The Dissertation Committee for Yuqi Zhou
certifies that this is the approved version of the following dissertation:

Real-time Grid Topology Modeling and Optimization for Power Transmission Systems

Committee:

Hao Zhu, Supervisor

Surya Santoso

Constantine Caramanis

Ufuk Topcu

Grani Hanasusanto

Real-time Grid Topology Modeling and Optimization for Power Transmission Systems

by

Yuqi Zhou

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2022

To my family.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my advisor, Dr. Hao Zhu. I am very fortunate to have worked with Dr. Zhu over the past few years. She had a very profound influence on me as a researcher, and I cannot thank her enough for her guidance and support.

My gratitude extends to my committee members who have offered generous help along the journey. I would like to thank Dr. Surya Santoso for his knowledge and guidance in the fields of power system operations and control. I am grateful to Dr. Grani Hanasusanto, for his remarkable teaching on optimization under uncertainty. I thank Dr. Constantine Caramanis and Dr. Ufuk Topcu for their constructive advice on my dissertation.

I would also like to thank the professors and researchers with whom I have collaborated. I am very grateful to my master's advisor Dr. Le Xie, for his guidance during my early stage as a researcher. I would want to thank Dr. Ahmed S. Zamzam and Dr. Andrey Bernstein from National Renewable Energy Laboratory for a wonderful summer internship during the pandemic. I thank Dr. Chaitanya Baone and Dr. David Ganger for the internship at Eaton research lab, working on the load forecasting project. I would also want to thank Dr. Deepjyoti Deka and Dr. Kaarthik Sundar from Los Alamos National Laboratory for collaborating on the power system resilience project.

I am indebted to my fellow students and friends from the University of Texas at Austin. I am especially grateful to Shanny Lin, Shaohui Liu, Kyung-bin Kwon, Jeehyun Park, Pablo Paz Salazar, Young-ho Cho, Fabricio Espinoza, Ayat Albuali, Bing Huang, Bowen Hua, Manuel Garcia, Quan Nguyen, Dawei Liang, for all the great memories that we have shared. I would also want to thank my friends and colleagues outside of UT Austin, Xinbo Geng, Bainan Xia, Yuanyuan Li, Tong Huang, Xiaolin Chen, Sadegh Modarresi, Xiangtian Zheng, Hao Huang, Yijing Liu, Yize Chen, Yuanyuan Shi, Ling Zhang, Wenqi Cui, Jip Kim and many more.

Last but not least, I would like to thank my parents. This dissertation would not be possible without their love and support. I dedicate this dissertation to them.

Real-time Grid Topology Modeling and Optimization for Power Transmission Systems

Publication No. _____

Yuqi Zhou, Ph.D.
The University of Texas at Austin, 2022

Supervisor: Hao Zhu

Accurately modeling and tactfully switching power grid topology are not only crucial for routine power system operational tasks but also play a critical role in system emergency responses under extreme events. The modern power grids have recently witnessed more frequent occurrences of unintentional topology changes. These changes can be caused by misoperations of substation protection systems, malicious cyber attacks, or natural disasters. Although strategically altering the grid topology through transmission switching can effectively relieve network congestion and thus has the potential to mitigate the impact of these events, the optimal decision is in general difficult to attain due to the uncertainty and variability therein. Therefore, this motivates us to devise efficient algorithms for achieving real-time power grid topology monitoring and optimization.

This dissertation first focuses on efficient modeling and monitoring of the bus split event, which is a type of grid topology change caused by circuit breakers in substations. We perform sensitivity analysis to evaluate the grid-wide impact of such events under the bus-branch representation, for which a synchrophasor data enabled identification problem is presented by matching the changes in the measurements. Inspired by this, we next explore the transmission switching problem that can incorporate the substation-level topology changes. Furthermore, to perform reliable and cost-effective transmission switching under the renewable uncertainty, we study the distributionally robust chance-constrained problem, which can provide superior robustness guarantees over the traditional chance-constrained formulation. Finally, to provide effective system responses under extreme weather events, we will also investigate scalable optimization and learning algorithms for quick power grid restoration.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Network topology plays an increasingly important role in power system operations. In particular, this can broadly cover system component protection and maintenance, daily transmission switching operations, long-term transmission planning, and even mitigation of natural disasters. For instance, network topology errors caused the 2003 Northeast blackout [260] in the United States, which was the world's second most widespread blackout in history. Malicious cyber intrusions are another example of how incorrect network topology information can pose threat to reliable power system operations, with a vivid example of the 2015/16 Ukrainian blackouts. On the other hand, the regional transmission organization (such as PJM [24]) can greatly benefit from strategically altering the network topology. The operations of transmission switching can not only reduce system operating costs but also relieve network congestion. Furthermore, as natural disasters are becoming a rising concern for reliable power system operations in recent years, topology control has also been used as an effective countermeasure to mitigate the impact of extreme events (e.g., [248]) to the power grid to prevent cascading failures.

To address these concerns and to provide more efficient topology control strategies, we will explore different aspects of topology monitoring and optimization in power system operations in this dissertation. In terms of system protection and maintenance, circuit breakers are often switched on/off to alter substation topology to isolate system components during maintenance. The configuration of the substation and circuit breaker status can affect the substation connectivity and thus have an impact on the overall power system topology. Misoperations and malicious compromises of circuit breakers can vary the grid topology, leading to bus faults or even bus split contingencies [124]. Modeling and monitoring of bus split events are in general difficult due to the change in the network topology. Capturing the circuit breaker statuses and substation connectivity can substantially increase the problem size and computational time. Due to these limitations, the problem of efficiently analyzing and identifying bus split events remains an open research area. Thus, we aim to develop an efficient modeling and monitoring framework for bus split events by incorporating them into the concise bus-branch representation.

Furthermore, tactfully switching the network topology can effectively reduce generation costs and relieve network congestion [93]. In recent years, the optimal transmission switching (OTS) problem has attracted interest in both algorithm design and practical implementations [87, 122, 256]. Increasing renewable integration in recent years challenges the efficient and reliable operations of power systems. As the OTS problem includes additional integer

decision variables on the transmission line status, it is more challenging to extend it to a stochastic/robust optimization framework compared to the optimal power flow (OPF) problem. Existing research works [74, 208, 233] have considered sample-based methods, which can cause the problem dimensions to quickly grow, making these approaches deficient for real-time OTS implementations. To this end, we aim to develop computationally efficient OTS algorithms that can provide robustness guarantees for power systems under uncertain renewables.

As extreme events such as natural disasters have greatly challenged reliable operation of power systems in recent years, utilizing topology control to enhance power system resilience has become an increasingly popular research area. Strategically reconfiguring the network topology and performing load shedding under contingencies can help mitigate the impact of these events to the power grid. To this end, we aim to develop efficient topology control solutions for real-time mitigation of natural disasters such as wildfires. This will enable quick system restoration under the contingencies and prevent potential cascading failures and even system blackouts.

Given the fast, complex process of cascading propagation, corrective actions such as optimal load shedding (OLS) are generally difficult to attain in large-scale power systems due to both computation complexity and communication latency. Therefore, we explore the potential of utilizing a scalable learning-for-OLS approach such that individual load centers can predict their own optimal decisions in a decentralized fashion. The scalable design of OLS

decision rules will enable each load center to quickly react to the contingency situations without requiring supervision from the control center, thus providing more prompt responses to unexpected system contingencies.

This dissertation is organized as follows: Chapter 2 focuses on the modeling and identification of the bus split event. We start with performing the linear sensitivity analysis for quick contingency impact evaluation by using matrix computation. After that, we further incorporate it into a security-constrained economic dispatch problem, which can help guarantee secure system operations against the bus split contingency. Particularly, the sensitivity analysis can also benefit the monitoring of bus split events. To identify the system topology under the bus split events using high-resolution synchrophasor data, we formulate it as a mixed integer optimization problem, which can achieve real-time implementation through tractable reformulation. Chapter 3 presents the OTS problem under uncertainty. To incorporate comprehensive substation level topology changes such as bus splits for improved operational efficiency, we first present an OTS problem which is modeled under the concise bus-branch model by utilizing the sensitivity analysis from the previous chapter. Next, to enable intermittent energy sources in the OTS problem, we propose a two-stage problem and utilize decision rule based approximation for its real-time implementation, which can significantly address the tractability and feasibility issues of the sample-based approaches. To further provide robustness in the OTS decision-making process, we have also studied the distributionally robust chance constrained (DRCC) formulation, which can ensure limited operation

constraint violations for any uncertainty within an ambiguity set. In Chapter 4, we mainly study the power system control problem under line outages caused by wildfires. We formulate the problem as a two-stage mixed-integer stochastic program and propose both preventive control and corrective control formulation to mitigate the impact of wildfires on the power grid. Both formulations are subject to efficient decomposition and relaxation techniques, thus can provide real-time mitigation solutions for system operators in face of fast-evolving wildfires. Lastly, in Chapter 5, we explore learning methods for the optimal load shedding (OLS) problem. Due to the communication latency under actual contingencies, we put forth a scalable decentralized learning framework for predicting the load shedding solution. Specifically, each load center only uses locally available measurements to predict local load shedding solutions. The scalable design of OLS decision rules enables load centers to quickly react to contingency situations without requiring the supervision from the control center.

Chapter 2

Monitoring of Substation Bus Split Events

This chapter presents the modeling and identification of bus split events. We will begin by performing the linear sensitivity analysis on the bus splits in Section 2.1. We incorporate bus splits into a security constrained economic dispatch problem in Section 2.2, which can help guarantee the secure system operations against this topology change. In Section 2.3, we focus on the identification problem of bus split events using high-resolution synchrophasor data.

2.1 Sensitivity Analysis of Bus Split Events

Security analysis in power system operations require accurate information of network topology and correct modeling of topology changes. One main function of substations is to report the status of circuit breakers (CBs) to the control center, which uses topology processing to obtain the real-time topol-

This chapter is based on the following publications to which the coauthors contributed equally:

Yuqi Zhou and Hao Zhu. "Bus split sensitivity analysis for enhanced security in power system operations." In Proc. North American Power Symposium, 2019.

Yuqi Zhou and Hao Zhu. "Efficient identification of bus split events using synchrophasor data." IEEE Transactions on Power Systems, 36(5):4800– 4808, 2021.

ogy [262, Ch. 12.7]. Due to increasing penetration of diverse energy resources, misoperations of grid protection systems occur more commonly, leading to unintentional CB actions [222]. Recently, malicious cyber intrusions have become another cause of unreliable CB status, with the vivid examples of the 2015/16 Ukrainian blackouts. Varying the status of CBs not only can result in the traditional contingency scenarios such as outages of lines or generation/loads, but also may give rise to electrical disconnection within a substation such as the *bus split* events. Therefore, it is imperative to consider the modeling of CB status changes and bus split contingencies for improving the security guarantees of power system operations.

The bus-branch representation is a concise and efficient modeling approach popularly used for power system optimization and planning. However, it lacks the detailed substation models, and thus cannot directly include CB switching actions and bus split. Hence, the node-breaker representation recently attracts more interest as an alternative; see e.g., [198, 205, 209]. Although it can provide the comprehensive substation-level representation, the node-breaker model typically involves sparse matrices of much larger dimension and results in non-convex power system optimization problems. Recent advances of automated topology processing for contingency analysis [110] has made it possible to quickly update the bus-branch model for given CB actions. Hence, it is possible now to efficiently model and analyze bus split events using the bus-branch model. As a contingency event, bus split is somewhat similar to line outage, for which sensitivity analysis has been the workhorse tool for ef-

ficient impact assessment and quick location identification [114, 299]. Inspired by this, we will aim to pursue the linear sensitivity analysis of bus split events.

This section aims to develop the sensitivity analysis framework for modeling bus split events using the bus-branch models. Towards this, we first introduce the dc linear power flow model and related matrix representations in Sec. 2.1.1. Linear sensitivity analysis of the post-split system is further considered in Sec. 2.1.2, by efficiently modifying the pre-split system to account for the additional bus.

2.1.1 System Modeling

We introduce the dc power flow model [239] for the modeling of bus split events. Consider a transmission system with $(N+1)$ buses collected in the set $\mathcal{N} := \{0, 1, \ldots, N\}$, and L transmission lines represented by the set $\mathcal{L} := \{(i,j)\}$. For each bus i, we use $\mathcal{N}_i$ to denote the set collecting its adjacent buses, and let θ_i denote its voltage angle, as well as g_i and d_i as its connected generation and load, respectively. Hence, the power injection per bus i is $p_i = g_i - d_i$. Without loss of generality (Wlog), we set bus 0 to be the reference angle bus with $\theta_0 = 0$. All non-reference phase angles are concatenated into vector $\boldsymbol{\theta} \in \mathbb{R}^N$; similarly for $\mathbf{p} \in \mathbb{R}^N$.

For each line (i,j), the power flow from bus i to bus j is denoted by f_{ij} and given by:

$$f_{ij} = \frac{1}{x_{ij}}(\theta_i - \theta_j) = b_{ij}(\theta_i - \theta_j), \forall (i,j) \in \mathcal{L} \tag{2.1}$$

where x_{ij} is the line reactance and its inverse equals to $b_{ij} = 1/x_{ij}$. Concatenating (2.1) into matrix form gives rise to the flow vector $\mathbf{f} \in \mathbb{R}^L$, as:

$$\mathbf{f} = \mathbf{K}\boldsymbol{\theta} \tag{2.2}$$

where matrix $\mathbf{K} \in \mathbb{R}^{L\times N}$ captures the network topology. The ℓ-th row of $\mathbf{K}$ corresponding to line (i, j) is given by $b_{ij}\left(\mathbf{e}_i - \mathbf{e}_j\right)^{\mathsf{T}}$, where $\mathbf{e}_i \in \mathbb{R}^N$ is the standard basis vector. Due to nodal flow conservation, one can sum up all the line flows to form the injected power $\mathbf{p}$, as given by:

$$\mathbf{p} = \mathbf{B}\boldsymbol{\theta} \tag{2.3}$$

where the *Bbus matrix* $\mathbf{B} \in \mathbb{R}^{N\times N}$ is invertible with each entry:

$$B_{ij} = \begin{cases} \sum_{k\in\mathcal{N}_i} b_{ik}, & \text{if } i = j \\ -b_{ij}, & \text{if } (i,j) \in \mathcal{L} \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

Therefore, matrix $\mathbf{B}$ can be also given by

$$\mathbf{B} = \sum_{(i,j)\in\mathcal{L}} b_{ij}(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^{\mathsf{T}}. \tag{2.5}$$

By solving for $\boldsymbol{\theta}$ in (3.3), one can write the line flow as $\mathbf{f} = (\mathbf{K}\mathbf{B}^{-1})\mathbf{p}$, with the coefficients in $(\mathbf{K}\mathbf{B}^{-1})$ termed as the injection shift factors (ISFs) that can transform from the injection $\mathbf{p}$ to line flow $\mathbf{f}$.

We are interested in the grid topology changes due to *bus split* within substations. Substations can be described as electrically connected nodes where multiple transmission lines terminate. For reliable and modular design, substations are equipped with switching equipment such as CBs and isolators

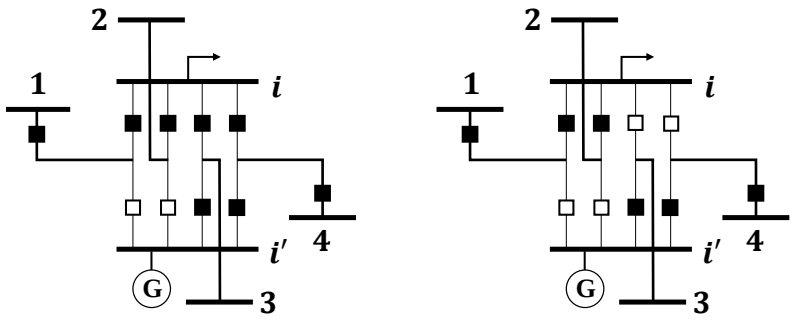


Figure 2.1: (Left) The original substation topology and (right) the new topology with two more breakers open.

to allow for flexible topology re-configurations during fault events [293]. Due to CB actions, the original bus may become electrically disconnected, commonly termed as "bus split." This type of topological changes is increasingly popular due to CB misoperations [140, 150] or malicious cyber intrusions [76, 244, 290]. Fig. 2.1 illustrates one such event, with solid squares representing closed breakers and hollow ones for open breakers. If the top right CBs open, bus i is split into two disconnected buses, i and i', in the same substation. Accordingly, the connectivity for transmission lines has changed, as well as for generation and load within this substation. Although the two buses are physically co-located in the same substation, they become electrically disconnected which affects the bus-branch model as shown in Fig. 2.2. Bus split events change the grid topology in a very different way from contingencies like line outages or generation/load disconnections. The ensuing section will first present the linearized sensitivity analysis, a powerful tool needed for various grid contingencies, for the bus split events.

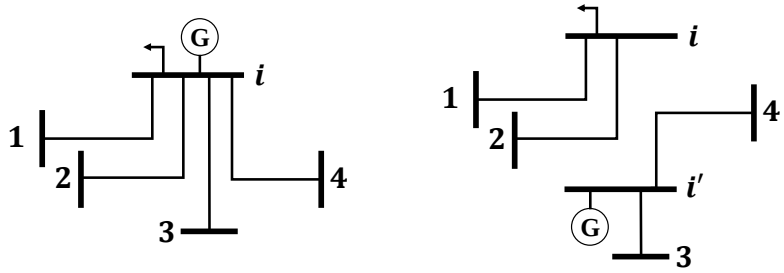


Figure 2.2: (Left) The original bus-branch model and (right) the updated model with reconnected lines/generation/loads due to bus split.

2.1.2 Linear Sensitivity Analysis

Consider the split of bus i leading to a new bus i', as illustrated by Fig. 2.3. After the split, let the subset $\mathcal{J} \subseteq \mathcal{N}_i$ collect the adjacent buses reconnected to the new bus i'. Moreover, denote $\tilde{\boldsymbol{\theta}} = \left[\boldsymbol{\theta}';\ \tilde{\theta}_{i'}\right] \in \mathbb{R}^{N+1}$ as the new angle vector after splitting; and similarly for $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{B}}$. The vector/matrix dimension increases by one due to the new bus i'. The goal is to update the new angle $\tilde{\boldsymbol{\theta}}$ and line flow $\tilde{\mathbf{f}}$. Note that bus splits may cause islanding in an interconnected system [240]. For simplicity, this paper assumes no island is formed from the split of bus i, as needed for typical analysis of line outages.

Notice that the $(N+1) \times (N+1)$ matrix $\tilde{\mathbf{B}}$ follows a block structure:

$$\tilde{\mathbf{B}} := \begin{bmatrix} \mathbf{B}' & \boldsymbol{\ell} \\ \boldsymbol{\ell}^\mathsf{T} & d \end{bmatrix} \tag{2.6}$$

where $\mathbf{B}'$ is the $N \times N$ sub-matrix, while vector $\boldsymbol{\ell}$ and scalar d capture the rest. Based on Fig. 2.3, the lines (i, j) with $j \in \mathcal{J}$ denote those lines that are reconnected to the new bus i' after the split. Similarly, we use buses $j \in \mathcal{J}'$ to match those lines (i, j) that remain connected to bus i after the split. Hence, the submatrix $\mathbf{B}'$ is formed by eliminating these reconnected lines from $\mathbf{B}$, as

given by [cf. (2.5)]

$$\mathbf{B}' = \mathbf{B} - \sum_{j \in \mathcal{J}} b_{ij} \left[(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T} - \mathbf{e}_j \mathbf{e}_j^\mathsf{T} \right]. \tag{2.7}$$

The reconnected lines $\{(i', j)\}$ also affect the rest of $\tilde{\mathbf{B}}$, as

$$\boldsymbol{\ell} = -\sum_{j \in \mathcal{J}} b_{ij} \mathbf{e}_j, \quad \text{and} \quad d = \sum_{j \in \mathcal{J}} b_{ij}. \tag{2.8}$$

With these definitions, one can find the inverse of $\tilde{\mathbf{B}}$ using the popular *matrix inverse lemma* [46, p. 650]. To this end, define the following *Schur complement* of the entry d to matrix $\tilde{\mathbf{B}}$ as

$$\mathbf{B}_d \coloneqq \mathbf{B}' - \left(d^{-1}\right) \boldsymbol{\ell} \boldsymbol{\ell}^\mathsf{T}. \tag{2.9}$$

Using $\mathbf{B}_d$, we can obtain the block structure of the inverse as

$$\tilde{\mathbf{B}}^{-1} = \begin{bmatrix} \mathbf{B}_d^{-1} & -\left(d^{-1}\right) \mathbf{B}_d^{-1} \boldsymbol{\ell} \\ -\left(d^{-1}\right) \boldsymbol{\ell}^\mathsf{T} \mathbf{B}_d^{-1} & (d^{-2}) \boldsymbol{\ell}^\mathsf{T} \mathbf{B}_d^{-1} \boldsymbol{\ell} + d^{-1} \end{bmatrix}. \tag{2.10}$$

Note that one can rewrite $\mathbf{B}_d = \mathbf{B} - d(\mathbf{u}\mathbf{u}^\mathsf{T})$ with the vector

$$\mathbf{u} \coloneqq \mathbf{e}_i + \left(d^{-1}\right) \boldsymbol{\ell}. \tag{2.11}$$

For example, if $|\mathcal{J}| = 1$, it becomes the case of single line reconnection as discussed in [295] where vector $\mathbf{u} = \mathbf{e}_i - \mathbf{e}_j$ with $\mathcal{J} = \{j\}$. Applying the Sherman-Morrison formula [232] leads to the inverse

$$\mathbf{B}_d^{-1} = (\mathbf{I} + \boldsymbol{\Delta}_\mathbf{u}) \mathbf{B}^{-1} = \mathbf{B}^{-1} + \frac{d \mathbf{B}^{-1} \mathbf{u}\mathbf{u}^\mathsf{T}}{1 - d \mathbf{u}^\mathsf{T} \mathbf{B}^{-1} \mathbf{u}} \mathbf{B}^{-1}, \tag{2.12}$$

with $\boldsymbol{\Delta}_\mathbf{u}$ capturing the fractional term above. Clearly, the inverse $\mathbf{B}_d^{-1}$ can be quickly formed with $\mathbf{B}^{-1}$ available, and so are the other blocks in (2.10).

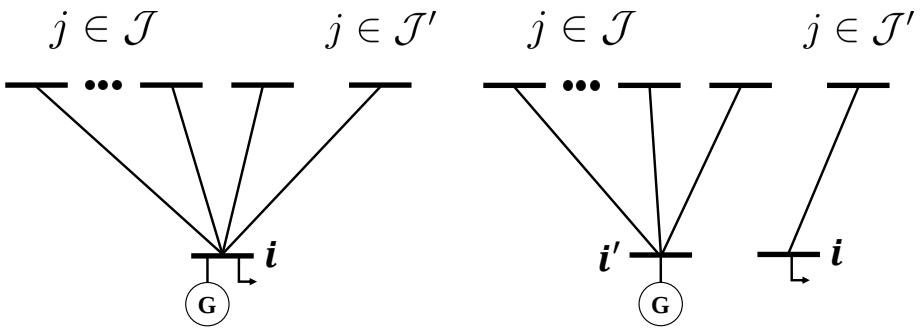


Figure 2.3: (Left) The original bus-branch model and (right) the updated model due to bus split.

The next step is to form the post-split power injection vector $\tilde{\mathbf{p}}$. Fig. 2.3 indicates that generation and load originally connected to bus i could become attached to the new bus i' due to the split. Let $\tilde{p}_i$ denote the total injected power to bus i' after the split, and thus the injected power to bus i becomes $(p_i - \tilde{p}_i)$. For example, Fig. 2.3 shows the case of $\tilde{p}_i = g_i$ and $(p_i - \tilde{p}_i) = -d_i$. In general, $\tilde{p}_i$ can also be the attached load, or, the combination of generation and load. Therefore, we can express the injected power as

$$\tilde{\mathbf{p}} = \begin{bmatrix} \mathbf{p} \\ 0 \end{bmatrix} + \begin{bmatrix} -\tilde{p}_i \mathbf{e}_i \\ \tilde{p}_i \end{bmatrix} = \begin{bmatrix} \mathbf{p} - \tilde{p}_i \mathbf{e}_i \\ \tilde{p}_i \end{bmatrix}. \tag{2.13}$$

Using (2.10) and (2.13), one can obtain the post-split $\tilde{\boldsymbol{\theta}}$ as

$$\begin{bmatrix} \boldsymbol{\theta}' \\ \tilde{\theta}_{i'} \end{bmatrix} = \tilde{\mathbf{B}}^{-1} \tilde{\mathbf{p}} = \begin{bmatrix} \boldsymbol{\theta}' \\ d^{-1} \left(-\boldsymbol{\ell}^\top \boldsymbol{\theta}' + \tilde{p}_i \right) \end{bmatrix} \tag{2.14}$$

where the angle vector for the original N buses is

$$\boldsymbol{\theta}' \coloneqq \mathbf{B}_d^{-1} \left(\mathbf{p} - \tilde{p}_i \mathbf{e}_i - d^{-1} \tilde{p}_i \boldsymbol{\ell} \right) \tag{2.15}$$

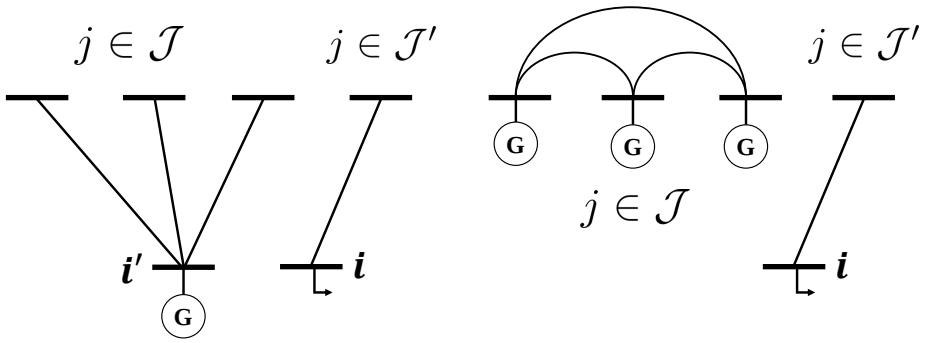


Figure 2.4: (Right) The equivalent bus-branch model after eliminating bus i' for (left) the system due to the bus split.

Although the solution $\tilde{\boldsymbol{\theta}}$ can be readily computed once obtaining (2.12), it is possible to better interpret it by developing an equivalent model for the post-split system. As illustrated in Fig. 2.4, the new bus i' can be eliminated by connecting every pair of buses in $\mathcal{J}$. For example, the new line connecting any buses $j_1, j_2 \in \mathcal{J}$ would have the equivalent line parameter $b_{j_1 j_2} := (d^{-1})\, b_{ij_1} b_{ij_2}$ by recalling d from (2.8). This follows from the well-known Kron's reduction approach [81]. Similarly, after eliminating bus i', its injected power $\tilde{p}_i$ is reallocated to each of its neighboring buses, with $d^{-1} b_{ij} \tilde{p}_i$ to bus $j \in \mathcal{J}$.

Interestingly, the equivalent topology in Fig. 2.4 exactly matches the matrix $\mathbf{B}_d$ as defined in (2.9). Substituting (2.7) into (2.9), one can decompose $\mathbf{B}_d = \mathbf{B}^{\mathcal{J}} + \mathbf{B}^a$, where the matrix

$$\mathbf{B}^{\mathcal{J}} := \mathbf{B} - \sum_{j \in \mathcal{J}} b_{ij} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T} \tag{2.16}$$

corresponds to the Bbus matrix with the outage of lines in $\{(i,j)\}_{j \in \mathcal{J}}$ from the

original system, while the remaining part

$$\mathbf{B}^a = \sum_{j\in\mathcal{J}} b_{ij}\mathbf{e}_j\mathbf{e}_j^\mathsf{T} - \left(d^{-1}\right)\boldsymbol{\ell}\boldsymbol{\ell}^\mathsf{T} \tag{2.17}$$

reflects the new equivalent lines among all buses in $\mathcal{J}$. Hence, the angle vector $\boldsymbol{\theta}'$ in (2.15) perfectly matches the equivalent system for the original N buses with updated topology/injection, as its total injection now includes the new ones at each $j \in \mathcal{J}$ due to eliminating bus i'. Using the updated angles in $\boldsymbol{\theta}'$, one can further recover the angle at bus i' by solving the total power flow on the eliminated lines in $\{(i', j),\ \forall j \in \mathcal{J}\}$ as

$$\tilde{p}_i = \sum_{j\in\mathcal{J}} b_{ij}\left(\tilde{\theta}_{i'} - \mathbf{e}_j^\mathsf{T}\boldsymbol{\theta}'\right), \tag{2.18}$$

leading to the same solution of $\tilde{\theta}_{i'}$ as in (2.14). This equivalencing analysis clearly explains how the post-split system can be related to the original one.

Upon solving for the post-split angle $\tilde{\boldsymbol{\theta}}$, one can formalize the sensitivity analysis in terms of the changes of bus angles and line flows, as given in the following proposition.

Proposition 2.1.1. *For the split of bus i with a re-connected injection of $\tilde{p}_i$, the phase angle difference at the original N buses can be written as*

$$\boldsymbol{\theta}' - \boldsymbol{\theta} = \boldsymbol{\Delta}_\mathbf{u}\boldsymbol{\theta} - (\mathbf{I} + \boldsymbol{\Delta}_\mathbf{u})\mathbf{B}^{-1}\left(\tilde{p}_i\mathbf{e}_i + d^{-1}\tilde{p}_i\boldsymbol{\ell}\right) \tag{2.19}$$

where d, $\boldsymbol{\ell}$, and $\boldsymbol{\Delta}_\mathbf{u}$ are given in (2.8) *and* (2.12)*. Accordingly, the line flow difference due to the bus split is [cf.* (2.2)*]*

$$\tilde{\mathbf{f}} - \mathbf{f} = \mathbf{K}\boldsymbol{\Delta}_\mathbf{u}\boldsymbol{\theta} - \mathbf{K}(\mathbf{I} + \boldsymbol{\Delta}_\mathbf{u})\mathbf{B}^{-1}\left(\tilde{p}_i\mathbf{e}_i + d^{-1}\tilde{p}_i\boldsymbol{\ell}\right). \tag{2.20}$$

Strictly speaking, (2.20) holds for the lines other than the reconnected ones $\{(i', j)\}_{\forall j \in \mathcal{J}}$, as the flows of the latter lines are related to $\tilde{\theta}_{i'}$. In addition, since matrix $\mathbf{\Delta_u}$ is a linear transformation of $\mathbf{B}^{-1}$ [cf. (2.12)], the matrix product $\mathbf{K\Delta_u}$ in (2.20) can be quickly computed using the ISF matrix $\mathbf{KB}^{-1}$.

Given the detailed post-split topology and injection, the sensitivity analysis in Proposition 4.3.1 enables explicit expression of the phase angle and line flow change. In addition to generalizing line outage sensitivity analysis [113], this result could benefit security-constrained economic dispatch by including bus split contingencies, which is the subject of the ensuing section.

2.2 Power System Operations Under Bus Splits

Limited by the bus-branch modeling approach, conventional power system analysis and operations rarely consider the status of CBs within substations for system monitoring and optimization. However, as the grid operations move closer to stability and capacity limits, it is increasingly crucial to incorporate the breaker models for enhanced security of grid operations. Here, we take the economic dispatch (ED) problem as an example to demonstrate the importance. ED aims to determine the most cost-effective generation output in real time, in order to serve total system load [262, Ch. 3]. There are several versions of this fundamental problem. One extension is to include the transmission line limits, termed as network-constrained (NC)ED. Furthermore, the guarantees of secure operations under potential contingencies such as line outages can be included as additional constraints, and therefore the

security-constrained (SC)ED. Using the linear sensitivity factors developed in Sec. 2.1.2, one can generalize the SCED to include the new type of bus-split contingencies.

2.2.1 Network Constrained Economic Dispatch

Let $\boldsymbol{g}, \boldsymbol{d} \in \mathbb{R}^{N+1}$ collect the generation (decision variables) and load (given inputs) at all buses, including the reference bus. With the corresponding generation cost given by vector $\mathbf{c} \in \mathbb{R}^{N+1}$, the NCED problem can be cast as the following:

$$\min_{\boldsymbol{g} \in \mathbb{R}^{N+1}} \quad \boldsymbol{c}^\mathsf{T} \boldsymbol{g} \tag{2.21a}$$

$$\text{s.t.} \quad \mathbf{1}^\mathsf{T} \boldsymbol{g} - \mathbf{1}^\mathsf{T} \boldsymbol{d} = 0 \tag{2.21b}$$

$$\underline{\boldsymbol{g}} \leq \boldsymbol{g} \leq \bar{\boldsymbol{g}} \tag{2.21c}$$

$$\underline{\mathbf{f}} \leq \mathbf{S}(\boldsymbol{g} - \boldsymbol{d}) \leq \bar{\mathbf{f}} \tag{2.21d}$$

where $[\underline{\boldsymbol{g}}, \bar{\boldsymbol{g}}]$ stands for the generation limits; and similarly for the line limits in $[\underline{\mathbf{f}}, \bar{\mathbf{f}}]$. The augmented ISF matrix $\mathbf{S} := \begin{bmatrix} \mathbf{0} & \mathbf{K}\mathbf{B}^{-1} \end{bmatrix}$ uses $\mathbf{0}$ to account for the reference bus. Note that the NCED problem given in (2.21) is a basic formulation and can be generalized to include quadratic (instead of linear) generation cost and variable (instead of constant) load demand.

2.2.2 Security Constrained Economic Dispatch

SCED extends NCED to account for the system security under contingency events. For simplicity, consider line outages only, and denote $\mathbf{S}^k :=$

$\begin{bmatrix}\mathbf{0} & \mathbf{K}^k(\mathbf{B}^k)^{-1}\end{bmatrix}$ as the updated ISF matrix under the outage of line $k \in \mathcal{L}$. Note that matrix $\mathbf{K}^k$ sets the k-th row of $\mathbf{K}$ to be all zeros, so the corresponding ISFs are zero for line k. Hence, the SCED problem includes an additional set of constraints for contingencies as:

$$\min_{\boldsymbol{g}\in\mathbb{R}^{N+1}} \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{g} \tag{2.22a}$$

$$\text{s.t.} \quad (2.21\text{b}) - (2.21\text{d}) \tag{2.22b}$$

$$\underline{\mathbf{f}} \leq \mathbf{S}^k(\boldsymbol{g} - \boldsymbol{d}) \leq \bar{\mathbf{f}},\ \forall k \in \mathcal{L} \tag{2.22c}$$

2.2.3 Breaker Security Constrained Economic Dispatch

Based on the SCED problem, we can incorporate bus-split contingencies as well. As mentioned in Sec. 2.1.2, the split of bus i is equivalent to the outage of line (i, j) with additional power transfer between bus i and j, indicated by Fig. 2.4. Hence, consider the subset $\mathcal{B} \in \mathcal{L}$ include all lines that can be affected by a certain bus split event. Therefore, the generalized breaker (B-)SCED formulation becomes:

$$\min_{\boldsymbol{g}\in\mathbb{R}^{N+1}} \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{g} \tag{2.23a}$$

$$\text{s.t.} \quad (2.21\text{b}) - (2.21\text{d}), (2.22\text{c}) \tag{2.23b}$$

$$\underline{\mathbf{f}} \leq \mathbf{S}^k(\boldsymbol{g} - \boldsymbol{d} - \tilde{p}_i\mathbf{e}_i + \tilde{p}_i\mathbf{e}_j) \leq \bar{\mathbf{f}},\ \forall k \in \mathcal{B} \tag{2.23c}$$

$$\underline{f_{ij}} \leq \tilde{p}_i \leq \overline{f_{ij}},\ \forall (i, j) \in \mathcal{B} \tag{2.23d}$$

where $\tilde{p}_i$ is an auxiliary variable that can be g_i, d_i, or the combination of the two, depending on the generation/load split. Note that the last constraint

(2.23d) corresponds to the limit on the new line (i', j), which equals to the injected power $\tilde{p}_i$ to bus i' due to power conservation. Given the CB model of all substations, all bus split events in $\mathcal{B}$ can be predetermined, which is usually a smaller subset of $\mathcal{L}$.

Clearly, the proposed B-SCED formulation extends the SCED one with additional security-related constraints on line flows. Since bus split is a common substation activity, adding related constraints can improve the security of power system operation and control.

2.2.4 Numerical Results

An illustrative 3-bus system and the modified IEEE 118-bus test case are used to study the security enhancement provided by B-SCED. The optimization problems have been implemented on a regular laptop in the MATLAB® R2018a simulator modeled by MATPOWER [303] and further solved by CPLEX solver (version 12.9).

2.2.4.1 Illustrative 3-Bus System

To demonstrate the impact of B-SCED, consider a simple 3-bus system with linear generation cost and other system parameters shown in Fig. 2.5. In this case, we set the line flow limit $\underline{\mathbf{f}} = -\bar{\mathbf{f}}$. One bus-split contingency at bus 3 is included for the study, which would result in a new bus 4 with injected power $\tilde{p}_3 = g_3$. Hence, the set of split lines is $\mathcal{B} = \{(3, 1)\}$. We have implemented the three ED problems for this system, all of which are linear

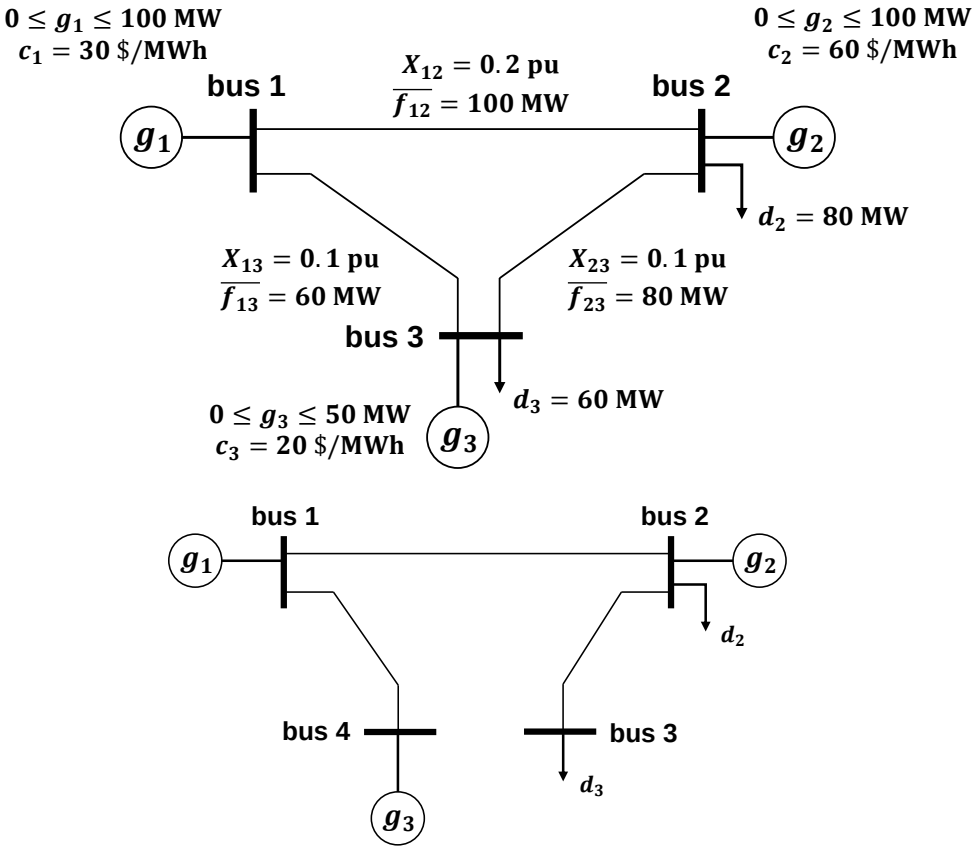


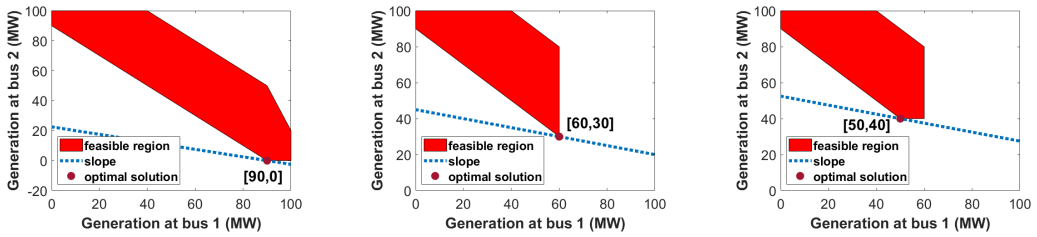
Figure 2.5: (Top) A toy 3-bus system with (bottom) its post-split topology.

Table 2.1: Optimal Dispatch and Total Generation Cost

	Optimal Dispatch $\boldsymbol{g}^*$ (MW)	Total Cost C^* (\$/h)
NCED	[90;0;50]	3700
SCED	[60;30;50]	4600
B-SCED	[50;40;50]	4900

programs (LPs) under linear generation costs. Table 2.1 lists the results of both the optimal generation $\boldsymbol{g}^*$ and the corresponding total cost, suggesting the trade-off for improved system security at the price of increased cost.

To better explain this trade-off, we have also plotted the feasible regions induced by each ED problem in Fig. 2.6. Since each ED problem is an LP, the feasible region is a polytope in the three-dimensional space. To simplify the visualization, we project it to the two-dimensional space by replacing g_3 with $(\mathbf{1}^T\boldsymbol{d} - g_1 - g_2)$ [cf. (2.21b)]. Thus, Fig. 2.6 provides the feasible region for the pair (g_1, g_2) only, along with the slope of the objective function on these two variables. From the feasible region of SCED to that of B-SCED, the additional bus-split contingency leads to further restriction of feasibility. This makes the optimal dispatch $\boldsymbol{g}^* = [60; 30; 50]$ MW for the SCED problem no longer feasible for B-SCED. Specifically, its corresponding line flows will be $[f_{12}; f_{23}; f_{41}] = [110; 60; 50]$ MW where $f_{12} = 110$ MW violates the line flow limit of $\overline{f_{12}} = 100$ MW. If this line flow limit is increased from 100 MW to 110 MW, then g_1 is able to provide another 10 MW to serve the load. This way, the optimal dispatch for B-SCED will coincide with the SCED one $\boldsymbol{g}^* = [60; 30; 50]$ MW. Otherwise, one has to increase the output of the most costly generator g_2 among the three, from 30 MW to 40 MW.



(a) (b) (c)

Figure 2.6: The feasible region and slope of objective function over (g_1, g_2) for (a) NCED; (b) SCED; and (c) B-SCED.

The 3-bus example shows that the bus-split contingencies can critically affect the security level provided by SCED solutions. This effect will likely increase for large-scale systems under more bus-split possibilities. Therefore, B-SCED could markedly improve the security of grid operations by accounting for the breaker actions, which will be studied next using a more practical test case.

2.2.4.2 IEEE 118-Bus System

In the 118-bus system, there are a total of 54 committed generators, all with quadratic cost function in the form of $c_i(g_i) = \alpha_i g_i^2 + \beta_i g_i$. We set the limits of transmission lines based on angular difference bounds, as given by $(\Delta\theta_{\max}/x_{ij})$. Infeasibility is a known issue of SCED for large systems; see e.g., [132,275]. This is mainly due to more significant change of power flow patterns under line outages. For example, for buses with very low vertex degree (number of incident lines), the outage of one incident line will induce transmission congestion in the local area, and thus putting higher stress on other connected lines. Such infeasibility issue can be addressed by either reducing the area load demand and other preventive or corrective control strategies as detailed in [275]. Similar issue may also occur in B-SCED of IEEE 118-bus case, which will be handled by load shedding at certain buses (bus 13, 20, 43, 45, and 78) to correct the feasible region under security constraints.

During the simulations we consider bus split at 4 substations (substation 6, 24, 55, and 100), all connected to both generator and load. All

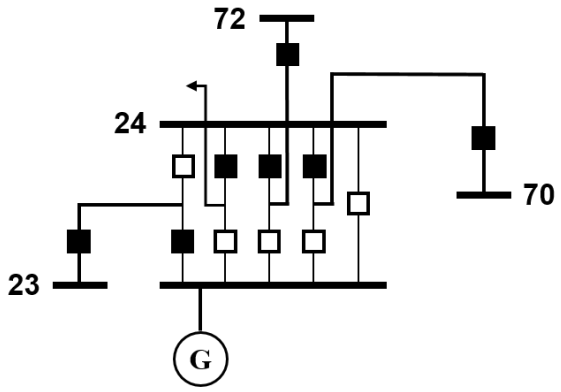


Figure 2.7: Bus split at substation 24.

bus-split events therein result in a single line switch in the bus-branch model, corresponding to lines (5,6), (23,24), (55,59) and (100,101). For example, bus split at substation 24 is represented by circuit breaker model shown in Fig. 2.7, where the original line (23,24) becomes connected to the new bus attached by the generator after the split.

Fig. 2.8 plots the optimal output of every generator in the 118-bus system under all three ED problems, along with the respective total generation cost. The total cost increases by around 8.4% (1.7%) from the base NCED (SCED) to B-SCED. However, the latter has provided higher security values in the event of either line-outage or bus-split contingencies.

For a given bus-split scenario, one can intuitively investigate its impact by evaluating the optimal solution of SCED using the bus-split constraints (2.23c)−(2.23d). If the solution by SCED satisfies both constraint, then the B-SCED problem will output the same dispatch solution as well. Otherwise, the B-SCED will use a higher-cost dispatch in order to meet these constraints.

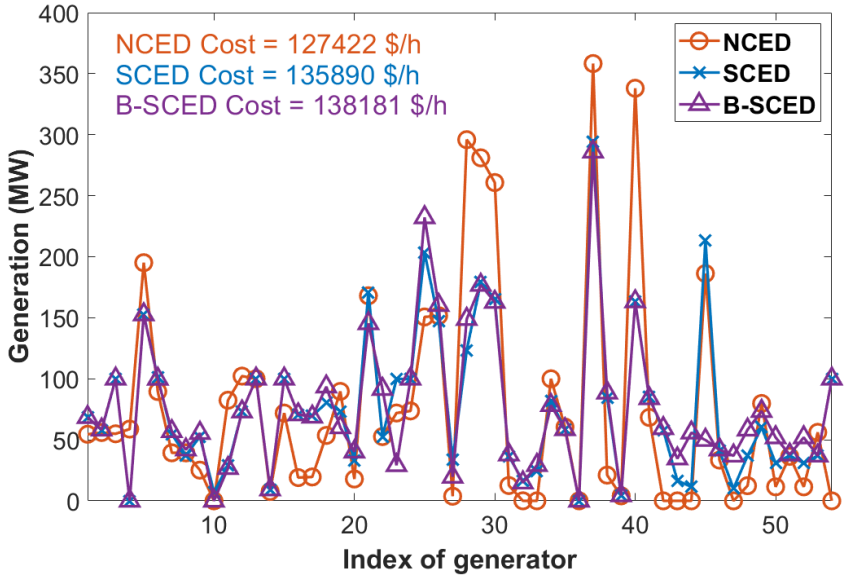


Figure 2.8: Comparison of the generator dispatch outputs under all three ED formulations.

For example, the SCED solution under the bus split at substation 100 violates the limits of five lines (92,94), (92,100), (101,102), (92,102) and (100,101). The first 4 lines correspond to the constraints in (2.23c) while the last one relates to the new line due to the split in (2.23d). Hence, for guaranteed security under the contingency of bus split at substation 100, it is necessary to increase the output of generators of higher costs. Compared to SCED, B-SCED under bus split at substation 100 has led to increased generation at buses 92, 99, 104, 105, 107, 110, and 112, all with incremental cost at 40 \$/MWh−42 \$/MWh, while decreasing the generation output of bus 100 with an incremental cost of 36.94 \$/MWh.

The economic impact of various bus split scenarios is investigated with

Table 2.2: Total Costs Under Various Bus Split Events

	Bus-split Substations	Total Cost C^* (\$/h)
SCED	-	135890
B-SCED	6, 24	135890
B-SCED	6, 24, 55	136420
B-SCED	6, 24, 55, 100	138181

the total costs listed in TABLE 2.2. Compared with SCED, bus splits only at substations 6 and 24 do not affect the dispatch solution at all. Including bus split at substation 55, and later substation 100, the total cost gradually increases. Our empirical observation suggests that the key factors on the cost increase are the post-split topology and the magnitude of the resultant $\tilde{p}_i$. Typically, the generator with large output under SCED dispatch is more likely to be affected by the bus split in terms of the feasible region. Thus, such economic studies can help to develop a quantitative measure for each bus-split event in terms of incremental cost, as the price paid for securing the grid against a contingency.

2.3 Identification of Bus Split Events Using Synchrophasor Data

Accurate information of grid topology is crucial for performing various power system operation and maintenance tasks. Recent reports [4,6,138] point out the increasing importance of developing efficient grid topology modeling and monitoring techniques that can include bus split events. Traditionally,

the grid topology is processed and updated by the energy management system (EMS) using the Supervisory Control and Data Acquisition (SCADA) inputs [9, Ch. 1]. A compact bus-branch model is typically obtained by examining the statuses of CBs and switching devices. *Topology errors* due to incorrect reports of CB statuses can be identified by the *generalized* state estimator using SCADA data; see e.g., [9, Ch. 8] and references therein. Recent deployment of phasor measurement units (PMUs) provides the synchrophasor data of high sampling rates and resolution, which can greatly facilitate the process of topology identification. Fast algorithms for identifying transmission line outages has been developed in e.g., [59, 242, 299], and also for distribution grids [17, 54, 75]. Unfortunately, the identification of CB status and bus split events has been rarely considered using synchrophasor data. A detailed node-breaker model is typically required to capture the CB statuses and substation connectivity, but this representation is known to increase the problem dimensionality and computational time [198]. Detecting CB actions from the synchrophasor data has been considered in [140], where lack of identifiability has been observed therein due to the complicated node-breaker model. Thus, the problem of efficiently analyzing and monitoring bus split events using the latest synchrophasor data still remains open.

The goal of this section is to develop an efficient monitoring framework for bus split events based on the concise bus-branch representation [296]. Towards this end, it is possible identify the bus split event by searching for the post-split scenario that best matches the actual synchrophasor measurements.

By enumerating all possible bus split scenarios, one can perform the sensitivity analysis for each scenario to compute the corresponding post-split phase angles. Nonetheless, the number of scenarios in this exhaustive search approach would grow exponentially with the number of substation connections, limiting it from real-time implementation. To address this complexity issue, we leverage the popular *McCormick relaxation* technique [177] to replace the *nonconvex* bilinear terms with equivalent *convex* linear inequalities. This reformulation leads to a tractable mixed-integer linear programming problem that is efficiently solvable for real-time identification. By systematically incorporating bus split events, this work directly addresses the lack of consideration for such events in the traditional grid analysis and monitoring framework.

2.3.1 Nonlinear Optimization Formulation

Increasing deployment of high-resolution sensors such as PMUs allows for real-time identification of anomalous events such as topology changes. It is thus possible to utilize synchrophasor data to efficiently identify the underlying CB statuses during bus split events. Based on the sensitivity analysis in Sec. 2.1.2, an intuitive solution could be the exhaustive search by enumerating all number of possible scenarios. However, the complexity order of this approach is not ideal for real-time identification and motivates us to consider more efficient alternatives.

As in the identification of line outages [242, 299], synchrophasor data provides the difference between pre- and post-event phase angle measure-

ments which can be used to locate bus splits. Consider again the split of bus i as in the sensitivity analysis. This candidate bus can be selected based on the bus locations where significant angle changes have been observed, as detailed later. Given bus i and its subset $\mathcal{J}$ that include all reconnected buses/generation/loads, the post-split phase angle $\tilde{\boldsymbol{\theta}}$ can be computed in (2.14). This suggests a brute-force solution by enumerating all possible scenarios of sets $\mathcal{J}$ and selecting the one best matching the measured changes. Towards this end, for the target bus i let us define the binary variables $\{z^j\}_{j\in\mathcal{N}_i}$ to indicate the post-event status for each line (i,j), such that

$$z^j = \begin{cases} 1, & \text{if bus } j \text{ re-connects to bus } i' \\ 0, & \text{if bus } j \text{ stays connected to bus } i \end{cases} \tag{2.24}$$

Similarly, let the binary variables $\{z^g,\ z^d\}$ denote the post-event connection status for the generation and load located at bus i. For simplicity, only one generator and one load are considered here, which can be extended to the case of multiple number of them. Given all these binary variables in vector $\boldsymbol{z} \in \{0,1\}^{(|\mathcal{N}_i|+2)}$, one can use the sensitivity analysis to compute the post-split angle $\tilde{\boldsymbol{\theta}}(\boldsymbol{z})$. Denoting the measured post-split angle as vector $\mathbf{c}$, the identification problem becomes to minimize the norm of mismatch error between the two, namely $\|\tilde{\boldsymbol{\theta}}(\boldsymbol{z}) - \mathbf{c}\|$. To address the potential approximation error of dc power flow model, a more popular error objective is to compare the *phase angle change* instead of the post-split phase angle itself; see e.g., [299]. Hence, the bus split identification problem becomes

$$\min_{\boldsymbol{z}\in\{0,1\}^{(|\mathcal{N}_i|+2)}} \quad \left\| \left(\tilde{\boldsymbol{\theta}}(\boldsymbol{z}) - \hat{\boldsymbol{\theta}} \right) - \boldsymbol{\delta} \right\| \tag{2.25}$$

where $\hat{\boldsymbol{\theta}} := [\boldsymbol{\theta}; \theta_i]$ is pre-split phase angle solution by the dc model and $\boldsymbol{\delta}$ is the observed phase angle difference by the PMUs. Common choice of norms such as L_1, L_2 or L_∞ can be used for quantifying the mismatch error in the objective (2.25). In this paper, we select the L_1 norm error which is known to enjoy a nice reformulation to linear objective function. Note that if only partial angles are monitored or the line flow data is also available, one can modify the objective function of (2.25) by extracting corresponding entries or adding a linear transformation of the angle difference vector. Such generalization will be investigated numerically in Sec. 2.3.5.2. Using (2.25), the naive exhaustive search approach works by enumerating all $2^{(|\mathcal{N}_i|+2)}$ scenarios of $\boldsymbol{z}$, and picking the one with the minimum error in (2.25). Such exponential complexity is not suitable for real-time implementation, and will be addressed here by developing a tractable solution.

The main challenge in directly optimizing for $\boldsymbol{z}$ in (2.25) lies in the post-split power flow model, as

$$\tilde{\boldsymbol{\theta}}(\boldsymbol{z}) = \left[\tilde{\mathbf{B}}(\boldsymbol{z})\right]^{-1} \tilde{\mathbf{p}}(\boldsymbol{z}). \tag{2.26}$$

The post-split Bbus matrix depends on the line status variables $\{z^j\}$, given by

$$\tilde{\mathbf{B}}(\boldsymbol{z}) = \bar{\mathbf{B}} + \sum_{j \in \mathcal{N}_i} z^j \ \mathbf{D}^j \tag{2.27}$$

where $\bar{\mathbf{B}} = \left[\mathbf{B}\ \mathbf{0}; \mathbf{0}^\mathsf{T}\right]$ corresponds to the original Bbus matrix for the augmented system including the new bus i', while the change due to each recon-

nected line is captured by

$$\mathbf{D}^j := b_{ij}\left[(\mathbf{e}_{i'}-\mathbf{e}_j)(\mathbf{e}_{i'}-\mathbf{e}_j)^\mathsf{T}-(\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^\mathsf{T}\right]. \tag{2.28}$$

Similarly, the post-split injection vector depends on the generation/load status variables, as

$$\tilde{\mathbf{p}}(\boldsymbol{z}) = \bar{\mathbf{p}} + z^g\boldsymbol{\delta}^g - z^d\boldsymbol{\delta}^d \tag{2.29}$$

where $\bar{\mathbf{p}} = [\mathbf{p}; 0]$ also augments the dimension of $\mathbf{p}$ to include the new bus i', while the other two vectors are given by

$$\boldsymbol{\delta}^g := g_i(\mathbf{e}_{i'}-\mathbf{e}_i) \text{ and } \boldsymbol{\delta}^d := d_i(\mathbf{e}_{i'}-\mathbf{e}_i). \tag{2.30}$$

Substituting (2.26) into the objective of (2.25) introduces the term $[\tilde{\mathbf{B}}(\boldsymbol{z})]^{-1}$, which is the inverse of a matrix function of binary variables. Furthermore, its multiplication with $\tilde{\mathbf{p}}(\boldsymbol{z})$ poses additional nonlinearity (bilinearity) to the problem. Both issues lead to the lack of tractability in solving (2.25).

2.3.2 Tractable Reformulation via McComick Relaxation

To tackle the bilinearity issues, we propose to adopt the *McCormick relaxation* technique [177], which is a powerful tool for dealing with bilinear terms of binary variables. Specifically, by introducing an additional matrix $\mathbf{X} \in \mathbb{R}^{(N+1)\times(N+1)}$ to represent $[\tilde{\mathbf{B}}(\boldsymbol{z})]^{-1}$, we can reformulate (2.25) as:

$$\min_{\boldsymbol{z},\mathbf{X}} \quad \left\|\mathbf{X}\tilde{\mathbf{p}}(\boldsymbol{z}) - \hat{\boldsymbol{\theta}} - \boldsymbol{\delta}\right\| \tag{2.31a}$$

$$\text{s.t.} \quad \tilde{\mathbf{B}}(\boldsymbol{z})\mathbf{X} = \mathbf{I}. \tag{2.31b}$$

This way, the matrix inversion is no longer needed, which is replaced by the bilinear products between the unknowns. Note that the bilinear constraint (2.31b) for enforcing the relation between $\mathbf{X}$ and $\tilde{\mathbf{B}}(\boldsymbol{z})$ now becomes [cf. 2.27]:

$$\bar{\mathbf{B}}\mathbf{X} + \sum_{j \in \mathcal{N}_i} \mathbf{D}^j \mathbf{Y}^j = \mathbf{I} \tag{2.32}$$

with the product $\mathbf{Y}^j := z^j \mathbf{X}$ defined for each line (i, j). Given the binary $z_j \in \{0, 1\}$ and the (m, n)-th entry $X_{mn} \in [X_{mn}^{\min}, X_{mn}^{\max}]$, each entry of $\mathbf{Y}^j$ can be written as:

$$Y_{mn}^j = z^j X_{mn}, \ \forall \ m, n. \tag{2.33}$$

Accordingly, the following four *linear* inequalities hold:

$$Y_{mn}^j \geq z^j X_{mn}^{\min}, \tag{2.34a}$$

$$Y_{mn}^j \geq X_{mn} + z^j X_{mn}^{\max} - X_{mn}^{\max}, \tag{2.34b}$$

$$Y_{mn}^j \leq z^j X_{mn}^{\max}, \tag{2.34c}$$

$$Y_{mn}^j \leq X_{mn} + z^j X_{mn}^{\min} - X_{mn}^{\min}. \tag{2.34d}$$

Each inequality in (2.34) can be verified by substituting (3.12). Interestingly, the set of four linear inequalities in (2.34) jointly also guarantees the validity of (3.12) for any binary z^j. To demonstrate this, first consider the case of $z^j = 0$. The two inequalities (2.34a) and (2.34c) jointly lead to $Y_{mn}^j = 0$, and thus (3.12) holds. Otherwise if $z^j = 1$, the other two inequalities (2.34b) and (2.34d) would enforce that $Y_{mn}^j = X_{mn}$. Hence, for binary z^j the set of inequalities in (2.34) is *equivalent to* the bilinear relation in (3.12). Reformulating

(3.12) using the linear inequalities in (2.34) is known as the *Mccormick relaxation* technique and has been usually used in power system topology designing problems [40, 196]. Thus, the original constraint (2.31b) can be replaced with (2.32) and (2.34a) - (2.34d).

For the objective function in (2.31a), we can introduce matrices $\mathbf{W}^p := z^p\mathbf{X}$ with $p \in \{g, d\}$ indicating either the generation or load, and similarly convert the matrix products into equivalent linear inequalities. We replace the resultant bilinear terms in (2.31a) for $p \in \{g, d\}$ using the follows:

$$W^p_{mn} \geq z^p X^{\min}_{mn}, \tag{2.35a}$$

$$W^p_{mn} \geq X_{mn} + z^p X^{\max}_{mn} - X^{\max}_{mn}, \tag{2.35b}$$

$$W^p_{mn} \leq z^p X^{\max}_{mn}, \tag{2.35c}$$

$$W^p_{mn} \leq X_{mn} + z^p X^{\min}_{mn} - X^{\min}_{mn}. \tag{2.35d}$$

Proposition 2.3.1. *Given the bounds* $[X^{\min}_{mn}, X^{\max}_{mn}]$ *for each entry* X_{mn}*, we can reformulate* (2.31) *as the following mixed-integer programming problem after substituting* $\tilde{\mathbf{p}}(\boldsymbol{z})$ *in* (2.29)*:*

$$\min_{\boldsymbol{z},\mathbf{X},\{\mathbf{Y}^j\},\{\mathbf{W}^p\}} \quad \|\mathbf{X}\bar{\mathbf{p}} + \mathbf{W}^g\boldsymbol{\delta}^g - \mathbf{W}^d\boldsymbol{\delta}^d - \hat{\boldsymbol{\theta}} - \boldsymbol{\delta}\| \tag{2.36a}$$

$$s.t. \quad (2.32),\ (2.34) \text{ and } (2.35),\ \forall\ (m, n). \tag{2.36b}$$

As mentioned earlier, the L_1 error norm in (2.36a) can lead to an equivalent linear objective cost; see e.g., [46, Ch. 4]. Meanwhile, all constraints in (2.36b) are linear, thanks to the Mccormick relaxation technique. Hence, the original problem (2.25) is converted to a mixed-integer linear program (MILP),

for which there exist several off-the-shelf efficient solvers such as CPLEX and Gurobi. This equivalent MILP reformulation (2.36) constitutes as a tractable solution for identifying the status of multiple connections within a specific substation. The setting of tight bounds $[X_{mn}^{\min}, X_{mn}^{\max}]$ for each system will be discussed soon.

2.3.3 Simplification with Matrix Sparsity

It is possible to further reduce the number of decision variables in (2.36) by taking advantage of the sparse structure of matrix $\mathbf{D}^j$. To this end, let us define the product $\mathbf{\Delta}^j := \mathbf{D}^j\mathbf{Y}^j = z^j\mathbf{D}^j\mathbf{X}$, which was used in (2.32). With the definition of $\mathbf{D}^j$ in (2.28), the product $\mathbf{\Delta}^j$ would involve only three rows of $\mathbf{X}$, namely the i-th, j-th and i'-th rows. Accordingly, it suffices to form the corresponding submatrix $\mathbf{Y}^j \in \mathbb{R}^{3\times(N+1)}$ as decision variables with $\forall n = 1, \ldots, N+1$:

$$Y_{1,n}^j = z^j X_{i,n} \ , \tag{2.37a}$$

$$Y_{2,n}^j = z^j X_{j,n} \ , \tag{2.37b}$$

$$Y_{3,n}^j = z^j X_{i',n} \ . \tag{2.37c}$$

As a result, matrix $\mathbf{\Delta}^j$ is a sparse matrix with all nonzero elements listed here:

$$\mathbf{\Delta}_{i,n}^j = b_{ij}\left(Y_{2,n}^j - Y_{1,n}^j\right), \tag{2.38a}$$

$$\mathbf{\Delta}_{j,n}^j = b_{ij}\left(Y_{1,n}^j - Y_{3,n}^j\right), \tag{2.38b}$$

$$\mathbf{\Delta}_{i',n}^j = b_{ij}\left(Y_{3,n}^j - Y_{2,n}^j\right). \tag{2.38c}$$

The same technique can be applied to reduce the number of variables in the product of $\mathbf{W}^p$ and $\boldsymbol{\delta}^p$ in (2.36a). For the example of generation, due to the definition of $\boldsymbol{\delta}^g$ in (2.30), the operation only depends on the i-th and i'-th columns of $\mathbf{X}$. Hence, one only needs to introduce a submatrix $\mathbf{W}^g \in \mathbb{R}^{(N+1)\times 2}$ as decision variables and form the product similar to (2.38). Thanks to the sparsity of matrix/vector in (2.36), we can reduce the number of decision variables from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, and same for the number of linear inequality constraints in (2.36b). This simplification step will allow for achieving the identification solution more efficiently by the MILP solvers.

2.3.4 Bounds on Continuous Variables

With the MILP reformulation given in (2.36), the problem remains to derive the upper/lower bounds of continuous variables in $\mathbf{X}$. In general, the tighter the bounds are, the faster the MILP can be solved; see e.g., [193]. Hence, we will aim to find reasonably good bounds for entries of $\mathbf{X}$.

To this end, first notice that each entry X_{mn} is non-negative; i.e., the lower bound $X_{mn}^{\min} = 0$ holds. This is because the Bbus matrix $\tilde{\mathbf{B}}(\boldsymbol{z})$ is an M-matrix [201] with its off-diagonal entries being non-positive and real eigenvalues being non-negative, and thus the entries of its inverse are non-negative. To obtain the upper bound $X_{mn}^{\max}$, it is possible to directly maximize the value of X_{mn} under the same constraints in the optimization problem (2.36), as given by

$$\max_{\boldsymbol{z},\mathbf{X},\{\mathbf{Y}^j\},\{\mathbf{W}^p\},X^{\max}} \quad X^{\max} \tag{2.39a}$$

$$\text{s.t.} \quad X_{mn} \leq X^{\max}, \ \forall (m,n) \tag{2.39b}$$

$$(2.32), \ (2.34) \text{ and } (2.35), \ \forall \ (m,n). \tag{2.39c}$$

Compared with (2.36), this new problem (2.39) differs only in the objective function and it is an MILP as well. Clearly, its optimum solution of $X^{\max}$ defines an upper bound for every entry of $\mathbf{X}$. Notice that to solve (2.39), we need to start with a rough estimate of the upper bound of X_{mn} to tackle the inequality constraints in (2.34) and (2.35). This is possible by using a very large value of the upper bound estimate which is refined by (2.39). Since this upper bound holds for any choice of $\boldsymbol{z}$, this refinement step can be performed off-line and does not affect the real-time identification time.

2.3.5 Numerical Results

We use the IEEE 14-bus and 300-bus test cases to demonstrate the identification of bus split events. The 14-bus case allows to better illustrate the system-wide effects of bus split, while the 300-bus case is used to provide quantifiable identification error performance. The optimization problems have been implemented on a regular laptop with Intel® CPU @ 2.60 GHz and 12 GB of RAM in the MATLAB® R2018a simulator with power system modeling provided by MATPOWER [303]. The reformulated MILP-based identification problems are computed by Gurobi.

2.3.5.1 IEEE 14-bus System

The IEEE 14-bus system consists of 20 lines and 5 conventional generators. The proposed identification algorithm has been shown effective for all possible bus split scenarios (excluding the islanding ones) for this system. All types of bus injections, generation only, load only, or the combination of both, have been considered as well. Instead of listing all the results, we pick the split of bus $i = 13$ as an illustrative example. The first neighbor buses of bus $i = 13$ is given by $\mathcal{N}_i = \{6, 12, 14\}$. The line $(13, 14)$ and load d_i are reconnected to the new bus $i' = 15$ after the bus split.

Upon solving the optimization problem (2.36) for every bus location, we plot the resultant minimum mismatch error in Fig. 2.9. As any but split at bus 8 leads to system islanding, it is excluded from the comparison. The mismatch error at bus 13 is the smallest among all possible buses, and thus it is identified as the correct location. Moreover, the optimal solution for this bus has $\boldsymbol{z} = [0; 0; 1]$, which correctly indicates that line $(13, 14)$ is reconnected to the new bus while $z^d = 1$ identifies the reconnected load. Using the optimal solution for each bus location, we can find the corresponding system-wide phase angle difference based on dc model, namely $[\tilde{\mathbf{B}}(\boldsymbol{z})]^{-1}\tilde{\mathbf{p}}(\boldsymbol{z}) - \hat{\boldsymbol{\theta}}$. Fig. 2.10 plots the resultant angle difference solutions for selected bus locations as compared to the actual values $\boldsymbol{\delta}$ from the ac power flow model as observed by PMUs. This comparison again confirms that the split at bus 13 has the closest match with the actual system responses.

It is worth mentioning that this test also points out how to efficiently

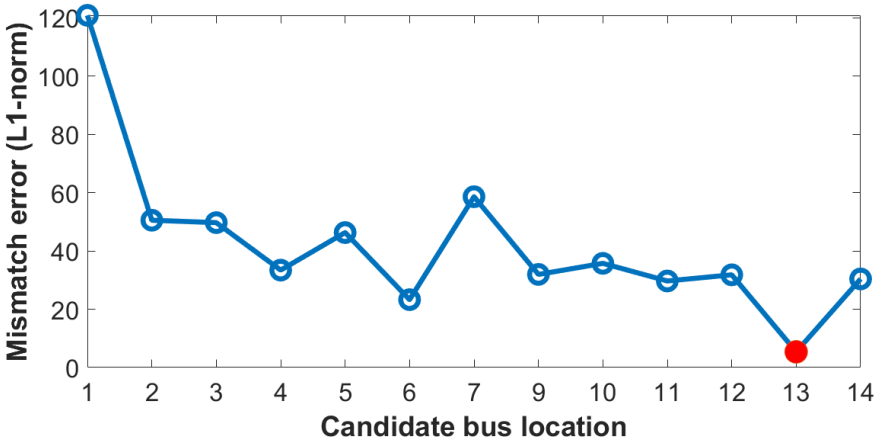


Figure 2.9: Comparison of the minimum error objective value achieved at each bus location, with the lowest value at bus 13 as the correct location.

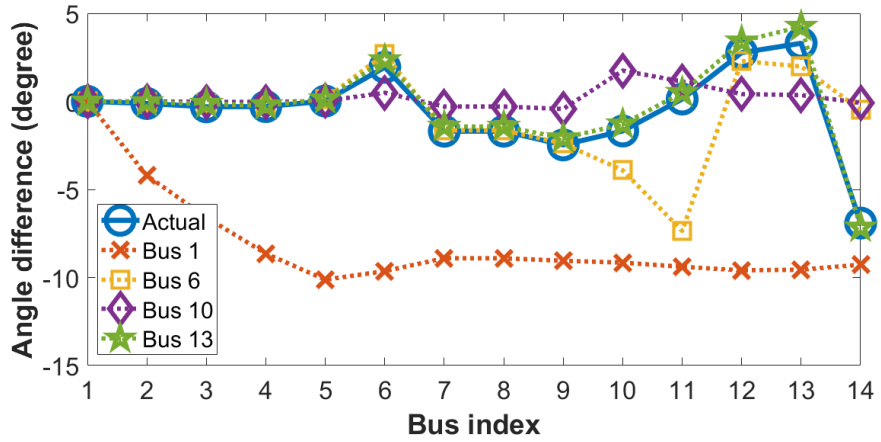


Figure 2.10: System-wide phase angle difference given by selected bus locations as compared to the actual values.

select candidate bus locations for larger-sized systems. As shown by Fig. 2.10, the actual phase angles at bus 13 and its neighbor buses $\mathcal{N}_{13} = \{6, 12, 14\}$ change more evidently than other buses. As the inverse of Bbus matrix is typically diagonally dominant, the effects of topology changes tend to reduce

from the change location to buses further away; see e.g., [300]. This localized impact property enables to adopt a simple yet effective approach to prioritize candidate bus locations based on their angle differences, which will turn out to be very useful for large systems as detailed soon.

2.3.5.2 IEEE 300-bus System

We have performed more comprehensive validations using the IEEE 300-bus system that provide quantifiable identification performance. We first investigate the bus split at the substation of bus $i = 120$, and then provide more quantitative results for different bus locations. Finally, the impact of partial observation is also studied.

Test Case 1: We pick the split of bus $i = 120$ to investigate. It has 3 neighboring buses, and is connected with both generation and load. Compared to the 14-bus system, there are a larger number of buses in this system, making it difficult to enumerate all the buses as candidate buses. Using the empirical insights earlier on, we use a heuristic selection scheme that ranks the buses based on the observed value of phase angle difference, namely $|\delta_i|$ provided by the synchrophasor data. If not all buses are equipped with PMUs, as later in Test Case 3, the neighboring buses of the highly-ranked buses would also be included as candidate locations. For the split of bus $i = 120$, the top six buses in the ranking are buses 120, 153, 151, 152, 155, and 154, most of which are co-located in the same area to bus 120. These six buses are selected to run the optimization problem. Accordingly, the bus with the smallest achievable

error objective value is deemed as the location of bus split, along with its corresponding optimal solution.

This identification process has been run for every possible topology scenario (every choice of binary $\boldsymbol{z}$) for the split of bus 120. The identification accuracy is evaluated based on the percentage of correctly recovered entries of vector $\boldsymbol{z}$, which is given in Fig. 2.11 for every topology scenario. For a majority of scenarios, the identification accuracy is perfect. This implies that the angle differences for each of these scenarios sufficiently differentiate from those of all other scenarios, and our proposed solution can effectively find the correct scenario. Nonetheless, this is not the case for scenario 3, where the connections for two out of the three lines have been erroneously identified. A closer look at this scenario reveals that the dc approximation error has led to that the mis-identified $\boldsymbol{z}$ solution has the smallest mismatch error with the nonlinear ac model. Overall, the proposed method can correctly find the optimal solution to (2.25), while the accuracy of latter may still depend on the approximation error of the dc model.

Test Case 2. We have further tested our proposed identification algorithm for the split events at 16 selected buses, giving rise to 116 topology scenarios in total. To quantify the identification accuracy for each bus, we average the percentage of correct recovery over all the scenarios under the split of that bus. The resultant accuracy results for the original 300-bus system are given in Fig. 2.12, along with those for a modified 300-bus system. Test Case 1 pointed out the impact of the dc power flow approximation, which is

highly related to the line resistance-reactance ratio [239]. Some of the transmission lines in the original 300-bus system have a quite high ratio value, which is uncommon for high-voltage grids. Hence, we reduce the resistance values therein by half to attain a modified system. Overall, our proposed algorithm has achieved effective identification of bus split events in all the 16 selected buses. On average, the original system gives an accuracy of 97.6%, with a minimum accuracy of 83.3% at bus 45. By decreasing the resistance values, the modified system enjoys an increased average accuracy of 99.2%. Hence, our identification algorithm can attain accurate results for practical systems.

To evaluate the computational complexity of the proposed algorithm, the mean and median values for the run-time of all topology scenarios per bus for the original 300-bus system are given in Fig. 2.13. For this 300-bus system, it takes around 100 seconds to execute the proposed identification algorithm using a standard computer. This run-time scales nicely with the size of system. Improved parameter settings (e.g., the upper bound $X_{mn}^{\max}$) and high-performance computing resources can further facilitate the implementation of the proposed algorithm for real-time monitoring of bus split events.

Test Case 3. This test investigates the effects of partial observation using the modified system of reduced resistance. As mentioned in Sec. **??**, the problem (2.25) can be easily modified to allow for partial angle or additional line flow measurements. Using the benchmark of full angle measurements, we compare three partial observation settings: 70% angle measurements, 85% angle measurements and a combination of 70% angle measurements and 50%

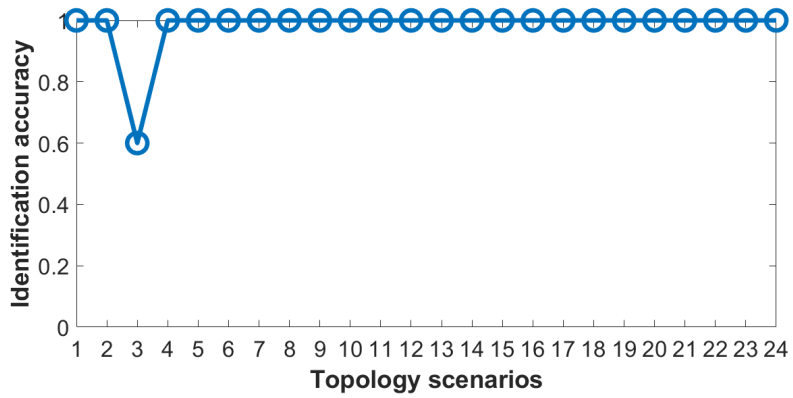


Figure 2.11: Identification accuracy of different scenarios for the bus split in Test Case 1.

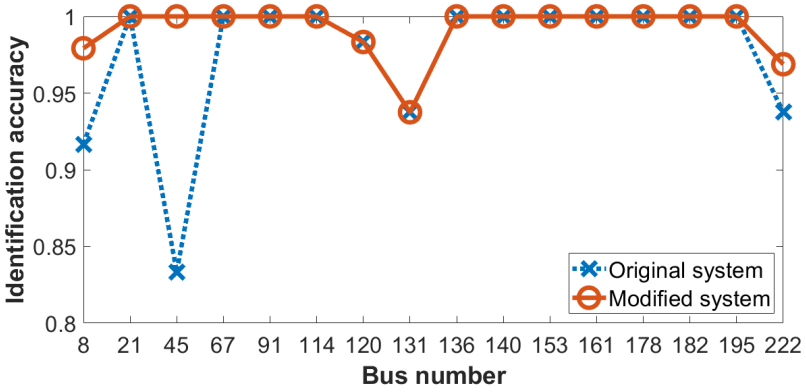


Figure 2.12: Identification accuracy of bus split at selected bus locations in Test Case 2.

line flow measurements. For the partial angle measurements, it suffices to use a selection matrix to extract the mismatch error vector in (2.25) for the metered locations. As for the line flow measurements, they can be incorporated into the error mismatch objective similar to angle measurements by recognizing the linear relation in (2.2). All the measurement locations have been randomly selected and the mismatch error vector is normalized to match the scaling

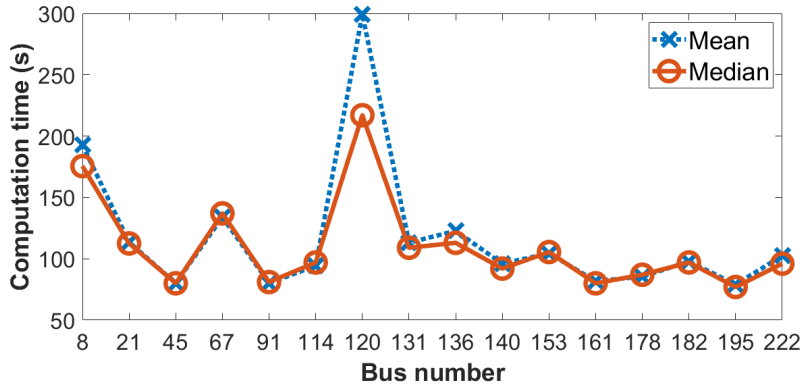


Figure 2.13: Computation time for identification for selected bus split locations in Test Case 2.

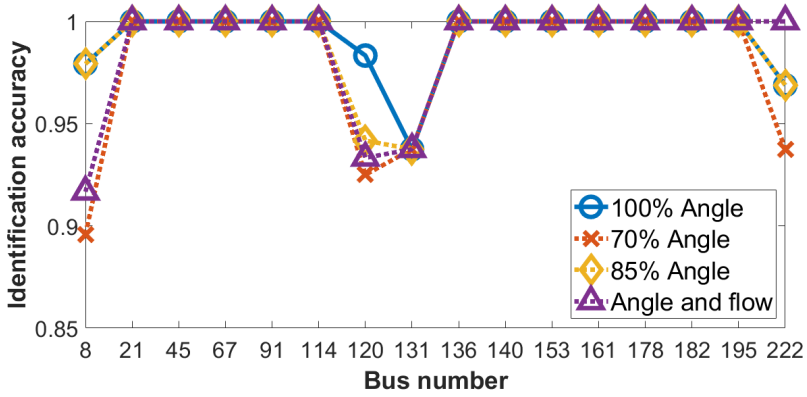


Figure 2.14: Identification accuracy for three partial observation cases compared to the benchmark in Test Case 3.

difference between angle and flow changes. Specifically, the ratio between the average of the absolute angle changes and that of the absolute line flow change is used to scale these two types of measurements.

Fig. 2.14 plots the identification accuracy for the partial observation settings along with the benchmark. All the three settings have achieved sat-

isfactory accuracy (on average 98.1%, 98.9% and 98.7% for each of the three settings), with a minimum accuracy of 89.6% at bus 8 under the 70% angle setting. Generally speaking, the accuracy consistently increases when there are more measurements of angles or line flows. Note that for the split events at bus 222, the inclusion of line flow data is very useful for achieving perfect identification results. At some other buses (bus 8 or 120), the additional angle information is more helpful for identifying certain split events. Therefore, the proposed algorithm can adapt to various measurement availability conditions and provide an accurate identification performance for large systems.

Chapter 3

Topology Optimization Under Uncertainty

This chapter focuses the optimal transmission switching (OTS) problem under uncertainty. We first present the system modeling and OTS formulation in Section 3.1. Using the analysis from Chapter 2, in Section 3.2, we incorporate substation bus splits into the OTS problem for improved operational efficiency. Lastly, to enable intermittent energy sources into the OTS problem, in Section 3.3, we mainly study the distributionally robust chance constrained (DRCC) formulation which can provide better robustness guarantee compared with the traditional chance constrained formulation.

3.1 System Modeling

Power system topology optimization is becoming increasingly important for efficient power system operations, thanks to its capability of effectively relieving network congestion and reducing generation cost. Utilizing

---

This chapter is based on the following publications to which the coauthors contributed equally:

Yuqi Zhou, Ahmed S Zamzam, Andrey Bernstein, and Hao Zhu. "Substation-level grid topology optimization using bus splitting." In American Control Conference (ACC), 2021.

Yuqi Zhou, Hao Zhu, and Grani A Hanasusanto. "Distributionally robust chance-constrained optimal transmission switching for renewable integration." IEEE Transactions on Sustainable Energy, 2022.

optimal transmission switching (OTS) to reduce the network dispatch cost has been first looked at in [93, 123, 141]. Related work on OTS can be roughly classified into the following categories. OTS under uncertain renewable generation [74, 208, 233]; OTS incorporating substation-level breaker activities [124, 224, 280]; AC based OTS [26, 104, 145, 236]; and fast computation of OTS [95, 164, 220].

In this work, we first look at the substation-level grid topology optimization that can incorporate bus splitting at high-voltage substations, which will be introduced in Section 3.2. After that, we also investigate the OTS that can account for uncertain renewables in real-time operation, which will be provided in Section 3.3.

Consider a transmission system with N buses collected in the set $\mathcal{N} := \{1, \ldots, N\}$ and L lines in $\mathcal{L} := \{(i, j)\} \subset \mathcal{N} \times \mathcal{N}$. For bus i, let θ_i be its phase angle and collect all the angles in $\boldsymbol{\theta} \in \mathbb{R}^N$. Similarly, let $\boldsymbol{g},\ \boldsymbol{d} \in \mathbb{R}^N$ denote the vectors of generation and load at all buses, respectively. Under the dc power flow model, line flows $\{f_{ij}\}$ which are collected in $\boldsymbol{f} \in \mathbb{R}^L$ are given by:

$$\boldsymbol{f} = \mathbf{K}\boldsymbol{\theta} \tag{3.1}$$

with the matrix $\mathbf{K} \in \mathbb{R}^{L \times N}$ mapping the phase angles to line flows. The row of $\mathbf{K}$ corresponding to line (i, j) is $b_{ij}(\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T}$, where b_{ij} is the inverse of the line (i, j) reactance. Furthermore, the network power flow conservation leads to:

$$\boldsymbol{p} = \mathbf{A}\boldsymbol{f} \tag{3.2}$$

where $\boldsymbol{p} = \boldsymbol{g} - \boldsymbol{d}$ is the net injection vector and $\mathbf{A} \in \mathbb{Z}^{N \times L}$ is the incidence matrix for the underlying graph $(\mathcal{N}, \mathcal{L})$. Substituting (3.1) into (3.2) yields the dc power flow model:

$$\boldsymbol{p} = \mathbf{B}\boldsymbol{\theta} \tag{3.3}$$

where the so-termed Bbus matrix $\mathbf{B} \in \mathbb{R}^{N \times N}$ is given by:

$$\mathbf{B} = \sum_{(i,j) \in \mathcal{L}} b_{ij} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^{\mathsf{T}}. \tag{3.4}$$

3.1.1 Optimal Transmission Switching

The optimal transmission switching (OTS) problem aims to determine the optimal grid topology while minimizing the total generation cost under given load $\boldsymbol{d}$. For simplicity, linear generation cost is considered and $\boldsymbol{c} \in \mathbb{R}^N$ denotes the known vector of generation cost coefficients. We introduce the vector of binary decision variables $\boldsymbol{z} \in \mathbb{R}^L$ to indicate the transmission line status (1: closed, 0: open). Under a maximum number of L_o open lines, the OTS problem is formulated as a mixed-integer linear program (MILP), given by

$$\min \quad \boldsymbol{c}^{\mathsf{T}} \boldsymbol{g} \tag{3.5a}$$

$$\text{s.t.} \quad \boldsymbol{g} \in \mathbb{R}^N, \boldsymbol{\theta} \in \mathbb{R}^N, \boldsymbol{f} \in \mathbb{R}^L, \boldsymbol{z} \in \mathbb{Z}^L \tag{3.5b}$$

$$\boldsymbol{g}^{\min} \leq \boldsymbol{g} \leq \boldsymbol{g}^{\max} \tag{3.5c}$$

$$\boldsymbol{\theta}^{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^{\max} \tag{3.5d}$$

$$\boldsymbol{D}_f^{\min} \boldsymbol{z} \leq \boldsymbol{f} \leq \boldsymbol{D}_f^{\max} \boldsymbol{z} \tag{3.5e}$$

$$\mathbf{A}\boldsymbol{f} = \boldsymbol{g} - \boldsymbol{d} \tag{3.5f}$$

$$\mathbf{K}\boldsymbol{\theta} - \boldsymbol{f} + \boldsymbol{D}_{\mathrm{M}}(\mathbf{1} - \boldsymbol{z}) \geq \mathbf{0} \tag{3.5g}$$

$$\mathbf{K}\boldsymbol{\theta} - \boldsymbol{f} - \boldsymbol{D}_{\mathrm{M}}(\mathbf{1} - \boldsymbol{z}) \leq \mathbf{0} \tag{3.5h}$$

$$\mathbf{1}^{\mathsf{T}}\boldsymbol{z} \geq L - L_o. \tag{3.5i}$$

The diagonal matrices $\boldsymbol{D}_f^{\min}$ and $\boldsymbol{D}_f^{\max}$ collect the given lower/upper limits of line flow, respectively. In addition, each diagonal entry of the diagonal matrix $\boldsymbol{D}_{\mathrm{M}}$ is a positive constant M_{ij} corresponding to line (i, j), which will be defined soon.

Constraints in (3.5) are discussed in detail here. Generation, phase angle and line flow limits are set in (4.1c)-(4.1e) according to system operating limits. Note that for any open line ($z_{ij} = 0$), its flow f_{ij} is set to zero in (4.1e). Constraint (4.1f) enforces network power balance in (3.2). Moreover, the pair of constraints (4.1g)-(4.1h) are introduced for the line flow model in (3.1). For any closed line ($z_{ij} = 1$), the two inequalities are equivalent to an equality constraint as in (3.1). Otherwise, under $z_{ij} = 0$ and thus $f_{ij} = 0$ [cf. (4.1e)], these two constraints respectively become $b_{ij}(\theta_i - \theta_j) + \mathrm{M}_{ij} \geq 0$ and $b_{ij}(\theta_i - \theta_j) - \mathrm{M}_{ij} \leq 0$, both of which hold for a sufficiently large constant M_{ij} and are thus redundant. This is known as the *Big-M* method [111], which is popular for handling constraints with binary variables. Typically, smaller M_{ij} values can reduce the computation time for (3.5). Thus, for each line (i, j), we set:

$$\mathrm{M}_{ij} \coloneqq b_{ij}\Delta\theta_{ij}^{\max} \tag{3.6}$$

where $\Delta\theta_{ij}^{\max}$ is a given line limit for angle stability [cf. (4.1d)]. Lastly, (4.1i) limits the total number of open lines.

3.2 Topology Optimization Using Substation Bus Split

Grid topology optimization is becoming increasingly important for efficient power system operations, thanks to its capability of effectively relieving network congestion and reducing generation costs. Varying the topology of power networks mainly relies on the operations of switching devices such as circuit breakers (CBs) within the electrical substations. The switching of CBs not only disconnects transmission lines and generation/load, but also can result in *bus splitting* [262, Ch. 11]. A comprehensive topology optimization framework that includes all types of topological changes can greatly enhance the benefits of reducing generation costs while improving the security of grid operations.

A majority of grid topology optimization work has mainly focused on the search of line-switching actions [93, 123, 208, 220, 297], and thus have overlooked the potentials of using bus-splitting operations. The switching of substation CBs was explored in [176, 224, 280] as a corrective measure for relieving localized grid stress caused by line overloads or voltage violations. These methods have been developed to target localized contigencies in power networks by analyzing a small subset of candidate CB actions, while not yet considering a global search for the economic benefits of the full grid. In [124], a topology optimization method based on generalized substation and CB modeling

was proposed to help reduce the total generation costs. Nonetheless, a pre-screening heuristic was utilized to address the scalability issue of the optimization problem therein to allow for real-time implementation. The optimality of the resultant topology solutions is questionable and the optimality gap is unclear. In fact, modeling the CB actions typically requires the detailed node-breaker representation of the power grid that includes the full list of substation components; see, e.g., [124, 196, 198, 205]. The complexity of this representation is the major cause of the lack of scalability as the resultant scenarios can be redundant. Thus, it still remains open to develop an efficient grid topology optimization algorithm that can account for substation-level topology change.

The goal of this work is to develop an efficient real-time topology optimization algorithm that can incorporate the substation-level topology change such as bus splitting. To address the scalability issue of node-breaker representation, we leverage an equivalent bus-branch model for the substation bus splitting. Hence, instead of explicitly modeling all the components within the substation, we can conveniently incorporate this concise equivalent model for bus splitting into a grid topology optimization formulation. To deal with the bilinear terms in the resultant formulation, we apply the McCormick relaxation technique and attain an *exact* mixed-integer linear program (MILP) reformulation, which can be efficiently solved for real-time implementation. Therefore, the main contribution of our work is to provide a tractable algorithm to effectively search for all possible topology change, both line switching and bus splitting, in order to attain the best grid-wide economic and security

benefits.

3.2.1 Nonlinear Optimization Formulation

In electrical networks, switching equipment such as CBs and isolators are usually installed in substations to allow for flexible network topology and emergency intervention. Under certain CB configurations, the substation bus can become electrically disconnected, commonly termed as "bus splitting" or "bus split." The occurrence of bus splitting is increasingly frequent due to misoperations of CBs [140, 150] or malicious cyberattacks [76, 133, 244, 290]. Fig. 3.1 shows an example bus split event for a specific node-breaker substation configuration (double bus double breaker arrangement). Solid (hollow) squares represent closed (open) breakers. If the circled CBs become open, bus i is split into two different buses, i and i'. Accordingly, transmission lines, generation and load can be reconnected to the new bus i'. Although the two buses are physically co-located in the same substation, they become electrically disconnected, leading to a different bus-branch model as shown in Fig. 3.2.

The grid-wide impact of the bus split topology change has been analyzed in [295] and is summarized in the following proposition.

Proposition 3.2.1. *Consider the split of bus i with a single line (i, j) and injection $\tilde{p}_i$ reconnected to the new bus i' (shown in Fig. 3.2). The post-split system is equivalent to having the opening of line (i, j) and an additional power transfer $\tilde{p}_i$ between buses i and j.*

The equivalent model for the post-split system is demonstrated in

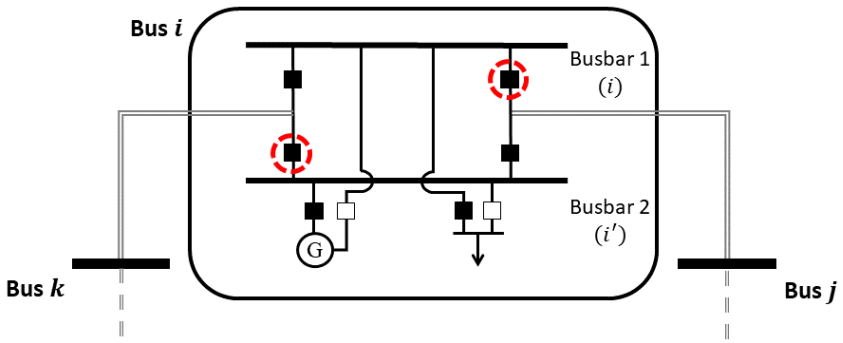


Figure 3.1: Opening of the circled breakers leads to a bus split event at bus i.

Fig. 3.3. Because the new bus i' connects only to bus j, it can be eliminated from the system by moving its connected injection ($\tilde{p}_i = g_i$ in this generation-only case) directly to bus j. Compared to the original system shown in Fig. 3.2, the equivalent system experiences the opening of line (i, j) in addition to a power transfer of $\tilde{p}_i$ between buses i and j. This equivalent model has also been verified by linear sensitivity analysis for the bus split events [295]. Proposition 4.3.1 is very useful for simplifying the incorporation of bus split events into the topology optimization problem, as discussed in the ensuing section.

The grid topology optimization problem aims to determine the optimal transmission grid topology with associated generation outputs in order to minimize the total generation cost. The feasible region of generation dispatch for this problem is the union of the sets of feasible solutions corresponding to each topology configuration. Thus, varying the grid topology will likely expand the overall feasible region and accordingly reduce the total generation

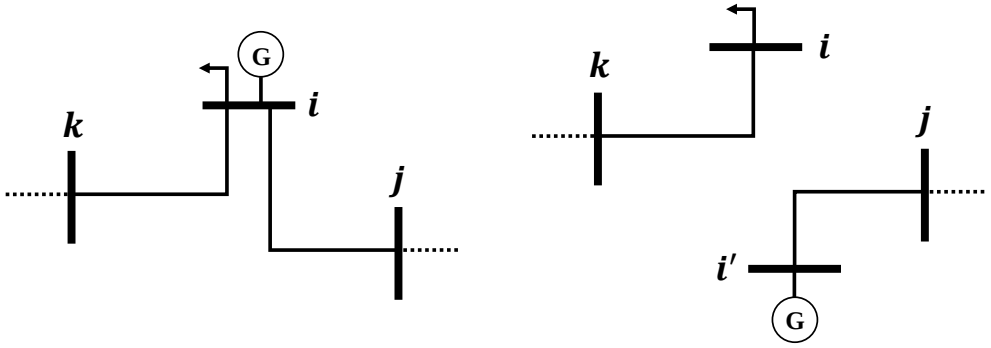


Figure 3.2: (Left) Original bus-branch model and (right) model with a bus split at bus i.

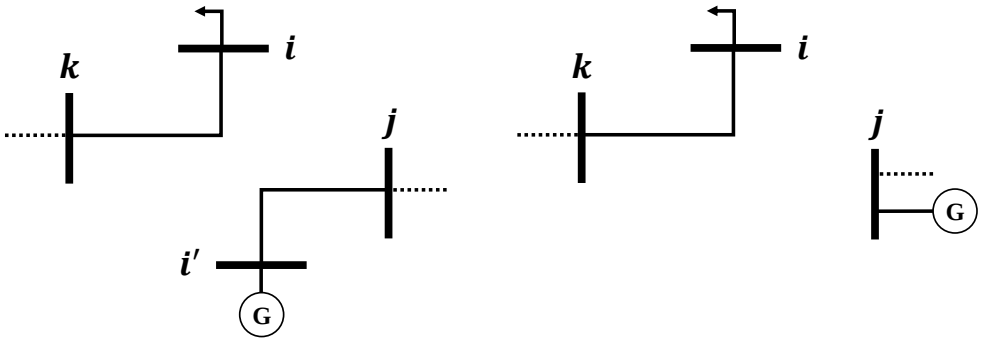


Figure 3.3: Equivalent reduced bus-branch model (right) for the (left) system with bus split.

costs [93,123]. Going beyond the traditional topology optimization framework, the inclusion of substation level bus split events allows for additional topology change, and hence it can further reduce grid congestion and improve the security of grid operations.

The nonlinear ac topology optimization formulation is well known to suffer from scalability issues, greatly challenging its real-time implementation [21, 236, 278].

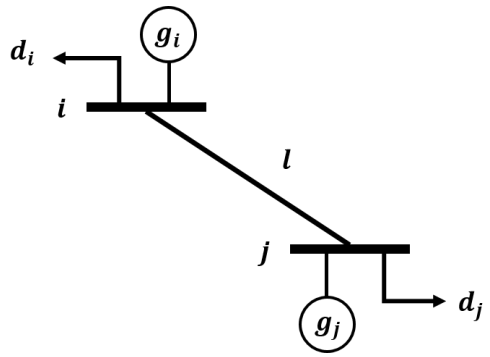


Figure 3.4: Line $\ell = (i, j)$ with connected generation/load for modeling power transfer between buses i and j.

Remark 3.2.1. In this work, we adopt the dc power flow model for formulating topology optimization problem. The proposed dc based model can be possibly generalized to the nonlinear ac formulation as well; see e.g., [145]. To corroborate the validity of the dc model, we will provide several numerical tests to assert the feasibility of the switching solutions under the ac power flow model.

3.2.2 Modeling of Power Transfer in Bus Splitting

We first discuss different scenarios of generation/load connection for modeling the power transfer in Proposition 4.3.1. To this end, consider a transmission line $\ell = (i, j)$ that connects to buses i and j, as shown in Fig. 3.4. Without loss of generality, the case of having both generation and load is assumed for the two buses. The bus split event can result in a model that is mathematically equivalent to a power transfer in between. For instance, the split of bus i can be associated with one of three power transfer scenarios from bus i to bus j, namely load only, generation only, and generation plus load.

To represent the change of power injection for all three scenarios, define the following $N \times 3$ matrix:

$$\boldsymbol{\Delta}_{\ell,i}(\boldsymbol{g}) = (\mathbf{e}_i - \mathbf{e}_j)\,[d_i \quad -g_i \quad d_i - g_i] \tag{3.7}$$

where $\mathbf{e}_i \in \mathbb{R}^{N\times 1}$ denotes the standard basis vector. Each column of $\boldsymbol{\Delta}_{\ell,i}(\boldsymbol{g})$ corresponds to one of three aforementioned scenarios under the split of bus i. Similarly, one can define this power injection matrix for the split of bus j, as:

$$\boldsymbol{\Delta}_{\ell,j}(\boldsymbol{g}) = (\mathbf{e}_j - \mathbf{e}_i)\,[d_j \quad -g_j \quad d_j - g_j]\,. \tag{3.8}$$

Notice that both $\boldsymbol{\Delta}_{\ell,i}(\boldsymbol{g})$ and $\boldsymbol{\Delta}_{\ell,j}(\boldsymbol{g})$ depend on the generation output $\boldsymbol{g}$, which is a decision variable. In what follows, we use $\boldsymbol{\Delta}_{\ell,i}$ to refer to $\boldsymbol{\Delta}_{\ell,i}(\boldsymbol{g})$ when the dependence on $\boldsymbol{g}$ is clear from the context.

Upon defining the power injection matrices, we are ready to formulate the topology optimization problem that includes the bus split operation. To this end, consider the linear generation cost model, with $\boldsymbol{c} \in \mathbb{R}^N$ collecting all the linear coefficients. The binary decision variable z_ℓ is introduced for each transmission line $\ell = (i, j)$ to indicate the *equivalent line status* (1: closed, 0: open), which will be explained in more detail after the formulation. The incident buses for line ℓ are collected in the set $\mathcal{N}_\ell$. A vector of binary variables $\boldsymbol{w}_{\ell,i} \in \{0, 1\}^3$ is used to select the power transfer scenario in case of a bus split at bus i, leading to an equivalent outage on line $\ell = (i, j)$. Similarly, vector $\boldsymbol{w}_{\ell,j} \in \{0, 1\}^3$ is defined to select the power transfer scenario in case of a bus split at bus j, with an equivalent outage on line $\ell = (i, j)$. Under a

maximum budget of s operations (either line switching or bus splitting), one can formulate the following optimization problem:

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^{\mathsf{T}}\boldsymbol{g} && \text{(3.9a)}\\
\text{over} \quad & \boldsymbol{\theta} \in \mathbb{R}^N, \boldsymbol{g} \in \mathbb{R}^N, \boldsymbol{f} \in \mathbb{R}^L, z_\ell \in \{0,1\},\ \forall \ell \in \mathcal{L} \\
& \boldsymbol{w}_{\ell,i} \in \{0,1\}^3, \boldsymbol{w}_{\ell,j} \in \{0,1\}^3,\ \forall \ell = (i,j) \in \mathcal{L} \\
\text{s.t.} \quad & \theta_i^{\min} \le \theta_i \le \theta_i^{\max},\ \forall i && \text{(3.9b)}\\
& g_i^{\min} \le g_i \le g_i^{\max},\ \forall i && \text{(3.9c)}\\
& f_\ell^{\min} z_\ell \le f_\ell \le f_\ell^{\max} z_\ell,\ \forall \ell && \text{(3.9d)}\\
& b_{ij}(\theta_i - \theta_j) - f_\ell + (1 - z_\ell)\mathrm{M}_\ell \ge 0,\ \forall \ell = (i,j) && \text{(3.9e)}\\
& b_{ij}(\theta_i - \theta_j) - f_\ell - (1 - z_\ell)\mathrm{M}_\ell \le 0,\ \forall \ell = (i,j) && \text{(3.9f)}\\
& \sum_\ell (1 - z_\ell) \le s && \text{(3.9g)}\\
& \mathbf{1}^{\mathsf{T}}\boldsymbol{w}_{\ell,i} + \mathbf{1}^{\mathsf{T}}\boldsymbol{w}_{\ell,j} \le 1 - z_\ell,\ \forall \ell = (i,j) && \text{(3.9h)}\\
& \sum_{\ell: i \in \mathcal{N}_\ell} \mathbf{1}^{\mathsf{T}}\boldsymbol{w}_{\ell,i} \le 1,\ \forall i && \text{(3.9i)}\\
& \mathbf{A}\boldsymbol{f} = \boldsymbol{g} - \boldsymbol{d} + \sum_{\ell=(i,j)} \boldsymbol{\Delta}_{\ell,i}\boldsymbol{w}_{\ell,i} + \sum_{\ell=(i,j)} \boldsymbol{\Delta}_{\ell,j}\boldsymbol{w}_{\ell,j} && \text{(3.9j)}\\
& f_\ell^{\min}\mathbf{1} \le \boldsymbol{\Delta}_{\ell,i}\boldsymbol{w}_{\ell,i} + \boldsymbol{\Delta}_{\ell,j}\boldsymbol{w}_{\ell,j} \le f_\ell^{\max}\mathbf{1},\ \forall \ell = (i,j) && \text{(3.9k)}
\end{aligned}
$$

We discuss the constraints for problem (3.9) here. Phase angle and generation limits are given in constraints (3.9b) - (3.9c). Line flow limits are given in (3.9d), while the flow f_ℓ is enforced to be zero when the line is open; i.e., $z_\ell = 0$. Constraints (3.9e) - (3.9f) are introduced for establishing the line flow model in (3.1), with the constant M_ℓ being sufficiently large. When the

line $\ell = (i, j)$ is closed, the two inequalities are equivalent to the dc power flow equation $f_\ell = b_{ij}(\theta_i - \theta_j)$. Otherwise, when the line is open $f_\ell = 0$ [cf. (3.9d)], the two constraints are guaranteed to be inactive for a large M_ℓ. This is called the *Big-M* method [111], which is often used to handle constraints with binary variables. For each line $\ell = (i, j) \in \mathcal{L}$, we set:

$$\mathrm{M}_\ell \coloneqq b_{ij} \Delta\theta_{ij}^{\max}, \tag{3.10}$$

where $\Delta\theta_{ij}^{\max}$ is a given upper bound for angle stability. Constraint (3.9g) limits the total number of operations including both line switching and bus splitting, and constraint (3.9h) further defines the operations for each line. Specifically, if $z_\ell = 0$ and $\mathbf{1}^\mathsf{T}\boldsymbol{w}_{\ell,i} + \mathbf{1}^\mathsf{T}\boldsymbol{w}_{\ell,j} = 0$, the operation is simply a line switching of $\ell = (i, j)$, i.e., deenergizing the line $\ell = (i, j)$. Otherwise, when $z_\ell = 0$ but $\mathbf{1}^\mathsf{T}\boldsymbol{w}_{\ell,i} + \mathbf{1}^\mathsf{T}\boldsymbol{w}_{\ell,j} \neq 0$, then one of the power transfer scenarios is selected after opening line ℓ, making it equivalent to a bus split at either end of line ℓ. The latter case utilizes the equivalent model of bus split events and does not actually deenergize line ℓ. Therefore, $z_\ell = 0$ itself cannot fully indicate the operation type (line switching or bus splitting) and is called the *equivalent line status* for this reason. Constraints (3.9h) - (3.9k) are introduced specifically for considering bus split events. Constraint (3.9h) limits the number of power transfers that can be selected for a bus split involving the opening of line $\ell = (i, j)$. When the line is closed ($z_\ell = 1$), no power transfer is allowed; Otherwise, when the line is open ($z_\ell = 0$), at most one (0 or 1) power transfer can be made, depending on whether it is line switching or bus splitting. Furthermore, notice that a single bus can be connected to multiple buses, but the power transfer

from bus i to other buses can be made only if the bus is split into two bus bars. Once bus i is split for a power transfer with one of its incident buses, no other power transfer can be made with other buses. Therefore, constraint (3.9i) limits the power transfer from each bus to be at most once due to the physical limit of the substation. Constraint (3.9j) enforces the network power balance in (3.2), where the injection also reflects any power transfer made because of the bus split. Last, (3.9k) guarantees that for the injection reconnected to the new bus i', the power flow on that incident line is not violating the transmission limit of line $\ell = (i, j)$.

3.2.3 Tractable Linear Reformulation

The main challenge of solving (3.9) lies in the nonlinearity of the constraints. Specifically, constraints (3.9j) and (3.9k) are bilinear in the decision variables, due to the multiplication terms, namely $\boldsymbol{\Delta}_{\ell,i}(\boldsymbol{g})\boldsymbol{w}_{\ell,i}$ and $\boldsymbol{\Delta}_{\ell,j}(\boldsymbol{g})\boldsymbol{w}_{\ell,j}$. To address these terms, we propose to adopt the *McCormick relaxation* technique to reformulate the problem that is amenable to off-the-shelf MILP solvers.

First, rewrite the following multiplication as:

$$\boldsymbol{\Delta}_{\ell,i}\boldsymbol{w}_{\ell,i} = \boldsymbol{\delta}_d^i \boldsymbol{w}_{\ell,i} d_i - \boldsymbol{\delta}_g^i \boldsymbol{w}_{\ell,i} g_i \tag{3.11}$$

where vectors $\boldsymbol{\delta}_d^i := (\mathbf{e}_i - \mathbf{e}_j)\begin{bmatrix}1 & 0 & 1\end{bmatrix}$ and $\boldsymbol{\delta}_g^i := (\mathbf{e}_i - \mathbf{e}_j)\begin{bmatrix}0 & 1 & 1\end{bmatrix}$. We define the product of $\boldsymbol{w}_{\ell,i}$ and g_i as:

$$\boldsymbol{y}_{\ell,i} = \boldsymbol{w}_{\ell,i} g_i, \quad \forall\ \ell. \tag{3.12}$$

Under the bounds $[g_i^{\min}, g_i^{\max}]$ for generation g_i, the following four *linear* inequalities hold:

$$\boldsymbol{y}_{\ell,i} \geq \boldsymbol{w}_{\ell,i} g_i^{\min} \tag{3.13a}$$

$$\boldsymbol{y}_{\ell,i} \geq \mathbf{1} g_i + \boldsymbol{w}_{\ell,i} g_i^{\max} - \mathbf{1} g_i^{\max} \tag{3.13b}$$

$$\boldsymbol{y}_{\ell,i} \leq \boldsymbol{w}_{\ell,i} g_i^{\max} \tag{3.13c}$$

$$\boldsymbol{y}_{\ell,i} \leq \mathbf{1} g_i + \boldsymbol{w}_{\ell,i} g_i^{\min} - \mathbf{1} g_i^{\min}. \tag{3.13d}$$

The inequalities (3.13a) - (3.13d) can be verified by substituting (3.12). Conversely, for any binary $\boldsymbol{w}_{\ell,i}$, the linear inequalities in (3.13) also guarantee the validity of (3.12). When the k-th entry in the binary $\boldsymbol{w}_{\ell,i}$ is equal to zero, the two inequalities (3.13a) and (3.13c) jointly force the k-th entry of $\boldsymbol{y}_{\ell,i}$ to be zero. Otherwise, when the k-th entry of $\boldsymbol{w}_{\ell,i}$ is equal to one, the inequalities (3.13b) and (3.13d) enforce that the k-th entry of $\boldsymbol{y}_{\ell,i}$ is equal to g_i. Due to the binary vector $\boldsymbol{w}_{\ell,i}$, the set of inequalities in (3.13) is equivalent to the bilinear relation in (3.12). Reformulating (3.12) with the linear inequalities in (3.13) is known as the *McCormick relaxation* technique, which has been popularly used in other problems of designing grid topology [27, 40].

Hence, the bilinear product as given in (3.11) can be *equivalently* replaced with:

$$\boldsymbol{\Delta}_{\ell,i} \boldsymbol{w}_{\ell,i} = \boldsymbol{\delta}_d^i \boldsymbol{w}_{\ell,i} d_i - \boldsymbol{\delta}_g^i \boldsymbol{y}_{\ell,i} \tag{3.14}$$

and the linear inequality constraints (3.13a) - (3.13d). Similarly, the bilinear product of $\boldsymbol{\Delta}_{\ell,j}$ and $\boldsymbol{w}_{\ell,j}$ can be directly given as:

$$\boldsymbol{\Delta}_{\ell,j} \boldsymbol{w}_{\ell,j} = \boldsymbol{\delta}_d^j \boldsymbol{w}_{\ell,j} d_j - \boldsymbol{\delta}_g^j \boldsymbol{y}_{\ell,j} \tag{3.15}$$

for similarly defined $\boldsymbol{\delta}_d^j$ and $\boldsymbol{\delta}_g^j$, together with the following four linear inequalities:

$$\boldsymbol{y}_{\ell,j} \geq \boldsymbol{w}_{\ell,j} g_j^{\min} \tag{3.16a}$$

$$\boldsymbol{y}_{\ell,j} \geq \mathbf{1} g_j + \boldsymbol{w}_{\ell,j} g_j^{\max} - \mathbf{1} g_j^{\max} \tag{3.16b}$$

$$\boldsymbol{y}_{\ell,j} \leq \boldsymbol{w}_{\ell,j} g_j^{\max} \tag{3.16c}$$

$$\boldsymbol{y}_{\ell,j} \leq \mathbf{1} g_j + \boldsymbol{w}_{\ell,j} g_j^{\min} - \mathbf{1} g_j^{\min}. \tag{3.16d}$$

Thus, we have reformulated the bilinear products $\boldsymbol{\Delta}_{\ell,i}\boldsymbol{w}_{\ell,i}$ and $\boldsymbol{\Delta}_{\ell,j}\boldsymbol{w}_{\ell,j}$ using linear constraints (3.14) and (3.15) followed by additional linear inequalities (3.13a) - (3.13d) and (3.16a) - (3.16d). Therefore, the equivalent topology optimization problem that incorporates substation bus splitting operation can be established in the following proposition.

Proposition 3.2.2 (bounds for big-M constants)**.** *The original nonlinear optimization problem* (3.9) *is equivalent to the following one:*

$$\min \quad \boldsymbol{c}^{\mathsf{T}} \boldsymbol{g} \tag{3.17a}$$

$$\begin{aligned} \textit{over} \quad & \boldsymbol{\theta} \in \mathbb{R}^N, \boldsymbol{g} \in \mathbb{R}^N, \boldsymbol{f} \in \mathbb{R}^L, z_\ell \in \{0,1\},\ \forall \ell \in \mathcal{L} \\ & \boldsymbol{w}_{\ell,i} \in \{0,1\}^3, \boldsymbol{w}_{\ell,j} \in \{0,1\}^3,\ \forall \ell = (i,j) \in \mathcal{L} \\ & \boldsymbol{y}_{\ell,i} \in \mathbb{R}^3, \boldsymbol{y}_{\ell,j} \in \mathbb{R}^3 \qquad\qquad \forall \ell = (i,j) \in \mathcal{L} \end{aligned}$$

$$\textit{s.t.} \quad (3.9\text{b}) - (3.9\text{i}), (3.13\text{a}) - (3.13\text{d}), (3.16\text{a}) - (3.16\text{d}) \tag{3.17b}$$

$$\begin{aligned} \mathbf{A}\boldsymbol{f} = \boldsymbol{g} - \boldsymbol{d} + & \sum_{\ell=(i,j)} (\boldsymbol{\delta}_d^i \boldsymbol{w}_{\ell,i} d_i - \boldsymbol{\delta}_g^i \boldsymbol{y}_{\ell,i}) \\ & + \sum_{\ell=(i,j)} (\boldsymbol{\delta}_d^j \boldsymbol{w}_{\ell,j} d_j - \boldsymbol{\delta}_g^j \boldsymbol{y}_{\ell,j}) \end{aligned} \tag{3.17c}$$

$$f_\ell^{\min}\mathbf{1} \le \boldsymbol{\delta}_d^i \boldsymbol{w}_{\ell,i} d_i - \boldsymbol{\delta}_g^i \boldsymbol{y}_{\ell,i} + \boldsymbol{\delta}_d^j \boldsymbol{w}_{\ell,j} d_j - \boldsymbol{\delta}_g^j \boldsymbol{y}_{\ell,j} \le f_\ell^{\max}\mathbf{1}, \quad \forall \ell = (i,j). \tag{3.17d}$$

This reformulated problem is a mixed-integer linear program (MILP) and can be efficiently solved by common optimization solvers such as CPLEX, MOSEK and Gurobi. Moreover, since this work directly considers a minimal set of decisions on line switching and power transfer as a result of substation bus splits, the dimension of binary decision variables has been greatly reduced from the original one under the detailed node-breaker representation. This reduction does not affect the solution quality, as one can easily recover the underlying CB status and thus determine the corresponding breaker actions. Therefore, our proposed approach has utilized the concise bus-branch representation to attain an efficient and effective transmission grid switching solution using substation-level actions.

3.2.4 Numerical Results

In this section, we first use the IEEE 14-bus system to mainly illustrate that bus split operations can be used to effectively relieve network congestion and to help address feasibility issues of the optimal power flow problem. After that, we perform the substation-level topology optimization on the larger sized IEEE 118-bus system to demonstrate the economic improvement on generation dispatch and to assess the computational complexity of the proposed topology optimization model. The optimization problems have been implemented on a regular laptop with Intel CPU @ 2.60 GHz and 12 GB of RAM

in the MATLAB R2018a simulator. The MILP-based optimization problems are computed using the CPLEX solver.

3.2.4.1 14-Bus System Test

The IEEE 14-bus system consists of 20 transmission lines and 5 conventional generators, and we use the ac power flow model to test the system. The system has been slightly modified to illustrate that the bus splitting can be used to relieve network congestion and thus can help with the feasibility issue of the optimal power flow problem. Specifically, we modify the transmission limits of lines $(2, 3)$ and $(3, 4)$ to be 100 MW and 10 MW, respectively. The maximum generation limit of the generator at Bus 3 is adjusted to be 20 MW, and all other network configurations are kept unaltered.

For Bus 3 in the original system, as shown in Fig. 3.5, a net load of at least $\boldsymbol{d}_3 - \overline{\boldsymbol{g}_3} = 74.2$ MW needs to be satisfied by the flows from line $(2, 3)$ and line $(3, 4)$. Due to the electrical characteristics of the lines, power flows on both lines are governed by the phase angle at Bus 3. Therefore, as we gradually increase the flow on line $(2, 3)$, the transmission limit on line $(3, 4)$ will be eventually violated before the sum of power flows on both lines meets the net load at Bus 3, leading to an *infeasible* solution to the optimal power flow problem. In order to relieve the congestion on line $(3, 4)$ in this scenario, solving the proposed topology optimization problem suggests performing a bus split at Bus 3 such that the generator is connected to bus bar $3'$ and the load is connected to bus bar $3''$. Essentially, this bus splitting decouples

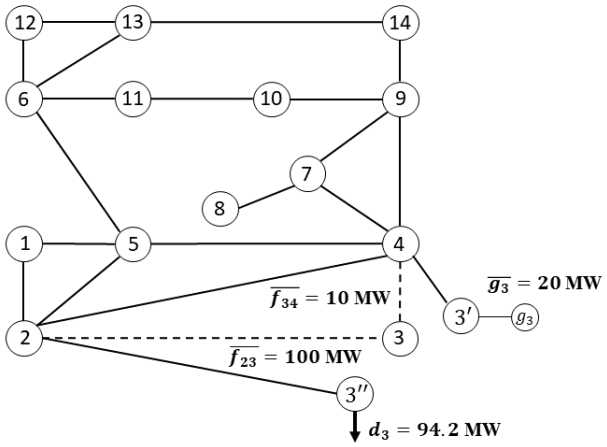


Figure 3.5: Bus splitting at Bus 3 in the IEEE 14-bus system.

the power flows on lines $(2, 3)$ and $(3, 4)$ by allowing each bus bar to have a different phase angle. After the bus splitting, the load at Bus 3 can be easily satisfied by the flow from line $(2, 3)$ without violating any transmission limit. Thus, the ac power flow model of the system for the updated system as shown in Fig. 3.5 becomes *feasible*. In fact, the bus split operation can be easily achieved through switching associated CBs at the substation. The case study on this small system indicates that similar to traditional line switching and load shedding, the operation of bus splitting can be also used to relieve network congestion and help with feasibility issues. Although the model used in (3.17) is a dc power flow model, the solution obtained which suggests a bus splitting at Bus 3 makes the problem feasible while considering the ac power flow model. Next, we will use a larger sized system to illustrate the enhanced economic benefit of the proposed substation-level topology optimization algorithm.

3.2.4.2 118-Bus System Test

The IEEE 118-bus test case is tested for the substation-level topology optimization. The system consists of 118 buses, 186 transmission lines and 19 committed conventional generators. To illustrate the improvement and economic benefits of the topology optimization by incorporating bus split events, we also tested the same system for the traditional topology optimization strategy [93]. This can be easily fulfilled by restricting $\boldsymbol{w}_{\ell,i}$ and $\boldsymbol{w}_{\ell,j}$ in (3.9) to be zero, which will exclude bus splits from consideration and allow for only line-switching operations. By doing so, constraints (3.9h), (3.9i) and (3.9k) always hold and are thus disabled. Meanwhile, (3.9j) becomes a linear constraint, which describes the network power balance without any power transfer. Accordingly, the optimization problem (3.9) itself constitutes an MILP that is readily solvable for common optimization solvers.

The comparison of total cost under different numbers of operations for line switching and breaker-level switching is given in Fig. 3.6 together with the benchmark cost for the system without any topology switching. Compared with the benchmark cost which involves no topology optimization, our proposed breaker-switching strategy achieves total savings of $14.1\% - 23.4\%$, depending on the number of operations; cf. Fig. 3.6. Meanwhile, compared with line switching, it provides additional cost savings of $4.9\% - 7.5\%$ correspondingly. Notice that these additional savings are obtained only by altering the status of several breakers at the substations, therefore the economic benefits are indeed attractive for system operators.

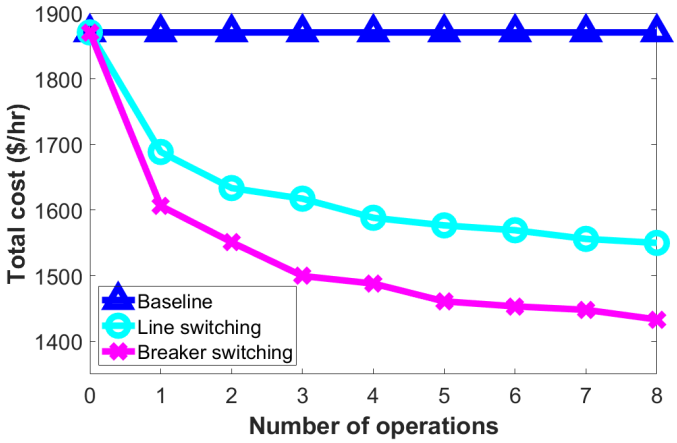


Figure 3.6: Comparison of total cost between line switching and breaker switching for the 118-bus system.

To compare the switching decisions provided by the traditional line switching and the proposed breaker switching, we list the topology optimization solutions for up to a maximum of $s = 5$ operations in Table 3.1. In the line-switching part, only operation on the transmission lines is allowed, which normally involves opening a pair of breakers at both ends of the line. In contrast, the breaker-switching strategy enables more complicated breaker operations that lead to not only line switching but also bus split events. Take $s = 3$ as an example, the line-switching scheme picks lines 41, 128 and 136 to open, whereas the proposed breaker switching suggests opening line 130 and performing bus splits at buses 77, 82 simultaneously. As the result of considering breaker-level operations, an additional reduction of 7.3% in the operational cost is achieved.

Additionally, we assessed the computation time for the two formula-

Table 3.1: Comparison of Switching Decisions for the 118-Bus System

s	Line Switching	Breaker Switching	Reduction
1	Line 128	Bus 82	4.9%
2	Lines 128, 136	Buses 77, 82	5.1%
3	Lines 41, 128, 136	Buses 77, 82 & Line 130	7.3%
4	Lines 119, 123, 124, 125	Buses 75, 77, 82 & Line 136	6.3%
5	Lines 118, 121, 131, 135, 149	Buses 77, 82 & Lines 123, 124, 125	7.3%

tions of the topology optimization scheme shown in Fig. 3.7. For up to $s = 8$ operations, on average the solving time for the substation-level topology optimization is 28.1% faster than the line-switching one. Notice that compared with the line-switching formulation, we further introduce variables such as $\boldsymbol{w}_{\ell,i}$, $\boldsymbol{w}_{\ell,j}$, $\boldsymbol{y}_{\ell}$ and $\boldsymbol{y}_{\ell,j}$, but the additional constraints (3.9h) - (3.9i) that incorporate bus split events can potentially facilitate the computation of the resultant MILP problem. In fact, when the number of operations increases, the additional savings on generation cost usually become less significant. In addition, excessive switching operations can lead to security and stability issues in transmission grids (see e.g., [93]). Therefore, in practice the number of operations is normally restricted to a small number. In the tested 118-bus system, until $s = 5$ operations, the proposed control scheme only requires less than 10 seconds to find the solution. The results imply the efficiency and scalability of the proposed optimization formulation for real-time implementation.

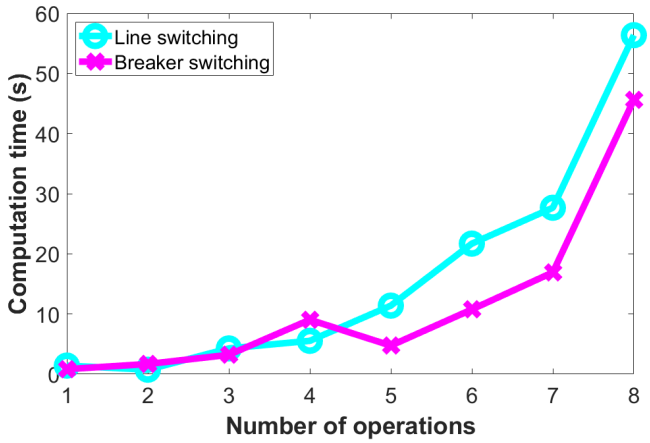


Figure 3.7: Comparison of computation time between line switching and breaker switching for the 118-bus system.

AC feasibility

It is certainly important to ensure that the switching decisions of the topology optimization problem (3.17) are feasible considering the ac power model. Thus, we verified the decisions of the breaker-level topology optimization shown in Table 3.1 using the ac power flow. Our results confirm that the ac power flow model remains feasible in the post-event systems for the proposed topology optimization.

3.3 Topology Optimization for Renewable Integration

Rising renewable penetration in recent years greatly challenges the efficient and reliable operations of power systems. With the increasing uncertainty from renewables, robust decision-making, such as grid topology optimization [93], is of great importance. Judicious line switching along with

generation dispatch can potentially reduce generation costs and renewable curtailment level, yet at possible violations of operational limits (e.g., line power flows). Thus, it is imperative to design optimal transmission switching (OTS) algorithms that ensure guaranteed robustness under the uncertain renewables.

The OTS problem has attracted high interest in recent years in its algorithm design and practical implementations (e.g., [87, 122, 145, 256]). The switching of transmission lines expands the feasible region for generation dispatch decisions and thus relieves grid congestion. Therefore, it can potentially adapt to the varying power transfer needed for renewable generation and reduce the level of renewable curtailment. For example, over 7000 MW of wind capacity was installed in Texas from 2006 to 2009, but major transmission congestion was experienced. Transmission constraints have resulted in excessive wind curtailment, as reported in [44]. Thus, efficiently solving OTS is of great importance to enhance the penetration level of renewable energy to the grid.

As the OTS problem includes additional integer decision variables, extending it to a stochastic/robust optimization framework is more difficult than that for optimal power flow (OPF). Similar to OPF, scenario-based approaches have been largely used to deal with the OTS problem under uncertainty. For example, stochastic topology optimization has been considered in [74, 233] based on wind uncertainty scenarios from known probabilistic models. In addition, the chance-constrained (CC) framework has been developed for the OTS problem in [208], aiming to attain guaranteed constraint violation for a

given uncertainty distribution using the sample-average approximation (SAA) approach. Nonetheless, constructing an accurate distribution for the uncertainty in energy resources can be extremely challenging in practice. Moreover, these approaches typically lead to a mixed-integer program (MIP) in which the problem dimensions quickly grow with the number of samples. This scalability issue results in high computational complexity and makes these scenario-based approaches sub-par for real-time OTS decision-making.

To tackle the scalability issue with scenario-based approaches, some robust/stochastic OTS work [73, 148, 154] invokes a repeating procedure of adding cuts to a master problem using sub-problem solutions. Nonetheless, their computational efficiency can still be problematic while the optimality guarantee is unclear. Instead, our earlier work [297] has proposed a robust OTS algorithm by using the linear decision rule (LDR) technique to approximate second-stage variables, seeking to maintain the operating limits for any uncertainty within a compact set. Even though LDR constitutes merely a linear approximation, the resultant mixed-integer linear program (MILP) has a fixed problem dimension and is efficient to solve. Nonetheless, its robustness under all possible uncertainty scenarios makes the solution unnecessarily conservative. In addition, recent work [189] has considered the distributionally robust chance-constrained (DRCC) OTS problem to account for the ambiguity of uncertainty distribution. Nonetheless, the linearized OTS model therein builds upon line outage sensitivity factors and cannot accurately include multiple, simultaneous topology changes.

Our work aims to provide computationally efficient algorithms for solving the DRCC-OTS problem by developing equivalent reformulations. Notably, we consider an *equivalent* linear reformulation for the integer line status and dc power flow variables. We analytically establish that for the two-stage OTS under linear generation response, the recourse actions (line flows and phase angles) can be represented as linear functions of the uncertainty variables. This linear OTS model is crucial for tractable DRCC reformulation through dualization. Compared to CC-OTS, the proposed DRCC framework seeks dispatch and switching decisions that are robust against the worst-case uncertainty distribution from within a prescribed ambiguity set. Thus, it greatly expands the possible probabilistic models, where variations of uncertainty distributions are common in real-world settings due to the lack of data samples or high variability. The DRCC approaches are of particular importance for enhancing renewable integration because they can provide guaranteed robustness performance as demonstrated by our numerical tests. To provide tractable DRCC-OTS solutions, this work considers moment-based (mean-mean absolute deviation) and distance-based (Wasserstein distance) ambiguity sets for the renewable uncertainty, both of which are amenable to linear reformulations.

The contribution of our work is three-fold.

- We put forth a two-stage DRCC-OTS problem under renewable uncertainty that models real-time linear adjustment of generation output.

- For the proposed two-stage OTS problem, we analytically establish an equivalent LDR-based reformulation by recognizing that recourse actions are exactly linear in the uncertainty for given first-stage decision variables.
- We are the first to construct scalable DRCC-OTS problems using the mean dispersion and the ∞-Wasserstein ambiguity sets, both leading to MILP reformulations through dualization-based analysis.

Numerical tests demonstrate the proposed DRCC-OTS solutions can effectively limit the constraint violations and reduce curtailment under renewable uncertainty, greatly improving the robustness guarantees over CC-OTS. Furthermore, the moment-based DRCC-OTS approach is scenario-free and efficiently solvable, hence very suitable for real-time grid operations.

3.3.1 Optimal Transmission Switching Under Uncertainty

This subsection formally presents the OTS problem under uncertainty as well as its chance-constrained solutions. We first discuss the model of uncertainty due to e.g., renewable generation or flexible demand. Let $\boldsymbol{\xi} \in \mathbb{R}^K$ stand for the uncertainty vector of the full system with its samples denoted by $\{\boldsymbol{\xi}^j\}_{j=1}^S$. We assume that $\boldsymbol{\xi}$ is bounded with a certain support set. In a data-driven setting, the support set can be estimated with high confidence from the samples under mild assumptions on the distribution (e.g., sub-Gaussian). For example, it can be the polytope formed by the convex hull of the samples [99]. The following general condition is assumed.

AS1. *The support set for* $\boldsymbol{\xi}$ *is compact and represented by a full-dimensional polytope* $\boldsymbol{\Xi} \coloneqq \{\boldsymbol{\xi} \in \mathbb{R}^K : \mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}\}$.

To incorporate the uncertainty into the OTS problem, we resort to a two-stage robust optimization by making a *here-and-now* decision while taking *recourse* or *wait-and-see* actions once the realizations of $\boldsymbol{\xi}$ are observed. Recourse functions are defined for the generation, angle, and line flow variables upon observing $\boldsymbol{\xi}$. For simplicity, we consider fast linear control for the generator recourse, similar to automatic generation control for quickly maintaining power balance [262, Ch. 9], as detailed here.

AS2. *The generation recourse actions follow a linear control mechanism that adjusts each dispatchable generator by a fixed percentage of instantaneous network-wide power imbalance. As the latter is equal to* $\mathbf{e}^\mathsf{T}\boldsymbol{\xi} = \xi_1 + \ldots + \xi_K$*, the generation adjustment becomes* $\boldsymbol{g}'(\boldsymbol{\xi}) = \boldsymbol{\gamma}(\mathbf{e}^\mathsf{T}\boldsymbol{\xi})$*, with vector* $\boldsymbol{\gamma} \in \mathbb{R}^N$ *collecting the linear coefficients to be determined.*

The flexibility of generation output is also limited by the committed reserves, and we use $\overline{\boldsymbol{r}}, \underline{\boldsymbol{r}} \in \mathbb{R}^N$ to denote the upper/lower reserve limits. In addition, changes of angles and line flows are respectively denoted by $\boldsymbol{\theta}'(\boldsymbol{\xi}) : \mathbb{R}^K \to \mathbb{R}^N$ and $\boldsymbol{f}'(\boldsymbol{\xi}) : \mathbb{R}^K \to \mathbb{R}^L$, both as recourse functions of $\boldsymbol{\xi}$. Inspired by the linearity of dc power flow, we will model them as linear functions, i.e., we have $\boldsymbol{\theta}'(\boldsymbol{\xi}) = \boldsymbol{Y}_\theta\boldsymbol{\xi}$ and $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_f\boldsymbol{\xi}$ with matrices $\boldsymbol{Y}_\theta \in \mathbb{R}^{N\times K}$ and $\boldsymbol{Y}_f \in \mathbb{R}^{L\times K}$ as decision variables. This approach is well known as the linear decision rule (LDR) scheme in two-stage robust optimization [152], which approximates

the recourse variables as affine functions of uncertainty. Interestingly, under (AS3)-(AS4) this linearized model turns out to be *exact* in representing the actual changes of angles and line flows at no modeling error, as detailed shortly. With vector $\boldsymbol{q}$ collecting the linear cost coefficients for generation adjustment, the OTS problem under uncertain $\boldsymbol{\xi}$ is cast as:

$$\min \quad \boldsymbol{c}^{\mathsf{T}}\boldsymbol{g} + \mathbb{E}[\boldsymbol{q}^{\mathsf{T}}\boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi}] \tag{3.18a}$$

$$\text{s.t.} \quad (4.1\text{b}) - (4.1\text{i}),\ \boldsymbol{\gamma} \in [0,1]^N,\ \mathbf{e}^{\mathsf{T}}\boldsymbol{\gamma} = 1 \tag{3.18b}$$

$$\boldsymbol{Y}_\theta \in \mathbb{R}^{N\times K},\ \boldsymbol{Y}_f \in \mathbb{R}^{L\times K} \tag{3.18c}$$

$$\underline{\boldsymbol{r}} \le \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \le \overline{\boldsymbol{r}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18d}$$

$$\underline{\boldsymbol{g}} \le \boldsymbol{g} + \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \le \overline{\boldsymbol{g}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18e}$$

$$\underline{\boldsymbol{\theta}} \le \boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi} \le \overline{\boldsymbol{\theta}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18f}$$

$$\underline{\boldsymbol{f}} \circ \boldsymbol{z} \le \boldsymbol{f} + \boldsymbol{Y}_f\boldsymbol{\xi} \le \overline{\boldsymbol{f}} \circ \boldsymbol{z} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18g}$$

$$\mathbf{A}(\boldsymbol{f} + \boldsymbol{Y}_f\boldsymbol{\xi}) = \boldsymbol{g} + \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} - \boldsymbol{d} - \mathbf{F}\boldsymbol{\xi} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18h}$$

$$\mathbf{K}(\boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi}) - \boldsymbol{f} - \boldsymbol{Y}_f\boldsymbol{\xi} + \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) \ge \mathbf{0} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.18i}$$

$$\mathbf{K}(\boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi}) - \boldsymbol{f} - \boldsymbol{Y}_f\boldsymbol{\xi} - \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) \le \mathbf{0} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}. \tag{3.18j}$$

In the following, we define $\mathbf{x}$ to be the vector of decision variables comprising the first-stage decisions $(\boldsymbol{g}, \boldsymbol{\theta}, \boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\gamma})$. Note that the coefficient $\boldsymbol{\gamma}$ is included for reducing the total cost. The second-stage decisions $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$ relate the angle and line flow adjustments to $\boldsymbol{\xi}$. Moreover, the transformation matrix $\mathbf{F} \in \mathbb{R}^{N\times K}$ in (3.19h) is a known mapping from $\boldsymbol{\xi} \in \mathbb{R}^K$ to the system dimension N. Basically, problem (3.19) aims to minimize the sum of total generation cost at the first-stage and the expected cost during real-time *recourse*

adjustment. Constraint (3.19d) imposes the operating reserve limits while the remaining ones (3.19e)-(3.19j) ensure that the system operating limits in the OTS problem (3.5) would still hold after recourse actions are taken [cf. (4.1c)-(4.1h)]. With the mean of uncertainty $\mathbb{E}[\boldsymbol{\xi}] = \boldsymbol{\mu}$ known, the term $\mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$ in (3.19a) simplifies to $\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\mu}$, which is a linear function of the unknown $\boldsymbol{\gamma}$.

By recognizing this linearity, we can establish the exactness of the modeling on $\boldsymbol{\theta}'(\boldsymbol{\xi}) = \boldsymbol{Y}_\theta\boldsymbol{\xi}$ and $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_f\boldsymbol{\xi}$, as follows.

Lemma 3.3.1. *Under (AS3)-(AS4), the adjustments on angle and line flow for fixed grid topology $\boldsymbol{z}$ and coefficients $\boldsymbol{\gamma}$ become exactly linear functions of $\boldsymbol{\xi}$, with both matrices $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$ uniquely determined by $\boldsymbol{z}$ and $\boldsymbol{\gamma}$.*

Proof: Under (AS4), the change of full network injection due to uncertainty is $(\boldsymbol{\gamma}\mathbf{e}^\mathsf{T} - \mathbf{F})\boldsymbol{\xi}$ [cf. (3.19h)]. Under the fixed topology of no islanding, the dc linear flow model [262, Ch. 4] states that changes of angle and line flow, namely $\boldsymbol{\theta}'$ and $\boldsymbol{f}'$, are linearly related to the change of injection, and thus to $\boldsymbol{\xi}$ as well. Lastly, the full-dimensionality of the support set $\boldsymbol{\Xi}$ in (AS3) further guarantees the uniqueness of $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$. $\square$

Lemma 3.3.2 ensures that the linear models in problem (3.19) yield the exact recourse values for the angle and line flow under any given first-stage decision variables of $\boldsymbol{z}$ and $\boldsymbol{\gamma}$. This is also the case under the optimal solutions of $\boldsymbol{z}$ and $\boldsymbol{\gamma}$. Hence, the LDR-based approximations $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_\theta\boldsymbol{\xi}$ and $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_f\boldsymbol{\xi}$ are a valid and exact model for these recourse variables, and problem (3.19) constitutes an *exact formulation* of two-stage OTS. This equivalence

differs from most LDR-based approaches, which attain only approximations [152].

This section formally presents the OTS problem under uncertainty as well as its chance-constrained solutions. We first discuss the model of uncertainty due to e.g., renewable generation or flexible demand. Let $\boldsymbol{\xi} \in \mathbb{R}^K$ stand for the uncertainty vector of the full system with its samples denoted by $\{\boldsymbol{\xi}^j\}_{j=1}^S$. We assume that $\boldsymbol{\xi}$ is bounded with a certain support set. In a data-driven setting, the set can be estimated with high confidence from the samples under mild assumptions on the distribution (e.g., sub-Gaussian). For example, it can be the polytope formed by the convex hull of the samples [99]. The following general condition is assumed.

AS3. *The support set for* $\boldsymbol{\xi}$ *is compact and represented by a full-dimensional polytope* $\boldsymbol{\Xi} := \{\boldsymbol{\xi} \in \mathbb{R}^K : \mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}\}$.

To incorporate the uncertainty into (3.5), we resort to a two-stage robust optimization by making a *here-and-now* decision while taking *recourse* or *wait-and-see* actions once the realizations of $\boldsymbol{\xi}$ are observed. Recourse functions are defined for the generation, angle, and line flow variables upon observing $\boldsymbol{\xi}$. For simplicity, we consider a linear response modeling for generator recourse actions, as motivated by frequency response and automatic generation control mechanisms [262, Ch. 9].

AS 4. *The generation recourse actions follow a linear response mechanism that adjusts each dispatchable generator by a fixed percentage of instantaneous*

network-wide power imbalance. As the latter is equal to $\mathbf{e}^\mathsf{T}\boldsymbol{\xi} = \xi_1 + \ldots + \xi_K$*, the generation adjustment becomes* $\boldsymbol{g}'(\boldsymbol{\xi}) = \boldsymbol{\gamma}(\mathbf{e}^\mathsf{T}\boldsymbol{\xi})$*, with vector* $\boldsymbol{\gamma} \in \mathbb{R}^N$ *collecting the linear coefficients to be determined.*

This linear policy has been widely adopted by various earlier work (e.g., [41,166,213]), as it can quickly restore the system-wide power imbalance. Specifically, the recourse actions are linear functions of total power mismatch, which can be quickly corrected by a proportional change from each generating unit. Such policy is very convenient to implement in practice as system-wide power mismatch is easily measured using frequency deviation. As a result, area-wide frequency responses require minimal communication overhead.

The flexibility of generation output is limited by the committed reserves, with $\overline{\boldsymbol{r}}, \underline{\boldsymbol{r}} \in \mathbb{R}^N$ denoting its upper/lower limits. Moreover, changes of angles and line flows are respectively denoted by $\boldsymbol{\theta}'(\boldsymbol{\xi}) : \mathbb{R}^K \to \mathbb{R}^N$ and $\boldsymbol{f}'(\boldsymbol{\xi}) : \mathbb{R}^K \to \mathbb{R}^L$, both as recourse functions of $\boldsymbol{\xi}$. Inspired by the linearity of dc power flow, we will model them as linear functions, i.e., we have $\boldsymbol{\theta}'(\boldsymbol{\xi}) = \boldsymbol{Y}_\theta \boldsymbol{\xi}$ and $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_f \boldsymbol{\xi}$ with matrices $\boldsymbol{Y}_\theta \in \mathbb{R}^{N \times K}$ and $\boldsymbol{Y}_f \in \mathbb{R}^{L \times K}$ as decision variables. This approach is well known as the linear decision rule (LDR) scheme in two-stage robust optimization [152], which approximates the recourse variables as affine functions of uncertainty. Interestingly, under (AS3)-(AS4) this linearized model turns out to be *exact* in representing the actual changes of angles and line flows at no modeling error, as detailed shortly. With vector $\boldsymbol{q}$ collecting the linear cost coefficients for generation adjustment, the OTS

problem under uncertain $\boldsymbol{\xi}$ is cast as:

$$\min \quad \boldsymbol{c}^{\mathsf{T}}\boldsymbol{g} + \mathbb{E}[\boldsymbol{q}^{\mathsf{T}}\boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi}] \tag{3.19a}$$

$$\text{s.t.} \quad (4.1\text{b}) - (4.1\text{i}),\ \boldsymbol{\gamma} \in [0,1]^N,\ \mathbf{e}^{\mathsf{T}}\boldsymbol{\gamma} = 1 \tag{3.19b}$$

$$\boldsymbol{Y}_\theta \in \mathbb{R}^{N\times K},\ \boldsymbol{Y}_f \in \mathbb{R}^{L\times K} \tag{3.19c}$$

$$\underline{\boldsymbol{r}} \le \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \le \overline{\boldsymbol{r}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19d}$$

$$\underline{\boldsymbol{g}} \le \boldsymbol{g} + \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \le \overline{\boldsymbol{g}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19e}$$

$$\underline{\boldsymbol{\theta}} \le \boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi} \le \overline{\boldsymbol{\theta}} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19f}$$

$$\underline{\boldsymbol{f}} \circ \boldsymbol{z} \le \boldsymbol{f} + \boldsymbol{Y}_f\boldsymbol{\xi} \le \overline{\boldsymbol{f}} \circ \boldsymbol{z} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19g}$$

$$\mathbf{A}(\boldsymbol{f} + \boldsymbol{Y}_f\boldsymbol{\xi}) = \boldsymbol{g} + \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} - \boldsymbol{d} - \mathbf{F}\boldsymbol{\xi} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19h}$$

$$\mathbf{K}(\boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi}) - \boldsymbol{f} - \boldsymbol{Y}_f\boldsymbol{\xi} + \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) \ge \mathbf{0} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} \tag{3.19i}$$

$$\mathbf{K}(\boldsymbol{\theta} + \boldsymbol{Y}_\theta\boldsymbol{\xi}) - \boldsymbol{f} - \boldsymbol{Y}_f\boldsymbol{\xi} - \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) \le \mathbf{0} \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}. \tag{3.19j}$$

In the following, we define $\mathbf{x}$ to be the vector of decision variables comprising the first-stage decisions $(\boldsymbol{g}, \boldsymbol{\theta}, \boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\gamma})$. Note that the coefficient $\boldsymbol{\gamma}$ is included for reducing the total cost. The second-stage decisions $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$ relate the angle and line flow adjustments to $\boldsymbol{\xi}$. Moreover, the transformation matrix $\mathbf{F} \in \mathbb{R}^{N\times K}$ in (3.19h) is a known mapping from $\boldsymbol{\xi} \in \mathbb{R}^K$ to the system dimension N. Basically, problem (3.19) aims to minimize the sum of total generation cost at the first-stage and the expected cost during real-time *recourse* adjustment. Constraint (3.19d) imposes the operating reserve limits while the remaining ones (3.19e)-(3.19j) ensure that the system operating limits in the OTS problem (3.5) would still hold after recourse actions are taken [cf. (4.1c)-

(4.1h)]. With the mean of uncertainty $\mathbb{E}[\boldsymbol{\xi}] = \boldsymbol{\mu}$ known, the term $\mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$ in (3.19a) simplifies to $\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\mu}$, which is a linear function of the unknown $\boldsymbol{\gamma}$.

By recognizing this linearity, we can establish the exactness of the modeling on $\boldsymbol{\theta}'(\boldsymbol{\xi}) = \boldsymbol{Y}_\theta\boldsymbol{\xi}$ and $\boldsymbol{f}'(\boldsymbol{\xi}) = \boldsymbol{Y}_f\boldsymbol{\xi}$, as follows.

Lemma 3.3.2. *Under (AS3)-(AS4), the adjustments on angle and line flow for fixed grid topology $\boldsymbol{z}$ and coefficients $\boldsymbol{\gamma}$ become exactly linear functions of $\boldsymbol{\xi}$, with both matrices $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$ uniquely determined by $\boldsymbol{z}$ and $\boldsymbol{\gamma}$.*

Proof: Under (AS4), the change of full network injection due to uncertainty is $(\boldsymbol{\gamma}\mathbf{e}^\mathsf{T} - \mathbf{F})\boldsymbol{\xi}$ [cf. (3.19h)]. Under the fixed topology of no islanding, the dc linear flow model [262, Ch. 4] states that changes of angle and line flow, namely $\boldsymbol{\theta}'$ and $\boldsymbol{f}'$, are linearly related to the change of injection, and thus to $\boldsymbol{\xi}$ as well. Lastly, the full-dimensionality of the support set $\boldsymbol{\Xi}$ in (AS3) further guarantees the uniqueness of $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_f$. $\square$

Lemma 3.3.2 ensures the linear models in problem (3.19) produce the exact recourse values for angle and line flow under given first-stage decision variables of $\boldsymbol{z}$ and $\boldsymbol{\gamma}$. Hence, our LDR approach yields an exact model for the recourse variables, and problem (3.19) constitutes an *equivalent* two-stage OTS formulation. This is a much stronger result than existing LDR solutions [152], including the earlier OTS application in [189].

3.3.1.1 Chance-Constrained (CC-) OTS

The chance-constrained (CC) formulation is popularly employed to deal with inequality constraints under uncertainty [159]. It ensures that constraints are satisfied with probability above a prescribed threshold. The relevant constraints from problem (3.19) can be collected in the following set:

$$\begin{aligned}\mathcal{I} = \Big\{ & \underline{\boldsymbol{r}} \leq \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \leq \overline{\boldsymbol{r}}, \\ & \underline{\boldsymbol{g}} \leq \boldsymbol{g} + \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi} \leq \overline{\boldsymbol{g}}, \\ & \underline{\boldsymbol{\theta}} \leq \boldsymbol{\theta} + \boldsymbol{Y}_{\theta}\boldsymbol{\xi} \leq \overline{\boldsymbol{\theta}}, \\ & \underline{\boldsymbol{f}} \circ \boldsymbol{z} \leq \boldsymbol{f} + \boldsymbol{Y}_{f}\boldsymbol{\xi} \leq \overline{\boldsymbol{f}} \circ \boldsymbol{z} \Big\}. \end{aligned} \tag{3.20}$$

These constraints correspond to the limits on reserve, generation, phase angle, and line flow as in (3.19d) - (3.19g), all of which are linear in $\boldsymbol{\xi}$. Note that the network power balance in (3.19h) and line flow relations in (3.19i) - (3.19j) are not part of the set (3.20). This is because they are used to determine the power flow and thus need to be satisfied strictly. Interestingly, they can be effectively reformulated by linear constraints without $\boldsymbol{\xi}$. For the semi-infinite equality constraint (3.19h), it reduces to a finite linear one, as stated in the following proposition.

Proposition 3.3.3. *Under (AS3), constraint* (3.19h) *is equivalent to:*

$$\mathbf{A}\boldsymbol{Y}_f = \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}} - \mathbf{F}. \tag{3.21}$$

Proof: Recalling $\mathbf{A}\boldsymbol{f} = \boldsymbol{g} - \boldsymbol{d}$ [cf. (4.1f)], we can rewrite (3.19h) as $(\mathbf{A}\boldsymbol{Y}_f - \boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}} + \mathbf{F})\boldsymbol{\xi} = \mathbf{0},\ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}$. This implies that the linear hull of $\boldsymbol{\Xi}$ should belong

to the null space of the linear operator $(\mathbf{A}\boldsymbol{Y}_f - \boldsymbol{\gamma}\mathbf{e}^\mathsf{T} + \mathbf{F})$. As $\boldsymbol{\Xi}$ spans the whole sample space under (AS3), the associated null space is empty and (3.21) holds accordingly. $\square$

For the inequality constraints (3.19i) and (3.19j), a well-known equivalence result in robust optimization [152] leads to a tractable constraint system, as described in the following proposition.

Proposition 3.3.4. *Under (AS3), the constraints* (3.19i) *and* (3.19j) *are respectively equivalent to:*

$$\boldsymbol{Y}_f = \mathbf{K}\boldsymbol{Y}_\theta + \boldsymbol{\Phi}_1^\mathsf{T}\mathbf{U},\ \mathbf{K}\boldsymbol{\theta} + \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) - \boldsymbol{\Phi}_1^\mathsf{T}\mathbf{t} \geq \boldsymbol{f} \tag{3.22a}$$

$$\boldsymbol{Y}_f = \mathbf{K}\boldsymbol{Y}_\theta - \boldsymbol{\Phi}_2^\mathsf{T}\mathbf{U},\ \mathbf{M} \circ (\mathbf{e} - \boldsymbol{z}) - \boldsymbol{\Phi}_2^\mathsf{T}\mathbf{t} + \boldsymbol{f} \geq \mathbf{K}\boldsymbol{\theta} \tag{3.22b}$$

where the matrices $\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2 \in \mathbb{R}_+^{L \times W}$ *collect the dual variables for constraints in* (3.19i) *and* (3.19j)*, respectively.*

Proof: This proposition can be viewed as a special case of [29, Thm. 3.2]. For any constraint of the form $\boldsymbol{h}^\mathsf{T}\boldsymbol{\xi} + m \geq 0, \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}$, under (AS3) it is equivalent to $0 \leq \min_{\boldsymbol{\xi}}\{\boldsymbol{h}^\mathsf{T}\boldsymbol{\xi} + m : \mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}\}$. The right-hand side expression is essentially a linear program, for which the equivalent dual problem under Slater's conditions becomes $0 \leq \max_{\varphi}\{-\mathbf{t}^\mathsf{T}\boldsymbol{\varphi} + m : \boldsymbol{\varphi} \geq \mathbf{0}, -\mathbf{U}^\mathsf{T}\boldsymbol{\varphi} = \boldsymbol{h}\}$, where $\boldsymbol{\varphi}$ is the vector of dual variables. For the maximum of the dual problem to be non-negative, the dual vector $\boldsymbol{\varphi} \geq \mathbf{0}$ has to satisfy $\mathbf{U}^\mathsf{T}\boldsymbol{\varphi} = -\boldsymbol{h}$ and $\mathbf{t}^\mathsf{T}\boldsymbol{\varphi} \leq m$. Thus, constraints (3.19i) and (3.19j) are rewritten into (3.22) using this equivalence. $\square$

Using Propositions 3.3.3 and 3.3.4, we can convert the remaining constraints in (3.19) to deterministic ones without $\boldsymbol{\xi}$ and accordingly formulate the CC-OTS problem as, follows:

$$\min \quad \boldsymbol{c}^{\mathsf{T}}\boldsymbol{g} + \mathbb{E}[\boldsymbol{q}^{\mathsf{T}}\boldsymbol{\gamma}\mathbf{e}^{\mathsf{T}}\boldsymbol{\xi}] \tag{3.23a}$$

$$\text{s.t.} \quad (3.19\text{b}), (3.19\text{c}), (3.21), (3.22\text{a}), (3.22\text{b}) \tag{3.23b}$$

$$\mathbb{P}\{\mathbf{a}_i(\mathbf{x})^{\mathsf{T}}\boldsymbol{\xi} \leq b_i(\mathbf{x})\} \geq 1 - \epsilon_i, \forall i \in \mathcal{I}. \tag{3.23c}$$

Here, the chance constraints (3.23c) guarantee that each inequality in (3.20) holds with a probability of at least $1 - \epsilon_i$, for a pre-specified tolerance level ϵ_i.

3.3.1.2 Benchmark Methods for CC-OTS

We present two benchmark methods for approximating the chance constraints (3.23c), which is the most critical step in solving (3.23). These two approximation methods give rise to mixed-integer problems and will be used to numerically compare with the proposed DRCC methods later on.

Sample-Average Approximation (SAA) Given independently and identically distributed (i.i.d.) uncertainty samples $\{\boldsymbol{\xi}^j\}$ with $j \in \mathcal{J} := \{1, \cdots, S\}$, the SAA approach [172] replaces the CC constraints (3.23c) with the sample-based empirical distribution $\hat{\mathbb{P}}$ that assigns equal mass to all samples.

Under the empirical distribution, the chance constraint (3.23c) is equivalent to the system of mixed-integer linear constraints:

$$w_i^j \in \{0, 1\}, \quad \forall i \in \mathcal{I},\, j \in \mathcal{J}. \tag{3.24a}$$

$$b_i(\mathbf{x}) - \mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi}^j + \mathrm{M}w_i^j \geq 0, \quad \forall i,\, j. \tag{3.24b}$$

$$\textstyle\sum_{j=1}^{S} w_i^j \leq S\epsilon_i, \quad \forall i. \tag{3.24c}$$

with a sufficiently large M. The binary decision variable w_i^j indicates whether the constraint $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi}^j \leq b_i(\mathbf{x})$ holds or not. If $w_i^j = 0$, (3.24b) is equivalent to $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi}^j \leq b_i(\mathbf{x})$ and the constraint holds; otherwise, (3.24b) becomes redundant for a large enough M. If each $\boldsymbol{\xi}^j$ is randomly sampled with an equal probability (i.e., $\mathbb{P}(\boldsymbol{\xi}^j) = 1/S$), constraint (3.24c) guarantees that the sample-based probability of violation ($\frac{1}{S}\sum_{j=1}^{S} w_i^j$) is not greater than the threshold ϵ_i. Under SAA, the resulting approximation of (3.23) is an MILP. Notice, however, that the number of constraints can grow quickly with the sample size S. Due to this scalability issue, the SAA approach will be mainly used for the small test case in numerical studies.

Gaussian Approximation This method assumes that $\boldsymbol{\xi}$ is a Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. For the chance constraint (3.23c) with a typical threshold of $\epsilon_i \leq \frac{1}{2}$, the Gaussian distribution leads to an equivalent second-order cone (SOC) constraint [46, Sec. 4.4]. To briefly introduce the basic idea of this method, we consider the variance of $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi}$ as denoted by σ^2 and constraint $i \in \mathcal{I}$ in (3.23c) now becomes

$$\mathbb{P}\left(\frac{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} - \mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\mu}}{\sigma} \leq \frac{b_i(\mathbf{x}) - \mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\mu}}{\sigma}\right) \geq 1 - \epsilon_i. \tag{3.25}$$

With $\Phi^{-1}(\epsilon_i)$ denoting ϵ_i-quantile of the standard normal distribution, (3.25) is equivalent to the following SOC constraint:

$$\boldsymbol{\mu}^{\mathsf{T}}\mathbf{a}_i(\mathbf{x}) + \Phi^{-1}(1-\epsilon_i)\left\|\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{a}_i(\mathbf{x})\right\|_2 \leq b_i(\mathbf{x}) \tag{3.26}$$

Under Gaussian approximation, the CC-OTS problem (3.23) becomes a mixed-integer SOCP (MISOCP). This method requires no sampling, yet its uncertainty model can be too restrictive for the renewable perturbations in practice.

3.3.1.3 Quantifying the Level of Renewable Curtailment

In practice, curtailment of renewable generation is used to avoid oversupply and to maintain constraint satisfaction [217]. Upon solving any CC-OTS problem with the optimal topology $\boldsymbol{z}^*$ and other values (denoted by *), one can apply the Monte Carlo method using a large number of uncertainty scenarios to obtain the average of resultant curtailment values. Specifically, for a given renewable scenario $\tilde{\boldsymbol{\xi}}$, we determine the renewable curtailment vector $\boldsymbol{\xi}_c \in \mathbb{R}_+^K$ in order to satisfy all network constraints, as given by

$$\begin{aligned}
\min \quad & \mathbf{1}^{\mathsf{T}}\boldsymbol{\xi}_c && \text{(3.27a)}\\
\text{s.t.} \quad & \boldsymbol{\xi}_c \in \mathbb{R}_+^K && \text{(3.27b)}\\
& \underline{\boldsymbol{\theta}} \leq \boldsymbol{\theta}^* + \mathbf{Y}_\theta^*(\tilde{\boldsymbol{\xi}} - \boldsymbol{\xi}_c) \leq \overline{\boldsymbol{\theta}} && \text{(3.27c)}\\
& \underline{\boldsymbol{f}} \circ \boldsymbol{z}^* \leq \boldsymbol{f}^* + \mathbf{Y}_f^*(\tilde{\boldsymbol{\xi}} - \boldsymbol{\xi}_c) \leq \overline{\boldsymbol{f}} \circ \boldsymbol{z}^*, && \text{(3.27d)}
\end{aligned}$$

where the network constraints in (3.20) have been simplified to the linear ones in (3.27) by fixing $\boldsymbol{z}^*$. Note the the curtailment criterion in (3.27a) is

essentially the L_1 norm of $\boldsymbol{\xi}_c$, which gives rise to an efficient linear program (LP) in (3.27). Other criteria such as L_2 norm can be used as well, at possibly increased computation complexity. By determining $\boldsymbol{\xi}_c$ using (3.27), the process for quantifying renewable curtailment boils down to computing the average of $(\mathbf{1}^\top \boldsymbol{\xi}_c)$ over a large number of renewable scenarios. This process serves to evaluate the impact of OTS solutions in terms of renewable curtailment level, which also applies to the DRCC-OTS solutions to be discussed soon. Note this evaluation is completed offline and does not affect the real-time computation of any OTS solution. As more frequent constraint violations naturally lead to higher renewable curtailment, the objective cost attained by (3.27) serves as an important criterion to evaluate the robustness performance of CC-OTS solutions, as shown by the numerical results in Section 4.4.

3.3.2 Distributionally Robust Chance-Constrained OTS

The distributionally robust optimization (DRO) framework has been recognized as a powerful yet potentially tractable approach to deal with uncertainty in energy systems [20, 157, 169, 259, 270, 288]. The DRO framework does not assume a particular probability distribution. Instead, it constructs an *ambiguity set* of plausible distributions that are consistent with the available statistical and structural information on uncertainty. A safe decision is then sought that is feasible to the chance constraints for all distributions within the ambiguity set. Hence, the framework mitigates data overfitting issues and yields superior performance in out-of-sample (OOS) tests.

To develop the DRO-based OTS formulation, consider the distributionally robust chance constraints (DRCC) for (3.20) as

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \leq b_i(\mathbf{x})\} \geq 1-\epsilon_i \ \ \forall i \in \mathcal{I}, \tag{3.28}$$

which require each chance constraint to be satisfied under all probability distributions $\mathbb{P} \in \mathcal{P}$. Typical ambiguity sets studied in related DRO-based power system decision-making problems fall into the following categories: i) moment-based ambiguity set [121, 281, 287], ii) distance-based ambiguity set [83, 115, 202] and iii) structural-based ambiguity set [156, 216]. We consider the DRCC reformulations using the moment-based ambiguity set (mean and mean absolute deviation) and the distance-based ambiguity set (Wasserstein distance). Both of them are amenable to mixed-integer linear programming reformulations.

3.3.2.1 Mean and Mean Absolute Deviation Ambiguity Set

The mean and mean absolute deviation (mean-MAD) ambiguity set [118] is defined as:

$$\mathcal{P}_1 := \{\mathbb{P} \in \mathcal{P}_0(\boldsymbol{\Xi}) : \mathbb{E}[\boldsymbol{\xi}] = \boldsymbol{\mu},\ \mathbb{E}[|\boldsymbol{\xi}-\boldsymbol{\mu}|] \leq \boldsymbol{\sigma}\}, \tag{3.29}$$

which includes all distributions with the mean equal to $\boldsymbol{\mu} \in \mathbb{R}^K$ and the mean absolute deviation bounded by $\boldsymbol{\sigma} \in \mathbb{R}^K_+$. Note that the absolute value and its inequality are both component-wise. This ambiguity set can be extended to impose certain dependence structures (see e.g., [261, Sec. 5]). Each worst-case

probability $\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \leq b_i(\mathbf{x})\}$ in (3.28) over the ambiguity set $\mathcal{P} = \mathcal{P}_1$ boils down to the following optimization problem:

$$Z_{\mathcal{P}_1} = \inf \quad \int \mathbb{1}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \leq b_i(\mathbf{x})\} v(\mathrm{d}\boldsymbol{\xi}) \tag{3.30a}$$

$$\text{s.t.} \quad v(\cdot) \in \mathcal{M}_+, \tag{3.30b}$$

$$\int v(\mathrm{d}\boldsymbol{\xi}) = 1, \tag{3.30c}$$

$$\int \boldsymbol{\xi} v(\mathrm{d}\boldsymbol{\xi}) = \boldsymbol{\mu}, \tag{3.30d}$$

$$\int |\boldsymbol{\xi} - \boldsymbol{\mu}| v(\mathrm{d}\boldsymbol{\xi}) \leq \boldsymbol{\sigma}, \tag{3.30e}$$

where $\mathbb{1}(\cdot)$ denotes the indicator function for the inequality constraint, while $\mathcal{M}_+$ defines the set of nonnegative measures. Constraints (3.35c) - (3.35e) are essentially the integral forms of (3.29). As the objective and constraint functions are all linear in the unknown measure $v(\cdot)$, the problem (3.30) is a convex semi-infinite linear program (SILP). If the DRCC (3.28) is feasible under $\mathcal{P}_1$, then we have $Z_{\mathcal{P}_1} \geq 1 - \epsilon_i$ for constraint $i \in \mathcal{I}$. By denoting $\alpha \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^K$, and $\boldsymbol{\kappa} \in \mathbb{R}^K_+$ as the dual variables of constraints (3.35c)-(3.35e), respectively, we can formulate the dual problem of (3.30) as:

$$\sup \quad \alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\mu} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\sigma} \tag{3.31a}$$

$$\text{s.t.} \quad \alpha \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^K,\ \boldsymbol{\kappa} \in \mathbb{R}^K_+, \tag{3.31b}$$

$$\mathbb{1}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \leq b_i(\mathbf{x})\} \geq \alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\xi} - \boldsymbol{\kappa}^\mathsf{T}|\boldsymbol{\xi} - \boldsymbol{\mu}| \ \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}. \tag{3.31c}$$

Strong duality holds as the ambiguity set $\mathcal{P}_1$ satisfies the Slater's condition [119] for the SILP (3.30). The semi-infinite constraint for the dual problem

(3.31) boils down to two cases according to the indicator $\mathbb{1}(\cdot)$ in (3.31c). Specifically, it equals to 0 for any $\boldsymbol{\xi}$ such that $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} > b_i(\mathbf{x})$, or 1 for any other choice of $\boldsymbol{\xi}$. These two cases can be reformulated using standard convex duality theory to arrive at the following equivalent linear constraints:

$$\alpha' + \boldsymbol{\beta}'^\mathsf{T}\boldsymbol{\mu} - \boldsymbol{\kappa}'^\mathsf{T}\boldsymbol{\sigma} \geq (1-\epsilon_i)\lambda' \tag{3.32a}$$

$$\alpha' + (\boldsymbol{\pi}'^\mathsf{T}_1 - \boldsymbol{\tau}'^\mathsf{T}_1)\boldsymbol{\mu} + \boldsymbol{\psi}'^\mathsf{T}_1\mathbf{t} \leq \lambda' \tag{3.32b}$$

$$\boldsymbol{\beta}'^\mathsf{T} + \boldsymbol{\tau}'^\mathsf{T}_1 = \boldsymbol{\pi}'^\mathsf{T}_1 + \boldsymbol{\psi}'^\mathsf{T}_1\mathbf{U} \tag{3.32c}$$

$$\boldsymbol{\pi}'^\mathsf{T}_1 + \boldsymbol{\tau}'^\mathsf{T}_1 = \boldsymbol{\kappa}'^\mathsf{T} \tag{3.32d}$$

$$\alpha' + (\boldsymbol{\pi}'^\mathsf{T}_2 - \boldsymbol{\tau}'^\mathsf{T}_2)\boldsymbol{\mu} + \boldsymbol{\psi}'^\mathsf{T}_2\mathbf{t} \leq b_i(\mathbf{x}) \tag{3.32e}$$

$$\boldsymbol{\beta}'^\mathsf{T} + \mathbf{a}_i(\mathbf{x})^\mathsf{T} + \boldsymbol{\tau}'^\mathsf{T}_2 = \boldsymbol{\pi}'^\mathsf{T}_2 + \boldsymbol{\psi}'^\mathsf{T}_2\mathbf{U} \tag{3.32f}$$

$$\boldsymbol{\pi}'^\mathsf{T}_2 + \boldsymbol{\tau}'^\mathsf{T}_2 = \boldsymbol{\kappa}'^\mathsf{T}. \tag{3.32g}$$

The dual variables $\boldsymbol{\pi}' \in \mathbb{R}^K_+$ and $\boldsymbol{\tau}' \in \mathbb{R}^K_+$ are introduced for the epigraph based constraints $\boldsymbol{\rho} \geq \boldsymbol{\xi} - \boldsymbol{\mu}$ and $\boldsymbol{\rho} \leq \boldsymbol{\xi} - \boldsymbol{\mu}$, respectively; the dual variables $\boldsymbol{\psi}' \in \mathbb{R}^W_+$ are assigned to the linear constraints of the support set $\boldsymbol{\Xi}$. Note that the dual variable $\lambda > 0$ corresponding to the new constraint $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} > b_i(\mathbf{x})$ introduces bilinearity in the original decision variables $\mathbf{x}$. We address this by dividing all constraints with λ and performing the change of variables for the primal variables in (3.31) as $\alpha' = \frac{\alpha}{\lambda} \in \mathbb{R}$ (similarly for $\boldsymbol{\beta}'$ and $\boldsymbol{\kappa}'$), and the dual variables as $\lambda' = \frac{1}{\lambda} \in \mathbb{R}_+$ (similarly for the aforementioned $\boldsymbol{\pi}'$ and $\boldsymbol{\tau}'$).

Proposition 3.3.5. *The DRCC-OTS problem under ambiguity set $\mathcal{P}_1$ is equivalent to the following optimization problem:*

$$\min \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{g} + \mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}] \tag{3.33a}$$

$$s.t. \quad \alpha' \in \mathbb{R},\ \boldsymbol{\beta}' \in \mathbb{R}^K,\ \boldsymbol{\kappa}' \in \mathbb{R}^K_+ \tag{3.33b}$$

$$\lambda' \in \mathbb{R}_+,\ \boldsymbol{\pi}' \in \mathbb{R}^K_+,\ \boldsymbol{\tau}' \in \mathbb{R}^K_+,\ \boldsymbol{\psi}' \in \mathbb{R}^W_+ \tag{3.33c}$$

$$(3.23b),\ (3.32a) - (3.32g). \tag{3.33d}$$

Thanks to all the linear constraints, the DRCC-OTS problem in (3.33) is an MILP. The DRCC-OTS significantly improves the scalability over the SAA-based MILP problem, as it effectively uses dualization techniques to attain a fixed set of linear constraints such that (3.28) holds for any distribution in $\mathcal{P}_1$. Therefore, the resulting problem (3.33) is *scenario-free* and of low computational complexity for efficient implementations in real time.

Incorporating Multimodality Information: To obtain less conservative solutions to the aforementioned DRCC model, one can further incorporate multimodality information of the uncertainty into the formulation. This additional structural information is particularly relevant to the problem studied in the paper, as it has been observed that wind energy data exhibits multimodal behavior [128, 257, 277]. To this end, we assume the actual distribution to be a mixture of m distinct distributions $\mathbb{P}_1, \cdots, \mathbb{P}_m$, with known probabilities $p_1, \cdots, p_m$, and each $\mathbb{P}_j$ has known mean and MAD values $(\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j)$. In this setting, the ambiguity set with multimodality information is given by

$$\mathcal{P}_1' := \sum_{j=1}^{m} p_j \mathcal{P}_1(\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j), \tag{3.34}$$

where $\mathcal{P}_1(\boldsymbol{\mu}_j, \boldsymbol{\sigma}_j)$ denotes the mean-MAD ambiguity set (3.29) with mean $\boldsymbol{\mu}_j$ and MAD $\boldsymbol{\sigma}_j$. Each worst-case probability $\inf_{\mathbb{P}\in\mathcal{P}_1'} \mathbb{P}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \le b_i(\mathbf{x})\}$ can be

cast as the following problem:

$$Z_{\mathbb{P}_1'} = \inf \quad \sum_{j=1}^{m} p_j \int \mathbb{1}\{\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} \leq b_i(\mathbf{x})\} v_j(\mathrm{d}\boldsymbol{\xi}) \tag{3.35a}$$

$$\text{s.t.} \quad v_j(\cdot) \in \mathcal{M}_+, \quad \forall j = 1, \cdots, m, \tag{3.35b}$$

$$\int v_j(\mathrm{d}\boldsymbol{\xi}) = 1, \quad \forall j = 1, \cdots, m, \tag{3.35c}$$

$$\int \boldsymbol{\xi} v_j(\mathrm{d}\boldsymbol{\xi}) = \boldsymbol{\mu}_j, \quad \forall j = 1, \cdots, m, \tag{3.35d}$$

$$\int |\boldsymbol{\xi} - \boldsymbol{\mu}_j| v_j(\mathrm{d}\boldsymbol{\xi}) \leq \boldsymbol{\sigma}_j, \quad \forall j = 1, \cdots, m. \tag{3.35e}$$

However, applying similar derivations as in (3.30)-(3.32) to the multimodal model leads to a non-convex problem, as there will be multiple bilinear products $\lambda_j \mathbf{a}_i(\mathbf{x})$ and $\lambda_j b_i(\mathbf{x}), \forall j = 1, \cdots, m$ for the different modes that cannot be handled simultaneously. To deal with this bilinearity, we propose to adopt the block coordinate descent (BCD) algorithm in [118, Sec. 5] for a tractable solution, as described in Algorithm 2.

For ease of exposition, we denote all the dual variables that are not directly coupled with $\mathbf{x}$ as $\Upsilon_j = \{\alpha_j, \boldsymbol{\beta}_j, \boldsymbol{\kappa}_j, \boldsymbol{\pi}_j, \boldsymbol{\tau}_j, \boldsymbol{\psi}_j\}, \forall j = 1, \cdots, m$. The BCD algorithm starts with an initial solution $\mathbf{x}^0$, which can be the optimal solution from the unimodality model (3.33). Per iteration t, an uncertainty quantification problem is solved to find the worst-case probability under the previous iterate $\mathbf{x}^{t-1}$. Once the dual multipliers λ_j, Υ_j are obtained, we fix λ_j and solve the DRCC-OTS to update the iterate $\mathbf{x}^t$. This iterative approach is repeated until the difference between consecutive objectives is below a prescribed convergence threshold ω. Note that the optimization problems solved

Algorithm 1 Block Coordinate Descent Algorithm

Input: $\mathbf{x}^0, p_j, \boldsymbol{\Xi}_j, \boldsymbol{\mu}_j, \boldsymbol{\sigma}_j, \forall j = 1, \cdots, m$
Output: $\mathbf{x}$

Initialization: Initial feasible solution $\mathbf{x}^0$
Get objective value f^0 using $\mathbf{x}^0$, set $t = 1$.
for $t = 1$ to $t_{\max}$ **do**
  Uncertainty Quantification: Find the optimal $\left(\Upsilon_j^*, \lambda_j^*\right)$ to the dual of (3.35) with input $\mathbf{x}^{t-1}$. Set $\lambda_j^t \leftarrow \lambda_j^*$.
  Policy Update: Fix λ_j^t, solve DRCC-OTS and obtain the optimal $(\mathbf{x}^*, \Upsilon_j^*)$. Set $\mathbf{x}^t \leftarrow \mathbf{x}^*$ and compute the objective value f^t.
  if $(|f^t - f^{t-1}| < \omega)$ **then**
    Stop the algorithm, and set $\mathbf{x} = \mathbf{x}^t$
  end if
end for
return $\mathbf{x}$

in each iteration are convex and can be solved efficiently. We will investigate this multimodality model in numerical tests, as well.

3.3.2.2 Wasserstein Ambiguity Set

The DRCC with the ambiguity set described in this section ensures the robustness against all probability distributions within a prescribed Wasserstein distance from the empirical distribution $\hat{\mathbb{P}}$. Compared with the mean-MAD criterion, the Wasserstein metric is purely data-driven and constructed using actual data samples. With more samples available, the latter better reveals the actual uncertainty distribution and thus can lead to less conservative DRCC solutions.

We adopt the ∞-Wasserstein ambiguity set which is known to enjoy a more tractable reformulation [267, 272]. The ∞-Wasserstein ambiguity set is defined as

$$\mathcal{P}_2 := \left\{ \mathbb{P} \in \mathcal{P}_0(\boldsymbol{\Xi}) : d_\infty(\mathbb{P}, \hat{\mathbb{P}}) \leq \delta \right\}, \tag{3.36}$$

where $\delta > 0$ is a given Wasserstein radius that determines the finite-sample performance guarantee of the DRCC problem; see e.g., [35]. The radius parameter δ depends on the number of sample S in a monotonically decreasing fashion. One choice of setting δ is $\delta = \eta S^{-\frac{1}{k \max\{K,2\}}}$ [35, Cor. 1], where K is the dimension of the uncertainty while k and η are problem-dependent constants. The ∞-Wasserstein distance between two distributions $\mathbb{P}_1$ and $\mathbb{P}_2$ is given by

$$\begin{aligned} d_\infty(\mathbb{P}_1, \mathbb{P}_2) := \inf \quad & \operatorname{ess\,sup} \|\tilde{\boldsymbol{\xi}}_1 - \tilde{\boldsymbol{\xi}}_2\| \\ \text{s.t.} \quad & \mathbb{P} \in \mathcal{P}_0(\mathbb{R}^K \times \mathbb{R}^K), \end{aligned} \tag{3.37}$$

where $\mathbb{P}$ is the joint distribution of $\tilde{\boldsymbol{\xi}}_1$ and $\tilde{\boldsymbol{\xi}}_2$ with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$, respectively. We use ess sup to denote the essential supremum of a function and $\|\cdot\|$ to denote a norm in $\mathbb{R}^K$. For each constraint $i \in \mathcal{I}$, suppose a big-M coefficient M_j exists to bound

$$\mathrm{M}_j \geq \max_{\mathbf{x}} \left\{ \mathbf{a}_i(\mathbf{x})^\mathsf{T} \boldsymbol{\xi}^j + \delta \|\mathbf{a}_i(\mathbf{x})\|_* - b_i(\mathbf{x}) \right\}, \ \forall j \in \mathcal{J}$$

where $\|\cdot\|_*$ is the corresponding dual norm. This way, the DRCC in (3.28) under $\mathcal{P}_2$ can be represented as the following mixed-integer constraints [267, Cor. 4]:

$$(3.24a), (3.24c) \tag{3.38a}$$

$$\delta\|\mathbf{a}_i(\mathbf{x})\|_* \leq b_i(\mathbf{x}) - \mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi}^j + \mathrm{M}_j w_i^j, \quad \forall i,\ j. \tag{3.38b}$$

Note that this reformulation mimics the SAA-based one in (3.24), by changing the lower bound of the right-hand side (RHS) of (3.38b) from 0 to $\delta\|\mathbf{a}_i(\mathbf{x})\|_*$, which acts as a regularizer. Intuitively, a smaller radius δ implies the restriction to distributions more similar to the empirical one $\hat{\mathbb{P}}$. Accordingly, the constraint (3.38b) becomes less restrictive. As δ decreases to 0, $\mathcal{P}_2$ reduces to the singleton $\hat{\mathbb{P}}$ itself and (3.38) becomes equivalent to the SAA approach.

Proposition 3.3.6. *The DRCC-OTS problem under ambiguity set $\mathcal{P}_2$ is equivalent to the following optimization problem:*

$$\min \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{g} + \mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}] \tag{3.39a}$$

$$s.t. \quad \text{(3.23b), (3.38a), (3.38b).} \tag{3.39b}$$

For better numerical tractability, we have picked the L_∞-norm as the ground metric for Wasserstein distance in (3.37), for which the dual norm is L_1 in (3.38b) and the problem (3.39) becomes an MILP. Due to the similarity to SAA, the DRCC-OTS under the Wasserstein metric also incurs the same complexity issue as the number of constraints grows with sample size $|\mathcal{J}|$. Nonetheless, the choice of ∞-Wasserstein ambiguity set already improves the tractability over the traditional Wasserstein metric as in [83,115]. Compared to the mean-MAD ambiguity set, the DRCC-OTS problem under the Wasserstein ambiguity set takes more computation time especially for large systems, but its data-driven feature makes the resulting solutions less conservative with sufficient number of data samples.

Table 3.2: Performance of Approaches A1-A4 in the IEEE 14-bus System

		A1 (CC-SAA)		A2 (CC-Gaussian)		A3 (DRCC-MAD)		A4 (DRCC-Wasserstein)	
	$1-\epsilon$	0.95	0.90	0.95	0.90	0.95	0.90	0.95	0.90
Switching Decision	$L_o=1$	[16]	[16]	[16]	[16]	[16]	[16]	[16]	[16]
	$L_o=2$	[9;18]	[9;18]	[9;18]	[9;20]	[9;20]	[9;20]	[9;20]	[9;20]
	$L_o=3$	[9;18;19]	[9;18;19]	[9;19;20]	[9;12;20]	[9;12;18]	[9;12;18]	[9;12;18]	[9;12;18]
Run Time (sec)	$L_o=1$	3.07	2.93	0.51	0.24	0.33	0.38	6.87	5.67
	$L_o=2$	1.73	2.29	0.25	0.43	0.30	0.38	11.52	9.05
	$L_o=3$	1.54	1.70	0.25	0.21	0.40	0.29	6.79	8.08
Out-of-sample Costs ($/$h$)	$L_o=1$	545.34	542.10	546.01	544.48	556.02	552.90	549.95	547.02
	$L_o=2$	518.95	516.57	519.76	518.36	529.91	526.28	525.36	524.76
	$L_o=3$	515.89	512.50	518.70	517.49	524.77	521.37	518.55	517.81
Average Violation Rates	$L_o=1$	0.0616	0.0896	0.0402	0.0412	0.0220	0.0232	0.0258	0.0278
	$L_o=2$	0.1054	0.1058	0.0472	0.0478	0.0238	0.0264	0.0294	0.0298
	$L_o=3$	0.1070	0.1184	0.0648	0.0676	0.0292	0.0304	0.0358	0.0364
Renewable Curtailment	$L_o=1$	2.4957	2.5637	2.4327	2.4381	2.2253	2.3218	2.0611	2.1042
	$L_o=2$	2.6626	2.7561	2.4350	2.5124	2.2671	2.3062	1.9863	2.3505
	$L_o=3$	2.5136	2.5815	2.4156	2.4268	2.0670	2.1367	2.0272	2.2341

Remark 3.3.1 (distributionally robust objective). We can also extend the DRCC formulations to include a distributionally robust objective function. To achieve this, the term $\mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$ in the objective functions of (3.33) and (3.39) can be changed to $\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}[\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$. The latter is equivalent to $\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}[\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$, in which $\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}[\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$ can be determined for a given ambiguity set $\mathcal{P}$ similar to the steps for analyzing the constraints. Note that this change only affects the coefficient for scaling the term $\boldsymbol{q}^\mathsf{T}\boldsymbol{\gamma}$. Thus, for simplicity, this work did not incorporate a DR cost objective, as $\mathbb{E}[\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$ can be viewed as a lower bound for $\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}[\mathbf{e}^\mathsf{T}\boldsymbol{\xi}]$.

3.3.3 Numerical Results

In this section, we present the numerical results validating the proposed DRCC-OTS methods using the IEEE 14-bus and 118-bus test cases. Other benchmark approaches are implemented too for performance compar-

isons in terms of robustness. For ease of exposition, we refer to all the tested approaches as the following:

- A1: Sample-average approximation benchmark (MILP)
- A2: Gaussian approximation benchmark (MISOCP)
- A3: DRCC under mean-MAD ambiguity set (MILP)
- A4: DRCC under ∞-Wasserstein ambiguity set (MILP)

We use the hourly wind power data from the ERCOT market from 2018 to 2020 by scaling it according to the size and load demand of the test systems. Due to the seasonality of wind patterns, its uncertainty may vary over the year. Thus, we have used data samples from all four seasons to build the ambiguity set and scenarios. To better compare the performance, we conduct out-of-sample (OOS) experiments by partitioning the dataset into training and testing samples. The optimal solutions are produced using the in-sample training dataset, while the costs and constraint violations are evaluated on the OOS testing dataset. The 14-bus case is used to test and compare all the CC approaches (A1, A2, A3, and A4). Due to the tractability issues of the scenario-based approaches, for the larger 118-bus case we mainly evaluate (A2) and (A3), with a robust benchmark of $\epsilon = 0$ (i.e., zero violation), all of which are scenario-free. Each individual CC tolerance level ϵ_i has been set to be the same value ϵ for simplicity. The test case parameters are obtained from MATPOWER, and the OTS problems (MILP, MISOCP) are solved using

Gurobi. The solver was set to utilize up to 12 available threads with a solution tolerance of $1e^{-2}$. All the numerical tests have been implemented on a regular laptop equipped with Intel® CPU @ 2.60 GHz and 16 GB of RAM using the MATLAB® R2020b simulator.

3.3.3.1 IEEE 14-Bus System Tests

The original IEEE 14-bus system consists of 20 lines and 5 conventional generators. We add 3 wind farms to the case, located at buses 3, 6, and 13, respectively. Given that the marginal gain reduces with more lines to switch, we have used a maximum of $L_o = 3$ opening lines. For the sample-based approaches (A1 and A4), increasing the sample size can lead to more accurate results at the cost of increased problem dimension and computation time. Therefore, we have used $S = 200$ samples for both A1 and A4. The Wasserstein radius δ is selected according to [35, Cor. 1] and tuned to comply with solutions from other approaches (A1, A2, and A3). By setting $1 - \epsilon = 0.95$ or $1 - \epsilon = 0.90$, we compare the optimal switching decisions, run times, OOS costs and constraint violation rates for A1-A4. The results are listed in Table 5.1. To evaluate the OOS testing performance, we used 5,000 random samples from the actual wind data and recorded the percentage of violated constraints by averaging over all testing samples. To avoid cases where the majority of line flow constraints are non-binding under uncertainty, we have slightly adjusted the line flow limits to increase the transmission congestion level as in [285].

The switching decisions tend to vary among the four approaches when $L_o = 2$ or 3. Interestingly, the switching decisions largely remain the same as the tolerance ϵ changes except for A2. Note that the tolerance ϵ more significantly affects the other decisions, namely the generation dispatch $\boldsymbol{g}$ and AGC coefficients $\boldsymbol{\gamma}$. This becomes clear when comparing the OOS costs, as discussed shortly. By and large, the run times of all approaches are very reasonable. Sample-based approaches (A1 and A4) take more time, while the scenario-free ones (A2 and A3) are much faster (within 1 second). For the sample-based A4, the ∞-Wasserstein metric makes its run time comparable to A1, while offering better DRO guarantees.

In terms of OOS performance, the DRCC approaches (A3 and A4) incur slightly higher total costs than the other two. This is expected as the DRCC approaches are designed to account for a variety of distributions in the ambiguity set. Between A3 and A4, the Wasserstein metric has lower OOS costs as its solutions are more data-driven and less conservative, as mentioned earlier. Note that although A3 and A4 produce exactly the same switching decisions, their OOS costs still differ due to their differences in the $\boldsymbol{g}$ and $\boldsymbol{\gamma}$ decisions. This difference can also be observed for all approaches with $L_o = 1$. Under fixed $(1 - \epsilon)$, the OOS costs generally are reduced as L_o increases, and a smaller $(1 - \epsilon)$ allows for more violations of constraints and thus lowers the total costs.

The comparisons on constraint violation and renewable curtailment in OOS testing are very important for evaluating the robustness performance.

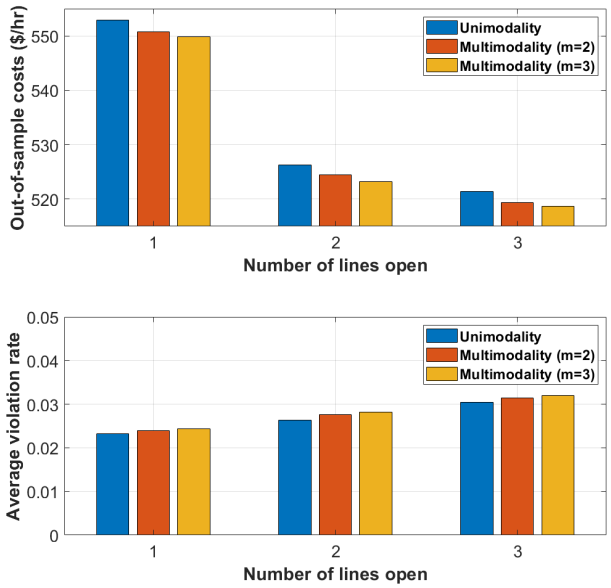


Figure 3.8: Comparisons of OOS costs and average violation rate for unimodality and multimodality models in A3 approach.

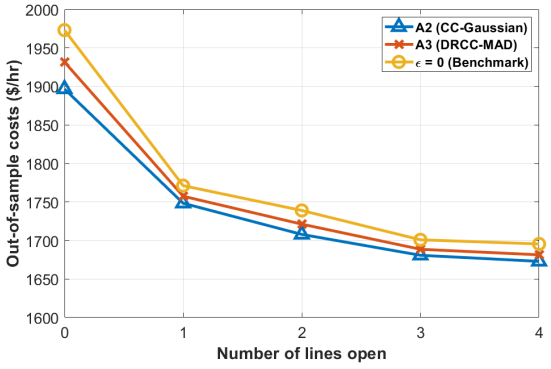
Ideally, the OOS violation rates should not exceed the pre-specified threshold ϵ. However, this is rarely the case for A1, because the SAA design relies on the approximation by the empirical distribution and cannot strictly enforce the robustness guarantees. In addition, A2 has one instance of exceeding $\epsilon = 0.05$ for the case of $L_o = 3$, which speaks to its disadvantage of solely relying on the assumption on Gaussian distributed uncertainty. Compared to A1 and A2, the proposed A3 and A4 have nicely maintained very low constraint violation rates for all choices of L_o, thanks to their DRCC based design principle. This is especially important for a smaller value of ϵ, where the robustness guarantees are more difficult to enforce. Using the renewable curtailment quantification approach in Sec. 3.3.1.3, we have shown the clear improvement of DRCC

approaches (A3 and A4) over CC approaches in reducing curtailment levels. As the former has demonstrated proved robustness guarantees, grid congestion is less likely to occur and so is the renewable curtailment. Fig. 3.8 further shows the trade-off between OOS costs and average violation rates attained by the multimodality model based mean-MAD approaches with $1 - \epsilon = 0.90$. Compared with the unimodality benchmark (A3), we increase the number of modes to be $m = 2$ or 3. We observe that including the multimodality information leads to a less conservative DRCC solution with decreasing OOS costs. Meanwhile, the average violation rates slightly increase with m as a trade-off.

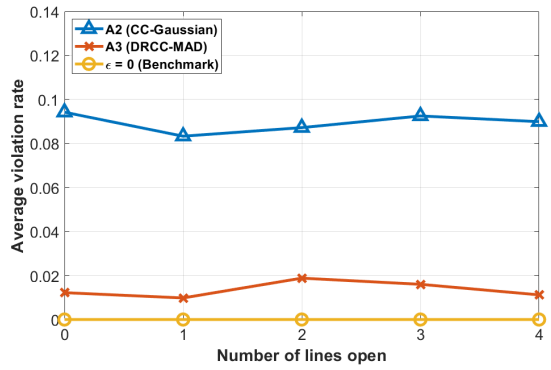
In summary, the proposed DRCC approaches demonstrate a graceful trade-off between the total cost and constraint satisfaction rate. They can reliably limit the occurrence of constraint violations and thus reduce the level of renewable curtailment, at some incremental cost.

3.3.3.2 IEEE 118-Bus System Tests

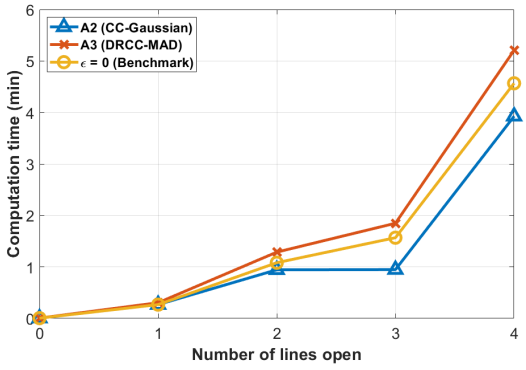
We have also tested the approaches on the IEEE 118-bus system, consisting of 186 lines and 19 conventional generators. Five wind farms have been added, which are located at buses 10, 23, 57, 62 and 86, respectively. Chance constraints have been applied on half of the line flow constraints with the other half strictly enforced. The sample-based methods (A1 and A4) are extremely inefficient for the mixed-integer formulation, especially for large systems (as high as 5-10 hours). Due to this scalability issue, we have only compared the



(a)



(b)



(c)

Figure 3.9: Comparisons of the (a) OOS costs; (b) average violation rates; and (c) computation time for the 118-bus system.

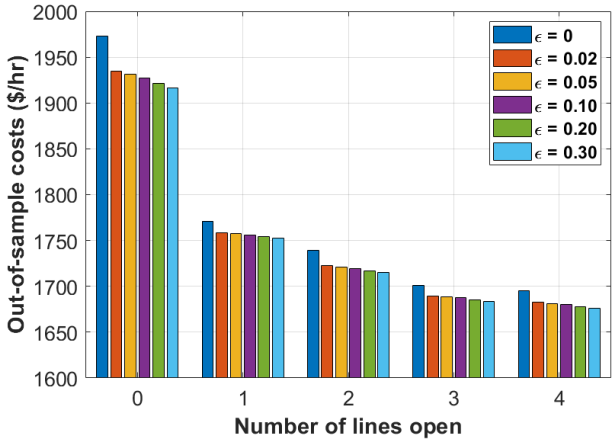


Figure 3.10: Comparisons of the OOS costs attained by A3 under different tolerance levels and L_o values.

scenario-free approaches A2 and A3, with a robust benchmark by setting the tolerance level $\epsilon = 0$.

First, we use the CC tolerance $1 - \epsilon = 95\%$ to compare the OOS performance under different L_o, as plotted in Fig. 3.9a. Overall, the OOS costs increase slightly from A2 to A3, and both are smaller than the benchmark cost. This trend is consistent with the average rate of constraint violations as shown in Fig. 3.9b. Compared with the benchmark, A3 achieves 1.0% cost reduction on average, while A2 achieves 1.8%. Notably, the constraint violation rates for the proposed A3 are nicely maintained around 0.02 which is smaller than the threshold $\epsilon = 0.05$, while those for A2 can go up to roughly 0.09 that exceeds the tolerance level. This large-system test again confirms the aforementioned improvement of the proposed DRCC-OTS approaches over A2 in terms of guaranteed constraint satisfaction. Moreover, we have compared the average

run time, as shown in Fig. 3.9c. In general, the run time of the proposed A3 is on par with the other two, with a moderate increase for larger L_o values. Lastly, we compare the OOS costs of the proposed A3 for different tolerance levels, by varying ϵ in the range of $0\% - 30\%$, as shown in Fig. 3.10. With fixed L_o, a larger ϵ value leads to gradually decreasing costs, by allowing higher occurrences of constraint violations. If we compare to the OOS costs of the benchmark approach ($\epsilon = 0$), the proposed DRCC-based A3 can attain lower costs with a roughly 1.1% reduction on average. Notice that the marginal gain of cost reduction is minimal at higher tolerance levels (ϵ increasing from 20% to 30%). Generally speaking, the range of $[5\%, 20\%]$ is deemed appropriate for ϵ in practical operations [83, 285].

In summary, the proposed DRCC approaches can effectively limit the occurrences of violating line flow constraints by accounting for the distributional ambiguity of uncertainty. In particular, the mean-MAD ambiguity criterion leads to a scenario-free, tractable MILP reformulation, with comparable complexity to the CC and benchmark approaches.

This work [298] considers the chance-constrained optimal transmission switching (CC-OTS) problem to account for renewable uncertainty in power systems. We proposed to simplify the two-stage OTS problem by establishing the equivalence of linear decision rules (LDR) based reformulation. Due to the lack of distributional knowledge on the uncertainty, we pursued a distributionally robust chance-constrained (DRCC) OTS paradigm that can ensure the guarantees over an ambiguity set of uncertainty distributions. Both

moment-based and distance-based ambiguity sets have been considered, leading to scalable MILP problems through dualization. Numerical tests validated the performance improvements of the proposed DRCC approaches over the CC alternatives in terms of guaranteed constraint violation rates. Between the two proposed DRCC-OTS approaches, the one using the mean-MAD ambiguity set brought lower computation complexity on par with other scenario-free approaches, while the one using the Wasserstein ambiguity led to less conservative solutions by adapting to the actual data samples. Future work includes reducing the complexity of scenario-based DRCC-OTS by simplifying the Wasserstein ambiguity set and developing machine learning approaches for accelerated OTS computations in real-time.

Chapter 4

Power System Control Under Wildfires

This chapter presents the optimal power system control problem under uncertain wildfires. We start by introducing the system modeling and the optimal power shut-off problem in Section 4.1. After that, we present the preventive control and corrective control formulation in Section 4.2 and 4.3, respectively. To enable real-time implementation, efficient decomposition and relaxation techniques have been applied for scalable reformulations. Lastly, numerical simulation results are provided in Section 4.4 to demonstrate the performance of the proposed algorithms.

Fast and effective wildfire mitigation has become an emerging challenge for reliable power system operations in recent years. Due to the climate change, wildfire ignitions have increased substantially in both numbers and severity. There were over 58000 wildfires in the United States in 2021, which has burned more than 7.13 million acres. From 2019 to 2021, in the U.S. there was a 16.9% increase in the wildfire cases and a 51.1% increase in the number of acres burned. The wildfires ignitions can be due to various causes such as lightning, arson, power lines, and campfire [248]. Once ignited, the wildfire

The work in this chapter is collaborative research with Los Alamos National Laboratory.

can spread very quickly (up to 10-15 mph), destroying infrastructure in its path. In particular, the power transmission lines located in high-risk areas are more vulnerable to spreading wildfires. This can include burning of wood poles, damage to steel towers and transmission lines. Furthermore, the dense smoke from nearby wildfires can contaminate the transmission line's insulating medium, which can eventually lead to line outages. Therefore, designing effective responses such as network topology control is of critical importance for system restoration tasks under the line outage contingencies.

In an effort to mitigate ongoing wildfires and also prevent future consequences, de-energization of system components (e.g., transmission lines, loads) has been recognized as a well-established solution, which is also known as the Public Safety Power Shutoff (PSPS). As shutting off power can bring economic impact and raise concerns over system reliability, strategically designing the optimal shutoff plan has been the focus of many research works in this area over the very recent few years. Depending on the optimization framework design, these problems generally include maximizing load delivery [64,147], minimizing system risks [18, 146, 211, 212, 243], and enhancing grid resilience [8, 188, 246]. Although these approaches are well-suited for either optimally switching lines under a given wildfire or designing a network topology in preparation for potential wildfires, the majority of existing optimization models is deterministic and has not considered the uncertainty of the spreading wildfire. Essentially, one critical issue is that the information of the fast-evolving wildfires (such as change of network topology) cannot be conveniently incorporated into the

current decision-making framework for real-time control and operations. Recently, a robust decision-making framework has been proposed [241] for highly renewable power systems to deal with wildfire threat. Nonetheless, a stochastic fire spread simulation needs to be done ahead of the attacker-defender model, thus it still raises concern on the efficiency of the proposed model for real-time grid control. Therefore, the problem of efficient real-time optimal topology control under the developing wildfires still remains open.

The goal of this work is to present a real-time optimal topology control strategy under uncertain line outages caused by wildfires. To incorporate the line outage uncertainty, we propose a two-stage stochastic mixed-integer program. The problem utilizes system available resources (generation ramping, load shedding, line switching, etc) to mitigate the evolving wildfires. It aims to minimize the total operating cost, to effectively recover network power balance and reduce load shedding. Depending on the stage of topology control variables (binary), we present both *preventive* and *corrective* control schemes for wildfire mitigation. Specifically, the preventive control seeks to find an optimal network topology in advance the wildfire while the corrective control enables flexible network configuration for each uncertain scenario individually. To allow for real-time implementation of both approaches, we have utilized Benders decomposition for efficiently solving the preventive control problem and applied relaxation techniques for the non-convex corrective control problem. Hence, our work directly addresses the lack of consideration of wildfire uncertainty in the existing framework and is able to provide a real-time net-

work topology re-configuration solution for efficient wildfire mitigation.

4.1 System Modeling

The topology control problem [93] aims to determine the optimal system topology that leads to the minimal total generation cost under given load $\boldsymbol{d}$. For simplicity, linear generation cost is considered and we use $\boldsymbol{c} \in \mathbb{R}^N$ to denote the known vector of generation cost coefficients. Compared with the generation dispatch problem, binary decision variables $z_\ell \in \{0,1\}, \forall \ell$ are further introduced here to allow for flexible transmission line switching (0: open, 1: closed). Thus, the problem is formulated as a mixed-integer linear program (MILP), as given by:

$$\min \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{g} \tag{4.1a}$$

$$\text{s.t.} \quad z_\ell \in \{0,1\},\ \forall \ell \tag{4.1b}$$

$$\theta_i^{\min} \le \theta_i \le \theta_i^{\max},\ \forall i \tag{4.1c}$$

$$g_i^{\min} \le g_i \le g_i^{\max},\ \forall i \tag{4.1d}$$

$$f_\ell^{\min} z_\ell \le f_\ell \le f_\ell^{\max} z_\ell,\ \forall \ell \tag{4.1e}$$

$$b_{ij}(\theta_i - \theta_j) - f_\ell + (1 - z_\ell)\mathrm{M}_\ell \ge 0,\ \forall \ell = (i,j) \tag{4.1f}$$

$$b_{ij}(\theta_i - \theta_j) - f_\ell - (1 - z_\ell)\mathrm{M}_\ell \le 0,\ \forall \ell = (i,j) \tag{4.1g}$$

$$\sum_{\ell=(i,j)} f_\ell - \sum_{\ell=(j,i)} f_\ell = g_i - d_i,\ \forall i \tag{4.1h}$$

$$\sum_{\ell}(1 - z_\ell) \le v \tag{4.1i}$$

where phase angle, generation and line flow constraints are provided in (4.1c)-(4.1e). The constraint (4.1e) enforces the line flow f_ℓ to be zero for any opening line ($z_\ell = 0$), and $f_\ell^{\min} \leq f_\ell \leq f_\ell^{\max}$ for the closed line ($z_\ell = 1$). Constraints (4.1f) and (4.1g) jointly ensures the dc power flow $f_\ell = b_{ij}(\theta_i - \theta_j), \forall \ell = (i,j)$ for all the closed lines ($z_\ell = 1$). On the other hand, under $z_\ell = 0$ and thus $f_\ell = 0$, the two constraints become $b_{ij}(\theta_i - \theta_j) + \mathrm{M}_\ell \geq 0$ and $b_{ij}(\theta_i - \theta_j) - \mathrm{M}_\ell \leq 0$, both of which trivially hold under a sufficiently large M_ℓ and become redundant. This is well-known as the *Big-M* method [111], which is a very useful technique for handling constraints with binary variables. For each line $\ell = (i,j)$, the big-M constant can be set to $\mathrm{M}_{ij} := b_{ij}\Delta\theta_{ij}^{\max}$ with a maximum limit $\Delta\theta_{ij}^{\max}$ according to angle stability. Furthermore, constraint (4.1h) enforces the network power balance [cf. (3.2)]. Lastly, for system stability concerns, constraint (4.1i) is imposed to restrict the total number of lines that can be switched off not to exceed the preset limit v.

Remark 4.1.1 (topology control modeling). The dc power flow model is adopted in this paper for formulating the topology control problem. As the dc model does not include voltage limits or other ac power flow considerations, to extend to the ac model one can use the relaxation-based formulation as introduced in [145]. In addition, post-selection ac power flow analysis [104] can be performed to verify the ac feasibility of resultant solutions.

Accurately modeling the line outages due to wildfire is extremely important for seeking the optimal mitigation strategy. The line outage happening on line ℓ due to wildfire can be modeled as a discrete uncertainty $\xi_\ell = \{0, 1\}$,

where $\xi_\ell = 0$ represents a line outage is triggered by the wildfire, and $\xi_\ell = 1$ represents the line remains intact. The probability of whether a line outage is triggered or not can be modeled as:

$$\mathbb{P}_{\xi_\ell=0} = p_\ell^o, \tag{4.2a}$$

$$\mathbb{P}_{\xi_\ell=1} = p_\ell^c \tag{4.2b}$$

with $p_\ell^o + p_\ell^c = 1$. In general, the probability p_ℓ^o, p_ℓ^c (open, closed) can depend on many environmental factors including temperature, wind, fuels and topography. For instance, a transmission line that locates close to a spreading wildfire, surrounded by dry vegetation will have a larger p_ℓ^o value, and thus a line outage is more likely to happen.

As the topology control includes integer decision variables, further introducing the discrete uncertainty ξ_ℓ leads to more complicated optimization problems. Specifically, when the number of scenarios is large, the resulting mixed-integer program can be difficult to solve for real-time operations. We will next present two topology control schemes, namely preventive control and corrective control, as effective wildfire mitigation solutions. The preventive control scheme seeks an optimal topology ahead of the observation of uncertainty, while the corrective control scheme can further allow for flexible network topology control according to each individual uncertainty realization. Notably, both of these methods are amenable to applying decomposition/relaxation techniques, thus leading to tractable real-time mitigation solutions.

4.2 Preventive Control Formulation

To incorporate the line outage uncertainty into the topology control, we formulate the problem as a two-stage stochastic program [229]. Accordingly, a *here-and-now* decision is sought on the first stage, and *wait-and-see* (or recourse) decisions are made upon observation of the uncertainty on the second stage. We propose to fully utilize available power system operations including topology control, generation ramping, load shedding as countermeasures to the ongoing wildfire. To this end, we first present the preventive control formulation as the following optimization problem:

$$\min \quad \boldsymbol{c}_1^\mathsf{T}\boldsymbol{g} + \mathbb{E}[\boldsymbol{c}_2^\mathsf{T}\boldsymbol{r}^k] + \mathbb{E}[\boldsymbol{c}_3^\mathsf{T}\boldsymbol{s}^k] \tag{4.3a}$$

$$\text{s.t.} \quad z_{\ell,\text{on}} \in \{0,1\}, z_{\ell,\text{off}} \in \{0,1\}, \ \forall \ell \tag{4.3b}$$

$$g_i^{\min} \le g_i \le g_i^{\max}, \ \forall i \tag{4.3c}$$

$$\theta_i^{\min} \le \theta_i^k \le \theta_i^{\max}, \ \forall i,k \tag{4.3d}$$

$$f_\ell^{\min}\xi_\ell^k(1-z_{\ell,\text{on}}) \le f_\ell^k \le f_\ell^{\max}\xi_\ell^k(1-z_{\ell,\text{on}}), \ \forall \ell,k \tag{4.3e}$$

$$f_\ell^{\min}\xi_\ell^k z_{\ell,\text{off}} \le f_\ell^k \le f_\ell^{\max}\xi_\ell^k z_{\ell,\text{off}}, \ \forall \ell,k \tag{4.3f}$$

$$b_{ij}(\theta_i^k-\theta_j^k) - f_\ell^k + (1-\xi_\ell^k+\xi_\ell^k z_{\ell,\text{on}}^k)\text{M}_\ell \ge 0, \ \forall \ell,k \tag{4.3g}$$

$$b_{ij}(\theta_i^k-\theta_j^k) - f_\ell^k - (1-\xi_\ell^k+\xi_\ell^k z_{\ell,\text{on}}^k)\text{M}_\ell \le 0, \ \forall \ell,k \tag{4.3h}$$

$$b_{ij}(\theta_i^k-\theta_j^k) - f_\ell^k + (1-\xi_\ell^k z_{\ell,\text{off}})\text{M}_\ell \ge 0, \ \forall \ell,k \tag{4.3i}$$

$$b_{ij}(\theta_i^k-\theta_j^k) - f_\ell^k - (1-\xi_\ell^k z_{\ell,\text{off}})\text{M}_\ell \le 0, \ \forall \ell,k \tag{4.3j}$$

$$\sum_{\ell=(i,j)} f_\ell^k - \sum_{\ell=(j,i)} f_\ell^k = g_i - r_i^k - (d_i - s_i^k), \ \forall i,k \tag{4.3k}$$

$$r_i^k \ge 0, \ \forall i,k \tag{4.3l}$$

$$r_i^k + g_i \le g_i^{\max},\ \forall i, k \tag{4.3m}$$

$$0 \le s_i^k \le d_i,\ \forall i, k \tag{4.3n}$$

$$\sum_{\ell} (z_{\ell,\text{on}} + z_{\ell,\text{off}}) \le v \tag{4.3o}$$

where the first-stage decision variables include generation and topology control decisions $\boldsymbol{x} = \{\boldsymbol{g}, z_{\ell,\text{on}}, z_{\ell,\text{off}}\}$. Second-stage decision variables include phase angle, line flow, generation ramping and load shedding $\boldsymbol{y} = \{\theta_i^k, f_\ell^k, r_i^k, s_i^k\}, \forall i$, for each uncertainty scenario $k \in \mathcal{K}$. The load d_i is provided by short-term load forecasting (e.g., [291]), and the mismatch will be handled by generation ramping s_i^k in the second stage. Note that different from the integer variable z_ℓ (*line status*) in the deterministic topology control problem (4.1), here we introduce $z_{\ell,\text{on}}$ and $z_{\ell,\text{off}}$ as *operations* on currently closed and open lines. Specifically, $z_{\ell,\text{on}} = 1$ means switching off the closed line ℓ, and $z_{\ell,\text{on}} = 0$ represents no operation on the line (remaining closed). Similarly, $z_{\ell,\text{off}} = 1$ means switching on the open line ℓ, while $z_{\ell,\text{off}} = 0$ means no operation on the line (remaining open). As frequent topology switching actions might be necessary for mitigating the developing wildfire, this change can easily adapt to various initial grid topology conditions. The optimization problem aims to minimize the total operation costs from generation $\boldsymbol{c}_1^\mathsf{T}\boldsymbol{g}$, expected generation ramping costs $\mathbb{E}[\boldsymbol{c}_2^\mathsf{T}\boldsymbol{r}^k]$, and expected load shedding costs $\mathbb{E}[\boldsymbol{c}_3^\mathsf{T}\boldsymbol{s}^k]$. The costs coefficients are designed such that $\boldsymbol{c}_{1,i} < \boldsymbol{c}_{2,i} \ll \boldsymbol{c}_{3,i},\ \forall i$, where generation should be prioritized before generation ramping and load shedding operations are activated. Load shedding has the highest operation costs among the three, as this will avoid load shedding as much as possible unless it is inevitable un-

der the contingency. The problem seeks an optimal topology control decision (indicated by $z_{\ell,\text{on}}$ and $z_{\ell,\text{off}}$) that can nicely balance generation, ramping and load shedding, leading to the lowest possible total operation costs.

The constraints for problem (4.3) are discussed in detail here. Constraint (4.3c) sets the generation limits. Phase angle limits for each scenario k are enforced in (4.3d). Furthermore, (4.3e) and (4.3f) are line flow constraints. The line switching operations $z_{\ell,\text{on}}$ and $z_{\ell,\text{off}}$ are coupled with the line outage uncertainty ξ_ℓ^k to represent the actual line status under both factors. Specifically, line ℓ is only connected when $\xi_\ell^k = 1$ and $1 - z_{\ell,\text{on}} = 1$ (or $z_{\ell,\text{off}} = 1$), and correspondingly line flow limits are enforced. Otherwise, the constraints become an equality constraint $f_\ell^k = 0$. Likewise, constraints (4.3g)-(4.3j) are big-M constraints, which only enforce power flow equations for connected lines [cf. (4.1f)-(4.1g)]. Network power balance is guaranteed in constraint (4.3k). Additionally, generation ramping constraints are cast in (4.3l)-(4.3m) and load shedding constraints are given in (4.3n). Lastly, the total number of line switching operations is restricted by constraint (4.3o).

4.2.1 Preventive Control via Benders Decomposition

The problem (4.3) is a two-stage stochastic mixed-integer program, where the constraints scale with number of scenarios k. This will increase the computational complexity significantly for large k, making it very difficult for real-time implementation. Since the second-stage constraints are repeated for different scenarios, some of these constraints might be unnecessary and

redundant for finding the optimal solution. This motivates us to consider decomposition techniques for more efficient implementation. To achieve this, let us consider the two-stage problem (4.3) in the following generalized form:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \ \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} + \mathbb{E}[Z(\boldsymbol{x},\boldsymbol{\xi})] \tag{4.4}$$

where

$$Z(\boldsymbol{x},\boldsymbol{\xi}) = \min \quad \boldsymbol{q}^{\mathsf{T}}\boldsymbol{y} \tag{4.5a}$$

$$\text{s.t.} \quad \boldsymbol{y} \in \mathbb{R}_{+}^{N_2} \tag{4.5b}$$

$$\boldsymbol{W}\boldsymbol{y} \geq \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi} + \boldsymbol{h}(\boldsymbol{x}) \tag{4.5c}$$

where $\boldsymbol{x}$ denotes all the first stage decision variables and $\boldsymbol{y}$ denotes the collection of second-stage decision variables. Given all scenarios $k \in \mathcal{K}$, an equivalent monolithic formulation can be given by combining (4.4) and (4.5), as the following:

$$\min \quad \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} + \sum_{k\in\mathcal{K}} p_k \boldsymbol{q}_k^{\mathsf{T}}\boldsymbol{y}^k \tag{4.6a}$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{y}^1, \boldsymbol{y}^2, \cdots, \boldsymbol{y}^k \in \mathbb{R}^{N_2} \tag{4.6b}$$

$$\boldsymbol{W}\boldsymbol{y}^k \geq \boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}),\ \forall k \in \mathcal{K} \tag{4.6c}$$

where $p_k = \mathbb{P}(\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi}^k)$ is the probability of line outage uncertainty $\boldsymbol{\xi}^k$. This minimization problem corresponds to our preventive control formulation as in (4.3). Recall that the second-stage problem is a linear program (LP). Thus, strong duality holds and the dual form of the second-stage problem for each contingency scenario k can be given as:

$$Z_d(\boldsymbol{x},\boldsymbol{\xi}^k) = \max \quad (\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^{\mathsf{T}}\boldsymbol{\pi}^k \tag{4.7a}$$

$$\text{s.t.} \quad \boldsymbol{\pi}^k \in \mathbb{R}^J \tag{4.7b}$$

$$\boldsymbol{q}_k \geq \boldsymbol{W}^\mathsf{T} \boldsymbol{\pi}^k \tag{4.7c}$$

For the dual problem (4.7), we can define set $\boldsymbol{\Pi}^k = \{\boldsymbol{\pi}^k \in \mathbb{R}^J : \boldsymbol{q}_k \geq \boldsymbol{W}^\mathsf{T}\boldsymbol{\pi}^k\}$ for each contingency k. This feasible set is a polytope with finitely many extreme points $\boldsymbol{\pi}_1^k, \boldsymbol{\pi}_2^k, \cdots, \boldsymbol{\pi}_S^k$. According to convex optimization theory [46], the optimal solution to linear program (4.7) belongs to the set $\boldsymbol{\Pi}^k$. Thus, problem (4.7) can be equivalently represented as $Z_d(\boldsymbol{x}, \boldsymbol{\xi}^k) = \max_{s\in\mathcal{S}}(\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^\mathsf{T}\boldsymbol{\pi}_s^k$, which is a piecewise linear function.

Hence, the problem (4.6) is equivalent to:

$$\min \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \sum_{k\in\mathcal{K}} p_k \max_{s\in\mathcal{S}} (\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^\mathsf{T}\boldsymbol{\pi}_s^k \tag{4.8a}$$

Epigraph reformulation can be further applied to this min-max problem. Specifically, we can introduce $\psi^k \in \mathbb{R}$ to replace the maximization term for each scenario k. By doing so, (4.6) is eventually reformulated to the following problem:

$$\min \quad \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + \sum_{k\in\mathcal{K}} p_k \psi^k \tag{4.9a}$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X},\ \psi^1, \psi^2, \cdots, \psi^k \in \mathbb{R} \tag{4.9b}$$

$$\psi^k \geq (\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^\mathsf{T}\boldsymbol{\pi}_s^k,\ \forall k \in \mathcal{K},\ s \in \mathcal{S} \tag{4.9c}$$

Compared with (4.6), this reformulated problem is clearly more efficient to solve. This is because (4.9) only uses the extreme points of the dual problem, instead of enforcing all the constraints in the primal. In addition, the number

of variables and constraints is greatly reduced in (4.9). Of course, the key to solving (4.9) is to find the extreme points $\boldsymbol{\pi}_s^k$ for each scenario k. To do this, we can utilize *Benders decomposition* [151] to efficiently search for these extreme points in a iterative way. The Benders decomposition technique has been popularly adopted in existing literature (e.g., [42, 78, 136, 163]) on power system topology design problems under uncertainty.

The decomposition algorithm works by repetitively adding cuts to the master problem using solutions from the sub-problem. The detailed steps of solving preventive control problem (4.3) using Benders decomposition are illustrated in Algorithm 2. The algorithm starts by solving the master problem with an initial set of extreme points. After that, we fix the solution of $\boldsymbol{x}$ and solve the second-stage sub-problem (4.7). As long as the condition $Z_d(\boldsymbol{x}, \boldsymbol{\xi}^k) \leq \hat{\psi}^k$ is not satisfied, we record the corresponding extreme points $\hat{\boldsymbol{\pi}}^k$ and add the cuts $(\psi^k \geq (\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^\mathsf{T}\hat{\boldsymbol{\pi}}^k)$ to the master problem repetitively to avoid same violations from future iterations. The algorithm is terminated until the iteration reaches the limit, or one can also utilize the following convergence criteria based on the difference between upper and lower bounds.

Remark 4.2.1 (convergence criteria). At each iteration, the objective value of master problem (4.9) provides a lower bound. Meanwhile, after solving sub-problem (4.7), we have $\boldsymbol{c}^\mathsf{T}\hat{\boldsymbol{x}} + \sum_{k\in\mathcal{K}} p_k Z_d(\hat{\boldsymbol{x}}, \boldsymbol{\xi}^k)$ that constitutes an upper bound. The algorithm is terminated when the difference of the two bounds fall below a convergence threshold ϵ.

It is worth mentioning that the first-stage solutions $\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{g}}, \hat{z}_{\ell,\text{on}}, \hat{z}_{\ell,\text{off}}\}$

from solving the master problem include both integer and continuous variables, can possibly cause infeasibility in the subsequent sub-problems. This concern can be conveniently addressed by adding feasibility cuts in addition to the aforementioned optimality cuts.

Remark 4.2.2 (feasibility cuts). To avoid infeasible second-stage sub-problems in Algorithm 2, feasibility cuts can be introduced when solving the master problem. Specifically, the cuts $0 \geq (\boldsymbol{T}(\boldsymbol{x})\boldsymbol{\xi}^k + \boldsymbol{h}(\boldsymbol{x}))^\mathsf{T}\boldsymbol{\omega}^k$ will be further added to provide feasibility guarantees, where $\boldsymbol{\omega}^k$ represents the extreme rays of the polytope $\boldsymbol{\Pi}^k$.

To summarize, the decomposition-based approach as in Algorithm 2 utilizes the solutions from the dual sub-problem to improve the searching of optimal decisions. Instead of solving the original large-scale mixed-integer program (4.3), it greatly reduces the number of variables and constraints. Notably, the master problem is more efficiently solvable using the extreme points and the sub-problems are simply linear programs (LPs). The Benders decomposition algorithm normally converges within a few iterations for many practical problems. Some earlier research (e.g., [173,210]) also provides various approaches for accelerating the Benders algorithm. Overall, the problem (4.3) can be solved efficiently using the proposed decomposition approach, thus making it suitable for real-time implementations.

atex@errorlgobble

Algorithm 2 Preventive Control via Benders Decomposition

Input: Problem parameters
Output: $z_{\ell,\text{on}}, z_{\ell,\text{off}}, \boldsymbol{g}$

Initialize $\boldsymbol{\Pi}^k = \varnothing$ (set of current extreme points $\boldsymbol{\pi}^k$).
for $t = 1$ to $t_{\max}$ **do**
  Solve master problem (4.9) using $\boldsymbol{\pi}^k \in \boldsymbol{\Pi}^k$ and obtain the optimal values of $\hat{\boldsymbol{x}}$, $\hat{\psi}^k$.
  Set $\boldsymbol{x} \leftarrow \hat{\boldsymbol{x}}$ and solve sub-problem (4.7).
  if $Z_d(\boldsymbol{x}, \boldsymbol{\xi}^k) \leq \hat{\psi}^k$ **then**
    Terminate algorithm and optimal solution $\boldsymbol{x}^* = \boldsymbol{x}$.
  end if
  Obtain the optimal solution $\hat{\boldsymbol{\pi}}^k$ to sub-problem (4.7).
  Adding cuts $\boldsymbol{\Pi}^k \leftarrow \boldsymbol{\Pi}^k \cup \{\hat{\boldsymbol{\pi}}^k\}$.
end for
return $z_{\ell,\text{on}}, z_{\ell,\text{off}}, \boldsymbol{g}$

4.3 Corrective Control Formulation

The preventive control allows to perform topology control prior to the wildfire contingency, yet it has not fully utilized the potential of network reconfiguration as response to the actual contingency. This inspires us to further consider the *corrective control* formulation, which moves the line switching operation to the second stage. Thus, the corrective control can be formulated as the following two-stage stochastic mixed-integer program:

$$\min \quad \boldsymbol{c}_1^\top \boldsymbol{g} + \mathbb{E}[\boldsymbol{c}_2^\top \boldsymbol{r}^k] + \mathbb{E}[\boldsymbol{c}_3^\top \boldsymbol{s}^k] \tag{4.10a}$$

$$\text{s.t.} \quad z^k_{\ell,\text{on}} \in \{0,1\}, z^k_{\ell,\text{off}} \in \{0,1\}, \ \forall \ell, k \tag{4.10b}$$

$$(4.3c) - (4.3d), (4.3k) - (4.3n) \tag{4.10c}$$

$$f_\ell^{\min} \xi_\ell^k (1 - z_{\ell,\text{on}}^k) \leq f_\ell^k \leq f_\ell^{\max} \xi_\ell^k (1 - z_{\ell,\text{on}}^k),\ \forall \ell, k \tag{4.10d}$$

$$f_\ell^{\min} \xi_\ell^k z_{\ell,\text{off}}^k \leq f_\ell^k \leq f_\ell^{\max} \xi_\ell^k z_{\ell,\text{off}}^k,\ \forall \ell, k \tag{4.10e}$$

$$b_{ij}(\theta_i^k - \theta_j^k) - f_\ell^k + (1 - \xi_\ell^k + \xi_\ell^k z_{\ell,\text{on}}^k)\text{M}_\ell \geq 0,\ \forall \ell, k \tag{4.10f}$$

$$b_{ij}(\theta_i^k - \theta_j^k) - f_\ell^k - (1 - \xi_\ell^k + \xi_\ell^k z_{\ell,\text{on}}^k)\text{M}_\ell \leq 0,\ \forall \ell, k \tag{4.10g}$$

$$b_{ij}(\theta_i^k - \theta_j^k) - f_\ell^k + (1 - \xi_\ell^k z_{\ell,\text{off}}^k)\text{M}_\ell \geq 0,\ \forall \ell, k \tag{4.10h}$$

$$b_{ij}(\theta_i^k - \theta_j^k) - f_\ell^k - (1 - \xi_\ell^k z_{\ell,\text{off}}^k)\text{M}_\ell \leq 0,\ \forall \ell, k \tag{4.10i}$$

$$\sum_\ell (z_{\ell,\text{on}}^k + z_{\ell,\text{off}}^k) \leq v,\ \forall k \tag{4.10j}$$

where first-stage decisions include generation only $\boldsymbol{x} = \{\boldsymbol{g}\}$. Second-stage decision variables include phase angle, line flow, generation ramping, load shedding as well as topology control $\boldsymbol{y} = \{\theta_i^k, f_\ell^k, r_i^k, s_i^k, z_{\ell,\text{on}}^k, z_{\ell,\text{off}}^k\}, \forall i$, for each uncertainty scenario $k \in \mathcal{K}$. Compared with the preventive control formulation, the major difference is that the integer decision variables are moved from the first stage to the second stage. In particular, the line switching operation decisions $z_{\ell,\text{on}}, z_{\ell,\text{off}}$ are now defined for each contingency k individually. This enables system operators to make on-target topology control decisions upon (or shortly before) realization of the actual contingency. However, this problem includes more variables and constraints than (4.3), and the two-stage stochastic program with integer recourse is in general challenging to solve [151], as the second-stage problems are non-convex. This makes algorithms such as Benders decomposition invalid, due to the difficulty in approximating $\mathbb{E}[Z(\boldsymbol{x}, \boldsymbol{\xi})]$ effectively.

To address the aforementioned issues due to integer recourse, we propose to utilize convex relaxation technique to solve the corrective control problem (4.10). The detailed steps are described in Algorithm 3. Specifically, we can relax the second-stage integer variables ($z_\ell^k \in (\{0,1\})$ to continuous ($z_\ell^k \in (0,1)$). Applying convex relaxation for the corrective control problem (4.10) leads to the following linear program (LP):

$$\min \quad \boldsymbol{c}_1^\mathsf{T}\boldsymbol{g} + \mathbb{E}[\boldsymbol{c}_2^\mathsf{T}\boldsymbol{r}^k] + \mathbb{E}[\boldsymbol{c}_3^\mathsf{T}\boldsymbol{s}^k] \tag{4.11a}$$

$$\text{s.t.} \quad z_{\ell,\text{on}}^k \in (0,1), z_{\ell,\text{off}}^k \in (0,1),\ \forall \ell, k \tag{4.11b}$$

$$(4.10\text{c}) - (4.10\text{j}) \tag{4.11c}$$

where the line switching decisions $z_{\ell,\text{on}}^k$ and $z_{\ell,\text{off}}^k$ are allowed to take any values between 0 and 1. Clearly, solving this LP problem provides a lower bound to the original problem (4.10) due to the convex relaxation. As we are primarily seeking line switching that can be either open ($z_\ell^k = 0$) or close ($z_\ell^k = 1$), the optimal solutions of relaxed $z_{\ell,\text{on}}^k$ and $z_{\ell,\text{off}}^k$ may not make sense in terms of actual operations. However, the solutions can suggest which lines have the potential to be operated. For instance, assuming $\xi_l^k = 1$, then under constraint (4.10d), when $z_{\ell,\text{on}}^k$ is close to 1 (switching off closed lines), the line flow is restricted to a small range around zero. On the other hand, when $z_{\ell,\text{on}}^k$ is close to 0 (no operations), the constraint is enforcing slightly restrictive line flow limits. For (4.10f) and (4.10g) (similarly (4.10h) and (4.10i)), if the big-M constant M_ℓ is too large, then the constraints are always redundant for any $z_\ell^k \neq 0$. Thus, it is essential to tighten the big-M constant M_ℓ to achieve better approximation performance.

Proposition 4.3.1 (bounds for big-M constants)**.** *We can formulate the following optimization problem to obtain the constant* M_ℓ^k *for each contingency* ξ_ℓ^k *individually:*

$$\min \quad \mathrm{M}_\ell^k \tag{4.12a}$$

$$s.t. \quad (4.10\mathrm{b}) - (4.10\mathrm{j}) \tag{4.12b}$$

$$\mathrm{M}_\ell^k \geq \|\boldsymbol{K}\boldsymbol{\theta}\|_\infty \tag{4.12c}$$

where we are seeking to obtain an upper bound for all lines $l = (i, j)$ *under each contingency* k*. The big-M constants* M_ℓ^k *can be updated using the optimal objective value of* (4.12)*.*

The problem (4.12) aims to find the smallest upper bound to $b_{ij}(\theta_i - \theta_j)$ for all lines l under scenario k. Note that the objective in (4.12) can subject to linear reformulation [46], and there are only one set of constraints for contingency ξ_ℓ^k here instead of replicating those for all scenarios as in (4.10), both of which make the problem efficiently solvable. Furthermore, this step of refining the big-M constants can be completed offline and does not affect the real-time computation of the corrective control method.

With the tightened big-M constants M_ℓ^k, let us again assume that $\xi_l^k = 1$. When $z_{\ell,\mathrm{on}}^k$ is close to 1, the line flow $f_l^k \approx 0$ [cf. (4.10d)]. Thus, constraints (4.10f) and (4.10g) jointly lead to:

$$-z_{\ell,\mathrm{on}}^k \mathrm{M}_\ell^k \lessapprox b_{ij}(\theta_i^k - \theta_j^k) \lessapprox z_{\ell,\mathrm{on}}^k \mathrm{M}_\ell^k \tag{4.13}$$

which provides a slightly restrictive approximation for $z^k_{\ell,\text{on}} \neq 1$. Due to how we refine the big-M constant, (4.13) always hold for $z^k_{\ell,\text{on}} = 1$, and thus become redundant. As $z^k_{\ell,\text{on}}$ is getting closer and closer to 1, the constraint is more likely to become redundant under (4.13). On the other hand, when $z^k_{\ell,\text{on}}$ is close to 0, it suggests the line to remain closed and (4.10f)-(4.10g) jointly become:

$$b_{ij}(\theta^k_i - \theta^k_j) - z^k_{\ell,\text{on}}\text{M}_\ell \leq f^k_\ell \leq b_{ij}(\theta^k_i - \theta^k_j) + z^k_{\ell,\text{on}}\text{M}_\ell \tag{4.14}$$

where instead of strictly enforcing power flow $f^k_\ell = b_{ij}(\theta^k_i - \theta^k_j)$, here the line flow is relaxed by $z^k_{\ell,\text{on}}\text{M}_\ell$. As $z^k_{\ell,\text{on}}$ gradually gets closer to 0, the constraint (4.14) becomes increasingly accurate to enforcing the power flow equation for line l. Note that all the discussions so far are based on the assumption that $\xi^k_l = 1$. When $\xi^k_l = 0$, the results become trivial as (4.10f)-(4.10g) boil down to $b_{ij}(\theta^k_i - \theta^k_j) - f^k_\ell + \text{M}_\ell \geq 0$ and $b_{ij}(\theta^k_i - \theta^k_j) - f^k_\ell - \text{M}_\ell \leq 0$, which always hold and thus become redundant.

4.3.1 Corrective Control via Convex Relaxation

We propose to utilize convex relaxation technique for efficiently solving corrective control problem (4.10), and the detailed steps are illustrated in Algorithm 3. In particular, we first tighten the big-M constants M^k_ℓ using Proposition 4.3.1. After that, we apply relaxation on the binary variables and solve LP problem (4.11), which gives us a lower bound to problem (4.10). Using the solutions from (4.11), we next rank the optimal solution of z_l in descending order ($\boldsymbol{z} = \{z_1, z_2, \cdots, z_L\}$). As we discussed, the solutions in fact provides the prioritized lines to be operated. Due to the physical restrictions

Algorithm 3 Corrective Control via Convex Relaxation

Input: Problem parameters
Output: $z^k_{\ell,\text{on}}, z^k_{\ell,\text{off}}, \boldsymbol{g}$

1: Obtain big-M constants using (4.12) offline.
2: Solve the relaxed problem (4.11) with updated big-M values.
3: Obtain the objective cost f_{lb}, which constitutes a lower bound for the original problem (4.10).
4: For each scenario k, record the optimal solution of z_l in descending order ($\boldsymbol{z} = \{z_1, z_2, \cdots, z_L\}$).
5: Set $\boldsymbol{z}^0 = \{z_1, \cdots, z_v = 1; z_{v+1}, \cdots, z_L = 0\}$, and set the initial upper bound $f_{\text{ub}} = f^0_{\text{ub}}$.
6: **for** $t = 1$ to $t_{\max}$ **do**
7: Policy update: Use z_{v+t} to replace switching candidates in $\boldsymbol{z}^{t-1}$. Find $\boldsymbol{z}^t$ with the smallest objective cost f^t_{ub}.
8: **if** $(f^t_{\text{ub}} < f_{\text{ub}})$ **then**
9: Record $\delta_{\text{ub}} = |f^t_{\text{ub}} - f^{t-1}_{\text{ub}}|$.
10: Set $f_{\text{ub}} \leftarrow f^t_{\text{ub}}$, $\boldsymbol{z} \leftarrow \boldsymbol{z}^t$.
11: **end if**
12: Obtain optimal generation: Update $\boldsymbol{z}$ in (4.10) and solve for $\boldsymbol{g}$ (linear program).
13: **if** $(\delta_{\text{ub}} < \epsilon)$ **then**
14: Terminate algorithm.
15: **end if**
16: **end for**
17: **return** $z^k_{\ell,\text{on}}, z^k_{\ell,\text{off}}, \boldsymbol{g}$

of the problem formulation, the optimal solutions of z_l (continuous) are very commonly distributed either near 0 or 1, which has also been confirmed in our numerical simulations later. Given the limit v on line switching operations, the initial candidates can be formed by rounding the largest v values to one ($z_1, \cdots, z_v = 1$) and all the other z_l to be zero ($z_{v+1}, \cdots, z_L = 0$). This

solution via convex relaxation can be used to re-solve problem (4.10), which constitutes an LP problem that gives us an upper bound to problem (4.10). To further improve the solution, a simple yet effective local searching algorithm can be incorporated. Specifically, for every iteration t, we use the next switching candidate z_{v+t} on the list to replace the current candidates and update the switching candidate set whenever the upper bound goes down. This searching procedures terminate when the iteration reaches the limit or the upper bound does not improve evidently ($\delta_{\text{ub}} = |f_{\text{ub}}^{t} - f_{\text{ub}}^{t-1}| < \epsilon$). Note that this algorithm by using convex relaxation only provides a tractable approximation to the original problem (4.10). It offers an approximated feasible solution, but does not necessarily guarantee the optimality. However, as in any approximation method, the gap between upper bound f_{ub} and lower bound f_{lb} can be very useful for evaluating the approximation accuracy of the feasible solution obtained from Algorithm 3.

Overall, the corrective control problem moves the topology control decisions to the second stage, thus enabling more flexible system responses for specific contingencies in real time. Due to the difficulty of exactly solving the resulting optimization problem with integer recourse, we propose an efficient algorithm through relaxing the integer variables and proactively improving the solutions. The proposed algorithm provides a tractable corrective control solution, with feasibility guarantees. It enables the system operators to make more on-target responses, especially when preventive control is not available and faster topology control responses are needed.

4.4 Numerical Results

4.4.1 RTS-GMLC System Tests

In this section, we use the RTS-GMLC system [25] to demonstrate the effectiveness of preventive/corrective control schemes under uncertain wildfire events. The RTS-GMLC system is illustrated in Fig. 4.1, which consists of 73 buses and 120 transmission lines. This synthetic system covers a geographical region of roughly 250×250 miles and can be utilized to represent an area in southwestern United States, which extends roughly from Los Angeles to Las Vegas in California. As the system includes detailed geographic information, other related information such as actual load, weather data (e.g., wind, temperature) can be conveniently incorporated for research on power system planning/operations under natural disasters [147, 211, 212, 274]. The wildfire can be generated using the real-time data provided by California Department of Forestry and Fire Protection, which has detailed geographic information of wildfire incidents in California. The location of these wildfires are then mapped into the RTS-GMLC system to generate the line outage uncertainty. The topology control problems have been implemented on a regular laptop equipped with Intel® CPU @ 2.60 GHz and 12 GB of RAM using the Julia language [39], and solved by the CPLEX optimization solver. The solver was set up to utilize up to 12 available threads with an optimality gap of 0.01%.

As mentioned earlier, the weights of generation, ramping and load shedding in the objective function are crucial for determining the optimal preventive/corrective control solutions. To avoid load shedding to the fullest extent,

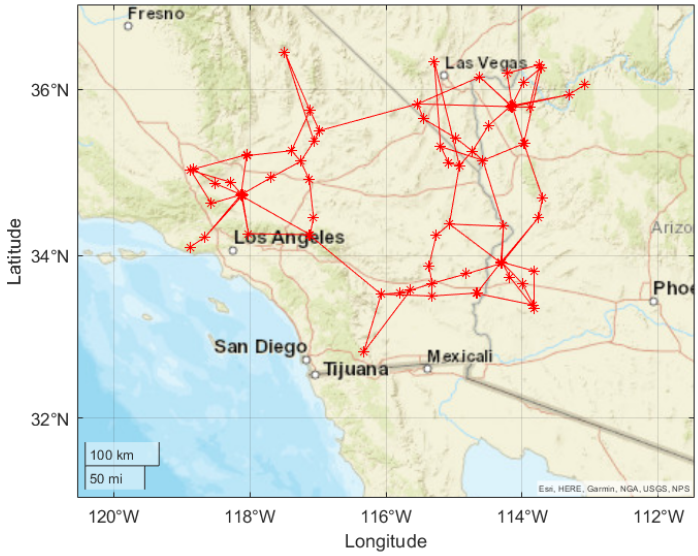


Figure 4.1: RTS 73-bus system with geographic information.

the design of cost coefficients should $\boldsymbol{c}_{1,i} < \boldsymbol{c}_{2,i} < \boldsymbol{c}_{3,i}$, $\forall i$, for both preventive control (4.3) and corrective control (4.10). We set the generation ramping cost $\boldsymbol{c}_2$ to be slightly higher than the generation cost $\boldsymbol{c}_1$. The relation between generation ramping costs and load shedding costs can be characterized by:

$$\boldsymbol{c}_3 = \alpha \cdot \boldsymbol{c}_2$$

where $\alpha \geq 1$ is a parameter that allows to adjust the selection of generation ramping and load shedding under the contingency. A large α value enables to fully prioritize generation ramping before evoking the load shedding, while smaller α values seek to strike a balance between the two actions.

We first fix the parameter $\alpha = 5$, and compare the performance of preventive/corrective control schemes together with a base case in which no

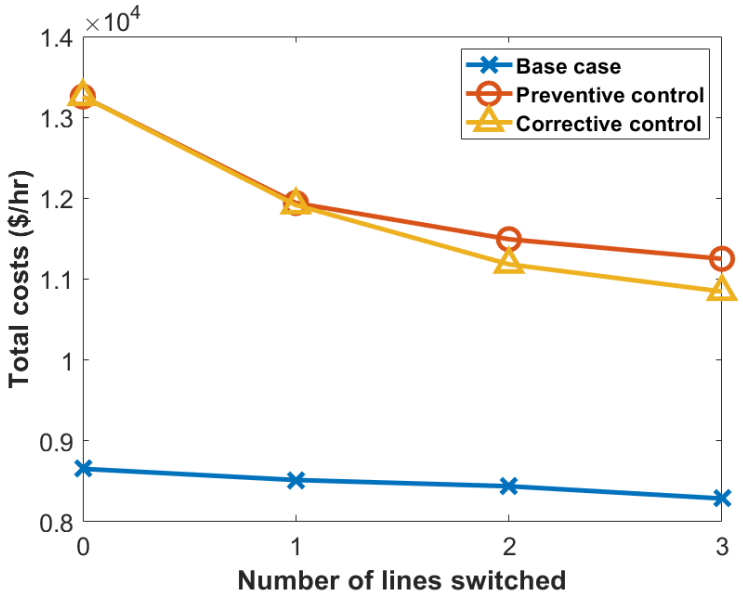


Figure 4.2: Comparison of total costs under different number of switching lines.

wildfire is involved. Earlier studies [122, 124, 207] have shown that the incremental reduction of total operating cost reduces rapidly when v reaches a certain level. In fact, practical choices of v are relatively small for large systems. Given that the marginal gain reduces as more transmission lines are switched, we have considered a maximum of $v = 3$ lines to switch for the numerical tests. The comparison of total costs under different switching limits v is illustrated in Fig. 4.2. The costs for the base case are the lowest among all three, as there is no wildfire or uncertainty in the system and thus generation ramping/load shedding are not required. When $v = 0$, preventive and corrective control have the same operating costs, since the line switching operation is not enabled yet. The increase on the total costs compared with the base

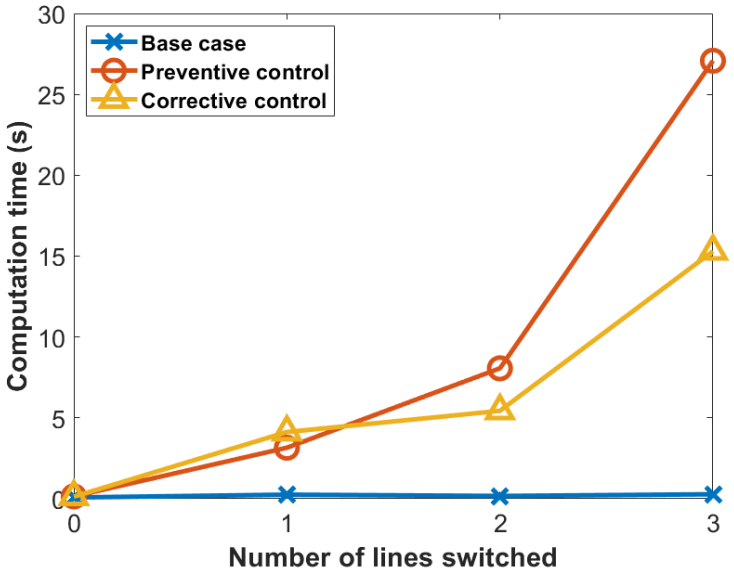


Figure 4.3: Comparison of computation time under different number of switching lines.

case is mainly due to generation ramping and load shedding in response to the wildfire. When $v \geq 1$, the corrective control starts to demonstrate lower costs than the preventive control. This is because the former approach allows to configure network topology separately for each individual uncertainty realization, instead of using a fixed topology uniformly for all scenarios. As v increases, the costs go down steadily. Through switching the transmission lines, the preventive and corrective control approaches reduce the total costs by 15.1% and 18.2%, respectively. We have also compared the computation time, as shown in Fig. 4.3. As v increases, the computation time goes up due to the increased complexity of solving resulting optimization problems. Overall, both preventive control and corrective control are solved efficiently using the

proposed algorithms. In particular, it takes less than 30 seconds to solve the preventive control problem, and less than 15 seconds to solve the corrective control problem. Notably, the preventive control using Benders decomposition provides *exact* solutions to solving the exact problem (4.3), at a much lower computational complexity. On the other hand, the corrective control via convex relaxation, although does not always guarantee optimality, provides a satisfactory feasible solution in a very efficient way.

To examine the trade-off between utilizing generation ramping and load shedding, we have performed more comprehensive numerical simulations with varying values of α. The value of α is varied from 1 to 10, and we have again tested preventive/corrective control and the base case with no wildfire. The comparison of costs under selected α values is shown in TABLE 4.1 and TABLE 4.2. In the base case, the generation ramping and load shedding are not needed under normal operating condition, and thus the solutions are not affected by varied α values. When line outages are introduced due to the uncertain wildfire, generation ramping and load shedding are required to recover the power flow balance. For both preventive and corrective control, when the number of lines switched is increased, the load shedding costs are reduced. This is because performing topology control can effectively relieve network congestion, and thus switching more lines can lead to reduced load shedding. In comparison, the generation ramping does not change evidently as v increases. In certain cases, the generation ramping even goes up when v increases. For example, under corrective control when $\alpha = 2$ and v changes

Table 4.1: Operating Costs Without Wildfire Events

Costs ($/hr)	$v = 0$	$v = 1$	$v = 2$	$v = 3$
Generation	8653	8515	8438	8286
Ramping	0	0	0	0
Shedding	0	0	0	0
Total	8653	8515	8438	8286

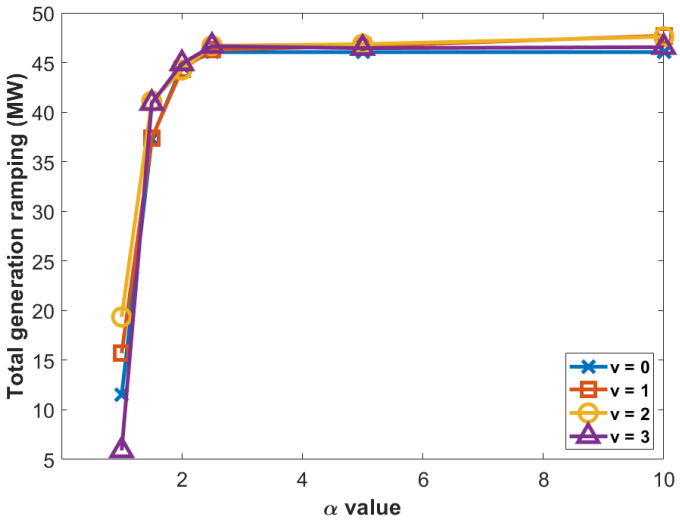
from 2 to 3, the generation ramping cost goes from 866 to 907. However, both generation and load shedding costs are reduced in this case, thus leading to a reduced total cost. Overall, under same α and v values, the total costs from corrective control are lower than that from preventive control. Similarly, the load shedding costs from corrective control are in general lower. Nevertheless, this is not always the case. For example, when $\alpha = 10, v = 3$, the corrective control gives a slightly higher shedding cost. This is because the corrective control find a topology control strategy that leads to much lower generation and ramping costs than the preventive control, which compensates for the increase in load shedding. Although with a higher load shedding cost, the corrective control still provides a lower total cost because of the reduction on generation and ramping. In fact, under certain system congestion patterns, it is also possible that the corrective control provides a higher total cost due to the convex relaxation.

TABLE 4.2 shows the costs for generation, ramping and shedding, which makes it easier for us to compare different approaches under the same α value. However, to better understand the effects of parameter α on the topology control problems, it is crucial to also track the actual change (in MW) of

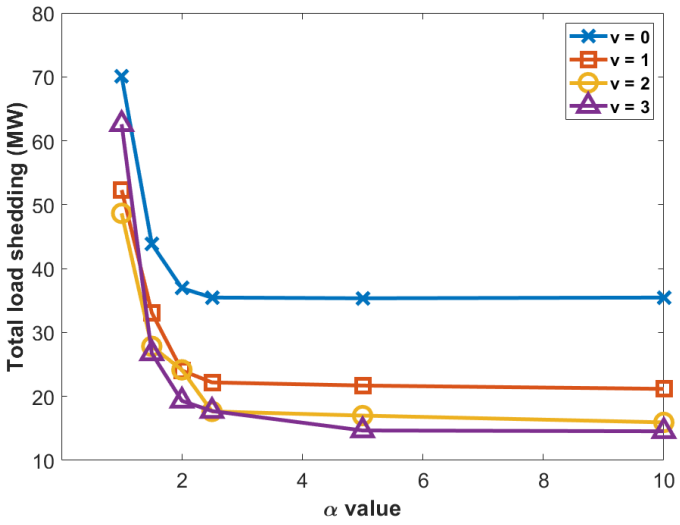
Table 4.2: Comparison of Operating Costs Under Different α Values for Preventive Control and Corrective Control

		Preventive Control				Corrective Control			
	Costs ($/hr)	$v = 0$	$v = 1$	$v = 2$	$v = 3$	$v = 0$	$v = 1$	$v = 2$	$v = 3$
$\alpha = 2$	Generation	8758	8769	8583	8586	8758	8523	8511	8442
	Ramping	893	887	885	898	893	751	866	907
	Shedding	1477	962	966	772	1477	1215	826	720
	Total	11129	10618	10434	10257	11129	10491	10204	10070
$\alpha = 5$	Generation	8790	8838	8856	8855	8790	8623	8617	8551
	Ramping	921	933	937	929	921	888	960	969
	Shedding	3545	2169	1699	1466	3545	2404	1605	1325
	Total	13257	11941	11493	11251	13257	11916	11183	10846
$\alpha = 10$	Generation	8790	8897	8986	8866	8790	8897	8620	8516
	Ramping	921	954	952	931	921	954	968	886
	Shedding	7091	4234	3186	2907	7091	4234	3129	3086
	Total	16803	14085	13125	12705	16803	14085	12718	12489

generation ramping and load shedding under varying α values. To this end, we have used a expanding set of α values to solve the optimization problems and recorded the solutions of total generation ramping and load shedding, as illustrated in Fig. 4.4. When the α value is small (e.g., $\alpha = 1$), the generation ramping is not fully utilized since it is equally expensive as the load shedding in the objective function. As α value gradually increases, the load shedding starts to dominate in the objective cost function. As a result, the total load shedding starts to decrease noticeably, at the expense of increased generation ramping. Under the same switching limit v, the decrease/increase patterns are very similar for load shedding/generation ramping. The changes are more evident when α is between 1 and 3, and starts to slow down after that. When α exceeds a certain threshold (e.g., $\alpha > 5$), the change is marginal and becomes barely noticeable. Under the same v, once α reaches the threshold the total load shedding does not reduce substantially. This is because the system already reaches its limit to minimize the load shedding, and the remaining load



(a)



(b)

Figure 4.4: Comparisons of total (a) generation ramping and (b) load shedding under different α values.

shedding (e.g., ≈ 35.5 MW under $v = 0$) is unavoidable under the contingency due to system physical constraints. Interestingly, a larger switching limit v generally leads to a reduced total load shedding, as shown in Fig. 4.4b. This is mainly because the topology control can optimally alter the feasible region for the resulting optimal power flow (OPF) problem, and thus effectively relieve network congestion. This also confirms the capability of performing topology control for effective wildfire mitigation.

In summary, the proposed topology control approaches can effectively reduce load shedding and save operating costs under line outages from uncertain wildfire. In particular, the preventive control problem can be implemented using efficient decomposition technique, which can provide the *exact solutions* quickly in advance to the line outage contingency. On the other hand, the corrective control implemented via convex relaxation offers an alternative to topology control in real time, which can provide more on-target mitigation solutions.

This work mainly considers the optimal topology control problem under line failures due to uncertain wildfires. To account for the uncertainty model of developing wildfires into the decision-making framework, we put forth a two-stage stochastic mixed-integer program. The optimization problem incorporates different scenarios of line outages due to the spreading wildfires, and seeks to obtain an optimal topology control strategy that provides the lowest operating costs. To enable line switching at different stages, we present both preventive and corrective control formulation, both of can be efficiently solv-

able using decomposition and relaxation based techniques. Numerical studies on the practical RTS-GMLC system have corroborated the performance of the proposed topology control algorithms for providing real-time wildfire mitigation solutions.

Chapter 5

Scalable Learning for Optimal Load Shedding

This chapter focuses on the scalable learning problem for performing optimal load shedding (OLS) under extreme events. We start by introducing the DC OLS problem in Section 5.1. After that, we further present the nonlinear AC OLS problem in Section 5.2. We propose a decentralized learning framework strategy for the OLS problem in Section 5.3. Lastly, numerical simulation results are provided in Section 5.4 to demonstrate the performance of the proposed learning algorithm.

Fast mitigation of power imbalance and operational limit violations during emergency events is of great importance for enhancing the resilience of power grids. To prevent potential cascading failures, load shedding is a commonly used emergency response action by adjusting the system operating point. Unlike normal operations, the decision making of load shedding under severe contingencies is timing-critical. Due to the computational com-

This chapter is based on the following publication to which the coauthors contributed equally:

Yuqi Zhou, Jeehyun Park, and Hao Zhu. "Scalable learning for optimal load shedding under power grid emergency operations." In IEEE Power & Energy Society General Meeting (PESGM), 2022.

plexity and communication latency concerns arising in attaining load shedding solutions, machine learning (ML) methods are uniquely positioned to enable timely emergency grid services thanks to their superior performance in real-time prediction.

To quickly restore power balance, traditional load shedding schemes perform a uniform percentage of reduction for all load centers based on the electric frequency deviation [67]. This proportional reduction is easy to implement, but fails to account for the heterogeneous effects of the contingency scenario across the grid such as severe congestion at certain locations. Recent advances in optimization-based load shedding schemes (e.g., [64,160,211]) can effectively mitigate the potential risks of cascading failures through strategically targeting locational congestion or stability concerns. Nevertheless, to solve the resultant optimal load shedding (OLS) problem in real time can be challenged by the underlying nonlinearity of AC power flow model. Recently, there is a surge of interest in adopting ML methods for power system decision making under normal conditions, particularly for the AC optimal power flow (AC-OPF) problem; see e.g., [22,77,112]. Similarly, to accelerate the AC-OLS solution, [142] has proposed to learn the percentage ratio of load shedding from the system-wide contingency information. In addition, a safe reinforcement learning approach has been developed in [251] to predict the dynamic load shedding policy from the overall system state, thus requiring the grid-wide information.

Although a centralized learning framework is suitable for normal grid

conditions, it fails to promptly react to the contingency situation due to the associated communication and response times. Therefore, existing ML-based solutions cannot cope with the real-time OLS needs where it is critical to implement timely corrective actions at distributed load centers.

This work [292] aims to develop a scalable learning-for-OLS framework such that individual load centers can predict their own optimal decisions in a decentralized fashion. To this end, we first formulate the OLS problem under the AC power flow model, as an extension to the AC-OPF. By determining the respective amount of load shedding at each bus, the AC-OLS problem aims to restore the system-wide power balance and mitigate the violations of operational limits. Upon solving this problem under a wide range of loading conditions and contingency scenarios, one can learn the decision rules from input system conditions to target load shedding actions through offline training. In order to attain high accuracy, we adopt the neural network (NN) model to construct nonlinear, expressive mappings from real-time measurements. One notable feature of our proposed approach is the *decentralized* design of the NN-based decision rules, which have been constructed solely based on *locally available measurements* at each load center. Thanks to the scalability of the decentralized information sharing, the offline training time to obtain individual decision rules is greatly reduced compared to a full-feedback one. During online emergency operations, these decisions rules further enable each load center to promptly react to the specific contingency from readily available local data.

5.1 DC Optimal Load Shedding Problem

We first formulate the optimal load shedding (OLS) problem under the linear dc power flow:

$$\min \quad \boldsymbol{c}^g(\boldsymbol{p}^g) + \boldsymbol{c}^s(\boldsymbol{p}^s) \tag{5.1a}$$

$$\text{s.t.} \quad \boldsymbol{p}^g \in \mathbb{R}^N, \boldsymbol{p}^s \in \mathbb{R}^N \tag{5.1b}$$

$$\underline{\boldsymbol{p}^g} \leq \boldsymbol{p}^g \leq \overline{\boldsymbol{p}^g} \tag{5.1c}$$

$$\underline{\boldsymbol{p}^s} \leq \boldsymbol{p}^s \leq \overline{\boldsymbol{p}^s} \tag{5.1d}$$

$$\underline{\boldsymbol{f}} \leq \boldsymbol{S}(\boldsymbol{p}^g - \boldsymbol{p}^d + \boldsymbol{p}^s) \leq \overline{\boldsymbol{f}} \quad (\underline{\boldsymbol{\mu}}, \overline{\boldsymbol{\mu}}) \tag{5.1e}$$

$$\mathbf{1}^\mathsf{T}\boldsymbol{p}^g = \mathbf{1}^\mathsf{T}\boldsymbol{p}^d - \mathbf{1}^\mathsf{T}\boldsymbol{p}^s \quad (\lambda) \tag{5.1f}$$

where $\boldsymbol{p}^g, \boldsymbol{p}^d, \boldsymbol{p}^s$ denote generation, load and load shedding, respectively. The DC optimal load shedding (DC-OLS) problem aims to minimize the total costs of generation and load shedding, as given by $\boldsymbol{c}^g(\boldsymbol{p}^g) + \boldsymbol{c}^s(\boldsymbol{p}^s)$. For each bus, the cost functions are designed such that $|c_i^g(p_i^g)| \ll |c_i^s(p_i^s)|$. Thus, the OLS solutions will prefer to fully utilize the generation resources before evoking load shedding. Constraints (5.1c) and (5.1d) enforce the limits for generation and load shedding. The line flow limits are enforced in (5.1e), and the system power balance is guaranteed in (5.1f).

5.1.1 Optimality Analysis on the DC OLS Problem

We use $\underline{\boldsymbol{\mu}}$ and $\overline{\boldsymbol{\mu}}$ to denote the dual variables for line flow constraints (5.1e), and λ to denote the dual variables for the power balance constraint (5.1f). To this end, the Lagrangian function for the DC-OLS problem can be

given as:

$$\mathcal{L} = \boldsymbol{c}^g(\boldsymbol{p}^g)+\boldsymbol{c}^s(\boldsymbol{p}^s) + (\underline{\boldsymbol{f}} - \boldsymbol{S}(\boldsymbol{p}^g - \boldsymbol{p}^d + \boldsymbol{p}^s))^\mathsf{T}\underline{\boldsymbol{\mu}} + (\boldsymbol{S}(\boldsymbol{p}^g - \boldsymbol{p}^d + \boldsymbol{p}^s) - \overline{\boldsymbol{f}})^\mathsf{T}\overline{\boldsymbol{\mu}} - \lambda(\mathbf{1}^\mathsf{T}\boldsymbol{p}^g - \mathbf{1}^\mathsf{T}\boldsymbol{p}^d + \mathbf{1}^\mathsf{T}\boldsymbol{p}^s) \quad (5.2)$$

According to the Lagrangian optimality condition, in general we have $\boldsymbol{p}$ (including $\boldsymbol{p}^g$ and $\boldsymbol{p}^s$) that satisfy the following:

$$\boldsymbol{p}^* = \underset{\underline{\boldsymbol{p}}\leq\boldsymbol{p}\leq\overline{\boldsymbol{p}}}{\arg\min}\ \boldsymbol{c}(\boldsymbol{p}) + \boldsymbol{S}^\mathsf{T}(\overline{\boldsymbol{\mu}}^* - \underline{\boldsymbol{\mu}}^*)\boldsymbol{p} - \lambda^*\mathbf{1}^\mathsf{T}\boldsymbol{p} = \underset{\underline{\boldsymbol{p}}\leq\boldsymbol{p}\leq\overline{\boldsymbol{p}}}{\arg\min}\ \boldsymbol{c}(\boldsymbol{p}) - \boldsymbol{\pi}^{*\mathsf{T}}\boldsymbol{p} \quad (5.3)$$

where $\overline{\boldsymbol{\mu}}^*, \underline{\boldsymbol{\mu}}^*$ and λ^* are the optimal dual variables. Furthermore, $\boldsymbol{\pi}^* = \boldsymbol{S}^\mathsf{T}(\overline{\boldsymbol{\mu}}^* - \underline{\boldsymbol{\mu}}^*) - \lambda^*\mathbf{1}^\mathsf{T}$ is known as the locational marginal price (LMP). Considering a quadratic cost for each bus, then the optimality condition for every single bus can be given as:

$$p_i^* = \underset{\underline{p_i}\leq p_i\leq\overline{p_i}}{\arg\min}\ c_i(p_i) - \pi_i{}^* p_i \quad (5.4)$$

Therefore, given the optimal solution of LMP $\pi_i{}^*$, the local optimal solution of p_i^* can be uniquely determined (assuming a quadratic cost function). This one-to-one mapping motivates us to consider learning local LMP before obtaining the optimal solutions of $\boldsymbol{p}^g$ and $\boldsymbol{p}^s$. The rationale behind it is that the LMP contains more information on the line congestion, and thus can help attain more accurate generation and load shedding solutions.

5.2 AC Optimal Load Shedding Problem

We next formulate the optimal load shedding (OLS) problem based on the nonlinear AC power flow. Consider a power grid with N buses and L

lines collected in the sets $\mathcal{N}$ and $\mathcal{L}$, respectively. Let $\boldsymbol{Y} = \boldsymbol{G} + j\boldsymbol{B}$ denote the network admittance matrix, where $\boldsymbol{G}$ and $\boldsymbol{B}$ are real and imaginary parts, respectively. For each bus i, the complex power output from its connected generation is denoted by $p_i^g + jq_i^g$, while load demand by $p_i^d + jq_i^d$. In addition, let V_i, θ_i denote the bus voltage magnitude and angle, respectively. Thus, the voltage phasor for each bus can be represented in polar form as $V_i \angle \theta_i$. Under the AC power flow, the complex power flow for line $(i, j) \in \mathcal{L}$ can be represented as [66]:

$$S_{ij} = Y_{ij}^* V_i^2 - Y_{ij}^* V_i V_j \angle(\theta_i - \theta_j) \tag{5.5}$$

where $(\cdot)^*$ denotes the conjugate of a complex number.

The AC-OLS problem is formulated similarly to AC optimal power flow (AC-OPF). The latter problem determines the optimal set-points under normal operations that account for system constraints on generation output, voltage, and line flows. Under emergency operations as a result of large-scale contingencies, the AC-OPF problem could become infeasible due to insufficient resources or transfer capability. Thus, corrective actions such as load shedding are typically carried out to maintain power balance and satisfy the system operation limits. With known load demand p_i^d and q_i^d per bus i, the OLS problem aims to determine the amount of load shedding denoted by p_i^s and q_i^s, as given by

$$\min \quad \sum_{i=1}^{N} c_i^g(p_i^g) + c_i^s(p_i^s) \tag{5.6a}$$

$$\text{s.t.} \quad p_i^g \in \mathbb{R}, q_i^g \in \mathbb{R}, V_i \in \mathbb{R}^+, \theta_i \in \mathbb{R}, \forall i \in \mathcal{N} \tag{5.6b}$$

$$p_i^s \in \mathbb{R}, q_i^s \in \mathbb{R}, \forall i \in \mathcal{N} \tag{5.6c}$$

$$\underline{p}_i^g \leq p_i^g \leq \overline{p}_i^g, \ \underline{q}_i^g \leq q_i^g \leq \overline{q}_i^g \tag{5.6d}$$

$$\underline{V}_i \leq V_i \leq \overline{V}_i, \ \underline{\theta}_i \leq \theta_i \leq \overline{\theta}_i \tag{5.6e}$$

$$0 \leq p_i^s \leq p_i^d, \ 0 \leq q_i^s \leq q_i^d \tag{5.6f}$$

$$|S_{ij}| \leq \overline{S}_{ij} \tag{5.6g}$$

$$\theta_{ij} = \theta_i - \theta_j \tag{5.6h}$$

$$p_i^g - p_i^d + p_i^s = \textstyle\sum_{j=1}^N V_i V_j \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}\right) \tag{5.6i}$$

$$q_i^g - q_i^d + q_i^s = \textstyle\sum_{j=1}^N V_i V_j \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}\right). \tag{5.6j}$$

The objective function in (5.6a) consists of the generation cost $c_i^g(\cdot)$ and load reduction cost $c_i^s(\cdot)$, which are typically (piece-wise) linear or convex quadratic functions. One can also extend the objective to incorporate the cost of reactive power reduction for each load. In general, the two cost functions c_i^g and c_i^s are designed such that for every bus $i \in \mathcal{N}$ we have $|c_i^g(p_i^g)| \ll |c_i^s(p_i^s)|$. Similar to the DC-OLS problem, the solutions will prefer to make use of available generation resources before triggering load shedding. The load shedding cost can vary from load centers by prioritizing critical loads with much higher costs than the others. Moreover, for the decision variables given in (5.6b) - (5.6c), constraints (5.6d) - (5.6f) enforce the upper/lower bounds for each of them based upon the resource budget or operational limits. The upper bounds for load shedding in (5.6f) can be lower than the total demand in case part of the load center constitutes as non-dispatchable critical loads.

Furthermore, the constraint (5.6g) enforces the line thermal limit for apparent power flow, while other line limits (e.g, line current, real power flow) can be similarly posed. Finally, the equality constraints (5.6h) - (5.6j) correspond to AC power flow equations.

The AC-OLS problem differs from the AC-OPF one mainly in the level of flexibility that each load center can provide. During the contingency conditions, corrective actions such as reducing load demand or reconnecting transmission lines are imperative in order to restore the power balance and increase the power transfer capability. It is worth pointing out that even though the OLS problem (5.6) is formulated to consider load shedding only, it can be generalized to include other corrective actions such as topology optimization and full de-energization of system components; see e.g., [64, 211].

Remark 5.2.1 (Solving AC-OLS). As a nonlinear program (NLP), the AC-OLS problem (5.6) is nonconvex and generally NP-hard, similar to the AC-OPF [53]. Various convex relaxation methods can be adopted to tackle the nonconvexity therein for OPF [144, 168] and similarly for OLS [64]. In general, open-source packages such as MATPOWER [303] and JuMP [63] are available to efficiently solve the OLS problem.

5.3 Learning the Scalable OLS Strategy

Although the centralized OLS problem can be solved by various optimization solvers, its implementation requires high rate of communications for the control center to acquire the system-wide information and dispatch

the emergency actions. Due to communication latency and quality issues, this centralized framework could affect the timeliness and effectiveness of corrective action responses at individual load centers, both critical for grid emergency operations. To enable fast and powerful load shedding actions during emergency events, we propose to develop a scalable OLS strategy by predicting the OLS decision using *locally available* measurements in real time. The key idea of the proposed framework is illustrated in Fig. 5.1, where each load center such as the one in bus 9 can directly form its own OLS decisions using the voltage phasor and power data collected by local meters. Under the physics-based power flow coupling, each contingency scenario leads to various level of changes in the available measurements at every load center. Thus, the latter can be used to infer the OLS solution for the specific contingency even without centralized information exchange.

5.3.1 Neural Network Design

Of course, the question becomes how to obtain such decentralized decision rules without using system-wide information. To this end, we utilize the basic feedforward neural network to obtain the decision rule $f_i(\boldsymbol{x}_i; \boldsymbol{\varphi}_i) \to \boldsymbol{y}_i$ for each load center i, that maps from local measurements $\boldsymbol{x}_i$ to its optimal decisions $\boldsymbol{y}_i$. Note that $\boldsymbol{\phi}_i$ denotes the NN parameters that will be specified later and learned during the training process. As the goal is to establish the decision rule $f_i(\cdot; \boldsymbol{\varphi}_i)$ for any possible system operating condition or contingency scenario, the offline training process builds upon generating a high number

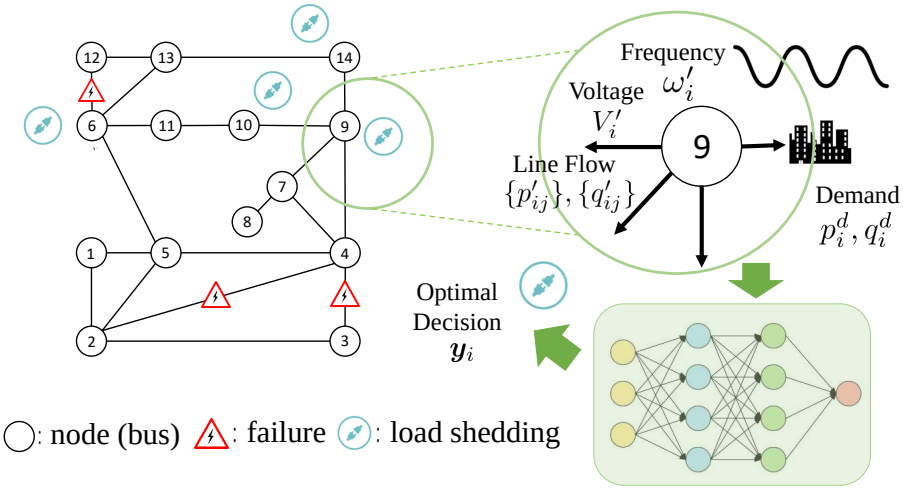


Figure 5.1: Illustration of scalable load shedding design on the IEEE 14-bus system.

of instances, each representing a specific loading condition and contingency scenario. As illustrated in Fig. 5.2, for each instance, the corresponding input feature $\boldsymbol{x}_i$ and target decision $\boldsymbol{y}_i$ can be respectively computed by solving the power flow and OLS problem (5.6). All these samples of $\{\boldsymbol{x}_i, \boldsymbol{y}_i\}$ will be used to train the decision rule f_i by learning its parameters $\boldsymbol{\varphi}_i$. When using f_i for online implementation, each control center can immediately use real-time local data $\hat{\boldsymbol{x}}_i$ to quickly obtain the decision as $\hat{\boldsymbol{y}}_i = f_i(\hat{\boldsymbol{x}}_i; \varphi_i)$. This constitutes the overall architecture of the proposed scalable OLS design, which leverages extensive offline computation and learning to empower the online decision-making process.

5.3.2 Scalable Learning of OLS Decisions

For accurate OLS prediction, the local input feature $\boldsymbol{x}_i$ should include all possible real-time measurements, such as

$$\boldsymbol{x}_i = \left[p_i^d, q_i^d, V_i', \{p_{ij}'\}, \{q_{ij}'\}, \omega_i'\right], \tag{5.7}$$

which represents local real/reactive power demand information, post-contingency voltage magnitude, all incident line flows, as well as the electric frequency. Note that symbols with $'$ here correspond to post-contingency values and differ from those in (5.6). Most of these values can be computed by open-source solvers such as MATPOWER through steady-state power flow simulations, except for the electric frequency ω_i'. As the latter is a very informative indicator of the overall power imbalance, one can approximate it using the difference between pre- and post-contingency steady-state power generation [103, Ch. 12]. In future, we plan to utilize dynamic power flow simulations such as the COSMIC tool [235] to improve the post-contingency modeling. As for the target decisions, we are primarily interested in predicting the load reduction solutions from (5.6), as given by:

$$\boldsymbol{y}_i = [p_i^s, q_i^s]. \tag{5.8}$$

General corrective decisions can be predicted as well by extending AC-OLS problem as mentioned earlier.

The proposed decentralized OLS design builds upon the strong correlation between load shedding decisions and local post-contingency data, as

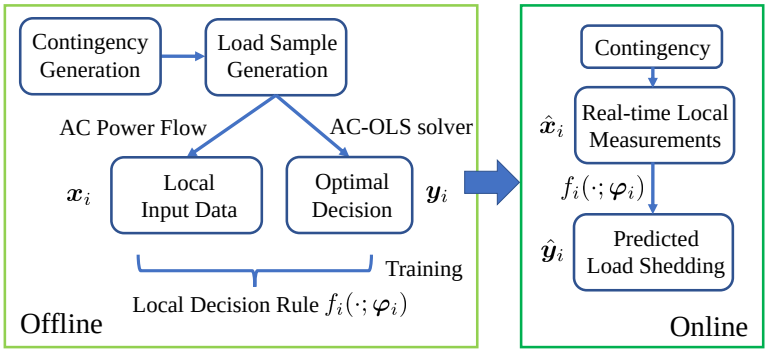


Figure 5.2: The overall architecture for the proposed scalable load shedding design with extensive offline studies and training to accelerate the online decision making.

observed in numerical tests later on. As the measurements in $\boldsymbol{x}_i$ can effectively reveal the effects of contingency at bus i, they are highly indicative of the corresponding optimal decision. For example, both V_i' and ω_i' are great stress indicator on the system loading conditions. Similarly, the line flows $\{p_{ij}', q_{ij}'\}$ indicate the change of power flow patterns due to contingency.

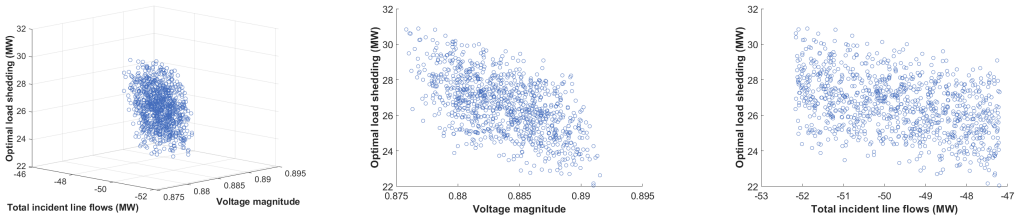
To obtain the OLS decision rule $f_i(\cdot;\boldsymbol{\varphi}_i)$ for load center i, the NN model consists of multiple fully-connected hidden layers between input $\boldsymbol{x}_i$ and output $\boldsymbol{y}_i$. With the first layer $\boldsymbol{z}_i^0 = \boldsymbol{x}_i$ incorporating the input feature, each layer k can be represented as:

$$\boldsymbol{z}_i^k = \sigma\left(\boldsymbol{W}^k \boldsymbol{z}_i^{k-1} + \boldsymbol{b}^k\right), \quad \forall k = 1, \ldots, K \tag{5.9}$$

where the final layer $\boldsymbol{z}_i^K \to \boldsymbol{y}_i$ predicts the output target. Thus, the NN parameters in $\boldsymbol{\varphi}_i$ include the weight matrices $\{\boldsymbol{W}^k\}$ and bias vectors $\{\boldsymbol{b}^k\}$ for the linear transformation per layer k. Each layer also uses a nonlinear activa-

tion function $\sigma(\cdot)$ to attain high-dimensional, expressive functional mapping that goes beyond linearity. Common choices of the activation function include sigmoid and ReLU. To determine $\boldsymbol{\varphi}_i$, we use the mean squared error (MSE) metric as the loss function to minimize, solved by popular NN training algorithms such as stochastic gradient descent.

Remark 5.3.1 (Safety of decentralized OLS). As a corrective action, the OLS decisions need to be effective and safe during online implementation. Such considerations can be further incorporated into offline training. For example, a weighted MSE metric could discourage larger prediction error for higher amount of load shedding amount. In addition, we can generalize it to a risk-aware learning framework using the conditional value-at-risk (CVaR) measure to reduce the worst-case prediction error; see e.g., [161].



(a) (b) (c)

Figure 5.3: The optimal load shedding decisions at bus 14 for double-line outages versus (a) both voltage magnitude V_i' and total incident line flows $\sum_i p_{ij}'$; (b) voltage magnitude V_i' only; and (c) total incident line flows $\sum_i p_{ij}'$ only.

5.4 Numerical Validations

This section presents numerical test results of the proposed decentralized OLS approach on the IEEE 14-bus system. The AC-OLS problem has been implemented using MATPOWER and solved by the primal-dual interior point method. Quadratic objective functions of $c_i^g(\cdot)$ and $c_i^s(\cdot)$ have been used for the cost of generation and shedding loads, respectively. The feedforward NN has been implemented with MATLAB® deep learning toolbox using the Bayesian regularization algorithm. The simulations are performed on a regular laptop with Intel® CPU @ 2.60 GHz and 16 GB of RAM.

5.4.1 IEEE 14-Bus System Tests

The IEEE 14-bus system is shown as in Fig. 5.1. It consists of 20 transmission lines and 5 conventional generators located at buses 1, 2, 3, 6 and 8. Given the initial load condition, the load shedding may be required more often at certain buses than the others. To this end, we mainly study the load centers located at buses 6, 9, 10, 11, 13 and 14, respectively. We consider line outage failures as the initial emergency events, while the proposed method is generalizable to other types of system contingencies as well. We have generated all $(N-1)$ contingency scenarios (single), and randomly selected $(N-2)$ and $(N-3)$ contingency scenarios (multiple). For simplicity, the scenarios that lead to system islanding are excluded here and will be studied in future. To encourage the occurrence of system emergency operations, we increase the original system loading to a total of 469 MW, under which the

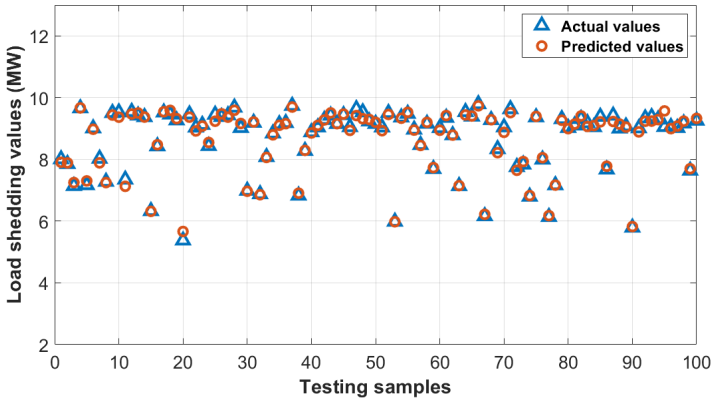
Table 5.1: RMSE Values Under Single Line Outage

Bus	Occurrence	Training [MW]	Testing [MW]
6	3.6%	0.1844	0.6091
9	5.3%	0.0483	0.0499
10	12.3%	0.2559	0.5707
13	5.3%	0.0402	0.0433
14	98.7%	0.4085	0.4253

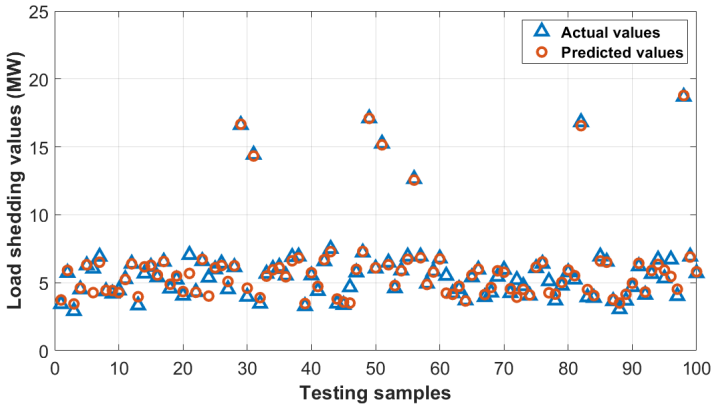
AC-OPF solution is closer to the infeasibility margin and load shedding is more likely to incur during contingencies. In addition to contingency sampling, we also randomly generate the load demand per bus to be $[95\%, 105\%]$ of its nominal value, to reflect its small variation in minute time-frame. For each contingency scenario, we generated a total of 1000 samples, a majority of which have experienced the occurrence of AC-OLS due to the stress of emergency conditions.

To demonstrate the correlation between local measurements and OLS decisions, Fig. 5.3 plots their relations at bus 14 under the outage of both line 2-3 and line 4-9. The cross-section scatter plots show that the optimal shedding amount increases as the voltage magnitude V_i' or total incident line flow decreases. This is because low voltage indicates system stress, while reduced incident line flow implies a change of power flow pattern, both calling for the need of load shedding. This observation supports to use local data to form OLS decisions. Note that for ease of exposition, only the predicted real power reduction amount will be presented.

The training is carried out using a feedforward NN with two hidden



(a) Load center at bus 10



(b) Load center at bus 14

Figure 5.4: Comparison of predicted and actual decision values for randomly selected samples under single line outage testing.

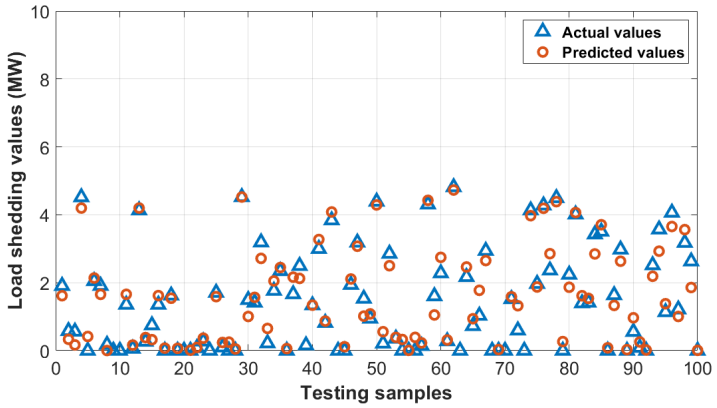
layers, each of which has 15 and 12 neurons, respectively. We split the data samples randomly into 80% and 20% for training and testing, separately performed on single and multiple line outages. On average, the offline training process takes around 16.5 and 38.7 seconds for single and multiple line outages, respectively. As the number of local inputs mainly depends on the incident topology of each load center, the scalability of the training process can be

Table 5.2: RMSE Values Under Multiple Line Outages

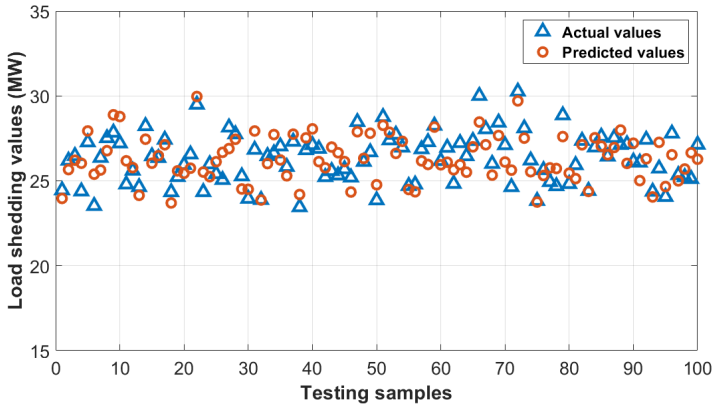
Bus	Occurrence	Training [MW]	Testing [MW]
6	8.0%	0.0957	0.4716
9	8.0%	0.0454	0.0490
10	52.7%	0.5269	0.8282
11	2.4%	0.0689	0.1518
13	11.5%	0.8140	1.4852
14	99.4%	0.8517	0.9636

guaranteed under the decentralized design. Table 5.1 lists the occurrence and prediction error of load shedding at the 5 load buses under single line outage scenarios. Note that bus 11 is not included as load shedding did not happen at this location. As OLS does not occur all the time per bus, we list the root mean square error (RMSE), given by $\sqrt{\frac{1}{S}\sum_{s=1}^{S}(p_i^s - \hat{p}_i^s)^2}$, for the samples where OLS has occurred. Fig. 5.4 also compares the predicted and actual load shedding values for selected testing samples at buses 10 and 14, both very likely to need load shedding during contingencies. Note that certain buses such as bus 6 have experienced noticeably higher testing error than the training one, and we will address this issue through regularization in future. Overall, the local predictions well match the actual values of AC-OLS solutions and attain satisfactory performance.

The performance of the proposed design for multiple line outage scenarios is presented in Table 5.2 and Fig. 5.5. As multiple line outages would increase the stress to the system, the occurrence of load shedding has increased with bus 11 experiencing load shedding in certain cases. By and large, the prediction performance remains good. However, the accuracy has reduced



(a) Load center at bus 10



(b) Load center at bus 14

Figure 5.5: Comparison of predicted and actual decision values for randomly selected samples under multiple line outage testing.

for certain buses compared to the single line outage results, mainly due to the vast variability of post-contingency conditions. Thus, while the simulation results have confirmed the validity of the proposed approach, we need to expand the variability of the training samples to enhance the expressiveness of the resultant NN models.

This work developed a decentralized framework for performing real-

time load shedding in order to prevent cascading propagation under emergency events. By solving the AC-OLS optimization problem for a multitude of contingency conditions, we put forth a learning-for-OLS framework that maps from each load center's local measurements to its own OLS decision. Clearly, this scalable design of OLS decision rules enables load centers to quickly react to the contingency situations without requiring the supervision from control center. Numerical results demonstrate the validity of the proposed design in terms of predicting the OLS solutions. For future work, we will extensively investigate the proposed design for different contingency conditions and types of systems, as well as improve the safety using risk-aware learning approaches.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Efficiently monitoring and controlling the grid topology are of great importance for maintaining secured power system operations. This work first establishes the modeling and monitoring framework for the bus split event, which is a grid topology change caused by the switching of electrical circuit breakers. From the control perspective, we examine the optimal transmission switching problem under uncertainty, which utilizes the flexibility of network topology to deliver more cost-effective power flow solutions.

Chapter 2 presents the sensitivity analysis framework for modeling the grid-wide impact of bus split events in order to enhance the security of power system operations. Using the popular bus-branch representation, we next perform the post-split system analysis via efficiently updating the changes in the system, which turns out to be equivalent to a line outage event followed by an additional pair of power transfer. This equivalent model allows to conveniently incorporate the bus split contingencies into the security-constrained economic dispatch problem to obtain more secure generation dispatch solutions. After that, we develop an efficient monitoring framework for power system bus split

events due to substation connectivity changes. The synchrophasor data enabled monitoring of bus split events was cast as an optimization problem, with binary variables indicating the connectivity of lines, generation, loads within a substation. To tackle the bilinear relations in the resultant problem, the McCormick relaxation technique has been leveraged to attain an *equivalent* MILP reformulation that is efficiently solvable for real-time identification.

Chapter 3 presents a substation-level network topology optimization formulation that can incorporate both line switching and bus splitting operations. To deal with the bilinearity in the nonlinear formulation, the McCormick relaxation has been utilized to devise a tractable MILP reformulation, which can be efficiently solved for real-time applications. Furthermore, to consider topology optimization under uncertain renewables, we consider the distributionally robust chance-constrained (DRCC) OTS that can ensure the guarantees over an ambiguity set of uncertainty distributions. Both moment-based and distance-based ambiguity sets have been considered, leading to scalable MILP problems through dualization. Between the two proposed DRCC-OTS approaches, the one using the mean-MAD ambiguity set incurred lower computation complexity on par with other scenario-free approaches, while the one using the Wasserstein ambiguity led to less conservative solutions by adapting to the actual data samples.

Chapter 4 considers the optimal topology control problem under line failures due to uncertain wildfires. To account for the uncertainty model of developing wildfires into the decision-making framework, we put forth a two-

stage stochastic mixed-integer program. The optimization problem incorporates different scenarios of line outages due to the spreading wildfires, and seeks to obtain an optimal topology control strategy that provides the lowest operating costs. To enable line switching at different stages, we present both preventive and corrective control formulation, both of can be efficiently solvable using decomposition and relaxation based techniques. Numerical studies on the practical RTS-GMLC system have corroborated the performance of the proposed topology control algorithms for providing real-time wildfire mitigation solutions.

In Chapter 5, we have developed a decentralized framework for performing real-time load shedding in order to prevent cascading propagation under emergency events. By solving the AC-OLS optimization problem for a multitude of contingency conditions, we put forth a learning-for-OLS framework that maps from each load center's local measurements to its own OLS decision. Clearly, this scalable design of OLS decision rules enables load centers to quickly react to the contingency situations without requiring the supervision from control center. Numerical results demonstrate the validity of the proposed design in terms of predicting the OLS solutions.

6.2 Future Work

The power system control problem under wildfires in Chapter 4 adopts a stochastic program formulation. However, given the fast evolving wildfires, the probability of each uncertainty realization can be difficult to obtain accu-

rately in real time. The existing preventive and corrective formulation only consider the ideal situation, but the optimal mitigation solutions may be rendered ineffective under actual wildfires. Therefore, for future work we would like to explore the robust optimization formulation for this problem.

Furthermore, the scalable learning of OLS problem in Chapter 5 currently considers a limited number of contingency scenarios for training. However, whether the OLS decision rules can remain effective when a new contingency scenario occurs is still unclear. Therefore, future work will address the generalizability and feasibility concerns of the scalable learning design.

Appendices

Appendix A

Proof of Proposition (3.3.5)

Here we will present the detailed steps to derive the equivalent reformulation from (3.31) to (3.32) for the problem (3.33). Using the definition of the indicator function $\mathbb{1}(\cdot)$, the semi-infinite constraint in problem (3.31) boils down to the following two cases:

$$\boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\xi} - \boldsymbol{\kappa}^{\mathsf{T}}|\boldsymbol{\xi} - \boldsymbol{\mu}| \leq 1, \quad \forall \boldsymbol{\xi} \tag{A.1a}$$

$$\boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\xi} - \boldsymbol{\kappa}^{\mathsf{T}}|\boldsymbol{\xi} - \boldsymbol{\mu}| \leq 0, \quad \forall \boldsymbol{\xi} : \mathbf{a}_i(\mathbf{x})^{\mathsf{T}}\boldsymbol{\xi} > b_i(\mathbf{x}) \tag{A.1b}$$

Specifically, the right hand side equals to 0 for any $\boldsymbol{\xi}$ such that $\mathbf{a}_i(\mathbf{x})^{\mathsf{T}}\boldsymbol{\xi} > b_i(\mathbf{x})$, or 1 for any other choice of $\boldsymbol{\xi}$. These two cases can be reformulated using standard convex duality theory [118]. Specifically, (A.1a) is equivalent to the following:

$$\sup \quad \boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\xi} - \boldsymbol{\kappa}^{\mathsf{T}}\boldsymbol{\rho} \leq 1 \tag{A.2a}$$

$$\text{s.t.} \quad \boldsymbol{\xi} \in \mathbb{R}^K \tag{A.2b}$$

$$\boldsymbol{\rho} \geq \boldsymbol{\xi} - \boldsymbol{\mu} \qquad (\boldsymbol{\pi}_1) \tag{A.2c}$$

$$\boldsymbol{\rho} \geq \boldsymbol{\mu} - \boldsymbol{\xi} \qquad (\boldsymbol{\tau}_1) \tag{A.2d}$$

Dualizing this optimization problem implies that there exists non-negative dual variables $\boldsymbol{\pi}_1 \in \mathbb{R}_+^K, \boldsymbol{\tau}_1 \in \mathbb{R}_+^K$ such that

$$\begin{aligned}\boldsymbol{\alpha} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\rho} + \boldsymbol{\pi}_1^\mathsf{T}(\boldsymbol{\rho} + \boldsymbol{\mu}) + \boldsymbol{\tau}_1^\mathsf{T}(\boldsymbol{\rho} - \boldsymbol{\mu}) - 1 \\ \leq \min_{\mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}} (-\boldsymbol{\beta}^\mathsf{T} + \boldsymbol{\pi}_1^\mathsf{T} - \boldsymbol{\tau}_1^\mathsf{T})\boldsymbol{\xi}\end{aligned} \tag{A.3}$$

We can dualize the right hand side again using the uncertainty support, and it leads to the following equivalent constraints:

$$\alpha + (\boldsymbol{\pi}_1^\mathsf{T} - \boldsymbol{\tau}_1^\mathsf{T})\boldsymbol{\mu} + \boldsymbol{\psi}_1^\mathsf{T}\mathbf{t} \leq 1 \tag{A.4a}$$

$$\boldsymbol{\beta}^\mathsf{T} + \boldsymbol{\tau}_1^\mathsf{T} = \boldsymbol{\pi}_1^\mathsf{T} + \boldsymbol{\psi}_1^\mathsf{T}\mathbf{U} \tag{A.4b}$$

$$\boldsymbol{\pi}_1^\mathsf{T} + \boldsymbol{\tau}_1^\mathsf{T} = \boldsymbol{\kappa}^\mathsf{T} \tag{A.4c}$$

where $\boldsymbol{\psi}_1 \in \mathbb{R}_+^W$ are introduced as the dual variables for the linear constraints ($\mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}$) for the support set $\boldsymbol{\Xi}$. Similarly, we can derive the equivalent constraints for (A.1b). The constraints (A.1b) are equivalent to the following:

$$\sup \quad \boldsymbol{\alpha} + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\xi} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\rho} \leq 0 \tag{A.5a}$$

$$\text{s.t.} \quad \boldsymbol{\xi} \in \mathbb{R}^K \tag{A.5b}$$

$$\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} > b_i(\mathbf{x}) \qquad (\lambda) \tag{A.5c}$$

$$\boldsymbol{\rho} \geq \boldsymbol{\xi} - \boldsymbol{\mu} \qquad (\boldsymbol{\pi}_2) \tag{A.5d}$$

$$\boldsymbol{\rho} \geq \boldsymbol{\mu} - \boldsymbol{\xi} \qquad (\boldsymbol{\tau}_2) \tag{A.5e}$$

Dualizing it implies that there exists non-negative dual variables $\lambda \in \mathbb{R}_+, \boldsymbol{\pi}_2 \in \mathbb{R}_+^K, \boldsymbol{\tau}_2 \in \mathbb{R}_+^K$ such that

$$\boldsymbol{\alpha} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\rho} + \boldsymbol{\pi}_2^\mathsf{T}(\boldsymbol{\rho} + \boldsymbol{\mu}) + \boldsymbol{\tau}_2^\mathsf{T}(\boldsymbol{\rho} - \boldsymbol{\mu}) - \lambda b_i(\mathbf{x})$$

$$\leq \min_{\mathbf{U}\boldsymbol{\xi}\leq\mathbf{t}} \ (-\boldsymbol{\beta}^\mathsf{T} + \boldsymbol{\pi}_2^\mathsf{T} - \boldsymbol{\tau}_2^\mathsf{T} - \lambda \mathbf{a}_i(\mathbf{x})^\mathsf{T})\boldsymbol{\xi} \tag{A.6}$$

We can dualize the right hand size again, which leads to the following equivalent constraints:

$$\alpha + (\boldsymbol{\pi}_2^\mathsf{T} - \boldsymbol{\tau}_2^\mathsf{T})\boldsymbol{\mu} + \boldsymbol{\psi}_2^\mathsf{T}\mathbf{t} \leq \lambda b_i(\mathbf{x}) \tag{A.7a}$$

$$\boldsymbol{\beta}^\mathsf{T} + \lambda \mathbf{a}_i(\mathbf{x})^\mathsf{T} + \boldsymbol{\tau}_2^\mathsf{T} = \boldsymbol{\pi}_2^\mathsf{T} + \boldsymbol{\psi}_2^\mathsf{T}\mathbf{U} \tag{A.7b}$$

$$\boldsymbol{\pi}_2^\mathsf{T} + \boldsymbol{\tau}_2^\mathsf{T} = \boldsymbol{\kappa}^\mathsf{T} \tag{A.7c}$$

where $\boldsymbol{\psi}_2 \in \mathbb{R}_+^W$ are introduced as the dual variables for the linear constraints $\mathbf{U}\boldsymbol{\xi} \leq \mathbf{t}$. Recall that the objective function (3.31a) also needs to satisfy:

$$\alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\mu} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\sigma} \geq 1 - \epsilon_i \tag{A.8}$$

Therefore, the original problem is equivalent to the following constraints:

$$\alpha + \boldsymbol{\beta}^\mathsf{T}\boldsymbol{\mu} - \boldsymbol{\kappa}^\mathsf{T}\boldsymbol{\sigma} \geq 1 - \epsilon_i \tag{A.9a}$$

$$\alpha + (\boldsymbol{\pi}_1^\mathsf{T} - \boldsymbol{\tau}_1^\mathsf{T})\boldsymbol{\mu} + \boldsymbol{\psi}_1^\mathsf{T}\mathbf{t} \leq 1 \tag{A.9b}$$

$$\boldsymbol{\beta}^\mathsf{T} + \boldsymbol{\tau}_1^\mathsf{T} = \boldsymbol{\pi}_1^\mathsf{T} + \boldsymbol{\psi}_1^\mathsf{T}\mathbf{U} \tag{A.9c}$$

$$\boldsymbol{\pi}_1^\mathsf{T} + \boldsymbol{\tau}_1^\mathsf{T} = \boldsymbol{\kappa}^\mathsf{T} \tag{A.9d}$$

$$\alpha + (\boldsymbol{\pi}_2^\mathsf{T} - \boldsymbol{\tau}_2^\mathsf{T})\boldsymbol{\mu} + \boldsymbol{\psi}_2^\mathsf{T}\mathbf{t} \leq \lambda b_i(\mathbf{x}) \tag{A.9e}$$

$$\boldsymbol{\beta}^\mathsf{T} + \lambda \mathbf{a}_i(\mathbf{x})^\mathsf{T} + \boldsymbol{\tau}_2^\mathsf{T} = \boldsymbol{\pi}_2^\mathsf{T} + \boldsymbol{\psi}_2^\mathsf{T}\mathbf{U} \tag{A.9f}$$

$$\boldsymbol{\pi}_2^\mathsf{T} + \boldsymbol{\tau}_2^\mathsf{T} = \boldsymbol{\kappa}^\mathsf{T} \tag{A.9g}$$

Notice that the dual variable $\lambda > 0$ corresponding to the constraint $\mathbf{a}_i(\mathbf{x})^\mathsf{T}\boldsymbol{\xi} > b_i(\mathbf{x})$ introduces bilinearity in the above formulation, due to $\lambda b_i(\mathbf{x})$ in constraint (A.9e) and $\lambda \mathbf{a}_i(\mathbf{x})^\mathsf{T}$ in constraint (A.9f). To address this, we divide

all the constraints with λ and redefine variables $\alpha' = \frac{\alpha}{\lambda} \in \mathbb{R}$, $\boldsymbol{\beta}' = \frac{\boldsymbol{\beta}}{\lambda} \in \mathbb{R}^K$, $\boldsymbol{\kappa}' = \frac{\boldsymbol{\kappa}}{\lambda} \in \mathbb{R}^K_+$, $\boldsymbol{\pi}' = \frac{\boldsymbol{\pi}}{\lambda} \in \mathbb{R}^K_+$, $\boldsymbol{\tau}' = \frac{\boldsymbol{\tau}}{\lambda} \in \mathbb{R}^K_+$, $\lambda' = \frac{1}{\lambda} \in \mathbb{R}_+$. Eventually, we arrive at the following equivalent *linear* constraints, as in (3.32):

$$\alpha' + \boldsymbol{\beta}'^{\mathsf{T}}\boldsymbol{\mu} - \boldsymbol{\kappa}'^{\mathsf{T}}\boldsymbol{\sigma} \geq (1 - \epsilon_i)\lambda' \tag{A.10a}$$

$$\alpha' + (\boldsymbol{\pi}'^{\mathsf{T}}_1 - \boldsymbol{\tau}'^{\mathsf{T}}_1)\boldsymbol{\mu} + \boldsymbol{\psi}'^{\mathsf{T}}_1\mathbf{t} \leq \lambda' \tag{A.10b}$$

$$\boldsymbol{\beta}'^{\mathsf{T}} + \boldsymbol{\tau}'^{\mathsf{T}}_1 = \boldsymbol{\pi}'^{\mathsf{T}}_1 + \boldsymbol{\psi}'^{\mathsf{T}}_1\mathbf{U} \tag{A.10c}$$

$$\boldsymbol{\pi}'^{\mathsf{T}}_1 + \boldsymbol{\tau}'^{\mathsf{T}}_1 = \boldsymbol{\kappa}'^{\mathsf{T}} \tag{A.10d}$$

$$\alpha' + (\boldsymbol{\pi}'^{\mathsf{T}}_2 - \boldsymbol{\tau}'^{\mathsf{T}}_2)\boldsymbol{\mu} + \boldsymbol{\psi}'^{\mathsf{T}}_2\mathbf{t} \leq b_i(\mathbf{x}) \tag{A.10e}$$

$$\boldsymbol{\beta}'^{\mathsf{T}} + \mathbf{a}_i(\mathbf{x})^{\mathsf{T}} + \boldsymbol{\tau}'^{\mathsf{T}}_2 = \boldsymbol{\pi}'^{\mathsf{T}}_2 + \boldsymbol{\psi}'^{\mathsf{T}}_2\mathbf{U} \tag{A.10f}$$

$$\boldsymbol{\pi}'^{\mathsf{T}}_2 + \boldsymbol{\tau}'^{\mathsf{T}}_2 = \boldsymbol{\kappa}'^{\mathsf{T}}. \tag{A.10g}$$

Bibliography

[1] Gurobi Optimizer. https://www.gurobi.com/products/gurobi-optimizer/.

[2] IBM CPLEX Optimizer. https://www.ibm.com/analytics/cplex-optimizer.

[3] Effect of wildfires on transmission line reliability. Technical report, California Public Utilities Commission, 2008.

[4] Node-breaker white paper. Technical report, WECC, Jan. 2014.

[5] Analysis of the cyber attack on the Ukrainian power grid. Technical report, Electricity Information Sharing and Analysis Center (E-ISAC), Mar. 2016.

[6] Proposal for development and use of node breaker topology representations for off-line study models. Technical report, NERC, Jul. 2019.

[7] CAISO 2020-2021 Transmission Plan. Technical report, California Independent System Operator, 2021.

[8] Michael Abdelmalak and Mohammed Benidris. Enhancing power system operational resilience against wildfires. *IEEE Transactions on Industry Applications*, 58(2):1611–1621, 2022.

[9] Ali Abur and Antonio Gomez Exposito. *Power system state estimation: theory and implementation*. CRC press, 2004.

[10] Ali Abur, Hongrae Kim, and MK Celik. Identifying the unknown circuit breaker statuses in power networks. *IEEE Trans. on Power Systems*, 10(4):2029–2037, 1995.

[11] Jamshid Aghaei, Ahmad Nikoobakht, Mohammad Mardaneh, Miadreza Shafie-khah, and João PS Catalão. Transmission switching, demand response and energy storage systems in an innovative integrated scheme for managing the uncertainty of wind power generation. *International Journal of Electrical Power & Energy Systems*, 98:72–84, 2018.

[12] Vladimir Agranat and Valeriy Perminov. Mathematical modeling of wildland fire initiation and spread. *Environmental Modelling & Software*, 125:104640, 2020.

[13] Shabbir Ahmed. Two-stage stochastic integer programming: A brief introduction. *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

[14] Hassene Aissi, Cristina Bazgan, and Daniel Vanderpooten. Min–max and min–max regret versions of combinatorial optimization problems: A survey. *European journal of operational research*, 197(2):427–438, 2009.

[15] Jose Ignacio Aizpurua, Victoria M Catterson, Ibrahim F Abdulhadi, and Maria Segovia Garcia. A model-based hybrid approach for circuit breaker prognostics encompassing dynamic reliability and uncertainty. *IEEE Trans. Systems, Man, and Cybernetics: Systems*, 48(9):1637–1648, 2017.

[16] Erling D Andersen and Knud D Andersen. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In *High performance optimization*, pages 197–232. Springer, 2000.

[17] Omid Ardakanian, Vincent WS Wong, Roel Dobbe, Steven H Low, Alexandra von Meier, Claire J Tomlin, and Ye Yuan. On identification of distribution grids. *IEEE Trans. on Control of Network Systems*, 6(3):950–960, 2019.

[18] Ayla Astudillo, Bai Cui, and Ahmed S Zamzam. Managing power systems-induced wildfire risks using optimal scheduled shutoffs. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2022.

[19] Alper Atamtürk and Muhong Zhang. Two-stage robust network flow and design under demand uncertainty. *Operations Research*, 55(4):662–673, 2007.

[20] Sadra Babaei, Ruiwei Jiang, and Chaoyue Zhao. Distributionally robust distribution network configuration under random contingency. *IEEE Trans. Power Systems*, 35(5):3332–3341, 2020.

[21] Yang Bai, Haiwang Zhong, Qing Xia, and Chongqing Kang. A two-level approach to AC optimal transmission switching with an accelerating technique. *IEEE Trans. Power Systems*, 32(2):1616–1625, 2016.

[22] Kyri Baker. Learning warm-start points for AC optimal power flow. In *IEEE International Workshop on MLSP*, 2019.

[23] Kyri Baker. Solutions of DC OPF are never AC feasible. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, pages 264–268, 2021.

[24] Pranavamoorthy Balasubramanian, Mostafa Sahraei-Ardakani, Xingpeng Li, and Kory W Hedman. Towards smart corrective switching: analysis and advancement of pjm's switching solutions. *IET Generation, Transmission & Distribution*, 10(8):1984–1992, 2016.

[25] Clayton Barrows, Aaron Bloom, Ali Ehlen, Jussi Ikäheimo, Jennie Jorgenson, Dheepak Krishnamurthy, Jessica Lau, Brendan McBennett, Matthew O'Connell, Eugene Preston, et al. The ieee reliability test system: A proposed 2019 update. *IEEE Transactions on Power Systems*, 35(1):119–127, 2019.

[26] Clayton Barrows, Seth Blumsack, and Paul Hines. Correcting optimal transmission switching for ac power flows. In *2014 47th Hawaii International Conference on System Sciences*, pages 2374–2379. IEEE, 2014.

[27] Mohammadhafez Bazrafshan, Nikolaos Gatsis, and Hao Zhu. Optimal power flow with step-voltage regulators in multi-phase distribution networks. *IEEE Trans. Power Systems*, 34(6):4228–4239, 2019.

[28] Aharon Ben-Tal, Boaz Golany, Arkadi Nemirovski, and Jean-Philippe Vial. Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management*, 7(3):248–271, 2005.

[29] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical programming*, 99(2):351–376, 2004.

[30] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[31] Aharon Ben-Tal and Arkadi Nemirovski. Robust convex optimization. *Mathematics of operations research*, 23(4):769–805, 1998.

[32] Patrizia Beraldi, Domenico Conforti, and Antonio Violi. A two-stage stochastic programming model for electric energy producers. *Computers & Operations Research*, 35(10):3360–3370, 2008.

[33] Arthur R Bergen and Vijay Vittal. *Power systems analysis, 2nd Edition*. Pearson, 2000.

[34] Dimitris Bertsimas, Shimrit Shtern, and Bradley Sturt. A data-driven approach for multi-stage linear optimization. *Available at Optimization Online*, 2018.

[35] Dimitris Bertsimas, Shimrit Shtern, and Bradley Sturt. Two-stage sample robust optimization. *Operations Research*, 2021.

[36] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

[37] Dimitris Bertsimas and Bartolomeo Stellato. Online mixed-integer optimization in milliseconds. *arXiv preprint arXiv:1907.02206*, 2019.

[38] Dimitris Bertsimas and Aurélie Thiele. Robust and data-driven optimization: modern decision making under uncertainty. In *Models, methods, and applications for innovative decision making*, pages 95–122. INFORMS, 2006.

[39] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

[40] Siddharth Bhela, Deepjyoti Deka, Harsha Nagarajan, and Vassilis Kekatos. Designing power grid topologies for minimizing network disturbances: An exact MILP formulation. In *Proc. American Control Conference (ACC)*, 2019.

[41] Daniel Bienstock, Michael Chertkov, and Sean Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *Siam Review*, 56(3):461–495, 2014.

[42] Silvio Binato, Mário Veiga F Pereira, and Sérgio Granville. A new benders decomposition approach to solve power transmission network design problems. *IEEE Transactions on Power Systems*, 16(2):235–240, 2001.

[43] Christian Bingane, Miguel F Anjos, and Sébastien Le Digabel. Tight-and-cheap conic relaxation for the AC optimal power flow problem. *IEEE Trans. Power Systems*, 33(6):7181–7188, 2018.

[44] Lori Bird, Jaquelin Cochran, and Xi Wang. Wind and Solar Energy Curtailment: Experience and Practices in the United States. Technical report, National Renewable Energy Laboratory (NREL), 2014.

[45] Seth Adam Blumsack. Network topologies and transmission investment under electric-industry restructuring. 2006.

[46] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[47] Christoph Buchheim and Jannis Kurtz. Robust combinatorial optimization under convex and discrete cost uncertainty. *EURO Journal on Computational Optimization*, 6(3):211–238, 2018.

[48] Giuseppe Carlo Calafiore. Multi-period portfolio optimization with linear control policies. *Automatica*, 44(10):2463–2473, 2008.

[49] Giuseppe Carlo Calafiore and Marco C Campi. The scenario approach to robust control design. *IEEE Transactions on automatic control*, 51(5):742–753, 2006.

[50] Marco C Campi and Simone Garatti. The exact feasibility of randomized solutions of uncertain convex programs. *SIAM Journal on Optimization*, 19(3):1211–1230, 2008.

[51] JM Canfield, RR Linn, JA Sauer, M Finney, and Jason Forthofer. A numerical investigation of the interplay between fireline length, geometry, and rate of spread. *Agricultural and Forest Meteorology*, 189:48–59, 2014.

[52] Eduardo Caro, Antonio J Conejo, and Ali Abur. Breaker status identification. *IEEE Trans. on Power Systems*, 25(2):694–702, 2009.

[53] Anya Castillo and Richard P O'Neill. Survey of approaches to solving the ACOPF. *Federal Energy Regulatory Commission, Tech. Rep*, 2013.

[54] Guido Cavraro and Vassilis Kekatos. Graph algorithms for topology identification using power grid probing. *IEEE control systems letters*, 2(4):689–694, 2018.

[55] Chao Chen, Jamie Twycross, and Jonathan M Garibaldi. A new accuracy measure based on bounded relative error for time series forecasting. *PloS one*, 12(3):e0174202, 2017.

[56] Xin Chen, Melvyn Sim, and Peng Sun. A robust optimization perspective on stochastic programming. *Operations Research*, 55(6):1058–1071, 2007.

[57] Xin Chen, Melvyn Sim, Peng Sun, and Jiawei Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56(2):344–357, 2008.

[58] Xin Chen, Melvyn Sim, Peng Sun, and Jiawei Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56(2):344–357, 2008.

[59] Yu Christine Chen, Taposh Banerjee, Alejandro D Dominguez-Garcia, and Venugopal V Veeravalli. Quickest line outage detection and identification. *IEEE Trans. on Power Systems*, 31(1):749–758, 2015.

[60] Terry L Clark, Janice Coen, and Don Latham. Description of a coupled atmosphere–fire model. *International Journal of Wildland Fire*, 13(1):49–63, 2004.

[61] Kevin A Clements and A Simões Costa. Topology error identification using normalized lagrange multipliers. *IEEE Trans. on power systems*, 13(2):347–353, 1998.

[62] Janice L Coen, Marques Cameron, John Michalakes, Edward G Patton, Philip J Riggan, and Kara M Yedinak. Wrf-fire: coupled weather–

wildland fire modeling with the weather research and forecasting model. *Journal of Applied Meteorology and Climatology*, 52(1):16–38, 2013.

[63] Carleton Coffrin, Russell Bent, Kaarthik Sundar, Yeesian Ng, and Miles Lubin. Powermodels. jl: An open-source framework for exploring power flow formulations. In *Power Systems Computation Conference (PSCC)*, pages 1–8, 2018.

[64] Carleton Coffrin, Russell Bent, Byron Tasseff, Kaarthik Sundar, and Scott Backhaus. Relaxations of AC maximal load delivery for severe contingency analysis. *IEEE Trans. Power Systems*, 34(2):1450–1458, 2018.

[65] Carleton Coffrin, Hassan L Hijazi, Karsten Lehmann, and Pascal Van Hentenryck. Primal and dual bounds for optimal transmission switching. In *2014 Power Systems Computation Conference*, pages 1–8. IEEE, 2014.

[66] Carleton Coffrin, Hassan L Hijazi, and Pascal Van Hentenryck. The QC relaxation: A theoretical and computational study on optimal power flow. *IEEE Trans. Power Systems*, 31(4):3008–3018, 2015.

[67] Charles Concordia, Lester H Fink, and Geroge Poullikkas. Load shedding on an isolated system. *IEEE Trans. Power Systems*, 10(3):1467–1472, 1995.

[68] Emil M Constantinescu, Victor M Zavala, Matthew Rocklin, Sangmin Lee, and Mihai Anitescu. A computational framework for uncertainty

quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Trans. Power Systems*, 26(1):431–441, 2010.

[69] IBM ILOG Cplex. V12. 1: User's manual for cplex. *International Business Machines Corporation*, 46(53):157, 2009.

[70] Constance Crozier, Kyri Baker, and Bridget Toomey. Feasible region-based heuristics for optimal transmission switching. *Sustainable Energy, Grids and Networks*, 30:100628, 2022.

[71] Emiliano Dall'Anese, Hao Zhu, and Georgios B Giannakis. Distributed optimal power flow for smart microgrids. *IEEE Trans. Smart Grid*, 4(3):1464–1475, 2013.

[72] Emiliano Dall'Anese, Kyri Baker, and Tyler Summers. Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Trans. Power Systems*, 32(5):3427–3438, 2017.

[73] Shahab Dehghan and Nima Amjady. Robust transmission and energy storage expansion planning in wind farm-integrated power systems considering transmission switching. *IEEE Transactions on Sustainable Energy*, 7(2):765–774, 2015.

[74] Payman Dehghanian and Mladen Kezunovic. Probabilistic decision making for the bulk power system optimal topology control. *IEEE Trans. Smart Grid*, 7(4):2071–2081, 2016.

[75] Deepjyoti Deka, Scott Backhaus, and Michael Chertkov. Estimating distribution grid topologies: A graphical learning based approach. In *2016 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2016.

[76] Deepjyoti Deka, Ross Baldick, and Sriram Vishwanath. One breaker is enough: Hidden topology attacks on power grids. In *Proc. IEEE PES General Meeting*, 2015.

[77] Deepjyoti Deka and Sidhant Misra. Learning for DC-OPF: Classifying active sets using neural nets. In *2019 IEEE Milan PowerTech*, pages 1–6, 2019.

[78] Andrés Delgadillo, José Manuel Arroyo, and Natalia Alguacil. Analysis of electric grid interdiction with line switching. *IEEE Transactions on Power Systems*, 25(2):633–641, 2009.

[79] Songyi Dian, Peng Cheng, Qiang Ye, Jirong Wu, Ruisen Luo, Chen Wang, Dafeng Hui, Ning Zhou, Dong Zou, Qin Yu, et al. Integrating wildfires propagation prediction into early warning of electrical transmission line outages. *IEEE Access*, 7:27586–27603, 2019.

[80] Tao Ding, Cheng Li, Yongheng Yang, Jiangfeng Jiang, Zhaohong Bie, and Frede Blaabjerg. A two-stage robust optimization for centralized-optimal dispatch of photovoltaic inverters in active distribution networks. *IEEE Trans. Sustainable Energy*, 8(2):744–754, 2016.

[81] Florian Dorfler and Francesco Bullo. Kron reduction of graphs with applications to electrical networks. *IEEE Trans. on Circuits and Systems I: Regular Papers*, 60(1):150–163, 2012.

[82] Mohamed Drissi. Modelling the spreading of large-scale wildland fires. *arXiv preprint arXiv:1402.6187*, 2014.

[83] Chao Duan, Wanliang Fang, Lin Jiang, Li Yao, and Jun Liu. Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric. *IEEE Trans. Power Systems*, 33(5):4924–4936, 2018.

[84] Iain Dunning, Joey Huchette, and Miles Lubin. JuMP: A modeling language for mathematical optimization. *SIAM review*, 59(2):295–320, 2017.

[85] P Dutta and M Kezunovic. Unified representation of data and model for sparse measurement based fault location. In *IEEE PES General Meeting*, pages 1–8, 2012.

[86] Erik Ela, Brendan Kirby, Eamonn Lannoye, Michael Milligan, Damian Flynn, Bob Zavadil, and Mark O'Malley. Evolution of operating reserve determination in wind power integration studies. In *Proc. IEEE PES General Meeting*, 2010.

[87] Hakan Ergun, Dirk Van Hertem, and Ronnie Belmans. Transmission system topology optimization for large-scale offshore wind integration. *IEEE Trans. Sustainable Energy*, 3(4):908–917, 2012.

[88] Tomaso Erseghe. Distributed optimal power flow using ADMM. *IEEE Trans. Power Systems*, 29(5):2370–2380, 2014.

[89] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, 2018.

[90] A Gómez Expósito and Antonio de la Villa Jaen. Reduced substation models for generalized state estimation. *IEEE Trans. on Power Systems*, 16(4):839–846, 2001.

[91] Thelma SP Fernandes, JR Lenzi, and Miguel A Mikilita. Load shedding strategies using optimal load flow with relaxation of restrictions. *IEEE Trans. Power Systems*, 23(2):712–718, 2008.

[92] Jean-Baptiste Filippi, Xavier Pialat, and Craig B Clements. Assessment of forefire/meso-nh for wildland fire/atmosphere coupled simulation of the fireflux experiment. *Proceedings of the Combustion Institute*, 34(2):2633–2640, 2013.

[93] Emily B Fisher, Richard P O'Neill, and Michael C Ferris. Optimal transmission switching. *IEEE Trans. on Power Systems*, 23(3):1346–1355, 2008.

[94] Aoife M Foley, Paul G Leahy, Antonino Marvuglia, and Eamon J McKeogh. Current methods and advances in forecasting of wind power

generation. *Renewable Energy*, 37(1):1–8, 2012.

[95] J David Fuller, Raynier Ramasra, and Amanda Cha. Fast heuristics for transmission-line switching. *IEEE Transactions on Power Systems*, 27(3):1377–1386, 2012.

[96] Jean Gallier. The schur complement and symmetric positive semidefinite (and definite) matrices. *Penn Engineering*, 2010.

[97] Rui Gao and Anton J Kleywegt. Distributionally robust stochastic optimization with Wasserstein distance. *arXiv preprint arXiv:1604.02199*, 2016.

[98] Manuel Garcia and Ross Baldick. Requirements for interdependent reserve types providing primary frequency control. *IEEE Transactions on Power Systems*, 37(1):51–64, 2021.

[99] Farhad Samadi Gazijahani and Javad Salehi. Robust design of microgrids with reconfigurable topology under severe uncertainty. *IEEE Trans. Sustainable Energy*, 9(2):559–569, 2017.

[100] Xinbo Geng and Le Xie. Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization. *Annual reviews in control*, 47:341–363, 2019.

[101] Angelos Georghiou, Daniel Kuhn, and Wolfram Wiesemann. The decision rule approach to optimization under uncertainty: methodology and applications. *Computational Management Science*, pages 1–32.

[102] Laurent El Ghaoui, Maksim Oks, and Francois Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations research*, 51(4):543–556, 2003.

[103] J Duncan Glover, Mulukutla S Sarma, and Thomas Overbye. *Power system analysis & design*. Cengage Learning, 2012.

[104] Evgeniy A Goldis, Xiaoguang Li, Michael C Caramanis, Aleksandr M Rudkevich, and Pablo A Ruiz. Ac-based topology control algorithms (tca)–a pjm historical data case study. In *2015 48th Hawaii International Conference on System Sciences*, pages 2516–2519. IEEE, 2015.

[105] Michael Gollner, Arnaud Trouve, Ilkay Altintas, Jessica Block, Raymond de Callafon, Craig Clements, Anna Cortes, Evan Ellicott, Jean Baptiste Filippi, Mark Finney, et al. Towards data-driven operational wildfire spread modeling: A report of the nsf-funded wifire workshop. Technical report, 2015.

[106] Antonio Gomez-Exposito and Ali Abur. *Power system state estimation: theory and implementation*. CRC press, 2004.

[107] Antonio Gómez-Expósito, Antonio J Conejo, and Claudio Cañizares. *Electric energy systems: analysis and operation*. CRC press, 2018.

[108] Chrysanthos E Gounaris, Wolfram Wiesemann, and Christodoulos A Floudas. The robust capacitated vehicle routing problem under demand uncertainty. *Operations Research*, 61(3):677–693, 2013.

[109] Gianpietro Granelli, Mario Montagna, Fabio Zanellini, Paola Bresesti, Riccardo Vailati, and Mario Innorta. Optimal network reconfiguration for congestion management by deterministic and genetic algorithms. *Electric power systems research*, 76(6-7):549–556, 2006.

[110] Santiago Grijalva and Ayusman Roy. Automated handling of arbitrary switching device topologies in planning contingency analysis. *IEEE Trans. Power Systems*, 28(2):1523–1530, 2012.

[111] Igor Griva, Stephen G Nash, and Ariela Sofer. *Linear and nonlinear optimization*, volume 108. Siam, 2009.

[112] Neel Guha, Zhecheng Wang, Matt Wytock, and Arun Majumdar. Machine learning for AC optimal power flow. *arXiv preprint arXiv:1910.08842*, 2019.

[113] Teoman Guler, George Gross, and Minghai Liu. Generalized line outage distribution factors. *IEEE Trans. on Power Systems*, 22(2):879–881, 2007.

[114] Jiachun Guo, Yong Fu, Zuyi Li, and Mohammad Shahidehpour. Direct calculation of line outage distribution factors. *IEEE Trans. on Power Systems*, 24(3):1633–1634, 2009.

[115] Yi Guo, Kyri Baker, Emiliano Dall'Anese, Zechun Hu, and Tyler Holt Summers. Data-based distributionally robust stochastic optimal power

flow - Part I: Methodologies. *IEEE Trans. Power Systems*, 34(2):1483–1492, 2018.

[116] Martin T Hagan and Mohammad B Menhaj. Training feedforward networks with the marquardt algorithm. *IEEE Trans. Neural Networks*, 5(6):989–993, 1994.

[117] Grani A Hanasusanto, Daniel Kuhn, and Wolfram Wiesemann. A comment on "computational complexity of stochastic programming problems". *Mathematical Programming*, 159(1-2):557–569, 2016.

[118] Grani A Hanasusanto, Vladimir Roitch, Daniel Kuhn, and Wolfram Wiesemann. A distributionally robust perspective on uncertainty quantification and chance constrained programming. *Mathematical Programming*, 151(1):35–62, 2015.

[119] Grani A Hanasusanto, Vladimir Roitch, Daniel Kuhn, and Wolfram Wiesemann. Ambiguous joint chance constraints under mean and dispersion information. *Operations Research*, 65(3):751–767, 2017.

[120] Lars Kai Hansen and Peter Salamon. Neural network ensembles. *IEEE Trans. pattern analysis and machine intelligence*, 12(10):993–1001, 1990.

[121] Ali Hassan, Robert Mieth, Deepjyoti Deka, and Yury Dvorkin. Stochastic and distributionally robust load ensemble control. *IEEE Trans. Power Systems*, 35(6):4678–4688, 2020.

[122] Kory W Hedman, Richard P O'Neill, Emily Bartholomew Fisher, and Shmuel S Oren. Optimal transmission switching: sensitivity analysis and extensions. *IEEE Trans. Power Systems*, 23(3):1469–1479, 2008.

[123] Kory W Hedman, Richard P O'Neill, Emily Bartholomew Fisher, and Shmuel S Oren. Optimal transmission switching with contingency analysis. *IEEE Transactions on Power Systems*, 24(3):1577–1586, 2009.

[124] Majid Heidarifar and Hassan Ghasemi. A network topology optimization model based on substation and node-breaker modeling. *IEEE Trans. Power Systems*, 31(1):247–255, 2015.

[125] Pierre Henneaux and Daniel S Kirschen. Probabilistic security analysis of optimal transmission switching. *IEEE Trans. Power Systems*, 31(1):508–517, 2015.

[126] Nicholas J Higham. Accuracy and stability of numerical algorithms, 2002.

[127] Wanshi Hong, Bin Wang, Mengqi Yao, Duncan Callaway, Larry Dale, and Can Huang. Data-driven power system optimal decision making strategy underwildfire events. Technical report, Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 2022.

[128] Qinghua Hu, Yun Wang, Zongxia Xie, Pengfei Zhu, and Daren Yu. On estimating uncertainty of wind energy with mixture of distributions. *Energy*, 112:935–962, 2016.

[129] Qiuhua Huang, Renke Huang, Weituo Hao, Jie Tan, Rui Fan, and Zhenyu Huang. Adaptive power system emergency control using deep reinforcement learning. *IEEE Trans. Smart Grid*, 11(2):1171–1182, 2019.

[130] Tong Huang, Meng Wu, and Le Xie. Prioritization of PMU location and signal selection for monitoring critical power system oscillations. *IEEE Trans. Power Systems*, 33(4):3919–3929, 2018.

[131] Gerd Infanger. Monte carlo (importance) sampling within a benders decomposition algorithm for stochastic linear programs. *Annals of Operations Research*, 39(1):69–95, 1992.

[132] Rabih A Jabr, Alun H Coonick, and Brian J Cory. A homogeneous linear programming algorithm for the security constrained economic dispatch problem. *IEEE Trans. Power Systems*, 15(3):930–936, 2000.

[133] Amir Abiri Jahromi, Anthony Kemmeugne, Deepa Kundur, and Aboutaleb Haddadi. Cyber-physical attacks targeting communication-assisted protection schemes. *IEEE Trans. Power Systems*, 35(1):440–450, 2019.

[134] Nasser Jaleeli, Louis S VanSlyck, Donald N Ewart, Lester H Fink, and Arthur G Hoffmann. Understanding automatic generation control. *IEEE Trans. Power Systems*, 7(3):1106–1122, 1992.

[135] Ruiwei Jiang, Jianhui Wang, Muhong Zhang, and Yongpei Guan. Two-stage minimax regret robust unit commitment. *IEEE Trans. Power*

Systems, 28(3):2271–2282, 2013.

[136] Ruiwei Jiang, Muhong Zhang, Guang Li, and Yongpei Guan. Benders' decomposition for the two-stage security constrained robust unit commitment problem. In *IIE Annual Conference. Proceedings*, page 1. Institute of Industrial and Systems Engineers (IISE), 2012.

[137] Wenyun Ju, Kai Sun, and Rui Yao. Simulation of cascading outages using a power-flow model considering frequency. *IEEE Access*, 6:37784–37795, 2018.

[138] John G Kassakian, Richard Schmalensee, Gary Desgroseilliers, et al. The future of the electric grid. Technical report, 2011.

[139] Andrew Keane, Michael Milligan, Chris J Dent, Bernhard Hasche, Claudine D'Annunzio, Ken Dragoon, Hannele Holttinen, Nader Samaan, Lennart Soder, and Mark O'Malley. Capacity value of wind power. *IEEE Trans. Power Systems*, 26(2):564–572, 2010.

[140] Vassilis Kekatos and Georgios B Giannakis. Joint power system state estimation and breaker status identification. In *Proc. North American Power Symp.*, 2012.

[141] Mojtaba Khanabadi, Hassan Ghasemi, and Meysam Doostizadeh. Optimal transmission switching considering voltage security and n-1 contingency analysis. *IEEE Transactions on Power Systems*, 28(1):542–550, 2012.

[142] Cheolmin Kim, Kibaek Kim, Prasanna Balaprakash, and Mihai Anitescu. Graph convolutional neural networks for optimal load shedding under line contingency. In *Proc. IEEE PES General Meeting*, pages 1–5. IEEE, 2019.

[143] Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

[144] Burak Kocuk, Santanu S Dey, and X Andy Sun. Strong SOCP relaxations for the optimal power flow problem. *Operations Research*, 64(6):1177–1196, 2016.

[145] Burak Kocuk, Santanu S Dey, and Xu Andy Sun. New formulation and strong misocp relaxations for ac optimal transmission switching problem. *IEEE Transactions on Power Systems*, 32(6):4161–4170, 2017.

[146] Alyssa Kody, Ryan Piansky, and Daniel K Molzahn. Optimizing transmission infrastructure investments to support line de-energization for mitigating wildfire ignition risk. *arXiv preprint arXiv:2203.10176*, 2022.

[147] Alyssa Kody, Amanda West, and Daniel K Molzahn. Sharing the load: Considering fairness in de-energization scheduling to mitigate wildfire ignition risk using rolling optimization. *arXiv preprint arXiv:2204.06543*, 2022.

[148] Akshay S Korad and Kory W Hedman. Robust corrective topology control for system reliability. *IEEE Transactions on Power Systems*, 28(4):4042–4051, 2013.

[149] George N Korres and Petros J Katsikas. Identification of circuit breaker statuses in WLS state estimator. *IEEE Trans. on Power Systems*, 17(3):818–825, 2002.

[150] GN Korres, PJ Katsikas, and GE Chatzarakis. Substation topology identification in generalized state estimation. *International Journal of Electrical Power & Energy Systems*, 28(3):195–206, 2006.

[151] Simge Küçükyavuz and Suvrajeet Sen. An introduction to two-stage stochastic mixed-integer programming. In *Leading Developments from INFORMS Communities*, pages 1–27. INFORMS, 2017.

[152] Daniel Kuhn, Wolfram Wiesemann, and Angelos Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209, 2011.

[153] Tian Lan and Garng M Huang. Transmission line switching in power system planning with large scale renewable energy. In *2015 First Workshop on Smart Grid and Renewable Energy (SGRE)*, pages 1–6. IEEE, 2015.

[154] Tian Lan, Zhangxin Zhou, Wenzong Wang, and Garng M Huang. Stochastic optimization for AC optimal transmission switching with generalized

benders decomposition. *International Journal of Electrical Power & Energy Systems*, 133:107140, 2021.

[155] Bowen Li, Ruiwei Jiang, and Johanna L Mathieu. Distributionally robust risk-constrained optimal power flow using moment and unimodality information. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 2425–2430. IEEE, 2016.

[156] Bowen Li, Ruiwei Jiang, and Johanna L Mathieu. Distributionally robust chance-constrained optimal power flow assuming unimodal distributions with misspecified modes. *IEEE Trans. Control of Network Systems*, 6(3):1223–1234, 2019.

[157] Peng Li, Ming Yang, and Qiuwei Wu. Confidence interval based distributionally robust real-time economic dispatch approach considering wind power accommodation risk. *IEEE Trans. Sustainable Energy*, 12(1):58–69, 2020.

[158] Xiao Li, H Vincent Poor, and Anna Scaglione. Blind topology identification for power systems. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 91–96. IEEE, 2013.

[159] Yang Li, Meng Han, Zhen Yang, and Guoqing Li. Coordinating flexible demand response and renewable uncertainties for scheduling of community integrated energy systems with an electric vehicle charging station: A bi-level approach. *IEEE Trans. Sustainable Energy*, 2021.

[160] Fu Lin and Chen Chen. An ADMM algorithm for load shedding in electric power grids. In *American Control Conf.*, pages 5002–5007. IEEE, 2016.

[161] Shanny Lin, Shaohui Liu, and Hao Zhu. Risk-aware learning for scalable voltage optimization in distribution grids. *Electric Power Systems Research*, 212:108605, 2022.

[162] Rodman Linn, Jon Reisner, Jonah J Colman, and Judith Winterkamp. Studying wildfire behavior using firetec. *International journal of wildland fire*, 11(4):233–246, 2002.

[163] Cong Liu, Mohammad Shahidehpour, and Lei Wu. Extended benders decomposition for two-stage SCUC. *IEEE Transactions on Power Systems*, 25(2):1192–1194, 2010.

[164] Cong Liu, Jianhui Wang, and James Ostrowski. Heuristic prescreening switchable branches in optimal transmission switching. *IEEE Transactions on Power Systems*, 27(4):2289–2290, 2012.

[165] Johan Löfberg. YALMIP: A toolbox for modeling and optimization in matlab. In *Proceedings of the CACSD Conference*, 2004.

[166] Álvaro Lorca, X Andy Sun, Eugene Litvinov, and Tongxin Zheng. Multistage adaptive robust optimization for the unit commitment problem. *Operations Research*, 64(1):32–51, 2016.

[167] Elizete Maria Lourenço, A Simões Costa, and Kevin A Clements. Bayesian-based hypothesis testing for topology error identification in generalized state estimation. *IEEE Trans. on power systems*, 19(2):1206–1215, 2004.

[168] Steven H Low. Convex relaxation of optimal power flow—part i: Formulations and equivalence. *IEEE Trans. Control of Network Systems*, 1(1):15–27, 2014.

[169] Xi Lu, Ka Wing Chan, Shiwei Xia, Bin Zhou, and Xiao Luo. Security-constrained multiperiod economic dispatch with renewable energy utilizing distributionally robust optimization. *IEEE Trans. Sustainable Energy*, 10(2):768–779, 2018.

[170] Miles Lubin, Yury Dvorkin, and Scott Backhaus. A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Trans. Power Systems*, 31(5):3840–3849, 2015.

[171] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

[172] James Luedtke, Shabbir Ahmed, and George L Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical programming*, 122(2):247–272, 2010.

[173] Thomas L Magnanti and Richard T Wong. Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. *Operations research*, 29(3):464–484, 1981.

[174] Jan Mandel, Jonathan D Beezley, and Adam K Kochanski. Coupled atmosphere-wildland fire modeling with wrf 3.3 and sfire 2011. *Geoscientific Model Development*, 4(3):591–610, 2011.

[175] Georgios Mavromatidis, Kristina Orehounig, and Jan Carmeliet. Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach. *Applied energy*, 222:932–950, 2018.

[176] Abdulhalem A Mazi, Bruce F Wollenberg, and Morten H Hesse. Corrective control of power system flows by line and bus-bar switching. *IEEE Trans. on Power Systems*, 1(3):258–264, 1986.

[177] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part I-Convex underestimating problems. *Mathematical programming*, 10(1):147–175, 1976.

[178] Alan W McMorran, Graham W Ault, Ian M Elders, Colin ET Foote, Graeme M Burt, and James R McDonald. Translating cim xml power system data to a proprietary format for system simulation. *IEEE Trans. on Power Systems*, 19(1):229–235, 2004.

[179] William Mell, Mary Ann Jenkins, Jim Gould, and Phil Cheney. A

physics-based approach to modelling grassland fires. *International Journal of Wildland Fire*, 16(1):1–22, 2007.

[180] L Mili, G Steeno, F Dobraca, and D French. A robust estimation method for topology error identification. *IEEE Trans. on Power Systems*, 14(4):1469–1476, 1999.

[181] Mohammad Sadegh Modarresi, Le Xie, Marco Claudio Campi, Simone Garatti, Algo Care, Anupam A Thatte, and PR Kumar. Scenario-based economic dispatch with tunable risk levels in high-renewable power systems. *IEEE Transactions on Power Systems*, 34(6):5103–5114, 2018.

[182] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, 2018.

[183] Alcir Monticelli. Electric power system state estimation. *Proceedings of the IEEE*, 88(2):262–282, 2000.

[184] Alcir Monticelli. *State estimation in electric power systems: a generalized approach*. Springer Science & Business Media, 2012.

[185] Daniel Morvan and Jean-Luc Dupuy. Modeling of fire spread through a forest fuel bed using a multiphase formulation. *Combustion and flame*, 127(1-2):1981–1994, 2001.

[186] Daniel Morvan and Jean-Luc Dupuy. Modeling the propagation of a wildfire through a mediterranean shrub using a multiphase formulation. *Combustion and flame*, 138(3):199–210, 2004.

[187] H Nazaripouya. Power grid resilience under wildfire: A review on challenges and solutions. In *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2020.

[188] Mostafa Nazemi and Payman Dehghanian. Powering through wildfires: An integrated solution for enhanced safety and resilience in power grids. *IEEE Transactions on Industry Applications*, 58(3):4192–4202, 2022.

[189] Mostafa Nazemi, Payman Dehghanian, and Miguel Lejeune. A mixed-integer distributionally robust chance-constrained model for optimal topology control in power grids with uncertain renewables. In *2019 IEEE Milan PowerTech*, pages 1–6. IEEE, 2019.

[190] HyungSeon Oh. A new network reduction methodology for power system planning studies. *IEEE Trans. Power Systems*, 25(2):677–684, 2010.

[191] Christos Ordoudis, Viet Anh Nguyen, Daniel Kuhn, and Pierre Pinson. Energy and reserve dispatch with distributionally robust joint chance constraints. Technical report, 2018.

[192] Shmuel Oren, K Hedman, and Richard P O'Neill. Optimal transmission switching: When economic efficiency and ftr markets collide. In

Economics of Energy Markets Conference, 2010.

[193] James Ostrowski, Miguel F Anjos, and Anthony Vannelli. Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Trans. on Power Systems*, 27(1):39–46, 2011.

[194] Bernardo K Pagnoncelli, Shabbir Ahmed, and Alexander Shapiro. Sample average approximation method for chance constrained programming: theory and applications. *Journal of optimization theory and applications*, 142(2):399–416, 2009.

[195] Anthony Papavasiliou, Shmuel S Oren, and Richard P O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Trans. Power Systems*, 26(4):2197–2206, 2011.

[196] Byungkwon Park and Christopher L Demarco. Optimal network topology for node-breaker representations with ac power flow constraints. *IEEE Access*, 8:64347–64355, 2020.

[197] Byungkwon Park, Jesse Holzer, and Christopher L DeMarco. A sparse tableau formulation for node-breaker representations in security-constrained optimal power flow. *IEEE Trans. on Power Systems*, 34(1):637–647, 2018.

[198] Byungkwon Park, Jesse Holzer, and Christopher L DeMarco. A sparse tableau formulation for node-breaker representations in security-constrained

optimal power flow. *IEEE Trans. Power Systems*, 34(1):637–647, 2019.

[199] Keith Parks, Yih-Huei Wan, Gerry Wiener, and Yubao Liu. Wind energy forecasting: A collaboration of the national center for atmospheric research (NCAR) and Xcel energy. Technical report, National Renewable Energy Lab (NREL), Golden, CO (United States), 2011.

[200] Pierre Pinson. Wind energy: Forecasting challenges for its operational management. *Statistical Science*, 28(4):564–585, 2013.

[201] Robert J Plemmons. M-matrix characterizations. I-nonsingular m-matrices. *Linear Algebra and its Applications*, 18(2):175–188, 1977.

[202] Bala Kameshwar Poolla, Ashish R Hota, Saverio Bolognani, Duncan S Callaway, and Ashish Cherukuri. Wasserstein distributionally robust look-ahead economic dispatch. *IEEE Trans. Power Systems*, 36(3):2010–2022, 2020.

[203] B Porterie, D Morvan, JC Loraud, and M Larini. Firespread through fuel beds: Modeling of wind-aided fires and induced hydrodynamics. *Physics of Fluids*, 12(7):1762–1782, 2000.

[204] Farzaneh Pourahmadi and Jalal Kazempour. Distributionally robust generation expansion planning with unimodality and risk constraints. *IEEE Trans. Power Systems*, 2021.

[205] Yemula Pradeep, P Seshuraju, Shrikrishna A Khaparde, and Rushikesh K Joshi. CIM-based connectivity model for bus-branch topology extraction and exchange. *IEEE Trans. on Smart Grid*, 2(2):244–253, 2011.

[206] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. Section 2.7. 1 sherman–morrison formula. *Numerical Recipes: The Art of Scientific Computing*, 2007.

[207] Feng Qiu and Jianhui Wang. Chance-constrained transmission switching with guaranteed wind power utilization. *IEEE Transactions on Power Systems*, 30(3):1270–1278, 2014.

[208] Feng Qiu and Jianhui Wang. Chance-constrained transmission switching with guaranteed wind power utilization. *IEEE Trans. on Power Systems*, 30(3):1270–1278, 2015.

[209] R Ramanathan and Brian Tuck. Contingency analysis using node/breaker model for operation studies. In *Proc. IEEE PES General Meeting*, 2015.

[210] Walter Rei, Jean-François Cordeau, Michel Gendreau, and Patrick Soriano. Accelerating benders decomposition by local branching. *INFORMS Journal on Computing*, 21(2):333–345, 2009.

[211] Noah Rhodes, Lewis Ntaimo, and Line Roald. Balancing wildfire risk and power outages through optimized power shut-offs. *IEEE Trans. Power Systems*, 2020.

[212] Noah Rhodes and Line Roald. Co-optimization of power line shutoff and restoration for electric grids under high wildfire ignition risk. *arXiv preprint arXiv:2204.02507*, 2022.

[213] Line Roald and Göran Andersson. Chance-constrained AC optimal power flow: Reformulations and efficient algorithms. *IEEE Trans. Power Systems*, 33(3):2906–2918, 2017.

[214] Line Roald, Sidhant Misra, Michael Chertkov, and Göran Andersson. Optimal power flow with weighted chance constraints and general policies for generation control. In *2015 54th IEEE conference on decision and control (CDC)*, pages 6927–6933. IEEE, 2015.

[215] Line Roald, Sidhant Misra, Thilo Krause, and Göran Andersson. Corrective control to handle forecast uncertainty: A chance constrained optimal power flow. *IEEE Trans. Power Systems*, 32(2):1626–1637, 2016.

[216] Line Roald, Frauke Oldewurtel, Bart Van Parys, and Göran Andersson. Security constrained optimal power flow with distributionally robust chance constraints. *arXiv preprint arXiv:1508.06061*, 2015.

[217] Michael Ross, Chad Abbey, Franois Bouffard, and Gza Jos. Multiobjective optimization dispatch for microgrids with a high penetration of renewable generation. *IEEE Trans. Sustainable Energy*, 6(4):1306–1314, 2015.

[218] Reuven Y Rubinstein and Dirk P Kroese. *Simulation and the Monte Carlo method*, volume 10. John Wiley & Sons, 2016.

[219] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[220] Pablo A Ruiz, Justin M Foster, Aleksandr Rudkevich, and Michael C Caramanis. Tractable transmission topology control using sensitivity analysis. *IEEE Transactions on Power Systems*, 27(3):1550–1559, 2012.

[221] Andrzej Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93(2):195–215, 2002.

[222] S. Santoso. Transmission line protection for enhanced security: Including circuit models in relays. Technical report, Dec. 2014.

[223] Chengcheng Shao, Xifan Wang, Mohammad Shahidehpour, Xiuli Wang, and Biyang Wang. Security-constrained unit commitment with flexible uncertainty set for variable wind power. *IEEE Trans. Sustainable Energy*, 8(3):1237–1246, 2017.

[224] Wei Shao and Vijay Vittal. Corrective switching algorithm for relieving overloads and voltage violations. *IEEE Trans. on Power Systems*, 20(4):1877–1885, 2005.

[225] Alexander Shapiro. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research*, 58(1):57–68, 2003.

[226] Alexander Shapiro. On complexity of multistage stochastic programs. *Operations Research Letters*, 34(1):1–8, 2006.

[227] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[228] Alexander Shapiro and Arkadi Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, pages 111–146. Springer, 2005.

[229] Alexander Shapiro and Andy Philpott. A tutorial on stochastic programming. *Manuscript Available Online*, 2007.

[230] Dmitry Shchetinin, Tomas Tinoco De Rubira, and Gabriela Hug. On the construction of linear approximations of line flow constraints for AC optimal power flow. *IEEE Transactions on Power Systems*, 34(2):1182–1192, 2018.

[231] Pouria Sheikhahmadi, Ramyar Mafakheri, Salah Bahramara, Maziar Damavandi, and João Catalão. Risk-based two-stage stochastic optimization problem of micro-grid operation with renewables and incentive-based demand response programs. *Energies*, 11(3):610, 2018.

[232] Jack Sherman and Winifred J Morrison. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127, 1950.

[233] Jiaying Shi and Shmuel S Oren. Wind power integration through stochastic unit commitment with topology control recourse. In *2016 Power systems computation conference (PSCC)*, pages 1–7. IEEE, 2016.

[234] Zhichao Shi, Hao Liang, Shengjun Huang, and Venkata Dinavahi. Distributionally robust chance-constrained energy management for islanded microgrids. *IEEE Trans. Smart Grid*, 10(2):2234–2244, 2018.

[235] Jiajia Song, Eduardo Cotilla-Sanchez, Goodarz Ghanavati, and Paul DH Hines. Dynamic modeling of cascading failure in power systems. *IEEE Trans. Power Systems*, 31(3):2085–2095, 2015.

[236] Milad Soroush and J David Fuller. Accuracies of optimal transmission switching heuristics based on dcopf and acopf. *IEEE Transactions on Power Systems*, 29(2):924–932, 2013.

[237] Robert F Stengel. *Optimal control and estimation*. Courier Corporation, 1994.

[238] Brian Stott and Of Alsac. Fast decoupled load flow. *IEEE Trans. Power Apparatus and Systems*, (3):859–869, 1974.

[239] Brian Stott, Jorge Jardim, and Ongun Alsaç. DC power flow revisited. *IEEE Trans. on Power Systems*, 24(3):1290–1300, 2009.

[240] Kai Sun, Da-Zhong Zheng, and Qiang Lu. Splitting strategies for islanding operation of large-scale power systems using OBDD-based methods. *IEEE Trans. Power Systems*, 18(2):912–923, 2003.

[241] Tomás Tapia, Álvaro Lorca, Daniel Olivares, Matías Negrete-Pincetic, et al. A robust decision-support method based on optimization and simulation for wildfire resilience in highly renewable power systems. *European Journal of Operational Research*, 294(2):723–733, 2021.

[242] Joseph Euzebe Tate and Thomas J Overbye. Line outage detection using phasor angle measurements. *IEEE Trans. on Power Systems*, 23(4):1644–1652, 2008.

[243] Sofia Taylor and Line A Roald. A framework for risk assessment and optimal line upgrade selection to mitigate wildfire risk. *Electric Power Systems Research*, 213:108592, 2022.

[244] Chee-Wooi Ten, Koji Yamashita, Zhiyuan Yang, Athanasios V Vasilakos, and Andrew Ginter. Impact assessment of hypothesized cyberattacks on interconnected bulk power systems. *IEEE Trans. Smart Grid*, 9(5), 2017.

[245] Kevin Tolhurst, Brett Shields, and Derek Chong. Phoenix: development and application of a bushfire risk management tool. *Australian journal of emergency management*, 23(4):47–54, 2008.

[246] Dimitris N Trakas and Nikos D Hatziargyriou. Optimal distribution system operation for enhancing resilience against wildfires. *IEEE Transactions on Power Systems*, 33(2):2260–2271, 2017.

[247] Nicolás Garcia Trillos and Dejan Slepčev. On the rate of convergence of empirical measures in ∞-transportation distance. *Canadian Journal of Mathematics*, 67(6):1358–1383, 2015.

[248] Daniel Alberto Zuniga Vazquez, Feng Qiu, Neng Fan, and Kevin Sharp. Wildfire mitigation plans in power systems: A literature review. *IEEE Transactions on Power Systems*, 2022.

[249] Andreas Venzke, Lejla Halilbasic, Uros Markovic, Gabriela Hug, and Spyros Chatzivasileiadis. Convex relaxations of chance constrained AC optimal power flow. *IEEE Trans. Power Systems*, 33(3):2829–2841, 2017.

[250] Maria Vrakopoulou, Kostas Margellos, John Lygeros, and Göran Andersson. A probabilistic framework for reserve scheduling and $N-1$ security assessment of systems with high wind power penetration. *IEEE Transactions on Power Systems*, 28(4):3885–3896, 2013.

[251] Thanh Long Vu, Sayak Mukherjee, Tim Yin, Renke Huang, Qiuhua Huang, et al. Safe reinforcement learning for emergency loadshedding of power systems. *arXiv preprint arXiv:2011.09664*, 2020.

[252] Can Wan, Zhao Xu, Pierre Pinson, Zhao Yang Dong, and Kit Po Wong. Optimal prediction intervals of wind power generation. *IEEE Trans. Power Systems*, 29(3):1166–1174, 2014.

[253] Cheng Wang, Rui Gao, Feng Qiu, Jianhui Wang, and Linwei Xin. Risk-based distributionally robust optimal power flow with dynamic line rating. *IEEE Trans. Power Systems*, 33(6):6074–6086, 2018.

[254] Qianfan Wang, Yongpei Guan, and Jianhui Wang. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Trans. Power Systems*, 27(1):206–215, 2011.

[255] Shiyuan Wang, Payman Dehghanian, and Bei Zhang. A data-driven algorithm for online power grid topology change identification with pmus. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2019.

[256] Yifan Wang, Shixin Liu, Jianhui Wang, and Bo Zeng. Capacity expansion of wind power in a market environment with topology control. *IEEE Trans. Sustainable Energy*, 10(4):1834–1843, 2018.

[257] Zhiming Wang and Weimin Liu. Wind energy potential assessment based on wind speed, its direction and power data. *Scientific reports*, 11(1):1–15, 2021.

[258] Joseph Warrington, Paul Goulart, Sébastien Mariéthoz, and Manfred Morari. Policy-based reserves for power systems. *IEEE Transactions*

on Power Systems, 28(4):4427–4437, 2013.

[259] Wei Wei, Feng Liu, and Shengwei Mei. Distributionally robust co-optimization of energy and reserve dispatch. *IEEE Trans. Sustainable Energy*, 7(1):289–300, 2015.

[260] David White, Amy Roschelle, Paul Peterson, David Schlissel, Bruce Biewald, and William Steinhurst. The 2003 blackout: solutions that won't cost a fortune. *The Electricity Journal*, 16(9):43–53, 2003.

[261] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.

[262] Allen J Wood, Bruce F Wollenberg, and Gerald B Sheblé. *Power generation, operation, and control*. John Wiley & Sons, 2013.

[263] D Woolnough, Vaithilingam Jeyakumar, and Guoyin Li. Exact conic programming reformulations of two-stage adjustable robust linear programs with new quadratic decision rules. *Optimization Letters*, 15:25–44, 2021.

[264] Felix F Wu and W-HE Liu. Detection of topology errors by state estimation (power systems). *IEEE Trans. Power Systems*, 4(1):176–183, 1989.

[265] Álinson S Xavier, Feng Qiu, and Shabbir Ahmed. Learning to solve large-scale security-constrained unit commitment problems. *INFORMS Journal on Computing*, 33(2):739–756, 2021.

[266] Le Xie, Yingzhong Gu, Xinxin Zhu, and Marc G Genton. Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch. *IEEE Trans. Smart Grid*, 5(1):511–520, 2013.

[267] Weijun Xie. On distributionally robust chance constrained programs with Wasserstein distance. *Mathematical Programming*, pages 1–41, 2019.

[268] Weijun Xie. Tractable reformulations of distributionally robust two-stage stochastic programs with ∞- Wasserstein distance. *arXiv preprint arXiv:1908.08454*, 2019.

[269] Weijun Xie. Tractable reformulations of two-stage distributionally robust linear programs over the type-∞ Wasserstein ball. *Operations Research Letters*, 48(4):513–523, 2020.

[270] Weijun Xie and Shabbir Ahmed. Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Trans. Power Systems*, 33(2):1860–1867, 2017.

[271] Weijun Xie, Shabbir Ahmed, and Ruiwei Jiang. Optimized bonferroni approximations of distributionally robust joint chance constraints. *Mathematical Programming*, pages 1–34, 2019.

[272] Weijun Xie, Jie Zhang, and Shabbir Ahmed. Distributionally robust bottleneck combinatorial problems: Uncertainty quantification and robust decision making. *Mathematical Programming*, pages 1–44, 2021.

[273] Mingyu Yan, Ning Zhang, Xiaomeng Ai, Mohammad Shahidehpour, Chongqing Kang, and Jinyu Wen. Robust two-stage regional-district scheduling of multi-carrier energy systems with a large penetration of wind power. *IEEE Trans. Sustainable Energy*, 10(3):1227–1239, 2018.

[274] Mengqi Yao, Meghana Bharadwaj, Zheng Zhang, Baihong Jin, and Duncan S Callaway. Predicting electricity infrastructure induced wildfire risk in california. *arXiv preprint arXiv:2206.02930*, 2022.

[275] Yaowen Yu and Peter Luh. Scalable corrective security-constrained economic dispatch considering conflicting contingencies. *International Journal of Electrical Power & Energy Systems*, 98:269–278, 2018.

[276] Xieshang Yuan, Naian Liu, Xiaodong Xie, and Domingos X Viegas. Physical model of wildland fire spread: Parametric uncertainty analysis. *Combustion and Flame*, 217:285–293, 2020.

[277] Nurseda Y Yürüşen and Julio J Melero. Probability density function selection based on the characteristics of wind speed data. In *Journal of Physics: Conference Series*, volume 753, page 032067. IOP Publishing, 2016.

[278] Ahmed S Zamzam and Kyri Baker. Learning optimal solutions for extremely fast AC optimal power flow. In *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2020.

[279] Ahmed S Zamzam, Nicholas D Sidiropoulos, and Emiliano Dall'Anese. Beyond relaxation and newton–raphson: Solving AC OPF for multi-phase systems with renewables. *IEEE Trans. Smart Grid*, 9(5):3966–3975, 2016.

[280] Fabrice Zaoui, Stéphane Fliscounakis, and Robert Gonzalez. Coupling opf and topology optimization for security purposes. In *15th Power Systems Computation Conference*, pages 22–26, 2005.

[281] Alireza Zare, CY Chung, Junpeng Zhan, and Sherif Omar Faried. A distributionally robust chance-constrained MILP model for multistage distribution system planning with uncertain renewables and loads. *IEEE Trans. Power Systems*, 33(5):5248–5262, 2018.

[282] Hui Zhang and Pu Li. Chance constrained programming for optimal power flow under uncertainty. *IEEE Trans. Power Systems*, 26(4):2417–2424, 2011.

[283] Nan Zhang, Benjamin D Leibowicz, and Grani A Hanasusanto. Optimal residential battery storage operations using robust data-driven dynamic programming. *IEEE Trans. Smart Grid*, 11(2):1771–1780, 2019.

[284] Yichi Zhang, Lingfeng Wang, Yingmeng Xiang, and Chee-Wooi Ten. Power system reliability evaluation with SCADA cybersecurity considerations. *IEEE Trans. Smart Grid*, 6(4):1707–1721, 2015.

[285] Yiling Zhang, Siqian Shen, and Johanna L Mathieu. Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. *IEEE Trans. Power Systems*, 32(2):1378–1388, 2016.

[286] Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse stochastic optimization with Wasserstein metric. *Operations Research Letters*, 46(2):262–267, 2018.

[287] Chaoyue Zhao and Ruiwei Jiang. Distributionally robust contingency-constrained unit commitment. *IEEE Trans. Power Systems*, 33(1):94–102, 2017.

[288] Weiye Zheng, Wanjun Huang, David J Hill, and Yunhe Hou. An adaptive distributionally robust model for three-phase distribution network reconfiguration. *IEEE Transactions on Smart Grid*, 12(2):1224–1237, 2020.

[289] Anping Zhou, Hefeng Zhai, Ming Yang, and You Lin. Three-phase unbalanced distribution network dynamic reconfiguration: A distributionally robust approach. *IEEE Transactions on Smart Grid*, 13(3):2063–2074, 2021.

[290] Yuqi Zhou, Jorge Cisneros-Saldana, and Le Xie. False analog data injection attack towards topology errors: Formulation and feasibility analysis. In *Proc. IEEE PES General Meeting*, 2018.

[291] Yuqi Zhou, Arun Sukumaran Nair, David Ganger, Abhinandan Tripathi, Chaitanya Baone, and Hao Zhu. Appliance level short-term load forecasting via recurrent neural network. In *2022 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2022.

[292] Yuqi Zhou, Jeehyun Park, and Hao Zhu. Scalable learning for optimal load shedding under power grid emergency operations. In *2022 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2022.

[293] Yuqi Zhou and Le Xie. Detection of bad data in multi-area state estimation. In *2017 IEEE Texas Power and Energy Conference (TPEC)*, pages 1–6. IEEE, 2017.

[294] Yuqi Zhou, Ahmed S Zamzam, Andrey Bernstein, and Hao Zhu. Substation-level grid topology optimization using bus splitting. In *2021 American Control Conference (ACC)*, pages 1–7. IEEE, 2021.

[295] Yuqi Zhou and Hao Zhu. Bus split sensitivity analysis for enhanced security in power system operations. In *Proc. North American Power Symp.*, 2019.

[296] Yuqi Zhou and Hao Zhu. Efficient identification of bus split events using synchrophasor data. *IEEE Transactions on Power Systems*, 36(5):4800–4808, 2021.

[297] Yuqi Zhou, Hao Zhu, and Grani A Hanasusanto. Transmission switching under wind uncertainty using linear decision rules. In *Proc. IEEE PES General Meeting*, 2020.

[298] Yuqi Zhou, Hao Zhu, and Grani A Hanasusanto. Distributionally robust chance-constrained optimal transmission switching for renewable integration. *IEEE Transactions on Sustainable Energy*, 2022.

[299] Hao Zhu and Georgios B Giannakis. Sparse overcomplete representations for efficient identification of power line outages. *IEEE Trans. Power Systems*, 27(4):2215–2224, 2012.

[300] Hao Zhu and Thomas J Overbye. Blocking device placement for mitigating the effects of geomagnetically induced currents. *IEEE Trans. on Power Systems*, 30(4):2081–2089, 2014.

[301] Rujie Zhu, Hua Wei, and Xiaoqing Bai. Wasserstein metric based distributionally robust approximate framework for unit commitment. *IEEE Trans. Power Systems*, 34(4):2991–3001, 2019.

[302] Ray D Zimmerman, Carlos E Murillo-Sánchez, and Robert J Thomas. MATPOWER's extensible optimal power flow architecture. In *Proc. IEEE PES General Meeting*, pages 1–7. IEEE, 2009.

[303] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Systems*, 26(1):12–19, 2010.