On Control and Optimization of Cascading Phenomena in a Class of Dynamic Networks

by

Ali Faghih

B.S., Electrical Engineering (2009),
University of Maryland, College Park
S.M., Electrical Engineering and Computer Science (2011),
Massachusetts Institute of Technology

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Signature redacted

Author .....
Department of Electrical Engineering and Computer Science
May 19, 2015

Signature redacted

Certified by ....
Munther A. Dahleh
William A. Coolidge Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Signature redacted

Accepted by ....
Professor Leslie A. Kolodziejski
Chair, Department Committee on Graduate Students

On Control and Optimization of Cascading Phenomena in a Class of Dynamic Networks

by

Ali Faghih

Submitted to the Department of Electrical Engineering and Computer Science
on May 19, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The primary goal of this thesis is to study transmission line reactance tweaking, as a mechanism for both post-disturbance control and pre-disturbance resilience enhancement in a transmission network, and develop an optimization framework for evaluating the efficacy of this mechanism in both scenarios. We start by developing a mixed-integer linear programming (MILP) formulation for tracking the redistribution of direct current (DC) flows and the graph-theoretic evolution of network topology over the course of cascading failures. Next, we propose a min-max setup for studying the impact of post-disturbance reactance tweaking on the resilience of the system to a worst-case N-k disturbance and devise a MILP reformulation scheme for the underlying bilevel nonconvex mixed-integer nonlinear program (MINLP) to facilitate the computation of its optimal solution. We then develop a MILP framework for computing the exact value of a tight upper bound on the efficacy of post-disturbance reactance tweaking among the set of all possible N-k disturbances for a given k and a given bus load scenario. Our numerical case study suggests that post-disturbance reactance tweaking, even on only a small number of lines, can considerably reduce the amount of load shed in some scenarios in the tested system. As for pre-disturbance resilience enhancement, we develop a MILP reformulation for approximating the bilevel MINLP that seeks to assess the efficacy of pre-disturbance reactance tweaking in reducing the number of lines that will fail over the propagation of cascading failures in the event of a worst-case-scenario N-k disturbance. We also give a MILP framework for computing an approximate upper bound on the efficacy of this mechanism among the set of all N-k contingencies for a given k. Our numerical case study suggests that pre-disturbance reactance tweaking on a few transmission lines can, in some cases, prevent the failure of multiple transmission lines over the course of cascading failures in the tested system.

Thesis Supervisor: Munther A. Dahleh
Title: William A. Coolidge Professor of Electrical Engineering and Computer Science

TO MY BELOVED PARENTS AND MY DEAREST SISTER,
AND TO ALL THOSE WHO HAVE TOUCHED MY SOUL WITH THE KINDNESS OF THEIR HEARTS...

Acknowledgments

I would like to wholeheartedly thank my thesis supervisor, Professor Munther Dahleh, for being a constant source of inspiration for me, and for his unwavering support, patience, invaluable advice, and generosity in sharing his deep knowledge. I am immensely grateful to Munzer for selflessly putting my growth as the first priority, giving me the freedom to explore the field, training me to become an independent thinker, and constantly encouraging me to dig deeper into my ideas. The impact he has had on my growth both as an academic and as an individual is way more profound than I can describe in words. Having Munzer as my role model for the past six years (and many more years to come) has been a life-changing experience for me. His brilliant critical-thinking skills and his ingenious approach to tackling difficult problems have provided me with an incredibly fulfilling learning experience; I constantly gained new perspectives into "the art of problem solving" as a result of interacting with Munzer. I consider myself extremely fortunate to have had the honor, and the privilege, of spending the past six years of my life under the tutelage of such an amazing, kindhearted, well-rounded, brilliant individual. I am forever indebted to Munzer for all he has done for me and for his tremendous impact on my life.

I would like to express my gratitude to my committee members, Professor George Verghese and Professor Emilio Frazzoli, for insightful conversations and wonderful advice. I would like to thank George for selflessly spending a lot of his precious time providing me with wonderful feedback regarding my model and my results. I have learned so much from interacting with George and his perspectives helped me gain a lot of valuable insight into the problem presented in this thesis and gave me new angles to evaluate my research problem from. I would like to thank Emilio for all the interesting discussions about the optimization-theoretic aspects of my model that gave me new perspectives into the problem.

I would like to thank Professor Luca Daniel for his wonderful support and advice in various stages of my graduate studies. In particular, I would like to express my gratitude to him for giving me the

wonderful opportunity to be his teaching assistant as I learned a lot from his teaching philosophy and his amazing relationship with the students. Luca inspired me with his fantastic teaching skills and I am indebted to him for all that I have learned from him. I would like to also thank Dr. Mardavij Roozbehani for mentoring me during my S.M. studies; I learned a lot from Mardavij and his guidance helped me mature much faster than I would otherwise have.

I am grateful to my academic counselor, Professor Asuman Ozdaglar, for her great advice throughout my time at MIT. I would like to thank Professor Terry Orlando for his support and advice in the early years of my graduate studies.

I would like to thank my undergraduate advisor, Professor Edo Waks, who provided me with great advice and constantly supported me throughout the period that I worked in his lab; working under the tutelage of Edo has had a wonderful impact on my life. I would also like to thank Professor Christopher Davis and Professor Nuno Martins for their advice and support during my undergraduate years.

I would like to also thank Albert and Alina, Munzer's assistants over the past few years, for being very friendly, scheduling everything meticulously, and being amazingly helpful.

I would like to thank my friends and lab-mates in the Laboratory for Information and Decision Systems (LIDS), who have been a major part of my life over the past 6 years at MIT, for their friendship and support. I have made tens of friends at LIDS; to name a few, I would like to thank, in alphabetical order: Amirali, Ammar, Christina, Diego, Donatello, Elaheh, Elie, Erich, Giancarlo, Hajir, Hamza, Kimon, Luis, Marzieh, Mesrob, Mihalis, Mike, Noel, Omer, Qingqing, Saverio, Shreya, Spyros, Yola, and Yongwhan. I would like to especially thank Elie, because aside from being an amazing friend/lab-mate and a great listener, he introduced me to some very cool topics in graph theory and networks, and motivated me to engage in numerous intriguing technical discussions. Outside of LIDS, I would like to thank Sahand, Shahriar, and Tristan for being amazing, loving friends and for keeping me company in various adventures of life.

I would like to express my deepest gratitude to my amazing parents and my wonderful sister for their unconditional love and support, for always believing in me, for constantly inspiring me to pursue my goals, and for always doing their absolute best in providing me with all the tools I need

in my pursuit of happiness. They each are my role models in their own way. My mother has always been an endless source of pure love to me; it is her limitless affection that fuels the warmth of my heart; her selflessness, patience, and incredibly kind heart have always been infinite sources of inspiration to me. My father has always inspired me to set my goals ambitiously and strive for perfection in whatever that I do; his determination in accomplishing his goals has always inspired me to work hard and constantly seek to excel at whatever I put my mind to. My sister, who is also my academic sibling from Munzer's group, has always been my best friend, has always been there for me, and has always inspired me to grow further by constantly being a step ahead of me in everything and providing me with wonderful advice. We have so many great memories together and her presence always warms my heart.

I would like to thank my aunt, Ladan, for her love and for everything she has done for me; I would not have been where I am today without her support. I would also like to thank my grandparents for their love and affection, and for always encouraging me to set the bar high in my academic endeavors. I shall cherish their love forever, regardless of whether they are warming my heart by sending me their love and best wishes, by traveling thousands of miles to attend my commencement ceremony, or by their loving memory.

I would like to also thank the National Science Foundation for supporting me in my first three years of graduate studies through the National Science Foundation Graduate Research Fellowship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over the past two decades, large-scale cascading failures in various power systems have highlighted the compelling need for substantially more powerful mechanisms for increasing power systems' resilience to disturbances. As a result, identifying various strategies and technologies that can potentially increase the resilience of a power system to cascading failures has been a core topic of interest in the power systems literature, with various studies focusing on developing effective analytical and computational tools for efficacy assessment, economic viability, engineering design, and implementation of these mechanisms.

Some resilience enhancement strategies focus on post-disturbance control with the goal of maximizing the amount of demand that can be satisfied after a disturbance while returning the system to reliable operating conditions. On the other hand, some (pre-disturbance) resilience enhancement strategies seek to prevent, or quickly eliminate the impact of, any contingencies that could otherwise compromise system reliability (e.g. by leading to cascading failures or causing issues with system frequency, etc.). In both post-disturbance control and pre-disturbance resilience enhancement, much attention has been given to the transmission system due to its significant impact on the propagation and control of cascading failures.

As far as post-disturbance control is concerned, transmission line switching, intentional islanding,

and controlled load-shedding have been studied extensively in the existing literature (e.g. see [64], [59], and [3] and the references therein). As for pre-disturbance resilience enhancement, problems such as transmission expansion planning (see e.g. [24]) and security-constrained economic dispatch (see e.g. [39]) have been the main focus of a large body of literature in the field of power systems. The existing literature suggest that these well-studied mechanisms can be quite effective. However, as the demand for electricity grows and uncertainty in the system increases, there seems to be a need for technologies that can allow the system to remain resilient even at higher loading levels and in the face of a broader array of contingencies. As power systems move towards deregulation, lowering costs motivates investment in technologies that allow more flexibility in the system; one set of technologies and devices that can provide flexibility in the system are the Flexible Alternating Current Transmission Systems (FACTS) (see e.g. [44]). According to [58], FACTS technologies can benefit power systems in various ways such as improving system stability, allowing improved voltage control, and increasing the loadability of the transmission network by creating the possibility to bring lines considerably closer to their thermal capacity thresholds.

In this thesis, we study "reactance tweaking" as a mechanism for both post-disturbance control and pre-disturbance resilience enhancement in a transmission network. The core idea behind post-disturbance reactance tweaking is simple: after the redistribution of power flows due to a disturbance, some lines become overloaded while some other lines have some residual capacity remaining; optimal tweaking of line reactances could reduce the flow in overloaded lines and direct it to those lines that have some unused capacity. Using the direct current (DC) approximation to alternating current (AC) power flows, we develop a framework for assessing the efficacy of this mechanism and our numerical case study suggests that changing the reactances of only a small number of transmission lines (within a reasonable range) to control the flow of power in the transmission network after a disturbance can, in some scenarios, considerably reduce the amount of load that needs to be shed in order to bring the system back to reliable operating conditions in the tested system (post-disturbance reactance tweaking is studied in Chapter 3). In particular, we first study the efficacy of post-disturbance reactance tweaking in response to a strategic disturbance that seeks to cause the wort-case-scenario load-shedding. We then develop a framework for computing

an exact upper bound on the efficacy of post-disturbance reactance tweaking in response to any $N-k$ contingency for a given k and a given load level (by an "$N-k$ contingency" we mean a disturbance that is initiated by the failure of up to k out of N transmission lines). Although the economic viability, engineering design, and implementation of a post-disturbance reactance-tweaking technology are not discussed in this thesis, our results can potentially motivate researchers to dig deeper into this idea and cover various aspects of such technology. This research is in line with the recent surge in both academic research and industrial investments related to the incorporation of power electronic technologies such as FACTS controllers into the transmission grid for more sustainable and reliable operation of the system (e.g. see [33] for a relevant discussion on FACTS).

The FACTS device most relevant to the study in this thesis appears to be the Thyristor-Controlled Series Compensator (TCSC). According to [44], TCSC provides flexible (and smooth) control of the impedance of a transmission line with considerably quicker response relative to traditional control technologies, and the utilization of this device in various applications ranging from enhancement of transient stability, damping inter-area oscillations, and preventing voltage collapse, to mitigating sub-synchronous resonance and improving system reliability has been studied in the literature (e.g. see [51], [65], [61], [17], [13]). The impact of TCSC on congestion and spot pricing has also been studied in [1], where the results suggest that TCSC could be useful in mitigating both congestion and losses. However, the application of TCSC that is most relevant to the work in this thesis is its use for mitigating transmission line overloads. This application of TCSC appears to have received less attention in the literature than the above-mentioned applications; nevertheless, multiple studies such as [44], [56], [54], [35], [58] have studied the value of TCSC in reducing line overloads from various aspects, mostly for the case of an $N-1$ contingency and particularly for the application of security-constrained economic dispatch. In the event of a single-component failure, these papers have suggested that TCSC can effectively enhance the system's security. Our approach to studying the impact of pre-disturbance reactance tweaking (reported in Chapter 4) is fundamentally different from the studies mentioned above as we shall give a bilevel (max-min) approach for studying whether reactance tweaking can reduce the number of lines that will fail over the propagation of cascading failures (i.e. after the initial disturbance) in the event of a worst-

case-scenario $N-k$ disturbance (for a given k), and also give a mixed-integer linear programming (MILP) framework for computing an approximate upper bound on the efficacy of this mechanism in response to the set of all $N-k$ contingencies for a given k and a given set of bus injections.

Models for resilience assessment of power systems have been proposed in various studies in the existing literature. Many of these models are in the form of bilevel (e.g. min-max) and trilevel (e.g. min-max-min) optimization problems (see, e.g. [15], [55], [64], [2], [12]) in which an adversarial agent and a system operator (SO) optimize their decisions strategically. In many such scenarios, a zero-sum Stackelberg game arises in which one of the two agents (i.e. either the system operator or the adversary) goes first and the other agent then goes next, and this sequential process can go for a number of steps, while both agents share the same objective function (i.e. one agent is minimizing the objective function while the opposing agent is maximizing it). Resilience is then typically measured by the severity of the worst-case disturbance despite the SO's efforts. Such problems are often referred to as "network interdiction" problems. For instance, in the version of the "$N-k$" problem that we shall use to find the worst-case "$N-k$" disturbance in the presence of post-disturbance reactance tweaking, we consider the case in which the adversary goes first by removing up to k lines, and the system operator goes next and responds to the disturbance by tweaking the line reactances and re-dispatching the system with the goal of maximizing the demand satisfied post-disturbance. Note that this Stackelberg game is zero-sum because both agents have the same objective function: the adversary seeks to minimize the demand satisfied post-disturbance by removing lines and the system operator seeks to maximize the demand satisfied post-disturbance by tweaking the line reactances and also adjusting the loads.

The literature on bilevel network interdiction problems with purely linear and continuous inner-level programs is extensive (by "inner-level program" we mean, for instance, the "max" part of the problem in a min-max setup). When the inner-level problem is bounded and feasible, a common technique is to reformulate the bilevel program into a one-level MILP by invoking the strong duality of the inner-level linear program (LP) and linearizing the integer$\times$continuous (or integer$\times$integer) bilinear terms (e.g. see [47]). Similar techniques have been used in [40] for

the case of a stochastic network interdiction problem as well. Approximate methods have also been employed in various works such as [12], in which a greedy-based algorithm is developed to lower the computational complexity. More sophisticated studies on network interdiction have employed graph-theoretic methods combined with nonlinear optimization methods to give more realistic models, especially using AC power flows ([52], [53], [28]). Particularly, [52] uses these techniques to give an approximation method for solving interdiction problems in large networks.

On the other hand, the literature on network interdiction problems with mixed-integer inner-level optimization problems is more recent and not as extensive. In terms of optimization theory, generally speaking, both analytical characterization and computational complexity of problems with mixed-integer inner-level programs are considerably more challenging because strong duality no longer holds and the integer inner-level variables add a new layer of non-convexity and NP-hardness to the problem. A majority of the previous work on interdiction problems with mixed-integer inner-level optimization has dealt with "transmission line switching"; some examples of works in this line of literature include [3], [26], and [64] in which different techniques varying from locally optimal heuristics (such as the genetic algorithm and multi-start Benders decomposition methods) to global optimization techniques (based on column-and-row generation which has been presented in [64]; this column-and-row generation framework shall be used extensively in this thesis as well) have been employed for solving the underlying optimization problem.

Aside from the above-mentioned static bilevel models, some other models consider the dynamic nature of cascading failures (see, e.g. [19], [18], [20], and [9]). Among these, the work in [9] is more focused on optimization theory and develops mixed-integer programming techniques for optimal transmission capacity expansion investment (i.e. they find minimal amounts of increases needed in line capacities to prevent cascading failures, using both static and multi-stage models). However, their model is quite different from ours and also, unlike the work in this thesis, they assume constant line reactances.

Aside from the literature on resilience assessment of power systems, the literature on, or related to, the modeling and characterization of cascading phenomena in networks is also extensive. Many

works have studied various models of cascades and contagion in general complex networks (see, for instance, [22], [50], [60], [45], [25], [46], [4], [27]), which may be applied to electricity networks to only abstractly model the propagation of cascading failures. Most of these works assume that a line or node failure in a network causes the failure of neighboring lines or nodes, often with some probabilistic rule (or in some cases with a deterministic local contagion rule, such as linear threshold models as in e.g. [42]). As discussed in [7], although such models of epidemics in complex networks pave the way for employing percolation-based tools to study the effects of cascading failures, they do not provide us with powerful tools that can be applied to actual large-scale cascades in real-world power systems. This is because the failure of a certain line can considerably impact a remote line in a power system and the cascading failures in electricity networks do not necessarily propagate in a neighborhood-limited manner. As shown in [7], if we use the linearized (DC) flow model to keep track of the evolution of cascading failures, we obtain failure spread characteristics that are considerably different from the dynamics described by the above-mentioned epidemic-based models. With that in mind, we are particularly interested in studying the evolution of dynamic DC flows in a capacity-constrained electricity network in the event of cascading failures. In other words, we use a DC-flow approximation of the AC flows in a power system.

Numerous works use DC or AC power flow equations to model and study various aspects of cascading failures in power systems (see e.g. [52], [28], [23], [11], [19]). A core focus of many of the studies in this stream of literature has been on identifying a small group of lines whose failure results in a power system that can no longer satisfy a prescribed minimum amount of demand (see e.g. [10], [23], [2]). To simplify numerical computations, many such studies use DC flow equations. However, some works such as [59] and [52] use the more realistic (but also more computationally challenging) AC power flows. Donde et al. [28] give a mixed-integer nonlinear program (MINLP) for identifying the largest blackout that can result from a certain number of line failures or to identify the smallest number of lines whose failure can cause a blackout with a specified severity, while including reactive power in their formulation. It is of note that similar power system contingency problems also arise in the context of transmission and generation expansion

planning and network design (see e.g. [24], [48], [49], [14], [21] and [41]) and sometimes with special attention to the accommodation of renewable energy integration (see e.g., [5]). Moreover, some works such as [55] and [30] consider "$N - k$" contingencies in optimal power flow models and some do the same in unit commitment problems [57].

Generally speaking, in the literature on resilience assessment of power systems to cascading failures, optimization-based static models are relatively well studied, but optimization-based multi-stage models (like the one in [9]) are scarce. This is somewhat justifiable because the SO quickly interferes with the cascading failures to mitigate their impact on the system. However, there could be enough time elapsing between the cascade's inception and the SO's response to allow the cascade to propagate for some time before the control mechanisms kick in. This could considerably change the nature of the problem and hence, multi-stage models generally appear to be more realistic. Therefore, other than the idea of reactance tweaking, one of the key features of this thesis is the development of a MILP framework for tracking the dynamic evolution of the network topology and redistribution of DC flows over the course of cascading failures (in Chapter 2).

1.1 Contributions

The contributions of this thesis can be summarized as follows:

(1) In Chapter 2, we develop a MILP model for tracking the evolution of network topology and redistribution of DC flows as cascading failures propagate over multiple stages.

(2) In Chapters 3 and 4, we build optimization frameworks for studying the two-agent (adversary vs. system operator) games corresponding to efficacy assessment of reactance tweaking as a post-disturbance control mechanism and a pre-disturbance resilience enhancement strategy, respectively. We also develop frameworks for studying how the efficacy of these mechanisms are upper bounded in response to the set of all $N - k$ contingencies for a given k. We present numerical experiments on a test case using our models, and our simulation results suggest that

both post-disturbance and pre-disturbance reactance tweaking could be quite effective in some scenarios in the tested system.

(3) We give a rigorous framework for deriving a MILP reformulation of the non-convex bilevel mixed-integer program underlying the efficacy assessment problems for post-disturbance reactance tweaking, which involve non-convexities stemming from both binary$\times$continuous and continuous $\times$ continuous bilinear terms in the inner-level problem. The continuous$\times$continuous bilinear terms in the post-disturbance reactance tweaking problem pose a significant computational burden for computing the globally optimal solution. However, we manage to develop exact reformulation techniques to solve this problem as a MILP by shifting the non-convexity from the continuous$\times$continuous bilinear terms to auxiliary binary variables that are introduced into the problem as linear terms, which allows us to then tailor existing MILP-based decomposition techniques (namely, column and row generation, see e.g. [62]) to our MILP reformulation and solve the underlying bilevel non-convex MINLP for post-disturbance reactance-tweaking to global optimality. We also derive a single-level MILP for computing an exact upper bound on the efficacy of post-disturbance reactance-tweaking, in response to the set of all $N-k$ scenarios for a given k and a given load level. Considering that a large number of variables and constraints could exist in the underlying MINLP for a real-world large power system, our MILP reformulation of the inner-level MINLP also allows for using the existing powerful MILP solvers which are "mature" in the sense that they are fast, robust, and capable of handling large problems with millions of variables (e.g. see [32] for a discussion of the advantages of MILP). As for pre-disturbance resilience enhancement, we develop an approximate MILP reformulation for solving the underlying bilevel MINLP, which makes it amenable to the above-mentioned column-and-row-generation scheme and allows us to solve the approximate problem to optimality. We also give a MILP formulation for approximating the upper bound (among the set of all N-k contingencies for a given k) on the number of lines that would otherwise have failed over the course of a cascading-failures event had we not tweaked the line reactances pre-disturbance.

Remark 1. *Please bear in mind that modeling power systems is only one application of DC flow*

networks, and hence, many of the analytical tools developed in this thesis potentially have applications beyond power systems. It is well known that DC power flows are the solution to a constrained quadratic program as shown below (see e.g. [43]):

$$\begin{aligned} \min_{\mathbf{p}} \quad & \tfrac{1}{2}\mathbf{p}^T X \mathbf{p} \\ s.t. \quad & M\mathbf{p} = \mathbf{b} \end{aligned} \tag{1.1}$$

where $\mathbf{p}$ *denotes the vector of power flows,* X *is a diagonal matrix whose diagonal elements are the line resistances,* M *is the node-branch incidence matrix (please see Definition 4 for a formal definition), and* $\mathbf{b}$ *is the vector of bus injections. This means that in any DC flow network, the flows in each line at each instant of time are the optimal solution to a constrained convex quadratic program, whose objective function is quadratic in the flows, subject to nodal flow conservation constraints; in the case of electric DC flows, this quadratic function quantifies the energy consumed by the flow network (see e.g. [43]). The dual variables (Lagrangian multipliers) corresponding to the flow conservation constraints are referred to as "voltages" in electricity networks. The well-known Kirchhoff's Voltage Law equations in electricity networks appear in the set of equations corresponding to the necessary and sufficient optimality conditions for solving the above-mentioned quadratic program. For a fixed reference voltage, the conservation of current and Kirchhoff's Voltage Law give the necessary and sufficient conditions for finding the optimal solution to this constrained quadratic program (please see [43] for further details). Therefore, any network that optimizes a similar positive-definite quadratic objective function subject to conservation of flow constraints can be thought of as a DC flow network and may be able to benefit from the tools developed in this thesis.*

1.2 Motivating Example

Given the non-trivial interplay between graph topology, bus injections, line impedances, and line capacities in determining (and controlling) the flow of power in the system, it is rather intuitive that aligning these parameters with each other can increase resilience. As discussed earlier, changing the graph topology via line switching and intentional islanding after a disturbance have been studied extensively in the literature as ways of controlling the graph topology. Load-shedding and re-dispatching are also the strategies used for controlling bus injections. This thesis focuses on controlling line impedances, specifically via tweaking the reactances of the transmission lines. Let us motivate our problem using the following simple DC circuit:

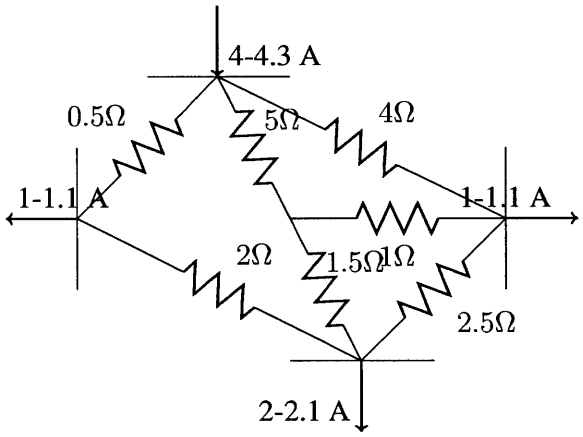


Figure 1-1: Simple DC circuit example prior to disturbance

In this circuit, we assume that the injection at the generation node is expected to vary between $4A - 4.3A$, and for this generation range, all lines will have plenty of residual capacity in them (please see the table below for line flows in the case when the injection is $4A$).

Table 1.1: Pre-Disturbance Flows (4A Generation Case) vs. Capacities

r	cap	|flow|
4Ω	2.1A	0.91A
0.5Ω	3.8A	2.41A
1Ω	2.1A	0.25A
2Ω	4.1A	1.41A
1.5Ω	1.5A	0.43A
2.5Ω	2A	0.16A
5Ω	1A	0.68A

Now if at an instant of time when the generation is for example $4A$, we remove the 4Ω line, then the topology changes and the new line flows are as shown in the table below (note the abundance of residual capacity in the 0.5Ω line):

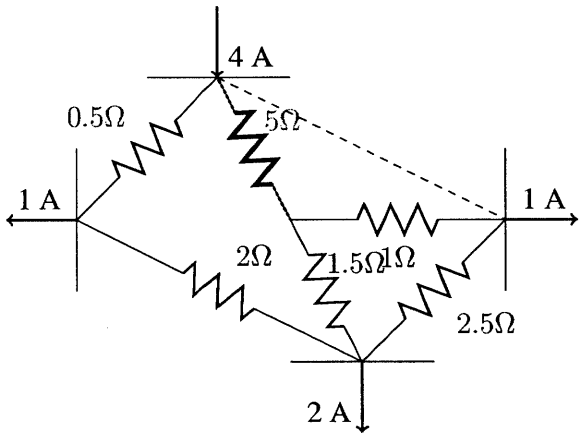


Figure 1-2: A line becomes overloaded post-disturbance in the simple DC circuit example

Table 1.2: Post-Disturbance Flows (4A Generation Case) vs. Capacities

r	cap	|flow|
4Ω	2.1A	$\times$
0.5Ω	3.8A	2.97A
1Ω	2.1A	0.81A
2Ω	4.1A	1.98A
1.5Ω	1.5A	0.22A
2.5Ω	2A	0.19A
5Ω	1A	1.02A

In order to prevent the 5Ω line from tripping due to overload, we now need to shed some of the injection at the generation node in order to bring the flow in the 5Ω line below its flow capacity. However, no load-shedding would have been necessary if we could increase the 5Ω line's resistance by 0.5Ω right after the initial disturbance.

Note that we expect impedance-tweaking to be a robust technology, in the sense that we expect it to be effective for a reasonably broad range of bus injections. For instance, if the injection is more than 4A in the above example (but still within the prescribed range of 4-4.3Ω), further resistance increase (only on the 5Ω line) is needed accordingly to prevent the tripping of the 5Ω line. Note that since the flow on each line of the network depends on the resistances of all lines in the system (as per Kirchhoff's laws), it would also have been possible to tweak the resistances of other lines in the above circuit in order to reduce the current flowing into the 5Ω line, and hence, we expect this tweaking technology to be fairly flexible in terms of the choice of its location in the transmission system. Alternatively, we could also have prevented the overloading of the 5Ω line by pre-disturbance tweaking of its resistance. If, for instance, we were aware that the failure of the 4Ω line would immediately lead to the overloading of the 5Ω line and that the failure of the 4Ω line is imminent, we could tweak the resistance of the 5Ω line prior to the failure of the 4Ω line so as to be immune to any potential contingencies on the 4Ω line. This is why we also study pre-disturbance tweaking of line reactances in Chapter 4.

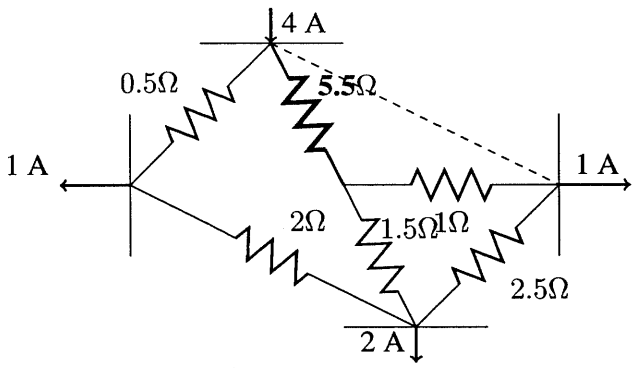


Figure 1-3: Tweaking the resistance of the 5Ω line (via a 0.5Ω increase) could have prevented its overloading in the simple DC circuit

Thus, the above simple example suggests that adjusting impedances of only a small fraction of lines post-disturbance or even pre-disturbance can potentially be effective.

1.3 Preliminaries

In this section we introduce the notation and nomenclature that shall be used jointly among all chapters (other variables and concepts shall be defined as they become introduced within their corresponding chapters).

1.3.1 Notation

- $\mathbb{R}$ The set of real numbers
- $\mathbb{Z}$ The set of integers
- $\mathbb{R}_+$ The set of positive real numbers
- $\mathbb{Z}_+$ The set of positive integers

- $\overline{\mathbb{R}}_+$ The set of non-negative real numbers
- $\overline{\mathbb{Z}}_+$ The set of non-negative integers
- $|\cdot|$ For a finite set X, its cardinality shall be denoted by $|X|$; however, for a vector or matrix X, $|X|$ shall denote the absolute value of X.
- The vector corresponding to any particular parameter is represented in bold-face. The i-th entry of a vector $\mathbf{x}$ is denoted by x_i.
- $\{\mathbf{x}\}$ With a slight abuse of notation, for a vector $\mathbf{x}$, we denote by $\{\mathbf{x}\}$ the set of all entries of $\mathbf{x}$.
- $diag(\mathbf{x})$ A diagonal matrix whose diagonal entries are the elements of a vector $\mathbf{x}$
- For a matrix (or vector) A, its transpose is denoted by A'.
- $\mathbf{1}$ Denotes a vector of all ones .
- $< A > \equiv < B >$ Denotes the equivalence of statements A and B.

1.3.2 Nomenclature

Network topology:

- $\mathcal{B}$ The finite set $\mathcal{B} = \{1, ..., n\}$ of buses (nodes)
- $\mathcal{E}$ The finite set $\mathcal{E} = \{1, ..., N\}$ of lines (edges)
- $(\mathcal{B}, \mathcal{E})$ This pair denotes a directed multigraph
- n The number of nodes (buses)
- N The number of lines. Each line is indexed by an integer in the set $\{1, \ldots N\}$.

- $s(L)$ For a line $L = (c, d, h) \in \mathcal{E}$, extending from bus $c \in \mathcal{B}$ to bus $d \in \mathcal{B}$ and indexed by h, we define $s(L) \stackrel{def}{=} c$ (i.e. the start bus of the line).

- $e(L)$ For a line $L = (c, d, h) \in \mathcal{E}$, extending from bus $c \in \mathcal{B}$ to bus $d \in \mathcal{B}$ and indexed by h, we define $e(L) \stackrel{def}{=} d$ (i.e. the end bus of the line).

- $\mathcal{E}_i$ The set of all lines incident to node i.

Transmission line characteristics:

- $\overline{cap}^a_{ijh}$ The active (operational) flow carrying capacity of line (i, j, h) (i.e. line (i, j, h) is guaranteed to be active as long as the power it carries is below $\overline{cap}^a_{ijh}$).

- $\overline{cap}^f_{ijh}$ The failing capacity threshold of line (i, j, h), i.e. line (i, j, h) is guaranteed to fail as soon as the power it carries reaches (or exceeds) $\overline{cap}^f_{ijh}$.

- $x^{\min}_{ijh}$ The minimum allowed reactance for line (i, j, h) after tweaking

- $x^{\max}_{ijh}$ The maximum allowed reactance for line (i, j, h) after tweaking

- x^o_{ijh} The value of the reactance of line (i, j, h) before tweaking (note that $x^o_{ijh} \in [x^{\min}_{ijh} \; x^{\max}_{ijh}]$ by assumption).

Power flow and cascading failures:

- T The time-horizon of the cascade, i.e. the number stages for which the cascading failures propagate before the system operator's response

- $\mathbf{b}^t$ The vector of net bus injections at time $t \leq T$

- $\mathcal{B}^o_{\pm}$ The set of all buses with non-zero injections at time zero

- θ^t The vector of all bus phase angles at time t

- I^t_{ijh} The DC approximation of the power flow in line (i, j, h) at time t

- z_{ijh}^t Binary variable indicating the failure state of line (i, j, h) at the end of time t ($z_{ijh}^t = 1$ if the line is active by the end of stage t, and $z_{ijh}^t = 0$ otherwise)

- s_{ijh}^t Binary variable that switches from 1 to zero at time t if line (i, j, h) trips at time t (if line (i, j, h) has already failed prior to stage t, then $s_{ijh}^t = 0$).

- $i \overset{t}{\leftrightarrow} j$ Indicates that an active path exists between buses i and j at the end of stage t

- $i \overset{t}{\nleftrightarrow} j$ Indicates no path exists between buses i and j at the end of stage t.

Definition 1. *Consider some bilevel program (e.g.* $\min_{x \in \tilde{X}} \max_{y \in Y} f(x, y)$*). Then, we shall refer to the lower-level decision (i.e.* $\max_{y \in Y} f(x, y)$ *in the above example) as the "inner-level problem" and the higher-level decision variables (x in the above example) as the "outer-level decision variables" and the constraints corresponding to the higher-level feasible set ($\tilde{X}$ in the above example) as the "outer-level constraints". If the problem is three-level (e.g. max-max-min) then we shall refer to the rightmost program (in the max-max-min example it would be the minimization level) as the "innermost-level problem".*

Definition 2. *A* connected network *is any network in which at least one path exists between each pair of nodes.*

Definition 3. *The indicator function* $\mathbb{1}_{\mathcal{X}}(c)$ *is a map from* $\mathbb{R} \to \{0, 1\}$ *such that* $\mathbb{1}_{\mathcal{X}}(c) = 1$ *if* $c \in \mathcal{X}$ *and* $\mathbb{1}_{\mathcal{X}}(c) = 0$ *otherwise.*

Definition 4. *Suppose we index all lines in* $\mathcal{E}$ *from 1 to* $|\mathcal{E}|$*. Then, the* incidence matrix, *M, is a* $|\mathcal{B}| \times |\mathcal{E}|$ *matrix in which every entry* $M_{BL} = 1$ *if* $B = s(L)$ *(i.e. B is the starting node of the line indexed by L) and* $M_{BL} = -1$ *if* $B = e(L)$ *(i.e. B is the ending node of L), and* $M_{BL} = 0$ *otherwise.*

Chapter 2

Modeling the Propagation of Cascading Failures

In this chapter, we shall develop a MILP model for tracking the evolution of network topology and redistribution of DC flows as cascading failures propagate over multiple stages.

We assume that the agent that initiates the cascading failures, whom we shall refer to as the "adversary", does so at time $t = 0$ by simultaneously removing up to k transmission lines from the grid. After that, the adversary takes no further action, and the cascading failures propagate for T discrete stages. We assume that the only source of the propagation of failures is the overflowing of new lines at each stage due to the redistribution of flows resulting from the collapse of the failed transmission lines from previous stages. Although other than the transmission lines one could investigate the failure of other types of components (e.g. generators and transformers) as the components of interest in the initiation and propagation of cascading failures, we shall focus on transmission lines as the only source of failures since they are known to be a major cause of cascading failures. It has been observed that a small number of line failures (e.g., 3-5) can cause severe blackouts, a prominent example of which is the Northeast Blackout of 2003 [52]).

The following simple circuit example illustrates the propagation of cascading failures as per our

model. The first figure below shows the pre-disturbance topology of the example network.

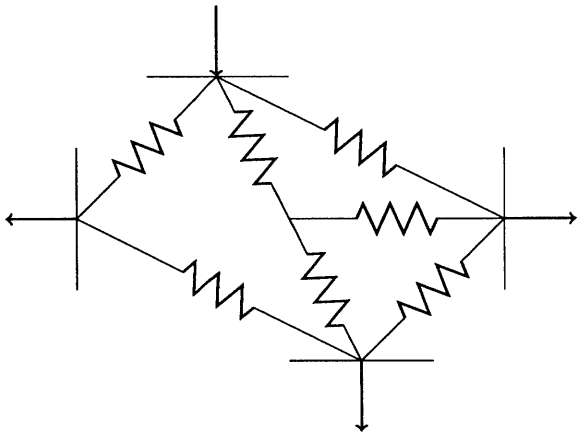


Figure 2-1: Network topology of the example system prior to disturbance

The adversary initiates the cascade by removing, say $k = 1$ line, at time 0, and does nothing afterwards.

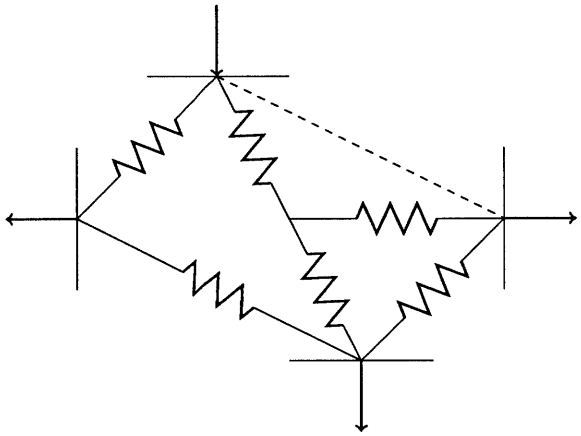


Figure 2-2: The removal of a line causes a disturbance in the example system

At t=1, using Kichhoff's Voltage Law (KVL) and Kirchhoff's Current Law (KCL), flows are computed for the new topology; some lines become overloaded (the overloaded lines are shown by thicker resistors in the figure below), and we set their tripping states to zero.

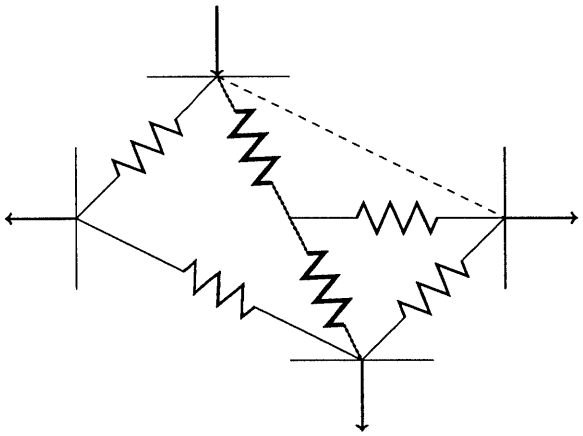


Figure 2-3: Two lines become overloaded in the example system at t=1

Note that at any $t \geq 0$ any disconnected line(s) will be considered failed even though they have not tripped; we set the failure state of these lines to zero. The disconnected line is shown using a dashed oval in the picture below.

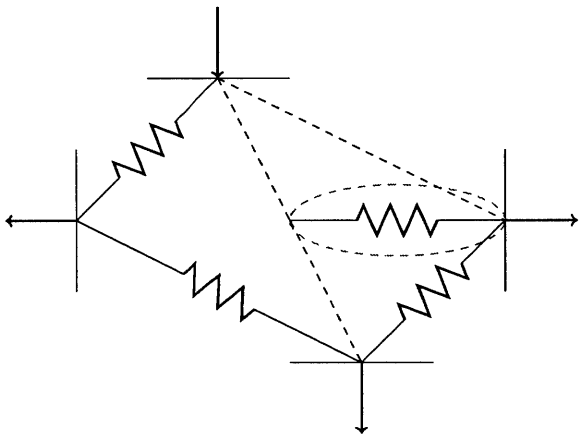


Figure 2-4: A line fails because of becoming disconnected in the example system at t=1

Now if the system operator (SO) has not yet responded to the disturbance and the cascade propagates for another stage, the line that was bridging one of the buses to the generator in the above example will fail.

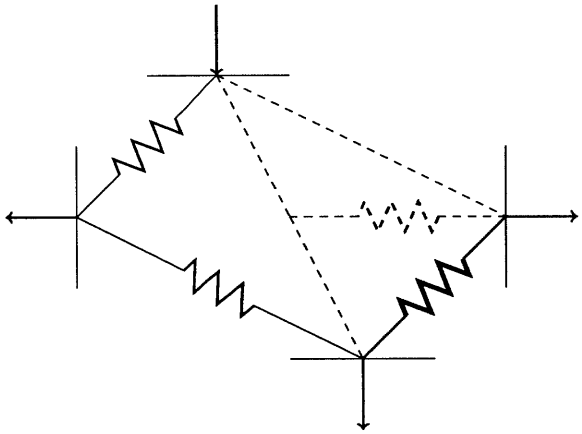


Figure 2-5: The delay in system operator's response allows the cascade to propagate for another stage

The initially connected graph splits into two "islands," and a demand bus becomes islanded away from the generator and loses all its demand.

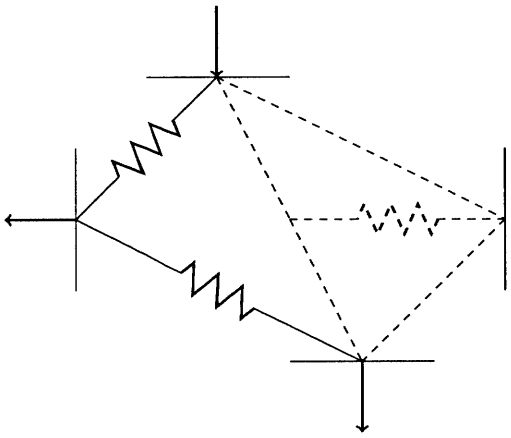


Figure 2-6: The example system splits into two islands

The assumptions made in our discrete-time model of cascading failures are summarized in the next section.

2.1 Model Assumptions

In our multi-stage model of cascading failures, we make the following assumptions:

1. We shall use the DC approximation to AC power flows. The equations for AC flows are (see e.g. [6]):

$$P_{ijh} = V_i^2 \rho_{ijh} - V_i V_j [\rho_{ijh} cos(\theta_i - \theta_j) + \varrho_{ijh} sin(\theta_i - \theta_j)]$$

$$Q_{ijh} = -V_i^2 \varrho_{ijh} - V_i V_j [\rho_{ijh} sin(\theta_i - \theta_j) - \varrho_{ijh} cos(\theta_i - \theta_j)]$$

   where P_{ijh} denotes the real power flow on line (i, j, h), Q_{ijh} denotes the reactive power flow on line (i, j, h), V_i denotes the voltage at node i, and θ_i denotes the phase angle at node i. Also, ρ (conductance) and ϱ (susceptance) are given by

$$\rho_{ijh} = r_{ijh}/(r_{ijh}^2 + x_{ijh}^2)$$

   and

$$\varrho_{ijh} = -x_{ijh}/(r_{ijh}^2 + x_{ijh}^2),$$

   where resistance is denoted by r_{ijh} and the reactance is denoted by x_{ijh}. The DC approximation assumes that the resistance of each line is negligible relative to the reactance, the voltage magnitude at each bus is equal to 1, and the phase angle difference $(\theta_i - \theta_j)$ across any line (i, j, h) has sufficiently small magnitude so that $cos(\theta_i - \theta_j) \approx 1$ and $sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$. Thus, by implementing these assumptions in the equations above for AC flow, we get $Q_{ijh} = 0$ and

$$P_{ijh} = (\theta_i - \theta_j)/x_{ijh},$$

   which is equivalent to $current_{ijh} = (voltage_i - voltage_j)/resistance_{ijh}$ in DC power flow equations.

2. After the cascade's initiation, flows are redistributed in discrete time steps (i.e. at each

discrete time-step, we compute the flows on all lines and remove any failed lines, and then move to the next discrete time-step and repeat the same thing for the new network topology).

3. Any line trips deterministically at time $t+1$ if its flow capacity threshold is exceeded at time t.

4. Line flow capacities remain the same before and after reactance tweaking.

5. Transmission lines are the only components that fail and the tripped lines are not back in service at the time of the SO's response.

6. We shall impose the constraint that all load and generation buses remain connected to each other for the duration of (but not necessarily after) the cascade. Thus, it is possible that at the time of the SO's response (i.e. at stage $T+1$), some demand and generation nodes have become islanded away from each other, but while the cascade is propagating (i.e. up until stage T) all the nodes with non-zero injections remain connected to each other. We make this assumption because we use the DC power flow model in our formulation, and imbalances between the amount of generation and the amount of demand can heavily impact the accuracy of the assumptions underlying the DC approximation to AC flows (unless $T = 0$, in which case the problem becomes static and such concerns will not matter, and hence, we allow the set of all possible disturbances for the special case of $T = 0$). Thus, we make this assumption in an effort to make our DC approximation more accurate. When working with AC power flows, if a bus with a considerable amount of demand or generation becomes islanded away from the rest of the network at the initiation of (or during) the cascade, the resulting imbalance between load and generation could lead to changes in system frequency. Such changes could considerably weaken the DC approximation in tracking how the AC power flows actually redistribute over the course of cascading failures, and hence, we make the above assumption to avoid such scenarios. Nevertheless, the graph-theoretic techniques that we develop below are general and can be extended and used to model settings in which demand-generation imbalances are used in the set of possible disturbances by the adversary.

2.2 Modeling the Evolution of the Graph Topology Prior to the System Operator's Response

We assume that each line can fail due to two reasons:

1) It trips (either because of the adversary's disturbance (at time $t = 0$ only) or because its flow exceeded its capacity limit (for $t \geq 1$)).

2) Its end buses become disconnected from the rest of the network (note that given our assumption of balanced load and generation, only those buses that have zero injection are allowed to become disconnected).

Thus, after the initiation of the cascade, at each stage the topology changes in two ways: 1) Removal of the tripped (i.e. overloaded) edges from the underlying graph; 2) Splitting of the initially connected graph into two or more islands (recall that we further enforce the constraint that buses with non-zero injections remain connected to one-another in the same island).

At time 0, the adversary trips (removes) up to k lines. The state of line (i, j, h) due to the attack at time 0 is binary: $s^0_{ijh} = 0$ if (i, j, h) is tripped by the adversary and $s^0_{ijh} = 1$ otherwise. Next, we check if the disturbance has disconnected any nodes from $\mathcal{B}^o_{\pm}$. If $y^0_i = 1$, nodes i is still connected to $\mathcal{B}^o_{\pm}$ and if $y^0_i = 0$ it's been islanded (in the remainder of the chapter, we shall set one node $g \in \mathcal{B}^o_{\pm}$ as the reference node, and determine the value of $\mathbf{y}^t$ based on whether each node is connected to g or not). The binary *failure state* of line (i, j, h) at stage 0, z^0_{ijh}, determines whether line (i, j) has failed (either due to tripping or due to islanding).

The nomenclature follows the same procedure for $t \geq 1$. The only distinction is that for $t \geq 1$, the binary *tripping state* of a line (denoted by s^t_{ijh}) will only change if line (i, j, h) trips due to overloading (and not a physical interference by the adversary) at stage t; this is because the adversary is no longer part of the game and the cascade is propagating purely due to the natural redistribution of flows following the initial topology-changing disturbance. Note also that z^t_{ijh}

takes into account both forms of line failure, and hence, it is updated at the end of stage t. Thus, $z_{ijh}^t \leq s_{ijh}^t$ for all t, and $s_{ij}^t \leq z_{ijh}^{t-1}$ for $t \geq 1$.

Proposition 1. *In order for the failure indicator* $\mathbf{z}^t$ *to take a value of 0 for either a disconnected or a tripped line, but* not *take on a value of 0 for an active line, it is necessary that the following hold:*

$$|M'|\mathbf{y}^t \geq \mathbf{z}^t \geq \mathbf{s}^t + |M'|\mathbf{y}^t - 2 \text{ and } \mathbf{s}^t \geq \mathbf{z}^t \; \forall t \geq 0.$$

Proof. Please see Section 3.5. □

Remark 2. *One advantage of having separate variables for line tripping* ($\mathbf{s}^t$) *and line failure* ($\mathbf{z}^t$) *is that we would be able to keep track of whether a line has tripped versus simply being islanded. If the* z *variable flips to zero for a line before* s *does, then we would know that the line has not tripped and its failure is due to islanding.*

Using the following theorem, we derive a full MILP approach for tracking whether a line/node is still connected to another node at the end of each stage.

Theorem 1. *Given a binary vector* $\mathbf{y}^t \in \{0,1\}^{|\mathcal{B}|}$ *with* $y_g^t = 1$, *we have the following:*

If for some $\mathbf{e}^t \in \{0,1\}^{|\mathcal{B}|}, \mathbf{p}^t \in \{0,1\}^{|\mathcal{B}|}, \mathbf{u}^t \in \{0,1\}^{|\mathcal{E}|}, \mathbf{w}^t \in \{0,1\}^{|\mathcal{E}|}$ *we have*

$$\begin{aligned}
&\mathbf{s^t} + \mathbf{w^t} - y_i^t \mathbf{1} \leq \mathbf{1}, \\
&-\mathbf{w^t} \leq M'\mathbf{p^t} \leq \mathbf{w^t}, \\
&p_g^t = 0; \;\; p_i^t = 1 - y_i^t, \\
&\sum_{\ell \in \mathcal{E}_n} u_\ell^t = 2e_n^t \; \forall n \neq g, n \neq i \\
&\sum_{\ell \in \mathcal{E}_g} u_\ell^t = y_i^t; \;\; \sum_{\ell \in \mathcal{E}_i} u_\ell^t = y_i^t, \\
&\mathbf{u}^t \leq \mathbf{x}^t; \;\; \mathbf{e}^t \leq \mathbf{y}^t; \;\; \mathbf{e}^t \leq y_i^t \mathbf{1}.
\end{aligned} \tag{2.1}$$

then $<g \overset{t}{\leftrightarrow} i> \equiv <y_i^t = 1>$ *and* $<g \overset{t}{\nleftrightarrow} i> \equiv <y_i^t = 0>$.

Proof. Please see Section 2.4. □

The key ideas in the above theorem are "cut removal" and "path existence"; when a node i has been islanded away from another node g, it means that an $i - g$ cutset (please see Section 2.4 for definition) has been removed from the graph; conversely, if i and g are connected, at least one path has to exist between them.

Although the above result provides a general framework for tracking the evolution of the topology, since we are focusing on the case of disturbances that do not cause imbalances between demand and generation, we shall impose for any bus $q \in \mathcal{B}_{\pm}^{o}$ the constraint that $y_q^t = 1$ for all $t \leq T - 1$, while leaving the y_j^t as free binary variables for all $j \notin \mathcal{B}_{\pm}^{o}$.

2.3 Modeling the Redistribution of Flows

It remains to describe how the redistribution of flows is tracked. To do this, we first compute flows and voltages for all active lines and nodes, respectively, at each stage using Kirchhoff's Current Law (KCL) and Kirchhoff's Voltage Law (KVL) for DC flows:

$$M\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \mathbf{b}^t \; \forall t \geq 1$$

where the above equality simply imposes at stage t that any line that has survived by the end of stage $t - 1$ is allowed to have a non-zero flow, and that the sum of incoming and outgoing flows at each node should equal zero (i.e. KCL); then, we write

$$\text{diag}(\mathbf{x}^o)\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \text{diag}(\mathbf{z}^{t-1})M'\theta \; \forall t \geq 1$$

where the above equality imposes KVL at stage t for the end nodes of any line that has survived by the end of stage $t-1$. As for tracking whether a line has overflown, we set

$$(z_{ijh}^{t-1}+s_{ijh}^{t}-1)|I_{ijh}^{t}| \leq \overline{cap}_{ijh}^{a} \; \forall (i,j,h)\in\mathcal{E}, \forall t \geq 1$$

where the above inequality ensures that any line that has not failed by the end of stage $t-1$ and has also not tripped at stage t should have a flow below its capacity limit, and the inequality becomes redundant if the line has failed by the end of $t-1$ or is tripping at t. Similarly, we write

$$z_{ijh}^{t-1}|I_{ijh}^{t}| \geq \overline{cap}_{ijh}^{f}(z_{ijh}^{t-1}-s_{ijh}^{t}) \; \forall (i,j,h)\in\mathcal{E}, \forall t \geq 1$$

to ensure that any line that trips at stage t must have first survived by the end of stage $t-1$ and then must have been overloaded at stage t (otherwise the constraint becomes redundant; recall that we also impose $z_{ijh}^{t-1} \geq s_{ijh}^{t}$).

Note that the absolute value terms and binary$\times$continuous product terms in the above equations and inequalities can be linearized using standard techniques that shall be described in the next chapter. Thus, what we have developed so far gives a MILP framework for tracking the multi-stage propagation of cascading failures as per our model.

The procedure shown in Figure 2-7 summarizes the approach described in this chapter for generating multi-stage constraints for $t \geq 1$.

Remark 3. *The cascade initiation mechanism described in this chapter is an example of the well-known "$N-k$" contingency framework (e.g. see [10]). Note that in our case k denotes the number of transmission lines that could fail at time 0, and according to [10], mostly small values of k, such as $k \in \{2,3,4,5\}$ are realistically of interest in power systems.*

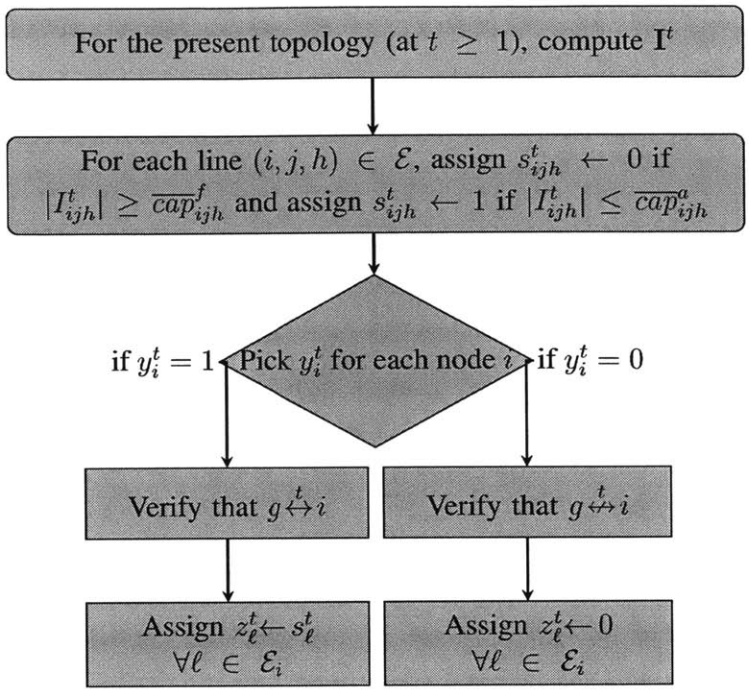


Figure 2-7: Procedure for generating constraints for tracking multi-stage evolution of the topology and flow redistributions for stage $t \geq 1$

2.4 Proofs

Proof of Proposition 1. There are two ways in which line (i, j, h) is considered failed: if it trips, or if its end buses get islanded. If a line (i, j, h) trips at time t, a value of zero will be assigned to s^t_{ijh}; thus, it is necessary to impose $s^t_{ijh} \geq z^t_{ijh}$ to ensure that any tripped line will also have $z^t_{ijh} = 0$. If the end nodes of a line (i, j, h) become islanded at time t, a value of zero will be assigned to y^t_i and y^t_j; thus, it is necessary to impose $y^t_i + y^t_j \geq z^t_{ijh}$ to ensure that $z^t_{ijh} = 0$ when the end nodes of the line have failed. However, it is also necessary to ensure that if none of the above scenarios happen, z^t_{ijh} will have no choice but to maintain $z^t_{ijh} = 1$. Thus, it is necessary to impose $z^t_{ijh} \geq s^t_{ijh} + y^t_i + y^t_j - 2$ to ensure that if $y^t_i = 1$ and $y^t_j = 1$ and $s^t_{ijh} = 1$, then $z^t_{ijh} = 1$. □

Definition 5. *A "*cut*" in a network $(\mathcal{B}, \mathcal{E})$ is a partition of $\mathcal{B}$ into two disjoint subsets (say, U and V). The cutset of a cut $C = (U, V)$ is then defined as the set of lines that have one end-node in U*

and their other end-node in V. For nodes $i \in \mathcal{B}$ and $g \in \mathcal{B}$, an $i-g$ cut is a cut in which i and g belong to two disjoint subsets of nodes as per the partition defined by the cut (e.g. $g \in U$ and $i \in U$).

Proof of Theorem 1. We are given a binary vector $\mathbf{y}^t \in \{0,1\}^{|\mathcal{B}|}$ with $y_g^t = 1$. If nodes i and g are connected (i.e. if $g \overset{t}{\leftrightarrow} i$), then a path of active lines has to exist between nodes g and i. Thus, a sequence of active nodes has to exist between i and g on this active path such that each of these nodes have exactly one line coming into them and another line coming out of them (i.e. a total of two distinct lines per node; we write this as $\sum_{\ell \in \mathcal{E}_n} u_\ell^t = 2e_n^t \; \forall n \neq g, i$, where e_n^t determines whether node n is on this path or not, and u_ℓ^t determines whether line ℓ is on this path or not), while i and g themselves exactly have one line connected to them from this path (we write this as $\sum_{\ell \in \mathcal{E}_g} u_\ell^t = 1$ and $\sum_{\ell \in \mathcal{E}_i} u_\ell^t = 1$); thus, if for some $\mathbf{e}^t \in \{0,1\}^{|\mathcal{B}|}$ and $\mathbf{u}^t \in \{0,1\}^{|\mathcal{E}|}$ we have

$$\begin{aligned} &\sum_{\ell \in \mathcal{E}_n} u_\ell^t = 2e_n^t \; \forall n \neq g, i \\ &\sum_{\ell \in \mathcal{E}_g} u_\ell^t = y_i^t; \;\; \sum_{\ell \in \mathcal{E}_i} u_\ell^t = y_i^t, \\ &\mathbf{u}^t \leq \mathbf{x}^t; \;\; \mathbf{e}^t \leq \mathbf{y}^t; \;\; \mathbf{e}^t \leq y_i^t \mathbf{1}. \end{aligned} \tag{2.2}$$

then $< y_i^t = 1 > \equiv < g \overset{t}{\leftrightarrow} i >$ because plugging $y_i^t = 1$ in the above equations guarantees that an active path exists between i and g. Note that the set of constraints on the last line above ensure that only active lines and nodes can be placed on the path between i and g, and that if $y_i^t = 0$ then all the above constraints become redundant.

Now if nodes i and g are not connected (i.e. if $g \overset{t}{\nleftrightarrow} i$), then a $g-i$ cutset must have been removed from the network. It has been shown in [52] that if we partition the graph into two sets of nodes, say V_0 and V_1 (one containing g and the other containing i; by assigning $p_g^t = 0$ and $p_i^t = 1$ we make sure of this), then by enforcing the constraint $-\mathbf{w^t} \leq M'\mathbf{p^t} \leq \mathbf{w^t}$ we can ensure that any line between V_0 and V_1 is labeled as a cut edge (i.e. if line (i,j,h) is a cut edge, we will have $w_{ijh}^t = 1$). To tailor this idea to the needs of our problem, we add the constraint that $\mathbf{s^t} + \mathbf{w^t} - y_i^t \mathbf{1} \leq \mathbf{1}$ to

ensure that when $y_i^t = 0$, only failed lines can be placed on the cutset, and when $y_i^t = 1$ these equations become redundant. Thus, if for some $\mathbf{p}^t \in \{0,1\}^{|\mathcal{B}|}$ and $\mathbf{w}^t \in \{0,1\}^{|\mathcal{E}|}$ we have

$$\begin{aligned} &\mathbf{s^t} + \mathbf{w^t} - y_i^t \mathbf{1} \leq \mathbf{1}, \\ &-\mathbf{w^t} \leq M'\mathbf{p^t} \leq \mathbf{w^t}, \\ &p_g^t = 0; \;\; p_i^t = 1 - y_i^t, \end{aligned} \tag{2.3}$$

then $< y_i^t = 0 > \equiv < g \overset{t}{\nleftrightarrow} i >$ because plugging $y_i^t = 0$ in the above equations guarantees that a $g - i$ cutset containing only failed lines exists in the graph. □

Chapter 3

Can Post-Disturbance Tweaking of Transmission Line Reactances Increase Resilience to Cascading Failures?

In this chapter we shall introduce and formulate two different measures of the efficacy of reactance tweaking in post-disturbance control, and develop tractable techniques for solving the optimization problems underlying those measures. We will conclude this chapter by presenting a case study.

In the literature on cascading failures, one of the most commonly used measures for evaluating the efficacy of a post-disturbance control mechanism is the severity of the worst-case scenario disturbance in the presence of the post-disturbance control mechanism of interest (e.g. see [64], [2], and the references therein for the case of post-disturbance "line switching"). Hence, for measuring the efficacy of post-contingency reactance tweaking in increasing system resilience, we start by developing a framework for measuring the worst-case-scenario efficacy of this mechanism.

To do so, we first define the notion of a "yield-minimizing $N-k$ contingency" (the term "yield" refers to the maximum amount of demand that the SO can satisfy post-disturbance by dispatching the system without violating any physical constraints), which deals with the following question:

In a power network with N lines, what set of at most k lines should the adversary cause to fail at time 0 so that their failure leads to a cascade that eventually results in the minimum yield? This implies that the adversary seeks to find a disturbance with the worst-case-scenario yield. Thus, our measure of the severity of the cascading failures is the amount of load that needs to be shed post-disturbance to bring the network back to reliable operating conditions while receiving help from reactance tweaking technologies. Next, as the second measure of efficacy, we shall consider the "general efficacy" of post-disturbance reactance-tweaking, which focuses on the efficacy of this strategy in increasing yield after *any* $N - k$ disturbance for a given k, without requiring that the disturbance was initiated by a strategic adversary.

Before we proceed to the formulation, we first establish a result which we shall use frequently as a building block for obtaining several of our main results. This result establishes, for a DC flow network, an upper bound on the node potentials (i.e. phase angles as per the DC approximation to AC flows). Given that the field of DC electric circuits has been around for many decades, one would expect that this result should already exist somewhere in the literature, possibly dating back to even decades ago. However, the author could not find any previous study in the existing literature that has proved this result, and Theorem 2 below has been derived independently by the author.

Theorem 2. *In any connected resistive network $(\mathcal{B}, \mathcal{E})$ in the set of all network topologies with $|\mathcal{B}|$ buses and $|\mathcal{E}|$ lines, if we fix $\theta_p = 0$ for some arbitrary bus $p \in \mathcal{B}$ and place a total of G units of positive net DC injection into the network and extract a net total of $-G$ units of current from any subset of nodes that do not have a positive net injection (in other words, if $G \triangleq \sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o)$), then we have:*

$$|\theta_i| \leq G \sum_{(p,q,h) \in \mathcal{E}} x_{pqh} \quad \forall i \in \mathcal{B}, \tag{3.1}$$

where x_{pqh} is the resistance of line (p, q, h). This upper bound is tight and is achieved by a line graph topology with a positive injection of magnitude G at one end and a negative injection of the same magnitude at its other end, with all other nodes having zero injection.

Proof. Please see Section 3.5. □

Remark 4. *Note that inequality (3.34) derived in the proof of Theorem 2 (presented in Section 3.5) gives a formula for computing an upper bound on the voltage difference (i.e. phase-angle difference as per the DC approximation to AC flows) between two arbitrary nodes in a connected DC flow network. As can be observed in inequality (3.34), if we think of line resistances as edge weights, then by computing weighted path lengths (preferably, weighted shortest-path lengths) in the corresponding weighted graph, one can use inequality (3.34) to compute bounds on nodal voltage differences.*

3.1 The System Operator's Response: Reactance-Controlled Yield Maximization

Before formally stating the optimization problem of interest, let us introduce the key nomenclature used in the formulation:

- $T+1$ The stage at which post-disturbance control is implemented by the system operator
- b_i^d The satisfied demand post-disturbance at bus i after the system operator's response
- b_i^g The amount of post-disturbance generation at bus i after the system operator's response
- b_i^{T+1} Net nodal injection at node i as per the SO's post-disturbance dispatch. Note that for the special case of $t = T+1$ (i.e. the case of post-disturbance dispatch), we have $b_i^{T+1} \triangleq b_i^g - b_i^d$ for each node i.
- $\overline{b}_i^d$ The amount of power demanded at bus i
- $\overline{b}_i^g$ The generation capacity at bus i
- τ_{ijh} Binary variable indicating whether the reactance of a line (i, j, h) at time $T+1$ is allowed to be tweaked ($\tau_{ijh} = 1$ if tweaking is allowed, and 0 otherwise).

- H The total number of lines on which reactance tweaking is allowed

Definition 6. *We define the "reactance-controlled yield" as the maximum amount of demand that the SO can satisfy post-disturbance by dispatching the system without violating any physical constraints while tweaking transmission line reactances:*

$$
\begin{aligned}
\max_{D} \ & \sum_{i\in\mathcal{B}} b_i^d && (3.2)\\
s.t. \ & I_{ijh}^{T+1} x_{ijh}^{T+1} = s_{ijh}^{T}(\theta_i^{T+1} - \theta_j^{T+1}),\ \forall (i,j,h) \in \mathcal{E} \\
& b_i^g - b_i^d = \sum_{(i,j,h)\in\mathcal{E}_i} I_{ijh}^{T+1},\ \forall i \in \mathcal{B} && (a)\\
& x_{ijh}^{T+1} - x_{ijh}^{o} - \tau_{ijh}(x_{ijh}^{\max} - x_{ijh}^{o}) \leq 0\ \forall (i,j,h)\in\mathcal{E} \\
& x_{ijh}^{o} - x_{ijh}^{T+1} + \tau_{ijh}(x_{ijh}^{\min} - x_{ijh}^{o}) \leq 0\ \forall (i,j,h)\in\mathcal{E} \\
& \sum_{(i,j,h)\in\mathcal{E}} \tau_{ijh} = H\ \forall (i,j,h) \in \mathcal{E} && (b)\\
& -\overline{cap}_{ij}^{a} \leq I_{ijh}^{T+1} \leq \overline{cap}_{ij}^{a},\ \forall (i,j,h) \in \mathcal{E} \\
& 0 \leq b_i^g \leq \overline{b}_i^g, \quad \forall i \in \mathcal{B} && (c)\\
& 0 \leq b_i^d \leq \overline{b}_i^d, \quad \forall i \in \mathcal{B} && (d)\\
& x_{ijh}^{T+1}, \theta_i^{T+1}, I_{ijh}^{T+1}, b_i^g, b_i^d \in \mathbb{R}, \forall i \in \mathcal{B}, (i,j,h) \in \mathcal{E} \\
& \tau_{ijh} \in \{0,1\}, \quad \forall (i,j,h) \in \mathcal{E}.
\end{aligned}
$$

where $D \triangleq \{\mathbf{x^{T+1}}, \theta^{\mathbf{T+1}}, \mathbf{I^{T+1}}, \mathbf{b}^g, \mathbf{b}^d, \tau\}$ *is the set of decision variables, and* $\mathbf{s}^T \in \{0,1\}^{|\mathcal{E}|}$ *is the binary vector of line failure states that is given to the problem as an input parameter. We shall interchangeably refer to the above reactance-controlled yield-maximization problem as "SO's post-disturbance control problem".*

In the above formulation, the objective is to maximize the amount of load served. The first two constraints are KVL and KCL for DC flows, the next three constraints ensure that the reactances of at most H lines can be tweaked within a prescribed range (and all the other lines will retain their original reactances), the sixth constraint ensures no line flow capacities are violated, the

seventh constraint ensures that the generation capacities of each bus are not exceeded, and the eighth constraint ensures that the served load at each bus will be between zero and the original demand.

Note that we could have allowed the lower limit in the seventh constraint in (3.2) to be a positive value instead of zero (for instance, to model ramp-down rates of generators or minimum generation requirements for operating generation units); that is a straightforward extension which requires the addition of a binary variable to model unit commitment for each generator (to maintain the possibility of disconnecting a generator from the system at the time of post-disturbance dispatch); however, in order to keep the model concise and have as few binary variables as possible in the above problem, we do not include this variable in our model. Also, in the above formulation, we could have imposed limits on the phase-angle difference across each line (which would have been a set of very simple linear constraints); however, to maintain our focus on the interaction between line capacities and the operating range of the reactance-tweaking devices, we have excluded this constraint from the above formulation.

Note that the set of decision variables in (3.2) contains some binary variables which introduce non-convexity in the problem but their continuous relaxation is linear. However, a more complicated source of non-convexity in the problem above is introduced by the continuous$\times$continuous bilinear terms in the first constraint (this continuous$\times$continuous bilinear term stems from the product of current and reactance in the KVL).

As shown in Theorem 3 below, we develop a reformulation scheme that turns the SO's post-disturbance control problem (3.2) into a MILP by shifting the non-convexity stemming from continuous$\times$continuous bilinear terms to auxiliary binary variables that we introduce into the problem in linear terms (along with some continuous auxiliary variables that are also introduced into the problem in linear terms).

Theorem 3. *If Γ^* is the globally optimal value of the objective function in the optimization problem*

(3.2), then

$$\Gamma^* = \max_{\omega} \sum_{i \in \mathcal{B}} b_i^d \tag{3.3}$$

$$\begin{aligned}
s.t.\quad & \boldsymbol{\mu} + \Upsilon(\mathbf{s}^T - 1) \leq M'\boldsymbol{\theta}^{T+1} \leq \boldsymbol{\mu} + \Upsilon(1 - \mathbf{s}^T); \\
& -diag(\mathbf{s}^T)\overline{\mathbf{cap}}^a \leq \boldsymbol{I}^{T+1} \leq diag(\mathbf{s}^T)\overline{\mathbf{cap}}^a; \\
& \boldsymbol{\mu} - diag(\boldsymbol{x}^o)\boldsymbol{I}^{T+1} - 2\Upsilon\boldsymbol{\tau} \leq 0; \\
& -\boldsymbol{\mu} + diag(\boldsymbol{x}^o)\boldsymbol{I}^{T+1} - 2\Upsilon\boldsymbol{\tau} \leq 0; \\
& diag(\mathbf{x}^{\min})\boldsymbol{\alpha} \leq \boldsymbol{\delta} \leq diag(\mathbf{x}^{\max})\boldsymbol{\alpha}; \\
& \boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{\varphi};\ \ \boldsymbol{I}^{T+1} = \boldsymbol{\beta} - \boldsymbol{\varphi}; \\
& \boldsymbol{\beta} - diag(\overline{\mathbf{cap}}^a)\boldsymbol{\upsilon} \leq 0;\ \ \boldsymbol{\varphi} - diag(\overline{\mathbf{cap}}^a)(\mathbf{1} - \boldsymbol{\upsilon}) \leq 0; \\
& \boldsymbol{\delta} = \boldsymbol{\phi} + \boldsymbol{\Phi};\ \ \boldsymbol{\mu} = \boldsymbol{\phi} - \boldsymbol{\Phi}; \\
& \boldsymbol{\Phi} - \Upsilon(\mathbf{1} - \boldsymbol{\psi}) \leq 0;\ \ \boldsymbol{\phi} - \Upsilon\boldsymbol{\psi} \leq 0; \\
& \boldsymbol{\phi} \geq 0;\ \ \boldsymbol{\Phi} \geq 0;\ \ \boldsymbol{\varphi} \geq 0;\ \ \boldsymbol{\beta} \geq 0; \\
& \textit{Constraints } (a),\ (b),\ (c),\ (d) \textit{ from Problem } (3.2); \\
& \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\varphi}, \boldsymbol{\mu}, \boldsymbol{\delta}, \boldsymbol{\Phi}, \boldsymbol{\phi}, \boldsymbol{I}^{T+1} \in \mathbf{R}^{|\mathcal{E}|}; \\
& \boldsymbol{\tau}, \boldsymbol{\psi}, \boldsymbol{\upsilon} \in \{0,1\}^{|\mathcal{E}|};\ \ \boldsymbol{\theta}^{T+1}, \boldsymbol{b}^g, \boldsymbol{b}^d \in \mathbb{R}^{|\mathcal{B}|}.
\end{aligned}$$

where $\Upsilon \triangleq G\sum_{(p,q,h)\in\mathcal{E}} x_{pqh}^{\max}$ *and* $\omega = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\varphi}, \boldsymbol{\mu}, \boldsymbol{\delta}, \boldsymbol{\Phi}, \boldsymbol{\phi}, \boldsymbol{I}^{T+1}, \boldsymbol{\tau}, \boldsymbol{\psi}, \boldsymbol{\upsilon}, \boldsymbol{\theta}^{T+1}, \boldsymbol{b}^g, \boldsymbol{b}^d\}$ *(note that* $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\varphi}, \boldsymbol{\mu}, \boldsymbol{\delta}, \boldsymbol{\Phi}, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\upsilon}$ *are all auxiliary variables).*

Proof. Please see Section 3.5. □

Remark 5. *In addition to making the problem amenable to column-and-row-generation (as we shall discuss later in this chapter) the above MILP reformulation also allows for using the existing powerful MILP solvers which are "mature" in the sense that they are fast, robust, and capable of handling large problems with millions of variables (e.g. see [32] for a discussion).*

By denoting the feasible set in (3.3) by Ω, the above can be written as

$$\max_{\omega \in \Omega} \sum_{i \in \mathcal{B}} b_i^d. \tag{3.4}$$

So far, we described how the SO re-dispatches the system at time $T+1$, using reactance tweaking (performed on up to H transmission lines) and load-shedding as control mechanisms, with the sole purpose of maximizing the amount of demand satisfied post-disturbance (regardless of generation cost) without violating any physical constraints such as line capacities or generator capacities. Note that due to the high costs of power electronics, we are interested in small values of H (e.g. 1-3).

Now that we have formulated the system operator's problem, we shall describe how the adversary plays in this zero-sum Stackelberg game.

3.2 The Adversary Minimizes the Maximum Post-Disturbance Yield: A Bilevel MINLP for the Worst-Case-Scenario

Given that the reactance-controlled yield problem is a maximization problem as defined in (3.2), the adversary's problem is a min-max bilevel program (since the adversary goes first and the SO goes next while sharing the same objective function, the problem of interest is a zero-sum Stackelberg game) shown below:

$$\min_{A \in \mathcal{A}} \max_{\omega \in \Omega} \sum_{i \in \mathcal{B}} b_i^d \tag{3.5}$$

where A is the vector of the adversary's decision variables, and $\mathcal{A}$ is the feasible set of the adversary's problem (which are defined explicitly in (3.6) below); this bilevel program, which is our

measure of the worst-case-scenario resilience, can now be written as

$$\min_{A} \ \max_{\omega\in\Omega} \ \sum_{i\in\mathcal{B}} b_i^d \tag{3.6}$$

$$\text{s.t.} \quad \mathbf{1}'\mathbf{s}^o \geq |\mathcal{E}| - k$$

$$M\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \mathbf{b}^o \ \forall t \geq 1$$

$$\text{diag}(\mathbf{x}^o)\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \text{diag}(\mathbf{z}^{t-1})M'\theta \ \forall t \geq 1$$

$$(z_{ijh}^{t-1}+s_{ijh}^t-1)|I_{ijh}^t| \leq \overline{cap}_{ijh}^a \ \forall (i,j,h)\in\mathcal{E}, \forall t \geq 1$$

$$z_{ijh}^{t-1}|I_{ijh}^t| \geq \overline{cap}_{ijh}^f (z_{ijh}^{t-1}-s_{ijh}^t) \ \forall (i,j,h)\in\mathcal{E}, \forall t \geq 1$$

$$|M'|\mathbf{y}^t \geq \mathbf{z}^t \geq \mathbf{s}^t + |M'|\mathbf{y}^t - 2 \ \ \forall t \geq 0$$

$$y_i^t = 1 \ \ \forall i \in \mathcal{B}_{\pm}^o, \forall t \leq T-1 \ (T \geq 1 \text{ only})$$

$$\mathbf{y}^{t-1} \geq \mathbf{y}^t, \ \ \forall t \leq T-1 \ (T \geq 1 \text{ only})$$

$$\mathbf{s}^t \geq \mathbf{z}^t \ \ \forall t \geq 0$$

$$\mathbf{z}^{t-1} \geq \mathbf{s}^t \ \ \forall t \geq 1$$

$$i \overset{t}{\leftrightarrow} j \ \ \forall \, i,j \in \mathcal{B}_{\pm}^o, 1 \leq t \leq T-1 \ (T \geq 1 \text{ only})$$

$$z_{ijh}^t, \ s_{ijh}^t, \ y_q^t \in \{0,1\}, \quad \forall (i,j,h) \in \mathcal{E}, \ \forall q \in \mathcal{B}.$$

where $A = \{\mathbf{s}^0, \ldots, \mathbf{s}^T, \mathbf{z}^0, \ldots, \mathbf{z}^{T-1}, \mathbf{y}^0, \ldots, \mathbf{y}^{T-1}, \mathbf{I}^1, \ldots, \mathbf{I}^T, \theta^1, \ldots, \theta^T, \mathcal{P}\}$. Here, $\mathcal{P}$ denotes the set of all variables used in keeping track of the connectivity of the buses (which are the auxiliary variables defined in Theorem 1) for all $t \in \{0, \ldots, T-1\}$. In (3.6), the first constraint requires that the adversary trips up to k lines at $t = 0$, the the second and third constraints are for KVL and KCL at each stage (discussed in Chapter 2), the fourth and the fifth constraints keep track (as discussed in Chapter 2) of the lines that have survived and failed, respectively, and the remainder of the constraints keep track of the connectivity (also ensuring that no failed line is reactivated) and evolution of the topology as established in Proposition 1 and Theorem 1.

The next step is to linearize all the binary×continuous product terms and absolute value terms in

problem (3.6) so that we can write the outer-level problem as a MILP. Doing so is a standard trick used frequently in the optimization literature. For some $\sigma \in \mathbb{R}$ and some $\zeta \in \{0,1\}$, if $|\sigma| \leq m$ for some m, then $-m\zeta \leq \hat{\sigma} \leq m\zeta$ and hence we can substitute $\zeta\sigma$ with $\hat{\sigma}$ where appropriate and add the constraints $-m\zeta \leq \hat{\sigma} \leq m\zeta$, $\hat{\sigma} \leq (1-\zeta)m + \sigma$, $\hat{\sigma} \geq (\zeta - 1)m + \sigma$ to the problem to complete this change of variables (this is a very standard trick in the literature and can be verified by inspection).

Dealing with the non-linearity of the absolute values is slightly trickier and hence, we describe it in the following lemma (note that Lemma 1 is standard in the optimization literature and its validity can be verified by inspection).

Lemma 1. *If* $|\mathbf{I}^t| \leq m$, *and the following hold:*

$$\begin{aligned}
&\hat{\mathbf{I}}^t = \mathbf{I}^t_+ + I^t_-; \\
&\mathbf{I}^t = \mathbf{I}^t_+ - \mathbf{I}^t_-; \\
&\mathbf{I}^t_- \leq m(1 - \mathbf{q}^t); \\
&\mathbf{I}^t_+ \leq m\mathbf{q}^t; \\
&\mathbf{I}^t_+ \geq 0;\ \mathbf{I}^t_- \geq 0, \mathbf{q}^t \in \{0,1\}^{|\mathcal{E}|}
\end{aligned} \tag{3.7}$$

then we have $|\mathbf{I}^t| = \hat{I}^t$.

Although both of the above linearization procedures are standard, for practical implementations we need to choose as small a value for m as possible, and hence, we first need to derive upper bounds for choosing the value of m. It is simple to choose the value of m for binary×continuous terms in which the continuous variable is the line flow, because the magnitude of the DC flow in a line can never exceed the total positive injection ($G = \sum_{i \in \mathcal{B}} b^o_i \mathbb{1}_{\mathbb{R}^+}(b^o_i)$) in the network (hence, setting $m = G$ will work just fine). Choosing the value of m for binary×continuous terms in which the continuous variable is θ is straightforward using the inequality we derived in Theorem 2. For the remainder of this chapter, we shall denote the set of the decision variables in the linearized

reformulation of the adversary's problem by a and the feasible set of this linearized reformulation by $\tilde{a}$, and hence, the bi-level MILP reformulation of the min-max problem can be written as:

$$\min_{a\in\tilde{a}} \max_{\omega\in\Omega} \sum_{i\in\mathcal{B}} b_i^d \tag{3.8}$$

3.3 Solving the Bilevel MINLP underlying the Min-Max Problem

Now that Problem (3.2) has been reformulated as a MILP, the min-max problem (3.6) can be solved to global optimality using a decomposition method involving column and row generation ([62], [63], [64]). Before we proceed to formally establishing how the column-and-row-generation framework can be tailored to our problem, we describe the core idea behind column and row generation in the simple example shown below (for a formal exposition of the theoretical foundations of the column-and-row-generation framework please refer to [63]).

Example 1. Consider the following bilevel MILP:

$$\begin{aligned} \min_{\substack{0\le p+q\le 3 \\ q\in\mathbb{R},p\in\{0,1\}}} \ \max_{x,y,z,w} \ & x+y \\ \text{s.t.} \ & x+y\le z+5;\ 2x+3y\le w+15 \\ & x+y\ge 0;\ x-y\le p;\ x,y\in\mathbb{R} \\ & z\le w;\ w,z\in\{0,1\} \end{aligned} \tag{3.9}$$

and note that the inner-level and the outer-level programs are both MILPs coupled through only a binary variable (i.e. through p). Now, let us define $\mathcal{I} \triangleq \{[0,0]', [0,1]', [1,1]'\}$ as the set of all possible outcomes for the vector $[z,w]'$ that are feasible for the above problem. The above can then

be rewritten as:

$$\min_{\substack{0\le p+q\le 3\\ q\in\mathbf{R},p\in\{0,1\}}} \max_{[z,w]'\in\mathcal{I}} \max_{x,y} x+y \tag{3.10}$$
$$\text{s.t.} \quad x+y\le z+5;\ 2x+3y\le w+15,$$
$$x+y\ge 0;\ x-y\le p;\ x,y\in\mathbb{R}.$$

Note that the "innermost-level" of the above problem is a linear program. If we denote by $\mathbf{c}([z,w]',p)$ the right-hand-side of the constraints of this LP (i.e., $\mathbf{c}([z,w]',p)=[z+5,w+15,0,p]'$), and denote by A the constraint matrix of this LP, then the above can be written as:

$$\min_{\substack{0\le p+q\le 3\\ q\in\mathbf{R},p\in\{0,1\}}} \max_{[z,w]'\in\mathcal{I}} \max_{x,y} [1,1]\begin{bmatrix}x\\y\end{bmatrix} \tag{3.11}$$
$$\text{s.t.} \quad A\begin{bmatrix}x\\y\end{bmatrix}\le \mathbf{c}([z,w]',p)$$
$$x,y\in\mathbb{R},$$

Since the innermost-level LP is feasible and is bounded, strong duality holds, and hence we can write the above problem as:

$$\min_{\substack{0\le p+q\le 3\\ q\in\mathbf{R},p\in\{0,1\}}} \max_{[z,w]'\in\mathcal{I}} \min_{\boldsymbol{h}} \mathbf{c}'([z,w]',p)\boldsymbol{h} \tag{3.12}$$
$$\text{s.t.} \quad A'\boldsymbol{h}=\begin{bmatrix}1\\1\end{bmatrix}$$
$$\boldsymbol{h}\ge 0,$$

where $\boldsymbol{h}$ is the vector of dual variables. Next, we create a copy of $\boldsymbol{h}$ for each element of $\mathcal{I}$ (each copy will be treated as a distinct decision variable; for instance, $\boldsymbol{h}^{0,0}$ which corresponds to element $[0,0]'$ of $\mathcal{I}$ is considered as a different decision variable than $\boldsymbol{h}^{0,1}$ which corresponds to element

$[0,1]'$ of $\mathcal{I}$). Then, enumerate over all elements of $\mathcal{I}$, and rewrite the above as:

$$\min_{\substack{0\le p+q\le 3\\ q\in\mathbb{R},p\in\{0,1\}}} \quad \min_{\boldsymbol{h}^{0,0},\boldsymbol{h}^{0,1},\boldsymbol{h}^{1,1}} \quad \ell \tag{3.13}$$

$$\begin{aligned}
\text{s.t.} \quad & \ell \geq \mathbf{c}'([0,0]',p)\boldsymbol{h}^{0,0} \\
& A'\boldsymbol{h}^{0,0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{0,0} \geq 0, \\
& \ell \geq \mathbf{c}'([0,1]',p)\boldsymbol{h}^{0,1} \\
& A'\boldsymbol{h}^{0,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{0,1} \geq 0, \\
& \ell \geq \mathbf{c}'([1,1]',p)\boldsymbol{h}^{1,1} \\
& A'\boldsymbol{h}^{1,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{1,1} \geq 0,
\end{aligned}$$

By merging the two minimizations, the above can be rewritten as a single-level program:

$$\begin{aligned}
\min_{p,q,\ell,\boldsymbol{h}^{0,0},\boldsymbol{h}^{0,1},\boldsymbol{h}^{1,1}} \quad & \ell \qquad (3.14)\\
\text{s.t.} \quad & \ell \geq \mathbf{c}'([0,0]',p)\boldsymbol{h}^{0,0}\\
& A'\boldsymbol{h}^{0,0} = \begin{bmatrix}1\\1\end{bmatrix}\\
& \boldsymbol{h}^{0,0} \geq 0,\\
& \ell \geq \mathbf{c}'([0,1]',p)\boldsymbol{h}^{0,1}\\
& A'\boldsymbol{h}^{0,1} = \begin{bmatrix}1\\1\end{bmatrix}\\
& \boldsymbol{h}^{0,1} \geq 0,\\
& \ell \geq \mathbf{c}'([1,1]',p)\boldsymbol{h}^{1,1}\\
& A'\boldsymbol{h}^{1,1} = \begin{bmatrix}1\\1\end{bmatrix}\\
& \boldsymbol{h}^{1,1} \geq 0,\\
& 0 \leq p+q \leq 3\\
& q \in \mathbb{R}, p \in \{0,1\}
\end{aligned}$$

Note in the constraints of the above problem that the terms $\mathbf{c}'([0,0]',p)\boldsymbol{h}^{0,0}$, $\mathbf{c}'([0,1]',p)\boldsymbol{h}^{0,1}$ and $\mathbf{c}'([1,1]',p)\boldsymbol{h}^{1,1}$ all contain bilinearities stemming from the product of the binary variable p and some dual variable (e.g. $\mathbf{c}'([0,0]',p)\boldsymbol{h}^{0,0} = 5h_1^{0,0} + 15h_2^{0,0} + ph_4^{0,0}$, which contains the bilinear term $ph_4^{0,0}$). But if we were given a bound on $\boldsymbol{h}^{z,w}$ for each element $[z,w]' \in \mathcal{I}$ and denoted this bound by $\overline{\boldsymbol{h}}^{z,w}$, then we could linearize all these bilinear terms by introducing auxiliary variables; e.g. we define auxiliary variable $u_i^{z,w}$ corresponding to each bilinear term $ph_i^{z,w}$, and write $0 \leq u_i^{z,w} \leq p\overline{h}_i^{z,w}$, $u_i^{z,w} \leq (1-p)\overline{h}_i^{z,w} + h_i^{z,w}$, $u_i^{z,w} \geq (p-1)\overline{h}_i^{z,w} + h_i^{z,w}$ (note that the dual variables in this example are all non-negative). Using this linearization of binary$\times$continuous terms, the above can

be rewritten as the single-level MILP below:

$$\min_{\substack{p,q,\ell,\boldsymbol{h}^{0,0},\boldsymbol{h}^{0,1},\boldsymbol{h}^{1,1}\\ u_4^{0,0},u_4^{0,1},u_4^{1,1}}} \quad \ell \tag{3.15}$$

$$\begin{aligned}
\text{s.t.} \quad & \ell \geq 5h_1^{0,0} + 15h_2^{0,0} + u_4^{0,0} \\
& 0 \leq u_4^{0,0} \leq p\overline{h}_4^{0,0} \\
& u_4^{0,0} \leq (1-p)\overline{h}_4^{0,0} + h_4^{0,0} \\
& u_4^{0,0} \geq (p-1)\overline{h}_4^{0,0} + h_4^{0,0} \\
& A'\boldsymbol{h}^{0,0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{0,0} \geq 0, \\
& \ell \geq 5h_1^{0,1} + 16h_2^{0,1} + u_4^{0,1} \\
& 0 \leq u_4^{0,1} \leq p\overline{h}_4^{0,1} \\
& u_4^{0,1} \leq (1-p)\overline{h}_4^{0,1} + h_4^{0,1} \\
& u_4^{0,1} \geq (p-1)\overline{h}_4^{0,1} + h_4^{0,1} \\
& A'\boldsymbol{h}^{0,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{0,1} \geq 0, \\
& \ell \geq 6h_1^{1,1} + 16h_2^{1,1} + u_4^{1,1} \\
& 0 \leq u_4^{1,1} \leq p\overline{h}_4^{1,1} \\
& u_4^{1,1} \leq (1-p)\overline{h}_4^{1,1} + h_4^{1,1} \\
& u_4^{1,1} \geq (p-1)\overline{h}_4^{1,1} + h_4^{1,1} \\
& A'\boldsymbol{h}^{1,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
& \boldsymbol{h}^{1,1} \geq 0, \\
& 0 \leq p+q \leq 3 \\
& q \in \mathbb{R}, p \in \{0,1\}
\end{aligned}$$

But we do not have to solve this MILP all at once; rather, we can decompose its feasible set and then iteratively add new decision variables and constraints (and hence, the name "column and row generation"). To do so, we first consider the above problem for only a subset of $\mathcal{I}$, e.g. only for the constraints and variables corresponding to $[0,0]'$; then, the decomposed problem would be:

$$\begin{aligned}
\min_{p,q,\ell,\boldsymbol{h}^{0,0},u_4^{0,0}} \quad & \ell \qquad (3.16)\\
\text{s.t.} \quad & \ell \geq 5h_1^{0,0} + 15h_2^{0,0} + u_4^{0,0}\\
& 0 \leq u_4^{0,0} \leq p\overline{h}_4^{0,0}\\
& u_4^{0,0} \leq (1-p)\overline{h}_4^{0,0} + h_4^{0,0}\\
& u_4^{0,0} \geq (p-1)\overline{h}_4^{0,0} + h_4^{0,0}\\
& A'\boldsymbol{h}^{0,0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\\
& \boldsymbol{h}^{0,0} \geq 0,\\
& 0 \leq p+q \leq 3\\
& q \in \mathbb{R}, p \in \{0,1\}
\end{aligned}$$

We call the above problem the "Master Problem"; since this problem has a less-constrained feasible set than the original problem but the same objective, it gives a lower bound on the globally optimal solution of the original problem. If we solve this "Mater Problem" to global optimality, its vector of optimal values of decision variables contains a value for p which we shall denote by p^*. We also denote the optimal objective value of the "Master Problem" by ℓ^*. We then form a "Subproblem", which is equivalent to the inner-level program in the original min-max problem that uses $p = p^*$ as

shown below:

$$\begin{aligned} \max_{x,y,z,w} \quad & x + y \\ \text{s.t.} \quad & x + y \leq z + 5 \\ & 2x + 3y \leq w + 15 \\ & x + y \geq 0 \\ & x - y \leq p^* \\ & z \leq w \\ & x, y \in \mathbb{R},\ w, z \in \{0, 1\} \end{aligned} \tag{3.17}$$

The optimal value of the above subproblem gives an upper bound on the optimal value of the original min-max problem. Thus, if the optimal value of this subproblem (i.e. $x^* + y^*$) is equal to the optimal value of the master problem (i.e. ℓ^*) we just solved in the previous step (i.e. if ℓ^*=$x^* + y^*$), then we are done because it means we have reached the optimal solution of the original min-max problem (recall that the master problem gives a lower bound on the optimal value of the min-max problem). If they are not equal, we add the vector $[z^*, w^*]$ (which we just obtained from solving the subproblem) to the subset of $\mathcal{I}$ with which we form the master problem in the next iteration (for instance, if $[z^*, w^*]$ was determined to be $[0, 1]$ by the subproblem, then we will form the new master problem using the subset $\{[0, 0]', [0, 1]'\}$ of $\mathcal{I}$). Next, we update the master problem by adding all the constraints and dual variables corresponding to $[z^*, w^*]$ to the master problem from the previous iteration to form a new master problem (note that we keep all the constraints from the master problem in the previous iteration in the updated master problem; so, the size of the master problem monotonically grows at each iteration). The above procedure is then repeated at each iteration until convergence is achieved. This process is depicted in the figure below.

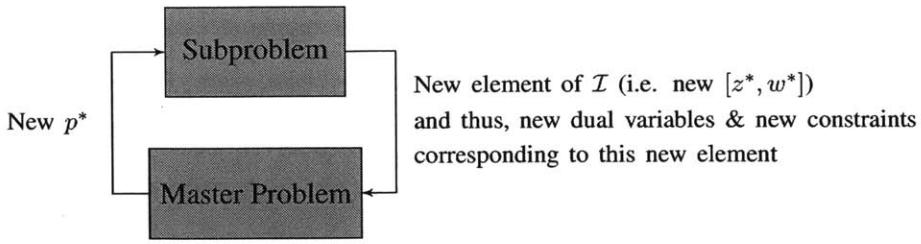


Figure 3-1: Block diagram for the example on column and row generation

Given the finiteness of the number of elements in $\mathcal{I}$, in order to clarify that this procedure converges to the globally optimal solution of the original min-max problem, it only remains to clarify that any repeated value of $[z^*, w^*]$ implies global optimality has been achieved (i.e. it only remains to show that this procedure will not cycle between non-globally-optimal solutions). This has been proved in [63] and [64] by the following argument: To see this, note that if at iteration k the value of $[z^*, w^*]$ obtained from solving the subproblem had already been encountered in some previous iteration t (i.e. $t < k$), it means that the set of dual variables and constraints corresponding to $[z^*, w^*]$ already exist in the Master Problem at iteration k, and hence, the master problem stays the same at $k+1$ and the optimal objective value of the Master Problem at iteration $k+1$ will remain the same as its optimal objective value at iteration k. Note that by strong duality, at iteration $k+1$, we have $UpperBound \leq x^* + y^* = \min_{\boldsymbol{h}^{z^*,w^*}} \mathbf{c}'([z^*, w^*]', p^*)\boldsymbol{h}^{z^*,w^*}$; but we also have that $LowerBound = \ell^* \geq \min_{\boldsymbol{h}^{z^*,w^*}} \mathbf{c}'([z^*, w^*]', p^*)\boldsymbol{h}^{z^*,w^*}$; thus, the lower and upper bounds are equal and we have reached the globally optimal solution (please see [63] or [64] for a formal proof).

The idea behind the column-and-row-generation procedure described in Example 1 above can thus be summarized as follows:

Consider the following min-max problem:

$$\min_{\mathbf{p}\in\{0,1\}^n} \quad \max_{\mathbf{z}\in\{0,1\}^m,\ \mathbf{x}\in\mathbb{R}^d} \quad \mathbf{c}'\mathbf{x} \tag{3.18}$$
$$\text{s.t.} \quad A\mathbf{x} \leq \mathbf{g}(\mathbf{z}, \mathbf{p})$$

where the vector $\mathbf{g}(\mathbf{z}, \mathbf{p})$ is a linear function of both $\mathbf{z}$ and $\mathbf{p}$ (also, n, m, and d are arbitrary positive

integers). By separating binary and continuous variables in the inner level we have:

$$\min_{\mathbf{p}\in\{0,1\}^n} \max_{\mathbf{z}\in\{0,1\}^m} \max_{\mathbf{x}} \mathbf{c}'\mathbf{x} \tag{3.19}$$
$$\text{s.t.} \quad A\mathbf{x} \leq \mathbf{g}(\mathbf{z},\mathbf{p})$$

By strong duality, this is equivalent to:

$$\min_{\mathbf{p}\in\{0,1\}^n} \max_{\mathbf{z}\in\{0,1\}^m} \min_{\boldsymbol{h}} \mathbf{g}'(\mathbf{z},\mathbf{p})\boldsymbol{h} \tag{3.20}$$
$$\text{s.t.} \quad A'\boldsymbol{h} = \mathbf{c}$$
$$\boldsymbol{h} \geq 0,$$

where $\mathbf{h}$ is the vector of dual variables. If we fix some disturbance $\mathbf{p}_0$, then:

$$\cancel{\min_{\mathbf{p}_0}} \max_{\mathbf{z}\in\{0,1\}^m} \max_{\mathbf{x}} \mathbf{c}'\mathbf{x} \tag{3.21}$$
$$\text{s.t.} \quad A\mathbf{x} \leq \mathbf{g}(\mathbf{z},\mathbf{p}_0)$$

This gives an **upper bound** on the original min-max. Refer to its optimal $\mathbf{z}$ as $\mathbf{z}_0^*$. Now fix $\mathbf{z}$ to $\mathbf{z}_0^*$ and solve for $\mathbf{p}$:

$$\min_{\mathbf{p}\in\{0,1\}^n} \cancel{\max_{\mathbf{z}=\mathbf{z}_0^*}} \min_{\boldsymbol{h}} \mathbf{g}'(\mathbf{z},\mathbf{p})\boldsymbol{h} \tag{3.22}$$
$$\text{s.t.} \quad A'\boldsymbol{h} = \mathbf{c}$$
$$\boldsymbol{h} \geq 0,$$

This gives a **lower bound** on the original min-max. Refer to its optimal $\mathbf{p}$ as $\mathbf{p}_1^*$. This time fix the

disturbance to $\mathbf{p}_1^*$:

$$\cancel{\min_{\mathbf{p}_1^*}} \max_{\mathbf{z}\in\{0,1\}^m} \max_{\mathbf{x}} \mathbf{c}'\mathbf{x} \tag{3.23}$$
$$\text{s.t.} \quad A\mathbf{x} \leq \mathbf{g}(\mathbf{z}, \mathbf{p}_1^*)$$

Refer to its optimal $\mathbf{z}$ as $\mathbf{z}_1^*$. Now fix the feasible set for $\mathbf{z}$ to $\{\mathbf{z}_0^*, \mathbf{z}_1^*\}$ and solve for $\mathbf{p}$:

$$\min_{\mathbf{p}\in\{0,1\}^n} \max_{\mathbf{z}\in\{\mathbf{z}_0^*,\mathbf{z}_1^*\}} \min_{\boldsymbol{h}} \mathbf{g}'(\mathbf{z}, \mathbf{p})\boldsymbol{h} \tag{3.24}$$
$$\text{s.t.} \quad A'\boldsymbol{h} = \mathbf{c}$$
$$\boldsymbol{h} \geq 0,$$

But the above can be written as a MILP; we can enumerate over $\mathbf{z}$ and write:

$$\min_{\ell,\mathbf{p},\boldsymbol{h}^0,\boldsymbol{h}^1} \ell$$
$$\text{s.t.} \quad \ell \geq \mathbf{g}'(\mathbf{z}_0^*, \mathbf{p})\boldsymbol{h}^0$$
$$A'\boldsymbol{h}^0 = \mathbf{c},$$
$$\boldsymbol{h}^0 \geq 0,$$
$$\ell \geq \mathbf{g}'(\mathbf{z}_1^*, \mathbf{p})\boldsymbol{h}^1$$
$$A'\boldsymbol{h}^1 = \mathbf{c}$$
$$\boldsymbol{h}^1 \geq 0$$
$$\mathbf{p} \in \{0,1\}^n$$

and if we derive a bound ($\overline{h}$) on h we can linearize ph by writing $0 \leq u \leq p\overline{h}$, $u \leq (1-p)\overline{h} + h$, $u \geq (p-1)\overline{h} + h$, so that $u \equiv ph$; thus, we are just solving a MILP again. This procedure shall be repeated iteratively as discussed in Example 1 above. Note that since the above master problem becomes more and more constrained as the number of iterations increases, the lower bound that it gives us on the original min-max problem is a non-decreasing function of the number of iterations.

3.3.1 Column and Row Generation for Solving the Min-Max Problem

Now, we formally establish how a column-and-row-generation scheme can be applied to our problem to obtain the globally optimal solution to problem (3.6), which entails deriving bounds on the dual variables in the innermost-level program in our problem. First, we turn Problem (3.6) into a large-scale single-level MILP. To do so, we start by separating the continuous variables in (3.3) from the integer ones. Let us denote the vector of integer decision variables in (3.3) by $\boldsymbol{\xi}$, the set of all $\boldsymbol{\xi}$ that maintain the feasibility of Problem (3.3) by $\mathcal{I}$, the vector of continuous decision variables in (3.3) by $\boldsymbol{\vartheta}$, and the set of the constraints in (3.3) that involve continuous variables (for a given $\boldsymbol{\xi}$) by $C(\boldsymbol{\xi})$; then we can write Problem (3.3) as:

$$\max_{\boldsymbol{\xi}\in\mathcal{I}} \max_{\boldsymbol{\vartheta}\in\mathcal{C}(\boldsymbol{\xi})} \sum_{i\in\mathcal{B}} b_i^d \tag{3.25}$$

and since the inner-level problem above is a linear program, we can write it in standard form:

$$\max_{\boldsymbol{\vartheta}\in\mathcal{C}(\boldsymbol{\xi})} \sum_{i\in\mathcal{B}} b_i^d \equiv \max_{C\boldsymbol{\vartheta}\leq \mathbf{c}_{\boldsymbol{\xi}}} \boldsymbol{f}'\boldsymbol{\vartheta} \tag{3.26}$$

where $\boldsymbol{f}'\boldsymbol{\vartheta} = \sum_{i\in\mathcal{B}} b_i^d$. Then we invoke strong duality as shown in the proposition below.

Proposition 2. *Strong duality holds for the inner-level problem in (3.25), and hence,*

$$\max_{\boldsymbol{\xi}\in\mathcal{I}} \max_{\boldsymbol{\vartheta}\in\mathcal{C}(\boldsymbol{\xi})} \sum_{i\in\mathcal{B}} b_i^d = \max_{\boldsymbol{\xi}\in\mathcal{I}} \min_{C'\boldsymbol{\lambda}_{\boldsymbol{\xi}}=\boldsymbol{f},\boldsymbol{\lambda}_{\boldsymbol{\xi}}\geq 0} \mathbf{c}'_{\boldsymbol{\xi}}\boldsymbol{\lambda}_{\boldsymbol{\xi}}. \tag{3.27}$$

Proof. Please see Section 3.5. □

Definition 7. *We define the set of all "bilinear terms" $P(g)$ of a multivariate polynomial $g(\cdot)$ as the set of all terms in $g(\cdot)$ in which strictly two variables have degree 1 and all other variables have degree 0. The set of all "bilinear forms" $R(g)$ of a multivariate polynomial $g(\cdot)$ is then obtained by dividing each element of $P(g)$ by its scalar coefficient (please see Example 2 below).*

Example 2. For instance, for $g(z, p_1, p_2, p_3, q_1, q_2) = z + 3zq_2 + 5p_1q_1 - 6p_2q_2 + p_3q_2z + q_2yp_1$, the set of all bilinear terms would be $P(g) = \{3zq_2, 5p_1q_1, -6p_2q_2, yp_1q_2\}$ and the set of all bilinear

forms would be $R(g) = \{zq_2, p_1q_1, q_2p_2, p_1q_2\}$. Note that bilinear forms that contain the same variables but different orders (e.g. p_1q_1 and q_1p_1) are considered identical.

Definition 8. *Consider two vectors* $\mathbf{p} = [p_1, p_2, \ldots, p_m]'$ *and* $\mathbf{q} = [q_1, q_2, \ldots, q_d]'$ *for some (finite)* $m, d \in \mathbb{Z}_+$. *Then, we define*

$$\mathcal{G}(g, \mathbf{p}, \mathbf{q}) \triangleq R(\mathcal{V}(\mathbf{p}, \mathbf{q})) \cap R(g(\mathbf{p}, \mathbf{q}))$$

where $\mathcal{V}(\mathbf{p}, \mathbf{q}) = \mathbf{1}'\mathbf{p}\mathbf{q}'\mathbf{1}$ *and* $g(\cdot)$ *is a multivariate polynomial (recall that* $R(\cdot)$ *is the set of bilinear forms). Also, denote by* F *a bijective map from* $\mathcal{G}(g, \mathbf{p}, \mathbf{q})$ *to a set* $\mathbf{J}$. *Then, we shall define the* substitution function $\mathcal{J}(g, \mathbf{p}, \mathbf{q}, \mathbf{J}, F)$ *as a function that replaces in* $g(\cdot)$ *every instance of each element of* $\mathcal{G}(g, \mathbf{p}, \mathbf{q})$ *with its corresponding element from* $\mathbf{J}$ *(as per map* F*); note that all the other terms in* $g(\cdot)$ *remain unchanged. Also,* $Y_{F,\mathbf{p}} : \mathbf{J} \to \{\mathbf{p}\}$ *is a surjective map that maps each element of* $\mathbf{J}$ *to its corresponding entry in* $\mathbf{p}$ *(as per the change of variables defined by map* F*); the surjective property is because each entry of* $\mathbf{p}$ *can appear in multiple different bilinear terms in* $g(\cdot)$ *(and hence in multiple elements of* $\mathcal{G}(g, \mathbf{p}, \mathbf{q})$*, and therefore, each entry of* $\mathbf{p}$ *can potentially correspond to more than one entry of* $\mathbf{J}$*; but the converse is not true because each element of* $\mathcal{G}(g, \mathbf{p}, \mathbf{q})$ *is a bilinear form that only contains one entry from* $\mathbf{p}$ *(multiplied by an entry from* $\mathbf{q}$*).*

Example 3. Given $g(z, p_1, p_2, p_3, q_1, q_2) = z + 3zq_2 + 5p_1q_1 - 6p_2q_2 + p_3q_2z + q_2yp_1$ as in Example 2, and assuming $\mathbf{p} = [p_1, p_2, p_3]'$ and $\mathbf{q} = [q_1, q_2]'$, we have

$$R(\mathcal{V}(\mathbf{p}, \mathbf{q})) = \{p_1q_1, p_1q_2, p_2q_1, p_2q_2, p_3q_1, p_3q_2\}$$

, and $\mathcal{G}(g, \mathbf{p}, \mathbf{q}) = \{p_1q_1, p_2q_2, p_1q_2\}$; then, by defining F such that $F(p_1q_1) = J_1$, $F(p_2q_2) = J_2$, $F(p_1q_2) = J_3$, we have $\mathcal{J}(g, \mathbf{p}, \mathbf{q}, \mathbf{J}, F) = z + 3zq_2 + 5J_1 - 6J_2 + p_3q_2z + yJ_3$, and for example, $Y_{F,\mathbf{p}}(J_1) = p_1$, $Y_{F,\mathbf{q}}(J_1) = q_1$, $Y_{F,\mathbf{p}}(J_3) = p_1$, and $Y_{F,\mathbf{q}}(J_2) = q_2$.

Next, by deriving a bound on the dual variables (i.e. $\boldsymbol{\lambda}_{\boldsymbol{\xi}}$) and enumerating over all elements of $\mathcal{I}$, we reformulate Problem (3.6) as a large-scale MILP as shown in the theorem below. Before proceeding to the statement of the theorem, recall from Section 1.3 that with a slight abuse of

notation, for a vector $\mathbf{x}$, we denote by $\{\mathbf{x}\}$ the set of all entries of $\mathbf{x}$.

Theorem 4. *If Λ^* is the globally optimal value of the objective function in the optimization problem (3.6), then*

$$\begin{aligned}
\Lambda^* = \min_{\eta} \ & \gamma \\
s.t. \ & \gamma \geq \mathcal{W}_{\boldsymbol{\xi}} \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& \boldsymbol{\lambda}_{\boldsymbol{\xi}} \geq 0 \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& C' \boldsymbol{\lambda}_{\boldsymbol{\xi}} = \boldsymbol{f} \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& 0 \leq \boldsymbol{\Psi}_{\boldsymbol{\xi}} \leq Y_{F,\mathbf{s^T}}(\boldsymbol{\Psi}_{\boldsymbol{\xi}}) V \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& \boldsymbol{\Psi}_{\boldsymbol{\xi}} \leq (1 - Y_{F,\mathbf{s^T}}(\boldsymbol{\Psi}_{\boldsymbol{\xi}})) V + Y_{F,\boldsymbol{\lambda}_{\boldsymbol{\xi}}}(\boldsymbol{\Psi}_{\boldsymbol{\xi}}) \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& \boldsymbol{\Psi}_{\boldsymbol{\xi}} \geq (Y_{F,\mathbf{s^T}}(\boldsymbol{\Psi}_{\boldsymbol{\xi}}) - 1) V + Y_{F,\boldsymbol{\lambda}_{\boldsymbol{\xi}}}(\boldsymbol{\Psi}_{\boldsymbol{\xi}}) \, \forall \boldsymbol{\xi} \in \mathcal{I} \\
& A \in \mathcal{A}
\end{aligned} \tag{3.28}$$

where $\mathcal{W}_{\boldsymbol{\xi}} = \mathcal{J}(Z, \mathbf{s}^T, \boldsymbol{\lambda}_{\boldsymbol{\xi}}, \{\boldsymbol{\Psi}_{\boldsymbol{\xi}}\}, F)$ for $Z(\mathbf{s}^T, \boldsymbol{\lambda}_{\boldsymbol{\xi}}) \triangleq \mathbf{c}'_{\boldsymbol{\xi}} \boldsymbol{\lambda}_{\boldsymbol{\xi}}$, $\boldsymbol{\Psi}_{\boldsymbol{\xi}} \in \mathbb{R}^{|(\mathcal{G}(Z, \mathbf{s}^T, \boldsymbol{\lambda}_{\boldsymbol{\xi}})|}$, $\eta = \{\gamma, A\} \cup (\cup_{\boldsymbol{\xi} \in \mathcal{I}} (\{\boldsymbol{\lambda}_{\boldsymbol{\xi}}\} \cup \{\boldsymbol{\Psi}_{\boldsymbol{\xi}}\}))$, F is a bijective map (as defined in Definition 8) from $\mathcal{G}(Z, \mathbf{s}^T, \boldsymbol{\lambda}_{\boldsymbol{\xi}})$ to $\{\boldsymbol{\Psi}_{\xi}\}$, and $V = \max(1, \sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$.

Proof. Please see Section 3.5. □

Now that we have turned Problem (3.6) into a large-scale single-level MILP, it can be solved to global optimality using column and row generation ([62] and [63]). To do so, we define a "decomposed master problem with input Q", which has the exact same form as (3.28) except that $\mathcal{I}$ shall be replaced by Q throughout (3.28). Note that $Q \subseteq \mathcal{I}$ in the decomposition process, and hence, the term "decomposed". The column-and-row-generation procedure (adopted from [64]), tailored to the global optimization of Problem (3.6), is outlined in Algorithm 1.

Using the results presented in [62] and [63] on the convergence of the column-and-row generation procedure (referred to as column-and-"constraint" generation in those papers), and by invoking strong duality in our reformulation for the continuous variables of the inner-level MILP and the

Algorithm 1 Column and Row Generation Procedure for Problem (3.6)

1: Initialize $\mathbf{s}^{T*}$ by setting the entries corresponding to up to k lines to zero, and setting all its other entries to 1.
2: Solve Problem (3.3) with $\mathbf{s}^{T*}$ as input
3: $Q \leftarrow$ vector of the optimal values of all integer variables in Step 2
4: Initialize $U \leftarrow \infty$, $L \leftarrow -\infty$
5: **while** $U - L >$ tolerance **do**
6: Solve the decomposed master problem with input Q and update $\mathbf{s}^{T*}$
7: $M^* \leftarrow$ optimal value of objective function in Step 6
8: Solve Problem (3.3) with $\mathbf{s}^{T*}$ as input
9: $S^* \leftarrow$ optimal objective function value in Step 8
10: $\mathcal{Z}^* \leftarrow$ vector of optimal values of all integer variables in Step 8
11: $Q \leftarrow \{\mathcal{Z}^*\} \cup Q$
12: $U \leftarrow \min(S^*, U)$
13: $L \leftarrow M^*$
14: **end while**

bounds we derived in Theorem 4 which gave us a valid single-level reformulation for forming the master problem, we have the following theorem.

Theorem 5. *Algorithm 1 converges in finite steps to a globally optimal solution of Problem (3.6).*

Proof. Please see Section 3.5. □

In practice, it would be very helpful to initialize the above algorithm with some thought, in an effort to achieve faster convergence. For instance, picking the initial disturbance ($\mathbf{s}^{T*}$) such that it removes k lines from a capacity-bottleneck of the system that separates a demand-rich part of the system from a generation-rich part of the system (if such bottleneck exists in the system of interest), generally appears to be a good starting point.

3.4 Maximum Efficacy in Reducing the Post-Disturbance Load Shed among the Set of All Possible N-k Scenarios

So far, we have assumed that the adversary picks its disturbance strategically with full awareness of the SO's post-disturbance problem and thus has the knowledge that a post-contingency reactance tweaking mechanism is in place. This worst-case-scenario approach for measuring the efficacy of reactance tweaking may not always be realistic, especially when the adversary is not strategic, as it may considerably underrate the efficacy of the post-disturbance control mechanism (this will become more evident in our numerical experiments in Section 3.6). Hence, in order to obtain a less extreme measure of efficacy, we introduce a new notion of efficacy that captures the capability of reactance tweaking in dealing with a broader range of disturbances.

Definition 9. *The* general efficacy *($E(\kappa)$) of post-disturbance reactance tweaking in response to a given disturbance κ is defined as the total savings (reduction) in the amount of load shed post-disturbance purely as the contribution of reactance tweaking. In other words, if $R(\kappa)$ denotes the amount of load satisfied post-disturbance for κ while using reactance tweaking as a control mechanism, and $Q(\kappa)$ denotes the amount of load satisfied post-disturbance without using reactance tweaking as a control mechanism for the same disturbance (i.e. κ), then $E(\kappa) = R(\kappa) - Q(\kappa)$.*

Definition 10. *The* constant-reactance post-contingency control *problem is the yield maximization problem in which the SO has no control over line reactances, and can be formulated as the*

following LP:

$$\begin{aligned}
\max_{d} \quad & \sum_{i\in\mathcal{B}} b_i^d && (3.29)\\
s.t. \quad & I_{ijh}^{T+1} x_{ijh}^o + \Upsilon(s_{ijh}^T - 1) \le (\theta_i^{T+1} - \theta_j^{T+1}),\ \forall (i,j,h) \in \mathcal{E} \\
& I_{ijh}^{T+1} x_{ijh}^o + \Upsilon(1 - s_{ijh}^T) \ge (\theta_i^{T+1} - \theta_j^{T+1}),\ \forall (i,j,h) \in \mathcal{E} \\
& b_i^g - b_i^d = \sum_{(i,j,h)\in\mathcal{E}_i} I_{ijh}^{T+1},\ \forall i \in \mathcal{B} \\
& -\overline{cap}_{ij}^a s_{ijh}^T \le I_{ijh}^{T+1} \le \overline{cap}_{ij}^a s_{ijh}^T,\ \forall (i,j,h) \in \mathcal{E} \\
& 0 \quad \le b_i^g \quad \le \overline{b}_i^g, \quad \forall i \in \mathcal{B} \\
& 0 \quad \le b_i^d \quad \le \overline{b}_i^d, \quad \forall i \in \mathcal{B} \\
& \theta_i^{T+1}, I_{ijh}^{T+1}, b_i^g, b_i^d \in \mathbb{R}, \forall i \in \mathcal{B}, (i,j,h) \in \mathcal{E}
\end{aligned}$$

where $d = \{\boldsymbol{\theta}^{T+1}, \mathbf{I}^{T+1}, \mathbf{b}^g, \mathbf{b}^d\}$. *We shall write this LP in standard form as*

$$\max_{W\boldsymbol{\varpi}\le\boldsymbol{\nu}} \boldsymbol{p}'\boldsymbol{\varpi} \tag{3.30}$$

Theorem 6. *For a given k, let us denote the set of all "$N-k$" disturbances by $\mathcal{K}(k)$. For a given vector of bus injections* $\mathbf{b}^o$*, the general efficacy of post-disturbance reactance tweaking in response to any $\kappa \in \mathcal{K}(k)$ is upper bounded as follows:*

$$\begin{aligned}
E(\kappa) \le \max_{\Pi} & \sum_{i\in\mathcal{B}} b_i^d - \mathcal{J}(\mathcal{U}, \mathbf{s}^T, \boldsymbol{\pi}, \{\boldsymbol{\Theta}\}, F) \\
s.t. \quad & \boldsymbol{\pi} \ge 0 \\
& W'\boldsymbol{\pi} = \boldsymbol{p} \\
& 0 \le \boldsymbol{\Theta} \le Y_{F,\mathbf{s}^T}(\mathbf{s}^T) V \\
& \boldsymbol{\Theta} \le (1 - Y_{F,\mathbf{s}^T}(\boldsymbol{\Theta})) V + Y_{F,\boldsymbol{\pi}}(\boldsymbol{\Theta}) \\
& \boldsymbol{\Theta} \ge (Y_{F,\mathbf{s}^T}(\boldsymbol{\Theta}) - 1) V + Y_{F,\boldsymbol{\pi}}(\boldsymbol{\Theta}) \\
& \omega \in \Omega; \quad a \in \tilde{a}
\end{aligned}$$

and this upper bound is tight. Here, $\Pi = \{A, \omega, \boldsymbol{\pi}, \Theta\}$, $\mathcal{U}(\mathrm{s}^T, \boldsymbol{\pi}) \triangleq \boldsymbol{\nu}'\boldsymbol{\pi}$, F *is a bijective map (as defined in Definition 8) from* $\mathcal{G}(\mathcal{U}, \mathrm{s}^T, \boldsymbol{\pi})$ *to* $\{\Theta\}$, $\Theta \in \mathbb{R}^{|(\mathcal{G}(\mathcal{U}, \mathrm{s}^T, \boldsymbol{\pi})|}$, *and* $V = \max(1, \sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$.

3.5 Proofs

Before we present the proofs, let us define some core concepts that we shall use in this section.

Definition 11. *The* weighted Laplacian *of a graph (also known as the* conductance matrix*) is defined as*

$$\overline{G} = M\tilde{C}M' \tag{3.31}$$

where M *is the incidence matrix, and* $\tilde{C}$ *is a diagonal matrix whose diagonal elements are the line conductances (conductance is the reciprocal of resistance).*

Definition 12 (see e.g. [34])**.** *The* resistance distance *(also known as* effective resistance*) between two arbitrary nodes* i *and* j*, which we shall denote by* $\overline{R}_{ij}$*, is the voltage difference across the buses* i *and* j *in a connected network when a unit current source is injected at* i *and collected at* j*. One way to compute resistance distances is to use the pseudo-inverse of the weighted Laplacian (denoted by* $\overline{G}^\dagger$*) as shown below (see e.g. [34] for details):*

$$\overline{R}_{ij} = \overline{G}^\dagger_{ii} + \overline{G}^\dagger_{jj} - 2\overline{G}^\dagger_{ij} \tag{3.32}$$

where $\overline{G}^\dagger_{ij}$ *denotes the* i, j *entry of* $\overline{G}^\dagger$*, and* $\overline{G}^\dagger$ *is computed as*

$$\overline{G}^\dagger = (\overline{G} + \mathbf{1}\mathbf{1}^t/|\mathcal{B}|)^{-1} - \mathbf{1}\mathbf{1}^t/|\mathcal{B}|. \tag{3.33}$$

Next, we establish two lemmas that are standard results in the literature.

Lemma 2 (see e.g. [34])**.** *Resistance distances satisfy the triangle equality:*

$$\overline{R}_{ik} \leq \overline{R}_{ij} + \overline{R}_{jk}.$$

Lemma 3 (see e.g. [31]). *If we denote the current flow in line (i, j, h) by I_{ijh}, then*

$$I_{ijh} = \frac{1}{2r_{ijh}} \sum_{m=1}^{|\mathcal{B}|} b_m(\overline{R}_{jm} - \overline{R}_{im})$$

where r_{ijh} is the resistance of line (i, j, h) and b_m is the net injection at bus m.

Proof of Theorem 2. Let us denote by r_{ijh} the resistance of the line (i, j, h) (note that r_{ijh} is equivalent to x_{ijh} used in the statement of this theorem). From Kirchhoff's Voltage Law, we have $\theta_i - \theta_j = I_{ijh} r_{ijh}$; thus, we can write the result in Lemma 3 above as

$$\theta_i - \theta_j = \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} b_m(\overline{R}_{jm} - \overline{R}_{im})$$

We have

$$\theta_i - \theta_j = \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} b_m(\overline{R}_{jm} - \overline{R}_{im}) \leq \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} |b_m(\overline{R}_{jm} - \overline{R}_{im})|$$

but

$$|b_m(\overline{R}_{jm} - \overline{R}_{im})| \leq |b_m| \, |\overline{R}_{jm} - \overline{R}_{im}|$$

and using the triangle inequality in Lemma 2 above (and keeping in mind that $\overline{R}_{ij} = \overline{R}_{ji}$), we have,

$$|\overline{R}_{jm} - \overline{R}_{im}| \leq \overline{R}_{ij}$$

Thus,

$$\frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} |b_m| \, |\overline{R}_{jm} - \overline{R}_{im}| \leq \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} |b_m| \overline{R}_{ij}$$

But $\overline{R}_{ij} \leq r_{ijh}$ (because $1/\overline{R}_{ij} = 1/r_{ijh} + 1/A$, for some $A \in \{\overline{\mathbb{R}}_+ \cup \infty\}$, and both $\overline{R}_{ij}$ and r_{ijh} are positive) and hence,

$$\theta_i - \theta_j = \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} b_m(\overline{R}_{jm} - \overline{R}_{im}) \leq \frac{1}{2} \sum_{m=1}^{|\mathcal{B}|} |b_m| r_{ijh}$$

If we denote $G \triangleq \sum_{i\in\mathcal{B}} b_i \mathbb{1}_{\overline{\mathbb{R}}_+}(b_i)$, then we have $\sum_{m=1}^{|\mathcal{B}|} |b_m| = 2G$, and hence can write the above as

$$\theta_i - \theta_j = \frac{1}{2}\sum_{m=1}^{|\mathcal{B}|} b_m(\overline{R}_{jm} - \overline{R}_{im}) \leq G r_{ijh}$$

Now take two arbitrary nodes p and q; since the network is assumed to be connected, a path has to exist between p and q. Let us denote the sequence of nodes on this path by $p, v_1, v_2, ..., v_L, q$. Then, we have:

$$\theta_p - \theta_q = (\theta_p - \theta_{v_1}) + (\theta_{v_1} - \theta_{v_2}) + \cdots + (\theta_{v_L} - \theta_q) \leq G(\overline{r}_{pv_1} + \overline{r}_{v_1v_2} + \cdots + \overline{r}_{qv_L}) \leq G \sum_{(i,j,h)\in\mathcal{E}} r_{ijh} \tag{3.34}$$

where $\overline{r}_{v_1v_2}$ (for instance) denotes the resistance of any arbitrary line between nodes v_1 and v_2. Thus, the upper bound on the voltage difference between any pair of nodes is maximized when the path between them is the longest possible. Hence, if we have $|\mathcal{B}|$ nodes and fix the voltage of some node to zero, then the maximum possible value of $|\theta_i|$ (i.e. voltage magnitude) at any node would be

$$|\theta_B| \leq G \sum_{(i,j,h)\in\mathcal{E}} r_{ijh} \;\; \forall B \in \mathcal{B}$$

It follows that this upper bound is tight and is achieved by a line graph with a positive injection of magnitude G at one end and a negative injection of the same magnitude at its other end, with all other nodes having a zero injection. Note that since in the statement of this theorem we are using the symbol for reactance instead of resistance (based on the DC approximation to AC power flows), the above inequality is written using x_{ijh} instead of r_{ijh}. □

Proof of Theorem 3. Consider the first constraint in Problem (3.2), and let us replace the term $I_{ijh}^{T+1} x_{ijh}^{T+1}$ in that constraint by μ_{ijh}; then, given the fact that the node voltage difference between two connected nodes (i.e. two nodes that are connected to each other via a path of active lines) is upper-bounded by Υ (as established in (3.34) in the proof of Theorem 2) and considering that s_{ijh}^T

is a binary constant, we can rewrite the first constraint in Problem (3.2) as

$$\mu_{ijh} + \Upsilon(s^T_{ijh} - 1) \leq \theta^{T+1}_i - \theta^{T+1}_j; \tag{3.35}$$

$$\mu_{ijh} + \Upsilon(1 - s^T_{ijh}) \geq \theta^{T+1}_i - \theta^{T+1}_j \tag{3.36}$$

Now, let us consider the third and fourth constraints in Problem (3.2). The goal of the third and the fourth constraints in Problem (3.2) is to establish that $x^{T+1}_{ijh} = x^o_{ijh}$ unless $\tau_{ijh} = 1$, in which case $x^{\min}_{ijh} \leq x^{T+1}_{ijh} \leq x^{\max}_{ijh}$. Using the change of variables above, and given the fact that $\mu_{ijh} = I^{T+1}_{ijh} x^{T+1}_{ijh} = I^{T+1}_{ijh} x^o_{ijh}$ unless $\tau_{ijh} = 1$, we can write

$$\boldsymbol{\mu} - diag(\boldsymbol{x}^o)\boldsymbol{I}^{T+1} - 2\Upsilon\boldsymbol{\tau} \leq 0; \tag{3.37}$$

$$-\boldsymbol{\mu} + diag(\boldsymbol{x}^o)\boldsymbol{I}^{T+1} - 2\Upsilon\boldsymbol{\tau} \leq 0; \tag{3.38}$$

so that whenever $\tau_{ijh} = 0$ we get $\mu_{ijh} = I^{T+1}_{ijh} x^o_{ijh}$ (note that the above inequalities follow from the fact that for any line (i, j, h) both $|\mu_{ijh}|$ and $|x^o_{ijh} I^{T+1}_{ijh}|$ are upper bounded by Υ, and hence, when $\tau_{ijh} = 1$ the above inequalities become redundant). Since $x^o_{ijh} \in [x^{\min}_{ijh} \;\; x^{\max}_{ijh}]$ by assumption, we can take care of the case of $\tau_{ijh} = 1$ by writing

$$x^{\min}_{ijh} |I^{T+1}_{ijh}| \leq |\mu_{ijh}| \leq x^{\max}_{ijh} |I^{T+1}_{ijh}|$$

so that x^{T+1}_{ijh} will have the freedom to range between $x^{\min}_{ijh}$ and $x^{\max}_{ijh}$ whenever $\tau_{ijh} = 1$ (recall that $\mu_{ijh} = I^{T+1}_{ijh} x^{T+1}_{ijh}$ by definition, and hence, $x^{\min}_{ijh} |I^{T+1}_{ijh}| \leq |\mu_{ijh}| \leq x^{\max}_{ijh} |I^{T+1}_{ijh}|$ is equivalent to $x^{\min}_{ijh} \leq x^{T+1}_{ijh} \leq x^{\max}_{ijh}$ whenever $I^{T+1}_{ijh} \neq 0$). The fifth constraint in the statement of the theorem is then obtained by using the change of variables $\delta_{ijh} \triangleq |\mu_{ijh}|$ and $\alpha_{ijh} \triangleq |I^{T+1}_{ijh}|$. It remains to linearize $|\mu_{ijh}|$ and $|I^{T+1}_{ijh}|$; the sixth, seventh, eighth, ninth, and tenth constraints in (3.3) (in the statement of the theorem) all correspond to the constraints and auxiliary variables required for linearizing both $|\mu_{ijh}|$ and $|I^{T+1}_{ijh}|$ as per the technique described in Lemma 1 for linearizing absolute values. □

Proof of Proposition 2. To verify that strong duality holds, note that the inner-level problem in (3.25) is bounded and feasible (because setting all decision variables to zero is a feasible solution); hence, strong duality holds and deriving the equations for the dual problem stated in the proposition is standard (e.g. see [8] for details on deriving the dual of a linear program). □

Proof of Theorem 4. Note that so far we have reformulated the original min-max problem (3.6) as a bilevel program in which the inner-level and the outer-level are both MILPs coupled through only a binary variable (i.e. through s^T). Using the result we just established in Proposition 2, we can write (3.6) as:

$$\min_{a \in \tilde{a}} \max_{\boldsymbol{\xi} \in \mathcal{I}} \min_{C'\boldsymbol{\lambda}_{\boldsymbol{\xi}} = \boldsymbol{f}, \boldsymbol{\lambda}_{\boldsymbol{\xi}} \geq 0} \mathbf{c}'_{\boldsymbol{\xi}} \boldsymbol{\lambda}_{\boldsymbol{\xi}}. \tag{3.39}$$

Next, we create a copy of $\boldsymbol{\lambda}$ for each element $\boldsymbol{\xi} \in \mathcal{I}$ (each copy will be treated as a distinct decision variable) and denote it by $\boldsymbol{\lambda}_{\boldsymbol{\xi}}$. Then, we enumerate over all elements of $\mathcal{I}$ (to eliminate the maximization level (i.e. the mid-level) of the above three-level problem), merge the two minimization levels, and rewrite the above as a large-scale single-level program (note that this step is a standard technique, see e.g. [64]):

$$\begin{aligned} \min_{\eta} \quad & \gamma \\ \text{s.t.} \quad & \gamma \geq \mathbf{c}'_{\boldsymbol{\xi}} \boldsymbol{\lambda}_{\boldsymbol{\xi}} \; \forall \boldsymbol{\xi} \in \mathcal{I} \\ & \boldsymbol{\lambda}_{\boldsymbol{\xi}} \geq 0 \; \forall \boldsymbol{\xi} \in \mathcal{I} \\ & C'\boldsymbol{\lambda}_{\boldsymbol{\xi}} = \boldsymbol{f} \; \forall \boldsymbol{\xi} \in \mathcal{I} \\ & a \in \tilde{a} \end{aligned} \tag{3.40}$$

where $\eta = \{\gamma, a\} \cup (\cup_{\boldsymbol{\xi} \in \mathcal{I}} \{\boldsymbol{\lambda}_{\boldsymbol{\xi}}\})$. However, this single-level program is not a MILP yet, because in the constraints $\gamma \geq \mathbf{c}'_{\boldsymbol{\xi}} \boldsymbol{\lambda}_{\boldsymbol{\xi}}$ we have some binary×continuous bilinear terms stemming from the product of some elements of s^T and some elements of $\boldsymbol{\lambda}_{\boldsymbol{\xi}}$ for all $\boldsymbol{\xi} \in \mathcal{I}$. To do so, we need a bound on the dual variables (i.e. $\boldsymbol{\lambda}_{\boldsymbol{\xi}}$) of our innermost level LP in (3.39). We claim that this upper bound

(which we denote by V) is $V = \max(1, \sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$; to verify this claim, observe that the LP in the innermost level of (3.39) maximizes the satisfied load post-disturbance; by definition, the dual variable corresponding to a constraint quantifies the amount of change in the optimal value of the objective function if we increase the bound in that constraint by one unit. With that in mind, if we look at any constraint in Problem (3.2) other than the constraint on the maximum nodal demands, increasing the limit of any other constraint in the problem can result in at most $\sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o)$ because no matter what, the objective only depends on the nodal demands, and the total maximum nodal demand is itself upper bounded by $\sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$. On the other hand, increasing the limit on each constraint that enforces maximum nodal demands can increase the objective funcion by at most one unit, and hence, $V = \max(1, \sum_{i \in \mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$. Now that we have a bound on the continuous variables that appear in bilinear terms, we can replace any bilinear term by an auxiliary variable (we denote the vector of these auxiliary variables by $\mathbf{\Psi}_{\boldsymbol{\xi}}$), and to complete the linearization and ensure that this auxiliary variable will be equal to the bilinear term, we further enforce:

$$
\begin{aligned}
&0 \leq \mathbf{\Psi}_{\boldsymbol{\xi}} \leq Y_{F,\mathbf{s^T}}(\mathbf{\Psi}_{\boldsymbol{\xi}}) V \ \forall \boldsymbol{\xi} \in \mathcal{I} \\
&\mathbf{\Psi}_{\boldsymbol{\xi}} \leq (1 - Y_{F,\mathbf{s^T}}(\mathbf{\Psi}_{\boldsymbol{\xi}}))V + Y_{F,\boldsymbol{\lambda}_{\boldsymbol{\xi}}}(\mathbf{\Psi}_{\boldsymbol{\xi}}) \ \forall \boldsymbol{\xi} \in \mathcal{I} \\
&\mathbf{\Psi}_{\boldsymbol{\xi}} \geq (Y_{F,\mathbf{s^T}}(\mathbf{\Psi}_{\boldsymbol{\xi}}) - 1)V + Y_{F,\boldsymbol{\lambda}_{\boldsymbol{\xi}}}(\mathbf{\Psi}_{\boldsymbol{\xi}}) \ \forall \boldsymbol{\xi} \in \mathcal{I}
\end{aligned}
$$

where $Y_{\cdot,\cdot}$ is defined in Definition 8 and F is a bijective map (as defined in Definition 8) from $\mathcal{G}(Z, \mathbf{s}^T, \boldsymbol{\lambda}_{\boldsymbol{\xi}})$ to $\{\mathbf{\Psi}_{\xi}\}$. □

Proof of Theorem 5. Since the Master Problem has a less-constrained feasible set than the single-level MILP reformulation of the original min-max problem but the same objective, it gives a lower bound on the globally optimal solution of the original problem. Conversely, the optimal value of the subproblem gives an upper bound on the optimal value of the original min-max problem. Given strong duality in our reformulation for the continuous variables of the inner-level MILP (i.e, Proposition 2), our single-level MILP reformulation for forming the master problem through the bounds we derived in Thorem 4, and the finiteness of the number of elements in $\mathcal{I}$, it only

remains to clarify that any repeated value of $\mathcal{Z}^*$ implies global optimality has been achieved (i.e. it only remains to show that this procedure will not cycle between non-globally-optimal solutions). Proving this is standard and has been presented in [62], [63], [64]; we shall repeat their proof technique here for completeness: To see this, note that if at iteration k the value of $\mathcal{Z}^*$ obtained from solving the subproblem had already been encountered in some previous iteration t (i.e. $t < k$), it means that the set of dual variables and constraints corresponding to $\mathcal{Z}^*$ already exist in the master problem at iteration k, and hence, the master problem stays the same at $k+1$ and the optimal objective value of the Master Problem at iteration $k+1$ will remain the same as its optimal objective value at iteration k. Note that by strong duality, at iteration $k+1$, we have $UpperBound \leq \sum_{i \in \mathcal{B}} b_i^{d*} = \min_{\boldsymbol{\lambda}_{\mathcal{Z}^*}} \mathbf{c}'(\boldsymbol{\lambda}_{\mathcal{Z}^*}, \mathbf{s}^{T*})\boldsymbol{\lambda}_{\mathcal{Z}^*}$; but we also have that $LowerBound = \gamma^* \geq \min_{\boldsymbol{\lambda}_{\mathcal{Z}^*}} \mathbf{c}'(\boldsymbol{\lambda}_{\mathcal{Z}^*}, \mathbf{s}^{T*})\boldsymbol{\lambda}_{\mathcal{Z}^*}$ from the master problem; thus, the lower and upper bounds are equal and we have reached the globally optimal solution. □

Proof of Theorem 6. Since the disturbance defined by the vector of decision variables a and their feasible set $\tilde{a}$ correspond to the set of all $N-k$ disturbances for a given k, we have

$$E(a) = R(a) - Q(a) \leq \max_{a \in \tilde{a}} \; R(a) - Q(a)$$

but by definition, $R(a)$ is the optimal solution to Problem (3.2) and $Q(a)$ is the optimal solution to Problem (3.29). Therefore, the above is equivalent to

$$E(a) = R(a) - Q(a) \leq \max_{a \in \tilde{a}} \; [\max_{\omega \in \Omega} \; \sum_{i \in \mathcal{B}} b_i^d - \max_{W\boldsymbol{\varpi} \leq \boldsymbol{\nu}} \boldsymbol{p}'\boldsymbol{\varpi}]$$

Using the same argument as in Proposition 2, we can show that strong duality holds for $\max_{W\boldsymbol{\varpi} \leq \boldsymbol{\nu}} \boldsymbol{p}'\boldsymbol{\varpi}$.

We shall denote the vector of dual variables of this problem by π. Then, we have

$$\begin{aligned}
\max_{a\in\tilde{a}}\ [\max_{\omega\in\Omega}\ \sum_{i\in\mathcal{B}} b_i^d - \max_{W\varpi\leq\nu} p'\varpi] &= \max_{a\in\tilde{a}}\ [\max_{\omega\in\Omega}\ \sum_{i\in\mathcal{B}} b_i^d - \min_{W'\pi=p,\pi\geq 0} \nu'\pi] \\
&= \max_{a\in\tilde{a}}\ [\max_{\omega\in\Omega}\ \sum_{i\in\mathcal{B}} b_i^d + \max_{W'\pi=p,\pi\geq 0} -\nu'\pi] \\
&= \max_{a\in\tilde{a}}\ [\max_{\omega\in\Omega, W'\pi=p,\pi\geq 0} (-\nu'\pi + \sum_{i\in\mathcal{B}} b_i^d)] \\
&= \max_{a\in\tilde{a},\omega\in\Omega, W'\pi=p,\pi\geq 0} (-\nu'\pi + \sum_{i\in\mathcal{B}} b_i^d)
\end{aligned} \tag{3.41}$$

However, this single-level program is not a MILP yet, because in the objective the term $\nu'\pi$ contains some binary$\times$continuous bilinear terms stemming from the product of some elements of s^T and some elements of π. To do so, we need a bound on the dual variables (i.e. π) of the LP corresponding to Problem (3.29). We claim that this upper bound (which we denote by V) is $V = \max(1, \sum_{i\in\mathcal{B}} b_i^o \mathbb{1}_{\mathbb{R}_+}(b_i^o))$ and the proof of this claim follows the exact same procedure as that of the bound (V) derived in the proof of Theorem 4, and is hence omitted. Now that we have a bound on the continuous variables that appear in bilinear terms, we can replace any bilinear term by an auxiliary variable (we denote the vector of these auxiliary variables by Θ), and to complete the linearization and ensure that this auxiliary variable will be equal to the bilinear term we further enforce:

$$\begin{aligned}
&0 \leq \Theta \leq Y_{F,\mathrm{s}^T}(\Theta)V \\
&\Theta \leq (1 - Y_{F,\mathrm{s}^T}(\Theta))V + Y_{F,\pi}(\Theta) \\
&\Theta \geq (Y_{F,\mathrm{s}^T}(\Theta) - 1)V + Y_{F,\pi}(\Theta)
\end{aligned}$$

where $Y_{.,.}$ is defined in Definition 8 and F is a bijective map (as defined in Definition 8) from $\mathcal{G}(\mathcal{U}, \mathrm{s}^T, \pi)$ to $\{\Theta\}$. □

3.6 Numerical Experiments

In this section, we shall present the results of our numerical experiments on the IEEE One-Area RTS-1996 test case system [36], which has 24 buses (11 generators and 16 buses with load) and 38 lines. The numerical experiments were performed in MATLAB, and Gurobi [37] was used for solving the MILPs. The operating range of the reactance tweaking device has been set to the operating range reported for TCSC in [33].

3.6.1 Efficacy in the Event of a Worst-Case-Scenario Disturbance

In the first set of numerical experiments, we focus on the worst-case-scenario efficacy of reactance tweaking in the RTS-1996 test system. In these experiments, we set $k \in \{2, 3\}$ and $H = 3$ (i.e. we allow tweaking the reactances of up to only three lines), and use a base load flow scenario adopted from the dataset for the IEEE One-Area RTS-1996 test case system [36]. At this load level, prior to the disturbance the system is serving 2850MW of load in total. Table 3.2 displays the results for this set of experiments (note that in this table, κ^* denotes the worst-case-scenario disturbance obtained from solving Problem 3.6).

The results in Table 3.2 suggest that reactance tweaking on up to three transmission lines cannot increase the resilience of the RTS-1996 system to a strategic adversary that seeks to cause the worst-case-scenario disturbance by removing two or three lines with full knowledge of the SO's post-disturbance response. Nevertheless, an important point to consider here is that reactance tweaking is meant for redistributing flows so as to prevent overloading of further lines, but it cannot increase the transmission capacity of the network. For example, when of the three lines that used to be connected to a load bus two of them have been removed by a disturbance, no matter how much we tweak the reactances, the maximum load that can be satisfied at that bus would be equal to the transmission capacity of the remaining line that is still connected to it; hence, a strategic adversary can take advantage of this fact and possibly choose a disturbance that thwarts

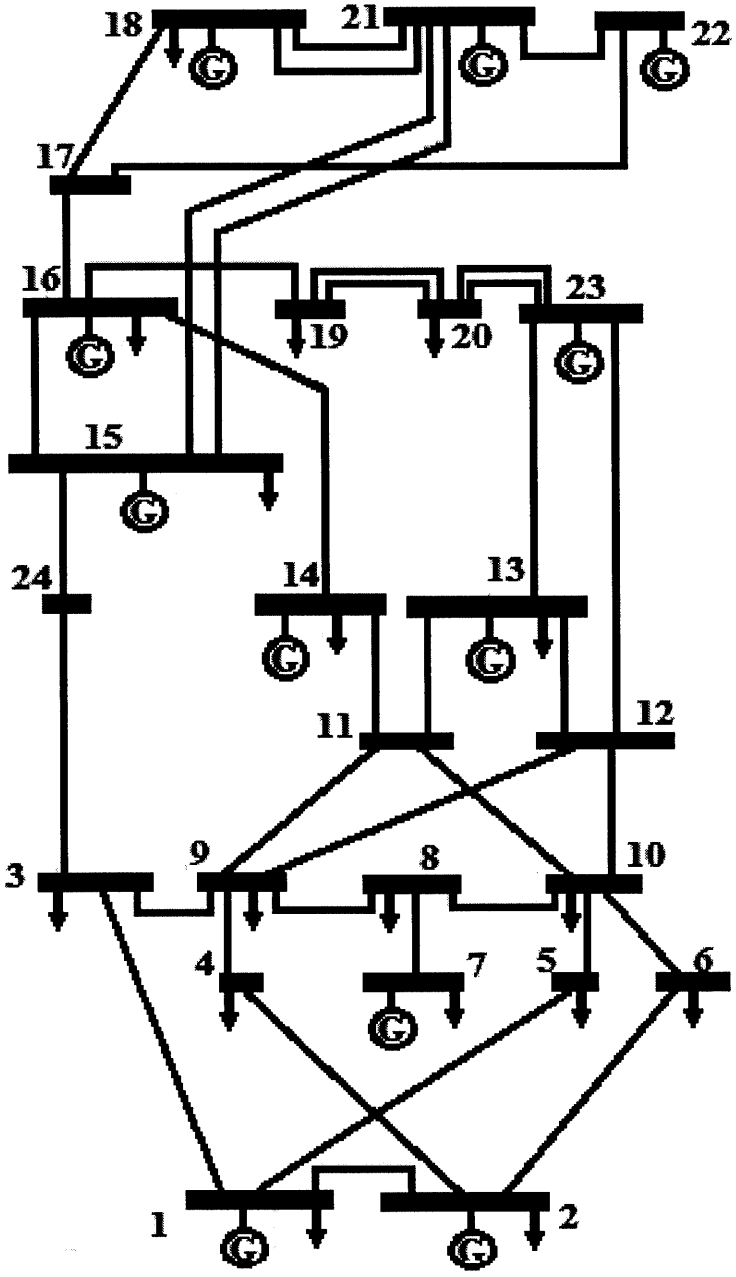


Figure 3-2: Diagram of the IEEE One-Area RTS-1996 Test System (based on the diagrams provided in [36] and [38]). Each Ⓖ denotes a generator and each ↓ denotes demand; the number written next to each bus indicates the index of the bus.

Table 3.1: Indexing the Transmission Lines for the One-Area RTS-1996 Test System (Based on Branch Data Provided in [36])

Line Index	Start Bus Index	End Bus Index
1	1	2
2	1	3
3	1	5
4	2	4
5	2	6
6	3	9
7	3	24
8	4	9
9	5	10
10	6	10
11	7	8
12	8	9
13	8	10
14	9	11
15	9	12
16	10	11
17	10	12
18	11	13
19	11	14
20	12	13
21	12	23
22	13	23
23	14	16
24	15	16
25	15	21
26	15	21
27	15	24
28	16	17
29	16	19
30	17	18
31	17	22
32	18	21
33	18	21
34	19	20
35	19	20
36	20	23
37	20	23
38	21	22

Table 3.2: Worst-Case-Scenario Efficacy (for $H = 3$)

k	T	$E(\kappa^*)$	% Load Satisfied by SO's Control
2	0	0	93.19 %
2	1	0	92.56%
3	0	0	89.16%
3	1	0	70.46%

any contribution from reactance tweaking. In our numerical experiments above, for instance, when k=3 and T=0, the adversary chooses to remove lines 29, 36, and 37, which results in the islanding of demand buses 19 and 20, and hence all the load at those two buses are inevitably lost and reactance tweaking will clearly not be able to make any difference; also, when k=3 and T=1, the adversary takes out lines 20, 21, and 23 at t=0, which leads to the tripping of lines 18, 6, and 7 as the cascade propagates at t=1; with the failure of these lines, the lower (demand-rich) and upper (generation-rich) parts of the system become detached from each other and a considerable amount of power is lost due to the lack of generation capacity in the detached lower part of the system besides the lack of transmission capacity, and hence, reactance tweaking will not be effective. Note, as an interesting side-observation in Table 3.2, that if the SO has a delay of even one stage in responding to the cascading failures (i.e. going from $T = 0$ to $T = 1$), the amount of load that needs to be shed post-contingency increases quite significantly for the case of $k = 3$, which reiterates the non-negligible difference between results obtained from a static model versus a multi-stage model.

3.6.2 General Efficacy

As mentioned earlier, the worst-case-scenario is an extreme case as disturbances are most often not strategic. Thus, to give a more general and practical measure of the efficacy of reactance tweaking in our numerical experiments, we next focus on the "general efficacy" of reactance tweaking (as defined in the previous section). In these experiments, we first set $k = 3$ and $H = 2$. In the first step, using the nominal load flow scenario that is provided in the dataset for the IEEE One-Area RTS-1996 test case system (which amounts to 2850MW in total load), we set $T = 1$ and compute (using the MILP derived in Theorem 6) that the total amount of savings in load-shedding

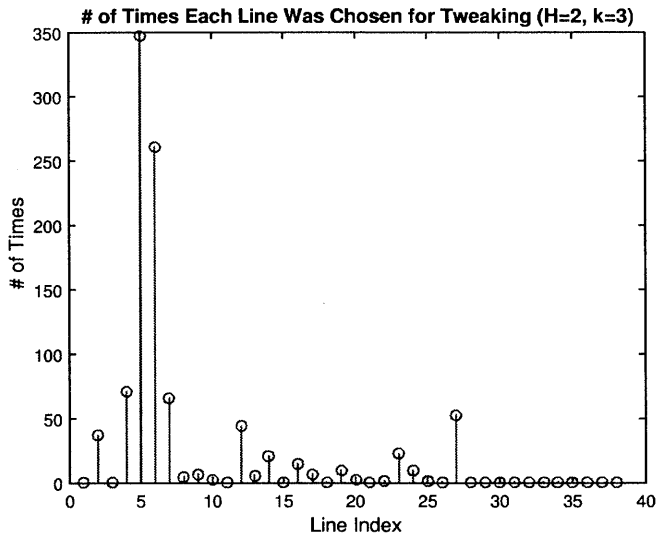


Figure 3-3: Case of H=2, k=3: Number of trials (out of 500 randomly sampled load levels) in which each line was chosen as one of the two optimal lines to tweak the reactance of, when computing the tight upper bound on the efficacy of reactance tweaking for that trial.

as a pure contribution of reactance tweaking (i.e. the parameter $E(\kappa)$) is tightly upper bounded by 60.7MW for this case (which is about 2.13% of the total demand); then, we set $T = 0$ for the same load level and the tight upper bound on $E(\kappa)$ drops to $42MW$. Next, we randomize the load at each demand bus such that it is sampled from a uniform distribution over the support $[L_i^N - 0.1L_i^N, L_i^N + 0.1L_i^N]$ where L_i^N is the nominal load (also referred to as the "base load") at bus i (i.e. we allow for up to 10% uncertainty in the load at each demand bus relative to the base load). For increased computation speed, we set $T = 0$ and run this experiment 500 times for 500 different realizations of the random vector of bus loads (we use the MILP derived in Theorem 6 to perform these computations). We observe that the value of the upper bound on $E(\kappa)$ among these 500 trials ranges from 34.3MW to 71.8MW with mean 45.9MW and standard deviation 6.8MW. This amount of contribution by reactance tweaking is quite remarkable and certainly promising.

Figure 3-3 shows the number of times (out of 500 randomly sampled load levels) that each line was chosen post-contingency as one of the two optimal lines to tweak the reactance of .

A first observation in Figure 3-3 is that the lines indexed from 1 through 14 (with the exception of 11) have been used more frequently than the rest of the lines. Looking at the line capacity data and the locations of lines $1 - 14$, an immediate observation is that almost all the lines indexed $1 - 14$, which are all located in the southern part of the power system, have much lower flow capacity than lines $15 - 38$, and hence, tweaking reactances in that part of the power system naturally makes intuitive sense because the line flows can become so close to capacity limits that even minor changes in reactances can redirect some of the flow to other lines and prevent overloading by taking advantage of all the residual capacity of the lines in the lower part of the system. Bearing in mind that the southern part of the system is demand-rich in the sense that a considerable portion of the total demand in the system is concentrated in the southern part, another interesting observation would be that the southern part of the system also quite frequently appears as the optimal location for reactance tweaking.

We repeat the above experiment for $H = 3$ and $H = 1$, while keeping all other parameters the same as above (especially, still $k = 3$, and we again use 500 random samples for the loads); the results are as follows:

- Increasing the number of tweaking devices (H) from 2 to 3 **does not** change the average bound on savings (i.e. average among 500 samples), and the same two lines (i.e. lines 5 and 6) are the modes.
- Decreasing the number of tweaking devices (H) from 2 to 1 decreases the average upper bound on savings among 500 samples to 33.9MW (i.e. the average bound on savings decreases by 12MW compared to the case of $H = 2$ which had an average savings of 45.9MW as reported above).

Figure 3-4 shows the number of times (out of 500 randomly sampled load levels) that each line was chosen post-contingency as the single optimal line to tweak the reactance of, when computing the tight upper bound on the efficacy of reactance tweaking for that trial.

Note in Figure 3-4 that for $H = 1$, line 5 is still the mode (recall that for $H = 2$, line 5 was the mode

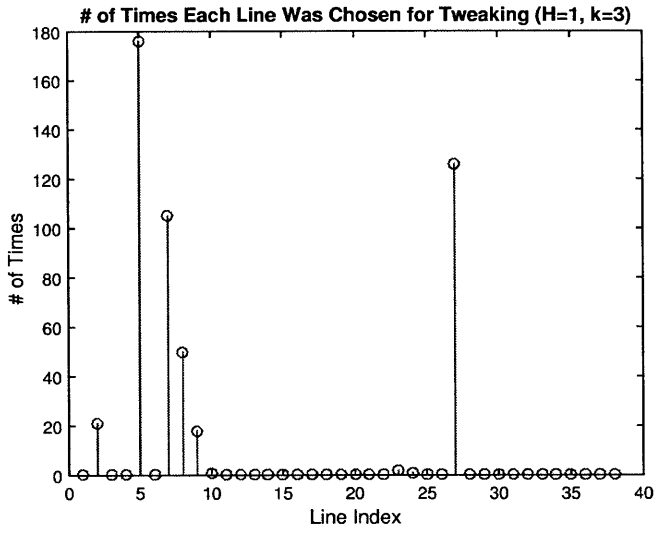
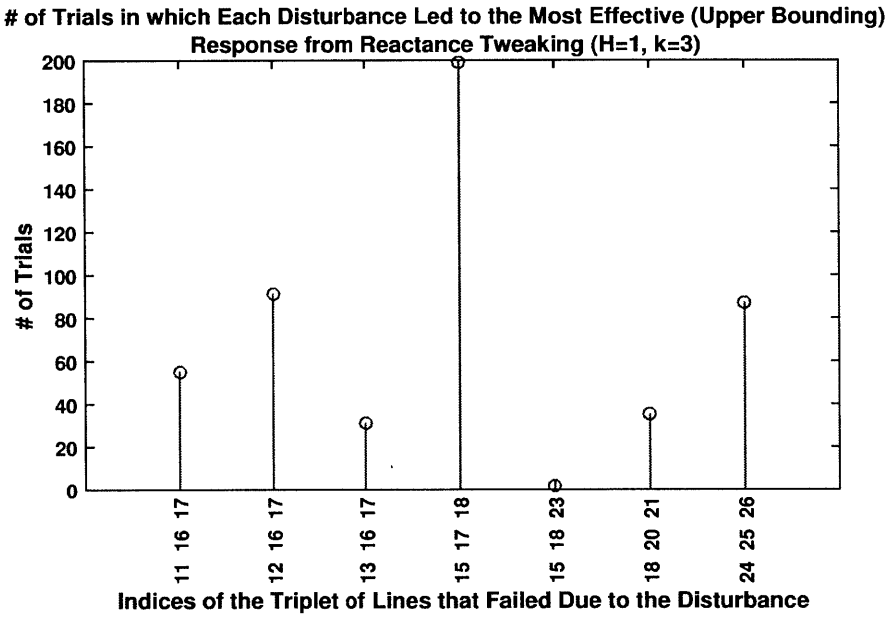


Figure 3-4: Case of H=1, k=3: Number of trials (out of 500 randomly sampled load levels) in which each line was chosen as the single optimal line to tweak the reactance of, when computing the tight upper bound on the efficacy of reactance tweaking for that trial.

but line 6 also had been selected as an optimal line to tweak in a large number of trials), but this time line 6 is never an optimal line for reactance tweaking in the 500 trials, and instead lines 7 and 27 have each appeared over 100 times as the optimal line to tweak the reactance of. This observation seems to potentially have non-trivial implications for optimal placement of reactance tweaking devices in the transmission system, especially when it come to deciding how many such devices are economically viable for the system. Although the problem of optimal placement of reactance-tweaking devices is not a topic of interest in this thesis and merits its own line of research, we will look more closely into it later in this chapter to motivate future research.

A side-observation made in the above set of simulations was that among the 500 trials for the various load levels, there seemed to be a small number of disturbances that appeared frequently as the N-3 disturbance that led to an upper-bounding efficacy from reactance tweaking. In fact, for $H = 1$ and $H = 2$, there were only 7 and 3 disturbances, respectively, that had led to the upper bounding performance from reactance tweaking in all 500 trials performed for each case (i.e. for the $H = 1$ case and for the $H = 2$ case); the corresponding plots are shown in Figure 3-5.



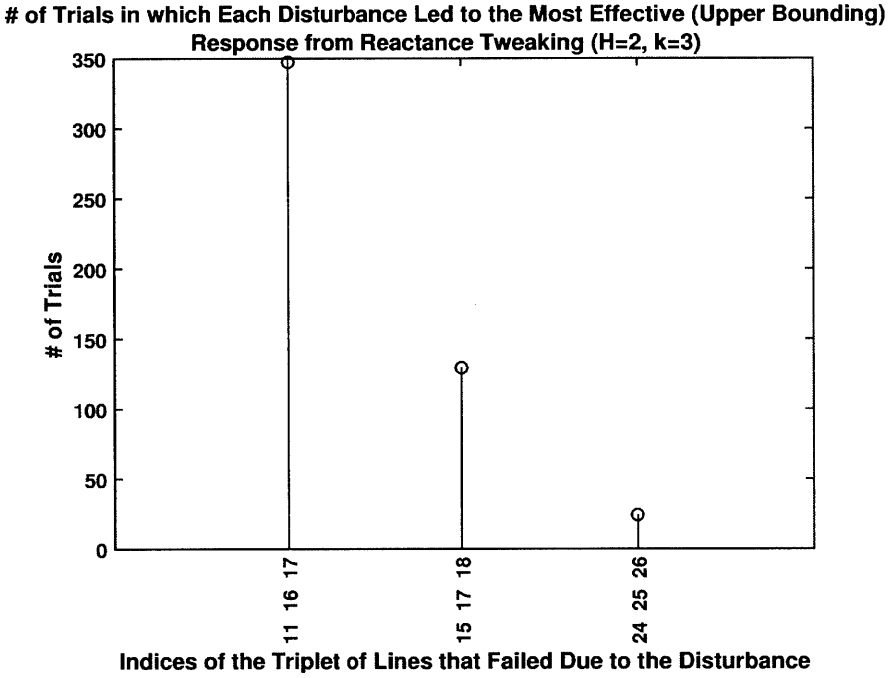


Figure 3-5: Cases of H=1 and H=2: Number of trials (out of 500 randomly sampled load levels) in which each disturbance led to maximal savings obtained from reactance tweaking.

Note in Figure 3-5 that the three disturbances appearing for the case of $H = 2$ as the disturbances in response to which reactance tweaking had the highest efficacy (for the 500 samples of load levels), namely the removal of lines $\{11, 16, 17\}$ or lines $\{15, 17, 18\}$ or lines $\{24, 25, 26\}$, are three of the disturbances that appeared very frequently in the upper bounding performance for the $H = 1$ case as well. This suggests that in One-Area RTS-1996, the 10% uncertainty that we allowed in the loads does not create much variety in the disturbances in response to which reactance tweaking is most effective. Also, the above results suggest that in One-Area RTS-1996, it is most often the same disturbances that lead to a maximally effective performance from reactance tweaking, regardless of whether we use reactance tweaking on one line or on two lines.

3.6.3 Further Observations/Motivating Future Directions

To dig deeper into the observations made in the previous subsection, we next fix the load to its base value (i.e. we no longer randomize the nodal demands), and instead, we try all possible N-3 contingencies (exhaustively) and compute the average value of the efficacy of reactance tweaking over all the contingencies for which reactance-tweaking is "noticeably" effective (i.e. leads to savings of at least 2-3MW in the amount of load shed post-disturbance), assuming we can place the tweaking device(s) anywhere we like for each contingency; the results are as follows:

- H=2 and H=3, base load case: 11 contingencies were "noticeably" mitigated by reactance tweaking; if we have the freedom to choose the optimal location of the devices independently for each contingency, the average value of savings among these 11 contingencies is about 18MW for both $H = 2$ and $H = 3$ (the number of times, among these 11 contingencies, that each line was chosen as an optimal location of the reactance-tweaking device for the $H = 2$ case is shown in Figure 3-6).

- H=2 and H=3, maximal load case: 45 contingencies were "noticeably" mitigated by reactance tweaking; if we have the freedom to choose the optimal location of the devices independently for each contingency, the average value of savings among these 45 contingencies

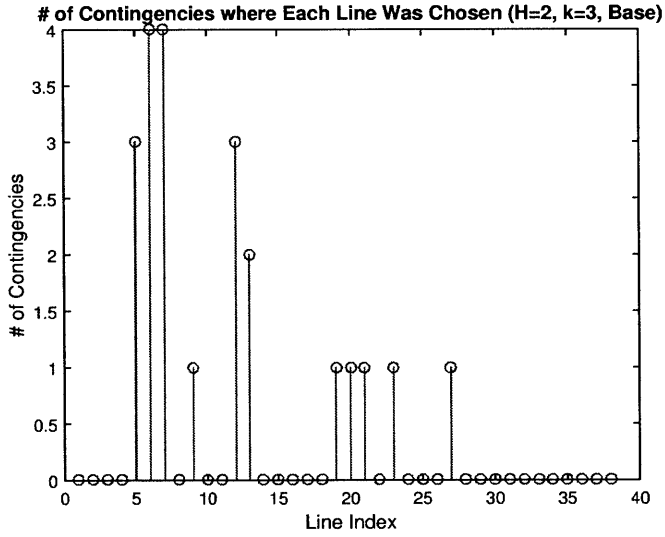


Figure 3-6: The number of times, among the 11 contingencies, that each line was chosen as one of the two optimal locations for the reactance-tweaking devices (for the case of $H = 2$, $k = 3$, and base load)

is about 19MW for both $H = 2$ and $H = 3$ (the number of times, among these 45 contingencies, that each line was chosen as an optimal location of the reactance-tweaking device for the $H = 2$ case is shown in Figure 3-7).

- H=1, base load case: 11 contingencies were "noticeably" mitigated by reactance tweaking; if we have the freedom to choose the optimal location of the devices independently for each contingency, the average value of savings among these 11 contingencies is about 14.7MW (the number of times, among these 11 contingencies, that each line was chosen as the optimal location of the reactance-tweaking device is shown in Figure 3-8).

- H=1, maximal load case: 45 contingencies were "noticeably" mitigated by reactance tweaking; if we have the freedom to choose the optimal location of the devices independently for each contingency, the average value of savings among these 45 contingencies is about 17.6MW in savings (the number of times, among these 45 contingencies, that each line was chosen as the optimal location of the reactance-tweaking device is shown in Figure 3-9).

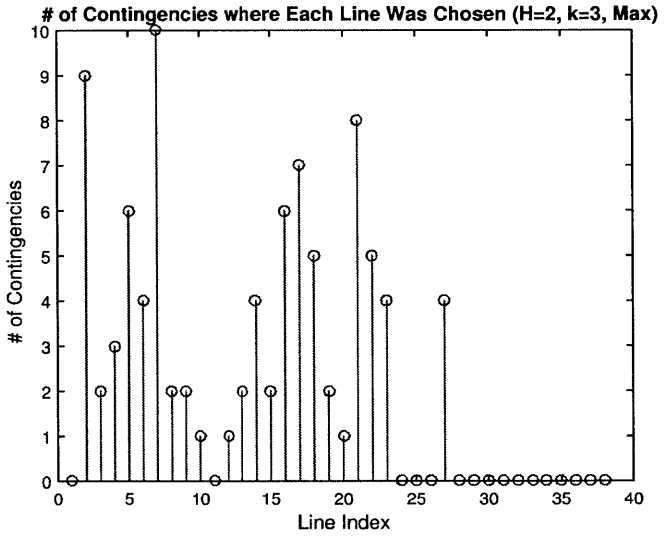


Figure 3-7: The number of times, among the 45 contingencies, that each line was chosen as one of the two optimal locations for the reactance-tweaking devices (for the case of $H = 2$, $k = 3$, and maximal load)

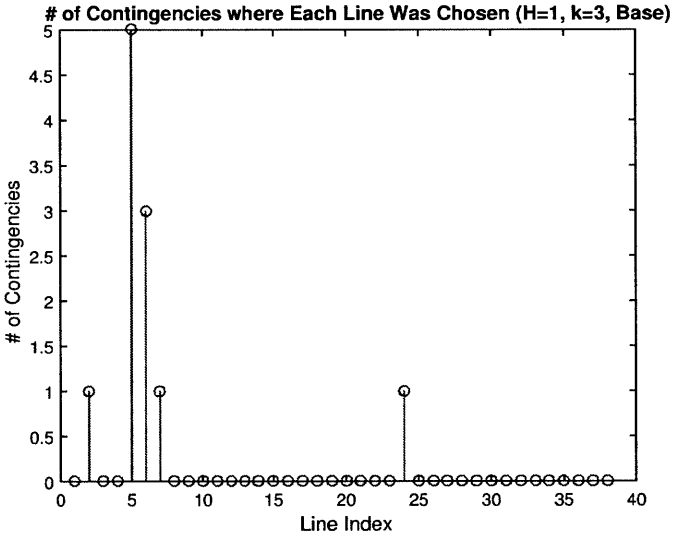


Figure 3-8: Total number of times, among the 11 contingencies, that each line was chosen as the optimal location of the reactance-tweaking device (for the case of $H = 1$ and base load)

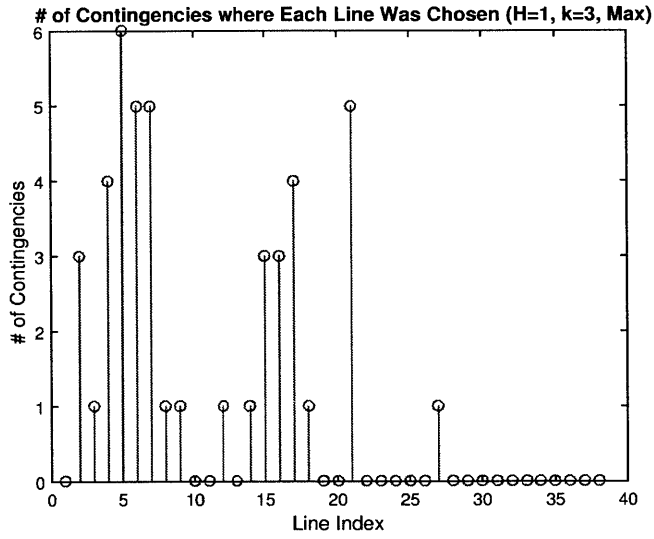


Figure 3-9: The number of times, among the 45 contingencies, that each line was chosen as the optimal location of the reactance-tweaking device (for the case of $H = 1$, $k = 3$, and maximal load)

The above results suggest that increasing the number of reactance-tweaking devices in One-Area RTS-1996 may not considerably change the efficacy of reactance tweaking if we have the freedom to change the location of the device(s) from contingency to contingency. This makes intuitive sense based on the same arguments as in the previous subsections regarding how the efficacy of reactance tweaking can be considerably limited by the lack of transmission capacity, and also given the fact that the flow in any line in a connected DC electric network depends on the resistances of all lines in the system; hence, in many scenarios, redirecting only a small amount of flow from some line(s) to other line(s) could be all that is feasible due the lack of residual capacity in the critical lines in the system, and this small amount of reactance-tweaking control may potentially be accomplished by a small number of reactance-tweaking devices with possibly a large number of options in terms of the location of the device(s). Comparing Figures 3-6 through 3-9 with each other reveals that depending on the load level and H, a broad array of lines could be desirable location(s) for reactance tweaking while some lines (such as line 5) appear frequently as an optimal location for both $H = 1$ and $H = 2$ cases and for both base and maximal load levels (note also in these figures that for the maximal load level, the desirable location for reactance tweaking devices

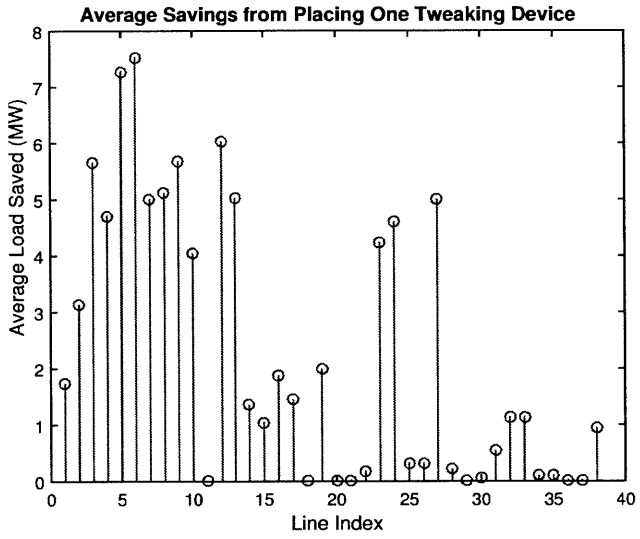


Figure 3-10: Average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective for the base load case) by placing a total of one reactance-tweaking device in the entire system, as a function of the line on which the device is placed.

is not as concentrated in the southern part of the system as it is for the base load level, and lines that are in the mid part have appeared multiple times as optimal locations in the above computations for the maximal load case). This intuition leads us to our next set of numerical simulations. This time, we will fix the location(s) of the reactance-tweaking device(s) and we no longer allow the location of the device(s) to be a decision variable, and instead, for each possible placement scenario for the device(s), we exhaustively compute the average savings over all N-3 contingencies to which reactance-tweaking is "noticeably" effective; namely, we average over the 11 contingencies we identified earlier (among the set of all $N-3$ contingencies for the base load case) that result in "noticeable" savings in load-shedding purely contributed by reactance tweaking).

Figure 3-10 shows the average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective for the base load case) by placing a total of one reactance-tweaking device in the entire system, as a function of the line on which the reactance-tweaking device is placed.

An immediate observation in Figure 3-10 is that although line 6 appears to be the location (for placing the reactance tweaking device) that yields the best performance in the corresponding numerical experiment, there are many other lines that are not very far below line 6 in terms of efficacy; this further strengthens the conjecture that reactance tweaking might generally be flexible with respect to the choice of the location of the device. Once again, the lines in the southern part of the One-Area RTS-1996 system appear to be the ones on which reactance tweaking seems most effective for the base load case, but some lines from the mid part of the system (such as lines 23, 24, and 27) also seem to provide reasonably good average savings in the numerical experiments corresponding to Figure 3-10.

Next, we look at the case of $H = 2$. Figure 3-11 shows, for the base load case, the average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective) by placing a total of two reactance-tweaking devices in the entire system, as a function of the pair of lines on which the devices are placed, for all possible $\binom{38}{2} = 703$ placement choices (the ordering of the location pairs on the horizontal axis of the plot has been sorted in descending order in terms of efficacy). An immediate observation in this figure is that the curve is relatively steep; by comparing this figure to Figure 3-10, we observe that the most effective placement for the case of H=2 leads to twice as much savings as the most effective placement for the case of H=1. Also, the 100-th most effective placement for the pair of reactance-tweaking devices is almost as effective as placing just one device only on line 6 (which was the best case for H=1); thus, it seems plausible to suggest that placing two reactance-tweaking devices in the system is notably more effective than placing only one reactance-tweaking device. Figure 3-12 also reveals that there are 400 possible ways of placing the two devices that would yield an average of at least $4MW$ in savings in the post-disturbance load shed (recall that the total demand in the base load case in One-Area RTS-1996 system is 2.85GW).

Next, we zoom into Figure 3-11 to give a closer look at the top 50 most effective choices for the locations of the two devices, so that the indices of the pairs of lines that provide the best 50 placement choices become available to the reader, and the result is shown in Figure 3-12. A

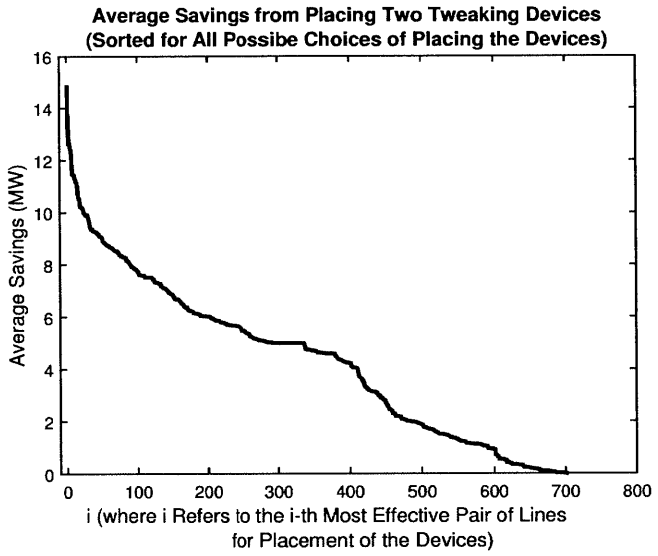


Figure 3-11: Average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective for the base load case) by placing a total of two reactance-tweaking devices in the entire system, as a function of the pair of lines on which the two devices are placed, sorted in descending order in terms of efficacy for all possible $\binom{38}{2} = 703$ placement choices.

quick observation in Figure 3-12 is that lines 5 and 6 appear in a vast majority of the top 50 pairs, and their combination (for placing the two devices) yields the highest "average" efficacy (in the "noticeably" effective sense described earlier in this subsection).

Next, we look at the case of $H = 3$. Figure 3-13 shows, for the base load case, the average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective) by placing a total of three reactance-tweaking devices in the entire system, as a function of the triplet of lines on which the devices are placed, for all possible $\binom{38}{3} = 8436$ placement choices (the ordering of the location triplets on the horizontal axis of the plot has been sorted in descending order in terms of efficacy). An immediate observation in this figure is that the curve is relatively steep, like it was for the case of H=2 shown in Figure 3-11. By comparing this figure to Figure 3-11, we observe that the most effective placement for the case of H=3 leads to only 3MW more savings compared to the most effective placement for the case of H=2; this suggests that the additional device seems to add marginal value to the efficacy of reactance tweaking. Figure 3-12 also reveals that there are 6500 possible ways of placing the two devices that would yield an

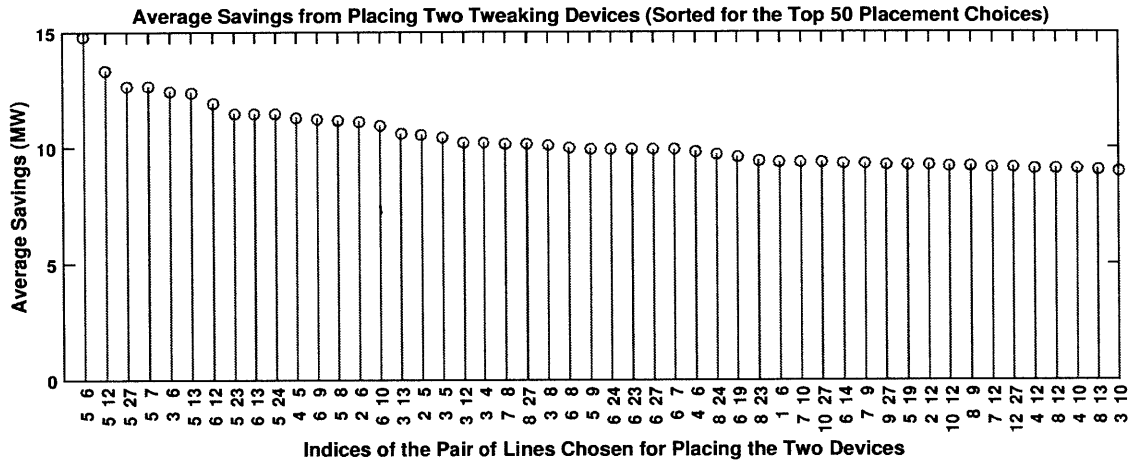


Figure 3-12: This figure is obtained by zooming into Figure 3-11 to give a closer look at the top 50 most effective choices for the locations of the two devices, so that the indices of the pairs of lines that provide the best 50 placement choices become available to the reader.

average of at least $4MW$ in savings in the post-disturbance load shed.

Now, we zoom into Figure 3-13 to give a closer look at the top 50 most effective choices for the locations of the three devices, so that the indices of the triplets of lines that provide the best 50 placement choices become available to the reader, and the result is shown in Figure 3-14. A quick observation in Figure 3-14 is that lines 5 and 6 appear in a vast majority of the top 50 triplets (and most often both lines 5 and 6 are together in the triplet of lines), and their combination (along with line 12) yields the highest "average" efficacy (in the "noticeably" effective sense). By comparing this figure to Figure 3-12, we observe that the 50-th most effective placement for the triplet of reactance-tweaking devices is almost as effective as placing just two devices in the most desirable location for H=2: one on line 5 and the other on line 6.

A main observation in the above set of simulations is that increasing the number of reactance-tweaking devices from 2 to 3 only changes the best-case average savings (in the "noticeably effective sense, for the set of N-3 contingencies and a base load case) by 3MW, while decreasing the number of devices from 2 to 1 reduces this average saving amount by more than 7MW. It thus appears that investing in 2 reactance-tweaking devices in this system might be a good decision.

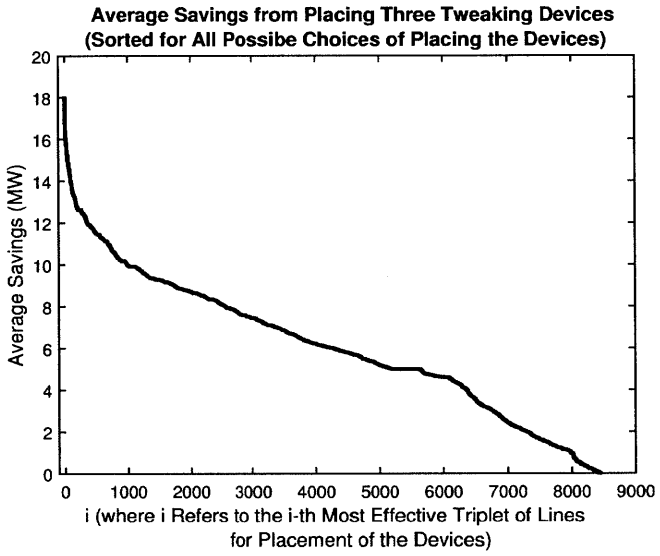


Figure 3-13: Average load saved (over the N-3 contingencies in response to which reactance tweaking is "noticeably" effective for the base load case) by placing a total of three reactance-tweaking devices in the entire system, as a function of the triplet of lines on which the three devices are placed, sorted in descending order in terms of efficacy for all $\binom{38}{3} = 8436$ possible placement choices.

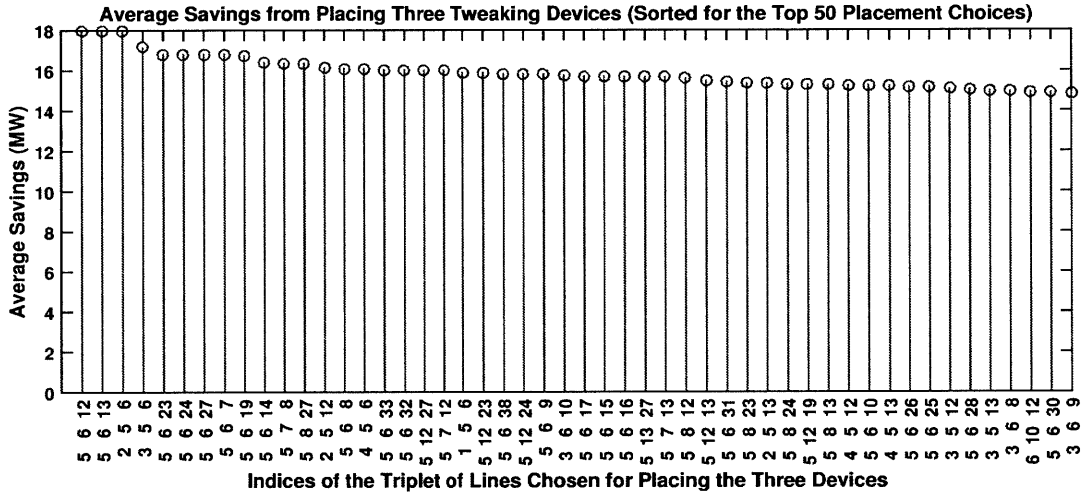


Figure 3-14: This figure is obtained by zooming into Figure 3-13 to give a closer look at the top 50 most effective choices for the locations of the three devices, so that the indices of the triplets of lines that provide the best 50 placement choices become available to the reader.

This is consistent with the computational results obtained in the previous subsection using the upper bound derived in Theorem 6, which also suggested that decreasing H from 2 to 1 could be unfavorable (as it considerably reduced the upper bound in those computations) while increasing H from 2 to 3 did not increase the upper bound. However, the above results suggest a fair amount of flexibility in terms of the placement of the device(s), while the computations with the upper bound in the previous subsection did not quite capture this property so well.

Assuming that average savings is the only criterion for choosing the locations of reactance-tweaking devices and assuming all contingencies have uniform probability, and disregarding other applications of the device, then the results in the above figures can give a good sense of where the "optimal" location(s) for the reactance-tweaking devices could be for the base load case in the One-Area RTS-1996 test system, in response to the set of all $N-3$ disturbances that can "noticeably" be mitigated by reactance tweaking. These results suggest that for the One-Area RTS-1996 base load case, the line(s) that appeared frequently as the upper-bounding locations for reactance tweaking (i.e. lines 5 and 6, as shown in Figures 3-3 and 3-4) in our computations in the previous subsection for the upper bound (derived in Theorem 6) on "general efficacy", were also the lines that eventually emerged as the "optimal" (in the exhaustive averaging sense mentioned above) placement locations for reactance tweaking. Please note that the problem of optimal placement of reactance-tweaking devices for post-disturbance control is not a topic of interest in this thesis as it merits its own line of research for the reasons mentioned above; thus, the purpose of presenting the above exhaustive simulations was to give some insight into the potential value of the upper bound derived in Theorem 6, because the above simulations suggest that this upper bound may be able to serve future research on this topic by providing a fast initial approximate of the lines on which reactance tweaking could be quite useful and also of the number of devices needed for the system (though, the author acknowledges that the observations made herein may not necessarily extend to systems other than the One-Area RTS-1996, and the author is solely sharing his observations). Please see Chapter 5 for a discussion of future directions that could potentially be relevant to the above observations.

Chapter 4

Can Pre-Disturbance Tweaking of Transmission Line Reactances Mitigate the Propagation of Cascading Failures?

So far, we focused on the case where the system operator responds to a disturbance by re-dispatching the system in order to maximize the amount of demand satisfied post-disturbance. Nevertheless, another application of a reactance-tweaking technology could be mitigating the propagation of cascading failures in the event of a multiple-line contingency, especially when the system operator cannot respond quickly to the disturbance and hence the cascading failures can propagate for a couple of stages. In other words, in the set of all possible $N - k$ scenarios for a given k, can we tweak line reactances at time $t = 0$ such that in the event of up to k outages at time $t = 1$, no more than q new outages happen at time $t = 2$? For the case of $k = 1$, this would be equivalent to the $N - 1$ security criterion discussed in the existing literature as explained in Chapter 1 (which is a very common security criterion in power systems nowadays) because the goal of the $N - 1$ criterion is to ensure that if we have one outage at time $t = 1$, we will have $q = 0$ outages at time $t = 2$. In this chapter, we seek to build an optimization framework for studying whether pre-disturbance tweaking of the reactances of a certain number of transmission lines (within a reasonable operating

range) can effectively reduce the number of failed lines over the propagation of cascading failures in the event of an $N - k$ scenario for an arbitrary k.

4.1 Efficacy in Mitigating the Worst-Case-Scenario

As in the previous chapter, we have two agents: A system operator (SO) and an adversary. At time $t = 0$, the system operator strategically tweaks the reactances of up to H lines, and then at time $t = 1$, the adversary removes up to k lines from the network, with the objective of initiating a cascade that will propagate for T discrete time steps (as per the rules and assumptions of cascade propagation defined in the previous chapters) such that the number of failed transmission lines at the end of the cascade (i.e. at $t = T + 2$) is maximized. Note that after tweaking line reactances at time $t = 0$, the system operator takes no further action. Similarly, after initiating a disturbance by removing up to k lines at time $t = 1$, the adversary takes no further action. Hence, from stage $t = 2$ onwards, it is only the flows that redistribute and cause more failures.

In order to evaluate the efficacy of the proposed pre-disturbance reactance-tweaking strategy in mitigating a worst-case-scenario disturbance, we need a max-min setup (zero-sum Stackelberg game) where the SO goes first by tweaking reactances and then the adversary takes action by causing a disturbance. We can write this bilevel program as:

$$\max_{\hat{d}\in\hat{D}} \min_{\hat{a}\in\hat{A}} \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2} \tag{4.1}$$

where, s_{ijh}^{T+2} denotes the binary state of transmission line (i, j, h) at time $T+2$ (hence, $\sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2}$ quantifies the number of transmission lines that have survived until time $T + 2$ which marks the final stage of cascading failures). All the variables and feasible sets in the above problem (i.e. $\hat{d}, \hat{D}, \hat{a}, \hat{A}$) shall be explicitly defined in the following two subsections.

The SO's Decision Variables and Constraints

The SO's problem is as follows:

$$\max_{\tilde{d}} \min_{\hat{a}\in\hat{A}} \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2} \tag{4.2}$$

$$\text{s.t.} \quad I_{ijh}^o x_{ijh} = \theta_i^o - \theta_j^o, \ \forall (i,j,h) \in \mathcal{E}$$

$$b_i^g - b_i^d = \sum_{(i,j,h)\in\mathcal{E}_i} I_{ijh}^o, \ \forall i \in \mathcal{B} \tag{S1}$$

$$x_{ijh} - x_{ijh}^o - \check{\tau}_{ijh}(x_{ijh}^{\max} - x_{ijh}^o) \leq 0 \ \forall (i,j,h) \in \mathcal{E}$$

$$x_{ijh}^o - x_{ijh} + \check{\tau}_{ijh}(x_{ijh}^{\min} - x_{ijh}^o) \leq 0 \ \forall (i,j,h) \in \mathcal{E}$$

$$\sum_{(i,j,h)\in\mathcal{E}} \check{\tau}_{ijh} = H \ \forall (i,j,h) \in \mathcal{E} \tag{S2}$$

$$-\overline{cap}_{ij} \leq I_{ijh}^o \leq \overline{cap}_{ij}, \ \forall (i,j,h) \in \mathcal{E} \tag{S3}$$

$$x_{ijh}, \theta_i^o, I_{ijh}^o, \in \mathbb{R}, \forall i \in \mathcal{B}, (i,j,h) \in \mathcal{E}$$

$$\check{\tau}_{ijh} \in \{0,1\}, \quad \forall (i,j,h) \in \mathcal{E}.$$

where $\tilde{d} \triangleq \{\mathbf{x}, \theta^{\mathbf{o}}, \mathbf{I}^{\mathbf{o}}, \check{\boldsymbol{\tau}}\}$ is the set of decision variables, and $\mathbf{x}^o$ is the vector of original line reactances which is given to the problem as an input parameter. In the above formulation, the objective is to maximize the number of lines that will survive the disturbance (which itself is another optimization problem $\min_{\hat{a}\in\hat{A}} \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2}$, namely the "adversary's problem"). The first two constraints are KVL and KCL for DC flows, the next three constraints ensure that the reactances of at most H lines can be tweaked within a prescribed range (and all the other lines will retain their original reactances), the sixth constraint ensures no line flow capacities are violated, and the seventh constraint ensures that the generation capacities of each bus are not exceeded (note that $\mathbf{b}^g$ and $\mathbf{b}^d$ are not decision variables; they are given parameters (possibly from solving the economic dispatch problem), and denote generation and demand at each bus, respectively).

The Adversary's Decision Variables and Constraints

The set of constraints for the adversary in this problem is very similar to the set of adversary's constraints from the previous chapter, with only one difference: we allow some uncertainty in the line reactances at each stage of the evolution of cascading failures that the adversary can exploit in order to come up with a better disturbance. This serves as a proxy for the uncertainty in the reactances implemented by the reactance-tweaking device (potentially stemming from implementation limitations of the reactance-tweaking device) and external factors such as frequency uncertainty that can impact the reactances of all lines during the propagation of cascading failures. Thus, the actual reactance $\hat{x}^t_{ijh}$ of line (i,j,h) at stage $t \geq 1$ is assumed to be $\hat{x}^t_{ijh} \triangleq x_{ijh} + \tilde{x}^t_{ijh}$, where $\tilde{x}^t_{ijh} \in [-\tilde{\mathcal{U}} \;\; \tilde{\mathcal{U}}]$ (for some $\tilde{\mathcal{U}} \in \mathbb{R}_+$) quantifies the uncertainty in the reactance of line (i,j,h), while x_{ijh} denotes the reactance that the SO had implemented prior to the disturbance. We shall assume that $\hat{x}_{ijh} \geq 2\,\tilde{\mathcal{U}}$ to ensure that the uncertain term does not dominate the actual reactance. The adversary's problem is stated below:

$$\min_{\hat{a}\in\hat{A}} \sum_{(i,j,h)\in\mathcal{E}} s^{T+2}_{ijh} \tag{4.3}$$

$$\text{s.t.} \quad \mathbf{1}'\mathbf{s}^1 \geq |\mathcal{E}| - k$$

$$M\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \mathbf{b}^g - \mathbf{b}^d \;\forall t \geq 2$$

$$\text{diag}(\hat{\mathbf{x}}^t)\text{diag}(\mathbf{z}^{t-1})\mathbf{I}^t = \text{diag}(\mathbf{z}^{t-1})M'\theta \;\forall t \geq 2 \qquad (A1)$$

$$\hat{x}^t_{ijh} = x_{ijh} + \tilde{x}^t_{ijh}, \quad \tilde{x}_{ijh} \in [-\tilde{\mathcal{U}},\tilde{\mathcal{U}}], \quad \forall (i,j,h) \in \mathcal{E} \qquad (A2)$$

$$(z^{t-1}_{ijh}+s^t_{ijh}-1)|I^t_{ijh}| \leq \overline{cap}^a_{ijh} \;\forall (i,j,h)\in\mathcal{E}, \forall t \geq 2$$

$$z^{t-1}_{ijh}|I^t_{ijh}| \geq \overline{cap}^f_{ijh}(z^{t-1}_{ijh}-s^t_{ijh}) \;\forall (i,j,h)\in\mathcal{E}, \forall t \geq 2$$

$$|M'|\mathbf{y}^t \geq \mathbf{z}^t \geq \mathbf{s}^t + |M'|\mathbf{y}^t - 2 \;\; \forall t \geq 2$$

$$y^t_i = 1 \;\; \forall i \in \mathcal{B}^o_{\pm}, \forall t \leq T+1; \;\; \mathbf{y}^{t-1} \geq \mathbf{y}^t, \;\; \forall t \leq T+1$$

$$\mathbf{s}^t \geq \mathbf{z}^t \;\; \forall t \geq 1 \;\; ; \mathbf{z}^{t-1} \geq \mathbf{s}^t \;\; \forall t \geq 2; \;\; i \overset{t}{\leftrightarrow} j \;\; \forall\, i,j \in \mathcal{B}^o_{\pm}, 1 \leq t \leq T+1$$

$$z^t_{ijh},\; s^t_{ijh},\; y^t_q \in \{0,1\}, \quad \forall (i,j,h) \in \mathcal{E},\; \forall q \in \mathcal{B}.$$

where $\hat{A} = \{\mathbf{s}^1, \ldots, \mathbf{s}^{T+2}, \mathbf{z}^1, \ldots, \mathbf{z}^{T+1}, \mathbf{y}^1, \ldots, \mathbf{y}^{T+1}, \mathbf{I}^2, \ldots, \mathbf{I}^{T+2}, \theta^2, \ldots, \theta^{T+2}, \hat{\mathbf{x}}^2, \ldots, \hat{\mathbf{x}}^{T+2}, \tilde{\mathbf{x}}^2, \ldots, \tilde{\mathbf{x}}^{T+2}, \mathcal{P}\}$. Here, $\mathcal{P}$ denotes the set of all variables used in keeping track of the connectivity of the buses (which are the auxiliary variables defined in Theorem 1) for all $t \in \{1, \ldots, T+1\}$. In (4.3), the first constraint requires that the adversary trips up to k lines at $t = 1$, the second and third constraints are for KVL and KCL at each stage, the fourth constraint defines the decomposition of line reactances into their certain and uncertain components, the fifth and the sixth constraints keep track of the lines that have survived and failed (respectively), and the remainder of the constraints keep track of the connectivity of the underlying graph (also ensuring that no failed line is reactivated) and evolution of the topology as established in Proposition 1 and Theorem 1.

4.2 Approximate Reformulation as a Bilevel MILP

Note that the above bilevel program has continuous$\times$continuous bilinear terms, and the two levels are coupled through the continuous variable x_{ijh}. A quick observation in the above formulation gives us an ingenious approach for giving a mixed-integer linear approximation of the above. Let us first approximate x_{ijh} using a finite number of its significant figures (via truncation) and denote this approximation by $\check{x}_{ijh}$ (i.e., if $x_{ijh} = 1.34267$ we can approximate it using only its first four significant figures and write $x_{ijh} \approx \check{x}_{ijh}$ where $\check{x}_{ijh} = 1.342$). This approximation is motivated by the fact that if the number of significant figures chosen for forming $\check{x}_{ijh}$ is aligned well with the magnitude of $\tilde{\mathcal{U}}$, then $\hat{x}^t_{ijh} \approx \check{x}_{ijh} + \tilde{x}^t_{ijh}$ can be a very good approximation. To see this, let us consider an example. Let us assume in this example that $\tilde{\mathcal{U}} = 0.01$ so that $\tilde{x}_{ijh} \in [-0.01 \;\; 0.01]$. Also, let us assume that x_{ijh} is set to 1.34267 as in the above example. Then, $\hat{x}^t_{ijh} = x_{ijh} + \tilde{x}^t_{ijh} \in [1.33267 \;\; 1.35267]$, while we have $\check{x}_{ijh} + \tilde{x}^t_{ijh} \in [1.332 \;\; 1.352]$ if we set $\check{x}_{ijh} = 1.342$.

Since $\check{x}_{ijh}$ has a finite number of digits, we can then discretize it using its binary expansion and write it as a finite sum:

$$\check{x}_{ijh} = \sum_{p=L}^{U} 2^p \check{\zeta}_{p;ijh}, \quad \check{\zeta}_{p;ijh} \in \{0,1\} \;\; \forall p$$

where L and U are in $\mathbb{Z}$, and their values are determined by the number of significant figures with which $\check{x}_{ijh}$ is defined. Then, if I^t_{ijh} is sufficiently small, we can approximately linearize the bilinear product $I^t_{ijh}x_{ijh}$ by first writing

$$I^t_{ijh}\hat{x}_{ijh} = I^t_{ijh}(x_{ijh} + \tilde{x}^t_{ijh}) \approx [\sum_{p=L}^{U} 2^p \check{\zeta}_{p;ijh} I^t_{ijh}] + I^t_{ijh}\tilde{x}^t_{ijh}$$

Note that since $\check{\zeta}_{p;ijh}$ is binary and I^t_{ijh} is bounded, $\check{\zeta}_{p;ijh}I^t_{ijh}$ can efficiently be linearized using the big-m technique described in the previous chapter. We now have all the tools to approximate the feasible set of Problem (4.3), which is the inner-level of our original max-min problem, by replacing x_{ijh} with $\sum_{p=L}^{U} 2^p \check{\zeta}_{p;ijh}$ in Problem (4.3). The bounds and techniques needed for linearizing all the binary$\times$continuous terms in this problem are the same as those used for Problem (3.6) in the previous chapter. To complete this approximation for the entire max-min problem (4.1), we shall adjust the outer-level (i.e. Problem (4.2)) to this approximation as well. To do so, for each line (i, j, h) we shall replace x_{ijh} with $\sum_{p=L}^{U} 2^p \check{\zeta}_{p;ijh} + \check{u}_{ijh}$ in (4.2) and add the constraint $\check{u}_{ijh} \in [-\tilde{\mathscr{U}} \quad \tilde{\mathscr{U}}]$ to the set of constraints in (4.2) and add $\check{u}_{ijh}$ and $\check{\zeta}_{p;ijh}$ to the set of decision variables in (4.2) (these decision variables and constraints shall be added for all p and (i, j, h)). We shall denote the approximate max-min problem we just descibed as follows:

$$\max_{\check{d}\in\check{D}} \min_{\check{a}\in\check{A}} \sum_{(i,j,h)\in\mathcal{E}} s^{T+2}_{ijh} \tag{4.4}$$

Let us denote by $\mathscr{B}$ the $|\mathcal{E}| \times |U - L + 1|$ matrix whose entry in the h-th row and p-th column is equal to $\check{\zeta}_{p;ijh}$. We shall denote by $2^{[LU]}$ the vector $[2^L, 2^{L+1}, \ldots, 2^U]'$. Using the same technique as in the proof of Theorem 3 in the previous chapter, we can prove that both the inner-level and the outer-level of the above approximation can each be converted to a MILP.

Theorem 7. *If $\check{\Gamma}^*$ is the optimal value of the objective function in the optimization problem (4.4), then*

$$\check{\Gamma}^* = \max_{\check{\omega}} \min_{\check{a}\in\check{A}} \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2} \tag{4.5}$$

$$\begin{aligned}
s.t.\quad & \check{\boldsymbol{\mu}} + (diag(\mathbf{I}^o)\mathscr{B})2^{[LU]} = M'\boldsymbol{\theta}^o; \\
& \check{\boldsymbol{\mu}} + (diag(\mathbf{I}^o)\mathscr{B})2^{[LU]} - diag(\boldsymbol{x}^o)\boldsymbol{I}^o - 2\Upsilon\check{\boldsymbol{\tau}} \leq 0; \\
& -\check{\boldsymbol{\mu}} - (diag(\mathbf{I}^o)\mathscr{B})2^{[LU]} + diag(\boldsymbol{x}^o)\boldsymbol{I}^o - 2\Upsilon\check{\boldsymbol{\tau}} \leq 0; \\
& diag(\mathbf{x}^{\min})\check{\boldsymbol{\alpha}} \leq \check{\boldsymbol{\delta}} + (diag(\check{\boldsymbol{\alpha}})\mathscr{B})2^{[LU]} \leq diag(\mathbf{x}^{\max})\check{\boldsymbol{\alpha}}; \\
& 0 \leq \check{\boldsymbol{\delta}} \leq \check{\boldsymbol{\alpha}}\tilde{\mathscr{U}}; \\
& \check{\boldsymbol{\alpha}} = \check{\boldsymbol{\beta}} + \check{\boldsymbol{\varphi}};\ \ \boldsymbol{I}^o = \check{\boldsymbol{\beta}} - \check{\boldsymbol{\varphi}}; \\
& \check{\boldsymbol{\beta}} - diag(\overline{\mathbf{cap}}^a)\check{\boldsymbol{\upsilon}} \leq 0;\ \ \check{\boldsymbol{\varphi}} - diag(\overline{\mathbf{cap}}^a)(1-\check{\boldsymbol{\upsilon}}) \leq 0; \\
& \check{\boldsymbol{\delta}} = \check{\boldsymbol{\phi}} + \check{\boldsymbol{\Phi}};\ \ \check{\boldsymbol{\mu}} = \check{\boldsymbol{\phi}} - \check{\boldsymbol{\Phi}}; \\
& \check{\boldsymbol{\Phi}} - \Upsilon(1-\check{\boldsymbol{\psi}}) \leq 0;\ \ \check{\boldsymbol{\phi}} - \Upsilon\check{\boldsymbol{\psi}} \leq 0; \\
& \check{\boldsymbol{\phi}} \geq 0;\ \ \check{\boldsymbol{\Phi}} \geq 0;\ \ \check{\boldsymbol{\varphi}} \geq 0;\ \ \check{\boldsymbol{\beta}} \geq 0; \\
& Constraints\ (S1),\ (S2),\ (S3)\ \textit{from Problem}\ (4.2); \\
& \check{\boldsymbol{\alpha}}, \check{\boldsymbol{\beta}}, \check{\boldsymbol{\varphi}}, \check{\boldsymbol{\mu}}, \check{\boldsymbol{\delta}}, \check{\boldsymbol{\Phi}}, \check{\boldsymbol{\Phi}}, \boldsymbol{I}^o \in \mathbf{R}^{|\mathcal{E}|}; \\
& \check{\boldsymbol{\tau}}, \check{\boldsymbol{\psi}}, \check{\boldsymbol{\upsilon}} \in \{0,1\}^{|\mathcal{E}|};\ \ \mathscr{B} \in \{0,1\}^{|\mathcal{E}|\times(U-L+1)};\ \ \boldsymbol{\theta}^o \in \mathbb{R}^{|\mathcal{B}|}.
\end{aligned}$$

where $\Upsilon \triangleq G\sum_{(p,q,h)\in\mathcal{E}} x_{pqh}^{\max}$ *and* $\check{\omega} = \{\mathscr{B}, \check{\boldsymbol{\alpha}}, \check{\boldsymbol{\beta}}, \check{\boldsymbol{\varphi}}, \check{\boldsymbol{\mu}}, \check{\boldsymbol{\delta}}, \check{\boldsymbol{\Phi}}, \check{\boldsymbol{\phi}}, \boldsymbol{I}^o, \check{\boldsymbol{\tau}}, \check{\boldsymbol{\psi}}, \check{\boldsymbol{\upsilon}}, \boldsymbol{\theta}^o\}$.

Proof. Please see Section 4.4. □

Note that the only bilinear terms in (4.5) are products of integer and continuous variables which can be linearized easily using the technique discussed in the previous chapter.

Using a similar technique as in Theorem 7, we can also linearize the continuous×continuous bilinear terms in the inner-level program in (4.4) (i.e. $\min_{\check{a}\in\check{A}} \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2}$). To do so, we shall replace

the constraints (A1) and (A2) in (4.3) with

$$diag(\mathbf{z}^{t-1})\tilde{\boldsymbol{u}} + (diag(\mathbf{z}^{t-1})diag(\mathbf{I}^t)\mathscr{B})2^{[LU]} = diag(\mathbf{z}^{t-1})M'\boldsymbol{\theta}^o$$

and

$$-|\mathbf{I}^t|\,\tilde{\mathcal{U}} \leq \tilde{\boldsymbol{u}} \leq |\mathbf{I}^t|\,\tilde{\mathcal{U}}$$

and then linearize all the absolute value terms and binary×continuous bilinear terms using the techniques described in the previous chapter (note that although $\mathscr{B}$ is not a decision variable in this inner-level program, we still need to linearize the product of its elements with the elements of $\mathbf{I}^t$ to simplify the column-and-row-generation procedure). We shall denote the set of decision variables and the feasible set of the linearized inner-level program by $\mathscr{A}$ and $\tilde{A}$, respectively.

To make this formulation amenable to row and column generation, we shall add the auxiliary continuous variable $\mathscr{X} \in \mathbb{R}$ and the constraint $\mathscr{X} = \sum_{(i,j,h)\in\mathcal{E}} s_{ijh}^{T+2}$ to the inner-level program so that we can write the objective as a function of a continuous variable rather than a sum of binary variables.

We shall denote the bilevel MILP obtained from performing all the above-mentioned linearizations and change of variables by

$$\max_{\tilde{d}\in\tilde{D}} \min_{\mathscr{A}\in\tilde{A}} \mathscr{X} \tag{4.6}$$

Column and Row Generation for Solving the Approximate Max-Min Problem

Once again, we shall apply a column-and-row-generation scheme (see e.g. [63], [62], [64]) to solve Problem 4.6, which entails deriving bounds on the dual variables in the innermost-level program in our problem. The procedure for doing so is similar to the one used in the previous chapter, which we shall repeat here for completeness. First, we turn Problem (4.6) into a large-scale single-level MILP. To do so, we start by separating the continuous variables in the outer level of (4.6) from the integer ones. Let us denote the vector of integer decision variables in the inner level of (4.6) by $\check{\boldsymbol{\xi}}$,

the set of all $\check{\boldsymbol{\xi}}$ that maintain the feasibility of the inner level of Problem (4.6) by $\check{\mathcal{I}}$, the vector of continuous decision variables in the inner level of (4.6) by $\check{\boldsymbol{\vartheta}}$, and the set of the constraints in in the inner level of (4.6) that involve continuous variables (for a given $\check{\boldsymbol{\xi}}$) by $\check{C}(\check{\boldsymbol{\xi}})$; then we can write Problem (4.6) as:

$$\max_{\tilde{d}\in\tilde{D}} \min_{\check{\boldsymbol{\xi}}\in\check{\mathcal{I}}} \min_{\check{\boldsymbol{\vartheta}}\in\check{C}(\check{\boldsymbol{\xi}})} \check{\mathscr{X}} \tag{4.7}$$

and since the inner-level problem above is a linear program, we can write it in standard form as $\min_{\check{\boldsymbol{\vartheta}}\in\check{C}(\check{\boldsymbol{\xi}})} \check{\mathscr{X}} \equiv \min_{\check{C}\check{\boldsymbol{\vartheta}}\leq\check{\mathbf{c}}_{\check{\boldsymbol{\xi}}}} \check{\boldsymbol{f}}'\check{\boldsymbol{\vartheta}}$ where $\check{\boldsymbol{f}}'\check{\boldsymbol{\vartheta}} = \check{\mathscr{X}}$. Then we invoke strong duality as shown in the proposition below.

Proposition 3. *Strong duality holds for the inner-level problem in (4.7), and hence,*

$$\min_{\check{\boldsymbol{\xi}}\in\check{\mathcal{I}}} \min_{\check{\boldsymbol{\vartheta}}\in\check{C}(\check{\boldsymbol{\xi}})} \check{\mathscr{X}} = \min_{\check{\boldsymbol{\xi}}\in\check{\mathcal{I}}} \max_{\check{C}'\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}=\check{\boldsymbol{f}},\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}\leq 0} \check{\mathbf{c}}'_{\check{\boldsymbol{\xi}}}\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}. \tag{4.8}$$

Proof. Please see Section 4.4. □

Let us denote by $\check{\boldsymbol{r}}$ the vector formed by concatenating all columns of $\mathscr{B}$.

Theorem 8. *If $\check{\Lambda}^*$ is the optimal value of the objective function in the optimization problem (4.6), then*

$$\begin{aligned}
\check{\Lambda}^* = \max_{\check{\eta}} \ & \check{\gamma} && (4.9)\\
s.t. \ & \check{\gamma} \leq \check{\mathcal{W}}_{\check{\boldsymbol{\xi}}} \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& \check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}} \leq 0 \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& \check{C}'\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}} = \check{\boldsymbol{f}} \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& 0 \geq \check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}} \geq -Y_{F,\check{\boldsymbol{r}}}(\check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}})N \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& \check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}} \leq (1 - Y_{F,\check{\boldsymbol{r}}}(\check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}}))N + Y_{F,\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}}(\check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}}) \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& \check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}} \geq (Y_{F,\check{\boldsymbol{r}}}(\check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}}) - 1)N + Y_{F,\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}}(\check{\boldsymbol{\Psi}}_{\check{\boldsymbol{\xi}}}) \ \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}} \\
& \tilde{d} \in \tilde{D}
\end{aligned}$$

where $\check{\mathcal{W}}_{\check{\xi}} = \mathcal{J}(\check{Z}, \check{\boldsymbol{r}}, \check{\boldsymbol{\lambda}}_{\check{\xi}}, \{\check{\boldsymbol{\Psi}}_{\check{\xi}}\}, F)$ *for* $\check{Z}(\check{\boldsymbol{r}}, \check{\boldsymbol{\lambda}}_{\check{\xi}}) \triangleq \check{\mathbf{c}}'_{\check{\xi}} \check{\boldsymbol{\lambda}}_{\check{\xi}}$, $\check{\boldsymbol{\Psi}}_{\check{\xi}} \in \mathbb{R}^{|(\mathcal{G}(Z, \check{\boldsymbol{r}}, \check{\boldsymbol{\lambda}}_{\check{\xi}})|}$, $\check{\eta} = \{\check{\gamma}, \tilde{d}\} \cup (\cup_{\check{\xi} \in \check{\mathcal{I}}}(\{\check{\boldsymbol{\lambda}}_{\check{\xi}}\} \cup \{\check{\boldsymbol{\Psi}}_{\check{\xi}}\}))$, *$F$ is a bijective map (as defined in Definition 8) from* $\mathcal{G}(\check{Z}, \check{\boldsymbol{r}}, \check{\boldsymbol{\lambda}}_{\check{\xi}})$ *to* $\{\check{\boldsymbol{\Psi}}_{\check{\xi}}\}$, *and N is the number of lines in the system (i.e. $N = |\mathcal{E}|$) pre-disturbance.*

Proof. Please see Section 4.4. □

Now that we have turned Problem (4.6) into a large-scale single-level MILP, it can be solved using a decomposition method involving column and row generation (e.g. see [62] and [63]), just like in the previous chapter. To do so, we again define a "decomposed master problem with input $\check{Q}$", which has the exact same form as (4.9) except that $\check{\mathcal{I}}$ shall be replaced by $\check{Q}$ throughout (4.9). Note that $\check{Q} \subseteq \check{\mathcal{I}}$ in the decomposition process, and hence, the term "decomposed". Let us denote by $\mathscr{B}^o$ the matrix containing binary variables corresponding to the binary expansion of the truncated approximation to $\mathbf{x}^o$. The column-and-row-generation procedure (adopted from [64]), tailored to Problem (4.6), is outlined in Algorithm 2.

Algorithm 2 Column and Row Generation for Problem (4.6)

1: Solve the inner-level of Problem (4.6) with $\mathscr{B}^o$ as input
2: $\check{Q} \leftarrow$ vector of the optimal values of all integer variables in Step 1
3: Initialize $UpperBound \leftarrow \infty$, $LowerBound \leftarrow -\infty$
4: **while** $UpperBound - LowerBound >$ tolerance **do**
5: Solve the decomposed master problem with input $\check{Q}$ and update $\mathscr{B}^*$
6: $\check{M}^* \leftarrow$ optimal value of objective function in Step 5
7: Solve the inner level of (4.6) with $\mathscr{B}^*$ as input
8: $\check{S}^* \leftarrow$ optimal objective function value in Step 7
9: $\check{\mathcal{Z}}^* \leftarrow$ vector of optimal values of all integer variables in Step 7
10: $\check{Q} \leftarrow \{\check{\mathcal{Z}}^*\} \cup \check{Q}$
11: $LowerBound \leftarrow \max(\check{S}^*, LowerBound)$
12: $UpperBound \leftarrow \check{M}^*$
13: **end while**

Same as in the previous chapter, using the results presented in [62] and [63] on the convergence of the column-and-row-generation procedure and the bounds we derived in Theorem 8 which gave us a valid single-level reformulation for forming the master problem, we have the following theorem.

Theorem 9. *Algorithm 2 converges in finite steps to the optimal solution of Problem (4.6).*

Proof. Please see Section 4.4. □

4.3 Upper Bound on Efficacy in Mitigating Any $N-k$ Scenario for a Given k

Consider the MILP in the inner-level of problem (4.6). If we turn $\mathscr{B}$ into a decision variable in this problem, the problem still remains as a MILP (because we have already linearized the product of its elements with the elements of $\mathbf{I}^t$). Let us denote the vector of the decision variables of this new problem by $\tilde{\mathscr{A}}_{\mathscr{B}}$ and its feasible set by $\tilde{A}_{\mathscr{B}}$. Then, we define the efficacy of pre-disturbance reactance tweaking in mitigating a disturbance (denoted by $\tilde{E}(\check{\kappa})$) as the number of lines that would otherwise have failed during the cascading failures had we not optimally tweaked the reactances prior to the disturbance $\check{\kappa}$. It follows that based on the approximate MILP technique presented in this chapter, the efficacy of pre-disturbance reactance tweaking in mitigating any $\check{\kappa}$ in the set of all $N-k$ scenarios for a given k is "approximately" upper bounded (i.e. upper bounded in the approximate sense presented in this chapter) as follows:

$$\tilde{E} \leq \max_{\hat{v}} \sum_{(i,j,h)\in\mathcal{E}} s^{T+2}_{ijh,\mathscr{B}} - \sum_{(i,j,h)\in\mathcal{E}} s^{T+2}_{ijh,o} \tag{4.10}$$

$$\text{s.t.} \quad \mathbf{s}^1_o = \mathbf{s}^1_{\mathscr{B}}$$

$$\mathscr{B}2^{[LU]} - \mathscr{B}^o 2^{[LU]} - diag(\check{\boldsymbol{\tau}})(\mathbf{x}^{\max} - \mathscr{B}^o 2^{[LU]}) \leq 0$$

$$-\mathscr{B}2^{[LU]} + \mathscr{B}^o 2^{[LU]} + diag(\check{\boldsymbol{\tau}})(\mathbf{x}^{\min} - \mathscr{B}^o 2^{[LU]}) \leq 0$$

$$\sum_{(i,j,h)\in\mathcal{E}} \check{\tau}_{ijh} = H \;\forall (i,j,h) \in \mathcal{E} \qquad (S2)$$

$$\mathscr{A}_o \in \tilde{A}(\mathscr{B}^o); \quad \tilde{\mathscr{A}}_{\mathscr{B}} \in \tilde{A}_{\mathscr{B}}; \quad \check{\tau}_{ijh} \in \{0,1\}, \quad \forall (i,j,h) \in \mathcal{E}.$$

where $\hat{v} = \{\tilde{\mathscr{A}}_{\mathscr{B}}, \tilde{\mathscr{A}}_o, \check{\tau}\}$, and $\tilde{A}(\mathscr{B}^o)$ denotes the feasible set of the inner level of Problem (4.4) with input $\mathscr{B}^o$ (recall that $\mathscr{B}^o$ is the matrix containing binary constants corresponding to the binary expansion of the truncated approximation of the original line reactances $\mathbf{x}^o$), and $\tilde{\mathscr{A}}_o$ denotes an exact copy of the vector of decision variables in the inner level of Problem (4.4). Thus, note that in the above formulation, $\mathscr{B}^o$ is a constant but $\mathscr{B}$ is a decision variable.

4.4 Proofs

Proof of Theorem 7. The proof of this theorem follows the exact same procedure as he proof of Theorem 3 (i.e. by first shifting the nonconvexity from continuous×continuous bilinear term to absolute values via a change of variables, and then linearizing the absolute values using standard techniques), and hence, the reader is referred to the proof of Theorem 3 for details on the derivation. Note that in this theorem $\check{\mu}_{ijh}$ is equivalent to the continuous×continuous bilinear term $I^o_{ijh}\check{u}_{ijh}$. □

Proof of Proposition 3. To verify that strong duality holds, note that the innermost-level problem in (4.7) is bounded and feasible (because by assumption, $\check{\mathcal{I}}$ only contains the vectors of integer variables that maintain the feasibility of the problem, and a disturbance of size zero (i.e. no disturbance) is always a feasible solution for the inner-level problem, so the feasible set is not empty); hence, strong duality holds and deriving the equations for the dual problem stated in the proposition is standard (e.g. see [8] for details on deriving the dual of a linear program). □

Proof of Theorem 8. This proof follows similar steps as in the proof of Theorem 4, but for completeness we present this proof below. Note that so far we have reformulated the max-min problem (4.6) as a bilevel program in which the inner-level and the outer-level are both MILPs coupled through only binary variables (i.e. through the entries of $\mathscr{B}$). Using the result we just established

in Proposition 3, we can write (4.6) as:

$$\max_{\tilde{d}\in\tilde{D}} \min_{\check{\xi}\in\check{\mathcal{I}}} \max_{\check{C}'\check{\lambda}_{\check{\xi}}=\check{f},\check{\lambda}_{\check{\xi}}\leq 0} \check{c}'_{\check{\xi}}\check{\lambda}_{\check{\xi}} \tag{4.11}$$

Next, we create a copy of $\check{\lambda}$ for each element $\check{\xi} \in \check{\mathcal{I}}$ (each copy will be treated as a distinct decision variable) and denote it by $\check{\lambda}_{\check{\xi}}$. Then, we enumerate over all elements of $\check{\mathcal{I}}$ (to eliminate the minimization level (i.e. the mid-level) of the above three-level problem), merge the two maximization levels, and rewrite the above as a large-scale single-level program (note that this step is a standard technique, see e.g. [64]):

$$\begin{aligned} \max_{\check{\eta}} \quad & \check{\gamma} \\ \text{s.t.} \quad & \check{\gamma} \leq \check{c}'_{\check{\xi}}\check{\lambda}_{\check{\xi}} \; \forall \check{\xi} \in \check{\mathcal{I}} \\ & \check{\lambda}_{\check{\xi}} \leq 0 \; \forall \check{\xi} \in \check{\mathcal{I}} \\ & \check{C}'\check{\lambda}_{\check{\xi}} = \check{f} \; \forall \check{\xi} \in \check{\mathcal{I}} \\ & \tilde{d} \in \tilde{D} \end{aligned} \tag{4.12}$$

where $\check{\eta} = \{\check{\gamma}, \tilde{d}\} \cup (\cup_{\check{\xi}\in\check{\mathcal{I}}}\{\check{\lambda}_{\check{\xi}}\})$. However, this single-level program is not a MILP yet, because in the constraints $\check{\gamma} \leq \mathbf{c}'_{\check{\xi}}\check{\lambda}_{\check{\xi}}$ we have some binary×continuous bilinear terms stemming from the product of some elements of $\mathscr{B}$ and some elements of $\check{\lambda}_{\check{\xi}}$ for all $\check{\xi} \in \check{\mathcal{I}}$. To do so, we need a bound on the magnitude of the dual variables (i.e. $\check{\lambda}_{\check{\xi}}$) of our innermost level LP in (4.11). We claim that this upper bound is equal to N (i.e. the number of lines in the network pre-disturbance); to verify this claim, observe that the LP in the innermost level of (4.11) minimizes the number of active lines post-disturbance; by definition, the dual variable corresponding to a constraint quantifies the amount of change in the optimal value of the objective function if we increase the bound in that constraint by one unit. Naturally, increasing the limit of any constraint in the problem can result in at most a change of N in the objective, because no matter what, we cannot change the number of active lines in the system by more than N. Now that we have a bound on the magnitude of the continuous variables that appear in bilinear terms, we can replace any bilinear term by an auxiliary

variable (we denote the vector of these auxiliary variables by $\check{\mathbf{\Psi}}_{\check{\xi}}$), and to complete the linearization and ensure that this auxiliary variable will be equal to the bilinear term we further enforce:

$$0 \geq \check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}} \geq -Y_{F,\check{\boldsymbol{r}}}(\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}})N \; \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}}$$
$$\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}} \leq (1 - Y_{F,\check{\boldsymbol{r}}}(\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}}))N + Y_{F,\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}}(\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}}) \; \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}}$$
$$\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}} \geq (Y_{F,\check{\boldsymbol{r}}}(\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}}) - 1)N + Y_{F,\check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}}}(\check{\mathbf{\Psi}}_{\check{\boldsymbol{\xi}}}) \; \forall \check{\boldsymbol{\xi}} \in \check{\mathcal{I}}$$

where $Y_{.,.}$ is defined in Definition 8 and F is a bijective map (as defined in Definition 8) from $\mathcal{G}(\check{Z}, \check{\boldsymbol{r}}, \check{\boldsymbol{\lambda}}_{\check{\boldsymbol{\xi}}})$ to $\{\check{\mathbf{\Psi}}_{\check{\xi}}\}$. □

Proof of Theorem 9. Since the Master Problem has a less-constrained feasible set than the single-level MILP reformulation of the original max-min problem but the same objective, it gives an upper bound on the globally optimal solution of the original problem (keep in mind that the Master Problem is a maximization problem, and hence, a less-constrained feasible set means a larger optimal objective). Conversely, the optimal value of the subproblem gives a lower bound on the optimal value of the original max-min problem. Given strong duality in our reformulation for the continuous variables of the inner-level MILP (i.e. Proposition 3) and our single-level MILP reformulation for forming the master problem through the bounds we derived in Theorem 8, and the finiteness of the number of elements in $\check{\mathcal{I}}$, the proof follows the exact same procedure as the techniques of [62], [63], [64] cited in the proof of Theorem 5 earlier in this thesis, with the only difference that this time the Master Problem gives an upper bound while the subproblem gives the lower bound (which is the opposite of the scenario in the proof of Theorem 5). □

4.5 Numerical Experiments

Our numerical experiments for a nominal load-flow scenario in the IEEE One-Area RTS-1996 test case system [36] reveal that pre-disturbance reactance tweaking on up to two transmission lines cannot reduce the number of lines that will fail at $t = 2$ when the adversary seeks to cause a

worst-case-scenario disturbance by removing up to two lines at time $t = 1$. For the same set of simulation parameters, the "approximate" upper bound on the efficacy is only 1 line, i.e. when we are allowed to tweak the reactances of up to two lines at $t = 0$, we can prevent the failure of at most one line at $t = 2$ among the set of all $N - 2$ disturbances that can be initiated at time $t = 1$; however, if we keep all parameters the same but this time consider the set of all $N-3$ disturbances, the simulation results show that we can save up to three lines from failing at $t = 2$ if we tweak the reactances of up to two lines at $t = 0$. Note that the numerical experiments were performed in MATLAB, and Gurobi [37] was used for solving the MILPs. In all the simulations, the operating range of the reactance tweaking device was set to the operating range reported for TCSC in [33].

Chapter 5

Conclusions, Discussion, and Future Work

In this thesis we studied transmission line reactance tweaking as a control mechanism for maximizing the yield post-disturbance and also as a pre-disturbance mechanism for mitigating the propagation of cascading failures. For post-disturbance control, we developed optimization formulations for assessing the efficacy of this mechanism using two different measures of system resilience. We gave a rigorous MILP reformulation scheme for the underlying bilevel nonconvex MINLP to facilitate the global optimization of the min-max problem corresponding to the post-disturbance response of a system operator (equipped with reactance tweaking) to a worst-case-scenario disturbance. We also derived a MILP reformulation for computing an exact upper bound on the amount of load saved post-disturbance as a pure contribution of reactance-tweaking in dealing with any $N-k$ contingency for a given k. A main feature of our model is its ability to track the multi-stage evolution of cascading failures before the SO's post-disturbance response, which as shown in our numerical experiment, can make a significant difference in the system yield. Our numerical case study suggests that reactance tweaking on only a small number of transmission lines can, in some scenarios, considerably reduce the amount of load shed post-disturbance in the tested system. As for pre-disturbance resilience enhancement, we developed a MILP approximation to the bi-level MINLP problem that reveals to us whether reactance-tweaking can reduce the number of lines that will fail over the propagation of cascading failures in the event of a worst-case-scenario $N-k$

disturbance. We also gave a MILP formulation for computing an approximate upper bound on the number of lines that would otherwise have failed over the course of a cascading-failures event had we not tweaked the line reactances pre-disturbance. Our numerical case study suggests that pre-disturbance tweaking of the reactances of only two lines can, in some scenarios, prevent the overloading of multiple lines after an $N-3$ disturbance in the tested system.

The economic viability, engineering design, and implementation of a post-disturbance reactance-tweaking technology would all be interesting future research problems, although TCSC seems to give a good starting point as a candidate device for such technology. Naturally, a reactance tweaking device would have applications way beyond post-disturbance control (some examples of the applications of TCSC were mentioned in Chapter 1); hence, the problem of optimal placement of such device that would also take into account its application in post-disturbance control among its other applications appears to be an interesting direction for future research. In particular, post-disturbance reactance tweaking seems to potentially be closely connected to the idea of congestion management via reactance tweaking; hence, one interesting area for future research appears to be the development of a framework for studying the connections between these two potential applications of reactance tweaking, possibly with a focus on whether the problems of optimal placement of this device for each of these two applications are interrelated. The tools developed in this thesis for studying post-disturbance reactance tweaking (in Chapter 3) mainly focused on devising tractable optimization frameworks for assessing the efficacy of reactance tweaking in a two-agent game setup and for computing an upper bound on the efficacy of this technology. Although the problem of optimal placement of reactance-tweaking devices for post-disturbance control was not a topic of interest in this thesis as this problem merits its own line of research for the reasons mentioned above, our numerical experiments suggested that for the base load case in the One-Area RTS-1996 system, there are multiple desirable locations for placing reactance-tweaking device(s) in response to $N-3$ contingencies. A major future direction for research could be to develop a systematic and comprehensive framework for identifying optimal locations for placement of reactance tweaking devices for post-disturbance control purposes, for instance by developing fast sampling techniques that allow for randomizing over various system parameters (especially, over bus injections and dis-

turbance size). Our numerical experiments in Chapter 3 suggested that in the One-Area RTS-1996 base load case, the line(s) that appear as the optimal locations for reactance tweaking in the computation of the upper bound (derived in Theorem 6) on "general efficacy" were also the lines that eventually emerged as the "optimal" (in the quasi-exhaustive averaging sense presented in Section 3.6.3) placement locations for reactance tweaking in response to $N-3$ contingencies; although this observation may not extend to systems other than One-Area RTS-1996, the bound that can be computed via Theorem 6 may serve future research on this topic by providing a fast initial estimate of the lines on which reactance tweaking could be most useful.

The results presented in Chapter 4 on pre-disturbance reactance tweaking are mainly meant to just scratch the surface and serve as a starting point to motivate future research on this topic. One of the computational results in Chapter 4 revealed that for the base load case of One-Area RTS-1996, there exists an N-3 contingency for which pre-disturbance reactance tweaking on only two lines can prevent the failure of three lines in the system as cascading failures propagate for one stage after that particular disturbance. Generally speaking, if the system operator expects that a particular contingency is imminent, pre-disturbance reactance tweaking may potentially be an effective strategy to mitigate the propagation of cascading failures. Some of the existing literature on pre-disturbance reactance tweaking were discussed in Chapter 1; however, future directions in this area seem abundant. To give an example, a major future step would be to devise a three-level optimization setup in which the system operator tweaks reactances pre-disturbance, then the adversary causes cascading failures by choosing from a set of possible disturbance scenarios, and finally the system operator acts again by re-dispatching the system post-disturbance with reactance tweaking as a control mechanism. This could provide a much broader idea regarding the efficacy of reactance tweaking in response to cascading failures. However, the issue of computational complexity poses a major challenge in this problem and dealing with it merits its own line of work, but it is hoped that the tools developed in this thesis provide a solid starting point for such research.

Once all these have been addressed, more detailed models that involve AC power flows, and hence can realistically take into account disturbances that cause imbalances between load and generation

for $T \geq 1$, would be quite useful. Combining the post-disturbance reactance-tweaking model with line switching would also be an interesting direction for future research to evaluate how these two methods can possibly complement each other. Another interesting path is to study the efficacy of reactance tweaking (and its optimal placement) in response to geographically correlated disturbances. Other potential extensions of the model include studying the impact of probabilistic line overloads on the evolution of cascading failures (especially, such model should seek to capture the non-uniformity in the time it takes different lines to trip depending on the level of their overload, and also should not assume that an overloaded line trips deterministically and instead should assign some randomness to the tripping of an overloaded line). Eventually, a model of reactance-tweaking efficacy that can also capture continuous-time dynamics of cascading failures would be a highly valuable future research direction.

A deeper look into the structure of the problem reminds us of the non-triviality of the interplay between graph topology, bus injections, line impedances, and line capacities in determining the resilience of the system to cascading failures. This interplay significantly complicates analytical characterization of many resilience assessment problems related to cascading failures in power systems. One direction for future research could be to focus on the impact of graph topology. Given the same graph topology, if we change the distribution of bus injections, line impedances, and line capacities across the network, the resilience of the system could dramatically change; thus, the picture does not seem to be ever complete without taking into account all the above-mentioned parameters that control the redistribution of flows in the system. However, by studying simple graph structures, and by allocating all other parameters systematically across the given network topology, it may be possible to gain some insight into the "average" impact of network topology on the efficacy of post-disturbance reactance tweaking (e.g. one could study whether graphs with higher clustering coefficients, on average, benefit more from reactance tweaking compared to graphs with lower clustering coefficients; or, one could study whether graphs with higher clustering coefficients, on average, require fewer reactance-tweaking devices). Some of the tools developed in this thesis can serve as building blocks for computational efforts in such future research directions. It has been suggested in [29] that a small subset of lines appear in a significantly larger number of

blackout scenarios than the majority of lines in the system. This suggests that there is a small number of lines in the system that are expected to be most critical in compromising the resilience of the system. With that in mind, another direction for future research could be to study whether the techniques developed by the literature focusing on the identification of these critical lines can be used to come up with fast techniques for identifying both the set of disturbances that reactance tweaking is not effective in responding to, and the set of disturbances that reactance tweaking is highly effective in responding to. More specifically, can we develop efficient algorithms for quickly identifying a set of critical lines whose failure is effectively dealt with via reactance tweaking? That may then allow for studying possible connections between graph topology, the number of critical lines in the network, and the portion of critical lines whose failure is effectively responded to via reactance tweaking.

Bibliography

[1] N. Acharya and N. Mithulananthan, "Influence of TCSC on congestion and spot price in electricity market with bilateral contract," *Elect. Power Syst. Res.*, vol. 77, pp. 1010-1018, 2007.

[2] J. M. Arroyo, "Bilevel programming applied to power system vulnerability analysis under multiple contingencies," *IET Gener Transm. Distrib.*, vol. 4, no. 2, pp. 178-190, 2010.

[3] J. M. Arroyo and F. Fernandez, "A genetic algorithm approach for the analysis of electric grid interdiction with line switching," in *Proc. IEEE 15th Int. Conf. Intelligent System Applications to Power Systems* (ISAP-09), pp. 1-6, 2009.

[4] C. Asavathiratham, S. Roy, B. Lesieutre, and G. Verghese, "The influence model," *Control Systems, IEEE*, vol.21, no.6, pp.52,64, 2001.

[5] R. Bent, A. Berscheid, A.L. Toole, "Generation and transmission expansion planning for renewable energy integration," In: *Proc. of Power Syst. Comp. Conf. (PSCC)*, Stockholm, Sweden, August 2011.

[6] A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2^{nd} Ed., Upper Saddle River, NJ: Prentice Hall, 2000.

[7] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman. "Power grid vulnerability to geographically correlated failures-analysis and control implications," *arXiv preprint*, arXiv:1206.1099, 2012.

[8] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.

[9] D. Bienstock and S. Mattia, "Using mixed-integer programming to solve power grid blackout problems", *Discrete Optimization*, vol. 4, pp. 115-141, 2007.

[10] D. Bienstock, and A. Verma. "The n-k problem in power grids: New models, formulations, and numerical experiments." *SIAM Journal on Optimization*, vol. 20, no. 5, pp. 2352-2380, 2010.

[11] D. Bienstock, "Optimal control of cascading power grid failures," In: *Proc. of IEEE CDC-ECC*, 2011.

[12] V.M. Bier, E.R. Gratz, N.J. Haphuriwat, W. Magua, K.R. Wierzbickiby, "Methodology for identifying near-optimal interdiction strategies for a power transmission system," *Reliability Engineering and System Safety*, vol. 92, 1155-1161, 2007.

[13] R. Billinton, M. Fotuhi-Firuzabad, and S. Faried, "Power system reliability enhancement using a thyristor-controlled series capacitor," *IEEE Transactions on Power Systems*, vol. 14, pp. 369-374, 1999.

[14] S. Binato, M. Pereira, S. Granville, "A new Benders decomposition approach to solve power transmission network design problems," *IEEE Transactions on Power Systems* , vol.16, no.2, pp.235-240, 2001.

[15] G. G. Brown, W. M. Carlyle, J. Salmeron, and K. Wood, "Analyzing the vulnerability of critical infrastructure to attack and planning defenses," in *Tutorials in Operations Research*. INFORMS, pp. 102-123, 2005.

[16] S. Burer and A. Letchford. "Non-convex mixed-integer nonlinear programming: A survey," *Surveys in Operations Research and Management Science*, vol. 17, no. 2, pp. 97-106, 2012.

[17] C. Canizares and Z. Faur, "Analysis of SVC and TCSC controllers in voltage collapse," *IEEE Transactions on Power Systems*, vol. 14, pp. 158-165, Feb. 1999.

[18] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, "Complex dynamics of blackouts in power transmission systems", *Chaos*, vol. 14, no. 3, pp. 643-652, 2004.

[19] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos*, vol. 12, no. 4, pp. 985-994, 2002.

[20] B.A. Carreras, D.E. Newman, I. Dobson, A.B. Poole, "Evidence for self organized criticality in electric power system blackouts", *IEEE Transactions on Circuits and Systems I*, vol. 51, no. 9, pp. 1733- 1740, 2004.

[21] M. Carrion, J.M. Arroyo, N. Alguacil, "Vulnerability-constrained transmission expansion planning: A stochastic programming approach," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1436-1445, 2007.

[22] D. P. Chassin and C. Posse, "Evaluating north american electric grid reliability using the Barabsi-Albert network model," *Physica A:Statistical Mechanics and its Applications*, vol.355, no. 2-4, pp. 667 - 677, 2005.

[23] R. Chen, A. Cohn, N. Fan, and A. Pinar, "N-k-epsilon survivable power system design," In: *Proc. PMAPS*, 2012.

[24] J. Choi, T.D. Mount, R.J. Thomas, "Transmission expansion planning using contingency criteria," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2249-2261, 2007.

[25] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Physical Review E*, vol. 69, no. 4, pp. 045104- 1-045104-4, 2004.

[26] A. Delgadillo, J. Arroyo, and N. Alguacil, "Analysis of electric grid interdiction with line switching," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 633-641, 2010.

[27] I. Dobson, B.A. Carreras, and D.E. Newman. "A loading-dependent model of probabilistic cascading failure." *Probab Eng Inform Sci* vol. 19, no.1, pp. 15-32, 2005.

[28] V. Donde, V. Lopez, B. Lesieutre, A. Pinar, C. Yang, and J. Meza, "Identification of severe multiple contingencies in electric power systems," *IEEE Transactions on Power Systems*, vol. 23, pp. 406-417, 2008.

[29] M. J. Eppstein and P. D. H. Hines. "A "Random Chemistry" Algorithm for Identifying Collections of Multiple Contingencies that Initiate Cascading Failure." *IEEE Transactions on Power Systems*, vol. 27, no. 3, 2012.

[30] N. Fan, H. Xu, F. Pan, P.M. Pardalos, "Economic analysis of the N-k power grid contingency selection and evaluation by graph algorithms and interdiction methods," *Energy Syst.,* vol. 2, no. 3-4, pp. 313-324, 2011.

[31] R. Fitzmaurice, *Cascading Failure in a Complex System Model for Power Systems: Operating and Planning Policy,* PhD Thesis, University College Dublin, 2010.

[32] B. Geißler, A. Martin, A. Morsi, and L. Schewe. "Using piecewise linear functions for solving MINLPs," In: *Mixed Integer Nonlinear Programming, The IMA Volumes in Mathematics and its Applications*; eds.: J. Lee and S. Leyffer, pp-287-314, Springer, New York, 2012.

[33] S. Gerbex, R. Cherkaoui, and A. J. Germond, "Optimal location of multi-type FACTS devices in a power system by means of genetic algorithms," *IEEE Transactions on Power Systems*, vol. 16, no. 3, pp.537-544, 2001.

[34] A. Ghosh, S. Boyd, and A. Saberi, "Minimizing effective resistance of a graph", *SIAM Review*, vol. 50, pp. 37-66, 2008.

[35] G. Glanzmann and G. Andersson, "Incorporation of N-1 security into optimal power flow for FACTS control," In: *Power Systems Conference and Exposition, IEEE PES*, pp. 683-688, 2006.

[36] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, C. Singh, "The IEEE Reliability Test System-1996. A report prepared by the

Reliability Test System Task Force of the Application of Probability Methods Subcommittee," *IEEE Transactions on Power Systems*, vol.14, no.3, pp.1010-1020, 1999.

[37] Gurobi Optimization, Inc. "Gurobi Optimizer Reference Manual Version 6.0.2." Houston, Texas: Gurobi Optimization, 2015.

[38] IEEE RTS-96 system diagram, available at http://www.ee.washington.edu/research/real/Library/Data/IEEE-RTS_3.jpg.

[39] R. A. Jabr, A. H. Coonick, B. J. Cory, "A homogeneous linear programming algorithm for the security constrained economic dispatch problem," *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp.930-936, 2000.

[40] U. Janjarassuk and J. Linderoth, "Reformulation and sampling to solve a stochastic network interdiction problem," *Networks*, vol. 52, no. 3, pp. 120-132, 2008.

[41] S. Jin, S.M. Ryan, J.-P. Watson, D.L. Woodruff, "Modeling and solving a large-scale generation expansion planning problem under uncertainty," *Energy Syst.*, vol. 2, no. 3-4, pp. 209-242, 2011.

[42] D. Kempe, J. Kleinberg, E. Tardos. "Maximizing the spread of influence through a social network." In Proc. 9th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 2003.

[43] C. Lai and S. Low. "The redistribution of power flow in cascading failures", In: *Communication, Control, and Computing, 51st Annual Allerton Conference on*, 2013.

[44] Y. Lu and A. Abur, "Static security enhancement via optimal utilization of thyristor-controlled series capacitor," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 324-329, 2002.

[45] A. Motter, "Cascade-based attacks on complex networks", *Physical Review E*, vol. 66, no. 6, 2002.

[46] A. Motter, "Cascade control and defense in complex networks," *Phys Rev Lett*, 93, 2004.

[47] A. Motto, J. Arroyo, and F. Galiana, "A mixed-integer LP procedure for the analysis of electric grid security under disruptive threat," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1357-1365, 2005.

[48] L. Moulin, M. Poss, C. Sagastizabal, "Transmission expansion planning with re-design," *Energy Syst.*, vol. 1, no. 2, pp. 113-139, 2010.

[49] G. C. Oliveira, S. Binato, L. Bahiense, L. Thome, and M.V. Pereira, "Security-constrained transmission planning: A mixed-integer disjunctive approach", In: *Proc. IEEE/Power Eng. Soc. Transmission and Distribution Conf.*, Sao Paulo, Brazil, 2004.

[50] G. A. Pagani and M. Aiello, "The power grid as a complex network: a survey," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 11, pp. 2688 - 2700, 2013.

[51] J. Paserba, N. Miller, E. Larsen, and R. Piwko, "A thyristor controlled series compensation model for power system stability analysis," *IEEE Transactions on Power Delivery*, vol. 10, pp. 1471-1478, 1995.

[52] A. Pinar, J. Meza, V. Donde, and B. Lesieutre, "Optimization strategies for the vulnerability analysis of the electric power grid," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1786-1810, 2010.

[53] A. Pinar, A. Reichert, and B.C. Lesieutre, "Computing criticality of lines in power systems," In: *IEEE Int. Symp. Circuits and Systems (ISCAS)*, New Orleans, LA, 65 - 68, 2007.

[54] D. Radu, and Y. Besanger, "Blackout prevention by optimal insertion of FACTS devices in power systems," In: *Future Power Systems, 2005 International Conference on*, pp. 1-6, 2005.

[55] J. Salmeron, K. Wood, R. Baldick, "Worst-case interdiction analysis of large-scale electric power grids," *IEEE Transactions on Power Systems*, vol. 24, no.1, pp. 96-104, 2009.

[56] S. N. Singhand and A. K. David,"Optimal location of FACTS devices for congestion management," *Elect. Power Syst. Res.*, vol. 58, no. 2, pp. 71-79, 2001.

[57] A. Street, F. Oliveira, J.M. Arroyo, "Contingency-constrained unit commitment with n-k security criterion: A robust optimization approach," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1581-1590, 2011.

[58] P. K. Tiwari and Y. R. Sood, "An efficient approach for optimal allocation and parameters determination of TCSC with investment cost recovery under competitive power market," *IEEE Transactions on Power Systems*, vol.28, no.3, pp. 2475-2484, 2013.

[59] P. A. Trodden, W. A. Bukhsh, A. Grothey, & K.I. McKinnon, "Optimization-based islanding of power networks using piecewise linear AC power flow." *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1212 - 1220, 2014.

[60] D. J. Watts, "A simple model of global cascades on random networks," *PNAS*, vol. 99, no. 9, pp. 5766-5771, 2002.

[61] N. Yang, Q. Liu, and J. D. McCalley, "TCSC control design for damping inter-area oscillations," *IEEE Transactions on Power Systems*, vol. 13, pp. 1304-1310, Nov. 1998.

[62] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, Vol. 41, no. 5, pp. 457-461, 2013.

[63] L. Zhao and B. Zeng, "An exact algorithm for two-stage robust optimization with mixed integer recourse problems," Univ. South Florida, 2011, available in optimization-online.

[64] L. Zhao and B. Zeng, "Vulnerability analysis of power grids with line switching," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2727-2736, 2013.

[65] W. Zhu, R. Spee, W. A. Mittelstadt, and D. Maratukulam, "An EMTP study of SSR mitigation using the thyristor controller series capacitor," *IEEE Transactions on Power Delivery*, vol. 10, pp. 1479-1485, 1995.