K_{e\hat{e}}$ can be computed as [76]

$$K_{e\hat{e}} = \frac{X_e}{X_{\hat{e}}} \cdot \frac{A_{iw} + A_{jz} - A_{jw} - A_{iz}}{X_e - (A_{ii} + A_{jj} - A_{ij} - A_{ji})},$$

where $A = (\overline{L})^{-1}$ is the inverse of the transmission network Laplacian matrix after removing the slack bus (as defined in Chapter 2). This formula only holds if the graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected, as otherwise its denominator is 0 by Corollary 2.4.6. Edges in $\mathcal{G}$ whose removal disconnects $\mathcal{G}$ into multiple connected components are known as *bridges* in the literature. We discuss this concept and its relationship to tree-partitions in more detail in Section 5.4.

Graphical Interpretation

Recall from Proposition 2.4.1 we know that elements of A are related to the distributions of different types of trees in the transmission network $\mathcal{G}$. We now follow the same approach and derive the following new formula for $K_{e\hat{e}}$:

Theorem 5.3.1. *Let $e = (i, j), \hat{e} = (w, z)$ be edges with $i, j, w, z \neq n$ such that $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected. Then $K_{e\hat{e}}$ is given by*

$$B_{\hat{e}} \times \frac{\sum_{E \in \mathcal{T}(\{i,w\},\{j,z\})} \chi(E) - \sum_{E \in \mathcal{T}(\{i,z\},\{j,w\})} \chi(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E} \setminus \{(i,j)\}}} \chi(E)}. \tag{5.4}$$

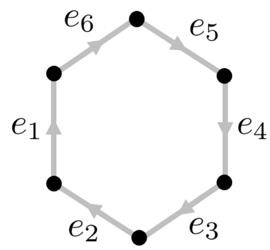


Figure 5.1: A ring network with clockwise orientation. Edge e_1 can only spread "negative" impacts to other lines.

Proof. This result follows from dividing the equation in Corollary 2.4.4 by the equation in Corollary 2.4.6. □

Similar to our discussion in Section 2.4 on generation shift sensitivity factors, each term in (5.4) also carries clear graphical meanings, as we now explain:

1. The numerator of (5.4) states that the impact of tripping e propagates to $\hat{e}$ through all possible trees that connect e to $\hat{e}$, counting orientation;
2. The denominator of (5.4) sums over all spanning trees of $\mathcal{G}$ that do not pass through $e = (i, j)$, and each tree of this type specifies an alternative path that power can flow through if (i, j) is tripped. When there are more trees of this type, the network has better ability in "absorbing" the impact of (i, j) being tripped, and the denominator of (5.4) precisely captures this effect by saying that the impact of e being tripped to other lines is inversely proportional to the sum of all alternative tree paths in the network;
3. The $B_{\hat{e}}$ constant in (5.4) captures the intuition that lines with larger reactance tend to be more robust against failures of other lines.

This graphical interpretation of (5.4) allows us to make general inferences on $K_{e\hat{e}}$ using only knowledge from the network topology. For example, in the ring network shown in Figure 5.1, by inspecting the graph we conclude that

$$K_{e_1 e_s} < 0, \quad s = 2, 3, 4, 5, 6,$$

as e_1 can only spread "negative" impacts to other lines (the positive term in (5.4) vanishes).

As an application of Theorem 5.3.1, we recover the following result from [41], whose original proof is much longer.

Corollary 5.3.2. *For adjacent lines $e = (i, j)$ and $\hat{e} = (i, k)$ with $i, j, k \neq n$, we have*

$$K_{e\hat{e}} \geq 0.$$

Proof. For such e and $\hat{e}$, the negative term in the numerator of (5.4) is over the empty set and thus equals to 0. □

Simple Loop Criteria

The formula (5.4) shows that the spanning forests

$$\mathcal{T}(\{i, w\}, \{j, z\}) \text{ and } \mathcal{T}(\{i, z\}, \{j, w\})$$

fully determine if $K_{e\hat{e}}$ is zero or not. In other words, whether tripping e has any impact on $\hat{e}$ depends on whether they can be connected via trees in $\mathcal{G}$. We now establish an equivalent criteria that is easier to verify from the graph.

If $K_{e\hat{e}} \neq 0$, then (5.4) implies that at least one of spanning forests set $\mathcal{T}(\{i, w\}, \{j, z\})$ and $\mathcal{T}(\{i, z\}, \{j, w\})$ is nonempty. Without loss of generality, let us assume

$$\mathcal{T}(\{i, w\}, \{j, z\}) \neq \emptyset.$$

For any element in $\mathcal{T}(\{i, w\}, \{j, z\})$ (see Figure 2.2), the tree containing $\{i, w\}$ induces a path from i to w, and the tree containing $\{j, z\}$ induces a path from j to z. By adjoining the edges $e = (i, j)$ and $\hat{e} = (w, z)$ to these two paths, we obtain a simple loop[1] containing both e and $\hat{e}$. As a result, $K_{e\hat{e}} \neq 0$ implies that we can find a simple loop in $\mathcal{G}$ which contains both e and $\hat{e}$.

The converse, unfortunately, in general does not hold because of certain systems with high symmetry. That is, there exist pathological systems where a simple loop containing e and $\hat{e}$ exists, yet $K_{e\hat{e}} = 0$. Nevertheless, for such systems, by perturbing the line susceptances B_e with an arbitrarily small noise (similar to Section 4.3), we can "break" the symmetry and show that $K_{e\hat{e}} \neq 0$ almost surely. The detailed technical treatments for this perturbation analysis are presented in Chapter 6.

The following proposition formally summarizes the discussions above[2]:

[1] A loop is simple if it visits each vertex at most once.

[2] We name this criteria the Simple Loop Criteria for two reasons: (a) The criteria is related to simple loops in $\mathcal{G}$; (b) This is a loop criteria that is simple.

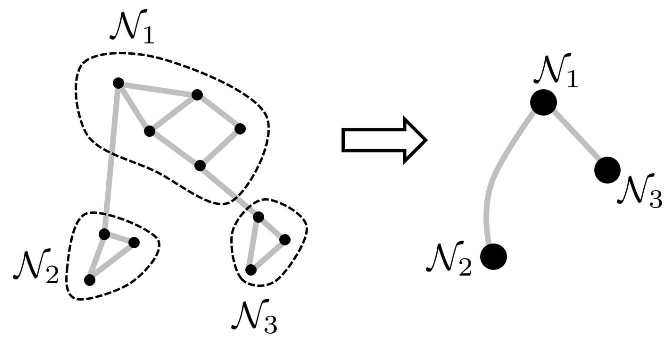


Figure 5.2: The construction of $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{P}$.

Proposition 5.3.3 (Simple Loop Criteria)**.** *For $e = (i, j), \hat{e} = (w, z) \in \mathcal{E}$ with $i, j, w, z \neq n$ such that $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \setminus \{e\})$ is connected, we have $K_{e\hat{e}} \neq 0$ "if" and only if there exists a simple loop in $\mathcal{G}$ that contains both e and $\hat{e}$.*

The "if" part of Proposition 5.3.3 should be interpreted as a probability one event under proper perturbations (see Chapter 6 for more details). This proposition shows that a simple loop containing e and $\hat{e}$ must exist if tripping e can possibly cause a successive failure of $\hat{e}$ since otherwise the branch flow on $\hat{e}$ is not impacted at all. As a result, by forming smaller management regions that are connected in a loop-free manner in the transmission network, we can prevent long-distance propagation of line failures. This motivates us to propose and study the *tree-partition* of a power grid, which we present in the next section.

5.4 Tree-partitions of Power Grids

In this section, we define the tree-partition motivated by the Simple Loop Criteria from the last section, discuss its uniqueness, and show that the "finest" tree-partition of a general graph can be computed in linear time.

For a power network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, a collection $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of subsets of $\mathcal{N}$ is said to form a *partition* of $\mathcal{G}$ if $\mathcal{N}_i \cap \mathcal{N}_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^{k} \mathcal{N}_i = \mathcal{N}$. For any partition, we can define a reduced multi-graph $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{G}$ as follows. First, we reduce each subset $\mathcal{N}_i$ to a super node (see Figure 5.2). The collection of all super nodes forms the node set for $\mathcal{G}_{\mathcal{P}}$. Second, we add an undirected edge connecting the super nodes $\mathcal{N}_i$ and $\mathcal{N}_j$ for each pair of $n_i, n_j \in \mathcal{N}$ with the property that $n_i \in \mathcal{N}_i$, $n_j \in \mathcal{N}_j$ and n_i and n_j are connected in $\mathcal{G}$. Note that multiple edges are added when multiple pairs of such n_i, n_j exist. Unlike the graph $\mathcal{G}$ to which we assign an arbitrary orientation (and thus is a directed graph), the reduced multi-graph $\mathcal{G}_{\mathcal{P}}$ is undirected.

Definition 5.4.1. *A partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$ is said to be a* **tree-partition** *if the reduced graph $\mathcal{G}_{\mathcal{P}}$ forms a tree. When this holds, the sets $\mathcal{N}_i$ are called the* **regions** *of $\mathcal{P}$. An edge $e = (w, z)$ with both endpoints inside $\mathcal{N}_i$ is said to be* **within** *$\mathcal{N}_i$. If e is not within $\mathcal{N}_i$ for any i, then we say e forms a* **bridge**[3].

Tree-partitions of a power network $\mathcal{G}$ are generally not unique. For instance, one can always collapse $\mathcal{G}$ into a single region with the partition $\mathcal{P}_0 = \{\mathcal{N}\}$, which is a trivial tree-partition of $\mathcal{G}$. This in particular yields a different tree-partition for the graph shown in Figure 5.2. Nevertheless, if we require the tree-partition to be as "fine" as possible, such a partition is unique.

More concretely, given a graph $\mathcal{G}$, we define a partial order $\succeq$ over the set of all tree-partitions of $\mathcal{G}$ (which is nonempty as it always contains the trivial partition $\mathcal{P}_0$) as follows: For two tree-partitions $\mathcal{P}^1 = \left\{\mathcal{N}_1^1, \mathcal{N}_2^1, \cdots, \mathcal{N}_{k_1}^1\right\}$ and $\mathcal{P}^2 = \left\{\mathcal{N}_1^2, \mathcal{N}_2^2, \cdots, \mathcal{N}_{k_2}^2\right\}$, we say $\mathcal{P}^1$ is *finer* than $\mathcal{P}^2$, denoted as $\mathcal{P}^1 \succeq \mathcal{P}^2$, if for any $i = 1, 2, \ldots, k_1$, there exists some $j(i) \in \{1, 2, \ldots, k_2\}$ such that $\mathcal{N}_i^1 \subset \mathcal{N}_{j(i)}^2$. That is, $\mathcal{P}^1$ is finer than $\mathcal{P}^2$ if each region in $\mathcal{P}^1$ is contained in some region in $\mathcal{P}^2$ (see Figure 5.3). It is routine to check that $\succeq$ defines a partial order over all possible tree-partitions of $\mathcal{G}$.

Definition 5.4.2. *A tree-partition $\mathcal{P}$ of $\mathcal{G}$ is said to be* **irreducible** *if $\mathcal{P}$ is maximal with respect to the partial order $\succeq$.*

In other words, an irreducible tree-partition $\mathcal{P}$ of $\mathcal{G}$ is a partition that cannot be reduced to a finer tree-partition.

Proposition 5.4.3. *For any graph $\mathcal{G}$, there exists a unique irreducible tree-partition.*

See Section 5.7 for a proof.

We remark that our proof of Proposition 5.4.3 not only shows that the irreducible tree-partition of $\mathcal{G}$ is unique, but also implies that the problem of computing this unique irreducible tree-partition reduces to finding all bridges of $\mathcal{G}$. As a result, we can adapt Tarjan's bridge-finding algorithm [66] to devise an algorithm that

[3] We remark that our definition of bridges agrees with the classical definition of bridges in graph theory (i.e., the removal of any such edge disconnects the original graph) in the sense that if the tree-partition $\mathcal{P}$ is irreducible (see Definition 5.4.2 later) any bridge defined in our sense is a bridge in the classical sense, and vice versa.

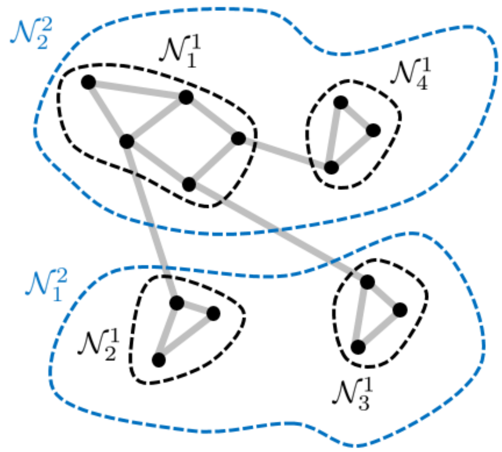


Figure 5.3: An illustration of the partial order $\succeq$ over tree-partitions. The partition $\mathcal{P}^1 = \left\{\mathcal{N}_1^1, \mathcal{N}_2^2, \mathcal{N}_3^1, \mathcal{N}_4^1\right\}$ is finer than $\mathcal{P}^2 = \left\{\mathcal{N}_1^2, \mathcal{N}_2^2\right\}$.

computes the irreducible tree-partition of $\mathcal{G}$ in $O(n + m)$ time. This is summarized in Algorithm 1. Interested readers are referred to the proof of Proposition 5.4.3 in Section 5.7 for more details on the algorithm.

Algorithm 1 Irreducible Tree Partition Finding Algorithm

1: Execute Tarjan's bridge-finding algorithm [66] on $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ to compute the set of bridges $\mathcal{E}_b$.
2: Remove edges in $\mathcal{E}_b$ from $\mathcal{E}$ to form the partitioned graph $(\mathcal{N}, \mathcal{E}\backslash\mathcal{E}_b)$.
3: Breadth-first search on the partitioned graph $(\mathcal{N}, \mathcal{E}\backslash\mathcal{E}_b)$ to compute its set of connected components $\mathcal{P} := \{C_1, C_2, \ldots, C_k\}$. Return $\mathcal{P}$.

5.5 Guaranteed Localization

In Section 5.4, we formally defined the tree-partitions of a power grid that are motivated by the Simple Loop Criteria from Proposition 5.3.3. In this section, we demonstrate with a small stylized example that tree-partitions help localize impacts of a line failure in power systems. Moreover, such failure localization does not necessarily come with increased line congestion if the tree-partition is properly chosen.

Failure Localization

Consider the double-ring network in Figure 5.4(a), which contains exactly one generator and one load bus (indicated by G and L, respectively). The original power flow on this network is also shown in Figure 5.4(a). As a comparison, we consider the left ring and the right ring as two separate regions, and switch off the upper

Figure 5.4: (a) A double-ring network. G is the generator bus and L is the load bus. Arrows represent the original power flow. (b) The new network after removing an edge. Arrows represent the new power flow.

tie-line to form a tree-partition. This new network and the redistributed power flow are shown in Figure 5.4(b).

It is easy to check that in Figure 5.4(a), there is a simple loop containing any pair of edges. By the Simple Loop Criteria we know that the impact of tripping any single line in this topology is global; that is, all the other branch flows are changed almost surely. Thus every other line is subject to potential successive failures. In fact, if we are allowed to adversarially choose the injection (over all buses) and line capacities, one can show that for any pair of edges $e, \hat{e}$ in Figure 5.4(a) there exists a scenario where the failure of e triggers the failure of $\hat{e}$.

In contrast, for the graph in Figure 5.4(b), there are only two possible simple loops given by the left and right ring, respectively. With the Simple Loop Criteria, we then see that the line failures inside each of these rings do not impact branch flows in the other. In other words, line failures are localized within their own region in Figure 5.4(b).

Such localization, however, only applies in one stage insofar as further failures may involve bridges in the graph, to which the Simple Loop Criteria no longer applies. In fact, as we show in Chapter 6, tripping a bridge always has a global impact under mild assumptions, and therefore failures may still propagate from one ring to the other after multiple steps. Another drawback of the topology in Figure 5.4(b) is the single-point vulnerability at the newly created bridge, whose failure disconnects the system into two islands. Nevertheless, by adopting a new control approach for fast-timescale frequency regulation, we can overcome this drawback (we present this new approach in Chapter 7).

Congestion Reduction

It is reasonable to expect that switching off lines to form a tree-partition may increase the stress on the remaining lines and, in this way, worsen the network congestion. In fact, one may expect that improved system robustness obtained by switching off lines *always* comes at the price of increased congestion levels. We now show that this is not necessarily the case, and demonstrate that if the lines to switch off are selected properly, it is possible to improve the system robustness and reduce the congestion simultaneously.

Indeed, by comparing the power flows in Figure 5.4(a) and (b), we see that forming a proper tree-partition as in Figure 5.4(b) can potentially remove the circulating flows and hence reduce the overall network congestion. In fact, for this specific example, the tree-partition as shown in Figure 5.4(b) minimizes the sum of absolute branch flows over all possible topologies on this network where G and L are connected.

More concretely, let

$$\mathfrak{C} := \{\mathcal{G}' = (\mathcal{N}, \mathcal{E}') : G \text{ and } L \text{ are connected in } \mathcal{G}'\}$$

be the collection of all topologies over $\mathcal{N}$ (which do not need to be a sub-graph of Figure 5.4(a)) such that G is connected to L. Let p' be the injection as shown in Figure 5.4(a), that is, $p'_G = d$, $p'_L = -d$ for some $d > 0$, and $p'_j = 0$ for other buses. For each $\mathcal{G}' = (\mathcal{N}, \mathcal{E}') \in \mathfrak{C}$, define the metric

$$\Psi(\mathcal{G}') := \sum_{e \in \mathcal{E}'} |f'_e|,$$

where f'_e is the branch flow on e under the injection p'. Then we have:

Proposition 5.5.1. *The graph in Figure 5.4(b) minimizes the sum of absolute branch flow* $\Psi(\cdot)$ *over* $\mathfrak{C}$.

Proof. For any $\mathcal{G}' = (\mathcal{N}, \mathcal{E}') \in \mathfrak{C}$, because of the conservation constraints in DC power flow equations (2.5), the injection p' and the branch flow f' can be considered as a single "flow" (see [75] for the rigorous definition of such "flow") from G to L with volume d. Thus, by the Max-Flow-Min-Cut Theorem we know for any cut over edges $E \subset \mathcal{E}'$,

$$\sum_{e \in E} |f_e| \geq d,$$

and therefore $\Psi(\mathcal{G}') \geq d$. For the graph shown in Figure 5.4(b), the sum of absolute branch flow is exactly d. The desired result then follows. $\square$

In other words, not only does this properly formed tree-partition help reduce the system congestion, it in fact achieves the best possible reduction in terms of $\Psi(\cdot)$. We revisit this phenomenon with case studies on more practical IEEE test systems in Chapter 6.

5.6 Conclusion

In this chapter, we studied the monotonicity and structural properties of power redistribution in a cascading failure process. We demonstrated that there is a rich collection of monotonicity one can explore in the Laplacian spectrum, and that the distributions of different types of trees in the transmission network determine possible patterns in power redistribution after a line failure. Our results motivate a novel approach via tree-partitions of the network to localize the impact of line failures. Moreover, switching off lines to create a tree-partition does not necessarily worsen the system congestion level if these lines are chosen properly.

5.7 Proofs

Proof of Proposition 5.2.2

Let the set of removed edges be

$$R_1 := \{(i_1, j_1), \dots, (i_s, j_s)\}$$

and denote the eigenvectors of $L_{\mathcal{G}}$ and $L_{\mathcal{H}}$ by $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_n$ respectively. To prove the inequality (5.1), we define

$$V_1 := \operatorname{span}\left(\{v_1, v_2, \dots, v_k\}\right)$$

and

$$W_1 := \operatorname{span}\left(\{w_k, w_{k+1}, \dots, w_n\}\right).$$

Then $\dim(V_1) = k$, $\dim(W_1) = n - k + 1$, which implies $\dim(V_1 \cap W_1) > 0$ and therefore we can find $\beta \in V_1 \cap W_1$. Without loss of generality, we can choose β such that $\|\beta\|^2 = \beta^T \beta = 1$. By the Courant-Fisher-Weyl variational formula, we then know

$$\begin{aligned}
\mu_k = \min_{v \in W_1} \frac{v^T L_{\mathcal{H}} v}{v^T v} \le \frac{\beta^T L_{\mathcal{H}} \beta}{\beta^T \beta} = \sum_{(i,j) \in \mathcal{E} \setminus R_1} w_{ij} (\beta_i - \beta_j)^2 \\
\le \sum_{(i,j) \in \mathcal{E}} w_{ij} (\beta_i - \beta_j)^2 = \frac{\beta^T L_{\mathcal{G}} \beta}{\beta^T \beta} \le \max_{v \in V_1} \frac{v^T L_{\mathcal{G}} v}{v^T v} \\
= \lambda_k.
\end{aligned}$$

To prove the inequality (5.2), define the vector spaces

$$R_2 := \left\{ x \in \mathbb{R}^n : x_{i_\tau} = x_{j_\tau}, \tau = 1, 2, \ldots, s \right\}$$

$$V_2 := \text{span}\left(\{v_{k-s}, v_{k-s+1}, \ldots, v_n\}\right)$$

and

$$W_2 := \text{span}\left(\{w_1, w_2, \ldots, w_k\}\right).$$

Then $\dim(R_2) \geq n - s$, $\dim(V_2) = n - k + s + 1$, $\dim(W_2) = k$. Therefore,

$$\begin{aligned} \dim(V_2 \cap W_2) &= \dim(V_2) + \dim(W_2) - \dim(V_2 + W_2) \\ &\geq n - k + s + 1 + k - n = s + 1, \end{aligned}$$

and thus $\dim(R_2 \cap V_2 \cap W_2) > 0$. So we can find $\gamma \in R_2 \cap V_2 \cap W_2$, and without loss of generality, we can assume $\|\gamma\|^2 = \gamma^T \gamma = 1$. Then again the Courant-Fisher-Weyl variational formula implies

$$\begin{aligned} \lambda_{k-s} &= \min_{v \in V_2} \frac{v^T L_{\mathcal{G}} v}{v^T v} \leq \frac{\gamma^T L_{\mathcal{G}} \gamma}{\gamma^T \gamma} = \sum_{(i,j) \in \mathcal{E}} w_{ij} (\gamma_i - \gamma_j)^2 \\ &= \sum_{(i,j) \in \mathcal{E} \setminus R_2} w_{ij} (\gamma_i - \gamma_j)^2 \qquad (5.5) \\ &= \frac{\gamma^T L_{\mathcal{H}} \gamma}{\gamma^T \gamma} \leq \max_{v \in V_1} \frac{v^T L_{\mathcal{H}} v}{v^T v} \\ &= \mu_k, \end{aligned}$$

where (5.5) is because $\gamma \in R_2$. □

Proof of Proposition 5.4.3

For $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, let $\mathcal{E}_b$ be the set of edges that all spanning trees of $\mathcal{G}$ pass through. In other words, $\mathcal{E}_b$ is the set of all bridges in classical graph theory. Let $\mathcal{G}_b = (\mathcal{N}, \mathcal{E} \setminus \mathcal{E}_b)$ denote the graph obtained from $\mathcal{G}$ by removing all edges in $\mathcal{E}_b$, and let $\mathcal{P}^* = \{C_1, C_2, \cdots, C_k\}$ be its connected components, where $k = |\mathcal{E}_b| + 1$. We claim $\mathcal{P}^*$ is a irreducible tree-partition of $\mathcal{G}$.

First we show that the reduced graph $\mathcal{G}_{\mathcal{P}^*}$ is a tree. Assume not, then there is a loop in $\mathcal{G}_{\mathcal{P}^*}$. Without loss of generality, let us assume the loop is $(C_1, C_2, \cdots, C_l)$ for some $2 \leq l \leq k$. Then, by the way we form $\mathcal{G}_{\mathcal{P}^*}$ there exist vertices $n_1^t, n_1^s, n_2^t, n_2^s, n_3^t, \cdots, n_l^t, n_l^s$ such that:

1. For each i, the vertices $n_i^s, n_i^t \in C_i$;

2. For each i, the edge $e_{i,i+_l 1} := (n_i^s, n_{i+_l 1}^t) \in \mathcal{E}$, where $+_l$ denotes the addition modulo l.

For any i, since C_i is connected, we can find a path P_i from n_i^t to n_i^s. It is then clear that the concatenated path

$$(P_1, e_{1,2}, P_2, e_{2,3} \dots e_{l-1,l}, P_l, e_{l,1})$$

forms a loop in the original graph $\mathcal{G}$. As a result, not all spanning trees pass through $e_{1,2}$ and thus $e_{1,2} \notin \mathcal{E}_b$, which leads to a contradiction.

Next we show $\mathcal{P}^*$ is irreducible. To do so, we prove that $\mathcal{P}^*$ is finer than any tree-partition $\mathcal{P} := \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_{k'}\}$ of $\mathcal{G}$. Consider a region in $\mathcal{P}^*$, say C_1. Since both $\mathcal{P}^*$ and $\mathcal{P}$ are partitions of $\mathcal{G}$, there must be some region in $\mathcal{P}$, say $\mathcal{N}_1$, such that $C_1 \cap \mathcal{N}_1 \neq \emptyset$. We claim $C_1 \subset \mathcal{N}_1$. Otherwise, there exists another region in $\mathcal{P}$, say $\mathcal{N}_2$, such that $C_1 \cap \mathcal{N}_2 \neq \emptyset$. Pick $n_1 \in C_1 \cap \mathcal{N}_1$ and $n_2 \in C_1 \cap \mathcal{N}_2$. Then $n_1 \neq n_2$ because $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$. Now since $n_1, n_2 \in C_1$, and C_1 does not contain any bridge (in classical graph theory sense), by Menger's Theorem [46], there exists a cycle (which is not necessarily simple) in C_1 containing both n_1 and n_2. By collapsing adjacent vertices in this cycle that belong to common regions, we can find regions $\mathcal{N}_{l_1}^1, \mathcal{N}_{l_2}^1, \cdots, \mathcal{N}_{l_{p_1}}^1, \mathcal{N}_{l_1}^2, \mathcal{N}_{l_2}^2, \cdots, \mathcal{N}_{l_{p_2}}^2$ so that the path from n_1 to n_2 in this cycle is given by

$$\left(P_1^1, e_{1,l_1}^1, P_{l_1}^1, e_{l_1,l_2}^1, \cdots, e_{l_{p_1},2}^1, P_2^1\right)$$

and the path from n_2 to n_1 in this cycle is given by

$$\left(P_2^2, e_{2,l_1}^2, P_{l_1}^2, e_{l_1,l_2}^2, \cdots, e_{l_{p_2},1}^2, P_1^2\right),$$

where $e_{i,j}^1, e_{i,j}^2$ are edges with source vertices in $\mathcal{N}_i$ and target vertices in $\mathcal{N}_j$ and P_i^1, P_i^2 are paths contained in $\mathcal{N}_i$. As a result, we see

$$\left(\mathcal{N}_1, \mathcal{N}_{l_1}^1, \cdots, \mathcal{N}_{l_{p_1}}^1, \mathcal{N}_2, \mathcal{N}_{l_1}^2, \cdots, \mathcal{N}_{l_{p_2}}^2\right)$$

forms a loop in $\mathcal{G}_{\mathcal{P}}$. This implies $\mathcal{P}$ is not a tree-partition, contradicting our assumption.

We thus have shown that $\mathcal{P}^*$ is a irreducible tree-partition of $\mathcal{G}$. Moreover, for any other irreducible tree-partition $\overline{\mathcal{P}}$, the above proof shows that

$$\mathcal{P}^* \succeq \overline{\mathcal{P}}.$$

Since $\overline{\mathcal{P}}$ is irreducible and thus maximal with respect to $\succeq$, we see $\overline{\mathcal{P}} = \mathcal{P}^*$. In other words, any irreducible tree-partition of $\mathcal{G}$ must coincide with $\mathcal{P}^*$. This completes our proof. $\square$

Chapter 6

FAILURE LOCALIZATION VIA TREE-PARTITIONS

In Chapter 5, we established the spectral representation of power redistribution that captures the Kirchhoff's Law in terms of tree distributions of the transmission network. This new representation motivated the tree-partition of power systems that can prevent the impact of failure from long-distance propagation.

In this chapter, we continue our study along this thread and prove that the tree-partition proposed in the last chapter provides a precise analytical characterization of line failure localizability, and that the tree-partition encodes rich information on how line failures can cascade. In Section 6.1, we focus on the single line failure case and present our formal characterization of localizability in Theorem 6.1.1, which summarizes the technical results in Sections 6.2 and 6.3. In particular, in Section 6.2, we characterize the power redistribution after the tripping of a non-bridge line and show that the impact of such failures only propagates within well-defined components, which we refer to as cells, inside the tree-partition regions. In Section 6.3, we consider the failure of bridge lines and prove that in normal operating conditions such failures propagate globally across the network and impact the power flow on all transmission lines. Later in Section 6.4, we extend the results from Theorem 6.1.1 to the case of multiple line failures, and show that non-bridge failures stay within the original cell until a bridge fails.

The characterization we provide in Theorem 6.1.1 yields many interesting insights for the planning and management of power systems and, further, suggests a new approach for mitigating the impact of cascading failures. Specifically, our characterization highlights that switching off certain transmission lines can lead to more, smaller regions/cells which localize line failures, thus making the grid less vulnerable to line outages. In Section 6.5, we illustrate this approach using the IEEE 118-bus test system. We demonstrate that switching off only a negligible portion of transmission lines can lead to significantly better control of cascading failures. Further, we highlight that this happens without significant increases in line congestion across the network.

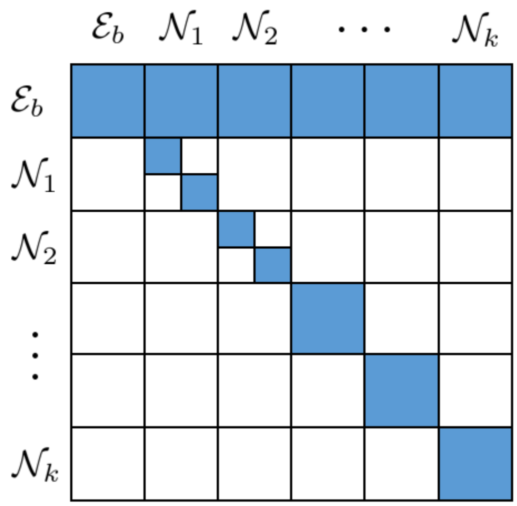


Figure 6.1: Non-zero entries of the $K_{e\hat{e}}$ matrix (as represented by the dark blocks) for a graph with tree-partition $\{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ and bridge set $\mathcal{E}_b$. The small blocks represent cells inside the regions.

6.1 Single Line Failure

In this section we state our main analytical result in the single line failure case. It summarizes the technical results in the two sections that follow.

Our result applies in contexts where the system is operating under *normal* conditions, i.e., when the following two assumptions are satisfied: (a) the injection is *island-free* (see Definition 6.3.1 for a formal definition); and (b) the grid is *participating* with respect to its power balance rule (see Definition 6.3.3 for a formal definition). Moreover, to address certain pathological cases, we add a perturbation drawn from certain probability measure μ to the line susceptances and assume μ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}_m$ on $\mathbb{R}^m$ (see Section 6.3).

Theorem 6.1.1. *For a power network operating under normal conditions, $K_{e\hat{e}} \neq 0$ "if" and only if:*

1. *$e, \hat{e}$ are within a common tree-partition region and $e, \hat{e}$ belong to the same cell; or*

2. *e is a bridge.*

The "if" part in the statement above should be interpreted as almost surely in μ (see Definition 6.2.4). This result highlights that, for a practical system, the tree-partition encodes rich information on how the failure of a line propagates through the network.

We emphasize that: (a) the condition that μ is absolutely continuous with respect to $\mathcal{L}_m$ is satisfied by almost all practical probability models for such perturbations (see Section 6.2); and (b) the conditions that the injection is island-free and the grid is participating are satisfied in typical operating scenarios (see Section 6.3). Therefore, the conditions posed in Theorem 6.1.1 are satisfied in practical settings.

Figure 6.1 shows how the tree-partition is linked to the sparsity of the $K_{e\hat{e}}$ matrix through Theorem 6.1.1. It suggests that, compared to a full mesh transmission network consisting of single region/cell, it can be beneficial to switch off certain lines so that more regions/cells are created and the impact of a line failure is localized within the cell in which the failure occurs. We study this network planning and design opportunity in Section 6.5.

In the next two sections we prove Theorem 6.1.1. We first characterize the power redistribution after the tripping of a non-bridge line in Section 6.2, and then consider the failure of bridge lines in Section 6.3. Later in Section 6.4, we generalize our single line failure characterization to the case of multi-line failures.

6.2 Non-Bridge Failures are Localizable

In this section, we characterize the power flow redistribution under the DC model when a non-bridge line is tripped and show that such failures are localized by the tree-partition regions. More specifically, we study how the tripping of a line $e \in \mathcal{E}$ impacts the branch flow on a different edge $\hat{e} \in \mathcal{E}$. Recall our mention in Section 5.1 that when e is not a bridge, the power flow change on $\hat{e}$ due to tripping e is given by

$$\Delta f_{\hat{e}} = K_{e\hat{e}} \times f_e.$$

The impact of the line failure of e can thus be characterized by the line outage distribution factor $K_{e\hat{e}}$.

Impact across Regions

To start with, we consider the case where $\hat{e}$ does not belong to the same region as e; that is, $\hat{e}$ either belongs to a different region or $\hat{e}$ is a bridge.

Proposition 6.2.1. *Consider a power network $\mathcal{G}$ with a tree-partition*

$$\mathcal{P} := \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}.$$

Let $e, \hat{e} \in \mathcal{E}$ be two different edges such that e is not a bridge. Then,

$$K_{e\hat{e}} = 0$$

for any $\hat{e}$ that is not in the same the region containing e.

Proof. If $\hat{e}$ is a bridge, then since no simple loop can contain a bridge, we know that $K_{e\hat{e}} = 0$ from the Simple Loop Criteria (Proposition 5.3.3).

If $\hat{e}$ is not a bridge, without loss of generality, assume e is within $\mathcal{N}_1$ and $\hat{e}$ is within $\mathcal{N}_2$. Since $\mathcal{P}$ is a tree-partition, we know that any path starting from a node in $\mathcal{N}_1$ and ending at a node in $\mathcal{N}_2$ must pass through all bridges in the path from $\mathcal{N}_1$ to $\mathcal{N}_2$ in the reduced graph $\mathcal{G}_{\mathcal{P}}$. As a result, any loop containing both e and $\hat{e}$ must pass through these bridges at least twice, and thus is not simple. By the Simple Loop Criteria, we then know that $K_{e\hat{e}} = 0$. □

This result implies that, when a non-bridge line e fails, any line $\hat{e}$ not in the same tree-partition region as e will not be affected, regardless of whether $\hat{e}$ is a bridge or not. In other words, non-bridge failures cannot propagate through the boundaries formed by the tree-partition regions of $\mathcal{G}$. This is a formal proof of the intuition that, under DC power flow model (which assumes zero line loss), the power flow on a bridge between two regions in a tree-partition depends only on the net generation-load imbalances in these regions, and therefore a line failure within a region will not change the flow on this bridge or line flows in the other region.

Impact within Regions

It is reasonable, based on physical intuition, to expect that the converse to the above result is also true. That is, if $e, \hat{e}$ belong to a common region (and thus e is not a bridge), we would expect $K_{e\hat{e}} \neq 0$. This, however, is not always the case for two reasons: (a) some vertices within a tree-partition region may "block" simple loops containing e and $\hat{e}$; and (b) the graph $\mathcal{G}$ may be too symmetric. We elaborate on these two scenarios separately in the following two subsections.

(a) Block Decomposition

To illustrate the issue described above, we use the following example to demonstrate that certain vertices within a tree-partition region may "block" simple loops containing e and $\hat{e}$ from being formed.

Example 1. *Consider a butterfly network shown in Figure 6.2(a) and pick $e = (i, j)$ and $\hat{e} = (w, z)$ from the butterfly wings. It is not hard to see that any loop containing e and $\hat{e}$ must pass through the body vertex c at least twice, and hence is not a simple loop. From the Simple Loop Criteria, we see that $K_{e\hat{e}} = 0$.*

Figure 6.2: (a) A butterfly network. (b) The block decomposition of the butterfly network into cells C_1 and C_2.

The issue with Example 1 is that the butterfly graph is not 2-connected. In other words, it is possible that the removal of a single vertex (in this case the body vertex c) can disconnect the original graph. We refer to such a vertex as a *cut vertex*, following graph-theoretic convention. From Example 1, we see that cut vertices may "block" simple loops from being formed.

Fortunately, we can precisely capture such an effect by decomposing each tree-partition region further through the classical *block decomposition* [33]. Recall that the block decomposition of a graph is a partition of its *edges* such that each partitioned component is 2-connected. See Figure 6.2(b) for an illustration. We refer to such components as *cells* to reflect the fact that they are smaller parts within a tree-partition region. Note that two different cells within a tree-partition region may share a common vertex, as the block decomposition is over graph edges. The block decomposition of a graph always exists and can be found in linear time [67].

Lemma 6.2.2. *Consider a power network $\mathcal{G}$ and let $e, \hat{e}$ be two distinct edges within the same tree-partition region but across different cells. Then $K_{e\hat{e}} = 0$.*

Proof. Let e be within the cell C_e and $\hat{e}$ be within the cell $C_{\hat{e}}$. It is a classical result that any path originating from a vertex within C_e to a vertex within $C_{\hat{e}}$ must pass through a common cut vertex in C_e [33]. As a result, it is impossible to find a simple loop containing both e and $\hat{e}$. By the Simple Loop Criteria, we then know $K_{e\hat{e}} = 0$. □

(b) Symmetry

Next, we demonstrate that graph symmetry[1] may also block the propagation of failures. Again, we illustrate the issue with a simple example.

Example 2. *Consider the complete graph on n vertices and pick $e = (i, j)$ and $\hat{e} = (w, z)$ such that e and $\hat{e}$ do not share any common endpoints. Assume the line susceptances are all* 1. *By symmetry, it is easy to see that there is a bijective correspondence between $\mathcal{T}(\{i, w\}, \{j, z\})$ and $\mathcal{T}(\{i, z\}, \{j, w\})$, and thus*

$$\sum_{E \in \mathcal{T}(\{i,w\},\{j,z\})} \chi(E) - \sum_{E \in \mathcal{T}(\{i,z\},\{j,w\})} \chi(E) = 0.$$

By Theorem 5.3.1, we then have $K_{e\hat{e}} = 0$.

A complete graph is 2-connected and thus forms a cell. Example 2 shows that even if the two edges $e, \hat{e}$ are within the same cell, when the graph $\mathcal{G}$ is rich in symmetries, it is still possible that a failure of e does not impact $\hat{e}$. Nevertheless, this issue is not critical as such symmetry almost never happens in practical systems because of heterogeneity in line susceptances. In fact, even if the system is originally symmetric, an infinitesimal change on the line susceptances is enough to break the symmetry, as we now show.

More formally, we adopt a form of perturbation analysis on the line susceptances similar to our discussions in Section 4.3. That is, instead of requiring the line susceptance to be fixed values B_e, we add a random perturbation $\omega = (\omega_e : e \in \mathcal{E})$ drawn from a probability measure μ. Such perturbations can come from manufacturing error or measurement noise. The perturbed system[2] shares the same topology (and thus tree-partition) as the original system, yet admits perturbed susceptances $B + \omega$. The randomness of ω implies the factor $K_{e\hat{e}}$ is now a random variable. Let $\mathcal{L}_m$ be the Lebesgue measure on $\mathbb{R}^m$. Recall that μ is *absolutely continuous* with respect to $\mathcal{L}_m$ if for any measurable set S such that $\mathcal{L}_m(S) = 0$, we have $\mu(S) = 0$.

Proposition 6.2.3. *Consider a power network $\mathcal{G}$ under perturbation μ and let $e, \hat{e}$ be two distinct edges within the same cell. If μ is absolutely continuous with respect to $\mathcal{L}_m$, then*

$$\mu\left(K_{e\hat{e}} \neq 0\right) = 1.$$

[1] By symmetry, we refer to graph automorphisms. The exact meaning of symmetry, however, is not important for our purpose.

[2] We assume the perturbation ensures $B_e + \omega_e > 0$ for any $e \in \mathcal{E}$ so that the new susceptance is physically meaningful.

See Section 6.7 for a proof.

Note that, by the Radon-Nikodym theorem [60], the probability measure μ is absolutely continuous with respect to $\mathcal{L}_m$ if and only if it affords a probability density function. In other words, there are no requirements on either the power or the correlation of the perturbation for Proposition 6.2.3 to apply. The only necessary condition is that the measure μ cannot contain Dirac masses. As a result, we see that for almost all practical probability models of such perturbation (e.g., truncated Gaussian noise with arbitrary covariance; bounded uniform distribution; truncated Laplace distribution), $K_{e\hat{e}} \neq 0$ for $e, \hat{e}$ within the same cell almost surely, no matter how small the perturbation is.

This perturbation approach is also useful for our results in the following sections. When we take this approach, our result often constitutes two directions: (a) an "only if" direction that should be interpreted as normal; and (b) an "if" direction that holds almost surely in μ. To simplify the presentation, we henceforth fix a perturbation μ that is absolutely continuous with respect to $\mathcal{L}_m$, and define the following:

Definition 6.2.4. *For two predicates p and q, we say that p* **"if" and only if** *q when both of the following hold:*

$$p \Rightarrow q, \qquad q \Rightarrow \mu(p) = 1.$$

We say the "if" is in μ-sense when we need to emphasize that the "if" direction only holds almost surely in μ.

6.3 Bridge Failures Propagate

The remaining case necessary to prove Theorem 6.1.1 is a characterization of the power flow redistribution when a bridge is tripped. Here, we show that such failures generally propagate through the entire network.

Extended $K_{e\hat{e}}$ and Island-free Grid

Recall that, when e is not a bridge, the branch flow change on $\hat{e}$ due to tripping e is given by

$$\Delta f_{\hat{e}} = K_{e\hat{e}} \times f_e. \tag{6.1}$$

When e is a bridge, tripping e disconnects the power grid into two islands, and the power in each connected component may not be balanced. Such power imbalance can be resolved by a power balance rule $\mathcal{R}$ (see [9, 11, 13] for examples of such rules), which together with the DC model uniquely determines the new branch flows (and thus the branch flow change $\Delta f_{\hat{e}}$). For the purpose of unified notation, we extend the definition of $K_{e\hat{e}}$ through (6.1) to the case where e is a bridge. Besides being related to the $A := (\overline{L})^{-1}$ matrix, the extended $K_{e\hat{e}}$ factor also depends on the power injection p and the power balance rule $\mathcal{R}$.

This $K_{e\hat{e}}$, of course, is only well-defined if $f_e \neq 0$. Since when $f_e = 0$ we clearly have $\Delta f_{\hat{e}} = 0$ for all the remaining lines, we will focus on the case $f_e \neq 0$ in this section. Indeed, power networks without micro-grids typically operate in "island-free mode", as islanding (i.e., isolating a part of the grid power flow from the rest of the network) poses a safety hazard to utility maintainence and repair personnel and potentially leads to damage of the infrastructure [59]. Formally we define the concept of island-free as follows:

Definition 6.3.1. *For a power network $\mathcal{G}$, an injection p is said to be* **island-free** *if under the injection p, the branch flow f in $\mathcal{G}$ satisfies $f_e \neq 0$ for any bridge e.*

Intuitively, island-free means that no part of the grid balances its own power. For a island-free grid, any bridge carries nonzero branch flow, and thus the extended $K_{e\hat{e}}$ is always well-defined.

Participating Bus

Consider an island $\mathcal{D} := (\mathcal{N}_{\mathcal{D}}, \mathcal{E}_{\mathcal{D}})$ created from removing the bridge e, and let u be the endpoint of e that belongs to $\mathcal{D}$. Tripping e from the grid effectively changes the injection at u by f_e, and the balancing rule $\mathcal{R}$ distributes such imbalance to a set of *participating* buses from $\mathcal{D}$ so that the total power imbalance f_e is canceled out. The rules studied in the literature [9, 11, 13, 62] are typically *linear* in the sense that, for any participating bus j, the injection adjustment Δp_j dictated by the rule $\mathcal{R}$ is linear in f_e. Denote the set of participating buses of $\mathcal{R}$ in $\mathcal{D}$ as $\mathcal{N}_{\mathcal{R}}$ and let $n_r = |\mathcal{N}_{\mathcal{R}}|$. The rule $\mathcal{R}$ can then be interpreted as a linear transformation from $\mathbb{R}$ to $\mathbb{R}^{n_r}$ given by

$$\mathcal{R}(M) = \left(\alpha_j M : j \in \mathcal{N}_{\mathcal{R}}\right),$$

where α_j are positive constants that sum to 1 (so that power balance is achieved after applying $\mathcal{R}$). Different rules correspond to different choices of participating buses and the constants α_j's. For instance, if the imbalance is uniformly absorbed by the generators as in [9, 62], we have $\mathcal{N}_{\mathcal{R}}$ to be the set of generators and $\alpha_j = 1/G$, where G is the number of generators in $\mathcal{D}$. As another example, if the imbalance is regulated through Automatic Generation Control (AGC), then we have $\mathcal{N}_\mathcal{R}$ to be the set of controllable generators and α_j to be the normalized generator participation factors.

Denote the injection adjustment over all buses in $\mathcal{D}$ by $\Delta p_\mathcal{D}$, which comes from tripping e (that effectively changes the injection at u by f_e) and also from the power balancing by $\mathcal{R}$. That is $(\Delta p_\mathcal{D})_j \neq 0$ only for the participating buses or if $j = u$. Let $\overline{C}$ be the matrix obtained from C by removing the slack bus, and A be the inverse of $\overline{L}$ (see Chapter 2 for more details on these notations). Then the branch flow changes on the remaining lines in $\mathcal{D}$ are given by

$$\Delta f_\mathcal{D} = B_\mathcal{D} \overline{C}_\mathcal{D}^T A_\mathcal{D} \Delta p_\mathcal{D},$$

where the matrices with subscript $\mathcal{D}$ refer to their submatrices corresponding to buses or lines in $\mathcal{D}$. We now determine when $(\Delta f_\mathcal{D})_{\hat{e}} = 0$ for a remaining line $\hat{e} \in \mathcal{E}_\mathcal{D}$, which in turn characterizes whether the extended $K_{e\hat{e}}$ is zero or not.

Proposition 6.3.2. *For $\hat{e} \in \mathcal{E}_\mathcal{D}$, we have $\Delta f_{\hat{e}} \neq 0$ "if" and only if $\exists j \in \mathcal{N}_\mathcal{R}$ such that there is a path in $\mathcal{D}$ from u to j containing $\hat{e}$.*

The "if" part in this result is in μ-sense as discussed in Section 6.2. See Section 6.7 for a proof.

Proposition 6.3.2 shows that the positions of participating buses in $\mathcal{D}$ play an important role in distributing power imbalance across the network. In particular, the power balancing rule $\mathcal{R}$ changes the branch flow for every edge that lies in a possible path from the failure point u to the set of participating buses. As a result, if $\hat{e}$ is a bridge that connects two islands $\mathcal{D}_1$ and $\mathcal{D}_2$ in $\mathcal{D}$, then assuming $u \in \mathcal{D}_1$, we see that $\Delta f_{\hat{e}} \neq 0$ "if" and only if $\mathcal{D}_2$ contains a participating bus since a path from u to any node in $\mathcal{D}_2$ must pass through $\hat{e}$. If $\hat{e}$ is not a bridge, then we can devise a simple sufficient condition on $\Delta f_{\hat{e}} \neq 0$ using *participating regions*, defined as follows:

Definition 6.3.3. *For a power grid $\mathcal{G}$ with tree-partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ operating under power balance rule $\mathcal{R}$, a region $\mathcal{N}_i$ with block decomposition $\left\{C_1^i, C_2^i, \cdots, C_{m_i}^i\right\}$ is said to be a* **participating region** *if $\mathcal{N}_\mathcal{R} \cap C_j^i$ contains a non-cut-vertex for $j = 1, 2, \ldots, m_i$. The grid $\mathcal{G}$ is said to be a* **participating grid** *if $\mathcal{N}_i$ is participating for $i = 1, 2, \ldots, k$.*

A typical power grid does not contain single-point vulnerabilities such as cut vertices, which often suggests that the tree-partition regions consist of single cells. In this case, a region is a participating region if there exists at least one bus in this region that participates in power balancing and is not the endpoint of a bridge, which is often satisfied. It is thus reasonable to assume most tree-partition regions are participating regions and hence most grids are participating. For instance, if all generators participate in AGC, and load-side participation is implemented at all load buses, then every bus in the network is a participating bus, and hence the grid is clearly participating.

Lemma 6.3.4. *If $\mathcal{N}_i$ is a participating region, then for any $\hat{e} \in \mathcal{N}_i$,*

$$\mu(\Delta f_{\hat{e}} \neq 0) = 1.$$

Proof. Let C be the cell that contains $\hat{e}$. Since $\mathcal{N}_i$ is a participating region, we know there exists a bus within C, say n_1, that participates in power balance and is not a cut vertex. Recall that any path from u to C must go through a common cut vertex in C [33], say n_e. Now by adding an edge between n_e and n_1 (if it does not exist originally), the resulting cell C' is still 2-connected. Thus there exists a simple loop in C' that contains the edge (n_e, n_1) and $\hat{e} = (w, z)$, which implies we can find two disjoint paths P_1 and P_2 connecting the endpoints of these two edges. Without loss of generality, assume P_1 connects n_e to w and P_2 connects n_1 to z. By concatenating the path from u to n_e, we can extend P_1 to a path $\tilde{P}_1$ from u to w, which is still disjoint from P_2. Now, by adjoining $\hat{e}$ to $\tilde{P}_1$ and P_2, we can construct a path from u to n_1 that passes through $\hat{e}$. By Proposition 6.3.2, we then know $\mu(\Delta f_{\hat{e}} \neq 0) = 1$. □

Given the above, we now state our main result for bridge failures.

Proposition 6.3.5. *Consider a participating power network $\mathcal{G}$ with island-free injection p . If e is a bridge of $\mathcal{G}$, then for any $\hat{e} \neq e$, we have*

$$\mu(K_{e\hat{e}} \neq 0) = 1.$$

Proof. If $\hat{e}$ is a bridge, denote the two connected components of $\mathcal{D}$ after removing $\hat{e}$ as $\mathcal{D}_1$ and $\mathcal{D}_2$, and without loss of generality assume $\mathcal{D}_1$ is originally connected to e in $\mathcal{G}$. It is easy to see that the branch flow change on $\hat{e}$ is given by

$$\Delta f_{\hat{e}} = \sum_{j \in \mathcal{D}_2} (\Delta p_{\mathcal{D}})_j \neq 0,$$

where the last $\neq$ is because the grid is participating and thus all tree-partition regions in $\mathcal{D}_2$ would adjust their aggregate injections (in the same "direction", as α_j's are positive).

If $\hat{e}$ is not a bridge, then Lemma 6.3.4 implies $\mu(\Delta f_{\hat{e}} \neq 0) = 1$. □

6.4 Generalization to Multi-line Failure

In this section, we generalize our results in previous sections to the case where multiple lines fail simultaneously. That is, we characterize the branch flow changes of remaining lines after a set E of lines are tripped from the system.

Non-cut Failure

Similar to the single-line failure case, the impact of tripping a set of lines propagates differently depending on whether the new graph $\mathcal{G}' := (\mathcal{N}, \mathcal{E} \backslash E)$ obtained from tripping E is connected. This motivates us to revisit the following definition that is well-known in graph theory literature.

Definition 6.4.1. *A set E is said to be a* ***cut set*** *of $\mathcal{G}$ if the graph $\mathcal{G}'$ obtained by removing E from $\mathcal{G}$ is not connected, or a* ***non-cut set*** *if it is not a cut set.*

In this subsection, we focus on the case where E is a non-cut set. Given the linear DC model (2.5) for power distribution, it is reasonable to conjecture that the impact of tripping a set of lines simultaneously is the linear sum of the impacts from tripping lines in this set separately. This, however, is not always true. For instance, given three edges $e_1, e_2, \hat{e}$, if we trip e_1, e_2 simultaneously, then the branch flow change in $\hat{e}$ is (see [71], for instance)

$$\Delta f_{\hat{e}} = \frac{K_{e_1\hat{e}}\left(f_{e_1} + K_{e_2e_1} f_{e_2}\right)}{1 - K_{e_1e_2}K_{e_2e_1}} + \frac{K_{e_2\hat{e}}\left(f_{e_2} + K_{e_1e_2} f_{e_1}\right)}{1 - K_{e_1e_2}K_{e_2e_1}}, \tag{6.2}$$

which is different from the direct linear sum

$$K_{e_1\hat{e}} f_{e_1} + K_{e_2\hat{e}} f_{e_2}.$$

More generally, given a non-cut set E of edges, it is known that [71]

$$\Delta f = B\overline{C}^T A C_E (I - B_E C_E^T A C_E)^{-1} f_E, \tag{6.3}$$

where $f_E := (f_e : e \in E)$, C_E and B_E are the submatrices of the incidence matrix C and the susceptance matrix B corresponding to E. Putting $\tilde{f}_E := (I - B_E C_E^T A C_E)^{-1} f_E$, for an edge $\hat{e}$, we then know from (6.3) that $\Delta f_{\hat{e}}$ is linear in $\tilde{f}_E$:

$$\Delta f_{\hat{e}} = B_{\hat{e}} \cdot \sum_{e \in E} D_{\hat{e},e} \tilde{f}_e, \tag{6.4}$$

where $D_{\hat{e},e}$ is the generation shift sensitivity factor between $\hat{e}$ and e as discussed in Chapter 2.

The formula (6.4) suggests that tripping E simultaneously from $\mathcal{G}$ can be interpreted as consisting of two steps: (a) first, the original flow f_E mixes to a vector of flow change $\tilde{f}_E$ according to $\tilde{f}_E = (I - B_E C_E^T A C_E)^{-1} f_E$, which captures the non-linearity as illustrated in (6.2); and (b) second, the mixed flow change $\tilde{f}_E$ propagates *linearly* to the remaining edges via (6.4). It is easy to see that $D_{\hat{e},e} \neq 0$ if and only if $K_{\hat{e}e} = 0$, and thus we can apply Theorem 6.1.1 to determine which terms in (6.4) vanish. In other words, the characterization of step (b) follows directly from our results in Section 6.2 regarding single line failures.

We are then left to characterize how $\tilde{f}_E$ is related to f_E in step (a). Towards this goal, we collect edges in E based on the cells they belong to and write $E = E_1 \cup E_2 \cup \cdots \cup E_k$ as its *cell decomposition*. That is, $E_i \subset \mathcal{C}_i$ for some cell $\mathcal{C}_i$ in $\mathcal{G}$, and $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$ if $i \neq j$. This decomposition is well-defined since a non-cut set E does not contain any bridge and thus any edge in E must belong to a certain cell.

Proposition 6.4.2. *Let $E = E_1 \cup E_2 \cup \cdots \cup E_k$ be its cell decomposition and put $m_i = |E_i|$. Then with proper permutation of rows and columns, the matrix $(I - B_E C_E^T A C_E)^{-1}$ is of the form:*

$$\begin{bmatrix} H_1 & 0 & \cdots & 0 \\ 0 & H_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_k \end{bmatrix},$$

where $H_i \in \mathbb{R}^{m_i \times m_i}$ for $i = 1, 2, \cdots, k$. Moreover, under a perturbation μ on line susceptances that is absolutely continuous with respect to $\mathcal{L}_m$, H_i consists of strictly nonzero values almost surely:

$$\mu((H_i)_{e_1 e_2} \neq 0) = 1, e_1, e_2 = 1, 2, \cdots, m_i.$$

The proof of this result is presented in Section 6.7. It shows that for $\tilde{e} \in E_i$, $f_{\tilde{e}}$ depends exactly on and only on the values of f_e for $e \in E_i$, in the sense that, changing f_e almost surely changes $f_{\tilde{e}}$ for $\tilde{e}$ in the same cell yet has no impact at all for edges in other cells.

Given $E = E_1 \cup E_2 \cup \cdots \cup E_k$ where $E_i \subset \mathcal{C}_i$, denote the branch flow change when we trip E_i simultaneously from the grid as Δf^{E_i}. Proposition 6.4.2 shows that the original flow f_E mixes within the corresponding cells, which together with (6.4) implies that

$$\Delta f^{E_i} = B_{\hat{e}} \cdot \sum_{e \in E_i} D_{\hat{e},e} \tilde{f}_{E_i} = B\overline{C}^T A C_{E_i} \tilde{f}_{E_i}$$

and therefore,

$$\Delta f = B\overline{C}^T A C_E \tilde{f}_E = \sum_{i=1}^{k} B\overline{C}^T A C_{E_i} \tilde{f}_{E_i} = \sum_{i=1}^{k} \Delta f^{E_i}.$$

In other words, the impact of tripping E simultaneously from the grid is the same as the aggregate impacts of tripping E_i separately from the grid, which forms a clear contrast to the non-linearity demonstrated by (6.2). Further, for a fixed cell, say C_1, the impact of tripping E_1 from the grid precisely consists of two steps: (a) f_{E_1} mixes within E_1 into a flow change $\tilde{f}_{E_1}$; and (b) then the impact of $\tilde{f}_{E_1}$ propagates to every edge in C_1 almost surely.

Formally, the above discussions can be summarized as the following theorem (its proof is presented in Section 6.7):

Theorem 6.4.3. *Let $K^E_{e\hat{e}}$ be the $(\hat{e}, e)$-th entry of $B\overline{C}^T A C_E (I - B_E C_E^T A C_E)^{-1}$. Then*

$$K^E_{e\hat{e}} \neq 0$$

"if" and only if e and $\hat{e}$ are within the same cell.

This result generalizes the non-bridge failure case of Theorem 6.1.1, and shows that such failures are localized within the corresponding cells (in one stage). Note that, unlike the single-line failure case where $K_{e\hat{e}} \neq 0$ automatically implies $f_{\hat{e}} \neq 0$, this is not the case when E consists of multiple lines because the impacts from different line failures can potentially cancel each other. Nevertheless, it can be shown that under mild conditions on the topology of $\mathcal{G}$, by adding a perturbation to the system injection p, such canceling almost never happens.

Cut Failure

Next we consider the case where E forms a cut set and thus the formula (6.3) no longer applies (the matrix $I - B_E C_E^T A C_E$ is not invertible in this case). The removal of E disconnects $\mathcal{G}$ into multiple islands that are potentially power imbalanced, and the power balancing rule $\mathcal{R}$ needs to be applied to cancel out such imbalances. For the purpose of presentation, let us focus on one island $\mathcal{D}$ thus created and denote its injection adjustment from $\mathcal{R}$ by $\Delta p_{\mathcal{D}}$. This adjustment $\Delta p_{\mathcal{D}}$ has non-zero components only at participating buses or buses that are endpoints of some edge in E (bridge flows are "absorbed" to its endpoints in $\mathcal{D}$).

Given the fixed island $\mathcal{D}$, put $B_{\mathcal{D}}, \overline{C}_{\mathcal{D}}$, to be the sub-matrices of $B, \overline{C}$ corresponding to buses and edges in this island, respectively, and let $E_{\mathcal{D}}$ be the set of edges in E that have both endpoints within this island. Note that $E_{\mathcal{D}}$ is a non-cut set of $\mathcal{D}$ since otherwise tripping E would disconnect $\mathcal{D}$ to multiple islands. Define $\overline{L}_{\mathcal{D}}$ to be the Laplacian matrix of $\mathcal{D}$ before $E_{\mathcal{D}}$ is tripped and also with a certain slack bus in $\mathcal{D}$ removed, and let $A_{\mathcal{D}} := (\overline{L}_{\mathcal{D}})^{-1}$. Put $K^{E_{\mathcal{D}}} := B_{\mathcal{D}}\overline{C}_{\mathcal{D}}^T A_{\mathcal{D}} C_{E_{\mathcal{D}}}(I - B_{E_{\mathcal{D}}} C_{E_{\mathcal{D}}}^T A_{\mathcal{D}} C_{E_{\mathcal{D}}})^{-1}$ to be the matrix from (6.3) with all the matrices replaced with their counterparts in this island. With all these notations, we characterize how Δf is related to $\Delta p_{\mathcal{D}}$, $K^{E_{\mathcal{D}}}$ and $E_{\mathcal{D}}$.

From (6.4), we know that when multiple lines are tripped from the grid simultaneously, the aggregate impact in general is different from the linear sum of tripping the lines separately. The impact from the balancing rule $\mathcal{R}$, though, turns out to be separable from the rest, as we now show.

Proposition 6.4.4. *Given the injection adjustment $\Delta p_{\mathcal{D}}$ and original flow $f_{E_{\mathcal{D}}}$ on $E_{\mathcal{D}}$, we have*

$$\Delta f_{\mathcal{D}} = B_{\mathcal{D}}\overline{C}_{\mathcal{D}}^T \tilde{A}_{\mathcal{D}} \Delta p_{\mathcal{D}} + K^{E_{\mathcal{D}}} f_{E_{\mathcal{D}}}, \tag{6.5}$$

where $\tilde{A}_{\mathcal{D}}$ is the pseudo-inverse of the Laplacian matrix of $\mathcal{D}$ after $\mathcal{E}_{\mathcal{D}}$ is tripped.

Proof. Denote the original injection over $\mathcal{D}$ by $p_{\mathcal{D}}$; then after tripping E the new injection on $\mathcal{D}$ is $p_{\mathcal{D}} + \Delta p_{\mathcal{D}}$. Therefore, the new power flow is given by

$$\tilde{f}_{\mathcal{D}} = B_{\mathcal{D}} C_{\mathcal{D}}^T \tilde{A}_{\mathcal{D}} (p_{\mathcal{D}} + \Delta p_{\mathcal{D}}).$$

Since $B_{\mathcal{D}} C_{\mathcal{D}}^T \tilde{A}_{\mathcal{D}} p_{\mathcal{D}}$ is simply the power flow on $\mathcal{D}$ after $\mathcal{E}_{\mathcal{D}}$ is tripped under the original injection $p_{\mathcal{D}}$, we see that

$$B_{\mathcal{D}} C_{\mathcal{D}}^T \tilde{A}_{\mathcal{D}} p_{\mathcal{D}} = f_{\mathcal{D}} + K^{E_{\mathcal{D}}} f_{\mathcal{E}_{\mathcal{D}}}.$$

The desired result then follows. □

The first term in (6.5) captures the impact of power imbalance, and is characterized by our discussions in Section 6.3. The second term in (6.5) reduces to the non-cut set case since $\mathcal{E}_{\mathcal{D}}$ is a non-cut set of $\mathcal{D}$. We thus see that a cut set failure impacts the branch flow on remaining lines in two independent ways: (a) via participating buses to distribute the power imbalances; and (b) via cells to mix and propagate original flows on the tripped lines. The formula (6.5) precisely captures the impact propagation through these two ways, which are fully characterized by our results in previous sections.

Localization Horizon

In this section, we summarize the results in the multi-line failure case and show that tree-partition localizes the impact of line failures until the grid is disconnected into multiple islands. More formally, given a cascading failure process described by $\mathcal{B}(\mathfrak{n})$, $\mathfrak{n} \in \mathcal{N}$, define

$$\mathfrak{T} := \min \{\mathfrak{n} \in \mathcal{N} : \mathcal{F}(\mathfrak{n}) \text{ is a cut set of } \mathcal{G}(\mathfrak{n})\}$$

to be the first stage where the grid is disconnected to multiple islands. Without loss of generality, assume the initial failure $\mathcal{B}(1)$ contains only one edge that belongs to a cell C. Then we know that:

Proposition 6.4.5. *For any* $\mathfrak{n} \leq \mathfrak{T}$*, we have*

$$\mathcal{F}(\mathfrak{n}) \subset C.$$

In other words, in a cascading failure process, the only way that a non-bridge failure can propagate to edges outside the cell that the original failure belongs to is to have the grid disconnected into multiple islands at a certain point of the cascades.

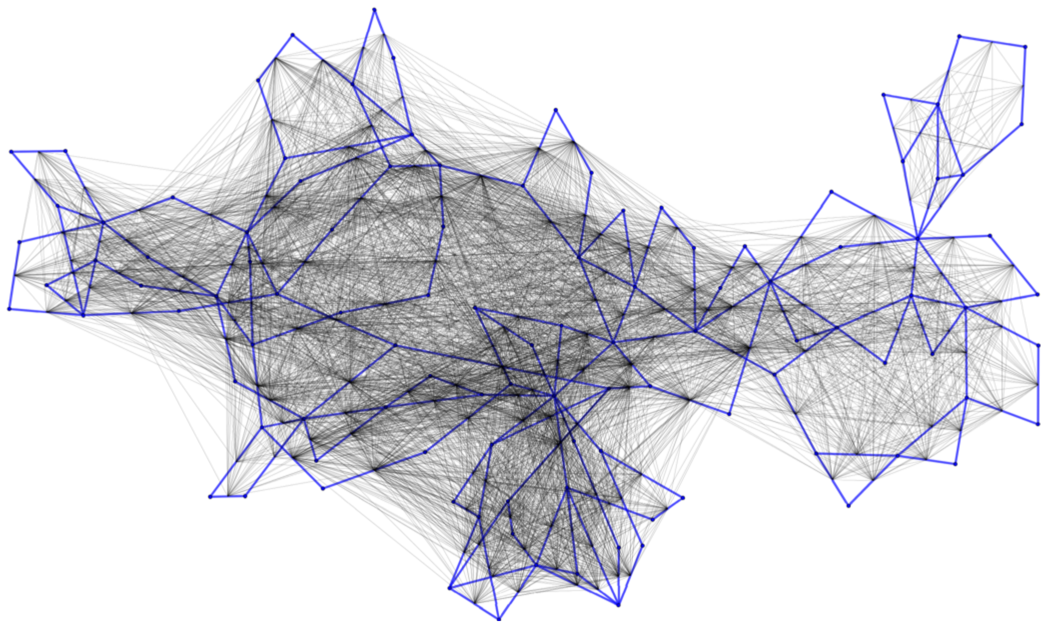
6.5 Case Studies

Our findings highlight a new approach for improving the robustness of the network. More specifically, Theorem 6.1.1 and the discussions in Section 6.2 suggest that it is possible to localize failure propagation by switching off certain transmission lines. This creates more, smaller areas where failure cascades can be contained. In this section, we consider the IEEE 118-bus test system to illustrate this approach.

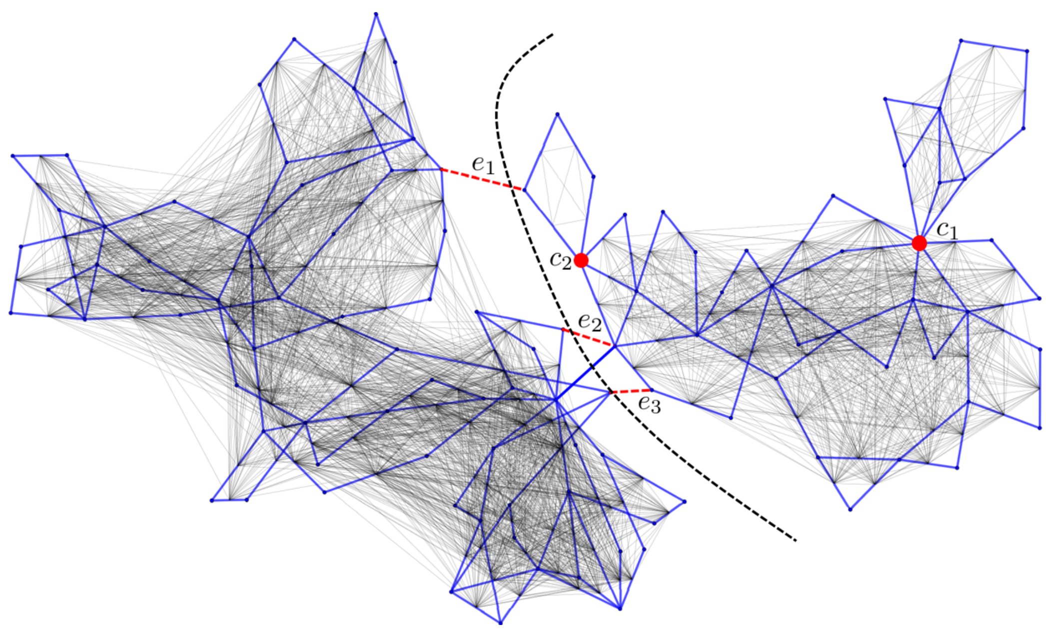
Influence Graph

In our experiments, the system parameters are taken from the Matpower Simulation Package [85], and we plot the influence graphs among the transmission lines to demonstrate how a line failure propagates in this network[3]. More specifically, in the

[3] The original IEEE 118-bus network has some trivial "dangling" bridges that we remove (collapsing their injections to the nearest bus) to obtain a more transparent influence graph.



(a) Original influence graph.



(b) The influence graph after switching off e_1, e_2 and e_3. The black dashed line indicates the failure propagation boundary defined by the tree-partition. The vertices c_1 and c_2 are cut vertices.

Figure 6.3: Influence graphs on the IEEE 118-bus network before and after switching off lines e_1, e_2 and e_3. Blue edges represent physical transmission lines and grey edges represent connections in the influence graph.

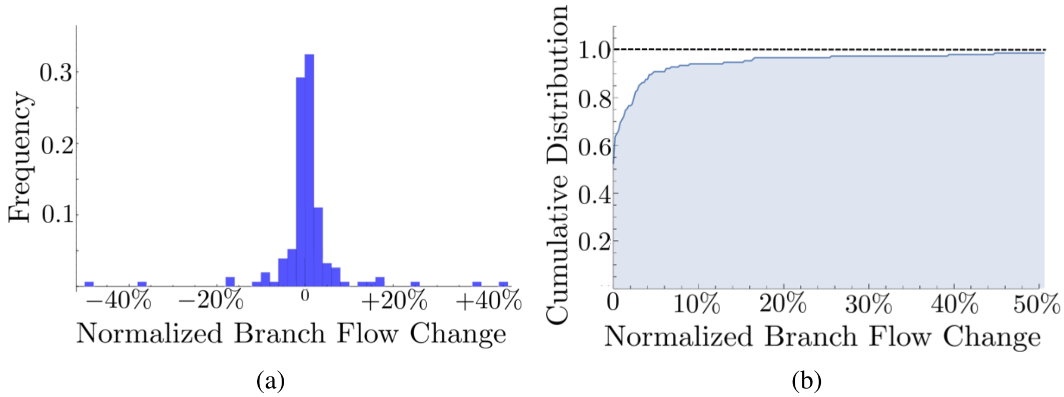


Figure 6.4: (a) Histogram of the normalized branch flow changes. (b) Cumulative distribution function of the positive normalized branch flow changes. Note that the curve intercepts the y-axis since 52.59% of the branch flows decrease.

influence graph we plot, two edges e and $\hat{e}$ are connected if the impact of tripping e on $\hat{e}$ is not negligible (we use $|K_{e\hat{e}}| \geq 0.005$ as a threshold). In Figure 6.3(a), we plot the influence graph of the original network. It can be seen that this influence graph is very dense and connects many edges that are topologically far away, showing the non-local propagation of line failures within this network.

Next, we switch off three edges (indicated as e_1, e_2 and e_3 in Figure 6.3(b)) to obtain a new topology that has a bridge and whose tree-partition now consists of two regions of comparable size. The new influence graph is shown in Figure 6.3(b). One can see that, compared to the original influence graph in Figure 6.3(a), the new influence graph is much less dense, and in particular, there are no edges connecting transmission lines that belong to different tree-partition regions. We also note that the network in Figure 6.3(b) contains two cut vertices (indicated by c_1 and c_2 in the figure, with c_2 being created when we switch off the lines). It can be checked that line failures are "blocked" by these cut vertices, which verifies our results in Section 6.2.

Congestion Management

It is also of interest to see how the network congestion is impacted by switching off these lines. To do so, we collect statistics on the difference between the branch flows in Figure 6.3(b) and those in 6.3(a). In Figure 6.4(a), we plot the histogram of such branch flow differences normalized by the original branch flow in Figure 6.3(a). It shows that roughly half (the exact percentage is 47.41%) of the transmission lines have higher congestion yet the majority of these branch flow increases are negligible.

To more clearly see how much the congestion worsens on these lines, we plot the cumulative distribution function of the normalized positive branch flow changes, which is shown in Figure 6.4(b). One can see from the figure that 90% of the branch flows increase by no more than 10%.

6.6 Conclusion

In this chapter, we provided a precise analytical characterization of line failure localizability in power systems. We demonstrated that the tree-partition of the transmission network graph encodes rich information on the regions that a line failure can impact. Further, using a case study on the IEEE 118-bus test network, we showed that switching off certain transmission lines can improve the grid robustness against line failures without significantly increasing line congestion.

6.7 Proofs

Proof of Proposition 6.2.3

Let C be the cell that e and $\hat{e}$ belong to and write $e = (i, j)$ and $\hat{e} = (w, z)$. Consider the polynomial in line susceptances $(B_e : e \in \mathcal{E})$ defined as

$$f(B) := \sum_{E \in \mathcal{T}\left(\{i,w\},\{j,z\}\right)} \chi(E) - \sum_{E \in \mathcal{T}\left(\{i,z\},\{j,w\}\right)} \chi(E).$$

We claim that f is not identically zero.

First, let C be a simple cycle in C that contains both e and $\hat{e}$. Such a cycle always exists as C is 2-connected by construction, and it is a classical result that any pair of edges are contained in a simple cycle for 2-connected graphs [33]. Therefore, we can find two disjoint paths P_1 and P_2 connecting the endpoints of e and $\hat{e}$. Without loss of generality, assume P_1 connects i to w and P_2 connects j to z. By iteratively adding edges from $\mathcal{G}$ to P_1 and P_2, we can extend P_1 and P_2 to a spanning forest of $\mathcal{G}$ consisting of exactly two trees. Moreover, the tree extended from P_1 contains $\{i, w\}$ and the tree extended from P_2 contains $\{j, z\}$. We thus have constructed an element of $\mathcal{T}\left(\{i, w\}, \{j, z\}\right)$. Denote this element by E_0.

Second, we show that

$$\mathcal{T}\left(\{i, w\}, \{j, z\}\right) \cap \mathcal{T}\left(\{i, z\}, \{j, w\}\right) = \emptyset.$$

Indeed, consider an element E_1 from $\mathcal{T}\left(\{i, w\}, \{j, z\}\right)$, which consists of two trees $\mathcal{T}_1$ and $\mathcal{T}_2$ with $\mathcal{T}_1$ containing $\{i, w\}$ and $\mathcal{T}_2$ containing $\{j, z\}$. If $E_1 \in \mathcal{T}\left(\{i, z\}, \{j, w\}\right)$,

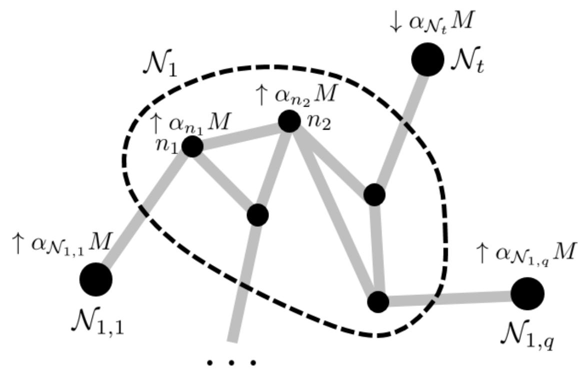


Figure 6.5: The localized graph $\mathcal{G}_{\mathcal{N}_1}$. $\mathcal{N}_t$ is the imaginary bus containing e and $\mathcal{N}_{1,i}$'s are remaining imaginary buses. The power adjustments from the power balance rule $\mathcal{R}$ are shown near each participating bus in reaction to a power loss of M.

then $\mathcal{T}_1$ must also contain z. However, this implies $z \in \mathcal{T}_1 \cap \mathcal{T}_2$, and thus $\mathcal{T}_1$ and $\mathcal{T}_2$ are not disjoint, contradicting the definition of $\mathcal{T}\left(\{i, w\}, \{j, z\}\right)$.

As a result, we see that the element E_0 constructed in our first step contributes a term to $\sum_{E \in \mathcal{T}(\{i,w\},\{j,z\})} \chi(E)$ but not to $\sum_{E \in \mathcal{T}(\{i,z\},\{j,w\})} \chi(E)$. Therefore $f(B)$ contains non-vanishing terms and is not identically zero.

It is well-known from algebraic geometry that the root set of a polynomial which is not identically zero has Lebesgue measure zero [31]. That is, we have

$$\mu\left(f(B+\omega)=0\right) = \mathcal{L}_m\left(f(B+\omega)=0\right) = 0,$$

where the first equality is because μ is absolutely continuous with respect to $\mathcal{L}_m$ (it is clear that the root set of the polynomial f is measurable since f is continuous).

Finally, by Theorem 5.3.1 we know $K_{e\hat{e}} = 0$ if and only if $f(B+\omega) = 0$. This then implies that

$$\mu(K_{e\hat{e}} \neq 0) = 1 - \mu(K_{e\hat{e}} = 0) = 1 - \mu\left(f(B+\omega)=0\right) = 1$$

and completes our proof. □

Proof of Proposition 6.3.2

By merging two tree-partition regions if necessary, we can assume $\hat{e}$ belongs to a certain tree-partition region, say $\mathcal{N}_1$. Let $n_1, n_2, \cdots, n_l$ be the set of participating buses in $\mathcal{N}_1$. By collapsing all regions other than $\mathcal{N}_1$, we can define a "localized" graph $\mathcal{G}_{\mathcal{N}_1}$ centered around $\mathcal{N}_1$ as shown in Figure 6.5, with the following constructions:

(a) For each bridge b incident to $\mathcal{N}_1$, we create an imaginary bus that aggregates all injections inside all tree-partition regions that can be reached by $\mathcal{N}_1$ through b *before* e is tripped, and this imaginary bus is connected to $\mathcal{N}_1$ via the corresponding bridge b. (b) If e is not directly incident to $\mathcal{N}_1$, then e must connect two regions that are collapsed to a common imaginary bus, say $\mathcal{N}_t$. If e is incident to $\mathcal{N}_1$ and its end point in $\mathcal{N}_1$ is w, then we let $\mathcal{N}_t$ be an additional imaginary bus that is connected to w to mimic the edge e (that is, before e is tripped this imaginary bus $\mathcal{N}_t$ supplies power f_e towards w and after e is tripped, $\mathcal{N}_t$ loses all its injection; moreover, the edge connecting $\mathcal{N}_t$ to n_e has susceptance B_e). Denote other imaginary buses as $\mathcal{N}_{1,1}, \mathcal{N}_{1,2}, \cdots, \mathcal{N}_{1,q}$.

Without loss of generality, let us assume because of the tripping of e, the aggregate power in $\mathcal{D}$ is in shortage of $M := f_e$. Then the enforcement of the power balance rule $\mathcal{R}$ would increase the power injection at each participating bus in $\mathcal{D}$, which translates to the power adjustment as demonstrated in Figure 6.5. Specifically, the injection at $\mathcal{N}_t$ would drop by $\alpha_{\mathcal{N}_t} M$ as the power flow from e is lost (the drop is generally not M unless e is directly incident to $\mathcal{N}_1$); the injections at $\mathcal{N}_{1,1}, \mathcal{N}_{1,2}, \cdots, \mathcal{N}_{1,q}$ increase by $\alpha_{\mathcal{N}_{1,i}} M$; and injections at the participating buses $n_1, n_2, \cdots, n_l$ in $\mathcal{N}_1$ increase by $\alpha_{n_i} M$. In other words, by rebalancing power according to the rule $\mathcal{R}$, we effectively shift the injections from $\mathcal{N}_t$ to $\mathcal{N}_{1,1}, \mathcal{N}_{1,2}, \cdots, \mathcal{N}_{1,q}$ and $n_1, n_2, \cdots, n_l$.

Let the set of edges in $\mathcal{G}_{\mathcal{N}_1}$ be $\mathcal{E}_1$. Pick $\mathcal{N}_t$ to be the slack bus in this localized graph $\mathcal{G}_{\mathcal{N}_1}$ and define an index set

$$\mathcal{I} := \left\{\mathcal{N}_{1,1}, \mathcal{N}_{1,2}, \cdots, \mathcal{N}_{1,q}, n_1, n_2, \cdots, n_l\right\}.$$

Write $\hat{e} = (w, z)$. For any $i \in \mathcal{I}$, let $g^i(B)$ be the following polynomial in line susceptances $(B_e : e \in \mathcal{E}_1)$:

$$\sum_{E \in \mathcal{T}_{\mathcal{E}_1}(\{i,w\},\{\mathcal{N}_t,z\})} \chi(E) - \sum_{E \in \mathcal{T}_{\mathcal{E}_1}(\{i,z\},\{\mathcal{N}_t,w\})} \chi(E),$$

where $\mathcal{T}_{\mathcal{E}_1}(\mathcal{N}_1, \mathcal{N}_2)$ is the set of spanning forests of $\mathcal{G}_{\mathcal{N}_1}$ consisting of exactly two trees that contain $\mathcal{N}_1$ and $\mathcal{N}_2$ respectively. By our graphical representation of the generation shift sensitivity factor defined in Chapter 2, the branch flow change on $\hat{e}$ coming from the power shift from $\mathcal{N}_t$ with amount $\alpha_i M$ towards i is given by

$$\Delta f_{\hat{e}}^i = \alpha_i M \times \frac{g^i(E)}{\sum_{E \in \mathcal{T}_{\mathcal{E}_1}} \chi(E)},$$

where $\mathcal{T}_{\mathcal{E}_1}$ denotes the set of all spanning trees of $\mathcal{G}_{\mathcal{N}_1}$. Put

$$g(B) := \sum_{i \in \mathcal{I}} \alpha_i g^i(B).$$

By linearity, we know

$$\Delta f_{\hat{e}} = M \times \frac{g(B)}{\sum_{E \in \mathcal{T}_{\mathcal{E}_1}} \chi(E)}.$$

If $g(B) \neq 0$, then we can find at least one i such that $\mathcal{T}_{\mathcal{E}_1}(\{i, w\}, \{\mathcal{N}_t, z\})$ or $\mathcal{T}_{\mathcal{E}_1}(\{j, z\}, \{\mathcal{N}_t, w\})$ is nonempty. Without loss of generality, assume

$$\mathcal{T}_{\mathcal{E}_1}(\{i, w\}, \{\mathcal{N}_t, z\})$$

is nonempty. Any element in $\mathcal{T}_{\mathcal{E}_1}(\{i, w\}, \{\mathcal{N}_t, z\})$ contains two trees containing $\{i, w\}$ and $\{\mathcal{N}_t, z\}$ respectively, and thus induces one path from w to i and another path from $\mathcal{N}_t$ to z. By adjoining (w, z) to these two paths, we can create a path from i to $\mathcal{N}_t$ that contains $\hat{e}$. It is easy to see that this path induces a path in the original graph $\mathcal{G}$ between u and a certain participating bus j. Therefore, $\Delta f_{\hat{e}} \neq 0$ only if $\exists j \in \mathcal{N}_{\mathcal{R}}$ such that there is a path in $\mathcal{D}$ from u to a participating bus j containing $\hat{e}$.

On the other hand, if there is a path P in $\mathcal{D}$ from u to a participating bus j containing $\hat{e}$, we claim $g(B)$ is not identically zero. Indeed, by a similar argument to the proof of Proposition 6.2.3, we know that for any $i, j \in \mathcal{I}$ (including the case $i = j$), the following is true:

$$\mathcal{T}_{\mathcal{E}_1}(\{i, w\}, \{\mathcal{N}_t, z\}) \cap \mathcal{T}_{\mathcal{E}_1}(\{j, z\}, \{\mathcal{N}_t, w\}) = \emptyset.$$

As a result, a term in $g^i(B)$ with positive coefficient is never canceled by a term in $g^j(B)$ with negative coefficient, and vice versa. Therefore, to show $g(B)$ is not identically zero, it suffices to show at least one term of $g^i(B)$ is not identically zero.

To do so, note that by removing $\hat{e}$ from P, we can create a path from u to one endpoint of $\hat{e}$, say w, and another path from z to j. This implies that in the localized graph Figure 6.5 we have a path from $\mathcal{N}_t$ to w and another path from z to some $i \in \mathcal{I}$. By iteratively adding edges, these two paths can be extended to an element in $\mathcal{T}_{\mathcal{E}_1}(\{j, z\}, \{\mathcal{N}_t, w\})$, which contributes to a term of $g^i(B)$ that is not identically zero.

Again by the classical algebraic geometry result asserting the root set of any polynomial that is not identically zero has Lebesgue measure zero [31], and because of the absolute continuity of μ, we know

$$\mu(\Delta f_{\hat{e}} = 0) = \mathcal{L}_m(g(B + \omega) = 0) = 0$$

and thus

$$\mu(\Delta f_{\hat{e}} \neq 0) = 1.$$

This completes our proof. $\square$

Proof of Proposition 6.4.2

Let $L^E := C_E^T A C_E$. First we claim that if $e_i \in E_i$, $e_j \in E_j$ and $i \neq j$, then $L^E_{e_i e_j} = 0$. To see this, note that

$$L^E_{e_i e_j} = D_{e_i, e_j},$$

where D_{e_i,e_j} is the generation shift sensitivity factor between e_i and e_j (see Chapter 2). It is easy to see that $D_{e_i,e_j} = 0$ if and only if $K_{e_i e_j} = 0$. In particular, if e_i and e_j are in different cells, then $D_{e_i,e_j} = 0$. The claim is then proved.

Note that both I and B_E are diagonal matrices. As a result, by permuting the edges according to the cells they belong to, we know that $I - B_E L^E$ is in the following block-diagonal form:

$$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix},$$

where $J_i \in \mathbb{R}^{m_i \times m_i}$ for $i = 1, 2, \cdots, k$. Moreover, J_i is invertible since $I - B_E L^E$ is invertible (see [64] for instance). This in particular implies that $(I - B_E L^E)^{-1}$ is of the form

$$\begin{bmatrix} H_1 & 0 & \cdots & 0 \\ 0 & H_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_k \end{bmatrix},$$

where $H_i \in \mathbb{R}^{m_i \times m_i}$ for $i = 1, 2, \cdots, k$.

Next we consider H_1 and fix $e_1, e_2 \in \{1, 2, \cdots, m_i\}$. Then

$$(H_1)_{e_1 e_2} = \frac{\det(J_1^{e_1, e_2})}{\det(J_1)},$$

where $J_1^{e_1,e_2}$ is the matrix obtained from J_1 by replacing the e_1-th column with a vector with value 1 at e_2-th component and 0 otherwise. Put J_1' to be the submatrix of J_1 obtained by removing the e_1-th column and e_2-th row from J_1. Then

$$\det(J_1^{e_1,e_2}) = (-1)^{e_1+e_2} \det(J_1').$$

The entries of J_1' are of the form $1 - B_{e_i} D_{e_i,e_i}$ or $B_{e_i} D_{e_i,e_j}$. Recall from Corollary 2.4.4 that for any pair of edges $e_i = (u, v)$, $e_j = (w, z)$, we have

$$D_{e_i,e_j} = \frac{1}{\sum_{E \in \mathcal{T}_{\mathcal{E}}} \chi(E)} \left(\sum_{E \in \mathcal{T}(\{u,w\},\{v,z\})} \chi(E) - \sum_{E \in \mathcal{T}(\{u,z\},\{v,w\})} \chi(E) \right).$$

Since all edges in E_1 are within the same cell, D_{e_i,e_j} is not identically zero for any $e_i, e_j \in E_1$ (including the case $e_i = e_j$). We now show that $\det(J_1')$ is a polynomial in B that is not identically zero. To do so, we first prove the following lemmas:

Lemma 6.7.1. *If $e_i \neq e_j$ are two different edges, then any term in the numerator of $B_{e_i} D_{e_i,e_j}$ does not contain B_{e_j}.*

Proof. The numerator of $B_{e_i} D_{e_i,e_j}$ is given by the difference of

$$B_{uv} \sum_{E \in \mathcal{T}(\{u,w\},\{v,z\})} \chi(E)$$

and

$$B_{uv} \sum_{E \in \mathcal{T}(\{u,z\},\{v,w\})} \chi(E).$$

When $e_i \neq e_j$, for each element of $E \in \mathcal{T}(\{u,w\},\{v,z\})$, adding $e_i = (u,v)$ to E induces a spanning tree of $\mathcal{G}$ that passes through $e_i = (u,v)$ but not $e_j = (w,z)$, and $d_E(u,w) < d_E(v,w)$, where $d_E(\cdot,\cdot)$ means the distance in terms of minimum number of hops in E. Conversely, for any spanning tree of $\mathcal{G}$ that passes through $e_i = (u,v)$ but not $e_j = (w,z)$, and $d_E(u,w) < d_E(v,w)$, removing e_i induces an element of $E \in \mathcal{T}(\{u,w\},\{v,z\})$.

A similar argument also applies to $\mathcal{T}(\{u,z\},\{v,w\})$. Therefore the denominator of $B_{e_i} D_{e_i,e_j}$ is exactly given as

$$\sum_{E \in \mathcal{T}_{e_i,-e_j}} \operatorname{sign}(E)\chi(E), \tag{6.6}$$

where $\mathcal{T}_{e_i,-e_j}$ is the set of spanning trees of $\mathcal{G}$ that pass through $e_i = (u,v)$ but not $e_j = (w,z)$, and

$$\operatorname{sign}(E) := \begin{cases} 1, & d_E(u,w) < d_E(v,w) \\ -1, & d_E(u,w) \geq d_E(v,w). \end{cases}$$

In particular, B_{e_j} does not appear in any of these terms. $\square$

Lemma 6.7.2. *If $e_i = e_j$, then the numerator of $B_{e_i} D_{e_i,e_j}$ is*

$$\sum_{E \in \mathcal{T}_{e_i}} \chi(E),$$

where $\mathcal{T}_{e_i}$ consists of all spanning trees of $\mathcal{G}$ that pass through e_i.

Proof. When $e_i = e_j$, D_{e_i,e_j} reduces to the effective reactance of e_i, and the result follows directly from Corollary 2.4.6. $\square$

Lemma 6.7.3. *Let $g_1, g_2, \cdots, g_{l_1}$ and $h_1, h_2, \cdots, h_{l_2}$ be functions in B of the form $B_{e_i} D_{e_i,e_j}$ with $e_i, e_j \in C_1$. Assume the e_j's for g_k are different over $k = 1, 2, \cdots, l_1$, and the e_j' for h_k are different over $k = 1, 2, \cdots, l_2$. Let q_1 be the number of g_k's with $e_i = e_j$, and q_2 be the number of h_k's with $e_i = e_j$. If $q_1 \neq q_2$, then for any fixed $a_1, a_2 \neq 0$, the following function*

$$f(B) := a_1 \prod_{k=1}^{l_1} g_k + a_2 \prod_{k=1}^{l_2} h_k$$

is not identically zero.

Proof. Without loss of generality, assume $l_1 \leq l_2$ and $q_1 < q_2$. Put $\zeta(B) = \sum_{E \in \mathcal{T}_\mathcal{E}} \chi(E)$, by collecting a common denominator, we then see that

$$f(B) = \frac{a_1 \zeta^{(l_2 - l_1)}(B) \tilde{g}(B) + a_2 \tilde{h}(B)}{\zeta^{l_2}(B)},$$

where $\tilde{g}(B)$ and $\tilde{h}(B)$ are homogeneous polynomials in B with order $l_1(n-1)$ and $l_2(n-1)$, respectively.

Let $\alpha(g) := (e_j : g_k = B_{e_i} D_{e_i,e_j}, e_i \neq e_j, k = 1, 2, \cdots, l_1)$ be the vector collecting all edges e_j corresponding to terms in g_k's with $e_i \neq e_j$, and define $\alpha(h)$ similarly. Since $q_1 < q_2$, we can find an edge $\tilde{e}_j$ in $\alpha(h)$ that is not in $\alpha(g)$. Without loss of generality, say this specific h is $h_{\tilde{k}}$. By Lemma 6.7.1, we know that the numerator of $h_{\tilde{k}}$ does not contain $B_{\tilde{e}_j}$. As a result, the order of $B_{\tilde{e}_j}$ is at most $l_2 - 1$ in all terms of $\tilde{h}(B)$.

Now, we claim that $\zeta^{(l_2 - l_1)}(B) \tilde{g}(B)$ contains a term where $B_{\tilde{e}_j}$ is of order l_2, which is strictly larger than $l_2 - 1$. This term cannot be canceled by any term from $\tilde{h}(B)$, and thus we know $f(B)$ is not identically zero. To show this claim, note that by expanding $\zeta^{(l_2 - l_1)}(B) \tilde{g}(B)$, we know each term in $\zeta^{(l_2 - l_1)}(B) \tilde{g}(B)$ is a product of terms from the factors involved. Thus it suffices to pick a term that contains $B_{\tilde{e}_j}$ from each of these factors.

First we consider the ζ factor. Since ζ sums over all spanning trees in $\mathcal{G}$, and any edge in $\mathcal{G}$ can be extended to a spanning tree of $\mathcal{G}$ by iteratively adding edges, we know there exists at least one term in ζ that contains $B_{\tilde{e}_j}$. We pick this term out for every ζ in $\zeta^{(l_2 - l_1)}(B) \tilde{g}(B)$, which multiplies to a term in which $B_{\tilde{e}_j}$ has order $l_2 - l_1$.

Next we consider g_k's with $e_i = e_j$. For such g_k, since e_i and $\tilde{e}_j$ are within the same cell, we can find a simple loop in this cell that contains both e_i and $\tilde{e}_j$. Removing an edge different from e_i and $\tilde{e}_j$ from this loop induces a path which, by iteratively adding edges, can be extended to a spanning tree that passes through e_j and $\tilde{e}_j$. This tree is an element of $\mathcal{T}_{e_i}$, and thus by Lemma 6.7.2 appears in the numerator of g_k. We pick this term for g_k, which contains $B_{\tilde{e}_j}$. Doing so for all g_k's of this type contributes $l_1 - q_1$ order of $B_{\tilde{e}_j}$.

Lastly, we consider g_k's with $e_i \neq e_j$. Since e_i and $\tilde{e}_j$ are within the same cell, we can find a simple loop in this cell that contains both e_i and $\tilde{e}_j$. If e_j is also in this loop, then we remove e_j. Otherwise, remove an edge different from e_i and $\tilde{e}_i$. This induces a path that contains e_i and $\tilde{e}_j$ but not e_j. By iteratively adding edges other than e_j, we can extend this path into a spanning tree of $\mathcal{G}$ that contains e_i and $\tilde{e}_j$ but not e_j, which in particular is an element of $\mathcal{T}_{e_i,-e_j}$. By (6.6) we then know that this tree appears in the numerator of g_k. We pick this term for g_k, which contains $B_{\tilde{e}_j}$. Doing so for all g_k's of this type contributes q_1 order of $B_{\tilde{e}_j}$.

In summary, we can find a term that contains $B_{\tilde{e}_j}$ in every factor of $\zeta^{(l_2-l_1)}(B)\tilde{g}(B)$. Multiplying all these terms together induces a term of $\zeta^{(l_2-l_1)}(B)\tilde{g}(B)$ where $B_{\tilde{e}_j}$ is of order l_2. Our claim then follows, and this completes the proof. $\square$

Recall that

$$\det(J_1') = \sum_{\sigma \in \mathcal{S}_{m_1}} \operatorname{sgn}(\sigma) \prod_{e \in E_1} (J_1')_{e\sigma(e)}, \tag{6.7}$$

where $\mathcal{S}_{m_1}$ is the symmetric group of order m_1 and $\operatorname{sgn}(\sigma)$ is the signature of σ.

Now, depending on whether $e_1 = e_2$, the matrix J_1' contains either $m_1 - 1$ or $m_1 - 2$ entries of the form $1 - B_{\tilde{e}_i} D_{\tilde{e}_i,\tilde{e}_j}$. When there are $m_1 - 1$ such entries, they must all appear on the diagonal of J_1', and thus multiplying all these terms induces a term in (6.7) that has the form

$$\prod_{k=1}^{m_1-1} B_{e_{i_k}} D_{e_{i_k},e_{i_k}}.$$

All other terms in $\det(J_1')$ contain at least one factor of the form $B_{\tilde{e}_i} D_{\tilde{e}_i,\tilde{e}_j}$ with $\tilde{e}_i = \tilde{e}_j$, and thus by Lemma 6.7.3, cannot cancel the above term. As a result, we know that $\det(J_1')$ is not identically zero.

When there are $m_1 - 2$ entries of the form $1 - B_{\tilde{e}_i} D_{\tilde{e}_i,\tilde{e}_j}$ in J_1', these entries must appear on the diagonal of J_1'. And thus multiplying them together with the remaining diagonal entry of J_1' induces a term in (6.7) that has exactly $m_1 - 1$ factors of the form $B_{\tilde{e}}D_{\tilde{e},\tilde{e}}$. Meanwhile, all other terms in $\det(J_1')$ contain at least two factors of the form $B_{\tilde{e}_i}D_{\tilde{e}_i,\tilde{e}_j}$ with $\tilde{e}_i \neq \tilde{e}_j$, and thus by Lemma 6.7.3, cannot cancel the above term. As a result, we know that $\det(J_1')$ is not identically zero.

In summary, we have shown that $\det(J_1')$ is a rational function that is not identically zero. Therefore we see

$$\mu(\det(J_1') = 0) = \mathcal{L}_m(\det(J_1') = 0) = 0,$$

or equivalently $(H_1)_{e_1e_2} \neq 0$ almost surely in μ. The desired result then follows. □

Proof of Theorem 6.4.3

Without loss of generality, assume $e \in E_1$. From Proposition 6.4.2, we see that the e-th column of $B\overline{C}^T AC_E(I - B_E C_E^T AC_E)^{-1}$ is given by

$$B\overline{C}^T AC_{E_1}(H_1)_e,$$

where $(H_1)_e$ is the column of H_1 corresponding to the edge e. As a result, we know

$$K_{e\hat{e}}^E = B_{\hat{e}} \cdot \sum_{e' \in E_1} D_{\hat{e},e'}(H_1)_{e'e}.$$

From the proof of Proposition 6.4.2, we can rewrite the above formula to

$$K_{e\hat{e}}^E = \frac{\sum_{e' \in E_1} B_{\hat{e}} D_{\hat{e},e'} \det(J_1^{e',e})}{\det(J_1)}, \tag{6.8}$$

where $J_1^{e',e}$ is the matrix obtained from J_1 by replacing the e'-th column with a vector with value 1 at e-th component and 0 otherwise.

If $\hat{e} \notin C_1$, then $D_{\hat{e},e'} = 0$ for all $e' \in E_1$ by Theorem 6.1.1. Thus we know $K_{e\hat{e}}^E = 0$.

If $\hat{e} \in C_1$, then we claim that the numerator of (6.8) is not identically zero. If the claim is true, then

$$\mu(K_{e\hat{e}}^E \neq 0) = \mathcal{L}_m(K_{e\hat{e}}^E \neq 0) = 1$$

and this completes the proof. We now show the claim indeed holds.

First we consider the case $\hat{e} \notin E_1$. In this case, note that the term in (6.8) corresponding to $e' = e$ contains one term that has exactly $m_1 - 1$ factors of the form $B_{\tilde{e}}D_{\tilde{e},\tilde{e}}$ ($\hat{e} \neq e'$ for all $e' \in E_1$ in this case), while all other terms in the numerator of (6.8) contains at most $m_1 - 2$ factors of the form $B_{\tilde{e}}D_{\tilde{e},\tilde{e}}$. From Lemma 6.7.3 we then see that the numerator of (6.8) is not identically zero.

Next we consider the case $\hat{e} \in E_1$ but $\hat{e} \neq e$. In this case, among all terms of (6.8), only the terms corresponding to $e' = \hat{e}$ and $e' = e$ contain one term that has exactly $m_1 - 1$ factors of the form $B_{\tilde{e}} D_{\tilde{e},\tilde{e}}$, and these two terms are

$$B_{\hat{e}} D_{\hat{e},e} B_{\hat{e}} D_{\hat{e},\hat{e}} \prod_{e'' \neq e,\hat{e}} B_{e''} D_{e'',e''}$$

and

$$B_e D_{e,\hat{e}} B_{\hat{e}} D_{\hat{e},\hat{e}} \prod_{e'' \neq e,\hat{e}} B_{e''} D_{e'',e''},$$

respectively, which do not cancel each other since $\hat{e} \neq e$. All other terms in the numerator of (6.8) contain at most $m_1 - 2$ factors of the form $B_{\tilde{e}} D_{\tilde{e},\tilde{e}}$. From Lemma 6.7.3 we then see that the numerator of (6.8) is not identically zero.

Finally we consider the case $\hat{e} = e$. In this case, among all terms of (6.8), only the term corresponding to $e' = e$ contains one term that has exactly m_1 factors of the form $B_{\tilde{e}} D_{\tilde{e},\tilde{e}}$, while all other terms in the numerator of (6.8) contain at most $m_1 - 2$ factors of the form $B_{\tilde{e}} D_{\tilde{e},\tilde{e}}$. From Lemma 6.7.3 we then see that the numerator of (6.8) is not identically zero.

Therefore the claim holds.

$\square$

Chapter 7

REAL-TIME OUTAGE MITIGATION

In this chapter, we leverage our results on fast-timescale swing dynamics and the localizability guarantees of tree-partitions to propose a distributed control strategy that operates on the frequency regulation timescale and offers provable failure mitigation properties and localization guarantees. Our control scheme ensures that failures do not propagate whenever there is a feasible way to avoid it (see Section 7.1 on the rigorous definition of such feasibility), and the impact of failures are localized as much as possible in a manner configurable by the system operator.

We introduce the main idea of our control design in Section 7.3, whose failure mitigation and localization guarantees are established by the technical results in Sections 7.4 and 7.5. The key piece of our control builds upon the so-called Unified Controller (UC), a novel design approach to frequency regulation [44, 82–84]. Our design revolves around the new and powerful properties that emerge when the regions that UC manages form a tree-partition. More specifically, in Section 7.4 we characterize how UC responds to an initial failure when it operates over a tree-partition, and prove that a non-critical failure is always mitigated and localized. Later in Section 7.5, we discuss how the tree-partition enables the system operator to explicitly specify the unfolding pattern of critical failures, and prove that UC can be extended to detect such scenarios as part of its normal operation.

In Section 7.6, we compare the proposed control strategy with classical Automatic Generation Control (AGC) using the IEEE 118-bus test system. We demonstrate that by switching off only a small subset of transmission lines and adopting UC as the fast timescale controller, one can significantly improve the system robustness to failures in terms of the $N-1$ security standard. Moreover, in a majority of the load profiles that are examined, our control strategy further localizes the impact of initial failures to the regions where they occur, leaving the operating points of all other control areas unchanged. Lastly, we highlight that when load shedding is inevitable, the proposed framework incurs significantly less load loss compared to AGC, in all of our case studies.

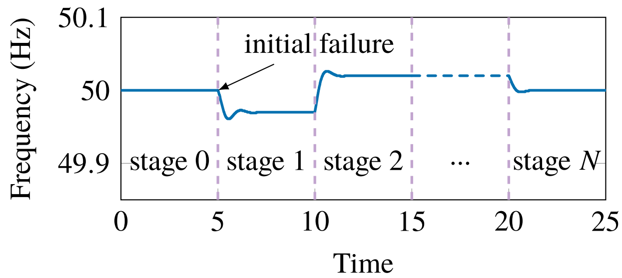


Figure 7.1: An illustration of the failure propagation model.

7.1 System Model

In this section, we extend the cascading failure model presented in Chapter 5 by adding fast-timescale swing dynamics, and discuss how the cascading process considered in Chapter 5 can be considered as a special case of this model.

Failure Occurrence and Propagation

In full generality, the control strategy that we introduce later applies to both generator failures and line failures. However, to simplify the presentation, in this chapter we focus only on line failures as the generalization to bus failures is straightforward.[1]

Recall that we describe the cascading failure process by keeping track of the set of failed lines $\mathcal{B}(\mathfrak{n})$ over $\mathfrak{n} \in \{1, 2, \cdots, N\}$. Instead of assuming that the power flow redistributes according to the DC model (2.5), we now assume that the system evolves according to the swing dynamics (2.1) and hence the equilibrium point is related to the design of the controllable injection d_j. Overloaded lines are tripped at slower timescales than the dynamics (2.1) and the cascade stages reflect this fact. Indeed, at each stage we assume that the system reaches the new steady state. The crux of our failure propagation model lies in the interplay between such slow-timescale line tripping process and the fast-timescale dynamics on system transient behavior described by (2.1), as illustrated in Figure 7.1.

More specifically, for each stage $\mathfrak{n} \in \{1, 2, \ldots, N\}$, the system evolves according to the dynamics (2.1) on the topology $\mathcal{G}(\mathfrak{n})$, and converges to an equilibrium point $x^*(\mathfrak{n}) = (\theta^*(\mathfrak{n}), \omega^*(\mathfrak{n}), d^*(\mathfrak{n}), f^*(\mathfrak{n}))$ that depends on $\mathcal{G}(\mathfrak{n})$. If all the branch flows $f^*(\mathfrak{n})$ are below the corresponding line ratings at equilibrium, then $x^*(\mathfrak{n})$ is a secure operating point and the cascade stops. Otherwise, let $\mathcal{F}(\mathfrak{n})$ be the subset of lines whose branch flows exceed the corresponding line ratings. The lines in $\mathcal{F}(\mathfrak{n})$ operate

[1] Our results readily apply to cases where the failure of a generator or substation can be emulated by the simultaneous failures of all the transmission lines connected to the corresponding bus.

above their safety limits in steady state, so by the end of stage $\mathfrak{n}$ they are overheated and tripped; i.e., $\mathcal{B}(\mathfrak{n}+1) = \mathcal{B}(\mathfrak{n}) \cup \mathcal{F}(\mathfrak{n})$. Line overloads during the transient phase before the system converges to $x^*(\mathfrak{n})$ are considered to be tolerable because the transient dynamics in (2.1) are not long enough to overheat a line [83] (lasting only seconds to a few minutes). This process repeats for stage $\mathfrak{n}+1$ and so on.

Recovering Previous Models

Our failure propagation model brings new perspectives to the commonly studied models in the literature, and reveals interesting insights into how certain critical limitations from previous work can be circumvented. In particular, the extra freedom in choosing d_j in the fast timescale dynamics (2.1) allows us to design and improve how the system reacts to line failures; thus achieving failure mitigation objectives directly using the well-known analytical tools from the frequency regulation literature.

As a first example to demonstrate this new approach, we show that, by adopting the classical droop control [8] as the dynamics for d_j in our framework, the cascading failure models from previous literature such as [62, 78] can be readily recovered. Indeed, as shown in [83], the closed-loop equilibrium of (2.1) under droop control is the unique[2] optimal solution to the following optimization:

$$\begin{aligned}
\min_{\omega,d,f,\theta} \quad & \sum_{j\in\mathcal{N}} \frac{d_j^2}{2K_j} + \frac{D_j w_j^2}{2} && \text{(7.1a)}\\
\text{s.t.} \quad & r - d - D\omega = Cf && \text{(7.1b)}\\
& f - BC^T\theta = 0 && \text{(7.1c)}\\
& \underline{d}_j \le d \le \overline{d}_j, \quad j \in \mathcal{N}, && \text{(7.1d)}
\end{aligned}$$

where K_j's are the generators' participation factors [8]. By plugging (7.1c) into (7.1b), it is routine to check that any feasible point $x = (\theta, \omega, d, f)$ of (7.1) satisfies $\sum_j r_j = \sum_j (d_j + D_j\omega_j)$. As a result, the Cauchy-Schwarz inequality implies that

$$\begin{aligned}
\left(\sum_{j\in\mathcal{N}} r_j\right)^2 &= \left[\sum_{j\in\mathcal{N}} \left(d_j + D_j\omega_j\right)\right]^2\\
&\le \sum_{j\in\mathcal{N}} \left(\frac{d_j^2}{2K_j} + \frac{D_j\omega_j^2}{2}\right) \sum_{j\in\mathcal{N}} \left(2K_j + 2D_j\right),
\end{aligned}$$

[2] Such uniqueness is up to a constant shift of all phase angles θ. See [83].

and equality holds if and only if

$$d_j = \frac{K_j}{\sum_j \left(K_j + D_j\right)} \sum_j r_j, \quad \omega_j = \frac{\sum_j r_j}{\sum_j \left(K_j + D_j\right)}. \tag{7.2}$$

Therefore, if the control limits (7.1d) are not active, (7.2) is always satisfied at the optimal point $x^* = (\theta^*, \omega^*, d^*, f^*)$.

Now consider a line e being tripped from the transmission network $\mathcal{G}$, and for simplicity assume the control limits (7.1d) are not active. If e is a bridge, the tripping of e results in two islands of $\mathcal{G}$, say $\mathcal{N}_1$ and $\mathcal{N}_2$, and two optimization problems (7.1) corresponding to $\mathcal{N}_1$ and $\mathcal{N}_2$ respectively. For $l = 1, 2$, $\sum_{j \in \mathcal{N}_l} r_j$ represents the total power imbalance in $\mathcal{N}_l$, and therefore (7.2) implies that droop control adjusts the system injection so that the power imbalance is distributed to all generators proportional to their participation factors in both $\mathcal{N}_1$ and $\mathcal{N}_2$. If e is not a bridge, then $\sum_{j \in \mathcal{N}} r_j = 0$ and thus (7.2) implies the system operating point remains unchanged. This control recovers exactly the failure propagation dynamics in [62, 78]. Moreover, one can show that this still holds when (7.1d) is active with a more involved analysis on the KKT conditions of (7.1).

We thus see that this droop control mechanism underlies some of the previous results in the literature on cascading failures in power systems. In particular, this suggests that, by using a different control design for d_j, we can obtain different and potentially better system behaviors after a line failure. For instance, it is shown in Chapter 6 that bridge failures under droop control have a global impact, while (as we outline in Section 7.4) the impact of bridge failures can in fact be localized using UC. Our new proposed control strategy leverages precisely this extra freedom in choosing the d_j's to offer stronger guarantees in both failure mitigation and localization compared to previous work.

Unified Controller (UC)

UC is a control approach recently proposed in the frequency regulation literature [44, 82–84]. Compared to classical droop control or Automatic Generation Control (AGC) [8], UC aims to achieve primary frequency control, secondary frequency control, and congestion management simultaneously at the frequency control timescale.

The key feature of UC that we use here is that the closed-loop equilibrium of (2.1) under UC solves the following optimization:

$$\min_{f,d,\theta} \quad \sum_{j\in\mathcal{N}} c_j(d_j) \tag{7.3a}$$

$$\text{s.t.} \quad r - d - Cf = 0 \tag{7.3b}$$

$$f = BC^T\theta \tag{7.3c}$$

$$ECf = 0 \tag{7.3d}$$

$$\underline{f}_e \le f_e \le \overline{f}_e, \quad e \in \mathcal{E} \tag{7.3e}$$

$$\underline{d}_j \le d_j \le \overline{d}_j, \quad j \in \mathcal{N}, \tag{7.3f}$$

where $c_j(\cdot)$'s are associated cost functions that penalize deviations from last optimal dispatch point (and hence attain minimum at 0), (7.3b) guarantees power balance at each bus, (7.3c) is the DC power flow equation, (7.3d) enforces zero area control error [8], (7.3e) and (7.3f) are the flow and control limits. The matrix E encodes control area information as follows: Given a partition $\mathcal{P}^{\mathrm{UC}} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$ that specifies the control areas in secondary frequency control, $E \in \{0,1\}^{|\mathcal{P}^{\mathrm{UC}}|\times|\mathcal{N}|}$ is defined by $E_{l,j} = 1$ if bus j is in region $\mathcal{N}_l$ and $E_{l,j} = 0$ otherwise. An edge $e \in \mathcal{E}$ is called a *tie-line* if its endpoints belong to different regions in $\mathcal{P}^{\mathrm{UC}}$ [8, 83]. As a result, the l-th row of $ECf = 0$ ensures that the branch flow deviations on the tie-lines connected to $\mathcal{N}_l$ sum to zero.

UC is designed so that its controller dynamics combined with the system dynamics (2.1) form a variant of projected primal-dual algorithms to solve (7.3). It is shown in [83, 84] that when the optimization problem (7.3) is feasible, under mild assumptions UC is globally asymptotically stable and converges to an optimal point of (7.3). Such an optimal point is unique (up to a constant shift of θ) if the cost functions $c_j(\cdot)$ are strictly convex. We refer readers to [83, 84] for its exact controller design and analysis.

7.2 Connecting UC to Tree-partition

We have introduced two distinct partitions of a power network so far: the tree-partition $\mathcal{P}^{\mathrm{tree}}$ and the control area partition $\mathcal{P}^{\mathrm{UC}}$. In general, $\mathcal{P}^{\mathrm{tree}}$ and $\mathcal{P}^{\mathrm{UC}}$ can be different. However, when they do coincide, the underlying power grid inherits analytical properties from both tree-partition and UC, making the system particularly robust against failures. Our proposed control strategy leverages this connection, as we present in more detail in Section 7.3, and we henceforth assume that $\mathcal{P}^{\mathrm{tree}} = \mathcal{P}^{\mathrm{UC}}$. Under this assumption, the bridges and the tie-lines of the power network $\mathcal{G}$ also coincide.

Definition 7.2.1. *Given a cascading failure process described by* $\mathcal{B}(\mathfrak{n})$*, with* $\mathfrak{n} \in \{1, 2, \ldots, N\}$*, the set* $\mathcal{B}(1)$ *is said to be its* **initial failure***.*

In a power system, it is reasonable to expect that different initial failures can have different levels of impact on the rest of the network. For instance, the disconnection of a single solar panel from the grid is unlikely to cause any disruption to the system operation, while the failure of a transmission line that connects a major generator to the grid may incur significant load shedding. We thus need to distinguish different types of failures and ensure that the proposed control scheme reacts accordingly.

Definition 7.2.2. *An initial failure* $\mathcal{B}(1)$ *is said to be* **critical** *if the UC optimization* (7.3) *is infeasible over* $\mathcal{G}(1) := (\mathcal{N}, \mathcal{E}\backslash\mathcal{B}(1))$*, or* **non-critical** *if it is not critical.*

To formally state our localization result, we define the following concept to clarify the precise meaning of a region being "local" with respect to an initial failure.

Definition 7.2.3. *Given an initial failure* $\mathcal{B}(1)$*, we say that a tree-partition region* $\mathcal{N}_l$ *is* **associated** *with* $\mathcal{B}(1)$ *if there exists an edge* $e = (i, j) \in \mathcal{B}(1)$ *such that either* $i \in \mathcal{N}_l$ *or* $j \in \mathcal{N}_l$*.*

As we discuss in Section 7.3, our control strategy provides strong guarantees in mitigation and localization for both non-critical and critical failures, in a way that only the operation of the associated regions are adjusted whenever possible.

7.3 Proposed Control Strategy

Our control strategy revolves around the new and powerful properties of the power system that emerge when the control areas that UC operates over form a tree-partition of the network. In this section, we outline how this strategy can be implemented, in both the planning phase, during which a tree-partition structure of the control areas should be created, and the operating phase, during which UC actively monitors and reacts to line failures. Figure 7.2 illustrates the sequence of events after an initial failure in the proposed control strategy.

Planning Phase: the Tree-partition of Control Areas

Power networks are often comprised of multiple control areas, each of which is managed by an independent system operator (ISO). Although these areas exchange power with each other as prescribed by economic dispatch, their operations are relatively independent and it is desirable to ensure that system disturbances in one

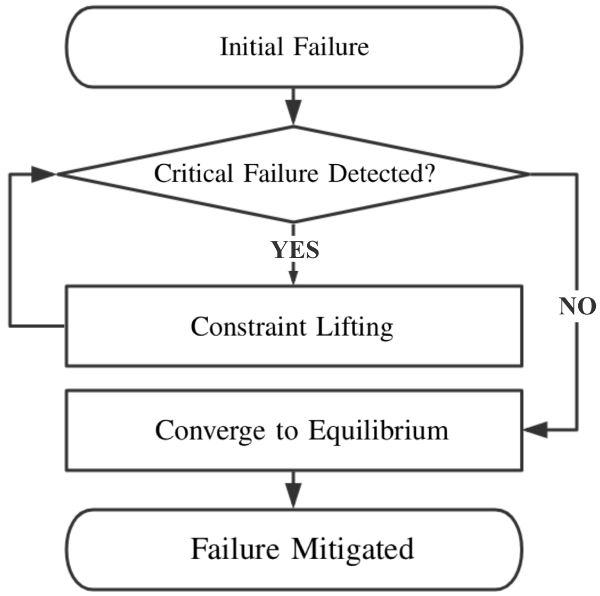


Figure 7.2: Flowchart of the events after an initial failure under the proposed control strategy.

area do not have a significant impact on the other areas. This is usually achieved via the zero area control error constraint in secondary frequency control [8], and is enforced in UC with (7.3d). As we mentioned in Section 7.2, such control areas typically do not form a tree-partition of the transmission network, since having redundant lines is believed to be a crucial part in maintaining $N-1$ security of the power system [8, 12, 34].

In order to implement our control strategy, we propose to create a tree-partition whose regions are precisely the control areas over which UC operates. This can be done by switching off a subset of the tie-lines so that the reduced multi-graph obtained from the control area partition forms a tree. The switching actions only need to be carried out in the planning phase, as line failures that occur during the operating phase do not affect the tree-partition already in place[3]. It is interesting to note that, when the subset of lines to switch off is chosen carefully, this action not only helps localize the impact of line failures, but can also improve the system reliability in the $N-1$ security sense. This seemingly counter-intuitive phenomenon is illustrated by our case studies in Section 7.6.

[3]In fact, in certain cases line failures lead to “finer” tree-partitions as more regions are potentially created when lines are removed from service.

Operating Phase: Extending the Unified Controller

Once a tree-partition is formed, the power network under UC operates as a closed-loop system and responds to disturbances such as transmission line failure or loss of generator/load in an autonomous manner. In normal conditions where the system disturbances are insignificant, UC always drives the power network back to an equilibrium point that can be interpreted as an optimal solution of (7.3). This is the case, for instance, when non-critical failures (see Definition 7.2.2) happen, and therefore such failures are always properly mitigated.

However, in extreme scenarios where a major disturbance (e.g., a critical failure) affects the system, the optimization problem (7.3) that UC aims to solve can be infeasible. In other words, it is physically impossible for UC to achieve all of its control objectives after such a disturbance. This causes UC to be unstable (see Proposition 7.5.1) and, further, leads to successive failures or even large scale outages. As such, there is a need to extend the version of UC proposed in [83, 84] with two features: (a) a critical failure detection component that monitors the system states and ensures UC is aware of such extreme situation promptly when it happens; and (b) a constraint lifting component that responds to critical failures by proactively relaxing certain goals that UC tries to achieve, and ensures system stability can be reached at minimal cost.

Our technical results in Section 7.5 suggest a way to implement both components as part of the normal operation of UC. System operators can prioritize different control areas by specifying the sequence of constraints to lift in response to extreme events. This allows the non-associated regions to be progressively involved and coordinated in a desired pattern when mitigating critical failures. We present and discuss some potential schemes in Section 7.5.

Guaranteed Mitigation and Localization

As we show in detail in Sections 7.4 and 7.5, our control strategy provides strong guarantees in mitigation and localization for both non-critical and critical failures. More specifically, the proposed control strategy ensures that: (a) non-critical failures are always fully mitigated by the associated regions, and the operating points for non-associated regions are not impacted at all; and (b) critical failures are guaranteed to be mitigated with certain constraints in (7.3) being lifted, in a progressive manner specified by the system operator. Thus the proposed strategy always prevents successive failures from happening, while localizing the impact of the initial failures as much as possible.

7.4 Localizing Non-critical Failures

In this section, we consider non-critical failures (as defined in Section 7.2), and prove that such failures are always fully mitigated within the associated regions.

We first characterize how the system operating point shifts in response to such failures. Recall that if an initial failure $\mathcal{B}(1)$ is non-critical, the UC optimization (7.3) is feasible and thus the new system operating point

$$x^*(1) := (\theta^*(1), \omega^*(1), d^*(1), f^*(1))$$

under UC control satisfies all the constraints in (7.3). In particular, none of the line limits in (7.3e) is violated at $x^*(1)$, i.e., $x^*(1)$ is a secure operating point and the cascade stops, namely $\mathcal{F}(1) = \emptyset$.

Lemma 7.4.1. *Given a non-critical initial failure $\mathcal{B}(1)$, the new operating point $x^*(1)$ prescribed by the UC satisfies $f_e^*(1) = 0$ for every bridge e.*

The above lemma shows that, in addition to the zero area control error constraints enforced by (7.3d), when the control areas that UC operates over form a tree-partition, UC further guarantees zero flow deviations on all tie-lines. This demonstrates how a tree-partition enables UC to achieve a stronger performance guarantee compared to its original form as proposed in [83, 84]. The following proposition is another result of this type, which clarifies how tree-partition brings localization properties to UC.

Proposition 7.4.2. *Assume $c_j(\cdot)$ is strictly convex and achieves its minimum at 0 for all $j \in \mathcal{N}$. Given a non-critical initial failure $\mathcal{B}(1)$, if a tree-partition region $\mathcal{N}_l$ is not associated with $\mathcal{B}(1)$, then $d_j^*(1) = 0$ for all $j \in \mathcal{N}_l$.*

The core idea underlying the proof of this proposition is easy to explain: Lemma 7.4.1 implies that the tie-line flows, which are the only coupling among the regions, are zero; thus the UC optimization (7.3) over different regions are totally "separated" and hence, the operating points for non-associated regions should remain unchanged. A rigorous proof is, however, more involved, requiring a technical result that relates the solution space of CBC^T to tree-partitions.

Lemma 7.4.3. *Let $\mathcal{P}^{\text{tree}} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_l\}$ be a tree-partition of $\mathcal{G}$ and consider a vector $b \in \mathbb{R}^{|\mathcal{N}|}$ such that $b_j = 0$ for all $j \in \mathcal{N}_1$ and $\sum_{j\in\mathcal{N}_k} b_j = 0$ for $k \neq 1$. Set*

$$\partial\mathcal{N}_1 := \{j : j \notin \mathcal{N}_1, \exists i \in \mathcal{N}_1 \text{ s.t. } (i,j) \in \mathcal{E} \text{ or } (j,i) \in \mathcal{E}\}$$

and $\overline{\mathcal{N}}_1 = \mathcal{N}_1 \cup \partial\mathcal{N}_1$. Then the linear system

$$CBC^T x = b \tag{7.4}$$

is solvable, and any solution x to (7.4) *satisfies $x_i = x_j$ for all $i, j \in \overline{\mathcal{N}}_1$.*

The set $\partial\mathcal{N}_1$ defined above are the "boundary" buses of $\mathcal{N}_1$ in $\mathcal{G}$ and $\overline{\mathcal{N}}_1$ can be interpreted as the closure of $\mathcal{N}_1$. It has a simple interpretation in the DC power flow context. Think of b as bus injections and x as the phase angles. Suppose that the injection at every node in $\mathcal{N}_1$ is zero and that the injections within every other region $\mathcal{N}_k$ are balanced (i.e., sum to zero). Then Lemma 7.4.3 says that the phase angles are the same at every node in $\overline{\mathcal{N}}_1$, i.e., the angle difference across every line in or incident to $\mathcal{N}_1$ is zero. This result only holds if the underlying regions form a tree-partition (its proof is presented in Section 7.8).

Proof sketch of Proposition 7.4.2. For the purpose of simplified notations, we drop the stage index (1) from x^* and denote $x^* = (\theta^*, \omega^*, d^*, f^*)$. To streamline the presentation, we only sketch the main ideas of the proof here, leaving the details to Section 7.8.

First, we construct a different point $\tilde{x}^*$ from x^* as follows: (a) replace d_j^* with 0 for all $j \in \mathcal{N}_l$; (b) replace f_e^* with 0 for $e \in \mathcal{E}$ that have both endpoints in $\mathcal{N}_l$; and (c) replace θ^* by a solution $\tilde{\theta}^*$ obtained from solving DC power flow equations with injections specified by $\tilde{d}^*$. Since $c_j(\cdot)$ attains its minimum at 0, $\tilde{x}^*$ achieves at least the same objective value (7.3a) as x^*. Thus $\tilde{x}^*$ must be an optimal point of (7.3), provided it is feasible.

Second, as the core step in the whole proof, we apply Lemma 7.4.3 to all regions of $\mathcal{P}^{\text{tree}}$ separately, and show that $\tilde{\theta}^*$ is consistent with the injections and branch flows specified by $\tilde{x}^*$. This, together with routine checks, allows us to prove the feasibility of the point $\tilde{x}^*$.

Finally, when the cost functions $c_j(\cdot)$ are strictly convex, the optimal solution to (7.3) is unique in d^* and f^* (θ^* is also unique up to a constant shift). We thus conclude that $\tilde{x}^* = x^*$ (up to a constant shift on θ). This completes the proof. □

This result reveals that, with the proposed control strategy, when the system converges to an equilibrium after a non-critical failure, the injections and power flows in the non-associated regions remain unchanged. In other words, our control scheme guarantees that non-critical failures in a control area do not impact the operations of other areas at all, achieving a stronger control area independence than that ensured by the zero control error requirement.

Unlike the tree-partition only approach in Chapter 6, bridge failures in this proposed control strategy are treated in exactly the same way as other lines, provided that they are non-critical. Furthermore, the impact of such bridge failures is localized to the associated regions. This contrast with the global impact of bridge failures in Chapter 6 demonstrates again the benefits of connecting UC to tree-partitions.

7.5 Controlling Critical Failures

We now consider the case where the initial failure is critical. This may happen when a major generator or transmission line is disconnected from the grid.

Unified Controller under Critical Failures

Since UC is a concept that emerged from the frequency regulation literature, the underlying optimization (7.3) is always assumed to be feasible in existing studies [44, 82–84]. As such, little is known about the behaviors of UC if this assumption is violated, which is the case when a critical failure happens. We now derive a result that closes this gap and characterizes the limiting behavior of UC in this setting.

In order to do so, we first need to formulate the exact controller dynamics of UC. Unfortunately, there is no standard way to do so as multiple designs of UC have been proposed in the literature [44, 82–84], each with its own strengths and weaknesses. Nevertheless, all of the proposed controller designs are (approximately) projected primal-dual algorithms to solve the underlying optimization (7.3), and satisfy the following assumptions:

UC1 : For all $j \in \mathcal{N}$, $\underline{d}_j \leq d_j(t) \leq \overline{d}_j$ is satisfied for all t. This is achieved either via a projection operator that maps $d_j(t)$ to this interval, or by requiring the cost function $c_j(\cdot)$ to approach infinity near these boundaries.

UC2 : Dual variables are introduced for constraints (7.3b)-(7.3e) and maintained throughout the operation. Denote these dual variables by λ_i for

$$i \in \left\{1, 2, \cdots, |\mathcal{N}| + 3\,|\mathcal{E}| + \left|\mathcal{P}^{\mathrm{UC}}\right|\right\}.$$

UC3 : The primal variables f, θ and the dual variables λ_i are updated by a primal-dual algorithm[4] to solve (7.3).

Proposition 7.5.1. *Assume UC1-UC3 hold. If* (7.3) *is infeasible, then there exists a dual variable λ_i such that:*

$$\limsup_{t\to\infty} |\lambda_i(t)| = \infty.$$

This result implies that after a critical failure, UC cannot drive the system to a proper and safe operating point. In fact, it always leads to instability in the system (certain dual variables can take arbitrarily large values). This drawback, however, when viewed from a different perspective, suggests a way to detect critical failures. More specifically, since Proposition 7.5.1 guarantees certain dual variables will become arbitrarily large in UC operation when (7.3) is infeasible, we can always set a threshold for the dual variables and raise an infeasibility warning if some of them exceed the corresponding thresholds. By doing so, critical failures can always be detected, and this happens in a distributed fashion in parallel to the normal operation of UC. Moreover, by setting tighter thresholds around the normal operating point, such failures can be detected more promptly.

Of course, this method is subject to false alarms, since non-critical failures may also cause relatively large dual variable values in transient state. There is an intrinsic tradeoff on the level of the thresholds to be applied, in the following sense: A tighter threshold allows critical failures to be detected more promptly, yet also leads to a larger false alarm rate. In practice, these thresholds should be chosen carefully by the operator in accordance to the specific system parameters and application scenarios.

Constraint Lifting as a Remedy

In the event of a critical failure, it is physically impossible for UC to simultaneously achieve all of its control objectives. Our discussion in the last subsection shows that, if UC still operates following its normal dynamics, the system is subject to instability and thus successive failures. In the worst case, this can lead to large scale outages.

[4]We do not consider the specific variants of the standard primal-dual algorithms that are proposed in different designs of UC, since the standard primal-dual algorithm is often a good approximation.

We can prevent this from happening by lifting certain constraints from UC. Without compromising the basic objective to stabilize the system, there are two ways to do so:

- The zero area control error constraints (7.3d) between certain control areas can be lifted. This in practice means the controller now gets more control areas involved to mitigate the failure.
- Certain load shedding can be applied, which in (7.3) is reflected by enlarging the range $[\underline{d}_j, \overline{d}_j]$ for the corresponding load buses.

By iteratively lifting the two types of constraints above, one can guarantee the feasibility of (7.3) and ensure that the system under the proposed control converges to a stable point, which in particular is free from successive failures. This, however, comes with the cost of potential load loss, and thus must be carried out properly. In practice, the iterative relaxation procedure can follow predetermined rules specified by the system operator to prioritize different objectives.

7.6 Case Studies

In this section, we evaluate the performance of the proposed control strategy on the IEEE 118-bus test system, which comprises of two control areas (as shown in Figure 7.3). The three dashed lines (15, 33), (19, 34), and (23, 24) are switched off whenever a tree-partition needs to be formed, and the new topology is referred to as the revised network.

$N-1$ Security

We first evaluate the system robustness to failures in terms of the $N-1$ security standard, where single line failure scenarios are examined. Those failure scenarios are created as follows: First, we generate 100 load injections by adding random perturbations (up to 25% of the base value) to the nominal load profile from [85] and then solve the DC OPF to obtain the corresponding generator operating points. Second, we iterate over every transmission line in the IEEE 118-bus test system as initial failures and simulate the cascading process thus triggered. This produces about 18,000 scenarios.

We implement both the proposed control strategy and the classical AGC [8] on the IEEE 118-bus testbed, and compare the average number of vulnerable lines across all the single line failure scenarios that lead to either successive failures or load

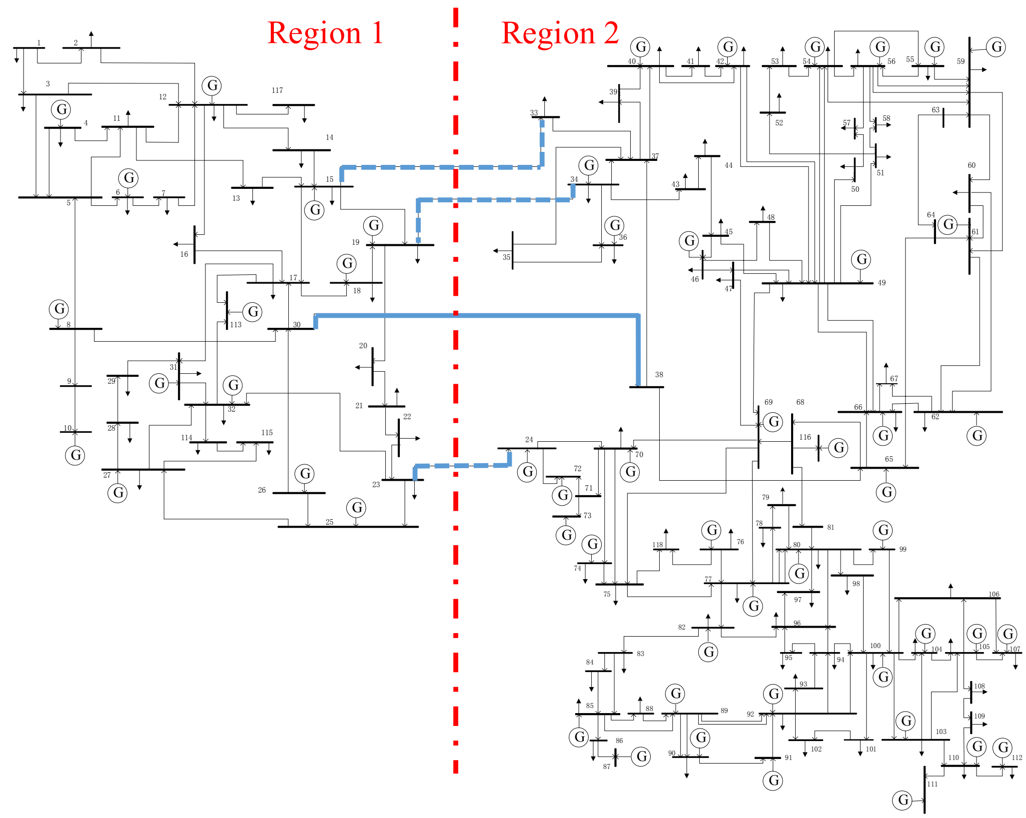


Figure 7.3: One line diagram of the IEEE 118-bus test system with two control areas. Dashed blue lines are switched off when a tree-partition needs to be formed.

shedding when they are tripped. In order to illustrate the improvements of the proposed control strategy in different levels of system congestion, we scale down the transmission line capacities to $\alpha = 0.9, 0.8, 0.7$ of the base values and collect statistics on the number of vulnerable lines in all these settings.

Our results are summarized in Figure 7.4. It can be seen that the proposed control incurs a far lower number of vulnerable lines in all cases compared to AGC, and this difference is particularly prominent when the system is congested. We highlight that this happens with the proposed control operating over the revised network, where some of the tie-lines are switched off and hence certain capacity is removed from the system. Moreover, the remaining tie-line (30, 38) in the revised network is never vulnerable under the proposed control.

Loss of Load and Disruption to System Operation

We now look at the load loss rate, defined as the ratio between the total loss of load with respect to the original total demand, of the system to evaluate how well

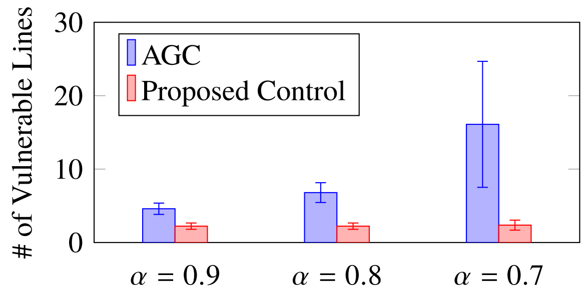


Figure 7.4: Number of vulnerable lines with respect to different levels of system congestion.

failures are mitigated in different settings. In this experiment, we scale down the generator capacities by 35% and the line capacities by 30%, so that the system is more susceptible to failures. In order to demonstrate how UC and tree-partition impact the system performance separately, we look at four different settings:

- AGC on the original network;
- AGC on the revised network;
- UC on the original network; and
- UC on the revised network.

Figure 7.5 plots the complementary cumulative distribution (CCDF) of the load loss rates across all of the failure scenarios in these settings.

As one can see from the figure, for both the original and revised networks, UC significantly outperforms AGC. In particular, the largest load loss rate for UC is less than 2% for both networks, while AGC can lead to loss rate up to 14% on the revised network and 21% on the original network. This demonstrates the benefits of using our control strategy to mitigate failures.

Although the performance of UC in terms of loss rate are roughly the same with or without tree-partition, there is a drastic difference when we look at how well the failure impacts are localized. In Figure 7.6, we plot the CCDF on the number of generators whose operating points are adjusted in response to the initial failures. It shows that the operation of much fewer generators is disrupted when the control areas that UC operates over form a tree-partition. This confirms our intuition and theoretical results about how a tree-partition structure helps localize failures.

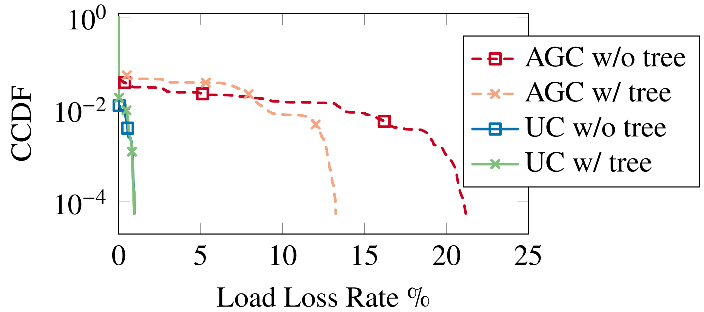


Figure 7.5: CCDF for load loss rate.

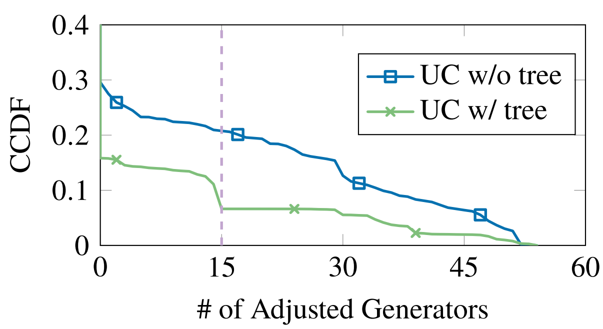


Figure 7.6: CCDF for generator response.

7.7 Conclusion

In this chapter, we proposed a control strategy that combines the concepts of the unified controller and the network tree-partition to mitigate and localize cascading failures in power system. Our case studies on the IEEE 118-bus test system show that the proposed control scheme greatly improves system robustness to cascading failures as compared to classical AGC. In particular, this new control prevents successive failures from happening while localizing the impacts of initial failures at the same time. Moreover, when load shedding is inevitable, the proposed strategy incurs significantly less load loss.

7.8 Proofs

Proof of Lemma 7.4.1

To simplify the notations, we drop the stage index (1) from x^* and denote $x^* = (\theta^*, \omega^*, d^*, f^*)$. Given a bridge $e = (j_1, j_2)$ of $\mathcal{G}$, removing e from $\mathcal{G}$ partitions $\mathcal{G}$ into two connected components, say C_1 and C_2. Without loss of generality, assume $j_1 \in C_1$ and $j_2 \in C_2$. For a region $\mathcal{N}_v$ from $\mathcal{P}$, we say $\mathcal{N}_v$ is within C_1 if for any $j \in \mathcal{N}_v$ we have $j \in C_1$. It is easy to check from the definition of tree-partitions that any region $\mathcal{N}_v$ from $\mathcal{P}$ is either within C_1 or within C_2, and e is the only edge in $\mathcal{G}$ that has one endpoint in C_1 and the other endpoint in C_2.

Let $\mathcal{P}'$ be the set of regions within C_1 from $\mathcal{P}$, and put $\mathbf{1}_{\mathcal{P}'} \in \{0,1\}^{|\mathcal{P}|}$ to be its characteristic vector (that is, the l-th component of $\mathbf{1}_{\mathcal{P}'}$ is 1 if $\mathcal{N}_l \in \mathcal{P}'$ and 0 otherwise). Given two buses i and j, we denote $i \to j$ if $(i,j) \in \mathcal{E}$ and $j \to i$ if $(j,i) \in \mathcal{E}$. With such notations, from (7.3d), we have

$$
\begin{aligned}
0 &= \mathbf{1}_{\mathcal{P}'}^T ECf^* \\
&= \sum_{l:\mathcal{N}_l \in \mathcal{P}'} \sum_{i \in \mathcal{N}_l} \left(\sum_{j:j\to i} f_{ji}^* - \sum_{j:i\to j} f_{ij}^* \right) \\
&= \sum_{i:i \in C_1} \left(\sum_{j:j\to i} f_{ji}^* - \sum_{j:i\to j} f_{ij}^* \right) \\
&= f_e^* + \sum_{i:i \in C_1} \left(\sum_{j:j\to i, j \in C_1} f_{ji}^* - \sum_{j:i\to j, j\in C_1} f_{ij}^* \right), \qquad (7.5)
\end{aligned}
$$

where (7.5) is because the only edge with one endpoint in C_1 and the other endpoint in C_2 is e. Note that

$$
\begin{aligned}
&\sum_{i:i \in C_1} \left(\sum_{j:j\to i, j \in C_1} f_{ji}^* - \sum_{j:i\to j, j\in C_1} f_{ij}^* \right) \\
&= \sum_{(i,j)\in \mathcal{E}_1} \left(f_{ij}^* - f_{ij}^* \right) \\
&= 0,
\end{aligned}
$$

where $\mathcal{E}_1$ is the set of edges with both endpoints in C_1. From (7.5), we see that $f_e^* = 0$.

Since the bridge e is arbitrary, we have thus proved the desired result.

□

Proof of Proposition 7.4.2

We now prove the core step as mentioned in the main body of this chapter. Denote $\tilde{x}^* = (\tilde{\theta}^*, \tilde{d}^*, \tilde{f}^*)$. From the way that $\tilde{x}^*$ is constructed, the constraints (7.3d) are easily seen to be satisfied. If we can show that $\tilde{f}^* = BC^T\tilde{\theta}^*$, then since $\tilde{\theta}^*$ is obtained by solving the DC power flow equations from $CBC^T\tilde{\theta}^* = r - \tilde{d}^*$, the constraints (7.3b) and (7.3c) are also satisfied. Now we show that $\tilde{f}^* = BC^T\tilde{\theta}^*$ indeed holds.

To do so, we first establish the following lemma:

Lemma 7.8.1. *For any tree-partition region $\mathcal{N}_z$ in $\mathcal{P}$, we have*

$$\sum_{j\in\mathcal{N}_z}\left(r_j - d_j^*\right) = \sum_{j\in\mathcal{N}_z}\left(r_j - \tilde{d}_j^*\right) = 0.$$

Proof. Let $\mathbf{1}_{\mathcal{N}_z} \in \mathbb{R}^{|\mathcal{N}|}$ be the characteristic vector of $\mathcal{N}_z$, that is, the j-th component of $\mathbf{1}_{\mathcal{N}_z}$ is 1 if $j \in \mathcal{N}_z$ and 0 otherwise. Summing (7.3b) over $j \in \mathcal{N}_z$, we have:

$$\sum_{j\in\mathcal{N}_z}\left(r_j - d_j^*\right) = \mathbf{1}_{\mathcal{N}_z}^T Cf = (ECf)_z = 0,$$

where $(ECf)_z$ is the z-th row of ECf.

For $\mathcal{N}_z$ that is different from $\mathcal{N}_l$, we have $\tilde{d}_j^* = d_j^*$ for any $j \in \mathcal{N}_z$ by construction. Thus for such $\mathcal{N}_z$ we also have

$$\sum_{j\in\mathcal{N}_z}\left(r_j - \tilde{d}_j^*\right) = 0.$$

For $\mathcal{N}_l$, since $\mathcal{N}_l$ is not associated with $\mathcal{B}(1)$, we have $r_j = 0$ for $j \in \mathcal{N}_l$. Moreover, by construction we also know that $\tilde{d}_j^* = 0$ for $j \in \mathcal{N}_l$. As a result

$$\sum_{j\in\mathcal{N}_l}\left(r_j - \tilde{d}_j^*\right) = 0.$$

This completes the proof. □

Now consider a region $\mathcal{N}_w$ that is different from $\mathcal{N}_l$. In this case, we do not change the injection from x^* when constructing $\tilde{x}^*$, thus $d_j^* - \tilde{d}_j^* = 0$ for all $j \in \mathcal{N}_w$. From Lemma 7.8.1, we see that $\sum_{j\in\mathcal{N}_z}\left(d_j^* - \tilde{d}_j^*\right) = 0$ for all z. Since d^* and θ^* conform to the DC power flow equations, we have

$$CBC^T\theta^* = r - d^*$$

and thus

$$CBC^T\left(\theta^* - \tilde{\theta}^*\right) = \tilde{d}^* - d^*.$$

By Lemma 7.4.3, we then have $\theta_j^* - \tilde{\theta}_j^*$ is a constant over $\overline{\mathcal{N}}_w$, and thus

$$\tilde{\theta}_i^* - \tilde{\theta}_j^* = \theta_i^* - \theta_j^*$$

for all $i, j \in \overline{\mathcal{N}}_w$. This in particular implies

$$\tilde{f}_e^* = f_e^* = B_e(\theta_i^* - \theta_j^*) = B_e(\tilde{\theta}_i^* - \tilde{\theta}_j^*)$$

for all $e = (i, j)$ such that $i \in \mathcal{N}_w$ or $j \in \mathcal{N}_w$.

Next, let us consider the region $\mathcal{N}_l$. In this region, we have $\tilde{d}_j^* = 0$ by construction. Moreover, since $\mathcal{N}_l$ is not associated with $\mathcal{B}(1)$, we know $r_j = 0$ for all $j \in \mathcal{N}_l$. Thus $r_j - \tilde{d}_j^* = 0$ for all $j \in \mathcal{N}_l$. Further, from Lemma 7.8.1 we have $\sum_{j \in \mathcal{N}_z} \left(r_j - \tilde{d}_j^* \right) = 0$ for all z. Thus by Lemma 7.4.3 and $CBC^T\tilde{\theta}^* = r - \tilde{d}^*$, we know $\tilde{\theta}_i^* = \tilde{\theta}_j^*$ for all $i, j \in \overline{\mathcal{N}_l}$. This implies that for any edge $e = (i, j)$ within $\mathcal{N}_l$, we have

$$\tilde{f}_e^* = 0 = B_e(\tilde{\theta}_i^* - \tilde{\theta}_j^*).$$

As a result, we see that $\tilde{f}_e^* = B_e(\tilde{\theta}_i^* - \tilde{\theta}_j^*)$ holds for all $e \in \mathcal{E}$. This completes the proof. □

Proof of Lemma 7.4.3

As we discussed in Chapter 2, the Laplacian matrix $L := CBC^T$ of a connected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ has rank $n - 1$, and $Lx = b$ is solvable if and only if $\mathbf{1}^T b = 0$, where $\mathbf{1}$ is the vector with a proper dimension that consists of ones. Moreover, the kernel of L is given by span$(\mathbf{1})$.

If $\mathcal{N}_1$ is the only region in $\mathcal{P}$, then $b = 0$ since $b_j = 0$ for all $j \in \mathcal{N}_1$. We thus know the solution space to $Lx = b$ is exactly the kernel of L, and the desired result holds.

If $\mathcal{N}_1$ is not the only region in $\mathcal{P}$, then we can find a bus that does not belong to $\mathcal{N}_1$, say bus z. Without loss of generality, assume the bus $z \in \mathcal{N}_k$ and corresponds to the last row and column in L. Consider a solution x to $Lx = b$. Since the kernel of L is span$(\mathbf{1})$, we can without loss of generality assume that the last component of x is 0. Let $\overline{L}$ be the submatrix of L obtained by removing its last row and last column, and similarly let $\overline{x}$ and $\overline{b}$ be the vectors obtained by removing the last component of x and b, respectively. Then $\overline{L}$ is invertible (see Chapter 2), and we have

$$\overline{L}\overline{x} = \overline{b}.$$

Denote the matrix obtained by deleting the l-th row and i-th column of $\overline{L}$ by $\overline{L}^{li}$, then by Proposition 2.4.2 we have

$$\det\left(\overline{L}^{li}\right) = (-1)^{l+i} \sum_{E \in \mathcal{T}(\{l,i\},\{z\})} \chi(E), \tag{7.6}$$

where $\chi(E) = \prod_{e \in E} B_e$ and $\mathcal{T}(\{l, i\}, \{z\})$ is the set of spanning forests of $\mathcal{G}$ that consists of exactly two trees containing $\{l, i\}$ and $\{z\}$, respectively. We refer readers to Chapter 2 for a detailed discussion on how to interpret these notations.

To state some useful results derived from (7.6), we introduce the following definition of directly connected regions:

Definition 7.8.2. *For a tree-partition $\mathcal{P} = \{\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_k\}$ of $\mathcal{G}$, we say $\mathcal{N}_v$ and $\mathcal{N}_w$ are* **directly connected without** $\mathcal{N}_l$ *if the path from $\mathcal{N}_v$ to $\mathcal{N}_w$ in $\mathcal{G}_\mathcal{P}$ does not contain $\mathcal{N}_l$.*

The path from $\mathcal{N}_v$ to $\mathcal{N}_w$ in the above definition is unique since $\mathcal{G}_\mathcal{P}$ forms a tree. As an example, in Fig. 5.2, $\mathcal{N}_1$ and $\mathcal{N}_2$ are directly connected without $\mathcal{N}_3$, yet $\mathcal{N}_2$ and $\mathcal{N}_3$ are not directly connected without $\mathcal{N}_1$.

In the following proofs we need to refer to paths in both the original graph $\mathcal{G}$ and the reduced graph $\mathcal{G}_\mathcal{P}$. To clear potential confusions, we agree to the following terminologies: Given two sets of nodes $\mathcal{N}_v$ and $\mathcal{N}_w$ (that can be different from the tree-partition regions in $\mathcal{P}$) of $\mathcal{G}$, a path in $\mathcal{G}$ from $\mathcal{N}_v$ to $\mathcal{N}_w$ refers to a path consisting of nodes (and lines) from the original graph $\mathcal{G}$ whose starting node belongs to $\mathcal{N}_v$ and ending node belongs to $\mathcal{N}_w$. Given two tree-partition regions $\mathcal{N}_v$ and $\mathcal{N}_w$, a path in $\mathcal{G}_\mathcal{P}$ from $\mathcal{N}_v$ to $\mathcal{N}_w$ refers to a path consisting of nodes (and lines) from the reduced graph $\mathcal{G}_\mathcal{P}$ whose starting node is $\mathcal{N}_v$ and ending node is $\mathcal{N}_w$. Since there is a natural correspondence between bridges in $\mathcal{G}$ and lines in $\mathcal{G}_\mathcal{P}$, if a line e in $\mathcal{G}_\mathcal{P}$ is contained in a path P in $\mathcal{G}_\mathcal{P}$, we also say the corresponding bridge $\tilde{e}$ from $\mathcal{G}$ is contained in P.

Lemma 7.8.3. *Assume $\mathcal{N}_2$ and $\mathcal{N}_k$ are not directly connected without $\mathcal{N}_1$. If $l_1, l_2 \in \mathcal{N}_2$ and $i \in \overline{\mathcal{N}}_1$, then*

$$\mathcal{T}(\{l_1, i\}, \{z\}) = \mathcal{T}(\{l_2, i\}, \{z\}).$$

Proof. The path from $\mathcal{N}_1$ to $\mathcal{N}_k$ in $\mathcal{G}_\mathcal{P}$ contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$. Denote this bridge as $\tilde{e}$ and let w be the endpoint of $\tilde{e}$ that is not in $\mathcal{N}_1$. Then it is easy to check that w is a cut node that any path from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ in $\mathcal{G}$ must contain.

Since $\mathcal{N}_2$ and $\mathcal{N}_k$ are not directly connected without $\mathcal{N}_1$, the path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}_\mathcal{P}$ passes through $\mathcal{N}_1$. In other words, any path in $\mathcal{G}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$ must pass through a certain node in $\mathcal{N}_1$, and thus contains a sub-path in $\mathcal{G}$ from $\mathcal{N}_1$ to $\mathcal{N}_k$. This implies that w is contained in any path in $\mathcal{G}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$.

Note that any tree containing $i \in \overline{\mathcal{N}}_1$ and $l_1 \in \mathcal{N}_2$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ and thus contains w. Further, any tree containing $l_2 \in \mathcal{N}_2$ and $z \in \mathcal{N}_k$ induces a path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}$, and thus also contains w. As a result, these two types of trees always share a common node w and cannot be disjoint:

$$\mathcal{T}(\{l_1, i\}, \{l_2, z\}) = \emptyset.$$

Similarly $\mathcal{T}(\{l_2, i\}, \{l_1, z\}) = \emptyset$. Therefore

$$\begin{aligned} \mathcal{T}(\{l_1, i\}, \{z\}) &= \mathcal{T}(\{l_1, l_2, i\}, \{z\}) \sqcup \mathcal{T}(\{l_1, i\}, \{l_2, z\}) \\ &= \mathcal{T}(\{l_1, l_2, i\}, \{z\}) \sqcup \mathcal{T}(\{l_2, i\}, \{l_1, z\}) \\ &= \mathcal{T}(\{l_2, i\}, \{z\}), \end{aligned}$$

where $\sqcup$ means disjoint union. The desired result then follows. □

Lemma 7.8.4. *Assume $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$. If $l \in \mathcal{N}_2$ and $i_1, i_2 \in \overline{\mathcal{N}_1}$, then*

$$\mathcal{T}(\{l, i_1\}, \{z\}) = \mathcal{T}(\{l, i_2\}, \{z\}).$$

Proof. The path from $\mathcal{N}_1$ to $\mathcal{N}_k$ in $\mathcal{G}_{\mathcal{P}}$ (denoted as P_1) contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$. Denote this bridge as $\tilde{e}$ and let w be the endpoint of $\tilde{e}$ that does not belong to $\mathcal{N}_1$. Then it is easy to check that w is a cut node that any path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_k$ must pass through.

We claim that if $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$, then any path from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ in $\mathcal{G}$ must also contain w. Indeed, suppose not, then the path from $\mathcal{N}_1$ to $\mathcal{N}_2$ in $\mathcal{G}_{\mathcal{P}}$ (denoted as P_2) contains a bridge in $\mathcal{G}$ that incidents to $\mathcal{N}_1$, and this bridge is different from $\tilde{e}$. If P_1 and P_2 do not have any common super nodes, then concatenating the two paths induces a path in $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{N}_2$ to $\mathcal{N}_k$ that passes through $\mathcal{N}_1$. In other words, the path from $\mathcal{N}_2$ to $\mathcal{N}_k$ in $\mathcal{G}_{\mathcal{P}}$ passes through $\mathcal{N}_1$, contradicting the assumption that $\mathcal{N}_2$ and $\mathcal{N}_k$ are directly connected without $\mathcal{N}_1$. Therefore, P_1 and P_2 share a common node, say $\mathcal{N}_3$. However, P_1 and P_2 induce two different sub-paths in $\mathcal{G}_{\mathcal{P}}$ from $\mathcal{N}_1$ to $\mathcal{N}_3$, contradicting the assumption that $\mathcal{G}_{\mathcal{P}}$ forms a tree. We thus have proved the claim.

Finally, note that any tree containing $i_1 \in \overline{\mathcal{N}}_1$ and $l \in \mathcal{N}_2$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_2$ and thus contains w. Further, any tree containing $i_2 \in \overline{\mathcal{N}}_1$ and $z \in \mathcal{N}_k$ induces a path in $\mathcal{G}$ from $\overline{\mathcal{N}}_1$ to $\mathcal{N}_k$ and thus contains w. Therefore these two types of trees always share a common node w and cannot be disjoint:

$$\mathcal{T}(\{l, i_1\}, \{i_2, z\}) = \emptyset.$$

Similarly $\mathcal{T}(\{l, i_2\}, \{i_1, z\}) = \emptyset$. As a result,

$$\begin{aligned}
\mathcal{T}(\{l, i_1\}, \{z\}) &= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \sqcup \mathcal{T}(\{l, i_1\}, \{i_2, z\}) \\
&= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \\
&= \mathcal{T}(\{l, i_1, i_2\}, \{z\}) \sqcup \mathcal{T}(\{l, i_2\}, \{i_1, z\}) \\
&= \mathcal{T}(\{l, i_2\}, \{z\}).
\end{aligned}$$

□

Now since $b_k = \overline{b}_k = 0$ for all $k \in \mathcal{N}_1$, by Cramer's rule, we have

$$x_i = \overline{x}_i = \frac{\sum_{l \notin \mathcal{N}_1} (-1)^{l+i} b_k \det\left(\overline{L}^{li}\right)}{\det\left(\overline{L}\right)} \tag{7.7}$$

for all i.

Let $\mathcal{P}_1$ be set of the regions in $\mathcal{P}$ that are directly connected to $\mathcal{N}_k$ without $\mathcal{N}_1$ and let $\mathcal{P}_2$ be the remaining regions. For a region $\mathcal{N}_l \in \mathcal{P}_1$, let

$$\chi(\mathcal{N}_l) := \sum_{E \in \mathcal{T}\left(\left\{\tilde{l}, i\right\}, \{z\}\right)} \chi(E),$$

where $\tilde{l}$ is an arbitrary bus in $\mathcal{N}_l$. $\chi(\mathcal{N}_l)$ is well-defined by Lemma 7.8.3. This, together with the assumption $\sum_{j \in \mathcal{N}_l} b_j = 0$, then implies

$$\begin{aligned}
\sum_{l \in \mathcal{N}_l} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right) &= \sum_{l \in \mathcal{N}_l} b_l \left(\sum_{E \in \mathcal{T}(\{l,i\},\{z\})} \chi(E) \right) = \sum_{l \in \mathcal{N}_l} b_l \chi(\mathcal{N}_l) \\
&= \chi(\mathcal{N}_l) \sum_{l \in \mathcal{N}_l} b_l \\
&= 0.
\end{aligned}$$

As a result

$$\begin{aligned}
\sum_{l \notin \mathcal{N}_1} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right) &= \sum_{\mathcal{N}_l \in \mathcal{P}_1} \sum_{l \in \mathcal{N}_l} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right) + \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{l \in \mathcal{N}_l} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right) \\
&= \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{l \in \mathcal{N}_l} (-1)^{l+i} b_l \det\left(\overline{L}^{li}\right) \\
&= \sum_{\mathcal{N}_l \in \mathcal{P}_2} \sum_{l \in \mathcal{N}_l} b_l \left(\sum_{E \in \mathcal{T}(\{l,i\},\{z\})} \chi(E) \right),
\end{aligned}$$

which by Lemma 7.8.4 takes the same value for all $i \in \overline{\mathcal{N}}_1$. In other words, the equation (7.7) takes the same value for all $i \in \overline{\mathcal{N}}_1$. This completes the proof. □

Proof of Proposition 7.5.1

First, let us put $x = [f, d, \theta] \in \mathbb{R}^{2|\mathcal{N}|+|\mathcal{E}|}$ to collect all the decision variables of the UC optimization (7.3) and rewrite it to a more generic form:

$$\begin{aligned}
\min_{\underline{d} \leq d \leq \overline{d}} \quad & c(d) && (7.8a)\\
\text{s.t.} \quad & Ax \leq g && (7.8b)\\
& Cx = h, && (7.8c)
\end{aligned}$$

where A, C, g, h are matrices (vectors) of proper dimensions from the optimization (7.3). Let λ_1, λ_2 be the corresponding dual variables to (7.8b) and (7.8c), respectively, and put $\lambda := [\lambda_1; \lambda_2]$ ($[\cdot;\cdot]$ here means matrix concatenation as a column), we can then write the Lagrangian for (7.8) as

$$\mathcal{L}(x, \lambda) = c(d) + \lambda_1^T (Ax - g) + \lambda_2^T (Cx - h).$$

Now by the assumption UC3, we know that:

$$\begin{aligned}
\dot{\lambda}_1 &= [Ax - g]^+_{\lambda_1} && (7.9a)\\
\dot{\lambda}_2 &= Cx - h, && (7.9b)
\end{aligned}$$

where the projection operator $[\cdot]^+_{\lambda_1}$ is defined component-wise by

$$([x]^+_{\lambda_1})_i := \begin{cases} x_i & \text{if } x_i > 0 \text{ or } \lambda_{1,i} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{7.10}$$

Consider two closed convex sets $S_1 = \{x | Ax \leq g, Cx = h\}$ and $S_2 = \{x | \underline{d} \leq d \leq \overline{d}\}$. If the optimization (7.3) is infeasible, then $S_1 \cap S_2 = \emptyset$, i.e., the sets S_1 and S_2 are disjoint. As a result, there exists a hyperplane that separates S_1 and S_2: $\exists p \in \mathbb{R}^{2|\mathcal{N}|+|\mathcal{E}|}, q \in \mathbb{R}$ such that

$$p^T x > q, \forall x \in S_1 \text{ and } p^T x \leq q, \forall x \in S_2.$$

This then implies the system

$$\begin{cases} Ax \leq g \\ Cx = h \\ p^T x \leq q \end{cases}$$

is not solvable. By Farkas' Lemma, we can then find w_1, w_2, w_3 of proper dimensions such that $w_1 \geq 0, w_3 \geq 0, A^T w_1 + C^T w_2 + p w_3 = 0$, and $g^T w_1 + h^T w_2 + q w_3 = \epsilon < 0$.

Define $z = [w_1; w_2]$. We then see that under the UC controller, we have for any t:

$$
\begin{aligned}
z^T \dot{\lambda}(t) &= w_1^T [Ax(t) - g]_\lambda^+ + w_2^T (Cx(t) - h) \\
&\geq w_1^T [Ax(t) - g]_\lambda^+ + w_2^T (Cx(t) - h) + w_3 (p^T x(t) - q) \qquad (7.11a) \\
&\geq w_1^T (Ax(t) - g) + w_2^T (Cx(t) - h) + w_3 (p^T x(t) - q) \qquad (7.11b) \\
&= \left(A^T w_1 + C^T w_2 + p w_3\right) x(t) - \left(w_1^T g + w_2^T h + w_3 q\right) \\
&= 0 - \epsilon \\
&> 0,
\end{aligned}
$$

where (7.11a) comes from $w_3 \geq 0$ and the assumption UC1, which ensures $x(t) \in S_2$ and thus $p^T x(t) - q \leq 0$, and (7.11b) comes from $w_1 \geq 0$ and the fact that $[x]_\lambda^+ \geq x$ for all x (the inequality is component-wise).

As a result, we see that

$$z^T \lambda(t) - z^T \lambda(0) > -\epsilon t$$

and thus

$$\lim_{t \to \infty} z^T \lambda(t) = \infty.$$

Finally, by noting

$$\lim_{t \to \infty} z^T \lambda(t) \leq w_1^T \limsup_{t \to \infty} |\lambda_1(t)| + |w_2|^T \limsup_{t \to \infty} |\lambda_2(t)|,$$

the desired result follows. $\square$

BIBLIOGRAPHY

[1] U.S.-Canada Power System Outage Task Force report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendation. 2004.

[2] Report of the enquiry committee on grid disturbance in Northern region on 30th July 2012 and in Northern, Eastern and North-Eastern region on 31st July 2012. Aug 2012.

[3] Cesar O Aguilar and Bahman Gharesifard. Graph controllability classes for the Laplacian leader-follower dynamics. *IEEE transactions on automatic control*, 60(6):1611–1623, 2014.

[4] Martin Andreasson, Dimos V Dimarogonas, Henrik Sandberg, and Karl H Johansson. Distributed PI-control with applications to power systems frequency control. In *American Control Conference (ACC), 2014*, pages 3183–3188. IEEE, 2014.

[5] Marian Anghel, Kenneth A Werley, and Adilson E Motter. Stochastic model for power grid dynamics. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, pages 113–113. IEEE, 2007.

[6] Qin Ba and Ketan Savla. A dynamic programming approach to optimal load shedding control of cascading failure in DC power networks. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3648–3653. IEEE, 2016.

[7] Ross Baldick, Badrul Chowdhury, Ian Dobson, Zhaoyang Dong, Bei Gou, David Hawkins, Henry Huang, Manho Joung, Daniel Kirschen, Fangxing Li, et al. Initial review of methods for cascading failure analysis in electric power transmission systems IEEE PES CAMS task force on understanding, prediction, mitigation and restoration of cascading failures. In *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–8. IEEE, 2008.

[8] Arthur R Bergen. *Power systems analysis*. Pearson Education India, 2009.

[9] Andrey Bernstein, Daniel Bienstock, David Hay, Meric Uzunoglu, and Gil Zussman. Power grid vulnerability to geographically correlated failures—analysis and control implications. In *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, pages 2634–2642. IEEE, 2014.

[10] Rajendra Bhatia. *Matrix Analysis*. Springer Verlag, New York, 1997.

[11] Daniel Bienstock. Optimal control of cascading power grid failures. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 2166–2173, Dec 2011.

[12] Daniel Bienstock and Sara Mattia. Using mixed-integer programming to solve power grid blackout problems. *Discrete Optimization*, 4(1):115 – 141, 2007. ISSN 1572-5286. doi: http://doi.org/10.1016/j.disopt.2006.10.007. Mixed Integer Programming IMA Special Workshop on Mixed-Integer Programming.

[13] Daniel Bienstock and Abhinav Verma. The $n - k$ problem in power grids: New models, formulations, and numerical experiments. *SIAM Journal on Optimization*, 20(5):2352–2380, 2010. doi: 10.1137/08073562X.

[14] Norman Biggs. *Algebraic Graph Theory*. Cambridge University Press, 2nd edition, 1993.

[15] Hedi Bouattour, John W Simpson-Porco, Florian Dörfler, and Francesco Bullo. Further results on distributed secondary control in microgrids. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 1514–1519. IEEE, 2013.

[16] Charles D Brummitt, Raissa M D'Souza, and Elizabeth A Leicht. Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences*, 109(12):E680–E689, 2012.

[17] Duncan S Callaway and Ian A Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, 2011.

[18] Benjamin A Carreras, Vickie E Lynch, Ian Dobson, and David E Newman. Critical points and transitions in an electric power transmission model for cascading failure blackouts. *Chaos: An interdisciplinary journal of nonlinear science*, 12(4):985–994, 2002.

[19] Seth Chaiken. A combinatorial proof of the all minors matrix tree theorem. *SIAM Journal on Algebraic Discrete Methods*, 3(3):319–329, 1982.

[20] Fan Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

[21] Tommaso Coletta and Philippe Jacquod. Performance measures in electric power networks under line contingencies. *arXiv preprint arXiv:1711.10348*, 2017.

[22] Thomas H Cormen, Clifford Stein, Ronald L Rivest, and Charles E Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001. ISBN 0070131511.

[23] Paolo Crucitti, Vito Latora, and Massimo Marchiori. A topological analysis of the Italian electric power grid. *Physica A: Statistical mechanics and its applications*, 338(1-2):92–97, 2004.

[24] Nair Maria Maia de Abreu. Old and new results on algebraic connectivity of graphs. *Linear Algebra and its Applications*, 423(1):53 – 73, 2007. ISSN 0024-3795. doi: http://dx.doi.org/10.1016/j.laa.2006.08.017.

[25] Eoin Devane, Andreas Kasis, Marina Antoniou, and Ioannis Lestas. Primary frequency regulation with load-side participation Part II: beyond passivity approaches. *IEEE Transactions on Power Systems*, PP(99):1–1, 2016. ISSN 0885-8950. doi: 10.1109/TPWRS.2016.2636284.

[26] Ian Dobson, Benjamin A Carreras, and David E Newman. A loading-dependent model of probabilistic cascading failure. *Probab. Eng. Inf. Sci.*, 19(1):15–32, January 2005. ISSN 0269-9648. doi: 10.1017/S0269964805050023.

[27] Ian Dobson, Benjamin A Carreras, David E Newman, and José M Reynolds-Barredo. Obtaining statistics of cascading line outages spreading in an electric transmission network from standard utility data. *IEEE Transactions on Power Systems*, 31(6):4831–4841, 2016.

[28] Florian Dörfler and Francesco Bullo. Spectral analysis of synchronization in a lossless structure-preserving power network model. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 179–184, Oct 2010. doi: 10.1109/SMARTGRID.2010.5622040.

[29] Florian Dörfler and Francesco Bullo. Kron reduction of graphs with applications to electrical networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1):150–163, Jan 2013. ISSN 1549-8328. doi: 10.1109/TCSI.2012.2215780.

[30] Florian Dörfler and Sergio Grammatico. Gather-and-broadcast frequency control in power systems. *Automatica*, 79:296–305, 2017.

[31] Herbert Federer. *Geometric measure theory*. Grundlehren der mathematischen Wissenschaften. Springer, 1969.

[32] Robert Grone, Russell Merris, and V S Sunder. The Laplacian spectrum of a graph. *SIAM Journal on Matrix Analysis and Applications*, 11(2):218–238, 1990. doi: 10.1137/0611016.

[33] Frank Harary. Graph theory. *Addison-Wesley*, 1969.

[34] Paul Hines and Sarosh Talukdar. Controlling cascading failures with cooperative autonomous agents. *International journal of critical infrastructures*, 3(1): 192, 2007.

[35] Paul Hines, Ian Dobson, and Pooya Rezaei. Cascading power outages propagate locally in an influence graph that is not the actual grid topology. *IEEE Transactions on Power Systems*, 32(2):958–967, 2017.

[36] Andreas Kasis, Eoin Devane, and Ioannis Lestas. Stability and optimality of distributed schemes for secondary frequency regulation in power networks. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3294–3299, Dec 2016. doi: 10.1109/CDC.2016.7798764.

[37] Andreas Kasis, Eoin Devane, Chrysovalantis Spanias, and Ioannis Lestas. Primary frequency regulation with load-side participation Part I: stability and optimality. *IEEE Transactions on Power Systems*, PP(99):1–1, 2016. ISSN 0885-8950. doi: 10.1109/TPWRS.2016.2636286.

[38] Bredan J Kirby. Spinning reserve from responsive loads. *Report of Oak Ridge National Laboratory*, 2003.

[39] Zhenning Kong and Edmund M Yeh. Resilience to degree-dependent and cascading node failures in random geometric networks. *IEEE Transactions on Information Theory*, 56(11):5533–5546, 2010.

[40] Prabha Kundur, Neal J Balu, and Mark G Lauby. *Power system stability and control*, volume 7. McGraw-hill New York, 1994.

[41] Chengdi Lai and Steven H Low. The redistribution of power flow in cascading failures. In *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1037–1044, Oct 2013. doi: 10.1109/Allerton.2013.6736639.

[42] Enrique Mallada. iDroop: A dynamic droop controller to decouple power grid's steady-state and dynamic performance. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 4957–4964. IEEE, 2016.

[43] Enrique Mallada, Changhong Zhao, and Steven H Low. Optimal load-side control for frequency regulation in smart grids. In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, pages 731–738. IEEE, 2014.

[44] Enrique Mallada, Changhong Zhao, and Steven H Low. Optimal load-side control for frequency regulation in smart grids. *IEEE Transactions on Automatic Control*, 62(12):6294–6309, 12 2017. doi: 10.1109/TAC.2017.2713529. URL `https://mallada.ece.jhu.edu/pubs/2017-TAC-MZL.pdf`.

[45] Dorian Mazauric, Saleh Soltan, and Gil Zussman. Computational analysis of cascading failures in power networks. In *Proceedings of the ACM SIGMETRICS/International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '13, pages 337–338, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1900-3. doi: 10.1145/2465529.2465752.

[46] Karl Menger. Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10 (1):96–115, 1927.

[47] Angel Molina-Garcia, François Bouffard, and Daniel S Kirschen. Decentralized demand-side contribution to primary frequency control. *IEEE Transactions on Power Systems*, 26(1):411–419, 2011.

[48] Daniel K Molzahn, Florian Dörfler, Henrik Sandberg, Steven H Low, Sambuddha Chakrabarti, Ross Baldick, and Javad Lavaei. A survey of distributed optimization and control algorithms for electric power systems. *IEEE Transactions on Smart Grid*, PP(99):1–1, 2017. ISSN 1949-3053. doi: 10.1109/TSG.2017.2720471.

[49] Dusko P Nedic, Ian Dobson, Daniel S Kirschen, Benjamin A Carreras, and Vickie E Lynch. Criticality in a cascading failure blackout model. *International Journal of Electrical Power & Energy Systems*, 28(9):627–633, 2006.

[50] Yutaka Ota, Haruhito Taniguchi, Tatsuhito Nakajima, Kithsiri M Liyanage, and Akihiko Yokoyama. An autonomous distributed vehicle-to-grid control of grid-connected electric vehicle. In *2009 International Conference on Industrial and Information Systems (ICIIS)*, pages 414–418, Dec 2009. doi: 10.1109/ICIINFS.2009.5429826.

[51] Fernando Paganini and Enrique Mallada. Global performance metrics for synchronization of heterogeneously rated power systems: The role of machine models and inertia. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 324–331, Oct 2017. doi: 10.1109/ALLERTON.2017.8262755.

[52] John Pang, Linqi Guo, and Steven Low. Optimal load control for frequency regulation under limited control coverage. In *10th Bulk Power Systems Dynamics and Control Symposium*, 2017.

[53] Richard Pates and Enrique Mallada. Decentralized robust inverter-based control in power systems. *arXiv preprint arXiv:1612.05812*, 2016.

[54] Mohammad Pirani, John W Simpson-Porco, and Baris Fidan. System-theoretic performance metrics for low-inertia stability of power networks. *arXiv preprint arXiv:1703.02646*, 2017.

[55] PNNL. Grid friendly controller helps balance energy supply and demand. URL `http://readthis.pnl.gov/MarketSource/ReadThis/B3099_not_print_quality.pdf`.

[56] Bala Kameshwar Poolla, Saverio Bolognani, and Florian Dörfler. Optimal placement of virtual inertia in power grids. *IEEE Transactions on Automatic Control*, 62(12):6209–6220, Dec 2017. ISSN 0018-9286. doi: 10.1109/TAC.2017.2703302.

[57] Mahshid Rahnamay-Naeini, Zhuoyao Wang, Nasir Ghani, Andrea Mammoli, and Majeed M Hayat. Stochastic analysis of cascading-failure dynamics in power grids. *IEEE Transactions on Power Systems*, 29(4):1767–1779, 2014.

[58] Mario A Rios, Daniel S Kirschen, Dilan Jayaweera, Dusko P Nedic, and Ron N Allan. Value of security: modeling time-dependent phenomena and weather conditions. *IEEE Transactions on Power Systems*, 17(3):543–548, 2002.

[59] M E Ropp, Miroslav Begovic, and A Rohatgi. Analysis and performance assessment of the active frequency drift method of islanding prevention. *IEEE Transactions on Energy conversion*, 14(3), Sep 1999.

[60] Halsey Royden and Patrick Fitzpatrick. *Real Analysis*. Prentice Hall, 2010. ISBN 9780131437470.

[61] Joe A Short, David G Infield, and Leon L Freris. Stabilization of grid frequency through dynamic demand control. *IEEE Transactions on power systems*, 22 (3):1284–1293, 2007.

[62] Saleh Soltan, Dorian Mazauric, and Gil Zussman. Analysis of failures in power grids. *IEEE Transactions on Control of Network Systems*, PP(99):1–1, 2015. ISSN 2325-5870. doi: 10.1109/TCNS.2015.2498464.

[63] Jiajia Song, Eduardo Cotilla-Sanchez, Goodarz Ghanavati, and Paul Hines. Dynamic modeling of cascading failure in power systems. *IEEE Transactions on Power Systems*, 31(3):2085–2095, 2016.

[64] B Stott, O Alsac, and FL Alvarado. Analytical and computational improvements in performance-index ranking algorithms for networks. *International Journal of Electrical Power & Energy Systems*, 7(3):154–160, 1985.

[65] Petr Šulc, Scott Backhaus, and Michael Chertkov. Optimal distributed control of reactive power via the alternating direction method of multipliers. *IEEE Transactions on Energy Conversion*, 29(4):968–977, Dec 2014. ISSN 0885-8969. doi: 10.1109/TEC.2014.2363196.

[66] Rober E Tarjan. A note on finding the bridges of a graph. *Information Processing Letters*, 2(6):160–161, 1974.

[67] Robert E Tarjan and Uzi Vishkin. An efficient parallel biconnectivity algorithm. *SIAM Journal on Computing*, 14(4):862–874, 1985.

[68] Emma Tegling, Bassam Bamieh, and Dennice F Gayme. The price of synchrony: Evaluating the resistive losses in synchronizing power networks. *IEEE Transactions on Control of Network Systems*, 2(3):254–266, Sept 2015. ISSN 2325-5870. doi: 10.1109/TCNS.2015.2399193.

[69] Andreia Sofia Teixeira, Pedro T Monteiro, João A Carriço, Mário Ramirez, and Alexandre P Francisco. Spanning edge betweenness. In *Workshop on mining and learning with graphs*, volume 24, pages 27–31, 2013.

[70] Danial Trudnowski, Matt Donnelly, and Eric Lightner. Power-system frequency and stability control using decentralized intelligent loads. In *2005/2006 IEEE/PES Transmission and Distribution Conference and Exhibition*, pages 1453–1459, May 2006. doi: 10.1109/TDC.2006.1668732.

[71] Konstantin S Turitsyn and P A Kaplunovich. Fast algorithm for $n-2$ contingency problem. In *2013 46th Hawaii International Conference on System Sciences*, pages 2161–2166. IEEE, 2013.

[72] Tyrone L Vincent, Kameswhar Poolla, Salman Mohagheghi, and Eilyan Bitar. Stability guarantees for primary frequency control with randomized flexible loads. In *2016 American Control Conference (ACC)*, pages 2328–2333, July 2016. doi: 10.1109/ACC.2016.7525265.

[73] Zhaojian Wang, Feng Liu, Laijun Chen, and Shengwei Mei. Distributed economic automatic generation control: A game theoretic perspective. In *2015 IEEE Power Energy Society General Meeting*, pages 1–5, July 2015.

[74] Zhifang Wang, Anna Scaglione, and Robert J Thomas. A Markov-transition model for cascading failures in power grids. In *2012 45th Hawaii International Conference on System Sciences*, pages 2115–2124. IEEE, 2012.

[75] Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, NJ, 1996.

[76] Allen J Wood, Bruce F Wollenberg, and Gerald B Sheblé. *Power Generation, Operation, and Control*. Wiley-Interscience, 2013.

[77] Zhao Xu, Jacob Ostergaard, and Mikael Togeby. Demand as frequency controlled reserve. *IEEE Transactions on Power Systems*, 26(3):1062–1071, Aug 2011. ISSN 0885-8950. doi: 10.1109/TPWRS.2010.2080293.

[78] Jun Yan, Yufei Tang, Haibo He, and Yan Sun. Cascading failure analysis with DC power flow model and transient stability analysis. *IEEE Transaction on Power System*, 30(1):285–297, 2015.

[79] Hongming Yang, C Y Chung, and Junhua Zhao. Application of plug-in electric vehicles to frequency regulation based on distributed signal acquisition via limited communication. *IEEE Transactions on Power Systems*, 28(2):1017–1026, May 2013. ISSN 0885-8950.

[80] Changhong Zhao, Ufuk Topcu, and Steven H Low. Frequency-based load control in power systems. In *American Control Conference (ACC), 2012*, pages 4423–4430. IEEE, 2012.

[81] Changhong Zhao, Ufuk Topcu, and Steven H Low. Swing dynamics as primal-dual algorithm for optimal load control. In *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, pages 570–575. IEEE, 2012.

[82] Changhong Zhao, Ufuk Topcu, Na Li, and Steven H Low. Design and stability of load-side primary frequency control in power systems. *Automatic Control, IEEE Transactions on*, 59(5):1177–1189, 2014.

[83] Changhong Zhao, Enrique Mallada, Steven H Low, and Janusz Bialek. A unified framework for frequency control and congestion management. In *Power Systems Computation Conference (PSCC), 2016*, pages 1–7. IEEE, 2016.

[84] Changhong Zhao, Enrique Mallada, Steven H Low, and Janusz Bialek. Distributed plug-and-play optimal generator and load control for power system frequency regulation. *International Journal of Electrical Power & Energy Systems*, 101:1–12, 2018.

[85] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems*, 26(1):12–19, 2011.